Liquefaction Hazard Mapping—Statistical and Spatial Characterization of Susceptible Units

Laurie G. Baise, M.ASCE[1]; Rebecca B. Higgins[2]; and Charles M. Brankman[3]

Abstract: In this paper, we present a three step method for characterizing geologic deposits for liquefaction potential using sample based liquefaction probability values. The steps include statistically characterizing the sample population, evaluating the spatial correlation of the population, and finally providing a local and/or global estimate of the distribution of high liquefaction probability values for the deposit. When spatial correlation is present, ordinary kriging can be used to evaluate spatial clustering of high liquefaction probability values within a geologic unit which in turn can be used in a regional liquefaction potential characterization. If spatial correlation is not present in the data, then a global estimate can be used to estimate the percentage of samples within the deposit which have a high liquefaction probability. By describing the liquefaction potential with a binomial distribution (high versus low), a global estimate can provide an estimate of the mean as well as uncertainty in the estimate. To demonstrate the method, we used a dense data set of subsurface borings to identify and characterize liquefiable deposits for hazard mapping in Cambridge, Mass.

DOI: 10.1061/(ASCE)1090-0241(2006)132:6(705)

CE Database subject headings: Liquefaction; Earthquakes; Seismic effects; Mapping; Statistics; Soil penetration tests.

Introduction

Regional liquefaction hazard mapping projects have predominantly relied on criteria that relate surficial geology to liquefaction susceptibility (Youd and Perkins 1978). Geologic units are identified by their age and depositional environment and then characterized in terms of their susceptibility. This method leads to the identification of large regions of susceptible material. The resulting maps show geologic units that likely contain liquefiable sediments but do not identify the location or extent of liquefiable sediments within the geologic unit. Therefore, within a susceptible unit, maybe only a very small area will actually liquefy given an earthquake. We set out to provide more information on the likelihood of liquefaction in a susceptible unit using global and local population estimates paired with a dense collection of geotechnical data.

Regional liquefaction susceptibility maps will never provide detailed enough information for absolute susceptibility at a site level and are not meant to, but a more thorough characterization than currently used will lead to a more accurate assessment of susceptibility or liquefaction potential. We propose a method for evaluating the distribution of liquefiable soils within a geologic unit. Our method combines data from the regional and local scale, where the regional data includes surficial geology and the local data includes soil sample data with calculated liquefaction probabilities. Using the local data, we propose a three step method to provide a global and when possible a local (spatial) estimate of the population of liquefiable materials in the geologic deposit. Our proposed method is demonstrated on a case study in Cambridge, Mass. where a dense collection of subsurface test borings was assembled to characterize potentially liquefiable materials. Additional information on the case study dataset and liquefaction hazard can be found in Baise and Brankman (2004).

Background

As liquefaction hazard mapping projects proliferate around the country and the world, the mapping method has remained relatively constant. Most of the existing liquefaction susceptibility maps are based solely on geology. For example, the liquefaction susceptibility maps for the San Francisco Bay Area provided on the Association of Bay Area Governments website ⟨http://quake.abag.ca.gov/⟩ are a direct interpretation of the susceptibility of surficial deposits based on surficial geology mapping of the region (Knudsen et al. 2000). Currently, many liquefaction mapping projects include the concurrent collection of subsurface data to provide local data for the susceptibility or liquefaction potential estimate. The subsurface data may include standard penetration test (SPT) *N*-values, cone penetrometer test (CPT) data, shear-wave velocity (Vs) data, soil descriptions (including grain size distributions), stratigraphy, and groundwater data. Generally, a scattered sample of subsurface data is collected in the susceptible unit and used to characterize that unit; however, the maps are still primarily based on surficial geology. Recent studies in Victoria, B.C.; Alameda, Berkeley, Emeryville, Oakland, and Piedmont, Calif.; San Francisco Bay area, Calif.; and Memphis and Shelby Counties, Tenn. have used subsurface test borings logs to supplement the characterization of susceptible

[1]Assistant Professor, Dept. of Civil and Environmental Engineering, Tufts Univ., 113 Anderson Hall, Medford, MA 02155.

[2]Senior Engineer, Haley & Aldrich, 465 Medford St., Suite 2200, Boston, MA 02129.

[3]Graduate Student, Dept. of Earth and Planetary Sciences, Harvard Univ., 20 Oxford St., Cambridge, MA; formerly, William Lettis and Associates, Inc., 1770 Botelho Dr., Suite 262, Walnut Creek, CA 94596.

Note. Discussion open until November 1, 2006. 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 June 24, 2004; approved on October 20, 2005. This paper is part of the ***Journal of Geotechnical and Geoenvironmental Engineering***, Vol. 132, No. 6, June 1, 2006. ©ASCE, ISSN 1090-0241/2006/6-705–715/$25.00.

deposits (Monahan et al. 1998; Knudsen et al. 2000; Monahan et al. 2000; Broughton et al. 2001; Holzer et al. 2002). Knudsen et al. (2000) used available boring logs to estimate the peak ground acceleration required to cause liquefaction [following the Seed and Idriss (1971) simplified approach]. The susceptibility category for geologic units was then derived from a combination of observations of historic liquefaction, peak ground acceleration necessary to cause liquefaction, ground water level, and geology-based susceptibility. Broughton et al. (2001) also used boring logs in their analysis of liquefaction susceptibility maps for Memphis and Shelby Counties, Tenn. Their maps were produced strictly by geologic methods and the analysis of boring logs [following the Seed and Idriss (1971) simplified approach] was used to verify the results in a qualitative way.

When regional liquefaction hazard mapping is attempted in a city, numerous locations of subsurface data are often available. Rather than 10 subsurface borings over a square kilometer, 1,000 subsurface borings are available. With a dense array of subsurface data, the characterization of units becomes more complete. The relative liquefaction hazard maps produced for Victoria, B.C. depended on the interpretation of stratigraphy derived from over 5,000 boring logs (Monahan et al., 2000). The hazard classification for the Victoria maps was based on an interpretation of the stratigraphy represented in the boring logs and a detailed analysis of 31 borings. The detailed analysis consisted of a combination of a probabilistic prediction of liquefaction using the Seed and Idriss (1971) simplified approach and a probability of liquefaction severity index which depends on depth and thickness of the liquefiable materials (Monahan et al. 1998, 2000). Six stratigraphic units were characterized using anywhere from 1 to 11 borings. The susceptibility classifications took into account the variability of investigated sites by setting a range of susceptibility rather than an absolute value: medium to very high or high to very high. In addition, Holzer et al. (2002) have completed liquefaction hazard maps for Alameda, Berkeley, Emeryville, Oakland, and Piedmont, Calif. using the liquefaction potential index (Iwasaki et al. 1978) based on 210 cone penetrometer tests. Holzer et al. (2002) applied the liquefaction potential index (LPI) to regional mapping by assigning approximate percentages of affected area for each geologic unit. The mapped percentages were found by finding the percentage of LPI values over 5 for each geologic unit.

