Geomechanics and Engineering, *Vol. 30, No. 3 (2022) 259-272*
https://doi.org/10.12989/gae.2022.30.3.259

Decision based uncertainty model to predict rockburst in underground engineering structures using gradient boosting algorithms

Richard Kidega[1], Mary Nelima Ondiaka[1], Duncan Maina[1], Kiptanui Arap Too Jonah[1] and Muhammad Kamran*[2]

[1]*Department of Mining and Mineral Processing Engineering, School of Mines and Engineering, Taita Taveta University, P.O. Box 635-80300 Voi, Kenya*
[2]*Department of Mining Engineering, Institute Technology of Bandung, Indonesia*

(*Received March 31, 2022, Revised June 20, 2022, Accepted July 5, 2022*)

Abstract. Rockburst is a dynamic, multivariate, and non-linear phenomenon that occurs in underground mining and civil engineering structures. Predicting rockburst is challenging since conventional models are not standardized. Hence, machine learning techniques would improve the prediction accuracies. This study describes decision based uncertainty models to predict rockburst in underground engineering structures using gradient boosting algorithms (GBM). The model input variables were uniaxial compressive strength (UCS), uniaxial tensile strength (UTS), maximum tangential stress (MTS), excavation depth (D), stress ratio (SR), and brittleness coefficient (BC). Several models were trained using different combinations of the input variables and a 3-fold cross-validation resampling procedure. The hyperparameters comprising learning rate, number of boosting iterations, tree depth, and number of minimum observations were tuned to attain the optimum models. The performance of the models was tested using classification accuracy, Cohen's kappa coefficient (k), sensitivity and specificity. The best-performing model showed a classification accuracy, k, sensitivity and specificity values of 98%, 93%, 1.00 and 0.957 respectively by optimizing model ROC metrics. The most and least influential input variables were MTS and BC, respectively. The partial dependence plots revealed the relationship between the changes in the input variables and model predictions. The findings reveal that GBM can be used to anticipate rockburst and guide decisions about support requirements before mining development.

Keywords: accuracy; gradient boosting algorithm; modelling; rockburst; sensitivity; specificity

1. Introduction

1.1 Backgroud

Rockburst is a major geohazard which affect the safety of numerous underground engineering structures and facilities for mining and civil works at large depths in high-stress conditions (He *et al*. 2012). Rockburst occurs when rock fragments and blocks are violently ejected from the surrounding rock, causing deformation and collapse of underground buildings, damage to equipment, and injury to mine personnel, with the possibility of fatalities. (He *et al*. 2012, Zhou *et al*. 2012). The rockburst hazards have a significant economic impact on both mining and civil projects, necessitating the development of effective methods to truly understand and predict their occurrences in order to decrease or eliminate the risk connected with them.

Predicting rockburst events requires a high level of accuracy, making it difficult to anticipate and hence resulting in fatal effects in numerous subterranean projects in China, Canada, South Africa, the United States, Australia, Switzerland, Peru, and elsewhere. In South Africa, 435 deaths occurred at Witwatersrand Mine in 1960, 224 people have died in the last 10 years in China (Wanger 2019, Chen *et al*. 2021, Dong *et al*. 2013, Adoko *et al*. 2013). These worst cases of rockburst demonstrate the global concern of conducting research that extends the understanding of the complexity of rockburst phenomena in underground projects prompting the requirement for efficient prediction before planning, designing, and operating underground facilities.

Accurate prediction of rockburst guides in decision-making to construct remedial support structures and improve strategies to avoid and/or control adversities in underground mines and civil operations (Zhou *et al*. 2006). The complicated and non-linear nature of the rockburst phenomena, which is attributed to the complex nature of geological structures, geometry, and rock mass mechanical parameter uncertainties and geometries, with limited reliable design guidelines and methodologies is a significant gap (Zhou *et al*. 2016, Albrecht and Potvin 2005). All of these limits the efficient application of various conventional methods including empirical and numerical methods to ascertain the occurrence in timing and location; hence, posing difficulty in its avoidance (Ahmad *et al*. 2021).

Conventional models used to predict rockburst occurrence in underground mines include experimental, empirical, and numerical techniques (Zhou *et al*. 2012, Zhou *et al*. 2006, Liu *et al*. 2019, Wang *et al*. 2018, Jia *et al*. 2020, Fan *et al*. 2019, Du *et al.* 2020). In most cases, the application of these models requires extensive skill on the

*Corresponding author, M.Sc.
E-mail: muhdkamran62@gmail.com

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

Table 1 State-of-the-art machine learning approaches proposed by the researchers for predicting rockburst

Machine Learning Algorithm	No of datasets	Input variables	Accuracy (%)	References
v-support vector regression	45	MTS, UCS, UTS, W_{et}, MTS /UCS, UCS/UTS	93.75	Zhu *et al.* (2008)
AdaBoost	36	MTS, W_{et}, MTS/UCS, UCS/UTS	87.8-89.9	Ge and Feng (2008)
Heuristic algorithms and support vector machines	132	MTS, UCS, UTS, W_{et}, MTS /UCS, UCS/UTS	66.67-90	Zhou *et al.* (2012)
Decision tree model	108 and 132	W_{et}, MTS /UCS, UCS/UTS	73-93	Pu *et al.* (2018)
J48	165	MTS, UCS, UTS, W_{et}	92.857	Ahmad *et al.* (2021)
Random forest	93	C_1, C_2, C_3, C_4, C_5, and C_6	80	Liang *et al.* (2020)
Adaboost	93	C_1, C_2, C_3, C_4, C_5, and C_6	66.66	Liang *et al.* (2020)
Gradient Boost Decision Tree	93	C_1, C_2, C_3, C_4, C_5 and C_6	76.67	Liang *et al.* (2020)
XGboost	93	C_1, C_2, C_3, C_4, C_5 and C_6	73.33	Liang *et al.* (2020)

Whereas C_1=Cumulative number of events, C_2=Logarithm of the cumulative release energy, C_3=Logarithm of the cumulative apparent volume, C_4=Event rate C4 (unit/day), C_5=Logarithm of the energy rate, C_6=Logarithm of the apparent volume rate, W_{et}=Strain energy storage index, UCS=Uniaxial Compressive strength, UTS=Uniaxial tensile strength of the rock; MTS=Maximum tangential stress and D=Depth of excavation

part of the practising engineer (Zhou *et al*. 2012, Zhou *et al*. 2006). Experimental techniques include triaxial testing, which cannot handle the non-linear nature of rockburst, in addition to the high cost and a lot of time required to set up a representative experiment to replicate field conditions (Qi *et al*. 2019, Lu *et al*. 2020, Miao *et al*. 2020, Zhai *et al* 2020).

Various researchers have proposed empirical methods for understanding and predicting rockburst (Hoek and Brown 2019, Lu *et al*. 2019). The common empirical models used to predict rockburst include; the energy release rate criterion, Hoek Brown failure criterion, Paul-Mohr-Coulomb criterion, stress criterion, strain energy criterion, rock brittleness criterion, tangential stress criterion, and energy release index criterion (Liu *et al*. 2019, Sun *et al*. 2016, Chen *et al*. 2016, Wang and Park 2001, Ma *et al*. 2018, Ma *et al*. 2018). These empirical methods are simple to apply but they cannot address the non-linear nature of rockburst (Zhou *et al*. 2006, Naji *et al*. 2018). The major advantages of empirical methods include the analytical simplicity, ease of interpretation and convenience to use since they are evaluated and visualized using simple equations, graphs, and curves (Zhou *et al*. 2006). However, their analytical simplicity is also a major limitation because underground mines require complex analysis that addresses a multifaceted geometry, thus restricting their efficiency for use in complex rockburst predictions (Zhou *et al*. 2006).

Similarly, many numerical models were proposed and applied in mining and other geotechnical fields to understand and/or predict rockburst occurrences in underground structures (Naji *et al*. 2019, Zhou *et al*. 2006, Naji *et al*. 2018, Zhang *et al*. 2013, Sun *et al*. 2007, Wang *et al* 2008). The main theories of numerical modeling in rockburst prediction include the rock failure process analysis theory (RFPA), discontinuous deformation analysis theory (DDA), variable weight theory, and matter element extension theory (Adoko *et al*. 2013, Kabwe and Wang 2015, Wang *et al* 2015). The difficulties exhibited by empirical approaches are overcome by numerical approaches, which use a constitutive model based on an appropriate failure criterion to predict rockburst (Zhou *et al*. 2006, Zhang *et al*. 2013). However, there is no precise constitutive model to accurately apply numerical methods in rockburst predictions (Zhou *et al*. 2012, Zhou *et al*. 2006). A comparison of numerical models with conventional empirical failure criteria helps to validate the actual circumstances that lead to the rockburst phenomenon (Naji *et al*. 2018). In addition, the lack of appropriate energy criteria and constitutive models to be used for precise prediction of rockburst using numerical models possess limitations on its use. These difficulties can be addressed by the use of machine learning (ML) modelling approaches because they use real cases of past rockburst events to forecast the future occurrence of the phenomena (Zhou *et al*. 2012, Adoko *et al*. 2013, Zhou *et al*. 2006, Pu *et al*. 2019, Pu 2019; He *et al*. 2015, Dong *et al*. 2013).

