Geotechnical Engineering Reliability: How Well Do We Know What We Are Doing?[1]

John T. Christian, Hon.M.ASCE[2]

Abstract: Uncertainty and risk are central features of geotechnical and geological engineering. Engineers can deal with uncertainty by ignoring it, by being conservative, by using the observational method, or by quantifying it. In recent years, reliability analysis and probabilistic methods have found wide application in geotechnical engineering and related fields. The tools are well known, including methods of reliability analysis and decision trees. Analytical models for deterministic geotechnical applications are also widely available, even if their underlying reliability is sometimes suspect. The major issues involve input and output. In order to develop appropriate input, the engineer must understand the nature of uncertainty and probability. Most geotechnical uncertainty reflects lack of knowledge, and probability based on the engineer's degree of belief comes closest to the profession's practical approach. Bayesian approaches are especially powerful because they provide probabilities on the state of nature rather than on the observations. The first point in developing a model from geotechnical data is that the distinction between the trend or systematic error and the spatial error is a modeling choice, not a property of nature. Second, properties estimated from small samples may be seriously in error, whether they are used probabilistically or deterministically. Third, experts generally estimate mean trends well but tend to underestimate uncertainty and to be overconfident in their estimates. In this context, engineering judgment should be based on a demonstrable chain of reasoning and not on speculation. One difficulty in interpreting results is that most people, including engineers, have difficulty establishing an allowable probability of failure or dealing with low values of probability. The $F-N$ plot is one useful vehicle for comparing calculated probabilities with observed frequencies of failure of comparable facilities. In any comparison it must be noted that a calculated probability is a lower bound because it must fail to incorporate the factors that are ignored in the analysis. It is useful to compare probabilities of failure for alternative designs, and the reliability methods reveal the contributions of different components to the uncertainty in the probability of failure. Probability is not a property of the world but a state of mind; geotechnical uncertainty is primarily epistemic, Bayesian, and belief based. The current challenges to the profession are to make use of probabilistic methods in practice and to sharpen our investigations and analyses so that each additional data point provides maximal information.

DOI: 10.1061/(ASCE)1090-0241(2004)130:10(985)

CE Database subject headings: Failures; Geotechnical engineering; Probability; Reliability; Statistics; Uncertainty analysis; Risk analysis.

Introduction

Uncertainty and reliability have a long history in geotechnical engineering. Even before there was a distinct discipline of geotechnical engineering, engineers who dealt with soils, rocks, and geological phenomena knew they were involved in an uncertain venture and that they had to provide for untoward developments. They have also been the focus of controversy. One of Terzaghi's most famous early papers (Terzaghi 1929) emphasizes the importance of minor geologic details—that is, features that differ from the expected or mean conditions. He criticized then current practice for "… blindly trusting in purely statistical relations with an extraordinary wide range of deviation to both sides from the average. As most of the textbooks fail to call the attention of the readers to the great uncertainty associated with the rules of design based on this practice, many engineers engaged in dam design have an exaggerated conception of the reliability of their methods of procedure and as a consequence, progress in this field came practically to a standstill." He recommended that designers "assume… the most unfavorable possibilities."

As the discipline developed, it became clear that it was seldom technically or economically possible to design for the most unfavorable possibilities, and Terzaghi himself proposed what he called the "learn-as-you-go" approach. Peck (1969) codified and expanded the approach, which he named the "observational method" and which is now an essential feature of geotechnical practice, especially for large or difficult projects. The observational method is a practical way to deal with uncertainty that is closely related to the techniques of Bayesian updating. The environmental management community has developed a similar technique, called "adaptive management," in which the design, construction, and operation of environmental remediation facilities are modified during the course of the project as additional observations become available. All of these approaches—the observa-

[1]The Thirty-Ninth Terzaghi Lecture presented at the ASCE 2003 Civil Engineering Conference and Exposition, Nashville, Tenn., November.

[2]Consulting Engineer, 23 Fredana Rd., Waban, MA 02468.

Note. Discussion open until March 1, 2005. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on November 18, 2003; approved on March 29, 2004. This paper is part of the ***Journal of Geotechnical and Geoenvironmental Engineering***, Vol. 130, No. 10, October 1, 2004. ©ASCE, ISSN 1090-0241/2004/10-985–1003/$18.00.

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

tional method, Bayesian updating, adaptive management—require that the project managers acknowledge uncertainty from the beginning and that the public accept the existence of uncertainty.

Although the geotechnical community long ago learned practical ways to deal with uncertainty, it has been reluctant to embrace the more formal and rational approaches of reliability theory while other fields of civil engineering have made major commitments to probabilistic approaches. Nevertheless, over the last ten or fifteen years several researchers have made major advances in applying probabilistic methods to geotechnical problems. The author is indebted to their work and to many discussions with most of them. In particular, he acknowledges his debt to, in alphabetical order, G. B. Baecher, C. A. Cornell, H. E. Einstein, M. E. Harr, F. L. Kulhawy, P. Lumb, W. H. Tang, E. Vanmarcke, D. Veneziano, S. G. Vick, and T. H. Wu. The present work has also benefited enormously from the collaboration with Professor Baecher that went into writing our recently published book (Baecher and Christian 2003b). Questions of risk and reliability have been the basis of at least three previous Terzaghi lectures (Casagrande 1965; Whitman 1984; S. Lacasse, 2001 Terzaghi Lecture, unpublished).

Uncertainty in Geotechnical Engineering

Blaise Pascal, the first major contributor to probability theory, set out in the 17th century the fundamental principle underlying reliability analysis: "We ought to fear or hope for an event not only in proportion to the advantage or disadvantage but also with some consideration of the likelihood of the occurrence" (Hacking 1975). Most people find this idea intuitively reasonable. We should concern ourselves with situations that have large consequences but also with those that are most likely to take place. How to distribute our attention over the full range of insignificant to significant events with small to large likelihood of occurrence is, of course, an essential engineering problem. Put another way, much of engineering is about how to deal with uncertainty, although one does not always need to understand uncertainty to deal with it.

Geotechnical engineers, like engineers in other disciplines, have developed several strategies for dealing with uncertainty. They include:

1. *Ignoring it*. While on its face such a head-in-the-sand approach would seem insupportable, it is surprisingly widespread. There are many stories of agencies and corporations that willfully ignored warnings that the assumptions underlying their decisions were fraught with uncertainty. Surely one of the earliest involved the British and Dutch governments, who in the 17th and 18th centuries sold annuities to finance their expenses (usually wars) and dismissed out of hand the objections of early statisticians that the annuities were actuarially unsound (Gigerenzer et al. 1989).
2. *Being conservative*. This is an obvious and frequently sound approach. Rather than get involved in the details of how often undesirable things might happen and what their consequences might be, the engineer makes the structure or system so robust that it will resist anything. While this works in many cases, it is usually expensive, it may drag the project out to unacceptable completion times, and in some cases it may simply not be possible. Eventually one must ask how conservative is conservative enough.
3. *Using the observational method*. The observational method has established itself as the preferred way for geotechnical engineers to deal with uncertainty in situations for which simple conservatism is unsatisfactory (Peck 1969). It involves (1) considering possible modes of unsatisfactory performance or other undesirable developments; (2) developing plans for dealing with each such development; (3) making field measurements during construction and operation to establish whether the developments are occurring; and (4) reacting to the observed behavior by changing the design or construction process. While the observational method has made it possible to carry out many projects that would have been impossible under conventional conservative procedures, it has limitations. The engineer must have access to the decision maker if the design or construction sequence is to be changed in mid-project, the usual applications do not consider explicitly the relative likelihood of the undesirable occurrences, and field measurements are expensive. Another factor limiting wider use of both the observational method and reliability-base approaches is that some regulatory agencies and the public often demand what they consider certainty at the outset of a project.
4. *Quantifying uncertainty*. This is the purpose of reliability approaches. Quantifying the uncertainty is consistent with the philosophy of the observational method; it might be considered a logical extension of the observational method that accommodates modern developments in probabilistic methods. It is central to this lecture.

Other disciplines have developed techniques that closely resemble the observational method with or without probabilistic input. An especially notable case is the already mentioned adaptive management approach widely used in environmental and ecological management. Whatever they are called, such methods require carefully thought out programs for field measurements and explicit determination of what the measurements are going to achieve. Leps (1987) explained

> Probably the only reasonable role for monitoring systems is to provide confirmation or denial of the routine performance characteristics anticipated in design. The often discussed role of providing advanced warning of impending failure is, in the present writer's judgment, simply impractical for several reasons, the most important of which being that it is usually totally impossible to pinpoint where failure may begin. The second and overwhelmingly important point is that if one actually thinks he knows where failure is most apt to occur, he is completely derelict if he has not provided a design which would eliminate such possibility.

Perhaps the first significant attempt to document an approach for dealing with uncertainty and risk in geotechnical engineering was Arthur Casagrande's 1964 Terzaghi lecture, which was published in 1965. He reported that he looked carefully into the various definitions of calculated risk that had been proposed over the years and settled on the following:

1. The use of imperfect knowledge, guided by judgment and experience, to estimate the probable ranges for all pertinent quantities that enter into the solution of a problem; and
2. The decision on an appropriate margin of safety, or degree of risk, taking into consideration economic factors and the multitude of losses that would result from failure.

That definition is adopted here.

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

Current Geotechnical Applications of Probabilistic Methods

Although many geotechnical professionals have been reluctant to embrace probabilistic methods, the techniques have found application in many areas. Prominent among these are

Design, Construction, and Operation of Offshore Platforms for Petroleum Industry

There is a large literature on the subject, and Suzanne Lacasse's Terzaghi lecture (S. Lacasse, 2001 Terzaghi Lecture, unpublished) provides an up-to-date summary of the state of the art. Among the seminal contributions have been those of Bea (1999) and the group at the Norwegian Geotechnical Institute (Lacasse and Nadim 1996).

Studies of Safety of Dams, Dikes, and Embankments

These range from detailed evaluations of the probability of failure or selection of appropriate factors of safety for particular slopes (Christian et al. 1994; El-Ramly et al. 2003b) to studies of various failure modes for dam systems (Von Thun 1996; Vick 2002). Studies of the stability of specific slopes, or sets of slopes, generally use one of the methods that lead to a reliability index β and a probability of failure p_f. The analysis of complete dam systems usually employs event trees or fault trees. Often the results of reliability analyses of individual components are used as input to the branches of the event or fault trees. It is worth noting that some agencies, such as the U.S. Bureau of Reclamation and BC Hydro, have embraced probabilistic safety analysis of their dams to the extent that it can be carried out defensibly, while others, such as the U.S. Federal Energy Regulatory Commission, remain adamantly opposed to probabilistic analysis. Other agencies, notably the U.S. Army Corps of Engineers, have adopted policies that fall somewhere between these two positions.

Probabilistic Seismic Hazard Analysis

Today almost all estimates of seismic hazard, whether developed for a specific project or presented in the form of maps for use in developing building codes, are based on probabilistic approaches. Most analyses use the basic approach developed by Cornell (1968), often with considerable elaboration to incorporate the elicitation of expert opinion (Budnitz et al. 1997, 1998). Although these analyses provide probabilistic descriptions of the seismic hazard, it is ironic that the results are usually used deterministically in subsequent engineering analyses.

Mining

Designs of open pit mine slopes and underground excavations have always involved tradeoffs between cost on the one hand and reliability on the other. Hoek (1998) provides a brief exposition of reliability methods suitable for underground openings. Riela et al. (1999) and Calderon et al. (2003) describe the application of reliability methods for studying the stability of open pit mines.

Nuclear Waste Repositories

Probabilistic estimates of potential future behavior of the waste repositories have been central to their evaluation. There is a large literature on the subject (U.S. Nuclear Regulatory Commission 1975).

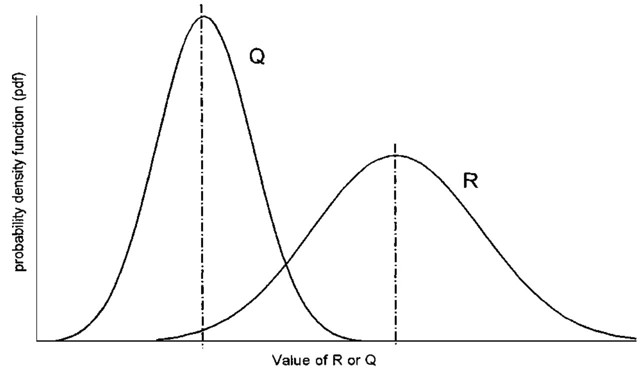

Fig. 1. General configuration in which load Q and resistance R are uncertain, where both have normal distributions, but that is not necessarily always case

Limit State Design or Load and Resistance Factor Design

These methodologies represent attempts to apply probabilistically based methods to routine design procedures. They have been used successfully in structural engineering, but their application in geotechnical engineering, especially foundation engineering, has been controversial. A great deal of research is underway, and there has been much discussion between researchers and practitioners. This is a topic that deserves extensive treatment, but space and time do not permit further discussion here.

