FROM PROBABILITY TO FUZZY SETS: THE STRUGGLE FOR MEANING IN GEOTECHNICAL RISK ASSESSMENT

Michael Oberguggenberger and Wolfgang Fellin
Faculty of Civil Engineering and Architecture, University of Innsbruck
Technikerstraße 13, 6020 Innsbruck, Austria
michael.oberguggenberger@uibk.ac.at, wolfgang.fellin@uibk.ac.at
T: ++43 512 507 6824, F: ++43 512 507 2941

ABSTRACT

The paper addresses basic issues concerning the interpretation of probability in the probabilistic safety concept. Using a simple geotechnical design problem we demonstrate that the failure probability depends in an extremely sensitive way on the choice of input distribution function. We conclude that the failure probability has no meaning as a frequency of failure. It supplies, however, a useful means for decision making under uncertainty. We suggest a number of alternatives, as interval probability or fuzzy sets, which serve the same purpose in a more robust way.

INTRODUCTION

This paper addresses the role of probabilistic modeling in geotechnical problems. There is a general awareness of the uncertainties in all questions of geotechnical engineering; by now, it has also become clear that the uncertainties themselves have to be subject to modeling. Common practice is to use a probabilistic set-up to achieve this task. We agree with EINSTEIN (4) that the probabilistic format supports decision making under uncertainty. It helps structuring the problem and aids in obtaining qualitative judgments.

The numerical values thus obtained, like failure probabilities and safety factors, play an important role in comparative and qualitative studies. The point we wish to make, however, is that these numerical values do not make quantitative assertions about reality.

In particular - contrary to common language -, the failure probability cannot be interpreted as a frequency of failure. To support our claims, we first enter a general discussion of the probabilistic format and its underlying assumptions in the second section. In the third section we present a simple example of a centrically loaded square footing to demonstrate that the failure probability may vary by orders of magnitude, namely between 10^{-11} and 10^{-3}, just by fitting different distributions (normal, lognormal, triangular) to the same input data obtained from laboratory tests. This effect has been observed by other authors as well (5). The required design dimensions of the footing are less sensitive to the input distribution, but still vary by a factor of 1.5 to 1.9. In the last section we discuss alternative concepts that are currently under investigation in reliability theory.

PROBABILISTIC MODELING

To set the stage, we briefly recall the format of the probabilistic safety concept. The vector R lumps together all variables describing the resistance of a structure, while S signifies the loads. An engineering model incorporating the structure as well as its geotechnical environment supplies the limit state function $g(R,S)$. Negative values of $g(R,S)$ correspond to failure. Modeling R and S as random variables, we can compute the failure probability

$$p_f = P\big(g(R,S) < 0\big)$$

provided the probability distributions of R and S and their parameters are known. However, the current codes employ critical values R_k and S_k (certain percentiles of R and S) and partial safety factors γ_R and γ_S, so that the designing engineer has to verify a relation of the type

$$R_k/\gamma_R \geq \gamma_S S_k. \tag{1}$$

In theory, the critical values and the partial safety factors are computed in such a way that (1) holds if and only if p_f attains a certain required value p_{fr}. For example, in the case of normally distributed resistance and loads and $g(R,S) = R - S$ one must choose

$$\gamma_R = \frac{1 - Q_p V_R}{1 - \beta\alpha_R V_R},\ \alpha_R = \frac{\sigma_R}{\sqrt{\sigma_R^2 + \sigma_S^2}},\ V_R = \frac{\sigma_R}{\mu_R}$$

where μ_R, σ_R are the mean and standart deviation of R, $p_{fr} = \Phi(-\beta),\ Q_p = \Phi(1 - p),\ R_k$ is the $(1-p)\cdot 100\%$ - percentile of R, and Φ the cumulative standard normal distribution. In practice, γ_R and γ_S are not computed but rather prescribed in the codes, where they entered as results of negotiations of the respective committees.

For a general description of the uncertainties involved in geotechnical modeling we refer e.g. to (4). Here we wish to emphasize a few aspects that are important for interpreting the results of the probabilistic method.

Uncertainty of non-probabilistic input. At the foremost place, the soil model has to be mentioned. Is it a continuum model with an elasto-plastic, a hypoplastic or another constitutive law, or is it a discrete model? Completely different sets of influence variables may arise in this way. Second, what is the failure mechanism? Is it assumed that failure occurs when averages of the parameters exceed certain values, or is failure assumed to be due to localized disturbances?

