Random Network Modeling for Determination of Representative Specimen Size of Compacted Clays
S. Thangavadivelu, L. N. Reddi, and S. Menon 1303

Evaluating Subsurface Uncertainty Using Zonal Kriging
W. L. Wingle and E. P. Poeter 1318

CASE HISTORIES

Practical Methods for Managing Uncertainties for Geosynthetic Clay Liners (Invited Paper)
D. E. Daniel and R. B. Gilbert 1331

Defects in Soft Clays: A Challenge to Site Characterization for Stability Analysis (Invited Paper)
G. A. Leonards and R. J. Deschamps 1347

The Reliability Analysis of a Major Dam Project
J. Barneich, D. Majors, Y. Moriwaki, R. Kulkarni, and R. Davidson 1367

Dealing with Uncertain and Highly Variable Geotechnical Conditions Beneath the Inco Smelter in Copper Cliff
K. J. Jansons 1383

Risk Assessment of Rockfall Hazard at Horse Mesa Dam: A Case History
P. M. Kandaris and K. M. Euge 1402

Deterministic and Probabilistic Analyses of Preload Design at a Hydraulic Fill Site
S. Thevanayagam, E. Kavazanjian, Jr., A. Jacob, S. Nesarajah 1417

Hazard Assessment of Debris Fans at Rico, Colorado
B. C. Wilbur 1432

Subject Index 1447

Author Index 1451

ORGANIZING AND EVALUATING UNCERTAINTY IN GEOTECHNICAL ENGINEERING

Robert V. Whitman[1]

ABSTRACT

Probabilistic methods are discussed with respect to four stages of a typical project: site characterization and evaluation, evaluation of designs, decision-making and construction control. Decision-making is discussed from the standpoint of acceptable risk. Unless clients or regulators are interested in quantifying risks as part of decision-making, engineers rely on traditional methods. Engineers hold the key to motivating clients to consider basing design on probabilistic thinking. More and better examples of applications of probabilistic methods are needed.

INTRODUCTION

Uncertainty is pervasive in geotechnical engineering. Practicing engineers of course worry about such questions as: Will we get the job?, When will it start?, Will the driller show up as promised?, When will we be paid?, etc. I am afraid this Conference won't help with those important uncertainties. We will have enough on our plate as we focus questions having to do with uncertainty in the engineering aspects of a project.

Engineers face uncertainties at all phases of a project: Is it possible that a site is so poor or contains unrevealed defects that make it unsuitable?; Is a field investigation adequate for characterizing the materials at a site?; What values should be assigned to soil parameters (strength, permeability, etc.) required for an analysis?; How accurate is an analysis leading to an important derived quantity (e.g. safety factor); and - ultimately - How

[1] Professor Emeritus, Massachusetts Institute of Technology, Cambridge, Mass.

confident are we that a proposed design is safe and adequate in other ways? There are also uncertainties as to just how well a design is being implemented during actual construction.

The founders and leaders of our profession have spoken and written extensively concerning the importance of recognizing uncertainties and taking them into account in design. Casagrande's well-known Terzaghi Lecture (1965) was specifically about "calculated risk", by which he meant very careful consideration of risk. Casagrande was not optimistic that risks could literally be calculated or even quantified.

During this conference, we *are* primarily interested in "quantifying" risk - that is, using numerical and analytical characterizations and methods to assist in making decisions concerning uncertainties such as those I just listed. In the years since Casagrande's lecture, there have in fact been advances in enumerating uncertainties in geotechnical engineering. There are new "tools" for use in guiding site exploration and characterization. There now are examples in which risks have been literally calculated as part of engineering projects, and also others where subjective judgments have been used to put numbers on risks. Increasingly, quantified risk is being used as the basis for engineering decision-making.

The subject matter of this Conference has been covered in the report "Probabilistic Methods in Geotechnical Engineering" (National Research Council [NRC], 1995). The 3rd Casagrande Lecture (Morgenstern, 1995) has dealt with managing risk in geotechnical engineering. I will mention a few other recent state-of-the-art papers, and I suspect that the plenary papers to this Conference, and other papers as well, will be state-of-the-art summaries covering portions of the subject. I will strive only to indicate the scope of useful probabilistic methods, referring to but a very limited segment of recent literature.

The comments I offer to begin this discussion are aimed at four questions:

* What do probabilists mean by all the words they use?

* How can probabilistic methods be used in geotechnical engineering?

* When, and for what type of projects, is it appropriate to use probabilistic methods?

* What should and can be done to encourage more use of probabilistic methods?

UNCERTAINTY WHEN TALKING ABOUT UNCERTAINTY

Fell (1994), in an excellent state-of-the-art paper concerning risk assessment relative to landslides, reports:

> Unfortunately, there are no generally accepted definitions of the terms used in risk assessment.....shortly after its formation in 1981, the United States Society for Risk Analysis established a committee to define risk. After 3 or 4 years of work the committee published a list of 14 candidate definitions and reported that it could not reach agreement. They recommended that a single definition of risk not be established but that everyone be free to define it as appropriate to his or her own work.

I certainly will not attempt to change this situation, but I do need to try to explain what I mean by various terms I will use.

I have already used the phrase *probabilistic methods*. This is a very loose concept, intended to cover a diverse range of techniques for expressing and dealing explicitly with uncertainty.

I have spoken of *quantifying risk*. This is meant to imply using numbers to express risk. Thus *quantified risk* is more explicit than Casagrande's *calculated risk*, which according to the American Heritage dictionary means "estimated with forethought". *Risk evaluation* is the process of arriving at *quantified risk*. Thus *evaluated risk* is the same as *quantified risk*.

However, *quantified risk* can be arrived at either by means of (a) theory and numerical calculations or (b) using subjective judgements. I have tried to think of simple phrases to distinguish between these two approaches, but I have been unsuccessful. Actually, both approaches may be used together in a *risk evaluation*.

I will strive to distinguish consistently between *hazard* and *risk*, using meanings that are widely if not universally accepted. When expressed in probabilistic terms, *hazard* expresses the likelihood that some event may occur. *Risk* expresses the likelihood that some loss occurs, and is often in the form of a product of *hazard* and a loss resulting from the event actually occurring. (Be warned, it is fairly certain that a strict probabilist will disagree with my use of the word *likelihood*. I use it merely to express a concept.)

Lastly, *probabilistic thinking* is another loose concept. It implies use of concepts from formal probability theory, reliability theory, statistics, etc., and possibly the actual use of some theoretical tools from these sciences - but without becoming a slave to formalism.

There is a moral: We must all be patient and diligent in our communications with each other, sparing no effort in our attempts to understand the message being sent.

PROBABILISTIC METHODS IN GEOTECHNICAL ENGINEERING

The scope of the papers being presented to this Conference makes clear the diversity of probabilistic methods that are being applied in geotechnical engineering. Different approaches are commonly used for different types of projects: landslides, foundations for offshore structures, environmental problems, etc. Different types of analysis are suitable for the several stages of a project.

For the purposes of my discussion, it is convenient to divide a project into four stages: (1) Site evaluation and characterization; (2) Design evaluation; (3) Decision-making, and (4) Construction control. The following sections discuss each of these stages. My aim is to indicate the types of questions that probabilistic methods can contribute at each stage, without going into technical detail. However it must be recognized that, from the standpoint of probabilistic methods, it may or may not be appropriate for these stages to proceed independently. For example, the choice of method for assessing the reliability of a design depends both upon the availability of data concerning the site and upon the way in which acceptable risk is judged.

SITE EVALUATION AND CHARACTERIZATION

A variety of probabilistic methods have been developed that can be useful during this stage of a project. The following discussion aims to suggest the range of possibilities, and is not exhaustive. Some - such as searches for "flaws" and construction of profiles - may be the end of using probabilistic methods for the project; others may aim to provide specific data required for further probabilistic analysis.

Designing a search to look for "flaws" at site

In this stage, the general nature of the site is established. A vital question is: Is the exploration program adequate to detect and reveal the extent of any flaws that might present unusual design problems or possibly make the site unsuitable for some intended purpose? Typical flaws are strata of especially weak or liquefiable soil or the

presence of an adversely sloping joint in rock, solution cavities, or channels of exceptionally high permeability. *Search theory* can be used to guide selection of patterns of borings and to help decide how many borings are necessary to reduce probability of an undetected flaw to below an acceptable limit. Application of search theory to geotechnical problems was pioneered by Baecher (1979) two decades ago. Halim and Tang (1993) present a recent contribution to this theory.

Characterizing variability over site

It is common practice to prepare soil profiles across a site, based upon soil types recorded in borings spaced some distance apart. Various probabilistic approaches can aid in this task, by identifying systematically possible correlations of soil types with depth and among boreholes. A mathematical technique known as Kriging is useful for this purpose. One common application is mapping of the elevation of the top of a soil type of particular interest. Nobre and Sykes (1992) provide one relatively recent example of using this approach to mapping the elevation of bed rock.

Spatial variability of soil properties, both horizontally and vertically, is an especially important problem. Accounting for this variability can be important in such diverse problems as estimating risks of slope failures with long embankments, study of the dispersion of plumes of pollutants, and evaluation of a site's resistance to liquefaction - all of which are mentioned subsequently. The work of Vanmarcke (1983) has provided a starting point for the characterization of spatial variability.

A very interesting application, described by Tang (1979), concerns the penetration of skirts for an offshore gravity platform. At the proposed site, layers of dense sand and hard clay occur at some locations, and the exact location at which the platform sets down cannot be controlled precisely. Cone penetration data were analyzed to take into account possible variations of skirt penetration resistance with depth and laterally. The analysis led to estimates for the mean$\pm$1s range for total resistance and the central 50% band for the unbalanced moment. Such information provides important guidance both for planning the sinking operation and monitoring and controlling the actual sinking.

Choosing values for material properties

A classic geotechnical engineering task is the selection of the value for a soil parameter to be used in some analysis. In some problems it is necessary and appropriate to rely upon judgement in making this decision. Increasingly, however, values are better selected systematically employing statistics and probability methods. For reliability analyses (to be discussed subsequently), it is at a minimum necessary to estimate the mean and standard deviation (or variance) for key parameters. In other situations (see subsequent discussion of codes), it may be desirable to select a value with some stated probability of not being below (or above as the case may be) the selected value.

There are several different types of uncertainties as regards the properties of soil as measured either in the field or in the laboratory: see Fig. 1.

- Data scatter, consisting of (1) real spatial variability and (2) random testing errors. Random testing errors should not be allowed to influence parameter selection; the magnitude of such errors should be identified and screened out of further analysis. Real spatial variability can be important, depending upon the distances over which it occurs compared to the scale of the project.

- Systematic errors, resulting from (1) errors because tests do not actually measure accurately the desired parameter, and (2) too few tests to average out random testing errors. The first of these difficulties (akin to model errors in analysis, to be discussed shortly) can be very important and require particular attention.

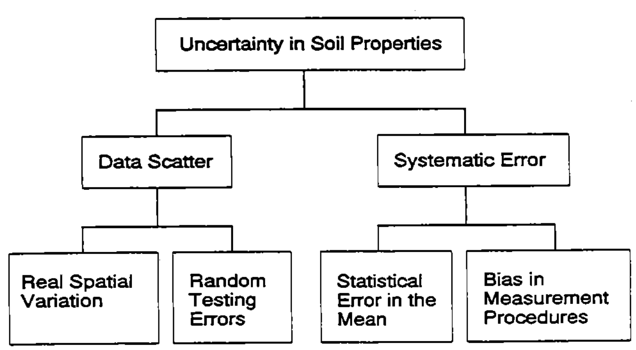


Figure 1 Categories of uncertainty in soil properties. (Christian *et al*, 1994)

The literature contains techniques for dealing with all of these difficulties in systematic ways. Christian *et al* (1994) give a brief summary as applied to problems with slope stability. The classic work by Vanmarcke (1977) provides methods for dealing with the possible consequences of spatial variability. The problem of deciding when enough tests have been performed to reduce random errors is similar to that of searching for "flaws". Wu *et al* (1989) present an interesting example of evaluating alternative exploration programs so as to choose the program that will most effectively define the site characteristic for a project. A reliability analysis (see below) provides the framework for the evaluation.

EVALUATION OF RISK

There is no one procedure for evaluating risk that is appropriate for all types of projects. The choice of method depends upon the approach that is most acceptable for the type of project, the data available, the degree to which there is reliance upon subjective judgement, and the criteria that will be used to judge whether the risk is acceptable or not.

Reliability analysis

A *reliability analysis* evaluates the probability that *capacity* (e.g. bearing capacity) exceeds *demand* (e.g. loading), where either or both capacity and demand are uncertain. This probability is called *reliability*, and probability of failure is (1 - reliability).

If probability distribution functions can be established for both capacity and demand, in principle an exact (but very tedious) calculation of reliability may be made. An alternative is to make many simulations, drawing random numbers to choose appropriate values for demand and capacity. Generally well-developed approximate methods are employed, which depend only upon characterization of capacity and demand by their means and standard deviations. Reliability analysis has often been applied to structural systems. There currently is considerable interest in the application of this approach to geotechnical engineering problems. A brief discussion of the method from the standpoint of geotechnical engineering appears in NRC (1995). Reliability-based design is the topic of the April 1996 Casagrande Memorial Lecture by Fred Kulhawy to the Boston Society of Civil Engineers Section of ASCE..

Since reliability analysis has frequently been used for the structural portions of large offshore structures, it is natural that the method has been applied to the foundations for such structures. The recent literature contains several good examples; a brief review of one will illustrate the key aspects of such an analysis.

Ronold and Bysveen (1992) evaluate the reliability of the foundation for a deep-water gravity platform resting on soft clay (Fig. 2). A standard for design of such structures is the worst 6-hour sea state during the lifetime of the platform. The largest significant wave height during this storm is uncertain. The capacity is determined from a stability analysis using the failure surface shown in Fig. 2, which was found to be the critical failure surface. The uncertainties concerning this capacity are:

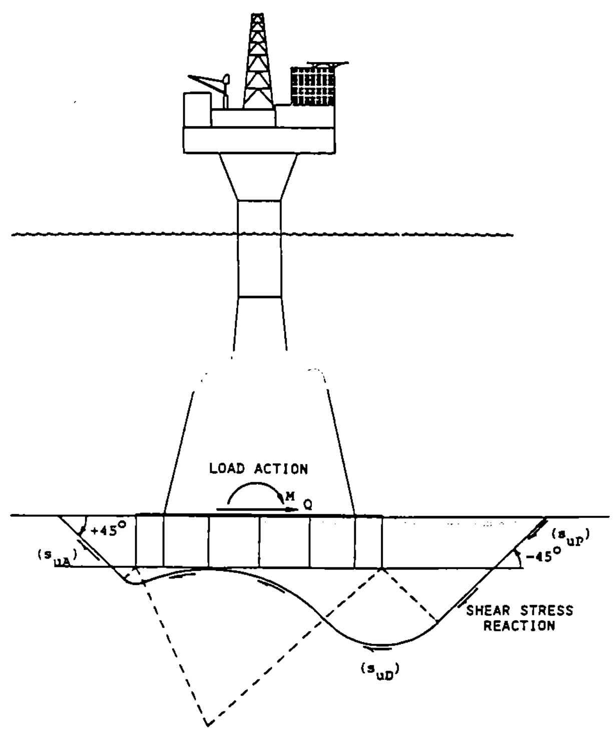


Figure 2 Deep-water platform on soft clay, with critical shear surface configuration through soil (Ronold and Bysveen, 1992)

- The undrained strength for active loading conditions, as evaluated from triaxial compression tests,

- The relation between undrained strength for active, plane strain and passive loading conditions (used for different portions of the failure surface),

- The effect of cyclic degradation of strength during cyclic loading, and

- Model error - how well the stability calculation gives the actual capacity, assuming strengths are known accurately.

Mean values and standard deviations were evaluated for all of these factors, using statistical techniques where adequate data existed (especially for active undrained strength) and subjective judgement otherwise (as for model error). Calculations then led to a probability-of-failure of 0.4×10^{-4}, which was deemed sufficiently small and in line with safety levels typical for good deterministic designs. (Note that this is the probability-of-failure given that the worst six-hour sea state actually occurs. The corresponding probability of failure during the lifetime of the structure is much less.) The dominating part of the total uncertainty came from uncertainty regarding the wave loading. Uncertainty in soil strength parameters was of little importance for the total uncertainty, except for that associated with cyclic degradation. Model uncertainties associated with the stability calculations were also found to contribute significantly to the total uncertainty. Nadim and Lacasse (1992) report a similar analysis for a jack-up platform, and also emphasize the importance of uncertainty in the loading.

The recent literature also contains reliability analyses for liners of landfills (Gilbert and Tang, 1995; Rowe and Fraser, 1995 - the latter using Monte Carlo simulations), for an anchored sheet-pile wall (Cherubini *et al*, 1992), and for stability of slopes (Christian et al, 1994). These analyses are characterized by careful at-tention to characterizing the uncertainty in the strength of soil, using statistical techniques plus some judgement as necessary, and to model errors. Uncertainty in the demand (i.e. loading) is typically less important in these studies. Some studies were primarily research to illustrate possible applications, while some have been used as input to actual decision-making.

I want to emphasize especially the importance of model errors: that is, potential errors in the deterministic calculations used to evaluate capacity for specified material properties. The uncertainties that can be associated with such calculations are often ignored or badly underestimated. By comparing predictions from a standard model for predicting flow through liners with actual measurements, Rowe and Fraser found it necessary to introduce a bias factor of 0.18 - that is, to account for a average error of more than a factor of 5! Lacasse and Nadim (1994) report upon the use of model tests to evaluate the mean and standard deviation for bearing capacity calculated by standard methods. Ronold and Bjerager (1992) suggest a method that may be used to evaluate model uncertainty from test programs. One warning that perhaps is obvious: A soil parameter (e.g. strength) must be chosen in the same way (e.g. the mean value) in a reliability analysis as when analyzing results of tests to calibrate the model used for calculations.

Model errors, in the broadest sense, are not just an affliction of probabilistic reliability analysis. Anytime a key feature of a problem is overlooked and not considered in decision-making, there is a model error. Morgenstern (1995) provides just such an example.

Event tree analysis

Fig. 3 shows part of an event tree used as part of a risk evaluation for an earth dam. The spillway for this dam was capable of passing large flood flows, but the unlined channel downstream of the spilling basin had experienced erosion. The concern was that headward erosion of this channel might lead to collapse of the stilling basin and thence to possible destruction of the spillway and/or erosion of the adjacent earthen embankment - and thence possible breaching of the dam. (This example is hypothetical although similar to a case from my files.) The "event" in this case is a flood flow of specified magnitude (actually, for a small range of flows centered on the probable maximum flood). Successive branch points account for the possibilities that:

- the downstream channel will erode back to the stilling basin, causing scour holes of different depths;

- the foundation for the stilling basin will collapse as a result of the scour hole;

- the collapse of the stilling basin will lead to undermining of the spillway and breaching of the dam;

- the collapse of the stilling basin will lead to erosion of the adjacent earthen embankment and breaching of the dam.

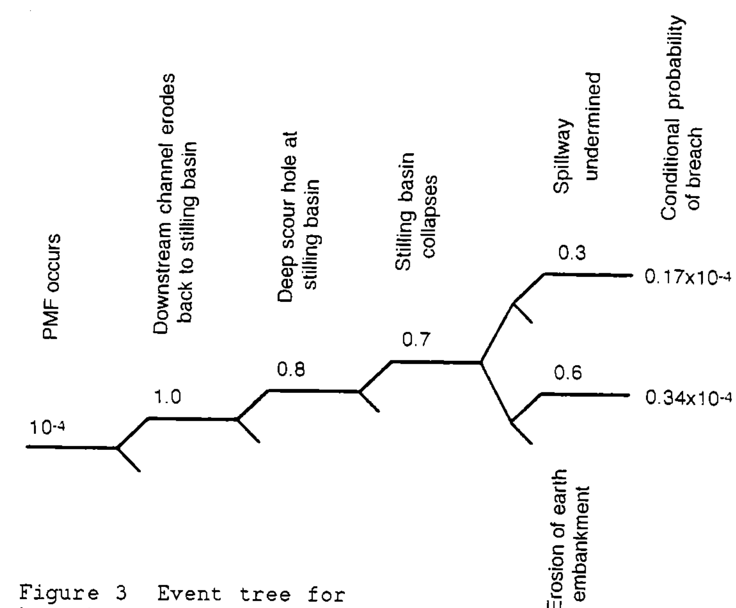


Figure 3 Event tree for breaching of earth dam

The probability of the flood flow was established from hydrologic studies. The probabilities at the branch points were estimated subjectively by engineers, based upon prior experience with erosion at the site, model hydraulic tests concerning erodibility of rip-rap, calculations concerning stability of a bulkhead wall at the toe of the stilling basin, and model hydraulic tests concerning scour-induced currents near the toe of the earthen embankment. Multiplying probabilities along the successive branch points gives the probability that this particular flood flow causes breaching of the dam. For this particular event tree, the probability of a failure is not much different from the probability of the initiating event. A set of similar event trees for different magnitudes of flood flows was used. For smaller and more likely flows, the probabilities of erosion, developing a scour pool, etc., are smaller. The sum of the probabilities from the set of event trees, which thus reflect the contributions from all magnitudes of flow, gave the overall probability of failure.

