

Epistemic uncertainty propagation in slope stability analysis and implications in safety margins

CE Valderrama *SRK Consulting (Chile) S.A., Chile*

M Cofré *SRK Consulting (Chile) S.A., Chile*

E Hormazábal *SRK Consulting (Chile) S.A., Chile*

R Álvarez *SRK Consulting (Chile) S.A., Chile*

Abstract

Stability analysis of slopes commonly deals with limited data relating to the characterisation of the material strength properties. This generates an uncertainty level that depends on the number of measurements used to determine them, and that is not related to the intrinsic variability of the material known as epistemic uncertainty. Considering a Mohr–Coulomb material characterised by its cohesion and friction angle, we describe a method to use interval estimation methods (confidence intervals) to estimate the level of epistemic uncertainty as a function of the number of available measurements (sample size), obtaining ranges of plausible values for the mean of the strength parameters. The calculated uncertainty in strength parameters is later propagated through a deterministic slope stability analysis, obtaining a relation between test sample size and the uncertainty in Factor of Safety (FoS) calculation, which can be compared with the typical safety margins used in slope designs.

Keywords: *uncertainty quantification, uncertainty propagation, design safety margins*

1 Introduction

In the context of the mining business, the main aim of a geotechnical slope design is to provide an acceptable degree of stability that minimise the risks of the personnel, equipment and economic impact due to a potential failure, being a fundamental step of this process the strength characterisation of the involved geomaterials.

These materials are affected by various factors during their formation process, such as properties of parent materials, weathering and erosion processes that vary spatially. Additionally, uncertainties are incorporated as a result of imperfect test equipment and procedural errors, transformation uncertainty related to the interpretation of results, and the statistical uncertainty arising from insufficient number of tests (Cao et al. 2017). Therefore, one of the major complications of the design process is dealing with the variability and uncertainty associated with natural variation of properties, and inaccuracy caused by lack of information on parameters or models (Baecher & Christian 2003).

There is extensive literature from various fields that recognises that unpredictability in a parameter or system results from the combined contribution of aleatory variability and epistemic uncertainty (Helton & Oberkampf 2004; Ang & Tang 2007), which are fundamentally different in nature (Bedi 2013).

The aleatory variability arises from the inherent randomness in the properties encountered in nature and therefore, it cannot be reduced by increasing the quantity of data. On the other hand, epistemic uncertainty derives from a lack of knowledge about the approximate value to use for a parameter that is assumed to have a fixed value in the context of a particular analysis (Helton 2011) and hence, it is dependent on the amount of data available.

As indicated by Bedi (2013), there is a necessity for divide variability and uncertainty. First, because of their very different characteristics, one is irreducible and objective and the other reducible and objective or

subjective. Second, because when both are inherent to the stability model, it is useful to know which part of unpredictability in the final result is due to each one, being possible to reduce or remove the part corresponding to the epistemic uncertainty by collecting more information about inputs.

For strength characterisation, the elimination of epistemic uncertainty requires a very large number of tests, e.g. direct shear tests with a good spatial sampling, which requires a great amount of time and money, which commonly are not available due to time and budget constraints, especially in the initial stages of projects.

For this reason, in this paper, we incorporate epistemic uncertainty in the stability analysis (which is done in a non-stochastic manner), analysing its uncertainty depending on the amount of available data to characterise strength with the goal of elucidating how to reduce the uncertainty with a few resources as possible, and making decisions such as how many tests to do characterise material's strength? In which materials is it better to focus on and carry out more tests to cut down the uncertainty? And to understand the impact of the uncertainty of the stability at each stage of the project.

This paper is organised as follows: Section 2 describes the methodology used to characterise and calculate the epistemic uncertainty related to lack of data, and how is after propagated in a deterministic slope stability analysis. Section 3 shows results about the size of the intervals for strength parameters and for a slope formed by a homogeneous Mohr–Coulomb material, under the assumptions that strength data distributes normal and non-normal. In Section 4, we discuss the relation between epistemic uncertainty and safety margin. Section 5 gives the conclusions and outlines future work.

2 Methodology

Precise probability distributions cannot adequately represent epistemic uncertainty, given that it is characterised by imprecision (Bedi 2013) and therefore, can be best represented by intervals and their generalisations (Baudrit & Dubois 2006; Dubois & Prade 2009; Dubois & Guyonnet 2011). Several frameworks have been developed to handle intervals which can be used to represent epistemic uncertainty as interval analysis (Kearfott & Kreinovich 1996; Jaulin et al. 2001), possibility theory (Zadeh 1978), evidence theory (Dempster 1968; Guan & Bell 1991) and imprecise probabilities (Walley 1991). In this paper, the intervals are calculated by a frequentist inference approach, the confidence intervals, and the uncertainty propagation is tackled by means of interval analysis and error analysis (Drosg 2007).

2.1 Epistemic uncertainty as confidence intervals

To obtain a set of complete information that allows for the total elimination of epistemic uncertainty is impossible or impractical (due to time and financial constraints) in the case of strength characterisation. In statistics terms, we always work with a given sample data with which we try to approximate the behaviour of the entire population. As the sample dataset increases in size, better is the approximation of the full population distribution.

During a deterministic stability analysis is common and logic to characterise the expected value of a strength parameter as the mean of its sample data. This is a point estimate or best guess of the real mean of the population. With the same sample data, we can also compute an interval that might contain the true value of the population mean, considering a confidence level that quantifies the level of confidence that the deterministic parameter is inside the interval, which is known as a confidence interval (Neyman 1937).

Therefore, the confidence interval can be used as an interval representation of the uncertainty of the unknown mean population, for a given confidence level. Mathematically, an interval is formulated on the assumption that a set U of possible values for a variable x is known but with no specified uncertainty structure within the set.

