



Optimizing Flyrock Forecasting in Open-Pit Blasting Using Hybrid Machine Learning Models

Yulin Zhang¹ · Yingui Qiu¹ · Kun Du¹ · Hoang Nguyen² · Danial Jahed Armaghani³ · Jian Zhou¹

Received: 21 October 2024 / Accepted: 14 June 2025

© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2025

Abstract

Flyrock occurrences pose significant environmental and safety risks in open-pit blasting, threatening equipment, personnel, and potentially causing fatal accidents. This study intends to reduce the risk of flyrock by developing multiple machine learning approaches to estimate its throw range, utilizing Random Forest (RF) as the primary model. Two optimization algorithms, the Light Spectrum Optimizer (LSO) and the Puma Optimizer Algorithm (POA), were employed to fine-tune the hyperparameters of the RF model. Researchers conducted field surveys of blasting events in quarries in Malaysia and the Ayazaga area of Istanbul to compile two databases. Based on these datasets, the reliability and practicality of the LSO-RF and POA-RF models were examined. For comparison, an unoptimized RF model was also created to estimate the flyrock throw distance. The performance of the models was assessed using four metrics: root mean square error (RMSE), mean absolute percentage error (MAPE), coefficient of determination (R^2), and variance accounted for (VAF). Results show that LSO-RF and POA-RF outperform the unoptimized RF model across all metrics. For the Malaysian dataset, LSO-RF achieves training/testing R^2 of 0.9661/0.9597, MAPE of 4.445/6.711, and RMSE of 12.081/16.091; POA-RF yields training/testing R^2 of 0.9662/0.9596, MAPE of 4.443/6.717, and RMSE of 12.075/16.094. For the Istanbul dataset, LSO-RF achieves training/testing R^2 of 0.9706/0.9698, MAPE of 5.369/6.512, and RMSE of 6.128/6.537; POA-RF yields training/testing R^2 of 0.9725/0.9680, MAPE of 6.006/7.554, and RMSE of 6.087/7.335. Performance differences between hybrid models are within 0.2% across datasets, indicating consistent reliability. Results from the engineering case validation demonstrate that the developed hybrid model exhibits strong generalization performance, with R^2 of 0.8778. SHAP-based sensitivity analysis identifies key parameters: hole diameter (HD) and maximum charge per delay (CD) for Malaysia, specific charge (q) and spacing-to-burden ratio (S/B) for Istanbul, all positively correlated with FD. An intelligent software tool integrating these models supports data-driven blasting parameter optimization, enhancing safety and efficiency by minimizing flyrock risks.

Highlights

- Optimized machine learning models predict flyrock throw range in open-pit blasting.
- Light Spectrum Optimizer and Puma Optimizer Algorithm enhance Random Forest performance.
- Field data from Malaysia and Istanbul validate model accuracy and reliability.
- Intelligent software tool developed for real-time flyrock prediction and blasting optimization.

✉ Jian Zhou
j.zhou@csu.edu.cn

Yulin Zhang
yl_zhang@csu.edu.cn

Yingui Qiu
225501014@csu.edu.cn

Kun Du
dukuncsu@csu.edu.cn

Hoang Nguyen
nguyenhoang@humg.edu.vn

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

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

² 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

³ School of Civil and Environmental Engineering, University of Technology Sydney, Ultimo, NSW 2007, Australia

Keywords Flyrock distance prediction · Open-pit blasting optimization · Light spectrum optimizer · Puma optimizer algorithm · Random forest · Interpretability analysis · Intelligent feedback system

1 Introduction

The drilling and blasting method is widely used for rock fragmentation in mining operations and tunnel construction. Nonetheless, current studies indicate that over 75% of the blasting energy is wasted, with merely 25% effectively utilized for rock fragmentation (Armaghani et al. 2020; Singh et al. 2016). In addition, blasting can result in negative environmental impacts, especially during extensive operations in open-pit mines (Fig. 1). Among the environmental concerns, flyrock poses one of the most hazardous and destructive threats, leading to damages, injuries, and fatal accidents. Hence, assessing the flyrock propagation distance is crucial for mitigating these adverse effects (Faradonbeh et al. 2016a, b; Ghasemi et al. 2014; Hudaverdi 2022).

Flyrock is characterized by the abrupt expulsion of rocks from the blasting site propelled by explosive energy, which can travel considerable distances and pose significant risks in quarries near residential areas (Armaghani et al. 2016). Beyond endangering workers and infrastructure, flyrock distance (FD) unpredictability stems from a complex interaction of controllable and uncontrollable factors. Controllable parameters, such as burden, sub-drilling, spacing, maximum charge weight per delay, explosive type, powder factor, hole diameter/depth, delay timing, and stemming length, are set by engineers during blast design (Hudaverdi & Agan 2023; Murlidhar et al. 2021). Conversely, uncontrollable factors—intrinsic to the rock mass and beyond engineering modification—include the rock mass rating, geological strength index, rock quality designation, and joint development degree (Koopalipoor et al. 2019). Accurate prediction of FD is, therefore, critical for optimizing blast patterns and identifying high-risk zones to minimize hazards.

According to research conducted by Zhou et al., (2019), the shape and size of flyrock can exhibit significant variability. Beyond identifying the influencing factors of flyrock, understanding the mechanisms of flyrock generation is crucial. The size of flyrock particles is directly proportional to the hole diameter; as the hole diameter increases, so does the size of the resulting flyrock. In addition, the initial ejection velocity of flyrock correlates with its size. Previous studies (Bhandari 1997; Khandelwal and Monjezi 2013; Jahed Armaghani et al. 2015; Hudaverdi & Akyildiz 2019) have classified the primary mechanisms of flyrock generation into three categories: face bursting, cratering, and rifling, as depicted in Fig. 1. Insufficient burden is the direct cause of face bursting, while a disproportionately small ratio of

stemming to hole diameter and inappropriate choice of stemming material can result in rifling and cratering phenomena.

Flyrock poses significant risks to infrastructure and human safety (Raina & Murthy 2016). Precise FD estimation ensures safe blast boundaries, reducing environmental harm and economic losses. Underestimating FD may compromise safety, while overestimating FD can lead to inefficient land use and resource wastage. Conversely, overestimating the FD can lead to unnecessary land resource underutilization and economic resource wastage. Hence, enhancing the capability to evaluate the flyrock ejection distance is essential for effective resource management. When formulating blasting plans, calculating the FD is vital to identify statutory hazardous areas and ensure the safe and efficient execution of blasting activities (Hasanipanah et al. 2017; Rezaei et al. 2023).

In recent years, numerous researchers have undertaken comprehensive studies to estimate or predict FD. These efforts aim not only to mitigate the environmental impact of blasting but also to delineate safe blasting zones. Currently, three predominant methods dominate the FD prediction research landscape: empirical formulas, statistical analysis, and artificial intelligence algorithms (Li, et al. 2023; Hudaverdi & Akyildiz 2019; Hudaverdi 2022; Zhou et al. 2024a, b). A review of previous studies reveals that numerous scholars have put forth empirical formulas and relationships to evaluate and predict blasting FD based on the blasting face. Lundborg et al. (1975) formulated a calculation relationship between hole diameter and FD. Gupta (1980) incorporated burden and stemming length into the empirical formula. In a separate study, McKenzie (2009) predicted the maximum ejection range of rocks based on the rock density, explosive properties, and blasting hole diameter. Moreover, after analyzing data from 95 blasting events across 4 open-pit mines in India, Trivedi et al. (2014) proposed an FD calculation equation. Ghasemi et al. (2012) introduced another empirical model utilizing dimensional analysis principles. However, accurately estimating FD is a complex phenomenon, and some studies have indicated that the predictive performance of empirical formulas is not satisfactory. The primary reason for this limitation is that these empirical equations consider only a limited set of influencing factors, whereas the prediction of FD often encompasses multiple variables (Yari, et al. 2023; Mohamad et al. 2012).

In addition to empirical formulas, many scholars have attempted to predict FD using statistical analysis (multiple regression) or simulation techniques (Monte Carlo simulation) (Li, et al. 2023; Hudaverdi 2022; Yari, et al. 2016;

Fig. 1 Schematic diagram flyrock generation: **A** mechanism of flyrock generation; **B** related parameters affecting the distance of flyrock

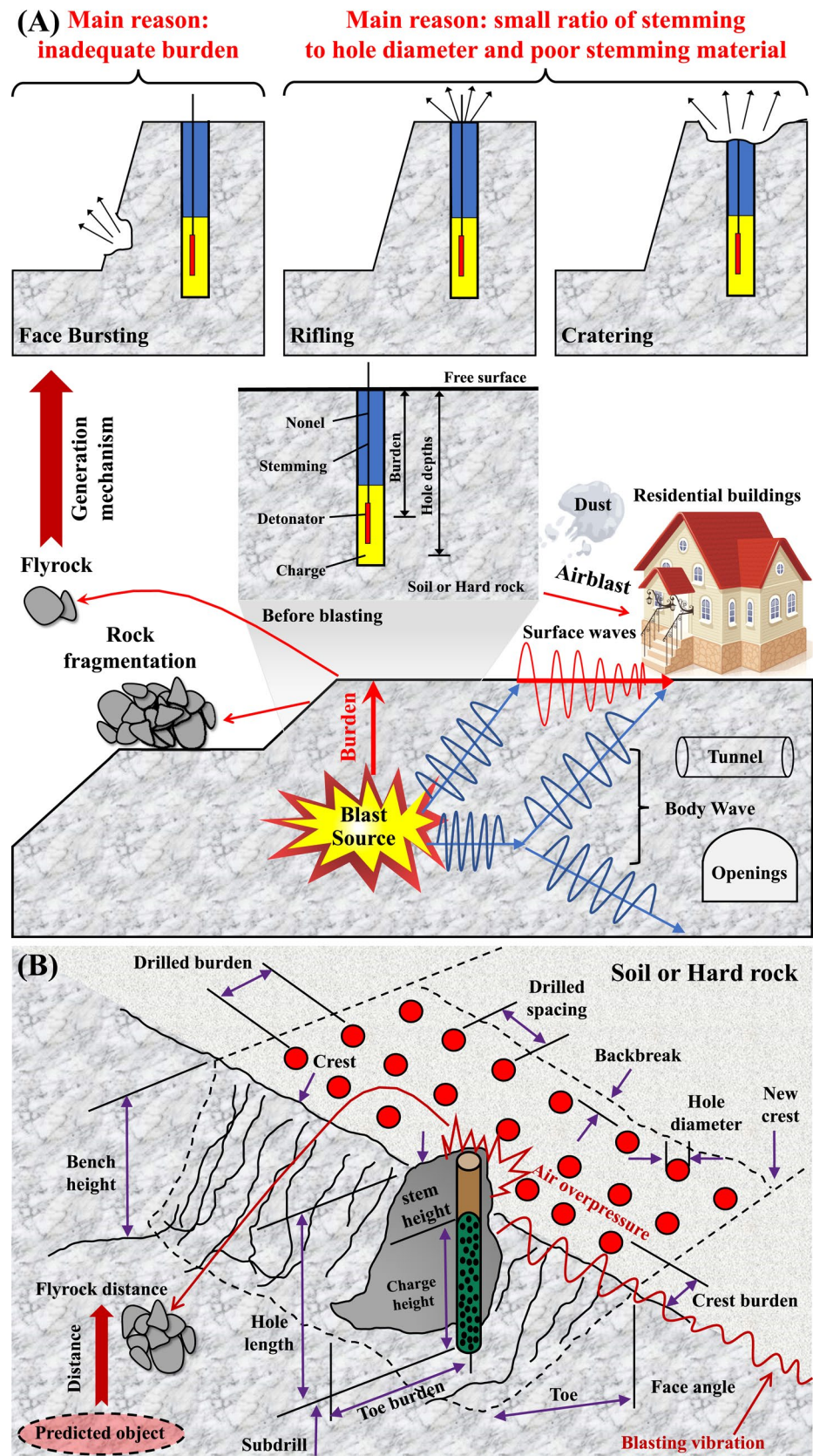


Table 1 The research results of AI technology to predict flyrock distance

Literatures	No. data	Influence factors		AI methods	Interpretability
		Controllable factors	Uncontrollable factors		
Monjezi et al. (2010)	250	HD, PF, SD, N, C, BS, ST	RD	ANN	No
Rezaei et al. (2011)	490	HD, S, PF, SD, C, B, ST	RD	FIS	No
Monjezi et al. (2012)	195	HD, S, PF, SD, D, B, ST, C	RMR	ANN-GA	No
Armaghani et al. (2014)	44	D, HD, S, B, ST, PF, c, Sb, N	RD	PSO-ANN	No
Jahed Armaghani et al. (2015)	166	C, DI, BS, ST	-	ANN, ANFIS	No
Faradonbeh et al. (2016a, b)	262	D, BS, ST, HD, C, PF	-	GP	No
Bakhtavar et al. (2017)	320	B, ST, HL, S, n, C, T, PF	RMR	H-DAFIS	No
Faradonbeh et al. (2018)	76	D, S, B, ST, PF	-	GEP	No
Lu et al., (2020)	82	S, B, T, PF	RD	ANN	No
Guo et al., (2021)	240	HD, C, B, S, ST, PF		DNN-WOA, ANN	No
Fattahi and Hasanipanah, (2022)	80	S, B, ST, PF	RD	ANFIS, ANFIS-ICA	No
Han et al., (2020)	262	HD, D, BS, ST, C, PF	-	RF, BN	No
Hudaverdi (2022)	109	B, B/D, ST/B, Sb/B, CW, PF, N, n, X _B , AMF	H/B	ANFIS	No
Li et al., (2023)	262	HD, D, BS, ST, C, PF	-	HHO-SVM	Yes

HL, hole length (m); S, spacing (m); B, burden (m); ST stemming (m); PF, powder factor (kg/m³); SD, specific drilling (m/m³); C, maximum charge per delay (kg); c, charge per delay (kg); D, hole diameter (mm); HD, hole depth (m); RD, rock density (g/cm³); BS, burden to spacing; AMF, Amount of the material in front of the face; N, number of rows; RMR, rock mass rating; Sb, sub-drilling (m/m³); DI, distance from the blasting face (m); q, specific charge (kg/ton); H, bench height (m); n, number of holes; CW, charge weight; XB mean in situ block size; T, delay time (ms); H/B, stiffness factor; H-DAFIS, hybrid dimensional analysis fuzzy inference system; GP, genetic programming; PSO, particle swarm optimization; GA, genetic algorithm; ICA, imperialist competitive algorithm; GEP, gene expression programming; ANN, artificial neural network; FIS, fuzzy inference system; DNN, deep neural network; WOA, whale optimization algorithm; ANFIS, adaptive neural fuzzy inference system; RF, Random Forest; HHO-SVM, Harris hawks optimization with multi-strategies-based support vector regression; BN, Bayesian network

Zhou et al. 2024a, b). However, these techniques exhibit notable drawbacks. First, as highlighted by Marto et al. (2014), if newly available data diverges from the original data used in model development, this significantly increases the unreliability of the output model. Second, according to Little and Blair (2010), a primary limitation of simulation techniques is their lack of dependence on historical databases to control or determine input distributions. Consequently, further research is needed to develop novel FD prediction techniques.

The interaction between various variables influencing blasting FD is intricate, characterized by nonlinear multivariate relationships. To mitigate the reliance on the complex principles of blasting, artificial intelligence (AI) models have been introduced. Unlike the aforementioned methods, AI approaches can estimate FD without a deep understanding of blasting principles (Zhou et al. 2024a, b; Shakeri et al. 2022; Bui et al. 2020; Ghojoghi et al. 2024). Owing to their streamlined modeling process, high prediction accuracy, and capability to address nonlinear problems, AI techniques have gained prominence in recent years for FD evaluation and have been widely adopted, as illustrated in Table 1. From Table 1, it is evident that FD estimation predominantly employs single AI models,

such as artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), support vector regression (SVR), back-propagation neural network (BPNN), gene expression programming (GEP), and deep neural network (DNN). However, these individual AI models often exhibit a low learning rate and are prone to getting ensnared in local minima. Furthermore, when confronting the complex challenges associated with blasting, the selection of optimal built-in parameters for AI models using conventional methods proves to be exceedingly challenging and time-consuming. Even with an optimal parameter combination, these AI models struggle to yield stable solutions and lack robust generalization capabilities (Zhou et al. 2020, 2023a, b; Rezaei et al. 2024; Qiu and Zhou 2024; Wang et al. 2024). Alternatively, the challenge of configuring built-in parameters can be reframed as an optimization problem. In recent years, several researchers have endeavored to enhance the efficacy of traditional FD models by leveraging metaheuristic algorithms. Following on-site investigations at two distinct quarries, Nikafshan Rad et al. (2020) integrated the genetic algorithm (GA) with the recurrent fuzzy neural network (RFNN) to devise the GA-RFNN intelligent model for FD prediction. Having gathered FD data from the Sungun copper mine, Monjezi et al. (2012)

sought to estimate the FD range utilizing the GA-ANN hybrid model. In another study, Li et al. (2023) compiled 262 sets of FD data from various quarries in Malaysia. Leveraging this dataset, the researchers employed the Harris hawks optimization (HHO) algorithm to regulate the hyperparameters of SVR, culminating in the development of the HHO-SVR model to effectively mitigate the detrimental effects of flyrock in open-pit mines.

