Soils and Rocks
An International Journal of Geotechnical and Geoenvironmental Engineering
www.soilsandrocks.com
ISSN 1980-9743
ISSN-e 2675-5475

Determination of the size of rock fragments using RVM, GPR, and MPMR

Pradeep Thangavel[1#], Pijush Samui[1]

Article

Keywords
Blasting
GPR
MPMR
RVM
Rock fragmentation

Abstract

For predicting the size of rock fragments during drilling and blasting operations, this article uses GPR, RVM, and MPMR. The current analysis makes use of a blast data set generated in a prior investigation. In this study, a portion of the blast data was utilized to train a model to determine the mean particle size arising from blast fragmentation for each of the similarity groups generated. The particle size was modeled as a function of seven different variables. The dataset contains information about the bench height and drilled burden ratio (H / B), spacing to burden ratio (S / B), burden to hole diameter ratio (B / D), stemming to burden ratio (T / B), powder factor (P_f), modulus of elasticity (E), and in-situ block size (X_B) are the input and output is X_{50}. By comparing forecasts with actual mean particle size values and predictions based on one of the most widely used fragmentation estimation techniques in the blasted literature, the capacity of the generated models may be established. The statistical parameters, actual vs predicted curve, Taylor diagram, error bar, and developed discrepancy ratio are used to analysis the performance of models. A comparative study has been carried out between the developed RVM, GPR, and MPMR. The results show the developed models have the capability for prediction of X_{50}. From these comparisons, the MPMR has the highest value with a high degree of precision and robustness in the size of rock fragments X_{50}.

1. Introduction

Rock mass is a heterogeneous material, and in blasting and drilling, the heterogeneity of the rock generates size distribution of fragmented rocks. The total economics of mine workings are heavily reliant on the estimation of blasted rock mass fragmentation. The cost of loading, transport, crushing, and milling operations can all be reduced dramatically by using blasting as a major fragmentation method. Blast fragmentation is primarily determined by the blast design as well as the qualities of the rock mass (Jug et al., 2017; Mohamed et al., 2019). Direct and indirect approaches are used to quantify the size distribution of shattered rock following blasting. In the direct procedure, the only methodology is fragment sieving analysis. Despite being the most accurate approach among others, it is not practicable due to its high cost and time requirements. As a result, observational, empirical, and digital approaches have been created as indirect methods. Researchers can use a range of existing tools and models to predict and process blasting findings, one of which is machine learning techniques, which is possibly the most extensively used way to estimate fragmentation after blasting. There are different empirical models available for the determination of the size distribution of rock fragments in the literature (Kuznetsov, 1973; Aler et al., 1996; Ozkahraman, 2006; Jethro et al., 2016). However, the available methods are not so reliable (Shi et al., 2012).

Singh et al. (2019) use dataset from forty open space bench explosion in the four open Indian mines to track blast-induced rock disintegration as a function of explosion parameters such as spacing, powder feature, hole size, weight, stemming depth, and hole bench height. Tao et al. (2020) investigated blast-induced rock fragmentation using a combination of analytical modelling, finite element simulation, and image recognition. They used sequential alterations in the model geometry to investigate the major impact of rock fragility and effective dimensions on fragment size distribution, demonstrating that the effect of fracture toughness on fragmentation is included in the effect of material size. The deep CNN was utilised by Yang et al. (2021) to automatically categorise rock fragment images taken by a timed capture camera. Bamford et al. (2021) discuss the implications of using deep learning models for the fragmentation of rock assessment. Using an end-to-end deep learning technique, convolution neural network architecture was trained to predict mean sizes of blasted rock fragments straight from a 2D image. His research examines the DNN model's accuracy and effectiveness as a tool for automated and rapid rock fragmentation analysis. Researchers have tried different numerical and Artificial Intelligence (AI)

#Corresponding author. E-mail address: tpradeep1125@gmail.com
[1]National Institute of Technology Patna, Department of Civil Engineering, Patna, Bihar, India.

Submitted on July 23, 2022; Final Acceptance on October 11, 2022; Discussion open until February 28, 2023.
https://doi.org/10.28927/SR.2022.008122

techniques for the determination of the size distribution of rock fragments (Zhang & Goh, 2013; Zhang et al., 2019, 2020a, b, c; Kumar et al., 2021; Li et al., 2022).

Relevance vector machine (RVM) is a type of soft computing method that combines the concepts of Markov property, automated relevance determination (ARD), Bayesian principle, and maximum probability into a probabilistic Bayesian learning framework (Kong et al., 2019). The functional forms of RVM and SVM are identical. The much more significant benefit of RVM over SVM is its ability to make probabilistic predictions. RVM's high sparseness also allows it to reduce the number of kernel functions utilized in computing, making it particularly suitable for online monitoring. SVM kernel functions must satisfy Mercer's criterion, which asserts that the related kernel matrix of a symmetric function is semi-positive. SVM kernel functions must satisfy Mercer's criterion, which asserts that the related kernel matrix of a symmetric function is semi-positive. RVM, on the other hand, has the advantage of being able to use any kernel function without having to satisfy Mercer's criterion (Samui, 2012; Li et al., 2017; Biswas et al., 2019; Kardani et al., 2021; Pradeep et al., 2021). In the domain of machine learning, GPR, a nonparametric Bayesian method fo regression, is causing a stir. GPR has various advantages, including the capacity to work with limited datasets and provide uncertainty measures on predictions (Chalupka et al., 2013; Caywood et al., 2017; Baiz et al., 2020). The minimax probability machine classification technique underpins MPMR, which is a regression method. There were no assumptions made in this model about the numerical distribution of the data. It is based on the probabilistic framework. It is been used brilliantly in a variety of engineering sectors (Strohmann & Grudic, 2003; Samui & Kim, 2017; Kumar et al., 2020).

