SSIM 2023 – PM Dight (ed.)

doi:10.36487/ACG_repo/2335_34

Data overload: Does more data really resolve uncertainty?

A Duran *PSM, Australia*

Abstract

The author presents a case study where ongoing drilling campaigns coupled with laboratory testing were utilised to assess the geotechnical characteristics along strike in an area of complex geology. While 'more' geotechnical logging is paramount in providing a spatial understanding and resolving uncertainties in geology and structure, the author's experience with the laboratory testing is somewhat contrary. For the case study, laboratory testing had been carried out on a campaign basis over an 11-year period, resulting in over 200 direct shear (DS) tests, 1,200 unconfined compressive strength (UCS) tests and 250 triaxial (TXL) tests.

This paper highlights issues that are encountered with such extensive laboratory data. UCS testing, using point values, has well-defined methodologies in statistical treatment and assessing uncertainty and reliability. However, tractable methods to deal with DS and TXL are not as well-defined. The paper highlights the issues and presents what appears to be increasing uncertainty with more data, contrary to the commonly held view that more testing is better.

Keywords*: uncertainty, laboratory testing, UCS testing, direct shear testing*

1 Introduction

This paper presents a case study relating to an open cut mining operation and the assessment of laboratory testing data utilised in determining strength parameters for pit slope design studies. The mining operation proceeded with a strike-advance approach and with geotechnical studies carried out as campaigns which progressively provided designs in advance of mining. The geotechnical studies comprised cored boreholes coupled with geotechnical logging, sampling and laboratory testing inclusive of more than 200 direct shear (DS) tests, 1,200 unconfined compressive strength (UCS) tests and 250 triaxial (TXL) tests.

This paper focuses on discussion of the interpretation of the UCS and DS laboratory testing. UCS testing, using point values, has well-defined methodologies in statistical treatment and assessing uncertainty and reliability. However, tractable methods to deal with DS and TXL are not as well-defined. The paper highlights the issues and presents what appears to be increasing uncertainty with more data, contrary to the commonly held view that more testing is better.

This paper presents a typical statistical approach that may be utilised for UCS testing and resolving key questions such as: 'Is there a spatial trend across the deposit? and What is the mean estimate and how reliable is that estimate?'

In contrast, while similar questions may be asked for DS testing, statistical treatment requires some assumptions in making the problem tractable.

This paper highlights the outcomes from the case study where there was extensive laboratory testing present and which ultimately highlighted that further testing, while appearing sensible as part of future geotechnical campaigns, is likely to provide negligible benefit to the project.

The author notes the first campaign has been designated as Y1, and with subsequent programs (not always carried out yearly) designated as Y2, Y3, Y4, Y5, Y8, Y9, Y10 and Y11.

2 Unconfined compressive strength testing

To highlight the approach utilised in the assessment of the UCS testing, this paper presents the results for the dominant sedimentary rock type encountered (sandstone). Table 1 provides a summary of the number of tests carried out with successive campaigns and summary statistics.

Table 1 Summary of unconfined compressive strength testing by campaign

Campaign	Number of tests	Mean (MPa)	Standard deviation (MPa)	Coefficient of variation
Y1	2	16.6	7.1	43%
Y3	15	30.0	19.0	63%
Y4	23	14.6	11.9	81%
Y5	3	10.6	4.0	38%
Y8	25	16.0	8.3	52%
Y9	46	14.6	9.1	62%
Y10	64	16.2	10.7	66%
Y11	16	13.3	7.1	53%

Table 2 provides a Student t-test pairwise comparison between progressive campaigns, noting that each campaign covered a strike length of nominally 300 m. The comparison in Table 2 is aimed at the key question of whether there was a change in intact strength along strike. A typical value accepted in a Student t-test, Cramer (1993), is that a P-value of less than 0.05 is considered significant and inferring a different mean between the two groups.

Of note is that the mean UCS from Y3 has a statistically significant different mean to majority of other campaigns (shaded in green in Table 2). The author largely assigns this difference to potential bias in sample selection during the Y3 campaign. Geotechnical logging and review of geophysics data for Y3 did not indicate differences in logging to suggest a higher strength as warranted in one area of the deposit.

