Applications of Information Theory in Rock Engineering

B Yang¹, D Elmo¹, and D Stead²

¹University of British Columbia, NBK Institute of Mining Engineering, Vancouver, Canada
²Simon Fraser University, Vancouver, Canada

delmo@mining.ubc.ca

Abstract. Rock engineering relies heavily on empirical systems to identify significant parameters influencing rock mass behaviour. The empirical and inductive nature of rock engineering design is such that it is not possible to eliminate uncertainty. One way of managing uncertainty during the design process is by collecting good quality data in a standardized and objective manner. However, difficulties arise when defining and determining what constitutes good quality data. We believe that information theory and the concept of Shannon’s entropy could be effectively used to better audit rock engineering data. This paper builds on established concepts by expanding and refining the application of information theory to rock mass classification systems, specifically the rock mass rating and the Q-system. One of the objectives is to provide and showcase a method whereby information auditing is used to flag uncertain (or poor quality) data. In the future it is not difficult to envision data collection processes that include improved core logging and data processing where imaging technologies are coupled with machine learning processing capability. Such an approach requires more quantitative and objective rock mass descriptions; in this context it easy to appreciate the role that information theory might have in the future in rock engineering.

1. Introduction
Uncertainty in rock engineering is unavoidable, whether geological, parameter, model, and/or human uncertainty. As such, it is imperative that rock engineers develop methods to manage uncertainty during the design process, especially as the digitalization trends increase. One such method is to collect data and then quantitatively determine what constitutes good quality data. Information theory, such as the concept of Shannon’s entropy [1], can be applied to rock engineering as a way to better audit rock engineering data and determine the quality of data. The field of information theory was originally developed for communications and has been extended to other fields, such as computer science and machine learning; however, its use in rock engineering has been limited. This paper will (i) provide a review of information theory concepts relevant to rock engineering and (ii) build upon the concepts and examples introduced in [2] by providing and showcasing a method whereby information auditing and assessment are used to flag uncertain (or poor quality) rock mass classification values, specifically the rock mass rating (RMR) and Q-system.
2. Review of information theory concepts

The field of information theory began with Claude Shannon’s 1948 paper titled “A Mathematical Theory of Communication”, where he introduced the fundamental concepts of information theory in the context of communication. One of the goals of the paper was to “find a measure of how much choice is involved in the selection of the event or how uncertain we are of the outcomes” [1]. This measure was defined as entropy, and is expressed as:

\[ H(X) = \sum_{i=1}^{n} p_i \log_2 \left( \frac{1}{p_i} \right) \]  

(1)

where \( H(X) \) is the entropy of a set of \( n \) possible events of random variable \( X \) whose probabilities of occurrence are \( p_1, p_2, \ldots, p_n \). As defined in Equation (1), entropy becomes a measure of the uncertainty associated with the random variable \( X \) and a higher entropy indicates that there is more uncertainty. Note that \( H(X) \) is not a function of the random variable \( X \); it is strictly a function of probability and the \( X \) denotes that it is the entropy of random variable \( X \).

The properties of entropy are outlined in [1] and are summarized below:

1. If one of the probabilities \( p_i \) is 1 and the remaining are 0 in a set of possible events, there is no entropy because we are certain of the outcome (i.e., \( H = 0 \)).
2. \( H \) can only be positive.
3. The most uncertain situation occurs when the probabilities in a set of possible events are the same (i.e., in a set of \( n \) events and subsequently \( n \) probabilities, \( H \) is maximum when the probability of each event is \( \frac{1}{n} \)). The maximum value of \( H \) is equal to \( \log(n) \).
4. The joint entropy of two discrete random variables \( X \) and \( Y \), where \( X \) has \( m \) possibilities and \( Y \) has \( n \) possibilities, is shown in Equations 2 and 3:

\[ H(X,Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} p_{(i,j)} \log_2 \left( \frac{1}{p_{(i,j)}} \right) \]  

(2)

\[ H(X,Y) \leq H(X) + H(Y) \]  

(3)

Where \( p_{(i,j)} \) is the joint probability of occurrence of \( i \) for random variable \( X \) and \( j \) for random variable \( Y \). If \( X \) and \( Y \) are independent, their joint entropy is the summation of their individual entropies (the maximum case in Equation 3).

