

## **CALIBRATION OF THE NEAR-SURFACE SEISMIC STRUCTURE IN THE SCEC COMMUNITY VELOCITY MODEL VERSION S**

Kim Olsen and Te-Yang Yeh

Department of Earth and Environmental Sciences  
San Diego State University

### **Abstract**

We have used 0-1 Hz 3D numerical wave propagation simulations and strong motion recordings for 7  $M_w$ 4.4-5.1 events around the greater Los Angeles basins to estimate the optimal thickness distribution for a generic low-velocity overlay taper (LVT) implemented in the top 1-2 km of the Statewide California Earthquake Center Community Velocity Model (CVM) version S4.26.M01. The optimized LVT model with spatially-varying thickness leads to a reduction of the Fourier Amplitude Spectrum bias by 39% from the CVM without LVT, and by 22% from the CVM including a LVT with a constant thickness of 600 m.

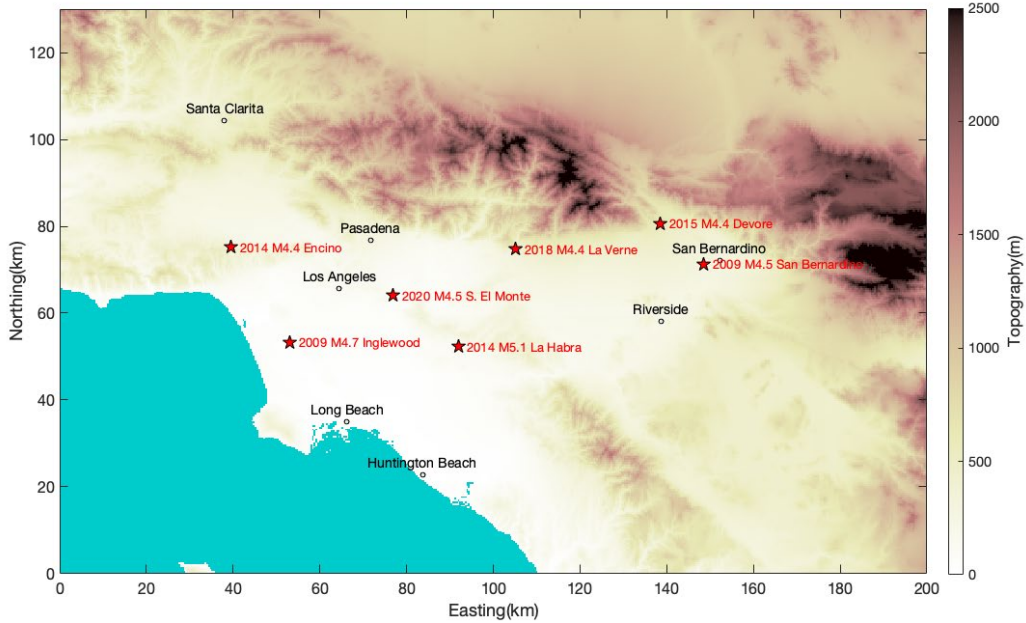
### **Introduction**

The near-surface seismic structure to a depth of about 1-2 km, particularly the shear wave velocity ( $V_s$ ), can strongly affect the seismic response. However, for tomographically derived models, and when constraints are missing from borehole studies, geotechnical measurements, and water and oil wells (such as at rock sites outside the sedimentary basins), the parameters of the shallow material are often poorly characterized. On the other hand, the  $V_s$  structure of the material deeper than about 1-2 km are typically reasonably-well determined by tomography studies.

Outside the basin areas, the Statewide California Earthquake Center (SCEC) Community Velocity Models (CVMs) generally default to regional tomographic estimates that do not resolve the uppermost  $V_s$  values. The SCEC Unified Community Velocity model (UCVM) software includes a method to refine the near-surface earth structure by applying a smooth generic overlay (hereafter referred to as a low-velocity taper, or LVT) that connects the time-averaged  $V_s$  in the top 30 m ( $V_{s30}$ ) to the velocity model at a depth of 350 m (Ely et al., 2010), which can be applied to any of the velocity models accessible through UCVM. However, Hu et al. (2022b) showed that 0-1 Hz 3D physics-based simulations in the Los Angeles area with the SCEC CVM-S4.26.M01 modified by a 350-m thick LVT significantly underpredict seismic recordings from the 2014  $M_w$ 5.1 La Habra earthquake. On the other hand, Hu et al. (2022b) showed that extending the LVT to a depth of 700-1,000 meters improved the fit between their synthetics and seismic data significantly, without compromising the fit at well constrained sites. In addition to recommending that the LVT depth be extended to deeper depths, Hu et al. (2022b) suggested that their results be checked using additional ground motion metrics and additional seismic recordings from earthquakes located at different azimuth relative to the basins, which has been shown to affect amplification patterns (Olsen, 2000). Moreover, they observed some indication

of spatial variability for the tapering depth of the LVT that, if accounted for, may further improve their method.

Here, we further analyze the near-surface velocities in the SCEC CVM-S4.26.M01 both inside and outside the greater Los Angeles basins, with the goal of improving the fit between synthetic and observed seismic data. Toward this goal, we estimate a spatially-variable distribution of optimal tapering depths through simulation of 0-1 Hz physics-based wave propagation for 7 well-recorded events with magnitudes between  $M_w 4.4$  and 5.1 (see Fig. 1).



**Figure 1.** Simulation domain. Locations of 7 events included in this study are shown by stars.

### Numerical Method

We carried out 0 - 1 Hz 3D numerical wave propagation simulations using the 4th-order accurate staggered-grid finite-difference numerical simulation code Anelastic Wave Propagation (AWP)-ODC which is GPU-enabled and highly scalable (Cui et al., 2013). Our simulation domain (200 km by 130 km) is shown in Fig. 1, covering the entire Los Angeles basin (LAB), San Fernando basin (SFB), San Gabriel basin (SGB), Chino basin (CB), and San Bernardino basin (SBB), as well as the surrounding areas. The model domain extends to a depth of 100 km. The material mesh is decomposed into three blocks with a factor-of-three coarsening using a discontinuous mesh (Nie et al., 2017). The curvilinear grid is utilized in the top mesh block (O'Reilly et al., 2021) to account for the topography defined from the 30 m-resolution digital elevation model by U.S. Geological Survey (2020). The minimum  $V_s$  included in the simulations is 180 m/s. See Table 1 for more details about the numerical simulations.

