geosciences

MDPI

Article

Rock Engineering Knowledge and Radical Uncertainty: From Empirical Methods to Professional Practice

Davide Elmo [1,*] **and Samantha Kenzie Adams** [2]

[1] NBK Institute of Mining Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada
[2] Faculty of Applied Science & Engineering, University of Toronto, Toronto, ON M5S 1A1, Canada; samantha.adams@mail.utoronto.ca
* Correspondence: davide.elmo@ubc.ca

Abstract

It is important for the rock engineering practice to acknowledge the difference between uncertainty, which diminishes with adequate data access, and radical uncertainty, which persists because critical features and failure mechanisms may remain undetected, not because of inadequate sampling but because they represent conditions that cannot be expected. Radical uncertainty represents an ontological feature of complex geological systems rather than a limitation of our current state of knowledge. The paper's central thesis is that current rock engineering practice has developed what we term the "epistemological three-body problem": the interaction between (i) inherent geological uncertainty that includes radical uncertainty (unknown unknowns), (ii) empirical methods that lack field-scale validation yet have gained professional acceptance through historical precedent, and (iii) regulatory frameworks that demand apparent certainty. We demonstrate this thesis through three interconnected arguments. First, we expose the epistemological and validation challenges inherent in widely adopted design methods. Second, we analyze how operational definitions, validation processes, and numerical modelling approaches may generate misleading precision rather than meaningful understanding of rock engineering problems, and third, we propose a framework for acknowledging and working within the boundaries of radical uncertainty. On this basis, we must acknowledge that rock engineering practice necessarily operates under a standard of a "balance of probabilities". Given the nature of radical uncertainty, professional practice should evaluate methods not by whether they eliminate uncertainty, but by whether they represent reasonable approaches to managing it.

Keywords: GSI; Hoek–Brown criterion; uncertainty; professional practice; synthetic rock mass modelling; calibration versus validation; rock engineering knowledge

Check for updates

Academic Editor: Mohamed Shahin

Received: 13 November 2025
Revised: 3 February 2026
Accepted: 4 February 2026
Published: 7 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

When studying rock mass behaviour, we assume a straightforward process: (i) collect data, (ii) make observations, and (iii) synthesize interpretations. This assumes theory-neutral observation, which is epistemologically naive. Without prior frameworks, how do we determine which observations are significant, which data to collect, and at what scale? Observation is never theory-neutral; we always bring conceptual frameworks that guide both our search strategies and interpretive processes.

When mathematical and validation frameworks prove insufficient, expert judgement is often invoked in rock engineering practice to fill the gap. This judgement represents

a form of implicit knowledge that cannot be easily formalized or transmitted through equations. This problem becomes acute when experts disagree about geological conditions or describe the same rock mass using different material parameters. Such disagreements reveal that our rock engineering knowledge may be more accurately described as informed opinion shaped by individual experience. Expert judgement becomes a necessary but epistemologically problematic component of rock engineering practice. Indeed, recent fatal ground-control incidents in mining operations worldwide demonstrate that despite decades of empirical refinement and increasing computational sophistication, our profession's methods for handling uncertainty may be fundamentally inadequate. This paper argues that rock engineering's empirical methods and professional practices systematically mishandle uncertainty by ignoring the distinction between uncertainty and radical uncertainty. Through this conflation, the profession claims a level of predictive capability that its epistemological foundations cannot support.

1.1. The Role of Radical Uncertainty in Rock Engineering Knowledge and Practice

Let us illustrate rock engineering knowledge as a cone that reduces uncertainty by capturing it (Figure 1). In our "cone" analogy, the process of gaining rock engineering knowledge is akin to studying the interior of a dark room through scattered pinholes (boreholes) and small surface openings (outcrops), inferring a three-dimensional structure from limited one-dimensional views. Data collection efforts aim to cut larger openings, and yet can never fully illuminate the interior.

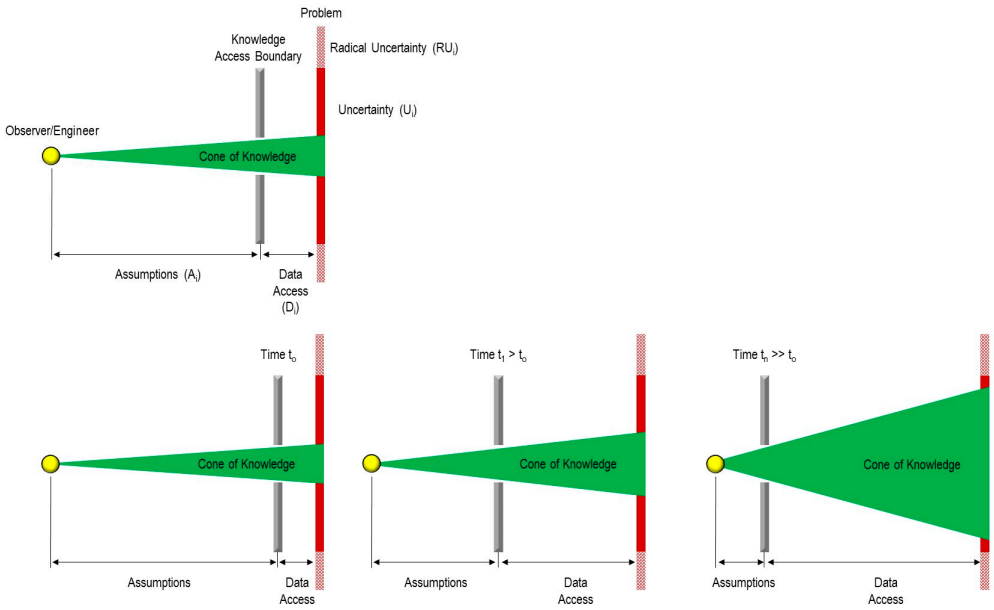


Figure 1. Rock engineering knowledge is imagined as a cone that reduces uncertainty by capturing it. Note the distinction between uncertainty and radical uncertainty. Radical uncertainty represents unknown unknowns that persist even with perfect data collection, distinguished from regular uncertainty, which can be reduced by additional information.

Our analogy emphasizes the difference between uncertainty, which diminishes with adequate data access (once sufficient data exist, additional observations do not materially change understanding or conclusions), and radical uncertainty [1], which persists because critical features may remain undetected, not due to inadequate sampling but because they represent unknown unknowns, a set of conditions that cannot be expected, because there has been no prior experience or theoretical basis for the conditions to occur. A single such observation can disproportionately alter aggregate knowledge regardless of sample size.

Therefore, fundamental limitations will always remain in making rational decisions. This requires shifting to a different epistemological goal, one that does not seek accuracy but seeks to manage uncertainty.

Figure 2 shows that radical uncertainty persists in geological systems and cannot be truly eliminated. It represents an ontological feature of complex geological systems rather than a limitation of our current state of knowledge. More importantly, from a risk perspective, radical uncertainty manifests as unknown unknowns. As such, radical uncertainty cannot be eliminated or fully quantified through statistical techniques. That is not to say that statistical methods have no role in rock engineering design. They remain valuable for characterizing what can be measured and modelled (e.g., variability within observed samples), but they cannot bridge the epistemic gap to what we have not observed and cannot validate.

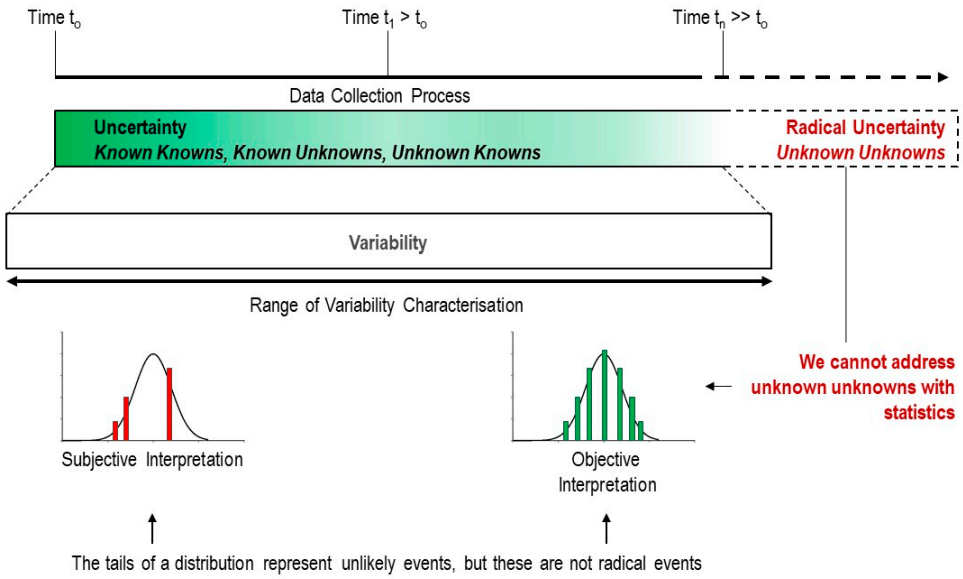


Figure 2. The difference between uncertainty and radical uncertainty and implications for statistical characterization.

The distinction between uncertainty and radical uncertainty is not the only challenge that practitioners face when excavating structures in rock media. Quoting [2], "*past observations may lead to the discovery of a theory, but the theory must predict the future*". This principle highlights a fundamental limitation in rock engineering practice, as our models and empirical methods often fail to achieve genuine predictive capability. Rather than addressing this fundamental limitation, we continue to add more data points to our databases and update our empirical criteria. But this process provides no certainty about future predictions, regardless of the accuracy we claim for the empirical parameters used in our design calculations. Additionally, since our empirical criteria can only be verified through observation, the practical constraints on conducting field-scale experiments necessarily limit our capacity to fully validate theoretical frameworks such as the Hoek–Brown criterion.

Following Einstein's [3] epistemological principles, theories require empirical validation rather than acceptance on purely a priori grounds. This paper argues that many methods commonly used in rock engineering practice derive much of their legitimacy from their historical role in the evolution of rock engineering. These systems have earned their position through decades of application and refinement. However, from an epistemological perspective, their continued validity depends on ongoing empirical evaluation rather than historical precedent alone. This suggests that their use should be accompanied by a critical assessment of their applicability to contemporary conditions and emerging

challenges. This places rock engineering knowledge on uncertain epistemological foundations, regardless of the mathematical sophistication we claim to adopt. Nonetheless, this limitation may not be decisive, since, given the inherent uncertainty of rock engineering knowledge (Figure 1), rock engineering predictions should never be framed in terms of certainty. Under absolute certainty, risk disappears entirely [1]. Either we possess complete knowledge, enabling failure-proof design, or we know that failure is inevitable, regardless of intervention, prompting alternative design approaches.

1.2. Objectives and Methodology

The paper's central thesis is that rock engineering has developed what we term the "epistemological three-body problem": the interaction between (i) inherent geological uncertainty that includes radical uncertainty (unknown unknowns), (ii) empirical methods that lack field-scale validation yet have gained professional acceptance through historical precedent, and (iii) regulatory frameworks that demand apparent certainty. Like the classical three-body problem in physics, no general analytical solution exists for this interaction. Yet the profession proceeds as if one does, encoding assumptions into increasingly sophisticated computational tools. What should readers do differently then? First, recognize that pursuing "certainty" through increasingly sophisticated empirical methods while ignoring validation gaps does not represent technical progress. Second, understand that regulatory frameworks demanding deterministic predictions are asking for something the epistemological structure cannot provide. This recognition itself represents a shift from viewing current challenges as solvable through better data or more complex models to understanding them as inherent to the system's architecture.

