18TH WORLD CONFERENCE ON EARTHQUAKE ENGINEERING
WCEE2024
MILAN

MILAN, ITALY
30TH JUNE - 5TH JULY 2024

www.wcee2024.it

A PARTIALLY NONERGODIC GROUND-MOTION MODEL FOR FOURIER AMPLITUDE SPECTRA FOR THE SAN FRANCISCO BAY AREA, CALIFORNIA, USA

M.P. Moschetti[1], E. M. Thompson[1], R. H. Peterson[1], J. Smith[1], B. Aagaard[1]

[1] U.S. Geological Survey, Golden, CO, USA, mmoschetti@usgs.gov

Abstract: *We develop a partially nonergodic ground-motion model (GMM) for Fourier amplitude spectra for the San Francisco Bay Area, California, USA, using the Bayless and Abrahamson (2019) GMM as a reference ergodic GMM and developing location-dependent adjustments to the predicted median and variance. We compile regional ground-motion data from moment magnitude (M_W) >3 earthquakes occurring during 2000–2022 for which magnitude information is available in the U.S. Geological Survey Comprehensive Catalog (Guy et al., 2015). The data set predominantly consists of records from M_W 3.5–4.5 earthquakes but includes three well-recorded $M_W > 5$ events. Ground-motion residuals are evaluated using the time-averaged shear-wave velocity in the top 30 m (V_{S30}) from the California-specific map of Thompson et al. (2018) and basin-depth site parameters from the seismic velocity model of Aagaard and Hirakawa (2021). The V_{S30} dependence and basin-depth scaling of the reference ergodic GMM of Bayless and Abrahamson (2019) are evaluated and modified with the updated data set. We compute maps of site adjustments using a varying-coefficient model that considers the spatial correlation structure and uncertainties at each observation location. The spatial covariance model is developed using ground-motion residuals that are standardized by the uncertainty model, which allows for consideration of the aleatory variability in developing the site adjustments. The covariance model is fit considering the means and standard deviations of the site terms at all locations. The use of partially nonergodic median adjustments results in modified variance components of the within-event variability. Due to the low number of large-magnitude earthquakes that control seismic hazard in the data set, we do not modify between-event variance; however, we present adjustments to site-to-site variability for use in partially nonergodic hazard assessments.*

1 Introduction

Regional and national seismic hazards assessments, such as the U.S. National Seismic Hazard Model (Petersen et al., 2023; Moschetti et al., 2024, commonly rely on empirical ground-motion models (GMMs) to predict the mean and standard deviation of the intensity of ground shaking for scenario earthquakes. Although seismologists and seismic hazard modelers have long recognized regional variations in features of earthquake ground motions, development of GMMs relies on the use of global data sets to provide sufficient constraint on ground motions from large-magnitude earthquakes and recordings at short distances. By pooling data from different regions, however, these GMMs make use of an implicit assumption of the ergodicity of earthquake ground motions; i.e., that a stochastic process developed from observations made across regions will describe the behavior that will be observed at one location over time (Anderson and Brune, 1999). The expansion of regional seismic networks over the past decades has revealed that repeatable site, path, and possibly source effects on earthquake ground motions can be resolved with sufficiently long observation times (Abrahamson

et al., 2019). Such highly regionalized models have been referred to as “nonergodic GMMs,” indicating that they make use of regional data sets that allow development of region-specific GMMs—or region-specific adjustments—and relax the ergodic assumption (e.g., Lavrentiadis et al., 2023). Following the terminology used in the seismic hazard community, we use the term “partially nonergodic GMM” to refer to those models that include spatially varying site-amplification functions, in addition to the attendant reductions in aleatory variability that result from improved ground-motion predictions.

In the seismically active western United States, nonergodic GMMs have been developed from observations and from three-dimensional (3D) ground-motion simulations (e.g., Landwehr et al., 2016; Sung and Abrahamson, 2022). To capture source, site, and path effects with higher spatial resolution, nonergodic GMMs make use of large data sets of small- to moderate-magnitude earthquakes. Recent observations and understanding of the frequency contributions of Fourier amplitudes (and phase) to oscillator periods for response spectral acceleration calculations indicate that the source spectrum can affect the recovery of repeatable effects, such as site terms (Bora et al., 2016; Stafford et al., 2017). Thus, some GMMs are now being developed for Fourier amplitudes (e.g., Sung et al., 2023). For seismic hazards applications, developers or users must then use random vibration theory to convert from Fourier amplitudes to response spectral acceleration (Boore, 1983; Lavrentiadis and Abrahamson, 2023). Fourier amplitudes have the benefit of not being affected by the spectral shape of the source and therefore do not exhibit magnitude dependence.

In this paper, we compile earthquake ground-motion data for the San Francisco Bay Area, California, USA, and develop a partially nonergodic GMM. We develop the site-specific adjustments using a varying coefficient model that accounts for the spatial correlation of ground-motion observations, as well the spatially varying uncertainties in site response at each recording location. The site-specific adjustments are referenced to the empirical Fourier GMM of Bayless and Abrahamson (2019; hereafter, BA18). The partially nonergodic GMM modifies predictions of site-response at each location using these empirical adjustments.

2 Data and Methods

Development of the partially nonergodic GMM comprises the compilation of a regional data set and development of the spatial prediction model to adjust the reference GMM.

2.1 Earthquake Ground-Motion Data

We initiate the ground-motion data set by defining the set of earthquakes from which to collect waveform data. We use the U.S. Geological Survey Comprehensive Catalog (ComCat) (Guy et al., 2015) to search for earthquakes with moment magnitudes (M_W) greater than 3.0 occurring January 1, 2000, through December 31, 2022, within the region defined by longitudes (-121.30˚ to -123.10°) and latitudes (36.80° to 38.60°). Ground-motion waveforms from strong-motion and broadband stations are compiled for all stations within 2° of all earthquake epicenters in the event set, with data for this region predominantly from the Northern California Earthquake Data Center (NCEDC) and the Center for Engineering Strong Motion Data (CESMD). We chose this search radius to minimize effects from paths with different anelastic attenuation, which becomes a controlling factor in ground-motion amplitudes beyond distances of about 100 km. Ground-motion processing uses the *gmprocess* code (Hearne et al., 2019), which implements automated ground-motion processing algorithms of the Next Generation Attenuation (NGA) projects (e.g., Goulet et al., 2021).

Processing steps are described more completely elsewhere (e.g., Rekoske et al., 2020) and have been applied in numerous regional studies and for real-time earthquake response. Briefly, the code requests time series, applies a baseline correction using the pre-event window, identifies clipped records using multiple algorithms (Kleckner et al., 2022), deconvolves instrument response using the station response files, selects high- and low-pass corner frequencies and filters waveforms based on user-input signal-to-noise-ratio (SNR) criteria (SNR>3), and applies a polynomial-fit-correction to the filtered (displacement) time series to apply an additional baseline correction. Time windows for the earthquake waveform and pre-event noise are based on travel time, phase pickers, duration models, and minimum phase speed estimates. Ground-motion intensity metrics, including spectral acceleration across a range of periods and Fourier spectra as a function of period, are computed from the processed waveforms.

