2018 SCEC Final Report for Award #18098

Toward Full Waveform Tomography Across California

Carl Tape, Principal Investigator
Geophysical Institute
University of Alaska Fairbanks
903 Koyukuk Drive
Fairbanks, AK, 99775
Phone: 907-474-5456, FAX: 907-474-7125
Email: `ctape@alaska.edu`

Clifford Thurber, Principal Investigator
Department of Geoscience
University of Wisconsin-Madison
1215 W. Dayton St.
Madison, WI 53706
Phone: 608-262-6027, FAX: 608-262-0693
Email: `cthurber@wisc.edu`

Total amount:	$ 60,000
UAF (Tape):	$ 30,000
UW (Thurber):	$ 30,000

2018 SCEC RFP

Proposal Category:	B: Integration and Theory
SCEC Science Priorities:	4a, 3a, 3b

Abstract

We seek to establish a complete, open-source software workflow for simulation-based seismic inversion in southern California. Our goal is to establish seismic imaging tools that are documented and usable by a broader group of seismologists. This would provide a sustainable pathway to validating and updating the SCEC Community Velocity Models (CVMs), such as CVM-H15.1 and CVM-S4.26, which are used for probabilistic ground motion simulation projects such as CyberShake. Earthquakes and ambient noise cross-correlations can both be used for seismic imaging. To start this effort, we focus on extracting the CVMs using UCVM software, choosing a small target set of earthquakes and ambient noise "master stations," performing moment tensor inversions, and establishing a semi-automated 3D synthetic inversion using a software package called seisflows. Also, using classical tomography approaches, we improve upon the tomographic models in central California and aim to use these models within future simulation-based inversions.

Intellectual Merit

We are promoting the use of highly accurate seismic wavefield simulations for the purposes of ground motion predictions and for improving our 3D tomographic models of southern California. Simulation- and adjoint-bases seismic inversion remains a formidable challenge, and we are seeking to reduce the effort by establishing a new workflow.

Broader Impacts

Our emphasis on open-source software is a key component that will allow the broadest range of seismologists to benefit from our work. In addition to using open-source software, we aim to provide suitable documentation, with the hope of attracting and sustaining users.

The project initiated a fruitful collaboration between UAF and U. Wisconsin and also fostered the development of three postdocs new to the SCEC community: Avinash Nayak, Ryan Modrak, and Vipul Silwal. Team members (Tape, Thurber, Nayak) participated in the SCEC CVM workshop in Pomona in October 2018, and also interacted with UCVM software developers over the course of the project.

Introduction

The SCEC Community Velocity Models (CVMs) are one of the key SCEC products, especially for their importance in research involving seismic ground motion simulations for earthquake hazard estimation (*Graves et al.*, 2011). Progress on improving the CVMs, however, has been limited over the past few years. At an October 2018 SCEC workshop, "The Next Leap Forward for SCEC Community Velocity Models," the status of the two southern California CVMs, CVM-S4.26 (*Lee et al.*, 2014) and CVM-H15.1 (*Shaw et al.*, 2015), was reviewed, and ideas for moving forward with these models were discussed in detail.

Among the action items from that workshop, and also indicated as a priority in the SCEC 2019 Science Plan, is the goal of developing a new, complete, open-source workflow for updating CVMs via full waveform tomography (FWT; also known as F3DT, full waveform inversion, adjoint tomography). The goal of this project was to develop such a workflow. Below we summarize the status and outcomes of the project tasks. The labels '[UAF]' = Tape + Silwal + Modrak and '[UW]' = Thurber + Nayak indicate the partitioning of tasks.

Progress on establishing initial models for full waveform tomography

1. Estimation of hypocenters and origin times [UW].
   The UW group has developed a high-quality San Francisco Bay area (SFBA) 3D V_S model and has extended their central California 3D V_P and V_S models to cover the entire SCEC-designated Central California Area (CCA) region (see 4). Therefore, the capability to provide accurate locations and origin times now exists for all the regions of interest in this project.

