Fourth Australia - New Zealand
Conference on Geomechanics,
Perth, 14-18 May 1984

Statistical Site Characterization

A. HALDAR
Assistant Professor, School of Civil Engineering, Georgia Institute of Technology, USA

SUMMARY The estimation of the in situ site parameters in any geomechanical problem influences the predictability of the model. When the parameters are estimated from the data from a few poorly sampled, widely spaced and often poorly logged borings, this adds another dimension to the problem. Soil deposits are inherently nonhomogeneous. The small-scale properties of this complex material need to be modeled accurately. Methods are discussed here to model the three-dimensional characteristics of a soil deposit statistically. A statistical site characterization model could complement the solution techniques for a complex geomechanical problem. Discussions are made here on how this method can be used in a liquefaction problem. The estimation of site-related parameters deserves more attention than it usually receives.

1 INTRODUCTION

Many sophisticated techniques are being developed in geomechanics to model material properties, constitutive relationships, soil behavior, etc. In many cases, improved finite element and boundary element techniques are being used for this purpose. In spite of all this sophistication, the site conditions may not be adequately represented in these models. Thus, the predictability of these models is limited to very special cases. The uncertainty in the estimation of parameters to represent site conditions deserves more attention than it receives.

The purposes of site investigation are well summarized by Sowers (1979). According to him, a complete investigation of underground conditions includes (1) the nature of the deposit (geology, recent history of filling, excavation and flooding, possibility of mineral exploitation); (2) depth, thickness, lateral extent, elevations and composition of each stratum and the geologic discontinuities; (3) groundwater elevations, their differences across the site, and their changes with time and environmental change; and (4) the engineering properties of the soil and rock strata that affect the performance of the structure.

The unique problem a geotechnical engineer faces is that he must work with a very complex material in "as is" condition and define its small-scale properties. This very important but difficult task is routinely performed following a few basic steps: (a) A feasibility study is made considering the basic structure to be built. This may include the general configuration, location, layout of the structure, availability of space for future expansion, even political considerations. (b) The next stage is a review of the available information and surface reconnaissance. (c) This stage generally consists of surface mapping, preliminary borings, initial laboratory testing and preliminary analysis. (d) In this stage, the data collected so far are synthesized. Also, this stage should involve borings to recover specialized samples, geophysical surveys, partial excavation (adits, trenches, calix holes, etc., if needed) and specialized testing.

All these steps are taken to quantify the geometric conditions and material properties of the site that can be used in an analysis and to avoid cases of "changed conditions" that were not expected from the site under consideration. However, there may be some deficiencies in the site characterization problem. Dowding (1978) observed that more test cylinders (6 times for the same volume) are taken for a conventional concrete building than for the geotechnical samples. He commented "the possible range in geotechnical properties is much greater than that of the structural properties. One must ask: why do we test man-made materials more intensively than geological material?" Osterberg (1978) added "In fact, since it is common (perhaps even the rule rather than the exception) that conditions revealed during excavation are not what was anticipated from the exploratory program, most site evaluations could be considered at least partial exploration failures."

The overall state of the art in the areas of exploration and characterization has been effectively stated by Underwood (1974) as "One attitude that has discouraged the writer over the past few years is the apparent hope for some new magical development that will fill in the gaps between a few poorly sampled, widely spaced and often poorly logged borings." He added "Field investigations are often hurriedly and carelessly conducted and the incomplete data is then carefully analyzed by precise (out to 8-digit accuracy) computer techniques which produce impressive but erroneous results, which in turn can lead to inaccurate design assumptions."

There is no doubt that advanced mathematical formulation is necessary to understand the geomechanical problems. However, the influence of the crude site information on the overall predictability of the problem needs some consideration. A statistical method is suggested here to resolve this problem.

2 PROBLEM DESCRIPTION

The general consensus in the profession is that the soil deposit could be extremely nonhomogeneous and very difficult to model for theoretical use. Often inferences are made from inadequate information obtained from limited samples due to the accessibility problem. The potential sources of uncertainties in the site evaluation arise due to the engineer's inability to visualize the subsoil, difficulty in choosing the appropriate scale for exploration, uncertainty as to what information to obtain, the need to interpolate or extrapolate point or line data over a vertical profile and lateral distance, and failure or inability to evaluate all information as it becomes available (Mitchell, 1978). A systematic evaluation of these sources of uncertainties is a necessity.

