Geotechnical Safety and Risk V
T. Schweckendiek et al. (Eds.)

© 2015 The authors and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License.
doi:10.3233/978-1-61499-580-7-159

Analyzing Spatial Variability of Geologic Profiles for Four Sites in Hong Kong

W. F. LIU, and Y. F. LEUNG
Department of Civil and Environmental Engineering,
The Hong Kong Polytechnic University, Hong Kong

Abstract. This paper investigates the spatial variability characteristics of geologic profiles, including variations in thickness of marine clay deposits and rockhead levels, based on borehole data obtained from four sites in Hong Kong. The numbers of boreholes are approximately 100 in two cases, while the other two cases comprise more than 300 boreholes each. The large volume of data allows comprehensive statistical analyses to identify the spatial correlation/variability in subsurface profiles using the Restricted Maximum Likelihood (REML) method. The Matérn Autocorrelation model is adopted for its flexible functional form, with the parameters optimized using the Differential Evolution algorithm, in order to maximize the log likelihood value in REML. This technique is used to evaluate the spatial variability characteristics of geologic profiles, including parameters such as the spatial dependence and scale of fluctuation at the four sites. The effects of irregular sampling pattern, sample domain scale and sampling density on these parameters are also discussed based on the analyses. In addition, the existence of faults in two of the sites is found to significantly affect the spatial variability of rockhead level, as indicated by the reduced scales of fluctuation and spatial dependence in areas intersected by faults.

Keywords. spatial variability, geologic profiles, irregular sampling pattern, geological faults

1. Introduction

The inherent spatial variability of soil properties and geologic profiles is one of the primary sources of geotechnical uncertainties that can exert significant impacts on designs and constructions in geotechnical engineering. Misunderstanding or ignoring spatial variability in the soil properties can lead to substantial risks in projects, as this may cause oversight of certain mechanisms in the behaviour of the geotechnical structures. To characterise the spatial variations of soil/rock properties, a number of researchers have discussed the use of the exponential, spherical or Gaussian autocorrelation models (e.g., Lark and Cullis, 2004; Minasny and McBratney, 2007; Pardo-Iguzquiza and Chica-Olmo, 2008; Santra et al. 2011). Some of them, however, have questioned the effectiveness of these models in providing accurate estimates of spatial variability features. Moreover, many previous studies have been conducted on regularly-spaced datasets, while studies on irregular sampling patterns are relatively limited.

In this paper, information of geologic profiles (thickness of marine clay deposits and level of moderately weathered granite) obtained from irregularly-spaced boreholes is analysed to reveal their spatial variability characteristics. The Matérn function (Matérn, 1960) is adopted to model the autocorrelation structures, owing to the flexibility of its functional form. Parameters of the Matérn function are optimized using a heuristic algorithm, known as the Differential Evolution, to maximize the log-likelihood value. This paper also discusses the impacts of sample domain size, sampling density and the existence of geologic features, such as faults, on the variability characteristics of geologic profiles.

2. Objectives of the Study

Borehole data obtained from four different sites in Hong Kong have been used for spatial variability studies. The main objectives of the current study are:

1. to characterise the spatial variability of the level of moderately decomposed granite/volcanic rocks and thickness of marine clay deposits at the sites;

2. to identify the effects of sample domain scale and sampling density on the spatial variability characteristics, under irregular sampling patterns; and
3. to identify the effects of geological faults on the spatial correlation of rockhead levels.

3. Site Description and Data Processing

Figure 1 shows the locations of the four study sites. The borehole records were obtained from geotechnical investigation reports of previous government projects in the areas, which were archived in the Civil Engineering Library of the Hong Kong Government. For the KLR site, the focus of the study is on the thickness of soft marine clay deposits, which posed engineering challenges to the design and construction of foundations and retaining walls. For the three other sites (HUH, CWE and CHE), the focus is on the level of moderately decomposed granite/volcanic rocks, referred to as Grade III material (GEO, 1988) and commonly taken as the ‘rockhead level’ in the local practice. In addition, several geological faults have been reported to intersect the sites of CWE and CHE.