Tools of Reliability Analysis

The tools available to the engineer for performing a reliability analysis fall into three broad categories. First are the methods of direct reliability analysis. These propagate the uncertainties in properties, geometries, loads, water levels, etc. through analytical models to obtain probabilistic descriptions of the behavior of a structure or system. The second includes event trees, fault trees, and influence diagrams, which describe the interaction among events and conditions in an engineering system. The third includes other statistical techniques. In particular, some problems are so poorly defined that it is useless to try to formulate mechanical models and the engineer must rely on simple statistics. Examples are extrapolation of landslide incidence in broad areas and studies of the behavior of Karst terrains. In practice, analysis of a specific system or structure usually involves a combination of methods appropriate to the problem at hand. Baecher and Christian (2003b) provide detailed descriptions of how these techniques operate.

Direct Reliability Analysis

If there are a loading Q and a resistance R, the margin of safety M is

$$M = R - Q \tag{1}$$

If both Q and R are uncertain, so is M (Fig. 1). Elementary probability theory then provides that the means (μ) and the standard deviations (σ) are related by

$$\mu_M = \mu_R - \mu_Q \tag{2}$$

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

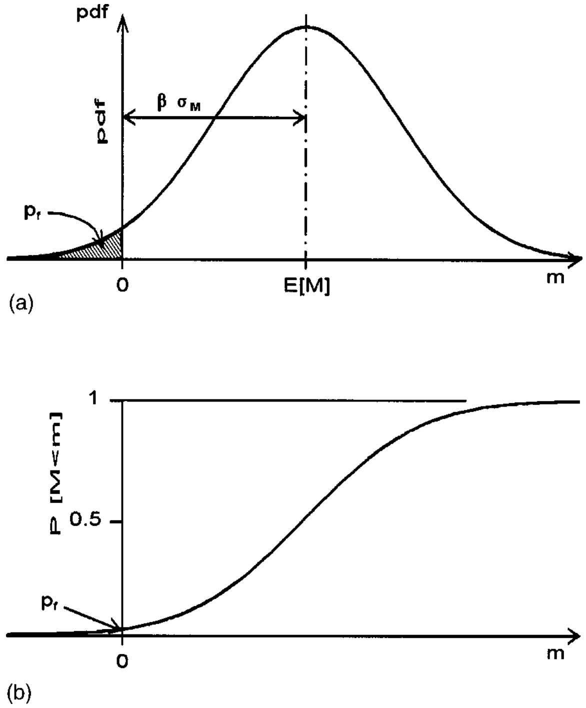

Fig. 2. Distribution of margin of safety $M(=R-Q)$: (a) probability density function and definition of reliability index β; (b) cumulative distribution function. Probability of failure is shaded area in (a) and intersection of cumulative distribution function with vertical axis in (b).

$$\sigma_M^2 = \sigma_R^2 + \sigma_Q^2 - 2\rho_{QR}\sigma_R\sigma_Q$$

in which ρ_{QR} = correlation between Q and R. If Q and R are not correlated, the last equation reduces to

$$\sigma_M^2 = \sigma_R^2 + \sigma_Q^2 \tag{3}$$

It is often more convenient to work with the logarithms of Q and R. Then the factor of safety F is the ratio R/Q, so

$$\ln F = \ln R - \ln Q \tag{4}$$

Since this is similar to Eq. (1), we can work with the logarithms of the variables provided we use ln F and the logarithms of the variables. It is customary to define λ and ζ as the mean and the standard deviation of the logarithms of a variable. It follows that, for any distribution (Aitchison and Brown 1969)

$$\zeta^2 = \ln\left(1 + \frac{\sigma^2}{\mu^2}\right)$$
$$\lambda = \ln\mu - \frac{1}{2}\zeta^2 \tag{5}$$

The essence of reliability methods is to recognize that the condition $M=0$ (or $\ln F=0$) corresponds to failure, so the problem is to find the probability that $M \leqslant 0$. As illustrated in Fig. 2, we now define a reliability index β as

$$\beta = \frac{\mu_M}{\sigma_M} \tag{6}$$

It follows that, if we are working with uncorrelated variables

$$\beta = \frac{\mu_R - \mu_Q}{\sqrt{\sigma_R^2 + \sigma_Q^2}} \tag{7}$$

or, for the case of the logarithms of uncorrelated variables

$$\beta = \frac{\lambda_R - \mu\lambda_Q}{\sqrt{\zeta_R^2 + \zeta_Q^2}} = \frac{\ln[(\mu_R/\mu_Q)\sqrt{(1+\Omega_Q^2)/(1+\Omega_R^2)}]}{\sqrt{\ln[(1+\Omega_R^2)/(1+\Omega_Q^2)]}} \tag{8}$$

in which Ω $(=\sigma/\mu)$ = coefficient of variation. As Fig. 2(a) shows, the probability of failure is the area under the probability density function of M lying to the left of $M=0$.

Now, if Q and R are both Normally distributed, so is M. Then it follows that the probability of failure is

$$p_f = \Phi(-\beta) \tag{9}$$

where Φ = cumulative distribution function (CDF) of the standard normal distribution and β is defined by Eq. (7). If Q and R are lognormally distributed, so is M, and Eq. (9) again applies, only with β defined by Eq. (8). In the past, evaluating the CDF required interpolation in tables, but today the CDF is a library function in spreadsheets and in mathematical software packages like *Mathcad* or *Matlab*.

While the model described by Figs. 1 and 2 and Eqs. (7)–(9) is conceptually straightforward, calculating the various means and standard deviations is anything but simple. Furthermore, distributions other than normal or lognormal arise often in practice. In situations where finite minimum and maximum values exist, one of the Beta distributions may be appropriate; problems involving recurrence of events usually lead to distributions like the exponential or Poisson. Methodologies based on normal or lognormal distributions must be modified when other distributions exist, but the underlying theory remains similar even while the details become more complicated.

Several methods for dealing with reliability models have evolved over the years:

First Order Second Moment Methods

The idea here is that, if we know the means and the variances (the second moments) of the variables that enter into the evaluation of a function such as M, we can estimate the mean and variance of M using only first order terms in a Taylor expansion (Cornell 1969)

$$\mu_M \approx M(\mu_{x_1}, \mu_{x_2}, \ldots, \mu_{x_n})$$
$$\sigma_M^2 \approx \sum_{i=1}^{n}\left(\frac{\partial M}{\partial x_i}\right)^2 \sigma_{x_i}^2 \tag{10}$$

in which the x_i = uncertain variables. Eq. (10) applies when the variables are uncorrelated; a somewhat more complicated expression is used when some of the variables are correlated. When it is difficult to evaluate the partial derivatives directly, central divided partial differences usually provide sufficient accuracy.

First Order Reliability Method

One shortcoming of the first order second moment (FOSM) approach is that the results depend on the particular values of the variables x_i at which the partial derivatives are calculated. Hasofer and Lind (1974) proposed to resolve this difficulty by evaluating the derivatives at the critical point on the failure surface. Finding this point usually requires iteration, but the process tends to converge rapidly. If the variables are all normalized by dividing them by their respective standard deviations, the distance between

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

the failure point and the point defined by their normalized means is the reliability index β. This method assumes normal distributions and must be modified to accommodate other distributions.

Point–Estimate Methods

The variance of a function—or any of its moments—is essentially the result of integration. Rosenblueth (1975, 1981) proposed that an accurate approximation is obtained by evaluating the function *M* at a set of discrete points and using those values to compute the desired moments. In practice, for uncorrelated variables, the points are usually taken at plus or minus one standard deviation from the mean of each of the variables. Other schemes can be used, especially when the variables are correlated or skewed. The method is a form of Gaussian quadrature (Christian and Baecher 1999).

Monte Carlo Simulation

Monte Carlo simulation enjoys a long history and a rich literature. Each continuous variable is replaced by a large number of discrete values generated from the underlying distribution; these values are used to compute a large number of values of function *M* and its distribution. The large numbers of computations once presented a constraint on the use of this method, but cheap modern computers have largely removed this obstacle. There are also several serious questions of convergence and of randomness in the generated variables. Several so-called variance reduction schemes can be effective in improving convergence and reducing computational effort. Fishman (1995) provides one of many treatments of the method. Monte Carlo simulation with variance reduction is particularly helpful in improving the accuracy of first order reliability method (FORM) results (Baecher and Christian 2003b).

Others

Perhaps the most significant methods other than those just described are the second order second moment and second order reliability method, which provide higher order approximations than those underlying FOSM and FORM. While these have found some applications in structural reliability studies, they have not found much application in geotechnical work.

Event Trees, Fault Trees, and Influence Diagrams

Event trees, fault trees, and influence diagrams are techniques for describing the logical interactions among a complex set of events, conditions, physical parameters, and physical states. In this context, there is no logical difference between an "event," such as the occurrence of an earthquake or of a large storm, and a "condition," such as the existence of a liquefiable layer of soil or the presence of erodible material in an earth dam.

Event trees (Fig. 3) start with an initiating event, such as, say, the occurrence of an earthquake. Then the analyst develops a set of events that could follow; say the peak ground acceleration could fall within a certain range. Associated with each range is a conditional probability; for example, for the range 0.05–0.10 g the conditional probability could be 25%. These events must be exhaustive—that is, all possible outcomes are included—and exclusive—that is, no possible result could fall within more than one outcome. The analysis then proceeds along each path to evaluate the next outcomes, and so on and so forth. At each stage the probabilities are conditional; that is, they are the probabilities of the current event if all preceding events in that branch have occurred. At the end of the tree, the probability of each outcome is simply the product of the conditional probabilities. Event trees

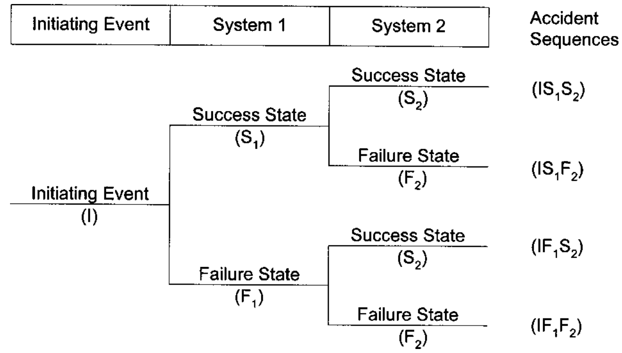

Fig. 3. Simple, generic event tree. Tree for actual situation would have many more branches (U.S. Nuclear Regulatory Commission 1975)

have been used to study the reliability of dams (Vick and Stewart 1996; Von Thun 1996), tank farms on liquefiable soil (T. W. Lambe & Associates 1982, 1989), and other engineered systems.

Fault trees (Fig. 4) start with the failure and work backward. The tree contains the conditions that must be met for the failure to occur. There are two basic situations. If all the conditions must be met, they are connected to the event by an "and" gate; if the event will occur if one or more of the conditions are met, they are connected by an "or" gate. The analyst develops the tree from the top down, moving from condition to condition. In the usual formulation, the conditions at each stage must be independent and must encompass all the conditions that could lead to the next stage. To compute the probability of failure, the analyst works from the bottom up. The effect of an "and" gate is that the probability of occurrence of a stage is the product of the probabilities for events feeding into the gate ($p=p_1p_2\cdots p_n$). The effect of an "or" gate is that the probability of occurrence of a stage is 1 minus the product of the probabilities of nonoccurrence of the events leading into the gate $[p=1-(1-p_1)(1-p_2)\cdots(1-p_n)]$. Fault trees have also been used in geotechnical practice (Van Zyl et al. 1996).

The influence diagram (Fig. 5) displays the relations between various events and conditions in a system. The direction of the arrows and other conventions represent the dependencies between the objects.

Other Techniques

Many other statistical and probabilistic tools exist, and most will find some applications in geotechnical engineering. One important case arises when the mechanics of a problem are not understood well enough to permit detailed modeling. The detailed mechanisms of failures of slopes along highway rights of way or in a Karst terrain are not really responsive to conventional slope stability analysis. The probability of failure is best estimated by compiling statistics on the numbers and magnitudes of failures that have been observed over time and developing a probability distribution that describes the observations.

Requirements for Reliability Analysis

An engineer faced with the task of evaluating the reliability of a facility, structure, or system must address four issues: the nature of the input uncertainties, the methodology for reliability analysis, the geotechnical analytical models, and interpreting the output.