Uncertainty of probabilistic input. The probabilistic input in the model consists in specifying the types of probability distributions of the variables under consideration as well as the parameters of the distributions. We first point out that probability plays several roles here. It has to describe as diverse things as measurement errors, random data fluctuations, lack of information, and it may even have the burden of compressing spatial variability into parameters of a model which neglects spatial variability.

Second, what *is* probability? In engineering modeling, at least three interpretations can be identified. There is *classical probability* which assigns probabilities from combinatorial considerations (for example, the assertion that the number k of successes in a sequence of *n* trials has a binominal distribution). Then there is the *frequentist interpretation* in which probability is an approximation to relative frequencies of outcomes in large samples. For the practical purpose of determining confidence intervals and performing statistical tests for the parameters of a single random variable, “large” means $n \geq 20$. This sample size is rarely available in engineering practice, and this makes *subjective probability* attractive in engineering: in this interpretation, probability is just a subjective measure of confidence in the available information. Using prior knowledge and Bayes’ rules, even the information of a sample of size $n = 1$ may be incorporated in the assessment of a certain parameter. It should be pointed out that *frequentist probability*, in its applications, has two facets: there is *individual* and *collective frequency*. For example, over the past decade 440 dam failures of large dams (more than 7.7 m high) have been recorded in the US. There are 75.000 dams (over 7.7 m) in the US (data according to (1)). That translates into a collective frequency of failure of about 1/1700 per year. Does that mean that a specific dam under consideration will fail once in the next 1700 years? Of course, for an assessment of the individual failure rate, individual data like local yearly precipitation averages etc. are needed. If these data are lacking, the collective frequency is a highly questionable estimate for the individual frequency. Often only a mixture of individual and collective data is available.

Summarizing, we note that the failure probability in the engineering safety concept is an amalgam of classical, subjective and frequentist bits of information. That observation does not diminish its importance and usefulness for comparative studies of scenarios. However, the failure probabilities obtained in different engineering projects cannot be related to each other, as they depend on many individual choices that had to be made along the way. All the more, they do not have a meaning as an expected frequency of failure for an individual project. They are what KLINGMÜLLER and BOURGUND have termed *operational probabilities* (7). There is one more point which emphasizes these assertions: The probability distributions of the input data have to be chosen as part of the modeling procedure. It is common statistical practice to fit them by some likelihood principles and accept them if they pass a number of statistical tests (Chi-square test of fit, Kolmogorov-Smirnov-test). We show in the next section that the choice of probability distribution, even among the standard types in use, has dramatic effects on the failure probability. Thus keeping the limit state model as well as all data fixed, while slightly changing same input distributions may (and usually does) completely change the value of p_f.

EXAMPLE

We consider a simple example of a centrically loaded square footing with the dimensions $a = b = 1.0$ m and $t = 0.5$ m in silt (Figure 1).

For the purpose of demonstration, we choose a soil which was very well tested. Twenty direct shear tests with disturbed samples of silt had been performed with equal initial conditions (2). Note that this is much more than usually available on construction site.

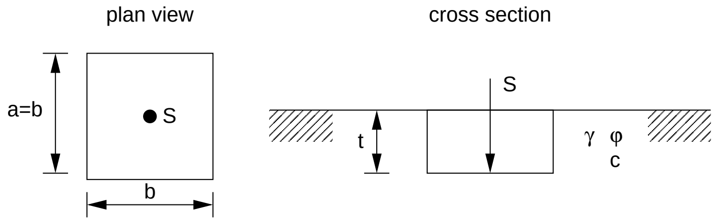

Figure 1: Bearing capacity example

The water content of each specimen was near the liquid limit of the silt. The unit weight was nearly the same in each experiment: $\gamma = 19.8$ kN/m^3. The shear velocity was constant and equal in each experiment. The cohesion resulted to $c = 0$ in each test. The obtained friction angles are listed in Table 1.

Table 1: $n = 20$ direct shear tests

friction angle φ [°]									
25.6	25.5	24	26	24.1	24	28.5	25.3	23.4	26.5
23.2	25	22	24	24.9	30	27	24.4	24.3	29.5

The characteristic soil parameters are defined in EC 7, part 1, 2.4.3.(5): "The characteristic value of soil or rock parameter shall be selected as a cautious estimate of the value effecting the occurrence of the limit state."

