Geophysical Journal International

Geophys. J. Int. (2022) **230,** 2162–2182
Advance Access publication 2022 May 10
GJI Seismology

https://doi.org/10.1093/gji/ggac174

0–5 Hz deterministic 3-D ground motion simulations for the 2014 La Habra, California, Earthquake

Zhifeng Hu [1,2] Kim B. Olsen [1] and Steven M. Day[1]

[1]*Department of Geological Sciences, San Diego State University, San Diego, CA* 92093*, USA. E-mail: zhh076@ucsd.edu*
[2]*Scripps Institution of Oceanography, University of California, San Diego, San Diego, CA* 92182*, USA*

Accepted 2022 April 29. Received 2022 April 8; in original form 2021 October 6

SUMMARY

We have simulated 0–5 Hz deterministic wave propagation for a suite of 17 models of the 2014 M_w 5.1 La Habra, CA, earthquake with the Southern California Earthquake Center Community Velocity Model Version S4.26-M01 using a finite-fault source. Strong motion data at 259 sites within a 148 km × 140 km area are used to validate our simulations. Our simulations quantify the effects of statistical distributions of small-scale crustal heterogeneities (SSHs), frequency-dependent attenuation $Q(f)$, surface topography and near-surface low-velocity material (via a 1-D approximation) on the resulting ground motion synthetics. The shear wave quality factor $Q_S(f)$ is parametrized as $Q_{S,0}$ and $Q_{S,0}f^{\gamma}$ for frequencies less than and higher than 1 Hz, respectively. We find the most favourable fit to data for models using ratios of $Q_{S,0}$ to shear wave velocity V_S of 0.075–1.0 and γ values less than 0.6, with the best-fitting amplitude drop-off for the higher frequencies obtained for γ values of 0.2–0.4. Models including topography and a realistic near-surface weathering layer tend to increase peak velocities at mountain peaks and ridges, with a corresponding decrease behind the peaks and ridges in the direction of wave propagation. We find a clear negative correlation between the effects on peak ground velocity amplification and duration lengthening, suggesting that topography redistributes seismic energy from the large-amplitude first arrivals to the adjacent coda waves. A weathering layer with realistic near-surface low velocities is found to enhance the amplification at mountain peaks and ridges, and may partly explain the underprediction of the effects of topography on ground motions found in models. Our models including topography tend to improve the fit to data, as compared to models with a flat free surface, while our distributions of SSHs with constraints from borehole data fail to significantly improve the fit. Accuracy of the velocity model, particularly the near-surface low velocities, as well as the source description, controls the resolution with which the anelastic attenuation can be determined. Our results demonstrate that it is feasible to use fully deterministic physics-based simulations to estimate ground motions for seismic hazard analysis up to 5 Hz. Here, the effects of, and trade-offs with, near-surface low-velocity material, topography, SSHs and $Q(f)$ become increasingly important as frequencies increase towards 5 Hz, and should be included in the calculations. Future improvement in community velocity models, wider access to computational resources, more efficient numerical codes and guidance from this study are bound to further constrain the ground motion models, leading to more accurate seismic hazard analysis.

Key words: Structure of the Earth; Numerical Modelling; Computational seismology; Earthquake ground motions; Seismic attenuation; Wave propagation.

1 INTRODUCTION

It is the ultimate goal for ground motion modellers to deliver their results to engineers and see their work used in applications beneficial for society, such as structural design. This is particularly useful in cases of infrequent observations, such as for large-magnitude events at short distances from the fault, where simulations may provide a viable alternative to data. Deterministic ground motion predictions, including features such as 3-D velocity structure and frequency-independent anelastic attenuation are now routinely produced for frequencies up to about 1 Hz with generally satisfactory fit to recorded data (e.g. Graves 1996; Olsen *et al.* 2009; Roten *et al.* 2012). These simulations have

proved to be useful in public earthquake emergency response and seismic hazard management (e.g. the ShakeOut scenario; Jones *et al.* 2008), and in complementing empirical ground motion prediction models in regions with sparse station coverage (e.g. Day *et al.* 2008).

While the results of these low-frequency simulations are promising, structural engineers frequently need ground motions with signal content up to 5 Hz and higher for design purposes. Hybrid techniques, combining deterministic low-frequency and stochastic high-frequency signals (e.g. Olsen & Takedatsu 2015; Graves & Pitarka 2016) can be used to generate synthetic seismograms with frequency content up to at least 10 Hz. However, simulating both the lower and higher frequency content using a deterministic approach (in the sense outlined in Section 2) has the potential to lower part of the epistemic uncertainty in the resulting ground motion estimates. In this study we investigate the feasibility of increasing the highest frequency for accurate deterministic ground motion predictions to 5 Hz, using simulations and data for the 2014 M_w 5.1 La Habra, CA, earthquake. The La Habra event was chosen due to an abundance of records available in the greater Los Angeles area, while ground motions can be considered linear due to its relatively small magnitude.

As frequencies increase above about 1 Hz, features with increasingly small length scales become important to realistically predict deterministic ground motions. For example, small-scale heterogeneities of both the source and surrounding medium (SSHs), on the order of tens to hundreds of metres, are expected to increasingly affect ground motion predictions at higher frequencies. Frequency-independent anelastic attenuation ($Q(f)$), often chosen as proportional to the local velocity structure (e.g. Graves & Pitarka 2010) is usually a good approximation for lower frequencies (up to ∼1 Hz; e.g. Liu *et al.* 1976; Fehler *et al.* 1992). However, models of frequency-independent anelastic attenuation appear to be inconsistent with seismic records at higher frequencies where regional studies indicate that larger Q values may be more appropriate in some regions (e.g. Withers *et al.* 2015). Finally, ground motion simulations often artificially truncate the lowest near-surface velocities due to computational limitations, which may be a reasonable approximation for lower frequencies (e.g. Olsen *et al.* 2003). However, stronger effects from the near-surface material emerge as frequencies increase and wavelengths decrease (e.g. Pitarka & Ichinose 2009; Imperatori & Mai 2013). Here, we quantify the effects of all of these features in our 3-D simulations of the La Habra event.

In southern California, two state-of-the-art 3-D velocity models, namely the Community Velocity Models (CVM) versions S and H, have been developed through the Southern California Earthquake Center (SCEC). These CVMs have been validated against ground motion data in a series of studies (e.g. Taborda *et al.* 2016; Savran & Olsen 2019; Lai *et al.* 2020). Ely *et al.* (2010) proposed a method to calibrate the near-surface material based on estimates of the time-averaged velocity in the upper 30 m (V_{S30}), and later improved by Hu *et al.* (2022) (this issue, companion paper), specifically for sites with poor constraints for shallow rock site velocities (see Section 2.6). Here, we use the SCEC CVM-S with the update by Hu *et al.* (2022).

The effects of irregular surface topography on ground motions play an increasingly large role as frequencies increase (e.g. Liu *et al.* 2020). In recent studies, theoretical and numerical methods have helped clarify the interaction between seismic waves and topography (mainly scattering and trapping of waves, e.g., Imperatori & Mai 2015; Takemura *et al.* 2015; Rodgers *et al.* 2018), as well as describing the characteristic effects on ground motions. Some of the most notable effects of topography found by these studies are listed in the following. (1) Amplification tends to occur at the top of relatively steep slopes for waves with comparable wavelength to the size of the topographic features; on the other hand, deamplification tends to occur at low-elevation areas (Trifunac & Hudson 1971; Boore 1972; Bouchon & Barker 1996; Spudich *et al.* 1996; Assimaki *et al.* 2005). Amplification can range up to a factor of 10 or more between the crest and base of a topographic feature (Davis & West 1973; Umeda *et al.* 1987; Geli *et al.* 1988; Gaffet *et al.* 2000). (2) The amplification at mountain tops is systematically larger for incident *S* compared with *P* waves, with diminishing effect when the slope decreases or the incidence angle increases (Bard 1982). (3) Body and surface waves are strongly scattered by irregular topography, thus reducing ground motion amplitudes while prolonging the shaking duration (Sánchez-Sesma & Campillo 1991; Lee *et al.* 2009). (4) Topography tends to disrupt the coherency of high-frequency ground motion and thereby distort the *S*-wave radiation pattern (Imperatori & Mai 2015). Notably, 3-D models of the topography are necessary to capture the amplification effects, as noted in two-dimensional simulation studies (Geli *et al.* 1988; Bouchon & Barker 1996). While geometrical proxies, such as smoothed curvature and relative elevation, have been explored to represent topographic effects (e.g. Maufroy *et al.* 2015; Rai *et al.* 2017), they require critical parameter constraints based on local velocity and target frequency, and are thus difficult to generalize for broad-band studies.

It should be noted that previous studies discussed above include only a subset of the model features deemed to be affecting high-frequency ground motions, or omitted validation of the results. In this study, we present the results of a suite of ground motion simulations for frequencies up to 5 Hz in the widely tested SCEC CVM-S4.26M01 including high-resolution topography, and compare to strong-motion data for the 2014 M_w 5.1 La Habra, CA, earthquake, in order to constrain the relative contributions from topography, near-surface velocities, SSHs, and $Q(f)$ on the ground motions.

The paper is organized as follows. We first describe the velocity model, simulation parameters, processing of the synthetic and recorded ground motions, and the source description. Then the relative effects of model features mentioned above are quantified through goodness-of-fit (GOF) measures between synthetics and data. Finally, we discuss future research directions based on our results.

2 MODEL FEATURES AND COMPUTATIONAL ASPECTS

In this section we describe the simulation method and model set-up, and summarize the features included in our model. In addition, we outline the processing parameters for simulations and data, and introduce our GOF measures to validate our simulations.

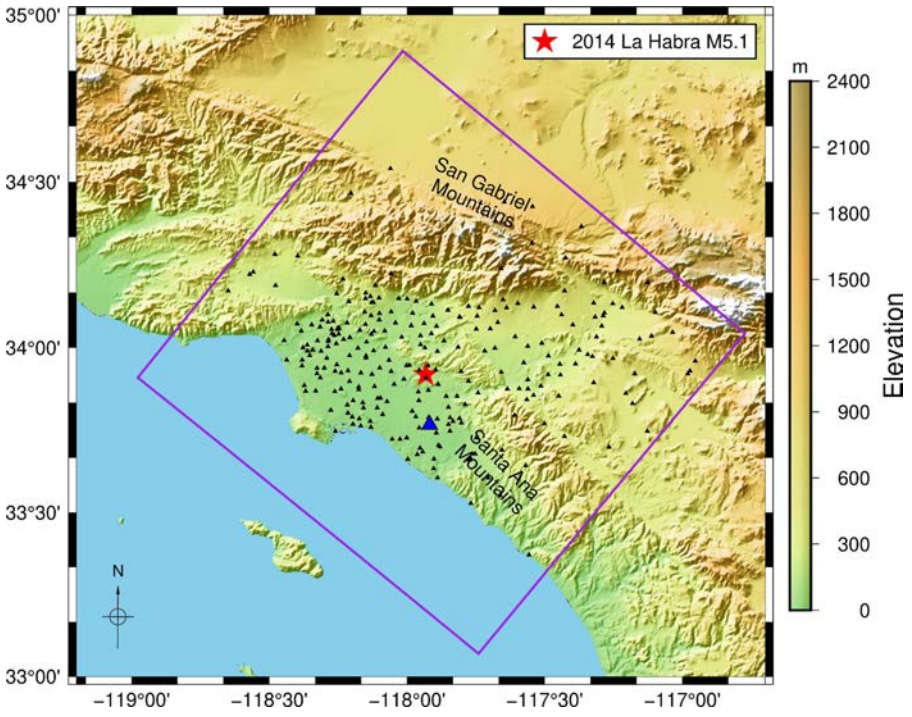

Figure 1. Simulation domain (purple solid rectangle) and locations of 259 strong motion stations used in this study (black triangles). The star denotes the epicentre of the La Habra earthquake. The blue triangle depicts the station CE13884 used to illustrate our 1-D approximation of the low-velocity material in the simulations (Fig. 4). All stations are located at least 4 km from the model boundaries to minimize artificial reflections from the sponge layers. The geographical coordinates of the corners of the simulated domain are listed in Table 1.

2.1 Numerical method for simulating ground motions

We use the staggered-grid finite-difference (FD) code AWP-ODC (Anelastic Wave Propagation, Olsen-Day-Cui, from the authors of the code, hereafter denoted by AWP; Cui *et al.* 2010), which is fourth-order accurate in space and second-order accurate in time, to generate ground motion predictions for the La Habra event. AWP has been adapted to GPU accelerators for kinematic sources (Cui *et al.* 2013) and provides support for frequency-dependent viscoelastic attenuation (Withers *et al.* 2015) and topography using a curvilinear grid (O'Reilly *et al.* 2022). Sponge layers following Cerjan *et al.* (1985) are implemented along the sides and bottom of the model to mitigate spurious energy reflected back to our computational domain.

The accuracy of AWP has been thoroughly verified. For example, large-scale earthquake simulations in realistic 3-D earth models with strong heterogeneities and complex finite-fault source descriptions (Bielak *et al.* 2010, 2016) revealed good agreement between AWP, another staggered-grid FD code and a finite-element code. The implementation of frequency-dependent anelastic attenuation was tested by Withers *et al.* (2015) against a frequency–wavenumber solution, and the accuracy of the curvilinear topography implementation (O'Reilly *et al.* 2022) was verified against SPECFEM3D (Komatitsch & Tromp 2002).

2.2 CVM

We used a model domain of lateral dimensions 148 km × 140 km, rotated 39.9° clockwise with a depth extent of 60 km (see Fig. 1). The mesh was extracted from the SCEC CVM-S4.26-M01, a version of the original CVM-S4 model (Magistrale *et al.* 2000; Kohler 2003) with iterative 3-D tomography updated inversions in Southern California (Lee *et al.* 2011). The SCEC Uniform Community Velocity Model software framework (V19.4; Small *et al.* 2017) was used for the extraction of seismic *P*-wave velocity (V_P), V_S and the material density. The choice of CVM-S4.26-M01 (hereafter abbreviated with CVM-S) for this study was based on the results by Taborda *et al.* (2016) who concluded from a comprehensive validation of four velocity models with 30 earthquakes in the greater Los Angeles region that this model consistently yielded the best fit to ground motion data using a variety of metrics. The model includes the near-surface V_S tapering method proposed by Hu *et al.* (2022), which improved the fit between 0 and 1 Hz synthetic and recorded ground motions for the La Habra event.