Liquefaction susceptibility or potential is generally assessed on two scales: regionally based on surficial geologic unit or locally based on sample data (SPT, CPT, or Vs). There have been numerous studies focused on local liquefaction hazard across a site (Robertson and Wride 1998; Holzer et al. 1999; Liao et al. 2002; Elkateb et al. 2003; Popescu et al. 2005). Holzer et al. (1999) evaluated several sites that experienced liquefaction during the 1994 Northridge earthquake using the LPI and other methods. Liao et al. (2002) used CPT data to investigate several sites in the New Madrid seismic zone where paleoliquefaction evidence was found. The lateral heterogeneity was included using cross sections for the site. Popescu et al. (2005) looked at the effect of three-dimensional soil variability on liquefaction potential using Monte Carlo simulations with random fields in a nonlinear finite element analyses. They found that three-dimensional soil variability is significant in evaluating liquefaction potential. Elkateb et al. (2003) used geostatistics and CPT data paired with stochastic Monte Carlo simulations to investigate the effect of soil variability on liquefaction hazard.

A major goal of this project is to incorporate both local and regional data in a single map. In this paper, we propose a method for quantifying liquefaction potential using statistical methods to connect local and regional data. The proposed method will provide an estimate of the extent of liquefiable materials across a geologic unit. Similar to Holzer et al. (2002), we recognize that liquefaction potential will have a distribution and that probability methods are appropriate. We also recognize that soil deposits from varying depositional environments and ages will have different distributions of liquefaction potential. In this study, we evaluate our proposed method with a dense data set in Cambridge where the surficial geology is artificial fill and the relevant substrata include marine, fluvial, and estuarine sands. Liquefaction has not been directly observed at the site; therefore the data set does not have the same sampling bias that data sets acquired for liquefaction studies might have (Toprak and Holzer 2003). However, we will not be able to calibrate our results to expected damage; rather, this study will focus on liquefaction potential as originally defined by Seed and Idriss (1971).

Method of Study

Because lithologic and engineering properties of sediments can vary significantly both laterally and with depth, it is necessary to integrate surface and subsurface data to realistically depict three-dimensional variations in liquefaction potential. The accurate extrapolation of these properties away from known data points (subsurface test borings) is an additional challenge; subsurface test boring data are usually unevenly distributed across the study area, and the natural variability of soil properties within a given geologic unit must be accounted for. We approach this issue with a combination of techniques. First, geologic units are defined on the basis of surficial mapping and geologist interpretation of subsurface test borings. This allows for a first-order division of soils into units with likely similar geologic and geotechnical properties. Next, we incorporate subsurface test boring data into the regional mapping using statistical, probabilistic, and geostatistical methods. We employ these techniques to assess the natural variability of properties within the geologic units. We investigate spatial variation as well as population distribution. Finally, we evaluate the relationship between sampling density and estimate uncertainty.

Liquefaction Calculations

Liquefaction potential refers to the relative resistance of soils to loss of strength due to an increase in pore water pressure caused by ground shaking. The degree of resistance is governed primarily by the soil's physical properties such as grain-size, density, and saturation. Liquefaction potential can be quantified according to the adjusted SPT blow count $(N_1)_{60}$ values. A significant amount of research on liquefaction potential has been based on the Seed and Idriss (1971) simplified procedure which was reviewed and updated in a workshop report summarized by Youd et al. (2001). These deterministic techniques have recently been adapted to a Bayesian framework to develop a probability of liquefaction function (Cetin et al. 2004). We have used the following probability of liquefaction function (Cetin et al. 2004) to represent liquefaction potential for a given sample

$$P[N_{1,60},\mathrm{CSR}_{eq},M_w,\sigma'_v,\mathrm{FC}] = \Phi\left[-\frac{N_{1,60}(1+0.004_1\mathrm{FC}) - 13.32\ln \mathrm{CSR}_{eq} - 29.53_3 \ln M_w - 3.70\ln\frac{\sigma'_v}{P_a} + 0.05\mathrm{FC} + 16.85}{2.70}\right] \quad (1)$$

where the input variables are corrected blowcount ($N_{1,60}$), fines content (FC); cyclic stress ratio (CSR_{eq}); moment magnitude (M_w); and vertical effective stress (σ'_v). This probability function is developed using a liquefaction field case history database where uncertainties are defined for all input parameters ($N_{1,60}$, CSR_{eq}, M_w, σ'_v, and FC) (Cetin et al. 2004). The CSR was calculated using the expected peak ground acceleration for 2% exceedence in 50 years according to the 2003 International Building Code for the maximum credible earthquake in Cambridge, Mass. The soil conditions at the site were classified as site class D. The resulting PGA for this analysis was 0.24 g. A moment magnitude of 6.5 was used for the maximum credible earthquake in Boston. The fines content was assumed based on soil classification. If a soil sample was above the measured water table for a boring (or noncohesionless), a probability of 0 was assigned. If the water table was not measured, the depth to water table was set to the ground surface which is a conservative value for this area.

Proposed Method for Liquefaction Potential

We use statistical, probabilistic, and geostatistical methods to combine geotechnical data (local liquefaction probability values based on SPT data) with regional geologic information for liquefaction potential evaluation. We treat a delineated geologic unit as a population that is sampled by subsurface test borings. Statistical methods can be used to characterize a unit; however, they assume that the unit is homogeneous. With a homogeneous assumption, there is no estimate of the location or dispersion of the liquefiable portion of the deposit. Geostatistical methods take into account spatial correlation within a deposit. The proposed method follows three steps: First, we characterize the liquefaction probability population statistically. Standard statistics (mean, standard deviation, etc.) as well as histograms are used to estimate the population variability of liquefaction probability values. Second, we use geostatistical methods to evaluate the spatial correlation of the data set to determine if the data should be treated spatially or as a homogeneous population. Third, if the data are spatially correlated, ordinary kriging is used to interpolate liquefaction probability values across the region with the goal of locally identifying zones of highly liquefiable soils. Otherwise, the population statistics are used to provide a global estimate of the distribution of liquefiable materials with a measure of estimate uncertainty.

In a case study for Cambridge, Mass., we evaluated sampling density by developing multiple random samples of varying size from the original data set. We evaluated samples comprised of 200, 100, 50, 25, and 10 borings each. For each sampling density, we took 100 random samples from the original data set. We calculated the statistics of each population sample and the mean statistics for a given boring sampling density. Using these results and assuming the population follows a binomial distribution, we determined the relationship between random sample size and uncertainty in population statistics (mean, distribution).

Step 1: Statistical Description of Liquefaction Probability

While mapping the surficial geology is a helpful first step in identifying potentially liquefiable units, a more useful quantitative analysis of the soil properties is possible when test boring data are available. With the goal of providing mapping criteria based on the statistical distribution of liquefaction probability, we explored the statistics of liquefaction probability within geologic units. Unfortunately, data from test borings are highly variable even within a single geologic unit. Therefore, many test borings are needed in order to fully classify the geologic unit—and that classification is more of a distribution than a single value.

In order to evaluate the range of liquefaction probability values within a geologic unit, we used histograms. Although the liquefaction probability calculations result in probabilities ranging from 0 to 1, the tendency is for low (near zero) and high values (near 1). As a result, the probability of liquefaction can be seen as a Bernoulli trial where outcomes are either high probability of liquefaction ($P>0.65$) or not high probability of liquefaction ($\leqslant 0.65$). The 65% cutoff for high liquefaction probability was based on recommendations by Chen and Juang (2000). By assuming that measuring the liquefaction probability of a soil sample is a Bernoulli trial, each trial within a given geologic deposit has a probability, p, of resulting in a high probability of liquefaction. If we can assume that each Bernoulli trial is independent then the sum of multiple Bernoulli trials results in a binomial distribution with parameters p and n, number of trials.

