UCLA
UCLA Previously Published Works

Title

A latent Gaussian process model for the spatial distribution of liquefaction manifestation

Permalink

https://escholarship.org/uc/item/2f93p946

Journal

Earthquake Spectra, 39(2)

ISSN

8755-2930

Authors

Bullock, Zach
Zimmaro, Paolo
Lavrentiadis, Grigorios
et al.

Publication Date

2023-05-01

DOI

10.1177/87552930231163894

Copyright Information

Peer reviewed

EARTHQUAKE SPECTRA

A latent Gaussian process model for the spatial distribution of liquefaction manifestation

Zach Bullock[1], Paolo Zimmaro M. EERI[2,3], Grigorios Lavrentiadis M. EERI[1,3], Pengfei Wang M. EERI[3], Olaide Ojomo M. EERI[4], Domniki Asimaki M. EERI[1], Ellen M Rathje M. EERI[4], and Jonathan P Stewart M. EERI[3]

Abstract

This paper presents a model for distributing zones of liquefaction and nonliquefaction for use in regional liquefaction risk analysis. There are two broad methodologies that have been used to evaluate liquefaction risk on the regional scale: (a) application of site-specific procedures using soil properties inferred from geology, or (b) application of geospatial proxies for liquefaction. The first approach will tend to predict similar liquefaction probabilities across broad areas with similar geology, water table depths, and shaking intensities. The second approach yields the probability of liquefaction, which can be interpreted as the portion of the area affected by liquefaction ($\%A_{liq}$). Neither approach, however, gives an informed prediction of the spatial distribution of liquefaction and the resulting displacements, which are particularly important for assessments of seismic risk for spatially distributed infrastructure systems. We propose a methodology for incorporating spatial correlation into a geospatial proxy for liquefaction to create maps of liquefaction and nonliquefaction for a given earthquake scenario. First, we describe a latent Gaussian process that is assumed to govern the spatial distribution of liquefaction. Next, a database of empirical observations of liquefaction is used to obtain the coefficients that describe that latent Gaussian process. The proposed model yields random realizations of maps of liquefaction and nonliquefaction conditioned on a map of $\%A_{liq}$. Such maps can be used to constrain the area over which displacements are estimated using soil properties inferred from geology and are therefore a critical component in reducing bias in assessments of liquefaction risk at the regional scale.

[1]California Institute of Technology, Pasadena, CA, USA
[2]University of Calabria, Rende, Italy
[3]University of California, Los Angeles, Los Angeles, CA, USA
[4]The University of Texas at Austin, Austin, TX, USA

Corresponding author:
Domniki Asimaki, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA.
Email: domniki@caltech.edu

Introduction

Background

Liquefaction has caused extensive damage to land, structures, and infrastructure systems in past earthquakes (e.g. Bray et al., 2014; Hamada et al., 1996; Van Ballegooy et al., 2014). Accordingly, liquefaction risk analysis is a critical tool for estimating the potential damage in future earthquakes and making decisions regarding mitigation strategies.

Although liquefaction consequences can be influenced by highly site-specific conditions (e.g. Beyzaei et al., 2020; Cubrinovski et al., 2019), the primary conditions necessary for liquefaction to occur are driven by geology and geography. Geology (i.e. age and depositional environment) determines whether granular, liquefaction-susceptible soils are present (e.g. Baise et al., 2006; Lewis et al., 1999; Obermeier et al., 1990; Youd and Hoose, 1978; Youd and Perkins, 1978). Geography (i.e. topography and proximity to water) determines the depth of the water table, *wtd* (e.g. Fan et al., 2013). These two factors combined determine whether saturated, liquefaction-susceptible soils are present at a given location. Because geology and geography determine the distribution of deposits that are prone to liquefaction and whether they are saturated, liquefaction tends to threaten multiple locations across a given region (i.e. on the scale of geologic units or bodies of water). Furthermore, because these factors are not rapidly evolving, liquefaction will tend to repeatedly affect a given region in any sufficiently strong earthquake if saturated, susceptible soil deposits are present (e.g. Quigley et al., 2013).

Liquefaction risk assessment at the regional scale is a key component in forecasting the impacts of liquefaction in future earthquakes because (a) liquefaction can occur over broad regions, and (b) the systems affected by liquefaction are likewise distributed over broad regions. Furthermore, the response and performance of systems affected by liquefaction may depend on the nature of its spatial distribution. For instance, we expect that a buried pipe would respond differently to crossing a single continuous 50 m wide polygon containing liquefaction-induced deformations than to crossing five discontinuous 10 m wide polygons. Figure 1 shows these two scenarios schematically. The length of the pipe exposed to liquefaction-induced deformations is the same in both cases (50 m), but the outcomes regarding pipeline performance are likely to be different.

The majority of procedures for liquefaction risk analysis are developed for use in a site-specific context (e.g. Boulanger and Idriss, 2016; Cetin et al., 2018; Moss et al., 2006). Applying these procedures in regional analysis requires inferring subsurface conditions that are typically too sparse to use directly (e.g. Holzer et al., 2006; Youd and Perkins, 1978). This approach will tend to give similar estimates of liquefaction hazard across broad areas of similar geology, conditioned on *wtd* and the shaking intensity in the given earthquake scenario. Recently, Greenfield and Grant (2020) used Gaussian fields to map site-specific observations into 3D independent random fields of groundwater depth, soil plasticity, and penetration resistance for different geologic units, and integrated the resulting models in a regional-scale probabilistic framework for liquefaction triggering. Still, their approach
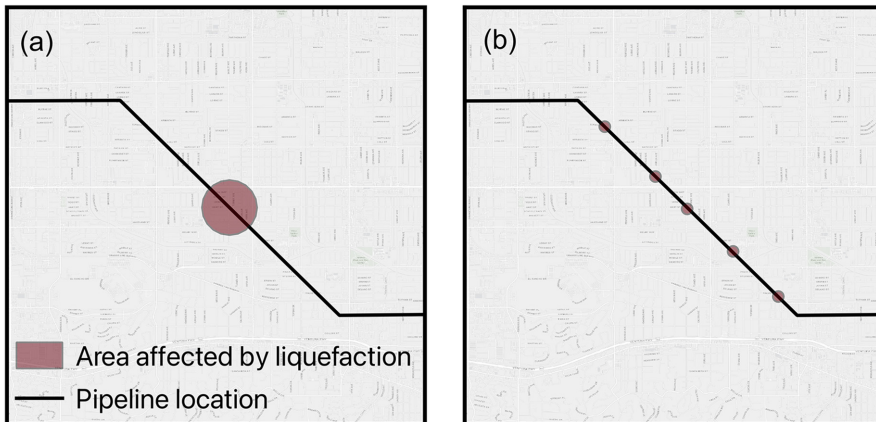
Figure 1. Schematic maps showing (a) one 50-m wide polygon of liquefaction-affected area and (b) five 10-m wide polygons of liquefaction-affected area relative to a hypothetical pipeline.

requires statistically significant site characterization data, which is rarely available on regional scales. On the opposite length scale end, geospatial proxies for liquefaction hazard (e.g. Rashidian and Baise, 2020; Zhu et al., 2015, 2017) provide an alternative tool for regional liquefaction risk assessment that uses inputs available from remote sensing, such as the topography at a given location, to predict the probability that a specific location will be affected by liquefaction (P_{liq}). Such proxies may still require inference of certain inputs or estimation using global models (e.g. *wtd*), but the uncertainty associated with these inferences is expected to be smaller than that associated with subsurface soil conditions. P_{liq} can be interpreted as the percent of the area covered by liquefaction ($\%A_{liq}$), as in the USGS ground failure product (Allstadt et al., 2021). Applying site-specific procedures regionally by inferring subsurface conditions also gives P_{liq}, and can account for both variability around those procedures' outputs as well as uncertainty associated with the inference of soil properties. Regardless of which of these two approaches (applying site-specific procedures with inferred inputs or applying a geospatial proxy with remotely available inputs) is used, no method exists for assigning liquefaction and nonliquefaction to specific locations aside from random assignment.

Objective

In this study, we develop an empirical procedure for creating maps of areas of liquefaction and nonliquefaction based on a latent Gaussian process. This procedure extends existing geospatial proxies to allow the assignment of liquefaction and nonliquefaction across regions while reflecting the spatial correlation structure observed in liquefaction manifestations in past earthquakes. The proposed methodology allows continuous spatial variation in $\%A_{liq}$.

Methods

Geospatial proxy for liquefaction

Geospatial proxies such as Zhu et al. (2015, 2017) are commonly used to analyze liquefaction risk at a regional scale. Zhu et al. (2017) proposed separate models for use in coastal

and non-coastal regions, where the coastal model is applied at sites less than 20 km from the coast. Both the coastal and non-coastal models apply as independent variables the time-averaged shear-wave velocity in the top 30 m at a given location ($V_{S,30}$), the total annual precipitation at that location (*precip*), and the peak ground velocity at that location in a given earthquake (*PGV*). The coastal model also requires the distance from the location of interest to the nearest river (d_r) and the distance from that location to the coast (d_c). The non-coastal model instead uses *wtd* at the location of interest and the distance from that location to the nearest body of water (d_w).

The Zhu et al. (2017) models give the probability of liquefaction (P_{liq}) at the location of interest in a given earthquake as their output. These models interpret P_{liq} as being equivalent to %A_{liq} based on the assumptions and procedure used in their development. The inputs for Zhu et al. (2017) can be obtained from broad, global models; specialized, regional models; or detailed measurement at the sites of interest. In this study, we apply regional models where available and global models elsewhere.

Wald and Allen (2007) developed a globally-applicable model for $V_{S,30}$ based on topography. Thompson et al. (2014) produced a $V_{S,30}$ map of California that incorporates observed values of $V_{S,30}$ as well as geology and topography. Foster et al. (2019) and Ahdi et al. (2017) applied a similar approach to develop maps of $V_{S,30}$ in New Zealand and the Pacific Northwest, respectively. Multiple studies have produced global maps of *precip*, including Daly et al. (1997), Kubota et al. (2007), and Adler et al. (2018). *PGV* can be obtained from shake maps for backward analysis of past earthquakes (e.g. Allen et al., 2008) or predicted using ground motion equations (e.g. Boore et al., 2014; Campbell and Bozorgnia, 2014) or seismic hazard maps (e.g. Petersen et al., 2020) for forward analysis. Fan et al. (2013) developed a globally applicable model for predicting *wtd* based on observations at over 1.6 million locations and used hydrological modeling to address gaps in the data. Region-specific models of *wtd* have been developed for some areas (e.g. Befus et al., 2020 for the California coast and Westerhoff et al., 2018 for New Zealand). Finally, d_r, d_c, and d_w can be obtained for a given location by calculating distances between that location and coastlines, rivers, and lakes.

