

# Can new technologies shake the empirical foundations of rock engineering?

**D Elmo** *University of British Columbia, Canada*

**D Stead** *Simon Fraser University, Canada*

**B Yang** *University of British Columbia, Canada*

**R Tsai** *University of British Columbia, Canada*

**Y Fogel** *University of British Columbia, Canada*

## Abstract

*The past decade has witnessed an increasing interest in applications of machine learning (ML) to solve mining and geotechnical problems. This is largely due to an increased use of high-level programming languages, development of user-friendly and open source ML libraries, improved computational power, and increased cloud storage capacity to handle large and complex data sets. The benefit of incorporating ML in rock engineering design are apparent, including the reduction in the time required to sort and characterise field data and the capability to find mathematical correlations between overly complex sets of input data. However, when applied to geotechnical engineering, the question arises as to whether ML can truly provide objective results. In geotechnical engineering, where the medium considered is typically heterogenous and only limited information is spatially available, experience and engineering judgement dominate the early stage of the design process. However, experience and engineering judgement alone cannot reduce data uncertainty. It is also true that the inherent variability of natural materials cannot be truly captured unless sufficient field data is collected in an objective manner.*

*This paper investigates the readiness of the technical community to integrate ML in rock engineering design at this time. To fully realise the potential and benefits of ML tools, the technical community must be willing to accept a paradigm shift in the data collection process and, if required, abandon empirical systems that are considered 'industry standards' by virtue of being commonly accepted despite acknowledging their limitations.*

**Keywords:** *cognitive biases, rock mass classification systems, uncertainty and variability*

## 1 Introduction

The term 'machine learning' (ML) somehow evokes modern images of autonomous machines capable of making decisions without the need of human intervention. In today's world of social media, machine learning is a 'buzzword'; a shortcut to otherwise complex algorithms such as naive Bayes, random forest (RF), artificial neural networks, and support vector machines. In this context, ML is not a novel approach. In fact, Zhang & Song (1991) discussed how neural networks could be applied to rock mechanics using either quantitative or qualitative data. The novelty rests with the recent surge of ML thanks to the wide availability of faster computers, high performing graphical processing units and open source deep learning libraries (Elmo & Stead 2020a).

It is important to note that ML techniques offer the ability to look for patterns and correlations but on their own, they do not represent a new physical model of rock mass behaviour (McGaughey 2020; Elmo & Stead 2020a). The best definition of ML techniques for geotechnical applications is possibly by Marcus (2017) who called ML a "passive dredging system" to help in finding new correlations between datasets.

ML techniques are numerous and include methods for regression, classification, clustering, association, and anomaly detection. The choice of a given ML technique would depend on the type, quality, and quantity of data available, and the use that engineers are expected to make of the predictions. The same discussion applies to the distinction between shallow and deep learning methods. The requirement of significantly large training data sets may make the use of deep learning method impractical if not impossible in geotechnical engineering; rock engineering often relies on data collection methods with a relatively high degree of subjectivity, and consequently there is the need to process, analyse and prepare design data sets in an effort to reduce human uncertainty. In the context of rock engineering, ML algorithms can be found in a variety of applications, such as site characterisation, tunnels and underground design, blasting, slope stability and geo-environmental studies (Morgenroth et al. 2019).

Major mining companies have already demonstrated and expressed their visions in developing a future in mining where ML is heavily involved in providing real time data analytics leading to optimised and timely decision-making. Example areas of study for such opportunity include preventative maintenance, material movement (haulage and scheduling) optimisation, blast fragmentation optimisation, and water quality assurance. However, the application of ML to geotechnical aspects of mining engineering, like ground control design and management, remains to be evaluated due to the inductive nature of the design process.

The hype generated in the media about ML has unfortunately and arguably created the myth of ML algorithms as tools that can solve any type of problem. However, one of the main shortcomings of ML for design applications is their black-box nature. For example, neural network models might give good and sometimes robust predictions, but the relationships between factors (weights) may be lost in the vast hidden layers of the neural network. Transparency should be the key requirement for applying ML predictions to rock engineering design since an engineering decision should be made with high confidence in both the input and the output of the model.