**Table 1.** Simulation parameters

|                              |                                                                                                                           |
|------------------------------|---------------------------------------------------------------------------------------------------------------------------|
| <b>Model dimensions</b>      | Top mesh: 6696x4320x416<br>Middle mesh: 2232x1440x480<br>Bottom mesh: 744x480x160                                         |
| <b>Grid spacings</b>         | 30 m: Free surface to 12.42 km depth<br>90 m: 12.21 km depth to 55.32 km depth<br>270 m: 54.69 km depth to 96.62 km depth |
| <b>Minimum Vs</b>            | 180 m/s                                                                                                                   |
| <b>Maximum frequency</b>     | 1 Hz                                                                                                                      |
| <b>Timestep</b>              | 0.0015 s                                                                                                                  |
| <b>Simulated time length</b> | 100 s                                                                                                                     |

### Velocity and Anelastic Attenuation Model

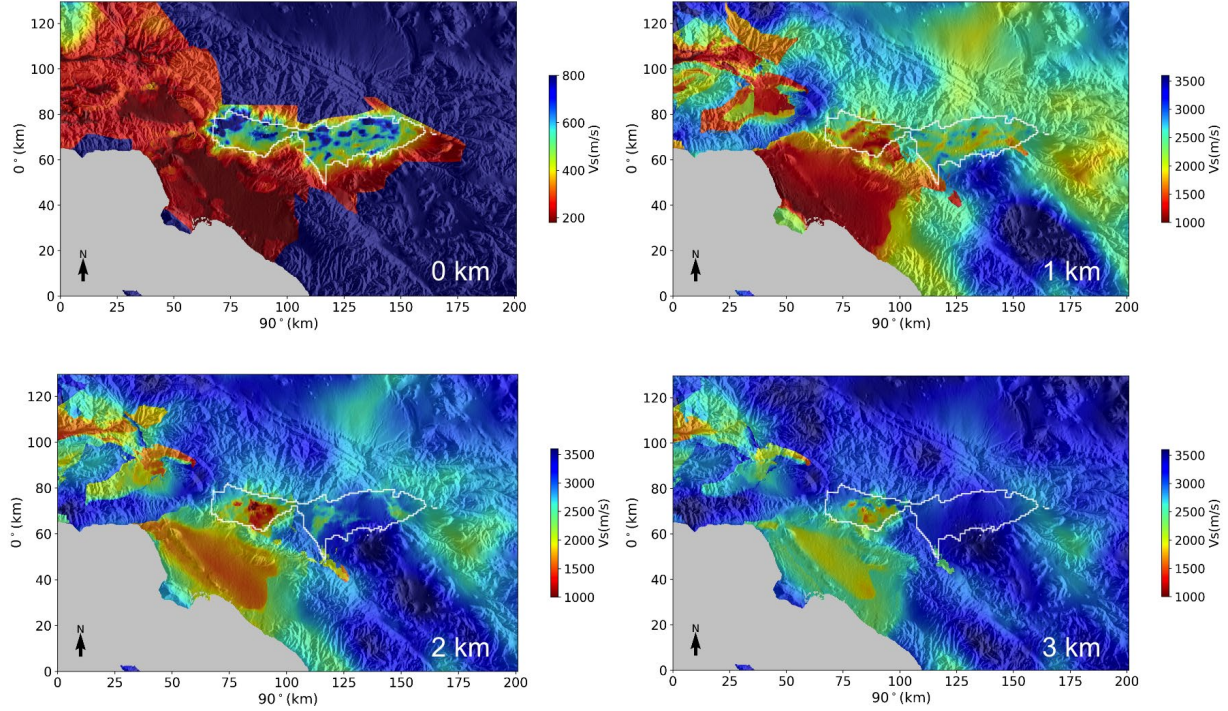
We start out with the SCEC CVM-S4.26.M01 (hereafter referred to as CVMSI) which was derived from full waveform tomography (Lee et al., 2014) and populated with geotechnical constraints inside the major basins (Small et al., 2017). Olsen and Yeh (2023) showed that a recent update on the velocity structures of San Gabriel, Chino, and San Bernardino basins (abbreviated here as SGSB) using ambient noise tomography (Li et al., 2023) incorporated into CVMSI improved the fit to the recordings for 7 small earthquakes. We therefore implement the SGSB model in CVMSI in the following tests. Fig. 2 shows horizontal slices of CVMSI including the SGSB. We adopt the same anelastic attenuation model as Hu et al. (2022b), that is,  $Q_S=0.1V_S$  and  $Q_P=2Q_S$ .

### LVT

We implement the LVT feature in CVMSI outside the major basins and the immediate surroundings in order to improve the description of shallow seismic structure using the formulation proposed by Ely et al. (2010) and considering the local  $V_{S30}$  information from Thompson (2018). The taper function proposed by Ely et al. (2010) is given by

$$\begin{aligned}
 z &= z'/z_T \\
 f(z) &= z + b(z - z^2) \\
 g(z) &= a - az + c(z^2 + 2\sqrt{z} - 3z) \\
 V_S(z) &= f(z)V_{ST} + g(z)V_{S30} \\
 V_P(x) &= f(z)V_{PT} + g(z)P(V_{S30}) \\
 \rho(z) &= R(V_P),
 \end{aligned} \tag{1}$$

where  $z$  is a normalized depth,  $z'$  is the depth,  $z_T$  is the tapering depth, and  $V_P$  and  $V_S$  are computed using linear combinations of the  $f(z)$  and  $g(z)$  functions along with  $V_{PT}$  and  $V_{ST}$  which are  $V_P$  and  $V_S$  in the original model at  $z_T$ , respectively.  $P$  and  $R$  are the  $V_P$  scaling law from Brocher (2005) and the Nafe-Drake law, respectively. For simplicity, we follow the approach of Hu et al. (2022b) using the coefficients  $a = 1/2$ ,  $b = 2/3$ , and  $c = 3/2$  in Eq. (1) consistent with Ely et al. (2010). In contrast to the approach of Hu et al. (2022a,b) in which the seismic properties computed by Eq. (1) only replace the original model values when the computed  $V_s$  is lower in order to preserve the lower  $V_s$ , we unconditionally overwrite the original model with the LVT.



**Figure 2.** Horizontal slices of shear wave speeds at different depths extracted from the combined model (CVM-S/SGSB). The white polygon outlines the surface projection of the imaging domain of Li et al. (2023). Depth of each slice is shown at the lower right corner.

In this formulation, the resulting  $V_{s30}$  values in the model modified by the LVT may deviate slightly from the actual  $V_{s30}$  value. To mitigate this problem we replaced the  $V_s$  profile in the top 30 m with the re-scaled  $V_s$  profile for generic rock sites from Boore and Joyner (1997), defined as

$$V_{s_c}(z) = (V_{s_{BJ1997}}(z)/617)V_{s30}, z < 30m,$$

where  $V_{s_c}$  is the corrected  $V_s$  profile,  $V_{s_{BJ1997}}$  is the  $V_s$  profile for generic rock sites from Boore and Joyner (1997), and  $V_{s30}$  is the targeted  $V_{s30}$  value to be matched. The scale factor (617 m/s) used here is the  $V_{s30}$  of the  $V_{s_{BJ1997}}$  profile, which is given by

$$V_{s_{BJ1997}}(z) = 245m/s, z \leq 1m$$

$$V_{s_{BJ1997}}(z) = 2206(z/1000)^{0.272} \text{ m/s}, 1 < z \leq 30 \text{ m},$$

where the depth is in meters. To avoid creating a velocity contrast at 30 m depth, we linearly transition  $V_{s_c}$  at a depth of 30 m to the existing  $V_s$  value at a depth of 60 m in the model.