Gaussian processes

Gaussian processes are collections of random variables where each variable corresponds to a location and/or a point in time. Essentially, a Gaussian process is a random field where each variable has a normal distribution. The distribution of a Gaussian process is then the joint distribution of all those (essentially infinite) random variables, and as such, it is a distribution over functions with a continuous domain. Gaussian processes can thus be seen as an infinite-dimensional generalization of multivariate normal distributions (Dudley, 1967).

Gaussian processes are said to be *stationary* if their properties (i.e. mean and standard deviation) remain constant over space and time. Gaussian processes are said to be *ergodic* if their variation over space and their variation over time are interchangeable (i.e. if their properties can be obtained using a sample over many locations at a single point in time or a sample at a single location over many points in time). The concept of ergodicity is frequently invoked in ground motion modeling (e.g. Anderson and Brune, 1999; Lavrentiadis et al., 2022), where the variability of ground motion across space (i.e. at multiple stations in multiple earthquakes in different locations) is assumed to reflect the variability of ground motion at a single location going into the future.

Describing a Gaussian process requires first obtaining a covariance function that describes the correlation among locations distributed in space. Formulating empirical semivariograms and fitting theoretical functions to the results is one approach to obtaining the covariance function. For a set of spatially distributed observations of some variable, z, the empirical semivariance, $\gamma(h)$, is given by Equation 1, where z_u is the value of z at location u and z_{u+h} is the value of z at a location separated from u by a distance of h:

$$\gamma(h) = \frac{1}{2} E[(z_u - z_{u+h})^2] \tag{1}$$

Theoretical semivariograms, denoted as $\hat{\gamma}(h)$, are fit to $\gamma(h)$ using nonlinear regression. This notation implies that we are treating the empirical semivariogram as the observed "truth," and the theoretical semivariogram as an estimator which we determine by regression. Several functional forms can be used for $\hat{\gamma}(h)$ that are valid for forward analysis using kriging, such as the exponential and Gaussian forms (Chiles and Delfiner, 2009; Isaaks and Srivastava, 1989), while linear combinations of valid semivariograms are also valid (Journel and Huijbregts, 1978).

We here model the theoretical semivariogram using a nested exponential model after Markhvida et al. (2018), given by Equation 2. Note that the sill value ($\gamma(100\,\text{m})$) of the empirical semivariograms is 1 because $z(s)$ has a standard normal distribution. The nugget ($\gamma(0\,\text{m})$), on the other hand, was fixed to zero because manifestations of liquefaction from a single earthquake on the same site (i.e. zero separation distance) are perfectly correlated:

$$\hat{\gamma}(h) = 1 - c_1 \exp\left[-3\left(\frac{h}{\ell_1}\right)\right] - c_2 \exp\left[-3\left(\frac{h}{\ell_2}\right)\right] \tag{2}$$

On the same time, this functional form for the semivariogram allows us to model separate short-distance and long-distance correlation structures, where the respective correlation lengths are ℓ_1 and ℓ_2. This feature of the form in Equation 2 allows the model to capture the influence of the size of individual manifestations (i.e. the scale of sand boils or clusters of sand boils) with ℓ_1 and the influence of the size of liquefiable deposits with consistent properties with ℓ_2. The relative importance of the short- and long-distance correlation is determined by c_1 (the weight on the short-distance correlation). c_2 is equal to $1 - c_1$ such that the semivariance at large distances is equal to unity.

Finally, we also provide the functional form for a nested Gaussian model, per Equation 3. This functional form is mean-squared differentiable, while the nested exponential (Equation 2) is not, which makes it more suitable for some risk analysis applications that adopt methods based on level-crossing statistics (Zimmaro et al., 2019). Figure 2 shows theoretical semivariograms of the forms given by Equations 2 and 3 for a given set of parameters ($c_1 = 0.7$, $c_2 = 0.3$, $\ell_1 = 20$ m, and $\ell_2 = 50$ m). The nested exponential form has a steeper slope for separating distances close to 0 m, but a shallower slope at long separating distances. Both forms reach $\gamma \approx 0.97$ at $h = \ell_2$.

$$\hat{\gamma}(h) = 1 - c_1 \exp\left[-3\left(\frac{h}{\ell_1}\right)^2\right] - c_2 \exp\left[-3\left(\frac{h}{\ell_2}\right)^2\right] \tag{3}$$
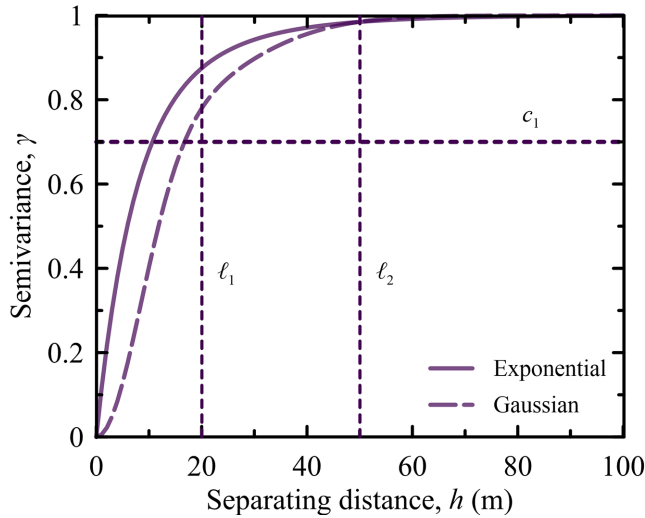
Figure 2. Nested exponential and Gaussian semivariograms with $c_1 = 0.7$, $c_2 = 0.3$, $\ell_1 = 20$ m, and $\ell_2 = 50$ m.

For the case of a standard normal random field (i.e. one with unit variance), the estimated correlation between the field values at two sites separated by h, denoted as $\hat{\rho}(h)$, is given by Equation 4:

$$\hat{\rho}(h) = 1 - \hat{\gamma}(h) \tag{4}$$

Liquefaction manifestation as a latent Gaussian process

This section interprets the spatial distribution of surficial liquefaction manifestation as being governed by a latent (i.e. hidden) Gaussian process. In the context of latent variable modeling (see, for example, Skrondal and Rabe-Hesketh, 2007), this Gaussian process is a *hypothetical construct*, meaning that it cannot be directly observed. Rather, we must measure it indirectly based on observations of liquefaction manifestation and our assumptions about its behavior.

We denote the sites of interest as s, where s_i is the i-th site; we denote surficial manifestations at the sites as m, where $m_i = 1$ if there are manifestations at s_i and zero otherwise, and we denote the latent Gaussian process as Z and a given realization of Z as $z(s)$ (denoted in Figure 3 as field value), whose cumulative probability of being below a threshold is equivalent to a positive observation of liquefaction. Our goal is to define Z such that Equation 5 is true at all s_i:

$$P(Z \leqslant z(s_i)) = \%A_{liq,i} \tag{5}$$

Liquefaction manifests at the surface at s_i (i.e. $m_i = 1$) if the inequality given by Equation 6 holds, where $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF) and $\%A_{liq,i}$ is the portion of the area liquefied estimated at s_i:

$$\Phi(z(s_i)) \leqslant \%A_{liq,i} \tag{6}$$

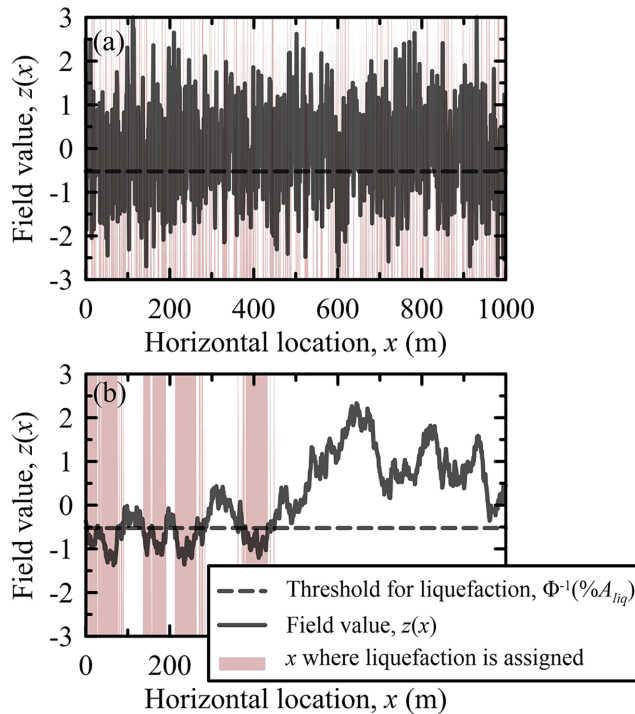If the inequality in Equation 6 holds, Equation 5 is also true.

Figure 3. A hypothetical horizontal cross-section showing areas assigned liquefaction according to Equation 6 with a constant $\%A_{liq}$ of 30% (a) with no correlation structure and (b) with a correlation structure.

This formulation assumes that the latent Gaussian process is stationary (i.e. we assume that Z is a standard normal variable at each location in s). Moreover, it assumes that P_{liq} at a given site in future earthquakes is equivalent to P_{liq} at many such sites in past earthquakes. This assumption is akin to the "ergodic assumption" commonly made in ground motion prediction (e.g. Anderson and Brune, 1999). Figure 3 demonstrates how this inequality determines the areas where liquefaction is assigned across a horizontal cross-section for a constant $\%A_{liq}$ of 30%. When z is less than $\Phi^{-1}(\%A_{liq})$, liquefaction is assigned. Two examples are shown: one without a correlation structure and one with a correlation structure like those derived subsequently.

In forward applications (i.e. simulating the effects of future earthquakes), a large number of realizations of $z(s)$ ensures that the probability of an individual site manifesting liquefaction is set to its $\%A_{liq,i}$ (per Equation 6). Furthermore, for a dense enough number of sampling points, the proportion of sites with a given $\%A_{liq,i}$ that manifest liquefaction in a single realization of $z(s)$ will also be $\%A_{liq,i}$. Both of these properties are desirable for simulating maps of the liquefied area for regional risk analysis because they are consistent with the assumption that P_{liq} and $\%A_{liq}$ are equivalent.

