Keynote Address

Use of probabilistic methods in geotechnical engineering

Jinsong Huang M.AGS M. ASCE

Discipline of Civil, Surveying and Environmental Engineering, College of Engineering, Science and Environment, The University of Newcastle, Callaghan NSW 2308, Australia; PH: +61 2 4921 5118; email: Jinsong.huang@newcastle.edu.au

ABSTRACT

Due to the intrinsic inhomogeneous nature of soils and rocks, the minimal site investigations, and the need to extrapolate available information over a large domain, geotechnical designs have inevitable uncertainties. To be conservative, geotechnical engineers traditionally use a safety factor to account for uncertainties. A more rigorous way of considering uncertainties is to use probabilistic methods. To promote the use of probabilistic methods in geotechnical engineering, this paper tries to address the following commonly encountered questions: 1) Why do we need to use probabilistic methods? 2) How can we use probabilistic methods if we don't have enough test data? 3) How much field/test data do we need? 4) How can we use multiple sources of information? 5) How can we use monitoring data to predict future performance?

Keywords: probabilistic methods, geotechnical engineering, limited data, Bayesian methods, prediction

1 INTRODUCTION

In recent years, there has been a remarkable increase in activity and interest in the use of probabilistic methodologies applied to more traditional areas of geotechnical engineering. This growth has manifested itself in many forms and spans both academe and practice within the geotechnical engineering community, for example, more dedicated conferences, short courses for practitioners, and new journals and books. However, there are still some reluctances. This paper tries to promote the use of probabilistic methods in geotechnical engineering by addressing some commonly encountered questions.

2 WHY DO WE NEED TO USE PROBABILISTIC METHODS?

2.1 Limited site investigation data

Soils and rocks in their natural state are among the most variable of all engineering materials, and geotechnical engineers must often "make do" with materials that present themselves at a particular site. In a perfect world with no economic constraints, we would drill numerous boreholes and take multiple samples back to the laboratory for measurement of standard soil properties such as permeability, compressibility, and shear strength. Armed with all this in-formation, we could then perform our design of a seepage problem, foundation, or slope and be very confident of our predictions. In reality we must usually deal with very limited site investigation data, and the traditional approach for dealing with this uncertainty in geotechnical design has been through the use of characteristic values of the soil properties coupled with a generous factor of safety.

If we were to plot the multitude of data from the hypothetical site investigation as a histogram for one of the properties, we would likely see a broad range of values in the form of a bell-shaped curve. The most likely values of the property would be somewhere in the middle, but a significant number of samples would display higher and lower values too. This variability inherent in soils and rocks suggests that geotechnical systems are highly amenable to a statistical interpretation. This is quite a different philosophy from the traditional approach mentioned previously. In the probabilistic approach, we input soil properties characterised in terms of their means and variances leading to estimates of the probability of failure or reliability of a design.

2.2 Probability of failure is more meaningful than factor of safety

It is common to apply the same factor of safety for different types of application without regard to the degree of uncertainty or failure consequence involved. In this case, the factor of safety can be misleading because a higher factor of safety doesn't necessarily mean the structure is safer.

Considering two examples of drained slope stability, they both have the same geometry as shown in Figure 1. The characteristic values of the soil properties in the two examples are shown in Eq. (1) and (2) respectively.

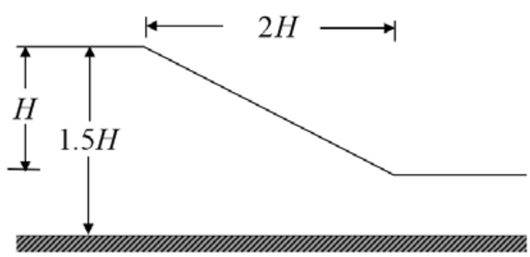

Figure 1. Slope profile

$$\phi' = 23^\circ \qquad \frac{c'}{\gamma H} = 0.048 \tag{1}$$

$$\phi' = 32^\circ \qquad \frac{c'}{\gamma H} = 0.048 \tag{2}$$

where c' and ϕ' are drained cohesion and friction angle respectively, and γ is unit weight of soil.

