Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly

doi: 10.3850/978-981-18-5182-7_03-010-cd

ISGSR 2022

Characterization of Vertical Spatial Variability of Soils Using CPTu Data Exploration

Izabela Couto Campello[1], Maria Das Graças Gardoni[2], Karla Cristina Araujo Pimentel[3], and Andre Assis[4]

[1]Klohn Crippen Berger, Brisbane, Australia.
E-mail: campelloizabela@gmail.com
[2]Department of Geotechnical Engineering, Federal University of Minas Gerais, Belo Horizonte, Brazil.
E-mail: gardoni@etg.ufmg.br
[3]VALE, Belo Horizonte, Brazil.
E-mail: kpimentel.maia@gmail.com
[4]Department of Civil and Environmental Engineering, University of Brasilia, Brasilia, Brazil.
E-mail: aassis.p@gmail.com

Abstract: Geotechnical uncertainties can be categorized into epistemic and aleatory uncertainties. Epistemic uncertainty is caused by the lack of knowledge or data. Aleatory uncertainty refers to the intrinsic randomness of a phenomenon. In geotechnical engineering, the aleatory uncertainty can be characterized by inherent spatial variability, which is the focus of this paper. The state-of-the-practice is to characterize this spatial variability using the scale of fluctuation or correlation length from a statistical interpretation of limited samplings. These descriptors characterize the distance over which the parameters of a soil or rock are similar or correlated. This paper presents a study case of a dam located in the northern region of Brazil. The undrained shear strength of the foundation soil was spatially modelled using CPTu profiles. The data were decomposed between trend and residuals and an autocorrelation function (ACF) was defined. The results were then compared against different theoretical autocorrelation models (ACM), and the vertical correlation length was estimated from the fitted ACM.

Keywords: Geotechnical site characterization; reliability; uncertainty; spatial variability; probabilistic models.

1 Uncertainty in Geotechnical Engineering

Geotechnical site characterization aims to delineate underground stratigraphy and estimate soil properties for geotechnical analysis and/or designs (Cao *et al*., 2017). This is a process that involves site reconnaissance, *in situ* investigation, laboratory testing, interpretation of observation data, and inferring soil properties and underground strata. Thus, uncertainties regarding the performance of a geotechnical structure, whether due to limited site investigation, natural soil variability, or limitations in calculation models, are inevitable.

From a geotechnical engineering perspective, uncertainties can be categorized into aleatory and epistemic. Aleatory uncertainty refers to the intrinsic randomness of a phenomenon and can be translated by inherent spatial variability. Epistemic uncertainty is caused by the lack of knowledge or data, such as measurement errors, limited data availability and uncertainties associated with transformation models.

According to Cao *et al*. (2017), soils and their properties are affected by various factors during their formation processes, such as weathering and erosion processes, transportation agents, and sedimentation conditions. Properties of geotechnical materials, therefore, vary spatially, which is usually known as "inherent spatial variability", as mentioned by Uzielli *et al*. (2006) and Cao *et al*. (2017), or simply "natural variability", as mentioned by Phoon & Retief (2016).

Both spatial variability and epistemic uncertainties can be incorporated into the estimate of soil properties. However, Cao *et al*. (2017) emphasize that epistemic uncertainties that arise from insufficient knowledge do not contribute to the actual response of geotechnical structures. In contrast, the spatial variability of soils significantly affects the response of geotechnical structures.

The spatial variability can be conveniently modeled by means of statistical and probabilistic tools. The most common techniques used to characterize soil variability comprise at least: (i) classical descriptive and inferential statistical analysis (e.g., mean, coefficient of variation (COV), and probability distribution function) and (ii) spatial correlation structure. The coefficient of variation is a standardized measure of dispersion, defined as the ratio of the standard deviation to the mean. The COV value can be defined according to the uncertainty type (i.e., statistical uncertainty and measurement error) or specified as 'total variability' (neglecting the source of uncertainties). It is extremely difficult in practice to evaluate the various sources of uncertainty separately. The spatial correlation structure can be specified from the correlation length or scale of fluctuation, θ, which corresponds to a measure of the distance within which soil properties of different points are significantly correlated. Points separated by a larger distance than θ will show little correlation.

