

# Probabilistic stability analyses for sedimentary deposits

A Duran Pells Sullivan Meynink Pty Ltd, Australia

## Abstract

*This paper presents several approaches utilised by the author in assessing slope designs, inclusive of Probabilities of Failure, in sedimentary strata. A common issue seen by the author in probabilistic analyses is the use of population statistics, which honour variability in point sampling, but do not reflect variability at the larger scale. This then results in overestimates of the Probability of Failure. Issues in assessing the variability in inputs for analyses are discussed. Two case studies are presented with focus and discussion on use of the appropriate variability in the respective analyses. The cases have considered the scale at which the data is collected, and, critically, the analysis methodology which influences the approach in selection of variability. The case studies have utilised a Monte Carlo approach and use of limit equilibrium stability analysis software. Recent trends in analysis methodology (surface response methodology) and emergence of improvements in software (which allow generation of random fields) suggest the field of probabilistic analysis has matured. However, without careful consideration to the key design parameters, probabilistic analysis may simply serve to provide what appears as more sophisticated results, but which offer no additional value in managing risk for a project.*

**Keywords:** probabilistic analyses, Monte Carlo, sedimentary strata

## 1 Variability in sedimentary deposits

Experience has shown that in sedimentary strata, there is typically a pronounced change in geotechnical character based on parent lithology. As such, review of data is typically best carried out based on key lithologies present at a site. Table 1 presents selected cases of the authors experience with unconfined compressive strength (UCS) test results in sedimentary deposits. The data covers eight different mining sites where testing of fresh rock was subdivided by key lithology. All the selected cases had more than 20 valid tests for each subdivision of lithology.

Figure 1 presents statistics associated with UCS testing of siltstones at Site 1. A key consideration not often appreciated is the difference between the variance in the population and the variance on the mean. The former defines the variability in the population, commonly referred to as the standard deviation (SD), which can be expressed as the unitless coefficient of variation (COV), i.e. the SD divided by the mean. Additional testing aids in better defining the value but does not significantly diminish the variability. The latter defines the uncertainty in the estimate of the mean, and additional testing significantly aids to diminish this uncertainty, albeit the change diminishes slowly, beyond about 20 tests in this example. The reader should note the SD on the mean is simply the square root of variance on the mean.

Table 1 highlights that COVs of 55–80% are typically common. These COVs, which are not uncommon for many mine projects, are distinctly at odds compared to the testing of very uniform rocks by initial researchers, for example the UCS testing of Deere and Miller (1966) and triaxial testing by Hoek (1983). Cargill and Shakoor (1990) provide examples of rock types commonly utilised by researchers with the majority of UCS testing having COVs in the range of 3–13%. To place this into perspective, Hess et al. (2002) present UCS testing from ordinary strength steel used in shipbuilding and presented COVs in the range of 3–11%.

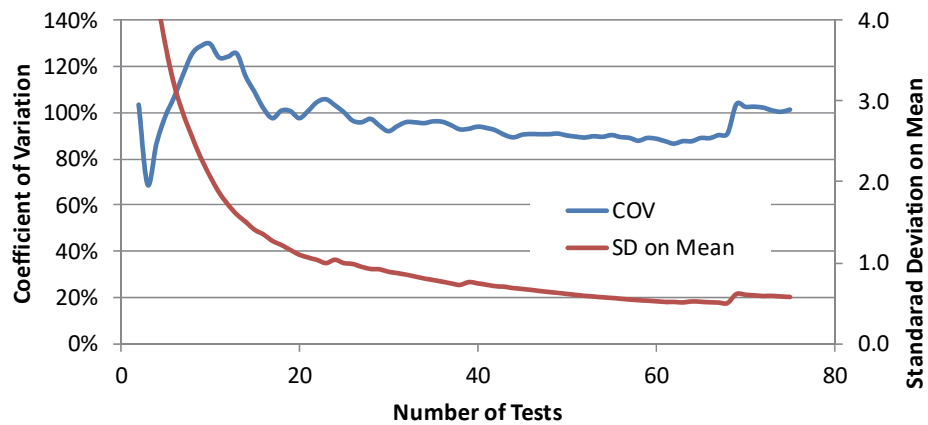


Figure 1 Change in coefficient of variation (COV) and standard deviation (SD) of mean with increasing testing for siltstone associated with Site 1

Table 1 Typical UCS and statistics associated with coal mine lithologies

Location	Key lithology	Mean UCS (MPa)	Number of tests	Standard deviation (MPa)	Coefficient of variation	Standard deviation on mean
Site 1	Siltstone	4.8	77	5.0	103%	0.6
Site 2	<b>Siltstone</b>	<b>17.4</b>	<b>21</b>	<b>15.6</b>	<b>89%</b>	<b>3.4</b>
	<b>Sandstone</b>	<b>35.7</b>	<b>26</b>	<b>26.3</b>	<b>74%</b>	<b>5.2</b>
	Carbonaceous sediments	27.4	68	17.2	63%	2.1
Site 3	Siltstone	33.5	200	19.8	59%	1.4
	Sandstone	52.1	205	32.5	62%	2.3
Site 4	Conglomerate	10	306	9.6	96%	0.5
	Siltstone	15.4	139	11.0	72%	0.9
	Sandstone	16.4	187	11.4	69%	0.8
Site 5	Sandstone	33.9	77	17.3	51%	2.0
Site 6	Coal	14.5	51	9.6	66%	1.3
	Claystone	9.6	40	6.9	72%	1.1
	Siltstone	15.3	173	12.6	82%	1.0
	Sandstone	26	58	24.1	93%	3.2
Site 7	<b>Claystone</b>	<b>9.8</b>	<b>41</b>	<b>9.5</b>	<b>97%</b>	<b>1.5</b>
	Siltstone	28.9	64	22.1	77%	2.8
	Sandstone	56.1	90	26.4	47%	2.8
Site 8	<b>Claystone</b>	<b>11.2</b>	<b>51</b>	<b>14.2</b>	<b>127%</b>	<b>2.0</b>
	Siltstone	23.9	35	13.9	58%	2.4
	Sandstone	34.3	81	19.5	57%	2.2

The studies of Gill et al. (2005a, 2005b) and Fillion and Hadjigeorgiou (2013) can be utilised to assist in determining the data confidence which is suitable for each project stage as indicated by Read (2009). Of the cases provided in Table 1, four (in bold) failed to provide what is considered as suitable for design and construction, with data confidence of the order of 60% even where 50 valid tests had been undertaken.