Various research studies have proposed and applied machine learning approaches to understand and predict rockburst (Ullah *et al* 2022, Zhou *et al*. 2012, Adoko *et al*. 2013, Zhou *et al*. 2006, Qi *et al*. 2019, He *et al*. 2015, Pohrt and Li 2014, Monjezi *et al*. 2011, Wang *et al*. 2009, Li *et al*. 2014). The most common advanced supervised machine learning methods used to predict rockburst include deep learning (Artificial neural network (ANN), support vector machine (SVM), random forest (RF), and gradient boosting algorithm (GBM) (Zhou *et al*. 2012, Adoko *et al*. 2013, Zhou *et al*. 2006, He *et al*. 2015, Dong *et al*. 2013). Table 1 indicates the state-of-the-art machine learning approaches proposed by the researchers for predicting rockburst. The ANN and SVM models cannot handle the mixed type of data, not effective in handling outliers and missing data values as the GBM (Qi *et al*. 2019, Keprate and Ratnayake 2017, Qi *et al*. 2018, Qi *et al*. 2018). The application of GBM has been applied in various rock engineering domains including; cement backfill strength (Qi *et al*. 2018), the fatigue life of small borehole piping (Keprate and Ratnayake 2017), circular model slope failure with remarkable performance (Zhou *et al*. 2019). Therefore, the GBM model stands to be the most appropriate especially in addressing rockburst-related problem analysis, since it hands a wide range of data types (Zhou *et al*. 2019). The key advantages offered by GBM are the robustness to

Table 2 Statistical description of the initial rockburst database

Variables	Uniaxial compressive strength	Uniaxial tensile strength	Maximum tangential stress	Depth of excavation	Stress ratio	Brittleness coefficient
Symbol	UCS	UTS	MTS	D	SR	BC
Unit	(MPa)	(MPa)	(MPa)	(m)	-	-
Mean	119.16	6.65	57.53	442.08	0.65	22.29
Standard Deviation	57.27	4.62	33.15	591.39	0.86	11.78
Minimum	18.32	0.38	2.6	0	0.05	4.48
Maximum	306.58	22.6	167.2	2520	5.26	80

handle outliers in output space, natural handling of data of mixed type, high predictive power and support for different loss functions (Keprate and Ratnayake 2017). The application of GBM is challenged by the limited capacity to scale a GBM model due to the sequential nature of boosting, and the GBM model is slow to train (Keprate and Ratnayake 2017). The computational cost of a GBM model is high due to the booting process. The strength of GBM models weights the weakness of GBM models, hence these form the core for the selection of GBM and the application for predicting rockburst in an underground engineering structures in this study.

One of the mistakes in choosing an optimum model is to develop and select a conceptual framework that answers the wrong question using the wrong input variables (Kuhn and Johnson 2013). Most researchers often tend to focus on the general model accuracy and unintentionally overlook the true nature of the problem to address the priority (Zhou *et al*. 2012, Adoko *et al*. 2013, He *et al*. 2015). Nonetheless, accuracy is very important, it only quantifies how well the model predicts the data. However, for the model to address a specific problem, a more direct evaluation criterion is required (Kuhn and Johnson 2013). In geo-hazard analysis like rockburst, the primary goal is to accurately predict at what predictor values rockburst events occur as opposed to at what value of predictors' does not cause rockburst. Considering a priori prediction outcome aims at minimizing the consequences of correct and incorrect prediction (Kuhn and Johnson 2013). This applies to rockburst prediction since the consequences of falsely predicting 'Yes' event worse compared to a false prediction of 'No' rockburst events. Therefore, the trade-off between model sensitivity and specificity is suited for analyzing rockburst phenomena to overcome the challenges exempted by general model accuracy. Most of the methodologies for predicting rockburst by various researcher does not inform readers about the nature and priority of the model prediction relating to the two-class rockburst responses because the majority researchers focused on model accuracy only.

Finally, because different models, including empirical and machine learning models of the same algorithm with different input variables, have different accurate prediction rates, the quality of feature engineering of the input variables used in rockburst prediction has an impact on the mode performances (Papadopoulos and Benardos 2021; Zhou *et al*. 2018). Most researchers have not precisely investigated the influence of the various combination of input variables in assessing the model prediction output despite the resultant change inaccuracy.

1.2 Significance of the study

This study develops an uncertainty based GBM model to predict rockburst considering two key aspects for effective application of the model in solving engineering problems. Firstly, the focus of this study is to optimize the model performance using the trade-off between model accuracy, sensitivity, and specificity. Secondly, to investigate how different combinations of rockburst governing variables as input variables affect the amount of input features. These two considerations resulted in two classes of assessment in this study; assessing model performance using different means of performance optimization and assessing the effect of the number of input variables using the different combinations of input variables. Therefore, the resultant model with optimal performance under both instances was considered best to be applied in rockburst prediction. This screen the most appropriate model optimization metric (general accuracy and receiver operating characteristic values from sensitivity and specificity values) and various combinations of input variables with the best performance in rockburst prediction analysis.

2. Material and methods

2.1 Data acquisition

A total of one hundred and ninety (190) datasets were used to develop the gradient boosting machine (GBM) models (Zhou *et al*. 2012, Shirani and Taheri 2019, Su *et al*. 2010, Adoko *et al*. 2013, Dong *et al*. 2013) data were obtained from the accessible published scientific literature on rockburst in underground tunneling projects' from 1994 to 2018. Table 2 indicates the statistical description of initial rockburst database. Fig. 1 depicts the Pearson rank matrices of correlations among the variables governing the 'Yes' and 'No' rockburst incidents, respectively. The variables are the uniaxial compressive strength (UCS); uniaxial tensile strength (UTS) of the rock; maximum tangential stress (MTS); depth of excavation (D); stress ratio (SR) and brittleness coefficient (BC).

2.2 Gradient boosting algorithm (GBM)

The Gradient boosting algorithm (GBM) was proposed by Friedman in 2001. It is a highly customizable ensemble supervised machine learning algorithm based on a decision

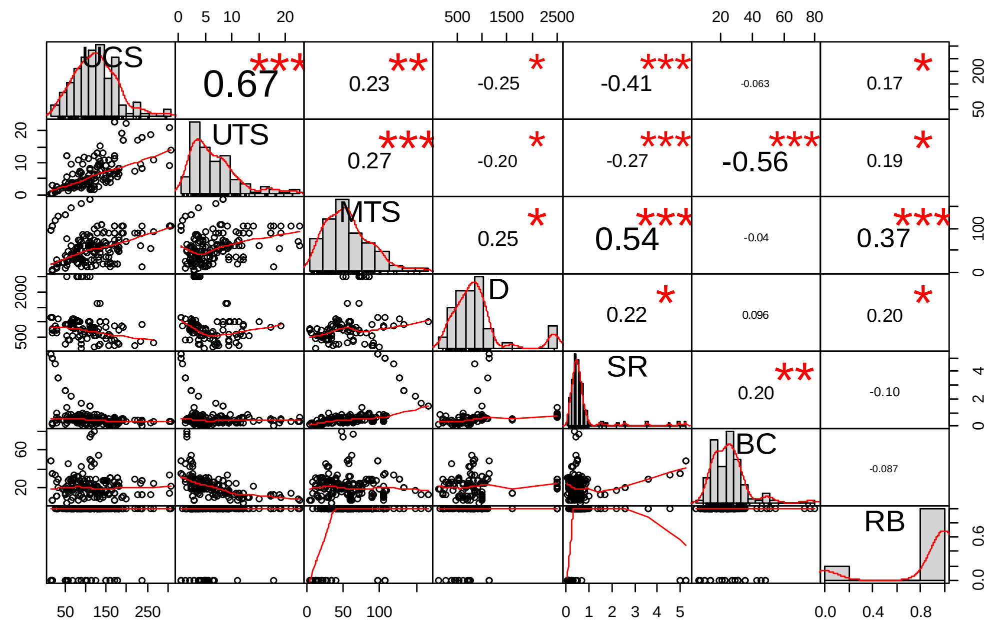

Fig. 1 Variable correlations for the ‘Yes’ and ‘No’ rockburst incidents

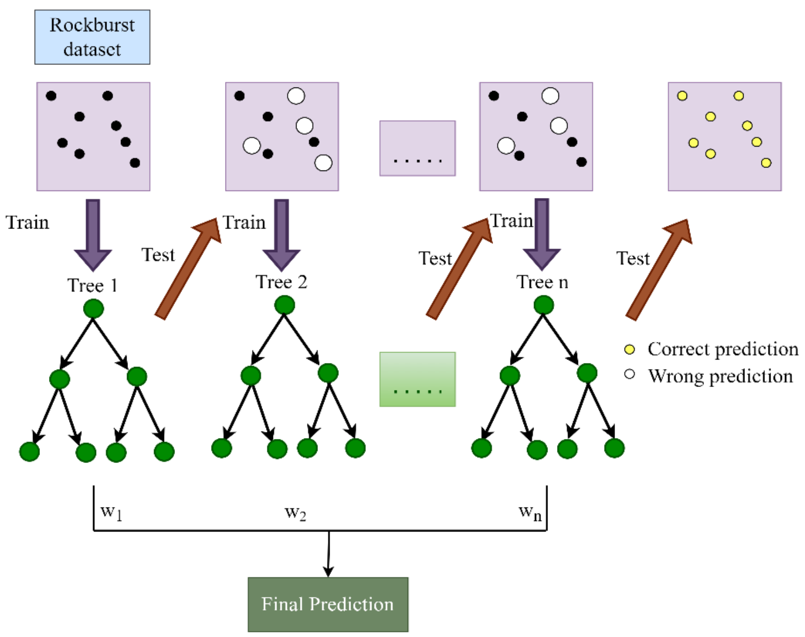

Fig. 2 Flowchart of GBM employed in this study

tree (Friedman 2001). The GBM builds models by consecutively fitting new models in a gradual additive sequence by minimizing the loss function to give a more accurate estimation of the output variables (Friedman 2001, Gollapudi 2016). The GBM is described below to show the sequential ensemble processes of weak learners based on decision trees to minimize the loss function. The flowchart of GBM is illustrated in Fig. 2.

