Ⓔ

An Empirical Model for the Interfrequency Correlation of Epsilon for Fourier Amplitude Spectra

by Jeff Bayless and Norman A. Abrahamson

Abstract   An empirical ground-motion model (GMM) is presented for the interfrequency correlation of normalized residuals, epsilon (ϵ), for smoothed Fourier amplitude spectra (FAS). The interfrequency correlation of ϵ (ρ_ϵ) model is developed for the smoothed effective amplitude spectrum (EAS), as defined by Pacific Earthquake Engineering Research Center (PEER; Goulet *et al.*, 2018). The EAS is the orientation-independent horizontal-component FAS of ground acceleration. Ground-motion data are from the PEER Next Generation Attenuation-West2 database (Ancheta *et al.*, 2014), which includes shallow crustal earthquakes in active tectonic regions. The normalized residuals are obtained from the Bayless and Abrahamson (2018b) GMM and are partitioned into between-event, between-site, and within-site components, and a model is developed for the total correlation between frequencies. The total correlation model features a two-term exponential decay with the natural logarithm of frequency. At higher frequencies, the model differs substantially from previously published models, in which the ground-motion smoothing technique used has a large effect on the resulting correlations. The empirical ρ_ϵ is not found to have statistically significant magnitude, distance, site parameter, or regional dependence, although potential regional variations should be studied further. The model is applicable for crustal earthquakes in active tectonic regions worldwide, for rupture distances of 0–300 km, **M** 3.0–8.0, and over the frequency range 0.1–24 Hz. Tables for the total correlation model coefficients and covariance matrices are provided in the Ⓔ supplemental content to this article.

Supplemental Content: Tables of model coefficients and effective amplitude spectrum (EAS) standard deviations, correlation matrices for each residual component, total empirical correlation, and the total correlation model.

Introduction

Residuals from empirical ground-motion models (GMMs, also known as ground-motion prediction equations) are typically partitioned into between-event residual (δB) and within-event residuals (δW) following the notation of Al Atik *et al.* (2010). For large number of recordings per earthquake, the between-event residual is approximately the average difference between the observed intensity measure (IM) from a specific earthquake and the IM predicted by the GMM. The within-event residual (δW) is the difference between the IM at a specific site for a given earthquake and the median IM predicted by the GMM + δB. By accounting for repeatable site effects, δW can further be partitioned into a site-to-site residual ($\delta S2S$) and the single-station within-event residual (δWS; also called the within-site residual; e.g., Villani and Abrahamson, 2015). Using this notation, the residuals take the following form:

$$Y_{es} = g(\mathbf{X_{es}}, \boldsymbol{\theta}) + \delta B_e + \delta S2S_s + \delta WS_{es} \quad (1)$$

$$\delta_{\text{total},es} = Y_{es} - g(\mathbf{X_{es}}, \boldsymbol{\theta}) = \delta B_e + \delta S2S_s + \delta WS_{es}, \quad (2)$$

in which Y_{es} is the natural logarithm of the recorded ground-motion IM for earthquake e and site s, $g(\mathbf{X_{es}}, \boldsymbol{\theta})$ is the median GMM, $\mathbf{X_{es}}$ is the vector of explanatory seismological parameters (magnitude, distance, site conditions, etc.), $\boldsymbol{\theta}$ is the vector of GMM coefficients, and $\delta_{\text{total},es}$ is the total residual.

The residual components δB, $\delta S2S$, and δWS are well represented as zero-mean independent normally distributed random variables with standard deviations τ, ϕ_{S2S}, and ϕ_{ss}, respectively (Al Atik *et al.*, 2010). GMM residual components are converted to epsilon (ϵ_B, ϵ_{S2S}, and ϵ_{WS}) by normalizing the residuals by their respective standard deviations. Because of the normalization, the random variables ϵ_B, ϵ_{S2S}, and ϵ_{WS} are represented by standard normal distributions (mean = 0 and variance = 1). If the total residual is used, then the resulting ϵ_{total} will, in general, not have zero mean
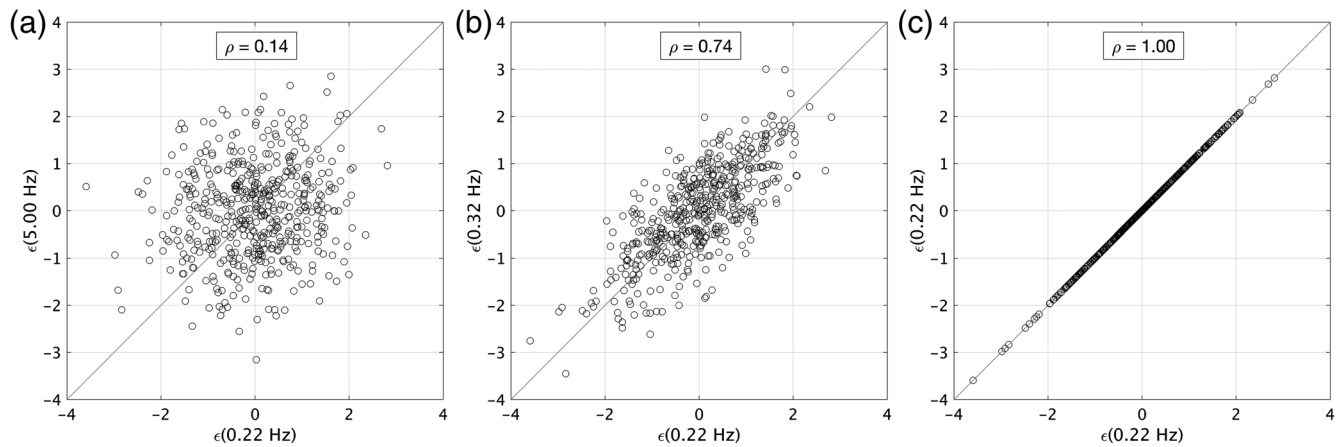
1058

Figure 1. ϵ values at pairs of frequencies, exhibiting the correlation dependent on frequency spacing. (a) $f_1 = 0.2$ Hz and $f_2 = 5.0$ Hz; (b) $f_1 = 0.2$ Hz and $f_2 = 0.3$ Hz; (c) $f_1 = 0.2$ Hz and $f_2 = 0.2$ Hz.

because of the uneven sampling of recordings per earthquake in the data set.

For a given recording, the values of ϵ at neighboring periods (T) are generally correlated. For example, if a ground motion is stronger than average at $T = 1.0$ s, then it is also likely to be stronger than expected at nearby periods, for example, $T = 0.8$ or 1.2 s; however, for a widely spaced period pair (e.g., $T = 10.0$ s compared with $T = 1.0$ s), we expect that the ϵ values will be weakly correlated. The interperiod (or equivalently, interfrequency) correlation coefficient ρ quantifies the relationship of ϵ values between periods for a given recording.

We present an empirical GMM for the interfrequency correlation of epsilon (ρ_ϵ) for smoothed Fourier amplitude spectra (FAS). The correlation model is based on recordings from the Pacific Earthquake Engineering Research Center (PEER) Next Generation Attenuation-West2 (NGA-West2) database (Ancheta *et al.*, 2014), which includes shallow crustal earthquakes in active tectonic regions. The normalized residuals (ϵ) are obtained from the GMM described in Bayless and Abrahamson (2018b, hereafter, BA18). Rather than the traditionally used response spectrum, BA18 is developed for the median and variance of the smoothed effective amplitude spectrum (EAS), as defined by PEER (PEER, 2015). The EAS is the orientation-independent horizontal-component FAS of ground acceleration, described in the following sections.

Paper Organization

In this article, we briefly review the correlation of epsilon, describe the ground-motion IM used, explain the reasoning behind selecting Fourier amplitudes as the IM, summarize the development of the correlation model and the

sensitivity of the correlation to database subsets, and compare the model with other published models.

Review of the Correlation of ϵ

The correlation coefficient of two random variables is a measure of their linear dependence. In this case, ϵ calculated from a large set of ground motions at different frequencies (f) are random variates. The correlation coefficient between $\epsilon(f_1)$ and $\epsilon(f_2)$ can be estimated using a maximum-likelihood estimator, the Pearson-product-moment correlation coefficient ρ (Fisher, 1958). The correlation coefficient for a sample of ϵ at frequencies f_1 and f_2 is given by the following equation:

$$\rho_{\epsilon(f_1),\epsilon(f_2)} = \frac{\mathrm{cov}(\epsilon(f_1),\epsilon(f_2))}{\sigma_{\epsilon(f_1)}\sigma_{\epsilon(f_2)}}$$
$$= \frac{\sum_{i=1}^{n}(\epsilon_i(f_1)-\overline{\epsilon(f_1)})(\epsilon_i(f_2)-\overline{\epsilon(f_2)})}{\sqrt{\sum_{i=1}^{n}(\epsilon_i(f_1)-\overline{\epsilon(f_1)})^2}\sqrt{\sum_{i=1}^{n}(\epsilon_i(f_2)-\overline{\epsilon(f_2)})^2}}$$

$$(3)$$

in which cov is the covariance, σ is the standard deviation, n is the total number of observations, i is the ith observation of ϵ, and $\overline{\epsilon(f_1)}$ and $\overline{\epsilon(f_2)}$ are the sample means of ϵ at frequencies f_1 and f_2, respectively. In our applications, $\bar{\epsilon}$ is equal to zero, indicating that the GMM is unbiased. The relation for $\rho_{\epsilon(f_1),\epsilon(f_2)}$ given in equation (3) is reciprocal; the correlation coefficient between two given frequencies is the same regardless of which frequency is the conditioning frequency.

