Proceedings of the XVIII ECSMGE 2024
GEOTECHNICAL ENGINEERING CHALLENGES
TO MEET CURRENT AND EMERGING NEEDS OF SOCIETY

ISBN 978-1-032-54816-6

ECSMGE 24
XVIII EUROPEAN CONFERENCE ON SOIL MECHANICS AND GEOTECHNICAL ENGINEERING

Relevance of correlation length in geotechnical engineering

Pertinence de la longueur de corrélation en ingénierie géotechnique

M. Lloret-Cabot*, K. Zhang, W. Zhang, D. Toll
Durham University, Durham, United Kingdom
**marti.lloret-cabot@durham.ac.uk (corresponding author)*

ABSTRACT: Soils are heterogeneous in nature and as such, their properties inherently exhibit a distinct spatial variation that reveals past information of their geological history. This spatial variability is an important source of geotechnical uncertainty and hence its proper characterisation is of importance to geotechnical design. Random field theory provides a consistent mathematical framework to account for this soil's heterogeneity and it is a powerful computational tool to perform reliability analyses of geotechnical structures when combined with the finite element method. A key parameter in these stochastic analyses is the correlation length (or sometimes also referred to as scale of fluctuation) because it controls variations of a soil property in a given spatial direction. This spatial statistic can be estimated from in-situ data (e.g., cone penetration test CPT) and the quality of such estimation is the main question tackled in this paper.

RÉSUMÉ: Les sols sont de nature hétérogène et, en tant que tels, leurs propriétés présentent intrinsèquement une variation spatiale distincte qui révèle des informations passées sur leur histoire géologique. Cette variabilité spatiale est une source importante d'incertitude géotechnique et sa caractérisation appropriée est donc importante pour la conception géotechnique. La théorie des champs aléatoires fournit un cadre mathématique cohérent pour tenir compte de l'hétérogénéité de ce sol et constitue un outil informatique puissant pour effectuer des analyses de fiabilité des structures géotechniques lorsqu'il est combiné avec la méthode des éléments finis. Un paramètre clé dans ces analyses stochastiques est la longueur de corrélation (ou parfois également appelée échelle de fluctuation) car elle contrôle les variations d'une propriété du sol dans une direction spatiale donnée. Cette statistique spatiale peut être estimée à partir de données in situ (par exemple, test de pénétration du cône CPT) et la qualité d'une telle estimation est la principale question abordée dans cet article.

Keywords: correlation length, scale of fluctuation, uncertainty, soil variability

1 INTRODUCTION

Due to the geological/engineering processes by which soil is formed/deposited, its property values vary with position, meaning that soil is not homogenous. More interestingly, if a given direction within a soil domain is considered, higher correlation values are likely to be found in the geotechnical property values (e.g., undrained shear strength, S_u) measured at points that are closer to each other (small lag distance τ) because the soil conditions at these closer points are likely to be quite similar. In contrast, values of such property measured at points that are further away from each other (larger lags τ) will likely show bigger differences, because of a lower correlation. This decay of correlation in the property values at pairs of points along a given direction is described by the correlation function $\rho(\tau)$ and the sharpness of such decay is controlled by a key statistical parameter known as the correlation length, θ, (or also referred to as scale of fluctuation, SOF).

Such representation of the spatial variation of a geotechnical property in a given direction is useful because it facilitates a consistent approach able to capture the spatial heterogeneous nature of a soil property. It is also useful because it allows to incorporate in a relatively easily way the spatial variability of a soil property in geotechnical design via the random finite element method (RFEM) (e.g., Fenton and Griffiths, 2003) which combines the finite element method with the random field generator LAS (Local Average Subdivision method proposed by Fenton & Vanmarcke, 1990).