Empirical data

Sources

This study uses observations of surficial liquefaction manifestations in past earthquakes to produce a model for their spatial correlation. Table 1 summarizes the data used. The

Table 1. Summary of the empirical data used in this study

Earthquake	Moment magnitude, M_W	Liquefaction points
1949 Olympia, WA[a]	7.1	151
1965 Puget Sound, WA[a]	6.7	223
1989 Loma Prieta, CA[b]	6.9	108
1994 Northridge, CA[c]	6.7	34
1999 Chi-Chi, Taiwan[d]	7.7	170
2001 Nisqually, WA[e]	6.8	68
2003 San Simeon, CA[f]	6.6	10
2008 Wenchuan, China[g]	7.9	116
2010 El-Mayor Cucapah, CA[h]	7.2	72
2010 Darfield, New Zealand[i]	7.1	1315
2011 Christchurch, New Zealand[i]	6.2	4786
2015 Gorkha, Nepal[j]	7.8	12
2016 Valentine's Day, New Zealand[i]	5.7	154
2019 Ridgecrest, CA[k]	7.1	190

[a]Chleborad and Schuster (1990) and Rogers et al. (1998); [b]Tinsley et al. (1998); [c]Stewart et al. (1996); [d]Chu et al. (2004); [e]Bray et al. (2001); [f]Holzer et al. (2005); [g]Cao et al. (2010); [h]McCrink et al. (2011); [i]Geyin et al. (2020); [j]Moss et al. (2015); [k]Zimmaro et al. (2020).

data were collected from Schmitt et al. (2017), Zimmaro et al. (2020), and Geyin et al. (2020).

The data from each earthquake detailed in Table 1 consist of the latitude and longitude coordinates of s, as well as m_i for all s_i. The Fan et al. (2013) water table model and the Daly et al. (1997) precipitation model are used to estimate *wtd* and *precip* in all regions, respectively. Fan et al. (2013) is used even when post-earthquake cone penetration test (CPT) data includes water table estimates so that *wtd* is consistent between the liquefaction (i.e. observed) and nonliquefaction (i.e. sampled) points in a given earthquake. Maps of *PGV* for all events are obtained from the ShakeMap Atlas Allen et al. (2008). The Thompson et al. (2014), Ahdi et al. (2017), Foster et al. (2019), and Kwok et al. (2018) models for $V_{S,30}$ are used for all locations in California, Washington, New Zealand, and Taiwan, respectively. The Wald and Allen (2007) model for $V_{S,30}$ is used in all other regions.

Per Table 1, there are limited instances of data from multiple earthquakes in the same region, with the exceptions of the Canterbury, New Zealand, and Puget Sound, Washington regions, each of which has three earthquakes in the database. Although there are five California earthquakes in the database, none affected the same region.

Balance and completeness

For the purposes of this study, the completeness and balance of the datasets are important. We call a dataset *complete* if it includes *all* locations with liquefaction manifestations in the corresponding earthquake. We call a dataset *balanced* relative to $\%A_{liq}$ if the ratio of the number of liquefaction and nonliquefaction points aligns with $\%A_{liq}$ (e.g. if 4 out of 10 s_i in a dataset have $m_i = 1$ where $\%A_{liq,i} = 0.4$ for all i). Equation 7 defines the balance of the database (denoted as B), where n is the number of sites in the database.

$$B = \frac{\sum_i m_i}{n} \tag{7}$$

We calculate $\%A_{liq,i}$ using Zhu et al. (2017) for all sites in all earthquakes. Based on the calculated $\%A_{liq,i}$, the databases in Table 1 are initially imbalanced relative to estimates of $\%A_{liq}$ for the corresponding earthquake. Therefore, we sample nonliquefaction points within the study area following an approach similar to that of Zhu et al. (2017). These points are located at least 20 m and at most 1000 m away from existing liquefaction points. This strategy assumes that the field observations include all surficial liquefaction manifestations that were generated in the given earthquake (i.e. that the datasets are complete). Put differently, this approach assumes that, for the studies listed in Table 1, liquefaction manifested at the ground surface *only* at the reported liquefaction points, and that any randomly sampled point is a true nonliquefaction point. We are generally confident in this assumption based on experience with post-earthquake reconnaissance. However, we should note that we are more confident in this assumption for more recent earthquakes where the reconnaissance benefited from geolocation and tracking of team members (e.g. Brandenberg et al., 2019). Other factors can also affect the validity of this assumption, such as certain areas being inaccessible to reconnaissance teams (e.g. private property or military installations). We do not address possible differences in our confidence in the earthquakes presented in Table 1 in this study.

Empirical approach to characterizing the latent Gaussian process

Let us now suppose that the latent Gaussian process for Z described above, exists. If Z were directly observable (i.e. if we could measure its value from past event data), the development of the geospatial model would have been straightforward. However, in this case, the available empirical dataset is binary (manifestation or no manifestation) and Z, which controls the manifestation of liquefaction, is hidden. For this reason, to develop the geospatial model for Z, we first perform a Monte Carlo simulation according to Equations 8 and 9, generating multiple realizations that both satisfy the underpinning assumptions (Equation 6) and are consistent with the empirical dataset. Using the Monte Carol realizations, we then determine the spatial correlation structure of $z(\mathbf{s})$: for each observation of manifestation or no manifestation of liquefaction in the empirical dataset, the random realizations are generated with:

$$u_i = \begin{cases} U(0, \%A_{liq,i}), & m_i = 1 \\ U(\%A_{liq,i}, 1), & m_i = 0 \end{cases} \tag{8}$$

$$z(s_i) = \Phi^{-1}(u_i) \tag{9}$$

where, $U(a, b)$ is a realization of a uniform random variable bounded from a to b; and $\Phi^{-1}(.)$ is the inverse standard normal CDF.

As mentioned above, Equations 8 and 9 create sets of random realizations of the latent variable $z(\mathbf{s})$ for all sites, based on: (a) whether or not liquefaction has been observed at a given site, and (b) the probability of liquefaction at that site (P_{liq}). Up to this point, no assumptions about the spatial distribution of Z have been made. Intuitively, if a site has liquefied, the range of $z(s_i)$ is not well constrained in the Monte Carlo simulations if P_{liq} is large, while $z(s_i)$ can only span a small range if P_{liq} is small. The opposite holds if the site has not liquefied. For instance, if a site has liquefied ($m_i = 1$), the first branch of Equation 8 gives the range of random realizations of u_i that increase proportionally to $\%A_{liq,i}$, which, in Zhu et al. (2017), is equivalent to P_{liq}. The opposite behavior is expressed in the second branch of Equation 8 for sites without liquefaction ($m_i = 0$). In both cases, Equation 9 is

then used to compute the corresponding random realizations for the latent variable $z(s_i)$ from which the geospatial model is derived.

Equations 8 and 9 enforce the assumption stated by the inequality in Equation 6. From the Probability Integral Transform (Casella and Berger, 2021), the $z(s)$ cumulative distribution ($\Phi(z(\mathbf{s}))$) follows a uniform distribution, from 0 to 1, if $z(\mathbf{s})$ is a standard normal random field (i.e. to have zero mean, unit variance, and to pass a test for normality such as Lilliefors, 1967). However, the sampled $\Phi(z(\mathbf{s}))$ is only uniformly distributed under certain conditions regarding the balance in the database (i.e. B per Equation 7) and the values of $\%A_{liq,i}$. Figure 4a demonstrates that a population of $z(s)$ is not normally distributed for $B > \%A_{liq}$. Figure 4b shows the variance of $z(s)$ for ranges of B and $\%A_{liq}$. Along the line defined by $B = \%A_{liq}$, the variance of $z(s)$ is approximately 1.0.

We address the issue of imbalanced data by sampling additional nonliquefaction points such that B is approximately equal to the average value of $\%A_{liq,i}$ in the database, which pushes $z(s)$ toward a standard normal distribution. We require that the sampled points are between 20 m and 1000 m from points in the database. Sampling additional nonliquefaction points assumes that surficial manifestations were not present in any location *not* identified as a liquefaction point.

We then formulate the empirical semivariogram of each $z(s)$ and fit its theoretical semivariogram using nonlinear regression with Equations 2 or 3 as the functional form. Because $z(s)$ has a standard normal distribution, the sill value of the empirical semivariograms is 1, and we assume that $c_1 + c_2 = 1$. Finally, because each $z(s)$ is created using realizations of uniform random variables (i.e. $\mathbf{u}$ per Equation 8), we repeat this process 1,000 times for each earthquake using Monte Carlo simulation and analyze the pooled results. Each Monte Carlo simulation is consistent with the assumed behavior of the latent Gaussian process as well as the $\%A_{liq_i}$ calculated using Zhu et al. (2017) at each site. Applying models in this framework derived from the empirical data in Table 1 to analyses of future earthquakes requires assuming some degree of ergodicity (i.e. that the $z(s)$ obtained for past earthquakes reflect future earthquakes in those regions or elsewhere).

Results

Table 2 provides the parameters for the fitted semivariograms (Equations 2 or 3 for the nested exponential and nested Gaussian models, respectively) calculated using 1,000 simulations of the latent Gaussian process for each individual earthquake, as well as for the earthquakes pooled together according to region.

Figure 5 shows the empirical and fitted nested exponential semivariograms for each earthquake in Table 1 grouped according to their regions. The results presented in Table 2 and Figure 5 indicate that for most earthquakes, the overall correlation length of liquefaction manifestation (i.e. the h where $\hat{\gamma}(h)$ is approximately 0.95 and $\hat{\rho}(h)$ is below 0.05) is between 300 and 500 m. These results are consistent with estimates of the correlation length of indices such as the liquefaction potential index (*LPI*; Iwasaki et al. (1978); $\ell \sim$ 500 m) and CPT cone resistance and sleeve friction ($\ell \sim$ 300 m) per Wang et al. (2017). Figure 6 shows the fitted and empirical semivariograms for each earthquake together, along with those of the results pooled together according to region.

