Investigating the Influence of Conditional Simulation on Small-Probability Failure Events Using Subset Simulation

van den Eijnden, Bram; Hicks, Michael; Vardon, Phil

Publication date
2017
Document Version
Accepted author manuscript
Published in
6th international symposium on geotechnical safety and risk

Investigating the influence of conditional simulation on small-probability failure events using subset simulation

Bram van den Eijnden,[1] Michael A. Hicks,[1] and Philip J. Vardon[1]

[1]Geo-Engineering Section, Department of Geoscience and Engineering, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, Netherlands;
e-mail: A.P.vandenEijnden@tudelft.nl; M.A.Hicks@tudelft.nl; P.J.Vardon@tudelft.nl

ABSTRACT

Spatial variation of soil strength parameters is a dominating uncertainty in slope stability analysis. This uncertainty can be accounted for in a stochastic description, based on a global geostatistical characterization of the soil strength parameters, which leads to a wide range of possible slope responses, of which only a small proportion typically concern slope failure. This paper investigates the effect of including additional data to reduce the range of possible scenarios in the stability analysis of slopes in spatially variable soils. Subset simulation, which is a technique to focus the random sampling of Monte Carlo analyses in the region of interest, is used here to address the already small probabilities of slope failure. The reduction in the range of possible scenarios with respect to both the probability of failure and modes of failure is then investigated. The strong reduction in the range of expected failure modes through conditional simulation demonstrates the uncertainty reduction and the relationship between the distribution of weaker zones in the slope and the development of sliding surfaces.

INTRODUCTION

Numerical stability analysis of slopes and embankments involves many forms of uncertainty that should be accounted for. These uncertainties can, for example, be related to the amount of available data, assumptions in the translation of this data into a characterizing model of the slope, or to the numerical methods for evaluating the response of a slope to specific loading conditions. Several of these uncertainties are addressed within the authors' research group, of which different contributions can be found in these proceedings. These contributions deal with the characterization and stability analysis of slopes and embankments in heterogeneous soils, taking account of heterogeneity as an explicit part of the model. Soil property characterization, including missing knowledge of the spatial variation (i.e. heterogeneity) and its characterization by (idealised) models, is here considered to be the dominant source of uncertainty. Structure specific spatial variability is addressed by Gast et al. (2017), who calibrated the models for characterizing spatial variability against detailed site investigation data.

Whereas the bias of simplified numerical models that take account of soil heterogeneity can be significant (see for example Varkey et al. (2017)), results obtained using more advanced numerical methods based on 3D finite element methods (e.g. Li et al. (2016c)) are generally considered more objective with respect to the numerical formulation within its domain of application. Note that, within this domain of validity of the finite element formulation, stability analysis only focusses on the initiation of slope failure, whereas the actual consequence of progressive failure requires numerical models that are more adequate for dealing with large deformations and structural reconstitution, such as the random material point method (Wang et al. 2016). This paper considers the reduction of uncertainty in the numerical simulation of slope stability by optimizing the use of available data. The focus here is on the reduction of uncertainty, although reducing the range of responses by including more of the available data can also reduce the calculated probability of failure.

When uncertainty in the slope characterization is accounted for by means of a number of stochastic variables and possible scenarios for slope response are random samples of these variables, the probability of failure can be expressed as a subdomain of the sampling space. For low (smaller than 10) numbers of variables, efficient integration methods exist to approximate the domain integral defining the probability of failure, among which are the FORM/SORM (Low 2014) and point estimate methods (Christian and Baecher 1999). For problems with more independent variables, only statistical methods such as Monte Carlo simulation (MCS) can be used. At lower levels of probability, MCS becomes inefficient in generating a sufficient number of relevant (failing) realizations and alternative simulation strategies are needed.

