Geomechanics and Engineering, *Vol. 39, No. 6 (2024) 547-561*
https://doi.org/10.12989/gae.2024.39.6.547

Evaluating blast-induced backbreak in open pit mines using the LSSVM optimized by the GWO algorithm

Niaz Muhammad Shahani[1,2], Xigui Zheng*[1,2,3,4], Patrick Siarry[5], Danial Jahed Armaghani[6] and Cancan Liu[1]

[1]School of Mines, China University of Mining and Technology, Xuzhou, 221116, Jiangsu Province, China
[2]Shanxi Guxian Jingu Coal Industry Co., Ltd Linfen, 041000, Shanxi Province, China
[3]School of Mines and Civil Engineering, Liupanshui Normal University, Liupanshui 553004, China
[4]Guizhou Guineng Investment Co., Ltd., Liupanshui 553600, China
[5]Université Paris-Est Creteil, LiSSi Lab, France
[6]School of Civil and Environmental Engineering, University of Technology Sydney, NSW, 2007, Australia

(*Received April 18, 2024, Revised November 25, 2024, Accepted December 12, 2024*)

Abstract. Backbreak, a recurring issue in blasting operations, causes mine wall instability, equipment failure, inappropriate disintegration, lower drilling efficiency, and increased cost of mining operations. This study aims to address these issues by developing a hybrid LSSVM-GWO model for predicting blast-induced backbreak in open pit mines. To evaluate the effectiveness of the proposed model, its predictive performance was compared with three convolutional models, such as the support vector machine, K-nearest neighbor, and the least square support vector machine. Results demonstrated that the LSSVM-GWO model outperformed the other three models, achieving coefficient of determination values of 0.998 and 0.997, mean absolute error values of 0.0068 and 0.1209, root mean squared error values of 0.0825 and 0.1936, and *a*20-index values of 0.99 and 1.01 for training and testing datasets, respectively. Furthermore, the SHAP machine learning technique was applied to evaluate the feature importance, revealing that the powder factor had the highest influence, while the burden exhibited the least impact on backbreak. Sensitivity analysis confirmed these findings, highlighting the robustness of the hybrid model. The study concludes that the LSSVM-GWO model significantly enhances the prediction and evaluation of backbreak in open pit mines, providing critical insights to improve blasting operations, reduce costs, and ensure mine safety.

Keywords: backbreak; blasting environmental issue; fracture mechanics; LSSVM- GWO; open-pit mines

1. Introduction

Blasting is widely considered the key technique for the disintegration of hard rock in mining activities, and the expense of rock mass degradation caused by blasting is gradually being acknowledged concerning mining efficiency and safety. The adverse consequences of blasting are unavoidable and cannot be entirely removed, although they can be reduced to a manageable level to prevent environmental harm. Backbreak in open pit mining is one of the adverse impacts that planners, designers, and environmentalists are concerned about. Backbreak is fragmented rocks extending beyond the blast pattern's array of holes (EL *et al*. 1995). Backbreak is an undesirable phenomenon in blasting operations that stems from mine wall instability, equipment failure, inappropriate disintegration, lower drilling efficiency, and other issues that raise the entire cost of a mining operation (Khandelwal *et al.* 2013). Backbreak caused by blasting can have a significant influence on slope stability. Bauer (1982) observed that uncontrolled backbreak led to a reduction in the overall pit slope angle, which would be necessary, leading to an increase in the stripping ratio. There will be more loose-face rock, and safety berms will be less productive. Total production costs will increase significantly due to the adverse effects of backbreak (Khandelwal *et al*. 2013, Scoble *et al*. 1997, Sayadi *et al*. 2013, Ebrahimi *et al*. 2016, Monjezi *et al*. 2010a, Ghasemi *et al*. 2016, Ghasemi 2017, Agrawal and Mishra 2018). Several studies have been conducted by different researchers to uncover the parameters that may impact the incidence of backbreak, as well as proposed preventive measures. When the burden or stemming rises, backbreak rises (Konya and Walter 1991, Kumar *et al*. 2021, Sharma *et al*. 2021, Monjezi and Dehghani 2008). According to (Ebrahimi *et al*. 2016), the length of the stemming, the burden, the spacing, and the depth of the hole are all important factors in generating a backbreak. Backbreak would spread because of the poor stiffness ratio and heavy burden (EL *et al*. 1995).

Previously, researchers attempted to use various traditional machine learning (ML) models, such as multiple regression analysis (MRA) (Khandelwal *et al*. 2013, Monjezi *et al*. 2010b), artificial neural networks (ANN) (Sayadi *et al*. 2013, Agrawal and Mishra 2018, Monjezi and Dehghani 2008, Esmaeili *et al*. 2012, Monjezi *et al*. 2013, Monjezi *et al*. 2014), neurogenetic method ANFIS (adaptive neuro-fuzzy inference system), as well as more advanced

*Corresponding author, Professor
E-mail: cumt_ckzxg@126.com

ISSN: 2005-307X (Print), 2092-6219 (Online)

techniques such as support vector machine (SVM) (Khandelwal *et al*. 2013, Mohammadnejad *et al*. 2013, Monjezi *et al*. 2010b, Kamran *et al*. 2022), extreme gradient boosting (XGBoost) (Du *et al.* 2024), fuzzy set theory (Monjezi *et al*. 2010b), and genetic programming (GP) (Saghatforoush *et al*. 2016, Mahmoodzadeh *et al*. 2024) to evaluate the backbreak. These models have shown varying levels of success in predicting backbreak, with hybrid models that combine ML with optimization algorithms, including the bee colony algorithm (BCA) (Sayadi *et al*. 2013), ant colony optimization (ACO), and the relevance vector regression-invasive weed optimization algorithm (RVR-IWO) (Fattahi and Bayat 2024), further enhancing model effectiveness. Similarly, (Faradonbeh *et al*. 2016) employed non-linear multiple linear regression (NLMR) and GP approaches for predicting backbreak with 35 datasets with input parameters of burden in m; spacing in m; stemming in m; powder factor in kg/m^3 and stiffness ratio in m/m. On the test dataset, the highest R^2 for NLMR prediction was 0.851 and 0.976 for GP. Likewise, (Sharma *et al*. 2022) used multivariate regression analysis (MVRA) and GP for backbreak prediction on 84 datasets, of which 56 datasets were utilized for train, 14 datasets for test, and 14 datasets for validation purposes. On the test dataset, the results exhibited the highest R^2 of 0.873 for MVRA prediction and 0.979 for GP prediction of Sungun Copper Mine; the highest R^2 of 0.63 for MVRA prediction and 0.816 for GP prediction of ASP Colliery. (Zhou *et al*. 2021) implemented two hybrid random forest (RF) models, the sine cosine algorithm (SCA)-RF and the Harris hawk optimizer (HHO)-RF for predicting backbreaks. According to their study, SCA-RF outperformed HHO-RF in terms of training and testing criteria, respectively. (Monjezi *et al*. 2008) estimated the influence of blasting patterns on the backbreak using ANNs. The obtained results were very helpful in reducing the impact of backbreak due to blasting patterns. Likewise, (Ebrahimi *et al*. 2016) developed fuzzy set theory and MVR algorithm to forecast backbreak. Thus, a fuzzy set theory prediction model was performed with the optimal results to minimize the impact of backbreak and enhance the efficiency of blasting. (Khandelwal *et al.* 2013) proposed an ML model, namely SVM to predict the behavior of blast-induced backbreak compared to MVRA. Based on their research, the SVM model revealed better results with R^2 of 987 compared to MVRA with R^2 of 0.89. (Ghasemi *et al.* 2016) used regression tree (RT) and ANFIS techniques for predicting backbreak in open pit mines with 35 datasets. On the whole dataset, the results showed the highest R^2 for RT prediction was 0.972 and 0.998 for ANFIS. (Ghasemi *et al.* 2017) developed particle swarm optimization (PSO) models in linear and quadratic forms for predicting backbreak due to bench blasting, so that the quadratic form of the PSO model performed better than linear PSO. (Hasanipanah *et al*. 2021) introduced an "uncertain rule-based fuzzy approach" approach to assess blast-induced backbreak. This method demonstrated robust results. Additionally, the researchers developed a PSO-ANFIS model for predicting blast-induced backbreak, achieving a higher R^2 value of 0.922. They concluded by recommending this optimized model as the most effective in predicting the provided data (Hasanipanah *et al*. 2017). (Kumar *et al*. 2013) developed linear regression (LR) and RF models to predict backbreak, with the highest R^2 of 0.879 for LR and 0.979 for RF. (Yari *et al*. 2024) comprehensively discussed a soft computing approach for assessing backbreak. While different ML models are anticipated to predict backbreak due to blasting, no or less robust metaheuristic models based on optimization algorithms have been established, particularly on the dataset used in this study. Thus, the hybrid LSSVM-GWO, a state-of-the-art model, is preferred in this study due to its effectiveness and accuracy in evaluating blast-induced backbreak, particularly in open pit mines, aligning with the study's main objective, and demonstrating superior performance compared to other studied models.

According to the above literature and the limitations of traditional prediction models, relying on a single model is less robust and cannot offer a perfect solution for all complex scenarios. The performance of such a model tends to vary with input features. As research progresses, novel techniques often emerge, superseding the limitations and shortcomings of earlier studies. For this reason, this study proposes the hybrid LSSVM-GWO model and evaluates its effectiveness by comparing its predictive performance with three convolutional models, namely SVM, k-nearest neighbor (KNN), and LSSVM, in solving optimization problems. The proposed model can significantly improve the performance evaluation of backbreak in open-pit mines and has important engineering applications, including blast designs and the enhancement of blasting efficiency in open-pit mining operations.