The probability mass function for the binomial distribution can be described by

$$p_X(x) = \frac{n!}{[x!(n-x)!]}p^x(1-p)^{n-x}$$

where p=probability of a soil sample having a high probability of liquefaction; and n=total number of soil samples evaluated. Here, x=number of soil samples that are found to be within the high probability category. The expected value of the binomial distribution is given by $E[x]=np$ and the variance is given by $\mathrm{Var}[x]=np(1-p)$. For large values of n, the binomial distribution can be approximated by the normal distribution. The normal approximation to the binomial distribution is given by

$$f_x(x) = \frac{1}{\sqrt{np(1-p)}}\Phi\left(\frac{x-np}{\sqrt{np(1-p)}}\right)$$

where $\Phi(\ )$=standard normal operator.

Step 2: Spatial Characterization of Liquefaction Probability

Geostatistical methods provide an analytical approach to explore spatial autocorrelation and provide an objective basis for deciding whether or not an observed spatial pattern is significantly different from random (O'Sullivan and Unwin 2003). We propose the use of the semivariogram to estimate the spatial correlation within

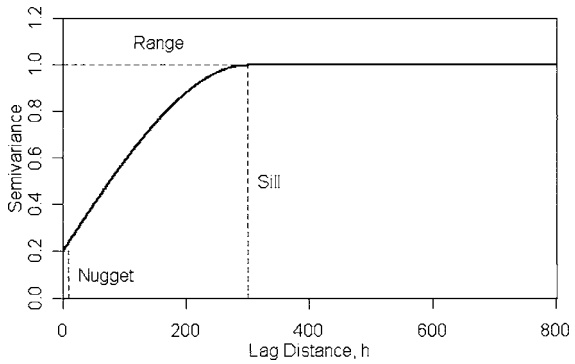


Fig. 1. Typical semivariogram with nugget, sill, and range labeled

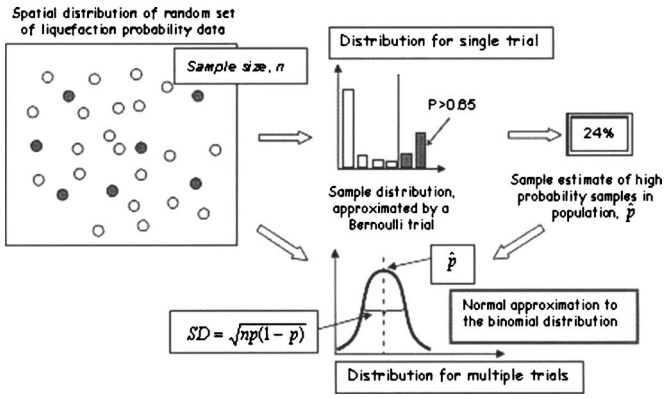


Fig. 2. Population sampling

the unit and when applicable (i.e., spatial correlation exists, see Step 3) ordinary kriging to provide an estimate of clustering of liquefiable materials.

The experimental semivariogram describes the spatial structure of the values at the sample locations, that is, the degree to which nearby locations have similar values, or do not (O'Sullivan and Unwin 2003). The semivariogram is a plot of the variance (one-half the mean squared difference) of paired sample measurements as a function of lag distance between the data points as shown in Fig. 1. The range value is the distance at which the semivariogram plateaus and corresponds to the distance over which sample points exhibit spatial autocorrelation. The plateau that the variogram reaches at the range is called the sill value. The sill value is equal to the variance of the population. The nugget value is the y-intercept of the semivariogram and provides a measure of the short-scale variability of the data set. Short-scale variability is often associated with sampling or measurement error and/or the inherent natural variability of the attribute. In an "ideal" situation, the nugget is zero, since multiple values measured at the same location are expected to be equal. However, with most natural data sets this is rarely the case. The ratio between the nugget and the sill is referred to as the relative nugget effect and provides one measure of spatial correlation.

Step 3: Local Interpolation or Global Population Characterization

The spatial structure determined in Step 2 using the semivariogram helps to determine which approach to take in characterizing the population. If the semivariogram is flat (or near flat) indicating that the relative nugget effect is near 100%, then the best estimate of the population is the global estimate. In this case, we will use the global mean and variance to describe the population using the normal approximation to the binomial distribution. If the relative nugget effect is low, than the spatial correlation can be used to characterize the population.

The global estimate of the population distribution will rely on the sample size and distribution as shown in Fig. 2. Rather than providing a mean value to characterize the liquefaction potential, the global estimate should include an estimate of the portion of the population that has a high probability of liquefaction $\hat{p}$ and an estimate of uncertainty for $\hat{p}$. Because the sample distributions were generally binomial, either high or low probability, the estimate uncertainty can be given by a confidence interval based on the binomial standard deviation $[\mathrm{SD}(x)=\sqrt{np(1-p)}]$ which is a function of sample size, n, and p.

If spatial correlation exists in the data set, we can use the semivariogram to estimate a continuous interpolated surface of liquefaction probabilities. This interpolated surface will help us decide if specific regions within a given unit are more susceptible than others. Kriging can be used to determine if the unit should be subdivided to better represent liquefaction potential. To predict values at unsampled locations, kriging methods use the semivariogram model to assign weights to the neighboring sample values. Kriging is often referred to as a "best linear unbiased estimator" (BLUE). It is "linear" because its estimates are weighted linear combinations of the available data; it is "unbiased" since it tries to have the mean residual equal to zero; and it is "best" because it aims at minimizing the variance of the errors (Isaaks and Srivastava 1989).

To estimate a value at an unsampled location, a weighted sum of the surrounding measured values is used according to the following equation

$$\hat{z}_s = w_1 z_1 + w_2 z_2 + \cdots + w_n z_n = \sum_{i=1}^{n} w_i z_i \tag{2}$$

where w_1 to w_n=set of weights applied to sample values, z_1 to z_n, in order to arrive at the estimated value, $\hat{z}_s$ (Isaaks and Srivastava 1989). The weights are assigned to surrounding values using the semivariogram model and the corresponding distance from the measured value to the prediction location. Ordinary kriging, a specific type of kriging that allows for local variation of the mean, is used in this study. Each estimate also has an associated standard error. When the standard error from the geostatistical estimate is less than the global population standard deviation, then the local estimate provides additional information and can be used to evaluate spatial clustering of high liquefaction probability soils.

The three-dimensional and anisotropic nature of soils can complicate the kriging algorithm. For this study, we propose evaluating the horizontal spatial correlation of the data set using the maximum probability of liquefaction value for each boring (within a geologic unit) in order to use two-dimensional geostatistics. An alternative to the maximum probability value would be to use three-dimensional geostatistics (Dawson and Baise 2005); however, we chose to use two-dimensional methods for

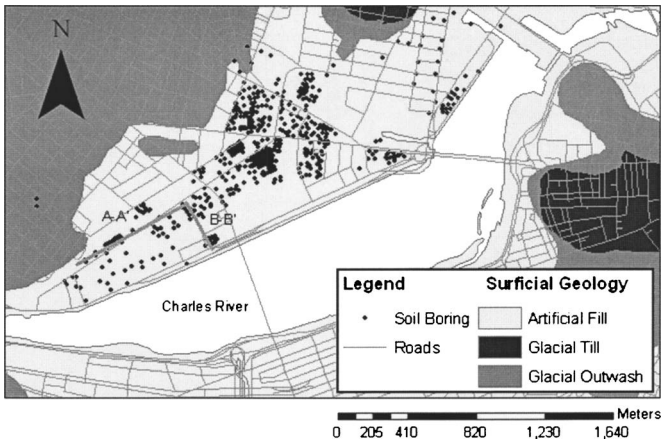


Fig. 3. Surficial geology and soil boring locations in Cambridge case study area

this application since the goal in liquefaction potential mapping is a two-dimensional map.