Any particular site exploration program, whether deterministic or probabilistic, must be specifically developed considering the potential construction problems and geotechnical conditions affecting the site. Conceptually, then, the cases that need to be considered are numerous. However, the question that often needs to be addressed is how to decide whether the site has been sufficiently explored with respect to a particular parameter, in terms of degree of accuracy, number of samples collected and number, arrangement, spacing and depth of boreholes. How are the parameters of the models estimated? Is it sufficient to use a point estimate based on one observation or mean, median or upper and lower bounds using multiple observations; or would it be preferable to estimate the parameter's characteristics over a volume considering spatial correlation, independence, etc.? What would be the appropriate statistical distribution of the parameters?

Some of the methods that can be used to address some of these questions are discussed in the following sections. In this paper, all discussion will be made considering a hypothetical site susceptible to liquefaction. Several models to evaluate liquefaction potential, both deterministic and probabilistic, were reviewed elsewhere by Haldar and Miller (1982). The purpose of this paper is not to discuss these methods, but to study how site-related parameters are used in these models. For a liquefaction problem, the relative density D_r and the cyclic strength parameter R are the two most important parameters (Haldar and Miller, 1982).

2.1 Objectives of Site Characterizations

As mentioned earlier, representative samples are taken from a site to characterize it. The most effective use of site information is the primary objective of a site characterization problem. From a geotechnical engineer's point of view, the following factors need to be considered: (i) the relative importance of the parameter being evaluated in the engineering analysis, (ii) the importance of the loss or error function, i.e., cost associated in case of failure or cost associated with overdesign, (iii) time considerations, (iv) the location of the structure relative to the zone where samples are being taken, and (v) accessibility considerations.

The purpose of sampling is to obtain estimates of parameters without observing or measuring every element of a soil deposit. From a practical point of view this is impossible without disturbing the adjacent elements. Thus, there must be some minimum distance between boreholes, as well as samples between points in a borehole. Obviously, there must be some upper bound for these estimates. The optimal distance would lie between these limits. The optimal sampling plan would also depend on whether the aim is to obtain (a) the highest precision for a fixed sampling plan or (b) the lowest sampling cost for a fixed precision. The precision is defined in terms of variance (dispersion from the mean value) of the parameter under consideration. For a given mean value, a smaller variance is always preferable.

2.2 Uncertainties in Site Characterization

The site characterization problem is complicated by the presence of the following major sources of uncertainties. The natural deposit is basically nonhomogeneous. In real cases, layer interface fluctuations and engineering soil properties usually exhibit considerable variation from point to point, even within a nominally homogeneous layer (Baecher, 1973; Vanmarcke, 1977). The other source of uncertainty is the limited availability of information, both in space and time, about the site subsurface conditions. The limited soil samples obtained from the site are tested under idealized test conditions. This increases the uncertainty level. Another important source of uncertainty is the difference between the measured and the actual field soil properties. If an indirect relationship is used to estimate a soil parameter (for example, estimation of relative density from the standard penetration values using Gibbs and Holtz's relationship), that adds another source of uncertainty (Haldar and Tang, 1979, 1981). Also, the interpretation of a test result can introduce some uncertainty in terms of biasedness.

2.3 Observational Method in Site Characterization

The first stage in any geotechnical problem is the development of a working hypothesis, i.e., the best estimate of the state of the soil deposit. These working hypotheses are confirmed or rejected as tests are made. This is an indirect way to model the inherent uncertainties. This leads to the observational method. Peck (1969) summarized the observational method as follows: (i) initial exploration to develop a working hypothesis, (ii) assessment of the most probable conditions and the most unfavorable conceivable deviations from these conditions, (iii) formalization of the most probable working hypothesis, (iv) selection of parameters to be observed, (v) estimation of these parameters using the available information, (vi) selection of probable course of action in case of significant deviation of the observational findings from the working hypothesis, (vii) measurement and evaluation of actual conditions, and (viii) appropriate modifications.

2.4 Sampling Plan

The basic objective in a sampling plan is to randomize collection of samples. This randomization can be theoretically done in four ways: simple random sampling, systematic random sampling, stratified random sampling and cluster random sampling. Simple random sampling basically indicates that each element in the sampled population has an equal probability of being sampled. In a systematic sampling plan the first sample is chosen at random and the subsequent samples are chosen in a methodical way, e.g., in a grid arrangement. The stratified random sampling plan can be used in a layered deposit where random sampling can be done for each layer to estimate the properties of that layer. As the name suggests, cluster sampling is done in a cluster of units rather than as individual elements. The details of these sampling plans cannot be discussed here for the sake of brevity. However, one sampling plan may have definite advantages over the others for a particular geotechnical problem. Thus, for each criterion (precision or cost), each sampling strategy can be studied and inferences can be made as to which one is preferable for the site under consideration.