2. Materials and methods

2.1 Dataset

Different multivariate features have been reported by (Shirani Faradonbeh *et al*. 2016, Ghasemi *et al*. 2016) to be used as input features for predicting blast-induced backbreak (m) in open pit mines. Based on their known significance in affecting blast-induced backbreak, input parameters, including burden (m), spacing (m), stemming (m), powder factor (kg/m^s), and stiffness ratio (m/m), were selected. These input variables have a direct influence on stability and fragmentation during blasting operations, making them useful indicators for accurate backbreak phenomenon prediction and evaluation. To ensure a robust and informed research approach, the selection of these five input variables was based on a comprehensive literature review and preliminary studies. These variables were chosen due to their demonstrated value in contributing to backbreak. For each developed model, a total dataset of 84 points, free of any missing values, outliers, or other anomalies, was equally split, with 80% used for training and 20% for testing purposes. The models employed in this study were initially trained with 70% of the dataset. However, they did not yield remarkable results. Consequently, this study proceeded to split the dataset as mentioned earlier and trained each model separately. Fig. 2

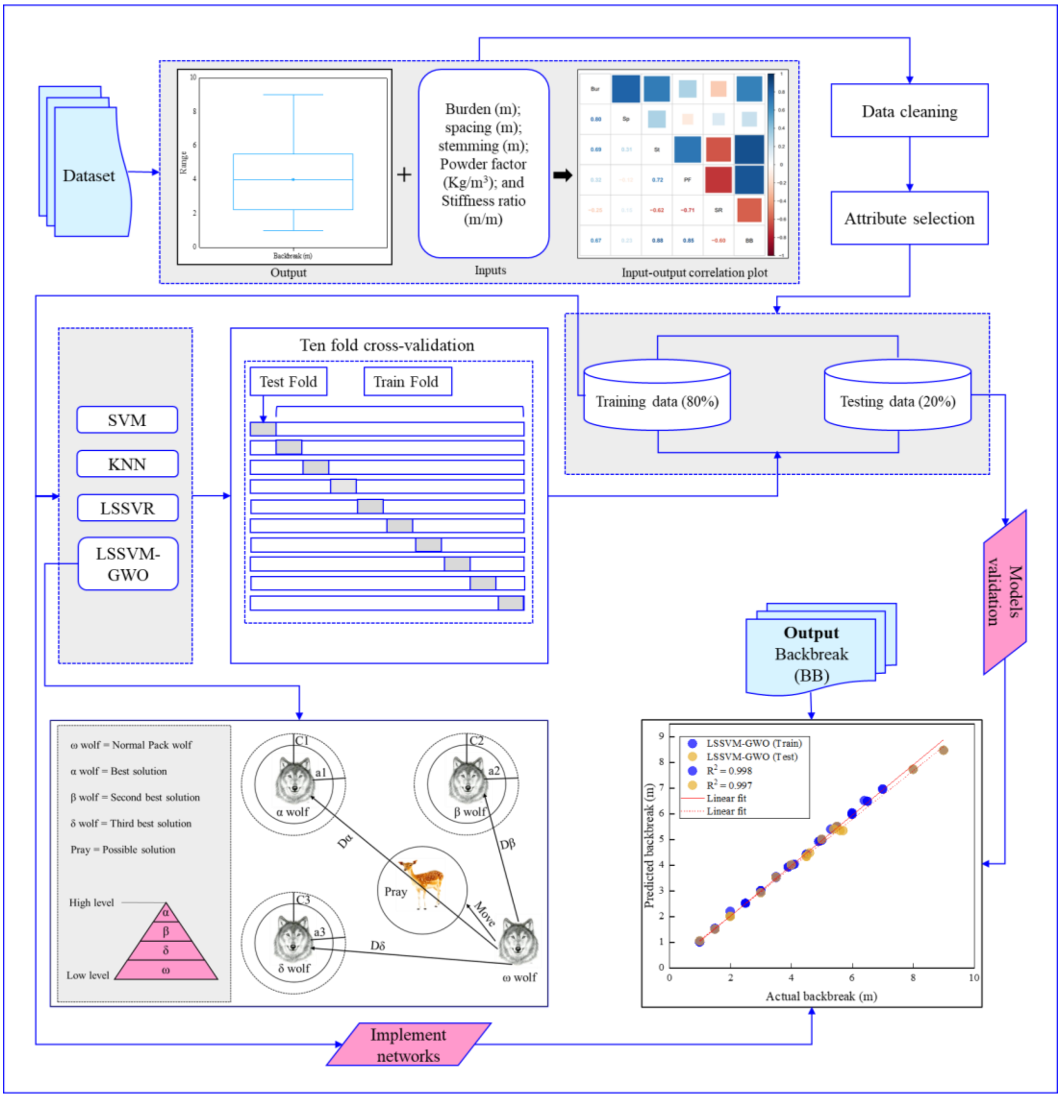


Fig. 1 Flowchart of the proposed methodology

Table 1 The statistical distribution of the actual dataset in this study

	Burden (m)	Spacing (m)	Stemming (m)	Powder factor (kg/m^3)	Stiffness ratio (m/m)	Backbreak (m)
Mean	3.45	4.45	3.70	0.49	2.96	4.02
Standard Deviation	0.72	0.83	0.62	0.24	0.58	1.95
Minimum	2	2.5	1.8	0.15	1.34	1
Maximum	5	5.6	4.5	1.1	4.5	9

shows the box plot of the input variables and output backbreak of the original dataset used in this study. "In Fig. 2, the legend of boxplots can be interpreted as follows: ▭ 25%~75%, ⌶ Range within 1.5IQR, — Median line, and ○ Outliers." Fig. 3 shows the correlation plot of different inputs and output backbreak of the actual dataset. As can be seen in Fig. 3, all parameters are positively correlated with backbreak, except for the stiffness ratio, which is negatively correlated with backbreak. Table 1 indicates the statistical distribution of the original dataset. Fig. 1 illustrates the flowchart of the utilized method.

2.2 Support vector machine

Vapnik *et al.* (1997) initially introduced SVM. Regression analysis and classification using hyperplane classifiers are two common applications of SVMs. The ideal hyperplane strengthens the separation of the two classes where the support vector is located (Sun *et al.* 2019). It proposes a kernel function and Vapnik's ε-insensitive loss function (Negara *et al.* 2017) to establish the forecast function from a high-extent feature space.

It employs a kernel function to map nonlinear input data into a feature space with a high dimensionality. It then seeks

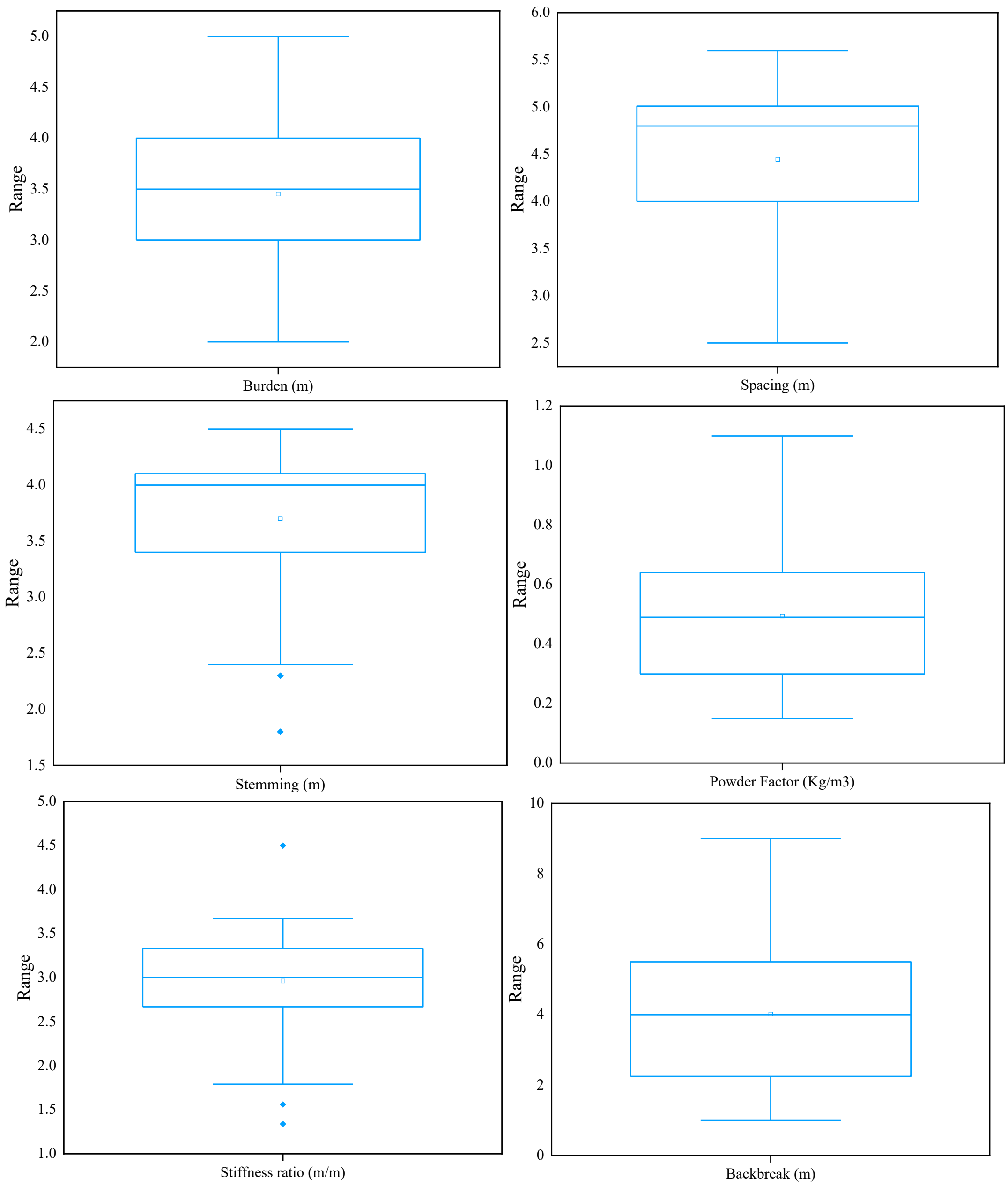


Fig. 2 The box plot of the original whole dataset

the best hyperplane to separate this data for a given dataset denoted as P = $\{(x_1, y_1), (x_2, y_2)\ldots(x_n, y_n)\}$, where $x_i \in R^n$ represents the inputs and $y_i \in R^n$ represents the output. This enables a linear regression function (Xu *et al*. 2022, Barzegar *et al*. 2016, Longjun *et al*. 2011) with the following characteristics to relate the initial input to the output

$$f(x) = M_v.\varphi(x) + l_b \quad (1)$$

here, $\varphi(x)$ represents the kernel function, while M_v stand for the weight vector and l_b is a biased term, respectively. To achieve the values of M_v and l_b, Cortes and Vapnik (1995) suggested the cost function must be decreased as in Eq. (2)

$$cost\ function = \frac{1}{2}M_v^2 + C\sum_{i=1}^{k}(\xi_i^- + \xi_i^+) \quad (2)$$

$$\text{Subject to:} \begin{cases} y_i - (M_v.\varphi(x_1) + l_b) \le \varepsilon_0 + \xi_i^+ \\ (M_v.\varphi(x_1) + l_b) - y_i \le \varepsilon_0 + \xi_i^- \\ \xi_i^-, \xi_i^+ \ge 0, i = 1, 2, \ldots., n \end{cases} \quad (3)$$

