Characterization of random fields at multiple scales: an efficient conditional simulation procedure and applications in geomechanics

Jack W. Baker and Andrew Seifried
Dept. of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305-4020, USA
Jose E. Andrade and Qiushi Chen
Department of Civil Engineering and Applied Mechanics, California Institute of Technology, Pasadena, CA 91125, USA

ABSTRACT: This paper describes an approach to account for the multi-scale nature of soil by using a multi-scale hierarchical Monte Carlo simulation framework. The behavior of particulate media, such as sands, is encoded at the granular-scale, and so methods for accurately predicting soil behavior must rely on methods for up-scaling such behavior across relevant scales of interest. Multi-scale analysis is known to be especially important under strain localization, penetration or liquefaction conditions, where a classical constitutive description may no longer apply. A probabilistic framework across multiple scales is needed to efficiently model and simulate multiscale fields of spatially varying material properties and to consistently compute the behavior of the material in a multi-scale model. From a material modeling standpoint, the multi-scale framework is facilitated here using a hierarchical conditional simulation procedure. With this approach, a more accurate material description at finer scales is pursued only when needed, such as in the presence of strong inhomogeneities. Monte Carlo simulation is used to simulate material properties at an initial coarse scale, and that initial simulation is adaptively refined at finer scale materials whenever necessary, conditional upon previously simulated coarse scale data. Here the background of the multiscale geomechanics motivation is summarized, the mathematics of this simulation approach is developed, and then several example calculations are shown to bring insights regarding the approach and its potential application in problems where multi-scale effects are important. Details regarding open-source software documenting these calculations are also provided.

1 INTRODUCTION

Many geotechnical engineering problems are multi-scale in nature because of inhomogeneities existing at different length-scales in geomaterials. Figure 1 shows length scales relevant to modeling of granular materials in civil engineering applications. Information pertaining to granular systems, including inhomogeneities, is encoded at the granular scale and propagated up to the field scale. For example, it is known that compactive shear bands in soils and rocks are zones of intense deformation where significant reductions in porosity are observed and hence strength and permeability characteristics in the medium are completely altered due to imposed deformations (Desrues and Viggiani 2004; Holcomb and Olsson 2003). These imposed changes are rather local and confined to the grain scale (deformation banding thickness ranges from 3-4 mean grain diameters in sandstones (Louis et al. 2006) to about 20 grain diameters in dense sands (Mühlhaus and Vardoulakis 1987)). However, their effect on the global performance of geosystems at the field scale, such as oil reservoirs, can be profound as it has been shown that these features can serve as flow barriers, reducing the effective permeability of the reservoir by orders of magnitude (Holcomb and Olsson 2003; Sternlof et al. 2006).

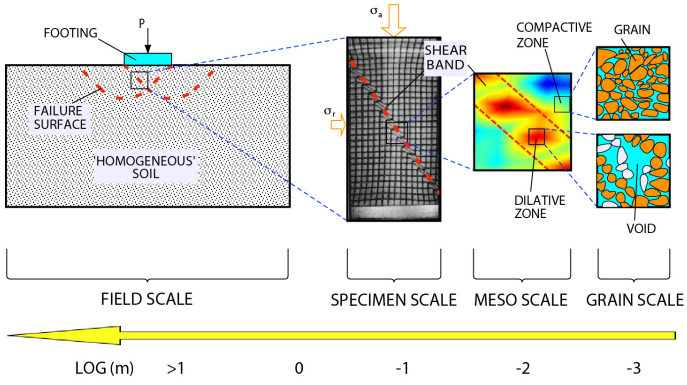

Figure 1. Illustration of the multi-scale nature of granular materials (from Andrade et al. 2008).

When characterizing fields of soil properties while accounting for uncertainty, consideration of spatial dependence is of great importance, and much effort has been devoted to this problem. Often, spa-

tially dependent random fields are modeled spectrally (e.g., Ghanem and Brzakala 1996; Popescu et al. 2005; Sudret and Der Kiureghian 2002). Another common approach is to use a correlation coefficient between the unknown values of a soil property at two points, and the correlation decreases with increasing distance between the points (Baise et al. 2006; Degroot and Baecher 1993; Goovaerts 1997). While spectral-based simulation approaches are often preferable for random field simulation due to their stability and computational tractability, here a sequential correlation-based approach is utilized. It is believed that this approach is valuable if one desires to do adaptive refinement as described below, because it is not necessary to specify a priori the locations requiring fine-scale resolution; one can simply add additional fine scale data, conditional upon all previously simulated data, as the need arises.

