Seismic tomography of the southern California crust based on spectral-element and adjoint methods

Carl Tape, Qinya Liu, Alessia Maggi, Jeroen Tromp

Geophysical Journal International

Geophys. J. Int. (2010) **180,** 433–462 doi: 10.1111/j.1365-246X.2009.04429.x

Seismic tomography of the southern California crust based on spectral-element and adjoint methods

Carl Tape,[1*] Qinya Liu,[2] Alessia Maggi[3] and Jeroen Tromp[4]

[1]*Seismological Laboratory, California Institute of Technology, Pasadena, CA* 91125, *USA. E-mail: carltape@gps.caltech.edu*
[2]*Department of Physics, University of Toronto, ON, Canada*
[3]*Institut de Physique du Globe, Université de Strasbourg, Strasbourg, France*
[4]*Department of Geosciences and Program in Applied & Computational Mathematics, Princeton University, Princeton, NJ, USA*

Accepted 2009 October 22. Received 2009 October 18; in original form 2009 July 31

SUMMARY

We iteratively improve a 3-D tomographic model of the southern California crust using numerical simulations of seismic wave propagation based on a spectral-element method (SEM) in combination with an adjoint method. The initial 3-D model is provided by the Southern California Earthquake Center. The data set comprises three-component seismic waveforms (i.e. both body and surface waves), filtered over the period range 2–30 s, from 143 local earthquakes recorded by a network of 203 stations. Time windows for measurements are automatically selected by the FLEXWIN algorithm. The misfit function in the tomographic inversion is based on frequency-dependent multitaper traveltime differences. The gradient of the misfit function and related finite-frequency sensitivity kernels for each earthquake are computed using an adjoint technique. The kernels are combined using a source subspace projection method to compute a model update at each iteration of a gradient-based minimization algorithm. The inversion involved 16 iterations, which required 6800 wavefield simulations. The new crustal model, $\mathbf{m}_{16}$, is described in terms of independent shear (V_S) and bulk-sound (V_B) wave speed variations. It exhibits strong heterogeneity, including local changes of ± 30 per cent with respect to the initial 3-D model. The model reveals several features that relate to geological observations, such as sedimentary basins, exhumed batholiths, and contrasting lithologies across faults. The quality of the new model is validated by quantifying waveform misfits of full-length seismograms from 91 earthquakes that were not used in the tomographic inversion. The new model provides more accurate synthetic seismograms that will benefit seismic hazard assessment.

Key words: Inverse theory; Body waves; Surface waves and free oscillations; Seismic tomography; Computational seismology; Crustal structure.

1 INTRODUCTION

Seismic tomography uses measurements between simulated (or 'synthetic') and observed seismic waveforms to obtain 3-D images of Earth's interior. A tomographic inversion is a minimization problem that begins with the specification of a measure of misfit between synthetic and observed seismograms. The accuracy of a tomographic inversion is limited by the accuracy of the synthetic seismograms. It is also limited by the accuracy of the sensitivity kernels (i.e. Fréchet derivatives), which reveal the sensitivity of each measurement to changes in the model. Highly accurate numerical methods, such as the spectral-element method (SEM) (e.g. Faccioli *et al.* 1997; Komatitsch & Tromp 1999), may now be used to compute synthetic seismograms in complex 3-D models at regional (Komatitsch & Vilotte 1998; Komatitsch *et al.* 2004) and global scales (Komatitsch & Tromp 2002a,b; Capdeville *et al.* 2003). Furthermore, the same numerical methods may be used to compute sensitivity kernels corresponding to these 3-D models at regional (Liu & Tromp 2006) and global scales (Liu & Tromp 2008). Equipped with these tools, Tape *et al.* (2009) demonstrated the feasibility of iteratively improving a complex 3-D crustal model of southern California. Their iterative inversion used traveltime measurements of body and surface waves from 52 000 three-component seismic records of 143 crustal earthquakes. This paper extends the results and analysis of Tape *et al.* (2009).

Tromp *et al.* (2005) emphasized the generality of using adjoint methods in seismic tomography, in terms of both the choice of model parameters and the choice of misfit function. One may choose model parameters that describe 3-D elastic structure (e.g. elastic tensor **c**, density ρ), 3-D attenuation, topography of

*Now at: Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA.

interfaces (e.g. Mohorovičić and basement surfaces), and earthquake sources (e.g. moment tensor **M**, hypocentre, origin time). The choice of misfit function—for example, waveform, traveltime, or amplitude differences—will affect the construction of 'adjoint sources'. Adjoint sources are source time functions that are placed at the stations to generate a wavefield that interacts with the regular wavefield to form the gradient of the misfit function for each earthquake (Tromp *et al.* 2005; Tape *et al.* 2007). The fundamental insight of using interactions between the regular wavefield emanating from the source and a secondary (adjoint) wavefield generated simultaneously at all the receivers is attributed to Claerbout (1971). This approach has been utilized in many studies (e.g. Chavent *et al.* 1975; Bamberger *et al.* 1982; McMechan 1982; Talagrand & Courtier 1987). Tarantola (1984) recognized that 'backprojecting' data, as suggested by Claerbout, should be viewed as just one step in an iterative inverse procedure. We use the term 'adjoint tomography' (Tape *et al.* 2007) for any tomographic inverse problem that employs adjoint methods (e.g. Tarantola 1984; Mora 1987; Luo & Schuster 1991; Igel *et al.* 1996; Pratt *et al.* 1998; Akçelik *et al.* 2003); see Tromp *et al.* (2005) and Virieux & Operto (2010) for additional references.

This paper is organized following the steps of a minimization problem.

(i) Specify an initial model described in terms of a set of earthquake source parameters and 3-D variations in density, shear wave and bulk-sound speeds (Section 2).
(ii) Specify a misfit function between observed waveforms and synthetic waveforms computed using 3-D wavefield simulations (Section 3).
(iii) Evaluate the misfit function for the initial model.
(iv) Compute the gradient (and Hessian, if feasible) of the misfit function for the initial model.
(v) Iteratively minimize the misfit function (Section 4).

In Section 5, we present the new crustal model and discuss implications for seismic hazard assessment. We provide a misfit analysis (Section 6) followed by a discussion of resolution and uncertainty (Section 7), and finish with a summary of key points (Section 8).

We present the new 3-D crustal model, $\mathbf{m}_{16}$, in a series of vertical cross-sections with corresponding full-length three-component seismograms for selected stations. Many new tomographic features are revealed, the most dramatic of which are the Coast Ranges and their numerous sedimentary basins, the southern San Joaquin basin, the crustal profile of the Mojave Desert region, the mid-crust of the western Transverse Ranges, and tomographic contrasts across faults, including the San Andreas.

2 INITIAL MODEL ($\mathbf{m}_{00}$)

We present the initial model and earthquake sources prior to discussing the measurement procedure (Section 3), although model and measurement choices are intertwined. The primary objective is to fit, via quantitative measurements, three-component, full-length synthetic seismograms to observed seismograms. To this end, we need to generate numerically accurate seismic waveforms using the best available descriptions of southern California structure and earthquakes.

We compute synthetic seismograms using the SEM. Due to the accuracy of the SEM (e.g. Komatitsch *et al.* 2004), the goodness of fit between observed and synthetic seismograms depends only on the following three factors.

(i) The quality of the tomographic model.
(ii) The quality of the earthquake source parameters.
(iii) The amount of noise in the observed seismograms.

Here we describe the structure and source parameters, followed by a description of the model vector **m** used in the tomographic inversion.

2.1 Initial 3-D crustal model

The initial 3-D model is provided by the Southern California Earthquake Center (www.scec.org) and was constructed using data and interpretations from several different seismic data sets: seismic reflection and industry well-log data to constrain the geometry and structure of major basins (Süss & Shaw 2003; Komatitsch *et al.* 2004; Lovely *et al.* 2006), receiver function data to estimate the depth to the Moho (Zhu & Kanamori 2000), and local earthquake data to obtain the 3-D background wave speed structure (Hauksson 2000; Lin *et al.* 2007b). The model is described in terms of shear wave speed (V_S) and bulk sound speed (V_B), which can be combined to compute compressional wave speed,

$$V_P^2 = \frac{4}{3} V_S^2 + V_B^2 \,. \tag{1}$$

We use a quality factor of $Q_\mu = 90$ within sedimentary basins, and no attenuation elsewhere (Komatitsch *et al.* 2004). The Moho surface provided by Zhu & Kanamori (2000) improved upon the surface wave-based waveform inversion results of Das & Nolet (1998); the most recent model is by Yan & Clayton (2007). The Moho surface remained fixed throughout our inversion; however, wave speed perturbations were allowed across the boundary.

As discussed in Tape *et al.* (2009), we make minor changes to the model presented in Süss & Shaw (2003), Komatitsch *et al.* (2004) and Lovely *et al.* (2006). The simulation region extends further to the west to include the Coast Ranges (Fig. 1). We implement an updated background model (Hauksson 2000; Lin *et al.* 2007b), and we obtain density (ρ) by empirically scaling V_P (Ludwig *et al.* 1970; Brocher 2005):

$$\begin{aligned}\rho(V_P) = {} & 1.6612\, V_P - 0.4721\, V_P^2 + 0.0671\, V_P^3 \\ & - 0.0043\, V_P^4 + 0.000106\, V_P^5,\end{aligned} \tag{2}$$

where ρ is in $\mathrm{g\,cm^{-3}}$ and V_P is in $\mathrm{km\,s^{-1}}$. Thus, density changes with each iteration via eq. (2).

2.2 Earthquake sources

The earthquakes in this study were modelled as point sources. Each source is described by ten parameters: origin time (one), hypocentre (three) and moment tensor (six). Most of the epicentres and origin times were previously determined using the relocation technique of Lin *et al.* (2007a), and these remain unchanged during the iterative improvement of the tomographic model. We performed numerous tests to determine the best focal mechanism for each earthquake, and we inverted for the focal mechanisms once in the initial 3-D model ($\mathbf{m}_{00}$) and again at the 12th iteration ($\mathbf{m}_{12}$) (Liu *et al.* 2004). At the first iteration, the consistency of the volumetric event kernels revealed that the seismogram misfit was dominated by structural perturbations. Therefore, we refrained from performing focal mechanism inversions, which require 6–7 wavefield simulations each, until the crustal model was substantially improved.

Details regarding the selection of events and the inversion for source parameters were presented in Tape *et al.* (2009, section S6).

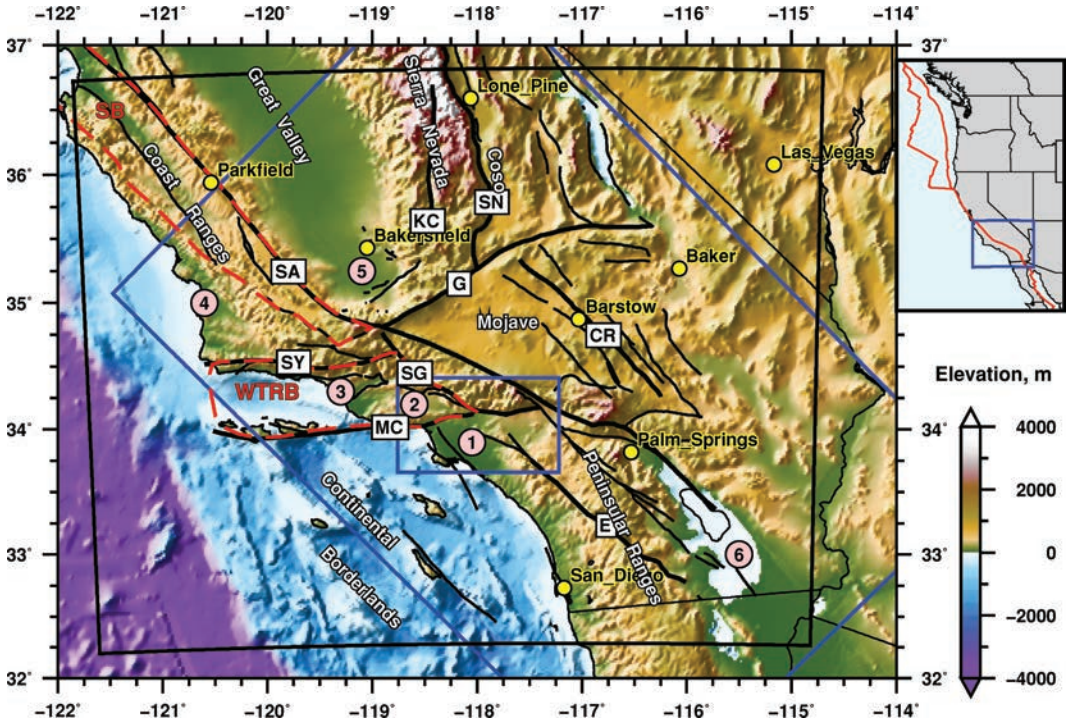


Figure 1. Map showing topography & bathymetry (Amante & Eakins 2008) and active faults (Jennings 1994). Labels 1–6 denote the sedimentary basins of (1) Los Angeles, (2) San Fernando, (3) Ventura–Santa Barbara, (4) Santa Maria, (5) southern San Joaquin and (6) the Salton trough, all of which have been active during the Neogene. Dashed red lines outline blocks that have undergone substantial Neogene motion: the Salinian block (SB) within the Coast Ranges, and the western Transverse Ranges block (WTRB). The black outline denotes the simulation region, which extends to 60 km depth. The oblique blue outline is the tomography region of Lin *et al.* (2007b); the inner blue outline is the tomography region of Chen *et al.* (2007). Faults labelled for reference in subsequent figures are drawn in bold black and labelled in the boxes: SA, San Andreas, KC, Kern Canyon, SN, Sierra Nevada, G, Garlock, CR, Camp Rock, SG, San Gabriel, SY, Santa Ynez, MC, Malibu Coast, E, Elsinore. The Camp Rock fault, as labelled, includes a connection of faults from north to south: Gravel Hills–Harper fault, Harper Lake fault, Camp Rock fault, and Emerson fault. The Malibu Coast fault is drawn in continuation to the west to mark the southern boundary of the WTRB; the Santa Ynez fault is drawn in continuation to the west to mark the northern boundary of the WTRB (Luyendyk *et al.* 1980). Inset map shows the plate boundary setting for western North America (Bird 2003).

We benefited from source characterizations from the following studies: Lin *et al.* (2007a), Lohman & McGuire (2007), Thurber *et al.* (2006), McLaren *et al.* (2008), Tan (2006), Hardebeck & Shearer (2003) and Clinton *et al.* (2006). We considered 234 events in total, 143 of which were used in the tomographic inversion. The remaining 91 events were used for an independent assessment of any given tomographic model (Tape *et al.* 2009). Both sets of earthquakes are shown in Fig. 2 along with the station coverage.

A summary of all 234 events is presented in Fig. 3. All earthquakes occurred between 1998 and 2009, with magnitudes between M_w 3.4 and M_w 5.4. The depths are between 1.5 and 21.0 km, with most earthquakes occurring in the upper 10 km. The majority of the earthquakes have half durations <0.4 s, which is small compared to the target measurement of $\sigma_0 = 1.0$ s (eq. 4). Using larger earthquakes, or a smaller value of σ_0, one would be wary of the approximation of the earthquake as a point source.

2.3 Model variables, model parametrization and model vector

The entries of the model vector $\mathbf{m}$ for the tomographic inversion describe the two continuous scalar fields $V_S(\mathbf{x})$ and $V_B(\mathbf{x})$, where $\mathbf{x}$ is a point in the volume. We parametrize these continuous fields using the same basis functions used in the numerical simulations, that is, Lagrange polynomials, similar to previous 2-D examples (Tape *et al.* 2007, Section 5.3). The simulation region is 639 km × 503 km at the surface and extends to 60 km depth (Fig. 1). The mesh contains 405 216 elements with a total of 27.5×10^6 global gridpoints and 50.6×10^6 local gridpoints. From the standpoint of inversion, the distinction between global and local gridpoints is only relevant if wave speed discontinuities exist in the model that are honoured by the mesh. For southern California, the Moho and parts of the basement surface are honoured (Komatitsch *et al.* 2004), and we therefore choose to use local gridpoints to parametrize the model.

Each local gridpoint in the numerical mesh, $\mathbf{x}_i$, has corresponding values of V_S and V_B that appear as entries of the model vector $\mathbf{m}$. Therefore, the model vector $\mathbf{m}$ has $2G$ entries, with G the number of local gridpoints in the mesh. We use a subscript to denote model iterations, such that $\mathbf{m}_{00}$ is the initial model and $\mathbf{m}_{16}$ is the final model. In Appendix B, we include details pertaining to the construction of both $\mathbf{m}$ and the covariance matrix, $\mathbf{C}_M$.

Based on the mesh described above, the average spacing between gridpoints at the surface is 450 m. The simulations use a time step of 0.011 s and are numerically accurate down to periods of approximately 2 s.

3 MISFIT FUNCTION

The formulation of the misfit function is a critical aspect of any tomographic inversion, and many choices are available (e.g. Tromp *et al.* 2005). We choose a traveltime misfit function that employs frequency-dependent measurements between observed and synthetic waveforms.

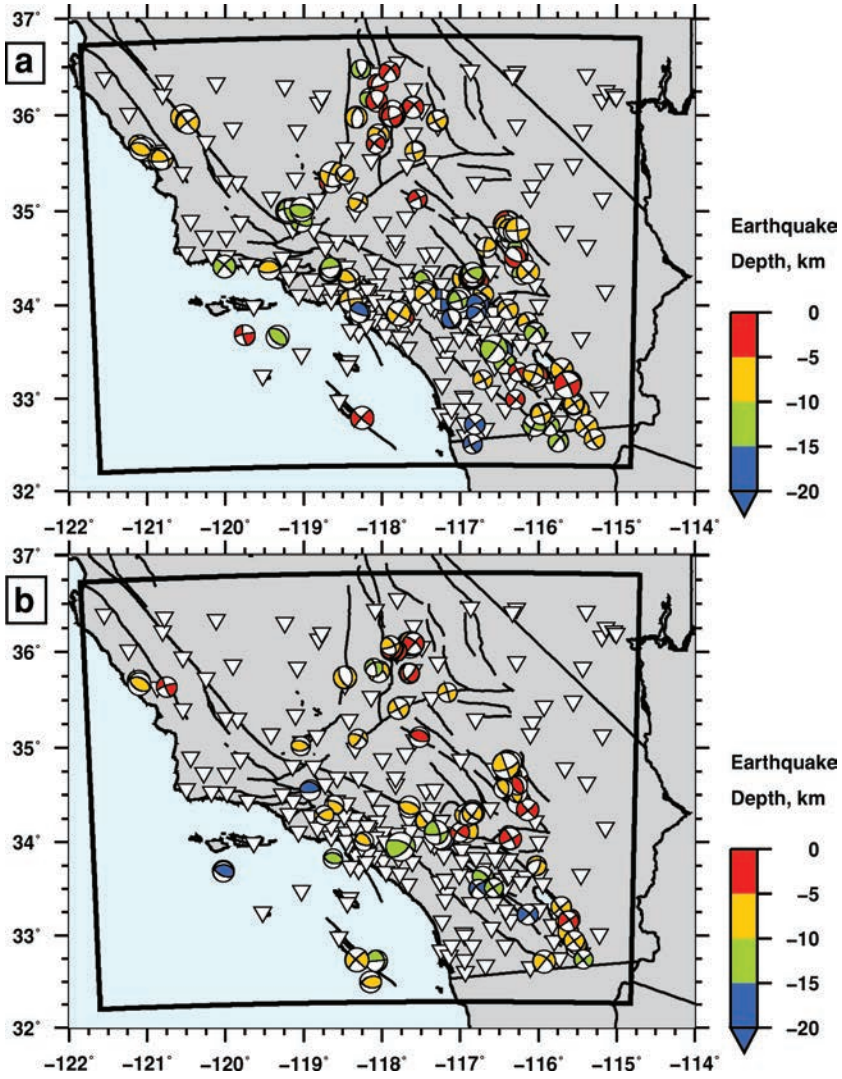


Figure 2. (a) Earthquakes (143) and stations (203) used in the tomographic inversion. The black boundary denotes the simulation region. (b) Extra earthquakes (91) used in validating the final tomographic model, but not used in the tomographic inversion. An earthquake not used in the tomographic inversion—or any future earthquake, for that matter—may be used to independently assess the misfit reduction from $\mathbf{m}_{00}$ to $\mathbf{m}_{16}$.

3.1 Collection of seismic waveforms

The computational cost of the tomographic technique scales linearly with the number of earthquakes, but is independent of the number of stations, components, or 'picks' (i.e. time windows for measurement). We collected seismic waveforms from 210 broad-band stations within the simulation region (Tables 1 and 2). We did not use available strong-motion stations, because these stations tend to be concentrated in populated regions where dense broad-band station coverage already exists. For each earthquake in the data set, the maximum number of recording stations is approximately 170. For earthquakes in the tomographic inversion, the maximum and minimum numbers of stations with measurements are 162 (event 14179736[1]) and 13 (event 9700049), respectively.

Over the course of analyzing tens of thousands of waveform comparisons for each model iteration, we discovered some systematic misfits that could not be attributed to source or structural effects, but only to station errors. We identified seismograms for seven station-epochs that produced what appeared to be a reversed polarity on either the horizontal components or on all three components: CRP.CI, HWB.AZ, BVDA2.AZ, PER.CI, BTP.CI, NSS2.CI and 109C.TA (see Tape 2009, Appendix E). Although the 'problematic' time intervals constitute a very small portion of the entire data set, the inclusion of such records would have a noticeable and detrimental impact on the tomographic inversion. For example, for a sinusoidal waveform, a polarity reversal corresponds to an apparent time shift of a half cycle that inadvertently could be interpreted as having arisen due to wave speed differences.

