International Journal of Mining Science and Technology 33 (2023) 555–571

Contents lists available at ScienceDirect

International Journal of Mining Science and Technology

journal homepage: www.elsevier.com/locate/ijmst

Rock mass structural recognition from drill monitoring technology in underground mining using discontinuity index and machine learning techniques

Alberto Fernández [a], José A. Sanchidrián [a,*], Pablo Segarra [a], Santiago Gómez [a], Enming Li [a], Rafael Navarro [b]

[a] Universidad Politécnica de Madrid – ETSI Minas y Energía, Spain
[b] Universidad de Salamanca – GIR Charrock, Spain

ARTICLE INFO

Article history:
Received 12 October 2022
Received in revised form 14 February 2023
Accepted 16 February 2023
Available online 30 March 2023

Keywords:
Drill monitoring technology
Rock mass characterization
Underground mining
Similarity metrics of binary vectors
Structural rock factor
Machine learning

ABSTRACT

A procedure to recognize individual discontinuities in rock mass from measurement while drilling (MWD) technology is developed, using the binary pattern of structural rock characteristics obtained from in-hole images for calibration. Data from two underground operations with different drilling technology and different rock mass characteristics are considered, which generalizes the application of the methodology to different sites and ensures the full operational integration of MWD data analysis. Two approaches are followed for site-specific structural model building: a discontinuity index (DI) built from variations in MWD parameters, and a machine learning (ML) classifier as function of the drilling parameters and their variability. The prediction ability of the models is quantitatively assessed as the rate of recognition of discontinuities observed in borehole logs. Differences between the parameters involved in the models for each site, and differences in their weights, highlight the site-dependence of the resulting models. The ML approach offers better performance than the classical DI, with recognition rates in the range 89% to 96%. However, the simpler DI still yields fairly accurate results, with recognition rates 70% to 90%. These results validate the adaptive MWD-based methodology as an engineering solution to predict rock structural condition in underground mining operations.

1. Introduction

A comprehensive knowledge and characterization of the rock mass is crucial in the design and planning stages of an underground mining operation. In short term planning, ore extraction activities are scheduled according to production requirements, their performance improvement being a key process for the subsequent downstream operations and ore processing. To this, production drilling can provide specific information of the structural condition of the rock mass, which further can be applied to adapt blast [1] and ground support designs [2].

Measurement-while-drilling (MWD) technology provides real-time measurements of operational parameters from drilling rigs, which can be regarded as the response of the rock mass during this operation [3] hence dependent on the geotechnic context of the mine. Relevant works in the area have studied the relations between drilling parameters and the characteristics of the rock mass, proposing analytical prediction models based on the characteristics of the site and the drill rigs used [4–7]. As these models have been proposed for the specific drilling parameters and geological context of the sites assessed, a universal model or formula based on drilling parameters that can be applied to any mine with relevant prediction ability does not exist. Hence, the development of a methodology that could be broadly integrated is of particular interest to extend the application of MWD technology in mining. Recent international research projects to develop sustainable and intelligent comprehensive rock mass characterization applying MWD technology [8–10] have pointed at machine learning (ML) techniques as cutting-edge solutions for data analysis. These techniques are increasingly being used in the analysis of drill monitoring datasets [11]. Additional geotechnical data from expert geological analysis to identify discontinuities is instrumental at this point in order to derive the complex relations between drilling parameters and rock mass characteristics.

The purpose of this work is the definition of a comprehensive methodology based on drilling monitoring data analysis to recognize with high accuracy the structural condition of the rock mass. The outcome of this process is a dynamic prediction model formulated for the specific characteristics of the mine site. The validation

* Corresponding author.
E-mail address: ja.sanchidrian@upm.es (J.A. Sanchidrián).

https://doi.org/10.1016/j.ijmst.2023.02.004
2095-2686/

of this methodology was carried out by its capacity to detect discontinuities from MWD records for two underground mines in which the boreholes were further inspected with two different systems, like optical televiewer and digital endoscope. Previous literature shows that no research has explored the application of a general model based on MWD parameters for different drilling technologies. To this, it should be added that no work to date has used production drilling data from underground mining to predict the occurrence of individual discontinuities nor the quality of the detection has been quantitatively rated. On the contrary, qualitative, and visual evaluation has been used in the validation of the classification models of the rock mass into zones with different fracture levels [12]. Regarding the validation context, most of the conclusions about categorization of drilling parameters for rock features recognition derive from tests in laboratory or in mine under a controlled space [13–16], the present work being the first to quantify and validate the performance of MWD-based models in production environments for two different underground operational contexts. In addition, the application of ML techniques to MWD data has never focused on the detection of individual discontinuities, in this case using geological input data for training, while model performance has been only presented to predict rock mass rating or ore grade by ML approaches [17,18].

A combination of a hybrid ensemble algorithm and a resampling technique has been applied [19] for classification of rock mass discontinuity traces with imbalanced categories in the datasets; this combination of techniques is followed in this work on drilling data, a novelty in this field.

2. Background

2.1. Drill monitoring

This technology was initially developed in the oil industry and later applied to exploration drilling with the objective of measuring mechanical properties of the rock, together with hole path data as azimuth and dip, through sensors installed near the bit, for depths of several kilometers [3]. Later, this logging technology was adapted to smaller borehole diameters and roto-percussive rigs used in civil construction works, as tunnelling [2,20–22], mining and quarrying [23] both open pit [18,24–26] and underground exploitations [6,12].

MWD technology provides a dense cloud of data samples when collected from several drillholes, allowing the on-site recognition of rock characteristics that further can be used for downstream operations. The interpretation of drilling data requires further processing of the recorded signals [27,28] to eliminate the effects unrelated to rock mass properties. To help in this, direct measurements of the rock characteristics like in-borehole inspection with optical televiewer or endoscope, analysis-while-drilling (AWD), or in-situ strength measurements such as the Schmidt hammer [24,28,29] may be used to calibrate an index based on the recorded signals.

In mining, MWD has been used to investigate rock mass condition through the recognition of discontinuities, fracturing density or different lithologies, and their effect on blast design or over or under-excavation results [21,26,29–31]. This is performed by combining drill parameters when the goal is the assessment of rock mass strength [32], or by including variations in the drilling signals, for analysis of the structural condition of the rock mass [7,27]. For rotary-percussive drilling, the parameters are classified as independent or dependent regarding the influences of the geological features of the rock in the drill rig [24,33,34]. The first type comprises parameters like feed pressure (FP), percussive pressure (PP) and rotation speed (RS), which are influenced only by drill rig capacity, drilling technique, and drill rig control system, while penetration rate (PR), rotation pressure (RP), damp pressure (DP) and water pressure (WP) depend on the previous group of parameters and also on how drilling reacts to the characteristics of the rock mass. These dependent drill parameters have been commonly used to correlate with the characteristics of the rock, although both independent and dependent drilling parameters respond to rock structural or strength changes [20,27], being their inclusion into an index relatively site dependent.

Insights about the use of drilling parameters for fracturing condition detection have pointed in some cases at FP as a good detector of anomalies or discontinuities in the rock [16]. An increase of the PP may indicate strong fracturing [20]. PR and RP are pointed as the parameters that respond with significant fluctuations in their signals when the drilling bit crosses a discontinuity [6]. WP also shows a large variability due to the presence of open fractures or inflow water zones [33]. It should be highlighted, however, that the responses and fluctuations in the drill parameters depend on the geotechnical context of the operation, and the characteristics of the drilling equipment.

To formulate a general model for fracturing level prediction, some authors combine with the same weights the variance of drill parameters as RP and PR [6,12] or FP, PP, and RP [23], into a single structural parameter called fracturing or discontinuity index (DI). To illustrate this, the more recent index [7] is defined at the sample point i as:

$$\mathrm{DI}_i = \sqrt{0.5\left(\frac{\mathrm{PR}_{\mathrm{var},i} - \overline{\mathrm{PR}_{\mathrm{var}}}}{std\mathrm{PR}_{\mathrm{var}}}\right)^2 + 0.5\left(\frac{\mathrm{RP}_{\mathrm{var},i} - \overline{\mathrm{RP}_{\mathrm{var}}}}{std\mathrm{RP}_{\mathrm{var}}}\right)^2} \quad \text{with } i = 1, 2, \cdots, L \tag{1}$$

where $\mathrm{PR}_{\mathrm{var},i}$ is PR variability; $\mathrm{RP}_{\mathrm{var},i}$ RP variability (calculated as the moving variance), at each sample point i; L the number of points; $\overline{\mathrm{PR}_{\mathrm{var}}}$ the mean of the PR variability; $std\mathrm{PR}_{\mathrm{var}}$ standard deviation of the PR variability; $\overline{\mathrm{RP}_{\mathrm{var}}}$ the mean of the RP variability; and $\mathrm{stdRP}_{\mathrm{var}}$ standard deviation of the RP variability.

Principal component analysis is applied in some works to extract a linear combination of drilling parameters, their variations and the DI to describe the response of the drill rig to discontinuities. The predictive capacity of these models is usually evaluated qualitatively from visual correlations between block models and geological mapping of the rock mass characteristics [6,7,12], and no quantitative measurement of the model performance is given. This prevents a sound comparison of the predictive accuracy of different models and limits their application to new datasets, so that they are seldom used for the design and execution of subsequent operations such as blast design and explosives charging.

2.2. Machine learning applications for drilling data

In recent years, ML techniques have been applied to solve highly complex nonlinear engineering problems in mining such as the selection of mining method, rock mechanics and blast design parameters and hazards evaluation [35]. In this line, most of the research on the application of ML for MWD data from mining aims to relate rock mass quality indices such as Rock Mass Rating (RMR), Geological Strength Index (GSI) or Barton's Q with drilling parameters [17,36,37] or with ore grade prediction [18].

The main aspects for the use of ML for MWD data are the number of features that defines the dimensionality of the data set and the quality of this data. Consequently, expert analysis from endoscope or televiewer records for structural assessment is suitable to be used as target for supervised learning. From these records, an imbalanced class distribution [38,39] may be obtained in rock mass structural categorization problems for a case e.g., of massive

rock with few discontinuities. As this will affect the model performance, especially to the prediction of less abundant classes, three approaches are generally followed: (i) re-sampling the original dataset by over-sampling the minority class, under-sampling the majority class [38,39], or blending under-sampling with over-sampling [40]; (ii) assigning weights to training examples [41]; or (iii) using ensemble learning methods by combining several models, as boosting-based or bagging algorithms, which resample the original data to provide balanced classes, and improve the performance of single classifiers [42,43].

For ML models, the characteristics of the datasets, and specifically the amount of data in each category are relevant for obtaining a low generalization error (i.e., accuracy of the algorithm to predict from unseen data) and a low variance when a different dataset is used. Most classification algorithms usually predict accurately samples of the controlling class, resulting in high prediction accuracies for training and testing, while samples of the less abundant classes will be misclassified. A good performance and generalization of the model is defined by the optimal capacity where both bias and variance are low [44]. For this, a sample dataset to train and validate the model and another to test or estimate its performance on unseen data are used; these sets correspond typically to 80% and 20% of the available data [45]. For this purpose, it is also important to provide a good validation scheme to evaluate ML models on a limited data sample. The k-fold cross-validation algorithm [46] is a common scheme used for model validation to increase the number of samples used for teaching the model and to reduce the randomness of the training and validation data selected. This contributes also to optimizing the bias-variance trade-off and preventing overfitting [47].

Among ML techniques, ensemble learning methods are being widely used for classifications problems, since they improve the predictive power of the learning systems by combining many basic models (weak learners), as decision trees or k-neighbors, and including random resampling of the data [48,49]. These approaches are usually even more effective when fast algorithms such as classification and regression trees are used [50]. Techniques as bagging or boosting have been used to generate prediction models based on drill features to assess rock mass quality [51,52]. Recent research work on different mining-related topics have applied the bagging technique through the Random Forest (RF) algorithm for regression and classification problems [19], the bagging technique having been pointed as very suitable for classification problems with binary data, showing robustness against class imbalance [43,53].

3. Data and model formulation

Fig. 1 summarizes the main steps of the process to generate the dynamic model for the prediction of the structural condition of the rock from MWD and geotechnical data.

Drill monitoring information is recorded in two underground mining operations (Lújar and Zinkgruvan) from production drilling, comprising MWD records and accompanying data from in-hole inspection. This database includes the signals generated from two rigs, an old retrofitted jumbo in Lújar and a modern unit with factory installed MWD technology in Zinkgruvan, which means different automation and control features of each rig control systems. Drilling data is processed according to current filtering and normalization methods [27,28], to further correlate it with the discontinuities observed in the same borehole. A novel method is proposed to describe quantitatively the condition of presence and absence of discontinuities as a binary sequence. These data are used to identify discontinuities using both a classical method based on the fracturing or discontinuity index and ML techniques. This allows to compare the classification accuracy in defining the optimum structural model of both approaches.