### Goodness-of-fit Measure

We quantify the mismatch between model and data using metrics in both time and frequency domains. In the time domain, we use the cumulative absolute velocity (CAV) given by

$$CAV = \sum_{i=1}^{i=N} |v_i| dt,$$

where  $v_i$  is the amplitude of the bandpass-filtered velocity waveform at time step  $i$ , and  $dt$  is the time step. It provides a single-value quantification of the cumulative waveform energy within a given frequency band. We quantify the bias between the observed and simulated CAV values as

$$CAV_{bias} = \log_{10} \frac{CAV_{model}}{CAV_{data}}. \quad (2)$$

In the frequency domain, we define the Fourier amplitude spectral (FAS) bias as

$$FAS_{bias}(f) = \log_{10} \frac{FAS_{model}(f)}{FAS_{data}(f)}, \quad (3)$$

where  $FAS_{model}(f)$  and  $FAS_{data}(f)$  are smoothed Fourier amplitude spectra of the simulated and observed waveforms, respectively. We use Konno-Ohmachi smoothing (Konno and Ohmachi, 1998) with a bandwidth of 40 to suppress excessive fluctuations in the FAS. Unlike the single-value CAV bias measure, the FAS bias given by Eq. (3) quantifies model bias as a function of frequency. To incorporate information from multiple events, we average FAS bias and CAV bias measures computed for all available events to form event-averaged values for sites that record two or more events. Sites that only record a single event are excluded for further analysis. The event-averaged FAS bias from Eq. (3) can be further simplified to a single-value measure to summarize the performance of the simulation over a specified frequency band by computing the mean value, that is

$$r = \frac{\sum_{i=1}^{i=N_f} FAS_{bias}[i]}{N_f}, \quad (4)$$

and alternatively, in an absolute-valued form

$$J = \frac{\sum_{i=1}^{i=N_f} |FAS_{bias}[i]|}{N_f}, \quad (5)$$

where  $N_f$  is the number of frequency points at which the FAS bias values are calculated and  $FAS_{bias}[i]$  is the FAS bias at the  $i$ -th frequency point.

To combine CAV and FAS bias measures for estimation of the optimal  $zT$ , we define a comprehensive model bias, given by

$$\xi = \frac{|CAV_{bias}^E| + |CAV_{bias}^N| + |CAV_{bias}^Z| + J^E + J^N + J^Z}{6}, \quad (6)$$

where the superscripts denote the components of the ground motions,  $CAV_{bias}$  values are computed using Eq. (2), and  $J$  values are given by Eq. (5). The use of absolute values in Eq. (6) is meant to provide a robust quantification of the goodness-of-fit of the simulations to the recordings without being biased by the sign of values.

### Optimization of LVT Model Parameters

Based on the surface  $V_s$  in CVMSI, Hu et al. (2022a,b) showed that implementation of the LVT is an effective way to remedy underprediction at type B sites for the 2014  $M_w 5.1$  La Habra event. Expanding the analysis to the 7 events included here confirm the underprediction at type B sites found by Hu et al. (2022a,b). Figs. 3 and 4 demonstrate that the models without an LVT underpredict the ground motions at type B sites while an LVT with a  $zT$  of 600 m - 1200 m effectively improves the underprediction, which is consistent with the findings of Hu et al. (2022a,b). Fig. 5 shows that the underprediction at type B sites is alleviated as  $zT$  increases, while clearly demonstrating spatial variation of the  $zT$ . With the LVT implemented both inside and outside the basins, Fig. 5 also suggests that some type A sites may be improved by an LVT.

Here, we expand the application of the LVT further to both site types A and B and allow for spatially-varying  $zT$ . We carefully examine the significance of the improvement made by the LVT by statistical testing. This approach enables us to address the underprediction at type B sites and to revise the velocity structures inside the major basins. Assuming that the response of the LVT is caused by purely local site effects, we study the seismic response of different sites independently. One major advantage of this approach is that it makes a high-dimensional nonlinear optimization problem solvable by carrying out a few numerical simulations for all 7 events with various (constant)  $zT$  values and estimating the optimal  $zT$  at each site individually. The optimization is done using grid-search, where the optimal  $zT$  value is searched among depths of 0 m (no LVT), 90 m, 150 m, 200 m, 250 m, 300 m, 450 m, 600 m, 750 m, 900 m, 1200 m, 1500 m, 1800 m, and 2100 m, followed by statistical testing to evaluate whether replacing the original model with an LVT is justified for the type A sites. The optimal  $zT$  is determined by minimizing  $\xi$  given by Eq. (6).

### Evaluating the Significance of LVT

The implementation of LVT is justified where ground motions are underestimated due to unrealistically high shear wave speeds near the surface (i.e., type B sites with surface  $V_s \geq 1000$  m/s). If the model bias is improved by using an LVT at type A sites ( $V_s < 1000$  m/s), on the other hand, we test the significance of the improvement since the generally well-constrained basin structures included in the reference model (CVMSI+SGSB) will be replaced. This procedure ensures that the LVT is implemented at type A sites only when substantial

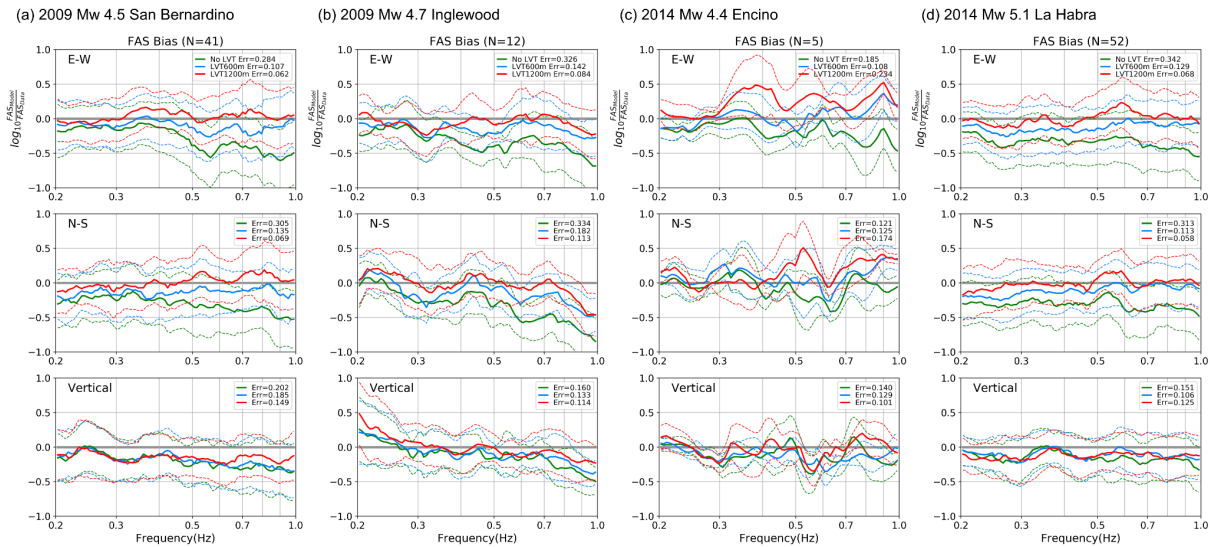
improvement of the FAS bias is obtained. By examining the absolute value of the FAS bias on all three components, we use the two-sample t-test (Snedecor and Cochran, 1989) to determine if the magnitude of FAS bias for the model including LVT is significantly smaller than that for the model without LVT. We form the statistical test by defining the null ( $H_0$ ) and alternative ( $H_a$ ) hypotheses,