3.2 Selection of period ranges

The quality of fit between observed and synthetic seismograms is strongly dependent on the frequency content of the seismic waves, because the overall quality of the model generally diminishes with shortening scalelength. We therefore examine multiple period ranges: 6–30, 3–30 and 2–30 s. Our choice of period ranges emphasizes fitting seismic waveforms in the period range 6–30 s, which for

[1] The event numbers listed in this paper are from the Southern California Earthquake Data Center.

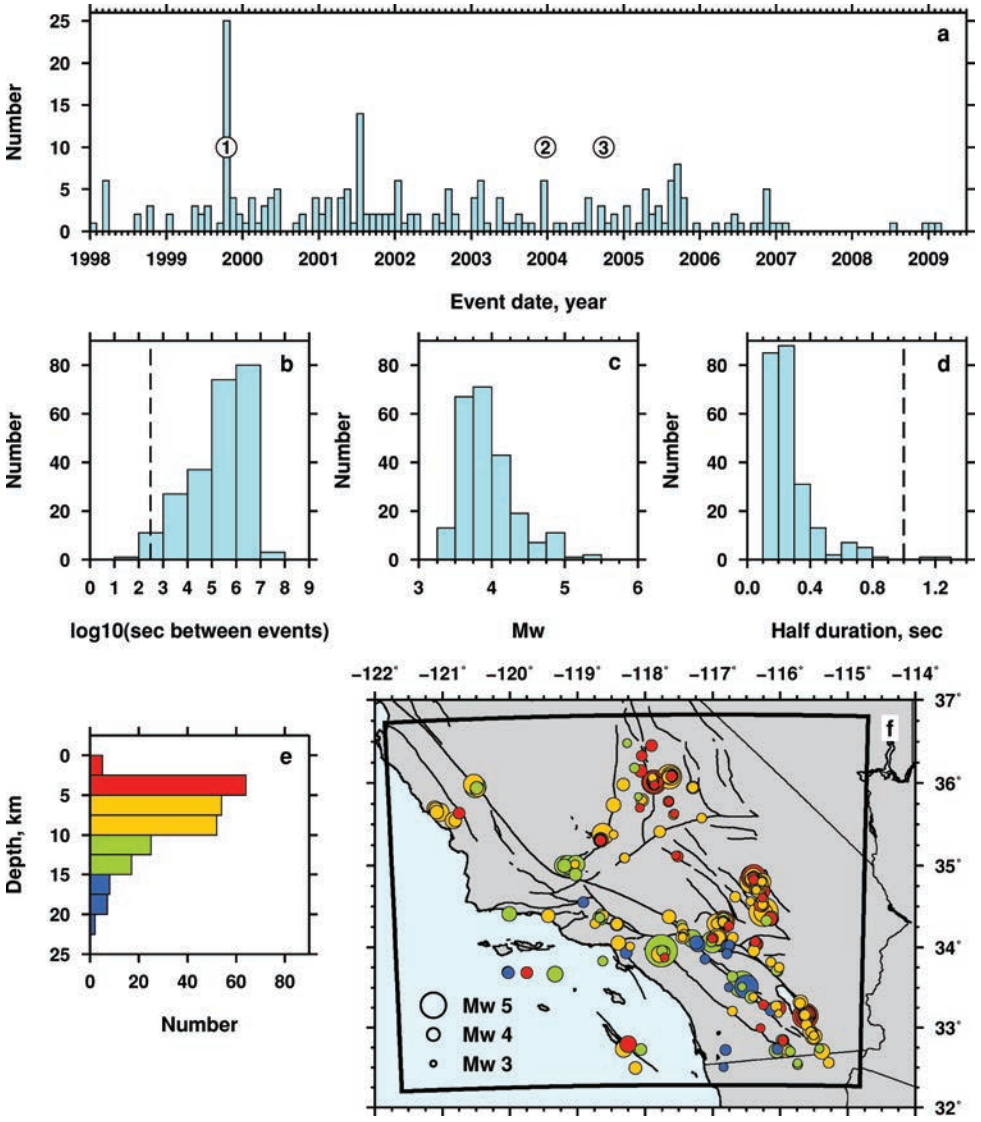


Figure 3. Summary of all 234 events considered in this study, including events used in the tomographic inversion (143) and 'extra' events (91). (a) Histogram of origin times of 234 events (bin-width = 1/12 yr). The circled numbers denote the origin times of three $M_w \geq 6$ earthquakes in southern California during this time interval: Landers 1999 October 16 M_w 7.1, San Simeon 2003 December 22 M_w 6.5, Parkfield 2004 September 28 M_w 6.0. These three earthquakes are not considered in this study, though several of their aftershocks are. (b) Histogram of the time interval between adjacent events, tabulated using $\log_{10}$ values. The dashed line denotes the maximum simulation length of 300 s ($\log_{10}$ 300 $\approx$ 2.48); in some cases, it is possible to consider events that occur close in time (and in space). (c) Histogram of moment magnitudes (M_w) (Kanamori 1977). (d) Histogram of half durations (τ_h) (Dahlen & Tromp 1998, p. 178). The dashed line denotes the target traveltime difference measurement value of $\sigma_0 = 1.0$ s (eq. 4) and indicates that the half durations are small relative to σ_0. (e) Histogram of depths. (f) Map showing epicentres, depths, and magnitudes of events. The depth colour scale is shown in (e).

crustal earthquakes in southern California, is dominated by surface waves. Furthermore, the relatively broad period range 2–30 s facilitates the identification of different waveforms on the basis of period content, as shown in Fig. 4(d). In these seismograms we identify short-period *P* and *S* arrivals (also identified in the 1-D synthetics in Fig. 4e) at 30 and 45 s, respectively, and a long-period waveform from 55 to 85 s that is the signature of a large sedimentary basin. If the seismograms were filtered over shorter periods, such as 2–6 s, the signature of the basin would be absent.

Table 3 summarizes measurements for the final tomographic model. As expected, seismograms filtered at longer periods (6–30 s) provide the most measurements (30 252). Seismograms containing the shortest period waveforms (2–30 s) are predominantly body waves and provide one-fourth of the measurements. Measurements of body waves are also common at longer periods (e.g. Fig. 9).

3.3 Selection of time windows

Our objective is to fit the entire seismogram, wiggle for wiggle. This includes body waves, surface waves and seismic waveforms that are not readily identifiable, such as the example shown in Fig. 5. In this example the observed seismogram exhibits resonance from both the Salton trough (near the source) and the Los Angeles basin (near the station). The overall characteristics of the observed waveform are nicely captured by the 3-D model ($\mathbf{m}_{16}$) but are virtually absent using a standard 1-D model. This three-component seismogram, like hundreds in the data set, does not contain readily identifiable waveforms such as 'the *P* wave' or 'the Love wave.' Nevertheless, it is possible to fit such waveforms, as long as a measurement between the synthetic and observed seismograms can be made.

Table 1. Seismograms used for measurements, sorted by network.

	Network	Data[a]	Stations	Seismograms	Windows
CI	Caltech Regional Seismic Network	SCEDC	171	58 463	68 800
AZ	ANZA Regional Network	IRIS	18	7469	8428
TA	USArray Transportable Array	IRIS	10	342	417
LB	Leo Brady Network	IRIS	1	342	394
BK	Berkeley Digital Seismic Network (BDSN)	NCEDC	3	188	228
NC	USGS Northern California Regional Network	NCEDC	2	66	74
NN	Western Great Basin/Eastern Sierra Nevada	UNR	4	44	61
SN	Southern Great Basin Network	UNR	1	27	29
			210	66 941	78 431

Note: The numbers consider only seismograms with measurements using $\mathbf{m}_{16}$ synthetics over all three period ranges and for all 234 earthquakes (Table 2).
[a]The data centre listed is the primary repository for each network's seismic waveforms: SCEDC, Southern California Earthquake Data Center; NCEDC, Northern California Earthquake Data Center; IRIS, Incorporated Research Institutions for Seismology; UNR, University of Nevada Reno.

Table 2. Summary of seismogram measurements for the final model $\mathbf{m}_{16}$.

	TOMO	EXTRA	COMBINED
Components (Z,R,T)	3	3	3
Networks	8	7	8
Earthquakes	**143**	**91**	**234**
Stations	203	200	210
Paths	12 583	4305	16 888
Seismograms (unique)	27 007	8013	35 020
Seismograms (total)	52 138	14 803	66 941
Windows (total)	61 673	16 758	78 431

Notes: 'TOMO' corresponds to the 143 earthquakes used in the tomographic inversion. 'EXTRA' corresponds to the 91 extra earthquakes not used in the tomographic inversion.'COMBINED' corresponds to the TOMO+EXTRA set of 243 earthquakes. The number of unique seismograms is indicated next to 'seismograms (unique)'. The number of total seismograms—including the three period ranges 6–30, 3–30 and 2–30 s—is indicated next to 'seismograms (total)'. The same is true for 'windows (total)'. A 'path' is a single source–station pair that has at least one measurement.

Table 3. Summary of seismogram measurement time windows for earthquakes used in the tomographic inversion ('TOMO').

	6–30 s	3–30 s	2–30 s	Total
Vertical (Z)	10 319	5623	4864	20 806
Radial (R)	9276	5443	4579	19 298
Transverse (T)	10 657	5684	5228	21 569
Total	30 252	16 750	14 671	61 673

Note: Each entry corresponds to the number of measurement windows for a particular period range (6–30, 3–30 and 2–30 s) for a particular component (Z, R and T).

We use an automated algorithm, FLEXWIN (Maggi *et al.* 2009), to select time windows for measurement for all seismograms in the data set. The objective of FLEXWIN is to select time windows within which there is quantifiable agreement between the observed and simulated waveforms. The algorithm requires several user parameters that need to be adjusted for a given data set. The specific set of user parameters for the initial model is listed in Maggi *et al.* (2009) and Table 3. In the tomographic inversion, as the model improved with each iteration, we made minor adjustments to the user parameters, as discussed in Tape *et al.* (2009, section S7) and listed in Tape *et al.* (2009, table S4).

There is a trade-off between the window-acceptance criteria of FLEXWIN and the amount of manual checking required. The two 'end members' are as follows.

(i) Use conservative window-acceptance criteria. This will result in many fewer window selections, and minimal manual checking is required. However, because less information in the seismograms is used, more iterations will be needed in the inversion.

(ii) Use liberal window-acceptance criteria. This will result in a maximal number of selected windows, but manual checking is needed to remove any undesirable picks. Fewer iterations in the inversion will be needed.

We used the latter approach, which allowed for the automatic selection of more windows, some of which contain exotic waveforms, such as the laterally reflected surface waves shown in Figs 14 and 26.

We emphasize that the selection of time windows is performed by FLEXWIN for each new set of synthetic seismograms for model $\mathbf{m}_k$ and for each of the three sets of bandpasses. The total number of seismograms 'seen' by FLEXWIN in this study is approximately 3.08 million: 3 (different bandpasses) $\times$ 141 (average number of broad-band stations available) $\times$ 3 (components) $\times$ 143 (earthquakes) $\times$ 17 (models $\mathbf{m}_{00}$ to $\mathbf{m}_{16}$). An automated procedure such as FLEXWIN makes it possible to manage such a large data set.

3.4 Misfit measures

Within each time window, we need to quantify the misfit between observed and synthetic waveforms. We consider two measures of misfit: a traveltime difference (F^{T}) and a waveform difference (F^{W}). We use the traveltime misfit measure within the tomographic inversion. However, to assess the misfit reduction we use the waveform misfit measure, because in many cases there is a waveform in the $\mathbf{m}_{16}$ synthetics to align with the data, but there is no corresponding waveform in the $\mathbf{m}_{00}$ synthetics (e.g. Fig. 10b).

Within each measurement window we choose to measure the frequency-dependent traveltime difference between observed and simulated seismic waveforms. For a single time window on a single seismogram, the misfit measure is given by

$$F_i^{\mathrm{T}}(\mathbf{m}) = \int_{-\infty}^{\infty} \frac{h_i(\omega)}{H_i} \left[\frac{\Delta T_i(\omega, \mathbf{m})}{\sigma_i} \right]^2 \mathrm{d}\omega, \tag{3}$$

where $\mathbf{m}$ is a model vector, $\Delta T_i(\omega, \mathbf{m}) = T_i^{\mathrm{obs}}(\omega) - T_i(\omega, \mathbf{m})$ denotes the frequency-dependent traveltime difference between

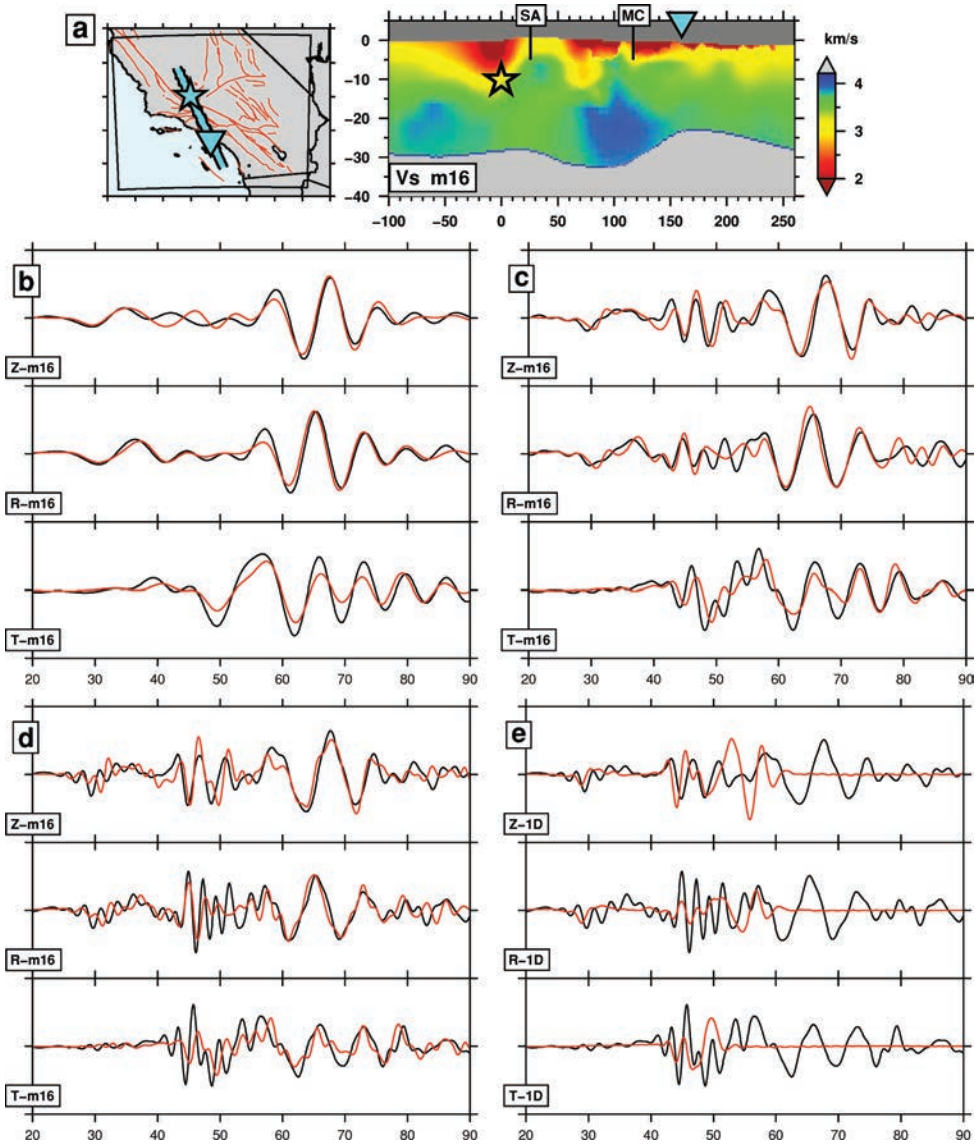


Figure 4. The frequency dependence of the seismic wavefield (Section 3.2). (a) Cross-section of the final V_S crustal model $\mathbf{m}_{16}$, containing the path from event 14186612 (⋆; M_w 4.4, depth 10.2 km) to station FMP.CI (∇; distance 159.6 km). SA, San Andreas fault; MC, Malibu Coast fault. (b) Data (black) and 3D synthetics for model $\mathbf{m}_{16}$ (red), bandpass filtered over the period range 6–30 s. Z, vertical component, R, radial component, T, transverse component. (c) Same as (b), but for the period range 3–30 s. (d) Same as (b), but for the period range 2–30 s. (e) Data (black) and 1-D synthetics (red), bandpass filtered over the period range 2–30 s.

observed and synthetic waveforms associated with the *i*th window, and σ_i is the estimated uncertainty associated with the traveltime measurement. The traveltime measurement is made using a multitaper method (e.g. Percival & Walden 1993; Laske & Masters 1996; Zhou *et al.* 2004), and $h_i(\omega)$ is a frequency-domain window with associated normalization constant $H_i = \int_{-\infty}^{\infty} h_i(\omega)\,\mathrm{d}\omega$. In the case of a frequency-independent measurement, $\Delta T_i(\omega, \mathbf{m})$ reduces to a cross-correlation traveltime measurement, and $F_i^{\mathrm{T}}(\mathbf{m}) = [\Delta T_i(\mathbf{m})/\sigma_i]^2$. The adjoint source associated with a multitaper misfit measurement is derived in Tape (2009, Appendix C).

The misfit function (eq. 3) contains an uncertainty estimate for each measurement (Appendix A). Because σ_i is in the denominator, it is necessary to provide a 'water-level' minimum value, σ_0, such that $\sigma_i \geq \sigma_0$. For the tomographic inversion, we chose

$$\sigma_0 = 1.0 \text{ s} \tag{4}$$

based on the range of M_w and the estimated uncertainties of source parameters, primarily hypocentre and origin time (e.g. Lin *et al.* 2007a). The tomographic inversion emphasizes fitting synthetic and observed waveforms whose traveltime differences exceed σ_0.

For a single time window on a single seismogram, the waveform misfit measure is given by

$$F_i^{\mathrm{W}}(\mathbf{m}) = \frac{\int_{-\infty}^{\infty} w_i(t)\,[d(t) - s(t,\mathbf{m})]^2\,\mathrm{d}t}{\left\{\int_{-\infty}^{\infty} w_i(t)\,[d(t)]^2\,\mathrm{d}t \int_{-\infty}^{\infty} w_i(t)\,[s(t,\mathbf{m})]^2\,\mathrm{d}t\right\}^{1/2}} \tag{5}$$

where $d(t)$ denotes the observed time-series, $s(t, \mathbf{m})$ the simulated time-series, $w_i(t)$ a time-domain window, and i the measurement window index. Eq. (5) contains the same normalization as the standard cross-correlation formula and has been used for source inversions (Zhu & Helmberger 1996, eq. 3).

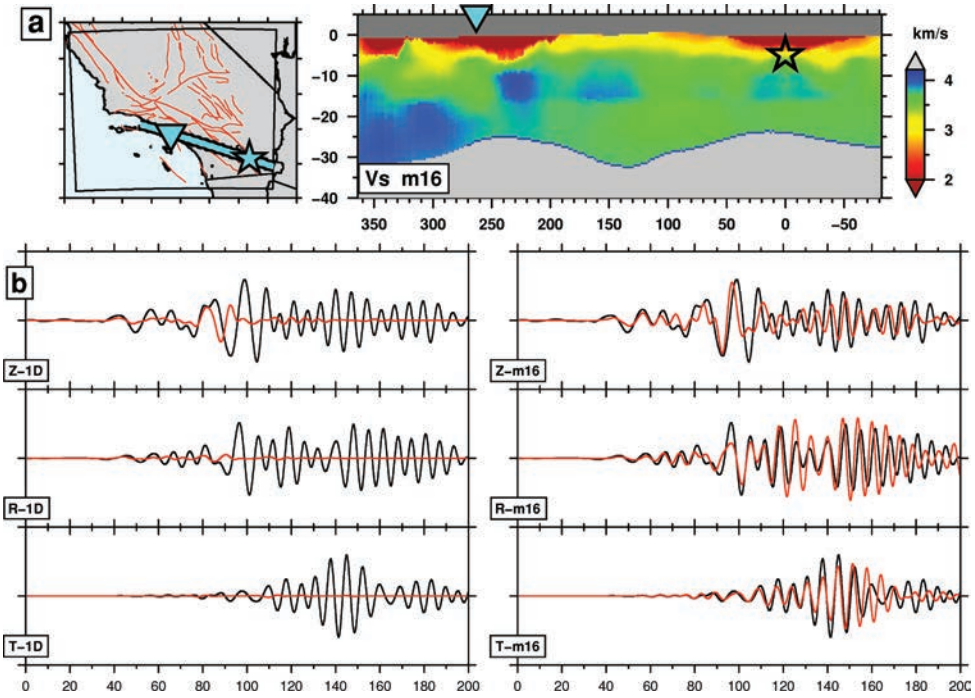


Figure 5. The influence of sedimentary basins on the seismic wavefield. (a) Cross-section of the final V_S crustal model $\mathbf{m}_{16}$, containing the path from event 14179736 ($\star$; M_w 5.0, depth 4.9 km), beneath the Salton trough, to station LAF.CI (∇; distance 263.5 km), within the Los Angeles basin. (b, left column) Data (black) and 1-D synthetics (red). (b, right column) Data (black) and 3-D synthetics for model $\mathbf{m}_{16}$ (red). The seismograms are bandpass filtered over the period range 6–30 s. Z, vertical component, R, radial component, T, transverse component.

Many other choices of misfit measures are possible, as exemplified in Tromp *et al.* (2005). For example, Luo & Schuster (1991) used a cross-correlation traveltime difference, while Fichtner *et al.* (2008, 2009) windowed and weighted waveforms both in the time and frequency domains prior to measuring phase and amplitude differences. Our choice to measure phase (or traveltime) differences was motivated by the success of such measurements in seismic tomography (e.g. Humphreys & Clayton 1990; Grand *et al.* 1997; Ritsema *et al.* 1999).

