Spatial characterization of V_s at Amherst NGES site from SCPT using Bayesian kriging

M. Uzielli, J. Facciorusso
Georisk Engineering S.r.l., Florence, Italy

P.W. Mayne
Georgia Institute of Technology, Atlanta USA

ABSTRACT: This paper illustrates a case study application of geostatistical kriging to shear wave velocity measurements from seismic piezocone (SCPT) testing in clay at the Amherst national geotechnical experimentation site (NGES) site. Analyses are performed in a pseudo-3D perspective, by which data are clustered into samples based on measurement depth. Bayesian kriging is performed at each depth cluster. The site-specific operational procedure is described, with particular reference to important aspects such as the identification of best-fit semivariogram parameters.

1 INTRODUCTION

Engineering soil properties exhibit variation from point to point even within a soil layer that is characterized as homogeneous. The variability in geotechnical data can be ascribed to both the soils and the investigators. Soils are natural materials, which are formed and continuously modified by complex geological processes, as discussed in detail by Hight & Leroueil (2003). The variety and complexity of such processes result in physical and mechanical heterogeneity and, consequently, in the variability of quantitative parameters. In any type of real-world geotechnical problem, it is impossible to obtain exhaustive values of data at every desired point because of practical and economical constraints. Hence, any engineering analysis must face the presence of uncertainties. *Aleatory uncertainty* parameterizes the real, inherent soil variability. *Epistemic uncertainty* exists as a consequence of the investigator‘s invariably limited information and imperfect measurement and modeling capabilities.

Proper interpretation and interpolation of the subsurface conditions to predict the unknown values in an area (or volume) of interest from the data observed at known locations is paramount for geotechnical characterization and design. It is advisable to address uncertainties as rationally as possible to pursue optimum cost-benefit ratio and system performance. Geostatistical kriging includes a family of stochastic interpolation methods which allow the estimation of soil properties at unsampled spatial locations from spatially referenced values. A noticeable advantage of kriging over other interpolation methods is the explicit quantification of the uncertainty (in the form of a kriging variance) associated with kriging estimates.

2 DESCRIPTION OF SITE

The National Geotechnical Experimentation Site (NGES) at Amherst, Massachusetts is located 1.5 km east from the old shore of Glacial Lake Hitchcock in Western Massachusetts. The site is underlain by a thick 30-m deposit of soft lacustrine varved clay where the upper 3 to 4 m consists of a desiccated crust. The ground water table typically occurs in the upper 2 m below ground surface and varies by as much as approximately 2 m throughout the year coinciding with changes in seasonal precipitation. Grain size data from hydrometer analyses of the bulk soil gives a clay fraction of approximately 65% and a silt fraction of 35%. Natural water content and void ratio for the bulk soil are typically lowest near the ground surface. There is a sharp transition in water content through the crust from an average of around 20% at 1 m to an average of around 62% below the crust. The void ratio data follow that of the water content with values around 1.8 below the crust. Test specimens from below the water table are usually found to be 100% saturated. The bulk soil classifies as ML (low plasticity silt), CL (low plasticity clay), MH (high plasticity silt), and CH (high plasticity clay) in the USCS system. The plasticity index, PI, averages 10 in the crust and around 20 below the crust with no significant trend with depth. The difference in PI is due to a change in liquid limit only, which increases from an average of 39 in the crust to an average of 51 below the crust, while the plastic limit is the same for both zones. The overconsolidation ratio decreases from approximately 10 at a depth of 4 m to approximately 2 at a depth of 14 m. Full details on compositional and index properties, as well as the results of laboratory and field tests are given by DeGroot & Lutenegger (2003).

