

International Forum on Engineering Decision Making

Second IFED Forum, April 26-29, 2006, Lake Louise, Canada

Accounting for soil spatial variability when assessing liquefaction risk

Jack W. Baker, Yahya Y. Bayraktarli, and Michael H. Faber
Swiss Federal Institute of Technology (ETH Zürich)

Abstract

Liquefaction triggering assessments are typically performed for only individual locations, providing little or no information in regard to the expected 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.

1. Introduction

Empirical models for the assessment of soil liquefaction potential are based on soil properties at a 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 as well as 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 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 approach has received relatively less attention for liquefaction problems, however, so the applicability for this specific problem will be considered 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 (although this is a challenging task in practice). The distribution of ground motion

intensity from potential future earthquakes can be obtained from probabilistic seismic hazard analysis results.

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 limited sample data.

2. 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 Figure 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. Then a liquefaction triggering criterion is evaluated at each location, based on the model parameters specified by the random fields. Finally, based on the field models and the liquefaction criterion, a probabilistic assessment of the liquefaction realizations may be performed. The individual components of this framework are described briefly in this section.

2.1. Modeling of liquefaction occurrence

The method chosen used to model liquefaction will determine which soil properties must be modeled, so it is useful to consider this model component first. The most common approach for modeling liquefaction occurrence in practice uses empirical criteria which relate measured soil parameters and observed occurrence (or non-occurrence) 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 is commonly used for modeling (e.g., Cetin et al. 2004; e.g., 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. These models can potentially capture more complex effects such as post-liquefaction soil behavior and the effect of nearby structures on liquefaction. These models are not utilized in this study, but they have been used in other studies that consider the seismic response of random-field soil models (e.g., Fenton and Vanmarcke 1998; Popescu et al. 2005). They could be incorporated in the approach described here if desired.

2.2. Probability distribution of soil properties

Before attempting to characterize the spatial dependence of soil properties it is necessary to first characterize the soil properties at a single point in terms of a joint probability distribution.

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, soil penetration resistance is measured by a different parameter obtained using that testing procedure, 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 (Fenton 1999a; Fenton 1999b; JCSS 2002; e.g., Jones et al. 2002; Phoon and Kulhawy 1999; Phoon et al. 2003). 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 on the site.

2.3. 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 (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 in more detail 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 can also be 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 (Uzielli et al. 2005). 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 practically quantified and in many applications it is deemed to be a sufficiently accurate representation of dependence (Goovaerts 1997).

2.4. Observations of 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, in addition to the sampled values, guides the selection of appropriate soil probability distributions discussed above. 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 adjacent locations. The ability to incorporate this information is an important feature of the proposed framework.

2.5. 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*) along with 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 disaggregation (Kramer 1996; McGuire 2004). Disaggregated 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, p195). An example of this joint density is shown in Figure 2, as obtained for a site in Los Angeles, using a site-specific hazard analysis. 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 (Wang and Takada 2005).

2.6. 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., Equation (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 0 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 $g(\mathbf{X}, \mathbf{u})$, may not (or can not) be known with certainty, they are best expressed as random variables. The spatial distribution of these random variables can be considered using random fields techniques (Fenton 1990; Vanmarcke 1983). In this view, regions of liquefaction correspond to excursions of random fields: a relatively well-studied topic.

Two primary approaches are available for considering excursions of random fields. The first approach uses analytical formulas obtained using random fields 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 2006). Note that these results are for excursions above some threshold value, but the liquefaction problem can be easily formulated in this way by considering excursions of $-g(\mathbf{X}, \mathbf{u})$ above zero rather than excursions of $g(\mathbf{X}, \mathbf{u})$ below zero. Analytical solutions are appealing because of their ability to explicitly provide a relationship between model parameters and computed values, but it is not clear whether the required assumptions will be 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 an important part of practical liquefaction susceptibility evaluations.

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 variable characterizations of all input variables, account for spatial dependence of the random variables, and are compatible with observed values obtained by sampling. The field of geostatistics considers approaches of this type (Goovaerts 1997). The assumptions required for this approach tend to be much less restrictive than with the random fields approach, and it has the important advantage (for this application) of incorporating observed data values at sampled locations. The relative disadvantage is that no analytical equations for excursion properties are available and, as with any Monte Carlo approach, the computational expense will be greater. Efficient public-source 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 in some applications considering spatially random soil, Monte Carlo simulations are performed in the frequency domain—this approach has some desirable features, but it is not able to incorporate observed values and so it not considered further here.

3. 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. This city experienced liquefaction failures during the 1999 Mw 7.4 Kocaeli earthquake, and extensive post-earthquake investigations have produced a large set of soil borings that are now available for analysis (Bray and Rodríguez-Marek 2004). A small individual site in this city, shown in Figure 3, 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 SPT tests are available and are used to constrain the uncertainty in soil properties at other locations within the site.