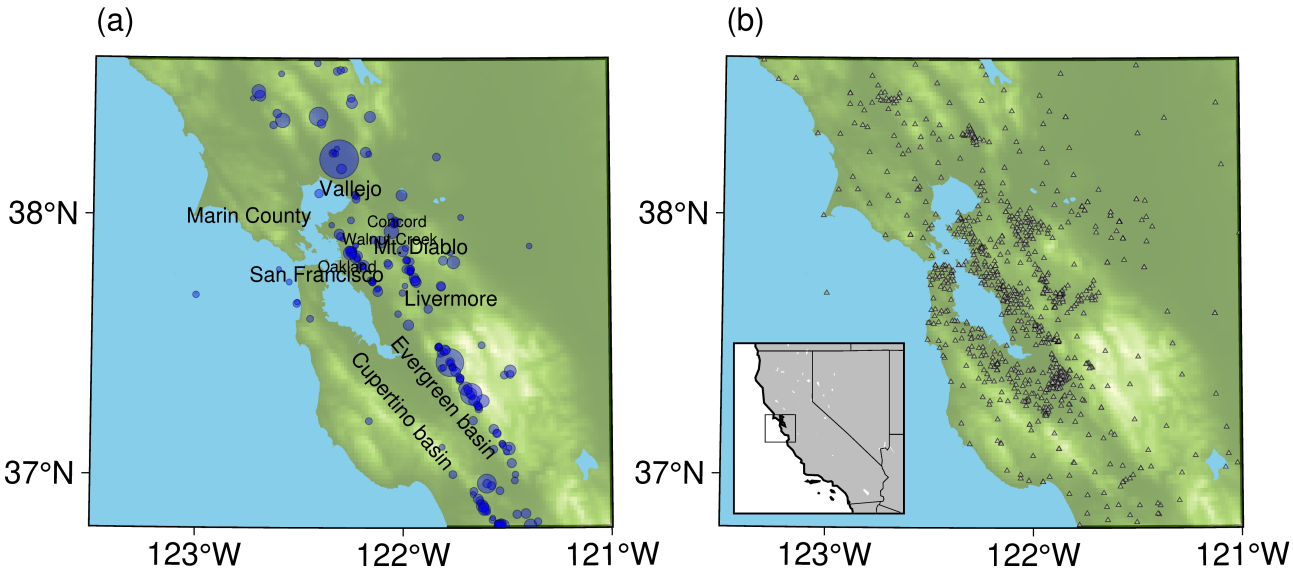

Figure 1. (a) Earthquake epicenters, with symbol size corresponding to earthquake magnitude and (b) station locations for records in the San Francisco Bay Area ground-motion data set.

The ground-motion intensities for this study are "effective amplitude spectra," which are the smoothed spectra obtained by computing the root-mean-square Fourier amplitude spectra from the two orthogonal horizontal components (Goulet et al., 2021). We measure Fourier amplitudes at 21 periods (0.01–10 s). This intensity metric for Fourier amplitudes is consistent with the predictions of the BA18 GMM. We assign site parameters using the time-averaged shear-wave velocity in the top 30 m (V_{S30}) model of Thompson et al. (2018) and the depths to 1 and 2.5 km/s shear-wave horizons (i.e., Z_1 and $Z_{2.5}$) from Aagaard and Hirakawa (2021).

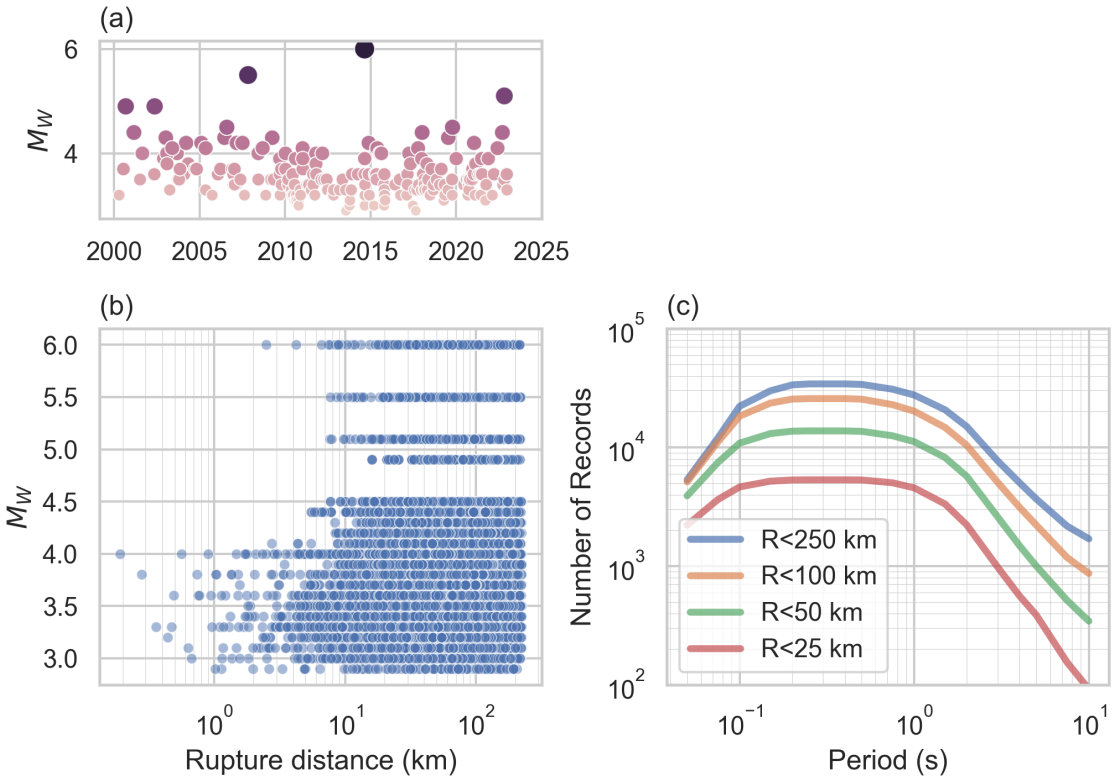

Figure 2. Depiction of the ground-motion data set. Panels depict (a) magnitudes as a function of time, (b) magnitudes as a function of rupture distance, and (c) the number of records that pass the signal-to-noise criterion as a function of period for several maximum distances.

The ground-motion data set consists of more than 34,000 records from more than 1,000 stations within the San Francisco Bay Area region. Although some of the records are from temporary seismic deployments, station density is high within much of the region. Earthquakes in this region predominantly occur along the northwest-striking, right-lateral strike slip faults in the region (Figure 1a). Event magnitudes range from M_W3 to 6. Station density is high within the San Francisco Peninsula, South Bay, and East Bay regions, but with sparse station coverage in the East Bay hills, east of the Hayward Fault (Figure 1b). Record density is high at distances greater than 10 km and magnitudes less than M_W4.5 (Figure 2a). Due to the SNR criterion for accepting intensity metrics, the number of records is greatest within the 0.2–1.0 s period band, with fewer records outside this period band. The full data set contains more than 10,000 records at 2 s period, dropping to about 4,000 records at 5 s period.

