

Multiscale Site-Response Mapping: A Case Study of Parkfield, California

by Eric M. Thompson, Laurie G. Baise, Robert E. Kayen,
Eugene C. Morgan, and James Kaklamanos

Abstract The scale of previously proposed methods for mapping site-response ranges from global coverage down to individual urban regions. Typically, spatial coverage and accuracy are inversely related. We use the densely spaced strong-motion stations in Parkfield, California, to estimate the accuracy of different site-response mapping methods and demonstrate a method for integrating multiple site-response estimates from the site to the global scale. This method is simply a weighted mean of a suite of different estimates, where the weights are the inverse of the variance of the individual estimates. Thus, the dominant site-response model varies in space as a function of the accuracy of the different models. For mapping applications, site-response models should be judged in terms of both spatial coverage and the degree of correlation with observed amplifications. Performance varies with period, but in general the Parkfield data show that: (1) where a velocity profile is available, the square-root-of-impedance (SRI) method outperforms the measured V_{S30} (30 m divided by the S -wave travel time to 30 m depth) and (2) where velocity profiles are unavailable, the topographic slope method outperforms surficial geology for short periods, but geology outperforms slope at longer periods. We develop new equations to estimate site response from topographic slope, derived from the Next Generation Attenuation (NGA) database.

Introduction

Earthquake hazard maps, such as the U.S. national seismic hazard maps (Frankel *et al.*, 1996, 2002; Petersen *et al.*, 2008) and rapid response maps (Wald *et al.*, 2005, 2006) are fundamentally linked to ground-motion prediction equations (GMPEs). In general, earthquake hazard maps use at least one GMPE to compute the intensity of ground shaking throughout the region of interest. Ground-motion prediction equations are generally a function of three types of explanatory variables: source, path, and site. In this article, we study the site term by comparing recorded ground motions with those predicted by previously developed GMPEs. We present (1) a new method for integrating multiple estimates of site response into a period-dependent map of site-response amplifications and (2) new equations for incorporating site effects into existing GMPEs using topographic slope, which we develop specifically for the purpose of incorporating site effects into earthquake hazard maps. Not only do we wish to extract as much information from the subsurface information as possible, but we also wish to infer as much information as possible in locations without subsurface data.

Douglas (2003) thoroughly described the history of GMPE development, including the various methods of incorporating site effects. Site effects have been most commonly included into GMPEs with multiplicative factors based on site

category (e.g., Trifunac, 1976). Some GMPE developers have also used S -wave velocity (V_S) profiles to eliminate the subjectivity of assigning site categories (e.g., Joyner and Fumal, 1984). Many of the most recently developed GMPEs use V_{S30} (30 m divided by the travel time of an S wave to 30 m depth) to model the site term (Chiou *et al.*, 2008). Three of the five recently developed Next Generation Attenuation (NGA) GMPEs use multiple explanatory variables: V_{S30} plus a soil depth parameter. Abrahamson and Silva (2008) and Chiou and Youngs (2008) define the soil depth parameter as the minimum depth at which $V_S \geq 1.0$ km/s ($Z_{1.0}$); Campbell and Bozorgnia (2008) define the soil depth parameter as the minimum depth at which $V_S \geq 2.5$ km/s ($Z_{2.5}$).

Despite the popularity of V_{S30} in GMPEs, the lack of V_{S30} measurements poses a major obstacle to incorporating site effects into earthquake hazard maps. Thus, researchers have developed correlations of V_{S30} with quantities that have greater spatial coverage, such as surficial geology (Wills and Clahan, 2006) and topographic slope (Wald and Allen, 2007). Other researchers have focused on the limitations of the ability of V_{S30} to model site response (Castellaro *et al.*, 2008). These limitations can be overcome by finding a more accurate site-response explanatory variable or including additional site-response variables into GMPEs to improve prediction accuracy.

The accuracy with which any variable can predict site response is a combination of how well that variable correlates with site response (i.e., model uncertainty) and the accuracy with which that variable can be estimated at a specific location (i.e., estimation uncertainty). For example, an estimate of V_{S30} from downhole logging will presumably correlate better with site response than qualitative categories of soil and rock because the V_{S30} model exhibits less model uncertainty. In contrast, the downhole estimate of V_{S30} will exhibit a stronger correlation with site response than the V_{S30} inferred from surficial geology or topographic slope because the latter estimates introduce additional estimation uncertainty (e.g., Moss, 2008). In this article, we develop a method for mapping site response that accounts for these two types of uncertainties.

To increase spatial coverage, one must generally sacrifice accuracy. For example, Wald and Allen (2007) developed a method for predicting site response from globally available topography data. Thus, the site-response explanatory variable (topographic slope) is known with uniform precision at every strong-motion station as well as any point on a map where an estimate of ground-shaking intensity is desired. The ubiquity of this explanatory variable, compared with direct measurements of V_{S30} , presumably can only be achieved by sacrificing some predictive accuracy.

Our first task is to assess the performance of alternative site-response explanatory variables, including V_{S30} , the dominant site frequency, topographic slope, surficial geology, the square-root-of-impedance (SRI) amplification, and the 1D plane SH -wave amplification. We use this information to demonstrate a framework for integrating alternative site-response explanatory variables into a unified, period-dependent site-response map. We judge the different variables on two different criteria: (1) the ability to explain GMPE site-response residuals and (2) data availability, both in terms of the number of strong-motion stations where the data are available and in terms of geographic coverage for creating earthquake hazard maps. This quantitative assessment of different site-response terms will aid future decisions regarding data collection efforts and GMPE development. Though the primary application of our results is for including site effects into hazard maps, the results are also useful for GMPE development because V_S profiles have not been measured at a large percentage of strong-motion stations used for GMPE regression. The same methods that we propose for including site effects into hazard maps can also be used to estimate the site effects at strong-motion stations.

The quantity and spatial density of the 52 spectral analysis of surface wave (SASW) V_S profiles collected by Thompson, Kayen, *et al.* (2010) at strong-motion stations in Parkfield, California, provide a unique opportunity to study site-response mapping. Although densely sampled V_S measurements have been collected in other regions (e.g., Holzer *et al.*, 2002), the Parkfield dataset is unique because of the number of locations that contain both a V_S profile and a strong-motion station within a relatively small region.

The 1983 M 6.36 Coalinga and the 2004 M 6.0 Parkfield earthquakes were recorded at most of these stations.

Our second task is to present new equations for incorporating site effects into GMPEs for use in seismic hazard maps. We conclude that topographic slope is appropriate for incorporating site response into hazard maps because (1) topographic slope provides global coverage and (2) topographic slope is able to predict site amplifications at least as accurately as surficial geology. We build upon the work of Wald and Allen (2007) by regressing topographic slope directly against site-response amplifications. For these regressions, we expand our sample size beyond the Parkfield region to include the NGA flatfile. This ensures that the site-response equations that we propose are valid under the same conditions as the NGA GMPEs.

Ground Motions and Site Characterization

Figure 1 shows the 52 strong-motion stations that we analyze in this study and the epicenters of the 1983 M 6.36 Coalinga and the 2004 M 6.0 Parkfield earthquakes. Each of the sites in Figure 1 is a strong-motion station where Thompson, Kayen, *et al.* (2010) also measured the V_S profile (see the Data and Resources section). All of the stations recorded the Parkfield earthquake except for fault zone

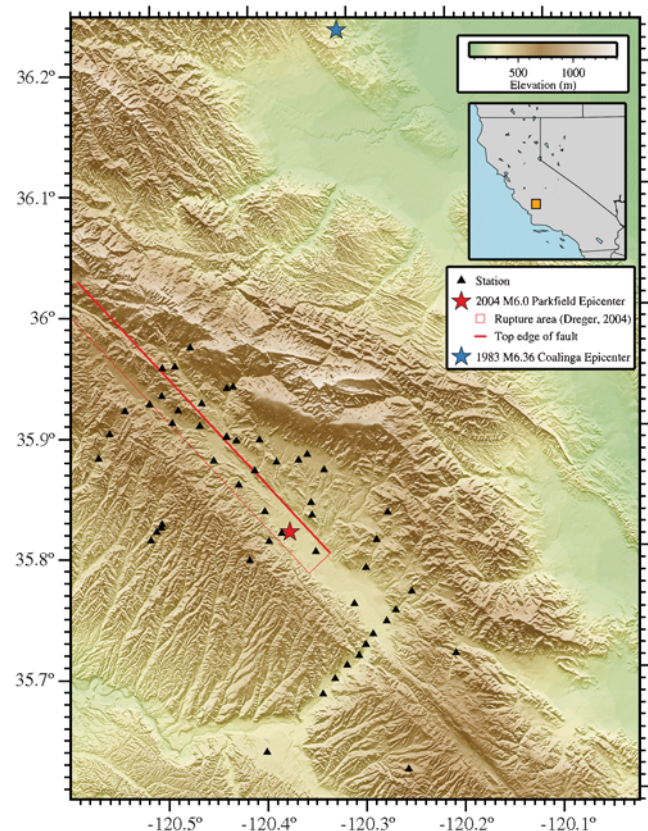


Figure 1. Vicinity map showing the location of the 52 strong-motion stations with SASW measurements, the epicenter, the finite fault rupture area of the 2004 Parkfield earthquake (Dreger, 2004), and the epicenter of the 1983 Coalinga earthquake.

16. The maximum acceleration at fault zone 16 exceeded the recording capacity of the station, so we cannot use it for this analysis. See [Shakal et al. \(2006\)](#) for a detailed analysis of the anomalously large ground motion at fault zone 16. A subset of 36 of these stations also recorded the 1983 **M** 6.36 Coalinga earthquake.

Table 1 gives station summary information for the strong-motion stations, including the latitude and longitude of the SASW survey, maximum depth (D_{\max}) of the SASW V_S profile, the period of the quarter-wavelength corresponding to D_{\max} for the SASW V_S profile (T_{\max}^{λ}), and the SASW V_{S30} . Table 1 also gives event-specific path information: the Joyner–Boore distance (R_{JB}) and the source-to-site azimuth (α) for the Coalinga and Parkfield earthquakes. The dominant site frequency (f_0 ; defined in the [Methods](#) section) is also provided in Table 1. Stations with missing values of R_{JB} and α do not have observed ground-motion records for that event. The source parameters for the Coalinga event are **M** 6.36 and rake angle of 90° (reverse faulting); the source parameters for the Parkfield event are **M** 6.0 and rake = 180° (right-lateral strike-slip faulting). The largest usable period (T_{\max}) is a function of the corner frequency of the high-pass filter (f_{\min}) applied to the Coalinga and Parkfield ground-motion records. For each record, we only use frequencies greater than the f_{\min} reported in the NGA flatfile.

The [Wills and Clahan \(2006\)](#) site characterization method requires a surficial geology map at the 1:250,000 scale. To apply the [Wills and Clahan \(2006\)](#) method, we digitized the 1:250,000 scale [Jennings \(1958\)](#) map of surficial geology that is available from the California Geological Survey. The [Wald and Allen \(2007\)](#) site characterization method requires Shuttle Radar Topography Mission (SRTM) 30 s global topography data (see the [Data and Resources](#) section).

Methods

Response Variable

Isolating site effects from the other processes that influence ground-shaking intensity is a fundamental challenge for site-response studies ([Steidl et al., 1996](#); [Boore and Joyner, 1997](#); [Abercrombie, 1997](#)). The most common approach, employed by [Borcherdt \(1970\)](#), compares the response at two stations for which the source and path effects can be assumed to be identical. Thus, any differences must be due to the local site conditions (i.e., site response). This was termed the reference site approach by [Field and Jacob \(1995\)](#). In terms of the pseudospectral acceleration (PSA), which is a function of the natural period (T) of a single-degree-of-freedom system (SDOF), the site-response amplification is

$$\ln[a(T)] = \ln[\text{PSA}(T)] - \ln[\text{PSA}^{\text{ref}}(T)], \quad (1)$$

where $\text{PSA}^{\text{ref}}(T)$ is the PSA at the reference site and $\text{PSA}(T)$ is the observed PSA at the location of interest (all of the PSA values in this article assume 5% viscous damp-

ing). Generally, one chooses a reference site with minimal site effects, termed a rock site. For example, [Borcherdt \(1994\)](#) defined site response as the ratio of the spectra of ground motions recorded on soil to the spectra of ground motions recorded on rock. A drawback to this approach is that the choice of the reference site can substantially influence the estimate of $a(T)$ because $\text{PSA}^{\text{ref}}(T)$ may not be known with certainty.

We wish to avoid the somewhat arbitrary choice of reference site because (1) there is no clear rock site that is appropriate for many of the Parkfield strong-motion stations and (2) a single reference site can have its own response that must be accounted for. In order to assess site-response explanatory variables within the GMPE framework, we define PSA^{ref} as the mean PSA computed from existing GMPEs for a reference rock condition. This approach was termed a nonreference site approach by [Field and Jacob \(1995\)](#) and assumes that the GMPEs accurately account for the path and source effects. We are comfortable with this assumption for this application because it is unlikely that the unaccounted-for source and path variability will create spurious correlations with any of the site-response explanatory variables that we consider in this article. Uncorrelated variability may reduce the degree of correlation between the site-response explanatory variables and the site-response amplification values, but will not favor one variable over another.

To calculate $\text{PSA}^{\text{ref}}(T)$, we assume a reference rock site of $V_{S30} = 760$ m/s and $Z_{1.0} = 50$ m. We define $\text{PSA}^{\text{ref}}(T)$ as the mean reference rock PSA of four NGA GMPEs after correcting for the event term: [Abrahamson and Silva \(2008\)](#), [Boore and Atkinson \(2008\)](#), [Campbell and Bozorgnia \(2008\)](#), and [Chiou and Youngs \(2008\)](#), labeled AS08, BA08, CB08, and CY08, respectively. The event term is the median residual for each GMPE for an individual earthquake. To determine the source and distance GMPE input parameters for the Coalinga earthquake, we use the information in the NGA flatfile for 35 of the 36 stations that recorded this earthquake. One station (853SC; Stone Corral 1E) is not available in the NGA flatfile, so we estimate the distance measures by linear extrapolation from stations Stone Corral 2E (852SC) and Stone Corral 3E (851SC), which are approximately evenly spaced in a line. The Parkfield earthquake is not included in the NGA database, and so we estimate the GMPE input parameters from the [Dreger \(2004\)](#) finite fault model (displayed in Fig. 1). Additional input parameters for the stations (such as the site coordinate, R_X) are computed using the procedure of [Kaklamanos et al. \(2011\)](#); we also used this methodology to determine unknown GMPE input parameters for the Parkfield earthquake.