For the purpose of determining the mean particle size X_{50} resulting from rock blast fragmentation, this article uses RVM, GPR, and MPMR. The dataset contains information about the ratio of bench height to drilled burden $/B$, ratio of stemming to burden (T / B), ratio of spacing to burden (S / B), ratio of burden to hole diameter (B / D), modulus of elasticity (E), powder factor (P_f), in-situ block size (X_B) are inputs and outputs are (X_{50}) in the database. The models are trained by using data of 70% and tested by 30% data. These models are performing under MATLAB software. The predicted mean size of rock (X_{50}) from models results is compared with actual data for analysis of the capacity of the model. For the comparative purpose statistical parameters, actual vs predicted curve, Taylor diagram, error bar, and DDR criteria are used in this article.

2. Details of data

The blast database created by Hudaverdi et al. (2011) (Kulatilake et al., 2010; Hudaverdi et al., 2011; Shi et al., 2012) is covered in this section. To create the blast database, data from previous blasts in various regions of the world. There are quarries in Istanbul, as well as mines in Spain called Enusa and Reocin. The Murgul Copper Mine in northeastern Turkey, with Mrica Quarry in Indonesia, Soma Basin in western Turkey, the Dongri–Buzurg mine in Central India, and the Akdaglar and Ozmert Quarries in northern Istanbul (Kulatilake et al., 2010). The ratio of bench height to drilled burden $\left(H / B\right)$, ratio of stemming to burden (T / B), ratio of spacing to burden (S / B), ratio of burden to hole diameter (B / D), modulus of elasticity (E), powder factor (P_f), in-situ block size (X_B) are inputs and outputs are X_{50} in the database. Figure 1 shows the scatter plot matrix for the original data set. Table 1 shows the statistical analysis of the input variables used to develop the models to predict fragmentation. The term "normalization" in statistics refers to the scaling down of a data collection so that the normalized data falls between 0 and 1. Such normalization approaches make it possible to compare matching normalized values from two or more separate data sets in a way that eliminates the impact of scale differences. To put it another way, a data set with large values may readily be compared to a data set with lower values. The data is split into two groups. A training dataset is necessary for the model to be trained. 70% of data sets are considered for training in this study. A testing dataset is necessary in order to estimate model performance. The remaining 30% is used as the testing data set in this study. The normalization equation is shown in Equation 1.

$$x_{Normalized} = \frac{\left(x_{Actual} - x_{mini}\right)}{\left(x_{maxi} - x_{mini}\right)} \tag{1}$$

3. Developed model details

3.1 Relevance Vector Machine (RVM)

RVM is introduced by Tipping (2000). It is constructed based on the Bayesian concept. In RVM, the Equation 2 represents the input and output relation.

$$t = \Phi w + \varepsilon \tag{2}$$

This article uses $H/B, S/B, B/D, T/B, P_f, E, and\ X_B$ as inputs of RVM. The output is X_{50}.

Table 1. Statistical analysis of the input parameters used to create models to predict fragmentation.

	S / B	H / B	B / D	T / B	P_f	X_B	E
Minimum	1	1.3	18	0.5	0.2	0.02	9.5
Maximum	1.7	6.8	39.5	4.7	1.3	2.3	60
Mean	1.2	3.3	27.4	1.3	0.5	1.1	29.5
Std deviation	0.1	1.6	4.8	0.7	0.2	0.5	17.9

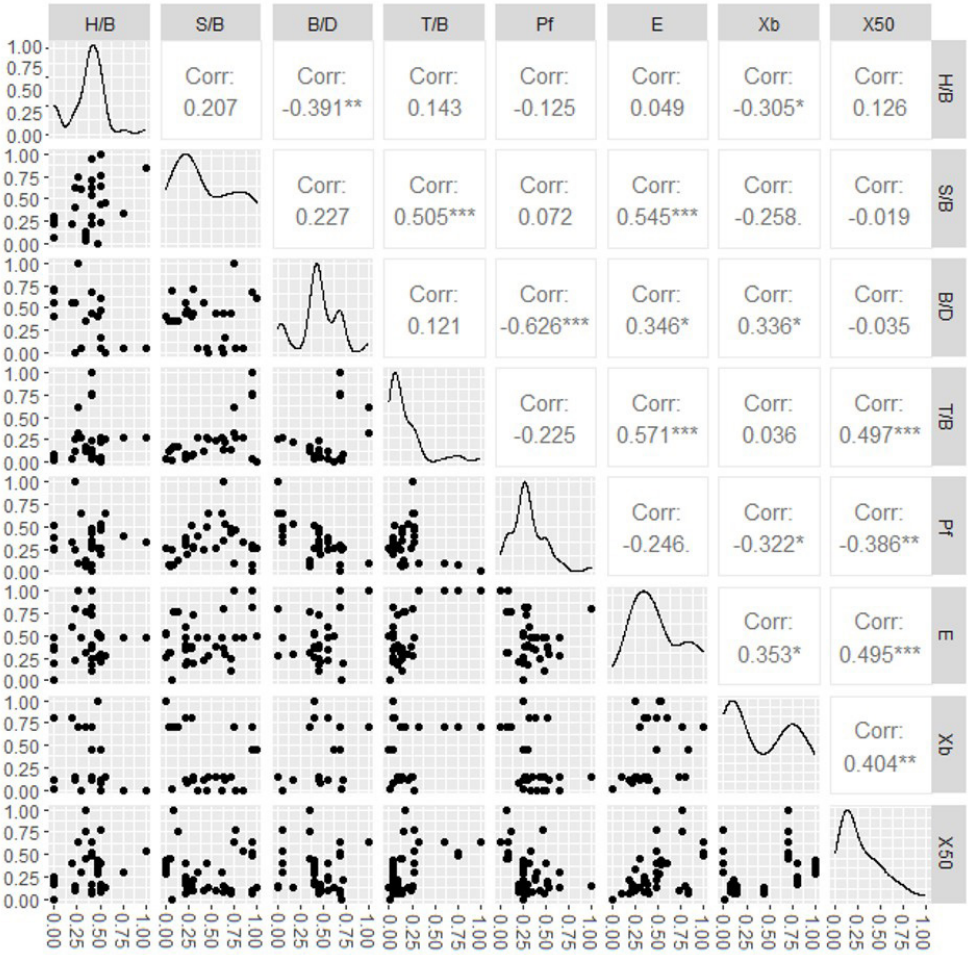

Figure 1. Correlation and scatter plot of data.