The width of the interval depends directly on the size of the sample (number of strength tests), being useful for the goals of this work. The interval's width is also conditional on the assumed statistical model or, in other words, on the probability distribution presumed for the population, which is commonly normal. For normally-distributed data, we can estimate the confidence interval, so uncertainty U as:

$$U = \left\{ x \mid \bar{x} - t_{\alpha} \times \frac{s}{\sqrt{N}} \leq x \leq \bar{x} + t_{\alpha} \times \frac{s}{\sqrt{N}} \right\} \quad (1)$$

where:

- \bar{x} = the sample mean of the strength parameter.
- s = the sample standard deviation of the strength parameter.
- N = the size of the sample.
- t_{α} = the inverse of standardised Student's-t distribution of $N - 1$ degrees of freedom and a confidence level of $CL = (1 - \alpha)$.

Equation 1 is the confidence interval estimation for a Student's-t distribution. A Student's-t distribution is more appropriate in this case because independently that we assume an underlying normal distribution for the population, we do not know the population standard deviation but the sample one. It is possible to see that interval's width (level of uncertainty U) is proportional to confidence level CL and to the sample variance s , but it is inversely proportional to the sample size N .

However, the assumption of normality is a strong postulate in some cases, even when probability plots and goodness of fit test has been done for model validation (a description of the problems associated with these procedures can be found in D'Agostino & Stephens 1986, Scholz 2005, and Loy et al. 2015). The worst consequence of an erroneous normal assumption is an underestimation of the confidence interval because normal distribution is a highly parametric model (i.e. more restrictive) which decreases the uncertainty (bias-variance trade-off).

For moderate departures for normality the Student's-t approach of Equation 1 is still effective; however, when there is evidence of severe skewness, existence of far outliers or others signals of non-normality, the aforementioned approach is not useful. It is recommended to use the bootstrap method for these cases, specifically the bias-corrected and accelerated bootstrapping (Efron 1987), but the choice of a particular bootstrap procedure depends on the features of each case. Bootstrap methods consist of estimating a characteristic of the unknown population (the mean in this work) by simulating the characteristic when the true population is replaced by an estimated one (Diciccio & Romano 1988). For the more common type of bootstrap (the non-parametric) several samples of the same size of the original are constructed by resampling with replacement. For each of these samples, the mean is calculated, a distribution of the means is computed, and later, confidence interval of this distribution can be obtained.

It is important to mention that for small samples (<10 tests) covering a restricted range, it is recommended to use Equation 1 because during bootstrapping, spurious structures in the sample may be faithfully reproduced in the simulated data.

2.2 Uncertainty propagation in slope stability analysis

By using the sample mean of the strength parameters in stability analysis, we obtain a best guess of the Factor of Safety (FoS), FS_{sample} . We do not know how well represents the real FS_{real} and therefore, there is uncertainty about the FS_{real} that depends on the level of uncertainty of the input parameters.

This dependency can be estimated in several manners. If we make a relationship between the characteristics of uncertainty and errors (such as Drogg 2007) who recommends refraining from using the term error, favoring the term uncertainty), it is possible to make the same treatment for both. With this consideration, the magnitude of uncertainty in the FoS, Δ_{FS} , generated by the uncertainty in the inputs can be calculated by the general law of error propagation, that for two inputs is (Drogg 2007):

$$(\Delta_{FS})^2 = (\partial FS/\partial x)^2 \times (\Delta x)^2 + (\partial FS/\partial y)^2 \times (\Delta y)^2 + 2(\partial FS/\partial x)(\partial FS/\partial y)\Delta x \times \Delta y \times k \quad (2)$$

where:

$FS(x, y)$ = safety factor as a function of input parameters x and y .

Δ = delta operator, representing the magnitude of uncertainty (interval width) U_{FS} .

k = factor representing the degree of correlation between Δx and Δy . $k = 1$ for correlated and $k = 0$ for non-correlated ($\Delta x, \Delta y$).

The need for an explicit expression of the function $FS(c, \phi)$ is clear from Equation 2, which in general only exists for idealised cases. Therefore, we use this approach to analyse the case of a dry slope conformed by a homogeneous Mohr–Coulomb material with strength parameters c and ϕ .

For the more common situation in which the $FS(c, \phi)$ function does not exist, we propose the use of interval analysis (Moore et al. 2009) for the estimation of the safety factor uncertainty, U_{FS} . Using this method, strength parameter intervals calculated as mentioned in Section 2.1, can be propagated through a model (FS calculation) given as a result another interval that bounds all possible values that the safety factor can take.

The number of intervals involved in calculations is equal to the number of strength parameters multiplied by the number of materials. The vertex method (Dong & Shah 1987) can be used to simplify computations that comprise several intervals, by performing a series of FS computations on the end points of each interval (vertices). In theory, for M intervals, 2^M computations are needed, and the output interval is given by:

$$U_{FS} = \left[\min_j (FS(c_j)), \max (FS(c_j)) \right], j = 1 \text{ to } 2^M \quad (3)$$

where:

FS = safety factor operator (given by a function or a black-box software).

c_j = j -th vertex, being each vertex a combination of intervals end points.

However, the FS operator is a monotonic function of its strength parameters. This characteristic makes possible to reduce the number of operations only to two, being the output interval equal to:

$$U_{FS} = [FS(\min(U_k)), FS(\max(U_k))], k = 1 \text{ to } M \quad (4)$$

where:

$FS(\min(U_k))$ = safety factor for lower bounds of all the strength parameters intervals.

$FS(\max(U_k))$ = safety factor for upper bounds of all the strength parameters intervals.

3 Results

3.1 Impact of sample size in uncertainty

It is important to show how the number of available tests that impact the uncertainty of the mean value. Figure 1 shows the magnitude of the uncertainty of the mean calculated with Equation 1 for different numbers of sample size (N) (from 5 to 40) and confidence levels (CL) (70, 80 and 90 %). The uncertainty is shown as sample standard deviation times, i.e. uncertainty divided by s . To see the importance of sample size in the mean uncertainty, consider two materials: the first with a sample size $N_1 = 5$ and the second with $N_2 = 10$, and sample standard deviations such that $s_2 = 1.5 s_1$. Without taking into account the sample size,

we can be tempted to assign more uncertainty to the material 2, but seeing Figure 2, we realise that the uncertainty U_1 (approximately equal to $2.0 s_1$, for 90% CL) is similar (or greater) than $U_2 = 1.1 s_2$.