Subset simulation was proposed as an improved version of MCS, based on Bayesian statistics theory, to efficiently address small probabilities in multivariate problems (Au and Beck 2001). This method was applied in slope stability analysis within the framework of the random finite element method (RFEM) by Li et al. (2016a). Modifications were proposed to improve the efficiency of the method, based on computationally less expensive surrogate models (Li et al. 2016b, Xiao et al. 2016) or indicative relations (Huang et al. 2016). A modification of the subset simulation algorithm itself was proposed in Eijnden and Hicks (2016), which involved changing from a probability-based subset selection to a threshold-based subset selection. This approach overcomes the time-consuming evaluation of the exact factor of safety for each realization, without relying on empirical indicator relations.

Here, subset simulation is used to investigate the probability and modes of failure in slopes with spatially varying shear strength, conditioned by CPT measurement data. Evaluation of the realizations failing without applying shear strength reduction allows the comparison of modes of failure of conditioned and unconditioned realizations of shear strength variability. Moreover, the resulting probabilities of failure are used to demonstrate that taking account of more of the available data can lead to a significant reduction in the calculated risk.

METHODS

With respect to the various types of uncertainty in an analysis, it is here assumed that the system uncertainty is controlled by the uncertainty in the spatial variation of the material strength and that any other uncertainty can be considered small in comparison. Hence, only uncertainty in the spatial variation of undrained shear strength c_u has been accounted for.

Random finite element method

The random finite element method (RFEM) (Griffiths and Fenton 2004) is a rigorous approach to reliability analysis that can be used for the stability analysis of slopes with uncertainty in the spatial variability of (strength) parameters. It combines random field theory with finite element analysis to evaluate the range of structural responses in MCS. Random fields are used as possible realizations of the spatial distribution of strength parameters, and the slope failure of a realization can be enforced by applying a shear strength reduction factor f_s. This factor is part of the analysis and, based on the choice of f_s, the slope will or will not fail. The smallest strength reduction factor leading to the failure of a specific realization of a slope defines the factor of safety *FOS*. *FOS* is therefore a realization-specific property and can be found by iteratively updating f_s in the analysis of a single realization. Here, realizations are only tested for slope failure ($FOS \leq f_s$) and computationally expensive iterations to find the exact value for *FOS* are avoided. Subset simulation (see below) is used to investigate the response of slopes with $FOS \leq 1.0$.

Conditional random field simulation

The actual field of spatially variable shear strength is generated by a deterministic transformation of a sample drawn from the standard normal sampling space. Although any algorithm for generating spatially correlated fields of strength parameters could theoretically be used as a transformation, covariance matrix decomposition (CMD) is used here. It decomposes the field into a minimum number of required variables without loss of accuracy.

Conditional random fields make use of additional available data to constrain the range of possible realizations. These data are typically the measurements from which the random field (spatial) statistics are derived. The conditional simulation of the variability in strength parameters in a Kriging-based formulation was used in Li et al. (2016c), for the conditioning of a 3D random field in slope stability analyses of slopes that were long in the third (out of plane) dimension. An alternative formulation is derived here, by preconditioning the algorithm for random field generation, and accounting for the local averaging in the discretization cells as well as the exact correlations between cell averages and point data (field measurements). For this purpose, a combined vector $\widetilde{\boldsymbol{Z}}$ of random field data (simulated) and CPT data (transformed data from in-situ measurements) is defined as:

$$\widetilde{\boldsymbol{Z}} = \begin{Bmatrix} \boldsymbol{Z}_f \\ \boldsymbol{Z}_{cpt} \end{Bmatrix} \tag{1}$$

with column vector $\boldsymbol{Z_f}$ being the standard normal transformation of the discretized random field of strength parameter $c_u(\vec{x})$ and $\boldsymbol{Z}_{cpt}$ being the standard normal transformation of the interpreted undrained shear strength profile of c_u derived from the CPT measurements. Covariance matrix $\tilde{\mathbf{C}}$ describes the autocovariance of the standard normal components of $\tilde{\boldsymbol{Z}}$. Direct integration of the correlation function over the random field discretization cells accounts for local averaging in the discretization cells and the method is independent of the correlation kernel $\rho(\tau)$ in which τ is a normalized distance. Covariance between the local averages of two cells, A and B, is given by the double integral of $\rho(\tau)$ over the cell domains Ω_A and Ω_B with cell volumes V_A and V_B:

$$C(\Omega_A, \Omega_B) = \frac{1}{V_A V_B} \iint_{\Omega_A \Omega_B} \rho(\tau) dV dV \tag{2}$$