3 PROBABILISTIC SITE CHARACTERIZATION

Sampling information obtained by using any of the sampling plans mentioned in Section 2.4 can be used to estimate the sample mean, variance, standard deviation or coefficient of variation (COV) (Haldar, 1981). However, these simple statistics may not be adequate to characterize a site. Even within a nominally homogeneous area, if the values of a parameter are plotted in the one vertical and two horizontal directions, some pattern of variation or fluctuation would emerge. The pattern of this fluctuation is of considerable interest. Thus, in addition to the standard statistics, the scale of fluctuation needs to be estimated.

3.1 Scale of Fluctuation

The scale of fluctuation shows relatively strong correlation of persistance from point to point. A small value implies rapid fluctuations about the average, while large values suggest that a slowly varying component is superimposed on the average value. The scale of fluctuation can also provide a host of practical information for the site exploration problem; for example, to avoid wasteful redundancy in information gathering, sampling distances should be chosen in such a way that they are large in comparison with the scale of fluctuation. On the other hand, when a soil property is being determined by two different tests, the locations of pairs of samples should be well within the correlation distance for maximum effectiveness.

3.2 Evaluation of Scale of Fluctuation

Several methods are available to estimate the scale of fluctuation. The scale of fluctuation, θ, in any direction (any of the two horizontal and one vertical) can be theoretically estimated by using the random field theory. It can be estimated from the information on variance function, correlation function and the unit area one-sided spectral density function (Vanmarcke, 1983).

3.2.1 Scale of fluctuation using variance function

Consider X(t) as a stationary random process. Averaging X(t) over a duration T, a family of moving average process $X_T(t)$ can be obtained as:

$$X_T(t) = \frac{1}{T}\int_{t-T/2}^{t+T/2} X(u)\ du \tag{1}$$

The variance of $X_T(t)$ can be shown to be:

$$\mathrm{Var}\ (X_T) = \Gamma(T)\ \sigma^2 \tag{2}$$

in which $\Gamma(T)$ is the variance function of X(t) and σ^2 is the point variance. Knowing the variance function, the scale of fluctuation θ can be estimated as

$$\theta = \lim_{T\to\infty} T\Gamma(T) \tag{3}$$

3.2.2 Scale of fluctuation using correlation function

Correlation function $\rho(\tau)$ represents the correlation structure of two points of X(t) separated by τ in a nondimensional way. The relationship between $\Gamma(T)$ and $\rho(\tau)$ can be shown to be

$$\Gamma(T) = \frac{2}{T}\int_0^T (1 - \frac{\tau}{T})\ \rho(\tau)\ d\tau \tag{4}$$

The scale of fluctuation in this case can be shown to be:

$$\theta = 2\int_0^\infty \rho(\tau)\ d\tau \tag{5}$$

3.2.3 Scale of fluctuation using one-sided spectral density function

When the stationary random process X(t) is represented in the frequency domain, it produces a spectral density function, $S(\omega)$. Since $S(\omega)$ is symmetric about ω=0, it can be expressed as a one-sided spectral density function $G(\omega)$ for ω>0. When $G(\omega)$ is normalized with respect to σ^2 mentioned earlier, it produces a unit area one-sided spectral density function, $g(\omega)$. In this case, the scale of fluctuation can be estimated as

$$\theta = \pi\ g(\omega)\ ;\ \text{when}\ \omega = 0 \tag{6}$$

3.2.4 Scale of fluctuation for geotechnical problems

For a geomechanics problem, (3),(5), and (6) cannot be used because $\Gamma(T)$, $\rho(\tau)$ and $g(\omega)$ may not be known. Information that is available is a series of boring data. In this case, the h-step variance function can be estimated as

$$\Gamma(h) = \frac{1}{(\ell-h+1)} \sum_{a=1}^{\ell-h+1} \left\{ \frac{1}{m}\sum_{j=1}^{m} \left[\sum_{i=a}^{h+a-1} u_j'(i) \right]^2 - \left[\frac{1}{m}\sum_{j=1}^{m} \sum_{i=1}^{h+a-1} u_j'(i) \right]^2 \right\} \tag{7}$$

in which ℓ is the number of layers in the deposit, h is the averaging interval, m is the number of borings, $u_j'(i)$ is the transformed soil property in the ith layer and jth boring = $u_j(i)$ - trend in (i), $u_j(i)$ is the soil property in the ith layer and jth boring.