3.5 Misfit function

The misfit function for the tomographic inversion is based on the individual traveltime misfit measure given in eq. (3). For a single earthquake, the misfit function is defined by

$$F_s^{\mathrm{T}}(\mathbf{m}) = \frac{1}{2}\frac{1}{N_s}\sum_{i=1}^{N_s} F_i^{\mathrm{T}}(\mathbf{m}) \,, \tag{6}$$

where N_s denotes the total number of measurement windows associated with source s. The ith window is identified by a source, a station, a component, a period range (e.g. 6–30 s), and a local window index. The local window index ranges up to the maximum number of windows for a particular filtered seismogram; for this data set, there are typically 0–2 windows per filtered seismogram.

The overall misfit function F^{T} is simply

$$F^{\mathrm{T}}(\mathbf{m}) = \frac{1}{S}\sum_{s=1}^{S} F_s^{\mathrm{T}}(\mathbf{m}) \,, \tag{7}$$

where S is the number of sources. We do not impose any manual weighting, but such choices would impact the tomographic inversion (e.g. Fichtner *et al.* 2009).

4 GRADIENT OF MISFIT FUNCTION AND ITERATIVE INVERSION PROCEDURE

The method we use to iterate from the initial model ($\mathbf{m}_{00}$) to the final model ($\mathbf{m}_{16}$) is adapted from the approach illustrated in Tape *et al.* (2007). The Tape *et al.* (2007) study involved a 2-D tomographic problem using only synthetic seismograms, whereas the current study is 3-D and uses real data. In the 3-D problem, the model vector $\mathbf{m}$ contains two variables that describe the structure, V_S and V_B. Our measure of misfit is a frequency-dependent multitaper traveltime difference, made in three overlapping period ranges (6–30, 3–30, 2–30), as discussed in Section 3.2.

A distinguishing feature of adjoint tomography is that the gradient of F_s^{T} (eq. 6), or the 'event kernel,' is computed from the interaction between two wavefields: the 'regular' forward simulation emanating from the source, and the 'adjoint' simulation emanating from the stations (Tarantola 1984; Tromp *et al.* 2005; Tape *et al.* 2007). For each iteration, there is a new set of event kernels, which are summed in some manner to produce the next model update. In Tape *et al.* (2007) we combined the event kernels by summing them, and then used a conjugate-gradient algorithm to obtain a model update. In this study we use a subspace projection technique to compute the model update. Instead of using a subspace of model parameter classes (Kennett *et al.* 1988; Sambridge *et al.* 1991), we use a subspace of sources, where we determine a linear combination of event kernels that exploits the features they have in common (Tape *et al.* 2010). This procedure provides a pre-conditioner for the gradient algorithm that increases convergence of the minimization problem. Not all 143 earthquakes were used in each iteration (Tape *et al.* 2009, table S3), because a satisfactory model update could generally be achieved with a well-chosen subset of event kernels.

For each model iteration we choose a step length such that the maximum local absolute values of $\ln(\mathbf{m}_{k+1}/\mathbf{m}_k)$ are less than approximately 0.10 (i.e. ± 10 per cent changes between successive iterations). We test the new model by computing new synthetics for 4–8 representative events that sample the volumetric model update. With the new synthetics and new measurements, we verify that the step length is not too large or too small. A large step means that the model will step *back* in the next iteration, while a small step means that additional (and unnecessary) steps will be taken.

4.1 Regularization

Our tomographic inversion contains three forms of regularization.

(i) The choice of bandwidth for the measurements.
(ii) The choice of smoothing of the event kernels.
(iii) The choice of regularization parameter in solving for the linear combination of event kernels to form the model update, which we address in Tape *et al.* (2010).

The choice of bandwidth for the measurements is an important but subtle form of regularization. By allowing shorter periods for measurement, we introduce more detailed features in the event kernels. With enough event kernels, we should be able to incorporate these detailed features directly into the model update. We adopted a conservative approach with the primary objective of not introducing any artificial short scalelength features. Thus, we smoothed each event kernel prior to combining them to form the model update.

We smoothed each event kernel via convolution with a Gaussian function, analogous to the 2-D examples in Tape *et al.* (2007). For the initial iteration, we chose a very conservative spherical Gaussian with a ∼28 km full-width to smooth the event kernels. By 'conservative,' we mean that the scalelength of the smoothing procedure is considerably larger than the minimum scalelength of variations within the event kernels, which are as short as 2 km.

As the model improved, we used less smoothing and therefore included more details of the event kernels. For the last eight iterations, the scalelengths of smoothing were 17.0 km in the horizontal direction and 2.8 km in depth.[2]

In the absence of smoothing, an additional form of regularization is the number of iterations that are performed. Additional iterations will ultimately introduce the finest details contained within the event kernels. In our approach, we have chosen to smooth the event kernels in order to avoid introducing unresolved finer structure in the tomographic model.

4.2 Computational demands

The tomographic inversion is computationally expensive due to the large number of simulations needed to evaluate the misfit function and its gradient at every iteration. The forward simulation technique is summarized in Section 2.3. For each earthquake we calculated up to five minutes of seismograms that are accurate down to a period of two seconds. For each earthquake we performed three simulations, one to evaluate the misfit function (eq. 6) and two to compute its gradient (the 'event kernel'). Each model iteration thus required $3S_k$ simulations, where S_k is the number of earthquakes used for the kth iteration. The total number of 168-core simulations used in producing model $\mathbf{m}_{16}$ was 6794. A summary of the simulations is presented in Tape *et al.* (2009, table S3). We used a total of 0.80 million CPU-hours for the 6794 simulations. This tally does not include computations devoted to smoothing the individual event kernels (Section 4.1).

[2] The scalelength is computed as $\Gamma = \sigma\sqrt{8}$, where σ defines the Gaussian, for example, for 1-D, $G(x) = (2\pi\sigma^2)^{-1/2}\ e^{-x^2/(2\sigma^2)}$. The listed values of Γ correspond to $\sigma_x = \sigma_y = 6$ km and $\sigma_z = 1$ km.

5 NEW MODEL ($\mathbf{m}_{16}$)

5.1 Presentation of tomographic models

We display the new model on both relative and absolute scales. The relative scale is described by the net model update, $\ln(\mathbf{m}_{16}/\mathbf{m}_{00})$, denoting entry-by-entry operations.[3] The net model update contains only features that are required by the data which are manifested in the event kernels. The absolute scale is described by $\mathbf{m}_{16}$, which represents the V_S and V_B wave speed values in units of km s^{-1}. This version of the model is relevant for geological and geodynamic interpretations. All cross-sections discussed below are of shear wave speed (V_S) models ($\mathbf{m}_{00}$ and $\mathbf{m}_{16}$). The bulk sound speed model is discussed in Section 5.4.

We present the tomographic models $\mathbf{m}_{00}$, $\mathbf{m}_{16}$, and $\ln(\mathbf{m}_{16}/\mathbf{m}_{00})$ in a series of horizontal and vertical cross-sections. For the vertical cross-sections, the emphasis is as much on the seismograms as it is on the details of the tomographic model. The term 'full waveform' refers to the record length from before the *P*-wave arrival (preferably from $t = 0$) until after the surface waves have qualitatively disappeared (this will depend on the particular bandpass). Fits to full waveforms on all three components provide compelling support for the inferred crustal structure along the corresponding path. However, we note that many documented wave propagation effects occur far from the vertical cross-section containing the source and station. In other words, one should not look strictly in the vertical cross-sections for explanations of all details in seismic waveforms (e.g. Fig. 27).

In each comparison of seismograms, the black time-series are the observed seismograms and the red time-series are the synthetic seismograms. The synthetics are computed for the initial 3-D model ($\mathbf{m}_{00}$), the final 3-D model ($\mathbf{m}_{16}$), and a standard 1-D model for southern California (Kanamori & Hadley 1975; Dreger & Helmberger 1990; Wald *et al.* 1995) (listed in Tape *et al.* 2009, table S5). Synthetics for the 1-D model were never used in the tomographic inversion, but they are helpful in identifying waveforms that are strongly influenced by 3-D structure. For illustrative purposes, we use one of three options for highlighting particular measurement windows.

(i) Plot the windows selected by FLEXWIN (Maggi *et al.* 2009): Figs 9, 10 and 26.
(ii) Plot no windows at all, to emphasize the full seismogram: Figs 4, 5, 8, 12, 13, 16 and 24.
(iii) Plot only particular windows for discussion in the text: Figs 11, 14, 15 and 26.
All seismograms—data and synthetics—are displacement records.

In the vertical cross-sections, topography is shown at the plotting resolution, which is coarser than the numerical mesh used in the simulations. The colour scale for the V_S models is fixed for all

[3] A first-order approximation of the natural logarithm leads to the formula for a percent difference, that is, $\ln(m_2/m_1) \approx (m_2 - m_1)/m_1$. This is the standard presentation of a tomographic model as a perturbation from a 1-D layered reference model (m_1).

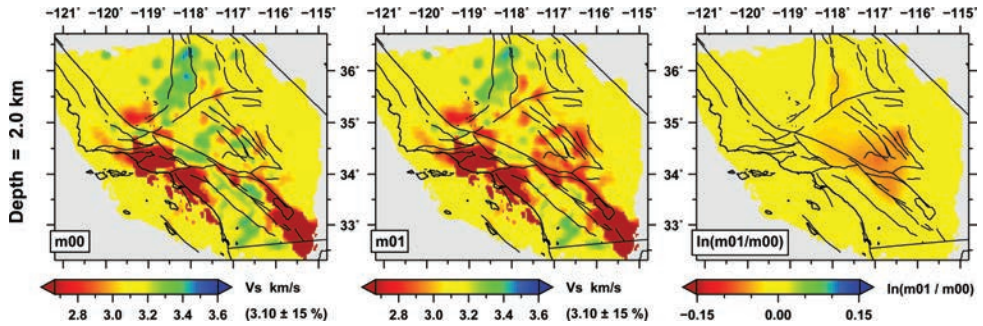


Figure 6. First iteration of the V_S tomographic model, showing a horizontal cross-section at a depth of 2 km for the initial model, $\mathbf{m}_{00}$, the first new model, $\mathbf{m}_{01}$, and the difference between the two models, $\ln(\mathbf{m}_{01}/\mathbf{m}_{00})$. Compare with Fig. 7(a).

cross-sections and ranges from 2.0 to 4.2 km s^{-1}. The colour scale is saturated dark grey for upper mantle values of $V_S \geq 4.2$ km s^{-1} in order to emphasize crustal wave speeds. The simulations extend to 60 km depth, but relatively few changes are made beneath the crust. The minimum wave speed is 0.6 km s^{-1} in the tomographic model.

5.2 Seismograms and cross-sections

Horizontal cross-sections convey lateral variations in the tomographic models. The first iteration, from $\mathbf{m}_{00}$ to $\mathbf{m}_{01}$, is shown in Fig. 6 as a cross-section at a depth of 2 km. The model update, $\ln(\mathbf{m}_{01}/\mathbf{m}_{00})$, is conspicuously simple: a single red blob covering the Mojave Desert, indicating that the wave speeds in this region of the initial model need to be decreased.

After 15 additional iterations, we obtain the results in Fig. 7, shown for depths of 2, 10 and 20 km. The net changes in the model are large, locally in excess of ±30 per cent, and are strongest in the upper 20 km. The deepest changes occur beneath the western Transverse Ranges. The strongest wave speed reductions occur within the Coast Ranges, the mid-crust of the Mojave, and the southern San Joaquin basin, where the anomaly is −35 per cent. The strongest wave speed increases occur beneath the Santa Monica Mountains and the western Transverse Ranges. We interpret Fig. 7 in detail in Section 5.3.

The lack of correspondence between $\ln(\mathbf{m}_{01}/\mathbf{m}_{00})$ (Fig. 6) and $\ln(\mathbf{m}_{16}/\mathbf{m}_{00})$ (Fig. 7a) serves as a reminder of how many iterations may be needed to obtain an accurate crustal model for the period range of interest. The only regions in common to both are the slow wave speeds in the eastern Mojave and Coso regions (Fig. 1).

5.2.1 Nine three-component seismograms

Each model $\mathbf{m}_k$ produces a different set of synthetic seismograms. We monitor the seismogram changes with each iteration, as illustrated in Tape *et al.* (2009, Fig. 1) and in Fig. 8. The following nine vertical cross-sections with corresponding seismograms are selected to demonstrate specific points relevant to the tomographic inversion (Figs 8–16).

Fig. 8 is an expanded version of Tape *et al.* (2009, Fig. 2). The observed three-component seismogram contains complex waveforms that are not readily identifiable (e.g. 'the Rayleigh wave'). By comparing the 1-D and 3-D synthetics, it is apparent that the Los Angeles and Ventura basins strongly influence the wavefield. Comparisons between synthetics from $\mathbf{m}_{00}$ and $\mathbf{m}_{16}$ reveal improvements in both the phase and amplitude of the waveforms, even though only phase measurements (traveltime differences) appear in the misfit function. This demonstrates that 3-D structural changes to the initial model induce additional focusing and amplification of the seismic wavefield. Furthermore, it suggests that relatively smooth tomographic perturbations can help improve the model in the vicinity of higher-resolution sedimentary basins.

The three observed seismograms in Fig. 9 exhibit the basic waveforms for the period range 6–30 s. The automated windowing code, FLEXWIN, picks five windows highlighting the *P* wave (Z and R components), Rayleigh wave (Z and R) and Love wave (T). (The *S* arrival is apparent on all three components at $t = 45$ s on the seismograms filtered over the period range 2–30 s.) In comparison to the 1-D synthetics, the synthetics for $\mathbf{m}_{00}$ exhibit better fits to the Rayleigh and Love phases. The synthetics for $\mathbf{m}_{16}$ capture the correct phase and amplitude of the Rayleigh wave, and exhibit a striking improvement over the $\mathbf{m}_{00}$ synthetics.

The earthquake shown in Fig. 9 occurred on the Garlock fault in January 2009 and was not used in the tomographic inversion, nor was any other comparable earthquake. Thus, in this case, the strong perturbations in the vicinity of the source are required by kernels associated with other earthquakes used in the tomographic inversion.

We note two points regarding the FLEXWIN measurement windows in Fig. 9. First, one might be wary of the large phase and amplitude misfits for the Rayleigh wave on the Z and R components of the 1-D and $\mathbf{m}_{00}$ seismograms. Based on this one seismogram alone, one would probably want to adjust the FLEXWIN parameters to exclude such measurement windows. Second, there is a considerable level of noise in these seismograms, evidenced by the non-zero amplitudes at the origin time of the earthquake ($t = 0$ s). Nevertheless, the seismic waveforms have a large enough signal-to-noise ratio to be selected by FLEXWIN.

After many iterations it is possible to fit seismic waveforms for which there were no previous measurements. Figs 10 and 11 show seismograms for an earthquake on the San Andreas fault near Parkfield, recorded at stations SMM (distance 89.0 km) and WGR (distance 199.8 km). For both stations the Rayleigh wave (Z and R) is not identifiable in the initial 3-D synthetics but is well matched in phase and amplitude by the final 3-D synthetics. This provides confidence in interpreting the slow wave speeds identified throughout the Coast Ranges and, in particular, near the San Andreas fault.

Despite the predominant 'slowing' of the initial 3-D model, as indicated in Fig. 7, there are some regions that require substantial increases in wave speeds. Fig. 12 shows such a path along coastal California, from the western Transverse Ranges to the vicinity of San Diego. The waveform fits for $\mathbf{m}_{16}$ are among the best in the entire data set and capture clear basin resonance effects, probably arising from the Ventura and Los Angeles basins. Fits in the period ranges 3–30 and 2–30 s are comparably good (Tape 2009, fig. 1.5).

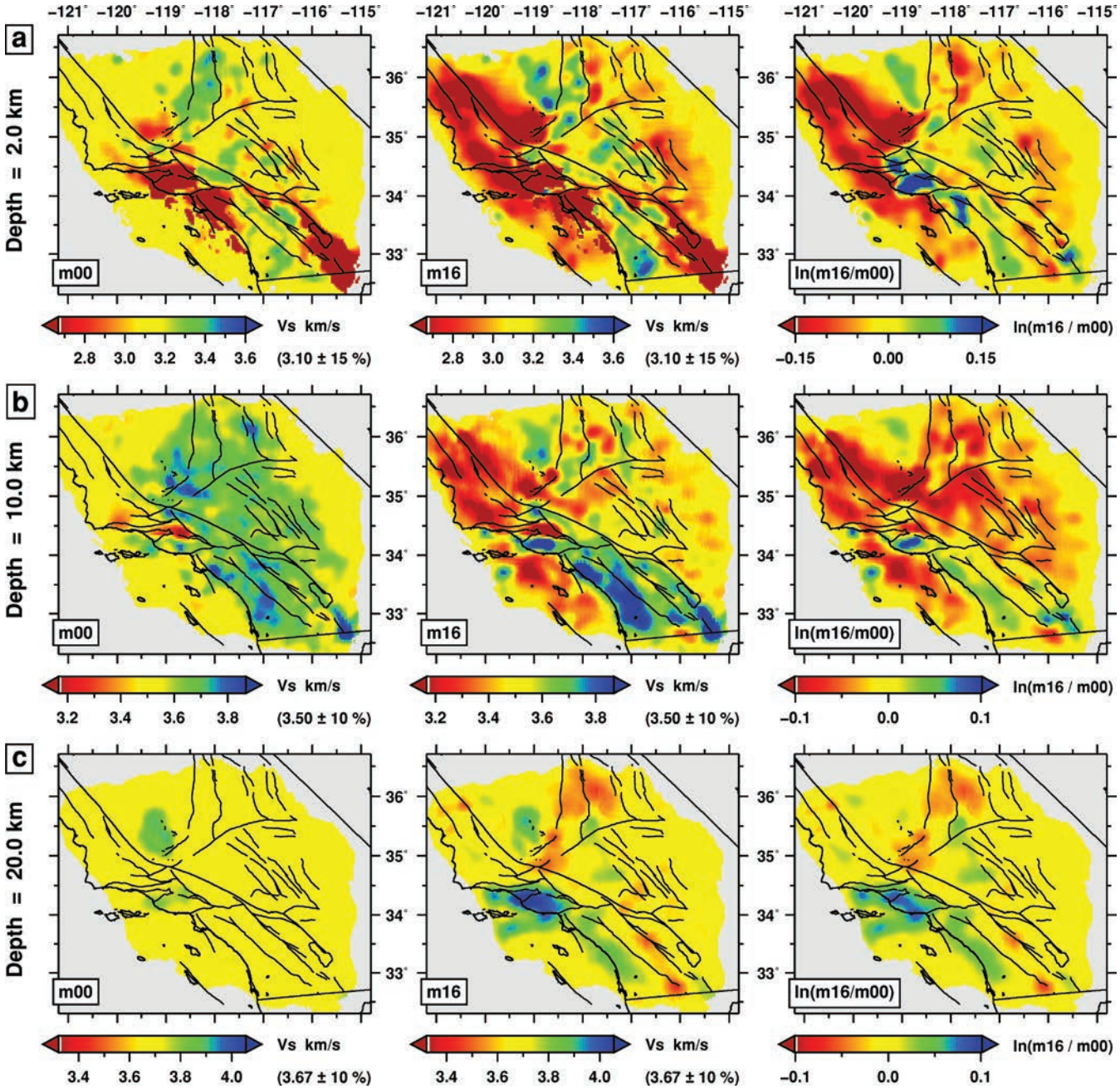


Figure 7. Horizontal cross-sections of V_S tomographic models, with active faults shown for reference. Refer to Fig. 1 for locations of principal faults and features. Left column shows the initial model, $\mathbf{m}_{00}$, centre column shows the final model, $\mathbf{m}_{16}$, and right column shows the difference between the two models, $\ln(\mathbf{m}_{16}/\mathbf{m}_{00})$. The mask covers regions of low sensitivity to changes in V_S. (a) V_S at 2 km depth. (b) V_S at 10 km depth. (c) V_S at 20 km.

An iterative tomographic inversion presents the challenge of fitting seismic waveforms while not 'unfitting' others. Fig. 13 shows a path that is very well fit in the initial 3-D model. Though not much improvement would be expected, we obtain somewhat better fits for the radial component. Although the before-and-after waveforms are similar, we still observe perturbations between −5 per cent and 15 per cent along the source–station path, clearly induced by other measurements (from other earthquakes).

We also capture wave propagation effects that occur far from the direct path between the earthquake and station. The seismograms in Fig. 14 contain shorter period waveforms (3–30 s) that strongly interact with adjacent basins. The three windows highlight three different Rayleigh-wave paths from the earthquake (near Hollywood) to station RVR. The second and third time windows contain Rayleigh waves interacting with the Los Angeles basin, south of the direct path between the source and station. Surface waves travelling multiple paths have been identified from lateral reflections (e.g. Stich *et al.* 2009) as well as from strong vertical heterogeneity (Savage & Helmberger 2004). Fig. 14 displays a definitive case of lateral reflections from a deep (>5 km) off-path sedimentary basin. Inclusion of such waveforms in the tomographic inversion shows we can increase the coverage by exploiting additional information already present in the seismograms.