A single integral is used for the covariance between a measurement point and a cell average value. Without prior knowledge, realizations of $\tilde{\boldsymbol{Z}}$ can be generated using covariance matrix decomposition:

$$\tilde{\boldsymbol{Z}} = \tilde{\mathbf{C}}^{\frac{1}{2}} \tilde{\boldsymbol{\xi}}, \quad \tilde{\mathbf{C}}^{\frac{1}{2}} = \boldsymbol{\Phi} \boldsymbol{\Lambda}^{\frac{1}{2}} \boldsymbol{\Phi}^{\mathrm{T}} \tag{3}$$

where $\tilde{\boldsymbol{\xi}}$ is a column vector of uncorrelated standard normal variables, $\boldsymbol{\Phi}$ is a matrix with the eigenvectors of $\tilde{\mathbf{C}}$ and $\boldsymbol{\Lambda}$ is a diagonal matrix with corresponding eigenvalues. An eigen decomposition is needed to derive the eigenvectors and eigenvalues. Partitioning of the covariance matrix leads to:

$$\tilde{\mathbf{C}} = \begin{bmatrix} \mathbf{C}_{\mathrm{ff}} & \mathbf{C}_{\mathrm{fc}} \\ \mathbf{C}_{\mathrm{cf}} & \mathbf{C}_{\mathrm{cc}} \end{bmatrix} \tag{4}$$

with subscript f referring to the discretized random field data and subscript c to CPT measurement data or any other available data. With prior knowledge of the CPT data, Kriging theory can be used to derive a direct expression for the discretized random field $\boldsymbol{Z}_f$:

$$\boldsymbol{Z}_f = \mathbf{C}_{\mathrm{fc}} \mathbf{C}_{\mathrm{cc}}^{-1} \boldsymbol{Z}_{cpt} + [\mathbf{C}_{\mathrm{ff}} - \mathbf{C}_{\mathrm{fc}} \mathbf{C}_{\mathrm{cc}}^{-1} \mathbf{C}_{\mathrm{cf}}]^{\frac{1}{2}} \boldsymbol{\xi}_f \tag{5}$$

Comparing Equation (5) with earlier work on conditional random fields (Journel and Huijbregts 1978; Eijnden and Hicks 2011, Lloret-Cabot et al. 2012), the term $\mathbf{C}_{\mathrm{fc}} \mathbf{C}_{\mathrm{cc}}^{-1} \boldsymbol{Z}_{cpt}$ is equal to the Kriged field based on the CPT profile data, as it defines the expectation of the normalized conditional random field $\mathrm{E}[\boldsymbol{Z_f}]$. The additional term $[\mathbf{C}_{\mathrm{ff}} - \mathbf{C}_{\mathrm{fc}} \mathbf{C}_{\mathrm{cc}}^{-1} \mathbf{C}_{\mathrm{cf}}]$ defines the remaining variation from the expected mean, following a normal distribution. The reduced uncertainty can be expressed in the standard deviation of the discretized conditional random field $\boldsymbol{Z}_f$:

$$\boldsymbol{\sigma}_f = diag([\mathbf{C}_{\mathrm{ff}} - \mathbf{C}_{\mathrm{fc}} \mathbf{C}_{\mathrm{cc}}^{-1} \mathbf{C}_{\mathrm{cf}}])^{1/2} \tag{6}$$

Subset simulation

When considering the (standard normal) samples characterizing random field realizations as points in standard normal sampling space, a failure domain can be defined as a subspace of this sampling space containing all possible realizations leading to failure for a given f_s. An efficient strategy of sampling from this subspace has the potential of decreasing the number of realizations required for a reliability analysis, compared with Monte Carlo simulation. In the context of slope stability analysis, subset simulation provides such a strategy.