The total residuals are correlated for a single earthquake, and this effect is removed by separating the residual components. To account for all residual terms, the total correlation is calculated as the following equation:

$$\rho_{\mathrm{total}}(f_1,f_2) = \frac{\rho_B(f_1,f_2)\tau(f_1)\tau(f_2) + \rho_{S2S}(f_1,f_2)\phi_{S2S}(f_1)\phi_{S2S}(f_2) + \rho_{WS}(f_1,f_2)\phi_{ss}(f_1)\phi_{ss}(f_2)}{\sigma(f_1)\sigma(f_2)},\qquad(4)$$

in which $\rho_B(f_1, f_2)$ is the correlation of the normalized between-event residuals, $\rho_{S2S}(f_1, f_2)$ is the correlation of the normalized site-to-site residuals, $\rho_{WS}(f_1, f_2)$ is the correlation of the normalized single-station within-event residuals, and σ is the total standard deviation. Confidence bounds on ρ are based on a variance stabilizing transformation of ρ, given in equation (5) (Kutner *et al.*, 2005). The variance z is given by equation (6), assuming that n is large enough so that z has an approximately normal distribution. The convergence of z to a normal distribution is very rapid as n increases (Bradley, 2011):

$$z = \tanh^{-1}(\rho) = \tfrac{1}{2}\ln\left(\tfrac{1+\rho}{1-\rho}\right) \qquad (5)$$

$$\mathrm{Var}(z) = \tfrac{1}{n-3}. \qquad (6)$$

Using a database of partitioned residuals, the calculation of $\rho_{\epsilon(f_1),\epsilon(f_2)}$ can be repeated for every frequency pair of interest. Figure 1 shows a graphical representation of this step at three example frequency pairs. The resulting correlation coefficients for each pair of frequencies can be saved as tables (e.g., Al Atik, 2011; Jayaram *et al.*, 2011; Abrahamson *et al.*, 2014; Akkar *et al.*, 2014; Azarbakht *et al.*, 2014) or can be empirically modeled. For modern GMMs, models of the correlation of ϵ are commonly created for the acceleration response spectrum or pseudospectral acceleration (PSA; e.g., Baker and Cornell, 2006; Baker and Jayaram, 2008; Goda and Atkinson, 2009; Cimellaro, 2013; Abrahamson *et al.*, 2014; Baker and Bradley, 2017). Recently, Stafford (2017) developed a correlation model for ϵ from FAS. This model and the development methodology are summarized and compared with our model in the Model Comparison section.

Motivation

The parameter ϵ is an indicator of the peaks and troughs at a given frequency in a spectrum, and ρ_ϵ characterizes the relative width of these extrema. As described in Bayless and Abrahamson (2018a), the width of peaks and troughs in ground-motion spectra has significance in risk assessments involving simulated ground motions because the variability in the dynamic structural response can be underestimated if the correlation in simulated ground motions is too low. Therefore, models for ρ_ϵ based on Fourier amplitudes can be used to evaluate and calibrate physics-based simulation methods. Correlation models based on FAS are preferable to response spectrum models for calibrating simulations because the FAS is a much more straightforward representation of the ground motion and is better understood by seismologists. In addition to calibrating the simulations, models for ρ_ϵ based on FAS can be combined with FAS-based ground-motion prediction models (e.g., Bora *et al.*, 2015, 2018; Bayless and Abrahamson, 2018b) to generate conditional

mean spectra for Fourier spectra or to conduct vector-valued probabilistic seismic hazard analyses for Fourier spectra (Abrahamson, 2006).

EAS Ground-Motion IM

The EAS is defined in Goulet *et al.* (2018) and is calculated for an orthogonal pair of FAS using the following equation:

$$\mathrm{EAS}(f) = \sqrt{\tfrac{1}{2}[\mathrm{FAS}_{\mathrm{HC1}}(f)^2 + \mathrm{FAS}_{\mathrm{HC2}}(f)^2]}, \qquad (7)$$

in which $\mathrm{FAS}_{\mathrm{HC1}}$ and $\mathrm{FAS}_{\mathrm{HC1}}$ are the FAS of the two orthogonal horizontal components of a three-component time series, and f is the frequency in hertz. The EAS is independent of the orientation of the instrument. Using the average power of the two horizontal components leads to an amplitude spectrum that is compatible with the use of random vibration theory (RVT) to convert Fourier spectra to response spectra. The EAS is smoothed using the $\log_{10}$-scale Konno and Ohmachi (1998) smoothing window, which has weights and window parameter defined as follows:

$$W(f) = \left(\tfrac{\sin(b\log(f/f_c))}{b\log(f/f_c)}\right)^4 \qquad (8)$$

$$b = \tfrac{2\pi}{b_w}. \qquad (9)$$

The smoothing parameters (W, f_c, b, b_w) are described by Kottke *et al.* (2018):

W is the weight defined at frequency f for a window centered at frequency f_c and defined by the window parameter b. The window parameter b can be defined in terms of the bandwidth, in $\log_{10}$ units, of the smoothing window, b_w.

The Konno and Ohmachi smoothing window was selected by PEER NGA-East (PEER, 2015) because it led to minimal bias on the amplitudes of the smoothed EAS compared with the unsmoothed EAS. The bandwidth of the smoothing window $b_w = 0.0333$ was selected such that the RVT calibration properties before and after smoothing were minimally affected (Kottke *et al.*, 2018). The smoothing of the EAS has a direct impact on ρ_ϵ. Using the smoothed EAS with the same smoothing bandwidth, we maintain consistency with the PEER database and with other PEER projects, including the NGA-East empirical FAS models (Goulet *et al.*, 2018) and the BA18 EAS model. The EAS models are processed by PEER following the procedure given by Kishida *et al.* (2016).

The correlation model presented in this article is based on the BA18 residuals and variance; therefore, the correlation model is for the interfrequency correlation of epsilon for the smoothed EAS ($\rho_{\epsilon,\mathrm{EAS}}$). For notational brevity, the
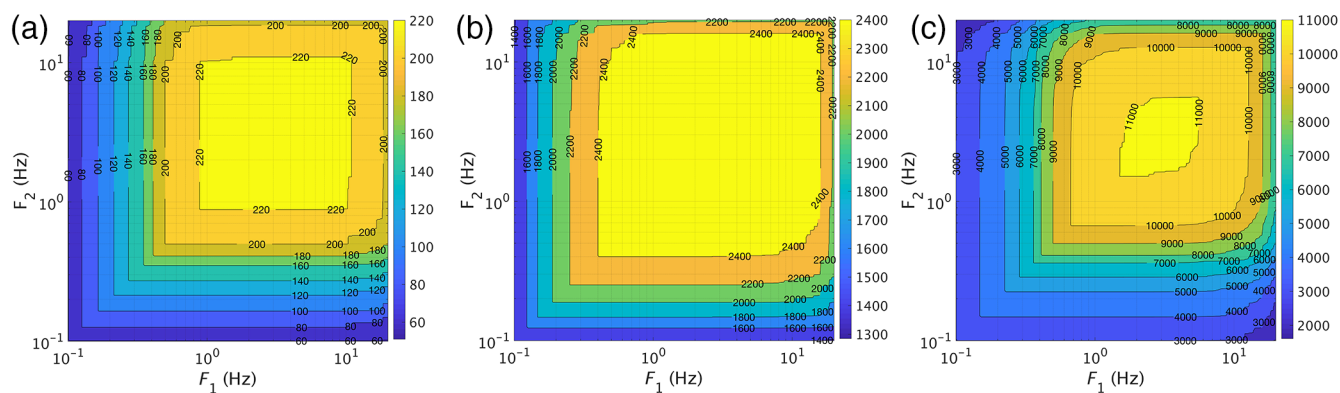
Figure 2. Contours showing the number of data points used in the analysis at each pair of frequencies. (a) The number of unique earthquakes for the between-event component, (b) the number of unique sites for the between-site component, and (c) the number of recordings for the within-site component. The color version of this figure is available only in the electronic edition.

EAS subscript is dropped hereafter and is implied unless noted otherwise. Similarly, if not stated explicitly, the term "interfrequency" is implied in all uses of the word "correlation" in this article because this is the only type of correlation we evaluate.

On the Selection of Fourier Amplitudes

In seismic hazard and earthquake engineering applications, the PSA of a 5% damped single degree of freedom oscillator (also referred to as an acceleration response spectrum) is a commonly used IM. PSA is useful for many applications; however, it has drawbacks. The FAS is a more direct representation of the frequency content of the ground motions than PSA and is better understood by seismologists. This leads to several advantages, both in the empirical modeling and in forward application.