3.1. Mine sites and data overview

Field measurements were gathered in Zinkgruvan and Lújar underground mines. Zinkgruvan is owned by Lundin Mining, it is a deposit located in the southern part of the Bergslagen province

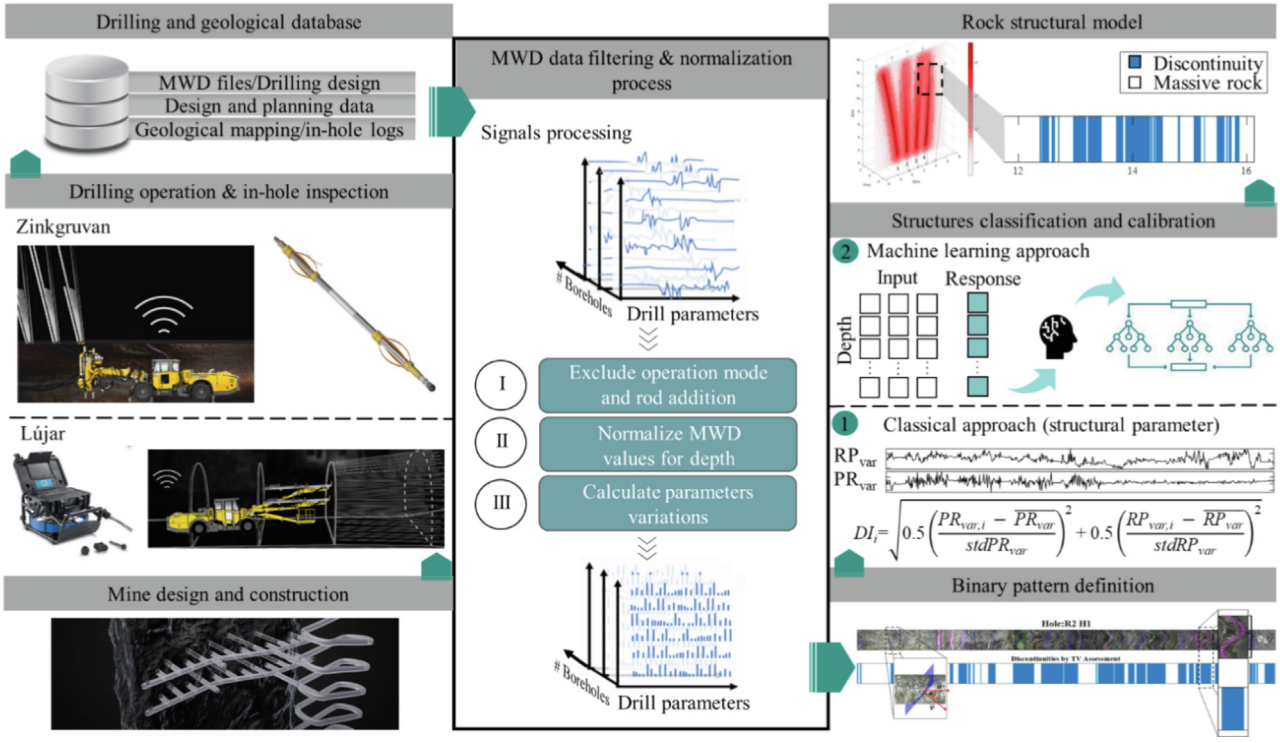

Fig. 1. Methodology for rock structural model generation.

of south-central Sweden. Long hole panel and sub-level bench stopping are used to mine zinc-lead and copper ores in a polymetallic deposit. The deposit comprises a stratiform, massive Zn-Pb deposit where the orebodies thickness ranges from 3 to 40 m. In the central part of the deposit the zinc-lead mineralization is stratigraphically underlain by a sub-stratiform copper stockwork. Lújar mine belongs to Minera de Orgiva and it is a narrow vein fluorite mine located in the Granada province, southern Spain. The mine is in the Alpujarride complex, composed mainly of carbonate formations and its stratigraphic series is composed, synthetically, of a basement of metamorphic rocks (quartzite, mica schist and phyllites) and a thick carbonate formation above them. The geological context in this site corresponds to mineralizations of sulphides of zinc, lead and fluorite, embedded in dolomite and calcite rock. The deposit is affected by two orogenic phases, resulting in cavities, faults and large open fractures that must be detected to adapt the explosive charging. The mining method is room and pillars adapted to the complex mineralization.

3.1.1. Zinkgruvan mine

Measurements in this operation correspond to a sublevel bench stope drilled with an Epiroc Simba E7C hydraulic long-hole production drill rig. Automatic drilling is conducted with 89 mm drill bits and 1.7 m long rods. The version of the control system is 4.14. The system provides PR (m/min), and a set of pressures, PP, FP, DP, RP, and WP (bar), at intervals of 2 cm. Fifteen boreholes from eight different rings were probed with an optical televiewer QL40 OBI-2G. The drilling pattern and the stope and access drift are shown in Fig. 2, where the holes logged with the televiewer are plotted in continuous colour lines and the rest in black dashed lines. The design of the stope consists of twenty-two rings with 4–6 upwards blastholes each, making up 93 boreholes, with inclinations with respect to the vertical axis of 0–48°. The length of the holes is about 15 m, with some boreholes having a shorter length, to fit the orebody. The nominal burden and spacing is 2.5 and 1.5 m, respectively. An emulsion explosive is used with variable density when fractured zones are recognized.

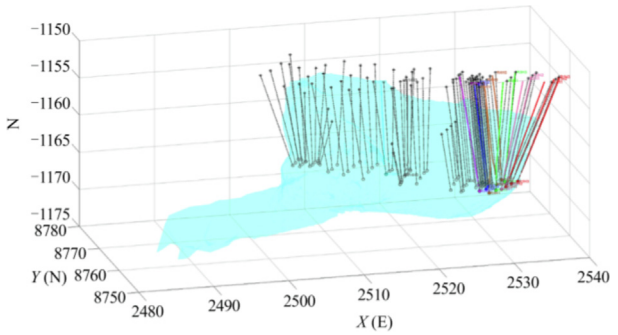

Fig. 2. 3D view of the drilling sketch from the production stope monitored in Zinkgruvan mine. Note: The orebody limits are showed in light blue.

Televiewer measurements were made from the collar of the hole up to 1.54 m from the hole collar, which corresponds to the length of the logging tool. The software WellCAD was used to identify three types of discontinuities that could affect the response of the drilling rig, namely, closed joints (CJ), shear zone (SZ) and change of lithology (COL), and three lithologies: zinc ore, probable ore, and sedimentary biotite gneiss. The latter has a uniaxial compressive strength (UCS) of 175 MPa (ranging from 100 to 275 MPa), and the ore 225 MPa (single value). In general, no mechanical differences occurred between ore and waste. The different discontinuities are marked in Fig. 3, where an image of each type is shown on the right part by way of example.

The joints mapped have small aperture, and a frequency that varies between 2.3 and 4 m^{-1}, with mean and standard deviation of (3.1±0.1) m^{-1}. The equation given by Priest and Hudson [54] to calculate the Rock Quality Designation-RQD index from the fracture frequency is applied. For this case, the resulting RQD is in the range 93.8%–97.6%, which indicates an excellent quality of the rock.

3.1.2. Lújar mine

A two-boom Atlas Copco 282 jumbo equipped with an in-house automatic recording system was used to drill 33 pseudo-horizontal boreholes in the level 70 of the mine. The jumbo is a rudimentary unit with a semi-automated control system, in which the feed pressure does not control the rotation pressure or the percussive pressure, that ultimately are controlled by the operator. The holes were drilled semi-automatically with a 4 m long initial rod and 3 m long additional rods. The system stores in a computer installed on the jumbo the following parameters: time, drilling length, hole ID, relative hole position, and three pressures (PP, FP, and RP), with a

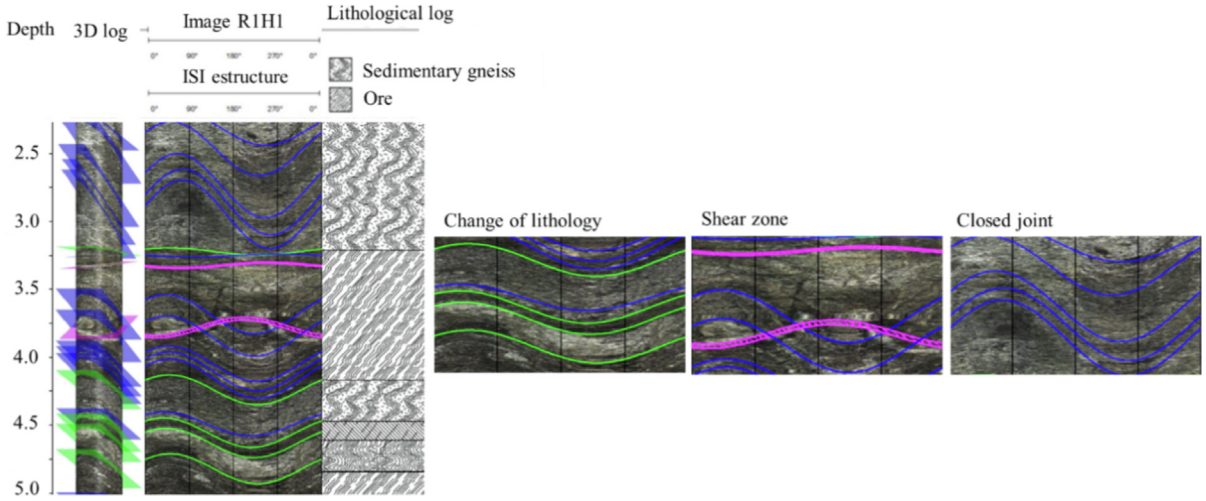

Fig. 3. 3D and unwrapped image of a section of hole 1 of ring 1 (R1H1). Note: Discontinuities are mapped in colors (COL in green, SZ in magenta, and CJ in blue).

sampling interval about 2 cm. The PR is calculated from the length and time recorded.

Most of the boreholes (25 holes) were drilled in small development heading blasts of nominal section 4×4 m. They were located in two development galleries identified as Zones A and B, see Fig. 4. The production holes have a diameter of 51 mm and a length in the range 3.5–4 m; they were logged with a digital endoscope and a measuring tape inside the borehole to reference the camera position. Pumped ANFO explosive is generally used though cartridges must be employed where cavities or faults are detected, in order to prevent explosive leakage and to improve blasting performance.

The rest of the boreholes (8 exploration drill holes) were drilled in Zone B (see red lines in Fig. 4) with a diameter of 63 mm; three of these holes have a length of about 16 m and the other five about 6 m. In these surveying holes, an endoscope manufactured by Forthaus Tech and provided with an encoder was used.

Rock strength data (UCS) ranks from 44 to 83 MPa for the waste and from 82 to 96 MPa for the ore. The structures, as fluorite occurrences and discontinuities, observed from the videos are shown in Fig. 5 over a geological profile of a half cast of the corresponding hole inspected after the blast. Small joints (red arrows in Fig. 5a) are not considered as their influence in MWD signals is limited in comparison with large discontinuities such as cavities or faults (Fig. 5b and f, respectively). The four discontinuity classes considered are: closed joint (CJ; Fig. 5g), open joint (OJ; Fig. 5c), change of lithology (COL; Fig. 5e), and cavity or fault (CAV; Fig. 5b and f).

Fracturing, cavities, and faults are abundant in some zones of the mine. The mean fracture frequency ranges from 6 to

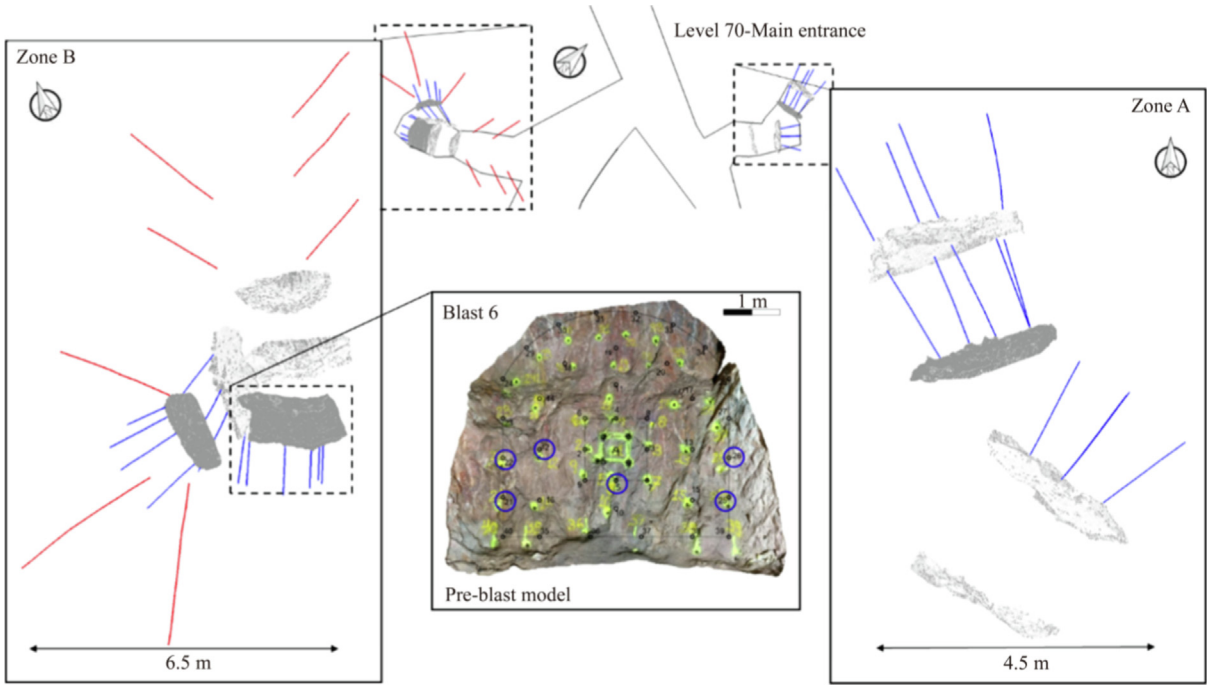

Fig. 4. Position of production (blue lines) and exploration (red lines) holes monitored in zone A (right) and B (left) and the detail of logged holes marked with blue circles for blast 6, at Lújar mine. Note: Grey clouds are the blast faces monitored.