2. Refinement of San Francisco Bay Area and Central California Area velocity models [UW].
   We have leveraged internal sources of support to continue work on body-wave, surface-wave, and joint inversions for the CCA, the subject of the UW-Madison 2017 SCEC project, and for the SFBA. Our goal with these efforts is to produce 3D Vp and Vs models at reasonably high resolution that can subsequently be used as starting models for F3DT following the workflow we are developing.

   Separate tomographic body-wave (Vp and Vs) and surface-wave (Vs) inversions were originally completed for the CCA for the 2017 SCEC project. We have substantially increased the size of the body-wave dataset using network picks and data from all major temporary deployments and recently deployed stations, and modified the tomography code tomoDD (*Zhang and Thurber*, 2003) to incorporate a constraint on the allowed Vp/Vs ratio. For Rayleigh waves in ANGF, we used our previous 3D Vs results derived from group velocity measurements to redefine measurement windows for estimating phase velocities in the period range 4 to 18 s. Inversion of the phase velocity data for phase velocity maps yielded a data misfit about half that for the group velocity data, indicating the higher quality of the phase velocity measurements. A preliminary joint inversion of the new datasets for Vp and Vs structure was carried out using the method of *Fang et al.* (2016). We are now exploring smoothing and damping parameter tradeoffs in order to determine the optimum values. Figure 1a-b show the improvement in the Vs model in terms of greater coherence in lower crustal features compared to the *Lin et al.* (2010) model on account of the expanded dataset. Figure 1c shows near-surface Vs from the joint inversion, featuring a low-velocity zone along the San Andreas fault and a clear extent of the Central Valley. Vs at the same depth from the CCA6 model (En-Jui Lee and co-workers) is shown in Figure 1d for comparison.

   Our SFBA tomography work has followed a similar path. A 3D Vs model from ambient noise data was produced first, and has been published (*Li and Thurber*, 2018). Following the assembly of a large body-wave dataset, a joint body-wave/surface-wave inversion was carried out using the (*Fang et al.*, 2016) algorithm to produce 3D models of Vp and Vs with a 0.05° horizontal node spacing and 1 to 10 km vertical node spacing. We have provided the results to Artie Rodgers of Lawrence Livermore National Laboratory so that he can carry out a validation analysis comparing synthetics generated from our tomographic results to synthetics generated with the USGS 3D Bay Area velocity model (*Brocher*, 2005; *Brocher et al.*, 2006).

We also worked with Hongjian Fang (currently at MIT) and others on an updated joint body-wave/surface wave inversion algorithm that solves simultaneously for Vp, Vs, and Vp/Vs with an optimal strategy. That work was recently published (*Fang et al.*, 2019) as SCEC Contribution No. 8909.

Progress on full waveform tomography workflow

1. Selection of target region and reference model [UAF, UW].
   A new workflow for FWT can best be benchmarked in a region that has a velocity model that has been extensively developed and broadly analyzed. Since FWT is a nonlinear process, testing a new workflow is best done with an initial model that fits existing data reasonably well, such as CVM-S4.26 (*Magistrale et al.*, 2000; *Lee et al.*, 2014) and CVM-H15.1 (*Süss and Shaw*, 2003; *Komatitsch et al.*, 2004; *Tape et al.*, 2009, 2010; *Shaw et al.*, 2015). We attempted to extract gridded models of CVM-S4.26 and CVM-H15.1 using the UCVM software (*Small et al.*, 2017). Some issues were encountered during the extraction of models from UCVM, and these issues were communicated to SCEC CME staff, which has led to updates in UCVM (2019-02-28 email from Phil Maechling).

2. Estimation of hypocenters and origin times.
   Waveform-relocated hypocenters and origin times in southern California (*Hauksson et al.*, 2012) are now updated on a quarterly basis (E. Hauksson, e-communication, 2018-11-07). These source parameters are sufficiently accurate for us to avoid introducing artifacts into the adjoint-based inversion (e.g., *Lin et al.*, 2007; *Tape et al.*, 2010), which targets periods 2 s and longer.