A set of realizations failing at an intermediate level of f_s is first generated. This set represents a subdomain of the sampling space with realizations $FOS < f_s^{(i)}$. Markov chain Monte Carlo simulation is then used to extend the number of realizations inside the conditional domain (i.e. $FOS < f_s^{(i)}$), after which a next subset selection if made for which $FOS < f_s^{(i+1)} < f_s^{(i)}$. This procedure is repeated until f_s = 1.00 is reached, after which the corresponding probability of failure is found as the product of conditional probabilities $P(FOS < f_s^{(i)} | FOS < f_s^{(i-1)})$.

The efficiency of subset simulation, measured in the required number of realizations to be analysed, converges proportionally to $\log(1/p_f)$, compared to $1/p_f$ for MCS. In this work, a modified version of subset simulation is applied, using performance-based subset selection rather than the classical probability-based subset selection. This modification overcomes the need for determining the true factor of safety of the realizations by predefining the performance threshold $f_s^{(i+1)}$ for subset selection and allowing a variation of conditional subset probabilities $P(FOS < f_s^{(i)} | FOS < f_s^{(i-1)})$ based on this performance. Details on subset simulation and its application to slope stability analysis can be found in Au and Beck (2001), Li et al. (2016a) and Eijnden and Hicks (2016).

EXAMPLE SIMULATION

An example case of a slope constructed in a cohesive soil is studied. The material behaviour is considered to be linear elastic perfectly plastic, and locally characterized by Young's modulus $E = 100$ kPa, Poisson ration $\nu = 0.3$, unit weight $\gamma = 17$ kPa and undrained shear strength c_u. The shear strength exhibits spatial variability, of which the (spatial) statistics can be estimated from cone penetration test (CPT) profiles (Lloret-Cabot et al. 2014), assumed here to be taken at the centre of the future slope face. This profile will serve as the additional information that constrains the variability of the shear strength in the domain of the analysis.

A lognormal distribution is assumed for c_u. The mean and standard deviation are 26.7 kPa and 6.7 kPa respectively, corresponding to a coefficient of variation $CoV = 0.25$. The horizontal and vertical scales of fluctuation are $\theta_x = 8$ m and $\theta_y = 1.25$ m and an exponential spatial correlation function is adopted:

$$\rho(\tau) = \exp(-2\tau), \qquad \tau = \sqrt{\left(\frac{x_B - x_A}{\theta_x}\right)^2 + \left(\frac{y_B - y_A}{\theta_y}\right)^2} \tag{7}$$

where τ is the normalized distance between any two points A and B. The geometry and boundary conditions of the slope are given in Figure 1; a slope of 5 m height at an angle of 45° is constructed in a soil layer of 10 m thickness. Vertical displacements on the sides of the domain are allowed, whereas displacements are fully constrained on the lower domain boundary. Loading is by the material self-weight under a quasi-static assumption. The slope can be brought to failure by applying a strength reduction factor f_s.

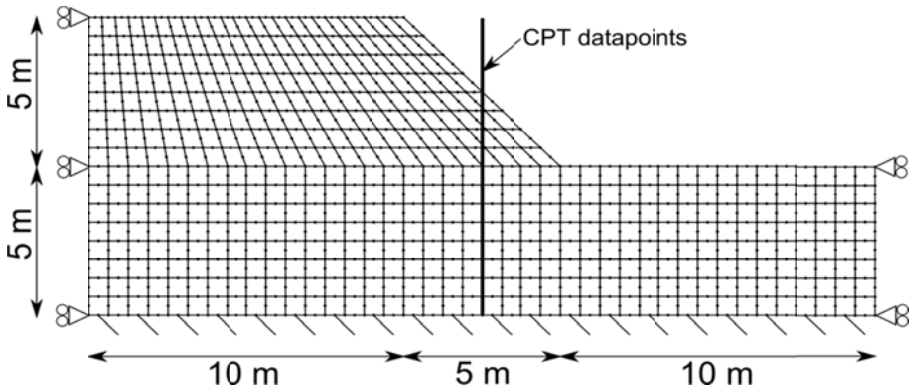

Figure 1: Slope geometry, boundary conditions and finite element mesh.