Used in this way, the event tree analysis is in effect a crude form of reliability analysis, where subjective judgments take the place of formal treatment of uncertainties. A similar example, concerning possible breach of a dike resulting from sinkhole collapse, is described by Vick and Bromwell (1989). The construction of an appropriate event tree is by itself an important exercise, which requires engineers to identify the sequence of events that might lead to a failure. Often the event tree is modified as the study progresses and certain sequences of events clearly become much less likely than others. Obviously, the numerical result is only as good as the subjective judgments, and engineers - who by nature typically are conservative - often need guidance in forming these judgments. Roberds (1990) has discussed methods for developing subjective probability assessments. One key is preparing tables for translating verbal statements concerning probability (e.g. "low") into numerical values.

Landslides

Reliability analysis is potentially appropriate for assessing the probability of failure of a particular slope - based upon geometry, shear strength and pore pressures. The major question always is whether or not the possible presence of weak, inclined strata has been taken into account properly.

There is an extensive literature concerning zoning for landslides and estimating possible slide-caused losses on a regional basis. Einstein (1988) presents an excellent, general summary of mapping techniques, especially those used in Europe. He deals with techniques used to assess probabilities of sliding, although the emphasis is upon general measures of likelihood rather than upon specific numerical measures. Fell (1994) has an excellent summary of methods for assessing probability of sliding and concerning allowable risks, together with numerical examples. Included in the methods is the use of historical data, relating risk of sliding to rainfall, and use of geomorphological and geotechnical data. Fell claims "...it is practical in many, if not all cases, to assign a probability to landsliding. In many cases it will be subjective, and approximate, but it is better than not trying."

One recent, interesting paper by Evans and Hungr (1993) assesses the rockfall hazard at the base of talus slopes. The analysis is primarily theoretical. One conclusion is that a strip development, 200 m along and 50 m from the margin of a talus slope, would be struck by a damaging boulder once in 95 years. Another interesting study (Bunce, 1994) deals with the risk to automobiles of being struck by rockfall onto highways.

Risks associated with earthquakes

Given the infrequent occurrence of earthquakes, it is not surprising that probabilistic methods play a major role for characterizing *demand;* e.g., the probability of exceeding some intensity of ground motion. There are debates as to whether a facility should be evaluated for a "500-year" or "2500-year" earthquake. But even with earthquake problems there is not universal acceptance of probabilistic methods. Ground shaking hazard maps, being prepared during 1996 by the U.S. Geological Survey as a basis for new building code provisions, are using probabilistic analysis for the Eastern United States but deterministic approaches for the West.

There is a school of thought holding that the design of critical projects (large dams, for example) should never be based upon probabilistic analysis of the earthquake threat, because there is so much uncertainty as to values for the parameters going into such analyses. This school argues that all such facilities should be designed for a *maximum credible earthquake*. However, the word *credible* itself implies some judgement involving uncertainty. In addition, methods used to evaluate ground motions caused by the maximum credible event generally are not based upon the absolutely largest values that have been measured. Thus the so-called deterministic methods for specifying earthquake motions for design involve some unquantified uncertainty.

Indeed there are uncertainties with regard to the accuracy of probabilistic ground shaking hazard analyses. Given the short history of direct experience with earthquakes in the United States, it would be overly bold to claim that a "2,500 year earthquake" can be evaluated precisely. However, this fact does not by itself mean that society should not decide to agree upon adoption of such events - calculated according to accepted rules - as a basis for design. At the same time, for some types of projects society may continue to opt to follow conservative, deterministic rules. The choice ultimately is made depending upon the costs of conservatism vs. the potential for consequences as a result of a failure.

Even when the ground shaking used for evaluation of a sit is the result of a probabilistic analysis, the subsequent assessment of *capacity* most commonly is deterministic. Liquefaction typically is evaluated using the well-known plot[2] in Fig. 4, which leads to a decision that a site either will or will not liquefy.

[2] An up-to-date version of this diagram has $(N_1)_{60}$ as the abscissa, as in Fig. 5.

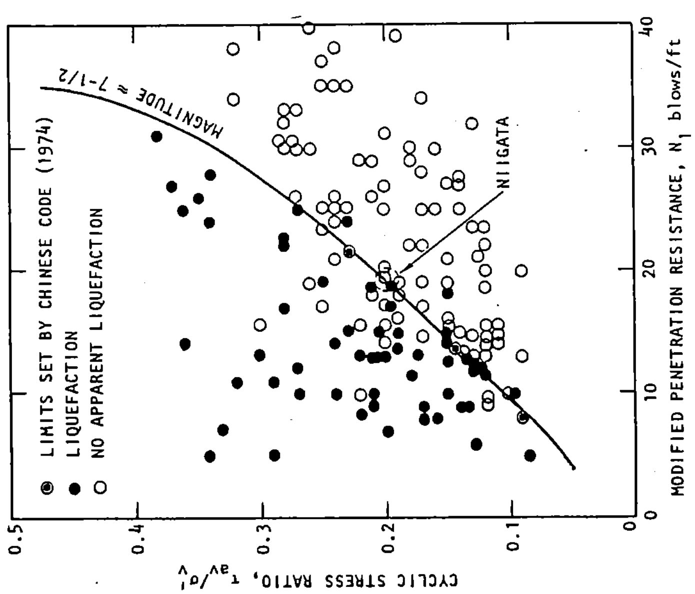


Figure 4 Empirical plot for evaluation of liquefaction. (From Liao *et al*, 1988, based upon work by H.B. Seed and others)

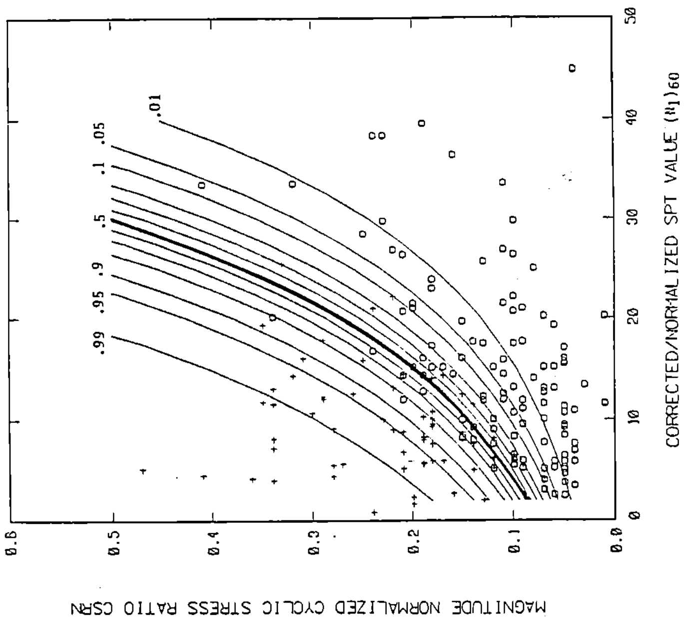


Figure 5 Probability of liquefaction, by regression of case studies. (Liao *et al*, 1988)

Liao *et al* (1988), using statistical methods (logis-tics) with many case histories where liquefaction did or did not occur, developed the plot in Fig. 5, which gives the probability of liquefaction for different combinations of shaking intensity and site resistance. It then is possible to multiply the probability of earthquake occurrence with the probability of liquefaction given the earthquake, so as to obtain an overall probability of liquefaction as a result of that particular level of shaking. The calculation can be repeated for different earthquakes with different probabilities of occurrence, and the individual results summed to give the overall probability of liquefaction.

There has been growing interest in this approach to evaluating the risk of liquefaction. In developing Fig. 5, no physical constraints were placed upon the combinations of shaking and blow count that might cause liquefaction. In addition, case studies where liquefaction did not occur are much scarcer in the literature than in actuality. Hence it seems likely that the probability of liquefaction is over-estimated for points lying well below the 50% probability curve. Loertscher and Youd, 1994) have applied logistics to study the influence of magnitude upon the probability of liquefaction. Fenton and Vanmarcke (1991) have analyzed implications of spatial variability at a site upon the site's susceptibility to liquefaction. Data are being accum-ulated concerning the consequences of liquefaction - e. g. the resulting lateral displacement or settlement - and this information potentially can be combined into an analysis leading to the probability of some damaging amount of movement. Thus probabilistic analysis of damage caused by liquefaction is a particularly fruitful area for research.

Environmental problems

I have no professional experience with environmental problems. Hence I can only offer a few observations based upon a superficial reading of a limited portion of the literature.

Although the word risk - in the form of information con-cerning the likelihood that a given degree of exposure to some substance will cause harmful consequences in humans - is commonly encountered when dealing with environmental problems, geoenvironmental engineers evaluating sites or designing waste repositories generally must follow very prescriptive rules with little opportunity for applying probabilistic methodologies. For example, the EPA scheme for ranking of hazardous sites uses a check-list of factors to establish a "likelihood of release value" that then goes into a simple equation and is combined with other similarly-evaluated factors. Thus probabilistic studies concerning hazardous sites are today mainly of value in pointing the way to more rational approaches that might appear in the future.

One major problem facing the analysis of pollutant movement through soils is the heterogeneity of typical soils. There is a long literature concerning the stochastic modeling of ground water flow (Thompson and Gelhar, 1990). In words from that paper:

> "... research...has been devoted to the development of more systematic and predictive modeling techniques which explicitly account for natural heterogeneity in a parsimonious statistical fashion.....these stochastic approaches are aimed at the quantitative description of bulk hydraulic behavior over large temporal and spatial scales while accounting for the influence of small-scale material variabilities."

In other words, the effect of randomness in local properties (such as permeability) upon the spreading of contaminant plumes is studied, and rules for accounting for these effects are developed. This particular paper uses multiple random-walk simulations.

These studies have assumed a medium that is "uniform" on a scale large compared to the local variability of soil properties. A paper published just prior to this Conference by researchers at the hosting university (Webb and Anderson, 1996) deals with the more difficult problem of large-scale heterogeneities - particularly those associated with braided stream channels. The practical inability to actually map such buried channels is accounted for by multiple simulations.

Here and there, there are isolated examples where probabilistic methods have been used to guide detailed choices during design of remediation measures. For, example, Massmann *et al* (1991) describe a study concerning a pumping scheme to extract contaminated ground water. The choice of a rate was optimized, using subjective judgments concerning the relative success of different pumping rates. Examples of reliability analysis applied to waste containment have been mentioned above.

ACCEPTABLE RISK

There is a considerable body of data that implicitly suggests acceptable risk. Fell (1994, quoting Reid, 1989) summarizes risk statistics for persons voluntarily or involuntarily exposed to various hazards, expressed as probability of death per person per year. They range from:

0.00014×10^{-3} for structural failure, through
0.009×10^{-3} for air travel and
0.3×10^{-3} for road accidents, to
1.9×10^{-3} for parachuting and
2.8×10^{-3} for deep sea fishing as an occupation.

For comparison, the average 30 year old male has, statistically, a chance of 10^{-3} of dying this year. It has been inferred that people, by their actions, implicitly accept a voluntary risk up to 10^{-3} and tolerate involuntary but recognized risks up to perhaps 10^{-5}. The tolerance for risks they suddenly discover or do not understand is lower yet. It is also well documented that risks that may affect a large number of people simultaneously (i. e air crashes) are less tolerable than risks of individual accidents.

Failure rates can be collected for classes of structures. For example the average annual failure rate of earthen dams, from all causes, is about 10^{-4}. By no means does this constitute an acceptable rate.

Another approach is to evaluate theoretically the risk of a common class of structures, such as steel-framed buildings. This was done when limit state codes were being developed for such buildings. The risk of failure, during the lifetime of a structure, implied by accepted designs was found to be on the order of 10^{-4}. (This is for any type of unsatisfactory behavior; the risk of a collapse would be less.)

The vexing question, of course, is: How can this information be used to establish allowable risks for specific projects? There is no general answer to this question. Fig. 6 reproduces a first attempt to assemble information to assist with such discussions and negotiations.[3] Relations of this general type have been developed in several countries. British Columbia Hydro, as part of an effort to review multi-hazard threats to dams and other facilities (Nielsen *et al*, 1994), have assembled the information shown in Fig. 7. The proposed criterion limits risk to any one individual to 10^{-4}/year, with smaller risk levels when potential multiple fatalities are involved. In addition, there is an maximum organizational financial risk for projects whose failure cost exceeds $100 million: The annual expected (i.e. best estimate) risk should not exceed $10,000/yr. BCHydro's approach to risk analysis for dams is

---

[3] Of all the figures with which my name is associated, this is perhaps the most often cited. I had a call about it as recently as the fall of 1995. Greg Baecher, from whom I originally borrowed the figure, laments that he seems best known for a figure he never published himself.

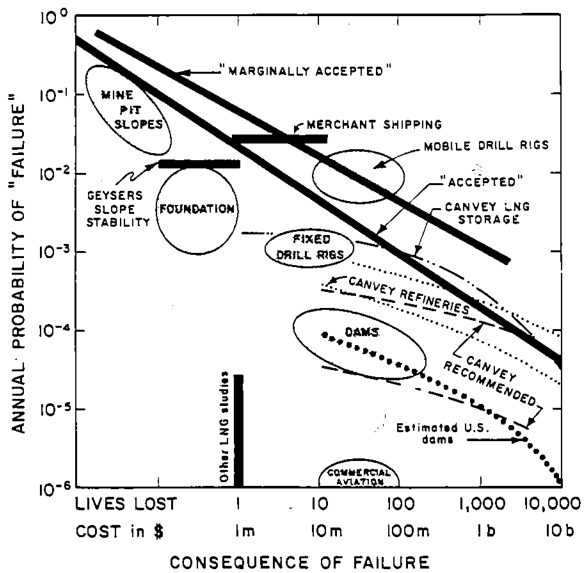


Figure 6 Risks for engineering projects (Whitman, 1984)

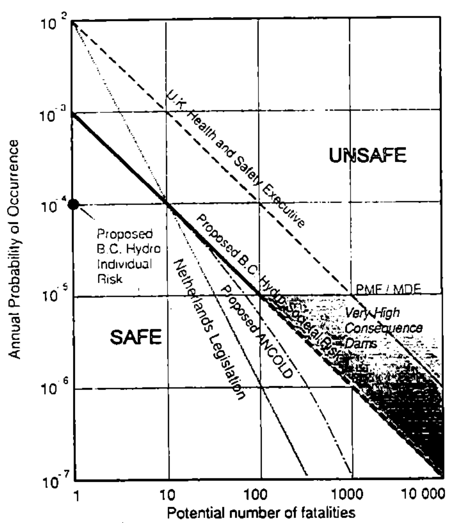


Figure 7 Various risk criteria (Courtesy of BCHydro)

discussed in a paper to this Conference by Vick and Stewart.

In a few problems, the risk calculated for an engineering solution may be compared with an absolute limit to risk. More commonly, quantified risk becomes a vehicle for communication between engineer and client or regulator - to express the degree of riskiness and to compare the relative risk among possible alternative solutions. In any project where geotechnical engineering deals with a significant portion of the effort (which will mean most important projects), it is vital that the geotechnical engineer be part of these communications - to understand how much geotechnical engineering solutions contribute to the overall risk and to explain how this risk might be altered. Obviously, in such discussions a geotechnical engineer must feel comfortable with probabilistic concepts.

Human errors

As mentioned above, there have been reliability analyses for classes of structures that have actually been built and are in service. It is typically found that the actual failure rates exceed predicted failure rates - perhaps by as much as two orders of magnitude. Further examination reveals that most of the failures are the result of human error; e. g. structures not built according to plans, materials not meeting specification, some loading not considered in the reliability analysis, etc. An obvious question then is: If failures result from oversights not considered in a reliability analysis, why perform such an analysis for judging adequacy of design? The answer is that engineers want to make sure that the probability of failure from things under their control is well less than the failure probability associated with things they cannot control. Not only is that in the best interest of the engineer, but for society as well.

CONSTRUCTION CONTROL

Controlling compaction using field sampling has long been a part of geotechnical engineering, and rules concerning sampling rates and criteria governing acceptability are to some extent influenced by statistical concepts. There continue to be new contributions motivated by new types of problems. For example, specifications typically require that the overall hydraulic conductivity of a compacted landfill liner be less than some specified limit. Benson, Zhai and Rashad (1994) have developed a procedure for selecting the number of samples that must be tested to ensure a high probability that this criterion is met, and Benson, Zhai and Wang (1994) demonstrate how borrow material can be evaluated for potential as a compacted liner. Quality control for membrane liners has also become a geotechnical engineering problem.

Morgenstern emphasizes, in his Casagrande Lecture, the importance of the *observational method*. For any who are not familiar with this concept, it implies adjusting construction procedures and details depending upon observations and measurements made as construction proceeds. Practicioners of this method understand the importance of identifying, in advance of construction, the range of possible soil conditions that may be encountered - and of having plans to cope with possible eventualities. Just describing the observational method suggests opportunities for using probabilistic concepts and methods. If the uncertainties and risks have been quantified before construction begins, then updated information obtained during construction can be used to revise risk estimates and to guide decisions made during construction.

WHEN TO USE PROBABILISTIC ANALYSIS

Ralph Peck participated in the workshop concerning Probabilistic Methods in Geotechnical Engineering (NRC, 1995) - as usual, a brave soul since almost all others attending were certified probabilists. Here is how he summarized the state-of-the-profession:

> We see geotechnical engineering as developing into two somewhat different entities: one part still dealing with traditional problems such as foundations, dams, and slope stability, and another part dealing with earthquake problems; natural slopes; and, most recently, environmental geotechnics. Practitioners in the first part have not readily adopted reliability theory, largely because the traditional methods have been generally successful, and engineers are comfortable with them. In contrast, practitioners in environmental geotechnics and to some extent in offshore engineering require newer, more stringent assessments of reliability that call for a different approach. Therefore, we may expect reliability methods to be adopted increasingly rapidly in these areas as confidence is developed. It is not surprising that those engineers working in environmental and offshore problems should be more receptive to new approaches, and it should not be surprising that there may be spillback into the more traditional areas.

It is difficult to improve upon this characterization of the present status of utilization of probabilistic methods. I would add that studies for evaluation and remediation of existing facilities - such as dams - originally designed by traditional approaches is a fertile field for risk evaluation..

I do want to suggest an alternate classification that look's to the future as well as the present. For the sake of simplification, I will divide geotechnical engineering problems into two broad categories.

1. Those where the client relies upon codes, regulations and "accepted practice" to ensure that he receives a satisfactory product. This category includes the vast majority of "routine" projects.

2. Those where the client, and/or a regulator, is actively in a discussion of potential risks and ultimately assumes whatever at least most of whatever risk is implied by the final choice of design. Such projects are characterized by either the impossibility of eliminating risks completely or by a very high cost the reduce risks to an insignificant level. Thus it is in the interest of the client to become actively engaged in decision-making. Projects of this type are less common, and typically are large in scale or involve uncommon types of buildings or facilities, or both. However, there is no reason why probabilistic methods cannot be utilized in traditional problems - if a client believes that doing so can be of potential benefit.

When involved in the first type of project, an engineer is unlikely to make use of quantified risk analysis - or of probabilistic thinking or statistics, except possibly in connection with planning details of site exploration and characterization or during construction control. However, probabilistic analysis can be of use when developing requirements of codes and regulations.

Involvement in the second type of project will certainly require an engineer to engage in probabilistic thinking. In some instances, acceptable risk may be specified numerically, and the engineer must choose a design and demonstrate that the specification is met. Even here the client, or at least a regulator, will be involved in a significant way, since seldom will there be clear, accepted procedures covering all aspects of the evaluation of risk. More likely, a number of design schemes will be discussed by the engineer and client, until the client (and likely the cognizant insurance company) is satisfied that there is an acceptable balance between cost and risk. Evaluation of risk, whether in quantitative terms or merely by words, becomes an important means of communication between client and engineer.

Role of probabilistic methodologies in code development

The NRC report suggests that probabilistic methods can be useful in the development of codes. Indeed, such methods have been used in the process of developing the limit state codes now common for the structural portions of buildings. During the past decade, there has been an effort to standardize codes within the European Community, and to bring geotechnics codes in line with the reliability-based approach to structural codes. These codes emphasize "limit states" and partial factors (akin to safety factors) applied to both loads and resistances. The approach often is referred to as Limit States Design (LSD).

In Canada, there has been considerable discussion re LSD. Geotechnical News for March 1995 has a piece entitled "Limit States Design on Trial", reporting upon a mock trial with arguments against and for LSD. The unanimous opinion of the Judges.....constitutes a recommendation to the profession and reads:

> "The Working Stress Design (WSD) approach is still the common and accepted touchstone for most geotechnical engineers. It marries experience to judgement. However, in itself, it does not wholly fit the need for a design approach consistent for both structural and geotechnical engineers. The LSD approach, when utilized in its broadest and most practical sense, namely using factored resistance rather than using partially factored strength parameters, goes far to meet this need and with time and accumulated experience by practicioners in both structures and foundations will improve its quality of practical applications.
>
> The Court concluded that provided the Factored Resistance Approach is used, the case against LSD was NOT proved."

In the United States, where for the most part geotechnical engineers are deeply suspicious of codes, there as yet has been relatively little debate on this matter.