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

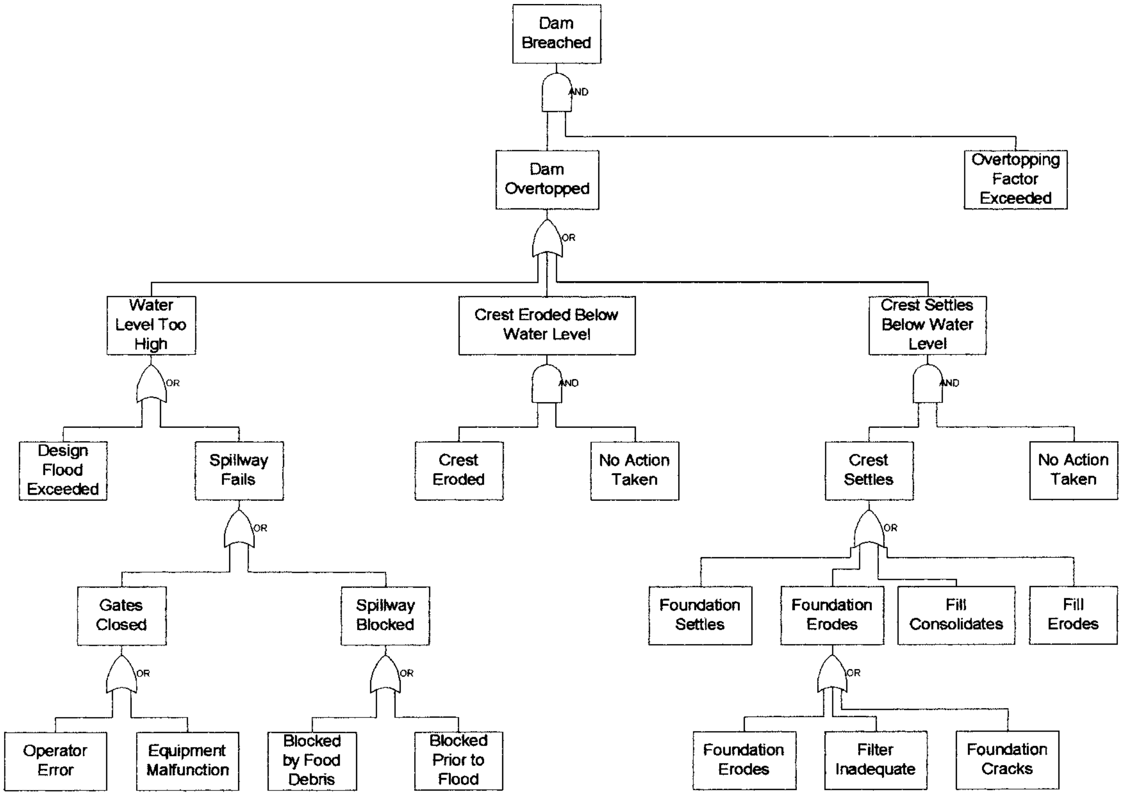

Fig. 4. Fault tree for analysis of dam breaching due to overtopping (after Parr and Cullen 1988)

The previous sections contain a brief description of the techniques for reliability analysis—techniques that are well established and thoroughly described in the literature. In general, the methodologies for reliability analysis are robust, their strengths and limitations well established.

The tools of geotechnical analysis for most practical problems are also well known. While some are well founded in theory and practice and will introduce little model error into the reliability calculations, others have large—and largely unknown—errors. For example, the widely used shallow bearing capacity equations depend on some dubious assumptions about plastic strains associated with the Mohr–Coulomb yield equation and on combining minimal solutions for three different factors (Christian and Urzua 1996). Furthermore, each of these factors is multiplied by up to five correction factors to account for shape and eccentricity, inclined loading, depth of embedment, base tilt, and ground slope (Meyerhof 1953, 1963; Vesic 1973, 1975). The uncertainties and ranges of validity in the three basic factors and fifteen correction factors are poorly understood, and it is far from clear that multiplying them together gives accurate results. Serious questions can be raised about the errors introduced by many other common analytical models. Suffice it to say that anyone using an analytical tool should be aware of the potential for error due solely to the inadequacies of the model.

Attention will now be directed at the first and fourth of the requirements for reliability analysis: input and output. The discussion of input requires some introduction to the meanings of uncertainty and probability and the difficulties involved in describing what we know. The discussion of output centers on how to understand what the output means and what can be learned from its details.

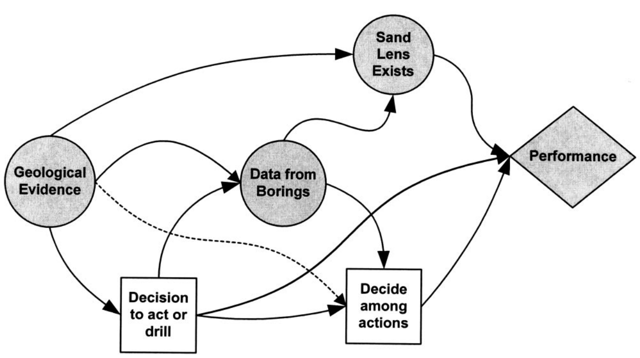

Fig. 5. Influence diagram for two-stage exploration decision, based on forthcoming Canadian Electricity Association Guide to Dam Safety Risk Management

Input—What are Uncertainty and Probability?

The input to any reliability analysis includes descriptions of the relevant parameters describing physical properties, loads, and geometry and of their uncertainties. Usually these are in the form of means and variances or standard deviations or probabilities of occurrence. However, before the engineer seizes values of the probabilistic parameters and leaps into the mechanics of a reliability analysis, he or she should have some understanding of the issues that have been raised about the nature of uncertainty and probability and how these issues affect the way one deals with uncertainty. Many of these issues arise again in interpreting the output of a reliability analysis.

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

Table 1. Terms Used in Literature to Describe Dual Meaning of Uncertainty, after Baecher and Christian (2003b)

Uncertainty due to naturally variable phenomena in time or space	Uncertainty due to lack of knowledge or understanding of nature	Reference citation
Aleatory uncertainty	Epistemic uncertainty	Hacking 1975; McCann 1999
Natural variability	Knowledge uncertainty	National Research Council 2000
Random or stochastic variability	Functional uncertainty	Stedinger et al. 1996
Objective uncertainty	Subjective uncertainty	Chow et al. 1988
External uncertainty	Internal uncertainty	Chow et al. 1988
Statistical uncertainty	Inductive probability	Carnap 1936
Chance	Probability	Poisson, Cournot (Hacking 1975)

Nature of Uncertainty

What exactly do we mean when we say that something is uncertain? Do we mean that the thing occurs at random in some unpredictable way, like the roll of a set of dice? That is, is the thing so unpredictable that additional knowledge or analysis will not affect our ability to estimate it?[1] This type of uncertainty is now known as aleatory, after the Latin word for gambler or dice thrower (Hacking 1975). Alternatively, we might mean that the thing is uncertain only in the sense that we do not know enough about it. For example, after a deck of playing cards is shuffled, the arrangement of the cards is fixed but unknown. We could discover the arrangement by simply examining each card in turn. However, that is precisely what we are not allowed to do, so the strategy in a game such as Bridge is to discover the arrangement by observation and induction. The uncertainty is due to lack of knowledge. This type of uncertainty is called epistemic, after the Greek word for knowledge (Hacking 1975). Table 1, based on a table compiled for analysis of flood risk (National Research Council 1995), presents seven pairs of alternate definitions proposed over the years (Baecher and Christian 2003b). The words "aleatory" and "epistemic" have achieved wide circulation and application, so they will be used here.

It will immediately be clear that the problem of establishing the geometry and properties of geologic deposits is closer to that of determining the arrangement of a deck of cards than it is to predicting the throw of a set of dice. Jensen (1997) was one of the first to point out the analogy between the configuration of geologic formations and the order of cards in a deck. In effect, the problem facing the geotechnical or geological engineer is epistemic rather than aleatory; it follows more from a lack of knowledge about materials and geometries than from inherent randomness in them.

Aleatory and epistemic uncertainties must be treated differently. If something is uncertain in the epistemic sense, the uncertainty may be reduced by additional information. Closer attention to the bidding and play of the hand in Bridge or additional exploration and testing in geotechnical engineering may reduce the epistemic uncertainty. It may not eliminate it, and the cost of reducing it below some level may not be worth it, but, in general, more information tends to reduce epistemic uncertainty. Conversely, more information will not reduce aleatory uncertainty, although it may establish more precisely the parameters governing that uncertainty. Veneziano (1995) has described the implications of the distinction between aleatory and epistemic uncertainty and how these affect the trade-offs that must be made in analysis.

Meaning of Probability

The mathematical theory of probability is an algebra that can be derived from three simple axioms

$$P[A] \geqslant 0$$

$$P[A] = 1 \text{ means } A \text{ is certain} \tag{11}$$

$$P[A \cup B] = P[A] + P[B] \text{ if } A \text{ and } B \text{ are mutually exclusive}$$

However, none of this describes what probability is. Does it describe the relative frequency with which something happens? Or does it describe the degree of belief that something happens or exists? The relative frequency view implies that there is some underlying frequency with which things happen and that repeated trials or experiments will reveal it. The degree-of-belief view argues that most important questions do not admit of repeated trials and that most practical applications of probabilistic methods employ probability as a measure of confidence in an uncertain outcome. The frequentist argues that probability is inherent in the state of nature and that the analyst's job is to estimate it. The adherent to the degree-of-belief school argues that probability is in the mind of the individual and the analyst's job is to elicit it.

It should be noted that it is possible for the two approaches to apply to the same transaction. The insurance company prices its products as a frequentist. It employs actuaries to calculate the rates of occurrence of various events from observed frequencies. Indeed, it has great difficulty pricing insurance for an event for which it does not have much actuarial data. On the other hand, the purchaser of insurance buys it on the basis of his or her degree of belief. Each of us has one life and a limited number of houses, cars, businesses, and so on. Our decisions whether to buy insurance, how much, and what sort are informed by our own particular circumstances, the exposure we are willing to undertake, and the steps we have taken to minimize risk. Thus, the insurance company is a frequentist, and we are degree-of-beliefers.

When the geotechnical engineer processes laboratory data from many tests to obtain estimates of the properties of geological materials, the engineer is acting like a frequentist. The results are often expressed as means and standard deviations, and there is an implication that the distributions of properties observed in the laboratory apply in the field. However, when carrying out an exploration program, geotechnical engineers are trying to sharpen their degree of belief in a model of the geologic conditions at the site. The author would argue that, in geotechnical engineering, the most important issues involve the engineer's degree of belief, especially when engineering judgment is employed.

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

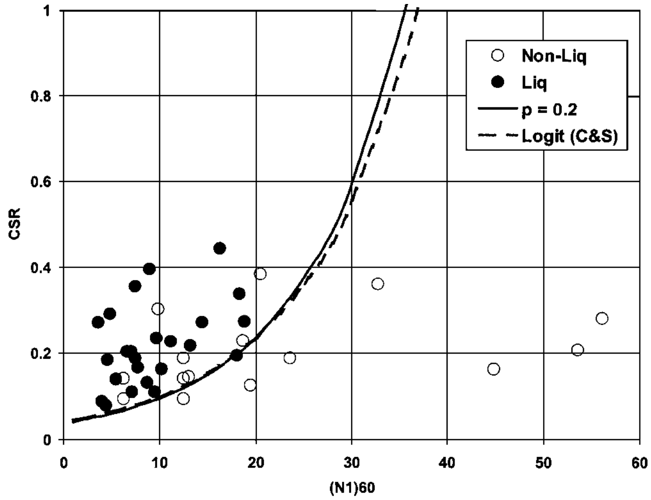

Fig. 6. Cyclic stress ratio versus normalized blow counts, instances of liquefaction and nonliquefaction, and 20% separation lines determined by discriminant analysis (solid line) and logistic regression (dashed line). Data are taken from database used by Christian and Swiger (1975, 1976).

Frequentist versus Bayesian Statistics

One outgrowth of the historical arguments between frequentist and degree-of-belief schools of probability is the distinction between frequentist (or classical) and Bayesian statistics.[2] Frequentist or classical statistics are described in most statistics textbooks and college courses. The essential thrust of classical statistics is to answer the question, "If a particular hypothesis is true, what is the probability that the data I have before me could have been generated?" In other words, it addresses the probability of the data given the state of nature, or, in mathematical notation, $P[\text{data}|\text{state of nature}]$. Geostatistics, logistic regression, and discriminant analysis are examples of classical statistical methods. Fig. 6 shows the 20% curves resulting from discriminant and logistic regression analysis of the modified ground acceleration and standard penetration data for liquefaction and nonliquefaction cases used by Christian and Swiger (1975, 1976).[3] The results are nearly identical, as they should be. Users first coming across this type of plot are inclined to believe that a site whose data fall below and to the right of the line has a 20% probability of liquefaction. This is precisely the wrong interpretation. The actual meaning of the plot is that, if a new site were to liquefy during an earthquake, there is 20% probability that its data would fall below and to the right of the curve. Similarly the probabilities associated with locations of curves in a geostatistical analysis are not the probabilities that the lines are located correctly but the probabilities that the data used in the analysis would be observed if the lines were correct (Baecher and Christian 2003a).

Bayesian analysis addresses the converse question, "If I have before me a set of data, what is now the probability that my view of the subject is true?" That is, it gives the probability of the state of nature given the data, or, in mathematical notation, $P[\text{state of nature}|\text{data}]$. The approach was first proposed by the Reverend Thomas Bayes in 1763 and independently discovered by Marquis Pierre Simon de Laplace in 1782 (Gelman et al. 1995). Bayes received the credit, but the version of the theory now commonly used is due to Laplace.[4] Sivia (1996) and Gelman et al. (1995), among others, have written excellent introductions to Bayesian analysis.