Table 2 Pairwise Student t-test of UCS testing by campaign

Campaign	Student t-test statistic – P-value							
	Y1	Y3	Y4	Y5	Y8	Y9	Y10	Y11
Y1	NA	0.11	0.77	0.43	0.94	0.76	0.95	0.63
Y3	0.11	NA	0.01	0.00	0.01	0.01	0.01	0.00
Y4	0.77	0.01	NA	0.26	0.64	0.98	0.59	0.66
Y5	0.43	0.00	0.26	NA	0.12	0.22	0.11	0.40
Y8	0.94	0.01	0.64	0.12	NA	0.50	0.96	0.27
Y9	0.76	0.01	0.98	0.22	0.50	NA	0.40	0.57
Y10	0.95	0.01	0.59	0.11	0.96	0.40	NA	0.20
Y11	0.89	0.00	0.68	0.40	0.27	0.57	0.20	NA

Figure 1 presents the progressive changes in mean UCS, standard deviation on mean and reliability (Fillion & Hadjigeorgiou 2013) of the UCS estimate with a progressive number of tests. Figure 1 indicates that beyond about 80 tests there was very little change in the mean UCS, with the standard deviation on mean nearing asymptote. By contrast, the reliability approached 80% (recommended minimum at operating phase [Read

2009]) after some 50 tests. This highlights there was almost no additional benefit in the further 100 tests that were undertaken from Y9 onwards. In this case 'more' provided negligible benefit.

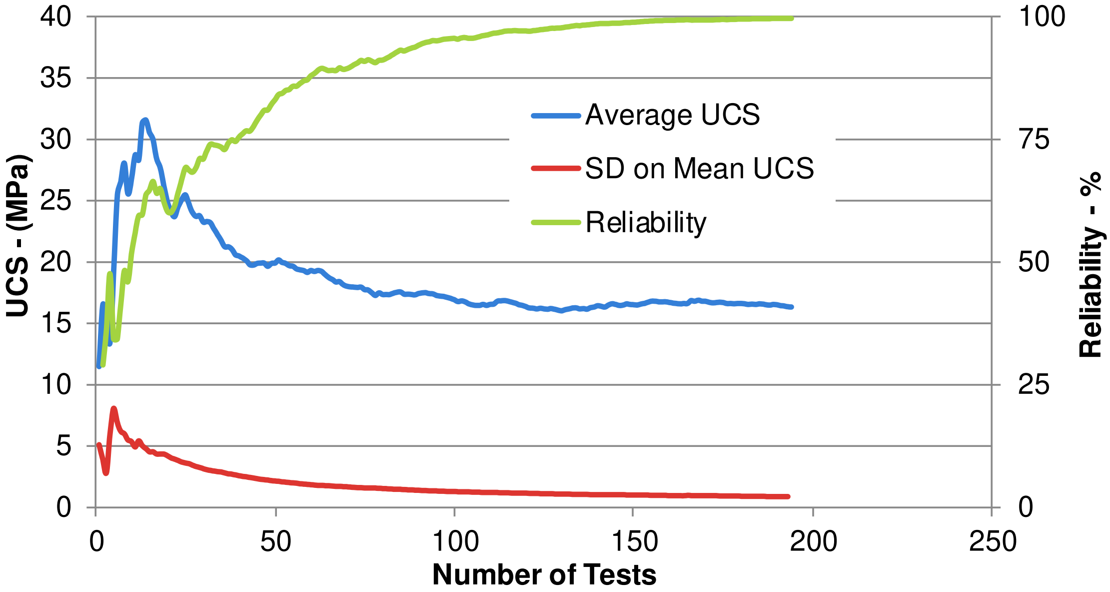


Figure 1 Progressive change in mean UCS, standard deviation on mean and reliability with a progressive number of tests

3 Direct shear testing

3.1 Overview

The author stresses that laboratory DS testing data requires careful consideration of the testing being accepted as valid. Moreover, the underlying premise used in the interpretation can have a profound impact on design and the appropriate design acceptance criteria (DAC) that is warranted. To highlight this aspect, three approaches have been utilised in the interpretation for the project data:

- Visual assessment of upper bound, lower bound and lower quartile strength using the entire dataset.
- Numeric approach.
- Statistical approach.