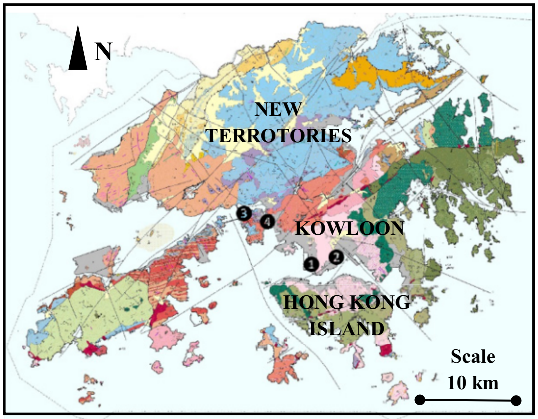

Figure 1. Site locations of the four cases: ❶ Kowloon Reclamation (KLR); ❷ Hung Hom Station (HUH); ❸ Cheng Wang Estate (CWE); ❹ Cheung Hong Estate (CHE).

Table 1 shows the sampling information for each of the four cases. To identify the effects of sample domain scale on the spatial correlation, two domain scales are specified for each site. The entire site is referred to as the “regional” scale area, where subsets of data are extracted to form the “local” scale area, so that similar sampling properties prevail in the two different domain scales. Also, for each case, different percentages of borehole data from the entire dataset were randomly extracted as a series of sub-datasets with lower sampling density, in order to identify the effects of sampling density on the spatial correlation characteristics of the geologic profiles. In addition, to understand the effects of faults, two local scale blocks were extracted from each of CWE and CHE, with geological fault crossing Block 1 but no faults reported in Block 2. Details of these analyses will be presented in later sections of the paper.

Table 1. Sample domain scale and sample size for each case

Case No.	Items	Sample Domain Scale (m×m)	Sample Size (No. of boreholes)
KLR	Regional	400×700	100
	Local	200×400	69
HUH	Regional	250×1400	88
	Local	120×800	52
CWE	Regional	600×600	312
	Local	400×300	207
CHE	Regional	1400×900	425
	Local	800×500	250

4. Autocorrelation Model by REML

The four cases involve approximately 100-450 boreholes with their locations irregularly spaced over the site. To ensure stationarity of the data, it is important to estimate and remove the trend in the data, such that the spatial correlation features are not masked by this deterministic component. In many previous studies, the trend component is determined by regression analysis using linear or polynomial functions (e.g., Campanella et al. 1987; Dasaka and Zhang, 2012), and the residuals are then analysed and presented using method of moments or semivariograms. The resulting autocorrelation function or semi-variograms are then fitted, typically with exponential, spherical or Gaussian functions.

In the current work, the Restricted Maximum Likelihood (REML) method (Stein, 1999; Lark and Cullis, 2004) is applied to simultaneously determine the optimal deterministic trend component and estimate the autocorrelation properties of residuals by maximizing the log-likelihood shown in Eq. (1):

$$L(\theta|Z) = -\frac{n-p}{2}log(2\pi) - \frac{1}{2}log|K| - \frac{1}{2}log|W| - \frac{1}{2}Z^T K^{-1} Q Z \quad (1)$$

where the observed values Z consist of deterministic trend component and residuals. K is the autocorrelation function matrix, M is the design matrix, then $W=M^T K^{-1} M$ and $Q=I-MW^{-1}M^T K^{-1}$.

