Geophysical Journal International

Geophys. J. Int. (2022) **230,** 2183–2198
Advance Access publication 2022 May 10
GJI Seismology

https://doi.org/10.1093/gji/ggac175

Calibration of the near-surface seismic structure in the SCEC community velocity model version 4

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

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

Accepted 2022 May 5. Received 2021 Octobor 6; in original form 2021 April 8

SUMMARY

The near-surface seismic structure (to a depth of about 1000 m), particularly the shear wave velocity (V_S), can strongly affect the propagation of seismic waves and, therefore, must be accurately calibrated for ground motion simulations and seismic hazard assessment. The V_S of the top (<300 m) crust is often well characterized from borehole studies, geotechnical measurements, and water and oil wells, while the velocities of the material deeper than about 1000 m are typically determined by tomography studies. However, in depth ranges lacking information on shallow lithological stratification, typically rock sites outside the sedimentary basins, the material parameters between these two regions are typically poorly characterized due to resolution limits of seismic tomography. When the alluded geological constraints are not available, models, such as the Southern California Earthquake Center (SCEC) Community Velocity Models (CVMs), default to regional tomographic estimates that do not resolve the uppermost V_S values, and therefore deliver unrealistically high shallow V_S estimates. The SCEC Unified Community Velocity Model (UCVM) software includes a method to incorporate the near-surface earth structure by applying a generic overlay based on measurements of time-averaged V_S in top 30 m (V_{S30}) to taper the upper part of the model to merge with tomography at a depth of 350 m, which can be applied to any of the velocity models accessible through UCVM. However, our 3-D simulations of the 2014 M_w 5.1 La Habra earthquake in the Los Angeles area using the CVM-S4.26.M01 model significantly underpredict low-frequency (<1 Hz) ground motions at sites where the material properties in the top 350 m are significantly modified by the generic overlay ('taper'). On the other hand, extending the V_{S30}-based taper of the shallow velocities down to a depth of about 1000 m improves the fit between our synthetics and seismic data at those sites, without compromising the fit at well-constrained sites. We explore various tapering depths, demonstrating increasing amplification as the tapering depth increases, and the model with 1000 m tapering depth yields overall favourable results. Effects of varying anelastic attenuation are small compared to effects of velocity tapering and do not significantly bias the estimated tapering depth. Although a uniform tapering depth is adopted in the models, we observe some spatial variabilities that may further improve our method.

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

1 INTRODUCTION

Ground motion amplification due to the near-surface structure is widely accepted and well studied (e.g. Gilbert *et al.* 1907; Field 2000), and needs to be incorporated in numerical simulations of earthquakes to produce accurate ground motion results. Theoretical analyses have shown that the near-surface shear wave velocity (V_S) can exert strong control on spectral amplification (Joyner *et al.* 1981; Boore & Joyner 1991; Anderson *et al.* 1996; Day 1996). The time-averaged shear wave velocity in the upper 30 m (V_{S30}) is routinely used as a representation of the site condition in ground motion prediction models and building codes (Borcherdt 1994; Bozorgnia *et al.* 2014; International Code Council 2014). Several methods have been proposed for estimating V_{S30} from topography (Wald & Allen 2007), supplemented with near-surface geological information (Thompson *et al.* 2014; Wills *et al.* 2015). However, despite the continuing advancement in the V_{S30}-based methodologies by the seismic hazard community (e.g. Thompson *et al.* 2014; Heath *et al.* 2020), estimating V_{S30} at high resolution remains a difficult task and it is noted that V_{S30} is not a good single proxy for the estimation of site amplification (e.g. Steidl 2000; Lee & Trifunac

2010; Shingaki *et al.* 2018). Other empirical methodologies provide additional predictive capability for shallow low-velocity amplification, normally constrained by sediment depth, which is parametrized using the depth to the 1 km s^{-1} (z_1) or 2.5 km s^{-1} ($z_{2.5}$) V_S horizon (e.g. Abrahamson *et al.* 2014; Boore *et al.* 2014; Campbell & Bozorgnia 2014). Nonetheless, these empirical methods, oftentimes dependent on V_{S30}, have similar limitations with V_{S30} that depth-dependent and lateral velocity variations are insufficiently accounted for.

While the current approximations to correct for site effects represent great progress in ground motion estimation, a fully physics-based approach to computing ground motion offers opportunities for further improvements. The physics-based approach entails difficult challenges as well, and remains a long-term goal. In such an approach, the full wavefield is computed deterministically, to maximum frequencies that are sometimes up to 5 Hz or higher, using a 3-D velocity model that includes observationally constrained heterogeneities (e.g. Savran & Olsen 2019; Withers *et al.* 2019; Hu *et al.* 2022). A necessary ingredient in producing accurate synthetic seismograms using physics-based simulations is an accurate velocity model for the model region. Community Velocity Models (CVMs) have been developed for such purpose, for example the Southern California Earthquake Center (SCEC) CVMs (Small *et al.* 2017), the Cascadia CVM (Stephenson *et al.* 2017) and the Subsurface Structure Model maintained by the Japan Seismic Hazard Information Station (Fujiwara *et al.* 2017). These velocity models are often generated by combining 3-D tomographic inversion from seismic waves (Tape *et al.* 2009, 2010; Lee *et al.* 2014) with shallow geotechnical information (e.g. V_{S30}). The spatial resolution of large-scale tomographic studies is generally limited by the density of ray paths, distribution of high-quality measurements, or intrinsic non-uniqueness of inversion, particularly in the upper ∼1000 m of the crust. For example, Lin *et al.* (2007) had a vertical grid spacing of 3 km and only resolved velocity contrasts over intervals larger than 1km in their tomographic inversion using *P*- and *S*-wave arrival times. In addition, the 3-D seismic waveform tomography conducted by Lee *et al.* (2014) reached at best 1 km resolution in the centre of the inverted region, and Qiu *et al.* (2019) found that the top 3 km was poorly constrained in their Eikonal tomography using ambient noise cross-correlations.

Effects of shallow velocity structure, for example from *S*-wave impedance and scattering, play a significant role in ground motion amplification and duration (e.g. Graves 1995; Anderson *et al.* 1996; Imperatori & Mai 2013). Specifically, the 1-D theoretical analysis by Day (1996) suggests that the smoothed amplification spectrum is principally determined by shallow V_S, above roughly the depth of half the smoothing bandwidth expressed as a wavelength. Over a ≈ 0.5 Hz bandwidth and typical Southern California rock site V_S values, the analysis predicts that the V_S structure above about 1000 m will have a disproportionately strong effect on ground motion. Therefore, resolving the shallow velocity structure is essential in accurate predictions of ground motions. In several SCEC CVMs, velocities and densities in the top 300 m within the basins are constrained by geotechnical and geophysical data, such as seismic reflection surveys, borehole seismic records and gravity data, and in the deeper basins are estimated either from empirical age- and depth-consolidation rules based on water and oil wells and geological studies, or sonic logs and reflection/refraction profiles from the oil industry (Magistrale *et al.* 1996, 2000; Süss & Shaw 2003). Unfortunately, outside and below the basins (typically rock sites), CVMs simply assign interpolated results from regional tomography studies. Additional data constraints on shallow velocity structure, including seismic refraction studies (e.g. Teague *et al.* 2018) or borehole logs (e.g. Steller 1996; Thompson *et al.* 2012) are rare in these regions.

Where location-specific constraints are lacking, previous studies have attempted to use generic models to bridge the gap between data constraints at shallow ($\lesssim$30 m) and deeper ($\gtrsim$1000 m) depths. For example, Boore & Joyner (1997) generated a continuous depth-dependent V_S function based on 3 different intervals. The V_S profile in the upper 30 m was constructed from interpolated shallow average arrival times. At depths below 4 km, V_S was estimated from the *P*-wave velocity (V_P), measured from earthquake location studies and velocity surveys, on the assumption of a fixed Poisson ratio at 0.25. Finally, the shallow and deeper V_S were connected using two power-law functions. Ely *et al.* (2010) proposed a generalized method that derives the surface V_S by linearly scaling V_{S30} and then interpolates velocities with depth until converging to the original tomography model at a certain depth, a scheme which has been implemented in the SCEC Unified Community Velocity Model (UCVM) software and can be applied to CVMs that are queried using UCVM. We will use the term 'tapering' to denote the replacement of (poorly constrained) site-specific CVM values by a generic function of depth that merges smoothly with the original CVM at some depth z_T. Ely *et al.* (2010) proposed a value of 350 m for z_T, based on qualitative comparison between synthetic and seismic records from the 2008 M_w 5.4 Chino Hills, CA, earthquake.

In this study, we quantify the accuracy of ground motion simulations based on comparisons to the 2014 M_w 5.1 La Habra, CA, earthquake, and interpret the results in terms of the representation of crustal V_S in the top 1000 m. The paper is organized as follows: we first briefly introduce our numerical approach to obtain the simulated ground motions, present an approximate 1-D analysis of site amplification to evaluate the potential to improve site amplification at poorly constrained sites, and finally evaluate different generic tapers using 3-D wave propagation simulations. The proposed tapering method amplifies the ground motions as the tapering depth increase, which generates up to 3 times (less than 10 per cent) increase in the spectral amplitudes at poorly (well) constrained locations, compared to the original (i.e. untapered) CVM. We also discuss the limitations of this study, in particular the neglect of spatial variation of the taper depth, which will be a future objective to investigate, using more validation metrics, additional seismic events, and higher frequencies.