Site-Response Explanatory Variables

The site-response explanatory variables that we analyze in this article can be categorized in a number of useful ways. We consider both vector and raster spatial data, as well as period-dependent and period-independent data. The only

Table 1
Summary of Station Information*

ID	Latitude (°)	Longitude (°)	D_{\max} (m)	T_{\max}^{λ} (s)	V_{S30} (m/s)	R_{JB} (km)		f_0 (Hz)	α (°)	
						Coalinga	Parkfield		Coalinga	Parkfield
807PAR	35.8985	-120.4329	120	1.00	261	27.96	0.29	1.49	73.55	-90.00
808PAR	35.8986	-120.4325	90	0.83	270	—	0.36	—	—	-90.00
809CHO	35.7328	-120.2893	110	1.16	173	43.83	8.24	3.34	49.68	13.26
810CHO	35.7430	-120.2753	40	0.65	233	42.76	7.95	0.74	48.48	4.30
811TUR	35.8779	-120.3599	70	0.31	907	—	3.83	1.43	—	-90.00
812CHO	35.8180	-120.3791	170	1.67	214	35.04	0.00	1.01	61.71	90.00
813CHO	35.7072	-120.3163	70	0.69	283	46.73	9.19	2.65	51.81	34.36
814VIN	35.9220	-120.5348	80	0.81	309	30.91	2.15	0.98	86.62	90.00
815CHO	35.6963	-120.3288	100	0.98	237	47.88	9.93	2.31	52.74	40.45
816CHO	35.7521	-120.2657	80	0.49	523	41.99	8.45	1.64	46.67	-8.24
817EAD	35.8952	-120.4229	130	0.89	384	—	0.74	0.62	—	-90.00
818PAR	35.9266	-120.4570	80	0.49	384	26.20	—	2.94	78.99	—
819VIN	35.9731	-120.4671	70	0.50	468	22.66	3.44	3.64	85.79	-90.00
820CHO	35.8701	-120.4051	70	0.79	297	29.91	0.20	0.77	67.84	-90.00
821TUR	35.8767	-120.3826	70	0.54	309	28.58	2.27	2.31	65.42	-90.00
822PAR	35.9080	-120.4595	50	0.73	246	28.11	0.00	0.98	77.31	90.00
823PAR	35.9211	-120.4813	120	0.93	308	28.00	0.00	0.81	81.19	90.00
824VCY	35.9399	-120.4247	120	0.50	657	—	3.83	2.50	—	-90.00
825VCN	35.9573	-120.4831	100	0.72	381	24.83	1.11	1.01	85.32	-90.00
826FZ	35.8955	-120.3989	120	0.59	542	27.10	2.44	2.67	68.92	-90.00
828VIN	35.8829	-120.5629	80	0.80	320	—	6.97	0.25	—	90.00
829VIN	35.9033	-120.5513	100	0.66	386	33.28	4.48	0.87	86.28	90.00
830CHO	35.7670	-120.2488	75	0.56	397	40.01	8.24	1.00	45.50	-24.52
831VIN	35.9336	-120.4974	90	0.85	284	27.72	0.00	0.94	84.31	90.00
832UPS	35.8214	-120.5059	80	0.73	358	—	7.76	1.33	—	90.00
833UPS	35.8248	-120.5011	80	0.68	417	—	7.30	1.25	—	90.00
834FZ	35.8791	-120.4461	150	1.05	372	30.43	0.00	0.69	73.26	90.00
835CHO	35.6842	-120.3418	100	0.94	252	49.40	11.16	2.13	53.76	46.17
836TUR	35.8824	-120.3510	100	0.52	467	—	4.66	5.00	—	-90.00
837CHO	35.7151	-120.3041	70	0.64	410	45.49	8.56	2.58	50.84	26.88
838CHO	35.6369	-120.3998	100	0.89	359	55.05	16.78	2.67	57.93	61.93
839CHO	35.7239	-120.2970	70	0.91	231	44.82	8.53	0.79	49.95	20.28
840RFU	35.6199	-120.2570	80	0.99	239	—	19.88	5.88	—	21.84
841KFU	35.7153	-120.2056	85	0.37	576	—	15.11	0.59	—	-8.46
842FFU	35.9108	-120.4873	70	0.75	227	—	0.00	5.88	—	90.00
843MFU	35.9564	-120.4959	90	0.56	398	—	0.37	2.63	—	-90.00
844JFU	35.9389	-120.4309	75	0.64	379	—	3.36	0.67	—	-90.00
845VIN	35.9268	-120.5100	80	0.49	439	29.01	0.09	2.94	84.76	90.00
846UPS	35.8277	-120.5001	75	0.73	342	—	6.81	0.91	—	90.00
847WFO	35.8145	-120.5118	90	0.64	447	—	8.51	0.62	—	90.00
848FZ	35.8586	-120.4214	80	0.87	267	31.64	0.00	0.21	68.90	90.00
849FZ	35.8360	-120.3958	100	1.49	221	33.42	0.00	0.37	64.58	90.00
850FZ	35.8019	-120.3450	70	1.01	212	36.14	0.00	2.67	56.06	90.00
851SC	35.8331	-120.2713	100	0.44	565	32.81	6.54	8.70	48.96	-82.30
852SC	35.8105	-120.2831	100	0.52	566	35.29	4.49	4.65	50.07	-62.49
853SC	35.7878	-120.2948	80	0.91	261	37.77	3.54	0.98	51.18	-25.97
854GH	35.7958	-120.4120	60	0.42	511	38.10	3.10	4.56	64.23	90.00
855GH	35.8113	-120.3922	40	0.50	290	35.93	0.57	0.26	62.88	90.00
856FZ	35.7580	-120.3075	60	1.06	178	41.04	4.91	1.79	51.88	17.34
857GH	35.8695	-120.3346	100	0.53	451	28.72	4.86	2.83	58.71	-90.00
858GH	35.8428	-120.3485	100	0.73	361	31.85	1.93	2.51	59.29	-90.00
859GFU	35.8324	-120.3477	50	0.27	558	—	1.31	5.88	—	-90.00

*Summary includes the latitude and longitude of the SASW survey, maximum depth (D_{\max}) of the SASW V_S profile, the period of the quarter-wavelength corresponding to D_{\max} for the SASW V_S profile (T_{\max}^{λ}), the SASW V_{S30} , and R_{JB} , f_0 , and α for the Coalinga and Parkfield earthquakes.

type of raster data that we use is topographic slope, and we use two types of vector data: point data (V_S profiles and strong-motion data) and polygon data (surficial geology). These data (with the exception of the strong-motion data)

can be easily processed to provide an estimate of V_{S30} , which is currently the most commonly used site-response explanatory variable. Because we discuss many different estimates of V_{S30} , it is convenient to use superscripts to specify the

different estimation techniques. For example, we refer to the estimate of V_{S30} from an SASW measurement as V_{S30}^{SASW} . We choose not to include the soil depth parameters in our analysis because only eight of the V_S profiles measured $Z_{1,0}$ and none measured $Z_{2,5}$. These explanatory variables were recently studied by [Kaklamanos and Baise \(2011\)](#). They found a high degree of uncertainty in current methods of estimating depth parameters at unsampled locations; they also found that the more complicated NGA GMPEs often do not have a predictive advantage over the simpler NGA models with fewer input parameters.

We compute the magnitude of the topographic slope (δ) with the method described by [Wald and Allen \(2007\)](#). The slope is estimated from SRTM 30 s global topography ([Farr and Kobrick, 2000](#)), so the pixel size is approximately 1 km by 1 km. From δ , we estimate V_{S30} (V_{S30}^δ) with the [Allen and Wald \(2009\)](#) empirical relationship for active tectonic regions. To estimate continuous values of V_{S30} , we linearly interpolate the V_{S30} that define the tabulated classes, as suggested by Allen and Wald (personal commun., 2010).

We refer to the estimate of V_{S30} predicted by the geologic classification method of [Wills and Clahan \(2006\)](#) as V_{S30}^{geo} . All but one of the 36 strong-motion stations that recorded the Coalinga earthquake are included in the NGA flatfile. The one station that is not in the flatfile is Stone Corral 1E (853SC in Table 1), which [Jennings \(1958\)](#) classified as Quaternary nonmarine terrace deposits (Qt). The [Wills and Clahan \(2006\)](#) classification that we assign to this station is Quaternary (Pleistocene) alluvium (Qoa). For all other locations considered in this study, we convert the [Jennings \(1958\)](#) geologic units to the [Wills and Clahan \(2006\)](#) classifications following Table 2.

[Lermo and Chávez-García \(1993\)](#) provided evidence that the horizontal-to-vertical spectral ratios (H/V) may provide valuable site-response information. This approach has become very popular due to the ease with which an estimate of the site conditions can be achieved. We use the procedure described by [Zhao et al. \(2006\)](#) to estimate the dominant site frequency (f_0) from H/V at the 36 stations in this article that recorded the 1983 Coalinga earthquake. Ideally, we would use a large number of earthquakes to achieve a more robust estimate of f_0 , but the vast majority of these stations only have records available for the Coalinga and Parkfield earthquakes. Horizontal-to-vertical spectral ratios are typically estimated from the Fourier spectrum (e.g., [Cadet et al., 2010](#)), but the [Zhao et al. \(2006\)](#) method uses the 5% damped PSA instead. The main advantage of using PSA over the Fourier spectrum is that no additional smoothing is required. In the numerator of the H/V PSA ratio, we use the geometric mean of the PSA of the horizontal components. For stations where multiple ground motions are available, we report f_0 as the geometric mean of the f_0 obtained from the different events.

The SASW V_S profiles include information that should be pertinent for site response besides the single value of V_{S30} . Thus, we investigate alternative data reduction methods to

Table 2
Relationship between the [Jennings \(1958\)](#) Geologic Units and the [Wills and Clahan \(2006\)](#) Geology Classifications

Jennings (1958)		Wills and Clahan (2006)
Symbol	Description	Symbol
Qal	Alluvium	Qal, thin
Qf	Fan deposits	Qal, thin
Qt	Quaternary nonmarine terrace deposits	Qoa
QP	Plio-Pleistocene nonmarine	QT
Pc	Undivided Pliocene nonmarine	QT
Pu	Upper Pliocene marine	Tsh
Pml	Middle and/or lower Pliocene marine	Tsh
Mc	Undivided Miocene nonmarine	Tsh
Mu	Upper Miocene marine	Tsh
Mv(r)	Miocene volcanic–rhyolite	Tv
Mm	Middle Miocene marine	Tsh
MI	Lower Miocene marine	Tsh
Oc	Oligocene nonmarine	Tss
E	Eocene marine	Tss
K	Undivided Cretaceous marine	Kss
Ku	Upper Cretaceous marine	Kss
KI	Lower Cretaceous marine	Kss
KJfv	Franciscan volcanic and metavolcanic rocks	KJf
KJf	Franciscan Formation	KJf
gr	Mesozoic granite rocks	xtaline
bi	Mesozoic basic intrusive rocks	xtaline
ub	Mesozoic ultrabasic intrusive rocks	xtaline
ls	Pre-Cretaceous metamorphic rocks–limestone or dolomite	xtaline

simplify the information in the V_S profile to a single value that can be used as an explanatory variable in regression analysis. This includes the SRI method and the 1D plane SH -wave computation. Unlike V_{S30} and f_0 , both of these alternative explanatory variables depend on period.

The ssRI method of estimating the site-response amplifications computes the amplification that a seismic wave will experience as it travels between two materials with different impedances, where the impedance is defined as the product of density and velocity: $\rho \times V$. The key assumption of this method is that the seismic energy is completely transmitted across the material interface. The SRI method ignores the effects of seismic reflections and refractions that will occur within a stack of laterally constant layers. [Joyner et al. \(1981\)](#) found that the impedance effect, as opposed to the effects of reflections and refractions, is the dominant process that contributes to site amplification. The method for estimating the period-dependent amplification with the SRI method [$a^{\text{SRI}}(T)$] has been extensively described in the relevant literature (e.g., [Joyner et al., 1981](#); [Boore, 2003](#); [Douglas et al., 2009](#); [Thompson, Baise et al., 2010](#)), so we do not reproduce it here. However, we note that the period of the amplification is related to the velocity profile through the quarter-wavelength approximation. We report maximum period (T_{max}^λ) for the SRI computation in Table 1. For example,

at site 807PAR, the D_{\max} of the V_S profile is 120 m, which results in $T_{\max}^{\lambda} = 1.0$ s. For periods greater than T_{\max}^{λ} , we use the amplification predicted at T_{\max}^{λ} . We hypothesize that $a^{\text{SRI}}(T_{\max}^{\lambda})$ will correlate better to the longer periods than V_{S30} because the degree of extrapolation is decreased. The periods of the quarter-wavelength for 30 m depth are $0.13 \leq T \leq 0.69$ s for these V_S profiles.

Accounting for the effect of reflections and refractions could potentially provide a more realistic model of $a(T)$ than $a^{\text{SRI}}(T)$. We accomplish this by assuming vertically incident plane SH -wave (the horizontally polarized component of the S wave) propagation through a laterally homogeneous medium (SH1D). The resulting amplifications $[a^{\text{SH1D}}(T)]$ are computed using the Thomson–Haskell technique (Thomson 1950; Haskell, 1953).

Both $a^{\text{SRI}}(T)$ and $a^{\text{SH1D}}(T)$ require a reference rock profile. Table 3 summarizes the reference rock profile that we use in this article. This model was selected such that $V_{S30} \approx 760$ m/s and $Z_{1.0} = 50$ m to be consistent with the reference condition of $a(T)$ in equation (1).

Assessing Explanatory Variables

One of the goals of this article is to assess the degree of correlation between the different explanatory variables and $\ln[a(T)]$. Pearson’s product moment correlation coefficient (r) is the most widely used measure of the linear association between two variables (Draper and Smith, 1981).

