Prediction of Fragment size using Regression Tree and Multi Linear Regression Model

V. K. Pisipati
Mining Engineering Dept., NIT Rourkela
Rourkela, India
pkarthikvenkat@gmail.com

S. K. Bisoyi
Mining Engineering Dept., NIT Rourkela
Rourkela, India
sbsunil3@gmail.com

Dr. B. K. Pal
Mining Engineering Dept., NIT Rourkela
Rourkela, India
bkpal@nitrkl.ac.in

***Abstract* — This paper is aimed to develop and compare precise and applicable models based on Regression Tree (RT) and Multiple Linear Regression (MLR) to predict Mean Fragment Size (MFS). In this regard, 35 blasting operations were investigated and the most influential factors on the fragmentation, i.e. Hole diameter, Depth of hole, Spacing, Burden, Stemming length and Specific charge were measured. Also, the Mean Fragment size values for the considered blasting events were carefully measured using WipFrag image analysis software. Regression Tree analysis was done using RapidMiner Studio and Multiple Regression analysis was done using computer-aided solution SPSS (Statistical Package for the Social Sciences) to analyze the data obtained from the study areas. Seven parameters were input into the regression tree and multiple linear regression analysis to generate the model. Mean fragment size (MFS) out of the seven input parameters was dependent variable and the remaining six such as Drill hole diameter (HD), Stemming length (ST), Burden (B), Spacing (S), Specific charge (Q_C) and Depth of hole (DH) were input as independent variables. The reliability of the developed models was checked using several performance indices, i.e. R^2, MEDAE and RMSE. It is found that the performance indices obtained by the RT model are better compared to the MLR model.**

Keywords— Fragmentation, Blasting parameters, Regression tree, Multiple Linear Regression.

I. INTRODUCTION

The blasting activities play an essential part in the financial matters of the mining industry. The blasted rock muckpile and fragment sizes are very important since they affect the downstream processes from hauling to grinding. To minimize the cost of production, optimal fragmentation from a properly designed blasting pattern has to be achieved [1]. Large fragments adversely affect the loading and hauling equipment and increase the frequency of sorting of oversize boulders and secondary blasting, thereby increasing the cost of mining. Similarly, generation of fines is also undesirable as involves excessive explosive consumption. It is, therefore, desirable to have a uniform fragment distribution, avoiding both fines and oversized fragments to optimize the overall cost of mining. The rock fragmentation obtained as an outcome of blasting operations said to be optimum when it contains the maximum percentage of fragments in the desired size range [2]. To achieve an optimum rock fragmentation a blast with optimized controllable parameters should be designed so that the effects of the uncontrollable parameters could be minimized. The controllable parameters for optimum fragmentation can be fixed after induction of trial blasts in a mine and quantification of fragmentation. Quantification of fragmentation refers to the measurement of fragmentation in order to predict the necessary corrections in the blast design. These corrections when applied to the blast design results in almost acceptable fragmentation [3].

II. REGRESSION TREE MODEL

Decision tree (DT) is one of the nonparametric classification methods which can introduce a pattern classification of observations utilizing a simple technique [4]. The developed model by RT is defined as a simple and understandable structure for decision-making. In classification, recognition and estimation, RT is considered as a simple method. Nevertheless, it can be used for solving problems and instead of some complicated techniques like ANN. DT is shown as a series of questions where every equation is defined by a parameter/variable. A typical tree in this technique is designed by roots, branches, leaves and nodes. A graph of DT comprises nodes and branches which are shown by circles and their connections, respectively [5]. In a DT process, a parameter is chosen as root or the first node and subsequently, the first node is divided into several internal nodes based on a series of features. Generally, a DT can be drawn from the top to down where the root of RT is located at the top. The end of a chain which comprises root, branch and node is named as leaf [6]. Splitting operation in DT is applied by one of the predictor/input parameters, and range of them is selected based on minimization of mean square error (MSE). When the output of a system is a discrete set of values, it is named as classification tree and when the output of a system is a real set of values, it is named as RT [7]. Figure 1 schematically indicates a typical structure of a DT.

