*Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)*
*Edited by* Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly
 *Published by* Research Publishing, Singapore.
doi: 10.3850/978-981-18-5182-7_00-004-cd

ISGSR 2022

# Bridging the Gap between Research and Practice

Stephen Buttling[1]

[1]National Geotechnical Consultants, Brisbane, Australia.
E-mail: stephen.buttling@ngconsult.com.au

**Abstract:** The gap between research in geotechnical safety and risk, and practice, is not just a gap, it is a yawning chasm. This keynote lecture will present a personal view, based on 50 years in the geotechnical industry, and call on first hand and second hand evidence gathered over those years, to show how wide and deep this chasm is, and will call for open discussion on what can and must be done to bridge the chasm, and make the practice of statistics part of the everyday work of all geotechnical engineers, so that proper assessments of uncertainty, risk, and reliability can also become a normal part of geotechnical practice. This is necessary for many reasons, but first and foremost is the misconception that deterministic analysis brings, that we are dealing with known fixed quantities.

Keywords: risk; reliability; education; design.

## 1 Introduction

I would first like to thank the organizers of this Symposium for the honor of inviting me to deliver this keynote lecture. I feel well out of my depth, and rather out of place here, since I claim no special skills in statistics, risk assessment or reliability analysis applied to geotechnical engineering, and so will not be able to talk about my latest research and discoveries or developments. However, I have become very interested in these topics over recent times in my career of over 50 years, and so would like to share with you some of my opinions that relate particularly to the teaching of statistics to engineers. These are based on my personal experiences and observations, so this will be a bit of a journey through my career, for which I hope you will forgive me.

## 2 History

I will start by very briefly going back to 1966, my last year at high school, and the last formal training in statistics which I can remember. It was something to do with Permutations and Combinations. I do not recall any introduction of statistics in my undergraduate studies at Imperial College. During the first 40 years of my career as a geotechnical engineer in the UK, South East Asia, and Australia I do not remember any specific involvement in statistics or statistical analysis, until about 9 years ago, when two totally unrelated projects brought these topics to my attention.

The first involved the peer review of a report on slope stability at an open pit mine. The young engineer author had included a section on Probability of Failure of the slope being analysed. This was the first time I had come across the concept, so I was curious, and asked him how he had carried out the probabilistic analysis. I was, therefore, somewhat horrified to find that he had used a grid and radius method to determine the minimum factor of safety, and that his probability of failure had been computed as the number of surfaces giving a Factor of Safety (FoS) < 1 as a function of the total number of surfaces analysed. When I explained that the Probability of Failure could be reduced simply by extending the grid to increase the number of circles analysed he realized the error of his ways, but he had opened up a whole new area of study for me! I then started to investigate how to carry out probabilistic analyses, and found that the popular slope stability program, SLOPE/W from Geostudio, which I used, had built in functionality to do this. At the same time I learned that it was risky to try to carry out probabilistic analyses without thinking carefully about what I was trying to do. For example, the software had built in functions for normal and log-normal distributions, but also for uniform and triangular distributions, and strange things called generalised spline functions. It also allowed normal distributions, for example, to be truncated, which would tend to distort the output. I found it hard to relate these modified distributions to the properties of natural materials.

The second topic involved a paper published in the Australian Geomechanics Journal, by Pollock et al (2011), with the title "Linking Limit Equilibrium Analysis and Landslide Risk Assessment". A very worthy topic, and one that I found interesting, but the paper contained a number of crucial mistakes. One was in the approach to Landslide Risk Assessment, which required an Annual Probability of Failure to be identified. The lack of a proper understanding of statistics allowed the authors to avoid the need for a time dependent variable in the analysis, and instead consider that a non-time dependent Probability of Failure could be equated to a failure during the design life, even when the former was simply based on the strength parameters being random variables. In the paper the authors correctly state that, given an annual probability of failure $= P$, and a design life of $y$, then the Probability of Failure during the design life will be $P_f = 1 - (1 - P)^y$. However, they then went on to suggest that, if they called their non-time dependent Probability of Failure the Probability of Failure during the design life, $P_f$ then they could determine the annual probability of failure from $P = 1 - \sqrt[y]{1 - P_f}$, which is clearly not the case. Another problem was the failure to understand that, without any spatial variability, which had not been allowed for, only one random variable was being assigned to the strength parameters for the whole 2-dirmensional slice. It was noted that, within SLOPE/W which had been used for the analyses, the option to include spatial variability was included, but it was not clear how important this would be to an analysis..

### 2.1 Normal and lognormal distributions

Somewhere along the way I had come to understand something about normal distributions, and accepted that natural random variables might well be normally distributed. On the other hand, I had difficulty with the notion that natural random variables might have a triangular or generalized spline distribution, although these might be implied by data with a small sample size. However, some experience on a project involving mixing of cement with the waste product of the alumina extraction process, known as red mud, has had a lasting impact. This was because I had always felt that a log-normal distribution was also likely to be a mathematical contrivance which was unlikely to fit natural behaviour. In this project a total of twelve blocks of material were mixed in slightly different ways, and tested by pushing a Wing Cone, a modified CPT, through the nominal 4 m depth of treated mud. The individual vertical profiles are shown in Figure 1.

.

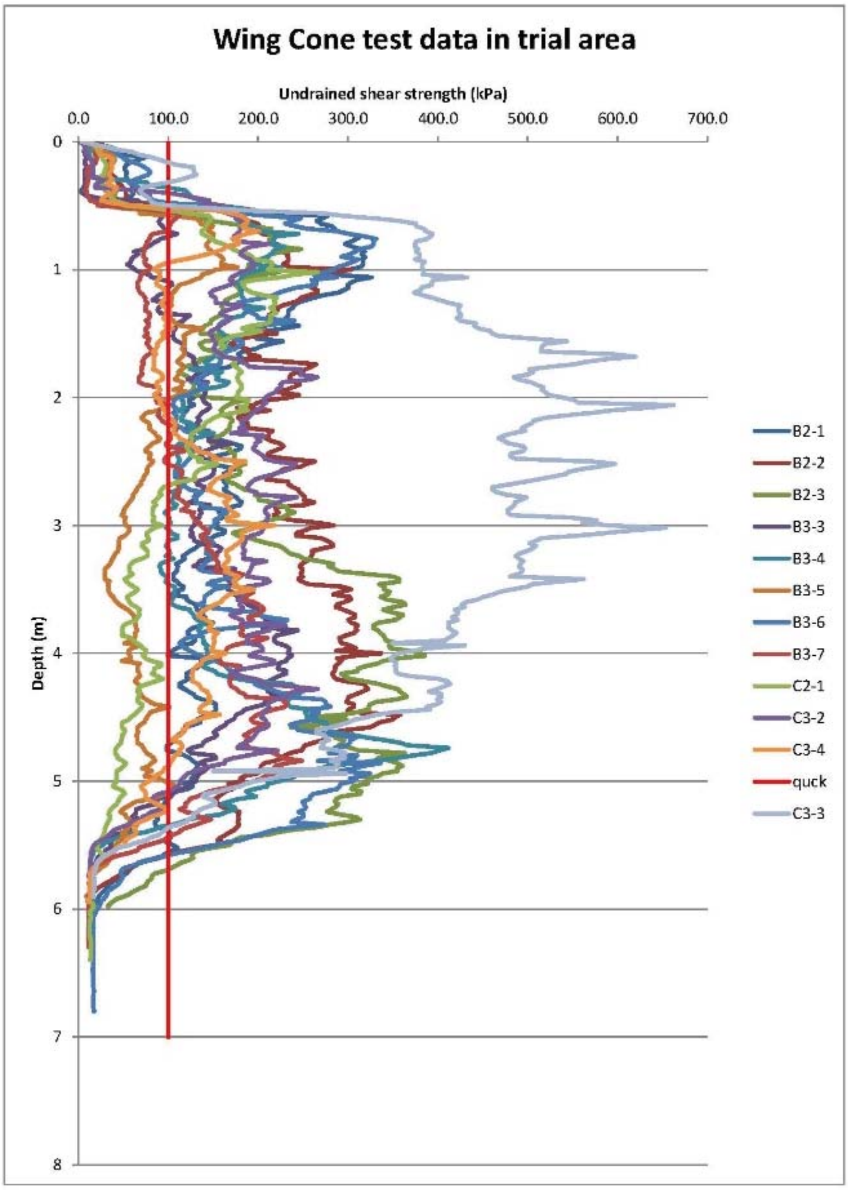


**Figure 1.** Vertical profiles of wing cone resistance

These show the sort of scatter which might be expected from a random variable, with no apparent trends. The same data grouped into bands and plotted as a Probability Distribution Function is shown in Figure 2, and this already suggests that the distribution might be lognormal. This was also supported by the fact that the target strength was about 100 kPa, but values of as much as over 600 kPa were recorded while clearly negative values were not possible.