Bayesian analysis starts with a prior probability or prior probability distribution; that is, the analyst must first estimate the state of nature before the new data are introduced. The data then provide an update to the probability of the state of nature. Additional data make possible further updates and better estimates of the state of nature. The basic idea is an extension of Jacques Bernoulli's comment, "Even the stupidest of men, by some instinct of nature, is convinced on his own that with more observations his risk of failure is diminished." (Bernoulli 1713) As De Finetti (1972) wrote: "Data never speak for themselves." They only tell us how to update what we thought before we saw the data to what we logically think afterwards. The procedure is most easily explained by a simple example that nonetheless illustrates some of the insights that arise from Bayesian analysis.

Consider the problem of determining whether a liquefiable zone exists under a proposed facility.[5] The field data are based on results of either the standard penetration test or cone penetration test, and the design earthquake has been specified in advance. The questions to be answered are

- What is the probability that a liquefiable zone exists?
- How is this probability affected by the results of successive borings?
- Are more borings justified?

Let the probability of finding the zone, if it exists, be 0.3 for any one boring; hence the probability of not finding it, if it exists, is 0.7. Also, it is possible to get a false positive when no liquefiable zone exists, so let the probability of the false positive be 0.05. This implies that the probability of not finding it if it does not exist is 0.95. If F indicates that the zone is found, E indicates that the zone exists, and a superposed bar indicates the complement, then the conventional probability notation is

$$P[F|E] = 0.3 \quad P[\overline{F}|E] = 0.7$$
$$P[F|\overline{E}] = 0.05 \quad P[\overline{F}|\overline{E}] = 0.95 \tag{12}$$

The basic form of Bayes' Theorem states that, if there is some prior estimate of the probability that the zone exists, $P_0[E]$, the posterior probability that it exists if the zone is "found" in one boring, $P_1[E|F]$, is

$$P_1[E|F] = \frac{P[F|E]P_0[E]}{P[F|E]P_0[E] + P[F|\overline{E}]P_0[\overline{E}]} \tag{13}$$

The posterior probability that it exists if it is not found in one boring is

$$P_1[E|\overline{F}] = \frac{P[\overline{F}|E]P_0[E]}{P[\overline{F}|E]P_0[E] + P[\overline{F}|\overline{E}]P_0[\overline{E}]} \tag{14}$$

Now, let us suppose that we are of two equal minds about whether or not the zone exists; we really do not know and would not be surprised to find that it does or does not exist. This is equivalent to

$$P[E] = P[\overline{E}] = 0.5 \tag{15}$$

Further, let the result of the first boring be that it "finds" the zone, but this could be a false positive. We want the probability that the zone exists if the boring seems to find it. Inserting the appropriate numbers into Eq. (13) gives

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

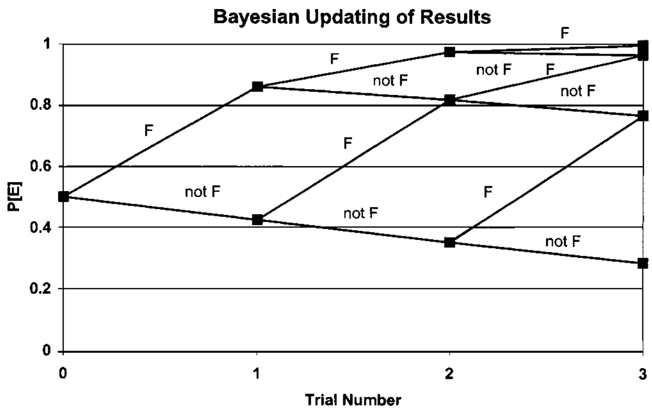

Fig. 7. Posterior probability of existence of liquefiable zone by Bayesian updating on basis of three borings when initial prior probability is 0.5. At each fork upper branch corresponds to "find' and lower to "not find."

$$P_1[E|F] = \frac{(0.3)(0.5)}{(0.3)(0.5) + (0.05)(0.5)} = 0.86 \tag{16}$$

which indicates a sharp increase in the degree of belief that the zone exists. If the boring had not found the zone, Eq. (14) would give

$$P_1[E|\overline{F}] = \frac{(0.7)(0.5)}{(0.7)(0.5) + (0.95)(0.5)} = 0.42 \tag{17}$$

indicating that the degree of belief in the existence of the zone has decreased, but not by much.

As results from additional borings are obtained, the probability of existence of the zone can be updated by treating the posterior result of the previous updating as the prior result for the next. Fig. 7 shows all possible results for three borings when the initial prior probability of existence is 0.5. Fig. 8 shows the corresponding results when the initial probability is 0.25. Four observations that conform to our intuitive experience are apparent:

1. The order of the results makes no difference;
2. Two or three positive results lead to near certainty that the zone exists for this set of parameters;
3. Two or three negative results reduce the belief that the zone exists, but not by much; and
4. As more data accumulate, the probabilities move from the prior assignment to values that reflect the data more strongly.

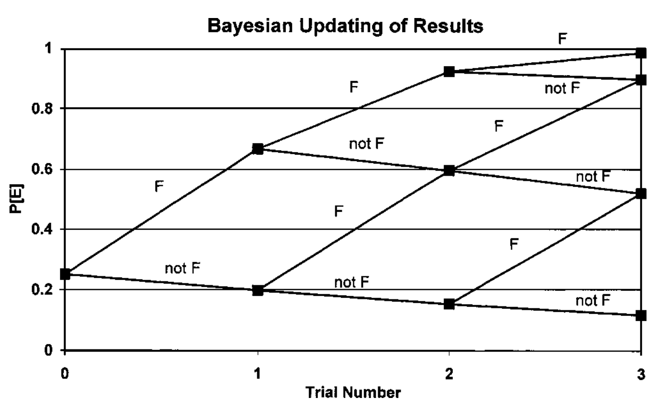

Fig. 8. Results for same analysis as that represented by Fig. 7 except that initial prior probability is 0.25

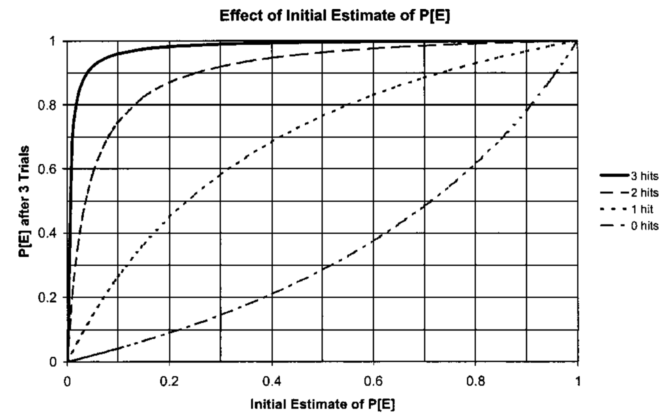

Fig. 9. Posterior probabilities of existence of liquefiable zone versus initial prior probabilities after three borings for all possible initial priors and all outcomes of boring program

The analysis can be repeated for a range of initial prior probabilities, with the results plotted in Fig. 9. The horizontal axis represents the initial probability; the vertical axis is the ultimate posterior probability after three borings. The four lines correspond to the four possible outcomes: three, two, one, or no hits. Again, if there are two or three hits, the data overwhelm the prior probabilities. If there are one or no hits, there is an effect on the posterior probabilities, but it is not nearly so strong. In particular, failure to find the zone in three borings does not lend much support to the belief that the zone does not exist. This conforms to the not-uncommon experience of encountering undesirable conditions during construction despite the exploration programs carried out during design.

Estimating Geotechnical Properties

A central problem facing the geotechnical engineer is to establish the properties of soils and rocks that will be used in analysis, whether that analysis is probabilistic or deterministic. Fig. 10 is a plot of the soil profile for one section of the James Bay dikes (Christian et al. 1994). An engineer wishing to estimate the vane shear strength of the Marine Clay (the middle layer) would be justified in choosing a value that was constant with depth and fell approximately at the mean of the measured data. (Of course, the value should be corrected for the effect of the plasticity index, but that is another issue.) The engineer would make such a choice regardless of whether the strength was to be used deterministically or probabilistically. There is some scatter about the mean in the data. The situation for the Lacustrine Clay (the lowest layer) is not so straightforward. The vane shear strength varies with depth, so there would be substantial scatter about a constant mean value. The engineer might choose a description of the strength that varied linearly with depth, or maybe a more complicated trend line such as a sine wave would be appropriate. The data would fall closer to the trend line, so the scatter about the trend would be reduced. Unfortunately, the uncertainty in the location of the trend is correspondingly increased. The scatter about the trend line is called data scatter, and the uncertainty in the location of the trend is called systematic error. The choice of the shape and location of the trend line is not an artifact of nature; it is a modeling decision made by the engineer. Thus, the separation between data scatter and systematic error is also a modeling decision. This is true even if the results are used entirely deterministically. Put another way,

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

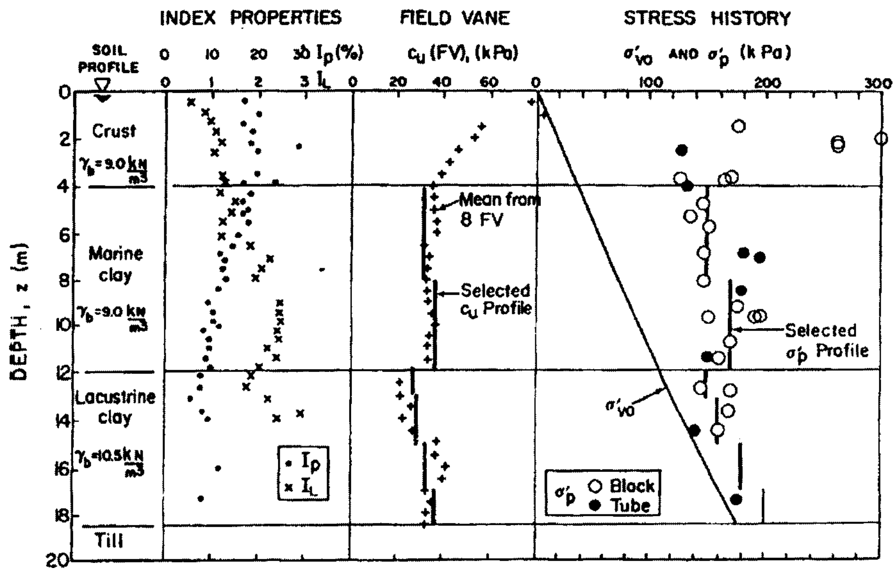

Fig. 10. Soil profile for James Bay dikes, after Christian et al. (1994)

the model for soil properties is a choice made by the modeler and is not a simple reflection of the realities of nature.

Fig. 11, also from Christian et al. (1994), depicts the dichotomy between data scatter and systematic error. It shows that data scatter can be further divided into actual spatial or temporal variation and random measurement error. It is desirable to remove the random measurement error from further analysis. The most common ways to do this are the method of moments and the method of maximum likelihood estimators, which have been described in detail by De Groot and Baecher (1993). The systematic error can also be divided into systematic error in the trend and bias in the measurement procedures. The classic example of the latter is the correction to the vane shear data to account for the plasticity index of the clay (Bjerrum 1972; Terzaghi et al. 1996).

It is important to bear in mind that data scatter and systematic error have different effects on a reliability analysis. In many problems, such as conventional slope stability analyses in which the contributions of shear strength are summed along a failure surface, the scatter in the value of the shear strength averages out, or nearly does so. The contribution of the scatter in the shear strength to the uncertainty of the result is thus greatly reduced as the geometry of the problem gets larger (Christian et al. 1994; El-Ramly et al. 2002, 2003a; Duncan et al. 2003). On the other hand, the systematic error propagates throughout the analysis.

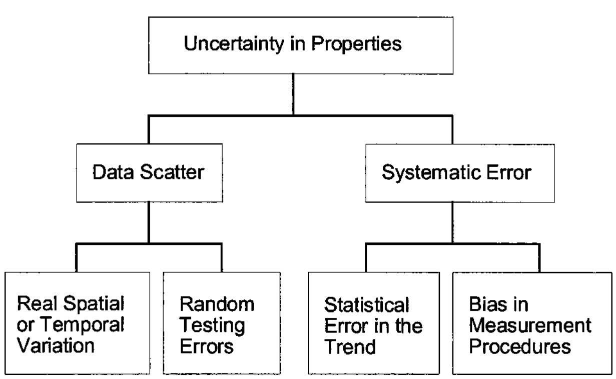

Fig. 11. Conceptual separation of uncertainty into its components for geotechnical applications (Christian et al. 1994)

There are other situations, such as those governed by a single plane of weakness or potential seepage path, in which the larger the volume the more likely the critical feature is to be found. In such cases the data scatter does not average out and is more important than the systematic error.

