

# DATA-DERIVED SITE RESPONSE AND ITS PREDICTABILITY USING ERGODIC AND SITE-SPECIFIC METHODS

Pengfei Wang and Jonathan P. Stewart

*Department of Civil & Environmental Engineering*

*University of California, Los Angeles*

## Abstract

We derive non-ergodic site response for California sites using an expanded version of the NGA-West2 database. We then investigate the degree to which different site response analysis methods capture observations. An ergodic site term provides a baseline against which other models are compared. Here we emphasize site-specific ground response analysis for sites with in situ  $V_S$  measurements. We describe the assignment of damping to individual soil layers using geotechnical models and site-specific spectral amplitude decay parameter  $\kappa$ . We provide data-model comparisons for cases in which ground response analyses provide variable levels of effectiveness.

## Introduction

Ergodic models for site response provide a mean estimate conditioned on certain site parameters (typically the time averaged shear wave velocity in the upper 30 meters of the site,  $V_{S30}$ , and basin depth). The ergodic estimate of site response includes all site amplification mechanisms (impedance, nonlinearity, resonance, two- and three-dimensional wave propagation in basins, etc.), but these effects are smoothed over a large number of sites with different characteristics. As such, the associated site-to-site uncertainties (denoted  $\phi_{S2S}$ ) are substantial, increasing mean or >50%tile ground motions at long return periods as derived from probabilistic seismic hazard analyses (PSHA) relative to what would be obtained with more accurate methods.

Site-specific or *non-ergodic* site response is intended to account for wave propagation mechanisms at a specific site that control site response. An unbiased estimate of site-specific site response, for example as derived from analysis of earthquake recordings, substantially reduces  $\phi_{S2S}$  (e.g., Rodriguez-Marek et al., 2014; Stewart et al. 2017). For sites without recordings, many projects seek to estimate site response using ground response analyses, which consider the effects of one-dimensional (1D) shear wave propagation and soil nonlinearity. Open questions related to this common practice are (1) How effective are such methods at capturing observed behavior, and how does this change with period?; and (2) What levels of epistemic uncertainty ( $\phi_{S2S}$ ), associated with wave propagation mechanisms not considered in 1D analysis, should be used in PSHA when site response is estimated from ground response analyses?

A sensible means by which to answer these questions is through comparisons of predictions of ground response analysis results to data. Not surprisingly, this general line of research contains

numerous contributions over many years, with a typical application taking various input motions, running them through 1D soil columns, and comparing resulting response spectra to those from recordings (e.g., Chang, 1996; Dickenson, 1994; Idriss, 1993). However, with the exception of vertical arrays, this research approach has a limited ability to answer the above questions, because predicted ground surface motions are strongly dependent on input motions, which are often highly uncertain. As a result, the effectiveness of the site response prediction is somewhat obscured.

The use of vertical arrays overcomes this problem because of the availability of recorded input motions, and has produced interesting findings that illustrate limitations, biases, and uncertainties associated with ground response analyses (e.g., Kaklamanos et al. 2013; Zalachoris and Rathje, 2015; Kaklamanos and Bradley, 2018; Afshari and Stewart, 2019). However, there are limitations associated with the use of vertical arrays to validate ground response analyses. First, the number of vertical arrays with sufficient ground motion recordings and site characterization is limited (but certainly growing with time). Second, vertical arrays only measure site response over the length domain of the array; as such they are not useful for evaluating long-period features that involve wavelengths longer than array dimensions. Third, the within-motion boundary condition that is used in analysis of vertical array data does not match that used in typical forward applications, in which outcropping input motions are selected.

To address these limitations, we suggest an alternative method for validating ground response analyses using data from surface-only instruments. The concept is to use recordings to infer the non-ergodic site response over a wide frequency range. The effectiveness of ground response analysis, and other methods, is then assessed by comparing predicted levels of site response against observation. This departs from the aforementioned prior work in that model effectiveness is not based on ground motions from a particular event (or series of events), but on the site amplification relative to a reference condition.

This work is in-progress, so final results are not provided here. We describe the approach, and summarize data assembled for this and related research. We then describe protocols that have been developed for applying ground response analyses at sites with  $V_s$  profiles but little of the other information typically required for such analyses (mainly, soil type and its variation with depth). Example results are presented and discussed.

## **Proposed Approach**

We suggest here a method that can be used to test the effectiveness of ground response analyses using the results of ground surface recordings. The method is substantially more robust when seismic velocity profiles are available at candidate sites, and we apply this constraint in the present work. The method has four components.

The first component is assembling the required data. If not already available from another project (such as NGA projects, which involve substantial data collection and synthesis), this is a substantial task. The information required is identical to that needed for ground motion model development, namely, a database that includes information on source attributes, site conditions at recording stations, and ground motions (with record-specific processing details). In this project, we supplemented the NGA-West2 database with additional sites and events, as described in the

next section. A need for the present work that is not shared in ground motion model development projects, is seismic velocity profiles at recording stations (particularly shear-wave velocity,  $V_s$ , vs depth profiles).

The second component consists of ground motion analyses targeted at extracting information on site responses at recording stations. The steps involved in developing these results are described elsewhere (Stewart et al. 2017), so the procedure is not repeated here. What these analyses provide is an estimate of a site term, denoted  $\eta_s$ , for each site and response spectral oscillator period. This site term represents the mean difference between a regionally-unbiased ground motion model and observed motions at the site. For weak shaking conditions that do not induce soil nonlinearity, the sum of  $\eta_s$  and the ergodic site term for the site ( $F_S$ ), comprises the mean non-ergodic site response ( $\mu_{lnY}$ ) relative to the ground motion model's reference condition:

$$\mu_{lnY} = F_S + \eta_s \quad [1]$$

The third component consists of predicting site response for each site in the data inventory using available information on site conditions. In the case of ground response analyses, a  $V_s$  profile is required, and borehole data indicating soil stratigraphy and soil type characteristics for each layer is also useful (for estimation of modulus reduction and damping relations). Other methods may require different information, such as peak frequency from H/V spectral ratios (e.g., Kwak et al, 2017; Hassani and Atkinson, 2016).

The fourth component involves model-to-data comparisons in the form of residuals analyses, which can be used to estimate model bias and uncertainty. These procedures, and the interpretation of results, will be presented in subsequent publications.

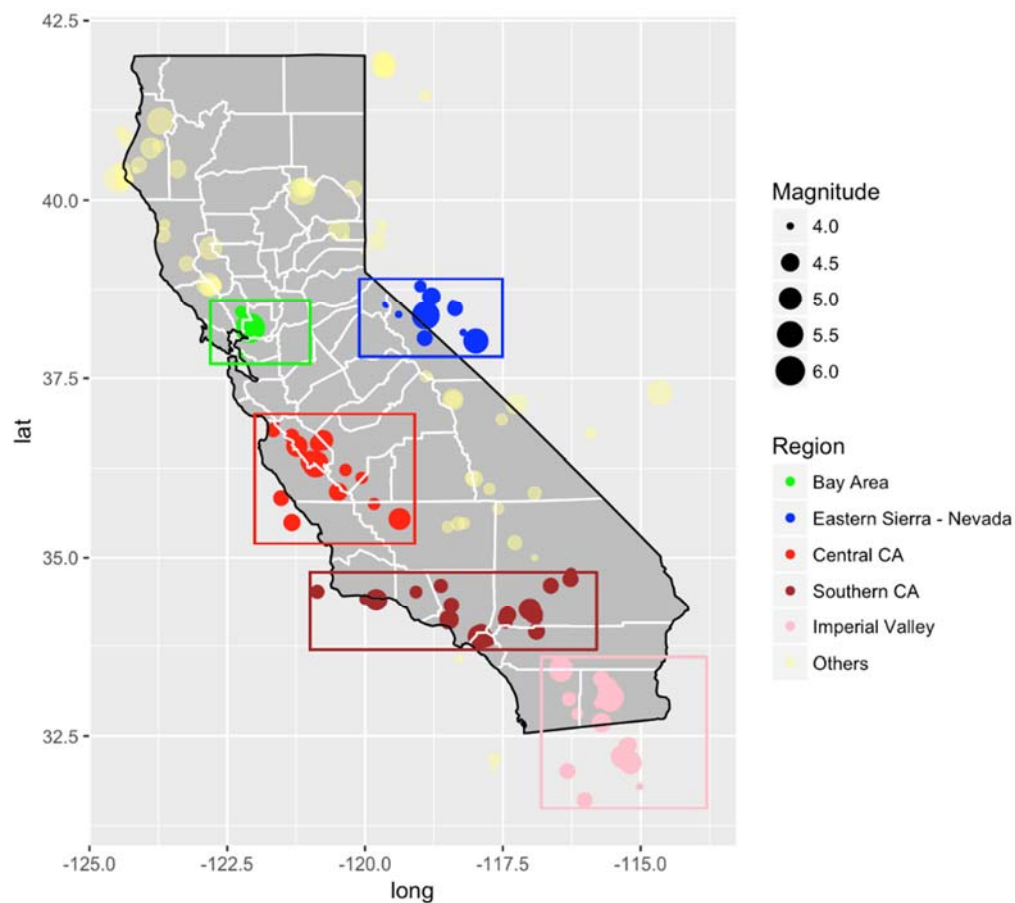
## Database

The approach described in the prior section requires a large database with many recordings for the second component (referred to here as *Full Database*). The database requirements for this component match those for ground motion model development projects. A subset of that database is used in components three and four (*Database Subset for Site Response Studies*).

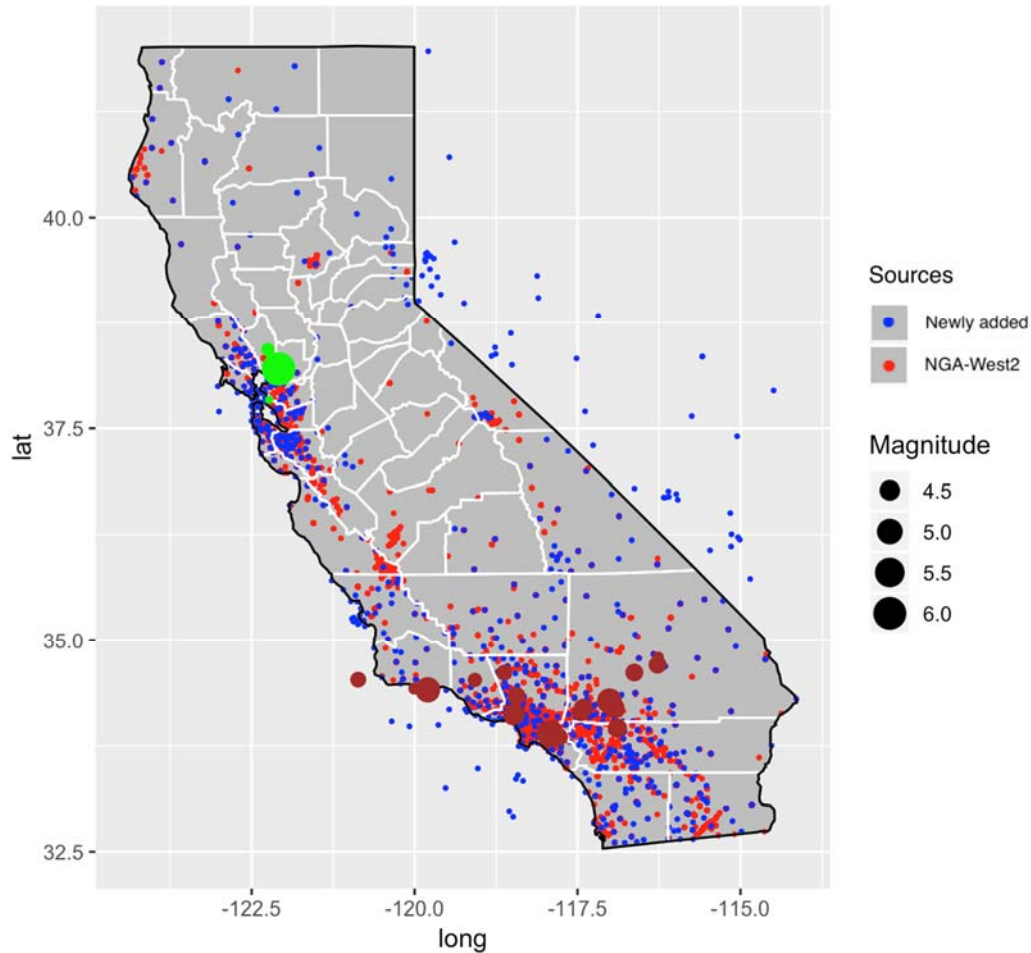
### Full Database

The database used in this study draws from an expanded version of the NGA-West2 database (Ancheta et al., 2014), which is a global database for active tectonic regions. There is a significant contribution of data from California to the NGA-West2 database (373 events, 1463 stations, 14231 recordings) over the time period 1938 to 2010. The site portion of the database (Seyhan et al. 2014) was developed to provide the principal site parameters used in model development –  $V_{S30}$  and various depth parameters denoted as  $z_x$ . As part of this project and other complimentary projects, we converted the spreadsheet files that comprised the original NGA-West2 flatfile (pertaining to sources, sites, and ground motions) into tables within a relational database, which is housed on a local server. Data modifications and additions are made within the relational database. The database is accessed using Python scripts within Jupyter notebooks on DesignSafe (Rathje et al. 2017).

We have identified earthquakes and recordings since 2011 in California, which significantly extend the NGA-West2 database. Figure 1 shows the locations of events sorted by magnitude, most of which occur in five main regions: Bay Area, Eastern Sierra and Nevada, central California, southern California, and Imperial Valley and northern Mexico. These five zones incorporate most of the urban areas in the state, and contain a large fraction of the ground motion stations. We focus here on the Bay Area and southern California regions. Moreover, since difficulties can be encountered in the analysis of site terms using small magnitude data, we only consider  $M \geq 4.0$  events (Stafford et al. 2017). The data from events within the Bay Area and southern California regions in Figure 1 is derived from 25 earthquakes that have produced about 9,300 three-component recordings within the distance cutoffs suggested by Boore et al. (2014). These data are screened for magnitude (requiring  $M > 4$ ), to remove duplicate recordings (e.g., seismometers and accelerometers at the same location), and to remove recordings that appear to be unreliable from instrument malfunctions or similar. This leaves about 5873 usable three-component records. Figure 2 shows the locations of these events and of the 1185 recording stations with recordings.

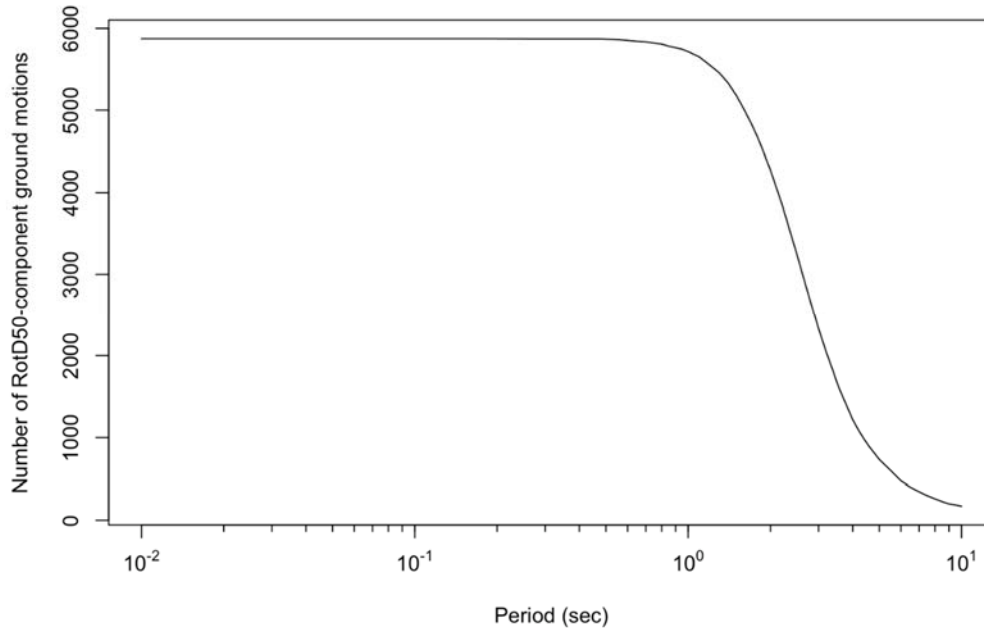


**Figure 1.** Locations of earthquakes in California and northern Mexico with  $M \geq 4.0$  since 2011 for which ground motion data has been compiled for addition to the NGA-West2 database



**Figure 2.** Map of California showing locations of considered earthquakes with  $M \geq 4.0$  since 2011 and locations of stations that recorded the events (blue – new stations, red – stations in NGA-West2 database)

Each of the three-component records has been processed according to standard protocols developed during Pacific Earthquake Engineering Research center (PEER)-NGA projects, as described in Ancheta et al. (2014). This processing provides a lowest usable frequency for each ground motion component. Horizontal ground motion components are combined to median-component (RotD50) as defined by Boore (2010) using the routines given in Wang et al. (2017). We take the lowest useable frequency for RotD50 as the higher of the two as-recorded values. Figure 3 shows the number of usable RotD50 horizontal-component ground motions as a function of oscillator period. The fall-off begins at about 1.0 sec and the data is reduced by 50% at 2.5 sec.



**Figure 3.** Number of usable RotD50-component ground motions as a function of oscillator period for the data added for the Bay Area and southern California regions.

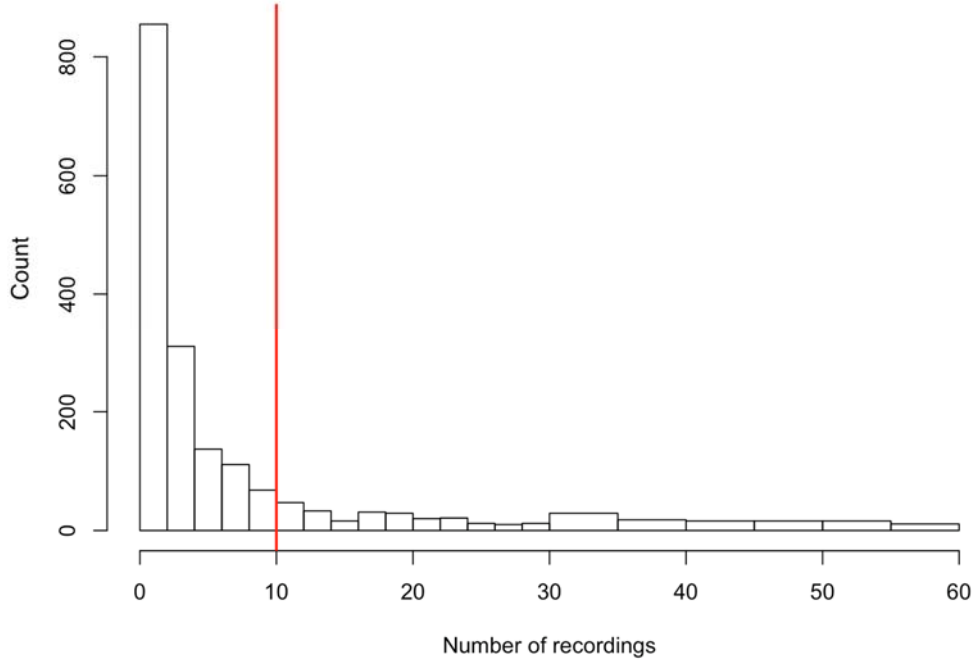
Considering both the NGA-West2 data and new data, there are 1818 recording sites shown in Figure 2. Of those, 1340 are sites that were included in the NGA-West2 site database. Hence, there are 478 new sites that require assignment of site parameters. We use measured  $V_S$  profiles to compute  $V_{S30}$  when available, and in the absence of this data, we use proxies (slope gradient – Wald and Allen (2007); terrain category – Yong et al., 2012 and Yong 2016; surface geology – originally by Wills and Clahan (2006) and Kriging interpolated by Thompson et al 2014, and later updated by Wills et al. 2015 and Thompson 2018. We have evaluated correlations among these proxies and used this information to develop model weights in a manner similar to that described in Kwok et al. (2018). This work will be documented subsequently, and resulted in the following weights:

- Surface geology with local data adjustment: 0.665
- Terrain categories: 0.323
- Surface gradient: 0.012

### Database Subset for Site Response Studies

A subset of the full database is applied for site response studies. The criteria used to define this subset are: (1) a minimum number of recordings per site of 10 (applied to ensure statistically robust estimates of site term,  $\eta_S$ ); (2) availability of a  $V_S$  profile for the site.

Figure 4 shows a histogram of the number of recordings at stations in the full database. Of the 1818 sites in the full database, 366 meet the minimum recordings/site criterion.



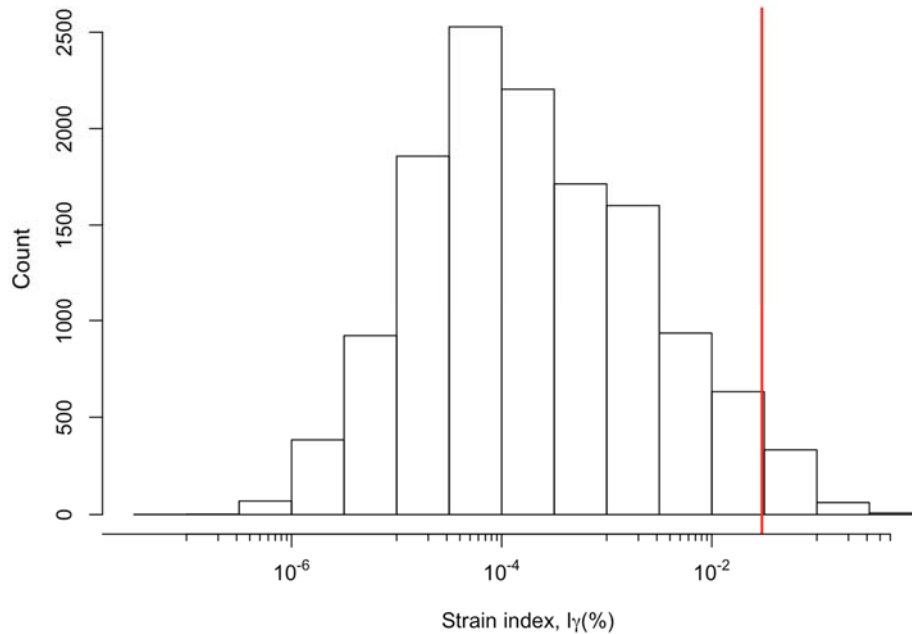
**Figure 4.** Histogram of number of recordings at stations in the full database. Ten is the minimum number of records/station for sites considered in the present research.

We performed a search for  $V_s$  profiles for each of the sites meeting the first criterion. This was done using the shear wave velocity profile database compiled for California by Ahdi et al. (2018). We find 149 sites with a  $V_s$  profile within around 200 m of the strong motion site. Many of these profiles are from Yong et al. (2013), which provides  $V_s$  profiles from various surface wave tests and H/V spectral ratios from microtremors. Of the 149 sites with  $V_s$  profiles, only 3 have a boring log that indicates stratigraphic details and soil/rock layer descriptions. This geotechnical data is needed to apply models for modulus reduction and damping as a function of shear strain.

Most of the recordings used in this research involve low ground motion amplitudes. Figure 5 shows a histogram of the ratio (strain index):

$$I_\gamma = \frac{PGV}{V_{S30}} \quad [2]$$

where  $PGV$  is from the surface recording and is taken from the RotD50 component. This ratio provides an index related to shear strain (Idriss, 2011; Kim et al, 2016), and can be used to judge the degree to which soil responses are likely to be affected by nonlinearity. As shown in Figure 5, 97% of ground motions in the subset have  $I_\gamma < 0.03\%$ . We conclude that the soil responses are predominantly in the linear range, meaning that modulus reduction is unity and damping is at the minimum value. We refer to the minimum damping from geotechnical models (Darendeli 2001 for soils with fines; Menq 2003 for granular soils) as  $D_{min}^L$ . As a result, the primary need for stratigraphic and material description information is to define  $D_{min}^L$  as a function of depth.



**Figure 5.** Histogram of strain index number of recordings at stations in the full database.

To derive  $D_{min}$  profiles for use in ground response analyses, the next two sections describe (1) how stratigraphy was inferred to enable estimates of  $D_{min}^L$  (for sites without borehole logs); and (2) how site spectral amplitude decay parameter ( $\kappa$ ) was measured from recordings and then interpreted to constrain small-strain damping. As such, these sections support the development of alternative damping profiles, each of which are being considered in the validation analyses.

### **Inference of Unit Weight and Material Damping**

Ground response analyses for linear conditions require shear wave velocity, unit weight, and  $D_{min}$  profiles. Shear wave velocity profiles are measured at each of the sites in the *Database Subset for Site Response Studies*. In most forward applications, geotechnical site characterization provides borehole logs that describe site stratigraphy and soil type information, which can be used to derive the input parameters used to predict unit weight and  $D_{min}^L$ . As described above, this is not the case for many of the sites considered in this research. This section describes how we estimate unit weight and soil parameters used to estimate material damping.

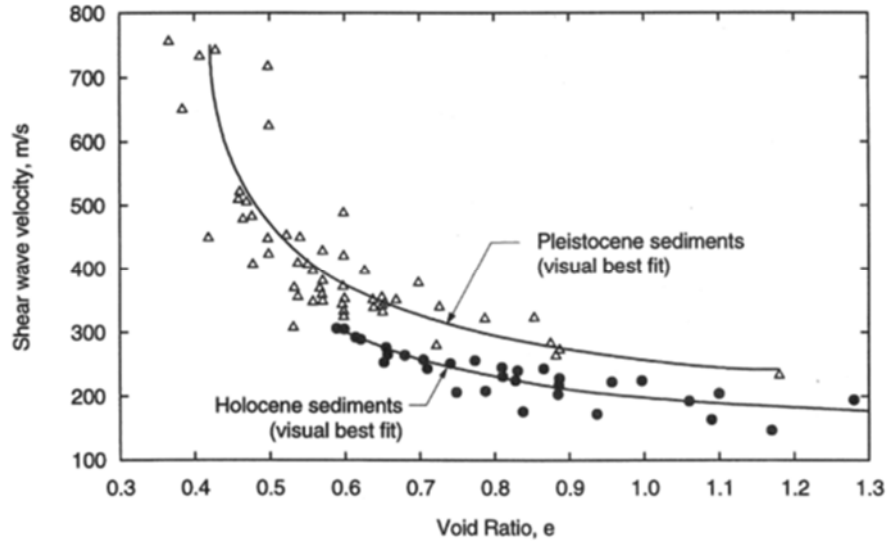
#### **Unit Weight**

For soil units, we estimate unit weight using phase relationships, which relate unit weight to void ratio, specific gravity, and saturation. Void ratio is taken from an empirical relationship with  $V_s$  shown in Figure 6 and given as (Rogers et al., 1985):



$$V_s = 42.9 + 94.1/e^2 \quad [3]$$

where  $V_s$  is in units of m/s.



**Figure 6.** Empirical relationships between void ratio, age, and shear-wave velocity for alluvial sediments in southern California (Fumal and Tinsley 1985). Eq. [3] fits the combined data (Rogers et al 1985)

Specific gravity is commonly taken as  $G_s = 2.7$ . Saturation ( $S$ ) is taken as 1.0 below the first depth where  $V_P$  exceeds 1500 m/s. Above that depth, or over the full depth where  $V_P$  data is absent, saturation is assumed as 50%. Total unit is then computed as:

$$\gamma = \frac{G_s \gamma_w}{1 + e} \left( 1 + \frac{eS}{G_s} \right) \quad [4]$$

where  $\gamma_w$  is the unit weight of water (10 kN/m<sup>3</sup>).

For rock units, we assigned unit weight based on  $V_s$  as follows:

$$\gamma = \begin{cases} 20 \text{ kN/m}^3, & \text{if } 450 < V_s < 700 \text{ m/s} \\ 22 \text{ kN/m}^3, & \text{if } V_s > 700 \text{ m/s} \end{cases}$$

### Stratigraphy and Soil Type to Estimate $D_{\min}$ in Soil Layers

Stratigraphic and soil type information is needed to apply the geotechnical model for  $D_{\min}$  estimation by Darendeli (2001), which is conditioned on plasticity index (PI), over-consolidation ratio (OCR), and mean effective stress. Effective stress can be calculated using unit weights from the prior section and water table depth (as applicable). PI and OCR are generally derived from laboratory tests on samples retrieved from the field.

We consider two types of available information as potentially useful to assign stratigraphy and soil type information – the mapped surface geology and the  $V_s$  profile. Surface geology is

used to estimate soil type near the ground surface. The  $V_s$  profile is used in combination with the surface unit assignment to estimate variations with depth.

Surface geology is taken from state-wide geologic maps by Wills and Clahan (2006) and Wills et al. (2015). We assume relationships between surface geological unit and PI/OCR, with details indicated in Table 1. Considerations in the development of the relationships in Table 1 include:

- Geologically young sediments (Holocene) are assumed to have low OCR, and older units are assumed to have relatively high OCRs. The rationale is that young deposits have relatively limited pseudo-overconsolidation from ageing and are unlikely to have experienced significant unloading from natural geological processes.
- Young sediments deposited in quiescent environments (e.g., bays, lakes, central/flat portions of alluvial basins) are assumed to be relatively fines- and clay-rich, thus having high PI. Young alluvial sediments deposited on steeper gradients are assumed to be relatively granular (PI = 0).
- Tertiary sedimentary bedrock units often carry information on rock type (e.g., shale, sandstone, etc.). We assume the bedrock units are similar to corresponding soil units (i.e., shale and sandstone interpreted as clay and sand, respectively).
- For pre-Quaternary units without information on material type or depositional environment, there is no basis for relatively coarse- or fine-grained behavior. We assume an intermediate condition in this case (roughly corresponding to low-plasticity silt).

Table 1 is organized in reference to 12 geological units that the stations in full database encountered and recommended by Wills and Clahan (2006): Qal1, Qal2, and Qal3 are relatively young alluvial sediments likely to be of Holocene age; Qoa is older alluvium of Pleistocene age; QT describes sediments in the early Pleistocene to Pliocene periods, for which the method of deposition is unknown; Tsh, Tss, and Tv comprise Tertiary age bedrock of consisting of shale, sandstone, or volcanic origin-materials (typically basalt or rhyolite), respectively; serpentine is a metamorphic rock of Tertiary age largely comprised of the clay mineral serpentinite; and Kss, Kjf, and crystalline are hard rock, typically of Cretaceous age.

Before assigning one of the hard rock classes (Kss, Kjf, crystalline), we perform a visual check of morphology using Google Earth™. When this check indicates that the surface appears to be soil, and if the velocity of the nearest-surface layer is compatible with soil, we assign a soil surficial unit and assign rock at greater depths where velocities become fast.

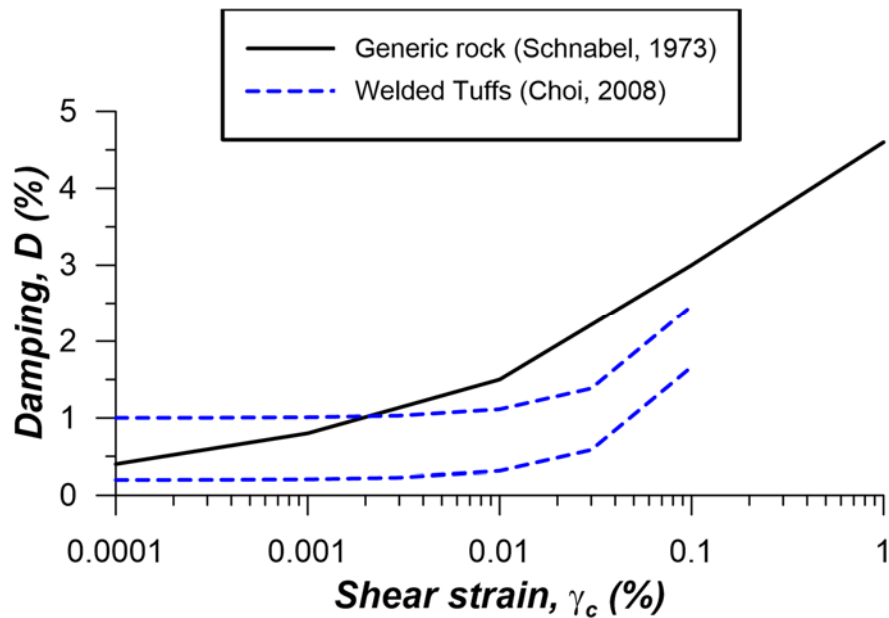
The soil property assignments in Table 1 apply for ground surface layers. The assignment of properties at depth is made in consideration of gradients in the  $V_s$  profile. If the surface layer consists of sediments or Tertiary rock, the soil index properties are not changed in successive layers absent sudden changes in velocity with depth. Sudden changes can trigger soil type changes – for example, when a granular layer is underlain by a much slower layer, the underlying unit is taken as clay. Similarly, when a fine-grained surface layer is underlain by a much stiffer layer, the underlying material is taken as granular. When a layer velocity exceeds 760 m/s, it is taken as rock. Figure 7 shows the flow chart used to assign soil type information as a function of depth.

**Table 1.** The list of 12 geological units and their corresponding PI and OCR. Ma indicates million years.

Geological age	Geol. unit	Description	Estimated Parameters
Holocene (< 0.011 Ma)	<b>Qal1</b>	Quaternary Holocene alluvium with flat gradients (< 0.5%).	PI = 30 OCR = 1.2
	<b>Qal2</b>	Quaternary Holocene alluvium with moderate gradients (0.5 - 2.0%).	PI = 10 OCR = 1.2
	<b>Qal3</b>	Quaternary Holocene alluvium with steep gradients (> 2%).	PI = 10 OCR = 1.2
Pleistocene (< 2.6 Ma)	<b>Qoa</b>	Quaternary Pleistocene alluvium. Soil composition unknown.	PI = 10 OCR = 2
Pliocene (2.6-5.3 Ma). Young era within the Tertiary.	<b>QT</b>	Quaternary to Tertiary deposits, including Saugus Fm. in So. CA, Paso Robles Fm. in central Coast Ranges, and Santa Clara Fm. in San Francisco Bay area. Soil composition unknown.	PI = 10 OCR = 2.5
Tertiary (2.6-66 Ma).	<b>Tsh</b>	Shale and siltstone units, such as the Repetto, Fernando, Puente, and Modelo Fms in So. CA.	PI = 15 OCR = 3
	<b>Tss</b>	Sandstone units, such as the Topanga Formation in So. CA and Butano Formation in San Francisco Bay area.	PI = 0 OCR = 3
	<b>Tv</b>	Volcanic units including the Conejo Volcanics in Santa Monica Mtns and the Leona Rhyolite in East Bay Hills.	PI = 15 OCR = 3
	<b>Serpentine</b>	Serpentine rock is clay-rich.	PI = 15 OCR = 3
Cretaceous	<b>Kss</b>	Cretaceous sandstone of the Great Valley Sequence	NA
	<b>Kjf</b>	Franciscan complex rocks, including mélangé, sandstone, shale, chert, and greenstone.	NA
	<b>crystalline</b>	Crystalline rocks, including Cretaceous granitic rocks, Jurassic metamorphic rocks, schist, and Precambrian gneiss.	NA

### $D_{min}$ in Firm Rock Layers

The Darendeli (2001) model cannot be used for pre-Tertiary rock (units Kss, Kjf, crystalline). Laboratory data on material damping for such materials is limited. A presumably judgement-based model was presented by Schnabel (1973) and has been widely used since that time. Choi (2008) performed testing on welded Bandelier Tuff and Topopah Spring Tuff and developed damping models. Models from these two sources are compared in Figure 8. The  $D_{min}$ -component from Topogah Spring Tuff is considered more representative of bedrock materials in our study region based on its unit weight (Bandelier Tuff has low unit weights). The  $D_{min}$  range for this material is about 0.2 – 1.0% (average = 0.3%). We have used the Choi model for the present work, but acknowledge that its use carries large uncertainty.



**Figure 8.** Comparison of rock damping model from Schnabel (1973) and range from Choi (2008).

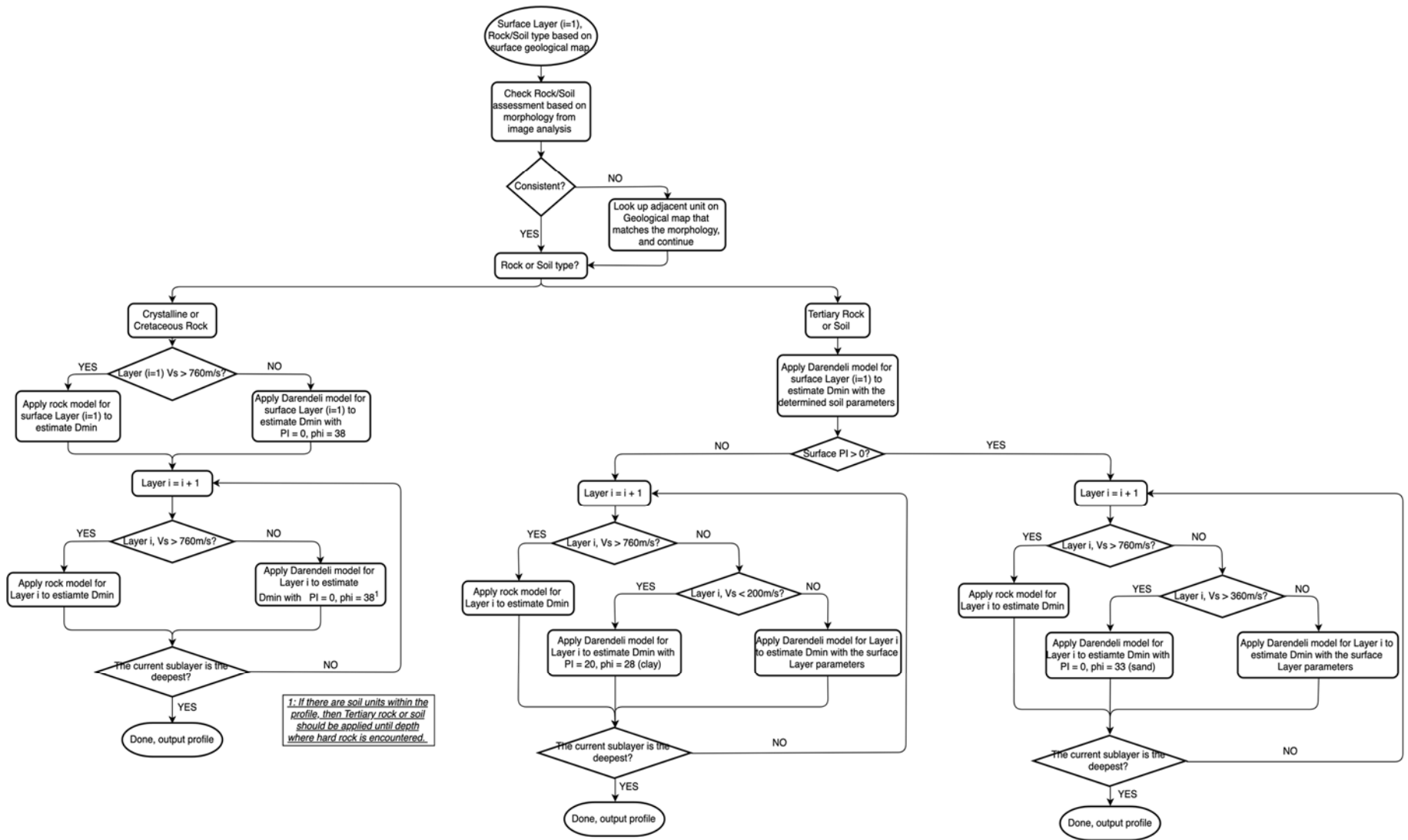


Figure 7. Flow chart used to assign soil type information as function of depth

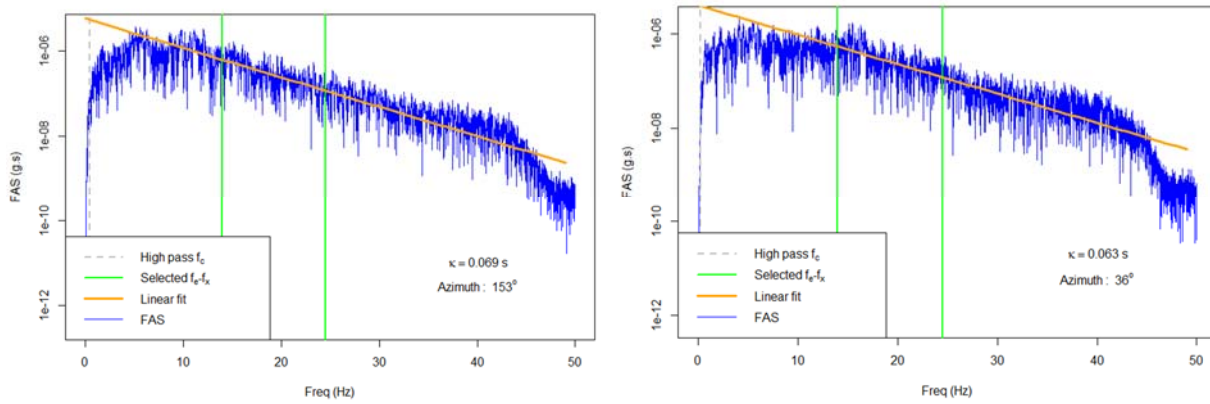
## $\kappa$ -Informed Damping Model

### Approach

Figure 9 shows Fourier amplitude spectra for ground motions at two examples sites in our database. The spectra show a characteristic feature, which is decay of Fourier amplitudes with increasing frequency for frequencies beyond the peak in the spectrum. This frequency-dependent decay can be described as:

$$D(f) = \exp(-\pi\kappa f) \quad [6]$$

where  $f$  is frequency in Hz and  $\kappa$  is a decay parameter that can be established through fits to data (e.g., Anderson and Hough, 1984).



**Figure 9.** Analysis of  $\kappa$  from recordings at two example sites (left: AZ.KNW, right: AZ.PFO)

The decay parameter arises from material damping and wave scattering that occurs on the wave path from source-to-site, often including appreciable contributions from site response. The path and site response contributions to  $\kappa$  combine as (adapted from Anderson 1991):

$$\kappa = \kappa_0 + \kappa_R R \quad [7]$$

where  $R$  is site-to-source distance,  $\kappa_R$  is the slope by which  $\kappa$  increases with distance, and  $\kappa_0$  represents the cumulative effect of damping and wave scattering through the soil column.

The relationship between  $\kappa_0$  and profile attributes can be expressed as (Hough and Anderson 1988; Chapman et al. 2003; Campbell 2009):

$$\kappa_0 = \int_0^{z_s} \frac{dz}{Q_{ef}(z)V_s(z)} \quad [8]$$

where  $z_s$  is the site column thickness (depth to reference crustal rock) and  $Q_{ef}(z)$  is the depth-dependent effective material quality factor, representing both the effects of frequency-dependent wave scattering and frequency-independent soil damping.  $Q_{ef}$  can be converted to an effective soil damping as follows (Campbell, 2009):

$$D_{ef}(\%) = \frac{100}{2Q_{ef}} \quad [9]$$

Measurements of  $\kappa$  from recordings can, in principal, inform levels of damping applied in ground response analyses as follows:

1. Measure  $\kappa$  for a set of sites from multiple earthquakes, as shown for example in Figure 9.
2. Develop a regionally appropriate model for  $\kappa_R$ .
3. Adjust each measured value of  $\kappa$ , for each event recorded at a given site, to estimate  $\kappa_0$  by re-arranging Eq. (7) as  $\kappa_0 = \kappa - \kappa_R R$ .
4. Since the soil/rock column thickness analyzed in ground response analysis is typically smaller than the full profile to reference crustal rock ( $V_S \approx 2.5\text{-}3$  km/s), adjust  $\kappa_0$  from Step (3) as,

$$\Delta\kappa = \kappa_0 - \kappa_{0,b} = \int_0^{z_p} \frac{2D_{ef}(z)}{100} \frac{dz}{V_S(z)} \quad [10]$$

where  $\kappa_{0,b}$  is the site decay parameter at the base of the profile and  $z_p$  is the depth of the analyzed soil column ( $z_p < z_s$ ).

5. Modify the laboratory damping with a profile-specific adjustment factor  $F_D$  to match  $\Delta\kappa$  from Step (5), which can be represented by re-writing Eq. (10) as:

$$\Delta\kappa = \int_0^{z_p} \frac{2D_{min}^L(z) \times F_D}{100} \frac{dz}{V_S(z)} \quad [11]$$

The depth-invariant value of  $F_D$  represents the means by which the field observations of kappa inform the damping model. In some cases,  $F_D$  may be unreasonably high. To constrain  $F_D$  so that it provides damping values within a realistic range, we have enforced a maximum value of  $F_D = 5$ . The approach maintains the scaling of damping with soil type and depth in the laboratory models, while adjusting for other effects encountered in field conditions (scattering).

Implementation of the above procedure requires several model components – distance correction term  $\kappa_R$  and site decay parameter for the base of profile condition  $\kappa_{0,b}$ . The following sub-sections describe the calculation of  $\kappa$  from recordings, models used for these components, and example results.

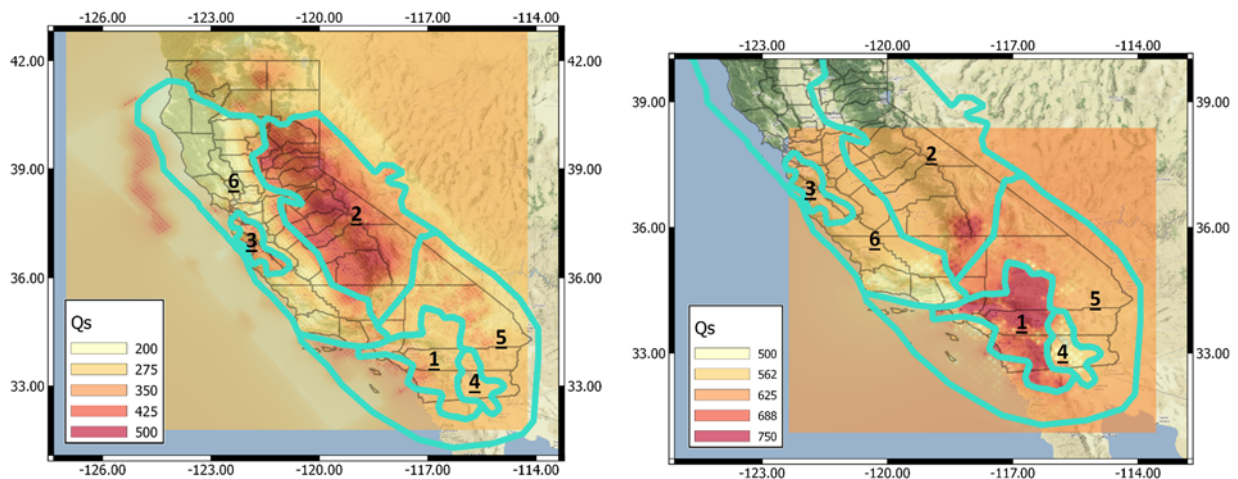
## Fitting of $\kappa$ from Ground Motions

We apply the  $\kappa$ -fitting procedures described in Afshari and Stewart (2019), which were adapted from Cabas et al. (2017) and Xu et al. (2020). The fit occurs over a range of frequencies from  $f_e$  to  $f_x$  (upper and lower bounds, respectively) that is selected for each record.