We demonstrate this thesis through three interconnected arguments. First, we expose the epistemological and validation challenges inherent in widely adopted design methods. Second, we analyze how operational definitions, validation processes, and numerical modelling approaches may generate misleading precision rather than meaningful understanding of rock engineering problems, and third, we propose a framework for acknowledging and working within the boundaries of radical uncertainty, as it cannot be eliminated through increasingly sophisticated analysis. In this context, it is worth mentioning that the domain of professional and regulatory implications is often neglected in academic literature, despite engineers' legal and ethical obligations when designs affect public interests.

Since our work is a methodological critique that analyzes the epistemological foundations of well-documented methods in the literature, we argue that additional empirical studies are not necessary to support the paper's objectives. Indeed, by systematically directing resources only toward practical outcomes, we reveal the troubling assumption that either we have solved the fundamental problems and all we need are minor tweaks, or that we will never solve them. Either way, we ensure we never escape our fundamental uncertainties. The problem is not that rock engineering remains empirical, but that in too many instances we have stopped being empirical, accepting historical correlations without the ongoing data collection that empiricism demands.

We have intentionally adopted a narrative style, as the paper's objectives are better served by discourse than by the conventional structure of technical papers.

2. Calibration and Validation Challenges in Rock Engineering Practice

Rock engineering inquiry should begin with experiments designed to answer specific questions. These experiments need not be confined to physical laboratory settings since the scales inherent in rock engineering problems necessitate large-scale field testing and synthetic experiments through numerical simulations. However, the fundamental question

is whether our observations represent knowledge or merely information. Observations divorced from conceptual frameworks remain mere data points in a database. Conversely, the validity of any theoretical framework depends entirely on the information we gather. This creates a circular dependency that requires both a quantitative structure from which to derive predictions and a theory of probability to assess the quality of these predictions.

Rock engineering's dependence on qualitative assessments, mistaken for quantitative measurements, creates inherent vulnerability to theoretical bias [4]. As Popper [5] argued, scientific knowledge should not depend on personal judgement but be evaluated solely through mathematics and logic. However, rock engineering necessarily incorporates professional expertise in interpreting qualitative geological features. This tension between scientific method and practical necessity shapes how the discipline develops and validates knowledge.

To understand calibration and validation challenges in rock engineering, we must first define these terms and recognize that their definitions vary across scientific and engineering disciplines. At the most fundamental level, calibration involves adjusting model parameters to match known observations or reference standards. It answers the question "Can we tune the model to reproduce what we have already observed?" Validation, conversely, tests whether a calibrated model can successfully predict independent observations not used during calibration, addressing the more demanding question "Does the model work for cases it has not seen before?" This fundamental distinction, however, manifests differently across disciplines, reflecting varied priorities and practical constraints (Table 1).

Table 1. Definition of calibration and validation across different disciplines.

Discipline	Calibration	Validation	Additional Layer
Measurement & Instrumentation	Adjusting an instrument against known standards (e.g., calibrating a scale with certified weights)	Confirming the instrument performs correctly across its operating range. The focus is more on equipment accuracy than predictive models	n/a
Computational Modelling	Adjusting parameters to match known behaviour	Comparing predictions to independent experimental/field data	Verification. It addresses whether equations are solved correctly
Machine Learning & Data Science	Fitting model parameters to training data (training, analogous to calibration).	Testing on a validation set during model development to tune hyperparameters	Testing. Final evaluation on completely independent test data (closest to validation in other fields)

Measurement and instrumentation fields interpret these terms through the lens of equipment accuracy rather than predictive modelling. This interpretation emphasizes reproducibility and accuracy in relation to external references rather than predicting new scenarios. Computational modelling in engineering and physics introduces an additional layer of complexity by distinguishing verification from both calibration and validation. This framework acknowledges that computational accuracy (verification) is conceptually distinct from parameter fitting (calibration) and predictive capability (validation). In machine learning and data science, the process is formalized into three distinct stages: training (analogous to calibration), validation, and testing. This three-tier structure reflects the field's emphasis on predictive capability and its ability to partition large datasets. The distinction between "validation set" and "test set" in machine learning represents a more granular approach than typically employed in traditional rock engineering practice.

In this paper, we use the following operational definitions:

- Verification refers to ensuring that a model correctly implements its intended mathematical formulation (i.e., solving the equations correctly).
- Calibration refers to adjusting model parameters to match observed behaviour within a specific context.
- Validation refers to demonstrating that a model's predictions reliably correspond to physical reality in conditions beyond those used for calibration (this requires independent field data).

Rock engineering practice occupies an uncomfortable middle ground among these disciplinary approaches. Unlike instrumentation calibration, our "measurements" of rock mass parameters involve qualitative geological assessments [6]. Unlike computational physics, we often conflate calibration with validation. Back-analysis illustrates this problem directly. Since back-analysis infers parameters from known failure conditions, it falls under the definition of calibration rather than validation, and yet it is often treated as the latter. For instance, when a slope fails, and engineers adjust strength parameters until numerical models reproduce the failure surface, this represents calibration. Yet such exercises are frequently cited as 'validating' both the parameters and the design method, a fundamental category error that pervades practice. More importantly, in rock engineering practice, validation does not operate in parallel; rather, it should be established independently for each site or condition, like a series circuit. And unlike machine learning, we rarely possess datasets large enough to partition into training, validation, and test subsets [7]. Similar validation challenges exist in other geoscience disciplines, particularly flood forecasting, where validation data (extreme events) are rare, and predictions must be made under uncertainty [8].

The distinction between methodological validation and professional acceptance may carry significant implications for professional practice, as discussed later in Section 6. What could, in principle, satisfy the professional standard of "balance of probabilities" for establishing reasonable practice may fall short of the scientific standard of "evidence beyond a reasonable doubt" required to claim genuine predictive validity.

Successful calibration does not guarantee a physically realistic mechanistic representation [7]. Consider two scenarios where numerical models successfully match observed deformation. In Scenario A, a continuum model with reduced modulus captures deformation by treating the rock mass as a degraded but homogeneous material. In Scenario B, a discontinuum model captures the same deformation through explicit slip along discrete fractures. While both achieve calibration success, they represent entirely different physical mechanisms. This distinction becomes critical when predicting future behaviour. This distinction is central to our argument about validation since successfully calibrating either model type to past behaviour does not validate the underlying mechanism, yet the mechanism determines future behaviour. In Scenario A, the model is fitting the data rather than capturing the underlying physics. One remedy involves continuous recalibration as new data emerge. However, two challenges persist: (i) practitioners may bypass mechanistic validation even when site-specific data exist, and (ii) the need for continuous calibration may indicate inadequate underlying relationships rather than merely incorrect parameters.

When considering computational modelling and data science, rigorous validation requires splitting available data into calibration and validation subsets before analysis begins. The calibration subset is used to develop the model, while the validation subset is used to independently assess its predictive performance. Applied to empirical rock engineering methods, this would involve deriving correlations from one data subset and then testing them against a reserved data set spanning different rock types, fracture configurations, and scales. Arguably, such a systematic validation process is largely absent in rock engineering practice.

3. Why "Extensively Applied" and "Validated" Are Not Synonyms

Empirical methods in rock engineering practice are often accepted as validated through widespread professional acceptance rather than through systematic testing against independent data [4,7]. This represents a departure from all the disciplinary frameworks described in the previous Section. This disciplinary peculiarity reflects a pragmatic adaptation to rock engineering's unique constraints (e.g., the difficulty of conducting large-scale field experiments, the scarcity of independent failure data, and the inability to derive predictive capability from case histories alone). However, it may represent an epistemological compromise that undermines the validity claims our profession routinely makes. Two critical examinations are presented to expose this problem.

3.1. The Inherent Uncertainty of Extensively Applied Empirical Correlations

In 1980, Hoek and Brown [9] provided the following formula as the basis for an empirical criterion to characterize rock mass strength:

$$\sigma_1 = \sigma_3 + \sqrt{m\sigma_c\sigma_3 + s\sigma_c^2} \quad (1)$$

where

σ_1 is the maximum principal stress at failure.

σ_3 is the minor principal stress applied to the specimen.

σ_c is the uniaxial compressive strength of the intact rock material in the specimen.

m and s are constants that depend on the properties of the rock mass and the degree to which it has been disturbed or fractured before being subjected to σ_1 and σ_3.

As the Hoek–Brown criterion evolved (see Equation (2) by [10]), the original parameter m was differentiated into m_i (subscript i for intact rock) and m_b (subscript b for broken or jointed rock mass). In Equation (2), the parameter σ_c was changed to σ_{ci} to represent the uniaxial compressive strength derived from fitting the Hoek–Brown failure envelope to a set of laboratory tests. Equation (2) also includes an exponent a to differentiate between intact rock (a equal to 0.5) and disturbed rock mass through the use of GSI (Geological Strength Index [11])

$$\sigma_1 = \sigma_3 + \sigma_{ci}\left(m_b\frac{\sigma_3}{\sigma_{ci}} + s\right)^a \quad (2)$$

$$m_b = m_i Exp\left(\frac{GSI - 100}{28 - 14D}\right) \quad (3)$$

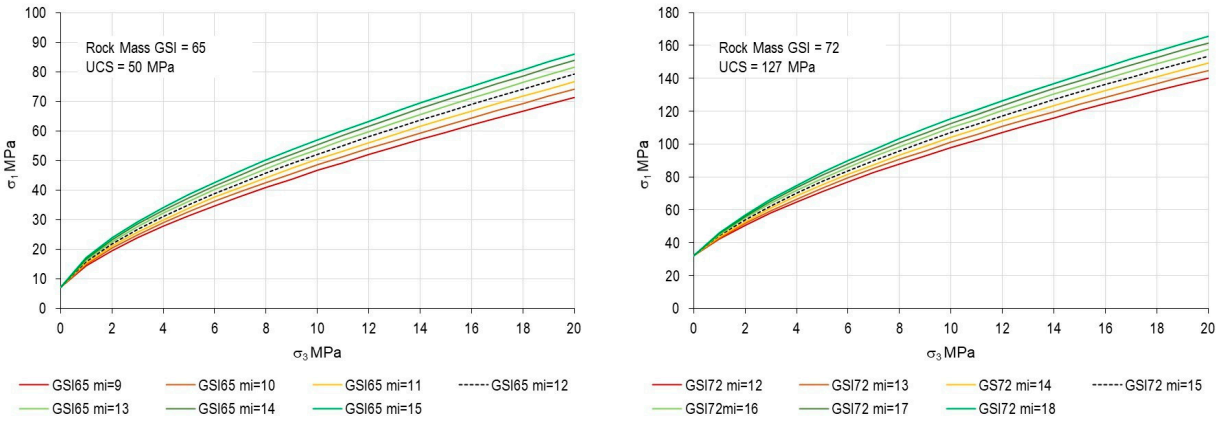
$$s = Exp\left(\frac{GSI - 100}{9 - 3D}\right) \quad (4)$$

$$a = \frac{1}{2} + \frac{1}{6}\left(e^{-\frac{GSI}{15}} - e^{-\frac{20}{3}}\right) \quad (5)$$

Practitioners generally agree that laboratory testing provides the most defensible basis for m_i determination. However, these tests are not routinely conducted for most projects. In the literature, there is debate over whether it is reasonable to rely on database correlations as estimates when site-specific data are unavailable [12]. Considerable effort has been devoted to developing precise m_i determination methods [12,13]. While we agree that testing should be the preferred method for determining m_i in intact rock samples, the practical significance of precise m_i determination is questionable given the absence of direct field-scale strength measurements and the criterion's sensitivity to GSI.