2 NUMERICAL APPROACH

We perform 0–1 Hz wave propagation simulations of the 2014 M_w 5.1 La Habra earthquake to explore the accuracy of ground motion predictions in terms of the shallow velocity structure. The simulations use the SCEC velocity model CVM-S4.26-M01 (hereafter referred to as CVM-S), which originates from the model CVM-S4 developed by Magistrale *et al.* (1996) and the CVM-S4.26 tomographic inversions

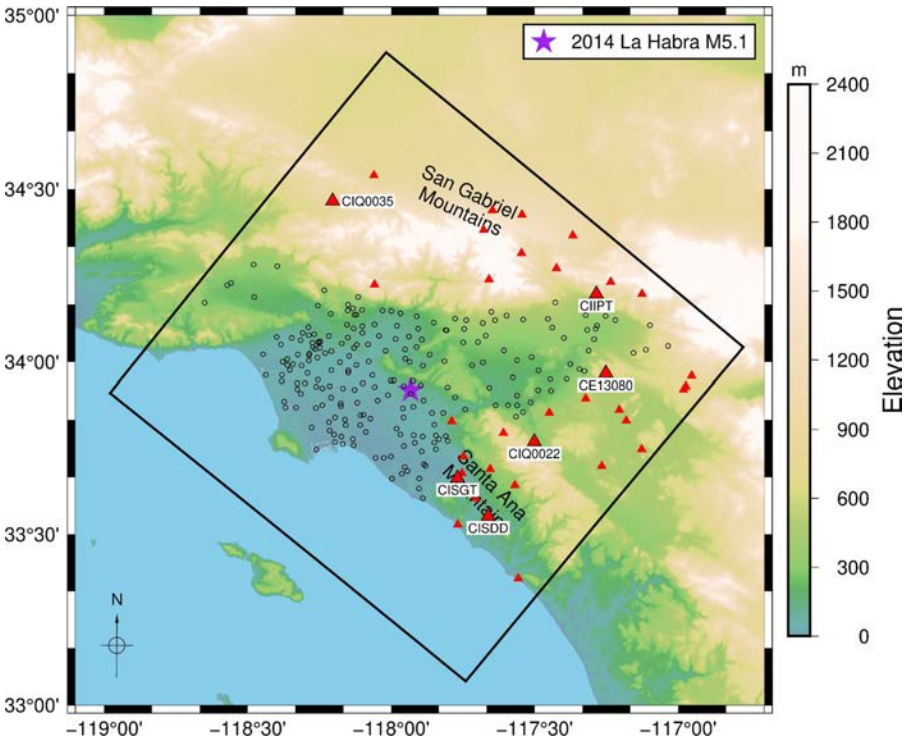

Figure 1. Simulation region for the La Habra event and locations of 259 strong ground motion stations (circles represent type A sites with surface V_S < 1000 m s^{-1} and red triangles represent type B sites with surface $V_S \geq$ 1000 m s^{-1}). The dimensions and coordinates of the simulated domain (black rectangle) are listed in Table 1. The named sites (triangles with black edge) are used for further comparison in Figs 2 and 11. The star depicts the epicentre of the La Habra earthquake.

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

Domain	
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	
Minimum V_S	500 m s^{-1}
Grid discretization	20/60 m
Number of cells	25.1 billion
Number of CPU processors	480
Number of GPU processors	960
Wall-clock time	1.5 hr
Temporal resolution	
Time discretization	0.001 s
Simulation time	120 s
Number of time steps	120 000

by Chen *et al.* (2007) and Lee *et al.* (2014). Fig. 1 shows the computational domain and strong motion seismic stations in the greater Los Angeles area used in this study. We discretize a 148 km × 140 km × 58 km region from CVM-S and the computational domain is rotated 39.9° clockwise to reduce the mesh size while optimizing the data coverage in our region of interest. Table 1 lists the simulation parameters.

The GPU-supported staggered-grid finite-difference code AWP-ODC (Anelastic Wave Propagation–Olsen, Day and Cui, from the authors of the code; Cui *et al.* 2010) with discontinuous mesh (Nie *et al.* 2017) was used for the simulations analysed in this study. We used

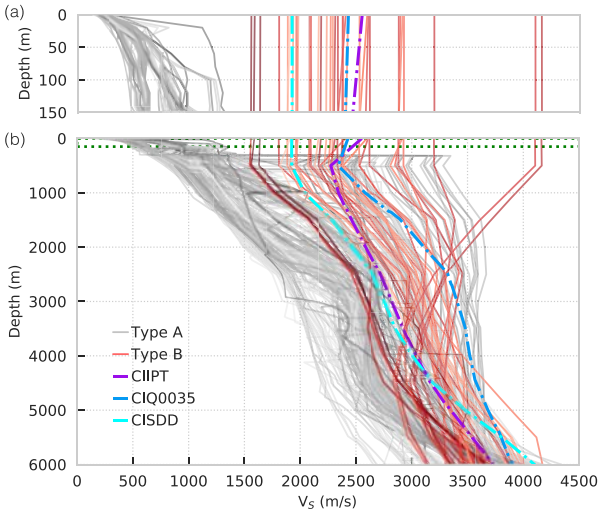

Figure 2. (a) Top 150 m and (b) 0–6000m V_S profiles at the 259 stations. The black and red curves represent type A (surface $V_S < 1000$ m s^{-1}) and type B (surface $V_S \geq 1000$ m s^{-1}) sites, respectively. Vertical profiles for three named sites (see Fig. 1) are identified using dash–dotted curves. The curves are colour coded with darker colours denoting sites with farther distance from the source.

spatial grid spacings of 20 and 60 m in the grid partitions above and below 7.5 km, respectively, and a minimum V_S of 500 m s^{-1}. To facilitate the use of these simulations in a companion, high-frequency study (Hu *et al.* 2022, this issue), we computed frequencies up to 5 Hz. However, we restrict our analysis to a maximum frequency of 1 Hz in this study, which precludes some of the complicating effects that may become important at higher frequencies, for example topography, frequency-dependent attenuation, etc. Anelastic attenuation is incorporated with the quality factors given by the linear velocity-dependent relation $Q_S = 0.1V_S$ (V_S in m s^{-1}) and $Q_P = 2Q_S$, as suggested by previous ground motion validation studies (e.g. Bielak *et al.* 2010; Withers *et al.* 2019). We applied sponge zones (Cerjan *et al.* 1985) with a width of 64 nodes at the exterior grid boundaries (except at the flat free surface) to limit artificial reflections.

The 2014 M_w 5.1 La Habra earthquake was well recorded by broad-band strong motion sensors. We selected 259 stations with epicentral distances up to 90 km and signal-to-noise ratios larger than 3 dB for our study. The assessment of the ground motion synthetics is made using the Fourier amplitude spectra (FAS) of accelerations at all 259 stations, and the goodness of fit to data is described by the FAS bias between model and data:

$$\text{Bias(frequency, site)} = \log_{10}\left(\frac{\text{FAS}_{\text{model}}}{\text{FAS}_{\text{data}}}\right). \quad (1)$$

We used a kinematic source description 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, fault area of 2.5 km × 2.5 km, and a hypocentral depth of 5 km (0.5 km below the top of the finite fault). The source was selected from comparison between synthetics and records 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 [R. Graves, personal communication, 2020; see Fig. 3 in Hu *et al.* (2022), this issue] using a series of 40 realizations with different random seeds, simulated in a smaller domain. The rupture duration is less than 2 s, and the source model was sampled at an interval of 0.001 s, identical to the time step used in our simulations.

The V_S profiles extracted from CVM-S beneath all 259 stations selected for the La Habra event are shown in Fig. 2. For most stations, the unmodified CVM-S gives low surface V_S (<500 m s^{-1}, see Fig. 2a), while a small portion (15 per cent) of the stations have significantly larger surface V_S (>1500 m s^{-1}, up to 4650 m s^{-1}). Such large V_S, typically representative of much larger depths, are highly unrealistic at the surface in western North America, even for rock sites in the presence of weathering (note that shallow velocities can be much higher in mid-continent and eastern North America where the surface weathered soils are stripped off by glacial erosion). Additionally, the fact that the V_S values remain constant between the surface and about 500 m depth (Fig. 2) indicates a poorly constrained near-surface V_S at these stations. We separate all sites into two classes: type A sites where CVM-S provides meaningful near-surface velocities based on geological and geophysical constraints, and type B sites where shallow velocities are typically higher than realistic near-surface velocities due to their derivation from relatively low-resolution seismic tomography. The two types of stations fall into two distinct CVM-S surface V_S classes, ≈200–300 and ≈1500–4650 m s^{-1}, respectively (Fig. 2). Type B represents sites with poor constraints in velocities and thus constitute our main target for calibration in this study (see Table 2 for type B site information from the original, untapered CVM-S). There are many indications that the near-surface V_S values at type B sites in CVM-S are anomalously high. CVM-S4 includes a geotechnical layer (GTL) from geological and geophysical data, which are typically confined to the top 300 m in basin areas only (Magistrale *et al.* 1996, 2000). CVM-S then incorporates the material properties from CVM-S4 and cuts off the GTL at 350 m depth when it is merged with the background tomographic inversion results from CVM-S4.26 (Lee *et al.* 2014), leading to a sharp contrast at that depth when the background model has