The hypothesis test for deviation of r from zero requires that the data follow the bivariate normal distribution (Fisz, 1963, section 9.9). Generally, the variables under consideration in this article may be assumed to be lognormally distributed; therefore, we apply the logarithmic transformation. The one-sided alternative hypothesis depends on the explanatory variable under consideration. For example, we expect $r < 0$ for the correlation between $\ln(V_{S30})$ and $\ln[a(T)]$. In contrast, we expect $r > 0$ for the correlation between $\ln[a^{\text{SRI}}(T)]$ and $\ln[a(T)]$.

Mapping Site Response

Some of the site-response explanatory variables that we consider in this article are available throughout the region of interest, namely V_{S30}^{geo} and δ . In contrast, explanatory variables derived from the SASW measurements $[V_{S30}^{\text{SASW}}, a^{\text{SRI}}(T), \text{ and } a^{\text{SH1D}}(T)]$ and f_0 are only available at 52 locations.

Table 3

Reference Rock Velocity and Attenuation Model for the SRI and SH1D Calculations

Thickness (m)	V_S (m/s)	Q_s
10	580	10
40	900	10
50	1200	20
100	1600	50
∞	1800	∞

If we wish to create maps of site-response amplification, then we must interpolate the explanatory variables that do not cover the extent of the region of interest. We use the ordinary kriging (OK) method as described by Cressie (1993) and employed by Thompson, Baise *et al.* (2010) to map site-response amplifications in the Kobe, Japan, sedimentary basin. Because of space limitations, we do not reproduce the mathematics of the OK method. Instead, we adopt the same notation as Cressie (1993) so that it can easily be used as a reference for the spatial statistics (i.e., kriging) material in this article.

Because our goal is to integrate different explanatory variables to predict $a(T)$, we are especially interested in the variance of the OK prediction (σ_k^2), which varies in space and increases with distance from the observed data. Kriging is a generalized least-squares regression algorithm that minimizes σ_k^2 . The kriging equations fundamentally depend on the spatial correlation structure of the data, which is typically represented by the variogram $[2\gamma(\mathbf{h})]$, where \mathbf{h} is the vector of Euclidean distances between two locations. One often analyzes the semivariogram $[\gamma(\mathbf{h})]$ to diagnose the spatial structure. In this article, we use the classical variogram estimator $[2\hat{\gamma}(\mathbf{h}); \theta]$, where θ is the vector of model parameters. We use a three-parameter semivariogram model that is termed the Whittle–Matérn model (Guttorp and Gneiting, 2005).

Regression Analysis

We apply the standard linear regression formulation

$$\ln[a(T)] = \beta_0 + \beta_1 X + \epsilon \quad (2)$$

for predicting $\ln[a(T)]$ from an explanatory variable X , where β_0 and β_1 are the parameters of the model (intercept and slope), and ϵ is a normally distributed random variable with mean zero and variance σ^2 . The predicted value of $\ln[a(T)]$ is

$$\ln[\hat{a}(T)] = b_0 + b_1 X, \quad (3)$$

where b_0 and b_1 are the least-squares estimates of β_0 and β_1 (Draper and Smith, 1981).

A key assumption of the linear model is that X is not a random variable, and thus has zero variance. Under this assumption, an estimate of the variance of a predicted value of $\ln[a(T)]$ (σ_p^2) is given by

$$s_p^2 = s^2 \left[1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right], \quad (4)$$

where X_0 is a particular value of X , X_i for $i = 1, \dots, n$ are the observations of X , \bar{X} is the sample mean of X_i , and an estimate of σ^2 (the variance of ϵ) is

$$s^2 = \frac{1}{n-2} \sum_{i=1}^n \{\ln[a(T)]_i - \ln[\hat{a}(T)]_i\}^2. \quad (5)$$

We use the logarithmic transformation of each of the explanatory variables [e.g., $X = \ln(V_{S30})$ rather than $X = V_{S30}$] because all of the explanatory variables are necessarily positive and the transformation removes much of the skewness in each of the variables. Again, we use superscripts to distinguish between the different methods of computing $\ln[\hat{a}(T)]$: for example, when $X = \ln(f_0)$ we write $\ln[\hat{a}^{f_0}(T)]$ for the resulting estimate of $\ln[a(T)]$.

For $X = \ln(V_{S30}^{\text{geo}})$, the reliability of the observations of X is variable. The standard deviation of $\ln(V_{S30}^{\text{geo}})$ clearly depends on geologic unit (see table 1 in Wills and Clahan, 2006). To account for this, we perform weighted least-squares regression, assuming that the observations of X are independent, but σ^2 is proportional to the variance of $\ln(V_{S30}^{\text{geo}})$. Thus, the weights equal $1/s_i^2$, where s_i^2 is set to

the square of the sample standard deviations of $\ln(V_{S30}^{\text{geo}})$ reported in table 1 of Wills and Clahan (2006).

Regression Predictions at Unsampld Locations

For explanatory variables estimated by kriging (X_k), which include point measurements such as f_0 and the variables derived from SASW profiles [V_{S30}^{SASW} , $a^{\text{SRI}}(T)$, $a^{\text{SH1D}}(T)$], the assumption that the variance of X is zero is severely violated. For these data, the variance of X increases with increasing distance to the observations. Thus, we must add another error term (ν) to equation (2) to predict $\ln[a(T)]$ at locations where X is a kriging estimate:

$$\ln[a_k(T)] = \beta_0 + \beta_1(X_k + \nu) + \epsilon, \quad (6)$$

where ν is a normally distributed random variable with mean zero and variance σ_k^2 (the variance of the kriging estimate). Equation (4) does not hold under these conditions, so we

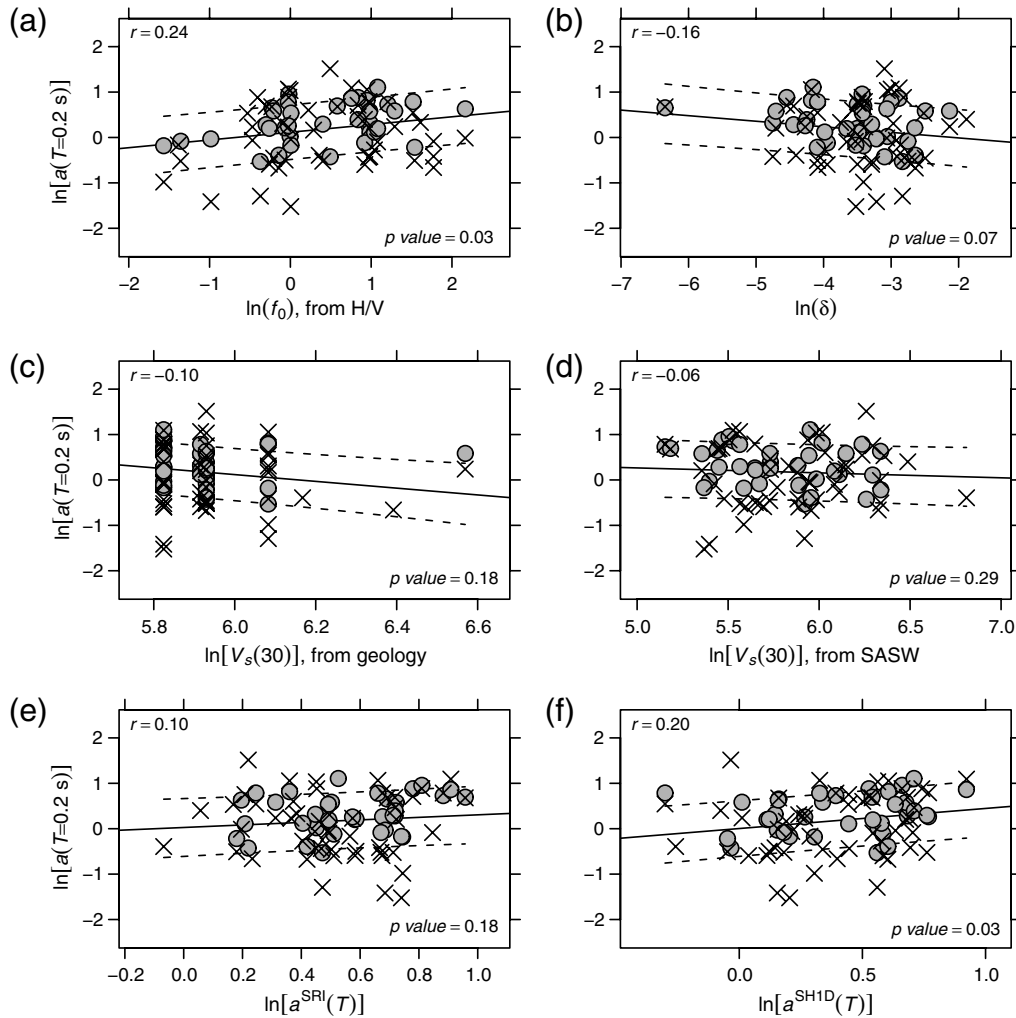


Figure 2. Correlation of $\ln[a(T = 0.2 \text{ s})]$ with (a) $\ln(f_0)$, (b) $\ln(\delta)$, (c) $\ln(V_{S30}^{\text{geo}})$, (d) $\ln(V_{S30}^{\text{SASW}})$, (e) $\ln[a^{\text{SRI}}(T)]$, and (f) $\ln[a^{\text{SH1D}}(T)]$. The 1983 Coalinga earthquake data are shown as circles and the 2004 Parkfield earthquake data are shown as crosses. The dashed lines indicate plus/minus one standard deviation prediction intervals.

estimate the variance of a prediction at an unsampled location ($s_{p,k}^2$) with the Monte Carlo method as follows:

1. For the first X_k , generate one sample of size 100,000 for each random variable in equation (6). The random variables are distributed according to

$$\begin{aligned}\beta_0 &\sim t(\mu_{\beta_0}, n-2, \sigma_{\beta_0}), \\ \beta_1 &\sim t(\mu_{\beta_1}, n-2, \sigma_{\beta_1}), \\ \nu &\sim \mathcal{N}(0, \sigma_k), \quad \text{and} \quad \epsilon \sim \mathcal{N}(0, \sigma),\end{aligned}\quad (7)$$

where $t(\mu, n_{df}, \sigma)$ is the t -distribution with mean μ , degrees of freedom n_{df} , and standard deviation σ ; and $\mathcal{N}(\mu, \sigma)$ is the normal distribution with mean μ and standard deviation σ . To generate the random sample, substitute the least-squares estimates of the means and standard deviations for the theoretical values in equation (7). The vectors of randomly generated samples in equation (7) are denoted \mathbf{b}_0 , \mathbf{b}_1 , \mathbf{n} , and \mathbf{e} , respectively.

2. Compute one sample of size 100,000 of $\ln[a_k(T)]$:

$$\ln[a_k(T)] = \mathbf{b}_0 + \mathbf{b}_1(X_k + \mathbf{n}) + \mathbf{e}. \quad (8)$$

3. Compute $s_{p,k}^2$ as the sample variance of $\ln[a_k(T)]$.
4. Repeat the previous steps for each value of X_k .

Note that we do not need to adjust the estimators of β_0 and β_1 to account for ν because the data used in the regressions are not kriged estimates.

Integrating Multiple Estimates of Site-Response Amplification

The method we propose for combining multiple estimates of $\ln[a(T)]$ relies on the prediction variance: $s_{p,k}^2$ for $\hat{a}^{f_0}(T)$ and $\hat{a}^{\text{SRI}}(T)$ are denoted $s_{f_0}^2(T)$ and $s_{\text{SRI}}^2(T)$, respectively; and s_p^2 for $\hat{a}^{\text{geo}}(T)$ and $\hat{a}^\delta(T)$ are denoted $s_{\text{geo}}^2(T)$ and $s_\delta^2(T)$, respectively. We combine four different estimates into

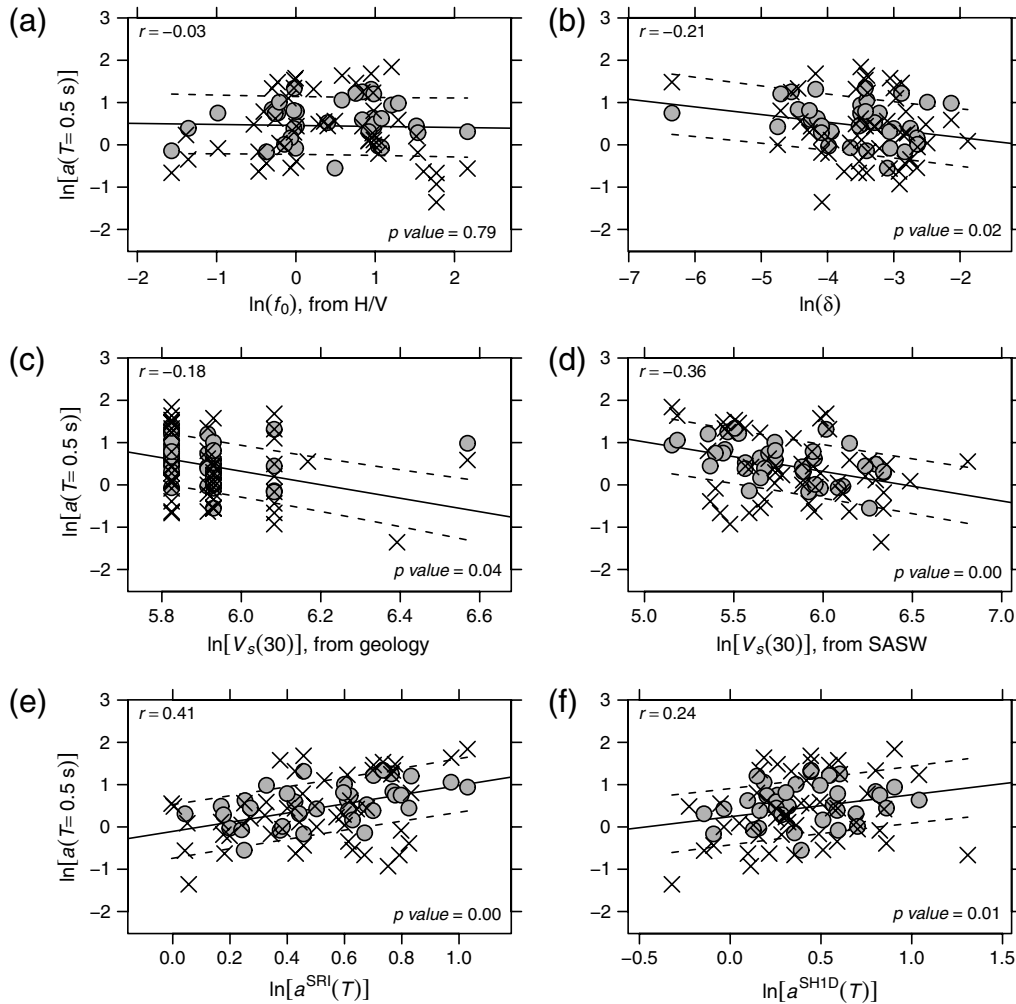


Figure 3. Correlation of $\ln[a(T = 0.5 \text{ s})]$ with (a) $\ln(f_0)$, (b) $\ln(\delta)$, (c) $\ln(V_{S30}^{\text{geo}})$, (d) $\ln(V_{S30}^{\text{SASW}})$, (e) $\ln[a^{\text{SRI}}(T)]$, and (f) $\ln[a^{\text{SHID}}(T)]$. The 1983 Coalinga earthquake data are shown as circles and the 2004 Parkfield earthquake data are shown as crosses. The dashed lines indicate plus/minus one standard deviation prediction intervals.