The work by Zhou et al. (2016) and Pu et al. (2019) are examples in which ML techniques have been applied, not to predict when a phenomena may occur, but to characterise the risk associated with the phenomena. These authors considered both the natural and human-induced settings behind underground rock bursting, events which can have multiple influencing factors that are not always fully understood but can lead to impending ground failure. Examples of influencing factors are exposure, stress change, seismic activity, porewater pressure, and ground disturbance activities nearby. Although time factor (exposure time or prediction of exact failure time) was not included in these studies, the authors provided the basis for creating a risk map of rockburst-prone environments.

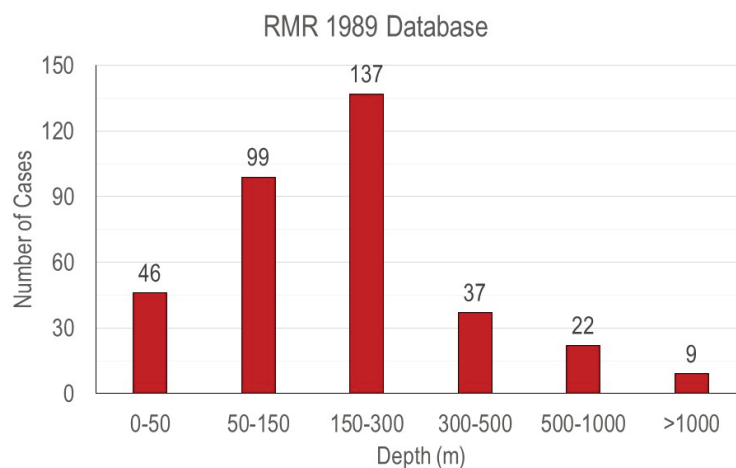
ML techniques are therefore of value when the primary goal of the model is to understand the influence of a certain parameter/feature (risk approach), and not the prediction itself (i.e. stability of final design), thus using a transparent approach that enables the user to easily isolate the weight assigned to every input data. More complex models reduce the control of the user and therefore are more suited to applications where there is a need to automate overly complex and repetitive tasks (e.g. analysing unstructured data such as images, videos, audios and text). In this case, ML may be better suited to a computational role rather than a decision-making role.

Ultimately, we need to understand that ML is not a tool that is supposed to code geological observations into precise numbers. Rather, it provides a method to compare and recognise geological patterns. In this context, ML becomes a repository of data, but it is the engineer that applies the data and not the machine. Therefore, there is the need to find a balance between engineering judgement (source of cognitive biases) and quantification of rock mass parameters to be used as training and validation data. The challenge is to develop algorithms capable of interpreting instances in which different qualitative assessments may be transformed into the same quantitative measurement, and whether such instances are geologically valid (Elmo & Stead 2020b).

## 2 Behavioural rock engineering

This paper proposes behavioural rock engineering as the study of rock engineering as it pertains to design decision-making processes made by individuals. Despite the recent introduction of a wide range of new technologies, to date, rock engineering remains an empirical discipline. Data and technology alone cannot reduce uncertainty since engineering judgement may be biased: our mind subconsciously interprets data by confirming instances that agree with our knowledge of a phenomenon and excludes data that do not agree with our assumptions; Taleb (2010) calls this process ‘narrative fallacy’. One of the major drawbacks of empirical knowledge and an inductive approach to design is that what we have learnt from the past, may not necessarily apply to new circumstances. This important aspect is often at danger of being ignored by industry professionals.

One clear example of narrative fallacy in rock engineering is given in Figure 1, which shows the number of cases and their depth below surface that form the database for the rock mass rating (RMR) system (Bieniawski 1989). If we were to use this database to train an ML algorithm, and subsequently apply the algorithm to predict the RMR of rock masses at depths of 1,000 m or greater we would not be able to trust the predictions since there is not data available to validate our predictions. To make matters worse, predictions would likely have to be based on 1D data (core logging) because of the lack of access to rock masses at depth. Despite best drilling practices, core samples are not undisturbed samples, and are likely to show a higher degree of fracturing due to handling, mechanical breaks, and de-stressing of weak veinlets. Classification systems strongly depend on fracture indices (i.e. rock quality designation (RQD), fracture frequency), and the accepted narrative of logging fractures as natural when in doubt is misleading. For instance, overestimating fracture frequency may produce conservative results for slope stability problems but would lead to poor predictions of rock mass fragmentation in the context of block cave mining.



