Liquefaction Risk Assessment Using Geostatistics to account for Soil Spatial Variability

Jack W. Baker, M.ASCE[1]; and Michael H. Faber[2]

Abstract: Liquefaction triggering assessments are often performed for individual locations, providing little information in regard to the expected spatial extent of liquefaction events. The present paper proposes a method to quantify the potential extent of liquefaction by accounting for spatial dependence of soil properties and potential future earthquake shaking. Random-field theory and geostatistics tools are used to model soil properties and earthquake shaking intensity; this approach facilitates incorporation of measurement results obtained at individual locations within the area of interest. An empirical liquefaction triggering criterion is then used to model liquefaction occurrence as a function of the random-field realizations. The framework components are briefly described and an example analysis is performed to illustrate the details of the approach. The area of liquefied soil under a building in Adapazari, Turkey, is considered in the example, conditional upon soil property measurements obtained from nearby standard penetration tests.

DOI: 10.1061/(ASCE)1090-0241(2008)134:1(14)

CE Database subject headings: Liquefaction; Statistics; Earthquakes; Soil properties.

Introduction

Empirical models for the assessment of soil liquefaction potential are based on soil properties at individual locations. To assess consequences of liquefaction as they relate to civil structures, however, it would be helpful to also understand the potential spatial extent of liquefaction events. This requires the spatial dependence of soil properties and ground motion intensity to be quantified and incorporated in an analysis framework. Dependencies within and among these properties affect the potential extent of liquefaction, and they also inform the analyst as to the uncertainty in soil properties at points near sampled locations where soil properties are understood in greater detail.

Tools developed in the field of geostatistics are applicable for incorporating spatial dependencies. This field has undergone significant development in the past few decades in fields such as mining, petroleum engineering, and hydrology, where there is a need to infer the spatial dependence of underground phenomena from a limited number of samples. The use of this approach for liquefaction problems approach has received some recent attention (e.g., Baise et al. 2006), but its interface with other tools such as seismic hazard analysis is considered more closely here.

Evaluation of liquefaction risk, even at individual locations, requires the use of several engineering models. Empirically developed criteria provide models for the probability of liquefaction occurring at a site for given values of the relevant soil properties and ground motion shaking intensity. Probability distributions for the needed soil properties can be obtained from published studies, measurements obtained at the site, and expert judgment by qualified geotechnical engineers. The probability distribution for ground motion intensity from potential future earthquakes can be obtained using probabilistic seismic hazard analysis.

The combination of these various models to assess the potential spatial extent of liquefaction is explained in this paper. A framework is described that combines available geotechnical, geostatistical, and seismic hazard models to produce informative assessments of potential liquefaction extent. Challenges for implementation lie primarily with obtaining appropriate characterizations of soil properties, rather than with the required computations. The approach promises to help analysts better understand liquefaction risks at a site, as well as increase the amount of insight provided by sample data.

Approach

Assessment of liquefaction requires models from geotechnical engineering and seismology, and accounting for spatial dependence requires additional tools from geostatistics. By incorporating all of the needed models, as illustrated schematically in Fig. 1, a framework can be used to consider spatial distribution of occurrence. First, random fields are established to represent the soil properties and ground motion parameters of interest in a given area. A liquefaction triggering criterion is then evaluated at each individual location, based on the model parameters specified by the random fields. Finally, a probabilistic assessment of the liquefaction realizations may be performed using the field models and liquefaction criterion. The individual components of this framework are described briefly in this section.

[1]Assistant Professor, Dept. of Civil and Environmental Engineering, Stanford Univ., 240 Terman Engineering Center, Stanford, CA 94305; formerly, Research Associate, ETH Zurich. E-mail: bakerjw@stanford.edu

[2]Professor, Chair of Risk and Safety, Swiss Federal Institute of Technology, ETH Zurich; Wolfgang-Pauli-Strauss 15, HIL E 23.3, CH-8093 Zurich, Switzerland.

Note. Discussion open until June 1, 2008. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on July 17, 2006; approved on April 20, 2007. This paper is part of the ***Journal of Geotechnical and Geoenvironmental Engineering***, Vol. 134, No. 1, January 1, 2008. ©ASCE, ISSN 1090-0241/2008/1-14–23/$25.00.

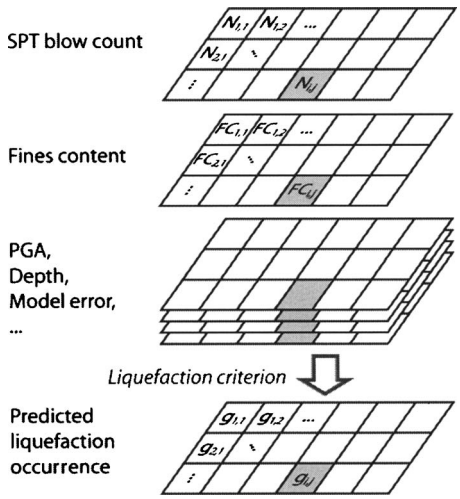

Fig. 1. Schematic illustration of the procedure for evaluating the extent of soil liquefaction as a function of random fields of soil properties and other model parameters

Modeling of Liquefaction Occurrence

The method chosen used to model liquefaction will determine which soil properties must be considered, so it is useful to consider this model component first. The most common approach for modeling liquefaction occurrence in practice uses empirical criteria that relate measured soil parameters and observed occurrence (or nonoccurrence) of liquefaction during past earthquakes. These criteria can be based on a variety of in situ soil properties obtained using various sampling methods. The standard penetration test (SPT) is the most common testing method, and data obtained using this method are often used for modeling (e.g., Cetin et al. 2004; Youd et al. 2001). Alternative models are based on data obtained from the cone penetration test (CPT), site shear wave velocity, or other testing methods (see, e.g., Kramer 1996). The framework proposed in this paper should be used with probabilistic liquefaction criteria, explicitly including model uncertainty, rather than deterministic factor-of-safety criteria, which neglect model uncertainty and often include an unquantified level of conservatism.

Some liquefaction evaluations are based on finite-element method approaches that attempt to model the physical phenomenon of liquefaction, rather than relying on empirical field observations. This approach can potentially capture more complex effects such as postliquefaction soil behavior and the interaction between liquefying soil and nearby structures. Models of this type have been used in other studies that consider seismic response of random-field soil models (e.g., Fenton and Vanmarcke 1998; Popescu et al. 2005), and they could be incorporated in the approach described here if desired.

Probability Distributions for Soil Properties

Liquefaction triggering criteria based on SPT data, typically, require knowledge of soil penetration resistance, stress conditions, fines content, and soil shear wave velocity. For criteria-based on the cone penetration test, details differ but the general approach is the same. Several past studies have provided guidance regarding appropriate probability distributions of these soil parameters, for various soil types (e.g., Fenton 1999a,b; JCSS 2002; Jones et al. 2002; Phoon and Kulhawy 1999a,b). Additional guidance may be available from judgment by geologists and geotechnical engineers familiar with the site, and from testing results at the site of interest. Guidance may sometimes come in the form of a range of possible values for the properties, and those ranges will need to be transformed into probability distributions for the approach used here. Empirical criteria are generally based on the soil values at only the critical (i.e., most liquefaction susceptible) layer of the soil, so that only a single value is needed for each surface location at the site.