$$\begin{aligned} H_0 &: \mu_1 = \mu_0 \\ H_a &: \mu_1 < \mu_0. \end{aligned}$$

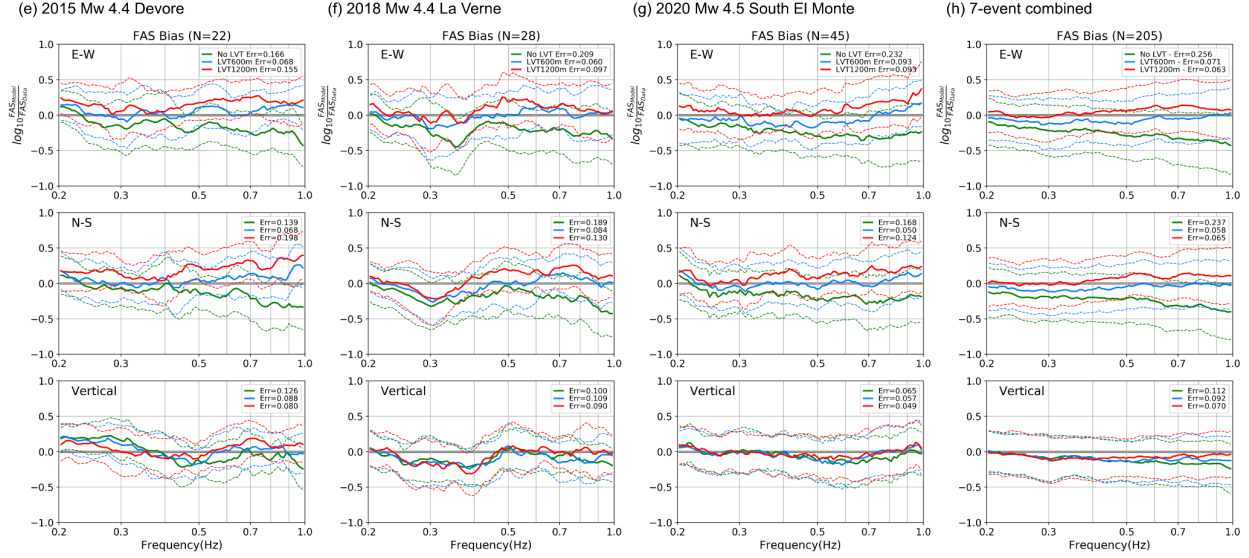
However, the FAS bias function is correlated with neighboring frequencies due to smoothing, which violates the assumption of independent samples. To meet this assumption, we downsampled the FAS bias prior to computing the test statistics based on analysis of the FAS bias autocorrelation function. The test statistic is computed as

$$T = \frac{\bar{Y}_1 - \bar{Y}_0}{\sqrt{s_1^2/N_1 + s_0^2/N_0}},$$

where  $\bar{Y}_0$  and  $\bar{Y}_1$  are mean values of the (downsampled) absolute event-averaged three component FAS bias for the model without LVT (CVMSI+SGSB) and the model including LVT with the optimal zT, respectively.  $s_1$  and  $s_0$  are standard deviations, and  $N_1$  and  $N_0$  are sample sizes after the downsampling. The model including LVT is rejected and the reference model (CVMSI+SGSB) is restored if the test fails to reject the null hypothesis ( $H_0$ ) at 95% confidence level.

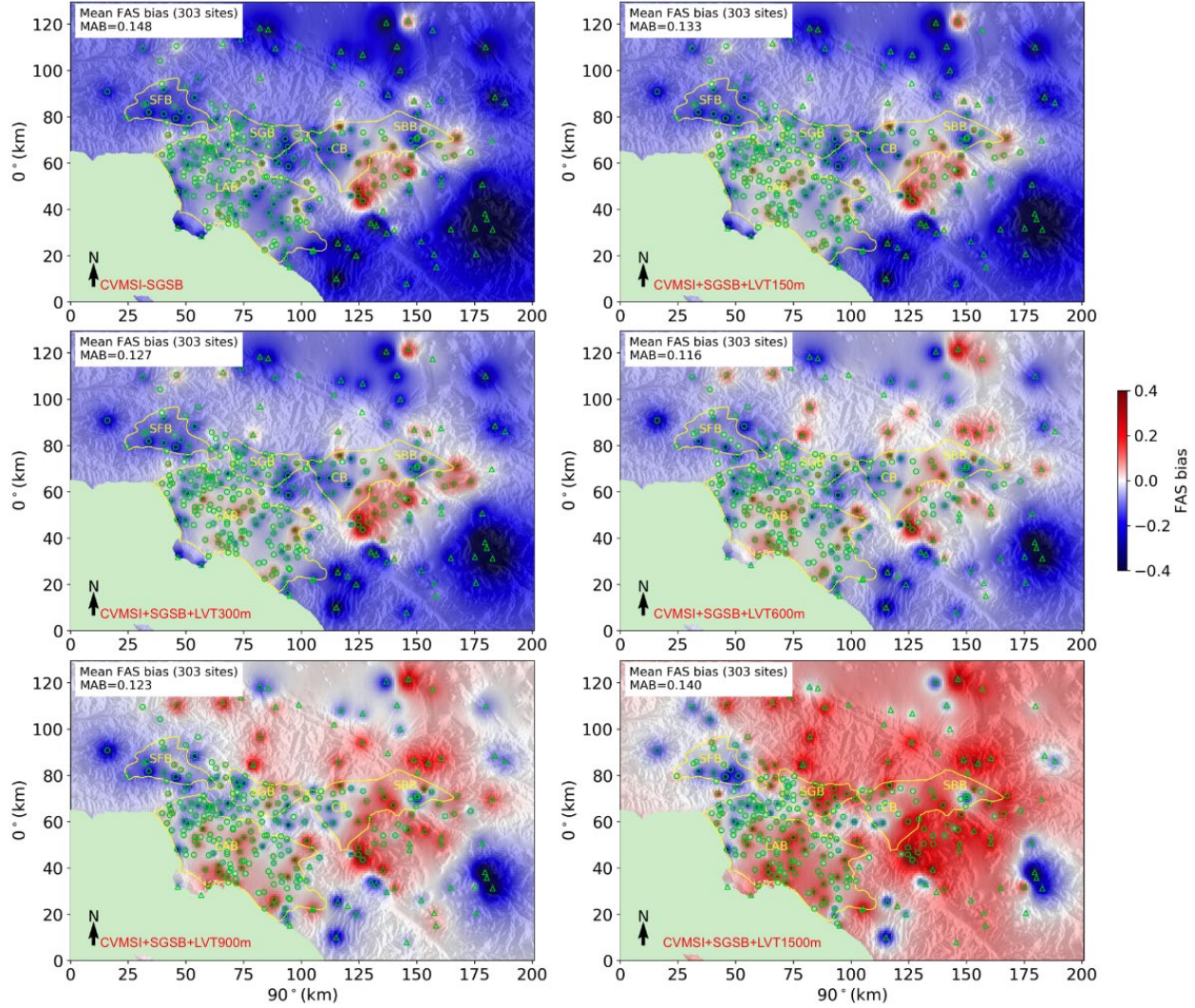


**Figure 3.** FAS bias between seismic recordings and simulations computed at type B sites using Eq. (3) for four events, including the model with no LVT (green), and two models with LVT using constant zT of 600 m (blue) and 1200 m (red). Thick and dashed curves depict mean and standard deviation, respectively.



