Research Article

Indirect Determination Approach of Blast-Induced Ground Vibration Based on a Hybrid SSA-Optimized GP-Based Technique

Zhaoxin Jiang,1 Hongyan Xu,2 Hui Chen,3,4,5 Bei Gao,2 Shijie Jia,2 Zhi Yu,5 and Jian Zhou5

1Xinjiang Xuefeng Sci-Tech (Group) Co., Ltd., Urumqi, Xinjiang 830047, China
2Xinjiang Xuefeng Blasting Engineering Co., Ltd., Urumqi, Xinjiang 830047, China
3School of Geology and Mines Engineering, Xinjiang University, Urumqi, Xinjiang 830047, China
4Key Laboratory of Environmental Protection Mining for Minerals Resources at Universities of Education Department of Xinjiang Uygur Autonomous Region, Urumqi, Xinjiang 830047, China
5School of Resources and Safety Engineering, Central South University, Changsha 410083, China

Correspondence should be addressed to Hui Chen; xjuchenhui@126.com

Received 23 December 2020; Revised 22 February 2021; Accepted 9 March 2021; Published 17 March 2021

Academic Editor: Haoyuan Hong

Copyright © 2021 Zhaoxin Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The accurate determination of blast-induced ground vibration has an important significance in protecting human activities and the surrounding environment. For evaluating the peak particle velocity resulting from the quarry blast, a robust artificial intelligence system combined with the salp swarm algorithm (SSA) and Gaussian process (GP) was proposed, and the SSA was used to find the optimal hyperparameters of the GP here. In this regard, 88 datasets with 9 variables including the ratio of bench height to burden (H/B) and the ratio of spacing to burden (S/B) were selected as the input variables, while peak particle velocity (PPV) was measured. Then, an ANN model, an SVR model, a GP model, an SSA-GP model, and three empirical models were established, and the predictive performance was evaluated by using the root-mean-square error (RMSE), determination coefficient ($R^2$), value account for (VAF), Akaike Information Criterion (AIC), Schwarz Bayesian Criterion (SBC), and the run time. After comparing, it is found that the proposed SSA-GP yielded a superior performance and the ratio of bench height to burden (H/B) was the most sensitive variable.

1. Introduction

With the development of science and technology, many rock fragmentation technologies have been invented and applied in the engineering area, but the blasting technique is still the method with the best economic benefit and the highest rock fragmentation efficiency both in the open-pit mines and underground mines [1]. However, a previous study [2] showed that more than half of the explosive energy was wasted, and blast-induced vibration, flyrock, blast-induced overpressure, etc. were caused during the blasting operation. Among them, blast-induced vibration is considered to be the most serious and most common blast-induced environment issue, so it is of great significance to predict and control the blast-induced vibration for avoiding the undesirable effects on the neighboring building and human activities, etc.

Usually, the blast-induced ground vibration is evaluated and measured by using peak particle velocity, frequency, and duration time. Among these three indexes, peak particle velocity (PPV) is the most common and widely used index in the previous literature [3–5]. Aiming at the accurate prediction of the peak particle velocity, some empirical equations were proposed and utilized in many cases such as Sungun Copper Mine, Iran [6], Miduk Copper Mine, Iran [7], Bakhtiar Dam, Iran [8], and Sinagreni Collieries Company Limited, India [9]. But, some research studies [10, 11] show that the predictive performance of these empirical models is not high when being utilized in peak particle velocity prediction.