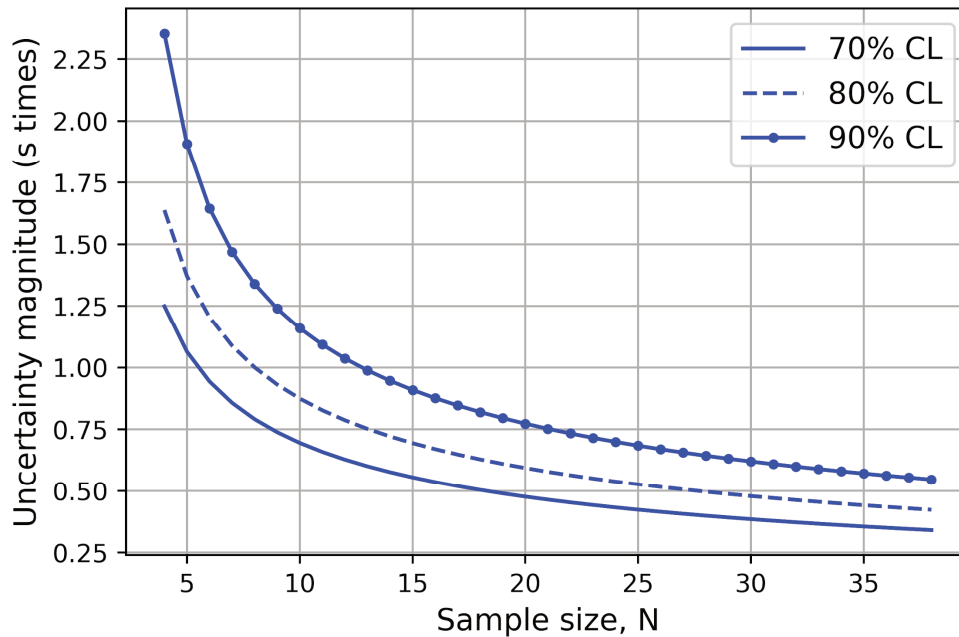


Figure 1 Influence of sample size on the mean uncertainty for different values of confidence level. Uncertainty in the y coordinate is divided by the sample standard deviation

With respect to the calculation of confidence intervals, in Section 2.1, two methods to calculate them were described, Student’s-t and bootstrapping. We generate random samples of 20 data from a normal distribution with mean 100 and standard deviation 10. We calculate 95% confidence interval using Student’s-t and bootstrap method. Table 1 shows that both methods give similar results when the distribution does not show significant skewness.

Table 1 Intervals obtained from Student’s-t and bootstrap compared with the true intervals for 95% confidence levels

Case	Lower level	Upper level
Normal distribution bootstrap	99.0	106.2
Normal distribution Student’s-t	95.3	104.7

Calculated confidence intervals can help to determine the minimum sample size for a required confidence level, as has been shown by Gill et al. (2005), Ruffolo & Shakoor (2009) and Fillion & Hadjigeorgiou (2013), who define this minimum sample size, N_{min} as:

$$N_{min} = \left(\frac{(p+1)}{(p-1)} \times t_{\alpha} \times \frac{s}{\bar{x}} \right)^2 \tag{5}$$

Where p is the precision index, defined as the ratio between upper and lower bounds of a confidence interval:

$$p = \frac{\bar{x} + t_{\alpha} \times \frac{s}{\sqrt{N}}}{\bar{x} - t_{\alpha} \times \frac{s}{\sqrt{N}}} \tag{6}$$

For permanent mine openings, Gill et al. (2005) propose to use a value of $p \leq 1.35$ with a confidence level of 95%, but Fillion & Hadjigeorgiou (2013) considered acceptable a confidence level of 80%. Therefore, the criteria used to define the minimum sample size is to ensure a maximum relative error ($|(1 - p)/(1 + p)|$) on the true mean. In this paper, we try to obtain how this accepted maximum relative error impact on safety factors, which can be done with the relation between p and uncertainty U :

$$U = 2\bar{x} \times \frac{p-1}{p+1} \quad (7)$$

Additionally, it is important to note that as N_{min} and p are functions of sample standard deviation, they also depend on the aleatory variability (sample standard deviation, s , is a point estimate of the population standard deviation, σ , which represents the aleatory variability). For this reason, for a targeted value of the precision index p , it is not possible to define a priori the minimum sample size, and for a given rock type the minimum sample size varies from one case to the other (Gill et al. 2005; Ruffolo & Shakoor 2009), being higher for samples with larger variability and vice versa.

3.2 Uncertainty in safety factors

Considering the slope stability problem, a slope of arbitrary height H and angle α , excavated in homogeneous, isotropic and dry ground, that obeys the Mohr–Coulomb shear failure criterion and characterised by unit weight γ , and uncertain values of cohesion, c , and friction angle, ϕ , it is considered. For this case, approximate explicit solutions are available for the FoS, as the one given by Carranza-Torres & Hormazábal (2018):

$$FS/\tan(\phi) = 1/\tan(\alpha) + g_1(\alpha)/X + g_2(\alpha)/X^{g^3(\alpha)} \quad (8)$$

where:

α = slope inclination.

X = dimensionless parameter equal to $\gamma H \tan(\phi)/c$.

$g_i(\alpha)$ = cubic polynomials function of slope inclination.

The expressions for the cubic polynomial $g_i(\alpha)$ have been omitted for simplicity.