Waveform fits for the shortest period range, 2–30 s, are shown in Figs 15 and 16. Most of the body wave pulses in Fig. 15 fit the observed pulses to within 1 s, the target value (eq. 4). The

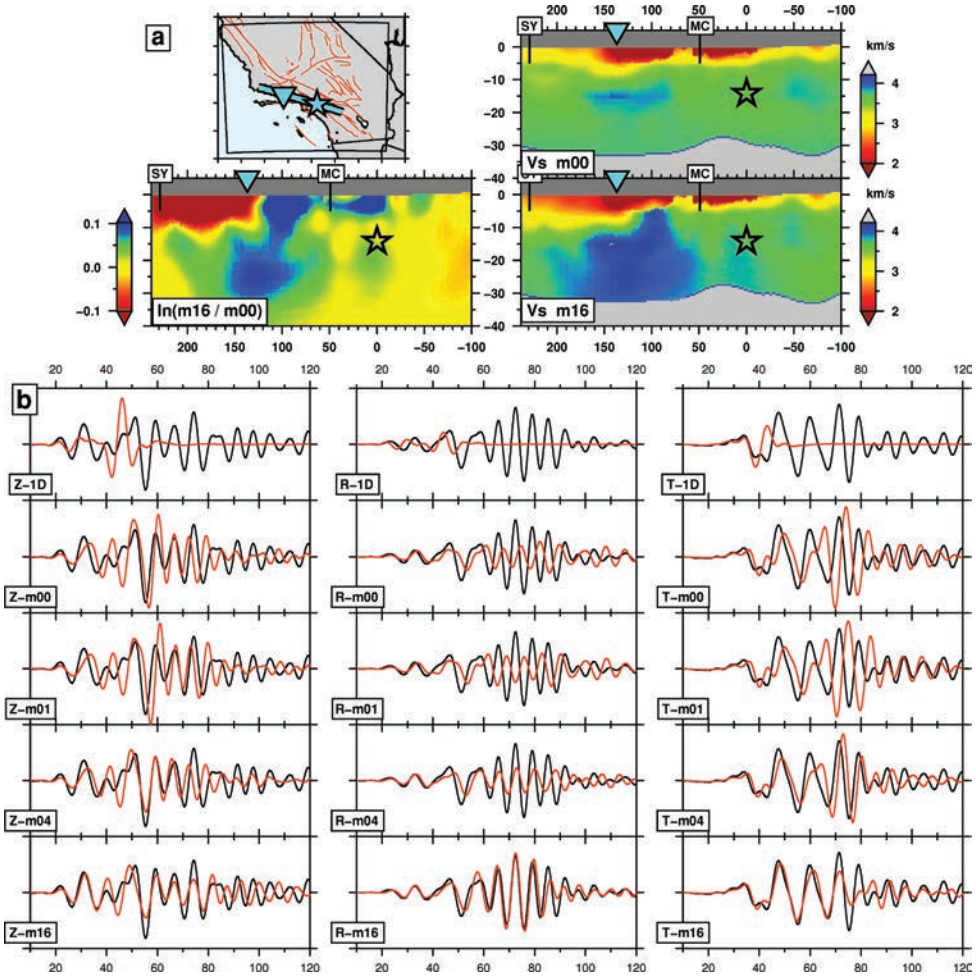


Figure 8. (a) Vertical cross-section of the V_S tomographic models for a path from event 14383980 ($\star$; M_w 5.4, depth 14.2 km), east of the Los Angeles basin, to station STC.CI (∇; distance 137.1 km), within the Ventura basin. Upper right panel is the initial 3-D model, $\mathbf{m}_{00}$, lower right panel is the final 3-D model, $\mathbf{m}_{16}$, and lower left panel is the difference between the two, $\ln(\mathbf{m}_{16}/\mathbf{m}_{00})$. The vertical exaggeration in these cross-sections, and all cross-sections in the paper, is 3.0. SY, Santa Ynez fault; MC, Malibu Coast fault. (b) Iterative three-component seismogram fits to data for models $\mathbf{m}_{00}$, $\mathbf{m}_{01}$, $\mathbf{m}_{04}$, and $\mathbf{m}_{16}$. Also shown are synthetic seismograms computed for a standard 1-D model. Synthetic seismograms (red) and observed seismograms (black), bandpass filtered over the period range 6–30 s. Left column, vertical component (Z); centre column, radial component (R); right column, transverse component (T). This earthquake was not used in the tomographic inversion.

highlighted downward pulse at 33 s on the transverse component $\mathbf{m}_{16}$ synthetic seismogram is not apparent in the corresponding $\mathbf{m}_{00}$ synthetic. The improved fits are partly due to structural perturbations that lie out of the vertical source–station plane.

Seismograms for a path crossing the entire Mojave are shown in Fig. 16. The $\mathbf{m}_{16}$ synthetic seismograms match the Love wave at 68 s (T), the *P* wave at 30 s (Z and R), the Rayleigh wave at 75 s (Z and R), and some additional complexity, particularly on the radial component. Some of this complexity is probably due to the excitation of the Los Angeles basin west of the source. The cross-section of $\ln(\mathbf{m}_{16}/\mathbf{m}_{00})$ reveals vertical variations with scalelengths of about 3 km in the vertical direction within the Mojave region. This demonstrates the resolving capability of frequency-dependent surface wave measurements.

5.3 Connections with geology and tectonics

A comprehensive interpretation of the new crustal model is beyond the scope of this paper. Such an interpretation of crustal wave speeds should take into account previous laboratory measurements of V_P and V_S of various crustal rocks under different pressure and temperature conditions (Christensen & Mooney 1995; Pellerin & Christensen 1998), as well as a geological and tectonic overview of the region (Burchfiel *et al.* 1992). In this section, we extend the interpretations of Tape *et al.* (2009) and address some of the key features of the crustal model.

5.3.1 *Wave speed contrasts across faults*

The middle panels of the horizontal cross-sections (Fig. 7) reveal lateral variations in the new crustal model. Because several active faults offset different lithologies at the surface, it is not surprising that the contrasts are also observed in the tomographic model. The magnitude and depth-extent of these contrasts are important for interpreting the structure of faults at depth. Furthermore, such wave speed contrasts are expected to influence the rupture of earthquakes (e.g. Ben-Zion & Andrews 1998).

At 2 km depth (Fig. 7a), the fastest regions (>3.5 km s^{-1}) occur in the Peninsular Ranges west of the Elsinore fault, and in the Sierra Nevada west of the Kern Canyon fault (Shapiro *et al.* 2005). The eastern front of the Sierra Nevada is marked by an eastward step in wave speed from about 3.5 to 2.8 km s^{-1} (Tape *et al.* 2009). At

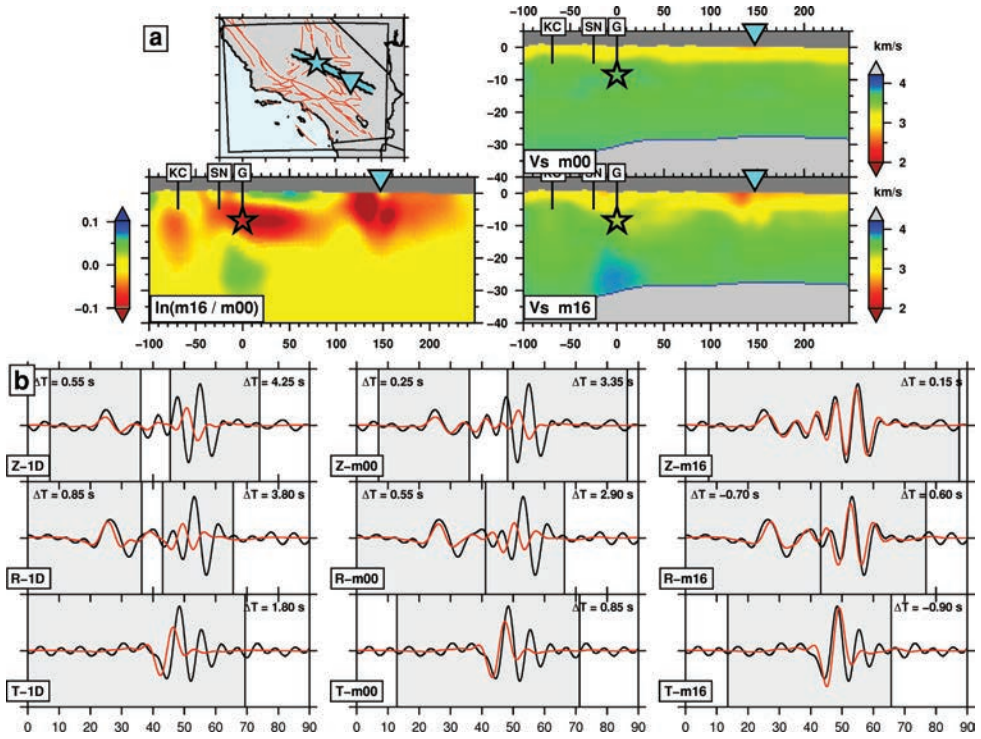


Figure 9. Vertical cross-sections and corresponding three-component seismograms (Section 5.2.1). (a) Cross-section of the V_S tomographic models for a path from event 14418600 ($\star$; M_w 3.9, depth 8.5 km), on the Garlock fault, to station HEC.CI (∇; distance 147.3 km), in the eastern Mojave Desert. Upper right panel is the initial 3-D model, $\mathbf{m}_{00}$, lower right panel is the final 3-D model, $\mathbf{m}_{16}$, and lower left panel is the difference between the two, $\ln(\mathbf{m}_{16}/\mathbf{m}_{00})$. KC, Kern Canyon fault; SN, Sierra Nevada fault; G, Garlock fault. (b) Synthetic seismograms (red) and observed seismograms (black) for the period range 6–30 s. The left column shows synthetics computed using a 1-D model for southern California. The centre column shows synthetics computed using our initial 3-D model $\mathbf{m}_{00}$. The right column shows synthetics computed using our final 3-D model $\mathbf{m}_{16}$. Z, vertical component, R, radial component, T, transverse component. In this example, the highlighted time windows, each with a corresponding ΔT traveltime difference, are automatically selected for measurement using the FLEXWIN algorithm (Maggi *et al.* 2009). This earthquake was not used in the tomographic inversion.

4 km depth (Fig. 27c), the northern segment of the San Jacinto fault separates faster wave speeds in the west (3.7 km s^{-1}) from slower wave speeds (3.3 km s^{-1}) beneath the San Bernadino basin.

At 10 km depth (Fig. 17) and south of about 34° latitude, the San Andreas fault separates a fast region (3.8 km s^{-1}) to the west from slower regions (3.4 km s^{-1}) to the northeast (Tape *et al.* 2009). The northwestern Mojave is slow (3.3 km s^{-1}) compared to faster material (3.6 km s^{-1}) in the southern Sierra Nevada, across the Garlock fault. Faults also partition the 50-km-scale variations in wave speeds along the longitudinal line 119°W (Fig. 17).

5.3.2 Deep sedimentary basins

At 2 km depth (Fig. 7a), large-scale slow regions (<2.8 km s^{-1}) reveal several known Neogene basins (Fig. 1a), most of which are present in the initial model. The new model illuminates the southern San Joaquin basin (SSJB; also known as the 'Maricopa subbasin'), which, according to seismic reflection profiles, extends to a depth of more than 8 km (Goodman & Malin 1992). The San Joaquin basin, as a whole, is a relic forearc basin associated with Farallon subduction and the Sierra Nevada arc. The anomalously deep portion at the southern end is probably a consequence of compression in the Big Bend region of the San Andreas fault, as well as delamination of the Sierra Nevada crust (Le Pourhiet *et al.* 2006; Fay *et al.* 2008).

The crustal model is able to match the arrival times of surface waves propagating from the San Simeon region (Hauksson *et al.* 2004; McLaren *et al.* 2008) in the western Coast Ranges, and *through* the SSJB to the eastern Mojave. We also capture lateral surface wave reflections at the southern boundary of the SSJB (Tape 2009, fig. 1.6), indicating the lateral sharpness of the wave speed contrast from the basin to the Transverse Ranges (Fig. 4a). The profile in Fig. 4(a) contains parts of three basins: SSJ, Ventura and Los Angeles.

The model is beginning to illuminate some of the basins in the offshore region of the Continental Borderlands (Legg 1991), a region that may accommodate 10–20 per cent of the relative motion between the Pacific and North America plates (Legg *et al.* 2007). In particular, we identify a clear extension of the Ventura basin into the Santa Barbara basin to the west (Fig. 7a). Additional off-shore seismic surveys (e.g. Nazareth & Clayton 2003) and ocean-bottom seismometers will help provide data needed to improve the structure model in the Borderlands.

5.3.3 Coast Ranges

The Coast Ranges, as a whole, form a strikingly slow (<3.0 km s^{-1}) anomaly (Fig. 7). In the uppermost 15 km, the Coast Ranges are tomographically similar to the San Joaquin basin and the Continental Borderlands (Tape *et al.* 2009). The Coast Ranges have experienced more than 25 km of east-west shortening, which has led to stacks of Miocene and younger sediments as much as 10 km thick, west of the San Andreas (Davis 1988; Namson & Davis 1990). The new crustal model is able to capture the signatures of these basins and sedimentary units at relatively long periods (6–30 s), evidenced by paths along the San Andreas fault (Figs 10 and 11).

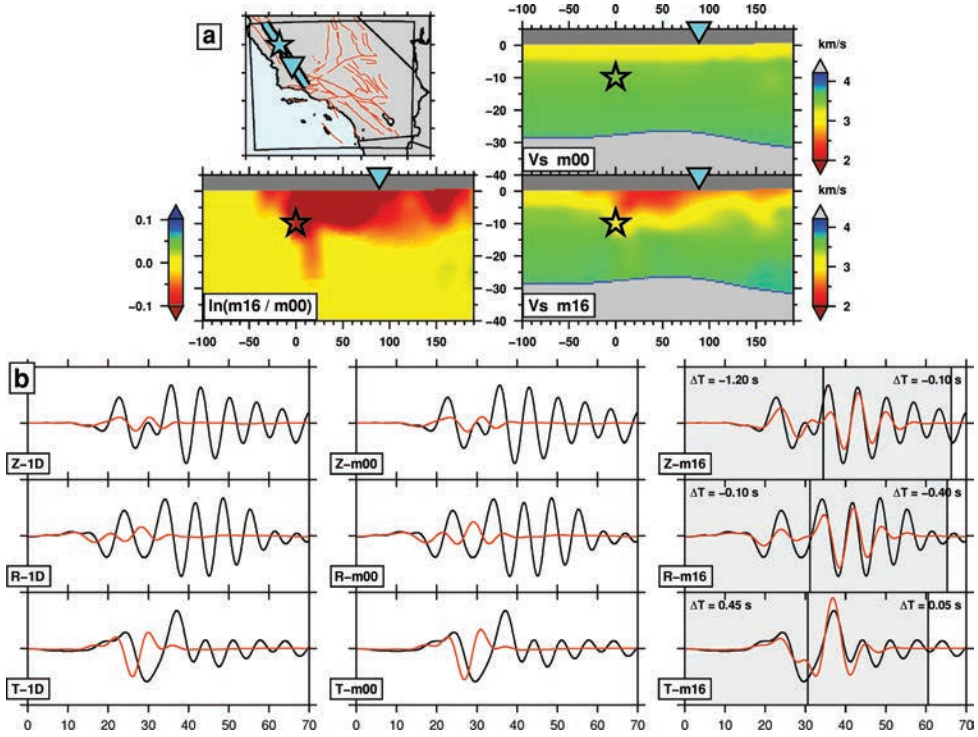


Figure 10. Vertical cross-sections and corresponding three-component seismograms (Section 5.2.1). (a) Cross-section of the V_S tomographic models for a path from event 14096196 on the San Andreas fault (⋆; M_w 4.6, depth 9.9 km) to station SMM.CI (∇; distance 89.0 km), west of the San Andreas fault. (b) Synthetic seismograms (red) and observed seismograms (black) for the period range 6–30 s. The grey windows of the $\mathbf{m}_{16}$ synthetics are the only quantifiably 'good' measurement windows for all three sets of seismograms (Maggi *et al.* 2009). See Fig. 9 caption for explanations of subplots.

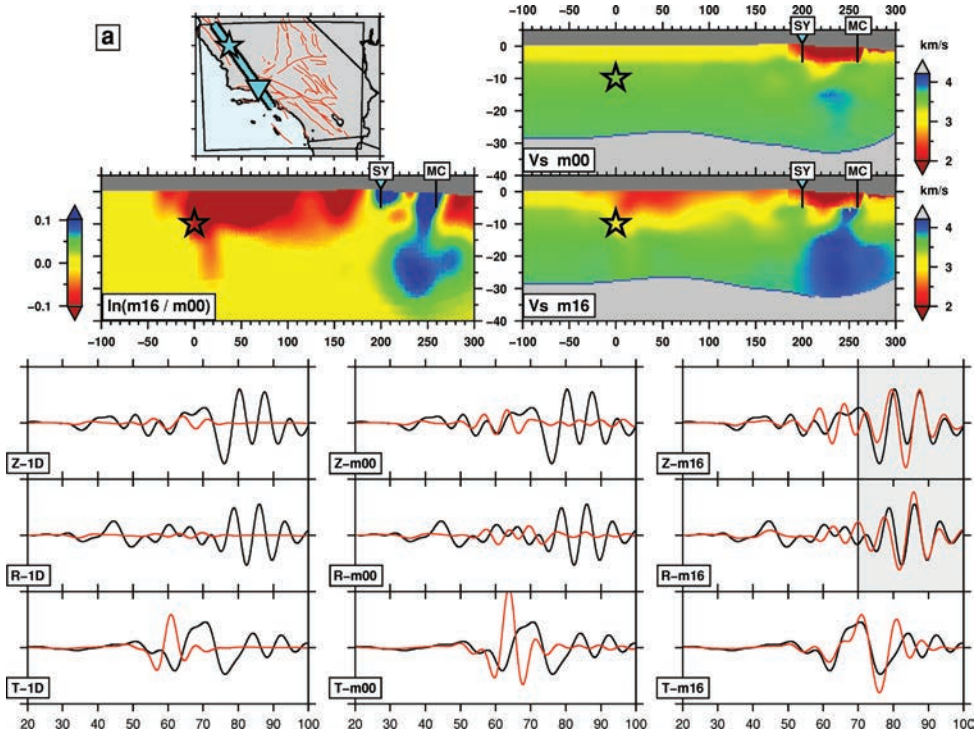


Figure 11. Vertical cross-sections and corresponding three-component seismograms (Section 5.2.1). (a) Cross-section of the V_S tomographic models for a path from event 14096196 (⋆; M_w 4.6, depth 9.9 km), on the San Andreas fault, to station WGR.CI (∇; distance 199.8 km), just north of the Ventura basin. SY, Santa Ynez fault; MC, Malibu Coast fault. (b) Synthetic seismograms (red) and observed seismograms (black) for the period range 6–30 s. The grey window highlights a Rayleigh wave discussed in Section 5.2.1. See Fig. 9 caption for explanations of subplots.

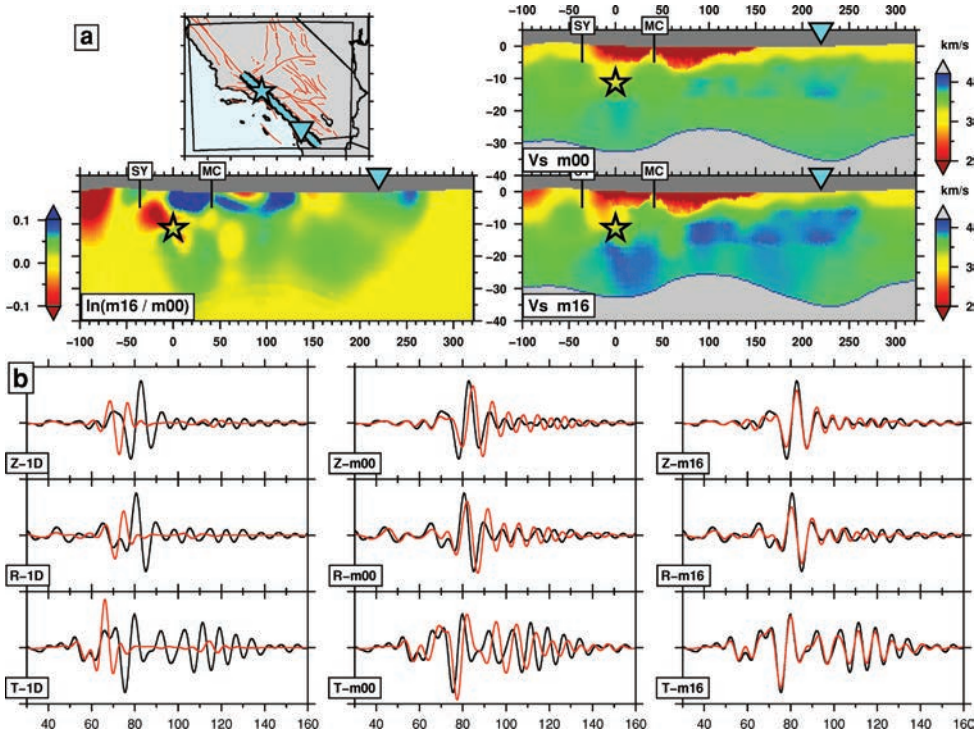


Figure 12. Vertical cross-sections and corresponding three-component seismograms (Section 5.2.1). (a) Cross-section of the V_S tomographic models for a path from event 9753485 (⋆; M_w 4.2, depth 11.4 km) to station DPP.CI (∇; distance 220.1 km), near San Diego. SY, Santa Ynez fault; MC, Malibu Coast fault. (b) Synthetic seismograms (red) and observed seismograms (black) for the period range 6–30 s. See Fig. 9 caption for explanations of subplots.

