Journal of Mines, Metals and Fuels, *70(4) : 203-213; 2022. DOI: 10.18311/jmmf/2022/30057*

Print ISSN : 0022-2755

Journal of Mines, Metals and Fuels

Contents available at: www.informaticsjournals.com/index.php/jmmf

Prediction of Peak Particle Velocity of Blast-induced Ground Vibrations using Boosted Regression Trees Authored

Rajesh Sonkar[1*], Prakash Y. Dhekne[1] and Narendra D. Londhe[2]

[1]Mining Engineering Department, National Institute of Technology, Raipur, Chhattisgarh – 492010, India; rajeshsonkariocl@gmail.com

[2]Electrical Engineering Department, National Institute of Technology, Raipur, Chhattisgarh – 492010, India

Abstract

Loosening of rockmass during its excavation in an infrastructure project is carried by rock blasting. The blast-induced ground vibrations pose a major challenge to the blasting engineers, whose main objective is to control their potential to cause any damage to the buildings in the vicinity. The research reported in this paper explains how the error in the prediction of the Peak Particle Velocity (PPV) by the United States Bureau of Mines (USBM)-based approach can be minimised using machine learning techniques. The complex correlation between the blast parameter and the PPV value has been modelled using the least square boosted decision tree approach after the selection of the best suitable feature has been selected based on the correlation matrix. The proposed model automatically maps the input blast feature (SD) with the target PPV values by aggregating the decision of various weak learners. The generalization of the proposed model has been validated through a 5-fold cross-validation approach using a dataset comprising of two hundred blast records generated by monitoring the blasts at International airport site, Navi Mumbai, India. The assessment of the prognostic ability of the proposed model demonstrates that it has outperformed the USBM-based approach for PPV prediction. The results establish that the predictions by the proposed model are closer to the measured values than the other regression models.

Keywords: *Blast-induced Ground Vibrations, Boosted Regression Tree, Linear Regression, PPV Prediction Model, Stepwise Regression*

1. Introduction

The development of infrastructure is necessary for boosting the economy. As a result, many infrastructure construction projects are underway, round the globe. Blasting is a widely accepted technique to loosen the hard rock before its excavation during the construction of infrastructure projects. When an explosive is detonated, it releases energy. The useful energy, which is nearly thirty per cent of the total energy released, is utilized for rock breakage whereas the remaining energy manifests itself as fly rock, air over pressure, blast-induced ground vibration, etc. The important parameters of blast-induced ground vibrations, to assess their potential for causing damage, are the PPV, acceleration, and frequency. However, the PPV is frequently used to estimate the intensity of the blast-induced ground vibrations and the investigators have extensively applied it for modelling (Amiri, 2016). The phenomenon of ground vibrations is annoying to the population residing in the houses near the blasting site. When the PPV of

**Author for correspondence*

Table 1. Important research on the AI based PPV prediction models

	Authors, Year	Algorithms/ Models	No. of datasets	Input variables	**Remarks**
1	Peng *et al.*, 2021	GBT and Optimized ANN Models	93	MC, T, U, B/S RQD, CF	ANN, ANN Forward, ANN-Backward ANN-PSO, ANN-GA, exhibited R= 0.879, 0.927, 0.941,0.945, 0.934, respectively.
2	Zhang *et al.*, 2020	PSO and XGBoost	175	MCPD, CF, R, B, and S	RMSE= 0.583, R^2 = 0.968, MAE=0.346, and VAF = 96.083%
3	Ding *et al.*, 2020	ICA and XGBoost	136	MCPD, R, B, CF, T, S and H	RMSE= 0.736, R^2= 0.988and MAE= 0.527
4	Qiu *et al.*, 2021a new intelligent method for predicting peak particle velocity (PPV	hybrid WOA-XGBoost, GWO-XGBoost and BO-XGBoost	150	d, D, B, S, MCPD, CL, R, BI, E, ν, P_v, VoD, ρ_e	RMSE, R^2, VAF, and MAE = (3.0538, 0.9757, 97.68, 2.5032), (3.0954, 0.9751, 97.62, 2.5189), and (3.2409, 0.9727, 97.65, 2.5867) for WOA-XGBoost, GWO-XGBoost, and BO-XGBoost models respectively.
5	Yang *et al.*, 2020	ANFIS optimized by PSO and GAs	86	B, S, T, CF, MCPD	ANFIS–PSO results in approximately 53% reduction in RMSE and 9% rise in R^2 And ANFIS–GA model revealed a fall of approximately 61% in RMSE and 10% improvement in R^2 compared to ANFIS.
6	Fattahi *et al.*, 2021	RVR Optimized by grey wolf optimization (GWO) and with BA	95	MCPD, B/S, T and R	R and MSE are (0.915, 7.920) and (0.867, 8.551) for RVR-GWO and RVR-BA models.
7	Dehghani *et al.*, 2021	GEP and TLBO	36	d, D, U, n, B, S, T, CL, CF	A comparative analysis of R^2, RMSE, MAPE showed the better performance of TLBO algorithm compared to the GEP.
8	Chen *et al.*, 2021	hybridizing FA, GA PSO with SVR ANN	95	MCPD, B/S, T, E, P_v, R	R^2 and RMSE (0.984, 0.614), (0.977, of 0.725), (0.964, 0.923),(0.957, 1.016), (0.936, 1.252), (0.925, 1.368), (0.924, 1.366) for MFA–SVR, PSO–SVR, FA–SVR, GA–SVR, GA–ANN and the PSO–ANN respectively. It confirms the advantage of MFA-SVR over other models.
9	Shang *et al.*, 2020	FA and ANN	83	MCPD, S, R, B, and CF	FA-ANN model RMSE= 0.464, MAE= 0.356, R^2= 0.966 and VAF= 96.620.
10	Ding *et al.*, 2021	bagged SVR optimized with FA	87	MCPD, R	R^2= 0.996, 0.896 and 0.828 for BSVR–FA, BPNN and RBFN respectively.
11	Zhou *et al.*, 2021	Jaya-XG boost	150	d, D, B, S, CL, Q, W, BI, E, ν, Pv, VOD, ρ_e	Jaya hybrid model obtained MAE (0.0008/3.2757), RMSE (0.0012/4.0884), VAF (99.99/95.73) and R^2 (1.0/0.9573).
12	Zhang *et al.*, 2020	Combination of FS and RFT	102	B/S, R, T, MCPD, CF, D	ANFIS models with R^2= 0.98, 0.93 and 0.92 for ANFIS models, SVM and hybrid models (PSO–ANN and ICA–ANN) for ANN-FIS models.