## 2 Study: Case 1

Case 1 involved slope design for an overburden sedimentary sequence. The sequence comprised nine lithological units; Unit 1 uppermost and Unit 9 the lowest in the sequence. The presence of low angle shears in the shale rich Unit 9, cross-cutting bedding, provided the potential for a block slide mechanism.

To address assessment for the slope design, a comprehensive geotechnical program was undertaken with geotechnical logging of the sequence. Laboratory testing was carried out which included triaxial testing, UCS and Brazilian testing of intact core, direct shear testing of defects and Atterberg limits testing on the shears and infills. Acoustic televiewer (ATV) logging was carried out to assist in understanding the orientation of the low angle shears.

### 2.1 Rock mass strength variability

The geotechnical logging was carried out with logging units based on intervals of core where similar intact strength, weathering/alteration, fracture frequency and character of defects were evident. These intervals of core varied in lengths less than 1 m to the entire length of the intersected lithological unit. Figure 2 presents the geological strength index (GSI) versus logged length for Unit 6.

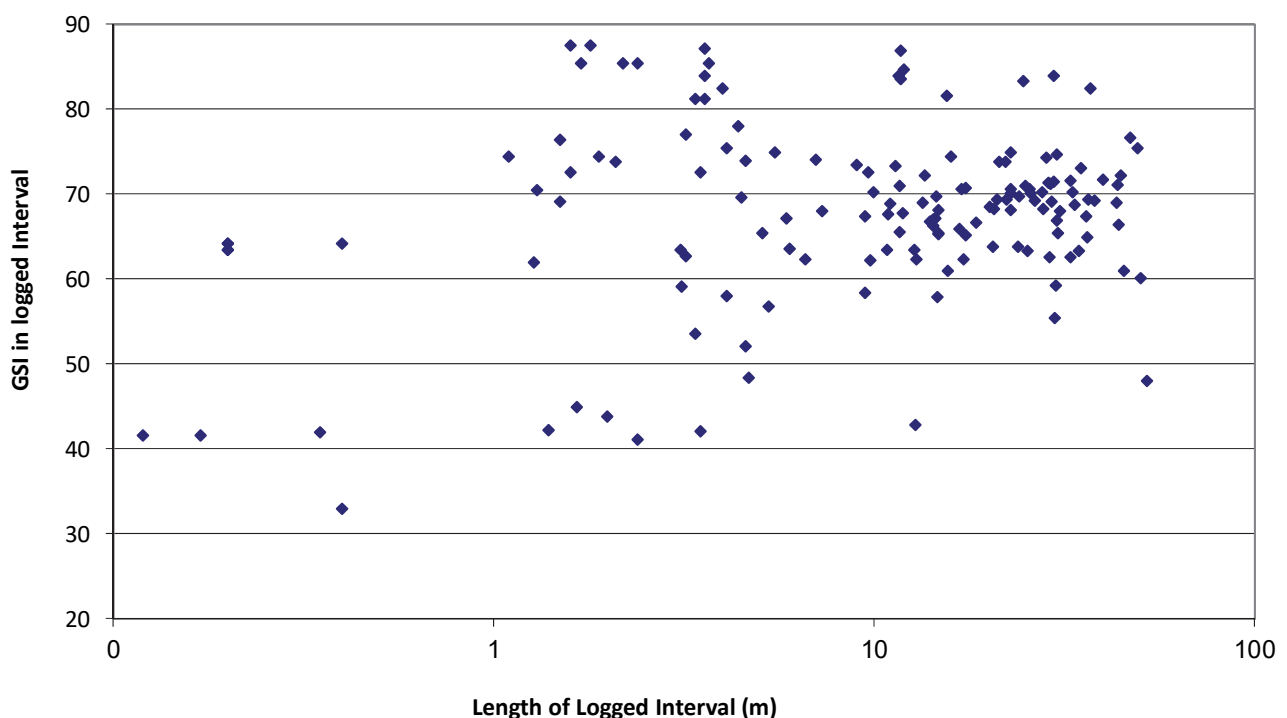


Figure 2 Geological strength index (GSI) in each logging interval plotted relative to length of logged interval

For the slope stability analyses, the manner in which the variability is applied was carefully considered in relation to the above. For the limit equilibrium software being utilised at the time, SLIDE V6 (Rocscience Inc. 2010), Monte Carlo analysis is based on assessing numerous simulations. For each simulation, each entire unit is assigned a single randomised value based on the provided mean and SD. The manner in which the variability is applied has a key bearing on the final estimates of probability. Hence, whilst it is appropriate to utilise the mean for each unit based on the logging and testing, the applied SD is critical.

For the UCS, and other parameters that vary from point-to-point such as density, the range in simulations of the Monte Carlo analysis should reflect the uncertainty in the mean estimate of the parameter rather than cover the range of values in the population. For example, and with reference to sandstone at Site 7 in Table 1, using the population statistics as input into a Monte Carlo analysis would result in UCS values between 29.7 and 82.5 MPa for 65% of the simulations. It would be more appropriate to consider that the variation in the mean estimate of UCS is 53.3 to 58.9 MPa for 65% of the simulations. This latter case provides significantly less variation in the analysis.

The triaxial testing was used to provide estimates of the Hoek–Brown parameter  $m_i$  as well as confirming the UCS test results. Figure 3 presents testing for Unit 7 and highlights a common issue encountered wherein there is a poor fit using RocLab (Rocscience Inc. 2004) or the equations provided by Hoek and Brown (1997). The approaches by Bewick et al. (2011) or Cammack and Duran (2015) can partly assist to reduce the uncertainty in interpreting testing. However, there remains residual uncertainty, partly owing to small datasets but largely driven by heterogeneity introduced in forming tractable spatially coherent geotechnical models. It is questionable whether the full uncertainty in the triaxial testing should be treated as variability in the modelling. The author believes that designers should adopt a variability that reflects the typical character rather than the extreme range. Moreover, one needs to consider that the testing may be biased owing to sample selection with either a tendency to sample weaker and more altered variations within a geotechnical unit wherein better quality rocks or the more competent pieces of core where in poorer quality rocks. This emphasises the need for good sample collection practices with photographs prior to wrapping, strength estimates and sample descriptions to allow appropriate interpretation post testing. For the example in Figure 3, the unit was typified by occasional zones of finer grained material which provided higher UCS values and could be associated with the very high triaxial test results. Ignoring these higher results, a manual fit was provided, with a UCS value in keeping with the mean of UCS testing.

For the rock mass strength, the methodology of Hoek et al. (2002) was utilised and the GSI assessed from the geotechnical core logging. Figure 4 presents GSI assessments for Unit 6 where about 400 m of core was logged and two histograms of GSI are presented. The top histogram provides the distribution of assessed GSI from 38 logging intervals. The data is presented on a length-weighted basis and indicates that about 80 m of core had a GSI between 70 and 75.