As previously discussed, the evolution in FD prediction has seen a transition from single AI models to hybrid models, underscoring the pivotal role of hybrid models in addressing FD challenges. The RF algorithm has been successfully employed to tackle a range of practical engineering issues, including tunnel water inflow prediction (Zhou et al. 2023a, b), wind pressure prediction (Lang et al. 2021), thickness prediction of excavation damage zone (Chen et al. 2023), tunneling-induced ground settlement prediction (Yang et al. 2023a, b), and soil liquefaction potential evaluation (Zhou et al. 2022a, b), it has exhibited outstanding performance in these domains. Evidently, the RF model has made substantial contributions to mining and tunnel engineering challenges, offering advantages such as interpretability analysis, robust generalization capability, and adeptness in handling high-dimensional data (Armaghani et al. 2014; Zhou et al. 2022a; Yang et al. 2023a, b; Chen et al. 2023; Zhang et al. 2024). The aforementioned analysis underscores the superiority of RF as an integrated model. Furthermore, based on the review of prior achievements (See Table 1), it is apparent that, to date, no RF-based hybrid models have been devised in the FD prediction field. Consequently, leveraging the advantages of RF and addressing the existing research gap, this study established two metaheuristic hybrid models, namely LSO-RF and POA-RF, to evaluate the blasting safety range in quarries located in Malaysia and Istanbul.

With advancements in the FD research field, sophisticated AI methods have gradually addressed the limitations and shortcomings of previous studies (Zhou et al. 2024a, 2023a; Yang et al. 2023a). In this context, the primary novelty of this study resides in the development of a hybrid RF model based on innovative metaheuristic algorithms, employed to assess the severity of blasting effects. To attain this objective, two intelligent metaheuristic optimization algorithms, namely LSO and POA, were utilized to fine-tune the configuration of RF's built-in parameters. LSO and POA are proficient in addressing nonlinear, high-dimensional problems through a relatively cost-effective computational process. In addition, datasets from two distinct regions were compiled, encompassing 262 sets of FD data from 6 quarries in Johor state, Malaysia, and 109 sets of FD data from the Ayazaga quarry in Istanbul. Moreover, the pre-established RF algorithm was benchmarked against all previously developed optimal models, the majority of which had not been

explored for this FD problem previously. These studies were conducted with the objective of overcoming the limitations of prior research, enhancing prediction accuracy, and bolstering model reliability. Consequently, this research aims to make significant contributions to mitigating the detrimental effects of blasting.

2 Case Study

The data utilized in this study comprise two distinct databases, designated as Database A and Database B. Sample A encompasses 262 sets of flyrock data sourced from 6 open-pit mines in Malaysia (Han et al. 2020; Li et al. 2023). Sample B, on the other hand, was acquired from a quarry in Istanbul, comprising a total of 109 sets of flyrock data (Hudaverdi 2022). The parameters associated with FD are illustrated in Fig. 1.

2.1 Project Overview of the Malaysian Sample

To accurately estimate the FD, this study investigated six open-pit mines in Malaysia: Taman Bestari, Putri Wangsa, Trans Crete, Ulu Tiram, Masai, and Ulu Choh, as depicted in Fig. 2A. According to publicly available information, these mines each conduct approximately 15 large-scale blasting events per month. Such large-scale and high-frequency blasting activities are prone to triggering significant flyrock incidents (Han et al. 2020; Li et al. 2023). Sample A collected in this study considered six design parameters related to FD. These parameters comprise hole diameter (HD), hole depth (D), burden to spacing ratio (B/S), stemming (ST), maximum charge per delay (CD), and powder factor (PF). The specific parameters are detailed in Table 2.

2.2 Project Overview of the Istanbul Sample

An additional investigation was carried out at the Ayazaga quarry in Istanbul, as illustrated in Fig. 2B. Sample B, collected for this study, considered 12 design parameters related to FD (Hudaverdi 2022). These parameters encompass burden (B), spacing (S), bench height (H), hole diameter (HD), stemming height (SH), sub-drilling (U), charge weight (CW), specific charge (q), mean in situ block size (XB), number of rows (NR), number of holes (NH), and amount of material in front of the face (AMF). The final parameter AMF denotes the face conditions, which are critical in rock blasting operations conducted through a free face. Face conditions directly influence both operational efficiency and environmental impacts. These conditions are classified into four categories (1 to 4) based on the quantity of material

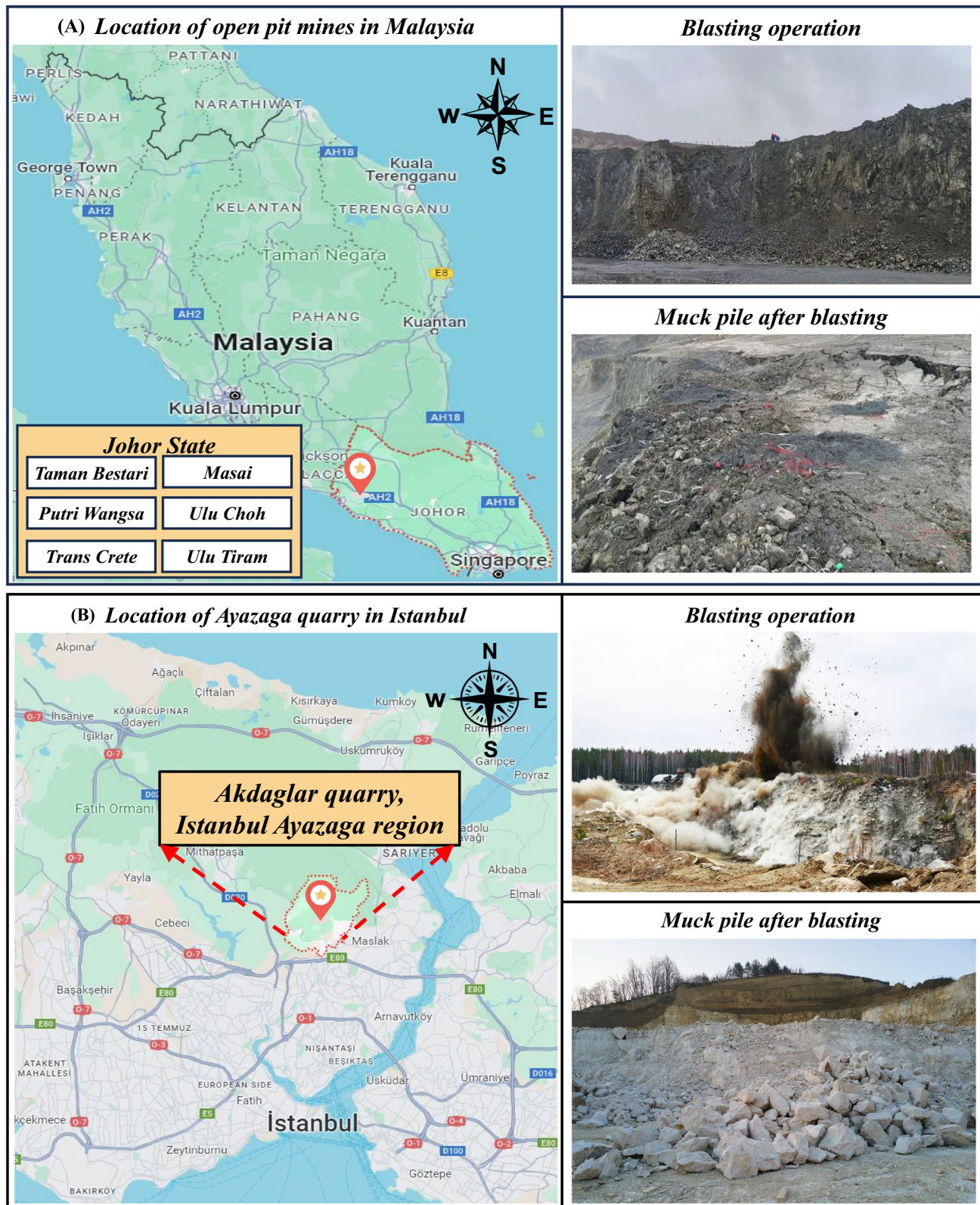


Fig. 2 Location of Asian cases

Table 2 Statistical analysis of parameters in Sample A

Parameter	Minimum	Maximum	Median	Mean		
				Initial sample	Training set	Test set
Inputs						
HD	75.00	150.00	115.00	115.81	117.20	110.23
D	10.00	29.00	20.00	19.76	20.04	18.62
B/S	0.40	0.95	0.75	0.70	0.70	0.69
ST	1.40	4.50	3.00	2.98	3.04	2.72
CD	47.71	593.76	201.56	257.63	266.15	223.24
PF	0.45	1.14	0.80	0.78	0.79	0.78
Outputs						
FD	67	354	189	204.58	206.77	195.74

Hole diameter (HD), powder factor (PF), max charge per delay (CD), burden to spacing ratio (B/S), stemming (ST), hole depth (D), flyrock distance (FD)

accumulated in front of the excavation face, with the classification system reflecting varying degrees of material obstruction affecting blasting performance. The specific parameters are detailed in Table 3.

2.3 Statistical Analysis of Datasets

The accuracy and efficacy of the FD evaluation model employing artificial intelligence algorithms are intrinsically linked to the volume and quality of the recorded blasting data. In this study, two FD databases from the Asian region were systematically compiled. The meticulousness of the data collection process culminated in the development of two comprehensive datasets encompassing a diverse array of blasting design parameters.

Prior to developing the hybrid model, the Pearson correlation analysis was conducted to identify relationships among input features and reduce computational redundancy. As a standard practice, features with correlation coefficients exceeding 0.9 were considered for elimination. The results of this analysis on the blasting FD dataset are presented in Fig. 3, which uses a matrix format to visualize pairwise relationships. The upper triangle uses a color gradient (red for positive correlations, blue for negative correlations), where intensity corresponds to correlation strength. The lower triangle directly lists numerical correlation values, enabling precise interpretation of variable interactions. This approach ensures transparency in feature selection while guiding model optimization. Clearly, as depicted in Fig. 3a, the correlations between the features are all below 0.9, obviating the need to consider the elimination of any features. Conversely, in Fig. 3b, features B and B/HD demonstrate a strong linear correlation, evidenced by a correlation coefficient of 1. Consequently, to mitigate unnecessary computational overhead due to

their high correlation, B and B/HD will be excluded at the model input stage.

3 Research Technique

The primary objective of this study is to enhance the predictive accuracy of FD in open-pit blasting by integrating advanced optimization algorithms with the RF model. While RF is inherently robust for regression tasks, its performance is highly sensitive to hyperparameter tuning, particularly for *n_estimators* and *max_depth*. To address this, two metaheuristic optimization algorithms (LSO and POA algorithms) were selected to systematically fine-tune these parameters. Below is the detailed rationale for their inclusion:

Reasons for Choosing LSO:

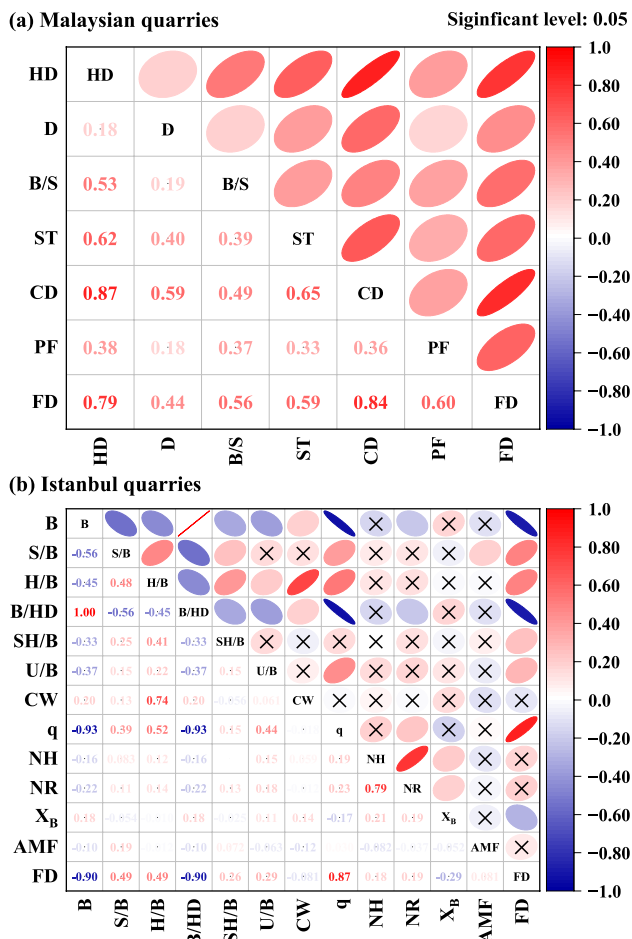
- LSO simulates light refraction and dispersion, enabling dynamic adjustment of search directions. This mechanism preserves population diversity while accelerating convergence, crucial for escaping local optima in hyperparameter space (Abdel-Basset et al. 2022; Thulasi and Sivamohan 2023).
- The algorithm incorporates adaptive scaling factors (e.g., inverse incomplete gamma function) to enhance exploration in early iterations and exploitation in later stages, ensuring efficient parameter optimization (Saber and Salem 2023).
- The vector-based light propagation model of LSO effectively handles high-dimensional parameter spaces, making it suitable for tuning *n_estimators* and *max_depth* of RF (Wang et al. 2025).

Reasons for Choosing POA:

Table 3 Statistical analysis of parameters in sample B

Parameter	Minimum	Maximum	Median	Mean		
				Initial sample	Training set	Test set
Inputs						
B	1.80	2.50	2.20	2.23	2.23	2.19
S/B	1.00	1.28	1.20	1.18	1.19	1.16
H/B	2.4	4.25	3.18	3.19	3.19	3.20
B/HD	20.22	28.09	24.72	25.01	25.11	24.62
SH/B	0.88	1.67	1.09	1.09	1.09	1.07
U/B	0.16	0.33	0.23	0.23	0.23	0.22
CW	17.97	40.17	26.43	27.25	27.37	26.79
q	0.47	0.98	0.69	0.67	0.67	0.71
NH	10	53	20	20.69	20.51	21.41
NR	1	4	2	2.21	2.16	2.41
X _B	0.21	3.11	1.17	1.32	1.35	1.21
AMF	1	4	1	1.61	1.69	1.32
Outputs						
FD	28	142	79	81.67	79.86	88.82

Hole diameter (HD), burden (B), number of rows (NR), stemming height (SH), spacing (S), specific charge (q), bench height (H), mean in situ block size (X_B), amount of the material in front of the face (AMF), sub-drilling (U), charge weight (CW), number of holes (NH), flyrock distance (FD)

**Fig. 3** Correlation analysis of each feature

- POA emulates puma hunting behavior, combining ambush, sprinting, and territorial exploration to dynamically switch between exploration and exploitation phases (Abdollahzadeh et al. 2024; Jiang et al. 2024). This adaptability aligns with the need to explore diverse hyperparameter configurations while intensifying search around promising regions.
- The intelligent phase change mechanism of POA, governed by scoring functions that reward recent improvements and penalize prolonged phase selection, prevents premature convergence and maintains solution diversity (Maurya et al. 2024).
- By integrating historical performance trends and territorial awareness, POA efficiently navigates complex landscapes, ensuring robust optimization hyperparameters of RF even in datasets with intricate feature interactions (Pravesjit and Kantawong 2025; Zaki et al. 2024).

3.1 Random Forest (RF)

Random Forest (RF) is an ensemble learning method that operates by constructing a multitude of decision trees during training and outputs the mode of the classes (for classification) or mean prediction (for regression) of the individual trees (Rigatti 2017; Biau and Scornet 2016). Each decision tree in the forest is constructed using a random subset of the training data and a random subset of the features. The principle behind RF lies in the wisdom of crowds. By aggregating predictions from multiple decision trees, RF reduces the risk of overfitting, as each tree is trained on a different subset of

the data and features (Belgiu and Drăguț, 2016). This diversity among the trees ensures that the model generalizes well to unseen data.