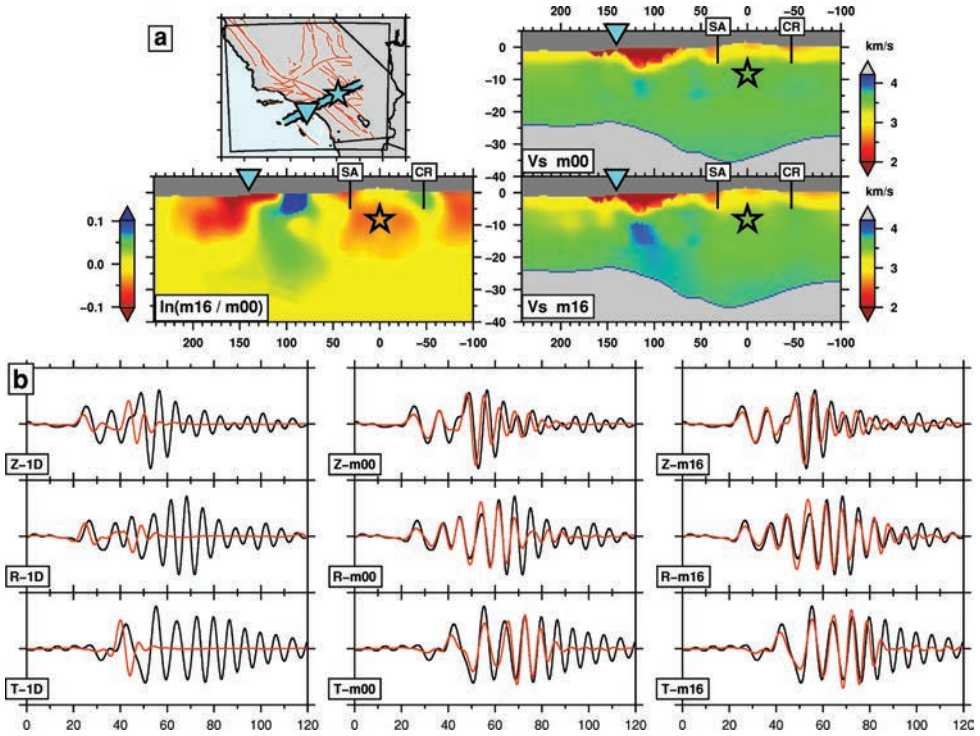


Figure 13. Vertical cross-sections and corresponding three-component seismograms (Section 5.2.1). (a) Cross-section of the V_S tomographic models for a path from event 9627721 (⋆; M_w 4.6, depth 8.3 km) to station FMP.CI (∇; distance 140.6 km), on the western margin of the Los Angeles basin. SA, San Andreas fault; CR, Camp Rock fault. (b) Synthetic seismograms (red) and observed seismograms (black) for the period range 6–30 s. See Fig. 9 caption for explanations of subplots.

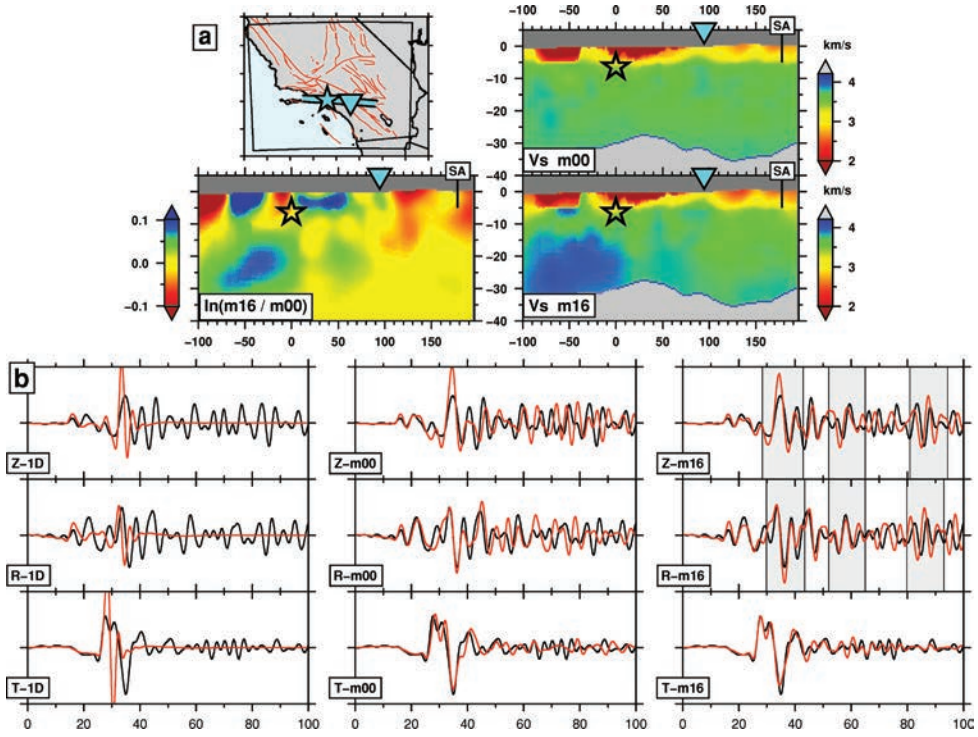


Figure 14. Vertical cross-sections and corresponding three-component seismograms (Section 5.2.1). (a) Cross-section of the V_S tomographic models for a path from event 9703873 near Hollywood ($\star$; M_w 4.2, depth 6.4 km) to station RVR.CI (∇; distance 94.6 km). SA, San Andreas fault. (b) Synthetic seismograms (red) and observed seismograms (black) for the period range 3–30 s. The grey time windows highlight three different Rayleigh wave arrivals discussed in Section 5.2.1. See Fig. 9 caption for explanations of subplots.

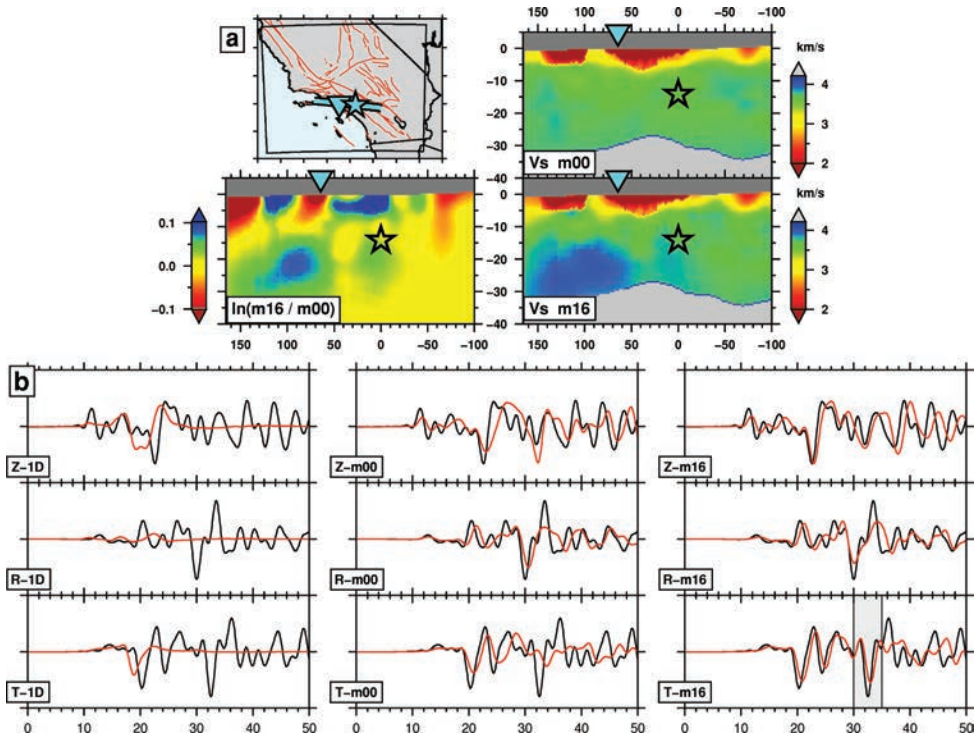


Figure 15. Vertical cross-sections and corresponding three-component seismograms (Section 5.2.1). (a) Cross-section of the V_S tomographic models for a path from event 14383980 ($\star$; M_w 5.4, depth 14.2 km), beneath the Chino Hills region of Los Angeles, to station SMS.CI (∇; distance 64.6 km), near the northwest margin of the Los Angeles basin. (b) Synthetic seismograms (red) and observed seismograms (black) for the period range 2–30 s. The grey window on the transverse component highlights a seismic waveform discussed in Section 5.2.1. See Fig. 9 caption for explanations of subplots. This earthquake was not used in the tomographic inversion.

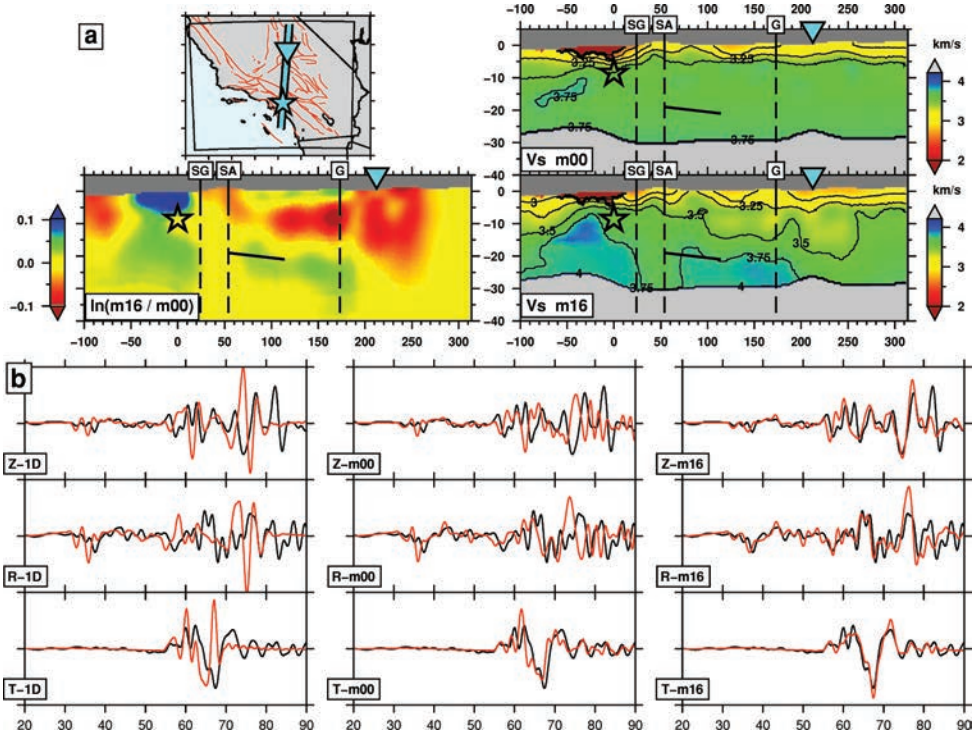


Figure 16. Vertical cross-sections and corresponding three-component seismograms (Section 5.2.1). (a) Cross-section of the V_S tomographic models for a path from event 9818433 (⋆; M_w 4.4, depth 8.6 km), east of the Los Angeles basin, to station CLC.CI (∇; distance 211.7 km), north of the Garlock fault. Vertical lines denote surface positions of faults: San Gabriel (SG), San Andreas (SA), and Garlock (G). Also shown for reference is a northward dipping layer identified in the wide-angle refraction survey of Fuis *et al.* (2001). Counter interval is 0.25 km s^{-1}. (b) Synthetic seismograms (red) and observed seismograms (black) for the period range 2–30 s. See Fig. 9 caption for explanations of subplots.

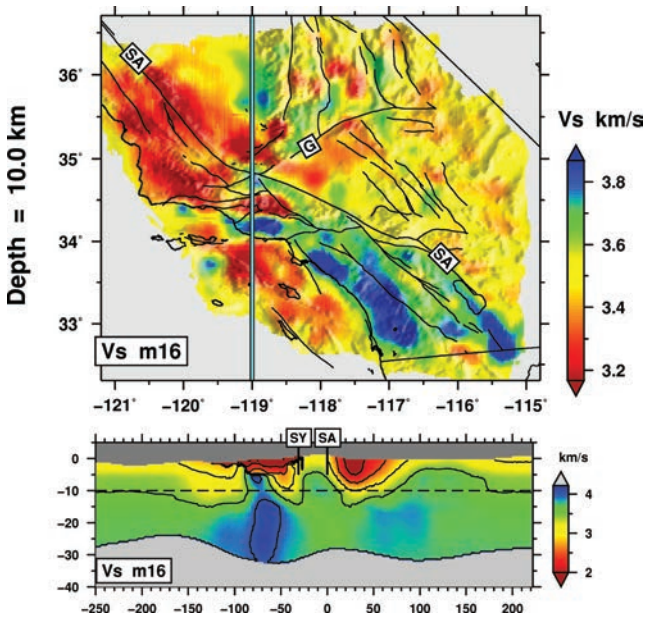


Figure 17. Top: Final V_S crustal model $\mathbf{m}_{16}$ at 10 km depth (middle panel of Fig. 7b), illuminated with shaded relief. Refer to Fig. 1 for locations of principal features. Bottom: Vertical cross-section along 119°W, looking west. Dashed line denotes 10 km depth; contour interval is 0.5 km s^{-1}. The variations in wave speed along longitude 119°W at 10 km depth are associated with, from north to south, the western Sierra Nevada (fast), the southern San Joaquin basin (slow), the San Emigdio Mountains (fast), the Ventura basin (slow), the Santa Monica Mountains (fast), and the Santa Monica basin (slow) (Tape *et al.* 2009). SA, San Andreas fault; G, Garlock fault; SY, Santa Ynez fault.

5.3.4 Western Transverse Ranges

At 20 km depth (Fig. 7c) the most conspicuous feature is the fast wave speed region (4.1 km s^{-1}) beneath the Ventura–Santa Barbara basin and the Santa Monica Mountains, also observed in Figs 8, 11 and 14. This region coincides with the surface expression of the western Transverse Ranges block (WTRB Luyendyk *et al.* 1980, Fig. 1). Tape *et al.* (2009) interpreted this feature as subduction-captured Farallon oceanic crust, which is believed to underlie the western portion of the WTRB (Nicholson *et al.* 1992, 1994). Below a depth of about 25 km, the anomaly could be attributed to a shallower crust than the ∼32 km (interpolated) value obtained by Zhu & Kanamori (2000) and used in our simulations. A north–south seismic refraction survey across the Santa Barbara Channel found that the crustal thickness beneath the WTRB decreases from 31 km near the coast to 23 km near Santa Cruz Island (Keller & Prothero 1987). However, it is also possible that there are fragments of oceanic crust beneath the WTRB (ten Brink *et al.* 2000) giving rise to higher wave speeds. Based on cross-sections such as the one in Fig. 8, we observe a strong vertical gradient in wave speed from the overlying Santa Barbara basin to the proposed oceanic crust; this is consistent with the model by Keller & Prothero (1987).

5.3.5 Mojave Desert region

Within a single vertical cross-section for a path traversing the Mojave region (Tape *et al.* 2009, Fig. 1), we identified reductions in wave speed due to both compositional and thermal features. Slow wave speeds near the White Wolf fault are associated with the southern San Joaquin sedimentary basin (Goodman & Malin 1992). Slow wave speeds east of the Camp Rock fault are associated with

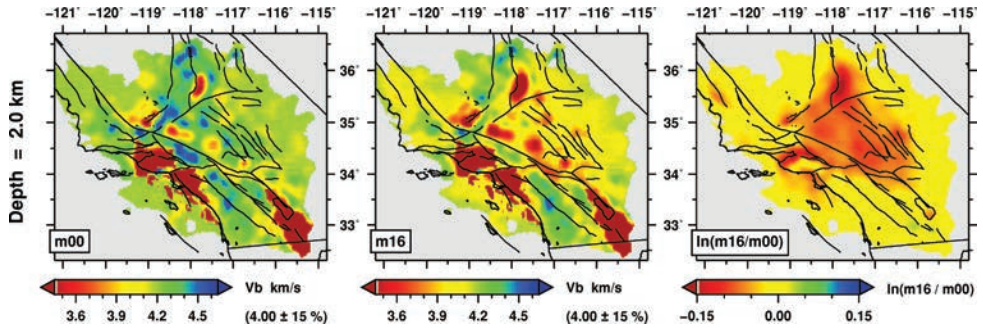


Figure 18. Horizontal cross-section at a depth of 2 km of bulk-sound speed V_B tomographic models, with active faults shown for reference. Refer to Fig. 1 for locations of principal faults and features. Left column shows the initial model, $\mathbf{m}_{00}$, centre column shows the final model, $\mathbf{m}_{16}$, and right column shows the difference between the two models, $\ln(\mathbf{m}_{16}/\mathbf{m}_{00})$.

Quaternary volcanism (e.g. Luffi *et al.* 2009, Fig. 1) manifested by higher heat flow (Bonner *et al.* 2003).

The north-south profile in Fig. 16 approximately coincides with the LARSE-I survey (Fuis *et al.* 2001, 2003). The tomographic inversion introduces a −10 per cent change in wave speed beneath the Coso geothermal region. The tomographic perturbation suggests the presence of a northward dipping interface at about 20 km depth, which approximately coincides with an interface identified in the LARSE-I survey.

We are unable to fit the arrival times of both Love and Rayleigh waves for dozens of paths across the Mojave region (Fig. 1a). This observation, known as the Love–Rayleigh discrepancy (Anderson 1961; Dziewonski & Anderson 1981), suggests the presence of radial anisotropy within the crust. Such an anisotropic unit would lie in the zone of dominant sensitivity of the surface waves in the data set, which is the upper 20 km.

Our preliminary simplified crustal column of the Mojave region west of the Camp Rock fault system (e.g. Fig. 16) is as follows.

(i) *0–6 km*: shallow (1–3 km deep) sedimentary basins (Murphy *et al.* 2008) and exhumed batholithic rocks of Mesozoic age (Grove *et al.* 2003).

(ii) *6–20 km*: approximately flat-lying schist, underplated during Farallon subduction ca. 80 Ma (Jacobson *et al.* 1996; Grove *et al.* 2003; Saleeby 2003).

(iii) *20–30 km depth*: underplated oceanic crust from relict Farallon subduction.

The COCORP seismic reflection surveys identified numerous dipping crustal reflectors at depths of 15 ± 6 km (Cheadle *et al.* 1986). Louie & Clayton (1987) interpreted a major reflector at 22 km depth within a simplified crustal column. The LARSE results reported wide-angle reflections at depths ranging from 20 to 24 km (Ryberg & Fuis 1998; Fuis *et al.* 2001, 2003, 2007). A detailed investigation is needed to reconcile reflection, refraction, tomographic and geological observations. For comparison, the seismic- and gravity-based model of Romanyuk *et al.* (2007) assumed a crustal column of primarily gneissic units.

5.4 Bulk sound speed model

Three horizontal cross-sections of the bulk sound speed (V_B) models are presented in Tape *et al.* (2009, fig. S9). In general, sensitivity to V_B is best at shallow depths (<10 km), and it is not as good as sensitivity to V_S.

At 2 km depth, the V_B perturbation, $\ln(\mathbf{m}_{16}/\mathbf{m}_{00})$, contains only longer scalelength variations (Fig. 18), and thus there are details in the initial 3-D model that are also present in the final model. Outside the Los Angeles basin, the slowest feature in the initial model is near Indian Wells Valley, just south of the Coso geothermal region (e.g. Hauksson & Unruh 2007). The model update involves a −15 per cent change in V_B to this (already slow) anomaly. Interestingly, in the central Mojave region at 2 km depth, the change to V_B is about −10 per cent (Fig. 18), whereas the change to V_S is about +5 per cent (Fig. 7a). These changes of opposite sign will lead to more pronounced changes in quantities such as Poisson's ratio and the V_P/V_S ratio (e.g. Christensen 1996). In order to enhance sensitivity to V_B in the inversion, one could preferentially weight the *P* waveforms.

5.5 Comparison with Chen *et al.* (2007)

Chen *et al.* (2007) obtained model LAF3D by performing a tomographic inversion for the greater Los Angeles region outlined in Fig. 1. The depth of their model volume was 26 km, and their data set included *P* and *S* traveltime measurements from 67 earthquakes recorded on 2000 seismograms. Synthetic seismograms were calculated using a finite-difference method. Their tomographic inversion was based on a 'scattering-integral' method, which (like adjoint methods) exploits the accuracy of the forward solver. They performed a single iteration of the model, but unlike a typical gradient-method iteration, they used the Hessian of the misfit function evaluated at the initial model. If the initial model is close to the solution, then the Hessian will provide a better model update than the gradient.

A direct comparison between our results and those of Chen *et al.* (2007) would involve generating synthetic seismograms for their LAF3D model and our model $\mathbf{m}_{16}$ using the same set of earthquakes, whether the ones used in their inversion, our inversion, or a separate set. Lacking a direct comparison, we nevertheless make some qualitative statements about the two models, at least in the volume that is common to both models. The two V_S models reveal a south-to-north increase in shear wave speed across the San Gabriel fault at shallow depths (upper 5 km) but a south-to-north decrease in wave speed at greater depths (below 10 km). The model of Chen *et al.* (2007) has a horizontal resolution of 4 km and contains some details that are not present in model $\mathbf{m}_{16}$. These additional details could be due to the enhanced horizontal resolution of Chen *et al.* (2007) or to less conservative choices of regularization.

5.6 Implications for seismic hazard assessment

The results presented in this paper demonstrate that moderate (M_w = 3.5–5.5) well-recorded earthquakes (Figs 2 and 3) can be used to make large, necessary changes to the crustal model

of southern California. As we have shown, waveforms from these earthquakes can be extremely complicated, even at relatively long periods (6–30 s). One example is the path from the Salton trough to the Los Angeles basin (Fig. 5), the same region covered by the TeraShake simulations (Olsen *et al.* 2006).