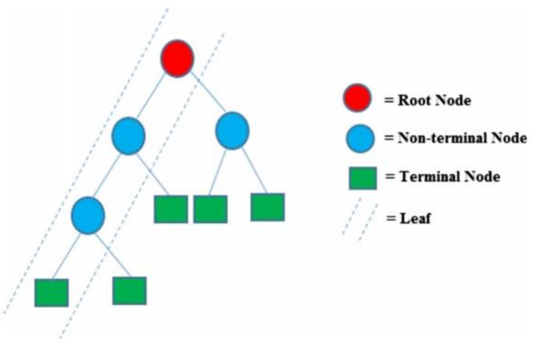

Fig 1: Structure of Decision Tree

Among all algorithms for determining DTs, classification and regression tree (CART) is considered as one of the famous ones and it has been widely utilized in approximating engineering problems (e.g. Tiryaki 2008). Hence, CART algorithm is applied for determining DT in the present study. CART algorithm was proposed by Breiman et al. (1984) [8] for determining purposes of DT. It is a rule-based method which can generate binary tree to recursively partition the predictor space into subsets. Breiman et al. (1984) mentioned that although CART algorithm was developed for quantitative variables, it can be utilized for any types of variables. It is worth mentioning that there are no initial assumptions regarding the relationships between parameters in CART algorithm. In this study, the output of the system is considered as real value, so that, RT should be applied to solve the problem. After the first node which is the root of the system (comprising of all data), each node is divided into two subsets. Each node is included till end leaf and model output can be predicted by end of each step. This process is repeated until the termination criteria are met.

III. Multiple Linear Regression Model

The MLR is one of the most well-known methods to fit a linear equation between one or more independent variables and one dependent variable. This method is extensively used to estimate some problems in the fields of rock and geotechnical engineering [9]. Generally, the MLR model can be formulated as follows:

$$Y= P+ P_1X_1+ P_2X_2+ P_3X_3+\ldots\ldots\ldots+ P_nX_n \qquad (1)$$

Where X_i (i=1,...., n) and Y are independent and dependent parameters, respectively. In addition, P_i (i=1,...., n) denote regression coefficients. In the present study, the MLR model was utilized and developed to predict mean fragment size. Developed MLR equation for the prediction of MFS is shown in equation (2).

IV. Image Analysis Using WipFrag

The WipFrag image analysis software uses the technique of analysis of a digital image of the blasted rock with granulometry system to predict the grain size distribution in the muck pile [10]. A camcorder acquires the images of the muck pile in the field. A scaling device is used in each view to reference the sizing. The muck pile is photographed or videotaped and this image is transferred to the WipFrag system. The broken rock image is transformed into a particle map or network. Network areas are converted into volumes and weights and the resulting data is displayed as a graph. The fidelity and speed of fragment edge detection allow fully automatic remote monitoring at a rate of one image per 3 to 5 seconds. More fragments are resolved, over a greater size range. WipFrag allows comparing the automatically generated net against the rock image. The fragment boundaries are analyzed efficiently using Edge Detection Variables (EDV). Inaccuracies can be corrected by manual editing with a mouse to improve edge detection. In the present study, using WipFrag image analysis software, 35 blasting muckpile photographs was analyzed in a system to obtain mean fragment size.

V. Experimental Site Details

To make accurate predictions of the fragment size, an applicable model based on multiple linear regression (MLR) is developed. In this regard, 35 blasting operations were investigated from two mines i.e. Dongri Buzurg mine, MOIL and Sonepur Bazari Project, ECL. The most influential factors on the fragmentation, i.e. hole diameter, blast-hole length, spacing, burden, stemming length, specific charge were measured.