It was recognised that for each simulation in the Monte Carlo limit equilibrium analysis, one randomised GSI would be applied to each unit. Hence, it would be appropriate that each simulated GSI be in keeping with the average GSI across the unit. As such, for Unit 6 where 13 boreholes traversed the unit, the mean GSI for each of these traverses was assessed. This assessment is shown in the bottom histogram of Figure 4, and where mean GSI values of 53, 60, 61, 69, 69, 69, 70, 70, 73, 77, 78, 82 and 84 were assessed from the 13 boreholes. During the Monte Carlo analyses, the population SD of nine GSI units was considered as appropriate. It is interesting to note that since the population SD is being utilised, additional drilling will improve but not significantly reduce the estimate of the uncertainty.

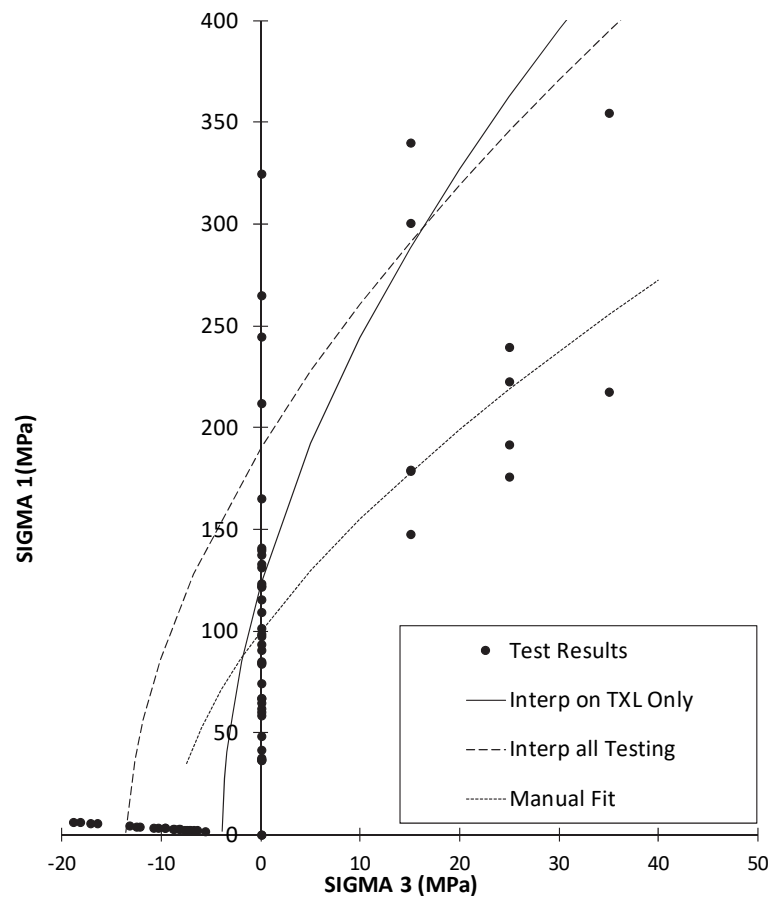


Figure 3 Interpretation of triaxial testing from Unit 7

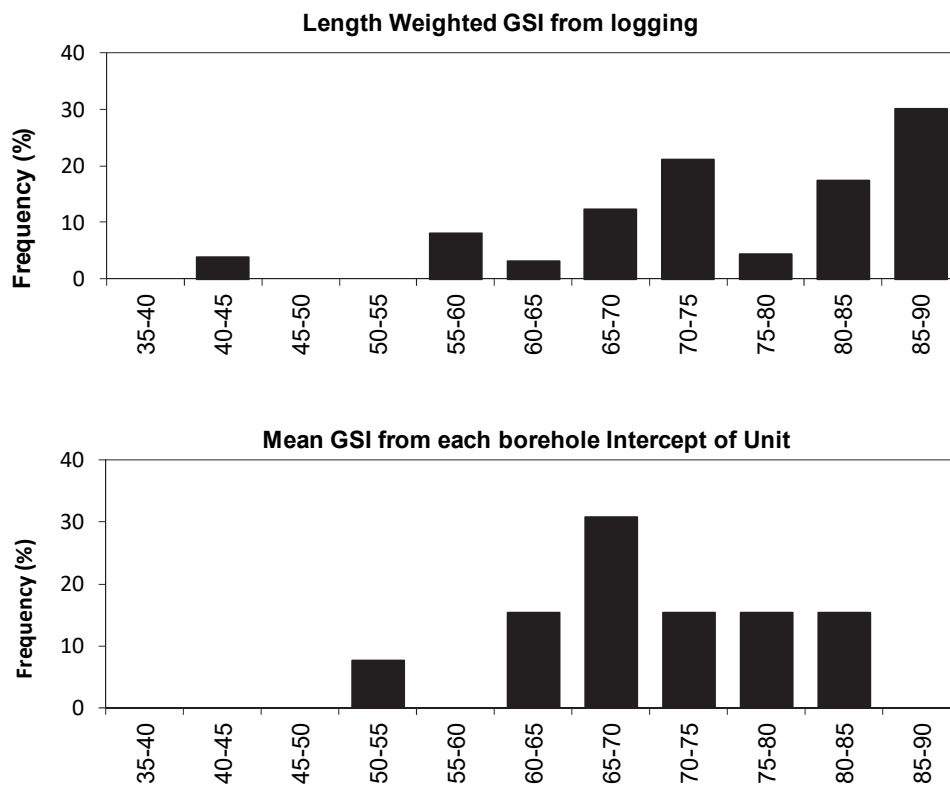


Figure 4 Distribution of geological strength index (GSI) for Unit 6. At top, for each logged interval. At bottom, the mean GSI over each length of borehole intersecting Unit 6

## 2.2 Variability in low angle shears

Figure 5 presents a stereographic projection of the low angle shears from ATV data. The variability in orientation was accounted for in the slope design by taking into consideration the slope aspect (i.e. direction to which the slope faces) and partly using the approach of McMahon (1971). The overall Probability of Failure (PF) of the slope is calculated using Equation 1.

$$PF = \sum (Pf \times Pu) \quad (1)$$

where:

$Pf$  = Probability of Failure for the dip of a given shear.

$Pu$  = probability of undercutting, incremental likelihood that a shear will occur at a given dip.