In the context of regression problems, RF exhibits several advantages. First, it can handle both linear and nonlinear relationships between features and the target variable. Unlike linear regression, which assumes a linear relationship (Zhou et al. 2021a, b), RF can capture complex interactions and nonlinearity in the data. Second, RF is robust to outliers and noise in the data. Since it aggregates predictions from multiple trees, outliers are less likely to have a significant impact on the overall predictions. Moreover, the random sampling of features reduces the influence of noisy or irrelevant features, enhancing the model's robustness. Therefore, RF technology is well suited to handle engineering problems related to blasting events (Speiser et al. 2019).

It is noteworthy that the RF model involves several built-in parameters that require careful consideration. In this study, two optimization algorithms, LSO and POA, were selected to fine-tune two critical parameters of RF: 'n_estimators' and 'max_depth'. The 'n_estimators' parameter denotes the number of trees in the RF model. If this parameter's value is too small, the hybrid model may not effectively fit the input parameters, resulting in suboptimal regression performance. Conversely, an excessively large value for this parameter can escalate the computational burden of the model, leading to a significant reduction in processing speed. The 'max_depth' parameter signifies the maximum depth of the trees generated in the RF model.

3.2 Light Spectrum Optimizer (LSO)

The Light Spectrum Optimizer (LSO) is a physics-inspired metaheuristic algorithm derived from the optical phenomenon of rainbow formation, where sunlight disperses into distinct colors through water droplets. The algorithm mimics the physical processes of light refraction, reflection, and dispersion to balance exploration and exploitation in optimization problems. By modeling light rays as candidate solutions, LSO leverages the refractive index variations (ranging from 1.331 for red to 1.344 for violet) and the angular deviations (40°–42°) observed in rainbows. This inspiration enables the algorithm to dynamically adjust search directions, preserving population diversity while accelerating convergence toward optimal solutions. The core innovation lies in integrating vector-based light propagation with adaptive control factors, such as the inverse incomplete gamma function, to enhance exploration in early iterations and exploitation in later stages, thus avoiding local optima (Abdel-Basset et al. 2022; Adam 2002).

1. Initialization

The LSO algorithm starts by generating an initial population of candidate solutions $\vec{x}^0$ within the search space bounds:

$$\vec{x}^0 = lb + RV_1 \times (ub - lb), \quad (1)$$

where RV_1 is a vector of uniform random values in $[0,1]$. This step ensures uniform exploration of the search space (Abdel-Basset et al. 2022; Yu et al. 2013).

2. Direction Calculation

Normal vectors for inner refraction $\vec{x}_{nA}$, inner reflection $\vec{x}_{nB}$, and outer refraction $\vec{x}_{nC}$ are computed using normalized vectors of a randomly selected solution $\vec{x}_t^*$, the current solution $\vec{x}_t^p$, and the global best solution $\vec{x}^*$:

$$\vec{x}_{nA} = \frac{\vec{x}_t^*}{\|\vec{x}_t^*\|}, \vec{x}_{nB} = \frac{\vec{x}_t^p}{\|\vec{x}_t^p\|}, \vec{x}_{nC} = \frac{\vec{x}^*}{\|\vec{x}^*\|} \quad (2)$$

The incident light ray $\vec{x}_{L0}$ is derived from the population mean (X_{mean}):

$$\vec{x}_{L0} = \frac{X_{\text{mean}}}{\|X_{\text{mean}}\|}, X_{\text{mean}} = \frac{\sum_{i=1}^N \vec{x}_i}{N} \quad (3)$$

where N is the population size.

3. Exploration Mechanism

New solutions are generated using either inner refraction $\vec{x}_{L1}$ or outer refraction $\vec{x}_{L3}$ based on a random probability P . The update equation incorporates adaptive scaling factors ϵ and a control factor GI :

$$\vec{x}_{t+1} = \vec{x}_t + \epsilon \cdot RV_1^n \cdot GI \cdot (\vec{x}_{L1} - \vec{x}_{L3}) \times (\vec{x}_{r1} - \vec{x}_{r2}) \quad (4)$$

where $\epsilon = a \cdot RV_3^n$, with a decaying linearly with iterations ($a = RV_2 \cdot (1 - \frac{t}{T_{\text{max}}})$). $GI = a \cdot r^{-1} \cdot P^{-1}(a, 1)$ (inverse incomplete gamma function). RV_1^n , RV_3^n are vectors of random values in $[0,1]$. r is a random scalar in $[0,1]$. $\vec{x}_{r1}$ and $\vec{x}_{r2}$ are randomly selected solutions.

4. Exploitation Mechanism

Solutions are updated by scattering around the global best $\vec{x}^*$ or current solution $\vec{x}_t$ using probabilistic rules:

$$\vec{x}_{t+1} = \begin{cases} \vec{x}_t + RV_3 \cdot (\vec{x}_{r1} - \vec{x}_{r2}) + RV_4^n \cdot (R < \beta) \cdot (\vec{x}^* - \vec{x}_t), & \text{if } R < P_e, \\ 2\cos(\pi \cdot r_1) \cdot \vec{x}^* \cdot \vec{x}_t, & \text{otherwise.} \end{cases} \quad (5)$$

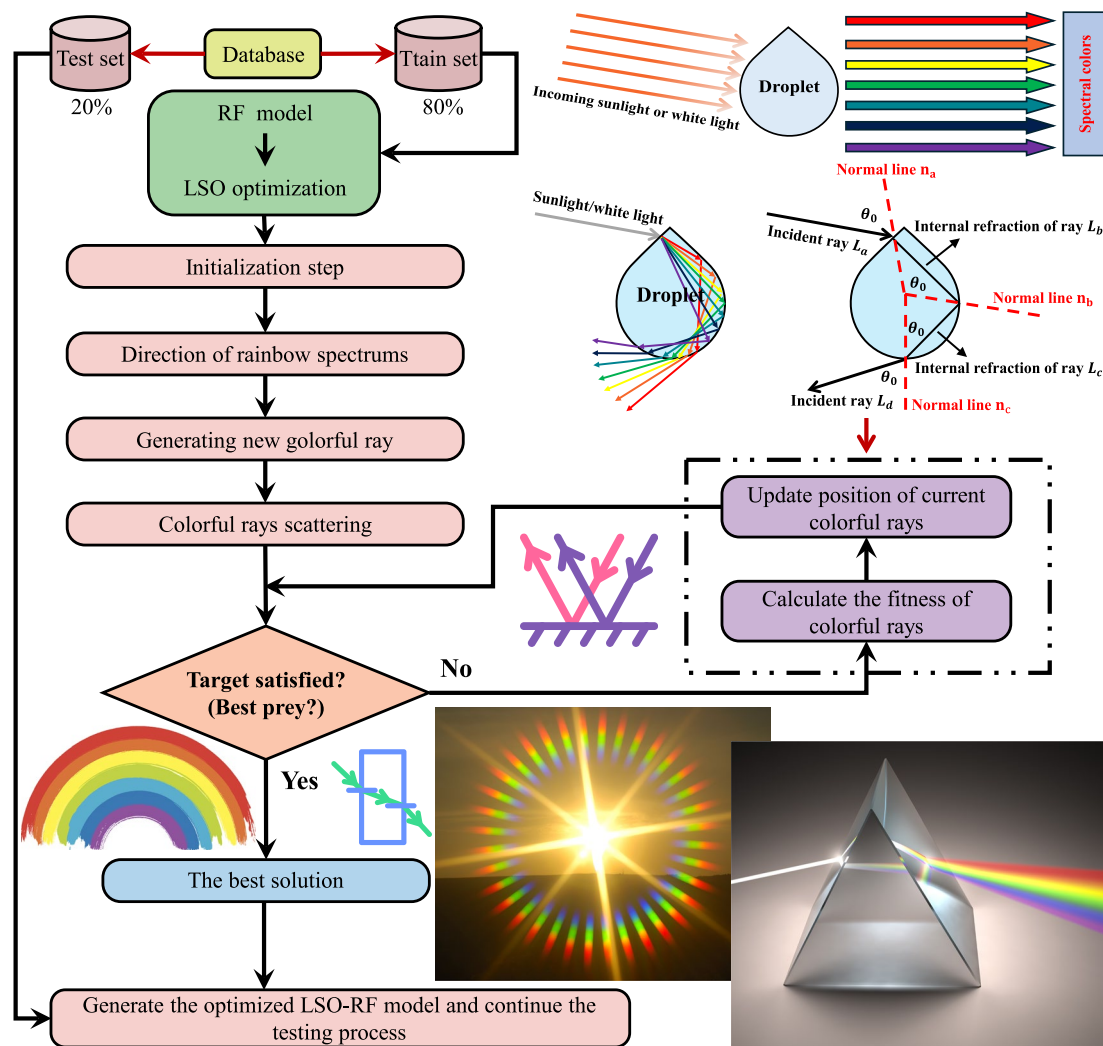


Fig. 4 The basic flow of the LSO operation strategy

Here: RV_3, RV_4^n are random values in $[0,1]$, β and P_e are predefined probabilities, r_1 is a random scalar in $[0,1]$, R is a random threshold.

5. Boundary Handling

Infeasible solutions are adjusted using a hybrid method combining boundary clamping and random regeneration, controlled by probability P_h :

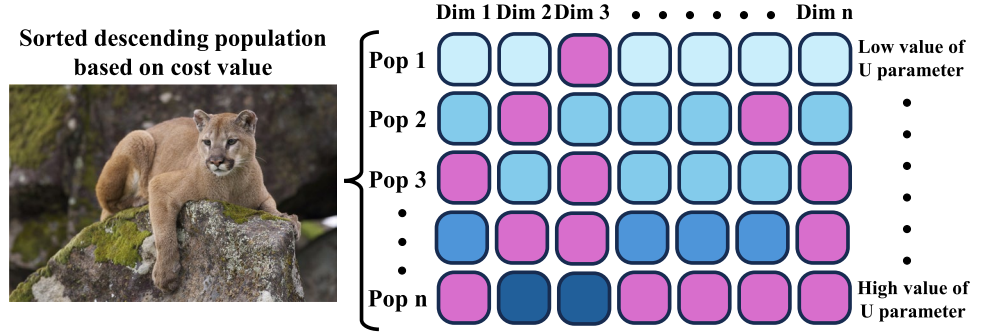
$$\vec{x}_{t+1} = \begin{cases} \text{Clamp to } [lb, ub], & \text{if } R < P_h, \\ \text{Randomly regenerate within } [lb, ub], & \text{otherwise.} \end{cases} \quad (6)$$

This ensures solutions remain within valid bounds while balancing convergence and diversity. To sum up, the basic workflow of LSO is shown in Fig. 4.

3.3 Puma Optimizer Algorithm (POA)

The Puma Optimizer (PO) is a bio-inspired metaheuristic algorithm derived from the hunting behavior and intelligence of pumas in nature. It mimics the adaptive strategies of pumas, such as ambush, sprinting, and territorial exploration, to balance exploration and exploitation in optimization problems. The core innovation lies in its intelligent phase change mechanism, which dynamically switches between exploration and exploitation based on performance metrics, ensuring adaptability to problem landscapes. This mechanism uses three scoring functions to prioritize phases: f_1 emphasizes intensification (rewarding recent improvements), f_2 focuses on historical performance trends, and f_3 promotes diversity by penalizing prolonged phase selection. By integrating these functions, PO adaptively adjusts its search strategy, leveraging puma-like memory and territorial awareness to escape local optima.

Fig. 5 The running program of the exploration stage



and converge efficiently (Abdollahzadeh et al. 2024; Drake et al. 2012).

1. Initialization

The PO algorithm begins by generating an initial population of candidate solutions $\vec{X}_i$ within the search space bounds (ub and lb):

$$\vec{X}_i = lb + rand \cdot (ub - lb) \quad (7)$$

where $rand$ is a uniform random vector in $[0, 1]$. This step ensures uniform exploration of the search space.

2. Phase Change Mechanism

PO dynamically switches between exploration and exploitation using three scoring functions:

Unexperienced phase (first 3 iterations): Simultaneously evaluate both phases using:

$$f1_{\text{Explor}} = PF_1 \cdot \frac{Seq_{\text{CostExplore}}^1}{Seq_{\text{Time}}}, f1_{\text{Exploit}} = PF_1 \cdot \frac{Seq_{\text{CostExploit}}^1}{Seq_{\text{Time}}} \quad (8)$$

where $Seq_{\text{CostExplore}}^1$ and $Seq_{\text{CostExploit}}^1$ track cost improvements, and $Seq_{\text{Time}} = 1$.

Experienced phase: After three iterations, select the dominant phase using:

$$F_i^{\text{phase}} = \alpha_i \cdot (f_1 + f_2) + \delta_i \cdot (lc \cdot f_3) \quad (9)$$

where f_1 rewards recent improvements, f_2 emphasizes historical performance trends, f_3 penalizes prolonged phase selection for diversity, α_i and δ_i balance intensification and diversity, $lc = \{ | \text{Cost}_{\text{old}} - \text{Cost}_{\text{New}} | \}$ tracks cost differences.

3. Exploration Mechanism

New solutions are generated using random jumps or neighboring interactions:

$$Z_{i,G} = \begin{cases} \text{RandDim} \cdot (ub - lb) + lb, & \text{if } rand_1 > 0.5, \\ X_{a,G} + G \cdot (X_{a,G} - X_{b,G}) + G \cdot \Delta, & \text{otherwise,} \end{cases} \quad (10)$$

where: $G = 2 \cdot rand_2 - 1$, $\Delta = (X_{a,G} - X_{b,G}) - (X_{c,G} - X_{d,G}) + (X_{e,G} - X_{f,G})$, $X_{a,G}, X_{b,G}, X_{c,G}, X_{d,G}, X_{e,G}, X_{f,G}$ are randomly selected solutions.

Solutions are updated with probability-based boundary handling:

$$X_{\text{new}} = \begin{cases} Z_{i,G}, & \text{if } rand_3 \leq U \text{ or } j = j_{\text{rand}}, \\ X_{a,G}, & \text{otherwise,} \end{cases} \quad (11)$$

where U adapts dynamically to control diversity ($U = U + p, p = \frac{NC}{N_{\text{pop}}}$). The program for the exploration phase is shown in Fig. 5.

4. Exploitation Mechanism

PO uses two hunting strategies inspired by pumas: Ambush or sprinting strategy:

$$X_{\text{new}} = \begin{cases} \frac{\text{mean}(S_{\text{total}}) \cdot X_i^r - (-1)^{\beta} \cdot X_i}{1 + \alpha \cdot rand_5} & \text{if } rand_4 \geq 0.5, \\ \text{Puma}_{\text{male}} + 2 \cdot rand_7 \cdot e^{rand_1} \cdot X_3^r - X_i & \text{if } rand_6 \geq L, \\ \frac{2 \cdot rand_8 \cdot (F_1 \cdot R \cdot X_i + F_2 \cdot (1 - R) \cdot \text{Puma}_{\text{male}})}{2 \cdot rand_3 - 1 + rand_2} - \text{Puma}_{\text{suspect}} & \text{otherwise} \end{cases} \quad (12)$$

where: $R = 2 \cdot rand_{11} - 1$, $F_1 = rand_{n3} \cdot \exp(2 - \frac{2 \cdot \text{Iter}}{\text{MaxIter}})$, $F_2 = w \cdot v^2 \cdot \cos(2 \cdot rand_{12} \cdot w)$, $\text{Puma}_{\text{male}}$ is the best solution, β, L , and α are control parameters.

5. Boundary Handling

Infeasible solutions are adjusted using a hybrid method:

$$X_{\text{new}} = \begin{cases} \text{Clamp to } [lb, ub], & \text{if } rand < P_h, \\ \text{Randomly regenerate within } [lb, ub], & \text{otherwise,} \end{cases} \quad (13)$$

where P_h controls the balance between convergence and diversity.

In summary, POA operation process is illustrated in Fig. 6.