If it is not possible to fit waveforms from moderate, point-source earthquakes, then it will not be possible to fit waveforms from large earthquakes with complex ruptures. By beginning to fit complex propagation paths for moderate earthquakes, we provide hope for accurately simulating larger earthquakes. For example, the transverse component of ground motion from a M_w 4.5 earthquake on the White Wolf fault (Tape *et al.* 2009, Fig. 1) does not match the data for simulations from the initial 3-D model, but begins to match the data for simulations from the final model. The new model therefore provides a better starting point for simulating larger earthquakes on the White Wolf fault, such as the 1952 M_w 7.2 Kern County earthquake (Jenkins & Oakeshott 1955; Ben-Menahem 1978). Similarly, the improvement in fits for the M_w 4.6 Parkfield earthquake (Figs 10 and 11) suggests that the new crustal model is more appropriate for simulations of larger earthquakes on this segment of the San Andreas fault, such as the 1857 M_w 7.9 Fort Tejon earthquake (Sieh 1978; Agnew & Sieh 1978). See Petersen & Wesnousky (1994) for an overview of estimated maximum magnitudes for faults in southern California.

In southern California, there are no high-quality seismic waveforms available for major ($M_w > 7$) earthquakes. However, strong ($M_w = 6$–7) earthquakes have been recorded well enough to determine rupture models, such as the 2004 September 28 M_w 6.0 Parkfield event (Custódio *et al.* 2005) and the 2003 December 22 M_w 6.5 San Simeon earthquake in the Coast Ranges (McLaren *et al.* 2008). These earthquakes present the formidable challenge of fitting near-source and regional waveforms that capture the complexities of both the rupture process and the heterogeneous structure of the crust. A more accurate 3-D crustal model will benefit the development of rupture models for strong earthquakes.

In northern California, similar efforts are underway to incorporate earthquake simulations in seismic hazard assessment. A 3-D model of the Bay Area region (Brocher *et al.* 2006; Jachens *et al.* 2006) has been used to model moderate earthquakes, strong earthquakes and scenario earthquakes (Rodgers *et al.* 2008; Aagaard *et al.* 2008a,b).

6 MISFIT ANALYSIS

Tape *et al.* (2009) assessed the quality of the new crustal model, $\mathbf{m}_{16}$, as well as the reduction in misfit from the initial 3-D model, $\mathbf{m}_{00}$, to $\mathbf{m}_{16}$. The analysis included an evaluation of both the traveltime measure of misfit used in the tomographic inversion as well as a simple waveform difference (not used in the tomographic inversion). A traveltime measure of misfit requires a degree of similarity between two waveforms that is not always present, and thus we cannot directly compare $F^{\mathrm{T}}(\mathbf{m}_{00})$ and $F^{\mathrm{T}}(\mathbf{m}_{16})$ for each $\mathbf{m}_{16}$ time window (e.g. Fig. 11). Of course, such a comparison is always possible using the waveform difference $F^{\mathrm{W}}(\mathbf{m})$.

6.1 Traveltime and amplitude misfit

The multitaper traveltime differences in the final model, $F^{\mathrm{T}}(\mathbf{m}_{16})$ defined in equation (3), have a standard deviation of less than 1 s for the entire data set. For the three period ranges, the patterns do not change appreciably, and all listed time shift standard deviations are <0.8 s (Tape *et al.* 2009, fig. S5).

In Fig. 19, we tabulate ΔT and $\Delta \ln A$ (eq. A2) measurement windows for 6–30 s seismograms for models $\mathbf{m}_{00}$ and $\mathbf{m}_{16}$; in other words, the comparison involves two different sets of windows. The reduction in misfit for ΔT is considerably better than for $\Delta \ln A$. One reason for this is that ΔT is used explicitly within the misfit function (eq. 7), whereas $\Delta \ln A$ is not. A second reason is that some amplitude misfit reduction has already occurred within the two source inversions, performed with $\mathbf{m}_{00}$ and $\mathbf{m}_{12}$ (Section 2.2). Thus the reduction in $\Delta \ln A$ misfit shown in Figs 19(d)–(f) is strictly due to wave speed changes. Explicit use of amplitude differences within the misfit function would help further reduce the amplitude differences; however, this presents the challenge of ascribing amplitude anomalies to differences in source parameters (e.g. magnitude, radiation pattern), wave speeds, and attenuation.

The largest (positive) traveltime differences in the initial data set are emphasized in Fig. 20. In particular, the $\mathbf{m}_{00}$ seismograms contain 201 windows with $\Delta T \geq 6$ s that are no longer present in the $\mathbf{m}_{16}$ seismograms. In hindsight, we know that these $\mathbf{m}_{00}$ outliers indicate that strong wave speed reductions are needed in distinct regions, including the southern San Joaquin basin, the mid-crust of the Mojave region, the upper crust of the Coast Ranges, and the upper crust of the Coso region.

6.2 Waveform misfit

We use a direct waveform difference, $F^{\mathrm{W}}(\mathbf{m})$ defined in (eq. 5), as the primary measure of misfit. In constructing the histograms in Tape *et al.* (2009, Fig. 3), we excluded all windows whose time shifts in *both* $\mathbf{m}_{00}$ and $\mathbf{m}_{16}$ are ≤ 1 s, the target measurement value in the tomographic problem (eq. 4). This procedure excludes only those windows that have not changed appreciably (for better or for worse). The waveform misfit measure is applied to either the portions of records within the FLEXWIN ($\mathbf{m}_{16}$) time windows (Tape *et al.* 2009, Figs 3a and c) or to the entire seismogram containing at least one (non-excluded) time window (Tape *et al.* 2009, Figs 3b and d), including time intervals before the expected *P*-wave arrival and after the surface wave arrivals. For a given set of windows, the number of seismograms containing windows will be less than (or equal to) the total number of windows. The number of seismograms listed includes (up to three) different bandpassed versions.

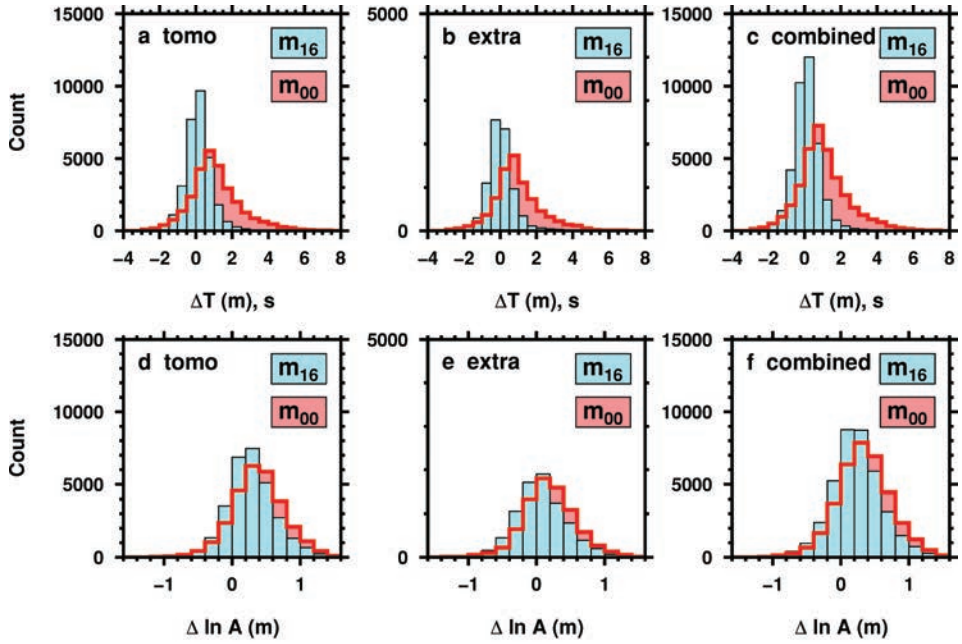
The waveform misfit analysis is shown in Tape *et al.* (2009, Fig. 3 and fig. S3). There are several comparisons to make among the subplots.

(i) There is a strong similarity between the earthquakes used in the inversion ('tomo') and the earthquakes not used in the inversion ('extra'), both in terms of the overall misfit and the misfit reduction.

(ii) The waveform misfit of full seismograms is reduced.

(iii) The waveform misfit of measurement windows is better than that computed for full seismograms, as expected.

(iv) For the measurement windows, neither the overall misfit nor the misfit reduction show a dependence on period range. This is not a one-to-one comparison, since the comparison is for different sets of windows, but it suggests that for many windows, such as those common to all three period ranges, the synthetic waveforms are capturing the dominant features of the wavefield. For the full seismograms, however, both the misfit and misfit reduction get progressively better from 2–30 to 3–30 to 6–30 s. This is because measurement windows selected from the 6–30 s records cover a greater portion of the full seismogram than those selected from the 3–30 and 2–30 s records, and thus we expect a greater misfit reduction.



			Windows (6–30 s)		Mean ± Std.	
			$\mathbf{m}_{00}$	$\mathbf{m}_{16}$	$\mathbf{m}_{00}$	$\mathbf{m}_{16}$
a	tomo	ΔT	28650	30252	1.10 ± 1.46 s	0.10 ± 0.76 s
b	extra	ΔT	7903	8083	0.92 ± 1.31 s	-0.02 ± 0.77 s
c	combined	ΔT	36553	38335	1.06 ± 1.43 s	0.08 ± 0.76 s
d	tomo	$\Delta \ln A$	28650	30252	0.38 ± 0.39	0.28 ± 0.36
e	extra	$\Delta \ln A$	7903	8083	0.19 ± 0.38	0.07 ± 0.38
f	combined	$\Delta \ln A$	36553	38335	0.34 ± 0.39	0.23 ± 0.37

Figure 19. Misfit analysis for traveltime and amplitude differences within measurement windows of the 6–30 s seismograms (Section 6.1). (a)–(c) Traveltime differences, ΔT, for models $\mathbf{m}_{00}$ (red) and $\mathbf{m}_{16}$ (blue). Fig. 19c is expanded in Fig. 20. (d)–(f) Same as (a)–(c), but for amplitude differences, $\Delta \ln A$. The reduction in misfit for $\Delta \ln A$ is less than that for ΔT, because only ΔT is explicitly used in the misfit function of the tomographic inversion. The inset table lists the mean and standard deviation for each distribution.

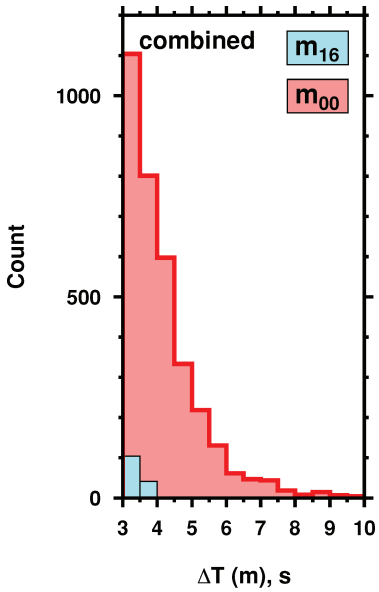


Figure 20. Zoom-in of Fig. 19(c), showing the reduction in ΔT from $\mathbf{m}_{16}$ to $\mathbf{m}_{00}$ for measurement windows within 6–30 s seismograms. There are 201 measurement windows for $\mathbf{m}_{00}$ with $\Delta T \geq 6$ s.

6.3 Waveform variance reduction

The typical formula for variance reduction, VR, is a first-order approximation of a logarithmic version, VRL (Appendix C). For the ith time window, we may first compute $F_i^{\mathrm{W}}(\mathbf{m}_{00})$ and $F_i^{\mathrm{W}}(\mathbf{m}_{16})$, followed by VR_i and VRL_i. For example, substituting eq. (5) into eq. (C2), the logarithmic variance reduction for a single window pick is

$$\mathrm{VRL}_i = \ln\left[\frac{F_i^{\mathrm{W}}(\mathbf{m}_1)}{F_i^{\mathrm{W}}(\mathbf{m}_2)}\right] = \ln\left\{\frac{\int_{-\infty}^{\infty} w_i(t)\,[d(t) - s(t,\mathbf{m}_1)]^2\,\mathrm{d}t}{\int_{-\infty}^{\infty} w_i(t)\,[d(t) - s(t,\mathbf{m}_2)]^2\,\mathrm{d}t}\right\}. \tag{8}$$

The logarithmic version of variance reduction provides a sensible formula, because VRL is approximately Gaussian distributed, whereas VR is not, as shown in Figs 21 and 22. We do not see a good reason to use the first-order approximation (VR), and thus the discussion emphasizes VRL.

Because VRL is approximately Gaussian, its centre corresponds to the mean value. We use the mean value to represent the overall variance reduction, either for the windows (Figs 22a and c) or for the full seismograms (Figs 22b and d). As expected from the distributions in Tape *et al.* (2009, Fig. 3), the variance reduction is larger

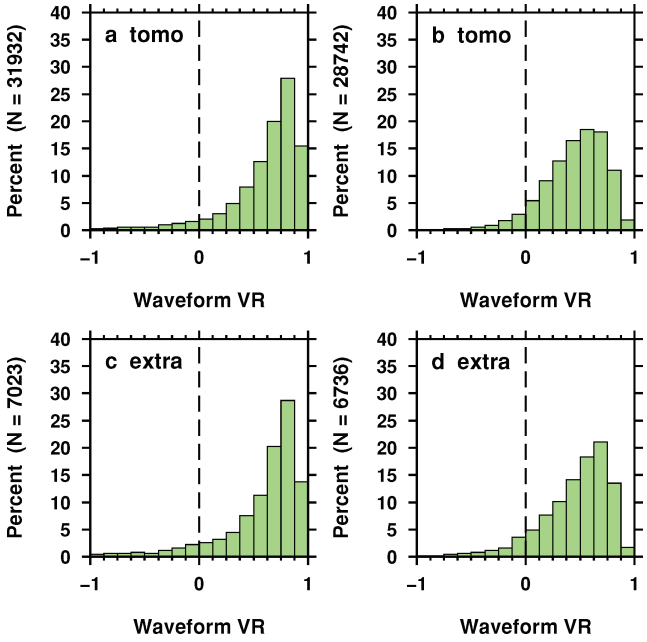


Figure 21. Histograms of variance reduction, VR, corresponding to the waveform misfits in Tape *et al.* (2009, Fig. 3). The range of permissible values of VR is from $-\infty$ to 1. A value of VR $= 0$ corresponds to no reduction in misfit. (a) Variance reduction from $\mathbf{m}_{00}$ to $\mathbf{m}_{16}$ for windows used in the inversion for seismograms in all three period ranges (6–30 s, 3–30 s, and 2–30 s). (b) Variance reduction for full seismograms containing at least one measurement window. (c)–(d) Same as (a)–(b), but for the set of extra earthquakes.

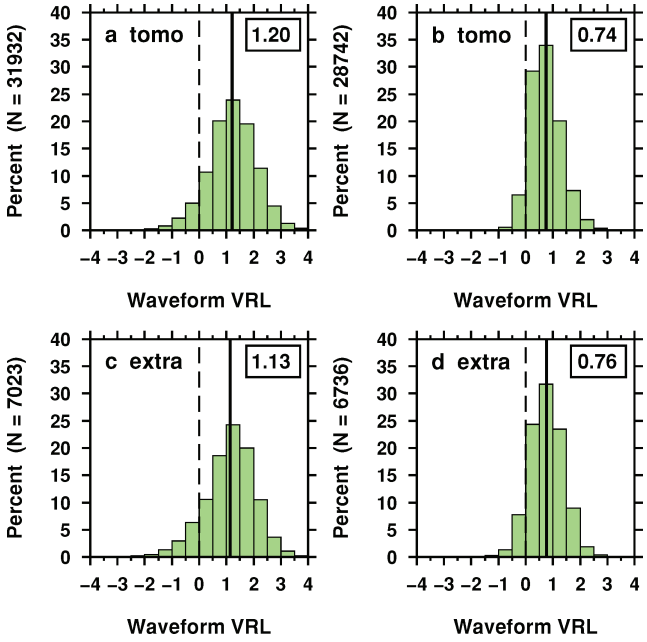


Figure 22. Same as Fig. 21, but showing logarithmic variance reduction, VRL, instead of variance reduction, VR. The quantities are approximately Gaussian distributions with the mean value denoted by the vertical black line and indicated by the boxed number. The range of permissible values of VRL is from $-\infty$ to ∞. A value of VRL $= 0$ corresponds to no reduction in misfit.

within the windows than within the full seismograms, and it is about the same for the earthquakes used in the tomographic inversion as it is for the extra earthquakes. For comparison, using eq. (C3), the VRL values of 1.20, 0.74, 1.13 and 0.76 (Fig. 22) correspond to VR values of 0.70, 0.52, 0.68 and 0.53.

The negative values of VRL (Fig. 22) correspond to seismic waveforms that show 'variance amplification,' that is waveforms in which $F^{\mathrm{W}}(\mathbf{m}_{16}) > F^{\mathrm{W}}(\mathbf{m}_{00})$. In Fig. 23, we have included two of the largest negative VRL full seismograms in the entire data set. In Figs 23(a)–(c), the transverse component is 'unfit' by $\Delta T = 1.1$ s, which barely exceeds the target measurement of 1.0 s (eq. 4), and the amplitude is too large as well. However, the phase of the Rayleigh wave is better fit by the $\mathbf{m}_{16}$ synthetics. In Figs 23(d)–(f), the transverse component is noticeably worse for $\mathbf{m}_{16}$ for the coda of the Love wave, which leads to the large negative VRL. However, the Rayleigh wave is much better fit in both phase and amplitude, indicating that $\mathbf{m}_{16}$ is probably the better model for this path. In short, the outliers with large negative waveform variance reductions are not a significant cause for concern.

6.4 Residuals

As with any inverse problem, the remaining residuals provide direction for additional improvements. The largest waveform residuals (Fig. 24) are associated with seismic wave propagation in the Coast Ranges–Great Valley region. In order to fit such waveforms, we must construct and incorporate sedimentary basin models for the Coast Ranges, Mojave Desert and Great Valley by using all available data. By 'available data,' we mean seismic reflection and refraction data and industry well logs (e.g. Süss & Shaw 2003). Gravity and magnetic data could provide additional constraints (e.g. Langenheim 1999; Romanyuk *et al.* 2007).

We identified subtle traveltime residuals that are likely due to the assumption of isotropy in our model (Section 5.3.5). In a future study we will attempt to improve fits to all waveforms—in particular, Love, Rayleigh and *S*—by using anisotropic sensitivity kernels (Sieminski *et al.* 2007).

7 MODEL UNCERTAINTY AND RESOLUTION

The emphasis in assessing the new crustal model is on misfit (Section 6). We do not estimate uncertainties of model parameters (V_{S} and V_{B}), nor do we perform a formal resolution analysis. A single resolution test would have a computational cost that is comparable to the tomographic inversion itself (Tape *et al.* 2007; Fichtner *et al.* 2009). Instead, we compute the composite volumetric sensitivity of all measurements (e.g. Chen *et al.* 2007), and we examine the scalelengths in the net model update, $\ln(\mathbf{m}_{16}/\mathbf{m}_{00})$.

7.1 Model uncertainty

As a first step in addressing uncertainties in the model parameters, we compute the composite volumetric sensitivity of all measurements (e.g. Chen *et al.* 2007). This is achieved by using the same procedure that was used to compute each event kernel (Tape *et al.* 2007), but omitting the traveltime measurement weight for each adjoint source (i.e. using a unit traveltime anomaly for all measurements). Because we have two inversion variables, V_{S} and V_{B}, we have two corresponding volumetric sensitivity kernels, which we refer to as 'coverage kernels' for brevity.

The coverage kernels for V_{S} and V_{B} are shown in Tape *et al.* (2009, figs S6 and S7). A horizontal cross-section of the V_{S} coverage

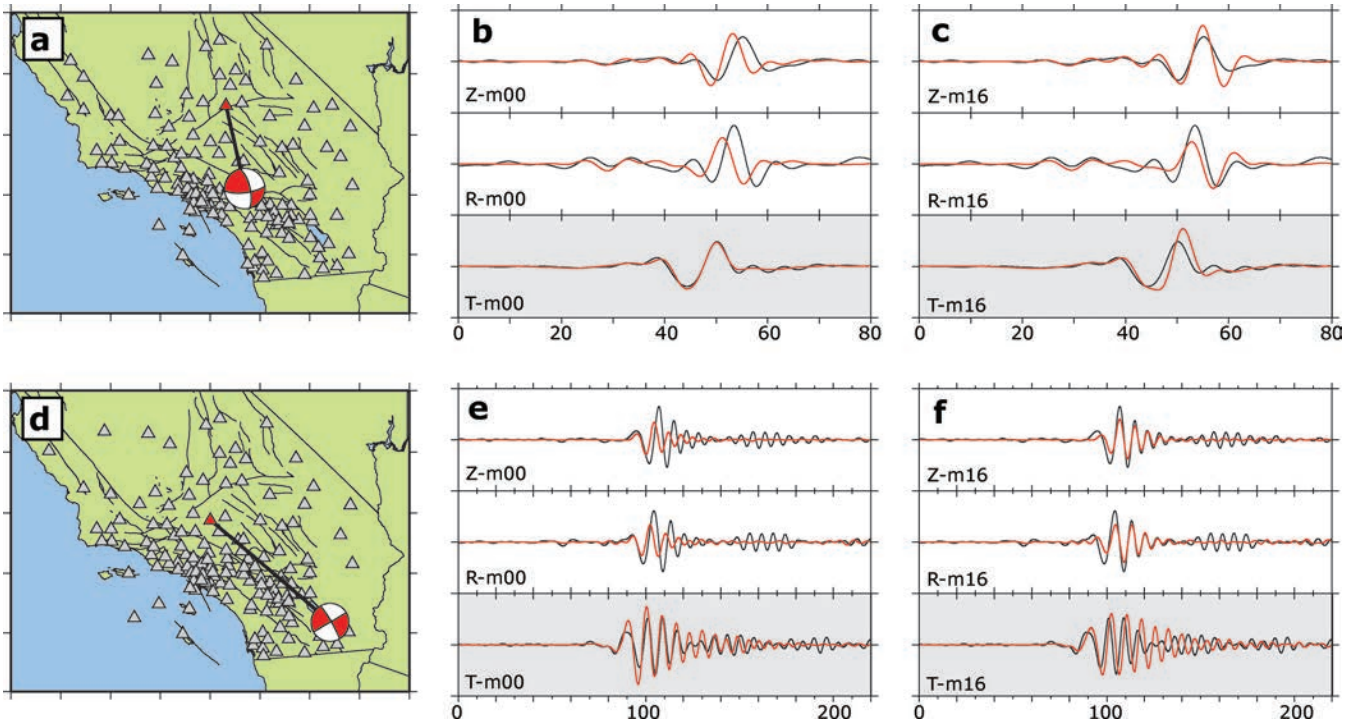


Figure 23. Two examples of the largest negative waveform variance reductions for all seismograms in the data set (Section 6.3). (a) Station coverage for 6–30 s seismograms for event 10370141 (M_w 4.4, depth 14.2 km), near the San Andreas fault, to station LRL.CI (distance 156.1 km), just north of the Garlock fault. (This earthquake was not used in the tomographic inversion.) (b) Seismogram fits for $\mathbf{m}_{00}$, bandpass filtered from 6–30 s. (c) Seismogram fits for $\mathbf{m}_{16}$, bandpass filtered from 6–30 s. For the highlighted transverse component, VRL $= -1.77$, corresponding to the largest increase in waveform misfit for the entire data set. (d) Station coverage for 6–30 s seismograms for event 14178236 (M_w 4.1, depth 4.0 km), near the Salton trough, to station EDW2.CI (distance 291.6 km) in the western Mojave. (e) Seismogram fits for $\mathbf{m}_{00}$, bandpass filtered from 6–30 s. (f) Seismogram fits for $\mathbf{m}_{16}$, bandpass filtered from 6–30 s. For the highlighted transverse component VRL $= -1.17$, corresponding to one of the largest increases in waveform misfit for the entire data set.