13	Zeng *et al*., 2021	Combinations of Boosted-CHAID and SVM Models with Various Kernels viz. SIG; POL; LIN; and RBF	166	B/S, R, T, MCPD, CF, S	all SVM models except SVMSIG outdid the ANN models while were hybridized with the Boosted-CHAID method.
14	Nguyen *et al*., 2020	SVR-Based PSO, GA, ICA, ABC	101	MCPD, R, B, and S	GA-SVR-RBF model was found to be the best technique for PPV prediction.

the ground vibrations exceeds the threshold then it may lead to damage to the structures, which causes avoidable litigations. It is therefore the objective of a blasting engineer to control the blast-induced ground vibrations so that they neither antagonise the nearby population nor have the potential to pose any danger to the nearby structures. Now-a-days, most of the construction sites are close to the human dwellings. Consequently, the PPV control is vital to safeguard the safety of neighbouring buildings and the dwellings. This requires a careful assessment of the PPV of the blast-induced ground vibrations due to a planned blast design. A lot of experimental equations have been developed by the researchers for the estimation of the PPV (Ambraseys and Hendron, 1968; Davies, 1964; Duvall and Petkof, 1959; Ghosh and Daemen, 1983; Gupta, 1988; Rai and Singh, 2004). Amongst these approaches, the USBM equation developed in (Duvall and Petkof, 1959) is popular for the prediction of the blast-induced ground vibrations. Over the last few decades, with the advent of fast computing machines, various Artificial Intelligence (AI) based models have been developed to accurately forecast the PPV of the blast-induced ground vibrations. Table 1 lists the important research on the AI-based PPV prediction models.

Note: where, GBT: Gradient boosted tree; PSO: Particle Swarm optimisation; ANN: Artificial Neural Network; XGBoost: Extreme gradient boost; ICA: Imperialistic competitive algorithm; WOA: whale optimisation algorithm; GWO: Grey wolf optimisation; BO: Bayesian optimisation; ANFIS; Adaptive neuro-inference fuzzy system; GA: Genetic algorithm; RVR: Relevance vector regression; BA: Bat-inspired algorithm; GEP: gene expression programming; TLBO: Teaching-learning based optimisation; FA: firefly algorithm; SVR: support vector regression; FS: Feature selection; RFT: random forest technique; CHAID: Chi-square automatic interaction detection; SIG: Sigmoid; PLO: Polynomial; LIN: Linear; ABC: Artificial bee colony; RBF: Radial basis function; B: Burden; S:Spacing; T: Stemming length; CF: Charger factor; MCPD: Maximum charge per delay; R: Distance of the seismograph from the blast site; H: Elevation of the seismograph from the blast site; d: diameter of blasthole; D: Depth of blast hole; BI: Blastability index; CL: Charge length; ν: Poisson's ratio; P_ν: Primary wave velocity; ρ_e: Density of the explosive; E: Young's modulus; n: Number of holes in arow; U: subgrade drilling; VoD: Velocity of detonation; R^2: Coefficient of determination; R: Coefficient of correlation: MAE: Mean absolute error; MSE: Mean square error; RMSE: Root mean square error; MAPE: Mean absolute percentage error; VAF: Variance accounted for.