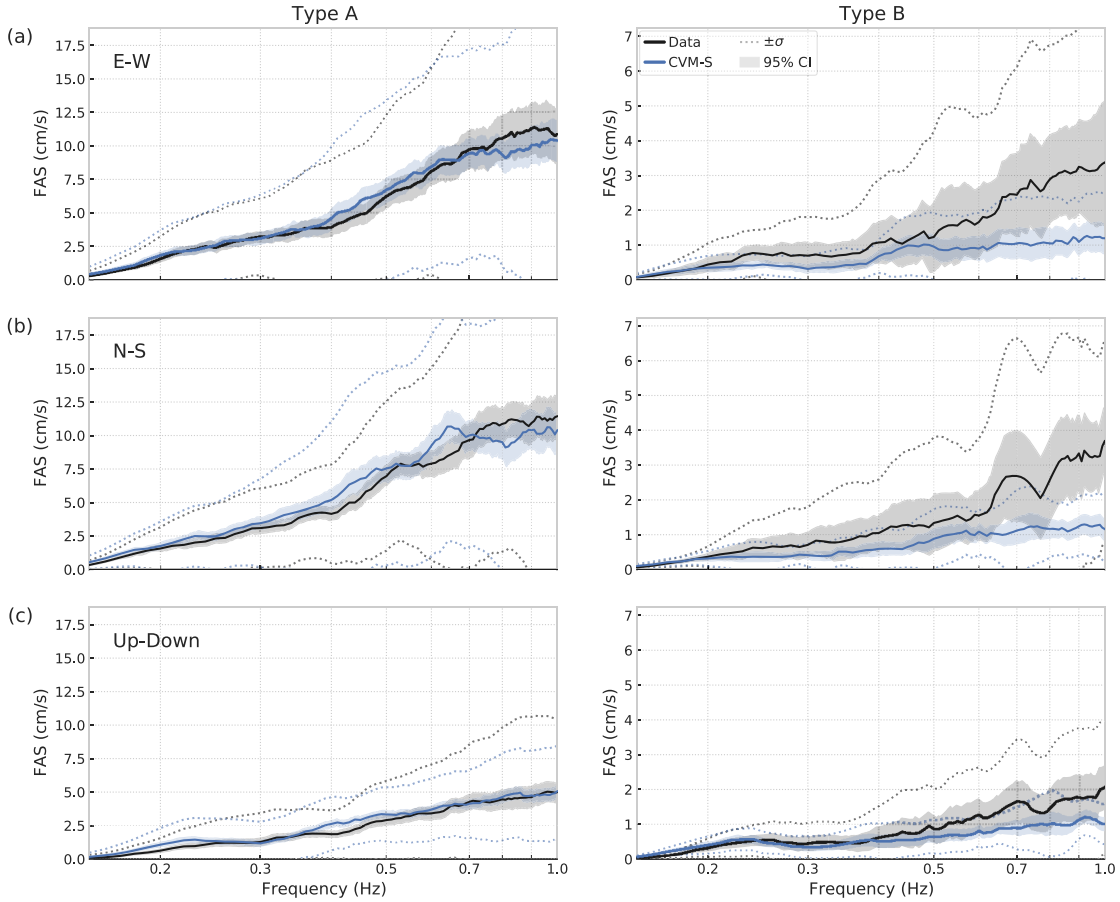

Figure 3. FAS derived from the records (black) and CVM-S (blue) for the (a) east–west component, (b) north–south component and (c) vertical component. The left- and right-hand columns represent type A and B sites, respectively. The solid line is the median FAS over the site group, the narrow band is the 95 per cent confidence interval of the median, and the dashed lines depict the standard deviation centred at the median.

high velocity (as at type B sites). Note that the mean values of each of the two groups of profiles become similar below a depth of 2000 m. The hypocentral-distance distributions for the respective site types (A and B) are similar, though type B sites are typically located outside the basins and thus of relatively larger distance.

3 SIMULATION RESULTS

We calculate the Fourier amplitude spectra (FAS) of accelerations for both recorded and synthetic time-series, both processed in the following way: (1) low-pass filtering with a corner frequency of 10 Hz using a fourth-order zero-phase Butterworth filter; (2) interpolating linearly to a uniform time step; (3) tapering at the last 2 s using the positive half of a Hanning window; and (4) padding with 5 s of zeros. Horizontal components for both data and synthetics were rotated to east–west (E–W) and north–south (N–S) directions, and the data were synchronized to the rupture time. Furthermore, if needed, records were padded with zeros to obtain a duration of 120 s relative to the rupture time. Finally, we calculated the FAS of the accelerations from the time derivatives of the velocities for the synthetics and records, which were bandpass filtered between 0.15 Hz and 1 Hz using a fourth-order zero-phase Butterworth filter. The lower cut-off frequency of 0.15 Hz was selected to avoid noise interference.

Fig. 3 shows a comparison of median FAS, taken over the two types of stations, of ground accelerations for synthetics from unmodified CVM-S and recordings. The FAS at type A sites are well predicted, especially below 0.7 Hz, with a small underprediction between 0.7 and 1 Hz on the horizontal components. At type B sites, however, significant underprediction is observed for all three components at frequencies as low as 0.2 Hz on the horizontal components. As the frequency increases toward 1 Hz, the difference in FAS between data and synthetics increases rapidly, leaving the FAS from the simulated results outside of the 95 per cent confidence interval of the data. The relatively good match at type A sites indicates that the source description is not likely to be a significant source of the misfit. Furthermore, more complicated path and site effects from topography, small-scale heterogeneities and frequency-dependent attenuation are expected to be negligible at frequencies below 1 Hz. For example, topographic relief mostly affects a frequency band that scales inversely with the characteristic dimensions of the relief (e.g. Boore 1972; Bouchon & Barker 1996; Durand *et al.* 1999), and that band is generally observed to be above 2 Hz (e.g. Pischiutta *et al.* 2010; Massa *et al.* 2014). Frequency dependence of anelastic Q (e.g. Liu *et al.* 1976; Fehler *et al.* 1992) and small-scale velocity

Table 2. Type B site information. The V_{S30} values are linearly interpolated laterally to the finite-difference grid points using the V_{S30} map from Thompson (2018). We extract the elevation data from the U.S. Geological Survey (USGS 2020).

Site name	Lon (°)	Lat (°)	R_{hypo} (km)	Surface V_S (m s^{-1})	V_{S30} (m s^{-1})	Elevation (m)
CISRN	−117.7894	33.8285	17.53	1908.44	351.90	212.32
CIQ0029	−117.7480	33.7270	27.87	2163.35	293.50	94.84
CE13220	−117.7527	33.6782	31.99	2090.55	351.90	70.28
CISTG	−117.7686	33.6640	32.61	1980.29	351.90	47.53
CE13441	−117.7752	33.6598	32.74	1934.54	447.28	45.87
CIPLS	−117.6091	33.7953	33.36	2234.93	351.90	1215.81
CE24399	−118.0582	34.2236	36.28	2597.01	710.10	1724.74
CIMWC	−118.0583	34.2236	36.28	2596.96	710.10	1727.73
CIQ0034	−117.6550	33.6900	36.56	2289.27	293.50	324.50
CIQ0009	−117.7050	33.6050	41.12	1885.39	351.90	106.23
CIQ0022	−117.5010	33.7700	43.52	2425.56	351.90	362.41
CIBFS	−117.6585	34.2388	44.01	2270.57	710.10	1301.77
NP707	−117.4490	33.8540	45.57	2913.91	293.50	407.88
CIQ0026	−117.5690	33.6430	45.83	2529.13	228.20	375.70
CIQ0005	−117.7680	33.5280	46.37	1961.18	710.10	42.60
CISDD	−117.6617	33.5526	48.13	1923.53	351.90	122.19
CIQ0038	−117.4280	33.7310	51.38	2926.22	293.50	416.98
CE13916	−117.3219	33.8946	56.71	2893.67	518.90	522.59
CITA2	−117.6782	34.3820	56.88	2381.10	351.90	2258.42
CILPC	−117.5464	34.3148	56.89	1970.65	351.90	1344.56
CICJM	−117.4245	34.2712	61.30	2404.97	228.20	1615.85
CE13080	−117.2519	33.9677	63.28	2607.63	518.90	542.10
CE23958	−117.6466	34.4393	63.84	2093.92	447.28	1236.29
CIQ0035	−118.2040	34.4660	66.03	2428.36	710.10	864.55
CE13096	−117.2664	33.6989	66.46	4105.94	518.90	426.84
CE23292	−117.5444	34.4257	66.98	1807.64	710.10	1211.92
CIIPT	−117.2850	34.1970	67.45	2552.71	228.20	945.86
CIPER	−117.2053	33.8616	67.67	2880.41	518.90	467.03
CIQ0028	−117.1810	33.8310	70.30	3197.89	518.90	461.22
CIQ0013	−118.0600	34.5400	70.31	2620.60	518.90	859.30
CE13927	−117.1731	33.9212	70.31	2377.65	351.90	494.08
CISOF	−117.5570	33.3710	70.35	2333.99	351.90	16.09
CILUG	−117.3668	34.3656	72.20	2080.27	513.69	1136.43
CISBPX	−117.2348	34.2324	73.34	2310.65	293.50	1872.13
CE13924	−117.1274	33.7475	76.98	4161.26	351.90	486.31
CIQ0049	−117.1260	34.1970	80.69	2184.63	710.10	1661.03
CIBBS	−116.9806	33.9214	88.03	1639.46	518.90	782.79
CE12919	−116.9726	33.9299	88.76	1559.19	518.90	795.50
CIQ0020	−116.9530	33.9600	90.66	1588.16	468.40	859.36

perturbations (e.g. Hartzell *et al.* 2010) are likewise expected to have limited effect on the results within the narrow band considered here. The large underprediction in Fig. 3, confined to type B sites and most significant at frequencies ≈ 0.5–1.0 Hz, is likely controlled by the artificially high shallow V_S in the CVM-S at those sites.

4 VELOCITY TAPER METHOD

Ely *et al.* (2010) proposed a method for tapering shallow velocities in SCEC CVMs. The method first multiplies the V_{S30} by a uniform constant (the coefficient a in equation eq. 2, determined by trial and error) to derive the surface V_S, which is used to infer V_P and density following the scaling laws of Brocher (2005). It should be noted that this method may not preserve the original V_{S30}, albeit the deviation is generally small. V_P, V_S and density at the transition depth are directly extracted from the velocity model. V_P and V_S are independently interpolated between the surface and the transition depth, and density is again calculated via the Nafe–Drake law (Ludwig *et al.* 1970). The revised velocities, as a function of depth, are obtained by:

$$\begin{aligned} z &= z'/z_T \\ f(z) &= z + b\left(z - z^2\right) \\ g(z) &= a - az + c\left(z^2 + 2\sqrt{z} - 3z\right) \\ V_S(z) &= f(z)V_{ST} + g(z)V_{S30} \\ V_P(z) &= f(z)V_{PT} + g(z)P\left(V_{S30}\right) \\ \rho(z) &= R\left(V_P\right) \end{aligned} \tag{2}$$

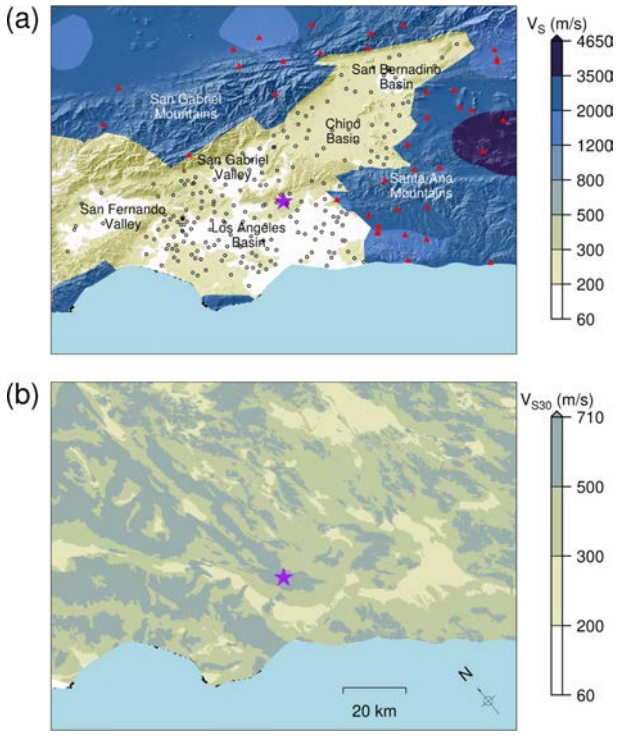

Figure 4. (a) Surface V_S extracted from CVM-S and (b) V_{S30} from Thompson (2018) in our model domain (values in the southwest corner are not available). The circles and red triangles in (a) depict type A and B sites, respectively, as in Fig. 1. The star denotes the epicentre of the La Habra event.