A reading of drafts of the geotechnics portion of the Eurocode actually reveals quite limited emphasis upon probabilistic methods and thinking. There is a section describing how partial resistance factors for pile design are selected based upon a pile-load test program and on the pile type adopted for design. (This section is discussed in NRC, 1995.)

Eurocode also introduces the concept of a *characteristic value,* described as follows:

> Characteristic values shall be selected with the intention that the probability of a more unfavorable value governing the occurrence of a limit state is not greater than 5%. For parameters for which the values governing field behavior are well established with little uncertainty, the characteristic value may be taken as the best estimate of the value in the field. Where there is greater uncertainty, the characteristic value is more conservative.....It might sometimes be helpful to use statistical methods. However, it is emphasized that this will rarely lead directly to characteristic values, since these depend upon an assessment of the field situation.

Safety factors (actually material resistance factors) are then to be applied to characteristic values. Eurocode suggests a factor of 1.2 to 1.25 for friction angle and 1.5 to 1.8 for cohesion - but provides exceptions and "outs". There are some complicated concepts here that require considerable thought.

The draft chapter on retaining walls contains primarily a lot of good, well-accepted advice concerning good design. Here and there is specific guidance re partial safety factors and load conditions (such as location of a water table). Simpson (1992) has written a very thoughtful critique concerning the use of partial factors for the design of retaining walls. In his conclusion he states:

> "At present, the best available tool.....is engineering judgement; there is danger that formal procedures, if they are sufficiently simple to be prescribed, will jettison valuable information. However, it is sensible to provide objective checks on judgement whenever possible..........the best way to combine these requirements is to make the designer directly responsible for design values adopted in the calculations. In addition, codes should specify how design values should be checked against values derived using characteristic values and partial factors. Both the characteristic and design values should be defined in terms of the expected probability that the values will occur in the field situation in such a way as to govern the occurrence of limit states. Numerical analysis of soil test results alone will often be an inadequate basis for selection of these values."

No matter how one feels concerning the wisdom of codes governing geotechnical practice, these are important ideas and questions. I hope that they will be discussed vigorously at this Conference.

FURTHERING USE OF PROBABILISTIC METHODOLOGIES

The NRC report's principal recommendation is that "education of new geotechnical engineers, as well as practicing engineers, in probabilistic methods should be undertaken". A retired professor cannot possibly resist commenting on such a matter.

In the MIT Department of Civil and Environmental Engineering we have for years required a subject in probability theory, taught by faculty from the Department. This requirement has proved less than a great success, primarily because it is rarely followed-up by applications in subsequent subjects, whether they be oriented to engineering science or to design. Actually, the same problem tends to exist with much of the material taught in other beginning subjects in engineering science. Observations such as these are stimulating a rethinking of engineering curricula, with a gradually increasing trend toward teaching material as it is needed for some application. This stimulates student interest in the basic material, but means that students may not get as thorough a grasp of basics and appreciate the potential application of this material beyond the particular context within which it is taught. In other words, the issue of teaching probabilistic methods to geotechnical engineers will inevitably be caught up in an ongoing debate concerning the pedagogy of teaching.

Morgenstern argues that it is the principles of *risk management* that should be taught at a fundamental level and then illustrated through applications. In his words, *risk management* relies on rational analyses and involves situation appraisal, problem and potential problem analyses and decision analyses. I agree with this perspective. I would prefer to see a subject covering the practice of risk management taught early in a curriculum, rather than a subject in probability theory. Of course, some basic concepts concerning probability must be incorporated into this type of subject.

The NRC report also urges that major geotechnical projects should involve a probability expert as part of the project team to provide opportunities for close interaction between that expert and the other team members. Following on from the thoughts in the last paragraph, I would recommend rather that an expert in risk management should be included as a member of the team. Of course this expert should be well-grounded in probability theory and its application.

Having made these arguments, I certainly do agree that

geotechnical engineers as a whole should become better versed in the important basic concepts of probability theory, reliability theory and risk analysis. The NRC report contains, in an appendix, an excellent primer concerning these matters. In my Terzaghi Lecture (Whitman, 1984), I urged the need for examples showing how risk can be quantified and used for decision-making. As has been noted above, a number of such examples has now been published. This is a good start, but more are needed - especially those illustrating clearly the role played by probabilistic methods in actual decision-making. The ready availability of examples will go a long way in piquing the interest of practicing engineers.

Perhaps the key question is: What can be done to interest clients in designs based upon probabilistic thinking rather than traditional approaches? I believe a client will always become interested in different approaches if it appears there is potential for financial benefit, for satisfying regulators, or even for public relations purposes. A client may receive such stimuli from various sources, but a key stimulator is the engineer for the project. A phrase in the middle of Peck's summary holds the key: Traditional methods are used ".....*because engineers are comfortable with them*." Clearly there will continue to be projects that are best engineered using traditional approaches. A challenge for this conference is to identify types of projects where the "spillback" predicted by Peck should begin to occur in the near future.

FINAL REMARKS

Traditional engineers worry rightly that too much emphasis upon analysis might drive out engineering judgement and lead to unsatisfactory designs. Wu *et al* (1975) long ago provided a thoughtful example of how a probabilistic analysis might go wrong: A slope that contains a plane of weakness but is assumed to be homogeneous when conducting a statistical analysis of data for strength. This potential problem exists with any analysis, whether deterministic or probabilistic, until it is well-calibrated to experience. This is why I have emphasized the importance of model errors. Thoughtful probabilists always emphasize that probabilistic methods do not replace traditional methods. Rather, probabilistic methods are tools that can effectively supplement traditional methods for engineering geotechnical projects, providing better insights into the uncertainties and providing an improved basis for interaction between engineers and decision-makers.

All this is particularly true concerning a methodology especially popular at the moment: reliability analysis. Christian *et al* (1994) offer some valuable insights concerning this tool:

> "Reliability analysis is especially useful in establishing design values for factor-of-safety representing consistent risks for different types of failures....The most effective applications of probabilistic methods are those involving relative probabilities of failure or illuminating the effects of uncertainties in the parameters."

Similar perspectives are needed concerning all of the existing and yet-to-be-developed probabilistic methods.

I hope I have helped you appreciate, in general terms, how the quantification of uncertainty can play a useful and important role in engineering projects. I also hope I have stimulated you to learn about the available examples of practical application. With Peck, I agree that probabilistic methods are now playing important roles in a number of engineering problems, and that there will be increasing spillover into problems now engineered by traditional methods. The continued challenge is to recognize problems in which probabilistic thinking can contribute effectively to the engineering solution - while at the same time not trying to force these new approaches into problems best engineered with traditional methods and viewpoints.

REFERENCES

Baecher, G.B., 1979: "Analyzing exploration strategies," *Site Characterization and Exploration,* C.H. Dowding, ed., ASCE.

Benson, C.H., H. Zhai and S. Rashad 1994: "Statistical sample size for construction of soil liners," *J. Geotechnical Engineering,* ASCE, 120(10): 1704-1724.

Benson, C.H., H. Zhai and X. Wang, 1994: "Estimating hydraulic conductivity of Clay liners," *J. Geotechnical Engineering,* ASCE, 120(2): 366-387.

Bunce, C.M., 1994: *Risk analysis for rock fall on highways,* M.S. thesis, Dept. Civil Engineering, University of Alberta, Edmonton, Alberta.

Casagrande, A: 1965: "The role of 'Calculated Risk' in earthwork and foundation engineering," *J. Soil Mechanics Division,* ASCE, 91(4): 1-40.

Cherubini, C., A Garrasi and C. Petrolla, 1992: "The reliability of an anchored sheet-pile wall embedded in cohesionless soil," *Canadian Geotechnical Journal,* 29(3): 426-435.

Christian, J.T., C.C. Ladd and G.B. Baecher, 1994: "Reliability applied to slope stability analysis," *J. Geotechnical Engineering,* ASCE, 120(12): 2180-2207.

Einstein, H.H., 1988: "Landslide risk evaluation procedure," *Proc. 5th Int. Sym.Landslides,* C. Bonnard, ed.; A.A. Balkema, 1075-1090.

Evans, S.G. and O. Hungr, 1993: "The assessment of rockfall hazard at the base of talus slopes," *Canadian Geotechnical Journal,* 30(4): 620-636.

Fell, R., 1994: "Landslide risk assessment and allowable risk," *Canadian Geotechnical Journal,* 31(2): 261-272.

Fenton, G.A. and E.H. Vanmarcke, 1991: "Spatial variation in liquefaction risk assessment," Geotechnical Engineering Congress, *Geotechnical Special Pub. 27,* ASCE, 1:595-607.

Geotechnical News, March 1995, 10-13.

Gilbert, R.B. and W.H. Tang, 1995: "Reliability-based design for waste containment systems", Geoenvironment 2000, *ASCE Geotechnical Spec. Pub. 46,* 499-513.

Halim, I.S. and W.H. Tang, 1993: "Site exploration strategy for geologic anomaly characterization," *J. Geotechnical Engineering,* ASCE, 119(2): 195-213.

Lacasse, S. and F. Nadim, 1994: "Reliability issues and future challenges in geotechnical engineering for off shore structures", *Proc. 7th Int. Conf. Behaviour of Offshore Structures,* Mass. Institute of Technology.

Loertscher, T.W. and T.L. Youd, 1994: "Magnitude scaling factors for analysis of liquefaction," *Proc. 5th U.S.-Japan Workshop on Earthquake Resistent Design of Lifeline Facilities and Countermeasures against Liquefaction,* Nat. Ctr.Earthquake Engineering Research Technical Report NCEER-94-0026, 703-715.

Liao, S.C., D. Veneziano and R.V. Whitman, 1988: "Regression models for evaluating liquefaction probability," *J.Geotechnical Engineering,* ASCE,114(4): 389-411.

Massmann, J., R.A. Freeze, L. Smith, T. Sterling and B. James, 1991: "Hydrogeologic decision analysis: 2. Applications to ground water contamination," *Ground Water,* 29(4): 536-548.

Morgenstern, N., 1995: "Managing risk in geotechnical engineering," *Proc. 10th Pan American Conference on Soil Mechanics and Foundation Engineering,* Vol. 4, in press.

Nadim, F. and S. Lacasse, 1992: "Probabilistic bearing capacity analysis of jack-up structures", *Canadian Geotechnical Journal,* 31(4): 580-588.

Nielsen, N., D. Hartford and T. MacDonald, 1994: "Selection of tolerable risk criteria in dam safety decision making," *Proc. Canadian Dam Safety Conference,* BiTech Publishers, Vancouver, BC.

Nobre, M.M. and J.F. Sykes, 1992: "Application of Bayesian Kriging to subsurface characterization," *Canadian*

Geotechnical Journal, 29(4): 589-598.

Reid, S.G., 1989: *Risk assessment research report R591,* School of Civil and Mining Engineering, Univ. Sydney, Australia, 1-46.

Roberds, W.J. 1990: "Methods for developing defensible subjective probability assessments," *Transportation Research Record,* 1288: 183-190.

Ronold, K.O. and P. Bjerager, 1992: "Model uncertainty representation in geotechnical reliability analyses", *J. Geotechnical Engineering,* ASCE, 118 (3): 363-376.

Ronold, K.O. and S. Bysveen, 1992: "Probabilistic stability analysis for deep-water foundation", *J. Geotechnical Engineering,* ASCE, 118 (3): 394-405.

Rowe, R.K. and M.J. Fraser, 1995: "Effect of uncertainty in the assessment of the potential impact of waste disposal facilities", Geoenvironment 2000, *ASCE Geotechnical Spec. Pub. 46,* 270-284.

Simpson, B. 1992: "Partial factors of safety for design of retaining walls," *Geotechnique,* 42(1): 131-136.

Tang, W.H. 1979: "Probabilistic evaluation of penetration resistance," *J. Geotechnical Engineering,* ASCE, 105(10): 1173-1191.

Thompson, A.F.B. and L.W.Gelhar, 1990: "Numerical simulation of solute transport in three-dimensional, randomly heterogeneous porous media," *Water Resources Research,* 26(10): 2541-2562.

Vanmarcke, E.H., 1977: "Reliability of earth slopes," *J. Geotechnical Engineering,* ASCE, 103(11): 1227-1246

Vanmarcke, E.H., 1983: *Random Fields, Analysis and Synthesis,* MIT Press, Cambridge MA, 382 pp.

Vick, S.G., 1992: "Risk in geotechnical practice", *Geotechnique and Natural Hazards,* BiTech Publishers, Richmond, British Columbia, 41-62.

Vick, S.G. and L.F. Bromwell, 1989: "Risk analysis for dam design in karst", *J. Geotechnical Engineering,* ASCE, 115(6): 819-835.

Vick, S.G. and R.A. Stewart, 1996: "Risk analysis in dam safety practice", This Conference.

Webb, E.K. and M.P. Anderson, 1996: "Simulation of preferential flow in three-dimensional, heterogeneous fields with realistic internal architecture," *Water Resources Research,* 32(3): 533-545.

Whitman, R.V. 1984: "Evaluating calculated risk in geotechnical engineering," *J. Geotechnical Engineering,* ASCE, 110(2): 145-188.

Wu, T.H., W.B. Thayer and S.S. Lin. 1975: "Stability of an embankment on clay," *J. Geotechnical Engineering,* ASCE, 101(9): 913-932.

Wu, T.H., O. Kjekstad, I.M. Lee and S. Lacasse. 1989: "Reliability analysis of foundation stability for gravity platforms in the North Sea," *Canadian Geotechnical Journal,* 26: 359-368.

ENGINEERING JUDGMENT IN THE EVOLUTION FROM DETERMINISTIC TO RELIABILITY-BASED FOUNDATION DESIGN

Fred H. Kulhawy[1], F. ASCE and Kok Kwang Phoon[2]

ABSTRACT: Engineering judgment has always played a dominant role in geotechnical design and construction. Until earlier this century, most of this judgment was based on experience and precedents. The role of judgment in geotechnical practice has undergone significant changes over the past 50 years as a result of theoretical, experimental, and field developments in soil mechanics and, more recently, in reliability theory. A clarification of this latter change specifically is needed to avoid misunderstanding and misuse of the new reliability-based design (RBD) codes. This paper first provides a historical perspective of the traditional factor of safety design approach. The fundamental importance of limit state design to RBD then is emphasized. Finally, an overview of RBD is presented, and the proper application of this new design approach is discussed, with an example given of the ultimate limit state design of drilled shafts under undrained uplift loading. Judgment issues from traditional approaches through RBD are interwoven where appropriate.

INTRODUCTION

Almost all engineers would agree that engineering judgment is indispensable to the successful practice of engineering. Since antiquity, engineering judgment has played a dominant role in geotechnical design and construction, although most of the early judgment was based on experience and precedents. A major change took place in engineering practice when scientific principles, such as stress analysis, were incorporated systematically into the design process. In geotechnical engineering, in particular, significant advances were made following World War II, primarily because of extensive theoretical, experimental, and field research. The advent of powerful and inexpensive computers in the last two decades has helped to provide further impetus to the expansion and adoption of theoretical analyses in geotechnical engineering

---

1 - Prof., School of Civil & Environ. Eng., Hollister Hall, Cornell Univ., Ithaca, NY 14853-3501

2 - Lect., Dept. of Civil Eng., Natl. Univ. of Singapore, 10 Kent Ridge Crescent, Singapore 0511

practice. The role of engineering judgment has changed as a result of these developments, but the nature of this change often has been overlooked in the enthusiastic pursuit of more sophisticated analyses. Much has been written by notable engineers to highlight the danger of using theory indiscriminately, particularly in geotechnical engineering (e.g., Dunnicliff & Deere 1984, Focht 1994). For example, engineering judgment still is needed (and likely always will be!) in site characterization, selection of appropriate soil/rock parameters and methods of analysis, and critical evaluation of the results of analyses, measurements, and observations. The importance of engineering judgment clearly has not diminished with the growth of theory and computational tools. However, its role has become more focused on those design aspects that remained outside the scope of theoretical analyses.

At present, another significant change in engineering practice is taking place. Much of the impetus for this innovation arose from the widespread rethinking of structural safety concepts that was brought about by the boom in post-World War II construction (e.g., Freudenthal 1947, Pugsley 1955). Traditional deterministic design codes gradually are being phased out in favor of reliability-based design (RBD) codes that can provide a more consistent assurance of safety based on probabilistic analyses. Since the mid-1970s, a considerable number of these new design codes have been put into practice for routine structural design, for example, in the United Kingdom in 1972 (BSI-CP110), in Canada in 1974 (CSA-S136), in Denmark in 1978 (NKB-36), and in the U.S. in 1983 for concrete (ACI) and in 1986 for steel (AISC). In geotechnical engineering, a number of RBD codes also have been proposed recently for trial use (e.g., Barker et al. 1991, Berger & Goble 1992, Phoon et al. 1995).

The impact of these developments on the role of engineering judgment might be analogous to that brought about by the introduction of scientific principles into engineering practice. In this continuing evolution, it must be realized that RBD is just another tool, but it is different from traditional deterministic design, even though the code equations from both methods have the same "look-and-feel". These differences can lead to misunderstanding and misuse of the new RBD codes. For these reasons, it is necessary to: (a) clarify how engineering judgment can be used properly so that it is compatible with RBD, and (b) identify those geotechnical safety aspects that are not amenable to probabilistic analysis. In this paper, an overview is given first of the traditional geotechnical design approach from the perspective of safety control. The philosophy of limit state design then is presented as the underlying framework for RBD. Finally, the basic principles of RBD are reviewed, and the proper application of this new design approach is described with an example of the ultimate limit state design of drilled shafts under undrained uplift loading. As noted prominently in the paper title, engineering judgment is interwoven throughout.

TRADITIONAL GEOTECHNICAL DESIGN PRACTICE

The presence of uncertainties and their significance in relation to design has long been appreciated (e.g., Casagrande 1965). The engineer recognizes, explicitly or otherwise, that there is always a chance of not achieving the design objective, which is to ensure that the system performs satisfactorily within a specified period of time. Traditionally, the geotechnical engineer relies primarily on factors of safety at the design stage to reduce the risk of potential adverse performance (collapse, excessive deformations, etc.). Factors of safety between 2 to 3 generally are considered to be adequate in most foundation design (e.g., Focht & O'Neill 1985). However, these values can be misleading because, too often, factors of safety are recommended without reference to any other aspects of the design computational process, such as the loads and their evaluation, method of analysis (i.e., design equation), method of property evaluation (i.e., how to select the undrained shear strength), and so on. Other important considerations that affect the factor of safety include variations in the loads and material strengths, inaccuracies in the design equations, errors arising from poorly supervised construction, possible changes in the function of the structure from the original intent, unrecognized loads, and unforeseen in-situ conditions. The manner in which these background factors are listed should not be construed as suggesting that the engineer actually goes through the process of considering each of these factors separately and in explicit detail. The assessment of the traditional factor of safety is essentially subjective, requiring only global appreciation of the above factors against the backdrop of previous experience.

The sole reliance on engineering judgment to assess the factor of safety can lead to numerous inconsistencies. First, the traditional factor of safety suffers from a major flaw in that it is not unique. Depending on its definition, the factor of safety can vary significantly over a wide range, as shown in Table 1 for illustration. The problem examined in Table 1 is to compute the design capacity of a straight-sided drilled shaft in clay, 1.5 m in diameter and 1.5 m deep, with an average side resistance along the shaft equal to 36 kN/m^2 and a potential tip suction of 0.5 atmosphere operating during undrained transient live loading. Five possible design assumptions are included. The first applies the factor of safety (FS) uniformly to the sum of the side, tip, and weight components; the second applies the FS only to the side and tip components; the third is like the first, but disregards the tip; the fourth is like the second, but disregards the tip; and the fifth is ultra-conservative, considering only the weight. It is clear from Table 1 that a particular factor of safety is meaningful only with respect to a given design assumption and equation.

Another significant source of ambiguity lies in the relationship between the factor of safety and the underlying level of risk. A larger factor of safety does not necessarily imply a smaller level of risk, because its effect can be negated by the presence of larger uncertainties in the design environment. In addition, the effect of the factor of safety on the underlying risk level also is dependent on the extent of conservatism in the selected design models and design parameters.

In a broad sense, these issues generally are appreciated by most engineers. They can exert additional influences on the engineer's choice of the factor of safety but, in the absence of a theoretical framework, it is not likely that the risk of adverse performance can be reduced to a desired level consistently. Therefore, the main weakness in traditional practice, where assurance of safety is concerned, can be attributed to the lack

TABLE 1. Design Capacity Example (Kulhawy 1984, p. 395)

Design Assumption	Design Equation	Q_{ud} (kN) for FS = 3	Q_u / Q_{ud} ("actual" FS)
1	$Q_{ud} = (Q_{su} + Q_{tu} + W) / FS$	170.7	3.0
2	$Q_{ud} - W = (Q_{su} + Q_{tu}) / FS$	214.2	2.4
3	$Q_{ud} = (Q_{su} + W) / FS$	108.9	4.7
4	$Q_{ud} - W = Q_{su} / FS$	152.4	3.4
5	$Q_{ud} = W / FS$	21.8	23.5

Note: Q_{su} = side resistance = 261.8 kN, Q_{tu} = tip resistance = 184.4 kN, W = shaft weight = 65.3 kN, Q_{ud} = design uplift capacity, FS = factor of safety, Q_u = available uplift capacity = $Q_{su} + Q_{tu} + W$ = 511.6 kN

of clarity in the relationship between the method (factor of safety) and the objective (reduce design risk). To address this problem appropriately, an essential first step is to establish the design process on a more logical basis, known as limit state design.