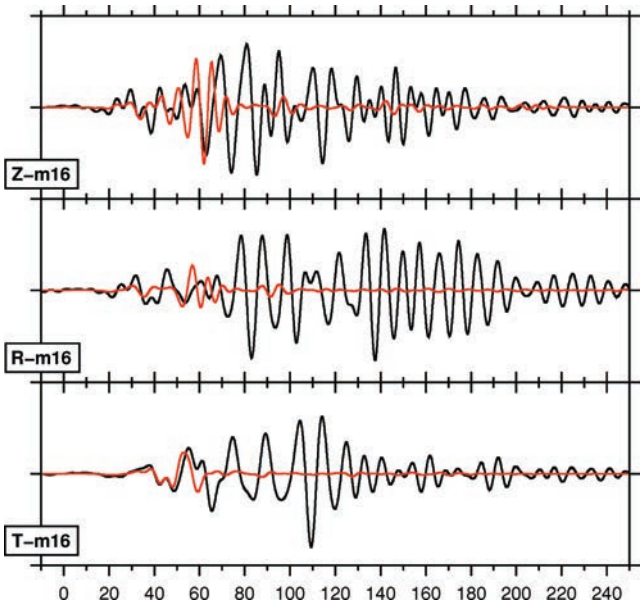


Figure 24. Extreme resonance within the San Joaquin basin captured by observed seismograms (black), but not by synthetic seismograms (red), for the period range 6–30 s. The earthquake is event 14096196 (⋆; M_w 4.6, depth 9.9 km), near the San Andreas fault at Parkfield, and the station is Bakersfield (BAK.CI; ∇; distance 199.8 km), within the San Joaquin basin. Fitting such seismic waveforms requires detailed 3-D models of sedimentary basins in the Great Valley–Coast Ranges region.

kernel is shown in Fig. 25. The left panel shows the field without a mask, and the right panel includes a mask. The threshold for the mask is given by a subjective value of $K_0 = 10^{-16}$ m^{-3}. In regions where the coverage kernel is less than this value, the tomographic model is masked out, as shown in Fig. 25. The coverage kernel decreases with depth (Tape *et al.* 2009, fig. S6), and has maximum sensitivity near the surface due to the sampling by shorter-period surface waves in the data set. The masks are applied to the V_S cross-sections shown in Fig. 7.

Coverage kernels, though they contain much more information than standard ray density maps, are an *ad hoc* representation of the uncertainties of model parameters, which we will address more formally in a future study. From a Bayesian perspective, the final model ($\mathbf{m}_{16}$) represents the mean model of a posterior distribution of possible models ($\mathbf{C}_M$) (Tarantola 2005). It is important that we construct a distribution of possible models that provides a guide for future studies that inevitably will include more stations and higher-frequency seismograms. The inclusion of additional data will then help reduce the distribution of all possible models.

7.2 Model resolution

A seismic tomographic study typically includes a resolution analysis that shows how well a model perturbation (e.g. a delta function or a checkerboard pattern) is expected to be resolved by the inversion procedure. If forward modelling is computationally inexpensive (e.g. ray theory), then it is possible to perform resolution tests with

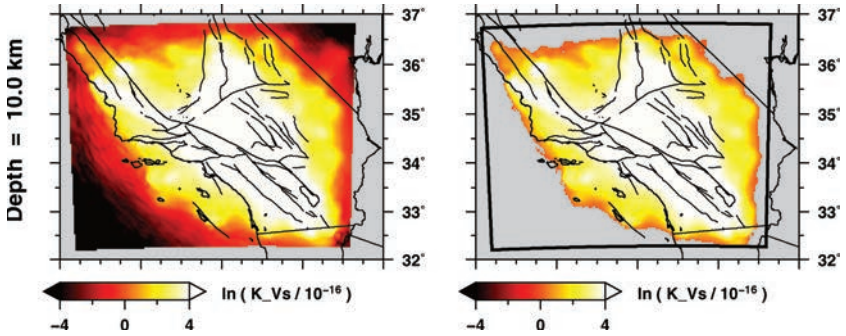


Figure 25. Volumetric coverage for the V_S tomographic model, plotted at 10 km depth (Section 7.1). The scalar field is computed as $\ln[K_{V_S}(\mathbf{x})/K_0]$, where $K_{V_S}(\mathbf{x})$ is the sum of all V_S kernels, and $K_0 = 10^{-16}$ m^{-3} is the threshold value that determines the mask shown in the right panel. The black outline denotes the simulation region, which extends to 60 km depth.

modest additional computations, but interpretations of such tests are complicated by the limited accuracy of the forward model embedded within the inverse problem. If the forward model is computationally expensive, such as a full-wavefield method, then resolution tests are as computationally expensive as the inversion itself (Tape *et al.* 2007); nevertheless, this has been achieved by Fichtner *et al.* (2009) and Chen *et al.* (2007). (Chen *et al.* (2007) performed a resolution test requiring no additional 3-D simulations, but it required access to all possible source and receiver Green's functions in the simulation volume.)

The 2-D synthetic resolution tests of Tape *et al.* (2007) used a realistic source–station geometry for the 3-D problem, and they suggest that the resolution of the new 3-D model should be as good as the scalelengths within the event kernels themselves. The 3-D model update, $\ln(\mathbf{m}_{16}/\mathbf{m}_{00})$, is constructed strictly from event kernels and therefore provides qualitative estimates of the resolvable scalelengths. In particular, the minimum scalelength is ∼3 km in the vertical direction (visible in Fig. 16a) and ∼20 km in the horizontal direction. These values correspond to the regions of best coverage; lower resolution is expected in regions of poor coverage, such as the Great Valley or near the boundaries of the simulation domain. We note that shorter scalelength features are present in the unsmoothed event kernels, but our choices of regularization (Section 4.1) prevent them from appearing in the model updates. In other words, we adopt a conservative approach that will introduce finer details into the model only if seismograms from many different earthquakes require them.

8 CONCLUSION

We have exploited the accuracy of the SEM within a tomographic inverse problem based on adjoint methods. After 16 iterations, we have obtained a model with local perturbations of ±30 per cent from the initial 3-D model. We have fit thousands of three-component seismograms that capture both the phase and amplitude of particle motion recorded at seismic stations. Direct waveform difference measurements of full-length seismograms that were never used in the inversion (Tape *et al.* 2009, Figs 3b and d) provide compelling support for the quality of the new crustal model.

The new tomographic model has been incorporated into the latest version of the Southern California Earthquake Center (SCEC) Community Velocity Model, CVM-H 6.0 (Plesch *et al.* 2009). It is available via download through the SCEC website (www.scec.org). A full set of vertical and horizontal cross-sections of the tomographic model is available from the interactive website at http://www.data.scec.org/research/carltape/socalm16.html.

We summarize several key points of our tomographic procedure and findings.

(i) An initial 3-D model is required to capture the general characteristics (e.g. shape, amplitude, duration) of three-component seismograms for the period range of interest (2–30 s). For seismogram comparisons in this paper, we have also included 1-D synthetics, which tend to fit body waves (e.g. Fig. 9) better than surface waves (e.g. Fig. 13).

(ii) For each iteration in the tomographic inversion, we obtain a new model of V_S and V_B, from which we compute a new set of synthetics seismograms. We subsequently run FLEXWIN (Maggi *et al.* 2009) to automatically pick a new set of measurement windows. Using the new measurements, we compute event kernels specific to the new model.

(iii) The net changes in the V_S model are large, locally in excess of ±30 per cent, but the changes during any one iteration are small, locally less than ±10 per cent (Figs 6 and 7). The net changes in the V_B model are predominantly negative and confined to the upper 15 km (Fig. 18).

(iv) The areas in the initial model that require changes are highly variable and generally unknown. For example, it takes more than 10 iterations to isolate the −35 per cent V_S anomaly related to the southern San Joaquin basin (Tape *et al.* 2009).

(v) Frequency-dependent surface waves are necessary to resolve crustal structure. Measurements over multiple period ranges allow us to apply changes with a vertical scalelength of 3 km (e.g. Fig. 16a).

(vi) Although only traveltime differences are used in the misfit function (eq. 6), amplitude differences also decrease in the final model, partly due to focusing and defocusing induced by the 3-D model (Figs 8–11).

(vii) Seismic arrivals absent in synthetics for the initial 3-D model appear in the final synthetics (Figs 10, 11 and 15).

(viii) The new model facilitates more accurate assessments of seismic hazard. We present examples of improvements of waveform fits from regions where major earthquakes ($M_w = 7$–8) have occurred in the past (e.g. Figs 5 and 10).

(xi) The new model illuminates several wave speed contrasts across faults (Section 5.3.1).

(x) The Coast Ranges, southern San Joaquin basin, and Continental Borderlands are characterized by predominantly slow wave speeds in the upper 15 km.

(xi) The relatively fast wave speeds beneath the western Transverse Ranges and the western Mojave may be the signatures of underplated oceanic crust related to Farallon subduction.

Finally, we summarize three important points regarding the assessment of the final model.

(i) The SEM is used to compute all synthetic seismograms, which therefore capture all the complexities of wave propagation that are afforded by the numerical technique. We assess the quality of any model, for example $\mathbf{m}_{00}$ or $\mathbf{m}_{16}$, by comparing its synthetic seismograms directly with observed seismograms.

(ii) We set aside 91 earthquakes to independently evaluate the misfit reduction from the initial 3-D model ($\mathbf{m}_{00}$) to the final 3-D model ($\mathbf{m}_{16}$) (Tape *et al.* 2009). The misfit reduction, measured using a direct waveform difference, is essentially the same for the 91 extra earthquakes as it is for the 143 earthquakes used in the inversion (Figs 22a and c). This demonstrates the high quality of both the new crustal model and the earthquake source parameters.

(iii) We observe a significant waveform misfit reduction in full-length seismograms, including portions of seismograms that were never used in the tomographic inversion (Figs 22b and d).

8.1 The upper 8 km

The seismic wavefield in southern California suggests a vertical transition in crustal complexity in the vicinity of 8 km depth. We use this depth to distinguish between 'shallow' (<8 km) and 'deep' (>8 km) earthquakes.

Surface waves from shallow earthquakes dominate the misfit for the initial 3-D model ($\mathbf{m}_{00}$) and dominate the residuals for the final 3-D model ($\mathbf{m}_{16}$). We find that the upper 8 km of the crust contain the strongest heterogeneity, primarily in the form of sedimentary basins. Although sedimentary basins do not cover all of southern California (Fig. 1), their signature is inescapable for most paths for the period range of interest (2–30 s). This is partly due to the concentration of people—and therefore, stations—in regions of actively forming basins with adjacent seismogenic faults.

Surface waves from deep earthquakes, which provide a weaker excitation of the shallowest crust, are fit reasonably well in the initial model. Similarly, body waves, which spend most of their time in the lower crust or upper mantle (for the paths in this study), are also fit reasonably well in the initial model. The relatively less heterogeneous lower crust might account for previous success in modelling body wave traveltimes from earthquakes at all depths (e.g. Hauksson 2000; Lin *et al.* 2007b), and in modelling full waveforms from deeper earthquakes (e.g. Hauksson *et al.* 2008).

8.2 Exotic seismic waveforms

The new crustal model captures several 'exotic' seismic waveforms (Fig. 14), for example, due to basin resonance and laterally reflected surface waves. We are therefore able to enhance tomographic images by delving deeper into the seismic records.

Consider the waveforms shown in Fig. 26 for a path from the San Jacinto fault to station EDW2 in the western Mojave. The three-component seismograms, filtered between 3 and 30 s, record structural complexity that lies off the direct path between source and station. Although the synthetic seismograms do not fit the full waveforms (in particular the body wave arrivals), there are isolated time windows of agreement on the vertical and radial components, suggesting the occurrence of secondary Rayleigh waves.

Using the approach of Liu & Tromp (2006), we can isolate the volumetric region of sensitivity for a particular waveform. This is shown in Fig. 27 for the three time windows highlighted in Fig. 26. The 'direct' Rayleigh wave ('1') in fact diffracts along the San Jacinto fault at the surface (Fig. 27d), thereby avoiding the tomographically slow anomaly at the convergence between the San Jacinto and San Andreas faults (Fig. 27b). At 4 km depth, the sensitivity is approximately situated on the direct path (Fig. 27e). The second time window ('2') is a Rayleigh wave that propagates through the north end of the Salton trough before turning west towards EDW2 (Fig. 27f). The third time window ('3') records (at least in part) a Rayleigh wave that reflects off a margin of the Salton trough, then follows a San Jacinto diffracted path (similar to '1') to EDW2 (Fig. 27g). From the perspective of seismic wave propagation, Figs 26 and 27 are a beautiful manifestation of crustal complexity. From the perspective of seismic tomography, we can use such measurements to sharpen the image of the southern California crust.

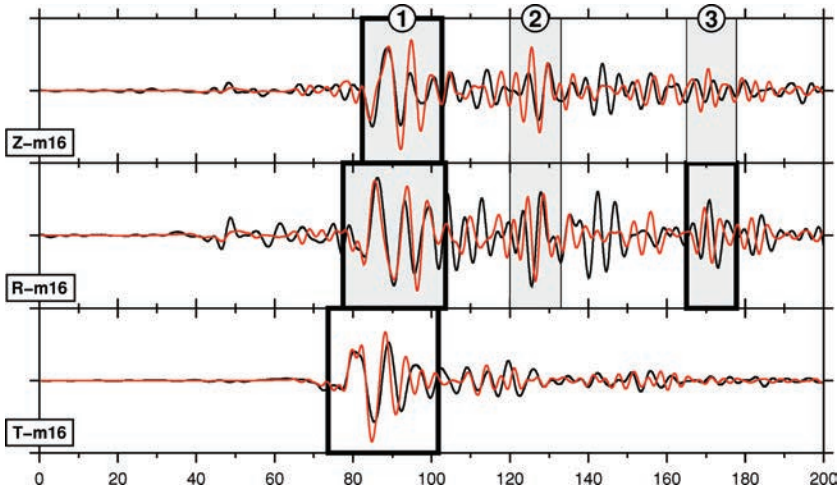


Figure 26. Strong lateral multipathing of Rayleigh waves identified in synthetic seismograms (red) and observed seismograms (black) for the period range 3–30 s for a path from event 14236768 (M_w 4.1, depth 3.8 km) on the San Jacinto fault to station EDW2.CI (distance 257.0 km) in the western Mojave. The shaded grey time windows denote three time intervals of the vertical (Z) and radial (R) components that arise from three different Rayleigh waves, as demonstrated in Fig. 27. For comparison, the black outlined time windows are the automated window picks from FLEXWIN (Maggi *et al.* 2009). The time interval between 135 s and 150 s contains an additional waveform that is not matched by the synthetic seismograms.

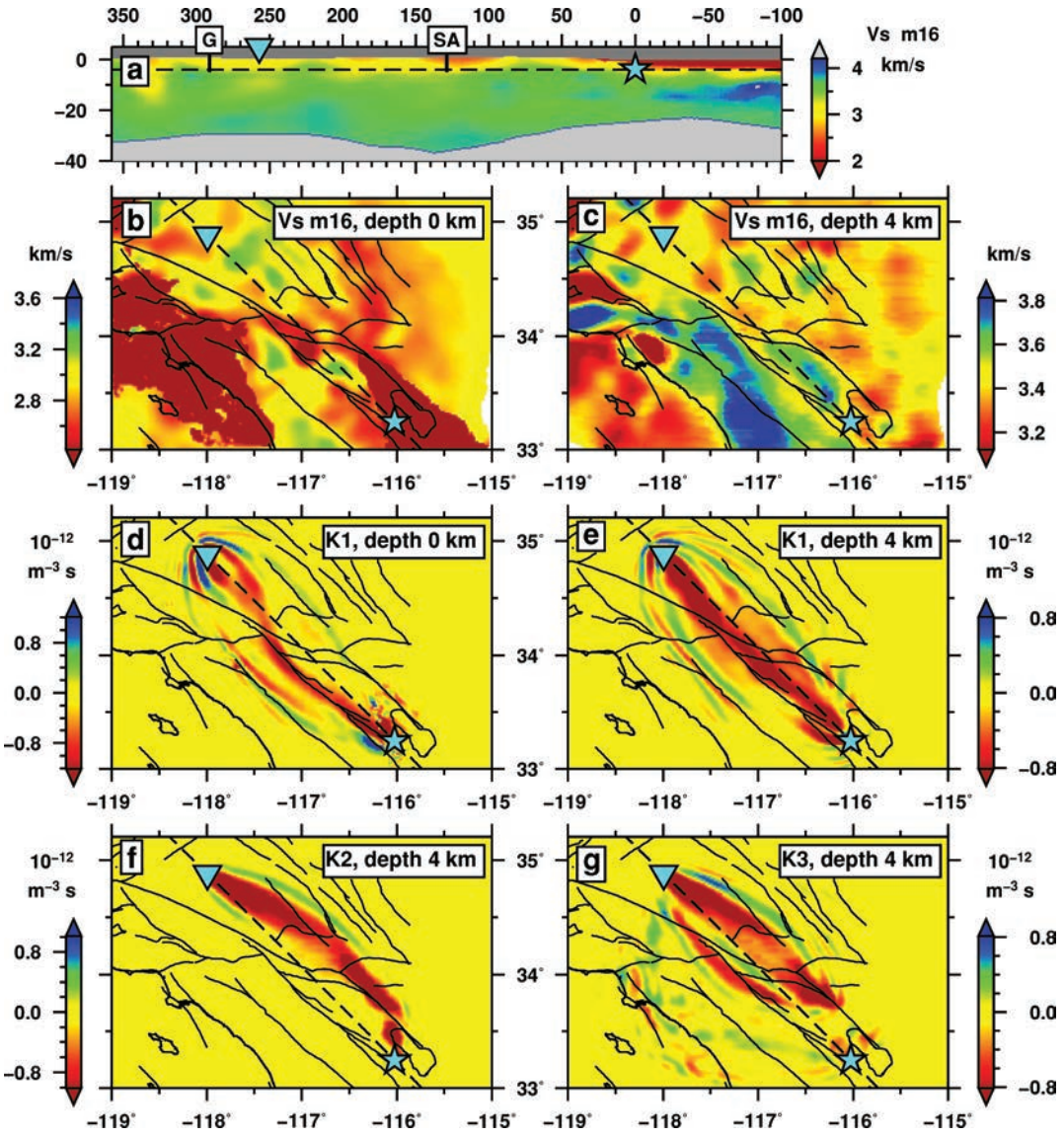


Figure 27. Sensitivity kernels for the three waveforms in Fig. 26, computed using the final model $\mathbf{m}_{16}$. See Section 8.2 for details. (a) Vertical cross-section of $\mathbf{m}_{16}$ containing the source ($\star$) and station (∇), as well as the intersections with the San Andreas (SA) and Garlock (G) faults. The dashed line denotes the 4 km depth of the horizontal cross-section in (c). (b) Horizontal cross-section of $\mathbf{m}_{16}$ at 0 km depth. (c) Horizontal cross-section of $\mathbf{m}_{16}$ at 4 km depth. (d) Horizontal cross-section of the V_S volumetric sensitivity kernel at 0 km depth, for time window 1 in Fig. 26. (e) Same as (d), but for a cross-section cut at 4 km depth. (f) Horizontal cross-section of the V_S kernel at 4 km depth, for time window 2 in Fig. 26. (g) Same as (f), but for time window 3.

