FACTA UNIVERSITATIS
Series: **Architecture and Civil Engineering** Vol. 8, N° 1, 2010, pp. 111 - 118
DOI: 10.2298/FUACE1001111D

MODELING OF SOIL PARAMETERS SPATIAL UNCERTAINTY BY GEOSTATISTICS

UDC 622.83:519.711+519.245(045)=111

Nebojša Davidović, Verka Prolović, Dragoslav Stojić

University of Niš, Faculty of Civil Engineering and Architecture, Serbia
E-mail: dnebojsa@gaf.ni.ac.rs

Abstract. *Geotechnical performance of "soil-structure" systems is strongly dependent on the properties of the soil and prediction of the performance of these systems in real conditions requires accurate modeling of soil parameters. With the help of high-speed computers, now it is possible to create advanced constitutive models, but large uncertainties and variations in soil properties could reduce the advantages gained by using such models. In this paper sources and types of uncertainty in geotechnical engineering practice are first presented, followed by a review of the basic concepts and terminology of geostatistics. Finally, procedures for quantification of uncertainty and for geostatistical estimation and simulation of spatially variable soil properties are presented.*

Key words: *Geostatistics, soil parameter, uncertainty, modeling, verification, validation, Kriging, Monte Carlo Method.*

1. INTRODUCTION

Prediction of the performance of "soil-structure" systems in real conditions requires accurate modeling of the geotechnical components in these systems. In recent years, a significant progress has been made in the field of constitutive modeling of materials. With the help of high-speed computers, now it is possible to create complex elasto-plasticity based constitutive models which give very accurate description of material behavior. However, the increasing number of model parameters, which are sometimes very dificult to obtain experimentally, have certain disadvantages. Uncertainty associated with the soil properties could reduce the advantages gained by using advanced and more sophisticated models. For example, Fig. 1 shows influence of material fluctuations on a bilinear elastic-plastic "stress-strain" behavior. The same material model could predict softening (*a*) or hardening (*b*), depending upon the material parameters considered in calibrating the model [1].

Received June, 2010

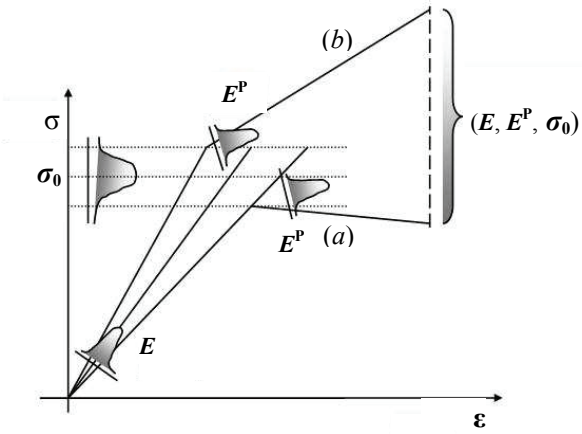

Fig. 1 Influence of material fluctuations on "stress-strain" behavior

During the process of computer-aided modeling it is necessary to implement verification and validation procedures. Verification represents the process of determining whether the products of a given phase of the software (model) development cycle fulfill the requirements established during the previous phase, and model accurately represents the author's concept. At the end of the software (model) development process, validation checks the degree to which a model is an accurate representation of the real world (Fig. 2).

The soil properties can be described using deterministic or probabilistic models. Deterministic models use a single discrete value for the parameter of interest. Probabilistic models describe parameters by using discrete statistical descriptors of probability distribution (density) functions.

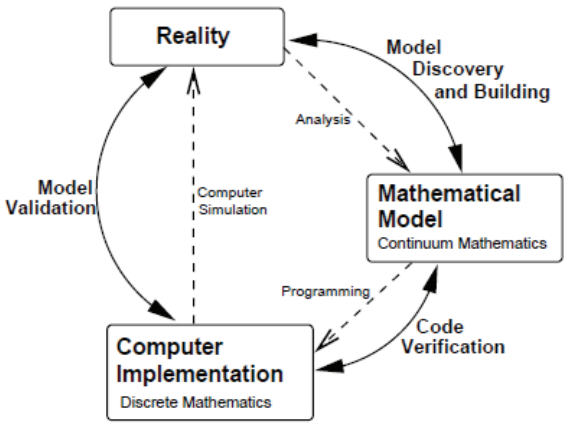

Fig. 2 Verification and validation in computer-aided modeling

EUROCODE 7 offers two different procedures of calibrating partial safety factors to incorporate soil variability in geotechnical design (Fig. 3). Procedure 1 involves the application of deterministic (historical and empirical) methods. Procedure 2 involves the application of semi-probabilistic or full probabilistic methods. The most general and accurate method is the full probabilistic analysis [2].