PHILOSOPHY OF LIMIT STATE DESIGN

The original concept of limit state design refers to a design philosophy that entails the following three basic requirements: (a) identify all potential failure modes or limit states, (b) apply separate checks on each limit state, and (c) show that the occurrence of each limit state is sufficiently improbable. Conceptually, limit state design is not new. It is just a logical formalization of the traditional design approach that would help to facilitate the explicit recognition and treatment of engineering risks. In recent years, the rapid development of RBD has tended to overshadow the fundamental role of limit state design. Much attention has been focused on the consistent evaluation of safety margins using advanced probabilistic techniques (e.g., MacGregor 1989). Although the achievement of consistent safety margins is a very desirable goal, it should not be overemphasized to the extent that the importance of the principles underlying limit state design become diminished.

This fundamental role of limit state design is particularly true for geotechnical engineering. The first step in limit state design, which involves the proper identification of potential foundation failure modes, is not always a trivial task (Mortensen 1983). This effort generally requires an appreciation of the interaction between the geologic environment, loading characteristics, and foundation response. Useful generalizations on which limit states are likely to dominate in typical foundation design situations are certainly possible, as in the case of structural design. The role of the geotechnical engineer in making adjustments to these generalizations on the basis of site-specific information is, however, indispensable as well. The need for sound engineering judgment in the selection of potential limit states is greater in foundation design than in structural design because in-situ conditions must be dealt with "as is" and might contain geologic "surprises". The danger of downplaying this aspect of limit state design in the fervor toward improving the computation and evaluation of safety margins in design can not be overemphasized.

The second step in limit state design is to check if any of the selected limit states has been violated. To accomplish this step, it is necessary to use a model that can predict the performance of the system from some measured parameters. In geotechnical engineering, this is not a straightforward task. Consider Fig. 1, which is the essence of any type of prediction, geotechnical or otherwise. At one end of the process is the forcing function, which normally consists of loads in conventional foundation engineering. At the other end is the system response, which would be the prediction in an analysis or design situation. Between the forcing function (load) and the system response (prediction) is the model invoked to describe the system behavior, coupled with the properties needed for this particular model. Contrary to popular belief, the quality of geotechnical prediction does not necessarily increase with the level of sophistication in the model (Kulhawy 1992). A more important criterion for the quality of geotechnical prediction is whether the model and property are calibrated together for a specific load and subsequent prediction (Kulhawy 1992, 1994). Reasonable predictions often can be achieved using simple models, even though the type of behavior to be predicted is nominally beyond the capability of the models, as long as there are sufficient data to calibrate these models empirically. However, these models then would be restricted to the specific range of conditions in the calibration process. Extrapolation beyond these conditions potentially can lead to erroneous predictions. Ideally, empirical calibration of this type should be applied judiciously by avoiding the use of overly simplistic models. Common examples of such an oversimplification are the sets of extensive correlations between the standard penetration test N-value and practically all types of geotechnical design parameters, as well as several design conditions such as footing settlement and bearing capacity. Although they lack generality, simple models will remain in use for quite some time because of our professional heritage that is replete with, and built upon, empirical correlations. The role of the geotechnical engineer in appreciating the complexities of soil behavior and recognizing the inherent limitations in the simplified models is clearly of considerable importance. The amount of attention paid to the evaluation of safety margins is essentially of little consequence if the engineer assesses the soil properties incorrectly or selects an inappropriate model for design.

Third, the occurrence of each limit state must be shown to be sufficiently impro-

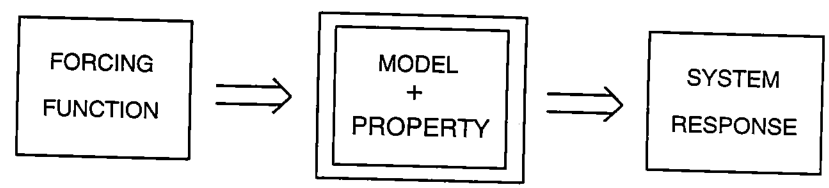


FIG. 1. Components of Geotechnical Prediction (Kulhawy 1994, p. 210)

bable. The philosophy of limit state design does not entail a preferred method of ensuring safety. Since all engineering quantities (e.g., loads, strengths) are inherently uncertain to some extent, a logical approach is to formulate the above problem in the language of probability. The mathematical formalization of this aspect of limit state design using probabilistic methods constitutes the main thrust of RBD. Aside from probabilistic methods, less formal methods of ensuring safety, such as the partial factors of safety method (e.g., Danish Geotechnical Institute 1985; Technical Committee on Foundations 1992), have also been used within the framework of limit state design.

In summary, the control of safety in geotechnical design is distributed among more than one aspect of the design process. Although it is important to consider the effect of uncertainties in loads and strengths on the safety margins, it is nonetheless only one aspect of the problem of ensuring sufficient safety in the design. The other two aspects, identification of potential failure modes and the methodology of making geotechnical predictions, can be of paramount importance, although they may be less amenable to theoretical analyses.

RELIABILITY-BASED DESIGN

Overview of Reliability Theory

The principal difference between RBD and the traditional design approach lies in the application of reliability theory, which allows uncertainties to be quantified and manipulated consistently in a manner that is free from self-contradiction. A simple application of reliability theory is shown in Fig. 2. Uncertain design quantities, such as the load (F) and the capacity (Q), are modeled as random variables, while design risk is quantified by the probability of failure (p_f). The basic reliability problem is to evaluate p_f from some pertinent statistics of F and Q, which typically include the mean (m_F or m_Q) and the standard deviation (s_F or s_Q). Note that the standard deviation provides a quantitative measure of the magnitude of uncertainty about the mean.

A simple closed-form solution for p_f is available if Q and F are both normally distributed. For this condition, the safety margin (Q - F = M) also is normally distributed with the following mean (m_M) and standard deviation (s_M) (e.g., Melchers 1987):

$$m_M = m_Q - m_F \tag{1a}$$

$$s_M^2 = s_Q^2 + s_F^2 \tag{1b}$$

Once the probability distribution of M is known, the probability of failure (p_f) can be evaluated as (e.g., Melchers 1987):

$$p_f = \mathrm{Prob}(Q < F) = \mathrm{Prob}(Q - F < 0) = \mathrm{Prob}(M < 0) = \Phi(-m_M / s_M) \tag{2}$$

in which $\mathrm{Prob}(\cdot)$ = probability of an event and $\Phi(\cdot)$ = standard normal cumulative

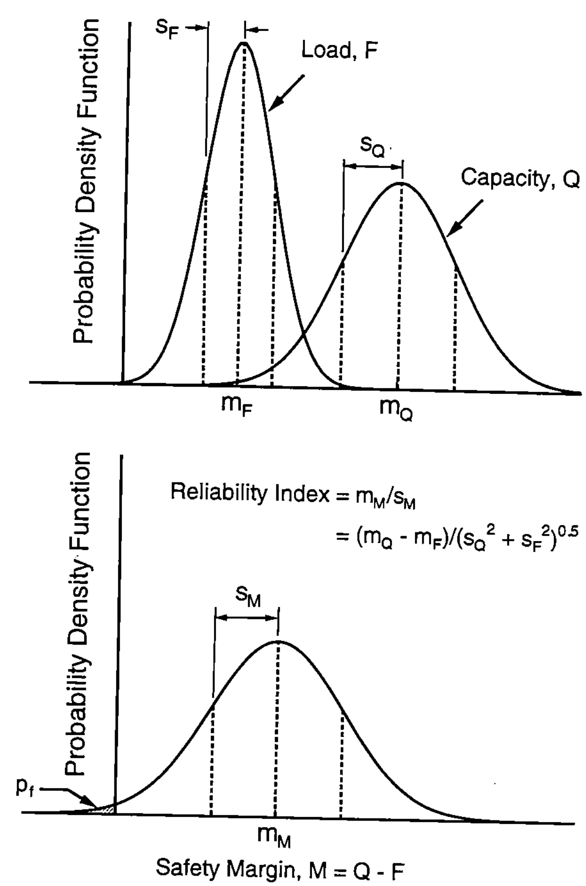


FIG. 2. Reliability Assessment for Two Normal Random Variables, Q and F

function. Numerical values for $\Phi(\cdot)$ are tabulated in many standard texts on reliability theory (e.g., Melchers 1987). The probability of failure is cumbersome to use when its value becomes very small, and this term carries the negative connotation of "failure". A more convenient (and perhaps more palatable) measure of design risk is the reliability index (β), which is defined as:

$$\beta = -\Phi^{-1}(p_f) \tag{3}$$

in which $\Phi^{-1}(\cdot)$ = inverse standard normal cumulative function. Note that β is not a new measure of design risk. It is just an alternative method for presenting p_f on a more convenient scale. A comparison of Eqs. 2 and 3 shows that the reliability index for the special case of two normal random variables is given by:

$$\beta = m_M / s_M = (m_Q - m_F) / (s_Q^2 + s_F^2)^{0.5} \tag{4}$$

The reliability indices for most structural and geotechnical components and systems are between 1 and 4, corresponding to probabilities of failure ranging from about 16 to 0.003%, as shown in Table 2. Note that p_f decreases as β increases, but the variation is not linear. A proper understanding of these two terms and their interrelationship is essential, because they play a fundamental role in RBD.

Simplified RBD for Foundations

Once a reliability assessment technique is available, the process of RBD would involve evaluating the probabilities of failure of trial designs until an acceptable target value is achieved. While the approach is rigorous, it is not suitable for designs that are conducted on a routine basis. One of the main reasons for this limitation is that the reliability assessment of realistic geotechnical systems is more involved than that shown in Fig. 2. The simple closed-form solution given by Eq. 2 only is applicable to cases in which the safety margin can be expressed as a linear sum of normal random variables. However, the capacity of most geotechnical systems is more suitably expressed as a nonlinear function of random design soil parameters (e.g., effective stress friction angle, in-situ horizontal stress coefficient, etc.) that generally are non-normal in nature. To evaluate p_f for this general case, fairly elaborate numerical procedures are needed, such as the First-Order Reliability Method (FORM). A general description of FORM for foundation engineering is given elsewhere (e.g., Phoon et al. 1995) and is beyond the scope of this paper. At the present time, it is safe to say that most geotechnical engineers would feel uncomfortable performing such elaborate calculations because of their lack of proficiency in probability theory (e.g., Whitman 1984).

All of the existing implementations of RBD are based on a simplified approach that involves the use of multiple-factor formats for checking designs. The three main

TABLE 2. Relationship Between Reliability Index and Probability of Failure

Reliability Index, β	Probability of Failure, $p_f = \Phi(-\beta)$
1.0	0.159
1.5	0.0668
2.0	0.0228
2.5	0.00621
3.0	0.00135
3.5	0.000233
4.0	0.0000316

Note: $\Phi(\cdot)$ = standard normal probability distribution

types of multiple-factor formats are: (a) partial factors of safety, (b) load and resistance factor design (LRFD), and (c) multiple resistance factor design (MRFD). Examples of these design formats are given below for the case of uplift loading of a drilled shaft:

$$\eta F_n = Q_u(c_n/\gamma_c, \phi_n/\gamma_\phi) \tag{5a}$$

$$\eta F_n = \Psi_u Q_{un} \tag{5b}$$

$$\eta F_n = \Psi_{su} Q_{sun} + \Psi_{tu} Q_{tun} + \Psi_w W \tag{5c}$$

in which η = load factor, F_n = nominal design load, Q_u = uplift capacity, c_n = nominal cohesion, ϕ_n = nominal friction angle, γ_c and γ_ϕ = partial factors of safety, Q_{un} = nominal uplift capacity, Q_{sun} = nominal uplift side resistance, Q_{tun} = nominal uplift tip resistance, W = shaft weight, and Ψ_u, Ψ_{su}, Ψ_{tu}, Ψ_w = resistance factors. The multiple factors in the simplified RBD equations are calibrated rigorously using reliability theory to produce designs that achieve a known level of reliability consistently. Details of the geotechnical calibration process are given elsewhere (e.g., Phoon et al. 1995).

In principle, any of the above formats or some combinations thereof can be used for calibration. The selection of an appropriate format is mostly unrelated to reliability analysis. Practical issues, such as simplicity and compatibility with the existing design approaches, are important considerations that will determine if the simplified RBD approach can gain ready acceptance among practicing engineers. At present, the partial factors of safety format (Eq. 5a) has not been used for RBD because of three main shortcomings. First, a unique partial factor of safety can not be assigned to each soil property, because the effect of its uncertainty on the foundation capacity depends on the specific mathematical function in which it is embedded. Second, indiscriminate use of the partial factors of safety can produce factored soil property values that are unrealistic or physically unrealizable. Third, many geotechnical engineers prefer to assess foundation behavior using realistic parameters, so that they would have a physical feel for the problem, rather than perform a hypothetical computation using factored parameters (Duncan et al. 1989, Green 1993, Been et al. 1993).

This preference clearly is reflected in the traditional design approach, in which the modification for uncertainty often is applied to the overall capacity using a global factor of safety (FS) as follows:

$$F_n = Q_{un} / FS \tag{6}$$

A comparison between Eq. 6 and Eqs. 5b and 5c clearly shows that the LRFD and MRFD formats are compatible with the preferred method of applying safety factors. In fact, the load and resistance factors in the LRFD format can be related easily to the familiar global factor of safety as follows:

$$FS = \eta / \Psi_u \tag{7}$$

The corresponding relationship for the MRFD format is:

$$FS = \eta / (\Psi_{su} Q_{sun} / Q_{un} + \Psi_{tu} Q_{tun} / Q_{un} + \Psi_w W / Q_{un}) \quad (8)$$

Although Eq. 8 is slightly more complicated, it still is readily amenable to simple calculations. These relationships are very important, because they provide the design engineer with a simple direct means of checking the new design formats against their traditional design experience.

RBD EXAMPLE

The development of a rigorous and robust RBD approach for geotechnical engineering design, which also is simple to use, is no trivial task. Since the early 1980s, an extensive research study of this type has been in progress at Cornell University under the sponsorship of the Electric Power Research Institute and has focused on the needs of the electric utility industry. Extensive background information on site characterization, property evaluation, in-situ test correlations, etc. had to be developed as a prelude to the RBD methodology. This work is summarized elsewhere (Spry et al. 1988, Orchant et al. 1988, Filippas et al. 1988, Kulhawy et al. 1992). Building on these and numerous studies by others, ultimate and serviceability limit state RBD equations were developed for drilled shafts and spread foundations subjected to a variety of loading modes under both drained and undrained conditions (Phoon et al. 1995). The results of an extensive reliability calibration study for ultimate limit state design of drilled shafts under undrained uplift loading are presented in Tables 3 and 4 and are to be used with Eqs. 5b (LRFD) and 5c (MRFD). All other limit states, foundation types, loading modes, and drainage conditions addressed have similar types of results, with simple LRFD and MRFD equations and single

TABLE 3. Undrained Ultimate Uplift Resistance Factors For Drilled Shafts Designed Using $F_{50} = \Psi_u Q_{un}$ (Phoon et al. 1995, p. 6-7)

Clay	COV of s_u (%)	Ψ_u
Medium	10 - 30	0.44
(mean s_u = 25 - 50 kN/m^2)	30 - 50	0.43
	50 - 70	0.42
Stiff	10 - 30	0.43
(mean s_u = 50 - 100 kN/m^2)	30 - 50	0.41
	50 - 70	0.39
Very Stiff	10 - 30	0.40
(mean s_u = 100 - 200 kN/m^2)	30 - 50	0.37
	50 - 70	0.34

Note: Target reliability index = 3.2

TABLE 4. Undrained Uplift Resistance Factors For Drilled Shafts Designed Using $F_{50} = \Psi_{su} Q_{sun} + \Psi_{tu} Q_{tun} + \Psi_w W$ (Phoon et al. 1995, p. 6-7)

Clay	COV of s_u (%)	Ψ_{su}	Ψ_{tu}	Ψ_w
Medium	10 - 30	0.44	0.28	0.50
(mean s_u = 25 - 50 kN/m^2)	30 - 50	0.41	0.31	0.52
	50 - 70	0.38	0.33	0.53
Stiff	10 - 30	0.40	0.35	0.56
(mean s_u = 50 - 100 kN/m^2)	30 - 50	0.36	0.37	0.59
	50 - 70	0.32	0.40	0.62
Very Stiff	10 - 30	0.35	0.42	0.66
(mean s_u = 100 - 200 kN/m^2)	30 - 50	0.31	0.48	0.68
	50 - 70	0.26	0.51	0.72

Note: Target reliability index = 3.2

corresponding tables of resistance factors. In these equations, the load factor is taken as unity to maintain statistical compatibility with the ASCE loading guide (Task Committee on Structural Loadings 1991), while the nominal load is defined as the 50-year return period load (F_{50}), which is typical for electrical transmission line structures. Note that the resistance factors depend on the clay consistency and the coefficient of variation (COV) of the undrained shear strength (s_u). The COV is an alternative measure of uncertainty that is defined as the ratio of the standard deviation to the mean. The clay consistency is classified broadly as medium, stiff, and very stiff, with corresponding mean s_u values of 25 to 50 kN/m^2, 50 to 100 kN/m^2, and 100 to 200 kN/m^2, respectively. Foundations are designed using these new RBD formats in the same way as in the traditional approach, with the exception that the rigorously-determined resistance factors shown in Tables 3 and 4 are used in place of an empirically-estimated factor of safety.

Target Reliability Index

Before applying these resistance factors directly in design, it is important to examine the target reliability index for which these resistance factors are calibrated. At the present time, there are no simple or straightforward procedures available to produce the "correct" or "true" target reliability index. However, important data that can be used to guide the selection of the target reliability index are the reliability indices implicit in existing designs (Ellingwood et al. 1980). An example of such data for ultimate limit state design of drilled shafts in undrained uplift is shown in Fig. 3, in which a typical range of COV of s_u, mean s_u normalized by atmospheric pressure (p_a), and global factor of safety are examined for a specific geometry. The reliability indices implicit in these existing global factor of safety designs lie in the approximate range of 2.6 to 3.7. A target reliability index of 3.2 is representative of this

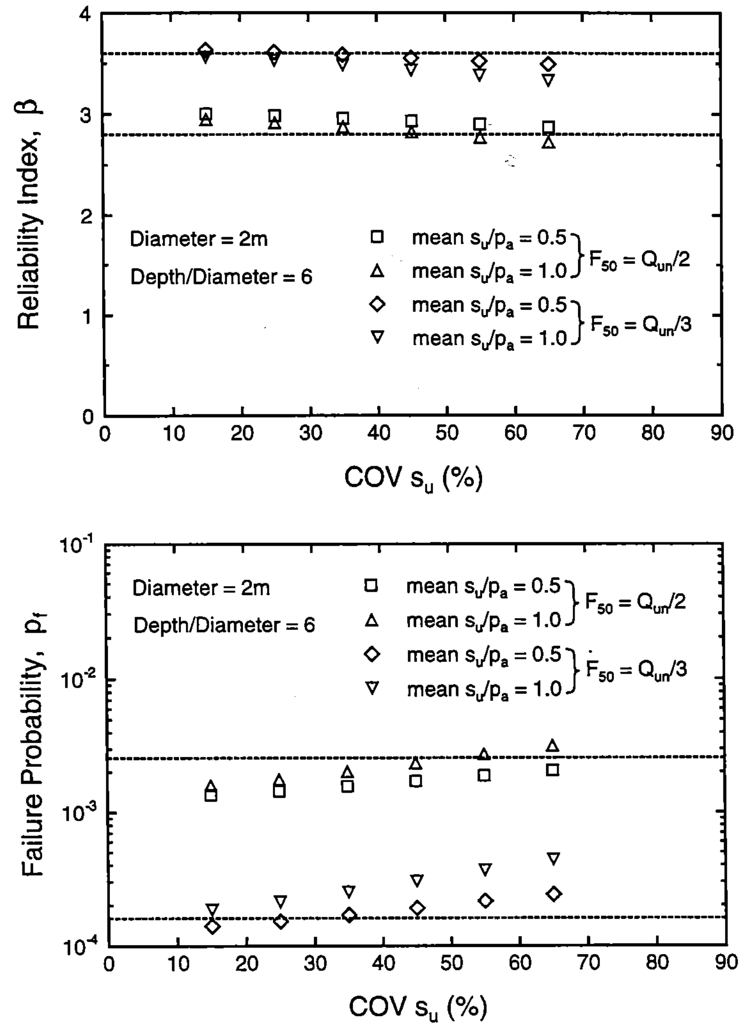


FIG. 3. Reliability Levels Implicit in Existing Ultimate Limit State Design of Drilled Shafts in Undrained Uplift

range. Similar ultimate limit state reliability studies for other parametric variations and loading modes under both drained and undrained conditions also strongly support the use of this target value (Phoon et al. 1995). While this approach is partially empirical, it possesses a major advantage of keeping the new design methodology com-

patible with the existing experience base.