Spatial dependence models for soils are addressed in the literature, but are less common than results for probability distributions of soil properties at a single point (Fenton 1999; Jaksa and Fenton 2000; Uzielli et al. 2005). Characterizing spatial dependence for general cases is difficult because the dependence model is dependent upon other modeling assumptions such as whether the mean of the soil property is homogeneous. Note that a linear correlation coefficient does not in general completely describe the stochastic dependence of two random variables, except for the case of joint normal distributions, but it is often all that can be quantified and in many practical applications has been observed to be a sufficiently accurate representation of dependence (Goovaerts 1997; K.K. Phoon 2006).

Most experience with random field models for soil properties is limited to a single spatial scale. Multi-scale models for some geotechnical properties have been proposed based on theoretical arguments (Fenton 1999; Taylor and Burrough 1986) but not implemented or validated within a Monte Carlo/Finite Element Analysis framework. Challenges remain regarding computationally inexpensive approaches for implementing Monte Carlo simulation with this model and transferring the results to the mechanical model, as will be discussed below. Once spatial dependence of the random field has been defined, Monte Carlo simulations can be generated using several methods. A sequential conditional simulation approach that has been used previously by the authors (Andrade et al. 2008; Baker and Faber 2008; Chen et al. 2010) is first briefly summarized.

2 CONDITIONAL SIMULATION OF DISCRETIZED RANDOM FIELDS AT A SINGLE SCALE

To perform a sequential conditional simulation, first an arbitrary location in the grid is selected and a simulation is generated from the standard Gaussian distribution. The needed conditional distribution is easy to compute when the field is Gaussian. Let Z_n denote the random field at the next location to be simulated, and $\mathbf{Z}_p$ denote the random field at all previously simulated locations. The joint distribution of Z_n and $\mathbf{Z}_p$ is given by

$$\left[\begin{array}{c} Z_n \\ \hline \mathbf{Z}_p \end{array}\right] \sim N\left(\left[\begin{array}{c} 0 \\ \hline 0 \end{array}\right], \left[\begin{array}{c|c} 1 & \Sigma_{np} \\ \hline \Sigma_{pn} & \Sigma_{pp} \end{array}\right]\right) \quad (1)$$

where $\sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes that the vector of random variables has a joint normal distribution with mean values $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ (note that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ have been partitioned in Equation 1, to clarify the matrix operations below). The covariance matrix is dependent upon the locations of the previously simulated data points and the model for spatial dependence.

Given this model, the distribution of Z_n, conditional upon the previously simulated data points, is given by

$$\left(Z_n \mid \mathbf{Z}_p = \mathbf{z}\right) \sim N\left(\Sigma_{np} \cdot \Sigma_{pp}^{-1} \cdot \mathbf{z}, 1 - \Sigma_{np} \cdot \Sigma_{pp}^{-1} \cdot \Sigma_{pn}\right) \quad (2)$$

where $\boldsymbol{z}$ is the vector of previously simulated numerical values. A value for Z_n is simulated from this conditional distribution, and this value is then treated as a fixed data point for later simulations at other locations (i.e., Z_n is included in the vector $\mathbf{Z}_p$ of Equation 1). The conditional simulation process is repeated until all values in the field have been simulated. This approach will produce a Gaussian random field with a mean of zero and unit standard deviation, which can then be transformed to have a specified probability distribution (Goovaerts 1997; Rosenblatt 1952). Note that this conditional simulation approach is ideally suited for multiscale analysis because it is possible to selectively produce fine-scale simulations that are conditional upon an original set of coarse-scale simulations.

The above equations apply to Gaussian random fields, and most random field simulations and characterization approaches have focused on Gaussian random fields, because of intractability associated with most non-Gaussian fields. Monte Carlo simulations of non-Gaussian fields are typically obtained using a post-processing transformation of a simulated Gaussian field. This approach is widely used today, is seen to be reasonable in many situations

(Goovaerts 1997; K.K. Phoon 2006), and has been adopted by the authors in past research (Andrade et al. 2008; Baker and Faber 2008; Chen et al. 2010).