Cambridge Case Study: Subsurface Test Boring Database

Data from subsurface test borings were entered into an electronic database in order to facilitate relational database management and allow for the flexibility of data input. For the Cambridge study, 715 borings were collected in and near the Cambridge fill unit along the shore of the Charles River. The boring locations from the compiled database are shown in Fig. 3. The database includes both general and geologic information gathered from subsurface explorations, such as project and drilling information, date and depth of test boring, ground surface elevation, depth to groundwater, depths and descriptions of stratigraphic units and samples, SPT N values, and x-y coordinate values. The soil samples are characterized by a brief soil type (i.e., sand, silt, silty sand, clay, etc.) and a detailed sample description. Geologic descriptions varied considerably and were therefore simplified to be more consistent throughout the region. The stratigraphic units for the area are artificial fill, fluvial and estuarine deposits, marine sand, marine clay, glacial till, and bedrock.

Case Study Results (Cambridge, Mass.)

Geologic Characterization

In addition to the boring locations and surficial geology, Fig. 3 shows the locations of two cross sections shown in Fig. 4. The artificial fill in Cambridge was placed over former tidal marshlands upon completion of a granite seawall in 1890 (Woodhouse et al. 1991). The artificial fill unit ranges from 0 to 8 m in thickness across the study region with a typical thickness of 3–4.5 m.

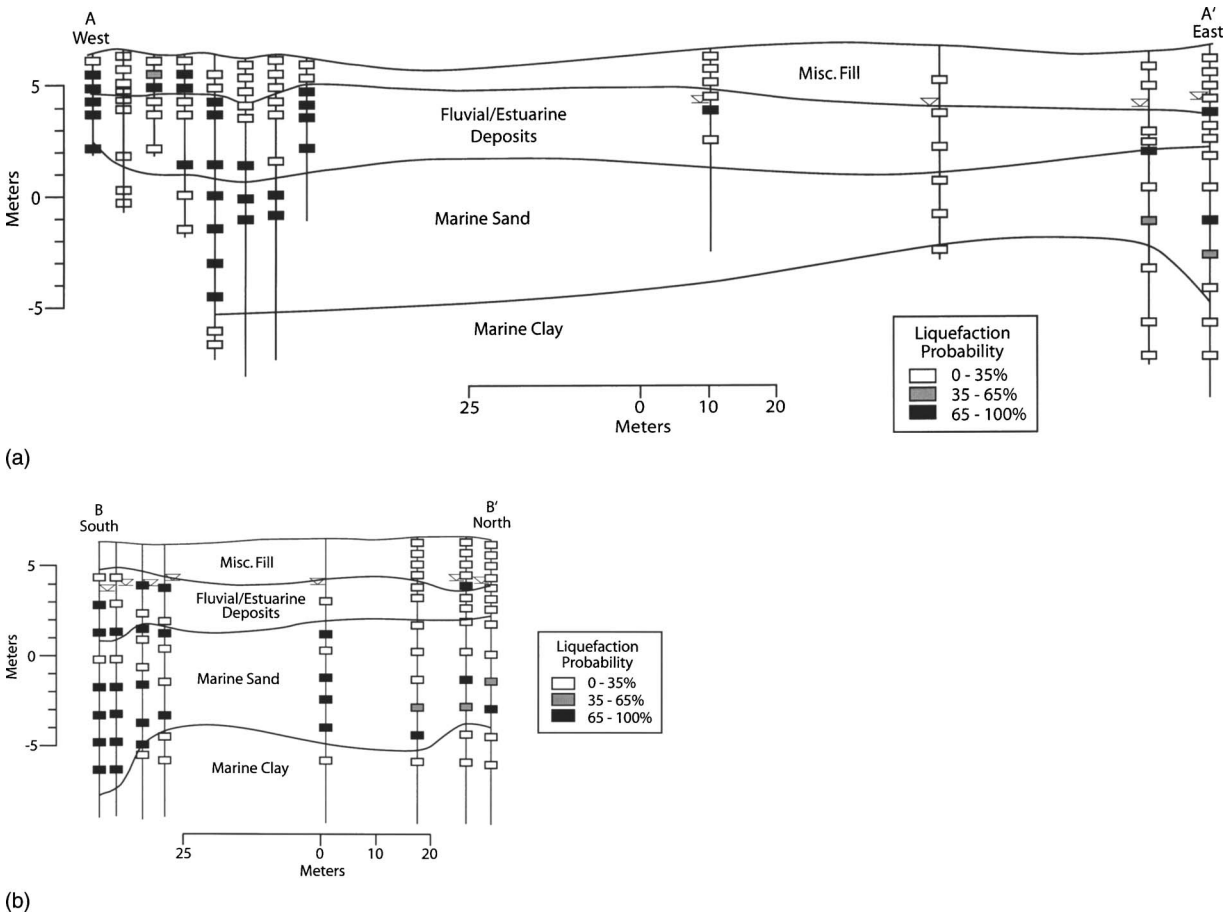


Fig. 4. Cross-section A-A′ and B-B′ as shown in Fig. 3

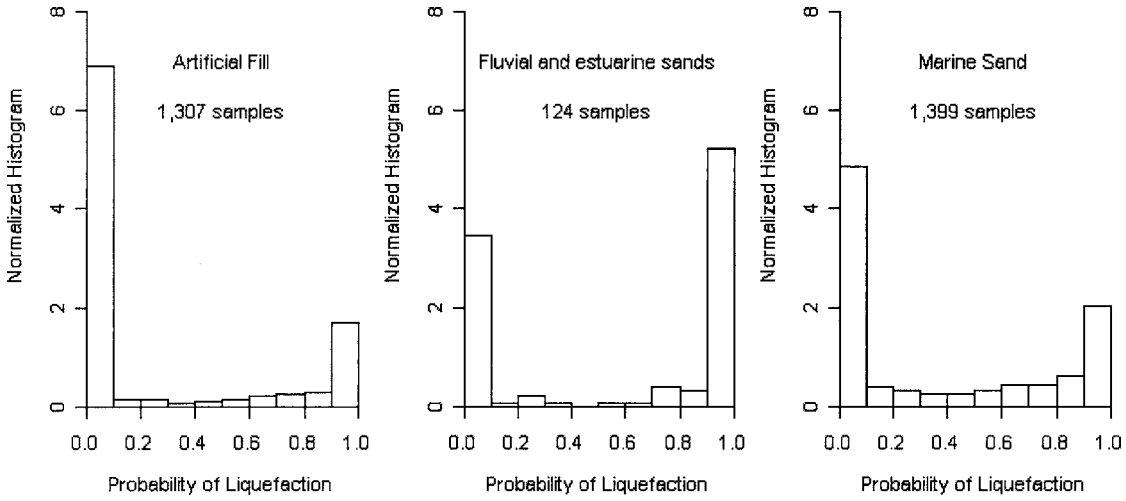


Fig. 5. Normalized histograms of liquefaction probability values for samples in the Cambridge study area

The bottom portion of the fill unit was obtained from the Charles River Basin and consists of silt, sand, and clay-sized particles. The fill was dredged from the river and pumped into the area between 1890 and 1899 (Horn and Lambe, 1964). Layers of artificial fill consisting of sand, silt, and clay-sized particles as well as building debris and trash were placed on top of the hydraulic fill at various times thereafter. The artificial fill unit in the Charles River Basin is underlain by Holocene fluvial and estuarine sediments ranging from sands and silty sands to organic silts and peats. The depositional environment of the artificial fill and Holocene fluvial and estuarine sediments paired with the relatively shallow groundwater table causes them to be potentially prone to liquefaction. Below the fluvial and estuarine sediments lies marine sand followed by a thick deposit of marine clay known as Boston Blue Clay. Glacial till and bedrock underlie the clay deposit.