So, $x=\left[H/B, S/B, B/D, T/B, P_f, E, X_B\right]$ and $y=\left[X_{50}\right]$

Where $\Phi=\left[\phi(x_1),\ldots,\phi(x_N)\right]$ and $\phi(x_n)=\left[K(x_n,x_1),\right.$ $\left.K(x_n,x_2),\ldots,K(x_n,x_M)\right]^T$, $K(x_n,x_1)$ is kernel function. ε follows the Gaussian distribution having mean zero and σ^2 variance. The likelihood of the complete dataset is given below:

$$p\left(\frac{t}{w},\sigma^2\right)=\left(2\pi\sigma^2\right)^{-\frac{N}{2}}\exp\left\{-\frac{1}{2\sigma^2 t-\Phi w^2}\right\} \quad (3)$$

The value of t can be determined by maximizing the above Equation 2. The maximization of Equation 2 can cause overfitting. Over the weights, automated relevance detection (ARD) prior is set to prevent overfitting.

$$p(w|\alpha)=\prod_{i=0}^{N}N\left(w_i|0,\alpha_i^{-1}\right) \quad (4)$$

Where the hyperparameter vector that specifies how far each weight can depart from zero. According to Bayes' rule, the combination of likelihood and prior is given by:

$$p\left(\frac{w}{t},\alpha,\sigma^2\right)=\frac{p\left(\frac{t}{w},\sigma^2\right)p\left(\frac{w}{\alpha}\right)}{p\left(\frac{t}{\alpha},\sigma^2\right)} \quad (5)$$

The posterior covariance (Σ) and mean (μ) are given below:

$$\Sigma=\left(A+\alpha^{-2}\Phi^T\Phi\right)^{-1} \quad (6)$$

$$\mu=\sigma^{-2}\sum\Phi^T t \quad (7)$$

Where $A=diag(\alpha)$

The details of RVM is given by Tipping (2000). Radial basis function $(x, x_i) = \exp\left\{-\frac{(x_i - x)(x_i - x)^T}{2\sigma^2}\right\}$, where the kernel function is the width of the radial basis function σ.

3.2 Gaussian Process Regression (GPR)

For output (y) prediction, the GPR model use the following model

$$y_i = f(x_i) + \varepsilon$$

In this article, GPR uses the same inputs and output as used by the RVM.

$$\text{So, } x = \left[\frac{H}{B}, \frac{S}{B}, \frac{B}{D}, \frac{T}{B}, P_f, E, X_B\right] \text{ and } y = \left[X_{50}\right]$$

The spreading of output (y_{N+1}) for a novel input vector (x_{N+1}) is represented by

$$\begin{pmatrix} y \\ y_{N+1} \end{pmatrix} \sim N(0, K_{N+1}) \tag{8}$$

The expression of K_{N+1} is given below:

$$K_{N+1} = \begin{bmatrix} [K] & \left[K(x_{N+1})\right] \\ \left[K(x_{N+1})^T\right] & \left[k(x_{N+1})\right] \end{bmatrix} \tag{9}$$

Where $K(x_{N+1})$ is the $N \times 1$ vector which covariances lies between training and the testing input, and $K(x_{N+1})$represent the auto covariance of the test input.

The distribution of y_{N+1} is Gaussian. The mean and variance of y_{N+1} are given below:

$$\mu = K(x_{N+1})^T K^{-1} y \tag{10}$$

$$\sigma^2 = k(x_{N+1}) - K(x_{N+1})^T K^{-1} K(x_{N+1}) \tag{11}$$

The GPR model uses the same training dataset, testing dataset, kernel function, and normalization technique as used by the RVM model. The program of GPR has been developed by MATLAB.

3.3 Minimax Probability Machine Regression (MPMR)

MPMR is developed by Lanckriet et al. (2003). In MPMR, the relation between input(x) and output(y) is given by the following equation.

$$y = \sum_{i=1}^{N} \beta_i K(x_i, x) + b \tag{12}$$

Where $K(x_i, x)$ is kernel function, β_i and b are output from the MPMR algorithm.

In this article, MPMR uses the same inputs and output as used by the RVM and GPR.

$$\text{So, } x = \left[\frac{H}{B}, \frac{S}{B}, \frac{B}{D}, \frac{T}{B}, P_f, E, X_B\right] \text{ and } y = \left[X_{50}\right]$$

MPMR is developed by constructing dichotomy classifier [24]. All of the regression data $+\varepsilon$ is shifted into one data set along the output. The second dataset is created by relocating all of the regression data $-\varepsilon$ down the output line. The regression surface is the categorization border between these two classes.

Both the RVM and the MPMR models employ the identical training dataset, testing dataset, kernel function, and normalization technique. The program of MPMR has been constructed using MATLAB.