5. Building off of point 3 above, “averaging” (or equalizing) the probabilities \( p_i \) increases \( H \).
6. The conditional entropy between two discrete random variables \( X \) and \( Y \) (similar to point 5) that are not necessarily independent measures how uncertain we are of \( Y \) when we know \( X \). The conditional entropy of \( Y \), which is the average entropy of \( Y \) for each value of \( X \), is shown in Equation (4) below.

\[ H_X(Y) = \sum_{ij} p_{(i,j)} \log_2 \left( \frac{1}{p_{(i,j)}} \right) = H(X,Y) - H(X) \]  

(4)

Where \( p_{(i,j)} = \frac{p_{(i,j)}}{\sum p_{(i,j)}} \). In other words, the conditional entropy of \( Y \) is also the difference between the conditional entropy of \( X \) and \( Y \) and the entropy of \( X \).

7. Having knowledge of random variable \( X \) will never increase the entropy (uncertainty) of random variable \( Y \). Rather, knowledge of random variable \( X \) will either a) decrease the entropy (uncertainty) of random variable \( Y \) if \( X \) and \( Y \) are not independent of each other or b) remain unchanged if \( X \) and \( Y \) are independent of each other.
Based off the above properties of $H$, the entropy of a system of variables that are all independent of one another is the summation of the entropy of each variable.

3. Information theory in rock engineering design

3.1. Review of previous work

The use of information theory in rock engineering design has so far been limited. [2] published a paper outlining various uses of information theory in rock engineering design, including the analysis of the uncertainty remaining about a variable being estimated with an empirical relation. They applied information theory to determine the most efficient site investigation methodology for rock mass characterization. The underlying theme of their applications is using Shannon’s concept of entropy to quantify the level of uncertainty in geotechnical parameters as a method of information auditing.

With respect to rock mass characterization, [2] provided an example of applying entropy to Bieniawski’s rock mass rating (RMR) system to determine the information content (i.e., how much uncertainty) of each RMR value. Using a Monte Carlo analysis, they were able to randomly generate combinations of RMR and compute their entropy. This allowed for the generation of a plot of the maximum and minimum entropy value of each RMR, which they recommended should be used as a reference for checking the entropy of a specific RMR combination against the overall range of entropies. Of significance is the variability in the maximum and minimum entropy values for RMR values; the variability is much greater for RMR values between 20 and 80. This can be attributed to the higher number of combinations (or pathways) for these RMR values.

While information theory has had limited direct use in rock engineering design, its prevalence in machine learning has resulted in more frequent indirect use as rock engineering moves to digitalization. Shannon’s entropy is commonly found in loss functions – a measure of how good a prediction is using a supervised machine learning model (i.e., how far the model prediction is from its label). Examples include cross-entropy and Kullback-Leibler divergence, both of which are employed in neural networks [3]. Additionally, a decision tree can use “information gain” to control how it makes decisions at its nodes; using “information gain” tells the decision tree to go with the split that decreases the entropy of the label before and after the split [4].

The increased use of machine learning in rock engineering has highlighted the importance of data quality and the need for a unified and quantitative method of determining the quality of our data. One solution is to utilize the concept of entropy from information theory to quantify the level of uncertainty in our measurements as done by [2].

3.2. Rock mass classification systems and implications for machine learning

Building upon the work by [2], a site-specific method for auditing rock mass classification values, specifically the rock mass rating (RMR) system proposed by [5], [6], and [7] and the Q-system by [8], is proposed in the following section. The RMR system proposed by is the summation of ratings given to 5 parameters:

1. Strength of intact rock
2. Rock quality designation (RQD)
3. Spacing of discontinuities
4. Condition of discontinuities
5. Groundwater

The difference between the 1973, 1976, and 1989 versions of RMR lie in the weights of the 5 parameters and the ratings used. The RMR value can also be adjusted for the discontinuity orientation with respect to the engineering structure (tunnels/mines, foundations, slopes) being designed.