BA18 illustrates that oscillators with different natural frequencies are controlled by different frequency ranges of the ground motion. At relatively higher oscillator frequencies, where there is little energy left to resonate the oscillator, the PSA ordinates are dominated by a wide-frequency band of the ground motion that ultimately equals the integration over the entire spectrum of the input ground motion (Bora *et al.*, 2016). The short-period PSA is then controlled by the dominant period of the input ground motion rather than the natural period of the oscillator. Therefore, as recognized by Carlton and Abrahamson (2014), at periods smaller than the period corresponding to the peak in the spectrum (T_p), the ϵ_{PSA} values will be more correlated with the ϵ_{PSA} values of T_p than for other periods with similar spacing. This effect can be observed as the reversal and increase in the Baker and Jayaram (2008) PSA correlation coefficients at short periods, which is discussed further in the Model Comparison section.

In summary, PSA provides the spectrum of peak response from a single degree of freedom system, which is influenced by a range of frequencies, and the breadth of that range is dependent on the oscillator period. The FAS provides a more direct representation of the frequency content

of the ground motions, and because the Fourier transform is a linear operation, the FAS is a more straightforward representation of the ground motion. As described previously, using FAS more easily facilitates future calibration of the interfrequency correlation of ground-motion simulation methods (e.g., Bayless and Abrahamson, 2018a) because it is better understood by seismologists and because the FAS correlation models are less complex.

Interfrequency Correlation Model

The subset of the NGA-West2 ground-motion database used to develop the model is described in BA18; the data used are dominated by California earthquakes but take advantage of crustal earthquake data worldwide to constrain the magnitude scaling and geometric spreading. In addition, a broader subset of data is used for testing regional variations of the correlation, as described further in the Dependence of the Correlation on Data Subsets section. We use the partitioned EAS residuals over the empirical frequency range of the BA18 model (0.1–24 Hz). The database accounts for the usable frequency range limitations of each record by applying recommended lowest and highest usable frequencies for response spectra determined from Abrahamson and Silva (1997). This usable frequency range accounts for the high-pass and low-pass filter frequencies of the two horizontal components, scaled by a factor of 1.25 to ensure that the filters do not have a significant effect on the response spectral values, as described in BA18. Retaining this usable frequency range maintains consistency with the response spectrum calculations. For each frequency pair, the records are only used in the correlation calculation if both frequencies fall within the usable range. The contour plot shown in Figure 2 displays the amount of earthquakes, sites, and records (i.e., ϵs) used at each pair of regression frequencies (F_1 and F_2).

To begin, we calculate the correlation coefficient ρ_ϵ for each of the normalized residual components (ϵ_B, ϵ_{S2S}, and ϵ_{WS}) at each pair of modeled frequencies. We calculate the total correlation (equation 4) using ρ_ϵ of each residual
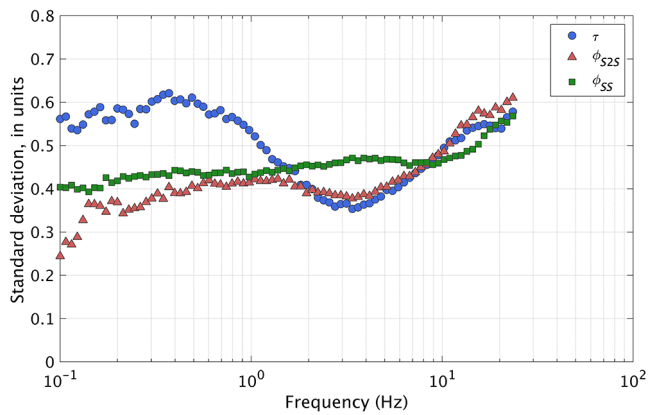
Figure 3. Standard deviation components of the Bayless and Abrahamson (2018b) effective amplitude spectrum ground-motion model (GMM). The color version of this figure is available only in the electronic edition.

comparable. As a result, the between-event correlation contributes significantly to the total correlation (equation 4). This is different from response spectra, in which the within-event standard deviation is often significantly larger than the between-event standard deviation, so the total correlations mostly reflect the within-event correlations (Stafford, 2017). The resulting correlations are presented as contours in Figure 4, in which F_1 and F_2 are any pair of frequencies. These figures are symmetric about the 1:1 line because correlation coefficient between two frequencies is the same regardless of which frequency is the conditioning frequency. The four correlation coefficient matrices shown in Figure 4 are provided in the Ⓔ supplemental content to this article.

To help visualize these results, Figure 5 deconstructs the ρ_ϵ contours from Figure 4 into five cross sections at conditioning frequencies: 0.2, 0.5, 2, 5, and 15 Hz. In this figure, solid lines are the ρ_ϵ cross sections, and dashed lines represent the 95% confidence interval of ρ_ϵ (Kutner *et al.*, 2005).

component and with the component standard deviations shown in Figure 3. Figure 3 shows that the between-event residual standard deviation (τ) is larger than the other two standard deviation components at frequencies below 1 Hz, and above 1 Hz, the values of all three components are

Between-Event Empirical Correlation

The between-event empirical ρ_ϵ cross sections are displayed in Figure 5a. Of the residual components, the
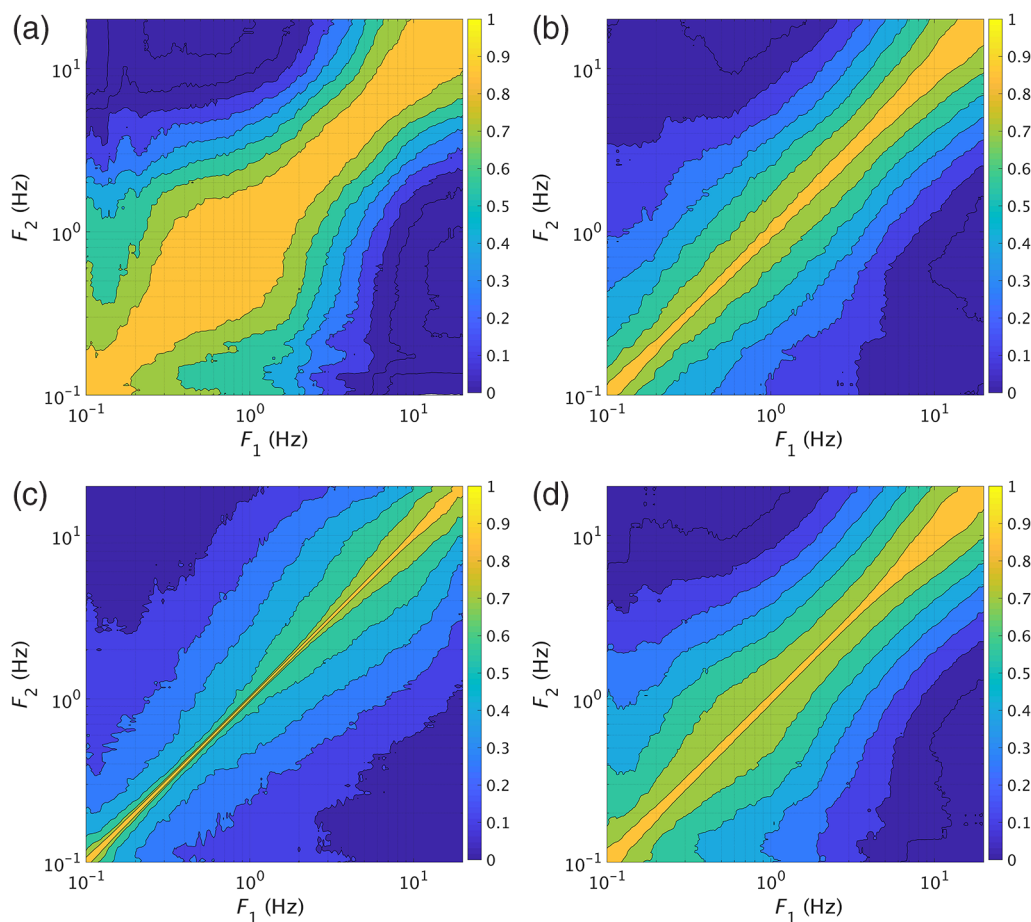