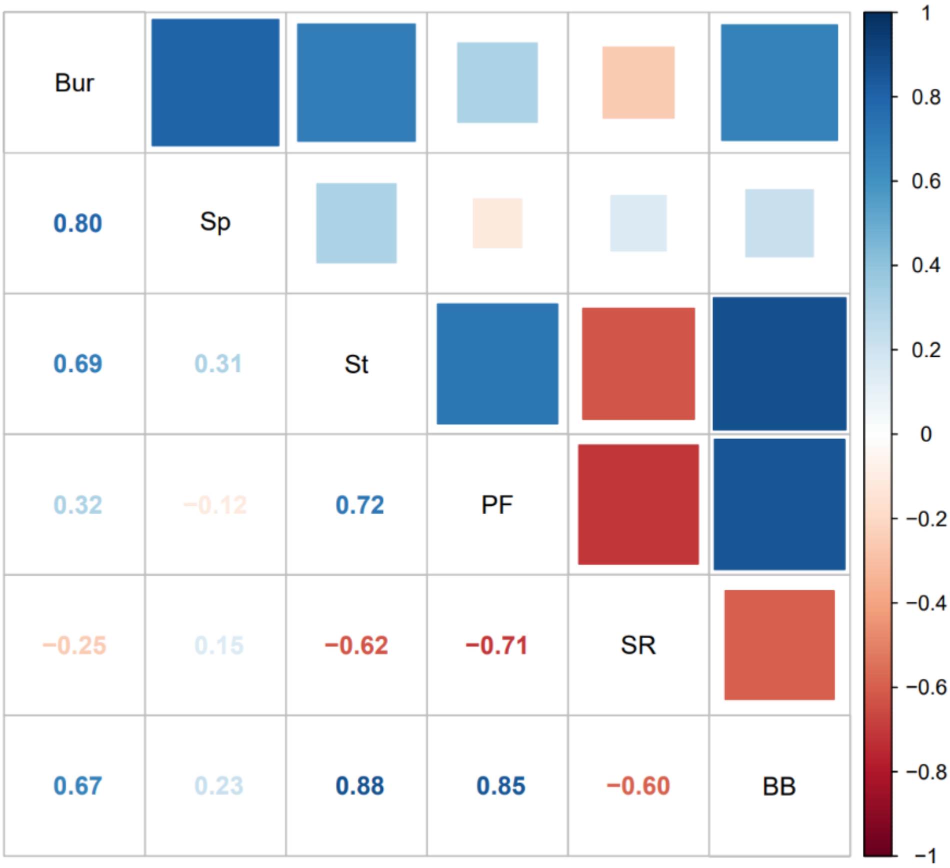


Fig. 3 Correlation plot of different inputs and output backbreak

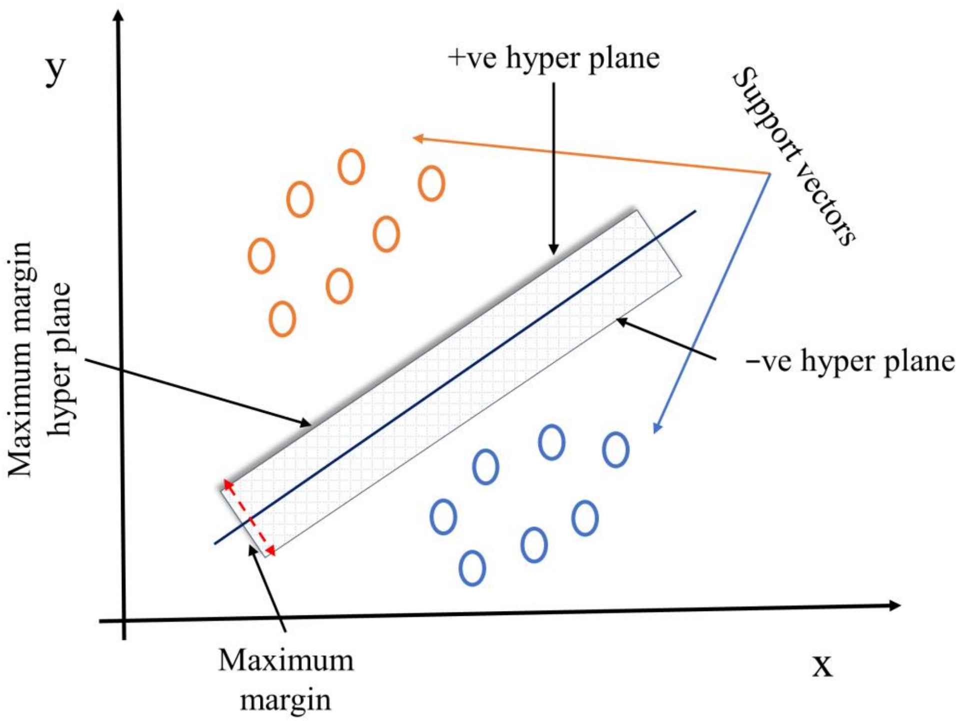


Fig. 4 The overall architecture of the support vector machine

Here, we have Lagrange multipliers denoted as ∞_i and ∞'_i, constrained within the range of $0 \leq \infty_i$ and $\infty'_i \leq$ C.

Meanwhile, the kernel function $\varphi(x_i, x_j)$ is also a part of this equation. The accomplishment of SVR depends on the latter option. In SVM, a wide range of kernel functions including linear, polynomial, sigmoid, gaussian, radial basis and exponential radial basis, have been examined (Barzegar *et al*. 2016). Fig. 4 illustrates the overall architecture of the SVM.

2.3 K-Nearest neighbors

The K-nearest neighbor (KNN) is considered a non-parametric classification method by Han and Kamber (2001). Both classification and regression problems can be solved using it. When given an uncertain instance (q) for classification, the KNN classifier looks for the k training instances in the pattern space that are most similar to q. These k training instances are the "nearest neighbors" of q. The class of q can then be computed by simple majority voting or by voting based on the distance from its KNN. In particular, for regression, we additionally determine the k instances of the dataset that are most similar to q. The value of q must then be predicted using their values. The easiest method is to take the average of the neighbors' values. Thus, KNN has two steps. The first step of KNN is to determine the nearest neighbors, and the second step is to utilize these neighbors to determine the class (in classification) or the actual value (in regression), or both.

2.4 Least square support vector machine

LSSVM was originally developed by (Suykenset *et al*. 1999) to tackle the SVM's computational intricacy Vapnik (2013). In LSSVM, an equality constraint replaces the inequality constraint in SVM, and a complex QP optimization problem is transformed into linear system equations. The model breakdown and prediction problems can be explained more quickly by this approach. The fundamental idea is as follows:

$\{(x_i, y_i)|\ i = 1,\dots,l\}, x_i \in R^n$ represents an input vector and $y_i \in R^n$ signifies matching outputs, given l sample points for training. In LSSVM, the approximation function can be specified as in Eq. (4)

$$y(x_i) = \omega^T\varphi(x_i) + b \tag{4}$$

The regression optimization problem is as in Eq. (5)

$$\min_{\omega,\xi,b}\ j(\omega,\xi) = \frac{1}{2}\omega^T\omega + \frac{g}{2}\sum_{i=1}^{l}\xi^2 \tag{5}$$

$$s.t.\ \ y_i(x_i) = \omega^T\varphi(x_i) + b + \xi, i = 1,\dots,l$$

where ω stands for the weight vector, $g \in R^+$ + is the penalty factor, ξ^i shows the error variable, b is the deviation, and $\varphi(\cdot)$ is a multi-scale spatial mapping. The above problems are solved by the Lagrange method as in Eq. (6)

$$Ł(\omega,\xi,b,\alpha) = \frac{1}{2}\omega^T\omega + \frac{g}{2}\sum_{i=1}^{l}\alpha_i\left(\omega^T\varphi(x_i) + b + \xi_i - y_i\right) \tag{6}$$

where α_i is the Lagrange multiplier. The conversion to a linear equation, based on Karush–Kuhn–Tucker (KKT) conditions, is as follows in Eq. (7) (Luo *et al*. 2019)

$$\begin{pmatrix} 0 & 1_l^T \\ 1_l & K + g^{-1}I_l \end{pmatrix}\begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix} \tag{7}$$

where $y = [y_1, y_2,\dots, y_l]^T, 1_l = [1,\ \dots, 1]^T, I_l$ represent the unit matrix of l_th order, $\alpha = [\alpha_1,\dots,\alpha_l]^T$ and the kernel function matrix denoted by K that meets Mercer's requirement as follows in Eq. (8)

$$K = \varphi(x_i)^T\varphi(x_j), \qquad (i,j) = 1,\dots,l \tag{8}$$

The radial basis function as the kernel function has been employed in this study due to its outstanding presentation and ability to generalize, as shown in Eq. (9) (Azimi *et al*. 2019)

$$K = K(x, x_i) = \exp^{\frac{-|x-x_i|^2}{2\sigma^2}} \tag{9}$$

The width of the kernel function is denoted by σ. Lastly, the LSSVM function is approximated as follows in Eq. (10)

$$y(x) = \sum_{i=1}^{l}\alpha_i K(x, x_i) + b \tag{10}$$

The penalty parameter $'g'$ and kernel width $'\sigma'$ have a significant impact on the LSSVM regression model's prediction accuracy and generalization ability. Therefore, these two parameters must be optimized.