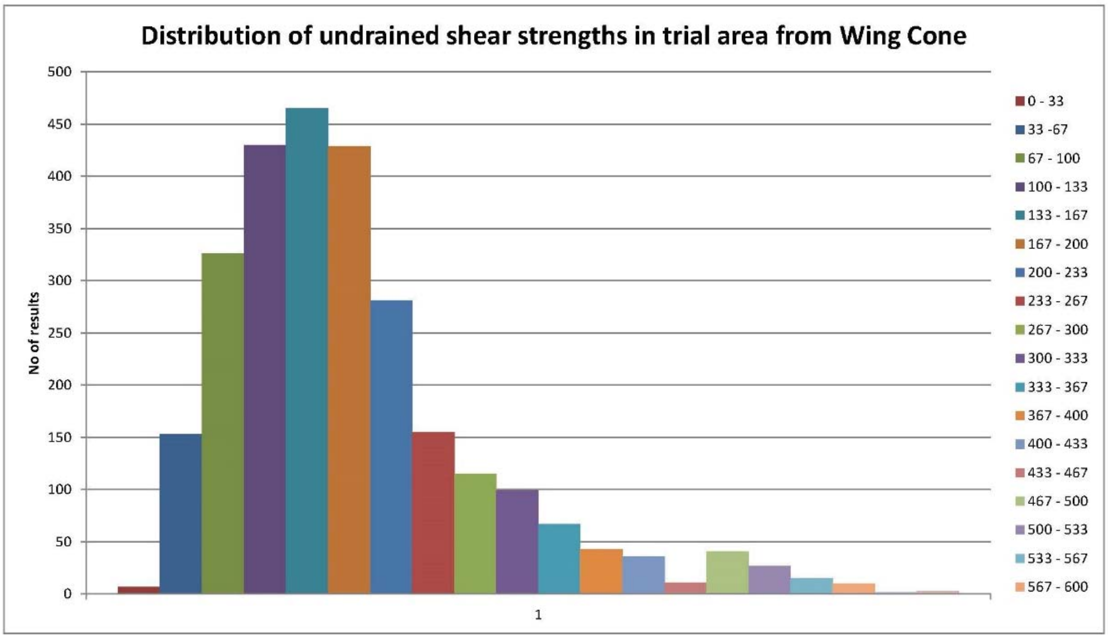


**Figure 1.** Probability Distribution Function for wing cone resistance data

Taking note of this the next chart in Figure 3 shows each of the twelve profiles with the data sorted and plotted as a CDF, together with a theoretical log-normal curve for each dataset.

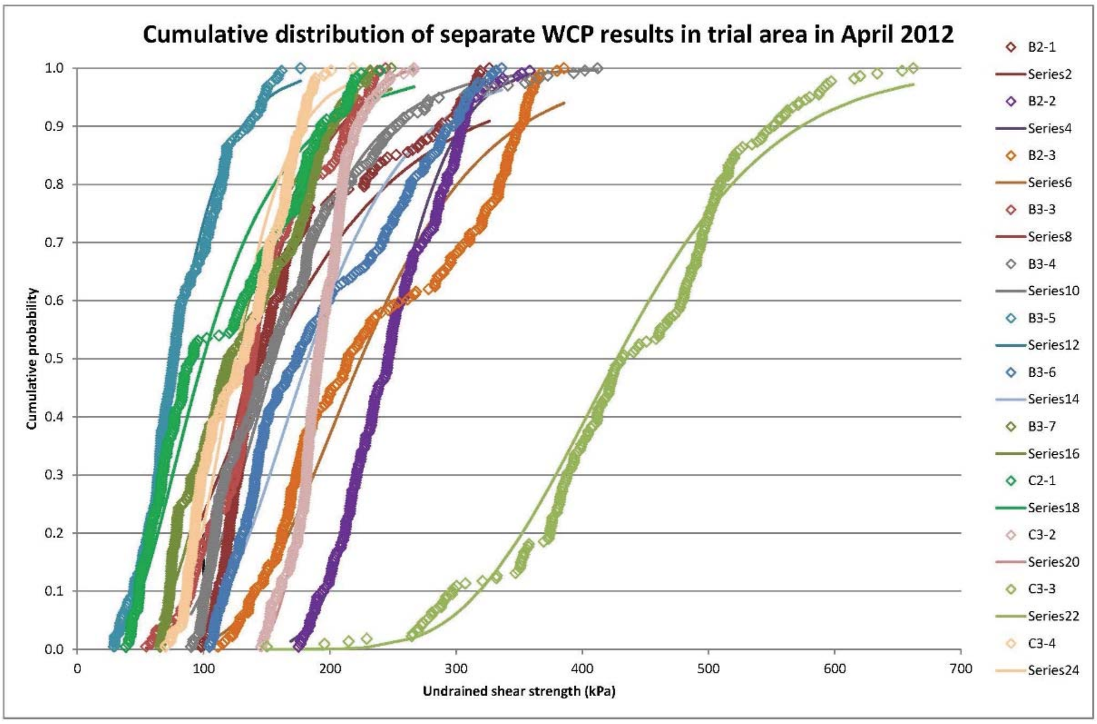


**Figure 3.** CDF plots of the twelve individual wing cone resistance profiles

These CDF plots may be considered to suggest a log-normal profile, but the scatter is very evident. It was therefore a great surprise to me to see what happened when all the data were combined into one CDF, as shown in Figure 4. It should be borne in mind that each of the twelve profiles was for a different mix, yet the combined data is so close to the theoretical log-normal CDF that there can be little doubt that a log-normal distribution represented an excellent fit to the real data.

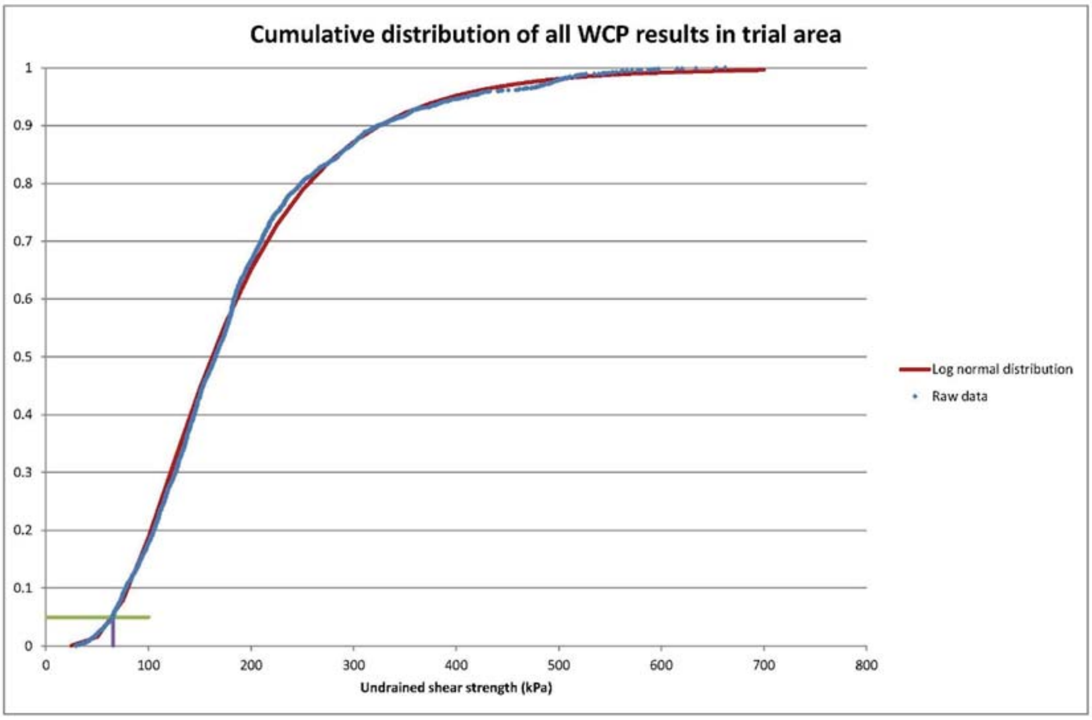


**Figure 4.** CDF of data from all twelve wing cone resistance profiles combined

## 2.2 Slope stability analysis

At about the same time I was involved in another project which included landslide potential. The consultant whose report I was reviewing had opted to carry out probabilistic analyses to assess landslide risk, using the guidelines for Australia prepared by the Australian Geomechanics Society (2007). They had also used the functionality available within SLOPE/W, but appeared to have used triangular probability distributions for the apparent cohesion and the angle of internal friction, which were quoted with minimum, maximum and mode values. I consider that such distributions are very unlikely to represent the true distributions of the random variables, and will therefore lead to distorted results in the distribution of factors of safety. In addition, the groundwater conditions were modelled using the pore pressure ratio, $r_u$, which expresses the pore water pressure as a fraction of the vertical total stress at the failure surface (bottom of the slice). This is sometimes used for modelling the stability of recent fill embankments, with a value of $r_u = 0.15$, but must be used with great caution. In this case a triangular distribution was again assumed, with 0.15 as the minimum value and 0.65 as the maximum value. For a soil with a weight density of 20 kN/m$^3$ this means that the phreatic surface would be above the ground surface, which was not physically possible with the existing slope.