According to the stability charts in Michalowski (2002), the factor of safety for the two slopes can be found as 1.5 and 2.0 respectively. We may think that the second slope is safer than the first one because of a higher factor of safety. However, if we use a probabilistic approach, we may find that the mean and standard deviation of factor of safety for the two slopes are $\mu_{FS} = 1.5,\ \sigma_{FS} = 0.18$ (slope 1), and $\mu_{FS} = 2.0,\ \sigma_{FS} = 0.5$ (slope 2). Assuming the factor of safety is normally distributed, one can easily calculate the corresponding probabilities of failure of the two slopes are 0.27% and 2.27% respectively. It is clear that the second slope is actually less safe than the first one because of a higher probability of failure. The main reason why factor of safety can be misleading is that it uses only one parameter to quantify safety level. In probabilistic approaches, however, two parameters are used to quantify not only the mean safety level but also the standard deviation.

The factor of safety cannot provide meaningful information for risk assessment. For the two examples we just considered, both factors of safety are bigger than one which means both slopes are safe. By looking at the factor of safety along, it is hard for engineers to choose among the two slopes.

On the contrary, when probabilities are coupled with consequences of design failure, we can assess the risk associated with the design. For the examples we just considered, if the construction cost and consequence of design failure are estimated as shown in Table 1, the risk of the two designs can be estimated as Risk= construction cost + probability of failure × consequence.

It can be seen from Table 1 that the second slope with a higher probability of failure has lower risk than the first slope. However, if the consequence increases from 10M to 100M, the first slope with a lower probability of failure will have a lower risk than the second slope. This is consistent with the common sense that for important project, we need to lower the probability of failure to reduce risk.

Table 1: Risk assessment of slope failure

	Slope 1		Slope 2	
Construction cost (million dollar)	1		0.5	
Probability of failure	0.27%		2.27%	
Consequence (million dollar)	10	100	10	100
Risk (million dollar)	1.027	1.27	0.727	2.77

3 HOW CAN WE USE PROBABILISTIC METHODS IF WE DON'T HAVE ENOUGH TEST DATA?

We actually have two schools of statistics. The first one is frequentist approach. Frequentist says that we need to have enough data to draw a statistical conclusion, for example, a probability distribution. This is the dominant statistical practice during the 20th century.

The other approach is Bayesian methods. Bayesian says that before we do any test, we have some prior knowledge about the test we are looking at. And data can be used to update our prior knowledge. Bayesian methods started getting popular in 21st century, although it was published 200 years ago. It is now widely used in computer science, medical science, social science and so on.

The Bayesian approach has special significance to engineering design where available information is invariably limited and subjective judgement is often necessary. In the case of parameter estimation, we may often have some knowledge (perhaps inferred intuitively from experience) of the possible values, or range of values, of a parameter; moreover, we may also have some intuitive judgment on the values that are more likely to occur than others.

In Bayesian statistics, site investigation information, laboratory test data, in situ test data and load tests results are used as evidence to update prior knowledge, predictions and designs. The posterior distribution of the parameters is calculated based on Bayes' theorem:

$$P(\theta \mid D) = \frac{L(D \mid \theta) P(\theta)}{P(D)} \tag{3}$$

where θ denotes the parameters to be updated, D the data collected, $P(\theta)$ the prior probability distribution of the parameters, $L(D \mid \theta)$ the likelihood probability distribution and $P(D)$ the probability distribution of the evidence.

To demonstrate how Bayes' theorem can be used when test data are limited, let's consider load testing of pile capacity. Before any test is performed, engineers can look in the literature or similar projects for the initial assessment of the pile capacity (y). For simplicity, let's assume the pile capacity is normally distributed with mean μ and standard deviation σ. The standard deviation σ can be estimated based on the variability of the site (e.g., Kay (1978)). The mean pile capacity μ is uncertain, and is initially estimated using the prior probability distribution expressed by

$$f'(\mu) \sim N(\mu_0, \sigma_0) \tag{4}$$

where $N()$ represents the Normal distribution, and μ_0 and σ_0 are the mean and standard deviation of μ (mean of mean pile capacity, standard deviation of mean pile capacity). The parameters μ_0 and σ_0 were assumed to be associated with prediction methods (a static or dynamic method).