2.2 Functional Form of Partially Nonergodic GMM and Ground-Motion Residuals

We define the partially nonergodic GMM relative to the ergodic GMM:

$$\mu_{nonerg}(M, R, V_{S30}, Z_1, x_S) = \mu_{erg}(M, R, V_{S30}, Z_1) + \delta S2S(\boldsymbol{x}_S) \quad \text{Eqn. 1}$$

Where μ_{nonerg} and μ_{erg} are the mean (natural log) predictions of the nonergodic and ergodic GMMs, respectively, which are functions of magnitude M, distance R, V_{S30}, Z_1, and site location x_S, which is implemented in our model as a function of coordinate location (i.e., longitude-latitude pairs). For this work, we use the ergodic GMM of BA18, which was developed from the NGA-West-2 data set (Ancheta et al., 2014), but which selectively used records from California and Nevada for constraining the anelastic attenuation terms. The spatially varying site adjustments (i.e., site-to-site terms) for the partially nonergodic GMM are given by the term $\delta S2S(x_S)$. These terms represent the spatially varying site-specific deviations from the GMM-predicted ground motions that are indicated by observations in the data set.

We develop the $\delta S2S(\boldsymbol{x_s})$ model by fitting a covariance model, at each period, to the within-event terms of the ground-motion residuals and their uncertainties at each station. We first compute residuals R_{es} for every observed ground motion of event e and station s (i.e., Y_{es}) as:

$$R_{es} = \ln(Y_{es}) - \mu_{erg}(M, R, V_{S30}, Z_1) \quad \text{Eqn. 2}$$

And partition the residuals using a mixed-effects regression that uses a grouping on unique events:

$$R_{es} = c + \delta B_e + \delta W_{es} \quad \text{Eqn. 3}$$

Where the regression provides a bias c, between-event terms δB_e, and within-event terms δW_{es} for all records.

We standardize the within-event terms using the within-event variability ϕ_{es}^{BA19} from the BA18 GMM and compute standardized site terms z_s and their standard deviations ϕ_{z_s} at each station as:

$$z_{es} = \frac{\delta W_{es}}{\phi_{es}^{BA19}} \quad \text{Eqn. 4}$$

$$z_s = \frac{1}{N}\sum_{e=1}^{N} z_{es} \quad \text{Eqn. 5}$$

$$\phi_{z_s} = std(z_{es})_e \quad \text{Eqn. 6}$$

Standardized residuals and their uncertainties are not commonly used in developing nonergodic GMMs, but this approach allows for development of GMM-consistent modifications that incorporate the modeled dependence of ground-motion variability on input parameters. For example, within-event variability of BA18 is magnitude dependent. Standardized residuals have been used in computing spatial adjustments for earthquake-response ground-motion maps (e.g., Engler et al., 2022).

2.3 Spatial Correlation Model and Spatially Varying Site-Adjustment Model

The spatial model δS2S(x_s) assumes that the site-adjustment factors are described by a Gaussian process, such that a finite collection of observations can be described by a multivariate normal distribution (MVN) that is described by a mean function $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, $\boldsymbol{Y}$ ~ MVN($\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$)**.**

Given observations $\mathbf{Y_2}$ and predictions $\mathbf{Y}_1$, such that $\boldsymbol{\mu} = \begin{bmatrix}\boldsymbol{\mu_1}\\ \boldsymbol{\mu_2}\end{bmatrix}$ and $\boldsymbol{\Sigma} = \begin{bmatrix}\boldsymbol{\Sigma_{11}} & \boldsymbol{\Sigma_{12}}\\ \boldsymbol{\Sigma_{22}} & \boldsymbol{\Sigma_{22}}\end{bmatrix}$, the conditional mean function and covariance matrix are:

$$\boldsymbol{\mu_{Y_1|Y_2}} = \mu_1 + \boldsymbol{\Sigma_{12}\Sigma_{22}^{-1}}(y_2 - \boldsymbol{\mu_2}) \quad \text{Eqn. 7}$$

$$\boldsymbol{\Sigma_{Y_1|Y_2}} = \boldsymbol{\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}} \quad \text{Eqn. 8}$$

Where $\boldsymbol{\Sigma_{k\,\ell}} = Cov(\boldsymbol{Y_k}, \boldsymbol{Y_\ell})$for $k, \ell = 1,2$. The covariance matrix defines a functional relationship between the observations in a spatial domain. We model the covariance through the Matérn covariance function (Stein 1999):

$$K(\boldsymbol{s_i}, \boldsymbol{s_j}) = \frac{\sigma^2}{2^{\nu-1}\,\Gamma(\nu)}\left(\kappa\left|\left|\mathbf{s_i} - \mathbf{s_j}\right|\right|\right)^{\nu} \mathrm{C}_\nu\left(\kappa\left|\left|\mathbf{s_i} - \mathbf{s_j}\right|\right|\right) \quad \text{Eqn. 9}$$

Where κ > 0 is a scaling coefficient, which we can interpret as a range parameter, ν > 0 is a smoothing parameter, σ^2 is the process variance, and C_ν is a modified Bessel function of second kind with order ν, with $i, j = 1, \ldots, n$. Next, we consider the inclusion of a nugget effect and standard deviations ϕ_{z_s} in our observations $\mathbf{Y_2}$. For the covariance of our observations, we have

$$\Sigma_{22} = K(\boldsymbol{s_i}, \boldsymbol{s_j}) + \tau^2\, W^{-1} \quad \text{Eqn. 10}$$

where τ^2 is our nugget variance and W is our weight matrix from our standard deviation, with

$$W = \begin{bmatrix} \phi_{z_1}^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \phi_{z_n}^2 \end{bmatrix} \quad \text{Eqn. 11}$$

We solve for unknown parameters using Maximum Likelihood Estimation (Gelfand et al, 2010), with the log-likelihood given by:

$$log\ell(\tau^2, \sigma^2, \theta;\, Y) = -\frac{n}{2}log(2\pi) - \frac{1}{2}log\big(det(\Sigma)\big) - \frac{1}{2}(Y-\mu)^T\Sigma^{-1}(Y-\mu) \quad \text{Eqn. 12}$$

3 Results

We present the results from mixed-effects regressions, in addition to the fit of and predictions from the spatial correlation in the following sub-sections.

3.1 Site Terms and Standardized Residuals

Site terms from mixed-effects regressions (Eqn. 3) are evaluated by computing the mean of within-event terms at each location, $\langle W_{es} \rangle_e$. Site terms are given as the (natural log) amplifications and are unitless. Results of the site terms are depicted in Figure 3.