The RFEM has been used extensively in the literature to assess the structural response of a geotechnical system (Fenton and Griffiths, 2003; Hicks and Spencer, 2010; Griffiths et al., 2011). Interestingly, using the RFEM in any of these geotechnical problems facilitates a probabilistic assessment of their structural performance because, rather than solving the problem for a single realisation (considering an homogeneous soil), the same problem can be solved for a finite number of realisations within a Monte Carlo approach, so that the response of the geotechnical system can be assessed in probabilistic terms, allowing for the description of failure in terms of a probability of failure (and not in terms of a single factor of safety, as historically done). In doing such probabilistic assessment, all these works above confirm the relevance of incorporating the spatial variation of soil properties in the geotechnical design of these structures and, more specifically, the relevance of the correlation length. Perhaps due to this importance, many studies in the literature have studied the concept of the correlation length (e.g., DeGroot and Baecher, 1993; Phoon and Kulhawy, 1999; Fenton, 1999ab; Jaksa, 1995; Uzielli et al., 2005; Cami et al., 2020; Lloret-Cabot et al., 2013, Ching et al. 2023) estimated from available in situ or artificially generated information (most commonly cone penetration test data, CPT). These studies, however, do not all use the same estimation method (not even the same theoretical correlation model), they often treat the trend in the data in different ways and they do not all estimate θ using the same specific in situ test or, if they do, the sampling interval is not necessarily the same. These differences in the available literature make challenging the comparison between the different estimated θs. Against this background, the current paper aims to provide some further insight (in terms of accuracy) of the computational performance of the most common estimation method of θ which consists in minimising the error between an assumed theoretical correlation model $\rho(\tau)$ (see Table 1) and the estimated experimental correlation function $\hat{\rho}(\tau)$ (calculated from the in-situ or artificially generated data). Various theoretical correlation models $\rho(\tau)$ are available in the literature but due to space constraints only the most common two are considered in this research. Their definition is given in Table 1.

Table 1. Most common theoretical correlation models.

Name	Expression
Markovian	$\rho(\tau) = \exp\left\{\frac{-2\lvert\tau\rvert}{\theta}\right\}$
Gaussian	$\rho(\tau) = \exp\left\{-\pi\left(\frac{\lvert\tau\rvert}{\theta}\right)^2\right\}$

Note: θ is the correlation length in a given direction and the τ lag distance

According to the method of moments, the experimental correlation function takes the following form:

$$\hat{\rho}(\tau_j) = \sum_{i=1}^{k-j} \frac{(X_i-\hat{\mu})(X_{i+j}-\hat{\mu})}{(X_i-\hat{\mu})^2}, \quad for\ j = 0, \ldots k-1 \qquad (1)$$

where $\hat{\rho}(\tau_j)$ is the experimental correlation function between two points separated by a lag distance τ_j, $\hat{\mu}$ is the experimental mean, k is the total number of measurement points and i is the total number of pairs of points (separated by τ_j).

Equation 1 can be re-written in terms of the experimental variance $\hat{\sigma}^2$ in the following two forms depending on wether an unbiased (Equation 2) or biased (Equation 3) estimator of the experimental variance $\hat{\sigma}^2$ is considered:

$$\hat{\rho}(\tau_j) = \frac{1}{\hat{\sigma}^2(k-j)} \sum_{i=1}^{k-j} (X_i - \hat{\mu})(X_{i+j} - \hat{\mu}), \quad for\ j = 0, \ldots k-1 \qquad (2)$$

$$\hat{\rho}(\tau_j) = \frac{1}{\hat{\sigma}^2 k} \sum_{i=1}^{k-j} (X_i - \hat{\mu})(X_{i+j} - \hat{\mu}), \quad for\ j = 0, \ldots k-1 \qquad (3)$$

Traditionally, Equation 2 has been most commonly used in the geotechnical literature. However, recent work by Cami et al. (2020) proposes using Equation 3 instead to prevent negative eigenvalues in the correlation matrices. In the context of the work presented here (primarily focussed on the estimation of the correlation length in the vertical direction from artificially generated CPT data), any of the three equations is likely to give very similar estimates because a CPT typically has a large number of data points available in the vertical direction (and the value of the experimental mean should be close to zero, when using de-trended and normalised data, Lloret-Cabot et al., 2014). This research has used Equation 3.

