Lawrence Berkeley National Laboratory
LBL Publications

Title

Performance evaluation of the USGS velocity model for the San Francisco Bay Area

Permalink

https://escholarship.org/uc/item/1gd908mx

Journal

Earthquake Spectra, 41(1)

ISSN

8755-2930

Authors

Pinilla-Ramos, Camilo

Pitarka, Arben

EERI, David McCallen M

et al.

Publication Date

2025-02-01

DOI

10.1177/87552930241270575

Peer reviewed

Performance evaluation of the USGS velocity model for the San Francisco Bay Area

Camilo Pinilla-Ramos[1] (iD)**, Arben Pitarka**[2] (iD)**,
David McCallen M. EERI**[1]**, and Rie Nakata**[1]

Abstract

In this study, we evaluated the performance of the United States Geological Survey velocity model developed for the San Francisco Bay Area (SFBA), version 21.1. The evaluation was performed through high-resolution three-dimensional physics-based ground motion simulations of seven small-magnitude earthquakes (ranging from magnitude 3.8 to 4.4) that occurred on the eastern side of the San Francisco Bay. The simulations were performed in the frequency range from 0 to 5 Hz with a minimum shear-wave velocity of 250 m/s, which allowed the capture of wave propagation effects of the near-surface soft materials that characterize local basins. Based on the direct comparison of Fourier amplitude spectra between recorded and simulated ground motions for more than 250 stations, we found that the velocity model generally performs well in the frequency range of 0.2–5 Hz. The median value of the Fourier amplitude residuals was found to be near zero for all seven earthquakes. The slight over-prediction of 0.2 log-natural units at frequencies above 3 Hz in our simulations was attributed to the potentially inaccurate representation of the source radiation pattern by a double-couple point source model, and simple representation of shallow small-scale underground structural complexity in the velocity model. Maps of spectral amplitude differences between the simulated and recorded data were used to identify areas responsible for systematic ground motion over-predictions or under-predictions. For example, while some sub-domains over soft sediments show over-prediction patterns, the block east of the Hayward fault is prone to exhibit patterns of under-prediction. These maps can be used to guide future refinements of the SFBA velocity model. Since our simulation methodology allows for the decoupling of the source and wave propagation effects, the ground motion data generated by our simulations can also be used to quantify the epistemic uncertainty due to the velocity model, in empirically based ground motion estimates for the SFBA.

[1]Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[2]Lawrence Livermore National Laboratory, Livermore, CA, USA

Corresponding author:
Camilo Pinilla-Ramos, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.
Email: camilopinillaramos@lbl.gov

Introduction

The San Francisco Bay Area (SFBA) is one of the most densely populated metropolitan regions on the West Coast of the United States. It is located in a complex geological environment forged by active tectonics, giving rise to regional natural hazards. Given its geologic features, a significant portion of the region's infrastructure, such as ports and bridges, is built over soft materials between very soft marine sediment and sandy fills placed to create artificial terrain on the sea and level streams at the foot of hills (Parks, 2019). These shallow materials can strongly affect the spatial amplification pattern of seismic waves (e.g. Ghofrani et al., 2013; Hartzell et al., 2016; Kramer, 1996; Lebrun et al., 2001; Rodgers et al., 2019), especially in areas with strong velocity contrasts between stiff rocks and soft basin sediments.

The Hayward fault is one of the main contributors to the seismic hazard in the SFBA, particularly on the eastern side of the San Francisco Bay (Aagaard et al., 2016). It is well known from probabilistic seismic hazard analysis that the seismic hazard of the Eastern Bay area is governed by earthquakes of magnitudes ranging from 6.8 to 7.2 on the Hayward fault at intermediate and high spectral frequencies (Field et al., 2015). The last large-magnitude earthquake that occurred on the Hayward fault was in 1868 (Toppozada et al., 2002). Lienkaemper et al. (2010) evaluated the recurrence interval between the last 12 large-magnitude events on the Hayward fault over the last 1900 years. They found that the mean recurrence interval varies between 138 and 161 years depending on the span of time analyzed, with a standard deviation around 60 years. The United States Geological Survey (USGS) has predicted a 72% probability of occurrence of a magnitude of at least 6.7 in the SFBA in the next 30 years, with the Hayward fault being the main contributor to the rate of occurrence (Aagaard et al., 2016).

The exponential growth of computational resources in the last decade has enabled numerically solving the elastodynamic wave equation for large domains while incorporating softer materials and has pushed physics-based ground motion modeling to higher frequencies. A key question related to these recent advancements in computational capabilities is how well we can predict ground motions at a specific site from a given set of earthquake scenarios. This question is especially important for large-magnitude events (above 6.5) and at short rupture distances (shorter than 20 km), where the recorded motion datasets lack observations (e.g. NGA-West2; Ancheta et al., 2014). The lack of such data also hinders the validation of velocity models that are used in physics-based simulations of strong ground motions. One option available for testing regional models consists of simulations of small local earthquakes. The large amount of available recorded data from small earthquakes in the SFBA has created the opportunity to test the performance of velocity models developed for the SFBA.

In this study, we investigate the performance of the USGS v21.1 SFBA community velocity model (USGS VM; Aagaard and Hirakawa, 2021) in simulations of ground motions from seven small local earthquakes on the Hayward and Calaveras faults, in the

frequency range 0–5 Hz. The moment magnitude of the seven earthquakes ranges from 3.8 to 4.4. The earthquakes were recorded by a dense network of stations throughout the SFBA. Because of the low ground motion amplitude, the soil response during these small earthquakes is expected to remain essentially linear at all recording sites. Due to their small rupture area, these small-magnitude earthquakes were modeled as pure double-couple point sources. Therefore, the differences between recorded and simulated waveforms can be mainly attributed to velocity model uncertainties associated with the complex shallow geology, including the near-surface material representation, and, to a lesser extent, to the seismic source parameterization, including epicenter location, depth, focal mechanism, and source-time function. The minimum shear-wave velocity is capped at 250 m/s in our simulations, and the adopted computational grid allows for accurate numerical modeling up to 5 Hz. The near-surface materials with relatively low velocity are expected to have a non-negligible impact on the simulated waveforms, especially at high frequencies (Hu et al., 2022; Taborda and Bielak, 2014). The anelastic wave propagation modeling was performed using SW4, a computer program based on an anelastic finite difference method of the fourth-order accuracy (Petersson and Sjögreen, 2012, 2015; Sjögreen and Petersson, 2012). SW4 allows the inclusion of surface topography through a near-surface curvilinear mesh.

The SFBA velocity model was originally created to simulate the main large-magnitude earthquakes that impacted the region in the last 120 years: the M7.9 1906 San Francisco and M6.9 1989 Loma Prieta earthquakes (Aagaard et al., 2008a, 2008b). Their approach for constructing the velocity model is based on the spatial distribution of the local geologic units, which is used to build a three-dimensional (3D) geologic model. The elastic material properties, such as shear and compressional wave velocities, density, and anelastic attenuation properties, were assigned based on empirical relationships from Northern California models applied on each geologic unit (e.g. Brocher, 2005, 2008). The anelastic properties are represented through quality factors. Several assessments of the velocity model based on first-wave arrival times have been used to refine the velocity model in the past, especially at low frequencies (below 0.5 Hz). After its first release, the velocity model underwent several modifications that improved its performance in several areas (Hirakawa and Aagaard, 2022). Figure 1 shows a schematic 3D view of the spatial distribution of the most important geologic units and blocks constituting the velocity model and the local fault system.

The velocity model performance was first evaluated through a qualitative comparison between recorded and simulated waveforms and then extended to a quantitative analysis. The overall modeling bias and its spatial variability were evaluated by comparing spectral amplitude residuals between recorded and simulated ground motions in the Fourier and spectral acceleration domains. The residuals' spatial distribution allowed for the recognition of systematic suboptimal prediction patterns of under- or over-prediction extracted from direct comparisons between recorded and simulated data at a large number of seismic stations across the SFBA. We also used the waveform duration as another ground motion parameter to evaluate the performance of the velocity model. The comparison of the waveform duration between recorded and simulated ground motions aids us in identifying areas where the velocity model may not fully represent small-scale geologic features.

Several authors (e.g. Hu et al., 2022; Savran and Olsen, 2019; Taborda and Bielak, 2014; Taborda et al., 2016) have adopted the Anderson's (2004) scale in quantitative evaluations of ground motion simulation performance. Anderson's scale provides a quantitative framework to evaluate the match between recorded and simulated waveforms.
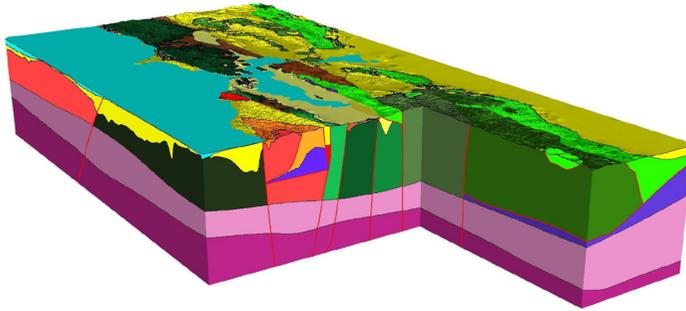
Figure 1. Conceptual and schematic 3D view of the San Francisco Bay Area geologic model (shared by Evan Hirakawa from the United States Geological Survey and available at https://www.usgs.gov/media/images/3d-geologic-model). The red lines display the main local faulting system in depth.

Nevertheless, its applicability is limited beyond a relative comparison between scores from different studies. Other authors have used goodness-of-fit plots to analyze the waveform matching in the spectral response and Fourier domains (Hernández-Aguirre et al., 2023; Pitarka et al., 2004; Smerzini et al., 2022). In contrast, the methodology used in our study quantifies the epistemic uncertainty induced by the velocity model, allowing direct estimation of the misfit between simulated and recorded data. This misfit is expected to be inherent to any simulated wave field propagated through this velocity model. Therefore, we can directly propagate the estimated epistemic uncertainty in seismic hazard analysis applications.

This article starts by describing the earthquakes used in this study and our approach to modeling the seismic sources. Then, we present the key features of the USGS velocity model, characterized by blocks delineated by the local faulting system. Once the seismic sources and the velocity model are described, we evaluated the velocity model performance using the aforementioned intensity measures that quantify the accuracy of the simulated ground motions. Finally, we discuss our results regarding the performance of the USGS velocity model and the potential implications of our findings on constraining wave path effects in ground motion simulations for large earthquakes in the SFBA.

Seismic source characterization

The seven small-magnitude earthquakes used in our simulations occurred between 2011 and 2021, with magnitudes ranging from 3.8 to 4.4. Six of the 7 events occurred on the Hayward fault, and one occurred on the Calaveras fault (Alum Rock event), close to the point of merging between these two faults. Simulating events on the western side of the San Francisco Bay would widen the azimuthal sampling. However, no events with magnitudes above 3.5 have occurred on the San Andreas fault in the last 20 years. Table 1 presents the location of each earthquake and Figure 2 shows a map of the earthquake epicenters, their focal mechanisms, and the stations used in our simulations.

The earthquakes were simulated by double-couple point sources. The data from Hirakawa and Aagaard (2022) were adopted for the source and focal mechanism characterization (strike, dip, rake, depth, epicenter location, and magnitude), except for the San Lorenzo earthquake, which occurred after the publication of their article. Hence, its double-couple source mechanism was characterized using the information reported by the

Table 1. Summary of the simulated earthquakes and source parameterization

Earthquake	Date	Latitude (°)	Longitude (°)	M_W	Depth (km)	Dip/strike/ rake (°)	f_c (Hz)	N_{rec}
Berkeley 2011	October 20, 2011	37.87	−122.253	3.94	8.0	88/144/176	1.8	124
El Cerrito	March 05, 2012	37.93	−122.29	4.08	8.3	80/145/140	3.0	91
Fremont	June 29, 2015	37.578	−121.974	4.0	8.0	80/155/151	1.5	140
Piedmont	August 17, 2015	37.84	−122.22	4.1	4.6	80/143/180	2.3	149
Berkeley 2018	January 04, 2018	37.855	−122.257	4.4	12.3	80/145/180	1.4	185
Alum Rock	April 16, 2018	37.43	−121.773	3.82	9.3	77/326/178	2.8	125
San Lorenzo	June 29, 2021	37.706	−122.122	3.9	9.2	73/154/166	2.5	130

f_c is the corner frequency and N_{rec} the number of recordings per earthquake.

Northern California Earthquake Data Center. Table 1 presents the source parameters adopted in modeling the seven earthquakes.

The recorded waveforms were downloaded using the mass-downloading module from the open-source ObsPy package written in Python (Beyreuther et al., 2010). A small number of recordings were discarded as being unreliable based on inconsistencies in the timing of waveform phases. The dataset used in our study has a total of 943 records over 7 events and 273 stations. The simulated and recorded waveforms were resampled to 0.01 s, and a Butterworth acausal bandpass filter of order 3 between 0.1 and 5 Hz was applied. After transforming the waveforms to the Fourier domain for the Fourier amplitude analysis, we smoothed them using a Hanning window of 13 points.

The Liu source-time function (STF; Liu et al., 2006) was adopted to model how the source releases energy over time. Being consistent with rupture dynamics, this slip rate function is not symmetric; it is characterized by a large initial peak and a gradual amplitude decay that represents the healing process of the rupture (e.g. Pitarka et al., 2021).

The corner frequency, f_c, is a fundamental parameter that controls the source's spectral shape. Therefore, its accuracy affects the simulation quality. Theoretically, f_c can be estimated directly from the Fourier amplitude spectra of recorded ground motions corrected for site and path effects (Kostrov and Das, 1988). Shearer et al. (2006) and Trugman et al. (2017) developed a methodology to identify f_c by removing systematic path and site effects in the Fourier domain. Using a similar technique, Trugman and Shearer (2018) estimated the f_c of earthquakes with magnitudes ranging from 1 to 4 that occurred in the SFBA. We adopted their estimated values for the events that occurred before 2018. We estimated f_c from the median value of Trugman and Shearer (2018) for the remaining earthquakes. Table 1 shows the corner frequency adopted in modeling the source of each earthquake. To illustrate the impact of the corner frequency on the wave field illumination, Figure 3 compares two Liu STFs with different corner frequencies (1 and 3 Hz) in the time and frequency domain. The amplitudes were normalized to enhance the differences between the two STFs. Under the same shear dislocation size, the 3-Hz STF releases the energy in a narrow time window, extending the radiated source energy to higher frequencies.
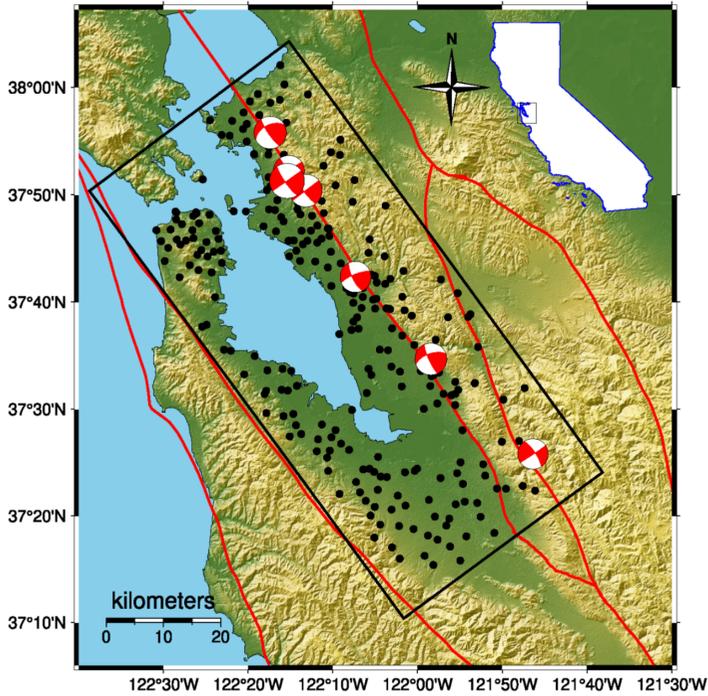
Figure 2. Map of the San Francisco Bay Area showing the locations of the stations (black dots), small earthquakes indicated by their focal mechanism, and the computational domain (black rectangle) used in this study. The red traces represent the regional fault system in the San Francisco Bay Area. The inset at the top-right corner shows the specific area of study with respect to the California state.
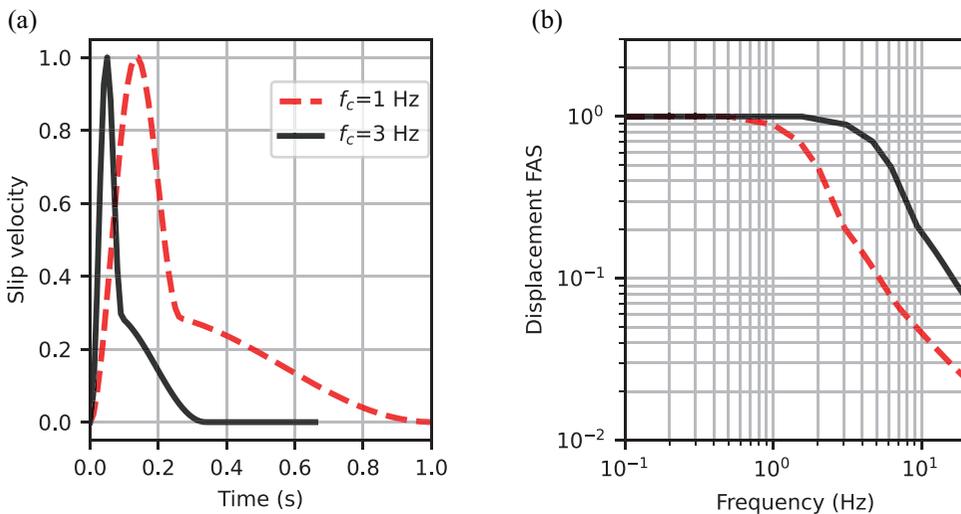