The value effecting the limit state is the shear strength of the soil $\tau_f = c + \sigma \cdot \tan\varphi$, with the normal stress σ. Therefore, the parameter whose distribution we analyze is the friction coefficient

$$\nu = \tan\varphi \ .$$

The mean value and the standard deviation of the sample are

$$\bar{\nu} = \frac{1}{n}\sum_{i=1}^{n} \nu_i = 0.474 \qquad , \qquad s_\nu = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\nu_i - \bar{\nu})^2} = 0.0452 \ .$$

There is some discussion in the literature, but no decisive conclusion, what distribution is appropriate for the friction coefficient. We consider some of those that have been proposed.

Normal distribution: The simplest and most common choice is the Gaussian normal distribution $\nu \sim \mathcal{N}(\mu_\nu, \sigma_\nu^2)$, with the estimations $\mu_\nu = \bar{\nu}$ and $\sigma_\nu^2 = s_\nu^2$. This distribution has the disadvantage that the friction coefficient can be negative, which is physically impossible. Though this happens only with low probability, a better choice may be the following lognormal distribution with two parameters, which is strictly nonnegative. Nevertheless, the normal distribution is often used, and can be seen as providing a conservative (high) estimate of the failure probability.

Lognormal distribution with two parameters: In this case the natural logarithm of the friction coefficient is assumed to be normally distributed: $\ln \nu \sim \mathscr{N}(\mu_{\ln \nu}, \sigma^2_{\ln \nu})$, with the estimations for the parameters

$$\mu_{\ln \nu} = \overline{\ln \nu} = \frac{1}{n}\sum_{i=1}^{n} \ln \nu_i \ , \qquad , \qquad \sigma^2_{\ln \nu} = \frac{1}{n-1}\sum_{i=1}^{n} (\ln \nu_i - \overline{\ln \nu})^2 \ .$$

This distribution is sometimes criticized to give too high probabilities for high friction coefficients.

Both distributions, normal and lognormal, pass the standard statistical tests (Table 3), but as we see in Figure 2 they do not match the peak of the histogram and its asymmetry very well. A better statistical fit is given by the two following distributions.

Lognormal distribution with three parameters: We shift the lognormal distribution by the lower limit of the friction coefficient ν_0: $\ln(\nu - \nu_0) \sim \mathscr{N}(\mu_{\ln(\nu-\nu_0)}, \sigma^2_{\ln(\nu-\nu_0)})$. The parameters are fitted with the maximum likelihood method. This distribution fits the histogram very well, see Figure 3.

Triangular distribution: The triangular distribution is seldom used in engineering, but it is simple and matches the histogram also well (Figure 3): $\nu \sim \mathscr{T}(l_\nu, m_\nu, u_\nu)$. The lower boundary l_ν, modal value m_ν and upper boundary u_ν are fitted with the maximum likelihood method.

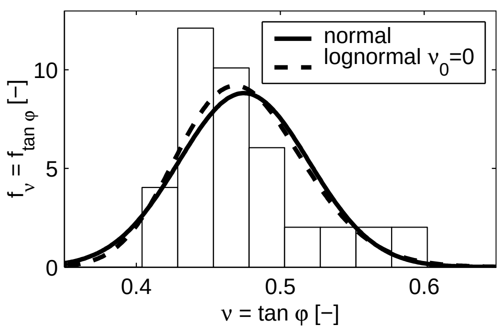

Figure 2: Fitted normal and two parameter lognormal distribution

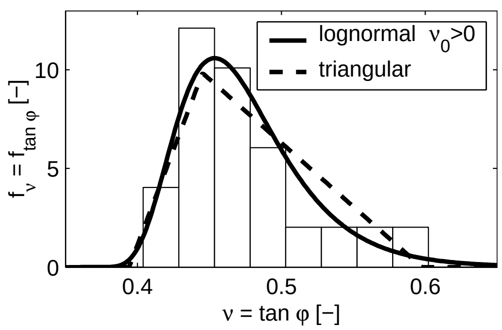

Figure 3: Fitted three parameter lognormal and triangular distribution

Table 2: Parameters of the distributions

normal:	$\mu_\nu = 0.474$	$\sigma_\nu = 0.0452$	
two parameter lognormal:	$\mu_{\ln \nu} = -0.749$	$\sigma_{\ln \nu} = 0.0923$	
three parameter lognormal:	$\mu_{\ln(\nu-\nu_0)} = -2.192$	$\sigma_{\ln(\nu-\nu_0)} = 0.3601$	$\nu_0 = 0.355$
triangular:	$l_\nu = 0.394$	$m_\nu = 0.445$	$u_\nu = 0.598$