3. Construction of database of ambient noise Green's functions (ANGF) [UW].
   An expansive database of frequency-dependent ANGF is needed for F3DT (*Lee et al.*, 2014). For the SFBA, the UW group recently produced a new ANGF catalog that yields about 30,000 Rayleigh-wave group velocity measurements from 174 stations over the period range of 0.5 to 14 seconds. For central California, we have compiled a new multi-component ANGF catalog using data of ∼730 temporary and permanent stations (∼430 3-component, ∼280 of them broadband) in the frequency range of 4 to 18 seconds. Rayleigh-wave dispersion measurements from these catalogs have been used for surface-wave tomography and joint body-wave/surface-wave tomography of the CCA region (*Nayak et al.*, 2018) and the SFBA (*Li and Thurber*, 2018) with the support of a 2017 SCEC award and other sources of support. Both of these ANGF catalogs are available as constraints for future F3DT work. We will also examine all available databases of ANGF (e.g., `https://ds.iris.edu/ds/products/ancc-ciei/`). Comparison of ANGF results for a set of the same station pairs allows assessment of ANGF quality and uncertainty and the effects of different processing approaches. Comparisons of a subset of our vertical-vertical component ANGFs involving permanent broadband stations and TA stations in Central California with ANGFs in the University of Colorado-Boulder catalog indicates that our ANGFs have slightly higher SNR at shorter periods and are effectively corrected for polarity errors. We are also collaborating with Lise Retailleau of Stanford University, who has agreed to share her catalog of ANGFs for the entire state of California.

4. Construction of database of earthquake waveforms [UAF].
   Earthquakes are selected from within a target simulation region (Figure 2). Waveforms are extracted and processed (including removal of instrument response) using the `pysep` package (`https://github.com/uafseismo/pysep`). (This is open-source but is mainly used by UAF seismologists only.)

5. Estimation of moment tensors [UAF, UW].
   We partition the simulation domain into 5 × 5 regions, with each of the 25 regions split at a depth of 12 km. Within each of the 50 bins, we select the largest-magnitude event within the depth range $M_{\rm w}$ 4.5–5.5 and between 1998–01-01 and 2018-09-01. This leads to 21 earthquakes (Figure 2).

   We use the cut-and-paste method, established by *Zhu and Helmberger* (1996), used in *Tape et al.* (2009) for the initial moment tensors, and recently adapted and applied in *Silwal and Tape* (2016). The approach assumes a 1D model (*Dreger and Helmberger*, 1990), and synthetic seismograms are calculated using a frequency-wavenumber method (*Zhu and Rivera*, 2002). We use body waves (periods 1.5–4 s) and surface waves (periods 16–40 s).

   The estimated moment tensors are shown in Figure 2. One of these earthquakes (2003-12-22 19:26:07, $M_{\rm l}$ 4.7) is an aftershock of the 2003 San Simeon earthquake and is not usable, due to poor signal-to-noise levels. Another earthquake (1999-05-14 07:54:02), near Palm Springs, does not fit the waveforms well using the waveform-relocated depth of 2.4 km. A search over depths should be performed for this event. It would be good to understand the discrepancy, given the good coverage (34 stations) for this event.

   We also have the capability to estimate moment tensors using a 3D reference model (*Liu et al.*, 2004; *Tape et al.*, 2010), but this is beyond the scope of this project.

6. Implementation of FWT for 10 earthquakes and 10 master stations [UAF, UW].
   Our primary goal is to train two teams and to establish the complete FWT workflow using established—though often in a continual state of development—open-source codes. We use the software package seisflows (*Modrak and Tromp*, 2016; *Rusmanugroho et al.*, 2017; *Modrak et al.*, 2018), which facilitates iterative inversions on an HPC cluster by tracking jobs, handling I/O, and performing an iterative optimization. Tape, Nayak, Silwal, and Modrak all have user accounts on the UAF chinook cluster, where Modrak is leading the task. We have successfully iterated an example 3D checkerboard model using seisflows. This involves generating synthetics in a simple model, then target synthetics in a model that has a checkerboard perturbation from the initial model, then trying to recover the checkerboard using wavefield simulations and volumetric sensitivity kernels. (In a real inversion, the target synthetics would be replaced by recorded data.)