Suppose n pile capacity tests have been performed and the results are denoted as the vector $\hat{\mathbf{y}}$. The posterior distribution of μ can be obtained by Bayes' rule:

$$f''(\mu|\hat{\mathbf{y}}) = \frac{P(\hat{\mathbf{y}}|\mu) f'(\mu)}{P(\mathbf{y})} \tag{5}$$

$$\propto \exp\left(- \frac{\left(\mu - \frac{n\sigma_0^2 \bar{y} + \sigma^2 \mu_0}{n\sigma_0^2 + \sigma^2} \right)^2}{2\left(\frac{\sigma_0^2 \sigma^2}{n\sigma_0^2 + \sigma^2} \right)} \right)$$

where $\bar{y}$ is the average value of $\hat{\mathbf{y}}$.

A detailed derivation of Eq. (5) can be found in Huang et al. (2016).

The posterior mean and standard deviation of μ are

$$\mu_1 = \frac{n\sigma_0^2 \bar{y} + \sigma^2 \mu_0}{n\sigma_0^2 + \sigma^2} \tag{6}$$

and

$$\sigma_1 = \frac{\sigma_0^2 \sigma^2}{n\sigma_0^2 + \sigma^2} \tag{7}$$

Let us assume the prior mean and standard deviation of μ are $\mu_0 = 1.3$ and $\sigma_0 = 0.5$ and suppose five load tests have been conducted. The average capacity of tested piles is $\bar{y} = 0.8$. According to Eqs. (6) and (7), the posterior mean and standard deviation of μ are

$$\begin{aligned} \mu_1 &= \frac{n\sigma_0^2 \bar{y} + \sigma^2 \mu_0}{n\sigma_0^2 + \sigma^2} \\ &= \frac{5 \times 0.5^2 \times 0.8 + 0.2^2 \times 1.3}{5 \times 0.5^2 + 0.2^2} \\ &= 0.82 \end{aligned}$$

and

$$\begin{aligned} \sigma_1 &= \frac{\sigma_0 \sigma}{\sqrt{n\sigma_0^2 + \sigma^2}} \\ &= \frac{0.5 \times 0.2}{\sqrt{5 \times 0.5^2 + 0.2^2}} \\ &= 0.09 \end{aligned}$$

The uncertainty of the mean capacity (μ) has been reduced significantly from $\sigma_0 = 0.5$ to $\sigma_1 = 0.09$ by the five load tests. It is interesting to note that even when we have only one test, we can still use Eqs. (6) and (7).

4 HOW MUCH FIELD/TEST DATA DO WE NEED?

Suppose we want to do an excavation on a site as shown in Figure 2(a) and the final slope profile is shown in Figure 2(b). The grayscale in Figure 2(b) represents different undrained strengths.

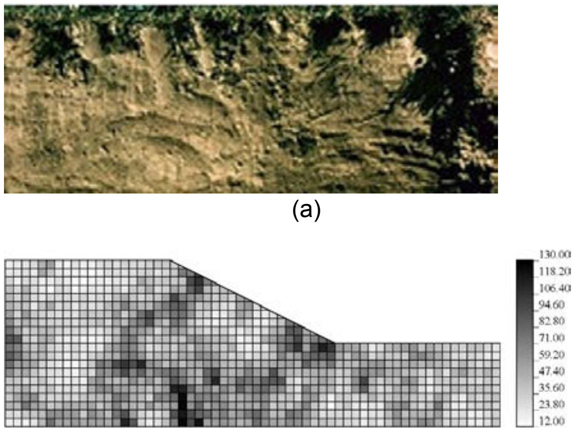

Figure 2. True in-situ distribution of undrained shear strength

Before we do the excavation or test, we actually don't know the undrained shear strength distribution in Figure 2(b). We can look in the literature for the statistics of the undrained shear strength of the soil in this particular site. And based on the statistics, we can use random field theory to guess the possible spatial distributions of the undrained shear strengths and perform a set of Monte Carlo simulations as shown in Figure 3. We generate a set of realisations of the slope, some of them will fail (e.g., Figure 3 a), some will not (e.g., Figure 3 b). If you divide the total number of failures by the total number of simulations, we can estimate the probability of failure, which is 17% in this case.