Table 3: Statistical tests of the distributions: computed values are below the 95%-fractile

	KS		χ^2	
	computed	95%-fractile	computed	95%-fractile
normal	0.162	0.294	5.12	5.99
two parameter lognormal	0.142	0.294	3.71	5.99
three parameter lognormal	0.113	0.294	1.99	3.84
triangular	0.154	0.294	3.48	3.84

Spatial and local variation

We mention briefly the well-known geotechnical spatial versus local variation problem. On the one hand the strength of the material cannot be measured exactly at each point, and on the other hand the soil is usually spatially inhomogeneous. Both effects contribute to the variability of the data. For the sake of simplicity we assume in this example $\tan\varphi$ to be homogeneously distributed in space, but otherwise a random variable with statistical fluctuations.

GRIFFITHS and FENTON (6) showed in the case of a bearing capacity problem in an undrained soil, that the spatial distribution of the undrained cohesion c_u leads to a reduced capacity factor N_c, because the failure plane tends to find its way through the weakest zones. The failure mechanism becomes asymmetric. The characteristic value of the cohesion effecting the occurrence of the limit state is therefore not the spatial mean value, it is somewhat less, depending on the practically unknown spatial correlation length.

The results in (6) let us presume that a calculation with spatially homogeneous soil but allowing statistical variation of the soil strength leads to an upper bound of the failure probability for small values of p_f (<0.05). Introducing a spatial variability of the soil parameters would bring additional unknown probability parameters into the calculation. It would not bring new effects for this presentation, except a prolongation of our list of failure probabilities.

Bearing capacity, load

We use the design bearing capacity due to Austrian code B4435-2 as loading of the footing. In a common interpretation the cautious estimate of the friction coefficient $\nu_k = \tan\varphi_k$ is the lower boundary of the 95% confidence interval for the mean value of $\nu = \tan\varphi$ (9): $\nu_k = \bar{\nu} - t_\gamma s_\nu/\sqrt{n}$, where t_γ is the γ-percentile of the STUDENT t-distribution with $(n-1)$ degrees of freedom. For $\gamma = 95\%$ and $n = 20$ we have $t_\gamma = 2.09$. Therefore the characteristic friction coefficient is $\nu_k = \tan\varphi_k = 0.454$.

The design values of the soil parameters are $\tan\varphi_d = \tan\varphi_k/\gamma_\varphi$, $\gamma_d = \gamma_k/\gamma_\gamma$, with $\gamma_\gamma = 1.0$ and $\gamma_\varphi = 1.3$. The design bearing capacity results to $Q_{f,d} = 101.93$ kN. The characteristic load is $S_k = Q_{f,d}/\gamma_S$ with $\gamma_S = 1.0$ according to Austrian code B4435-2. We interpret this characteristic load as 95%-fractile of a normally distributed load with a coefficient of variation $V_S = 0.1$ according to EC 1, part 1, 4.2 (4): $S_k = \mu_S + k_N \sigma_S$, with $k_N = 1.645$. Thus the mean value and the standard deviation of the load is

$$\mu_S = \frac{Q_k}{1+1.645 V_S} = 87.53 \text{ kN} \qquad , \qquad \sigma_S = V_S \mu_S = 8.75 \text{ kN} . \tag{2}$$