Friedman’s gradient boosting algorithm (GBM)

Inputs:

input data $(x, y)j = n$ where $x =$ influencing variables and y = response variables and n = number of samples.

number of iterations(m)

choice of the loss-function (y, f)

choice of the base-learner model $h(x, \theta)$

Algorithm: Source: (Friedman 2001, Natekin and Knoll 2013)

Algorithm: Gradient Boost	
Stage 1	Initialize (F_o) with a constant
Stage 2	**for** t = 1 to M **do**
Stage 3	Compute the negative gradient gt (x)
Stage 4	Fit new base-learner function h(x_i , t)
Stage 5	Find the best gradient descent step-size gt: g_t=arg mingi=1N.ψ[yi, F_t-1 xi+gh(x_i , t)]
Stage 6	Update the function estimate: $F_t \leftarrow$ F_t-1+ gh(x_i , t)
Stage 7	**end for**

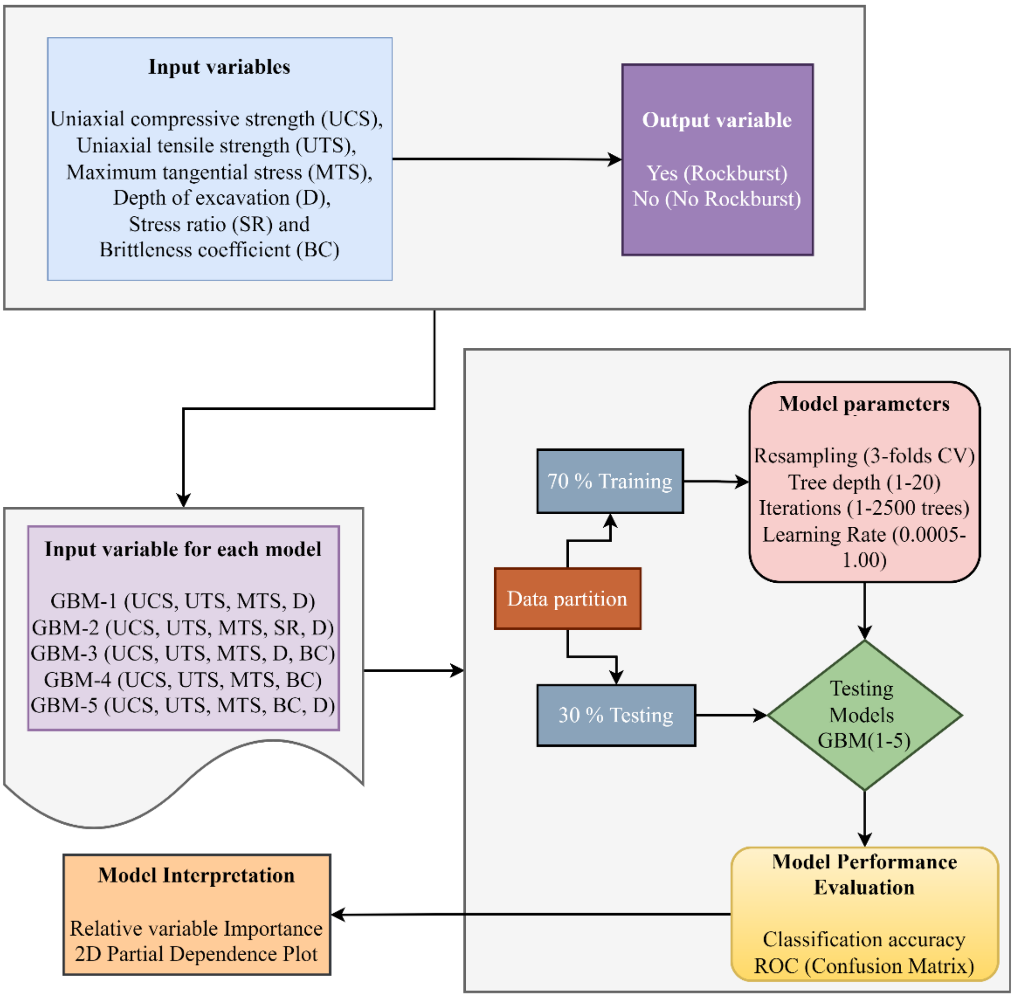

Fig. 3 Summary of modelling framework

The computational processes for learning the data, testing the model performance and interpreting the GBM model was performed in the R-Studio Software platform. All rockburst data process has been performed using *gbm* R-software packages to derive a relationship between the reaction (rockburst) and the influencing variables.

2.3 Design and Development of the Models

Fig. 3 illustrates the modelling framework applied in this study. The model was implemented using six input variables, that is UCS, UTS, MTS, D, SR, BC, and one output variable, that is rockburst incident classified as 'Yes' or 'No'. In this study, input variables are also referred to as 'features'.

The missing data were not imputed, missing data on the D has been addressed by the MTS since MTS and D are directly related to each other (Bruning *et al*. 2018). Also, the data were not normalized because tree-based models including GBM are notably insensitive to the characteristic of the data and are effective in handling missing, sparse, skewed, continuous, and categorical data without the need for pre-processing. Furthermore, the models can perform implicit feature selection, which is beneficial for many real-world modelling problems (Kuhn and Johnson 2013).

Table 3 The design of model input combinations

GBM Models	Number of input variables	Description of input variables
GBM-1	4	UCS, UTS, MTS and D
GBM-2	5	UCS, UTS, MTS, D and SR
GBM-3	5	UCS, UTS, MTS, D and BC
GBM-4	5	UCS, UTS, MTS, SR and BC
GBM-5	6	UCS, UTS, MTS, D, SR and BC

Table 3 shows the input variables for each of the five models implemented in this study. The variation in the number of input variables is designed to investigate the effect of combining different input variables in fitting a GBM model that predicts rockburst incidents. The data were randomly partitioned into training and validation (70%) and out-of-sample testing (30%) sets based on the method outlined by Dobbin and Simon (Dobbin and Simon, 2011). The training dataset was used to build and validate the model that classified rockburst incidents while the testing set was used to determine the capacity of the model in estimating the rockburst incidents on previously unseen data.

Random resampling was set to a 3-fold cross-validation, a technique found to be computationally efficient and

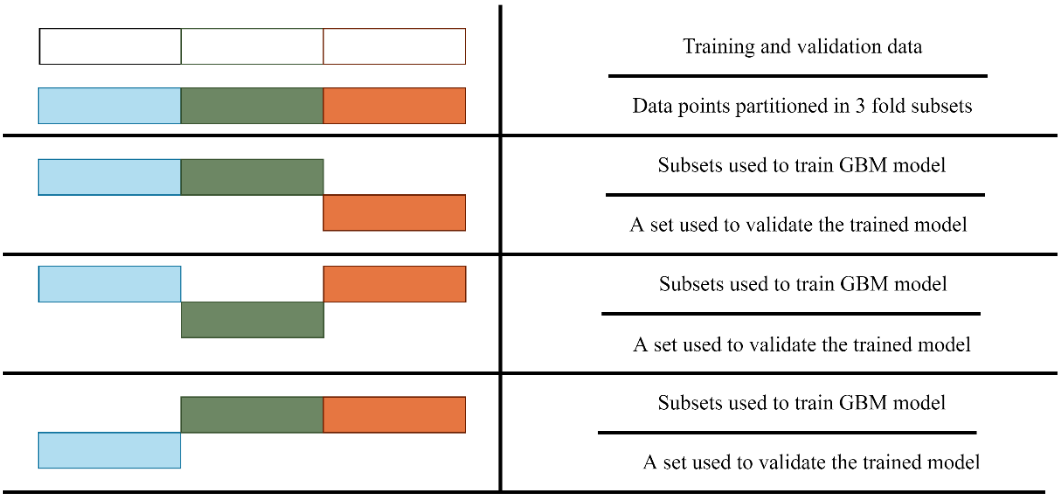

Fig. 4 Summary of a 3-fold cross-validation resampling method

performs well under uncertainty with less possibility of overfitting the model (Zhou *et al*. 2019, Kuhn and Johnson 2013). Fig. 4 summarizes the 3-fold cross-validation resampling method used during model training in this study. The data were randomly partitioned into three equal subsets or folds. The model was trained on 2-subsets of the data and the error was validated with the remaining 1-subset (or hold-out subset). This procedure was repeated until all the data subsets were used to train, and/or validate the model. An optimal model retained was based on the combination of input variables and the hyper-parameters of the entire training with the best performance estimated.

The model hyper-parameters, that islearning rate (v), number of boosting iterations or decision Trees (M), the tree depth (J), and the number of minimum observations in nodes were tuned during training to control the complexity and prevent overfitting the GBM models (Kuhn and Johnson 2013). The learning rate was set from 0.005 to 1.0, the minimum value selected to be within the 0.005 for small datasets with less than 500 data points (Friedman 2002).

The number of boosting iterations was set from 1 to 2500 in increments of 50 for the convenience, which determined the step-wise grid search for the optimum combination of the set hyper-parameter values and ensemble convergence of the predictive model. The range of values set for the tree depth was 1 to 20, which defined the highest level of variable interactions during model training. The number of minimum observations in a node ranged from 5 to 20 in increments of 5 for convenience. The optimal values of the hyper-parameters were selected based on the best-performing models in the preliminary modelling experiments.

Fig. 5 illustrates the steps involved in implementing the five models using different combinations of input variables, tuning and selecting the optimal values of hyper-parameters. The data were resampled, and the model fitted using the training hold-out datasets and repeated for all hold-out data. The resample results were aggregated in the performance profile and the best tuning parameters were chosen to be used in fitting the final model to the whole training dataset.