Figure 3. Example of Liu source-time functions for two corner frequencies: 1 and 3 Hz. (a) Slip velocity of the source-time function in the time domain. (b) Displacement spectra of the source-time function in the Fourier domain. The source-time function amplitude was normalized in the time and Fourier domain.
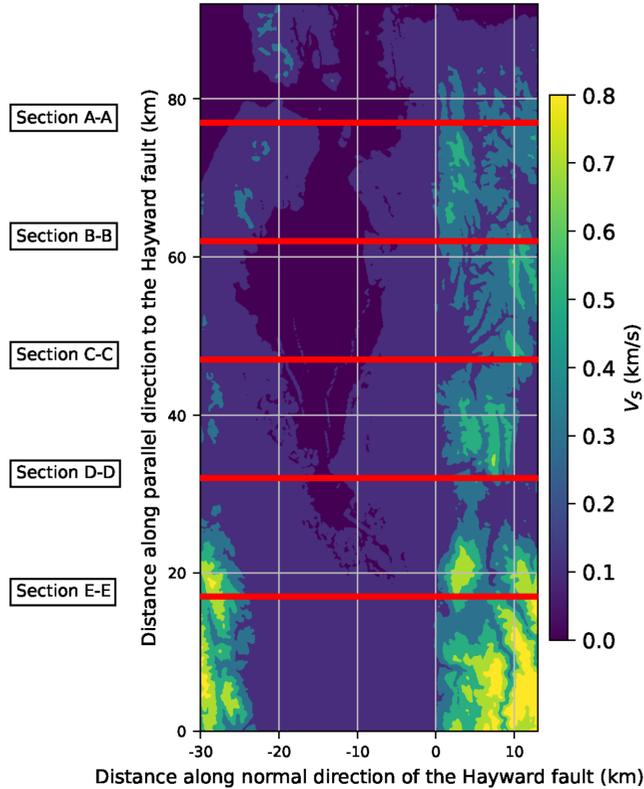
Figure 4. Map view of the free surface shear-wave velocity in the simulation domain, extracted from the United States Geological Survey velocity model. The thick E-W red lines indicate the location of the velocity model cross-sections shown in Figure 5.

Velocity model

In our simulations, we used a sub-domain of the detailed USGS v21.1 SFBA VM (Aagaard and Hirakawa, 2021). This sub-domain includes the sedimentary basin and the hills on the east and west sides of the Bay Area. The free surface shear-wave velocity is shown in Figure 4. The Hayward fault causes an abrupt change in the materials' seismic properties along most of the fault trace, which can be seen in the map view of the free surface shear-wave velocity. Toward the west of the Hayward fault, the surficial materials are softer, representing soft marine sediments and artificial fills. The eastern side of the Hayward fault has stiffer materials that have been uplifted by the local tectonic activity.

The SFBA region has complex geological structures that are expected to have a significant effect on wave propagation across the Bay Area. The northern area is characterized by two blocks separated by the Hayward fault, as it is shown in the cross-sections A-A, B-B, and C-C in Figure 5. The block located west of the Hayward fault has a shallow layer of soft Quaternary marine sediments, underlaid by sedimentary basins that extend up to 1 km depth (see cross-sections B-B and C-C) and then by sequences of sandstones and sedimentary metamorphic rocks, the Franciscan complex (Bailey et al., 1964; Phelps et al., 2008). The eastern block is constituted mainly by sequences of sedimentary rocks deposited in the late Mesozoic Era, the Great Valley complex (Bartow and Nilsen, 1990). The
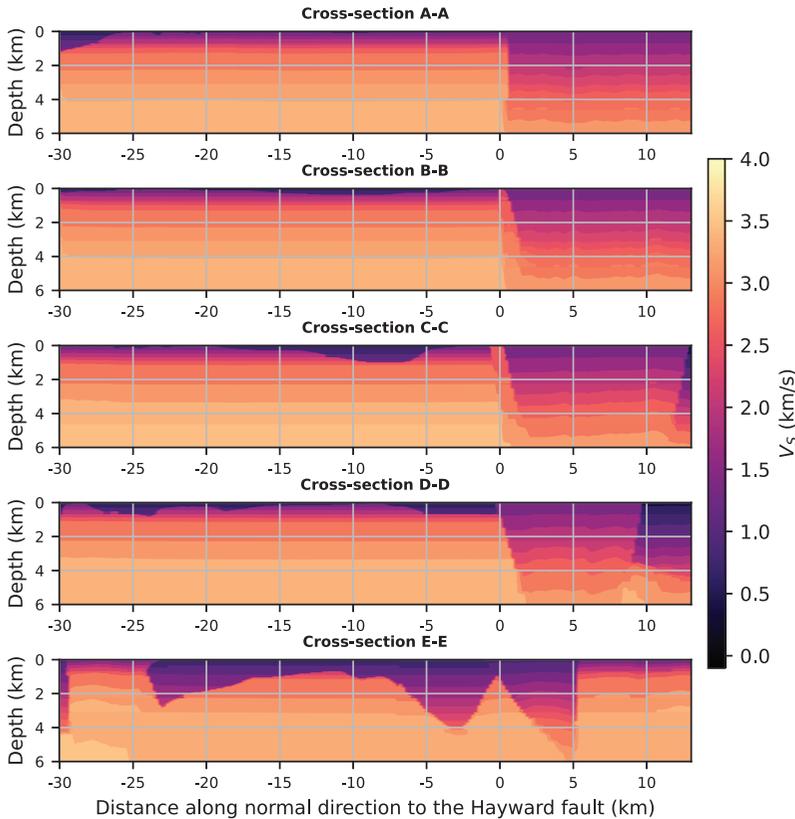
Figure 5. Northeast-southwest vertical cross-sections of the velocity model. The location of the cross-sections is indicated in Figure 4. The Hayward fault can be identified at 0 normal distance, through the strong velocity contrast between the western and eastern block. The Calaveras fault can be identified in cross-sections C-C and D-D as the dipping block around coordinate 10 km. Despite the simulations being run with topography, these cross-sections do not include the topography.

shallowest layers in the eastern block are characterized by stiffer materials than those in the western block. Due to this geological contrast, the eastern block tends to be stiffer at the surface than the western block, with shear-wave velocities usually above 500 m/s, but at larger depths, it has lower shear-wave velocities than the western block.

As shown in the cross-sections D-D and E-E in Figure 5, the local geology in the southern area of the western block is more complex than that in the northern area. The Calaveras fault bounds the eastern block from the east, shortening its width toward the south as seen in the cross-sections C-C, D-D, and E-E. The San Jose and Santa Clara regions, shown in the cross-section E-E, between horizontal coordinates -24 to 5 km, are underlaid by deeper sedimentary basin structures with an irregular geometry. The effects of these structures on the simulated waveforms are mostly manifested by basin wave reverberations and secondary surface waves that increase the waveform complexity and ground motion duration, as Frankel et al. (2001) showed by deploying a seismic array in the southern Bay Area.
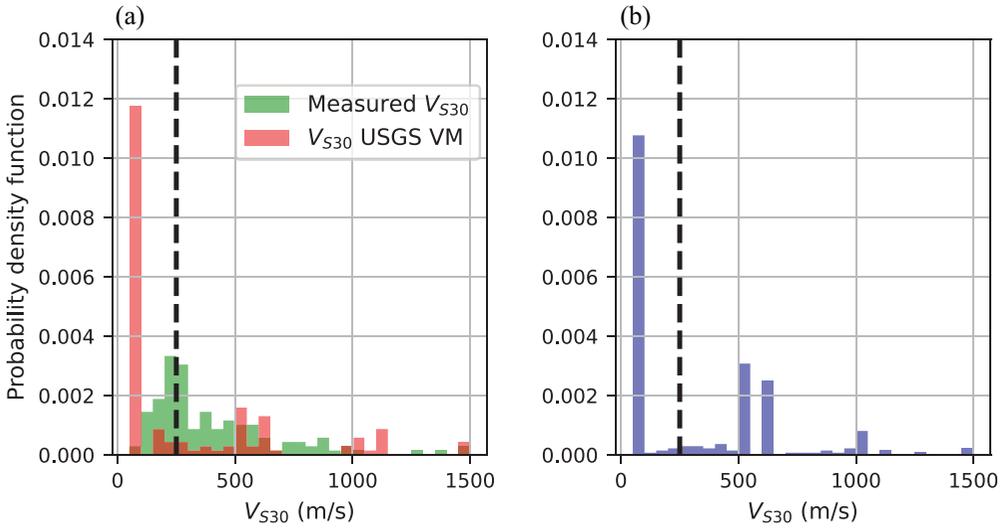
Figure 6. (a) United States Geological Survey velocity model-based (red bars) and measured (Tehrani et al., 2022; green bars) V_{S30} at locations throughout our computational domain. (b) V_{S30} computed at seismic stations considered in our simulations. The dashed line indicates the minimum V_s used in our simulations.

In addition to the potential misrepresentation of the deep underground structure, another source of epistemic uncertainty in community velocity models is the simplification of the near-surface materials (e.g. Taborda and Bielak, 2014). The seismic properties of near-surface materials at shallow depths are usually represented by the time-averaged shear-wave velocity in the first 30 m, V_{S30} (Kamai et al., 2016). To a first approximation, evaluating the differences between the V_{S30} in the velocity model and the estimated or measured one provides insights into how well the near-surface material properties are represented in our simulations. To gain insight into such differences, in Figure 6a, we compared V_{S30} values measured at sites within our computational domain (Tehrani et al., 2022) with those derived from the USGS velocity model at the same location. Conditioned by the near-surface grid spacing, the V_{S30} in our model was computed by sampling shear-wave velocity values in sublayers of 5 m thickness between the surface and a 30-m depth. The similarity of the V_{S30} distribution estimated throughout the computational domain and that estimated at seismic stations only, shown in Figure 6, demonstrates that the station-based sampling of the V_{S30} provides a good spatial representation of the V_{S30} distribution throughout the entire computational domain. The comparison of the V_{S30} distributions shown in Figure 6 suggests that the USGS velocity model needs improvements to better represent the near-surface shear-wave velocity, especially in the Bay basin where the modeled shear-wave velocity is extremely low. Nevertheless, because we cap the minimum shear-wave velocity at 250 m/s, such low shear-wave velocity values do not significantly affect our simulations.

Our ground motion simulations were performed with SW4, a parallelized computer program that employs an elastic finite difference method to solve the elastic wave equation in heterogeneous media. SW4 allows for the inclusion of a curvilinear mesh with vertical refinement to model the surface topography (Petersson and Sjögreen, 2023). Our simulations included the surface topography, which has a resolution of 100 m at the surface and

a minimum shear-wave velocity of 250 m/s. A minimum grid spacing of 6.25 m was required to ensure proper numerical accuracy for frequencies up to 5 Hz, thus ensuring 8 points per wavelength in all our mesh. The wave attenuation is modeled through a linear visco-elastic model characterized by the quality factors (Q_p and Q_s for compressional and shear waves, respectively). The quality factors were obtained directly from the characterization included in the velocity model by the USGS. Under these model settings, the wall time of one simulation performed on 128 nodes on Perlmutter, a 70-PF (Peta flops) Graphics Processing Unit (GPU)-accelerated super-computer at the National Energy Research Super Computer Center at Lawrence Berkeley National Laboratory, is 2.5 h.

Velocity model performance

Qualitative analysis

Our qualitative analysis of the velocity model performance in ground motion simulations of recorded earthquakes was focused on characterizing the waveform fit in terms of amplitude and wave phases by using three categories: good, fair, and poor. They were based on three criteria. The first criterion is the similarity of wave phases, the second is the amplitude match in the time and frequency domain, and the third criterion is the waveform and amplitude fit considering all three components of ground motions. A good waveform fit means that, overall, the recorded and simulated ground motions satisfy all three criteria. A fair fit means that at most two of the three criteria are satisfied. A poor fit is a case where two or more criteria are not satisfied.

As illustrated in Figure 7, the quality of the simulated waveforms at a given station (J056) is event-dependent, implying a wave path dependency. The waveform fit at this station is good for four earthquakes and fair for the three others. This site is located on the San Francisco Peninsula, which, as shown later, is a region that manifests azimuthally dependent wave path effects, especially at high frequencies. Nevertheless, the simulated ground motions at this station capture most of the recorded waveform complexities associated with basin effects along wave paths crossing the San Francisco Bay. Figure 8 demonstrates that for a given seismic source, in this case the Berkeley 2018 earthquake, the waveform fit varies from good to poor depending on the station location relative to the seismic source. A similar trend is also seen in the frequency content of the waveforms, as shown in Figure 9, where we compare the recorded and simulated Fourier amplitude spectra of the waveforms shown in Figure 8.

As a first-order comparison, the spatial pattern of the waveform matching can be used as an indicator of the quality of the velocity model along paths across the SFBA. This is demonstrated in Figure 10, which displays the simulated ground motion quality at all recording stations for all seven earthquakes. Although the overall spatial pattern of the simulated ground motion quality varies from one earthquake to the other, a close inspection reveals areas where the simulated ground motion quality is consistent among all scenarios. For example, the quality of the simulated waveforms is good at most of the stations located on the west edge of the bay, including some sites in the San Francisco Peninsula, for all the earthquakes. This demonstrates that the model is capable of reproducing wave propagation effects along paths across the bay. A similar observation can be made for most of the sites located on the east side of the bay, where the number of sites with good or fair quality waveform fit is much larger than those with a poor fit. It is important to note
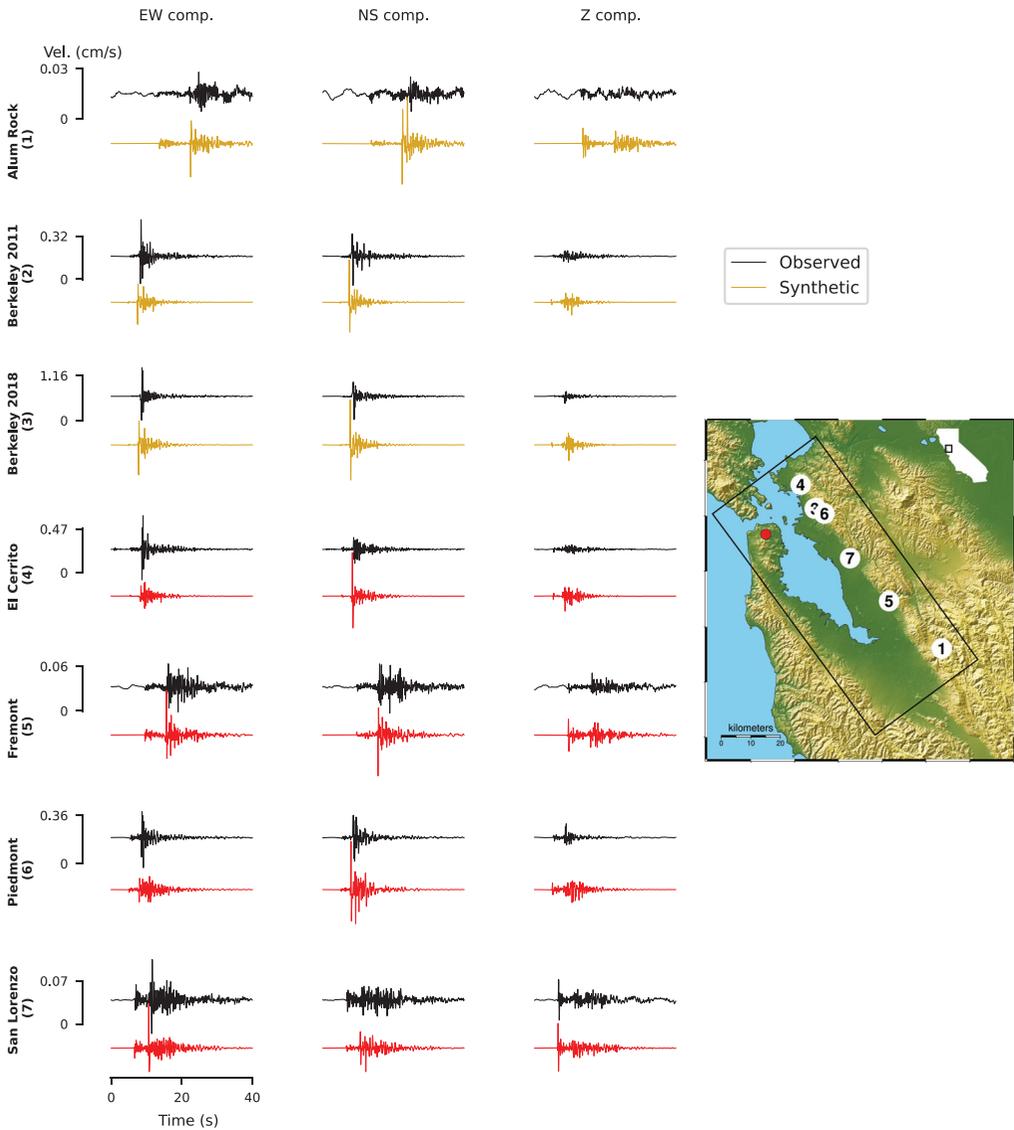
Figure 7. Waveform match quality at station J056. Comparison of recorded (black) and simulated (different colors) waveforms of ground motion velocity for the seven earthquakes, band-passed filtered between 0.2 and 5 Hz. The earthquake's name is indicated on the left of each trace. Red color indicates good, and yellow indicates fair waveform match quality. The amplitude for the plots was normalized within each earthquake, showing the maximum value between the six waveforms of each earthquake in the vertical axis.