Other important data to consider include the failure rates estimated from actual case histories. However, these failure rates can not be used directly for assessing the target reliability level because the theoretical probability of failure obtained from reliability theory usually is one order of magnitude smaller than the actual failure rate (e.g., CIRIA 1977, Livingstone 1989). This result is not surprising, because the safety of a design is not affected by uncertainties underlying design calculations alone. It also can be compromised severely by factors such as poor construction and human errors. An example of empirical rates of failure for civil engineering facilities and the related costs of failure is given in Fig. 4. For foundations, the empirical rate of failure lies between 0.1 and 1%. This failure rate implies a theoretical probability of failure in the range of 0.01 to 0.1%. Therefore, In terms of the reliability index, the currently accepted risk level is between 3.1 and 3.7. A target reliability index of 3.2 also is consistent with this range.

The above discussion highlights some of the important considerations that are involved in the determination of the target reliability index. It is apparent that the target reliability index can not be adjusted casually without extensive prior calibration with existing practice. Different target reliability indices can be used for design, but specific guidelines always should be given on the conditions for which each target val-

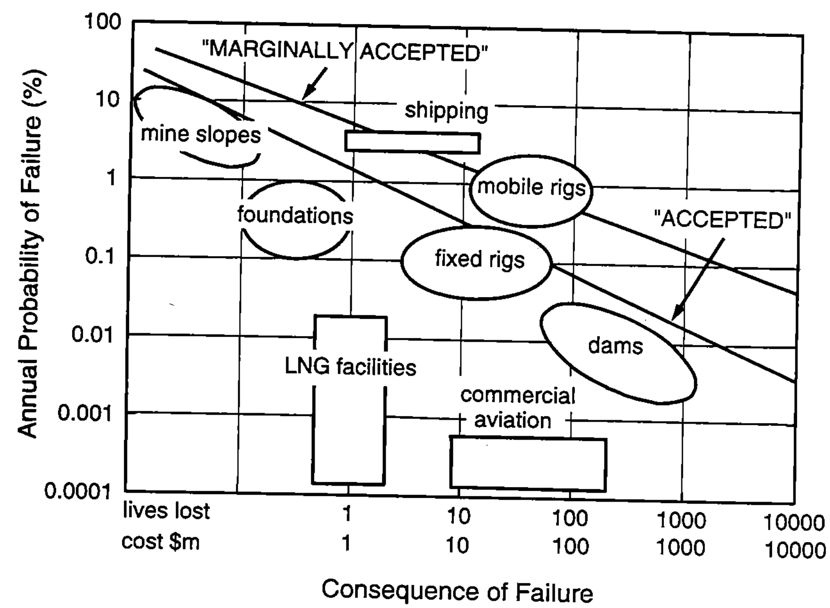


FIG. 4. Empirical Rates of Failure for Civil Engineering Facilities (Baecher 1987, p. 49)

ue applies. An example of a specific area in which a different target reliability index should be used is for serviceability limit state design (Phoon et al. 1995).

In the absence of specific guidelines, it might be possible for engineers to adjust the target reliability index to reflect some design conditions that already have been accounted for in the calibration of the load and resistance factors. For example, an engineer might be tempted to use a different target reliability index for drained and undrained analysis, because the uncertainty in the in-situ horizontal stress coefficient might be judged to be higher than that in the undrained shear strength. Such intuitive adjustment of the safety level based on judgment alone is normal in the traditional factor of safety design approach and for some rather simplified geotechnical RBD approaches that have been suggested (e.g., Task Committee on Structural Loadings 1991). However, in a rigorous RBD approach, the difference between drained and undrained analyses already has been accounted for rationally in the resistance factors, and further adjustment of the target level would amount to "double-counting". Errors of this type are to be expected in the absence of proper guidance because the typical RBD code user is not familiar with the details underlying the reliability calibration process. A proper appreciation of the target reliability index selection process particularly is important, because the target reliability index often has been mistaken (incorrectly) as the RBD equivalent of the empirical factor of safety.

Definition of Nominal Component

Aside from careful selection of the target reliability index, it also is important to define and understand precisely the nominal components shown in Eq. 5 . The level of safety in a design clearly is governed by the product of the load and resistance factors and their respective nominal components. Two foundations can have widely different safety levels, even though the same set of resistance factors is applied, because one design might be based on average soil parameters while another could have accrued additional safety by using highly conservative soil parameters. This important aspect is not sufficiently well-emphasized in the RBD literature (CIRIA 1977, Been & Jefferies 1993).

The definitions of nominal soil strengths in the simplified RBD formats ideally should be consistent with those used in traditional foundation design practice. However, the existing procedures for selecting nominal soil strengths are not well-defined, and they certainly are not followed uniformly by all engineers. Some engineers use the mean value, while others use the most conservative of the measured soil strength (e.g., Whitman 1984). Many other guidelines and rules-of-thumb exist in the literature. For example, Terzaghi and Peck (1948) recommended that the measured strength within a significant depth from each boring should be averaged, and then the smallest average should be used for design.

An alternative definition for the nominal value is based on the concept of an exclusion limit. The definition of a 5% exclusion limit is given in Fig. 5. However, the use of a small exclusion limit probably is not appropriate for foundation design because of several reasons. First, the amount of data required for the reliable determination of a 5 to 10% exclusion limit typically is much larger than the number of measurements taken

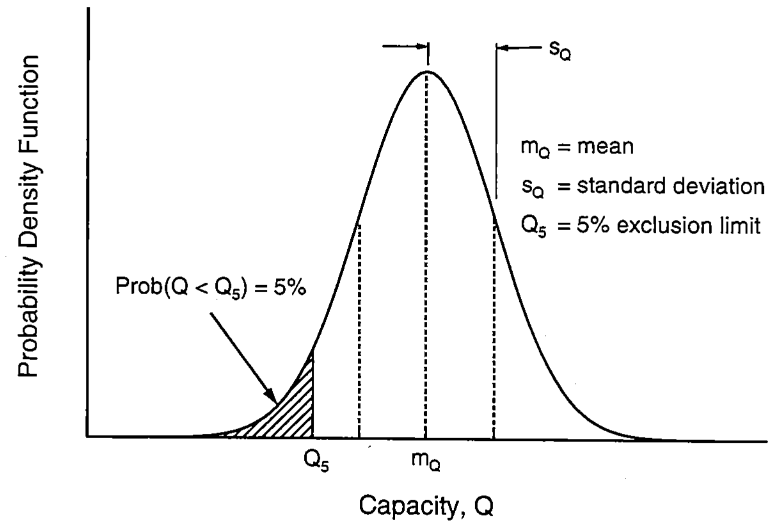


FIG. 5. Definition of 5 Percent Exclusion Limit

for a routine project (Been & Jefferies 1993). Second, the exclusion limit requires probability computations that are not currently performed in most foundation design. The main purpose of using a simplified RBD approach is to relieve practicing engineers from unfamiliar probability calculations so that they can focus on the real geotechnical aspects of the problem. Use of the exclusion limit introduces unnecessary complications and partially undermines the objective of a simplified RBD approach. Third, the exclusion limit concept is less intuitive than that for the mean value.

It is safe to say that most foundation engineers would feel more comfortable using the mean value, because they have a physical feel for that concept from their past experiences of working with actual measured soil strength parameters. Regardless of the choice, it is important to emphasize that the definition of nominal values can not be left to the judgment of each individual engineer, as is the case in traditional practice, if a uniform reliability level is to be maintained. In our opinion, all nominal soil parameters should be defined at the mean because of simplicity and compatibility with most existing foundation design practice.

The other important nominal component is the load, which geotechnical engineers normally do not investigate in detail. Loading agendas can be rather complicated, so it is necessary at least to appreciate what these values mean. In many codes these days, loads are specified using the concept of a return period. As an example, the ASCE loading guide for electrical transmission line structures (Task Committee on Structural Loading 1991) establishes the design loads for wind and other weather-related events at a return period of 50 years. The annual probability of exceeding the 50-year return period load is 1 in 50 or 2%. Other criteria are used by other organizations and for different types of structures and loads. The resistance factors used in RBD generally are related to the loading model and the definition of the nominal load as well.

Calibrations

Eqs. 5b and 5c and the corresponding Tables 3 and 4 look simple and just as easy to use as traditional design practice, which is the intent of any new or alternative design approach. However, there the similarities end. With RBD, rigorous calibrations of the design equations and all the input terms are performed to achieve a target reliability index. Specified within this approach are the nominal loads and resistances, target reliability index, design equation, and resistance factors, all calibrated together rigorously over a range of parameters using the First-Order Reliability Method (FORM). For the cases shown in Tables 3 and 4, the calibration parameter ranges were as follows: wind speed = 30 to 50 m/s, shaft diameter = 1 to 3 m, shaft depth/diameter = 3 to 10, mean s_u = 25 to 200 kN/m^2, and COV of mean s_u = 10 to 70%. For each combination of parameters, unique resistance factors apply. However, it is impractical to list all of these factors, because to do so would require literally dozens of tables. Instead, the results were scrutinized carefully, and it was found that the resistance factors could be grouped or "averaged" quite effectively over three ranges of s_u and three ranges of COV of s_u, as given in Tables 3 and 4.

The reliability indices for foundations obtained in this manner can not achieve the target reliability index exactly because the same resistance factor is applied to a range of undrained shear strengths. However, with the three groupings selected, it was possible to reduce the average deviation from the target reliability index to a minimum, as shown in Table 5 and Fig. 6. A comparison of the average deviations produced by the LRFD and MRFD formats also indicates that the MRFD format provides better reliability control. This observation is applicable to the other loading cases as well.

TABLE 5. Average Deviation From Target Ultimate Resistance Reliability Index For Drilled Shafts In Undrained Uplift (Phoon et al. 1995, p. 6-8)

Clay	COV of s_u (%)	Average Reliability Deviation	
		LRFD[a]	MRFD[b]
Medium	10 - 30	0.031	0.023
(mean s_u = 25 - 50 kN/m^2)	30 - 50	0.042	0.034
	50 - 70	0.053	0.042
Stiff	10 - 30	0.047	0.037
(mean s_u = 50 - 100 kN/m^2)	30 - 50	0.068	0.054
	50 - 70	0.087	0.063
Very Stiff	10 - 30	0.072	0.051
(mean s_u = 100 - 200 kN/m^2)	30 - 50	0.102	0.074
	50 - 70	0.125	0.082

Note: Target reliability index = 3.2
a - designed using $F_{50} = \Psi_u Q_{un}$
b - designed using $F_{50} = \Psi_{su} Q_{sun} + \Psi_{tu} Q_{tun} + \Psi_w W$

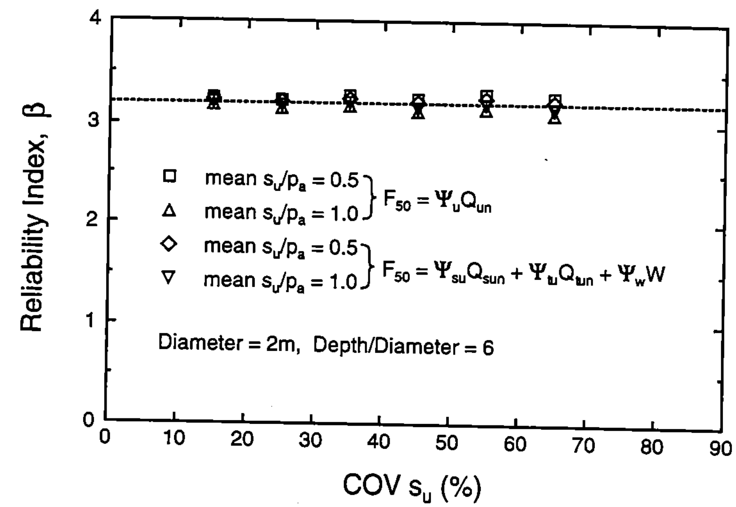


FIG. 6. Performance of Ultimate Limit State RBD Format for Drilled Shafts in Undrained Uplift

With these calibrations apparently so all-encompassing, one can ask the question whether engineering judgment is being usurped. The answer is an unequivocal "no". Instead, RBD focuses our efforts and judgment on the important issues. First, it forces us to agree on the load model, target reliability index, and design equation to use (at least for the time being). Second, it focuses our energies on evaluating the mean material properties and the variability (COV) in these properties for a given design situation. Specific guidance on these issues is beyond the scope of this paper, but detailed recent discussions by the authors are given elsewhere (Kulhawy 1992, Phoon et al. 1995, Phoon & Kulhawy 1996, Kulhawy & Trautmann 1996). Evaluation of the mean and COV for a particular boundary condition (shear, plane strain, extension, etc.) requires a careful assessment of all site, geologic, exploration, and testing variables. And finally, given that the design engineer knows explicitly what is included in RBD, the design engineer then can enhance or modify the calculation results to include the intangible and/or unforeseen issues noted previously.

SUMMARY

Judgment has, and probably always will, play a critical role in geotechnical engineering design, especially during the evolution from traditional deterministic design to new concepts of reliability-based design (RBD). In traditional geotechnical foundation design, the risk of adverse performance has been controlled by an empirical factor of safety at the design stage. However, this traditional design approach does not ensure a consistent level of safety, because the empirical factor of safety is not well-defined, and its relationship to its underlying uncertainties is ambiguous. To address these problems in a more realistic fashion, an essential first step is to adopt limit state design, which is intimately related to RBD. On one hand, the philosophy of limit states represents a logical and systematic approach to the process of engineering design. Conversely, the formalization of one aspect of this whole process, which is the application of reliability theory to ensure that the occurrence of limit states is sufficiently improbable, constitutes the main thrust of RBD. From this perspective, limit state design represents a more fundamental approach. Undue emphasis on RBD at the expense of the other design aspects clearly must be avoided.

In this paper, a general overview of reliability theory and a simplified RBD design approach have been presented. The load and resistance factor design (LRFD) and multiple resistance factor design (MRFD) formats were shown to be suitable for reliability calibration because they provide the design engineer with a simple direct means of checking the new design formats against their traditional design experience. Generally, the MRFD format is to be preferred. The proper use of these simplified RBD formats was discussed with reference to the ultimate limit state design of drilled shafts under undrained uplift loading. The two important aspects of this new design approach that can not be left entirely to the routine judgment of the design engineer are: (a) the selection of the target reliability index and (b) the definition of the nominal quantities in the design equations.

The applications of these new concepts were explored in some detail, and it was stressed that judgment still has a very important role in the design process. However, the judgment issues shift largely from assessing empirical factors to defining material characteristics and uncertainties explicitly and to judging intangibles and unknowns implicitly. This process puts design within a more rigorous and consistent framework.

ACKNOWLEDGMENTS

This paper is based largely on research results from RP1493, sponsored by the Electric Power Research Institute. The current EPRI project manager is A. Hirany. C. H. Trautmann and M. D. Grigoriu have been active and valuable contributors to this effort for over ten years, and their contributions are greatly appreciated.

REFERENCES

ACI Comm. 318 (1983), "Building Code Requirements for Reinforced Concrete", *ACI 318-83*, American Concrete Inst., Detroit, 111 p.

American Institute of Steel Construction (1986), *Load & Resistance Factor Design Manual of Steel Construction,* AISC, Chicago, 1114 p.

Baecher, GB (1987), "Geotechnical Risk Analysis User's Guide", *FHWA/RD-87-011*, Fed. Hwy. Admin., McLean, 55 p.

Barker, RM, Duncan, JM, Rojiani, KB, Ooi, PSK, Tan, CK & Kim, SG (1991), "Manuals for Design of Bridge Foundations", *NCHRP Report 343,* Trans. Res. Board, Washington, 308 p.

Been, K & Jefferies, MG (1993), "Determination of Sand Strength for Limit State Design", *Intl. Symp. Limit State Design in Geotech. Eng. (1),* Danish Geotech. Inst., Copenhagen, 101-110.

Been, K, Clark, JI & Livingstone, WR (1993), "Verification & Calibration Studies for New CAN/CSA-S472 Fndns. of Offshore Structures", *Can. Geotech. J.,* 30(3), 515-525.

Berger, JA. & Goble, GG (1992), "Reliability-Based Specification for Driven Piles", *Proc.,* 6th Spec. Conf. Probabilistic Mechanics & Struc. & Geotech. Reliability, Denver, preprint.

British Standards Institution (1972), "Code of Practice for Structural Use of Concrete", *CP110 (Pt 1),* BSI, London, 154 p.

Canadian Standards Assn. Comm. on Design of Light Gauge Steel Structural Members (NC Lind, Chair) (1974), "Cold-Formed Steel Structural Members", *Std. S136*, CSA, Rexdale, 43 p.

Casagrande, A (1965), "Role of Calculated Risk in Earthwork & Foundation Engineering", *J. Soil Mech. & Fndns. Div. (ASCE),* 91(SM4), 1-40.

Construction Industry Research & Information Assn. (1977), "Rationalization of Safety & Serviceability Factors in Structural Codes", *Rpt. 63*, CIRIA, London, 226 p.

Danish Geotechnical Institute (1985), "Code of Practice for Foundation Engineering", *Bull. 36,* DGI, Copenhagen, 53 p.

Duncan, JM, Tan, CK, Barker, RM & Rojiani, KB (1989), "Load & Resistance Factor Design for Bridge Foundations", *Symp. Limit States Design in Fndn. Eng.,* Can. Geotech. Soc., Toronto, 47-63.

Dunnicliff, J & Deere, DU (1984), *Judgment in Geotechnical Engineering: The Professional Legacy of Ralph B. Peck,* Wiley, New York, 332 p.

Ellingwood, B, Galambos, TV, MacGregor, JG & Cornell, CA (1980), "Development of Probability-Based Load Criterion for American National Standard A58", *Spec. Pub. 577,* Natl. Bureau Stds., Washington, 222 p.

Filippas, OB, Kulhawy, FH & Grigoriu, MD (1988), "Reliability-Based Foundation Design for Transmission Line Structures: Uncertainties in Soil Property Measurement" *Rpt. EL-5507(3),* Electric Power Res. Inst., Palo Alto, 94 p.

Focht, JA, Jr (1994), "Lessons Learned From Missed Predictions", *J. Geotech. Eng. (ASCE),* 120(10), 1653-1683.

Focht, JA, Jr & O'Neill, MW (1985), "Piles & Other Deep Foundations", *Proc.,* 11th Intl. Conf. Soil Mech. & Fndn. Eng. (1), San Francisco, 187-209.

Freudenthal, AM (1947), "Safety of Structures", *Trans. (ASCE),* 112, 125-159.

Green, R (1993), "LSD Code for Bridge Foundations", *Intl. Symp. Limit State Design in Geotech. Eng. (2),* Danish Geotech. Inst., Copenhagen, 459-468.

Kulhawy, FH (1984), "ASCE Drilled Shaft Standard: University Perspective", *Analysis & Design of Pile Fndns.,* Ed. JR Meyer, ASCE, New York, 390-395.

Kulhawy, FH (1992) "On the Evaluation of Static Soil Properties" *Stability & Performance of Slopes & Embankments II (GSP 31)* Ed. RB Seed & RW Boulanger, ASCE, New York, 95-115.

Kulhawy, FH (1994), "Some Observations on Modeling in Foundation Engineering", *Proc.*, 8th Intl. Conf. Computer Methods & Advances in Geomech. (1), Ed. HJ Siriwardane & MM Zaman, Morgantown, 209-214.

Kulhawy, FH, Birgisson, B & Grigoriu, MD (1992), "Reliability-Based Foundation Design for Transmission Line Structures: Transformation Models for In-Situ Tests" *Rpt. EL-5507(4)*, Electric Power Res. Inst., Palo Alto, 113 p.

Kulhawy, FH & Trautmann, CH (1996), "Estimation of In-Situ Test Variability", *this conference proceedings.*

Livingstone, WR (1989), "CSA Code for Design, Construction & Installation of Fixed Offshore Structures", *Symp. Limit States Design in Fndn. Eng.*, Can. Geotech. Soc., Toronto, 77-89.

MacGregor, JG (1989), "Problem from Viewpoint of Structural Engineer", *Symp. Limit States Design in Fndn. Eng.*, Can. Geotech. Soc., Toronto, 1-16.

Melchers, RE (1987), *Structural Reliability, Analysis & Prediction*, Wiley, New York, 400 p.

Mortensen, K (1983), "Is Limit State Design a Judgment Killer?", *Bull. 35*, Danish Geotech. Inst., Copenhagen, 16 p.

Nordic Comm. on Building Regulations (1978), "Recommendations for Loading & Safety Regulations for Structural Design", *Rpt. 36*, NKB (Nordic Comm. on Bldg. Regs.), Copenhagen.

Orchant, CJ, Kulhawy, FH & Trautmann, CH (1988), "Reliability-Based Foundation Design for Transmission Line Structures: Critical Evaluation of In-Situ Test Methods" *Rpt. EL-5507(2)*, Electric Power Res. Inst., Palo Alto, 214 p.

Phoon, KK, Kulhawy, FH & Grigoriu, MD (1995), "Reliability-Based Design of Foundations for Transmission Line Structures", *Rpt TR-105000*, Electric Power Res. Inst., Palo Alto, 380 p.

Phoon, KK & Kulhawy, FH (1996), "On Quantifying Inherent Soil Variability", *this conference proceedings.*

Pugsley, A (1955), "Report on Structural Safety", *Struc. Eng.*, 33(5), 141-149.

