

A frequency-dependent ground-motion spatial correlation model of within-event residuals for Fourier amplitude spectra

Earthquake Spectra

1–25

© The Author(s) 2021

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/8755293020981995

journals.sagepub.com/home/eqs

Nan Wang, M.EERI^{1,2} , Kim B Olsen¹, and Steven M Day¹

Abstract

Ground motion time series recorded at stations separated by up to about 50 km show a frequency-dependent spatial coherency structure, and the corresponding ground motion intensity measures are found to be correlated. As omitting this correlation can result in underestimation of seismic losses in risk analysis, it is critical to quantify the spatial correlation structure for ground motion Fourier spectra estimated at different sites during a single event within a region. Toward this goal, we have developed an empirical frequency-dependent spatial correlation model for the within-event residuals of effective Fourier amplitude spectra from the Pacific Earthquake Engineering Research Center (PEER) Next Generation Attenuation (NGA) West2 database. The correlation model shows slower decrease of the spatial correlation with distance at lower frequencies compared with higher frequencies, in agreement with the underlying ground motion data, and no significant dependence on the magnitude of the earthquakes is observed. We use this empirical model to incorporate frequency-dependent spatial correlation into a hybrid deterministic-stochastic broadband ground motion generation module, which successfully generates synthetic time series for seven western US earthquakes with frequency-dependent spatial correlation that closely mimics that of the empirical model. Furthermore, the method also significantly improves the correlation for spectral accelerations, cumulative absolute velocities, and Arias intensities, compared with that derived from the original broadband module.

¹Department of Geological Sciences, San Diego State University, San Diego, CA, USA

²Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA

Corresponding author:

Nan Wang, Department of Geological Sciences, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA.

Email: naw021@ucsd.edu

Keywords

Spatial correlation, Fourier amplitude spectra, within-event residual, stochastic simulation, broadband ground motion simulation

Date received: 23 March 2020; accepted: 9 November 2020

Introduction

Ground motion time series recorded from earthquakes reveal a spatial coherency structure at stations separated by up to a few tens of kilometers, causing intensity measures such as peak ground velocities, peak ground accelerations, and peak spectral accelerations (i.e. response spectra) to be correlated (e.g. Abrahamson et al., 1991; Bolt et al., 1982; Bycroft, 1980; Der Kiureghian, 1996; Hao et al., 1989; Harichandran and Vanmarcke, 1986). A number of studies have been carried out over the past decades addressing the spatial correlation of ground motions (e.g. Boore et al., 2003; Esposito and Iervolino, 2011; Goda and Hong, 2008; Heresi and Miranda, 2019; Jayaram and Baker, 2009; Kawakami and Mogi, 2003; Loth and Baker, 2013; Markhvida et al., 2018; Sokolov et al., 2012; Wang and Takada, 2005; Wesson and Perkins, 2001). In general, these previous studies have investigated correlations between spectral accelerations for a range of periods, using earthquake records from different locations.

Seismic loss estimation in a region with exposed infrastructure is used by earthquake insurance companies to estimate expected damage in future catastrophes. The accuracy of the insured loss estimates in a region is critically dependent on the spatial correlation between the ground motion intensities at different sites during a single event, which can be significant at distances up to 50 km (e.g. Abrahamson et al., 1991; Bolt et al., 1982; Bycroft, 1980; Der Kiureghian, 1996; Hao et al., 1989; Harichandran and Vanmarcke, 1986). Strong motion data naturally include such correlation, but is often available in insufficient amounts for loss analysis, in particular for large events and close to the causative fault. Instead, numerical simulations can provide key information for seismic hazard analysis. Seismic hazard assessment has benefited from recent advances in simulation methods due to improved source characterization, accuracy of numerical methods, and availability of powerful computational resources. However, while ground motion simulations produced from complex 3D rupture and crustal models may include realistic spatial correlation structure (e.g. Withers et al., 2019), those obtained by more simplified deterministic simulations (e.g. 1D) and, in particular, by stochastic approaches (e.g. Atkinson et al., 2009; Beresnev and Atkinson, 1997; Boore, 2003, 2009; Motazedian and Atkinson, 2005), oftentimes do not. For example, many broadband simulation methods (e.g. Atkinson and Assatourians, 2015; Crempien and Archuleta, 2015; Graves and Pitarka, 2015; Olsen and Takedatsu, 2015), which have been tuned to produce good agreement with median spectral acceleration from strong motion data, have received less attention to their spatial correlation behavior. The importance of including spatial correlation in ground motion simulations has been illustrated by many studies (e.g. Jayaram and Baker, 2010; Miller and Baker, 2015) for loss estimates, clearly showing that simulations without spatial correlation can result in an underestimation of seismic risk.

Pseudo-spectral acceleration (PSA) has traditionally been the preferred metric in earthquake engineering, and many studies have proposed spatial correlation models for PSA. However, each PSA ordinate depends (nonlinearly) on ground motion amplitudes over a range of frequencies, and therefore a correlation model for PSA does not provide a direct

means to impose the correlation structure on numerically simulated time histories (or on other ground-motion metrics derived therefrom). On the contrary, the Fourier amplitude spectrum (FAS) provides a straightforward means to incorporate an empirical correlation model into simulated ground motion time histories, through frequency-domain multiplication, but its empirical estimation is complicated by the fact that its value depends upon recording-instrument orientation. The effective amplitude spectrum (EAS), defined in the next section, avoids this complication, and bears a simple relationship to the FAS. Therefore, the aim of this study is to (1) develop a new, empirical frequency-dependent spatial correlation model of EAS and (2) describe and demonstrate its implementation into numerically simulated ground motion. In that implementation, the empirical EAS correlation model is used to generate separate but correlated FAS adjustments to the two horizontal components at a given site. Specifically, we use the findings for inter-frequency correlation by Wang et al. (2019) to generate correlated horizontal-component FAS residuals with correlation coefficient 0.7.

Within-event residual of the EAS

FAS, the amplitude spectrum of Fourier transform of the acceleration time series, depends on the recording instrument's orientation. Such dependency may cause an undesirable bias in applications of the calculated FAS values. On the contrary, the EAS defined by Goulet et al. (2018) as follows:

$$EAS(f) = \sqrt{\frac{1}{2} [FAS_{HC1}^2(f) + FAS_{HC2}^2(f)]} \quad (1)$$

is rotation independent, and will therefore be used as the intensity measure for our empirical model development. As discussed later, we then use the EAS correlation model, with the method of Wang et al. (2019), to generate FAS adjustments to simulated time histories. In Equation 1, FAS_{HC1} and FAS_{HC2} are the FAS of two orthogonal as-recorded horizontal components acceleration time series, and f is the frequency in Hertz. The EAS is smoothed by the \log_{10} -scale Konno and Ohmachi (1998) smoothing window (e.g. Kottke et al., 2018):

$$W(f) = \left(\frac{\sin(b \log(f/f_c))}{b \log(f/f_c)} \right)^4. \quad (2)$$

Here, W is a weight at frequency f designed for a window with center frequency f_c , and $b = \frac{2\pi}{b_w} = 60\pi$, where b_w is the smoothing window bandwidth in \log_{10} units (see Kottke et al., 2018, for more details on the smoothing technique). Note that the smoothing of the EAS can have a direct impact on the correlation. A comparison of the models of Stafford (2017) and Bayless and Abrahamson (2019) indicates that it is possible that smoothing contributes to larger inter-frequency correlations. PEER NGA-East (Pacific Earthquake Engineering Research Center (PEER), 2015) selected the Konno and Ohmachi (1998) type of smoothing window which leads to minimal bias on the amplitudes of the smoothed EAS compared with the unsmoothed EAS. The parameter b (188.5 in our study) was selected such that the random vibration theory (RVT) calibration properties after the smoothing were minimally affected (Kottke et al., 2018). Using the smoothed EAS with the same b , we maintain consistency with the PEER database as well as with other PEER projects, including the NGA-East empirical FAS models (Goulet et al., 2018) and the Bayless and Abrahamson (2018a) EAS model.

Following the notation by Al Atik et al. (2010) we define y_{es} as the natural logarithm of the ground-motion intensity measure observed at station s during earthquake e as follows:

$$y_{es} = \mu_{es} + \delta B_e + \delta W_{es} \quad (3)$$

where μ_{es} is the mean prediction of the natural logarithm of the intensity measure, δB_e is the between-event (or inter-event) residual representing the average shift of the observed ground motion for an individual earthquake e from the population mean prediction, and δW_{es} is the within-event residual (or intra-event residual), depicting the misfit between an individual observation at station s from the earthquake-specific mean prediction. The between-event residual includes average source effects (over all azimuths) such as stress drop and spatial and temporal variation of fault slip that are not captured by magnitude, faulting style, and source depth in the mean prediction. The within-event residual comprises azimuthal variations in source, path, and site effects that reflect the influence of factors such as crustal heterogeneity, deeper geological structure, and near-surface layering that cannot be captured by a simple distance metric and a site-classification based on the average shear-wave velocity (Al Atik et al., 2010). These residuals are normally distributed with zero mean and are uncorrelated with each other. Following these definitions, the normalized EAS within-event residual, ε , at station s during earthquake e is calculated as a function of frequency f as follows:

$$\varepsilon(f) = \frac{\delta W_{es}(f)}{\varphi(f)} = \frac{\ln EAS_{es}(f) - \mu_{\ln EAS_{es}}(f) - \delta B_e}{\varphi(f)}, \quad (4)$$

where φ is the standard deviation of δW_{es} , and ε is standard normally distributed.