that the simplification of the source radiation pattern applied to the earthquake's point source modeling can impact the simulation quality. The waveform quality at sites within the epicentral area and close to the nodal planes is sensitive to the focal mechanism, where a small misrepresentation of the focal mechanism can significantly affect the ground motion amplitude.
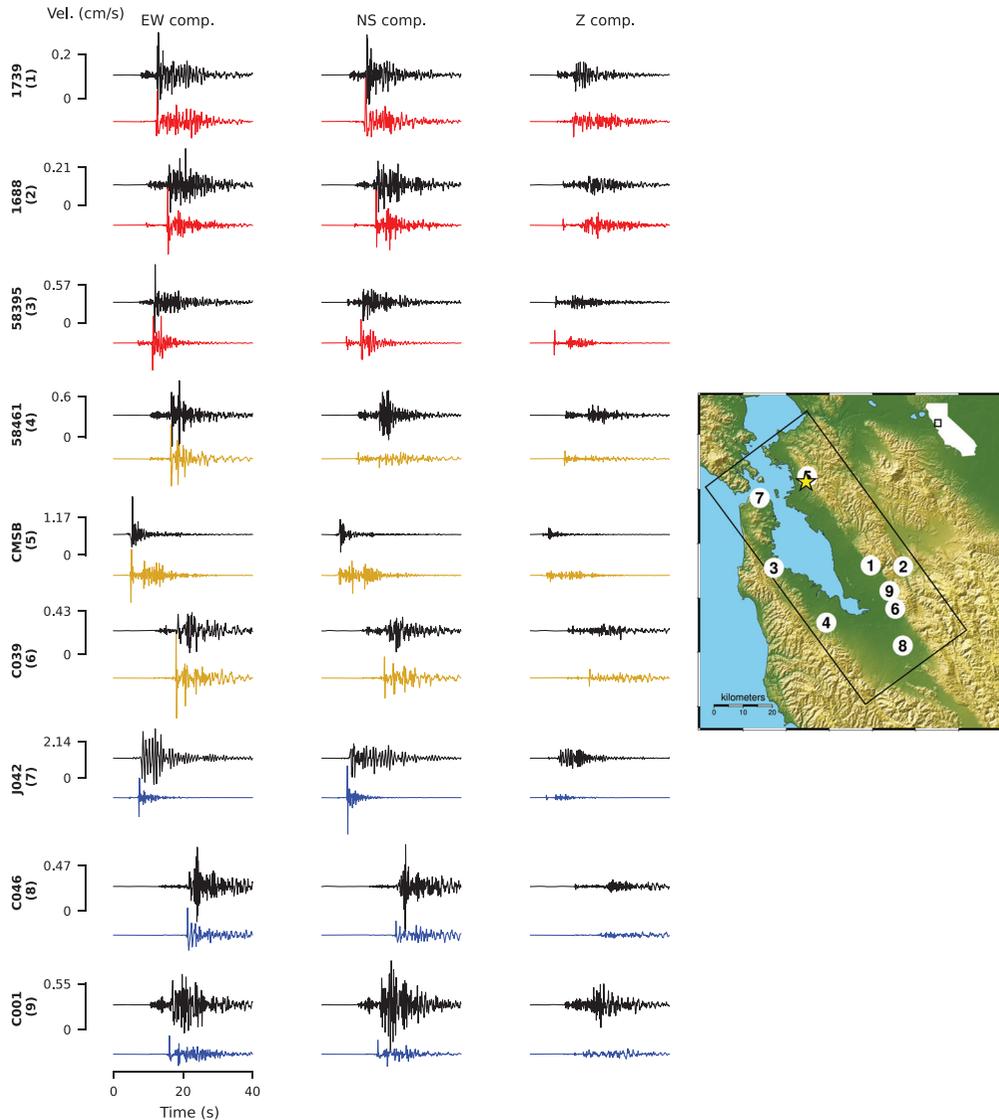
Figure 8. Comparison of recorded (black) and simulated (different colors) waveforms of ground motion velocity band-passed filtered between 0.2 and 5 Hz for the Berkeley 2018 earthquake at 9 stations. The station name is indicated on the left of each trace. Red, yellow, and blue colors correspond to good, fair, and poor waveform matching quality, respectively. The amplitude for the plots was normalized within each station, showing the maximum value between the six waveforms of each earthquake in the vertical axis.

Overall velocity model performance

We evaluated the overall systematic bias between recorded and simulated data in the Fourier domain in the frequency range of 0–5 Hz to provide a quantitative assessment of the simulation performance. The Fourier amplitudes allow for a direct comparison of the frequency content of the recorded and simulated waveforms related to both source and wave propagation effects. The residuals between recorded and simulated waveforms in the
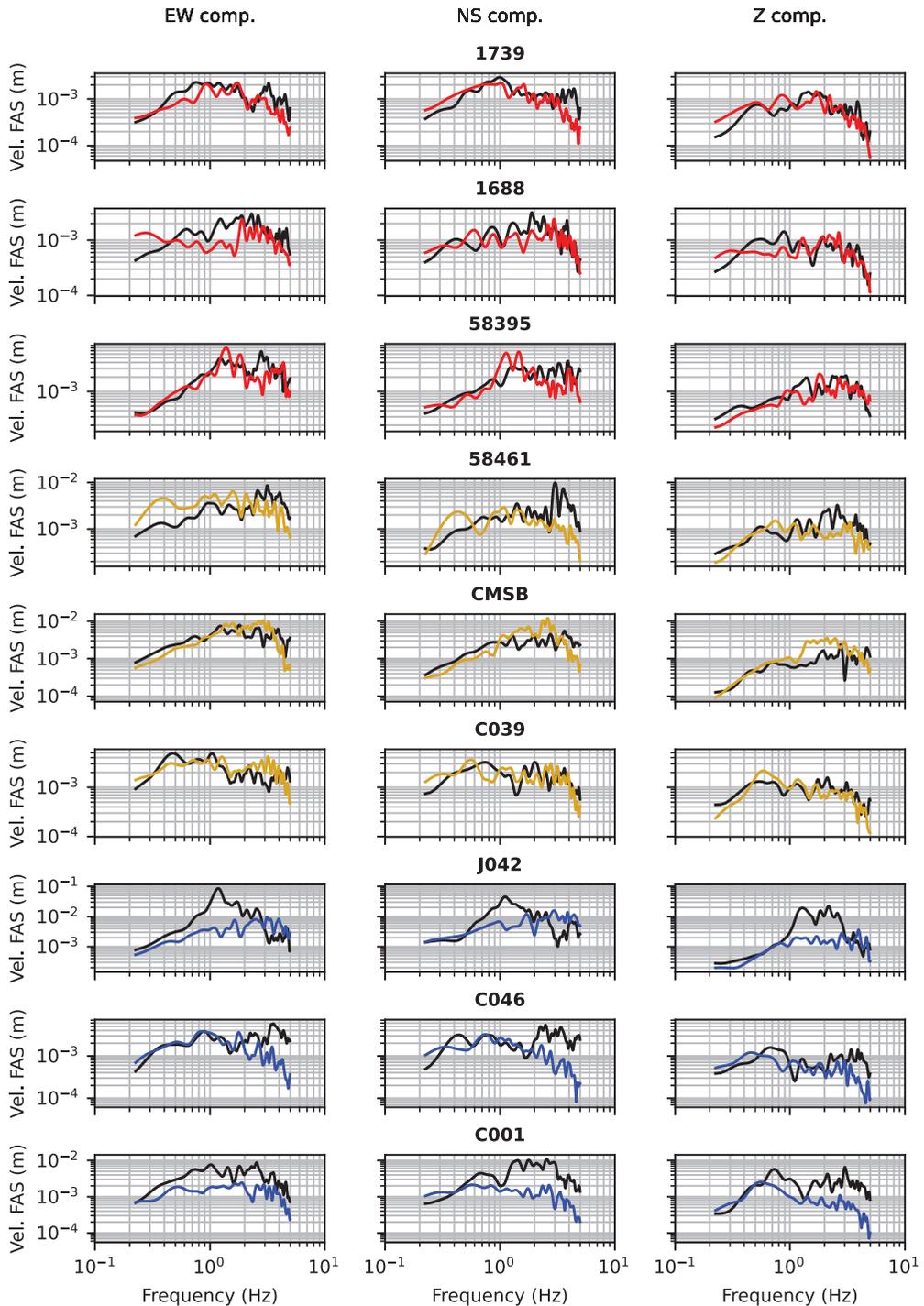
Figure 9. Comparisons of recorded (black) and simulated (different colors) Fourier amplitude spectra of ground motion for the Berkeley 2018 earthquake at the 9 stations shown in Figure 8. The station name is indicated above the central plot of each row. Red, yellow, and blue colors correspond to good, fair, and poor waveform matching quality, respectively.
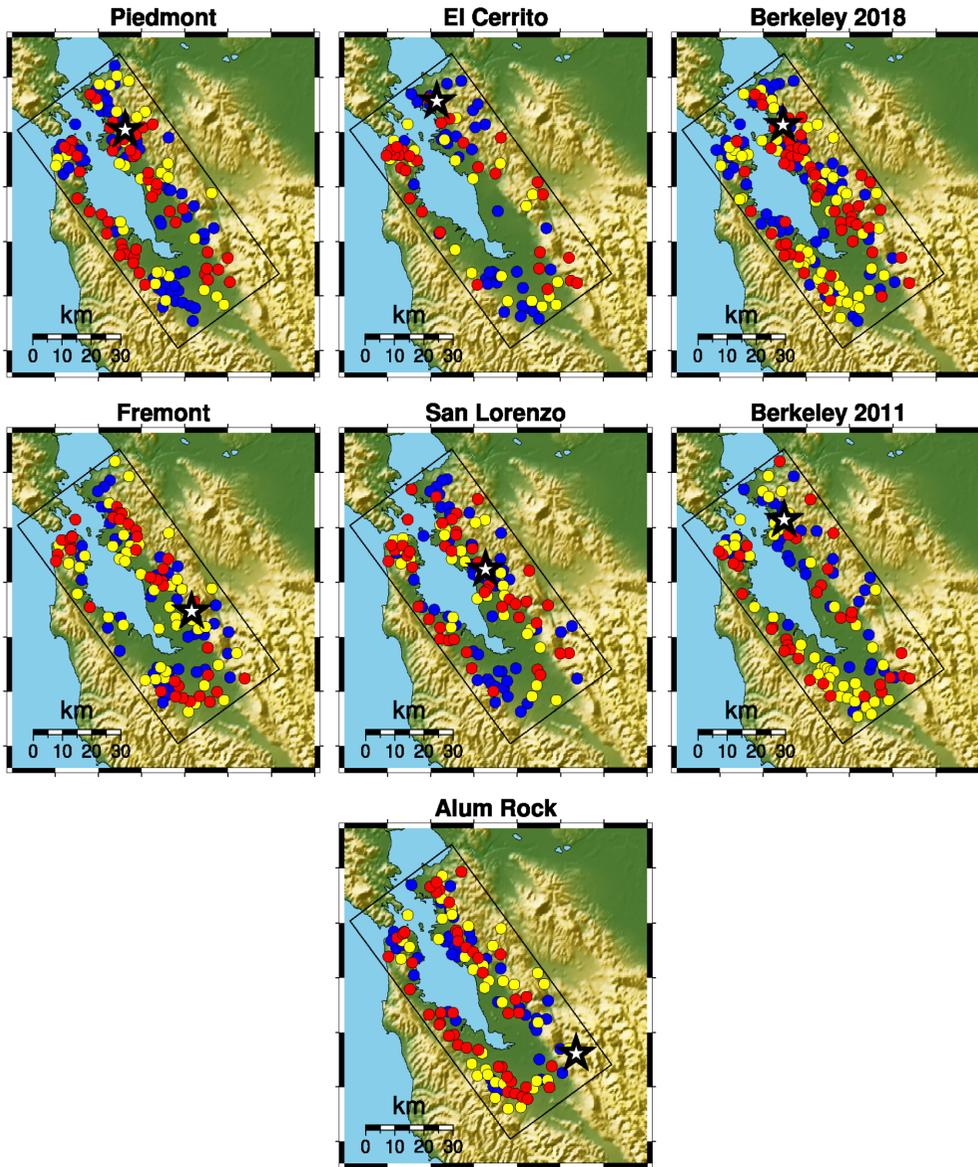
Figure 10. Waveform matching quality obtained for seven earthquakes at seismic stations indicated by colored circles. Red, yellow, and blue colors correspond to good, fair, and poor quality, respectively. The earthquake epicenter is indicated by the black star, and the earthquake's name is indicated at the top of each panel.

log-natural scale are shown in Figure 11 for the horizontal component for each earthquake, adopting the format described in Goulet et al.'s (2015) study. Berkeley 2018, El Cerrito, and San Lorenzo show a near-zero bias between 0.2 and 5 Hz. The residuals of Piedmont, Fremont, Berkeley 2011, and Alum Rock fluctuate around zero, with a maximum over-prediction peak of 0.4 log-natural units (LN units) at frequencies above 3 Hz. The same comparison of residuals for the vertical component is shown in Figure 12. The vertical component shows mean residuals with larger between- and within-event
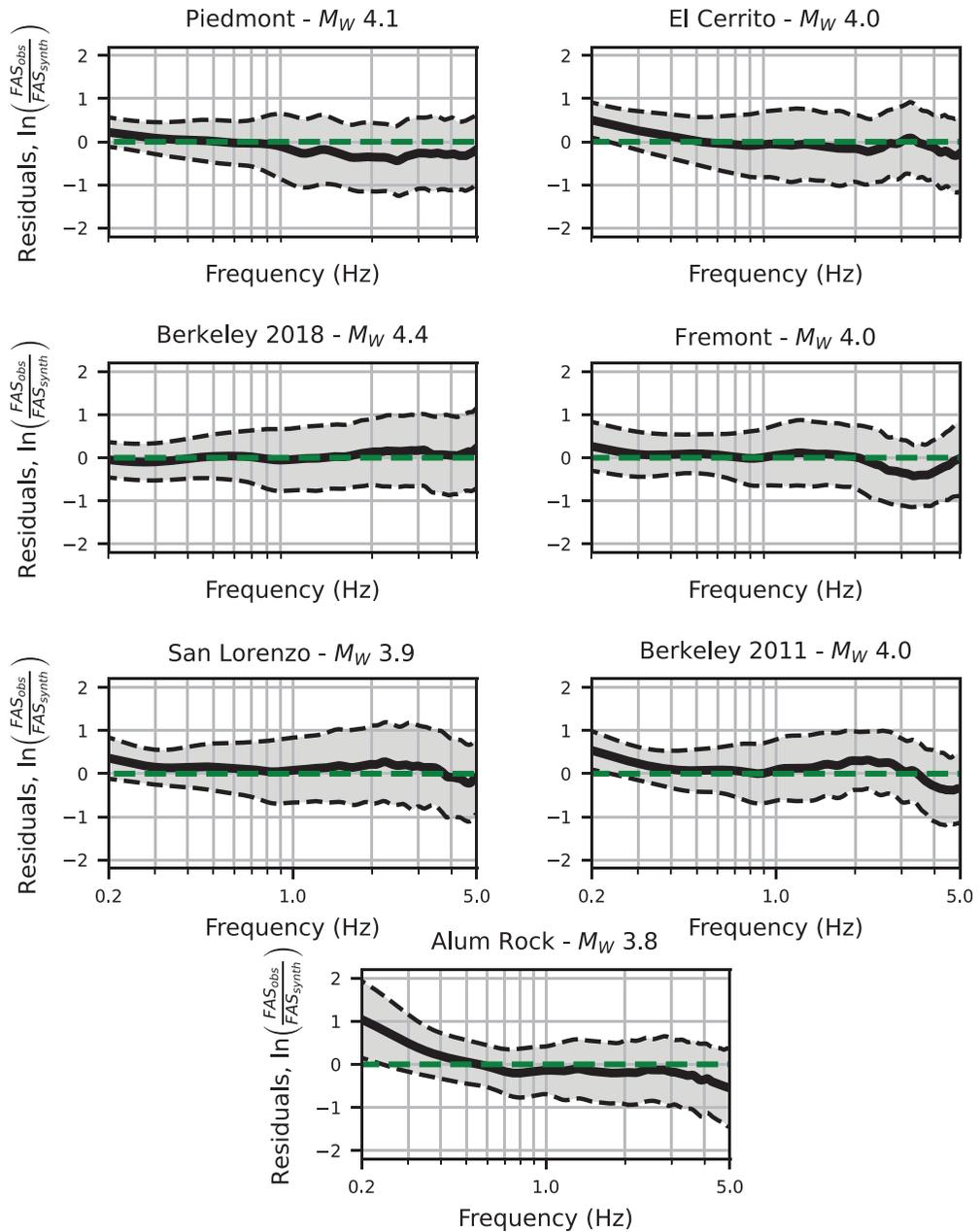
Piedmont - M_W 4.1