**Figure 4.** Same as Fig. 3, but for three other events (e-g), as well as the average over all 7 events (h).

The resolution of tomographically-derived velocity models generally improves at depths larger than about 1-2 km, suggesting that LVTs may be needed primarily in the shallow crust. For this reason, another t-test is carried out for both site types when  $zT$  is estimated to be deeper than 1,200 m, in addition to the statistical test of significance of LVT described above for the type A sites. Following the same test procedure, the best-fit model (one that gives the lowest  $\xi$ ) with a  $zT$  value deeper than 1,200 m is tested against the best-fit model found with a  $zT$  value shallower than 1200 m. If the test rejects the presence of a LVT with a  $zT$  value deeper than 1,200 m, the best-fit model found with  $zT < 1,200$  m is taken as the final estimate. This test confirms whether a deeper  $zT$  provides meaningful improvement on the FAS bias.

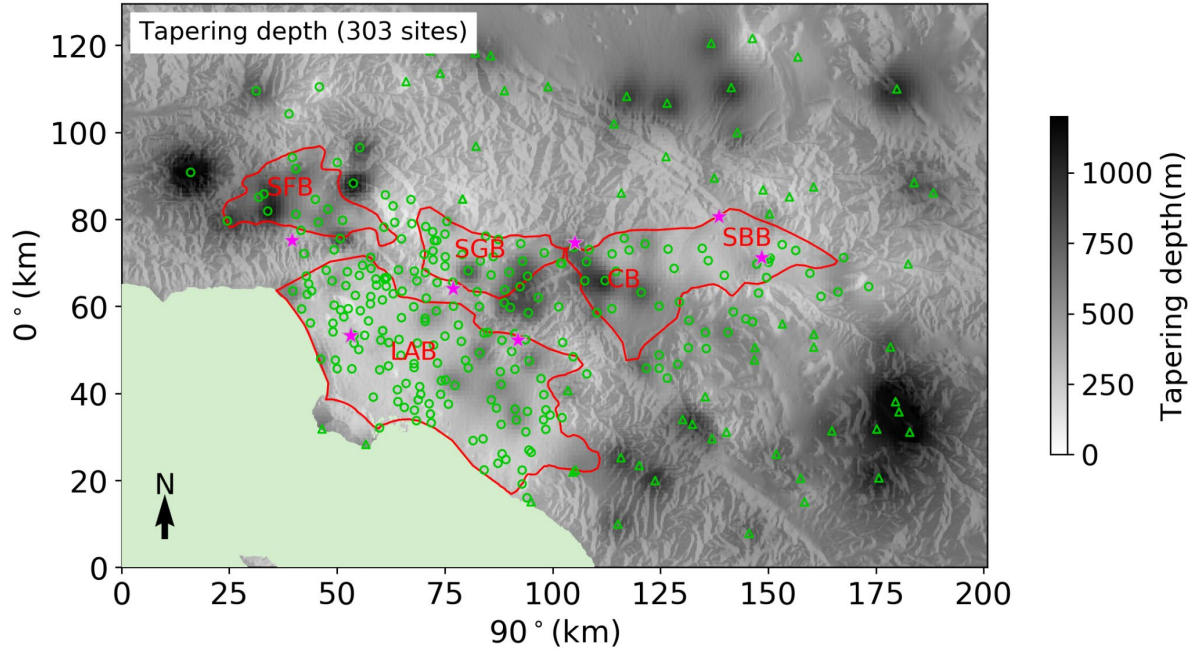


**Figure 5.** Interpolated map of event-averaged mean FAS bias,  $r$ , computed using Eq. (4) for different  $zT$  values (labeled near the lower left corner). Circles and triangles are locations of type A (surface  $V_s < 1000$  m/s) and type B (surface  $V_s \geq 1000$  m/s) sites, respectively. The polygons depict approximate outlines of the San Fernando basin (SFB), Los Angeles basin (LAB), San Gabriel basin (SGB), Chino basin (CB), and San Bernardino basin (SBB). The ‘MAB’ value shown in the text box is the mean absolute bias, averaging the  $|r|$  values from all sites.