The outcome was that the Reliability Index for one analysed slope was found to be 0.512. As with the paper by Pollock et al (2011), this report assumed a design life of 50 years, so the Probability of Failure resulting from the Reliability Index was 30.7% and this was assumed to be the Probability of Failure during the Design Life. Using the same formula as proposed by Pollock et al, they came up with an Annual Probability of Failure of 0.0073, which has no meaning since it does not include consideration of any time dependent data such as a ground water table with a known Annual Exceedance Probability. Thus, in the space of about 2 years, I came across direct evidence of different engineers trying to carry out probabilistic analysis of slope stability, but without the necessary understanding of the basic statistical theory required.

That situation then led me down the path of trying to learn more about probabilistic analysis, and how it should be carried out. There was a lot of reading, and spatial variability was clearly a major issue. It was therefore somewhat of a surprise that SLOPE/W provided the option, but it was almost greyed out and was not an essential part of the analysis, while its major competitor, SLIDE from Rocscience, also provided probabilistic functionality but without any spatial variability until 2019. I thought about the probabilistic analysis of a slope with a single homogeneous material, and realized that I was only analyzing my uncertainty as to what deterministic parameter I should be using, which is not really what I was expecting from a probabilistic analysis. It also occurred to me that if there was a 10% Probability of Failure of a 2D slice, then it must have a finite thickness and the slices on either side would have a 90% probability of not failing, meaning that the subject slice would not actually fail. This would require several adjacent slices to fail simultaneously, reducing the probability of failure to a much lower number, but still basically irrelevant as it is extremely unlikely that a whole 2D section will have uniform properties.

This then led me to analyse a simple slope as a uniform homogeneous material, but then to break it into individual "squares" or "cubes". Initially these were of 4 m side, and then of 2 m side. The geometry had to be input by hand, based around the critical circle from the homogeneous analysis, but the soil properties were cloned to make them identical in every square. Running the probabilistic analysis then randomly applied a different strength to each square. Because of the averaging that this introduced, the Probability of Failure dropped to 10% of its initial value for the 4 m squares, and to 10% of that value for 2 m squares. This was visualized as shown in Figure 5, but it should be noted that the critical surface is unlikely to be circular when a failure path is seeking the line of least resistance through the weakest areas.

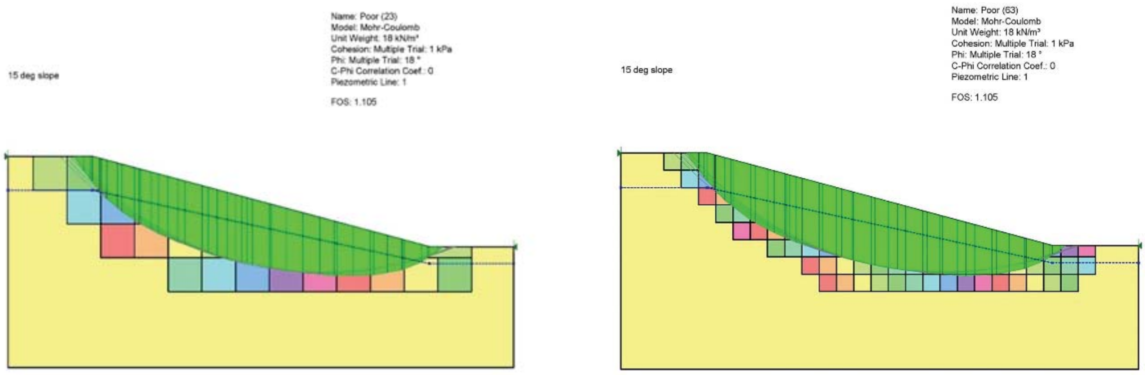


**Figure 5.** Probabilistic analysis for uniform slope with grids of squares with random properties

**Table 1.** Results of analyses with different sizes of zone with uniform properties

| Model | Probability of failure |
|---|---|
| Uniform layer | 29.35% |
| 4 m squares | 2.77% |
| 2 m squares | 0.28% |
| Vary every 4 m | 1.51% |
| Vary every 2 m | 0.14% |
| Vary every slice | 0.1% |

My increasing interest in statistics led to the organization of a workshop in Brisbane at the end of 2016, at which Vaughan Griffiths presented on Quantitative Risk Assessment in Geotechnical Engineering on one day, and we held a session with several local contributions on the second day. I was keen to learn about and be able to use random fields, so bought a copy of Fenton & Griffiths (2008) and have been trying to work my way through it.

### 2.3 Earth retaining structures

At around this time I was one of a team of about 20 engineers of who had volunteered through our professional society, Engineers Australia, to try to write a good practice guide for the design and construction of embedded retaining walls, which were not adequately covered by our Australian Standard published in 2002. The plan was not to write a standard, and not to reinvent the wheel, largely adapting the principles of a UK publication, CIRIA Report C580 and subsequently C760, to Australian conditions. We compiled an agreed Table of Contents, and assigned tasks to members, but were thwarted when Engineers Australia made our flag carrier redundant. Nothing happened until early 2017 when I was here in Newcastle, speaking with Professor Huang Jinsong, and discussed the earth retaining structures issue with another professor familiar with Australian Standards. The result was that I submitted a proposal to completely revise the 2002 standard, and this was approved in September 2017, just as I went to Seoul for the 19th ICSMGE. There I heard much about reliability and probabilistic methods, and had the opportunity to speak to Peter Day and Farouk Nadim, as well as Vaughan Griffiths, and decided that our revision to the Australian Standard should pave the way for reliability-based design which I felt must come within the lifetime of the revised standard. I therefore attended my first Technical Committee meeting of Standards Australia in June 2018, and met a structural engineer who had been involved with writing the 2002 version. It soon came out that he was fully aware of reliability requirements, had done a lot of work with the Australian Building Control Board which produces the National Construction Code (NCC), and knew that the NCC was stipulating minimum

Reliability Index values for structures and buildings. It seemed as if everything might fall into place, but I also spoke to a senior geotechnical engineer colleague, educated in Europe with a PhD and with about 15 years post-graduate experience, and he was very reluctant to embrace reliability-based design as it was completely outside of his experience.

### 2.4 Undergraduate teaching

It was also about this time that one of my daughters was studying civil engineering at the Queensland University of Technology. The course included a module on statistics and, as seems to be the norm, this was taught by the Mathematics Department. This brings with it the problem that, at any time, they can be teaching students from many different disciplines, which means that the theory and the examples used are not necessarily well suited to engineering students. The anecdote which I remember clearly was to do with Conditional Probability. The example in the assignment talked about a disease and a test, requesting the probability of having the disease given a positive result, or of not having the disease given a negative result. The follow up question was "without changing the test, how can the accuracy be improved?" Since the accuracy cannot be changed without changing the test, one has to assume that the question was supposed to be about the reliability, and hence a second test was required. We are still left with the problem that it was very easy for engineering students to consider that this topic, and therefore most of the statistics course, was irrelevant so just a unit to be passed and forgotten about. It is noteworthy that exactly the same topic in Ang & Tang is introduced with respect to a concrete batching plant, where the relevance to engineering is obvious.

I did discuss this aspect with a very bright Mathematics post-graduate here at Newcastle University while on a visit, and he tried to assure me that bright students would make the link for themselves, from a non-engineering application to an engineering application for any given topic. I fear that, with the opening up of university education to a very much wider group in the last 50 years, this is no longer the case. Only a very few very bright students will be able to make those links, unless it is explained in some detail that they need to do so and how to go about it. I have talked to many practicing geotechnical engineers, and most do not see any application for statistics in their normal working days, a frequent reason being given that they do not have enough data. This is presumed to be due to the fact that the traditional teaching approach links statistics to dealing with large amounts of data, and does not pass on the processes to be used where there is little or no data. Similarly, most standard statistics courses probably pay little regard to dealing with extreme events, which are frequently of importance to civil engineers.

### 2.5 FORM and Excel

Meanwhile, having been introduced to FORM by Vaughan Griffiths, I came across the work of Low Bak Kong and Wilson Tang, promoting FORM and the use of Excel spreadsheets and the Solver routine to do the computation. I worked my way through many of the examples for different applications, not always successfully, but got a reasonable understanding of the process. I was amazed to discover that a spreadsheet could be used to carry out a slope stability analysis using the method of slices, and determine the critical non-circular slip surface. This still seems to be a very well-guarded secret.

### 2.6 Characteristic values

In early 2017 I was in discussion with Trevor Orr about selection of characteristic values. My interest in this had been sparked by the survey data published in Andrew Bond and Andrew Harris's book Decoding Eurocode 7. As many of you are probably aware the authors sent 5 sets of geotechnical data to about 100 practicing geotechnical engineers and engineering geologists, asking them to select characteristic values in accordance with EC7, which specifies "a cautious estimate of the value affecting the limit state". The first two data sets were for SPT N values and undrained shear strength of stiff London Clay, the second two were for vane shear tests and triaxial tests on softer Singapore Marine clay, and the last for SPT N values in the Thames Gravel in London. The results were quite shocking to me, and showed that selected characteristic values ranged by a factor of up to 3 or even 5 between lowest and highest. It must be remembered that the design value, $P_d$, of a parameter for a geotechnical design is given by.

$$P_d = \gamma P_k \tag{1}$$

where $P_k$ is the characteristic value of the parameter and $\gamma$ is the partial factor, which can be considered to vary between 0 and 1, and is typically defined to two decimal places. How can any sort of consistency be expected in a design process being carried out by different engineers if the characteristic values selected are so far apart, to be multiplied by a reducing factor which has such apparent precision?