El Cerrito - M_W 4.0

Berkeley 2018 - M_W 4.4

Fremont - M_W 4.0

San Lorenzo - M_W 3.9

Berkeley 2011 - M_W 4.0

Alum Rock - M_W 3.8

Figure 11. Comparison of the within-event Fourier amplitude residuals for the horizontal component of the ground motion averaged over the recording stations for each event. The solid line is the bias expressed by the log-natural of the amplitude ratio between the recorded and simulated motions as a function of frequency. The shaded area corresponds to the 16th–84th residuals' percentile. Each panel corresponds to each of the seven earthquakes. The name of the earthquake is indicated on top of each panel.

fluctuations around zero; nevertheless, the maximum within-event bias does not exceed 0.5 LN units. Some events show a break point in the mean residual at low frequencies, with a trend toward under-prediction in the bias (positive bias), as for example Alum
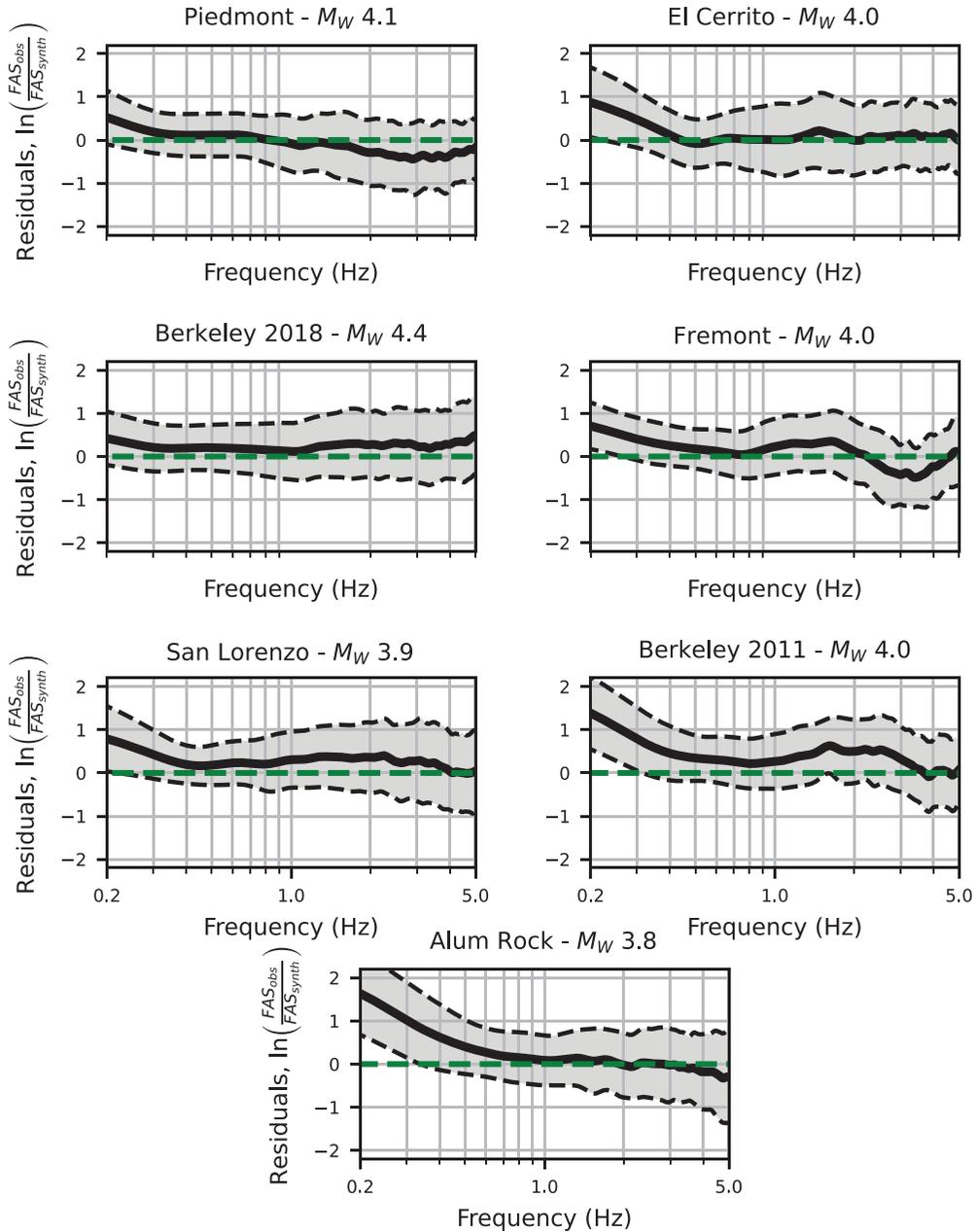
Figure 12. Same as Figure 11, but for the vertical component.

Rock with a bias of 1 and 1.5 LN units for the horizontal and vertical components, respectively. This mismatch is significant also for the El Cerrito, San Lorenzo, and Berkeley 2011 earthquakes. Because this mismatch is earthquake-dependent and is more pronounced at long-distance stations where the source-related long-period waves have relatively low amplitudes, our interpretation is that it is due to large-amplitude long-period ambient noise at the time when these three earthquakes occurred. This earthquake-
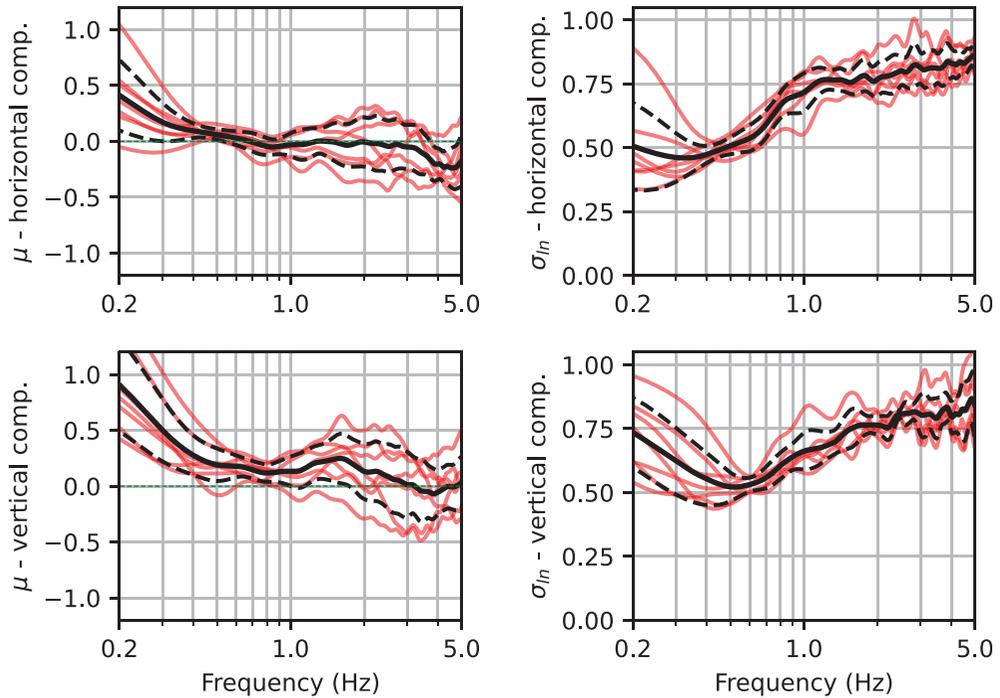
Figure 13. Mean (left panels) and standard deviation (right panels) of the within-event Fourier amplitude residuals averaged over seven simulated earthquakes. The upper and lower panels correspond to the horizontal and vertical components, respectively. The black solid lines are the mean of the 7 events, and the dashed lines correspond to the 16th–84th percentile. The red traces show the individual standard deviations for each earthquake.

dependent mismatch has also an impact in the standard deviation, which tends to increase below 0.4 Hz.

It is important to note that, as shown in Figure 13, the simulations of the horizontal ground motion, when averaged across all seven earthquakes, have a near zero bias up to 3 Hz. This figure shows the mean and standard deviation of the horizontal and vertical components for the seven simulated earthquakes. The vertical component's mean bias is centered at zero except for a frequency band between 1 and 3 Hz, where a slight underprediction of around 0.25 LN units appears. We believe that the mismatch between the recorded and simulated ground motion at frequencies below 0.5 Hz, observed in almost all considered earthquakes, is caused by the relatively high local ambient noise.

The standard deviation of both components increases at higher frequencies, reflecting the impact of the mismatch between recorded and simulated ground motion due to small-scale heterogeneities, especially near the free surface (Lu and Ben-Zion, 2022). This behavior was also observed by Taborda et al. (2016) in their broadband simulations of the Chino-Hills earthquake in the Los Angeles basin using a 3D regional velocity model. They showed that the standard deviation between the recorded and simulated ground motions increases at higher frequencies. Another interesting finding in our analysis is that the standard deviation of the residuals is similar among all considered earthquakes, especially for the vertical component; the within-event standard deviation converges to a narrow band

among the simulation of the 7 events. This implies that the accuracy obtained in our simulations and modeling methodology is stable among the 7 events.

Spatial distribution of spectral acceleration residuals

In this section, we analyze the spatial distribution of the spectral acceleration residuals. Maps of the spectral acceleration residuals for the horizontal component at 0.5, 1.0, and 4.0 Hz (2.0, 1.0, and 0.25 s) are illustrated in Figures 14 to 16, respectively. These maps show clear patterns of under- and over-predictions within each event at three different spectral frequencies. For example, a roughly circular region surrounding the San Lorenzo earthquake epicenter shows consistent patterns of under-prediction. On the other hand, the southernmost corner of our domain consistently displays over-predictions, suggesting that the soft sediments in the southwestern part of the SFBA can be singled out as areas where the model needs to be improved.

The spatial correlation of the residuals is a manifestation of poorly characterized path effects. For example, the West San Jose area (southwest corner of the San Francisco basin) exhibits strong path effects for different earthquakes at these three frequencies. The simulation of the San Lorenzo earthquake produces the largest over-prediction in this area, with a source site azimuthal variation between 150° and 180°. In contrast, the simulation of the Alum Rock earthquake produces the largest under-prediction with azimuths varying between 215° and 270°. This is consistent with Stidham et al. (1999) and Dolenc and Dreger (2005), who using 3D physics-based ground motion simulations and seismic array analysis, respectively, found strong basin effects in this region of the SFBA. The San Francisco Peninsula also experiences azimuth-dependent path effects. They are observed mostly at high frequencies. Ground motions of events located to the east of the basin are over-predicted at frequencies above 4 Hz (e.g. San Lorenzo and Fremont earthquakes), and for events located to the south, along the Hayward fault, the over-prediction gradually becomes relatively small (e.g. Alum Rock earthquake), and we can even observe a slight under-prediction of the ground motion amplitude. It appears at high frequencies, suggesting shallow geologic structures may cause this ground motion azimuthal variability.

To get a deeper insight into these 3D path effects, in Figure 17 we show the example of waveforms recorded and simulated at station 1796 in Santa Clara for the seven earthquakes. The recorded waveforms from earthquakes coming from the north to northwest exhibit phases with comparable amplitudes after the main shear-wave phase arrivals. In contrast, the waveforms coming from the northeast to the east (Alum Rock and Fremont) induce a smaller waveform reverberation in the recorded waveforms. The simulated data captured some of these phases after the main shear wave arrival but lacked the recorded waveforms' complexities and amplitudes. This effect can also be observed in the simulated Fourier amplitude spectra (Figure 18). The shape of the simulated spectra from several events tends to be similar, while the recorded spectral shapes show more complex event-dependent variations, as relative spectral amplitude fluctuations between the recorded and simulated waveforms at a fixed frequency for different source–receiver pairs exist. The amplitude of the recorded waveforms at a given frequency is larger than the simulated waveforms for some earthquakes, while the opposite behavior is observed for others. A better representation in the velocity model of the 3D structures that induce these complexities is necessary for reproducing these recorded path effects.

The difference in the residual distributions obtained for the Piedmont and Berkeley 2018 and 2011 earthquakes indicates that earthquakes in similar locations but with
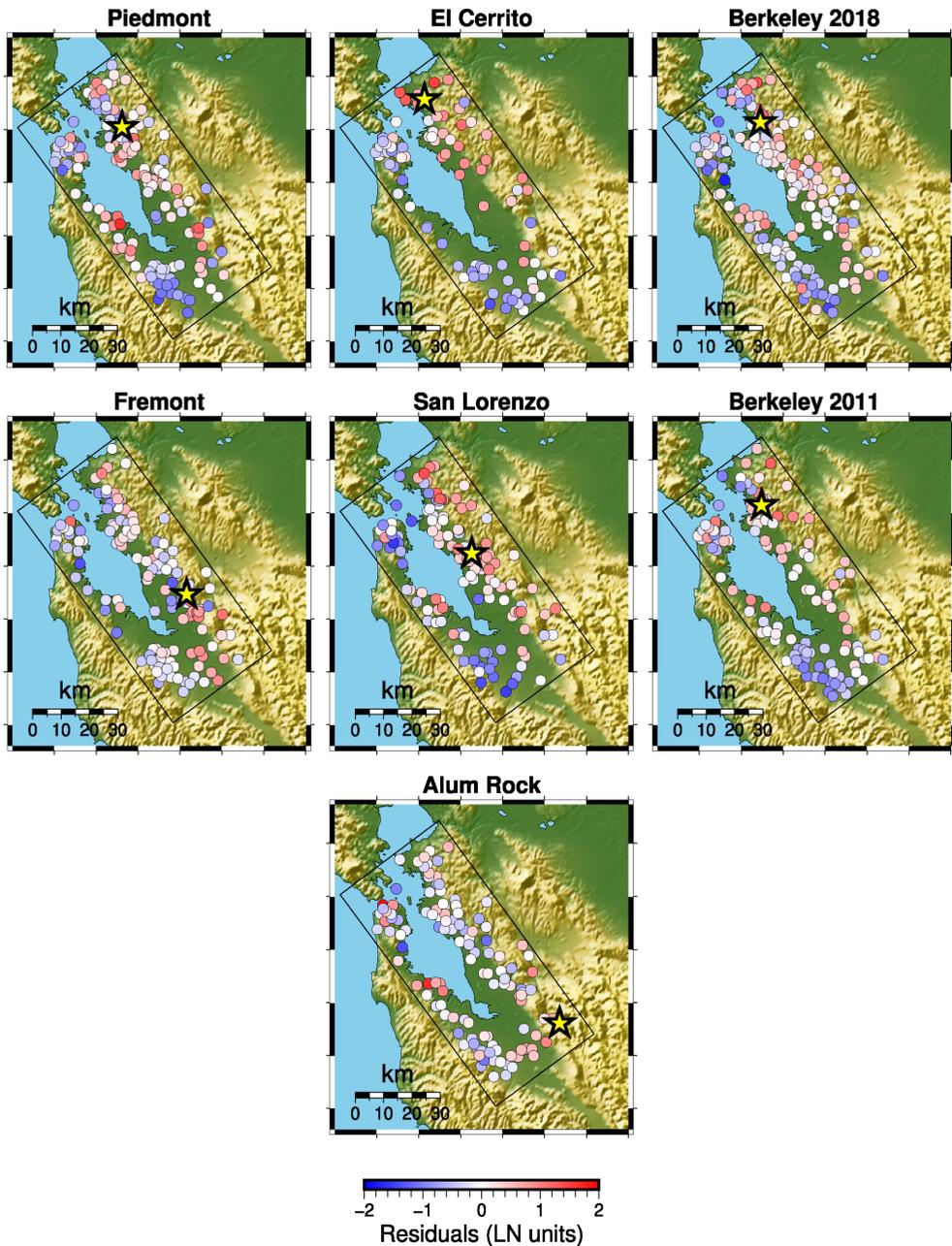
Figure 14. Total spectral acceleration residuals at 0.5 Hz computed at each station indicated by colored circles. The colored circles represent the residual and the yellow star the epicenter location. LN = log-natural.