Semivariogram analysis

A semivariogram (γ) characterizes the strength of statistical dissimilarity as a function of distance and is often used to describe spatially distributed random variables in geostatistics (see Appendix 1 for a summary of semivariograms). Under the stationary and isotropic assumptions, the semivariograms are independent of the locations and offset direction of the site pair but depend on the distance between the sites. The empirical semivariogram matrix for ε at each frequency pair (f_i, f_j) can be summarized by an isotropic semivariogram matrix (Γ) as a function of separation distance h :

$$\Gamma(h) = \gamma_{f_i, f_j}(h) = \begin{bmatrix} \gamma_{f_1, f_1}(h) & \cdots & \gamma_{f_1, f_n}(h) \\ \vdots & \ddots & \vdots \\ \gamma_{f_n, f_1}(h) & \cdots & \gamma_{f_n, f_n}(h) \end{bmatrix}, \quad (5)$$

where matrix element γ_{f_i, f_j} is as defined, in terms of ε , by Equation 25 in Appendix 1. Similarly, the empirical isotropic covariance matrix (C) can be written as a function of separation distance h as follows:

$$C(h) = c_{f_i, f_j}(h) = \begin{bmatrix} c_{f_1, f_1}(h) & \cdots & c_{f_1, f_n}(h) \\ \vdots & \ddots & \vdots \\ c_{f_n, f_1}(h) & \cdots & c_{f_n, f_n}(h) \end{bmatrix}, \quad (6)$$

and we have

$$C(h) = C(0) - \Gamma(h). \quad (7)$$

Empirical frequency-dependent spatial correlation model of within-event residuals

Data sources

In this study, the frequency-dependent spatial correlation model is developed from EAS values for recorded ground motions in the PEER NGA West2 database (Ancheta et al., 2014). The NGA-West2 database includes shallow crustal earthquakes with $M > 3$ in active tectonic regions (dominated by California and Nevada earthquakes). The normalized within-event EAS residual, ε , was determined from the Bayless and Abrahamson (2018a) ground motion model. This model was calculated from the individual EAS values and the earthquake-specific smoothed EAS median value at each station for each recorded event. The within-event residuals used in this study are obtained from Bayless and Abrahamson (2018a). For more details on the ground-motion database and data selection criteria, see Bayless and Abrahamson (2018a) and Abrahamson et al. (2014).

Semivariograms γ_{f_i, f_j} of ε were calculated for each pair of frequencies, f_i and f_j (at frequency points 0.1–1 Hz with a spacing of 0.1 Hz, and 1–23 Hz with a spacing of 1 Hz) and as a function of h from 0 to 120 km with a bin size equal to 2 km. Figure 1 shows an example of the semivariogram produced for the frequency pair $f_1 = f_2 = 1$ Hz computed from the data.

Linear model of coregionalization

Previous studies (e.g. Wang and Takada, 2005) have observed an exponential decay of the ground motion spatial correlation, suggesting that the semivariogram can be well fit using an exponential model. For this reason, we assume a functional form of the semivariogram with the general behavior:

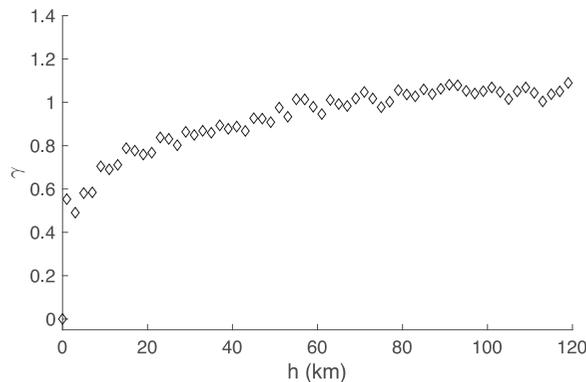


Figure 1. Semivariogram as a function of h at the frequency pair $f_1 = f_2 = 1$ Hz computed from the dataset.

$$\gamma(h) = S \left[1 - \exp\left(-\frac{3h}{R}\right) \right], \quad (8)$$

where S is the sill that represents the asymptotic value of $\gamma(h)$ as h goes to infinity, and R is the range that represents the distance at which the value of $\gamma(h)$ equals 95% of the sill. In the multivariate case (as in Equation 25 in Appendix 1), this general behavior would imply a semivariogram at a given frequency pair f_i and f_j of the form:

$$\gamma_{ij}(h) = S_{ij} \left[1 - \exp\left(-\frac{3h}{R_{ij}}\right) \right]. \quad (9)$$

However, it has been shown that (for within-event residuals of spectral accelerations) R_{ij} varies at different frequencies such that lower frequencies tend to have larger ranges than do higher frequencies (Loth and Baker, 2013). To better represent these frequency dependencies, we followed an approach similar to that of Loth and Baker (2013), using a nested semivariogram model (a linear combination of single semivariogram models):

$$\gamma_{ij}(h) = P_{ij}^1 \left(1 - \exp\left(-\frac{3h}{R_1}\right) \right) + P_{ij}^2 \left(1 - \exp\left(-\frac{3h}{R_2}\right) \right) + P_{ij}^3. \quad (10)$$

Combining all elements γ_{ij} , we obtain the linear model of coregionalization:

$$\Gamma(h) = \mathbf{P}^1 \left(1 - \exp\left(-\frac{3h}{R_1}\right) \right) + \mathbf{P}^2 \left(1 - \exp\left(-\frac{3h}{R_2}\right) \right) + \mathbf{P}^3, \quad (11)$$

where \mathbf{P}^1 and \mathbf{P}^2 are coregionalization matrices corresponding to the short-range and long-range models, respectively. Note that the third term, the coregionalization matrix \mathbf{P}^3 in Equation 11 corresponds to the nugget effect as follows:

$$\gamma(h) = \begin{cases} 0 & \text{if } h=0 \\ S & \text{if } h>0 \end{cases}, \quad (12)$$

which can be used to represent discontinuity of the semivariogram at separation distances larger than zero. Ranges $R_1 = 10\text{km}$ and $R_2 = 100\text{km}$ provide a reasonable fit to the data and are adopted in our model. The coregionalization matrices, \mathbf{P}^1 , \mathbf{P}^2 and \mathbf{P}^3 , which are symmetric and semipositive definite, are estimated from the empirical semivariogram data by the procedure given in the next section.

Empirical frequency-dependent spatial correlation model for covariance

We use the Goulard–Voltz algorithm (Goulard and Voltz, 1992) to develop our frequency-dependent spatial correlation model for covariance. The iterative algorithm, commonly used to fit a linear model of coregionalization with semipositive definite coregionalization matrices, uses a least square fitting technique to find the coregionalization matrices that minimize the weighted sum of squares as follows:

$$WSS = \sum_{k=1}^K \omega_k \|\hat{\Gamma}(h_k) - \Gamma(h_k)\|^2 = \sum_{k=1}^K \omega_k \sum_{i,j=1}^N [\hat{\gamma}_{ij}(h_k) - \gamma_{ij}(h_k)]^2, \quad (13)$$

where $\hat{\Gamma}(h_k)$ and $\hat{\gamma}_{ij}(h_k)$ represent the semivariogram values computed from the model, and $\Gamma(h_k)$ or $\gamma_{ij}(h_k)$ represent the semivariogram values computed from the empirical data at h_k , the center of the k th bin. ω_k is a positive weight at h_k , which is defined as $\omega_k = \frac{1}{h_k}$ in this study.

Let us denote $\left(1 - \exp\left(-\frac{3h}{R_1}\right)\right)$ by $g^1(h)$, $\left(1 - \exp\left(-\frac{3h}{R_2}\right)\right)$ by $g^2(h)$ and 1 by $g^3(h)$. Equation 11 can then be written as follows:

$$\Gamma(h) = \sum_{l=1}^L \mathbf{P}^l g^l(h), L=3. \quad (14)$$

The Goulard–Voltz algorithm is now executed using the following steps:

1. Initialize the coregionalization matrices \mathbf{P}^l , $l=1, 2, 3$ in this study.
2. Iterate from (a) to (c):
 - (a) Compute WSS with the current coregionalization matrices.
 - (b) For each l :
 - (b1) Compute the new coregionalization matrix $\tilde{\mathbf{P}}^l$ as follows:

$$\tilde{\mathbf{P}}^l = \frac{\sum_{k=1}^K \omega_k g^l(h_k) \left[\hat{\Gamma}(h_k) - \sum_{u=1, u \neq l}^L \mathbf{P}^u g^u(h_k) \right]}{\sum_{k=1}^K \omega_k [g^l(h_k)]^2} \quad (15)$$

- (b2) Decompose $\tilde{\mathbf{P}}^l$ as $\tilde{\mathbf{P}}^l = \mathbf{Q}_l \mathbf{\Lambda}_l \mathbf{Q}_l^T$ where $\mathbf{Q}_l \mathbf{Q}_l^T$ is an identity matrix and $\mathbf{\Lambda}_l$ is a diagonal matrix.
 - (b3) Change all the negative diagonal elements of $\mathbf{\Lambda}_l$ to zero to obtain $\mathbf{\Lambda}_l^+$ (this step is applied for ensuring semipositive definiteness of each coregionalization matrix).
 - (b4) Update $\tilde{\mathbf{P}}^l$ as a semipositive definite matrix $\tilde{\mathbf{P}}^l = \mathbf{Q}_l \mathbf{\Lambda}_l^+ \mathbf{Q}_l^T$.
 - (c) Compute WSS with the updated coregionalization matrices and loop over (a) to (c) until the difference of the WSS value from (a) and (c) is smaller than a positive prespecified value.

The empirical semivariogram and the estimated multivariate semivariogram model are shown in Figure 2, for example frequency pairs. Noting that

$$C(0) = \lim_{h \rightarrow +\infty} \Gamma(h) = \mathbf{P}^1 + \mathbf{P}^2 + \mathbf{P}^3, \quad (16)$$

the correlation matrix in Equation 6 can be derived from Equation 7 as

$$C(h) = \mathbf{P}^1 \exp\left(-\frac{3h}{R_1}\right) + \mathbf{P}^2 \exp\left(-\frac{3h}{R_2}\right) + \mathbf{P}^3 I_{\{h=0\}} \quad (17)$$

where

$$I_{\{h=0\}} = \begin{cases} 1 & h=0 \\ 0 & h \neq 0 \end{cases} \quad (18)$$

is the indicator function.