2.5 Grey wolf optimization

Mirjalili *et al*. (2014) proposed the grey wolf optimization (GWO) algorithm, a novel approach to swarm intelligence. Several studies have successfully used and applied the GWO method (Aljarah *et al*. 2019, Faris *et al*. 2016, Maroufpoor *et al*. 2019, Tikhamarine *et al*. 2019b). The GWO algorithm was inspired by the social hunting behavior of grey wolves in nature. The pack of wolves was divided into four classes to map the social hierarchy of grey wolves, with Alpha (α) being the most fitting, followed by Beta (β) and Delta (δ), and the rest of the pack being designated as Omega (ω). The GWO algorithm is based on a mathematical model of the grey wolves' hunting process (victim tracking, encircling, and attacking). Fig. 5 shows the graphical representation of the social hierarchy and the process of a location update. Table 2 shows the pseudo code of the GWO algorithm. The encircling behavior of the grey wolves towards their prey can be computed as follows

$$\vec{D} = \left|\vec{C} * \vec{X}_p(t) - \vec{X}_p(t)\right| \tag{11}$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} * \vec{D} \tag{12}$$

$$\vec{A} = 2\vec{a} * \vec{r_1} - \vec{a} \tag{13}$$

$$\vec{C} = 2 * \vec{r}_2 \tag{14}$$

where X is the grey wolf's vector location, X_p is the prey's vector location, D = distance between X and X_p, t = the number of the existing iteration, and A and C are the analogous components-wise multiplication.

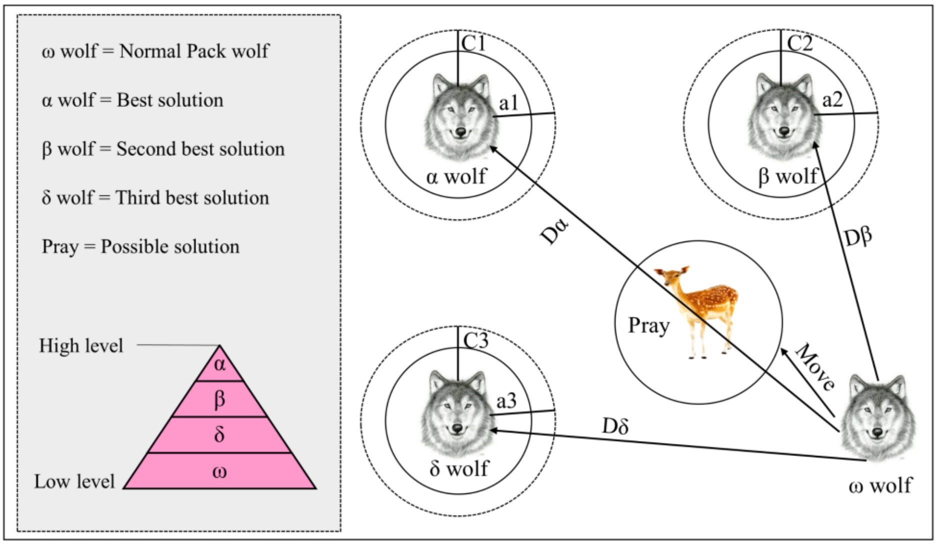


Fig. 5 Graphical representation of the social hierarchy and the process of a location update

Table 2 Pseudo code of the GWO algorithm

Algorithm: GWO Algorithm
initialization of the grey wolf population *Xi (i = 1, 2, …, n)* initialization of *a, A and C* computation of the fitness value of each wolf in the population Xα = the best solution Xβ = the second best solution Xδ = the third best solution ***while*** *(t<maximum number of cycles)*   **for** each wolf   updating the position by equation 21   **end for**   updating *a, A and C*   computation of the fitness value of all wolves in the population   updating the positions of *Xα, Xβ and Xδ*   *t = t+1* **end while** return Xα

Eq. (15) to Eq. (19) show how grey wolves update their positions of α, β, and δ wolves to mimic the grey wolves hunting behavior. It is widely believed that the wolves α, β, and δ are closest to the prey and attract the other wolves to the position of the prey. The grey wolf population can use the following formulae to determine the position of prey

$$\overrightarrow{D_\alpha} = \left|\overrightarrow{C_1} * \vec{X}_\alpha(t) - \vec{X}\right| \tag{15}$$

$$\overrightarrow{D_\beta} = \left|\overrightarrow{C_2} * \vec{X}_\beta(t) - \vec{X}\right| \tag{16}$$

$$\overrightarrow{D_\delta} = \left|\overrightarrow{C_3} * \vec{X}_\delta(t) - \vec{X}\right| \tag{17}$$

$$\overrightarrow{X_1} = \vec{X}_\alpha(t) - \overrightarrow{A_1} * \overrightarrow{D_\alpha} \tag{18}$$

$$\overrightarrow{X_2} = \vec{X}_\beta(t) - \overrightarrow{A_2} * \overrightarrow{D_\beta} \tag{19}$$

$$\overrightarrow{X_3} = \vec{X}_\delta(t) - \overrightarrow{A_3} * \overrightarrow{D_\delta} \tag{20}$$

Eq. (21) uses the positions received from Eq. (18) to Eq. (20) to adjust the wolves' next position

$$\vec{X}(t+1) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3} \tag{21}$$

where X(t+1) represents the position for the next iteration. Utilizing Eq. (21) to determine a fresh position for the alpha wolves, compels the Omega wolves to adjust their positions, facilitating convergence with the prey.

2.6 Lease Square Support Vector Machine-Grey Wolf Optimization (LSSVM-GWO)

The LSSVM model's parameters 'g' and 'σ' are crucial for predicting accuracy. A high value of the penalty parameter 'g' result in exceptionally high accuracy for the training data but lower accuracy for the test data, whereas a low value of 'g' result in poor performance due to poor functionality of the model (Robles-Rodriguez *et al*. 2020). Furthermore, an extremely large value of kernel factor 'σ'

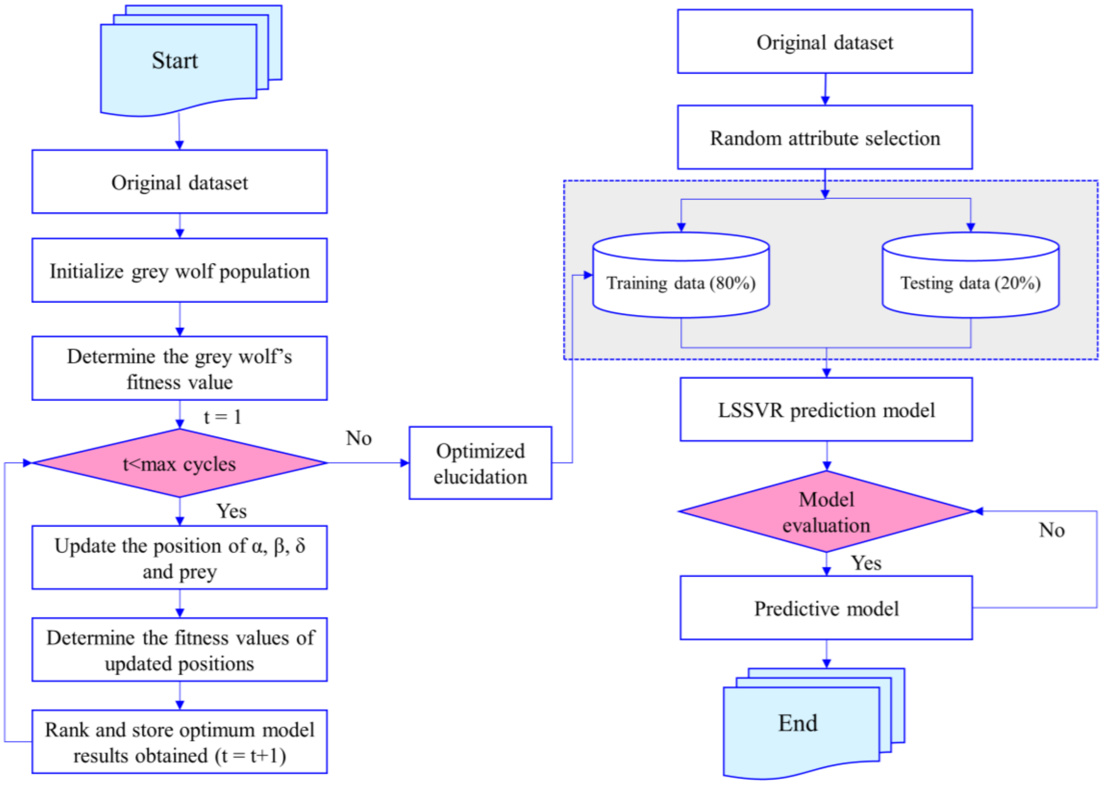


Fig. 6 The hybrid LSSVM-GWO procedure

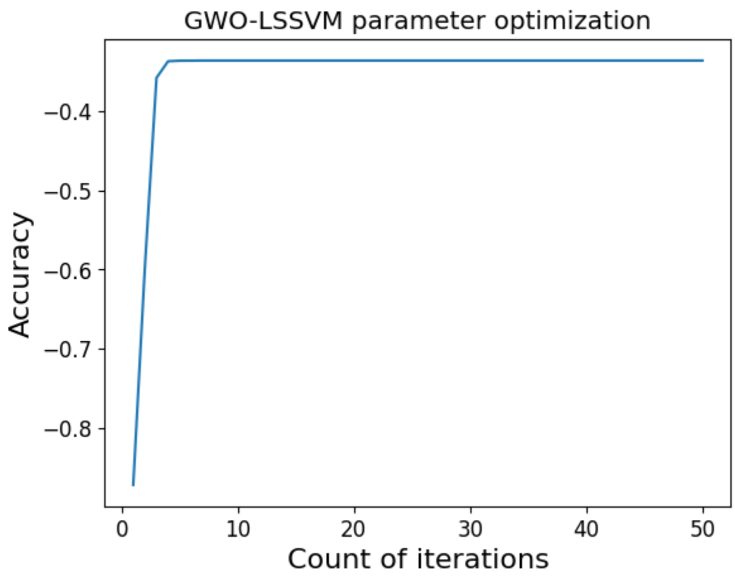


Fig. 7 LSSVM-GWO parameter optimization

can lead to overfitting, while a small value can cause underlearning issues. The influence of a single training sample on other examples is defined by the kernel width 'σ'. Therefore, the values of the LSSVM model parameters 'g' and 'σ' must be prudently chosen. Several studies have employed optimization techniques, such as the PSO algorithm, to find the optimal values of important parameters in regression models by Zhu and Zhu (2018). In this research, an effective hybrid LSSVM-GWO model is developed to determine the optimal parameters of the LSSVM model, as shown in Fig. 6, which also illustrates the hybrid LSSVM-GWO procedure. Fig. 7 illustrates the parameter optimization of LSSVM-GWO, with the best C and best gamma identified as 3.83282 and 0.041540, respectively. The LSSVM model has already demonstrated superior performance compared to the other models, such as KNN and SVM in this study. In an effort to enhance its robustness, it was exclusively optimized using the GWO optimization algorithm.