Another problem arises from the use of small numbers of test results. Much of statistical theory is based on the Law of Large Numbers, which can be summarized in a mathematically nonrigorous way by the statement that, if there is a large enough number of data points, statistical properties can be estimated with an arbitrary degree of accuracy. In the real world, and certainly in geotechnical engineering, there are often far from enough data to satisfy the conditions of the Law of Large Numbers. Tversky and Kahneman (1971) observed that, despite the fact that people often do not have enough data to make valid inferences, they behave as though they did. They called this the "Law of Small Numbers." Consider a data set consisting of six values of shear wave velocity: 229, 224, 229, 217, 200, and 241 m/s. For these data the sample mean is 223 m/s, the standard deviation is 13.9 m/s, and the standard error of the mean is 5.7 m/s. In fact, these are not measured values, but the first six values created by a random number generator from an underlying normal distribution with mean of 240 m/s and standard deviation of 24 m/s. Fig. 12 compares the underlying distribution with a normal distribution inferred from the observed values. It is clear that the inferences drawn from the small sample of six values are not valid. Unfortunately, the same problem of inadequate numbers of data arises in many geotechnical problems, except that the underlying distribution is not known. Basing estimates of geotechnical properties on small numbers of data points, which is the case in many geotechnical projects, can lead to significant and unknown biases in those estimates. This is true regardless of whether the estimates are used in probabilistic or deterministic analyses.

Statistical sampling theory provides some guidance when one is dealing with small numbers of data points. A well-known result is that the standard error of the mean or the standard deviation of the estimate of the mean equals the standard deviation of the sample divided by the square root of the number of data points. However, this applies only in a statistical sense. In the present example, the standard error is 5.7. It is clear, however, that the actual mean of the underlying distribution does not fall within the

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

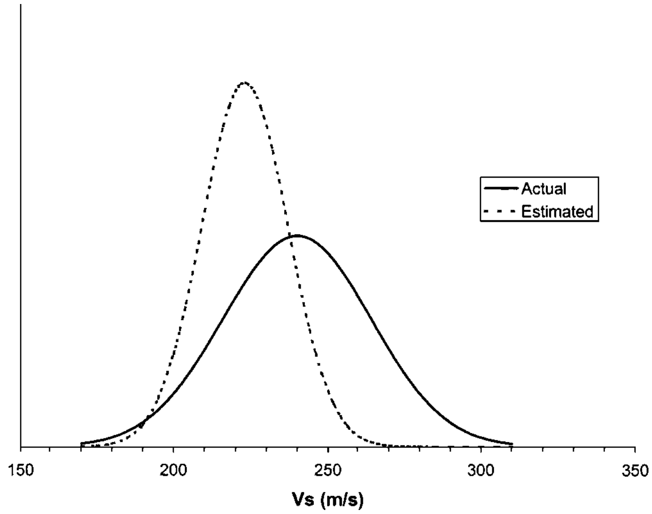

Fig. 12. Actual underlying probability distribution function for shear wave velocity example and normal distribution estimated from six data points

range 223±6 m/s. The actual results in a specific case may not conform to statistical expectations. Anomalies can and do occur.

Expert Elicitation and Engineering Judgment

In view of the limited number of field and experimental data usually available, the geotechnical engineer often has to rely on the opinions of experts and engineering judgment to establish the values and ranges of engineering properties. Obtaining relevant information from experts, or, to use the technical term of art, "elicitation" of expert opinion, has been the subject of extensive study in the management and psychological communities. Morgan and Henrion (1990) and Vick (2002) provide accessible summaries of the issues. These can become quite complicated, so that elicitation of expert opinion is seldom the straightforward process imagined by those who have never worked on it. As evidence of this statement, the Senior Seismic Hazard Analysis Committee report (Budnitz et al. 1997) dealing largely with eliciting expert input for seismic hazard analysis runs to 256 pages plus seven appendices totaling over 850 additional pages. The utility of the methodology is called into question by the need for such voluminous explication.

The first problem is identifying an expert. Who is an expert, and how well qualified is the expert? Obviously, the expert's own opinion of his or her own worth may be too high or too low, so procedures have to be developed to establish the range of the expert's expertise. Furthermore, an expert trained in one discipline may not appreciate the statistical implications of an opinion. Some feedback and iteration is needed to address this problem. In one of the early probabilistic seismic hazard evaluations for a nuclear facility, one of the evaluators for the U.S. Nuclear Regulatory Commission asked one the experts on seismicity, "Do you realize that your model implies that there ought to be a magnitude 5 earthquake or higher at the plant boundary every 10 years?" The expert replied that he was not aware of the implication and did not believe that such a series of events would occur. The two of them then worked out a probabilistic description of the seismicity that was more consistent with the expert's real opinions about the seismicity. In a basic sense, the expert had not understood the question he was being asked.[6]

The literature on eliciting expert opinion generally arrives at two conclusions. First, real experts tend to be good at estimating mean or median values or trends. That is, they get the expected values right. Furthermore, the average of the opinions of several experts tends to be even better. Second, experts are usually too confident in their estimates and tend to underestimate the uncertainty in their estimates.

The last points are illustrated nicely by results published by Hynes and Vanmarcke (1976). An embankment had been built north of Boston, Mass., for a highway project that was later abandoned. In 1974 a team from the Massachusetts Institute of Technology placed additional fill on the embankment to bring it to failure in conjunction with an international workshop at which seven acknowledged experts were invited to make predictions of the behavior of the embankment. Each was asked to predict how much additional fill it would take to cause the embankment to fail and to provide a range within which the expert's confidence of the failure was 50%, also known as the interquartile range The results, modified from Hynes and Vanmarcke's paper, are in shown Fig. 13. The large square points are the experts' best estimates; the vertical lines are the interquartile ranges. The dashed line represents the actual amount of fill that caused failure, 18.7 ft. The average of the seven experts' best estimates was 15.6 ft. This is a good estimate of the actual event, especially since the critical parameter leading to failure is the total height of the embankment, not the last increment. However, the figure also shows that in no case did the actual amount of fill to cause failure fall within an expert's 50% confidence limits. Pure chance would predict that, if the 50% confidence estimates really represent the uncertainties in the experts' judgments, half the vertical lines (i.e., 3 or 4) would intersect the observed value of 18.7 ft. Thus, the experts performed well on the average, but each expert was too confident of his own estimate.

Fig. 14 shows the results when the audience was asked to estimate the required additional fill. Twenty-six people submitted estimates. Once again, the best estimates were distributed approximately evenly about the actual result, but in this case sixteen of the 50% confidence estimates intersected the observed value. Thus, the audience, which had much less time to do its work, managed to include the correct value within the interquartile range 62% of the time. In this case, the experts performed less well than their audience. It is not clear why this is so. The interquartile rages in Figs. 13 and 14 are approximately equal, so the experts and the audience were equally confident of their estimates. The reason for the better performance of the audience cannot be that it was more humble and less confident. Since there were no detailed studies of the psychology of the experts or audience, the reason for the discrepancy must remain a mystery.

Another interesting result appears in Fig. 15. The seven experts were asked to provide, in addition to the interquartile range, the minimum and maximum values of the additional height of fill. In only three cases did the actual additional height of fill fall within an expert's minimum to maximum range. In many cases the minimum to maximum range is virtually identical to or less than the interquartile range. Hynes and Vanmarcke concluded, "It is clear that there are wide differences among engineers in the way they interpret the terms "minimum" and "maximum." These widely used terms are essentially meaningless unless related to relative likelihood or probability."

Kondziolka and Kandaris (1996) described another study of expert elicitation in geotechnical engineering. Nine engineers of various degrees of expertise were asked to design six transmission tower footings against uplift, and the footings were then built and

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

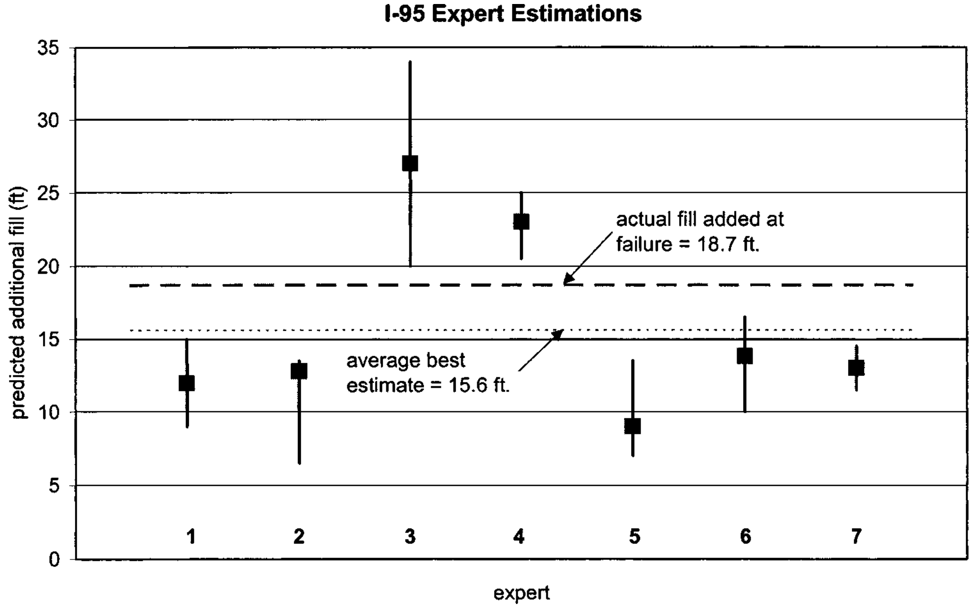

Fig. 13. Seven experts' estimates of additional height of fill to cause failure of I-95 embankment. Square points are experts' best estimates, and vertical bars are their 50% confidence bounds (Hynes and Vanmarke 1976).

tested to failure. The design capacity is designated P, and the actual failure load Q. Fig. 16 presents the results in terms of P/Q for each designer. The square points are the averages over all six footings, and the lines indicate the range of results. A value of P/Q of unity indicates exact prediction of the actual result; values less than unity are conservative in the sense that the predicted capacity is less than that observed. Kondziolka and Kandaris numbered the designers in decreasing order of average goodness of their predictions. Except for the first three designers, they tended to be quite conservative. The range of values of P/Q for each designer is large. For example, the range for the best designer, number 1, is from 0.67 (33% conservative) to 1.2 (20% unconservative). Table 2 presents the experience and education of the participants. It is not clear how this information correlates with a particular designer's performance. The two best performers had 14 and 30 years of experience, respectively, but the worst had 22. Designers with advanced degrees fell into Positions 4–8. As a general rule, the more boxes checked in Table 2, the better the designer performed. However, the third best designer had limited experience and no advanced degrees. One of the lessons to be learned from this example is that it is difficult to predict an expert's performance on the basis of credentials and experience.

These and similar results are relevant to the question of how much reliance should be placed on engineering judgment. There are those who argue that, in the last analysis, judgment is the basis for all geotechnical engineering and that, from the start of their careers, engineers should be encouraged to use it. Others argue that judgment must be based on something other than intuition.

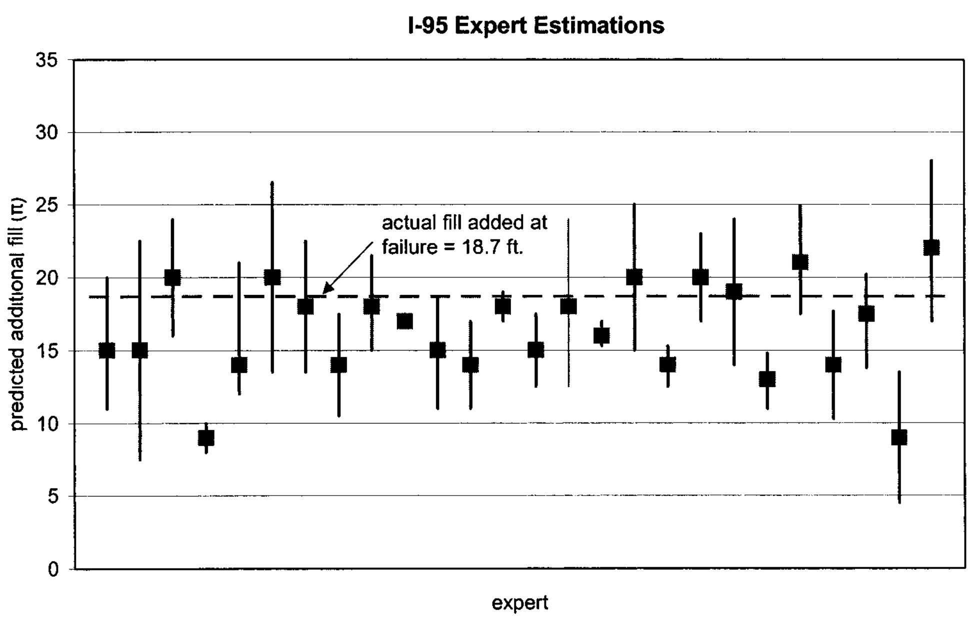

Fig. 14. Audience's estimates of additional height of fill to cause failure of I-95 embankment. Square points are audience members' best estimates, and vertical bars are their 50% confidence bounds (Hynes and Vanmarke 1976).