The  $Pf$  for a given dip of shear plane was assessed through Monte Carlo limit equilibrium stability analyses for a block slide mechanism and with variability applied in rock mass strengths. These  $Pf$  analyses involved 10 analyses with shear plane angles varying from  $0^\circ$  to  $45^\circ$  in  $5^\circ$  increments. A window of  $60^\circ$  was utilised in the McMahon analysis to assess  $Pu$ . An example slope aspect of  $140^\circ$  and a  $60^\circ$  window is indicated in Figure 5. It was recognised that all flat dipping shears (i.e. those with dips less than  $5^\circ$ ) would provide a similar case to that of a horizontal shear and the window in the  $Pu$  was increased to  $180^\circ$ .

PF acceptance criteria of 1–3% was utilised for the overall slope design in keeping with the guidelines of Kirsten (1983). Figure 6 provides the results of the overall analysis and highlighted the sensitivity to groundwater. Significant risk was indicated for selected slope aspects even with proposed depressurisation measures.

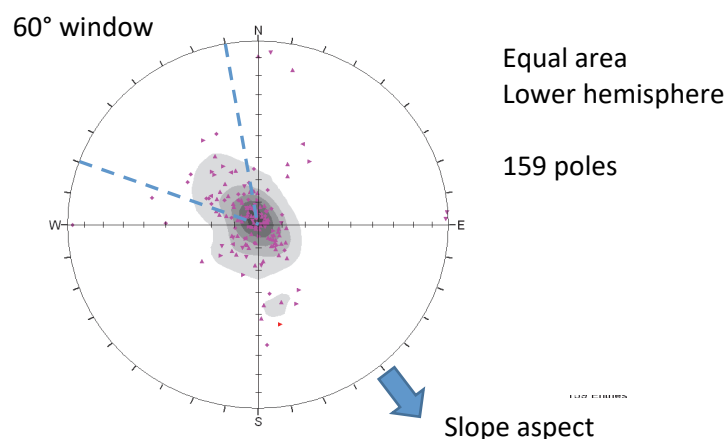


Figure 5 Stereographic projection of low angle shears and concept of slope aspect and window

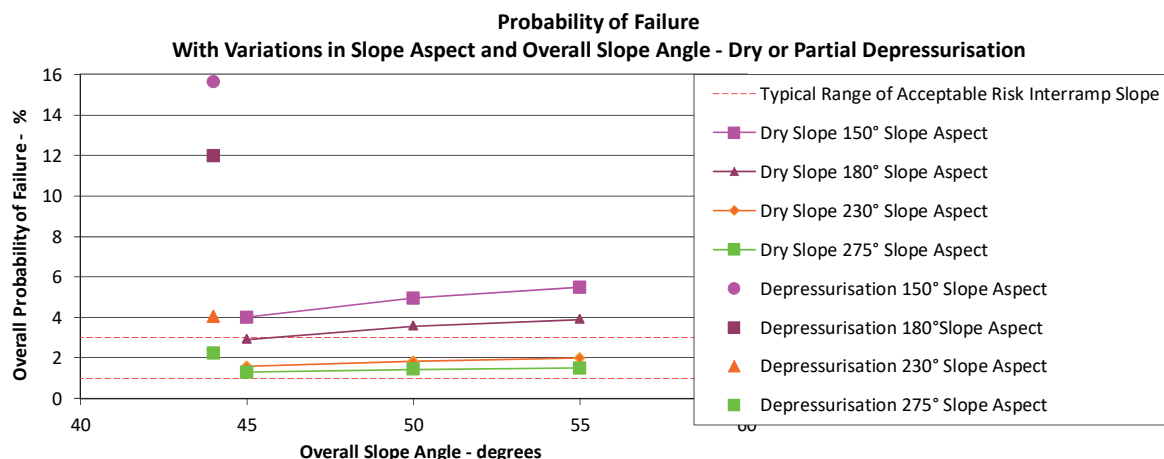


Figure 6 Overall Probability of Failure assessed for Case 1

### 3 Study: Case 2

Case 2 presents what could be termed a typical coal mine highwall with an intercalated sequence of coal seams, claystones, siltstones and sandstones. Unlike the previous case, the overall slope can be seen as comprising a significant number of individual beds. Typically, coal mines provide designations to coal seams which allow correlation within the overall stratigraphic package through geophysical signatures. Whilst the coal seams themselves can remain continuous over significant horizontal distances the inter-burdens between seams may laterally change in lithological character owing to changes in depositional environment. As such, a sandstone bed in one location may vary laterally to a siltstone in the next adjacent borehole. Figure 7 presents the relative proportion of sandstone, siltstone and claystone encountered within inter-burdens in a multi-seam coal deposit. The data highlights that there are some inter-burdens which are predominantly comprised of one lithology but in a broader sense, groups can be defined where individual inter-burden units have similar distribution of lithologies (Figure 8).