3 DESCRIPTION OF DATA SET

A database consisting of V_s measurements from 15 pseudo-interval SCPT tests is available from Hegazy (1998). These SCPT soundings were conducted in a 60 m x 60 m area shown in Fig. 1 (X and Y are the relative Cartesian coordinates used in the study). Sounding depths slightly exceed 14 m. The depth interval between consecutive measurements is approximately 0.9 m. As shown in Fig. 2, 16 data sets corresponding to distinct analysis depths (D01-D16) can be identified. Table 1 reports the numerosity of the data sets (column 2), along with second-moment statistics of depth and V_s (columns 3 to 6). Geostatistical characterization of the Amherst NGES site was performed in a pseudo-3D perspective, i.e. 2-D kriging was performed on hori-

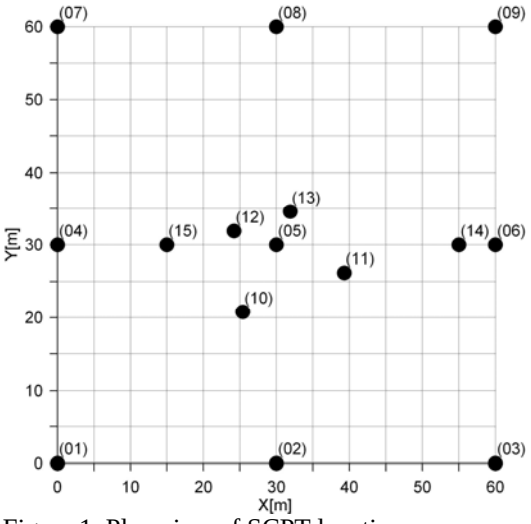

Figure 1. Plan view of SCPT locations

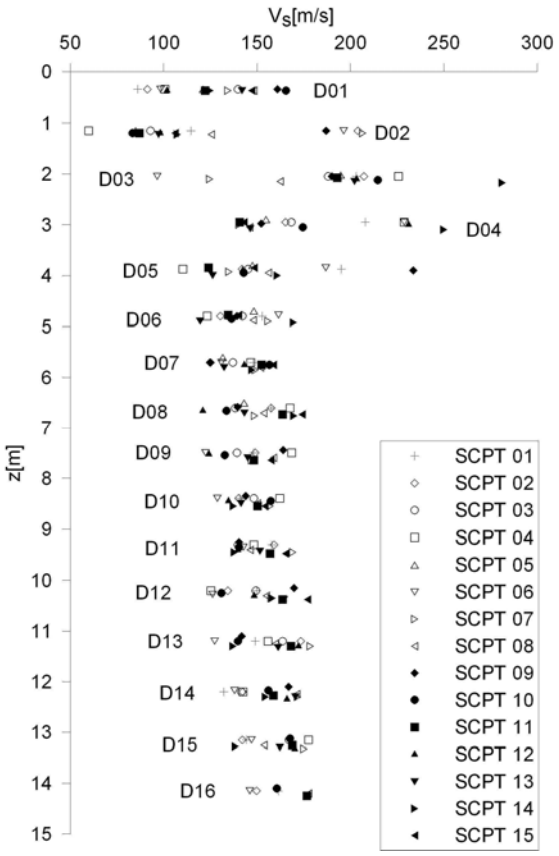

Figure 2. SCPT measurements of V_s at Amherst NGES site

Table 1. Second-moment statistics of depth-wise data sets D01-D16

	N	$\bar{z}$ [m]	COV(z)	$\bar{V}_s$ [m/s]	COV(V_s)
D01	15	0.36	0.04	125.55	0.20
D02	15	1.18	0.03	123.33	0.40
D03	15	2.09	0.02	191.85	0.22
D04	15	2.99	0.02	178.50	0.22
D05	15	3.90	0.01	153.03	0.20
D06	15	4.82	0.01	142.39	0.10
D07	15	5.74	0.01	142.35	0.08
D08	15	6.65	0.01	149.98	0.10
D09	15	7.55	0.01	145.58	0.09
D10	15	8.45	0.01	146.21	0.06
D11	15	9.36	0.01	149.33	0.07
D12	15	10.26	0.01	149.67	0.11
D13	15	11.20	0.01	157.22	0.10
D14	14	12.21	0.01	154.66	0.09
D15	13	13.19	0.01	159.44	0.08
D16	12	14.18	0.01	164.93	0.06

