EPOXI: Comet 103P/Hartley 2 observations from a worldwide campaign

Robert R. Howell
University of Wyoming, rhowell@uwyo.edu

Follow this and additional works at: http://repository.uwyo.edu/geology_facpub

Publication Information
Seismological Research Letters

This copy is for distribution only by the authors of the article and their institutions in accordance with the Open Access Policy of the Seismological Society of America.

For more information see the publications section of the SSA website at www.seismosoc.org
Testing Waveform Predictions of 3D Velocity Models against Two Recent Los Angeles Earthquakes

by En-Jui Lee, Po Chen, and Thomas H. Jordan

INTRODUCTION

Nearly half of the national seismic risk is located in Southern California, and about one-fourth is concentrated in Los Angeles County alone (Federal Emergency Management Agency [FEMA], 2000). To assess the seismic hazards that drive this risk, we must forecast the strong ground motions that are likely to be produced by large fault ruptures. The standard probabilistic seismic-hazard model of California calculates shaking intensities according to an ensemble of ground-motion predictions equations (Petersen et al., 2008). These empirical equations have a high aleatory variability, primarily because they do not model much of the ground-motion variance caused by 3D crustal heterogeneities (Strasser and Bommer, 2009).

One approach for improving hazard estimates is to simulate the wave propagation through realistic 3D crustal models (Olsen et al., 1995, 2006; Komatitsch et al., 2004; Graves et al., 2008). If seismograms can be simulated for sufficiently large samples of probable ruptures, then hazard curves of exceedance probability versus seismic intensity can be constructed from the calculated site response. Graves et al. (2010) have developed a software platform, CyberShake, that combines seismic reciprocity with highly optimized anelastic wave propagation codes to reduce the time of this calculation to manageable levels. Low-frequency (<0.5 Hz) CyberShake hazard models, each comprising about 240 million synthetic seismograms, have been computed for the Los Angeles region (Graves et al., 2010; Wang et al., 2013; Wang and Jordan, 2014).

The ability to simulate seismograms depends on having accurate models of 3D crustal structure. Crustal models that specify the lateral and depth variations of seismic velocities have been developed for Southern California by synthesizing constraints from geologic studies, well logs, seismic reflection and refraction surveys, and earthquake tomography (Hauksson, 2000; Magistrale et al., 2000; Kohler et al., 2003; Süss and Shaw, 2003). To support simulation-based hazard analysis, the Southern California Earthquake Center (SCEC) has released iterative refinements to these 3D structures in two series of community velocity models (CVM), CVM-S (Magistrale et al., 1995, 2006), and CVM-H (Süss and Shaw, 2003). Recent improvements have been made using the nonlinear techniques of full-3D tomography, to the CVM-H series by Tape et al. (2010) and to the CVM-S series by Chen et al. (2007) and Lee et al. (2014).

These models demonstrate the extreme heterogeneity of the southern California crust, which contains a number of fault-bounded blocks, some with deep sedimentary basins (Plesch et al., 2007). At depths of 4–6 km, for example, the seismic velocities are observed to change by more than 100% across horizontal distances of less than 10 km. Figure 1 compares sections through three of the SCEC models—CVM-S4, CVM-S4.26, and CVM-H11.9—with 2D tomographic models derived from a dense collection of reflection/refraction data along the two lines of the Los Angeles region seismic experiments (LARSE), LARSE I (Lutter et al., 1999) and LARSE II (Lutter et al., 2004). The models show significant differences at the basin scales, as well as deeper in the crust (Fig. 1).

In this paper, we compare the low-frequency seismograms observed in two recent Los Angeles earthquakes, 17 March 2014 Encino (Ms 4.4) and 29 March 2014 La Habra (Ms 5.1), with the synthetic seismograms computed from three SCEC CVMs being used in CyberShake and other earthquake simulations. Because the data from the recent events were not used to derive the models, they provide prospective tests of the models’ forecasting skill. The two earthquakes occurred near the axes of LARSE I and II lines (Fig. 1), which we use as additional structural information in the model comparisons.