The fitted and empirical semivariograms are similar for most earthquakes in the study. Two outliers (the 2001 Nisqually and 2015 Gorkha earthquakes) have longer correlation
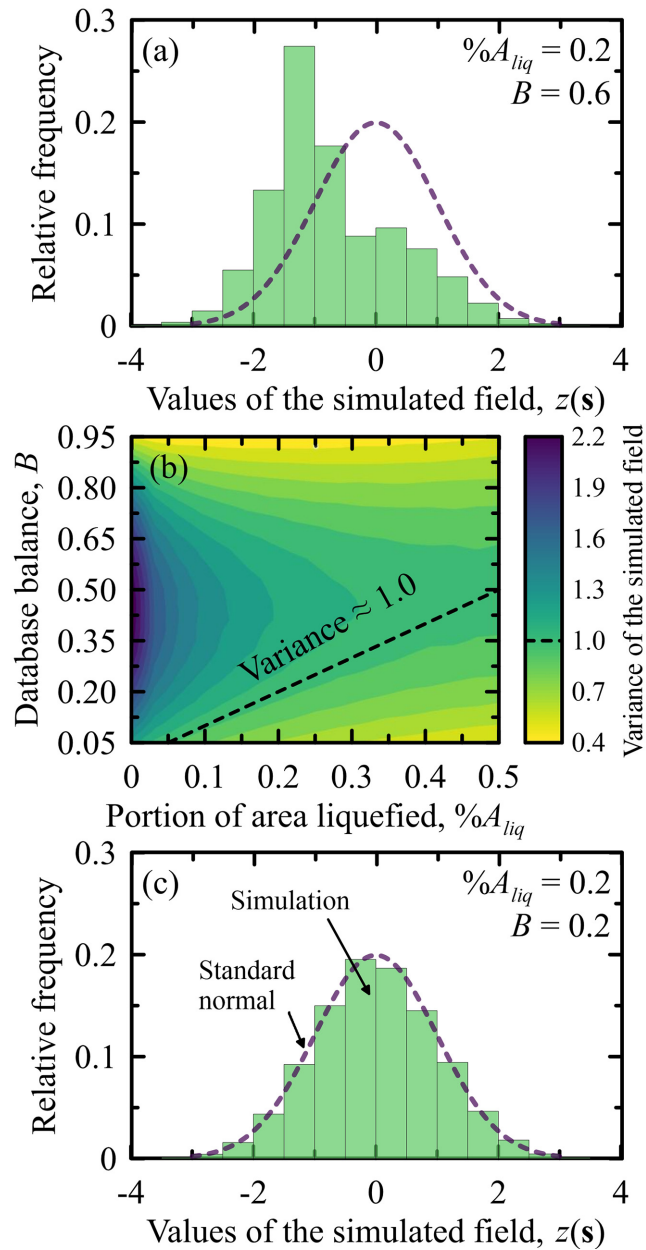
Figure 4. Demonstration of the effects of database balance including (a) a histogram of a simulated field for a poorly balanced database, (b) the observed variance of the simulated field as a function of database balance and the portion of the area liquefied, and (c) a histogram of a simulated field for a properly balanced database.

lengths of 800 to 900 m. The databases of observations of liquefaction manifestations for these two earthquakes are small relative to many of the other earthquakes in the database, per Table 1.
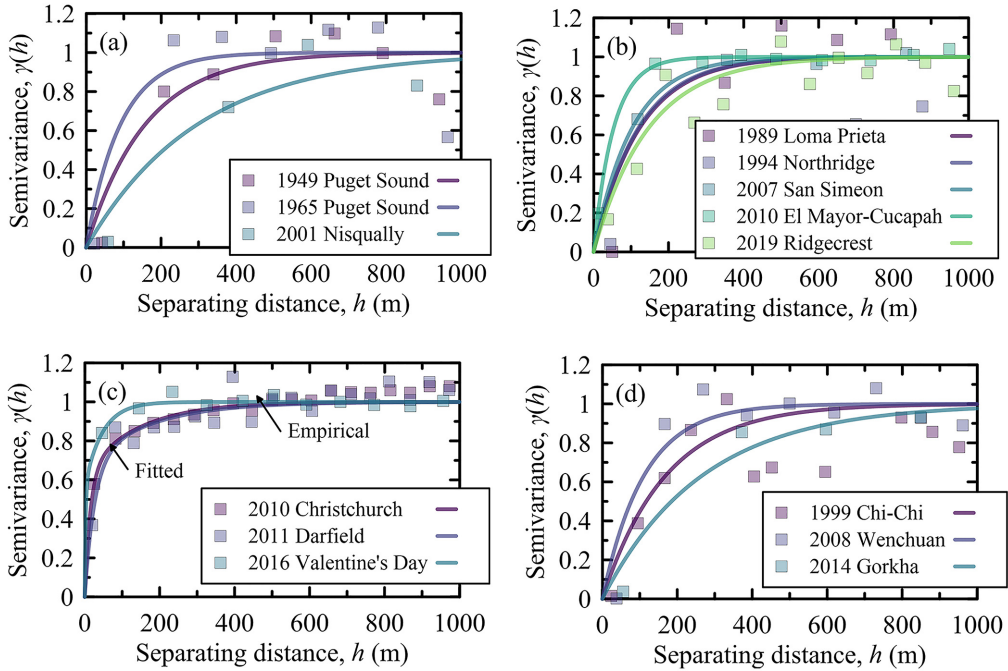
Figure 5. Empirical and fitted semivariograms for earthquakes in (a) Washington, (b) California, (c) Canterbury and (d) other global locations.

Table 2. Parameters for the fitted semivariograms for each earthquake and region

| Earthquake	Exponential			Gaussian		
	c_1	ℓ_1 (m)	ℓ_2 (m)	c_1	ℓ_1 (m)	ℓ_2 (m)
1949 Olympia	1.00	473	–	1.00	296	–
1965 Puget Sound	1.00	275	–	1.00	171	–
1989 Loma Prieta	1.00	365	–	1.00	212	–
1994 Northridge	1.00	355	–	1.00	262	–
1999 Chi-Chi	1.00	506	–	1.00	278	–
2001 Nisqually	1.00	892	–	1.00	564	–
2003 San Simeon	1.00	306	–	1.00	190	–
2008 Wenchuan	1.00	333	–	1.00	203	–
2010 El-Mayor Cucapah	1.00	159	–	1.00	44	–
2010 Darfield	0.64	46	431	0.80	69	788
2011 Christchurch	0.68	77	508	0.80	75	872
2015 Gorkha	1.00	799	–	1.00	481	–
2016 Valentine's Day	0.57	8	140	0.80	32	310
2019 Ridgecrest	1.00	436	–	1.00	271	–
Washington	1.00	386	–	1.00	206	–
California	1.00	468	–	1.00	301	–
Canterbury	0.82	66	435	0.89	67	755
Other regions	1.00	445	–	1.00	268	–
Global	0.87	123	501	0.88	113	778
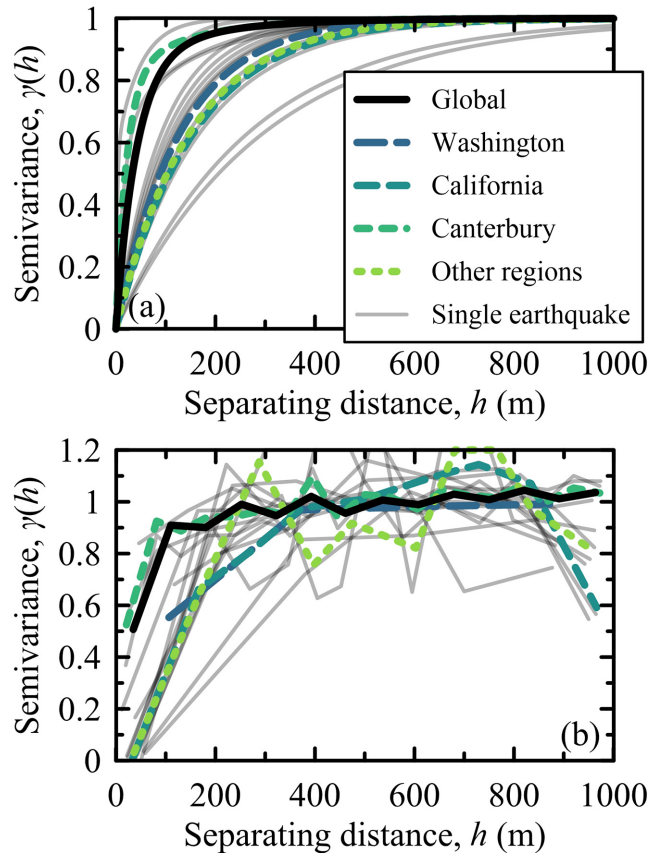
Figure 6. (a) Fitted and (b) empirical semivariograms for calculated for each earthquake and for the earthquakes in each region pooled together.

Correlation structures with multiple scales (i.e. significant values of both ℓ_1 and ℓ_2) are only identified for the three Cantabrian earthquakes. This structure is more apparent for the earthquakes with very many (i.e. > 1000) recorded liquefaction points (the 2010 Darfield and 2011 Christchurch earthquakes). It is unclear whether this structure emerges for these two cases because (a) liquefaction was particularly widespread in those earthquakes, (b) some feature of the geology in the area around Christchurch differentiates it from other regions globally, or (c) the data for those two earthquakes are the most robust. Additional empirical observations are needed to determine which possible explanation of this result is correct.

The models presented above are developed without consideration of surficial geology. To examine the error introduced by this assumption, we repeat the analysis above for the earthquakes in California and restrict the non-liquefaction points to be sampled in geologic units containing quaternary alluvium or artificial fill over intertidal mud according to the Wills et al. (2015) geologic map (for all other geologic units, liquefaction is considered to be absent). Table 3 reports the correlation lengths obtained in the conditioned analysis using the Gaussian functional form (Equation 3), which we denote as ℓ_1^*. The correlation lengths conditioned on susceptible geology differ by 10% or less compared to the original

Table 3. Parameters for the fitted semivariograms for each earthquake located in California: estimates of ℓ_1^* are conditioned on sampling points only in susceptible surficial geology

Earthquake	ℓ_1 (m)	ℓ_1^* (m)
1989 Loma Prieta	212	208
1994 Northridge	262	261
2003 San Simeon	190	190
2010 El-Mayor Cucapah	44	45
2019 Ridgecrest	271	244
California	301	310

values. The distance restrictions placed on sampling the non-liquefaction points (i.e. that they are between 20 m and 1000 m away from the observed liquefaction points) may affect this result. Further investigation is needed to determine how the correlation structure of $z(s)$ is affected by conditioning on geology, including separating the liquefaction manifestation length-scales for Quaternary alluvium and artificial fill, which in principle could be derived from densely sampled observational data. In its current form, the model should be considered agnostic to the underlying geologic units.

Finally, the models presented above can also be used to develop probabilistic estimates of the size of features or zones of features (features are hereby identified as areas connecting adjacent grid points that are assigned liquefaction). Figure 7 shows exceedance probability curves for the size of liquefied zones for two values of $\%A_{liq}$. The exceedance probability curves in Figure 7 assume a large area with a constant $\%A_{liq}$ for clarity. To obtain the exceedance probability curves in Figure 7, we perform Monte Carlo simulations of $\hat{z}(s)$ for a 2D grid of s at a regular spacing and with a constant $\%A_{liq}$; $\hat{z}(s)$ denotes realizations of Z that are generated using the smoothed fitted semivariogram $\hat{\gamma}(h)$.