3.2.5 Some practical approaches

The concept of coefficient of correlation between values of u(t) at two points can be used to estimate the scale of fluctuation for geotechnical applications. When points are located at very small intervals, the correlation coefficient will be close to 1, and it usually decays as the distance increases. For an assumed theoretical correlation model (Vanmarcke, 1977), the scale of fluctuation can be estimated. Another approximate method can be used to estimate θ if a reasonably complete record of u(t) is available. It is based on the approximate relationship between the scale of fluctuation and the average distance, $\bar{d}$, between the intersections of the fluctuating property, u(t), and its mean. The average distance between mean crossings is approximately (Vanmarcke, 1977):

$$\bar{d} \simeq \frac{\pi}{2} \cdot \theta \simeq 1.25\ \theta \tag{8}$$

A deposit could have three different scales of fluctuation in the three different directions. In many sites, the scales of fluctuation in the two horizontal directions would be the same. The information on the scale of fluctuation could be used to estimate the statistics of spatially averaged soil properties.

4 SPATIALLY AVERAGED SOIL PROPERTIES

Consider a site being investigated for liquefaction potential. It is very likely that a very loose pocket of sand is located during the subsurface investigation. It is known that to cause noticeable damage to a structure located on the site, a sufficient volume of sand needs to liquefy. The evaluation of the site considering the loose pocket of sand is thus obviously incomplete. Some sort of spatially averaged soil properties in the critical soil volume needs to be used for this purpose.

Consider a statistically homogeneous soil parameter, such as relative density, for the liquefaction problem. The value of the soil parameter at a location, x , y, and z from a referenced origin can be represented as U(x, y, z). The spatial average of the soil parameter u(x, y, z) over a volume Δv, $u_{\Delta v}$, can be estimated as

$$u_{\Delta v} = \frac{1}{\Delta v} \int_{\Delta x} \int_{\Delta y} \int_{\Delta z} u(x,y,z)\ dxdydz \quad (9)$$

in which $\Delta v = \Delta x \cdot \Delta y \cdot \Delta z$ and Δx, Δy, and Δz are the length of the soil volume in the x, y, and z directions, respectively. For a statistically homogeneous soil deposit the point mean $\bar{u}$ and variance Var(u) can be estimated from the field observation. The spatial mean, $\bar{u}_{\Delta v}$, and the spatial variance, $Var(u_{\Delta v})$, can be shown to be

$$\bar{u}_{\Delta v} = \bar{u} \quad (10)$$

and

$$Var(u_{\Delta v}) = \Gamma_u^2 (\Delta v)\ Var(u) \quad (11)$$

in which $\Gamma_u^2 (\Delta v)$ is the variance function. It describes the decay of the variance of the spatial average as the averaging dimensions increase. If the correlation structure of u(x, y, z) is separable, then (11) reduces to

$$Var(u_{\Delta v}) = \Gamma_u^2 (\Delta x)\ \Gamma_u^2 (\Delta y)\ \Gamma_u^2 (\Delta z)\ Var(u) \quad (12)$$

in which $\Gamma_u^2 (\Delta x)$, $\Gamma_u^2 (\Delta y)$, and $\Gamma_u^2 (\Delta z)$ are the variance functions in the X, Y, and Z directions, respectively.

The variance function can be calculated from the information on the scale of fluctuation in a given direction. For all practical purposes, the variance function in the X direction can be estimated as:

$$\Gamma_u^2 (\Delta x) = 1.0;\ \Delta x \leq \theta_{u_x} \quad (13)$$

$$= \frac{\theta_{u_x}}{\Delta x};\ \Delta x > \theta_{u_x}$$

in which θ_{u_x} is the scale of fluctuation of u in the X direction. The variance functions in the other two directions can similarly be estimated from the knowledge of the corresponding scale of fluctuation.