3.4 Evaluation of models

The model's accuracy was explained using a variety of statistical methodologies. The parameters are determination coefficient (R^2), Nash-Sutcliffe efficiency (*NS*), Root mean square error (*RMSE*), Weighted mean absolute percentage error (*WMAPE*), Variance Account Factor (*VAF*), Performance index (*PI*) (Wong, 1985), Willmott's Index of agreement (*WI*) (Willmott, 1984), Mean absolute error (*MAE*) (Chai & Draxler, 2014), Mean Bias Error (*MBE*), Expanded uncertainty (U_{95}) (Behar et al., 2015), and t-statistic (*t stat*) (Stone, 1994).

$$R^2 = \frac{\sum_{i=1}^{N}(d_i - d_{mean})^2 - \sum_{i=1}^{N}(d_i - y_i)^2}{\sum_{i=1}^{N}(d_i - d_{mean})^2} \tag{13}$$

$$WMAPE = \frac{\sum_{i=1}^{n}\left|\frac{d_i - y_i}{d_i}\right| \times d_i}{\sum_{i=1}^{n} d_i} \tag{14}$$

$$NS = 1 - \frac{\sum_{i=1}^{n}(d_i - y_i)^2}{\sum_{i=1}^{n}(d_i - d_{mean})^2} \tag{15}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i - y_i)^2} \tag{16}$$

$$VAF = \left(1 - \frac{var(d_i - y_i)}{var(d_i)}\right) \times 100 \tag{17}$$

$$PI = adj.R^2 + (0.01 \times VAF) - RMSE \tag{18}$$

$$WI = 1 - \left[\frac{\sum_{i=1}^{N} (d_i - y_i)^2}{\sum_{i=1}^{N} \left\{ |y_i - d_{mean}| + |d_i - d_{mean}| \right\}^2} \right] \quad (19)$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |(y_i - d_i)| \quad (20)$$

$$MBE = \frac{1}{N} \sum_{i=1}^{N} (y_i - d_i) \quad (21)$$

$$U_{95} = 1.96 \left(SD^2 + RMSE^2 \right)^{\frac{1}{2}} \quad (22)$$

$$t\,stat = \sqrt{\frac{(N-1)MBE^2}{RMSE^2 - MBE^2}} \quad (23)$$

4. Results and discussion

4.1 Parameter and evaluation of models

The models are trained by using training data set under adjusting model parameters by trial and error method. The achievement of RVM model depends on the right selection of value of σ. The design value of σ has been determined using the trial and error method. The developed RVM model gives greatest performance shown in Table 1. The value of w has been represented in Figure 2. In Figure 2 it is clearly shown that 22 training datasets have non-zero w. So, number of relevance vector is 22. Figures 3-5 demonstrate the performance of training dataset and testing dataset. This article uses Determination coefficient (R^2) to asses the performance of the developed RVM, GPR and MPMR models. For a good model, the value of R^2 should be close to one. Figures 3-5 show that the value of R for both the training and testing datasets is close to one. For the prediction of X_{50}, the constructed RVM yields the following equation (Equation 24).

$$\varepsilon x = \sum_{i=1}^{90} w_i \exp\left\{ -\frac{(x_i - x)(x_i - x)^T}{0.08} \right\} \quad (24)$$

For GPR model the design values of error ε and kernel function σ have been considered by the approach of trial and error method. Therefore, the developed GPR proves his ability for prediction of X_{50}. Similarly for MPMR, Figure 5 illustrates the performance of training and testing for the MPMR model. It is also clear from figure that the value of R^2 is close to one for training as well as testing datasets. The model tuning parameters are shows in the Table 2.

Table 2. Model tuning parameter.

	RVM	GPR	MPMR
ε	-	0.001	0.004
σ	0.2	0.5	0.7

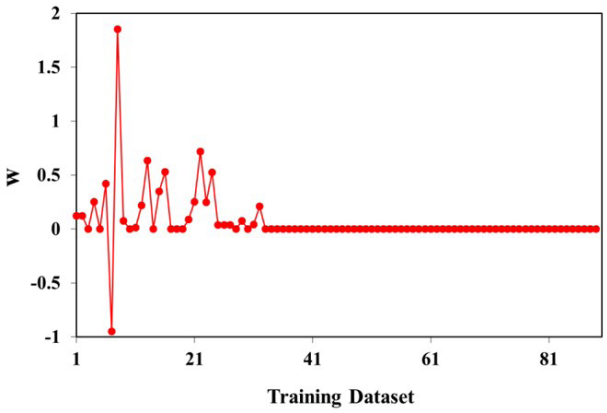

Figure 2. Weights Vs number of data for RVM.

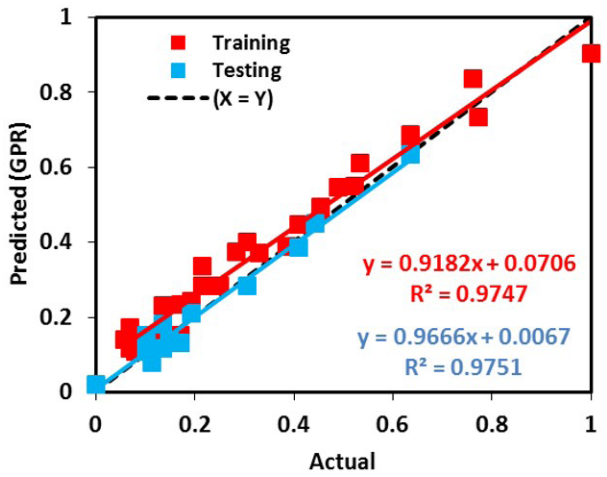

Figure 3. Actual Vs predicted plot for GPR.

For the training and testing datasets, a scatter plot is created showing actual against predicted values. Scatter plot measure the prediction capacity of the developed models using the target value and actual value. The point on the line $(y = x)$denotes the predictive model's perfect prediction value. Similarly, a point close to the line denotes a model forecast that is accurate. Figures 3-5, depicts a graphical depiction of the actual and anticipated value performance for the training datasets and testing datasets. According to this graph, all three models are the best, especially RVM achieving $R^2 = 0.99$ in the training and testing stages of the model.

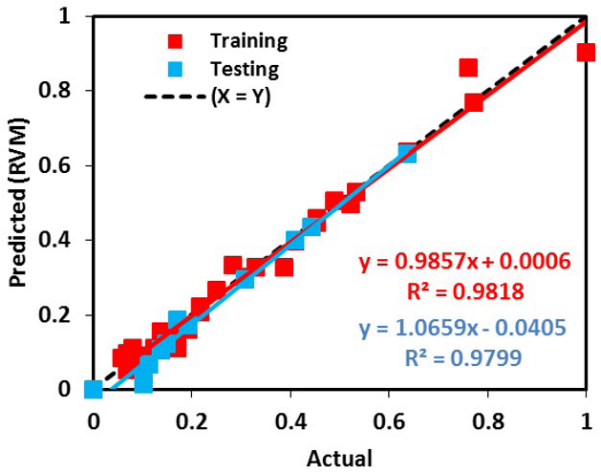

Figure 4. Actual Vs predicted plot for RVM.