Spry, MJ, Kulhawy, FH & Grigoriu, MD (1988), "Reliability-Based Foundation Design for Transmission Line Structures: Geotechnical Site Characterization Strategy" *Rpt. EL-5507(1)*, Electric Power Res. Inst., Palo Alto, 142 p.

Task Comm. on Structural Loadings (J. D. Mozer, Chair) (1991), "Guidelines for Electrical Transmission Line Structural Loading", *Manual & Report on Eng. Practice 74*, ASCE, New York, 1991, 139 p.

Technical Comm. on Foundations, *Canadian Foundation Engineering Manual*, 3rd Ed, Can. Geotech. Soc., Richmond, 512 p.

Terzaghi, K & Peck, RB (1948), *Soil Mechanics in Engineering Practice*, Wiley, New York, 384 p.

Whitman, RV (1984), "Evaluating Calculated Risk in Geotechnical Engineering", *J. Geotech. Eng. (ASCE)*, 110(2), 145-188.

UNCERTAINTIES IN CHARACTERISING SOIL PROPERTIES

Suzanne Lacasse and Farrokh Nadim[1], MASCE

Abstract

The paper presents a review of the uncertainties in characterising soil properties, including spatial variability and measurement methods. It stresses the importance of characterising these uncertainties for design. The usefulness of determining and, where possible reducing, the uncertainties is illustrated with examples from actual case studies. To characterise the uncertainties in a soil property, the engineer needs to combine, in addition to the actual data, knowledge about the quality of the data, knowledge on the geology, and most importantly engineering judgement. When sufficient data are available, site description strategy, with the identification of correlation structure and stochastic interpolation to estimate a soil property is recommended. The geotechnical parameters for analysis are then more clearly defined. The added knowledge obtained from systematic uncertainty assessments should lead to safer and more economical designs. Accounting for uncertainties in soil properties and the calculation model is a necessary complement to more conventional deterministic analyses. A documentation that all possible aspects have been considered and dealt with is essential today, as quality assurance requirements grow. One challenge is to balance technology with complexity, given the budget and the consequences of the project. Accounting for uncertainties in an analysis and their effects on the response contributes to an improved understanding of the complexity of the problem to model.

Introduction

In recent years, new and creative solutions have been developed for geotechnical design, and calculation methods have been improved. Yet the characterisation and reduction of uncertainties still is an area where few are working, even though as early as 1982 Einstein and Baecher stated the following words of wisdom:

[1] Norwegian Geotechnical Institute, P.O. Box 3930 Ullevål Hageby, N-0806 Oslo, Norway

«In thinking about sources of uncertainty in engineering geology,
one is left with the fact that uncertainty is inevitable.
One attempts to reduce it as much as possible, but it must ultimately be faced.
It is a well recognised part of life for the engineer.
The question is not whether to deal with uncertainty, but how?»

The characterisation of uncertainties has the potential of benefiting to a great extent geotechnical solutions by making them explicit and documented, less uncertain, and the basis for a growing data base of well documented case studies.

This overview of characterisation of uncertainties in soil properties is divided into five main sections, with examples from actual geotechnical applications from NGI's files. The sources and types of uncertainty are first reviewed, then the statistical treatment of geotechnical data and their spatial variability and the uncertainties due to the measurement of soil properties are discussed. One cannot do a detailed review of sources of uncertainties without also looking into model uncertainty and its consequence for calculations. The importance of characterising uncertainties in the soil properties for design is also focused on.

As some of the terminology of statistical analysis may be new, Appendix I defines some of the terms used in this paper. References are listed in Appendix II.

Sources and types of uncertainty

Uncertainty modelling of the variables entering an analysis, whether probabilistic or any other type of analysis, requires collection of data, evaluation of the data set(s), selection of a "model" to represent the data, estimation of the uncertainty(ies) in this model and its significant characteristics, and a verification of the assumptions made. The evaluation of the data set(s) needs recognition of the type of uncertainties, whether the variables are dependent or independent, whether the observations are independent and whether the uncertainties noted are the result of a combination of uncertainties.

Uncertainties associated with a geotechnical problem can be divided into two groups: aleatory (inherent or natural) uncertainty and epistemic (due to lack of knowledge) uncertainty. Human error, which is not covered herein, would fall into a third category. Within a geological layer, the soil properties can be affected by both aleatory and epistemic uncertainties:

- Aleatory uncertainty represents the natural randomness of a property. The variation in the ocean wave height or wind force and the spatial variation in a soil property are aleatory uncertainties. Aleatory uncertainty cannot be reduced or eliminated.
- Epistemic uncertainty represents a range of values that can be reduced by collecting more information, improving the measurement method(s) or improving the calculation method(s). Epistemic uncertainty can be reduced, perhaps eliminated. The epistemic uncertainty can be either statistical, measurement-related and/or model-related. Statistical uncertainty is due to limited information such as limited number of observations. Measurement uncertainty is due to for example imperfections of an instrument or of a method to register a quantity. Model uncertainty is due to idealisations made in the physical formulation of the problem.

Statistical uncertainty is present because the parameters are estimated from a limited set of data, and is affected by the type of estimation technique used. Measurement uncertainty is described in terms of accuracy and is affected by bias (systematic error) and by precision (random error). It can be evaluated from data provided by the manufacturer, laboratory tests and/or scaled tests. Model uncertainty is defined as the ratio of the actual quantity to the quantity predicted by a model. A mean value different from 1.0 expresses a bias in the model, while the standard deviation expresses the variability in the predictions by the model.

Uncertain soil properties and model uncertainty are best defined as random variables described by a mean and a standard deviation (or coefficient of variation) and a probability distribution function. Figure 1a compares a soil property described deterministically and a soil property described with its uncertainty. In practice, none of us ever determines a deterministic (punctual) value for a soil property. Armed with engineering judgement, we select an appropriate characteristic value, on the basis of the available data, the expected range of values for this property, the type of problem to be analysed and our experience. Mentally we establish a possible range of values, and select either a most probable value or a somewhat conservative value.

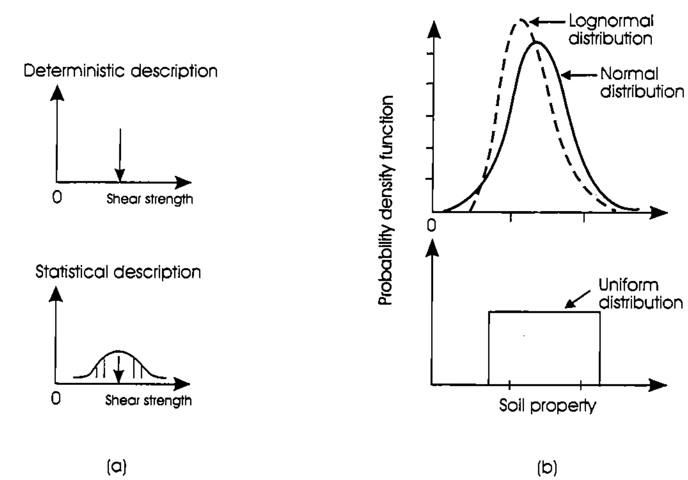


Figure 1. Deterministic and statistical description of soil property

Figure 1b shows typical probability distribution functions used in geotechnical problems. The normal and lognormal are the most common; the lognormal is often used to characterise variables that do not take negative values. A uniform distribution may also be adequate for an equally likely range of values. These distributions are simple and require little work except the use of standard statistical tables.

Reliability analyses (e.g. Lacasse and Nadim, 1994; Nadim et al, 1994) show that uncertainties on different soil properties affect differently the reliability of geotechnical analyses. It is therefore important that the uncertainties in analysis parameters be adequately quantified and their effect carefully evaluated, if a reliable safety margin or safety factor is to be established. Examples are given later in the paper.

Statistical treatment of geotechnical data

An important part of a reliability analysis is the statistical treatment of the data used to evaluate soil parameters. The statistical estimates give a mean value and an estimate of the uncertainty in the data. The statistical estimates should be combined with engineering judgement to select the parameters for design, and should consider both the quality of data and the geologic evidence. Statistics apply within 'homogeneous' layers, and it is important to identify the main soil or geologic layers before performing statistical calculations. Statistics can also be used to identify layer boundaries.

Statistical methods belong to either traditional statistics or geostatistical approaches. Table 1 summarises different types of statistical approaches that can be applied. The list is not exhaustive, but presents techniques that have been found to have successful geotechnical applications. If a soil parameter is obtained from a complex calculation with many random parameters, e.g. the cyclic shear strength of a clay under random loading, one should consider using simulation methods, for example Monte-Carlo simulation with Latin Hypercube sampling.

Table 1 Statistical tools and their application to geotechnical analysis

Method	Application	Recommendation
Short-cut estimates (Baecher, 1985; Snedecor and Cochran, 1964)	-For cases where little data are available - Useful for "symmetrical" data -Gives bound for standard deviation	-Use if few data -Use to check variance
Mean, variance, histograms, probability density (Ang and Tang, 75)	-Best method -Distribution function from probability plots and/or goodness-of-fit tests	-Use whenever possible
Geostatistics (stochastic interpolation) (Matheron, 1963; Nadim,1988)	-To do statistical site description -Applies to all soil characteristics -Need many data points -Software adapted to geotechnical data -Enhanced data presentation	-Use whenever enough data are available

To estimate the variance with only a limited number of data, one can use short-cut estimates, as suggested by Krumbein and Greyhill (1965) or Snedecor and Cochran (1964). The standard deviation is obtained by multiplying the range of the

available values by weighting factors shown in Table 2. The range is defined as the difference between the largest and smallest values in the data population. For example, for five data points, the standard deviation would be the range multiplied by 0.43. This approach is a good estimator for symmetric data populations.

Table 2. Short-cut estimates for limited data set (Snedecor and Cochran, 1964)

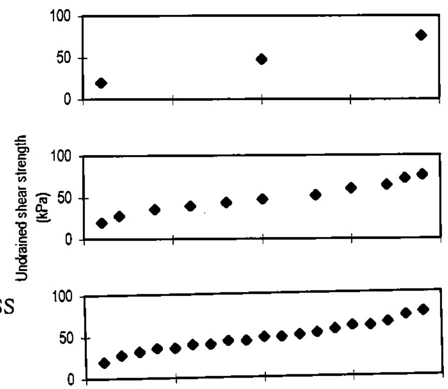
# points	Weighting factor	# points	Weighting factor
1	---	11	0.315
2	0.886	12	0.307
3	0.591	13	0.300
4	0.486	14	0.294
5	0.430	15	0.288
6	0.395	16	0.283
7	0.370	17	0.279
8	0.351	18	0.275
9	0.337	19	0.271
10	0.325	20	0.268

Figure 2 presents the results of in situ and laboratory tests on a low plasticity lightly overconsolidated clay. The data are restricted to one 5 meter layer with approximately constant undrained shear strength in this layer. At first, the results of three field vane tests from one boring were available. The data had a range of 56 kPa with a mean of 48 kPa. Three piezocone profiles were then run and the shear strength values were taken at the depths of the field vane tests. The undrained shear strengths were obtained with a cone factor of 14, based on a calibration with field vane and direct simple shear test results (Aas et al, 1986). Direct simple shear tests, consolidated to the preconsolidation stress and unloaded to the in situ stresses, yielded the s_u-values in Fig. 2. For a low plasticity clay, the field vane and direct simple strength generally compare well, and the data set can be made into a consistent population for statistical analysis. Figure 2 also gives the standard deviation from traditional statistics and with the short-cut estimates. With a large number of data points, the standard deviations are quite comparable. The figure also shows the distribution of the measurements made. Even with few data in the population, the distribution is close to normal or lognormal.

It is often the relation between two parameters which is of relevance for a calculation, for example correlation between corrected cone penetration resistance and undrained shear strength from triaxial compression tests via a cone factor, or variation of undrained shear strength with depth. Ang and Tang (1975) described in detail the mechanics of linear regression analysis. The goodness-of-fit is given by how close the correlation coefficient is to unity.

The standard deviation is of great importance for the evaluation of the uncertainties and their consequences. Figure 3 illustrates the standard deviation for a normally distributed variable: 68% of the data fall within one standard deviation about the mean, and nearly 100% of the data fall within 3 standard deviations. A coefficient of variation (standard deviation divided by mean value) of 0.30 therefore implies that the variable can take values 90-100% lower or higher than the mean value. This range is very large, and it is important to consider whether the standard deviation/ coefficient of variation arrived at is representative of the actual range of values over which the soil property is expected to vary.

Measurements	Undrained shear strength s_u (kPa)	Mean, μ Standard deviation, σ
FV (3 data points)		μ = 48 kPa σ = 28 kPa $\sigma_{shortcut}$ = 33 kPa
FV + CPTU (11 data points)		μ = 49 kPa σ = 17.8 kPa $\sigma_{shortcut}$ = 17.6 kPa
FV + CPTU + DSS (20 data points)		μ = 48 kPa σ = 14.7 kPa $\sigma_{shortcut}$ = 15.2 kPa



FV Field vane test CPTU Piezocone penetration test DSS Direct simple shear test

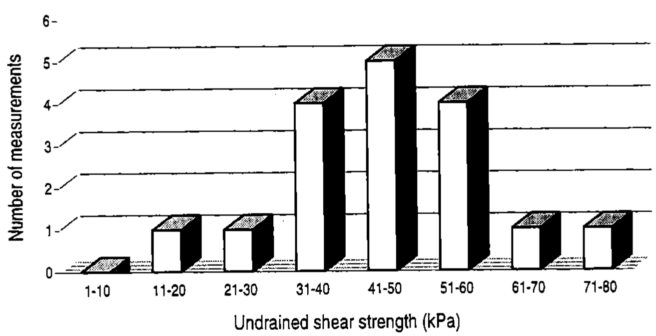


Figure 2. Comparison of shortcut method with actual mean and standard deviation Norwegian lightly overconsolidated clay of low plasticity

Results of cone penetration test are well suited for statistical analysis because penetration tests generate a lot of data. Filtering and smoothing techniques have been used (Vivatrat, 1978; Harder and von Bloh, 1988) to give unbiased average representations. Mortensen et al (1991) also used smoothing techniques to obtain a frequency histogram for cone resistance and for the cone factor correlating cone resistance and vane shear strength in clay tills. The quality of the correlation should be corrected by setting "quality criteria", for example by rejecting anomalous data points or data that do not fit in a specified criteria.

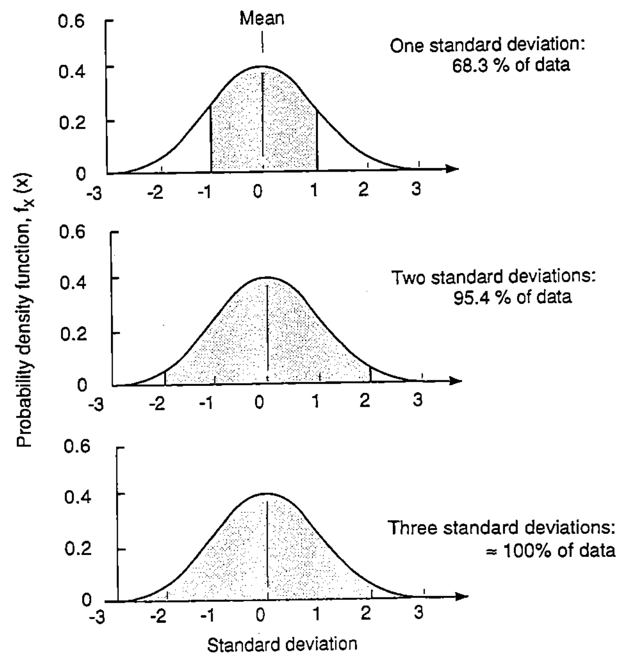


Figure 3. Uncertainty in soil property and range of values for normal variable

Spatial variability

Autocorrelation

For any extensive volume of natural soil layer, the characteristics fluctuate spatially. There is a greater tendency for the properties to be similar in value at closely neighbouring points than at widely spaced points. Soil properties are expected to show dependence both laterally and with depth.

The variation of a soil property in space is illustrated in Fig. 4a as a function of a trend, T(x), and residuals, $\varepsilon(x)$ (Vanmarcke, 1977, 1984; DeGroot and Baecher, 1993). The residuals over a large volume of soil are believed to have zero mean. The trend function is obtained by regression analysis; the residuals are correlated, unless the data are very widely spaced. If a soil property is assumed to be controlled by a random process, the spatial variation can be modelled as the sum of a trend and a residual $Y(x) = T(x) + \varepsilon_r(x)$, where $Y(x)$ is the measurement at location X, $T(x)$ the trend and $\varepsilon_r(x)$ the residual (deviation about trend).

A strong spatial correlation structure appears as a waviness about the trend. The larger the width, length or depth over which a parameter is averaged, the more the fluctuations tend to cancel each other in the process of "spatial averaging". This correlation structure can be important both for improving estimates at unsampled locations and for assessing a soil parameter. The degree of spatial correlation can be expressed through an autocovariance function, C(r), where r is the vector of separation distance between two points. The normalised form of the autocovariance function $\{C(r)/C(0)\}$ is known as the autocorrelation function, where C(0) is the autocovariance function at a distance r of 0 (i.e. the variance of the data).

The three autocovariance functions most often used to model soil properties are shown in Fig. 4b. In the exponential models, the distance at which the autocovariance function [C(r)] decays to a value of 1/e (where e is the base of the natural logarithm) is called the autocorrelation distance, r_o. This length is a measure of the extent of the spatial correlation. Figure 4c illustrates how the autocovariance distance and the variance influence the fluctuation of a soil property (DeGroot and Baecher, 1993).

As mentioned earlier, the uncertainties associated with soil characteristics are generally attributed to two primary sources: (1) inherent (natural soil) variability, and (2) sampling and testing errors (identified as "noise"). The soil variability component and noise could be separated with the use of an autocorrelation function. Nadim (1988) and Keaveny et al (1989) derived autocorrelations functions in the vertical and horizontal directions for various sets of in situ and laboratory data.

Figure 5 gives an example of the autocorrelation structure of cone penetration data at a depth of 9 m in a dense sand. The scale of fluctuation is the distance within which a soil property shows strong correlation and is related to the autocorrelation distance, r_o, (Vanmarcke, 1984). In Fig. 5, the autocorrelation function had the form $C(r) = 0.99\ e^{-(r/37.5)^2}$, where r and 37.5 are in meters. When extrapolating the curve back to a distance r of zero, the closeness of the factor 0.99 to unity indicates that there was little measurement noise in the recorded data.

Properly accounting for the variability in soil properties when predicting geotechnical performance may reduce substantially the uncertainties in soil parameters and therefore the uncertainties associated with a design. Unfortunately, one is never able to gather enough subsurface data to get an exact picture of the variation of a soil property for an engineering structure. One must therefore interpolate the soil properties within a large volume.

Traditional methods of interpolation used in geotechnical engineering give little regard to the uncertainties associated with soil properties and that soil properties exhibit a spatial correlation structure. A stochastic interpolation technique well-

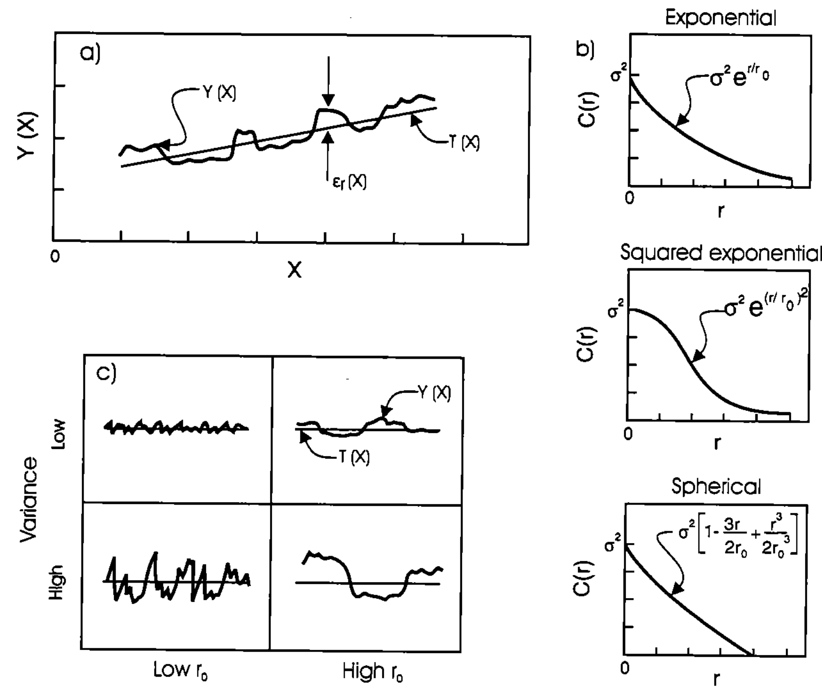


Figure 4. Spatial variability of soil property
(a) Trend and residuals (b) Autocovariance functions (c) Autocovariance distances

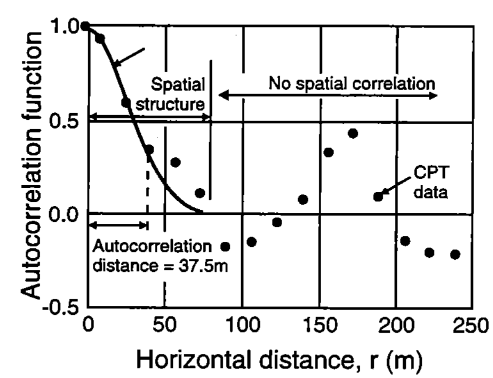


Figure 5. Autocorrelation structure of 17 cone penetration test data in dense sand

suited for this is the «kriging» approach (Matheron, 1963). It accounts for the uncertainties associated with the soil properties and minimises the variance in the interpolated data. To do the kriging interpolation, one needs to first identify the spatial structure of the soil characteristics or the autocorrelation function. An analysis method is described in Nadim (1988).