For the same slope, if we perform a cone penetration test as indicated by the red box in Figure 4, we then know the undrained shear strength of this column. Now we have statistics of the undrained shear strengths plus some test data. Based on the statistics and data, we can perform a set of conditional Monte Carlo simulations (e.g., Yang et al. 2019). Conditional means that the values of the known undrained shear strengths will not change from simulation to simulation, but fixed. From a set of conditional Monte Carlo simulations, we can estimate the probability of failure, which is 6% in this case.

If we perform a second CPT, we will know two columns of undrained shear strengths as shown in Figure 5. We can perform another set of conditional Monte Carlo simulations, and estimate the probability of failure, which is 0.7%. If we perform a third CPT, the estimated probability of failure is essentially close to zero as shown in Figure 6. If we keep doing test, eventually we will know the undrained shear strength distribution everywhere. And it turns out that the slope is actually safe. we can see based on three CPTs, we can reach a very reliable prediction of the slope stability. Please note that this conclusion only holds for this particular slope. For a different slope, the slope may fail, and we may need different number of CPTs to find it out. But the same methodology can be used to study how much tests do we need to draw a reliable prediction.

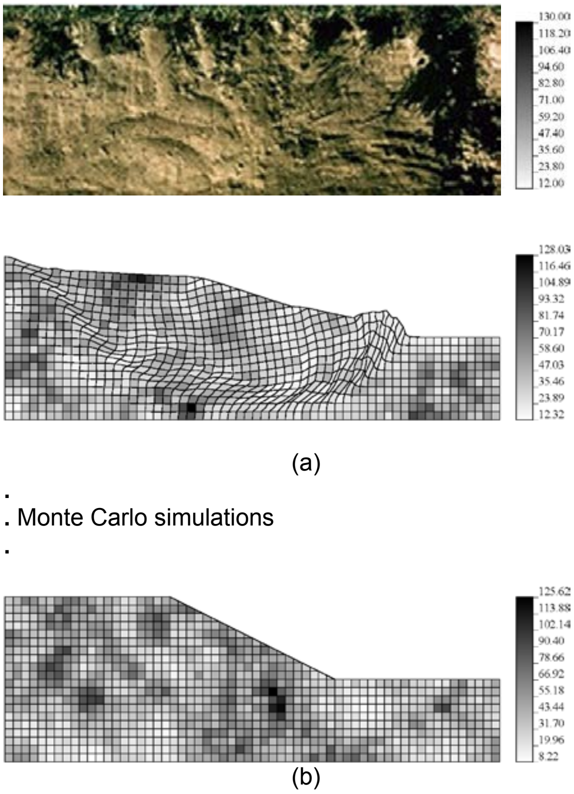

Figure 3. Unconditional Monte Carlo simulations based on the statistics in the literature

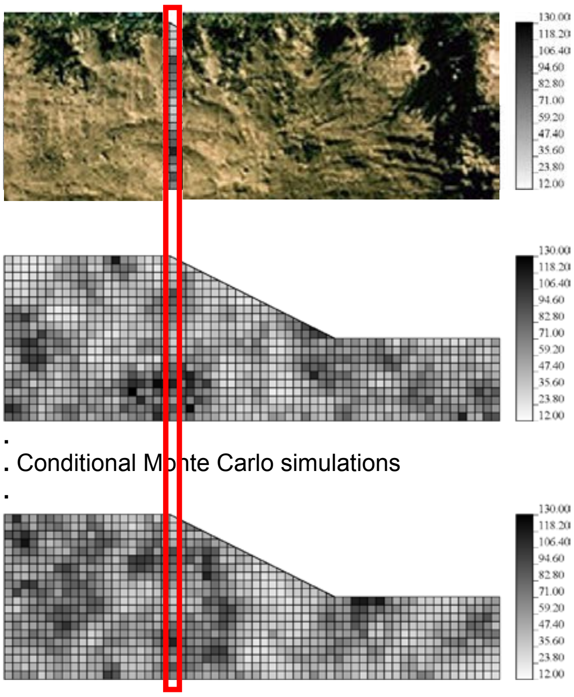

Figure 4. Conditional Monte Carlo simulations based on the statistics and one CPT test