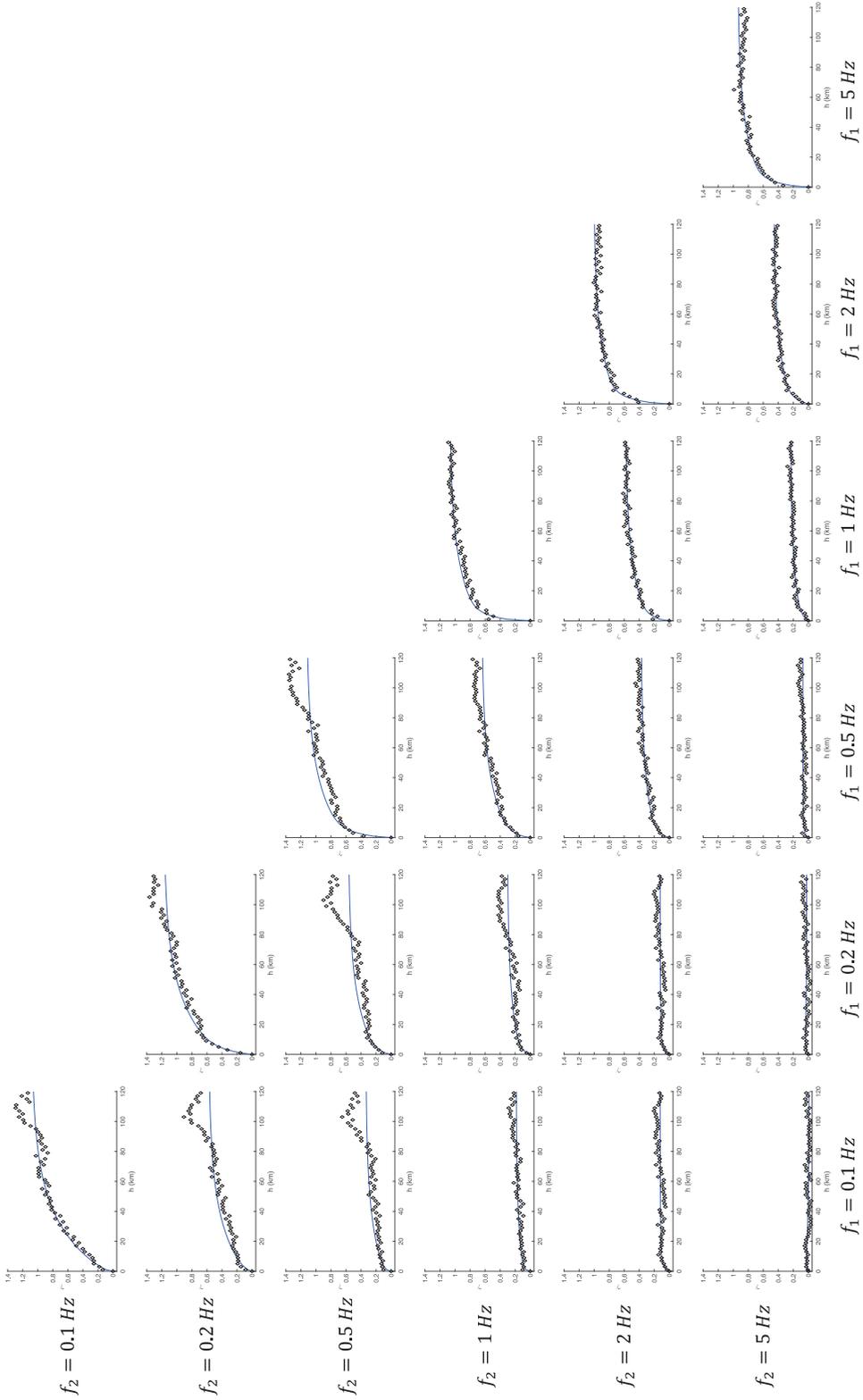


Figure 2. Empirical semivariograms (diamonds) and the fitted multivariate semivariogram model (solid lines) at frequency pairs for 0.1, 0.2, 0.5, 1, 2 and 5 Hz.

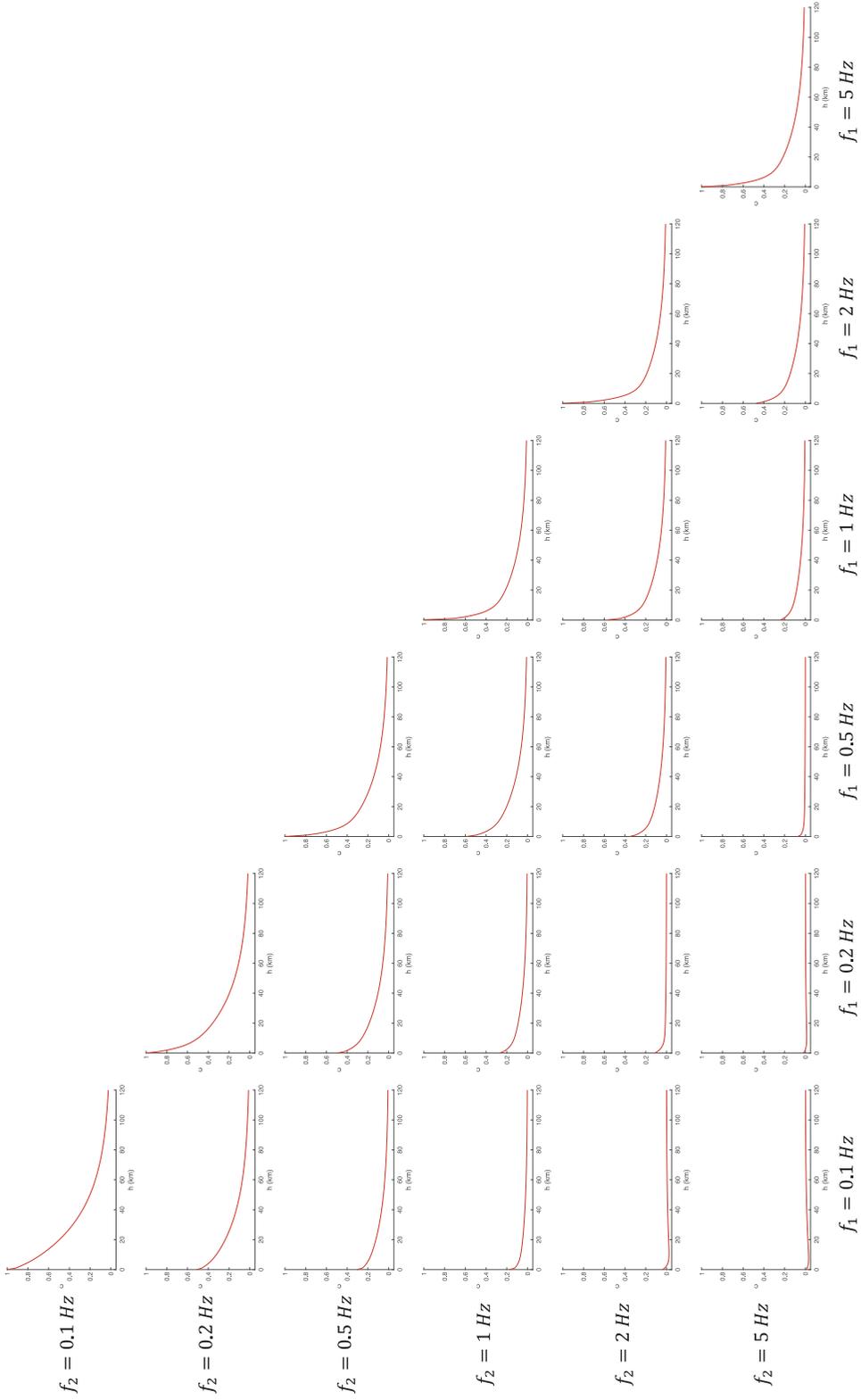


Figure 3. The resulting correlation coefficient model at frequency pairs for 0.1, 0.2, 0.5, 1, 2, and 5 Hz.

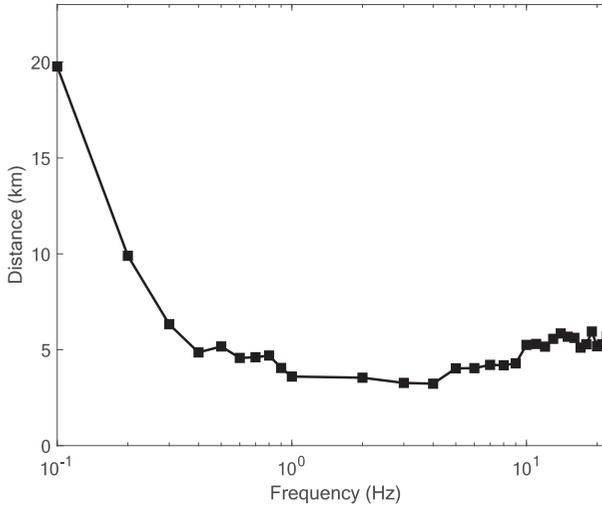


Figure 4. Separation distances where the correlation coefficient drops to 0.5 using pairs of two identical frequencies.

The resulting coregionalization matrices with each element corresponding to a pair of 32 frequency points 0.1–1 Hz with a spacing of 0.1 Hz and 1–23 Hz with a spacing of 1 Hz are provided in the Supplemental material of this article. Note that the coregionalization matrices after computation from the Goulard–VOLTZ algorithm are normalized as follows:

$$\frac{P_{ij}^h}{\sqrt{P_{ii}^1 + P_{ii}^2 + P_{ii}^3} + \sqrt{P_{jj}^1 + P_{jj}^2 + P_{jj}^3}} \quad (19)$$

The correlation coefficient model is shown in Figure 3, for example frequency pairs. Figure 4 shows the separation distances where the correlation coefficient of frequency pairs with two identical frequencies drop to 0.5. It can be seen that, as expected, the correlation at lower frequencies (<0.4 Hz) persists to larger separation distances, compared with higher frequencies. This is expected, as lower frequency signals are less affected by smaller-scale crustal features (e.g. topographic relief, velocity perturbations) that tend to control the variation of more high-frequency (HF) motion. The small increase (about 2 km) in the separation distances at frequencies larger than 4 Hz in Figure 4 is discussed in the “Discussion” section.