zontal planes corresponding to analysis depths. Average values of measurement depths were taken as nominal depths for each data set D01-D16; these are reported in column 3. The assignment of a unique depth for each data set was deemed acceptable given the small values of the COVs of depth (column 4). The scatter in V_s values is significantly larger in the fill and crust (D01-D05), where COVs of V_s (column 6) range from 0.20 to 0.40, than in the silty clay and clay (D06-D16), where COV(V_s) never exceeds 0.11. It may be noted in Fig. 2 that none of the soundings yields consistently higher or lower values in comparison with other soundings.

4 CHARACTERIZATION OF SPATIAL CORRELATION

In order to apply the kriging interpolation technique successfully, the spatial structure of soil variability must be identified. The method requires a description of how a measurement taken at a point in the soil mass is correlated with the measurements taken at other points. By "correlation", it is meant that the properties of adjacent soil elements tend to be more similar in magnitude in comparison with those of widely separated elements.

Uzielli et al. (2007) provided a state-of-the-art review on the characterization of spatial variability in geotechnical engineering. The most frequently used statistical function for the description of spatial correlation in geostatistics is the *semivariogram*. The semivariogram measures the average dissimilarity between data pairs separated by a lag h. The sample semivariogram is computed as half the average squared distance between the components of every data pair:

$$\gamma(h) = \frac{1}{2Q_h} \sum_{(i,j)\left|h_{ij} \le h+\Delta h\right.} \left(V_{s,i} - V_{s,j}\right)^2 \quad (1)$$

where Q_h is the number of data pairs locations spatially separated by $h+\Delta h$ (at most), ξ_i and ξ_j are the data values at i-th and j-th location respectively. The issue of anisotropy in spatial correlation was not addressed due to the small sample size of data sets. Omnidirectional semivariograms were calculated at lag intervals of 10 m, from 10 m to 60 m. A tolerance of Δh=5 m on separation distance was defined to allow inclusion of a greater number of pairs, i.e. to obtain a clearer semivariogram structure. This tolerance is commonly allowed in geostatistical analyses, especially for dataset with few samples (Isaaks & Srivastava, 1989). Once sample semivariograms have been calculated at each depth, depth-specific semivariogram models able to describe the spatial correlation structure are required to compile the covariance matrices which enter the kriging algorithm (e.g. Goovaerts, 1997). Here, the *spherical model*, the *exponential model* and the *Gaussian model* were considered. Details of such models can be found in dedicated textbooks (e.g. Goovaerts, 1997).

Table 2. Results of spatial correlation analysis

	C_0 [m/s]^2	BFM	a [m]	$C+C_0$ [m/s]^2	quality	$C_0/(C+C_0)$
D01	20	S	10	670	1	0.03
D02	30	G	40	3230	1	0.01
D03	50	G	40	2250	1	0.02
D04	40	E	20	1640	2	0.02
D05	30	G	70	1630	2	0.02
D06	20	S	60	270	3	0.07
D07	20	E	40	130	3	0.15
D08	20	E	20	220	3	0.09
D09	20	S	20	180	2	0.11
D10	20	S	20	120	2	0.17
D11	20	E	20	120	2	0.17
D12	20	S	20	320	2	0.06
D13	20	S	60	300	2	0.07
D14	20	G	50	240	1	0.08
D15	20	G	30	220	2	0.09
D16	20	E	20	140	1	0.14