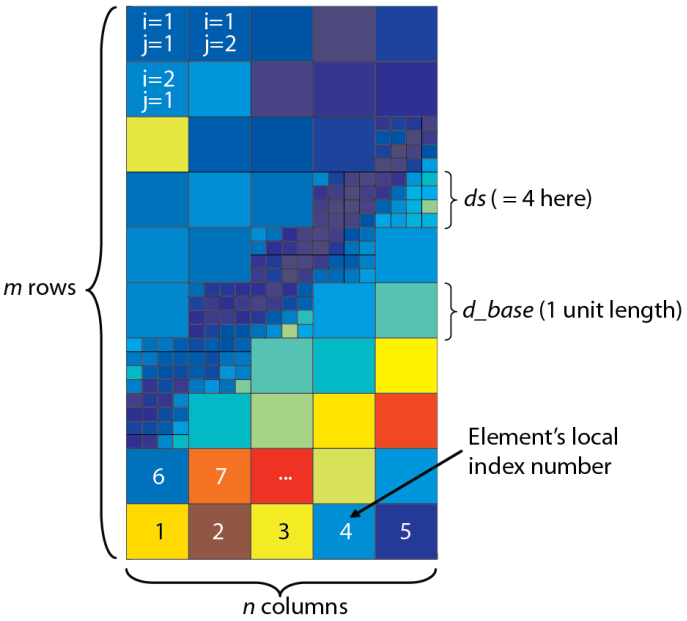

Figure 2. Geometry and notation for refined meshing of multiscale simulations.

3 MULTISCALE SIMULATIONS

When extending this procedure to simulations at multiple scales (i.e., multiple levels of spatial discretization), the challenge is to maintain spatial dependence across several scales, as illustrated schematically in Figure 2. For later finite element analysis, it is desired that the numerical value at each discretized grid point represents the average property value over the spatial extent of that grid point. Because the spatial extent of the grid points varies with level of discretization, the needed probability distributions and spatial correlation structures varies with refinement scale. The relationships between scales can be determined, however, and are documented for this approach in a previous publication by the authors (Chen et al. 2010).

Once the relevant probability distributions and correlation structures between scales have been determined, it is relatively straightforward to apply the approach of Section 2 and input the needed means and covariances appropriate for each grid point depending upon its scale. The book-keeping in the simulation software is somewhat more complex, but an example source code is provided to demonstrate the implementation of the proposed procedure (see *http://www.stanford.edu/~bakerjw/random-fields.html*).

Given that a simulation procedure is available, the challenge is then to use this approach at multiple scales in a computationally efficient manner. A brute force technique would be to simulate everywhere at the finest scale of interest and then up-sample these simulations in regions of lesser interest (i.e., regions where the simulated properties do not vary significantly, or where the field's properties will not significantly affect the calculations of interest). This approach may be feasible because the random field simulation step is generally less expensive than the finite element analysis step, so the computational burden of simulating data and then later discarding it may not be significant (Durlofsky et al. 1997; Tureyen and Caers 2004, 2005). When the scales of interest vary over several orders of magnitude, however, then the computational expense of directly simulating all properties at the fine scale, only to then discard most of them, may become significant.

An alternative that will be considered is to first produce simulated properties at the coarsest scale of interest and then iteratively simulate at finer scales, conditional on the coarse-scale random field simulations. This iterative simulation approach has recently been previously used by the authors (Chen et al. 2010). The advantage of this approach is that fine-scale simulations are produced only at locations where they are needed. For example, if the coarse scale random field simulation produces a region of weak or potentially unstable material, then additional simulations at progressively finer scales can `fill in' more detail at this critical region. Or if the finite element analysis for soil instabilities indicates that strains are localizing in a particular band (as in Figure 1), then supplemental random field simulations can be generated and the finite element mesh can be updated to incorporate the needed fine-scale details in this particular region. The above approach is well suited for this adaptive refinement, so the only further input needed is how to identify the locations where refinement would produce the most benefit for later analyses.

3.1 *Pre-specified refinement geometry*

One option, used by Chen et al. (2010) is to prespecify the geometry where high resolution is needed based on knowledge of the problem to be studied. For example, Figure 3 shows a coarse scale and refined geometry for a footing analysis problem. It is known a priori that the soil directly underneath the footing has the strongest influence on the capacity of the footing, and so that region is refined in Figure 3b.