A relatively coarse mesh is used (the ratio between random field discretization cell and minimum scale of fluctuation is 0.25) and the analysis is performed under plane strain conditions, rather than as a 3D analysis. Although this simplification influences the quantitative results of the analysis, it has no influence on the conceptual performance of the proposed methods of subset simulation or conditional simulation. CPT data were generated artificially along a vertical profile of measurement points at intervals of 2 cm following the geostatistical characteristics introduced above. Figure 2 gives five c_u-profiles, used here as examples of possible profiles of undrained shear strength interpreted from CPT data.

Slope stability analyses are here performed within the framework of RFEM, by analysing the stability (failure or no failure) of each realization against a predefined strength reduction factor f_s. Figure 3 gives an example of a realization with spatially variable shear strength, conditioned by CPT 4. The field matches the CPT profile at the conditioning locations.

SIMULATION RESULTS

First, a reference case was investigated, in which the above mentioned slope was analysed without conditioning of the random field. The probability of failure as a function of the strength reduction factor f_s was evaluated using subset simulation, by reducing the strength reduction factor incrementally to f_s = 1.00. 500 failing slope realizations were generated at each subset level, for which the target conditional probability p_0 was set at 0.2. This corresponds to an expected 2500 analyses per subset. The relationship between p_f and f_s was then compared with simulations in which the random fields were conditioned by the CPT profiles from Figure 2.

These profiles are synthetic realizations in accordance with Equations 3-4, with Equation 4 reducing to $\boldsymbol{Z}_{cpt} = \mathbf{A}_{cc}\boldsymbol{\xi}_{cpt}$ for standard normal profiles $\boldsymbol{Z}_{cpt}$. Figure 4 shows that using the additional information available from a single CPT profile can lead to a significant reduction in the probability of failure (from $p_f \approx 0.0005$ for unconditional simulations to $p_f << 1\times10^{-7}$ for conditional simulations). A reduction in the probability of failure is not guaranteed and profiles can exist that increase the probability of failure when used to condition the random field. However, the introduction of additional data reduces the uncertainty in the response, and this is likely to result in a lower probability of failure.

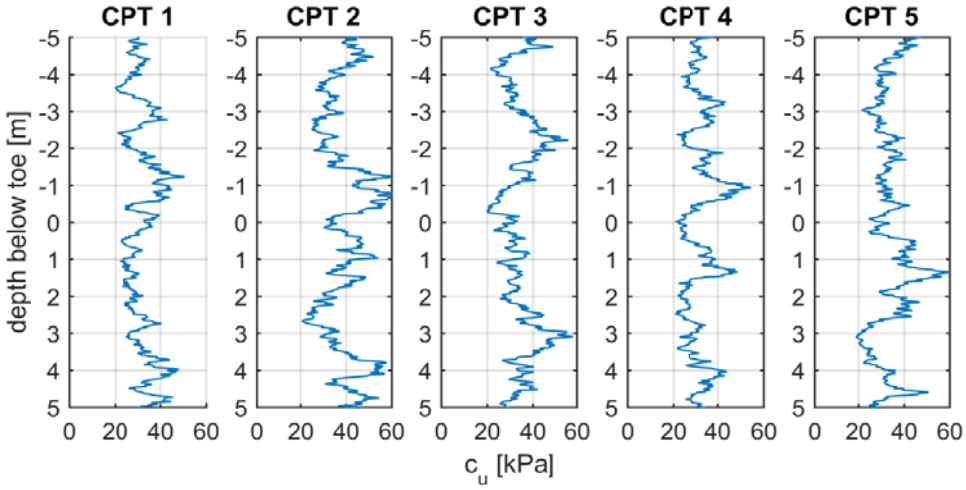

Figure 2: Simulated c_u profiles for random field conditioning, as could be interpreted from CPT data.