A. Dongri Buzurg Mine (MOIL)

Dongri Buzurg mine is located in the village Dongri Buzurg, Balapur Hamesha and Kurmuda in Tumsar Tahsil under Bhandara district of Maharashtra state. It is in the Northeastern part of Bhandara district, in the state of Maharashtra and it is about 120 Kms from Nagpur. The manganese ore horizon occurs as a continuous bed at the stratigraphic contact of overlying Sitasongi formation and the underlying Munsar formation, on the reversed limb of a regional anticline within the Balapur Hamesha leasehold area of the mine. Considering the mineralization and disposition of manganese ore, it has been proposed to work with diesel hydraulic shovel and rear dumper combination. Horizontal slicing method of mining has been adopted in this mine for both extraction of ore and development. A diesel hydraulic backhoe in combination with existing 35T rear dumper and diesel-hydraulic backhoe in combination with 60T rear dumper have been proposed. Drill hole diameter of 110mm has been proposed for muck generation and 320 HP Dozers are proposed for bench preparation. Auxiliary equipment has also been proposed to ease the mining operations. Figure 2 shows the overview of Dongri Buzurg Mine, MOIL, India.

Fig 2: Overview of Dongri Buzurg Mine, MOIL, India.

B. *Sonepuri Bazari Project (ECL)*

Sonepur Bazari Project of Eastern Coalfields Limited is located in the Eastern part of Raniganj Coalfields. The Grand Trunk Road passes at 14km west of the project. Four coal seams viz. R-IV, R-V, R-VI and R-VII are mainly exposed in the mine. Presently, seams R-V and R-VI are being extracted by the opencast method of mining. The mine is producing about 3.5Mt of coal and removal of overburden is about 12 million cubic meters. The average stripping ratio of the mine is 4.72 m^3 per tonne coal produced. The total reserve of the project is 188.26 Mt. Figure 3 shows the overview of Sonepur Bazari Project, Eastern Coalfields Limited, CIL, India.

Fig 3: Overview of Sonepur Bazari Project, Eastern Coalfields Limited, CIL, India

VI. Prediction of Mean Fragment Size using RT and MLR Model

To make accurate predictions of the mean fragment size (MFS), an applicable model based on Regression tree (RT) and Multiple Linear Regression (MLR) are developed. In this regard, 35 blasting operations were investigated and the most influential factors on the fragmentation, i.e. hole diameter, blast-hole length, spacing, burden, stemming length, specific charge were measured. Also, the mean fragment size values for the considered blasting events were carefully measured using WipFrag image analysis software. Regression Tree analysis was done using RapidMiner Studio and Multiple Linear Regression analysis was done using computer-aided solution SPSS (Statistical Package for the Social Sciences) to analyze the data obtained from the study areas. Mean fragment size (MFS) out of the seven input parameters was dependent variable and the remaining six such Drill hole diameter (HD), Stemming length (ST), Burden (B), Spacing (S), Specific charge (Q_C) and Depth of Hole (DH) were input as independent variables. Descriptive statistics of parameters are given in tableI1 below.

TABLE I. Descriptive Statistics of Parameters

Parameter	**Mean**	**Std. Deviation**	**N**
MFS (mm)	129.006	90.174	35
HD (mm)	174.690	69.999	35
ST (m)	3.680	0.984	35
B (m)	3.643	1.204	35
S (m)	4.257	1.452	35
Q_C (Kg/m^3)	0.518	0.188	35
DH (m)	9.477	3.478	35

A. *Prediction of Mean Fragment Size using RT*

This section presents the modeling procedure of the RT model to predict the mean fragment size. As the first step of simulation works, the most influential parameters on Mean fragment size (MFS) i.e. Drill hole diameter (HD), Stemming length (ST), Burden (B), Spacing (S), Specific charge (Q_C) and Depth of Hole (DH) are selected. In this study, the developed models by RT were built with help of RapidMiner Studio 9.0. Two termination criteria including the number of interval and max tree depth were considered in designing RT models. Figure 4 displays the developed RT model to predict the mean fragment size showing and Specific charge (Q_C) parameter as a root node.