In considerations in relation to the proposed revision of the Australian Standard for Earth Retaining Structures, AS 4678, it was noted that from the previous committee Hausmann et al (1996) refer to the characteristic value as "a cautious estimate of the mean". This was amplified to our new Technical Committee by another member from the previous committee, who was adamant that the intention of the current version had been that the characteristic value was an estimate of the mean, rather than the "cautious estimate" required by Eurocodes or AS 2159. The actual wording in the definitions is that "the characteristic value is a cautious estimate, that is, close to but not greater than, the mean value". Having given the matter a lot of thought, we have therefore decided, for the proposed revision to AS 4678, to adopt the inferior estimate of the mean as the characteristic value. This automatically takes account of the amount of data, and the scatter of the data, in estimating the mean, but it is hoped that it will lead to a much narrower range of selected values by different engineers using the same dataset, which means more consistency. It is then suggested that a range of other partial factors are applied to this characteristic value to achieve the design value, and to avoid any lack of conservatism by the use of the mean value. These could include partial factors on resistance, based on how the strength data has been obtained, i.e. reflecting data quality. These factors would be closer to unity for data based on high quality laboratory tests, smaller for quality in situ tests, smaller again for tests such as SPT, and smaller again for data looked up in a publication. Similarly, a partial factor can be applied depending on the expected size of the failure surface. If this is significant such that the mean parameter value is representative of variation along the failure surface, then the partial factor can be close to unity, but if a small failure surface such as for an anchor fixed length could be affected by a local zone of lower strength, then a lower partial factor will be appropriate.

### 2.7 Prediction events

If I may I will just digress for a while onto a slightly different topic, but one which turned out to be very closely related as I thought about it more. This starts off with Pile Prediction Events, which have been going on in an organized if sporadic manner for over 20 years now as far as I personally have been involved. The first of this series was based in Orlando in 2002 (Fellenius et al, 2004) and, like all of the ones to do with piling, Bengt Fellenius was very closely involved. The format is similar for all. A set of geotechnical data is provided for a site, much as would be available for a design. The geometry of a pile is defined, and contributors are invited to predict and submit a load/settlement curve, and usually to suggest where the ultimate pile resistance might be placed on that curve. For the first event there were 33 contributors, but there were several practical problems with static load testing in both compression and tension, which reduced the benefit to be gained from the exercise. The second event took place associated with the 2nd International Conference on Site Characterization, ISC'2, in Porto in 2004 (ISC, 2004). Several pile types were involved, and Figure 6 shows the range of predictions for a bored pile, while Figure 7 shows the range for a Continuous Flight Auger (CFA) pile.

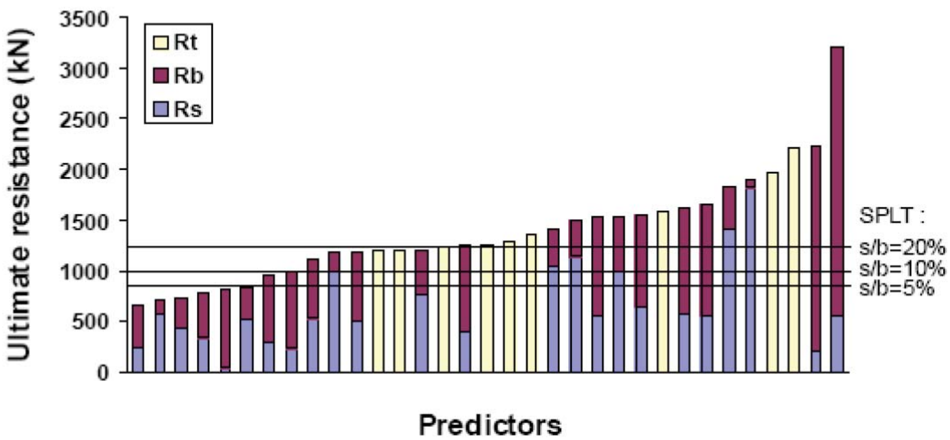


**Figure 6.** Ultimate resistance predictions for bored pile E9 at ISC2

It is immediately noticeable that over 50% of the predictions overestimated the bored pile capacity, while for the CFA pile this increased to about 67%. Since this is potentially unsafe it is a cause for some concern.

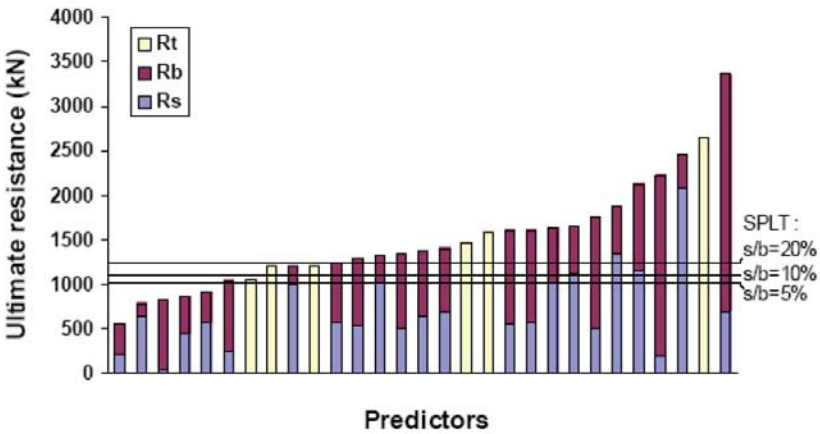


**Figure 7** Ultimate resistance predictions for CFA pile T1 at ISC2

The next event was in Edmonton in Canada in 2011 (Fellenius, 2013), and once again this was beset by static load test problems which meant that the ultimate capacity was never approached. Figure 8 shows the measured load settlement in green, with the blue lines representing predictions and the red dots on each prediction showing where they thought the ultimate capacity should be defined. The scatter in the predictions is huge, and even more worrying is the inability to agree on a definition of ultimate capacity, or even the settlement at which the ultimate capacity is reached. It can be seen that these range between about 6mm and 220 mm for a test on a 400 mm diameter CFA pile. In this and the similar charts to follow, there is no need to take in the detail, but the similarity in trends is important.

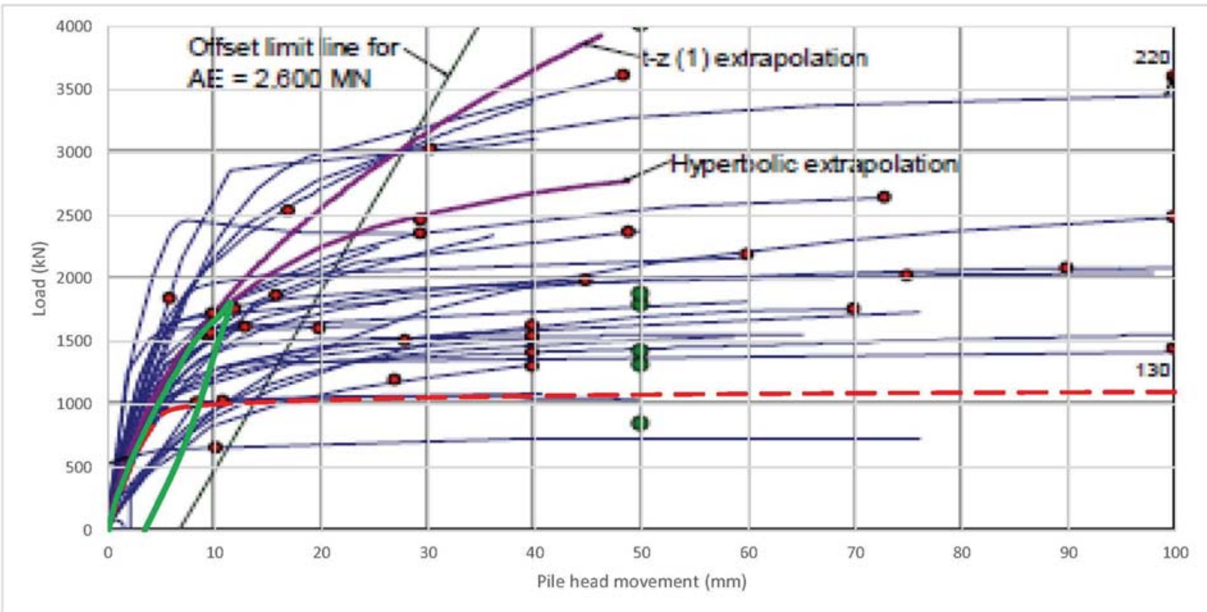


**Figure 8.** Predicted and measured load/settlement curves and selected ultimate pile capacities for Edmonton test on 400 mm CFA pile

There were then four events over the next 6 years, all organized in South America, three in Bolivia and one in Brazil (Fellenius and Terceros, 2014), (Tereceros and Felenius, 2015). The first was in 2013 in Bolivia, and the summarised predictions are shown in Figure 9. These have a very similar scatter in load/settlement relationship, but also again in selection of the deflection at which the ultimate resistance occurs. In Figure 9 the measured load/settlement is shown in red, and both this and Figure 8 show that a significant number of contributors overpredicted the pile capacity.