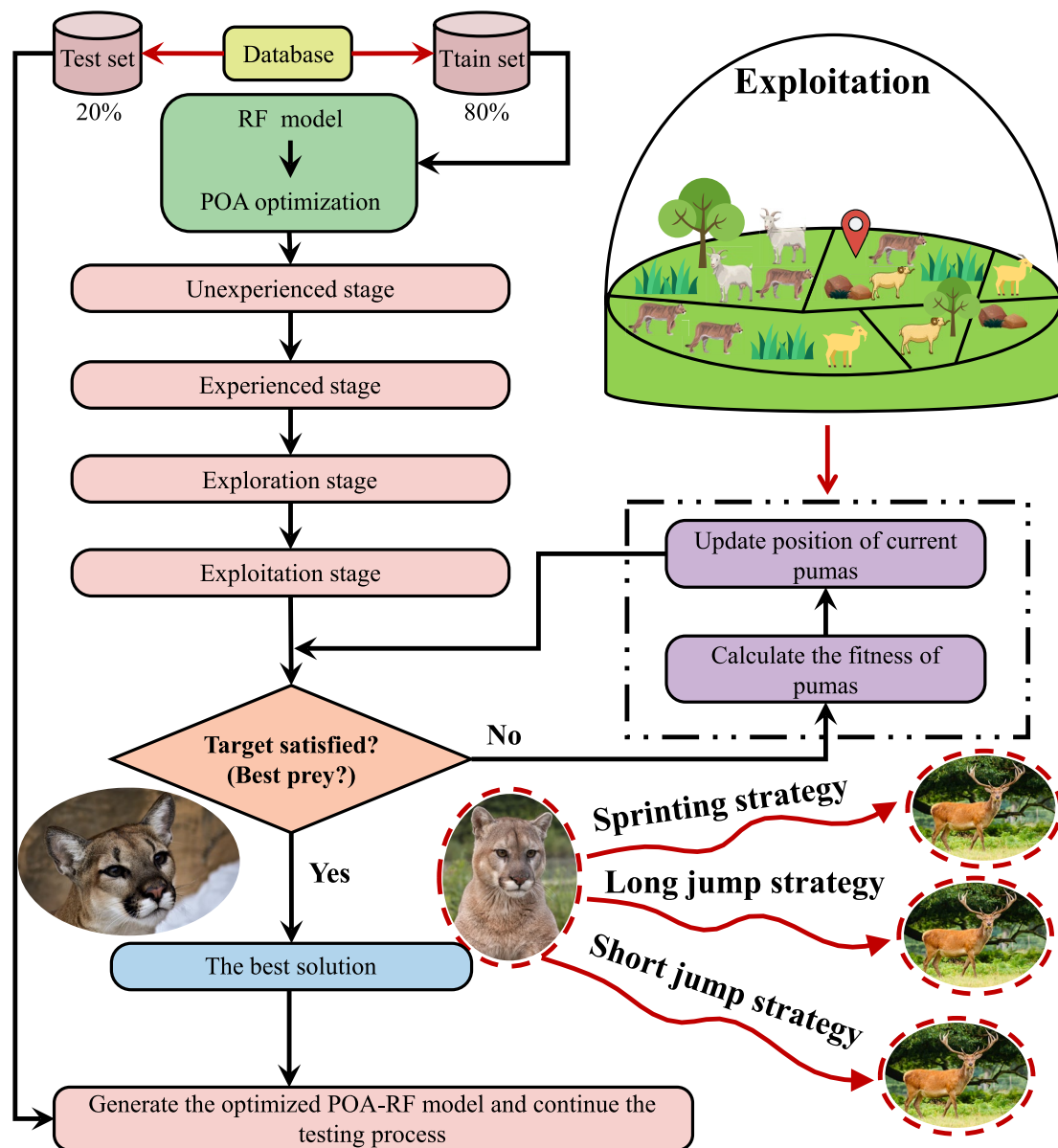


Fig. 6 The basic flow of the POA operation policy

3.4 RF-Based Hybrid Models

In the realm of RF, the model's generalization performance and predictive capability are significantly impacted by the configurations of the two parameters, 'n_estimators' and 'max_depth' (Zhou et al. 2022a, b; Zhou et al. 2022a). To enhance the RF model's estimation accuracy for FD, two metaheuristic algorithms, LSO and POA, were utilized. The resulting intelligent models are denoted as LSO-RF and POA-RF, respectively. As previously mentioned, both algorithms are inspired by social phenomena, but their specific heuristic principles differ. The inherent exploration and exploitation mechanisms of these algorithms substantially

enhance the robustness and diversity in searching for the optimal solution within the search space.

When LSO and POA are synergistically integrated with RF, they constitute a comprehensive and adaptive hyperparameter tuning approach, leading to the development of a more effective and generalized intelligent model. While the inspiration principles underlying LSO and POA differ, both are fundamentally algorithms developed based on social phenomena. Consequently, the parameter adjustment process for both LSO and POA algorithms is essentially analogous. The flowchart of the hybrid prediction model is illustrated in Fig. 7.

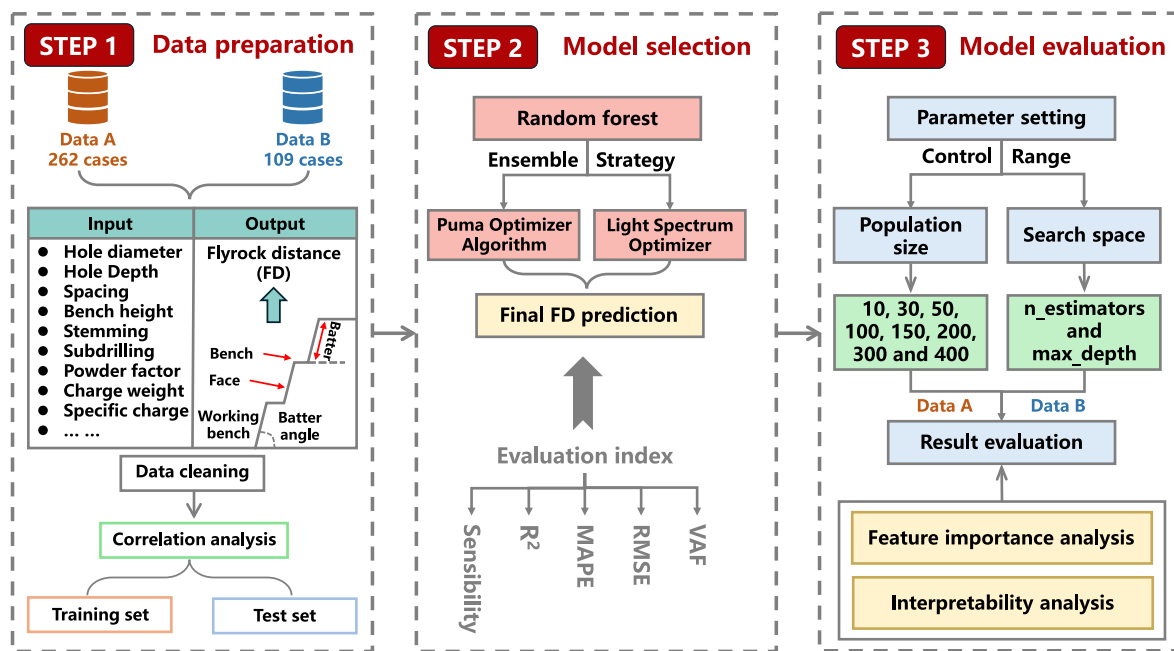
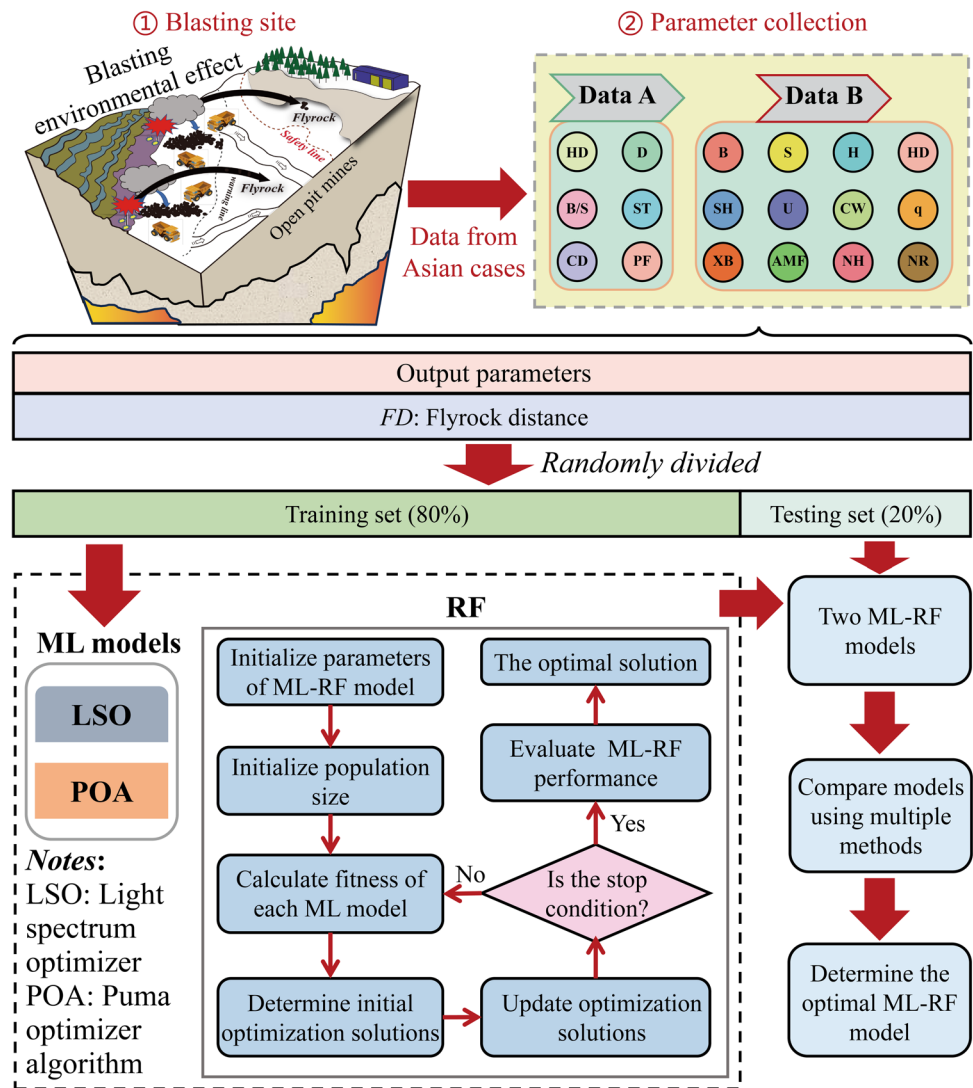


Fig. 7 The operational framework of hybrid intelligence models

- (1) **Database Construction and Preprocessing:** The flyrock data collected from the two regions, Malaysia and Istanbul, were organized and compiled to establish a specialized database. Subsequently, utilizing the RF algorithm as a foundation, Pearson correlation analysis was conducted on the collected FD data. This method systematically eliminates highly correlated outliers and features, thereby preventing computational redundancy. Following the correlation analysis, the dataset was randomly partitioned into a training set and a test set to facilitate the development of the hybrid regression model. It is crucial to ensure consistency between the training and test sets when constructing various combined regression models to enhance the model's generalization capability.
- (2) **Hyperparameter Setting:** The optimization algorithms LSO and POA represent three emerging optimization algorithms, which explore the optimal solution by simulating the mechanism of light refraction or group predation behavior. In addition, they offer significant advantages. These algorithms are not controlled by input parameters but only manage the population size and parameter range. For instance, before executing the POA optimization algorithm, it is essential to set the population size of Pumas and the search space. Consequently, a comparative analysis was conducted on the LSO-RF and POA-RF intelligent models by configuring a series of population sizes to determine the optimal population size. It is crucial to note that the population size range for each hybrid model remains consistent, specifically at 10, 30, 50, 100, 150, 200, 300, and 400. After determining the population size, setting the search range of the RF built-in parameters becomes particularly pivotal as it directly influences the evaluation performance of the entire FD model. In this study, the range for 'n_estimators' is set between 1 and 800, while the range for 'max_depth' is established between 1 and 30. Importantly, once the hyperparameter range is defined, it is imperative for all optimization models to maintain consistency. Furthermore, the iteration number's magnitude directly impacts the computational speed of the optimization model. Therefore, the iteration number for both optimization models is set to 300.
- (3) **Cycle Iteration and Output of the Optimal Solution:** This process primarily encompasses the initialization phase, cost function calculation, and solution output. Initially, the initial population is established based on the sizes determined by LSO and POA. Subsequently, the performance of the hybrid model is assessed based on the cost computed by the cost function to ascertain if it fulfills the specified criteria. In this study, the mean absolute error (MAE) metric is employed as the target

Fig. 8 The whole process of estimating the range of flyrock throw



cost function to analyze the convergence properties of each model. Ultimately, the parameter configuration for the FD model exhibiting optimal performance is determined.

4 Model Evaluation Index

This study utilizes four performance evaluation metrics to validate the predictive capability of the models, namely the root mean square error (RMSE), the coefficient of determination (R^2), the mean absolute percentage error (MAPE), and the variance accounted for (VAF) (Li et al. 2023; Hudaverdi 2022; Han et al. 2020; Dai et al. 2022; Armaghani et al. 2022). Furthermore, the MAE metric is employed to compute the fitness value, assessing the convergence properties of each hybrid model. The specific formulas for these validation metrics are illustrated in Fig. 8.

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2}$$

This metric measures the goodness of fit of a regression model, with values ranging between [0, 1]. A higher R^2 indicates a better fit, as it represents the proportion of the variance in the dependent variable that is predictable from the independent variables.

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

RMSE quantifies the magnitude of deviation between observed y_i and predicted $\hat{y}_i$ values. It provides a measure of the average error in the same units as the dependent variable, with lower values indicating better predictive accuracy.

Table 4 The prediction performance of Sample A

Population size	R ²	Rank	MAPE	Rank	VAF	Rank	RMSE	Rank	Total
Training (LSO-RF)									
10	0.96537	1	4.4751	1	96.2802	1	12.1916	1	4
30	0.96539	2	4.4750	2	96.2826	2	12.1877	2	8
50	0.96561	3	4.4592	3	96.3053	4	12.1503	4	14
100	0.96562	4	4.4562	4	96.3042	3	12.1521	3	14
150	0.96612	8	4.4451	5	96.3475	8	12.0809	8	29
200	0.96590	6	4.4430	6	96.3281	6	12.1128	6	24
300	0.96594	7	4.4421	7	96.3329	7	12.1048	7	28
400	0.96588	5	4.4420	8	96.3265	5	12.1154	5	23
Testing (LSO-RF)									
10	0.95951	3	6.7192	4	94.6667	1	16.2010	1	9
30	0.95955	5	6.7190	5	94.6786	2	16.1783	2	14
50	0.95949	2	6.7386	1	94.6965	4	16.1232	4	11
100	0.95952	4	6.7254	2	94.7011	5	16.1084	5	16
150	0.95965	7	6.7105	8	94.7181	6	16.0910	8	29
200	0.95966	8	6.7120	6	94.7139	8	16.1005	7	29
300	0.95962	6	6.7108	7	94.7132	7	16.1011	6	26
400	0.95941	1	6.7206	3	94.6859	3	16.1503	3	10
Training (POA-RF)									
10	0.96562	3	4.4624	4	96.3019	3	12.1560	3	13
30	0.96567	5	4.4626	3	96.3021	4	12.1559	4	16
50	0.96564	4	4.4604	5	96.3041	5	12.1523	5	19
100	0.96587	6	4.4554	6	96.3243	6	12.1192	6	24
150	0.96524	2	4.4732	2	96.2748	2	12.2004	2	8
200	0.96511	1	4.4903	1	96.2712	1	12.2063	1	4
300	0.96615	8	4.4431	8	96.3509	8	12.0752	8	32
400	0.96596	7	4.4541	7	96.3298	7	12.1101	7	28
Testing (POA-RF)									
10	0.95982	8	6.7039	7	94.7032	5	16.1369	4	24
30	0.95974	6	6.7016	8	94.6973	4	16.1302	5	23
50	0.95949	3	6.7111	5	94.6853	3	16.1457	3	14
100	0.95980	7	6.7172	3	94.7313	8	16.0772	8	26
150	0.95869	2	6.7174	2	94.5937	1	16.2723	1	6
200	0.95863	1	6.7069	6	94.6217	2	16.2008	2	11
300	0.95962	4	6.7169	4	94.7165	7	16.0937	7	22
400	0.95967	5	6.7200	1	94.7096	6	16.1117	6	18

Bold values indicate the optimal performance metrics and corresponding population sizes of LSO-RF and POA-RF models in training and testing

$$VAF = [1 - \frac{\text{var}(y_i - x_i)}{\text{var}(y_i)}] \times 100$$

VAF represents the percentage of the variance in y_i that is captured by the model. It indicates how much of the variability in the data is explained by the relationship with x_i , with higher percentages signifying greater explanatory power.

$$MAPE = \frac{1}{n} \sum_i \left(\frac{|y_i - x_i|}{y_i} \right) \times 100$$

MAPE measures the average percentage error between predicted x_i and observed y_i values. It is a relative measure of error, making it useful for comparing models across different scales, with lower values indicating better performance.