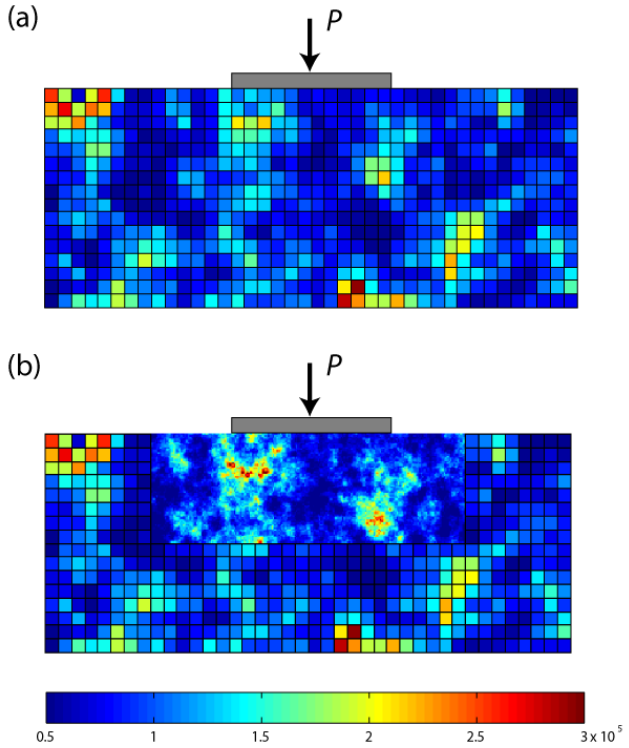

Figure 3. Refinement of a porosity simulation to improve footing reliability calculations, in the case where the region to be refined is known prior to simulation. (a) Initial coarse scale simulation. (b) Simulation with refined grid in area known to affect footing capacity (adapted from Chen et al. 2010).

3.2 *Adaptive refinement*

In some other applications, the regions that would benefit from refinement may not be known prior to simulation. In these cases, an adaptive refinement technique may be advantageous. In regions where the simulated grid takes a relatively constant value, finite element calculations will not benefit from a finer mesh. On the other hand, regions with large gradients of property values will be more accurately represented in finite element calculations if a finer mesh is used. For this reason, a potentially useful simulation technique is to generate a coarse scale field of values, and then refine regions with strong heterogeneity. Other regions, whose refinement should not affect the resulting predicted mechanical behavior significantly, can be left at the original coarse scale.

To quantify heterogeneity within a block that is a candidate for refinement, a heterogeneity index can be defined as the maximum absolute difference in property values among grid points within that block. Mathematically, an index, I, for a grid point, can be computed as

$$I = \max_{\substack{\delta=-1,0,1 \\ \varepsilon=-1,0,1}} \left\{ \left| z_{i,j} - z_{i+\delta,j+\varepsilon} \right| \right\} \qquad (3)$$

where $z_{i,j}$ is the property value at grid location i,j and the location indices i and j are as denoted in Figure 2. Grid points can be refined if $I>threshold$, where the *threshold* value can be varied depending upon the heterogeneity level of interest or the fraction of grid points that are desired to be refined. To illustrate, Figure 4 shows a coarse scale simulation, and Figure 5, Figure 6 and Figure 7 show simulations refined using the criterion of equation 3 with the heterogeneity threshold level varied so that 10%, 20% and 30% of the grid points are refined.

Figure 4. Original coarse scale simulation.

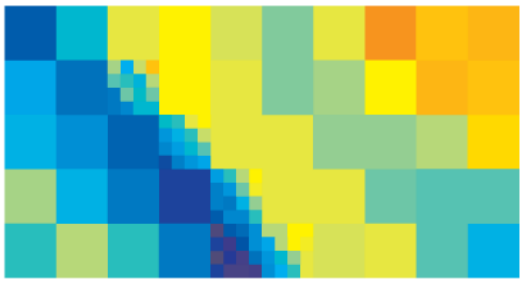

Figure 5. Simulation with 10% of grid points refined (maximum difference criterion).