Measurement uncertainty can be a significant component of total uncertainty, and should not be neglected in principle. Moss (2008) parameterized uncertainty in V_s measurements by a coefficient of variation on the order of 0.01–0.03 for SCPT testing. The upper-bound value of 0.03 is adopted herein as pseudo-interval SCPT is less robust than true-interval SCPT. In the present analysis, measurement uncertainty is accounted for in the spatial correlation analysis phase, by setting a non-zero semivariance at zero lag, i.e. assigning a non-zero semivariogram nugget C_0. At each depth, the value of the semivariogram nugget is obtained by multiplying the assumed coefficient of variation of measurement uncertainty COV_m by a "characteristic value" of V_s, given by the sample mean of V_s measurements at that depth, $m(V_s)$ plus 3 times the sample standard deviation of V_s, $s(V_s)$:

$$C_0 = \frac{1}{2}\mathrm{COV}_m^2\left(V_s\right)\cdot\left[m\left(V_s\right)+3\cdot s\left(V_s\right)\right]^2 \quad (2)$$

In the hypothesis of Gaussianity of data sets, such characteristic value corresponds to the 99.9th percentile, i.e. essentially to a cautious upper-bound value. Semivariogram model fitting should rely on experience rather than on automatized procedures (Isaaks & Srivastava, 1989). Semivariogram model fitting is not a straightforward operation, especially in the presence of small data sets as in the present case. A number of undesirable effects may blur and introduce noise into the experimental semivariogram, thus hindering the identification of the underlying spatial correlation structure (Isaaks

& Srivastava, 1989). Several checks and auxiliary operations, including the calculation of general and pairwise relative semivariograms (e.g. Goovaerts, 1997) were conducted to enhance the analysis of spatial correlation. Details of auxiliary procedures are not reported herein for sake of brevity. Table 2 reports the best-fit semivariogram model (S: spherical; G: Gaussian; E: exponential) and parameters (range a; nugget C_0; sill C_0+C) for each analysis depth, along with the subjective assessment of the quality of the fit (1: good; 2: fair; 3: poor). The last column in Table 2 reports the relative nugget, which expresses the relative magnitude between measurement uncertainty (and small-scale variation) and spatial variability of V_s. It is evident that measurement uncertainty is relatively less relevant in the upper crust, where inherent variability is more pronounced. Results of kriging making use of higher-quality model fitting can be expected to be more reliable. An example of semivariogram fitting (for D10) is shown in Fig. 3.

5 BAYESIAN KRIGING OF SHEAR WAVE VELOCITY

In the context of kriging and in a pseudo-3D perspective, V_s is modeled, at each analysis depth, as a random function $V_s(A)$ of a two-dimensional space A which corresponds to a horizontal plane positioned at the nominal analysis depth. At any spatial location on such plane, the random function can be expressed as the sum of its expectation (or *trend*), $E[V_s(A)]$, and a zero-mean set of residuals, $w(A)$, such that:

$$V_s(A) = E\left[V_s(A)\right] + w(A) \tag{3}$$

In Bayesian kriging (Omre, 1987; Omre & Halvorsen, 1989), it is possible to assign prior estimates to the trend and to the uncertainty associated with the trend itself. Subsequently, the kriging algorithm allows updating of such estimates based on the N available observations of V_s at each analysis depth. Bayesian kriging is a more general case of simple and universal kriging, and has been shown to provide more precise predictions than the latter two algorithms (see e.g. Omre & Halvorsen, 1989). The Bayesian kriging predictor of V_s for any arbitrary location $A_0=(X_0,Y_0)$ has the form:

$$\hat{V}_s(A_0) = \sum_{i=1}^{N}\left[\alpha_{i0}\cdot w(A_i)\right] + E\left[V_s(A_0)\right] \tag{4}$$

The weights α_{i0}; i=1,…,N to be applied to the residuals at the N sampled locations w(A_i) are determined through a procedure described in detail by Omre & Halvorsen (1989). The trend function is updated automatically by the algorithm as well. The kriging algorithm provides reliable solutions if $V_s(A)$ and the vector of trend coefficients which define trend values (not addressed in detail herein; the reader is referred to Omre & Halvorsen, 1989) are at least approximately jointly Gaussian. Normal probability plots were obtained for each analysis depth to allow subjective visual assessment of the adequateness of sets of trend coefficients and data observations in terms of Gaussianity. Given the small sample size of data sets at each of the analysis depths, visual inspection of normal probability plots is at least as reliable as more refined normality tests (Thode, 2002). An example is shown in Fig. 4 for observations at D10. Overall, sets of observations D01-D16 were assessed to be acceptably Gaussian, while sets of trend coefficients did not display a Gaussian behavior.