2.3 Small-scale heterogeneities

Small-scale crustal heterogeneities (SSHs, on the order of tens to hundreds of metres) are known to exist in nature (e.g. Savran & Olsen 2016) but are insufficiently resolved in state-of-the-art velocity models. Instead, SSHs are commonly included in numerical simulations via statistical models of property fluctuations (e.g. Imperatori & Mai 2013; Savran & Olsen 2019). Here, we superimpose a statistical model of

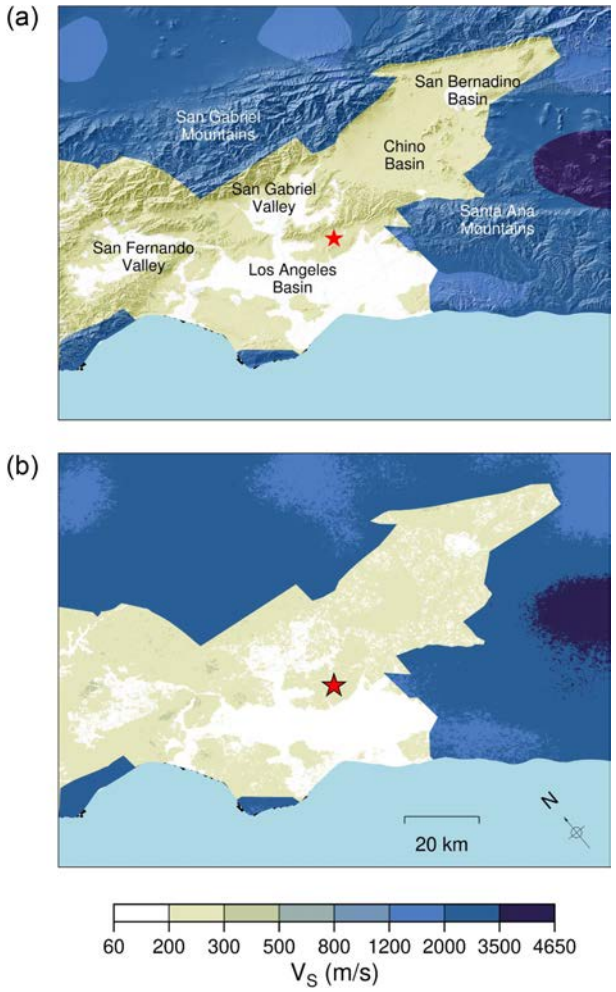

Figure 2. Illustration of the imprint of small-scale heterogeneities at the surface. (a) V_S extracted from the CVM-S. (b) Same as (a) but superimposed with a statistical model of heterogeneities with vertical and horizontal correlation lengths of 100 m and 500 m, respectively, Hurst number of 0.05 and standard deviation of 5 per cent. Topography is removed in (b) for clarity. The epicentre for the La Habra earthquake is depicted with a star.

velocity and density perturbations onto CVM-S, defined via a von Kármán autocorrelation function (Frankel & Clayton 1986):

$$\Phi_{v,a}(r) = \sigma^2 \frac{2^{1-v}}{\Gamma(v)} \left(\frac{r}{a}\right)^v K_v \left(\frac{r}{a}\right) \tag{1}$$

which has Fourier transform:

$$P(k) = \frac{\sigma^2 (2\sqrt{\pi}a)^E \Gamma(v + E/2)^{v+E/2}}{\Gamma(v)(1 + k^2 a^2)} \tag{2}$$

in which k is the wave number and E is the Euclidean dimension, Γ denotes the Gamma function, and K stands for the modified Bessel function of the second kind with order ν. The parameters of the von Kármán autocorrelation function include correlation length a, standard deviation σ and Hurst number ν. This approach generates a random field with zero mean following a Gaussian distribution, and the desired standard deviation is guaranteed by scaling the random variable at each computational node. We used a fixed Hurst number of 0.05 and introduced elliptical anisotropy with a ratio of horizontal-vertical correlation lengths of 5, and tested correlation lengths between 100 and 5000 m, and standard deviations of 5 and 10 per cent, based on previous studies in southern California (e.g. Nakata & Beroza 2015; Savran & Olsen 2016). In our model, the random perturbations extend to a depth of 7.5 km, and then linearly taper to a standard deviation of 0 at 10 km depth, following the results by Olsen *et al.* (2018). Fig. 2 shows an example realization of small-scale heterogeneities, compared to the original CVM-S in terms of V_S at the surface.

2.4 Topography

While the greater Los Angeles region, including the epicentral area of the La Habra event, is characterized by relatively flat topographic relief, the San Gabriel and Santa Ana Mountains bound the area to the north and east, respectively (see Fig. 1). To quantify the effects of topography on ground motions from the La Habra event, we use the curvilinear grid approach by O'Reilly *et al.* (2022). In this version of

Table 1. Simulation parameters used for the deterministic ground motion simulations of the 2014 La Habra earthquake.

Model	
Topography	Yes
Length	147.840 km
Width	140.400 km
Depth	58.000 km
Northwest corner	−118.0154409, 34.8921683
Southwest corner	−118.9774168, 33.9093124
Southeast corner	−117.7401908, 33.0695780
Northeast corner	−116.7729754, 34.0429241
Spatial resolution	
Maximum frequency	5 Hz
Minimum V_S	500 m s^{-1}
Points per minimum *S*-wavelength	5
Grid discretization	20 m
Number of cells	150 486 336 000
Number of CPU processors	756
Number of GPU processors	1512
Wall-clock time	5 hr
Temporal resolution	
Time discretization	0.0006 s
Simulation time	120 s
Number of time steps	200 000

AWP, surface topography is incorporated by stretching the computational grid in the vertical direction, while keeping the horizontal grid spacing unchanged, so that the surface grid locations conform to the shape of the topography. We include surface topography into our model domain via the 1/3 arc-second resolution Digital Elevation Model in southern California from the U.S. Geological Survey (USGS 2020).

2.5 Anelastic attenuation

Anelastic attenuation is needed for accurate simulation of seismic wave propagation through earth models at distances further than the dominant wavelength to account for dissipation of the mechanical energy. Frequency-independent attenuation (constant Q), resulting in identical seismic energy loss per cycle across a frequency bandwidth, has successfully been used to validate ground motion recordings for frequencies up to about 1 Hz (e.g. Olsen *et al.* 2003; Graves & Wald 2004). However, as frequencies increase above about 1 Hz, data often support frequency-dependent Q (e.g. Raoof *et al.* 1999; Eberhart-Phillips *et al.* 2014; Wang & Shearer 2017). To address these findings, Withers *et al.* (2015) developed an efficient coarse-grained memory variable approach to model frequency-dependent attenuation using a power-law formulation

$$Q(f) = Q_0 \left(\frac{f}{f_0}\right)^{\gamma}, \quad f \geq f_0, \tag{3}$$

where Q_0 is a frequency-independent Q value applied for $f < f_0$, which is used in our simulations.

A widely used parametrization of Q_0 is proportional to the local seismic velocity, with separate values $Q_{P,0}$ and $Q_{S,0}$ for V_P and V_S quality factors, respectively, producing expected stronger attenuation for lower velocity material, as pointed out by Hauksson & Shearer (2006). Taborda & Bielak (2014) revised the formula expressed by Brocher (2008) and applied a sixth-order polynomial function for $Q_{S,0}$ from V_S, and $Q_{P,0} = \frac{3}{4}(V_P/V_S)^2 Q_{S,0}$. Here, we test a variety of these parametrizations of Q_0 for the La Habra event. See also Table 2 for a list of the Q models tested, and the comparison of proposed Q_0–V_S relations in Supporting Information Fig. S1.

2.6 Near-surface geotechnical layer

CVM-S includes geotechnical data which integrates geology and geophysics data from surficial and deep boreholes, oil wells, gravity observations, seismic refraction surveys and empirical rules calibrated based on ages and depth estimates for geological horizons in southern California (Magistrale *et al.* 1996, 2000). While recent validation studies, such as Taborda *et al.* (2016), have shown that the basin structure included in CVM-S is reasonably accurate, unrealistically large surface rock site velocities (see Fig. 2) motivated the method by Ely *et al.* (2010) to reduce the V_S in the top 350 m based on available V_{S30} values. Hu *et al.* (2022) proposed a method to optimally taper Vs30 information to the velocities constrained by tomography in the top ∼1 km where lithological information is often poorly characterized. At this depth range, CVM-S delivers unrealistically high shallow Vs estimates for many sites outside the sedimentary basins. Here, Hu *et al.* (2022) tested different tapering depths (z_T) and showed that z_T of about 1000 meters generates the best fit between synthetics and seismic data for the 2014 La Habra earthquake, without compromising the fit at well-constrained sites. In contrast, they found that the widely used method

Table 2. Summary and main features of the models used in this study.

Model ID	Topography	$Q(f)$ *	SSH†	V_S taper depth (m)	GOF (0.15–2.5 Hz)	GOF (2.5–5 Hz)	Figures where shown
1	Yes	$Q_S = 0.1V_S f^{0.6}$	No	1000	5.24	4.01	6–9, 15, S5
2	Yes	$Q_S = 0.1V_S f^{0.3}$	No	1000	5.36	4.64	10, 11, S6
3	Yes	$Q_S = 0.1V_S$	No	1000	FFCC675.43	4.71	7–9, 14, S7
4	Yes	$Q_S = 0.075V_S f^{0.4}$	No	1000	FFCC675.43	68CBD04.75	7, 8, 13
5	Yes	$Q_S = 0.05V_S f^{0.6}$	No	1000	5.26	68CBD04.75	7, S8
6	Yes	$Q_S = Q_{\text{seg}}(V_S) f^{0.4}$ **	No	1000	FFCC675.42	4.70	S9
7	Yes	$Q_S = Q_{\text{poly}}(V_S) f^{0.4}$ ***	No	1000	5.31	68CBD04.79	S10
8	Yes	$Q_S = 0.1V_S f^{0.6}$	No	0	5.22	4.04	9, 15, S11
9	Yes	$Q_S = 0.1V_S$	No	0	5.37	4.45	9, 14, S12
10	Yes	$Q_S = Q_{\text{seg}}(V_S) f^{0.4}$ **	$\sigma = 5\%, a = 100m$	1000	FFCC675.42	68CBD04.79	S4, S13
11	Yes	$Q_S = Q_{seg}(V_S) f^{0.4}$**	$\sigma = 5\%, a = 500m$	1000	5.34	4.62	S4, S14
12	Yes	$Q_S = Q_{seg}(V_S) f^{0.4}$**	$\sigma = 10\%, a = 500m$	1000	5.24	4.32	S4, S15
13	Yes	$Q_S = 0.1V_S f^{0.3}$	$\sigma = 5\%, a = 5000m$	1000	5.20	4.38	10, 11, S4, S16
14	No	$Q_S = 0.1V_S f^{0.6}$	No	1000	5.08	4.00	15, S17
15	No	$Q_S = 0.1V_S f^{0.6}$	No	0	5.06	3.99	15, S15
16	No	$Q_S = 0.1V_S$	No	1000	5.34	68CBD04.78	14, S19
17	No	$Q_S = 0.1V_S$	No	0	5.16	4.52	A1, S20

The largest GOFs are marked by tan (0.15–2.5 Hz) and turquoise (2.5–5 Hz) background colours.
* $Q_P = 2Q_S$.
** $Q_{\text{seg}}(V_S) = 0.075V_S$, for $V_S <= 1500$ m s^{-1}; $Q_S = 0.2V_S - 187.5$, for $V_S > 1500$ m s^{-1}.
*** $Q_{\text{poly}}(V_S) = 10.5 - 16V_S + 153V_S^2 - 103V_S^3 + 34.7V_S^4 - 5.29V_S^5 + 0.31V_S^6$, for V_S in km s^{-1}, see Taborda & Bielak (2014).
†When included, hurst number = 0.05 and horizontal to vertical correlation length ratio = 5; a represents the vertical correlation length.

by Ely *et al.* (2010) using a z_T of 350 m significantly underpredicts low-frequency (<1 Hz) ground motions at sites subject to the generic overlay. Finally, while a uniform tapering depth was tested in the models, the authors suggested that some spatial variability in z_T may further improve their method. We have implemented the recommendations by Hu *et al.* (2022) in the simulations presented here.

2.7 Ground motion simulations

Table 1 lists the parameters used in our simulations. All simulations have the same duration of 120 s and resolve wave propagation up to $f_{\max} = 5$ Hz by at least 5 points per minimum S-wavelength. We use AWP-topo that supports a uniform regular, curvilinear mesh to model wave propagation in composite models with support for topography, with V_S less than 500 m s^{-1} set to (hereafter referred to as clamped at) 500 m s^{-1} to reduce computational cost (additional approximations to account for the effects of material with V_S lower than 500 m s^{-1} are described in Section 3.1). V_P in the low-velocity material is determined by the V_P/V_S ratio from (the un-clamped) CVM-S, and the density is unchanged. We use a kinematic source generated following Graves & Pitarka (2016), which creates finite-fault rupture scenarios with stochastic characteristics optimized for California events. The focal mechanism was taken from the U.S. Geological Survey (strike = 233°, dip = 77°, rake = 49°; USGS 2014) with a moment magnitude 5.1.

2.8 Data processing

259 strong-motion seismic stations were used to validate the simulations. The strong motion recordings (velocity time-series) are obtained from SCEC (F. Silva, personal communication, 2020), with hypocentral distances up to 90 km and signal-to-noise ratios above 3 dB. The processing procedure include the following steps: (1) low-pass filtering of the time-series below 10 Hz using a zero-phase filter; (2) interpolating the time-series linearly to a uniform time step; (3) tapering off at the last 2 s using the positive half of a Hanning window; (4) zero padding the last 5 s; (5) filtering the seismograms to the desired frequencies, for bandwidths 0.15–1, 0.15–2.5, 1–2.5 and 2.5–5 Hz and (6) converting velocities to accelerations by a time derivative. Except for the initial 10 Hz low-pass filter, all filters used a low-cut frequency of 0.15 Hz to avoid noise interference (in the data). Fourth-order Butterworth filters were used in all cases. Finally, our horizontal synthetic seismograms were rotated 39.9° counter-clockwise to provide east-west and north-south polarizations consistent with the recorded seismograms. The same filters were applied to records and synthetics for proper comparison.