**Figure 1 Rock mass rating cases as a function of depth**

Subjectivity of data interpretation, site specific geological conditions, and specific project characteristics requires ML algorithms to be trained and validated against the specific context of the project for which they are being developed. The physical ML algorithms would therefore not be transferable from project to project. The conceptual idea would be transferable, but the question arises as to how validate an ML algorithm for a new project that falls outside our current experience.

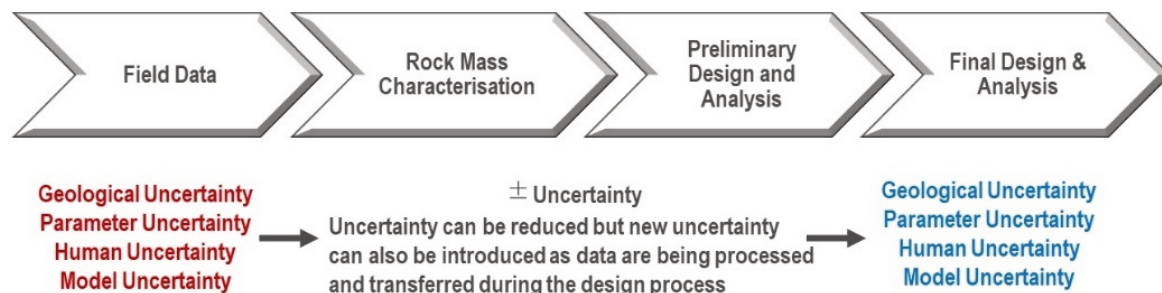
The problem of narrative fallacy is also related to another important aspect of rock engineering that is the definition of industry standards. It is not uncommon in rock engineering practice to refer to empirical methods as industry standards. Whereas the Collins Dictionary defines an industry standard as an “established standard, norm, or requirement in a particular area of business” (Industry standard 2020), the use of the word ‘established’ provides temporal constraints (i.e. the standard has been in existence for a long time), but it does not necessarily imply that a standard is the best technical solution, nor does it imply that the standard is correct. Indeed, industry standards should be subjected to continuous revision and well-informed improvements (Yang et al. 2020). Behavioural rock engineering teaches us that revisions of

established empirical methods are not so immediate, and they may often not welcome criticism; “Our engineering habits form slowly, and once formed are slow to change” Tye (1944).

The RQD (Deere et al. 1969) is a typical example of industry standard used in some of the most common rock mass classification systems whose validity has been challenged over the years. For instance, consider the following quotes: “RQD is not suited to form the basis for an engineering classification system of all rock masses, in terms of stability and support requirements” (Heuze 1971) and “Incorporation of RQD within the RMR and Q classification systems was a matter of historical development, and its incorporation into rock mass classifications is no longer necessary” Pells et al. (2017). Note the author of the RMR system (Bieniawski 1989) was also a co-author of this paper.

Most importantly, the formulation of RQD and the assumed 10 cm threshold is based on a somehow subjective decision rather than a true geological causation. “4 inches (10 cm) threshold was chosen after considerable deliberation as being a “reasonably” lower limit for a fair quality rock mass containing 3 or 4 joint sets of close to moderate spacing” (Deere & Deere 1989). Note that a significant limitation of the testing originally performed by Deere in 1969 was the small sample size and their limited variability; only 11 sites were tested, six sites were predominantly gneiss, and the remaining sites consisted of either limestone, sandstone, siltstone, rhyolite and dacite, and some schist. Data from one site was excluded as being an outlier since it did not match the relationship between RQD and velocity indexes that was derived based on data from the other sites (Yang et al. 2020).

It may be difficult to escape narrative fallacy in rock engineering since design practice is driven by data interpretation and reduction (i.e. the process of assigning numbers to geology). Applications of ML algorithms need to acknowledge the consequences of behavioural rock engineering, or risk to propagate known and unknown uncertainty in the data analysis process (Figure 2).