Search ranges for  $f_e$  and  $f_x$  are taken as 10-18Hz and 22-28Hz, respectively, each with 0.5Hz increments. For each possible combination of  $f_e$  and  $f_x$ ,  $\kappa$  is computed for combinations of the two horizontal components rotated to various azimuths. The variability of  $\kappa$  with azimuth is computed for each  $f_e$ - $f_x$  combination, which is expressed as a coefficient of variation (COV). We seek the combination of  $f_e$  and  $f_x$  that minimizes the azimuthal variability, and then take  $\kappa$  as the median. The Fourier amplitude spectra for two example sites shown in Figure 9 are for the azimuths and frequency ranges identified using this process.

## Analysis of Path- and Site Contributions to $\kappa$

Rates of crustal attenuation vary spatially due to variations in geologic conditions. Conditions producing relatively fast ground motion attenuation rates (i.e., low crustal quality factor,  $Q$ ) would be expected to increase  $\kappa_R$ . Insight into spatial variations of attenuation rates are provided by maps of frequency-independent  $Q$  (denoted  $Q_s$ ) by Eberhart-Phillips (2016) for Northern California and Hauksson and Shearer (2006) for Southern California. Figure 10 shows maps of California indicating variations of  $Q_s$  at a depth of 10 km from the two sources. There are systematic differences between  $Q_s$ , with southern California values being higher.



**Figure 10.** Spatial variation of frequency-independent quality factor ( $Q_s$ ) for California as derived from two models at a depth of 10 km. Zones considered in mixed-effects analysis of path  $\kappa$  are shown.

Allowing for differences in  $Q_s$  between the two sources, and considering both maps, we have assigned six zones of approximately uniform  $Q_s$ , with the intention of computing  $\kappa_R$  separately for each zone. If the value of  $\kappa_R$  for zone  $i$  is taken as  $\kappa_{R,i}$ , then the  $\kappa_R R$  term in Eq. (7) is computed as:



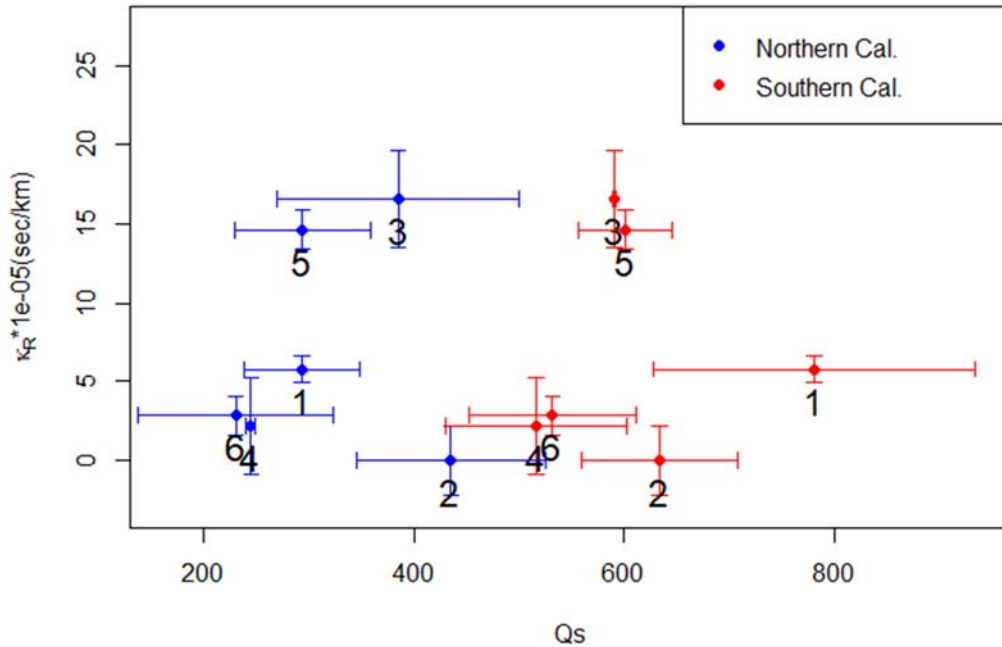
$$\kappa_R R = \sum_i \kappa_{R,i} R_i \quad [12]$$

where  $R_i$  is the path length (between source and site) through zone  $i$ . Distance  $R_i$  is zero if the path does not go through the zone  $i$ .

We use mixed-effects regression (more specifically, random intercept model) to obtain  $\kappa_{0,j}$  for each station  $j$  and  $\kappa_{R,i}$  for each zone  $i$ , by adapting Eq. (7) as follows:

$$\kappa_{k,j} = \kappa_{0,j} + \sum_{i=1}^n \kappa_{R,i} R_{k,i} \quad [13]$$

where  $\kappa_{k,j}$  is the measured  $\kappa$  from recording  $k$  at station  $j$ ,  $\kappa_{0,j}$  is the site-specific decay parameter at station  $j$ ,  $n$  is the number of zones in California, and  $R_{k,i}$  is the source-to-site path length for recording  $k$  that goes through zone  $i$ . Station terms  $\kappa_{0,j}$  are taken as random effects and path terms  $\kappa_{R,i}$  as fixed effects. Eq. (13) is solved using an equivalent matrix form. Regressions are performed in R [packages *nlme* (Pinheiro et al. 2019) or *lme4* (Bates et al., 2015)] using the full database (i.e., NGA-West2 stations and records in California as augmented here). The resulting  $\kappa_R$  values are shown for each zone in Figure 11, where they are plotted against the Qs values from the two references. The error bars shown in the figure indicate the estimation error for  $\kappa_R$  from the regressions and the within-zone ranges of Qs. Regressions provided negative  $\kappa_R$  in Zone 2, which is plotted instead at zero.



**Figure 11.** Variation of  $\kappa_R$  with average Qs within the six zones shown in Figure 10. Average Qs is taken from both Eberhart-Phillips (2016) for Northern California and Hauksson and Shearer (2006) for Southern California

## Base of Profile Site Decay Parameter, $\kappa_{0,b}$

The base of profile site decay parameter  $\kappa_{0,b}$  is needed to estimate the change in site kappa over the profile depth ( $\Delta\kappa$ ) using Eq. (10). Because the sites considered in this research are surface-only instruments,  $\kappa_{0,b}$  cannot be measured (i.e., from a downhole instrument) but instead is estimated from models. Several such models were considered.

Silva et al. (1998) used California data to relate  $\kappa_0$  to  $V_{S30}$ ,

$$\kappa_0 = \begin{cases} 0.008 \text{ sec,} & V_{S30} > 1500 \text{ m/s} \\ 0.020 \text{ sec,} & 760 < V_{S30} < 1500 \text{ m/s} \\ 0.030 \text{ sec,} & 360 < V_{S30} < 760 \text{ m/s} \end{cases} \quad [14]$$

Van Houtte et al. (2011) and Xu et al. (2000) used larger databases from the KiK-net array in Japan and NGA-West data to derive empirical relationships between  $\kappa_0$  and  $V_{S30}$ . Van Houtte et al. (2011) proposed:

$$\ln\kappa_0 = 3.490 - 1.062\ln V_{S30} \quad [15]$$

Xu et al. (2020) proposed:

$$\ln\kappa_0 = \begin{cases} k_1(\ln V_1)^2 + k_2\ln V_1 + k_3, & 100\text{m/s} < V_{S30} < V_1 \\ k_1(\ln V_{S30})^2 + k_2\ln V_{S30} + k_3, & V_1 < V_{S30} < V_2 \\ k_1(\ln V_2)^2 + k_2\ln V_2 + k_3, & V_2 < V_{S30} < 3000\text{m/s} \end{cases} \quad [16]$$

where  $k_1=0.18$ ,  $k_2=1.816$ ,  $k_3=-7.38$ ,  $V_1=155$  m/s, and  $V_2=2000$  m/s. The units of  $\kappa_0$  are sec in both Eq. (15) and (16).

We apply the Van Houtte et al. (2011) relationship in the present work. To obtain  $\kappa_{0,b}$ , we estimate the  $V_{S30}$  corresponding the base of the soil column by projecting vertically (constant velocity) the  $V_S$  at the deepest portion of the profile. We then enter this value into Eq. [15] to compute  $\kappa_{0,b}$ . Results of this process for the two example sites are shown in Table 2, as are derived values of  $\Delta\kappa$  and  $F_D$  using Eqs [10-11].