Table 1 shows that a plenty of research is already carried out with the sole objective of convalescing the accurateness in the prediction of the PPV of the ground vibrations. Recently, the research is directed towards the application of machine learning techniques in ground vibration predictions. Further, it is also observed that the accuracy of the prediction can be improved if the clustered data is used for prediction rather than the raw data gathered from the blasts. In India, the regulatory authority, Directorate General of Mines Safety, (DGMS) has adopted the USBM equation for the prediction of PPV of the ground vibration for the mining operations and prescribes the use of the same in mine blasting operations. On similar lines, many infrastructure projects also use the same equation for the prediction purpose. Although the equation is generally used for PPV prediction, it might provide slightly lower performance due to noise present in the data. Therefore, in this study, we have proposed a machine learning-based approach that directly learns input (Scaled distance) output (PPV) mapping from the raw data. The data sets in the present research have been generated through the blasting operations carried out at the site of international Airport, Navi Mumbai, India.

1.1 Site of Study

Excavation is going for the land development work at the site of international airport at Navi Mumbai, India. The area extends 2 km in north-south and 700 m in east-west. The excavation site is shown in Figure 1. The site consists of various hillocks that reach a height of up to 50 m. The rockmass at the site comprises of Basalt and certain areas of Amygdaloidal Basalt. Table 2 list the geotechnical parameters of the rocks. Table 3 gives an overview of the excavation method using drilling and blasting. There are numerous structures nearby to the site of excavation. The structures either belong to the contractors who have been engaged for the excavation or to the villagers residing in the vicinity. The structures belonging to the contractors have been constructed mainly by using bricks and cement mortar whereas those belonging to the villagers have been constructed using the mud and bricks. The structures belonging to the contractors are within 200 m from the site of blasting whereas those belonging to the villagers are at 150 m to 220 m from the blast site. In addition to these, there are two temples at 75 m and 150 m from the site of the blasts. The temples are constructed using brick and cement. It is evident that the structures belonging to the villagers are weak because of the nature of the building material and the temples are the sensitive structures. This calls for a good vibration control so that no damage is inflicted upon the structures belonging to the villagers and the sensitive structures.

Table 2. The geotechnical parameters of the rocks

Parameters	**Basalt**	**Amygdaloidal Basalt**
Category of rock	Very Good	Good
RQD, %	75-90	50-75
Average joint spacing, mm	60-200	60-200
Average density, g/cc	2.8	2.5
UCS, M Pa	75-125	50-80

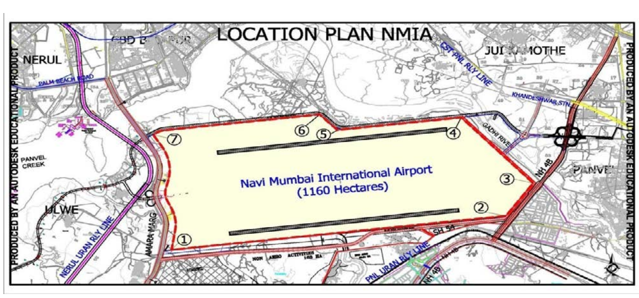


Figure 1. Excavation site at International airport at Navi Mumbai, India (Sonkar et al., 2021).

Table 3. An overview of the excavation method using drilling and blasting

Division of area	**The area is divided in four zones and each of them is called as package I, package II, package III and package IV**
Bench height, m	6 to 12
No of holes fired in a round	25 to 100
Blast holes diameter, mm	110
Average burden, m	3 to 3.5
Average spacing, m	3.5 to 4
Stemming, m	3
Subgrade drilling, % of bench height	5 to 10
Average charge ANFO, kg	Mostly in the range of 38.78 to 94.18
Hole to hole delay, ms	17
Down-the-hole delay, ms	250
Row to row delay, ms	25 between the first and second row; 43 between the second and third row.
Blasting technique	Muffled blasting
Excavator bucket capacity, m^3	0.9
Dumper capacity, m^3	8

1.2 Development of Predictor Equation

The blast geometry parameters affect the outcome of a blast (Sonkar *et al.*, 2021). The important geometry parameters for a blast i.e. spacing, burden, stemming height hole diameter, depth of hole including sub-grade drilling, along with the maximum charge per delay have been noted for two-hundred blasts and fifty of the same (25%)have been measured to ensure the exactness of the data. Instantel make Micromate seismograph has been used to measure the PPV of the blast-induced vibrations. At the time of the readings, a minimum count of 1 mm/s was set, and the seismograph's distance from the blast site was recorded using GPS.