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

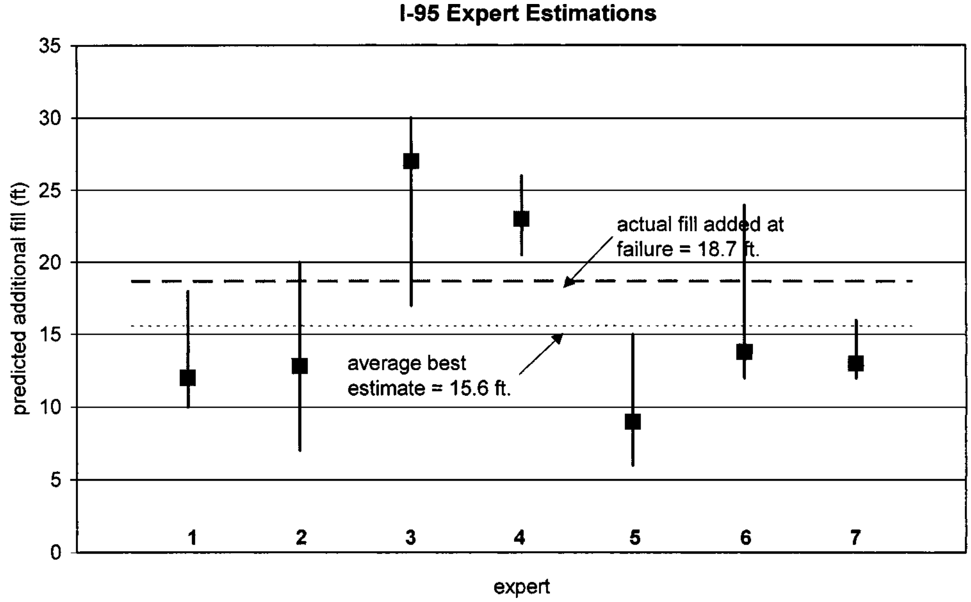

Fig. 15. Seven experts' estimates of additional height of fill to cause failure of I-95 embankment. Square points are experts' best estimates, and vertical bars are their minimum and maximum estimates (Hynes and Vanmarke 1976).

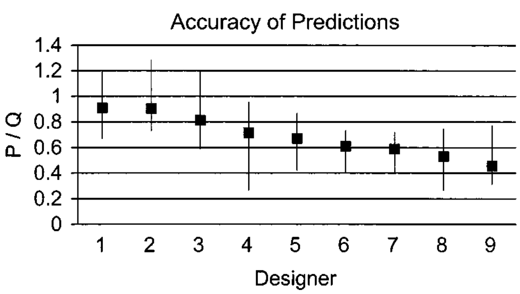

Fig. 16. Ratio of predicted uplift capacity (P) to measured uplift capacity (Q) for six transmission tower foundations evaluated by nine designers. Each square point represents average of six values of P/Q for each designer. Vertical bar represents range of each designer's results. Plotted from results presented by Kondziolka and Kandaris (1996).

For example, (Hartford 2000) proposes that engineering judgment must be based on a chain of reasoning explicitly laid out for inspection. Studies like the two just described and the literature of expert elicitation indicate that good judgment requires not only knowledge and experience but also evaluated experience. It also requires that the engineer be able to demonstrate how the clear chain of reasoning led to the conclusions. That is, it requires that the expert not only have devoted time and effort to learn the topic at hand but also have studied the results of earlier predictions and evaluated what worked and what did not.[7]

Selecting Parametric Values

The above examples suggest certain conclusions about selecting parametric values, whether they will be used in probabilistic or deterministic analyses. Among these are:

1. Dividing uncertainty between spatial and systematic components is fundamentally a modeling choice and not a fact of nature.
2. Spatial and systematic uncertainties contribute differently to

Table 2. Experience and Education of Participants in Transmission Tower Foundation Uplift Project, after Kondziolka and Kandaris (1996)

	Designer								
Experience or education	1	2	3	4	5	6	7	8	9
Previous full scale uplift test experience	x	—	—	x	—	—	—	—	—
Previous uplift foundation design experience	x	x	—	x	x	x	x	—	—
Transmission line tower project experience	x	x	x	—	—	—	—	—	—
Regional geotechnical design experience	x	x	x	x	x	x	—	—	—
Previous involvement with geotech investigations	x	x	—	x	x	x	x	—	x
Drilled pier foundation construction experience	x	x	x	x	x	x	x	—	x
Professional engineer	x	x	x	x	x	x	—	—	—
Bachelor's degree	x	x	x	x	x	x	x	x	—
Master's degree	—	—	—	x	x	x	x	x	—
Doctorate	—	—	—	x	—	—	x	—	—
Years of experience	14	30	8	5	10	10	5	0	22

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

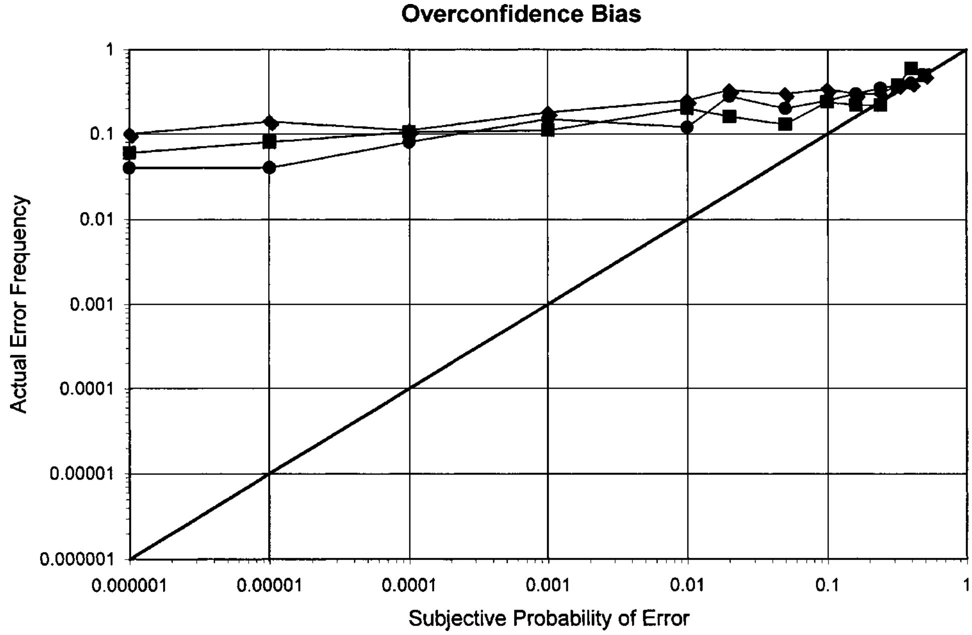

Fig. 17. Actual error frequency versus subjective estimate of probability of error for three groups of subjects asked questions of varying difficulty on variety of subjects. Original results developed by Fischhoff et al. (1997) and also presented by Vick (2002).

uncertainty analyses; in particular, spatial uncertainties tend to average out.

3. Values computed from small samples can be misleading.
4. Experts tend to be more confident than they should be; that is, they underestimate the variances.
5. Engineering judgment is invaluable if it is based on evaluated experience and a demonstrable chain of reasoning; it should never be used as a euphemism for speculation or guessing.

Output—Interpreting Results

After a reliability analysis is complete, the results, like any analysis, must be interpreted. What do the probabilities mean? How are they to be used? What does the analysis reveal about the importance of the various parameters and their uncertainties?

Absolute Probability of Failure

Most people do not understand what a probability means, especially if it is a small probability. Fischoff et al. (1997). developed the original form of Fig. 17; it has been reproduced by other authors (Vick 2002). Three groups of subjects were asked questions of varying difficulty and asked to provide estimates of the probability of error in their answers. Fig. 17 compares these estimated, subjective probabilities of error with the actual frequency of error. When the actual error rates were greater than 0.2, the subjective estimates agreed well with the actual rates. However, when the actual error rates fell below 0.2, the subjective rates dropped precipitously. At the extreme left side of the figure, when the actual rates were between 0.04 and 0.1, the subjects thought their error rates were 10^{-6}. In words, the subjects were overconfident by 5 orders of magnitude! In a similar vein, people are notoriously more frightened of accidents on commercial airliners than on the highways, despite an abundance of widely publicized data that show that air travel is many times safer than driving on the highways.

Two mechanisms for presenting the results of probabilistic analyses in a form that can be grasped intuitively and used in decision-making are the $f-N$ and $F-N$ diagrams, the latter being the cumulative form of the latter. Fig. 18 is a typical example of the $f-N$ diagram (Baecher 1982). The plot has on the horizontal axis either cost in dollars or lives lost.[8] The vertical axis is the observed annual frequency of the losses for various activities. Both axes are logarithmic. The results plot along a broad swath running from the upper left (small costs and high frequency of failure) to lower right (high costs and low frequency of failure). This is an experimental result; it reveals the rates of failure and costs that society—or at least some operating part of society—has

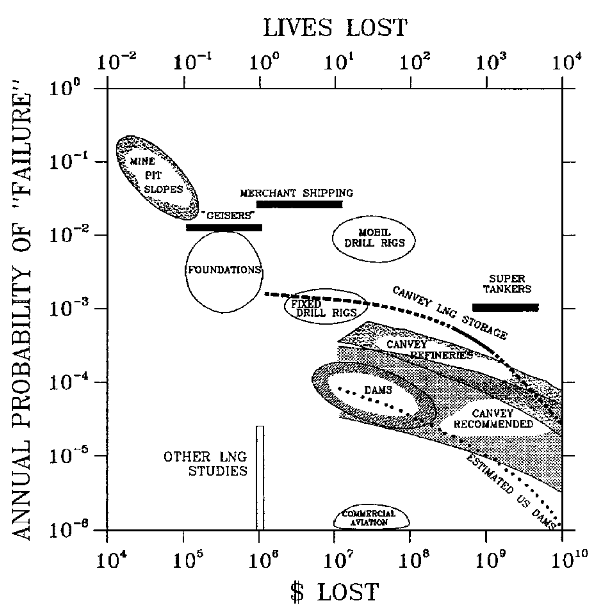

Fig. 18. One version of $f-N$ plot annual risk cost or number of lives. In this plot both cost and lives are shown; it is customary to use one or the other rather than both on same plot (Baecher 1982).

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

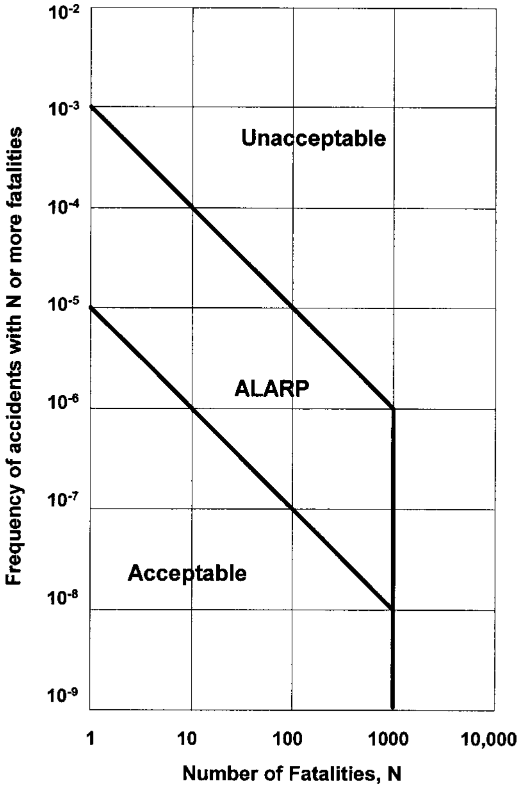

Fig. 19. *F–N* diagram adopted by Hong Kong Planning Department for planning purposes (Hong Kong Government Planning Department 1994). "ALARP" stands for "as low as reasonably practicable."

implicitly found acceptable. Of course, asked in a referendum what risks are acceptable, society might choose a different acceptable risk, but Fig. 18 represents what we live with now. Some activities, such as commercial aviation, fall well below the trend, and this conforms to the general perception that commercial air travel is relatively safe. Others, such as mobile drilling rigs, fall above the trend; it is not surprising that mobile drilling rigs are dangerous places to work. Any other risk can be plotted on the same figure to see how it compares to other activities. For example, there have been at this writing 22 years of National Aeuronautics and Space Administration space shuttle missions, two of which failed costing seven lives each and untold dollars. If one plots a point for seven lives and an annual failure rate of 0.09, the point falls well above the trend line. This confirms that astronauts on the space shuttle are exposed to high risks. It also indicates that the technology is not yet so reliable that people from the general public, such as schoolteachers, should be invited to participate.