where z' is the depth, z_T is the transition (taper) depth, z is a normalized depth, and $f(z)$ and $g(z)$ are functions defined for formulating the resulting V_P and V_S. V_{ST} and V_{PT} are the S- and P-wave velocities extracted from the velocity model at z_T, respectively, and P and R are the Brocher (2005) V_P scaling law and Nafe–Drake law, respectively. The coefficient a controls the ratio of surface V_S to V_{S30}, and b and c control the overall and near-surface curvatures, respectively. The method generates a profile as a function of depth minimally parametrized by V_{S30}, properties at the transition depth and three empirical coefficients only, which greatly simplifies the introduction of the model modifications into the velocity mesh. The coefficients ($a = 1/2$, $b = 2/3$, $c = 3/2$) proposed by Ely *et al.* (2010) are calibrated to match the generic rock profiles of Boore & Joyner (1997) and Magistrale *et al.* (2000).

V_{S30} is one of the key parameters, along with z_T, controlling the profile generated using the method proposed by Ely *et al.* (2010). The V_{S30} values adopted by Ely *et al.* (2010) were obtained from the geology-based V_{S30} map of Wills & Clahan (2006) for California and the topography-based estimations by Wald & Allen (2007) outside California. Thompson *et al.* (2014) proposed a V_{S30} map for California based on regression kriging to incorporate multiple constraints from geology, topography and site-specific V_{S30} measurements at various spatial scales based on the method by Wills & Clahan (2006), and later updated to the V_{S30} map by Wills *et al.* (2015). This approach is adopted by the U.S. Geological Survey (Thompson 2018) for California, which we adopt to calibrate near-surface velocities in CVM-S.

Fig. 4 shows a comparison of surface V_S values extracted from CVM-S to the V_{S30} values from Thompson (2018) in our model domain. For type B sites, it is clear that surface velocities are unrealistically high compared to the V_{S30} values. This discrepancy motivated the V_S tapering method by Ely *et al.* (2010), which replaces the original velocities from the surface to the transition depth z_T, while leaving velocities below z_T unchanged. Note, that the Ely *et al.* (2010) method does not necessarily maintain low velocities in the original model, which are always overwritten by the calculated profile. For type B sites, where the surface velocities in CVM-S are typically much larger than the corresponding V_{S30} values, this velocity tapering works as intended to lower unrealistically large shallow V_S values. However, for type A sites, the benefits of this method are less clear, as the shallow V_S values in CVM-S are close to, or sometimes smaller than the V_{S30} from Thompson (2018). In addition, the near-surface velocities at type A sites are derived from a combination of detailed well logs and other geotechnical information (Magistrale *et al.* 2000; Small *et al.* 2017), often different and likely more accurate than the result of the Ely *et al.* (2010) GTL.

For the reasons mentioned above, we propose and test the following variant of the Ely *et al.* (2010) method for assigning the shallow velocities in our model domain. The velocity mesh is queried from CVM-S [without the Ely *et al.* (2010) GTL] using UCVM, and we apply the tapering method described by eq. (2) above a specified depth z_T, which has two input values, namely the material properties at z_T and the V_{S30} value. The former is extracted from the velocity mesh, and for the latter we adopt V_{S30} values from Thompson (2018) across the entire domain. Between the surface and z_T, we replace the V_S values in CVM-S by eq. (2) whenever the former exceeds the latter. Fig. 5 illustrates the application of our method on two representative profiles by averaging two groups of velocity profiles for type A and type B sites. At type B sites, V_S is reduced (relative to original CVM-S values) at all depths above z_T. At type A sites, the effect is more variable. Since CVM-S already includes a low-velocity GTL (from geological measurements and borehole data) down to 350 m in part of California (note the abrupt

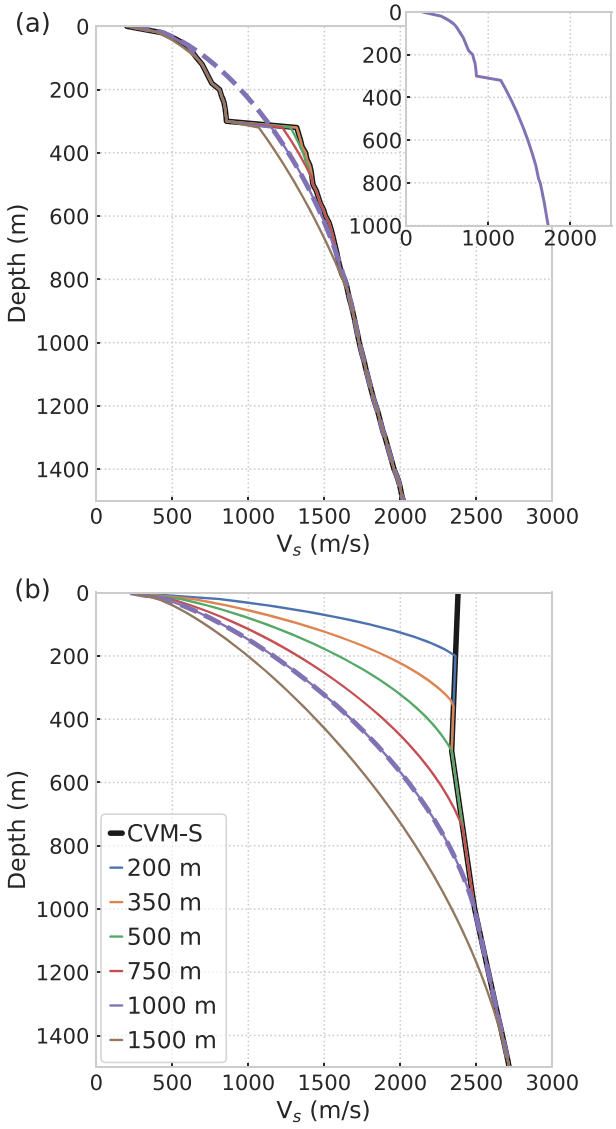

Figure 5. Representative V_S profiles for (a) type A sites and (b) type B sites from CVM-S. The thick black curves depict the averaged velocity profiles for all 220 type A and 39 type B sites directly extracted from the original [without the Ely *et al.* (2010) GTL] CVM-S. The thin lines show the V_S profiles resulting from our proposed method for different z_T depths between 200 and 1500 m. The dashed curve shows the V_S profile calculated using tapers from eq. (2) tapers with our preferred z_T of 1000 m (note that because the tapers are applied as upper bounds to V_S, they typically only affect the type A V_S structure at depths exceeding 350 m, where the GTL in CVM-S terminates and causes the abrupt discontinuity). The inset figure in (a) shows the resulting profile with 1000 m z_T.

discontinuity in Fig. 5a), the typical type A profile is mostly unaffected when z_T is small (e.g. the two smallest z_T values for the profiles shown in Fig. 5). For larger values of z_T, type A sites velocities are typically reduced only for depths between 350 m and z_T. Note that, because we impose eq. (2) as an upper bound rather than as an equality, we can apply the method to CVM-S without explicitly identifying type A and type B sites *a priori* (e.g. we do not have to worry about eq. (2) inadvertently overwriting low sediment velocities at type A sites, as illustrated by the purple dashed curve in Fig. 5).

5 SH1D THEORETICAL ANALYSIS

Before performing computationally expensive 3-D numerical simulations, we use a theoretical approach to estimate the threshold depth z_T. We model vertically incident *SH* waves in a horizontally layered half-space (hereafter referred to as 'SH1D' modelling) to obtain a preliminary estimate of the effects of the velocity taper. SH1D is widely used in theoretical (e.g. Day 1996) and numerical (e.g. Thompson *et al.* 2012) analysis of elastic site response.

Fig. 6 shows the 1-D site amplification functions for tapered V_S profiles from CVM-S (shown as FAS ratios of the tapered to untapered case), grouped into type A and type B sites. The curves indicate the effect of tapering depth z_T on site amplification (in the 1-D approximation). As the tapering depth increases, the shallow velocity decreases, generating larger amplification. For type A sites, amplification (i.e. FAS ratio greater than one) is only obtained for z_T larger than 350 m. Amplification is generated below 0.5 Hz and above $\approx$ 0.8 Hz, with de-amplification

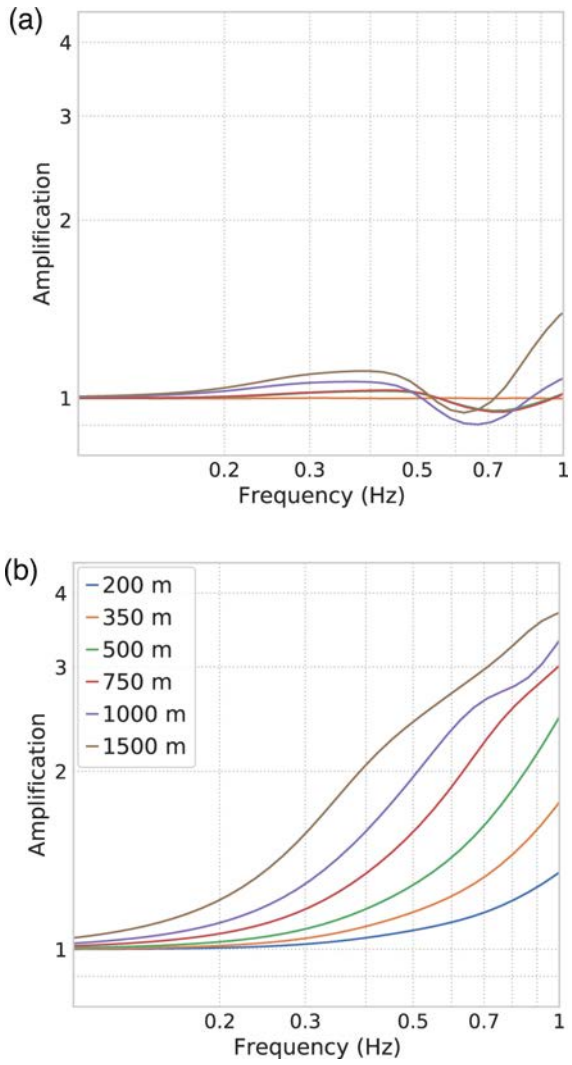

Figure 6. SH1D response for the refined profiles using a series of z_T depths for average (a) type A and (b) type B sites, divided by the response obtained with the averaged type A and type B profiles from CVM-S, respectively.

in between these frequencies. Deeper tapering depths tend to further decrease near-surface velocities and produce stronger amplification. The narrow band of de-amplification appears only for profiles with tapering depth greater than 350 m, where the sharp discontinuity starts to subside. The de-amplification band gets narrower as the tapering depth increases, leaving the profiles smoother near 350 m depth. We therefore attribute the de-amplification seen in some of the type A sites to the removal of the velocity contrast at 350 m depth. In general, the changes at type A sites are relatively small, with less than 10 per cent amplification or de-amplification for tapering shallower than 1000 m. The amplification at type B sites, on the other hand, increases monotonically with the tapering depth, as expected from the pattern of velocity reduction in Fig. 5. The type B amplifications can be quite large; for example, they exceed a factor of 2 in the 0.5–1.0 Hz range for a taper depth of 1000 m.