An appropriate polynomial order (but not the coefficients) of the trend is first determined based on the variation of raw data. To reduce the computational demands, relatively low-order (≤3rd order) polynomials are specified for trend components. Meanwhile, the Matérn model, which has a flexible functional form, is adopted to simulate the autocorrelation structure, as well as the corresponding autocorrelation distance parameter 'r', the smoothness parameter 'ν' and spatial dependence 's', as shown in Eq. (2):

$$F(h) = \frac{1}{2^{\nu-1}\,\Gamma(\nu)} \left(\frac{h}{r}\right)^{\nu} K_{\nu}\left(\frac{h}{r}\right) \quad (2)$$

where h is the separation distance, K_ν is a modified Bessel function of the second kind of order ν, and Γ is the gamma function. The Matérn function represents a generalized function with its shape controlled by the smoothness parameter, ν. For example, it corresponds to the exponential function when $\nu=$ 0.5, and is equivalent to the Gaussian function when ν approaches infinity (Minasny and McBratney, 2007). In the current study, parameters of the Matérn function are optimised using the Differential Evolution (DE) (Storn and Price, 1997) algorithm, in order to maximize the log-likelihood value. The DE is conceptually similar to other evolutionary algorithms, which is not prone to converging at local maxima/minima, and is recently applied in other engineering problems (e.g. Leung et al. 2015).

5. Results and Discussions

5.1. Effects of Sample Domain Scale on Spatial Variability Characteristics

Previous studies (e.g. DeGroot and Baecher, 1993) described that under regular sampling patterns, the spatial variability characteristics of soil properties vary with the scale of the sampling site, with a larger sample domain scale associated with a larger scale of fluctuation. In the current study, the four cases are analysed to identify the effects of sample domain scale on spatial variability features of rockhead levels and thickness of marine clay deposits, under irregular sampling patterns. For each case study, the entire site was considered as the "regional" scale site, and a relatively small subcase was extracted as the "local" scale site. During the data extraction process, each local scale domain was ensured to possess similar sampling property (e.g., sampling density and COV of sample distance) to the corresponding regional scale.

The above-mentioned REML approach with the Matérn model was applied to estimate the spatial variability of each subcase. By comparing the results of regional scale and local scale of each case, the differences in the scale of fluctuation values 'δ' and the spatial dependence values 's' reflect the impacts of the sample domain scale on spatial variability characteristics. The results are summarized in Table 2, and the autocorrelation curves for all regional scale and local scale cases are shown in Figure 2.

Table 2. Spatial variability characteristics of four cases with regional and local domain scales

Case No. \ Items		Polynomial order of Trend	Scale of fluctuation (δ) (m)	Spatial dependence (s)
KLR	Regional	Linear	23.9	1.0
	Local	Linear	22.3	1.0
HUH	Regional	Quadratic	94.4	0.86
	Local	Quadratic	89.2	0.76
CWE	Regional	Quadratic	171.6	0.75
	Local	Quadratic	96.2	0.46
CHE	Regional	Quadratic	216.4	0.95
	Local	Quadratic	136.9	0.88

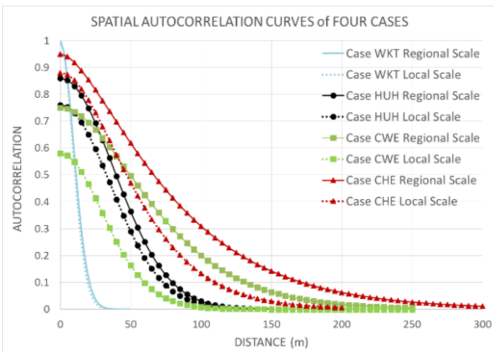

Figure 2. Autocorrelation curves of four cases with regional and local domain scales.

Figure 3 shows the influence of sample domain scale on the scale of fluctuation value δ, and spatial dependence value s. It is observed that values of δ and s are more similar in the two domain scales for KLR and HUH, but vary significantly for CHE and CWE. This may be attributed to the fact that CHE and CWE are crossed by geological faults. These local features may not be revealed fully in the 'regional' scale where the entire dataset is included in the analyses of spatial variability. The influence of faults will be discussed further in Section 5.3.

Based on the results of the four cases, the domain scale affects the spatial variability characteristics, and the impacts are substantial when local geological features are identified. It is therefore important that local scale effects are considered separately in these situations.