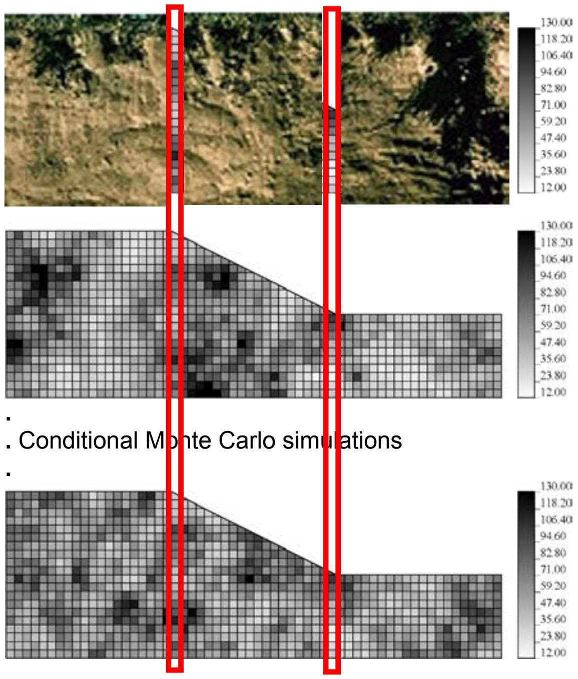

Figure 5. Conditional Monte Carlo simulations based on the statistics and two CPT tests

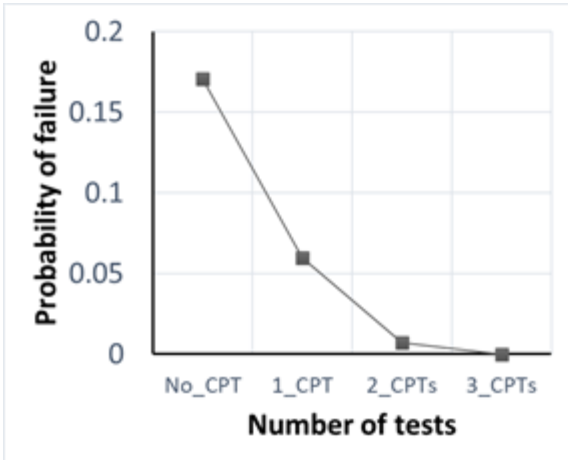

Figure 6. Probability of failure verse number of CPT tests

Generally speaking the approaches that we use to assess the safety of structures can be summarised in Table 2. It can be seen from Table 2 that probabilistic approaches can use test data directly, thus can reduce the uncertainty in our prediction. Recently Machine learning approaches become popular in geotechnical engineering. It is noted however, that Machine learning approaches use test data only, while probabilistic approaches can be combined with mechanical models (e.g., finite element model).

Table 2: Risk assessment of slope failure

Approach	Input data
Deterministic	Characteristic values
Unconditional probabilistic analysis	Mean, standard deviation, spatial correlation structure
Conditional probabilistic analysis	Mean, standard deviation, spatial correlation structure, test data
Machine learning	Test data

5 HOW CAN WE USE MULTIPLE SOURCES OF INFORMATION?

Site characterisation is a fundamental step of collecting geotechnical information for proper design, construction and long term performance of all types of civil and geotechnical structures. Current site investigation practice in geotechnical engineering often involves two steps: 1) a geophysical site investigation or a desk-top study of existing geophysical data; and 2) a dedicated *in situ* geotechnical investigation including several boreholes and laboratory testing at the intended location of the future structure. Each type of investigation explores a specific volume of the subsoil and has different degrees of uncertainty. Geophysics provides a wide variety of tools which can help to identify the subsoil stratigraphy. Data are obtained on a two- (or three-) dimensional section of the ground, often with low resolution. On the other hand, *in situ* tests are performed along a one-dimensional line with depth and laboratory tests performed at point locations within the ground. These tests provide direct measurements of physical and mechanical properties, but cover only a small volume of the subsoil. In current engineering practice, the integration of geophysical data with geotechnical data is done manually, often by visual inspection based largely on engineering judgement and experience, which does not only introduce additional human error, but also causes loss of information. For example, the geophysical data is used to find the desirable locations of geotechnical tests, but often ignored in deriving geotechnical properties.