1	Test	Train	Train	Train	Train	Train	Train	Train	Train	Train
2	Train	Test	Train	Train	Train	Train	Train	Train	Train	Train
3	Train	Train	Test	Train	Train	Train	Train	Train	Train	Train
4	Train	Train	Train	Test	Train	Train	Train	Train	Train	Train
5	Train	Train	Train	Train	Test	Train	Train	Train	Train	Train
6	Train	Train	Train	Train	Train	Test	Train	Train	Train	Train
7	Train	Train	Train	Train	Train	Train	Test	Train	Train	Train
8	Train	Train	Train	Train	Train	Train	Train	Test	Train	Train
9	Train	Train	Train	Train	Train	Train	Train	Train	Test	Train
10	Train	Train	Train	Train	Train	Train	Train	Train	Train	Test

Fig. 8 The 10-Fold Cross-Validation

2.7 K-fold cross-validation

Different methods, including the simple substitution method (Braga-Neto *et al*. 2004a) bootstrap method by Efron and Tibshirani (1994) holdout method by Pham (2002), and the bolstered method [48], have been employed to validate the regression model. In this study (Braga-Neto *et al*. 2004b), k-fold cross-validation (CV) was used to validate the training data, and k was chosen as 10 based on the number of datasets and suggestions. The training data were allocated into 10 folds for hyperparameter tuning. As shown in Fig. 8, one fold was used for validation and nine folds were used for training. The average of the results of 10 rounds constitutes the model's ultimate result. The problem of overfitting can be solved by using this approach.

3. Models evaluation

Performance metrics are the key indicators to assist in model evaluation. The optimal model is considered to have the largest coefficient of determination (R^2) (modified after (Cai *et al*. 2022, Shahani *et al*. 2021)), the smallest MSE (Wu *et al*. 2024), MAE (Ahmed *et al*. 2023), and RMSE (Zeng *et al*. 2021), as well as appropriate *a*20-index (Shahani *et al*. 2022a, Shahani *et al*. 2022b) values. Each developed model's performance in predicting backbreak was assessed employing the evaluation metrics below.

$$R^2 = 100\left[\frac{\sum_{i=1}^{n}(b_o-\bar{b}_o)(b_p-\bar{b}_p)}{\sqrt{\sum_{i=1}^{n}(b_o-\bar{b}_o)(b_p-\bar{b}_p)}^{2}}\right]^2 \quad (22)$$

$$MAE = \frac{\sum_{i=1}^{n}\left|b_o-b_p\right|}{N} \quad (23)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(b_o-b_p)^2}{N}} \quad (24)$$

$$a20-index = \frac{m20}{N} \quad (25)$$

where, b_o and b_p represent the original and predicted backbreak; $\bar{b}_o$ and $\bar{b}_p$ are the average of the actual and predicted backbreak, $m20$ represents the dataset with a rate original/predicted backbreak between 0.80 and 1.20, and N signifies the dataset's number.

4. Results and discussion

The main objective of this research is to examine the capability of different ML prediction models, i.e., SVM, KNN, LSSVM, and LSSVM-GWO in predicting blast-induced backbreak. The collected actual and predicted output values were subsequently compared and graphed, facilitating an in-depth performance and the investigation of correlations within the constructed models. Different performance indices, namely R^2, MAE, RMSE, and *a*20 index, were employed as performance measures to determine the final result. These indices were utilized to analyze and contrast the predicted models, aiding in the identification of the optimal model for accurate data prediction. The overall dataset of 84 samples was allocated to 80% (67 samples) to train the models and 20% (17 samples) to test the models.

Fig. 9 demonstrates the scatter plots of backbreak prediction by SVM, KNN, LSSVM, and hybrid LSSVM-GWO regression models under the train and test criteria, respectively. In Fig. 9, the R^2 values predicted by the SVM, KNN, LSSVM, and LSSVM-GWO regression models are 0.945, 0.960, 0.999, and 0.998 and 0.941, 0.962, 0.963, and 0.997 under the training and the testing criteria, respectively. Due to the strong accuracy demonstrated by the LSSVM model, this study has employed the GWO optimization technique to further enhance the predictive capability of LSSVM in evaluating blast-induced backbreak in open pit mines.

Table 3 demonstrates the performance metrics, i.e., R^2, MAE, RMSE, and *a*20-index of the developed SVM, KNN, LSSVM, and LSSVM-GWO regression models computed by Eq. 22 to 25. In this study, according to the developed regression models, LSSVM-GWO outperformed other

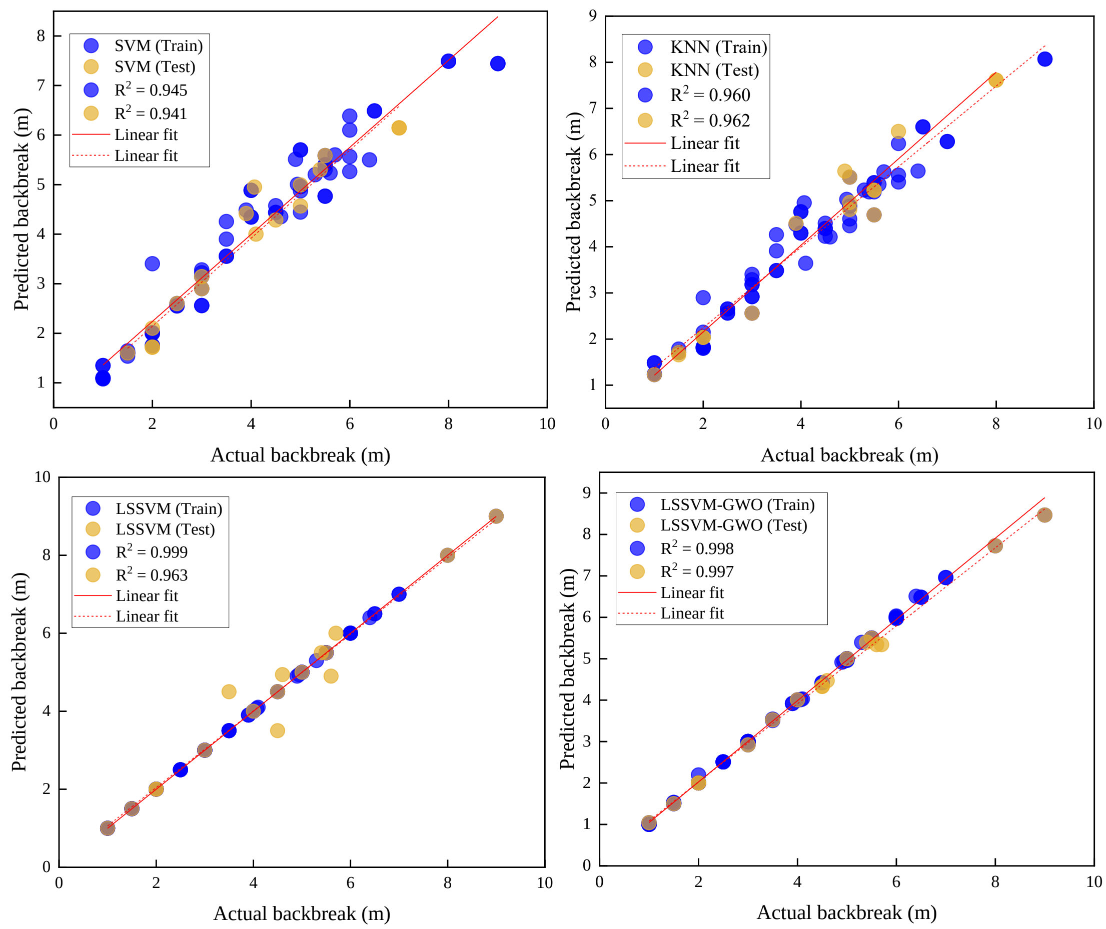


Fig. 9 Scatter plot of backbreak prediction by SVM, KNN, LSSVM, and GWO-LSSVM regression models.

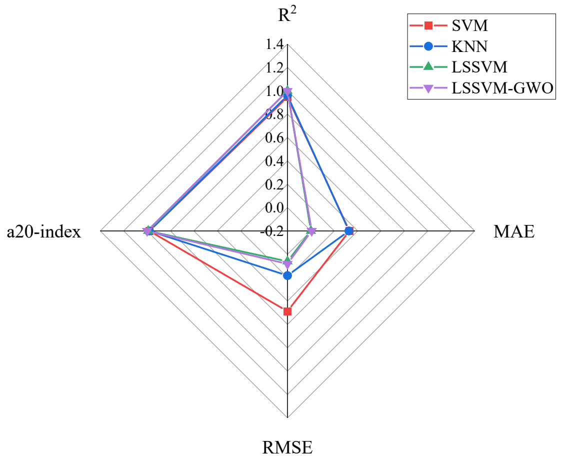
Table 3 Performance metrics of the developed regression models

Model	Training				Testing			
	R^2	MAE	RMSE	*a*20-index	R^2	MAE	RMSE	*a*20-index
SVM	0.945	0.3298	0.48850	0.98	0.941	0.3030	0.4184	1.02
KNN	0.96	0.3256	0.18258	0.98	0.962	0.3428	0.18691	0.98
LSSVM	0.999	0.00	0.06065	0.99	0.963	0.2023	0.14394	0.99
LSSVM-GWO	0.998	0.0068	0.08250	0.999	0.997	0.1209	0.19360	1.01

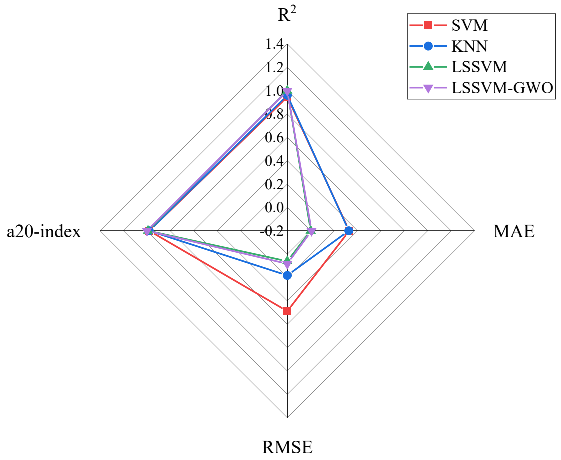
models with R^2 = 0.998 and 0.997, MAE = 0.0068 and 0.1209, RMSE = 0.0825 and 0.1936, and *a*20-index = 0.99 and 1.01 for backbreak prediction under training and testing criteria. Therefore, LSSVM-GWO is a suitable ML method that can be accurately utilized to evaluate backbreak, as depicted in Fig. 10.