By using Equation 2 and assuming correlation between c and ϕ (justified by extensive literature, e.g. Di Matteo et al. 2013), after some algebraic manipulation we can obtain the following expression for the magnitude of FS uncertainty U_{FS} :

$$U_{FS} = \left| \left(\frac{1}{\tan(\alpha)} + \frac{(1-g^3)g^2}{X} \right) U_{\tan(\phi)} + \frac{1}{\gamma H} \left(g_1 + \frac{g_2 g^3}{X^{g^3-1}} \right) U_c \right| \quad (9)$$

where:

$U_{\tan(\phi)}$ = uncertainty magnitude of the friction angle's tangent.

U_c = uncertainty magnitude of the cohesion.

In Equation 9, the value of X is estimated with the interval average values of c and ϕ . The function of Equation 9 is linear with respect to the uncertainties of the strength parameters and a decreasing function of the parameters α , H and γ . This is an interesting characteristic, which means that for more critical slopes (greater α , H and γ) the uncertainty in the safety factor is lower. Of course, the last is valid only when c and ϕ are not functions of H , as happens in linearisation of Hoek–Brown parameters.

Figure 2 shows how the uncertainty in the strength parameters is propagated to uncertainty in safety factor, for several slope angles (20, 30, 40 and 50°) and values of X . For example, for a slope such that $X = 30$, with an inclination of 40 degrees, we obtain approximately a factor 5 in the y axis of the bottom figure, then, for an uncertainty of 30 kPa in the mean cohesion, we will have an uncertainty $U_{FS} = 0.03 \cdot 5.6 = 0.17$.

Figure 2 also gives us more information about the uncertainty propagation in FoS calculation. As was mentioned, the uncertainty propagated is greater for gently inclined slopes. Considering that high values of X indicates a ground that is predominantly frictional, while low values are predominantly cohesive, it is interesting to note that the uncertainty in friction angle has more impact in the safety factor of cohesive grounds. On the contrary, uncertainty in cohesion has more impact on frictional grounds.

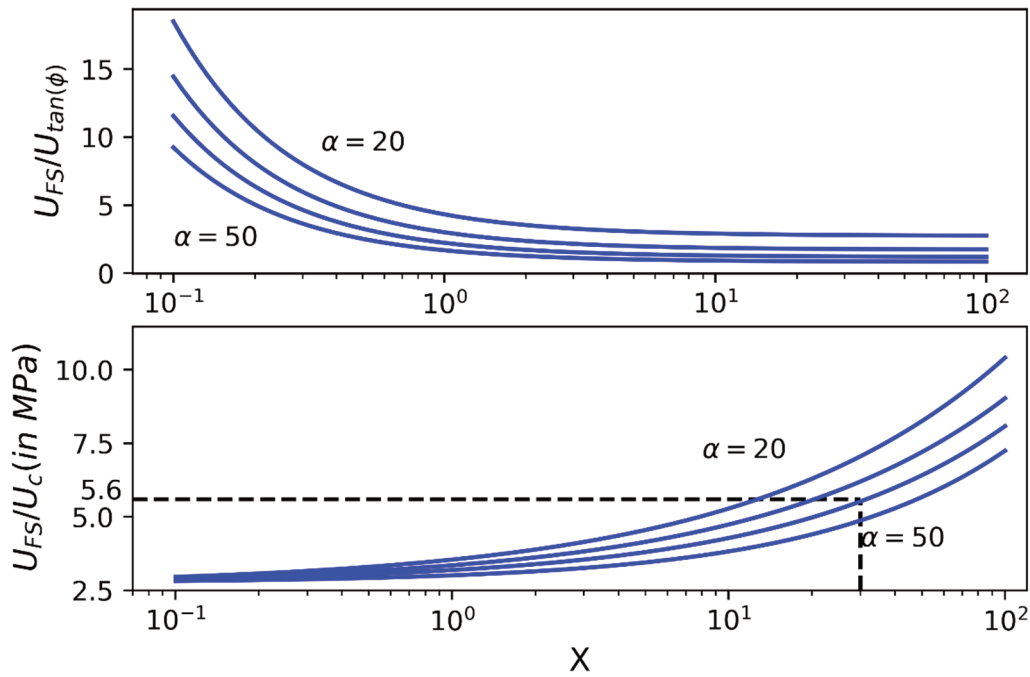


Figure 2 For different slope angles and values of the dimensionless factor X, it is shown how much of the uncertainty of friction angle and cohesion is propagated to uncertainty of the safety factor. The uncertainty in FoS is equal to the uncertainty of the parameter multiplied by the values of the curves

To relate these results with the work of Gill et al. (2005), Ruffolo & Shakoor (2009) and Fillion & Hadjigeorgiou (2013), we assume that c and ϕ have the same p , and replacing Equation 7 into Equation 9 to obtain the uncertainty of the safety factor in terms of the precision index and the sample mean of the strength parameters:

$$U_{FS} = 2 \frac{(p-1)}{(p+1)} \times \left| \left(\frac{1}{\tan(\alpha)} + \frac{(1-g^3)g^2}{X} \right) \overline{\tan(\phi)} + \frac{1}{\gamma H} \left(g1 + \frac{g2 g3}{X g^{3-1}} \right) \bar{c} \right| \quad (10)$$

In Equation 10, uncertainty is limited to a percentage of the mean value of the strength parameters, different to Equation 9, where the magnitude of the uncertainty is an explicit input. For a precision index of 1.35 (as recommended by Gill et al. 2005), Figure 3 shows the uncertainty in safety factor, U_{FS} , for several slope angles (20, 30, 40, and 50°) and values of X showing U_{FS} as the number of times the mean value of the strength parameters.

Figure 3 shows similar behaviour to that observed in Figure 2. Considering the criteria of Gill et al. (2005), a slope with $X = 30$, with an inclination of 40 degrees, and mean cohesion of 300 kPa can have an uncertainty of $U_{FS} = 0.3 \cdot 1.7 = 0.51$ in the FoS, which means that the criteria of Gill et al. (2005) can be not enough in some situations. Therefore, to determine the minimum sample size, it is recommended to use the method proposed by Gill et al. (2005) followed by a simple check such as the one presented in this paper.

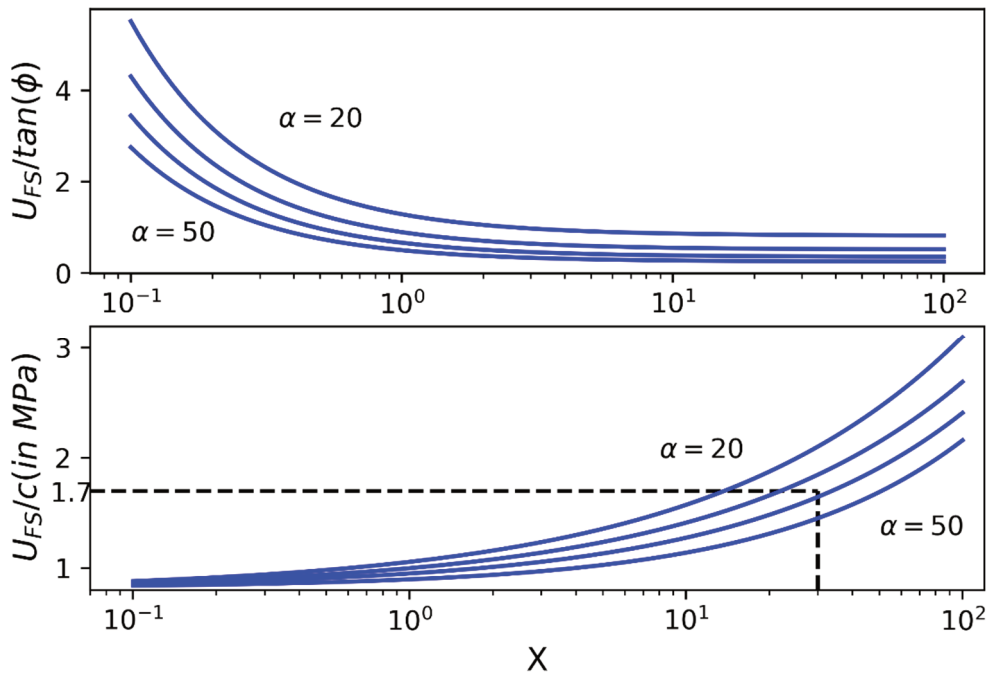


Figure 3 For different slope angles and values of the dimensionless factor X, it is shown the uncertainty of the safety factor with respect to the mean values of friction angle and cohesion, considering an uncertainty in strength parameters given by a precision index of 1.35. The uncertainty in FoS is equal to the mean value of the parameter multiplied by the values of the curves

To illustrate the application of the described concepts and the use of interval analysis, the following case of slope geometry and ground properties is considered:

- Height of slope, $H = 300$ m.
- Angle of slope, $\alpha = 48^\circ$.
- Unit weight of the ground, $\gamma = 24\text{kN/m}^3$.
- Friction angle of the ground, normal distributed with mean $\phi = 35^\circ$ and coefficient of variation (COV) = 20%.
- Cohesion of the ground, normal distributed with mean $c = 300$ kPa and COV = 20%.

We assume that between distributions of c and ϕ , a correlation exists and can be characterised by a Pearson’s correlation coefficient of $\rho = -0.6$. From these distributions, we generate a random sample with size $N = 30$. Later, we calculate uncertainty in mean parameters taking sub-samples of sizes $N = 5, 8, 10, 13, 15, 18, 20, 23, 25, 28$ and 30. For each of these samples, we also calculate the FoS intervals by using the interval method (Equation 4):

$$FS_{min} = FS(c_{min}, \phi_{min}) \quad FS_{max} = FS(c_{max}, \phi_{max}) \tag{11}$$

Figure 4 shows a scatter plot of the sample of 30 data obtained for the correlated normal distributions, and the histograms of the cohesion at the top, and the friction angle at the right side.

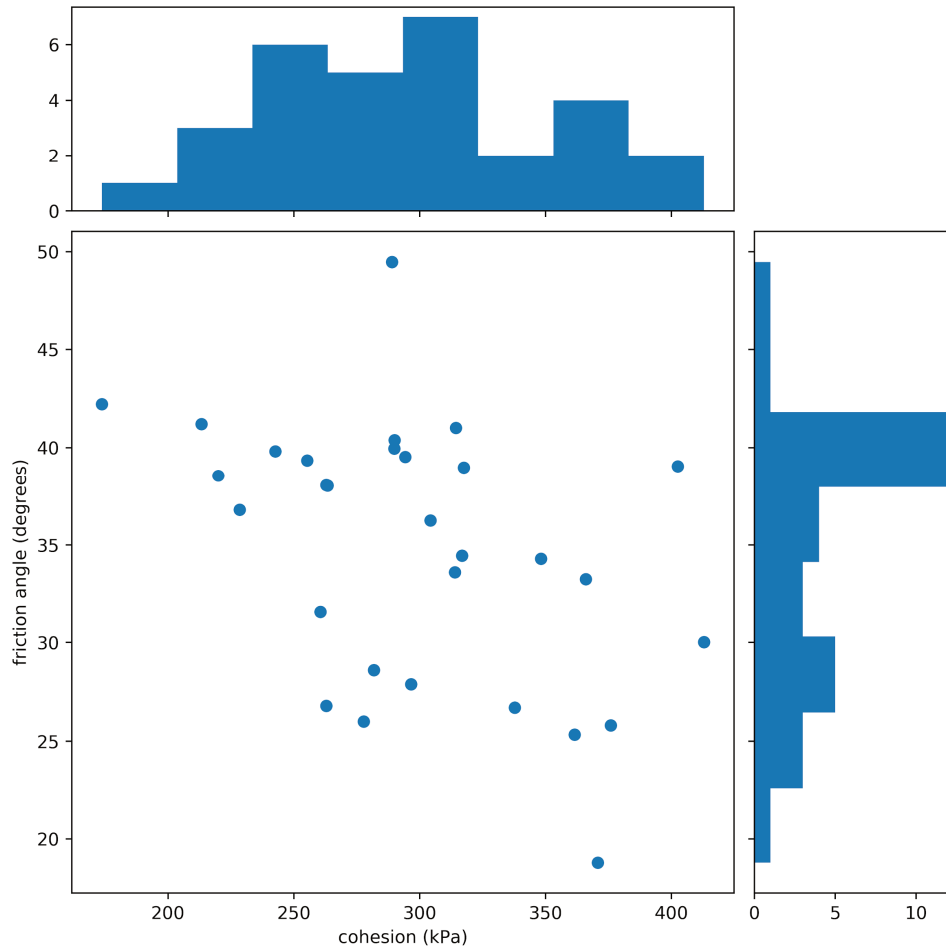


Figure 4 Scatter plot of simulated strength properties cohesion and friction angle, and their respective histograms

With the data shown in Figure 4, uncertainty intervals were constructed for a confidence level of 85% (confidence level for slope operation) using the Student's-t method (Equation 1). Figure 5 shows these intervals as error bar plots, where the central line represents the sample mean and vertical lines the uncertainty range.

As expected, it is possible to see that uncertainty intervals are decreasing in size as the sample size increases; however, this decreasing behaviour slows down due to the asymptotic behaviour shown in Figure 1, in fact, the last error bars look the same size.

Then, the asymptotic behaviour (proportional to $1/\sqrt{N}$) is an important characteristic because it indicates that for samples of 15 or 20 samples (see Figure 1) there is not a great gain with additional test, at least in terms of the uncertainty of the mean value. In fact, a convergence of the sample mean to the true mean ($c = 300$ kPa and $\phi = 35^\circ$) is observed in both cases, however, in a real case, this convergence could not be observed due to bad spatial sampling, this could be an indicator that additional tests are needed. It was also observed that for the correlation coefficient ρ , as greater is $|\rho|$, slow is the convergence to the true mean.

Safety factor intervals are shown in Figure 6. Additionally, the real safety factor $FS = 1.19$ is shown with a solid red line. In this case, an overestimation of the safety factor is observed for samples size smaller than 15 samples but due to the stochastic nature of the sample generation, the results eventually could have shown also an underestimation. However, a convergence of safety factor calculated with sample mean parameters to the real FoS occur, the same behaviour observed for the strength parameters. As sample size increases, the uncertainty of safety factor decreases, similar to Figure 6.

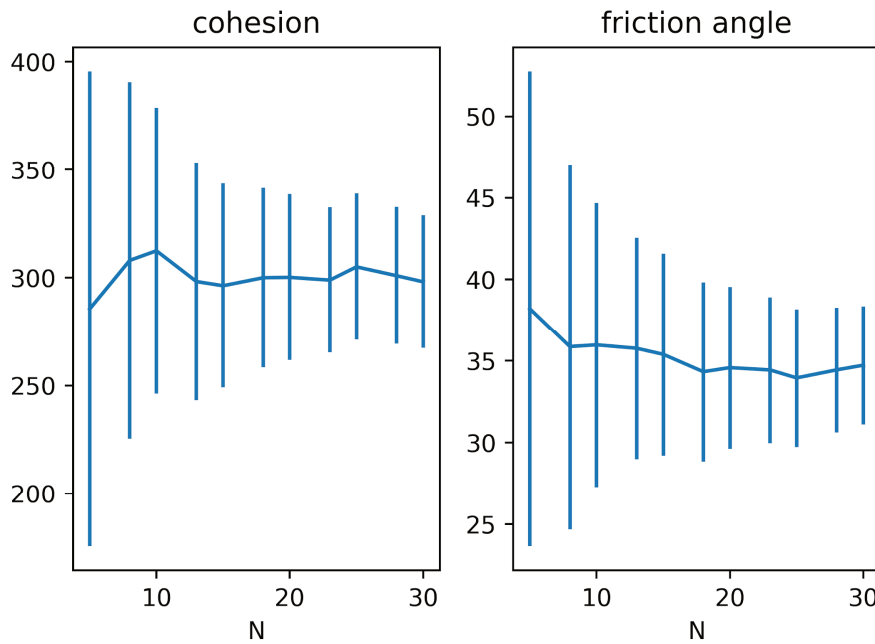


Figure 5 Error bar plots representing the uncertainty in the mean values of cohesion and friction angle. The central line shows the sample mean for different sample sizes

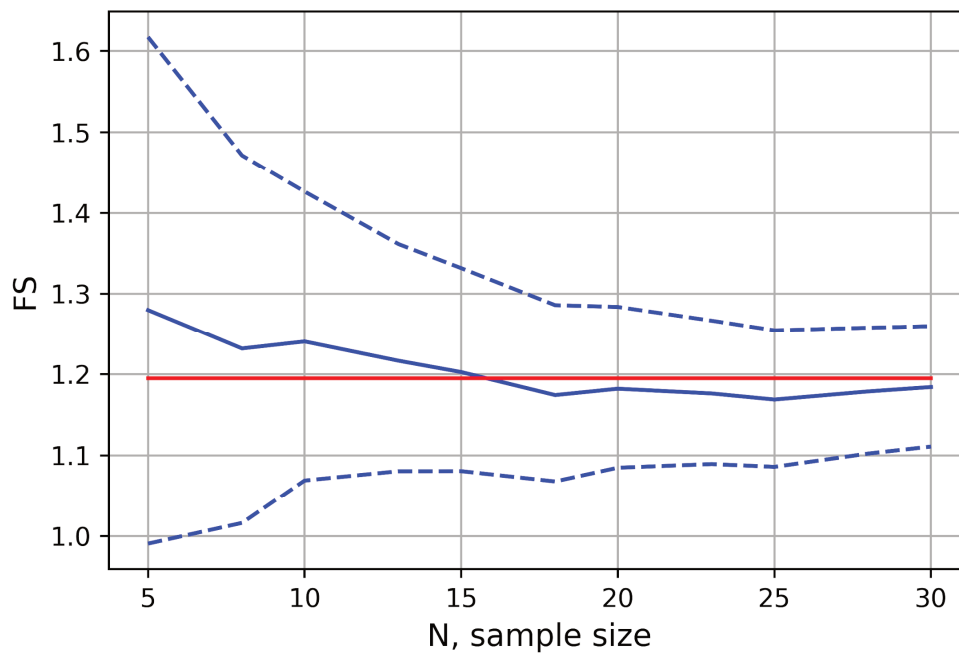


Figure 6 Factor of Safety intervals for different sample sizes with a 95% confidence level. Dashed lines show the lower and upper limits for the FoS interval, while the solid black line represents the FS for the sample mean properties. Red shows the real FS = 1.19 for the assumed distributions

4 Epistemic uncertainty and safety margin

The safety factor method is the most common and oldest method to evaluate the geotechnical design of a slope (Wyllie 2018). Its versatility is because for different levels of importance of a slope given by its risk level, either by damage severity or by frequency of failure events, it is possible to define different acceptability criteria for the safety factors. In recent years, probabilistic methods have been increasingly used in slope design. These methods are based on the calculation of the Probability of Failure (PoF) of the slope but the application of the epistemic uncertainty to the probability of failure calculation is an ongoing work.

Since safety and economic risk need to be considered and the risk tolerance of the companies is different, there is limited literature to define acceptability criteria. A useful starting point are the guidelines provided by Read & Stacey (2009), which take values widely used in the industry, as $FS = 1.2$ to 1.3 for non-critical slopes and $FS = 1.3$ to 1.5 for critical slopes. These values are based on the accumulated practical experience or simply to leave a safety margin of 20 to 50% between the representative value of the strength and the stress load, considering that there is a variability in the strength properties, geological conditions that are not identified and limitations in the analysis and modelling of site conditions, however, as was indicated in Section 3, the uncertainty of the strength mean values depends on the number of available tests with which it was defined and the required confidence level.

With respect to the confidence level of the information used to evaluate the designs, it depends on the stage of the project, which is shown in Table 2. In this manner, a potential benefit of the use of the concepts presented in this paper is the definition of sampling standards, both in information survey (drilling campaign and mapping) and in characterisation (laboratory tests), to diminish the uncertainty in the estimation of safety factor or to ensure the safety margin by comparing the acceptability criterion with the lower limit of the safety factor. This can be done for example, with the quantification of margins and uncertainty method (QMU) (Helton 2011); a decision-support methodology designed for complex technical decisions based on the quantification of the ratio (or distance) of the design margin to the model output uncertainty.

Table 2 Suggested confidence levels for the geotechnical information depending on the project stage (Read & Stacey 2009)

Project stage	Confidence level
Conceptual	Greater than 30%
Pre-feasibility	40–65%
Feasibility	60–75%
Design and construction	70–80%
Operation	Greater than 80%

To divide the uncertainty in aleatory variability and epistemic uncertainty also allows improvement of the classification of the former one, commonly described by means of the coefficient of variation. Generally, coefficient of variation less than 10% are considered low and the variability of the parameter is classified as acceptable, while coefficients greater than 30% are considered high, being the parameter's variability considered as excessive. However, these classifications do not consider the sample size used to calculate the mean and standard deviation and therefore, it is common to find inconsistencies as materials with poor sampling classified with low variability, and the contrary for well-sampled materials.

For the PoF approach, which involves the aleatory variability, we can consider also the uncertainty of standard deviation of the sample, making the presented procedure extendible to PoF calculation. This is ongoing work.

5 Conclusion

This paper shows the importance of epistemic uncertainty in the analysis of deterministic slope stability analyses. We outline two manners to quantify the level of uncertainty in strength parameters, both by means of two intervals: Student's-t confidence intervals and bootstrap intervals.

For sample sizes consisting of less than 10 tests, the Student's-t method is recommended. For sample size greater than 10 tests, this method is also recommended except in case of severe skewness, existence of far outliers or other signals of non-normality. In these cases, the bootstrap interval method is suggested.

After these intervals are generated, we show the use of error propagation law in a conceptual analysis of how uncertainty in strength parameters is propagated through the safety factor calculation. We can observe

that the friction angle propagation is maximum for predominantly cohesive soils, and vice versa with cohesion in predominantly frictional soils. While unlikely, the most critical situations are cohesive soils with high variability in friction angle and frictional soils with high variability in cohesion.

We show how the uncertainty limit proposed by Gill et al. (2005) in input data by means of the precision index might not be enough in term of the safety factor uncertainty.

Through an example, we show the entire process: interval construction for the parameters, and uncertainty propagation in safety factor calculation, for different sample sizes. We observe the expected decreasing size of the uncertainty but also an asymptotic behaviour in which additional information does not result in improved outcomes.

We discuss how this information can be incorporated in current safety factor's criteria, where (QMU) appears as promising.

This work was done with the goal of know better how to reduce the uncertainty with less resources. In Section 1, we presented questions to be answered:

1. How many tests are required to characterise mean strength parameters? The minimum sample can change from one to another rock, due to the uncertainty depends on the aleatory variability. To determine the minimum sample size, it is recommended to use the method proposed by Gill et al. (2005), followed by a simple check about the effect of parameters uncertainty over safety factor uncertainty, as the one presented in this paper.
2. In which materials is better to carry out more tests to cut down the uncertainty in safety factor? Figure 1 shows that sample standard deviation on its own is not a credible parameter to define the uncertainty because uncertainty depends strongly on the number of tests used to determine this standard deviation. The calculation of the uncertainty interval for a strength parameter allows us to determine the materials for which it is necessary to carry out more tests.
3. Finally, to understand the impact of the uncertainty of the stability at each stage of the project, we can first observe Table 2 in conjunction with Figure 1. Evidently, when a higher confidence level is required, greater samples sizes are needed, but also, when a high level of confidence is required, the effect of an insufficient characterisation is more severe.

References

- Ang, A & Tang, W 2007, *Probability Concepts in Engineering – Emphasis on Applications in Civil & Environmental Engineering*, 2nd edn, John Wiley & Sons, Hoboken.
- Baecher, G & Christian, J 2003, *Reliability and statistics in Geotechnical Engineering*, Chichester, Wiley.
- Baudrit, C & Dubois, D 2006, 'Practical representations of incomplete probabilistic knowledge', *Computational Statistics & Data Analysis*, vol. 51, no. 1, pp. 86–108.
- Bedi, A 2013, A proposed framework for characterizing uncertainty and variability in rock mechanics and rock engineering, PhD thesis, Imperial College London, London.
- Cao, Z, Wang, Y & Li, D 2017, *Probabilistic approaches for geotechnical site characterization and slope stability analysis*, Springer-Verlag, Berlin.
- Carranza-Torres, C & Hormazabal, E 2018, 'Computational tools for the determination of factor of safety and location of the critical circular failure surface for slopes in Mohr-Coulomb dry ground', *Proceedings of the 2018 International Symposium of Slope Stability in Open Pit Mining and Civil Engineering*, Seville.
- D'Agostino, R & Stephens, M 1986, *Goodness of fit techniques*, 1st edn, Marcel Dekker Inc, New York.
- Dempster, AP 1968, 'A generalization of Bayesian inference', *Journal of the Royal Statistical Society*, vol. 30, pp. 205–247.
- Diciccio, TJ & Romano, JP 1988, 'A review of bootstrap confidence intervals', *Journal of the Royal Statistical Society*, vol. 50, no. 3, pp. 338–354.
- Di Matteo, L, Valigi, D & Ricco, R 2013, 'Laboratory shear strength parameters of cohesive soils: variability and potential effects on slope stability', *Bulletin of Engineering Geology and the Environment*, vol. 72, no. 1, pp. 101–106.
- Dong, W & Shah, HC 1987, 'Vertex method for computing functions of fuzzy variables', *Fuzzy Sets and Systems*, vol. 24, no. 1, pp. 65–78.
- Drosg, M 2007, *Dealing with uncertainties - A guide to error analysis*, Springer-Verlag, New York.
- Dubois, D & Guyonnet, D 2011, 'Risk-informed decision-making in the presence of epistemic uncertainty', *International Journal of General Systems*, vol. 40, no. 2, pp. 145–167.

- Dubois, D & Prade, HM 2009, 'Formal representations of uncertainty', *Decision-making process*, ISTE & Wiley, London, pp. 85–156.
- Efron, B 1987, 'Better bootstrap confidence interval', *Journal of the American Statistical Association*, vol. 82, pp. 171–185.
- Fillion, MH & Hadjigeorgiou, J 2013, 'Reliability of strength estimates based on limited laboratory data', 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, https://doi.org/10.36487/ACG_rep/1308_05_Hadjigeorgiou
- Gill, DE, Corthesy, R & Leite, MH 2005, 'Determining the minimal number of specimens for laboratory testing of rock properties', *Engineering Geology*, vol. 78, pp. 53–67.
- Guan, JW & Bell, DA 1991, *Evidence theory and its applications*, North-Holland, New York.
- Helton, JC & Oberkampf, WL 2004, 'Alternative representations of epistemic uncertainty', *Reliability Engineering & System Safety*, vol. 85, no. 1–3, pp. 1–10.
- Helton, JC 2011, 'Quantification of margins and uncertainties: conceptual and computational basis', *Reliability Engineering & System Safety*, vol. 96, pp. 1034–1052.
- Jaulin, L, Kieffer, M, Didrit, O & Walter, E 2001, *Applied interval analysis*, Springer-Verlag, New York.
- Kearfott, RB & Kreinovich, V 1996, *Application of interval computations*, Kluwer Academic Publishers, Boston.
- Loy, A, Follet, L & Hofmann, H 2015, 'Variations of Q-Q plots: The power of your eyes!', *The American Statistician*, vol. 70, pp. 202–214.
- Moore, RE, Kearfott, RB & Cloud, MJ 2009, *Introduction to interval analysis*, Society for Industrial and Applied Mathematics, Philadelphia.
- Neyman, J 1937, 'Outline of a theory of statistical estimation based on the classical theory of probability', *Philosophical Transactions of the Royal Society of London, Series A, Mathematical and Physical Sciences*, vol. 236, no. 767, pp. 333–380.
- Read, J & Stacey, P 2009, *Guidelines for Open Pit Slope Design*, CRC Press, Boca Raton.
- Ruffolo, RM & Shakoor, A 2009, 'Variability of unconfined compressive strength in relation to number of test samples', *Engineering Geology*, vol. 108, pp. 16–23.
- Scholz, F 2005, 'Non-parametric tail extrapolation', Boeing Information & Support Services ISSTECH 95-014, Seattle, https://pdfs.semanticscholar.org/905a/ef3747263589c2e0a4b48b53dbf76d82d966.pdf?_ga=2.227757022.290737029.1582785789-2133554030.1580356856
- Walley, P 1991, *Statistical reasoning with imprecise probabilities*, Chapman and Hall, London.
- Wyllie, D 2018, *Rock Slope Engineering: Civil Applications*, 5th edn, CRC Press, Boca Raton.
- Zadeh, LA 1978, 'Fuzzy sets as a basis for a theory of possibility', *Fuzzy Sets and Systems*, vol. 1, pp. 3–28.