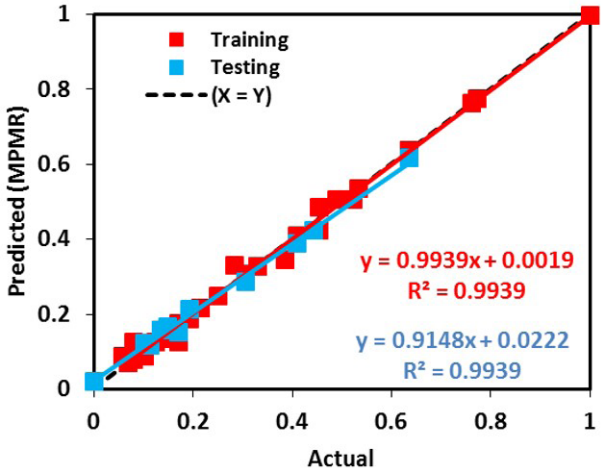

Figure 5. Actual Vs predicted plot for MPMR.

4.2 Statistical parameter

Table 3 shows the statistical parameters of the proposed models. In the rock sample, all of the models achieve above a 95% level of correlation. In every case of a rock sample, the models outperformed the humans. To account for the higher efficiency of the models, RMSE, MAE, and MBE should be near to 0, R^2 should be close to 1, and VAF should be close to 100. As a result, the improved fit of all of the models is confirmed. The degree of error in model predictions is measured by WI, which runs from 0 to 1. Its values near 1 are the most advantageous for good models. The model with the greater value is superior. All models have the best value based on the limits and range of parameters in these tables. MPMR values are best then RVM better than GPR. More detail of parameters are referred in Kardani et al. (2021).

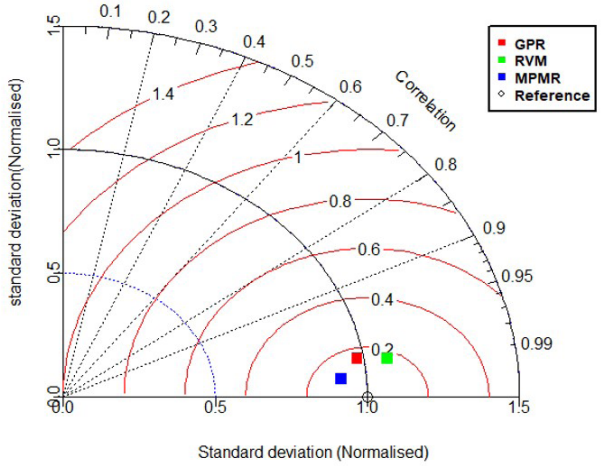

Figure 6. Taylor diagram for the dataset (training).

4.3 Taylor diagram

The mathematical diagrams are used to show which of various model's accuracy in single 2D, Taylor created this figure in 1994 (Taylor, 2001) to make comparing different models easier. The Pearson correlation coefficient, the root-mean-square error (RMSE), and the standard deviation are used to assess the degree of correspondence between the modeled and observed behavior in terms of three statistics. This diagram (Figures 6-7) is presented in this article using the GPR, MPMR, and RVM models. All models are performing well in the training and testing stages. When compared to GPR and MPMR, RVM is the best.

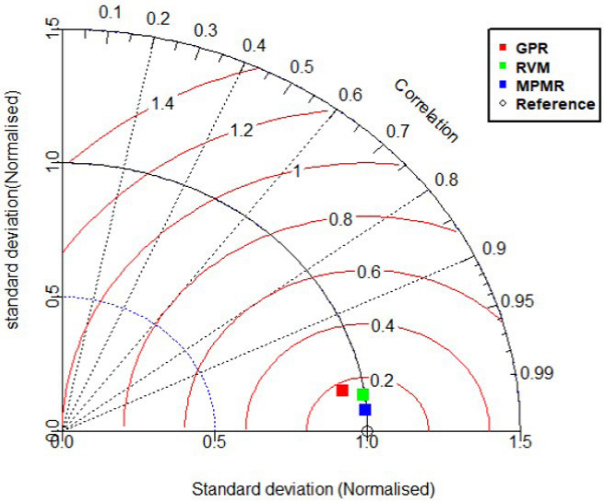

Figure 7. Taylor diagram for the dataset (testing).

4.4 Error bars

This section studies the error of predicted data in each model for the purposes of comparison of the model. The error bar diagrams are used to display the error level in models. Maximum, mean, and minimum values are also

shown in Figures 8-10. In the GPR model has been described most of the values are negative and the range of -0.1 to 0.1. In RVM error values are in the range of -0.1 to 0.1 and the maximum, mean, and minimum values are also better than GPR. MPMR error values are comparatively best because the range of -0.06 to 0.06 and other maximum, mean and minimum values also. Therefore, the solution of this study MPMR has been the robust model.

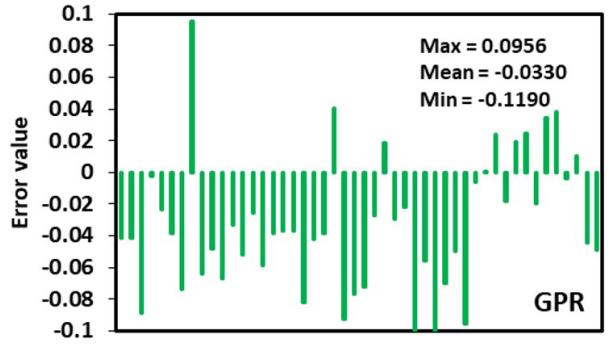

Figure 8. Error bars for GPR model.