The Q-system proposed by [8] is calculated using Equation 5 below:
Where RQD is the rock quality designation, \( J_n \) is the joint set number, \( J_r \) is the joint roughness number, \( J_a \) is the joint alteration number, \( J_w \) is the joint water reduction factor, and \( SRF \) is the stress reduction factor.

As outlined by [2], there are several ways to obtain the same RMR value. This also applies to the Q-system; each RMR or Q value has several different combinations of the ratings of their respective parameters. Because both the RMR and Q-system are a system of variables, the entropy of each RMR or Q value is the summation of the entropy of their respective parameters. For RMR and Q values with multiple combinations, they will have multiple different entropy values corresponding to a specific combination.

The site-specific method for auditing RMR and Q values is outlined below:

1. Determine all possible combinations of RMR and Q, keeping in mind the geological constraints and the discrete nature of the ratings of the parameters.
2. Determine the discrete probability distributions of the ratings for each parameter in the site data. Update the combinations of RMR and Q determined from the previous step by removing combinations with a probability of zero.
3. Determine the entropy of each classification system parameter based off of the probabilities determined in the previous step.
4. Add the entropy of the parameters of each classification system parameter to determine the overall entropy of that specific classification value. This step has been simplified by assuming that all parameters are independent of one another, resulting in the maximal entropy for that combination.
5. Once the entropy for all combinations has been determined, the maximum and minimum entropy for each combination are plotted, similar to what was done in Mazzoccola et al. (1997).
6. The entropy of the site data can be compared against the plot of maximum and minimum entropy for all possible combinations at that site to determine if it plots close to the maximum or minimum entropy. Those site data-points that plot close to the maximum entropy have a higher uncertainty and should be examined closely and used with caution.

An example analysis using field mapping data was performed and is shown below. Combinations for RMR76 and Q were generated using a Python script, keeping in mind the relationship between RQD and discontinuity spacing, the relationship between RQD and the number of joint sets, the rock wall contact, and the discrete nature of the classification ratings.

Figure 1 show the distributions of all possible combinations of RMR76 and Q with respect to their respective rock mass classes. Similar to the combination analysis in [2], the distribution of RMR follows a normal distribution, with “Fair” and “Good” rock mass classes having the highest number of combinations. Unlike RMR, the Q-system follows a negative exponential distribution, with “Extremely poor” and “Very poor” rock mass classes having the highest number of combinations. Using site-specific data to determine the discrete probabilities of the ratings of each RMR and Q parameter, the plots in Figure 1 are then updated to reflect the site conditions by showing the distribution of all possible combinations of RMR76 and Q for this example site, shown in Figure 2. The site-specific data used were obtained from window mapping and include detailed rock mass classification information. The next step is to calculate the entropy of each parameter rating of all possible RMR76 and Q combinations for the example site by using Equation 1. The entropy of each RMR76 and Q value is the summation of the entropy of their respective parameter ratings.
Following the calculation of the entropy of the RMR76 and Q values, the maximum and minimum entropy for each value are plotted, as shown in Figures 3 and 4. The entropy of RMR and Q values obtained during the site investigation can be checked against these plots to determine their entropy relative to the maximum and minimum entropy for that specific classification value. RMR or Q values with a high entropy (close to the maximum entropy for that specific value) have a high uncertainty; at least one of their parameters has a high entropy (unexpected value). These unexpected RMR and Q values are “flagged” and should be 1) re-examined to identify the parameter(s) with high entropy and determine if the unexpectedness is due to human/lab errors instead of unexpected rock mass conditions and 2) treated with caution if the high entropy is due to unexpected rock mass conditions.

