Engineering Geology 220 (2017) 110–122

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo

Spatially correlated multiscale V_{s30} mapping and a case study of the Suzhou site

Wenxin Liu[a,b], Qiushi Chen[a,*], Chaofeng Wang[a], C. Hsein Juang[a], Guoxing Chen[b]

[a]Glenn Department of Civil Engineering, Clemson University, Clemson, SC29634, USA
[b]Institute of Geotechnical Engineering, Nanjing Tech University, Nanjing210009, China

ARTICLE INFO

Article history:
Received 11 October 2016
Received in revised form 21 January 2017
Accepted 22 January 2017
Available online 31 January 2017

Keywords:
Shear-wave velocity
V_{s30}
Spatial variability
Random field model
Monte Carlo simulation

ABSTRACT

The average shear-wave velocity in the first 30 m of subsoil, V_{s30}, is a key indicator of site response affecting the ground-motion amplification for many earthquake engineering applications. Mapping of V_{s30} over a large region is commonly done through proxy-based models correlating V_{s30} with geological or topographic information. In this paper, a multiscale random field-based framework is presented and applied to mapping V_{s30} over extended areas. This framework accounts for spatial variations of V_{s30} values across different length scales and is able to adaptively refine around areas of high interest while maintaining consistent description of spatial dependence. In the case study site, Suzhou City, a total of 309 shear-wave velocity measurements are compiled and used to calculate V_{s30} values, from which the statistical and spatial parameters for the random field model are inferred. USGS topography-based V_{s30} data are also collected and used as secondary information to improve the accuracy of predictions. The random field models are coupled with Monte Carlo simulations to obtain a multiscale V_{s30} map and its associated uncertainties at the Suzhou site. The new V_{s30} map is then applied to site classification and amplification factor characterization in the studied region to demonstrate its applications.

1. Introduction

The time average shear-wave velocity in the first 30 m of subsoil, denoted as V_{s30}, is an important site parameter used in estimating site response, classifying sites in recent building codes and loss estimation (Boore, 2004). Because of its importance and effectiveness as a site parameter for site response prediction, the NGA-West2 project (Ancheta et al., 2014; Seyhan et al., 2014) made a project-level decision to compile a site database in terms of V_{s30}. The U.S. Geological Survey (USGS) earthquake hazard program also provides and maintains a global V_{s30} map server. Site databases in terms of V_{s30} give useful site information that allows engineers to choose appropriate site conditions for various design and analysis purposes.

While the V_{s30} can be computed directly given a shear-wave velocity measurement, such geophysical measurements are typically very sparse. Therefore, various descriptors or quantitative metrics of site condition have been proposed for the purpose of estimating V_{s30} in the absence of geophysical measurements. For instance, Wald and Allen (2007) proposed a technique to derive first-order site-condition maps directly from topographic data, where the V_{s30} values are correlated with the topographic slope. Wills and Clahan (2006) and Wills and Gutierrez (2008) grouped shear-wave velocity data by corresponding geologic units to determine the shear-wave velocity characteristics of each geologic unit. Then, the geologic unit designation and shear-wave velocity characteristics are applied to sites without shear-wave velocity data. This revised geologic designation improves the previous geology-based V_{s30} method by Wills et al. (2000) and Wills and Silva (1998). In addition, geology-topography hybrid (Scasserra et al., 2009) and geomorphometry-based proxy relationships (Yong et al., 2012) have been proposed for estimating V_{s30}.

A major limitation of proxy-based methods is that, while initially derived from or constrained by observed V_{s30} values, these methods do not directly incorporate the V_{s30} measurements into the generated site condition map. This, along with the increasing amount of available direct geophysical measurement data, motivates the application of geostatistical methods to V_{s30} and site condition mapping. Examples of recent work along this line include the work of Thompson and his coworkers (Thompson et al., 2014, 2011, 2010), where a new map of V_{s30} for California is developed accounting for geology, topography and most importantly, site-specific V_{s30} measurements. The geostatistical approach of regression kriging (RK) is applied to combine these constraints to predict V_{s30}. This approach allows the resulting V_{s30} map to be locally refined to reflect the rapidly

* Corresponding author.
E-mail address: qiushi@clemson.edu (Q. Chen).

http://dx.doi.org/10.1016/j.enggeo.2017.01.026

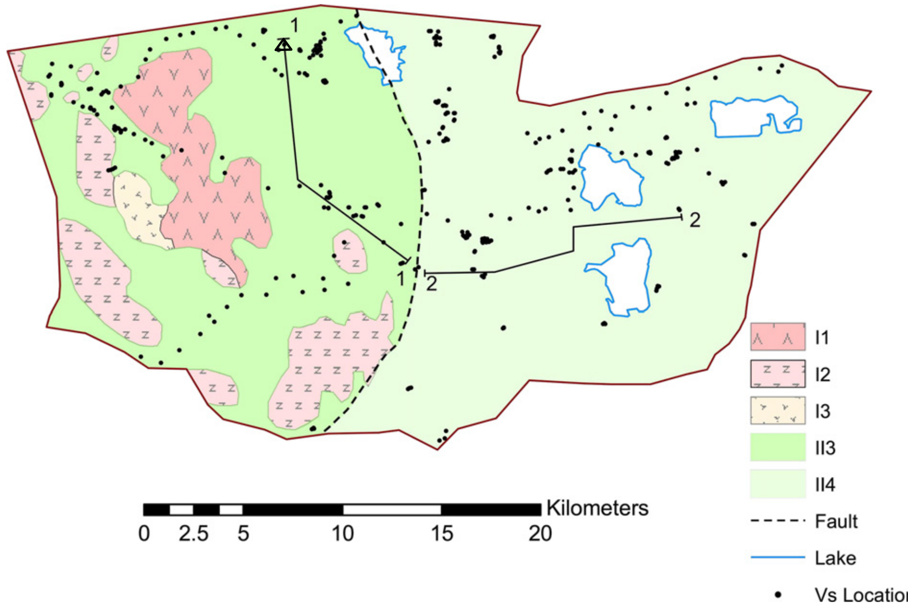
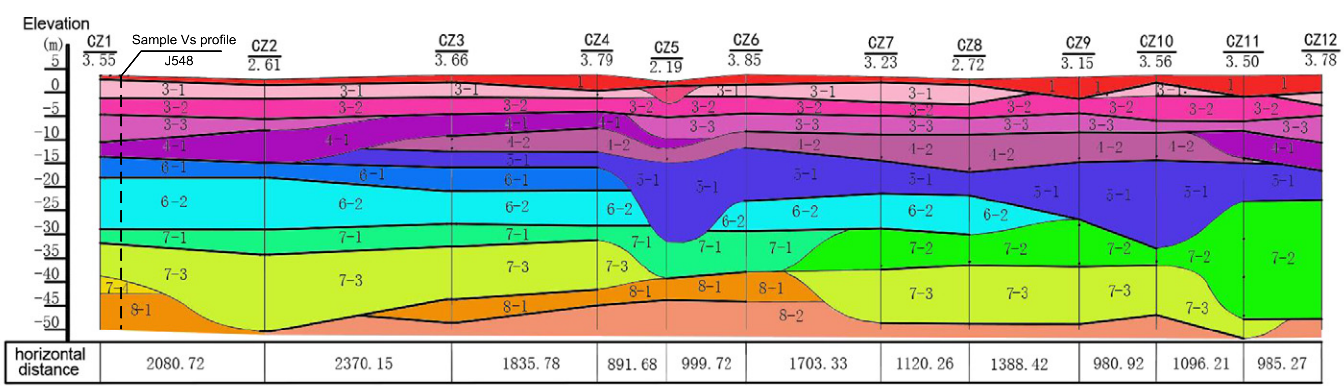
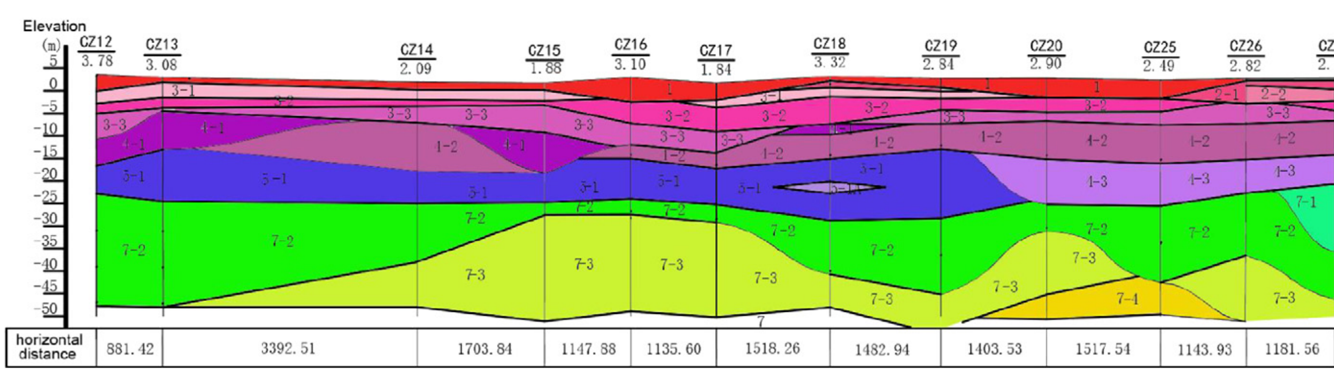


Fig. 1. Surficial geology map of the Suzhou site and locations of shear-wave velocity measurements (black dots in the figure). II_3 is the Taihu alluvial plain; II_4 is the lake-swamp plain; I_1, I_2 and I_3 are outcrops with different rock types. Cross sections 1-1 and 2-2 are used to plot example soil profiles for the top 50 m. The little triangle shows the location of the sample V_s profile in Fig. 3.

expanding database of V_{s30} measurements. Yong et al. (2013) and Wald et al. (2011) applied the kriging-with-a-trend method to mapping V_{s30}, where the baseline model was derived from topographic slope. Also, Lee and Tsai (2008) established the spatial relationship between the shear-wave velocity (V_s) and the N value of the standard penetration test (SPT-N) and adopted the kriging with varying local means to update the V_{s30} maps in Taiwan. Thompson et al. (2007) modeled the horizontal variability of near-surface soil shear-wave velocity in the San Francisco Bay Area using geostatistical methods.