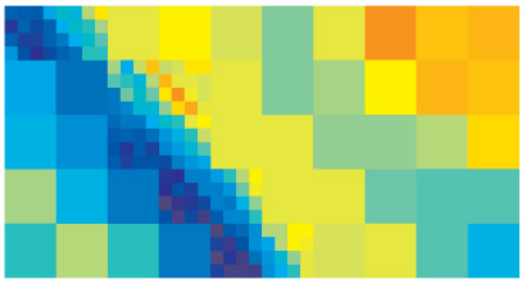

Figure 6. Simulation with 20% of grid points refined (maximum difference criterion).

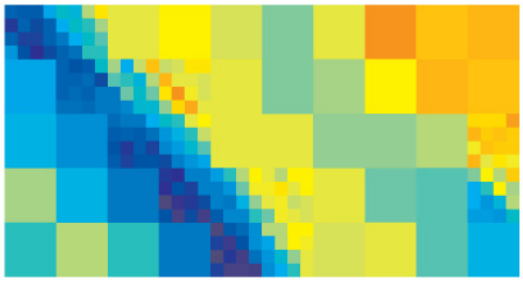

Figure 7. Simulation with 30% of grid points refined (maximum difference criterion).

An alternative metric to quantify heterogeneity is the sample variance of the grid points within a block. This can be computed as

$$I = \frac{1}{9} \sum_{\substack{\delta=-1,0,1 \\ \varepsilon=-1,0,1}} \left(z_{i+\delta,j+\varepsilon} - \bar{z} \right)^2 \quad (4)$$

where

$$\bar{z} = \frac{1}{9} \sum_{\substack{\delta=-1,0,1 \\ \varepsilon=-1,0,1}} z_{i+\delta,j+\varepsilon} \quad (5)$$

Figure 8, Figure 9 and Figure 10 show simulations with 10%, 20% and 30% of the grid points refined, respectively, based on identifying the points with the largest inhomogeneities using equation 4. If more broad scale variation is of interest, the variance over more than just adjacent grid points can be computed as follows

$$I = \frac{1}{n} \sum_{\substack{\delta=-k,..,k \\ \varepsilon=-k,..,k}} \left(z_{i+\delta,j+\varepsilon} - \bar{z} \right)^2 \quad (6)$$

where k is an integer specifying the size of the block to be considered, n is the number of grid points in the block, and

$$\bar{z} = \frac{1}{n} \sum_{\substack{\delta=-k,0,k \\ \varepsilon=-k,0,k}} z_{i+\delta,j+\varepsilon} \quad (7)$$

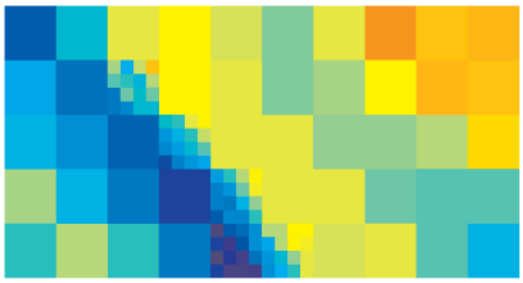

Figure 8. Simulation with 10% of grid points refined (variance criterion).

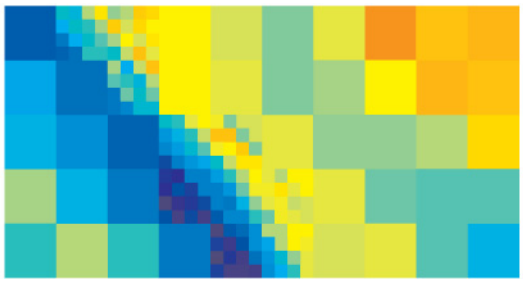

Figure 9. Simulation with 20% of grid points refined (variance criterion).

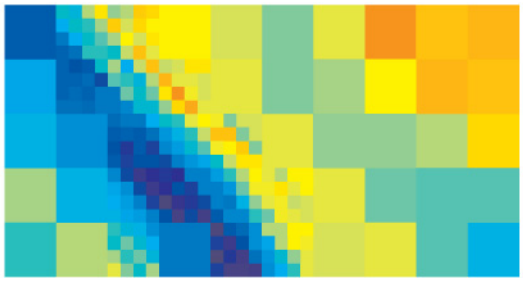

Figure 10. Simulation with 30% of grid points refined (variance criterion).