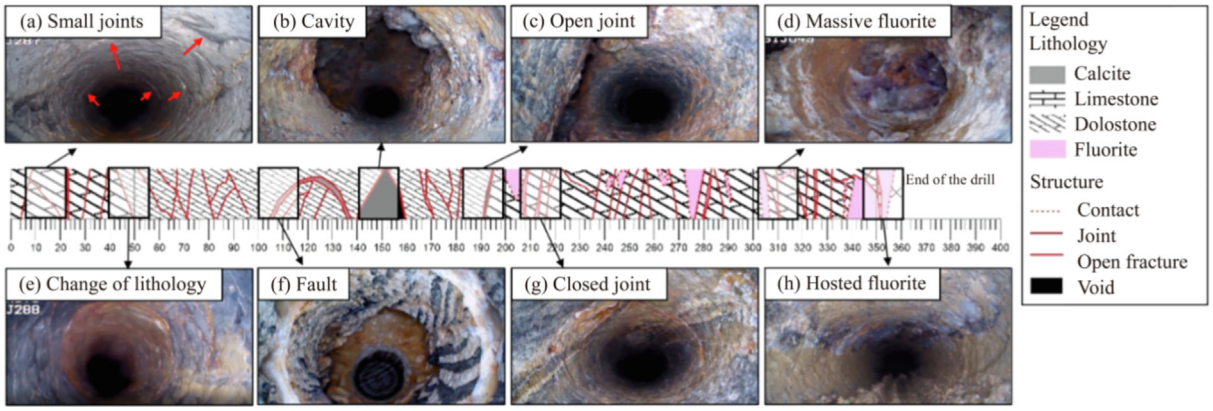

Fig. 5. Geological profile of the half cast of hole 30 (Blast No. 8) mapped after blasting in Lújar mine. Note: A representative image of the different discontinuities identified from optical endoscope videos are shown in the top and bottom image stripes.

20.8 m^{-1} with mean and standard deviation (13.4±5.9) m^{-1}. The resulting RQD in the range 36.1%–94.7% suggests zones with a fair rock quality, due to large discontinuities and strong fracturing like cavities and faults, and zones with excellent rock quality where massive rock (MR) is predominant.

3.2. Drilling data

In the case of Zinkgruvan, the MWD data recorded by the Rig Control System (RCS) for the 15 holes logged with the televiewer correspond to 10068 records. In Lújar, 9,830 drilling records are obtained in the 33 holes logged with the optical endoscope.

The probability density distributions of raw MWD parameters for both operations are represented in Fig. 6. Most of PR records from Lújar mine are concentrated at low values, and high values up to 2 m/min, are associated to a cavity or a fault that does not present resistance to the bit advance. FP in Lújar shows a bimodal distribution, about 32 and 47 bars, which may indicate two different rock mass strengths, this being a subject of further study on lithology recognition. For Zinkgruvan, the damper and flushing systems are started before the bit advances through the rock mass involving then DP and WP above zero. In addition, most of PP and DP are grouped at high values, but there are also lower pressures possibly related with the action of the control system. FP, RP, and PR include zero in the distribution due to either a rod addition or the presence of discontinuities. The automation system on the rig strives to keep the DP constant by the action of a two-step feed cylinder with a small and large area. The small area of the feed cylinder aims to reduce FP at the beginning of drilling with a new rod in order to prevent an increase of hole deviation when a rod is added, leading to the bi-modal distribution for FP in Fig. 6. In general, levels of drill parameters are lower in Lújar than Zinkgruvan reflecting differences in the drill rigs, their control system and the geomechanical characteristics of the rock mass.

To correlate drilling signals with borehole logs, the operational and mechanical effects in MWD data that could bias the recognition of discontinuities or lithologies are removed following the filtering and normalizing steps [27,28], see the central in box of Fig. 1.

3.3. Definition of a binary pattern of the structural rock characteristics

To correlate discontinuities recognized in the borehole walls with the response of the drill rig, a binary sequence (presence/absence) identified as D_{obs} is built [55]; the categories are assigned as value 1 for discontinuities (DISC) and 0 for massive rock (MR), as function of the borehole depth using the same resolution as the MWD system. The binary sequence criterion lumps the different categories of discontinuities in a single one, this way building a category with sufficient members compared with the MR class; even with this grouping, the discontinuity class is outnumbered by the MR class. The depth, dip, azimuth, and aperture (if applicable) of three types of discontinuities (i.e., CJ, SZ and COL) are only available for televiewer logs in Zinkgruvan. The drill response to inclined and closed discontinuities is not restricted to a single point associated with the depth of the center of the discontinuity (d_{cp}) but to the whole intersection length (d_b) with the borehole, to which a sequence of ones is assigned (Fig. 7):

$$d_b = d_{cp} \pm \frac{\phi_h}{2} \cdot \tan\alpha \tag{2}$$

where ϕ_h is the hole diameter and α is the angle of the normal of the discontinuity with the borehole axis (Fig. 7). A comparison between a section of a televiewer log with the corresponding binary sequence D_{obs}, in which zeros or massive rock are plotted in white

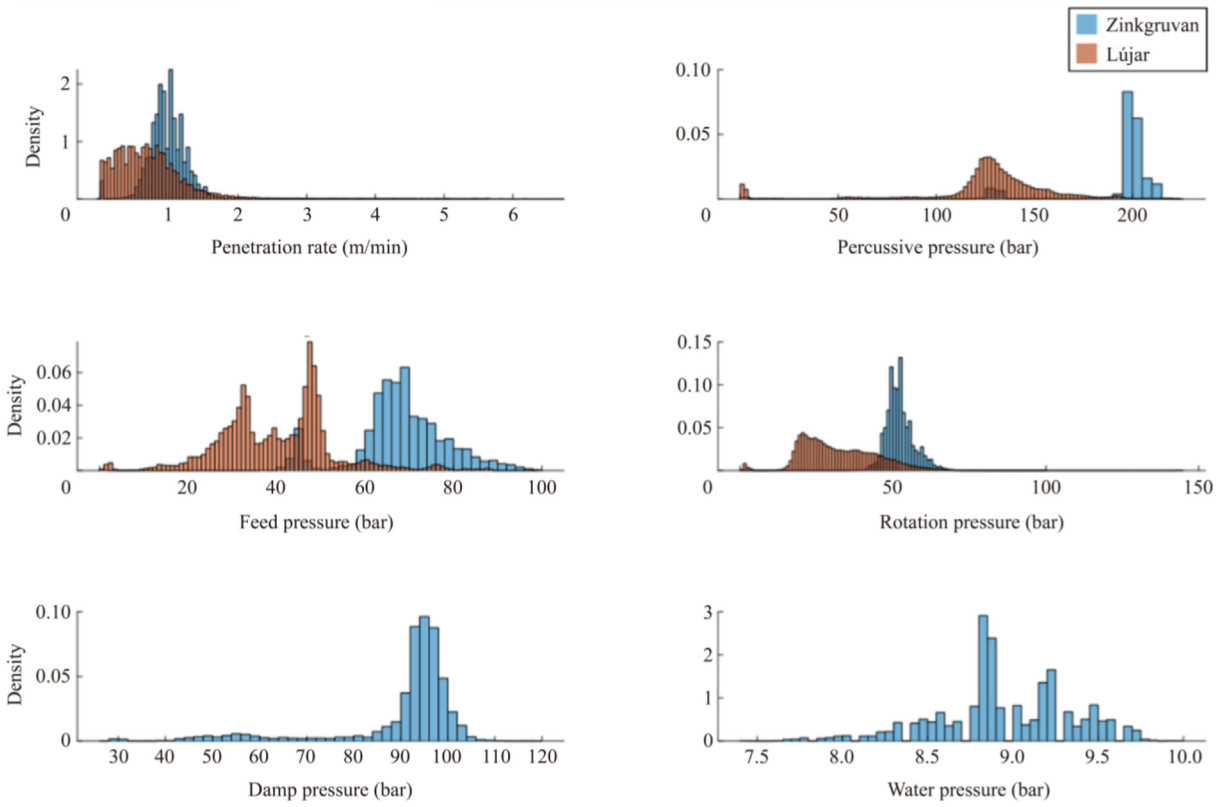

Fig. 6. Normalized histograms of unprocessed MWD data from Lújar and Zinkgruvan data sets. Note: The darker area in histograms corresponds to the overlapping between them.

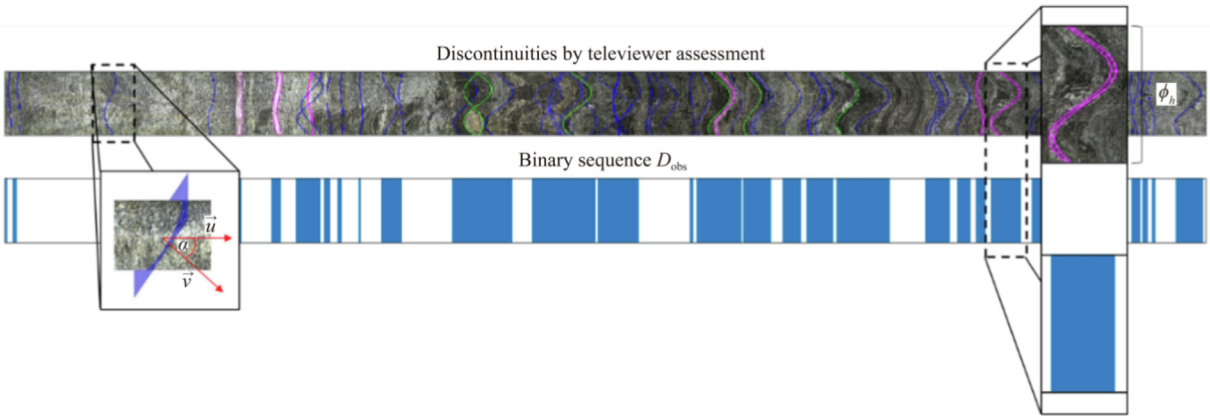

Fig. 7. Discontinuities by TV assessment for Hole 1, Ring 2 (R2H1). Note: Televiewer image up and binary sequence D_{obs} down, calculated with Eq. (2).

and ones or discontinuities in blue, is shown in Fig. 7. The zoom on the right outlines a shear zone (magenta trace) observed in the televiewer log and the corresponding fracture zone in the binary sequence.

The total number of ones (or samples treated as DISC) is 49% of the total samples logged with the MWD system in Zinkgruvan. Isolated and pseudo closed horizontal discontinuities have likely a limited influence in the response of the drill rig and are discarded for the analysis. To define a criterion to filter out these discontinuities, a comparison of the D_{obs} sequence with the MWD signals shows that discontinuities in D_{obs} that are separated (i.e., distance between the end and the beginning of the influence area of two adjacent discontinuities) at least by 20 cm and have a width (d_b) below 20 cm have a limited influence in the rig response. An example of this discontinuity type is the zoomed joint in the initial part (left zone) of the borehole in Fig. 7, with an influence width of 12 cm.

Eq. (2) cannot be applied to data from Lújar mines because the dip and azimuth of discontinuities cannot be measured from the endoscope videos. The size of the discontinuities in the binary sequence D_{obs} (or number of ones) is obtained by visual assessment of the discontinuity width (or the affected hole length) in the recorded videos, so that the effect of dip differences between discontinuities is not included in the resulting model. Minor discontinuities with negligible influence in the drill rig cannot be detected in the videos due to the lower resolution of the endoscope in comparison with the optical televiewer. This involves that less samples are classified as joints, just a 10% of the total samples, than mainly correspond to larger faults and cavities. As minor discontinuities are not marked in the videos, no filtering is required to remove them.