Across the maps—for all periods and many regions—we find that site terms are close to zero for many sites, indicating that the site-response model of BA18, which uses site parameters V_{S30} and Z_1, provides good predictive skill. However, there are areas where the site-response models do not agree with observations. At 0.1 s period, we observe underpredictions (positive residuals) through much of the San Francisco Peninsula and near the Hayward Fault; ground motions are overpredicted (negative residuals) in the northern part of the Evergreen basin and in the Livermore basin. At longer periods ($T \geq 1\ s$) similar spatial patterns emerge, with varying amplitudes. Ground motions near San Francisco and within the San Francisco Peninsula are overpredicted. Overpredictions also emerge in the East Bay, west of the Hayward Fault. The northern Evergreen basin remains anomalously overpredicted, and the Cupertino basin becomes increasingly underpredicted with increasing period.

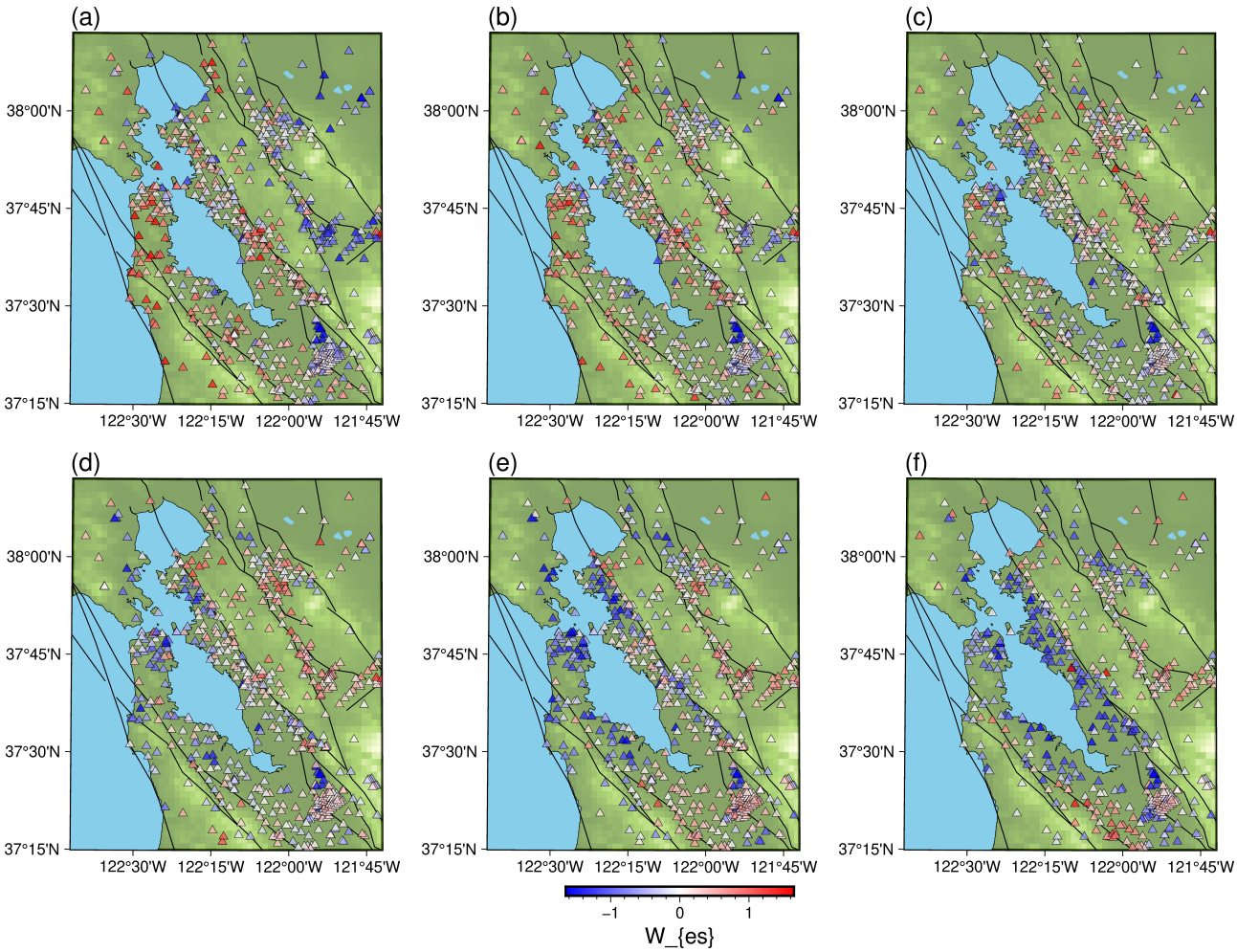

Figure 3. Maps of site terms $\langle W_{es} \rangle_e$ at periods (a) 0.1 s, (b) 0.2 s, (c) 0.5 s, (d) 1.0 s, (e) 2.0 s, and (f) 5.0 s.

Standard deviations in the site terms are computed to show the variations in the within-event residuals that exist at different stations, $std(W_{es})$. Such variations include many of the ground-motion effects that are not modeled in the empirical GMM, such as variability in site response due to azimuth, incidence angle, or other features, deviations of specific source-site path terms from the average path attenuation, and finite-source effects (e.g., Seekins and Boatright, 2010; Hirakawa et al., 2023). Site-specific standard deviations of the within-event residuals show strong spatial patterns that range from about 0.3–0.7 natural log units, indicating significant differences in ground-motion variability among the sites. These effects are particularly notable because the data set primarily comprises records from small- to moderate-magnitude earthquakes for which the finite-source effects (e.g., rupture directivity) are likely to be small or negligible in the period band of this analysis. Notable features include high standard deviations at short periods (0.1–0.2 s) in the East Bay region near the Hayward and Calaveras faults and near San Francisco, persistent low standard deviations in the South Bay region across a broad period band (0.2–2 s), reduced station-to-station variability near 0.5–1 s period, and increasing variability for periods greater than 1 s.