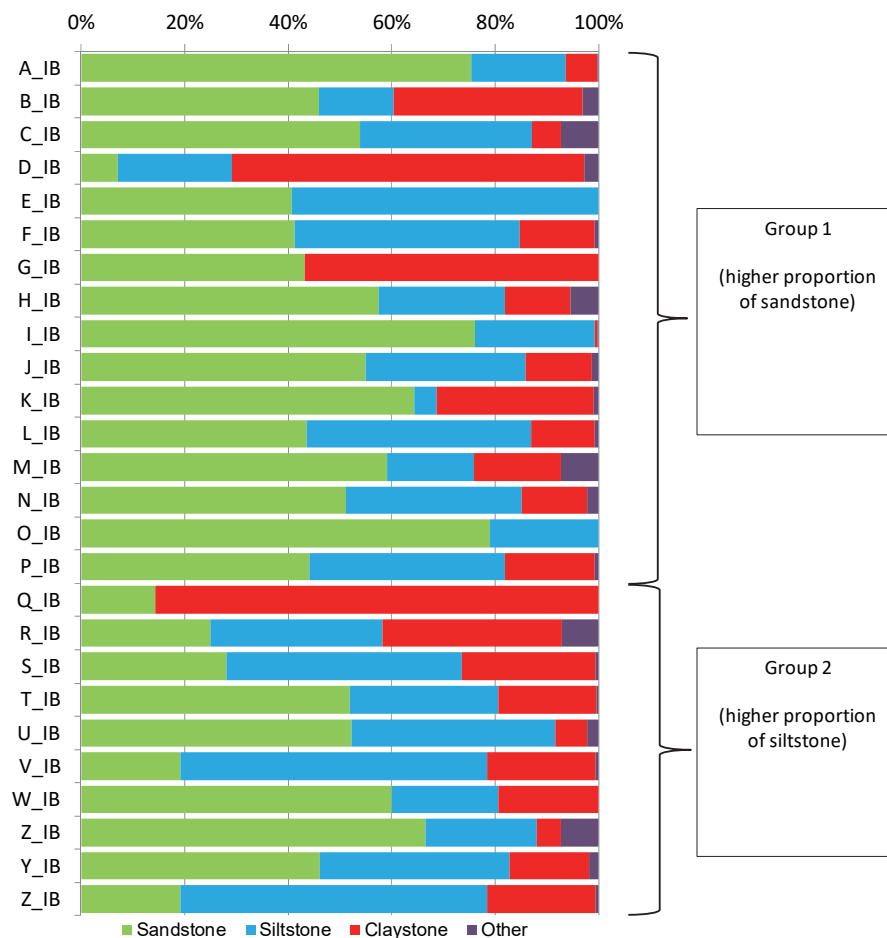
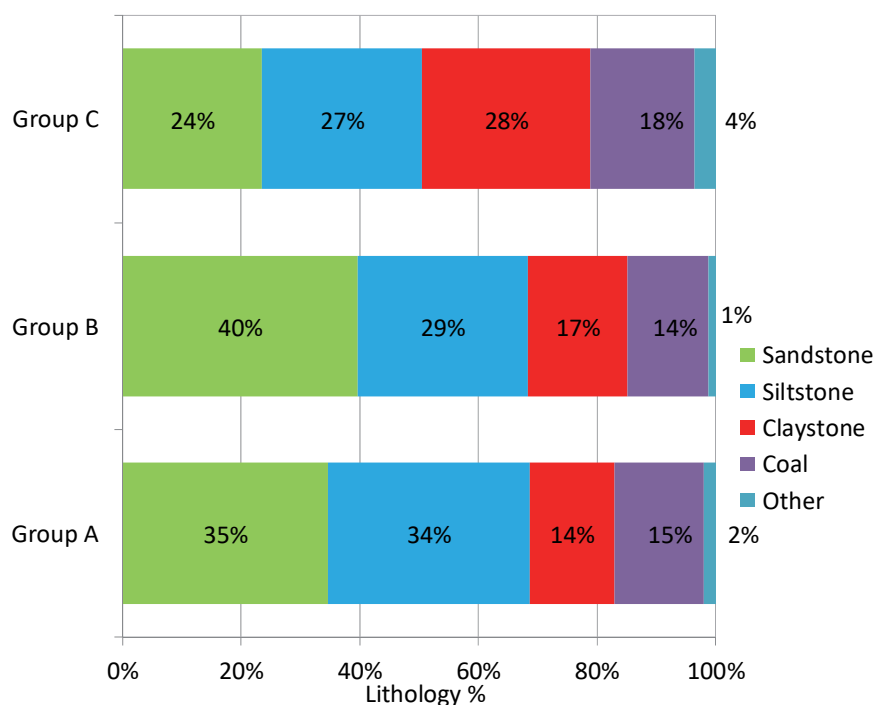


Figure 7 Distribution of lithologies within inter-burdens



**Figure 8** Example of overall distribution of lithologies within stratigraphic groups

For overall scale pit stability, a failure path is likely to crosscut the rock mass with key lithologies intersected in proportion to that seen in each stratigraphic group. As such, in order to simplify stability analyses, the slope can be subdivided by stratigraphic groups and with rock mass strengths assigned to each stratigraphic group in proportion to the key lithologies. Such an approach is appropriate for a deterministic analysis but then brings to question as to how it should be addressed with a probabilistic analysis.

Three approaches are presented to highlight the issues. The first approach has utilised the concept in Case 1 and applied stratigraphic groups. The second approach has utilised a ‘quasi’ detailed stratigraphy. The third approach has blindly utilised the population statistics.

For brevity, this paper only presents a failure mechanism which considers rock mass failure through an overall circular mechanism. The Monte Carlo limit equilibrium stability analyses utilised an ‘overall’ search to assess critical failure mechanism, mean Factor of Safety (FS) and PF. The analyses utilised 200 slices and 100,000 simulations. The groundwater conditions were kept fixed in the analyses.

### 3.1 Approach One

The first approach has considered the proportion of key lithologies in each stratigraphic group. Table 2 provides a summary of the pertinent inputs for estimating rock mass strength parameters for each of the key lithologies. The proportion of key lithology types was assessed on a borehole-by-borehole basis, whilst the GSI was based on the average in each borehole intersecting the stratigraphic group. For both parameters, the population SD is considered appropriate in the stability analyses. This approach is in keeping with that discussed for Case 1. The UCS testing was subdivided by key lithology, and the SD of the mean is considered appropriate at the overall scale. Variability in the parameter  $m_i$  was based on experience by the author with triaxial testing measurements in coal rocks. The example provided in Cammack and Duran (2015) highlighted that minor variations in  $m_i$  can be expected.



Table 2 Parameters Case 2

Group	Lithology	Number of holes	Mean GSI	GSI SD	Proportion (%)	Proportion SD	Mean UCS (MPa)	UCS SD mean (MPa)	Mean $m_i$	$m_i$ SD
1	Sandstone	6	59	6.9	43	11	26.7	2.5	13	0.6
2	Sandstone	5	53	6.6	40	9				
1	Siltstone	6	49	8.2	30	7	13.7	1.0	11	0.5
2	Siltstone	5	49	9.9	36	5				
1	Claystone	6	47	8.2	15	3	9.9	0.7	7	0.4
2	Claystone	5	48	8.1	10	12				
1	Coal	6	52	2.9	12	3	14.6	1.2	16	1.0
2	Coal	5	49	4.1	15	8				

A Monte Carlo approach was utilised to assess inputs for each group that would be utilised in the stability analyses and the following was carried out for each simulation:

- GSI,  $m_i$  and UCS were assigned for each lithology using the statistical inputs in Table 2.
- Above step provided four rock mass shear strength envelopes, one for each lithology (Figure 9).
- The proportion of each of the four lithologies in the stratigraphic group for this simulation was assigned using the lithology distributions in Table 2.
- As each lithology proportion was independently sampled, the sum would not equal 100% and hence, sampled proportions were standardised to sum to 100%.
- An overall composite rock mass shear strength was evaluated for the simulation based on the proportion of each lithology present in the Group (Figure 9).
- A process was utilised that applied iterative GSI,  $m_i$  and UCS values to provide a best-fit to the overall generalised Hoek–Brown (GHB) rock mass strength envelope (Figure 9).
- 500 simulations were carried out of the above for each group and the best-fit distributions of GSI,  $m_i$  and UCS were used as input for the limit equilibrium stability analyses.