Example of geostatistical analysis

Geostatistical analysis is illustrated with an example. In the neighbourhood of a shallow foundation, several cone penetration tests were run. Figure 6 presents the locations of some of the available soundings and cone penetration test results (cone resistance q_c versus depth.)

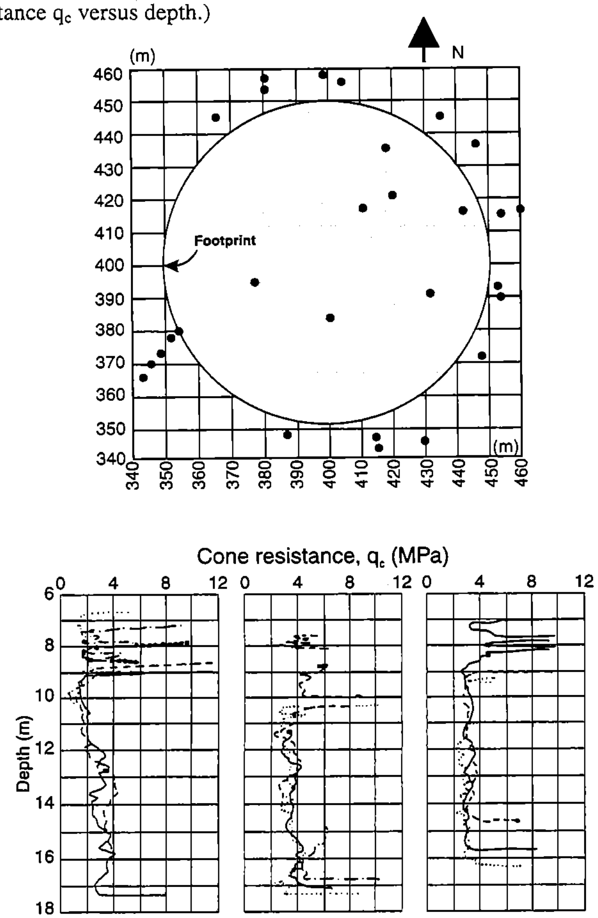


Figure 6. Soil investigation, geostatistics example

The soil profile consists of a top sand (7-10 m) over a relatively weaker clay, partly laminated, with varying shear strength. In this case, the location of the weaker clay layer, and the undrained shear strength for the weaker material were the main variables that conditioned the feasibility of the foundation. Using minimum values would result in large added costs.

Figure 7 presents the contours of cone penetration resistance at each meter between 7 and 12 m. At a depth of 9 m, the mean determined by kriging is significantly higher (4.6 MPa) than what had originally been assumed in design at that depth (about 2 MPa) based on the cone penetration profiles similar to those in Fig. 6. The 3-D plot also shows that there does not seem to be a continuous layer with weaker shear strength directly beneath the structure. At 9 m, the best fit autocorrelation function had the exponential form $C(r) = 1.0\ e^{-(r/9.6)}$, where r and 9.6 are in meters. The fact that the curve extrapolates back to unity at distance zero indicates no noise in the measurements.

Christian et al. (1994) also presented recent examples of spatial correlation analysis estimating the level of random noise in the soil property data.

The three-dimensional graphical representation in Fig. 7, obtained with fairly simple computer programs such as developed by Nadim (1988) and a graphics package, provides improved insight into the possible variation in the cone resistance and the most likely value beneath the foundation.

In this particular case study, the geostatistical analysis enabled the designers to adjust the assumed position of the clay layer below the depth that had originally been assumed and to use slightly higher cone penetration resistance values, and therefore higher shear strength in design.

Several workers studied the spatial autocorrelation structure of cone resistances. A brief review suggests the autocorrelation distances in Table 3.

Table 3. Autocorrelation distance for cone penetration resistance

Soil	Direction	Autocorr'on distance (m)	Reference
Offshore soils	Horizontal	30	Høeg and Tang(1976); Tang(1979)
Offshore sand	Horizontal	14-38	Keaveny et al (1989)
Silty clay	Horizontal	5-12	Lacasse and Lamballerie (1995)
Clean sand	Vertical	3	Alonzo and Krizek (1975)
Mexico clay	Vertical	1	Alonzo and Krizek (1975)
Clay	Vertical	1	Vanmarcke (1977)
Sensitive clay	Vertical	2	Chiasson et al (1995)
Silty clay	Vertical	1	Lacasse and Lamballerie (1995)

Spatial averaging to reduce uncertainty

The effect of the spatial variability on the computed performance (probability of failure, reliability index, margin of safety) is reduced because the variability is averaged over a volume, and only the averaged contribution to the uncertainty is of

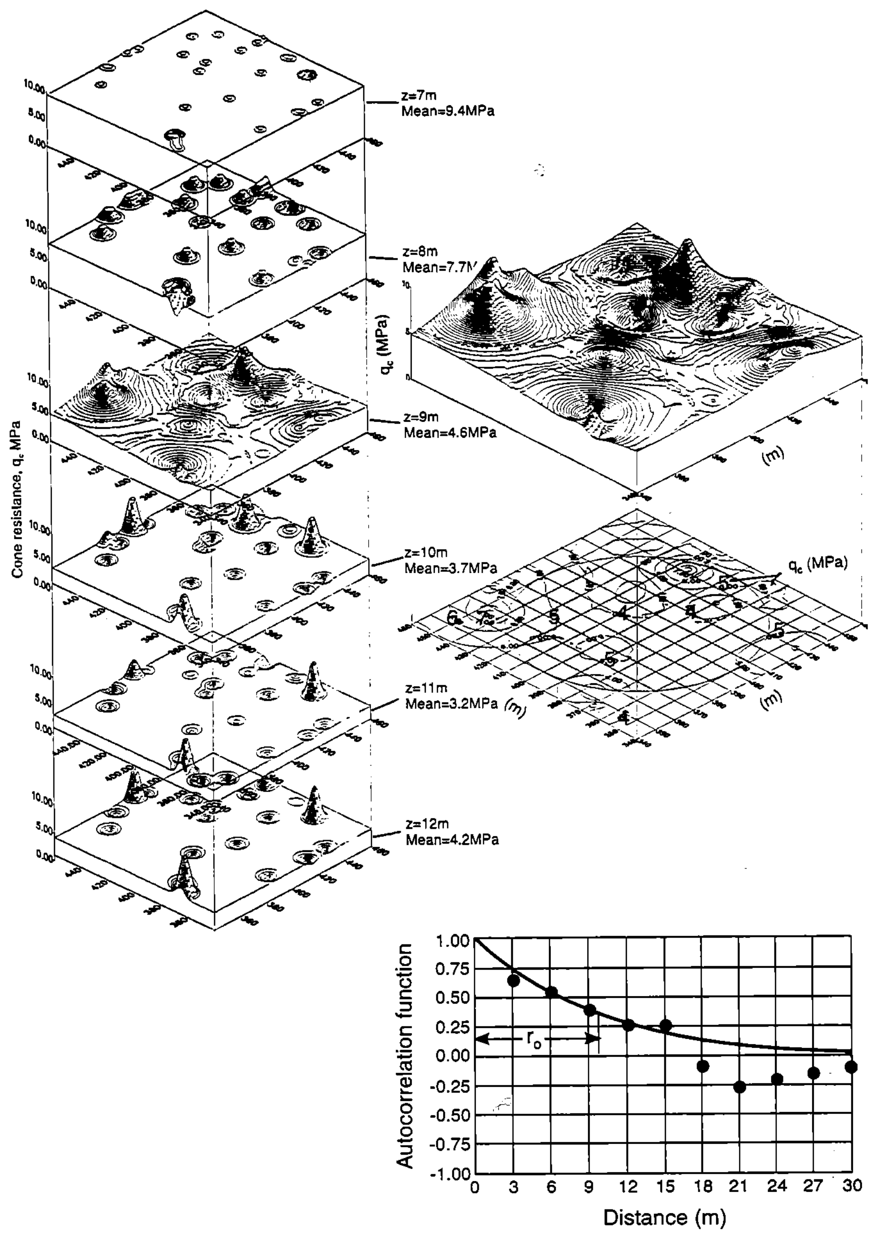


Figure 7. Results of geostatistical analysis at different depths

interest. Figure 8 illustrates how one obtains the reduction factor to account for spatial averaging as a function of the fluctuation distance.

The fluctuation distance is the distance over which a soil property shows relatively strong correlation from point to point. The fluctuation distance is between 1.4 and 2.0 times the autocorrelation distance for the exponential, squared exponential and spherical autocorrelation functions (Vanmarcke, 1977; 1984).

The reduction factor is obtained from the square root of the autocovariance function. The reduction in variance can be a factor of as much as 0.4 to 0.8, when the property has spatial structure. Vanmarcke (1984) suggested that the variance reduction for most autocorrelations functions used in geotechnical engineering could be approximated by a unique curve, which results in a simple relation between reduction factor and distance over which the soil parameter is averaged.

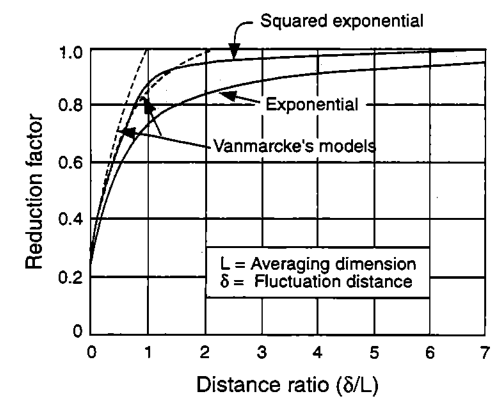


Figure 8. Reduction of variance with spatial averaging

<u>Uncertainties due to measurement of soil properties</u>

Figure 2 has already given an example of the uncertainties due to different measurements of a soil property. The first rule, when determining the uncertainties related to a soil property and using statistical methods, is to ensure that consistent data populations are used. Major uncertainties have been introduced in the past because of inconsistent data sets. The inconsistency can originate from different soils, different stress conditions, different test methods, stress history, different codes of practice, testing errors or imprecision that are not reported, different interpretations of the data, sampling disturbance, ...etc.

A review was made on test results in NGI's files and data available from the literature. Suspicious data were eliminated. The uncertainties, in terms of coefficient of variation (COV), shown in Table 4 were obtained. The probability distribution function arrived at are also shown.

Table 4. Coefficient of variation of different soil properties

Soil property	Soil type	Prob.distr. function	Mean	COV
Cone resistance	Sand Clay	LN N/LN	*	*
Undrained shear strength, s_u **	Clay (triax) Clay (index su) Clayey silt	LN LN N	*	5 - 20% 10 - 35% 10 - 30%
Ratio s_u/σ'_{vo}	Clay	N/LN	*	5 - 15%
Plastic limit	Clay	N	0.13-0.23	3 - 20%
Liquid limit	Clay	N	0.30-0.80	3 - 20%
Submerged unit weight	All soils	N	5-11 (kN/m^3)	0 - 10%
Friction angle	Sand	N	*	2 - 5%
Void ratio, porosity, initial void ratio	All soils	N	*	7 - 30%
Overconsolidation ratio	Clay	N/LN	*	10 - 35%

N/LN Normal and lognormal distribution
* Values are site- and soil type-dependent
** Undrained shear strength is anisotropic and depends on the type of stresses imposed. The coefficient of variation for good quality tests (consolidated triaxial compression/ extension, direct simple shear, true triaxial, plane strain) is expected to be 5-20%. For extension tests, because of generally fewer data available and at times more difficult testing conditions, the coefficient of variation may be higher.

Figure 9 presents an example of the undrained shear strength profile. A standard deviation of 15% about the mean includes most of the data, except for two low values below 100 m due to sampling disturbance at large depth. The results are for unconsolidated-undrained laboratory tests, which do not correct for any effects of sampling disturbance. Most geotechnicians have less difficulty determining the mean and range of the normalised undrained shear strength ratio s_u/σ'_{vo} of clays than the mean and range of the undrained shear strength s_u, especially when this property varies between 50 and 200 kPa. In the analysis of the uncertainty in Fig. 9, the estimates were done on the basis of the undrained shear strength ratio s_u/σ'_{vo} of the clay instead of the property s_u. The shear strength data set was also dependent on depth. Soil parameters and their uncertainty, when evaluated, should therefore be expressed in terms of the variables that are easiest to quantify.

The correlation structure of the undrained shear strength of a clay using different data populations is presented in Fig. 10. The undrained shear strength normalised with the in situ effective vertical stress is considered for this lightly overconsolidated clay. On the left side of the figure, results from tests consolidated to the in situ effective stresses and to stresses past the preconsolidation stress are mixed. On the right side, only the p'$_o$-consolidated tests are included. It was not

possible to obtain a reasonable autocorrelation function in the case of the «mixed» data, whereas a correlation structure is obtained for the p'$_o$- data set.

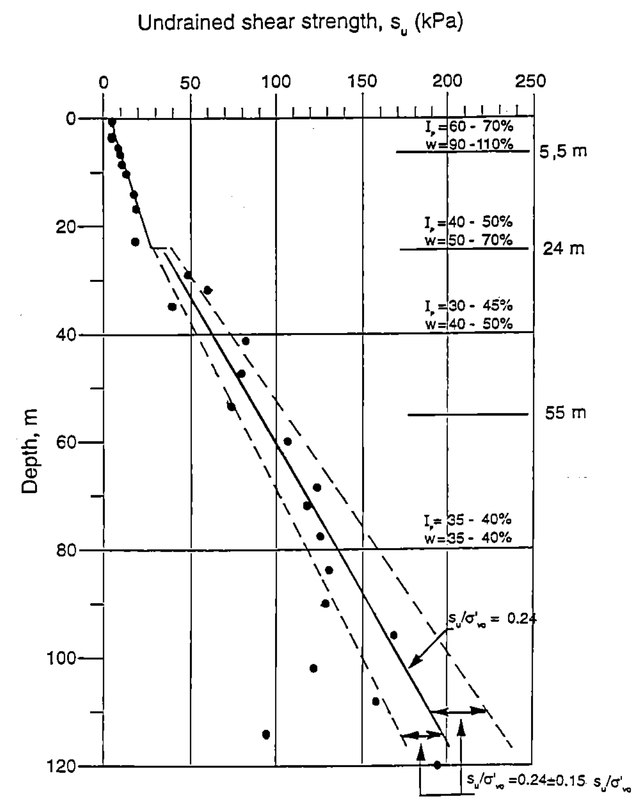
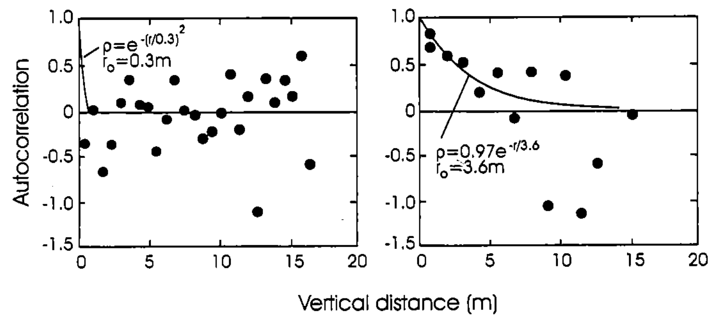


Figure 9. Uncertainty in undrained shear strength ratio with depth (UU tests)

Figure 11 illustrates the importance of lumping together only consistent data populations. The diagrams present results of field vane test at over forty sites on clays with varying plasticities The data are plotted in terms of normalised undrained shear strength ratio as a function of the plasticity index. The diagram on the left shows all the data (Aas et al. 1986). It is practically impossible to establish a relationship except that the strength ratio appears to increase with plasticity. If the data are then qualified in terms of the overconsolidation ratio of the clay tested, trends can easily be established. [These data also form the basis for the correction of the field vane strengths as a function of plasticity index proposed by Bjerrum (1972) and followed up by Aas (Aas et al, 1986)].



(a) All data (b) Only p'$_o$-consolidated data

Figure 10. Correlation structure with inconsistent and consistent data set

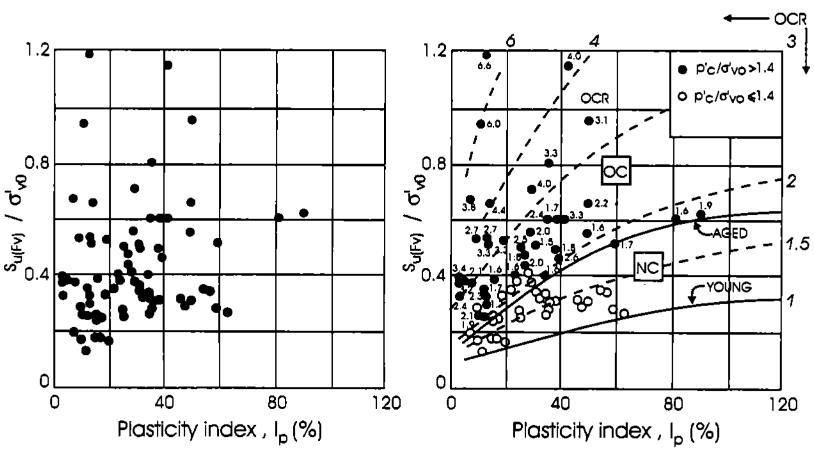


Figure 11. Selection of comparable databases

Model uncertainty

Including model uncertainty is not only necessary, but is absolutely more rational than ignoring it. One of the main reasons to place focus on model uncertainty is that it is generally large and it can be reduced.

Uncertainty that can arise from the choice of a calculation model is illustrated in Fig. 12. Current axial capacity calculation methods used for offshore piles have been derived predominantly from load tests on small piles. Penetration depth, pile length, pile diameter and ultimate load for the largest piles in the reference database are much smaller for the test piles than for those currently used. The uncertainty due to the

calculation model is therefore large because the reference database of pile load tests applies to different pile and load conditions than used in design. The linear extrapolation implied when using the calculation models has by no means been verified. The uncertainty due to this extrapolation needs to be included in the estimation of the possibility of a failure.

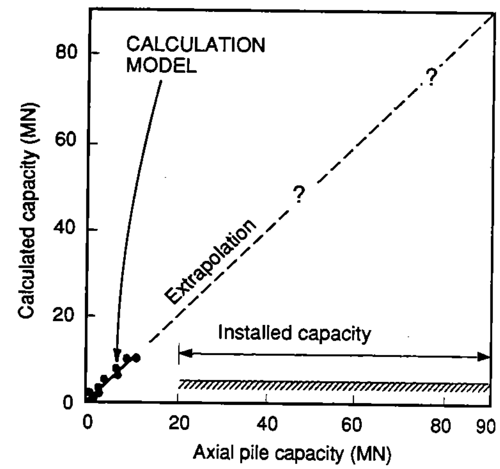


Figure 12. Extrapolation implied by calculation model for pile capacity

Model uncertainty is defined with a mean and a coefficient of variation, and usually a normal or log-normal distribution. Model uncertainty is best included in one of three ways: (1) factor on each random variable in the analysis, (2) factor on friction (each layer) and end bearing components and (3) global factor on the equation describing failure.

Model uncertainty can be evaluated from comparisons between model tests and deterministic calculations, pooling of expert opinions, case studies of prototypes or other model tests, results from the literature, and naturally engineering judgement. To estimate model uncertainty, the relevant mechanisms should first be identified. For example, pooling of 30 international experts on pile design gave the consensus that the currently most used pile design method [API RP2A method (API, 1993)] is conservative in medium dense to very dense sand (Lacasse and Goulois, 1989). In dense sands, uncertainties (coefficients of variation of 25% or more) were associated with most of the empirical design factors entering the calculation formulas.

On the other hand, when calculation models have been checked with repeated model tests under different modes of failure and give bias as those shown in Table 5, the model uncertainty is quite small.

Several factors affect model uncertainty. These are listed in Table 6.