Model implementation and demonstration

In this section, we provide a concise summary of how the models presented in this paper can be applied, and validate the outcomes of the method by comparing results of the simulation to observations from the 1989 Loma Prieta and 2001 Nisqually earthquakes.

Model implementation

Implementing the latent Gaussian process model to produce spatially correlated maps of the liquefied area in a past or future earthquake proceeds as follows:

1. Select the lateral extents of the region of interest.
2. Divide the region of interest into a grid with spacing at least one tenth of the correlation length obtained from Table 2.
3. Define s by locating the centroid of each differential element within the grid (for example square, rectangular, hexagon).
4. Estimate $\%A_{liq_i}$ for each s_i using a geospatial proxy (e.g. Zhu et al., 2017) or inference based on geology (e.g. Youd and Perkins, 1978).
5. Assemble a matrix, $\boldsymbol{h}$, where h_{ij} is the distance separating locations s_i and s_j in m.
6. Construct the covariance matrix for a standard normal random field, $\hat{z}(s)$ using Equations 2 or 3 and the appropriate coefficients from Table 2 for the region of interest.
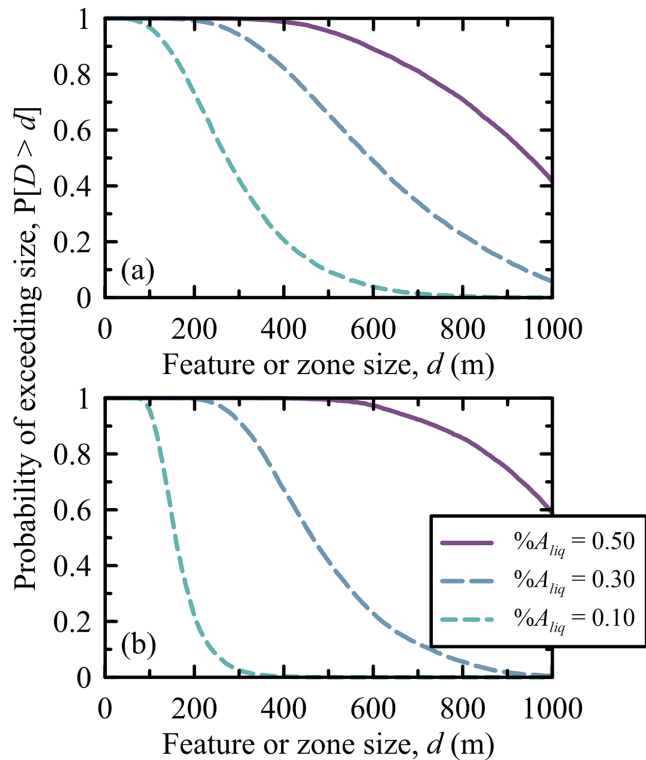
Figure 7. Exceedance probability curves for two dimensional feature or zone size for two values of $\%A_{liq}$ for (a) California and (b) Canterbury.

7. Generate the desired number of realizations of $\hat{z}(s)$.
8. For each realization, assign liquefaction to those grid polygons where $\Phi(\hat{z}(s_i)) \leqslant \%A_{liq,i}$ (Equation 6).

Case studies

This section applies the latent Gaussian process model to develop maps of the liquefied area in Seattle, WA, in the 2001 Nisqually earthquake and San Francisco and Monterey Bay, CA, in the 1989 Loma Prieta earthquake. For all cases, we use a hexagonal grid at 100 m spacing (roughly 1/4 of ℓ_1 for these regions). Figure 8 shows maps of *PGV* for these case studies as obtained from ShakeMaps and Figure 9 shows maps of $\%A_{liq}$ for these case studies as derived from the Zhu et al. (2017) model.

Figures 10 and 11 show maps of $\hat{z}(s)$ for these case studies including and excluding spatial correlation, respectively. Per Equation 6, low (negative) and high (positive) values of $\hat{z}(s)$ are relatively likely and unlikely, respectively, to be liquefaction locations, depending on the values of $\%A_{liq_i}$. The coefficients from Table 2 for Washington are used for Seattle and those for California are used for San Francisco and Monterey Bay. It should be noted that the realizations of $\hat{z}(s)$ with and without spatial correlation (Figures 10 and 11 correspondingly) are operators that translate the percent area liquefied ($\%A_{liq_i}$) maps into
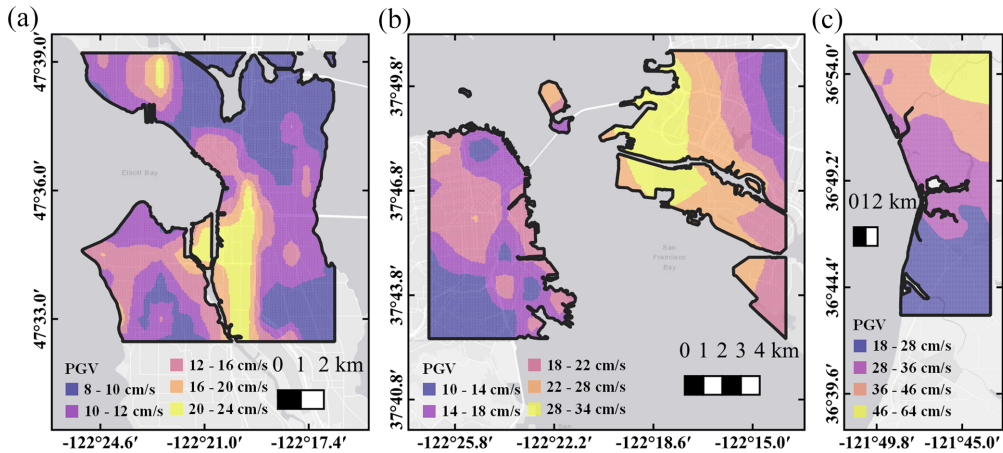
Figure 8. Peak ground velocity maps for (a) Seattle in the 2001 Nisqually earthquake, (b) San Francisco and Oakland in the 1989 Loma Prieta earthquake, and (c) Monterey Bay in the 1989 Loma Prieta earthquake.
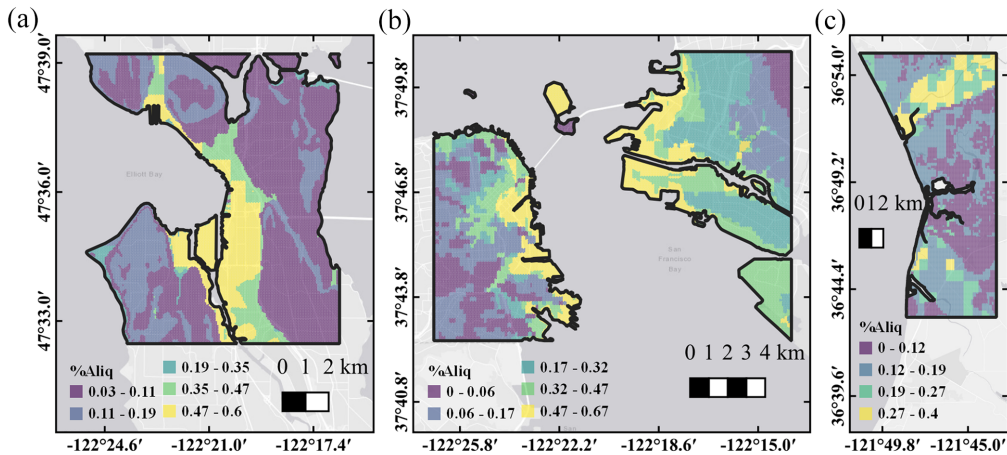

Figure 9. Percent area liquefied ($\%A_{liq_i}$) maps for (a) Seattle in the 2001 Nisqually earthquake, (b) San Francisco and Oakland in the 1989 Loma Prieta earthquake, and (c) Monterey Bay in the 1989 Loma Prieta earthquake. $\%A_{liq_i}$ calculated using Zhu et al. (2017).

spatially correlated maps and uncorrelated maps (Figures 12 and 13 correspondingly) of the liquefied area, and as such have no physical meaning.

Next, Figures 12 and 13 show maps of the liquefied area corresponding to the $\hat{z}(s)$ fields from Figures 10 and 11. Qualitatively, we see that the density of liquefaction observations in both the correlated and uncorrelated maps (Figures 12 and 13) reflects the spatial distribution of the $\%A_{liq}$ predictions. Furthermore, the predictions of $\%A_{liq}$ reflect the different geologic conditions (e.g. Treasure Island is a hydraulic fill, and Yerba Buena island is a rock outcrop). In the uncorrelated maps the liquefaction points do not tend to form continuous polygons corresponding to liquefaction features such as lateral spreads. However,
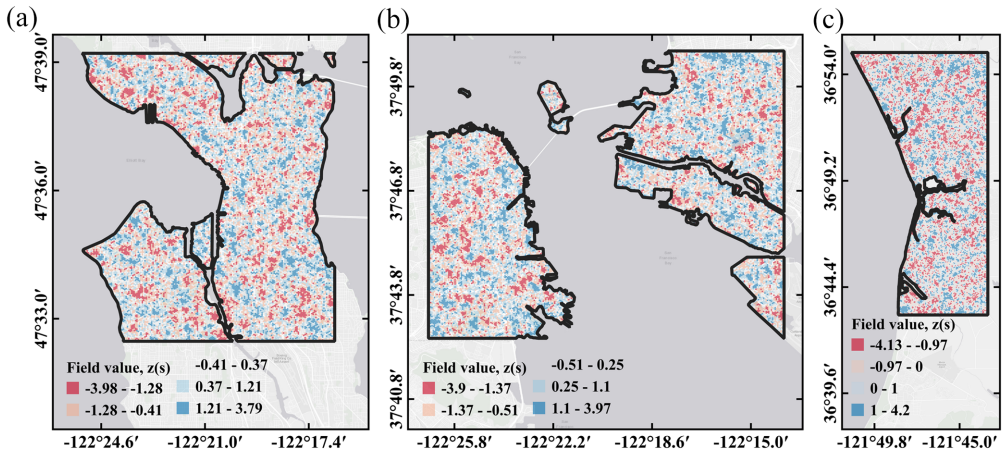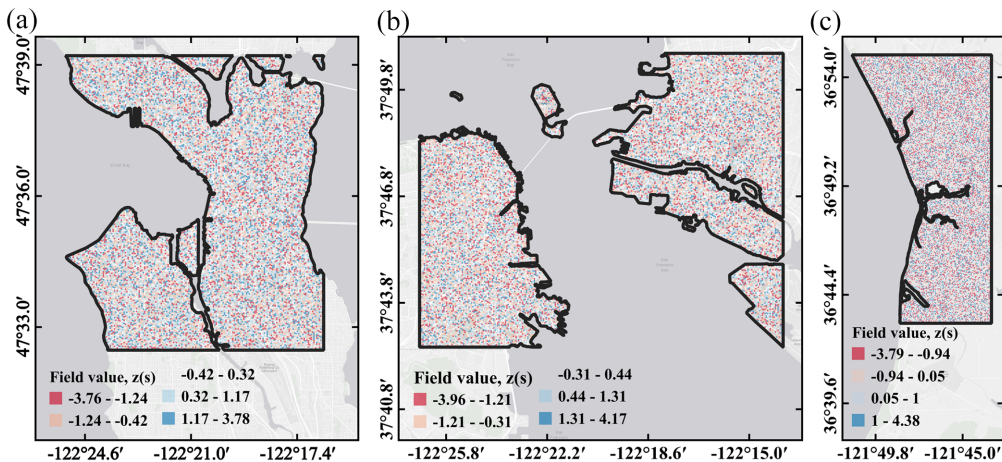
Figure 10. Spatially correlated standard normal random field maps, depicting field value ($\hat{z}(\mathbf{s})$) for (a) Seattle in the 2001 Nisqually earthquake, (b) San Francisco and Oakland in the 1989 Loma Prieta earthquake, and (c) Monterey Bay in the 1989 Loma Prieta earthquake.