Figure 4. Empirical ρ_ϵ contours, showing (a) the between-event component, (b) the between-site component, (c) the within-site component, and (d) the total. The color version of this figure is available only in the electronic edition.
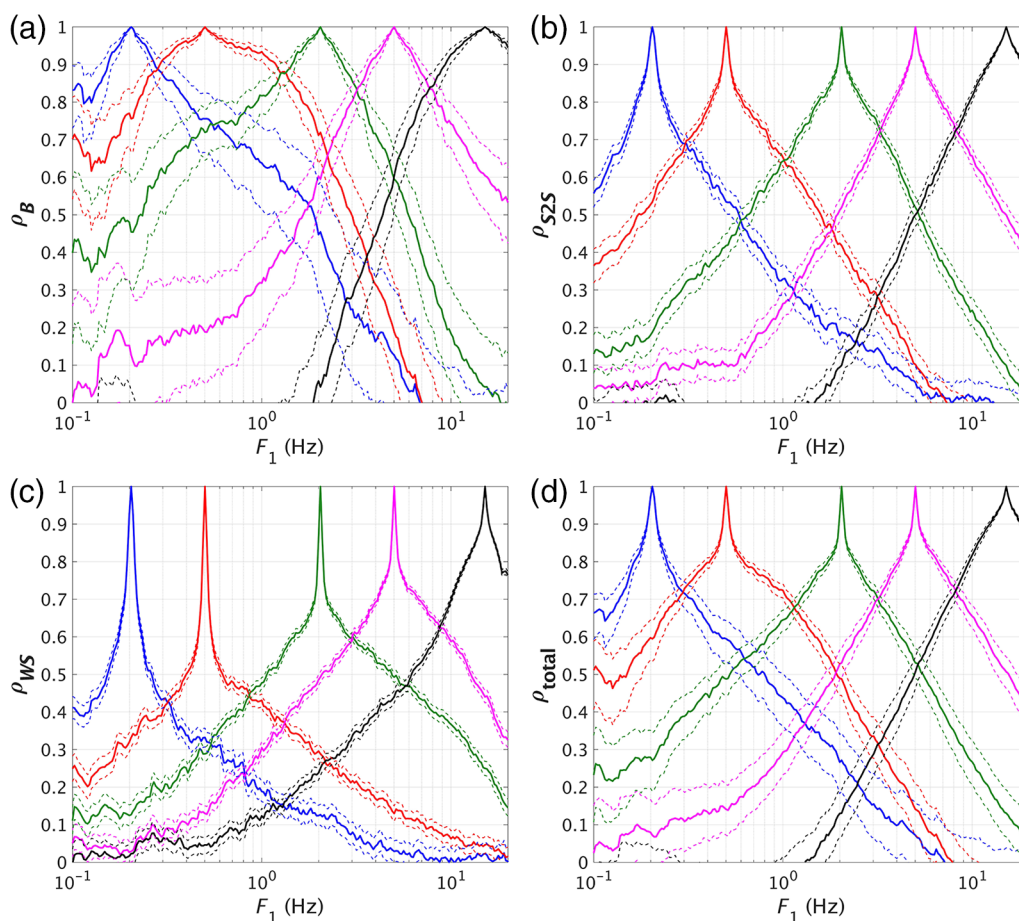
Figure 5. Empirical ρ_ϵ cross sections versus frequency at conditioning frequencies 0.2, 0.5, 2, 5, and 15 Hz (solid lines), with 95% confidence bounds on ρ (dashed lines), for (a) the between-event component, (b) the between-site component, (c) the within-site component, and (d) the total correlation. The color version of this figure is available only in the electronic edition.

confidence intervals on these correlation coefficients are the widest because there are the fewest samples of the between-event terms (earthquakes) for calculating ρ_ϵ. Figure 5a shows that the between-event ρ_ϵ contributes significantly to the $\rho_{\epsilon,\text{total}}$, and some frequency dependence exists. The between-event ρ_ϵ physically relates to source effects (e.g., stress drop), which drive ground motions over a broad frequency range and thus lead to relatively broad interfrequency ρ_ϵ.

Stafford (2017) observed minor magnitude dependence of the between-event empirical ρ_ϵ and attributed these to the variations in the source corner frequency for events of the same magnitude, concluding that larger magnitude events should exhibit stronger interfrequency correlations over a broader range of frequencies than smaller magnitude events. We chose not to model any magnitude dependence into our ρ_ϵ model; our reasoning is described in the Dependence of the Correlation on Data Subsets section.

Between-Site Empirical Correlation

The between-site residual represents the systematic deviation of the observed amplification at a site from the median amplification predicted by the model using a V_{S30}-

based site classification (Al Atik *et al.*, 2010). Therefore, the between-site ρ_ϵ represents the interfrequency correlation of the systematic site amplification deviations. The between-site empirical ρ_ϵ cross sections are displayed in Figure 5b. These correlations are generally not as strong as the between-event empirical ρ_ϵ but still contribute significantly to the total correlation. The shape of the ρ_ϵ cross sections does not vary strongly with conditioning frequency, especially below 5 Hz. At frequencies above 5 Hz, the ρ_ϵ cross sections broaden mildly.

Within-Site Empirical Correlation

The within-site residual component represents the remaining residual after partitioning the random effects for the event and the site. The within-site empirical ρ_ϵ cross sections are shown in Figure 5c. The confidence intervals on these correlation coefficients are close fitting because there are many samples of the within-site residuals for calculating ρ_ϵ. The within-site ρ_ϵ cross sections are characterized by a steep decay at frequencies very close to the conditioning frequency followed by a relatively flat slope at frequencies farther away from the conditioning frequency. In general, the
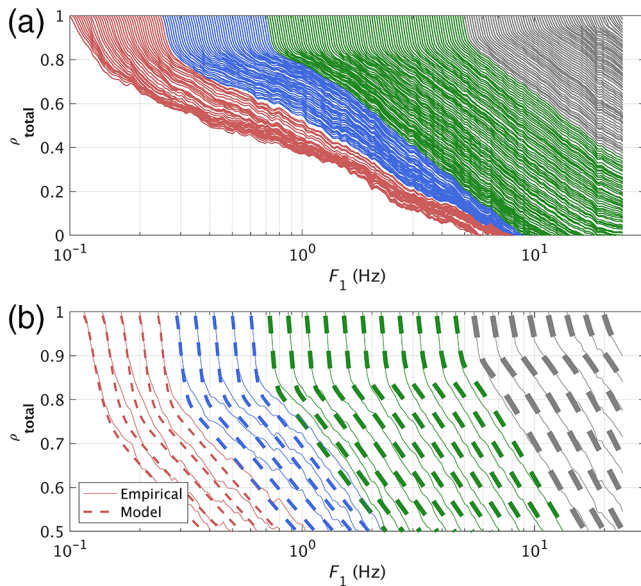
Figure 6. (a) Empirical total correlation coefficients for 239 conditioning frequencies. (b) A subset of the total empirical correlation coefficients (solid lines), along with the model (dashed lines). The color version of this figure is available only in the electronic edition.

strength of the within-site-component interfrequency correlations is substantially lower than the other residual components.

Total Correlation Model

The total interfrequency ρ_ϵ cross sections, calculated using equation (4), are shown in Figure 5d. Some frequency dependence is observed; if the correlations were independent of the conditioning frequency, all the contour lines in Figure 4d would be parallel. The contours are not parallel, indicating a minor frequency dependence of the interperiod correlations. For example, the 0.2 Hz correlation cross section in Figure 5d drops off more rapidly moving away from the conditioning frequency and has a different overall shape than the cross section conditioned at 15 Hz. The broader correlations at high frequencies are the result of the log-scale smoothing window used on the EAS. Initially, we modeled the correlations independently of the conditioning frequency with an exponential drop off in log-frequency space. The simplicity of this approach had a few advantages. First, it guarantees to produce a positive definite covariance matrix, which is a favorable feature for simulating realizations of ground motions. Second, a simple model was a good starting point for evaluating the correlation in ground-motion simulations. The frequency-independent model was a good fit to the empirical correlations on average but was generally too broad at low frequencies and too narrow at high frequencies; therefore, we decided to fit the total interfrequency ρ_ϵ with a slightly more complex frequency-dependent model. The frequency-dependent model allows for more robust evaluations

of the simulations and for future applications of the model to incorporate correlations as similar to the data as possible.

Figure 6 shows the total ρ_ϵ contours (Fig. 5d) in yet another manner; only the upper triangular part of the symmetric correlation matrix is plotted. Each line in Figure 6a is the empirical total correlation coefficients for one of 239 conditioning frequencies, indicated by the frequency with correlation value 1. Each of these correlation contours is fit independently to develop the correlation model. Figure 6b shows a subset of the total empirical correlation coefficients along with the empirical model. The model contours in Figure 6b are assigned different line weights subjectively to identify frequency ranges with significantly different shapes.

The total ρ_ϵ empirical model takes the form shown in the following equations:

$$\rho_{\epsilon,\text{total,Model}}(f_1,f_2) = \tanh[A(f_m)e^{B(f_m)\times f_r} + C(f_m)e^{D(f_m)\times f_r}] \tag{10}$$

$$f_r = \left| \ln\left(\frac{f_1}{f_2}\right) \right| \tag{11}$$

$$f_m = \min(f_1, f_2) \tag{12}$$

$$\text{If } f_1 = f_2, \rho_{\epsilon,\text{total,Model}}(f_1,f_2) = 1, \tag{13}$$

in which f_1 and f_1 are the two frequencies considered; tanh is the hyperbolic tangent; A, B, C, and D are frequency-dependent constants; f_r is the absolute value of the natural log-ratio of the two frequencies; and f_m is the minimum of the two frequencies. The model in equation (10) is a two-term exponential decay with the natural logarithm of frequency. Two exponential terms are required to model the shape of the correlation cross sections (e.g., Fig. 5d), which starts off with a steep decay at frequencies very close to the conditioning frequency and then flattens as the log ratio of frequencies increases. Equation (10) includes the hyperbolic tangent operator because the regression is performed on Fisher-transformed values of the correlation (equation 5), which results in approximately normally distributed variables z. This transformation emphasizes the fit to the higher correlation values, which we are most concerned about modeling accurately. The Fisher transformation is undefined for $\rho_\epsilon = 1$, so we force the correlation model to be unity when $f_1 = f_2$ (equation 13). The frequency dependence of coefficients A, B, C, and D is shown in Figure 7, and values are given in the Ⓔ supplemental content to this article.