2.9 Intensity measures

We use seven different intensity measures to characterize the performance of our ground motion models for the La Habra earthquake, namely peak ground velocity (PGV), peak ground acceleration (PGA), energy duration (DUR), cumulative energy (ENER), response spectral acceleration (RS), smoothed Fourier amplitude spectrum (FAS), as well as Arias intensity (AI). We computed the RS at frequencies linearly spaced from 0.2 to 5 Hz. Cumulative energy is calculated as ENER $= \int V(t)^2 \mathrm{d}t$, where $V(t)$ is the particle velocity and t is time. Both ENER

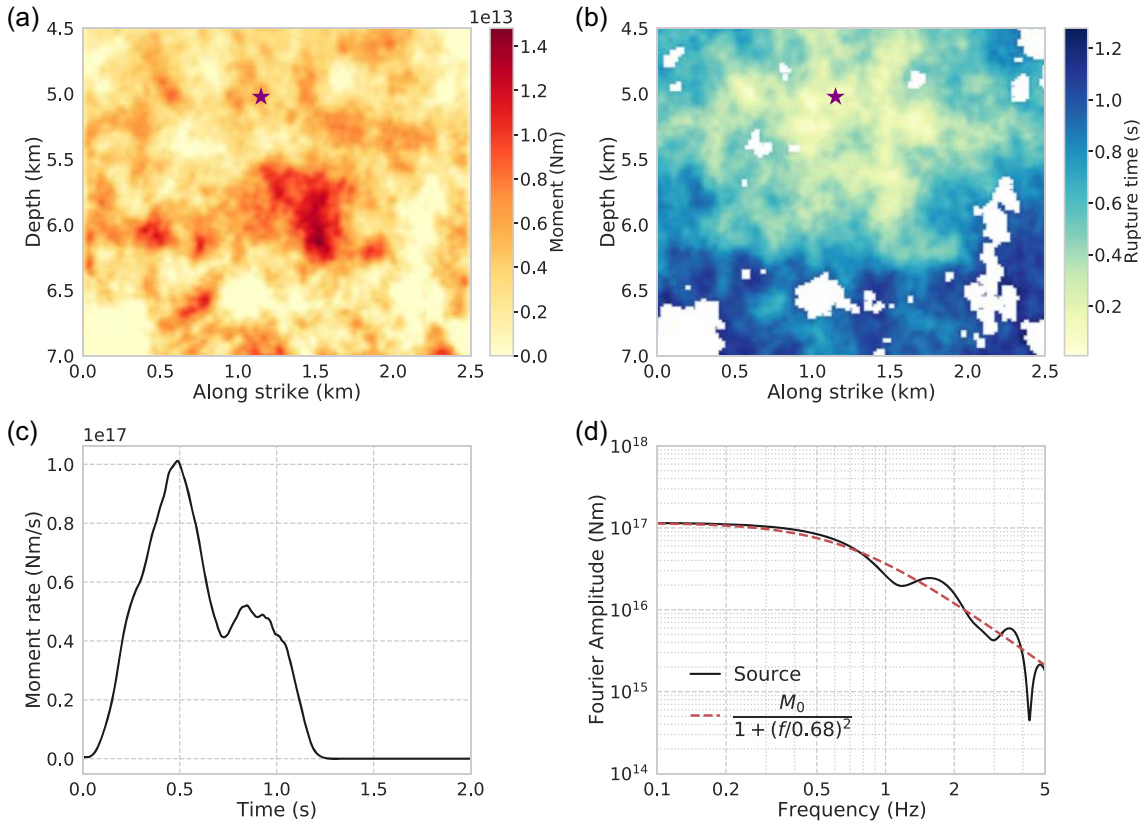

Figure 3. Description of the selected source model used in this study. (a) Moment and (b) rupture time distribution across the fault, (c) sum of the moment rates for all subfaults and (d) Fourier amplitude spectrum. A Brune-type ω^{-2} decay source (Brune 1970) that fits the source spectrum is superimposed in (d) for reference.

and DUR are defined on the interval between the arrival of 5 and 95 per cent of the total energy. Arias intensity (Arias 1970) is defined as $AI = \frac{\pi}{2g}\int a(t)^2 \mathrm{d}t$, where $a(t)$ is the acceleration time-series and g is the gravitational acceleration.

2.10 Source models

Due to the stochastic characteristics of the kinematic source generator by Graves & Pitarka (2016), a series of 40 source realizations following the focal mechanism discussed in Section 2.7 with different random seeds were evaluated based on comparisons between spectral accelerations from records at stations with epicentral distance of 31 km or less (R. Graves, personal communication, 2020). The 40 source models, using a fault area of 2.5 km × 2.5 km, were simulated using CVM-S and evaluated based on the average absolute bias between synthetics and data up to 5 Hz using a weighted average of two metrics: (1) the median pseudospectral acceleration (PSA) rotated over all azimuths (RotD50) at stations within 31 km and (2) the PSA for the north–south and east–west components separately within 13 km of the source, from which we selected the three best performing source descriptions with hypocentral depths at 5, 5.5 and 6 km (see Supporting Information Fig. S2). The rupture duration of the source descriptions is less than 2 s, and they were sampled at an interval of 0.0006 s, identical to the time step used in our simulations. The three sources tend to generate overall similar patterns of PGV within a given bandwidth, such as low (<2.5 Hz) or high (>2.5 Hz) frequencies (see Supporting Information Fig. S3). Based on this result, we carry out our ground motion simulations with the source model with the best fit to the data (Source 1; see Fig. 3) only in order to limit the computational expense. We note that this source displays a Brune-type high-frequency omega-squared fall-off, with realistic values of corner frequency (0.68 Hz) and stress drop (about 10 MPa).

3 RESULTS

Using the source model described in Section 2.10, we carried out 17 simulations in an attempt to determine the model parameters that generate the best fit to the strong motion data for the La Habra earthquake. An overview of the 17 models used in our study, as well as the figures showing the corresponding results, can be found in Table 2. In this section, we discuss the relative effects of the model parameters as obtained from our simulations.

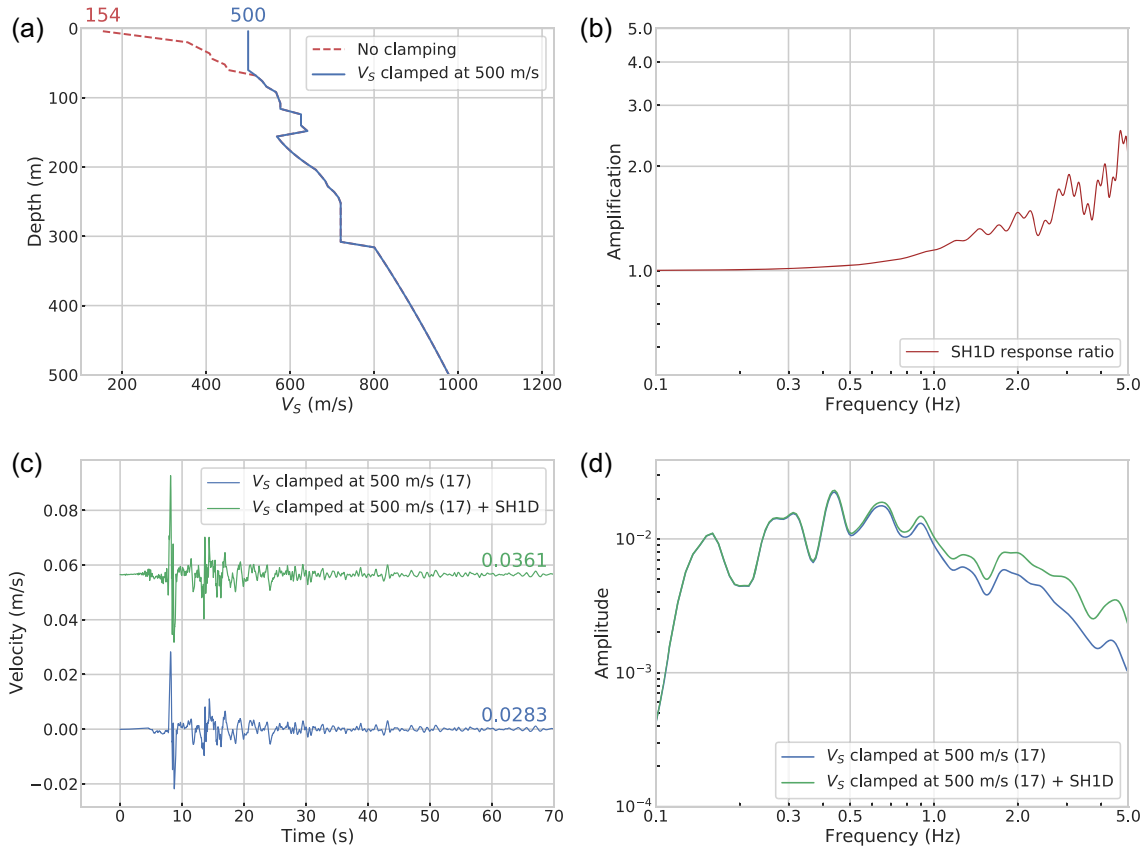

Figure 4. Illustration of the SH1D method used to include the effects of material with V_S less than 500 m s^{-1} in our 3-D simulations for an example site (CE13884, see Fig. 1) in the LA basin. (a) V_S profile extracted from CVM-S (red dashed curve) and V_S clamped at 500 m s^{-1} (blue). (b) SH1D response ratio between the CVM-S profile and that with V_S clamped at 500 m s^{-1}. (c) Synthetics from a 3-D simulation with V_S clamped at 500 m s^{-1}, with and without the SH1D response ratio. The peak velocities are listed to the right of the traces. (d) Smoothed Fourier amplitude spectra corresponding to the waveforms in (c).

3.1 Low-velocity near-surface material

Southern California features several low-velocity basins where the lowest V_S in CVM-S can be much lower than the minimum value of 500 m s^{-1} that we imposed in our models (see Fig. 2). As previous studies have pointed out, soft soils, characterized by lower V_S, can generate significant amplification of ground motions (e.g. Anderson & Hough 1984). Reducing the minimum V_S will, however, increase the computational cost for the series of 3-D numerical simulations needed in our analysis beyond the available resources. For example, simulations in models with minimum V_S at 200 m s^{-1} instead of 500 m s^{-1} requires about 40 times more node hours for a single simulation with constant grid spacing.

For this reason, we use a computationally much less expensive 1-D method, which models vertically incident *SH* waves in a horizontally layered half-space (e.g. Day 1996; Thompson *et al.* 2012), to account for effects of the material with V_S less than 500 m s^{-1}. At each site, we calculated the response from two 1-D models, one using the velocity profile from our models, and the other using the same profile but with V_S clamped at 500 m s^{-1}. In this way, the FAS ratio of the two 1-D results characterizes the effects of the material with V_S less than 500 m s^{-1}, and is then superimposed upon our 3-D simulations (with minimum V_S clamped at 500 m s^{-1}) by multiplication in the frequency domain. Because the SH1D method considers SH waves only, we will apply this calibration to horizontal components only.

Fig. 4 illustrates the results of applying the SH1D method low-velocity correction for an example site (CE13884, $V_{S,\,\mathrm{min}}$=154 m s^{-1}). The two profiles show similar SH1D responses below about 0.3 Hz, above which the SH1D response ratio slowly trends upward with frequency, depicting the amplification from the material with V_S less than 500 m s^{-1}. The PGV of the horizontal synthetic with the correction is increased by up to 32 per cent (largest near 5 Hz) relative to that with V_S clamped at 500 m s^{-1} (Fig. 4c). However, the correction leaves the shape of the waveform almost unchanged. The smoothed Fourier spectra (Fig. 4d) further suggests that such clamping of V_S may be reasonable for frequencies up to 0.8–1 Hz.

We have verified that the 1-D correction is reasonably accurate using a 3-D simulation with minimum V_S of 200 m s^{-1}, see Appendix A. For example, the approximate solution is within about 10 per cent and 25 per cent of the 3-D response on the horizontal components at frequencies less than 2.5 Hz and for 2.5–5 Hz, respectively, while the vertical component is largely unaffected. We do not expect a significant bias of the resulting optimal $Q(f)$ models from the 1-D approximation, which is applied to all of our 3-D simulations.

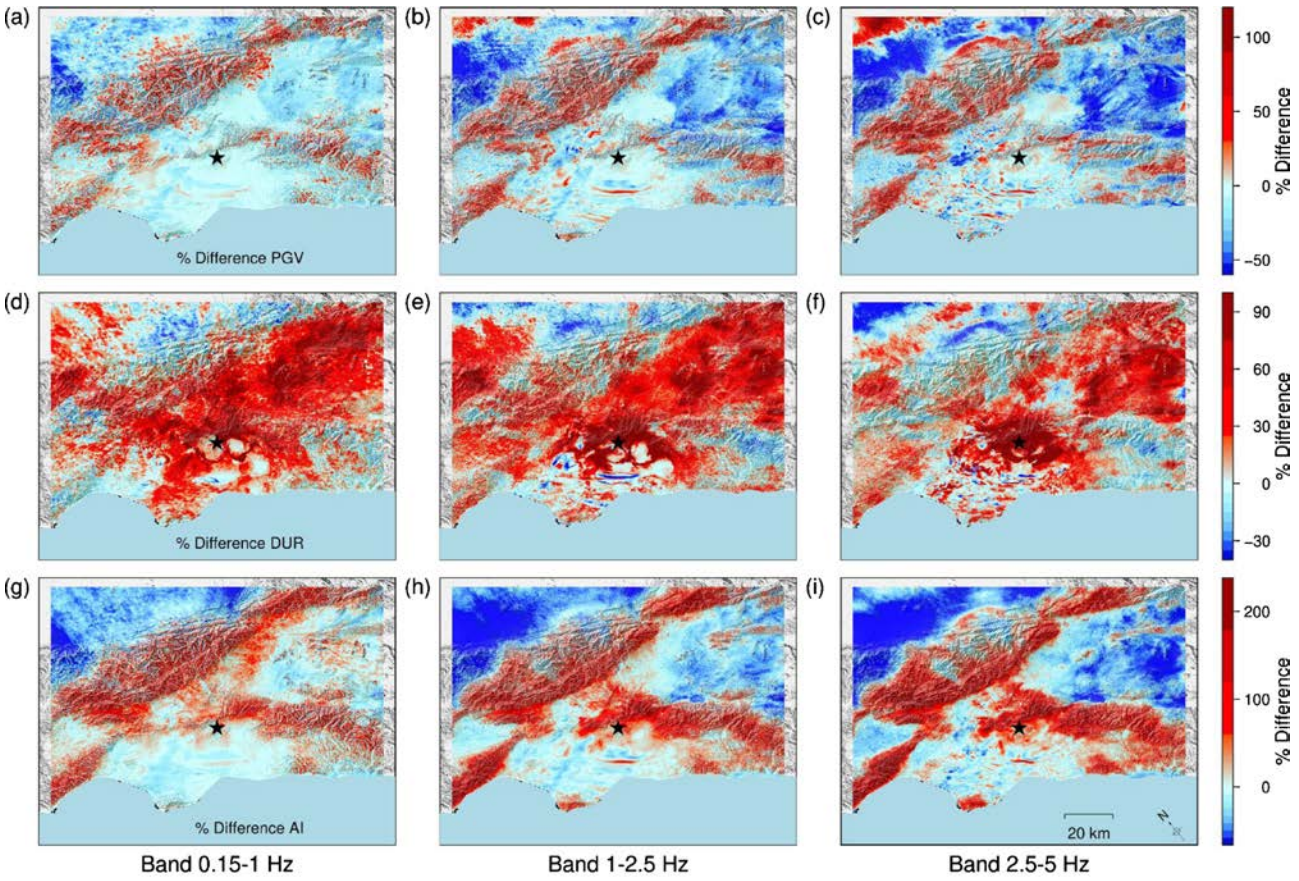

Figure 5. Per cent difference of PGV (top row), DUR (second row), and AI (bottom row) at the surface determined by the model with topography (Model 1) and the model without topography (Model 14) for (left-hand column) 0.15–1 Hz, (centre column) 1–2.5 Hz and (right-hand column) 2.5–5 Hz. Positive (negative) values coloured in red (blue) indicate amplification (deamplification). The star denotes the epicentre of the La Habra event.