Within this context, several papers have been published in recent years covering approaches for spatial variability modeling and its effects on slope stability (Lloret-Cabot *et al*., 2014; Liu *et al*., 2017; Oguz *et al*., 2018; Krogt *et al*., 2018; Jiang *et al*., 2018; Chakraborty & Dey, 2019; Ching & Phoon, 2019; Qi & Liu, 2019). Despite its relevance, the application of probabilistic analysis accounting for spatial variability in dam engineering and geotechnical practice is still limited.

This paper describes one of the existing methods for modeling spatial variability and presents a study case of a dam for which the correlation length was calculated.

2 Modeling Spatial Variability

According to Uzielli *et al*. (2006), second-moment statistics (i.e., mean and standard deviation) alone are unable to describe the spatial variation of soil properties, whether measured in laboratory or in-situ. Two sets of measurements may have similar second-moment statistics and statistical distributions but could display substantial differences in spatial distribution.

Among the existing techniques, decomposition is the most widely used in geotechnical engineering, as highlighted by Uzielli *et al*. (2006). Decomposition consists of separating spatial variability into two parts: a trend function $t(z)$ and a set of residuals about the trend $w(z)$, as shown in Eq. (1). The separation of spatial variation between trend and residuals is a technique for modeling soil variability.

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

in which:
$\psi(z)$ is the soil property at a location z;
$t(z)$ is a trend function of z, characterized deterministically by an equation;
$w(z)$ is a fluctuating component, or residual variation of z, statistically characterized by a random variable, usually with zero mean and non-zero variance.

A trend is estimated from the adjustment of mathematical functions, such as regression analysis. Although the trend is a deterministic function, statistical uncertainty is also accounted for due to the limited size of the data set.

According to El-Ramly (2001), the higher the number of the trend function parameters, the higher the uncertainty in the estimation of these parameters. Baecher & Christian (2003) stated that the trend function should be kept as simple as possible, without injustice to the data set and without ignoring the geological history.

Residuals represent the part of spatial variability that cannot be explained by a relatively simple function of the reference spatial coordinate, as mentioned by Uzielli *et al*. (2006). They are usually a zero-mean set that, when plotted against the spatial coordinate, fluctuates around the mean value.

If two points of a soil deposit, i and j, are close to each other, the residuals w_i and w_j are strongly correlated. It is assumed, therefore, that the association between the residuals increases as the separation distance (or lag) between the points decreases. The autocorrelation function (ACF) is, in this context, a mathematical tool to describe the variation of the spatial correlation as a function of the spatial separation distance between two points.

As described by Baecher & Christian (2008), the autocorrelation is typically assumed to be the same everywhere within a deposit, which corresponds to the assumption of stationarity. In the geotechnical literature, stationarity is sometimes referred to as statistical homogeneity. If the autocovariance function depends only on the absolute separation distance and not on direction, the random field is said to be isotropic.

The state-of-the-practice is to characterize this spatial variability using the scale of fluctuation, which describes the distance over which the parameters of a soil or rock are similar or correlated (Cami *et al*., 2020). The physical meaning of the scale of fluctuation, or correlation length, is illustrated in Figure 1. For a particular material, the friction angle is a normally distributed variable, with a mean of 30° and minimum and maximum values of 27° and 33°, respectively. In practice, it is unlikely that at a given point, the material has a low friction angle (27°) and that at an adjacent point the value will jump to 33°. It is more realistic, therefore, to predict that adjacent points have more similar parameters, which leads to the concept of the correlation length.