3.4. Classical approach: Structural factor

To formulate the best prediction model, a re-parametrization of the DI (Eq. (1)) that defines the optimal combination of drill parameters and their weight in the model may be done as follows:

$$\mathrm{DI}_i = \sqrt{\sum_{j=1}^{n} \beta_j \left(\frac{var_i(\mathrm{MWD}_j) - \overline{var_{\mathrm{MWD}_j}}}{std(var_{\mathrm{MWD}_j})} \right)^2} \quad \text{with } i = 1, 2, \ldots, L \tag{3}$$

where $var_i(\mathrm{MWD}_j)$ is the moving variance at the sample point i calculated through a sliding window of the j^{th} MWD parameter (calculated with MATLAB's instruction *movvar*) with a length of 14 samples (equivalent to a depth of 26 cm; note that the MWD resolution is constant at both mines); $\overline{var_{\mathrm{MWD}_j}}$ the mean of the moving variance for the j^{th} MWD parameter regarding the total length of the borehole; $std(var_{\mathrm{MWD}_j})$ the standard deviation of the moving variance for the j^{th} MWD parameter regarding the total length of the borehole; n the number of MWD parameters; β_j the weight of the j^{th} MWD parameter, determined as those yielding a higher fractures recognition; $\sum_{j=1}^{n} \beta_j = 1$. Eq. (1) is a particular case of Eq. (3) with n=2 and β_1=β_2=0.5; and L the number of records for each borehole.

Eq. (3) is applied to combinations of two, three and four (i.e., n=2, 3 or 4) MWD parameters. These are PR and three pressures (PP, DP, and RP) for Zinkgruvan data; WP and FP are discarded to simplify the analysis: WP is sensitive to water inflow, which is not observed for the discontinuities marked due to the absence of large or wide-open discontinuities [33]; FP is biased by the effect of the two-step cylinder that controls its level so as to prevent hole deviation. In Lújar, all four recorded parameters (PR, PP, FP, RP) are used.

After works by Schunnesson [20] and Ghosh [6], in which the major peaks in the DI signals could be related with the presence of discontinuities, a binary sequence (hereinafter D_{DI}) is created from the resulting DI. For this, values of the peaks over a threshold are defined as ones (discontinuity) and the rest as zeros (massive rock). This threshold T is taken as a percentile of the distribution of the DI for all boreholes monitored in each site. Percentiles are varied from 50 to 95 in steps of 5; the values providing a better recognition rate are selected.

To find the DI (i.e., the optimum combination of MWD parameters and their β_j weights, and the threshold percentile T, which leads to the best recognition of the discontinuities observed in the borehole walls), similarity measurements between D_{obs} and D_{DI} sequences are used [56]. This type of analysis has been used to assess binary data in different areas as in pattern recognition, information retrieval, statistical analysis, and data mining, and in specific for rock mechanics or related areas [57].

To account for the uncertainty in the hole logs (mainly in the exact position associated to a discontinuity) and for the fact that the response of the drill rig is not instantaneous, and it is also affected by conditions at previous depths, the cross-correlation between the binary sequences D_{obs} and D_{DI} is calculated considering offset lags of up to 10 positions. This allows a tolerance between both sequences, which enables a maximum gap in the position of one series over the other of 20 cm, according to the precision of the MWD system in both sites. For the relative position between both series where the maximum correlation is obtained, the number of samples matched between the binary sequences D_{DI} and D_{obs} for massive rock (MR_{SM}) and for discontinuity ($DISC_{SM}$) are calculated. From them, the 'simple matching' similarity criteria by Sokal and Michener [58] is applied:

$$S_T = \frac{MR_{SM} + DISC_{SM}}{TOT_S} \quad (4)$$

where TOT_S is the total number of samples in either D_{DI} or D_{obs}.

The similarity index in Eq. (4) handles the double-zero state (e.g., absence of a discontinuity or massive rock in both objects D_{obs} and D_{DI}) in the same way than the double-one scenario (e.g., presence of a discontinuity in both objects D_{obs} and D_{DI}), which avoids any bias in the performance of the DI for discontinuity recognition.

3.5. Machine learning approach

The definition of the binary sequence from in-hole images converts the recognition of the structural condition in each borehole into a classification problem, for which the most suitable classification algorithm must be selected. A scheme is defined for data set validation and hyper-parameters optimization, resulting in a hybrid ensemble model.

3.5.1. Input data and resampling process

The input parameters considered for ML are the length of the hole, all drilling parameters (after processing), and their variances. This makes up thirteen features for Zinkgruvan and nine features for Lújar. The model outputs or target values are the binary sequence determined from hole logs. As discussed in Section 2.2, it is important that the categories targeted (i.e., presence or absence of discontinuities) have the same probability of occurrence to train and test the model with an equitable quantity of data for each class [59]. This is not a problem for Zinkgruvan, as the number of samples classified as massive rock and as discontinuity is even (49% and 51%, respectively; see Section 3.3). For this site, samples from 14 holes are randomly split into training and testing sets with 7476 and 1869 samples (ratio 8:2), respectively. Measurements from one hole (R16 H3), selected randomly, are left out to validate the model.

In the case of Lújar, about 90% of the samples are classified as massive rock, representing an imbalanced class distribution. Under sampling of the original dataset could be an option to balance both classes. For this, only 10% of the samples (same number as DISC class) of MR could be used, which could bias the original data set that includes several areas of massive rock along the holes, limiting the generalization performance of the model for new data with these characteristics. As an alternative, a systematic re-sampling technique [60,61] is applied: The samples where a significant variation in the response of drill parameters may take place, are duplicated and added to the original set as new data set. This enables the model to identify the boundary between discontinuities and massive rock.

3.5.2. Classifier: Random Forest (RF)

RF is a classifier composed by an aggregation of tree-structured classifiers or classification and regression trees (CARTs), that are trained following a bagging process [62]. In the bagging method, the dataset is randomly sampled-with-replacement (bootstrapping) to form a group of subsets (bootstrap sets) of the same size, which are trained on different features, leading to the same number of different strong decision trees or CART [63]. The CART is based on nodes which represent a classification criterion, following a splitting process that is optimized to obtain the best goodness of fit, and ultimately deliver a prediction through the individual contribution of all the nodes [19]. During the process, some of the samples in the set are not considered for the training, called out-of-bag (OOB) samples, that are used to further validate the final prediction obtained as combined output, this being the class selected by the higher number of CARTs. A general structure of this method is illustrated in the ensemble classifier section of Fig. 8.

The performance of the RF algorithm depends on the hyper-parameters that control the characteristics of the trees and the bagging process. These are the number of learners (i.e., number of decision trees in the model), the maximum number of splits, (i.e., the depth of the decision tree or number of divisions, that controls the complexity of the ensemble), and the minimum observations per leaf (number of nodes at the end of the tree). Using the default values specified in suitable software packages works relatively well, but results can be improved by tuning them [64]. This optimization can be set by the user or automatically driven by tuning strategies, namely: Bayesian optimization (BO) [65], grid search [66] or random search [67].

3.5.3. Data structure and performance metrics

According to the characteristics of the data and the classification target, RF is the ensemble method selected (Section 2.2); other ensemble methods like the subspace k-nearest-neighbour (KNN) provide slightly worse results and are omitted here. After selecting the model to train with the input data, a validation scheme is chosen to estimate the predicting ability of the model trained. Considering the size of the dataset and the number of sub-datasets used in previous studies [68,69], a random k-fold cross-validation scheme with ten folds is applied. The proposed flowchart for the optimum ML classification model is described in Fig. 8.

The BO algorithm is used to tune the hyper-parameters (i.e., number of learners, maximum number of splits, and minimum observations per leaf) that minimize the cross-validation loss or error. Then, the general performance of the model is evaluated through the classification accuracy, calculated from the weighted average classification loss for the k-fold cross validation, after tuning the hyper-parameters with the optimization function. It is defined as:

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN) \cdot 100 \quad (5)$$

where TP is True Positives (i.e., number of samples of discontinuities correctly predicted); TN the True Negatives (i.e., number of samples of massive rock correctly predicted); FP the False Positives (i.e., number of samples in D_{DI} classified wrongly as discontinuities); and FN the False Negatives (i.e., number of samples in D_{DI} classified wrongly as massive rock). Note that Eq. (4) is the same to Eq. (5), but the last one does not account for the uncertainty in the position as was done by the similarity index in the binary sequences.

Other metrics used to evaluate the model performance for each class are presented in a confusion matrix:

(1) The True Positive (TPR or recall number) and False Negative Rates (FNR=1−TPR): They describe the proportion of correctly and incorrectly classified observations, respectively, per true class.
(2) The Positive Predicted Values (PPV or precision) and False Discovery Rate (FDR=1−PPV): They represent the proportion of correctly and incorrectly classified observations, respectively, per predicted class.
(3) The F1 score: It is a measure of test accuracy very useful when there is an imbalanced data problem to analyse the performance of the model for each category. It is calculated from the recall (TPR) and the precision (PPV):

$$\text{F1 score} = 2 \cdot (TPR \cdot PPV)/(TPR + PPV) \quad (6)$$

The last metric considered is the Area Under the Curve (AUC) that indicates the model performance to distinguish between positive (discontinuities) and negative (massive rock) classes; the higher AUC the better is the classification ability of the model. AUC is defined from the plot of the TPR versus the FPR showing

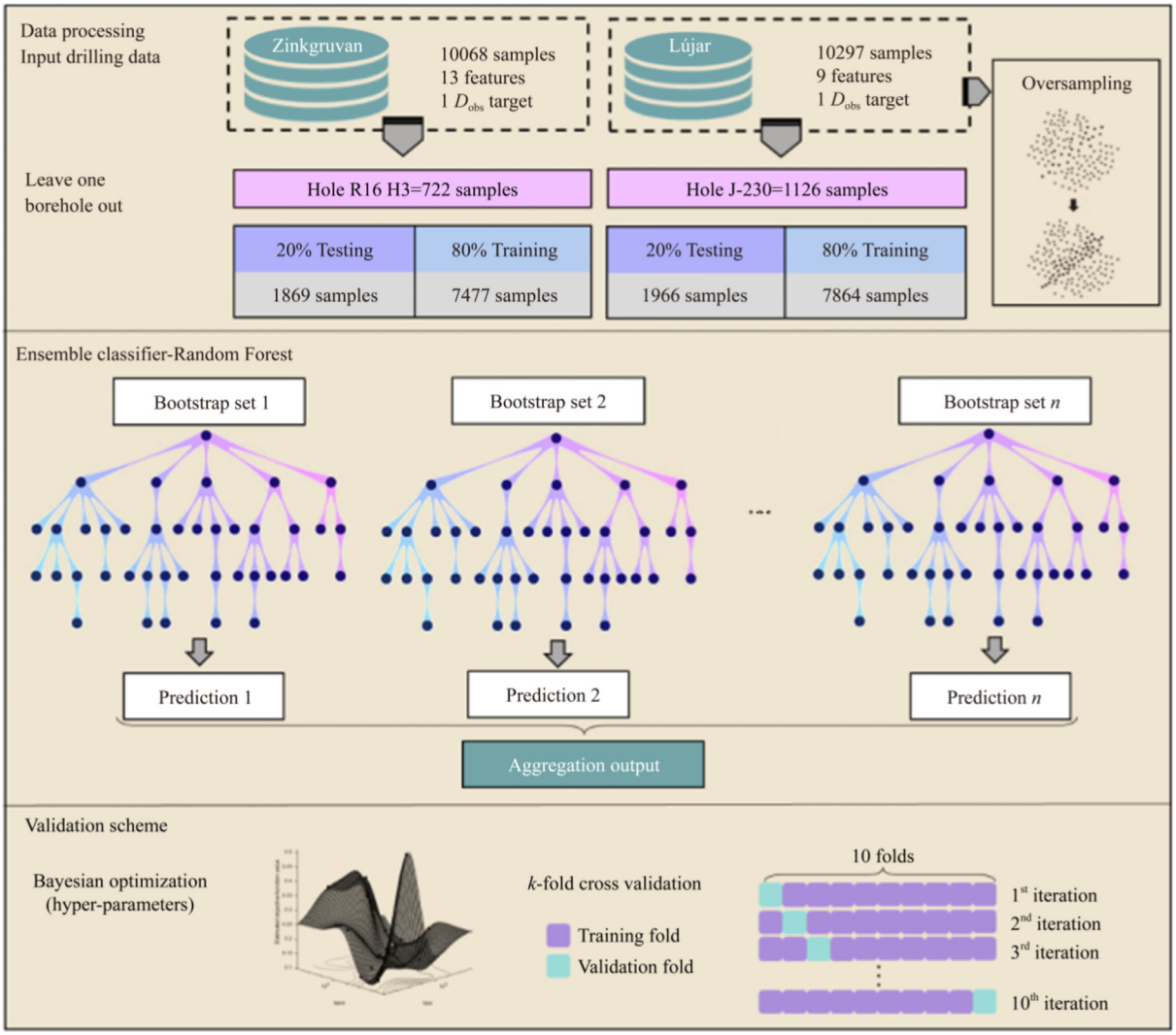

Fig. 8. Flowchart of the classification model used.

the performance of the classification model at different classification thresholds.

4. Results

4.1. Discontinuity index calibration

Table 1 shows the characteristics (i.e., drilling parameters, weights, and threshold) of the best similarity indices between the binary sequences D_{obs} and D_{DI} obtained through Eq. (3) for each mine. The similarity indices are sorted in descending order and correspond to the mean of the match accuracy from all boreholes monitored in Zinkgruvan (15) and Lújar (33). This ranges from 60.5% to 69.3% in Zinkgruvan and from 90.0% to 91.9% in Lújar.