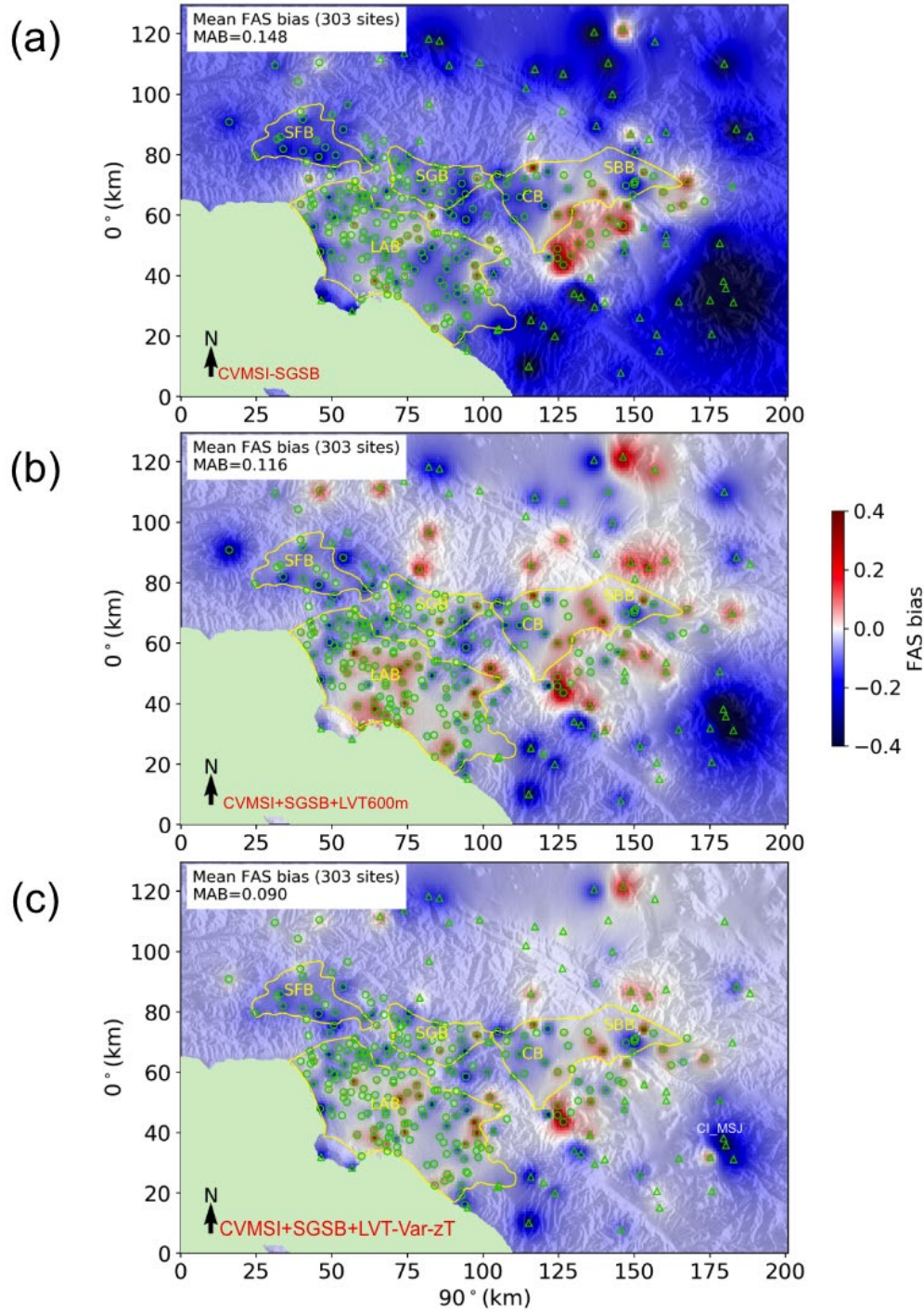
### Spatially-varying Tapering Depth

Fig. 6 shows the interpolated map of estimated  $zT$  within the simulation domain obtained by grid search along with the statistical test. Our analysis generally suggests very small  $zT$  values inside the major basins, except near the edges of the basins (e.g., western edge of the Chino basin and southern edge of the San Gabriel basin), where LVTs with thicknesses of 600 m or larger improve the waveform fit significantly. Large spatial variation is obtained in the  $zT$  values for type B sites (surface  $V_s > 1000$  m/s), from 300 - 450 m in the San Gabriel Mountains to +900 m to the north into the high deserts, as well as very deep  $zT$  (2100 m) to the west of Mt. San Jacinto

near the southeast corner of the domain. The mean  $zT$  value for type B sites is approximately 680 m, close to the optimal constant  $zT$  of 600 m. When implementing the interpolated  $zT$  map (Fig. 7), we obtain a reduction of FAS bias by 22% on average (Fig. 7c) as compared to that of the constant  $zT=600$  m model (Fig. 7b).



**Figure 6.** Optimal LVT tapering depth from minimizing the event-averaged model bias  $\xi$  (Eq. (6)) over the 7 events (stars). The polygons depict approximate outlines of the San Fernando basin (SFB), Los Angeles basin (LAB), San Gabriel basin (SGB), Chino basin (CB), and San Bernardino basin (SBB).



**Figure 7.** Interpolated map of event-averaged mean FAS bias ( $r$ , see Eq. (4)) for (a) CVMSI without LVT, (b) CVMSI including LVT with the best-fit constant  $zT = 600$  m, and (c) CVMSI including LVT with spatially-varying  $zT$  (shown in Fig. 6). The polygons depict approximate outlines of the San Fernando basin (SFB), Los Angeles basin (LAB), San Gabriel basin (SGB), Chino basin (CB), and San Bernardino basin (SBB). The ‘MAB’ value shown in the text box is the mean absolute bias, averaging the  $|r|$  values from all sites.

### Discussion and Conclusions

We have used 0 - 1 Hz 3D physics-based wave propagation simulations in the greater Los Angeles area for 7  $M_w$  4.4-5.1 events in CVMSI including an LVT with various constant tapering depths to estimate the  $zT$  value at each recording site providing an optimal GOF to data. The resulting spatial distribution of  $zT$  shows strong spatial variation, where  $zT$  ranges from nearly 0 m (no LVT) inside the major basins to more than 1500 m in some areas outside the basins. If confirmed by a student t-test at basin sites and for depths larger than 1,200 m, the irregularly spaced distribution of  $zT$  was then interpolated into a regular grid and implemented in the ground motion simulations. The average reduction of FAS bias from simulations of the 7 events in CVMSI with the optimized, spatially-variable LVT  $zT$  distribution is 22% relative to that from the model with constant  $zT = 600$  m, and 39% relative to that from the model without an LVT.

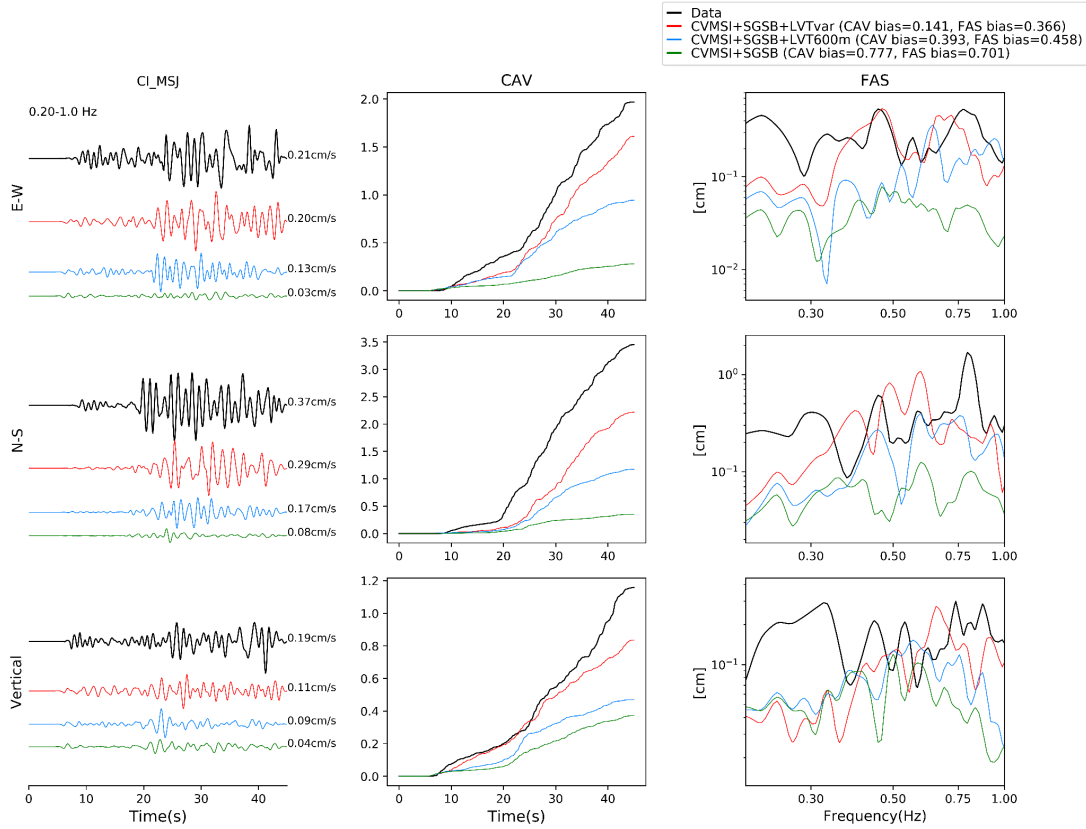
We obtain the largest reduction in misfit between simulations and data outside the basins. At sites where the optimal LVT depth is shallow (i.e., a few hundred meters), the need for the taper may be due to the lack of a weathering layer in the CVMSI. However, the need for much larger  $zT$  values, in some cases larger than 1,200 m, is likely caused by low resolution of the tomographically-derived velocity model, in particular in areas where the station density is low. On the other hand, our optimal tapering depths at type A sites (surface  $V_s < 1,000$  m/s) are generally small, suggesting (as expected) that the near-surface velocity structure for the most parts is well-constrained inside the major basins in Southern California (Lee et al., 2014; Small et al., 2017; Li et al., 2023).