Nowadays, artificial intelligence (AI) technology is developing rapidly and many new artificial intelligence technologies such as ANN, SVR, and RF are being
proposed and applied in practical cases such as predicting the field capacity in disaster management problems [12–18], mining production problems [19], etc. In practice, artificial intelligence (AI) technology is also being used to predict PPV values in many studies. For example, Arthur et al. [20] checked the efficiency of the Wavelet Neural Network (WNN) in predicting peak particle velocity, and satisfying prediction results were obtained, which proves the strong prediction capability of the Wavelet Neural Network (WNN). Monjezi et al. [21] used an artificial neural network (ANN) model which includes an input layer, two hidden layers, and an output layer to evaluate the ground vibration based on the Siabbishe project, which demonstrates the effectiveness of using an ANN to predict the blast-induced ground vibration. Li et al. [22] proposed two hybrid models using biogeography-based optimization (BBO), deterministic optimization algorithm (DIRECT), and artificial neural network, namely, BBO-ANN and DIRECT-ANN, and the generalization capability was found to be better than other prediction models. This study shows the use of an optimization algorithm to improve the prediction performance of the prediction model. Sheykhi et al. [23] combined the fuzzy C-means clustering (FCM) and support vector regression (SVR) to develop an accurate prediction model based on a database from the Sarcheshmeh copper mine, and the model performance of this new proposed hybrid model is introduced in this paper. In addition to the abovementioned research studies, many studies [24–30] were conducted by using artificial intelligence (AI) technology to predict the PPV for vibration control.

Although many studies were conducted previously, there is not a model that is suitable for all areas because every engineering site has its characteristics. Meanwhile, artificial intelligence (AI) technology is an evolving technology, and the predictive models in the artificial intelligence (AI) area are gradually updated and developed. The salp swarm algorithm (SSA) is a newly proposed metaheuristic algorithm, and this algorithm was checked using 7 unimodal test functions and 6 multimodal benchmark functions. Comparing with particle swarm optimization (PSO), gravitational search algorithm (GSA), bat algorithm (BA), firefly algorithm (FA), and genetic algorithm (GA), SSA shows superior and steady performance on average. Meanwhile, the successful application in solving the classical engineering design problem [31], airfoil design for aero vehicles problem [31], the strength of fiber-reinforced cemented paste backfill [32], and the compressive strength of concretes [33] also proves the high capacity of SSA. Besides, the Gaussian process was found to be an effective prediction technique after be used in landslide cases [34], broken rock zone prediction [35], carbon dioxide emission prediction [36], etc. Meanwhile, the conclusion that the GP performance can be improved after using the metaheuristic algorithm has been proved after many combinations such as PSO-GP [35] and GA-PSO [34] were proposed and tested. However, as far as the authors know, the combination of the SSA and Gaussian process (GP) has not to be proposed and tried in predicting PPV.

This paper proposes a new machine learning model, namely, SSA-GP, the Gaussian process (GP) model was optimized by the salp swarm algorithm (SSA), and the modeling process and the application of using the SSA-GP to predict the blast-induced ground vibration were introduced. This is an innovative work as the SSA-GP has not been analyzed and tried in predicting PPV before. The model development process and this application will promote the application of the artificial technique in solving mining and geotechnical problems and will be helpful for blast-induced ground vibration controlling.

2. Materials and Methods

2.1. Collected Database. To develop a high-precision evaluation method, a database generated by Hudaverdi [37] was used here for constructing the empirical, ANN, SVR, GP, and SSA-GP model. A total of 88 sets of data were monitored at the Akdaglar Quarry in northern Istanbul, Turkey, and several parameters including burden (B), spacing (S), bench height (H), stemming (T), subdrilling (U), hole diameter (D), power factor (PF), the distance between the monitoring station and blasting point (D), the weight of explosive charge detonated per delay (W), and peak particle velocity (PPV) were recorded. Among these parameters, burden is the distance between each explosive charge or the distance between the explosive charge to free face. Spacing is the distance between the two blastholes in a row. Bench height is the distance from the surface to hole bottom. Stemming is the distance from the top of the charging section to the surface. Subdrilling is the distance from the bottom of the bench to the bottom of blasthole. Power factor means the explosive charge per cubic meter. After reviewing the study of Hudaverdi [37], the ratio of bench height to burden (H/B), the ratio of spacing to burden (S/B), the ratio of burden to hole diameter (B/D), the ratio of subdrilling to burden (U/B), the ratio of stemming to burden (T/B), power factor (PF), the distance between the monitoring station and blasting point (D), and the weight of explosive charge detonated per delay (W) were selected as the input variables, while PPV was selected as the output variable. To describe the collected database, a scatter plot which shows the data distribution using diagonal, the bivariate scatter plots using the bottom of diagonal, the correlation, and significance level using the top of diagonal was plotted as shown in Figure 1. It should be noted that the symbol of "∗∗∗," "∗∗," and "∗" are, respectively, corresponding to the p values with the range of [0, 0.001], [0.001, 0.01], and [0.01, 0.05]. Meanwhile, the coefficient of relationship between the B/D and PF was found to be −0.91, so PF was removed from the input variable after reviewing the study of Rovini et al. [38].