Several researchers (Ghasemi *et al*. 2016, Faradonbeh *et al*. 2016, Sharma *et al*. 2021) have predicted backbreak by applying different ML regression models, i.e., NLMR, GP, RT, ANFIS, and MRVA, using the same dataset, which can be found in the online literature. However, this study explains a more optimized performance than previously published work by developing a hybrid LSSVM-GWO model.

SHAP (Shapley Additive Explanations), originating from game theory, is an ML technique that calculates importance values for each feature, facilitating the understanding of the impact of each feature on model predictions. In Figs. 11(a) and 11(b), the evaluation of the importance of individual variables' features is illustrated using SHAP values for each data point at the test dataset. This visual representation establishes the connection between each feature value and its corresponding SHAP value. Specifically, focusing on the key feature 'stemming,' it is evident that a high feature value corresponds to a high SHAP value, and vice versa. This observation elucidates that elevating the feature value leads to an increase in the output backbreak. Therefore, when the SHAP value is high, it indicates a propensity to elevate the output probability or value.

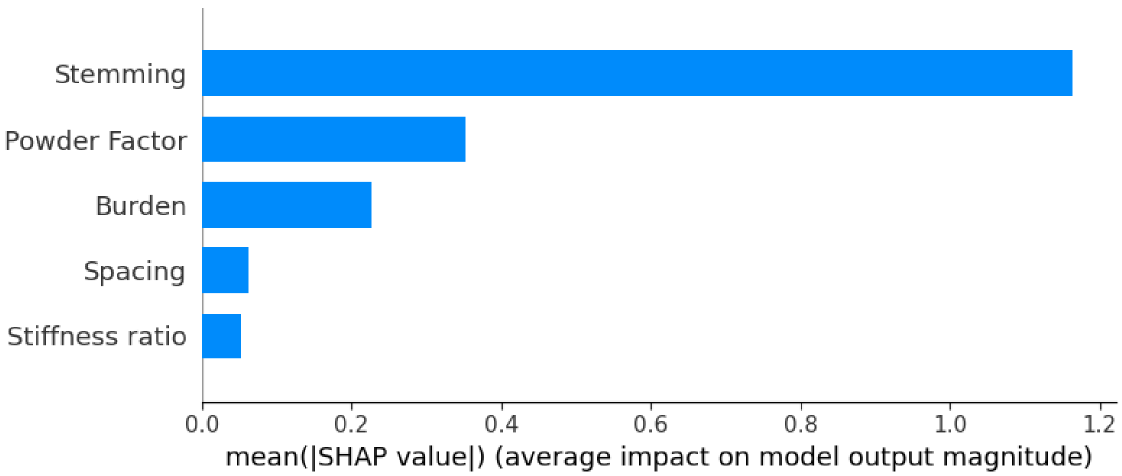


(a) Train

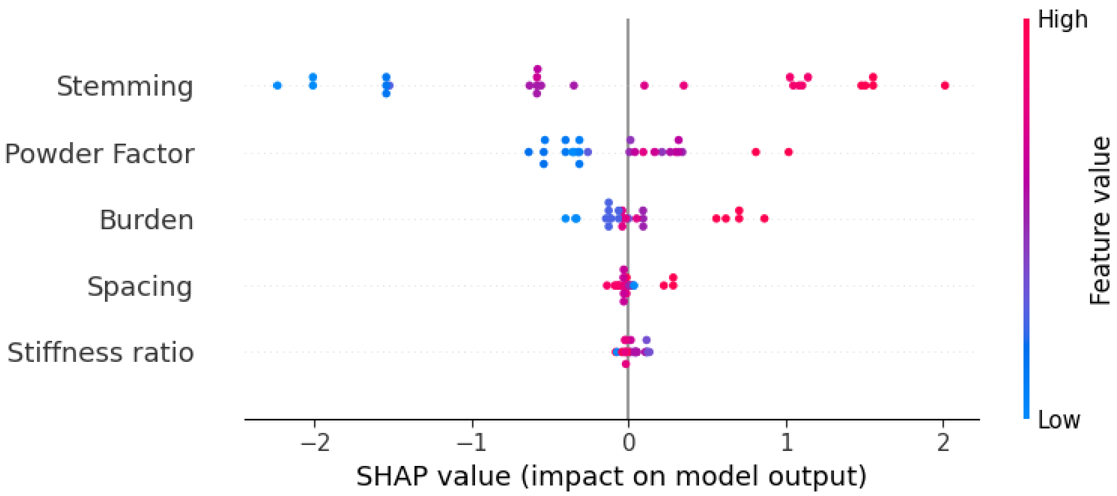


(b)Test

Fig. 10 Radar plot of performance metrics of the developed regression models



(a)



(b)

Fig. 11 SHAP values plot for each input feature variable that impacts the output backbreak at the test dataset. (a) The bar plot of the mean absolute SHAP values and (b) The SHAP value plot for the degree of positive or negative influence of each individual measurement on the prediction. Warmer colors indicate higher observed values for that measurement, while cooler colors represent lower values for that measurement

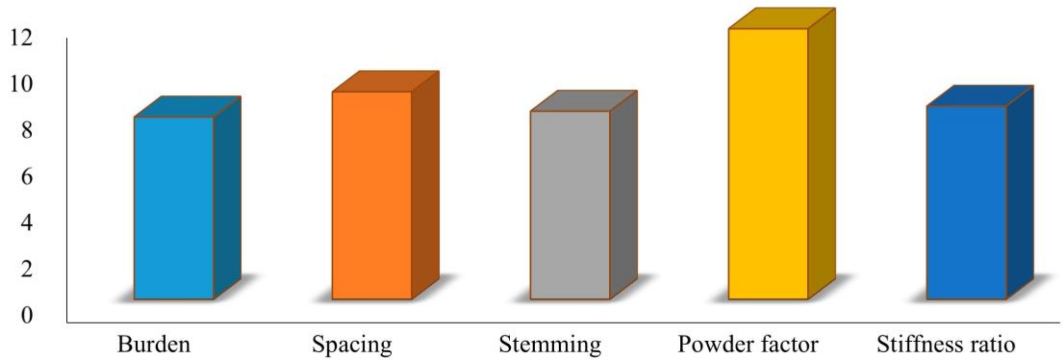


Fig. 12 Sensitivity analysis of the developed LSSVM-GWO model in this study on the testing dataset

5. Sensitivity analysis

Accurate analysis of the crucial parameters that greatly impact rock blast-induced backbreak is necessary, and this can pose challenges in blast design. Therefore, in this study, the cosine amplitude technique (Ji and Liang 2017, Momeni *et al*. 2014) is employed to calculate the relative influence of the input parameters on the output. Eq. 26 represents the utilized technique

$$r_{ij} = \frac{\sum_{k=1}^{n}(x_{ik}y_{jk})}{\sqrt{\sum_{k=1}^{n}x_{ik}^2\sum_{k=1}^{n}y_{jk}^2}} \qquad (26)$$

where x_i stands for input values and y_j represents output values. The parameter n displays the number of datasets during the testing purpose. The variable r_{ij} which falls within the range of 0 and 1, and provides further information about the accuracy between each variable and the target. Based on Eq. (26), if the value of r_{ij} for a parameter is 0, it specifies that there is little to no important correlation between that parameter and the target. Conversely, when r_{ij} is equal to 1 or around 1, it shows a substantial relationship that can strongly impact the backbreak.

Fig. 12 shows the influence of the input parameters, i.e., burden, spacing, stemming, powder factor, and stiffness ratio on the output backbreak at the test data. Thus, according to the sensitivity analysis, the powder factor is the most influential parameter, while the burden is the least influential parameter on the backbreak. Due to the strong accuracy exhibited by the LSSVM-GWO model on the testing dataset, sensitivity analysis was the only one performed with coefficient values of burden = 7.86, spacing = 8.95, stemming = 8.11, powder factor = 11.67 and stiffness ratio = 8.33.

6. Model repeatability and reproducibility analysis

Fig. 13 illustrates the repeatable results with error bars for the SVM, KNN, LSSVM, and LSSVM-GWO models across three distinct tests under the same testing conditions, using R^2, MAE, RMSE, and a20-index, providing valuable insights into their repeatability and reproducibility on the testing dataset. Among the models, LSSVM-GWO demonstrated the highest level of stability, consistently delivering strong performance across all tests. The model maintained robust R^2 values, along with low MAE, RMSE, and *a*20-index scores, suggesting its excellent predictive accuracy and reliability under diverse testing conditions. LSSVM also exhibited solid repeatability and reproducibility, showing stable and high performance throughout the tests. On the other hand, SVM and KNN showed greater variability in performance, particularly with MAE and RMSE, indicating that their results were more susceptible to fluctuations in experimental conditions. Despite this, both SVM and KNN still yielded reliable predictions, though with somewhat reduced repeatability and reproducibility compared to LSSVM and LSSVM-GWO. Therefore, LSSVM-GWO proved to be the most reliable and consistent model, offering superior stability and performance across all conducted tests

7. Conclusions

7.1 Findings

This study presents a hybrid model, LSSVM-GWO, for predicting blast-induced backbreak in open-pit mines. The model's performance was evaluated using 84 datasets, with a training-to-testing ratio of 80% and 20%, respectively. A 10-fold iterative cross-validation method was employed to ensure robust results. Among the models tested, SVM, KNN, LSSVM, and LSSVM-GWO, the latter outperformed the others in terms of predictive accuracy, achieving an R^2 of 0.997, MAE of 0.1209, RMSE of 0.1936, and *a*20-index of 1.01 on the test datasets. SHAP analysis provided deeper insights into the influence of input features on backbreak predictions, highlighting the significant role of the powder factor in blast-induced backbreak. The LSSVM-GWO model's superior accuracy, with notable sensitivity coefficients for burden, spacing, stemming, powder factor, and stiffness ratio, demonstrates its potential for practical use in optimizing blast designs and improving blasting efficiency in open-pit mining operations.