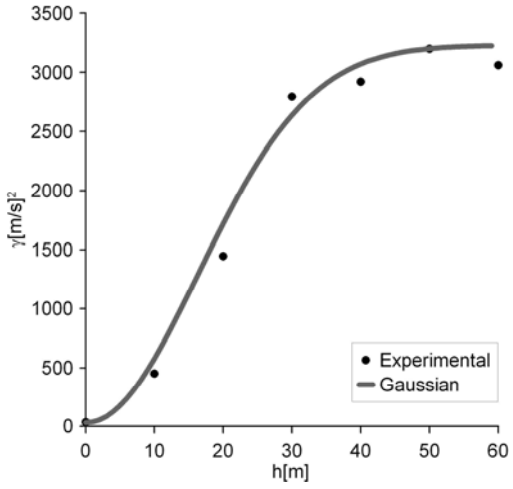

Figure 3. Fitting of a Gaussian semivariogram model to empirical semivariance values at D10

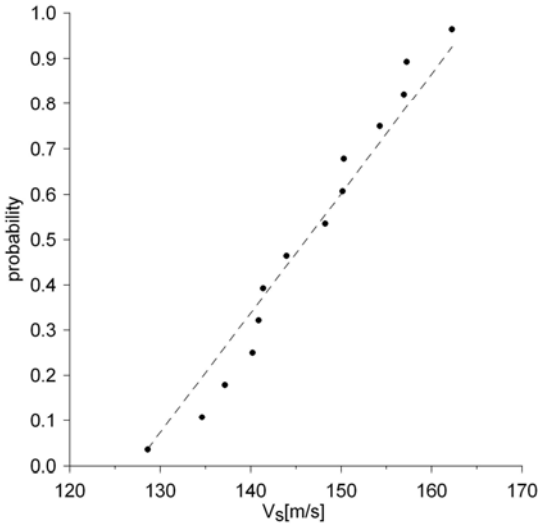

Figure 4. Normal probability plot for Vs measurements at D10

Kriging estimates can be obtained at any desired unsampled target location. Here, a mesh of target locations at horizontal and vertical intervals of 0.5 m was defined in the intervals $X \subset [0,60]$, $Y \subset [0,60]$ identically for all data sets D01-D16. The set of target locations also comprised the sampled locations in order to verify that the output kriging variance at such locations is null. The total number of target points at each depth of analysis is 4225. A second-degree polynomial function was selected as trend surface typology.

Example outputs of the Bayesian kriging algorithm are shown for the upper crust (D04) and the varved clay (D10) in Fig. 5 and Fig. 6, respectively. Each figure contains: (a) the kriging prediction; (b) the associated kriging variance; (c) the "updated" second-degree polynomial trend function; and (d) the residuals, corresponding to the difference between the kriged values and the polynomial trend. A comparative analysis of such results allows appreciation of the different degree of spatial variability between the upper crust and the underlying varved clay. It is also possible to appreciate the capability of the kriging algorithm to quantify the uncertainty associated with the kriged estimates, and, in relation to kriging variance, to note: (1) the higher degree of uncertainty in the D04 estimates, due to the larger scatter in measured values; (2) the spatial variation, for both D04 and D10, of kriging uncertainty, which is larger with increasing distance of target locations from sampled location; and (3) the null value of kriging estimation variance at sampled locations. The kriging variance of V_s as obtained herein could be used for second-moment probabilistic characterization and design, and could also provide a useful indication for the location of additional measurements should the variance exceed a preset desired or imposed threshold. A 3-D map of V_s (and its associated variance) can be reconstructed simply by merging the kriging estimates from the 16 data sets on which individual analyses were performed.