Figures 3 and 4 on the following page show that the greatest difference in entropy for an RMR or Q value corresponds to those values with the greatest number of combinations. For RMR, the greatest variation in entropy can be found in the “Fair” rock mass class, while the greatest variation in entropy for Q can be found in “Poor” and “Fair” rock mass classes. Additionally, both Figures 3 and 4 show oscillations in both the minimum and maximum entropy values for an RMR or Q value; however, Figure 4 (Q values) shows much greater oscillations. The oscillations are greatest for entropy values of Q values between 1 – 10, which corresponds to the same range with the greatest difference in entropy. These oscillations indicate that even within rock mass classes, certain rock mass classification values are much more uncertain than the others.
Incorporating this information auditing process during the design process could make it easier to flag uncertain data, which is becoming increasingly important as rock engineering moves to machine learning and other advanced data analysis techniques. One of the major limitations of any computer model, especially machine learning models, is the data quality. In the context of machine learning, a model trained with poor quality data will output poor quality results; the success and quality of the model is dependent on the quality of the data used. As a result, it is imperative that rock engineers focus on collecting better quality data in a less subjective manner. The information auditing process outlined above provides a method to quantify uncertainty, making it easier to find and resolve uncertain and poor quality data during pre-processing.

Additional benefits of using entropy to quantify the uncertainty in rock mass classification values include identifying “extreme” (i.e., unexpected) rock mass conditions and shifting the focus from the classification value to the actual parameters in classification systems.
3.3. Limitations in rock engineering

The method and analysis outlined in the previous section are only applicable when RMR and the Q-system are used in a fully standardized manner consistent with how they were originally devised. Variations in how RMR and the Q-system are applied – for example treating RMR rating values as continuous instead of discrete, interpolating between ratings, or using company guidelines that introduce modifications to the original system – render the entropy approach, along with machine learning algorithms that use it, void. The fault lies with the inherent subjectivity which is present from rock mass characterisation through rock mass classification. A clear example is given by personal preferences that indicates which version of the RMR table is to be used; under these circumstances different entropy values would be determined for the same rock mass when using RMR76 and RMR89. Because of the entropy is a measure of uncertainty, in this case the difference between the entropy values associated to RMR76 and RMR89 (for a given rock mass) would then be a manifestation of human uncertainty, impossible to remove independently of the quality of the data that have been collected.

4. Conclusions

The shift to digitalisation in rock engineering has highlighted the importance of data quality and the ability to determine the quality of the data in a quantitative manner. This paper has presented an updated approach for auditing of rock mass classification values to help identify uncertain (poor quality) data by using the concept of Shannon’s entropy from information theory. By quantifying the uncertainty of rock mass classification values, uncertain data can be flagged to be (i) re-examined to identify the parameter(s) with high entropy and determine if the unexpectedness is due to human/lab errors instead of unexpected rock mass conditions and (ii) treated with caution if the high entropy is due to an unexpected rock mass. This method also emphasizes the importance of focusing on the parameters in rock mass classification systems rather than solely on the classification value, allowing for a better understanding of the rock mass quality. As data collection processes improve and begin to include imaging technologies and machine learning, it is necessary for rock mass descriptions to become more quantitative and less subjective; in this context it easy to appreciate the role that information theory might have in future rock engineering.

References

[1] Shannon CE 1948 *The Bell System Technical Journal* 27 p 379
[2] Mazzoccola DF, Millar DL and Hudson JA 1996 *Geotechnical and Geological Engineering* 15 p 145
[3] MacKay DJC 2004 *Information Theory, Inference, and Learning Algorithms* Cambridge University Press
[4] Shalev-Shwartz S and Ben-David S 2014 *Understanding Machine Learning Algorithms: From Theory to Algorithms* Cambridge University Press
[5] Bieniawski ZT 1973 *Transactions of the South African Institution of Civil Engineers* 15 p 335
[6] Bieniawski ZT 1976 *Exploration for Rock Engineering, Proceedings of the Symposium* ed ZT Bieniawski Cape Town 1 p 97
[7] Bieniawski ZT 1989 *Engineering Rock Mass Classifications* New York Wiley
[8] Barton NR, Lien R and Lunde, J 1974 *Rock Mechanics* 6 p 189