Note that when $h=0$ in Equation 17, the correlation model becomes $C(0) = P^1 + P^2 + P^3$, which represents the inter-frequency correlation at a single site. Hence, the presented frequency-dependent spatial correlation model also includes the inter-frequency correlation simultaneously. A comparison of the regressed within-event inter-frequency correlation model with the empirical within-event inter-frequency correlation from Bayless and Abrahamson (2019) is shown in Figure 5. The presented model compares well with the empirical inter-frequency correlation, especially at higher correlation values.

Inclusion of frequency-dependent spatial correlation into ground motion simulation

We demonstrate and validate our spatial correlation approach on the San Diego State University Broadband Ground Motion Generation Module (hereafter the “SDSU

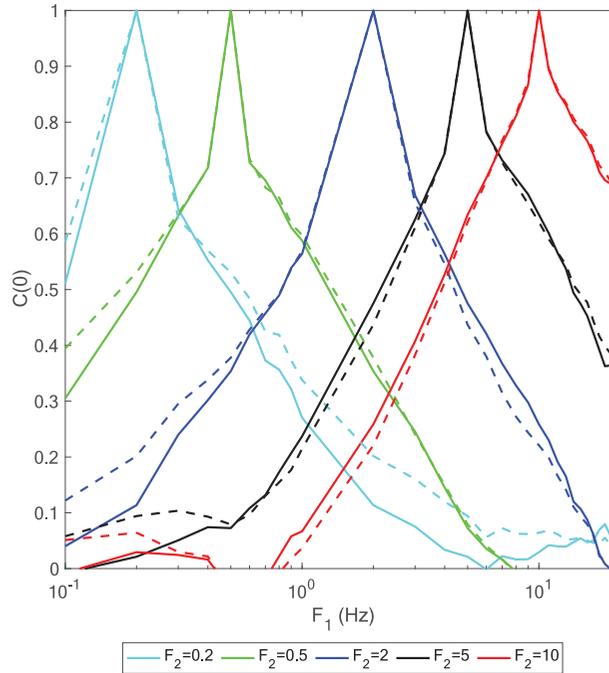


Figure 5. Comparison of the within-event inter-frequency correlation model after setting $h = 0$ in Equation 17 (solid lines) and the Bayless and Abrahamson (2019) empirical within-event inter-frequency correlation coefficients (dashed lines), at reference frequencies 0.2, 0.5, 2, 5, and 10 Hz.

Module”) (Mai et al., 2010; Mena et al., 2010; Olsen and Takedatsu, 2015). The SDSU Module is a hybrid method merging deterministic low-frequency (LF) synthetics and HF stochastic contributions designated as scattergrams. The SDSU Module is implemented on the Southern California Earthquake Center (SCEC) Broadband Platform (BBP), using a number of source realizations (e.g. 50) from the Graves and Pitarka (2015) kinematic source generator to generate the LFs. The HF scattergrams are simulated for each component of ground motion based on the multiple scattering theory by Zeng et al. (1991, 1993). The seismic-scattering wave energy appears after the direct P-wave arrival time, calculated using 3D ray tracing (Hole, 1992). The scattergrams are then convolved with an appropriately magnitude-scaled source time function, assuming that the scattering operators and moment release originate throughout the fault, starting at the hypocenter (Olsen and Takedatsu, 2015).

The SDSU Module passed the SCEC BBP validation exercise (Dreger et al., 2015; Goulet et al., 2015), which assessed ground motion simulations on the basis of their median pseudo-spectral acceleration (PSA) predictions for a specified set of earthquakes in western and eastern United States and Japan, as well as on their degree of agreement with median estimates from the NGA Ground Motion Prediction Equations (GMPEs). Thus, the method has undergone thorough calibration for PSA using GMPEs and strong motion data. However, this validation exercise did not extend to validation of prediction variability measures, and the current SDSU Module (i.e. current as of the above-referenced validation exercise) does not generate time history sets with significant spatial correlation. For example, Figure 6 (top) shows the resulting spatial correlation coefficients for the Loma

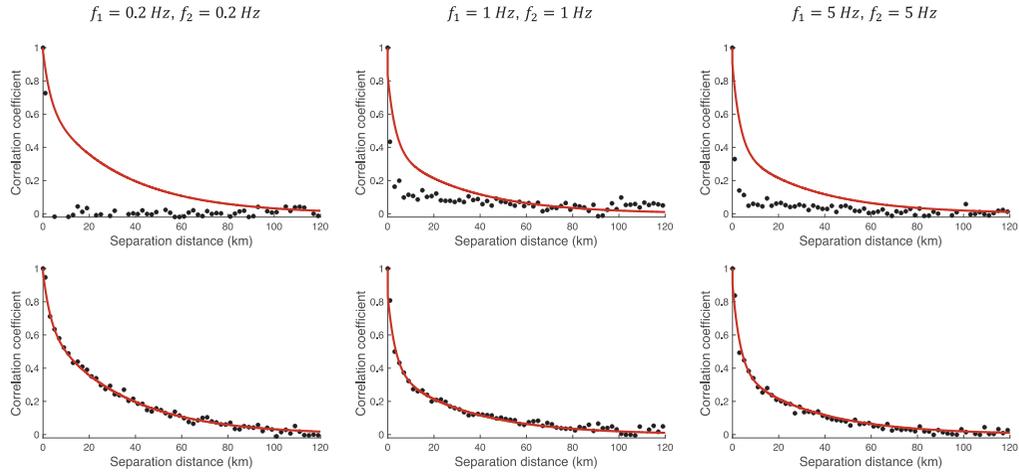


Figure 6. Comparison of the spatial correlation coefficients of ε for EAS at the reference frequency pairs $f_1 = f_2 = 0.2\text{ Hz}$ (left), $f_1 = f_2 = 1\text{ Hz}$ (middle), and $f_1 = f_2 = 5\text{ Hz}$ (right) from the proposed model (red lines) and the SDSU Module before (top) and after (bottom) applying our method (dots) for the Loma Prieta earthquake with 50 source realizations.

Prieta earthquake using the current SDSU Module synthetics compared with the correlation model presented here (Equation 17). It is clear that the spatial correlation coefficients of ε are significantly lower than the empirical model value for station separation distances larger than 1 km. In the following, we implement a post-processing procedure for introducing spatial correlation in SDSU Module synthetic time histories, and show that the results match our specified empirical EAS correlation model.

Our implementation approach for the frequency-dependent spatial correlation is an extension of that developed by Wang et al. (2019) for incorporating inter-frequency correlation. The spatial correlation model in Equation 17 is developed for the within-event residual of the orientation-independent EAS, while the SDSU Module simulations generate separate components of ground motion. For this reason, we apply the EAS frequency-dependent spatial correlation model to the FAS of each of the two horizontal components generated by the method. The resulting synthetic time histories are then found to include correlations in agreement with the EAS correlation model, provided the FAS adjustments made to the two individual horizontal components at each station are suitably correlated. We use a correlation coefficient of 0.7 for the two FAS component adjustments at the same station, a value recommended by Wang et al. (2019) from their study on inter-frequency correlation. The procedure is described in detail in Appendix 2.

We illustrate our method using 50 source realizations for the Loma Prieta earthquake obtained by the kinematic source generator module by Graves and Pitarka (2015). These 50 source realizations have variations in hypocenter locations and slip distributions that are represented by the between-event residual. Here, we refer to each of the 50 source realizations as a separate event. For each event, we generate ten simulations with imposed within-event frequency-dependent spatial correlation at all the stations. The ten simulations differ by the random variables (R_{HC1} and R_{HC2}) in Step (2) in Appendix 2. The mean of the ten simulations and their within-event residuals are computed for each event, respectively. The within-event residual of all the 50 events are then pooled together at the

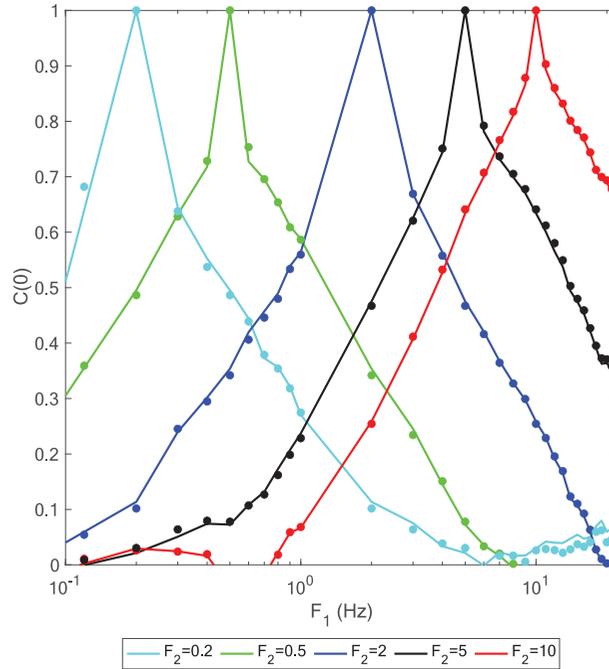


Figure 7. Comparison of the inter-frequency correlation resulting from the presented model setting $h = 0$ in Equation 17 (solid lines) and the correlation coefficients of ε for EAS from the SDSU Module after applying our method to the Loma Prieta earthquake with 50 source realizations (dots) at reference frequencies 0.2, 0.5, 2, 5, and 10 Hz.

corresponding frequencies and stations. Note that, at each station and each frequency, the sampled ε of within-event residual has a length of 500 (50 events by ten simulations). A total of 40 stations are used for the Loma Prieta earthquake in our simulations.

Figure 6 (bottom) shows the spatial correlation coefficients of EAS from 50 source realizations of the Loma Prieta earthquake generated from the SDSU Module with the implementation of our spatial correlation method, at example frequency pairs. In contrast to the low interstation correlation obtained from the current version of the module, the correlation implementation step results in correlation of the synthetics that very closely follows the empirical model, with significant correlation persisting to distances of ~ 50 km. Figure 7 shows the inter-frequency correlation coefficients of EAS from 50 source realizations of the Loma Prieta earthquake generated from the SDSU Module with the implementation of our frequency-dependent spatial correlation model, at 5 reference frequencies. This verifies that the presented frequency-dependent spatial correlation model can address both the spatial correlation and the inter-frequency correlation simultaneously.