B. *Prediction of Mean Fragment Size using MLR*

This section presents the modeling procedure of the MLR model to predict the mean fragment size. As the first step of simulation works, the most influential parameters on Mean fragment size (MFS) i.e. Drill hole diameter (HD), Stemming length (ST), Burden (B), Spacing (S), Specific charge (Q_C) and Depth of Hole (DH) are selected. In this study, the developed models by MLR were built with help of computer-aided solution SPSS (Statistical Package for the Social Sciences). Mean fragment size (MFS) out of the seven input parameters was dependent variable and the remaining six such as Drill hole diameter (HD), Stemming length (ST), Burden (B), Spacing (S), Specific charge (Q_C) and Depth of Hole (DH) were input as independent variables. Table II displays the developed MLR model to

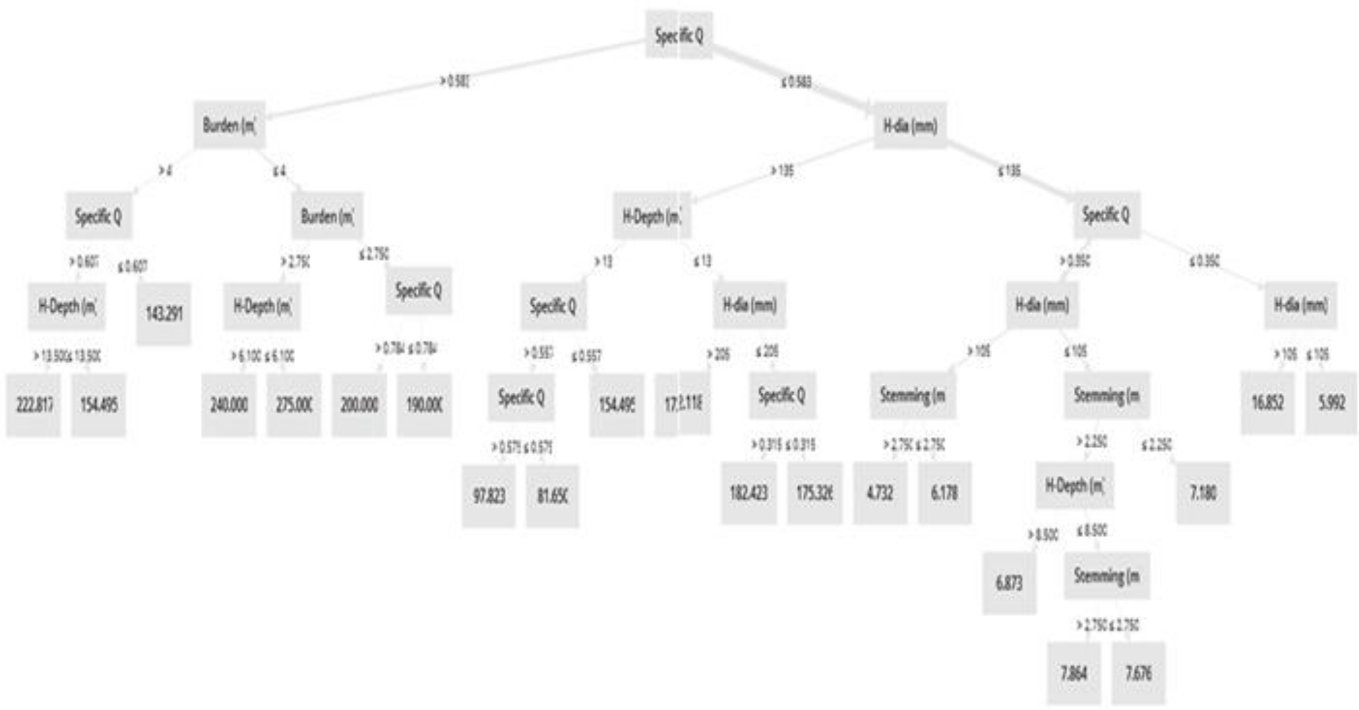

Fig 4: Structure of the developed regression tree model

predict mean fragment size and estimation multiple linear regression model for the Mean fragment size (MFS) is written as in equation (2):

Table 2: Computation for the model, Dependent variable: MFS

Parameter	Unstandardized Coefficients		t	Sig.	Tolerance
	B	Std. Error			
(Constant)	-255.173	42.131	-6.057	0.000	
HD	-1.449	0.455	-3.181	0.004	0.032
ST	20.179	20.933	0.990	0.331	0.076
B	71.614	19.404	3.691	0.001	0.059
S	44.406	22.299	1.991	0.056	0.031
Q_C	452.520	58.136	7.784	0.000	0.269
DH	-13.045	4.060	-3.213	0.003	0.162