a variance-weighted estimate: $\ln[\hat{a}^{\text{vwe}}(T)]$. This is simply a weighted average of the different estimates of $\ln[a(T)]$, where each estimate is given a weight equal to the reciprocal of the variance of the estimate

$$\ln[\hat{a}^{\text{vwe}}(T)] = \left\{ \ln[\hat{a}^{f_0}(T)]s_{f_0}^2(T)^{-1} + \ln[\hat{a}^{\text{SRI}}(T)]s_{\text{SRI}}^2(T)^{-1} \right. \\ \left. + \ln[\hat{a}^{\text{geo}}(T)]s_{\text{geo}}^2(T)^{-1} + \ln[\hat{a}^{\delta}(T)]s_{\delta}^2(T)^{-1} \right\} \\ / [s_{f_0}^2(T)^{-1} + s_{\text{SRI}}^2(T)^{-1} + s_{\text{geo}}^2(T)^{-1} + s_{\delta}^2(T)^{-1}]. \quad (9)$$

Note that equation (9) combines four specific explanatory variables to obtain the new variance-weighted estimate. Available data vary by region, so this equation can simply be modified to reflect the available data. For example, one may have only measurements of V_{S30} , and thus cannot compute $a^{\text{SRI}}(T)$. If this is the case, V_{S30} can simply be substituted

for $a^{\text{SRI}}(T)$. Also, surficial geology maps or ground motions to derive f_0 may not be available. In such cases, the $a^{\text{geo}}(T)$ and $a^{f_0}(T)$ terms can simply be dropped.

Regression of Topographic Slope with NGA Database

Wald and Allen (2007) correlated topographic slope with V_{S30} and recommended that the site-response amplifications then be computed from V_{S30} using the Borchardt (1994) equations. We refer to this procedure, with the Allen and Wald (2009) update, as WA07. We hypothesize that a substantial improvement can be achieved by correlating $\ln[a(T)]$ directly with $\ln(\delta)$ using the NGA ground motion database.

Of the four NGA GMPEs used in the definition of $\text{PSA}^{\text{ref}}(T)$, BA08 has the most restrictive data selection criteria. We use the BA08 subset of the database, which includes many records in the NGA flatfile, including aftershocks, singly recorded earthquakes, and a number of other criteria (see Boore and Atkinson, 2007, for the complete list).

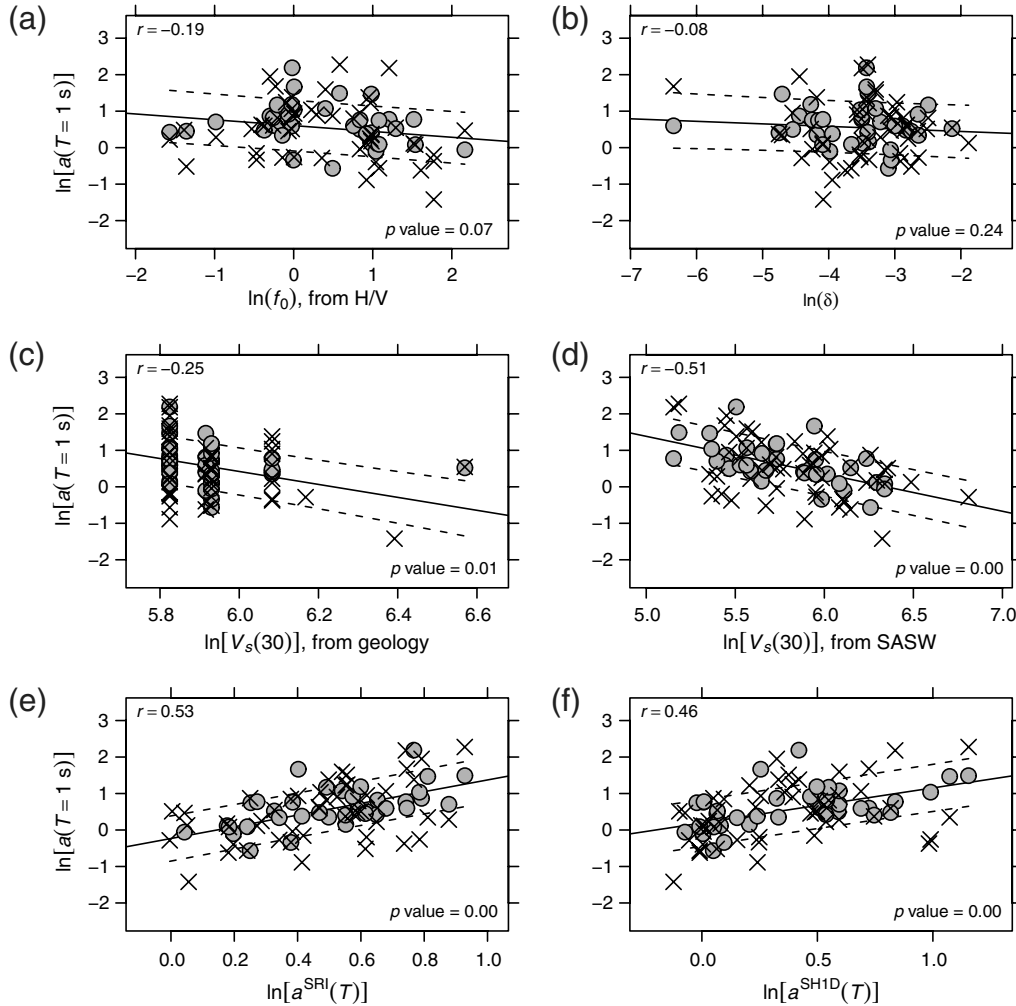


Figure 4. Correlation of $\ln[a(T = 1.0 \text{ s})]$ with (a) $\ln(f_0)$, (b) $\ln(\delta)$, (c) $\ln(V_{S30}^{\text{geo}})$, (d) $\ln(V_{S30}^{\text{SASW}})$, (e) $\ln[a^{\text{SRI}}(T)]$, and (f) $\ln[a^{\text{SH1D}}(T)]$. The 1983 Coalinga earthquake data are shown as circles and the 2004 Parkfield earthquake data are shown as crosses. The dashed lines indicate plus/minus one standard deviation prediction intervals.

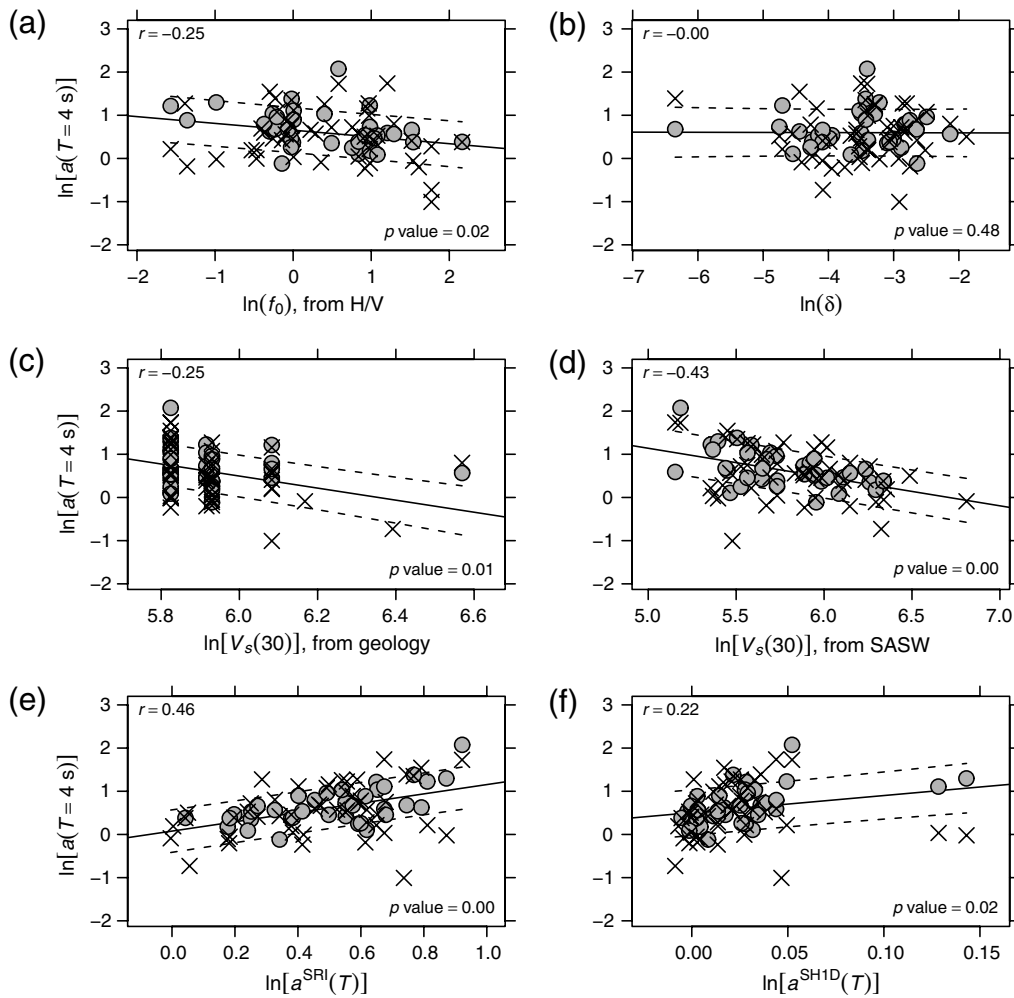


Figure 5. Correlation of $\ln[a(T = 4.0 \text{ s})]$ with (a) $\ln(f_0)$, (b) $\ln(\delta)$, (c) $\ln(V_{s30}^{\text{geo}})$, (d) $\ln(V_{s30}^{\text{SASW}})$, (e) $\ln[a^{\text{SRI}}(T)]$, and (f) $\ln[a^{\text{SH1D}}(T)]$. The 1983 Coalinga earthquake data are shown as circles and the 2004 Parkfield earthquake data are shown as crosses. The dashed lines indicate plus/minus one standard deviation prediction intervals.

The NGA database contains records that span a wider range of PSA, so it is important to add another term to the linear model (equations 2 and 3) to account for the dependence of $\ln[a(T)]$ on the intensity of the reference ground motion (we avoid the term nonlinear because of its different meanings in the context of linear regression and soil dynamics):

$$\ln[a(T)] = \beta_0 + \beta_1 \ln(\delta) + \beta_2 \ln[\text{PSA}^{\text{ref}}(T)] + \epsilon, \quad (10)$$

$$\ln[\hat{a}(T)] = b_0 + b_1 \ln(\delta) + b_2 \ln[\text{PSA}^{\text{ref}}(T)], \quad (11)$$

where β_2 is the coefficient to account for the ground-motion intensity and b_2 is the least-squares estimate of β_2 .

Table 4
Correlation Coefficients for the Different Site-Response Explanatory Variables as a Function of PSA Period

T (s)	$\ln(f_0)$	$\ln(\delta)$	$\ln(V_{s30}^{\text{geo}})$	$\ln(V_{s30}^{\text{SASW}})$	$\ln[a^{\text{SRI}}(T)]$	$\ln[a^{\text{SH1D}}(T)]$
0.05	0.19	-0.22	-0.11	-0.06	0.14	0.06
0.10	0.21	-0.20	-0.12	-0.01	0.09	0.02
0.20	0.24	-0.16	-0.10	-0.06	0.10	0.20
0.40	0.10	-0.21	-0.14	-0.24	0.28	0.14
0.50	-0.03	-0.21	-0.18	-0.36	0.41	0.24
0.75	-0.17	-0.16	-0.26	-0.49	0.55	0.53
1.00	-0.19	-0.08	-0.25	-0.51	0.53	0.46
2.00	-0.20	0.01	-0.19	-0.42	0.44	0.16
4.00	-0.25	-0.00	-0.25	-0.43	0.46	0.22

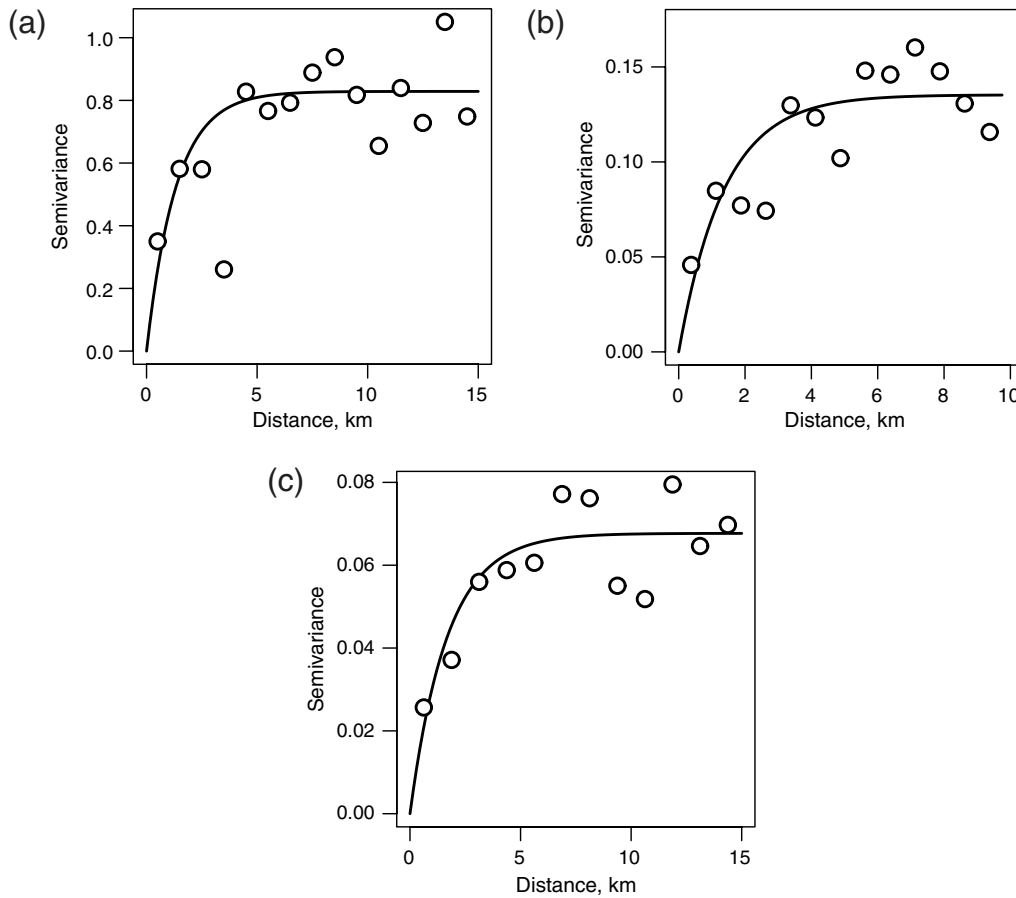


Figure 6. (a) Semivariogram of $\ln(f_0)$ for the 51 stations that recorded the 1983 Coalinga and/or the 2004 Parkfield earthquakes. (b) Semivariogram of $\ln(S_{S30})$ at the 52 SASW sites. (c) Semivariogram of $\ln[a^{SRI}(T = 0.5 \text{ s})]$ at the 52 SASW sites. The points are $\hat{\gamma}(\mathbf{h})$ and the lines are $\gamma(\mathbf{h}; \theta)$.