Using data for Middleton mine [14] and an undisclosed mine location [15], Figure 3a,b shows how the parameter m_i affects the shape of the resulting Hoek–Brown failure envelope for different GSI ($\pm$5) and σ_{ci} (assumed equal to the average UCS). For the two cases under consideration, the influence of m_i variations is most significant at low confining stresses,

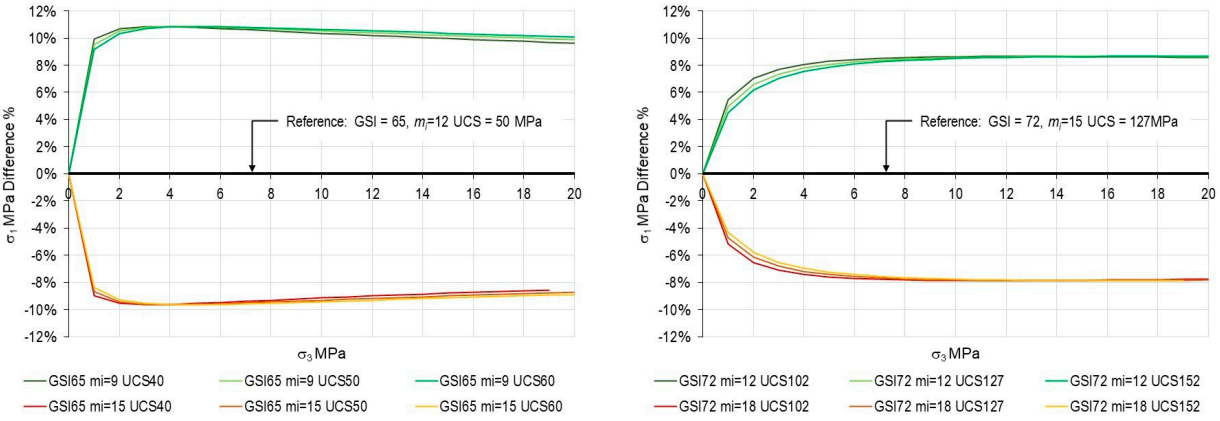
ranging from 0 to 2 MPa. While this analysis examines the compressive stress region, comparable effects would be anticipated in the tensile regime.



(**a**) (**b**)

Figure 3. Hoek–Brown curves for two different rock masses as a function of the parameter m_i: (**a**) UCS = 50 MPa, GSI = 65, and $m_i = 12 \pm 3$; (**b**) UCS = 127 MPa, GSI = 72, and $m_i = 15 \pm 3$. (**a**) Data from Middleton mine [14] and (**b**) data from an undisclosed mine location [15]. Note that for each case, the impact of varying the parameter m_i is rather minimal at low confinements (σ_3).

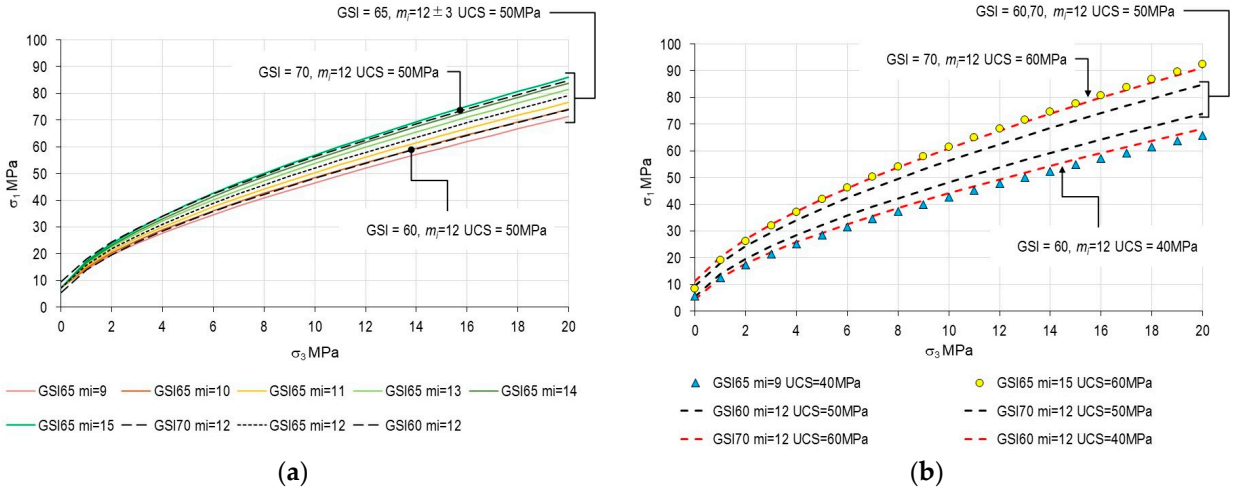
At the same time, Figure 4a,b shows that σ_1 predictions differ by less than 11% between the extreme cases of m_i ($m_i = 12 \pm 3$ and $m_i = 15 \pm 3$, respectively) and assuming that σ_{ci} will also vary by ±20%. These results suggest that for blocky and very blocky rock masses, the influence of m_i variations may be less pronounced than the literature focused on intact rock or massive rock masses would imply. The emphasis on precise m_i determination may, therefore, stem from the engineering profession's drive to assign numerical precision to geological parameters, rather than from demonstrated sensitivity of design outcomes to m_i values. These results are a good example of the mathematical optimism that pervades rock engineering practice [16], as if the assumption that adding analytical complexity to empirical relationships will overcome their foundational limitations as site-specific and context-dependent parameters.



(**a**) (**b**)

Figure 4. (**a**) Sigma 1 percentage difference using UCS = 50 MPa, GSI = 65, and $m_i = 12$ as reference, with cases for UCS = 50 ± 10 MPa, GSI = 65, and $m_i = 12 \pm 3$, and (**b**) Sigma 1 percentage difference using UCS = 127 MPa, GSI = 72, and $m_i = 15$ as reference, with cases for UCS = 127 ± 25 MPa, GSI = 72, and $m_i = 15 \pm 3$. Data from Middleton mine [14] and (**b**) data from an undisclosed mine location [15].

What remains less explored is the extent to which accurate m_i estimation affects practical design decisions. Figure 5 demonstrates that pursuing m_i precision offers little practical value. For the Middleton mine rock mass, the range of failure envelopes produced by varying m_i from 9 to 15 (m_i = 12 ± 3) at GSI = 65 can be equivalently reproduced by holding m_i constant at 12 and varying GSI by just ±5. Furthermore, the full spectrum of σ_1 corresponding to the scenarios shown earlier in Figure 3a can be reproduced by varying GSI by ±5 and σ_{ci} by 20% while still holding m_i at 12. This range of variability is not unusual in rock engineering practice and is representative of the rock mass at Middleton mine [14,17]. Figure 5 shows that GSI uncertainty can exert a comparable or greater influence on calculated strength parameters than m_i. This is particularly significant, given that GSI is not a quantitative measurement but a qualitative assessment, which is necessarily subject to observer interpretation and geological judgement.



(**a**) (**b**)

Figure 5. (**a**) Rock mass conditions corresponding to UCS = 50 MPa, GSI = 65, and m_i = 12 ± 3 and UCS = 50 MPa, GSI = 65 ± 5, and m_i = 12 (black dashed lines). (**b**) Rock mass conditions corresponding to UCS = 50 MPa, GSI = 65 ± 5, and m_i = 12, and UCS = 50 ± 10 MPa, GSI = 65 ± 5, and m_i = 12 (lines) compared to the extreme range of conditions expected at Middleton mine (triangle and circle symbols).

In this context, the proliferation of GSI quantification methods in the literature [18] may reflect the profession's difficulty in accepting that some parameters resist quantification, regardless of analytical sophistication. Perhaps we should acknowledge that, for blocky to very blocky rock masses (30 < GSI < 70), the precise determination of mi is subordinate to the degree of natural fracturing, as further discussed in the following Section.

3.2. How Limited Data Have Become Established Practice

What constitutes accuracy for empirically derived parameters and for qualitative assessments? Furthermore, with reference to the definitions presented earlier in Table 1, how can parameters be meaningfully calibrated to case studies when their underlying formulations have never been validated through field-scale testing of jointed rock masses and when quantification methods for qualitative indices yield potentially inconsistent results? Addressing these questions requires acknowledging that rock engineering practice frequently relies on "established standards" despite gaps in their validation. This raises questions about whether validation requirements are being applied consistently across established and new (emerging) methodologies.

The definitions of m_b and s (Equations (3) and (4)) originated from triaxial testing of Panguna andesite [9]. Despite being derived from a restricted dataset, Equations (3)

and (4) are now applied universally across rock engineering practice. However, none of the Panguna andesite specimens represented actual large-scale jointed rock masses; instead, they consisted of reconstituted samples (571 mm diameter) or small-scale rock specimens (152 mm diameter). This matters mechanically because laboratory-scale intact specimens fail through propagation of new fractures, whereas rock mass behaviour at the engineering scale may be dominated by pre-existing discontinuities, block interlocking, and structure-controlled failure modes. The extrapolation from the Panguna andesite specimen to universal applicability across all rock types and scales thus assumes that parameters derived from one failure mechanism can characterize an entirely different one (e.g., discontinuity-controlled mass behaviour).

The relationship between the m_b/m_i ratio and rock mass quality was inferred from correlations with estimated RMR_{76} [19] and Q values [20]. It was later extrapolated to all rock types beyond the original dataset, with GSI replacing RMR_{76}. Hoek and Brown [9] explicitly acknowledged the scarcity of reliable field-scale testing data needed to validate this extrapolation. Recently, ref. [21] demonstrated that a fundamental issue arises from the assumed universality of Equations (3) and (4): the same exponential relationship applies to all rock types, meaning that different rock masses may experience proportionally identical strength reductions relative to their input m_i and GSI values (Figure 6).