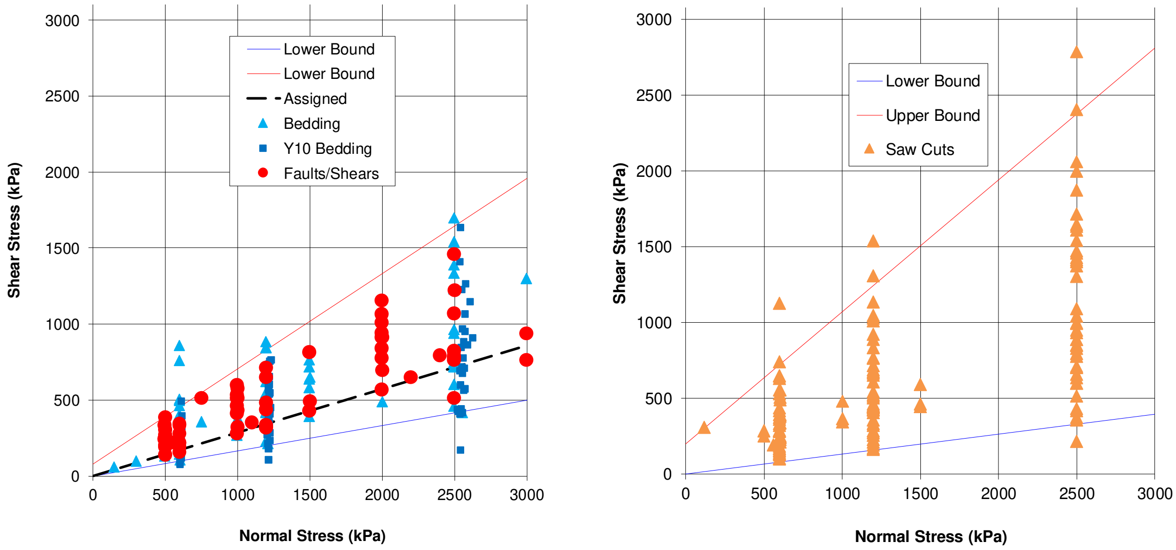
Regardless of approach, it is considered good practice to carefully review outliers and consider defect type/s where grouping of data is carried out. A key aim of the DS testing was to provide an estimate of the shear strength of bedding partings within the sedimentary sequence. As such, test results of bedding partings and faults/shears near parallel to bedding were grouped for consideration.

Table 3 provides an overview of the DS testing that was grouped, with testing comprising multi-stage testing of each sample. Of note is that in the Y8 and Y9 campaigns there were a significant number of DS tests on saw-cut defects. While such testing is advocated by some authors – Hoek (2014) indicating a reliable estimate of basic friction angle can be provided by testing of saw cuts, and Barton (1976) though advocating saw cuts stresses a requirement for sand blasting of the surface. However, in the author's opinion, this is often not the case. Figure 2a indicates the testing of natural defects. In comparison, Figure 2b provides the testing of saw cuts.

Table 3 Summary of direct shear testing by campaign

Campaign	Number of tests – bedding	Number of tests – faults/shears	Number of tests – saw cuts
Y3	3	9	0
Y4	2	7	0
Y8	17	2	13
Y9	2	3	33
Y10	26	0	0

The testing of saw cuts highlights several aspects of concern. Firstly, a much higher scatter in the results provides a very poor indicator of basic friction angle. A similar scatter is evident in testing of saw cuts presented within Alejano et al. (2012), with Hencher (2012) expressing similar concern to the author of such wide scatter. Secondly, nominally 50% of the test results provide high inferred cohesions. Thirdly, some of the test results indicate very low friction angles, as low as 8°. Both the second and third aspects are considered unrepresentative for essentially clean surfaces. The author considers that results from saw-cut surfaces are highly dictated by the process utilised in the laboratory and may provide a very poor representation of field scale behaviour. As such, saw-cut test results need to be utilised with appropriate review. For this paper, the author has disregarded the saw-cut test results in the subsequent analysis of data.