3.1. 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

$$g(\mathbf{X}, \mathbf{u}) = g(N_{1,60}, CSR_{eq}, M_w, FC, \sigma'_v, \varepsilon) = N_{1,60}(1 + 0.004FC) - 13.32 \ln CSR_{eq} - 29.53 \ln M_w - 3.70 \ln \frac{\sigma'_v}{P_a} + 0.05FC + 44.97 + \varepsilon_L \quad (1)$$

where $N_{1,60}$ is the corrected SPT blow count, CSR_{eq} is the equivalent cyclic stress ratio, M_w is the moment magnitude of the earthquake, FC is fines content, σ'_v is effective vertical stress, and ε_L is a random variable representing model uncertainty. In the above notation, \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 Equation (1), a few further relationships are needed. The calculation for equivalent cyclic stress ratio is

$$CSR_{eq} = 0.65 \left(\frac{PGA}{g} \right) \left(\frac{\sigma_v}{\sigma'_v} \right) r_d \quad (2)$$

where g is the acceleration of gravity, σ_v is total vertical stress, and r_d is 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

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

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

The evaluation of liquefaction triggering for a single point thus requires the evaluation of Equations (1) through (3), and thus the input parameters in these equations must be modeled as random fields. The parameter models have been used in this calculation, 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

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

where d is the 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 measured values from throughout the city (it had a mean of 6.5, standard deviation of 5.6, and was strongly skewed to the right). Note that homogeneity has been assumed in order to estimate the probability distribution from a sample of measurements. Average shear wave velocity, $V_{s,12m}^*$; was deterministically defined as 150 m/s, on the basis of spectral analysis of surface waves data at the site (Bray and Rodríguez-Marek 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 throughout the city.

As described earlier and illustrated in Figure 2, the joint rate density of PGA and M (i.e., the joint probability density function multiplied by a rate of occurrence), is obtained from the results of a site-specific probabilistic seismic hazard analysis. The discrete numerical representation of this distribution will be utilized directly for these properties, rather than a parametric distribution.

For the purposes of illustration, all random variables except $N_{1,60}$ were assumed to be perfectly dependent at the spatial scale of Figure 3 (i.e., the variables were assumed to take the same value at each point in the example site). The spatial dependence of $N_{1,60}$ was treated more rigorously, as explained in the following section.

3.2. Spatial dependence of soil properties

The mean value of $N_{1,60}$ is assumed to be known and constant throughout the study site (prior to consideration of measurements from nearby samples). 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 which vary smoothly over the study site, and could 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 joint dependence between properties at two points, and the analytical equations are very tractable. The soil properties do not necessarily have Gaussian distributions, however. In order to take advantage of the desirable properties of multivariate Gaussian models, the data of interest is transformed using a normal-score mapping. To perform this transform, the complimentary distribution function (CDF) of the true soil values is generated (using either an empirical CDF from observed data values or the CDF corresponding to an assumed or fitted probability 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. This is expressed mathematically by

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

where y is the original data from a distribution represented by the CDF $F(y)$, $\Phi^{-1}(\cdot)$ is the inverse of the standard Gaussian CDF and z is the transformed data. This transformation by definition produces variables that marginally have a standard Gaussian distribution. After verifying that the transformed data is reasonably represented by a multivariate Gaussian distribution (Goovaerts 1997, p271), this transformed data is used for statistical estimation and simulation. By inverting Equation (5) with the simulated values, one obtains simulated data having the originally specified marginal distribution. Other algorithmic details are provided in detail elsewhere (Deutsch and Journel 1997; Goovaerts 1997).

The normal-score transformed data is then 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}

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

where $Z(\mathbf{u})$ is the distribution of the (normal-score-transformed) random variable at location \mathbf{u} . The vector distance \mathbf{h} accounts for both length and direction. Isotropy is assumed for this application, so that the semivariogram is a function of separation length ($|\mathbf{h}|$) 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 corrected SPT blow count, $N_{1,60}$

$$\gamma(h) = \begin{cases} 1.5\left(\frac{h}{25}\right) - 0.5\left(\frac{h}{25}\right)^3 & \text{if } h \leq 25 \\ 1 & \text{if } h > 25 \end{cases} \quad (7)$$

where h is the (directionally independent) scalar separation length in units of meters. This functional form is referred to 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., Stanford 2006).

Once spatial dependence of the considered soil parameters has been defined, realizations of the soil properties can be generated using a sequential simulation approach. Consider the example site from Figure 3. Soil properties are desired for a discretized field of 120 x 200 elements, each with dimension 0.25 meters x 0.25 meters. Thus, the goal is to generate joint realizations of $N_{1,60}$ values for these 24,000 elements, consistent with the observed marginal distributions and spatial dependence of this property.