SCEC COMMUNITY VELOCITY MODELS

Both CVM-S and CVM-H were constructed by embedding detailed basin structure models within the regional 3D seismic travel-time tomography model of Hauksson (2000). In CVM-S, the seismic velocities within major basins were determined mainly from the age and depths of the sediments using empirical relations. The latest official release is version 4 (CVM-S4), which includes a geotechnical layer constrained by sonic log data (Magistrale et al., 2000), a variable-depth Moho determined from receiver functions (Zhu and Kanamori, 2000), and an upper-mantle velocity model from Moho to about 100 km depth (Kohler et al., 2003).

In CVM-H, seismic structures of major basins were determined from a large number of sonic logs and reflection/refraction profiles from the oil industry (Süss and Shaw, 2003). The crustal structure in CVM-H was improved through 16
iterations of full-3D tomography based on the adjoint-wavefield method (AW-F3DT; Tape et al., 2009, 2010). The latest official release was in November 2011 (CVM-H11.9), which includes a geotechnical layer constrained by the near-surface ($V_{S30}$) shear velocities (Ely et al., 2010), a variable-depth Moho (Yan and Clayton, 2007), and an upper-mantle velocity model determined from finite-frequency teleseismic surface-wave tomography (Prindle and Tanimoto, 2006).

CVM-S4.26 is the 26th iterate of a full 3D tomographic (F3DT) inversion procedure (Lee et al., 2014). The procedure started with CVM-S4 and successively improved the fit to datasets that eventually included about 550,000 differential wave-form measurements at frequencies up to 0.2 Hz, obtained from about 38,000 earthquake seismograms and 12,000 ambient-noise Green’s functions. Navigation through this nonlinear iterative process involved two types of F3DT inversion methods: the AW-F3DT, which backpropagates the misfits between observed and synthetic seismograms from the receivers to image structures (Tarantola, 1984; 1988; Pratt, 1990; Tromp et al., 2005), and the scattering-integral method (SI-F3DT),

**Figure 1.** Comparisons of $P$-wave velocity models along the LARSE-I (left) and LARSE-II (right) profiles. From top to bottom, we show the 2D tomography models obtained using controlled-source travel-time data and the cross-section views through CVM-S4.26, CVM-S4, and CVM-H11.9. The vertical axes on the velocity model plots have been exaggerated by a factor of 2.0 for the LARSE-I profile and a factor of 2.5 for the LARSE-II profile. The geographic maps show the locations of the LARSE-I (left) and LARSE-II (right) profiles and the epicenters of the Encino (yellow star) and the La Habra (red star) earthquakes. WF, Whittier fault; SMFZ, Sierra Madre fault zone; SAF, San Andreas fault; SSF, Santa Susana fault; SGF, San Gabriel fault; GF, Garlock fault.
which calculates and stores the sensitivity kernels of each misfit measurement and solves the Gauss–Newton normal equation using the least-squares algorithm (Zhao et al., 2005, Zhao et al., 2006; Chen et al., 2007).

In each inversion step, synthetic seismograms for the updated model were calculated using the Olsen (1994) fourth-order staggered-grid finite-difference code, which has been optimized for massively parallel computations (Cui et al., 2010), and differential waveform measurements were made between the observed seismograms and these synthetics, accounting for the nonlinearity of the inversion. High structural resolution was obtained by included frequency-dependent, phase-coherent measurements of various seismic phases on all three components of ground motion. Lee et al. (2014) describe CVM-S4.26 and demonstrate its excellent fit to observed seismograms from a large number of well-recorded earthquakes. The seismograms from the two Los Angeles events analyzed here were not used in the F3DT inversion.