Spatial Dependence of Soil Properties

Spatial dependence is used to quantify the relationship between soil properties at multiple site locations. This knowledge is used to answer the question, "given knowledge of soil properties at one boring, how much is the uncertainty at other locations reduced?" Typically, dependence is modeled using 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). It should be noted that in many cases, the correlation coefficient is actually computed for transformed data rather than the original data, as will be explained below. Spatial dependence models are addressed in the literature, but they are less common than results for probability distributions of soil properties at a single point (Fenton 1999b; Jaksa and Fenton 2000; Uzielli et al. 2005). Characterizing spatial dependence for general cases is difficult because the spatial dependence model is dependent upon other modeling assumptions such as whether the soil properties are homogeneous. Dependence can be estimated using empirical data, so datasets containing a large number of soil borings are helpful for developing this component of the framework. It is important to note that a linear correlation coefficient does not 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; Phoon 2006).

Measured Soil Properties

On-site tests are a part of many liquefaction assessments. The results obtained using methods such as the SPT, CPT, and spectral analysis of surface waves (SASW) help to identify the geology of the site, which guides the selection of appropriate soil probability distributions. But the observations also provide more precise information about soil properties at the specific locations that were sampled and, due to the spatial dependence, increased information about soil properties at nearby locations. The ability to incorporate this information is an important feature of the proposed framework.

Ground Motion Intensity

Soil properties quantify the resistance of a site to liquefaction, but the loads applied to the site also affect the potential occurrence of liquefaction. For assessment of the probability of a site liquefying, it is necessary to compute the rates of occurrence of all ground motion intensities of interest. For many empirical liquefaction occurrence models, ground motion intensity is measured using a combination of peak ground acceleration (PGA) and the earthquake's magnitude (M), which is a proxy for the ground motion duration.

Information regarding recurrence of PGA and M can be obtained using probabilistic seismic hazard analysis (PSHA), along with results from deaggregation (Kramer 1996; McGuire 2004). Deaggregated PSHA provides rates of exceedance of given levels of PGA along with the conditional distribution of causal earthquake magnitudes associated with exceedance of each PGA level—this can be converted into joint rates of occurrence of discretized PGA and M values (Bazzurro 1998, p. 195). Standard PSHA provides the needed joint density for only a single location. Ground motion intensities can with reasonable accuracy be assumed to be perfectly dependent over scales of a few hundred meters, but for regional assessments over scales of kilometers, spatial variation will need to be considered (e.g., Wang and Takada 2005).

Spatial Distribution of Liquefaction Occurrence

Liquefaction criteria can often be formulated as limit state functions referring to soil properties at a single location [see, e.g., Eq. (1) below]. Thus, at each location, it is possible to compute a probability of liquefaction. The ground motion intensity and soil properties at nearby locations are likely to be dependent, however, and so it may also be of interest to quantify the regions of potential liquefaction. The vector $\mathbf{u}$ designates the coordinates of a location in the site, and at each location $\mathbf{u}$, the limit state function $g(\mathbf{X},\mathbf{u})$ takes a value, depending upon the values of the model parameters $\mathbf{X}$ at that location (one such function will be discussed in detail below). Because the model parameter values such as soil properties are spatially dependent, the values of $g(\mathbf{X},\mathbf{u})$ are also spatially dependent. Values of $g(\mathbf{X},\mathbf{u})$ less than zero indicate liquefaction, and so regions where $g(\mathbf{X},\mathbf{u})$ is less than zero are regions where liquefaction will occur. Because the model parameters, and thus $g(\mathbf{X},\mathbf{u})$, are not known with certainty, they are best expressed as random variables. The spatial distribution of these random variables can be considered using random-field techniques (Fenton 1990; Vanmarcke 1983). In this view, regions of liquefaction correspond to excursions of random fields. It should be noted that while this formulation is compatible with the use of empirical liquefaction criteria, it does not recognize that liquefaction occurrence at one location has a causal effect on liquefaction triggering at adjacent locations; this challenge will be addressed in more detail in the Discussion section below.

Two primary approaches are available for considering excursions of random fields. The first approach uses analytical formulas obtained using random-field theory (Adler 1981; Faber 1989; Vanmarcke 1983). Under certain conditions, closed-form equations for some properties of $g(\mathbf{X},\mathbf{u})$ can be obtained. Nearly all results have been obtained for cases where $g(\mathbf{X},\mathbf{u})$ is a Gaussian field (i.e., all sets of points in the field are defined by joint Gaussian distributions), or a type closely related to a Gaussian field. Some stationarity restrictions are generally also required. Assuming these requirements are met, results can be obtained for the expected number of excursions in a region and the expected area of each excursion. The probability of an excursion at a site can also be computed, although this is a limiting value for low excursion probabilities (Adler and Taylor 2007). Analytical solutions are appealing because of their ability to explicitly provide a relationship between model parameters and computed values, but the required assumptions are often not met for typical liquefaction assessment problems. Further, this approach does not easily allow for incorporation of observed values at individual locations obtained by soil sampling, which is often an important part of practical liquefaction susceptibility evaluations.

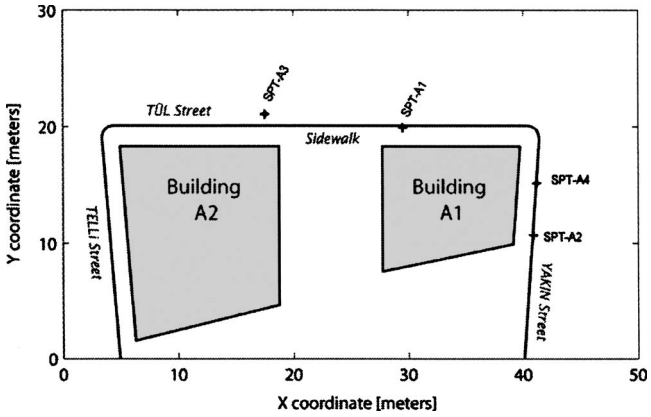

Fig. 2. Example site in Adapazari (adapted from Bray et al. 2000)

The second approach for incorporating spatial variability is to use Monte Carlo methods to simulate potential realizations of the random-field $g(\mathbf{X},\mathbf{u})$ that are consistent with the random-field characterizations of all input variables. The field of geostatistics considers approaches of this type (Goovaerts 1997). This approach is not limited by the assumptions required to obtain analytical solutions, and it has the important advantage (for this application) of incorporating observed data values at sampled locations. No analytical solutions for excursion properties are available with this approach, and the computational expense may be significant in some cases. Efficient simulation algorithms are available, however, and computational expense associated with generating the simulations is often much less than the expense associated with analyzing the results, as is the case here (Deutsch and Journel 1997). Note that Monte Carlo simulations of random fields are often performed in the frequency domain. Frequency-domain approaches have some desirable features, but are not able to incorporate observed values and so are not considered further here.