different hypocentral depths can result in different wave propagation effects. This interesting observation can be made by analyzing the residual distribution from the previous maps and Figures 17 and 18, showing that the path effects are not only azimuthally dependent but also event-depth dependent. The close proximity of their epicenter could lead to
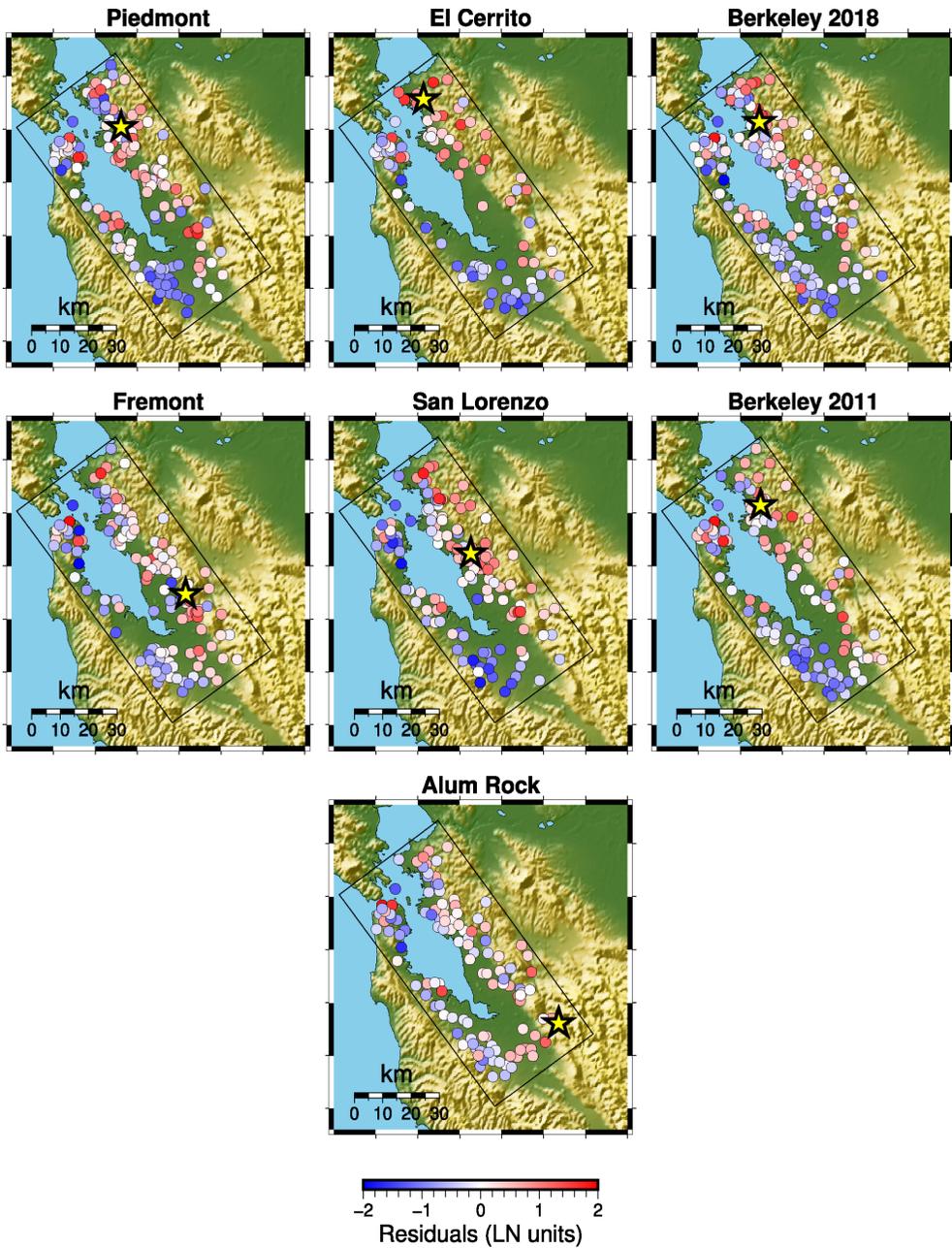
Figure 15. Total spectral acceleration residuals at 1.0 Hz computed at each station indicated by colored circles. The colored circles represent the residual and the yellow star the epicenter location. LN = log-natural.

the assumption that their impulsive response should be similar. However, the difference in their hypocentral depths (ranging from 4.5 to 12.3 km) impacts the wave fields, showing poorly represented depth-dependent path effects.
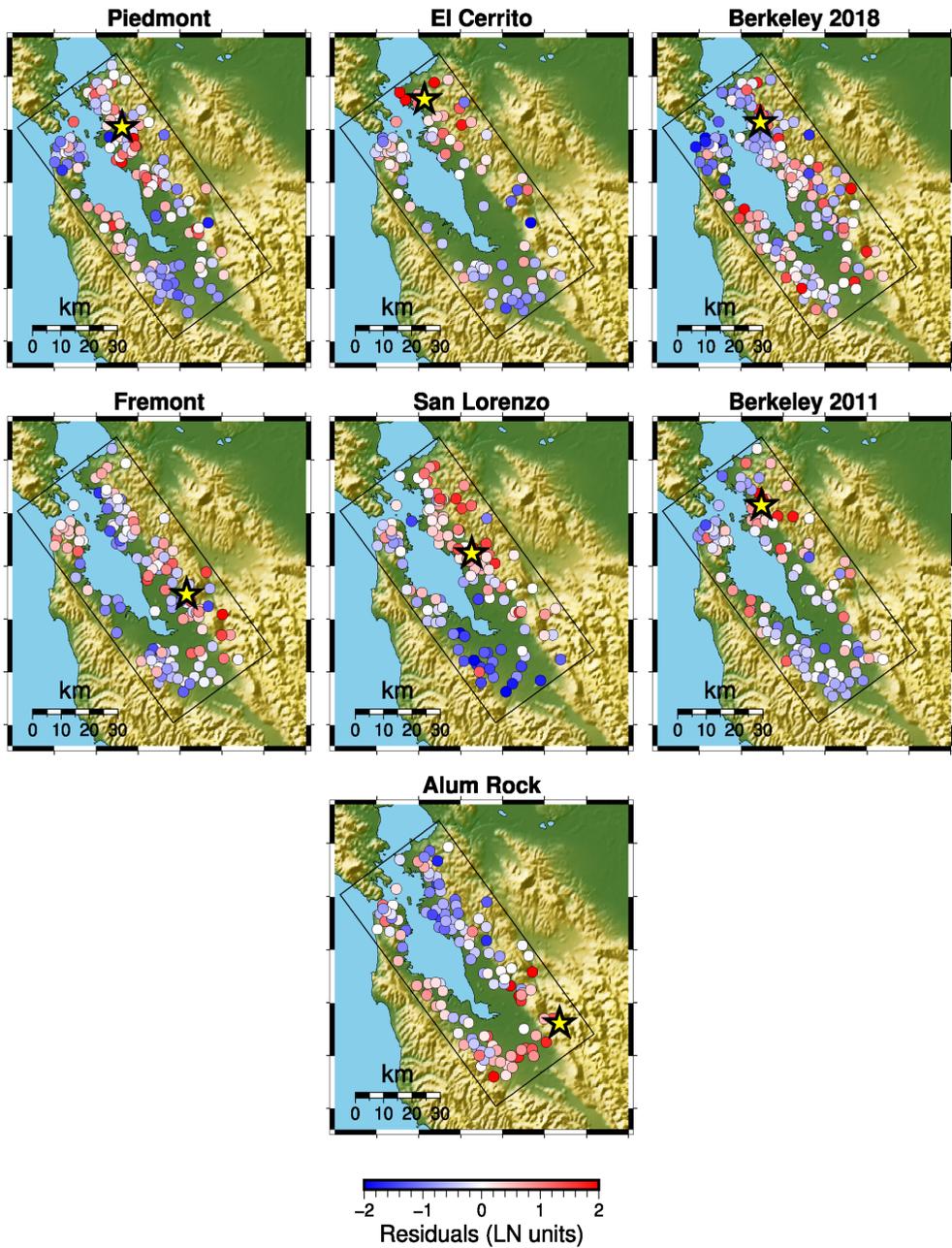
Figure 16. Total spectral acceleration residuals at 4.0 Hz computed at each station indicated by colored circles. The colored circles represent the residual and the yellow star the epicenter location. LN = log-natural.

Modeling of spatial patterns of suboptimal predictions

One of the valuable features of this extensive database of small event observations is the opportunity to assess areas for potential improvement in the existing regional velocity
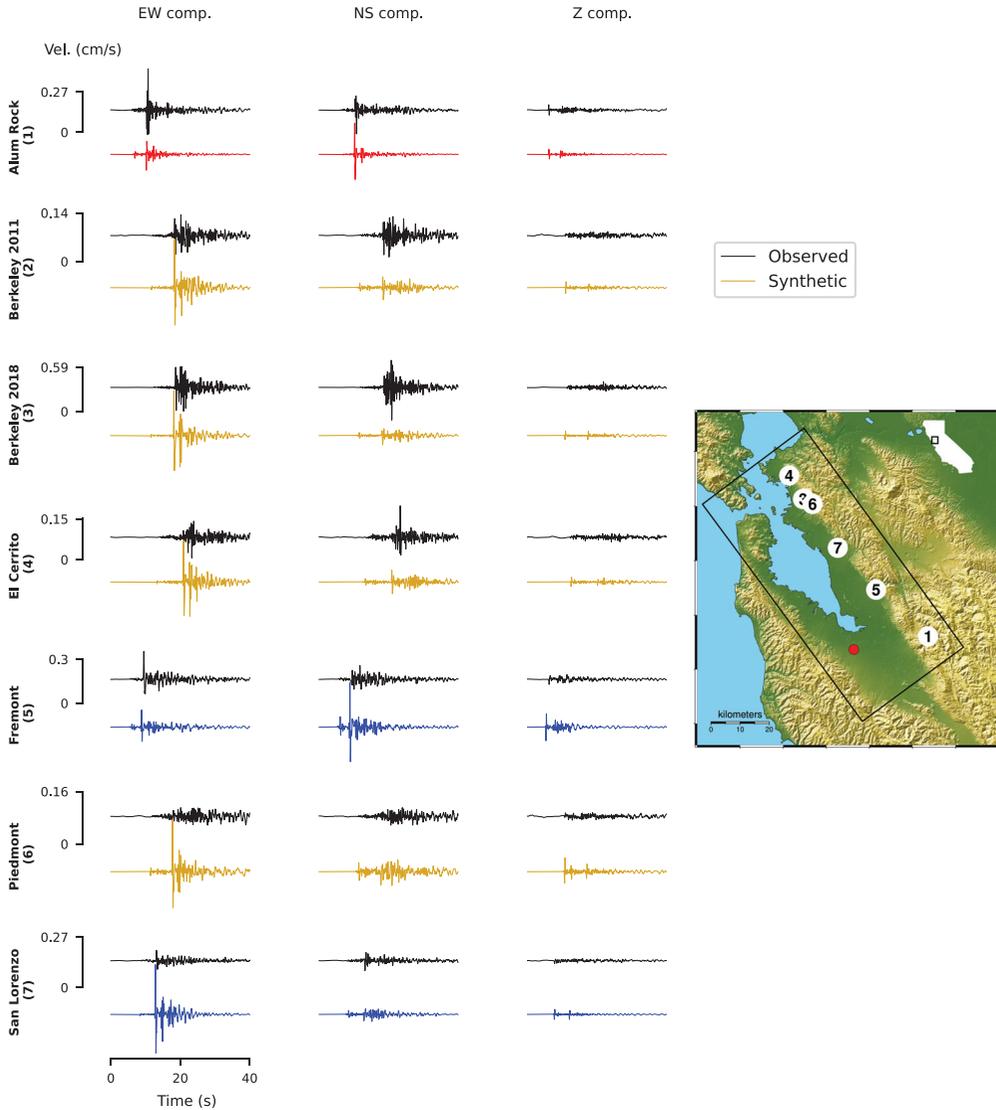
Figure 17. Comparison of recorded (black) and simulated (different colors) time histories of ground motion velocity band-passed filtered between 0.2 and 5 Hz at station 1796 (red circle in the map) obtained for the seven earthquakes. The earthquake name is indicated on the left of each trace. Red, yellow, and blue colors correspond to good, fair, and poor waveform matching quality, respectively. Numbered white circles in the map indicate the earthquake's epicenter location. The amplitude for the plots was normalized within each earthquake, showing the maximum value between the six waveforms of each earthquake in the vertical axis.

model. The spectral acceleration residuals between recorded and simulated ground motions are spatially correlated, which allows us to model their spatial patterns. We estimate the median spatial distribution of these residuals to evaluate suboptimal prediction patterns in our computational domain (i.e. areas of systematic under- or over-prediction).
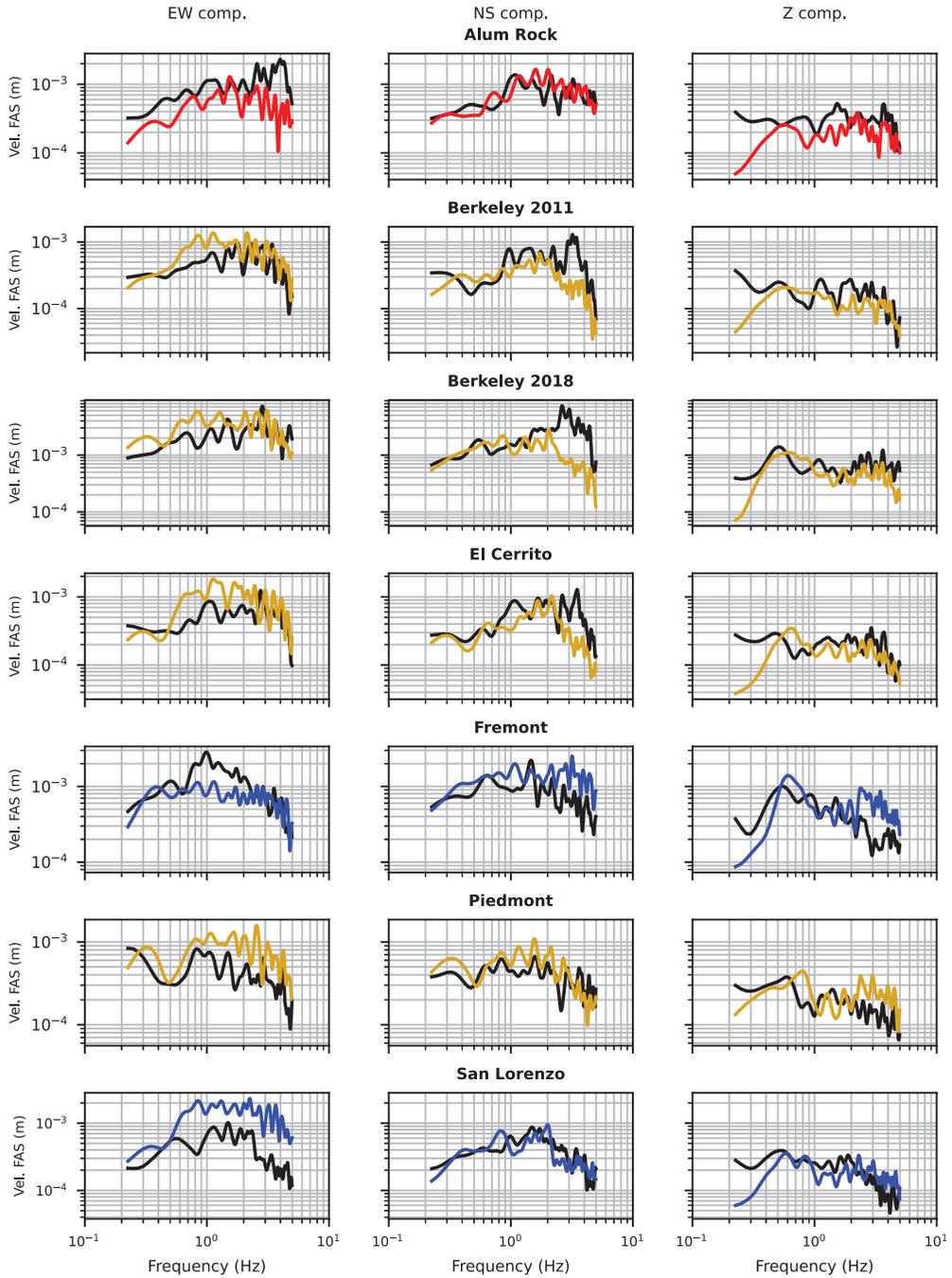
Figure 18. Comparisons of recorded (black) and simulated (different colors) Fourier amplitude spectra of ground motion at station 1796 obtained for seven earthquakes. The earthquake name is indicated on top of each panel. Red, yellow, and blue colors correspond to good, fair, and poor waveform matching quality, respectively.