CMT INVERSION

The two 2014 Los Angeles earthquakes were well recorded by three-component broadband seismic stations of the Southern California Seismic Network (SCSN). Figure 2 gives the locations and source mechanisms determined by SCSN from focal mechanism (FM) using P-wave polarities and S/P amplitude ratios and centroid moment tensor (CMT) waveform inversions using the 1D crustal model of Dreger and Helmberger (1993) (e.g., Clinton et al., 2006; Hutton et al., 2010).

We revised the CMT solutions by applying the fast waveform-inversion technique of Lee et al. (2011). Synthetic seismograms were computed to 0.2 Hz from the 3D crustal model CVM-S4.26 for all broadband stations by a reciprocity-based finite-difference method (Zhao et al., 2006), and the optimal CMT parameters were found by a hierarchical grid-search algorithm that minimized the travel-time and amplitude differences for a selected set of observed waveforms.

For the Encino earthquake, our centroid location is 0.9 km shallower than the original SCSN-FM hypocenter and 1 km deeper than the SCSN-CMT hypocenter (Fig. 2). For the La Habra earthquake, our centroid location is at the same depth as the SCSN-CMT, which places at 2.5 km shallower than the SCSN-FM hypocenter. The Encino centroid time stayed about the same, but the La Habra centroid time increased by nearly a full second, perhaps indicative of this larger event’s finite-source duration.

The waveform fits to the Encino earthquake do not vary much among the source models, although the improved CMT does provide the best overall fit. In the case of the shallower La Habra earthquake, however, the waveform amplitudes computed from the original SCSN-FM solution tend to underestimate the observations, whereas the revised CMT yields a much better match (Fig. 3). The revised CMT provides substantially better waveform fits than either of the SCSN solutions, and this is true for the other two models, CVM-S4 and CVM-H11.9, as well as CVM-S4.26. Therefore, we have used this source model in all of our waveform comparisons.

WAVEFORM PREDICTION TEST

We tested the waveform predictions of the three CVMs against more than 900 three-component broadband seismograms recorded from the Encino and La Habra earthquakes. All synthetic and observed seismograms were band-pass filtered using a Butterworth filter with corners at 0.02 and 0.2 Hz. We measured the difference between an individual observed seismogram $u_k(t)$ and its corresponding synthetic $\tilde{u}_k(t)$ within the time window $[t_k, t'_k]$ by the relative waveform misfit (RWM) statistic, defined by the integral

$$ \text{RWM} = \sqrt{\int |u_k(t) - \tilde{u}_k(t)|^2 dt} $$

<table>
<thead>
<tr>
<th>Event</th>
<th>Source</th>
<th>Origin Time</th>
<th>Lon.</th>
<th>Lat.</th>
<th>Depth</th>
<th>Mw</th>
<th>CMT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encino</td>
<td>Full-Wave</td>
<td>2014/03/17 13:25:36.85</td>
<td>-118.492</td>
<td>34.139</td>
<td>9.0 km</td>
<td>4.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SCSN-FM</td>
<td>2014/03/17 13:25:36.90</td>
<td>-118.486</td>
<td>34.135</td>
<td>9.9 km</td>
<td>4.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SCSN-CMT</td>
<td>2014/03/17 13:25:37.00</td>
<td>-118.486</td>
<td>34.135</td>
<td>8.0 km</td>
<td>4.39</td>
<td></td>
</tr>
<tr>
<td>La Habra</td>
<td>Full-Wave</td>
<td>2014/03/29 04:09:42.97</td>
<td>-117.930</td>
<td>33.922</td>
<td>5.0 km</td>
<td>5.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SCSN-FM</td>
<td>2014/03/29 04:09:42.01</td>
<td>-117.944</td>
<td>33.919</td>
<td>7.5 km</td>
<td>5.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SCSN-CMT</td>
<td>2014/03/29 04:09:42.00</td>
<td>-117.944</td>
<td>33.919</td>
<td>5.0 km</td>
<td>5.10</td>
<td></td>
</tr>
</tbody>
</table>