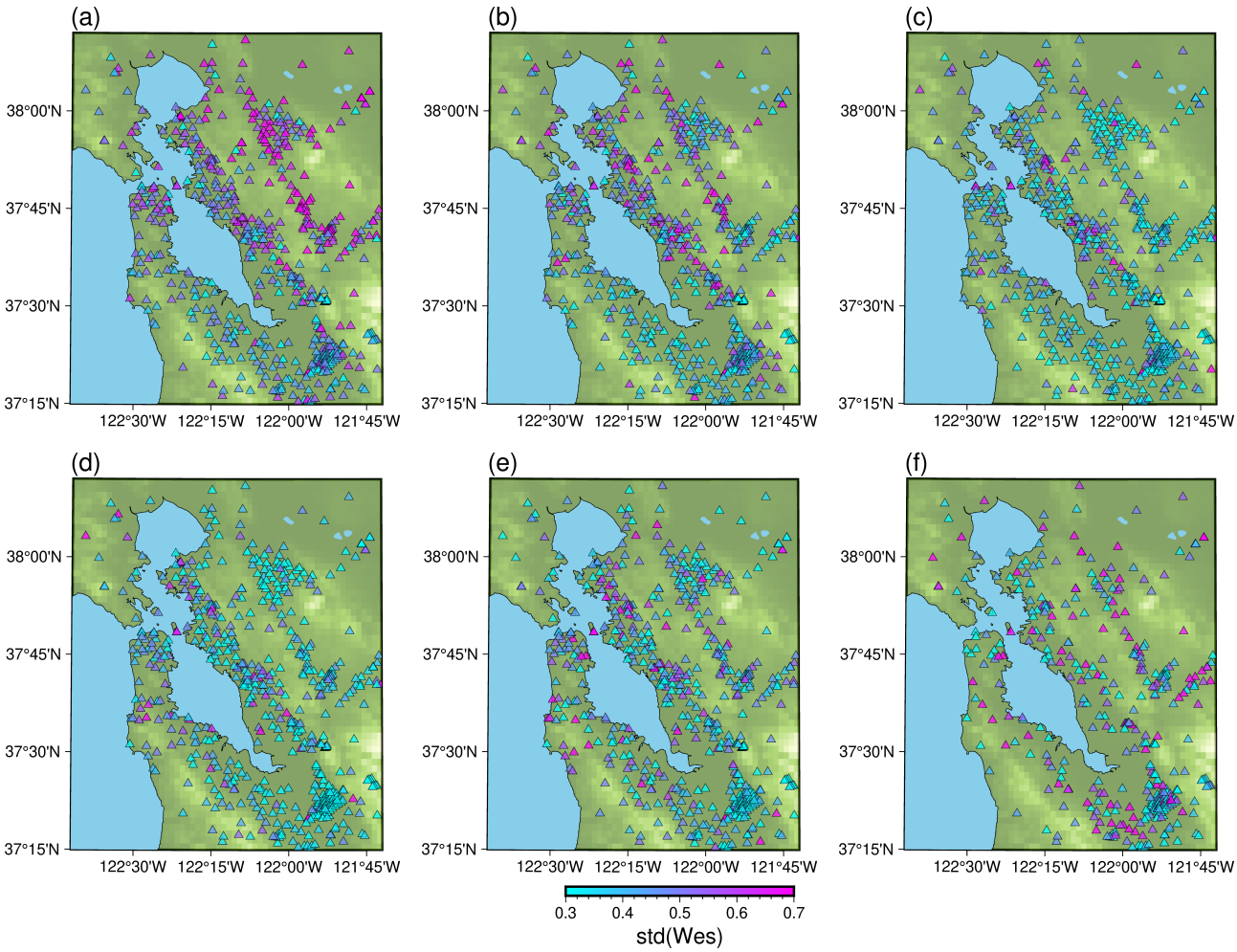

Figure 4. Maps of standard deviations of the within-event terms at each station, $std(W_{es})$ at periods (a) 0.1 s, (b) 0.2 s, (c) 0.5 s, (d) 1.0 s, (e) 2.0 s, and (f) 5.0 s.

We compute standardized residuals (Eqn. 4) and compare the distributions to the standard normal distribution (Figure 5). We find broad agreement between the standardized residuals and the normal distribution indicating that the variability of within-event residuals agrees with the model predictions of BA18 on average.

3.2 Spatial Correlation Model and Partially Nonergodic Site Adjustments

We compute correlation lengths for the covariance function (Eqn. 9) using maximum-likelihood estimation with the *fields* package for spatial statistics (Nychka et al., 2021). Correlation lengths are determined from (Eqn. 11) by first solving analytically for σ^2, then numerically maximizing the other two parameters, including correlation length, θ. We note that correlation lengths typically increase with increasing period, with mean estimates ranging from about 1.5 km to 5.8 km. This increase in correlation lengths is consistent with other estimates from alternative data sets (e.g., Loth and Baker, 2013) and with the physical mechanism of increasing seismic wavelengths for increasing period. Loth and Baker (2013) used data from eight large-magnitude (M>6) and no earthquakes from northern California.

We compute the conditional mean and conditional covariance matrix (Eqns. 7, 8) using the covariance functions (Eqn. 9) and a smoothness parameter of $\nu = 1$. However, there is no closed-form solution for this covariance function, which is possible with other choices of smoothness. For example, a Matérn with $\nu = ½$ gives an exponential covariance function. We further modify the covariance matrix by the measured variance in the standardized site terms (Eqn. 10). The modified covariance matrix accounts for the variation between sites in their standardized ground-motion residuals, which we interpret to have physical causes that are not modelled in GMMs (Eqn. 10).

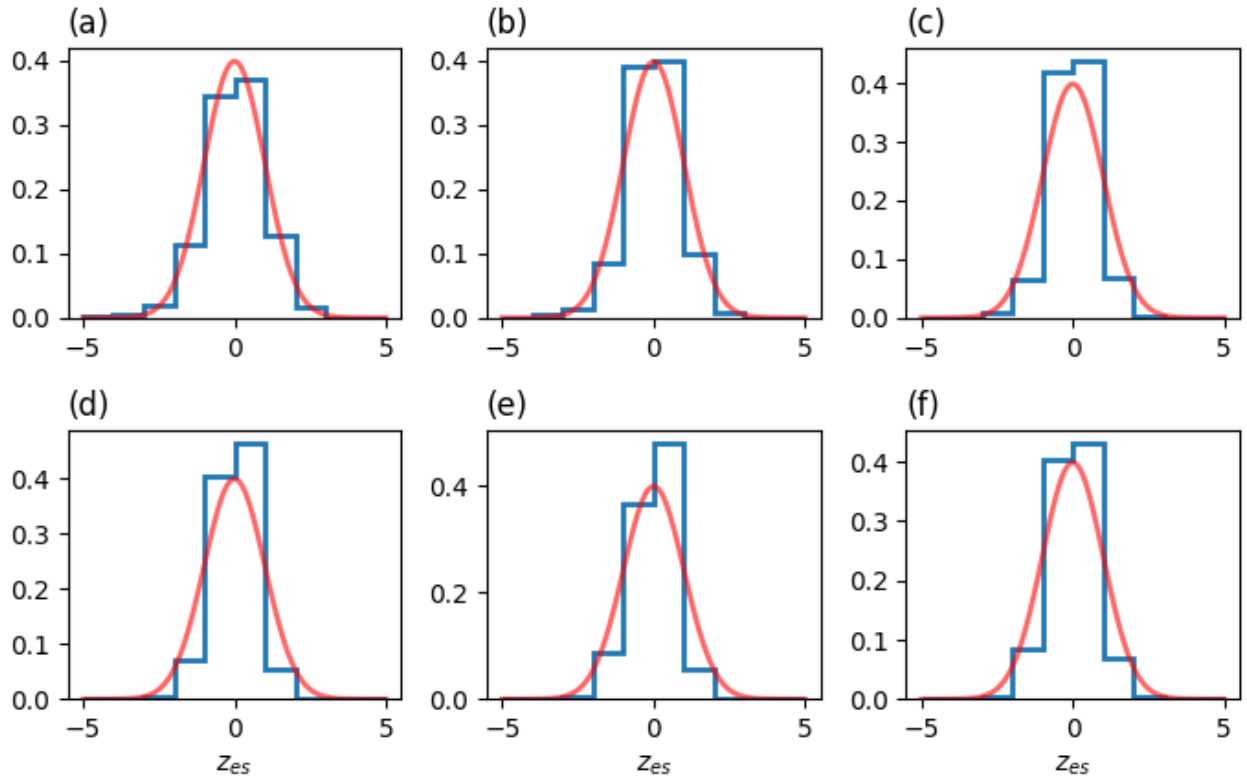

Figure 5. Comparisons of the distribution of standardized residuals z_{es} (blue) with standard normal distributions (red) at periods (a) 0.1 s, (b) 0.2 s, (c) 0.5 s, (d) 1.0 s, (e) 2.0 s, and (f) 5.0 s.