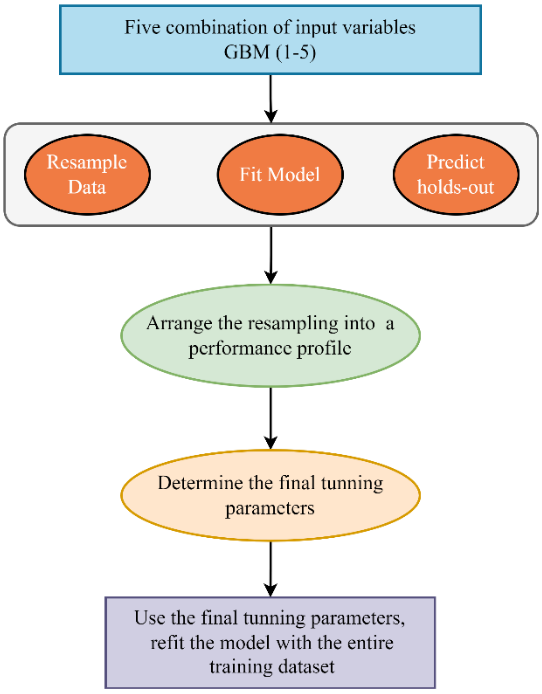

Fig. 5 A schematic illustration of the hyper-parameter tuning process

The established hyper-parameters' optimal values were used in training and validating the models in this study. A total of one hundred and fifty (150) trial models were generated based on the values of tuned parameters to find an appropriate combination and range of hyper-parameters to manage the learning process and achieve optimal GBM model performance. The grid search method iterates over all hyper-parameter ranges and combinations until optimal performance is obtained, at which point the parameter is declared optimised (Zhou *et al*. 2019). The best-performing model was selected for subsequent model testing analysis to ascertain potential application in predicting rockburst incidents in underground excavation projects.

2.4 Evaluation of the models

Predicted	Observed	
	Event (Yes)	Non-event (No)
Event (Yes)	TP	FP
Non-event (No)	FN	TN

Fig. 6 Confusion matrix for a binary rockburst classification problem

2.4.1 Accuracy of the models

The classification accuracy of the models was calculated using Eq. (1). The Cohen's kappa coefficient (k) was determined and used to assess the agreement between the observed and predicted binary rockburst incident cases as shown in Eq. (2).

$$Accuracy = \frac{Correctly\ predicted\ class}{Total\ testing\ class} * 100\% \quad (1)$$

$$Cohen's\ kappa\ coefficient\ (k) = \frac{O-E}{1-E} * 100 \quad (2)$$

Whereas O *is* the relative observed accuracy of the two-class raters and E is the expected accuracy of chance agreement in the prediction. The k is a chance agreement or probability estimate whose values range from -100 to 100 or 0 to 100 depending on the classification learning problem. The value of $k = -100$ defines inverse prediction, $k = 0$ defines chance or no agreement, and $k = 100$ defines perfect agreement between the observed and predicted classes (Kuhn and Johnson 2013).

2.4.2 Sensitivity and specificity of the models

The sensitivity and specificity of the models were determined using the true positive (TP), false positive (FP), true negative (TN), and false-negative (FN) values. A confusion matrix was developed to represent the 'Event' and 'Non-Event' attributes referring to the occurrence of rockburst and 'no occurrence of rockburst' respectively as shown in Fig. 6. For a two-class problem, a model's sensitivity and specificity characterize the efficacy of an event and non-event. (Kuhn and Johnson 2013).

$$Sensitivity = \frac{number\ of\ cases\ with\ 'yes'\ rockburst\ and\ predicted\ as\ 'yes'}{Number\ of\ cases\ having\ 'Yes'\ rockburst} \quad (3)$$

$$Specificity = \frac{number\ of\ cases\ with\ 'No'\ rockburst\ and\ predicted\ as\ 'No'\ rockburst}{Number\ of\ cases\ having\ 'No'\ rockburst} \quad (4)$$

2.4.3 Efficiency of the models

On the testing dataset, the models with the best prediction of "Yes" cases of rockburst indicated the models efficiency. The performance of the models to efficiently predict the rockburst incidents were assessed using the receiver operating characteristic (ROC) score. A comparison between ROC score and accuracy metrics has been carried out to select the suitable metric that resulted in the optimal classifier. In addition to the confusion matrix, ROC score of true positive (TP) and false positive (FP) rates for each model were determined and generated for both ROC and accuracy metrics.

2.4.4 Selection of the optimum model

The accuracy determines the agreement between the observed and predicted binary classifiers without distinguishing the misclassification error. The criteria used to select the optimum model considers the accuracy, sensitivity, specificity, and efficiency (ROC) scores of the five models developed. Both the accuracy and ROC are critically important because the available rockburst criteria have not been adequately considered. The model that exhibited the highest ROC on the testing dataset was selected for subsequent application in predicting the rockburst potential.

2.4.5 Relative influence of input variables based on the optimum model

The relative influence of each input variable in predicting rockburst incidents was determined based on the optimum model selected. The significance was presented using a two-dimensional (2-D) space bar plot of input variables versus the percentage of relative influence.

2.4.6 Partial dependence relationships

Eq. (5) (Hastie *et al*. 2009) was used to determine the contribution of independent input variables, which are UCS, UTS, MTS and D in predicting rockburst using partial dependence plots (PDPs) proposed by (Friedman 2001).

$$\hat{f_S} = \frac{1}{N}\sum_{i=1}^{N} f^*\ (x_s\ , x_{ci}\ ) \quad (5)$$

Where N is the number of samples in the training set and x_{Ci} through x_{cn} are observed values of x_c from the training set.

For the best model developed, the PDPs were used to show marginal effects on rockburst predictions based on the changes in the values of the independent features. The PDPs apply to independent input variables (Friedman 2001, Khoda and Ahmed 2021), thus the contribution of calculated input variables, which are SR and BC were not investigated. The R-Software (Version 3.6.2) was used to create the models and perform statistical computations and visualization using a supervised gradient boosting algorithm (GBM) algorithm. The software was installed on an Acer computer, Intel (R) Celeron (R) CPU N2830 dual-core of 2.16GHz, 4.00 GB (3.89 GB usable) RAM, and 64-bit Windows operating system.

3. Results and discussion

3.1 Development of the models

Table 4 presents the optimum model hyper-parameters, and their trends with `five different combinations of input variables. The different combinations of input variables affect the loss function of models (Zhou *et al*. 2021), which

Table 4 Summary of hyper-parameters optimization results

GBM Optimisation matrices	Hyper parameters	Range of values	Interval of values	Optimal values for each GBM Model				
				GBM 1	GBM 2	GBM 3	GBM 4	GBM 5
Classification Accuracy	Learning rate (v)	0.005,1.0	0.001	0.050	0.045	0.035	0.045	0.038
	No. of boosting iterations (M)	1, 2500	50	700	550	600	700	350
	Tree depth (J)	1, 20	1	9	10	9	5	10
	minimum observation in nodes	5, 20	5	5	15	5	20	5
Receiver Operating Characteristic values (ROC)	Learning rate (v)	0.005,1.0	0.001	0.455	0.455	0.450	0.020	0.045
	No. of boosting iterations (M)	1, 2500	50	1350	2300	2200	1500	2250
	Tree depth (J)	1, 20	1	9	9	5	5	6
	minimum observation in nodes	5, 20	5	5	5	5	5	5

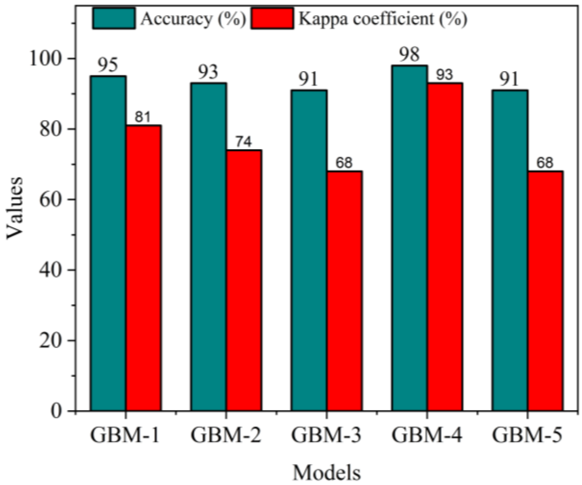

Fig. 7(a) Performance of the models optimised by using ROC metrics

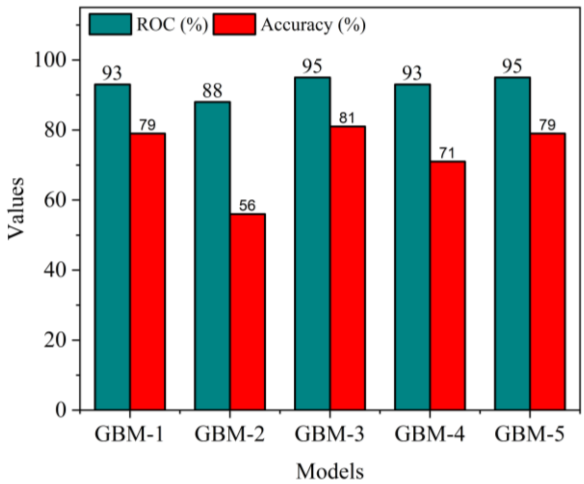

Fig. 7(b) Performance of the models optimised by using accuracy metrics

would explain the variations in the hyper-parameters that regularized the prediction of rockburst incident cases in this study. The hyper-parameters have been presented based on the accuracy and efficiency of the retained models in classifying the rockburst incidents, the latter measure being the receiver operating characteristic (ROC) test. The random resampling process placed 133 data values in the training and validation datasets while 57 data values were placed in the out-of-sample testing dataset.