Numerical simulation

The simulations were done with the aid of a FORTRAN-Program using the random number generators for a uniform distribution $\mathscr{U}(l,u)$ (for high number of calls $N > 10^8$) and a normal distribution $\mathscr{N}(0,1)$ published in (8). The realizations of the friction coefficient for the various distributions were calculated as follows:
normal distribution: $\nu = \mu_\nu + \sigma_\nu \cdot \mathscr{N}(0,1)$
two parameter lognormal distribution: $\nu = \exp\left(\mu_{\ln \nu} + \sigma_{\ln \nu} \cdot \mathscr{N}(0,1)\right)$
three parameter lognormal distribution: $\nu = \exp\left(\mu_{\ln(\nu-\nu_0)} + \sigma_{\ln(\nu-\nu_0)} \cdot \mathscr{N}(0,1)\right) + \nu_0$
triangular distribution:

$$\nu = \begin{cases} l_\nu + \sqrt{\mathscr{U}(0,1) \cdot (u_\nu - l_\nu) \cdot (m_\nu - l_\nu)} & \text{if} \quad \mathscr{U}(0,1) \leq \frac{m_\nu - l_\nu}{u_\nu - l_\nu} \\ u_\nu - \sqrt{(1-\mathscr{U}(0,1)) \cdot (u_\nu - l_\nu) \cdot (u_\nu - m_\nu)} & \text{if} \quad \mathscr{U}(0,1) > \frac{m_\nu - l_\nu}{u_\nu - l_\nu} \end{cases}$$

A series of numerical simulations were carried out. In each simulation N independent realizations of the friction coefficient and the load

$$S = \mu_S + \sigma_S \cdot \mathscr{N}(0,1)$$

were computed.

With each realization of the friction coefficient $\nu = \tan\varphi$ a realization of the resistance

$$R = Q_f = ab(\gamma b N_\gamma s_\gamma + \gamma t N_q s_q) , \tag{3}$$

with $N_q = \frac{1+\sin\varphi}{1-\sin\varphi}\, e^{\pi \tan\varphi}$, $N_\gamma = (N_q - 1)\tan\varphi$, $s_q = 1 + \frac{b}{a}\sin\varphi$, and $s_\gamma = 1 - 0.3\frac{b}{a}$ was calculated.

The number of failure events $Z = R - S < 0$ was counted in H_f. The relative frequency provides an estimate for the failure probability $p_f \approx H_f/N$. The approximative failure probabilities are listed in Table 4. The value of p_f varies in between 10^{-3} and 10^{-11} for different distributions of $\tan\varphi$. For curiosity we mention that transforming the data for $\tan\varphi$ in Table 1 into a histogram for R via (3) and fitting a normal distribution to R, we get $p_f = 2 \times 10^{-2}$. These variations of several orders of magnitude were also confirmed by analytical estimates.

If we design the dimension $a = b$ of the footing with the given failure probability $p_f = 10^{-6}$ (probability index $\beta = 4.7$, EC 1, Table A.2) we get the varying areas of the footing in Table 5. The minimal and the maximal footing area differ by a factor of 1.9. For a failure probability of $p_f = 10^{-4}$ this factor is 1.5.

Table 4: Failure probabilities p_f in m simulations with N realizations

distribution	p_f	$N \times m$
normal	0.81×10^{-3}	$10^8 \times 4$
two parameter lognormal ($\nu_0 = 0$)	1.1×10^{-4}	$10^8 \times 4$
three parameter lognormal ($\nu_0 > 0$)	1.0×10^{-9}	$10^{11} \times 5$
triangular	1×10^{-11}	$10^{11} \times 9$

Table 5: Dimensions of the footing for fixed failure probability $p_f = 10^{-6}$

distribution	$a = b$ [m]	difference [%]	$A = a \cdot b$ [m^2]	difference [%]
normal	1.26	35	1.59	85
two parameter lognormal ($\nu_0 = 0$)	1.11	19	1.23	43
three parameter lognormal ($\nu_0 > 0$)	0.95	2	0.90	5
triangular	0.93	–	0.86	–

ALTERNATIVES AND CONCLUSIONS

As demonstrated in the previous section, the failure probability is highly sensitive to changes in the input distribution parameters. A more robust measure as a basis for decisions appears desirable, in particular, a measure that would allow to account for model dependence and non-probabilistic uncertainties as well. Current trends in reliability and decision theory aim at relaxing same axioms of probability theory. This goes under the heading of “imprecise probability” and subsumes interval probability, convex sets of probability measures, random sets, plausibility and belief functions, fuzzy sets, and more (see (3) for a review). We exemplify two of these concepts in the case of the situation of the previous section. The most conservative estimate for the failure probability is obtained by admitting all probability distributions that could feasibly produce the data from Table 1. These could e.g. be modeled by the set of all probability distributions concentrated in the interval $[\underline{\varphi},\ \overline{\varphi}] = [20^\circ,\ 32^\circ]$. Still assuming that the load S is normally distributed according to (2), it can be shown that the lower and upper failure probabilities are obtained by

$$\begin{aligned} \underline{p}_f &= \inf\{P(Z<0) : \varphi \in [20,32]\} = 0, \\ \overline{p}_f &= \sup\{P(Z<0) : \varphi \in [20,32]\} = 0.00235. \end{aligned}$$

Less conservative estimates would be obtained by restricting the set of admitted input distributions.