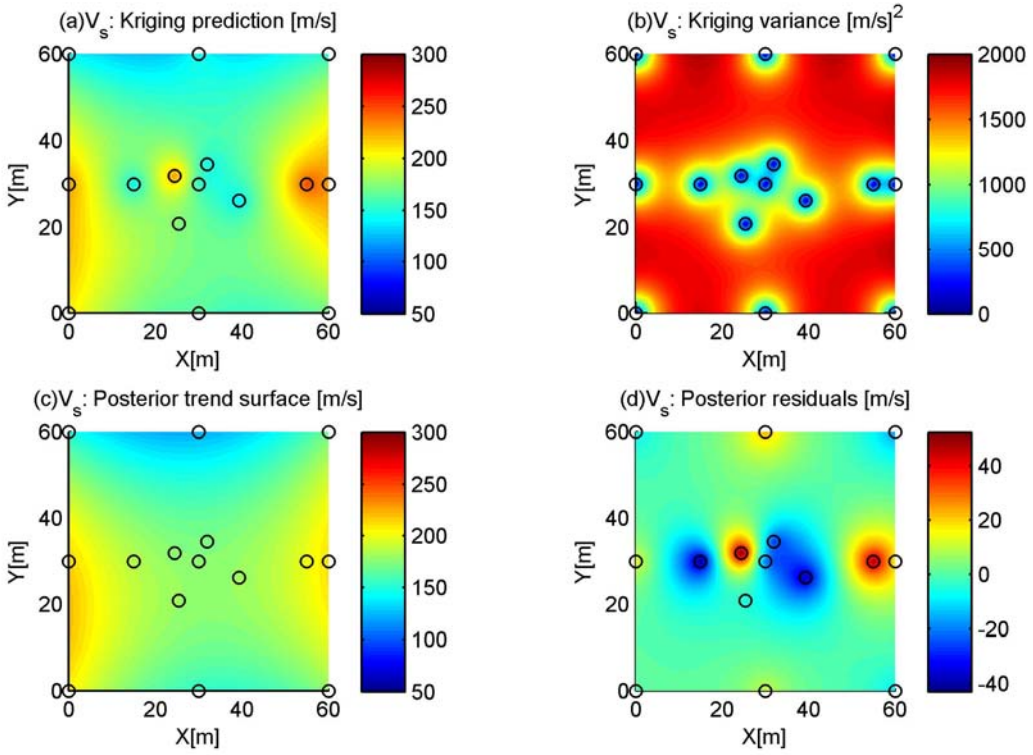

Figure 5. Example output of Bayesian kriging: upper crust (D04)