Table 5 The prediction performance of Sample B

Population size	R ²	Rank	MAPE	Rank	VAF	Rank	RMSE	Rank	Total
Training (LSO-RF)									
10	0.97064	4	5.3687	8	96.9996	6	6.1280	6	24
30	0.97117	5	6.0780	1	96.8985	4	6.2302	4	14
50	0.96879	2	5.9781	5	96.7441	1	6.3974	1	9
100	0.97137	6	6.0195	2	96.9495	5	6.1778	5	18
150	0.96982	3	5.9745	6	96.8618	3	6.2702	3	15
200	0.96872	1	5.6918	7	96.7758	2	6.3510	2	12
300	0.97256	8	5.9907	3	97.0423	8	6.0843	8	27
400	0.97243	7	5.9871	4	97.0369	7	6.0915	7	25
Testing (LSO-RF)									
10	0.96978	5	6.5118	7	96.9254	6	6.5368	6	24
30	0.96955	4	7.5644	2	96.2906	3	7.2622	3	15
50	0.97220	7	6.7859	5	96.8539	5	6.5817	5	22
100	0.96932	3	7.2666	4	96.3294	4	7.1594	4	15
150	0.97173	6	6.3474	8	96.9848	7	6.3830	7	28
200	0.97365	8	6.5169	6	97.2028	8	6.1543	8	30
300	0.96871	2	7.4773	3	96.2532	2	7.2626	2	9
400	0.96753	1	7.5829	1	96.1195	1	7.3929	1	4
Training (POA-RF)									
10	0.96699	3	6.6053	1	96.5486	3	6.5710	3	10
30	0.97253	8	6.0061	6	97.0397	8	6.0872	8	32
50	0.96684	1	6.4509	2	96.5479	2	6.5719	2	7
100	0.97174	6	6.0450	5	96.9564	6	6.1716	6	23
150	0.96693	2	6.2747	3	96.5437	1	6.5755	1	7
200	0.97101	4	6.0819	4	96.9052	4	6.2223	4	16
300	0.97214	7	5.9803	8	97.0142	7	6.1145	7	29
400	0.97140	5	6.0056	7	96.9443	5	6.1852	5	22
Testing (POA-RF)									
10	0.97790	7	5.7060	7	97.4283	7	5.9230	7	28
30	0.96795	3	7.5539	2	96.1607	3	7.3349	3	11
50	0.97804	8	5.5381	8	97.5262	8	5.8042	8	32
100	0.96948	5	7.5487	3	96.2972	4	7.2346	4	16
150	0.97683	6	6.4664	6	97.3630	6	6.0045	6	24
200	0.96913	4	7.3105	5	96.3002	5	7.1798	5	19
300	0.96678	1	7.6055	1	96.0520	1	7.4499	1	4
400	0.96760	2	7.4693	4	96.1490	2	7.3441	2	10

Bold values indicate the optimal performance metrics and corresponding population sizes of LSO-RF and POA-RF models in training and testing

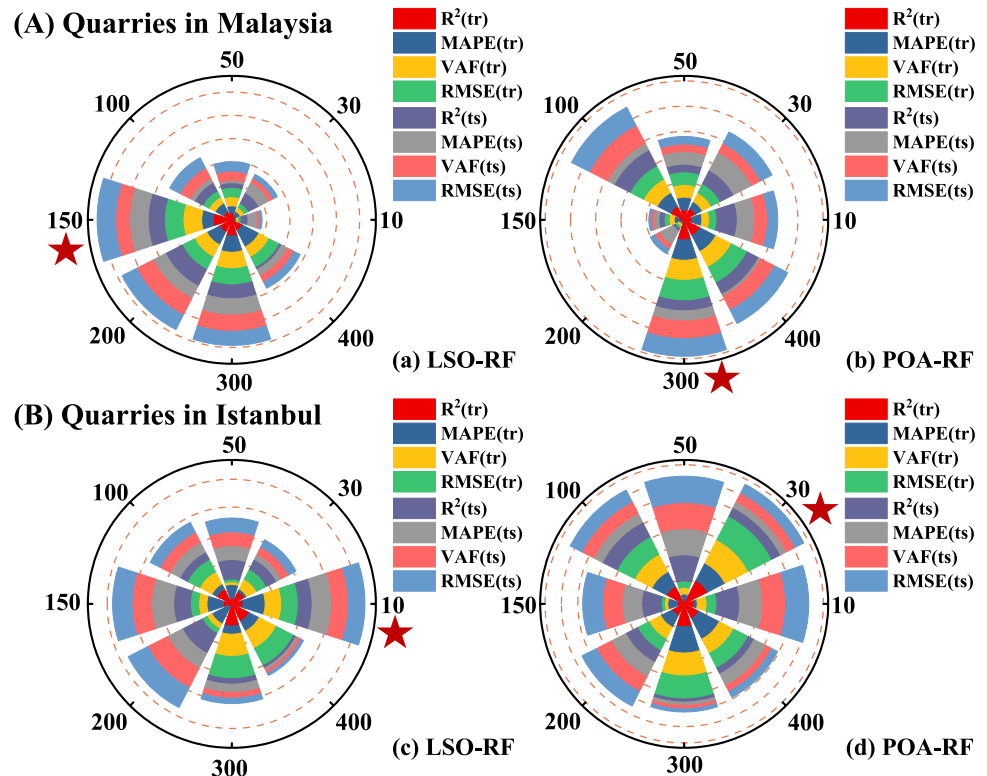
5 Comparative Assessment of Model Performance and Interpretability Analysis

5.1 Parameter Configuration

To augment the sophistication of the FD prediction models in blasting research, this study opted for an optimization strategy for decision-making. This approach was selected to enhance performance by integrating the RF model into the existing framework. The integrated model is denoted as the hybrid intelligent model, specifically LSO-RF and POA-RF.

In addition, to facilitate direct comparison with the optimal model, an unoptimized RF model was also established. In the modeling process, two databases acquired from field surveys were utilized, namely Sample A and Sample B.

It is worth noting that adjusting the population size significantly impacts the computational cost of the model and its capability to search for the global optimal solution. Both excessively large and overly small population sizes can impose limitations. A population size set too large will substantially increase the computational cost of the model, rendering it impractical for application in real-world mining and

Fig. 9 Comprehensive score of all optimization strategies**Table 6** Optimal hyperparameter configuration of hybrid models in different databases

Hyperparameter	Sample A	Sample B
LSO-RF		
Optimal swarm size	150	10
n_estimators	667	9
max_depth	6	11
POA-RF		
Optimal swarm size	300	30
n_estimators	669	48
max_depth	6	11

tunnel engineering sites. Conversely, a population size set too small may lead to unstable fitness values, resulting in the model being trapped in local optima. To comprehensively compare the regression performance of the two hybrid models, this study selected eight consistent population sizes, specifically 10, 30, 50, 100, 150, 200, 300, and 400 (Zhou et al. 2022a, b; Qiu et al. 2022; Zhou et al. 2022a), to construct the hybrid regression models in both LSO-RF and POA-RF. As observed from Tables 4 and 5, the performance of the two hybrid regression techniques under various population sizes has achieved a satisfactory level. Nonetheless, considering all models solely from a practical standpoint is unrealistic. Therefore, it is imperative to comprehensively evaluate the various indicators of the two intelligent models post-training

and to evaluate them collectively. The integrated scoring situation is depicted in Fig. 9. This scoring method aids in identifying the optimal population size. In addition, Table 6 consolidates all the optimal population values for LSO-RF and POA-RF, and the optimal hyperparameter configurations produced by the optimal model are also statistically documented.

5.2 Hybrid Model Performance Analysis

To predict the FD range, this study systematically analyzed and compared the regression performance of the LSO-RF and POA-RF models using Sample A and Sample B during both the training and testing phases. The objective was to develop intelligent evaluation models that meet actual engineering requirements. In addition, it was crucial to determine whether LSO and POA enhance the RF regression performance and, if so, to what extent. These questions were addressed by incorporating an unoptimized RF model in the performance analysis process.

During the training phase, all hybrid models generated eight fitness function curves that varied with the number of iterations. The fitness function curves produced by the two hybrid regression models based on Sample A and Sample B are illustrated in Fig. 10. The upper section of the figure presents the fitness curves of LSO-RF and POA-RF derived from training with Sample A data. Conversely, the lower section of the figure depicts the fitness curves related to

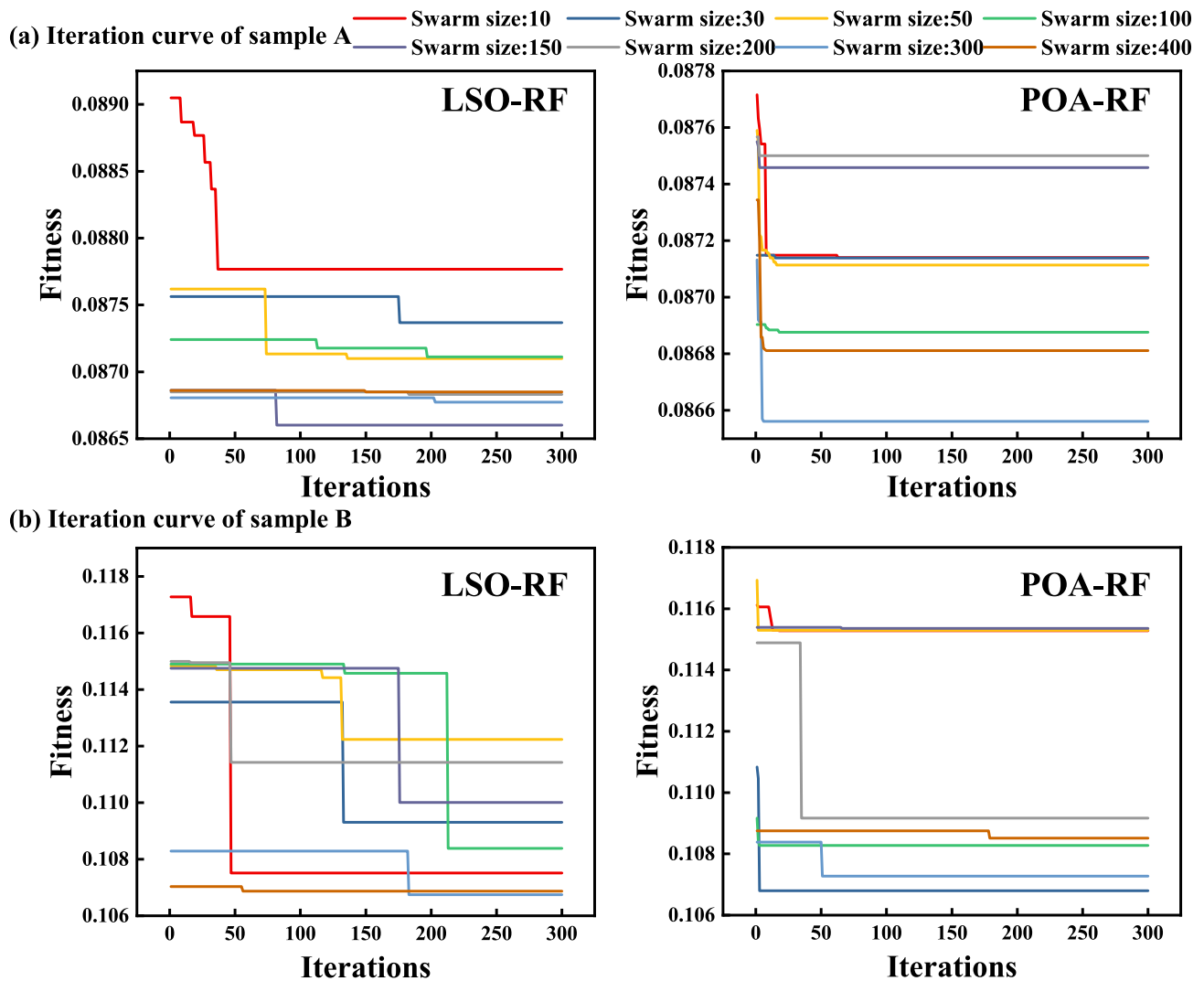


Fig. 10 Fitness function curves of all optimization strategies

Sample B data. From the figure, it is evident that when the optimization algorithm is POA, the fitness curves for different population sizes converge more rapidly. For the majority, the fitness curves essentially achieve convergence after 50 iterations. In contrast, the convergence pattern of the LSO-RF model is rather erratic, and its convergence process lags behind that of the POA-RF model.

While comparing hybrid model performance is critical, evaluating the unoptimized RF model remains essential to determine whether optimization strategies enhance performance and by what magnitude. Under consistent dataset conditions (Sample A and Sample B), this study, therefore, conducts a comprehensive comparison between the optimized LSO-RF/POA-RF models and the baseline RF model, as presented in Table 7. Analysis of evaluation indicators reveals that both LSO-RF and POA-RF outperform the unoptimized RF across all metrics, demonstrating their

superior feasibility and reliability in handling FD datasets compared to the single estimator. This comparative framework ensures rigorous validation of optimization effectiveness while maintaining methodological transparency. In addition, presented in Table 7, in the training phase, the POA-RF model slightly outperforms the LSO-RF model in both Sample A and Sample B. However, in the testing phase, the situation is entirely reversed, with the LSO-RF model slightly surpassing the POA-RF model. Overall, these differences are not very significant. Therefore, the disparities between the POA-RF and LSO-RF models are minimal, and their performances are nearly identical. In addition, the R^2 values for both models exceed 0.95 in all phases for both Sample A and Sample B, indicating the excellent performance of these two intelligent models in handling FD data.

Figure 11A illustrates the different error distribution states for Sample A and Sample B, providing a more

Table 7 The comparison results optimized models and pre-developed model

Models	R ²	MAPE	RMSE	VAF
Sample A (training)				
RF	0.92376	35.8682	17.4537	92.3761
LSO-RF (150)	0.96612	4.4451	12.0809	96.3475
POA-RF (300)	0.96615	4.4431	12.0752	96.3509
Sample A (testing)				
RF	0.93989	39.7111	16.4618	93.9898
LSO-RF (150)	0.95965	6.7105	16.0910	94.7181
POA-RF (300)	0.95962	6.7169	16.0937	94.7165
Sample B (training)				
RF	0.94539	60.7390	8.2653	94.5390
LSO-RF (10)	0.97064	5.3687	6.1280	96.9996
POA-RF (30)	0.97253	6.0061	6.0872	97.0397
Sample B (testing)				
RF	0.94718	60.3677	8.4434	94.7189
LSO-RF (10)	0.96978	6.5118	6.5368	96.9254
POA-RF (30)	0.96795	7.5539	7.3349	96.1607

intuitive representation of the error distribution between the estimated and observed values. A histogram in the graph that is closely centered around zero and shows a narrower distribution indicates good regression performance of the model (Chen et al. 2023). Clearly, for both Sample A and Sample B, the error histograms of the RF model are widely distributed and far from zero. This suggests that the pre-developed RF model may have significant prediction errors in some extreme cases. In contrast, the POA-RF and LSO-RF models display the opposite distribution, underscoring the importance and necessity of optimization strategies and the reliability and stability of the hybrid models. In addition, it can be observed that the distribution shapes of the POA-RF and LSO-RF models are largely similar in both Sample A and Sample B at any stage, further indicating minimal performance differences between the two hybrid models. For a more comprehensive assessment of the stability of various models in predicting FD in blasting scenarios, this study utilized the Taylor diagram to evaluate the performance of all models (Taylor 2001), as depicted in Fig. 11B. In the Taylor diagram, proximity of a model's point to the reference point (purple point) signifies its superior performance (Qasem et al. 2019). For Sample A, it is evident that the points corresponding to RF(tr) and RF(ts) are the most distant from the reference point, whereas the remaining points largely overlap and closely align with the reference point. Similarly, for Sample B, the point representing RF(ts) exhibits the greatest distance from the reference point, followed by RF(tr), while the other points exhibit overlap. These findings underscore the significance of optimization techniques and validate the accuracy and reliability of the optimized models.