The characteristics of the DI are different in each site. For Zinkgruvan, the best recognition, 69% on average, is obtained with the combination of the variabilities of PR and RP weighted by 60% and 40% respectively, and a threshold of 65%, equivalent to a DI of 0.5. This result is in line with the fracturing parameter defined by Ghosh et al. [6] and Navarro et al. [7,12], but with a slightly higher contribution of PR. Slightly smaller similarity indices are obtained when PP is added to the previous parameters (PR and RP), and when two parameters DP and RP are combined; despite that the two-step feed cylinder tries to keep the DP constant, this parameter is included in seven of the combinations shown in Table 1. In general, the best results are obtained when RP is considered, while the lower accuracies are obtained when it is removed.

In Lújar, the DI recognizes in the best case near 92% of the discontinuities. This corresponds to the combination of the variability of FP and PR with weights of 60% and 40%, respectively. The larger weight in FP with respect to the other parameter may be explained by the effect of the cavities and faults in the advance of the drill bit. Similar indices, above 91%, are obtained when variability in RP is combined with variability in FP and/or PR, individually or both together. The inclusion of the variability in PP also leads to good similarity indices above 90% probably related with the existence of large discontinuities causing oscillations in all the parameters when crossed by the drilling. Despite of the limitations of the control system of the drill rig, this does not hide the response of the drill rig to discontinuities.

To graphically show the recognition capacity of the DIs with the highest mean similarity index, the binary sequences D_{obs} and D_{DI} for the holes with the worst and best results are plotted in Figs. 9 and 10 for Zinkgruvan and Lújar, respectively; the televiewer logs are also shown for visual comparison in Zinkgruvan. Here, the DI provides a similarity index between 60% (hole R3H1, worst case) to 79% (hole R1H3, best case). The visual correlation of discontinuities observed in the borehole walls and those identified by the DI

Table 1
Summary results for the definition of the discontinuity index.

Drilling parameters (MWD_j)	Weights (β_j)	Threshold T (DI_T value)	Similarity index (S_T, %)
Zinkgruvan			
PR, RP	**60–40**	**65 (0.5)**	**69.3**
PR, RP, PP	70–20–10	80 (15.9)	67.7
RP, DP	90–10	70 (0.6)	67.3
RP, PP	90–10	80 (1.0)	66.9
PR, DP, RP	50–30–20	75 (5.3)	66.8
PP, PR, DP, RP	50–30–10–10	75 (4.7)	65.7
PR, DP	90–10	70 (0.6)	65.6
RP, PP, DP	80–10–10	90 (11.8)	64.5
PR, PP	90–10	85 (0.9)	64.0
DP, PP	90–10	90 (1.4)	61.4
PR, PP, DP	80–10–10	95 (1,276.0)	60.5
Lújar			
FP, PR	**60–40**	**95 (2.5)**	**91.9**
PR, FP, RP	40–40–20	95 (19.3)	91.5
RP, PR	60–40	95 (2.6)	91.5
FP, RP	80–20	95 (2.3)	91.3
PP, PR	70–30	95 (2.4)	91.3
FP, PP	80–20	95 (2.2)	91.1
PR, RP, PP	60–30–10	95 (16.2)	91.0
RP, PP	80–20	95 (2.1)	90.8
PP, FP, RP, PR	50–30–10–10	95 (147.9)	90.0
PP, RP, FP	70–20–10	95 (43,170.0)	90.0
FP, PP, PR	80–10–10	95 (61.2)	90.0

shows that almost all the zones with grouped discontinuities are identified in borehole R1H3, and that massive rock zones at the beginning and at the end of the hole are correctly recognized; mismatching is mainly related with differences between the extension of the discontinuities calculated from in-hole logs (probably the effect of the dip is smaller than Eq. (2) shows) and the response in drill parameters (records with DI values over the threshold). For borehole R3H1, several discontinuities and massive rock zones are poorly discriminated; however, in general the DI identifies numerous discontinuities for this hole that are in line with the televiewer assessment. Considering the excellent quality of the rock in this mine, changes in the mechanical response of the rock to the drilling are more apparent when discontinuities are grouped, for which the DI model provides a better recognition. Conversely, recognition of several isolated discontinuities is difficult because each one of them may not individually trigger a noticeable response in the drill rig. This is in line with the approach followed by other authors in which zones from massive to highly fractured zones are the goal.

For Lújar, the best DI provides a recognition of 80% (hole B7-H19) to 99% (hole B5-H11). For hole B7-H19, the DI recognizes the CAV at 1.6 m and one OJ at 2.2 m, while the other four OJ at 2.1, 2.7, 3.1, and 3.5 m are incorrectly assigned as massive rock. Probably the OJ discontinuities misclassified by the DI did not cause a significant mechanical response of the drill. For hole B5-H11, the unique discontinuity observed from the endoscope videos (a fault at the end of the hole) has been correctly identified, and zones of massive rock are correctly estimated. A false positive (i.e., narrow discontinuity defined by the DI value over the threshold, which does not appear in the video) is predicted at 3.3 m.

To further analyse the performance of the best DIs (see rows marked in bold in Table 1) in each site, the fractions of recognized and non-recognized discontinuities (blue and orange bars, respectively) of each type is shown in Fig. 11. The analysis is performed comparing each record in the same position of the D_{DI} and D_{obs} sequences as defined by the optimum offset lag (Section 3.4). The labels over each bar are the similarity index for each class (i.e., structures from DI that match those from optical inspection). In both operations, massive rock (MR) is the class with higher recognition accuracy. This indicates a good recognition of zones with limited variability or noise in the drilling signals and the absence of discontinuities.

In general, discontinuities are better identified in Zinkgruvan than in Lújar. For Zinkgruvan, SZ and CJ are similarly detected (the similarity index is near 45%), while the identification of COL is poor. Such result may be explained by the fact that shear zones and joints cause a larger variability in PR and RP than when there is a change in the lithology as this does not involve a significant dif-

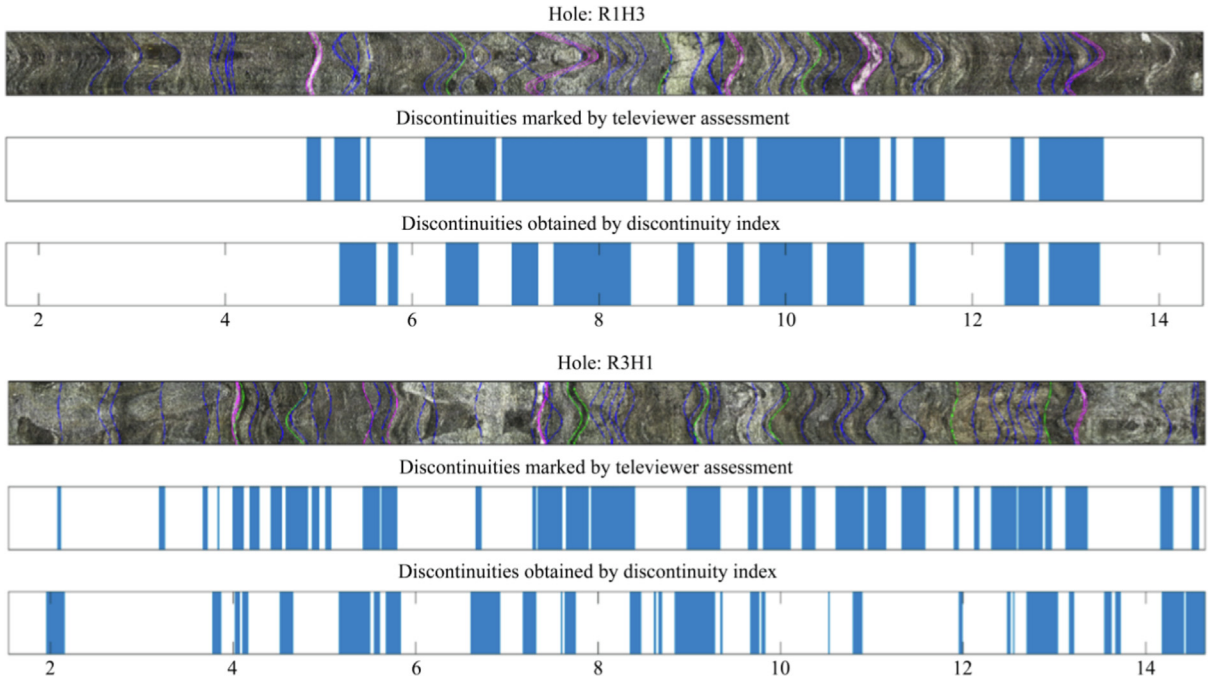

Fig. 9. Televiewer log and binary sequences D_{obs} and D_{DI} for holes R1H3 (S_T=79%) and R3H1 (S_T=60%) of Zinkgruvan. Note: Parameters marked in bold in Table 1 are used to calculate DI.

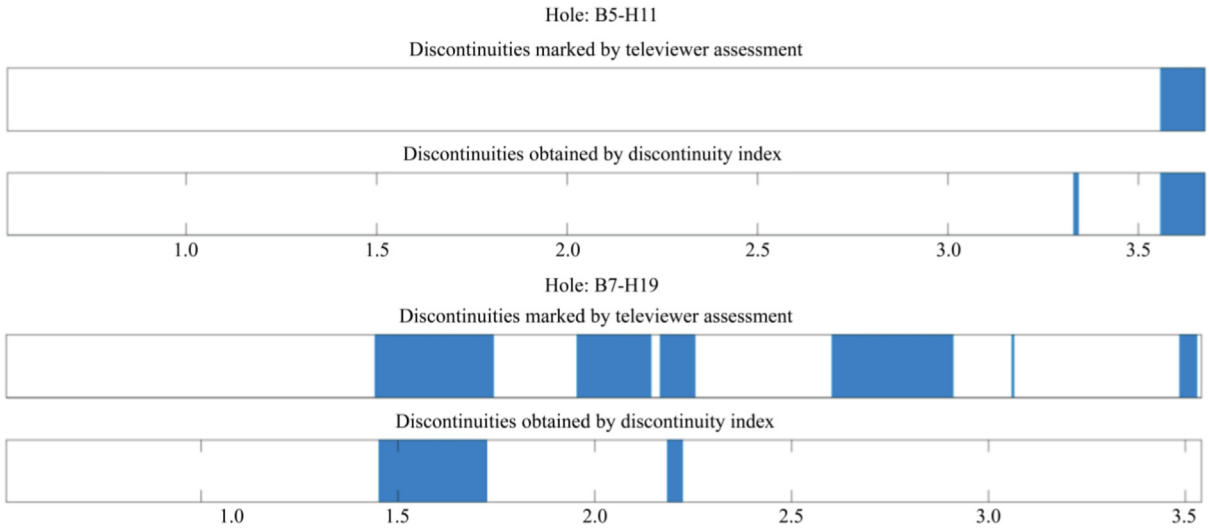

Fig. 10. Televiewer log and binary sequences D_{obs} and D_{DI} for holes B5-H11 (S_T=99%) and B7-H19 (S_T=80%) from Lújar. Note: Parameters marked in bold in Table 1 are used to calculate DI.

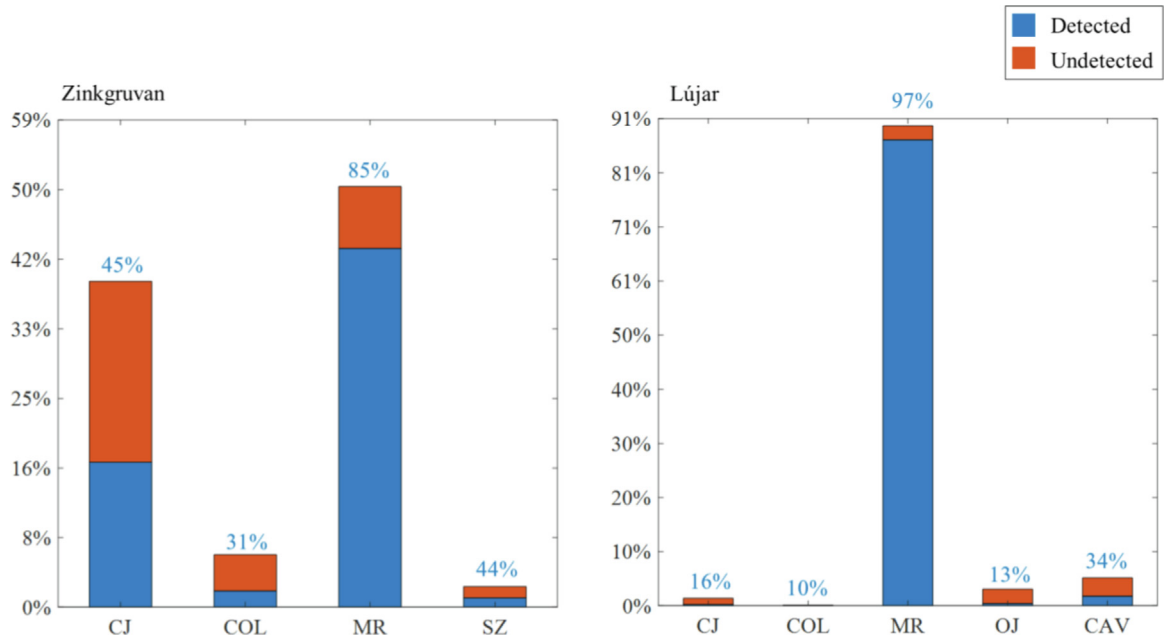

Fig. 11. Frequency of the different structural characteristics. Note: The labels above each bar are the similarity indices for the recognition of each type of structure from the DI, left, based on the variability of PR and RP in Zinkgruvan; right, based on the variability of FP and PR in Lújar. (See the rows in bold in Table 1 for more details about the DI).

ference in the strength of the corresponding rocks. The poor identification of discontinuities and the similar number of samples of DISC and MR in the sequence D_{obs}, leads to a moderate overall accuracy, up to 70%.