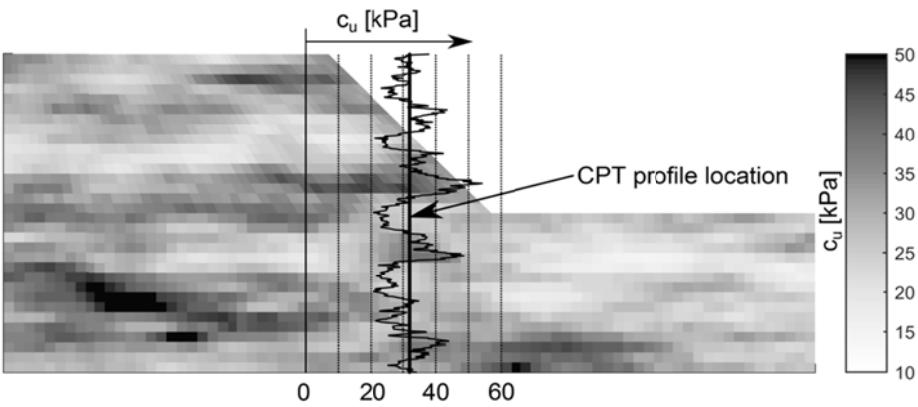

Figure 3: Typical realization of a c_u random field conditioned by a single CPT profile.

The resulting realizations of slopes failing at $f_s = 1.00$ were used to study the modes of failure. The geometry of the sliding body was determined by applying the K-means clustering method to the displacement field (Huang et al. 2013). The depth of the corresponding sliding surface is then used to characterize the mode of failure of a realization. The distributions of the depth of the sliding surfaces are given in Figure 5 together with the corresponding CPT profiles. A strong

influence of conditioning can be observed in the sliding depth, as the distributions show a strong clustering around the weakest parts of the profiles. This confirms that the sliding surface is attracted by the weaker parts of the domain.

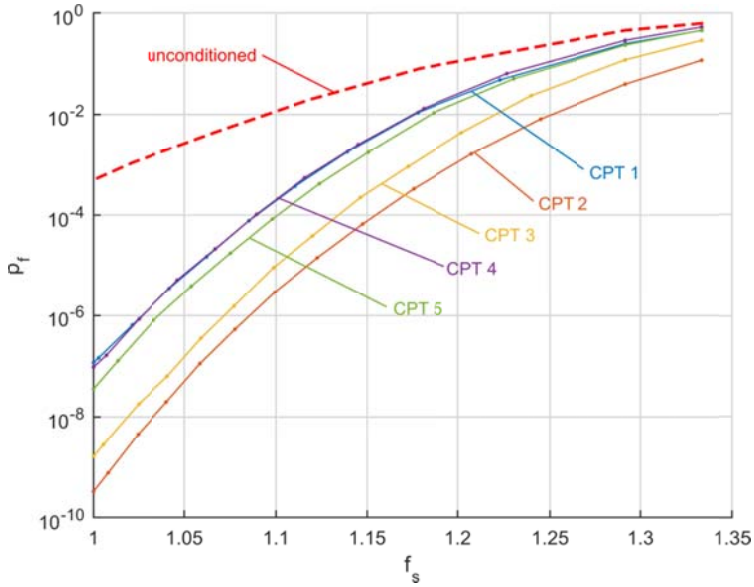

Figure 4: Probability of failure p_f as a function of the strength reduction factor f_s for simulations with different conditioning CPT profiles.