Table 3 provides the best-fit GSI,  $m_i$  and UCS parameters for the two stratigraphic groups. Review indicated a very strong correlation coefficient of  $m_i$  and UCS relative to GSI. However, this correlation could not be accounted for in the stability analyses using the available software at the time. A Monte Carlo analysis in SLIDE V6 (Rocscience Inc. 2010) using these parameters and for the slope geometry being considered indicated a mean FS of 1.26 and a PF of 0.093% (Figure 10).

Table 3 Best-fit parameters Approach One

Group	Mean GSI	GSI SD	Mean UCS (MPa)	UCS SD (MPa)	$m_i$	$m_i$ SD
1	58.0	6.9	16.4	1.9	10.9	0.7
2	53.5	6.2	16.3	1.8	11.8	0.5

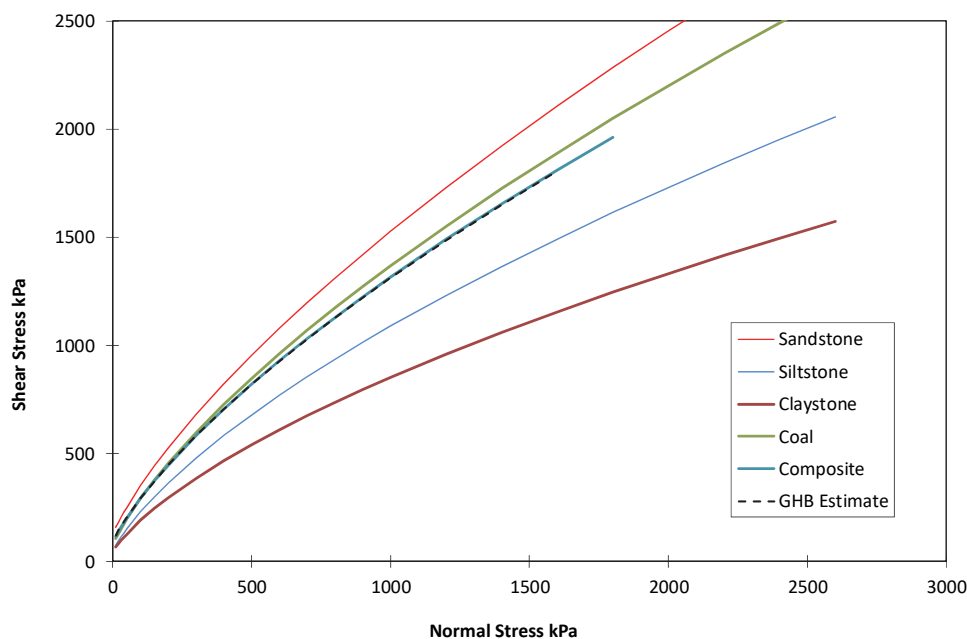


Figure 9 For one simulation, rock mass strength for each lithology, composite for group and best-fit generalised Hoek–Brown (GHB) envelope and with composite and GHB rock mass strength envelopes largely overlying

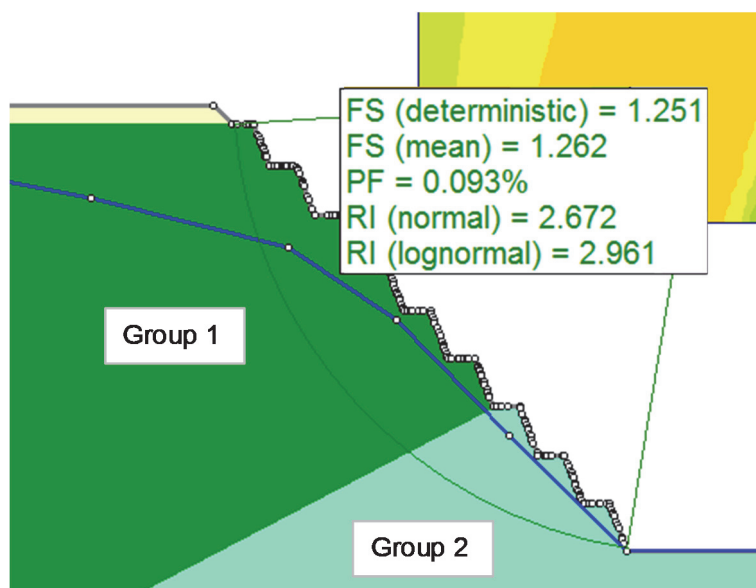


Figure 10 Results of probabilistic analysis using Approach One for Case 2

### 3.2 Approach Two

The second approach has been utilised to emphasise the significant averaging that takes place owing to the intercalated nature encountered in coal mine strata. Utilising one borehole in proximity to a proposed highwall, a layered model was created that reflected the encountered lithologies; albeit somewhat simplified with a uniform dip. The model resulted in creation of over 300 layers.

A Monte Carlo analysis in SLIDE V6 (Rocscience Inc. 2010) was carried out using parameters in Table 2 and indicated a mean FS of 1.22 and a PF of 0.001%.

Even with this very low PF, what is not immediately appreciated is that during each simulation, all layers associated with one lithology are provided with equivalent parameters (i.e. each lithology layer is

dependent, Approach Two-a). As such, to correctly address the analysis it would be appropriate that each layer is independently provided with a randomised strength at each simulation. For the software being utilised, this requires that each individual layer in the model be assigned as a different material type. Whilst feasible to implement, it is considerably time consuming. The problem was simplified so that for each key lithology, four different material layers were assigned in each group and with the material layers interspersed in the created model (Approach Two-b). This analysis indicated a mean FS of 1.19 and a PF of 0.000% (Figure 11). The creation of multiple layers within a model essentially results in a gross-averaging effect as per the law of large numbers and has a significant impact on the PF and in line with the findings of Cylwik et al. (2018).

The mean FSs for Approach Two were slightly lower than using Approach One and is attributed to possibly introducing a higher proportion of claystone layers in the created model than noted in Table 2.