Dependent Variable: MFS

$$MFS = -1.449(HD) + 20.719(ST) + 71.614(B) + 44.406(S) + 452.520(Q_c) - 13.045(DH) - 255.173 \quad (2)$$

Explanation of Parameters and their Coefficients:

a) Stemming length (ST) (20.719): This value indicates that as the stemming length increases by one unit, the mean fragment increases by 20.719 units. This interpretation is true only if the effects of other parameters are held constant. The t-test (0.990) associated with this value shows that Stemming length as a significant effect on mean fragment size.

b) Burden (B) (71.614): This value indicates that as the burden increases by one unit, the mean fragment increases by 71.614 units. This interpretation is true only if the effects of other parameters are held constant. The t-test (3.691) associated with this value shows that it is significant.

c) Spacing (S) (44.406): This value indicates that as the spacing increases by one unit, the mean fragment increases by 44.406 units. This interpretation is true only if the effects of other parameters are held constant. The t-test (1.991) associated with this value shows that it is significant. The spacing between the holes is making a significant contribution to the model.

d) Specific charge (Q_C) (452.520): This value indicates that as the specific charge increases by one unit, the mean fragment increases by 452.520 units. This interpretation is true if the effects of other parameters are held constant. The t-test (07.784) associated with this value shows that it is significant.

e) Depth of hole (DH) (-13.045): This value indicates that as the borehole depth increases by one unit, the mean fragment decreases by 13.045units. This interpretation is true only if the effects of other parameters are held constant. The t-test (-3.213) associated with this value shows that it is significant. Borehole depth is making significant contribution to the model.

VII. EVALUATION OF THE PREDICTIVE MODELS

This section is aimed to measure the accuracy of the proposed predictive models. In this regard, RT and MLR models are developed in this paper. In these models, the measured mean fragment size values are considered to be the product of the six input parameters, namely HD, S, B, ST, Q_C and DH. Some performance indices, i.e. median absolute error (MEDAE), R^2 and Root Mean Square error (RMSE), were used for evaluating the accuracy of models. Theoretically, RMSE, MEDAE and R^2 equal to 0, 0 and 1 indicate the best approximation. Performance indices of the predictive models for datasets are given in Table III below.

Table III: Performance indices of the predictive models for datasets

Predictive Model	Network Results		
	MEDAE	**R^2**	**RMSE**
RT	16.34	0.941	21.618
MLR	25.33	0.885	30.183

In this Table, when considering the obtained results of the RMSE for the RT model and MLR model, the values are 21.618 and 30.183 respectively. These values reveal a higher accuracy of the RT model. On the other hand, when considering the obtained results of the R^2 for the RT model and MLR model, the values are 0.941 and 0.885 respectively. These values demonstrate higher conformity of the RT model. In order to have a better comparison, the values of measured and predicted mean fragment size using MLR and RT models are plotted for all datasets, as shown in Figs. 5 and 6. What is clear from Table 4 and Figs. 5 and 6 is that the performance capacity of the RT model is higher than the MLR model.