For Lújar, CAV are the discontinuities better recognized. Despite that the similarity index for this class is low, most of the cavities (75%) are detected but their width is narrower and/or they are slightly shifted compared with D_{obs}, which reduce the recognition percentage when the discontinuities are compared in each record. The other discontinuity types, CJ, OJ and COL do not lead to significant variations in the drilling signals and are not well recognized, resulting in low similarity indices. The limited influence of these discontinuities in the drill rig response will likely mean a minor effect blasting-wise, and they could probably be treated as massive rock in terms of blast design and explosive charging patterns. The smaller number of samples of discontinuities in comparison with samples of massive rock in D_{obs}, involves a higher overall similarity in comparison with Zinkgruvan.

4.2. ML classification

Table 2 summarizes the main performance metrics of the resampling process and hybrid ensemble algorithm BO-RF developed for both mines. The validation accuracy is similar for the training dataset and the testing one in both sites, which shows the consistency of the model when applied to unseen data. The prediction ability of the model in Lújar is high, near 96%, while for Zinkgruvan is near 90%. These results correspond to the best combination of hyper-parameters, defined by the lower cross-validation loss obtained from the BO technique applied. The value

Table 2
Summary of RF model result for the classification of discontinuities.

Site	Validation accuracy		AUC	Hyper-parameters		
	Training (%)	Testing (%)		Number of learners	Max number of splits	Min observations per leaf
Zinkgruvan	88.7	88.6	0.95	235	3471	1
Lújar	96.3	96.9	0.97	27	6751	1

Note: The hyper-parameters correspond to the optimal combination from thirty iterations for BO technique.

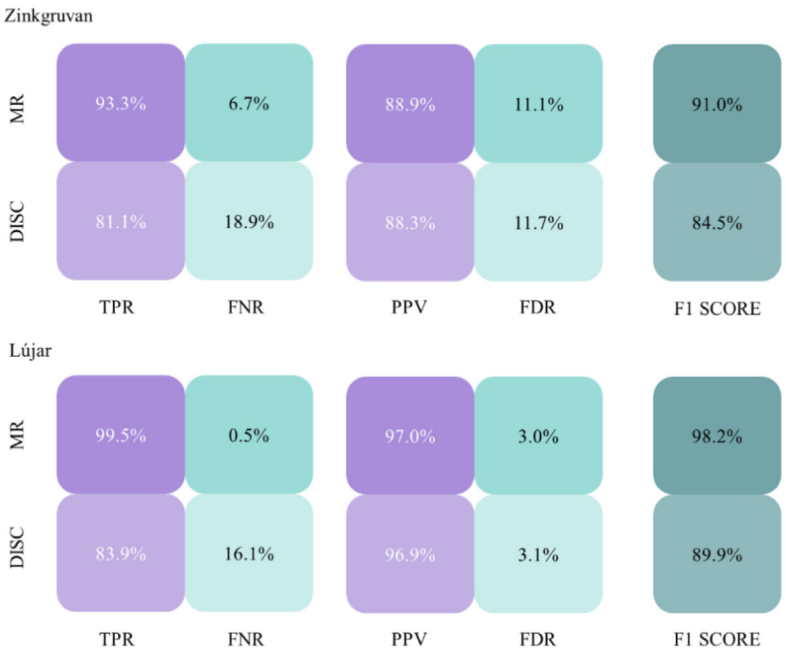

Fig. 12. Confusion matrices for testing results using BO-RF classifier for MR and DISC categories.

of the AUC is high in both sites, indicating that the models distinguish properly between massive rock and discontinuities. The resulting classification models are composed by a large number of complex decision trees (see the large number of learners and splits in Table 2) indicating complex relations between the drilling parameters and their variations.

The confusion matrices obtained for the testing set is presented in Fig. 12 for both mines. The high F1 score for both mines, above 85%, indicates that the chosen ML model handles both classes (DISC and MR) properly. Specially in Lújar, the precision values (PPV and FDR) are excellent for both classes indicating a trustable classification. The massive rock class is nearly completely recognized with a high recall (TPR) and F1 scores (near 100%), while the discontinuity class presents a slightly lower recall value (near 84%), probably affected by structures that trigger a similar response in drill parameters than massive rock samples. In Zinkgruvan, the model performance is slightly worse than in Lújar; FDR value is higher, about 11% for both classes, but MR is still well

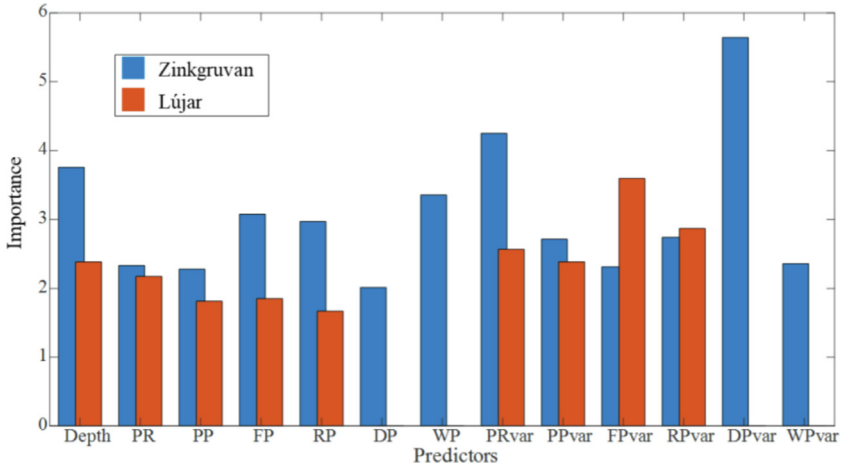

Fig. 13. Predictor's importance graph for RF model in Zinkgruvan and Lújar mines.

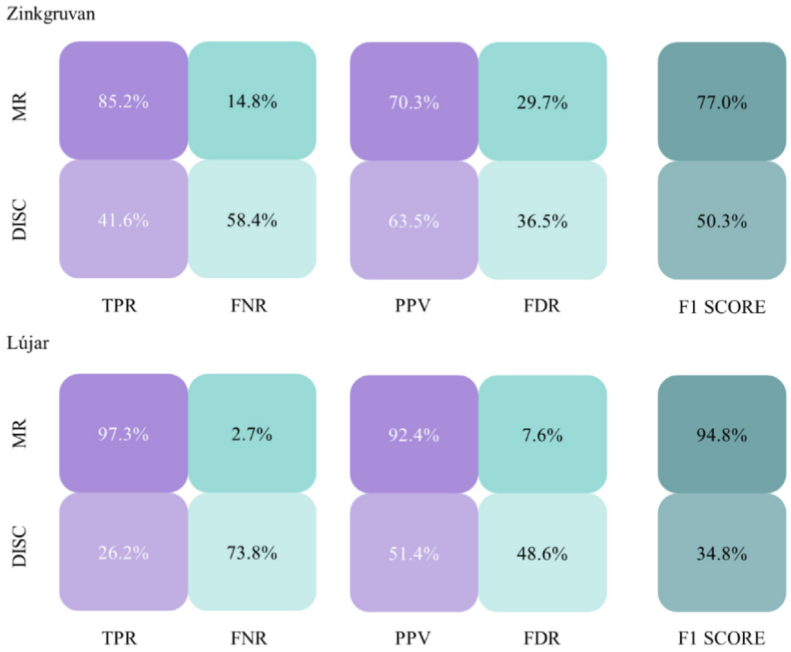

Fig. 14. Confusion matrix for MR and DISC categories from the DI built from variability of PR and RP in Zinkgruvan (top) and from variability of FP and PR in Lújar (bottom). Note: The same drillholes used to train and test the ML model are considered.

recognized. In both mines, the classification ability of discontinuities has improved significantly with respect to the classical approach.

The relative contribution of each parameter (predictor) in the classification model (i.e., how relevant is an input parameter to predict rock mass conditions) is represented in Fig. 13. It shows the weighted average importance from each tree learner of the ensemble; higher values indicate a larger influence of an input parameter in the predictions. In Zinkgruvan (blue bars), variations in DP and PR present the higher importance for the model. The last parameter was included in the DI with the highest similarity index (Table 1), while oscillations in DP were probably not considered in this DI due to the effect of the two-step feed cylinder. The complex relations defined by the BO-RF algorithm overcome this spurious effect, making DP variation the most important feature for the class prediction. Among direct drill parameters in both mines, the measured depth presents the higher importance, which reflects that the ML model considers the process as a regionalized phenomenon within a continuous medium; this means that is more probable to find the same class for close (i.e., same depth) records. In addition, WP presents the second higher importance suggesting probably that some of the structures marked in televiewer logs have an aperture that is not identified in the images, but that affects the pressure of the water flush. For Lújar (orange bars), FP variation has the higher importance, in agreement with the results from the classical DI approach. Also, variations in the other three parameters (RP, PR, and PP) present the next higher importance, supporting the high similarity indices obtained when these features are considered in the DI. PR is the direct drilling parameter with higher relevance, in line with the high rates observed when a cavity or a fault are crossed by a drillhole.

5. Comparative analysis and discussion

A calibration procedure of the DI has been followed aiming to maximize the similarity index between the binary sequences from observation of discontinuities with optical systems, and from the calculations with drilling data. For both mines data, the variation of the parameters related with the response in the rock, as PR and RP, explain well the presence of discontinuities despite the differences in the rock conditions and drilling rigs characteristics. RP probably accounts for the presence of small discontinuities which are the prevalent ones observed in Zinkgruvan. In Lújar, variations in the FP are also included in the DI with best performance. This parameter is also pointed as the one with largest importance in the predictions from ML model (Fig. 13). This result may be explained by the presence of large cavities and faults, and the automatization level of the RCS, which does not control properly the RP or the PR through the thrust (FP). That is why this pressure, commonly a rig control parameter, oscillates with the presence of large discontinuities.

DP variation in Zinkgruvan, not considered in the DI due to the spurious fluctuations caused by the two-step feed cylinder, is the most important parameter in the BO-RF ensemble model. The trained BO-RF algorithm seems to be able to filter this effect and include this parameter in the prediction of discontinuities. This result would suggest developing a filtering routine to include this parameter in the classical model.

The better performance of the ML approach may be due at least partially to the use of not only variances of drill parameters but also to the values of the parameters themselves. The presence of a discontinuity may result in a progressive increase or decrease of the values, making it relevant to include also the drilling parameters. However, both approaches point out the effect of the discontinuities on the variability of the recorded signals. To compare the predictions obtained from ML classification model with the classical approach, the confusion matrix is calculated for the prediction classes obtained with the best DI in both mines (see Fig. 14 and parameters marked in bold in Table 1).

In line with Fig. 12, the classical approach recognizes significantly better massive rock than discontinuities in both sites, resulting in a very good precision (PPV) and prediction (TPR) in Lújar mine for massive rock, where the corresponding F1 score is near 95%. The performance in Zinkgruvan is worse, as with the DI, and a lower F1 score for massive rock (77%) is obtained. The perfor-

mance of the DI in Lújar to predict discontinuities is, on the contrary, poor (F1 value close to 35%), while in Zinkgruvan it is higher and more balanced between the two classes. A comparison of Figs. 12 and 14 shows that the classification ability of the DI is worse than ML. This different performance is most likely related to the use of more inputs and complex relations in ML than in the DI (that essentially comprises a linear combination of variations of drilling parameters).

To further compare the performance of the BO-RF classification model versus the classical approach, the similarity indices between D_{obs} and the binary sequence of the holes left out for validation (R16H3 in Zinkgruvan and J-230 in Lújar) have been calculated for both BO-RF and DI models. The discontinuity maps from optical logging and those calculated from BO-RF and DI for Zinkgruvan's hole R16H3 are plotted in Fig. 15. The classification accuracy is similar for both approaches; the similarity index is 79.6% (F1 score is 85.4% for MR and 65.1% for DISC) for the ML model compared with 71.3% (F1 score is 80.9% for MR and 34.4% for DISC) for the DI calculated with the parameters in bold in Table 1. However, ML recognizes better the discontinuities (the F1 score is almost twice) than the DI. The ML model identifies more discontinuities than the classical model but is not precise in their position and extension; the ML results are, however, better in overall terms.

