Computers and Geotechnics 79 (2016) 159–172

Contents lists available at ScienceDirect

Computers and Geotechnics

journal homepage: www.elsevier.com/locate/compgeo

Research Paper

Uncertainty reduction and sampling efficiency in slope designs using 3D conditional random fields

Y.J. Li, M.A. Hicks *, P.J. Vardon

Section of Geo-Engineering, Department of Geoscience & Engineering, Faculty of Civil Engineering & Geosciences, Delft University of Technology, Delft, The Netherlands

A R T I C L E I N F O

Article history:
Received 21 December 2015
Received in revised form 28 April 2016
Accepted 30 May 2016

Keywords:
Conditional random fields
Kriging
Reliability
Sampling efficiency
Spatial variability
Uncertainty reduction

A B S T R A C T

A method of combining 3D Kriging for geotechnical sampling schemes with an existing random field generator is presented and validated. Conditional random fields of soil heterogeneity are then linked with finite elements, within a Monte Carlo framework, to investigate optimum sampling locations and the cost-effective design of a slope. The results clearly demonstrate the potential of 3D conditional simulation in directing exploration programmes and designing cost-saving structures; that is, by reducing uncertainty and improving the confidence in a project's success. Moreover, for the problems analysed, an optimal sampling distance of half the horizontal scale of fluctuation was identified.

1. Introduction

Soil properties exhibit three dimensional spatial variability (i.e. heterogeneity). In geotechnical engineering, a site investigation may be carried out, and the data collected and processed in a statistical way to characterise the variability [1–10]. The outcomes of the statistical treatment, e.g. the mean property value, the standard deviation or coefficient of variation, and the spatial correlation distance, may be used as input to a geotechnical model capable of dealing with the spatial variation (e.g. a random field simulation). However, when it comes to making use of the field data, there arises the question: How can we make best use of the available data? The idea is to use the data more effectively, so that it is worth the effort or cost spent in carrying out the investigation, as well as the additional effort in post-processing the data. The aim of this paper is to contribute towards answering this question.

For example, cone penetration tests (CPTs) are often carried out in geotechnical field investigations, in order to obtain data used in implementing the design of a structure. The amount of data from CPT measurements is often larger than from conventional laboratory tests. This is useful, as a large database is needed to accurately estimate the spatial correlation structure of a soil property. For example, Fenton [3] used a database of CPT profiles from Oslo to estimate the correlation statistics in the vertical direction, and Jaksa et al. [5] used a database from Adelaide to estimate the correlation distances in both the vertical and horizontal directions.

In geotechnical engineering, a substantial amount of numerical work has been done using idealised 2D simulations based on collected in-situ data (e.g. [4]), although a 3D simulation would be preferable due to site data generally being collected from a 3D space. However, there are relatively few studies simulating the effect of 3D heterogeneity due to the high computational requirements. Examples include the effect of heterogeneity on shallow foundation settlement [11–13], on steady state seepage [14–16], on seismic liquefaction [17] and on slope reliability [18–27].

The above investigations all used random fields to represent the soil spatial variability and the finite element method to analyse geotechnical performance within a Monte Carlo framework, a form of analysis sometimes referred to as the random finite element method (RFEM) [28]. However, they did not make use of the spatial distribution of related measurement data to constrain the random fields. In other words, for those applications that are based on real field data, many realisations not complying with the field data at the measurement locations will be included in the simulation, which, in turn, will result in an exaggerated range of responses in the analysis of geotechnical performance.

Studies on conditional simulations are available in geostatistics in the field of reservoir engineering [29]. However, there are not many studies dealing with soil spatial variability in geotechnical engineering that utilise conditional simulation (some 2D exceptions

* Corresponding author.
E-mail addresses: y.li-4@tudelft.nl (Y.J. Li), m.a.hicks@tudelft.nl (M.A. Hicks), p.j.vardon@tudelft.nl (P.J. Vardon).

http://dx.doi.org/10.1016/j.compgeo.2016.05.027
0266-352X/

include, e.g., [6,30–32]). This is partly due to the smaller amount of data generally available in geotechnical engineering, and partly due to there often not being a computer program specially implemented for those situations where there are sufficient data (e.g. CPT, vane shear test (VST)), especially in 3D. However, unconditional random fields can easily be conditioned to the known measurements by Kriging [29,33]. Hence, following the previous 2D work of Van den Eijnden and Hicks [31] and Lloret-Cabot et al. [30], this paper seeks to implement and apply conditional simulation in three dimensional space, in order to reduce uncertainty in the field where CPT measurements are carried out.

Usually, site investigation plans are designed to follow some regular pattern. For example, a systematic grid of sample locations is generally used, due to its simplicity to implement [5]. Moreover, although there are various sampling plans in terms of layouts, it is found that systematically ordered spatial samples are superior in terms of the quality of estimates at unsampled locations [34]. Therefore, this paper will be devoted to implementing a 3D Kriging algorithm for sampling schemes following a regular grid. This will then be combined with an existing 3D random field generator to implement a conditional simulator. However, extension to irregular sampling patterns is straightforward based on the presented framework.