Because $\delta = 0$ at five of the sites in our analysis and $\ln(0)$ is undefined, we apply a minimum value such that $\delta \geq 5 \times 10^{-4}$. We select this value because it is approximately the minimum nonzero value of δ in these data. Note that the problem of handling $\delta = 0$ does not arise for Wald and Allen (2007) because they use cutoff values to define site categories rather than regression analysis; thus, they do not require the natural logarithm of slope.

This method is only valuable if it improves upon WA07. Similar to the approach by Stewart *et al.* (2003), we compare the effectiveness of different site-response metrics. To assess model uncertainty, we measure the accuracy of the different models with the root mean square error (rmse):

$$\text{rmse}(T) = \sqrt{\frac{1}{n} \sum_{i=1}^n \{\ln[a(T)]_i - \ln[\hat{a}(T)]_i\}^2}. \quad (12)$$

We define a series of different rmse values for different methods of computing $\ln[\hat{a}(T)]$: $\text{rmse}^\delta(T)$ indicates that $\ln[\hat{a}(T)]$ is computed from equation (10), which are the amplifications predicted by slope when regressed against the NGA database; $\text{rmse}^{\text{WA07}}(T)$ indicates that $\ln[\hat{a}(T)]$ is computed with the WA07 method, applying the Borchardt (1994) equations

to V_{S30}^δ ; and $\text{rmse}^{\text{BA08}}(T)$ for when $\ln[\hat{a}(T)]$ is computed from the BA08 equations with the V_{S30} values specified in the NGA flatfile. These V_{S30} values are a combination of V_{S30}^{geo} and site-specific geophysical measurements. The site term in the BA08 equations is a modification of the Choi and Stewart (2005) equations. In this way, we consider $\text{rmse}^{\text{BA08}}(T)$ as the rmse standard by which we judge the other models.

Results

Correlation of Site-Response Residuals

The site-response residuals for $T = 0.2, 0.5, 1.0, 4.0 \text{ s}$ are shown in Figures 2, 3, 4, and 5, respectively. Figures 2a–f, 3a–f, 4a–f, and 5a–f are for $X = \ln(f_0), \ln(\delta), \ln(V_{S30}^{\text{geo}}), \ln(V_{S30}^{\text{SASW}}), \ln[a^{SRI}(T)]$, and $\ln[a^{\text{SH1D}}(T)]$, respectively. The value of r is reported in the top-left corner of each subfigure, and the p value for the null hypothesis that $r = 0$ is reported in the bottom-right corner. Table 4 gives values of r for additional periods. Note that outliers, such as the very small value of δ in Figures 2–5b, have a substantial influence on r .

Figures 2–5 and Table 4 show clear trends in terms of the performance of the different explanatory variables. Site

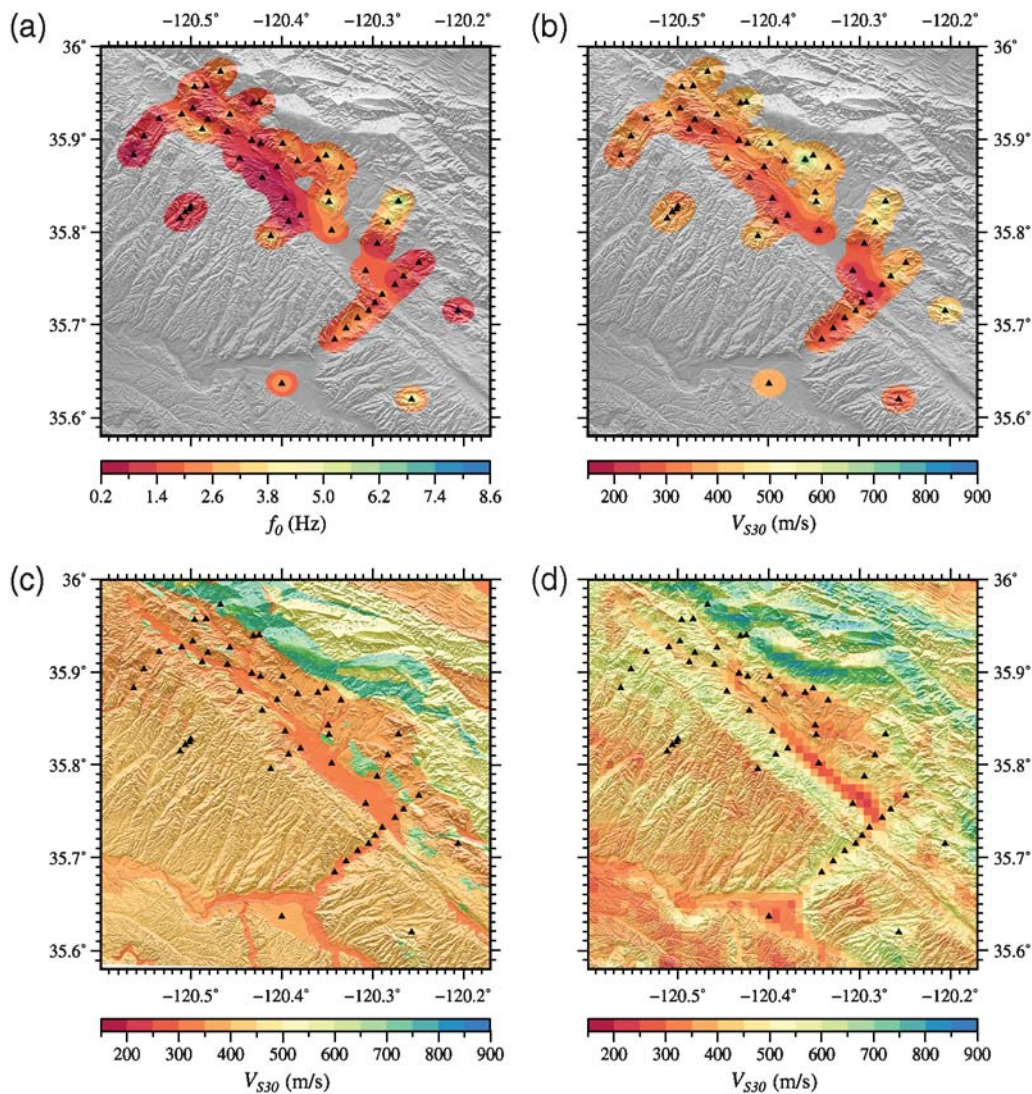


Figure 7. Maps of alternative site-response explanatory variables: (a) the dominant frequency f_0 (Zhao *et al.*, 2006), (b) V_{S30} estimated from SASW measurements (Thompson, Kayen, *et al.*, 2010), (c) V_{S30} estimated from surficial geology (Wills and Clahan, 2006), and (d) V_{S30} estimated from topographic slope (Wald and Allen, 2007; Allen and Wald, 2009).

response is positively correlated with f_0 at small periods and negatively correlated with f_0 at larger periods. Because of this gradual transition, both r and b_1 become statistically insignificant for intermediate periods (see the Discussion section for more details).

In contrast, $|r|$ for all of the SASW-derived explanatory variables is greater at larger periods. The period-dependence of $|r|$ for $\ln(\delta)$ and $\ln(V_{S30}^{\text{geo}})$ are mirror images: $|r|$ for $\ln(\delta)$ is greatest at short periods and becomes insignificant at periods of approximately 1 s and longer, while $|r|$ for $\ln(V_{S30}^{\text{geo}})$ is small for short periods and becomes larger at longer periods.

For our subsequent analysis, we eliminate all but one of the explanatory variables derived from the SASW data because the information contained in these variables are nearly redundant. We easily dismiss $\ln[a^{\text{SH1D}}(T)]$ because $|r|$ is generally smaller than $\ln(V_{S30}^{\text{SASW}})$ and $\ln[a^{\text{SRI}}(T)]$. Note that $\ln(V_{S30}^{\text{SASW}})$

and $\ln[a^{\text{SRI}}(T)]$ perform nearly equally well at most periods. We drop $\ln(V_{S30}^{\text{SASW}})$ from our subsequent analysis because $|r|$ is slightly greater for $\ln[a^{\text{SRI}}(T)]$. We are then left with four independent explanatory variables that we use to map $a(T)$, each exhibiting distinct spatial characteristics: $\ln(f_0)$, $\ln(\delta)$, $\ln(V_{S30}^{\text{geo}})$, and $\ln[a^{\text{SRI}}(T)]$. Although $\ln[a^{\text{SRI}}(T)]$ generally exhibits the best correlation with $\ln[a(T)]$, the other explanatory variables have distinct advantages. For example, the accuracy of both $\ln[a^{\text{SRI}}(T)]$ and $\ln(f_0)$ will diminish as distance increases from the observed values. Thus, $\ln(\delta)$ and $\ln(V_{S30}^{\text{geo}})$ should be included for locations distant from the SASW V_S profiles and strong-motion stations.

Mapping Explanatory Variables

The raster and polygon data (δ and V_{S30}^{geo}) do not require further analysis for mapping because these values are

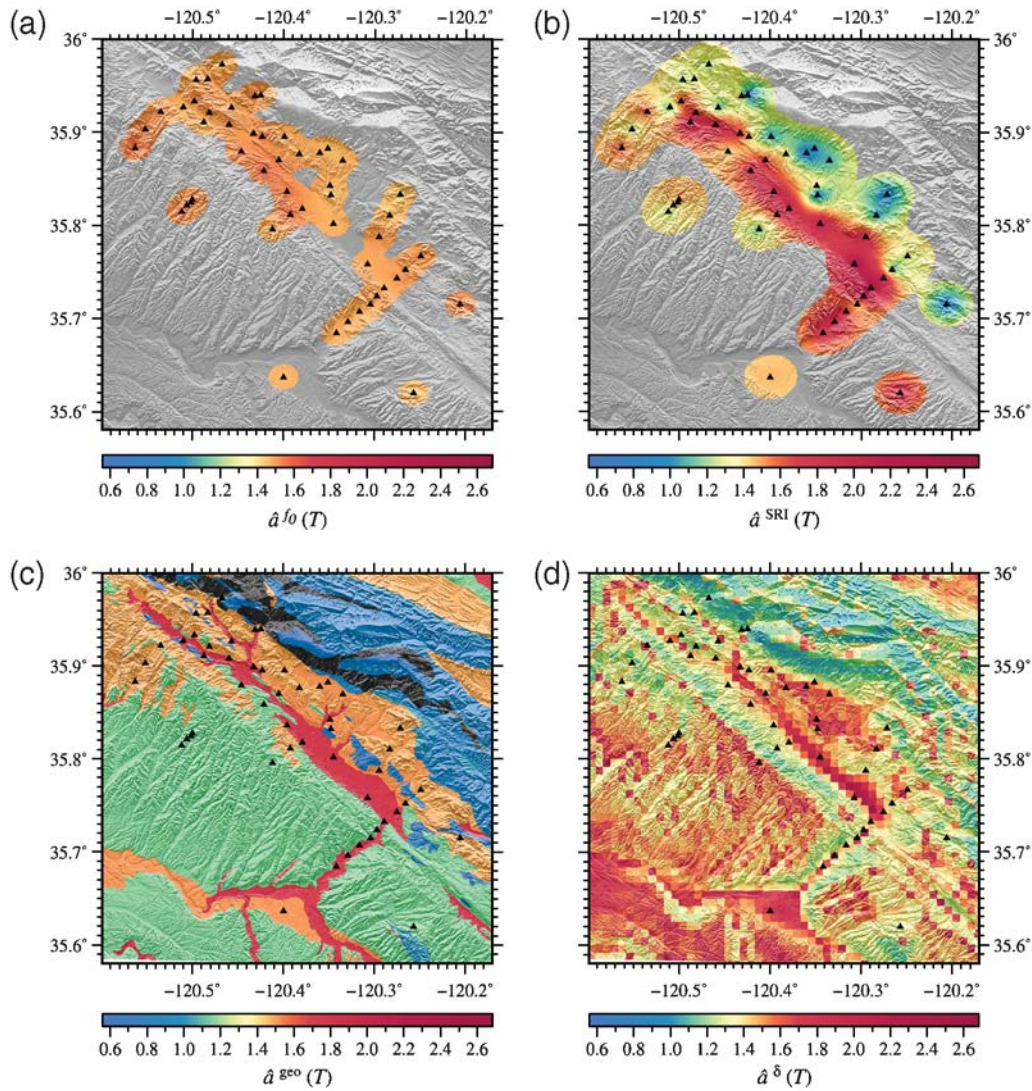


Figure 8. Maps of four different estimates of $a(T = 0.5 \text{ s})$: (a) $\hat{a}^{f_0}(T)$, (b) $\hat{a}^{S^{\text{RI}}}(T)$, (c) $\hat{a}^{S^{\text{geo}}}(T)$, and (d) $\hat{a}^{\delta}(T)$.

spatially continuous throughout the region. In contrast, we must interpolate the point data [f_0 and $a^{S^{\text{RI}}}(T)$]. Reliable interpolation with the kriging method is only possible if $\gamma(\mathbf{h})$ increases with \mathbf{h} , demonstrating that points that are closer together are more similar than those that are farther apart. Figure 6 shows $\hat{\gamma}(\mathbf{h})$ and $\gamma(\mathbf{h}; \theta)$ for f_0 in 6a, the average S -wave slowness (S_S ; inverse of V_S) to 30 m depth (S_{S30}) in 6b, and $a^{S^{\text{RI}}}(T = 0.5 \text{ s})$ in 6c.