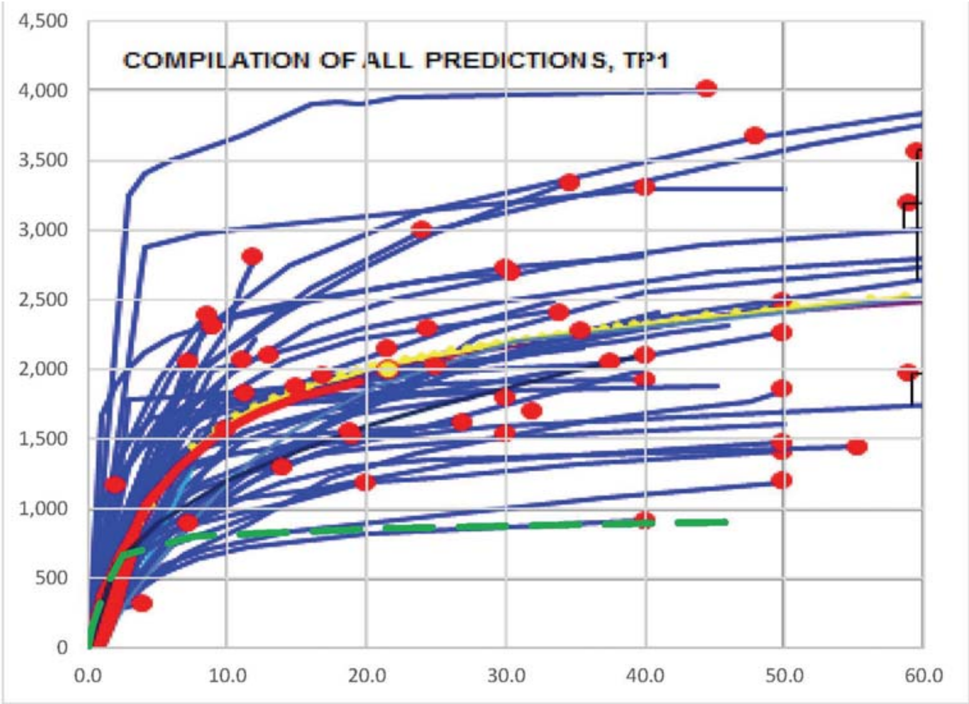


**Figure 2.** Predicted and measured load/settlement curves and ultimate pile capacities from Bolivia 2013 for a 400 mm pile bored under bentonite

In the second event in Bolivia in 2015, TP1 was a 600 mm diameter bored pile, with a total length of 16.4 m. This time there were 10 predictors, with the results shown in Figure 10.

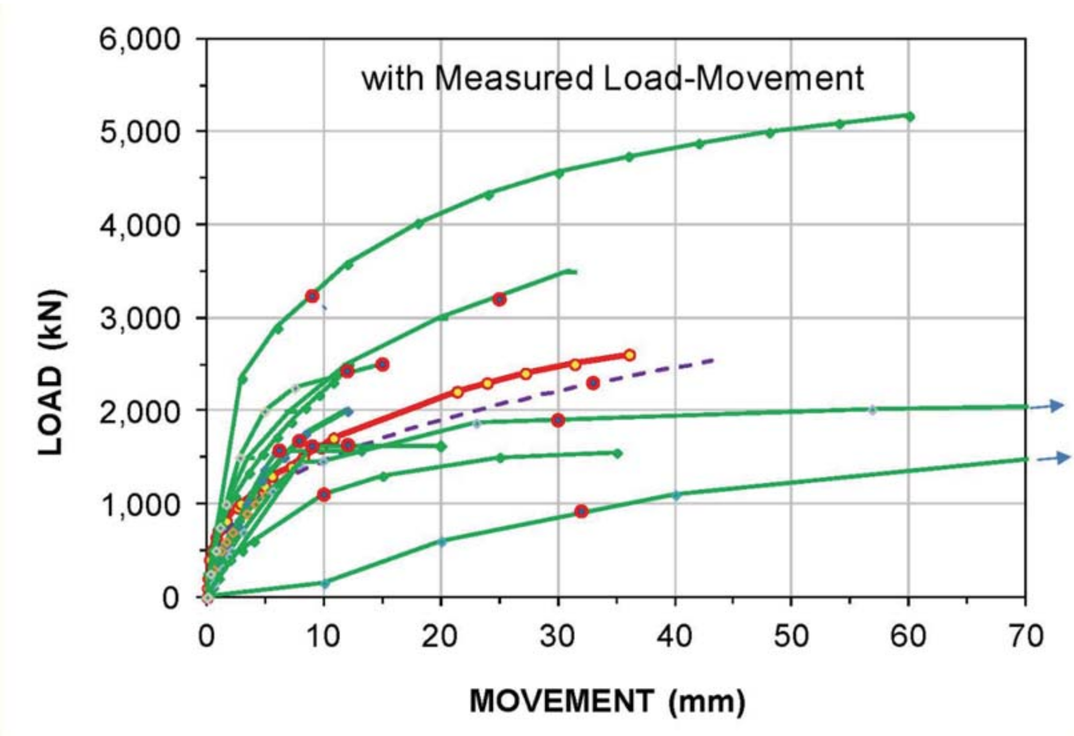


**Figure 10.** Predicted and measured load/settlement curves and ultimate pile capacities from Bolivia 2015 for a 600 mm bored pile

Despite the publication of the results from two years earlier, the range and scatter are very much the same, with some conservative and some non-conservative estimates.

Also in 2015 an event was held in Brazil, supported by the ISSMGE with an International Organizing Committee of experts, and a Scientific Committee of established international names in the field of piling.

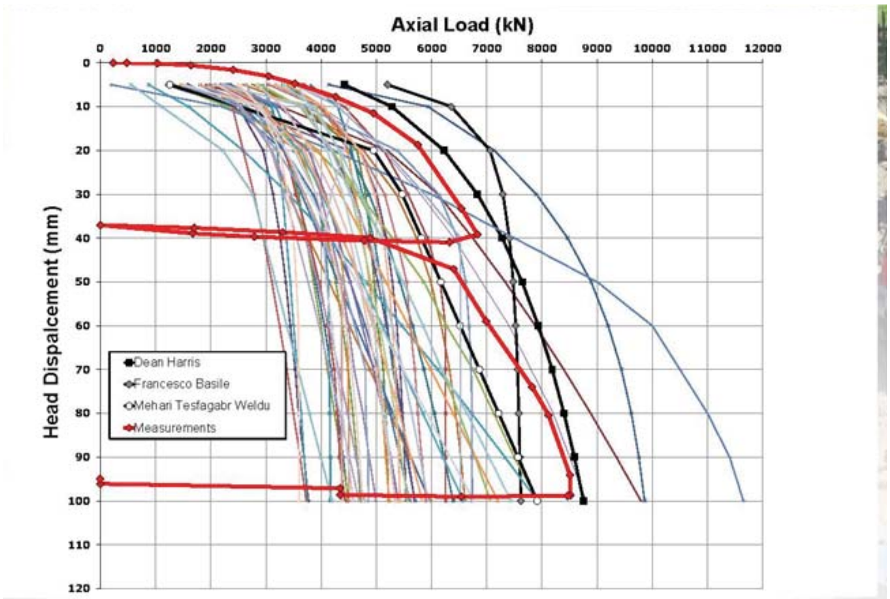


**Figure 11.** Predicted and measured load/settlement curves from Brazil 2015 for test on 1 m pile bored under bentonite

The big difference this time was that, although the scatter in results was similar to other events as seen in Figure 11, out of 72 contributions the best were close to the measured data, with 96% of predictors under, and only three over predicting performance. The same was not true when the event returned to Bolivia in 2017 (Fellenius et al, 2017), despite having had two previous events there. The results are shown in Figure 12 with similar scatter to Figure 9.