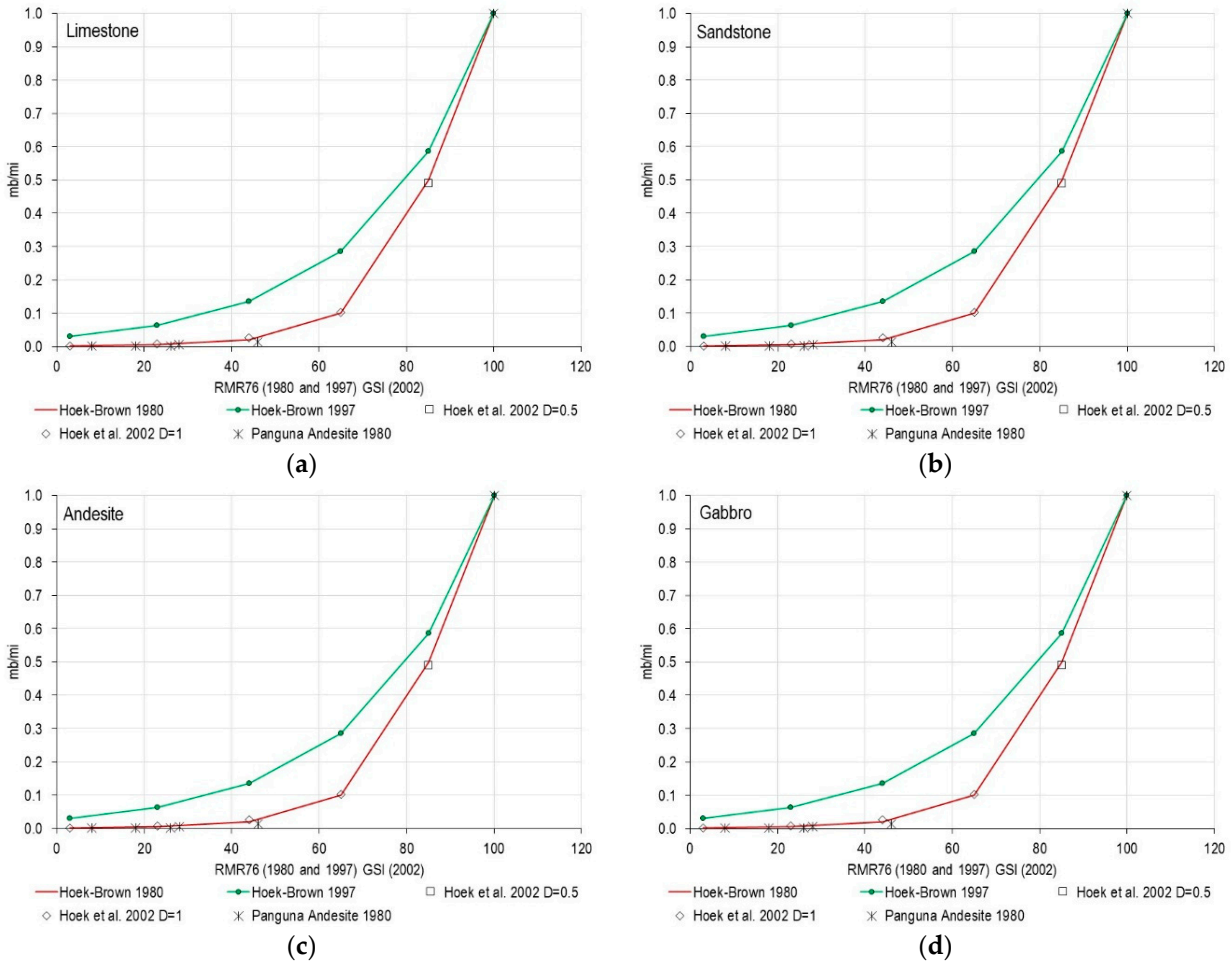


Figure 6. m_b/m_i relationships with rock mass quality for different rock types: (**a**) Limestone, (**b**) Sandstone, (**c**) Andesite and (**d**) Gabbro(based on data published originally in [9]). It becomes apparent that the same exponential relationship is invoked for all rock types, despite a lack of field validation, implying that different rock masses may experience proportionally identical strength reductions relative to their input m_i and GSI values. Modified from [21].

Figure 7 illustrates how the universal empirical relationships (3) and (4) give rise to the "false similarity problem" identified by [22], where fundamentally different rock masses may yield identical strength parameters. For example, while m_b values of 3.86 (GSI = 62, mi = 15) and 3.79 (GSI = 44, mi = 30) differ by less than 2%, they represent dramatically different geological conditions. Although the m_b values for the four scenarios shown in Figure 7 are nearly identical, the failure envelopes diverge only when GSI exceeds 80. While Figure 7 might suggest an inverse proportional relationship between m_i and GSI, this apparent correlation is merely an artifact of Equation (3), which mathematically couples these parameters.

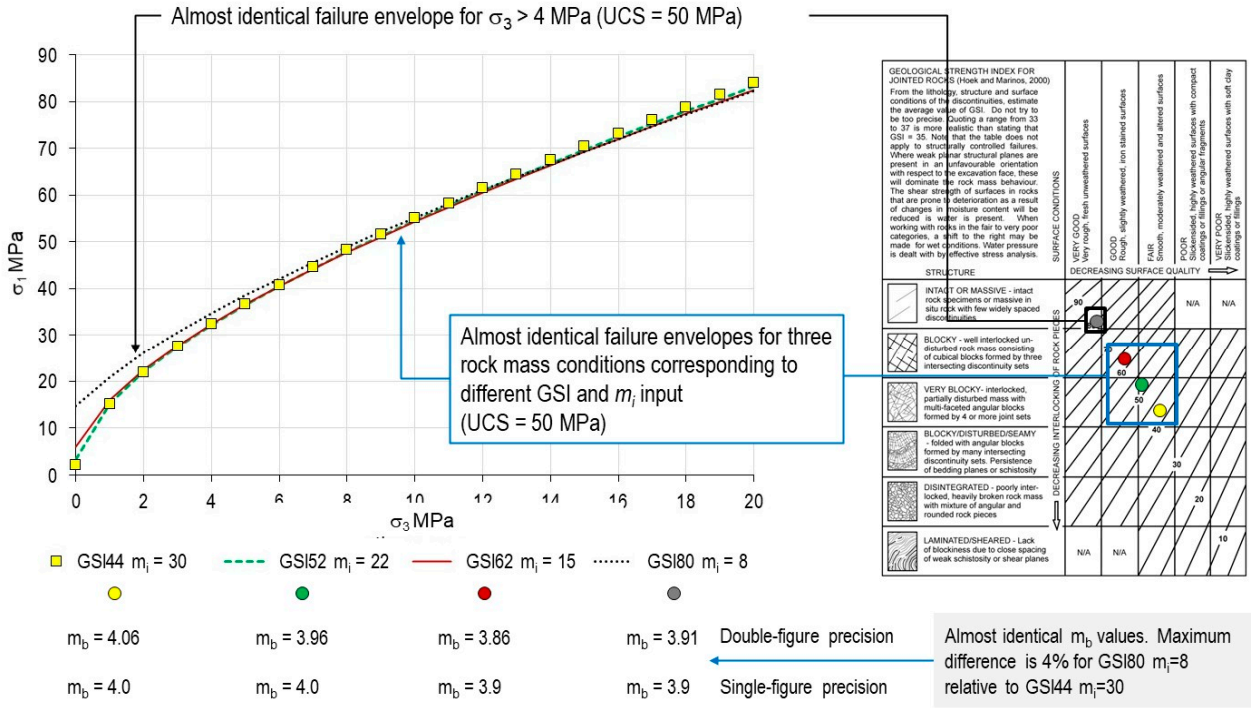


Figure 7. Example of rock mass conditions yielding almost identical Hoek–Brown curves despite representing dramatically different geological conditions. The Figure shows how empirical parameters derived from geological complexity cannot always be reverse-engineered to reveal the original geological understanding.

Conversely, a practitioner presented with the failure envelopes in Figure 7, and given only UCS and m_i values, would not be able to distinguish whether the underlying rock mass is massive, blocky, or very blocky. This aligns with conclusions by [4], who argue that reducing geological complexity to empirical parameters can remove information that cannot always be reconstructed. The professional implications are profound. Technical reviewers may not be able to work backward from reported parameters to verify the geological understanding that informed them, limiting the effectiveness of independent review processes.

One might object that Figure 7 presents purely conceptual scenarios and that practitioners should rely on site-specific mapped GSI values and measured m_i parameters. However, this objection exposes a fundamental challenge. Suppose we insist that m_i and GSI must be determined based on specific site conditions because rock masses are too variable for generalized relationships. How can we simultaneously accept that Equations (3) and (4), which transform those site-specific inputs into strength parameters, are universally applicable across all geological conditions and field sites? Clearly, the assumption that

for blocky to very blocky rock masses, the parameter m_i remains independent of natural fracturing deserves scrutiny.

If m_i relates to fracture mechanics and crack initiation processes [14], one would expect it to vary with fracture connectivity, which governs stress localization through block interlocking mechanisms. Using data from [14], Figure 8 shows that when fitting Hoek–Brown curves to simulated biaxial test data (synthetic rock mass, SRM, models), the derived m_i parameter varies with fracture intensity. While reasonable curve-fitting can be achieved by constraining m_i to the intact rock value (m_i = 12) and adjusting only GSI, this approach yields systematically different GSI estimates compared to allowing both parameters to vary (GSI range of [38, 79] vs. [50, 75], respectively, with reference to Figure 8).

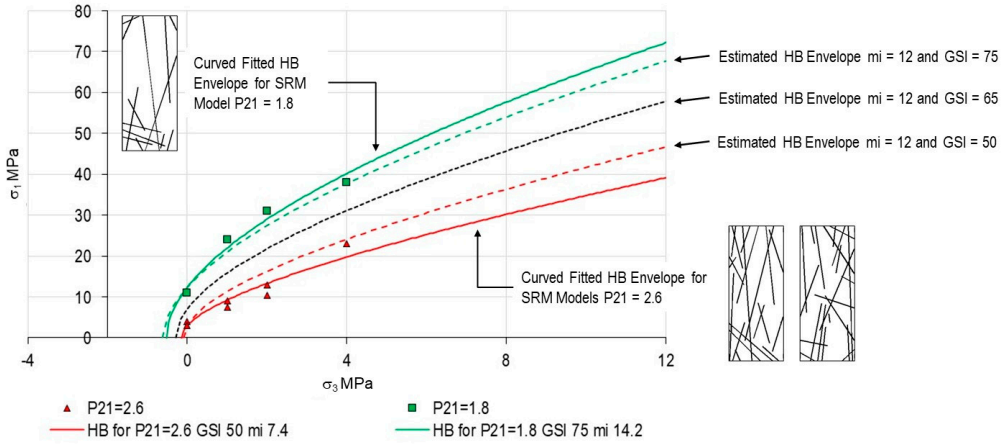


Figure 8. Hoek–Brown curves for SRM models of Middleton mine (2.8 m × 7 m dimensions). Comparison between curves for which both m_i and GSI must be fitted simultaneously to reproduce the simulated behaviour, and curves with equivalent GSI but for which m_i is considered as a material constant independent of structural context (m_i = 12, as per intact rock assumption).

This creates a methodological impasse rooted in the false similarity problem discussed above: either (i) we treat m_i as a material constant independent of structural context, requiring only GSI calibration to match SRM results, or (ii) we acknowledge that both m_i and GSI must be fitted simultaneously to reproduce the simulated behaviour. The former approach (i) derives from the original 1980–1995 conceptual framework, which was developed without large-scale validation, and assumes that m_i remains invariant regardless of fracture network characteristics. The latter (ii) represents a more rigorous interpretation of curve-fitting to simulated results and aligns with the original Hoek–Brown [9] formulation (Equation (1)), where the material parameter m should be empirically derived from tests that include structural components. If m_i varies mechanistically with fracture intensity in SRM models that explicitly simulate brittle fracture processes, should we force it to remain constant to preserve a continuum-based interpretation not supported by these mechanistic simulations?

To address this question, Figures 9–11 present the results of a series of SRM models for three lithologies (designated A, B, and C, with reference to the geological units in [16]). Details of the specific finite–discrete element method (FDEM) approach used in the simulations are provided in [14]. Relevant material properties are listed in Tables 2 and 3. Note that the m_i parameter in Table 2 was determined from triaxial tests of intact rock samples.