Before the model development, both the input variables and output variables should be scaled into the range of 0 to 1 for avoiding the effect of the variable with the large number on the variable with the small number [12]. After scaling, 80% of the database which is normally called training datasets will be utilized to train and validate the prediction
model, the remained 20% of the database which is normally called testing datasets will be used to check the model performance [39, 40], and the consistency of data distribution of these two datasets can reduce the impact of the data partitioning process on model performance.

2.2. Empirical Techniques. The empirical equation of the United States Bureaus of Mines (USBM) [41, 42] is the most common and widely used technique for PPV prediction, and that equation can be described as

$$PPV = a \left( \frac{D}{W} \right)^{-b},$$

(1)

where $a$ and $b$ are the constants that are affected by the mine geology; $D$ is the distance between the monitoring station and blasting point; and $W$ is the weight of explosive charge detonated per delay.

In addition to the abovementioned empirical formula, two equations proposed by Davies et al. [43] and the Bureau of Indian Standard [44] were also selected.

Davies et al.:

$$PPV = aD^{-b}W^c.$$  
(2)

Bureau of Indian Standard:

$$PPV = a \left( \frac{W}{D^{2/3}} \right)^b.$$  
(3)

2.3. Artificial Neural Network (ANN). The ANN model is an important branch of the machine learning (ML) technique and is inspired by the human brain [45, 46]. With the help of computer calculation, many problems including blast-induced rock movement [47–49], blast-induced overpressure [50], rockburst [51], flyrock [52], and rock fragmentation [53, 54] can be solved by learning message from the input variables and using these messages to predict the output variables. After reviewing previous studies [55, 56], multilayer perception (MLP) which is composed of input layers, hidden layers, and output layers is the best type of neural network among many artificial neural networks. In the multilayer perception model, the layer in the multilayer perception consists of one or more nodes, and the line between the nodes show the information exchange.

To train the neural network, an efficient learning algorithm should be selected. According to the study of Jahed Armaghani et al. [57], Dreyfus [58], and Pedrycz et al. [59], the back-propagation (BP) algorithm is the most competent learning algorithm for MLP neural networks. During the process of the BP algorithm, the message learned from the input variables is used to predict the output variables, then the predicted values of the output variable will be compared with the actual values, and the error will be calculated [60]. After that, the error will be sent back for updating the individual weights in the neural network, and that process will be repeated until the error of the predict value and actual value reach a defined level [61].

---

**Figure 1:** Scatter plot of the established database.
2.4. Support Vector Regression (SVR). SVR was successfully applied in solving engineering problems including hydrocarbon reservoir prediction [62], the thermal conductivity of frozen soil prediction [63], and rock mass parameter [64] prediction. That algorithm is developed from the statistical learning theory, and the input variables were reflected into a high-dimensional space by using the kernel function [65, 66].

The kernel function provided by support vector regression (SVR) mainly includes three types such as the linear kernel, polynomial kernel, and radial primary kernel function [67]. Among these kernel functions, the radial primary kernel function was found to be the most efficient kernel function for higher predictive performance [68].