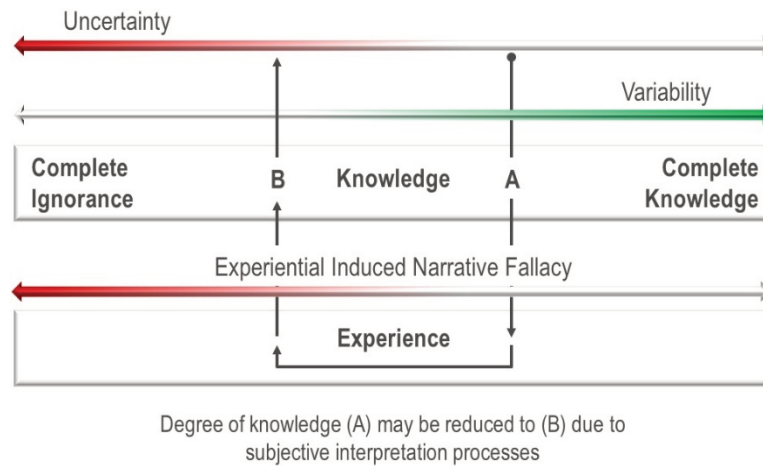


**Figure 2 Propagation of uncertainty in the design process**

## 2.1 Experience versus knowledge

Rock engineering projects require information about the intact rock, natural discontinuities, in situ stresses, and hydrogeological conditions. Data collection in rock engineering becomes a process of putting numbers to geology (Hoek 1999): qualitative geological descriptions are translated into engineering quantities. RQD, joint conditions in the RMR system, the parameters  $J_n$ ,  $J_r$ ,  $J_a$  in the Q-system (Barton et al. 1974) and the geological strength index (GSI) classification system (Hoek et al. 1995) are prime examples of a process of quantification of qualitative assessments. Derived quantities are not measurable properties and they may change depending on the engineering judgement and experience of the person collecting the data. Experience and engineering judgement therefore may introduce a sort of artificial variability, which is the product of human uncertainty.

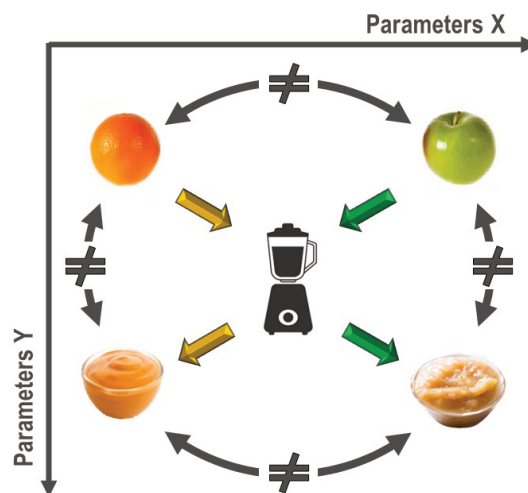
For experience to be considered a synonym of knowledge, it would require experience to be a process by which uncertainty is always reduced as more experience is gained. This is not possible in rock engineering design because of the cognitive biases that permeate the data collection and characterisation processes (Elmo & Stead 2020b), and the lack of truly quantitative measurements that can capture the highly variable nature of rock masses (Figure 3). Nonetheless, experience and engineering judgements still retain a critical role in the validation of the predictions made by ML algorithms.



**Figure 3** Role of narrative fallacy in rock engineering

## 2.2 The apple versus orange problem

Let us imagine a system to define the quality of a rock mass based on two key variables; note that each variable could, in principle, be a combination of multiple conditions (e.g. joint conditions, number of joint sets, fracture frequency and structural character). We call this the ‘apple versus orange problem’ which is represented in Figure 4: the blender represents variable geological conditions. The objective of the ML algorithm would be that of analysing all the different combinations of the parameters collected in the field and synthesise a unique set of X-Y variables to differentiate rock mass conditions. The rock type (apple or orange) could be known relatively easily, but the challenge remains for the ML algorithm to recognise which degree of blending of parameters would not be geologically sound.