The distributions of sliding depths from unconditional simulations are included in Figure 5 for comparison. Where unconditional simulation shows predominantly shallow modes of failure, simulations conditioned by additional data can lead to distributions of sliding depths well outside the unconditional probability distribution. The lower graphs of Figure 5 show a similar comparison based on sliding volume; depending on the conditioning profile, the expected sliding mass can be much larger than for an unconditional analysis. However, these larger sliding volumes have a much lower probability of occurrence (see Figure 4).

CONCLUSION

The simulation of slope failure at small failure probabilities using a new subset simulation strategy was used to demonstrate the effect that uncertainty reduction by conditional simulation can have on the predicted modes of failure and the calculated probability of failure. The strong correlation between the depth of the sliding surfaces and weak zones in the different conditioning profiles demonstrates the tendency for sliding surfaces to seek out the weaker parts of the slope. Using several artificial conditioning profiles, it was demonstrated that conditional simulation can significantly reduce the calculated probability of failure, well below probability levels generally accessible by MCS. It can be concluded that for a full risk assessment, where these probabilities are linked with consequences (e.g. sliding volumes), more efficient simulation approaches such as subset simulation become indispensable in evaluating the correct modes of failure.

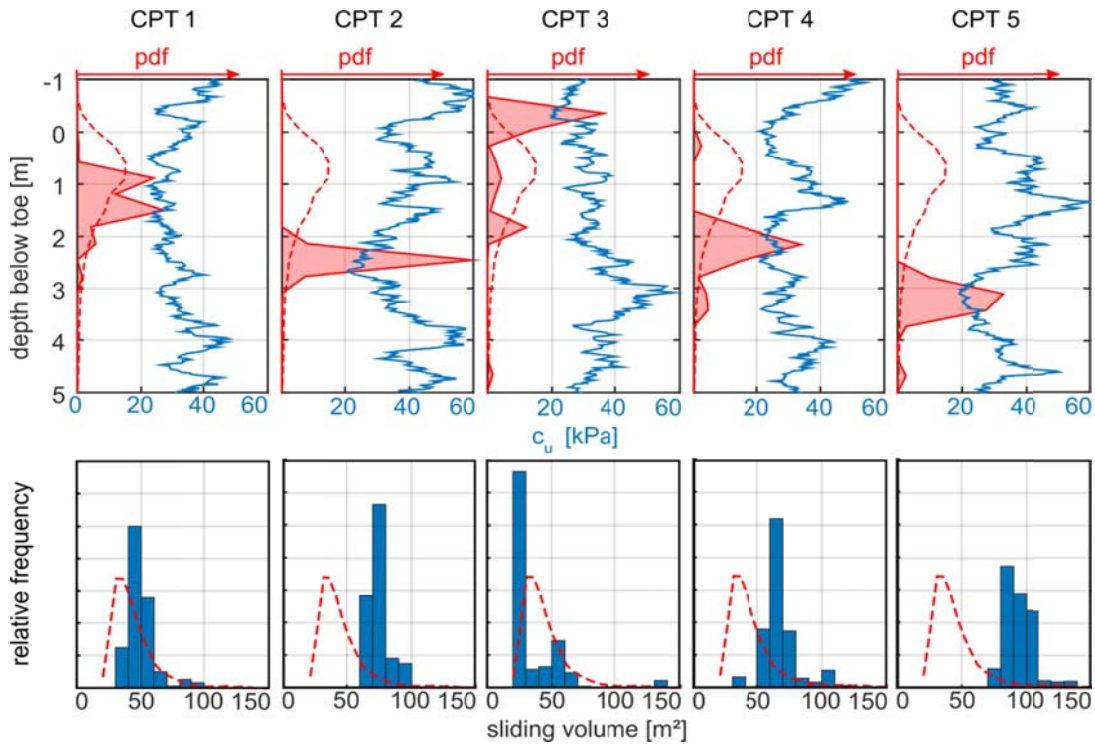

Figure 5: Top: Distributions of slip surface depth at $f_s = 1.0$ for simulations conditioned by CPT profiles. Bottom: Distributions of sliding volume. Dashed line - - - - indicates the unconditioned simulation results (from Eijnden and Hicks (2016)).