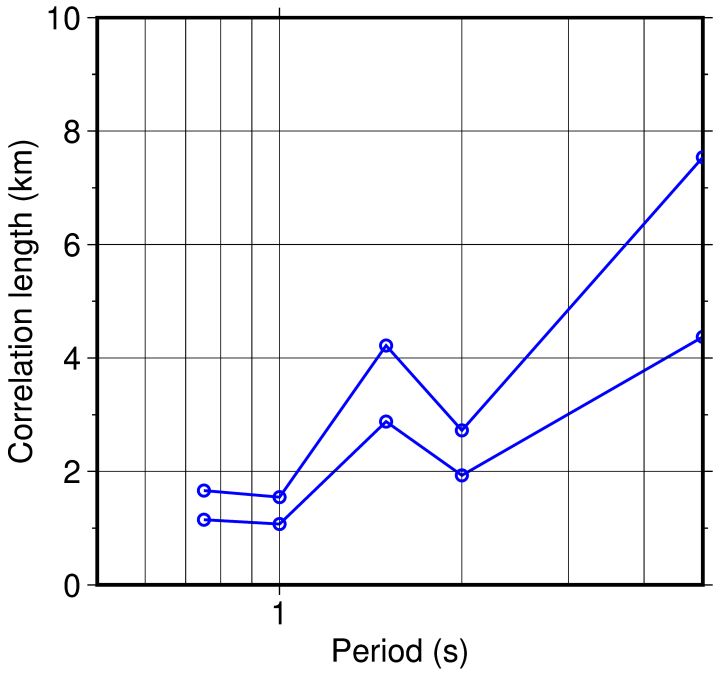

Figure 6. Correlation lengths for the covariance function computed from Fourier amplitudes of ground-motion data. Blue lines define the 95-percent confidence interval for the correlation lengths.

We present example results of the conditional means and variances at 2.0 s period (Figure 7). The conditional mean values correspond to the spatially varying, partially nonergodic adjustments, relative to the prediction of

the empirical GMM. Negative values identify regions where the observations indicate lower-than-predicted site response (or predicted ground motions), and positive values identify where the observations indicate higher-than-expected site response. Because the GMMs already include the effects of V_{S30} and Z_1, the spatially varying adjustments highlight areas where the GMM predictions can be improved.

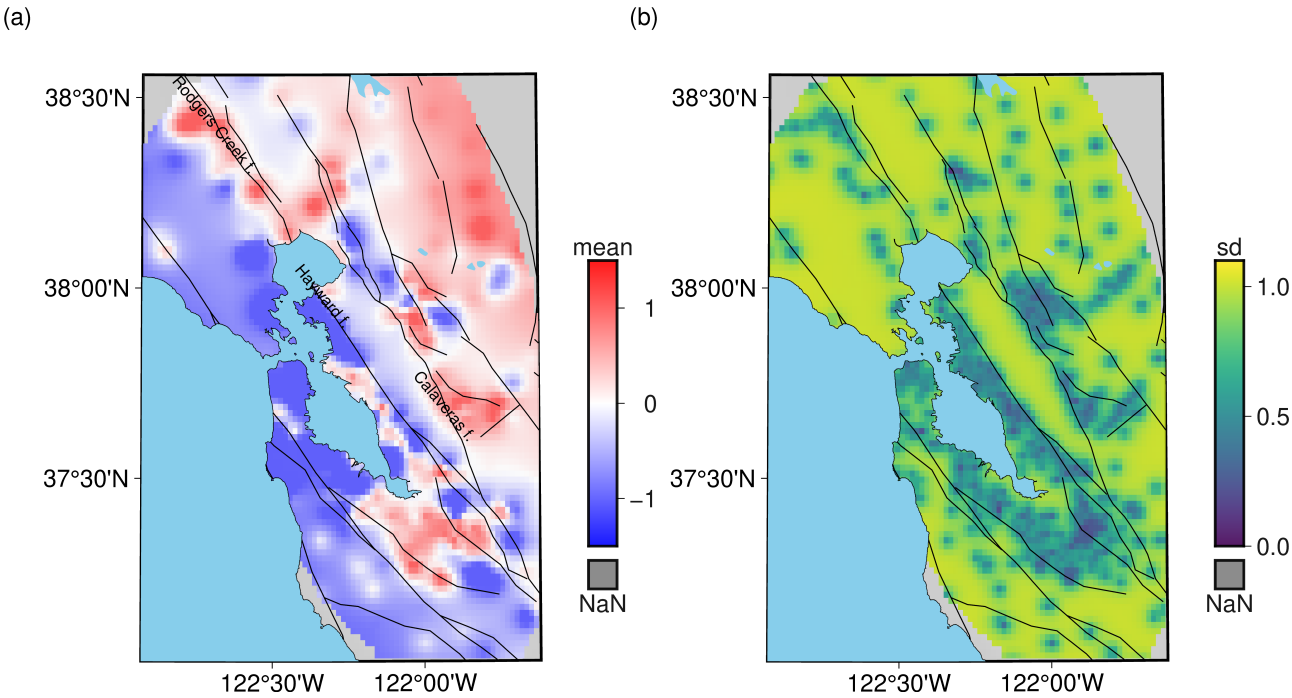

Figure 7. (a) Conditional means and (b) conditional variances at 2.0 s period. The conditional means and variances correspond to the spatially varying adjustments and uncertainties inferred from regional data. Adjustments are given as natural log modifications to the ergodic GMM.

Significant deviations from the GMM predictions occur at numerous sites in the region. Overpredictions at 2.0 s period occur throughout the San Francisco Peninsula, west of the Hayward Fault in the East Bay, through much of Marin County, and in pockets throughout the region, notably at locations along the Hayward Fault, near Vallejo, and in the Mt. Diablo region. Underpredictions occur through much of the southeast Bay region west of the Hayward Fault, through much of the South Bay, along the Rodgers Creek fault, and in pockets through the East Bay, including near Walnut Creek, Concord, and north of Livermore. The urbanized sedimentary basins in the East and South Bay are notably underpredicted at 2.0 s period, indicating that the observation exceed the GMM predictions that use site parameters V_{S30} and Z_1.