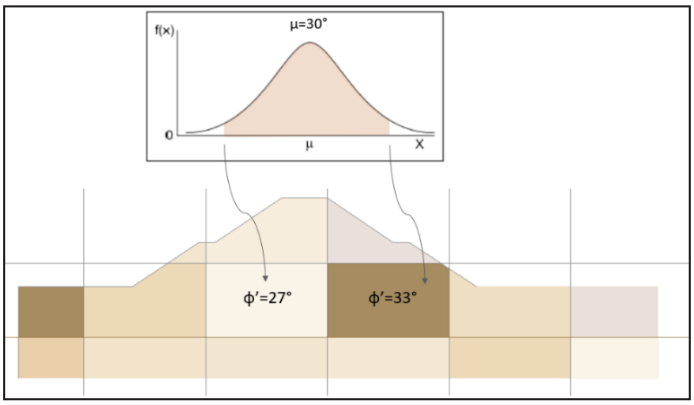

Figure 1. Random field and correlation length.

3 Study Case

The study case corresponds to a dam located in the northern region of Brazil, where the geology is characterized by the occurrence of bauxite deposits and sedimentary rocks, modified over time by weathering processes. A site investigation program was undertaken on the foundation of a tailings dam, including several laboratory and field tests (vane shear and CPTu tests). This paper is focused on the CPTu probing for characterizing the vertical spatial variability. A total of 10 boreholes pushed through the residual soils (named the clayey foundation) were considered in this study. The interpreted undrained shear strength (s_u) is plotted *versus* depth in Figure 2.

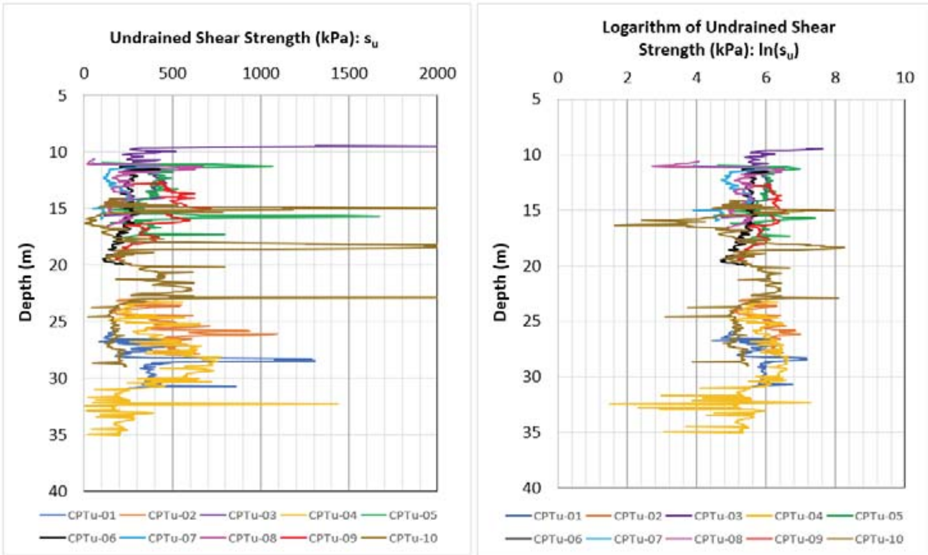

Figure 2. Clayey foundation undrained shear strength.

The piezocone measurements were recorded every 5 cm, but as the top of the clayey foundation layer is not uniform, the readings correspond to different depths. Therefore, depths ending in 30.11 m to 30.15 m, for example, were grouped as 30.15 m, and depths ending in 30.16 m to 30.20 m were grouped as 30.20 m. Finally, as the results present a wide scale, ranging from 4 kPa to 3949 kPa, the logarithm of the undrained shear strength, $\ln(s_u)$, was considered for better visual representation.

The data was initially decomposed into a trend function and a fluctuating component (residual variation). Figure 3 shows the logarithm of s_u at each depth. Three fitting curves were used to represent the data: linear, polynomial order 2, and polynomial order 3. The functions and the respective R^2 are described in Eq. (2).