The resulting binary sequences for hole J-230 in Lújar are plotted in Fig. 16; images from the most significant discontinuities observed by optical inspection are shown in the top plot. The results for both techniques are very similar; similarity index for ML is 92.2% (F1 score is 95.8% for MR and 21.1% for DISC), while it is 92.6% (F1 score is 95.9% for MR and 39.2% for DISC) for the classical model. The binary sequences obtained with both models recognize the two large cavities at 7.1 and 12.7 m. The void located at about 2 m could correspond to a weathering of the walls of the hole by the effect of the drilling, indicating a soft rock without a strong fluctuation in drilling parameters. Conversely, the open joint at 10.7 m is clearly a structural condition that is not detected by any of the models. In the case of the DI analysis, this could be attributed to small fluctuations on the signal when this discontinuity is crossed compared with the two cavities identified, while for the ML analysis, it could be due to the small number of examples of the open joint category, that prevents the model from learning this class correctly.

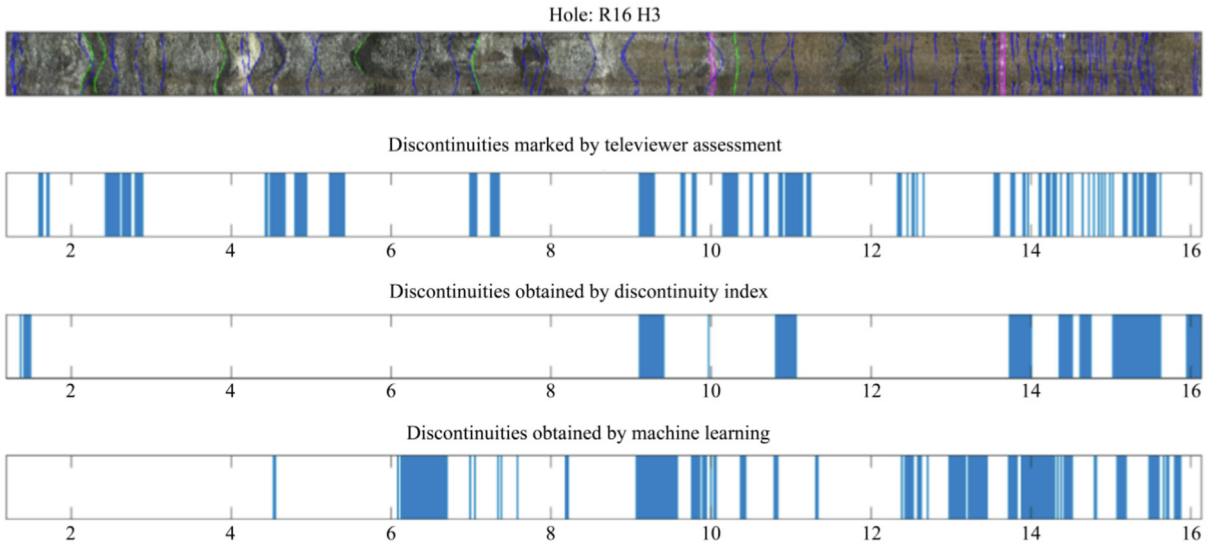

Fig. 15. Televiewer log and binary vectors for TV, DI and ML analysis for testing data in hole R16 H3, Zinkgruvan.

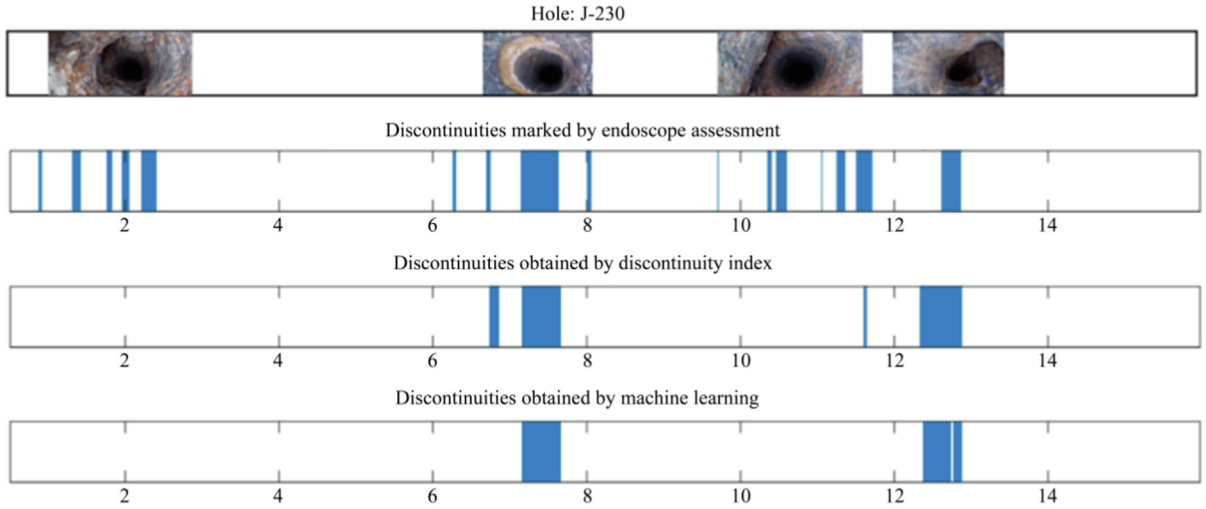

Fig. 16. Endoscope images and binary vectors for optical inspection, DI and ML analysis for testing data in hole J-230, Lújar.

Both DI and ML approaches provide better results for Lújar mine than for Zinkgruvan despite that the control unit of the drill rig in Lújar is very limited. Such differences may reflect a smaller effect of inclined discontinuities in the MWD response than what is considered by Eq. (2), that involves a wide zone of discontinuity influence in the D_{obs} sequence, that is used to calibrate both DI and machine learning. Overall, the classical analysis is a simple calculation with which it is possible to obtain a reasonable recognition accuracy of the discontinuities and massive rock classes in both operations, while the ML analysis shows better prediction results, at the cost of using more parameters and complex relationships.

Regarding the performance of the model, the quality of the MWD geological accompanying data is key for a precise identification of each discontinuity class and its position along the hole. A significant limitation of this task is the difficulty to differentiate structural discontinuities from voids and fractured zones worn by the action of the drill itself. A wrong classification of some category or discontinuity feature, like aperture, affects the selection of parameters and their weights in the classical approach, and the model performance in the ML approach is reduced if trained with quite different values for the same response. This occurs in ML model in Zinkgruvan in which WP has a relevant importance in the classification ability though no visual open discontinuities are apparent from televiewer logs.

The results obtained for both models are probably as good as can be since accuracy in the testing process is limited by: (1) The difficulty in the recognition of discontinuities for which the drill may not be sensitive, as the structural condition of each discontinuity is often barely assessed by the visual evaluation of the optical logs, so that discontinuities that may not cause a mechanical response of the drill (i.e., a response similar to massive rock class, a small variation or noise in the signal) could be marked; (2) Errors in the exact position and extension of the discontinuities: Despite the high-resolution image of the televiewer, an offsetting error in the collar could shift all the marked discontinuities from their real position. For the endoscope, the encoder resolution is 10 cm, that is five times the MWD resolution, introducing bias in the exact extension and position of the discontinuities marked; this includes the unknown effect in the response of the drill rig between inclined and pseudo-horizontal discontinuities; (3) Different discontinuity classes are grouped in only one category in the binary approach, so different magnitudes of response shown in the signals may confuse the recognition in the model training.

These limitations could be eased with: (1) A broader dataset to define more rock conditions, certainly enabling an increase of the similarity index and the model accuracy; (2) Marking only major structural changes i.e., cavities, faults or highly fractured zones, for which drill parameters may be sensitive; those discontinuity types would likely be the ones with e.g., a major role in rock breakage by blasting; (3) Studying the performance of the techniques proposed in a multiclassification scenario, for which it is necessary to have a reasonable quantity of samples from each category assessed.

The main advantage of the methodology proposed is the adaptive structuring of the classification model, regarding the drilling parameters available on site, and the training process based on the mine's own rock mass characteristics.

The testing of the methodology presented in two operational environments with very different conditions of rock mass quality and drilling technology used, added to the fact that imbalanced sets are a prevalent scenario for rock mass classification, support the use of this approach as engineering practice for any underground mining operation. For the application of the hybrid ensemble ML model and the resampling technique, a sufficient database of geological and drilling data is required, that must be customized to the appropriate data formats.

The application of the solution proposed in a continuous excavation cycle requires an initial commissioning period to consolidate a comprehensive database, comprising drill rig and geotechnical features of the operation. When a continuous implementation of this methodology is planned, it is relevant to consider enlarging the geological data as the rock mass characteristics changes or a different drilling unit is used in the operation. With this calibration and validation steps, the rock structural model generated is very likely to present a good predictive accuracy and high confidence in their results, as shown in this research. The validated virtual core obtained from the MWD parameters response, in which the structures that occur naturally in the rock are recognized, can be used e.g., to estimate rock quality indices used in blast design and/or guiding blast charging adaptations.

6. Conclusions

Drilling data from different drill rigs (with different automation levels and control features), and borehole logs made with different rock structural mapping techniques (logging with optical televiewer and with digital endoscopes) were gathered in two underground mines with different rock mass characteristics. The following findings and conclusions can be drawn:

(1) A comprehensive methodology for discontinuity recognition from MWD data is developed and validated for implementation in an underground mining operational environment. The recognition of discontinuities has been quantified by introducing classification performance metrics that compares their existence at a certain location in a borehole as from the drilling parameters and from in-borehole televiewer or video footage inspection.

(2) Both a classical linear combination and ML techniques have been applied to MWD data to build predictive models of discontinuities, obtaining high classification accuracies. The selection of drill parameters or features to be included in the analysis is crucial for a successful result, the optimal combination being site specific, dependent on the rock mass properties and the control system of the drilling equipment.

(3) A hybrid ensemble Bayesian optimization technique to optimize a Random Forest model, BO-RF, is validated as an ideal model to achieve the classification of rock discontinuities despite the existing imbalance of categories in the datasets used. The resampling technique, the hyper-parameters optimization algorithm, and the 10-fold cross-validation, are proven as a sound combination of techniques to apply in rock type classification from drilling parameters. Given that imbalanced distributions are a prevalent scenario for the rock structural classification problem, the application of the ML technique proposed is generalizable for any underground mining operation, provided that a sufficient database of geological and drilling data exists.

(4) The predictive capacity of the models is quantified by introducing classification performance metrics, adapted for both approaches. This enables a reliable assessment of the results, rather than the classical qualitative assessments based on visual comparison between actual and predicted rock mass conditions. The prediction accuracy is rated for each class included in the model, which allows it to detect the classes poorly predicted so as to increase the number of samples in the database.

(5) Using the ML approach, the general classification accuracy from the resulting site-specific models is in excess of 90% for both mines, using nine direct or derived drill parameters for Lújar and thirteen for Zinkgruvan. Consistent results are

obtained when models are validated with external data, with overall recognition accuracies about 80% in Zinkgruvan and 90% in Lújar.

(6) For the classical approach, in Zinkgruvan the DI that provides better recognition results is a combination of PR and RP variability, while in Lújar, the best DI is obtained by a combination of FP and PR variabilities. Applying these combinations, the resulting similarity index is over 90% in Lújar, where large discontinuities, such as voids and faults, are predominant. In Zinkgruvan, where only small discontinuities are observed, the prediction is worse, in the range 70%–80%.

(7) The application of ML yields a higher classification accuracy than the DI through the use of more drilling parameters and more complex and unknown relations between them. However, the simplicity of the classical model makes it a practical alternative to predict the structural condition along the borehole. In general, the ML model defines drilling variations as most useful predictors for structural conditions, much like the classical model, fully based on parameters variation.

The results from this study encourage the application of a similar methodology to generate a site-dependent rock strength condition model and to quantify its performance. It should be noted that a good discrimination of the discontinuity types that cause a mechanical response in the drill bit as it passes through them is important in order to achieve good recognition algorithms. The bias in the in-hole logging process inevitably affects the validation of the model performance, and consequently a more precise logging method, non-operation-disruptive, is worthy of being studied.

Acknowledgments

This work has been conducted under the illuMINEation project, funded by the European Union's Horizon 2020 research and innovation program under grant agreement (No. 869379). The authors would like to thank EPIROC, Lundin Mining and Minera de Órgiva for providing data and facilitating the work in their premises, and for allowing its publication. The work of Enming Li is supported by the China Scholarship Council (No. 202006370006).