2 ACCURACY ASSESSMENT

The main goal of this investigation is to assess the accuracy of using two of the most common theoretical correlation models to estimate the vertical correlation length from CPT data (see Table 1). This section aims to address this issue by using an equivalent numerical strategy to that proposed in Lloret-Cabot et al. (2014). Rather a two-dimensional (2-D) random fields as used in Lloret-Cabot et al. (2014), this work uses LAS to generate 1-D random fields as a suitable representation of de-trended and normalised CPT tip resistances with known zero mean (μ = 0) and unit variance ($\sigma^2 = 1$), and with a vertical correlation length of $\theta_v = 1.5$m (e.g, Fenton, 1999b). The total length considered is D = 20m and the sampling distance d_y is 0.1m. A typical example is illustrated in Figure 1.

From this random field data, the experimental correlation function $\hat{\rho}(\tau)$ can be estimated using Equation 3 as illustrated in Figure 2. The optimal value $\widehat{\theta_v}$ can be then estimated by best fitting $\hat{\rho}(\tau)$ to a theoretical correlation structure $\rho(\tau)$. A number n of estimated $\widehat{\theta_v}$ are obtained by repeating this process from j = 1 to n, (where n = 1,000 corresponds to the number of realisations considered here). In order to assess the accuracy of the correlation model used, the obtained value of $\widehat{\theta_v}$ (averaged over the n estimated correlation lengths) is compared against the known value of θ_v used to generate each CPT.

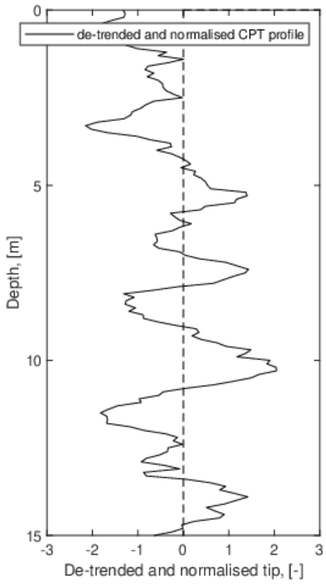

Figure 1. Typical CPT profile generated using LAS.

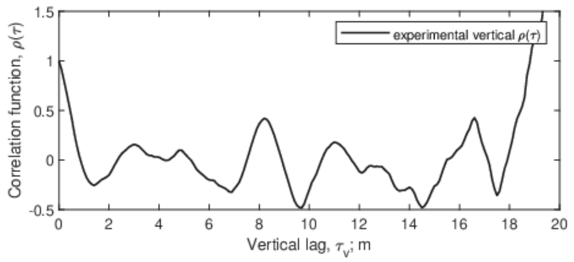

Figure 2. Typical experimental correlation structure.

The accuracy of estimating θ_v when using the approach discussed earlier with a Markovian or a Gaussian correlation structure is illustrated in Figures 3 and 4 respectvely. Each figure includes results for n = 1,000 realisations. In the figures, the representation of $\rho(\tau)$ with the corresponding optimized θ_v is represented as a thick solid line whereas the averaged experimental correlation model is indicated by a dashed thick line. The thinner continuous lines represent each of the individual $\hat{\rho}(\tau)$ estimated for each CPT. It is perhaps worth noting here that practically the same result for the optimal $\widehat{\theta_v}$ is obtained by best fitting the averaged experimental correlation structure and the theoretical one. Inspection of the figures shows that a very similar performance is achieved by the two correlation structures assumed, with an optimal $\widehat{\theta_v}$ of 1.55m and 1.54m for Markovian and Gaussian, respectively, corresponding to relative errors of 3.33% and 2.67%, which are values in line with similar analysis published in the literature (e.g., Lloret-Cabot et al., 2014). This similarity suggests that both theoretical correlation functions are equally valid for tackling this problem.