Fig. 4 shows two cross sections across the site which identify the potentially liquefiable units: artificial fill, fluvial and estuarine deposits, and marine sand. The liquefaction probability categories are also shown. The water table was often found at around 2–2.5 m depth near the boundary between artificial fill and the fluvial and estuarine deposits. The liquefaction potential of the artificial fill depends strongly on the location of the ground water table. The deposit is generally loose and therefore potentially liquefiable; however, because it is not always saturated, the liquefaction potential is greatly reduced. The fluvial and estuarine deposits are generally saturated and relatively loose; therefore the liquefaction potential of the deposit depends on the material composition. The areas of high liquefaction probability are composed of silty sands whereas the low probability regions are primarily organic silts and peats. For the remainder of this paper, the fluvial and estuarine deposits are separated into fluvial and estuarine sands and fluvial and estuarine organic deposits. The liquefaction potential is only evaluated for the fluvial and estuarine sands. The marine sand varies in density and has both loose and dense regions as seen in cross section A-A′.

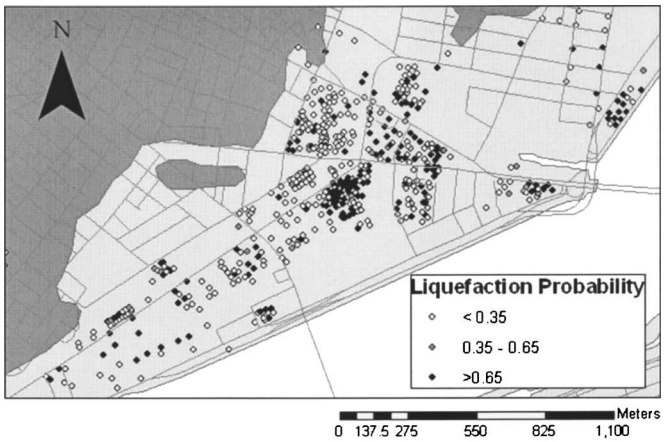
Table 1. Comparison of Sample Percentages

Geologic unit	Number of samples	High probability $p>65\%$ (%)	Medium probability $35\%<P<65\%$ (%)	Low probability $P<35\%$ (%)
Artificial fill	1,307	23.5	4.3	72.2
Fluvial and estuarine sands	124	59.7	2.4	37.9
Marine sands	1,399	33.5	9.0	57.5

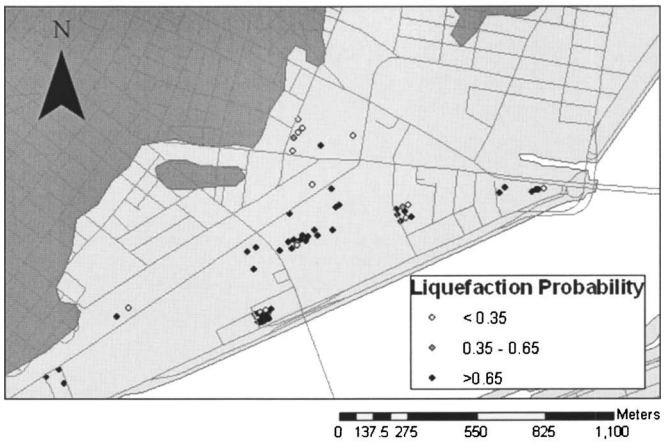
Statistical Description of Liquefaction Probability

To look at the distribution of liquefiable materials in the sands, histograms of probability of liquefaction values in the artificial fill, the fluvial and estuarine sands, and marine sand are shown in Fig. 5. The large number of zero probability values in the fill is primarily associated with unsaturated samples in the upper portion of the fill as well as some high blow counts. In the marine sand, the zero probability values are most related to high blow count values or cohesive materials. All three histograms have a bimodal shape where most samples are either zero probability or have a probability near one. As discussed above, the data can be described by a binomial distribution. We used three categories to describe the probability of liquefaction levels—high, medium, and low—according to Table 1. The cutoff for high probability is 65% and the cutoff for low probability is 35%. Only the high probability cutoff is used when the data are evaluated as a binomial distribution.

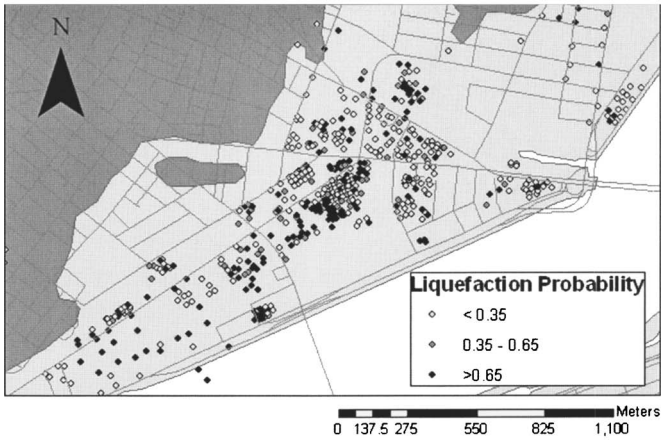
The artificial fill, although placed loosely, is primarily unsaturated, leading to a large percentage of samples with low liquefaction probability (72%). The tidal fluvial and estuarine sands are expected to be relatively loose as a result of their placement or deposition, leading to high percentages of high liquefaction probability samples (60%). If we look at the entire fluvial and estuarine deposit including cohesive soils, only 17% of the samples have a high probability for liquefaction. The marine sand's overall liquefaction potential is similar to and somewhat higher than the artificial fill (34% high for marine sand versus 24% high for fill). These results indicate that the susceptible layers should be evaluated as three distinct units. The artificial fill and the marine sand are pervasive across the area and their liquefaction potential will affect all sites within the study area. The fluvial and estuarine sands are not pervasive; however, when present, the probability of liquefaction is very high.



(a)



(b)



(c)

Fig. 6. Maximum liquefaction probability for (a) artificial fill; (b) fluvial and estuarine stands; and (c) marine sands per boring

Spatial Characterization of Liquefaction Probability

In order to further characterize the population, we needed to evaluate the spatial distribution of liquefaction probability. If soil sample values of liquefaction probability are spatially correlated then we can use geostatistical interpolation to locally map the liquefaction probability across the region. On the other hand, if the individual soil sample values of liquefaction probability are not spatially correlated, then the global characterization should proceed based on population statistics.