The spatially varying variances provide information on the precision of these adjustments and have important implications for probabilistic seismic hazard analyses. These variances correspond only to the within-event (i.e., ϕ) component of the variance. The mapped values indicate large variations in space that may be due to geologic structure or structural gradients or to other unknown causes. Conditional variances range from about 0.1–1.0. Where observational data constrains the site amplifications, there are significant improvements in the precision of the ground-motion predictions. Because observations predominantly include small- to moderate-magnitudes, and we focus the model development on station-specific adjustments, the model yields improved station-specific linear amplification models. Such data sets can also be extended to capture repeatable, path-specific trends. Whether nonergodic adjustments to event terms may yield benefits remains an open area of research (e.g., Hardebeck, 2020). The authors are unaware of attempts to constrain nonergodic, nonlinear behavior.

To adjust GMMs for the nonergodic adjustments described here, users must select a within-event standard deviation value (ϕ) to unstandardize the adjustments and compute the partially nonergodic ground-motion adjustment at each station. This step consists of defining a target value for the within-event variability that can be used to scale the standardized site model. There is some ambiguity in this step because the within-event variability ϕ can be a function of magnitude, distance, V_{S30}, or other parameters. We identify two approaches

to implement the standardized residuals: (1) forward calculations—whether deterministic or probabilistic—can use scenario-dependent values for the within-event variability (i.e., $\phi_{es} = f(M, R, V_{S30}, ...)$ to compute event-specific adjustments, $\delta_{GMM,es} = z_s \phi_{es}$. For probabilistic calculations, this will require looping over all sources that contribute to seismic hazard at a particular site and modifying the predicted ground motions using these adjustments. (2) Alternatively, one can select a representative site-scenario value for the within-event variability ($\overline{\phi_{es}}$) and use this value to compute a single GMM adjustment value, $\delta_{GMM,s} = z_s \overline{\phi_{es}}$. Approaches to determining the representative site-scenario values could be determined from the mean of the within-event variability values from the data set, from the mean variability value from a probabilistic hazard analysis, or from representative or controlling variability values from a seismic hazard analysis (e.g., from disaggregation). To clarify, most current models make implicit use of the first option provided here (i.e., mean of within-event variability values from the data set) in constructing partially nonergodic adjustments.

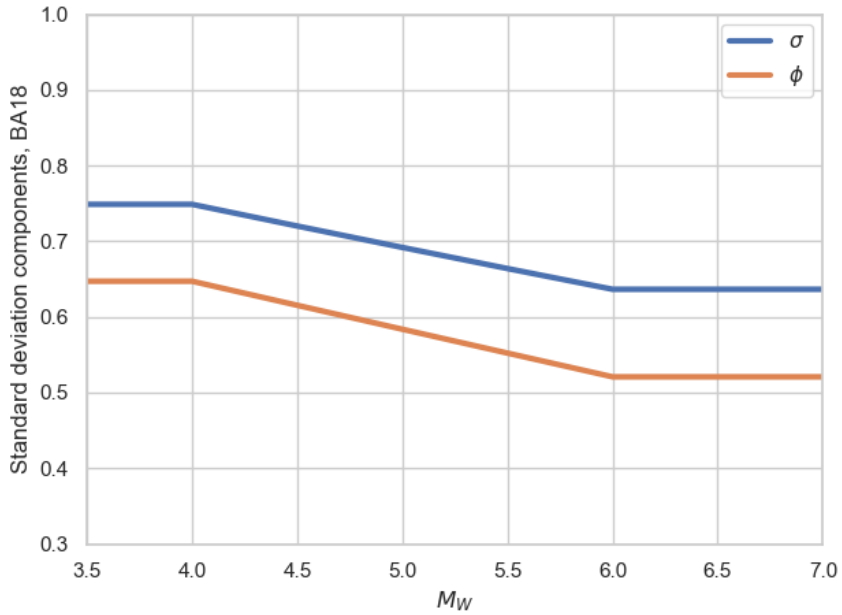

Figure 8. Aleatory variability components BA18. Total standard deviation (σ) and within-event variability (ϕ) are plotted as a function of moment magnitude.

The impact of using event-site-specific variability values on the partially nonergodic adjustments is unclear, and further work would be beneficial to understand the impact of this choice. The BA18 has a magnitude-dependent aleatory variability, which is primarily controlled by the within-event variability. Between-event variability also has a weak magnitude dependence, but due to its lower amplitudes compared to within-even variability, this effect has a much smaller effect on total standard deviation. If the development of the partially nonergodic GMM did not consider the modelled higher variability of small-magnitude data ($M_W \lesssim 4.5$)—which dominate the ground-motion data set due to the Gutenberg-Richter magnitude-frequency distribution—the resulting uncertainties would be inflated, relative to what is produced using standardized residuals. As a consequence, seismic hazard assessments would use a larger than appropriate aleatory variability, contributing to high ground motions at low probabilities of exceedance. Although BA18 includes only magnitude dependence of the aleatory variability, other GMMs from the NGA-West-2 GMM suite (Bozorgnia et al., 2014) have magnitude, distance, and V_{S30} dependence. Inconsistencies between the recorded data sets used to develop partially nonergodic GMMs and the conditions or parameters used to compute seismic hazard may cause unintended artifacts in the resulting models.

4 Conclusions

Seismic hazard analyses are increasingly taking advantage of regional data sets to modify empirical GMMs and account for regional nonergodic site and path effects. The use of spatially varying site adjustments, and the resulting average reductions in aleatory variability, are referred to as partially nonergodic GMMs. We

compile and process earthquake ground-motion data from earthquakes that occurred during 2000–2022 in the San Francisco Bay Area to develop a partially nonergodic GMM for Fourier amplitude spectra. The partially nonergodic adjustments are developed as a varying coefficient model that makes use of standardized residuals and accounts for the modelled variations in ground-motion variability with magnitude. For GMMs for which the aleatory variability models depend on site and propagation-path parameters (such as V_{S30} and R_{rup}), development of partially nonergodic adjustments without consideration of potential inconsistencies between the recorded data sets and the distributions used in hazard analyses may produce artifacts. We focus the model development here by using the ground-motion predictions at 2.0 s period. The resulting spatial adjustments show ground-motion anomalies exceeding a factor of two from the median predictions at numerous sites within the studied region. Noteworthy deviations include overpredictions in the San Francisco Peninsula and Marin County, as well as the East Bay region west of the Hayward Fault and north of Oakland; adjustments indicate underpredictions by the median model in the South Bay, within the basins east of the Calaveras fault, and along the Rodgers Creek fault. The spatial adjustments include updated estimates of the within-event variabilities that could be used for probabilistic seismic hazard analysis or for deterministic predictions of fractiles other than the median value. Spatially varying uncertainties are lower where observations are available and taper to the GMM-predicted value away from these locations. The reductions in within-event variability can be substantial at some locations, and further work would be beneficial to evaluate the use of observations for constraining future ground motions from the large-magnitude earthquakes that are relevant to seismic hazard assessments.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

5 References

Aagaard, B.T. & Hirakawa, E.T. (2021). San Francisco Bay region 3D seismic velocity model v21.1: U.S. Geological Survey data release, doi:10.5066/P9TRDCHE.