Here, we must explain why we display the $\hat{\gamma}(\mathbf{h})$ and $\gamma(\mathbf{h}; \theta)$ for S_{S30} rather than V_{S30} in Figure 6b. Although V_{S30} is displayed in Figure 7, we do not krig V_{S30} directly; instead we krig S_{S30} and convert the result to V_{S30} . This is analogous to how V_{S30} is computed from a 1D V_S profile: S_{S30} is computed and then converted to V_{S30} . See discussions in Brown *et al.* (2002) and Boore and Thompson (2007) for the motivation of using S_S as the primary variable of interest. Figure 6 demonstrates that kriging is an effective method for spatially interpolating these variables because $\hat{\gamma}(\mathbf{h})$ increases with \mathbf{h} .

Figure 7 compares maps of the different explanatory variables, converted to V_{S30} when applicable. The three different maps of V_{S30} (Fig. 7b–d) use a single color palette. This figure fulfills two purposes: (1) visualize the spatial variability of the different explanatory variables and (2) contrast the different methods of V_{S30} estimation. Although $a^{S^{\text{RI}}}(T)$ is not included, the spatial structure of $a^{S^{\text{RI}}}(T)$ is very similar to V_{S30}^{SASW} , with minor variations as a function of T . For the f_0 and V_{S30}^{SASW} maps, we use only the kriged estimates where σ_k^2 is less than the sample variance of the measured data.

Maps of $\hat{a}(T = 0.5 \text{ s})$ are displayed in Figure 8 for the four different explanatory variables. Each subfigure uses the same color palette to visually highlight the differences. Analogous maps can easily be created for other periods, but we only display one period as an illustrative example due to space constraints. The estimates $\hat{a}^{f_0}(T)$ and $\hat{a}^{S^{\text{RI}}}(T)$ show detailed variability near the stations, but cannot accurately

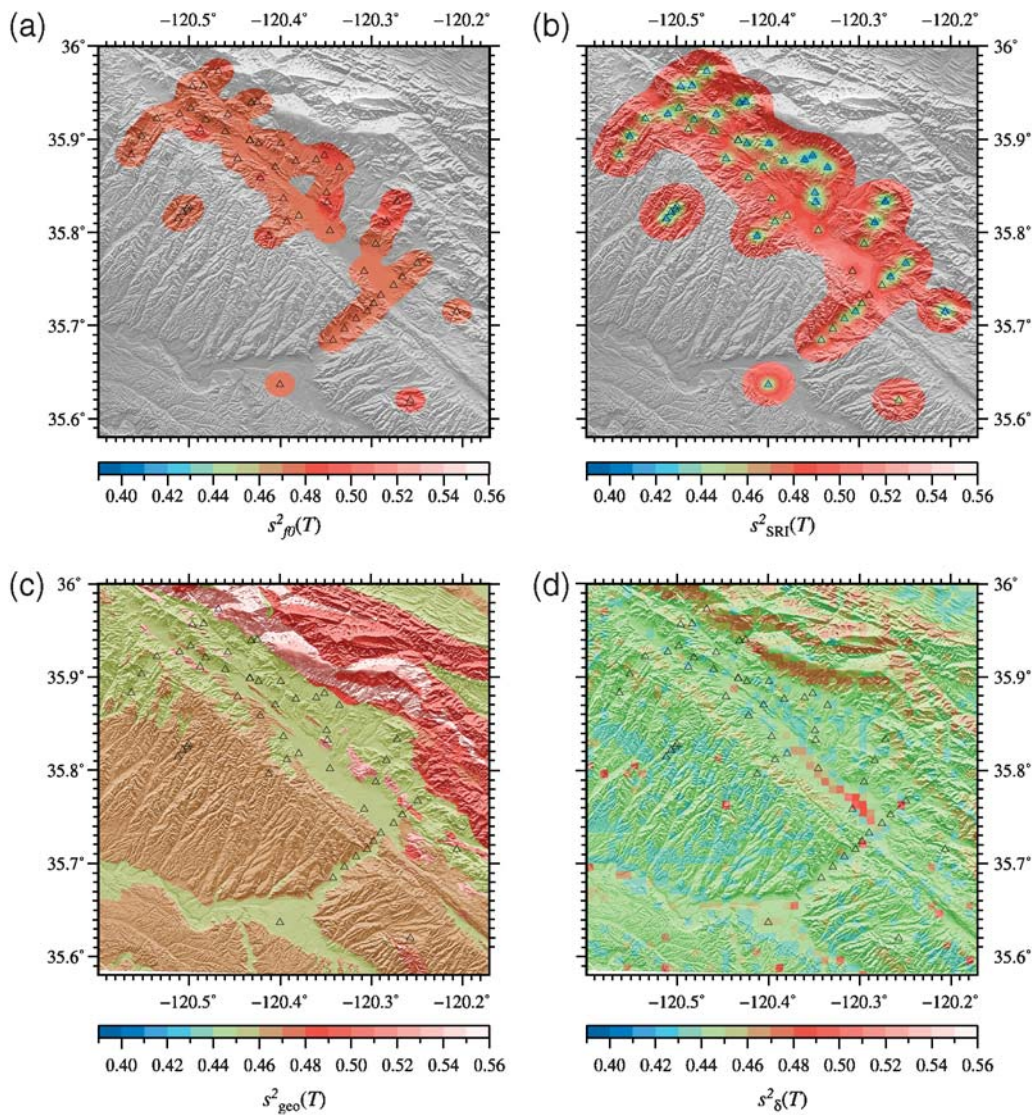


Figure 9. Maps of the variance of the $a(T = 0.5 \text{ s})$ estimates in Figure 8: (a) $s^2_{f_0}(T)$, (b) $s^2_{\text{SRI}}(T)$, (c) $s^2_{\text{geo}}(T)$, and (d) $s^2_{\delta}(T)$. We use open triangles for the stations in this map because the variance changes significantly over short distances in the vicinity of each station.

predict the amplification throughout most of the mapped region. In contrast, the resolution of $\hat{a}^{\text{geo}}(T)$ and $\hat{a}^{\delta}(T)$ does not vary with distance from the data. This highlights the major challenge that must be addressed to map site response: the performance of the different explanatory variables depends both on location (distance to the observed data) as well as T (as summarized in Table 4).

Figure 9a–d plots $s^2_{f_0}(T)$, $s^2_{\text{SRI}}(T)$, $s^2_{\text{geo}}(T)$, and $s^2_{\delta}(T)$, respectively, again using a single color palette for all of the maps. These figures show that for $T = 0.5 \text{ s}$, $\hat{a}^{\text{SRI}}(T)$ has the lowest prediction variance at the strong-motion station locations, which is consistent with the fact that it has the largest value of $|r|$ at this period (see Table 4). The second lowest prediction variance is achieved by $\hat{a}^{\delta}(T)$, which is again consistent with the r values reported in Table 4. The prediction variance of $\hat{a}^{\text{geo}}(T)$ is clearly a function of geologic unit. The prediction variance of $\hat{a}^{f_0}(T)$ is

relatively large for this period, but performs better at all other periods. The poor performance of $\hat{a}^{f_0}(T)$ for $T = 0.5 \text{ s}$ can also be seen in Figure 8, which shows that it predicts a nearly constant value of $a(T)$, reflecting the low value of r . It is also important to note that although the prediction variance is smallest for $\hat{a}^{\text{SRI}}(T)$ at the station locations, the variance increases rapidly with distance, highlighting the value of the slope and geology methods.

Figure 10a–d plots $\hat{a}^{\text{vwe}}(T)$ for $T = 0.2 \text{ s}$, $T = 0.5 \text{ s}$, $T = 1.0 \text{ s}$, and $T = 4.0 \text{ s}$, respectively. The striking trend in these maps is the trade-off in the dominant contribution between slope and geology as a function of period. Because the prediction variance of the point data increases so rapidly as a function of distance from the stations, the maps are dominated by the geology and slope methods. At short periods, slope controls $\hat{a}^{\text{vwe}}(T)$, and geology has more influence at longer periods.

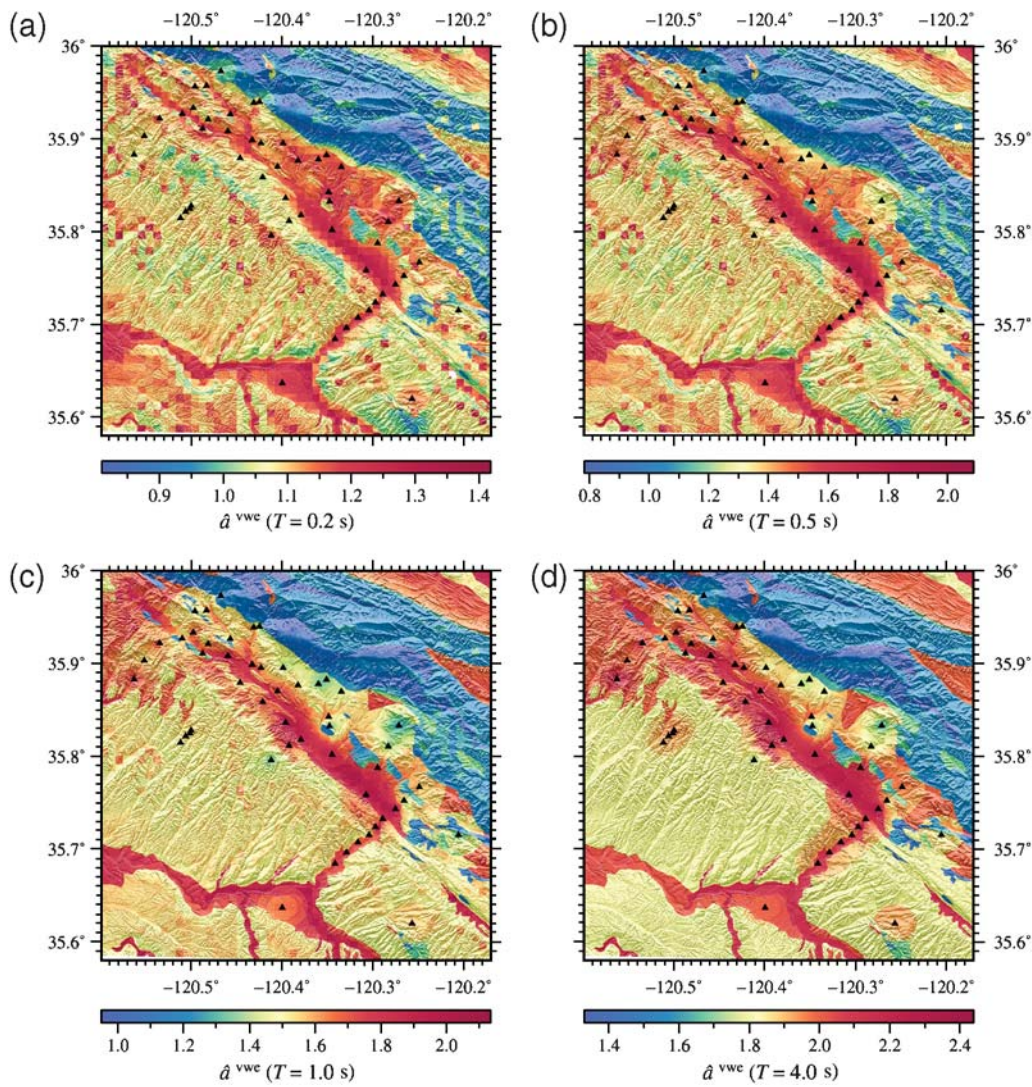


Figure 10. Maps of $\hat{a}^{vwc}(T)$ for (a) $T = 0.2$ s, (b) $T = 0.5$ s, (c) $T = 1.0$ s, and (d) $T = 4.0$ s.

NGA Database and Topographic Slope

We assess the performance of the regression of $\ln[a(T)]$ against $\ln(\delta)$ for the NGA database (equation 10) in Table 5, which includes the number of records that fulfill the BA08 record selection criteria, the estimates of the regression parameters and their p values, and the rmse for the different estimation methods for the periods used by BA08. These data are more easily visualized in Figure 11, which plots b_1 , b_2 , and the rmse for the different estimation methods as a function of T .

Figure 11a shows the trends of the model parameters as a function of period. As expected, Figure 11a shows that the values of b_1 and b_2 are all negative, indicating that amplification is inversely related to topographic slope and the reference ground-motion intensity. Also, we see that the magnitude of b_2 decreases as T increases, indicating that the degree of soil nonlinearity is larger for smaller period motions.

Figure 11b shows the trends of the model uncertainty as a function of period. The rmse data show that equation (10)

accounts for more variability in the site-response term than WA07, particularly at longer periods ($T \geq 0.3$ s). Figure 11b also shows that both slope methods (WA07 and equation 10) slightly outperform the BA08 model for $T \leq 0.2$ s. At other periods, the BA08 model generally has smaller rmse than the slope methods. $\text{rmse}^\delta(T)$ is marginally smaller than $\text{rmse}^{\text{WA07}}(T)$ for $T \leq 0.3$ s and shows more improvement over $\text{rmse}^{\text{WA07}}(T)$ at longer periods. Equation (10) performs poorest relative to the BA08 model for $0.75 \leq T \leq 1$ s, but at other periods $\text{rmse}^\delta(T)$ is either comparable to (within 3%) or smaller than $\text{rmse}^{\text{BA08}}(T)$.