Example Application and Algorithmic Details

To illustrate the approach and provide some algorithmic details, the probabilistic modeling of liquefaction at an example site in Adapazari, Turkey, is considered. Extensive postearthquake investigations following the 1999 magnitude 7.4 Kocaeli earthquake have produced a large set of soil borings that are now available for analysis (Bray et al. 2004). SPT logs and other relevant information are available on a dedicated website (Bray et al. 2000, http://peer.berkeley.edu/publications/turkey/adapazari). A small individual site in this city, shown in Fig. 2, is considered here. Of particular interest is the probability that a specified fraction of the area under one of the buildings will liquefy during a future earthquake. Results from four nearby SPT tests are available and are used to constrain the uncertainty in soil properties at other locations within the site.

Note that while the results below use measured data when possible to illustrate the interface with real engineering data, several important model parameters were unavailable and thus assumed. Further, the complete distribution of potential ground motion intensities will be considered for the risk analysis computation. Therefore, the calculations should be interpreted as a demonstration of the relevant calculations, rather than a validation or case-study exercise.

Soil Properties and Liquefaction Criterion

The SPT-based empirical liquefaction criterion proposed by Cetin et al. (2004) is used for this example. It can be expressed using the following limit state function:

$$g(\mathbf{X},\mathbf{u}) = g(N_{1,60},\mathrm{CSR}_{\mathrm{eq}},M_w,\mathrm{FC},\sigma_v',\varepsilon) = N_{1,60}(1+0.004\mathrm{FC}) - 13.32\ln\mathrm{CSR}_{\mathrm{eq}} - 29.53\ln M - 3.70\ln\frac{\sigma_v'}{P_a} + 0.05\mathrm{FC} + 16.85 + \varepsilon_L \quad (1)$$

where $N_{1,60}$=corrected SPT blow count; $\mathrm{CSR}_{\mathrm{eq}}$=equivalent cyclic stress ratio; M=moment magnitude of the earthquake; FC=fines content; σ_v'=effective vertical stress; and ε_L=random variable representing model uncertainty. In the earlier discussion, $\mathbf{X}$ was used to refer to the vector of these variables. Function values of less than zero indicate occurrence of liquefaction. Given that some or all of the model parameters will be not perfectly known (and that the exact value of ε_L is never known) it is not possible to make a deterministic prediction of liquefaction occurrence. By characterizing the uncertainty in the predictor variables, however, one can compute the probability of liquefaction.

To evaluate Eq. (1), a few further relationships are needed. The calculation for equivalent cyclic stress ratio is

$$\mathrm{CSR}_{\mathrm{eq}} = 0.65\left(\frac{\mathrm{PGA}}{g}\right)\left(\frac{\sigma_v}{\sigma_v'}\right)r_d \quad (2)$$

where g=acceleration of gravity; σ_v=total vertical stress; and r_d=the nonlinear shear mass participation factor. A variety of models have been proposed for r_d, but the model of Cetin et al. (2004) is adopted here, as it was used in the development of the above criterion

$$r_d = \frac{\left[1+\frac{-23.013+2.949a_{\max}+0.999M_w+0.0525V_{s,12\ \mathrm{m}}^*}{16.258+0.201e^{0.341(-d+0.0785V_{s,12\ \mathrm{m}}^*+7.586)}}\right]}{\left[1+\frac{-23.013+2.949a_{\max}+0.999M_w+0.0525V_{s,12\ \mathrm{m}}^*}{16.258+0.201e^{0.341(0.0785V_{s,12\ \mathrm{m}}^*+7.586)}}\right]} + \varepsilon_{r_d} \quad (3)$$

where $V_{s,12\ \mathrm{m}}^*$=representative shear wave velocity over the top 12 m at the site; and ε_{r_d}=Gaussian random variable representing model error. Eq. (3) is only valid when the critical liquefaction layer is in the top 20 m, as is always the case in this example, but Cetin et al. (2004) provide another equation for depths greater than 20 m.

The evaluation of liquefaction triggering for a single point thus requires the evaluation of Eqs. (1)–(3), and the input parameters in these equations are modeled as random fields. The parameter models have been chosen for this application based on empirical observations at and around the site, and literature guidance where appropriate. To simplify the illustration here, a single soil layer is assumed to be the critical liquefaction susceptible layer for the entire site.

Cetin et al. (2004) specify the following distributions for the model errors: ε_L has a Gaussian distribution with a mean of zero and a standard deviation of 2.7, and ε_{r_d} is Gaussian with zero mean and standard deviation equal to

$$\sigma_{\varepsilon_{r_d}} = \begin{cases} d^{0.85}\cdot 0.0198 & \text{if } d \le 12\ \mathrm{m} \\ 12^{0.85}\cdot 0.0198 & \text{if } d > 12\ \mathrm{m} \end{cases} \quad (4)$$

where d=depth of the critical liquefaction susceptible layer.

Other random variables for soil properties were defined as follows: the distribution of $N_{1,60}$ was defined by an empirical distribution of 312 measured values from throughout the city (it had a mean of 6.5, a standard deviation of 5.6, and a strong skew to the right). Note that homogeneity has been assumed in order to estimate the probability distribution from this sample of measurements. Average shear wave velocity, $V_{s,12\ \mathrm{m}}^*$, was deterministically defined as 150 m/s, on the basis of spectral analysis of surface waves data at the site (Bray et al. 2004). Fines content was modeled as beta distributed with parameters a=2.9 and b=7.3 estimated from a maximum likelihood fit to 33 measured values from throughout the city.

For the purposes of illustration, all random variables except $N_{1,60}$ were assumed to be perfectly dependent at the spatial scale of Fig. 2 (i.e., within one simulation, the variables take the same value at each location in the example site). The spatial dependence of $N_{1,60}$ was treated more rigorously, as explained in the following section.

Spatial Dependence of Soil Properties

The mean value of $N_{1,60}$ is assumed to be known and constant throughout the study site. This assumption, believed to be reasonable here, allows for the use of so-called simple Kriging when developing estimates of the joint probability distributions at multiple points in the site. Other Kriging approaches allow for local variations in mean values, or for mean values that vary smoothly over the study site, and can also be utilized within this framework if deemed appropriate.

The stochastic dependence between soil properties at any two points is modeled using a covariance function. For Gaussian data, this fully describes the joint dependence between properties at two points, and the analytical equations are very tractable. Soil properties do not generally have Gaussian distributions, however. To take advantage of the desirable properties of multivariate Gaussian models, the data of interest are transformed using a normal-score mapping. This transform requires the complimentary distribution function (CDF) of the true soil values, which is obtained using either an empirical CDF from observed data values or the CDF corresponding to an appropriate parametric distribution. Each potential value of the soil property is then mapped to a value such that the CDF of the original soil property has the same fractile value as the transformed value does with regard to a standard Gaussian CDF (Goovaerts 1997; Phoon 2006; Rosenblatt 1952). This is expressed mathematically by

$$z = \Phi^{-1}[F(y)] \quad (5)$$

where y=original data from a distribution represented by the CDF $F(y)$; $\Phi^{-1}(\cdot)$=inverse of the standard Gaussian CDF; and z=transformed data. This transformation by definition produces variables that marginally have a standard Gaussian distribution. After verifying that the transformed data are reasonably represented by a multivariate Gaussian distribution, these transformed data are used for statistical estimation and simulation.