After constructing the AI models, evaluating their generalization performance becomes a pivotal and straightforward task. A lack of generalization ability in a model is often characterized by overfitting, where the model's complexity increases, and it learns unnecessary noise (Qiu and Zhou 2023a, b). Thus, preventing overfitting is imperative. In this study, the overfitting phenomena of all estimators were analyzed, as illustrated in Fig. 12. Deviation of an estimator from the 1:1 black line (baseline) indicates the presence of overfitting. For Sample A, the unoptimized RF model demonstrates the most pronounced overfitting. However, upon employing optimization algorithms such as LSO-RF (150) and POA-RF (300), the extent of overfitting notably decreased. In contrast, for Sample B, all models closely align with the baseline, signifying robust generalization performance across all estimators. In summary, compared to a standalone regression model, both LSO and POA methods effectively mitigate overfitting in estimating FD and enhance the overall generalization capability of the model.

5.3 Compared With Previous Literature

Comparing the results with previous related literature is a crucial approach to assessing and demonstrating the performance, effectiveness, and innovation of the ensemble models in this study. By contrasting with the interpretability outcomes presented in Table 1 of existing studies, it is evident that the ensemble models in this research exhibit high interpretability, unveiling underlying relationships among the primary features. This underscores the novelty of the current study. In addition, benchmarking the model results against those in existing literature enables the validation of the newly developed model's performance and determines its superiority. This comparison aids in establishing the efficacy and practicality of the model. Hasanipanah et al. (2017) estimated FD using various AI techniques, achieving R² values between 0.86 and 0.96. Faradonbeh et al. (2016a) evaluated the FD problem using the GP method and attained an accuracy of 0.908. Both Faradonbeh et al. (2018) and Nikafshan Rad et al. (2020) employed the gene expression programming and genetic algorithm artificial neural network approaches, yielding R² values of 0.924–0.944. In the present study, whether employing the POA-RF or LSO-RF model for FD prediction across the two samples, the accuracy surpassed 0.95. This clearly indicates the commendable performance of the ensemble models.

5.4 Sensitivity Analysis Based on Ensemble Models

Ensemble models with high complexity and robust predictive performance often suffer from reduced interpretability (Zhou et al. 2022a, b; Cao et al. 2018). To interpret the FD evaluation results of the RF model and quantify the influence

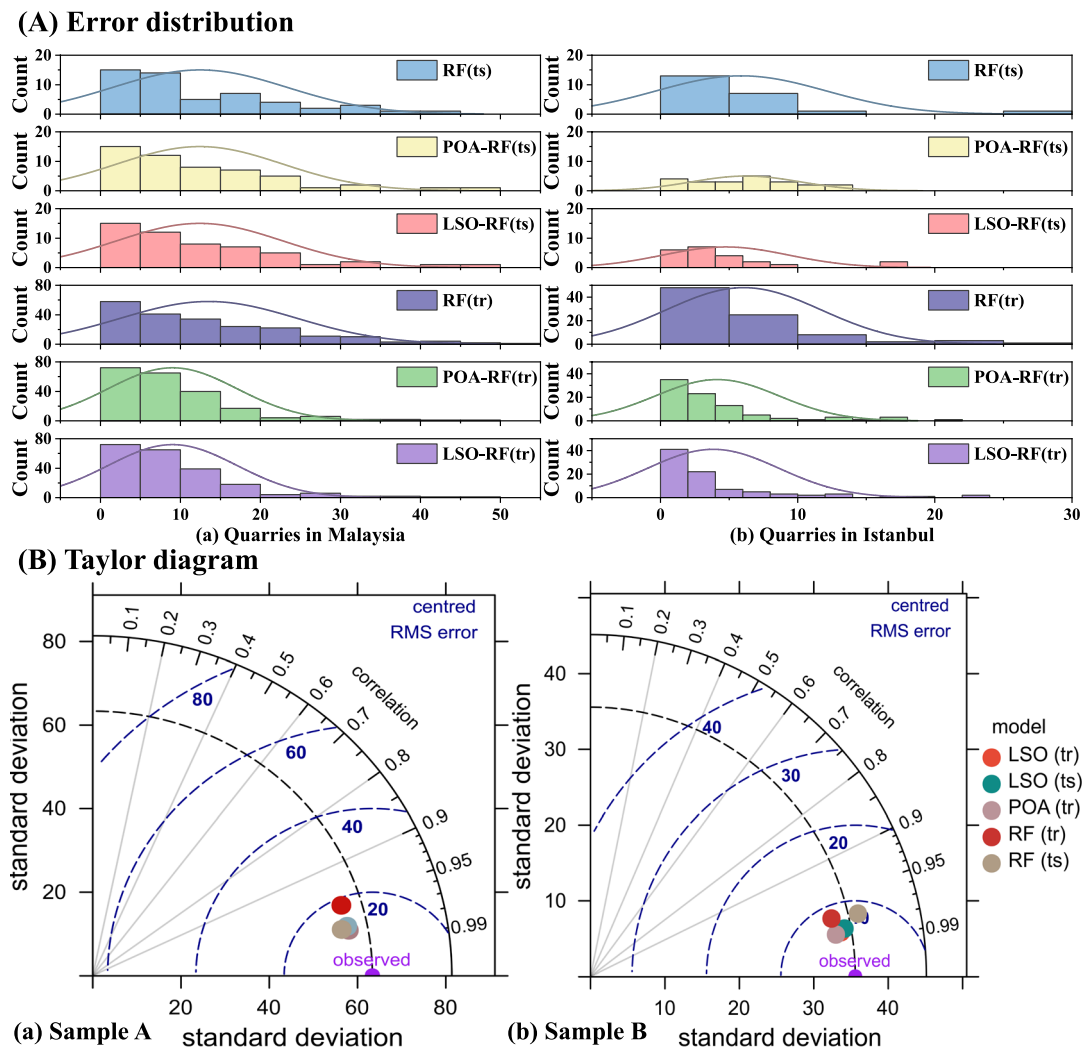
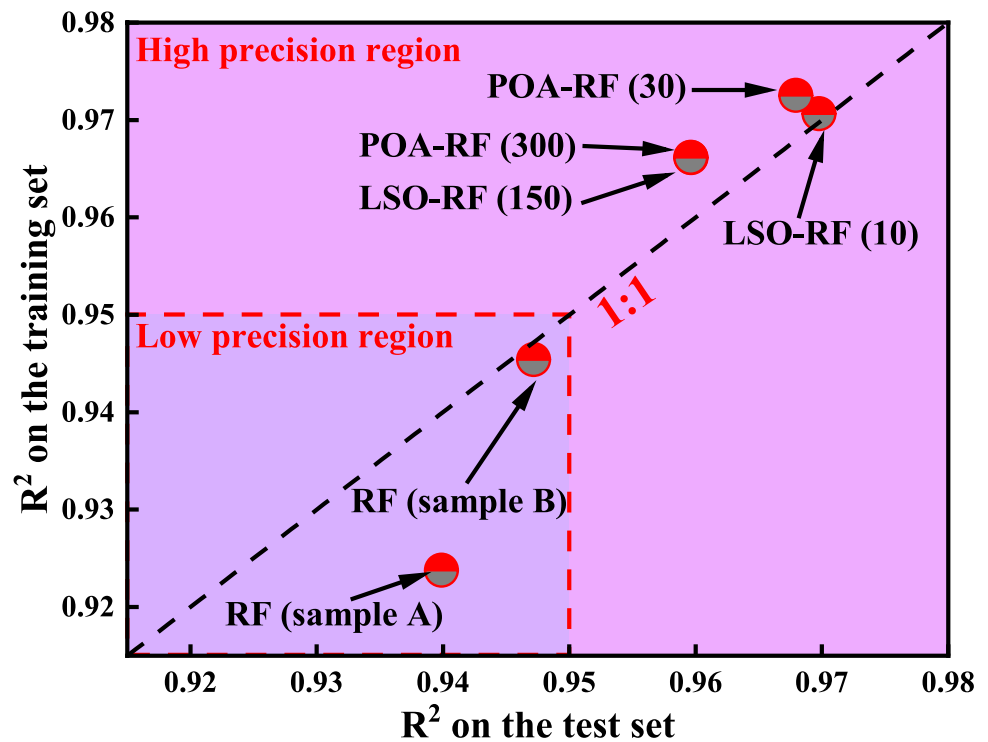


Fig. 11 Statistical chart of error distribution and Taylor chart for all models

of the pertinent input parameters, this study employed the SHapley Additive exPlanations (SHAP) method. Shapley is a concept derived from cooperative game theory, and SHAP is an interpretability technique developed based on this principle. Notably, SHAP can directly illustrate the impact of the cumulative feature attribute values in the input sample on the target variable (Zhou et al. 2021a; Qiu and Zhou 2023a, b). In other words, it can elucidate the contribution of each feature value to the model's output metrics. Furthermore, by examining the SHAP values, researchers can discern to some extent whether the contribution has a positive or negative effect. In addition, after identifying the most influential features affecting the target variable, the SHAP method can effectively capture the interactions between these features based on Shapley values. This approach unveils concealed relationships among features, thereby augmenting the local interpretability of the model.

In this section, the regression features of the two FD event datasets are interpreted using Shapley values. Figure 13 illustrates the SHAP values of each feature in the two FD datasets, where positive values indicate a positive effect on the model evaluation, and negative values suggest a negative impact on FD estimation. In Fig. 13A, the features HD and CD exert the most substantial influence on the FD output results of Sample A events. In addition, as the hole diameter and max charge per delay increase, the positive effect on the FD prediction also escalates, implying a heightened risk of flyrock hazards. An augmentation in the hole diameter and max charge per delay could result in an increased amount of explosive and deeper blast hole, amplifying the blasting energy and potentially extending the FD range. This observation aligns with the real-world scenario. Examining Fig. 13B, it is evident that the parameters q and S/B hold the most significant sway over the FD output results of Sample B events. Specific charge and

Fig. 12 Overfitting analysis of all models

spacing-to-burden ratio positively influence the FD range. An escalation in q enhances the blasting energy, which might consequently augment the FD. Similarly, a critical parameter (S/B) in blasting engineering significantly influences flyrock behavior by governing explosive energy distribution, rock fragmentation, and fragment dynamics. A higher S/B ratio often leads to insufficient stress wave overlap, causing uneven energy release and localized overpressure that accelerates fragments into long-distance flyrock, while a lower ratio may induce premature gas breakthrough, generating high-velocity projectiles.

Figure 14 illustrates an interaction analysis of the most critical features in each sample. Observing Fig. 14a, a positive correlation is evident between the hole diameter and max charge per delay in predicting FD for Sample A. Specifically, an increase in the hole diameter positively influences the FD range estimation. However, this positive effect diminishes when the max charge per delay is relatively small. Overall, the combined influence of these two pivotal features in Sample A is distinctly positive. The overall trend depicted in Fig. 14b parallels that of Fig. 14a, where the two primary features positively impact FD prediction. The distinguishing factor, however, is the elevated values of the q and S/B in Sample B, both of which positively influence the FD prediction.

6 Design Optimization Feedback of Quarry Blasting Effects

To operationalize the iterative optimization of blasting parameters, an intelligent software tool was developed to integrate predictive models with a feedback-driven design framework. The software interface (Fig. 15) is structured to not only facilitate data input/output but also to visualize the analytical logic governing flyrock prediction and parameter optimization. Key components of the interface are aligned with the interpretability of model and decision-making workflow:

(1) Parameter Panel and Input Interface:

The left-hand section of the interface allows users to input both controllable parameters (e.g., HD, B/S, CD) and site-specific conditions (e.g., rock mass characteristics, bench geometry). These inputs are directly mapped to the features identified by the SHAP sensitivity analysis (Sect. 5.4), ensuring that critical factors like HD and q —which exert the most significant influence on FD—are prioritized in the input workflow.

(2) Function Panel and Model Configuration:

The central function panel enables users to select between hybrid models (LSO-RF/POA-RF) and configure optimization parameters (e.g., swarm size, iteration count). This

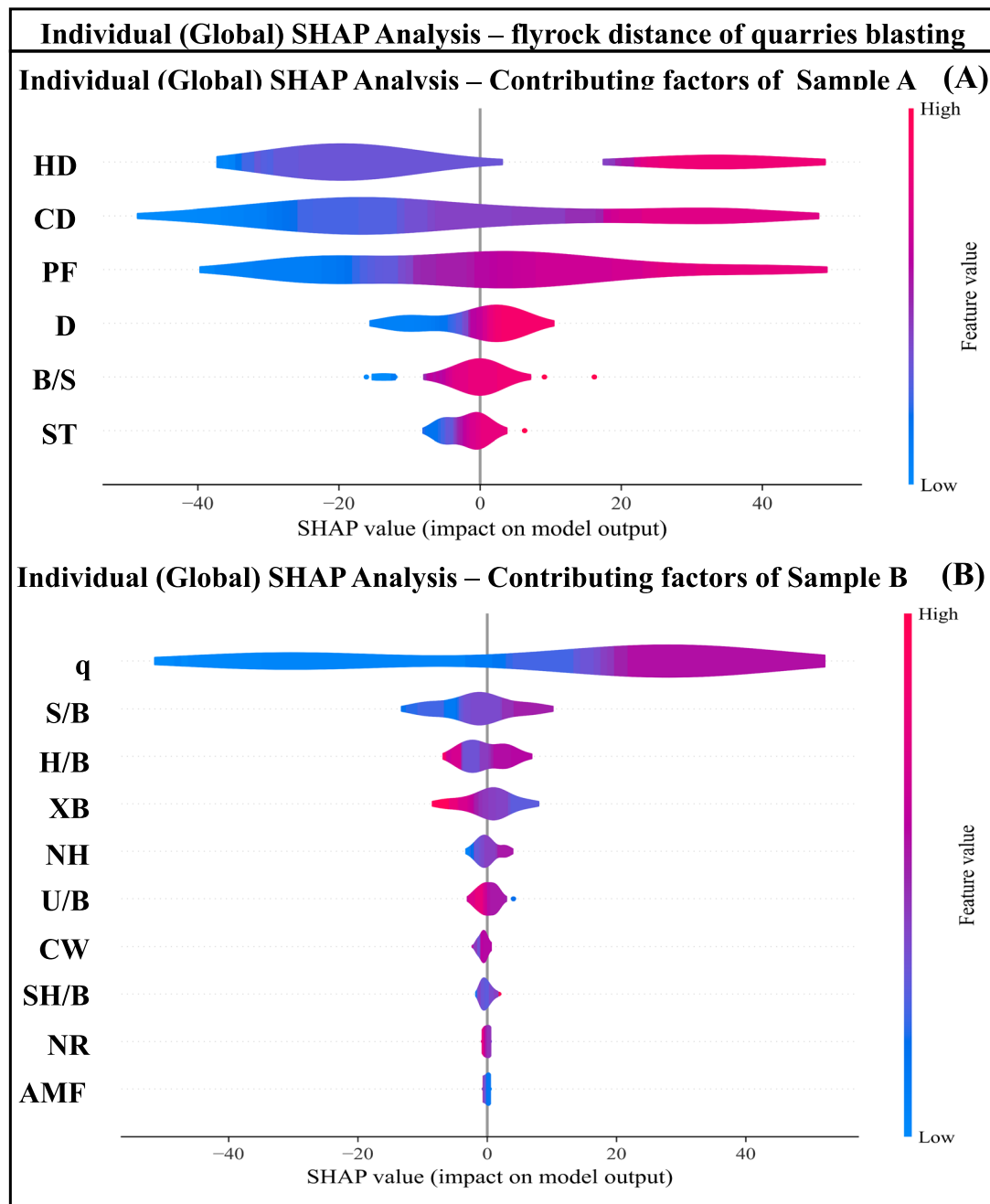


Fig. 13 Global interpretability for two samples based on SHAP values

panel explicitly reflects the hyperparameter optimization logic discussed in Sect. 3.4, where metaheuristic algorithms (LSO/POA) refine RF parameters ($n_estimators$, max_depth) to maximize prediction accuracy. The “Feature Importance” visualization, linked to SHAP values, dynamically updates to show how input parameters contribute to FD predictions, bridging the gap between raw data and model interpretability.