The literature has a mention of various models for the PPV prediction of the blast-induced vibrations. But the site owners use the USBM model for PPV prediction among the models available in the literature. Equation 1 represents the mathematical relation between PPV and scaled distance factor for the said regressive model that is shown below:

$$ppv = k\left(\frac{D}{\sqrt{Q}}\right)^{-b} \quad (1)$$

where, D represents the seismograph's distance from the blast site, Q denotes the maximum charge per delay and the site constants b and k are associated to characteristics of local rock that are evaluated by regression analysis. The ratio $D/\sqrt{Q}$ is called as scaled distance factor (SD). Based on the wide use of the USBM predictor, the values of the site constants from the 149 recorded blasts have been evaluated and the final regressive equationis shown by Equation 2 (Sonkar *et al.*, 2021).

$$ppv = 108.08\ SD^{-1.014} \quad (2)$$

2. Methodology

This section presents a detailed description of the proposed approach for the PPV prediction using boosted decision trees (BRT) approach. The overall block diagram of the proposed technique is depicted in Figure 2. The proposed approach consists of five stages namely, data matrix preparation, data normalization, BRT model training, validation using k-fold cross-validation and PPV values prediction. In the 1st stage, the data matrix has been prepared in which row (r) indicates the number of blast samples and column (c) indicates the corresponding features. The final column contains the continuous PPV values which are utilized as target values. Further, in the 2nd stage, the best feature column has been selected by analyzing the correlation matrix. The feature vector has been partitioned into 5 folds in which data from the 4th and 5th fold are used for training and validation respectively. During this cross-validation phase, the hyper parameters of the BRT model are fine-tuned to reduce the prediction loss (i.e., Mean Square Error (MSE)). Once the MSE loss is minimized the BRT model is used for the PPV prediction in the final stage.

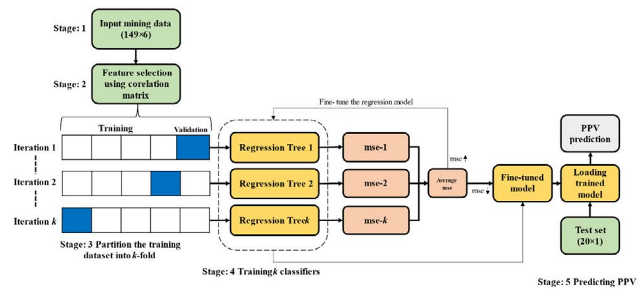


Figure 2. Pictorial representation of PPV prediction using the proposed BRT approach.

2.1 Feature Selection for Predictive Modelling

Feature selection is one of the crucial prerequisites to train a robust machine learning model. To validate which feature (i.e., blast parameter) is correlated with PPV a correlation matrix has been analysed as shown in Figure 3.

Algorithm 1. Least Square Boosting Regression Trees (Freund *et al.*, 2017)

M= number of iterations, initialize $\hat{r}^0 = y, \hat{\beta}^0 = 0, l = 0$

Step 1:For $0 \le l \le M$ **do**:

Step 2: Find the covariate j_l&u_{jl}as follows:

$$\hat{u}_m = argmin\left(\sum_{i=1}^{n}\left(\hat{r}_i^l - x_{im}u\right)^2\right) for\, m = 1,\ldots,p,\ j_l \epsilon \ \mathrm{argmin}\left(\sum_{i=1}^{n}\left(\hat{r}_i^l - x_{im}\hat{u}_m\right)^2\right)$$

Step 3:Updating the current residuals and regression coefficient:

$$\hat{r}^{l+1} \leftarrow \hat{r}^l - \epsilon X_{jl}\hat{u}_{jl}$$

$$\hat{\beta}_{jl}^{l+1} \leftarrow \hat{\beta}_{jl}^l + \epsilon\, \hat{u}_{jl} \;\&\; \hat{\beta}_j^{l+1} \leftarrow \hat{\beta}_j^l,\ j \ne j_l$$

It can be observed that the values of SD are highly correlated (i.e., negative correlation) with PPV values. A negative correlation means that if one variable increases another decrease. Therefore, in this study to train the BRT model we have selected only the SD parameter.

2.2 Least Square Boosted Regression Trees (BRT) for PPV Prediction

Boosting is one of the most popular techniques for classification and regression task. This technique improves the performance of the model by combining multiple weak models. As compared to conventional regression models (i.e., linear regression) the BRT-based approach combines different weak models to improve the perdition performance. In the BRT model, the first step is to select the proper devisor which can separate the dependent variable. Further, the strong divisor split the results into 2 classes. Moreover, the node of BRT is split into 2 domains by utilizing recursive grouping of the points and ultimately separating those domains. This division of nodes stops till the nodes are uniform.

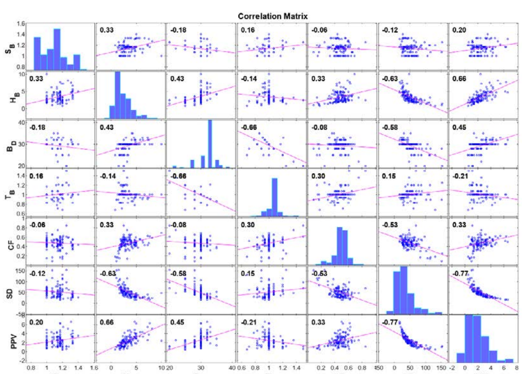


Figure 3. Correlation matrix among pairs of blast parameters.