Discussion

Alternative Site-Response Variables

This article has quantified the performance of many different commonly used site-response explanatory variables using the Parkfield dataset of strong-motion records and velocity measurements. These variables should not only be

Table 5
Summary of the Regression for Equation (10) and Comparison with Alternative Models*

T, s	No. of Records	Intercept		Slope		Intensity		rmse		
		b_0	p value	b_1	p value	b_2	p value	δ	WA07	BA08
PGA	1574	-0.214	0.000	-0.049	0.000	-0.091	0.000	0.467	0.473	0.486
PGV	1574	-0.000	1.000	-0.145	0.000	-0.054	0.000	0.487	0.523	0.488
0.010	1574	-0.211	0.000	-0.049	0.000	-0.090	0.000	0.467	0.474	0.486
0.020	1574	-0.207	0.000	-0.048	0.000	-0.089	0.000	0.468	0.474	0.487
0.030	1574	-0.199	0.000	-0.044	0.000	-0.089	0.000	0.472	0.478	0.491
0.050	1574	-0.196	0.000	-0.035	0.000	-0.095	0.000	0.484	0.491	0.501
0.075	1574	-0.172	0.000	-0.027	0.002	-0.096	0.000	0.487	0.494	0.499
0.100	1574	-0.150	0.001	-0.026	0.002	-0.096	0.000	0.496	0.500	0.506
0.150	1574	-0.104	0.019	-0.027	0.002	-0.095	0.000	0.499	0.503	0.504
0.200	1574	-0.088	0.051	-0.028	0.001	-0.101	0.000	0.512	0.519	0.511
0.250	1574	-0.090	0.051	-0.045	0.000	-0.094	0.000	0.521	0.536	0.515
0.300	1574	-0.081	0.096	-0.054	0.000	-0.089	0.000	0.544	0.565	0.532
0.400	1574	-0.049	0.321	-0.063	0.000	-0.076	0.000	0.546	0.579	0.529
0.500	1573	-0.065	0.205	-0.083	0.000	-0.066	0.000	0.557	0.570	0.539
0.750	1570	-0.161	0.003	-0.128	0.000	-0.054	0.000	0.579	0.585	0.555
1.000	1560	-0.220	0.000	-0.152	0.000	-0.049	0.000	0.581	0.589	0.559
1.500	1546	-0.311	0.000	-0.191	0.000	-0.040	0.000	0.570	0.589	0.554
2.000	1474	-0.399	0.000	-0.222	0.000	-0.034	0.000	0.567	0.597	0.567
3.000	1329	-0.504	0.000	-0.241	0.000	-0.042	0.000	0.538	0.585	0.557
4.000	1076	-0.465	0.000	-0.244	0.000	-0.036	0.001	0.543	0.603	0.568

*Summary shows the number of records available at each period, parameter estimates and their p values, and rmse for equation (10) (δ), the Wald and Allen (2007) model (WA07), and the BA08 model.

judged in terms of correlations with observed amplifications, but also in terms of data availability; for some applications, an approximate model that is more widely applicable may be more useful to the scientific and engineering community than a more accurate model that requires inputs that are difficult and expensive to estimate. However, engineering projects should always rely on detailed geotechnical and geophysical investigations. For such applications, we show that the SRI method can provide a more accurate estimate of the amplification than published correlations of V_{S30} for the Parkfield dataset.

Although Joyner *et al.* (1981) showed that the SRI amplification is the dominant process that contributes to site response, it is surprising that the more complex SH1D calculation is less accurate because the SH1D calculation accounts for the reflections and refractions at impedance contrasts. If the effects of reflections and refractions were minor but still present, then we expect to see a minor improvement over the SRI amplifications. Brown *et al.* (2002) suggested that the reason for the poor performance of SH1D arises from the sensitivity of the calculation to the details of the V_S profile, and these details are associated with

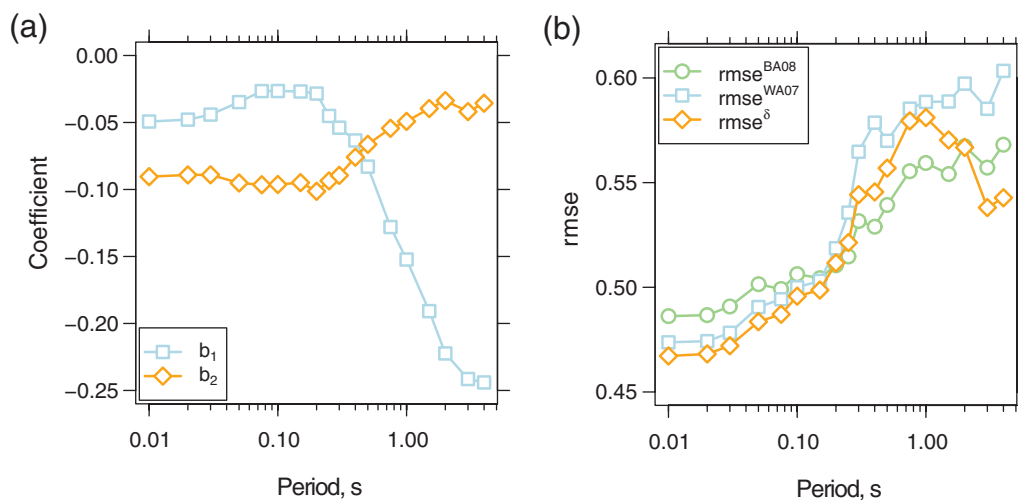


Figure 11. (a) The values of b_1 and b_2 , and (b) the rmse values reported in Table 5 as a function of period for equation 10 (rmse^δ), the Wald and Allen (2007) method ($\text{rmse}^{\text{WA07}}$), and the Boore and Atkinson (2008) equations ($\text{rmse}^{\text{BA08}}$).

a large amount of estimation uncertainty. Thus, the model uncertainty for the SH1D method may be smaller than the SRI method, but the uncertainty of estimating the input parameters results in larger estimation uncertainty for the SH1D model than for the SRI model. An alternative but related explanation is that the assumption of laterally constant layering is severely violated. The natural heterogeneity of geologic materials scatters the wavefield, which diminishes the resonances that result from the interference of reflected and refracted waves (Thompson *et al.*, 2009). This interpretation allocates more of the uncertainty to the modeling assumptions and less to the uncertainty of the input parameters.

Because of space limitations, we focus on a single period in Figures 8 and 9. These figures visualize the method that we propose, but are not helpful for judging the overall performance of one explanatory variable over another because of the strong period-dependence of $|r|$ reported in Table 4. We see that $s_{f_0}^2(T)$ substantially exceeds $s_{\text{SRI}}^2(T)$ and $s_{\delta}^2(T)$ at the station locations in Figure 9. Table 4, however, indicates that $|r|$ for $\ln(f_0)$ is larger at other periods, which will decrease the variance of $\hat{a}^{f_0}(T)$. At most periods $\ln(f_0)$ underperforms $\ln[a^{\text{SRI}}(T)]$ (except for $T \leq 0.20$ s), but f_0 is an easier explanatory variable to obtain for GMPE development: $\ln[a^{\text{SRI}}(T)]$ requires a V_S profile, which is expensive and time-consuming, while the computation of f_0 can be automated and only requires that one or more previous earthquakes have been recorded at the station of interest. Thus, f_0 may provide valuable information where $a^{\text{SRI}}(T)$ is not available.

Dominant Site Frequency

The period dependence of b_1 for $X = \ln(f_0)$ is worth discussing further. Note that b_1 is closely related to r (Table 4) through a positive proportionality constant (though the magnitude of the constant will vary with period). Thus, it is clear that $b_1 < 0$ for larger periods and $b_1 > 0$ for smaller periods. One might expect that $b_1 < 0$ at all periods because larger values of f_0 are associated with stiffer soils, which are in turn associated with less site amplification. This is true for b_1 at $T \geq 0.75$ s, but there is a clear gradual increase in b_1 as T decreases; for $T \leq 0.4$ s, $b_1 > 0$.

This trend can be explained by defining the difference between the dominant period of the soil ($T_0 = 1/f_0$) and the natural period of the SDOF system (T) for which the PSA is computed: $\Delta T = T - T_0$. As the absolute value of ΔT decreases, the resonance of the soil will align with the resonance of the SDOF, so $a(T)$ will increase.

Assume $\Delta T < 0$ for all n observations of $a(T)$. Under this condition, ΔT decreases as T_0 increases (and f_0 decreases); thus, $a(T)$ increases as T_0 increases. The inequality $\Delta T < 0$ is more likely to be fulfilled for small values of T .

Now assume $\Delta T > 0$ for all n observations of $a(T)$. Under this condition, ΔT increases as T_0 increases (and f_0 decreases); thus, $a(T)$ decreases as T_0 increases. The

inequality $\Delta T > 0$ is more likely to be fulfilled for large values of T .

It follows from these two scenarios that $b_1 < 0$ is expected for large values of T , and $b_1 > 0$ is expected for small values of T . If the inequalities $\Delta T > 0$ or $\Delta T < 0$ are not true for all n observations of $a(T)$, then the two previously described trends will both be present and have offsetting effects. The trend of r with T for $X = \ln(f_0)$ in Table 4 qualitatively supports this explanation. We plot b_1 against the number of sites that fulfill the inequality $\Delta T > 0$ in Figure 12. The trend is what we would expect: as the number of sites where $\Delta T > 0$ increases, the value of b_1 decreases. Note that $b_1 \approx 0$ where the number of sites for which $\Delta T > 0$ is approximately half the total number of sites.

Topographic Slope

The ability of topographic slope to model the site-response amplifications across a broad range of periods is the most surprising result in this article. Figure 11 shows that directly regressing $\ln(\delta)$ against the observed amplifications in the NGA flatfile is generally more accurate than the WA07 method. Table 5 and Figure 11 shows that the rmse of the WA07 model is smaller than the BA08 model for small periods ($T \leq 0.2$ s). The relatively poor performance of equation (10) at $0.75 \leq T \leq 1.0$ s is offset by (1) the gain in spatial coverage and (2) the savings in terms of time and cost of field investigations that are required to achieve the V_{S30} values in the NGA database.

The comparison of the slope-based site-response methods with the V_{S30} -based BA08 amplifications is not

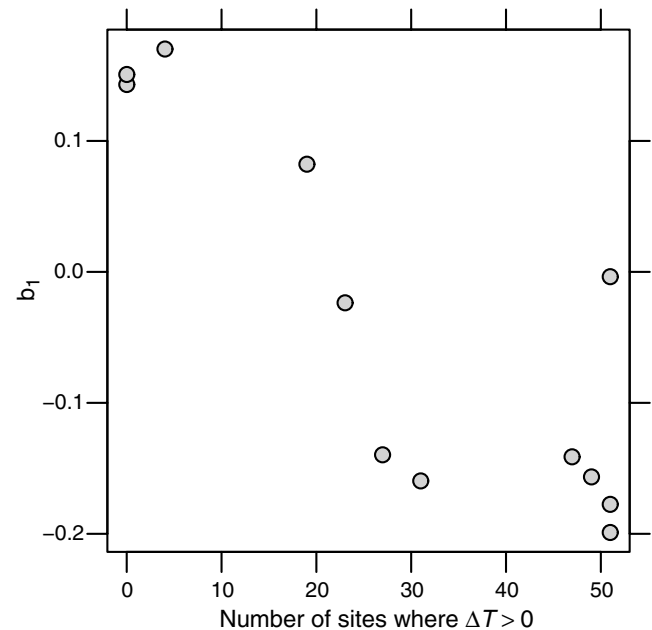


Figure 12. b_1 for $X = \ln(f_0)$ versus the number of sites that fulfill the inequality $\Delta T > 0$.

straightforward, because site response is estimated from many different sources, including site-specific V_S profiles and the Wills and Clahan (2006) site conditions map. In fact, only about 35% of the NGA records are associated with site-specific measurements of the V_S profile, while the rest are inferred (Chiou *et al.*, 2008). Moss (2008) showed that the uncertainty of these different estimates of V_{S30} are highly variable. The Parkfield data are consistent with the Moss (2008) conclusions: V_{S30} estimated from an SASW survey is a more accurate predictor of site response than the surficial geology. Thus, the rmse reported for the BA08 model is a mixture of the V_{S30} model uncertainty with a wide range of estimation uncertainties. Thus, BA08 will perform best when the estimate of V_{S30} is derived from a site-specific measurement of the V_S profile. If these data were available at all stations in the NGA flatfile, then the rmse values reported in Table 5 for BA08 would certainly be substantially smaller. But this level of accuracy is not feasible for hazard mapping, and so the more pertinent comparison for mapping applications is between the amplifications predicted by the slope methods and the amplifications predicted by BA08 when the V_{S30} values are inferred from the Wills and Clahan (2006) model. This would clearly increase the estimation uncertainty; thus, this would increase the $\text{rmse}^{\text{BA08}}$ values reported in Table 5, further emphasizing the value of the slope-based site-response estimation methods.

Conclusions

A wide variety of methods have been proposed for mapping site-response amplifications. Each has been developed with a different purpose: Thompson, Baise *et al.* (2010) sought to characterize a single urban region, while Wills and Clahan (2006) sought a map with coverage of the entire state of California, and Wald and Allen (2007) sought an estimate of site response for the entire globe. Each of these methods uses a different explanatory variable as a site-response proxy. The methodology that we propose takes advantage of the inherent benefits of each of these different methods of mapping site response. Addressing the estimation and model uncertainty of these widely varying methods is possible with the aid of the densely spaced strong-motion stations in Parkfield.

Our analysis confirms that the Wald and Allen (2007) method of correlating site response with topographic slope performs surprisingly well considering the simplicity of the method and the spatial coverage that it attains. In Parkfield, topographic slope correlates better with the site-response amplifications than surficial geology for $T \leq 0.5$ s. When more accurate results can be achieved with local geophysical surveys of the near-surface V_S profiles, linking the V_S profiles to site-response amplifications with the SRI method improves the correlation with observed site-response amplifications compared with V_{S30} . Although this improvement is modest, the effort that is required to achieve this gain is minimal and thus we feel that the effort is warranted.