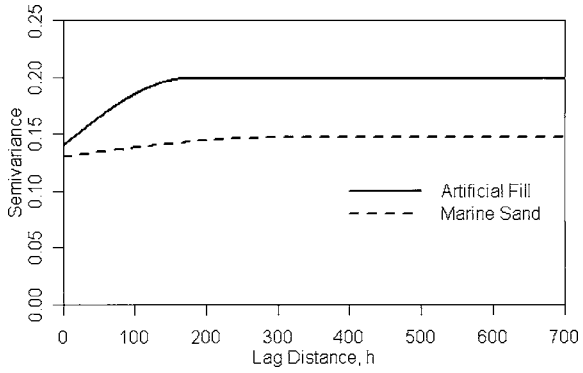


Fig. 7. Experimental semivariograms for maximum liquefaction probability per boring in artificial fill and marine sand

The spatial distribution of liquefaction probability values for the artificial fill, fluvial and estuarine sands, and marine sands are shown in Figs. 6(a–c), respectively. Because each boring has multiple samples, the plots show the maximum liquefaction probability value within the geologic layer for each boring.

Fig. 7 presents the semivariograms generated for the maximum liquefaction probability per boring for both the artificial fill and the marine sand. The fluvial and estuarine sands are too intermittent for an estimate of spatial correlation. The range value for the fill data set is approximately 150 m, indicating that beyond 150 m, there is no particular spatial structure in the data. For the semivariogram the sill value for the artificial fill is approximately 0.2. The nugget value for the semivariogram is approximately 0.14, relatively high compared to the magnitude of the sill, leading to a relative nugget effect of 70%. The minimum spacing of borings is near 10 m with few pairs while at 20 m spacing many boring pairs exist. Therefore the large nugget may include some spatial correlation that cannot be resolved below 20 m.

As compared to the fill unit, the spatial correlation of liquefaction probability in the marine sand is far less. As shown in Fig. 7, the semivariogram is relatively flat indicating little spatial correlation (relative nugget effect=88%). The nugget for the marine sand semivariogram is similar to that for the artificial fill deposit. The lower sill exhibited by the marine sand as compared to the artificial fill indicates that the overall variability of liquefaction probability is less. A comparison of the spatial structure of the two units indicates that the fill unit exhibits more spatial correlation than the marine sand based on a lower relative nugget effect. The relatively high nugget value for both units is indicative of a large amount of short-scale variation associated with the data. Since the liquefaction probability values are largely based on the density of the sample, the large nugget values for both units are likely associated with the high degree of inherent variability and error associated with blow count measurements, as well as the large degree of spatial variability in the vertical and horizontal directions. As a result, spatial correlation is limited to less than 150 m in the fill unit and there is virtually no spatial correlation in the marine sand. The relatively flat semivariogram leads to the conclusion that the homogeneous layer assumption is more appropriate for the marine sand than a spatially correlated model.

Local Interpolation or Global Population Characterization

For the fluvial and estuarine sands, the global population estimate is appropriate since the deposit is not spatially continuous. The

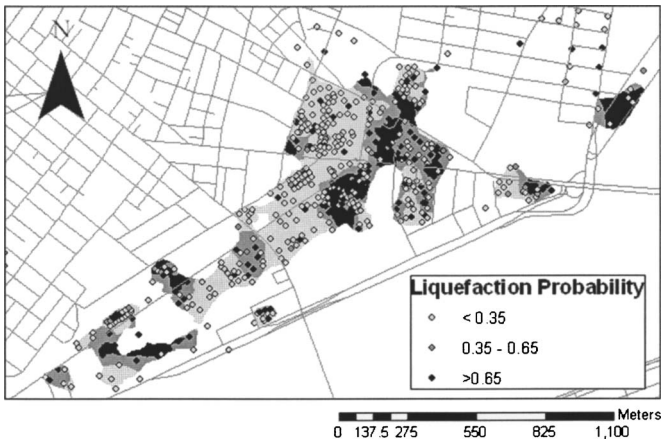


Fig. 8. Interpolated map of liquefaction probability in the artificial fill using ordinary kriging. The shaded areas on the figure show the area where the local estimate (ordinary kriging) has a smaller estimate variance than the population estimate.

marine sand has limited spatial correlation with a relative nugget of 88%, so we opted to use global characterization for the liquefaction susceptibility. The spatial correlation in the artificial fill was also limited with a relatively high nugget value and short range value (150 m) for a site that has a dimension of $2{,}500 \times 1{,}000$ m and typical boring spacing of 30–50 m. However, we evaluated the assessment of liquefaction probability using geostatistical interpolation for the artificial fill by comparing the global population variance to the estimated local variance from ordinary kriging.

Fig. 8 shows the results of ordinary kriging of liquefaction probability in the artificial fill. The colored areas in the figure show the regions where the local estimate variance is smaller than the global estimate variance. Even though the relative nugget effect was high (70%), variance reduction was achieved by the local estimates near the central part of the data set. In the regions without a local estimate, the global population estimate is more reliable than the local estimate. From Fig. 8, we can now say something about the spatial distribution of liquefiable materials in the artificial fill. Several zones of high liquefaction probability are shown on the map. These zones range from less than 100 m across to approximately 300 m across. This result indicates that the liquefiable materials are colocated within the artificial fill and will present a hazard during the design earthquake.

Now that we can see the spatial distribution of the susceptible samples, the regional classification is more clear. A summary of the population characterization for the three geologic units is given in Table 2. In the artificial fill (23.5% of samples have a high probability for liquefaction for the design earthquake), we expect several 100–300 m zones of liquefiable materials scattered throughout the fill unit. By evaluating the population using the normal approximation to the binomial distribution, we can say that the high probability estimate falls between 20.9 and 26.1% with 95% confidence. The dense data set with 1,307 samples in the artificial fill provides this tight confidence interval. According to the local spatial estimates, the entire unit will not liquefy, but liquefiable zones cover single sites as well as entire city blocks. For the estuarine and fluvial sands, the unit is not pervasive across the site. However, where it exists it is very likely to liquefy (59.7% of samples with high probability) during the design earthquake. The 95% confidence interval for the fluvial and estuarine sands (52–78%) is broader as a result of the low sample number (124). Again the zones of material are small but will likely be large enough to impact overlying structures. Finally, the marine sand is also highly liquefiable (33.5% of samples with high probability) with zones of collocated liquefiable materials that may impact sites throughout the mapped geologic unit. The 95% confidence interval for marine sand is similar to that for artificial fill (30.4–36.6%).