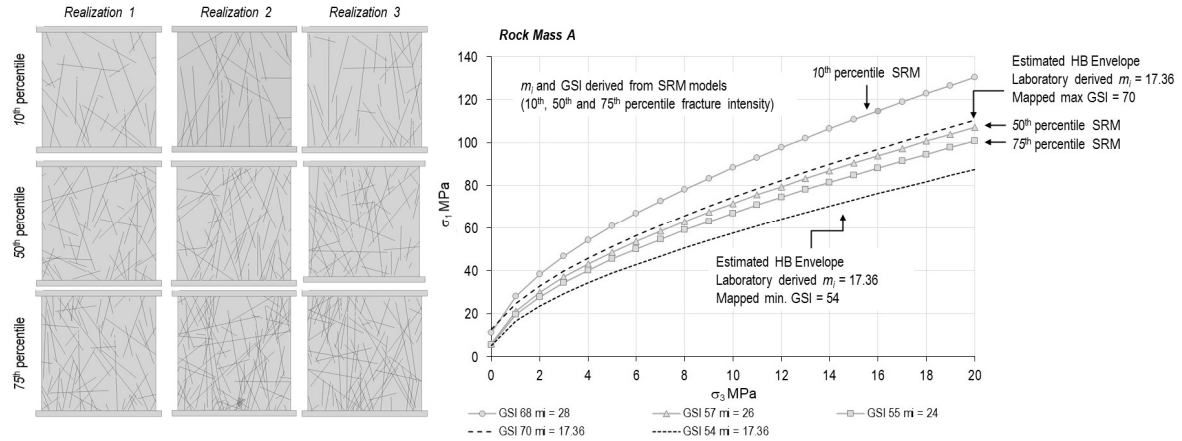


Figure 9. Hoek–Brown curves derived based on a series of biaxial tests for Rock Mass A (data from [15]) for different fracture intensities and multiple DFN realizations. The Hoek–Brown curves corresponding to the intact rock m_i value and the mapped GSI are also included.

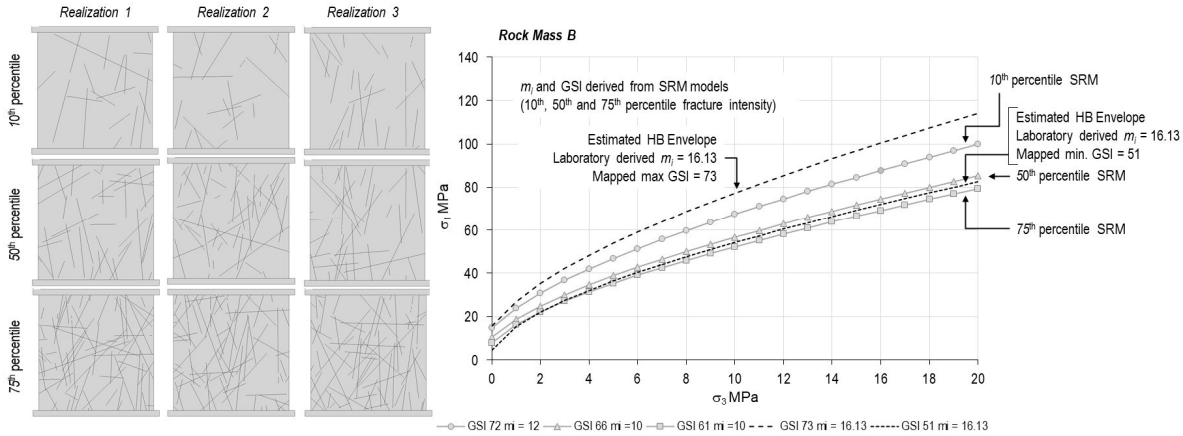


Figure 10. Hoek–Brown curves derived based on a series of biaxial tests for Rock Mass B (data from [15]) for different fracture intensities and multiple DFN realizations. The Hoek–Brown curves corresponding to the intact rock m_i value and the mapped GSI are also included.

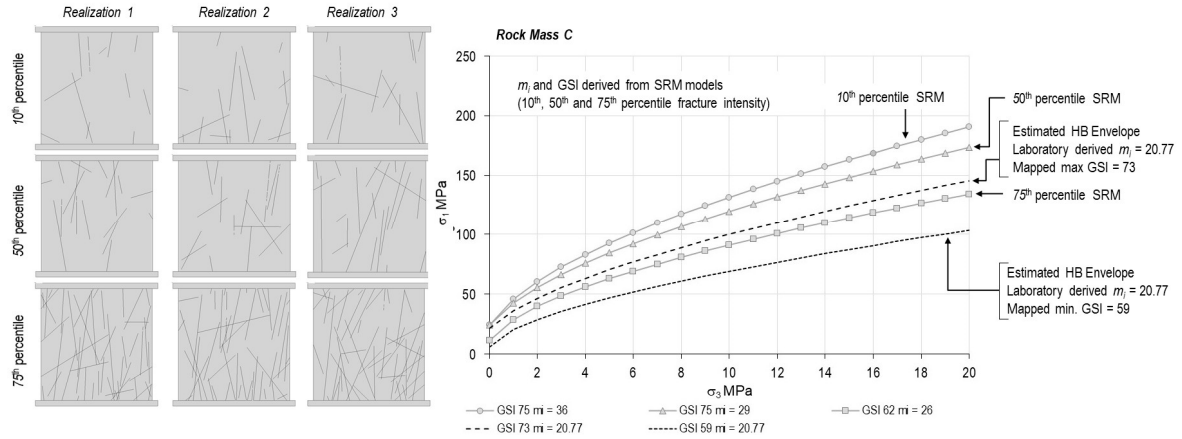


Figure 11. Hoek–Brown curves derived based on a series of biaxial tests for Rock Mass C (data from [15]) for different fracture intensities and multiple DFN realizations. The Hoek–Brown curves corresponding to the intact rock m_i value and the mapped GSI are also included.

Table 2. Material properties for SRM models A, B, and C.

Property	Rock Mass A	Rock Mass B	Rock Mass C
Density (ton/m^3)	2.7	2.6	2.6
Uniaxial compressive strength, UCS (MPa)	67.2	69.9	96.50
Indirect tension, σ_t (MPa)	2.4	3.1	3.9
Hoek & Brown m_i (laboratory data)	17.3	16.1	20.7
Young's Modulus, E (GPa)	20.0	29.5	37.1
Poisson ratio	0.21	0.21	0.21
Cohesion (MPa) *	9.5	10.1	12.5
Friction angle *	57	57	60
Fracture energy G_f (J/m^2)	6.0	6.7	8.4

* Calculated in RSData, with envelope range for 200 m depth.

Table 3. Properties for pre-existing and induced fractures in SRM models A, B, and C.

Pre-Existing Fractures (DFN Traces)	**Rock Mass A**	**Rock Mass B**	**Rock Mass C**
Cohesion (MPa)	0.5	0.5	0.5
Friction angle (degrees)	41	41	41
Normal stiffness (GPa/m)	100	50	50
Shear stiffness (GPa/m)	10	5	5
New fracture properties	**Rock Mass A**	**Rock Mass B**	**Rock Mass C**
Cohesion (MPa)	0.0	0.0	0.0
Friction angle (degrees)	31	031	31
Normal stiffness (GPa/m)	35	25	50

The Hoek–Brown curves for the SRM models in Figures 9–11 combine biaxial results from three DFN realizations at the 10th, 50th, and 75th percentile fracture intensities. While the results broadly validate the Hoek–Brown framework, the relationship is not uniform. Rock Mass B, for instance, exhibits a minimal strength difference between the 50th and 75th percentile fracture intensities, which is consistent with observations by [23] that fracture intensity (whether areal fracture intensity, P_{21}, or volumetric fracture intensity, P_{32}) is insufficient to characterize rock mass strength across different structural configurations. Rock mass strength and interlocking are fundamentally intertwined. Interlocking describes how block geometry, relative to the stress orientation, controls stress transfer pathways and governs strength mobilization. Therefore, fracture intensity alone cannot capture the impact of interlocking because it ignores the mechanical interaction between blocks.

This highlights a fundamental gap between field-based GSI and SRM-derived GSI estimates. Field-based GSI values transform the qualitative assessment of rock mass structure and surface conditions into equivalent continuum properties via Equation (2); therefore, they do not account for interlocking directly. Conversely, the SRM-derived GSI values mechanistically reflect how block geometry relative to stress orientation mobilizes strength through mechanical interlocking.

Figure 12 compares field-based GSI with SRM-derived GSI estimates using the two fitting approaches introduced earlier: (i) GSI fitted with m_i constrained to laboratory values, and (ii) GSI and m_i fitted simultaneously. The former systematically performs worse than the latter when compared to field-based GSI values. The results thus suggest that treating m_i as independent of rock mass structure (blocky to very blocky conditions) requires reconsideration.

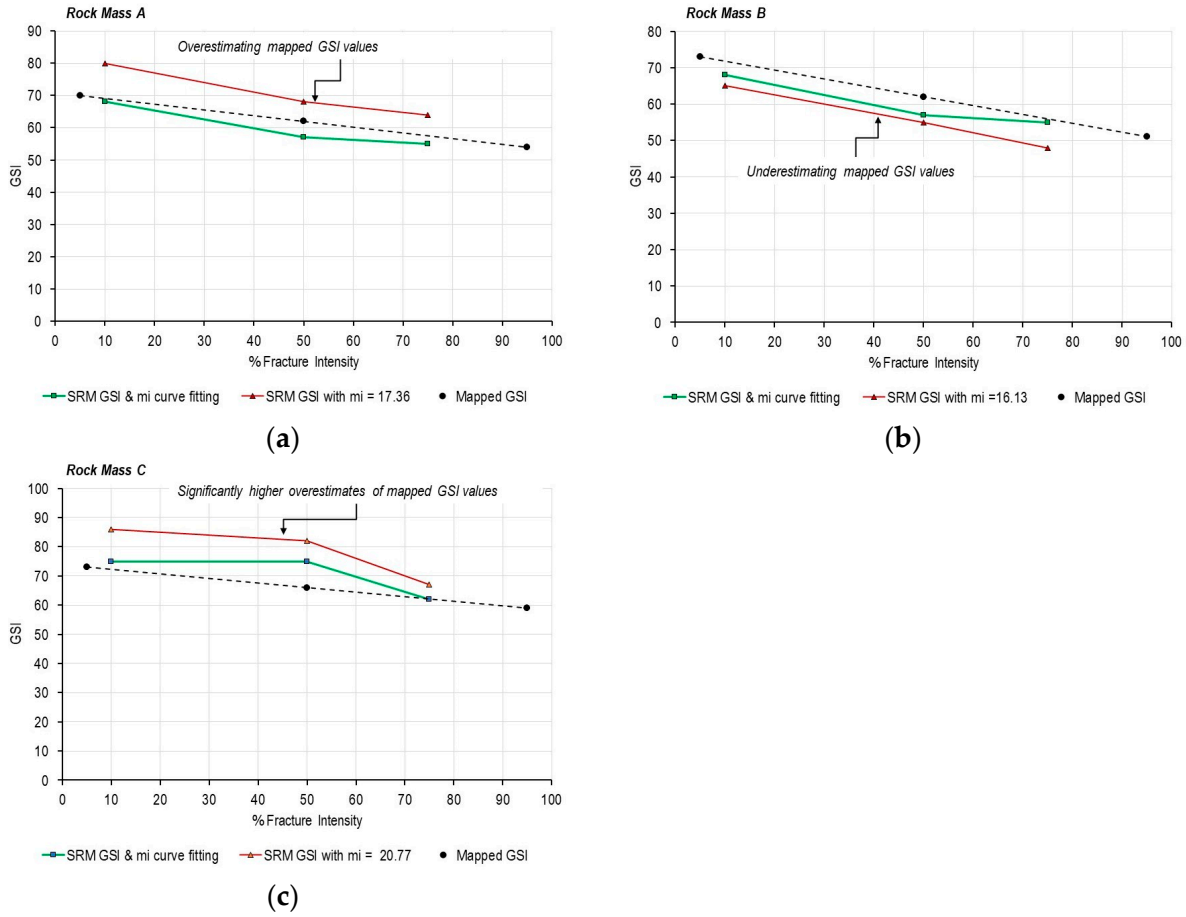


Figure 12. Comparison between field-based GSI (dashed line) and SRM-derived GSI estimates using the two fitting approaches introduced in the text: (i) GSI fitted with m_i constrained to laboratory values (red line), and (ii) GSI and m_i fitted simultaneously (green line). (**a**) Results for Rock Mass A in Figure 9; (**b**) results for Rock Mass B in Figure 10; and (**c**) results for Rock Mass A in Figure 11.