   Our ongoing efforts focus on: (1) generalizing seisflows to accommodate a set of different misfit functions (such as the open-source pyflex and pyadjoint packages for selecting windows, makeing phase-based measurements, and making adjoint sources), (2) making minor adjustments to the UCVM-extracted SCEC CVM models, (3) using the set of 10–20 earthquakes and 10–20 master stations from ANGF within the southern Califoria inversion.

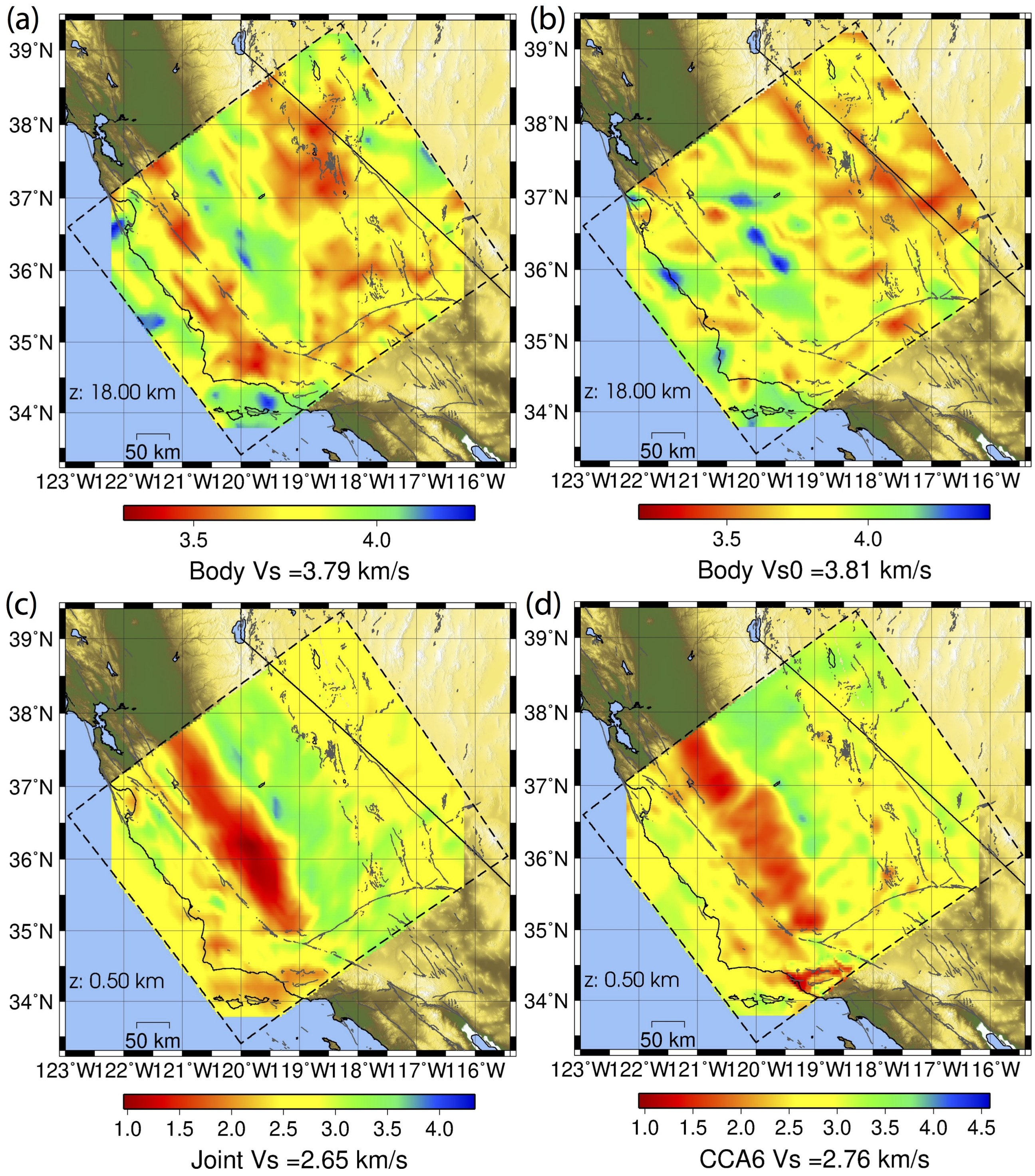

Figure 1: (a) Vs at depth 18 km estimated using the revised body-wave dataset and TOMODD with the Vp/Vs ratio constraint, compared to the *Lin et al.* (2010) model shown in (b). (c) Vs at depth 0.5 km from joint inversion of the revised body-wave dataset and ambient noise derived Rayleigh-wave phase velocity dispersion measurements compared to model CCA6 shown in (d).