Table 2. Global and Local Liquefaction Susceptibility Estimates

	Global estimates		
Geologic unit	Mean, μ (%)	Standard deviation, σ (%)	Local estimation?
Artificial fill	23.5	1.3	Yes
Fluvial and estuarine sands	59.7	6.7	No
Marine sand	33.5	1.6	No

Random Sampling

The characterizations presented above include 715 borings. A collection of that many borings densely spaced is not realistic for many liquefaction mapping projects; therefore one goal of this study was to determine how many borings would be necessary to characterize a regional geologic unit. Our study is based on three geologic units present in the case study region: artificial fill, fluvial and estuarine sands, and marine sands. By selecting a random sample of borings and then statistically quantifying the results of the liquefaction probability for that population, we can validate that the normal approximation to the binomial distribution adequately characterizes the population. We will assume that the entire sample of 715 borings is enough to provide an accurate estimate of the population; and therefore we set out to confirm the relationship between sample size and confidence in estimating the distribution of the liquefaction probability. As discussed earlier, each of the random samples can be described by a Bernoulli trail and repeated trials can be described by the binomial distribution, where p fully characterizes the distribution. As discussed before, for a large number of trials, n, the binomial distribution can be approximated by the normal distribution. By taking repeated random samples from the original population, we can estimate p and confirm that $\mathrm{SD}(x)=\sqrt{np(1-p)}$ can be used to describe the estimate uncertainty. Fig. 2 illustrates how each random sample has a binomial distribution and how for repeated samples the probability density function tends towards a normal distribution which can be used to estimate the percentage of high probability samples within the unit and the estimate variance as it relates to the sample size.

In order to validate the relationship between the number of samples needed and the confidence in the estimate of population distribution for all three geologic units, random samples of decreasing size were taken from the original population of 715 borings. Fig. 9 shows the results for 100 random samples of 10 borings, 50 borings, and 200 borings as histograms of estimated p, the percentage of samples within the high probability categories for the artificial fill. The estimated binomial distribution based on a single realization is shown (using the average p for 100 random samples). Out of 100 random samples of 10 borings, the estimated population of high probability values ranges from 0 to 57% with a mean of 24%. As expected, the histogram for

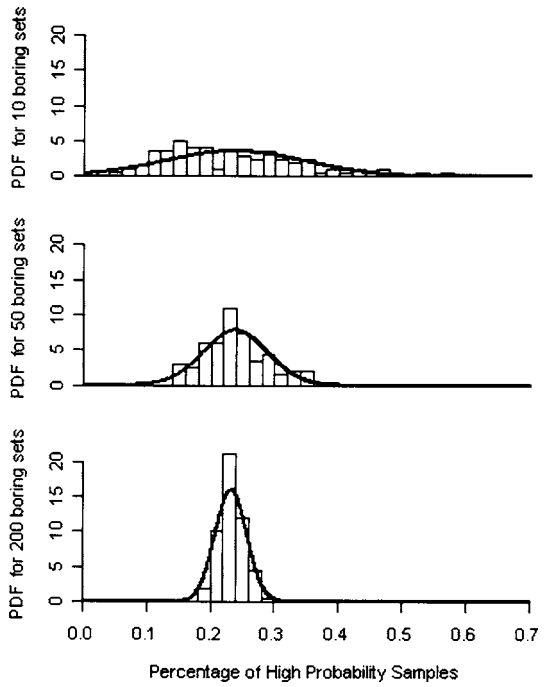


Fig. 9. Probability density functions (based on histograms) of average sample distributions for high probability of liquefaction for 100 random sets of 10, 50, and 200 borings for the artificial fill. The estimated normal approximation to the binomial distribution (based on the sample size and mean p for 100 random sets) is shown.

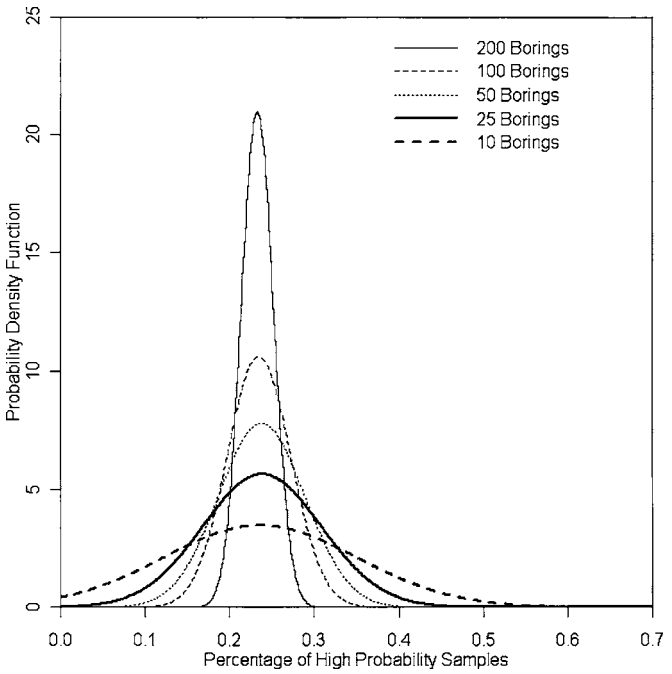


Fig. 10. Probability density functions using the normal approximation to the binomial distribution based on estimated p and n for 100 random sets of 10, 25, 50, 100, and 200 borings for the marine sand.

estimated population mean is significantly narrower with a 200 boring sample (standard deviation=1.9%) rather than a 10 boring sample (standard deviation=11%). Table 3 summarizes the results for all three geologic units for five sample sizes (200, 100, 50, 25, and 10 borings). The binomial estimates of the distribution based on the mean high probability percentage and the average number of samples per boring set are also shown. For the marine sand, the binomial estimates tend to be lower than the estimates based on 100 random sets as the population size gets smaller. Overall, the agreement between the binomial estimates and the estimates based on 100 random sets is good. For artificial fill, the standard deviation jumps from 5% with 50 borings to 12% for 10 borings.

The population of mean probability estimate can be approximated by the normal approximation to the binomial distribution, especially for the larger boring samples as shown in Fig. 9. For the marine sand, the distribution for probability categories from different samples sizes are shown in Fig. 10. The standard deviations for the estimates range from 4.2% for 100 boring samples to 10% for 25 boring samples.

In terms of characterizing the population, not all 715 borings are necessary to provide a reasonable estimate of the population. By describing the population of high liquefaction probability soil samples as a binomial distribution, a direct evaluation of estimate uncertainty can be made. It is important to be cognizant of the increase in population estimate uncertainty as it relates to decreasing sample size. The estimate uncertainty is a function of p and n, where p determines location of the estimate and both p and n affect distribution width.

Table 3. Comparison of Mean Population Estimates for 100 Random Samples and Assumed Binomial Distribution (Based on Sample Size and Sample Mean)

	Artificial fill					Marine sand				
	Results from 100 random samples			Assumed binomial distribution		Results from 100 random samples			Assumed binomial distribution	
	Number of samples	High probability $P>65\%$		High probability $P>65\%$		Number of samples	High probability $P>65\%$		High probability $P>65\%$	
Sample size	μ	μ (%)	σ (%)	μ (%)	σ (%)	μ	μ (%)	σ (%)	μ (%)	σ (%)
200 borings	369	23.3	1.9	23.3	2.5	390	33.1	2.9	33.1	2.9
100 borings	183	23.5	3.8	23.5	3.6	194	33.5	4.2	33.5	4.2
50 borings	92	23.8	5.1	23.8	5.1	99	32.8	6.4	32.8	5.8
25 borings	45	23.8	7.0	23.8	7.3	49	34.3	9.6	34.3	8.4
10 borings	19	23.6	11.4	23.6	11.1	21	34.4	13.6	34.4	12.8

Regional Mapping Criteria

One goal of this research is to integrate data from test borings with surficial geology to produce a two-dimensional map of liquefaction potential. By characterizing susceptible units according to probability distributions and spatial variability, we have gathered insight into the variability of liquefiable materials in a deposit and how best to characterize that deposit. When a limited number of borings is available, it is important to consider the limited confidence in estimating the population distribution. Using the normal approximation to the binomial distribution, the estimate uncertainty can be evaluated using n and an estimate of p. When insufficient boring data are available, criteria based on surficial geology should be used. When possible, liquefaction hazard mapping criteria should be linked with an estimate of how much of the unit is liquefiable by developing an estimate of the population probability distribution and the spatial distribution of liquefiable materials.