(3) Output Panel and Feedback Loop:

The right-hand output section displays predicted FD alongside evaluation metrics (R^2 , RMSE, MAPE) and fitness function curves, mirroring the model convergence analysis in Sect. 6.2. Crucially, it incorporates a “Feedback Optimization” module that suggests parameter adjustments based on predefined safety thresholds. For example, if predicted FD exceeds acceptable limits, the software

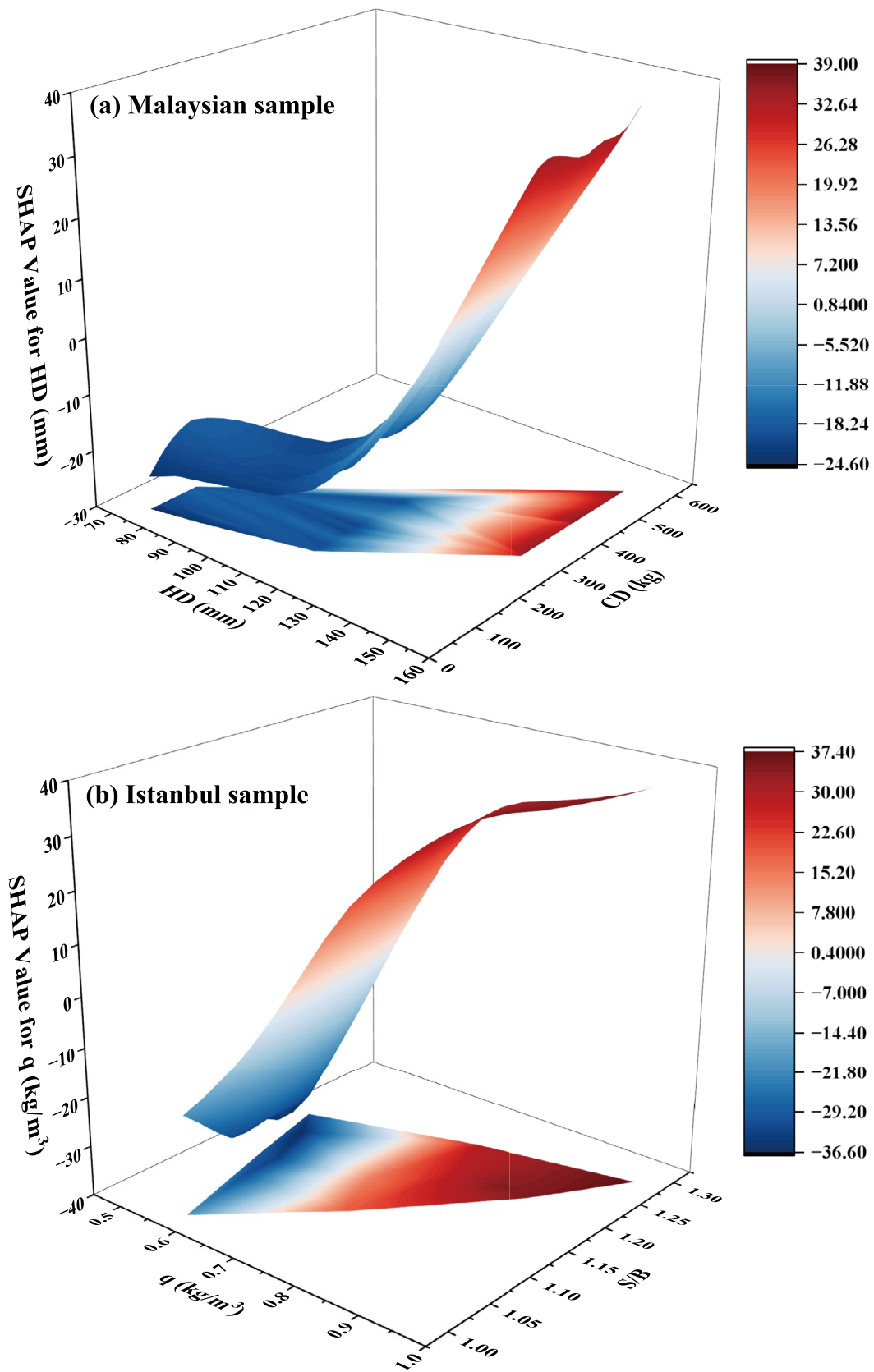


Fig. 14 Analyze the interactions between major variables based on SHAP values

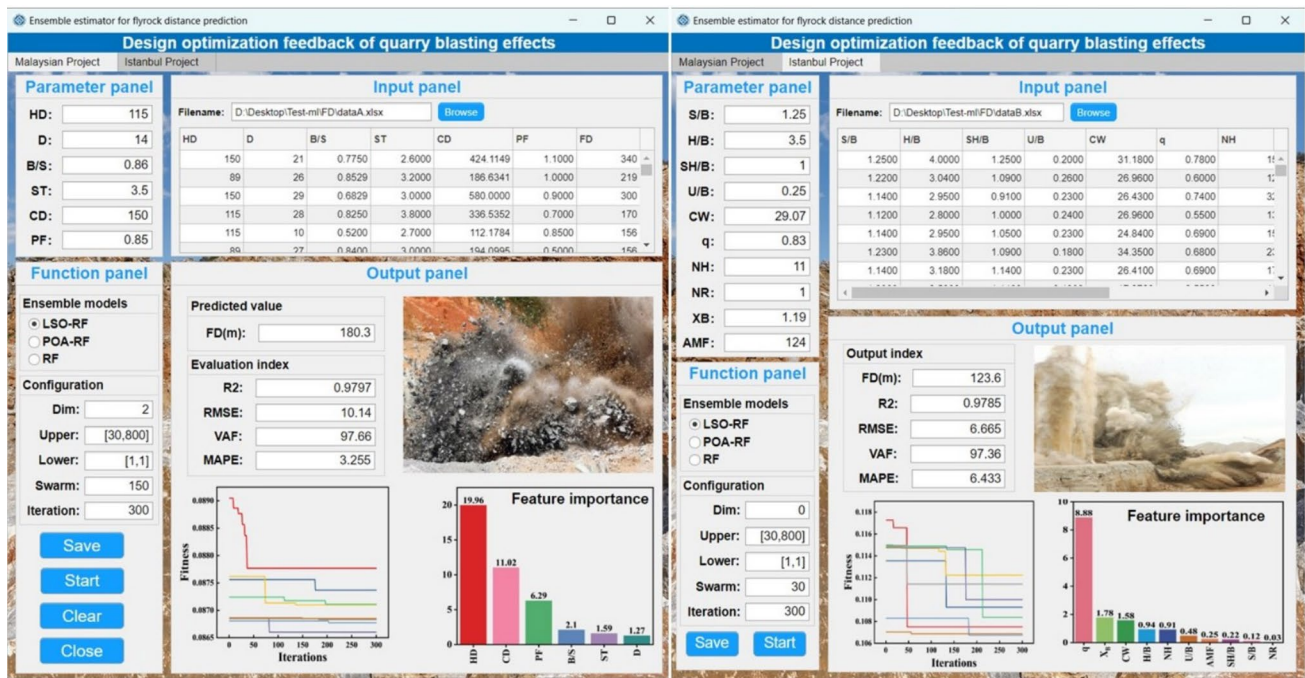
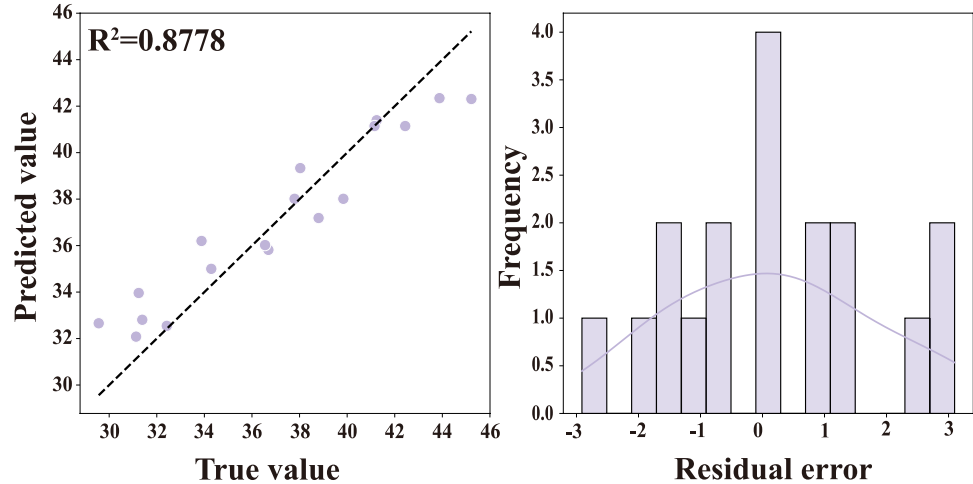


Fig. 15 Flyrock forecast optimization feedback software

Fig. 16 Engineering verification of blasting flyrock intelligent prediction model

Case verification



highlights influential parameters (e.g., reducing CD or increasing ST) using SHAP-based interactions (Fig. 14), guiding users to iteratively refine blasting designs. This creates a closed-loop system where model predictions inform parameter modifications, which are then re-fed into the model for validation, ensuring alignment with safety and efficiency goals.

(4) Workflow Integration:

The interface logic follows a structured analytical pipeline:

- Data import and preprocessing: Users upload blasting datasets (via “Browse”) that undergo Pearson correlation analysis (Sect. 2.3) to eliminate redundant features, as visualized in Fig. 3.
- Model training and validation: The “Save” function stores trained models with optimized hyperparameters (Table 6), while the “Start” button initiates predictions using the selected algorithm, generating outputs that reflect the convergence behavior of model (Fig. 10).
- Iterative refinement: The “Clear” and “Close” functions support workflow continuity, allowing users to test multiple parameter scenarios and compare model per-

formances (e.g., LSO-RF vs. POA-RF) across different blasting sites, as demonstrated in the cross-database validation (Sect. 5.2).

This design ensures that the software not only explicitly illustrates input–output relationships but also systematically integrates the analytical rigor of hybrid models—encompassing feature importance assessment, hyperparameter optimization, and feedback-driven adaptive mechanisms—thereby translating the study’s methodological advancements into practical utility within a user-centric operational framework.

7 Project Verification

Based on the production data and on-site measurement data from the Heshangqiao Quarry in Anhui Province, China, a comprehensive blasting flyrock database was established. The key parameters covered include, but are not limited to, blasting environment data such as hole diameter, hole spacing, hole depth, burden, stemming height of blast holes, charge length, toe burden, charge per hole, and Proctor coefficient. In this study, standard blasting design documents for the Heshangqiao Project of Ma Steel Nanshan Mine were systematically collected and organized weekly, with blasting design parameters input in batches in a timely manner to facilitate subsequent data retrieval and processing.

In open-pit mining blasting engineering, the accurate prediction of blasting flyrock is a core objective for ensuring safe production. Traditional empirical formulas, limited by the coupling effects of geological conditions and parameters, struggle to achieve prediction accuracy that meets the demands of complex engineering projects. This study, based on the preliminarily constructed intelligent prediction optimal hybrid model (POA-RF), takes the Heshangqiao Iron Mine in Ma'anshan, Anhui Province, as the engineering background. Through verification with measured data, the reliability and generalization ability of the model are validated, providing empirical support for the engineering application of intelligent prediction technologies. Figure 16 presents the engineering verification results of the blasting flyrock prediction model at Heshangqiao Iron Mine. Using the best-performing POA-RF model to predict flyrock distance, this study achieves an R^2 value of 0.8778, indicating a high degree of agreement between predicted and measured values. The residual distribution exhibits symmetric characteristics, primarily concentrated around 0, which demonstrates that the model provides consistent estimates of flyrock distance and can serve as a reliable basis for determining safety alert ranges. The engineering verification results confirm that the intelligent prediction model

proposed in this study demonstrates excellent performance in real-world mining scenarios.

8 Discussion and Limitation Analysis

The advancement and application of ensemble models, notably LSO-RF and POA-RF, have substantially enhanced the predictive accuracy of FD in quarry blasting. The incorporation of metaheuristic optimization algorithms into the RF model has markedly improved FD’s predictive performance. Comparative analysis indicates that both LSO-RF and POA-RF consistently outperform the unoptimized RF model across two datasets, achieving R^2 values exceeding 0.95. These results underscore not only the effectiveness and robustness of the proposed ensemble models but also their superior performance compared to existing AI-based methods in the literature. The sensitivity analysis yielded valuable insights into the relative importance of various input parameters for FD prediction. In Sample A, HD and CD were identified as the most influential factors, with HD being the predominant predictor. In contrast, for Sample B, the q and B/S were deemed pivotal parameters. These findings not only validate the models’ consistency and reliability but also emphasize the importance of considering these key parameters in blast design and optimization. In addition, the SHAP interpretability analysis clarified the relationships and interactions among the input variables, enhancing the model’s local interpretability and providing deeper insights into the factors influencing FD prediction.

However, the study has several limitations that warrant discussion. First, the reliance on two specific datasets from open-pit mines in Malaysia and Istanbul may constrain the generalizability of the proposed models to other geological and operational settings. Future research should validate the models using more diverse and extensive datasets to enhance their applicability and robustness. Second, the optimization algorithms LSO and POA, although effective, require meticulous tuning of population size and hyperparameter range, potentially escalating computational complexity and time consumption. Further research is needed to devise more efficient and user-friendly optimization strategies to streamline the model training process. Lastly, the feedback design employed in the intelligent flyrock prediction software, while innovative, necessitates rigorous validation in real-world blasting scenarios to ensure its practicality and efficacy in optimizing blasting parameters and effectively mitigating flyrock hazards.

In conclusion, the proposed ensemble models, coupled with advanced optimization techniques and interpretability analysis, have significantly elevated the accuracy and reliability of FD prediction in blasting engineering. Nonetheless, to fully leverage the potential of these models and

the intelligent feedback design to enhance the safety and efficiency of quarry blasting, ongoing research efforts are essential to address the identified limitations and validate the models' performance across a broader range of operational contexts.

9 Conclusion

In the realm of mining engineering, blasting technology stands as a pivotal engineering technique, integral to ore extraction and underground construction projects. Nonetheless, conventional blasting techniques frequently face limitations stemming from inadequate data acquisition and quality evaluation, which can result in unstable blasting effects and associated safety hazards. This study endeavors to leverage advanced methods and technologies to predict FD, establish a comprehensive analysis platform for evaluating blasting effects, and develop a feedback function for optimizing blasting design based on this platform. The objective is to enhance efficiency, mitigate risks, and furnish decision-makers with more comprehensive and reliable information.

In this study, ensemble models, specifically LSO-RF and POA-RF, were developed and applied to improve the prediction accuracy of FD in quarry blasting. By integrating metaheuristic optimization algorithms with the RF model, the predictive performance of FD was markedly enhanced, achieving R^2 values exceeding 0.95 across two datasets from open-pit mines in Malaysia and Istanbul. The sensitivity analysis and SHAP interpretability analysis offered valuable insights into the relative importance of various input parameters and clarified the relationships and interactions among the input variables. This enriched the model's local interpretability and deepened the understanding of the factors influencing FD prediction.

An intelligent flyrock prediction software, equipped with a feedback design, was developed to continuously optimize blasting parameters, effectively mitigating the hazards associated with flyrock and ensuring optimal blasting outcomes. However, the study identified several limitations, including the restricted generalizability of the models to diverse geological and operational contexts, the computational complexity associated with the optimization algorithms, and the necessity for rigorous validation of the feedback design in real-world blasting scenarios.

In conclusion, the proposed ensemble models, integrated with advanced optimization techniques and interpretability analysis, have substantially improved the accuracy and reliability of FD prediction in blasting engineering. To fully leverage the potential of these models and the intelligent feedback design in enhancing the safety and efficiency of quarry blasting, ongoing research efforts are essential to

address the identified limitations and validate the models' performance across a broader range of operational contexts.

Acknowledgements This research is partially supported by the National Natural Science Foundation of China (52474121, 42177164) and the Outstanding Youth Project of Hunan Provincial Department of Education (23B0008).

Author contribution Yulin Zhang: Data curation; Formal analysis; Investigation; Methodology; Software; Visualization; Visualization; Writing—original draft. Yingui Qiu: Formal analysis; Investigation; Methodology; Writing—review & editing. Kun Du: Investigation; Writing—review & editing. Hoang Nguyen: Formal analysis; Writing—review & editing. Danial Jahed Armaghani: Data curation; Writing—review & editing. Jian Zhou: Conceptualization; Methodology; Validation; Funding acquisition; Supervision; Writing—review & editing.

Data availability The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Conflict of 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.