(a) Cross-section 1-1



(b) Cross-section 2-2

Fig. 2. Example soil profiles in the top 50 m for the cross-sections 1-1 and 2-2 shown in Fig. 1.

Table 1
Explanation of soil type numbers used in Fig. 2.

Number	Soil type	Property	Number	Soil type	Property
1	ground fill	loose	5–1A	clay	medium dense
2–1	silt clay	plastic	5–2	fine	dense
2–2	silt clay	soft plastic	6–1	clay	hard plastic
3–1	clay	hard plastic	6–2	silt clay	plastic
3–2	silt clay	plastic	7–1	silt clay	soft plastic
3–3	silt	medium dense	7–2	silt with fine	dense
4–1	silt clay	flow plastic	7–3	silt clay	soft plastic
4–2	sandy clay	medium dense	7–4	silt	dense
4–3	silt	dense	8–1	silt clay	hard plastic
5–1	silt with clay	flow plastic	8–2	silt clay	plastic

In this paper, a multiscale random field-based approach is presented and applied to mapping V_{s30} over an extended region. Unlike existing geostatistical methods for V_{s30} mapping, the presented approach explicitly accounts for the spatial variability of V_{s30} across different length scales and incorporates the compiled database of direct geophysical measurements and proxy-based V_{s30} values. High resolution predictions of V_{s30} can be obtained by adaptively refining coarse-scale values into finer scales in areas where deemed necessary while retaining appropriate spatial correlation, which is a particular useful feature for analyzing fine scale quantities of interest, such as estimation of uncertainties. Coupled with Monte Carlo simulations, the multiscale random field models also allow the quantification of uncertainties in the V_{s30} maps. The resulting V_{s30} maps preserve known V_{s30} data, uphold appropriate spatial correlation and have multiscale resolutions with information on associated uncertainties.

The order of presentation of this paper goes as follows: Section 2 summarizes the engineering geology, field data and secondary V_{s30} data of the Suzhou site; In Section 3, key components of the developed geostatistical tools for mapping V_{s30} are presented; Statistical and spatial characterizations of the known V_{s30} data will be discussed in detail in Section 4; In Section 5, new V_{s30} maps will be represented and applications of those new V_{s30} maps will be discussed in Section 6.

2. The Suzhou site: engineering geology and field data

Suzhou is a populous city on the alluvial plain of the Yangtze River Delta in the southeast of Jiangsu Province, China. In this section, the engineering geology and field data of the Suzhou site are briefly summarized. The dominating alluvial deposits beneath the studied

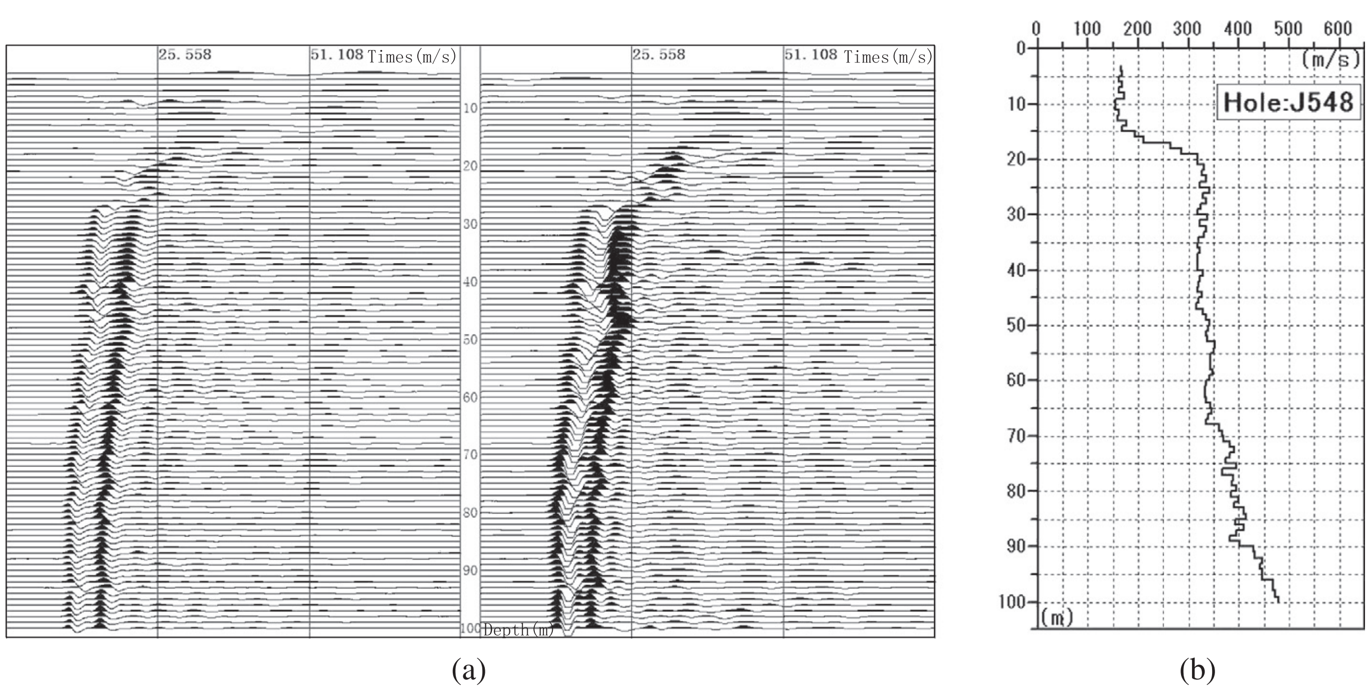


Fig. 3. Sample shear-wave velocity data obtained from the suspension P-S velocity logging method: (a) depth sequential waveform arrivals; (b) shear wave velocity (V_s) versus depth.

Table 2
Summary of soil parameters obtained from borehole samples.

	ρ_{sat} (g/cm^3)	ρ_d (g/cm^3)	LL	PL
Min	1.73	1.14	22.9	11.5
Max	2.96	2.59	70.1	34.4
Mean	2.81	1.51	35.6	20.1

site are soft and sensitive. In addition to geotechnical engineering challenges associated with construction on soft soil, long-period ground motions of far earthquakes may also cause serious damage to engineering projects in this area (Zhan et al., 2009).

2.1. Engineering geology

The studied area of Suzhou City is covered by Quaternary deposits of fluvial, lake, lagoon and marine origins. Most of the area is a combination of a lacustrine plain and delta plain. Some layers of the lake and river deposits are rich in over-consolidated clay. Most of the lagoonal and marine deposits, however, consist of soft clays, which are dark in color and rich in organic matters. Fig. 1 shows the boundaries of the studied area, the major surficial geology units and locations of shear-wave velocity measurements. As shown in Fig. 1, the western portion of the studied area belongs to the Taihu alluvial plain (II_3) with interspersed outcrops (I_1, I_2 and I_3). The eastern portion belongs to the lake-swamp plain (II_4). Almost all of the shear-wave velocity measurements were taken in the geological units II_3 and II_4. Example profiles of the top 50 m soil are plotted in Fig. 2 (a) for the Taihu alluvial plain (II_3) (cross-section 1-1 in Fig. 1) and in Fig. 2 (b) for the lake-swamp plain (II_4) (cross-section 2-2 in Fig. 1), respectively. Explanations of the soil type number are summarized in Table 1.

2.2. Field data

The field data compiled for this study consists of shear-wave velocity measurement data and soil parameters from lab tests on samples collected at boreholes throughout the studied site. The Institute of Earthquake Engineering for Jiangsu Province, China, performed 309 shear-wave velocity tests in the Suzhou site using the suspension P-S velocity logging method. The suspension P-S logging system uses a probe that contains a source and two receivers spaced 1 m apart. The probe is lowered into the borehole to a specified depth, where the source generates a pressure wave in the borehole fluid to be received by the receivers. The elapsed time between arrivals of the waves at the receivers is used to determine the average velocity of a 1-meter-high column of soil around the borehole. An example sequential waveform arrival along depth profile is shown in Fig. 3 (a) and the corresponding shear-wave velocity profile is shown in Fig. 3 (b). The location of this profile is marked in Fig. 1 as an triangle. In general, the shear-wave velocity profile corresponds well with the expected soil conditions. For the top 20 m, the shear wave velocity is relatively small (around 150 m/s), which corresponds to the soft soil layers (types 3-1 to 4-1 in Table 1). When the depth reaches below 20 m, the shear wave velocity increases significantly (to around 350 m/s) and remains constant from 20 to 50 m, which corresponds well to the relatively hard soil layers (type 6-1 to 8-1 in Table 1).

Soil samples were also collected at selected boreholes and analyzed to obtain various soil parameters of interest including the saturated density (ρ_{sat}), the dry density (ρ_d), the liquid limit (LL) and the plastic limit (PL). Table 2 summarizes ranges of soil parameters obtained from borehole samples. The water table is found to be at 1.35 to 1.97 m below ground surface.