We further examined the effects on the resulting synthetics from applying these velocity tapers; we did so by combining the 3-D simulations (which used the original, untapered CVM) with the SH1D amplification results, as shown in Fig. 7. At every station, and for each tapered profile, we calculated the FAS of the 3-D simulated acceleration, divided by the FAS of the recorded acceleration, and multiplied by the SH1D amplification of the tapered profile. We then averaged across all sites in type A and B site groups, respectively. As expected, type A sites show limited effects of superimposing the velocity tapering, and larger z_T generally yields more amplification, with the exception of $z_T < 500$ m for 0.7–1 Hz. Unlike the muted effects of the velocity tapering for type A sites, the range of tapering depths produce much greater amplification effects for type B sites. Deeper tapering depths generate a favourable fit for 0.2–0.4 Hz, but tend to overpredict above 0.4–0.5 Hz for $z_T \geq 750$ m. On the other hand, in the 0.5-1.0 Hz range, taper depths in the range 350–750 m appear more favourable. These estimates from SH1D encourage us to explore tapering depths in the range 350–1000 m using 3-D simulations in the next section.

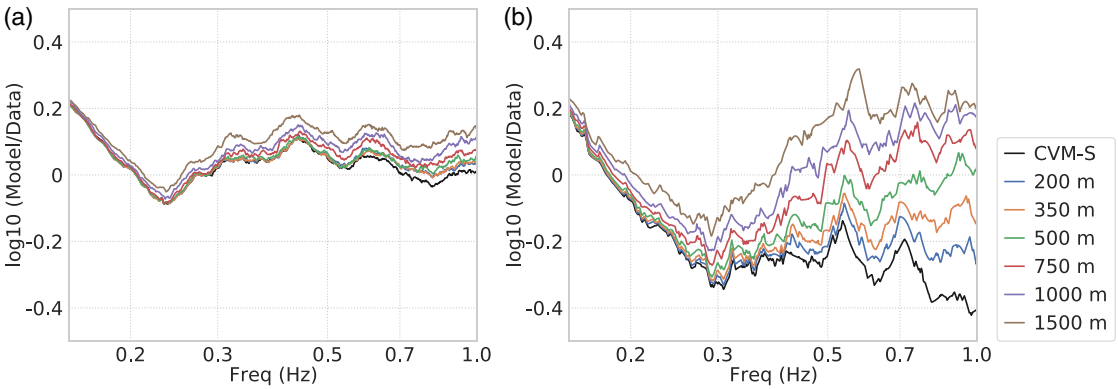

Figure 7. Bias of FAS for the two horizontal components averaged over all (a) type A and (b) type B sites from CVM-S at all 259 stations, superimposed with the corresponding SH1D response. The black curves denote CVM-S and other labelled curves represent various tapering depths using SH1D results.

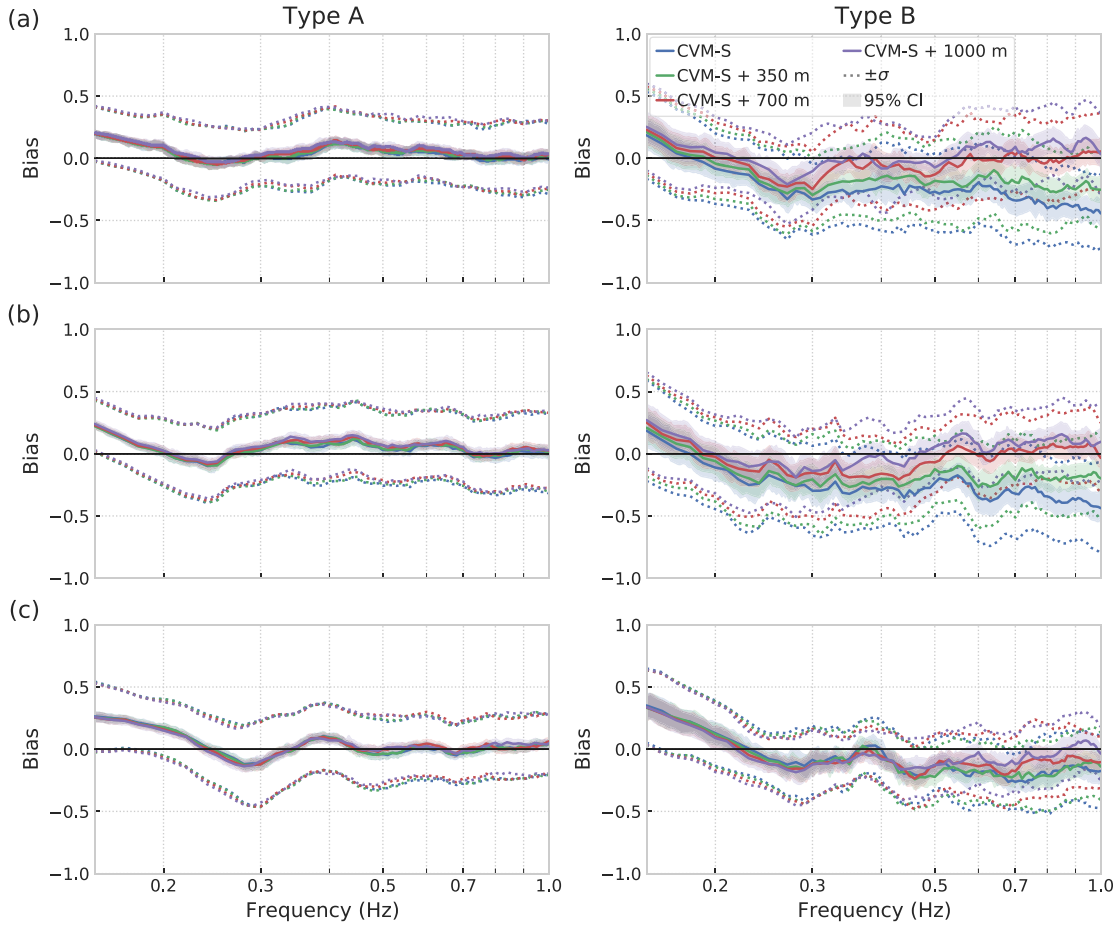

Figure 8. Bias of FAS on the (a) east–west, (b) north–south and (c) vertical components, calculated from 3-D simulations in CVM-S, and in CVM-S with tapering depths of 350, 700 and 1000 m. A positive (negative) value depicts overprediction (underprediction). The left-hand (right-hand) column shows type A (B) sites. The solid line is the median of FAS, 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.

6 3-D NUMERICAL SIMULATIONS

We carried out three additional 3-D simulations using CVM-S, with z_T of 350 m (Ely *et al.*, 2010), 700 m and 1000 m, respectively. Fig. 8 shows the resulting FAS bias for all three components at both type A and B sites. There is a negligible visual difference at type A sites for all components because our velocity tapering method leaves the original low velocities virtually unchanged. On the other hand, these models show significant differences at type B sites, where the original near-surface V_S are deemed too large. Both the 700 and 1000 m tapering models eliminate the critical underprediction from CVM-S for the two horizontal components. However, between these two models, the 1000 m model produces a better fit below 0.5 Hz while slightly overpredicting above 0.5 Hz for the horizontal components. Effects of the velocity

Table 3. Average FAS bias for all three components from simulations with various models.

Model	Type A sites				Type B sites			
	East–west	North–south	Vertical	Average	East–west	North–south	Vertical	Average
CVM-S*	0.034	0.044	0.009	**0.029**	-0.277	-0.261	-0.136	**-0.225**
CVM-S + 350 m*	0.040	0.048	0.009	**0.033**	-0.171	-0.153	-0.138	**-0.154**
CVM-S + 700 m*	0.055	0.062	0.018	**0.045**	-0.020	-0.015	-0.087	**-0.041**
CVM-S + 1000 m*	0.065	0.073	0.020	**0.053**	0.048	0.055	-0.055	**0.016**
CVM-S + 350 m + $Q_S = 0.05V_S$ †	−0.039	−0.027	-0.091	**-0.052**	-0.064	-0.052	-0.156	**-0.091**
CVM-S + 1000 m + $Q_S = 0.15V_S$ †	0.085	0.091	0.061	**0.080**	-0.135	-0.120	-0.105	**-0.120**

* $Q_S = 0.1V_S$; $Q_P = 2Q_S$.
† $Q_P = 2Q_S$.