References

[1] Rai P, Schunnesson H, Lindqvist PA, Kumar U. Measurement-while-drilling technique and its scope in design and prediction of rock blasting. Int J Min Sci Technol 2016;26(4):711–9.
[2] van Eldert J, Funehag J, Saiang D, Schunnesson H. Rock support prediction based on measurement while drilling technology. Bull Eng Geol Environ 2021;80(2):1449–65.
[3] Desbrandes R, Clayton R. Measurement while drilling. In: Developments in Petroleum Science. Amsterdam: Elsevier; 1994:251–279.
[4] Wang HT, He MM, Zhang ZQ, Zhu JW. Determination of the constant mi in the Hoek-Brown criterion of rock based on drilling parameters. Int J Min Sci Technol 2022;32(4):747–59.
[5] He MM, Li N, Zhang ZQ, Yao XC, Chen YS, Zhu CH. An empirical method for determining the mechanical properties of jointed rock mass using drilling energy. Int J Rock Mech Min Sci 2019;116:64–74.
[6] Ghosh R, Gustafson A, Schunnesson H. Development of a geological model for chargeability assessment of borehole using drill monitoring technique. Int J Rock Mech Min Sci 2018;109:9–18.
[7] Navarro J, Seidl T, Hartlieb P, Sanchidrián JA, Segarra P, Couceiro P, Schimek P, Godoy C. Blastability and ore grade assessment from drill monitoring for open pit applications. Rock Mech Rock Eng 2021;54(6):3209–28.
[8] SLIM project. Sustainable low impact mining solution for exploitation of small mineral deposits based on advanced rock blasting and environmental technologies. Funding from the European Union's Horizon 2020. Grant agreement No. 730294. Coordinated by Universidad Politécnica de Madrid. 2016-2020. DOI: 10.3030/730294.
[9] IlluMINEation Project. IlluMINEation—Bright concepts for a safe and sustainable digital mining future. Funding from the European Union's Horizon 2020. Grant agreement No. 869379. Coordinated by Montanuniversitaet Leoben. 2020-2023. DOI: 10.3030/869379.
[10] DIGIECOQUARRY project. Innovative digital sustainable aggregates systems. Funding from the European Union's Horizon 2020. Grant agreement No. 101003750. Coordinated by Asociación Nacional De Empresarios Fabricantes De Áridos. 2021-2025. DOI: 10.3030/101003750.
[11] Isheyskiy V, Sanchidrián JA. Prospects of applying MWD technology for quality management of drilling and blasting operations at mining enterprises. Minerals 2020;10(10):925.
[12] Navarro J, Schunnesson H, Ghosh R, Segarra P, Johansson D, Sanchidrián JÁ. Application of drill-monitoring for chargeability assessment in sublevel caving. Int J Rock Mech Min Sci 2019;119:180–92.
[13] Khanal M, Qin J, Shen BT, Dlamini B. Preliminary investigation into measurement while drilling as a means to characterize the coal mine roof. Resources 2020;9(2):10.
[14] Park J, Kim K. Use of drilling performance to improve rock-breakage efficiencies: A part of mine-to-mill optimization studies in a hard-rock mine. Int J Min Sci Technol 2020;30(2):179–88.
[15] Liu WP, Rostami J, Keller E. Application of new void detection algorithm for analysis of feed pressure and rotation pressure of roof bolters. Int J Min Sci Technol 2017;27(1):77–81.
[16] Peng SS, Tang D, Sasaoka T, Luo Y, Finfinger G, Wilson G. A method for quantitative void/fracture detection and estimation of rock strength for underground mine roof. In: Proceedings of the 24th International Conference on Ground Control in Mining. Morgantown: ICGCM; 2005, p.187–95.
[17] Galende-Hernández M, Menéndez M, Fuente MJ, Sainz-Palmero GI. Monitor-while-drilling-based estimation of rock mass rating with computational intelligence: The case of tunnel excavation front. Autom Constr 2018;93:325–38.
[18] Liaghat S, Gustafson A, Johansson D, Schunnesson H. Ore grade prediction using informative features of MWD data. Mining Goes Digita 2019:226–34.
[19] Chen JY, Huang HW, Cohn AG, Zhang DM, Zhou ML. Machine learning-based classification of rock discontinuity trace: SMOTE oversampling integrated with GBT ensemble learning. Int J Min Sci Technol 2022;32(2):309–22.
[20] Schunnesson H. RQD predictions based on drill performance parameters. Tunn Undergr Space Technol 1996;11(3):345–51.
[21] Navarro J, Sanchidrián JA, Segarra P, Castedo R, Costamagna E, López LM. Detection of potential overbreak zones in tunnel blasting from MWD data. Tunn Undergr Space Technol 2018;82:504–16.
[22] Navarro J, Sanchidrian JA, Segarra P, Castedo R, Paredes C, Lopez LM. On the mutual relations of drill monitoring variables and the drill control system in tunneling operations. Tunn Undergr Space Technol 2018;72:294–304.
[23] Navarro J, Segarra P, Sanchidrián JA, Castedo R, Pérez-Fortes AP, Natale M, López LM. Application of an in-house MWD system for quarry blasting. In: Proceedings of the 12th International Symposium on Rock Fragmentation by Blasting – Fragblast 12, Luleå, Sweden, 11-13 June 2018; p. 203–7.
[24] Scoble MJ, Peck J, Hendricks C. Correlation between rotary drill performance parameters and borehole geophysical logging. Min Sci Technol 1989;8(3):301–12.
[25] Manzoor S, Liaghat S, Gustafson A, Johansson D, Schunnesson H. Establishing relationships between structural data from close-range terrestrial digital photogrammetry and measurement while drilling data. Eng Geol 2020;267:105480.
[26] Vezhapparambu VS, Ellefmo SL. Estimating the blast sill thickness using changepoint analysis of MWD data. Int J Rock Mech Min Sci 2020;134:104443.
[27] Schunnesson H. Rock characterisation using percussive drilling. Int J Rock Mech Min Sci 1998;35(6):711–25.
[28] van Eldert J, Schunnesson H, Saiang D, Funehag J. Improved filtering and normalizing of measurement-while-drilling (MWD) data in tunnel excavation. Tunn Undergr Space Technol 2020;103:103467.
[29] Li SJ, Feng XT, Wang CY, Hudson JA. ISRM suggested method for rock fractures observations using a borehole digital optical televiewer. Rock Mech Rock Eng 2013;46(3):635–44.
[30] Babaei Khorzoughi M, Hall R, Apel D. Rock fracture density characterization using measurement while drilling (MWD) techniques. Int J Min Sci Technol 2018;28(6):859–64.
[31] Scoble MJ, Peck J. A technique for ground characterization using automated production drill monitoring. Int J Surf Min Reclam Environ 1987;1(1):41–54.
[32] Kalantari S, Hashemolhosseini H, Baghbanan A. Estimating rock strength parameters using drilling data. Int J Rock Mech Min Sci 2018;104:45–52.
[33] Schunnesson H, Elsrud R, Rai P. Drill monitoring for ground characterization in tunnelling operations. In: Proceedings of International Symposium on Mine Planning and Equipment Selection (MPES 2011). Almaty, Kazakhstan: National Center on Complex Processing of Mineral Raw Materials of the Republic of Kazakhstan; 2011. p.731–44.
[34] Ghosh R, Schunnesson H, Kumar U. The use of specific energy in rotary drilling: The effect of operational parameters. In: Proceedings of International Symposium on the Application of Computers and Operations Research in the Mineral Industry. Fairbanks, AK, USA: Society for Mining, Metallurgy, and Exploration, Inc.; 2015.p.713–23.
[35] Jang H, Topal E. A review of soft computing technology applications in several mining problems. Appl Soft Comput 2014;22:638–51.
[36] Kadkhodaie-Ilkhchi A, Monteiro ST, Ramos F, Hatherly P. Rock recognition from MWD data: A comparative study of boosting, neural networks, and fuzzy logic. IEEE Geosci Remote Sens Lett 2010;7(4):680–4.
[37] Basarir H, Wesseloo J, Karrech A, Pasternak E, Dyskin A. The use of soft computing methods for the prediction of rock properties based on

measurement while drilling data. In: Proceedings of the Eighth International Conference on Deep and High Stress Mining. Perth: Australian Centre for Geomechanics; 2017.p.537–51.
[38] Kubat M, Matwin S. Addressing the curse of imbalanced training sets: One-sided selection. In: Proceedings of the Fourteenth International Conference on Machine Learning. Nashville, Tennesse, USA: Morgan Kaufmann Publishers; 1997.p.179–86.
[39] Lewis DD, Catlett J. Heterogeneous uncertainty sampling for supervised learning. In: Machine Learning Proceedings 1994. Amsterdam: Elsevier; 1994. p. 148–56.
[40] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique. Jair 2002;16:321–57.
[41] Domingos P. MetaCost: A general method for making classifiers cost-sensitive. In: Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining. New York: ACM; 1999.p.155–64.
[42] Breiman L. Bagging predictors. Mach Learn 1996;24(2):123–40.
[43] Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F. A review on ensembles for the class imbalance problem: Bagging-, boosting-, and hybrid-based approaches. IEEE Trans Syst Man Cybern C Appl Rev 2012;42(4):463–84.
[44] Zhou ZH. Machine Learning. Singapore: Springer Singapore; 2021.
[45] Koopialipoor M, Armaghani DJ, Hedayat A, Marto A, Gordan B. Applying various hybrid intelligent systems to evaluate and predict slope stability under static and dynamic conditions. Soft Comput A Fusion Found Methodol Appl 2019;23(14):5913–29.
[46] Stone M. Cross-validation and multinomial prediction. Biometrika 1974;61 (3):509–15.
[47] Cawley GC, Talbot NLC. On over-fitting in model selection and subsequent selection bias in performance evaluation. J Mach Learn Res 2010;11:2079–107.
[48] Melville P, Mooney RJ. Diverse ensembles for active learning. In: Proceedings of the 21st International Conference on Machine Learning. New York: ACM; 2004.p.74-81.
[49] Opitz D, Maclin R. Popular ensemble methods: An empirical study. Jair 1999;11:169–98.
[50] Zhou ML, Chen JY, Huang HW, Zhang DM, Zhao S, Shadabfar M. Multi-source data driven method for assessing the rock mass quality of a NATM tunnel face via hybrid ensemble learning models. Int J Rock Mech Min Sci 2021;147:104914.
[51] Gurina E, Klyuchnikov N, Zaytsev A, Romanenkova E, Antipova K, Simon I, Makarov V, Koroteev D. Application of machine learning to accidents detection at directional drilling. J Petroleum Sci Eng 2020;184:106519.
[52] Hegde C, Wallace S, Gray K. Using trees, bagging, and random forests to predict rate of penetration during drilling. In: Proceedings of the 2015 SPE Middle East Intelligent Oil & Gas Conference & Exhibition. Abu Dhabi, UAE: Society of Petroleum Engineers; 2015.
[53] Singhal Y, Jain A, Batra S, Varshney Y, Rathi M. Review of bagging and boosting classification performance on unbalanced binary classification. In: Proceedings of the 2018 IEEE 8th International Advance Computing Conference (IACC). Greater Noida, India: IEEE; 2019.p.338–43.
[54] Priest SD, Hudson JA. Discontinuity spacings in rock. Int J Rock Mech Min Sci Geomech Abstr 1976;13(5):135–48.
[55] Legendre P, Legendre L. Numerical Ecology. 2nd ed. Amsterdam: Elsevier; 1998.
[56] Choi SS, Cha SH, Tappert C. A survey of binary similarity and distance measures. J Syst Cybern Inf 2010;8:43–8.
[57] Yong R, Ye J, Liang QF, Huang M, Du SG. Estimation of the joint roughness coefficient (JRC) of rock joints by vector similarity measures. Bull Eng Geol Environ 2018;77(2):735–49.
[58] Sokal R, Michener C. A statistical method for evaluating systematic relationships. Univ Kans Sci Bull 1958;38:1409–38.
[59] Das B, Krishnan NC, Cook DJ. Handling class overlap and imbalance to detect prompt situations in smart homes. In: Proceedings of the 2013 IEEE 13th International Conference on Data Mining Workshops. Dallas TX, USA: IEEE; 2014.p.266–73.
[60] Cochran WG. Sampling Techniques. 3rd ed. John Wiley & Sons; 1977.
[61] Sain H, Purnami SW. Combine sampling support vector machine for imbalanced data classification. Procedia Comput Sci 2015;72:59–66.
[62] Breiman L. Random forests. Mach Learn 2001;45(1):5–32.
[63] Quinlan JR. Bagging, boosting, and C4.5. In Proceedings of the 13th National Conference on Artificial Intelligence. Portland: AAAI Press; 1996,p.725–30.
[64] Probst P, Wright MN, Boulesteix AL. Hyperparameters and tuning strategies for Random Forest. Wiley Interdisciplinary Rev Data Mining Knowl Discov 2019;9 (3):1–15.
[65] Hernández-Lobato JM, Gelbart MA, Adams RP, Hoffman MW, Ghahramani Z. A general framework for constrained Bayesian optimization using information-based search. J Mach Learn Res 2016;17:5549–601.
[66] Shekar BH, Dagnew G. Grid search-based hyperparameter tuning and classification of microarray cancer data. In: Proceedings of the 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP). Gangtok, India: IEEE; 2019.p.1–8.
[67] Bergstra J, Bengio Y. Random search for hyper-parameter optimization. J Mach Learn Res 2012;13:281–305.
[68] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 14th international joint conference on Artificial intelligence. Montreal: Morgan Kaufmann Publishers Inc; 1995. p.1137–43.
[69] James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning. NY: Springer New York; 2013.