In the simulations including the optimized LVT depth distribution, one can still notice areas with larger bias values, such as the underpredicted area near the western foothills of Mt. San Jacinto (~180 km along E-W, ~35 km along N-S in Fig. 7c), or the overpredictions in the Corona area (~130 km along E-W, ~40 km along N-S in Fig. 7c). In these areas, implementation of an LVT fails to improve the accuracy of the reference model, suggesting that a more sophisticated refinement needs to be implemented. Nevertheless, substantial improvement in the fit between simulations and recordings in both time and frequency domains are obtained at some of these sites using the optimized spatially-varying  $zT$  model. As an example, Fig. 8 shows the improvement in waveform, CAV and spectral amplitude at CI\_MSJ (see Fig. 7c for location) for the optimal  $zT$  value of 2,100 m (largest value tested).

The accuracy of the shallow structure is critical for seismic hazard assessment. As conventional seismic hazard analyses are typically accomplished using ground motion models (GMMs), we compare the spectral acceleration values (RotD50 with 5% damping) predicted by our numerical simulations and by four NGA-West2 GMMs (ASK14, Abrahamson et al., 2014; BSSA14, Boore et al., 2014; CB14, Campbell and Bozorgnia, 2014; and CY14, Chiou and Youngs, 2014) to those calculated from the seismic recordings. We use  $V_{s30}$  values from Thompson (2018), and  $Z1.0$  and  $Z2.5$  values extracted from CVMSI for site corrections in GMMs, when applicable. We summarize the misfit between the observations and predictions (physics-based simulations and GMMs) for each model by computing the mean absolute log ratio, given by

$$\frac{\sum_{i=1}^{i=N} \left| \log_{10} \frac{SA_{pred,i}}{SA_{obs,i}} \right|}{N}, \quad (7)$$

where  $N$  is the number of stations, and  $SA_{pred,i}$  and  $SA_{obs,i}$  are predicted and observed 5% damped rotD50 spectral acceleration values at the  $i$ -th station at a given period. Table 2 lists the misfit values computed using Eq. (7) for the 2014  $M_w 5.1$  La Habra event for periods of 2, 3, and 4 s, along with the mean misfit value across the three periods. Our physics-based simulations obtained using the optimized spatially-varying LVT result in smaller misfit values as compared to those from the GMMs. CB14 provides the best-fit GMM results for the 2014  $M_w 5.1$  La Habra event and closest to those from the physics-based simulations.



**Figure 8.** Comparison of waveforms, cumulative absolute velocity (CAV), and FAS at CI\_MSJ (see Fig. 7c for location) between seismic recordings and simulations with (red) the model using the optimized spatially-varying LVT  $zT$ , (blue) the model including LVT with constant  $zT = 600$  m, and (green) the reference CVMSI model (no LVT included).

**Table 2.** Misfit between predicted and observed rotD50 at three different periods for the 2014 Mw5.1 La Habra event, computed using Eq. (7).

|                | Simulation<br>No LVT | Simulation<br>LVT<br>zT = 600 m | Simulation<br>LVT<br>Variable zT | GMM<br>ASK14 | GMM<br>BSSA14 | GMM<br>CB14 | GMM<br>CY14 |
|----------------|----------------------|---------------------------------|----------------------------------|--------------|---------------|-------------|-------------|
| T = 2 s        | 0.205                | 0.200                           | 0.177                            | 0.200        | 0.312         | 0.184       | 0.232       |
| T = 3 s        | 0.205                | 0.190                           | 0.175                            | 0.214        | 0.308         | 0.186       | 0.250       |
| T = 4 s        | 0.199                | 0.176                           | 0.165                            | 0.228        | 0.291         | 0.192       | 0.277       |
| <b>Average</b> | 0.203                | 0.189                           | 0.172                            | 0.214        | 0.304         | 0.187       | 0.253       |

Fig. 9 compares the bias of predicted spectral acceleration at a period of 2 s as a function of rupture distance in a form of

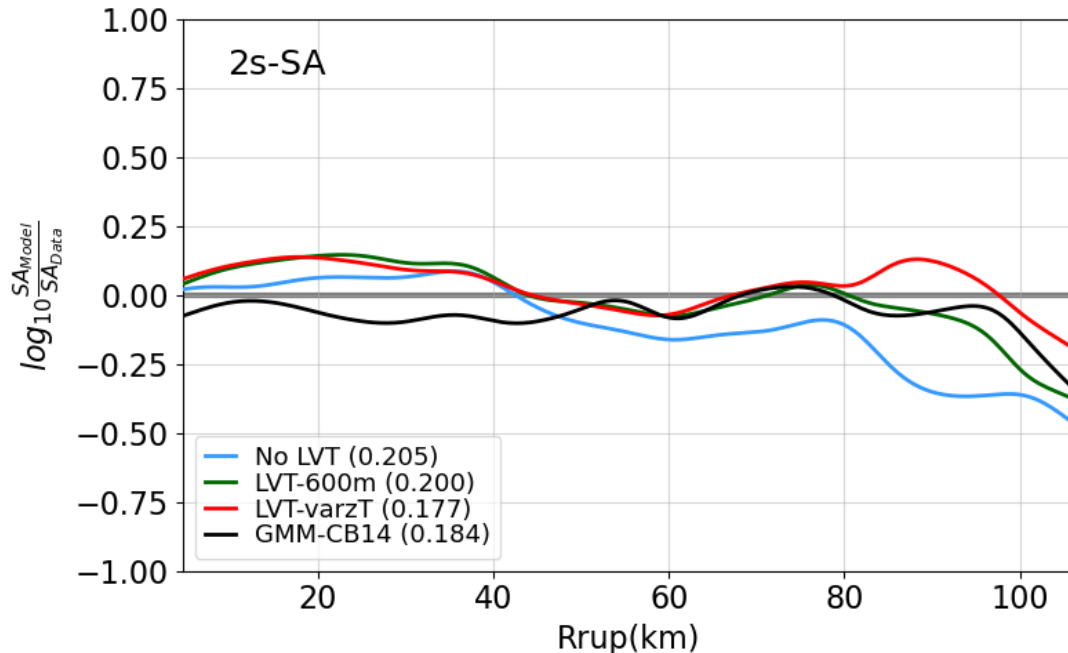
$$\log_{10} \frac{SA_{pred}}{SA_{obs}},$$

where  $SA_{pred}$  and  $SA_{obs}$  are predicted (simulations and CB14) and observed values, respectively. The comparison shows that the optimal LVT model generates generally unbiased results throughout all distance ranges, whereas the simulation with a constant zT of 600 m underpredicts at distances larger than 80 km from the source. This finding suggests that the spatially-varying zT model has a larger impact as the distance increases, particularly for sites outside the basins. The physics-based simulation with no LVT (CVMSI+SGSB only) underpredicts even more, starting at distances of 40 km away from the source. CB14 underpredicts the values from the recordings at distances larger than 80 km, similar to the accuracy of the physics-based simulation with a constant LVT (zT=600m). This result suggests that the generic site corrections for GMM can be further improved for sites outside the major basins.