2.5. SSA-GP

2.5.1. Salps Swarm Algorithm (SSA). Normally, the hyperparameter of the prediction model such as ANN, RF, SVR, and GP is obtained from the trial-and-error method, but this method is time consuming and it is hard to find the optimal hyperparameter combination when there are several in the prediction model. The emergence of a metaheuristic algorithm provides a good way to solve this problem, and by using these algorithms, the optimal hyperparameter combination can be found after hundreds of iterations. In this area, many metaheuristic algorithms were inspired by the natural phenomenon such as animal behavior (Harris Hawks Optimization Algorithm (HHO) [69], Grey Wolf Optimizer Algorithm (GWO) [70, 71], etc.) and evolutionary behavior (Differential Evolution Algorithm (DE) [72], Genetic Algorithm (GA) [73, 74], etc.). The salp swarm algorithm is a metaheuristic algorithm inspired by the behavior of the salp swarm in the ocean [31]. Like the behavior of jellyfish, the salp moves forward by pumping water from the salp body. Meanwhile, the salp always forms a swarm in the deep ocean, and some research studies [31, 75] show that the swarming behavior (shown in Figure 2) of the salp can help the salp determine the location of food resource fastly and accurately. According to the research of Mirjalili et al. [31], Faris et al. [76], and Sayed et al. [77], the salp chains consist of a leader salp (the first salp in the salp chain) and follower salps (rest of the salps in the salp chain). The lead salp controls the moving direction, and the follower salps will follow the leader salp during the movement of the salp chain.

The updating rule of the position of the leader and follower salp is given by [31]

\[
x_j^i = \begin{cases} F_j + c_1 \left((ub_j - lb_j)c_2 + lb_j\right), & c_3 \geq 0, \\ F_j - c_1 \left((ub_j - lb_j)c_2 + lb_j\right), & c_3 < 0, \end{cases}
\]

with

\[
c_1 = 2 \exp \left(-\frac{4l^2}{L^2}\right),
\]

where \(x_j^i\) and \(F_j\) are, respectively, the position of the leader salp in the salp chain and the position of food resource; \(ub_j\) and \(lb_j\) are, respectively, the upper bound and lower bound of \(j_{th}\) dimensional; \(c_1, c_2, c_3\) are random numbers; and \(l\) and \(L\) are, respectively, the current and maximum number of iteration.

After obtaining the position of the leader salp in the salp chain, the position of the following salps can be expressed as follows:

\[
x_j^i = \frac{1}{2} \left(x_j^i + x_j^{i-1}\right),
\]

where \(x_j^i\) represents the position of \(i_{th}\) salp in the \(j\) dimension and \(i\) is the number greater than or equal to 2.

2.5.2. Gaussian Process (GP). The Gaussian process (GP) can find a relationship between the input variable value and the output variable value of the training datasets, and the predicted output variable value of the testing datasets can be calculated using the created relationship. According to the study of Yu et al. [35], Arthur et al. [79], and Fang et al. [36], a GP model can be determined by using the mean function \(m(x)\) and covariance function \(k(x, x')\), and that model \(f(x)\) is defined as follows:

\[
f(x) \sim \text{GP}(m(x), k(x, x')),
\]

subjected to

\[
\begin{align*}
   & \{m(x) = E[f(x)], \\
   & k(x, x') = E[[f(x) - m(x)][f(x') - m(x')]].
\end{align*}
\]

Assume \(D = (x_i, y_i)\) is the training dataset of the Gaussian model, where \(x_i\) and \(y_i\) are, respectively, the input and output vector.

Then, the standard linear regression model is determined using the following formula:

\[
y = f(x) + \varepsilon,
\]

where \(\varepsilon\) is an independent random variable, \(\varepsilon \sim N(0, \sigma_n^2)\), where \(\sigma_n^2\) is the variance.

Given the new testing input \(x'\) and the training set \(D\), the goal of the GP model is to calculate the \(y'\) using the posterior probability formula.

The parameters in the mean function \(m(x)\) and covariance function \(k(x, x')\) constitute the hyperparameters of the Gaussian process model, which are the only parameters to be determined in the Gaussian process.

2.5.3. Hybrid of the SSA and GP. During the development of the GP model, the combination of hyperparameters is a very important issue for achieving higher predictive performance. Normally, the optimal combination of hyperparameters in the Gaussian process was found by using the conjugate gradient algorithm, but it is easy to fall into local optimum and that algorithm is strongly dependent on the initial value.

After scaling the collected datasets into the range of 0 to 1, the collected database will be randomly split into the training (80%) and testing datasets (20%). The SSA algorithm was used to find the optimal combination of
hyperparameters by learning knowledge from training datasets, and the obtained hyperparameter combination will be used to establish the optimal GP model (see Figure 3). In this process, the 10-fold cross-validation method was applied to calculate the fitness value for hyperparameter searching and optimal GP model development. During the 10-fold cross-validation method, the training datasets were split into 10 folds; 9 of them will be utilized to develop the prediction model, and the remained one will be used to test the performance of that model. After 10 runs, each dataset will have the opportunity to be learned and tested, and the overfitting and underfitting phenomenon can be avoided. Finally, the testing datasets will be used to evaluate the established optimal GP model by using performance evaluation metrics.