Summary and Conclusions

We outline a three step method for characterizing geologic units for liquefaction potential. The three steps include: statistical characterization of the distribution, spatial characterization of the population, and global and local (when possible) characterizations. When spatial correlation is sufficient, local spatial interpolation can be used to determine collocation of liquefiable materials. Either way, a global characterization can be made from the sample population. By characterizing liquefaction probabilities as a binomial distribution (high or not high), we can provide a direct evaluation of estimate uncertainty as a function of sample size (and percent high probability). In addition, we validated the effect of sample density by taking a series of 100 random boring samples for 10, 25, 50, 100, and 200 borings. Because each random sample could be evaluated as a Bernoulli trial where p is the probability of a soil sample having a high liquefaction probability, repeated samples follow a binomial distribution. The results were evaluated using the normal approximation to the binomial distribution. The binomial distribution for a particular unit and sample size provides useful information on the effect of sample size and estimate uncertainty.

The liquefaction potential in the Cambridge case study was evaluated for three distinct units. The loosest material was the fluvial and estuarine sands in the fluvial and estualine deposit below the artificial fill. This unit has a high probability of liquefaction for the design earthquake (59.7% of samples with $\sigma=6.7$); however, this unit is not pervasive across the site. The artificial fill is also highly susceptible to liquefaction with 23.5% of the samples having a high probability of liquefaction and $\sigma=1.3$. In terms of the samples in the marine sand, 33.5% of the samples have a high probability of liquefaction for the design earthquake and $\sigma=1.6$. All three deposits are expected to liquefy in significant zones during the design earthquake. These zones will range in size from a single site to a city block (as seen in Figs. 6 and 8).

The results of this study illustrate the large degree of variability of liquefaction potential within geologic units and highlight the need for an evaluation of spatial and population variability when developing regional liquefaction hazard maps. The sampling density study provides a method for evaluating the liquefaction potential estimate uncertainty associated with a given sample size. Future liquefaction hazard mapping projects need to address the issues of sample size and geologic variability in order to make the resulting maps accurate and informative.

Acknowledgments

This work was funded through USGS Award No. 02HQGR0036 and 02HQGR0040 through the USGS National Earthquake Hazard Reduction Program NEHRP program. The writers would also like to thank our friends at Haley & Aldrich, Inc. as well as Greg Knight and Michael Parkin at the Massachusetts Institute of Technology for their assistance in the collection of subsurface test boring data for this study. This paper was also significantly improved as a result of comments from three anonymous reviewers.

References

Baise, L. G., and Brankman, C. M. (2004). "Technical report for regional liquefaction hazard mapping in Boston, Massachusetts: Collaborative effort between William Lettis & Associates and Tufts University." *Rep. No. Award No. 02HQGR0036 and 02HQGR0040*, USGS, Reston, Va.

Broughton, A. T., Van Arsdale, R. B., and Broughton, J. H. (2001). "Liquefaction susceptibility mapping in the city of Memphis and Shelby County, Tennessee." *Eng. Geol.*, 62(1–3), 207–222.

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.

Chen, C. J., and Juang, C. H. (2000). "Calibration of SPT- and CPT-based liquefaction evaluation methods." *The 2000 Geotechnical Specialty Conf. Innovations and Applications in Geotechnical Site Characterization*, American Society of Civil Engineers, Reston, Va., 49–64.

Dawson, K. M., and Baise, L. G. (2005). "Three-dimensional liquefaction potential analysis using geostatistical interpolation." *Soil Dyn. Earthquake Eng.*, 25(5), 369–381.

Elkateb, T., Chalaturnyk, R., and Robertson, P. K. (2003). "Simplified geostatistical analysis of earthquake-induced ground response at the Wildlife Site, California, U. S. A." *Cognition*, 40(1), 16–35.

Holzer, T. L., Bennett, M. J., Noce, T. E., Padovani, A. C., and Tinsley, J. C. III. (2002). "Liquefaction hazard and shaking amplification maps of Alameda, Berkeley, Emeryville, Oakland, and Piedmont, California: A digital database." *Rep. No. United States Geologic Survey Open-file Rep. 02–296*, United States Geologic Survey, Reston, Va.

Holzer, T. L., Bennett, M. J., Ponti, D. J., and Tinsley, J. C. I. (1999). "Liquefaction and soil failure during 1994 Northridge earthquake." *J. Geotech. Geoenviron. Eng.*, 125(6), 438–452.

Horn, H. M., and Lambe, T. W. (1964). *Settlement of buildings on the MIT Campus*, Massachusetts Institute of Technology Report, Cambridge, Mass.

Isaaks, E. H., and Srivastava, R. M. (1989). *Applied geostatistics*, 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.

Knudsen, K. L., Sowers, J. M., Witter, R. C., Wentworth, C. M., and Helley, E. J. (2000). "Description of mapping of quaternary deposits and liquefaction susceptibility, nine-county San Francisco Bay Region, California." *Rep. No. United States Geologic Survey Open-File Report 00–444*, United States Geologic Survey, Reston, Va.

Liao, T., Mayne, P. W., Tuttle, M. P., Schweig, E. S., and Van Arsdale, R. B. (2002). "CPT site characterization for seismic hazards in the New Madrid seismic zone." *Soil Dyn. Earthquake Eng.*, 22(9–12), 943–950.

Monahan, P. A., Levson, V. M., Henderson, P., and Sy, A. (2000). "Rela-

tive liquefaction hazard map of Greater Victoria, British Columbia; Geoscience Map 2000-3a." British Columbia, Ministry of Energy and Mines, Geologic Survey Branch, ⟨http://www.em.gov.bc.ca/DL/Hazards/liquefaction.pdf⟩ (April, 5, 2004).

Monahan, P. A. et al. (1998). "Seismic microzonation mapping in Greater Victoria, British Columbia, Canada." *Geotechnical earthquake engineering and soil dynamics III, Geotechnical special publication, 75*, American Society of Civil Engineers, Seattle, 128–140.

O'Sullivan, D., and Unwin, D. (2003). *Geographic information analysis*, Wiley, Hoboken, N.J.

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

Robertson, P. K., and Wride, C. E. (1998). "Evaluating cyclic liquefaction potential using the cone penetration test." *Can. Geotech. J.*, 35(3), 442–459.

Seed, H. B., and Idriss, I. M. (1971). "Simplified procedure for evaluating soil liquefaction potential." *J. Soil Mech. Found. Div.*, 97(9), 1249–1273.

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

Woodhouse, D. et al. (1991). "Geology of Boston, Massachusetts, United States of America." *Assn. Eng. Geolog., Bull.*, 28(4), 375–512.

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.

Youd, T. L., and Perkins, D. M. (1978). "Mapping liquefaction-induced ground failure potential." *J. Geotech. Eng. Div., Am. Soc. Civ. Eng.*, 104(4), 433–446.