We decoupled the residuals from the observation at site "*s*" from the earthquake "*e*" into three main terms as:

$$r_{es} = \ln(Y_{es}) - \ln(\hat{Y}_{es}) = \delta B_e + \delta S(\vec{X}) + \delta E_{es}, \qquad (1)$$

in which Y_{es} and $\hat{Y}_{es}$ are the recorded and simulated ground motion spectral amplitude, respectively. δB_e is an event term that centers the data with respect to some of the systematic unmodeled or mismodeled source effects that manifest in all the recorded waveforms inside an event, quantifying the variance induced by our source characterization with respect to these patterns. $\delta S(\vec{X})$ is a spatially varying random variable that models systematic spatial differences between recorded and simulated data at location $\vec{X}$, based on the residuals obtained at each station, providing a first-order modeling misfit in our simulations due to mostly wave propagation. Under this formulation, the variability that cannot be explained by $\delta S(\vec{X})$ moves to δE_{es}, including systematic unconstrained path effects which $\delta S(\vec{X})$ cannot capture because of its azimuth-independent formulation. In addition, because of our source formulation, it cannot reproduce finite source effects, such as directivity, for example, which are tracked by δE_{es}. In the "Discussion" section, we refer deeper to the trade-off between source characterization and wave propagation.

Equation 1 was solved using the Gaussian process module of pymc, a Bayesian modeling package written in Python (Salvatier et al., 2016). $\delta S(\vec{X})$ is modeled following a Gaussian process regression (Rasmussen and Williams, 2005), allowing for building a spatially correlated Gaussian field. The covariance matrix of $\delta S(\vec{X})$ has a site-dependent exponential-quadratic kernel. This kind of regression also provides a correlation length and a correlation structure of the Gaussian field, allowing us to extrapolate the available data to other areas, quantifying the epistemic uncertainty in the extrapolation. For instance, predictions made on regions without data nearby (e.g. beyond two correlation lengths from a point with data) have a $\delta S(\vec{X})$ equal to 0. Figure 19 displays maps of the resulting Gaussian fields for three different spectral frequencies: 0.5, 1.0, and 4.0 Hz (2.0, 1.0, and 0.25 s). The upper row presents the $\delta S(\vec{X})$ median. The lower row illustrates the epistemic uncertainty in the model inference, including the extrapolation outside areas with data. This analysis shows systematic over-prediction patterns in a significant portion of the western block of the Hayward fault, specifically in the sediments of the San Francisco Bay. This systematic ground motion over-prediction suggests that the shear-wave velocity assigned to most of these sites is considerably lower than the actual geologic structure, or there are inconsistencies between the velocity model and the local shallow geology. In addition, the Gaussian fields show systematic under-prediction in a major fraction of the eastern block of the Hayward fault.

Treasure and Yerba Buena islands are not characterized in the velocity model, keeping the same velocity structure as the nearby bay sediments (see Figure 4). The impact of this characterization is shown in Figure 19 with systematic patterns of over-prediction. The velocity model should be updated in this area by including the buried geology structures surrounding these islands, which give rise to the Yerba Buena outcrop.

The USGS velocity model performance will improve by explaining the sources of the discrepancies captured by $\delta S(\vec{X})$. The benefits of this potential improvement are quantified in Figure 20, showing the reduction in the within-event standard deviation induced by the variance explained by $\delta S(\vec{X})$. The dashed lines represent the within-event standard deviation, ϕ, for each individual event, and the black line is their median. ϕ can be computed as the standard deviation of the within-event residual maps shown in Figures 14 to 16. The
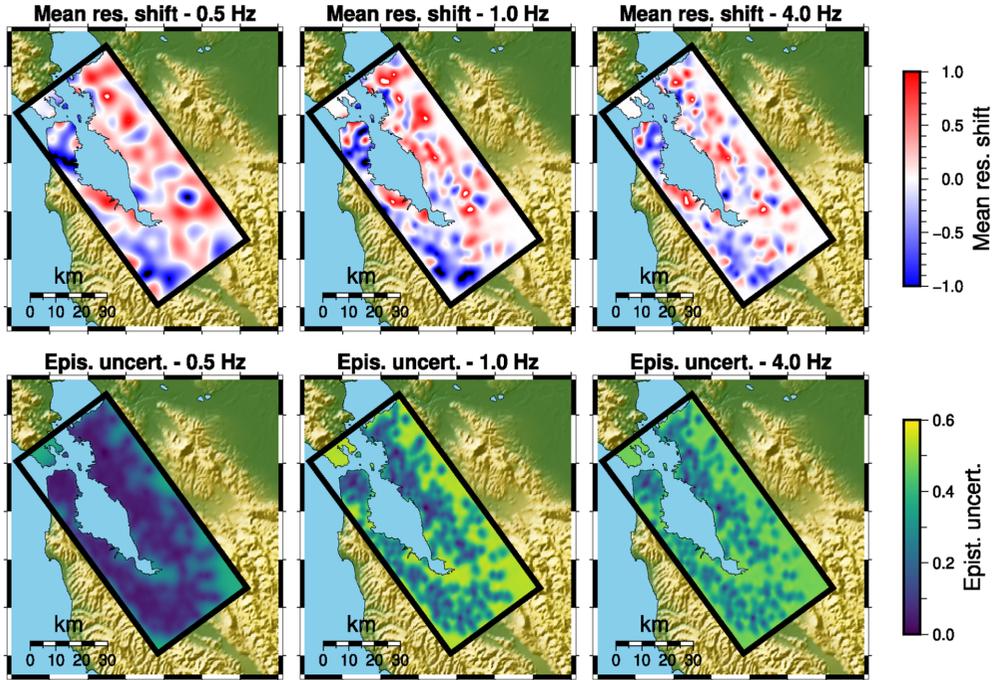
Figure 19. Inference of the spatially varying term, $\delta S(\vec{X})$. The upper row is the median of the spectral response amplitude difference in log-scale between the recorded and simulated data. Blue means over-prediction and red under-prediction. The lower row presents the epistemic uncertainty in the extrapolation to regions without data.

red line is the standard deviation of δE_{es}, σ_0, representing the remaining unexplained standard deviation after applying $\delta S(\vec{X})$ to the residuals. Because of the azimuthal-independent formulation of $\delta S(\vec{X})$, this variable explains local systematic differences between recorded and simulated data at location $\vec{X}$, not including any information about the source location or azimuthal source site dependence. Nevertheless, the standard deviation can decrease to values near 0.4 LN units after explaining the variability in the residuals by including $\delta S(\vec{X})$ and δB_e. This is an interesting result: although the Gaussian field developed in this section only considers the site location in the covariance kernel, solving as a first order the differences identified by $\delta S(\vec{X})$ would significantly reduce the velocity model's epistemic uncertainty and improve the simulations' accuracy.

The hyper-parameters from the Gaussian process regression are shown in Table 2. The correlation length ranges from 4.36 km at 0.25 Hz to 0.76 km at 4 Hz, showing that the higher the frequency, the smaller the correlation length. The standard deviation of δB_e, τ, increases with increasing the frequency. This implies that the actual source mechanisms and STFs become more complex at high frequencies, existing features manifested in the actual source and observed in all the recordings inside an event that our source characterization based on a point-source double-couple mechanism with a Liu STF is not fully capturing.

The spatially varying systematic misfit, $\delta S(\vec{X})$, shows an inter-period correlation, being stronger at intermediate to low frequencies. If a site exhibits under-prediction at 0.5 Hz, it

Table 2. Summary of the parameters obtained from the Gaussian process regression

Spectral frequency (Hz)	ρ (km)	τ (log-natural units)	σ_0 (log-natural units)
0.25	4.36	0.095	0.431
0.5	3.70	0.118	0.408
0.75	1.94	0.174	0.383
1	1.69	0.207	0.408
2	1.27	0.27	0.504
4	0.76	0.329	0.625

ρ is the correlation length. τ and σ_0 are the standard deviations of δB_e and δE_{es}, respectively.
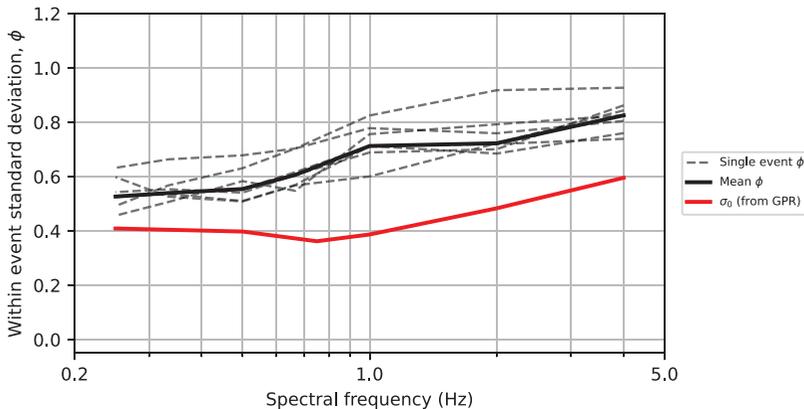

Figure 20. Within-event standard deviation for the spectral acceleration, ϕ, in log-natural units before and after applying $\delta S(\vec{X})$. The dashed black lines correspond to ϕ of each individual event, and the black line is the median ϕ. The red line represents the reduction on ϕ because of the variance explained by $\delta S(\vec{X})$. GPR = Gaussian Process Regression.

will likely have the same behavior at 1 Hz. At 4 Hz, this correlation still exists but is weaker, driven by the smaller correlation length. This feature should be useful for constraining the velocity model in future refinements.

Waveform duration analysis

Complex geologic structures cause wave scattering, which can create secondary wave phases in addition to primary source-generated body waves and surface waves. Most of these phases, also known as coda waves, arrive after the primary shear waves, thus increasing the duration of ground motions. The normalized Arias intensity is often used to quantify how the total seismic energy is spread over time, indirectly quantifying the presence of these additional phases (Pinilla-Ramos et al., 2023). We adopt D_{5-95} to estimate the waveform duration, which measures the time elapsed between the 5% and 95% of the normalized Arias intensity energy in the waveforms. Thus, D_{5-95} can be used to evaluate differences in waveform duration between recorded and simulated waveforms, attributable mainly to additional phases generated by complex wave propagation patterns not captured by the simulations. Figure 21 compares recorded and simulated waveforms from the Berkeley 2018 earthquake for stations J037 and C005, at the edge of the Bay basin and at
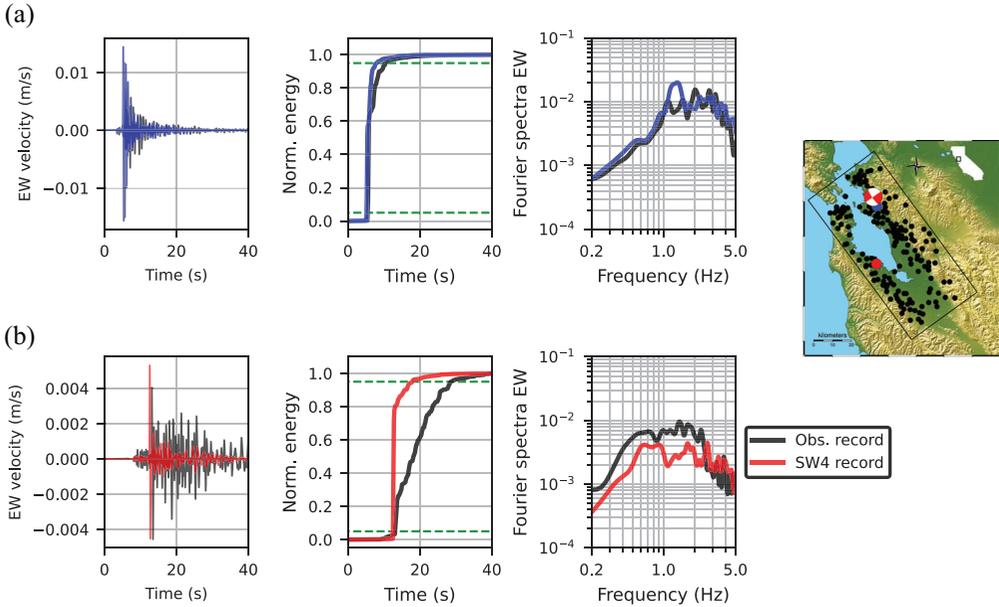
Figure 21. Example of waveforms' duration at two sites for East-West (EW) component. (a) Station C005 (blue color) shows a similar duration between simulated and recorded waveforms. (b) Station J037 (red color) develops strong surface wave amplitude in the recorded waveforms after the arrival of the shear-wave phases. The simulated waveforms do not develop this more complex impulsive response pattern.

the Berkeley hills, respectively. The simulated waveforms from station C005 represent a case where the simulated and recorded waveforms show a similar duration, with the simulated being slightly longer (2.8 vs 4.9 s). This case represents an example of a good match in which the Husid plots share a similar shape (central plots in Figure 21). In contrast, the simulated waveforms on station J037 contain relatively small amplitude phases arriving after the direct shear-wave phases, while the recorded waveforms contain late-arrival phases with amplitudes as high as 75% of that of the direct shear-wave phase, increasing the duration. Despite the simulated and recorded waveforms having similar amplitudes at the time of arrival of the first shear waves, the recorded waveforms have larger amplitudes in the low-frequency Fourier domain, induced by these late-manifested phases. For this station, D_{5-95} is 15.6 and 5.4 s for the recorded and the simulated waveforms, respectively.

We evaluated spatial patterns of D_{5-95} residuals to identify areas with significant differences in the duration between recorded and simulated data. The D_{5-95} residuals are defined as:

$$r_{es} = \ln\left(\frac{D_{5-95-\text{obs},\,es}}{D_{5-95-\text{synth},\,es}}\right), \tag{2}$$

in which $D_{5-95-\text{obs},\,es}$ and $D_{5-95-\text{synth},\,es}$ are the D_{5-95} for the recorded and simulated waveforms, respectively, at station "*s*" and event "*e*." We defined the residuals as the ratio in the log-natural scale to be consistent with Pinilla-Ramos et al. (2023), which models the significant duration using a power normal distribution, adopting a transformation closer to a log-normal than to a normal distribution. We modeled r_{es} using a Gaussian process
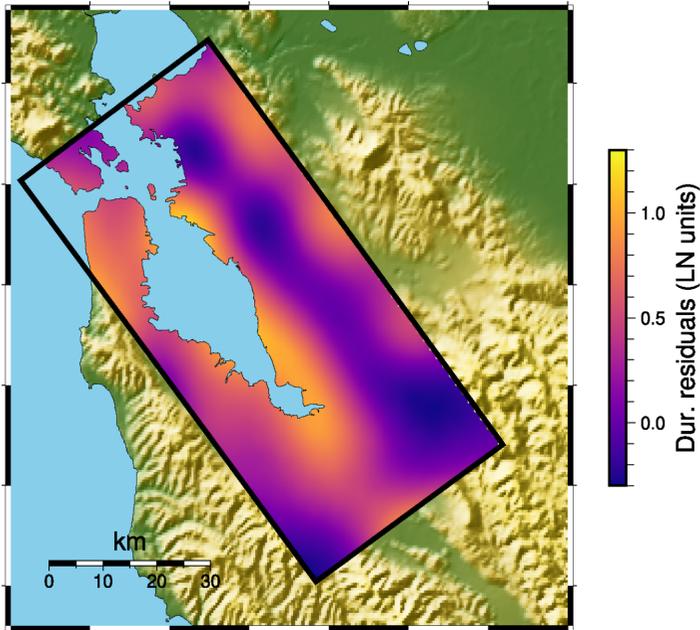
Figure 22. Spatial distribution of the D_{5-95} residuals in log-scale (log-natural units) obtained from the Gaussian process regression. Colors close to purple show areas where D_{5-95} is similar between the simulated and recorded waveforms. Colors close to yellow represent areas where the recorded waveforms exhibit a longer duration.

regression. In the "Modeling of spatial patterns of suboptimal predictions" section, we explained the details of our Gaussian process regression model. The resulting residuals' Gaussian field is displayed in Figure 22, where it is clear that almost all sites have a log ratio above zero, except for a few that have rather small negative values. A positive value entails that the duration of the recorded data is longer than the duration of the simulated records, while a negative value indicates the opposite. The spatial patterns of the residuals, together with the data from Figure 4, show that stiff sites tend to have small ratios, primarily located at the Hayward fault's eastern side and the western Santa Clara hills. Regions located on the bay sediments reach D_{5-95} ratios ranging between 2.5 and 3.3 (0.9–1.2 in LN units), showing that in these areas, the simulated waveforms are missing energy after the arrival of the main direct shear-wave phases.