The sequential simulation approach is appealing because it is easy to generate samples at a single element, conditional upon the values that sampled (or previously simulated) elements at surrounding locations take. Thus, the joint 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 has 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 is given by

$$\begin{bmatrix} Z_1 \\ \mathbf{Z}_{orig} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} 1 & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \right) \quad (8)$$

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 $\boldsymbol{\Sigma}$. The vector \mathbf{Z}_{orig} represents the original data values at the sampled locations, and $\mathbf{0}$ is a vector of zeros having the same size as \mathbf{Z}_{orig} (i.e., the mean vector, which is equal to zero in this case because of the normal-score transform). 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 Equation (7), noting that the covariance between locations with a separation distance h is equal to $1 - \gamma(h)$. Note that 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

$$(Z_1 | \mathbf{Z}_{orig} = \mathbf{z}) \sim N(\boldsymbol{\Sigma}_{12} \cdot \boldsymbol{\Sigma}_{22}^{-1} \cdot \mathbf{z}, 1 - \boldsymbol{\Sigma}_{12} \cdot \boldsymbol{\Sigma}_{22}^{-1} \cdot \boldsymbol{\Sigma}_{21}) \quad (9)$$

where \mathbf{z} is the vector of $N_{1,60}$ values at the sampled locations. Note that \mathbf{Z}_{orig} is a random variable representing our model for uncertain soil parameters prior to sampling, and the sampling procedure reveals 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}_{orig} of Equation (8) for the subsequent simulations of Z_2, Z_3, \dots).

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 Equation (5) for each location in the field (i.e., by performing the inverse of the normal-score transform). 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 also been developed. One extension that may be of interest for liquefaction problems is the simulation of random fields for correlated parameters. Once the dependence between the two parameters has been estimated, a procedure similar to the one above is used, where each location in the site is visited, but now 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 easily 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 in which values for each location are simulated is not important, but to ensure sufficient variety among a finite number of samples, usually the order of locations is randomized and a different order is used for each sample. Also, note that when simulating the last value in the field, Equation (9) implies that all 23,999 previous values should be used for conditioning, and this requires inversion of a Σ_{22} covariance matrix with size 23,999 x 23,999. In practice, however, the effect of distant data points is ‘screened’ by the influence of the closest data. Thus the number of conditioning points can be significantly reduced, which decreases the computational expense of the procedure without practically affecting the resulting conditional distribution. These and other implementation issues are addressed by Goovaerts (1997). 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 Figure 4. In Figure 5, the mean and standard deviation of 1000 simulations generated using this method are shown. In Figure 5b, 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.

3.3. Probability of liquefaction for a given ground motion intensity

To perform the evaluation of liquefaction occurrence, the procedure illustrated in Figure 1 is used. The site of interest is divided into discrete cells. For each soil or ground motion property used in Equation (1), a matrix of values is generated representing that variable’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 others will take different values in each cell. The illustrations in Figure 4 are simply graphical displays of these matrices. Some of the variables can also be specified deterministically, if their uncertainty is not expected to affect the results.

Once all the matrices are generated, Equation (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 Figure 6. To generate Figure 6, a magnitude 7.4 earthquake with a PGA of 0.3 g was assumed; variation in ground motion intensity will be considered below. The plotted examples correspond to the $N_{1,60}$ values shown in Figure 4, but there is not a one-to-one mapping between the two pictures because of the uncertainty in the other model parameters, which were also simulated to generate Figure 6.

The graphical illustrations are interesting, but they must be summarized if a large number of simulations are to be considered. Any function of interest can be evaluated for each of the simulations, and if the function output is a scalar then it can be easily summarized numerically or graphically. For example, liquefaction occurrence under Building A1 may be of interest. Then 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

$$P(Y > y | pga, m) = \int_{\boldsymbol{\eta}} I(h(\mathbf{g}(\boldsymbol{\eta} | pga, m) > y)) f_{\boldsymbol{\eta}}(\mathbf{e}) d\mathbf{e} \quad (10)$$

where $\mathbf{g}(\boldsymbol{\eta} | pga, m)$ is the limit state function from Equation (1), evaluated for $PGA=pga$ and magnitude= m , but now written in bold to denote that it is a vector output corresponding to the 24,000 $\mathbf{g}(\boldsymbol{\eta} | pga, m)$ values at the site. The random variable $\boldsymbol{\eta}$ represents the vector of all input soil variables at each location in the site, and $f_{\boldsymbol{\eta}}(\mathbf{e})$ is the 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 $h(\cdot)$ is the 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 Equation (10) involves considering all possible realizations of soil properties $\boldsymbol{\eta}$ at the site, but numerical integration is not possible because $\boldsymbol{\eta}$ is very high dimensional (i.e., it contains 4 soil properties * 24,000 locations in this example). For this reason, the random field simulation approach described above is used to evaluate this integral. The results from this calculation are shown in Figure 7. For the given earthquake loading and soil property distributions, there is approximately a 50% probability that some portion of the soil will liquefy, but only a 25% probability that more than 50% 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.

3.4. Incorporating multiple ground motion intensity levels

The result in Figure 7 was obtained by assuming that a magnitude 7.4 earthquake occurred and caused a peak ground acceleration of 0.3 g . But without knowing the probability that this event, and others, will occur, it is difficult or impossible to evaluate whether this performance is acceptable. For this reason, ground motion hazard is incorporated in the analysis. The mean rate density illustrated in Figure 2 provides the joint probability density function of magnitude and PGA , multiplied by the rate of occurrence of any considered event. 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. The alternative approach used here, which addresses this shortcoming, is to use Monte Carlo simulation for the soil properties and numerical integration for the ground motion intensity.