ACKNOWLEDGMENTS

We thank Andreas Fichtner and Guust Nolet for reviews that improved this manuscript. This study would not have been possible without either seismic waveforms or parallel computing facilities. Seismic waveforms were provided by the Southern California Earthquake Data Center (SCEDC), the Northern California Earthquake Data Center (NCEDC), the Incorporated Research Institutions for Seismology (IRIS) and the University of Nevada Reno (Glenn Biasi) (see Table 1). All earthquake simulations were performed on the CITerra Dell cluster at the Division of Geological & Planetary Sciences (GPS) of the California Institute of Technology (Caltech). Egill Hauksson, Jeanne Hardebeck, Guoqing Lin, Don Helmberger and Shengji Wei provided unpublished earthquake source parameters. Paul Friberg executed the earthquake source inversions for the initial model. We thank Malcolm Sambridge for discussions and for suggesting the use of a source subspace projection method to compute the model update in the tomographic inversion. We acknowledge many valuable discussions with Albert Tarantola during his extended visits to Caltech (2007) and Princeton (2009). We utilized SAC and GMT software for processing seismograms and for plotting the figures (Goldstein *et al.* 2002; Wessel & Smith 1991). The Fortran90 software packages SPECFEM3D and FLEXWIN are available for download from the Computational Infrastructure for Geodynamics (www.geodynamics.org). Faria Chowdhury of the Southern California Earthquake Data Center constructed the interactive website (Section 8) for displaying cross-sections of the tomographic model. We acknowledge support by the National Science Foundation under grant EAR-0711177. This research was supported by the Southern California Earthquake Center. SCEC is funded by

NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number for this paper is 1297.

REFERENCES

Aagaard, B.T. *et al.*, 2008a. Ground-motion modeling of the 1906 San Francisco earthquake. Part II: gound-motion estimates for the 1906 earthquake and scenario events, *Bull. seism. Soc. Am.,* **98**(2), 1012–1046.

Aagaard, B.T. *et al.*, 2008b. Ground-motion modeling of the 1906 San Francisco earthquake. Part I: validation using the 1989 Loma Prieta earthquake, *Bull. seism. Soc. Am.,* **98**(2), 989–1011.

Agnew, D.C. & Sieh, K.E., 1978. A documentary study of the felt effects of the great California earthquake of 1857, *Bull. seism. Soc. Am.,* **68**(6), 1717–1729.

Akçelik, V. *et al.*, 2003. High resolution forward and inverse earthquake modeling on terascale computers, in *Proceedings of the ACM/IEEE Supercomputing Conference (SC'03)*, doi:10.1109/SC.2003.10056.

Amante, C. & Eakins, B.W., 2008. ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis, National Geophysical Data Center, NESDIS, NOAA, U.S. Dept. Commerce, Boulder, CO, USA.

Anderson, D.L., 1961. Elastic wave propagation in layered anisotropic media, *J. geophys. Res.*, **66**(9), 2953–2963.

Bamberger, A., Chavent, G., Hemon, C. & Lailly, P., 1982. Inversion of normal incidence seismograms, *Geophysics*, **47**(5), 757–770.

Ben-Menahem, A., 1978. Source mechanism of the 1906 San Francisco earthquake, *Phys. Earth planet. Inter.,* **17,** 163–181.

Ben-Zion, Y. & Andrews, D.J., 1998. Properties and implications of dynamic rupture along a material interface, *Bull. seism. Soc. Am.,* **88**(4), 1085–1094.

Bird, P., 2003. An updated digital model of plate boundaries, *Geochem. Geophys. Geosyst.,* **4,** 1027, doi:10.1029/2001GC000252.

Bonner, J.L., Blackwell, D.D. & Herrin, E.T., 2003. Thermal constraints on earthquake depths in California, *Bull. seism. Soc. Am.,* **93**(6), 2333–2354.

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

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

Burchfiel, B., Cowan, D.S. & Davis, G.A., 1992. Tectonic overview of the Cordilleran orogen in the western United States, in *The Cordilleran Orogen: Conterminous U.S.*, Vol. G-3: The Geology of North America, pp. 407–479, eds Burchfiel, B.C., Lipman, P.W. and Zoback, M.L., Geol. Soc. Am., Boulder, CO, USA.

Capdeville, Y., Chaljub, E., Vilotte, J.P. & Montagner, J.P., 2003. Coupling the spectral element method with a modal solution for elastic wave propagation in global earth models, *Geophys. J. Int.,* **152,** 34–67.

Chavent, G., Dupuy, M. & Lemonnier, P., 1975. History matching by use of optimal theory, *Soc. Petrol. Eng. J.,* **15**(1), 74–86.

Cheadle, M.J. *et al.*, 1986. The deep crustal structure of the Mojave Desert, California, from COCORP seismic reflection data, *Tectonics*, **5**(2), 293–320.

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.

Christensen, N.I., 1996. Poisson's ratio and crustal seismology, *J. geophys. Res.,* **101**(B2), 3139–3156.

Christensen, N.I. & Mooney, W.D., 1995. Seismic velocity structure and composition of the continental crust: a global view, *J. geophys. Res.,* **100**(B7), 9761–9788.

Claerbout, J.F., 1971. Toward a unified theory of reflector mapping, *Geophysics*, **36**(3), 467–481.

Clinton, J.F., Hauksson, E. & Solanki, K., 2006. An evaluation of the SCSN moment tensor solutions: robustness of the M_w magnitude scale, style of faulting, and automation of the method, *Bull. seism. Soc. Am.,* **96**(5), 1689–1705.

Custódio, S., Liu, P. & Archuleta, R.J., 2005. The 2004 M_w6.0 Parkfield, California, earthquake: inversion of near-source ground motion using multiple data sets, *Geophys. Res. Lett.,* **32,** L23312, doi:10.1029/2005GL024417.

Dahlen, F.A. & Baig, A.M., 2002. Fréchet kernels for body-wave amplitudes, *Geophys. J. Int.,* **150,** 440–446.

Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology*, Princeton Univ. Press, Princeton, NJ, USA.

Das, T. & Nolet, G., 1998. Crustal thickness map of the western United States by partitioned waveform inversion, *J. geophys. Res.,* **103**(B12), 30 021–30 038.

Davis, T.L., 1988. Structure of the Cuyama Valley, Caliente Range, and Carrizo Plain and its significance to the structural style of the southern Coast Ranges and western Transverse Ranges, in *Tertiary Tectonics and Sedimentation in the Cuyama Basin, San Luis Obispo, Santa Barbara, and Ventura Counties, California*, pp. 141–158, ed. Bazeley, W.J.M., Pacific Section, S.E.P.M., Los Angeles, CA, USA.

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

Dziewonski, A. & Anderson, D., 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.*, **25,** 297–356.

Faccioli, E., Maggio, F., Paolucci, R. & Quarteroni, A., 1997. 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method, *J. Seism.*, **1,** 237–251.

Fay, N.P., Bennett, R.A., Spinler, J.C. & Humphreys, E.D., 2008. Small-scale upper mantle convection and crustal dynamics in southern California, *Geochem. Geophys. Geosyst.,* **9**(8), Q08006, doi:10.1029/2008GC001988.

Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2008. Theoretical background for continental- and global-scale full-waveform inversion in the time–frequency domain, *Geophys. J. Int.,* **175,** 665–685.

Fichtner, A., Kennett, B.L.N., Igel, H. & Bunge, H.-P., 2009. Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophys. J. Int.,* **179,** 1703–1725.

Fuis, G.S., Ryberg, T., Godfrey, N.J., Okaya, D.A. & Murphy, J.M., 2001. Crustal structure and tectonics from the Los Angeles basin to the Mojave Desert, southern California, *Geology*, **29**(1), 15–18.

Fuis, G.S. *et al.*, 2003. Fault systems of the 1971 San Fernando and 1994 Northridge earthquakes, southern California: relocated aftershocks and seismic images from LARSE II, *Geology*, **31**(2), 171–174.

Fuis, G.S., Kohler, M.D., Scherwath, M., ten Brink, U., Van Avendonk, H.J.A. & Murphy, J.M., 2007. A comparison between the transpressional plate boundaries of South Island, New Zealand, and southern California, USA: The Alpine and San Andreas Fault systems, in *A Continental Plate Boundary: Tectonics at South Island, New Zealand*, Vol. 175: Geophys. Monogr. Series, pp. 307–327, eds Okaya, D., Stern, T. and Davey, F., American Geophysical Union, Washington, DC.

Goldstein, P., Dodge, D., Firpo, M. & Minner, L., 2002. SAC2000: Signal processing and analysis tools for seismologists and engineers, in *International Handbook of Earthquake and Engineering Seismology, Part B*, Vol. 81: International Geophysics Series, pp. 1613–1614, eds Lee, W.H.K., Kanamori, H., Jennings, P.C. and Kisslinger, C., Academic Press, London.

Goodman, E.D. & Malin, P.E., 1992. Evolution of the southern San Joaquin basin and mid-Tertiary "transitional" tectonics, central California, *Tectonics*, **11**(3), 478–498.

Grand, S.P., van der Hilst, R.D. & Widiyantoro, S., 1997. Global seismic tomography: a snapshot of convection in the Earth, *GSA Today*, **7**(4), 1–7.

Grove, M., Jacobson, C.E., Barth, A.P. & Vucic, A., 2003. Temporal and spatial trends of Late Cretaceous–early Tertiary underplating of Pelona and related schist beneath southern California and southwestern Arizona, in *Tectonic evolution of northwestern Mexico and the southwestern USA*,

pp. 381–406, eds Johnson, S.E., Paterson, S.R., Fletcher, J.M., Girty, G.H., Kimbrough, D.L. and Martin-Barajas, A., *Geol. Soc. Am.*, Boulder, CO, USA, Special Paper 388.

Hardebeck, J.L. & Shearer, P.M., 2003. Using S/P amplitude ratios to constrain the focal mechanisms of small earthquakes, *Bull. seism. Soc. Am.*, **93**(6), 2434–2444.

Hauksson, E., 2000. Crustal structure and seismicity distribution adjacent to the Pacific and North America plate boundary in southern California, *J. geophys. Res.*, **105**(B6), 13 875–13 903.

Hauksson, E. & Unruh, J., 2007. Regional tectonics of the Coso geothermal area along the intracontinental plate boundary in central eastern California: three-dimensional V_p and V_p/V_s models, spatial-temporal seismicity patterns, and seismogenic deformation, *J. geophys. Res.*, **112,** B06309, doi:10.1029/2006JB004721.

Hauksson, E., Oppenheimer, D. & Brocher, T.M., 2004. Imaging the source region of the 2003 San Simeon earthquake within the weak Franciscan subduction complex, central California, *Geophys. Res. Lett.*, **31,** L20607, doi:10.1029/2004GL021049.

Hauksson, E. *et al.*, 2008. Preliminary report on the 29 July 2008 Mw 5.4 Chino Hills, eastern Los Angeles basin, California, earthquake sequence, *Seism. Res. Lett.*, **79**(6), 855–866.

Humphreys, E.D. & Clayton, R.W., 1990. Tomographic image of the southern California mantle, *J. geophys. Res.*, **95**(B12), 19 725–19 746.

Igel, H., Djikpéssé, H. & Tarantola, A., 1996. Waveform inversion of marine reflection seismograms for *P* impedance and Poisson's ratio, *Geophys. J. Int.*, **124,** 363–371.

Jachens, R., Simpson, R., Graymer, R., Wentworth, C. & Brocher, T., 2006. Three-dimensional geologic map of northern and central California: a basic model for supporting earthquake simulations and other predictive modeling (abstract), *Seism. Res. Lett.*, **77,** 270.

Jacobson, C.E., Oyarzabal, F.R. & Haxel, G.B., 1996. Subduction and exhumation of the Pelona-Orocopia-Rand schists, southern California, *Geology*, **24**(6), 547–550.

Jenkins, O.P. & Oakeshott, G.B., 1955. *Earthquakes in Kern County, California During 1952*, State of California Natural Resources, Division of Mines, San Francisco, CA, USA, Bulletin 171.

Jennings, C.W., 1994. Fault activity map of California and adjacent areas, with locations and ages of recent volcanic eruptions, Calif. Div. Mines and Geology, Geologic Data Map No. 6, map scale 1:750,000.

Kanamori, H., 1977. The energy release in great earthquakes, *J. geophys. Res.*, **82,** 2981–2987.

Kanamori, H. & Hadley, D., 1975. Crustal structure and temporal velocity change in southern California, *Pure appl. Geophys.*, **113,** 257–280.

Keller, B. & Prothero, W., 1987. Western Transverse Ranges crustal structure, *J. geophys. Res.*, **92**(B8), 7890–7906.

Kennett, B.L.N., Sambridge, M.S. & Williamson, P.R., 1988. Subspace methods for large inverse problems with multiple parameter classes, *Geophys. J. Int.*, **94,** 237–247.

Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophys. J. Int.*, **139,** 806–822.

Komatitsch, D. & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophys. J. Int.*, **149,** 390–412.

Komatitsch, D. & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation—II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophys. J. Int.*, **150,** 308–318.

Komatitsch, D. & Vilotte, J.-P., 1998. The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, *Bull. seism. Soc. Am.*, **88**(2), 368–392.

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

Langenheim, V.E., 1999. Gravity and Aeromagnetic Models along the Los Angeles Region Seismic Experiment (Line 1), California, Open-File Report 99-388.

Laske, G. & Masters, G., 1996. Constraints on global phase velocity maps from long-period polarization data, *J. geophys. Res.*, **101**(B7), 16 059–16 075.

Le Pourhiet, L., Gurnis, M. & Saleeby, J., 2006. Mantle instability beneath the Sierra Nevada Mountains in California and Death Valley extension, *Earth planet. Sci. Lett.*, **251,** 104–119.

Legg, M.R., 1991. Developments in understanding the tectonic evolution of the California Continental Borderland, in *From Shoreline to Abyss*, pp. 291–312, ed. Osborne, R.H., SEPM, Tulsa, Oklahoma, USA.

Legg, M.R., Goldfinger, C., Kamerling, M.J., Chaytor, J.D. & Einstein, D.E., 2007. Morphology, structure and evolution of California Continental Borderland restraining bends, in *Tectonics of Strike-Slip Restraining and Releasing Bends*, pp. 143–168, eds Cunningham, W.D. and Mann, P., Geol. Soc. London, Special Publication 290.

Lin, G., Shearer, P.M. & Hauksson, E., 2007a. 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., Shearer, P.M., Hauksson, E. & Thurber, C.H., 2007b. A three-dimensional crustal seismic velocity model for southern California from a composite event method, *J. geophys. Res.*, **112,** B11306, doi:10.1029/2007JB004977.

Liu, Q. & Tromp, J., 2006. Finite-frequency kernels based on adjoint methods, *Bull. seism. Soc. Am.*, **96**(6), 2383–2397.

Liu, Q. & Tromp, J., 2008. Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods, *Geophys. J. Int.*, **174,** 265–286.

Liu, Q., Polet, J., Komatitsch, D. & Tromp, J., 2004. Spectral-element moment tensor inversions for earthquakes in southern California, *Bull. seis. Soc. Am.*, **94**(5), 1748–1761.

Lohman, R.B. & McGuire, J.J., 2007. Earthquake swarms driven by aseismic creep in the Salton Trough, California, *J. geophys. Res.*, **112,** B04405, doi:10.1029/2006JB004596.

Louie, J.N. & Clayton, R.W., 1987. The nature of deep crustal structures in the Mojave Desert, California, *Geophys. J. R. astr. Soc.*, **89,** 125–132.

Lovely, P., Shaw, J.H., Liu, Q. & Tromp, J., 2006. A structural V_P model of the Salton Trough, California, and its implications for seismic hazard, *Bull. seism. Soc. Am.*, **96**(5), 1882–1896.

Ludwig, W.J., Nafe, J.E. & Drake, C.L., 1970. Seismic refraction, in *New Concepts of Sea Floor Evolution*, Vol. 4: The Sea: Ideas and Observations on Progress in the Study of the Seas, pp. 53–84, ed. Maxwell, A.E., Wiley-Interscience, New York.

Luffi, P., Saleeby, J.B., Lee, C.-T.A. & Ducea, M.N., 2009. Lithospheric mantle duplex beneath the central Mojave Desert revealed by xenoliths from Dish Hill, California, *J. geophys. Res.*, **114,** B03202, doi:10.1029/2008JB005906.

Luo, Y. & Schuster, G.T., 1991. Wave-equation traveltime inversion, *Geophysics*, **56,** 645–653.

Luyendyk, B.P., Kamerling, M.J. & Terres, R., 1980. Geometric model for Neogene crustal rotations in southern California, *Geol. soc. Am. Bull.*, **91,** 211–217.

Maggi, A., Tape, C., Chen, M., Chao, D. & Tromp, J., 2009. An automated time-window selection algorithm for seismic tomography, *Geophys. J. Int.*, **178,** 257–281.

McLaren, M.K., Hardebeck, J.L., van der Elst, N., Unruh, J., Bawden, G.W. & Blai, J.L., 2008. Complex faulting associated with the 22 December 2003 M_W 6.5 San Simeon, California, earthquake, aftershocks, and postseismic surface deformation, *Bull. seism. Soc. Am.*, **98**(4), 1659–1680.

McMechan, G.A., 1982. Determination of source parameters by wavefield extrapolation, *Geophys. J. R. astr. Soc.*, **71,** 613–628.

Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data, *Geophysics*, **52**(9), 1211–1228.

Murphy, J.M., Fuis, G.S., Ryberg, T., Lutter, W.J., Catchings, R.D. & Goldman, M.R., 2008. Detailed P- and S-wave velocity models along the LARSE II transect, southern California, in *EOS, Trans. Am. geophys. Un.*, **89,** Fall. Meet. Suppl., Abstract S11C–1763.

Namson, J. & Davis, T.L., 1990. Late Cenozoic fold and thrust belt of the southern Coast Ranges and Santa Maria basin, California, *Am. Assoc. Petrol. Geol. Bull.*, **74**(4), 467–492.

Nazareth, J.J. & Clayton, R.W., 2003. Crustal structure of the Borderland-Continent Transition Zone of southern California adjacent to Los Angeles, *J. geophys. Res.*, **108**(B8), 2404, doi:10.1029/2001JB000223.

Nicholson, C., Sorlien, C.C. & Luyendyk, B.P., 1992. Deep crustal structure and tectonics in the offshore Santa Maria Basin, California, *Geology*, **20**(3), 239–242.

Nicholson, C., Sorlien, C.C., Atwater, T., Crowell, J.C. & Luyendyk, B.P., 1994. Microplate capture, rotation of the western Transverse Ranges, and initiation of the San Andreas transform as a low-angle fault system, *Geology*, **22**(6), 491–495.

Olsen, K.B. *et al.*, 2006. Strong shaking in Los Angeles expected from southern San Andreas earthquake, *Geophys. Res. Lett.*, **33,** L07305, doi:10.1029/2005GL0254720.

Pellerin, C.L.M. & Christensen, N.I., 1998. Interpretation of crustal seismic velocities in the San Gabriel–Mojave region, southern California, *Tectonophysics*, **286,** 253–271.

Percival, D. & Walden, A., 1993. *Spectral Analysis for Physical Applications*, Cambridge U. Press, Cambridge, UK.

Petersen, M.D. & Wesnousky, S.G., 1994. Fault slip rates and earthquake histories for active faults in southern California, *Bull. seism. Soc. Am.*, **84**(5), 1608–1649.

Plesch, A., Tape, C., Shaw, J.H. & members of the USR working group, 2009. CVM-H 6.0: Inversion integration, the San Joaquin Valley and other advances in the community velocity model, in *2009 Southern California Earthquake Center Annual Meeting, Proceedings and Abstracts*, Vol. 19, pp. 260–261.

Pratt, R.G., Shin, C.S. & Hicks, G.J., 1998. Gauss–Newton and full Newton methods in frequency–space seismic waveform inversion, *Geophys. J. Int.*, **133,** 341–362.

Ritsema, J., van Heijst, H.J. & Woodhouse, J.H., 1999. Complex shear velocity structure imaged beneath Africa and Iceland, *Science*, **286,** 1925–1928.

Rodgers, A., Petersson, A., Nilsson, S., Sjögreen, B. & McCandless, K., 2008. Broadband waveform modeling of moderate earthquakes in the San Francisco Bay area and preliminary assessment of the USGS 3D seismic velocity model, *Bull. seism. Soc. Am.*, **98**(2), 969–988.

Romanyuk, T., Mooney, W.D. & Detweiler, S., 2007. Two lithospheric profiles across southern California derived from gravity and seismic data, *J. Geodyn.*, **43,** 274–307.

Ryberg, T. & Fuis, G.S., 1998. The San Gabriel Mountains bright reflective zone: possible evidence of young mid-crustal thrust faulting in southern California, *Tectonophysics*, **286,** 31–46.

Saleeby, J., 2003. Segmentation of the Laramide Slab—evidence from the southern Sierra Nevada region, *Geol. soc. Am. Bull.*, **115**(6), 655–668.

Sambridge, M.S., Tarantola, A. & Kennett, B.L.N., 1991. An alternative strategy for non-linear inversion of seismic waveforms, *Geophys. Prospect.*, **39,** 723–736.

Savage, B. & Helmberger, D.V., 2004. Complex Rayleigh waves resulting from deep sedimentary basins, *Earth planet. Sci. Lett.*, **218,** 229–239.

Shapiro, N.M., Campillo, M., Stehly, L. & Ritzwoller, M.H., 2005. High-resolution surface-wave tomography from ambient seismic noise, *Science*, **307,** 1615–1618.

Sieh, K.E., 1978. Slip along the San Andreas fault associated with the great 1857 earthquake, *Bull. seism. Soc. Am.*, **68**(5), 1421–1448.

Sieminski, A., Liu, Q., Trampert, J. & Tromp, J., 2007. Finite-frequency sensitivity of surface waves to anisotropy based upon adjoint methods, *Geophys. J. Int.*, **168,** 1153–1174.

Stich, D., