Use of Bayesian Hierarchical Models to Estimate Geotechnical Parameters for Tailings

Holly Williams & Osvaldo Nicolás Ledesma
SRK Consulting, Vancouver, BC, Canada

GeoCalgary 2022 October 2-5
Reflection on Resources

ABSTRACT
There is a growing shift toward incorporating uncertainty through reliability-based approaches in the geotechnical industry. Typical frequentist approaches to statistics rely on large sample theory to quantify uncertainty. Bayesian approaches, on the other hand, can deal with uncertainties related to small sample sizes or limited knowledge. They can also incorporate many different types of information like engineering judgment and experience at similar sites. This makes it a natural fit to many geotechnical applications, including tailings dam design, where practitioners need to make design decisions based on a limited of number of samples. Hierarchical, or multi-level, Bayesian models have the additional benefit of providing a formal framework to incorporate related – but not identical – information. For example, information from other similar tailings facilities can still be useful even if it is not as directly relevant as data from the specific facility of interest.

This paper presents an example case study showing how a Bayesian hierarchical model can be used to quantify uncertainty and define effective friction angle estimates to use in stability assessments. In this example, consolidated undrained triaxial compression tests are available from four different facilities spread across one mine site. All the available data is used to quantify uncertainty across the site, within each facility, and when making predictions for a fifth facility where no data is available. Upon receiving new data for the fifth facility, the predictions are updated. While the presented case study provides an example for estimating effective friction angles, the same basic theory and model framework can easily be extended to any other parameter of interest.

RÉSUMÉ

1 INTRODUCTION

Selecting appropriate input parameter assumptions is a key step in any geotechnical design. Many practitioners apply a deterministic approach using conservative values to estimate a factor of safety (FoS) and then assess whether that FoS meets a defined threshold. Recently, there has been a growing shift in the industry to apply reliability-based approaches to estimate probabilities of failure. In either case, there is a need to incorporate uncertainty and apply engineering judgement to select appropriate parameter values.

Traditional frequentist statistics are applied in many other fields to quantify uncertainty and define distributions of parameters. These methods rely on large sample theory and assume there are enough samples to represent the true population. Geotechnical applications often deal with small sample sizes. Therefore, practitioners also need to rely on other sources of often subjective information that could include experience at similar sites or projects.

Bayesian approaches offer an advantage over frequentist statistics in that they provide a formal basis for incorporating multiple types of information (e.g., observed drilling or test data and engineering judgement), can deal with small sample sizes, and allow updating of parameters as new information becomes available. A special type of Bayesian models, called hierarchical or multi-level, have the additional benefit of providing a formal framework to incorporate related but not identical information. For example, available measurements of a parameter from multiple projects or sites should be incorporated as related information but should not carry the same importance as measurements from within the specific site of interest. For these reasons, Bayesian approaches are highly suitable to many geotechnical applications.

2 OVERVIEW OF BAYESIAN METHODOLOGY

2.1 General Bayesian Theory

In any statistical framework, there is typically a parameter of interest to estimate and available observations (data). In a Bayesian context, the parameter is linked to data using the general form of Bayes equation shown as Equation 1:

$$P(\theta|D) = [P(D|\theta)P(\theta)] / P(D) \quad [1]$$

where θ is the parameter of interest and D is the observed data.

The term $P(\theta)$, called the prior, represents our understanding of the parameter before seeing any data. This is where engineering judgement can be incorporated to represent knowledge about the parameter. The term $P(D|\theta)$, called the likelihood, ~~and~~ represents the probability of observing the data if prior understanding was correct. The result, $P(\theta|D)$, is the posterior and represents the updated probability distribution of the parameter of interest given the observed data. The denominator, $P(D)$, is the unconditioned probability of the data. Since it does not depend on the parameter of interest, θ, it is a constant that is simply used as a normalization factor. Therefore, Equation 1 can also be written as:

$$P(\theta|D) \propto P(D|\theta)\, P(\theta) \quad [2a]$$

or

$$\text{posterior} \propto \text{likelihood x prior} \quad [2b]$$

Solving the Bayesian formula can be very difficult, especially as the number of parameters being evaluated increases and when applying prior assumptions or likelihoods that do not necessarily easily combine with closed form solutions. Therefore, simulation methods such as Markov Chain Monte Carlo (MCMC) are typically used to approximate the posterior distribution. A more comprehensive description of MCMC methods and associated sampling algorithms is provided in Hastings (1970) and Metropolis (1953).

2.2 Bayesian Hierarchical Models

The framework of any Bayesian model is dependent on how data is considered. One approach treats all data the same, pooling it to estimate a single overall parameter as shown schematically in Figure 1a. The other approach treats all data independently to estimate values for each location as shown in Figure 1b. Reality is likely somewhere in between. Data may not be identical for every location or may not be entirely independent of data from another location. This is where a multilevel or hierarchical model is useful. In a hierarchical framework, there are multiple levels of parameters of interest that we are trying to estimate. In this case, the top level is an overall parameter that describes all tailings materials; however, we recognize there is also variability between locations (Figure 1c). Therefore, we say each location has its own value that is related to the overall parameter for all locations. As we run the model, the overall parameter acts as a prior for the parameter at each location. However, since the overall parameter is also unknown, it also needs a prior and is fit based on the values from all other locations. In other words, the fitting process is not sequential. Instead, we are fitting the overall and individual parameters simultaneously.

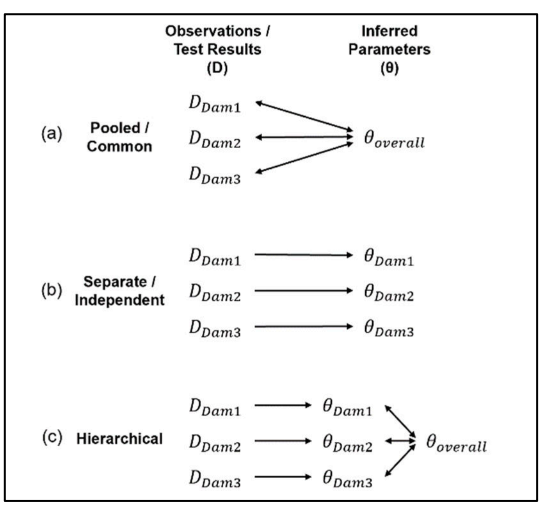

Figure 1. Comparison of Bayesian model frameworks

This approach has few key benefits including that it provides a formal basis for

- showing that predictions for locations with no available data have more variance than predictions for locations with data,
- borrowing information from locations that are data-rich to constrain locations that are data-poor, and
- accounting for sources of variability that may be important, even if the cause is not known (i.e., the variability between dams could be due to a changes in deposition, mineralogical characteristics of the mined ore, proportions of fines, etc.; however, we do not need to determine the exact cause to recognize and account for variability between locations).

Table 1 shows a simple comparison of the three main model frameworks. In the hierarchical model, the cross-location variability is captured in the first level and the variability within locations is captured in the second level.

Table 1. Comparison of Bayesian model frameworks

	Pooled / common	**Separate / Independent**	**Hierarchical**
Key idea	Fit one overall value using all observations	Fit unique value for each location	Fit overall value and individual values for each location simultaneously
Independence of observations	All observations are **identical**	Observations from each location are **independent** of other locations	Observations between locations are **related**
Making predictions at a new[1] site	Same predictions for all locations[2]	Not possible[3]	Uses overall value for predictions at new locations
Variability	overall variability in all observations	variability **within** locations	variability **across** and **within** locations

[1]A "new" location refers to a facility where there is no available information.
[2]The same value is predicted regardless of location.
[3]This is because there is no statistical basis for making a prediction at a new location if observations at any one location are assumed to be completely independent from observations at other sites

3 CASE STUDY

3.1 Background

The case study presented in this work is based on a real design project. Site-specific details have been omitted due to confidentiality restrictions; however, real laboratory testing results were used in all the analyses. The problem statement is also simplified for brevity.

The TSFs at the site were all built using a similar construction procedure. Tailings were initially deposited behind a compacted earth fill dam that was then raised in an upstream fashion (i.e., each dam raise was built on top of the previous raise and was partially founded on the newly deposited tailings). The resulting downstream slope of these raises is, approximately, 1.75H: 1.0 V. An example simplified cross section is shown in Figure 2.

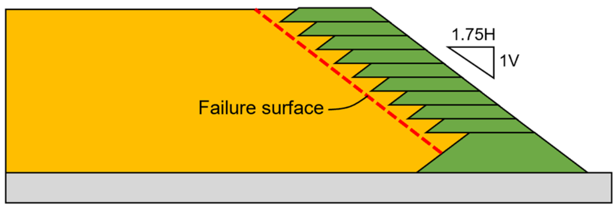

Figure 2. Example typical cross-section

Preliminary stability analyses indicated that none of the TSFs meet the required minimum factor of safety (FoS) of 1.5 for long-term conditions and static loads as recommended by current best practice guidance such as CDA (2013, 2019). Therefore, remediation efforts such as buttressing are likely needed for all TSFs.

The goal of the analysis is to make decisions on which TSF(s) to prioritize for remediation considering estimated deterministic FoS and probabilities of failure (PoF) while also accounting for the varying vulnerability of the different facilities.

3.2 Available Information and Approach

FoS and PoF estimates are calculated using a simplified infinite slope limit equilibrium analysis. Since there is no permanent phreatic level within the tailings, the stability is ~~therefore~~ solely controlled by the effective friction angle of the stored tailings at the contact between the stored tailings and the upstream dam shell.

Several site investigations have been completed, including performing consolidated undrained triaxial tests on tailings samples from the different TSFs.

The case study analysis is completed in two parts. In Part 1, laboratory strength tests are available for only four of the five TSFs. In Part 2, laboratory strength tests for TSF 4 became available, and the analysis is updated using this new information.

For Part 2, there are 36 consolidated undrained triaxial test results available at the site, with almost half of the tests coming from TSF 5, as summarized in Table 2. Each triaxial test result corresponds to one consolidated undrained triaxial compression at a constant confining stress (i.e., one observation).

Table 2. Available triaxial stage test results

TSF	**Number of Observations**	
	Part 1	**Part 2**
1	3	3
2	3	3
3	9	9
4	-	6
5	15	15

4 METHODOLOGY

Models for all three main types of Bayesian frameworks (pooled, separate, and hierarchical) were developed for comparative purposes. For all models, the strength of the tailings is characterized by its effective friction angle (ϕ) at the critical state. Therefore, the parameters of interest are the effective friction angles for each dam, which are determined based on the results of consolidated undrained triaxial compression tests (the observations).

For each model, the principal stresses at failure for each test specimen are inputted as the data (or evidence) and posterior samples of the parameter of interest, ϕ, are collected through Markov Chain Monte Carlo (MCMC) simulation methods. If enough samples are collected and the model is a well-specified model, values of ϕ will converge toward the range of values that best describe the data while also factoring in the prior assumptions.

4.1 Modelling Tools and Diagnostics

The Python library PyStan (Carpenter et al., 2017) was used to perform Hamiltonian Monte Carlo with No-U-Turn Sampling MCMC method (Hoffman and Gelman, 2014). Four chains were run for every model and the $\hat{R}$ diagnostic was used to assess model performance. $\hat{R}$ compares between-chain and within-chain variances (see Gelman and Rubin (1992) for details). High values indicate non-convergence and therefore a target $\hat{R}$ value of 1.0 ±0.05 was used to verify model convergence and stability.

4.2 Prior Selection

The first step in any Bayesian framework is to define the prior assumptions before seeing any data. In this case, we need to define priors for the effective friction angle and uncertainty (standard deviation).

The prior distribution has been defined in terms of the tangent of the friction angle, tan(ϕ), since stability (or FoS) is linearly related to tan(ϕ). Based on experience, we expect tailings effective friction angles to be restricted between 20° and 45°. We capture these assumptions by assuming tan(ϕ) is normally distributed with tan(20°) and tan(45°) representing the lower and upper bounds of a 99.7% confidence interval. This results in an assumed prior mean of 0.667 (or about 33.7°) and a standard deviation of 0.106 representing six standard deviations between tan(20°) and tan(45°).

A prior assumption also needs to be defined for the standard deviation or variance parameter. The half-cauchy distribution was selected as a weakly informative prior, with a location parameter equal to 0.106. Gelman (2006) and Polson and Scott (2012) describe the benefits of using the half-cauchy distribution for variance parameters in hierarchical models.

An alternative approach could have chosen completely uninformative uniform priors bounding the friction angle to be between 0° and 90°, and the standard deviation to be between 0 and infinity. Instead, both priors apply weakly informative assumptions to incorporate some judgement while still being open-ended enough to not overly bias or influence the results.

4.3 Pooled Model

For the pooled model, we assume there is only one effective friction angle, ϕ, that can be estimated from all our observed principal stresses at failure at all locations. We also know there is some inherent variability or error in the relationship linking tan(ϕ) to our measurements.

This assumption means we have two parameters that are unknown and need priors, tan(ϕ), and the standard deviation, σ. The governing prior and likelihood equations are as follows:

Priors:

$$\mu_{\tan(\phi)} \sim \mathcal{N}(0.667, 0.106) \quad [3]$$

$$\sigma_{\tan(\phi)} \sim \text{Half-Cauchy}(0, 0.106) \quad [4]$$

Likelihood:

$$y \sim \mathcal{N}(\mu_{\tan(\phi)}, \sigma_{\tan(\phi)}) \quad [5]$$

Where y is the measured tan(ϕ) for each principal stress at failure.

One posterior distribution is generated, representing the overall tan(ϕ) for all TSFs.

4.4 Separate Model

For the separate model, we assume there is a unique effective friction angle, ϕ, for each location that we estimated using only the information available at that location and ignoring information from other locations.

Like the pooled model, we know there is some inherent variability or error in the relationship linking tan(ϕ) to our measurements and represent it using the same normal distribution assumption. However, in this case, we also assume that the standard deviation, σ, is unique to each TSF.

Priors:

$$\mu_{\tan(\phi),\, TSF} \sim \mathcal{N}(0.667,0.106) \quad [6]$$

$$\sigma_{\tan(\phi),\, TSF} \sim \text{Half-Cauchy}(0,0.106) \quad [7]$$

Likelihood:

$$y_{TSF} \sim \mathcal{N}(\mu_{\tan(\phi),\, TSF}, \sigma_{\tan(\phi,\, TSF)}) \quad [8]$$

Multiple posterior distributions are generated, one for each TSF.

4.5 Hierarchical Model

The hierarchical model is similar to the separate model in that we assume there is a unique tan(ϕ) for each TSF; however, there are a few key differences. First, we introduce an extra level to our model by assuming each locations' tan(ϕ) is related to one common overall effective friction angle for all TSFs. We also assume the standard deviation, σ, in the relationship linking tan(ϕ) estimates to our measurements is shared across all TSFs. All the available information for all TSFs is used as evidence for estimating overall values. Each location's tan(ϕ) uses information from its own location observations as evidence, but it also uses the overall mean and standard deviation as a prior. This is how information gets shared across TSFs. As evidence increases, it gets weighted higher; however, when there is little evidence, the prior information from all other dams has a stronger contribution.

Priors:

$$\mu_{\tan(\phi)} \sim \mathcal{N}(0.667,0.106) \quad [9]$$

$$\sigma_{\tan(\phi)} \sim \text{Half-Cauchy}(0,0.106) \quad [10]$$

$$\mu_{\tan(\phi),\, TSF} \sim \mathcal{N}(\mu_{\tan(\phi)}, \sigma_{\tan(\phi)}) \quad [11]$$

$$\sigma_{\tan(\phi),\, TSF} \sim \text{Half-Cauchy}(0,0.106) \quad [12]$$

Likelihood:

$$y_{TSF} \sim \mathcal{N}(\mu_{\tan(\phi),\, TSF}, \sigma_{\tan(\phi,\, TSF)}) \quad [13]$$

This model includes one extra posterior distribution compared to the separate model since one unique tan(ϕ) is generated for each location, as well as one for representing the overall tan(ϕ) for all the TSFs. This overall tan(ϕ) can be used to make out-of-sample predictions at a new TSF.

4.6 Factor of Safety and Probability of Failure Calculations

The 33rd percentile value of tan(ϕ) generated from the posterior distribution was used as the key input for deterministic factor of safety (FoS) calculations. As previously mentioned, FoS was calculated using infinite slope limit equilibrium analysis. The PoF is calculated as follows:

$$PoF=P[FoS\leq 1]=P[\tan(\phi)\leq\tan(\beta)] \quad [14a]$$

$$= \Phi[(\tan(\beta)-\mu_{\tan(\phi)})/\sigma_{\tan(\phi)}] \quad [14b]$$

where tan(β) is the slope of the downstream dam shell, Φ is the is the standard normal cumulative distribution, and $\mu_{\tan(\phi)}$ and $\sigma_{\tan(\phi)}$ are the mean and standard deviation values for tan(ϕ) determined from the Bayesian modelling.

5 RESULTS

All results in this section show only one value for the pooled models, and do not include any values for TSF 4 in the Part 1 separate models. This is because pooled models only calculate one value that applies to all TSFs, and separate models do not allow out-of-sample predictions for TSFs with no data (Part 1 does not include any tests at TSF 4).

Table 3 shows the mean and standard deviation values for tan(ϕ) determined from the Bayesian modelling.

Table 3. Posterior tan(ϕ) results

Statistical Model	Part	tan(ϕ), mean ± standard deviation				
		TSF 1	TSF 2	TSF 3	TSF 4	TSF 5
Pooled	1	0.708 ± 0.133				
	2	0.704 ± 0.121				
Separate	1	0.693 ±0.023	0.799 ±0.051	0.656 ±0.042	-	0.712 ±0.039
	2	0.693 ±0.022	0.800 ±0.050	0.654 ±0.042	0.689 ±0.035	0.712 ±0.037
Hierarchical	1	0.696 ±0.019	0.742 ±0.055	0.684 ±0.032	0.705 ±0.060	0.712 ±0.029
	2	0.698 ±0.018	0.734 ±0.048	0.687 ±0.030	0.698 ±.024	0.709 ±0.028

Table 4 shows the resulting effective friction angle ϕ values and confidence intervals. Note that while tan(ϕ) results follow an approximate normal distribution characterized by a mean and standard deviation, the effective friction angle ϕ does not, since it is a non-linear transformation of tan(ϕ).

Figure 3 shows the same information as Table 4 and includes the equivalent 33rd percentile value.

Table 6 shows the calculated FoS and PoF results.

6 DISCUSSION

6.1 Differences Between Model Frameworks

Figure 4 shows the posterior distributions of the estimated effective friction angles for TSF 4. The top plot shows the results for Part 1, before any tests were available at TSF 4, and the bottom plot shows the results after incorporating TSF-specific tests.

Table 4. Friction angle (ϕ) results and confidence intervals

Statistical Model	Part	Friction angle, mean (95% confidence interval)				
		TSF 1	TSF 2	TSF 3	TSF 4	TSF 5
Pooled	1	35.30 (45.24 - 25.35)				
	2	35.15 (43.73 - 25.87)				
Separate	1	34.72 (36.69 - 32.99)	38.55 (41.66 - 34.78)	33.27 (36.4 - 29.66)	-	35.47 (38.21 - 32.62)
	2	34.72 (36.38 - 32.89)	38.7 (41.65 - 35.14)	33.18 (36.43 - 29.79)	34.57 (37.12 - 31.83)	35.45 (38.42 - 32.55)
Hierarchical	1	34.86 (36.36 - 33.49)	36.58 (40.41 - 33.55)	34.36 (37.08 - 31.95)	35.16 (39.92 - 30.75)	35.32 (37.67 - 33.23)
	2	34.90 (36.31 - 33.57)	36.27 (39.80 - 33.3)	34.49 (36.72 - 31.90)	34.90 (36.64 - 32.93)	35.33 (37.60 - 33.31)

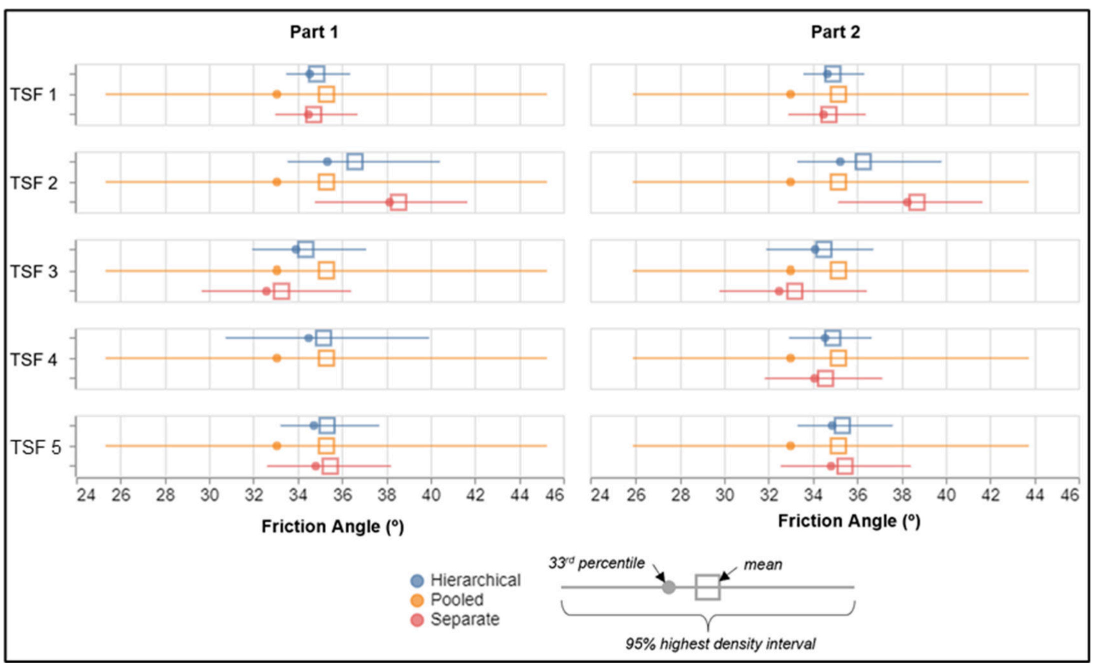

Figure 3. Friction angle results and confidence intervals

Table 5. FoS and PoF results

Statistical Model	Part	FoS (PoF)				
		TSF 1	TSF 2	TSF 3	TSF 4	TSF 5
Pooled	1			1.14 (1.52E-01)		
	2			1.14 (1.37E-01)		
Separate	1	1.20 (6.26E-08)	1.36 (4.06E-06)	1.12 (2.20E-02)	-	1.22 (7.26E-05)
	2	1.20 (1.64E-08)	1.36 (2.42E-06)	1.11 (2.47E-02)	1.18 (3.91E-04)	1.22 (1.56E-04)
Hierarchical	1	1.20 (2.76E-11)	1.26 (9.63E-04)	1.17 (2.18E-04)	1.19 (1.30E-02)	1.22 (1.05E-06)
	2	1.21 (1.02E-12)	1.25 (3.53E-04)	1.18 (5.85E-05)	1.20 (6.68E-08)	1.22 (4.48E-07)

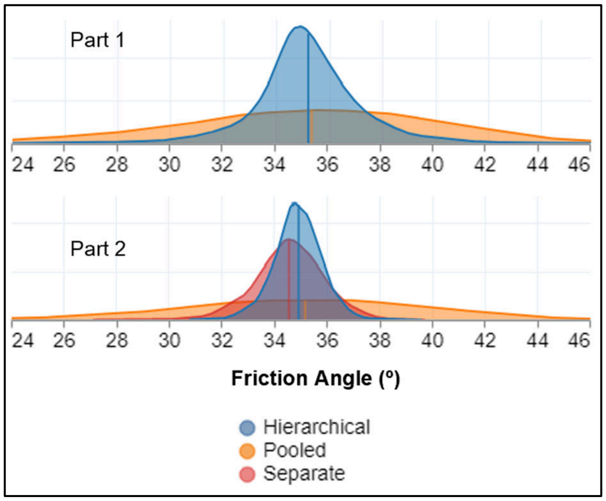

Figure 4. Posterior distributions of friction angles, ϕ, for TSF 4

The pooled model results (shown in orange in Figure 4) have the largest variance and uncertainty. As shown in Figure 3, the predictions are also identical for every dam and centered around the average of all the available data across all dams. Comparing Part 1 and Part 2, adding the new Dam 4 information has a very minor impact of slightly reducing uncertainty, but the results are otherwise relatively insensitive to new information.

There are no separate model results (shown in red in Figure 4) available for TSF 4 in Part 1 since the framework does not have a mechanism to make predictions for TSFs with no information. In Part 2, we see that TSF 4 results are slightly lower compared to the pooled models that use all the data.

The hierarchical model results (shown in blue in Figure 4) have the lowest variance and uncertainty. In Part 2, we also see the hierarchical model mean lies in between the means of the pooled and separate models. This is a common characteristic of hierarchical models and is generally referred to as “shrinkage towards the mean”. Both the reduced variance and shrinkage effects are a result of the model’s ability to incorporate all available information while weighting TSF-specific information as more relevant. The hierarchical model also has much lower uncertainty compared to the pooled model in Part 1. This is because the pooled lumps all the data together and looks at the overall variance. The decreased variance in the hierarchical model is due to the data being grouped by TSF and accounting for both within-TSF and between-TSF variances.

Hierarchical models have clear benefits of incorporating all available information, reducing uncertainty, and allowing relative comparisons between TSFs when compared to the other model frameworks. For these reasons, the hierarchical model is adopted the preferred method and only these results will be presented in subsequent sections.

6.2 Factor of Safety Versus Probability of Failure

Figure 5 shows the FoS results for all dams for Parts 1 and 2. There are only very minor differences between Part 1 and 2 since the overall means do not change very much. Clearly none of the dams meet the minimum target FoS of 1.5.

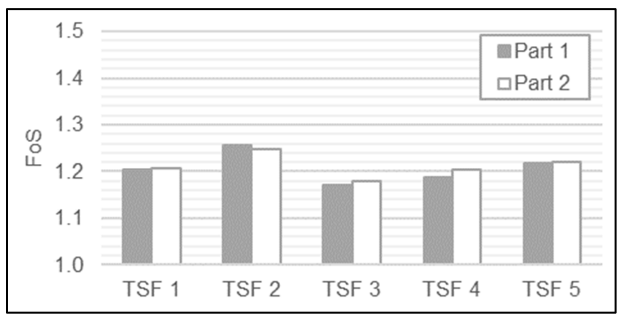

Figure 5. Calculated factor of safety

Looking at this information alone makes it very difficult to make decisions on which TSF(s) need to be prioritized for remediation. Deterministic FoS values do not include any consideration of uncertainty and the fact that different TSFs have varying levels of information.

PoF estimates, on the other hand, do incorporate uncertainty and are better suited to performing relative comparisons and determining prioritizations. Figure 6 shows the PoF results for Parts 1 and 2. PoFs decrease in Part 2 compared to Part 1 for all TSFs due to new information being available. For most TSFs, the change is minor, but at TSF 4 the difference is significant since it went from having no information in Part 1 to having site specific information in Part 2. This results in major changes in the overall patterns:

- In Part 1, TSF 4 has the highest PoF, but this is mainly due to it having the most uncertainty since no TSF-specific information is available.
- In Part 2, we add information for TSF 4 and then see the PoF drop multiple orders of magnitude below TSF 2 and 3 and comparable to TSF 5.

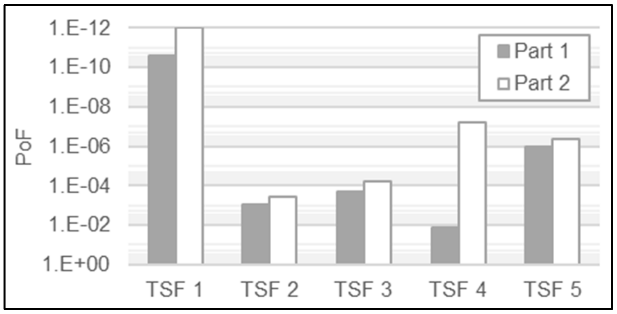

Figure 6. Calculated probability of failure

Revisiting Figure 4 and comparing the TSF 4 results for Part 1 and 2, shows this reduction in PoF is not due to a change in our best estimate of the mean. (We see the mean reduces slightly or gets worse.) Instead, it is due to the large reduction in uncertainty as a result of having dam-specific information. This aligns with what we would expect using our engineering judgement and helps highlight the value of getting more information.

6.3 Implications on Engineering Decisions

The key question is how to use all this information to make engineering decisions on which TSF(s) to prioritize for stabilization. If one only considers deterministic FoS estimates, it is very difficult to make any decision because all TSFs have similar values. If forced to decide, one might choose to prioritize TSF 3 because it has the lowest FoS value. If we only looked at PoF estimates, TSF 2 and 3 have the lowest values and might both be selected for priority stabilization.

However, if we also inspect the overall results (Figure 7), we see TSF 2 has higher uncertainties compared to other TSFs. This is because the limited available test results at TSF 2 plot significantly higher than the averages at other TSFs. This suggests there could be value in performing additional tests to confirm whether these higher strengths are true and potentially narrow the uncertainty.

Therefore, the recommended approach could include stabilizing TSF 3 as an immediate priority and conducting additional tests at TSF 2 before deciding on the next TSF to stabilize.

7 SUMMARY AND CONCLUSIONS

Bayesian approaches offer many advantages and are a natural fit to many geotechnical applications. This case study provides one example comparing three types of Bayesian modelling frameworks to define and quantify the uncertainty of geotechnical parameters to use in stability analyses for TSFs. Ultimately, the results were used to prioritize remediation efforts.

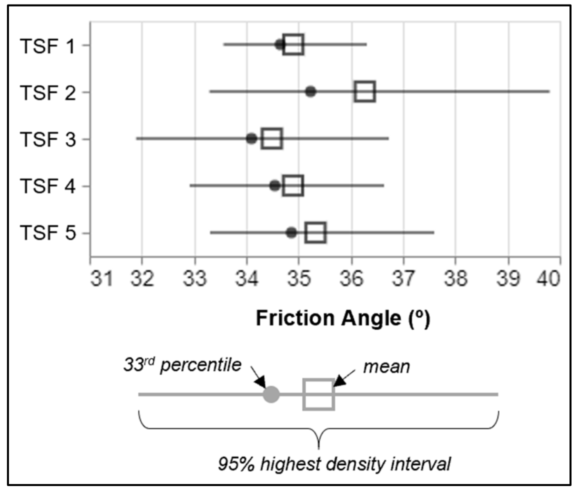

Figure 7. Posterior distributions of Part 2 hierarchical model friction angles, ϕ, for all TSFs

The key findings were as follows:

- Bayesian approaches offer an advantage over frequentist statistics in that they provide a formal basis for incorporating multiple types of information such as observed drilling or test data and engineering judgement and can deal with small sample sizes. In this case, available test data included consolidated undrained triaxial compression tests from multiple TSFs and judgement was incorporated through the assignment of priors representing expected ranges of effective friction angles.
- Hierarchical models have clear benefits of incorporating all available information, reducing uncertainty, and allowing relative comparisons between dams when compared to other Bayesian model frameworks.
- Updating the analysis after receiving additional test results highlighted the value of receiving new information.
- Insights from the full results can be used to make practical engineering recommendations. In this case, the recommendation included prioritizing one TSF for immediate remediation and selecting another to be investigated further.

While the case study presented in this paper focused on one practical application of estimating effective friction angles, the same basic theory and model framework can easily be extended to any other geotechnical parameter of interest.

8 REFERENCES

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. 2017. Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1), 1–32.

Canadian Dam Association (CDA). 2013 Dam Safety Guidelines. *Canadian Dam Association*, Markham, Ontario, Canada, 2013 Revision.

Canadian Dam Association (CDA). 2019 Dam Safety Guidelines. *Canadian Dam Association*, Markham, Ontario, Canada, 2019 Draft Revision.

Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, Vol. 1, N. 3, pp. 515–533

Gelman, A. and Rubin, D. B. 1992. Inference from iterative simulation using multiple sequences. *Statistical science*, 457–472.

Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57, No. 1, 97–109.

Hoffman, M.D., and Gelman, A. 2014. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15: 1351-1381.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. 1953. Equation of state calculations by fast computing machines. *The journal of chemical physics* 21, No. 6, 1087–1092.

Polson, N.G. and Scott, J.G.2012), On the Half-Cauchy Prior for a Global Scale Parameter. *Bayesian Analysis*, Vol. 7, No. 4, pp. 887-902