Figure 8 shows one example component of synthetic time histories of acceleration and FAS at a station (8001-CLS) for the Loma Prieta earthquake before and after implementing the proposed spatial correlation model. The “uncorrelated” case here in Figure 8 is computed when the off-diagonal correlation terms of the correlation matrixes (Step (5) in Appendix 2) are being set to zero. The comparison shows that the resulting correlation has subtle effects on the time domain. Similar comparisons at other 39 stations are provided in Figure S1, available in the Supplemental material of this article. Note that the

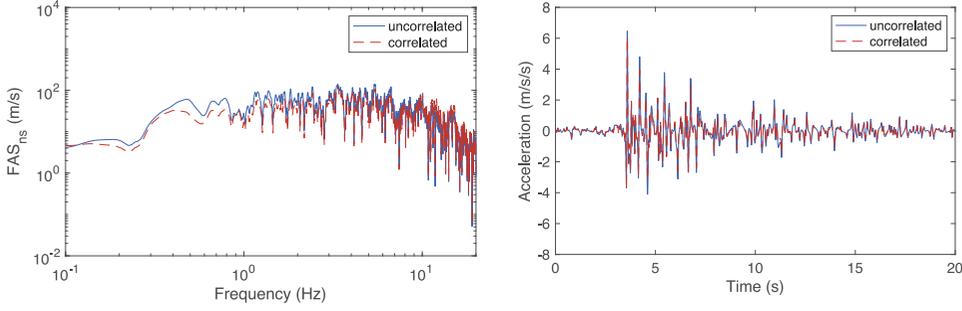


Figure 8. Examples of the north–south component of FAS (left) and acceleration time histories (right) for one simulation of the Loma Prieta earthquake at station 8001-CLS after (red dashed line) and before (blue solid line) applying our method to implement the proposed spatial correlation model.

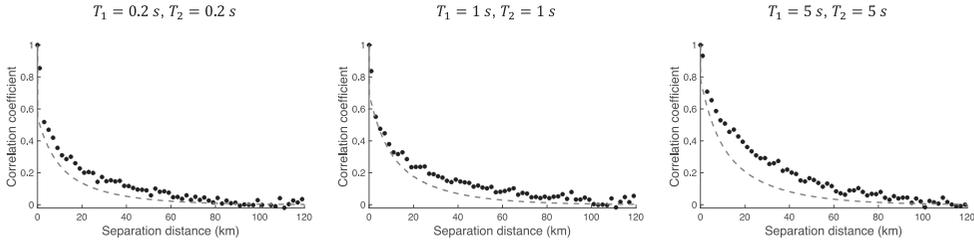


Figure 9. Comparison of the spatial correlation coefficients of ε for the spectral accelerations at reference period pairs (left) $T_1 = T_2 = 0.2$ s, (middle) $T_1 = T_2 = 1$ s, and (right) $T_1 = T_2 = 5$ s from the Loth and Baker (2013) model (dashed lines) and the SDSU Module after applying our method (dots) to the Loma Prieta earthquake synthetic seismograms with 50 source realizations.

FAS correlation adjustment, which is done with an assumption of zero phase adjustment (as is also the case for the inter-frequency correlation analysis, for example, Bayless and Abrahamson, 2018b; Stafford, 2017; Wang et al., 2019), does not lead to any visibly anomalous behavior in the time domain, such as non-causality. For this reason, we did not embrace the additional complexity of doing the FAS adjustments in the form of a causal filter. The resulting spatial correlation coefficients from six other western U.S. earthquakes considered in the SCEC broadband validation exercise (the 1992 M7.2 1992 Landers, the 1994 M6.7 Northridge, the 1986 M6.1 North Palm Springs, the 1987 M5.9 Whittier, the 2008 M5.4 Chino Hills, and the 2007 M5.5 Alum Rock earthquakes) are provided in Figures S2–S7, available in the Supplemental material of this article.

Comparison with other correlation models

Loth and Baker (2013) regressed a within-event spatial correlation model for spectral accelerations based on recordings of 8 earthquakes from the PEER NGA database. The assumptions of stationarity and isotropy are present in both the presented model and the Loth and Baker (2013) model. While the regression model in Equation 11 has the same format as the regression model function of Loth and Baker (2013) model, the range parameters R_1 and R_2 are chosen differently than in the Loth and Baker PSA model, to better fit to the empirical correlation of EAS. Figure 9 shows the resulting cross-correlation

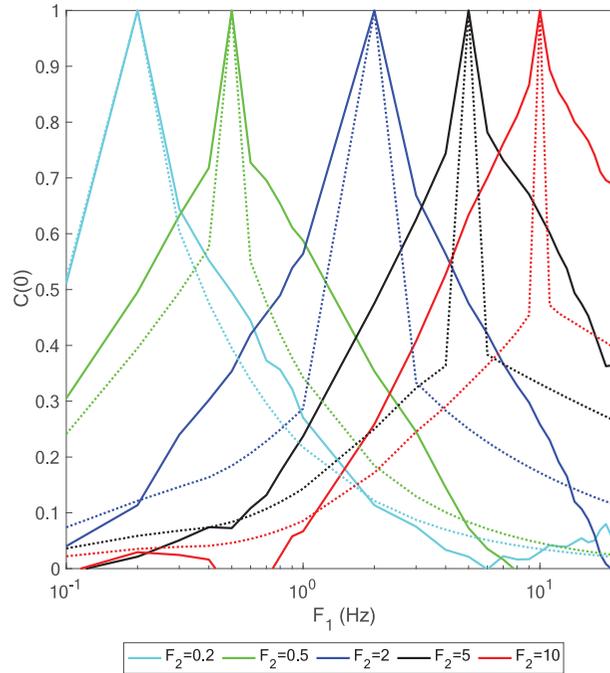


Figure 10. Comparison of the within-event inter-frequency correlation model after setting $h=0$ in Equation 17 (solid lines) and Stafford (2017) within-event inter-frequency correlation model (dotted lines), at reference frequencies 0.2, 0.5, 2, 5, and 10 Hz.

coefficients of response spectral accelerations for the Loma Prieta earthquake synthetics after applying the EAS spatial correlation implementation method, indicating that incorporating the empirical spatial correlation into the EAS of ground motion simulations can also lead to an improvement of the spatial correlation in spectral accelerations. The comparison shows that the Loth and Baker (2013) model's correlation coefficients decay slightly faster than the simulated correlation coefficients. These differences are likely caused by a combination of the following factors. (1) The Loth and Baker (2013) model is regressed for spectral accelerations using a smaller database with 2080 recordings from 8 earthquakes while the present model is regressed from a much larger database with 13,346 recordings from 232 earthquakes. (2) The values of the modeled covariance matrices of Loth and Baker (2013) are obtained by averaging all the fitted coregionalization matrices over various earthquakes while the presented model fits the covariance matrices once after pooling the residuals from all the earthquakes together. (3) The smoothing technique applied in the EAS dataset, which may have increased the correlation, as described previously. One other possible cause of these differences is that the simulations use the actual median as the referencing median to compute the within-event residuals instead of the median ground motion models of PSA used by Loth and Baker (2013). A direct comparison of the two models is provided in Figure S8, available in the Supplemental material of this article.

Stafford (2017) developed inter-frequency correlation models for FAS. A comparison of the presented within-event inter-frequency correlation model ($h=0$ in Equation 17) with Stafford's (2017) within-event inter-frequency correlation model is shown in Figure 10.

The within-event inter-frequency correlation of Stafford (2017) shown here is computed as the combination of their between-site and within-site correlation components. Stafford's (2017) correlation model shows lower correlation and faster decay at higher reference frequency than the model developed in this study and the Bayless and Abrahamson (2019) empirical correlation. These differences are likely caused by a combination of the following factors. (1) The different ground motion component used. The use of both as-recorded FAS horizontal components in Stafford (2017) is a key difference from this study which uses an orientation-independent horizontal component, EAS. (2) The different database and median ground motion model used. Stafford (2017) used a subset of the PEER NGA West1 database to develop the correlation model and used a FAS ground motion model adapted from Yenier and Atkinson (2015) to compute the residual. (3) The different smoothing technique applied. Stafford (2017) used unsmoothed FAS and this study uses smoothed EAS. By averaging the EAS in frequency windows, it is possible that the smoothing could increase the correlation between adjacent frequencies. In this study, the smoothed EAS is chosen to maintain consistency with prior studies, as mentioned before.

Discussion

In this study, we regressed the spatial correlation model at 32 frequency points from 0.1 to 23 Hz, which sufficed to illustrate the efficacy of the implementation of the method. However, if the spatial correlation is needed at additional frequency pairs, a straightforward two-dimensional interpolation of the coregionalization matrices \mathbf{P}^1 , \mathbf{P}^2 , and \mathbf{P}^3 can be applied.

We observe no meaningful event-size dependence of the spatial correlation of earthquake ground motion for the magnitude range in our analysis, as shown by Figure 11, where results are binned by magnitude. Some apparent variation of the spatial correlation with magnitude for smaller sample partitions from the full database is due to unbalanced sampling of earthquake magnitude at a particular distance or frequency. This is consistent with the empirical model of inter-frequency correlation for the EAS residual by Bayless and Abrahamson (2019) that also showed no statistically significant magnitude dependence.