In this study, the blast-induced ground vibration models were developed in the MATLAB environment, and the computation code was programmed using MATLAB language and implemented on the computer with Intel (R) Core (TM) i7-7500U CPU running at 2.70 GHz and 2.90 GHz.

3. Results and Discussion

3.1. Results of Empirical Models. For the empirical model, two constants are determined by using the MATLAB code based on 70 training datasets which are the same as those used in the proposed ANN, SVR, GP, and SSA-GP model. After regressing, these fitted equations were obtained and are as follows:

\[
PPV = 167.8 \times \left( \frac{D}{W} \right)^{-0.9846}, \tag{10}
\]

\[
PPV = 370.6 \times D^{-1.21} \times W^{0.5221}, \tag{11}
\]

\[
PPV = 1.901 \times \left( \frac{W}{D^{2/3}} \right)^{1.215}. \tag{12}
\]

After formula regression, the testing datasets (18 blasting datasets) were predicted using the abovementioned evolution formula, the model performance of equation (10) was found to be an \( R^2 \) of 0.64 and 0.67, RMSE of 4.92 and 3.94, and VAF of 70.45 and 75.36, the model performance of equation (11) was found to be an \( R^2 \) of 0.70 and 0.75, RMSE of 4.44 and 3.43, and VAF of 70.44 and 74.74, and the model performance of equation (12) was found to be an \( R^2 \) of 0.46 and 0.47, RMSE of 6.00 and 4.97, and VAF of 45.92 and 47.67.

3.2. Results of the ANN Model. To check the performance of using the ANN model to forecast the peak particle velocity, various ANN models should be built for optimal ANN model development. In the ANN model, the number of hidden layers and hidden nodes significantly affect the model performance due to the different weights and biases. An ANN model with too many hidden layers and hidden nodes may lead to the excessive learning of the training datasets, while too few of them will lead the ANN model to not be able to effectively learn the knowledge from the training datasets. Here, we constructed 17 ANN models with 1 hidden layer and a different number of hidden nodes ranging from 1 to 17 after following the suggestion of Mohamad et al. [80] and Hecht-Nielsen [81]. With the help of the 10-fold cross-validation method, an ANN model with 1 hidden layer and 3 hidden neurons in this hidden layer shows the best results (\( R^2 \) of 0.85 and 0.83, RMSE of 3.14 and 2.78, and VAF of 85.29 and 83.86 for training and testing datasets) and determined as the prediction model to evaluate peak particle velocity in this research.

3.3. Results of the SVR Model. Similar to the development of the ANN model, the blast datasets were converted, and then, a grid search method (GSM) was used to search the optimal hyperparameter combination for PPV prediction. In the SVR
model, two hyperparameters including penalty factor \((C)\) and gamma \((g)\) in the RBF kernel should be determined. Here, the search scope of \(\log 2(C)\) and \(\log 2(g)\) are \(-10.0\) to \(10.0\), the cross-validation method was applied to guide hyperparameter selection, and the hyperparameter is optimal when the fitness value is the smallest. As a result, the optimal combination of \(C\) and \(g\) was found with \(C = 10.556\) and \(g = 0.435\). It is found that the \(R^2\) values are 0.87 and 0.84, the RMSE values are 2.91 and 2.71, and the VAF values are 87.32 and 84.34 for training and testing datasets, respectively.

### 3.4. Results of the GP Model

Using the conjugate gradient algorithm, the hyperparameter combination of an original Gaussian process model was searched after learning knowledge from the training datasets. After training, the Gaussian process model yielded a prediction performance of \(R^2 = 0.86\), RMSE = 3.00, and VAF = 86.46 for training datasets and \(R^2 = 0.87\), RMSE = 2.43, and VAF = 87.40 for testing datasets.