Discussion

This study focuses on the performance evaluation of the USGS velocity model using physics-based ground motion modeling of earthquakes in the SFBA. In our analysis, we tried to identify the cause of the misfit between the recorded and simulated ground motions, which can mainly be linked to the velocity model, earthquake source, and material stress–strain characterizations adopted in our simulations. Conceptually, we can decouple the variance of the systemic misfit between simulated and recorded waveforms as:

$$\sigma^2 = \sigma^2_{VM} + \sigma^2_{SC} + \sigma^2_{NL}, \tag{3}$$

in which σ_{VM} corresponds to the epistemic uncertainty or misfit induced by the velocity model. σ_{SC} represents the source characterization variability or its modeling misfit. σ_{NL} is the overall misfit between the actual stress–strain material behaviors compared to how we are modeling them. Because of the low amplitude induced by small-magnitude earthquakes, the constitutive models adopted can be represented by a linear elastic stress–strain relationship, leading to assume σ_{NL} equal to zero. This allows the removal of one potential source of epistemic uncertainty, which becomes more important at ground motions with larger amplitudes, where materials can reach non-linear deformation regimes, especially on materials near the surface. Then, we parameterized the seismic source independently of the velocity model to reduce the trade-off between misrepresentations of the source characterization and the velocity model, bounding σ_{SC}. Thus, this methodology aids us in the direct identification and quantification of the epistemic uncertainty induced by the wave propagation of simulated waveforms in this velocity model, σ_{VM}. As a result of this analysis, we found that the velocity model has an overall good performance, with a slight systematic over-prediction at high frequencies and a shorter duration of the simulated waveforms compared to the recorded waveforms in the softer sediments of the Bay basin. The combination of the USGS VM for the SFBA and our source parametrization shows that our input parameters are well chosen and have physical sense. In addition, as the simulated earthquakes' hypocenters cover a wide range of the Hayward fault surface, our analysis shows that this velocity model is suitable for simulating wave propagation under elastic deformation regimes of large-magnitude earthquakes in the Hayward fault for a frequency range from 0.2 to 5 Hz. Nevertheless, because of the innate trade-off between wave propagation and source characterization in seismological applications, we broaden the discussion by including elements that can be investigated more deeply in further research to improve the modeling of small-magnitude earthquakes.

Use of spatial patterns of suboptimal predictions to improve the velocity model

The spectral amplitude residuals' maps identify features that can be enhanced in future versions of the velocity model, quantifying the benefits of these improvements. For instance, the soft sediments of the Bay basin tend to produce systematic over-prediction of the recorded ground motion, suggesting that the model needs improvements in areas with very soft surface sediments. Moreover, the western block exhibits over-prediction of spectral acceleration, especially in Santa Clara, San Jose, Fremont, and several areas of the SFBA. This could also be due to differences in the actual shallow velocity structure, but this time with the velocity model having lower values. In addition, the seismic properties in the regions surrounding Treasure Island and Yerba Buena Island, where the current USGS velocity model does not include the rock outcrop of these islands, need to be added.

We modeled the residuals using Gaussian process regressions, a stochastic machine learning technique (Rasmussen and Williams, 2005) that allows us to track these spatial patterns. By explaining the sources of the differences our model identified, being tracked by $\delta S(\vec{X})$, the accuracy of the simulations can be significantly increased. Furthermore, the spatial distribution of the duration's residuals aids in identifying where the geometry of shallow and deep geologic structures is well constrained and where it needs refinements. The geologic units' geometry of the eastern block seems to be well constrained. In spite of the velocity model including complex buried structures in the Bay basin, these basins still do not capture the whole range of complexities of the actual geology. Therefore, these maps of the spatial distribution of residuals provide valuable information for new refinements and a methodology to evaluate their impacts.

Future refinements of the velocity model can take advantage of the fact that in California, V_{S30} can explain a significant portion of the local site amplification (Pinilla-Ramos et al., 2022) since it is highly correlated with the velocity profile at depth (Boore et al., 2011; Kamai et al., 2016). The observed discrepancies on V_{S30} provide space to evaluate local and regional correlation structures between surficial values of shear-wave velocity (i.e. near the surface), represented by V_{S30}, with the shear-wave velocity profile at depth, as the methodology developed by Tehrani et al. (2022). Including these correlations can provide invaluable information for future velocity model refinements. This is supported by the results of our Gaussian process regressions, which, being its formulation only site-dependent, can explain an important portion of the mismatch between recorded and simulated data. This suggests that modifications of the velocity model based on one-dimensional (1D) velocity profiles can improve the velocity model performance at this stage.

Corner frequency

The corner frequency, f_c, is a fundamental parameter of the STF used in our simulations and controls the frequency content of the generated seismic energy. Following the Hanks and Thatcher (1972) model, the corner frequency is proportional to the cubic stress drop and inversely proportional to the cubic moment magnitude. Due to the non-uniqueness of the proposed methods for estimating the f_c and its trade-off with wave propagation effects, including high-frequency attenuation and local site effects, the reported stress drop estimates for recorded earthquakes are often not precise or unique (Abercrombie, 2021; Chang et al., 2023). Trugman and Shearer (2018) estimated the uncertainty on f_c for small events in the SFBA, showing that the standard deviation of f_c suggests a non-negligible uncertainty. To demonstrate the effects of the f_c uncertainty on simulated ground motions, we performed a sensitivity analysis of the mean bias in the Fourier amplitude domain. We varied f_c in the 1–5 Hz range in simulations of the Berkeley 2018 and Piedmont earthquakes. This sampling of f_c is inside the 5th–95th confidence interval estimated by Trugman and Shearer (2018). The results of the sensitivity analyses are displayed in Figure 23, which shows that the bias at high frequencies strongly depends on the chosen f_c, impacting the overall performance of the simulations. Nevertheless, by using the Trugman and Shearer (2018) model for estimating the corner frequency (Table 1), we obtained a low bias in the residuals suggesting that the adopted values for f_c produce ground motions that, on average, fit the recorded data. Additional simulations of recorded local earthquakes are necessary to better understand the trade-off between the source and wave propagation effects, especially at high frequencies.

Earthquake location

As was shown in the "Velocity model" section, there is a strong impedance contrast between the western and the eastern block of the Hayward fault. Because of this lateral contrast, the wave field has multiple paths compared to a 1D velocity model (Baise et al., 2003; Stidham et al., 1999). Head waves traveling through an interface with strong impedance contrasts will arrive earlier in the softer block (Allam et al., 2012) compared to a 1D velocity model, which, together with the manifestation of multiple paths, add complexities to the estimation of the event location. Nevertheless, the waveform comparison between recorded and simulated data does not show a significant bias in the arrival time, meaning that, on average, the sources' locations and the velocity model adopted provide
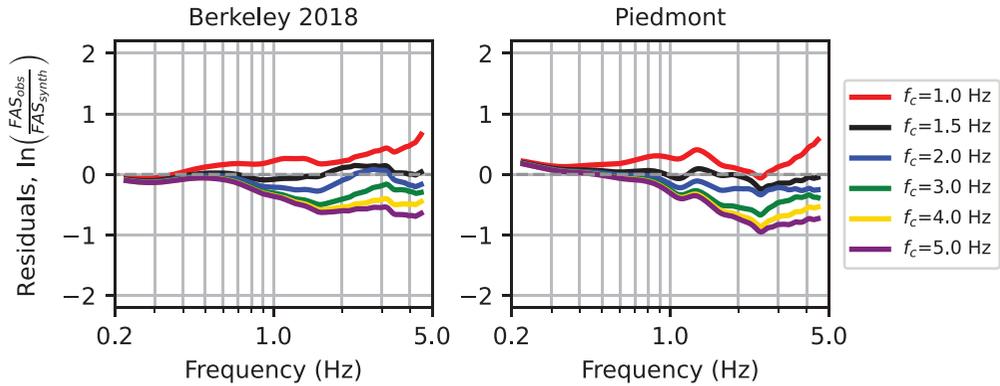
Figure 23. Sensitivity analysis performed on the corner frequency, f_c, for the Berkeley 2018 and Piedmont earthquakes. The dashed horizontal gray line represents the zero bias. f_c strongly impacts the bias obtained in the mean Fourier amplitude residuals.

good results. This indicates that the added complexities to the estimation of the event location have little effect in this case. This is consistent with Hirakawa and Aagaard (2022), who updated this version of the velocity model based on inversions of the arrival time of waveforms from small-magnitude earthquakes through simulations using SW4.

Stochastic variability in the velocity model

We showed that the velocity model has an overall good performance, with a slight systematic over-prediction at high frequencies (above 3 Hz) and a shorter duration of the simulated waveforms compared to the recorded waveforms. One possibility for improving the performance at high frequencies is the inclusion of stochastic variability in the velocity model. This leads to more realistic waveforms (Savran and Olsen, 2019; Pitarka and Mellors, 2021), being able to reconcile this bias observed above 3 Hz and improving the waveform matching (Graves and Pitarka, 2016). Adding stochastic variability in the geologic structure will aid in reconciling some differences observed in the waveform duration by adding more phases induced by internal reflection effects due to a more diffuse wave propagation regime. Natural materials have inherent variabilities in their properties spread out in space, and data are needed to constrain the stochastic perturbation generator models. Consequently, the inclusion of stochastic small-scale variability in velocity models is expected to improve the quality of the simulated ground motion on a broad frequency range (e.g. Graves and Pitarka, 2016). Despite the fact that it is appealing to control the over-prediction at high frequencies by increasing the velocity model randomization, we should consider that the randomization of the velocity model cannot be blindly increased outside the boundaries of real variabilities in natural materials. The stochastic variability decreases the spatial coherence of the waveforms, hindering the phase matching in the time domain. As shown by Abrahamson et al. (2022), an increment of the velocity model randomization increases the standard deviation of the phase difference in the frequency domain, decreasing the waveform coherence. They state that randomization excess in the velocity model can reduce the waveform coherence, with a consequent increase in the within-event standard deviation.

Source radiation pattern

The variance decomposition in the "Modeling of spatial patterns of suboptimal predictions" section shows that the between-event standard deviation increases at higher frequencies. This is consistent with the theoretical framework of source characterization. Point-source mechanisms provide a proper source characterization when the wave field is observed in wavelengths much larger than the event size (Aki and Richards, 2002; Kostrov and Das, 1988). Even moderate to large earthquakes at teleseismic distances can be represented as point-source double-couple mechanisms without missing general information about the earthquake at the observed wavelengths. However, for wavelengths comparable to the rupture dimensions, details of the dynamic rupture process of the fault could become important. The smallest wavelength we solved is 50 m, while the ruptures' dimensions can range between 1.0 and 2.5 km (based on the scaling equations from Leonard (2010)); therefore, the finiteness of the rupture at this scale could matter. In addition, because of the plastic damage-breakage process developed during the earthquake rupture, the source radiation pattern could have a significant non-double-couple component (isotropic and Compensated Linear Vector Dipole (CLVD)), especially for crack-type ruptures, as small earthquakes are (Ben-Zion and Ampuero, 2009; Lyakhovsky et al., 2016). In these cases, the shear-wave radiation pattern is influenced by the damage process in the whole source volume, leading to a more complex mechanism than the classical deviatoric rose of a double-couple mechanism (Kurzon et al., 2022; Lyakhovsky et al., 2016). More in-depth investigations of observed source radiation patterns are required to provide robust constraints to modeling the frequency-dependent radiation pattern and facilitate the inclusion of non-double-couple components in source models used in ground motion simulations of small earthquakes. For instance, we should quantify the benefits of adopting a frequency-dependent radiation pattern, including non-double-couple components, and how much the accuracy increases.

Misfit contributions between source characterization and wave propagation

In the evaluation of the accuracy of simulation model results, a key point is the extent to which mismodeled effects in the source characterization contribute to the misfit between recorded and simulated waveforms and, thus, what error is introduced by adopting a point-source double-couple mechanism with a Liu STF, emulating a symmetric circular crack. Studies of source radiation patterns of small earthquakes (Takenaka et al., 2003; Trugman et al., 2021; Vidale, 1989) suggest that the radiation pattern is predominantly a double couple for close source–receiver distances and maximum frequency ranging from 3.5 to 16 Hz, depending on the author. Based on an extensive dataset of earthquakes in Southern California, Trugman et al. (2021) showed that the focal mechanism maintains about 90% of its shear component up to 16 Hz for distances smaller than 35 km. Based on these findings, it is reasonable to assume that in the frequency range 0–3 Hz the differences between recorded and simulated waveforms tend to be dominated by wave propagation effects, with a lesser contribution from the mismodeled radiation pattern effects. Nevertheless, it is acknowledged that wave propagation regimes in a diffuse media lead to a loss of the source radiation pattern toward the manifestation of non-double-couple components, the effects of which are embedded in wave field observational data.

Going forward, it will be desirable to move toward the parameterization of stochastic perturbations of the velocity structure to obtain an enhanced representation of the high-frequency range of the wave field in 3D physics-based regional-scale simulations. This is a

key step to include in future developments for this kind of simulation, especially when pushing toward the even higher frequencies enabled by computational advancements. Similarly, source directivity could be important as we move toward higher frequency simulations. Even ruptures of small earthquakes are not symmetrical and exhibit some tendency to nucleate at the edge of rupture areas (Brietzke and Ben-Zion, 2006). This asymmetry induces a zenithal and azimuthal dependency on how an observer perceives the wave field's low-frequency amplitude and corner frequency (Kaneko and Shearer, 2014, 2015), indicating that the corner frequency is not only a source parameter but also an observer parameter. For practical purposes, this manifests as local variations of the corner frequency from the observer's standpoint, which can have an impact on the local illumination provided by the simulated source with respect to the actual observations. The degree to which source characterization in a simulation model influences the misfit between recorded and simulated data should be a focus of future research to facilitate a more comprehensive understanding of simulation data misfit and an appropriate parametric representation of earthquake sources.

Conclusions

Our ground motion simulations of seven small-magnitude earthquakes recorded in the SFBA suggest that, overall, the SFBA USGS velocity model v21.1 performs well in capturing the recorded wave path effects. The comparison of the residuals in the Fourier amplitude domain obtained for each earthquake indicates that the bias between the recorded and simulated ground motion averaged over more than 200 stations is close to zero in the frequency range 0.25–3 Hz. A slight over-prediction of 25% was observed between 3 and 5 Hz. The standard deviation of the residuals between recorded and simulated data ranges from 0.48 LN units at 0.25 Hz to 0.8 LN units at 5 Hz. The relatively low bias suggests that, in general, the performance of our source parameterization, in addition to the velocity model used in our simulations, is acceptable. We concluded that most of the misfit between the recorded and simulated waveforms was caused by the inadequate wave propagation effects due to velocity model inaccuracies in several areas.

The spectral acceleration residuals were found to be spatially correlated between and within each earthquake. We modeled the residuals using Gaussian process regressions, a stochastic machine learning technique (Rasmussen and Williams, 2005) that allows us to track these spatial patterns. The outputs of these regressions can aid in the identification of sub-domains that may need future refinements in new versions of the velocity model. For instance, the spectral acceleration residuals in the eastern block show systematic under-predictions, especially in the hills east of the Hayward fault. On the contrary, the western block exhibits over-prediction of spectral acceleration, especially in Santa Clara, San Jose, Fremont, and several other areas of the SFBA.

In addition to the comparison of the spectral amplitude, we also looked at comparisons of the waveform duration between simulated and recorded waveforms, represented by the normalized Arias intensity parametrization D_{5-95}. Due to its high sensitivity to wave propagation effects, the waveform duration is a very important ground motion intensity parameter that can be used to characterize the quality of the velocity model. Being directly linked to wave scattering caused by underground geologic complexities, it can also be used to identify regions where the velocity model needs modifications. We found that, on average, the recorded waveforms exhibit longer duration than the simulated waveforms at soft soil basin sites. In contrast, the simulated waveforms on stiff sites have a similar waveform

duration compared to that of the recorded waveforms, implying that the underground geologic complexities in these latter areas are well characterized in the model.