3.2 Topography

In this section we investigate the effects of topography, which are often ignored in numerical simulations of wave propagation (e.g. Graves & Wald 2004; Olsen *et al.* 2006; Savran & Olsen 2019). We note that the curvilinear grid in the topography model distorts the Cartesian grid in the vertical direction to be identical to the non-topography model at any given depth below sea level. Fig. 5 shows the per cent difference between models with and without topography for PGV, DUR and AI for bandwidths of 0.15–1, 1–2.5 and 2.5–5 Hz. It is clear that topography complicates the wavefield pattern significantly, even at frequencies below 1 Hz, in terms of DUR and thus AI. Consistent with previous studies (e.g. Hartzell *et al.* 1994; Lee *et al.* 2009), we observe weak deamplification of PGV below 1 Hz in basin areas, while mountain peaks and ridges may amplify PGV by up to 50 per cent. In addition, we find that PGV is reduced by about 30 per cent in the Chino Basin and northwest of San Gabriel Mountains. These results are in agreement with Ma *et al.* (2007) who found that the San Gabriel Mountains scatter surface waves from a northern rupture on the San Andreas Fault and reduce the PGVs in the LAB by up to 50 per cent. We interpret these results as shielding and focusing effects on the front (e.g. the south slope of the San Gabriel Mountains) and back (e.g. the north slope of the San Gabriel Mountains) sides of the mountains, respectively, which become more significant at higher frequencies, in agreement with Liu *et al.* (2020). In addition, as frequencies increase above 2.5 Hz, we observe a clear pattern of 'amplification–deamplification–amplification' along the N–S (short) axis of the San Gabriel Mountains, which is predicted in the numerical experiments by Liu *et al.* (2020).

It is particularly notable that DUR within 10 km of the source is strongly increased for both low and high frequencies, mostly to the north (northwest end of the Santa Ana Mountains). Here, topography seems to act as a significant source of scattering that increases the wave duration on the sides of the mountain facing the incoming wavefields, while DUR is reduced on the 'back' sides of the mountain as seen from the source location. At further distance from the source, our results show a clear negative correlation between the effects on PGV amplification and DUR lengthening, suggesting that topography redistributes seismic energy from the large-amplitude first arrivals to the adjacent coda waves. These results are in agreement with Lee *et al.* (2009) who noted that the effects from topography can interfere with those from path and directivity.

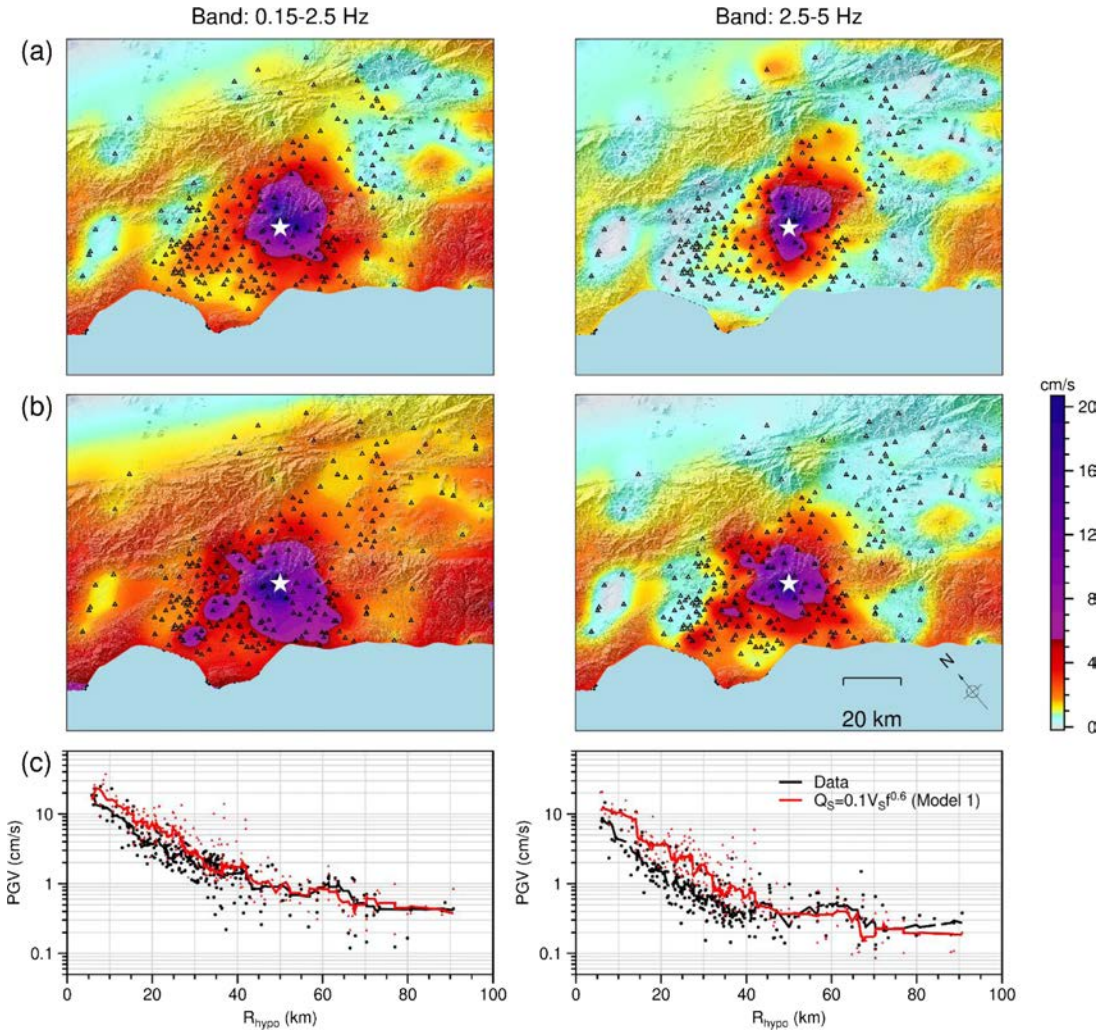

Figure 6. Comparison of interpolated PGVs measured at 259 stations, depicted by triangles, for (a) data, (b) synthetics using Model 1 (including topography, 1000 m shallow velocity taper and frequency-dependent attenuation, see Table 2). The star denotes the epicentre of the La Habra earthquake. (c) PGV as a function of $R_{\rm hypo}$ for data and synthetics. The left- and right-hand columns show band-limited results for 0.15–2.5 and 2.5–5 Hz, respectively.

3.3 Anelastic attenuation

We tested a range of different parameter settings for the frequency-dependent anelastic attenuation implementation in AWP (see Table 2), with their efficacy reported in conjunction with other features in this study. Note that all of our models use $Q_P/Q_S = 2$, an assumption often used in southern California (e.g. Withers *et al.* 2019). Fig. 6 illustrates the areal pattern and distance dependence of PGVs with Model 1 using $Q_S = 0.1 V_S f^{0.6}$, one of the models providing favourable fit to data for simulations of the 2008 Chino Hills earthquake by Savran & Olsen (2019). While the simulated PGVs provide an overall good fit to data for 0-2.5 Hz as reported by Savran & Olsen (2019), the PGVs for distances less than about 40 km significantly overpredict the data for 2.5–5 Hz.

Fig. 7 summarizes some of the effects of our tests by a comparison between the FAS and corresponding bias from Models 1, 3, 4 and 5 (see Table 2). Among the four attenuation models in Fig. 7, $Q_S = 0.075 V_S f^{0.4}$ (Model 4) fits the vertical component of the data the best, and $Q_S = 0.1 V_S$ (Model 3) provides the best fit to the FAS of the data for the horizontal components. $Q_S = 0.05 V_S f^{0.6}$ (Model 5) underpredicts the FAS of the data between about 0.7 and 2 Hz on both horizontal and vertical components. A power-law exponent of 0.6 overpredicts the high-frequency FAS on the horizontal components, particularly for Model 1, while the exponent of 0.4 used for Model 4 provides a much improved fit to data. Thus, a scaling of $Q_{S,0}$ with V_S by factors of 0.05 and lower as well as higher than 0.1, and using γ larger than 0.6 appear to worsen the fit to data. Two additional models with sharper increase in $Q_{S,0}$ for higher V_S as compared to lower V_S (segmented, Model 6, and polynomial, Model 7; Supporting Information Fig. S1) do not significantly improve the fit to data (see Supporting Information Figs S9 and S10)

The differences in the fit for horizontal and vertical components mentioned above (Fig. 7) may indicate a different (and possibly depth-dependent) Q_P/Q_S relationship, as the vertical component is generally dominated by P waves. However, another reason for the discrepancy

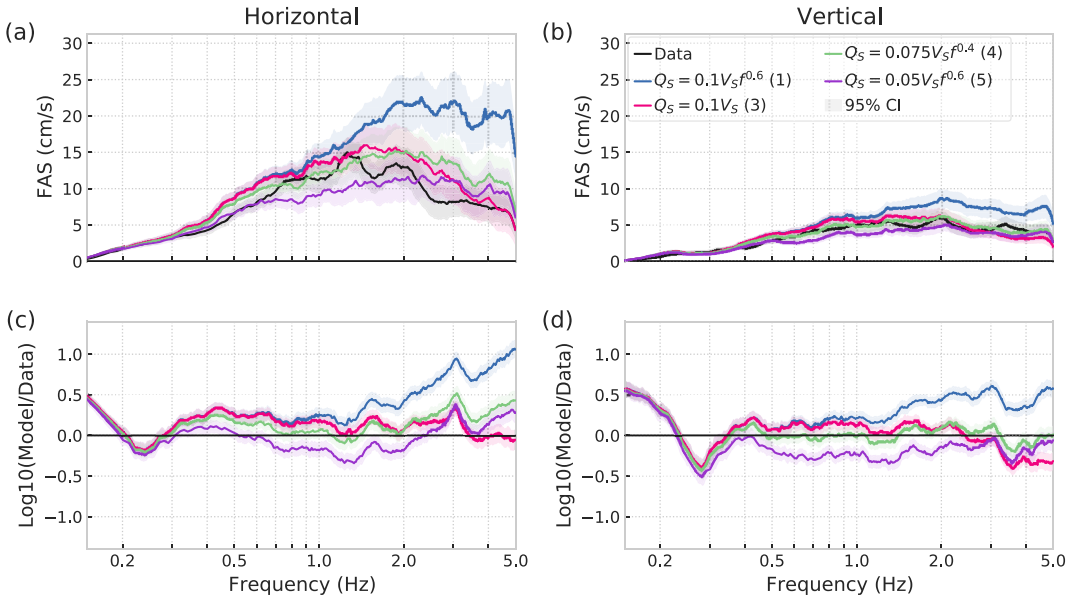

Figure 7. FAS computed from records and models with various attenuation models (enclosed numbers in parentheses in the legend indicate model IDs, see Table 2). The left-hand (right-hand) column shows results for the horizontal (vertical) components. The top row shows the FAS amplitudes and the bottom row shows the FAS bias between models and records, calculated as the 10-based log between simulations and data. The solid lines depict the median FAS over all 259 stations. The shading shows the 95 per cent confidence interval (CI).

may be the application of the 1-D correction in the horizontal direction only. While we have focused on optimizing Q_S in this study, we recommend exploration of additional Q_P/Q_S models as future work.

As for FAS, the distance decay of PGVs varies significantly between various $Q(f)$ models (see Fig. 8). $Q_S = 0.1V_S f^{0.6}$ (Model 1) strongly overpredicts the high-frequency (>2.5 Hz) PGVs in the basins towards the south and west, while providing a fairly good match at distances greater than 40 km. The models $Q_S = 0.075V_S f^{0.4}$ (Model 4) and $Q_S = 0.1V_S$ (Model 3), on the other hand, generate moderate overprediction in the near-source regions while underpredicting the PGVs at farther distances.

The presence of realistic shallow low velocities at rock sites is crucial in determining the best-fitting Q models. Fig. 9 shows a comparison of PGV and DUR for models with and without shallow velocity taper. As the 1000 m taper lowers the shallow velocities, the PGVs are increased as required by data, despite lower Q values following the Q_S–V_S relations, in particular beyond 30–40 km, where most rock sites are located. Note also that γ less than 0.6 are needed to aleviate significant overprediction of DUR.

3.4 SSHs

Fig. 10 shows the effects on PGV and DUR from adding a von Kármán distribution of SSHs with $\sigma = 5$ per cent and horizontal correlation length of 5000 m compared to a reference model without SSHs. Our results indicate that realistic distributions of SSHs generate more strongly scattered ground motions as compared to those by topography with stronger effects at farther distance and at higher frequencies. These findings are in general agreement with Savran & Olsen (2019), who studied a smaller region up to 2.5 Hz, roughly in the centre of our simulated domain.