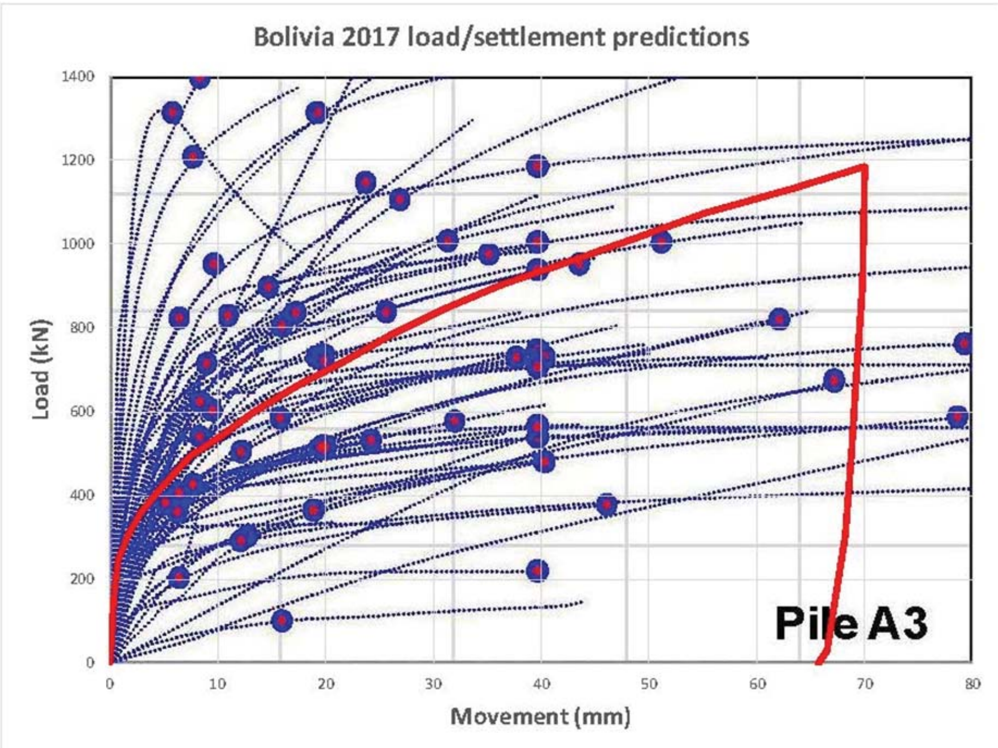


Figure 12. Predicted and measured load/settlement curves for Bolivia 2017 test on 620 mm diameter bored pile

These illustrations have related only to piling, but the same sort of exercise has also been attempted in Australia with regard to pad footings. The first was in 2017 and concerned a 1 m square footing on the soft marine clay at the Ballina Test Site, supervised by the University of Newcastle. Detailed high quality soil investigation data was made available, including both laboratory and in situ tests, and details on the proposed rate of loading. There were 50 predictors who responded to the invitation. Figure 13 shows the excavation for the footing and the completed footing before testing.

**Figure 13.** Careful excavation for 1 m square pad footing on soft marine clay and the cast footing prior to test

The results are compiled in Figure 14, on which the measured ultimate capacity of 205 kN is also marked.

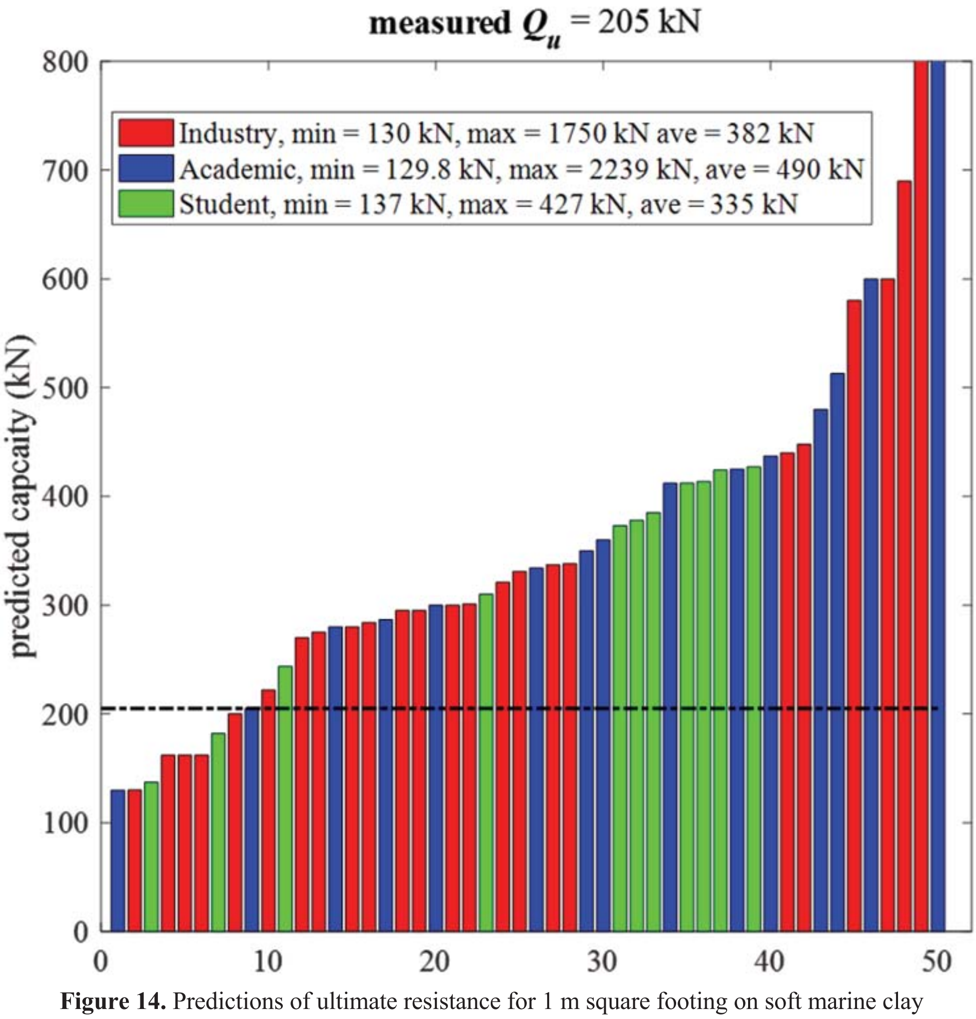


**Figure 14.** Predictions of ultimate resistance for 1 m square footing on soft marine clay

What is of most concern to me is the number of predictions which exceeded the measured value, in some cases by factors of close to 10. In the end 82% of the submissions overestimated the ultimate resistance. Another similar exercise was carried out by the University of Western Australia in Perth, and therefore on sand. This time detailed CPT data was provided, with the information that the footing was to be loaded to 180 kN, with requests to predict the settlement which would be measured. The prepared footing is shown in Figure 15, and the results in Figure 16.

**Figure 15.** 1 m square footing on sand for settlement prediction

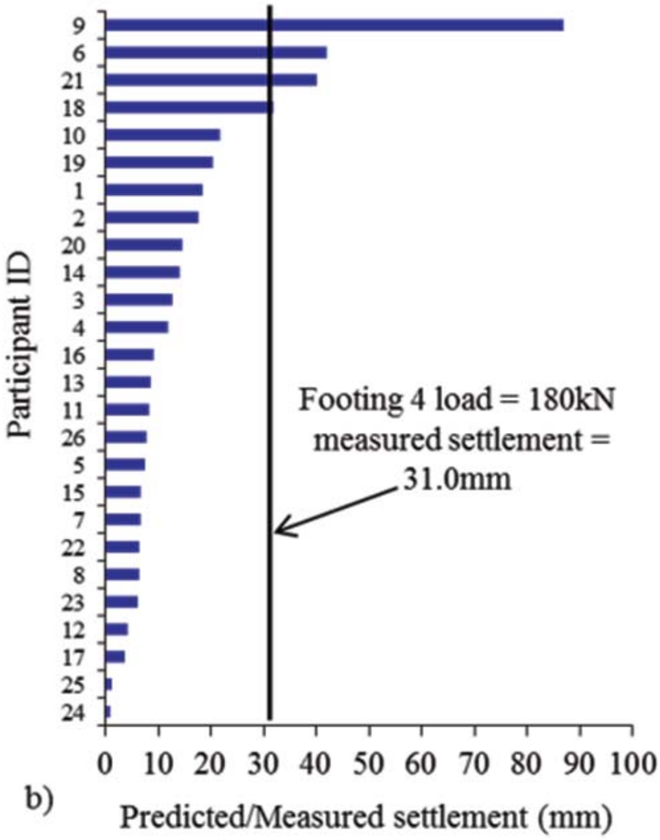


**Figure 16.** Submitted predictions of footing settlement

This time, out of 26 submissions, one was very close and 3 more over-predicted, which means were conservative. This means that 85% underpredicted the settlement, about half of those by more than 50%.

### 2.8 ISSMGE Survey

To bring this topic up to date we need to consider the work done for the 20th ICSMGE in Sydney last May. This provided the results to the question asked by ISSMGE members from the Asian Region in June 2018, "Are we overdesigning?" (Day & Briaud, 2022). Real soil data from the two National Geotechnical Experimentation Sites at Texas A&M University was provided, and 10 simple geotechnical problems on clay or sand were set. The survey was organized by ISSMGE Regions, held online in 2020 and 2021, and nearly 240 responses were received. Of these, one set of responses was received from ISSMGE Technical Committee TC304 "Engineering Practice of Risk Assessment and Management", which must be very familiar to many members of this audience, who submitted a number of reliability-based solutions. It was a very interesting exercise, and study of the keynote paper is recommended. Here I will only extract two illustrations, Figure 18 and Figure 19, for a bored pile in clay and a bored pile in sand respectively, since these can also be compared with the pile prediction figures shown above. Both piles were 0.76 m diameter and 10 m in length, and contributors were asked to predict the ultimate capacity of the piles, the base and shaft resistance, and the load/settlement curve up to 80% capacity.