The total ρ_ϵ empirical model contours and cross sections are shown in Figure 8. The empirical model compares favorably with the empirical correlations, especially at high correlation values, which are emphasized in the regression using the Fisher transformation (equation 5).
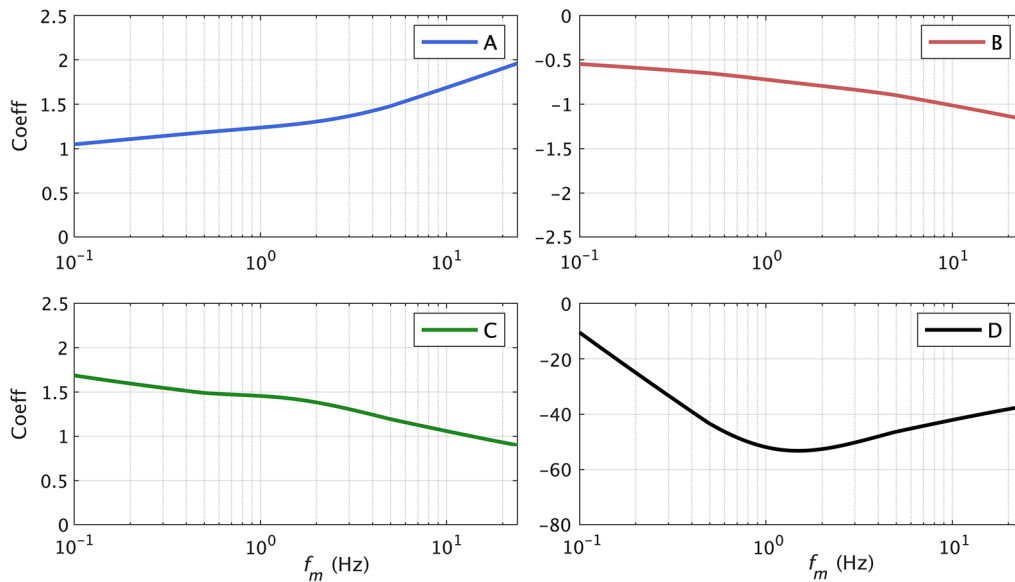
Figure 7. Frequency dependence of the $\rho_{\epsilon,\text{total}}$ empirical model coefficients A, B, C, and D. f_m is the minimum of the two frequencies considered (equation 10). The color version of this figure is available only in the electronic edition.
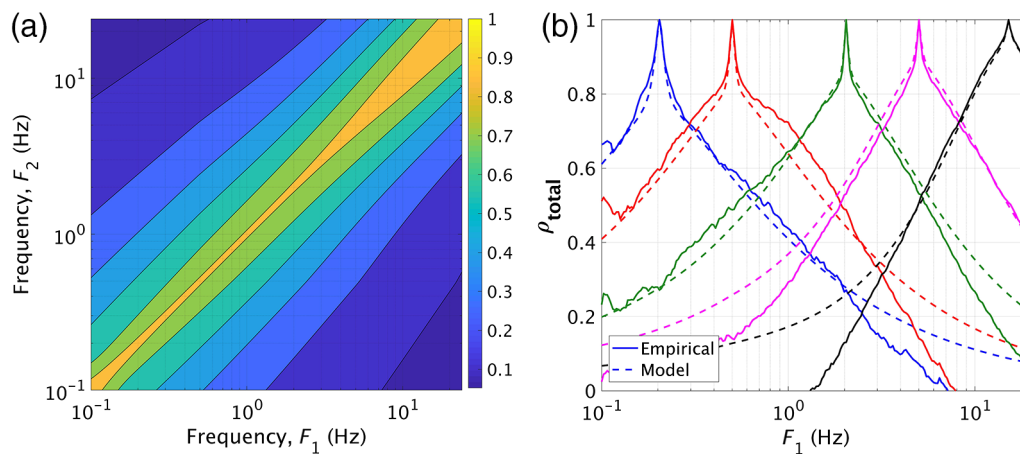

Figure 8. (a) Model $\rho_{\epsilon,\text{total}}$ contours; (b) model $\rho_{\epsilon,\text{total}}$ cross sections (dashed lines) compared with empirical cross sections (solid lines). The color version of this figure is available only in the electronic edition.

Range of Applicability

The total ρ_ϵ empirical model is developed using the PEER NGA-West2 EAS database (Ancheta *et al.*, 2014) and was not found to have strong magnitude, distance, site parameter, or regional dependence (discussed further later). Therefore, the model is applicable for crustal earthquakes in active tectonic regions worldwide. The model is applicable for rupture distances of 0–300 km, *M* 3.0–8.0, and over the frequency range 0.1–24 Hz. At frequencies outside this range, the model has not been tested. If extrapolation is required, we recommend using the values for coefficients A, B, C, and D at either $f = 0.1$ or 24 Hz for extrapolating to lower and higher frequencies, respectively. Tables for the total ρ_ϵ model coefficients and covariance matrices are provided in the Ⓔ supplemental content to this article.

Model Comparison

In this section, the model is compared with two other empirical models for ρ_ϵ.

Comparison with Baker and Jayaram (2008)

Baker and Jayaram (2008) developed an interperiod correlation model for within-event ϵ based on PSA using the NGA-West1 database (Chiou *et al.*, 2008). Using an updated database, Baker and Bradley (2017) confirmed that the updated correlations were largely consistent with the Baker and Jayaram (2008) model. In PSA GMMs, the within-site and between-site residuals are usually combined when within-event correlations are computed, and the within-event standard deviation is often significantly larger than the
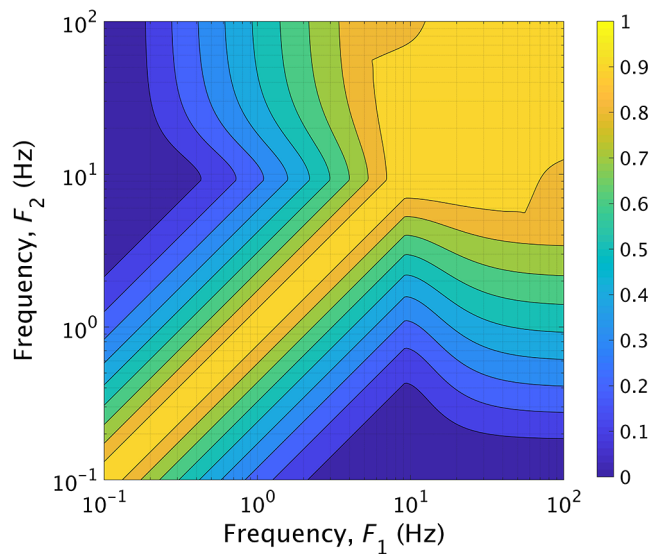
Figure 9. Baker and Jayaram (2008) pseudospectral acceleration correlation model contours, developed from the within-event residuals of Next Generation Attenuation-West1 GMMs. The color version of this figure is available only in the electronic edition.

between-event standard deviation, so the total correlations mostly reflect the within-event correlations (Stafford, 2017). The Baker and Jayaram (2008) contours derived from within-event ϵ for PSA are shown in Figure 9. An important difference between the contours derived from EAS (Fig. 4) and those from PSA is the behavior at high frequencies. The PSA contours in Figure 9 broaden substantially at high frequencies (short periods); this is because of the wide-frequency range that influences the short-period PSA, as discussed previously. The EAS contours do not behave this way because the Fourier transform operation at each frequency bin is independent of neighboring bins. For frequencies below about 10 Hz, the Baker and Jayaram (2008) model is independent of the conditioning frequency.

Comparison with Stafford (2017)

Stafford (2017; hereafter, S17) used a subset of the NGA-West1 database to develop models for the interfrequency ρ_ϵ and variance of FAS. S17 modeled the FAS using two approaches: first by adapting the Yenier and Atkinson (2015) FAS model to the data and second by performing a regression to the data with a simple GMM at each frequency independently. Like this study, S17 partitioned the residuals into between-event, between-site, and within-site components. S17 used unsmoothed FAS ordinates in the model development, which is an important distinction from the approach presented here and has an effect on the resulting models, as shown in the subsequent sections. In addition, the S17 model used both as-recorded horizontal components of the ground motions, as opposed to an orientation-independent horizontal component such as the EAS used here.