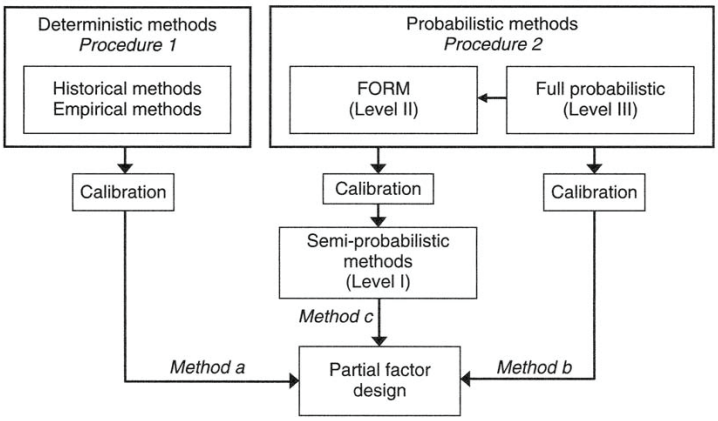

Fig. 3 Consideration of soil variability in geotechnical design in EUROCODE 7

It is important to emphasize that implementation of these methods does not preclude the effects of uncertainty of principal influential factors, but facilitates their quantification, and therefore, their inclusion in the modeling process. The first step in this process is identification of potential sources of uncertainty in the values of geotechnical soil parameters [3].

2. Sources of Uncertainty in Geotechnical Soil Parameters

Uncertainty in geotechnical soil properties can be grouped into random (*aleatory*[1]) and subjective (*epistemic*[2]) uncertainty (Fig. 4). Aleatory uncertainty represents the natural randomness of a soil property and is a function of the spatial variability of the soil property. This type of uncertainty cannot be reduced or eliminated. Epistemic uncertainty results from a lack of information and shortcomings in measurement and calculation[4]. Sampling uncertainty is present because the parameters are estimated from a limited set of data, while testing uncertainty is due to imperfections of an instrument or of a method to register a quantity. Epistemic uncertainty can be reduced by collecting more information or improving measurement methods. Human error would be considered a third source of uncertainty, however, as it is difficult to isolate, its effects are usually included in compilations of statistics on aleatory uncertainty.

[1] lat. *aleator* – gambler
[2] gr. *επιστημη* – knowledge

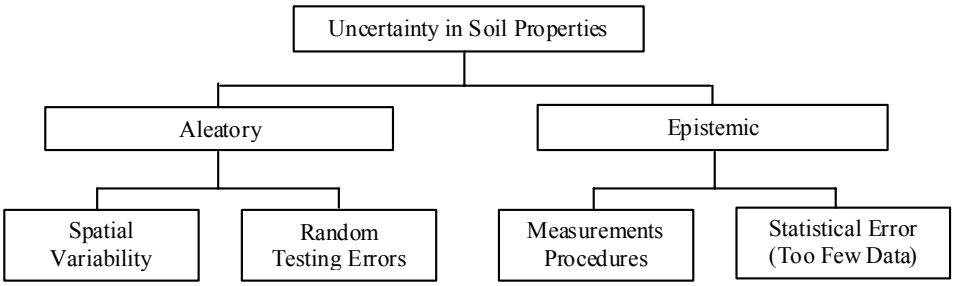

Fig. 4 Sources of uncertainty in geotechnical soil parameters

As an example, Standard Penetration Test (SPT) procedure includes sources of aleatory uncertainty in the measured SPT resistance occurring due to the natural variability of the soil deposit and random testing errors (such as that caused by a presence of an erratic boulder). Sources of epistemic uncertainty could include non-standard equipment and insufficient data to form reasonable statistics (such as one boring over a large site) [5].

Over the years geotechnical engineers have developed several strategies for dealing with uncertainty. They include: (1) "by ignoring"; (2) "by being conservative"; (3) "by using observational method" or (4) "by quantifying the uncertainty" [6]. Quantification or mathematical description of uncertainty is usually done within the framework of probability theory, where random soil parameters are modeled as random variables. Geotechnical engineering by its nature demands tools, which can improve design and help handle the large uncertainties inherent in soil properties better. One such tool, which can be regarded as a set of statistical estimation tools for modeling the spatial variability of a quantity, is geostatistics.

3. Modeling by Geostatistics

Geostatistics, as a methodology for estimating recoverable reserves in mining deposits, was mathematically formalized by French professor Georges Matheron 1963, inspired by the pioneering work of South African mining engineer D.G. Krige in the 1950's. Estimation of recoverable reserves has to be made based on very sparsely sampled information. The ratio of the volume of samples from exploration boreholes to the volume of an ore deposit is on the order of $1x10^{-9}$. Yet, recoverable reserves have to be reliably estimated based on this information, and decisions made on investing large sums of money into developing the deposit. Geotechnical engineering has similar concerns. The volume of samples extracted for characterizing soil masses constitutes only a minute fraction of the volume of material that impacts engineering behavior. The engineering properties of soil masses are heterogeneous, varying from point to point. But as a rule, geotechnical engineers assume properties to be the same throughout a domain. It is however well known that use of averaged parameter values can lead to conclusions, which significantly differ from actual behavior, and it is acknowledged that more accurate knowledge of the spatial distribution of material properties promotes safe and economic design.

The spatial distribution of geotechnical properties in natural soil deposits is difficult to predict deterministically. Limited sampling further complicates prediction of soil proper-

ties. Conventional statistical analysis of core samples from a site investigation program might show that measured cohesion values can be described by a normal distribution. However, this distribution only describes the population of values gathered; it does not offer any information about which zones are likely to have high cohesion values and which areas low values.

Geostatistical analysis examines spatial relationships in addition to establishing statistical distribution of data. In the given example, geostatistics can reveal how cohesion varies over distance, and can predict localized zones of high and low cohesion values. Application of geostatistics will lead to more ready incorporation of the inherent uncertainty of soil masses into numerical models and into the design process in general.

Phoon and Kulhawy (1999) [7] proposed a simple model for description of spatial variation of soil properties:

$$\xi(z) = t(z) + w(z) + e(z) \tag{1}$$

where: z - depth; $\xi(z)$ – in situ soil property; $t(z)$ – deterministic trend function; $w(z)$ – deviation from trend (random component); $e(z)$ – measurement error. These components are presented in Fig. 5. The trend (t) can be determined by fitting a smooth deterministic function (straight line, parabola, exponential) to the data. Measurement error (e), whether it be from laboratory or field measurements, can introduce additional variability into soil properties. The scale of fluctuation (θ) describes the spatial fluctuation of the property of interest about the trend. A parameter with a short scale of fluctuation changes rapidly with position, while a long scale of fluctuation indicates low spatial variability of the property. For both laboratory and field measured data, the scale of fluctuation in the horizontal direction is much higher than in the vertical direction [8].

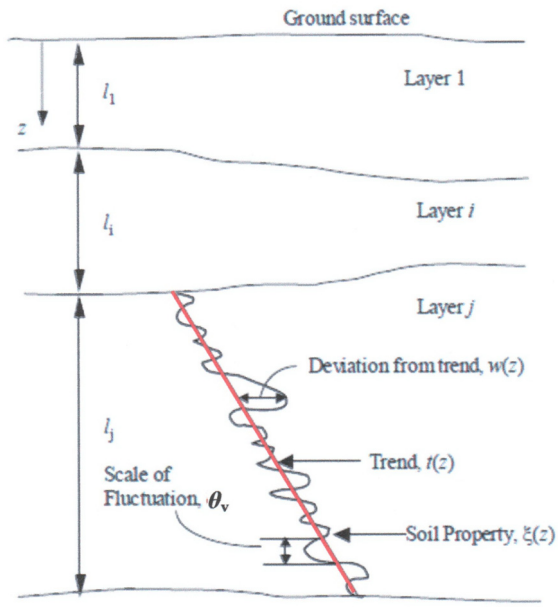

Fig. 5 Spatial variation of soil properties

Among the spatial interpolation (geostatistical estimation) techniques, a process called *Kriging* [9] is the most useful. If soil property data are available at specific locations, it is possible to estimate the value of that property at other locations through Kriging. The goal of Kriging is to predict the average value of X at specific point of soil mass. If $X_1, X_2, ..., X_n$ are known values of parameter at the points $x_1, x_2, ..., x_n$ defined by coordinates, then the estimated value of parameter at point x is given by:

$$\hat{X}(x) = \sum_{i=1}^{n} \beta_i X_i \tag{2}$$

where β_i are weights applied to the respective values X_i, such that: $\sum_{i=1}^{n} \beta_i = 1{,}0$.

If the point x is close to point x_k, then the weight β_k (associated with known value X_k) would be high. However, if $X(x)$ and X_k are in different soil layers, then β_k should be small. The weights β_i are determined by using correlation coefficient between two points since this reflects not only distance but also the effects of differing geologic units [10]. In Fig. 6 is given a simple example of Kriging procedure of estimation of the value of the parameter in the point x which is equally distant from the points x_1, x_2 and x_3 in which the values of the parameters are known and amount to 1, 2 and 3. In the case (a) the distance between the points x_2 and x_3 is 20 m, and the value of the parameter in the point x estimated by Kriging is 1,767. In the case (b) the points x_2 and x_3 are at the distance of 174 m, and the estimated value (by Kriging) of the parameter in the point x is 2,000.

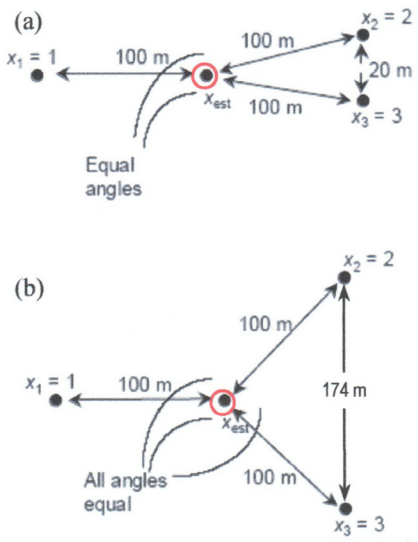

Fig. 6 Kriging estimation of parameter value at point x

However, if soil property data is not available at specific locations, but "only" distribution and autocorrelation function for the data are known, geostatistical simulation techniques can be used to generate the values of that property at other location. The value estimated by Kriging procedure then can be viewed as one of ghe values of he function of a

random variable. Geostatistic simulation, applying the *Monte Carlo method* (MCM) [11] generates other values and take into account only those meeting previously defined criteria. In the general case, the MCM procedure can be divided into the following steps:

- choice of a model that will produce a deterministic solution to a problem,
- identification of all the independent variables and choice of input parameters whose variability will be modeled by the probability distribution functions,
- repeated estimation of input parameter values which are in accordance with the corresponding distribution functions,
- repeated solving of the problem using the deterministic model,
- determination of statistic indicators of the calculated value of X (mean value μ_X, standard deviation σ_X, distribution function $f_X(x)$).

Solution of this simulation is in principle "sensitive" to a number of iterations. The higher the number of iterations, the more accurate the solution. The required number of iterations increases by geometrical progression, with the increase of the number of input parameters and required level of reliability of solution. For example if the required level of reliability is 90%, the required number of iterations for one variable (input parameter) will be 67, for two variables 4521, and for three variables 304007. However, the practice demonstrated that in almost all cases, even in those with multiple input parameters and high required reliability level, the solution becomes "stable" after first several thousand iterations [12]. With computers becoming faster and faster, MCM is becoming increasingly attractive.

Geostatistic simulation can help geotechnicians to assess the risk introduced to calculations by the indeterminacy of soil parameters. The spatial arrangement of realizations (obtained values) of variables (parameters) obtained during simulations can be introduced to numerical models and used for estimation of this risk.

4. Conclusions

For geotechnical engineering, one of the most significant potentials of geostatistic analysis is creation of relatively simple but accurate models of inhomogeneous material on the basis of a limited range of data obtained from the available samples from the given location. It is known that the most challenging task in exploration of some location is defining of such a sampling program, which would incur minimum cost and simultaneously facilitate the best insight of the conditions in the underground soil. This task is more easily done if potentials of geostatistic in defining of the optimal distribution of sampling/testing points are used, in order to obtain as much information at as low a price as possible.

Geostatistic analysis provides a 3-dimensional visualization of spatial variability of soil parameters, through creation of maps with data on the parameter values and their distribution across the surface or in the mass of the soil. By drawing contour maps of standard deviations of estimated values of this parameter at the places where the samples have not been taken and tested, zones of increased uncertainty (due to higher standard deviations) can be observed. On the basis of this knowledge, a geotechnical engineer can ask for the additional samples of these zones, with the goal of a more complete analysis of soil conditions.