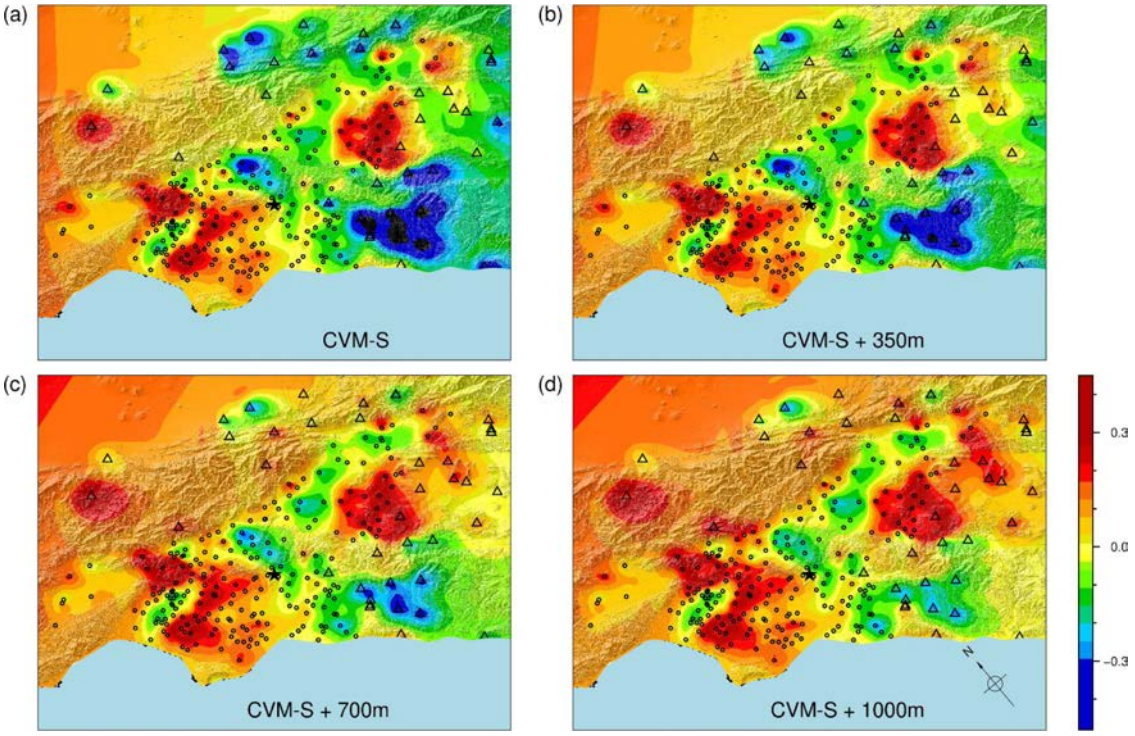

Figure 9. Maps of interpolated log10-based FAS bias between four 3-D models and data: (a) CVM-S, and CVM-S with tapering depths of (b) 350 m, (c) 700 m and (d) 1000 m, calculated from the synthetics and records at 259 stations. The warm (cool) colours represent overprediction (underprediction). The circles (triangles) depict type A (B) sites. Note the log10-based colour bar.

taper are smaller for the vertical component, where the 1000 m model again is superior with a slight underprediction above 0.2 Hz. Thus, the 1000 m velocity tapering model provides the best fit across almost the entire frequency band.

It is helpful to quantify the results using a single goodness of fit (GOF) metric, which we define as the average of the median bias over a group of sites:

$$\mathrm{GOF(model,\ component)} = \frac{1}{\mathrm{nfreq}} \sum_{i=1}^{\mathrm{nfreq}} \mathrm{median(Bias(frequency}_i,\ \mathrm{site)}), \tag{3}$$

where nfreq is the number of discrete frequencies in the FAS calculation. We prefer using the median of the bias over stations to minimize the effects of outliers. Table 3 lists the bias for different components and the single GOF value, averaged for the three components, for various tapering models. Resulting biases are −0.22, −0.154, −0.041 and 0.016 (corresponding to amplitude underpredictions of 40, 30, 10 per cent, and overprediction of 4 per cent) for taper depths of 0, 350, 700 and 1000 m, respectively.

Fig. 9 shows a map of interpolated horizontal FAS bias for CVM-S and our three models with tapering depths of 350, 700 and 1000 m. The large basin areas with very low near-surface velocities, including the central Los Angeles and Chino basins, show almost no variability among these models. Despite the fairly large spatial variability, the median FAS ratio in the basins is generally small and insensitive to the tapering depth. More importantly, we see significant improvement of the bias outside of the basins. For example, the Santa Ana mountains (see Fig. 1 for location) suffer strong underpredictions in CVM-S, and the GOF improves as the tapering depth increases to 1000 m.

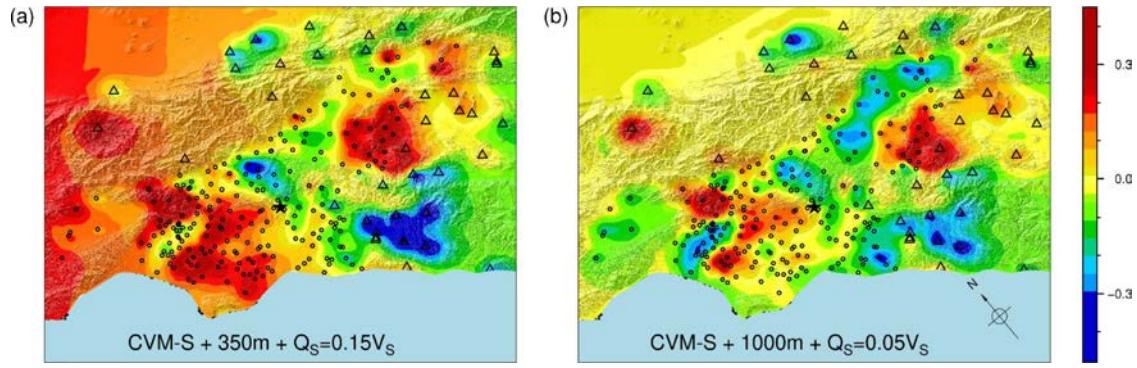

Figure 10. Maps of interpolated log10-based FAS bias for two 3-D CVMs and data. (a) CVM-S with velocity tapering depth of 350 m and $Q_S = 0.15V_S$, and (b) CVM-S with velocity tapering depth of 1000 m and $Q_S = 0.05V_S$. Warm (cool) colours represent overprediction (underprediction). Circles depict type A sites and triangles show type B sites.

7 DISCUSSION AND CONCLUSIONS

A taper depth z_T of 1000 m provides the largest improvement in FAS bias (as measured by GOF, eq. 3) for type B sites that we could find using a single generic taper over the region of interest (Fig. 1). This value is substantially different than the 350 m value of Ely *et al.* (2010). The spatial distribution of FAS bias in Fig. 10 suggests that additional improvements may be possible by permitting spatial variations of the V_S tapering. For example, the areas of type B sites with relatively small underprediction remaining in the preferred 1000 m tapering model suggests the need for V_S modification to even larger depths. However, the eastern termination of the San Gabriel Mountains (see Fig. 1 for location) shows underprediction for the 350 m model and overprediction for the 700 and 1000 m models, which indicates the need for tapering with z_T in the 350–700 m range. We also note a slightly degraded fit at type B sites east of the Chino Basin (see Fig. 4 for location). Possible reasons for this degraded fit include the fact that this area features a relatively abrupt separation of regions with low and high velocities right at the boundary, which complicates a laterally homogeneous tapering depth. In addition, the accuracy of the topography-based V_{S30} estimation in this area may be decreased by relatively high elevation and limited surface topographic slopes, which may induce less accurate topography-based V_{S30} estimation.

The optimal tapering depth may also be affected by the anelastic attenuation. We parametrize anelastic attenuation as a function of local V_S, a commonly accepted procedure for ground motion estimation (e.g. Olsen *et al.* 2003; Savran & Olsen 2019; Lai *et al.* 2020). Our choice of the relation $Q_S = 0.1V_S$ (V_S in m s^{-1}; $Q_P = 2Q_S$) relation is based on the results from modelling the 2008 M_w 5.4, Chino Hills, earthquake by Savran & Olsen (2019). However, to examine whether the overprediction in the valleys (underpredictions in the mountain areas) diminishes with lower (higher) Q, we tested two additional models: (1) CVM-S with z_T of 1000 m using $Q_S = 0.05V_S$ and (2) CVM-S with z_T of 350 m with $Q_S = 0.15V_S$ (see Fig. 10). Although reducing the overprediction in the valleys, the $Q_S = 0.05V_S$ model reduces the FAS below the levels observed in the data as distance increases. In addition, the $Q_S = 0.05V_S$ model increasingly degrades the fit on the vertical component at type A sites (see Table 3), as the frequency increases toward 1 Hz (see Supporting Information Fig. S1) and likely beyond. On the other hand, the $Q_S = 0.15V_S$ model increases the FAS, mostly above the basins and provides little improvement at type B sites. For these reasons, we prefer the $Q_S = 0.1V_S$, $Q_P = 2Q_S$ model.

Fig. 11 compares recorded and simulated three-component FAS and cumulative kinetic energy for a subset of 6 out of 39 type B sites in our model domain. Here, cumulative energy is defined as $\int_0^T v(t)^2 \mathrm{d}t$, where $v(t)$ is particle velocity as a function of time and T is duration. The six sites are selected throughout the domain, representing a broad range of fit between synthetics and data. Both metrics show amplification on the horizontal components from the velocity tapering that reduces the shallow velocities. The amplification on the vertical component generally increases less with the tapering depth (e.g. sites CIQ0022 and CISDD), indicating that the horizontal components are more sensitive to site amplification effects, in agreement with previous studies (e.g. Bonilla *et al.* 1997; Gülerce & Abrahamson 2011). At almost all type B sites, CVM-S underpredicts the FAS below 1 Hz compared with the records, while the tapered models mitigate the underprediction. Note, however, that different sites show very different peak amplification frequencies, which complicates the definition of a single, domain-wide z_T. For example, sites CIQ0022 and CE13080 have similar surface V_S (see Table 2) in the Supporting Information, but CIQ0022 shows significant underprediction and CE13080 shows overprediction for the 1000 m tapering model. Also, the low-frequency (0.2–0.3 Hz) FAS peaks present at site CIIPT, CISTG, CIQ0022 and CISDD are likely due to other local site effects. Supporting Information Fig. S2 shows the FAS bias at all type B sites from different models as a function of surface V_S. All models show the trend that the bias increases with surface V_S, indicating that sites with larger surface V_S need relatively weaker velocity reduction or shallower velocity tapering. In summary, our method introduces first-order improvement in overall type B site amplification, while additional fine-tuning of local amplification requires additional work. Such analysis should include additional GOF metrics, as well as using simulations of multiple events.