Our methodology for assessing the velocity model performance is helpful in trying to decouple the three primary sources of uncertainty: the rupture characterization, velocity model, and constitutive models. Since we simulated small-magnitude earthquakes, and assuming that the materials would remain within a linear deformation regime, we could adopt a linear elastic stress–strain relationship. This avoids the increase of epistemic uncertainty due to testing different non-linear constitutive models. We obtained a quantitative evaluation of the velocity model performance by directly comparing recorded and simulated waveforms. This methodology provides a basis for propagating epistemic uncertainties in seismic hazard analysis applications based on 3D regional physics-based simulations.

Data and resources

The simulation of recorded ground motions from these earthquakes at the SFBA stations was performed at the National Energy Research Scientific Computing Center under the framework of the EQSIM project (McCallen et al., 2021a, 2021b) and at the Oak Ridge Leadership Computing Facility under the DOE Innovative and Novel Computational Impact on Theory and Experiment program. The simulated and recorded waveforms are available in Xarray format in the electronic supplement. The USGS velocity model for the SFBA is available on the USGS website.

ORCID iDs

Camilo Pinilla-Ramos (iD) https://orcid.org/0000-0002-4034-6618
Arben Pitarka (iD) https://orcid.org/0000-0001-6026-8859

References

Aagaard B and Hirakawa E (2021) San Francisco Bay region 3D seismic velocity model v21.*1*. DOI: 10.5066/P9TRDCHE.

Aagaard B, Blair J, Boatwright J, Garcia S, Harris R, Michael A, Schwartz D and DiLeo J (2016) Earthquake outlook for the San Francisco Bay region 2014–2043. DOI: 10.3133/fs20163020.

Aagaard B, Brocher T, Dolenc D, Dreger D, Graves R, Harmsen S, Hartzell S, Larsen S and Zoback M (2008a) Ground-motion modeling of the 1906 San Francisco earthquake, part I: Validation using the 1989 Loma Prieta earthquake. *Bulletin of the Seismological Society of America* 98(2): 989–1011.

Aagaard B, Brocher T, Dolenc D, Dreger D, Graves R, Harmsen S, Hartzell S, Larsen S, McCandless K, Nilsson S, Petersson A, Rodgers A, Sjögreen B and Zoback M (2008b) Ground-motion modeling of the 1906 San Francisco earthquake, part II: Ground-motion estimates for the 1906 earthquake and scenario events. *Bulletin of the Seismological Society of America* 98(2): 1012–1046.

Abercrombie R (2021) Resolution and uncertainties in estimates of earthquake stress drop and energy release. *Philosophical Transactions of the Royal Society A* 379(2196): 20200131.

Abrahamson N, Pinilla-Ramos C, Tehran P, Feenstra P and Krimotat A (2022) Modeling of vertical component ground motion for soil-structure-interaction analyses. In: *26th International Conference on Structural Mechanics in Reactor Technology*, Potsdam, 10–15 July.

Aki K and Richards P (2002) *Quantitative Seismology*. Dulles, VA: University Science Books.

Allam A, Ben-Zion Y and Peng Z (2012) Seismic imaging of a bimaterial interface along the Hayward Fault, CA, with fault zone head waves and direct P arrivals. *Pure and Applied Geophysics* 171: 2993–3011.

Ancheta T, Darragh R, Stewart J, Seyhan E, Silva W, Chiou B, Wooddell K, Graves R, Kottke A, Boore DM, Kishida T and Donahue J (2014) NGA-West2 database. *Earthquake Spectra* 30(3): 989–1005.

Anderson J (2004) Quantitative measure of the goodness-of-fit of synthetic seismograms. In: *Oral Presentation at 13th World Conference on Earthquake Engineering*, Vancouver, BC, Canada, 1–6 August.

Bailey E, Irwin W and Jones D (1964) *Franciscan and Related Rocks, and Their Significance in the Geology of Western California (California Division of Mines and Geology Bulletin 183)*. Denver, CO: United States Geological Survey.

Baise L, Dreger DS and Glaser SD (2003) The effect of shallow San Francisco Bay sediments on waveforms recorded during the MW 4.6 Bolinas, California, earthquake. *Bulletin of the Seismological Society of America* 93(1): 465–479.

Bartow J and Nilsen T (1990) Review of the Great Valley sequence, eastern Diablo Range and northern San Joaquin Valley, Central California. DOI: 10.32375/1990-GB65.22.

Ben-Zion Y and Ampuero J (2009) Seismic radiation from regions sustaining material damage. *Geophysical Journal International* 178(3): 1351–1356.

Beyreuther M, Barsch R, Krischer L, Megies T, Behr Y and Wassermann J (2010) ObsPy: A python toolbox for seismology. *Seismological Research Letters* 81(3): 530–533.

Boore D, Thompson E and Cadet H (2011) Regional correlations of vs30 and velocities averaged over depths less than and greater than 30 meters. *Bulletin of the Seismological Society of America* 101(6): 3046–3059.

Brietzke G and Ben-Zion Y (2006) Examining tendencies of in-plane rupture to migrate to material interfaces. *Geophysical Journal International* 167(2): 807–819.

Brocher T (2005) Compressional and shear wave velocity versus depth in the San Francisco Bay Area, California: Rules for USGS Bay Area velocity model 05.0.0.

Brocher T (2008) Compressional and shear-wave velocity versus depth relations for common rock types in Northern California. *Bulletin of the Seismological Society of America* 98: 950–968.

Chang H, Abercrombie R, Nakata N, Pennington C, Kemna K, Cochran E and Harrington R (2023) Quantifying site effects and their influence on earthquake source parameter estimations

using a dense array in Oklahoma. *Journal of Geophysical Research: Solid Earth* 128: e2023JB027144.

Dolenc D and Dreger D (2005) Microseisms observations in the Santa Clara Valley, California. *Bulletin of the Seismological Society of America* 95(3): 1137–1149.

Field E, Biasi G, Bird P, Dawson T, Felzer K, Jackson D, Johnson K, Jordan T, Madden C, Michael A, Milner K, Page M, Parsons T, Powers P, Shaw B, Thatcher W, Weldon RW II and Zeng Y, California Earthquake Probabilities (2015) *UCERF3: A New Earthquake Forecast for California's Complex Fault System.* Reston, VA: US Geological Survey.

Frankel A, Carver D, Cranswick E, Bice T, Sell R and Hanson S (2001) Observations of basin ground motions from a dense seismic array in San Jose, California. *Bulletin of the Seismological Society of America* 91(1): 1–12.

Ghofrani H, Atkinson G and Goda K (2013) Implications of the 2011 M 9.0 Tohoku Japan earthquake for the treatment of site effects in large earthquakes. *Bulletin Earthquake Engineering* 11: 171–203.

Goulet C, Abrahamson N, Somerville P and Wooddell K (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.

Graves R and Pitarka A (2016) Kinematic ground-motion simulations on rough faults including effects of 3D stochastic velocity perturbations. *Bulletin of the Seismological Society of America* 106(5): 2136–2153.

Hanks T and Thatcher W (1972) A graphical representation of seismic source parameters. *Journal of Geophysical Research* 77: 4393–4405.

Hartzell S, Leeds A, Ramirez-Guzman L, Allen J and Schmitt R (2016) Seismic site characterization of an urban sedimentary basin, Livermore Valley, California: Site response, basin-edge-induced surface waves, and 3D simulations. *Bulletin of the Seismological Society of America* 106(2): 609–631.

Hernández-Aguirre VM, Paolucci R, Sánchez-Sesma FJ and Mazzieri I (2023) Three-dimensional numerical modeling of ground motion in the Valley of Mexico: A case study from the Mw3.2 earthquake of July 17, 2019. *EERI Earthquake Spectra* 39: 2323–2351.

Hirakawa E and Aagaard B (2022) Evaluation and updates for the USGS San Francisco Bay Region 3D seismic velocity model in the East and North Bay portions. *Bulletin of the Seismological Society of America* 112(4): 2070–2096.

Hu Z, Olsen K and Day S (2022) 0–5 Hz deterministic 3-D ground motion simulations for the 2014 La Habra, California, Earthquake. *Geophysical Journal International* 230(3): 2162–2182.

Kamai R, Abrahamson N and Silva W (2016) VS30 in the NGA GMPEs: Regional differences and suggested practice. *Earthquake Spectra* 32(4): 2083–2108.

Kaneko Y and Shearer P (2014) Seismic source spectra and estimated stress drop derived from cohesive-zone models of circular subshear rupture. *Geophysical Journal International* 197(2): 1002–1015.

Kaneko Y and Shearer P (2015) Variability of seismic source spectra, estimated stress drop, and radiated energy, derived from cohesive-zone models of symmetrical and asymmetrical circular and elliptical ruptures. *Journal of Geophysical Research: Solid Earth* 120(2): 1053–1079.

Kostrov B and Das S (1988) *Principles of Earthquake Source Mechanics.* Cambridge: Cambridge Monographs on Mechanics and Applied Mathematics.

Kramer S (1996) *Geotechnical Earthquake Engineering.* Hoboken, NJ: Prentice-Hall.

Kurzon I, Lyakhovsky V, Sagy A and Ben-Zion Y (2022) Radiated seismic energy and source damage evolution from the analysis of simulated dynamic rupture and far-field seismograms. *Geophysical Journal International* 231(3): 1705–1726.

Lebrun B, Hatzfeld D and Bard P (2001) Site effect study in urban area: Experimental results in Grenoble (France). *Pure Applied Geophysics* 158: 2543–2557.

Leonard M (2010) Earthquake fault scaling: Self-consistent relating of rupture length, width, average displacement, and moment release. *Bulletin of the Seismological Society of America* 100(5): 1971–1988.

Lienkaemper J, Williams P and Guilderson T (2010) Evidence for a twelfth large earthquake on the southern Hayward fault in the past 1900 years. *Bulletin of the Seismological Society of America* 100(5A): 2024–2034.

Liu P, Archuleta R and Hartzell S (2006) Prediction of broadband ground-motion time histories: Hybrid low/high-frequency method with correlated random source parameters. *Bulletin of the Seismological Society of America* 96: 2118–2130.

Lu Y and Ben-Zion Y (2022) Validation of seismic velocity models in southern California with full-waveform simulations. *Geophysical Journal International* 229(2): 1232–1254.

Lyakhovsky V, Ben-Zion Y, Ilchev A and Mendecki A (2016) Dynamic rupture in a damage-breakage rheology model. *Geophysical Journal International* 206(2): 1126–1143.

McCallen D, Petersson A, Rodgers A, Pitarka A, Miah M, Petrone F, Sjögreen B, Abrahamson N and Tang H (2021a) EQSIM—A multidisciplinary framework for fault-to-structure earthquake simulations on exascale computers part I: Computational models and workflow. *Earthquake Spectra* 37(2): 707–735.

McCallen D, Petrone F, Miah M, Pitarka A, Rodgers A and Abrahamson N (2021b) EQSIM—A multidisciplinary framework for fault-to-structure earthquake simulations on exascale computers, part II: Regional simulations of building response. *Earthquake Spectra* 37(2): 736–761.

Parks M (2019) *Engineering Properties and Geologic Setting of Old Bay Clay Deposits, Downtown San Francisco, California.* PhD Dissertation, Civil and Environmental Engineering Department, University of California, Berkeley, Berkeley, CA.

Petersson N and Sjögreen B (2012) Stable and efficient modeling of anelastic attenuation in seismic wave propagation. *Communications in Computational Physics* 12(1): 193–225.

Petersson N and Sjögreen B (2015) Wave propagation in anisotropic elastic materials and curvilinear coordinates using a summation-by-parts finite difference method. *Communications in Computational Physics* 299: 820–841.

Petersson N and Sjögreen B (2023) *User's Guide to SW4, Version 3.0.* Livermore, CA: Lawrence Livermore National Laboratory.

Phelps G, Graymer R, Jachens R, Ponce D, Simpson R and Wentworth C (2008) *Three-Dimensional Geologic Map of the Hayward Fault Zone, San Francisco Bay Region, California.* Reston, VA: US Geological Survey.

Pinilla-Ramos C, Abrahamson N and Kayen R (2022) Estimation of site terms in ground-motion models for California using horizontal-to-vertical spectral ratios from microtremor. *Bulletin of the Seismological Society of America* 112(6): 3016–3036.

Pinilla-Ramos C, Abrahamson N, Kayen R, Phung V and Castellanos-Nash P (2023) Ground-motion model for significant duration constrained by seismological simulations. *Bulletin of the Seismological Society of America* 114(2): 1015–1032.

Pitarka A and Mellors R (2021) Using dense array waveform correlations to build a velocity model with stochastic variability. *Bulletin of the Seismological Society of America* 111: 2021–2041.

Pitarka A, Graves R and Somerville P (2004) Validation of a 3D velocity model of the Puget Sound region based on modeling ground motion from the 28 February 2001 Nisqually earthquake. *Bulletin of the Seismological Society of America* 94(5): 1670–1689.

Pitarka A, Graves R, Irikura K, Miyakoshi K, Wu C, Kawase H, Rodgers A and McCallen D (2021) Refinements to the Graves–Pitarka kinematic rupture generator, including a dynamically consistent slip-rate function, applied to the 2019 Mw 7.1 Ridgecrest earthquake. *Bulletin of the Seismological Society of America* 112(1): 287–306.

Rasmussen C and Williams C (2005) *Gaussian Processes for Machine Learning.* Cambridge, MA: The MIT Press.

Rodgers A, Petersson A, Pitarka A, McCallen D, Sjögreen B and Abrahamson N (2019) Broadband (0–5 Hz) fully deterministic 3D ground-motion simulations of a magnitude 7.0 Hayward fault earthquake: Comparison with empirical ground-motion models and 3D path and site effects from source normalized intensities. *Seismological Research Letters* 90(3): 1268–1284.

Salvatier J, Wiecki T and Fonnesbeck C (2016) Probabilistic programming in python using PyMC3. *PeerJ Computer Science* 2: e55.

Savran W and Olsen K (2019) Ground motion simulation and validation of the 2008 Chino Hills earthquake in scattering media. *Geophysical Journal International* 219(3): 1836–1850.

Shearer P, Prieto G and Hauksson E (2006) Comprehensive analysis of earthquake source spectra in southern California. *Journal of Geophysical Research: Solid Earth* 111(B6): 2006JGRB..111.6303S.

Sjögreen B and Petersson N (2012) A fourth order accurate finite difference scheme for the elastic wave equation in second order formulation. *Journal of Scientific Computing* 52(1): 17–48.

Smerzini C, Vanini M, Paolucci R and Traversa P (2022) Validation of regional physics-based ground motion scenarios: The case of the Mw 4.9 2019 Teil earthquake in France. *Research Square*. DOI: 10.21203/rs.3.rs-1552665/v1.

Stidham C, Antolik M, Dreger D, Larsen S and Romanowicz B (1999) Three-dimensional structure influences on the strong-motion wavefield of the 1989 Loma Prieta earthquake. *Bulletin of the Seismological Society of America* 89(5): 1184–1202.

Taborda R and Bielak J (2014) Ground-motion simulation and validation of the 2008 Chino Hills, California, earthquake using different velocity models. *Bulletin of the Seismological Society of America* 104(4): 1876–1898.

Taborda R, Azizzadeh-Roodpish S, Khoshnevis N and Cheng K (2016) Evaluation of the southern California seismic velocity models through simulation of recorded events. *Geophysical Journal International* 205(3): 1342–1364.

Takenaka H, Mamada Y and Futamure H (2003) Near-source effect on radiation pattern of high-frequency S waves: Strong SH–SV mixing observed from aftershocks of the 1997 Northwestern Kagoshima, Japan, earthquakes. *Physics of the Earth and Planetary Interiors* 137: 31–43.

Tehrani H, Lavrentiadis G, Seylabi E, McCallen D and Asimaki D (2022) Towards a three-dimensional geotechnical layer model for northern California. https://earthquake.usgs.gov/cfusion/external_grants/reports/G21AP105 (accessed 29 August, 2024).

Toppozada T, Branum D, Reichle M and Hallstrom C (2002) San Andreas fault zone, California: M 5.5 earthquake history. *Bulletin of the Seismological Society of America* 92(7): 2555–2601.

Trugman DT, Chu S and Tsai V (2021) Earthquake source complexity controls the frequency dependence of near-source radiation patterns. *Geophysical Research Letters* 48(17): e2021GL095022.

Trugman DT and Shearer PM (2018) Strong correlation between stress drop and peak ground acceleration for recent M1–4 earthquakes in the San Francisco Bay Area. *Bulletin of the Seismological Society of America* 108(2): 929–945.

Trugman DT, Dougherty SL, Cochran ES and Shearer PM (2017) Source spectral properties of small to moderate earthquakes in southern Kansas. *Journal of Geophysical Research: Solid Earth* 122(10): 8021–8034.

Vidale J (1989) Influence of focal mechanism on peak accelerations of strong motions of the Whittier Narrows, California, earthquake and an aftershock. *Journal of Geophysical Research: Solid Earth* 94(B7): 9607–9613.