Figure 11. Uncorrelated standard normal random field maps, depicting field value ($\hat{z}(\mathbf{s})$) for (a) Seattle in the 2001 Nisqually earthquake, (b) San Francisco and Oakland in the 1989 Loma Prieta earthquake, and (c) Monterey Bay in the 1989 Loma Prieta earthquake.

such features are obtained by application of the spatial correlation model, which produces the correlated map (Figure 12) that has clearer distinctions between areas of liquefaction and nonliquefaction.

We can obtain a suite of maps like those in Figure 12 for each region by generating multiple realizations of $\hat{z}(\mathbf{s})$ in Monte Carlo simulation. Each map can be used to evaluate performance metrics of interest (e.g. disruption of an infrastructure system), the results of which can then be aggregated across the map suite to evaluate probabilities (e.g. probability of failure of an infrastructure system).
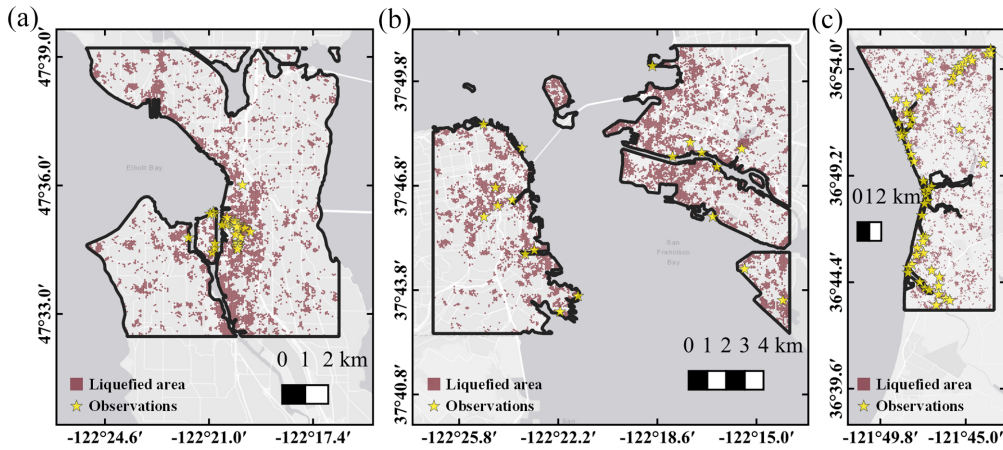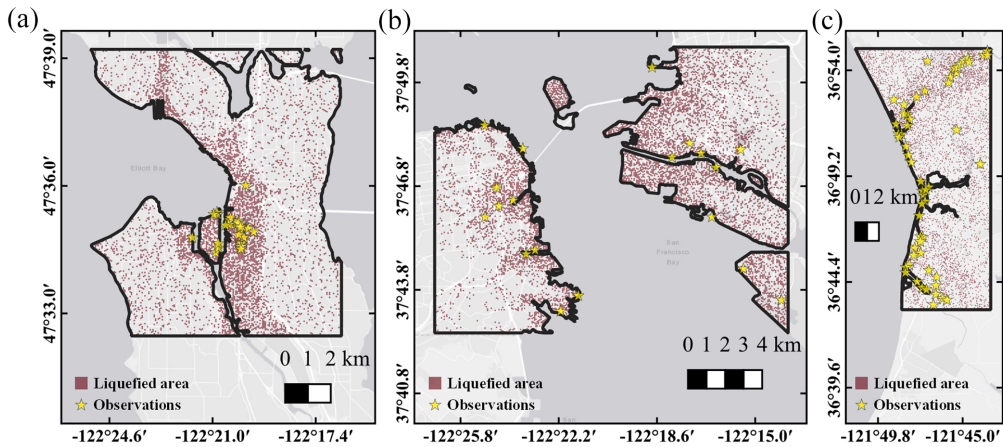
Figure 12. Spatially correlated maps of the liquefied area for (a) Seattle in the 2001 Nisqually earthquake, (b) San Francisco and Oakland in the 1989 Loma Prieta earthquake, and (c) Monterey Bay in the 1989 Loma Prieta earthquake.

Figure 13. Uncorrelated maps of the liquefied area for (a) Seattle in the 2001 Nisqually earthquake, (b) San Francisco and Oakland in the 1989 Loma Prieta earthquake, and (c) Monterey Bay in the 1989 Loma Prieta earthquake.

To verify that the Gaussian process accurately captures the predictions of $\%A_{liq}$ from the geospatial model, Figure 14 shows the simulated $\%A_{liq}$ (i.e. the area of cells marked as "liquefied areas" in Figures 12 and 13 divided by the total area) binned according to the predicted $\%A_{liq}$ (i.e. Figure 9). Results from 10 random field realizations are shown for both approaches. Either approach generates maps of the liquefied area that are consistent with $\%A_{liq}$.

Across a sufficiently large number of realizations, the true positive rate at sites with manifestations is equal to $\%A_{liq,i}$ at those sites, as dictated by Equation 6. The false positive rate at sites *without* manifestations is likewise equal to $\%A_{liq,i}$ at those sites. These
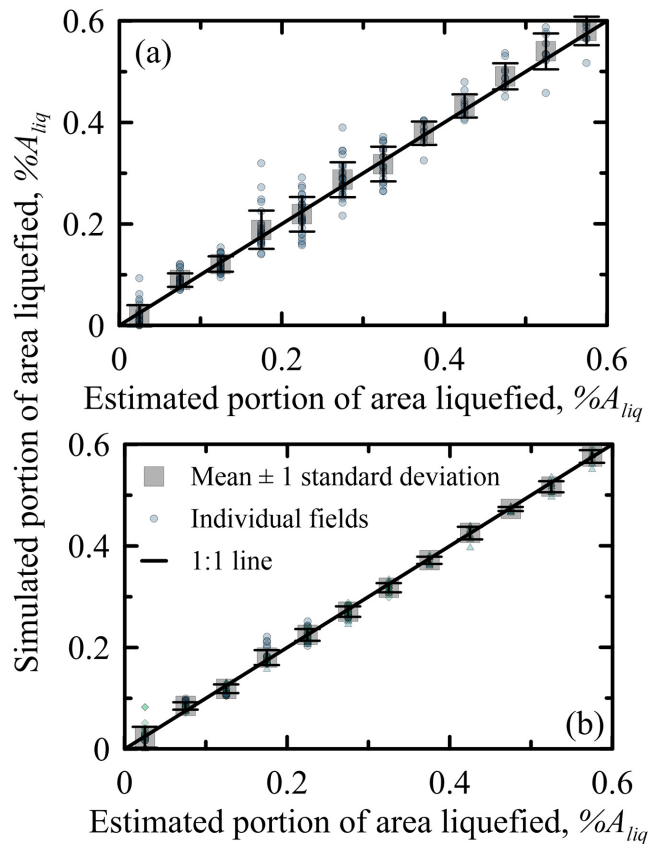
Figure 14. Comparison of the predicted and simulated $\%A_{liq}$ values for 10 random field realizations for each case study area for (a) the case including spatial correlation and (b) the uncorrelated case.

identities are true for the implementations with and without spatial correlation. However, the case including spatial correlation is more likely to correctly identify clusters of sites with manifestations. Figure 15 shows the true positive rate conditioned on a nearby true positive (i.e. how likely it is for the model under evaluation to have correctly predicted that a site has truly liquefied given that a nearby site has also liquefied) for the cases with and without spatial correlation for 1,000 realizations of the Monterey Bay and Puget Sound study areas. These areas are used for this analysis because they include more observations (46 and 37 points, compared to 16 in the San Francisco and Oakland study area). The conditional true positive rate is much higher at short distances when including spatial correlation. In the case without correlation, predictions at each site are independent, and the true positive rate is $\%A_{liq,i}$ regardless of whether a site with a true positive is located nearby.