▲ Figure 2. Comparisons of source parameters among full wave solutions and the results from SCSN catalogs for the Encino and La Habra earthquakes. Full-Wave, our full-wave CMT solution; SCSN-FM, SCSN focal mechanism catalog; SCSN-CMT, SCSN centroid moment tensor catalog.
The RWM statistic can be calibrated by considering a Gaussian wavelet of the form

\[ \tilde{u}(t) = A e^{-\sigma^2 (t - \tau_p)^2 / 2} \cos \omega_0 (t - \tau_p). \]  

In the narrowband limit \((\sigma/\omega_0 \ll 1)\), the RWM between this wavelet and one of equal amplitude but phase shifted by \(\omega_0 \Delta \tau_p = 1\) is nearly unity (0.92), and the maximum value of RWM = 4 is reached when \(\omega_0 \Delta \tau_p = \pi\); that is, when the two wavelets are perfectly out of phase. Similarly, two wavelets of equal phase but with different amplitudes, \(A_1\) and \(A_2\), yield a RWM just over unity (1.08) when \(\omega_0 \Delta \tau_p \equiv \ln(A_2/A_1) = 1\). RWM increases quadratically with the logarithm of the amplitude ratio, reaching 4 when \(\omega_0 \Delta \tau_p \approx 0.56\pi\). (The notation is from Gee and Jordan [1992], who generalized the phase delay \(\Delta \tau_p\) and amplitude reduction time \(\Delta \tau_q\) to be frequency-dependent data functionals.) Roughly speaking, RWM < 1 indicates a good waveform fit, whereas RWM > 1 indicates a poor fit.

We computed a station misfit by averaging \(RWM_k\) over all (up to three) components at that each station and interpolated these averages to obtain the maps in Figure 4. About 150 stations distributed across all of southern California were used for each earthquake. Histograms of the station RWMs are plotted below the maps for each model. Labeled on each histogram are its median value (mRWM) and median absolute deviation (MAD), statistics that provide robust measures of the location and spread of a univariate distribution (Hoaglin et al., 1983).

Figure 4 shows overall waveform fits to observed seismograms of the two Los Angeles earthquakes for the three CVMs. The mRWM and MAD of CVM-S4.26 are substantially smaller than those of CVM-S4 and CVM-H11.9. CVM-S4 was the starting model of our F3DT inversion procedure and the fact that CVM-S4.26 provides substantially better fit to all observed seismograms of the two earthquakes not included in our inversion is a strong evidence that our F3DT inversion procedure has been highly effective in improving the accuracy of the crustal model for the entire southern California. In the Los Angeles basin (LAB) region, where seismic hazard is high, CVM-S4.26 provides better waveform fits than both CVM-S4 and CVM-H11.9, suggesting CVM-S4.26 might have a more accurate large-scale LAB structure. For CVM-S4.26 the level of fit is equally good for both earthquakes, whereas for CVM-S4 and CVM-H11.9 the level of fit is poorer for the La Habra earthquake than that of the Encino earthquake.

Examples of the observed and the synthetic seismograms computed using the three CVMs for both earthquakes are shown in Figures 5 and 6. In general, synthetic seismograms computed using CVM-S4.26 provide substantially better fit to both the phases and the amplitudes of the observed seismograms than the other two CVMs. Synthetics computed using CVM-H11.9 generally have better fits than those computed using CVM-S4.