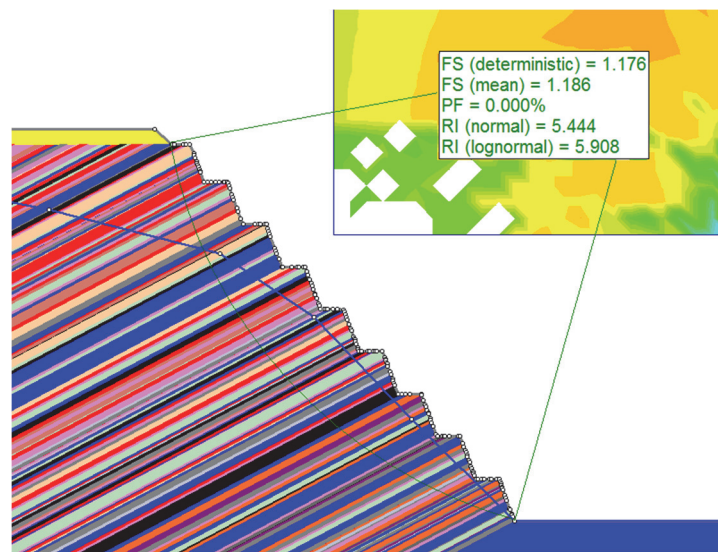


Figure 11 Results of probabilistic analysis using Approach Two-b for Case 2

### 3.3 Approach Three

The third approach has been to blindly utilise the population statistics. In order to remove the complexity with distribution based on key lithologies, the population statistics for UCS and GSI were assessed ignoring lithology. Table 4 provides a summary of the GSI,  $m_i$  and UCS used in this approach. Note the  $m_i$  parameters are in keeping with Table 3.

Table 4 Population statistics Approach Three

Group	Mean GSI	GSI SD	Mean UCS (MPa)	UCS SD (MPa)	$m_i$	$m_i$ SD
1	51.4	15.9	16.4	16.0	11	1.5
2	50.6	14.2				

A Monte Carlo analysis in SLIDE V6 (Rocscience Inc. 2010) was carried out using parameters in Table 4. This indicated a mean FS of 1.21 and a PF of 17.2%.

### 3.4 Overview of different approaches

Of particular note is the typically perceived PFs associated with a given FS. Wesseloo and Read (2009) suggested that a FS in the range of 1.2 to 1.3 is typically equitable to a PF of 15–20%. Table 5 provides a summary of the mean FSs and PFs associated with each approach. Whilst very similar, mean FSs (1.19 to 1.26) are indicated for all approaches there are significant differences in PFs (near zero to 17%) and the SD of FS.

**Table 5** Overview of results from different approaches for Case 2

Approach	Mean FS	SD FS	PF (%)	Comments
One	1.26	0.09	0.093	
Two-a	1.22	0.06	0.001	Independent layers
Two-b	1.19	0.03	0.000	Four material types in each group
Three	1.21	0.23	17.2	

Approach Three, which has utilised the population statistics, provides a PF in keeping with Wesseloo and Read (2009) for a slope where the consequence of failure is not serious. However, it needs to be questioned as to whether the approach realistically addresses the actual variability present in the slope. For Approach Three, there are simulations where a low GSI,  $m_i$  and UCS are applied to the entire slope. Comparison to the actual data, which shows point-to-point variability throughout the rock mass, suggests realisations with low parameters applied to the overall slope as unrealistic. For example, there are simulations where GSI values less than 30 are sampled from the population and applied to the overall slope. This may be unrealistic as the slope would invariably also include inter-burdens of higher GSI and give the rock mass an average GSI of more than 30 along the failure surface.

The large difference in the PFs, as shown by this author for the different approaches in Case 2, is in keeping with Rocscience (2018a) and Renani et al. (2018). Rocscience (2018b) discusses this aspect as part of the introduction of random fields in their Slide2 (Rocscience 2018c) limit equilibrium stability analysis software and which also highlights the limitations of Approach Three. Renani et al. (2018) provides similar findings by comparing results from limit equilibrium analysis (Approach Three) and results of numerical modelling wherein strengths were randomised for each element at each simulation (i.e. in keeping with random fields). The latter providing lower PF and Renani et al. (2018) indicated one example where the difference in the SD of FS was an order of magnitude between analytical approaches. The example by the author in Table 5 suggests similar differences in the SD of FS.

## 4 Discussion

The author highlights that alternatives to using Monte Carlo could be considered but have to be weighed up against the tractability of analysis effort and the rigour of the results. Either the point estimate method, as presented by Rosenbleuth (1975), or the surface response methodology, as discussed in Chiwaye and Stacey (2010), would be an option for Approach One, time consuming for Approach Two, and although both are possible for Approach Three, it would be arguable as to the validity of the results.

Considering the findings of Renani et al. (2018), Rocscience (2018a), and the author's results of Approaches One and Two in Table 5, one has to question the quasi-relationship between FS and PF that has been widely held by geotechnical practitioners. This then also poses a challenge to the emergence of a stronger preference to PF as acceptance criteria, particularly if PF analyses can result in low standard of deviations in FS and therefore, potentially suggest very low risk. It could be argued that the quasi-relationship between FS and PF has been somewhat verified through large-scale experience to date. However, this may be a function of other factors not fully captured in the analyses and which have a more significant bearing on stability such as the influence of complex mechanisms, groundwater and structure rather than rock mass strengths alone. Sensitivity analyses may be just as useful in that decision-makers may not fully appreciate the intricacies and limitation of PF analyses.

McMahon (1985) posed three types of uncertainty, with Type 3 comprising bias and highlighting that a key risk for any design is the bias in parameter selection (Figure 12). To emphasise this aspect, consider a theoretical slope designed with two different views on input parameters. For view 1, the designer considered a given PF as the acceptance criteria and providing a high FS (essentially the far right in Figure 12). However, owing to bias in the input estimates, the true FS of the design is just above one (far

left in Figure 12). View 2, using more conservative parameters, resulted in a different design, provided a lower FS and a higher PF but was potentially less at risk. Reliance on either FS or PF does capture this bias and may ultimately pose the real measure of risk.

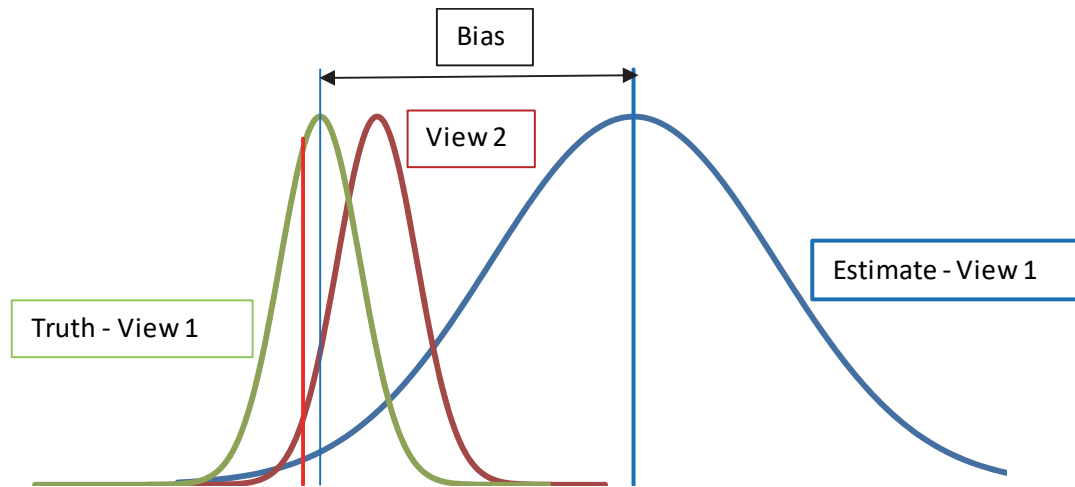


Figure 12 Risk of bias (after McMahon 1985)