Several organizations that deal with public policy and safety have adopted *F–N* plots as aids in decision-making. Fig. 19 is the version adopted by the Hong Kong Planning Department (Hong Kong Government Planning Department 1994). Fig. 20 was developed in the Netherlands (Versteeg 1987). The Australia New Zealand Committee on Large Dams (ANCOLD 1994) proposed Fig. 21. Von Thun (1996) presented a somewhat more complicated figure proposed for the Bureau of Reclamation. Three points of clarification must be emphasized. First, the vertical axis in Fig. 18 is the number of events occurring in a year, but the vertical axis in Figs. 19–21 is the annual rate of occurrence of *N* or more events. Second, locations of the lines separating the regions are not the same in all the figures; the locations reflect negotiations among the designers of the figures. Third, current practice is not

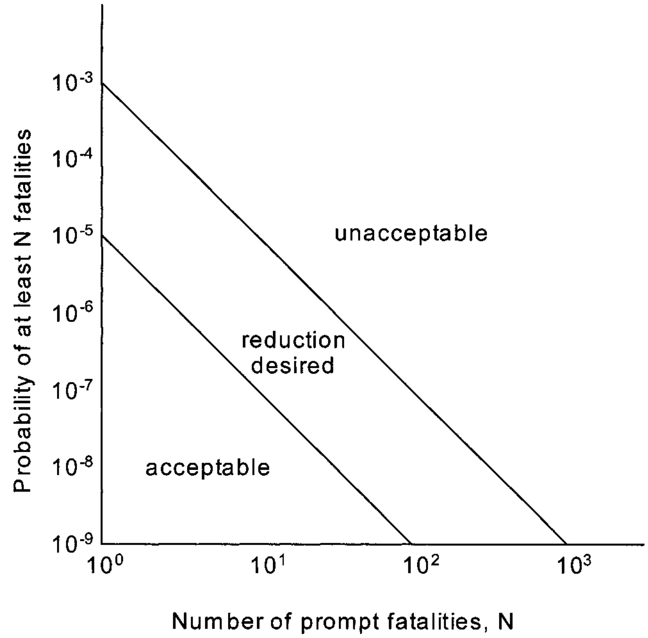

Fig. 20. *F–N* diagram proposed for Netherlands for planning and design (Versteeg 1987). "Prompt fatalities" is term used in original reference and refers to failures that occur in short term rather than because of lingering effects.

to be bound by the bright lines separating the regions but to use them as guidelines. Regardless of how the figures are developed, they are convenient tools for comparing the results of reliability analyses with acceptable levels of risk.

Whenever one is working with a computed probability of failure, it must be borne in mind that the probability of failure computed with best estimates of the statistical parameters is likely to

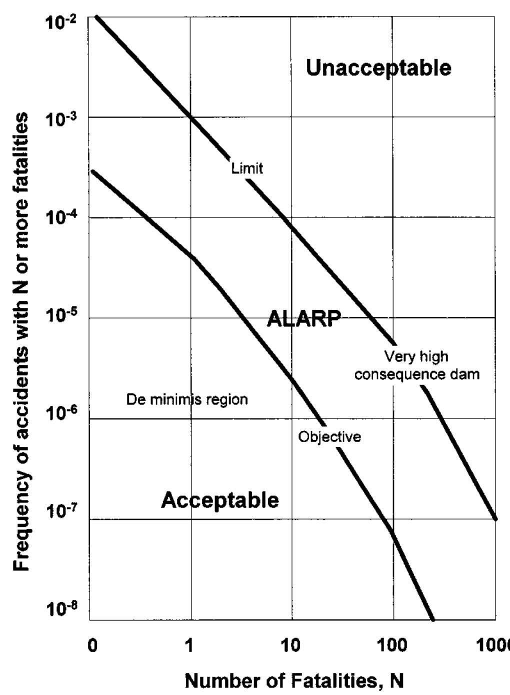

Fig. 21. *F–N* diagram proposed by Australia New Zealand Committee on Large Dams (ANCOLD 1994). "ALARP" stands for "as low as reasonably practicable."

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

Table 3. Comparative Probabilities of Failure for James Bay Dikes, after Christian et al. (1994)

Case	$E[F]$	p_f
H=6 m, single stage	1.58	2.5×10^{-2}
H=12 m, single stage	1.53	4.7×10^{-3}
H=23 m, multiple stage	1.50	7.1×10^{-4}

be a lower bound. The computed value necessarily does not include the effects of factors that were not included in the analysis. Since things that were ignored during design and construction cause many failures, engineers should be wary of placing too much confidence in the absolute values of the computed probabilities of failure. As Leps' (1987) statement quoted earlier emphasizes, if the designer knows of something that could cause a failure, he or she should fix it. Of course, a probability of failure computed with conservatism piled on conservatism is not necessarily a lower bound, but it is also not clear what the computed probabilities mean.

Comparative Probability of Failure

It is often more useful to compare probabilities of failure for different alternative courses of action than to rely on the absolute probability of failure. In the previous sentence the word "compare" is used deliberately. As Gigerenzer (2002) stresses, relative probabilities can be misleading. He gives the example of a screening procedure that reduces the risk of dying of breast cancer from 4 per 1,000 patients to 3 per 1,000. The absolute effect is to reduce the risk by 1 per 1,000, but the relative effect is 25% reduction in risk. He writes, "Relative risks do not carry information about the absolute benefits of treatment."

Christian et al. (1994) give an example of the use of comparative probabilities. Three heights of dikes were proposed for the James Bay project: 6, 12, and 23 m. The first two are single stage dikes; the last, a composite dike built in stages. Table 3 gives the estimated factors of safety and probabilities of failure for the three designs. Although the factors of safety are similar, the probabilities of failure are quite different. The 23 m dike has a lower probability of failure, which is not reflected in the factor of safety. Another way to look at these results is to consider the desirable target probabilities of failure. On the basis of the historical behavior of dikes of this type, an annual probability of failure of 0.001 was established as a reasonable target for typical dikes such as the 12 m dikes. The consequences of failure for the lower 6 m dikes would be smaller, so a larger probability of failure of 0.01 was chosen. The greater size and importance of the 23 m multistage dike led to a reduced target of 0.0001. Working through the analytical results led to the target estimated factors of safety corresponding to these target probabilities and listed in Table 4. These could then be used for design calculations. It is interesting that the values of the target factors of safety are in inverse order to what would be expected intuitively. The reason for this reversal of expectations is primarily that the spatial contributions to the uncertainty average out over the larger failure surfaces that apply for the higher dikes. This reduces the uncertainty in the factor of safety. Such might not be the case for another design problem.

Table 4. Target Design Factors of Safety for James Bay Dikes, after Christian et al. (1994)

Case	Target p_f	Target FS
H=6 m, single stage	0.01	1.63
H=12 m, single stage	0.001	1.53
H=23 m, multiple stage	0.0001	1.43

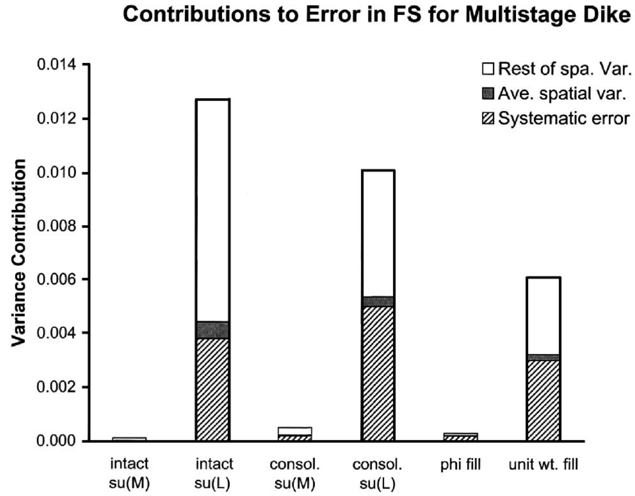

Fig. 22. Contributions of each of six different uncertain factors to variance of factor of safety of the 23 m high multistage dike for James Bay project (Christian et al. 1994)

Contributions of Components

Another useful result from reliability analyses is the contribution of the various factors to the probability of failure. Fig. 22, also from the James Bay study (Christian et al. 1994), shows the contribution of each of six uncertain factors contributing to the variance in the factor of safety for the 23 m high dike. The factors are the shear strength of the intact marine clay, the intact lacustrine clay, the consolidated marine clay, and the consolidated lacustrine clay; the friction angle of the fill; and the unit weight of the fill. Each bar has three parts: the contribution of the systematic error, the contribution of the spatial error after it has been averaged over the failure surface, and the additional contribution of the spatial error that is removed by averaging. The plot shows clearly that the strength of the lacustrine clay and the unit weight of the fill contribute much more to the variance of the factor of safety than the other three factors. One implication is that to reduce the uncertainty in the factor of safety, and hence the probability of failure, the engineer would be well advised to concentrate on improving the knowledge of the strength of the lacustrine clay and the unit weight of the fill.

Factor of Safety and Reliability

The preceding paragraphs could give the erroneous impression that there is an inherent conflict between approaches using a factor of safety and those based on reliability theory. This is not the case. The factor of safety is a value computed by well-known methods that provides a measure of the expected performance of a slope. For other problems, other computed values are appropriate, such as estimated settlement, uplift pressures, bending moments, and so forth. Reliability theory does not invalidate such calculations. It extends them by giving them a context and by giving additional information to help the engineer interpret the results.

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

Conclusions

Many conclusions can be drawn from the applications of probability to geotechnical problems. This paper has concentrated on the imperfections in our knowledge and how they affect our ability to make decisions. It is clear that our knowledge of the geological and environmental factors affecting geotechnical engineering is imperfect and that it will remain so. Although modern developments in remote sensing and information technology promise to ameliorate this situation, we are not likely ever to have as much or as reliable information as we would like to have. However, we have to proceed with our projects. The first step is to recognize the extent of our ignorance and to understand whence it arises. We can reduce uncertainty by obtaining more information, especially when the search or more information is guided by a rational understanding of the nature of uncertainty and its impact on our decisions. Many practical tools—the observational method, adaptive management—have been developed to deal with uncertainty in the engineering project.

Probabilistic methods provide a powerful tool for dealing with these issues, a tool that is finding increasing application in practice. Many of its insights apply to deterministic methods as well, whether or not they are formally recognized. Most of the currently available tools for applying probabilistic methods to engineering can be placed in one of two categories—logic trees and direct reliability methods. The details of these techniques are widely available, and the methods themselves have found application across all engineering fields, including some in geotechnical engineering.

While the tools themselves are increasingly well known, the underlying nature of uncertainty, the meaning of probability, and the differences between frequentist and Bayesian statistics are not. There are also problems in estimating geotechnical parameters. We usually deal with an inadequate number of data points, and it is important to separate the spatial from the systematic contributions. Exerts are often used to elucidate such questions, but a large body of experience from other fields as well as geotechnical engineering indicates that it is difficult to elicit information from experts and that experts are often too confident of their estimates.

Using the output of probabilistic analyses is hindered by the well-established fact that people, including engineers, have a lot of trouble understanding small probabilities. In recent years, the $f-N$ and $F-N$ diagrams have proven to be useful tools for describing the meaning of probabilities and risks in the context of other risks with which society is familiar. Computed absolute probabilities may not include all contributions; an effective approach is to compare probabilities of different options or alternatives. Probabilistic methodologies also provide insight into the relative contributions of different parameters to the uncertainty of the result and thus give guidance for where further investigations will be most fruitful.

There are three important conclusions with which the paper closes:

- Probability in geotechnical engineering is not a property of the world but a state of mind;
- Thus, geotechnical and geological uncertainty is belief-based and necessarily Bayesian; and
- The current challenge to the geotechnical engineering profession is how to use probabilistic methods in practice.

To return to the question posed in the subtitle to this paper, there is a lot that we do not know about what we are doing. The best engineers have always brought a healthy skepticism to projects that deal with geology and the environment. Modern probabilistic methods now provide an additional tool for describing and dealing with that uncertainty.

Acknowledgments

Many engineers and researchers, including the reviewers, have contributed to developing probabilistic methods for geotechnical engineering; some are mentioned in the first section of this paper. The writer acknowledges his debt to them. In the case of the present paper and the lecture on which it is based, the author acknowledges the helpful criticisms of Gregory B. Baecher, Desmond N. D. Hartford, William F. Marcuson, III, James K. Mitchell, and Alfredo Urzua.

Endnotes

[1]It might be argued that, if we knew enough about the linear and angular velocities of the dice, their inertia, the rebounding characteristics of the dice and the table, and so on, we would be able to predict the outcome of any throw. However, this is so impractical that the expressions "throw of the dice" and "crap shoot" have entered the language as synonyms for totally random events.

[2]It should be noted that it is possible to apply Bayesian methods when probability is defined by relative frequency or classical methods to degree-of-belief probability. However, to avoid excessive and extraneous complication, the presentation follows the line that frequentist definitions of probability tend toward classical statistics while degree-of-belief approaches are more congenial with Bayesian approaches. This is also the historical distinction.