Let's suppose the input matrix is $X= [X_1,\ldots\ldots X_p]$ and target vector is y, $\boldsymbol{\beta}$= coefficient of regression. The predicted value of the target is given by X$\boldsymbol{\beta}$ and the residual (i.e., error) is denoted by r = y-X$\boldsymbol{\beta}$. The algorithm for the BRT is as follows:

The LS-Boost algorithm is based on the least square fitting approach which starts from the null model with residual $\hat{r}^0 = y$. The LS-Boost algorithm finds a covariate j_l at the lth iteration. As a result, there are maximum decreases in the regression fit to the current residuals (Freund *et al*., 2017). If the univariate fit is represented by $X_{jl}\, u^{i}_{jl}$ for current residuals (for corresponding j_l) the updating of the residuals happens using step 3 (1st equation) of algorithm 1. Moreover, the j_l^{th} the regression coefficient of the model is updated using step 3 (2nd equation) of algorithm 1. To enhance the generalization of any machine learning model the hyper parameters should be optimized. During the cross-validation process, the hyper parameters of the BRT model have been fine-tuned for optimum performance. Table 4 shows the optimized parameters for the proposed BRT model for which the MSE loss is minimized. Once the model parameters have been fine-tuned, the model has been used to predict PPV from an unseen test set (Figure 2).

2.2.1 Regression Analysis

In the recent past, various studies have been conducted to predict the blast level induced by the vibration. Particle velocity is used extensively in blasting seismology since it provides quite reliable results. To evaluate the regression model two regression losses namely, MSE and Root Mean Square Error (RMSE) are utilized. The equations of regression loss are as follow:

Table 4. Hyperparameter details of optimized BRT model

hyperparameters	Values
Ensemble method	LS Boost
Learning rate	0.1
Number of base learners	30
Minimum leaf size	8

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right)^2 \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - x_i\right)^2} \tag{4}$$

where, n = total number of samples, y_i and x_i represent the actual and predicted PPV values for i^{th} sample.

3. Implementation Details

All experiments are completed in an HP workstation with 64 GB of RAM. Data standardization and model training has been carried out using MATLAB programming language (Mathworks, 2016). To validate the performance of each model a 5-fold cross-validation technique is used. In this work, two experiments have been carried out: (1) PPV prediction using BRT models and (2) PPV prediction using different machine learning models. The detailed analysis of the PPV prediction result is presented in successive subsections.

4. Results

4.1 Results of the Proposed BRT Approach

The proposed BRT-based model has been validated using 5-fold cross-validation and the mean *MSE* and *RMSE* are tabulated in Table 5. The fold-wise results show that the proposed model has a minimum loss for the 3rd fold. Overall, the model has achieved 0.30 and 0.53 of mean MSE and mean *RMSE* respectively. The low error rate indicates that the model is robust for the unseen validation set.

Table 7. Comparison of the proposed BRT model with standard USBM equation

	Calculated scaled distance factor, m kg$^{-1/2}$	**Measured *PPV*, mm/s**	**Predicted *PPV*, mm/s by USBM Equation 2**	**Predicted *PPV*, mm/s by BRT model**
1	22.70	5.72	4.66	5.19
2	47.34	1.40	1.98	2.14
3	47.18	1.40	1.99	2.14
4	56.65	2.85	1.61	1.59
5	53.68	1.10	1.71	1.89
6	38.10	1.91	2.55	2.53
7	37.53	1.94	2.60	2.53
8	47.34	1.40	1.98	2.14
9	47.18	1.40	1.98	2.14
10	57.63	2.86	1.57	1.59
11	53.68	1.10	1.71	1.89
12	35.20	2.30	2.79	2.97
13	47.20	1.40	1.98	2.14
14	56.60	2.80	1.60	1.59
15	56.67	2.50	1.60	1.59
16	53.60	1.20	1.71	1.89
17	35.60	2.40	2.76	2.97
18	32.66	2.41	3.05	3.32
19	28.09	4.83	3.64	3.64
20	33.63	4.19	2.95	3.32
MSE			0.64	**0.63**
RMSE			0.80	**0.79**

Further, to visualize the predicted PPV values using the proposed approach we have also plotted a graph between actual and predicted values. For this purpose, 100 samples are used and passed through the trained BRT model. Figure 4 shows that the predicted values are very close to actual PPV values which demonstrates that the proposed model precisely predicts the PPV values.