The normal-score transformed data are used to estimate spatial dependence, using an empirical semivariogram (Goovaerts 1997). The semivariogram, denoted $\gamma(\mathbf{h})$ is equal to half of the variance of the increment in data points separated by a distance $\mathbf{h}$

$$\gamma(\mathbf{h}) = \tfrac{1}{2}\mathrm{Var}[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})] \quad (6)$$

where $Z(\mathbf{u})$ = distribution of the (normal-score-transformed) random variable at location $\mathbf{u}$. The vector distance $\mathbf{h}$ accounts for both length and direction. Because the models in this application consider only variation of critical values in a horizontal plane, and not variation with depth, isotropy is assumed for this application, so that the semivariogram is a function of separation length only, but if appropriate the semivariogram can be a function of orientation as well. This semivariogram is often used in geostatistics instead of a covariance, because it requires second-order stationarity of only the increments and not the underlying process, but the two can be used interchangeably in nearly all applications. Here, a semivariogram function of the following form was chosen for the (transformed) corrected SPT blow count, $N_{1,60}$:

$$\gamma(h) = \begin{cases} 1.5\left(\frac{h}{c}\right) - 0.5\left(\frac{h}{c}\right)^3 & \text{if } h \leq c \\ 1 & \text{if } h > c \end{cases} \quad (7)$$

where h = (directionally independent) scalar separation length; and c = so-called range parameter that indicates the scale at which spatial dependence is significant. This functional form is referred to as a spherical model. Basic tools for empirical semivariogram analysis are available in many GIS software packages, as well as stand-alone geostatistics tools (e.g., SCRF 2006). In the following results, c values of 10, 25, and 50 m are considered, to show the effect of spatial dependence on calculated results. While characterization of spatial correlation remains a challenge, due to limited data and limited reporting in the scientific literature, these c values are in the range consistent with empirical results obtained from the 312 sampled sites in the region.

Once the spatial dependence has been defined, realizations of $N_{1,60}$ can be generated using a sequential simulation approach. Consider the example site from Fig. 2. Soil properties are desired for a discretized field of 120×200 elements, each with dimension 25×25 cm. Simulations of $N_{1,60}$ values for these 24,000 elements are needed that are consistent with the observed marginal distributions and the spatial dependence of this property, as well as the four observed values.

The sequential simulation approach is appealing because it is easy to generate samples for a single element, conditional upon the values of sampled elements at surrounding locations. Realizations of the $N_{1,60}$ values at the site of interest can be generated using a series of successive conditional simulations (Goovaerts 1997). First, the conditional distribution at an arbitrary unsampled location, $\mathbf{u}^{(1)}$ is determined, conditional upon values of the originally sampled data points. Because the data have been normal-score transformed, this conditional distribution is easy to compute. The joint distribution of $Z^{(1)}$ and the values of the sampled data points are given by

$$\begin{bmatrix} Z^{(1)} \\ \mathbf{Z}^{(\cdot)} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ \mathbf{0} \end{bmatrix}, \left[\begin{array}{c|c} 1 & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right] \right) \quad (8)$$

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 Eq. (8) to clarify the matrix operations below). The vector $\mathbf{Z}^{(\cdot)}$ represents the original (transformed) data values at the sampled locations; and $\mathbf{0}$ is a vector of zeros having the same size as $\mathbf{Z}^{(\cdot)}$. The covariance matrix is dependent upon the locations of the original and simulated data points. Each element of the matrix can be computed by evaluating Eq. (7), noting that the covariance between locations with a separation distance h is equal to $1-\gamma(h)$. Note that all means are equal to zeros and all variances are equal to 1 because of the normal-score transform.

Given this model for the joint distribution, the distribution of Z_1 conditional upon the original data points is given by

$$(Z^{(1)} | \mathbf{Z}^{(\cdot)} = \mathbf{z}) \sim N(\boldsymbol{\Sigma}_{12} \cdot \boldsymbol{\Sigma}_{22}^{-1} \cdot \mathbf{z},\ 1 - \boldsymbol{\Sigma}_{12} \cdot \boldsymbol{\Sigma}_{22}^{-1} \cdot \boldsymbol{\Sigma}_{21}) \quad (9)$$

where $\mathbf{z}$ = vector of $N_{1,60}$ values at the sampled locations. Note that $\mathbf{Z}^{(\cdot)}$ = random variable representing the model for uncertain soil parameters prior to sampling, and sampling has revealed their values $\mathbf{z}$. A value for $Z^{(1)}$ 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., $Z^{(1)}$ is included in the vector $\mathbf{Z}^{(\cdot)}$ of Eq. (9) for the subsequent simulations of $Z^{(2)}, Z^{(3)}, \ldots$].

The process is repeated at each location in the region, and at each location a conditional distribution is computed based on the values of the original data points plus the previously simulated data points. Once values have been simulated for all locations, the resulting field can be transformed back to the original probability distribution by inverting Eq. (5) for each z value. The resulting set of values represents one realization of the soil properties at the site of interest. The simulated field will always agree with observed values at sampled locations, and at other locations it will be consistent with the specified stochastic properties of the field.

Many generalizations of this basic approach have been developed (e.g., Deutsch and Journel 1997; Goovaerts 1997). One important extension is the simulation of vector-valued random fields. A procedure similar to the one above is used, but now at each location in the site a vector containing each parameter of interest is simulated, conditional upon the values of parameters previously observed or simulated. The approach is not used here, but it will be incorporated in future work using this framework.

A few issues relating to the practical implementation of this approach should be noted briefly. Mathematically, the order of locations at which values are simulated is not important, but usually a unique order is used for each sample to ensure sufficient variety among a finite number of samples. Note also that Eq. (9) implies all 23,999 previous values should be used for conditioning when simulating the last value in the field, and this requires inversion of a $\boldsymbol{\Sigma}_{22}$ covariance matrix with size $23{,}999 \times 23{,}999$. In practice, however, the distant values are "screened" because the closest values are dominant. Thus, the number of conditioning points can be significantly reduced, decreasing the computational expense of the procedure without practically affecting the resulting conditional distribution. An implementation of the algorithm, which addresses these and other practical issues, is available as part of the open-source GSLIB software package (Deutsch and Journel 1997). Example simulations of $N_{1,60}$ values generated using this approach and software package are shown in Fig. 3. In Fig. 4, the mean and standard deviation of 1,000 simulations generated using this method are shown. In Fig. 4(b), it can be seen that the standard deviation of $N_{1,60}$ values is lowest near the sampled locations, reflecting the effect of spatial dependency. In the next section, these simulations of soil properties will be used to evaluate liquefaction phenomena of engineering interest.