The above figures show the refinement results for a relatively small field, but the conditional simulation and adaptive refinement procedure scales relatively easily to larger fields. Figure 11 shows a simulation of a 30x20 field with 30% of grid points refined using the variance criterion This simulation took twenty minutes on a standard desktop computer using a Matlab implementation, but the computational expense could be reduced through further optimization of the code. The source code for the current implementation is provided at *http://www.stanford.edu/~bakerjw/random-fields.html*.

When performed over several spatial scales, the computational expense associated with computing heterogeneity scores for all potential blocks can be substantial. It is expected that this can be reduced if needed using approximate search/optimization approaches, and statistical prediction of heterogeneity scores using simplified results such as the heterogeneity score from a subsample of grid points within a block.

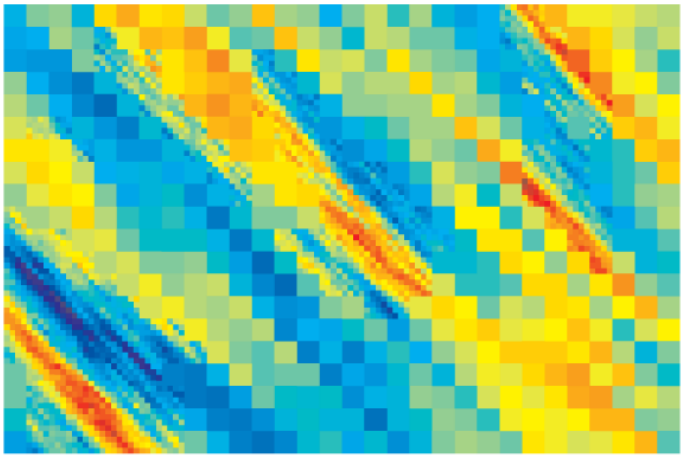

Figure 11. Adaptively refined grid at a slightly larger scale.

The above figures and equations illustrated heterogeneity calculations for a single property, to schematically demonstrate the calculations for a simple example. The same procedures can be performed if the geomechanics problem of interest is dependent upon a vector of spatially varying proper-

ties (such as permeability and strength). If more than one property is to be considered, then a weighted sum of differences or variances across each element in the vector can be used as the heterogeneity metric (Tureyen and Caers 2005).

Finally, it should also be noted that the above adaptive refinement approach is easily implemented into the finite element software developed by the authors (Chen et al. 2010). Another refinement technique considered by the authors was to refine the simulations after an initial finite element calculation was performed, and to refine in regions of stress, strain or deformation concentrations. This approach is still under consideration by the authors, due to the challenges involved in implementation.

4 CONCLUSIONS

This manuscript outlined a modeling approach used to simulate spatially correlated fields of random variables, and refine the discretization adaptively (i.e., without pre-specifying the refinement locations prior to beginning the simulations). The procedure relies on transforming the random variables of interest to Gaussian random variables, and then relying on the properties of Gaussian fields to simulate additional values conditional upon previously specified variable values. The procedure has been implemented for the analysis of geomechanics problems, but is applicable to other problems where spatially varying random variables are used as inputs. The software used to produce these results is easily adaptable for other geometries and refinement strategies, and the source code has been provided on the authors' website (*http://www.stanford.edu/~bakerjw/random-fields.html*) for use by other researchers.

5 ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation under NSF grant number CMMI 0727121 to Stanford University and grant number CMMI 0726908 to Northwestern University. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

6 REFERENCES

Adler, R. J. (1981). *The Geometry of Random Fields*. J. Wiley, Chichester; New York.

Andrade, J. E., Baker, J. W., and Ellison, K. C. (2008). "Random porosity fields and their influence on the stability of granular media." *International Journal for Numerical and Analytical Methods in Geomechanics*, 32(10), 1147-1172.

Baise, L. G., Higgins, R. B., and Brankman, C. M. (2006). "Liquefaction Hazard Mapping---Statistical and Spatial Characterization of Susceptible Units." *Journal of Geotechnical and Geoenvironmental Engineering*, 132(6), 705-715.

Baker, J. W., and Faber, M. (2008). "Liquefaction risk assessment using geostatistics to account for soil spatial variability." *Journal of Geotechnical and Geoenvironmental Engineering*, 134(1), 14-23.