References

- Abdel-Basset M, Mohamed R, Sallam KM, Chakraborty RK (2022) Light spectrum optimizer: a novel physics-inspired metaheuristic optimization algorithm. *Mathematics* 10(19):3466
- Abdollahzadeh, B., Khodadadi, N., Barshandeh, S., Trojovský, P., Gharehchopogh, F. S., El-kenawy, E. S. M., & Mirjalili, S. (2024). Puma optimizer (PO): A novel metaheuristic optimization algorithm and its application in machine learning. *Cluster Computing*, 1–49.
- Adam JA (2002) The mathematical physics of rainbows and glories. *Phys Rep* 356(4–5):229–365
- Armaghani DJ, Hajihassani M, Mohamad ET, Marto A, Noorani SA (2014) Blasting-induced flyrock and ground vibration prediction through an expert artificial neural network based on particle swarm optimization. *Arab J Geosci* 7:5383–5396
- Armaghani DJ, Mahdiyar A, Hasanipanah M, Faradonbeh RS, Khandelwal M, Amnieh HB (2016) Risk assessment and prediction of flyrock distance by combined multiple regression analysis and Monte Carlo simulation of quarry blasting. *Rock Mech Rock Eng* 49:3631–3641
- 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 Env* 79(8):4369–4385
- Armaghani, D. J., Harandizadeh, H., Momeni, E., Maizir, H., & Zhou, J. (2022). An optimized system of GMDH-ANFIS predictive model by ICA for estimating pile bearing capacity. *Artificial Intelligence Review*, 1–38.
- Bakhtavar E, Nourizadeh H, Sahebi AA (2017) Toward predicting blast-induced flyrock: a hybrid dimensional analysis fuzzy inference system. *Int J Environ Sci Technol* 14:717–728
- Belgiu M, Drăguț L (2016) Random forest in remote sensing: a review of applications and future directions. *ISPRS J Photogramm Remote Sens* 114:24–31

- Bhandari, S. (1997). Engineering rock blasting operations.
- Biau G, Scornet E (2016) A random forest guided tour. *TEST* 25:197–227
- Bui XN, Nguyen H, Le HA, Bui HB, Do NH (2020) Prediction of blast-induced air over-pressure in open-pit mine: assessment of different artificial intelligence techniques. *Nat Resour Res* 29(2):571–591
- Cao Y, Ji H, Zhang T (2018) Blasting control and monitoring system for safety improvement during blasting operation. A case study in Guilaizhuang gold mine. *Mining Science*, 25.
- Chen Y, Yong W, Li C, Zhou J (2023) Predicting the Thickness of an Excavation Damaged Zone around the Roadway Using the DA-RF Hybrid Model. *CMES-Computer Modeling in Engineering & Sciences*, 136(3).
- Dai Y, Khandelwal M, Qiu Y, Zhou J, Monjezi M, Yang P (2022). A hybrid metaheuristic approach using random forest and particle swarm optimization to study and evaluate backbreak in open-pit blasting. *Neural Computing and Applications*, 1–16.
- Drake JH, Özcan E, Burke EK (2012) An improved choice function heuristic selection for cross domain heuristic search. In *Parallel Problem Solving from Nature-PPSN XII: 12th International Conference, Taormina, Italy, September 1-5, 2012, Proceedings, Part II* 12 (pp. 307–316). Springer Berlin Heidelberg.
- Faradonbeh RS, Jahed Armaghani D, Monjezi M (2016a) Development of a new model for predicting flyrock distance in quarry blasting: a genetic programming technique. *Bull Eng Geol Env* 75:993–1006
- Faradonbeh RS, Armaghani DJ, Monjezi M, Mohamad ET (2016b) Genetic programming and gene expression programming for flyrock assessment due to mine blasting. *Int J Rock Mech Min Sci* 88:254–264
- Faradonbeh RS, Armaghani DJ, Amnieh HB, Mohamad ET (2018) Prediction and minimization of blast-induced flyrock using gene expression programming and firefly algorithm. *Neural Comput Appl* 29:269–281
- Fattahi H, Hasanipanah M (2022) An integrated approach of ANFIS-grasshopper optimization algorithm to approximate flyrock distance in mine blasting. *Eng Comput* 38(3):2619–2631
- Ghasemi E, Sari M, Ataei M (2012) Development of an empirical model for predicting the effects of controllable blasting parameters on flyrock distance in surface mines. *Int J Rock Mech Min Sci* 52:163–170
- Ghasemi E, Amini H, Ataei M, Khalokakaie R (2014) Application of artificial intelligence techniques for predicting the flyrock distance caused by blasting operation. *Arab J Geosci* 7:193–202
- Ghojoghi E, Farsangi MAE, Mansouri H, Rashedi E (2024) Prediction and minimization of blasting flyrock distance, using deep neural networks and gravitational search algorithm, JAYA, and multi-verse optimization algorithms. *Heliyon*, 10(19).
- Guo H, Zhou J, Koopialipoor M, Jahed Armaghani D, Tahir MM (2021) Deep neural network and whale optimization algorithm to assess flyrock induced by blasting. *Engineering with Computers* 37:173–186
- Gupta RN (1980) Surface blasting and its impact on environment. Impact of mining on environment. Ashish Publishing House, New Delhi, pp 23–24
- Han H, Jahed Armaghani D, Tarinejad R, Zhou J, Tahir MM (2020) Random forest and bayesian network techniques for probabilistic prediction of flyrock induced by blasting in quarry sites. *Nat Resour Res* 29:655–667
- Hasanipanah M, Faradonbeh RS, Armaghani DJ, Amnieh HB, Khandelwal M (2017) Development of a precise model for prediction of blast-induced flyrock using regression tree technique. *Environmental Earth Sciences* 76:1–10
- Hudaverdi T (2022) Prediction of flyrock throw distance in quarries by variable selection procedures and ANFIS modelling technique. *Environmental Earth Sciences* 81(10):281
- Hudaverdi T, Agan Y (2023) Application of intelligent models for flyrock prediction considering design parameters and bench face characteristics. *Mining, Metallurgy & Exploration* 40(6):2331–2347
- Hudaverdi T, Akyildiz O (2019) A new classification approach for prediction of flyrock throw in surface mines. *Bull Eng Geol Env* 78:177–187
- Jahed Armaghani D, Hajihassani M, Monjezi M, Mohamad ET, Marto A, Moghaddam MR (2015) Application of two intelligent systems in predicting environmental impacts of quarry blasting. *Arab J Geosci* 8:9647–9665
- Jiang L, Zhang Z, Lu L, Shang X, Wang W (2024) Nonparametric modelling of ship dynamics using puma optimizer algorithm-optimized twin support vector regression. *J Marine Sci Eng* 12(5):754
- Khandelwal M, Monjezi M (2013) Prediction of backbreak in open-pit blasting operations using the machine learning method. *Rock Mech Rock Eng* 46:389–396
- Koopialipoor M, Fallah A, Armaghani DJ, Azizi A, Mohamad ET (2019) Three hybrid intelligent models in estimating flyrock distance resulting from blasting. *Eng Comput* 35:243–256
- Lang L, Tiancai L, Shan A, Xiangyan T (2021) An improved random forest algorithm and its application to wind pressure prediction. *Int J Intell Syst* 36(8):4016–4032
- Li C, Zhou J, Du K, Armaghani DJ, Huang S (2023) Prediction of flyrock distance in surface mining using a novel hybrid model of harris hawks optimization with multi-strategies-based support vector regression. *Nat Resour Res* 32(6):2995–3023
- Little TN, Blair DP (2010) Mechanistic Monte Carlo models for analysis of flyrock risk. *Rock fragmentation by blasting*. <https://www.tnrc.com.au/wp-content/uploads/2014/02/LittleBlairFlyrock.pdf>
- Lu X, Hasanipanah M, Brindhadevi K, Bakhshandeh Amnieh H, Khalafi S (2020) ORELM: a novel machine learning approach for prediction of flyrock in mine blasting. *Nat Resour Res* 29:641–654
- Lundborg N, Persson A, Ladegaard-Pedersen A, Holmberg R (1975) Keeping the lid on flyrock in open-pit blasting. *Eng Min J* 176:95–100
- Marto A, Hajihassani M, Jahed Armaghani D, Tonnizam Mohamad E, Makhtar AM (2014) A novel approach for blast-induced flyrock prediction based on imperialist competitive algorithm and artificial neural network. *Sci World J* 2014(1):643715
- Maurya P, Tiwari P, Pratap A (2024) Puma optimizer technique for optimal planning of different types of distributed generation units in radial distribution network considering different load models. *Electrical Engineering*, 1–52.
- McKenzie CK (2009) Flyrock range and fragment size prediction. <http://docs.isee.org/ISEE/Support/Proceed/General/09GENV2/09v206g.pdf>
- Mohamad ET, Armaghani DJ, Noorani SA, Saad R, Alvi SV, Abad NK (2012) Prediction of flyrock in boulder blasting using artificial neural network. *Electron J Geotech Eng* 17:2585–2595
- Monjezi M, Bahrami A, Varjani AY (2010) Simultaneous prediction of fragmentation and flyrock in blasting operation using artificial neural networks. *Int J Rock Mech Min Sci* 47(3):476–480
- Monjezi, M., Amini Khoshalan, H., & Yazdian Varjani, A. (2012). Prediction of flyrock and backbreak in open pit blasting operation: a neuro-genetic approach. *Arabian Journal of Geosciences*, 5(3).
- Murlidhar BR, Nguyen H, Rostami J, Bui X, Armaghani DJ, Ragam P, Mohamad ET (2021) Prediction of flyrock distance induced by mine blasting using a novel Harris Hawks optimization-based multi-layer perceptron neural network. *J Rock Mech Geotech Eng* 13(6):1413–1427
- Nikafshan Rad H, Bakhshayeshi I, Wan Jusoh WA, Tahir MM, Foong LK (2020) Prediction of flyrock in mine blasting: a new computational intelligence approach. *Nat Resour Res* 29:609–623
- Pravesjit, S., & Kantawong, K. (2025). Enhancing the Puma Optimizer Algorithm for Optimization Problems. In *2025 Joint International*

- Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON) (pp. 583–587). IEEE.
- Qasem SN, Samadianfard S, Kheshtgar S, Jarhan S, Kisi O, Shamshirband S, Chau KW (2019) Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural networks in arid and humid climates. *Eng Appl Comput Fluid Mech* 13(1):177–187
- Qiu Y, Zhou J (2023a) Short-term rockburst damage assessment in burst-prone mines: an explainable XGBOOST hybrid model with SCSO algorithm. *Rock Mech Rock Eng* 56(12):8745–8770
- Qiu Y, Zhou J (2023b) Short-term rockburst prediction in underground project: insights from an explainable and interpretable ensemble learning model. *Acta Geotech* 18(12):6655–6685
- Qiu Y, Zhou J (2024) Novel rockburst prediction criterion with enhanced explainability employing CatBoost and nature-inspired metaheuristic technique. *Underground Space* 19:101–118
- Qiu Y, Zhou J, Khandelwal M, Yang H, Yang P, Li C (2022) Performance evaluation of hybrid WOA-XGBoost, GWO-XGBoost and BO-XGBoost models to predict blast-induced ground vibration. *Eng Comput* 38:4145–4162
- Raina AK, Murthy VMSR (2016) Prediction of flyrock distance in open pit blasting using surface response analysis. *Geotech Geol Eng* 34:15–28
- Rezaei M, Monjezi M, Varjani AY (2011) Development of a fuzzy model to predict flyrock in surface mining. *Saf Sci* 49(2):298–305
- Rezaei M, Monjezi M, Matinpour F, Bolbanabad SM, Habibi H (2023) Simulation of induced flyrock due to open-pit blasting using the PCA-CART hybrid modeling. *Simul Model Pract Theory* 129:102844
- Rezaei, M., Ahmadi, S. R., Nguyen, H., & Armaghani, D. J. (2024). Improved determination of the S-wave velocity of rocks in dry and saturated conditions: Application of machine-learning algorithms. *Transportation Geotechnics*, 101371.
- Rigatti SJ (2017) Random forest. *J Insur Med* 47(1):31–39
- Saber S, Salem S (2023) An improved light Spectrum optimizer for parameter identification of triple-diode PV model. *Sustain Mach Intell J* 4:5–1
- Shakeri J, Amini Khoshalan H, Dehghani H, Bascompta M, Onyelowe K (2022) Developing new models for flyrock distance assessment in open-pit mines. *J Mining Environ* 13(2):375–389
- Singh PK, Roy MP, Paswan RK, Sarim MD, Kumar S, Jha RR (2016) Rock fragmentation control in opencast blasting. *J Rock Mech Geotech Eng* 8(2):225–237
- Speiser JL, Miller ME, Tooze J, Ip E (2019) A comparison of random forest variable selection methods for classification prediction modeling. *Expert Syst Appl* 134:93–101
- Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. *J Geophys Res* 106(D7):7183–7192
- Thulasi T, Sivamohan K (2023) LSO-CSL: Light spectrum optimizer-based convolutional stacked long short term memory for attack detection in IoT-based healthcare applications. *Expert Syst Appl* 232:120772
- Trivedi R, Singh TN, Raina AK (2014) Prediction of blast-induced flyrock in Indian limestone mines using neural networks. *J Rock Mech Geotech Eng* 6(5):447–454
- Wang Z, Zhou J, Du K, Khandelwal M (2024) Enhanced multi-task learning models for pile drivability prediction: Leveraging metaheuristic algorithms and statistical evaluation. *Transp Geotech* 47:101288
- Wang X, Gu L, Lin Y, Mao Z, Yuan Y, Niu C, Yang Q (2025) Advanced multispectral thermometry method based on an improved light spectrum optimizer for reliable temperature and emissivity measurement. *Opt Express* 33(6):13594–13612
- Yang P, Yong W, Li C, Peng K, Wei W, Qiu Y, Zhou J (2023a) Hybrid random forest-based models for earth pressure balance tunneling-induced ground settlement prediction. *Appl Sci* 13(4):2574
- Yang P, Li C, Qiu Y, Huang S, Zhou J (2023b) Metaheuristic optimization of random forest for predicting punch shear strength of FRP-reinforced concrete beams. *Materials* 16(11):4034
- Yari M, Bagherpour R, Jamali S, Shamsi R (2016) Development of a novel flyrock distance prediction model using BPNN for providing blasting operation safety. *Neural Comput Appl* 27:699–706
- Yari M, Armaghani DJ, Maraveas C, Ejlali AN, Mohamad ET, Asteris PG (2023) Several tree-based solutions for predicting flyrock distance due to mine blasting. *Appl Sci* 13(3):1345
- Yu H, Xu F, Tropea C (2013) Optical caustics associated with the primary rainbow of oblate droplets: simulation and application in non-sphericity measurement. *Opt Express* 21(22):25761–25771
- Zaki DA, Hasanien HM, Alharbi M, Sun C (2024) Precise three-diode photovoltaic model for photovoltaic modules based on Puma optimizer. *Ain Shams Eng J* 15(12):103170
- Zhang YL, Qiu YG, Armaghani DJ, Monjezi M, Zhou J (2024) Enhancing rock fragmentation prediction in mining operations: a hybrid GWO-RF model with SHAP interpretability. *J Central South Univ* 31(8):2916–2929
- Zhou J, Aghili N, Ghaleini EN, Bui DT, Tahir MM, Koopialipoor M (2020) A Monte Carlo simulation approach for effective assessment of flyrock based on intelligent system of neural network. *Eng Comput* 36:713–723
- Zhou J, Dai Y, Khandelwal M, Monjezi M, Yu Z, Qiu Y (2021a) Performance of hybrid SCA-RF and HHO-RF models for predicting backbreak in open-pit mine blasting operations. *Nat Resour Res* 30:4753–4771
- Zhou J, Qiu Y, Khandelwal M, Zhu S, Zhang X (2021b) Developing a hybrid model of Jaya algorithm-based extreme gradient boosting machine to estimate blast-induced ground vibrations. *Int J Rock Mech Min Sci* 145:104856
- Zhou J, Huang S, Zhou T, Armaghani DJ, Qiu Y (2022a) Employing a genetic algorithm and grey wolf optimizer for optimizing RF models to evaluate soil liquefaction potential. *Artif Intell Rev* 55(7):5673–5705
- Zhou J, Huang S, Qiu Y (2022b) Optimization of random forest through the use of MVO, GWO and MFO in evaluating the stability of underground entry-type excavations. *Tunn Undergr Space Technol* 124:104494
- Zhou J, Zhang Y, Li C, Yong W, Qiu Y, Du K, Wang S (2023a) Enhancing the performance of tunnel water inflow prediction using random forest optimized by grey wolf optimizer. *Earth Sci Inf* 16(3):2405–2420
- Zhou J, Wang Z, Li C, Wei W, Wang S, Armaghani DJ, Peng K (2023b) Hybridized random forest with population-based optimization for predicting shear properties of rock fractures. *J Comput Sci* 72:102097
- Zhou J, Zhang Y, Qiu Y (2024a) State-of-the-art review of machine learning and optimization algorithms applications in environmental effects of blasting. *Artif Intell Rev* 57(1):1–54
- Zhou J, Yang P, Yong W, Khandelwal M, Huang S (2024b) A comparative analysis of hybrid RF models for efficient lithology prediction in hard rock tunneling using TBM working parameters. *Acta Geophys* 72(3):1847–1866

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.