**Figure 4** Schematic of a classification system controlled by two sets of parameters (X and Y)

The challenge is to develop an algorithm capable of interpreting instances in which different qualitative assessments may be transformed into the same quantitative measurement. Indeed, when considering existing classification systems, the same RMR value could refer to very different ratings for RQD, strength, spacing and joint conditions, or the same GSI value could represent massive to very blocky rock masses (Figure 5a). However, it is safe to assume that those rock masses would behave very differently under loading. When the problem is not reduced to just a single number, and greater emphasis is placed on the geological observations of rock mass characteristics, then the risk of representing different geological conditions with the same rating would be minimum. The problem arises when ML algorithms take charge and isolate the numbers from the underlying geology, with the risk that different rock masses would be assigned the same mechanical behaviour, as in the case of a ML algorithm attempting to quantify GSI (Figure 5b).

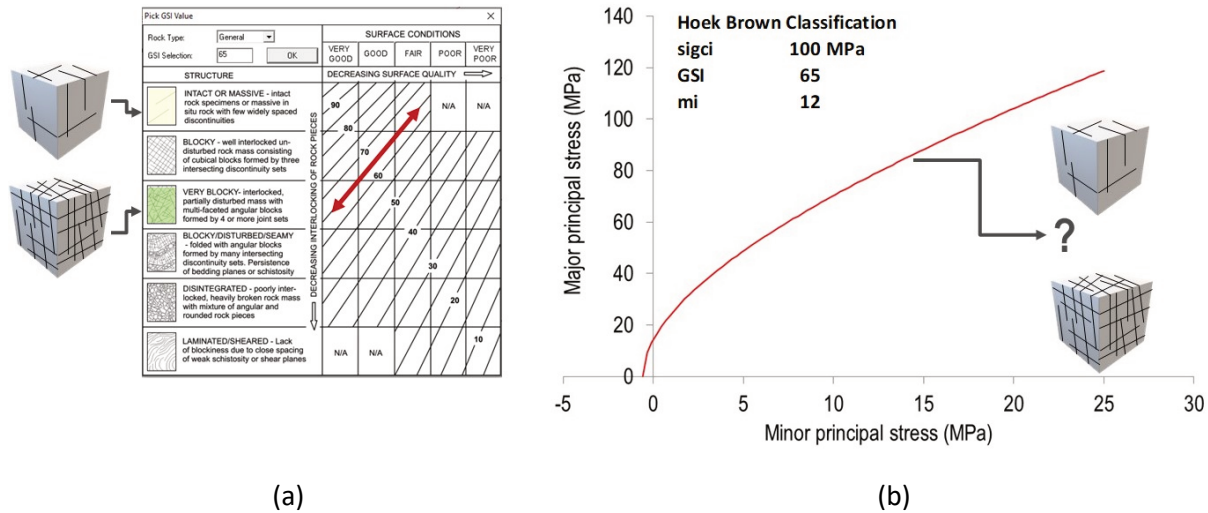


Figure 5 Relationship between geological strength index (GSI) and rock mass behaviour

### 3 A quantitative approach to rock mass characterisation

In an attempt to reduce the impact of qualitative measurements, Elmo et al. (2020) developed the network connectivity index (NCI) system (Figure 6), which is a method to quantify rock mass quality in the form of a potential function, relating observed rock mass conditions to induced rock mass damage under loading. The NCI system also addresses the challenges encountered using new technologies to acquire larger and better quality datasets of key geological parameters that are more appropriate for statistical analysis.

Note that the NCI system was primarily developed to address cognitive biases shaping the commonly accepted methods to measure and characterise rock bridges and rock bridge strength, which are based on geological conditions that are seldom encountered in the field (Elmo et al. 2018). Accordingly, the NCI system is primarily designed to work in combination with numerical analysis of brittle fracturing. NCI could potentially be used as a stand-alone classification system to provide an equivalent GSI rating. However, in its current version, the NCI system relies on a quantitative interpretation of qualitative characteristics of fracture surfaces and therefore, it would be subjected to the same cognitive biases as those affecting other classification systems. It is recommended to constrain the estimated equivalent GSI indicated in Figure 6 by using actual measurements of the NCI parameter to account for the irreversibility problem affecting rock mass classification system (Elmo & Stead 2020b).