In summary, our optimized, spatially-varying LVT model substantially improves the accuracy of 0 - 1 Hz predicted ground motions in Southern California compared to the CVMSI reference model (no LVT) as well as the CVMSI with a constant zT. Future work should attempt to calibrate the optimal LVT depths for higher frequencies, as well as improve the formulation of the near-surface structure in order to approximate more complicated shallow seismic structures, including constant-velocity layers, to further improve the ground motion predictions.

### Acknowledgements

This research was supported by the California Geological Survey (Award #1022-003). Computational resources for the simulations were obtained at the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract Number DE-AC05-00OR22725.



**Figure 9.** Smoothed log10-based bias values between rotD50 spectral accelerations at a period of 2s for models and recordings as a function of rupture distance using physics-based simulations (blue) with no LVT, (green) an LVT with constant  $zT = 600$  m, and (red) LVT with spatially-varying  $zT$ , as well as those from (black) CB14.

## References

- Abrahamson, N. A., W. J. Silva, and R. Kamai, 2014, Summary of the ASK14 Ground Motion Relation for Active Crustal Regions, *Earthq. Spectra* 30, no. 3, 1025–1055, doi: 10.1193/070913EQS198M.
- Ajala, R., and P. Persaud, 2021, Effect of Merging Multiscale Models on Seismic Wavefield Predictions Near the Southern San Andreas Fault, *J. Geophys. Res. Solid Earth* 126, no. 10, e2021JB021915, doi: 10.1029/2021JB021915.
- Boore, D. M., 2010, Orientation-Independent, Nongeometric-Mean Measures of Seismic Intensity from Two Horizontal Components of Motion, *Bull. Seismol. Soc. Am.* 100, no. 4, 1830–1835, doi: 10.1785/0120090400.
- Boore, D. M., and W. B. Joyner, 1997, Site amplifications for generic rock sites, *Bull. Seismol. Soc. Am.* 87, no. 2, 327–341, doi: 10.1785/BSSA0870020327.
- Boore, D. M., J. P. Stewart, E. Seyhan, and G. M. Atkinson, 2014, NGA-West2 Equations for Predicting PGA, PGV, and 5% Damped PSA for Shallow Crustal Earthquakes, *Earthq. Spectra* 30, no. 3, 1057–1085, doi: 10.1193/070113EQS184M.
- Bozorgnia, Y. et al., 2014, NGA-West2 Research Project, *Earthq. Spectra* 30, no. 3, 973–987, doi: 10.1193/072113EQS209M.
- Brocher, T. M., 2005, Empirical Relations between Elastic Wavespeeds and Density in the Earth's Crust, *Bull. Seismol. Soc. Am.* 95, no. 6, 2081–2092, doi: 10.1785/0120050077.
- Campbell, K. W., and Y. Bozorgnia, 2014, NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra, *Earthq. Spectra* 30, no. 3, 1087–1115, doi: 10.1193/062913EQS175M.
- Chiou, B. S.-J., and R. R. Youngs, 2014, Update of the Chiou and Youngs NGA Model for the

- Average Horizontal Component of Peak Ground Motion and Response Spectra, *Earthq. Spectra* 30, no. 3, 1117–1153, doi: 10.1193/072813EQS219M.
- Cui, Y. et al., 2013, Physics-based seismic hazard analysis on petascale heterogeneous supercomputers, 1–12.
- Ely, G. P., T. Jordan, P. Small, and P. J. Maechling, 2010, A VS30-derived nearsurface seismic velocity model, American Geophysical Union, Fall Meeting 2010, San Francisco, CA.
- Hu, Z., K. B. Olsen, and S. M. Day, 2022a, 0–5 Hz deterministic 3-D ground motion simulations for the 2014 La Habra, California, *Earthquake, Geophys. J. Int.* 230, no. 3, 2162–2182, doi: 10.1093/gji/ggac174.
- Hu, Z., K. B. Olsen, and S. M. Day, 2022b, Calibration of the near-surface seismic structure in the SCEC community velocity model version 4, *Geophys. J. Int.* 230, no. 3, 2183–2198, doi: 10.1093/gji/ggac175.
- Konno, K., and T. Ohmachi, 1998, Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor, *Bull. Seismol. Soc. Am.* 88, no. 1, 228–241, doi: 10.1785/BSSA0880010228.
- Lee, E.-J., P. Chen, T. H. Jordan, P. Maechling, M. Denolle, and G. C. Beroza, 2014, Full-3-D tomography for crustal structure in Southern California based on the scattering-integral and the adjoint-wavefield methods, *J. Geophys. Res. Solid Earth* 119, no. 8, 6421–6451, doi: 10.1002/2014JB011346.
- Li, Y., V. Villa, R. W. Clayton, and P. Persaud, 2023, Shear Wave Velocities in the San Gabriel and San Bernardino Basins, California, *J. Geophys. Res. Solid Earth* 128, no. 7, e2023JB026488, doi: 10.1029/2023JB026488.
- Nie, S., Y. Wang, K. B. Olsen, and S. M. Day, 2017, Fourth-Order Staggered-Grid Finite-Difference Seismic Wavefield Estimation Using a Discontinuous Mesh Interface (WEDMI) Fourth-Order Staggered-Grid Finite-Difference Seismic WEDMI, *Bull. Seismol. Soc. Am.* 107, no. 5, 2183–2193, doi: 10.1785/0120170077.
- Olsen, K. B., 2000, Site Amplification in the Los Angeles Basin from Three-Dimensional Modeling of Ground Motion, *Bull. Seismol. Soc. Am.* 90, no. 6B, S77–S94, doi: 10.1785/0120000506.
- Olsen, K.B. and T.-Y. Yeh, 2023, Calibration of the near-surface seismic structure in the SCEC community velocity model version S, in: *Procs. SMIP24 seminar on utilization of strong-motion data*, California Department of Conservation, California Geological Survey, Oct. 19, 2023, pp 80-94.
- O'Reilly, O., T. Yeh, K. B. Olsen, Z. Hu, A. Breuer, D. Roten, and C. A. Goulet, 2021, A High-Order Finite-Difference Method on Staggered Curvilinear Grids for Seismic Wave Propagation Applications with Topography, *Bull. Seismol. Soc. Am.* 112, no. 1, 3–22, doi: 10.1785/0120210096.
- Small, P., D. Gill, P. J. Maechling, R. Taborda, S. Callaghan, T. H. Jordan, K. B. Olsen, G. P. Ely, and C. Goulet, 2017, The SCEC Unified Community Velocity Model Software Framework, *Seismol. Res. Lett.* 88, no. 6, 1539–1552, doi: 10.1785/0220170082.
- Snedecor, G. W., and W. G. Cochran, 1989, Statistical Methods, eight edition, Iowa State Univ. Press Ames Iowa, 1191, no. 2.
- Thompson, E. M., 2018, An Updated Vs30 Map for California with Geologic and Topographic Constraints, US Geol. Surv. Data Release, doi: 10.5066/F7JQ108S.
- U.S. Geological Survey, 2020, 3D Elevation Program 1-Meter Resolution Digital Elevation Model (Published 20200330).