Probability of Liquefaction for a Given Ground Motion Intensity

To perform the evaluation of liquefaction occurrence, the procedure illustrated in Fig. 1 is used. The site of interest is divided into discrete cells. For each soil or ground motion property used in Eq. (1), a matrix of values is generated representing that vari-

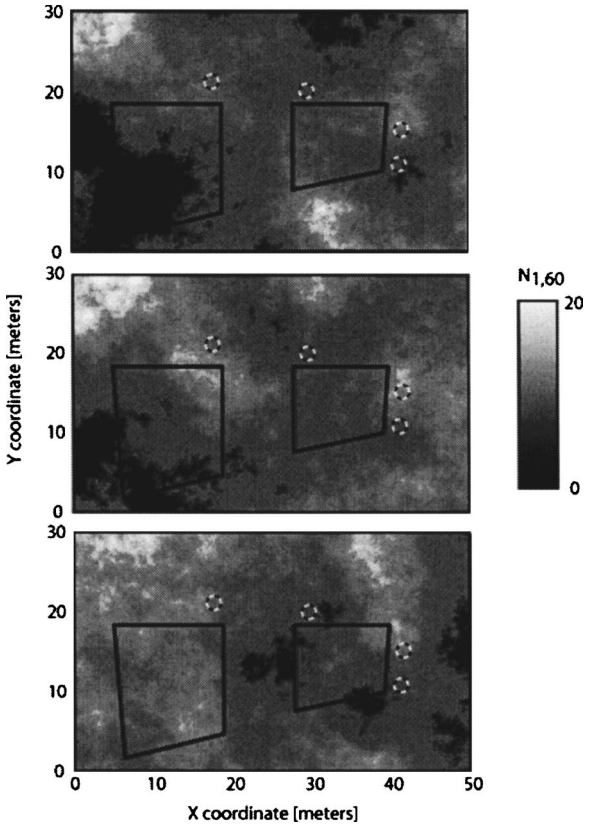

Fig. 3. Three conditional simulations of corrected SPT blow count ($N_{1,60}$) values. Locations with measured $N_{1,60}$ values are circled.

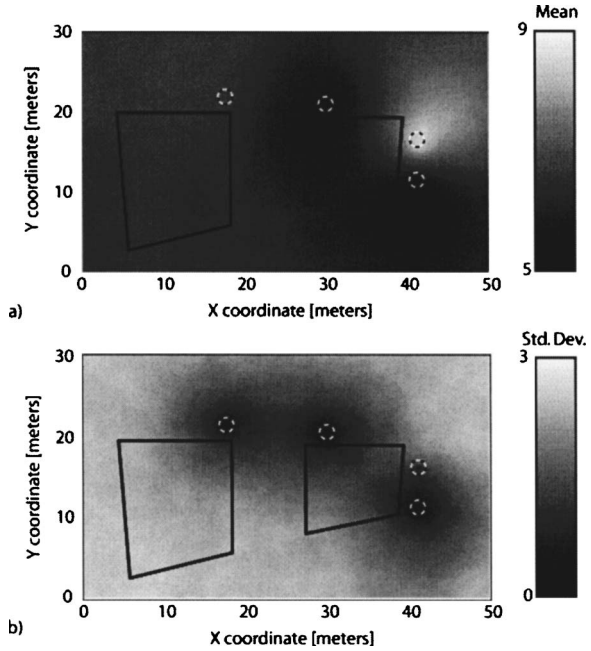

Fig. 4. (a) Mean; (b) standard deviation, of simulated $N_{1,60}$ values. Locations with measured $N_{1,60}$ values are circled.

able's value at each cell. When measurements of the property are available at some locations, then the measured values are input into the matrix, and the remaining cells are simulated conditional on those values, using the approach described above. Some variables such as earthquake magnitude will take the same value in each cell, while the values of other variables will vary. The illustrations in Fig. 3 are simply graphical displays of these matrices. Some variables can also be specified deterministically if their uncertainty is not expected to affect the results.

Once all the matrices are generated, Eq. (1) can be evaluated at each location in the site to determine whether liquefaction has occurred at the given location and for the given realization of model parameters. By repeating this simulation and evaluation procedure multiple times, the probabilistic behavior of the extent of liquefaction can be evaluated. To illustrate, example evaluations of liquefaction extent are shown in Fig. 5. To generate Fig. 5, a magnitude 7.4 earthquake with a PGA of $0.3g$ was assumed. While the magnitude was chosen to match that of the 1999 event that occurred nearby, the results are not intended to be a validation of the model for that event. The plotted examples correspond to the $N_{1,60}$ values shown in Fig. 3, but there is not a one-to-one mapping between the two pictures because of the uncertainty in the other model parameters used to generate Fig. 5.

The graphical illustrations are useful, but they must be summarized if a large number of simulations are to be efficiently considered. Any function of interest can be evaluated for each of the simulations and then summarized numerically or graphically. For example, liquefaction occurrence under Building A1 may be of interest. It is simple to define a variable Y representing the fraction of liquefied area under Building A1, and measure the liquefied area associated with each Monte Carlo simulation. This can be formulated mathematically using the following equation:

$$P(Y > y|\text{pga},m) = \int_{\mathbf{X}} I\{h[\mathbf{g}(\mathbf{X}|\text{pga},m) > y]\} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (10)$$

where $\mathbf{g}(\mathbf{X}|\text{pga},m)$=limit-state function from Eq. (1), evaluated for PGA=pga and magnitude=m, with g now written in bold to denote that it is a vector output corresponding to the limit-state values at all 24,000 site locations. The random variable $\mathbf{X}$ represents the vector of all input soil variables at each location in the site; and $f_{\mathbf{X}}(\mathbf{x})$=joint probability density function of the variables. The function $h(\cdot)$ produces a scalar output for a given realization of liquefaction extent; in this example $h(\cdot)$=fraction of liquefied area under Building A1. The distribution of values that $h(\cdot)$ takes is represented by the random variable Z, and the indicator function $I[h(\cdot) > y]$ takes value 1 if $h(\cdot) > y$, and 0 otherwise.

Evaluating Eq. (10) involves considering all possible realizations of soil properties $\mathbf{X}$ at the site, but numerical integration is not possible because $\mathbf{X}$ is very high dimensional (i.e., it contains four soil properties $\times$24,000 locations in this example). For this reason, the simulation approach described above is used to evaluate this integral. The results from this calculation are shown in Fig. 6. One hundred simulations of $N_{1,60}$ values were generated using Eq. (7) with c values of 10, 25, and 50 m, to show the effect of spatial dependence on calculated results. The calculations take only a few minutes on a desktop computer, so the computational expense for simulation should not be unreasonable for many applications.

For a c value of 25 m, there is approximately a 50% probability that some portion of the soil will liquefy, but only a 15% probability that more than half of the area under the building will liquefy. Thus, the implied reliability of the building with respect to liquefaction is dependent upon the liquefied area which is required to cause building failure. For all three values of the correlation range parameter c, the mean liquefied area from this event is approximately 10% of the building area. Small c values are

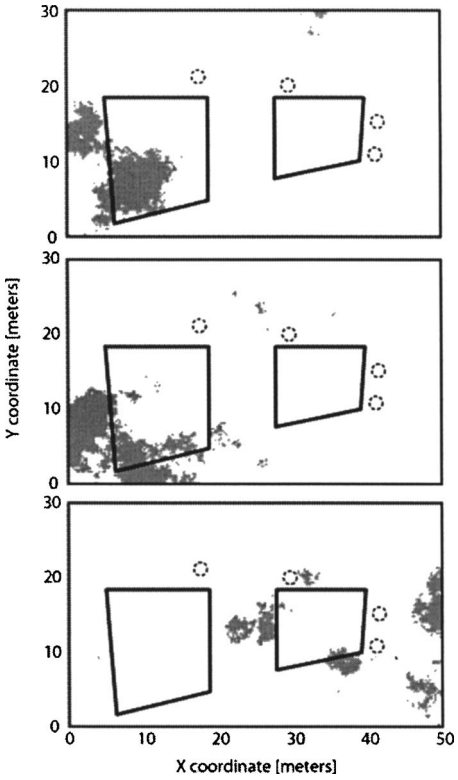

Fig. 5. Locations of liquefaction triggering, as computed from the conditional simulations of site soil properties, and given a magnitude 7.4 earthquake with a PGA of 0.3g. Liquefied regions are shaded, and locations with measured $N_{1,60}$ values are circled.

associated with a higher probability of observing small liquefied areas and a lower probability of observing large liquefied areas. These results make clear the effect of spatial dependencies on the potential extent of liquefaction.

Incorporating Multiple Ground Motion Intensity Levels

The results in Fig. 6 were obtained by inputting fixed values for earthquake magnitude and peak ground acceleration. But without knowing the rate of occurrence of this specific event (and others), it is not clear whether this performance is acceptable (Kramer et al. 2006). For this reason, ground motion hazard results are incorporated in the analysis. The mean rate density from PSHA provides the joint probability density function of magnitude and PGA, multiplied by the rate of occurrence of earthquake events. The most direct way to incorporate this information is to sample PGA and M values from this joint distribution, and use them in the Monte Carlo scheme discussed in the previous section. The difficulty with this approach is that events with small M and/or PGA values are orders of magnitude more likely to occur than the largest events, and so the great majority of simulations will be for small events. An alternative approach proposed here is to use Monte Carlo simulation for the soil properties and numerical integration for the ground motion intensity. This approach is effective because the ground motion intensity variables contribute significantly to variability in the final output, and a large set of other parameters cannot be integrated over numerically (Gentle 2003).

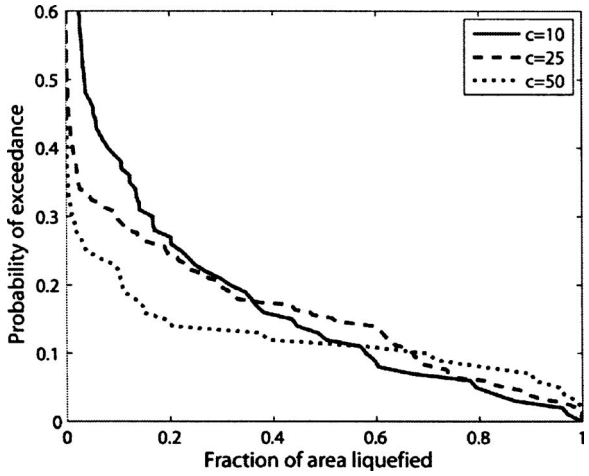

Fig. 6. Probability of exceedance versus fraction of the area under Building A1 that liquefies, given a magnitude 7.4 earthquake with a PGA of 0.3g. Results are shown for three different models of $N_{1,60}$ spatial correlation.

The first step in the proposed approach is to perform an evaluation conditional upon a given PGA and M value, as was done in the Eq. (10). To complete the evaluation, one must then consider the entire range of relevant PGA and magnitude values using the following equation:

$$\lambda(Y > y) = \int_{\text{PGA}} \int_{M} P(Y > y|\text{pga},m)\text{MRD}_{\text{PGA},M}(\text{pga},m)\text{dpga}\,\text{d}m \tag{11}$$

where $P(Y>y|\text{pga},m)$ comes from Eq. (10); and $\text{MRD}_{\text{PGA},M}(pga,m)$=mean rate density of PGA and M. The mean rate density used in this example calculation is shown in Fig. 7. It was obtained by processing the results from a site-specific hazard analysis with deaggregation (Bazzurro 1998, p. 195). By integrating over all PGA and M values using the total probability theorem, it is possible to obtain $\lambda(Y>y)$, the rate of exceedance of the variable of interest, Y, over some specified threshold value y. There are only two scalar variables to integrate over in this case, so numerical integration is feasible. Results from Eq. (11) are shown in Fig. 8. Fig. 8 was generated using the soil properties

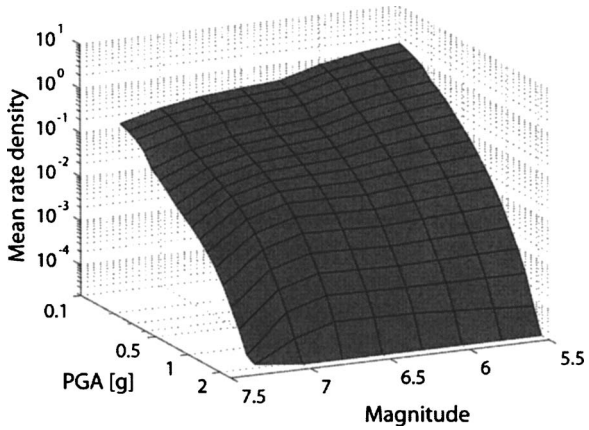

Fig. 7. Mean rate density of peak ground acceleration and earthquake magnitude, obtained from probabilistic seismic hazard analysis

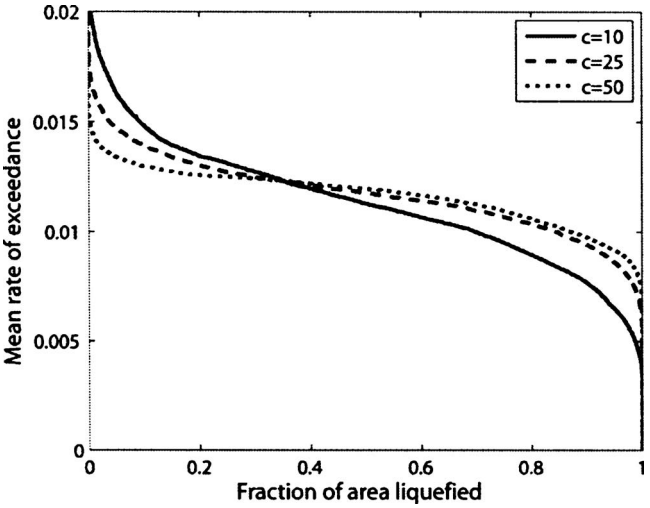

Fig. 8. Rate of exceedance versus fraction of the area under Building A1 that liquefies, considering all possible ground motion intensities as specified by the ground motion hazard. Results are shown for three different models of $N_{1,60}$ spatial correlation.

model discussed above, and using 20 PGA values and 10 M values for the numerical integration of Eq. (11).

Again, the effect of spatial correlation is seen clearly when considering three different values for the range parameter c. This parameter again affects the relative ratio of small liquefied regions and large liquefied regions. While the specific results depend upon the random variable assumptions used here for illustration, they nonetheless illustrate the potentially useful information that can be obtained using the approach. The estimated rate of exceedance of a given liquefied area can be computed explicitly, considering uncertainties and spatial dependencies in soil properties and ground motion intensity.