Another family of velocity models for southern California, CVM-H, was originally developed by Süss & Shaw (2003) and later improved by Plesch *et al.* (2007, 2009, 2011) with incorporation of the tomography results from Tape *et al.* (2009, 2010). CVM-H supports the option to include the Ely *et al.* (2010) GTL, with a default transition depth of 350 m, across the entire domain. Taborda *et al.* (2016) performed 3-D deterministic simulations of small earthquakes in California and showed that CVM-S consistently provides overall superior fit to records as

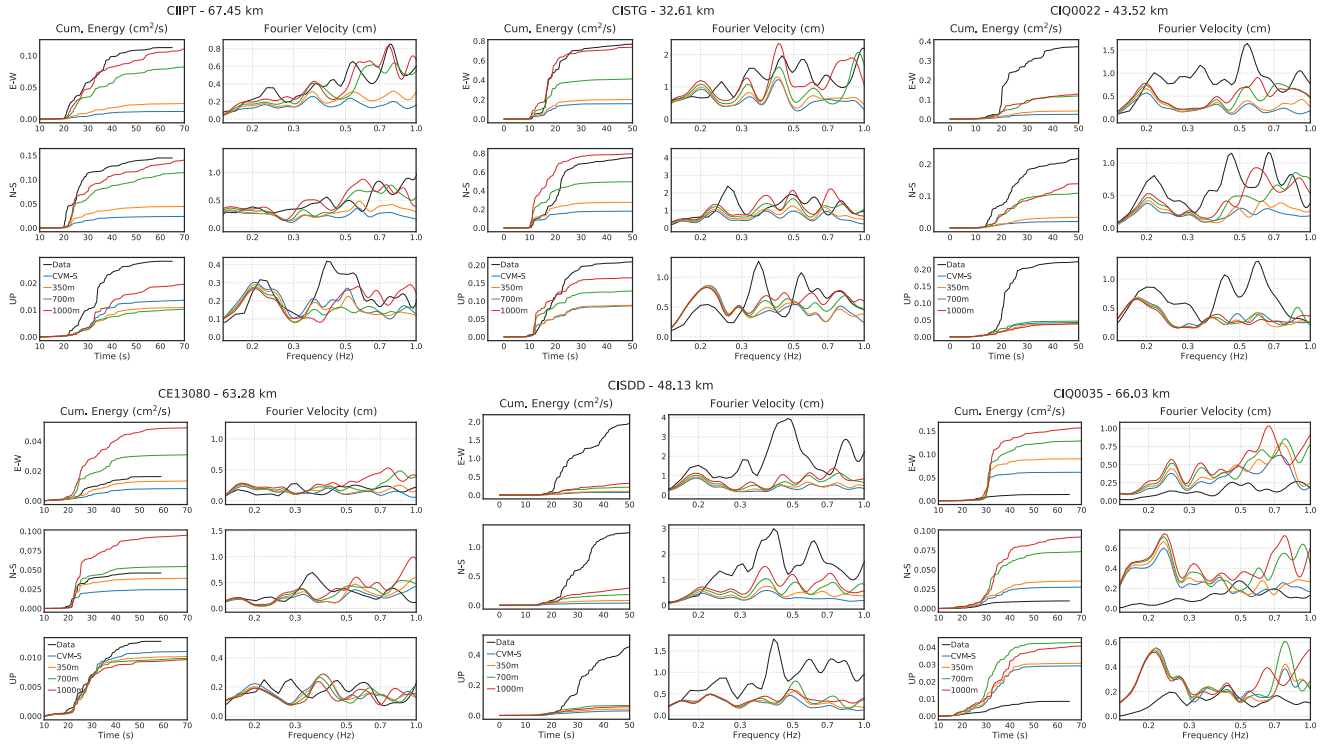

Figure 11. Cumulative kinetic energy and Fourier velocity spectra at six type B sites. The subtitles show the names of the sites and their hypocentral distances. See Fig. 1 for site locations.

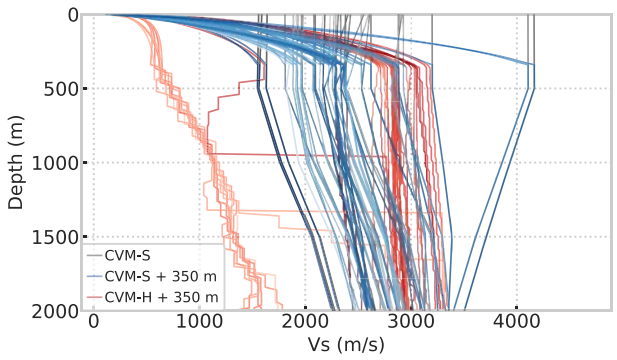

Figure 12. Type B site V_S profiles from CVM-S, and CVM-S and CVM-H with (default) Ely *et al.* (2010) GTL taper depth of 350 m.

compared to CVM-H for frequencies up to 1 Hz, with or without the Ely *et al.* (2010) GTL (z_T of 350 m). They also noted that the addition of the Ely *et al.* (2010) GTL generally improved the GOF. Fig. 12 shows a comparison of V_S profiles from CVM-S and CVM-H at the recording stations located on type B sites with records from the La Habra event in our model domain. We note that CVM-H includes topography and the material properties are defined relative to sea level. Here, we rely on the built-in algorithm of UCVM to convert depth to elevation by querying CVM-H by depth to keep consistency with CVM-S, which operates by depth from the free surface. It is clear that the most poorly constrained site profiles from CVM-S and CVM-H that include the Ely *et al.* (2010) GTL with a transition depth of 350 m are similar. We therefore expect that CVM-H can benefit from a deeper tapering depth (≈1000 m) at type B sites, similar to what provided significant improvement for CVM-S.

In Fig. 13, we uniformly sampled 300 locations in California and queried their velocity profiles from CVM-S and CVM-H (note that the two models default to a background model outside their respective boundaries). The majority of type A sites are located in the west and south and most type B sites in the east. Similar to our simulation domain (Fig. 2), the state-wide sampling shows that most type B sites lack sufficient resolution in the top 1000 m, likely requiring calibration of the shallow velocities. For this reason, the application of our proposed tapering depth will likely improve the accuracy of ground motion simulations in other regions of California.

While our analysis was limited to frequencies below 1 Hz, the proposed tapering of the near-surface V_S structure for type B sites in CVM-S will likely benefit future ground motion predictions to even higher frequencies. As the available computational resources increase, ground motion simulations can be extended to higher frequencies, where model features such as topography, small-scale crustal heterogeneities and

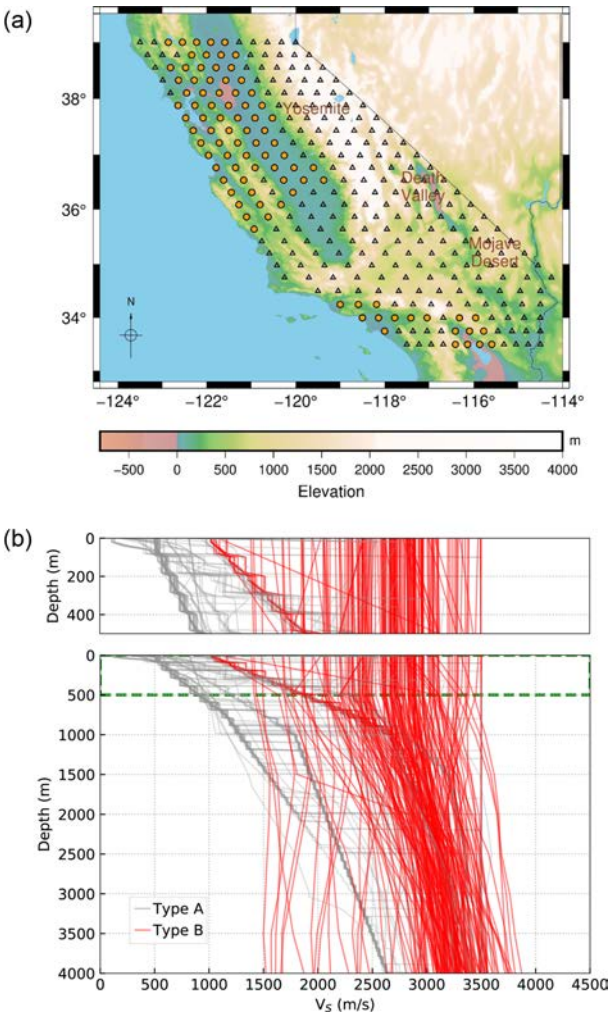

Figure 13. (a) V_S profile sample locations in California. Circles denote type A sites and triangles denote type B sites, and (b) extracted V_S profiles. The top panel zooms into the top 500 m.

frequency-dependent attenuation play an increasingly large role. However, unless the underlying velocity model is sufficiently accurate, these model features may cause trade-offs in the results. For example, unrealistically large near-surface velocities, if present, may trade-off with attenuation quality factors to compensate for the underprediction. To further isolate shallow low-velocity effects and examine the efficacy of our velocity tapering method, higher-frequency simulations for multiple earthquakes that resolve the velocity structure in different directions and paths are demanded, along with validations via more ground motion metrics.

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 11782). SCEC is funded by the National Science Foundation (NSF) Cooperative Agreement EAR-1600087 and the U.S. Geological Survey (USGS) Cooperative Agreement G17AC00047. This work was initiated from the 'High-F' collaboration in SCEC. We thank Robert Graves for providing the source models, and Fabio Silva for providing the station records of the 2014 La Habra earthquake. We appreciate the constructive reviews from Erdinc Saygin and one anonymous reviewer leading to an improved paper.

DATA AVAILABILITY

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 intensity measures 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

Abrahamson, N.A., Silva, W.J. & Kamai, R., 2014. Summary of the ASK14 ground motion relation for active crustal regions, *Earthq. Spectra,* **30**(3), 1025–1055.

Anderson, J.G., Lee, Y., Zeng, Y. & Day, S., 1996. Control of strong motion by the upper 30 meters, *Bull. seism. Soc. Am.,* **86**(6), 1749–1759.

Bielak, J. *et al.*, 2010. The ShakeOut earthquake scenario: verification of three simulation sets, *Geophys. J. Int.,* **180**(1), 375–404.