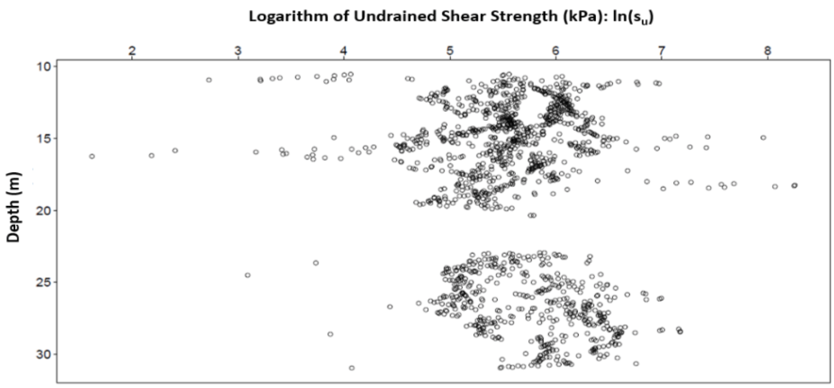

Figure 3. Variation of the mean value of undrained shear strength with depth.

$$Su_{avg}(z) = 5.131 + 0.025z \text{ (linear) } R^2 = 0.19$$

$$Su_{avg}(z) = 6.343 - 0.104z + 0.003z^2 \text{ (order 2) } R^2 = 0.26 \quad (2)$$

$$Su_{avg}(z) = 3.864 + 0.297z - 0.017z^2 + 0.0003z^3 \text{ (order 3) } R^2 = 0.29$$

in which:
z is depth and $Su_{avg}(z)$ is the mean of $\ln(s_u)$ at depth z.

The low R^2 values suggest a weak correlation between the undrained shear strength and depth. As there is a significant increase in R^2 from the linear approximation to the polynomial order 2, but not from the polynomial order 2 to the polynomial order 3, the polynomial order 2 was adopted as representative. The residuals for each observation of each borehole were then standardized using Eq. (3).

$$\overline{w}(z) = \frac{Su\ (z)\ -\ Su_{avg}(z)}{\sigma(z)} \tag{3}$$

The methodology to calculate the correlation length consists of determining the sample autocorrelation function (ACF) and, subsequently, adjusting a theoretical autocorrelation model (ACM) to the defined ACF. Hence, the correlation length can be estimated from the fitted ACM parameter. According to Chakraborty & Dey (2019), by minimizing the mean square error between the estimated ACF and the ACM, the parameter a of correlation structure is estimated from the standardized error.

The autocorrelation function (ACF) of the standardized residuals, for each of the 10 boreholes, is shown in Figure 4**Figure 4.** Sample autocorrelation function (ACF) based on 10 boreholes.

. As aforementioned, the ACF is a measure of the correlation between two observations in space (or time) separated by units of space (or time); it is calculated by dividing the covariance of measurements made at points distant from each other by the variance of the property. Lag is the number of space units that separate data from spatial series. When the distance is zero, the maximum value is 1, as it represents the variable correlated with itself.

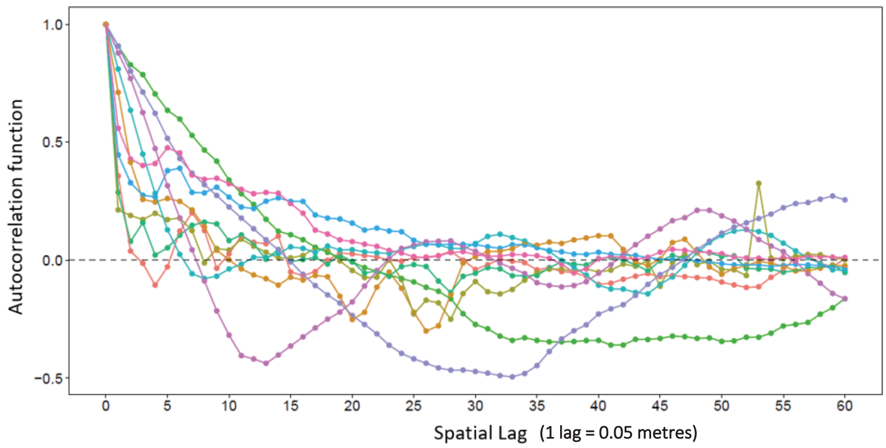

Figure 4. Sample autocorrelation function (ACF) based on 10 boreholes.