3.2 Performance evaluation of the models

Figs. 7(a) and 7b present the performance of the five models based on the testing dataset. The highest (98%) and least (88%) accurate models were GBM-4 and GBM-2, respectively. Similarly, the *k* of GBM-4 is higher than that of GBM-2 by 37%.

The model accuracy results demonstrate that removing, adding, and/or binning input variables improves model performance, which corroborates the findings by (Kuhn and Johnson 2013). Similarly, changes in the combination of input variables have been found to improve the risk predictions (Zhou *et al*. 2021). However, care should be taken since some variables would cripple the model performance by increasing the variance and affecting interpretability (Kuhn and Johnson 2013). The performance improvement (differences in accuracy and *k*) of the GBM models were due to the omission and/or inclusion of the independent variable D and feature-engineered variables consisting of the SR and/or BC.

In this study, selective removing, adding/or and binning inputs variables using conventional rockburst prediction criteria like SR and BC significantly improved the model classification accuracy and *k*. The resulting *k* of 0.71 attained by GBM-4 shows a substantially better agreement between the predicted and observed values than the *k* of 0.543, which stipulates a good model that can be applied in prediction. The highest (98%) accuracy and *k* (93%) was attained by GBM-4 trained by optimizing model ROC metrics; however, the same GBM-4 which was trained by optimizing model accuracy attained an accuracy value of 93% and *k* of 71%. In general, models optimized using ROC metric attained much higher performance on the test datasets as compared to the models optimized by accuracy metric during the process of training, iteration and evaluations.

All the *k* values show a substantial agreement between the predicted and observed values according to (Zhou *et al*. 2021, Kuhn and Johnson 2013) except the value of 0.56 which exhibits a weak agreement between the predicted and observed values.

Fig. 8 shows the changes in the sensitivity and specificity of the five gradients boosting machine (GBM) models developed in this study. The values determine the accuracy of the model's performance with the value 1.0 representing the ideal prediction performance (Zhou *et al*.

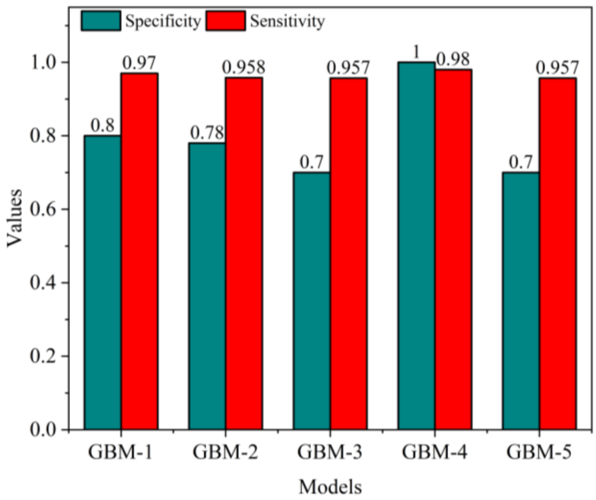

Fig. 8 Sensitivity and specificity of the GBM models

Table 5 Confusion matrices of the GBM-1

Predicted	Optimised metric accuracy		Optimised metric ROC	
	Observed		Observed	
	No	Yes	No	Yes
No	8	1	8	1
Yes	3	45	2	46

Table 6 Confusion matrices of the GBM-2

Predicted	Optimised metric accuracy		Optimised metric ROC	
	Observed		Observed	
	No	Yes	No	Yes
No	6	3	7	2
Yes	4	44	2	46

Table 7 Confusion matrices of the GBM-3

Predicted	Optimised metric accuracy		Optimised metric ROC	
	Observed		Observed	
	No	Yes	No	Yes
No	8	1	7	2
Yes	2	46	3	45

Table 8 Confusion matrices of the GBM-4

Predicted	Optimised metric accuracy		Optimised metric ROC	
	Observed		Observed	
	No	Yes	No	Yes
No	6	3	8	1
Yes	1	47	0	48

Table 9 Confusion matrices of the GBM-5

Predicted	Optimised metric accuracy		Optimised metric ROC	
	Observed		Observed	
	No	Yes	No	Yes
No	7	2	7	2
Yes	1	47	3	45

2019). The GBM-4 was the best performing model with sensitivity and specificity values of 1.00 and 0.957, respectively. While GBM-3 and GBM-5 were the least performing models that attained the same sensitivity and specificity values of 0.957 and 0.70, respectively.

The GBM-4 demonstrates the highest generalization capability in predicting rockburst occurrence while the GBM-3 and 5 show the least prediction capability as shown in Fig. 8. The GBM-1 was the second-best performing model, while the third one was GBM-2, with sensitivity and specificity of 0.8 and 0.78, 0.979 and 0.958, respectively. The performance of GBM-3 and 5 show similar loss functions.

The GBM-1 predicted 93.75% and 95.83% of TP cases, and each 88.89% of TN cases using the optimized accuracy and ROC metrics, respectively (Table 5). The GBM-2 predicted 91.67% and 96.83% of TP cases, and 66.67% and 77.78% of TN cases using optimized accuracy and ROC metrics, respectively (Table 6). The GBM-3 predicted 95.83% and 93.75% of TP cases, and 88.89% and 77.78% of TN cases using the optimized accuracy and ROC metrics, respectively (Table 7). The GBM-4 predicted 97.92% and 100% of TP cases, and 66.67% and 88.89% of TN cases using the optimized accuracy and ROC metrics, respectively (Table 8). Finally, GBM-5 predicted 97.92% and 93.75% of TP cases, and each of 77.78% of TN cases using the optimized accuracy and ROC metrics, respectively (Table 9). Hence, the model's performance evaluation demonstrate that the GBM-4 effectively predicted rockburst incidents on the testing dataset.

The optimum TP and TN cases prediction obtained is 100% and 88.98% by ROC optimized metrics. Therefore, it is clear that the performance of the GBM models is best when the receiver operating characteristic (ROC) curve is used as the optimization metric. It can be observed that true positive, true negative, false positive, and false negative for model four are; 48, 8, 0, and 1 respectively for the 57 test dataset. The means of optimizing a model performance (either accuracy or ROC) plays a critical role in the efficiency of the model, this study shows that model optimized by ROC (maximum TP prediction is 100%) is superior in predicting rockburst occurrence, compared to a model trained by optimizing accuracy (maximum TP prediction is 97.92). This difference in the model performance shows that the criteria for optimizing models affect their performance on a new dataset.

It is very important to make a key distinction about the type of error according to the problem being addressed (Kuhn and Johnson 2013). In rockburst prediction, the cost of errors made by the models is different on the production and safety of underground facilities. Therefore, in situations where the costs of errors are different, accuracy may not measure the prominence model characteristics (Kuhn and Johnson 2013). For example, the cost of misclassifying 'Yes' incidents is the loss of lives, damage to equipment, and negative effects on mine life and production. However, the cost of misclassifying 'No' incidents have no lethal effects.

Fig. 9 presents the relative influence of input variables of the selected model. The MTS was the most influential variable (36.13%) followed by the SR (18.41%), UCS

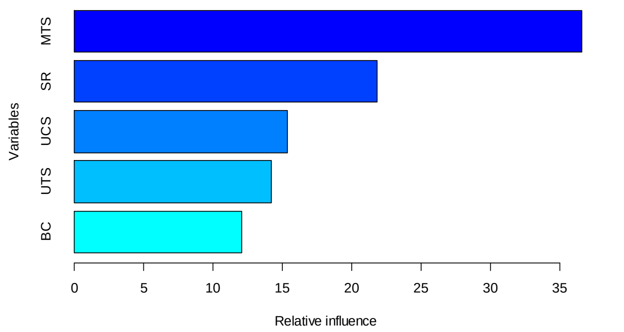

Fig. 9 Relative influence of the input variables for the GBM-4

(17.93%), UTS (14.16%), and BC (12.37%). The high relative influence of MTS corroborates findings by (Zhou *et al*., 2019; Zhao *et al*, 2013) that showed the variable as the most influential factor in predicting rockburst.

Furthermore, the local shape and stress concentration factor around the excavation relates to MTS, and thus the risks of rockburst in underground excavations (Keneti and Sainsbury 2018). The consistency of these results with (Zhou *et al*. 2021, Zhao *et al*. 2013, Keneti and Sainsbury 2018) confirms the validity and reliability of the GBM-4 in predicting rockburst in underground engineering projects.

The SR (or stress condition factor) controls the static loading to the intact strength of the rock that can result in rock failure and is a factor of safety in underground excavations (Zhou *et al*. 2006). The stress condition is the most important factor after MTS in contributing to rockburst (Zhou *et al*. 2021) which is corroborated in this study.

The UCS and UTS of rocks accounts for the ground support system capacity that controls the potential of rock mass failure that occurs if stress conditions exceed the UCS of the rock mass (Zhou *et al*. 2006; Zhou *et al*. 2021). Also, rockburst is more sensitive to UCS and UTS (Zhou *et al*. 2021). Therefore, UCS is a critical factor in the design of underground structures and assessing their potential to fail. The BC is the least sensitive variable towards predicting rockburst occurrence, which corroborates scientific findings (Zhou *et al*. 2021).

An increase in the demand for minerals and underground spaces has led to an increase in underground projects for mineral extraction and construction of civil works, respectively. As the D increase, the stress level exerted on the excavation surface also increases (Szwedzicki 2018). Thus, the D is a critical factor in any underground excavation project but is excluded as an input variable in the selected GBM-4, which could be attributed to the high correlation with MTS in non-rockburst incidents. A study (Kuhn and Johnson 2013) demonstrated that fitting a model with input variables having high association negligibly improves model performance.