In addition to investigating the convergence of the estimation method to the true value of the vertical correlation length (Figures 3 and 4), it is also convenient to study the response of each individual estimated value of θ_v, as illustrated in Figure 5 in the form of a histogram.

Figure 5 includes also the true and estimated values of the vertical correlation length which are indicated, respectively, by a solid and a dashed line. For completeness the 90% confidence interval of LogNormal probability distribution fit is also plotted in the figures showing a reasonably good representation of the estimated θ_v. Inspection of the figure confirms that both approaches provide very similar results and, in particular, both show a right-skewed distribution which, interestingly, resembles well the histograms obtained from estimations of θ_v using real CPT data (e.g., Cami et al. 2020). Finally, it is worth noting that about two thirds of the estimated values of θ_v below the true value (1.5m) in agreement with the results from Nie et al. (2015) which illustrates well the challenge in estimating this parameter (as, clearly, the number of CPTs available in a specific site will be much less and hence the magnitude of the error might become quite important).

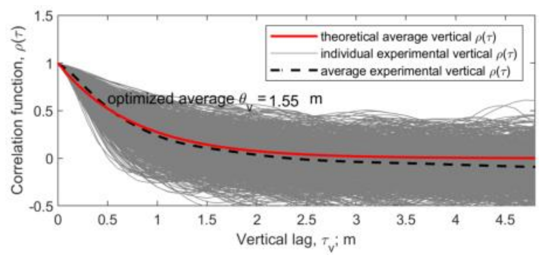

Figure 3. Estimation of the vertical correlation length with Markovian correlation structure and n = 1,000 realisations.

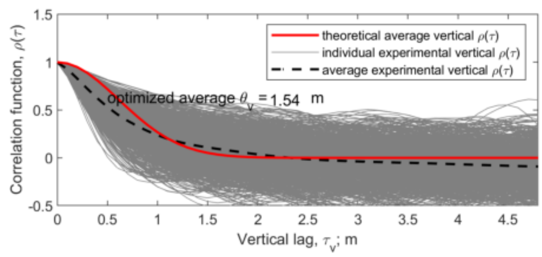

Figure 4. Estimation of the vertical correlation length with Gaussian correlation structure and n = 1,000 realisations.

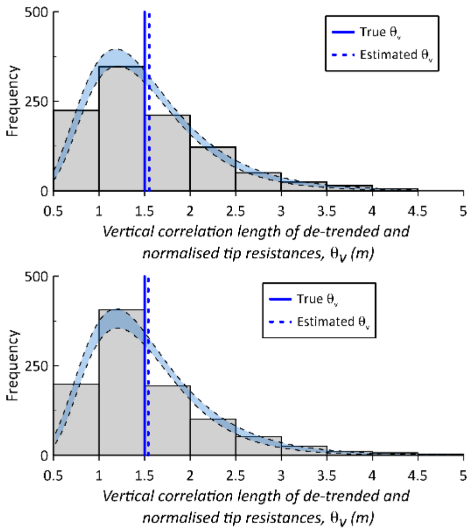

Figure 5. Histogram of estimated correlation lengths Markovian (top) or Gaussian (bottom).

3 CONCLUSIONS

The importance of the correlation length has been discussed in the context of probabilistic geotechnical analyses and its estimation has been assessed through a numerical investigation using artificially generated CPT data. The analyses showed that a very good accuracy is generally achieved when using a substantial number of realisations (1,000) demonstrating the applicability of the methods. However, a more detailed analysis on the individual estimated values of θ_v showed great variability which highlights the challenge and the potential error involved in the estimation process.