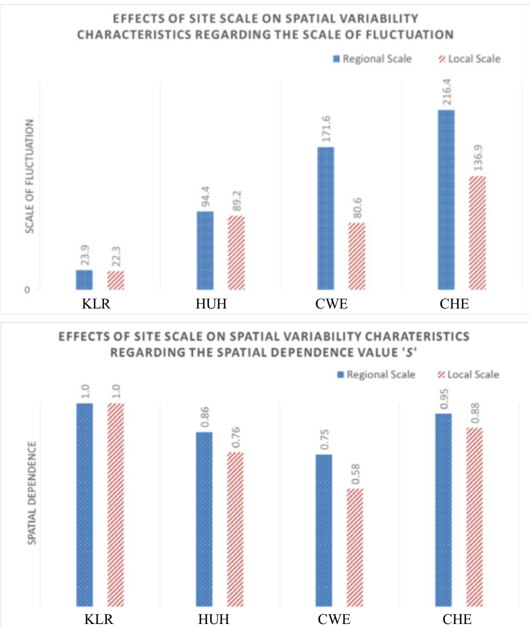

Figure 3. Scale of fluctuation δ, and spatial dependence value s, of four cases with regional and local scales.

5.2. Effects of Sampling Density on Spatial Variability Characteristics

Previous studies involving regular sampling patterns concluded that the scale of fluctuation generally increases with the individual sample spacing (Cafaro and Cherubini, 2002; Dasaka and Zhang, 2012). This section investigates the effects of sampling density on the spatial variability characteristics under irregular sampling patterns, which are common in geotechnical exploration programs.

Borehole data from two cases, namely KLR and HUH, are analysed for this purpose. For each of the two cases, subsets of different percentages, ranging from 40% to 90%, of the original dataset were extracted randomly to form sub-datasets, which have similar domain scales and therefore lower sampling densities compared to the entire dataset. For each subset percentage (40%-90%, with an interval of 10%), 420 repetitions are performed, resulting in a total of 2520 realizations for each of KLR and HUH. Since each realization consists of a random population of data extracted from the original dataset, the mean sampling densities are different even among realizations with the same subset percentage.

The REML analyses with DE optimisation of the Matérn function were conducted for the 2520 realizations of both cases, and the resulting estimates of scales of fluctuation δ are shown in two different forms in Figures 4 and 5.

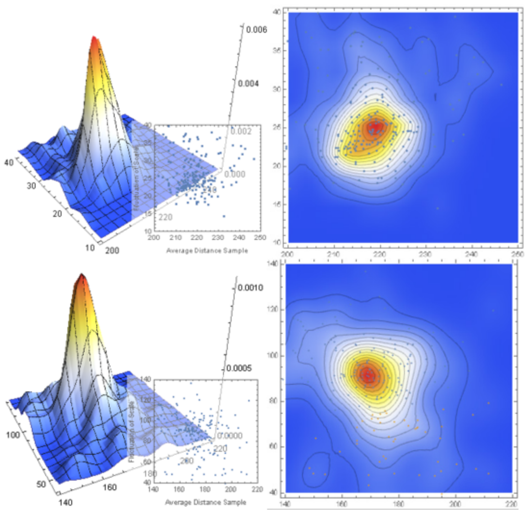

Figure 4. Probability density distributions of δ against average borehole spacing for KLR *(top)* and HUH *(bottom)*.

Figure 4 plots the values of δ against average borehole separation distances in the realizations, also represented in contours and 3-D plots of the probability density distributions (variations in s are not substantial). For both KLR and HUH, there is no obvious relationship between the estimated δ (of marine clay thickness and bedrock level, respectively) and average borehole spacing. Instead, a large proportion of the estimated δ is concentrated in a relatively narrow range.

As the realizations involve smaller percentages of the original dataset, a wider distribution of δ estimates can be expected. This is illustrated in Figure 5, which also shows the approximate boundaries of the 10th, 25th, 75th and 90th percentiles of δ estimates for each subset percentage of KLR, normalized by the δ value estimated using the complete dataset (δ_{100}). The results may be interpreted as the impacts of imperfect accessibility to the dataset on subsequent estimates of spatial variability characteristics. For example, if the analyst only had access to half (0.5) of all the borehole information, there would be roughly a 50% chance that the estimated δ value would fall between 0.9 and 1.35 times of δ_{100}, or 80% chance that the estimated δ value falls between 0.75 and 1.75 times of δ_{100}.