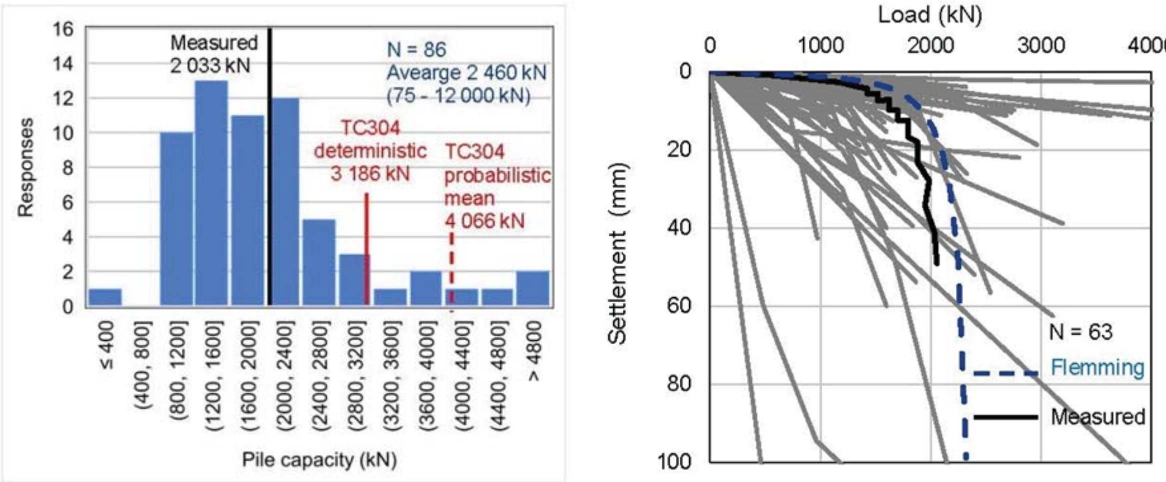


**Figure 18.** Predictions of capacity and load/settlement curves for bored pile in clay from ISSMGE survey

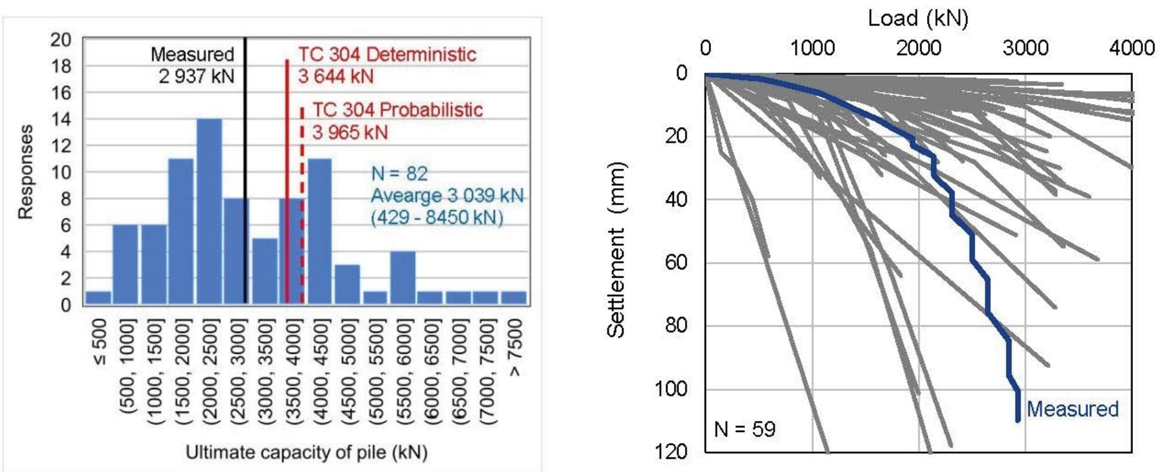


**Figure 19** Predictions of capacity and load/settlement curves for bored pile in sand from ISSMGE survey

It is quite clear that the calculations for load/settlement curves and for determination of ultimate pile resistance cover a similarly wide range to those carried out in the pile prediction symposia. The survey covered a total of ten problems, some to do with prediction, therefore concentrating on selection of data and choice of design model, while the remainder were to do with design so as to introduce the concept of safety. Also analysed in the survey were the occupations and age ranges of respondents. One of the main conclusions of the study was that, while the average of all contributions for each problem was within ± 30% of the measured or best estimate answer, the range was far too wide. Also that, as a result of that finding, the ISSMGE should pay as much attention to the application of existing knowledge as to the creation of new knowledge. It was not until I started drafting this talk that I realized another very significant fact. This is that, with over 550 contributions from all around the world, all relevant occupations and all age ranges, only the TS 304 submission acknowledged that, in dealing with soil data we are concerned with random variables. All of the other contributors treated the data deterministically, enshrining the concept that the design values are discrete numbers, and that the "answer" is also a discrete number, and I myself was guilty of the same charge. It has been realized that this is wrong in concept, and that the data should have been used to estimate the range of probable values, then carry out analyses across that range and determine not "an answer", but a probability distribution function of the answer. Especially bearing in mind that, in the cases where there was a measured result to compare with, then similar tests at an adjacent location would almost certainly have produced a different result, there is a much better chance of a correlation between a measured result and a pdf of predicted results. It is a cause for concern that only TC 304 out of all the submissions in all of these events decided to use a probabilistic approach.

## 2.9 Characteristic values again

Just returning to the issue of characteristic values briefly, in the light of what I have just said about the practice of geotechnical engineers, I have looked again at the definitions in AS 2159:2009, AS 5100.3:2017 and BS EN 1990:2002. All of these refer to the characteristic value, or the "value … used in design", just as if it is a single value. There is reference to "the potential variability of the parameter values", and to "variability of the measured property values", but there is still a very strong implication that this representative value is a single number, and no requirement to assess and make use of the uncertainty associated with that number. This must surely be incorrect, and I believe that there needs to be some formal way in which the variability or uncertainty in the data is acknowledged and accounted for, not only in the selection of the characteristic value but also in its use.

### 2.10 Calibration of partial factors

While planning this keynote talk, I was also grappling with trying to carry out the calibration to determine the partial factors required for limit state design of earth retaining structures for insertion into the draft standard to achieve prescribed levels of reliability. Reliability Indices for Australian Buildings are prescribed in the NCC, and application of the current version of the earth retaining structures standard has become a legal requirement under the NCC (ABCB, 2019). This has happened without the reliability produced by compliance with the existing standard being checked, which makes it hard to calibrate against something that has not been calibrated itself, while at the same time trying to allow for better definition of uncertainty. Unfortunately I found it impossible to locate engineers among the volunteers on the committee who are willing and able to help. In the end most of them did not even seem to understand what was required, even when presented with a proposal for the calibration work to be done here at University of Newcastle. They were unwilling to consider trying to raise funds, and seemed to think that any University could do this sort of work for nothing. This was compounded by the fact that the administrative body, Standards Australia, had no idea of what was involved and were not willing to assist in any way whatever.

All of this brings me back to my topic from the beginning, which was the education of engineers, especially civil engineers and even more especially geotechnical engineers, in statistics. I have been a practicing geotechnical consulting engineer in Australia for the past 16 years, and in that time I have come into contact with very many other practicing geotechnical engineers. A very small proportion of them have a sound or better grasp of statistics as applied to civil and geotechnical engineering, and most of those are here at this Symposium. The vast majority see no application of statistics in their work, and this must be because the subject is being inadequately taught.

### 2.11 Practicing geotechnical engineers' experience of statistics

I am aware of another anecdote, and many of you may also have heard this because it comes from one of your senior members. He is an internationally renowned expert in geotechnical risk and statistics, and he advised me that he had a plan to teach a course at his university. This was apparently very strongly opposed by the mathematics department, who took the matter to the highest available level of dispute resolution within the university system. I am not even sure how it ended but it demonstrates the extreme feelings involved in this subject, and yet we must do something. In the last 6 months I have had the experience of reading a report written by a senior geotechnical engineer, in which I noted that the data in relation to hydraulic conductivity had not been presented. Noting that hydraulic conductivity is a notoriously difficult parameter to measure, with values often covering several orders of magnitude, I was surprised by his response that he had not presented the data because it was "too variable". This can only reflect his concept that all geotechnical parameters are numbers, which are supposed to have unique values. This was an engineer with 15 or 20 years of experience, both in Australia and overseas.