Bonilla, L.F., Steidl, J.H., Lindley, G.T., Tumarkin, A.G. & Archuleta, R.J., 1997. Site amplification in the San Fernando Valley, California: variability of site-effect estimation using the S-wave, coda, and H/V methods, *Bull. seism. Soc. Am.,* **87**(3), 710–730.

Boore, D.M., 1972. A note on the effect of simple topography on seismic *SH* waves, *Bull. seism. Soc. Am.,* **62**(1), 275–284.

Boore, D.M. & Joyner, W.B., 1991. Estimation of ground motion at deep-soil sites in eastern North America, *Bull. seism. Soc. Am.,* **81**(6), 2167–2185.

Boore, D.M. & Joyner, W.B., 1997. Site amplifications for generic rock sites, *Bull. seism. Soc. Am.,* **87**(2), 327–341.

Boore, D.M., Stewart, J.P., Seyhan, E. & Atkinson, G.M., 2014. NGA-West2 equations for predicting PGA, PGV, and 5 per cent damped PSA for shallow crustal earthquakes, *Earthq. Spectra,* **30**(3), 1057–1085.

Borcherdt, R.D., 1994. Estimates of Site-dependent response spectra for design (methodology and justification), *Earthq. Spectra,* **10**(4), 617–653.

Bouchon, M. & Barker, J.S., 1996. Seismic response of a hill: the example of Tarzana, California, *Bull. seism. Soc. Am.,* **86,** 66–72.

Bozorgnia, Y. *et al.*, 2014. NGA-West2 research project, *Earthq. Spectra,* **30**(3), 973–987.

Brocher, T.M., 2005. Empirical relations between elastic wavespeeds and density in the Earth's crust, *Bull. seism. Soc. Am.,* **95**(6), 2081–2092.

Campbell, K.W. & Bozorgnia, Y., 2014. NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5 per cent damped linear acceleration response spectra, *Earthq. Spectra,* **30**(3), 1087–1115.

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.

Chen, P., Zhao, L. & Jordan, T.H., 2007. Full 3D tomography for the crustal structure of the Los Angeles region, *Bull. seism. Soc. Am.,* **97**(4), 1094–1120.

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,* Dallas, TX , pp. 1–20, IEEE.

Day, S.M., 1996. RMS response of a one-dimensional half-space to *SH*, *Bull. seism. Soc. Am.,* **86**(2), 363–370.

Durand, S., Gaffet, S. & Virieux, J., 1999. Seismic diffracted waves from topography using 3-D discrete wavenumber-boundary integral equation simulation, *Geophysics,* **64**(2), 572–578.

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.

Field, E.H., 2000. A Modified ground-motion attenuation relationship for southern California that accounts for detailed site classification and a basin-depth effect, *Bull. seism. Soc. Am.,* **90,** S209–S221.

Fujiwara, H., Kawai, S., Hao, K.X., Morikawa, N. & Azuma, H., 2017. J-SHIS—an integrated system for sharing information on national seismic hazard maps for Japan, in *16th World Conference on Earthquake, 16WCEE 2017,* Santiago, Chile, pp. 7.

Gilbert, G.K., Holmes, J., Humphrey, R.L., Sewell, J. & Soule, F., 1907. *The San Francisco earthquake and fire of April 18, 1906, and their effects on structures and structural materials,* doi:10.3133/b324.

Gülerce, Z. & Abrahamson, N.A., 2011. Site-specific design spectra for vertical ground motion, *Earthq. Spectra,* **27**(4), 1023–1047.

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., 1995. Preliminary analysis of long-period basin response in the Los Angeles region from the 1994 Northridge earthquake, *Geophys. Res. Lett.,* **22**(2), 101–104.

Hartzell, S., Harmsen, S. & Frankel, A., 2010. Effects of 3D random correlated velocity perturbations on predicted ground motions, *Bull. seism. Soc. Am.,* **100**(4), 1415–1426.

Heath, D.C., Wald, D.J., Worden, C.B., Thompson, E.M. & Smoczyk, G.M., 2020. A global hybrid V_{S30} map with a topographic slope-based default and regional map insets, *Earthq. Spectra,* **36**(3), 1570–1584.

Hu, Z., Olsen, K.B. & Day, S.M., 2022. 0-5 Hz deterministic 3D ground motion simulations for the 2014 La Habra, California, earthquake, *Geophys. J. Int*.

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.

International Code Council, 2014. *2015 IBC International Building Code,* International Code Council.

Joyner, W.B., Warrick, R.E. & Fumal, T.E., 1981. The effect of quaternary alluvium on strong ground motion in the Coyote Lake, California, earthquake of 1979, *Bull. seism. Soc. Am.,* **71**(4), 1333–1349.

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), doi:10.1029/2020JB019663.

Lee, E.-J., Chen, P., Jordan, T.H., Maechling, P.B., Denolle, M. A.M. & Beroza, G.C., 2014. Full-3-D tomography for crustal structure in Southern California based on the scattering-integral and the adjoint-wavefield methods, *J. geophys. Res.,* **119**(8), 6421–6451.

Lee, V.W. & Trifunac, M.D., 2010. Should average shear-wave velocity in the top 30m of soil be used to describe seismic amplification?, *Soil Dyn. Earthq. Eng.,* **30**(11), 1250–1258.

Lin, G., Shearer, P.M., Hauksson, E. & Thurber, C.H., 2007. A three-dimensional crustal seismic velocity model for southern California from a composite event method, *J. geophys. Res.,* **112**(B11), doi: 10.1029/2007JB004977.

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.

Ludwig, W., Nafe, J. & Drake, C., 1970. Seismic Refraction,*The Sea,***4,**53–84.

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.

Massa, M., Barani, S. & Lovati, S., 2014. Overview of topographic effects based on experimental observations: meaning, causes and possible interpretations, *Geophys. J. Int.,* **197**(3), 1537–1550.

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.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.

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.

Plesch, A., Suess, M., Munster, J., Shaw, J.H., Hauksson, E. & Tanimoto, T., USR Working Group, 2007. A new velocity model for southern California: CVM-H 5.0, *SCEC Annual Meeting, Proceedings and Abstracts,***17,**159.

Plesch, A., Tape, C. & Shaw, J.H., USR Working Group, 2009. CVM-H 6.0: Inversion integration, the San Joaquin valley and other advances in

the community velocity model, *SCEC Annual Meeting, Proceedings and Abstracts,* **19,** 260.

Plesch, A., Tape, C., Graves, R.W., Shaw, J.H. & Ely, G.P., 2011. Updates for the CVM-H including new representations of the offshore Santa Maria and San Bernardino basin and a new Moho surface, *SCEC Annual Meeting, Proceedings and Abstracts,* **21,** 214.

Qiu, H., Lin, F.-C. & Ben-Zion, Y., 2019. Eikonal tomography of the southern California Plate Boundary region, *J. geophys. Res.,* **124**(9), 9755–9779.

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.

Shingaki, Y., Goto, H. & Sawada, S., 2018. Evaluation performance for site amplification factors: S-wave impedance vs. V30, *Soils Foundations,* **58**(4), 911–927.

Small, P., Gill, D., Maechling, P.J., Taborda, R., Callaghan, S., Jordan, T.H., Olsen, K.B., Ely, G.P. & Goulet, C., 2017. The SCEC unified community velocity model software framework, *Seismol. Res. Lett.,* **88**(6), 1539–1552.

Süss, M.P. & Shaw, J.H., 2003. *P* wave seismic velocity structure derived from sonic logs and industry reflection data in the Los Angeles basin, California: 3-D velocity structure in Los Angeles basin, *J. geophys. Res.,* **108**(B3).

Steidl, J.H., 2000. Site response in southern California for probabilistic seismic hazard analysis, *Bull. seism. Soc. Am.,* **90,** S149–S169.

Steller, R., 1996. *New borehole geophysical results at GVDA,NEES@ UCSB Internal Report.*.

Stephenson, W.J., Reitman, N.G. & Angster, S.J., 2017. *Cascadia subduction zone for 3D earthquake ground motion simulations, Version 1.6,Update for Open-File Report 2007–1348,* No. 2017-1152.

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.

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2009. Adjoint tomography of the southern California crust, *Science,* **325**(5943), 988–992.

Tape, C., Liu, Q., Maggi, A. & Tromp, J., 2010. Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.,* **180**(1), 433–462.

Teague, D.P., Cox, B.R. & Rathje, E.M., 2018. Measured Vs. predicted site response at the Garner Valley Downhole Array considering shear wave velocity uncertainty from borehole and surface wave methods, *Soil Dyn. Earthq. Eng.,* **113,** 339–355.

Thompson, E.M., 2018. *An Updated Vs30 Map for California with Geologic and Topographic Constraints, US Geol. Surv. Data Release,* **10,** F7JQ108S.

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.

Thompson, E.M., Wald, D.J. & Worden, C.B., 2014. A VS30 map for California with geologic and topographic constraints, *Bull. seism. Soc. Am.,* **104**(5), 2313–2321.

USGS, 2014. Earthquake Events, Focal Mechanism.

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

Wald, D.J. & Allen, T.I., 2007. Topographic slope as a proxy for seismic site conditions and amplification, *Bull. seism. Soc. Am.,* **97**(5), 1379–1395.

Wills, C.J. & Clahan, K., 2006. Developing a map of geologically defined site-condition categories for California, *Bull. seism. Soc. Am.,* **96,** 1483–1501.

Wills, C.J., Gutierrez, C.I., Perez, F.G. & Branum, D.M., 2015. A next generation V_{S30} map for California based on geology and topography, *Bull. seism. Soc. Am.,* **105**(6), 3083–3091.

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. Bias of FAS of the (a) east–west, (b) north–south and (c) vertical components, calculated from 3-D simulations in CVM-S with V_S tapering depths of 350 and 1000 m along with attenuation models $Q_S = 0.05V_S$, $Q_S = 0.1V_S$ and $Q_S = 0.15V_S$. A positive (negative) value means overprediction (underprediction). The left-hand (right-hand) columns show type A (B) sites. The solid line is 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.

Figure S2. Averaged FAS bias for frequencies between 0.15 and 1 Hz at poorly constrained sites plotted as a function of site surface V_S for (a) three-component average, (b) east–west, (c) north–south and (d) vertical components. The shades represent 95 per cent confidence intervals estimated using bootstrap.

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.