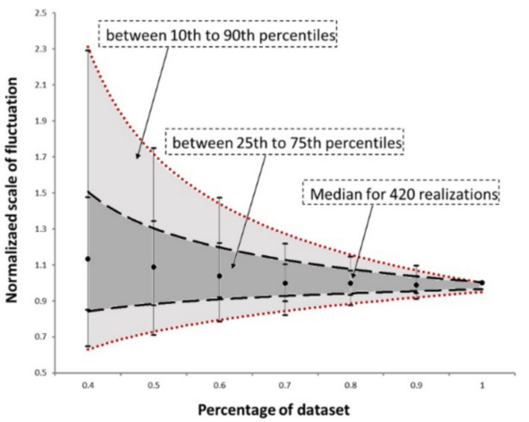

Figure 5. Estimates of δ, normalized by δ value of full dataset, with different subset percentages at KLR.

Figure 5 may be interpreted from a different, though less rigorous, perspective. Comparing the distribution of δ estimates for the KLR dataset, it may be postulated that if the number of boreholes are doubled (from 0.5 to 1.0), there is an approximately 50% chance that the revised δ estimate will differ within the range of +11% to -26%, or 80% chance that the revised value will differ by +33% to -43%. However, it should be noted that this is strictly correct only when the original dataset is sufficiently large. Further work is required to investigate the validity of this proposition.

5.3. Effects of Geological Faults on Spatial Variability Characteristics

Geological faults often form discontinuities which may have a significant influence on the mechanical behaviour, such as the strength and deformation of soil and rock masses. In order to enhance the understanding of the impacts of geological faults on spatial variability of Grade III rockhead levels, four blocks were extracted from the local scale sites of CWE and CHE. Each case contained two blocks with the same sample domain scale and a similar sampling density. The borehole location plots and location of the blocks for each case are illustrated in Figure 6.

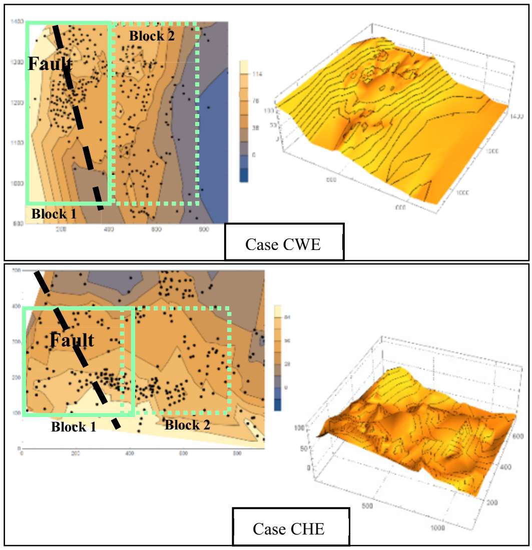

Figure 6. Extraction of blocks for study of effects of faults in CWE and CHE.

One NW-SE fault cuts through the northwest corner of the CWE site, while CHE is intersected by a fault in the western area. As shown in Figure 6, Block 1 of the two cases are designed to be intersected by the faults, while Block 2 is assumed to be free of influence by faults. Since the two blocks in each case have similar domain scales and sampling densities, by comparing the results from the two blocks of each case, the impacts of the faults can be illustrated.

Table 3 shows the comparisons of spatial variability characteristics of the two blocks for both cases. According to these results, faults have significant impacts on the spatial variability structure, and are associated with reductions in scale of fluctuations δ and spatial dependence value s, which imply higher levels of uncertainties in the rockhead levels. Together with discussions in Section 5.1, it is recommended that 'local' scale analyses be performed separately whenever local geological features are identified at the project site.