In this study, the effect of the D on rockburst incidents might be explained as indirectly represented by the MTS in the best-performing model. The high MTS is mainly experienced at deep levels of excavations, however, there are exceptional cases of rockburst occurrence where high MTS have experienced at shallow depths <200 owing to the influence of adjacent excavation, topography, excavation methods, and seismic events (Larsson 2004; Wang *et al*. 2020).

Fig. 10 shows single-variable partial dependence plots (PDP) for the GBM-4. The PDPs depict curvilinear negative relationships between the input variables and the rockburst.

The contribution of uniaxial compressive strength (UCS) on rockburst prediction shows an increase in the negative probability of correctly predicting the 'Yes' events with an increase in UCS as depicted in Fig. 10. The prediction trend principally decreases to a peak between 100 MPa and 150 MPa then decreases with increasing UCS to about 200 MPa. When the UCS exceeds 200 MPa, the prediction score of rockburst exhibits a constant function

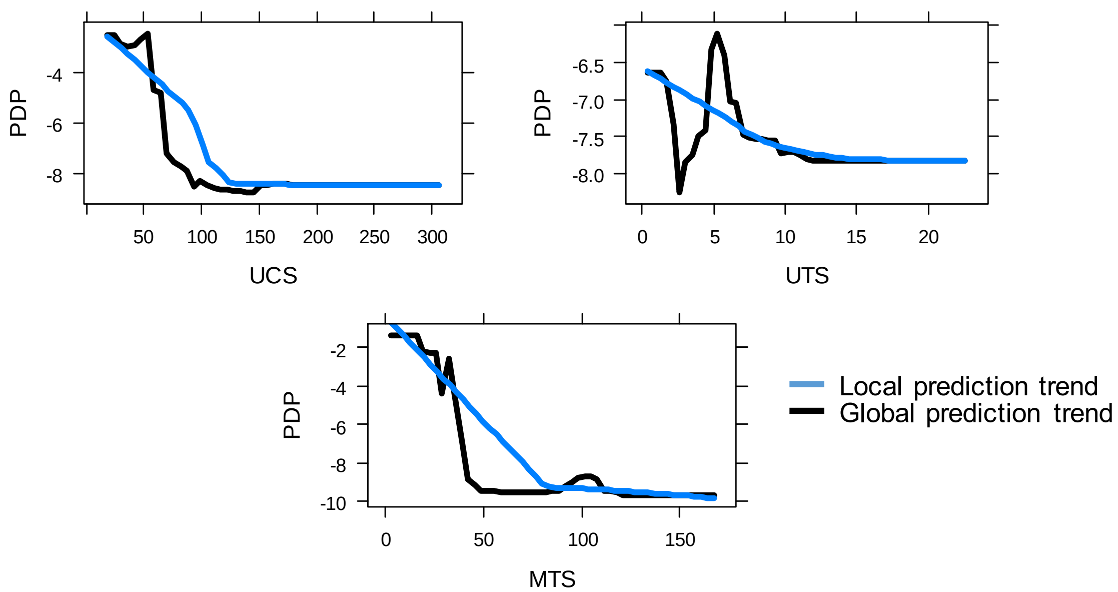

Fig. 10 The changes in UCS, UTS and MTS versus PDP rockburst predictions

demonstrating no further influence on the prediction. The ability of a rock mass to store energy is positively related to the UCS, and rockburst events, therefore the potential of rockburst increases with an increase in UCS (Zeng *et al*. 2020). Generally, GBM-4 partial dependence plot shows the higher rockburst event with increasing values of UCS.

The rockburst prediction score shows a sinusoidal negative marginal effect with increasing UTS as indicated in Fig. 10. Rock masses are known to be very weak when subjected to UTS compared to MTS and UCS making underground structures susceptible to fail exceeded (Wang and Park 2001), which account for the sinusoidal trend in rockburst prediction. However, a detailed investigation involving a controlled experiment is needed to explain the causal effect.

The slope exhibited by maximum tangential stress (MTS) shows trends similar to the UCS as shown in Fig. 10. These negative probability of correctly predicting 'Yes' events relationship effects could be masking stronger interaction effects with other rockburst governing variables. Scientific reports reveal that the hard and brittle rock masses under high-stress conditions have a high potential of experiencing rockburst (Zhai *et al*. 2020) supplemented with other internal and external variables governing rockburst (Szwedzicki 2018).

The intensity of stress-dependent rockburst occurrence is determined by the rock type, lithology, and loading conditions (Zhai *et al*. 2020) and increases with an increase in MTS (Li *et al*. 2017, Weng *et al* 2017, Li 2017). Overall, the PDPs demonstrate that rockburst is a multivariate phenomenon that requires the assessment of governing variables to enhance prediction accuracy. Rockburst has been explained by the interaction of multiple governing variables (Zhai *et al*. 2020, Szwedzicki 2018, Li 2017) which explains the sinusoidal curves in the prediction shown when single variable is investigated.

The empirical models demonstrate relationships between the MTS, UCS, and UTS (Ma *et al*. 2015) as well as SR and BC (Wang and Park 2001) in rockburst predictions. The model deployed to develop PDPs comprised the SR and BC variables that have not been investigated. However, the PDPs indicates that the smaller values of SR indicates rockburst incidence as opposed to larger values of BC. These are rational variations that would be a reasonable indicator of the similarity between the trends shown by the MTS and UCS on rockburst prediction compared to the UTS.

3.3 Application of the proposed model in underground engineering structures

Rockburst is mostly site-dependent and specific due to the heterogeneity of the rock mass and geological complexity (Zhou *et al.* 2021, Wang 2020). Rockburst occurrence is influenced by the UCS and UTS, underground excavation geometry, the partner and orientation of in-situ and mining-induced stress, and the method of mine excavation (Zhou *et al*. 2019, Akram *et al*. 2018, Brady and Brown 2007)

The GBM-4 developed and selected for analysis in this study explains rockburst in three major aspects. First, the model successfully classified the 'Yes' and 'No' rockburst incidents based on data use. Secondly, the relative importance of the variables that govern rockburst was determined to show a high and least influence of MTS and BC, respectively. Lastly, the partial relationship between changes in the values of measurable variables to the occurrence of rockburst was simulated to show that the potential of rockburst events increase with an increase in values of UCS, UTS and MTS. In this regard, the model's high prediction of yes-rockburst incidents (98%) and no-rockburst incidents (100%) on an out-of-sample dataset. Hence the decision based uncertainty model develops in this study demonstrates a capability to estimate the rockburst potential in underground engineering structures.

Moreover, for the effective application of this model, the variables governing rockburst require experimental determination from existing underground projects. This way, the model could be used to predict the potential of rockburst incidents using historical information. In the longer-term, modelling would aid in proper underground excavation sequences, geometry and design of structural support systems that reduce or eliminate rockburst occurrence.

Above all, a good understanding of the influential variables would guide decisions in the utilization of resources during geotechnical investigations and planning for the development of underground projects.

4. Conclusions

The uncertainty based GBM was effectively applied to predict rockburst incidents in underground excavation projects. Several GBM models were built using different combinations of input variables comprising the uniaxial compressive strength (UCS), tensile strength (UTS), tangential stress (MTS), depth of excavation (D), stress ratio (SR), and brittleness coefficient (BC). One model was retained, in this study referred as GBM-4 having classification accuracy, *k*, sensitivity and specificity values of 98 %, 93%, 1.00 and 0.957 respectively by optimizing model ROC metrics.. The variables that influence rockburst were found to be BC (12.37%), UTS (14.16%), UCS (17.93%), SR (18.13%), and MTS (37.13%). Despite the influence of D on stress intensity on the excavation openings, this variable was not amongst the features that produced the best model. The knowledge about the relative importance of the influential variables governing rockburst would be valuable in their assessment and monitoring during the development of underground excavation projects.

The gradient boosting algorithm (GBM) classification model was trained and tested using a dataset comprising one hundred ninety (190) data values. In the future, it is recommended that a large dataset, with an increased number of input variables should be investigated. The governing variables used as input variables were limited to UCS, UTS, MTS, D, SR and BC. Additional internal variables such as energy index and external variables such as the layout and structure of the underground facility, construction factors, excavation methods, topography and geological formations, the location of the underground facilities with the geological structures, high sidewall effects, and the intersection of caverns should be considered for inclusion in future rockburst prediction models. In addition, the researchers should accelerate the modelling of more complex geological conditions using more advanced modelling methods to sufficiently generate and provide more data to be used in machine learning to precisely predict the rockburst potential.

Acknowledgement

Richard Kidega, Mary Nelima Ondiaka, Duncan Maina, Kiptanui Arap Too Jonah are grateful to The German Academic Exchange Service (DAAD) In-Country/In-Region 2018/2020 scholarship, to study at Taita Taveta University Voi Kenya as a Centre of Excellence for Mining, Environmental Engineering & Resource Management (CEMEREM) Project scholar.

References

Adoko, A.C., Gokceoglu, C., Wu, L. and Zuo, Q.J. (2013), "Knowledge-based and data-driven fuzzy modeling for rockburst prediction", *Int. J. Rock Mech. Min. Sci*., **61**, 86-95. https://doi.org/10.1016/j.ijrmms.2013.02.010.

Ahmad, M., Hu, J.L., Hadzima-Nyarko, M., Ahmad, F., Tang, X.W., Rahman, Z.U., Nawaz, A. and Abrar, M. (2021), "Rockburst hazard prediction in underground projects using two intelligent classification techniques: A comparative study", *Symmetry*, **13**(4), 632. https://doi.org/10.3390/sym13040632.

Akram, M.S., Mirza, K., Zeeshan, M., Ali, M. and Ahmed, L. (2018), "Geotechnical investigation and prediction of rock burst, squeezing with remediation design by numerical analyses along headrace tunnel in Swat Valley, Khyber Pakhtunkhwa, Pakistan", *Open J. Geology*, **8**(10), 965-986. https://doi.org/10.4236/ojg.2018.810058.