More recently I was in discussion with another PhD qualified geotechnical engineer, also with about 15 years of experience, who is developing an interest in statistics and probabilistic analysis. I mentioned the importance of spatial variability to him, and was surprised that he did not easily differentiate between spatial variability and borehole spacing. Just another example of the poor education of our geotechnical engineers in statistics.

One more example is a colleague who is also an internationally recognised professor of geotechnical engineering, specializing over a long and successful career in piling, pile groups, and offshore geotechnics. We actually wrote a paper together recently about probabilistic design of pile groups, based on him writing a Python script front end for his well-established spreadsheet based program. He is a very clever fellow, and coped well with the mathematics, statistics and programming involved, but confessed that the probabilistic approach was all new to him. How much more so to normal geotechnical engineers pursuing their careers in consulting offices around the world.

I have a final anecdote with regard to the general lack of awareness of the skills associated with statistics within the geotechnical engineering community. This is an example of which I confess to have only limited knowledge, but I consider it to be important. It involves an attempt by the Queensland Department of Main Roads, as they were back in 2005, to introduce probabilistic analysis into geotechnical design. This was in place when I arrived in Australia in 2006, and my work at that time did not involve QDMR projects so I did not come face to face with the methods. I am aware that it included application to slope stability and that the book, Soil Strength and Slope Stability, by Duncan and Wright was one that was specified. What I do know is that by the time I moved to work for GHD at the beginning of 2011 the idea had been dropped. It has never been made public what were the reasons for this, and I tried to get some speakers from QDTMR to come to this symposium to talk about the experience, but that was not possible. I therefore assume, based on all the other things that I have already said about the understanding of statistical concepts by geotechnical engineers, that we were just not ready as an industry to make such a step change.

## 3 The way forward

This Symposium appears to someone on the outside like me to be a type of exclusive club. You are all unquestionably experts in your fields, and you meet up every two years to discuss your latest progress and where you will concentrate your attention next. Your work is invaluable, but it is just not reaching the mainstream of geotechnical practice. I therefore appeal to you, because many of you are directly involved in education, to come up with new ways to disseminate your work to a wider audience. This must start with a change in the teaching of statistics to civil engineers and especially geotechnical engineers, so that the relevance of the subject cannot be lost to the majority. Since it appears likely that Mathematics Departments are likely to be desperate for teaching hours for some time to come, it is also likely that they will need to teach the courses, but should only be allowed to do so if they accept the necessary curriculum changes to make the subject of obvious relevance to civil engineers. The use of Ang & Tang as the textbook would probably be a suitable step, with appropriate reference to Baecher and Christian (2003).

Then we need to find ways in which the valuable research work that has been and is being done can be accessible to a wider audience. Currently there are too many "obscure" publications where useful practical work can be found. With the massive influence of the world wide web, and on-line learning, maybe we need to make more short courses and webinars available, and also attractive to engage with, but we need to be very careful about the content. There will need to be quite a wide range of entry levels, to ensure that participants are neither bored nor out of their depth, and I am learning through experience that some special skills are required. For example, when talking to a slide in PowerPoint, the cursor is often not necessary, or it is easy to follow. On the other hand, when using a spreadsheet such as Excel live on screen, it is a different matter. The operator knows where the cursor is, knows where it is going, and has control of the mouse to know when and which direction it will move. The viewers know only where the cursor is, and have no idea of in which direction or how fast or far it will move. This makes it very difficult to follow, and easy for viewers to get lost, without the operator being aware at all. These skills will probably get developed as we engage more in developing on-line learning, but we need to watch out for them now. There has been a very good development in the form of a course on probabilistic methods that was widely available, but I must confess that I was unable to complete the course, because of errors in the teaching material and in the assignments. That was very unfortunate for me.

## 4 Conclusions

So to a few conclusions:

1 There is a drastic need for engineers in general, and geotechnical engineers in particular, to have a much better understanding of statistical processes and how they help us to understand and interpret geotechnical data.

2 There is an equal need for geotechnical engineers to understand that geotechnical data are not discrete numbers as implied by a single characteristic value, but are random variables, especially with spatial variability, and that gaining an appreciation of that variability is the key to deriving characteristic values with their associated uncertainty, and to the carrying out appropriate analyses.

3 There is a need for all of the excellent research that is being done in the field of statistics in geotechnical engineering to be disseminated much more widely than it is at present, so that this symposium does not continue to be an exclusive club.

## References

ABCB (2019). National Construction Code – Building Code Australia – 2019. *ABCB, Canberra, Australia*.

Australian Geomechanics Society (2007). Landslide Risk Management. *Australian Geomechanics Journal.* 42 (1)

Baecher, G.B. and Christian, J.T. (2003). Reliability and Statistics in Geotechnical Engineering. *Wiley, Chichester, UK*.

Bond, A. and Harris, A. (2008). Decoding Eurocode 7. *Taylor & Francis, London and New York*.

CIRIA. (2003). Embedded retaining walls – guidance for economic design. Report C580. *CIRIA, London*.

CIRIA. (2017). Guidance on embedded retaining wall design. Report C760. *CIRIA, London*.

Day, P. and Briaud, J-L. (2022). Are we overdesigning: Results of ISSMGE Survey. *Proc. 20th International Conference on Soil Mechanics and Geotechnical Engineering.* Sydney, Australia.

Fellenius, B.H., Hussein, M., Mayne, P., and McGillivray, R.T. (2004). Murphy's Law and the pile prediction event at the 2002 ASCE GeoInstitute's Deep Foundations Conference. *Deep Foundation Institute Meeting on Current Practice and Future Trends in Deep Foundations*. Vancouver, 29-43.

Fellenius, B.H. (2013). Capacity and load-movement of a CFA pile: A prediction event. *ASCE GeoInstitute Geo Congress, San Diego. ASCE Geotechnical Special Publication*, GSP 229, 707-719.

Fellenius, B.H. and Terceros, M.H. (2014). Response to load of four different types of bored piles. *Proc. International Conference on Piling and Deep Foundations.* Stockholm.

Fellenius, B.H., Massarsch, K.R., and Tereceros, M.H. (2017). The Bolivian Experimental Site for Testing Piles (B.E.S.T.) Test Results. *Volume 3 of Proc. of the 3rd Bolivian International Conference on Deep Foundations.* Santa Cruz, Bolivia.

Fenton, G.A. and Griffiths, D.V. (2008). Risk Assessment in Geotechnical Engineering. *Wiley, New Jersey*.

Hausmann, M.R., Shirley, A.F. and Boyd, M. (1996). New Draft AS – Earth Retaining Structures (including reinforced soils). *Proc., 7th ANZ Conference on Geomechanics,* Adelaide.

ISC. (2004). International Prediction Event – Behaviour of bored, CFA and driven piles in residual soil, ISC'2 Experimental Site. ISC, Porto and Lisboa.

Low, B.K. (1996) Practical Probabilistic Approach using Spreadsheet. *ASCE Geotechnical Special Publication No 58, Proc. Uncertainty in the Geologic Environment – From Theory to Practice.* Madison, Wisconsin. 2, 1284-1302.

Low, B.K. (2003) Practical probabilistic slope stability analysis. *Proc. Soil and rock America 2003, 12th Panamerican Conference on Soil Mechanics and Geotechnical Engineering and 39th US Rock Mechanics Symposium.* MIT, Cambridge. 2, 2777-2784.

Low, B.K. (2005) Reliability-based design applied to retaining walls. *Géotechnique*, 55 (1), 63-75.

Low, B.K. (2010) Slope reliability analysis some insights and guidance from practitioners. *Proc. 17th Southeast Asian Geotechnical Conference.* Taipei, Taiwan.231-234

Low, B.K., Lacasse, S. and Nadim, F. (2007). Slope reliability analysis accounting for spatial variation. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards.* 1:4, 177-189. DOI: 10.1080/17499510701772089

Low, B.K. and Tang, W.H. (1997). Automated reliability-based design of footing foundations. *Proc. 7th International Conference on Structural safety and Reliability, ICOSSAR '97.* Kyoto, Japan. 3, 1837-1840.

Low, B.K. and Tang, W.H. (1997). Efficient reliability evaluation using spreadsheet. *ASCE Journal of Engineering Mechanics.* 123 (7), 749-752.

Low, B.K. and Tang, W.H. (1997). *Probabilistic slope analysis using Janbu's Generalised Procedure of Slices.*

Low, B.K. and Tang, W.H. (1997). *Reliability analysis of reinforced embankments on soft ground.*

Low, B.K. and Tang, W. (2007) Efficient spreadsheet algorithm for first-order reliability method, *ASDCE Journal of Engineering Mechanics,* 133 (12), 1378-1387

Standards Australia. (2002). Earth retaining structures, AS 4678:2002. *Standards Australia, Sydney*.

Standards Australia. (2009). Piling – Design and installation, AS 2159:2009. *Standards Australia, Sydney*.

Tereceros, M.H. and Fellenius, B.H. (2015). A Prediction Challenge to All Delegates to the 2nd CFPB, Santa Cruz, Bolivia. *Proc, 2nd Congreso Internacional de Fondazione's Profundas de Bolivia*. Santa Cruz, Bolivia.