### 3.5. Results of the SSA-GP Model

For developing the SSA-GP model, the number of salps in the salp chain and the maximum iteration should be determined first. Then, 9 SSA-GP models with the number of salps ranging from 20 to 300 and maximum iteration of 500 were established, the fitness curve calculated from the 10-fold cross-validation method of each SSA-GP model was recorded, and the hyperparameter combination is the best when the fitness value is smallest. After setting parameters for the SSA algorithm, the salps find the optimal location for the GP models with the corresponding parameters. Model results show that the number of salps does not have a significant effect on the fitness curve, and then, 20 and 500 were selected to be the optimal number of slaps and a maximum number of iterations.

After determining the hyperparameters of the SSA-GP model, the optimal SSA-GP model was established and evaluated by using \(R^2\), RMSE, and VAF. After predictive performance evaluation, the results show that the measured PPV values agree well with the predicted PPV values with \(R^2 = 0.88\), RMSE = 2.78, and VAF = 88.38 for the training datasets and \(R^2 = 0.89\), RMSE = 2.25, and VAF = 89.37 for testing datasets.

### 3.6. Performance of Various Models

For the comparison of the model performance, three performance metrics including \(R^2\), RMSE, and VAF were applied, and a prediction model can be considered as the best model when \(R^2 = 1\), RMSE = 0, and VAF = 100. Meanwhile, the value of these performance metrics can be calculated using the following formula [32, 82–87]:

\[
R^2 = 1 - \frac{\sum_{i=1}^{N} (y - y_{pre})^2}{\sum_{i=1}^{N} (y - \bar{y})^2},
\]

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{N} (y - y_{pre})^2}{N}},
\]

\[
VAF = 100 \times \left(1 - \frac{\text{var}(y - y_{pre})}{\text{var}(y_{pre})}\right),
\]

\[\text{(13)}\]

\[\text{(14)}\]

\[\text{(15)}\]
where \( N \), \( \bar{y}, y \), and \( y_{\text{pre}} \) are the number of datasets, the average PPV values, the actual PPV values, and the predicted PPV values, respectively.

Besides the statistical criteria shown above, the Akaikie Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC) were also utilized after reviewing the study of Agoubi and Kharroubi [88] and Phiri et al. [89].

\[
\text{AIC} = \ln \left( \frac{\sum_{i=1}^{N} \epsilon_i^2}{N} \right) + \frac{2k}{N} \\
\text{SBC} = \ln \left( \frac{\sum_{i=1}^{N} \epsilon_i^2}{N} \right) + k \ln \left( \frac{N}{N}_i \right),
\]

where \( \epsilon \) is the residual value and \( k \) is the number of estimated coefficients.

For comparing the prediction performance of the empirical model 1, empirical model 2, empirical model 3, ANN model, SVR model, GP model, and SSA-GP model, the datasets (88 measured PPV datasets) were compared with the 88 predicted PPV datasets, as shown in Figure 4. It shows that the predicted blast-induced ground vibration value obtained from the empirical model 1, empirical model 2, empirical 3, ANN model, SVR model, GP model, and SSA-GP model and the measured blast-induced ground vibration shows the same data distribution trend. Meanwhile, the errors between these measured 88 datasets and the predicted 88 PPV datasets obtained from the empirical model 1, empirical model 2, empirical 3, ANN model, SVR model, GP model, and SSA-GP model were plotted in Figure 5. It can be seen that error of the SSA-GP in each sample is smaller. The performance shows that the predicted PPV values provided by SSA-GP are closer to the measured PPV values than the PPV values provided by the other 6 prediction models, while the difference between the measured PPV value and the value predicted by the empirical model 3 is the biggest.

By comparing the results shown in Table 1, the proposed SSA-GP model yield better prediction performance in PPV prediction due to the higher \( R^2 \) and VAF value, lower RMSE value, and higher summing ranking value obtained from the ranking method proposed by Zorlu et al. [90]. Meanwhile, the prediction performance of empirical model 3 is the worst among these 7 prediction models. Notably, the VAF value of the empirical model 3 in PPV prediction was only 45.92 and 47.67 for training datasets and testing datasets, whereas these values of the proposed SSA-GP model were 88.38 and 89.37. The values of those performance metrics prove that the proposed SSA-GP model is the most dominant model for PPV prediction.