Table 5. Results of the proposed model for *PPV* prediction using 5-fold cross-validation

Fold	**MSE**	**RMSE**
1	0.24	0.49
2	0.43	0.65
3	0.15	0.39
4	0.25	0.50
5	0.41	0.64
mean	**0.30**	**0.53**

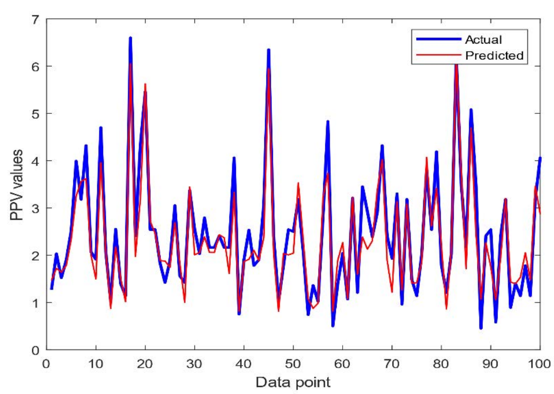


Figure 4. Plot of actual vs predicted PPV response values.

4.2 Results of PPV Prediction using Various Regression Model

In the previous section, we have used the BRT model for the PPV prediction. In this section, we have analyzed the effect of different regression models on PPV prediction. For this purpose, three different regression models have been trained namely, linear regression, stepwise linear regression and boosted regression trees using the same validation approach. First, the simplest models called linear regression and its variant have been analyzed which has achieved MSE of 0.64 and 0.64 respectively. Note that the linear regression model works better if there is a linear relationship between variables. In case the data is correlated in a very complex way, linear regression might not work in such scenarios. Hence, we have implemented and evaluated the decision tree-based model. The proposed BRT model has been evaluated and compared with other regression models. The comparison between the performances of these models are tabulated in Table 6. From Table 6,it can be noticed that the proposed BRT model has overall achieved significantly better results compared to other two regression models. This demonstrates the superiority of the proposed BRT model for PPV prediction.

Table 6. Comparative analysis of the proposed model with different models for *PPV* prediction

Regression model	**Validation**		**Testing**	
	MSE	**RMSE**	**MSE**	**RMSE**
Linear	0.64	0.80	1.24	1.11
Stepwise Linear	0.64	0.80	1.24	1.11
Boosted Trees	**0.30**	**0.53**	**0.63**	**0.79**

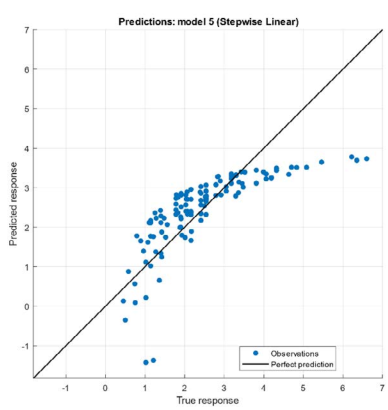
4.3 Comparison of the Proposed BRT Model with USBM Equation for PPV Prediction

In this section, the performance of the proposed BRT model is compared with the performance of the standard USBM equation. To validate the performance of the proposed BRT model, 2 experiments are carried out. In the first experiment, the PPV values have been predicted using Equation 2 (see section 2). To obtain the PPV equation first a clustering algorithm has been implemented to cluster the homogeneous parameters. Moreover, based on the clustered data a PPV prediction equation (i.e., Equation 2) is created. The measured and predicted PPV values are tabulated in Table 7. The clustering USBM approach has obtained MSE and RMSE of 0.64 and 0.80 respectively. Further, in the second experiment, we have utilized the same test dataset and predicted the PPV values using the proposed BRT model.

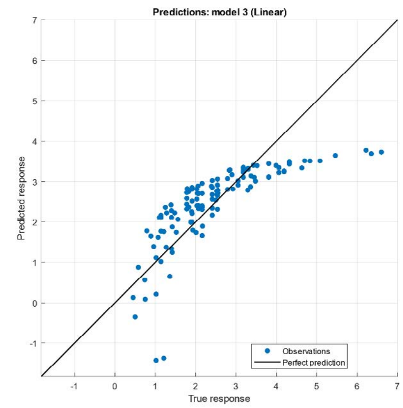
From Table 7, it can be observed that the proposed BRT model has achieved MSE and RMSE of 0.63 and 0.79 respectively which are better as compared to clustering and USBM-based approach for PPV prediction. It is worth mentioning that the USBM based approach first utilizes a clustering approach and then creates the predictive equation based on a homogeneous group. Therefore, the process of PPV prediction using k-mean clustering and regression-based approach is cumbersome. Whereas, in this work, the BRT-based model does not require any preprocessing and directly learns the mapping between input peta meters and the PPV values which is more efficient as compared to the clustering and regression-based approach. Moreover, the results of the proposed BRT model demonstrate that the machine learning-based model can learn better feature mapping as compared to the traditional approach. Therefore, in the future machine learning-based approaches could be utilized to predict the PPV value.