(a) (b)

Figure 2 Total direct shear testing to Y10 with (a) natural defects at left and (b) saw cuts at right

3.2 Visual approach

Figure 2 provides an overview of the visual assessment approach whereby the selected dataset of testing is utilised in assigning an upper bound, a lower bound and a lower quartile strength (noted as assigned). The approach is one that was utilised by mentors in the consultancy firm where the author initially commenced in the 1980s, and was considered an appropriate approach in selecting a shear strength to provide robust design.

While it can be argued the approach can be unduly biased by individual test results (either high or low), the approach requires that outliers are reviewed and, where appropriate, discounted from the selection of upper and lower bounds. The adopted strength is selected based on a combination of visual fit to the data coupled with a nominal count of test results above and below the adopted strength, so as to achieve a nominally lower quartile. For this project, zero cohesion was considered appropriate for design.

Table 4 provides a progression by campaign of the upper bound, lower bound and adopted shear strength utilising the visual approach. Of note is that as more testing became available the bounds gradually changed, with the adopted value remaining constant until Y9. Thereafter a significant amount of testing on bedding defects in Y10 resulted in a change to the lower bound and the adopted strength. A check-analysis of large-scale stability utilising the adopted strength from Y10 in Table 4 yielded unrealistic results and validated that the adopted value to Y9 remained as an appropriate design strength.

Table 4 Visual assessment of shear strength by campaign

Campaign (number of cumulative tests)	Lower bound cohesion – friction angle	Upper bound cohesion – friction angle	Adopted cohesion – friction angle
Y3 (12)	0 kPa – 15°	55 kPa – 28°	0 kPa – 18°
Y4 (21)	0 kPa – 15°	100 kPa – 28°	0 kPa – 18°
Y8 (40)	0 kPa – 14°	80 kPa – 32°	0 kPa – 18°
Y9 (45)	0 kPa – 14°	80 kPa – 32°	0 kPa – 18°
Y10 (71)	0 kPa – 9.5°	80 kPa – 32°	0 kPa – 16°

Unlike the case above for UCS testing wherein more data can be considered superfluous, for this case additional DS testing resulted in greater uncertainty.

3.3 Numeric approach

The numeric approach tabulates the cohesion and friction angle of each sample; discarding values of cohesion or friction angle which are considered unrepresentative and then calculating an overall average cohesion and friction angle. A key issue with this approach is selecting what is considered as a reasonable cutoff for high cohesions and which requires judgement.

Somewhat similarly to that noted for saw-cut defects, key issues with the numeric approach for the natural defects included nominally 40% of the test results provide high inferred cohesions while nominally 10% of results indicate very low friction angles of 9° or less. The latter results are not considered realistic as the author's experience in similar deposits is that 11° to 12° is considered as a realistic lower bound friction angle based on back-analysis of large-scale failures. Table 5 provides an overview of the indicated defect shear strength using the numeric approach.

Table 5 Numeric assessment of shear strength by campaign

Campaign (number of cumulative test)	Mean strength cohesion – friction angle
Y3 (12)	37 kPa – 20°
Y4 (21)	43 kPa – 19°
Y8 (40)	44 kPa – 19°
Y9 (45)5	50 kPa – 19°
Y10 (71)	88 kPa – 19°

Of critical note are the significant cohesions assigned to bedding defects, which increased over time. While such cohesions may provide limited impact on the large scale, they can have profound impact on stability analyses at the bench and multi-bench scale. Also of note are the marginally higher friction angles than those achieved using the visual approach.

From a design perspective it would be prudent to use a higher DAC when a numeric approach is utilised in assigning shear strength, in recognition that design strength could be considered as slightly non-conservative; particularly if significant cohesions are adopted.

For the numeric approach, a similar finding to the visual approach is evident. Additional DS testing resulted in greater uncertainty, owing to the increasing cohesion with subsequent campaigns.