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 Equation (10). To complete the evaluation, one must then consider the entire range of relevant PGA and magnitude values using the following equation

$$\lambda(Y > y) = \int \int_{PGA, M} P(Y > y | pga, m) MRD_{PGA, M}(pga, m) dpga dm \quad (11)$$

where $P(Y > y | pga, m)$ comes from Equation (10) and $MRD_{PGA, M}(pga, m)$ is the mean rate density of PGA and M , as illustrated in Figure 2. 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 Equation (11) are shown in Figure 8. The figure was generated using the soil properties model discussed above, and then considering 200 PGA/M combinations in the numerical integration of Equation (11): 20 PGA values and 10 M values. Although ground motions in the intensity range of interest occur approximately once every 8 years, some level of liquefaction under the building occurs only once every 50 years approximately. Liquefaction of the total area under the building is predicted to occur approximately once every 100 years. While these results depend upon the hypothetical random variable assumptions used here for illustration, they nonetheless illustrate the potentially useful information that could be obtained using the approach.

This combination of numerical integration and simulation, which is similar to the stratified sampling Monte Carlo technique, is useful in this case 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).

4. Discussion

The framework described above has great potential for describing the spatial distribution of soil liquefaction. Significant challenges remain, however, with respect to modeling assumptions and practical estimation challenges. 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 carefully before a model is assumed (Rackwitz 2000). The specification of which parameters are ergodic and non-ergodic will be important for test planning, and non-ergodicity of parameters may cause serious difficulties for parameter estimation. Note also that in the above approach, all probabilistic parameters are assumed to be known with certainty (e.g., the soil parameters are random variables, but their means and variances are assumed to be perfectly known). In practice this will not be the case, because the distribution parameters will be estimated from expert judgment and/or finite data samples. Uncertain probabilistic parameters can be incorporated by performing the above calculations conditional upon the probabilistic parameters, and then considering all possible values of the probabilistic parameters through another loop of either numerical integration or Monte Carlo simulation to randomize the probabilistic parameter values.

A further challenge for random field characterization is the presence of soil layers. If 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 will likely not hold for all 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 which provides a representation of the phenomenon being studied, and it may not be feasible to develop a reasonable training image for soil layering in many cases 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 post-triggering 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 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. Empirical modeling of interactions between structures and potentially liquefiable soil has also received somewhat limited attention (Rollins and Seed 1990). Models which use finite element analysis to model liquefaction promise to address this issue more completely (Fenton and Vanmarcke 1998; Steve Koutsourelakis 2002), but also require more computational expense and analyst time to model the site. The authors are not aware of any studies of this types which consider spatially variable soil models that are 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 does variability of the peak ground acceleration values needed above. The extent to which these additional challenges will limit the use this method with finite-element-based evaluations is not yet known.

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 was used, and further work is needed 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. A structure may fail even if the soil does not liquefy, and the interaction between soil response and structural response is sometimes important as well. In order to make fully informed decisions for managing liquefactions

risks, it would be helpful to consider a system which incorporates all potential failure mechanisms (e.g., structural collapse, liquefaction, bearing failure due to soil-structure-interaction). A number of active research fields are aiding in progress towards this goal.

5. Summary

A framework has been proposed for evaluating occurrence of liquefaction probabilistically, conditional upon observations of soil properties obtained from site samples. The framework incorporates several model components 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 shaking. These model components are all available presently, although they have been developed independently and have not been previously combined in the form seen here.

The numerical procedures have been outlined to demonstrate the feasibility of the proposed approach. Software tools for seismic hazard analysis and geostatistics facilitate the needed computations, and allow calculations of this type to be performed without great effort. Estimating the needed probability distributions for soil properties at an arbitrary 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 conventional criteria. The simultaneous consideration of random (unknown) soil properties and random future earthquake shaking is an improvement over many current assessments, which 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 towards considering costs and benefits explicitly when making engineering decisions regarding liquefaction risk.