[3]This database has been superceded by many more observations since the analyses were first carried out, but the point about the meaning of the curves remains valid.

[4]This is one instance of *Stigler's Law of Eponymy*, which states in its simplest form, "No scientific discovery is named after its original discoverer" (Stigler 1999).

[5]A similar analysis applies to many other exploration problems; the liquefiable zone problem is chosen for convenience.

[6]This story was told to the author by the National Research Council evaluator.

[7]Many engineers have learned to their sorrow that relying simply on "engineering judgment" in an adversarial proceeding can lead to embarrassing cross examination.

[8]The plot has both cost and lives lost axes because some of the original references wrote about costs and others about lives lost. In a particular application one should use one or the other, but not both.

References

Aitchison, J., and Brown, J. A. C. (1969). *The lognormal distribution: With special reference to its uses in economics*, Cambridge University Press, Cambridge, U.K.

Australian, New Zealand Committee on Large Dams (ANCOLD). (1994). "Guidelines on Risk Assessment 1994." Brisbane, Australia.

Baecher, G. B. (1982). "Statistical methods in site characterization." *Updating subsurface samplings of soils and rocks and their in-situ testing*, Engineering Foundation, Santa Barbara, Calif., 463–492.

Baecher, G. B., and Christian, J. T. (2003a). "Discussion of 'Evaluating site investigation quality using GIS and geostatistics' by R. L. Parsons and J. D. Frost." *J. Geotech. Geoenviron. Eng.,* 129(9), 867–868.

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

Baecher, G. B., and Christian, J. T. (2003b). *Reliability and statistics in geotechnical engineering*, Wiley, Chichester, U.K.
Bea, R. G. (1999). "Characteristics of a platform in the Mississippi River Delta." *J. Geotech. Geoenviron. Eng.*, 124(8), 729–738.
Bernoulli, J. (1713). *Ars Conjectandi*, quoted by Luigi Luca Cavalli-Sforza, *Genes, People, and Languages*, North Point Press, New York, 2000.
Bjerrum, L. (1972). "Embankments on soft ground: State-of-the-art report." *Specialty Conf. on Performance of Earth and Earth-Supported Structures*, ASCE, New York, 1–54.
Budnitz, R. H., Apostolakis, G., Boore, D. M., Cluff, L. S., Coppersmith, K. J., Cornell, C. A., and Morris, P. A. (1997). "Recommendations for probabilistic seismic hazard analysis: Guidance on uncertainty and use of experts, NUREG/CR-6372." U.S. Nuclear Regulatory Commission, Washington, D.C.
Budnitz, R. H., Apostolakis, G., Boore, D. M., Culff, L. S., Coppersmith, K. J., Cornell, C. A., and Morris, P. A. (1998). "Use of technical expert panels: Applications to probabilistic seismic hazard analysis." *Risk Anal*, 18(4), 463–469.
Calderon, A., Catalan, A., and Karzulovic, A. (2003). "Modelo de riesgo geotecnico y bases geotecnicas, Plan Minero CBV-2003, Mina Chiquicamata." CODELCO, Chile.
Carnap, R. (1936). "Testability and meaning." *Philos. Sci.*, 3, 420.
Casagrande, A. (1965). "The role of the 'calculated risk' in earthwork and foundation engineering." *J. Soil Mech. Found. Div.*, 91(4), 1–40.
Chow, V. T., Maidment, D. R., and Mays, L. W. (1988). *Applied hydrology*, McGraw-Hill, New York.
Christian, J. T., and Baecher, G. B. (1999). "Point-estimate method as numerical quadrature." *J. Geotech. Geoenviron. Eng.*, 125(9), 779–786.
Christian, J. T., Ladd, C. C., and Baecher, G. B. (1994). "Reliability applied to slope stability analysis." *J. Geotech. Eng.*, 120(12), 2180–2207.
Christian, J. T., and Swiger, W. F. (1975). "Statistics of liquefaction and SPT results." *J. Geotech. Eng. Div., Am. Soc. Civ. Eng.*, 101(11), 1135–1150.
Christian, J. T., and Swiger, W. F. (1976). "Closure to statistics of liquefaction and SPT results." *J. Geotech. Eng. Div., Am. Soc. Civ. Eng.*, 102(12), 1279–1281.
Christian, J. T., and Urzua, A. (1996). *Productivity tools for geotechnical engineers*, Magellan Press, Newton, Mass.
Cornell, C. A. (1968). "Engineering seismic risk assessment." *Bull. Seismol. Soc. Am.*, 58, 1583–1606.
Cornell, C. A. (1969). "Structural safety specifications based on second-moment reliability analysis." Dept. of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Mass.
De Finetti, B. (1972). *Probability, induction and statistics: The art of guessing*, Wiley, New York.
De Groot, D. J., and Baecher, G. B. (1993). "Estimating autocovariances of in situ soil properties." *J. Geotech. Eng.*, 119(1), 147–166.
Duncan, J. M., Navin, M., and Wolff, T. F. (2003). "Discussion of 'Probabilistic slope stability analysis for practice' by El-Ramly, Morgenstern, and Cruden." *Can. Geotech. J.*, 40(4), 848–850.
El-Ramly, H., Morgenstern, N. R., and Cruden, D. M. (2002). "Probabilistic slope stability analysis for practice." *Can. Geotech. J.*, 39(3), 665–683.
El-Ramly, H., Morgenstern, N. R., and Cruden, D. M. (2003a). "Closure to discussion by Duncan, Navin, and Wolff of 'Probabilistic slope stability analysis for practice'." *Can. Geotech. J.*, 40(4), 851–855.
El-Ramly, H., Morgenstern, N. R., and Cruden, D. M. (2003b). "Probabilistic stability analysis of a tailings dyke on presheared clay-shale." *Can. Geotech. J.*, 40(1), 192–208.
Fischhoff, B., Slovic, P., and Lichtenstein, S. (1997). "Knowing with certainty: The appropriateness of extreme confidence." *J. Exp. Psychol.*, 3(4), 552–564.
Fishman, G. S. (1995). *Monte Carlo: Concepts, algorithms, and applications*, Springer, New York.
Gelman, A., Carlin, J. B., Stern, H. B., and Rubin, D. B. (1995). *Bayesian data analysis*, Chapman & Hall/CRC, Boca Raton, Fla.
Gigerenzer, G. (2002). *Calculated risks*, Simon & Schuster, New York.
Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., and Kruger, L. (1989). *The empire of chance: How probability changed science and everyday life*, Cambridge University Press, New York.
Hacking, I. (1975). *The emergence of probability*, Cambridge University Press, Cambridge, U.K.
Hartford, D. N. D. (2000). "Judged values and value judgements in dam risk assessment: A personal perspective." *ANCOLD Bull., 114*, 78–86.
Hasofer, A. M., and Lind, N. (1974). "Exact and Invariant Second-Moment Code Format." *J. Eng. Mech.*, 100(1), 111–121.
Hoek, E. (1998). *Rock engineering: The application of modern techniques to underground design*. Brazilian Rock Mechanics Committee (CBMR), Brazilian Tunneling Committee (CBT), Brazilian Society for Soil Mechanics and Geotechnical Engineering (ABMS), Sao Paulo, Brazil.
Hong Kong Government Planning Department. (1994). "Potentially hazardous installations." *Hong Kong planning standards and guidelines*, Chap. 11, Hong Kong.
Hynes, M., and Vanmarke, E. (1976). "Reliability of embankment performance predictions." *Proc., Engineering Mechanics Division Specialty Conf.*, University of Waterloo Press, Waterloo, Ont., Canada.
Jensen, J. L. (1997). *Statistics for petroleum engineers and geoscientists*, Prentice Hall, Upper Saddle River, N.J.
Kondziolka, R. E., and Kandaris, P. M. (1996). "Capacity predictors for full scale transmission line test foundations." *Uncertainty in the geologic environment: From theory to practice*, ASCE, Reston, Va. 695–709.
Lacasse, S., and Nadim, F. (1996). "Uncertainties in characteristic soil properties." *Uncertainty in the geological environment: From theory to practice*, ASCE, Reston, Va. 49–75.
Leps, T. M. (1987). "Comments on lessons learned." *Eng. Geol. (Amsterdam)*, 24(1–4) 94–94.
McCann M. (1999). "National performance of dams program." ⟨http://npdp.stanford.edu/⟩.
Meyerhof, G. G. (1953). "The bearing capacity of foundations under eccentric and inclined loads." *Proc., 3rd Int. Conf. on Soil Mechanics and Foundation Engineering*, Zurich, Switzerland, 440–445.
Meyerhof, G. G. (1963). "Some recent research on the bearing capacity of foundations." *Can. Geotech. J.*, 1(1), 16–26.
Morgan M. G., and Henrion, M. (1990). *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*, Cambridge University Press, New York.
National Research Council (NRC). (1995). *Flood risk management and the American River basin: An evaluation*, National Academies Press, Washington, D.C.
National Research Council. (2000). *Risk analysis and uncertainty in flood damage reduction studies*, National Academy Press, Washington, D.C.
Parr, N. M., and Cullen, N. (1988). "Risk management and reservoir maintenance." *J. Inst. Water Environ. Manage.*, 2.
Peck, R. B. (1969). "Ninth Rankine lecture: Advantages and limitations of the observational method in applied soil mechanics." *Geotechnique*, 19(2), 171–187.
Riela, J., Urzua, A., Christian, J. T., Karzulovic, A., and Flores, G. (1999). "Sliding rock wedge reliability analysis of chuquicamata mine slopes." *Proc., 11th Pan-American Conf. on Soil Mechanics and Geotechnical Engineering*, Foz do Iguacu, Brazil.
Rosenblueth, E. (1975). "Point estimates for probability moments." *Proc. Natl. Acad. Sci. U.S.A.*, 72(10), 3812–3814.
Rosenblueth, E. (1981). "Two-point estimates in probabilities." *Appl. Math. Model.*, 5(2), 329–335.
Sivia, D. S. (1996). *Data analysis: A Bayesian tutorial*, Oxford University Press, Oxford, U.K.
Stedinger, J. R., Heath, D. C., and Thompson, K. (1996). "Risk assessment for dam safety evaluation: Hydrologic risk." USACE Institute for Water Resources, Washington, D.C.

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003

Stigler, S. M. (1999). *Statistics on the table: The history of statistical concepts and methods*, Harvard University Press, Cambridge, Mass.
T. W. Lambe & Associates. (1982). "Earthquake risk to patio 4 and site 400." Longboat Key, Fla.
T. W. Lambe & Associates. (1989). "Earthquake risk analysis for KSS tank area." Longboat Key, Fla.
Terzaghi, K. (1929). "Effect of minor geological details on the stability of dams." *Technical Publication No. 215*, American Institute of Mining and Metallurgical Engineers, New York, 31–44.
Terzaghi, K., Peck, R. B., and Mesri, G. (1996). *Soil mechanics in engineering practice*, Wiley, New York.
Tversky, A., and Kahneman, D. (1971). "Belief in the 'law of small numbers'." *Psychol. Bull.*, 76, 105–110.
U.S. Nuclear Regulatory Commission. (1975). "Reactor safety study: An assessment of accident risks in U.S. commercial nuclear power plants." Washington, D.C.
Van Zyl, D., Miller, I., Milligan, V., and Tilson, W. J. (1996). "Probabilistic risk assessment for tailings impoundment founded on Paleokarst." *Uncertainty in the geologic environment: From theory to practice*, ASCE, Reston, Va.
Veneziano, D. (1995). "Uncertainty and expert opinion in geologic hazards." Massachusetts Institute of Technology, *The Earth, Engineers, and Education: A Symposium in Honor of Robert V. Whitman*, Cambridge, Mass., 102–124.
Versteeg, M. (1987). "External safety policy in the Netherlands: An approach to risk management." *J. Hazard. Mater.*, 17, 215–221.
Vesic, A. S. (1973). "Analysis of ultimate loads of shallow foundations." *J. Soil Mech. Found. Div.*, 99(1), 45–73.
Vesic, A. S. (1975). "Bearing capacity of shallow foundations." *Foundation engineering handbook*, H. F. Winterkorn and H.-Y. Fang, eds., Van Nostrand–Reinhold, New York, 121–147.
Vick, S. G. (2002). *Degrees of belief: Subjective probability and engineering judgment*, ASCE, Reston, Va.
Vick, S. G., and Stewart, R. A. (1996). "Risk analysis in dam safety practice." *Uncertainty in the geologic environment: From theory to practice*, ASCE, Reston, Va., 586–603.
Von Thun, J. L. (1996). "Risk assessment of Nambe Falls Dam." *Uncertainty in the geologic environment: From theory to practice, GSP 58*, ASCE, Reston, Va., 604–635.
Whitman, R. V. (1984). "Evaluating calculated risk in geotechnical engineering." *J. Geotech. Eng.*, 110(2), 143–188.

J. Geotech. Geoenviron. Eng., 2004, 130(10): 985-1003