2.3. Calculation of V_{s30} at measurement locations

Given the shear-wave velocity measurement data, a time-averaged shear-wave velocity to a profile depth z, denoted as V_{sz}, can be calculated at each measurement location as

$$V_{sz} = \frac{z}{\Delta t_z} \tag{1}$$

$$\Delta t_z = \int_0^z \frac{dz}{V_s(z)} \tag{2}$$

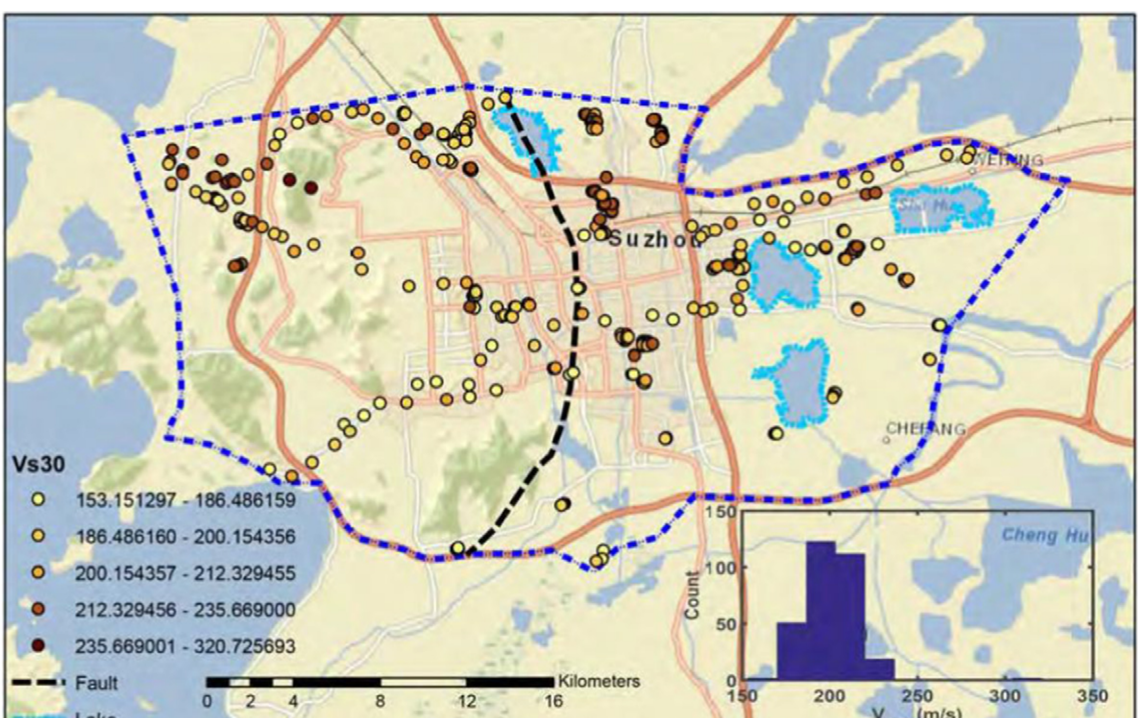


Fig. 4. Map of V_{s30} measurements in Suzhou City, with histogram inset.

where Δt_z is the travel time for shear waves from depth z to the ground surface; $V_s(z)$ is the shear-wave velocity at depth z; the integral is usually evaluated in practice through summation across velocities taken as constant within depth intervals. When the shear-wave velocity profile extends to depths of 30 m or greater, z is taken as 30 m, and the resulting velocity is V_{s30}. When $z < 30$ m, V_{s30} cannot be calculated directly and various correlations between $V_s(z)$ and V_{s30} have been developed to estimate V_{s30} (Boore, 2004; Boore et al., 2011). For this study, all shear wave velocity measurements reach over 30 m.

Fig. 4 plots the V_{s30} values at 309 measurement locations as well as their histogram (the inset).

Those V_{s30} values shown in Fig. 4 are only available at locations with measured shear-wave velocity profiles. To estimate and map V_{s30} values across the region of interest, geostatistical tools and multiscale random field models will be developed and presented in Section 3. Statistical and spatial characterization of the known V_{s30} will be discussed in Section 4.

2.4. Secondary V_{s30} data

In addition to the calculated V_{s30} values at measurement locations, proxy-based V_{s30} values are also collected in this study from the U.S. Geological Survey (USGS) global V_{s30} map server (http://earthquake.usgs.gov/hazards/apps/vs30/). Those V_{s30} values are based on a simplified approach that correlates V_{s30} value with the topographic slope (Wald et al., 2004; Allen and Wald, 2009). Such secondary V_{s30} data are necessary because almost all V_{s30} measurements (307 out of 309) are within the Taihu alluvial plain (II_3) and the lake-swamp plain (II_4), i.e., within relatively soft soils. There is little information on V_{s30} values in hilly areas (I_1, I_2 and I_3). The USGS V_{s30} data will be used to improve V_{s30} predictions in hilly areas, which will be discussed in more detail in Section 5.

Fig. 5 plots the USGS V_{s30} data along with its histogram. It is clear from the map that the hilly areas in the western part of the city have much higher V_{s30} values. Moreover, in the alluvial plain, the mean of the USGS V_{s30} is 219 m/s and the minimum is 180 m/s. The mean of the measured V_{s30} values is 200 m/s and the minimum is 153 m/s. Distributions of the USGS and measurement V_{s30} values have also been compared. In general, it is found that the USGS V_{s30} values tend to predict a higher estimate in the alluvial plain.

3. Geostatistical approach to characterize spatial variability across scales

In this section, key components of the developed geostatistical tools and random field-based models to map V_{s30} are presented. The rational behind a geostatistical approach is the fact that the measured soil parameters at one location are more similar to those at neighboring locations than those further away, i.e., soil parameters are spatially correlated. It is desirable to characterize the spatial structure of soil parameters of interest to improve the accuracy of predictions at unsampled locations.

In this study, a form of covariance called the semivariogram is used to describe the spatial structure, which is equal to one half of the variance of two random variables separated by a distance $\boldsymbol{h}$ as

$$\gamma(\boldsymbol{h}) = \frac{1}{2}\mathrm{Var}[Z(\boldsymbol{u}) - Z(\boldsymbol{u}+\boldsymbol{h})] \quad (3)$$

where $Z(\boldsymbol{u})$ is the variable under consideration at location $\boldsymbol{u}$ and $Z(\boldsymbol{u}+\boldsymbol{h})$ is the lagged version of the variable.

Under the condition of second-order stationarity, the semivariogram is related to the spatial correlation $\rho(\boldsymbol{h})$ by

$$\rho(\boldsymbol{h}) = 1 - \frac{\gamma(\boldsymbol{h})}{\mathrm{COV}(\boldsymbol{0})} \quad (4)$$

where COV($\boldsymbol{0}$) is the covariance at $\boldsymbol{h} = \boldsymbol{0}$. The semivariogram $\gamma(\boldsymbol{h})$ is typically preferred by the geostatistics community because it only requires the increment $Z(\boldsymbol{u}) - Z(\boldsymbol{u}+\boldsymbol{h})$ to be second-order stationary, which is a weaker requirement than the second-order stationarity of the variable itself. In the following examples, the spatial structure of

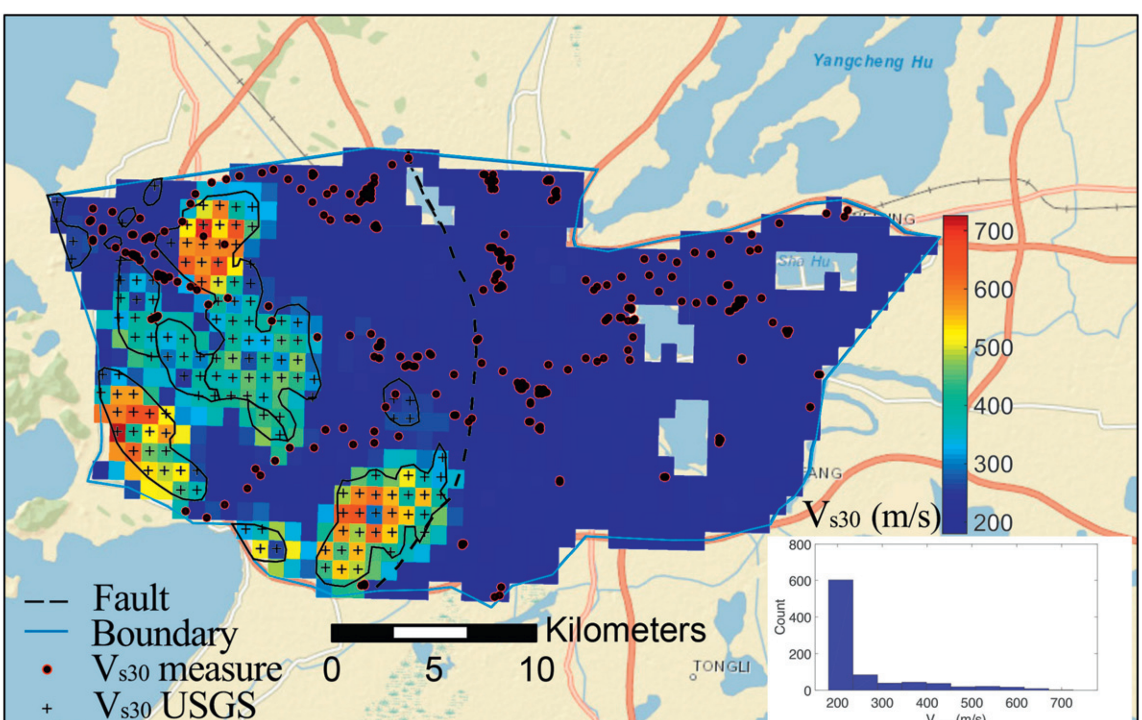


Fig. 5. USGS global slope-based V_{s30} data: map of the V_{s30} values in Suzhou City and the corresponding histogram (inset).

the soil parameter under consideration (i.e., the V_{s30} value) is characterized by the semivariogram model, which can be converted to ρ and implemented within a random field model.

To account for the multiscale nature of soil variability, Chen et al. (2012) and Baker et al. (2011) extended the definition of spatial correlation to multiple scales based on the notion that material properties at the coarser scale are the arithmetically averaged values of the properties over corresponding areas at the finer scale. Such notion is formally similar to the block kriging (Goovaerts, 1997) but with a different intention to consistently and adaptive refine a coarse scale random field. The multiscale random field allows a higher resolution field to be adaptively generated around areas of high interest.

In this work, two scales of interest are considered and all the subsequent developments apply to variables following the standard Gaussian distribution, i.e., variables after the normal score transformation. The variable of interest Z_I^c at the coarse scale is defined as the arithmetically averaged fine scale values over corresponding areas as (Chen et al., 2012)

$$Z_I^c = \frac{1}{N}\sum_{i=1}^{N} Z_{i(I)}^f \tag{5}$$

where the superscripts "c" and "f" correspond to coarse and fine scales, respectively; N is the number of fine scale elements within a corresponding coarse scale area (element) I.

Defining the variable of interest at the fine scale and using the relation of Eq. (5), the expression for the mean, the variance and the spatial correlation of coarse scale variables of interest can be explicitly derived. The mean of a coarse scale element Z_I^c can be derived by taking the expectation of Eq. (5) as

$$\mu_{Z^c} = E[Z_I^c] = \frac{1}{N}\sum_{i=1}^{N} \mu_{Z_{i(I)}^f} = 0 \tag{6}$$

where $\mu_{Z_{i(I)}^f}$ is the mean at the fine scale, which equals to zero for variables following the standard Gaussian distribution. Accordingly, if the variance of the fine scale variable is unity, the coarse scale variance, denoted as $\sigma^2_{Z^c}$, can be computed as

$$\sigma^2_{Z^c} = E\left[\left(Z_I^c\right)^2\right] - 0 = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} \rho_{Z_i^f,Z_j^f}\sigma_{Z_i^f}\sigma_{Z_j^f} \tag{7}$$

where $\rho_{Z_i^f,Z_j^f}$ is the correlation between two fine scale element i and j with variance $\sigma^2_{Z_i^f}$ and $\sigma^2_{Z_j^f}$, respectively.

The covariance between any two elements Z_i and Z_j within the random field is defined as

$$\mathrm{COV}[Z_i, Z_j] = \rho_{Z_i Z_j}\sigma_{Z_i}\sigma_{Z_j} \tag{8}$$

The correlations between all considered scales can be calculated by rearranging the definition of covariance such that

$$\rho_{Z_i,Z_j} = \frac{\mathrm{COV}[Z_i, Z_j]}{\sigma_{Z_i}\sigma_{Z_j}} \tag{9}$$

where Z_i and Z_j are two elements within the random field at any scale with variance $\sigma^2_{Z_i}$ and $\sigma^2_{Z_j}$. By making appropriate substitutions at each scale using Eqs. (8) and (9), the correlation between elements at different scales can be obtained as (Chen et al., 2015, 2016)

$$\rho_{Z_I^c,Z_{II}^c} = \frac{\sum_{i=1}^{N}\sum_{k=1}^{N}\rho_{Z_{i(I)}^f,Z_{k(II)}^f}}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{Z_{i(I)}^f,Z_{j(I)}^f}}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{Z_{i(II)}^f,Z_{j(II)}^f}}} \tag{10}$$

$$\rho_{Z^f,Z_I^c} = \frac{\sum_{i=1}^{N}\rho_{Z^f,Z_{i(I)}^f}}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\rho_{Z_{i(I)}^f,Z_{j(I)}^f}}} \tag{11}$$

where the Roman numerals *I, II*... are used for the coarse scale element number; $\rho_{Z_I^c,Z_{II}^c}$ is the correlation between two coarse-scale elements I and II; ρ_{Z^f,Z_I^c} is the correlation between a fine-scale element and a coarse scale element I; $\rho_{Z_{i(I)}^f,Z_{k(II)}^f}$ is the correlation between a fine element i and a fine element k, which belong to two different coarse scale elements I and II, respectively. Given the correlation ρ between elements at different scales, the corresponding covariances COV can be easily obtained via Eq. (8).

Once the covariance COV between any two elements at any scale in the random field is determined, a conditional sequential simulation approach is taken for the simulation procedure. The process simulates each value individually, conditional upon all known data and previously simulated values. Using such a process, the conditional distribution of the next value to be simulated in the random field, denoted as Z_n, is given by a univariate normal distribution with the updated mean and the variance as

$$(Z_n|\boldsymbol{Z}_p) \sim N\left(\boldsymbol{\Sigma}_{np}\cdot\boldsymbol{\Sigma}_{pp}^{-1}\cdot\boldsymbol{Z}_p, \sigma_n^2 - \boldsymbol{\Sigma}_{np}\cdot\boldsymbol{\Sigma}_{pp}^{-1}\cdot\boldsymbol{\Sigma}_{pn}\right) \tag{12}$$

where $\boldsymbol{Z}_p$ is a vector of all known or previously simulated points; $\boldsymbol{\Sigma}_{np}, \boldsymbol{\Sigma}_{pp}, \boldsymbol{\Sigma}_{pn}$ are covariance matrices; σ_n^2 is the covariance of the next simulated point; the subscription "p'' and "n'' refer to the "previous" simulated point(s) and the "next" point to be simulated, respectively.

Eq. (12) means that the unknown value Z_n at an unmeasured location can be drawn from the conditional normal distribution with the mean $\boldsymbol{\Sigma}_{np}\cdot\boldsymbol{\Sigma}_{pp}^{-1}\cdot\boldsymbol{Z}_p$ and the variance $\sigma_n^2 - \boldsymbol{\Sigma}_{np}\cdot\boldsymbol{\Sigma}_{pp}^{-1}\cdot\boldsymbol{\Sigma}_{pn}$. Once Z_n is generated, it is inserted into the "previous" vector, i.e., $\boldsymbol{Z}_p$, upon which the "next" unknown value at another unsampled location will be generated. Such process is repeated until all locations within a random field are simulated. A key advantage of such conditional simulation is that it preserves the field data in the random field. Moreover, as pointed out by Baker et al. (2011), such a simulation approach is particular suitable for an adaptive refinement process, where additional fine-scale simulations can be progressively added in the random field in locations deemed necessary.

4. Data inference - statistical and spatial characterizations of the known V_{s30} data

The multiscale random field models require as inputs the statistical distributions and the spatial structures of the variable under consideration. In the Suzhou site, a total of 309 V_{s30} values are obtained from direct shear-wave velocity measurements. Fig. 6 plots the histogram of the 309 V_{s30} measurements. Among those 309 V_{s30} measurements, 307 measurements are located in the two dominating surficial geological units: the Taihu alluvial plain (II_3) and the lake-swamp plain (II_4) as shown in Fig. 1. Those V_{s30} measurements are grouped by geological units II_3 and II_4 to see whether significant differences exist. Table 3 summarizes the statistical characteristics (e.g., mean, variance, maximum, upper quantile, median, lower quantile, minimum) of the two groups. As can be seen from Table 3, the statistical characteristics do not differ significantly between the two dominant surficial geological units. In subsequent

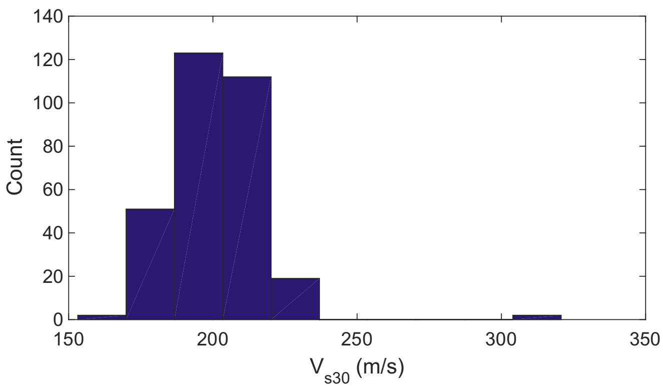


Fig. 6. Histogram of all 309 V_{s30} values calculated from shear-wave velocity measurements.

Table 3
Statistical characteristics of the known V_{s30}.

Statistical parameter	II_3	II_4	Combined II_3 and II_4
Data count	143	164	307
Mean	198	202	200
Variance	205	216	192
Maximum	236	233	236
Upper quantile	208	212	210
Median	196	203	200
Lower quantile	188	193	191
Minimum	172	153	153

characterizations and examples, geologic units II_3 and II_4 are grouped together in random field models. In the outcrop areas (I_1, I_2 and I_3), no direct shear-wave velocity measurement is available. The USGS proxy-based V_{s30} data are collected (refer to Fig. 5) and incorporated as known data in those outcrop areas in subsequent random field simulations.

Fig. 7 plots all measurement data projected in the east–west and north–south directions along with the trend lines. The trend line along the west–east direction is almost a horizontal line, indicating little trend in this direction. On the other hand, Fig. 7 (b) shows slightly increased V_{s30} values from north to south. However, the change is still relatively mild to make any significant impact. It should be pointed out that 307 of the 309 V_{s30} measurements are in the Taihu alluvial plain (II_3) and lake-swamp plain (II_4). So, the trend analysis reveals the trend (or no trend) of V_{s30} in those geological units only.

The empirical or sample semivariogram of V_{s30} measurements are also computed to infer their spatial structure in the studied

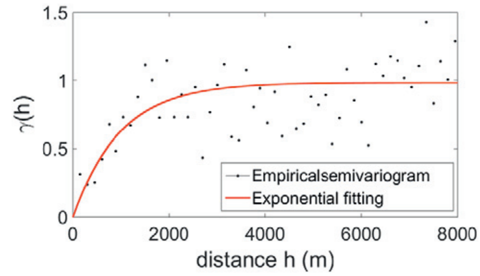


Fig. 8. Empirical and fitted semivariogram based on known V_{s30} at measurement locations.

region. The empirical semivariogram, denoted as $\hat{\gamma}(\boldsymbol{h})$, is calculated as (Goovaerts, 1997)

$$\hat{\gamma}(\boldsymbol{h}) = \frac{1}{2N(\boldsymbol{h})} \sum_{\alpha=1}^{N(\boldsymbol{h})} [Z(\boldsymbol{u}_\alpha) - Z(\boldsymbol{u}_\alpha + \boldsymbol{h})]^2 \tag{13}$$

where $N(\boldsymbol{h})$ is the number of pairs of data ($Z(\boldsymbol{u}_\alpha)$ and $Z(\boldsymbol{u}_\alpha + \boldsymbol{h})$) separated by a vector distance $\boldsymbol{h}$.

To facilitate the incorporation of the semivariogram into random field models, the empirical semivariogram is typically fitted by a basic semivariogram model or a linear combination of several basic semivariogram models that are permissible (Goovaerts, 1997). Fig. 8 plots the empirical semivariogram model as well as the fitted exponential model of the form

$$\gamma(h) = \omega \left[1 - \exp\left(-\frac{3h}{a}\right)\right] + \tau \tag{14}$$

where h is a scalar measure of the separation distance between a pair of points; a is the range, i.e., the distance at which the semivariogram levels off and beyond which the semivariance is constant; $\omega+\tau$ is the sill, which is the constant semivariance beyond the range. The fitted range for this study site is 2973 m and the sill is 0.9833.

5. V_{s30} mapping of the Suzhou site

With the inferred model parameters, the known V_{s30} at measurement locations and the secondary V_{s30} information in the outcrop areas (I_1, I_2, I_3), the multiscale random field models are used to generate V_{s30} maps of the Suzhou site. An initial coarse grid with an element size of 500 × 500 m is used. Lakes are excluded from the V_{s30} maps. The new maps account for and preserve the site-specific shear-wave velocity measurements and the inherent multiscale soil spatial structure. When coupled with Monte Carlo

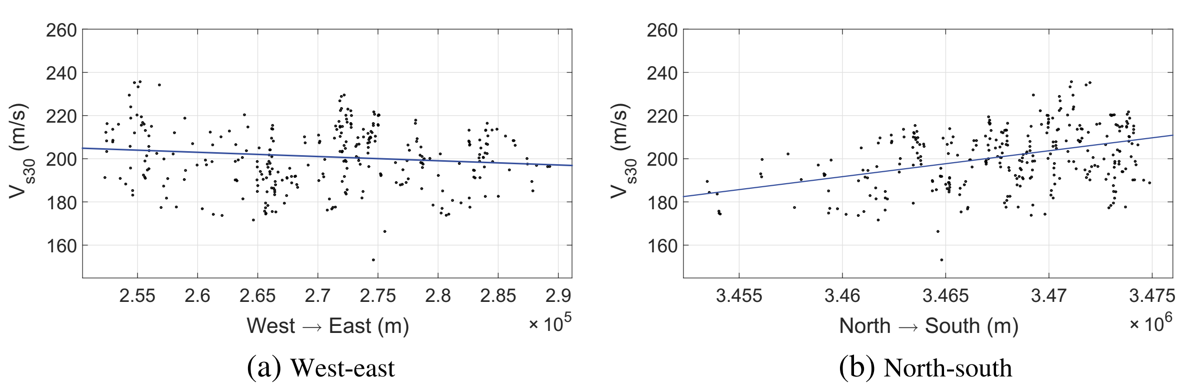
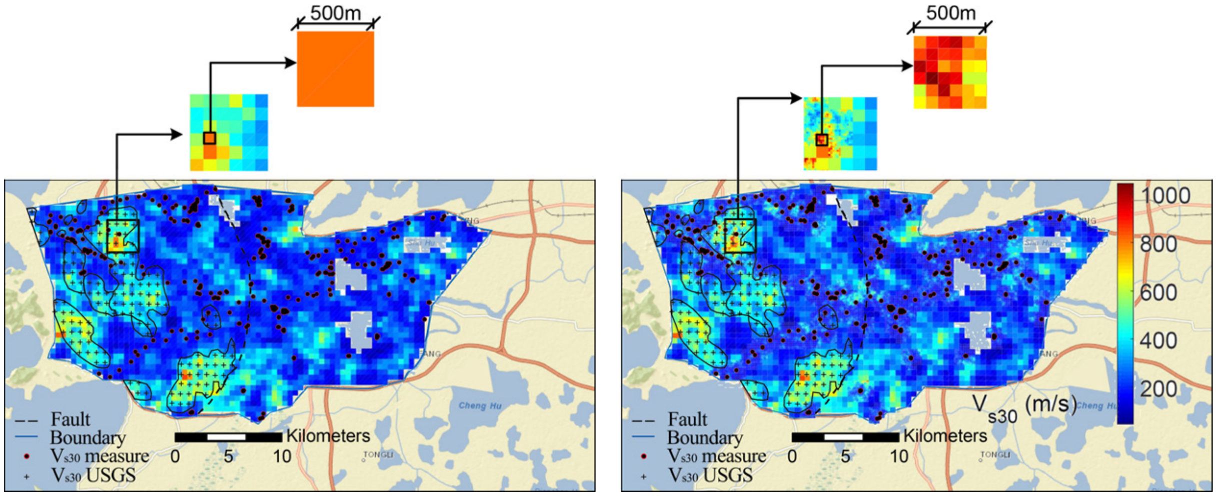


Fig. 7. Trend of the known V_{s30} values at measurement locations along (a) the west–east direction and (b) the north–south direction.



(a) Single scale (b) Multiscale

Fig. 9. Sample random field realizations of V_{s30} in Suzhou site.

simulations, uncertainties associated with the generated V_{s30} maps can also be estimated. The generated V_{s30} maps will be compared with the available topography-based V_{s30} map obtained from the U.S. Geological Survey global V_{s30} database.

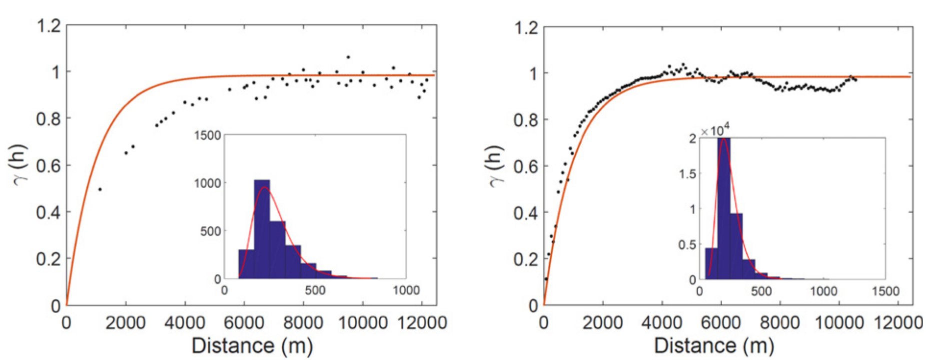
5.1. Random field realizations of V_{s30}

A typical set of V_{s30} realizations (single and multiscale) is shown in Fig. 9. In the multiscale realization, each coarse grid neighboring a measurement location is refined into 36 fine scale elements, where high resolution V_{s30} are generated through the multiscale model described in Section 3. Such fine scale field enables predictions across different scales and can facilitate estimation of uncertainties at much finer scales without sacrificing computation efficiency. The secondary V_{s30} data from USGS, placed on a grid with a spacing of 800 m, are incorporated as known point data values in the random fields in the outcrop areas. It should be noted that for the current study, the amount of the secondary data is fixed. A preliminary work (Liu et al., 2017) is undergoing to investigate the effect of secondary data on the spatial structure and the distribution of the resulting V_{s30} realizations.

The corresponding histograms and empirical semivariograms of the simulated V_{s30} are shown in Fig. 10. Both single and multi-scale simulations preserve the statistical characterizations and the spatial structures of V_{s30} inferred from the known measurement data.

Coupling the random field model with Monte Carlo simulations, the expected V_{s30} values across the Suzhou site as well as the associated uncertainties can be obtained. Maps of the expected V_{s30} values, averaged from 1000 independent Monte Carlo simulations, are shown in Fig. 11 (a) and (b). An obvious trend manifested in the map is that high V_{s30} values occur in the southern and western part of the city, especially the hilly areas. Low values are common in the northern and eastern part, which are consistent with the trends observed in the measurement data and the knowledge about the geology of this studied area. It should be noted that, in the current study, geological boundaries are not explicitly incorporated in the data reference or in random field simulations.

One of the strengths of the proposed method is its ability to estimate uncertainties associated with generated V_{s30} maps. To quantify uncertainties, coefficients of variation (COV) from 1000 independent Monte Carlo simulations are calculated at each location and plotted in Fig. 11 (c) and (d). As shown in the figure, the COVs are generally very small and approach zero around locations with measurement data. It is interesting to note that the uncertainties associated with single scale map are smaller compared to the multiscale counterpart, especially around locations with known data. Recall that the coarse (single) scale field can be seen as the average of the corresponding



(a) Single scale (b) Multiscale

Fig. 10. Semivariograms and histograms (the insets) of simulated V_{s30} from one set of random field realizations in Suzhou site. Black dots are the empirical semivariogram and the red solid line is the specified exponential model. The red solid line in the histogram inset is the fitted probability density function. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

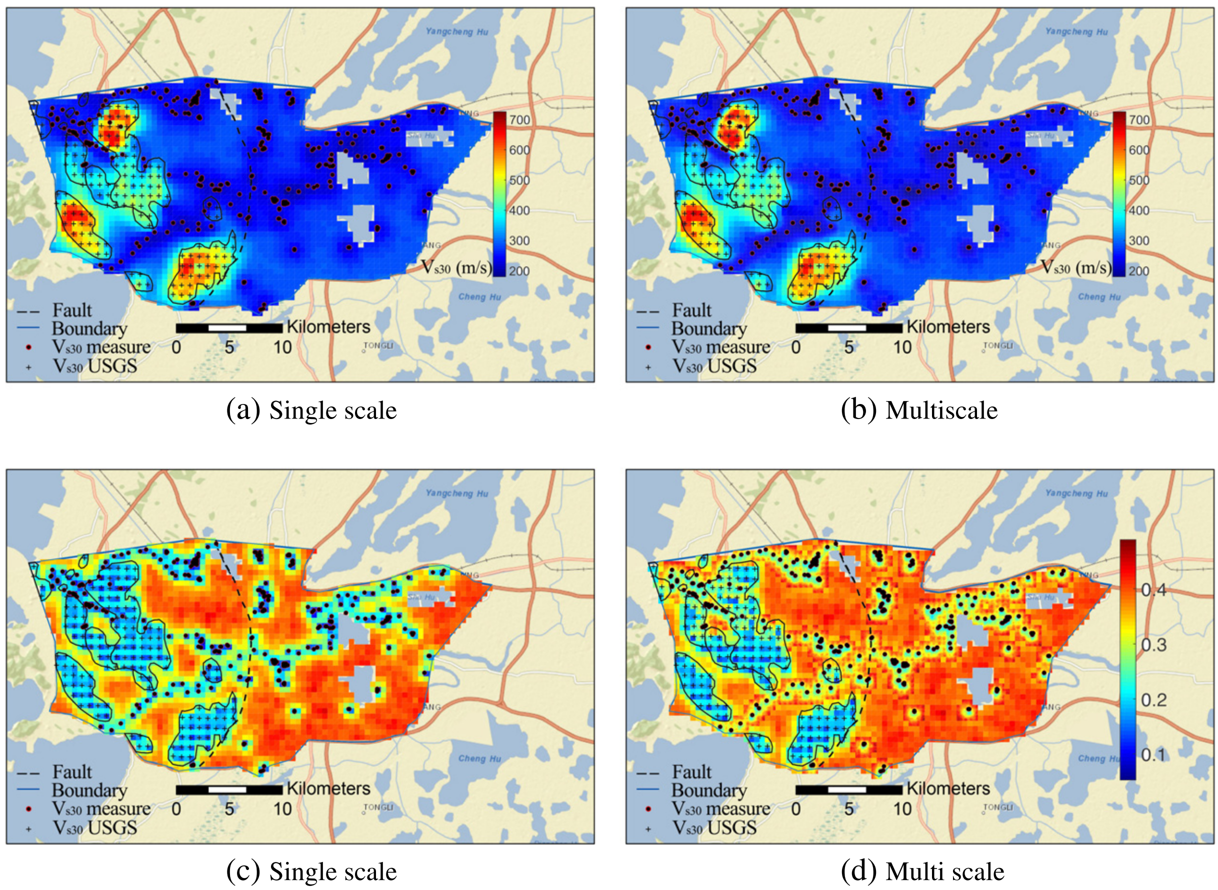


Fig. 11. Expected V_{s30} values and associated uncertainties (coefficient of variations) at the Suzhou site.

fine (multi) scale realizations and such averaging process results the reduced uncertainties observed in Fig. 11 (c) and (d).

The empirical semivariograms of the predicted V_{s30} values are calculated and shown in Fig. 12 along with the error bars indicating ±one standard deviation. It can be seen from Fig. 12 that the specified exponential spatial structure, which is inferred from measurement data, is preserved well in the simulations. It is noted that the spatial structures, quantified here by the semivariogram, are different between single and multiscale. This is because the coarse (single) scale spatial correlation is derived based on the notion that a coarse scale element is the average of the corresponding fine scale element. This averaging of the fine scale points will effectively increase the correlation of a given distance relative to the fine scale. This effect has been previously reported and studied in Chen et al. (2012).

5.2. Comparison with USGS V_{s30} maps

The newly generated multiscale random field-based V_{s30} maps incorporate and preserve the site-specific shear wave velocity measurement data and their spatial dependency. To understand the

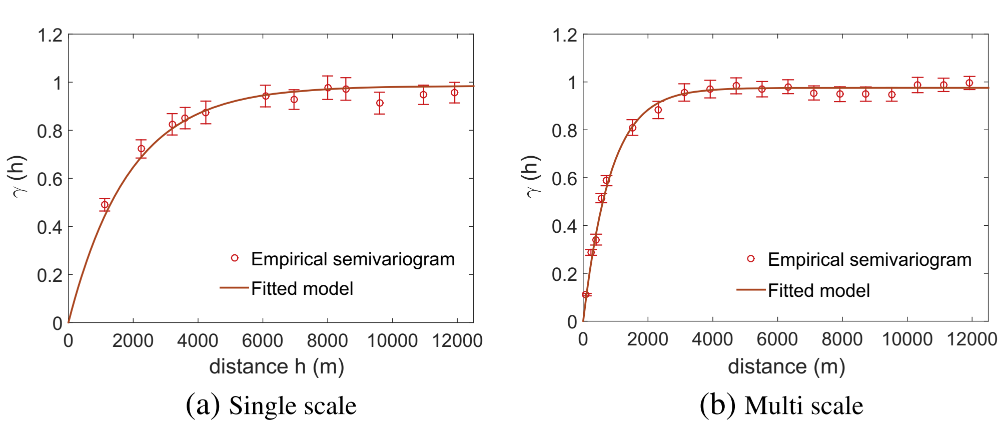
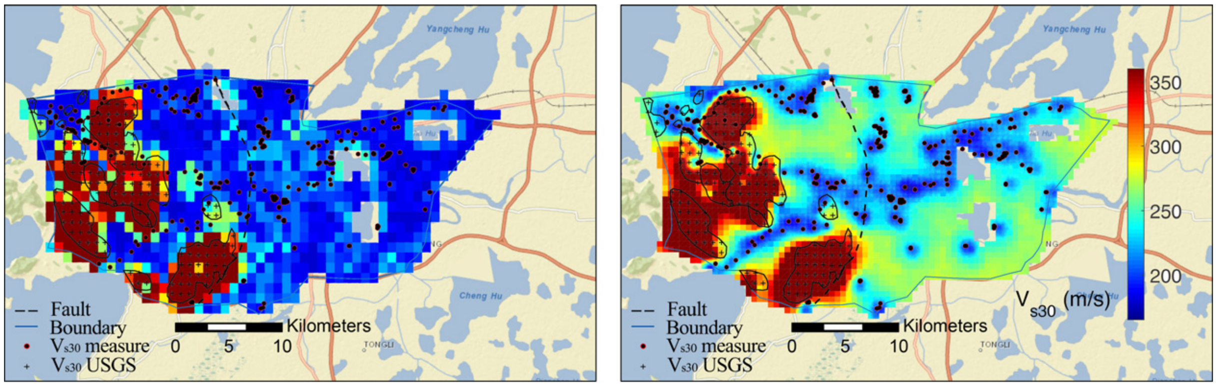


Fig. 12. Empirical semivariograms of predicted V_{s30}. Error bars indicate ±one standard deviation.



(a) USGS (b) Multiscale map

Fig. 13. Comparison of V_{s30} maps: (a) USGS topography-based proxy; (b) current study.

effect of local measurement data and spatial dependency on V_{s30} mapping, Fig. 13 plots side-by-side the V_{s30} map from the current study and the one from the USGS global V_{s30} map server. Note that the upper limit of the color map is set to V_{s30} = 360 m/s, which corresponds to the upper bound of the NEHRP site class D (refer to Table 4). Since most of the Suzhou site has soft soil with relatively low V_{s30} values, such scale makes the difference among two maps more distinguishable. As can be seen from Fig. 13, while both maps capture the general trend of high V_{s30} values in the western hilly area and low V_{s30} values in the eastern region, the current map has significantly higher resolution and has captured the transition from hilly to plain region fairly well. The current V_{s30} map captures a northeast–southwest band with low V_{s30}, as reflected from the V_{s30} measurement data, which is missed in the proxy-based USGS map. Moreover, the current map precisely preserves the known V_{s30} values at measurement locations and provides multiscale resolution, which contains small-scale V_{s30} information. Such information can be used to estimate uncertainties at a much higher resolution without sacrificing the overall computational efficiency.

To quantify the performance of the proxy-based USGS map, the difference between USGS V_{s30} values and the measured V_{s30} normalized by the measured V_{s30} value is calculated and the histogram of all 309 data is plotted in Fig. 14. As shown in Fig. 14, many of the normalized differences are within 0 to 40% range with a few points indicating over 100% difference.

6. Applications of the new V_{s30} maps

V_{s30} is a key indicator of site response in many earthquake engineering applications, such as ground-motion prediction equations, site classification, and earthquake hazard maps. In this section, two of the applications of the newly generated V_{s30} maps will be presented: V_{s30}-based site classification in Section 6.1 and the estimation of site amplification factors in Section 6.2.

Table 4
NEHRP site class and corresponding V_{s30} range.

Site class	Description	V_{s30}
A	Hard rock	>1500 m/s
B	Firm to hard rock	760 to 1500 m/s
C	Dense soil, soft rock	360 to 760 m/s
D	Stiff soil	180 to 360 m/s
E	Soft clay	<180 m/s
F	Soil requiring site specific evaluation	–

6.1. V_{s30}-based site classification

The National Earthquake Hazards Reduction Program (NEHRP) classifies a site into 5 groups and provides the range of V_{s30} values for each class as shown in Table 4. Given a V_{s30} map, the site of interest can be classified based on V_{s30} values following the NEHRP criteria.

Fig. 15 shows the site classification maps for the Suzhou site based on the new multiscale random field-based V_{s30} and the USGS proxy-based V_{s30} maps. The classification map of Fig. 15 (a) shows that most of the studied region can be classified as NEHRP soil type D, where V_{s30} ranges from 180 to 360 m/s. In the hilly area in the western part, the site is classified as soil type C with V_{s30} values ranging from 360 to 760 m/s. This is consistent with the known engineering geology of this region previously described in Section 2.

The site classification shown in Fig. 15 (a) is based on the expected V_{s30} values averaged from 1000 Monte Carlo simulations as previously shown Fig. 11 (b). To quantify the associated uncertainties in the site classification, upper and lower bound site classification maps are also generated by using ±one standard deviation of the expected V_{s30} values. The results are shown in Fig. 16. Compared to the mean V_{s30}-based site classification shown in Fig. 15 (a), most of the hilly areas in the western part of the city remain in the site class C, but the eastern plain changes to site E when the lower bound (mean minus

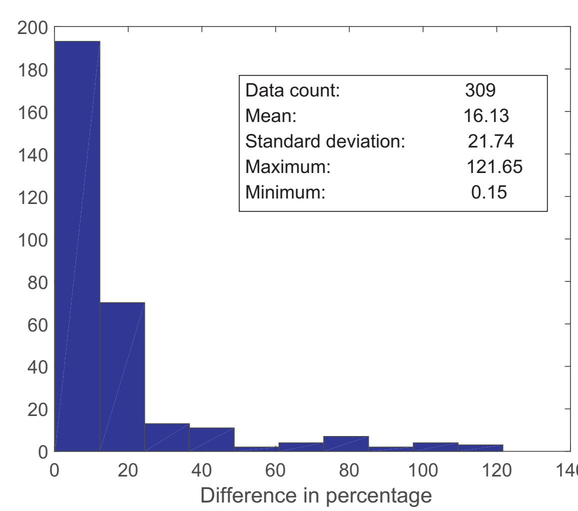
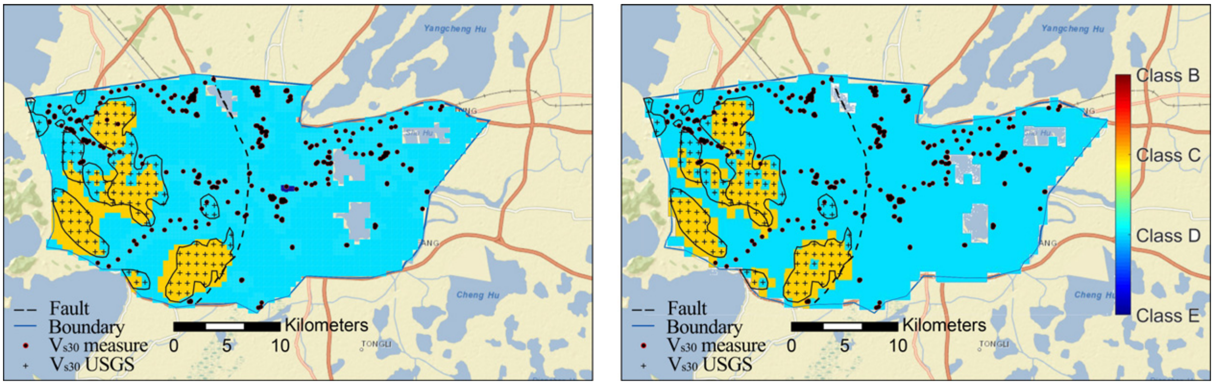
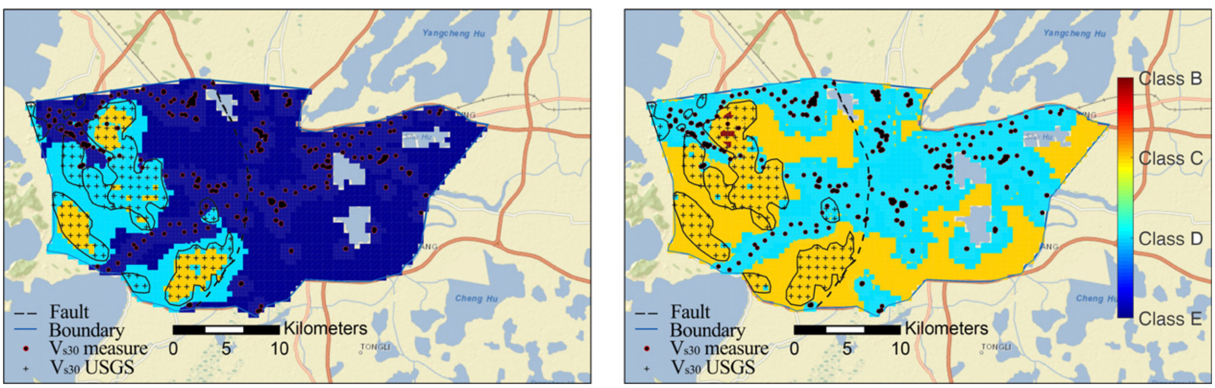


Fig. 14. Statistic characteristic of the difference (in percentage) between USGS V_{s30} prediction and known V_{s30} at 309 measurement locations.



(a) Multiscale (a) USGS

Fig. 15. V_{s30}-based NEHRP site classification (Table 4): (a) based on the new multiscale V_{s30} map; (b) based on the USGS V_{s30} map.



(a) Lower bound (b) Upper bound

Fig. 16. Uncertainties associated with the site classification maps based on expected V_{s30} values ±one standard deviation: (a) lower bound (mean −one standard deviation); (b) upper bound (mean +one standard deviation).

one standard deviation) V_{s30} map is used, which is considered to be a more conservative estimation.

6.2. Amplification factor mapping

The second application of the new V_{s30} map is the estimation and mapping of site amplification factors. Among various commonly used models for estimating site amplification factor, the model by Choi and Stewart (2005) is used in this work to illustrate the application. In the Choi and Stewart (2005) model, the model for estimating the amplification factor F_{ij} is expressed as

$$\ln(F_{ij}) = c\ln\left(\frac{V_{s30_{ij}}}{V_{\mathrm{ref}}}\right) + b\ln\left(\frac{\mathrm{PHA}_{r_{ij}}}{0.1}\right) + \eta_i + \epsilon_{ij} \tag{15}$$

where PHA_r is the peak horizontal acceleration for the reference site condition and is expressed in the unit of the gravitational acceleration g; b is a function of the regression parameters as given in Eq. (6.2); c and V_{ref} are the regression parameters; η_i is a random effect term for the i-th earthquake event with zero median and a standard deviation denoted as τ; ϵ_{ij} represents the intra-event model residual for the j-th motion in i-th earthquake event, which has a median near zero for well-recorded events with a standard deviation denoted as σ.

The variation of model parameter b is described in the following model (Choi and Stewart, 2005):

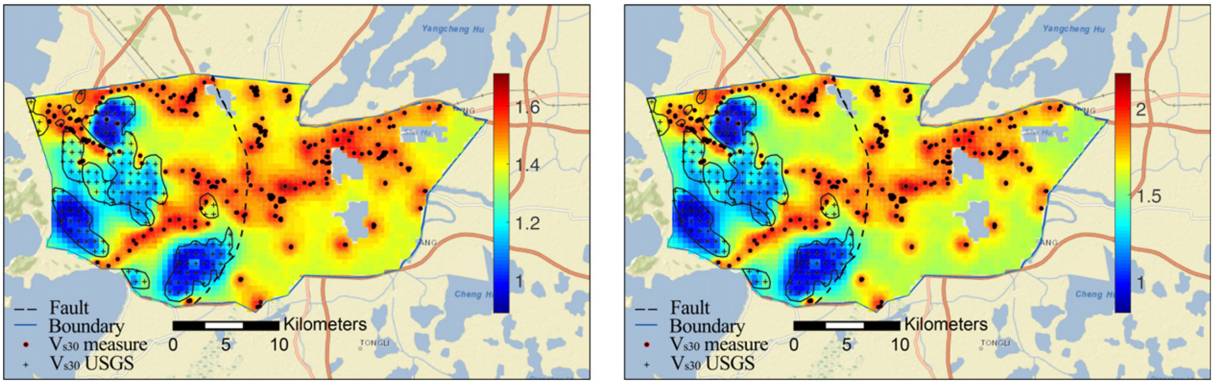
$$\begin{aligned}
&b = b_1 \qquad \text{Category E}\\
&b = b_2 + (V_{s30} - b_v)^2\frac{b_1 - b_2}{(180 - b_v)^2} \qquad 180 < V_{s30} < b_v\ (\mathrm{m/s})\\
&b = b_2 \qquad b_v < V_{s30} < 520\ (\mathrm{m/s})\\
&b = b_2 - (V_{s30} - 520)\frac{b_2}{240} \qquad 520 < V_{s30} < 760\ (\mathrm{m/s})\\
&b = 0 \qquad V_{s30} > 760\ (\mathrm{m/s})
\end{aligned} \tag{16}$$

where the units of V_{s30} are in m/s; b_1, b_2 and b_v are model parameters. For this reference model, Abrahamson and Silva (1997) provided values of site factor model parameters from regression analysis, which are summarized in Table 5.

With the amplification model Eq. (15), Eq. (6.2) and the fitting parameters in Table 5, site factors F_a (corresponding to a low-period

Table 5
Regression parameters for site amplification factors after Abrahamson and Silva (1997).

Parameter	b_1	b_2	b_v	c	V_{ref} (m/s)	τ	σ
F_a (0.3)	−0.41	−0.11	300	−0.46	532	0.35	0.54
F_v (1.0)	−0.39	0.02	300	−0.69	519	0.41	0.55



(a) T=0.3s (b) T=1.0s

Fig. 17. Maps of amplification factors in Suzhou City based on the Choi and Stewart (2005) model: (a) F_a ($T = 0.3$ s) and (b) F_v ($T = 1.0$ s).

range with $T = 0.1$–0.5 s) and F_v (corresponding to a mid-period range with $T = 0.4$–2.0 s) are calculated based on an assumed PHA$_r$ of 0.1 g. Results of the site factors are plotted in Fig. 17 for F_a ($T = 0.3$ s) and F_v ($T = 1.0$ s). Fig. 17 shows that most of the eastern and central areas have relatively high amplification factors with a maximum of 1.7 for $T = 0.3$ s and 2.2 for $T = 1.0$ s, which correlates well with the softer soils (NEHERP site classes D and E, refer to Fig. 15 (a)).

7. Conclusions

In this work, a multiscale random field-based framework is presented to map V_{s30} values over extended areas. The random field model explicitly accounts for the spatial variability of V_{s30} across different scales while incorporates and preserves measured V_{s30} data. The framework is applied to map V_{s30} over the Suzhou site, where 309 shear-wave velocity measurements and topography-based V_{s30} values are compiled. Monte Carlo simulations are coupled with the random field model to quantify uncertainties of the generated multiscale V_{s30} map. The new map is then applied to site classification and amplification factor characterization in the studied region. In summary, it is found that:

1. Quantitatively consistent V_{s30} estimates over different length scales over the entire studied region can be obtained using the multiscale random field model. The resulting map has multiscale resolutions and is particular convenient to incorporate and preserve local measurement data into a regional V_{s30} map.
2. Comparison of the new V_{s30} map with existing USGS topography-based V_{s30} map shows that the new V_{s30} map provides more accurate and more detailed V_{s30} values, especially in the eastern plain region of the studied site because of the incorporated local V_{s30} measurements and their spatial dependency.
3. Uncertainties associated with the new V_{s30} map are quantified in terms of the coefficient of variation (COV) calculated from Monte Carlo simulations. In general, the COVs approach zero around locations with measurement data and gradually increase in areas without any known V_{s30} values. COVs in single scale random field map are found to be slightly smaller when compared to the multiscale counterpart.
4. The site application map based on the newly generated V_{s30} map shows that relatively stiff soil (NEHRP site class C) is found in the northwestern part of the city and the soil tends to be softer in the southeastern region (NEHRP site class D and E). This trend in the soil type correlates well with the calculated amplification factor map, where high amplification factors are predicted in the southeastern part of the city, indicating potential seismic amplification effect in this region.

Acknowledgments

The first co-author wishes to acknowledge the financial support provided by the Glenn Department of Civil Engineering through the Glenn Professorship awarded to the fourth co-author and also acknowledge the support of the National Natural Science Foundation of China (51378258). Clemson University is acknowledged for the generous allotment of compute time on the Palmetto high-performance computing cluster.

References

Abrahamson, N, Silva, WJ, 1997. Empirical response spectral attenuation relations for shallow crustal earthquakes. Seismol. Res. Lett. 68 (1), 94–127.
Allen, TI, Wald, DJ, 2009. On the use of high-resolution topographic data as a proxy for seismic site conditions v_{s30}. Bull. Seismol. Soc. Am. 99 (2A), 935–43.
Ancheta, T, Darragh, R, Stewart, J, Seyhan, E, Silva, W, Chiou, B, Wooddell, K, Graves, R, Kottke, A, Boore, D, Kishida, T, Donahue, J, 2014. NGA-West2 database. Earthq. Spectra 30 (3), 989–1005.
Baker, JW, Seifried, A, Andrade, JE, Chen, Q, 2011. Characterization of random fields at multiple scales: an efficient conditional simulation procedure and applications in geomechanics. . Applications of Statistics and Probability in Civil Engineeringpp. 347–8.
Boore, DM, 2004. Estimating $\bar{V}_{s30}$ (or NEHRP site classes) from shallow velocity models (depths ¡30 m). Bull. Seismol. Soc. Am. 94 (2), 591–7.
Boore, DM, Thompson, EM, Cadet, H, 2011. Regional correlations of v_{s30} and velocities averaged over depths less than and greater than 30 meters. Bull. Seismol. Soc. Am. 101 (6), 3046–59.
Chen, Q, Seifried, A, Andrade, JE, Baker, JW, 2012. Characterization of random fields and their impact on the mechanics of geosystems at multiple scales. Int. J. Numer. Anal. Methods Geomech. 36 (2), 140–65.
Chen, Q, Wang, C, Juang, CH, 2015. CPT-based evaluation of liquefaction potential accounting for soil spatial variability at multiple scales. J. Geotech. Geoenviron. 04015077,
Chen, Q, Wang, C, Juang, CH, 2016. Probabilistic and spatial assessment of liquefaction-induced settlements through multiscale random field models. Eng. Geol. 211, 135–49.
Choi, Y, Stewart, JP, 2005. Nonlinear site amplification as function of 30 m shear wave velocity. Earthq. Spectra 21 (1), 1–30.
Goovaerts, P, 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.
Lee, C, Tsai, B, 2008. Mapping v_{s30} in Taiwan. Terr. Atmos. Ocean. Sci. 19 (6), 671–82.
Liu, W, Wang, C, Chen, Q, Chen, G, Juang, CH, 2017. Multiscale random field-based shear wave velocity mapping and site classification. Proceedings of the Geo-Risk 2017 Conference, Denver, Colorado.
Scasserra, G, Stewart, JP, Kayen, RE, Lanzo, G, 2009. Database for earthquake strong motion studies in Italy. J. Earthq. Eng. 13 (6), 852–81.
Seyhan, E, Stewart, JP, Ancheta, TD, Darragh, RB, Graves, RW, 2014. NGA-West2 site database. Earthq. Spectra 30 (3), 1007–24.
Thompson, E, Wald, DJ, Worden, C, 2014. A v_{s30} map for California with geologic and topographic constraints. Bull. Seismol. Soc. Am. 104 (5), 2313–21.

Thompson, EM, Baise, LG, Kayen, RE, 2007. Spatial correlation of shear-wave velocity in the San Francisco Bay Area sediments. Soil Dyn. Earthq. Eng. 27 (2), 144–52.
Thompson, EM, Baise, LG, Kayen, RE, Morgan, EC, Kaklamanos, J, 2011. Multiscale site-response mapping: a case study of Parkfield, California. Bull. Seismol. Soc. Am. 101 (3), 1081–100.
Thompson, EM, Baise, LG, Kayen, RE, Tanaka, Y, Tanaka, H, 2010. A geostatistical approach to mapping site response spectral amplifications. Eng. Geol. 114 (3), 330–42.
Wald, D, Earle, P, Quitoriano, V, 2004. Topographic slope as a proxy for seismic site correction and amplification. EOS. Trans. AGU 85 (47), F1424.
Wald, DJ, Allen, TI, 2007. Topographic slope as a proxy for seismic site conditions and amplification. Bull. Seismol. Soc. Am. 97 (5), 1379–95.
Wald, DJ, McWhirter, L, Thompson, E, Hering, AS, 2011. A new strategy for developing V_{s30} maps. Proc. 4th Int. Effects of Surface Geology on Seismic Motion Symp.
Wills, C, Clahan, K, 2006. Developing a map of geologically defined site-condition categories for California. Bull. Seismol. Soc. Am. 96 (4A), 1483–501.
Wills, C, Gutierrez, C, 2008. Investigation of geographic rules for improving site-conditions mapping. . California Geologic Survey Final Technical Report (07HQGR0061)
Wills, C, Petersen, M, Bryant, W, Reichle, M, Saucedo, G, Tan, S, Taylor, G, Treiman, J, 2000. A site-conditions map for California based on geology and shear-wave velocity. Bull. Seismol. Soc. Am. 90 (6B), S187–S208.
Wills, CJ, Silva, W, 1998. Shear-wave velocity characteristics of geologic units in California. Earthq. Spectra 14 (3), 533–56.
Yong, A, Hough, SE, Iwahashi, J, Braverman, A, 2012. A terrain-based site-conditions map of California with implications for the contiguous United States. Bull. Seismol. Soc. Am. 102 (1), 114–28.
Yong, A, Martin, A, Stokoe, K, Diehl, J, 2013. ARRA-funded v_{s30} measurements using multi-technique approach at strong-motion stations in California and central-eastern United States. . Technical Report, US Geological Survey
Zhan, JY, Chen, GX, Liu, JD, 2009. Empirical relationship between shear wave velocity and soil depth on deep soft sites in urban area of Suzhou City. World Earthq. Eng. 25 (2), 11–7.