The lag represents the distance, in 0.05 meters, between two measurements from a borehole (i.e., lag = 3 indicates that two observations are vertically 0.15 meters apart). A total of 60 lags (3 meters) were considered in the analysis since beyond this lag the average autocorrelation of the measurements is approximately zero.

The calculated sample autocorrelation function (ACF) was adjusted to three theoretical models of autocorrelation (ACM), namely: Exponential, Exponential squared, and Markov of second order. The fitted curves are shown in Figure 5.

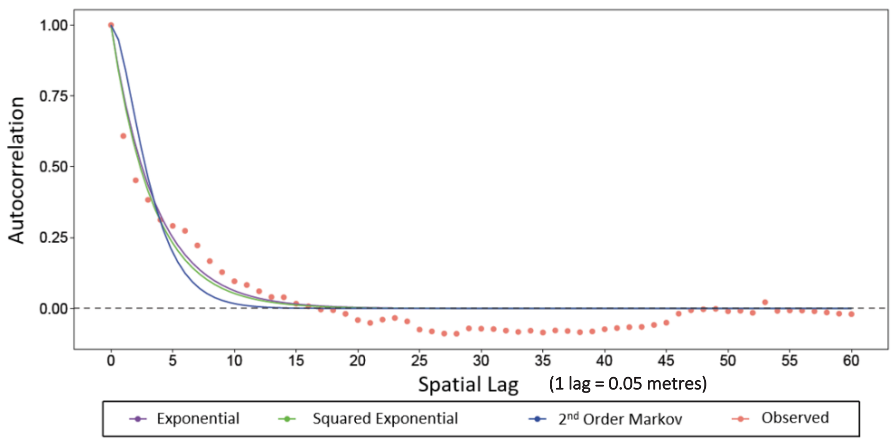

Figure 5. Fitted autocorrelation model (ACM).

The errors between the estimated and observed values were then calculated by the root-mean-square error (RMSE), which is the standard deviation of the residuals (prediction errors). By minimizing the RMSE between the estimated ACF and the ACM, the correlation was estimated by means of the parameter a.

Table 1 presents the results for the three theoretical curves, showing the estimated parameter a, the correlation length in meters, the RMSE and the R^2.

Table 1. Correlation length based on the fitted theoretical ACM.

ACM	Correlation function	a	Correlation Length (m) [1]		RMSE	R^2
Exponential	$y = e^{-ax}$	0.277	$\Theta = 0.05 \times 2/a$	0.36	0.028	0.886
Squared Exponential	$y = e^{-(x/a)^2}$	3.404	$\Theta = 0.05 \times a\sqrt{\pi}$	0.30	0.033	0.866
Second Order Markov	$y = e^{-ax}(1 + ax)$	0.602	$\Theta = 0.05 \times 4/a$	0.33	0.030	0.876

Note: The correlation length is defined as $2/a$, $a\sqrt{\pi}$ and $4/a$, respectively, for the three ACM functions. The factor 0.05 was inserted to convert the scale to meters, since the observations were made every 0.05 m.

In all cases, the vertical correlation length was between 0.3 m and 0.4 m. Since the exponential model presented a higher R^2 and a lower RMSE, the standardized error is recommended to be modeled as an exponential random field. The estimated correlation length is applicable to foundation soils with geology and stratigraphy profile similar to where the dam is located.

According to the literature values presented by Uzielli *et al.* (2006), vertical correlation lengths (scale of fluctuation) for undrained shear strength (from laboratory testing and Vane Shear Test) vary between 0.8 m to 8.6 m, while q_c (from CPT) range from 0.1 m to 3.0 m.

Once the correlation structure is derived, the results can be used along with other statistical parameters (e.g., mean, standard deviation, and probability distribution function) to carry out probabilistic stability analyzes and estimate the annual probability of failure.

4 Conclusions

In recent years, several papers have been published highlighting the importance of considering uncertainties, especially spatial variability, in the probabilistic slope stability analyses. The state-of-practice of spatial variability characterization considers the scale of fluctuation descriptor.