Network connectivity is a measurable parameter well known in the discrete fracture network (DFN) community (e.g. Xu et al. 2006; Alghalandis et al. 2015). Building on the concept of network connectivity and the work by Elmo (2006) and Elmo & Stead (2020a), NCI combines  $P_{21}$ , number of fracture intersections per area ( $l_{20}$ ) and number of fractures per area ( $P_{20}$ ) into an index that can be easily measured from sampling of 2D rock exposures or derived from 3D DFN models. The basic principle driving NCI is that the longer the average fracture trace length and the greater the number of fracture intersections, the blockier the rock mass. NCI (Figure 6) is expressed as:

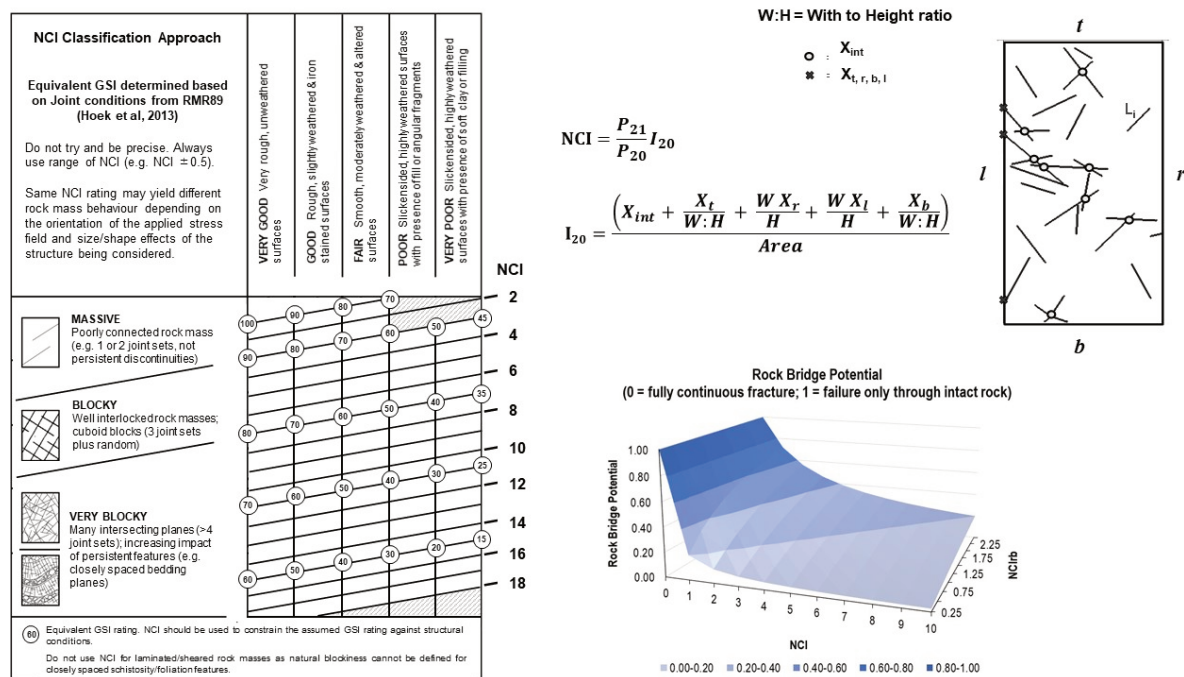
$$NCI = \frac{P_{21}}{P_{20}} I_{20} \tag{1}$$

$$I_{20} = \frac{\left( X_{int} + \frac{X_t}{W:H} + \frac{W X_r}{H} + \frac{W X_l}{H} + \frac{X_b}{W:H} \right)}{Area} \tag{2}$$

where:

- $P_{21}$  = areal fracture intensity.
- $P_{20}$  = areal fracture density.
- $l_{20}$  = areal fracture intersection density.

The parameter  $I_{20}$  in the NCI formulation must be corrected for censoring effects and shape effects (i.e. width to height ratio of mapping windows).  $X_t$ ,  $X_r$ ,  $X_l$ ,  $X_b$  and  $X_{int}$  are the number of intersections on the top, right, left, bottom of the rock mass domain and the internal intersections, respectively. NCI measurements could be obtained from sampling of 2D rock exposures (remote sensing tools are well-suited to the measurement of NCI) or derived from 3D DFN models generated based in 1D information (e.g. core logging). The basic principle of the NCI is that the longer the average fracture trace length and the greater the number of fracture intersections, the blockier the rock mass. A relatively low NCI rating implies that rock bridge failure occurs by connecting existing fractures, while for a high NCI rating, rock bridge failure may only occur in the form of intra-blocks damage.