4.4 Regression Analysis Using Scatter Plot

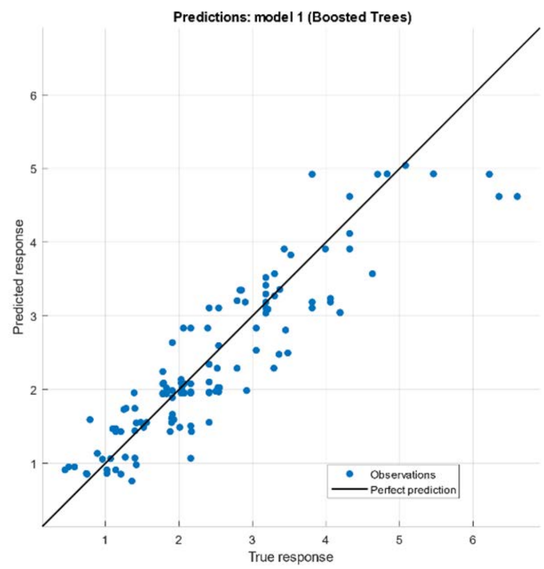
Further, the predicted vs. actual plot is also analyzed in order to check the various model performance as shown in Figure 5. The plot demonstrates that how well the model has made the predictions for different target PPV values. The actual and predicted values will overlap for a regression model that achieves perfect results. For such a model, all the data points should lie on a diagonal line. The error is calculated by taking the difference between the vertical distance from the diagonal lines to any data points. Note that for a best-performing regression model, the data points are scattered nearby the diagonal lines. Based on the plot, it is seen that the data points are not symmetrically scattered nearby the diagonal lines. Hence for these models, the PPV prediction error is high. Contrary to model (a) and model (b), the data points are scattered approximately symmetrically nearby the diagonal lines for the BRT model (i.e., model (c)). This demonstrates that the BRT model has lower error as compared to other regression models. From the scatter plot, it can also be observed that the BRT-based model is the best performing model for PPV prediction as compared to other regression models.



(a)



(b)



(c)

Figure 5. Plot of predicted vs actual for (a) Step-wise Linear regression, (b) Linear regression, and (c) Boosted Regression Trees.

5. Conclusion

In this work, a self-created blast dataset has been utilized which is acquired at the site of International Airport, Navi Mumbai. To model the complex correlation between the blast parameter and the PPV value, a decision tree-based approach called the least square boosted decision tree has been utilized. Before, training the BRT model the best suitable feature has been selected based on correlation matrix and then passed through the BRT model. The proposed model automatically maps the input blast feature (SD) with the target PPV values by aggregating the decision of various weak learners. To validate the generalization of the proposed model a 5-fold cross-validation approach is used. Moreover, the model has been also validated on a separate unseen test set to ensure robustness. Besides, the predictive performance of the proposed model has been compared with PPV prediction obtained using the standard USBM approach. The results demonstrate that the proposed model has outperformed the USBM based approach for PPV prediction. Further, to validate the robustness of the proposed BRT model a comparative analysis is carried out with other regression models. The results demonstrate that the proposed model significantly improves the prediction accuracy as compared to other regression models.

6. Acknowledgments

Authors are thankful to Mr. Manoj Zhade, M/s TIPL-JMM (JV) Package 1 and 2, Mr. Dheeraj Gupta, M/s Gayatri Package 3 and Mr. Charudatta Deshmukh, M/s GVK, Package 4 for giving permission to monitor the blasts at the Navi Mumbai International Airport site, Mumbai, India. The authors are also thankful to Mr. Vipin Gupta, Team Leader, Consulting Engineers Group, International Airport, Navi Mumbai for facilitating the field studies.

7. Funding

No funding was received to assist with the preparation of this manuscript.

8. Competing Interests

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

9. Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Rajesh Sonkar. The first draft of the manuscript was written by Rajesh Sonkar and all authors commented on the previous version of the manuscript. All authors read and approved the final manuscript.

10. References

1. Amiri, M., Amnieh, H.B., Hasanipanah, M., & Mohammad Khanli, L. (2016). A new combination of artificial neural network and K-nearest neighbors models to predict blast-induced ground vibration and air-overpressure. *Engineering with Computers*, 32(4): 631–644. https://doi.org/10.1007/s00366-016-0442-5
2. Ambraseys, N.R., & Hendron, A.J. (1968). *Dynamic behavior of rock masses, rock mechanics in engineering practice* (KG Stagg and OC Zienkiewicz, eds.)
3. Davies, B., Farmer, I.W., & Attewell, P.B. (1964). Ground vibration from shallow sub-surface blasts. Engineer, 217pp.
4. Duvall, W.I., & Petkof, B. (1959). *Spherical propagation of explosion-generated strain pulses in rock*. US Department of the Interior, Bureau of Mines
5. Ghosh, A., & Daemen, J.J.K. (1983). *A simple new blast vibration predictor (based on wave propagation laws)*. In: The 24th US symposium on rock mechanics (USRMS)
6. Gupta, R.N., Roy, P.P., & Singh, B. (1988). *On a blast induced blast vibration predictor for efficient blasting*. In: Proceedings of the 22nd International Conference of Safety in Mines, Beijing, China, pp. 1015–1021.
7. Rai, R., & Singh, T. (2004). A new predictor for ground vibration prediction and its comparison with other predictors. Indian Journal of Engineering and Materials Sciences, 11: 178–184.
8. Peng, K., Zeng, J., Armaghani, D.J., *et al.* (2021). A novel combination of gradient boosted tree and optimized ANN models for forecasting ground vibration due to quarry blasting. *Natural Resources Research*, 30: 4657–4671. https://doi.org/10.1007/s11053-021-09899-1
9. Zhang, X., Nguyen, H., Bui, X.N., *et al.* (2020). Novel soft computing model for predicting blast-induced ground vibration in open-pit mines based on particle swarm optimization and XGBoost. *Natural Resources Research*. 29: 711–721. https://doi.org/10.1007/s11053-019-09492-7
10. Ding, Z., Nguyen, H., Bui, X.N., *et al.* (2020). Computational intelligence model for estimating intensity of blast-induced ground vibration in a mine based on imperialist competitive and extreme gradient boosting algorithms. *Natural Resources Research*, 29: 751–769. https://doi.org/10.1007/s11053-019-09548-8
11. Qiu, Y., Zhou, J., Khandelwal, M., *et al.* (2021). Performance evaluation of hybrid WOA-XGBoost, GWO-XGBoost and BO-XGBoost models to predict blast-induced ground vibration. *Engineering with Computers*. https://doi.org/10.1007/s00366-021-01393-9
12. Yang H, Hasanipanah M, Tahir MM, & Bui DT. (2020). Intelligent prediction of blasting-induced ground vibration using ANFIS optimized by GA and PSO. *Natural Resources*