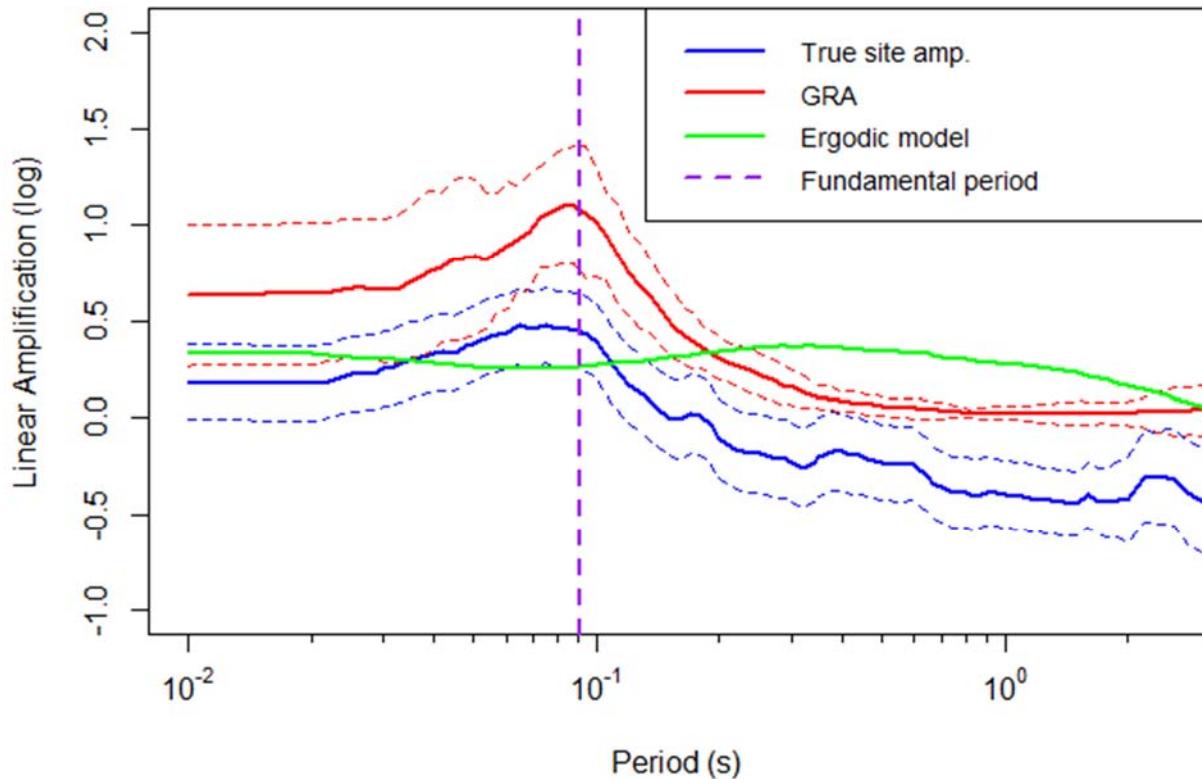
**Table 2.** Site kappa results for the AZ.KNW and AZ.PFO sites

Site	$\kappa_0$ (sec)	$\kappa_{0,b}$ (sec)	$\Delta\kappa$ (sec)	$F_D$
Keenwild Fire Station, Mountain Center, CA (AZ.KNW)	0.059	0.009	0.050	5
Pinyon Flats Observatory, CA (AZ.PFO)	0.050	0.006	0.044	5

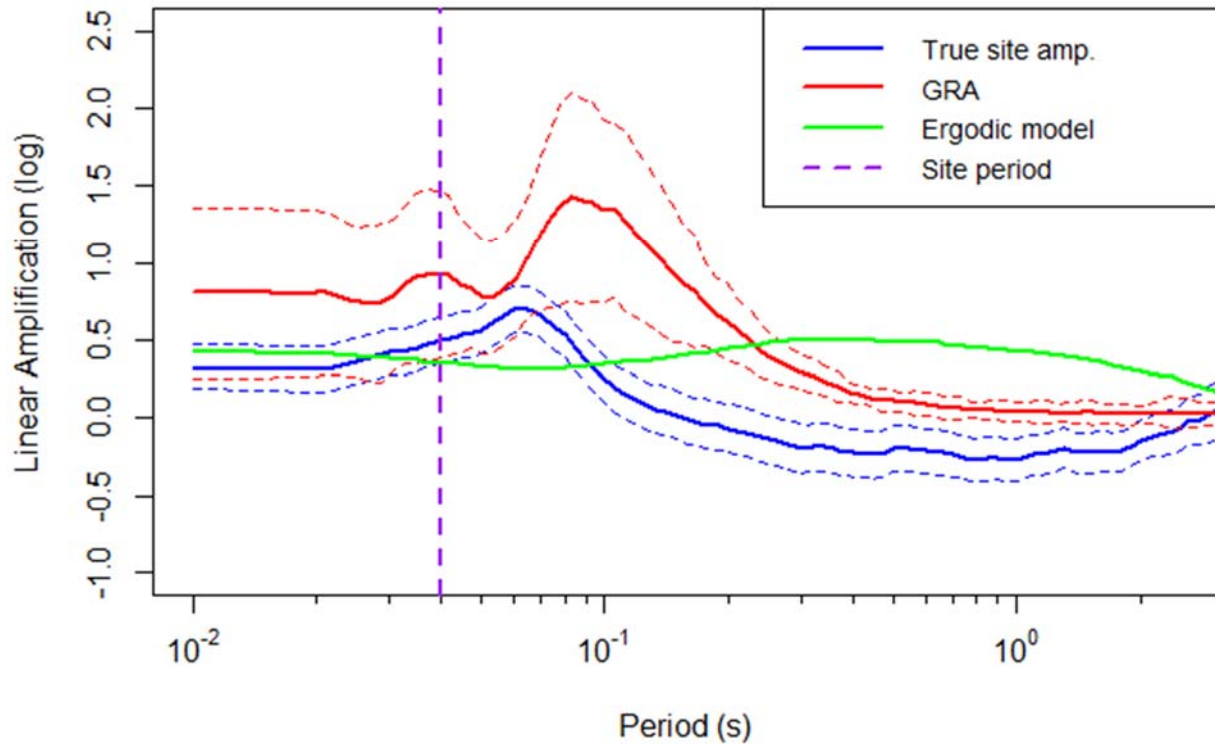
### Data-to-Model Comparisons for Example Sites

Figure 12 compares the data-derived site amplification (labelled as “true site amp.”) with estimates derived from ground response analyses (“GRA”) and from an ergodic model (Seyhan and Stewart 2014) for the AZ.KNW site. The ground response analysis results shown here use laboratory-based damping models. The site exhibits a peak in the amplification at about 0.09 sec. This occurs because AZ.KNW is a rock site with a shallow surficial soil layer. Ground response analyses are able to capture this feature and provide a good representation of the frequency-dependent shape of the site amplification. In contrast, the ergodic model significantly misfits the observed site amplification. Figure 13 shows similar results for the AZ.PFO site, where ground response analyses misfit the data, but still arguably improve upon the predictions of an ergodic model.

The effects of a  $\kappa$ -informed damping model are modest for the two example sites examined here. The impacts of alternate damping models is being explored by applying these procedures to additional sites with thick soil columns.



**Figure 12.** Non-ergodic site response at the AZ.KNW site, compared with site response predictions obtained with use of ground response analysis and an ergodic model. The maximum period used in the plots is the median of the maximum usable periods from data processing. The ground response model provides a good estimate of the shape of the amplification function in this case.



**Figure 13.** Non-ergodic site response at the AZ.PFO site, compared with site response predictions obtained with use of ground response analysis and an ergodic model. The ground response model provides a relatively poor estimate of the shape of the amplification function in this case.

## Conclusions

This research has the broad objective of investigating the effectiveness of ground response analysis, and other methods of site response analysis, through comparisons to true site response as established from analysis of recordings. Effectiveness is judged, in this context, through bias and site-to-site uncertainty of predicted site response.

We describe a new approach using non-ergodic site responses derived from surface-only instruments as the basis for validation studies. This paper has the main objective of describing the methodology, particularly with regard to procedures that implement ground response analyses given limited available information. We illustrate the approach using two sites with shallow soil layers overlying firm rock materials. In these cases, ground response analyses provide improved estimates relative to ergodic models. The work for this project is ongoing and full results will be presented in a later report.


## Acknowledgments

Funding for this study is provided by California Strong Motion Instrumentation Program, California Geological Survey, under Agreement Number 1016-985, as well as the US Geological Survey under contract number G17AP00018. This support is gratefully acknowledged. The work

presented here represents the views and opinions of the authors and does not reflect the policy, expressed or implied, of the US Government.