The extensive reconnaissance survey after the 2019 Ridgecrest (Zimmaro et al., 2020) is believed to have mapped all locations of liquefaction manifestation within certain study areas located within or adjacent to the two dry lakes (Searles Lake and China Lake; their specific study locations shown in their Figures 4, 5 and 8). Accordingly, for this event, a similar comparison can be performed both for the true positive rate conditioned on true positive, as well as, the true negative rate conditioned on true negative (Figure 16). The case with spatial correlation has a higher true positive and true negative at short separation
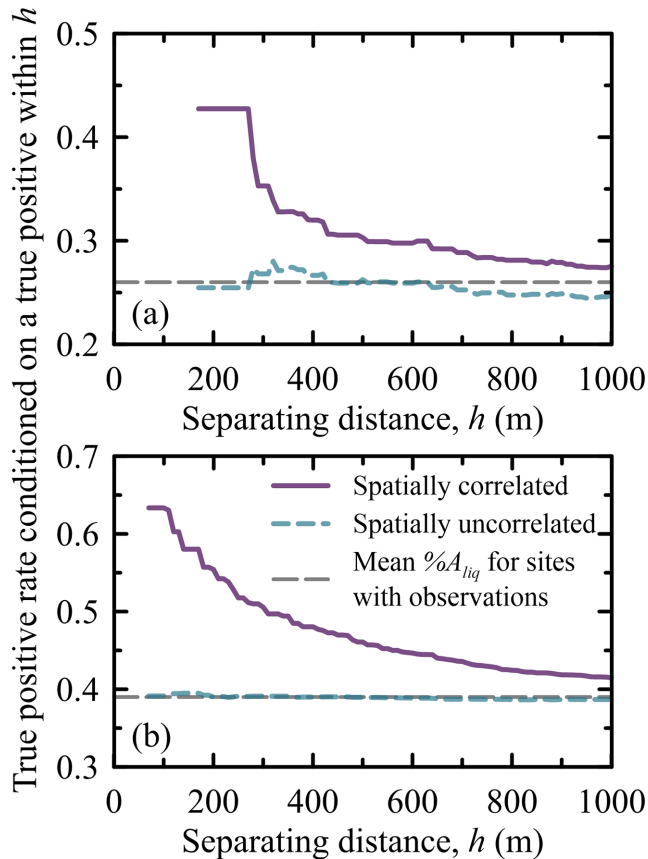
Figure 15. The true positive rate for sites in the (a) Monterey Bay and (b) Puget Sound case study areas conditioned on a true positive within a threshold separating distance.

distances, meaning that it more effectively predicts the liquefaction and no-liquefaction clusters. The case with no spatial correlation treats every site independently and so the conditional true positive and true negative rate are equal to $\%A_{liq}$ and $1 - \%A_{liq}$ regardless of their proximity to nearby location with liquefaction or no-liquefaction manifestation, respectively. Figure 16 indicates a more rapid rate of decay of true positive probabilities with separation distance than is seen in Figure 15 for Monterey Bay and Puget Sound. We believe this is caused by the potentially liquefiable regions being separated by bands of non-liquefiable material, which are associated with the depositional environment of the dry lakebeds where the reconnaissance was performed. Within such setting, the decay rate is highly directional, being slow parallel to the old shoreline and rapid in the perpendicular direction. Since the semi-variogram model does not capture these directional features, a directional averaging occurs that produces, in aggregate, a relatively show decay rate.

Concluding remarks

This study proposes a latent Gaussian process model for generating maps of the liquefied area in future earthquakes. The model uses a geospatial proxy model for liquefaction and a standard normal random field ($\hat{z}(s)$) to assign liquefaction or non-liquefaction to an arbitrary set of locations in a given earthquake or earthquake scenario. As implemented here,
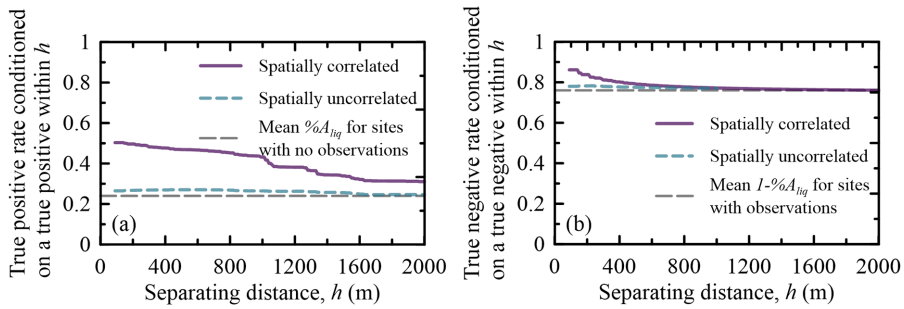
Figure 16. Ridgecrest (a) true positive rate conditioned on true positive and (b) true negative conditioned on a true negative within a threshold distance.

the Zhu et al. (2017) geospatial proxy model is applied, although this feature is modular and other models could be used for particular applications if they are judged to be more suitable for the application region.

The semi-variogram coefficients needed to simulate $\hat{z}(s)$ are estimated using databases of observations of surficial liquefaction manifestations from 14 historical earthquakes. At present, the models are strictly compatible only with the Zhu et al. (2017) geospatial proxy, but the approach described in this study could be applied to produce models for $\hat{z}(s)$ for another proxy. Whether the semi-variogram applied here could be used with another geospatial proxy depends mainly on the accuracy and spatial resolution of that proxy. If $\%A_{liq}$ is mapped using a methodology with a similar level of accuracy and indices that are broadly consistent spatially (such as surface geology) the present semi-variogram model could be used, albeit with additional epistemic uncertainty.

The latent Gaussian process models developed in this study allow users to assign liquefaction and nonliquefaction to a grid of locations in a manner that is consistent with both (a) the expected portion of area liquefied ($\%A_{liq}$) at those locations and (b) the spatial distribution of liquefaction manifestations observed in past earthquakes. The models could be improved by incorporating liquefaction inventories from additional earthquakes, and this work highlights the need for collecting rich inventories of both liquefaction and nonliquefaction points in future reconnaissance efforts.

ORCID iDs

Paolo Zimmaro https://orcid.org/0000-0002-3544-5961
Grigorios Lavrentiadis https://orcid.org/0000-0001-6546-1340
Pengfei Wang https://orcid.org/0000-0002-2844-4557
Domniki Asimaki https://orcid.org/0000-0002-3008-8088
Jonathan P Stewart https://orcid.org/0000-0003-3602-3629

References

Adler RF, Sapiano MR, Huffman GJ, Wang J-J, Gu G, Bolvin D, Chiu L, Schneider U, Becker A, Nelkin E, Xie P, Ferraro R and Shin D-B (2018) The Global Precipitation Climatology Project (GPCP) monthly analysis (new version 2.3) and a review of 2017 global precipitation. *Atmosphere* 9(4): 138.

Ahdi SK, Stewart JP, Ancheta TD, Kwak DY and Mitra D (2017) Development of VS profile database and proxy-based models for V_{S30} prediction in the Pacific Northwest region of North America. *Bulletin of the Seismological Society of America* 107(4): 1781–1801.

Allen TI, Wald DJ, Hotovec A, Lin K, Earle P and Marano K (2008) *An atlas of ShakeMaps for selected global earthquakes*. Technical report 2008-1236, August. Reston, VA: United States Geological Survey (USGS).

Allstadt KE, Thompson EM, Jibson RW, Wald DJ, Hearne M, Hunter EJ, Fee J, Schovanec H, Slosky D and Haynie KL (2021) The US Geological Survey ground failure product: Near-real-time estimates of earthquake-triggered landslides and liquefaction. *Earthquake Spectra* 38: 5–36.

Anderson JG and Brune JN (1999) Probabilistic seismic hazard analysis without the ergodic assumption. *Seismological Research Letters* 70(1): 19–28.

Baise LG, Higgins RB and Brankman CM (2006) Liquefaction hazard mapping—Statistical and spatial characterization of susceptible units. *Journal of Geotechnical and Geoenvironmental Engineering* 132(6): 705–715.

Befus KM, Barnard PL, Hoover DJ, Hart JF and Voss CI (2020) Increasing threat of coastal groundwater hazards from sea-level rise in California. *Nature Climate Change* 10(10): 946–952.

Beyzaei CZ, Bray JD, Cubrinovski M, Bastin S, Stringer M, Jacka M, Van Ballegooy S, Riemer M and Wentz R (2020) Characterization of silty soil thin layering and groundwater conditions for liquefaction assessment. *Canadian Geotechnical Journal* 57(2): 263–276.

Boore DM, Stewart JP, Seyhan E and Atkinson GM (2014) NGA-West2 equations for predicting PGA, PGV, and 5% damped PSA for shallow crustal earthquakes. *Earthquake Spectra* 30(3): 1057–1085.

Boulanger RW and Idriss I (2016) CPT-based liquefaction triggering procedure. *Journal of Geotechnical and Geoenvironmental Engineering* 142(2): 04015065.

Brandenberg SJ, Wang P, Nweke CC, Hudson K, Mazzoni S, Bozorgnia Y, Hudnut KW, Davis CA, Ahdi SK, Zareian F, Fayaz J, Koehler RD, Chupik C, Pierce I, Williams A, Akciz S, Hudson MB, Kishida T, Brooks BA, Gold RY, Ponti DJ, Scharer K, McPhillips D, Duross C, Ericksen T, Hernandez J, Patton J, Olson B, Dawson TE, Treiman J, Blake K, Buchhuber J, Madugo CLM, Sun J, Donnellan A, Lyzenga G and Conway E (2019) Preliminary report on engineering and geological effects of the July 2019 Ridgecrest earthquake sequence. Technical report, Geotechnical Extreme Event Reconnaissance (GEER) Association. DOI: 10.18118/G6H66K

Bray JD, Cubrinovski M, Zupan J and Taylor M (2014) Liquefaction effects on buildings in the central business district of Christchurch. *Earthquake Spectra* 30(1): 85–109.

Bray JD, Sancio RB, Kammerer AM, Merry S, Rodriquez-Marek A, Khazai B, Chang S, Bastani A, Collins B, Hausler E, Dreger D, Perkins W and Nykamp M (2001) Some observations of geotechnical aspects of the February 28, 2001, Nisqually earthquake in Olympia, South Seattle, and Tacoma, Washington. Technical report. Available at: https://apps.peer.berkeley.edu/publications/nisqually/geotech/index.html. (accessed January 20, 2021)

Campbell KW and Bozorgnia Y (2014) NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. *Earthquake Spectra* 30(3): 1087–1115.

Cao Z, Hou L, Xu H and Yuan X (2010) Distribution and characteristics of gravelly soil liquefaction in the Wenchuan M_s 8.0 earthquake. *Earthquake Engineering and Engineering Vibration* 9(2): 167–175.

Casella G and Berger RL (2021) *Statistical Inference*. Boston, MA: Cengage Learning.

Cetin KO, Seed RB, Kayen RE, Moss RE, Bilge HT, Ilgac M and Chowdhury K (2018) SPT-based probabilistic and deterministic assessment of seismic soil liquefaction triggering hazard. *Soil Dynamics and Earthquake Engineering* 115: 698–709.

Chiles J-P and Delfiner P (2009) *Geostatistics: Modeling Spatial Uncertainty*, vol. 497. Hoboken, NJ: John Wiley & Sons.