ACKNOWLEDGEMENTS

This research is supported under project number 13864 by the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

Au, S. K. and Beck, J. L. (2001). "Estimation of small failure probabilities in high dimensions by subset simulation." *Probabilistic Engineering Mechanics*, 16(4), 263–277.

Christian, J. T. and Baecher, G. B. (1999). "Point-estimate method as numerical quadrature." *Journal of Geotechnical and Geoenvironmental Engineering*, 125(9), 779-786.

Eijnden, A. P. van den and Hicks, M. A. (2011). "Conditional simulation for characterizing the spatial variability of sand state." In: *Proc 2nd Int symp on computational geomechanics*, Dubrovnik, Croatia, pp 288–296.

Eijnden, A. P. van den and Hicks, M. A. (2016). "Efficient subset simulation for evaluating the modes of improbable failure." *Submitted to Computers and Geotechnics*.

Gast, T. de, Vardon, P. J., Hicks, M. A. (2017). "Estimating spatial correlations under man-made structures on soft soils." *Accepted for ISGSR 2017.*

Griffiths, D.V., and Fenton, G.A. (2004). "Probabilistic slope stability analysis by finite elements." *Journal of Geotechnical and Geoenvironmental Engineering*, 130(5), 507-518.

Journel, A. G. and Huijbregts, C. J. (1978). *Mining geostatistics*, Academic press.

Huang, J., Fenton, G., Griffiths, D. V., Li, D. Q., and Zhou, C. (2016). "On the efficient estimation of small failure probability in slopes." *Landslides*, DOI 10.1007/s10346-016-0726-2.

Huang, J., Lyamin, A. V., Griffiths, D. V., Krabbenhoft, K., and Sloan, S. W. (2013). "Quantitative risk assessment of landslide by limit analysis and random fields." *Computers and Geotechnics*, 53, 60-67.

Li, D. Q., Xiao, T., Cao, Z. J., Zhou, C. B., and Zhang, L. M. (2016a). "Enhancement of random finite element method in reliability analysis and risk assessment of soil slopes using Subset Simulation." *Landslides*, 13(2), 293-303.

Li, D. Q., Xiao, T., Cao, Z. J., Phoon, K. K., and Zhou, C. B. (2016b). "Efficient and consistent reliability analysis of soil slope stability using both limit equilibrium analysis and finite element analysis." *Applied Mathematical Modelling*, 40(9), 5216-5229.

Li, Y. J., Hicks, M. A., and Vardon, P. J. (2016c). "Uncertainty reduction and sampling efficiency in slope designs using 3D conditional random fields." *Computers and Geotechnics*, 79, 159-172.

Lloret-Cabot, M., Hicks, M. A., and Eijnden, A. P. van den (2012). "Investigation of the reduction in uncertainty due to soil variability when conditioning a random field using Kriging." *Géotechnique Letters,* 2, 123-127.

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

Low, B. K. (2014). "FORM, SORM, and spatial modeling in geotechnical engineering." *Structural Safety*, 49, 56-64.

Varkey, D., Hicks, M. A., Vardon, P. J., (2017). "Influence of spatial variability of shear strength parameters on 3D slope reliability and comparison with an analytical method." *Accepted for ISGSR 2017.*

Wang, B., Hicks, M.A., and Vardon, P.J. (2016). "Slope failure analysis using the random material point method." *Géotechnique Letters*, 6, 113-118.

Xiao, T., Li, D. Q., Cao, Z. J., Au, S. K., and Phoon, K. K. (2016). "Three-dimensional slope reliability and risk assessment using auxiliary random finite element method." *Computers and Geotechnics*, 79, 146-158.