Chen, Q., Seifried, A., Andrade, J. E., and Baker, J. W. (2010). "Characterization of random fields and their impact on the mechanics of geosystems at multiple scales." *International Journal for Numerical and Analytical Methods in Geomechanics (in review)*, ((in review)).

Degroot, D., and Baecher, G. B. (1993). "Estimating autocovariance of in-situ soil properties." *Journal of Geotechnical Engineering*, 119(GT1), 147-166.

Desrues, J., and Viggiani, G. (2004). "Strain localization in sand: an overview of the experimental results obtained in Grenoble using stereophotogrammetry." *International Journal for Numerical and Analytical Methods in Geomechanics*, 28(4), 279–321.

Deutsch, C. V., and Journel, A. G. (1997). *GSLIB geostatistical software library and user's guide*. Oxford University Press, New York.

Durlofsky, Jones, R. C., and Milliken, W. J. (1997). "A nonuniform coarsening approach for the scale-up of displacement processes in heterogeneous porous media." *Advances in Water Resources*, 20(5-6), 335-347.

Fenton, G. A. (1999). "Random Field Modeling of CPT Data." *J. Geotech. and Geoenvir. Engrg.*, 125(6), 486-498.

Gentle, J. E. (2003). *Random number generation and Monte Carlo methods*. Springer-Verlag, New York.

Ghanem, R., and Brzakala, W. (1996). "Stochastic Finite-Element Analysis of Soil Layers with Random Interface." *Journal of Engineering Mechanics*, 122(4), 361-369.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. Applied geostatistics series., Oxford University Press, New York.

Holcomb, D. J., and Olsson, W. A. (2003). "Compaction localization and fluid flow." *Journal of Geophysical Research*, 108(B6), 2290.

Jaksa, M. B., and Fenton, G. A. (2000). "Random Field Modeling of CPT Data." *J. Geotech. and Geoenvir. Engrg.*, 126(12), 1212-1216.

Louis, L., Wong, T., Baud, P., and Tembe, S. (2006). "Imaging strain localization by X-ray computed tomography: discrete compaction bands in Diemelstadt sandstone." *Journal of Structural Geology*, 28(5), 762–775.

Mühlhaus, H. B., and Vardoulakis, L. (1987). "The thickness of shear bands in granular materials= L'épaisseur des bandes de cisaillement des matériaux granulaires." *Geotechnique*, 37(3), 271–283.

Phoon, K. (2006). “Modeling and Simulation of Stochastic Data.” *GeoCongress 2006*, ASCE, Atlanta, Georgia, USA, 17.

Popescu, R., Deodatis, G., and Nobahar, A. (2005). “Effects of random heterogeneity of soil properties on bearing capacity.” *Probabilistic Engineering Mechanics*, 20(4), 324-341.

Rosenblatt, M. (1952). “Remarks on a multivariate transformation.” *Annals of Mathematical Statistics*, 23, 470-472.

Spanos, P. D., and Zeldin, B. A. (1998). “Monte Carlo Treatment of Random Fields: A Broad Perspective.” *Applied Mechanics Reviews*, 51(3), 219-237.

Sternlof, K. R., Karimi-Fard, M., Pollard, D. D., and Durlofsky, L. J. (2006). “Flow and transport effects of compaction bands in sandstone at scales relevant to aquifer and reservoir management.” *Water Resources Research*, 42(7), 7425.

Sudret, B., and Der Kiureghian, A. (2002). “Comparison of finite element reliability methods.” *Probabilistic Engineering Mechanics*, 17(4), 337-348.

Taylor, C. C., and Burrough, P. A. (1986). “Multiscale sources of spatial variation in soil. III. Improved methods for fitting the nested model to one-dimensional semivariograms.” *Mathematical Geology*, 18(8), 811-821.

Tureyen, O., and Caers, J. (2004). “A parallel scheme for multi-scale data integration.” *Geostatistics Banff 2004*, 10.

Tureyen, O., and Caers, J. (2005). “A parallel, multiscale approach to reservoir modeling.” *Computational Geosciences*, V9(2), 75-98.

Uzielli, M., Vannuchi, G., and Phoon, K. (2005). “Random field characterization of stress-normalized cone penetration testing parameters.” *Geotechnique*, 55(1), 3-20.

Vanmarcke, E. (1983). *Random fields, analysis and synthesis*. MIT Press, Cambridge, Mass.