As shown in Figure 4, the correlation at lower frequencies (<0.4 Hz) persists to larger separation distances, compared with higher frequencies, which is expected. Moreover, a small increase of the separation distances at frequencies larger than 4 Hz (meaning that the correlation decreases more slowly with distance at higher frequencies) is observed in Figure 4. We also observed such trend at periods smaller than 0.2 s (i.e. frequencies larger than 5 Hz) in the Loth and Baker (2013) model (for spectral accelerations) which is based on recordings of 8 earthquakes with magnitudes between 6 and 7.6 from the PEER NGA database as shown in Figure S9. We find that the increase in our EAS correlation model at frequencies larger than 4 Hz is mostly dominated by the data from earthquakes with magnitudes between 6 to 7, while the trend is not obvious for other magnitudes (M3-6 and M7-8), as shown in Figure S10. This result warrants further investigation of the data screening procedure or the median ground motion models (used as reference to calculate the residuals). In any case, the trend is relatively small (separation distances increased by ~ 2 km from 4 to 23 Hz when the correlation coefficient equals 0.5), and further decreases as the correlation coefficient increases until no noticeable increase is observed when the correlation coefficient equals 0.7, as shown in Figure S11. Thus, the trend does not affect our

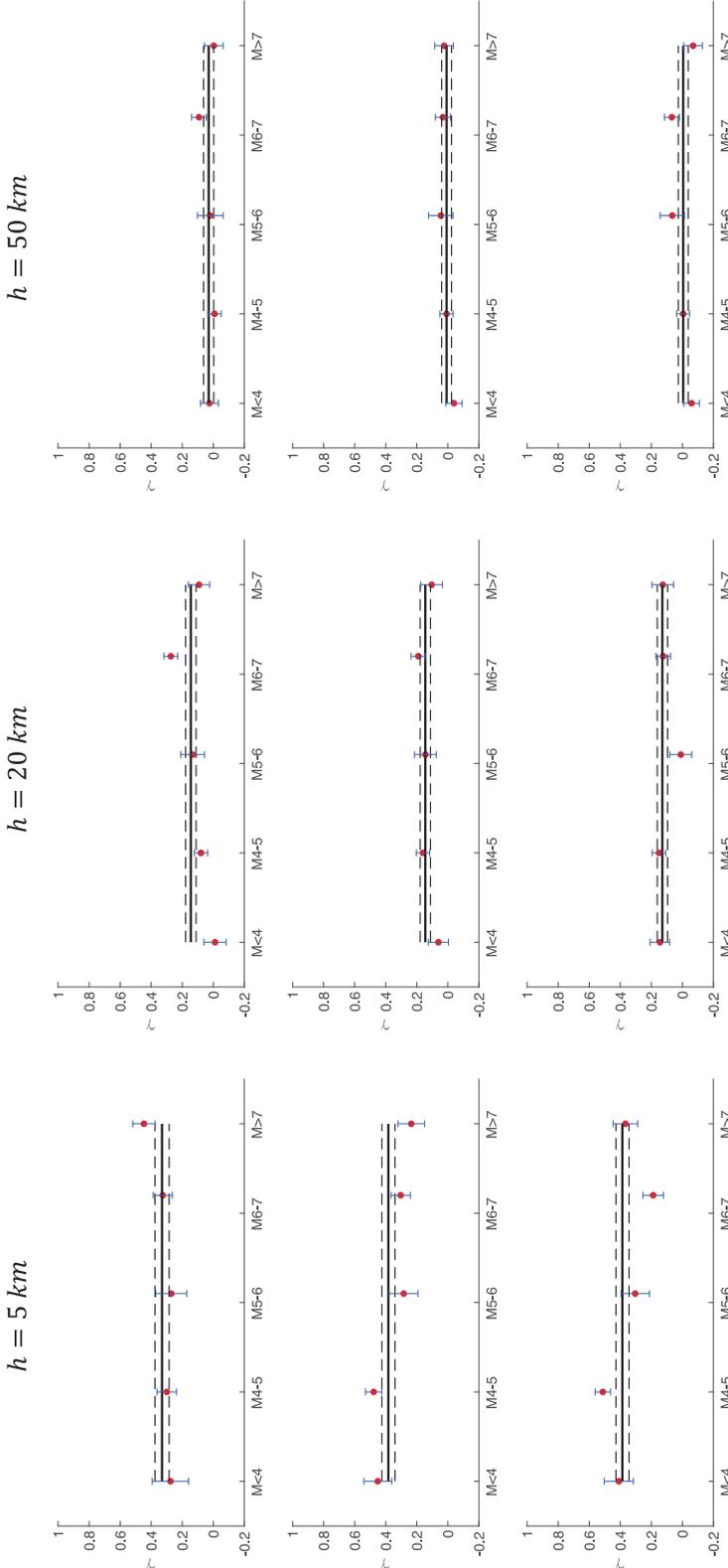


Figure 11. Empirical semivariograms for the full (solid lines) and partial (dots) datasets binned by magnitude ($M < 4$, $4 \leq M < 5$, $5 \leq M < 6$, $6 \leq M < 7$, $M \geq 7$) at the frequency pairs (top row) $f_1 = f_2 = 1\text{ Hz}$, (middle row) $f_1 = f_2 = 5\text{ Hz}$ computed at separation distances (left column) 5 km , (center column) 20 km , and (right column) 50 km . Dashed lines and vertical bars represent 99% confidence intervals of the semivariograms for the full and partial datasets, respectively.

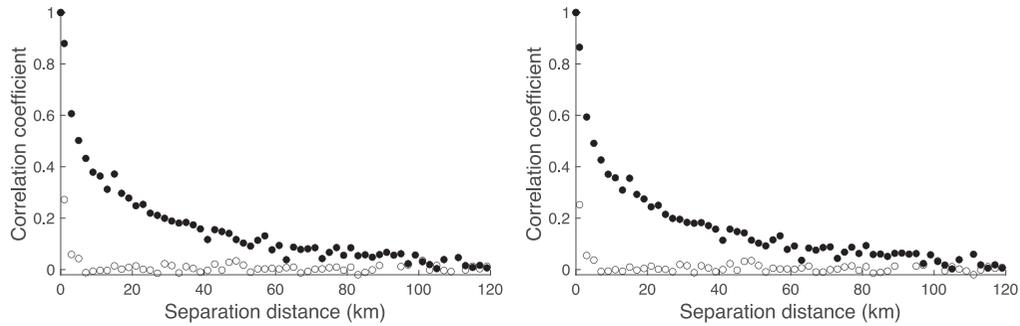


Figure 12. Spatial correlation coefficients of ε for the cumulative absolute velocity (left) and Arias intensity (right) from the SDSU Module before (hollow dots) and after (solid dots) applying our method to the Loma Prieta earthquake synthetic seismograms with 50 source realizations.

overall result that the empirical correlation at lower frequencies persists to larger separation distances. Figures S9–S11 are available in the Supplemental material of this article.

In our analysis, we use a linear coregionalization model to generate our spatial correlation model because of its efficiency using the Goulard-Voltz algorithm, its applicability to a broad frequency range, and its simplicity in the implementation approach. However, other models or regression methods such as the Markov-type screening hypothesis model (Goda and Hong, 2008) and the principal components semivariogram model (Markhvida et al., 2018) may also provide adequate implementation support for the frequency-dependent spatial correlation model.

We have applied the spatial correlation technique to the SDSU Module, which (otherwise) produces broadband synthetic time series with little spatial correlation for either low-frequency deterministic or high-frequency stochastic components. For this reason, the empirical spatial correlation matrices were applied directly to the broadband synthetics as a post-processing procedure. If the method is applied to synthetic time histories from other ground motion generation methods (such as the ones using 3D rupture and crustal models) that already include a certain level of inherent spatial correlation, the procedure should be adjusted to avoid possible double counting. An example of such adjustment is to generate spatially correlated residuals in the process such that the SRSS (square root of the sum of the squares), which consists of the residuals' spatial covariance as well as the inherent spatial covariance, becomes the desired total value that matches the empirical spatial correlation model developed in this study.

The fact that the SDSU Module correlation implementation allows the correlated synthetics to essentially replicate the empirical correlation for PSA suggests that one could now use the correlated SDSU Module synthetics (with the present EAS correlation model implemented) to generate any other ground motion metrics with a valid “empirical” correlation, such as Arias intensity and Cumulative Absolute Velocity. Figure 12, as an example, shows the improved spatial cross-correlation coefficients of cumulative absolute velocity and Arias intensity (combining two orthogonal horizontal components using the arithmetic mean, Travararou et al., 2003) for the Loma Prieta earthquake synthetics after applying the spatial correlation implementation method. This also suggests that the correlated SDSU Module synthetics may provide a means for deriving correlation models for other ground motion metrics. However, further study addressing the spatial variation in the duration of ground motion might be needed to justify such a procedure.

Bayless and Abrahamson (2018a) used a mixed-effect regression in their EAS dataset (as used in our analysis). This regression should include a mean bias term in Equation 3 (e.g. Afshari and Stewart, 2016; Lee et al., 2020). Omission of this mean bias term could cause the within-event (and the between-event) residuals to be altered and lead to a net increase in the estimated spatial correlations. However, we checked that the overall mean bias is negligible in our analysis, justifying leaving out the bias term.

Conclusion

We have developed a frequency-dependent spatial correlation model of ε (the normalized within-event residual) of EAS from the PEER NGA-West2 database, in order to mimic the spatial correlation of observed ground motion data between stations, critical for applications such as seismic risk analysis. The spatial correlation coefficients at lower frequencies decrease more slowly with distance than those at higher frequencies, with no significant dependence on the magnitude of the earthquakes observed. The empirical spatial correlation model of ε is regressed for a linear coregionalization model of semivariograms using the Goulard-Voltz algorithm. We implement the frequency-dependent spatial correlation into ground motion simulations via the SDSU Module on the SCEC BBP. Our method makes use of a two-dimensional Gaussian random variable that has a correlation matrix corresponding to the developed spatial cross-correlation model. The EAS correlation calculated from sets of spatially distributed, two-component synthetic seismograms using our method closely match the empirical EAS correlation model, and the correlation in spectral accelerations from these seismograms is also significantly improved. Because the correlated synthetics successfully replicate empirically derived correlations of spectral acceleration, we suggest that the correlated synthetics from our model could also provide an efficient means for deriving spatially correlated models for other ground motion metrics.

Acknowledgments

The authors thank Jeff Bayless for providing the EAS dataset used in the analysis. They also thank Jeff Bayless and Samuel Shen for their advice and comments during this research project, and are grateful to Brendon Bradley, one anonymous reviewer, the Associate Editor, and the Editor-in-Chief for useful review comments and suggestions that helped improve the article.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Southern California Earthquake Center (SCEC; Contribution Number 10911). SCEC is funded by the National Science Foundation (NSF) Cooperative Agreement EAR-1600087 and the U.S. Geological Survey (USGS) Cooperative Agreement G17AC00047.