REFERENCES

Cami, B., S. Javankhoshdel, K. K. Phoon, and J. Ching. (2020). Scale of fluctuation for spatially varying soils: Estimation methods and values. *ASCE-ASME J. Risk Uncertainty Eng. Syst. Part A: Civ. Eng.*, 6(4): 03120002. https://doi.org/10.1061/AJRUA6.0001083.

Ching, J., M. Uzielli, K. K. Phoon and X. Xiaojun. (2023). Characterization of Autocovariance Parameters of Detrended Cone Tip Resistance from a Global CPT Database. *J. Geotech. Geoenviron. Eng.*, 149(10): 04023090. https://doi.org/10.1061/JGGEFK.GTENG-11214.

DeGroot, D. J., and G. B. Baecher. (1993). Estimating autocovariances of in-situ soil properties. *J. Geotech. Eng.*, 119(1):147–166. https://doi.org/10.1061/(ASCE)0733-9410(1993)119:1(147).

Fenton, G. A. and E. H. Vanmarcke. (1990). Simulation of random fields via local average subdivision. *J. Engineering Mechanics*, 116:(8), 1733-1749. https://doi.org/10.1061/(ASCE)0733-9399(1990)116:8(1733).

Fenton, G. A. (1999a). Estimation for stochastic soil models. *J. Geotech. Geoenviron. Eng.*, 125 (6): 470–485. https://doi.org/10.1061/(ASCE)1090-0241(1999)125:6(470).

Fenton, G. A. (1999b). Random field modeling of CPT data. *J. Geotech. Geoenviron. Eng.*, 125(6): 486–498. https://doi.org/10.1061/(ASCE)1090-0241(1999)125:6(486).

Fenton, G. A., and D. V. Griffiths. (2003). Bearing capacity prediction of spatially random c-φ soils. *Can. Geotech. J.*, 40(1): 54–65. https://doi.org/10.1139/t02-086.

Griffiths, D. V., J. Huang, and G. A. Fenton. (2011). Probabilistic infinite slope analysis. *Comput. Geotech.*, 38(4): 577–584. https://doi.org/10.1016/j.compgeo.2011.03.006.

Hicks, M. A., and W. A. Spencer. (2010). Influence of heterogeneity on the reliability and failure of a long 3D slope. *Comput. Geotech.*, 37 (7–8): 948–955. https://doi.org/10.1016/j.compgeo.2010.08.001.

Jaksa, M. (1995). The influence of spatial variability on the geotechnical design properties of a stiff, overconsolidated clay. *Ph.D. dissertation*, Dept. of Civil and Env. Engineering, Univ. of Adelaide.

Lloret-Cabot, M., G. A. Fenton, and M. A. Hicks. (2014). On the estimation of scale of fluctuation in geostatistics. *Georisk: Assess. Manage. Risk Eng. Syst. Geohazards*, 8(2): 129–140. https://doi.org/10.1080/17499518.2013.871189.

Lloret-Cabot, M., M. A. Hicks, and J. D. Nuttall. (2013). Investigating the Scales of Fluctuation of an Artificial Sand Island. In *Proceedings of the Int. Conf. GEO-INSTALL*, Hicks et al. (eds.), Rotterdam: CRC Press Taylor and Francis Group 192–197

Nie, X., Huang, H., Liu, Z., et al. (2015). Scale of fluctuation for geotechnical probabilistic analysis. In Schweckendiek et al. (eds.), Proc. *ISGSR2015: Geotechnical Safety and Risk V*, IOS Press, 816-821.

Phoon, K. K., and F. H. Kulhawy. (1999). Characterization of geotechnical variability. *Can. Geotech. J.*, 36(4): 612–624. https://doi.org/10.1139/t99-038.

Uzielli, M., G. Vannucchi, and K. K. Phoon. (2005). Random field characterisation of stress-normalised cone penetration testing parameters. *Géotechnique*, 55(1): 3–2. https://doi.org/10.1680/geot.2005.55.1.3.