Table 5. Comparisons of calculated and measured bearing capacities*

Structure	Type of loading	Bias calculated/measured failure loads
Shallow foundation	Static failure, test 1	0.98-1.01
	Cyclic failure, test 2	0.99-1.15
	Cyclic failure, test 3	1.16-1.17
	Cyclic failure, test 4	1.06-1.23
Tension leg platform	Static failure, test 1	1.00
	Cyclic failure, test 2	1.06
	Cyclic failure, test 3	1.06
	Cyclic failure, test 4	1.02

* (From Dyvik et al, 1989; Andersen et al, 1989; Andersen et al, 1992b; Dyvik et al, 1993; Andersen et al, 1993; also summarised in Lacasse and Nadim, 1994)

Table 6. Factors affecting model uncertainty

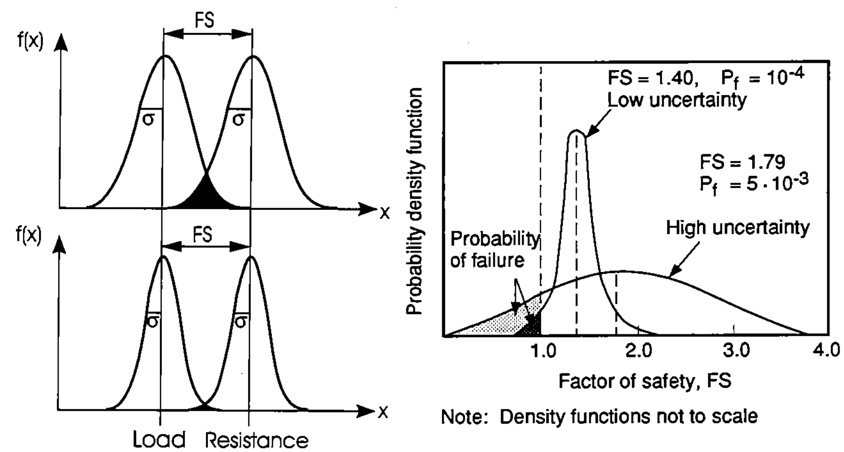
Property/calculations	Factors of influence
Undrained shear strength (clay)	• sampling disturbance • test method and scale of laboratory/in situ tests • spatial variability, anisotropy • rate of loading
Friction angle (sand)	• reconstitution of test specimen • density, test method and scale of laboratory test
Pile capacity	• skin friction assumption • limiting values for skin friction and end bearing • subdivision in soil layers • pile installation, residual stresses and plug condition • reconsolidation, rate of loading, cyclic loading, scour • stiffness of pile, pile length, single pile vs pile group • extrapolation from reference database to prototype
Shallow foundations	• position of critical slip surface • modelling of static and cyclic load history • strain-softening and/or progressive failure • testing procedures in reference tests • scale effect, rate of shear and stress conditions • redistribution of stresses and anisotropy • plane strain versus 3-D model, stiffness of structure • model of soil profile and drainage assumptions

Importance of characterising uncertainties for use in design

Figure 13a illustrates the difference between safety factor (SF) and probability of failure (P_f). Safety factor is the ratio of resistance to load, safety margin is the difference between resistance and load. For the two cases of load and resistance in the figure, the safety factor is the same, but the probability of failure is very different. The probability of failure is given by the size of the intersecting area between the resistance and the load. It represents the likelihood of a failure for a combination of the load and resistance parameters, each of the parameters being defined as random variables.

Three geotechnical examples illustrate the importance of characterising uncertainties in geotechnical analysis. The experience offshore, is that the loads and the model uncertainty, to a large extent are at least as important, or even more important, contributors to the overall uncertainty, and therefore, the probability of failure.

There are also other good examples to be found in the literature. Christian et al. (1994) presented a case where accounting for the uncertainties in the soil properties and the model uncertainty in the analysis of a slope stability provided a more meaningful measure of the safety margin than the deterministic factor of safety. The uncertainty in the soil properties represented a major contributor to the resulting safety.



(a) Safety factor vs probability of failure (b) Reliability of a pile foundation
(FS = factor of safety, σ = standard deviation)

Figure 13. Safety factor and probability of failure

Example 1. Reliability of pile foundation

Figure 13b presents the results of the reliability analysis of the most loaded pile in a jacket installed in 1976 and reanalysed in 1989 after a new soil investigation and new calculations of the environmental and gravity loads had been completed. The newer deterministic analysis gave a low safety factor (FS). The added information reduced the factor of safety, but at the same time reduced the uncertainty in both soil and load parameters.

The pile with a safety factor of 1.40 (in 1989) is nominally safer than the pile with a safety factor is 1.79 (in 1976). The 1989 analyses show that the pile, although with a lower safety factor, had higher safety margin than perceived at the time of design. The lower uncertainty in the parameters in 1989 led to a reduction in the probability of failure P_f by a factor of 2.

Factor of safety is not a sufficient indicator of safety margin because the uncertainties in the analysis parameters affect probability of failure. The uncertainties are not explicit in the deterministic calculation of safety factor.

Example 2. Stability of gravity foundations.

Reliability analyses were run for a shallow foundation offshore on a uniform soft plastic clay site. The slip surfaces in Fig. 14 were analysed. Spatial variability, which reduced the uncertainty in the averaged soil strength along the failure surface, was included. The coefficient of variation of the extreme environmental loads was taken as 15%; horizontal load and moment were taken as perfectly correlated.

The reliability analyses indicated that the slip surface with highest probability of failure was different from the critical slip surface obtained from deterministic analyses. Higher probability of failure corresponds to lower reliability index. Reliability index is the distance (in a 3-D multi-function representation) between the most probable response and the response for the random variables causing failure.

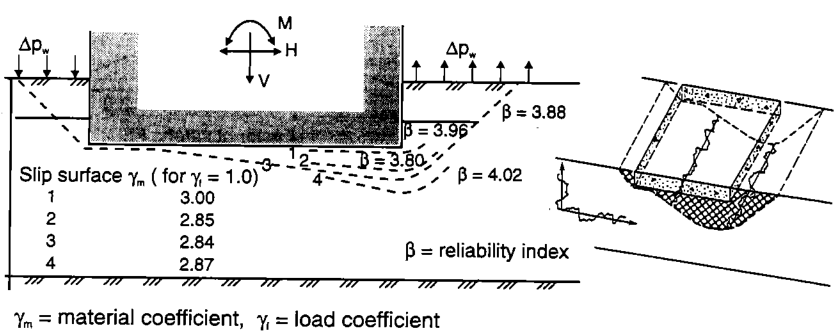


Fig. 14. Results of probabilistic analysis of bearing capacity of shallow foundation

This discrepancy has been often observed for different soil profiles. It illustrates the important effect of the uncertainty in the analysis parameters on factor of safety. It is often due to the different uncertainties associated with the passive and active strengths to use in the analysis. In the analyses of gravity structures on soft and stiff clays, model uncertainty was often one of the most significant uncertain variables. Changing the probability distribution of the soil parameters from normal to lognormal had only a modest effect on the computed probability of failure.

Example 3. Analysis of shallow foundation with two approaches

Probabilistic stability analyses were done using the «mobilised friction angle» approach (an effective stress approach) and the «available shear strength» approach (based on the undrained shear strength of the soil). The two approaches define factor of safety as follows: either as the ratio between the undrained shear strength and the shear stress mobilised for equilibrium, or as the ratio between the tangent of the characteristic friction angle and the tangent of the friction angle being mobilised at equilibrium. Both analysis methods are allowed in the Norwegian code of practice.

Shallow foundations on two soil types were considered: a contractive soil (loose sand, normally consolidated clay) and a dilative soil (dense sand, heavily overconsolidated clay). The «true» safety margin for the foundations should be independent of the method of analysis.

Table 6 presents the results of the calculations. Depending on soil type, the computed nominal probability of failure differed appreciably for the two approaches. The probabilistic and deterministic results showed significant differences, especially for the dilative soil, as the uncertainties in the soil properties interacted differently in each approach. For the «mobilised friction angle» approach, uncertainties in friction angle ϕ', cohesion, pore pressure parameter and submerged unit weight were considered. For the «available shear strength» approach, uncertainties in undrained shear strength and submerged unit weight were included. To «calibrate» the two analysis methods, a model uncertainty factor would have to be included.

Table 6. Results of stability analyses with two approaches (Nadim et al., 1994)

Soil type	Analysis method	Factor of safety	Probability of failure
Contractive	Mobilised friction angle	1.9	1.7×10^{-5}
	Available shear strength	1.4	2.5×10^{-3}
Dilative	Mobilised friction angle	1.4	6.7×10^{-3}
	Available shear strength	1.5	2.3×10^{-6}

Recommendation

It is important to document in our analysis what is accounted for and what is neglected, for example with a checklist, and to assess in such cases the risk and probability of non-performance involved. In the present days of increasing requirements from quality assurance and certification, a documentation that all possible aspects have been considered and dealt with in some manner is essential.

One important challenge for the engineer is combining the uncertainty components and doing a reliability analysis on which he can base his decisions. A reasonable approach could be to (1) determine the traditionally required reliability, (2) introduce new information, (3) calibrate the existing codes of practice, and (4) change the traditionally required reliability for the effect of new information and other unknowns that should be included in the analysis.

Figure 15 compares schematically the capacity obtained from a traditional design approach and a refined design approach. In both cases, the physical capacity is the same. The traditional design approach gives a calculated and design capacity, and a traditional safety margin (e.g. using the safety factor prescribed in codes). The refined design approach also gives a calculated and design capacity, a refined safety margin and calibrated safety factor, and an indication of the 'unnecessary' safety margin implied by the traditional approach. The aim, as we improve our design methods and reduce the uncertainties in soil properties, is to have the ranges *a*, *b* and *c* in Fig. 15 as narrow as possible. It then would be feasible to document why the factor of safety should be changed, because of increased information or increased accuracy of the calculation. An essential condition is the consistent definition of the characteristic values of the analysis parameters.

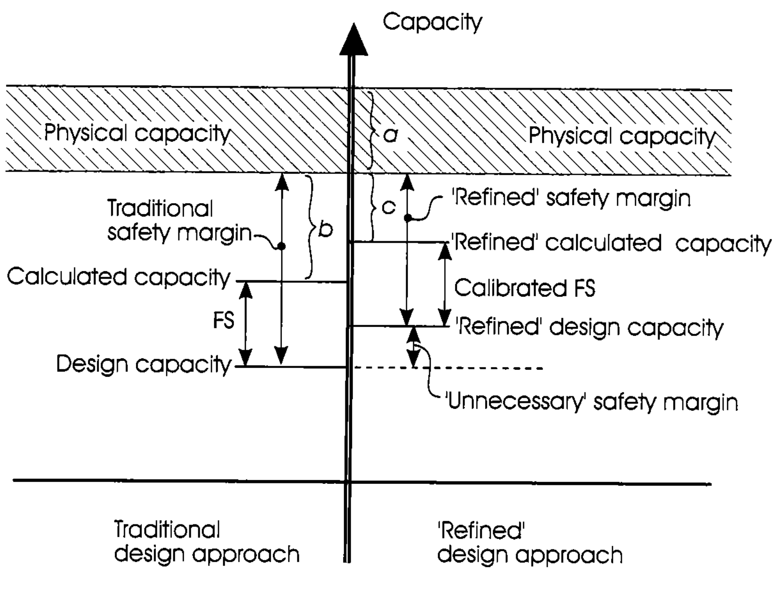


Figure 15. Reduction of uncertainty and «unnecessary» safety margin

It is necessary in a design to find the balance between appropriate technology and adequate complexity, given the budget and the consequences of the project. Approximating methods should be used, but not without understanding the complexity of the problem to model. In this respect, accounting for uncertainties and their effects on the response contributes to this understanding.

Conclusions

Basic statistics represent a useful means to establish the mean and variance of a soil property. In geotechnical problems that involve large soil volumes, the "spatially averaged" properties govern foundation analysis, as local fluctuations average over a large soil volume. Neglecting uncertainties does not mean that they do not exist. To characterise the uncertainties in soil properties, the engineer needs to combine, in addition to the actual data, knowledge on the quality of the data, knowledge about the geology and layering, especially engineering judgement; because in geotechnics, there will never be enough data.

Statistics and reliability analysis are a complement to more conventional deterministic analyses. The authors perceive as necessary to account for the uncertainties in the soil properties and the calculation model. The added knowledge on the soil properties and their uncertainties by doing a systematic uncertainty assessment should lead to safer and more economical designs.

The challenge faced by the engineer is a decision on how to account for the uncertainties in his design and taking the responsibility for the effects of these uncertainties on the soundness of the design.

When a sufficient quantity of data are available, site description strategy, with the identification of the correlation structure within a layer and stochastic interpolation to estimate a soil property is recommended. The geotechnical parameters for analysis are then more clearly defined, and at times the statistical uncertainties can be reduced. Geostatistics provide additional information to guide the interpretation of soil profiles and the variability in the soil properties. The analysis will also provide an estimate of the noise in the measurements and help discern actual in situ trends from anomalies.

Difficulties arise in the practical assessment of geotechnical uncertainties. There are often not enough data to evaluate the autocorrelation distance statistically. Model uncertainty, generally very important, is difficult to establish, although this is gradually improving as more well documented relevant case studies become available. The engineer still needs to rely heavily on engineering judgement. Using engineering judgement is not necessarily a problem. However it is essential to indicate where engineering judgement is used. Engineering judgement is one of the main contributions of the qualified geotechnical engineer. If uncertainty analysis and reliability concepts promote the need for analysing and documenting systematically the uncertainties and the need to treat each of them explicitly and consistently, the authors believe that this part of engineering science will have made an important contribution to practice.

Appendix I - Terms from statistical analysis

Autocorrelation distance	Distance expressing the rate of decay of the autocorrelation function. The autocorrelation distance depends on the autocorrelation function used (exponential, triangular, spherical, etc).
Autocorrelation function	Function describing the correlation of the residuals about a trend
Coefficient of variation	Ratio of standard deviation to mean value
Correlation	Mutual dependency between two or more variables
Fluctuation distance	Distance over which a soil property shows relatively strong correlation from point to point.
Histogram	Graphical representation of a range of measured or observed values and of how frequently these values occur (also called frequency diagram)
Linear regression	Linear relation between two random variables expressed in terms of mean and variance of one random variable as a function of the other variable
Mean	Measure of the most likely value of a random variable (also called average)
Median	Value of a variable at which values above and below it are equally probable
Population	Set of data points considered
Probability distribution	Law for describing the probability associated with each of the values of a random variable
Probability of failure	Probability that failure will occur under a combination of random load and resistance parameters
Random process	Process associated with the numerical outcome of random variable(s)
Random variable	Variable which exhibits scatter or dispersion and which value cannot be predicted with certainty; for each outcome of a random variable is associated a numerical value

Reliability index	Measure of distance between most probable response and the load and resistance parameter combinations causing failure; for a linear function and a Gaussian variable, the reliability index is the inverse of the normal distribution function
Residual	Algebraic measure of distance between the value of a data point and the value of the trend at the same location
Safety margin	Difference between physical resistance and maximum load effect (safety margin = resistance - load)
Scale of fluctuation	Distance within which a soil property shows correlation between two values of a random variable
Standard deviation	Measure of dispersion or variability of a random variable and of the closeness of the values; the standard deviation is the square root of the variance
Trend	Direction or tendency of a pair of variables (often the slope of a function)
Variance	Measure of dispersion or variability of a random variable, and of the closeness of the values; dispersion is taken with respect to the mean value

Appendix II - References

Aas, G., Lacasse, S., Lunne ,T. and Høeg, K. (1986). Use of In Situ Tests for Foundation Design on Clay. *Proc. ASCE In Situ '86*. Blacksburg, VA, USA. pp. 1-30.

Alonzo, E.E. and. Krizek, R.J. (1975). Stochastic formulation of soil properties, Proc. *2nd Conf on Application of Probability and Statistics to Soil and Structural Eng*, Aachen, Germany, Vol. II, pp. 9-32.

Andersen, K.H., Dyvik, R., Lauritzsen, R, Heien, D., Hårvik, L. and Amundsen, T. (1989). Model Tests of Gravity Platforms. II: Interpretation. ASCE. *Journal of Geotechnical Engineering*. V 115, No 11. pp. 1550-1568.

Andersen, K.H., Dyvik, R. and Schrøder, K. (1992b). Pull-Out Capacity Analyses of Suction Anchors for Tension Leg Platforms. *Proc BOSS'92 Conf.* London, U.K. V 2, pp. 1311-1322.

Andersen, K.H., Dyvik, R., Schrøder, K., Hansteen, O.E. and Bysveen, S. (1993). Field Tests of Anchors in clay II: Predictions and Interpretation. ASCE. *Journal of Geotechnical Engineering*. V 119, No 10, pp. 1532-1549.

Andersen, K.H., Lauritzsen, R., Dyvik, R. and Aas P.M. (1988). Cyclic Bearing Capacity Analysis for Gravity Platforms; Calculation Procedure, Verification by Model Tests, and Application for the Gullfaks C Platform. *Proc BOSS'88 Conf.* Trondheim, Norway. V 1, pp. 311-325.

Ang, A.H.S. and Tang, W. (1975). *Probability concepts in enginering planning and design. Volume I Basic Principles*. John Wiley and Sons, New York, USA, 409p.

API (American Petroleum Institute) (1993). *Recommended Practice for Planning, Designing and Constructiong Fixed Offshore Platforms - Load and resistance Factor Design.* API RP2A-LRFD 1st ed. Dallas, Texas, USA. Sept. 1993.

Baecher, G. (1985). Geotechnical error analysis. Proc. MIT Special Summer Course: Recent Developments in measurement and modelling of clay behaviour for foundation design. MIT, Cambridge, USA, Lecture 16.

Bjerrum, L. (1972). Embankments on Soft Ground. SOA Report. *ASCE Spec. Conf. Performance of Earth and Earth-Supported Structures*. Purdue Univ., Lafayette, Indiana. Vol. 2, pp. 1-54.

Chiasson, P., Lafleur, J., Soulié, M. and Law, K.T. (1995). Characterizing spatial variability of a clay by geostatistics. *Canadian Geotechnical Journal*, Vol. 32, No. 1, pp. 1-10.

Christian, J.T., Ladd, C.C. and Baecher, G.B. (1994). Reliability Aplied to Slope Stability Analysis. *ASCE*, , Vol. 120, No. 12, pp. 2180-2207.

De Groot, D.J. and Baecher, G. (1993). Extimating autocovariance of in-situ soil properties. *ASCE, Journal of Geotechnical Engineering*, Vol. 119, No. 1, pp. 147-167.

Dyvik, R., Andersen, K.H., Madshus, C. and Amundsen, T.(1989). Model Tests of Gravity Platforms I: Description. ASCE. *Journal of Geotechnical Engineering.* V 115, No 11. pp. 1532-1549.

Dyvik, R., Andersen, K.H., Hansen, S.B. and Christophersen, H.P. (1993). Field Tests of Anchors in Clay I: Description. ASCE. *Journal of Geotechnical Engineering*. V 119, No 10, pp. 1515-1531.

Einstein, H.H. and Baecher, G.B. (1982). Probabilistic and statistical methods in engineering geology. I. Problem statement and introduction to solution. *Rock Mechanics, Supp.* 12, pp.47-61.

Harder, H. and van Bloh, G. (1988). Determination of representative CPT-parameters. Proc. *Penetration testing in the UK*, Thomas Telford, London, pp. 237-240.

Høeg, K. and Tang, W. (1976). Probabilistic considerations in the foundation engineering of offshore structures. Proc. *2nd ICOSSAR*, Aachen, Germany, 29 p. (also published in *Norwegian Geotechnical Institute* Publ. # 120, Oslo, Norway.

Keaveny, J., Nadim, F. and Lacasse, S. (1989). Autocorrelation functions for offshore geotechnical data, Proc. *5th ICOSSAR*, San Francisco, USA, pp.263-270.

Krumbein, W.C. and Grayhill, F.A. (1965). *An introduction to statistical models in geology*, McGraw Hill, New York, 45 p.

Lacasse, S. and de Lamballerie, J.Y. (1995). Statistical Treatment of CPT Data. *Proc. CPT'95*. Linkøping, Sweden (in print).

Lacasse, S. and Goulois, A. (1989). Uncertainty in API Parameters for Predictions of Axial Capacity of Driven Pile in Sand. 21st *OTC* Houston, Texas, USA. Paper 6001, pp. 353-358.

Lacasse, S. and Nadim, F. (1994). Reliability Issues and Future Challenges in Geotechnical Engineering for Offshore Structures. Plenum presentation *7th Int. Conf. Behaviour of Off-Shore Structures BOSS'94*, MIT. Mass. USA.

Matheron, G. (1963). Principles of geostatistics, *Economic Geology*, Vol. 58, pp. 1246-1266.

Mortensen, J.K., Hamsen, G. and Sørensen, B. (1991). Correlation of CPT and field vane tests for clay tills. *Bulletin* 7, Danish Geotechnical Society, 62 p.

Nadim, F. (1988). Geotechnical site description using stochastic interpolation. *10th NGM-Conf*, Oslo, pp. 158-162.

Nadim, F., Lacasse, S. and Guttormsen, T.R. (1994). Probabilistic foundation stability analysis: Mobilized friction angle vs available shear strength approach. *Proc. Stuctural Safety and Reliability*, (Schueller, Shinozuka and Yao eds.) Balkema, Rotterdam, NL. pp.2001-2008.

Snedecor, G.W. and Cochran, W.G. (1964). *Statistical methods*. Univ of Iowa Press, 140 p.

Tang. W. (1979). Probabilistic evaluation of penetration resistance, *ASCE*Vol. 105, No. GT10, pp.1173-1191.

Vanmarcke, E. (1977), Probabilistic modelling of soil profiles, *ASCE, Journal of Geotechnical Engineering Division*, Vol. 103, No. GT11, pp. 1227-1246.

Vanmarcke, E. (1984), *Random fields*, MIT Press, Cambridge, Mass, USA, 382 p.

Vivatrat, V. (1978). Cone penetration in clays. PhD Thesis, MIT, Cambridge, Mass. USA.