Figure 10 summarizes the S17 model for the total ρ_ϵ of the unsmoothed FAS over the frequency range 0.1–24 Hz, assuming a source corner frequency of 0.08 Hz. In Figure 10b, the S17 total ρ_ϵ model cross sections are compared with the total ρ_ϵ model developed here. The S17 total ρ_ϵ model features more frequency dependence and at high frequencies has a much stronger decay of the correlations in the vicinity of the conditioning frequency than the model developed here. The large differences in high-frequency ρ_ϵ models are likely the result of the different smoothing techniques used. The smoothing averages the EAS in log-spaced frequency bands, which increases the correlation between frequencies. As mentioned previously, the smoothing is done to maintain consistency with the PEER database and with models developed in other studies using the PEER database. At frequencies below 0.2 Hz, the S17 exponential decay near the conditioning frequency is slightly weaker than the model developed here, but the differences are small. These differences can be attributed to the combination of differences described previously: ground-motion component, database, smoothing technique, and GMM used for computing the residuals.
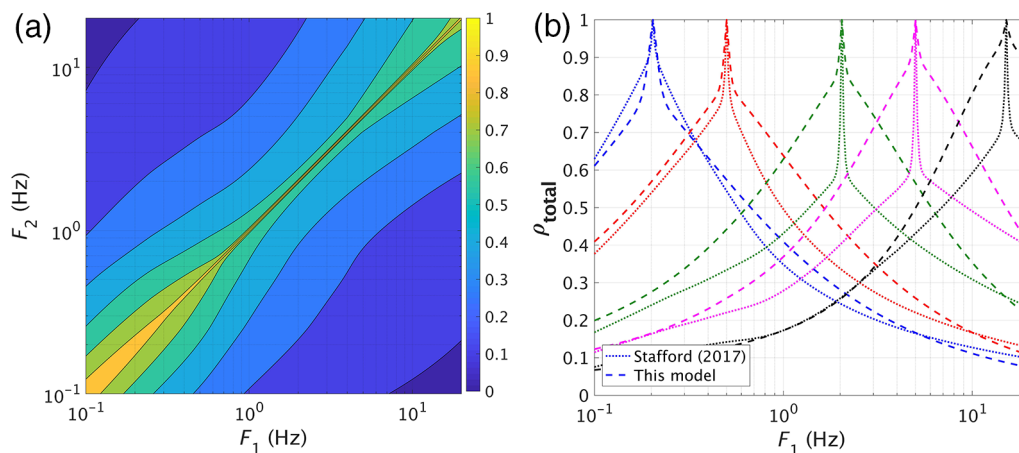

Figure 10. (a) Stafford (2017) $\rho_{\epsilon,\text{total}}$ contours using $f_c = 0.08$ Hz. (b) Comparison of the two $\rho_{\epsilon,\text{total}}$ model cross sections at five conditioning frequencies. The color version of this figure is available only in the electronic edition.

Table 1
Data Subsets Analyzed to Investigate ρ_ϵ Dependence on Seismological Parameters

Parameters	Bins
M	< 4.0, 4.0–5.0, 5.0–6.0, 6.0–7.0, > 7.0
R_{rup} (km)	0–15, 15–30, 30–50, 50–75, 75–100
V_{S30} (m/s)	< 300, 300–500, 500–700, > 700
Region	Western North America (WNA; primarily California), all non-WNA, Japan, Taiwan, China, Mediterranean

S17 observed minor magnitude dependence of the between-event ρ_ϵ and attributed these to the variations in the source corner frequency for events of the same magnitude, concluding that larger magnitude events should exhibit stronger interfrequency correlations over a broader range of frequencies than smaller magnitude events. We chose not to model any magnitude dependence into our total correlation model; our reasoning is described in the following section.

Dependence of the Correlation on Data Subsets

There have been conflicting conclusions published about the sensitivity of response spectra correlation coefficients to the ground-motion database subsets. Azarbakht *et al.* (2014), using PSA and the NGA-West1 database, concluded that the within-event correlation coefficients had meaningful dependencies on the causal magnitudes and distances of the recordings. This conclusion differs from those made by several other published studies, including Baker and Cornell (2006), Baker and Jayaram (2008), Carlton and Abrahamson (2014), and Baker and Bradley (2017). Baker and Bradley (2017) investigated the dependence of PSA interperiod correlations on binned data subsets of the PEER NGA-West2 database. They concluded that the correlations show no systematic trends with causal magnitude, distance, or V_{S30}. This was the same conclusion made by Baker and Jayaram (2008), which was developed using the NGA-West1 database. Carlton and Abrahamson (2014) concluded that the robustness of generic correlation models for PSA is a result of their dependence on spectral shape rather than tectonic region. Stafford (2017), working with unsmoothed FAS, found weak magnitude dependence on the between-event interperiod correlations, attributing these to the variations in the source corner frequency for events with the same magnitude. Stafford (2017) did not observe systematic dependence of the between-site or within-site residual correlations on causal magnitude or distance.

To investigate the dependence of the correlations on different seismological parameters, we recalculate the total ρ_ϵ of the EAS for the subsets of the data. The subsets are created by binning residuals based on magnitude, distance, V_{S30}, and earthquake region. The complete list of residual subsets analyzed is given in Table 1. For each data subset listed in Table 1, we calculate ρ_ϵ for each component of the residuals and review the ρ_ϵ contours and cross sections. As expected, deviations from the full database ρ_ϵ occur, but we do not find any systematic differences based on this qualitative assessment.

The dependence is also investigated more methodically by following the procedure taken by Baker and Bradley (2017). With this routine, the total ρ_ϵ from each data subset is calculated using the GMM residuals from that subset. The total ρ_ϵ for the subset is then compared with the ρ_ϵ from the full database. The results of this procedure are summarized in Figure 11, in which the total empirical ρ_ϵ for four frequency pairs are shown for the 20 subsets from Table 1. In Figure 11, the full database total ρ_ϵ for each frequency pair is shown with the solid, horizontal line, with dashed lines representing the lower and upper bounds for 95% confidence intervals of these coefficients (Kutner *et al.*, 2005). Solid circles are the total ρ_ϵ calculated for each indicated data subset, and the triangles indicate 95% confidence intervals of those coefficients. The 95% confidence intervals represent the statistical uncertainty in the correlation coefficients due to the finite number of samples and the standard deviation of the samples. If the confidence intervals of two groups do not overlap, then the differences in the correlation coefficients of the two groups are statistically significant at the 95% confidence level.

Figure 11a shows the magnitude-binned results, which reveal no systematic trends. The largest magnitude bin suffers from the smallest sample size, especially for the between-event terms, and has the largest variations from the full database ρ_ϵ. But at each frequency pair, the 95% confidence intervals for the binned data overlap with those for the full database, indicating that the two are not statistically significantly different for this bin. Figure 11b shows the distance-binned results, which also have overlapping confidence intervals for each frequency pair and bin, revealing no apparent dependence of ρ_ϵ on distance. Figure 11c shows the V_{S30}-binned results. The $V_{S30} > 700$ m/s bin has the largest deviations from the full database, but no systematic, statistically significant dependencies are observed. Figure 11d shows the results for the regional data subsets, in which deviations from the full database are stronger than any of the other data subsets examined. Figure 11c are from the California model residuals, but correlations in Figure 11d are from residuals for a larger subset of the full NGA-West2 database. The regional subsets have overlapping confidence intervals with the full database for each frequency pair except for the Japan subset coefficient at 1 and 4 Hz. The Western North America (WNA), all non-WNA, Japan, and Taiwan regions contain a substantial number of events and recordings in this analysis. The China and Mediterranean regions have the smallest sample sizes, as indicated by the wide confidence intervals, such that their deviations from the full database correlation coefficients are likely not significant, but this should be investigated in the future using more data for each region.