The structures around the LAB region in the three CVMs can be tested by examining the seismograms from the source–station paths crossing the basin. Examples of such basin paths are those from the Encino epicenter to stations LAF and SDD (Fig. 5) and those from the La Habra epicenter to LAF and SDD (Fig. 6). For the Mojave Desert region, we can examine...
Figure 4. Spatial distributions of the average RWM values for the three CVMs interpolated from the average RWM values calculated at about 150 stations. White solid line, boundaries of CVM-S4.26; white dashed lines, boundaries of CVM-H11.9; black solid lines, coastlines and active faults in southern California; black dashed lines, LARSE-I and LARSE-II profiles; stars, epicenters of the earthquakes; white circles, broadband stations from which the waveforms were used for RWM calculations. The corresponding RWM histograms are shown below the maps and the vertical dashed lines on the histograms indicate the mRWM values.
The paths from the Encino epicenter to stations TEH, CLC, RRX, and TUQ (Fig. 5) and from the La Habra epicenter to stations CLC and TUQ (Fig. 6). For structures in the western Transverse Ranges and the Coast Ranges, we can examine the paths from the Encino epicenter to station MPP (Fig. 5) and from the La Habra epicenter to station SMR. Paths crossing the Peninsular Ranges region include those from the Encino epicenter to stations SOL and JEM (Fig. 5) and from the La Habra epicenter to stations BOR and SOL (Fig. 6). The station RXH is located in the Salton trough and seismograms from both earthquakes to this station are useful for examining structures in and around the Salton trough region. For almost every source–station path we have examined, synthetic seismograms computed using CVM-S4.26 provide better fits to observed seismograms than those computed using the other two CVMs.

**Comparisons Along the LARSE Lines**

The epicenter of the Encino event lies very close to the LARSE-II profile. The source–station paths between the Encino event and stations BTP and TEH lie almost along the LARSE-II profile (Figs. 1, 4a–c, and 5). At both stations, synthetics computed using CVM-S4.26 provide the best fit to observed seismograms than those computed using the other two CVMs (Figs. 4a–c and 5). A 2D P-wave velocity model was obtained by Lutter et al. (2004), who inverted a dense collection of active-source travel-time data. We have digitized and recolored their tomography map and a comparison with the P-wave velocities in the three CVMs along the LARSE-II profile is shown on the right side of Figure 1. In general, the P-wave velocities in CVM-S4.26 have the highest correlation with the 2D model of Lutter et al. (2004). In CVM-S4, the
$P$-wave velocities below the Santa Clarita Valley from about 3 km depth to about 6 km depth are too low, the velocities inside the Antelope Valley are too high and the velocities beneath the Antelope Valley from about 6 km depth to 8 km depth are too low, compared with those in Lutter et al. (2004).

In CVM-H11.9, the San Fernando Valley and the Santa Clarita Valley are shown as a single large basin extending down to about 6–8 km depths, both the thickness and the shape of the basin are inconsistent with those in Lutter et al. (2004).

The $P$-wave velocities inside the Antelope Valley are too high, and the velocities beneath the Antelope Valley are too low in CVM-H11.9.

The epicenter of the La Habra event lies very close to the LARSE-I profile, and the source–station path between the epicenter and the station ADO lies almost along the LARSE-I profile. For this source–station path, synthetics computed using CVM-S4.26 provide the best fit to the observed seismograms than the other two CVMs (Figs. 4g–l and 6). Lutter et al. (1999) derived a 2D $P$-wave velocity model through ray-theoretic travel-time tomography using a dense collection of active-source data. Their model has been digitized, recolored, and displayed on the left side of Figure 1, together with the $P$-wave velocities in the three CVMs along the LARSE-I profile. The velocities in CVM-S4.26 show the highest correlation with the model of Lutter et al. (1999). In CVM-S4, the velocities are too low underneath the San Gabriel Valley at 6–7 km depth, the velocities are too high at shallow depths in the Mojave Desert and are too low beneath the Mojave Desert at about 6–7 km depth. In CVM-H11.9, the velocities are too low both inside and beneath the San Gabriel Valley, too high at shallow depths in the Mojave Desert, and too low at larger depths beneath the Mojave Desert.
DISCUSSIONS AND CONCLUSION

These results show that, for the two recent Los Angeles earthquakes, the waveform predictions of CVM-S4.26 are significantly better than those of CVM-S4 or CVM-H11.9. We emphasize, however, that this validation applies only to waveforms that have been low-pass filtered below the frequency range of interest for most earthquake-engineering and risk-analysis applications. Experiments are underway to extend the validation tests to higher frequencies for these and other well-recorded earthquakes using misfit measures of engineering interest, such as the goodness-of-fit criteria (e.g., Kristeková et al., 2009; Taborda and Bielak, 2013).