3.4 Statistical approach

While there is extensive literature on using the statistical approach in assessing UCS testing, including Gill et al. (2005a, 2005b) and Fillion & Hadjigeorgiou (2013), this is not the case for DS testing.

Two statistical concepts have been utilised in assessing the DS testing: firstly, the approach as presented within Maldonado & Dight (2020); and secondly, linear regression of the data coupled with consideration of the 95% confidence limits of the regression.

For the former approach, peak shear strength values at normal shear stresses of nominally 2,000 kPa (commonly last stage in Y3, Y4 and Y8) or 2,500 kPa (commonly last stage in Y9 and Y10) were treated as point statistics. As indicated by Maldonado & Dight (2020), 'in order to combine the effects of both friction and apparent cohesion in just one variable'. Figure 3 presents the distribution of shear strengths using this approach for Y9 and Y10, with Table 6 providing a summary of assessed friction angles with variation by campaign under the assumption of zero cohesion.

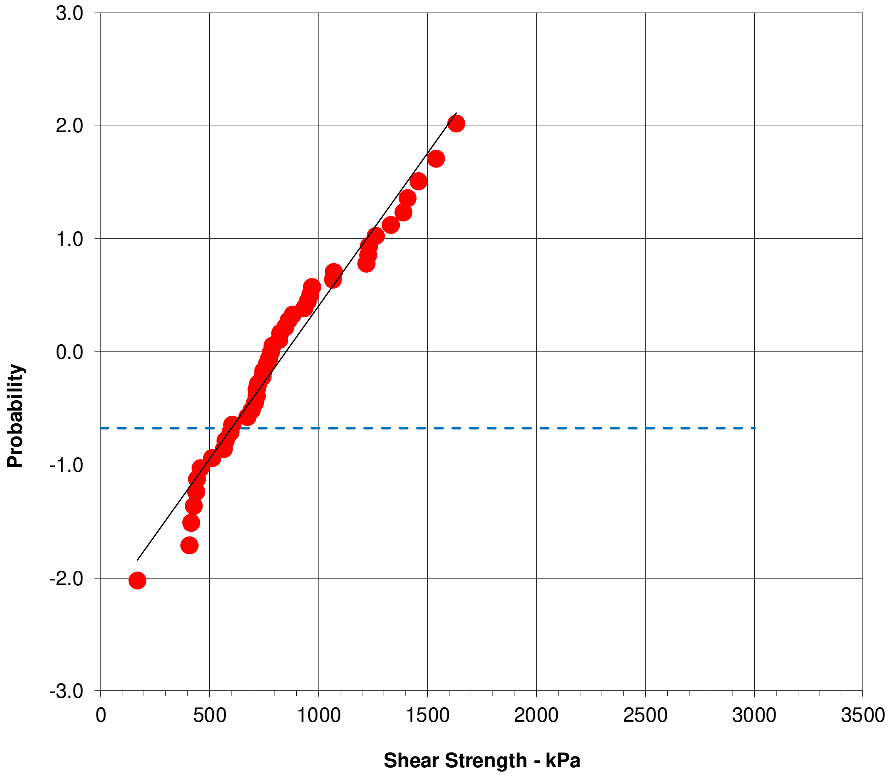


Figure 3 Distribution of peak shear strength values at normal testing of 2,500 kPa for Y9 and Y10. Note the Y-axis is presented in terms of probability. The dashed line is the 25th percentile

Table 6 Statistical assessment of shear strength by campaign

Campaign (number stage results)	Mean effective friction angle	25th percentile effective friction angle	Visual approach adopted friction angle
Y3 (7)	24	18	18
Y4 (11)	24	18.5	18
Y8 (17)	18	16	18
Y9 (22)	18	16.5	18
Y10 (46)	17.5	13.5	16

Of note is the comparison between the 25th percentile by the statistical approach and the adopted strength by the visual approach (shaded in blue in Table 6). The statistical approach is driven by the data and with a gradual reduction in shear strength as more data becomes available. In comparison to the visual approach, there is very good agreement initially but with a divergence from Y8. The divergence is primarily a result of the visual approach rejecting the absolute lower bound stage test results (i.e. discarding results where the friction angle is below 11°) while the statistical approach honours results as low as 9°. The implications of this latter aspect highlight that with further additional testing, there is possibility of the 25th percentile result being further reduced by outliers.