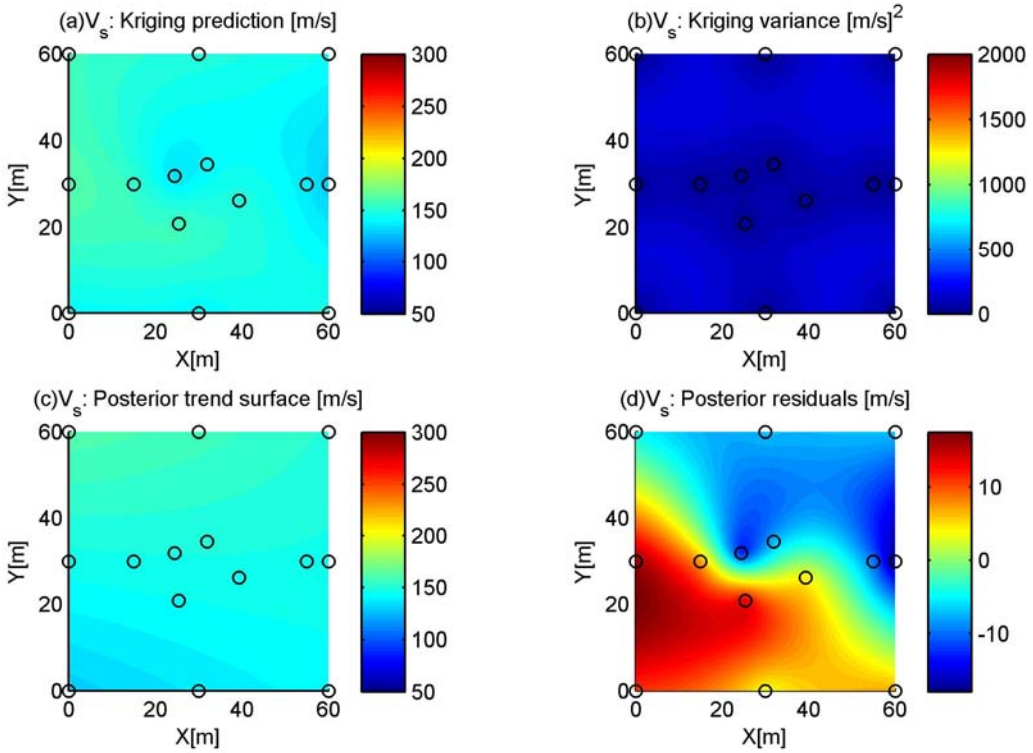

Figure 6. Example output of Bayesian kriging: varved clay (D10)

6 CONCLUDING REMARKS

An example application of Bayesian kriging of shear wave velocity data from SCPT was illustrated herein. The utility of the kriging outputs is manifold: it provides optimized interpolation with respect to assumed site-specific spatial variability, while explicitly quantifying the uncertainty associated with the interpolation. These features are especially notable in the increasingly important perspective of uncertainty-based geotechnical analyses, which are preliminary to the implementation of evolutionary design codes.

REFERENCES

DeGroot, D.J. & Lutenegger, A.J. 2003. Geology and engineering properties of Connecticut Valley varved clay. In T.S. Tan, K.K. Phoon, D.W. Hight & S. Leroueil (eds.), *Characterisation and engineering properties of natural soils*; *Proceedings of the first International Workshop on Characterisation and Engineering Properties of Natural Soils*. Vol. 1 (IS-Singapore): 695-724. Lisse: Swets & Zeitlinger.

Goovaerts, P. 1997. *Geostatistics for natural resources evaluation.* New York: Oxford University Press.

Hegazy, Y.A. 1998. Delineating geostratigraphy by cluster analysis of piezocone data. *PhD Dissertation*, Civil & Environmental Engineering, Georgia Institute of Technology, Atlanta, GA: 464 p.

Hight, D.W. & Leroueil, S. 2003. Characterisation of soils for engineering practice. In T.S. Tan, K.K. Phoon, D.W. Hight & S. Leroueil (eds.), *Characterisation and engineering properties of natural soils*; *Proceedings of the first International Workshop on Characterisation and Engineering Properties of Natural Soils*, Vol. 1 (IS-Singapore): 255-360. Lisse: Swets & Zeitlinger.

Isaaks, E.H. & Srivastava, R.M. 1989. *An introduction to applied geostatistics.* New York: Oxford University Press.

Moss, R.E.S. 2008. Quantifying measurement uncertainty of thirty-meter shear wave velocity. *Bulletin of the Seismological Society of America* 98(3): 1399-1411.

Omre, H. 1987. Bayesian kriging – Merging observations and qualified guesses in kriging. *Mathematical Geology* 19(1): 25-39.

Omre, H. & Halvorsen, K.B. 1989. The Bayesian bridge between simple and universal kriging. *Mathematical Geology* 21(7): 767-786.

Thode, H.C. (Jr.) 2002. *Testing for normality*. New York: Marcel Dekker.

Uzielli, M., Lacasse, S., Nadim, F. & Phoon, K.K. 2007. Soil variability analysis for geotechnical practice. In T.S. Tan, K.K. Phoon, D.W. Hight & S. Leroueil (eds.), *Proceedings of the Second International Workshop on Characterisation and Engineering Properties of Natural Soils*. Vol. 3, (IS-Singapore), Taylor & Francis, London: 1653-1752.