Abrahamson, N. A., Kuehn, N. M., Walling, M., & Landwehr, N. (2019). Probabilistic seismic hazard analysis in California using nonergodic ground-motion models. *Bulletin of the Seismological Society of America*, *109*(4), 1235-1249.

Ancheta, T. D., Darragh, R. B., Stewart, J. P., et al. (2014). NGA-West2 database. *Earthquake Spectra*, *30*(3), 989-1005.

Anderson, J. G. & Brune, J. N. (1999). Probabilistic seismic hazard analysis without the ergodic assumption. *Seismological Research Letters*, *70*(1), 19-28.

Bayless, J. & Abrahamson, N. A. (2019). Summary of the BA18 ground-motion model for Fourier amplitude spectra for crustal earthquakes in California. *Bulletin of the Seismological Society of America*, *109*(5), 2088-2105.

Boore, D. M. (1983). Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. *Bulletin of the Seismological Society of America*, 73(6):1865-1894.

Bora, S. S., Scherbaum, F., Kuehn, N., & Stafford, P. (2016). On the relationship between Fourier and response spectra: Implications for the adjustment of empirical ground-motion prediction equations (GMPEs). *Bulletin of the Seismological Society of America*, *106*(3), 1235-1253.

Bozorgnia, Y., Abrahamson, N. A., Atik, L. A., et al. (2014). NGA-West2 research project. *Earthquake Spectra*, *30*(3), 973-987.

Engler, D. T., Worden, C. B., Thompson, E. M., & Jaiswal, K. S. (2022). Partitioning ground motion uncertainty when conditioned on station data. *Bulletin of the Seismological Society of America*, *112*(2), 1060-1079.

Gelfand, A. E., Diggle, P., Guttorp, P., & Fuentes, M. (Eds.). (2010). *Handbook of spatial statistics*. CRC press.

Goulet, C. A., Kishida, T., Ancheta, T. D., et al. (2021). PEER NGA-east database. *Earthquake Spectra*, *37*(1_suppl), 1331-1353.

Guy, M.R., Patton, J.M., Fee, J., Hearne, M., Martinez, E., Ketchum, D., Worden, C., Quitoriano, V., Hunter, E., Smoczyk, G., & Schwarz, S. (2015). National Earthquake Information Center systems overview and integration. U.S. Geological Survey Open-File Report 2015-1120.

Hardebeck, J. L. (2020). Are the stress drops of small earthquakes good predictors of the stress drops of moderate-to-large earthquakes? *Journal of Geophysical Research: Solid Earth*, *125*(3), e2019JB018831.

Hearne, M., E. M. Thompson, H. Schovanec, J. Rekoske, B. T. Aagaard, & C. B. Worden (2019). USGS automated ground motion processing software, USGS software release, doi: 10.5066/P9ANQXN3.

Hirakawa, E., Parker, G. A., Baltay, A., & Hanks, T. (2023). Rupture directivity of the 25 October 2022 Mw 5.1 Alum Rock earthquake. *The Seismic Record*, *3*(2), 144-155.

Kleckner, J. K., Withers, K. B., Thompson, E. M., Rekoske, J. M., Wolin, E., & Moschetti, M. P. (2022). Automated detection of clipping in broadband earthquake records. *Seismological Research Letters*, *93*(2A), 880-896.

Landwehr, N., Kuehn, N. M., Scheffer, T., & Abrahamson, N. (2016). A nonergodic ground-motion model for California with spatially varying coefficients. *Bulletin of the Seismological Society of America*, *106*(6), 2574-2583.

Lavrentiadis, G. & Abrahamson, N.A. (2023). A nonergodic spectral acceleration ground motion model for California developed with random vibration theory. *Bull Earthquake Eng,* 21, 5265-5291. doi:10.1007/s10518-023-01689-9.

Lavrentiadis, G., Abrahamson, N. A., Nicolas, K. M., Bozorgnia, Y., Goulet, C. A., Babič, A., & Walling, M. (2023). Overview and introduction to development of nonergodic earthquake ground-motion models. *Bulletin of Earthquake Engineering*, *21*(11), 5121-5150.

Loth, C. & Baker, J. W. (2013). A spatial cross-correlation model of spectral accelerations at multiple periods. *Earthquake Engineering & Structural Dynamics*, *42*(3), 397-417.

Moschetti, M.P., B. Aagaard, S. Ahdi, J. Altekruse, O. Boyd, A. Frankel, J. Herrick, M. D. Petersen, P. Powers, S. Rezaeian, A. Shumway, J. Smith, W. Stephenson, E. Thompson, & K. Withers. (2024). The 2023 U.S. National Seismic Hazard Model: Ground-motion characterization for conterminous U.S. *Earthquake Spectra*. doi: 10.1177/8755293031223995.

Nychka, D., Furrer, R., Paige, J., & Sain, S. (2021). fields: Tools for spatial data. R package version 15.2, https://github.com/dnychka/fieldsRPackage.

Petersen M.D., Shumway A.M., Powers P.M., et al. (2023). The 2023 U.S. 50-State National Seismic Hazard Model: Overview and implications. *Earthquake Spectra*, doi:10.1177/87552930231215428.

Rekoske, J. M., Thompson, E. M., Moschetti, M. P., Hearne, M. G., Aagaard, B. T., & Parker, G. A. (2020). The 2019 Ridgecrest, California, earthquake sequence ground motions: Processed records and derived intensity metrics. *Seismological Research Letters*, *91*(4), 2010-2023.

Seekins, L. C. & Boatwright, J. (2010). Rupture directivity of moderate earthquakes in northern California. *Bulletin of the Seismological Society of America*, *100*(3), 1107-1119.

Stafford, P. J., Rodriguez-Marek, A., Edwards, B., Kruiver, P. P., & Bommer, J. J. (2017). Scenario dependence of linear site-effect factors for short-period response spectral ordinates. *Bulletin of the Seismological Society of America*, *107*(6), 2859-2872.

Stein, M. L. (1999). *Interpolation of spatial data: some theory for kriging*. Springer Science & Business Media.

Sung, C. H. & Abrahamson, N. (2022). A partially nonergodic ground-motion model for Cascadia interface earthquakes. *Bulletin of the Seismological Society of America*, *112*(5), 2520-2541.

Sung, C. H., Abrahamson, N. A., Kuehn, N. M., Traversa, P., & Zentner, I. (2023). A nonergodic ground-motion model of Fourier amplitude spectra for France. *Bulletin of Earthquake Engineering*, *21*(11), 5293-5317.

Thompson, E.M. (2018). An updated Vs30 map for California with geologic and topographic constraints (ver. 2.0, July 2022): U.S. Geological Survey data release, doi:10.5066/F7JQ108S.