Przybilla *et al.* (2009) used elastic radiative transfer theory to show that the directional dependence of scattering can be identified by ak, where a is the correlation length and k is the wave number. For $ak \approx 1$, waves interact with a heterogeneous medium most intensively because the wavelength and correlation length are on the same order. When $ak \gg 1$, waves are predominantly scattered in the forward direction, which generates focusing in the early arrivals and leads to larger peak amplitudes, and vice versa for $ak \ll 1$. Fig. 11, showing the case with $a = 5000$ m and $\sigma = 5$ per cent, indicates a regime in between forward and backward scattering, with an increase of PGVs and lengthening of durations averaged across the region of less than 10 per cent. The variability of the ground motions introduced by the SSHs range from 20 per cent to 41 per cent for PGVs and durations respectively, at one standard deviation, increasing with frequency. The results for additional realizations with various standard deviations and correlation lengths of the random field show that larger standard deviations and correlation lengths tend to yield stronger SSH effects (see Supporting Information Fig. S4).

4 DISCUSSION AND CONCLUSIONS

In this study, we have explored the effects of a series of different model features on the resulting ground motions for the 2014 M5.1 La Habra, CA, earthquake. However, trade-offs among the parameters complicate determining a unique set of model parameters creating a best fit to the

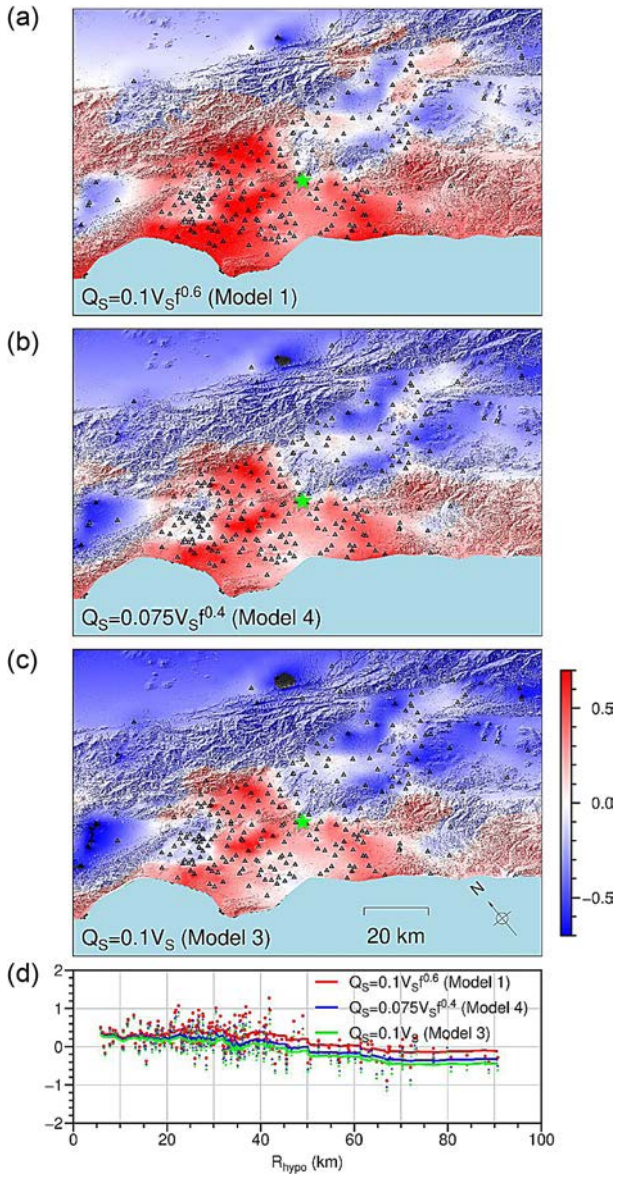

Figure 8. (a––c) Spatial distribution of the three-component bias for PGV, band-pass filtered between 2.5 and 5 Hz. The bias values are computed as the base 10 logarithm of the ratio between simulations and records at each strong motion site. Positive (negative) values represent overprediction (underprediction). (d) Moving average of the bias of PGV using a 20-point window from the three Q models (see Table 2) shown in (a)–(c) versus hypocentral distance.

data. In order to quantitatively rate the performance of the different model features, we used a modified subset of the GOF metrics proposed by Anderson (2004) and Olsen & Mayhew (2010) for the comparison of broad-band seismic traces (0 to 10+ Hz). We define the GOF score for each metric as

$$G_{\text{metric}} = 10\,\text{erfc}\left(\frac{2|x-y|}{x+y}\right), \tag{4}$$

where x and y are two positive scalars from the selected metrics, and erfc is the Error function. G_{metric} is computed for each metric and combined into a weighted average using all three components. We used weights of 0.5, 0.5, 1, 0.5, 0.5, 1 and 1 for PGV, PGA, DUR, AI, ENER, RS and FAS, respectively, where the reduced weights are chosen due to correlation between metrics (Olsen & Mayhew 2010). The GOF score for the entire simulation is calculated as the average across all 259 stations (G_{station}). GOF values between two signals above 4.5 and 6.5 are considered fair and very good fits, respectively.

The GOF scores for the models (see Table 2) explored in this study are shown in Fig. 12 (see also Fig. 13 and Supporting Information Figs S5–S20 for results for individual models). Note that the model ID does not represent any ranking of their GOF against seismic records. The models generally achieve GOF in the range of 5.1–5.4 for the low frequencies (<2.5 Hz), and about an increment lower for the high frequencies (>2.5 Hz), due to increased uncertainty in the source description and model parameters of the former.

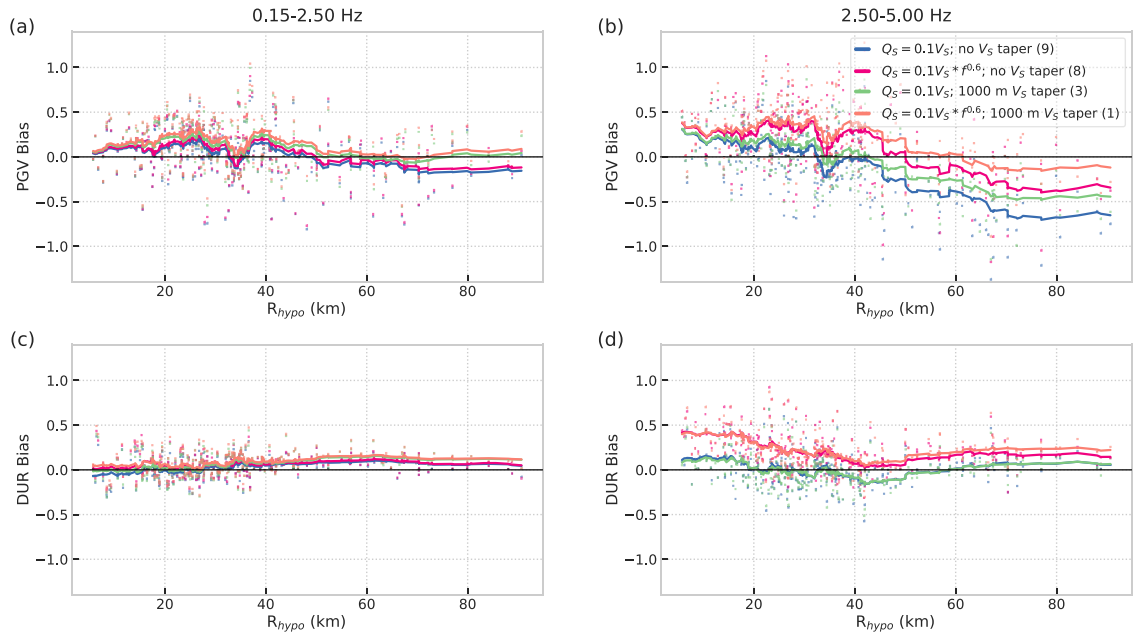

Figure 9. Bias of (a and b) PGV and (c and d) DUR for passbands (left) 0.15–2.5 Hz and (right) 2.5–5 Hz at all 259 stations. The bias is calculated in the same way as for Fig. 7. The solid lines depict the moving average of the bias using a 20-point window for each of the Q models (model IDs enclosed in parentheses in the legend; see Table 2) versus hypocentral distance. 'Taper' refers to the maximum V_S modification depth described by Hu *et al.* (2022).

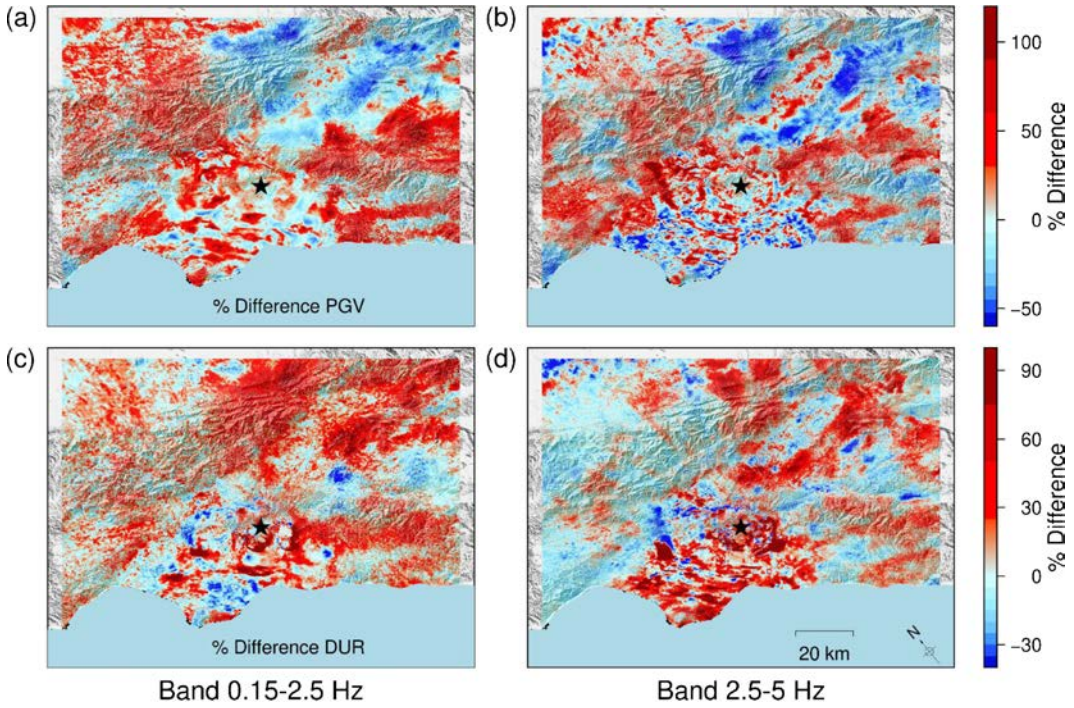

Figure 10. Difference in (top row) PGV and (bottom row) DUR from Model 13, including SSH with $\sigma = 5$ per cent and $a = 5000$ m, versus Model 2 (no SSHs). Left-hand (right-hand) columns show results for bandwidths 0.15–2.5 Hz (2.5–5 Hz). The star depicts the epicentre.

The highest GOF value for models including for 0–5 Hz include Models 2–4, 6–7, 10 and 16, favouring $Q_{S,0}/V_S$ values of 0.075–0.1 and γ values of 0–0.6. However, the models with frequency-independent Q (see Fig. 9, Models 3 and 9) tend to generate too large attenuation for the higher frequencies at distances larger than 40 km, and we recommend using $\gamma > 0$ for future simulations in the Los Angeles area, which is in agreement with the inverted value of 0.4 from the study by Lin & Jordan (2018) using P and S waves in Southern California. The presence of topography generally produces higher GOF values than models without topography (except Model 16) for 2.5–5 Hz. In general,

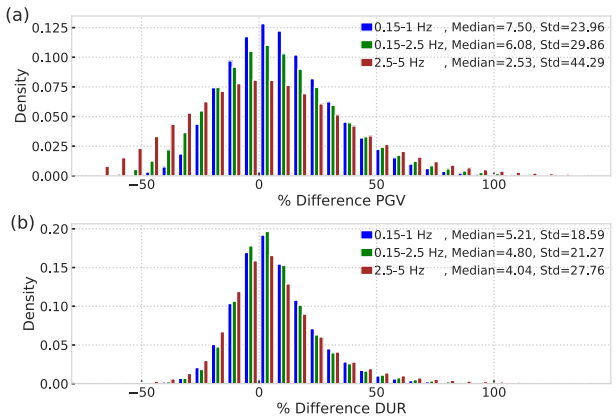

Figure 11. Probability density histogram of the difference between Model 13 (including SSH with $\sigma = 5$ per cent and $a = 5000$ m) and Model 2 (no SSHs). The definition of per cent difference (x-axis) is the same as in Fig. 10.

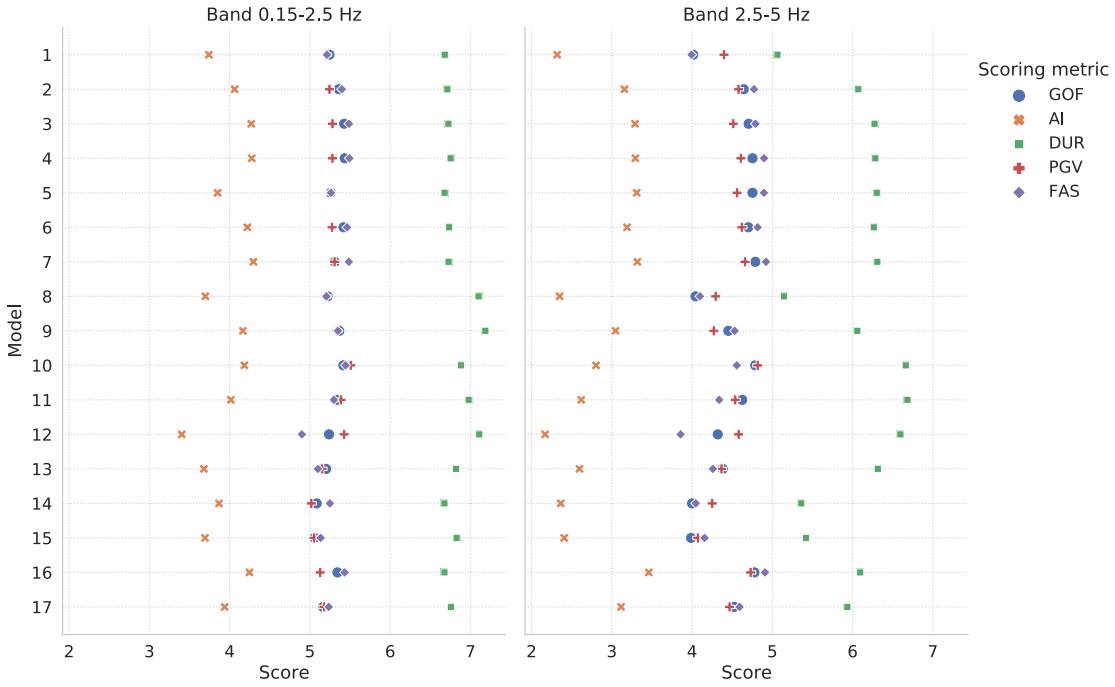

Figure 12. GOF scores for a subset of the metrics used in this study, for frequency bands 0.15–2.5 Hz and 2.5–5 Hz. Model IDs are listed in Table 2.

the models with taper depth of 1000 m yield better GOF, as compared to the models with shallower taper depth or no velocity taper. The simulations with SSHs constrained by borehole data (Models 10–13) fail to significantly improve the GOF.