## 5 Conclusion

In sedimentary strata, the intercalated nature on inter-burdens can result in very low calculated Probabilities of Failure. The author considers that a focus on careful selection of input parameters to avoid bias, coupled with the critical influence of groundwater and structure in the designs, provides the cornerstone to mitigating risks of instability with slope designs.

For the results of probabilistic analyses to have meaning, they need to appropriately address variability at the scale in question. The approach of assuming the population variance and applying this variability to the entire slope as part of simulations in the analysis is inappropriate. This can result in PF values, which appear to offer risk management but have negligible bearing to the real measure of risk.

## References

- Bewick, RP, Kaiser, PK & Valley, B 2011, 'Interpretation of triaxial testing data for estimation of the Hoek-Brown strength parameter  $m_i$ ', *Proceedings of the 45th U.S. Rock Mechanics/Geomechanics Symposium*, American Rock Mechanics Association, Alexandria.
- Cammack, R & Duran, A 2015, 'A review of methods for assessing the Hoek-Brown  $m_i$  constant from triaxial testing', *Proceedings of the 13th ISRM International Congress of Rock Mechanics*, International Society for Rock Mechanics and Rock Engineering, Lisbon.
- Cargill, JS & Shakoor, A 1990, 'Evaluation of empirical methods for measuring unconfined compressive strength of rock', *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, vol. 27, issue 6, pp. 495–503.
- Chiwaye, HT & Stacey, TR 2010, 'A comparison of limit equilibrium and numerical modelling approaches to risk analysis for open pit mining', *Journal of the Southern African Institute of Mining and Metallurgy*, vol. 110, no. 10, pp. 571–580.
- Cylwik, SD, Beck, JA & Ryan, TM 2018, 'The uncertainty of rock mass shear strength estimates: how to incorporate the reduction in variance due to spatial averaging for use in probabilistic analysis', *Proceedings of Slope Stability 2018*, Asociacion Nacional de Ingenieros de Minas and Colegio Oficial de Ingenieros de Minas del Sur, Seville.
- Deere, DU & Miller, RP 1966, *Engineering Classification and Index Properties for Intact Rock*, University of Illinois, Urbana.
- Fillion, MH & Hadjigeorgiou, J 2013, 'Reliability of strength estimates based on limited laboratory testing', in PM Dight (ed.), *Proceedings of the 2013 International Symposium on Slope Stability in Open Pit Mining and Civil Engineering*, Australian Centre for Geomechanics, Perth, pp. 163–176.
- Gill, DE, Corthésy, R & Leite, MH 2005a, 'Determining the minimal number of specimens for laboratory testing of rock properties', *Journal of Engineering Geology*, vol. 78, issue 1, pp. 29–51.
- Gill, DE, Corthésy, R & Leite, MH 2005b, 'A statistical approach for determining practical rock strength and deformability values from laboratory tests', *Journal of Engineering Geology*, vol. 78, issue 1, pp. 53–67.
- Hess, PE, Bruchman, D, Assakkaf, IA & Ayyub, BM 2002, 'Uncertainties in material and geometric strength and load variables', *Naval Engineers Journal*, vol. 114, no. 2, pp. 139–166.
- Hoek, E 1983, 'Strength of jointed rock masses', *Géotechnique*, vol. 23, no. 3, pp. 187–223.

- Hoek, E & Brown, ET 1997, 'Practical estimates of rock mass strength', *International Journal of Rock Mechanics and Mining Sciences*, vol. 34, no. 8, pp. 1165–1186.
- Hoek, E, Carranza-Torres, CT & Corkum, B 2002, 'Hoek-Brown failure criterion – 2002 edition', in R Hammah, W Barden, J Curran & M Telesnicki (eds), *Proceedings of the Fifth North American Rock Mechanics Symposium*, University of Toronto Press, Toronto, pp. 267–273.
- Kirsten, HAD 1983, 'Significance of probability of failure in slope engineering', *The Civil Engineer in South Africa*, vol. 25, no. 1, January, pp. 17–27.
- McMahon, BK 1971, 'Statistical methods for the design of rock slopes', *Proceedings First Australia New Zealand Conference Geomechanics*, vol. 1, pp. 314–321.
- McMahon, BK 1985, *Geotechnical Design in the Face of Uncertainty: EH Davis Memorial Lecture*, Australian Geomechanics Society, Barton, p. 38.
- Read, J 2009, 'Data uncertainty', in J Read & P Stacey (eds), *Guidelines for Open Pit Slope Design*, CRC Press, Boca Raton, pp. 213–220.
- Renani, HR, Martin, CD, Varona, P & Lorig, L 2018, 'Probabilistic stability analysis of slopes in highly heterogeneous rock masses', *Proceedings of Slope Stability 2018*, Asociacion Nacional de Ingenieros de Minas and Colegio Oficial de Ingenieros de Minas del Sur, Seville.
- Rocscience Inc. 2004, *RocLab*, computer software, Rocscience Inc., Toronto
- Rocscience Inc. 2010, *SLIDE*, version 6, computer software, Rocscience Inc., Toronto
- Rocscience Inc. 2018a, 'Tutorial 33: spatial variability', *Slide 2018 Tutorial Manual*, Rocscience Inc., Toronto.
- Rocscience Inc. 2018b, 'Tutorial 34: spatial variability multi material', *Slide 2018 Tutorial Manual*, Rocscience Inc., Toronto.
- Rocscience Inc. 2018c, *Slide2*, computer software, Rocscience Inc., Toronto
- Rosenblueth, E 1975, 'Point estimates for probability moments', *Proceedings of the National Academy of Sciences*, vol. 72, no. 10, pp. 3812–3814.
- Wesseloo, J & Read, J 2009, 'Acceptance criteria', in J Read & P Stacey (eds), *Guidelines for Open Pit Slope Design*, CRC Press, Boca Raton, pp. 221–236