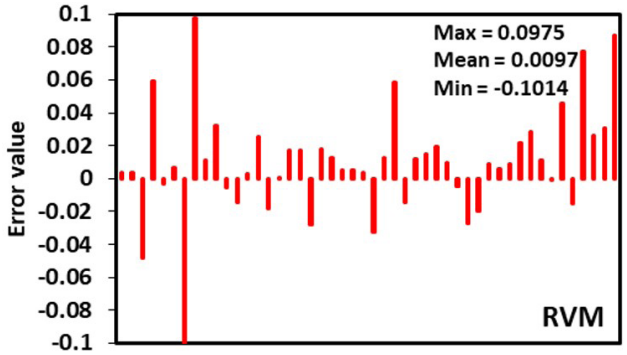

Figure 9. Error bars for RVM model.

4.5 DDR criterion

Developed discrepancy ratio also used in this paper and it was proposed by Noori et al. (2010). Evaluation have been dependent on MSE and R^2. DDR values were obtained by Equation 25. The standard error indexes a mean error value but does not provide information about the error distribution. As a result, the model's efficiency during the development phase must be assessed through the use of the dataset. Figures 11-12 shows the DDR results obtained for

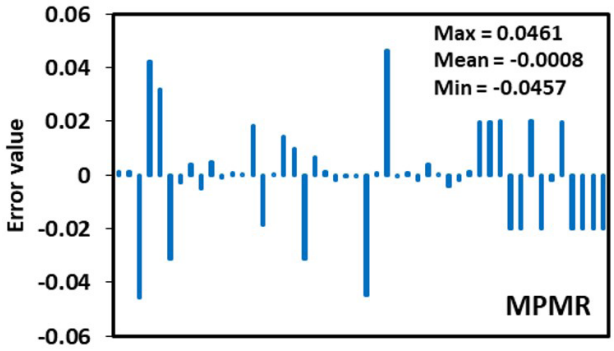

Figure 10. Error bars for MPMR model.

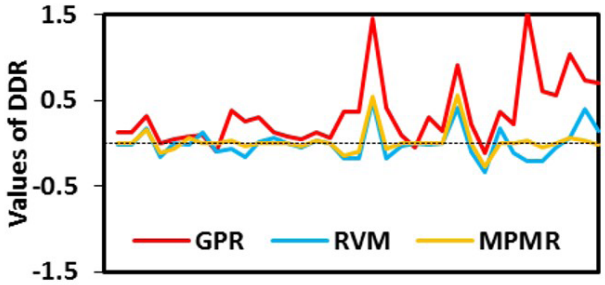

Figure 11. DDR values for train data.

Table 3. Evaluated statistical parameter values.

Parameters	GPR		RVM		MPMR		Ideal value
	Train	Test	Train	Test	Train	Test	
R^2	0.9747	0.9751	0.9818	0.9799	0.9939	0.9939	1
WMAPE	0.1780	0.0941	0.0714	0.1265	0.0356	0.0765	0
NS	0.9350	0.9720	0.9816	0.9409	0.9939	0.9863	1
RMSE	0.0609	0.0274	0.0324	0.0398	0.0186	0.0192	0
VAF	97.1459	97.2211	98.1804	97.0058	99.3929	98.6462	100
PI	1.8787	1.8684	1.9265	1.8837	1.9677	1.9477	2
MAPE	35.5901	15.2365	12.2803	24.2228	6.9645	10.3958	0
WI	0.9826	0.9929	0.9953	0.9867	0.9985	0.9963	1
MAE	0.0544	0.0227	0.0218	0.0305	0.0109	0.0185	0
MBE	0.0456	-0.0025	-0.0038	-0.0280	0.0000	0.0018	0
U95	0.2335	0.1717	0.2433	0.1934	0.2423	0.1580	0
t-sta	6.5916	0.3017	0.6848	3.2728	0.0023	0.3185	Smaller value

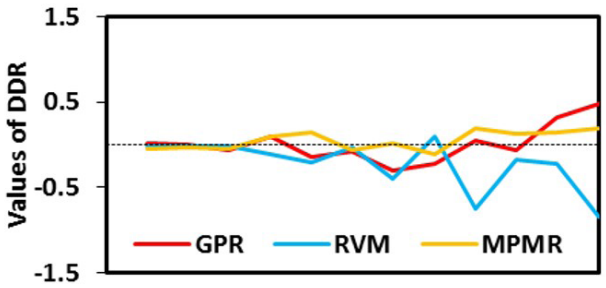

Figure 12. DDR values for test data.

all three models in the both (train and test) stages. The DDR figures shows the less deviation models (i.e.) MPMR curve nearly equal to zero line. When compared to the efficiency of other models based on DDR index, it reveals that the MPMR model is best.

$$DDR = \left(\frac{Estimated\ value}{Actual\ value} \right) - 1 \tag{25}$$

5. Conclusion

The machine learning methods were developed for predicting the rock fragmentation X_{50} due to drilling and blasting operations by using GPR, SVM, and MPMR models. A predicted all models were developed using factors such as blast design parameters, explosive parameters, modulus of elasticity, and in-situ block size. The developed models were trained using 90 training data and performance was tested by 13 testing data. The models were successfully demonstrated for predicting the rock fragmentation X_{50}. The performances were evaluated by using statistical parameter, Actual vs, predicted curve, Taylor diagram, Error bar diagram , and developed discrepancy ratio. All statistical parameter values of models were attained result within the ideal limit. Actual vs predicted show the accuracy of predicted value. Taylor diagram deals with three parameters like correlation, standard deviation, and RMSE in single 2D graph. In this diagram, the MPMR model behaved admirably. Error value also used to compared the models, MPMR have been reached very lesser range of error (-0.06 to 0.06). DDR values also showed nearer to the 0, in the case of MPMR. From this evaluation study, all models were performed well especially for MPMR has been performed best than other two models. The MPMR has been attained best accuracy of predicted value, $R^2 = 0.99$ in training and testing. Hence, the MPMR has been chosen as a robust model for predicting of rock fragmentation X_{50}. Expanding the blast databases that will be used to build the fragmentation prediction models outlined in this paper, as well as analyzing additional rock attributes of the rock mass that will be exposed to blasting if such information is available, could be part of future work.