Hu *et al.* (2022; this issue, companion paper) showed that reducing shallow velocities in CVM-S at poorly constrained sites in the greater Los Angeles area by a generic taper function based on V_{S30} was able to improve the fit of ground motion synthetics to data for the La Habra earthquake, particularly in regions constrained by limited geological information. However, their tests were restricted to frequencies below 1 Hz and models with a flat free surface. Here, we examine the efficacy of the method for frequencies up to 5 Hz, while adding topography to the models from Hu *et al.* (2022). As in their study we divided the 259 strong motion recording sites into two groups: type A sites representing sites with good geological constraints, and type B sites with poor geological constraints, characterized by unrealistically large surface V_S (>1000 m s^{-1}) and nearly constant V_S in the top 500–1000 m in CVM-S. Fig. 14 shows the median FAS for both types of sites from various models (see Table 2). As observed by Hu *et al.* (2022), type A sites are largely unaltered by the shallow velocity tapering methodology (Model 3 versus Model 9), while the original CVM-S (Model 9) significantly underpredicts the FAS at type B sites up to 5 Hz. On the other hand, including the 1000 m V_S taper results in a much improved fit to data for type B sites, with the 1-D correction included, for the entire bandwidth up to 5 Hz.

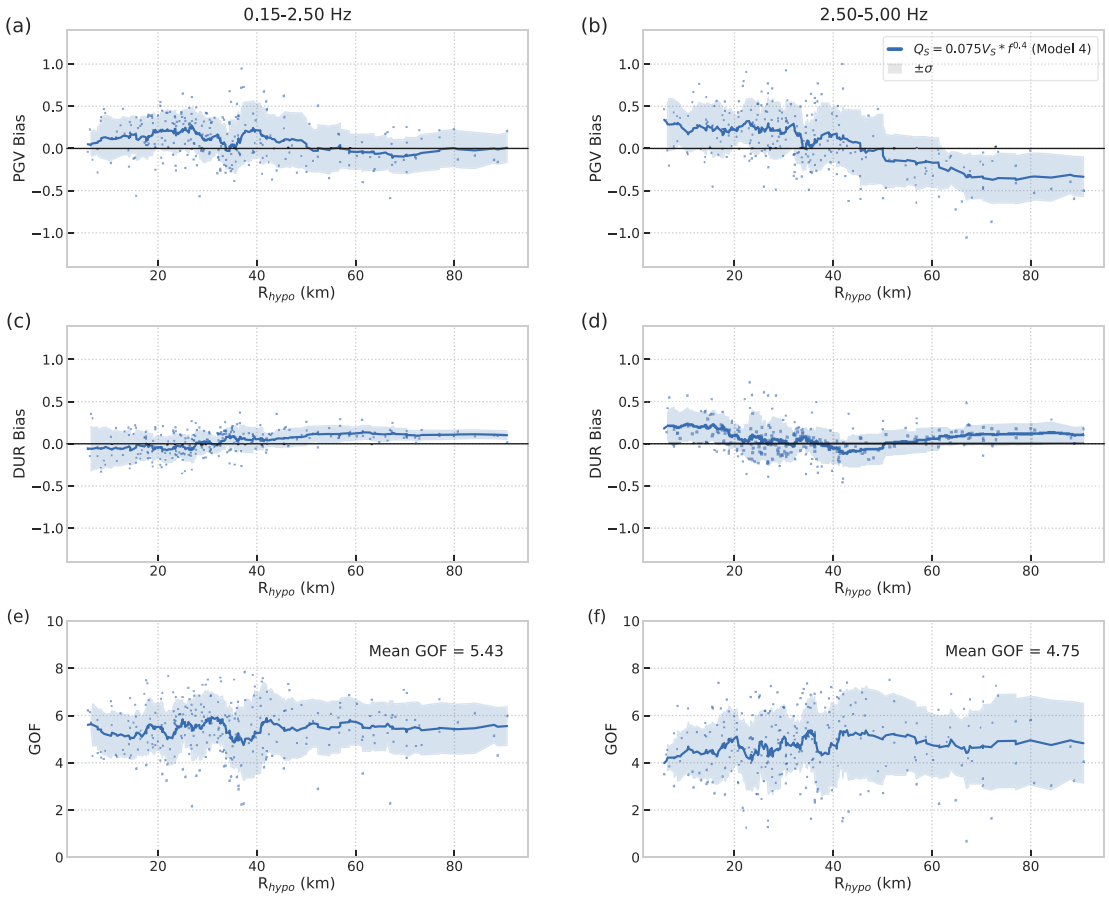

Figure 13. Bias of (top row) PGV, (centre row) DUR and (bottom row) GOF for bandwidths (left-hand column) 0.15–2.5 Hz and (right-hand column) 2.5–5 Hz at all 259 stations for Model 4 (see Table 2 for a list of model features). The bias is calculated in the same way as for Fig. 9. The solid line depicts the moving average using a 20-point window. The shading denotes the standard deviation centred at the mean.

Topography appears to cause negligible effects in terms of FAS (see Fig. 14). Previous studies attempting to capture the effects of topography on ground motions and establish proxies to characterize such effects typically have used simple homogeneous models of earth material, for example Maufroy *et al.* (2015) and Rai *et al.* (2017). These studies found that topographic curvature is a good proxy characterizing irregular surfaces for evaluating topographic effects. However, assessment of topographic effects on ground motions is complicated by amplification due to the presence of a shallow weathering layer of low velocities, typically present in mountain regions, but omitted in these earlier studies. We re-assess these findings including the amplification effects from the modification of near-surface material at type B sites proposed by Hu *et al.* (2022). We calculated the smoothed curvature of topography with a smoothing window of 640 m. Here, steeper relief is characterized by larger curvature values, while flatter regions are of curvature close to zero. Fig. 15 shows the per cent difference in PGV caused by including topography for varying curvature in the simulated region. We show the response of two models, one with 1000 m V_S taper (Hu *et al.* 2022) and one without (original CVM-S). The model without the near-surface low velocities introduced by the V_S taper method tends to reduce the PGVs by about up to 40 per cent for most curvatures below 2.5 Hz, with a broader spectrum of de-amplification (up to 75 per cent) and amplification (up to 40-100 per cent, most pronounced for the largest curvatures) at frequencies between 2.5 and 5 Hz. These trends, however, become much more notable when the shallow low velocities are present, where steep topography (e.g. mountain summits and local steep hills) increases PGVs. This result may partly explain why in previous studies, omitting the near-surface model complexity, topographic effects on ground motions tend to underestimate the amplification at mountain tops compared to observations (e.g. Pischiutta *et al.* 2010; Lovati *et al.* 2011).

The accuracy of the source description is critical for obtaining reliable estimates of the parameters controlling model features such as $Q(f)$. The overprediction of the PGVs for near-source epicentral distances and distance decay faster than that for data at further distances (see Figs 8, 9 and 13) may at least partly be explained by uncertainty in the source moment and/or fault area. For example, the most commonly used empirical magnitude-area scaling relations suggest a source area about 60 per cent larger than that used in our study (Leonard 2010). While the moment is likely relatively well constrained, a somewhat larger fault area may decrease near-source PGVs and modify the pattern of near- and intermediate-field ground motions, e.g., due to directivity effects. While exploration of source variability in this study was limited by computational resources, we recommend using ensembles of sources to further examine the model features in future work, including some generated by different approaches (e.g. Savran & Olsen 2020).

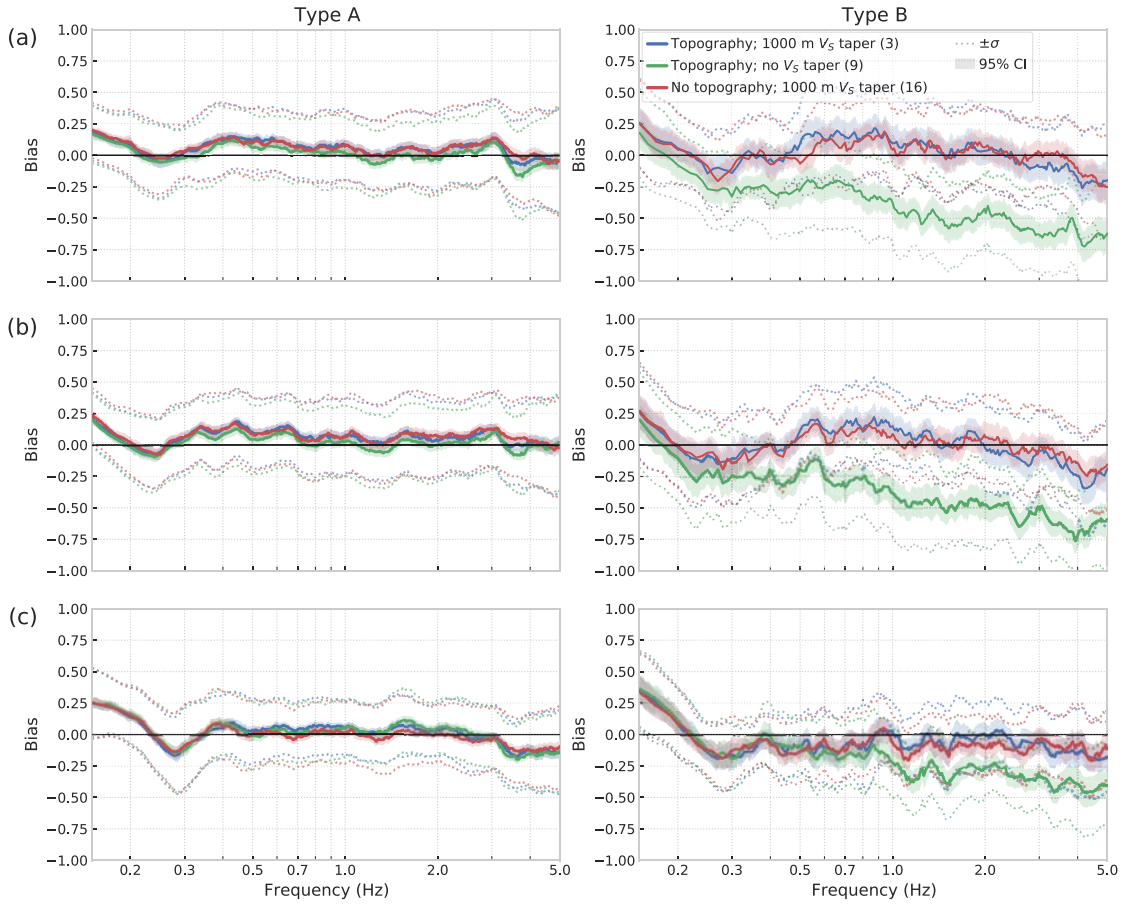

Figure 14. Bias of FAS on the (a) east–west, (b) north–south and (c) vertical components, calculated from models with and without topography and 1000 m velocity taper. The models have the same attenuation rule of $Q_S = 0.1V_S$, and their model IDs are enclosed in the parentheses in the legend. A positive (negative) value depicts overprediction (underprediction). The left- and right-hand columns show type A (low near-surface V_S) and B (high near-surface V_S) sites, respectively. The solid lines depict the median of FAS, where the narrow band is the 95 per cent confidence interval of the median, and the dashed lines depict the standard deviation centred at the median.

The consistent overprediction of the PGVs within 40 km for all Q models (see e.g. Fig. 9) may be caused by the omission of the near-surface low-velocity material (<500 m s^{-1}) during the inversion for the optimal source description, used in this study. Another potential cause of the near-source overprediction could be a very low Q in the shallow sediments (more abundant in the near-source area as compared to larger epicentral distances), as proposed by Hough *et al.* (1988) and examined in the numerical simulations by Withers *et al.* (2019). The presence of such thin, near-surface layer with very low Q would also alleviate at least part of the overprediction of the duration (DUR), as obtained for most models. On the other hand, it is intriguing that the PGVs predicted by all tested Q models appear to decay faster than data beyond 40 km, roughly at the boundaries of the LAB. This observation may suggest unrealistically large contrasts in the shear impedance near the basin boundaries in the CVM, causing excessive entrapment of waves (see Fig. 2). Imperatori & Gallovič (2017) came to a similar conclusion and was able to alleviate the trend by smoothing the velocity model. Another explanation for the underprediction of ground motions beyond epicentral distances of 40 km could be a different Q_S–V_S relation in the surrounding mountain areas compared to that in the sedimentary basins. We recommend further research into these discrepancies.

In addition to the source description, inaccuracies in the velocity structure further complicates estimation of the optimal model parameters, in particular for the Q model, as pointed out by Savran & Olsen (2019) and Lai *et al.* (2020). As a result, findings from previous studies may have produced biased values of γ. For example, Withers *et al.* (2015) estimated higher γ values (near 0.8) from their modelling of the Chino Hill earthquake. However, this estimate was obtained with a CVM-S with near-surface rock velocities biased high, as well as smaller source–station distances and lower maximum frequency. Savran & Olsen (2019) used γ equal to 0.6, but with limited high-frequency resolution up to 2.5 Hz. Nevertheless, future improvement in community velocity models, wider access to computational resources, more efficient numerical codes and guidance from this study are bound to further constrain the ground motion models, leading to more accurate seismic hazard analysis.