6. References

- Adler, R. J. (1981). *The geometry of random fields*, J. Wiley, Chichester [Eng.] ; New York.
- Adler, R. J., and Taylor, J. E. (2006). *Random Fields and Geometry*, (in press, preprint at <http://iew3.technion.ac.il/~radler/publications.html>).
- Bazzurro, P. (1998). *Probabilistic seismic demand analysis*, Stanford University, Stanford, CA.
- Bray, J. D., and Rodríguez-Marek, A. (2004). "Characterization of forward-directivity ground motions in the near-fault region." *Soil dynamics and earthquake engineering*, 24(11), 815-828.
- Bray, J. D., et al. Documenting Incidents of Ground Failure Resulting from the August 17, 1999 Kocaeli, Turkey Earthquake. Pacific Earthquake Engineering Research Center website. <http://peer.berkeley.edu/turkey/adapazari/>.
- Cetin, K. O., et al. (2004). "Standard Penetration Test-Based Probabilistic and Deterministic Assessment of Seismic Soil Liquefaction Potential." *Journal of Geotechnical and Geoenvironmental Engineering*, 130(12), 1314-1340.
- Degroot, D. J., and Baecher, G. B. (1993). "Estimating autocovariance of in-situ soil properties." *Journal of Geotechnical Engineering*, 119(GT1), 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. (1989). "Structural Reliability Theory, Extremes of Nonhomogeneous Gaussian Random Fields." *Ph.D. Thesis*.
- Fenton, G. A. (1990). "Simulation and analysis of random fields." *Ph.D. Thesis*.
- Fenton, G. A. (1999a). "Estimation for Stochastic Soil Models." *Journal of Geotechnical and Geoenvironmental Engineering*, 125(6), 470-485.
- Fenton, G. A. (1999b). "Random Field Modeling of CPT Data." *Journal of Geotechnical and Geoenvironmental Engineering*, 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-Verlag, New York.
- Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*, Oxford University Press, New York.
- Iwasaki, T., Tatsuoka, F., Tokida, K., and Yasuda, S. "Estimation procedure of liquefaction potential and its application to earthquake resistance design." *U.S.-Japan Panel on Wind and Seismic Effects, UJNR*.
- Jaksa, M. B., and Fenton, G. A. (2000). "Random Field Modeling of CPT Data." *Journal of Geotechnical and Geoenvironmental Engineering*, 126(12), 1212-1216.
- Joint Committee on Structural Safety. (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." Pacific Earthquake Engineering Research Center, University of California at Berkeley, Berkeley, California.
- Kramer, S. L. (1996). *Geotechnical earthquake engineering*, Prentice Hall, Upper Saddle River, N.J.
- Krishnan, S., and Journel, A. G. (2003). "Spatial Connectivity: From Variograms to Multiple-Point Measures." *Mathematical Geology*, 35(8), 915-925.
- McGuire, R. K. (2004). *Seismic Hazard and Risk Analysis*, Earthquake Engineering Research Institute, Berkeley.
- Phoon, K.-K., and Kulhawy, F. H. (1999). "Evaluation of geotechnical property variability." *Canadian Geotechnical Journal*, 36, 625-639.
- Phoon, K.-K., Quek, S.-T., and An, P. (2003). "Identification of Statistically Homogeneous Soil Layers Using Modified Bartlett Statistics." *Journal of Geotechnical and Geoenvironmental Engineering*, 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 modelling." *Computers and Geotechnics*, 26(3-4), 199-223.
- Rollins, K. M., and Seed, R. B. (1990). "Influence of buildings on potential liquefaction damage." *Journal of Geotechnical Engineering*, 116(2), 165-185.
- Stanford Center for Reservoir Forecasting. (2006). The Stanford Geostatistical Modeling Software (S-GeMS). <http://sgems.sourceforge.net/>.
- Steve Koutsourelakis, J. H. P., George Deodatis, (2002). "Risk assessment of an interacting structure-soil system due to liquefaction." *Earthquake Engineering & Structural Dynamics*, 31(4), 851-879.
- Toprak, S., and Holzer, T. L. (2003). "Liquefaction Potential Index: Field Assessment." *Journal of Geotechnical and Geoenvironmental Engineering*, 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 Press, Cambridge, Mass.
- Wang, M., and Takada, T. (2005). "Macrospectral 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." *Journal of Geotechnical and Geoenvironmental Engineering*, 127(10), 817-833.

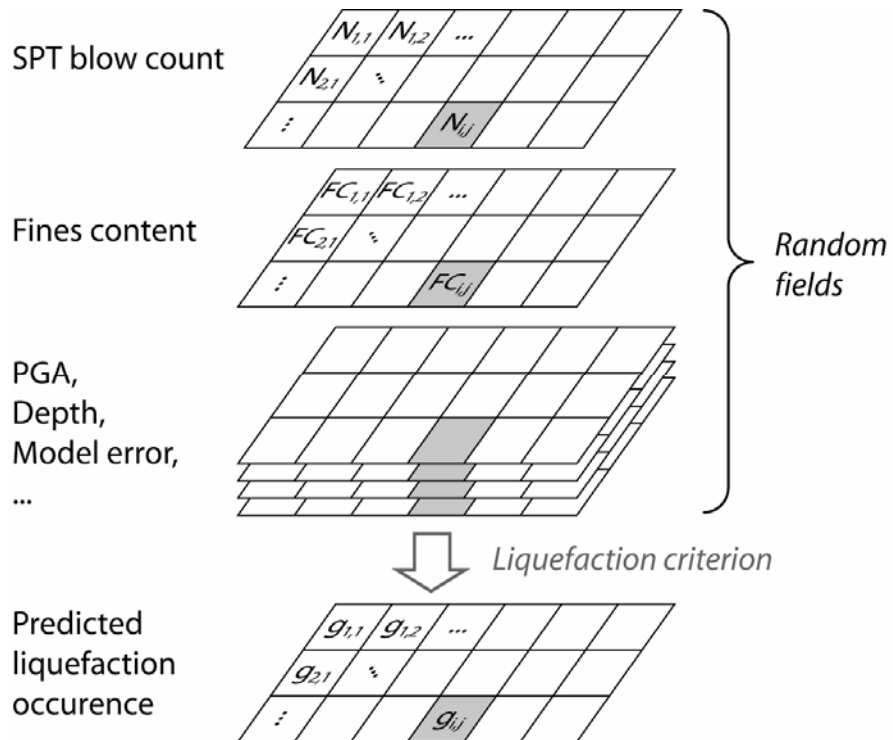


Figure 1: Schematic illustration of the procedure for evaluating the extent of soil liquefaction as a function of random fields of soil properties and other model parameters..

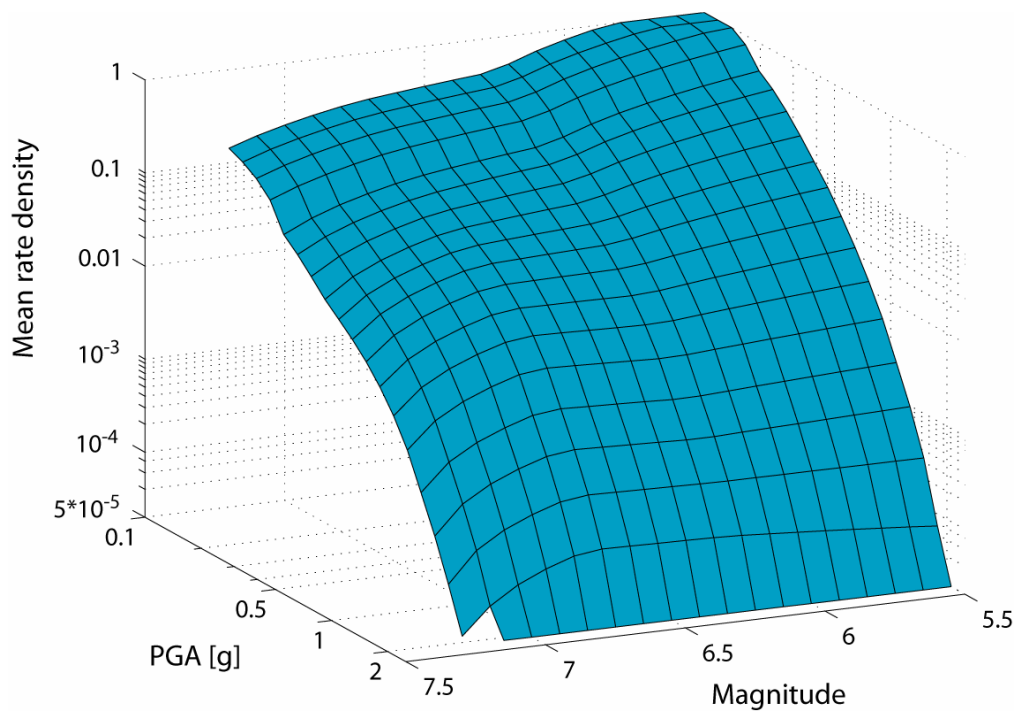


Figure 2: Example mean rate density of PGA and magnitude.

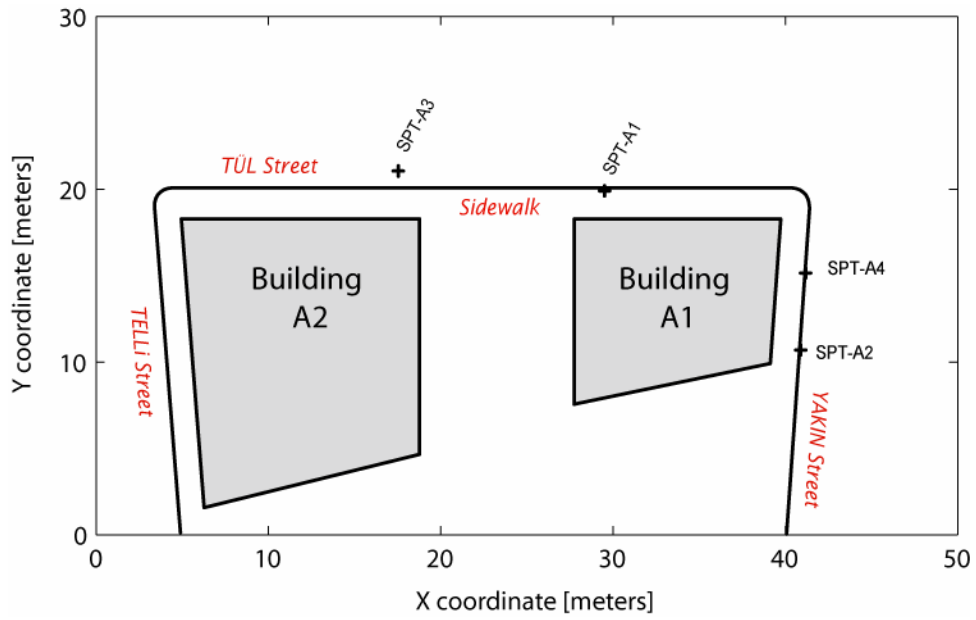


Figure 3: Example site in Adapazari (adopted from Bray et al. PEER website)