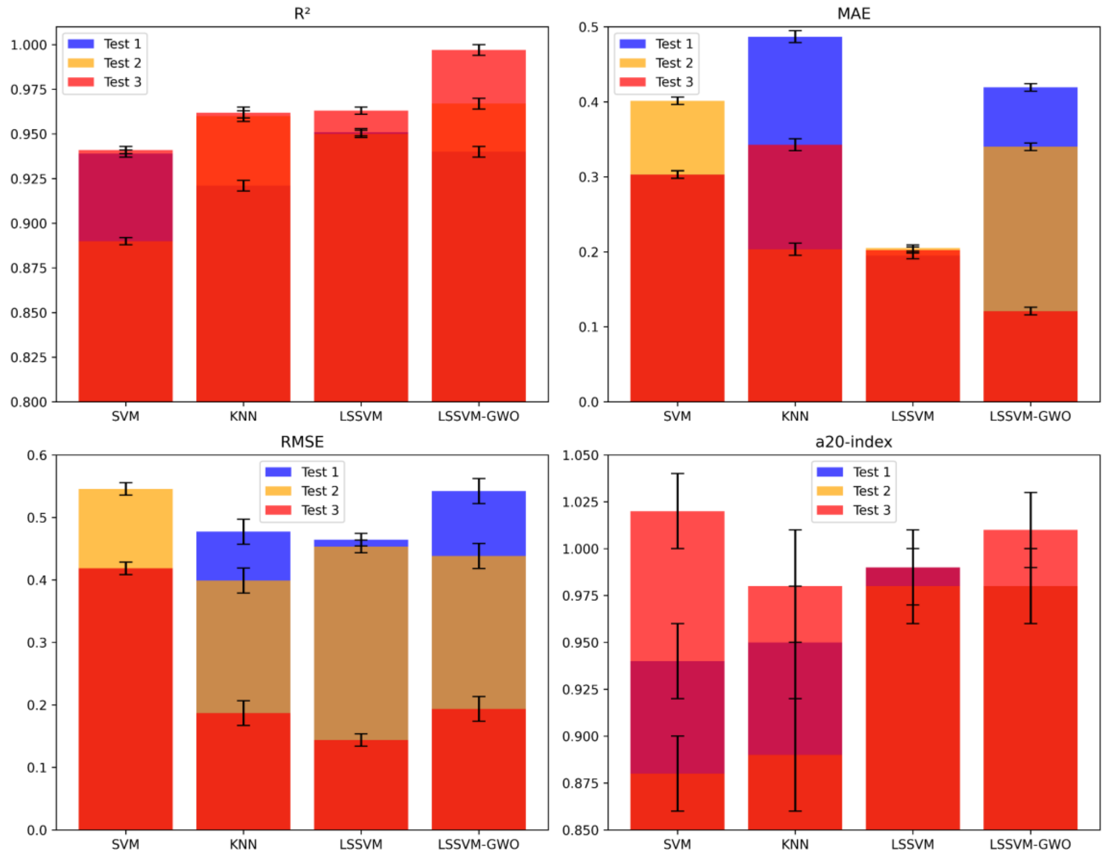


Fig. 13 Repeatable results with error bars for the developed models across three distinct tests under the same testing conditions on the testing dataset

7.2 Research Limitations and Recommendations for Future Research

While promising, the LSSVM-GWO model has several limitations that need to be addressed in future research. Firstly, its performance may be compromised when dealing with noisy or incomplete input data, which could affect its accuracy and generalizability. Additionally, the model was specifically developed for open-pit mining, and its applicability across other mining contexts, including underground mining with varying geological conditions, and input variations (e.g., rock types and blasting conditions), requires further validation. Moreover, the model's performance might decrease when applied to large datasets or when hyperparameters are not optimally tuned. There is also room for improvement regarding feature selection, as not all potentially influential factors were considered in this study. Future research should focus on improving data quality handling, investigating additional features, and optimizing hyperparameters to enhance the model's robustness. Furthermore, the model should be tested in different mining environments to assess its adaptability, and integration with other machine learning techniques could potentially boost its predictive power. Finally, implementing the model in real-time monitoring systems would enable proactive adjustments in blast designs, ensuring safer and more efficient mining operations.

7.3 Implications of the study

This study provides significant implications for the field of mining engineering. The LSSVM-GWO model offers a robust, data-driven approach to predict backbreak, which could be directly applied to improve safety and efficiency in blast design. By adjusting key parameters, such as the powder factor and burden, mining engineers can minimize the risk of backbreak and enhance blasting efficiency. Additionally, the model's integration with real-time monitoring systems could help mitigate risks proactively, ultimately improving both the economic and safety issues of mining operations.

Acknowledgments

This study was funded by the Guizhou Provincial Education Department's (Hundred Schools Thousands of Enterprises Science and Technology Research List) Project([2024]013) and the Qiankehezhongyindi([2024]039).

References

Agrawal, H. and Mishra, A.K. (2018), "Probabilistic analysis on scattering effect of initiation systems and concept of modified charge per delay for prediction of blast induced ground vibrations", *Measurement*, **130**, 306-317. https://doi.org/10.1016/j.measurement.2018.08.032.

Ahmed, H.U., Mostafa, R.R., Mohammed, A., Sihag, P. and Qadir, A. (2023), "Support vector regression (SVR) and grey wolf optimization (GWO) to predict the compressive strength of GGBFS-based geopolymer concrete." *Neural Comput. Appl*., **35**(3), 2909-2926. https://doi.org/10.1007/s00521-022-07724-1.

Aljarah, I., Faris, H., Mirjalili, S., al-Madi, N., Sheta, A. and Mafarja, M. (2019), "Evolving neural networks using bird swarm algorithm for data classification and regression applications", *Cluster Comput*., **22**, 1317-1345. https://doi.org/10.1007/s10586-019-02913-5.

Azimi, H., Bonakdari, H. and Ebtehaj, I. (2019), "Design of radial basis function-based support vector regression in predicting the discharge coefficient of a side weir in a trapezoidal channel", *Appl. Water Sci*., **9**, 78. https://doi.org/10.1007/s13201-019-0961-5.

Barzegar, R., Sattarpour, M., Nikudel, M.R. and Moghaddam, A.A. (2016), "Comparative evaluation of artificial intelligence models for prediction of uniaxial compressive strength of travertine rocks, case study: Azarshahr area, NW Iran", *Model. Earth Syst. Environ*., **2**(2), 76. https://doi.org/10.1007/s40808-016-0132-8.

Bauer, A. (1982), "Wall control blasting in open pits, CIM Special 30, Canadian Institute of Mining and Metallurgy", In 14th Can, rock mechanics symposium, 3-10.

Braga-Neto, U., Hashimoto, R., Dougherty, E.R., Nguyen, D.V. and Carroll, R.J. (2004a), "Is cross-validation better than resubstitution for ranking genes?", *Bioinformatics*, **20**(2), 253-258. https://doi.org/10.1093/bioinformatics/btg399.

Braga-Neto, U. and Dougherty, E. (2004b), "Bolstered error estimation", *Pattern Recogn.*, **37**, 1267-1281. https://doi.org/10.1016/j.patcog.2003.08.017.

Cai, M., Hocine, O., Mohammed, A. S., Chen, X., Amar, M.N., and Hasanipanah, M. (2022), "Integrating the LSSVM and RBFNN models with three optimization algorithms to predict the soil liquefaction potential", *Eng. Comput*., **38**(4), 3611-3623. https://doi.org/10.1007/s00366-021-01392-w.

Cortes, C. and Vapnik, V. (1995), "Support-vector networks", *Mach. Learn.*, **20**(3), 273-297. https://doi.org/10.1007/BF00994018.

Du, X., Ma, X., Dong, C. and Nikkhoo, M.S. (2024), "Estimating pile setup parameter using XGBoost-based optimized models", *Geomech. Eng*., **36**(3), 259-276. https://doi.org/10.12989/gae.2024.36.3.259.

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

Efron, B. and Tibshirani, R.J. (1994), "An Introduction to the Bootstrap", *CRC press: Boca Raton, FL, USA*, ISBN 0412042312.

EL, J.C.L. and Carcedo, A. (1995), "Drilling and blasting of rocks", *CRC Press*.

Esmaeili, M., Osanloo, M., Rashidinejad, F., Bazzazi, A.A. and Taji, M. (2012), "Multiple regression, ANN and ANFIS models for prediction of backbreak in the open pit blasting", *Eng. Comput.*, **30**, 549-558. https://doi.org/10.1007/s00366-012-0298-2.

Faris, H., Aljarah, I. and Mirjalili, S. (2016), "Training feedforward neural networks using multi-verse optimizer for binary classification problems", *Appl. Intell.*, **45**, 322-332. https://doi.org/10.1007/s10489-016-0767-1.

Fattahi, H. and Bayat, N. (2024), "Developing drilling rate index prediction: A comparative study of RVR-IWO and RVR-SFL models for rock excavation projects", *Geomech. Eng*., **36**(2), 111-119. https://doi.org/10.12989/gae.2024.36.2.111.

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

Ghasemi, E., Amnieh, H.B. and Bagherpour, R. (2016), "Assessment of backbreak due to blasting operation in open pit mines: a case study", *Environ. Earth Sci.*, **75**, 552. https://doi.org/10.1007/s12665-0165354-6.

Han, J. and Kamber, M. (2001), DATA MINING: CONCEPTS AND TECHNIQUES, Morgan Kauffman, San Francisco.

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

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

Ji, X. and Liang, S.Y. (2017), "Model-based sensitivity analysis of machining-induced residual stress under minimum quantity lubrication", *Proc. Inst. Mech. Eng., Part B*., **231**(9), 1528-1541. https://doi.org/10.1177/0954405415601802s.