**Figure 6 The network connectivity index system and rock bridge potential (modified from Elmo & Stead 2020b)**

Compared to other classification systems, the NCI provides a better indicator of rock mass quality. In the NCI system, because rock mass behaviour is related to the characteristics of the fracture network, a massive and a blocky rock mass would have very different NCI values. In combination with geomechanical models, the NCI approach allows us to characterise whether rock mass behaviour is largely a stress driven damage accumulation process (e.g. spalling) or a combination of stress driven failure and sliding along existing discontinuities by considering NCI values calculated pre- and post-failure ( $NCI$  and  $NCI_d$ , respectively). In this case, the rock bridge potential is given by the ratio of  $NCI_{rb}$  ( $NCI$  calculated for induced fractures) to  $NCI_d$ .

Note that NCI is not scale invariant and it could not be otherwise since rock mass properties are not scale invariant. To capture the representative elementary volume of the rock mass, NCI should be calculated for exposures of  $10 \text{ m}^2$  or higher. ML algorithms offer the ability to quickly calculate  $NCI$ ,  $NCI_{rb}$  and  $NCI_d$ , for multiple models, and in addition study the spatial variability of where induced fracturing is occurring in the different models and at what stage of loading, thus adding an extra layer of information to characterise the failure process of a rock mass.

## 4 Conclusion

In the past decade there has been an increasing interest in applications of ML to solve mining and geotechnical problems. High-level programming languages and the development of user-friendly and open source libraries have contributed to the increased applications of a variety of ML algorithms, which are well-suited to characterise field data and could be used to find mathematical correlations between complex

datasets. However, the question arises as to whether ML could or should be integrated within empirical schemes, which may include a degree of experiential fallacy. This key topic has been discussed within the proposed framework of ‘behavioural rock engineering’. The need for a paradigm shift is emphasised including a critical reappraisal of empirical systems that, although considered ‘industry standards’ by virtue of being commonly accepted, have known and important limitations.

It is not difficult to envision a significantly improved core logging and data processing approach in the future where imaging technologies such as automated core scanning are coupled with ML processing capability. However, such an approach demands more quantitative rock mass descriptions, which would enable reduction in the considerable bias introduced by human subjectivity. The NCI system and the concept of a rock mass ellipsoid discussed in this paper are examples of methods developed in the context of research for new, more quantitative methods to describe rock mass conditions. Note that the NCI system described in this paper is not intended to be yet another attempt to quantify the GSI classification system. The NCI rather builds on the framework of the GSI system focussing on the two descriptive parameters (structural characters and joint conditions) and is intended to be used in combination with geomechanical numerical models (synthetic rock mass models) to define a rock mass quality pre- and post-failure.

## Acknowledgement

The authors would like to acknowledge that this paper is based on a discussion on the role of cognitive biases in rock engineering that forms the core of a recent manuscript (Elmo & Stead 2020b), in which the authors first introduced the concept of behavioural rock engineer. The section included in this paper about NCI is a revised version of the one presented in Elmo et al. (2020).