Declaration of interest

The authors confirm that this research work content has no conflict of interest.

Authors' contributions

Pradeep Thangavel: conceptualization, developing models, visualization, and writing – original draft. Pijush Samui: conceptualization, data curation, methodology, and supervision.

List of symbols

x_{Actual}	actual data set
x_{Mini}	dataset's minimum value
x_{max}	dataset's maximum value
w	weight vector
ε	noise vector
N	number of datasets
α	hyperparameter vector
d_i	observed i^{th} value
y_i	predicted i^{th}
d_{mean}	average of the observed value
SD	standard deviation
K_{N+1}	covariance matrix

References

Aler, J., Mouza, J., & Arnould, M. (1996). Measurement of the fragmentation efficiency of rock mass blasting and its mining applications. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, *33*, 125-139. http://dx.doi.org/10.1016/0148-9062(95)00054-2.

Baiz, A.A., Ahmadi, H., Shariatmadari, F., & Torshizi, M.A.K. (2020). A Gaussian process regression model to predict energy contents of corn for poultry. *Poultry Science*, *99*, 5838-5843. http://dx.doi.org/10.1016/j.psj.2020.07.044.

Bamford, T., Esmaeili, K., & Schoellig, A.P. (2021). A deep learning approach for rock fragmentation analysis. *International Journal of Rock Mechanics and Mining Sciences*, *145*, 104839. http://dx.doi.org/10.1016/j.ijrmms.2021.104839.

Behar, O., Khellaf, A., & Mohammedi, K. (2015). Comparison of solar radiation models and their validation under Algerian climate - the case of direct irradiance. *Energy Conversion and Management*, *98*, 236-251. http://dx.doi.org/10.1016/j.enconman.2015.03.067.

Biswas, R., Samui, P., & Rai, B. (2019). Determination of compressive strength using relevance vector machine and emotional neural network. *Asian Journal of Civil Engineering*, *20*, 1109-1118. http://dx.doi.org/10.1007/s42107-019-00171-9.

Caywood, M.S., Roberts, D.M., Colombe, J.B., Greenwald, H.S., & Weiland, M.Z. (2017). Gaussian process regression for predictive but interpretable machine learning models: an example of predicting mental workload across tasks. *Frontiers in Human Neuroscience*, *10*, 1-19. http://dx.doi.org/10.3389/fnhum.2016.00647.

Chai, T., & Draxler, R.R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*, 1247-1250. http://dx.doi.org/10.5194/gmd-7-1247-2014.

Chalupka, K., Williams, C.K.I., & Murray, I. (2013). A framework for evaluating approximation methods for Gaussian process regression. *Journal of Machine Learning Research*, *14*, 333-350.

Hudaverdi, T., Kulatilake, P.H.S.W., & Kuzu, C. (2011). Prediction of blast fragmentation using multivariate analysis procedures. *International Journal for Numerical and Analytical Methods in Geomechanics*, *35*(12), 1318-1333.

Jethro, M.A., Ajayi, O.D., & Elijah, O.P. (2016). Rock fragmentation prediction using Kuz-Ram Model. *Journal of Environment and Earth Science*, *6*, 110-115.

Jug, J., Strelec, S., Gazdek, M., & Kavur, B. (2017). Fragment size distribution of blasted rock mass. *IOP Conference Series. Earth and Environmental Science*, *95*, 042013. http://dx.doi.org/10.1088/1755-1315/95/4/042013.

Kardani, N., Pradeep, T., Samui, P., Kim, D., & Zhou, A. (2021). Smart phase behavior modeling of asphaltene precipitation using advanced computational frameworks: ENN, GMDH, and MPMR. *Petroleum Science and Technology*, *39*(19-20), 804-825. http://dx.doi.org/10.1080/10916466.2021.1974882.

Kong, D., Chen, Y., Li, N., Duan, C., Lu, L., & Chen, D. (2019). Relevance vector machine for tool wear prediction. *Mechanical Systems and Signal Processing*, *127*, 573-594. http://dx.doi.org/10.1016/j.ymssp.2019.03.023.

Kulatilake, P.H.S.W., Qiong, W., Hudaverdi, T., & Kuzu, C. (2010). Mean particle size prediction in rock blast fragmentation using neural networks. *Engineering Geology*, *114*, 298-311. http://dx.doi.org/10.1016/j.enggeo.2010.05.008.

Kumar, M., Samui, P., Kumar, D., & Zhang, W. (2021). Reliability analysis of settlement of pile group. *Innovative Infrastructure Solutions*, *6*, 24. http://dx.doi.org/10.1007/s41062-020-00382-z.

Kumar, S., Rai, B., Biswas, R., Samui, P., & Kim, D. (2020). Prediction of rapid chloride permeability of self-compacting concrete using Multivariate Adaptive Regression Spline and Minimax Probability Machine Regression. *Journal of Building Engineering*, *32*, 101490. http://dx.doi.org/10.1016/j.jobe.2020.101490.

Kuznetsov, V.M. (1973). The mean diameter of the fragments formed by blasting rock. *Soviet Mining Science*, *9*, 144-148. http://dx.doi.org/10.1007/BF02506177.

Lanckriet, G.R.G., Ghaoui, L., Bhattacharyya, C., & Jordan, M.I. (2003). A robust minimax approach to classification. *Journal of Machine Learning Research*, *3*, 555-582. http://dx.doi.org/10.1162/153244303321897726.

Li, N., Nguyen, H., Rostami, J., Zhang, W., Bui, X., & Pradhan, B. (2022). Predicting rock displacement in underground mines using improved machine learning-based models. *Measurement*, *188*, 110552. http://dx.doi.org/10.1016/j.measurement.2021.110552.