Based on our analysis of the data subsets, we do not detect any conclusive systematic relationships between ρ_ϵ and the seismological parameters reviewed. We do find that the largest differences in correlation coefficients occur at
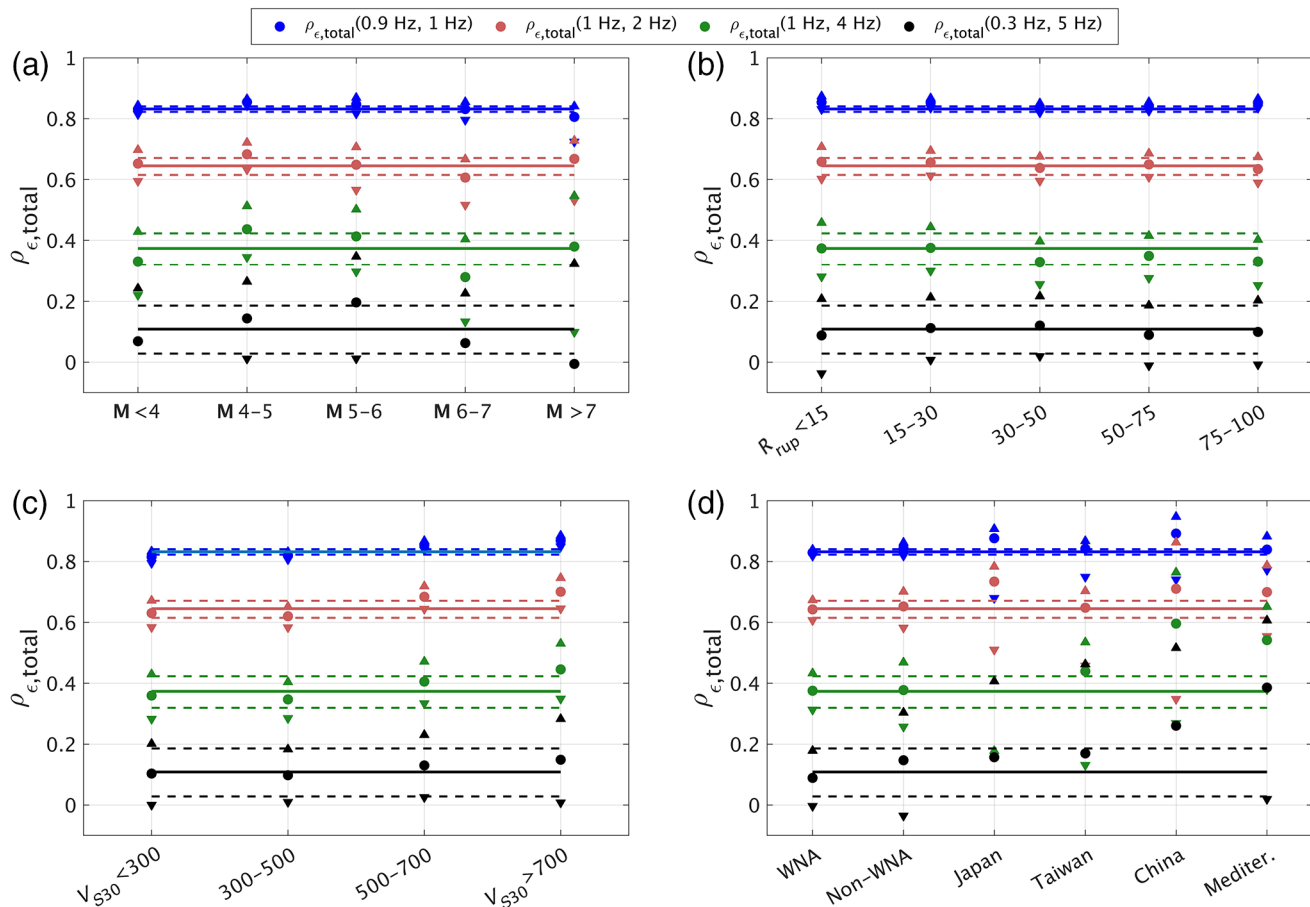
Figure 11. Total ρ_ϵ for four frequency pairs (identified in the legend) for data binned by (a) magnitude, (b) rupture distance, (c) site V_{S30}, and (d) earthquake region. The full database total ρ_ϵ for each frequency pair is shown with the solid horizontal line, and dashed lines represent the lower and upper bounds for 95% confidence intervals of these coefficients (Kutner *et al.*, 2005). Solid circles are the total ρ_ϵ calculated for each indicated data subset, and triangles indicate 95% confidence intervals of these coefficients. The color version of this figure is available only in the electronic edition.

widely spaced frequencies when ρ_ϵ values are low. This is an expected feature because of the heteroskedastic (nonconstant standard deviation) nature of the correlation coefficients. Correlation coefficients with values close to zero have a larger standard deviation than coefficients with values close to 1, meaning that the confidence intervals for low correlation coefficients are wider. This effect can be observed in Figure 11a, in which the 95% confidence intervals are tight for the 0.9 and 1 Hz pair coefficient and wide for the 0.3 and 5 Hz pair coefficient. As a result, differences between ρ_ϵ at low values are not usually significant. In addition, in practice, we care most about the frequency ranges with high correlations because these are related to the width of peaks and troughs in the spectra, and the wider frequency pairs with low correlations are not as impactful. Therefore, we conclude it is neither practical nor necessary to include dependencies on the reviewed seismological parameters in our interfrequency EAS correlation model. This conclusion agrees with the Baker series of conclusions for PSA (Baker and Cornell, 2006; Baker and Jayaram, 2008; Baker and Bradley, 2017) and with Carlton and Abrahamson (2014).

Correlation of Select Well-Recorded Events

Because no systematic relationships between ρ_ϵ and magnitude, distance, or site parameter were observed, the interfrequency correlation should approximately agree with the empirical model for a given event or set of events. To test this, we use the residuals from nine events identified by the Southern California Earthquake Center (SCEC) Broadband Platform (BBP) validation project (Dreger *et al.*, 2015) to calculate the interfrequency ρ_ϵ and compare with the empirical model. The SCEC BBP is a collaborative software development project, with the objective to integrate complex scientific codes for generating broadband ground-motion simulations for earthquakes. A key part of the SCEC process is to validate the simulations against data from well-recorded earthquakes, as described in the Dreger *et al.* (2015) validation exercise. The nine events from active crustal regions for validating the simulations against data are 2008 Chino Hills, 2007 Alum Rock, 1987 Whittier Narrows, 1986 North Palm Springs, 1994 Northridge, 1989 Loma Prieta, 1992 Landers, 2000 Tottori, and 2004 Niigata (Goulet *et al.*, 2015).
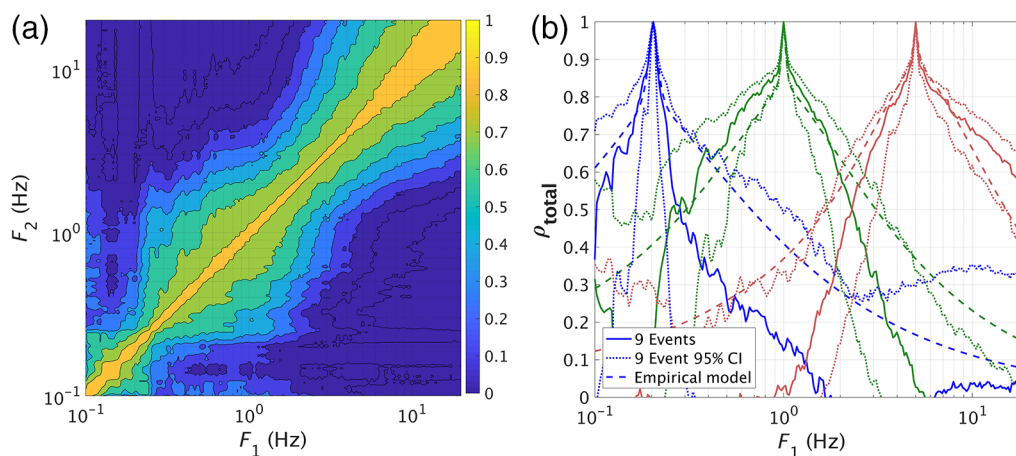
Figure 12. $\rho_{\epsilon,\text{total}}$ calculated from the nine Southern California Earthquake Center validation events (Goulet *et al.*, 2015). (a) Contours over the 1–24 Hz range. (b) Cross sections at three conditioning frequencies (solid lines), with 95% confidence intervals (CIs) for $\rho_{\epsilon,\text{total}}$ (dotted lines) compared with empirical model for $\rho_{\epsilon,\text{total}}$ (dashed lines). The color version of this figure is available only in the electronic edition.

Figure 12a shows the $\rho_{\epsilon,\text{total}}$ contours derived from residuals for these nine events, and Figure 12b compares the $\rho_{\epsilon,\text{total}}$ cross sections with the empirical model. Figure 12 supports the hypothesis that $\rho_{\epsilon,\text{total}}$ should approximately agree with the empirical model for a given event or set of events. In this case, departures from the model are observed, especially for the cross section conditioned at 0.2 Hz, but the 95% confidence bounds on $\rho_{\epsilon,\text{total}}$ indicate that differences are not statistically significant because these enclose the model over most frequencies.

Conclusions

The empirical model for the interfrequency correlation of the EAS developed in this study is applicable to shallow crustal earthquakes in active tectonic regions worldwide for rupture distances of 0–300 km, magnitudes of 3.0–8.0, and frequencies of 0.1–24 Hz. We do not find statistically significant magnitude, distance, site parameter, or regional dependence of the correlation, although potential regional variations should be studied further with more data from additional regions.

The ρ_{ϵ} empirical model developed here has several applications in practice. The correlation model can be used to define the interfrequency correlation in stochastic ground-motion simulation methods such as Boore (2003). Stafford (2017) and Bayless and Abrahamson (2018a) give examples of this procedure, using their respective correlation models, to modify the point-source stochastic simulation method to generate simulated acceleration time series with realistic ρ_{ϵ}. This correlation model is also appropriate for use in evaluation and validation studies of the interfrequency correlations from physics-based numerical simulations for ground motions from finite-fault earthquakes (Bayless and Abrahamson, 2018a). In such studies, the standard EAS approach for smoothing the

FAS needs to be used for the simulations to be consistent with the ρ_{ϵ} empirical model.