5 RISK OF LIQUEFACTION

The probability that the soil volume Δv will liquefy is given by the probability of the event $\tau_{R_{\Delta v}} \leq \tau_{A_{\Delta v}}$. Since the statistics of $\tau_{A_{\Delta v}}$ and $\tau_{R_{\Delta v}}$ cannot be adequately defined beyond the first two moments, for simplicity, lognormal distributions can be prescribed for $\tau_{A_{\Delta v}}$ and $\tau_{R_{\Delta v}}$ in estimating the probability of liquefaction. The probability of liquefaction is thus given by the following:

$$p_{f_{\Delta v}} = 1 - \Phi\left(\frac{\ln\left[\frac{\bar{\tau}_{R_{\Delta v}}}{\bar{\tau}_{R_{\Delta v}}} \sqrt{\frac{1+\Omega^2_{\tau_{A_{\Delta v}}}}{1+\Omega^2_{\tau_{R_{\Delta v}}}}} \right]}{\sqrt{\ln\left[(1+\Omega^2_{\tau_{R_{\Delta v}}})(1+\Omega^2_{\tau_{A_{\Delta v}}})\right]}} \right) \quad (14)$$

where $\Phi(\ )$ is the standard normal cumulative distribution function.

Detailed procedures for estimating the mean and COV of $\tau_{A_{\Delta v}}$ and $\tau_{R_{\Delta v}}$ cannot be presented here due to lack of space. These are discussed in detail by Haldar and Miller (1982) elsewhere.

A site in Niigata, Japan, which liquefied during the 1964 earthquake is considered. The magnitude of the earthquake was 7.5 and the site experienced 0.16 g maximum ground acceleration (Haldar and Miller, 1982). Considering a liquefied volume of sand as 61 m x 61 m x 1.5 m, the risk of liquefaction can be estimated as 0.98. Details of the site conditions and calculation procedures cannot be given here due to lack of space.

6 CONCLUSIONS

The estimation of parameters in any geomechanical problem is important and needs serious consideration. The predictability of the method depends on this often overlooked step, particularly when the parameters are estimated from the data from a few poorly sampled, widely spaced and often poorly logged borings. Methods are discussed here to model three-dimensional characteristics of a soil deposit statistically. A statistical site characterization model could complement a geomechanical problem.

7 ACKNOWLEDGEMENTS

This material is based upon work partly supported by the National Science Foundation under Grant No. CEE-8111691. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the writer and do not necessarily reflect the views of the National Science Foundation.

8 REFERENCES

Baecher, G.B. (1973). Site Exploration: A Probabilistic Approach. Thesis (Ph.D). Massachusetts Institute of Technology.

Dowding, C.H. (1978). Perspectives and Challenges of Site Characterization. Proc. Site Characterization and Exploration, Northwestern University, Illinois, pp. 10-35.

Haldar, A. (1981). Chapter 6 - Statistical Methods in Numerical Methods in Geomechanics. Holland, D. Reidel Publishing Company.

Haldar, A. and Miller, F.J. (1982). Probabilistic Evaluation of Damage Potential in Earthquake-Induced Liquefaction in a 3-D Soil Deposit. Report No. SCEGIT-101-82, Georgia Institute of Technology, Atlanta, Georgia.

Haldar, A. and Tang, W.H. (1979). Uncertainty Analysis of Relative Density. Journal of the Geotechnical Engineering Division, ASCE, Vol. 105, No. GT7, pp. 899-904.

Haldar, A. and Tang, W.H. (1981). Statistical Study of Uniform Cycles in Earthquake Motions. Journal of the Geotechnical Engineering Division, ASCE, Vol. 107, No. GT5, pp. 577-589.

Mitchell, J.K. (1978). In Situ Techniques for Site Characterization. Proc. Site Characterization and Exploration, Northwestern University, Illinois, pp. 107-129.

Osterburg, J.O. (1978). Failures in Exploration Programs. Proc. Site Characterization and Exploration, Northwestern University, Illinois, pp. 3-9.

Peck, R.B. (1969). Advantages and Limitations of the Observational Method in Applied Soil Mechanics. Geotechnique 19, No. 2, pp. 171-187.

Sowers, G.F. (1979). Introductory Soil Mechanics and Foundations: Geotechnical Engineering. 4th ed. New York, Macmillan Publishing Co. Inc.

Underwood, L.B. (1974). Exploration and Geologic Prediction for Underground Works. Proc. Specialty Conference on Subsurface Exploration for Underground Excavation and Heavy Construction, Henniker, New Hampshire, pp. 65-83.

Vanmarcke, E.H. (1977). Probabilistic Characterization of Soil Profiles. Journal of the Geotechical Engineering Division, ASCE, Vol. 103, No. GT11, pp. 1227-1246.

Vanmarcke, E.H. (1983). Random Fields Analysis and Synthesis. Cambridge, Massachusetts, The MIT Press.