Research. 29: 739–750. https://doi.org/10.1007/s11053-019-09515-3

13. Fattahi, H., & Hasanipanah, M. (2021). Prediction of blast-induced ground vibration in a mine using relevance vector regression optimized by metaheuristic algorithms. *Natural Resources Research*, 30: 1849–1863. https://doi.org/10.1007/s11053-020-09764-7
14. Dehghani, H., Shokri, B.J., Mohammadzadeh, H., *et al*. (2021). Predicting and controlling the ground vibration using Gene Expression Programming (GEP) and Teaching–Learning-Based Optimization (TLBO) algorithms. *Environmental Earth Sciences*, 80. https://doi.org/10.1007/s12665-021-10052-7
15. Chen, W., Hasanipanah, M., Rad, H.N., *et al*. (2021). A new design of evolutionary hybrid optimization of SVR model in predicting the blast-induced ground vibration. *Engineering with Computers*, 37: 1455–1471. https://doi.org/10.1007/s00366-019-00895-x
16. Shang, Y., Nguyen, H., Bui, X.N., *et al*. (2020). A novel artificial intelligence approach to predict blast-induced ground vibration in open-pit mines based on the firefly algorithm and artificial neural network. *Natural Resources Research*, 29: 723–737. https://doi.org/10.1007/s11053-019-09503-7
17. Ding, X., Hasanipanah, M., Rad, H.N., Zhou, W. (2021). Predicting the blast-induced vibration velocity using a bagged support vector regression optimized with firefly algorithm. *Engineering with Computers,* 37: 2273–2284. https://doi.org/10.1007/s00366-020-00937-9
18. Zhou, J., Qiu, Y., Khandelwal, M., *et al* (2021). Developing a hybrid model of Jaya algorithm-based extreme gradient boosting machine to estimate blast-induced ground vibrations. *International Journal of Rock Mechanics and Mining Sciences*. 145pp. https://doi.org/10.1016/j.ijrmms.2021.104856
19. Zhang, H., Zhou, J., Armaghani, D.J., Tahir, M.M., Pham, B.T., & Huynh, V.V. (2020). A combination of feature selection and random forest techniques to solve a problem related to blast-induced ground vibration. *Applied Sciences*, 10(3): 869. https://doi.org/10.3390/app10030869
20. Zeng, J., Roussis, P.C., Mohammed, A.S., *et al*. (2021). Prediction of peak particle velocity caused by blasting through the combinations of boosted-chaid and SVM models with various kernels. *Applied Sciences*. 11(8): 3705. https://doi.org/10.3390/app11083705
21. Nguyen, H., Choi, Y., Bui, X. N., & Nguyen-Thoi, T. (2020). Predicting blast-induced ground vibration in open-pit mines using vibration sensors and support vector regression-based optimization algorithms. *Sensors*, 20(1): 132. https://doi.org/10.3390/s20010132. PMid:31878226. PMCid:PMC6983179
22. Sonkar, R., Dhekne, P.Y., Londhe, N.D. (2021). Improvement in the prediction of peak particle velocity of blast-induced ground vibrations using K-means clustering. *Arabian Journal of Geosciences*, 14pp. https://doi.org/10.1007/s12517-021-08620-z
23. Freund, R.M., Grigas, P., Mazumder, R. (2017). A new perspective on boosting in linear regression via subgradient optimization and relatives. *Annals of Statistics*.
24. Mathworks. (2016). MATLAB - Mathworks - MATLAB & Simulink. Available from: www.mathworks.com