The three CVMs tested in this study are all isotropic. For both CVM-S4 and CVM-S4.26, the RWM values do not show significant variations with respect to the component. However, the RWM values for CVM-H11.9 show that the misfits on the transverse components are substantially larger than those on the vertical and the radial components (Fig. 7). Example seismograms of such component-dependent misfit for CVM-H11.9 can be seen at stations JEM, MSC, RXH, SOL, and IRM on Figure 5 and at stations SPG2, IRM, EDW2, BOR, TUQ, and

![Figure 7](image-url)
MCT on Figure 6. One hypothesis for this observation is that there exists a significant level of radial anisotropy in the southern California crust, therefore synthetics computed using purely isotropic velocity models are not capable of fitting all three components equally well (Tape et al., 2010). We do not favor this hypothesis based on waveform comparisons shown in this study, as synthetics computed using the isotropic CVM-S4.26 can fit all three components equally well for the majority of the source–station paths we have examined. In Chen et al. (2007), a very small (<1%) but statistically significant level of crustal anisotropy was detected in the LAB region using S waves at frequencies ranging from 0.2 to 1 Hz. One possible explanation proposed in Chen et al. (2007) was the vertical variations in elastic moduli in thin, isotropic layers in the sediments (Backus, 1962). We expect this basin layering effect to be significant at basin scales but not at the crustal scale at frequencies below 0.2 Hz. An in-depth study of this issue will likely require a much larger dataset that provides better sampling of the crust.

An accurate estimate of hypocenter is critical for both geologic interpretation of active faults and ground-motion predictions for seismic-hazard assessments. The SCSN special report of the $M_w$ 5.1 La Habra earthquake suggests that the La Habra earthquake sequence might be associated with the Puente Hills thrust, a blind thrust in the Los Angeles region. The La Habra earthquake is located in the center of the seismic network, so the SCSN-CMT solution is well resolved. However, an inaccurate hypocenter estimate may lead to inappropriate understanding and interpretation of the active fault. For the La Habra earthquake, two foreshocks preceded the mainshock and a few hundred aftershocks occurred in 15 km distance from the mainshock epicenter. Accurate hypocenters are important for analyzing this foreshock–mainshock–aftershock sequence. The initial SCSN report placed the hypocenter of the mainshock at about 7.5 km depth, which is about 2.5 km deeper than the centroid locations in the CMT solutions determined by us and also by a refined SCSN analysis. Synthetic seismograms computed using the hypocenter in the initial SCSN report substantially underestimate the observed surface-wave amplitudes at frequencies below 0.2 Hz (Fig. 3). The updated SCSN report has moved the hypocenter to 4.8 km depth. An accurate 3D velocity model that correctly accounts for the basin structures in southern California can be important for improving the accuracy of the hypocenters.

The waveform prediction tests presented in this paper are important for evaluating the accuracy of the different SCEC CVMs for a variety of applications, such as long-term seismic-hazard assessments based on the CyberShake platform. Results from this study indicate that among the three SCEC CVMs, CVM-S4.26 is the most accurate model for predicting low-frequency ($\leq 0.2$ Hz) seismograms. Efforts are currently underway to further improve CVM-S4.26 by assimilating ground-motion observations at frequencies up to 0.5 Hz through our F3DT inversion procedure. With the continued development of the 3D crustal structure model, we should be able to conduct simulation-based seismic-hazard analysis at higher frequencies for earthquake-engineering applications in southern California in the near future.

ACKNOWLEDGMENTS

This research was supported by the Southern California Earthquake Center (SCEC), which is funded by the National Science Foundation and U.S. Geological Survey. This research used resources of the Argonne Leadership Computing Facility at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357. En-Jui Lee is supported by the SCEC. Po Chen is supported jointly by the School of Energy Resources and the Department of Geology and Geophysics at the University of Wyoming.

REFERENCES