The second approach noted above is that of accepting a linear regression of the data. Furthermore, the author has also utilised the 95% confidence limits of the linear regression to bound potential interpretations. While strict statistics argue that any line between the bounds provides an acceptable result, the author has utilised a line of best fit through each confidence bound to provide upper and lower bounds (Figure 4).

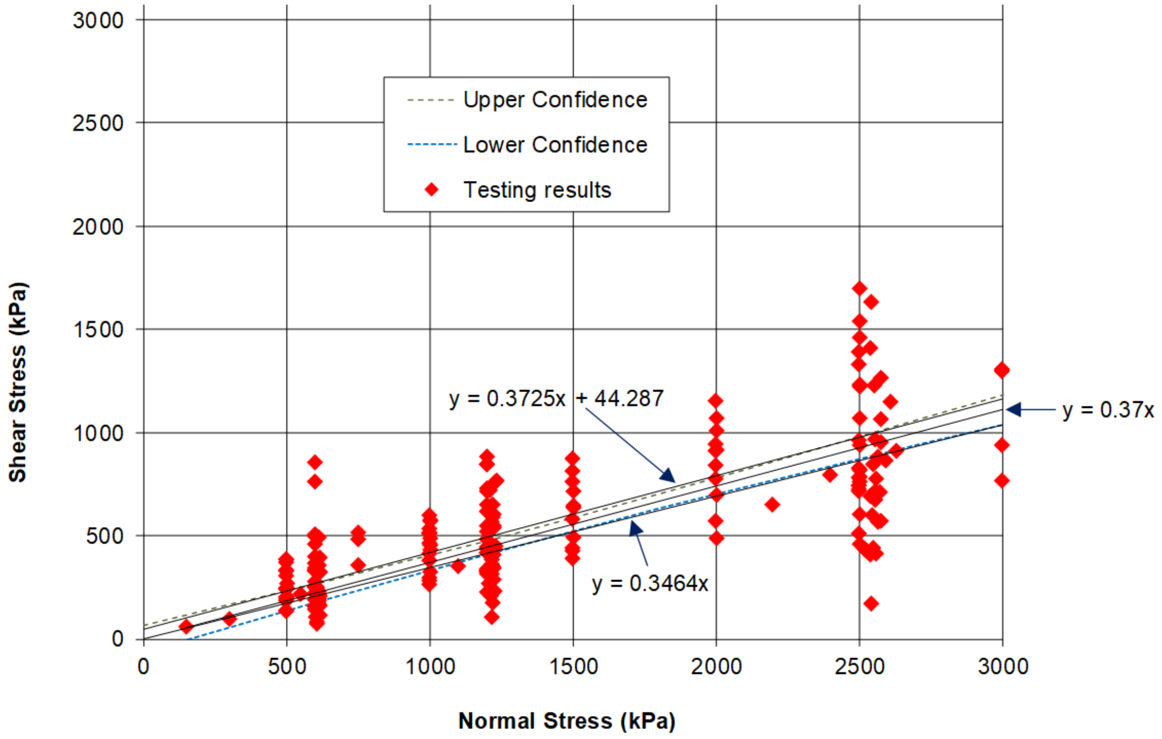


Figure 4 Linear regression of testing results to Y10 with upper and lower confidence bounds (dashed lines) and with linear regression of the confidence bounds

Table 7 provides a summary of the strengths resulting from a linear regression assessment. In keeping with the numeric assessment, the results are driven by data and indicate high cohesion values in the range of 65 to 120 kPa, which are considered unreasonable for bedding defects in a sedimentary setting. While it is appreciated that the results may reflect the available data, they do not appear to be sensible values for use as a design basis unless one completely ignores the cohesion values and adopts friction angle only.