Critically, field-based GSI values are not exempt from uncertainty, as they are subject to qualitative judgement and rely on the observer's interpretation, geological experience, and the inherent difficulty of transforming qualitative assessments into quantitative values. The question is not whether field-based GSI values are "accurate" and SRM-derived values are "inaccurate" or vice versa. Instead, both approaches carry inherent uncertainties. Field-based GSI values, often derived from core logging rather than mapping of rock exposures, reflect a subjective assessment of rock mass quality. In contrast, SRM-derived GSI values reflect uncertainties in DFN representation and mechanical modelling. Neither field-based GSI nor SRM-derived values can eliminate radical uncertainty. The profession's repeated attempts to eliminate uncertainty through methodological elaboration, whether by quantifying GSI, refining m_i determination methods, or increasing DFN model complexity, reflect the difficulty of accepting that some geological characteristics remain fundamentally indeterminate (i.e., radically uncertain), regardless of analytical investment. The objective of either approach should be establishing a reliable range of GSI conditions that brackets plausible behaviour, rather than pursuing illusory precision.

3.3. The Representative Elementary Volume: Conceptual Limitations and Validation Challenges

The Representative Elementary Volume (REV) concept, widely used in rock engineering practice, is purported to establish the dimensions above which rock mass properties become scale-invariant. However, this interpretation fundamentally misconstrues what REV represents and conflates distinct concepts (structural homogenization, size effects, and

mechanical behaviour) into a single, convenient abstraction. The conventional understanding suggests that REV defines a critical volume beyond which rock mass strength stabilizes, implying that larger volumes exhibit consistent properties while smaller volumes show scale-dependent variation. This interpretation has become so ingrained that REV and "size effects" are now treated as synonymous in much of the literature. Yet this equivalence rests on a conceptual interpretation of what is actually being observed in numerical models.

Classical illustrations [10] and modelling of REV (e.g., [24–27]) suggest that as the problem scale increases, rock mass behaviour becomes less sensitive to the presence of adversely oriented individual discontinuities. This observation pertains to the statistical averaging of structural variability, rather than fundamental changes in material strength. As initially conceived by Bear [28] for porous media, REV represents the volume at which statistical homogeneity emerges, where properties can be characterized by their statistical distribution rather than deterministic values. However, rock engineering practice has transformed this statistical concept into a deterministic threshold, assuming that above the REV, rock mass properties (e.g., rock mass strength, deformation modulus) become constant single values. Indeed, results of synthetic rock mass models are often interpreted as rock mass strength vs. model size to determine the "average rock mass strength" at a given size, implying that variance disappears rather than stabilizes. This represents a fundamental misapplication of the REV concept.

In his discussion of scale effects, Pinto da Cunha [29] referred to geometrically homothetical samples of the same rock or rock mass, subjected to similar loading conditions. The use of the appropriate terminology merits clarification. Scale effects refer to mechanical characteristics that change when a system is proportionally altered across different hierarchical levels, maintaining geometric similarity (homothetic samples). Size effects, conversely, refer to how absolute dimensions influence behaviour independent of proportionality. In rock engineering discourse, "size effects" remains the prevalent terminology for three practical reasons [23]: (i) the dimensional difference between laboratory specimens and in situ rock masses is substantial rather than incremental; (ii) the literature typically presents strength as a function of absolute dimensions rather than dimensionless scaling ratios; and (iii) rock engineering problems are not necessarily scalable. For example, pillar dimensions are dictated by orebody geometry, design constraints, and economic considerations rather than simple geometric proportionality.

Central to understanding scale effects is the concept of homothety, that is, a mathematical transformation that proportionally scales all geometric features of an object about a fixed centre point by a constant ratio k. However, the term "geometrical homothety" introduces a crucial distinction that is systematically overlooked in rock engineering applications. True geometric homothety requires that the scaling function apply not only to external dimensions but to all structural features at all scales, including mineral grains, microcracks, macroscopic discontinuities, and the statistical distributions governing their occurrence. A genuinely homothetical rock mass model of twice the linear dimension would contain discontinuities of twice the length, spaced at twice the interval, with twice the aperture, embedded in a matrix with grain sizes doubled and microdefect spacing doubled accordingly. Real engineering problems do not meet these conditions [23].

SRM and DFN studies claiming to demonstrate size effects only scale the external problem boundaries. This represents partial homothety. While the model boundaries are scaled up, the natural fracture network continues to obey its own spatial distribution function. This distinction is not purely linguistic, as significant effects emerge when critically oriented structures are included or excluded stochastically as the model volume changes, introducing variability unrelated to scale. Similarly, attempts to define the REV of a rock mass using only geometric descriptors (e.g., volumetric fracture intensity P_{32} and volumetric fracture

count P30) decouple structures (natural fracture network) from mechanisms, ignoring how fracture networks translate into mechanical behaviour. Geometric scale-invariance does not guarantee mechanical scale-invariance because geometric characterization methods (e.g., the vertical axis in the GSI table, P_{32}, P_{30}, etc.) ignore loading conditions, which fundamentally influence how structure translates into mechanical behaviour. Indeed, GSI and similar rock mass classification systems establish isotropic conditions a priori.

While SRM studies are often used to demonstrate apparent size effects (as modelled volumes decrease, calculated rock mass strength increases), this interpretation may confuse mechanisms with dimensions [23]. Indeed, Bewick and Elmo [23] demonstrated behaviour contrary to conventional size effect theory, in which rock mass strength either remains scale-invariant (models without DFNs) or actually decreases with decreasing size (models with homothetical DFNs).

Applied to rock masses, this means that REV addresses the conditions under which continuum approximations become statistically reasonable, rather than the scale or size at which rock mass strength increases or decreases. The concept of rock mass complexity, introduced by [30], fundamentally concerns the limits of applicability of homogenized failure criteria, such as the Hoek–Brown criterion, rather than actual physical changes in rock mass strength with volume. This discussion reveals an additional paradox of the validation of rock engineering problems raised in Section 3. We cannot validate scale effect relationships in rock masses because there are no truly homothetical rock mass samples at the field scale. Laboratory studies can approach full homothety only within narrow size ranges and only for intact rock. Field observations necessarily sample different structural configurations rather than scaled versions of identical structures. This impossibility of validating scale effects reveals a manifestation of radical uncertainty unique to rock engineering. We cannot know whether the scale effects we observe in numerical models and attribute to rock mass behaviour actually exist as intrinsic properties, or whether they are artifacts of our inability to maintain geometric similarity across scales. This is a question that even more sophisticated testing or computational power cannot resolve. Each SRM model that claims to demonstrate scale effects actually captures the emergent behaviour of different structural configurations at different sizes, not the scaling of identical structures. This exemplifies how radical uncertainty in rock engineering extends beyond missing data to encompass the fundamental impossibility of validating basic conceptual frameworks, such as REV, through direct observation.

4. Rock Engineering and the Challenge of Operational Definitions

Bridgman [31] challenged how scientists conceptualize measurement and meaning. Bridgman argued that scientific concepts should be defined entirely through the operations used to measure them, which he named "operational definitions." For Bridgman, the concept of length, for example, means nothing more than the set of operations by which length is measured. This operationalism emerged from his recognition that classical physics had relied on concepts that could not be operationally defined, resulting in conceptual confusion.

In rock engineering, Bridgman's operationalism poses profound challenges. While Bridgman's strict operationalism has been critiqued for being overly restrictive [32], a modified operationalist lens reveals critical issues. Consider the concept of rock mass strength. What operations define this concept? We cannot directly measure rock mass strength in situ without destroying the very structure we seek to analyze. Instead, as discussed in Section 3, we infer rock mass strength from indirect measurements and observations, including combinations of laboratory tests on intact rock samples, back-analysis of failures, empirical correlations with classification indices, and SRM models. However, each of

these operations measures and observes something different. Laboratory tests measure the strength of intact rock under controlled conditions that deviate from the field stress states. Back-analysis assumes our numerical models can correctly simulate actual failure mechanisms. Classification systems attempt to quantify qualitative assessments through a series of rating schemes. And yet, even when combined, none of these operations offers a direct measure of rock mass strength.

The operational definitions framework requires that conceptual models be verified against the physical operations that define our measurements. This creates what we might call the validation paradox, since conceptual failure models can be verified only through observed failures, which is precisely what engineering design seeks to prevent. For instance, our conceptual models of slope failure incorporate assumptions about failure surfaces, strength parameters, and triggering mechanisms. To verify these models operationally, we would need to observe actual slope failures under controlled conditions. However, such validation would require:

- Deliberately inducing failures in prototype slopes (ethically and practically unacceptable).
- Waiting for natural failures to occur (temporally impractical for design purposes).
- Relying on historical failures (which introduces temporal and contextual uncertainties).

This paradox means that our most critical design concepts cannot be operationally verified. Nonetheless, when rock engineering models successfully reproduce past failures, we often interpret this as validation. This exemplifies a "post hoc ergo propter hoc" fallacy, which assumes that because models match past failures, they can predict those failures. Bridgman's operationalism rejects such reasoning because it conflates correlation with causation and retrofitting with prediction. Accurate operational validation would require that models, developed independently of the phenomena they claim to explain, successfully predict future observations [1]. In rock engineering, this standard is rarely met. Instead, we adjust model parameters until they reproduce known outcomes, then assume these calibrated models possess predictive validity.

The simplification from real rock mass to equivalent continuum models further complicates the problem. Figure 13 illustrates the progression from a real rock mass scenario to an equivalent continuum model. As models become simplified, from realistic, undulating fractures through increasingly simplified DFN representations to fully isotropic continuum models, mechanistic accuracy decreases, and conversely, model uncertainty gets exposed. How can we validate a fully isotropic continuum model against a highly anisotropic, discontinuous rock mass? Matching observed outcomes through calibration proves only that parameters can be adjusted to fit data, not that the simplified model captures the actual failure mechanism. We could argue that the more we abstract away from mechanistic reality, the less meaningful our validation processes become.

In rock engineering practice, unforeseen conditions often become the default explanation when reality diverges from prediction. However, this attribution carries significant implications depending on what type of uncertainty these unforeseen conditions represent:

- If they represent uncertainty (whether epistemic or aleatoric), this means that these conditions are unforeseen due to inadequate data collection and characterization. From a legal perspective, this amounts to acknowledging that we failed to recognize that we had not collected sufficient information.
- If they represent radical uncertainty, this means no one could claim they would have acted differently, since the conditions would not have been known to them either.