An even more robust non-probabilistic approach would employ fuzzy sets to describe the parameter uncertainties. The construction of a fuzzy set describing input variability may be based on an expert’s risk assessment of possible ranges of the angle of internal friction, say. An example of a triangular fuzzy number is given in Figure 4.

This says, e.g., that the expert had assigned the degree of possibility $1/2$ that $\nu = \tan\varphi$ is in the range of $[0.415, 0.52]$ and possibility 1 that it has the value 0.44. The

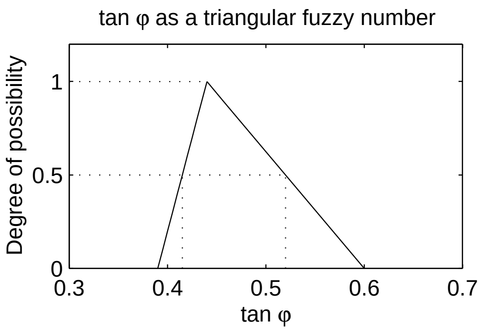

Figure 4: Triangular fuzzy number for $\tan\varphi$

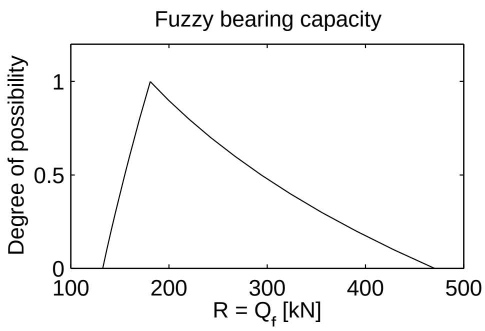

Figure 5: Fuzzy number for bearing capacity R

computation of the fuzzy set describing the resistance is done according to the rules of fuzzy set theory (intervals of equal possibility in input and output correspond to each other). The result is depicted in Figure 5 and allows, e.g., to estimate the degree of possibility that $R < S$ for any given value of S (even fuzzy S as well). It appears that such considerations may be equally useful as a framework for decision making under uncertainty.

REFERENCES

[1] G.B. Baecher and J.T. Christian. The practice of risk analysis and the safety of dams. *University of Maryland (available electronically at* `http://www.glue.umd.edu/~gbaecher/papers.d/Cairo Paper.pdf`*)*, 2001.

[2] E. Bösinger. Serienuntersuchungen zum Vergleich verschiedenener Rahmenschergeräte. Institut für Bodenmechanik und Felsmechanik der Universität Karlsruhe, persönliche Mitteilung, 1996.

[3] G. De Cooman, T. Fine, and T. Seidenfeld, editors. *ISIPTA'01, Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications*. Shaker Publishing BV, Maastricht, 2001.

[4] H.H. Einstein. Quantifying uncertain engineering geologic information. *Felsbau*, 19(5):72–84, 2001.

[5] I. Elishakoff. What may go wrong with probabilistic methods? In I. Elishakoff, editor, *Whys and Hows in Uncertainty Modelling: Probability, Fuzziness and Anti-Optimization*, volume 388 of *CISM Courses and Lectures*, pages 265–283. Springer, Wien, New York, 1999.

[6] D.V. Griffiths and A. Fenton. Bearing capacity of spatially random soil: the undrained clay Prandtl problem revisited. *Geotechnique*, 51(4):351–359, 2001.

[7] O. Klingmüller and U. Bourgund. *Sicherheit und Risiko im Konstruktiven Ingenieurbau*. Verlag Vieweg, Braunschweig/Wiesbaden, 1992.

[8] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes*. Cambridge University Press, 2nd edition, 1992.

[9] H.R. Schneider. Definition and determination of characteristic soil properties. In *Proceedings of the Fourteenth International Conference on Soil Mechanics and Foundation Engineering*, pages 2271–2274. Balkema, 1997.