The implemented approach has been applied to two idealised slope stability examples. The first demonstrates how the approach may be used to identify the best locations to conduct borehole testing, and thereby allow an increased confidence in a project's success or failure to be obtained. While it is very important to pay sufficient attention to the required intensity of a site investigation (i.e. the optimal number of boreholes) with respect to the site-specific spatial variability, as highlighted by Jaksa et al. [12], the first example starts by focusing on the optimum locations for carrying out site investigations for a given number of boreholes, before moving on to consider the intensity of testing. The second example compares different candidate slope designs, in order to choose the best (most cost-effective) design satisfying the reliability requirements.

For simplicity, this paper focuses on applications involving only a single soil layer (i.e. a single layer characterised by a statistically homogeneous undrained shear strength), although the extension to multiple soil layers is straightforward. Moreover, the effect of random variation in the boundary locations between different soil layers can also be easily incorporated by conditioning to known boundary locations (e.g. corresponding to where the CPTs have been carried out).

2. Theory and implementation

2.1. Conditioning

A conditional random field, which preserves the known values at the measurement locations, can be formed from three different fields [28,35,36]:

$$Z_{rc}(\mathbf{x}) = Z_{ru}(\mathbf{x}) + (Z_{km}(\mathbf{x}) - Z_{ks}(\mathbf{x})) \tag{1}$$

where $\mathbf{x}$ denotes a location in space, $Z_{rc}(\mathbf{x})$ is the conditionally simulated random field, $Z_{ru}(\mathbf{x})$ is the unconditional random field, $Z_{km}(\mathbf{x})$ is the Kriged field based on measured values at $\mathbf{x}_i (i = 1, 2, \ldots, N)$, $Z_{ks}(\mathbf{x})$ is the Kriged field based on unconditionally (or randomly) simulated values at the same positions $\mathbf{x}_i (i = 1, 2, \ldots, N)$, and N is the number of measurement locations.

The unconditional random field can be simulated via several methods [37]; for example, interpolated autocorrelation [38], covariance matrix decomposition, discrete Fourier transform or Fast Fourier transform, turning bands, local average subdivision (LAS), and Karhunen–Loeve expansion [39], among others. The LAS method [40] is used in this paper. The Kriged fields are obtained by Kriging [41], which has found extensive usage in geostatistics [42,43]. The LAS and Kriging methods are briefly reviewed in the following sections.

2.2. Anisotropic random field generation using 3D LAS

The LAS method [40,44] is used herein to generate the unconditional random fields, using statistics (i.e. mean, variance and correlation structure) based on the observed field data. The LAS method proceeds in a recursive fashion, by progressively subdividing the initial domain into smaller cells, until the random process is represented by a series of local averages. The major advantage is its ability to produce random fields of local averages whose statistics are consistent with the field resolution; that is, it maintains a constant mean over all levels of subdivision, and ensures reduced variances as a function of cell size based on variance reduction theory [45], taking account of spatial correlations between local averages within each level and across levels.

The following covariance function is used in the subdivision process:

$$C(\boldsymbol{\tau}) = C(\tau_1, \tau_2, \tau_3) = \sigma^2 \exp\left(-\frac{2|\tau_1|}{\theta_1} - \sqrt{\left(\frac{2\tau_2}{\theta_2}\right)^2 + \left(\frac{2\tau_3}{\theta_3}\right)^2}\right) \tag{2}$$

where σ^2 is the variance of the soil property, $\boldsymbol{\tau}$ is the lag vector, and θ_1, θ_2 and θ_3, and τ_1, τ_2 and τ_3 are the respective scales of fluctuation and lag distances in the vertical and two lateral coordinate directions, respectively. Herein, an isotropic random field is initially generated by setting $\theta_1 = \theta_2 = \theta_3 = \theta_{iso}$; i.e. so that θ_{iso} equals the horizontal scale of fluctuation, θ_h. This field is then squashed in the vertical direction to give the target vertical scale of fluctuation, θ_v. The 3D LAS implementation of Spencer [25] has been used in this paper, and the reader is referred to Spencer [25] and Hicks and Spencer [19] for more details. Note also that a truncated normal distribution has been used to describe the pointwise variation in material properties [19].

2.3. Kriging

In contrast to conventional deterministic interpolation techniques, such as moving least squares and the radial point interpolation method, Kriging incorporates the variogram (or covariance) into the interpolation procedure; specifically, information on the spatial correlation of the measured points is used to calculate the weights. Moreover, standard errors of the estimation can also be obtained, indicating the reliability of the estimation and the accuracy of the prediction. Kriging is a method of interpolation for which the interpolated values are modelled by a Gaussian process governed by prior covariances and for which confidence intervals can be derived. While interpolation methods based on other criteria need not yield the most likely intermediate values, Kriging provides a best linear unbiased prediction of the soil properties (Z) between known data [43,46] by assuming the stationarity of the mean and of the spatial covariances, or variograms. A brief review is first given to facilitate understanding of the implementation.

Suppose that $Z_1, Z_2, \ldots, Z_N$ are observations of the random field $Z(\mathbf{x})$ at points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ (i.e. $Z_i = Z(\mathbf{x}_i)$ $(i = 1, 2, \ldots, N)$). The best linear unbiased estimation (i.e. $\hat{Z}$) of the soil property at some location $\mathbf{x}_0$ is given by

$$\hat{Z}(\mathbf{x}_0) = \sum_{i=1}^{N} \lambda_i Z_i = \sum_{i=1}^{N} \lambda_i(\mathbf{x}_0) Z(\mathbf{x}_i) \tag{3}$$