A more scientific integration of geophysical data with geotechnical data can be done based on Bayes' theorem. Figure 7 shows a one dimensional cone tip resistance profile obtained by a CPT test. On top of that, it also shows a two dimensional shear wave velocity profile obtained by multi channel surface wave (MASW) test. If we can use CPT and MASW test results together, it will help to reduce the uncertainties in our model.

Huang et al. (2018) developed a Bayesian updating approach, which can combine multiple sources of information automatically. If we use the CPT test result only, we can use random field theory to estimate the mean, standard deviation and spatial correlation length using the CPT results and then use conditional random fields to guess the possible soil profiles. Figure 8 shows the soil profile based on conditional random field and the CPT test only. It can be seen from Figure 8 that at the place where we have CPT test, the mean cone tip resistance is the same as we tested by CPT and the standard deviation of cone tip resistance is low. However, at the places where we don't have any test, the mean cone tip resistance is similar to and informed only by the CTP test results, and the standard deviation of cone tip resistance is large. This means that we have less confidence on the profile of cone tip resistance based on one CPT only.

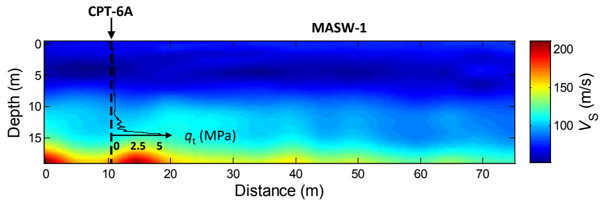

Figure 7. Geotechnical CPT test and geophysical multi channel surface wave test

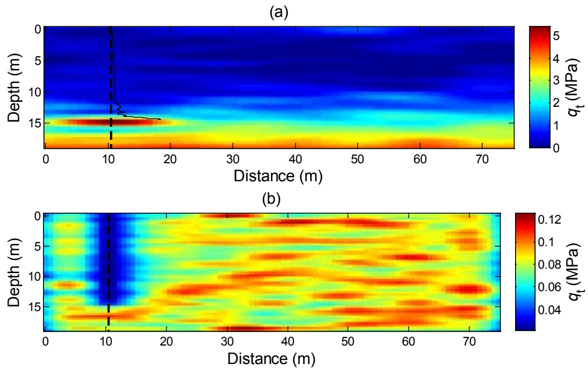

Figure 8. Soil profile based on one CPT test, a) mean cone tip resistance, b) standard deviation of cone tip resistance.

If we use the CPT and MASW test results together, we can have a higher confidence on the cone tip resistance profile. Figure 9 shows the soil profile based on Bayesian updating and both CPT and MASW test results. It can be seen from Figure 9 that at the place where we have CPT test, the mean cone tip resistance is still the same as we tested by CPT and the standard deviation of cone tip resistance is low. However, at the places where we don't have CPT test, the mean cone tip resistance is informed by the MASW results, and the standard deviation of cone tip resistance is reduced.

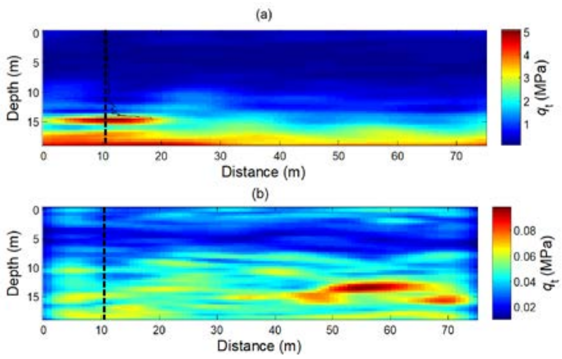

Figure 9. Soil profile based on CPT and MASW test results, a) mean cone tip resistance, b) standard deviation of cone tip resistance.

6 HOW DO WE USE MONITORING DATA TO PREDICT FUTURE PERFORMANCE?

In design stage, we predict the safety level of the structures based in lab and field test results. After construction, the behaviour of the structure most of the time is different from our prediction. For important projects such as large dams and bridges, we usually install some sensors or devices to monitor the behaviour of the structure. In this case, we need to use the monitoring data to reassess the safety level of the structure. This is called back analysis.