Chleborad AF and Schuster RL (1998) *Ground Failure Associated with the Puget Sound Region Earthquakes of April 13, 1949, and April 29, 1965*. Reston, VA: United States Geological Survey (USGS).

Chu DB, Stewart JP, Lee S, Tsai J-S, Lin P, Chu B, Seed RB, Hsu S, Yu M and Wang MC (2004) Documentation of soil conditions at liquefaction and non-liquefaction sites from 1999 Chi–Chi (Taiwan) earthquake. *Soil Dynamics and Earthquake Engineering* 24(9–10): 647–657.

Cubrinovski M, Rhodes A, Ntritsos N and Van Ballegooy S (2019) System response of liquefiable deposits. *Soil Dynamics and Earthquake Engineering* 124: 212–229.

Daly C, Taylor G and Gibson W (1997) The PRISM approach to mapping precipitation and temperature. In: *Proceedings of the 10th AMS conference on applied climatology*, Reno, NV, 20–23 October.

Dudley R (1967) The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *Journal of Functional Analysis* 1(3): 290–330.

Fan Y, Li H and Miguez-Macho G (2013) Global patterns of groundwater table depth. *Science* 339(6122): 940–943.

Foster KM, Bradley BA, McGann CR and Wotherspoon LM (2019) A V_{s30} map for New Zealand based on geologic and terrain proxy variables and field measurements. *Earthquake Spectra* 35(4): 1865–1897.

Geyin M, Maurer BW, Bradley BA, Green RA and van Ballegooy S. CPT-based liquefaction case histories compiled from three earthquakes in Canterbury, New Zealand. *Earthquake Spectra*. 2021; 37(4): 2920–2945. DOI: 10.1177/8755293021996367

Greenfield MW and Grant A (2020) Probabilistic regional-scale liquefaction triggering modeling using 3D Gaussian processes. *Soil Dynamics and Earthquake Engineering* 134: 106159.

Hamada M, Isoyama R and Wakamatsu K (1996) Liquefaction-induced ground displacement and its related damage to lifeline facilities. *Soils and Foundations* 36: 81–97.

Holzer TL, Bennett MJ, Noce TE, Padovani AC and Tinsley JC III (2006) Liquefaction hazard mapping with LPI in the Greater Oakland, California, area. *Earthquake Spectra* 22(3): 693–708.

Holzer TL, Noce TE, Bennett MJ, Tinsley JC III and Rosenberg LI (2005) Liquefaction at Oceano, California, during the 2003 San Simeon earthquake. *Bulletin of the Seismological Society of America* 95(6): 2396–2411.

Isaaks EH and Srivastava MR (1989) *Applied Geostatistics*. Oxford: Oxford University Press.

Iwasaki T, Tatsuoka F, Tokida K-I and Yasuda S (1978) A practical method for assessing soil liquefaction potential based on case studies at various sites in Japan. In: *Proceedings of the 2nd international conference on microzonation for safer construction, research and application*, *ISSMGE*: San Francisco, CA, 26 November–1 December, vol. 2, pp. 885–896.

Journel AG and Huijbregts CJ (1978) *Mining Geostatistics*, vol. 600. Cambridge, MA: Academic Press.

Kubota T, Shige S, Hashizume H, Aonashi K, Takahashi N, Seto S, Hirose M, Takayabu YN, Ushio T, Nakagawa K, Iwanami K, Kachi M and Okamoto K (2007) Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation. *IEEE Transactions on Geoscience and Remote Sensing* 45(7): 2259–2275.

Kwok OLA, Stewart JP, Kwak DY and Sun P-L (2018) Taiwan-specific model for Vs30 prediction considering between-proxy correlations. *Earthquake Spectra* 34(4): 1973–1993.

Lavrentiadis G, Abrahamson NA, Nicolas KM, Bozorgnia Y, Goulet CA, Babič A, Macedo J, Dolšek M, Gregor N, Kottke AR, Lacour M, Liu C, Meng X, Phung V-B, Sung C-H and Walling M (2022) Overview and introduction to development of non-ergodic earthquake ground-motion models. *Bulletin of Earthquake Engineering*. Epub ahead of print 17 August. DOI: 10.1007/s10518-022-01485-x.

Lewis M, Arango I, Kimball J and Ross T (1999) Liquefaction resistance of old sand deposits. *In: Proceedings of the 11th Pan American conference on soil mechanics and geotechnical engineering*, Foz do Iguaçu, Brazil, 8–12 August, pp. 821–829. San Paulo, Brazil: ABMS.

Lilliefors HW (1967) On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association* 62(318): 399–402.

McCrink TP, Pridmore CL, Tinsley JC III, Sickler RR, Brandenberg SJ and Stewart JP (2011) Liquefaction and other ground failures in Imperial County, California, from the April 4, 2010, El Mayor-Cucapah earthquake. Technical report, US Geological Survey (USGS), Reston, VA.

Markhvida M, Ceferino L and Baker JW (2018) Modeling spatially correlated spectral accelerations at multiple periods using principal component analysis and geostatistics. *Earthquake Engineering & Structural Dynamics* 47(5): 1107–1123.

Moss RE, Seed RB, Kayen RE, Stewart JP, Der Kiureghian A and Cetin KO (2006) CPT-based probabilistic and deterministic assessment of in situ seismic soil liquefaction potential. *Journal of Geotechnical and Geoenvironmental Engineering* 132(8): 1032–1051.

Moss RE, Thompson EM, Scott Kieffer D, Tiwari B, Hashash YM, Acharya I, Adhikari BR, Asimaki D, Clahan KB, Collins BD, Dahal S, Jibson RW, Khadka D, Macdonald A, Madugo CLM, Mason HB, Pehlivan M, Rayamajhi D and Uprety S (2015) Geotechnical effects of the 2015 magnitude 7.8 Gorkha, Nepal, earthquake and aftershocks. *Seismological Research Letters* 86(6): 1514–1523.

Obermeier SF, Jacobson RB, Smoot JP, Weems RE, Gohn GS, Monroe JE and Powars DS (1990) Earthquake-induced liquefaction features in the coastal setting of South Carolina and in the fluvial setting of the New Madrid seismic zone. Technical report, United States Geological Survey (USGS), Reston, VA.

Petersen MD, Shumway AM, Powers PM, Mueller CS, Moschetti MP, Frankel AD, Rezaeian S, McNamara DE, Luco N, Boyd OS, Rukstales Ks, Jaiswal KS, Thompson EM, Hoover SM, Clayton BS, Field EH and Zeng Y (2020) The 2018 update of the US National Seismic Hazard Model: Overview of model and implications. *Earthquake Spectra* 36(1): 5–41.

Quigley MC, Bastin S and Bradley BA (2013) Recurrent liquefaction in Christchurch, New Zealand, during the Canterbury earthquake sequence. *Geology* 41(4): 419–422.

Rashidian V and Baise LG (2020) Regional efficacy of a global geospatial liquefaction model. *Engineering Geology* 272: 105644.

Rogers AM, Walsh TJ, Kockleman WJ and Priest GR (1998) *Assessing earthquake hazards and reducing risk in the Pacific Northwest*. Professional paper 1560, pp. 373–440. Reston, VA: United States Geological Survey (USGS).

Schmitt R, Tanyas H, Nowicki Jessee MA, Zhu J, Biegel KM, Allstadt KE, Jibson RW, Thompson EM, Van Westen CJ, Sato HP, Wald DJ, Godt JW, Gorum T, Xu C, Rathje EM and Knudsen KL (2017) *An Open Repository of Earthquake-Triggered Ground-Failure Inventories* (version 3.0, September 2020). Reston, VA: United States Geological Survey (USGS).

Skrondal A and Rabe-Hesketh S (2007) Latent variable modelling: A survey. *Scandinavian Journal of Statistics* 34(4): 712–745.

Stewart JP, Seed RB and Bray JD (1996) Incidents of ground failure from the 1994 Northridge earthquake. *Bulletin of the Seismological Society of America* 86(1B): S300–S318.

Thompson EM, Wald DJ and Worden C (2014) A V_{s30} map for California with geologic and topographic constraints. *Bulletin of the Seismological Society of America* 104(5): 2313–2321.

Tinsley JC III, Egan JA, Kayen RE and Bennett MJ (1998) *Liquefaction and associated effects*. Professional paper 1551-287. Reston, VA: United States Geological Survey (USGS).

Van Ballegooy S, Malan P, Lacrosse V, Jacka M, Cubrinovski M, Bray J, O'Rourke T, Crawford S and Cowan H (2014) Assessment of liquefaction-induced land damage for residential Christchurch. *Earthquake Spectra* 30(1): 31–55.

Wald DJ and Allen TI (2007) Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America* 97(5): 1379–1395.

Wang C, Chen Q, Shen M and Juang CH (2017) On the spatial variability of CPT-based geotechnical parameters for regional liquefaction evaluation. *Soil Dynamics and Earthquake Engineering* 95: 153–166.

Westerhoff R, White P and Miguez-Macho G (2018) Application of an improved global-scale groundwater model for water table estimation across New Zealand. *Hydrology and Earth System Sciences* 22(12): 6449–6472.

Wills C, Gutierrez C, Perez F and Branum D (2015) A next generation V_{S30} map for California based on geology and topography. *Bulletin of the Seismological Society of America* 105(6): 3083–3091.

Youd TL and Hoose SN (1978) *Historic Ground Failures in Northern California Triggered by Earthquakes*, vol. 993. Washington, DC: US Government Printing Office (GPO).

Youd TL and Perkins DM (1978) Mapping liquefaction-induced ground failure potential. *Journal of the Geotechnical Engineering Division* 104(4): 433–446.

Zhu J, Baise LG and Thompson EM (2017) An updated geospatial liquefaction model for global application. *Bulletin of the Seismological Society of America* 107(3): 1365–1385.

Zhu J, Daley D, Baise LG, Thompson EM, Wald DJ and Knudsen KL (2015) A geospatial liquefaction model for rapid response and loss estimation. *Earthquake Spectra* 31(3): 1813–1837.

Zimmaro P, Nweke CC, Hernandez JL, Hudson KS, Hudson MB, Ahdi SK, Boggs ML, Davis CA, Goulet CA, Brandenberg SJ, Hudnut KW and Stewart JP (2020) Liquefaction and related ground failure from July 2019 Ridgecrest earthquake sequence. *Bulletin of the Seismological Society of America* 110(4): 1549–1566.

Zimmaro P, Stewart JP, Brandenberg SJ, Kwak DY and Jongejan R (2019) Multi-hazard system reliability of flood control levees. *Soil Dynamics and Earthquake Engineering* 124: 345–353.