The methodology used for spatial correlation modeling was the decomposition technique, which consists of separating spatial variability into a trend function and a set of residuals about the trend. The residuals increase as the separation distance (or lag) between two points decreases. The autocorrelation function (ACF) is a mathematical tool to describe the variation of the spatial correlation as a function of the spatial separation distance between two points. After determining the sample autocorrelation function (ACF), a theoretical autocorrelation model (ACM) is fitted to the ACF. Hence, the correlation length can be estimated from the fitted ACM parameter.

This paper used 10 CPTu boreholes for characterizing the vertical spatial variability of the undrained shear strength. The calculated sample autocorrelation function (ACF) was adjusted to three theoretical models of autocorrelation (ACM), namely: Exponential, Exponential squared, and Markov of second order. In all cases, the vertical correlation length was between 0.3 m and 0.4 m. The spatial correlation structure is an essential parameter used to characterize soil variability.

References

Baecher, G.B. and Christian, J.T. (2003). Reliability and Statistics in Geotechnical Engineering, *Wiley, Chichester, U.K.*

Baecher, G.B. and Christian, J.T. (2008). Reliability-Based Design in Geotechnical Engineering Computations and Applications, Chapter 2, *Taylor & Francis, edited by Kok-Kwang Phoon, Oxon, U.K.*

Cami. B. Javankhoshdel. S. Phoon. K-K. Ching. J. (2020). Scale of fluctuation for spatially varying soils: estimation methods and values. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems Part A Civil Engineering*. v. 6. issue 4. DOI: 10.1061/AJRUA6.0001083

Cao. Z.; Li. D.; Wang. Y. (2017). Probabilistic Approaches for Geotechnical Site Characterization and Slope Stability Analysis. *1 ed. Heidelberg: SPRINGER*. 190 p.

Chakraborty. R.; Dey. A. (2019). Stochastic Modeling of the Spatial Variability of Soil. *Sustainable Civil Infrastructures.* p. 144–155. DOI: 10.1007/978-3-030-01926-6_11

Ching. J.; Phoon. K. K. (2019). Impact of Autocorrelation Function Model on the Probability of Failure. *Journal of Engineering Mechanics.* v. 145. n. 1. DOI: 10.1061/(ASCE) EM.1943-7889.0001549

El-Ramly. H. (2001). Probabilistic Analyses of Landslide Hazards and Risks: Bridging Theory and Practice. 420 p. *PhD Thesis Civil and Environmental Department. University of Alberta. Edmonton*.

Jiang. S-H.; Huang. J.; Huang. F.; Yang. J.; Yao. C.; Zhou. C-B. (2018). Modelling of spatial variability of soil undrained shear strength by conditional random fields for slope reliability analysis. *Applied Mathematical Modeling.* v. 63. p. 374-389. DOI: 10.1016/j.apm.2018.06.030

Krogt. M. G. Van Der.; Schweckendiek. T.; Kok. M. (2018). Uncertainty in spatial average undrained shear strength with a site-specific transformation model. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards.* DOI: 10.1080/17499518.2018.1554820

Liu. W. F.; Leung. Y. F.; Lo. M. K. (2017). Integrated framework for characterization of spatial variability of geological profiles. *Canadian Geotechnical Journal.* v. 54. p. 47-58. DOI: 10.1139/cgj-2016-0189

Oguz. E. A.; Huvaj. N.; Griffiths. D. V. (2018). Vertical spatial correlation length based on standard penetration tests. *Marine Georesources & Geotechnology.* v. 37. p. 45-56. DOI: 10.1080/1064119X.2018.1443180

Phoon. K. K.; Retief. J. V. (2016). Reliability of Geotechnical Structures in ISO2394. *London: CRC PRESS - Taylor & Francis Group.* 249 p.

Uzielli. M.; Lacasse. S.; Nadim F.; Phoon. K. K. (2006). Soil Variability Analysis for Geotechnical Practice. In: 2nd *International Workshop on Characterization and Engineering Properties of Natural Soils. Singapore*.

M. Lloret-Cabot. G.A. Fenton & M.A. Hicks (2014). On the estimation of the scale of fluctuation in geostatistics. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*. 8:2. 129-140. DOI: 10.1080/17499518.2013.871189.