Another example where back analysis becomes important is embankment built on soft soils. Due to variations in a soft soil profile and its properties, the actually settlement behaviour is usually different from the performance anticipated in design. If the decision on the settlement behaviour (i.e., quicker or slower than expected) can be made sooner, then a smaller amount of surcharge is required and road construction can start earlier, which leads to significant financial benefits.

Table 3 lists some of the most commonly used back analysis methods. The simplest back analysis is manual calibration where engineers try to minimise the different between observation D and model output $g(\theta)$. Least square method is probably the most widely used method for back analysis because it always provides a solution no matter how complicated the problem is. Least square is actually a special case of the Maximum Likelihood method where measurement error is normally distributed with zero mean. In contrast to the Maximum Likelihood method, Maximum A Posterior method not only consider the likelihood, but also the prior distribution of the parameters. Bayesian back analysis is the most generic type of back analysis. Instead of solving for the set of parameters that maximum posterior probability density as in the Maximum A Posterior method, Bayesian updating samples the whole posterior probability density function. The most commonly used method for this purpose is the Markov chain Monte Carlo method.

Table 3: Commonly used methods for back analysis

Method	Formula
Manual calibration	$\text{Min } D-g(\theta)$
Least square	$\text{Min}\left(D-g(\theta)\right)^2$
Maximum Likelihood	$\text{Max } L\left(\frac{D-g(\theta)-\mu_e}{\sigma_e}\right)$
Maximum A Posterior	$\text{Max } L\left(\frac{D-g(\theta)-\mu_e}{\sigma_e}\right)P(\theta)$
Bayesian updating	$P(\theta\|D)\propto L\left(\frac{D-g(\theta)-\mu_e}{\sigma_e}\right)P(\theta)$

Bayesian updating can use prior knowledge (e.g., lab and field test results) and monitoring data in a rigorous framework. It can also consider measurement errors and provide confidence interval on the predictions. A proof-of-concept application of Bayesian updating for embankment settlement prediction can be found in Kelly et al. (2015). Zheng et al. (2018) used Bayesian updating with laboratory data, field test data, and monitoring data to yield accurate predictions during the construction and consolidation periods for the test embankment built at Ballina, New South Wales, Australia. The results show that surface settlement can be well predicted using 116 days of observed settlements, while the pore pressure can be predicted using 292 days of pore pressure measurements. The predictions are shown to converge to the field measurements, regardless of some assumptions about the measurement errors. It is also demonstrated that incorporating more monitoring data into the Bayesian updating process enables more accurate predictions.

7 CONCLUSIONS

Deterministic factor of safety approach has been used in geotechnical engineering for many decades. Although probabilistic approaches can provide more information for risk assessment, there are still quite some resistances to the use of probabilistic approaches. The most common excuse for not using probabilistic approach is lack of data. This paper has shown that Bayesian approaches can be used even when we have only one test result. Bayesian approaches are also useful for combining different types of test results and back analysis based on monitoring data. It is anticipated that probabilistic approaches will become more and more popular in geotechnical engineering to supplement, not replace the factor of safety approach.

8 ACKNOWLEDGEMENTS

The research supports by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (project DP190101592) and Linkage funding scheme (project LP200100367) are acknowledged.

REFERENCES

Huang, et al. (2016). "Updating reliability of single piles and pile groups by load tests." Computers and Geotechnics 73: 221-230.

Huang, et al. (2018). "Probabilistic characterization of two-dimensional soil profile by integrating cone penetration test (CPT) with multi-channel analysis of surface wave (MASW) data." Canadian Geotechnical Journal 55(8): 1168-1181.

Kay (1978). "Safety Factor Evaluation for Single Piles in Sand." Journal of the Geotechnical Engineering Division-Asce 104(1): 148-149.

Kelly, et al. (2015). "Bayesian updating for one-dimensional consolidation measurements." Canadian Geotechnical Journal 52(9): 1318-1330.

Michalowski (2002). "Stability Charts for Uniform Slopes." Journal of Geotechnical and Geoenvironmental Engineering 128(4): 351-355.

Yang, et al. (2019). "Importance of soil property sampling location in slope stability assessment." Canadian Geotechnical Journal 56(3): 335-346.

Zheng, et al. (2018). "Embankment prediction using testing data and monitored behaviour: A Bayesian updating approach." Computers and Geotechnics 93: 150-162.