Li, S., Zhao, H., & Ru, Z. (2017). Relevance vector machine-based response surface for slope reliability analysis. *International Journal for Numerical and Analytical Methods in Geomechanics*, *41*, 1332-1346. http://dx.doi.org/10.1002/nag.2683.

Mohamed, F., Riadh, B., Abderazzak, S., Radouane, N., Mohamed, S., & Ibsa, T. (2019). Distribution analysis of rock fragments size based on the digital image processing and the Kuz-Ram model Cas of Jebel Medjounes Quarry. *Aspects in Mining & Mineral Science*, *2*, 325-329. http://dx.doi.org/10.31031/amms.2019.02.000545.

Noori, R., Khakpour, A., Omidvar, B., & Farokhnia, A. (2010). Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Expert Systems with Applications*, *37*, 5856-5862. http://dx.doi.org/10.1016/j.eswa.2010.02.020.

Ozkahraman, H.T. (2006). Fragmentation assessment and design of blast pattern at Goltas limestone quarry, Turkey. *International Journal of Rock Mechanics and Mining Sciences*, *43*, 628-633. http://dx.doi.org/10.1016/j.ijrmms.2005.09.004.

Pradeep, T., Bardhan, A., & Samui, P. (2021). Prediction of rock strain using soft computing framework. *Innovative Infrastructure Solutions*, *7*, 37. http://dx.doi.org/10.1007/s41062-021-00631-9.

Samui, P. (2012). Application of relevance vector machine for prediction of ultimate capacity of driven piles in cohesionless soils. *Geotechnical and Geological Engineering*, *30*, 1261-1270. http://dx.doi.org/10.1007/s10706-012-9539-9.

Samui, P., & Kim, D. (2017). Minimax probability machine regression and extreme learning machine applied to compression index of marine clay. *Indian Journal of Geo-Marine Sciences*, *46*(11), 2350-2356.

Shi, X.Z., Zhou, J., Wu, B.B., Dan, H., & Wei, W. (2012). Support vector machines approach to mean particle size of rock fragmentation due to bench blasting prediction. *Transactions of Nonferrous Metals Society of China*, *22*(2), 432-441. http://dx.doi.org/10.1016/S1003-6326(11)61195-3.

Singh, B.K., Mondal, D., Shahid, M., Saxena, A., & Roy, P.N.S. (2019). Application of digital image analysis for monitoring the behavior of factors that control the rock fragmentation in opencast bench blasting: a case study conducted over four opencast coal mines of the Talcher

Coalfields, India. *Journal of Sustainable Mining*, *18*, 247-256. http://dx.doi.org/10.1016/j.jsm.2019.08.003.

Stone, R.J. (1994). A nonparametric statistical procedure for ranking the overall performance of solar radiation models at multiple locations. *Energy*, *19*, 765-769. http://dx.doi.org/10.1016/0360-5442(94)90014-0.

Strohmann, T., & Grudic, G.Z. (2003). A formulation for minimax probability machine regression. *Advances in Neural Information Processing Systems*, *15*, 1-8.

Tao, J., Yang, X.G., Li, H.T., Zhou, J.W., Qi, S.C., & Lu, G.D. (2020). Numerical investigation of blast-induced rock fragmentation. *Computers and Geotechnics*, *128*, 103846. http://dx.doi.org/10.1016/j.compgeo.2020.103846.

Taylor, K.E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research, D, Atmospheres*, *106*(D7), 7183-7192.

Tipping, M.E. (2000). The relevance vector machine. *Advances in Neural Information Processing Systems*, *12*, 652-658.

Willmott, C.J. (1984). On the evaluation of model performance in physical geography. In G.L. Gaile & C.J. Willmott (Eds.), *Spatial statistics and models* (pp. 443-460). Dordrecht: Springer. https://doi.org/10.1007/978-94-017-3048-8_23.

Wong F.S. (1985). Slope reliability and response surface method. *Journal of Geotechnical Engineering*, *111*, 32-53.

Yang, Z., He, B., Liu, Y., Wang, D., & Zhu, G. (2021). Classification of rock fragments produced by tunnel boring machine using convolutional neural networks. *Automation in Construction*, *125*, 103612. http://dx.doi.org/10.1016/j.autcon.2021.103612.

Zhang, W., Zhang, R., Wang, W., Zhange, F., & Goh, A.T.C. (2019). A multivariate adaptive regression splines model for determining horizontal wall deflection envelope for braced excavations in clays. *Tunnelling and Underground Space Technology*, *84*, 461-471. http://dx.doi.org/10.1016/j.tust.2018.11.046.

Zhang, W., Zhang, R., Wu, C., Goh, A.T.C., & Wang, L. (2020a). Assessment of basal heave stability for braced excavations in anisotropic clay using extreme gradient boosting and random forest regression. *Underground Space*, *7*(2), 233-241. http://dx.doi.org/10.1016/j.undsp.2020.03.001.

Zhang, W., Zhang, R., Wu, C., Goh, A.T.C., Lacasse, S., Liu, Z., & Liu, H. (2020b). State-of-the-art review of soft computing applications in underground excavations. *Geoscience Frontiers*, *11*, 1095-1106. http://dx.doi.org/10.1016/j.gsf.2019.12.003.

Zhang, W.G., & Goh, A.T.C. (2013). Multivariate adaptive regression splines for analysis of geotechnical engineering systems. *Computers and Geotechnics*, *48*, 82-95. http://dx.doi.org/10.1016/j.compgeo.2012.09.016.

Zhang, W.G., Li, H.R., Wu, C.Z., Li, Y.Q., Liu, Z.Q., & Liu, H.L. (2020c). Soft computing approach for prediction of surface settlement induced by earth pressure balance shield tunneling. *Underground Space*, *6*(4), 353-363. http://dx.doi.org/10.1016/j.undsp.2019.12.003.