Additionally, we expect that better performance can be achieved with the SRI method for the longer periods with deeper V_S profiles. Though the V_S profiles that we have in Parkfield generally extend to 100 m and deeper, it is clear from the T_{max}^{λ} values in Table 1 that the SRI method still requires extrapolation for periods greater than 1 s.

We also find that f_0 provides complementary information to the V_S profile, especially for small periods. Interestingly, we see that f_0 exhibits a direct relationship with amplification for small periods and an inverse relationship with amplification for large periods.

We show that the Wald and Allen (2007) method agrees well with the observed site response in the NGA flatfile. Building upon the work of Wald and Allen (2007), we develop new equations for estimating site response from topographic slope.

Data and Resources

V_S profiles at the strong-motion stations were measured with the SASW method and are available at <http://gdc.cee.tufts.edu> (last accessed May 2010). More details can be found in Thompson, Kayen, *et al.* (2010). We received the 1983 Coalinga and 2004 Parkfield earthquake ground motions from the Consortium of Organizations for Strong Motion Observation Systems at <http://db.cosmos-eq.org> (last accessed May 2010) and the National Strong-Motion Project <http://nsmr.wr.usgs.gov/> (last accessed May 2010). Shuttle Radar Topography Mission (SRTM) 30 s global topography data (Farr and Kobrick, 2000) are available at <http://www2.jpl.nasa.gov/srtm> (last accessed May 2010). The NGA flatfile is the database used by the NGA GMPE developers and is available at <http://peer.berkeley.edu/nga/flatfile.html> (last accessed May 2010). We computed the topographic slope from the SRTM data with Generic Mapping Tools (GMT; Wessel and Smith 1991) as described by Wald and Allen (2007) and Allen and Wald (2009). All other computations in this article were completed with the open source software R (R Development Core Team 2010). We computed the NGA GMPEs with the Kaklamanos *et al.* (2010) R package “nga”.

Acknowledgments

This research is funded by National Earthquake Hazards Reduction Program (NEHRP) Award #G09AP00062. We thank Rich Vogel and David Garman for valuable discussions and support on statistical modeling. Discussions with David Wald and Trevor Allen improved our implementation of their method for predicting V_{S30} from topographic slope. Reviews by Jon Stewart, Dave Boore, and an anonymous reviewer have substantially improved this article.

References

- Abercrombie, R. E. (1997). Near-surface attenuation and site effects from comparison of surface and deep borehole recordings, *Bull. Seismol. Soc. Am.* **87**, no. 3, 731–744.
- Abrahamson, N., and W. Silva (2008). Summary of the Abrahamson and Silva NGA ground-motion relations, *Earthquake Spectra* **24**, no. 1, 67–97.

- Allen, T. I., and D. J. Wald (2009). On the use of high-resolution topographic data as a proxy for seismic site conditions (V_{S30}), *Bull. Seismol. Soc. Am.* **99**, no. 2A, 935–943, doi [10.1785/576.0120080255](https://doi.org/10.1785/576.0120080255).
- Boore, D. M. (2003). Simulation of ground motion using the stochastic method, *Pure Appl. Geophys.* **160**, no. 3–4, 635–676.
- Boore, D. M., and G. Atkinson (2007). *Boore-Atkinson NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters*, Pacific Earthquake Engineering Center, Berkeley, California.
- Boore, D. M., and G. M. Atkinson (2008). Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s, *Earthquake Spectra* **24**, no. 1, 99–138.
- Boore, D. M., and W. B. Joyner (1997). Site amplifications for generic rock sites, *Bull. Seismol. Soc. Am.* **87**, no. 2, 327–341.
- Boore, D. M., and E. M. Thompson (2007). On using surface-source downhole-receiver logging to determine seismic slownesses, *Soil Dyn. Earthquake Eng.* **27**, no. 11, 971–985.
- Borcherdt, R. D. (1970). Effects of local geology on ground motion near San Francisco Bay, *Bull. Seismol. Soc. Am.* **60**, no. 1, 29–61.
- Borcherdt, R. D. (1994). Estimates of site-dependent response spectra for design (methodology and justification), *Earthquake Spectra* **10**, 617–654.
- Brown, L. T., D. M. Boore, and K. H. Stokoe (2002). Comparison of shear-wave slowness profiles at 10 strong-motion sites from noninvasive SASW measurements and measurements made in boreholes, *Bull. Seismol. Soc. Am.* **52**, 289–300.
- Cadet, H., P.-Y. Bard, and A. Rodriguez-Marek (2010). Defining a standard rock site: Propositions based on the KiK-net database, *Bull. Seismol. Soc. Am.* **100**, no. 1, 172–195, doi [10.1785/0120090078](https://doi.org/10.1785/0120090078).
- Campbell, K., and Y. Bozorgnia (2008). NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s, *Earthquake Spectra* **24**, no. 1, 139–71.
- Castellaro, S., F. Mulargia, and P. L. Rossi (2008). V_{S30} : Proxy for seismic amplification?, *Seismol. Res. Letters* **79**, no. 4, 540–543, doi [10.1785/gssrl.79.4.540](https://doi.org/10.1785/gssrl.79.4.540).
- Chiou, S.-J. B., and R. R. Youngs (2008). An NGA model for the average horizontal component of peak ground motion and response spectra, *Earthquake Spectra* **24**, no. 1, 173–215.
- Chiou, B., R. Darragh, N. Gregor, and W. Silva (2008). NGA project strong-motion database, *Earthquake Spectra* **24**, no. 1, 23–44.
- Choi, Y., and J. P. Stewart (2005). Nonlinear site amplification as function of 30 m shear wave velocity, *Earthquake Spectra* **21**, no. 1, 1–30, doi [10.1193/1.1856535](https://doi.org/10.1193/1.1856535).
- Cressie, N. A. C. (1993). *Statistics for Spatial Data*, Wiley, New York.
- Douglas, J. (2003). Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectral ordinates, *Earth-Science Reviews* **61**, no. 1–2, 43–104.
- Douglas, J., P. Gehl, L. F. Bonilla, O. Scotti, J. Regnier, A.-M. Duval, and E. Bertrand (2009). Making the most of available site information for empirical ground-motion prediction, *Bull. Seismol. Soc. Am.* **99**, no. 3, 1502–1520.
- Draper, N. R., and H. Smith (1981). *Applied Regression Analysis*, Wiley, New York, 709 pp.
- Dreger, D. S. (2004). *09/28/2004 Preliminary slip model*, Berkeley Seismological Laboratory, University of California, Berkeley, available at Finite-Source Rupture Model Database, <http://www.seismo.ethz.ch/srcmod/> (last accessed Dec. 2010).
- Farr, T. G., and M. Kobrick (2000). Shuttle Radar Topography Mission produces a wealth of data, *Eos Trans. AGU* **81**, 583–585.
- Field, E. H., and K. H. Jacob (1995). A comparison and test of various site-response estimation techniques, including 3 that are not reference-site dependent, *Bull. Seismol. Soc. Am.* **85**, no. 4, 1127–1143.
- Fisz, M. (1963). *Probability Theory and Mathematical Statistics*, Third Ed., Wiley, New York, 677 pp.
- Frankel, A., C. Mueller, T. Barnhard, D. Perkins, E. V. Leyendecker, N. Dickman, S. Hanson, and M. Hopper (1996). National Seismic Hazard Maps: Documentation June 1996, *U.S. Geol. Surv. Open-File Rept.* 96-532, 110.
- Frankel, A. D., M. D. Petersen, C. S. Mueller, K. M. Haller, R. L. Wheeler, E. V. Leyendecker, R. L. Wesson, S. C. Harmsen, C. H. Cramer, D. M. Perkins, and K. S. Rukstales (2002). Documentation for the 2002 update of the national seismic hazard maps, *U.S. Geol. Surv. Open-File Rept.* 02-420, 39.
- Guttorp, P., and T. Gneiting (2005). On the Whittle–Matérn correlation family, Technical Report 81, University of Washington, NRCSE Technical Report Series.
- Haskell, N. A. (1953). The dispersion of surface waves on multilayered media, *Bull. Seismol. Soc. Am.* **72**, 17–34.
- Holzer, T. L., M. J. Bennett, T. E. Noce, A. C. Padovani, and J. C. Tinsley (2002). Liquefaction hazard and shaking amplification maps of Alameda, Berkeley, Emeryville, Oakland, and Piedmont, California: A digital database, *U.S. Geol. Surv. Open-File Rept.* 02-296.
- Jennings, C. W. (1958). Geologic atlas of California, San Luis Obispo sheet, *Calif. Div. Mines Geol. Geologic Atlas Map GAM 018*, scale 1:250,000.
- Joyner, W. B., and T. E. Fumal (1984). Use of measured shear-wave velocity for predicting geologic site effects on strong ground motion, in *Proc. of the 8th World Conf. on Earthquake Engineering*, volume 2, pp. 777–783.
- Joyner, W. B., R. E. Warrick, and T. E. Fumal (1981). The effect of quaternary alluvium on strong ground motion in the Coyote Lake, California, Earthquake of 1979, *Bull. Seismol. Soc. Am.* **71**, no. 4, 1333–1349.
- Kaklamanos, J., and L. G. Baise (2011). Model validations and comparisons of the next generation attenuation of ground motions (NGA-West) project, *Bull. Seismol. Soc. Am.* **101**, no. 1, 160–175.
- Kaklamanos, J., L. G. Baise, and D. M. Boore (2011). Estimating unknown input parameters when implementing the NGA ground-motion prediction equations in engineering practice, *Earthquake Spectra* **27**, no. 3, in press.
- Kaklamanos, J., D. M. Boore, E. M. Thompson, and K. Campbell (2010). Implementation of the Next Generation Attenuation (NGA) ground-motion prediction equations in Fortran and R, *U.S. Geol. Surv. Open-File Rept.* 2010-1296, 47.
- Lermo, J., and F. J. Chávez-García (1993). Site effect evaluation using spectral ratios with only one station, *Bull. Seismol. Soc. Am.* **83**, no. 5, 1574–1594.
- Moss, R. E. S. (2008). Quantifying measurement uncertainty of 30-meter shear-wave velocity, *Bull. Seismol. Soc. Am.* **98**, no. 3, 1399–1411.
- Petersen, M. D., A. D. Frankel, S. C. Harmsen, C. S. Mueller, K. M. Haller, R. L. Wheeler, R. L. Wesson, Y. Zeng, O. S. Boyd, D. M. Perkins, N. Luco, E. H. Field, C. J. Wills, and K. S. Rukstales (2008). Documentation for the 2008 update of the United States National Seismic Hazard Maps, *U.S. Geol. Surv. Open-File Rept.* 2008-1128, 61.
- R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>, ISBN 3-900051-07-0 (last accessed May 2010).
- Shakal, A. F., H. R. Haddadi, and M. J. Huang (2006). Note on the very-high-acceleration fault zone 16 record from the 2004 Parkfield earthquake, *Bull. Seismol. Soc. Am.* **96**, (4, Part B Suppl. S), S119–S128, doi [10.1785/0120050828](https://doi.org/10.1785/0120050828).
- Steidl, J. H., A. G. Tumarkin, and R. J. Archuleta (1996). What is a reference site?, *Bull. Seismol. Soc. Am.* **86**, no. 6, 1733–1748.
- Stewart, J. P., A. H. Liu, and Y. Choi (2003). Amplification factors for spectral acceleration in tectonically active regions, *Bull. Seismol. Soc. Am.* **93**, no. 1, 332–352.
- Thompson, E. M., L. G. Baise, R. E. Kayen, and B. B. Guzina (2009). Impediments to predicting site response: Seismic property estimation and modeling simplifications, *Bull. Seismol. Soc. Am.* **99**, no. 5, 2927–2949.

- Thompson, E. M., L. G. Baise, R. E. Kayen, Y. Tanaka, and H. Tanaka (2010). A geostatistical approach to mapping site response spectral amplifications, *Eng. Geol.* **114**, no. 3–4, 330–342.
- Thompson, E. M., R. E. Kayen, B. Carlin, and H. Tanaka (2010). Surface wave site characterization at 52 strong motion recording stations affected by the Parkfield M 6.0 Earthquake of 28 September 2010, *U.S. Geol. Surv. Open-File Rept. 2010-1168*.
- Thomson, W. T. (1950). Transmission of elastic waves through a stratified solid, *J. Appl. Phys.* **21**, 89–93.
- Trifunac, M. D. (1976). Preliminary analysis of peaks of strong earthquake ground motion—Dependence of peaks on earthquake magnitude, epicentral distance, and recording site conditions, *Bull. Seismol. Soc. Am.* **66**, no. 1, 189–219.
- Wald, D. J., and T. I. Allen (2007). Topographic slope as a proxy for seismic site conditions and amplification, *Bull. Seismol. Soc. Am.* **97**, no. 5, 1379–1395.
- Wald, D. J., P. S. Earle, K.-W. Lin, V. Quitoriano, and B. Worden (2006). Challenges in rapid ground motion estimation for the prompt assessment of global urban earthquakes, *Bull. Seismol. Soc. Am.* **81**, no. 3–4, 273–281.
- Wald, D. J., B. C. Worden, V. Quitoriano, and K. L. Pankow (2005). *ShakeMap Manual: Technical Manual, User's Guide, and Software Guide*, U.S. Geological Survey, Techniques and Methods 12A1, <http://pubs.usgs.gov/tm/2005/12A01/> (last accessed May 2010).
- Wessel, P., and W. H. F. Smith (1991). Generic mapping tools, *Eos Trans. AGU* **72**, 441.
- Wills, C. J., and K. B. Clahan (2006). Developing a map of geologically defined site-condition categories for California, *Bull. Seismol. Soc. Am.* **96**, no. 4, 1483–1501.
- Zhao, J., K. Irikura, J. Zhang, Y. Fukushima, P. Somerville, A. Asano, Y. Ohno, T. Oouchi, T. Takahashi, and H. Ogawa (2006). An empirical site-classification method for strong-motion stations in Japan using H/V response spectral ratio, *Bull. Seismol. Soc. Am.* **96**, no. 3, 914–925.

Department of Civil and Environmental Engineering
 Tufts University
 Medford, Massachusetts 02155 USA
 eric.thompson@tufts.edu
 (E.M.T., L.G.B., E.C.M., J.K.)

Coastal and Marine Geology
 United States Geological Survey
 Menlo Park, California 94025 USA
 (R.E.K.)

Manuscript received 5 August 2010