Discussion

The framework described can characterize the potential spatial distribution of soil liquefaction. Significant challenges remain, however, with respect to modeling assumptions and parameter estimation. The model for spatial dependence of soil properties depends solely on the correlation coefficient between normal-score transformed values. This approach has been seen to provide good results in a variety of mining and petroleum engineering applications (Goovaerts 1997), but its validity should still be verified when possible. Probabilistic properties such as ergodicity and homogeneity should also be considered when choosing a model (Phoon et al. 2003; Rackwitz 2000). The specification of which parameters are ergodic and nonergodic will be important for test planning, and nonergodicity of parameters may cause difficulties for parameter estimation.

A further challenge for random-field characterization is the presence of soil layers. When the soil is composed of several discrete layers with significantly differing properties, then modeling may become more complicated. Empirical liquefaction criteria often depend only upon soil values from the most susceptible layer. In the example calculation above, measured soil values in critical layers appeared to be well modeled by assuming that it all came from the same population, but this conclusion may not hold for other sites. Improved methods for dealing with layered, or otherwise structured, random fields are in active development (e.g., Krishnan and Journel 2003), but their applicability for this problem has yet to be determined. In particular, approaches of the type cited require a training image that provides a representation of the phenomenon being studied, and it may not be feasible to develop a reasonable training image for soil layering without performing extensive sampling. Modifications to the model used above should nonetheless be considered if they are deemed important and can be characterized.

The use of empirical liquefaction criteria with the framework, while appealing because it is simple and used often in practical evaluations, has some weaknesses. Empirical models generally claim to address only the initial triggering of liquefaction and not provide information about posttriggering behavior—although it has been suggested that some criteria may also indicate liquefaction severity (Iwasaki et al. 1978; Toprak and Holzer 2003). The physical process of interest is caused by buildup of pore water pressure due to dynamic excitation, and so liquefaction of soil at one location will generally affect the behavior of the surrounding soil. This may be an important phenomenon when modeling the spatial extent of liquefaction. Using a coarser discretization of the site may avoid the implied independence of liquefaction triggering at adjacent cells for a given set of input parameters, thus reducing the potential errors introduced by this approach. Relatively few empirical-type models are available for interactions between structures and potentially liquefiable soil (e.g., Rollins and Seed 1990), although this effect would be possible to incorporate in the above approach. Models that use finite-element analysis to model liquefaction promise to address these issues more rigorously (Fenton and Vanmarcke 1998; Koutsourelakis et al. 2002; Popescu et al. 2005), but also require more computational expense and analyst time to model the site. The writers are not aware of any finite-element studies that consider spatially variable soil models conditional upon nearby observed values. Finite-element models could be used in the above framework, after several modifications were made. First, the random fields for the soil would need three dimensions rather than the two dimensions used above; this poses no problems for the geostatistics algorithms, but estimating and specifying the needed random-field parameters will be more challenging. Second, ground motion time histories will be needed to represent the ground motion input, rather than just PGA/M values—this will make it more difficult to consider all possible ground motions, and will likely prohibit the use of the proposed stratified sampling technique to reduce computational expense. Finally, spatial variability of ground motion will require more careful consideration for small sites because incoherence of the ground motions will need to be considered, and incoherence occurs at smaller spatial scales than variability of the peak ground acceleration values needed above. It is not yet known to what extent these additional challenges will limit the adoption of this method for use with finite-element-based evaluations.

Geotechnical engineering assessments often must incorporate information from varying sources and of varying quality, and this framework is no different. In the example application, only SPT data were used, and further work will address the need to simultaneously incorporate other data sources such as nearby CPT tests. This task is challenging because the different information sources will generally not describe the same soil properties, and liquefaction criteria based on the various soil properties may not be consistent.

Finally, it is relevant to note that liquefaction is only one of several mechanisms by which earthquake ground motions can

cause an engineering system to fail. In order to make fully informed decisions for managing liquefactions risks, it would be helpful to consider a system that incorporates all potential failure mechanisms (e.g., structural collapse, liquefaction, and bearing failure due to soil–structure interaction). A number of active research fields are aiding in progress toward this goal.

Summary

A framework has been proposed for evaluating the potential spatial distribution of liquefaction occurrence, conditional upon observations of soil properties obtained from site samples. The framework incorporates several engineering models that are rarely used together. Geostatistics tools are used to model uncertain soil properties, conditional upon observed values obtained from samples at a few locations in the area of interest. Probabilistic seismic hazard analysis is used to compute the distribution of intensity of future ground motion shaking. An empirical liquefaction triggering criterion is used to model liquefaction occurrence as a function of soil properties and ground motion intensity. These model components are all available presently, although they have been developed independently and have not previously been combined in the form seen here.

The numerical procedures have been outlined to demonstrate the feasibility of the proposed approach. Existing software tools for seismic hazard analysis and geostatistics facilitate the needed computations, and allow calculations of this type to be performed without great effort. An example calculation has been performed to illustrate the details of the approach and demonstrate the type of information that can be obtained. Estimating the needed probability distributions for soil properties at a given site will likely prove the greatest challenge for implementation of this approach.

By considering treating soil properties, ground motion shaking, and liquefaction triggering probabilistically, this approach allows for a more complete evaluation of liquefaction risk than is possible using more simplified criteria. The simultaneous consideration of random (unknown) soil properties and random future earthquake shaking is an improvement over many current assessments that only consider a single level of ground motion intensity from a scenario earthquake. By accounting for a range of possible ground motions, annual rates of liquefaction can be computed, making it possible to produce design projects that have uniform levels of risk. Further, by considering spatial dependence of soil properties, the extent of liquefied area under a building can be characterized and used for decision making. With these assessment approaches, progress can be made toward considering costs and benefits explicitly when making engineering decisions regarding liquefaction risk.

Acknowledgments

This study was supported by the Managing Earthquake Risks Using Condition Indicators project of the Swiss National Science Foundation. The writers thank Yahya Bayraktarli for managing the soil properties database used for estimation of soil property probability distributions and spatial correlations, and the anonymous reviewers whose insights significantly improved this paper.

Notation

The following symbols are used in this paper:

$\mathrm{CSR}_{\mathrm{eq}}$ = equivalent cyclic stress ratio;
d = depth of the critical liquefaction susceptible layer;
$f_{\eta}(\mathbf{e})$ = joint PDF of all input soil variables at all locations;
$F(\cdot)$ = cumulative distribution function of a given soil property;
FC = soil fines content;
g = acceleration of gravity;
$g(\cdot)$ = liquefaction limit state function;
$g(\cdot,\mathbf{u})$ = liquefaction limit state function, evaluated at location $\mathbf{u}$;
$\mathbf{h}$ = vector distance between two locations;
$h(\cdot)$ = function that produces a scalar output from a given realization of liquefaction extent at the site (e.g., fraction of area liquefied);
$I(\cdot)$ = indicator function, taking value 1 if the argument is true and 0 otherwise;
M = earthquake moment magnitude;
$\mathrm{MRD}_{\mathrm{PGA},M}(\cdot)$ = mean rate density of PGA and M;
$N(\boldsymbol{\mu},\boldsymbol{\Sigma})$ = random vector having a Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$;
$N_{1,60}$ = soil SPT blow count;
PGA = peak ground acceleration;
r_d = nonlinear shear mass participation factor;
$\mathbf{u}$ = location within the site;
$V^*_{s,12\ \mathrm{m}}$ = representative shear wave velocity over the top 12 m of the soil;
$Z(\mathbf{u})$ = distribution of the normal-score-transformed soil variable at location $\mathbf{u}$;
ε_L = model uncertainty in the liquefaction triggering criterion;
ε_{r_d} = model uncertainty in the r_d prediction;
$\Phi^{-1}(\cdot)$ = inverse of the standard Gaussian cumulative distribution function;
$\gamma(\cdot)$ = semivariogram;
$\lambda(Y>y)$ = annual rate of exceedance of Y over some specified threshold y;
σ_v = soil vertical stress; and
σ'_v = soil effective vertical stress.