Geostatistics is an applied discipline, its development has been the work of mining engineers, petroleum engineers, soil scientists, hydrologists as well as statisticians, so there are successful applications to a variety of fields. As example, application of geostatistical analysis to Channel Tunnel Project enabled the careful assessment of geological risks, and

was used to improve the originally proposed alignment of the tunnel [13]. If there were any doubters to the usefulness of geostatistics for geotechnical engineering, this project should have helped ease their fears.

REFERENCES

1. Jeremić, B.: Computational Geomechanics: Lecture Notes. Department of Civil and Environmental Engineering, University of California, Davis, 2006.
2. Huber, M., Moellmann, A., Vermeer, P.A., Bárdossy, A.: Contributions to probabilistic soil modelling. Proceedings of the 7th International Probabilistic Workshop, Delft, 2009, pp. 1-12.
3. Davidović, N.: Kvantifikacija neodređenosti i promenljivosti glavnih uticajnih faktora u geotehničkim računskim analizama stabilnosti. Zbornik radova sa Naučno-stručnog savetovanja "Geotehnički aspekti građevinarstva", Soko Banja, 2007, str. 67-73.
4. Lacasse, S., Nadim, F.: Uncertainties in characterizing soil properties. Uncertainty in the Geologic Environment: From Theory to Practice, Proceeding of Uncertainty '96, Geotechnical Special Publication No. 58, 1996.
5. Jones, A. L., Kramer, S. L., Arduino, P.: PEER Report 2002/16, Estimation of Uncertainty in Geotechnical Properties for Performance-Based Earthquake Engineering. Pacific Earthquake Engineering Research Center, College of Engineering, University of California, Berkeley, 2002.
6. Christian, J. T.: Geotechnical engineering reliability: How well do we know what we are doing? Journal of Geotechnical and Geoenvironmental Engineering 130(10), 2004, pp. 985-1003.
7. Phoon, K. K., Kulhawy, F. H.: Characterization of geotechnical variability. Canadian Geotechnical Journal 36(4), 1999, pp. 612-624.
8. Baker, J., Calle, E., Rackwitz, R.: Joint committee on structural safety probabilistic model code, section 3.7: Soil properties, updated version, 2006.
9. Baecher, G. B., Christian, J. T.: Reliability and Statistics in Geotechnical Engineering, John Wiley & Sons. Chichester, West Sussex, England; Hoboken, NJ., 2003.
10. Fenton, G. A.: Data analysis / Geostatistics. Probabilistic Methods in Geotechnical Engineering, ASCE GeoLogan '97 Conference, Logan, Utah, 1997, pp. 51-73.
11. Hutchinson, S., Bandalos, D.: A Guide to Monte Carlo Simulation Research for Applied Researchers. Journal of Vocational Education Research 22, 1997, pp. 233-245.
12. Davidović, N.: Primena probabilističkog koncepta u računskoj analizi stabilnosti kosina. Zbornik radova sa Naučno-stručnog savetovanja "Geotehnički aspekti građevinarstva", Kopaonik, 2005, str. 223-228.
13. Blanchin, R., Chilès, J. P.: The Channel Tunnel: Geostatistical prediction of the geological conditions and its validation by the reality. Mathematical Geology 25, 1993, pp. 963-974.

MODELIRANJE PROSTORNE NEODREĐENOSTI PARAMETARA TLA PRIMENOM GEOSTATISTIČKIH METODA

Nebojša Davidović, Verka Prolović, Dragoslav Stojić

Geotehničko ponašanje sistema "tlo-konstrukcija" u velikoj meri zavisi od karakteristika tla, tako da predviđanje ponašanja ovih sistema u realnim uslovima zahteva precizno modeliranje parametara tla. Uz pomoć sve moćnijih kompjutera sad je moguće kreirati napredne konstitutivne modele, ali jako izražena neodređenost i promenljivost osobina tla može da umanji prednosti korišćenja ovakvih modela. U radu su prvo prikazani izvori i kategorije neodređenosti u geotehničkoj inženjerskoj praksi, a zatim su predstavljeni osnovni koncept i terminologija geostatističkih metoda modeliranja. Na kraju su prikazani postupci za kvantifikaciju neodređenosti, kao i za geostatističku procenu i simulaciju prostorne promenljivosti parametara tla.

Ključne reči: *Geostatistika, parametar tla, neodredjenost, modeliranje, verifikacija, validacija, Kriging, Monte Carlo Metod*