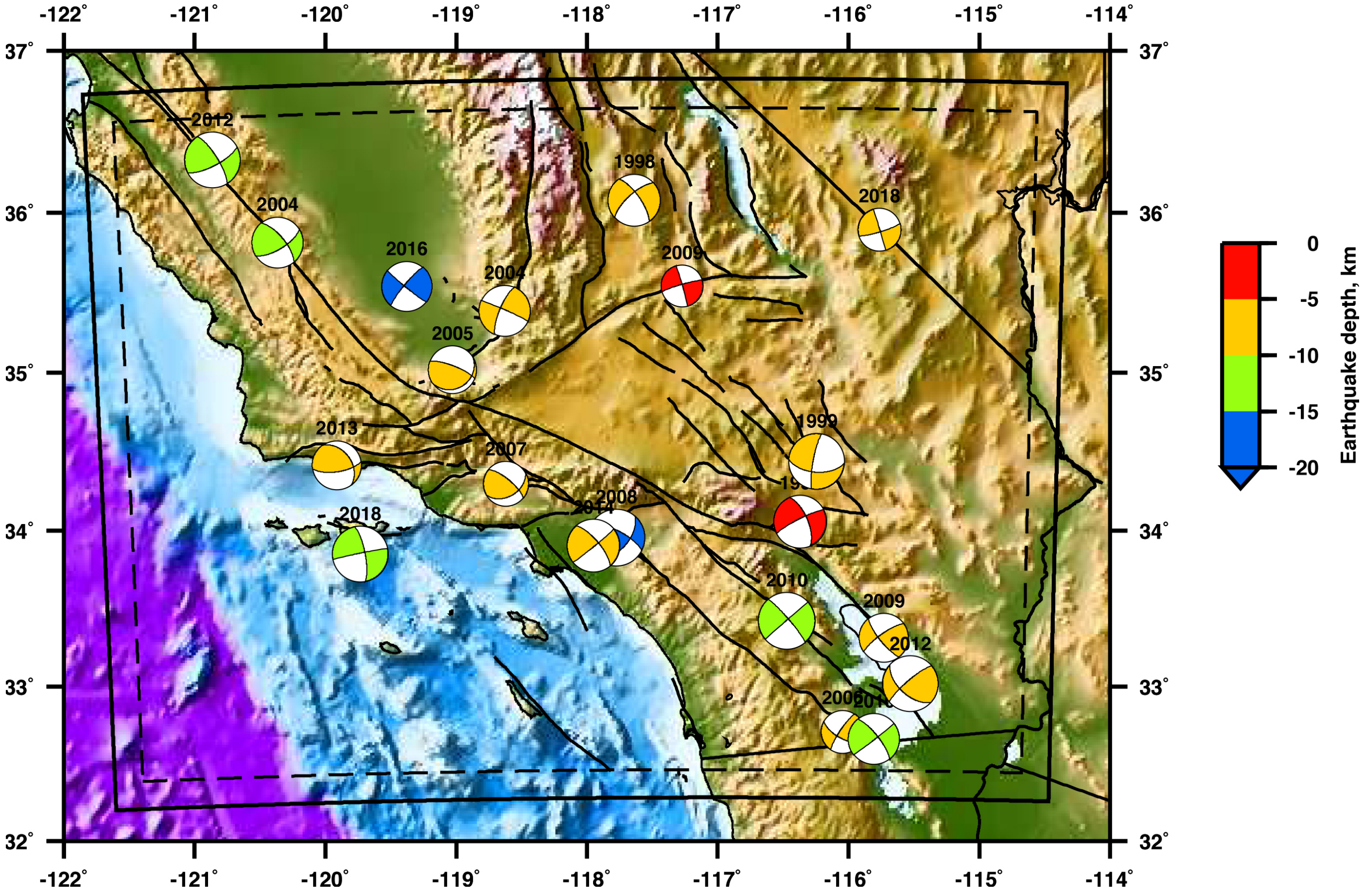

Figure 2: Seismic moment tensors estimated by Vipul Silwal for 20 example events. The events are selected starting with all M_l 4.5–5.5 events between 1998-01-01 and 2018-09-01. A declustering approach is used to obtain a set of events; where adjacent events appear, on event is shallower than 12 km, while the other is deeper. The outer box is the simulation boundary; the inner dashed box is a buffer zone to avoid events near the boundary. The hypocenters and origin times are from the relocated catalog of *Hauksson et al.* (2012) (and later updates). The magnitude and mechanism are estimated using body waves and surface waves (*Zhu and Helmberger*, 1996; *Silwal and Tape*, 2016).

References

Brocher, T., B. Aagaard, R. Simpson, and R. Jachens (2006), The new USGS 3D seismic velocity model of northern California (abstract), *Seismol. Res. Lett.*, *77*, 271.

Brocher, T. M. (2005), Compressional and Shear Wave Velocity Versus Depth in the San Francisco Bay Area, California: Rules for USGS Bay Area Velocity Model 05.0.0, Open-File Report 05-1317.

Dreger, D. S., and D. V. Helmberger (1990), Broadband modeling of local earthquakes, *Bull. Seismol. Soc. Am.*, *80*(5), 1162–1179.