References

Adler, R. J. (1981). *The geometry of random fields*, Wiley, New York.
Adler, R. J., and Taylor, J. E. (2007). *Random fields and geometry*, Springer, New York.
Baise, L. G., Higgins, R. B., and Brankman, C. M. (2006). "Liquefaction hazard mapping—Statistical and spatial characterization of susceptible units." *J. Geotech. Geoenviron. Eng.*, 132(6), 705–715.
Bazzurro, P. (1998). "Probabilistic seismic demand analysis." Ph.D. thesis, Stanford University Press, Stanford, Calif.
Bray, J. D., et al. (2000). "Documenting incidents of ground failure resulting from the August 17, 1999 Kocaeli, Turkey earthquake." Pacific Earthquake Engineering Research Center, ⟨http://peer.berkeley.edu/publications/turkey/adapazari/⟩ (Dec. 28, 2006).
Bray, J. D., et al. (2004). "Subsurface characterization at ground failure sites in Adapazari, Turkey." *J. Geotech. Geoenviron. Eng.*, 130(7), 673–685.
Cetin, K. O., et al. (2004). "Standard penetration test-based probabilistic

and deterministic assessment of seismic soil liquefaction potential." *J. Geotech. Geoenviron. Eng.*, 130(12), 1314–1340.

Degroot, D. J., and Baecher, G. B. (1993). "Estimating autocovariance of in situ soil properties." *J. Geotech. Engrg.* 119(1), 147–166.

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

Faber, M. H. (1989). "Excursions of Gaussian random fields in structural reliability theory." Ph.D. thesis, Univ. of Aalborg, Aalborg, Denmark.

Fenton, G. A. (1990). "Simulation and analysis of random fields." Ph.D. thesis, Princeton University Press, Princeton, N.J.

Fenton, G. A. (1999a). "Estimation for stochastic soil models." *J. Geotech. Geoenviron. Eng.*, 125(6), 470–485.

Fenton, G. A. (1999b). "Random field modeling of CPT data." *J. Geotech. Geoenviron. Eng.*, 125(6), 486–498.

Fenton, G. A., and Vanmarcke, E. H. (1998). "Spatial variation in liquefaction risk." *Geotechnique*, 48(6), 819–831.

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

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

Iwasaki, T., Tatsuoka, F., Tokida, K., and Yasuda, S. (1978). "A practical method for assessing soil liquefaction potential based on case studies at various sites in Japan." *Proc., 2nd Int. Conf. on Microzonation*, San Francisco, 885–896.

Jaksa, M. B., and Fenton, G. A. (2000). "Discussion of Random Field Modeling of CPT Data by G. A. Fenton." *J. Geotech. Geoenviron. Eng.*, 126(12), 1212–1214.

Joint Committee on Structural Safety (JCSS). (2002). "JCSS probabilistic model code, Section 3.7: Soil properties." ⟨http://www.jcss.ethz.ch/⟩

Jones, A. L., Kramer, S. L., and Aduino, P. (2002). "Estimation of uncertainty in geotechnical properties for performance-based earthquake engineering." *PEER 2002-16*, Pacific Earthquake Engineering Research Center, Univ. of California at Berkeley, Berkeley, Calif.

Koutsourelakis, S., Prévost, J. H., and Deodatis, G. (2002). "Risk assessment of an interacting structure-soil system due to liquefaction." *Earthquake Eng. Struct. Dyn.*, 31(4), 851–879.

Kramer, S. L. (1996). *Geotechnical earthquake engineering*, Prentice-Hall, Upper Saddle River, N.J.

Kramer, S. L., Mayfield, R. T., and Anderson, D. G. (2006). "Performance-based liquefaction evaluation: Implications for codes and standards." *Proc., 8th U.S. National Conf. on Earthquake Engineering*, San Francisco.

Krishnan, S., and Journel, A. G. (2003). "Spatial connectivity: From variograms to multiple-point measures." *Math. Geol.*, 35(8), 915–925.

McGuire, R. K. (2004). *Seismic hazard and risk analysis*, Earthquake Engineering Research Institute, Berkeley, Calif.

Phoon, K. K. (2006). "Modeling and simulation of stochastic data." *Proc., GeoCongress 2006*, Atlanta, 17.

Phoon, K.-K., and Kulhawy, F. H. (1999a). "Characterization of geotechnical variability." *Can. Geotech. J.*, 36, 612–624.

Phoon, K.-K., and Kulhawy, F. H. (1999b). "Evaluation of geotechnical property variability." *Can. Geotech. J.*, 36, 625–639.

Phoon, K.-K., Quek, S.-T., and An, P. (2003). "Identification of statistically homogeneous soil layers using modified Bartlett statistics." *J. Geotech. Geoenviron. Eng.*, 129(7), 649–659.

Popescu, R., Prevost, J. H., and Deodatis, G. (2005). "3D effects in seismic liquefaction of stochastically variable soil deposits." *Geotechnique*, 55(1), 21–31.

Rackwitz, R. (2000). "Reviewing probabilistic soils modeling." *Comput. Geotech.*, 26(3–4), 199–223.

Rollins, K. M., and Seed, R. B. (1990). "Influence of buildings on potential liquefaction damage." *J. Geotech. Engrg.*, 116(2), 165–185.

Rosenblatt, M. (1952). "Remarks on a multivariate transformation." *Ann. Math. Stat.*, 23, 470–472.

Stanford Center for Reservoir Forecasting (SCRF). (2006). "The Stanford geostatistical modeling software (S-GeMS)." ⟨http://sgems.sourceforge.net/⟩.

Toprak, S., and Holzer, T. L. (2003). "Liquefaction potential index: Field assessment." *J. Geotech. Geoenviron. Eng.*, 129(4), 315–322.

Uzielli, M., Vannuchi, G., and Phoon, K.-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, Cambridge, Mass.

Wang, M., and Takada, T. (2005). "Macrospatial correlation model of seismic ground motions." *Earthquake Spectra*, 21(4), 1137–1156.

Youd, T. L. et al. (2001). "Liquefaction resistance of soils: Summary report from the 1996 NCEER and 1998 NCEER/NSF workshops on evaluation of liquefaction resistance of soils." *J. Geotech. Geoenviron. Eng.*, 127(10), 817–833.