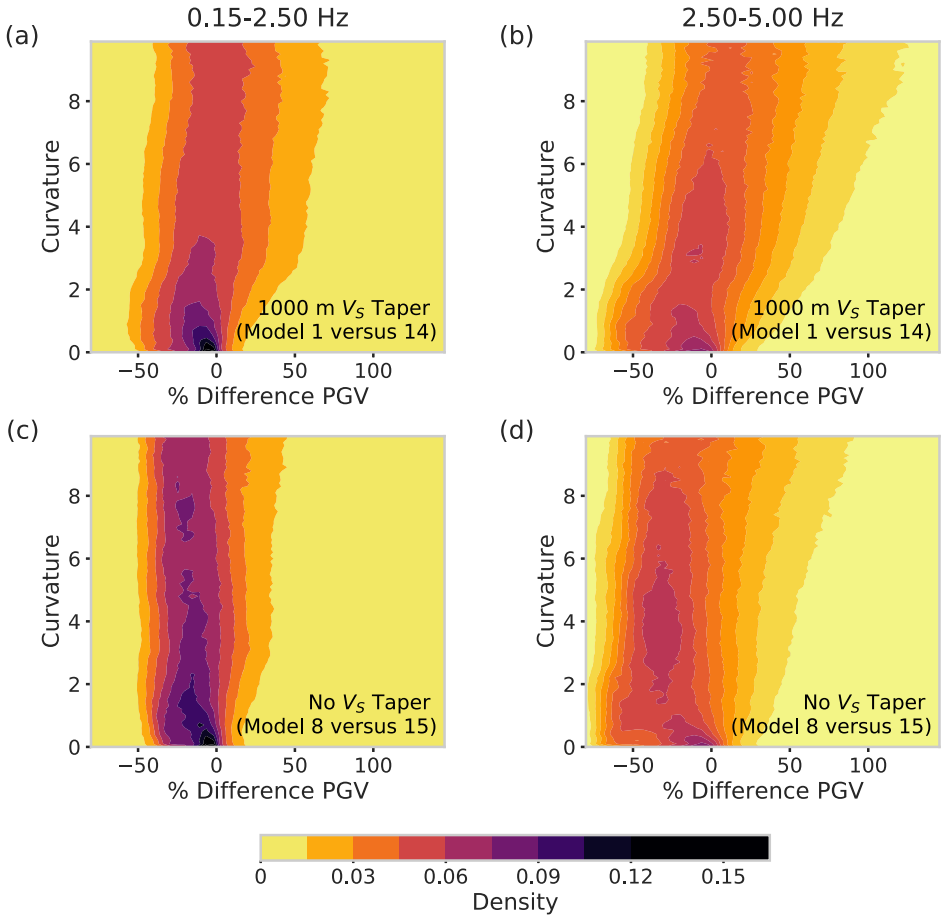

Figure 15. Density of PGV change for models with topography (Model 1 and Model 8) relative to models without topography (Model 14 and Model 15) for bandwidths of (left-hand column) 0.15–2.5 Hz and (right-hand column) 2.5–5 Hz, and models with (top row) and without (bottom row) modified shallow velocities (Hu *et al.* 2022). The *y*-axis depicts topographic curvature smoothed using a 2-D window of dimensions 640 m × 640 m. Values towards the top right (bottom left) denote strong amplification at steep topography (deamplification at flat topography). Note that density intervals do not correspond to constant bin sizes.

ACKNOWLEDGMENTS

This research was supported through the U.S. Geological Survey Earthquake Hazards Program (award #G19AS00021), as well as the Southern California Earthquake Center (SCEC; Contribution Number 11781). SCEC is funded by the National Science Foundation (NSF) Cooperative Agreement EAR-1600087 and the U.S. Geological Survey (USGS) Cooperative Agreement G17AC00047. We thank Robert W. Graves for providing the source models, Fabio Silva for providing the station records of the 2014 La Habra earthquake and the 'High-F' collaboration in SCEC, from which this work was initiated. Thorough reviews by Frantisek Gallovic, Alice-Agnes Gabriel and Duncan Agnew greatly improved our paper.

DATA AVAILABILITY

The digital elevation model is downloaded from the national map maintained by USGS at https://apps.nationalmap.gov/downloader/ (last accessed October 2020). The UCVM program used to extract velocity meshes can be obtained from SCEC at https://github.com/SCECcode/UCVMC (last accessed December 2020). The simulations were performed on Summit at the Oak Ridge Leadership Computing Facility in Tennessee. Most of the data-processing work was done using Python and the Generic Mapping Tools package (https://www.generic-mapping-tools.org, last accessed April 2021). Configuration parameters, input files and synthetic IMs for this work are available at https://doi.org/10.5281/zenodo.6403870. AWP-ODC is open source and freely available at https://github.com/HPGeoC/awp-odc-os.

REFERENCES

Anderson, J.G., 2004. *Quantitative Measure of the Goodness-of-Fit of Synthetic Seismograms,* pp. 243, Earthquake Engineering Research Institute.

Anderson, J.G. & Hough, S.E., 1984. A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, *Bull. seism. Soc. Am.,* **74**(5), 1969–1993.

Arias, A., 1970. Measure of earthquake intensity. *Seismic Design for Nuclear Power Plants,* 438–483.

Assimaki, D., Kausel, E. & Gazetas, G., 2005. Soil-dependent topographic effects: a case study from the 1999 Athens earthquake, *Earthq. Spectra,* **21**(4), 929–966.

Bard, P.-Y., 1982. Diffracted waves and displacement field over two-dimensional elevated topographies, *Geophys. J. Int.,* **71**(3), 731–760.
Bielak, J. *et al.*, 2010. The ShakeOut earthquake scenario: verification of three simulation sets, *Geophys. J. Int.,* **180**(1), 375–404.
Bielak, J. *et al.*, 2016. Verification and validation of high-frequency (f_{max} = 5 Hz) ground motion simulations of the 2014 M 5.1 La Habra, California, earthquake, in *AGU Fall Meeting 2016,* Abstract S33G–04.
Boore, D.M., 1972. A note on the effect of simple topography on seismic SH waves, *Bull. seism. Soc. Am.,* **62**(1), 275–284.
Bouchon, M. & Barker, J.S., 1996. Seismic response of a hill: the example of Tarzana, California, *Bull. seism. Soc. Am.,* **86,** 66–72.
Brocher, T.M., 2008. Compressional and shear-wave velocity versus depth relations for common rock types in northern California, *Bull. seism. Soc. Am.,* **98**(2), 950–968.
Brune, J.N., 1970. Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. geophys. Res.,* **75**(26), 4997–5009.
Cerjan, C., Kosloff, D., Kosloff, R. & Reshef, M., 1985. A nonreflecting boundary condition for discrete acoustic and elastic wave equations, *Geophysics,* **50**(4), 705–708.
Cui, Y. *et al.*, 2010. Scalable earthquake simulation on petascale supercomputers, in *2010 ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis,* pp. 1–20, IEEE.
Cui, Y. *et al.*, 2013. Physics-based seismic hazard analysis on petascale heterogeneous supercomputers, in *SC'13: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis,* pp. 1–12, IEEE.
Davis, L.L. & West, L.R., 1973. Observed effects of topography on ground motion, *Bull. seism. Soc. Am.,* **63**(1), 283–298.
Day, S.M., 1996. RMS response of a one-dimensional half-space to SH, *Bull. seism. Soc. Am.,* **86**(2), 363–370.
Day, S.M., Graves, R., Bielak, J., Dreger, D., Larsen, S., Olsen, K.B., Pitarka, A. & Ramirez-Guzman, L., 2008. Model for basin effects on long-period response spectra in Southern California, *Earthq. Spectra,* **24**(1), 257–277.
Eberhart-Phillips, D., Thurber, C. & Fletcher, J.B., 2014. Imaging *P* and *S* attenuation in the Sacramento-San Joaquin Delta region, Northern California, *Bull. seism. Soc. Am.,* **104**(5), 2322–2336.
Ely, G., Small, P., Jordan, T.H., Maechling, P.J. & Wang, F., 2010. A Vs30-derived Near-surface Seismic Velocity Model,*AGU Fall Meeting,* Abstract S51A–1907.
Fehler, M., Hoshiba, M., Sato, H. & Obara, K., 1992. Separation of scattering and intrinsic attenuation for the Kanto-Tokai region, Japan, using measurements of S-wave energy versus hypocentral distance, *Geophys. J. Int.,* **108**(3), 787–800.
Frankel, A. & Clayton, R.W., 1986. Finite difference simulations of seismic scattering: Implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity, *J. geophys. Res.,* **91**(B6), 6465–6489.
Gaffet, S. *et al.*, 2000. A site effect study in the Verchiano valley during the 1997 Umbria-Marche (Central Italy) earthquakes, *J. Seismol.,* **4**(4), 525–541.
Geli, L., Bard, P.-Y. & Jullien, B., 1988. The effect of topography on earthquake ground motion: a review and new results, *Bull. seism. Soc. Am.,* **78**(1), 42–63.
Graves, R. & Pitarka, A., 2016. Kinematic ground-motion simulations on rough faults including effects of 3D stochastic velocity perturbations, *Bull. seism. Soc. Am.,* **106**(5), 2136–2153.
Graves, R.W., 1996. Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences, *Bull. seism. Soc. Am.,* **86**(4), 1091–1106.
Graves, R.W. & Pitarka, A., 2010. Broadband ground-motion simulation using a hybrid approach, *Bull. seism. Soc. Am.,* **100,** 2095–2123.
Graves, R.W. & Wald, D.J., 2004. Observed and simulated ground motions in the San Bernardino basin region for the Hector Mine, California, earthquake, *Bull. seism. Soc. Am.,* **94**(1), 131–146.
Hartzell, S.H., Carver, D.L. & King, K.W., 1994. Initial investigation of site and topographic effects at Robinwood Ridge, California, *Bull. seism. Soc. Am.,* **84**(5), 1336–1349.
Hauksson, E. & Shearer, P.M., 2006. Attenuation models (Q_P and Q_S) in three dimensions of the southern California crust: inferred fluid saturation at seismogenic depths: Q_P and Q_S 3-D models of Southern California, *J. geophys. Res.,* **111**(B5), doi:10.1029/2005JB003947.
Hough, S.E. *et al.*, 1988. Attenuation near Anza, California, *Bull. seism. Soc. Am.,* **78**(2), 672–691.
Hu, Z., Olsen, K.B. & Day, S.M., 2022. Calibration of the Near-surface Seismic Structure in the SCEC Community Velocity Model Version 4, *Geophys. J. Int,* doi:10.1093/gji/ggac175.
Imperatori, W. & Gallovič, F., 2017. Validation of 3D velocity models using earthquakes with shallow slip: case study of the 2014 M_W 6.0, South Napa, California, Event, *Bull. seism. Soc. Am.,* **107,** 1019–1026.
Imperatori, W. & Mai, P., 2015. The role of topography and lateral velocity heterogeneities on near-source scattering and ground-motion variability, *Geophys. J. Int.,* **202**(3), 2163–2181.
Imperatori, W. & Mai, P.M., 2013. Broad-band near-field ground motion simulations in 3-dimensional scattering media, *Geophys. J. Int.,* **192**(2), 725–744.
Jones, L.M. *et al.*, 2008. The ShakeOut Scenario: Effects of a Potential M7.8 Earthquake on the San Andreas Fault in Southern California, U.S. Geological Survey Open-File Report 2008-1150.
Kohler, M.D., 2003. Mantle heterogeneities and the SCEC Reference Three-Dimensional Seismic Velocity Model Version 3, *Bull. seism. Soc. Am.,* **93**(2), 757–774.
Komatitsch, D. & Tromp, J., 2002. Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophys. J. Int.,* **149**(2), 390–412.
Lai, V.H., Graves, R.W., Yu, C., Zhan, Z. & Helmberger, D.V., 2020. Shallow basin structure and attenuation are key to predicting long shaking duration in Los Angeles basin, *J. geophys. Res.,* **125**(10).
Lee, E.-J., Chen, P., Jordan, T.H. & Wang, L., 2011. Rapid full-wave centroid moment tensor (CMT) inversion in a three-dimensional earth structure model for earthquakes in Southern California: rapid full-wave CMT inversion, *Geophys. J. Int.,* **186**(1), 311–330.
Lee, S.-J., Chan, Y.-C., Komatitsch, D., Huang, B.-S. & Tromp, J., 2009. Effects of realistic surface topography on seismic ground motion in the Yangminshan Region of Taiwan based upon the spectral-element method and LiDAR DTM, *Bull. seism. Soc. Am.,* **99,** 681–693.
Leonard, M., 2010. Earthquake fault scaling: self-consistent relating of rupture length, width, average displacement, and moment release, *Bull. seism. Soc. Am.,* **100,** 1971–1988.
Lin, Y.-P. & Jordan, T.H., 2018. Frequency-dependent attenuation of *P* and *S* waves in Southern California, *J. geophys. Res.,* **123**(7), 5814–5830.
Liu, H.-P., Anderson, D.L. & Kanamori, H., 1976. Velocity dispersion due to anelasticity; implications for seismology and mantle composition, *Geophys. J. Int.,* **47**(1), 41–58.
Liu, Z., Shang, C., Huang, L., Liang, J. & Li, J., 2020. Scattering of seismic waves by three-dimensional large-scale hill topography simulated by a fast parallel IBEM, *Earthq. Eng. Eng. Vib.,* **19**(4), 855–873.
Lovati, S., Bakavoli, M., Massa, M., Ferretti, G., Pacor, F., Paolucci, R., Haghshenas, E. & Kamalian, M., 2011. Estimation of topographical effects at Narni ridge (Central Italy): comparisons between experimental results and numerical modelling, *Bull. Earthq. Eng.,* **9**(6), 1987–2005.
Ma, S., Archuleta, R.J. & Page, M.T., 2007. Effects of large-scale surface topography on ground motions, as demonstrated by a study of the San Gabriel Mountains, Los Angeles, California, *Bull. seism. Soc. Am.,* **97**(6), 2066–2079.
Magistrale, H., McLaughlin, K. & Day, S., 1996. A geology-based 3D velocity model of the Los Angeles basin sediments, *Bull. seism. Soc. Am.,* **86**(4), 1161–1166.
Magistrale, H., Day, S., Clayton, R.W. & Graves, R., 2000. The SCEC Southern California Reference Three-Dimensional Seismic Velocity Model Version 2, *Bull. seism. Soc. Am.,* **90,** S65–S76.
Maufroy, E., Cruz-Atienza, V.M., Cotton, F. & Gaffet, S., 2015. Frequency-scaled curvature as a proxy for topographic site-effect amplification and ground-motion variability, *Bull. seism. Soc. Am.,* **105**(1), 354–367.