ORCID iD

Nan Wang  <https://orcid.org/0000-0002-8028-0043>

Data and resources

Analyses and graphics production were performed using the numeric computing environment MATLAB (www.mathworks.com, last accessed March 2020). All data are from the Pacific Earthquake Engineering Research Center (PEER) Next Generation Attenuation (NGA) West2 database (Ancheta et al., 2014). Simulations in this paper are performed on the SCEC BBP.

Supplemental material

Supplemental material for this article is available online.

References

- Abrahamson N, Schneider J and Stepp J (1991) Spatial coherency of shear waves from the Lotung, Taiwan large-scale seismic test. *Structural Safety* 10(1–3): 145–162.
- Abrahamson NA, Silva WJ and Kamai R (2014) Summary of the ASK14 ground motion relation for active crustal regions. *Earthquake Spectra* 30(3): 1025–1055.
- Afshari K and Stewart JP (2016) Physically parameterized prediction equations for significant duration in active crustal regions. *Earthquake Spectra* 32(4): 2057–2081.
- Al Atik L, Abrahamson N, Bommer JJ, Scherbaum F, Cotton F and Kuehn N (2010) The variability of ground-motion prediction models and its components. *Seismological Research Letters* 81(5): 794–801.
- Ancheta T, Darragh RB, Stewart JP, Seyhan E, Silva WJ, Chiou BSJ, Wooddell KE, Graves RW, Kottke AR, Boore DM, Kishida T and Donahue JL (2014) NGA-West2 database. *Earthquake Spectra* 30(3): 989–1005.
- Atkinson GM and Assatourians K (2015) Implementation and validation of EXSIM (a stochastic finite-fault ground-motion simulation algorithm) on the SCEC broadband platform. *Seismological Research Letters* 86(1): 48–60.
- Atkinson GM, Assatourians K, Boore DM, Campbell K and Motazedian D (2009) A guide to differences between stochastic point-source and stochastic finite-fault simulations. *Bulletin of the Seismological Society of America* 99(6): 3192–3201.
- Bayless J and Abrahamson NA (2018a) *An empirical model for Fourier amplitude spectra using the NGA-West2 database*. Report no. 2018/07, January. Berkeley, CA: Pacific Earthquake Engineering Research Center, University of California.
- Bayless J and Abrahamson NA (2018b) Evaluation of the inter-period correlation of ground motion simulations. *Bulletin of the Seismological Society of America* 108(6): 3413–3430.
- Bayless J and Abrahamson NA (2019) An empirical model for the interfrequency correlation of epsilon for Fourier amplitude spectra. *Bulletin of the Seismological Society of America* 109(3): 1058–1070.
- Beresnev IA and Atkinson GM (1997) Modeling finite-fault radiation from the ω^n spectrum. *Bulletin of the Seismological Society of America* 87(1): 67–84.
- Bolt B, Loh CH, Penzien J, Tsai YB and Yeh YT (1982) *Preliminary report on the SMART-1 strong motion array in Taiwan*. Report no. UCB/EERC-82/13. Berkeley, CA: Earthquake Engineering Research Center, University of California.
- Boore DM (2003) Simulation of ground motion using the stochastic method. *Pure and Applied Geophysics* 160(3–4): 635–676.
- Boore DM (2009) Comparing stochastic point-source and finite-source ground-motion simulations: SMSIM and EXSIM. *Bulletin of the Seismological Society of America* 99(6): 3202–3216.
- Boore DM, Gibbs JF, Joyner WB, Tinsley JC and Ponti DJ (2003) Estimated ground motion from the 1994 Northridge, California, earthquake at the site of the Interstate 10 and La Cienega Boulevard bridge collapse, West Los Angeles, California. *Bulletin of the Seismological Society of America* 93(6): 2737–2751.
- Bycroft GN (1980) *El Centro, California, differential ground motion array*. Report no. 80-919, 21 August. California: US Geological Survey.

- Crempien JG and Archuleta RJ (2015) UCSB method for simulation of broadband ground motion from kinematic earthquake sources. *Seismological Research Letters* 86(1): 61–67.
- Der Kiureghian A (1996) A coherency model for spatially varying ground motions. *Earthquake Engineering and Structural Dynamics* 25(1): 99–111.
- Dreger DS, Beroza GC, Day SM, Goulet CA, Jordan T, Spudich P and Stewart JP (2015) Validation of the SCEC broadband platform V14.3 simulation methods using pseudospectral acceleration data. *Seismological Research Letters* 86(1): 39–47.
- Esposito S and Iervolino I (2011) PGA and PGV spatial correlation models based on European multievent datasets. *Bulletin of the Seismological Society of America* 101(5): 2532–2541.
- Goda K and Hong H (2008) Spatial correlation of peak ground motions and response spectra. *Bulletin of the Seismological Society of America* 98(1): 354–365.
- Goulard M and Voltz M (1992) Linear coregonalization model: Tools for estimation and choice of cross-variogram matrix. *Mathematical Geology* 24(3): 269–286.
- Goulet C, Abrahamson N, Somerville P and Wooddell KE (2015) The SCEC Broadband Platform validation exercise: Methodology for code validation in the context of seismic-hazard analyses. *Seismological Research Letters* 86(1): 17–26.
- Goulet CA, Kottke A, Boore DM, Bozorgnia Y, Hollenback J, Kishida T, Der Kiureghian A, Ktenidou OJ, Kuehn NM, Rathje EM, Thompson EM and Wang XY (2018) Effective Amplitude Spectrum (EAS) as a metric for ground motion modeling using Fourier amplitudes. In: *Seismology of the Americas Meeting*, Miami, FL, 14–17 May.
- Graves R and Pitarka A (2015) Refinements to the Graves and Pitarka (2010) broadband ground-motion simulation method. *Seismological Research Letters* 86(1): 75–80.
- Hao H, Oliveira C and Penzien J (1989) Multiple-station ground motion processing and simulation based on SMART-1 array data. *Nuclear Engineering and Design* 111(3): 293–310.
- Harichandran R and Vanmarcke E (1986) Stochastic variation of earthquake ground motion in space and time. *Journal of Engineering Mechanics* 112(2): 154–174.
- Heresi P and Miranda E (2019) Uncertainty in intraevent spatial correlation of elastic pseudo-acceleration spectral ordinates. *Bulletin of Earthquake Engineering* 17(3): 1099–1115.
- Hole JA (1992) Nonlinear high-resolution three-dimensional seismic travel time tomography. *Journal of Geophysical Research: Solid Earth* 97(B5): 6553–6562.
- Jayaram N and Baker J (2009) Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering and Structural Dynamics* 38(15): 1687–1708.
- Jayaram N and Baker JW (2010) Efficient sampling and data reduction techniques for probabilistic seismic lifeline risk assessment. *Earthquake Engineering and Structural Dynamics* 39(10): 1109–1131.
- Kawakami H and Mogi H (2003) Analyzing spatial intraevent variability of peak ground accelerations as a function of separation distance. *Bulletin of the Seismological Society of America* 93(3): 1079–1090.
- Konno K and Ohmachi T (1998) Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor. *Bulletin of the Seismological Society of America* 88(1): 228–241.
- Kottke A, Abrahamson NA, Boore DM, Bozorgnia Y, Goulet C, Hollenback J, Kishida T, Der Kiureghian A, Ktenidou AJ, Kuehn N, Rathje EM, Silva WJ, Thompson E and Wang X (2018) *Selection of random vibration procedures for the NGA-East project*. Report no. 2018/05, November. Berkeley, CA: Pacific Earthquake Engineering Research Center, University of California.
- Lee RL, Bradley BA, Stafford PJ, Graves RW and Rodriguez-Marek A (2020) Hybrid broadband ground motion simulation validation of small magnitude earthquakes in Canterbury, New Zealand. *Earthquake Spectra* 36(2): 673–699.
- Loth C and Baker JW (2013) A spatial cross-correlation model of spectral accelerations at multiple periods. *Earthquake Engineering and Structural Dynamics* 42(3): 397–417.
- Mai PM, Imperatori W and Olsen KB (2010) Hybrid broadband ground-motion simulations: Combining long-period deterministic synthetics with high-frequency multiple S-to-S backscattering. *Bulletin of the Seismological Society of America* 100(5A): 2124–2142.

- Markhvida M, Ceferino L and Baker JW (2018) Modeling spatially correlated spectral accelerations at multiple periods using principal component analysis and geostatistics. *Earthquake Engineering and Structural Dynamics* 47(5): 1107–1123.
- Mena B, Mai PM, Olsen KB, Purvance MD and Brune JN (2010) Hybrid broadband ground-motion simulation using scattering Green's functions: Application to large-magnitude events. *Bulletin of the Seismological Society of America* 100(5A): 2143–2162.
- Miller M and Baker J (2015) Ground-motion intensity and damage map selection for probabilistic infrastructure network risk assessment using optimization. *Earthquake Engineering and Structural Dynamics* 44(7): 1139–1156.
- Motazedian D and Atkinson GM (2005) Stochastic finite-fault modeling based on a dynamic corner frequency. *Bulletin of the Seismological Society of America* 95(3): 995–1010.
- Olsen KB and Takedatsu R (2015) The SDSU broadband ground-motion generation module BBtoolbox version 1.5. *Seismological Research Letters* 86(1): 81–88.
- Pacific Earthquake Engineering Research Center (PEER) (2015) NGA-East: Median ground-motion models for the Central and Eastern North America region. Report no. 2014/05, April. Berkeley, CA: Pacific Earthquake Engineering Research Center, University of California.
- Sokolov V, Wenzel F, Wen K-L and Jean W-Y (2012) On the influence of site conditions and earthquake magnitude on ground-motion within-earthquake correlation: Analysis of PGA data from TSMIP (Taiwan) network. *Bulletin of Earthquake Engineering* 10(5): 1401–1429.
- Stafford PJ (2017) Interfrequency correlations among Fourier spectral ordinates and implications for stochastic ground-motion simulation. *Bulletin of the Seismological Society of America* 107(6): 2774–2791.
- Travasarou T, Bray JD and Abrahamson NA (2003) Empirical attenuation relationship for Arias intensity. *Earthquake Engineering and Structural Dynamics* 32: 1133–1155.
- Wang M and Takada T (2005) Macrospatial correlation model of seismic ground motions. *Earthquake Spectra* 21(4): 1137–1156.
- Wang N, Takedatsu R, Olsen KB and Day SM (2019) Broadband ground-motion simulation with interfrequency correlations. *Bulletin of the Seismological Society of America* 109(6): 2437–2446.
- Wesson R and Perkins D (2001) Spatial correlation of probabilistic earthquake ground motion and loss. *Bulletin of the Seismological Society of America* 91(6): 1498–1515.
- Withers KB, Olsen KB, Day SM and Shi Z (2019) Validation of deterministic broadband ground motion and variability from dynamic rupture simulations of buried thrust earthquakes. *Bulletin of the Seismological Society of America* 109(1): 212–228.
- Yenier E and Atkinson GM (2015) An equivalent point-source model for stochastic simulation of earthquake ground motions in California. *Bulletin of the Seismological Society of America* 105(3): 1435–1455.
- Zeng Y, Aki K and Teng T (1993) Mapping of the high-frequency source radiation for the Loma Prieta earthquake, California. *Journal of Geophysical Research: Solid Earth* 98(B7): 11981–11993.
- Zeng Y, Su F and Aki K (1991) Scattering wave energy propagation in a random isotropic scattering medium: 1. Theory. *Journal of Geophysical Research: Solid Earth* 96(B1): 607–619.