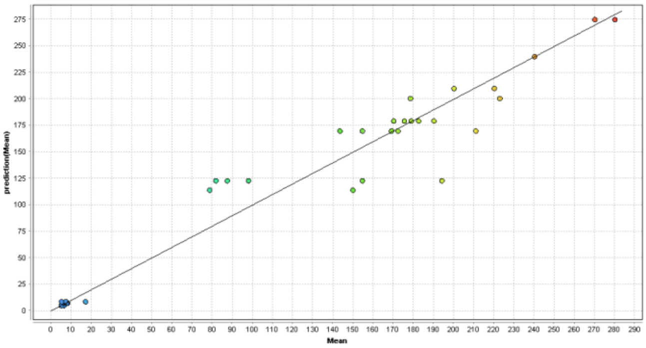

Fig 5: Measured versus predicted mean fragment size values by RT model

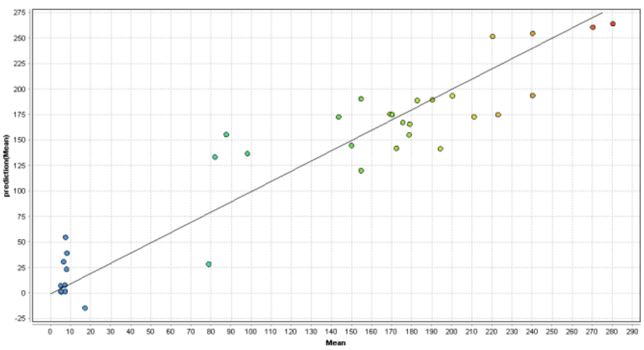

Fig 6: Measured versus predicted mean fragment size values by MLR model

VIII. CONCLUSIONS

In the present paper, RT and MLR models were developed for the prediction of mean fragment size caused by blasting in Dongri Buzurg mine, MOIL and Sonepur Bazari Project, ECL. The objective of a blasting is to generate a suitable muck pile having a suitable size distribution of the rock that can be efficiently loaded, transported and milled. The blasting operation affects all the other secondary activities, and the ultimate goal is to achieve the lowest costs of exploitation and processing. For this aim, 35 blasting events in the mines were investigated and the most influential parameters on the mean fragment size, namely HD, S, B, ST, Q_C and DH, were measured. The reliability of the developed models was checked using several performance indices, i.e. R^2, MEDAE and RMSE. Considering only results of R^2, values of 0.941 and 0.885 were obtained for datasets of RT model and MLR model respectively, that reveal the higher performance of this model in predicting mean fragment size, In addition to the R^2, MEDAE and RMSE values of proposed models were compared according to the accuracy of them. It is found that the performance indices obtained by the RT model are better compared to the MLR model. The end results indicated that both models were capable of predicting the mean fragment size; however, the most accurate result can be obtained by using the RT model.

REFERENCES

[1] Sharma, P. D., Blast Fragmentation Appraisal Means to Improve Cost-Effectiveness in Mines, Mining and Blasting Wordpress Journal, pp.1-14, 2010.

[2] Tekniska, H., International Symposium on Rock Fragmentation by Blasting, Australasian Institute of Mining and Metallurgy, Society for Experimental Mechanics (U.S.), pp. 312-316, 2009.

[3] Hartman, H. L., Mutmansky, J.M., Introductory to Mining Engineering, John Wiley & Sons publications, second edition, New Jersey (US), pp. 421-430, 2002.

[4] Tomczyk, A. M., Ewertowski M, Planning of recreational trails in protected areas: application of regression tree analysis and geographic information systems, pp.129–139, 2013.

[5] Vega, F. A., Matias, J. M., Andrade, M. L., Reigosa, M. J., Covelo, E. F., Classification and regression trees (CARTs) for modeling the sorption and retention of heavy metals by soil. J Hazard Mater 167, pp. 615–624, 2009.

[6] Tiryaki, B., Predicting intact rock strength for mechanical excavation using multivariate statistics, artificial neural networks, and regression trees, Eng Geology, pp.51–60, 2008.

[7] Lewis, R. J., An introduction to classification and regression tree (CART) analysis. In: Annual meeting of the society for academic emergency medicine in San Francisco, California, pp. 1–14, 2000.

[8] Breiman, L., Freidman, J., Olshen, R., Stone, C., Classification and regression trees, Wadsworth, Belmont, pp. 1-25, 1984.

[9] Gulden, K. U., Nese, G., A study on multiple linear regression analysis, Procedia-Social and behavior sciences, ELSEVIR publications, pp. 1-7, 2013.

[10] Maerz, N. H., Palangio, T. C. and Franklin, J. A., WipFrag image based granulometry system, Proceedings of the FRAGBLAST 5 Workshop on Measurement of Blast Fragmentation, Montreal, Quebec, Canada, pp. 91-99, 1996.