Nakata, N. & Beroza, G.C., 2015. Stochastic characterization of mesoscale seismic velocity heterogeneity in Long Beach, California, *Geophys. Suppl. Mon. Not. R. astr. Soc.,* **203**(3), 2049–2054.

Nie, S., Wang, Y., Olsen, K.B. & Day, S.M., 2017. Fourth-order staggered-grid finite-difference seismic wavefield estimation using a discontinuous mesh interface (WEDMI), *Bull. seism. Soc. Am.,* **107**(5), 2183–2193.

Olsen, K. & Takedatsu, R., 2015. The SDSU broadband ground-motion generation module BBtoolbox version 1.5, *Seismol. Res. Lett.,* **86**(1), 81–88.

Olsen, K. *et al.*, 2009. ShakeOut-D: ground motion estimates using an ensemble of large earthquakes on the southern San Andreas fault with spontaneous rupture propagation, *Geophys. Res. Lett.,* **36**(4), doi:10.1029/2008GL036832.

Olsen, K.B. & Mayhew, J.E., 2010. Goodness-of-fit criteria for broadband synthetic seismograms, with application to the 2008 M_w 5.4 Chino Hills, California, earthquake, *Seismol. Res. Lett.,* **81**(5), 715–723.

Olsen, K.B., Day, S.M. & Bradley, C.R., 2003. Estimation of Q for long-period (>2 sec) waves in the Los Angeles basin, *Bull. seism. Soc. Am.,* **93**(2), 627–638.

Olsen, K.B. *et al.*, 2006. Strong shaking in Los Angeles expected from southern San Andreas earthquake, *Geophys. Res. Lett.,* **33**(7), L07305.

Olsen, K.B., Begnaud, M., Phillips, S. & Jacobsen, B.H., 2018. Constraints of crustal heterogeneity and $Q(f)$ from regional (<4 Hz) wave propagation for the 2009 North Korea nuclear test, *Bull. seism. Soc. Am.,* **108,** 1369–1383.

O'Reilly, O., Yeh, T., Olsen, K.B., Hu, Z., Breuer, A., Roten, D. & Goulet, C.A., 2022. A high-order finite-difference method on staggered curvilinear grids for seismic wave propagation applications with topography, *Bull. seism. Soc. Am.,* **112**(1), 3–22.

Pischiutta, M., Cultrera, G., Caserta, A., Luzi, L. & Rovelli, A., 2010. Topographic effects on the hill of Nocera Umbra, central Italy: topographic effects on Nocera Umbra hill, *Geophys. J. Int.,* **182**(2), 977–987.

Pitarka, A. & Ichinose, G., 2009. Simulating forward and backward scattering in viscoelastic 3D media with random velocity variations and basin structure, US Geol. Surv. Tech. Rep.

Przybilla, J., Wegler, U. & Korn, M., 2009. Estimation of crustal scattering parameters with elastic radiative transfer theory, *Geophys. J. Int.,* **178**(2), 1105–1111.

Rai, M., Rodriguez-Marek, A. & Chiou, B.S., 2017. Empirical terrain-based topographic modification factors for use in ground motion prediction, *Earthq. Spectra,* **33**(1), 157–177.

Raoof, M., Herrmann, R. & Malagnini, L., 1999. Attenuation and excitation of three-component ground motion in southern California, *Bull. seism. Soc. Am.,* **89**(4), 888–902.

Rodgers, A.J., Pitarka, A., Petersson, N.A., Sjögreen, B. & McCallen, D.B., 2018. Broadband (0 Hz) ground motions for a magnitude 7.0 Hayward Fault earthquake with three-dimensional structure and topography: an M 7 Hayward Fault earthquake, *Geophys. Res. Lett.,* **45**(2), 739–747.

Roten, D., Olsen, K.B. & Pechmann, J.C., 2012. 3D simulations of M 7 earthquakes on the Wasatch Fault, Utah, Part II: broadband (0 Hz) ground motions and nonlinear soil behavior, *Bull. seism. Soc. Am.,* **102**(5), 2008–2030.

Sánchez-Sesma, F.J. & Campillo, M., 1991. Diffraction of *P*, *SV*, and Rayleigh waves by topographic features: a boundary integral formulation, *Bull. seism. Soc. Am.,* **81**(6), 2234–2253.

Savran, W. & Olsen, K., 2016. Model for small-scale crustal heterogeneity in Los Angeles basin based on inversion of sonic log data, *Geophys. J. Int.,* **205**(2), 856–863.

Savran, W.H. & Olsen, K.B., 2019. Ground motion simulation and validation of the 2008 Chino Hills earthquake in scattering media, *Geophys. J. Int.,* **219**(3), 1836–1850.

Savran, W.H. & Olsen, K.B., 2020. Kinematic rupture generator based on 3-D spontaneous rupture simulations along geometrically rough faults, *J. geophys. Res.,* **125**(10), e2020JB019464.

Small, P. *et al.*, 2017. The SCEC unified community velocity model software framework, *Seismol. Res. Lett.,* **88**(6), 1539–1552.

Spudich, P., Hellweg, M. & Lee, W. H.K., 1996. Directional topographic site response at Tarzana observed in aftershocks of the 1994 Northridge, California, earthquake: implications for mainshock motions, *Bull. seism. Soc. Am.,* **86,** S193–S208.

Taborda, R. & Bielak, J., 2014. Ground-motion simulation and validation of the 2008 Chino Hills, California, earthquake using different velocity models, *Bull. seism. Soc. Am.,* **104**(4), 1876–1898.

Taborda, R., Azizzadeh-Roodpish, S., Khoshnevis, N. & Cheng, K., 2016. Evaluation of the southern California seismic velocity models through simulation of recorded events, *Geophys. J. Int.,* **205**(3), 1342–1364.

Takemura, S., Furumura, T. & Maeda, T., 2015. Scattering of high-frequency seismic waves caused by irregular surface topography and small-scale velocity inhomogeneity, *Geophys. J. Int.,* **201**(1), 459–474.

Thompson, E.M., Baise, L.G., Tanaka, Y. & Kayen, R.E., 2012. A taxonomy of site response complexity, *Soil Dyn. Earthq. Eng.,* **41,** 32–43.

Trifunac, M.D. & Hudson, D.E., 1971. Analysis of the Pacoima dam accelerogram—San Fernando, California, earthquake of 1971, *Bull. seism. Soc. Am.,* **61**(5), 1393–1411.

Umeda, Y., Kuroiso, A., Ito, K. & Muramatu, I., 1987. High accelerations produced by the Western Nagano Prefecture, Japan, earthquake of 1984, *Tectonophysics,* **141**(4), 335–343.

USGS, 2014. Earthquake Events, Focal Mechanism.

USGS, 2020. *National Seismic Hazard Maps - Elevation Products (3DEP)*.

Wang, W. & Shearer, P.M., 2017. Using direct and coda wave envelopes to resolve the scattering and intrinsic attenuation structure of Southern California: resolve the attenuation structure, *J. geophys. Res.,* **122**(9), 7236–7251.

Withers, K.B., Olsen, K.B. & Day, S.M., 2015. Memory-efficient simulation of frequency-dependent q, *Bull. seism. Soc. Am.,* **105**(6), 3129–3142.

Withers, K.B., Olsen, K.B., Day, S.M. & Shi, Z., 2019. Ground motion and intraevent variability from 3D deterministic broadband (0–7.5 Hz) simulations along a nonplanar strike-slip fault, *Bull. seism. Soc. Am.,* **109**(1), 229–250.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. Shear wave quality factor (Q_S) plotted against V_S (m s^{-1}) for several attenuation models widely used in the literature (e.g. Olsen *et al.* 2003; Taborda & Bielak 2014; Savran & Olsen 2019; Withers *et al.* 2019) and investigated here. The inset figure in the upper left corner zooms into $V_S \leq 1600$ m s^{-1}, denoted by the dashed black box. Note that these Q_S relations are valid for constant Q models, or frequency-dependent Q models for frequencies below 1 Hz.

Figure S2. Description of three candidate source models used in this study. Top: slip distribution (shading) for sources 1, 2 and 3 (a–c), characterized by their hypocentral depths of 5, 5.5 and 6 km, respectively. Contours depict rupture times at 0.4 s interval starting from 0. (d) Sum of the moment rates for all subfaults and (e) Fourier amplitude spectra. Source 1 is the default source model used elsewhere in this paper.

Figure S3. PGVs for sources 1, 2 and 3 (from left to right; see Supporting Information Fig. S2). The top and bottom rows represent the bandpass filtered results for 0.15–2.5 and 2.5–5 Hz, respectively. The star denotes the epicentre of the La Habra event.

Figure S4. Effect of SSHs on PGVs illustrated by probability density histograms of the PGV difference between models with (Models 10–13) and without (Models 2 and 6) SSHs. The definition of per cent difference (x-axis) is the same as in Fig. 11.
Figure S5. Bias of (top row) PGV and (middle row) DUR and (bottom row) GOF for bandwidths of (left-hand column) 0.15–2.5 Hz and (right-hand column) 2.5–5 Hz at all 259 stations for Model 1 (see Table 2 for model features). The bias is calculated in the same way as in Fig. 9. The solid line depicts the moving average of the bias of PGV using a 20-point window versus hypocentral distance. The shading denotes the standard deviation centred at the mean.
Figure S6. Same as Fig. S5, but for Model 2.
Figure S7. Same as Fig. S5, but for Model 3.
Figure S8. Same as Fig. S5, but for Model 5.
Figure S9. Same as Fig. S5, but for Model 6.
Figure S10. Same as Fig. S5, but for Model 7.
Figure S11. Same as Fig. S5, but for Model 8.
Figure S12. Same as Fig. S5, but for Model 9.
Figure S13. Same as Fig. S5, but for Model 10.
Figure S14. Same as Fig. S5, but for Model 11.
Figure S15. Same as Fig. S5, but for Model 12.
Figure S16. Same as Fig. S5, but for Model 13.
Figure S17. Same as Fig. S5, but for Model 14.
Figure S18. Same as Fig. S5, but for Model 15.
Figure S19. Same as Fig. S5, but for Model 16.
Figure S20. Same as Fig. S5, but for Model 17.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX A: 3-D SIMULATION WITH MINIMUM V_S OF 200 m s^{-1}

In this study, we constrained our 3-D simulations to using $V_S > 500$ m s^{-1} due to limitations in computational resources and added 1-D corrections for the effects of the material with $V_S \leq 500$ m s^{-1}. In order to assess the effect of these 1-D corrections, we used a 3-D simulation of the La Habra earthquake with CVM-S and a minimum V_S of 200 m s^{-1} obtained from a discontinuous mesh (DM) version of AWP (Nie *et al.* 2017). The simulated domain (Table 1) is discretized into three partitions: (1) dx = 8 m from the surface to 1472 m, (2) dx = 24 m between 1472 m and 10 336 m and (3) dx = 72 m at deeper levels. The improved efficiency of the DM approach allows us to lower the minimum V_S to 200 m s^{-1} and retain 5 points per minimum S-wavelength. The reason for not using the more efficient DM code for other simulations in this study is that the code is currently limited to a flat free surface condition.

Fig. A1 shows the FAS bias of two 3-D simulations for Model 17 (see Table 2) with minimum V_S of 200 and 500 m s^{-1}, both using a flat free surface boundary condition. Since the lower velocity material has very little effect on the vertical component, we apply the correction to the horizontal components only. The SH1D correction (with the minimum V_S clamped at 200 m s^{-1}) is applied to the horizontal components of the 3-D simulation with minimum V_S of 500 m s^{-1}. For the horizontal components, the corrected simulation matches the minimum V_S of 200 m s^{-1} simulation fairly well, with less than 10 per cent difference in terms of the median bias below about 2.5 Hz. The moderate overprediction for higher frequencies up to 5 Hz (<about 25 per cent) is likely caused by vertical resonance effects primarily in the 1-D model.

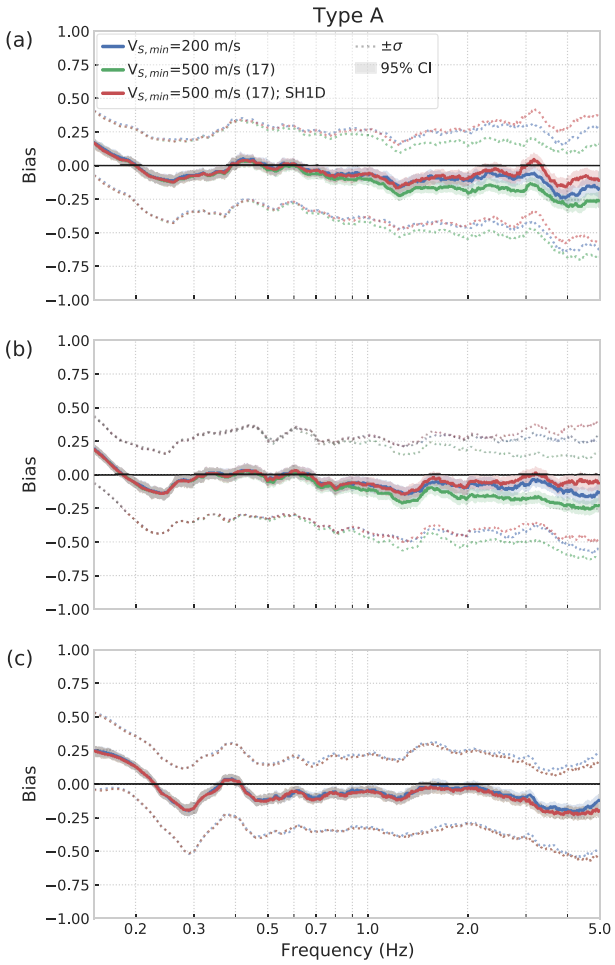

Figure A1. FAS bias between data and synthetics with minimum V_S clamped at 200 m s^{-1} (blue), 500 m s^{-1} (green) and 500 m s^{-1} with 1-D correction (red) for (a) E–W, (b) N–S and (c) vertical component (the 1-D correction is only applied to the horizontal components, and thus the red and green curves coincide in the vertical component). A positive (negative) bias depicts overprediction (underprediction). The solid lines show the median FAS bias over all 259 stations, shading depicts the 95 per cent confidence interval (CI) and the dashed lines denote one standard deviation centred at the median.