Of note is the comparison of the Y10 numeric assessment against the corresponding linear regression, with very similar values indicated; albeit that the results of the numeric assessment are dictated by the selection of which cohesions are considered as high and thus were discarded in the assessment process.

Table 7 Linear regression assessment of shear strength by campaign

Campaign (number of cumulative test)	Lower confidence cohesion – friction angle	Upper confidence cohesion – friction angle	Linear regression of test results cohesion – friction angle
Y3	0 kPa – 20°	134 kPa – 20°	66 kPa – 20°
Y4	0 kPa – 18.5°	134 kPa – 19°	81 kPa – 19°
Y8	66 kPa – 18°	179 kPa – 19°	122 kPa – 18.5°
Y9	58 kPa – 18.5°	166 kPa – 19°	113 kPa – 19°

3.5 Triaxial data

While triaxial data is not discussed herein, a similar approach to what is discussed for DS testing could be utilised.

4 Conclusion

The following conclusions can be drawn from the assessment of laboratory testing for the project presented:

- Unless there is strong evidence from geotechnical logging to suggest a change in intact strength, additional UCS testing for a unit/rock type appears to offer negligible benefit once about 50 samples have been tested. More testing has negligible benefit in reducing uncertainty in UCS results.
- The author’s experience with testing of saw cuts is in accord with Hencher (2012), and contrary to the recommendations of Hoek (2004) and Barton (1976), wherein saw cuts can provide a large scatter in results and provide a poor basis in estimating basic friction angle.
- For the project in question there appeared to be very limited benefit in additional DS testing for a unit/rock type once 50 multi-stage tests were completed. Moreover, for the project, additional DS testing provided wider scatter in results and an increase in uncertainty.
- For assessment of DS testing, the author advises caution on the use of a numeric or linear regression assessment approach, particularly if high cohesions are indicated and contrary to expectation based on the defect types being considered. If such an approach is adopted, a higher DAC should be utilised in recognition that design strength could be considered as slightly non-conservative.
- Statistical assessment as indicated by Maldonado & Dight (2020) provides a useful approach but requires cognisance of low rogue test results.
- While the author advocates the visual approach, it is also prone to low rogue results and can require the careful deliberation of an experience-based lower bound.

References

Alejano, LR, Gonzalez, J & Muralha, J 2012 ‘Comparison of different techniques of tilt testing and basic friction angle variability assessment’, *Rock Mech Rock Eng,* vol. 45, pp. 1141–1143.

Barton, N 1976, ‘The shear strength of rock and rock joints’, *International Journal of Rock Mechanics & Mining Sciences*, vol. 13, pp. 255–279.

Cramer, H 1999, Mathematical Methods of Statistics, Princeton University Press, Princeton.

Fillion, M-H & Hadjigeorgiou, J 2013, 'Reliability of strength estimates based on limited laboratory data', in PM Dight (ed.), *Slope Stability 2013: Proceedings of the 2013 International Symposium on Slope Stability in Open Pit Mining and Civil Engineering*, Australian Centre for Geomechanics, Perth, pp. 163–176, https://doi.org/10.36487/ACG_rep/1308_05_Hadjigeorgiou

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.

Hencher, SR 2012, 'Discussion of Alejano, Gonzalez and Muralha', *Rock Mechanics and Rock Engineering*, vol. 45, pp. 1137–1139.

Hoek, E 2004, *Rock Engineering*, Course notes, Canada, 313 p.

Maldonado, A & Dight, PM 2020, 'The shear strength of bedding partings in shales of the Pilbara: the similarity of non-dilatational angles, mineralogy relationships, and nominal roughness', in PM Dight (ed.), *Slope Stability 2020: Proceedings of the 2020 International Symposium on Slope Stability in Open Pit Mining and Civil Engineering*, Australian Centre for Geomechanics, Perth, pp. 551–564, https://doi.org/10.36487/ACG_repo/2025_33

Read, J 2009, 'Data uncertainty', in J Read & P Stacey (eds), *Guidelines for Open Pit Slope Design*, CRC Press, Boca Raton.