Fang, H., H. Zhang, H. Yao, A. Allam, D. Zigone, Y. Ben-Zion, C. Thurber, and R. D. van der Hilst (2016), A new algorithm for three-dimensional joint inversion of body wave and surface wave data and its application to the Southern California plate boundary region, *J. Geophys. Res. Solid Earth*, *121*, 3557–3569, doi:10.1002/2015JB012702.

Fang, H., H. Yao, H. Zhang, C. Thurber, Y. Ben-Zion, and R. D. van der Hilst (2019), $V_\mathrm{p}/V_\mathrm{s}$ tomography in the southern California plate boundary region using body- and surface-wave traveltime data, *Geophys. J. Int.*, *216*, 609–620, doi:10.1093/gji/ggy458.

Graves, R., et al. (2011), CyberShake: A physics-based seismic hazard model for southern California, *Pure App. Geophys.*, *168*, 367–381.

Guo, B. (2018), Application of state of the art seismic techniques to California, Ph.D. thesis, University of Wisconsin–Madison.

Hauksson, E., W. Yang, and P. M. Shearer (2012), Waveform relocated earthquake catalog for southern California (1981 to June 2011), *Bull. Seismol. Soc. Am.*, *102*(5), 2239–2244, doi:10.1785/0120120010.

Komatitsch, D., Q. Liu, J. Tromp, P. Süss, C. Stidham, and J. H. Shaw (2004), Simulations of ground motion in the Los Angeles basin based upon the spectral-element method, *Bull. Seismol. Soc. Am.*, *94*(1), 187–206, doi:10.1785/0120030077.

Lee, E.-J., P. Chen, T. H. Jordan, P. B. Maechling, M. A. 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*, 6421–6451, doi:10.1002/2014JB011346.