## References

- Afshari, K., Stewart, J. P. (2019) Insights from California Vertical Arrays on the Effectiveness of Ground Response Analysis with Alternative Damping Models. *Bull. Seismol. Soc. Am.* DOI: 10.1785/0120180292
- Ahdi, SK, S Sadiq, O Olhan, Y Bozorgnia, YMA Hashash, DY Kwak, D Park, A Yong, JP Stewart (2018). Development of a United States community shear wave velocity profile database, in Geotechnical Engineering and Soil Dynamics V: Seismic Hazard Analysis, Earthquake Ground Motions, and Regional-Scale Assessment, June 10-13, 2018, Austin, TX, ASCE Geotechnical Special Publication No. 291, S.J. Brandenberg and M.T. Manzari (eds.), pp 330-339.
- Ancheta, T.D., Darragh, R.B., Stewart, J.P., Seyhan, E., Silva, W.J., Chiou, B.S.-J., Wooddell, K.E., Graves, R.W., Kottke, A.R., Boore, D.M., Kishida, T., and Donahue, J.L., (2014). NGA-West2 database, *Earthquake Spectra*, **30**, 989-1005.
- Chapman, M. C., P. Talwani, and R. C. Cannon (2003). Ground-motion attenuation in the Atlantic Coastal Plain near Charleston, South Carolina, *Bull. Seism. Soc. Am.* **93**, 998–1011.
- Anderson, J. G., and S. E. Hough (1984). A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, *Bull. Seis-mol. Soc. Am.* **74**, 1969–1993.
- Anderson, J. G. (1991). A preliminary descriptive model for the distance dependence of the spectral decay parameter in southern California, *Bull. Seism. Soc. Am.* **81**, 2186–2193.
- Bates D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software.* 67, 1-48.
- Boore D.M. (2010). Orientation-independent, nongeometric-mean measures of seismic intensity from two horizontal components of motion, *Bull. Seismol. Soc. Am.*, **100**, 1830–1835.
- Boore, D. M., Stewart, J. P., Seyhan, E., and Atkinson, G. M, (2014). NGA-West 2 equations for predicting PGA, PGV, and 5%-damped PSA for shallow crustal earthquakes, *Earthquake Spectra*. **30**, 1057–1085.
- Cabas, A., Rodriguez-Marek, A. and Bonilla, L.F. (2017). Estimation of site-specific Kappa ( $\kappa_0$ )-consistent damping values at KiK-net sites to assess the discrepancy between laboratory-based damping models and observed attenuation (of seismic waves) in the field, *Bull. Seism. Soc. Am.* **107**, 2258-2271
- Campbell, K.W., (2009). Estimates of shear-wave Q and  $\kappa_0$  for unconsolidated and semiconsolidated sediments in Eastern North America. *Bull. Seismol. Soc. Am.*, **99**, 2365-2392.
- Chang, S.W. (1996). Seismic response of deep stiff soil deposits, *Ph.D. Thesis*, Univ. of California, Berkeley.
- Choi WK (2008) Dynamic properties of ash-flow tuffs. *Ph.D. thesis*, Department of Civil Engineering, University of Texas, Austin, Texas.
- Darendeli, M. B., (2001). Development of a new family of normalized modulus reduction and material damping curves, *PhD Thesis*, Department of Civil Engineering, University of Texas, Austin, TX.
- Dickenson, S.E. (1994). The dynamic response of soft and deep cohesive soils during the Loma Prieta earthquake of October 17, 1989, *Ph.D. Thesis*, Univ. of California, Berkeley.

- Eberhart-Phillips, D. (2016). Northern California seismic attenuation: 3D QP and QS models. *Bull. Seismol. Soc. Am.* **106**, 2558-2573.
- Fumal, TE and JC Tinsley (1985). Mapping shear wave velocities of near surface geologic materials, *Prof. Paper 1360*, US Geological Survey.
- Hassani, B., and G. M. Atkinson (2016). Site effects model for central and eastern North America based on peak frequency, *Bull. Seismol. Soc. Am.*, **106**, 2197–2213.
- Hauksson, E., and P. Shearer (2006). Attenuation models (QP and QS) in three dimensions of the southern California crust: Inferred fluidsaturation at seismogenic depths, *J. Geophys. Res.* **111** (B05302).
- Hough, S. E., and J. G. Anderson (1988). High-frequency spectra observed at Anza, California: Implications of Q structure. *Bull. Seismol. Soc. Am.* **78**, 692–707.
- Idriss, I. M., 2011. Use of Vs30 to represent local site conditions, in *4th LASPEI/IAEE International Symposium Effects of Surface Geology on Strong Ground Motions*, Santa Barbara, CA.
- Idriss, I.M. (1993). Assessment of site response analysis procedures, *Report to National Institute of Standards and Technology*, Gaithersburg, Maryland, Dept. of Civil & Environmental Eng., Univ. of California, Davis.
- Kaklamanos, J, BA Bradley, EM Thompson, LG Baise. (2013). Critical parameters affecting bias and variability in site-response analyses using KiK-net downhole array data, *Bull. Seismol. Soc. Am.*, **103**, 1733–1749.
- Kaklamanos, J, B Bradley, 2018. Challenges in predicting seismic site response with 1D analyses: Conclusions from 114 KiK-net vertical seismometer arrays, *Bull. Seismol. Soc. Am.*, **108**, 2816-2838.
- Kim, B., Y. M. A. Hashash, J. P. Stewart, E. M. Rathje, J. A. Harmon, M. I. Musgrove, K. W. Campbell, and W. J. Silva (2016). Relative differences between nonlinear and equivalent-linear 1D site response analyses. *Earthq. Spectra*, **32**, 1845–1865.
- Kwak, DY, JP Stewart, S.-J. Mandokhai, D Park (2017). Supplementing VS30 with H/V spectral ratios for predicting site effects, *Bull. Seism. Soc. Am.*, **107**, 2028-2042.
- Kwok, AOL, JP Stewart, DY Kwak, P-L Sun (2018). Taiwan-specific model for VS30 prediction considering between-proxy correlations, *Earthquake Spectra*, **34**, 1973-1993.
- Menq, F., (2003). Dynamic properties of sandy and gravelly soils, *PhD Thesis*, Department of Civil Engineering, University of Texas, Austin, TX.
- Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team (2019). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-140.
- Rodriguez-Marek, A., Rathje, E. M., Bommer, J. J., Scherbaum, F., and Stafford, P. J.. (2014). Application of single-station sigma and site response characterization in a probabilistic seismic hazard analysis for a new nuclear site, *Bull. Seismol. Soc. Am.*, **104**, 1601–1619.
- Rathje EM, Dawson C, Padgett JE, Pinelli JP, Stanzione D, Adair A, Arduino P, Brandenburg SJ, Cockerill T, Dey C, Esteva M. DesignSafe: new cyberinfrastructure for natural hazards engineering. *Natural Hazards Review*. 2017 Feb 21;**18**(3):06017001.
- Rogers, A. M., J. C. Tinsley, and R. D. Borchardt (1985). Predicting relative ground response, *U.S. Geol. Surv. Profess. Paper 1360*, 221-248. 

- Schnabel PB (1973) Effects of local geology and distance from source on earthquake ground motions. *Ph.D. thesis*, University of California, Berkeley.
- Seyhan, E, JP Stewart, TD Ancheta, RB Darragh, and RW Graves (2014). NGA-West 2 site database. *Earthquake Spectra*, **30**, 1007-1024.
- Silva, W., R. B. Darragh, N. Gregor, G. Martin, N. Abrahamson, and C. Kircher (1998). Reassessment of site coefficients and near-fault factors for building code provisions, Technical Report Program Element II: 98-HQ-GR-1010, Pacific Engineering and Analysis., El Cerrito, California USA.
- Stafford, J.S., Rodriguez-Marek, A., Edwards, B., Kruiver, P.P., and Bommer, J.J., (2017). Scenario dependence of linear site-effect factors for short-period response spectral ordinates. *Bull. Seismol. Soc. Am.*, **107**, 2859–2872.
- Stewart, JP, K Afshari, CA Goulet (2017). Non-ergodic site response in seismic hazard analysis, *Earthquake Spectra*, **33**, 1385-1414.
- Thompson E. M., D. J. Wald, C. B. Worden. (2014) A Vs30 Map for California with Geologic and Topographic Constraints. *Bull. Seismol. Soc. Am.*, **104**, 2313–2321.
- Thompson, E.M., (2018), An Updated Vs30 Map for California with Geologic and Topographic Constraints: U.S. Geological Survey data release. DOI: 10.5066/F7JQ108S.
- Van Houtte, C., Drouet, C., Cotton, F. (2011). Analysis of the origins of  $\kappa$  (kappa) to compute hard rock to rock adjustment factors for GMPEs. *Bull. Seismol. Soc. Am.*, **101**, 2926–2941.
- Wald, D. J., and T. I. Allen (2007). Topographic slope as a proxy for seismic site conditions and amplification, *Bull. Seismol. Soc. Am.*, **97**, 1379–1395.
- Wang, P., Stewart, J.P., Bozorgnia, Y., Boore, D.M., and Kishida, T., (2017). “R” Package for computation of earthquake ground-motion response spectra. *PEER Report No. 2017/09*, Pacific Earthquake Engineering Research Center, Berkeley, CA.
- Wills, C. J., and K. B. Clahan (2006). Developing a map of geologically defined site-condition categories for California, *Bull. Seismol. Soc. Am.*, **96**, 1483–1501.
- Wills, C. J., C. I. Gutierrez, F. G. Perez, D. M. Branum. (2015). A Next Generation VS30 Map for California Based on Geology and Topography. *Bull. Seismol. Soc. Am.*, **105**, 3083–3091.
- Xu, B., EM Rathje, YMA Hashash, JP Stewart, KW Campbell, and WJ Silva (2020).  $\kappa_0$  for soil sites: Observations from Kik-net sites and their use in constraining small-strain damping profiles for site response analysis, *Earthquake Spectra*, <https://doi.org/10.1177/8755293019878188>.
- Yong, A., A. Martin, K. Stokoe, and J. Diehl (2013). ARRA-funded V S30 measurements using multi-technique approach at strong-motion stations in California and central-eastern United States, *U.S. Geol. Surv. Open-File Rept. 2013-1102*, 65 pp.
- Yong, A., S. E. Hough, J. Iwahashi, and A. Braverman. (2012). A terrain- based site-conditions map of California with implications for the contiguous United States, *Bull. Seismol. Soc. Am.* **102**, 114–128.
- Yong, A., (2016). Comparison of measured and proxy-based  $V_{S30}$  values in California, *Earthquake Spectra*. **32**, 171-192.
- Zalachoris, G, EM Rathje, 2015. Evaluation of one-dimensional site response techniques using borehole arrays, *J. Geotech. Geoenviron. Eng.*, **141**, 04015053.