Albrecht, J. and Potvin, Y. (2005), "Identifying the factors that control rockburst damage to underground excavations. In Identifying the factors that control rockburst damage to underground excavations", Australian Centre for Geomechanics, 519-528. https://doi.org/10.36487/acg_repo/574_56.

Brady, B.H.G. and Brown, E.T. (2007), "Rock mechanics and mining engineering", *Rock Mechanics for Underground Mining*, Springer, Dordrecht.

Bruning, T., Karakus, M., Akdag, S., Nguyen, G.D. and Goodchild, D. (2018), "Influence of deviatoric stress on rockburst occurrence: An experimental study", *Int. J. Min. Sci. Technol*., **28**(5), 763-766. https://doi.org/10.1016/j.ijmst.2018.08.005.

Chen, B., Gu, C., Bao, T., Wu, B. and Su, H. (2016), "Failure analysis method of concrete arch dam based on elastic strain energy criterion", *Eng. Fail. Anal*., **60**, 363-373. https://doi.org/10.1016/j.engfailanal.2015.11.045

Chen, Y., Zhang, J., Zhang, J., Xu, B., Zhang, L. and Li, W. (2021), "Rockburst precursors and the dynamic failure mechanism of the deep tunnel: A review", *Energi*., **14**(22), 7548. https://doi.org/10.3390/en14227548.

Dobbin, K.K. and Simon, R.M. (2011), "Optimally splitting cases for training and testing high dimensional classifiers", *BMC Med. Genom*., **4**(1), 1-8. https://doi.org/10.20944/preprints202007.0753.v1.

Dong, lj., Li, X.B. and Peng, K. (2013) "Prediction of rockburst classification using random forest", *Trans. Nonferrous Met. Soc. China*, **23**(2), 472-477. https://doi.org/10.1016/s1003-6326(13)62487-5.

Du, F., Wang, K., Guo, Y., Wang, G., Wang, L. and Wang, Y. (2020), "The mechanism of rockburst-outburst coupling disaster considering the coal-rock combination: An experiment study", *Geomech. Eng*., **22**(3), 255-264. https://doi.org/10.12989/gae.2020.22.3.255.

Fan, J., Chen, J., Jiang, D., Wu, J., Shu, C. and Liu, W. (2019), "A stress model reflecting the effect of the friction angle on rockbursts in coal mines", *Geomech. Eng*., **18**(1), 21-27. https://doi.org/10.12989/gae.2019.18.1.021.

Friedman, J.H. (2001), "Greedy function approximation: a gradient boosting machine", *Ann. Statist*., **11**(6), 332. https://doi.org/10.1214/aos/1013203451.

Friedman, J.H. (2002). "Stochastic gradient boosting", *Comput. Sstat. Data Anal*., **38**(4), 367-378.

Ge, Q. and Feng, X. (2008), "Classification and prediction of rockburst using AdaBoost combination learning method", *Rock Soil Mech., Wuhan*, **29**(4), 943.

Gollapudi, S. (2016), *Practical Machine Learning*, Packt Publishing Ltd.

Hastie, T., Tibshirani, R. and Friedman, J. (2009), *The Elements of Statistical Learning Data Mining, Inference, and Prediction*, Springer.

He, M., e Sousa, L.R., Miranda, T. and Zhu, G. (2015), "Rockburst laboratory tests database-application of data mining techniques", *Eng. Geol*., **185**, 116-130. https://doi.org/10.1016/j.enggeo.2014.12.008.

He, M., Xia, H., Jia, X., Gong, W., Zhao, F. and Liang, K. (2012), "Studies on classification, criteria and control of rockbursts", *J. Rock Mech. Geotech. Eng*., **4**(2), 97-114. https://doi.org/10.3724/sp.j.1235.2012.00097.

Hoek, E. and Brown, E.T. (2019), "The Hoek-Brown failure criterion and GSI-2018 edition", *J. Rock Mech. Geotech. Eng*., **11**(3), 445-463. https://doi.org/10.1016/j.jrmge.2018.08.001.

Jia, C., Wang, H., Sun, X., Yu, X. and Luan, H. (2020), "Study on rockburst prevention technology of isolated working face with thick-hard roof", *Geomech. Eng*., **20**(5), 447-459. https://doi.org/10.12989/gae.2020.20.5.447.

Kabwe, E. and Wang, Y. (2015), "Review on rockburst theory and types of rock support in rockburst prone mines", *Open J. Saf. Sci. Technol*., **5**(04), 104. https://doi.org/10.4236/ojsst.2015.54013.

Keneti, A. and Sainsbury, B.A. (2018), "Review of published rockburst events and their contributing factors", *Eng. Geology*, **246**, 361-373. https://doi.org/10.1016/j.enggeo.2018.10.005.

Keprate, A. and Ratnayake, R.C. (2017), "Using gradient boosting regressor to predict stress intensity factor of a crack propagating in small bore piping", *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, December. https://doi.org/10.1109/ieem.2017.8290109.

Khoda Bakhshi, A. and Ahmed, M.M. (2021), "Utilizing black-box visualization tools to interpret non-parametric real-time risk assessment models", *Transportmetrica A: Trans. Sci*., **17**(4), 739-765. https://doi.org/10.1080/23249935.2020.1810169.

Kuhn, M. and Johnson, K. (2013), "Book review: Max Kuhn and Kjell Johnson, Applied predictive modeling, New York, Springer", *Biometric*., **74**(1), 383-383. https://doi.org/10.1111/biom.12855.

Larsson, K. (2004), *Seismicity in Mines*, Birkhäuser.

Li, N. (2017), "Predicting underground tunnel hazards using machine learning techniques", Doctoral Dissertation, Caminos. https://doi.org/10.20868/upm.thesis.48301.

Li, T., Ma, C., Zhu, M., Meng, L. and Chen, G. (2017), "Geomechanical types and mechanical analyses of rockbursts", *Eng. Geology*, **222**, 72-83. https://doi.org/10.1016/j.enggeo.2017.03.011.

Li, Z., Li, B., Han, X. and Song, W. (2014), "Tunnel rockburst proneness study based on AHP-FUZZY method and field test", *Elec. J. Geotech. Eng.*, **19**, 117-128

Liang, W., Sari, A., Zhao, G., McKinnon, S.D. and Wu, H. (2020). Short-term rockburst risk prediction using ensemble learning methods", *Nat. Hazard.*, **104**(2), 1923-1946. https://doi.org/10.1007/s11069-020-04255-7.

Liu, X., Xia, Y., Lin, M. and Benzerzour, M. (2019), "Experimental study of rockburst under true-triaxial gradient loading conditions", *Geomech. Eng.*, **18**(5), 481-492. https://doi.org/10.12989/gae.2019.18.5.481.

Lu, J., Yin, G., Gao, H., Li, X., Zhang, D., Deng, B., Wu, M. and Li, M. (2020), "True triaxial experimental study of disturbed compound dynamic disaster in deep underground coal mine", *Rock Mech. Rock Eng.*, **53**(5), 2347-2364. https://doi.org/10.1007/s00603-019-02041-x.

Lu, Z., Ju, W., Gao, F., Feng, Y., Sun, Z., Wang, H. and Yi, K. (2019), "A new bursting liability evaluation index for coal-the effective elastic strain energy release rate", *Energi.*, **12**(19), 3734. https://doi.org/10.3390/en12193734.

Ma, C.S., Chen, W.Z., Tan, X.J., Tian, H.M., Yang, J.P. and Yu, J.X. (2018), "Novel rockburst criterion based on the TBM tunnel construction of the Neelum-Jhelum (NJ) hydroelectric project in Pakistan", *Tunnel. Underg. Space Technol.*, **81**, 391-402. https://doi.org/10.1016/j.tust.2018.06.032.

Ma, T.H., Tang, C.A., Tang, L.X., Zhang, W.D. and Wang, L. (2015), "Rockburst characteristics and microseismic monitoring of deep-buried tunnels for Jinping II Hydropower Station". *Tunnel. Underg. Space Technol.*, **49**, 345-368. https://doi.org/10.1016/j.tust.2015.04.016.

Ma, Y., Liu, C., Wu, F. and Li, X. (2018), "Rockburst characteristics and mechanisms during steeply inclined thin veins mining: a case study in Zhazixi Antimony mine, China", *Shock Vib.*, **2018**. Article ID 3786047. https://doi.org/10.1155/2018/3786047.

Miao, C., Sun, X., Zhang, Y., Wang, J., Zhang, J., Song, P. and Li, G. (2020), "Experimental study on the strain rockburst of calcareous sandstone containing joint surface", *Arab. J. Geosci.*, **13**(19), 1-10. https://doi.org/10.1007/s12517-020-06035-w.

Monjezi, M., Ghafurikalajahi, M. and Bahrami, A. (2011), "Prediction of blast-induced ground vibration using artificial neural networks", *Tunnel. Underg. Space Technol.*, **26**(1), 46-50. https://doi.org/10.1016/j.tust.2010.05.002.

Naji, A.M., Rehman, H., Emad, M.Z. and Yoo, H. (2018), "Impact of shear zone on rockburst in the deep neelum-jehlum hydropower tunnel: A numerical modeling approach", *Energi.*, **11**(8), 1935. https://doi.org/10.3390/en11081935.

Naji, A.M., Rehman, H., Emad, M.Z., Ahmad, S., Kim, J.J. and Yoo, H. (2019), "Static and dynamic influence of the shear zone on rockburst occurrence in the headrace tunnel of the Neelum Jhelum hydropower project, Pakistan", *Energi.*, **12**(11), 2124. https://doi.org/10.3390/en12112124.