Li, P., and C. Thurber (2018), Rayleigh wave group velocity and shear wave velocity structure in the San Francisco Bay region from ambient noise tomography, *Geophys. J. Int.*, *213*, 1599–1607, doi:10.1093/gji/ggy095.

Lin, G., P. M. Shearer, and E. Hauksson (2007), Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005, *J. Geophys. Res.*, *112*, B12309, doi:10.1029/2007JB004986.

Lin, G., C. H. Thurber, H. Zhang, E. Hauksson, P. M. Shearer, F. Waldhauser, T. M. Brocher, and J. Hardebeck (2010), A California statewide three-dimensional seismic velocity model from both absolute and differential times, *Bull. Seismol. Soc. Am.*, *100*(1), 225–240.

Liu, Q., J. Polet, D. Komatitsch, and J. Tromp (2004), Spectral-element moment tensor inversions for earthquakes in southern California, *Bull. Seismol. Soc. Am.*, *94*(5), 1748–1761, doi:10.1785/012004038.

Magistrale, H., S. Day, R. W. Clayton, and R. Graves (2000), The SCEC Southern California reference three-dimensional velocity model Version 2, *Bull. Seismol. Soc. Am.*, *90*(6B), S65–S76.

Modrak, R., and J. Tromp (2016), Seismic waveform inversion best practices: regional, global, and exploration test cases, *Geophys. J. Int.*, *206*, 1864–1889, doi:10.1093/gji/ggw202.

Modrak, R. T., D. Borisoc, M. Lefebvre, and J. Tromp (2018), SeisFlows—Flexible waveform inversion software, *Computers & Geosciences*, *115*, 88–95, doi:10.1016/j.cageo.2018.02.004.

Nayak, A., C. Thurber, H. Fang, X. Zeng, and H. Zhang (2018), Surface-Wave and Body-Wave Tomography for Central California, Abstract S21B-06 presented at 2018 Fall Meeting, AGU, Washington, D.C., 10-14 Dec.

Rusmanugroho, H., R. Modrak, and J. Tromp (2017), Anisotropic full-waveform inversion with tilt-angle recovery, *Geophysics*, *82*(3), 1–17, doi:10.1190/geo2016-0025.1.

Shaw, J. H., et al. (2015), Unified Structural Representation of the southern California crust and upper mantle, *Earth Planet. Sci. Lett.*, *415*, 1–15, doi:10.1016/j.epsl.2015.01.016.

Silwal, V., and C. Tape (2016), Seismic moment tensors and estimated uncertainties in southern Alaska, *J. Geophys. Res. Solid Earth*, *121*, 2772–2797, doi:10.1002/2015JB012588.

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*(6), 1539–1552, doi:10.1785/0220170082.

Süss, M. P., and J. H. Shaw (2003), P-wave seismic velocity structure derived from sonic logs and industry reflection data in the Los Angeles basin, California, *J. Geophys. Res.*, *108*(B3), 2170, doi:10.1029/2001JB001628.

Tape, C., Q. Liu, A. Maggi, and J. Tromp (2009), Adjoint tomography of the southern California crust, *Science*, *325*, 988–992, doi:10.1126/science.1175298.

Tape, C., Q. Liu, A. Maggi, and J. Tromp (2010), Seismic tomography of the southern California crust based on spectral-element and adjoint methods, *Geophys. J. Int.*, *180*, 433–462, doi:10.1111/j.1365-246X.2009.04429.x.

Zhang, H., and C. H. Thurber (2003), Double-difference tomography: the method and its application to the Hayward fault, California, *Bull. Seismol. Soc. Am.*, *93*(5), 1875–1889, doi:10.1785/0120020190.

Zhu, L., and D. Helmberger (1996), Advancement in source estimation techniques using broadband regional seismograms, *Bull. Seismol. Soc. Am.*, *86*(5), 1634–1641.

Zhu, L., and L. A. Rivera (2002), A note on the dynamic and static displacements from a point source in multilayered media, *Geophys. J. Int.*, *148*, 619–627, doi:10.1046/j.1365-246X.2002.01610.x.