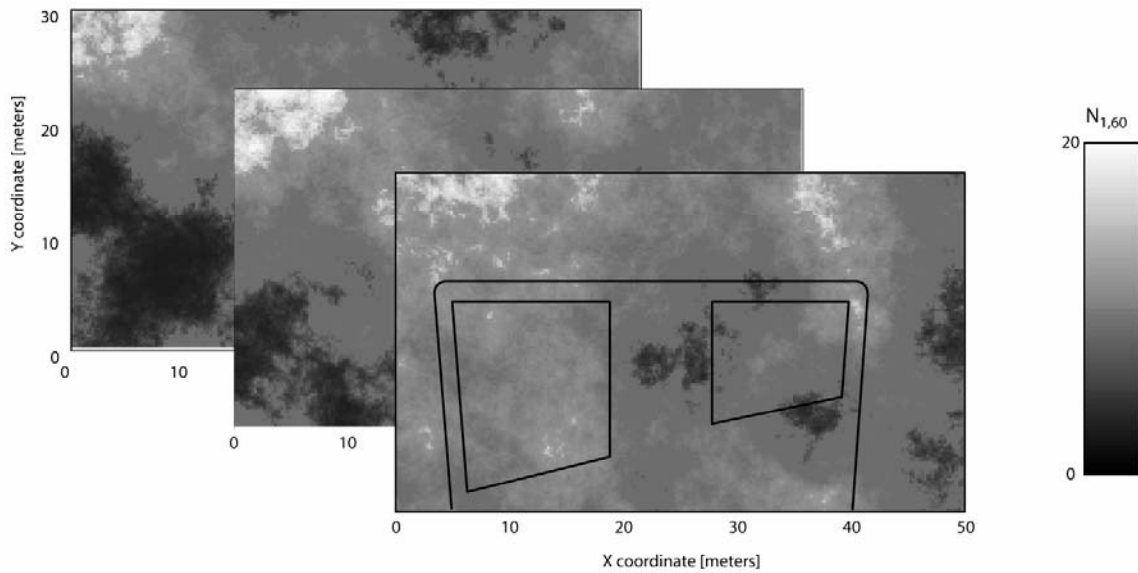


Figure 4: Three conditional simulations of corrected SPT blow count ($N_{1,60}$) values.

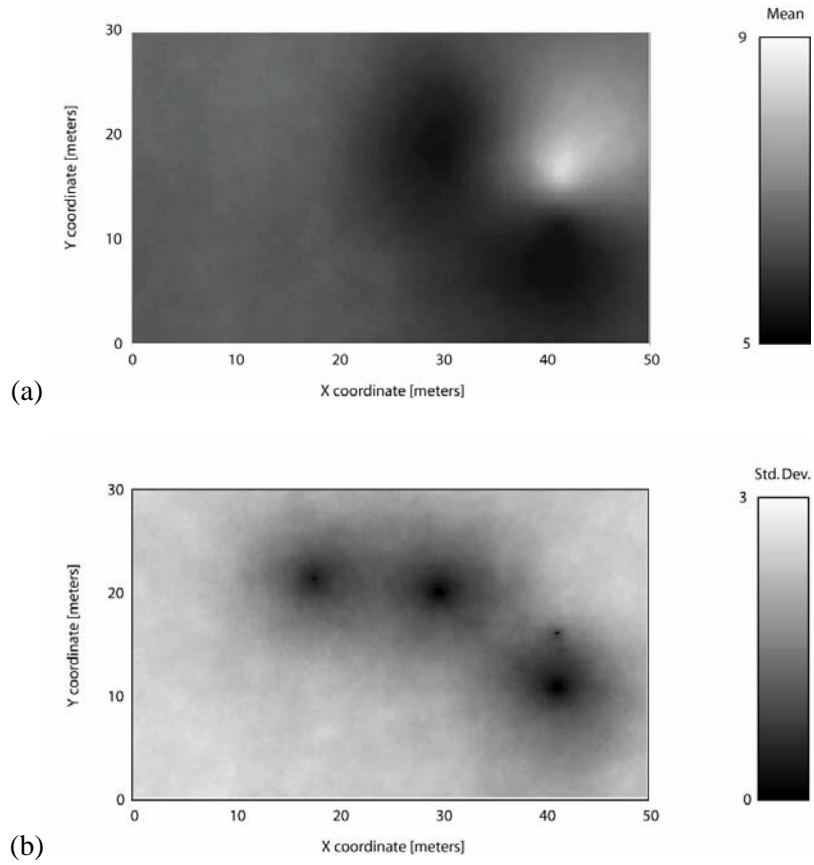


Figure 5: (a) Mean, and (b) standard deviation, of simulated $N_{1,60}$ values.