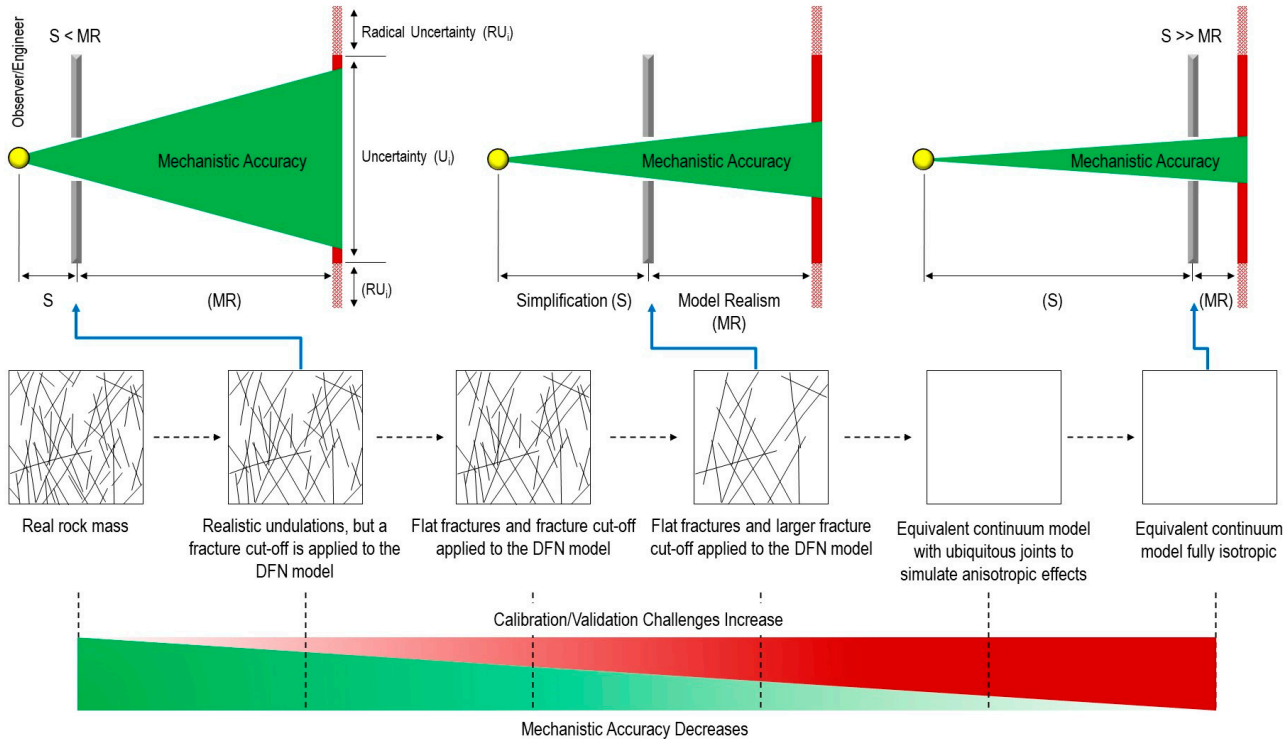


Figure 13. The validation paradox in rock mass modelling. As models progress from mechanistic realism (MR) to simplification (S), calibration/validation challenges increase while mechanistic accuracy decreases, and the impact of model uncertainty increases.

The authors suggest adopting what could be called "pragmatic operationalism", explicitly acknowledging that many of the measurements used in rock engineering design are proxies rather than direct observations of the phenomena we seek to understand. This includes key parameters such as rock mass strength and the deformability modulus. This approach would emphasize understanding the limitations and assumptions embedded in our operational definitions rather than treating them as windows onto physical reality. It is important to clarify what pragmatic operationalism offers and what it does not. The purpose of this framework is to establish a way of thinking about validation and uncertainty that differs from current practice. Pragmatic operationalism is not a set of standardized procedures or prescriptive checklists; rather, it is a conceptual framework focused on communicating uncertainty rather than measuring certainty. What constitutes appropriate implementation of pragmatic operationalism depends entirely on specific project contexts (e.g., geological conditions, regulatory environments, available data, consequences of failure, and stakeholder risk tolerance). In the context of pragmatic operationalism, an evaluation of confidence refers to an assessment of its degree of uncertainty, rather than certainty. For example, when using a factor of safety approach, we tend to use a larger margin, not because we are certain of our inputs, but because we acknowledge the uncertainty in our estimates and their variability. Accordingly, a factor of safety of 1.3 based on well-validated methods within their calibration range carries different epistemic weight than a factor of safety of 1.5 derived from extrapolated empirical relationships.

5. Uncertainty and Professional Responsibility in Rock Engineering Practice

Sections 2–4 have revealed significant challenges in rock engineering knowledge. We now address important practical implications. The challenge for practitioners is to manage professional responsibility and risk while acknowledging that our most widely used empirical and numerical methods often yield results that cannot be validated against field-scale reality. At the same time, rock engineering practice demands apparent certainty, driven by pressure related to professional liability and safety considerations. Consequently, empirical rock mass classification methods, which are intended to characterize the variability of rock mass conditions, are instead used as quantitative and supposedly precise measurement tools to derive rock mass properties.

The Large Open Pit (LOP) guidelines [33] illustrate how rock engineering practice prescribes probability thresholds that represent epistemic uncertainty rather than actual failure frequency. Kay and King [1] would refer to these probability thresholds as subjective probability. For instance, a 50% probability of failure threshold for a bench design does not imply that half of the benches will fail; instead, it acknowledges that the design can proceed despite knowledge limitations, with the understanding that operational management and monitoring will complement analytical predictions. They represent a form of pragmatic operationalism already embedded in our design practice. We proceed with structured professional judgement despite radical uncertainty, relying on operational experience, observational methods, and adaptive management to bridge the gap between our analytical predictions and the complex geological reality. In the authors' view, making this pragmatism explicit, rather than hiding it behind the pursuit of seemingly precise calculations (see Section 3), could better serve the profession.

5.1. The Challenge of Dual Uncertainty

When rolling a die, we face uncertainty about the outcome but possess certainty about the constraints, as the result will necessarily fall between one and six. This represents bounded uncertainty. We do not know the specific outcome, but we are aware of the possibilities and their associated probabilities. Rock engineering offers no such epistemological comfort, as it confronts what might be termed dual uncertainty:

- First, we remain uncertain about our inputs. What are the actual rock mass properties at depth? How do joint properties vary spatially? What is the true in situ stress state? These input uncertainties reflect not only measurement limitations, but also fundamental constraints imposed by observing three-dimensional geological structures with one-dimensional sampling.
- Second, even if we could magically eliminate all input uncertainty and know exact geological conditions, we would still face output uncertainty: What will actually happen when we excavate? Which failure mechanism will dominate? How will the rock mass respond to changing stress conditions over time? Will progressive failure occur, and if so, at what rate? This output uncertainty exists because geological systems exhibit emergent behaviour, scale-dependent mechanisms, and time-dependent processes that cannot be fully predicted even with perfect knowledge of the initial conditions.

Dual uncertainty distinguishes geological systems from most other engineering domains. A structural engineer designing a steel frame faces uncertainty in material properties and loading conditions, but, given those inputs, the output behaviour is well-constrained by established theory. The steel's yield strength may vary within a range, but its fundamental mechanical behaviour is known and reproducible. In contrast, rock mass behaviour involves applying empirically derived approximations of complex, scale-dependent phenomena to uncertain parameters. The constitutive laws we use remain empirical. Furthermore, the interaction between input and output uncertainty creates non-linear epistemological

effects. Small uncertainties in input parameters can produce disproportionately large uncertainties in predicted outcomes, not just through sensitivity analysis but through qualitative changes in failure mode. A rock mass that appears stable under one set of assumptions within our uncertainty bounds might fail catastrophically under a slightly different but equally plausible set of assumptions.

5.2. Professional Practice vs. Uncertainty and Radical Uncertainty

This distinction between uncertainty and radical uncertainty has profound implications for engineering practice. The standard paradigm of "reduce uncertainty through better measurement" has fundamental limits in rock engineering practice. No amount of additional boreholes can eliminate the fact that we are sampling a three-dimensional, heterogeneous geological body through one-dimensional windows. No improvement in laboratory testing can resolve scale-dependent mechanisms where small specimens do not represent large rock masses. And no refinement of numerical models can overcome our uncertainty about both the inputs we provide and the mechanisms by which those inputs are translated into outputs. As illustrated earlier in Figure 2, our models attempt to predict rock mass behaviour as if unknown knowledge follows the identical distributions we have established for known knowledge. But if we cannot fully comprehend the present state of a rock mass, our predictions about its future behaviour rest on even more unstable foundations.

5.3. Professional Practice vs. Linguistic Imprecision

Linguistic imprecision compounds the challenges discussed in previous Sections. Claims of "accurate rock mass strength" or "accurate determination" of empirical parameters pervade rock engineering practice and the literature. Yet claims of accurate rock mass strength require closeness to an actual value that cannot be measured a priori; at best, we can model rock mass behaviour and derive an apparent strength, but this approach encounters the same validation challenges described in Section 3. Where the "true value" of geological parameters cannot be determined independently, claims of accuracy represent a category error. We cannot be accurate about quantities that lack objective physical meaning or cannot be measured in the field (e.g., rock bridges [4]).

The term "reliable" adds another layer of complexity. Reliability refers to a system or process that consistently yields predictable and repeatable results. Critically, our design methods can be reliable even if they are not entirely accurate. For example, empirical strength formulae are often reliable because engineers using the same formula and inputs consistently obtain the same result. Yet, as discussed in Section 3, empirical relationships carry uncertainty due to being extrapolated beyond the limited calibration data and specific field conditions. Figure 14 maps this problem in a reliability–uncertainty space. At the beginning of a project, rock engineering practice occupies a space between the upper-right and upper-left quadrants (R_1U_1 and R_1U_2, respectively), characterized by varying reliability and high uncertainty. In principle, given a set of initial estimates, we can reliably determine the rock mass strength and, with it, calculate a factor of safety; however, the validity of our assumptions means that this reliable calculation may be consistently incorrect. The ideal state (quadrant R_2U_2, lower-right) requires both procedural consistency and epistemic confidence in the form of adequate data access and validated assumptions. Our cone of rock engineering knowledge will not capture the entire set of uncertainty, and radical uncertainty persists as a second layer superimposed on the reliability–uncertainty space.

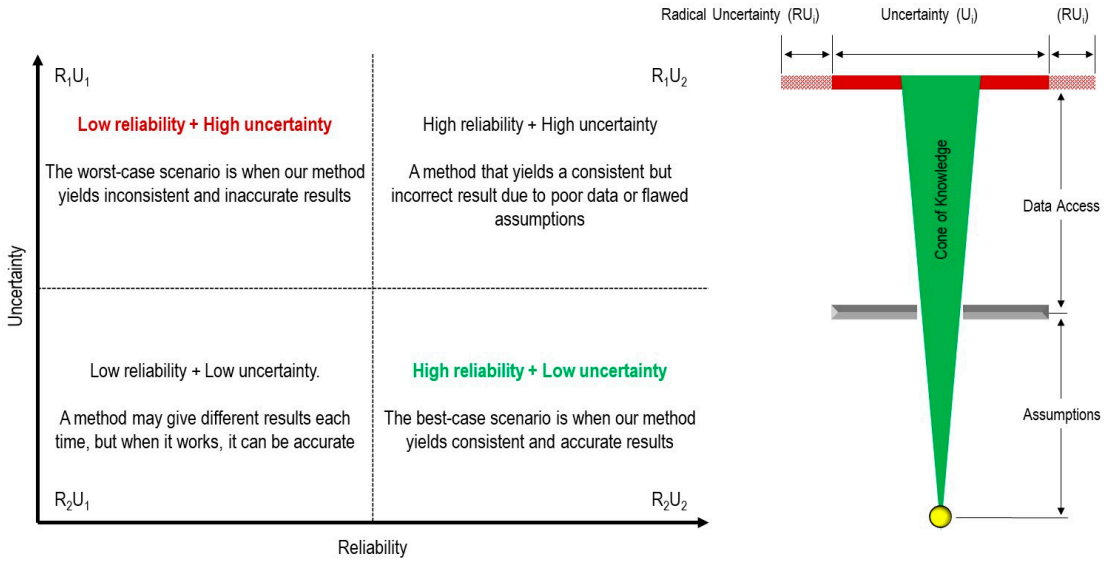


Figure 14. Reliability–uncertainty space. The cone of knowledge illustrates how uncertainty diminishes with access to data. However, radical uncertainty persists beyond data-driven reduction.