Also, the PPV value compassion shown in Figure 6 expresses that the errors between the actual PPV value and the value predicted by the SSA-GP model are smaller, which also means that the difference between actual PPV values and the predicted values obtained from the ANN model, SVR model, GP model, empirical model 1, empirical model 2, and empirical model 3 is greater. Besides the prediction model provided in this study, the same database was studied by Hudaverdi [37] using multivariate analysis, and the prediction performance of SSA-GP was found to be superior.

Besides \( R^2 \), RMSE, and VAF, the AIC and SBC values were also calculated using equations (12) and (13), the AIC value of the empirical model 1, empirical model 2, empirical model 3, ANN, SVR, GP, and SSA-GP is 6.13, 1.04, −1.09, 2.45, 1.15, −2.80, and −3.46, and the BIC value of the empirical model 1, empirical model 2, empirical model 3, ANN, SVR, GP, and SSA-GP is 6.33, 1.23, −0.89, 2.65, 1.35, −2.60, and −3.26, respectively. Normally, the smaller the AIC and SBC value, the better the suitability of the prediction model. Therefore, SSA-GP was found to be the best prediction model among these 7 models.

The run time of developing a prediction model was recorded to check the time complexity of the prediction model. In this investigation, the run time from data preprocessing to dataset prediction was recorded, and the run time of the ANN model, SVR model, GP model, and SSA-GP model was 0.20, 31.57, 0.46, and 405.24 seconds, respectively.

It is found that the SVR model and SSA-GP model take a longer time than ANN model and GP model, and the reason for this is that the hyperparameter optimization process of the SVR model and SSA-GP model was contained in the model development process. Although more time needs to be used in SSA-GP development, it can be accepted when comparing with the model performance improvement.

3.7. Performance of Dependency Analysis. After model development and performance comparison, a performance dependency analysis was carried out to analyze the effect of the number of training datasets (NTD) on the model performance. In this analysis, 7 ANN models, 7 SVR models, 7 GP models, and 7 SSA-GP models with the NTD equal to 10, 20, 30, 40, 50, 60, and 70 were developed, and 10 unused datasets in the collected database were used to check the model performance. It can be seen from Figure 7 that the prediction accuracy evaluated by the \( R^2 \) value of these models gradually increases and finally tend to be stable when NTD is larger than 40. This phenomenon shows that the developed prediction model can provide stable results when used in the engineering site.

3.8. Sensitivity Analysis. To identify the most sensitive parameter, a sensitivity analysis was carried out and a cosine amplitude method [91] was utilized. In this method, the sensitive value is 0 when that variable is least important and is 1 when that variable is most important. After calculating, the sensitive value of the ratio of bench height to burden (H/B), the ratio of spacing to burden (S/B), the ratio of burden to hole diameter (B/D), the ratio of subdrilling to burden (U/B), the ratio of stemming to burden (T/B), the distance between the monitoring station and blasting point (D), and the weight of explosive charge detonated per delay (W) is 0.8519, 0.8503, 0.8459, 0.8379, 0.8514, 0.8249, and 0.6712, respectively. The results show that the ratio of bench height to burden (H/B) is the most sensitive parameter among these variables. Meanwhile, there is little difference between the ratio of bench height to burden (H/B), the ratio of spacing to burden (S/B), the ratio of burden to hole diameter (B/D), the ratio of subdrilling to burden (U/B), the ratio of stemming to burden (T/B), and the distance between the monitoring station and blasting point (D).
burden (T/B), and the distance between the monitoring station and blasting point (D). Normally, the distance between the monitoring station and blasting point (D) and the weight of explosive charge detonated per delay (W) were considered in the empirical prediction model, but the sensitivity analysis shows that more variables should be considered in future investigation.