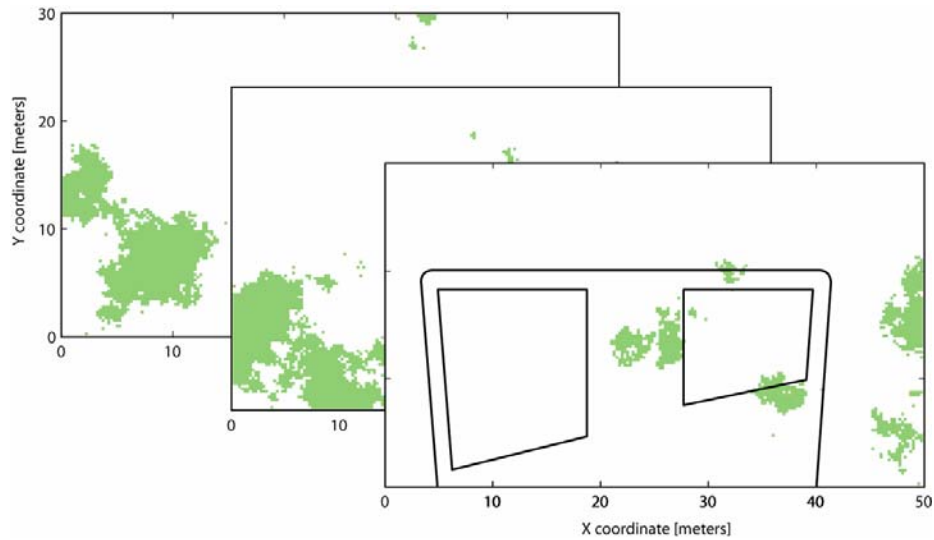


Figure 6: Locations of liquefaction triggering, as computed from the conditional simulations of site soil properties, and given a Magnitude 7.4 earthquake with a PGA of 0.3 g. Liquefied regions are shaded.

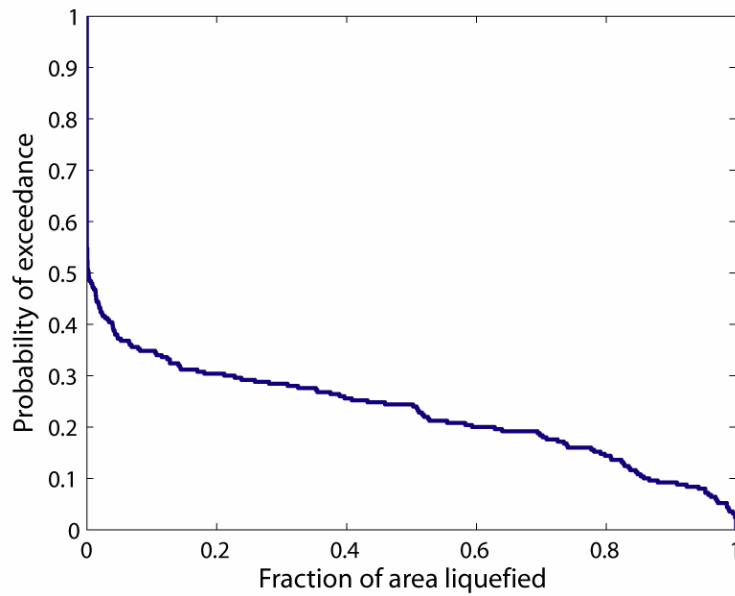


Figure 7: Probability of exceedance versus fraction of the area under Building A1 that liquefies, given a Magnitude 7.4 earthquake with a PGA of 0.3 g.

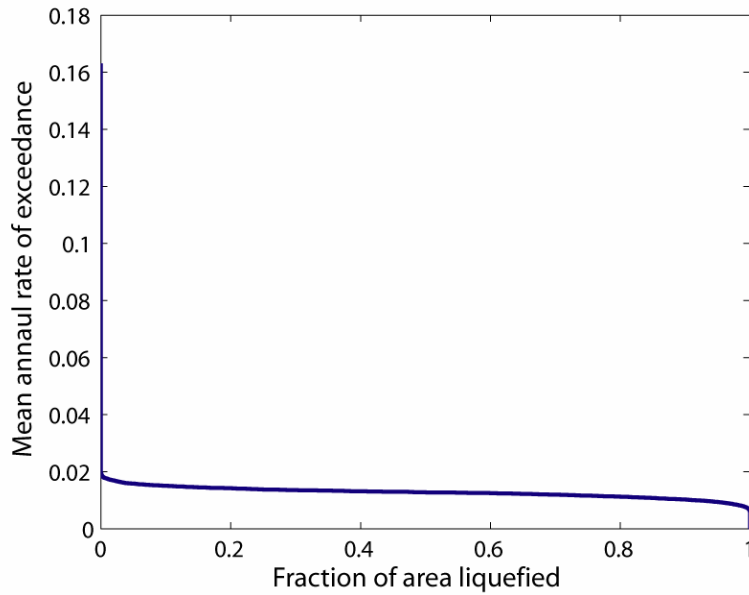


Figure 8: Rate of exceedance versus fraction of the area under Building A1 that liquefies, considering all possible ground motion intensities as specified by the ground motion hazard.