Table 3. Effects of geological faults on spatial variability characteristics

Case No. \ Items		Geology	Scale of Fluctuation (δ) (m)	Spatial Dependence (s)
CWE	Block 1	Fault	80.6	0.58
	Block 2	No Fault	147.4	0.69
CHE	Block 1	Fault	93.4	0.73
	Block 2	No Fault	122.5	0.89

6. Conclusions

The spatial variability characteristics of Grade III rockhead levels and thickness of marine clay deposits are studied using the borehole data collected from four sites in Hong Kong. Effects of the sample domain scale, sampling density and geological faults on spatial variability characteristics have been discussed based on a series of analyses. The following summarizes the current work and the conclusions:

- The REML approach is coupled with the Matérn model, with parameters optimized using the Differential Evolution algorithm. This leads to a robust technique to estimate the autocorrelation structure of the dataset.
- The spatial variability of geologic profiles is affected by the sampling domain scale. This is likely due to the fact that details of local geological features are not revealed when a large domain scale is considered. On the contrary, there is not an obvious relationship between the scale of fluctuation and sampling density.
- Geological faults can affect the spatial variability characteristics of bedrock levels. In particular, scale of fluctuation and spatial dependence reduce with existence of faults, which imply a higher uncertainty that should be taken into consideration in engineering practice.

7. Acknowledgements

The work presented in this paper is financially supported by the Research Grants Council of the Hong Kong Special Administrative Region (Project No. 25201214). Also, the authors would like to acknowledge the permission of the Civil Engineering and Development Department, the Government of Hong Kong Special Administrative Region, to present analyses of the data obtained from the Civil Engineering Library.

References

Cafaro, F., Cherubini, C. (2002). Large sample spacing in evaluation of vertical strength variability of clayey soil, *Journal of Geotechnical and Geoenvironmental Engineering* **128** (7), 558–68.

Campanella, R. G., Wickremesinghe, D. G., Robertson, P. K. (1987). Statistical treatment of cone penetrometer test data, *Proc. of the 5th Int. Conf. on Application of Statistics and Probability, AA Balkema, Rotterdam.*

Dasaka, S. M., Zhang, L. M. (2012). Spatial variability of in situ weathered soil. *Geotechnique* **62** (5), 375–84.

DeGroot, D. J., & Baecher, G. B. (1993). Estimating autocovariance of in-situ soil properties, *Journal of Geotechnical Engineering* **119**(1), 147-166.

GEO (1988). *GEOGUIDE 3: Guide to Rock and Soil Descriptions*. Geotechnical Engineering Office, Civil Engineering Department, Government of Hong Kong.

Lark, R. M., Cullis, B. R. (2004). Model-based analysis using REML for inference from systematically sampled data on soil, *European Journal of Soil Science* **55**, 799-813.

Leung, Y.F., Klar, A., Soga, K., Hoult, N.A. (2015). Multi-objective foundation optimization with consideration of superstructure stiffness and pile reuse. Under review.

Matérn, B. (1960). Stochastic models and their application to some problems in forest surveys and other sampling investigations. *Meddelanden från Statens Skogsforskningsinstitut*, **49** (5), 1-144.

Minasny, B., McBratney. A. B. (2005). The Matérn function as a general model for soil variograms, *Geoderma* **128** (3-4), 192–207.

Minasny, B., McBratney, A. B. (2007). Spatial prediction of soil properties using EBLUP with the Matérn covariance function, *Geoderma* **140**(4), 324-336.

Pardo-Iguzquiza, E., Chica-Olmo, M. (2008). Geostatistics with the Matern semivariogram model: A library of computer programs for inference, kriging and simulation, *Computers & Geosciences* **34**, 1073-1079.

Santra, P., Das, B. S., Chakravarty, D. (2012). Spatial prediction of soil properties in a watershed scale through maximum likelihood approach, *Environmental Earth Sciences* **65** (7), 2051-2061.

Stein, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*, New York: Springer-Verlag.

Storn, R. and Price, K. (1997). Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization* **11** (4), 341-359.