3.9. Superiority and Limitations. After studying, a novel SSA-GP model was proposed and utilized to predict the blast-induced ground vibration caused by blast operation in Akdaglar Quarry. This method is inexpensive and has high precision, and the related modeling process of SSA-GP can guide the development of other hybrid models. Compared with three empirical prediction models, ANN model, and
Figure 6: Continued.
The proposed SSA-GP model can accurately predict the blast-induced ground vibration, and the accuracy of that model can be easily improved by enlarging the collected database or collecting more variables. Although some achievements have been obtained, some limitations were also found and should be considered in future investigations. First, the developed SSA-GP model in this study is a black box method and may be difficult for mining engineers when comparing with an explicit equation. Then, only a small database with 88 datasets and 9 variables was collected from the previous study, which may affect the precision of the developed model and the sensitivity analysis results. Meanwhile, the use of the ratio of two parameters and the lack of consideration of geology and explosive parameters may make the prediction a bias, so these problems should be considered in future investigations. So, the collection of a bigger database with more datasets and more variables such as rock type and detonation delay time can be considered in future investigation. Also, developing an SSA-GP model needs more time when comparing with the ANN, SVR, and GP model, but it is meaningful. Last, only the GP model was optimized by the SSA in this paper. Although the SSA is a newly proposed metaheuristic algorithm and the optimization performance was proved when comparing with PSO, GSA, BA, FA, and GA, in the previous study, a more detailed comparison of using the SSA to optimize the hyperparameters of the prediction model such as the ANN, SVR, or RF or the comparison of using the different metaheuristic algorithm to optimize the GP model is meaningful.

\[ y = 0.8x \]

\[ y = 1.2x \]

\[ V_{AF} = 89.37 \]

\[ RMSE = 2.25 \]

\[ R^2 = 0.88 \]

\[ R^2 = 0.89 \]

\[ V_{AF} = 88.38 \]

\[ RMSE = 2.78 \]

\[ R^2 = 0.89 \]

\[ RMSE = 2.25 \]

Figure 6: Measured versus predicted PPV value: (a) empirical model 1; (b) empirical model 2; (c) empirical model 3; (d) ANN model; (e) SVR model; (f) GP model; and (g) SSA-GP model.

Figure 7: Performance dependency analysis results.
As a machine learning model, the development of the prediction model is based on the collected database, so the developed SSA-GP model is only suitable for Akdaglar Quarry. But, the mining engineers in other mines can use the modeling process provided in this paper for developing their prediction model. Nevertheless, this study is a powerful supplement to the blast-induced ground vibration prediction and can be utilized to control the blast-induced hazard.

4. Conclusions
For an open-pit mine, blasting is a very important part of the mining process, and the quality of blasting has a significant impact on the mining economy. However, some unwanted effects, especially blast-induced vibration, can usually be found due to the wastage of explosive energy and lead to safety risks to the surrounding resident’s lives and buildings. It is, therefore, stated that the high-precision prediction of peak particle velocity (PPV) is meaningful and can provide help for both mining engineers and the government.

For the easy-operated, inexpensive, and accurate blast-induced ground vibration determination, a hybrid SSA-optimized GP-based model, namely, SSA-GP, was proposed based on an 88-dataset database. During the model development process, the training datasets (70 datasets) were learned with the help of the 10-fold cross-validation method and SSA optimization method, then the testing datasets (18 datasets) were predicted, and the performance was checked using some performance metrics. Then, its performance was analyzed and compared with the empirical, ANN, SVR, and GP models. The comparison results show that the proposed SSA-GP yields the promising reliability with an $R^2$ of 0.88 and 0.89, RMSE of 2.78 and 2.25, and VAF of 88.38 and 89.37 for training and testing datasets and the smallest AIC and SBC value. Performance dependency analysis results show that the developed prediction model can provide a stable prediction performance when forecasting the unused datasets. Sensitivity analysis shows that more parameters should be considered in the future empirical prediction model development rather than the explosive charge of each delay and the distance from the blast block to the monitoring station. Through the acquired results, the superior optimization ability of the SSA was verified, and that algorithm has the potential to be applied in other prediction issues.

Data Availability
Previously reported data were used to support this study and are available at [10.1016/j.soildyn.2012.08.002]. These prior studies (and datasets) are cited at relevant places within the text as references.

Conflicts of Interest
The authors declare that there are no conflicts of interest regarding the publication of this paper.

References


Advances in Civil Engineering 13