Kamran, M., Shahani, N.M. and Armaghani, D.J. (2022), "Decision support system for underground coal pillar stability using unsupervised and supervised machine learning approaches", *Geomech. Eng*., **30**(2), 107-121. https://doi.org/10.12989/gae.2022.30.2.107.

Khandelwal, M. and Monjezi, M. (2013), "Prediction of backbreak in open-pit blasting operations using the machine learning method", *Rock Mech Rock Eng.*, **46**, 389-396. https://doi.org/10.1007/s00603-012-0269-3.

Konya, C.J. and Walter, E.J. (1991), ROCK BLASTING AND OVERBREAK CONTROL. United States, Federal Highway Administration.

Kumar, S., Mishra, A.K. and Choudhary, B.S. (2022), "Prediction of back break in blasting using random decision trees", *Eng. Comput.*, **38**, 1185-1191. https://doi.org/10.1007/s00366-020-01280-9.

Longjun, D., Xibing, L., Ming, X. and Qiyue, L. (2011), "Comparisons of random forest and support vector machine for predicting blasting vibration characteristic parameters", *Procedia Eng.*, ***26***, 1772-1781. https://doi.org/10.1016/j.proeng.2011.11.2366.

Luo, C., Huang, C., Cao, J., Lu, J., Huang, W., Guo, J. and Wei, Y. (2019), "Short-term traffic flow prediction based on least square support vector machine with hybrid optimization algorithm", *Neural Process. Lett*., **50**, 2305-2322. https://doi.org/10.1007/s11063-019-09994-8.

Mahmoodzadeh, A., Ibrahim, H.H., Flaih, L.R., Alanazi, A., Alqahtani, A., Alsubai, S., Kahla, N.B and Mohammed, A.H. (2024), "A gene expression programming-based model to predict water inflow into tunnels", *Geomech. Eng*., **37**(1), 65-72. https://doi.org/10.12989/gae.2024.37.1.065.

Maroufpoor, S., Maroufpoor, E., Bozorg-Haddad, O., Shiri, J. and Yaseen, Z.M. (2019), "Soil moisture simulation using hybrid artificial intelligent model: hybridization of adaptive neuro fuzzy inference system with grey wolf optimizer algorithm", *J. Hydrol.*, **575**, 544-556. https://doi.org/10.1016/j.jhydrol.2019.05.045.

Mirjalili, S., Mirjalili, S.M. and Lewis, A. (2014), "Grey wolf optimizer", *Adv. Eng. Softw.*, **69**, 46-61. https://doi.org/10.1016/j.advengsoft.2013.12.007.

Mohammadnejad, M., Gholami, R., Sereshki, F. and Jamshidi, A. (2013), "A new methodology to predict backbreak in blasting operation", *Int. J. Rock Mech. Min. Sci.*, **60**, 75-81.

https://doi.org/10.1016/j.ijrmms.2012.12.019.

Momeni, E., Nazir, R., Armaghani, D.J. and Maizir, H. (2014), "Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN", *Measurement*, **57**, 122-131. https://doi.org/10.1016/j.measurement.2014.08.007.

Monjezi, M., Hashemi Rizi, S.M., Majd, V.J. and Khandelwal, M. (2014), "Artificial Neural Network as a Tool for Backbreak Prediction", *Geotech. Geol. Eng.*, **32**, 21-30. https://doi.org/10.1007/s10706-013-9686-7.

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

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

Monjezi, M., Amini Khoshalan, H. and Yazdian Varjani, A. (2010b), "Prediction of flyrock and backbreak in open pit blasting operation: A neuro-genetic approach", *Arab. J. Geosci.*, **5**, 441-448. https://doi:10.1007/s12517-010-0185-3.

Monjezi, M. and Dehghani, H. (2008), "Evaluation of effect of blasting pattern parameters on backbreak using neural networks", *Int. J. Rock Mech. Min. Sci.*, **45**(8), 1446-1453 https://doi.org/10.1016/j.ijrmms.2008.02.007.

Negara, A., Ali, S., AlDhamen, A., Kesserwan, H. and Jin, G. (2017), "Unconfined compressive strength prediction from petrophysical properties and elemental spectroscopy using support-vector regression", *In SPE annual technical symposium and exhibition*, OnePetro, Kingdom of Saudi Arabia.

Pham, D.L. (2002), "Fuzzy clustering with spatial constraints. Proceedings. 2002 International Conference on IEEE, Rochester, NY, USA, 2, 22-25.

Robles-Rodriguez, C.E., Bideaux, C., Roux, G., Molina-Jouve, C. and Aceves-Lara, C.A. (2020), "Soft-sensors for lipid fermentation variables based on PSO Support Vector Machine (PSO-SVM)", In distributed computing and artificial intelligence, *Proceedings of the 13th International Conference, Salamanca*, Spain, 175-183.

Saghatforoush, A., Monjezi, M., Faradonbeh, R.S. and Armaghani, D.J. (2016), "Combination of neural network and ant colony optimization algorithms for prediction and optimization of flyrock and back-break induced by blasting", *Eng. Comput.*, **32**, 255-266. https://doi.org/10.1007/s00366-015-0415-0.

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

Scoble, M.J., Lizotte, Y.C., Paventi, M. and Mohanty, B.B. (1997), "Measurement of blast damage", *Min. Eng.*, **49**, 103-108.

Shahani, N.M., Zheng, X., Liu, C., Hassan, F.U. and Li, P. (2021), "Developing an XGBoost Regression Model for Predicting Young's Modulus of Intact Sedimentary Rocks for the Stability of Surface and Subsurface Structures", *Front. Earth Sci.*, **9**, 761990. https://doi.org/10.3389/feart.2021.761990.

Shahani, N.M., Zheng, X., Guo, X. and Wei, X. (2022a), "Machine learning-based intelligent prediction of elastic modulus of rocks at thar coalfield", *Sustainability*, **14**(6), 3689. https://doi.org/10.3390/su14063689.

Shahani, N.M., Kamran, M., Zheng, X. and Liu, C. (2022b), "Predictive modeling of drilling rate index using machine learning approaches: LSTM, simple RNN, and RFA", *Pet. Sci. Technol.*, **40**(5), 534-555. https://doi.org/10.1080/10916466.2021.2003386.

Sharma, M., Agrawal, H. and Choudhary, B.S. (2022), "Multivariate regression and genetic programming for prediction of backbreak in open-pit blasting", *Neural Comput. Appl.*, **34**, 2103-2114. https://doi.org/10.1007/s00521-021-06553-y.

Sharma, M., Choudhary, B.S., Kumar, H. and Agrawal, H. (2021), "Optimization of delay sequencing in multi-row blast using single hole blast concepts", *J. Inst. Eng. (India): D*, **102**, 453-460. https://doi.org/10.1007/s40033-021-00270-5.

Shirani, F.R., Monjezi, M. and Armaghani, D.J. (2016), "Genetic programing and non-linear multiple regression techniques to predict backbreak in blasting operation", *Eng. Comput.*, **32**(1), 123-133. https://doi.org/10.1007/s00366-015-0404-3.

Stone, M. (1994), "Cross-validatory choice and assessment of statistical predictions", *J. R. Stat. Soc., ser. B Methodol.*, **36**(2), 111-147.

Suykens, J.A. and Vandewalle, J. (1999), "Least squares support vector machine classifiers", *Neural Process. Lett.*, **9**, 293-300. https://doi.org/10.1023/A:1018628609742.

Sun, J., Zhang, J., Gu, Y., Huang, Y., Sun, Y. and Ma, G. (2019), "Prediction of permeability and unconfined compressive strength of pervious concrete using evolved support vector regression", *Constr. Build. Mater.*, **207**, 440-449. https://doi.org/10.1016/j.conbuildmat.2019.02.117.

Tikhamarine, Y., Souag-Gamane, D. and Kisi, O. (2019b), "A new intelligent method for monthly streamflow prediction: hybrid wavelet support vector regression based on grey wolf optimizer (WSVR–GWO)", *Arab. J. Geosci.*, **12**, 540-520. https://doi.org/10.1007/s12517-019-4697-1.

Vapnik, V. (2013), *The nature of statistical theory*, Springer Science & Business Media: New York, NY, USA.

Vapnik, V., Golowich, S.E. and Smola, A. (1997), "Support vector method for function approximation, regression estimation, and signal processing", *Adv. Neural Inf. Process. Syst.*, 281-287.

Wu, L., Li, M., Zhang, J., Wang, Z., Yang, X. and Bian, H. (2024), "Deep learning-based AI constitutive modeling for sandstone and mudstone under cyclic loading conditions", *Geomech. Eng.*, **37**(1), 049-64. https://doi.org/10.12989/gae.2024.37.1.049.

Xu, C., Amar, M.N., Ghriga, M.A., Ouaer, H., Zhang, X. and Hasanipanah, M. (2022), "Evolving support vector regression using Grey Wolf optimization; forecasting the geomechanical properties of rock', *Eng. Comput.*, **38**, 1819-1833. https://doi.org/10.1007/s00366-020-01131-7.

Yari, M., Khandelwal, M., Abbasi, P., Koutras, E.I., Armaghani, D. J. and Asteris, P.G. (2024), "Applications of soft computing methods in backbreak assessment in surface mines: A comprehensive review", *CMES-Comput. Model. Eng. Sci.*, **140**(3). 2207-2238. https://doi.org/10.32604/cmes.2024.048071.

Zeng, F., Nait Amar, M., Mohammed, A. S., Motahari, M. R., and Hasanipanah, M. (2021), "Improving the performance of LSSVM model in predicting the safety factor for circular failure slope through optimization algorithms", *Eng. Comput.*, **38**(3), 1755-1766. https://doi.org/10.1007/s00366-021-01374-y.

Zhou, J., Dai, Y., Khandelwal, M., Monjezi, M., Yu, Z. and Qiu, Y. (2021), "Performance of hybrid SCA-RF and HHO-RF models for predicting backbreak in open-pit mine blasting operations", *Nat. Resour. Res.*, **30**(6), 4753-4771. https://doi.org/10.1007/s11053-021-09929-y.

Zhu, X. and Zhu, Z. (2018), "The generalized predictive control of bacteria concentration in marine lysozyme fermentation process", *Food Sci. Nutr.*, **6**, 2459–2465. https://doi.org/10.1002/fsn3.850.

GC