Naji, A.M., Rehman, H., Emad, M.Z., Ahmed, S., Kim, J.J. and Yoo, H. (2019), "Rockburst evaluation in complex geological environment in deep hydropower tunnels. In Tunnels and Underground Cities", *Engineering and Innovation meet Archaeology, Architecture and Art*, CRC Press.

Natekin, A. and Knoll, A. (2013), "Gradient boosting machines, a tutorial", *Front. Neurorobot.*, **7**, 21. https://doi.org/10.3389/fnbot.2013.00021.

Papadopoulos, D. and Benardos, A. (2021), "enhancing machine learning algorithms to assess rock burst phenomena", *Geotech. Geolog. Eng.*, **39**(8), 5787-5809. https://doi.org/10.1007/s10706-021-01867-z.

Pohrt, R. and Li, Q. (2014), "Complete boundary element formulation for normal and tangential contact problems", *Phys. Mesomech.*, **17**(4), 334-340. https://doi.org/10.1134/s1029959914040109.

Pu, Y. (2019), "Machine learning approaches for long-term rock burst prediction", Doctor of Philosophy, University of Alberta. https://doi.org/10.7939/r3-7dwn-5c22.

Pu, Y., Apel, D.B. and Lingga, B. (2018), "Rockburst prediction in kimberlite using decision tree with incomplete data", *J. Sustain. Min.*, **17**(3), 158-165. https://doi.org/10.1016/j.jsm.2018.07.004.

Pu, Y., Apel, D.B., Liu, V. and Mitri, H. (2019), "Machine learning methods for rockburst prediction-state-of-the-art review", *Int. J. Min. Sci. Technol.*, **29**(4), 565-570. https://doi.org/10.1016/j.ijmst.2019.06.009.

Qi, C., Chen, Q., Fourie, A. and Zhang, Q. (2018), "An intelligent modelling framework for mechanical properties of cemented paste backfill", *Miner. Eng.*, **123**, 16-27. https://doi.org/10.1016/j.mineng.2018.04.010.

Qi, C., Fourie, A. and Chen, Q. (2018), "Neural network and particle swarm optimization for predicting the unconfined compressive strength of cemented paste backfill", *Constr. Build. Mater.*, **159**, 473-478. https://doi.org/10.1016/j.conbuildmat.2017.11.006.

Qi, C., Fourie, A., Chen, Q. and Zhang, Q. (2018), "A strength prediction model using artificial intelligence for recycling waste tailings as cemented paste backfill", *J. Clean. Prod.*, **183**, 566-578. https://doi.org/10.1016/j.jclepro.2018.02.154.

Qi, C., Tang, X., Dong, X., Chen, Q., Fourie, A. and Liu, E. (2019), "Towards intelligent mining for backfill: A genetic programming-based method for strength forecasting of cemented paste backfill", *Miner. Eng.*, **133**, 69-79. https://doi.org/10.1016/j.mineng.2019.01.004.

Shirani Faradonbeh, R. and Taheri, A. (2019), "Long-term prediction of rockburst hazard in deep underground openings using three robust data mining techniques", *Eng. Comput.*, **35**(2), 659-675. https://doi.org/10.1007/s00366-018-0624-4.

Su, G., Zhang, Y. and Chen, G. (2010), "Identify rockburst grades for Jinping II hydropower station using Gaussian process for binary classification", *2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering*, **2**, 364-367. https://doi.org/10.1109/cmce.2010.5609934.

Sun, J.S., Zhu, Q.H. and Lu, W.B. (2007), "Numerical simulation of rock burst in circular tunnels under unloading conditions", *J. China Univ. Min. Technol.*, **17**(4), 552-556. https://doi.org/10.1016/s1006-1266(07)60144-8.

Sun, X., Xu, H., Zheng, L., He, M. and Gong, W. (2016), "An experimental investigation on acoustic emission characteristics of sandstone rockburst with different moisture contents", *Sci. China Technol. Sci.*, **59**(10), 1549-1558. https://doi.org/10.1007/s11431-016-0181-8.

Szwedzicki, T. (2018), *Rock Mass Response to Mining Activities: Inferring Large-scale Rock Mass Failure*, CRC Press. https://doi.org/10.1201/9781315112336.

Ullah, B., Kamran, M. and Rui, Y. (2022), "Predictive modeling of short-term rockburst for the stability of subsurface structures using machine learning approaches: T-SNE, K-Means clustering and XGBoost", *Math.*, **10**(3), 449. https://doi.org/10.3390/math10030449.

Wang, C., Chuai, X., Shi, F., Gao, A. and Bao, T. (2018), "Experimental investigation of predicting rockburst using Bayesian model", *Geomech. Eng.*, **15**(6), 1153-1160. https://doi.org/10.12989/gae.2018.15.6.1153.

Wang, C., Wu, A., Lu, H., Bao, T. and Liu, X. (2015), "Predicting rockburst tendency based on fuzzy matter-element model", *Int. J. Rock Mech. Min. Sci.*, **75**, 224-232. https://doi.org/10.1016/j.ijrmms.2015.02.004.

Wang, J.A. and Park, H.D. (2001), "Comprehensive prediction of rockburst based on analysis of strain energy in rocks", *Tunnel. Underg. Space Technol.*, **16**(1), 49-57. https://doi.org/10.1016/s0886-7798(01)00030-x.

Wang, J.L., Chen, J.P., Yang, J. and Que, J.S. (2009), "Method of distance discriminant analysis for determination of classification of rockburst", *Rock Soil Mech.*, **30**(7), 2203-2208.

Wang, M., Zhou, J.W., Shi, A.C., Han, J.Q. and Li, H.B. (2020), "Key factors affecting the deformation and failure of surrounding rock masses in large-scale underground powerhouses", *Adv. Civil Eng.*, **2020**, Article ID 8843466. https://doi.org/10.1155/2020/8843466.

Wang, Y.H., Chen, L.W. and Shen, F. (2008), "Numerical modeling of energy release in rockburst", *Rock Soil Mech.-Wuhan*, **29**(3), 790.

Weng, L., Huang, L., Taheri, A. and Li, X. (2017), "Rockburst characteristics and numerical simulation based on a strain energy density index: A case study of a roadway in Linglong gold mine, China", *Tunnel. Underg. Space Technol.*, **69**, 223-232. https://doi.org/10.1016/j.tust.2017.05.011.

Zeng, Y., Chen, J., Li, D. and Li, X. (2020) "Study on rock burst behavior and tendency identification of surrounding rocks in hard and brittle formations of deep and ultra-deep wells", *IOP Conf. Ser.: Earth Environ. Sci.*, **570**(3), 032056. https://doi.org/10.1088/1755-1315/570/3/032056.

Zhai, S., Su, G., Yin, S., Zhao, B. and Yan, L. (2020), "Rockburst characteristics of several hard brittle rocks: A true triaxial experimental study", *J. Rock Mech. Geotech. Eng.*, **12**(2), 279-296. https://doi.org/10.1016/j.jrmge.2019.07.013.

Zhang, C., Feng, X.T., Zhou, H., Qiu, S. and Wu, W. (2013), "Rockmass damage development following two extremely intense rockbursts in deep tunnels at Jinping II hydropower station, southwestern China", *Bul. Eng. Geol. Environ.*, **72**(2), 237-247. https://doi.org/10.1007/s10064-013-0470-y.

Zhang, P., Yi, C.P., Nordlund, E., Shirzadegan, S., Nyberg, U., Malmgren, L. and Nordqvist, A. (2013), "Numerical back-analysis of simulated rockburst field tests by using coupled numerical technique", *Ground Support 2013: Proceedings of the Seventh International Symposium on Ground Support in Mining and Underground Construction*, May.

Zhao, Z.N., Feng, X.T., Chen, B.R., Feng, G.L. and Chen, T.Y (2013), "Study of relativity between rockburst and microseismic activity zone in deep tunnel", *Rock Soil Mech.*, **34**(2), 491-497.

Zhou, J., Guo, H., Koopialipoor, M., Jahed Armaghani, D. and Tahir, M.M. (2021), "Investigating the effective parameters on the risk levels of rockburst phenomena by developing a hybrid heuristic algorithm". *Eng. Comput.*, **37**(3), 1679-1694. https://doi.org/10.1007/s00366-019-00908-9.

Zhou, J., Li, E., Yang, S., Wang, M., Shi, X., Yao, S. and Mitri, H.S. (2019), "Slope stability prediction for circular mode failure using gradient boosting machine approach based on an updated database of case histories", *Saf. Sci.*, **118**, 505-518. https://doi.org/10.1016/j.ssci.2019.05.046.

Zhou, J., Li, X. and Mitri, H.S. (2016), "Classification of rockburst in underground projects: Comparison of ten supervised learning methods", *J. Comput. Civil Eng.*, **30**(5), 04016003. https://doi.org/10.1061/(asce)cp.1943-5487.0000553.

Zhou, J., Li, X. and Mitri, H.S. (2018), "Evaluation method of rockburst: State-of-the-art literature review", *Tunnel. Underg. Space Technol.*, **81**, 632-659. https://doi.org/10.1016/j.tust.2018.08.029.

Zhou, J., Li, X. and Shi, X. (2012), "Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines", *Saf. Sci.*, **50**(4), 629-644. https://doi.org/10.1016/j.ssci.2011.08.065.

Zhou, J., Shi, X. and Li, X. (2016), "Utilizing gradient boosted machine for the prediction of damage to residential structures owing to blasting vibrations of open pit mining", *J. Vib. Control*, **22**(19), 3986-3997. https://doi.org/10.1177/1077546314568172.

Zhu, Y.H., Liu, X.R. and Zhou, J.P. (2008), "Rockburst prediction analysis based on v-SVR algorithm", *J. China Coal Soc.*, **33**(3), 277-281.

CC