## References

- Alghalandis, YF, Dowd, PA & Xu, C 2015, ‘Connectivity field: a measure for characterising fracture networks’, *Mathematical Geosciences*, vol. 47, issue 1, pp. 63–83.
- Barton, N, Lien, R & Lunde, J 1974, ‘Engineering classification of rock masses for the design of tunnel support’, *Rock Mechanics*, vol. 6, pp. 189–236.
- Bieniawski, ZT 1989, *Engineering Rock mass classification*, Wiley, New York.
- Deere, DU & Deere, DW 1989, *Rock Quality Designation (RQD) After Twenty Years*, Rocky Mountain Consultants, Inc, Longmont, report prepared for Department of the Army, US Army Corps of Engineers, Washington.
- Deere, DU, Merritt, AH & Coon, RF 1969, *Engineering classification of in-situ rock*, Air Force Weapons Laboratory, Air Force Systems Command, Kirtland Air Force Base, New Mexico.
- Elmo, D 2006, *Evaluation of a Hybrid FEM/DEM Approach for Determination of Rock Mass Strength Using a Combination of Discontinuity Mapping and Fracture Mechanics Modelling, with Emphasis on Modelling of Jointed Pillars*, PhD thesis, University of Exeter, Exeter.
- Elmo, D & Stead, D 2020a, ‘Disrupting rock engineering concepts: Is there such a thing as a rock mass digital twin and are machines capable of “learning” rock mechanics’, in PM Dight (ed.), *Proceedings of the 2020 International Symposium on Slope Stability in Open Pit Mining and Civil Engineering*, Australian Centre for Geomechanics, Perth, pp. 565–576, [https://doi.org/10.36487/ACG\\_repo/2025\\_34](https://doi.org/10.36487/ACG_repo/2025_34)
- Elmo, D & Stead, D 2020b, ‘The role of behavioural factors and cognitive biases in rock engineering’, *Rock Mechanics and Rock Engineering*, submitted for publication.
- Elmo, D, Donati D & Stead D, 2018, ‘Challenges in the characterization of rock bridges’, *Engineering Geology*, vol. 245, pp. 81–96.
- Elmo, D, Stead, D & Yang, B 2020, ‘Disrupting the concept of rock bridges’, *Proceedings of the 52nd International Symposium on Rock Mechanics*, American Rock Mechanics Association, Golden, <https://www.onepetro.org/conference-paper/>
- Heuze, FE 1971, ‘Sources of errors in rock mechanics field measurements and related solutions’, *International Journal of Rock Mechanics and Mining Sciences*, pp. 297–310.
- Hoek, E 1999, ‘Putting numbers to geology, an engineer’s viewpoint’, *Quarterly Journal of Engineering Geology*, vol. 32, pp. 1–19.
- Hoek, E, Kaiser, PK & Bawden, WF 1995, *Support of Underground Excavations in Hard Rock*, A.A. Balkema, Rotterdam.
- Industry Standard 2020, *Collinsdictionary.com dictionary*, <https://www.collinsdictionary.com/dictionary/english/industry-standard>
- Marcus, G 2017, ‘Artificial intelligence is stuck. Here’s how to move it forward’, *New York Times*, <https://www.nytimes.com/2017/07/29/opinion/sunday/artificial-intelligence-is-stuck-heres-how-to-move-it-forward.html>
- McGaughey, J 2020, ‘Artificial intelligence and big data analytics in mining geomechanics’, *Journal of the Southern African Institute of Mining and Metallurgy*, vol. 15, pp. 15–21.
- Morgenroth, J, Khan, UT & Perra, M 2019, ‘An overview of opportunities for machine learning methods in underground rock engineering design’, *Geosciences*, vol. 9, issue 12.



- Pells, PJ, Bieniawski, ZT, Hencher, SR & Pell, SE 2017, 'Rock quality designation (RQD): time to rest in peace', *Canadian Geotechnical Journal*, vol. 54, pp. 825–834.
- Pu, Y, Apel, D, Liu, V & Mitri, H 2019, 'Machine learning methods for rockburst prediction-state-of-the-art review', *International Journal of Mining Science and Technology*, vol. 29, issue 4, pp. 565–570, <https://doi.org/10.1016/j.ijmst.2019.06.009>
- Taleb, N 2010, *The Black Swan: The Impact of the Highly Improbable*, Random House, New York.
- Tye, W 1944, 'Factor of safety – Or of Habit', *Journal of Royal Aeronautical Society*, vol. 48, pp. 487–494.
- Xu, C, Dowd, PA & Fowell, RJ 2006, 'A connectivity index for discrete fracture networks', *Mathematical Geology*, vol. 38, issue 5, pp. 611–634.
- Yang, B, Elmo, D & Stead, D, 2020, 'Questioning the use of RQD in rock engineering and its implications for future rock slope design', *Proceedings of the 52nd International Symposium on Rock Mechanics*, American Rock Mechanics Association, Golden, <https://www.onepetro.org/conference-paper/>
- Zhang, Q & Song, J 1991, 'The application of machine learning to rock mechanics', *Proceedings of the 7th ISRM Congress*, International Society for Rock Mechanics and Rock Engineering, Lisbon.
- Zhou, J, Li, X & Mitri, H 2016, 'Classification of rockburst in underground projects: comparison of ten supervised learning methods', *Journal of Computer in Civil Engineering*, vol. 30, issue 5.