Data and Resources

Analyses and graphics production were performed using the numeric computing environment MATLAB (www .mathworks.com, last accessed January 2019). All ground-motion data are from the Pacific Earthquake Engineering Research Center (PEER) Next Generation Attenuation-West2 (NGA-West2) database (Ancheta *et al.*, 2014).

Acknowledgments

References

Abrahamson, N. A. (2006). Seismic hazard assessment: Problems with current practice and future developments, *Proceedings, First European Conf. on Earthquake Engineering and Seismology*, Geneva, Switzerland, 17 pp.

Abrahamson, N. A., and W. J. Silva (1997). Empirical response spectral attenuation relations for shallow crustal earthquakes, *Seismol. Res. Lett.* **68,** no. 1, 94–127.

Abrahamson, N. A., W. J. Silva, and R. Kamai (2014). Summary of the ASK14 ground motion relation for active crustal regions, *Earthq. Spectra* **30,** 1025–1055.

Akkar, S., M. A. Sandıkkaya, and B. Ö. Ay (2014). Compatible ground-motion prediction equations for damping scaling factors and vertical-to-horizontal spectral amplitude ratios for the broader Europe region, *Bull. Earthq. Eng.* **12,** 517–547, doi: 10.1007/s10518-013-9537-1.

Al Atik, L. (2011). Correlation of spectral acceleration values for subduction and crustal models, *COSMOS Technical Session*, Emeryville, California, 4 November 2011, 13 pp.

Al Atik, L., N. A. Abrahamson, F. Cotton, F. Scherbaum, J. J. Bommer, and N. Kuehn (2010). The variability of ground-motion prediction models and its components, *Seismol. Res. Lett.* **81,** no. 5, 794–801.

Ancheta, T. D., R. B. Darragh, J. P. Stewart, E. Seyhan, W. J. Silva, B. S.-J. Chiou, K. E. Wooddell, R. W. Graves, A. R. Kottke, D. M. Boore, *et al.* (2014). NGA-West2 database, *Earthq. Spectra* **30,** 989–1005.

Azarbakht, A., M. Mousavi, M. Nourizadeh, and M. Shahri (2014). Dependence of correlations between spectral accelerations at multiple periods on magnitude and distance, *Earthq. Eng. Struct. Dyn.* **43,** 1193–1204, doi: 10.1002/eqe.2393.

Baker, J. W., and B. A. Bradley (2017). Intensity measure correlations observed in the NGA-West2 database, and dependence of correlations on rupture and site parameters, *Earthq. Spectra* **33,** no. 1, 145–156.

Baker, J. W., and C. A. Cornell (2006). Correlation of response spectral values for multi-component ground motions, *Bull. Seismol. Soc. Am.* **96,** no. 1, 215–227.

Baker, J. W., and N. Jayaram (2008). Correlation of spectral acceleration values from NGA ground motion models, *Earthq. Spectra* **24,** 299–317, doi: 10.1193/1.2857544.

Bayless, J., and N. A. Abrahamson (2018a). Evaluation of the inter-period correlation of ground motion simulations, *Bull. Seismol. Soc. Am.* **108,** no. 6, 3413–3430, doi: 10.1785/0120180095.

Bayless, J., and N. A. Abrahamson (2018b). An empirical model for Fourier amplitude spectra using the NGA-West2 database, *PEER Report No. 2018/07*, Pacific Earthquake Engineering Research Center, University of California, Berkeley, California.

Boore, D. M. (2003). Simulation of ground motion using the stochastic method, *Pure Appl. Geophys.* **160,** 635–675.

Bora, S. S., F. Cotton, and F. Scherbaum (2018). NGA-West2 empirical Fourier and duration models to generate adjustable response spectra, *Earthq. Spectra* **35,** no. 1, 61–93.

Bora, S. S., F. Scherbaum, N. Kuehn, and P. J. Stafford (2016). On the relationship between Fourier and response spectra: Implications for the adjustment of empirical ground-motion prediction equations (GMPEs), *Bull. Seismol. Soc. Am.* **106,** no. 3, doi: 10.1785/0120150129.

Bora, S. S., F. Scherbaum, N. Kuehn, P. J. Stafford, and B. Edwards (2015). Development of a response spectral ground-motion prediction equation (GMPE) for seismic hazard analysis from empirical Fourier spectral and duration models, *Bull. Seismol. Soc. Am.* **105,** no. 4, 2192–2218.

Bradley, B. A. (2011). Empirical correlation of PGA, spectral accelerations and spectrum intensities from active shallow crustal earthquakes, *Earthq. Eng. Struct. Dyn.* **40,** 1707–1721, doi: 10.1002/eqe.1110.

Carlton, B., and N. A. Abrahamson (2014). Issues and approaches for implementing conditional mean spectra in practice, *Bull. Seismol. Soc. Am.* **104,** no. 1, 503–512. doi: 10.1785/0120130129.

Chiou, B. S.-J., R. B. Darragh, N. Gregor, and W. J. Silva (2008). NGA project strong-motion database, *Earthq. Spectra* **24,** 23–44.

Cimellaro, G. P. (2013). Correlation in spectral accelerations for earthquakes in Europe, *Earthq. Eng. Struct. Dyn.* **42,** 623–633, doi: 10.1002/eqe.2248.

Dreger, D. S., G. C. Beroza, S. M. Day, C. A. Goulet, T. H. Jordan, P. A. Spudich, and J. P. Stewart (2015). Validation of the SCEC Broadband Platform V14.3 simulation methods using pseudospectral acceleration data, *Seismol. Res. Lett.* **86,** no. 1, doi: 10.1785/0220140118.

Fisher, R. A. (1958). *Statistical Methods for Research Workers*, 13th Ed., Hafner, Edinburgh, London.

Goda, K., and G. M. Atkinson (2009). Probabilistic characterization of spatially correlated response spectra for earthquakes in Japan, *Bull. Seismol. Soc. Am.* **99,** 3003–3020, doi: 10.1785/0120090007.

Goulet, C. A., N. A. Abrahamson, P. G. Somerville, and K. E. Wooddell (2015). The SCEC Broadband Platform validation exercise: Methodology for code validation in the context of seismic-hazard analyses, *Seismol. Res. Lett.* **86,** no. 1, doi: 10.1785/0220140104.

Goulet, C. A., A. Kottke, D. M. Boore, Y. Bozorgnia, J. Hollenback, T. Kishida, A. Der Kiureghian, O. J. Ktenidou, N. M. Kuehn, E. M. Rathje, *et al.* (2018). Effective amplitude spectrum (EAS) as a metric for ground motion modeling using Fourier amplitudes, *2018 Seismological Society of America Annual Meeting*, Miami, Florida, 14–17 May 2018.

Jayaram, N., J. W. Baker, H. Okano, H. Ishida, M. W. McCann Jr., and Y. Mihara (2011). Correlation of response spectral values in Japanese ground motions, *Earthq. Struct.* **2,** no. 4, 357–376.

Kishida, T., O. Ktenidou, R. B. Darragh, and W. J. Silva (2016). Semi-automated procedure for windowing time series and computing Fourier amplitude spectra for the NGA-West2 database, *PEER Report No. 2016/02*, Pacific Earthquake Engineering Research Center, University of California, Berkeley, California.

Konno, K., and T. Ohmachi (1998). Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor, *Bull. Seismol. Soc. Am.* **88,** 228–241.

Kottke, A., E. Rathje, D. M. Boore, E. Thompson, J. Hollenback, N. Kuehn, C. A. Goulet, N. A. Abrahamson, Y. Bozorgnia, and A. Der Kiureghian (2018). Selection of random vibration procedures for the NGA east project, *PEER Report No. 2018/05*, Pacific Earthquake Engineering Research Center, University of California, Berkeley, California.

Kutner, M., C. Nachtsheim, J. Neter, and W. Li (2005). *Applied Linear Statistical Models*, McGraw-Hill/Irwin, New York, New York, 1396 pp.

Pacific Earthquake Engineering Research Center (PEER) (2015). NGA-East: Median ground-motion models for the central and eastern North America region, *PEER Report No. 2014/05*, Pacific Earthquake Engineering Research Center, University of California, Berkeley, California.

Stafford, P. J. (2017). Inter-frequency correlations among Fourier spectral ordinates and implications for stochastic ground-motion simulation, *Bull. Seismol. Soc. Am.* **107,** no. 6, 2774–2791.

Villani, M. A., and N. Abrahamson (2015). Repeatable site and path effects on the ground-motion sigma based on empirical data from southern California and simulated waveforms from the CyberShake platform, *Bull. Seismol. Soc. Am.* **105,** doi: 10.1785/0120140359.

Yenier, E., and G. M. Atkinson (2015). An equivalent point-source model for stochastic simulation of earthquake ground motions in California, *Bull. Seismol. Soc. Am.* **105,** no. 3, 1435–1455.

Department of Civil and Environmental Engineering
University of California, Davis
3021 Ghausi Hall, One Shields Avenue
Davis, California 95616 U.S.A.
jrbayless@ucdavis.edu