5.4. The Algorithmic Amplification of Uncertainty

Professional and research practices are increasingly seeking to add an even more complex layer through AI-driven characterization methods, potentially compounding rather than resolving these foundational epistemological challenges. For instance, when datasets used for training purposes consist of subjectively derived parameters, AI systems inherit our epistemic uncertainties and systematic biases. Furthermore, due to their inherent nature, many AI systems can mask the subjective foundations of their predictions, creating an illusion of objectivity that may be more dangerous than an explicit acknowledgment of uncertainty.

While theoretically we accumulate vast databases of data, rework our empirical correlations, and develop increasingly sophisticated numerical models, each apparent advance in knowledge can reveal new depths of geological complexity, and it does not protect our design from radical uncertainty.

Our numerical models represent the manifestation of what we believe we know. Yet their inability to predict unknown unknowns (radical uncertainty) suggests that our confidence in computational tools may itself be a form of cognitive bias. When dealing with incompletely validated models and methods, encoding fundamentally limited human knowledge into systems that appear cognitively superior to humans creates an epistemological illusion, in which algorithmic complexity appears to transcend the cognitive limitations in the underlying models and methods.

6. Conclusions

This paper has examined significant challenges, including the conflation of calibration and validation (Section 2) and the use of empirical relationships that have not been validated against field-scale experiments (Section 3). Additionally, the paper has discussed radical uncertainty and the limitations of operational definitions (Section 4). These challenges create profound professional implications (Section 5).

Rock engineering practice faces both practical difficulties (e.g., the challenges of large-scale testing of jointed rock masses) and an understandable professional reluctance to undertake studies whose outcomes might require a fundamental reassessment of established practice. The absence of systematic validation for commonly adopted design methods and practices may suggest an inadvertent approach that favours maintaining

established methods over subjecting them to the same validation protocols demanded of new approaches. Validation protocols should apply uniformly. If new methodologies require validation across diverse conditions, established methods should also be subjected to ongoing empirical scrutiny and not accepted on an a priori basis, particularly when they are applied beyond their original context.

Rock mass strength manifests as a range of conditional responses that vary with scale, stress path, and time-dependent processes. This range reflects not only measurement uncertainty but also the inherent variability of natural rock mass systems. A critical challenge underlies all rock mass modelling, as it remains fundamentally impossible to verify a priori whether our results accurately represent the future. Unlike laboratory specimens, which can be subjected to strength testing at scales relevant to engineering applications, we cannot directly test rock masses or build prototypes of our excavations. Even when physical models are constructed at considerable expense, as in the case of the model used prior to the Vajont Dam disaster of 1963, they may offer no advantage over numerical models when founded on incorrect assumptions.

On this basis, we must acknowledge that rock engineering practice necessarily operates under a standard of "balance of probabilities". Given the nature of radical uncertainty, professional practice should evaluate methods not by whether they eliminate uncertainty, but by whether they represent reasonable approaches to managing it. The path forward requires what might be termed "epistemological integrity", i.e., aligning our claims with our actual knowledge, our professional communications with our methodological limitations, and our validation demands with consistent principles. This does not mean abandoning current practice but rather practising with explicit acknowledgment of its empirical foundations.

The implications of the arguments raised in this paper extend beyond theoretical concerns. As mining operations advance to greater depths and into more complex geological environments, we risk amplifying epistemological errors on an unprecedented scale. Indeed, algorithms learning from subjectively derived parameters and unvalidated correlations will encode and perpetuate our current misunderstanding of uncertainty. While this paper is not about algorithms and AI (artificial intelligence), the growing adoption of AI-assisted design tools in rock engineering demands critical attention, as we risk encoding and amplifying the contradictions and knowledge limitations already embedded in design methods commonly accepted by researchers and practitioners. Computational sophistication risks creating false confidence in parameters that remain fundamentally non-unique and empirically unvalidated.

Finally, we need to acknowledge that rock engineering practice demonstrates remarkable success in managing uncertainty through multiple defensive layers: conservative empirical correlations, factors of safety, observational methods, and adaptive management. These pragmatic approaches work precisely because practitioners implicitly recognize that radical uncertainty exists, as unexpected geological conditions or failure mechanisms may emerge despite our best efforts at characterization.

Author Contributions: Conceptualization, D.E. and S.K.A.; methodology, D.E. and S.K.A.; formal analysis, D.E. and S.K.A.; investigation, D.E. and S.K.A.; resources, D.E.; data curation, D.E. and S.K.A.; writing—original draft preparation, D.E.; writing—review and editing, D.E. and S.K.A.; visualization, D.E.; supervision, D.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created in this study. The modelling results presented in Section 3 are based on papers by the first author. These are included in the references.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

GSI	Geological strength index
RMR	Rock mass rating
DFN	Discrete fracture network
SRM	Synthetic rock mass
REV	Representative elementary volume
AI	Artificial intelligence

References

1. Kay, J.; King, M. *Radical Uncertainty: Decision-Making Beyond the Numbers*; WW Norton: New York, NY, USA, 2020; p. 384.
2. Dougherty, E.R. *The Evolution of Scientific Knowledge: From Certainty to Uncertainty*; SPIE Press: Bellingham, WA, USA, 2016.
3. Einstein, A. Autobiographical notes. In *Albert Einstein: Philosopher Scientist*; Schilpp, P.A., Ed.; Harper and Row: New York, NY, USA, 1959.
4. Elmo, D.; Stead, D. The Role of Behavioural Factors and Cognitive Biases in Rock Engineering. *Rock Mech. Rock Eng.* **2021**, *54*, 2109–2128. [CrossRef]
5. Popper, K. *The Logic of Scientific Discovery*; Routledge: Abingdon-on-Thames, UK, 1959; p. 545.
6. Yang, B.; Mitelman, A.; Elmo, D.; Stead, D. Why the future of rock mass classification systems requires revisiting its empirical past. *Q. J. Rock Eng. Hydrogeol.* **2021**, *55*, qjegh2021–qjegh2039.
7. Yang, B. Examining the Reliability of Integrating Machine Learning with Rock Mass Characterization and Classification Data. Doctoral Thesis, University of British Columbia, Vancouver, BC, Canada, 2024.
8. Sivapalan, M.; Bloschl, G.; Zhang, L.; Vertessy, R. Downward approach to hydrological prediction. *Hydrol. Process.* **2003**, *17*, 2101–2111. [CrossRef]
9. Hoek, E.; Brown, E.T. *Underground Excavations in Rock*; Institution of Mining and Metallurgy: London, UK, 1980.
10. Hoek, E.; Brown, E. Practical estimates of rock mass strength. *Int. J. Rock Mech. Min.* **1997**, *34*, 1165–1186. [CrossRef]
11. Hoek, E. Strength of rock and rock masses. *ISRM News J.* **1994**, *2*, 4–16.
12. Marinos, V.; Carter, T. Integrating GSI and mi for Reliable Rockmass Strength Estimation: Integrating GSI and mi. *Rock Mech. Rock Eng.* **2025**, *58*, 11217–11260. [CrossRef]
13. Cai, M. Practical Estimates of Tensile Strength and Hoek–Brown Strength Parameter mi of Brittle Rocks. *Rock Mech. Rock Eng.* **2010**, *43*, 167–184. [CrossRef]
14. Elmo, D. Evaluation of a Hybrid FEM/DEM Approach for Determination of Rock Mass Strength Using a Combination of Discontinuity Mapping and Fracture Mechanics Modelling, with Particular Emphasis on Modelling of Jointed Pillars. Doctoral Dissertation, University of Exeter, Exeter, UK, 2006.
15. Elmo, D.; Yang, B.; Stead, D.; Rogers, S. A discrete fracture network approach to rock mass classification. In *Challenges and Innovations in Geomechanics, Proceedings of the 16th International Conference of IACMAG-Volume 1, Turin, Italy, 5–8 May 2021*; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; pp. 854–861.
16. Barton, N. *Reflections on Unrealistic Continuum Modelling*; NB&A: Oslo, Norway, 2025; 30p.
17. Pine, R.; Harrison, J.P. Rock mass properties for engineering design. *Q. J. Eng. Geol. Hydrogeol.* **2003**, *36*, 5–16. [CrossRef]
18. Yang, B.; Elmo, D. Why engineers should not attempt to quantify GSI. *Geosciences* **2022**, *12*, 417. [CrossRef]
19. Bieniawski, Z.T. Rock mass classification in rock engineering. In *Exploration for Rock Engineering*; Bieniawski, Z.T., Ed.; Balkema: Cape Town, South Africa, 1976; pp. 97–106.
20. Barton, N.; Lien, R.; Lunde, J. Engineering classification of rock masses for the design of tunnel support. *Rock Mech.* **1974**, *6*, 189–236. [CrossRef]
21. Elmo, D.; Zoorabadi, M. Examining the case for accuracy and precision when determining the Hoek-Brown parameters. *Rock Mech. Rock Eng.* 2025, *under review*.
22. Hadjigeorgiou, J.; Harrison, J.P. Uncertainty and sources of error in rock engineering. In Proceedings of the 12th ISRM International Congress on Rock Mechanics, Beijing, China, 16–21 October 2012; pp. 2063–2067. [CrossRef]
23. Bewick, R.P.; Elmo, D. Size effect and rock mass strength. *Can. Geotech. J.* **2025**, *62*, 1–18. [CrossRef]
24. Esmaieli, K.; Hadjigeorgiou, J.; Grenon, M. Estimating geometrical and mechanical REV, based on synthetic rock mass models at Brunswick Mine. *Int. J. Rock Mech. Min. Sci.* **2010**, *47*, 915–926. [CrossRef]
25. Stavrou, A.; Vazaios, I.; Murphy, W.; Vlachopoulos, N. Refined approaches for estimating the strength of rock blocks. *Geotech. Geol. Eng.* **2019**, *37*, 5409–5439. [CrossRef]

26. Li, Y.; Wang, R.; Chen, J.; Zhang, Z.; Li, K.; Han, K. Scale dependency and anisotropy of mechanical properties of jointed rock masses: Insights from a numerical study. *Bull. Eng. Geol. Environ.* **2023**, *82*, 114. [CrossRef]
27. Yiouta-Mitra, P.; Dimitriadis, G.; Nomikos, P. Size effect on triaxial strength of randomly fractured rockmass with discrete fracture network. *Bull. Eng. Geol. Environ.* **2023**, *82*, 8. [CrossRef]
28. Bear, J. *Dynamics of Fluids in Porous Media*; Courier Corporation: North Chelmsford, MA, USA, 2013.
29. Pinto da Cunha, A. *Scale Effects in Rock Masses*; CRC Press: London, UK, 1993; p. 366.
30. Erharter, G.; Elmo, D. Is Complexity the Answer to the Continuum vs. Discontinuum Question in Rock Engineering? *Rock Mech. Rock Eng.* **2025**, *58*, 12695–12713. [CrossRef]
31. Bridgman, P.W. *The Logic of Modern Physics*; Macmillan: New York, NY, USA, 1927.
32. Chang, H. Operationalism. In *The Stanford Encyclopedia of Philosophy*; The Metaphysics Research Lab, Stanford University: Stanford, CA, USA, 2019.
33. Read, J.; Stacey, P. *Guidelines for Open Pit Slope Design*; CSIRO Publishing: Melbourne, VIC, Australia, 2009.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.