Appendix I

Multivariate semivariogram and covariance

A semivariogram characterizes the strength of statistical dissimilarity as a function of distance and is often used to describe spatially distributed random variables in geostatistics. The semivariogram is defined as follows:

$$\gamma(s_x, s_y) = \frac{1}{2} E \left[\left(Z(s_x) - Z(s_y) \right)^2 \right], \quad (20)$$

where $E[\]$ denotes the expectation, $Z(s_x)$ and $Z(s_y)$ are random variables at site location s_x and s_y , respectively, and $\gamma(s_x, s_y)$ is the value of the semivariogram for $Z(s_x)$ and $Z(s_y)$. The covariance of two random variables $Z(s_x)$ and $Z(s_y)$ is defined as follows:

$$c(s_x, s_y) = \text{cov}(s_x, s_y) = E[(Z(s_x) - E[Z(s_x)])(Z(s_y) - E[Z(s_y)])]. \quad (21)$$

When empirically estimating the semivariogram or covariance of ground motion observations, the stationary and isotropic assumptions usually need to be established due to the absence of enough data to constrain the additional parameters resulting from a non-stationary and non-isotropic model. Under the stationary and isotropic assumptions, the semivariogram and covariance are independent of the locations and offset direction of the site pair but depend on the distance between the sites. Denoting the separation distance as h , we can write the semivariogram as follows:

$$\gamma(h) = \frac{1}{2} E[(Z(s_x) - Z(s_{x+h}))^2] \quad (22)$$

and the covariance as follows:

$$c(h) = E[(Z(s_x) - E[Z(s_x)])(Z(s_{x+h}) - E[Z(s_{x+h})])]. \quad (23)$$

Here, note that $E[Z(s_x)] = E[Z(s_{x+h})]$ under the assumption of stationarity, are constant at all sites. The relationship between $\gamma(h)$ and $c(h)$ is given as

$$c(h) = c(0) - \gamma(h). \quad (24)$$

For a given set of ground motion observations, the values of ε at nearby stations are correlated and the similarity decreases as the separation distance increases. It can also be shown that ε at neighboring frequencies (f) are probabilistically correlated and are weakly correlated if the frequency pair are widely separated (Bayless and Abrahamson, 2018a). To calculate the semivariogram of ε at multiple frequency pairs, a multivariate semivariogram is used in this study. Denoting $Z_i = \varepsilon(f_i)$ and $Z_j = \varepsilon(f_j)$, we can write the multivariate semivariogram for frequency pair (f_i, f_j) as follows:

$$\gamma_{ij}(h) = \frac{1}{2} E[(Z_i(s_x) - Z_i(s_{x+h}))(Z_j(s_x) - Z_j(s_{x+h}))], \quad (25)$$

where $Z_i(s_x)$ represents the ε at station s_x at frequency f_i . γ_{ij} can then be estimated using the following:

$$\gamma_{ij}(h) = \frac{1}{2N_{ij,h}} \sum_{k=1}^{N_{ij,h}} [(Z_i(s_{k,x}) - Z_i(s_{k,x+h}))(Z_j(s_{k,x}) - Z_j(s_{k,x+h}))], \quad (26)$$

where $N_{ij,h}$ represents the total number of observations of ε at the frequency pair (f_i, f_j) with a separation distance h .

Appendix 2

Here, we describe the implementation approach of our frequency-dependent spatial correlation into the SDSU Module. The current implementation is focused on only two horizontal components; however, this approach generalizes to the vertical component (once the vertical component correlation is defined). The steps are as follows:

1. Take the Fourier transform of the two horizontal components of the synthetic ground motion time series at all stations, and let the station number be m . For each component, let the number of frequency points be n , the Fourier amplitude and phase at the i th frequency be $Amp_{mean}(i)$ and $Ph_{mean}(i)$, respectively;
2. For the two horizontal components 1 and 2, sample normally distributed vector-valued random variables R_{HC1} and R_{HC2} , respectively, with zero mean, constant standard deviation, σ (0.5 for the Loma Prieta event, which is consistent with the original BBP value), and size n at all stations. R_{HC1}^c and R_{HC2}^c are correlated with a correlation coefficient $\rho_R = 0.7$ (Wang et al., 2019), and can be generated by the following steps:

(a) Express covariance matrix C of the two components:

$$C = \begin{bmatrix} 1 & \rho_R \\ \rho_R & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0.7 \\ 0.7 & 1 \end{bmatrix}; \quad (27)$$

(b) Apply the Cholesky decomposition of covariance matrix C and obtain a 2-by-2 upper triangular matrix U as:

$$C = U^T U; \quad (28)$$

(c) Right multiply matrix $[R_{HC1}, R_{HC2}]$ by U so that the resulting two new random variables R_{HC1}^c and R_{HC2}^c have correlation coefficient ρ_R equal to 0.7:

$$[R_{HC1}^c, R_{HC2}^c] = [R_{HC1}, R_{HC2}]U, \quad (29)$$

where $[R_{HC1}^c, R_{HC2}^c]$ and $[R_{HC1}, R_{HC2}]$ are n -by-2 matrices with R_{HC1}^c or R_{HC1} as the first columns and R_{HC2}^c or R_{HC2} as the second columns, respectively. Only the upper triangular matrix featuring the correlation between the two columns of the matrix $[R_{HC1}^c, R_{HC2}^c]$ is used here (Wang et al., 2019).

3. Repeat step (2) three times to generate three sets of independent standard normal random variables $[R_{HC1}^{1c}, R_{HC2}^{1c}]$, $[R_{HC1}^{2c}, R_{HC2}^{2c}]$, $[R_{HC1}^{3c}, R_{HC2}^{3c}]$. For each component, obtain three independent sets of n -by- m random variables Rs_{HC1}^1 , Rs_{HC1}^2 or Rs_{HC1}^3 by combining the vector R_{HC1}^{1c} , R_{HC1}^{2c} or R_{HC1}^{3c} for the first component at all the stations. Similarly, obtain Rs_{HC2}^1 , Rs_{HC2}^2 and Rs_{HC2}^3 for the second component.

The following steps are then the same for the two components, so the “HC1” and “HC2” subscripts are dropped for notational brevity and Rs^1 , Rs^2 , Rs^3 refer to either of the two components of the three n by m random variables, if not specified.

4. Calculate m by m matrices D^1 and D^2 with each element representing the cross-correlation at different station pairs (S_x, S_y) , that correspond to the coregionalization model factors $\exp\left(-\frac{3h}{R_1}\right)$ and $\exp\left(-\frac{3h}{R_2}\right)$ in model $C(h)$:

$$D_{xy}^l = \exp\left(-\frac{3h_{xy}}{R_l}\right), l = 1, 2; \quad (30)$$

5. Apply the Cholesky decomposition to $\mathbf{P}^1, \mathbf{P}^2, \mathbf{P}^3$ to get lower triangular matrices $\mathbf{K}^1, \mathbf{K}^2, \mathbf{K}^3$, and to $\mathbf{D}^1, \mathbf{D}^2$ to get upper triangular matrices $\mathbf{L}^1, \mathbf{L}^2$:

$$\mathbf{P}^1 = \mathbf{K}^1 (\mathbf{K}^1)^T, \mathbf{P}^2 = \mathbf{K}^2 (\mathbf{K}^2)^T, \mathbf{P}^3 = \mathbf{K}^3 (\mathbf{K}^3)^T, \quad (31)$$

and

$$\mathbf{D}^1 = (\mathbf{L}^1)^T \mathbf{L}^1, \mathbf{D}^2 = (\mathbf{L}^2)^T \mathbf{L}^2. \quad (32)$$

6. Compute

$$S = S^1 + S^2 + S^3 = \mathbf{K}^1 R_s^1 \mathbf{L}^1 + \mathbf{K}^2 R_s^2 \mathbf{L}^2 + \mathbf{K}^3 R_s^3, \quad (33)$$

such that S is a matrix of random variables with rows corresponding to different frequencies and columns corresponding to different stations, and S following the correlation model $C(h)$. Note that, here, for the corresponding frequency points outside the 0.1–23 Hz range, $S = (R_s^1 + R_s^2 + R_s^3)/3$;

7. For all the stations, take the exponential of the corresponding column of $S(S^{col})$, and multiply $\exp(S^{col})$ with the station's Amp_{mean} to compute the Fourier amplitude of the new ground motion synthetics, Amp_{new} , as

$$Amp_{new}(i) = Amp_{mean}(i) \exp S_i^{col}; \quad (34)$$

8. Calculate the new ground-motion time series by applying the inverse Fourier transform to the amplitude spectrum obtained in (5) and phase spectrum from (1).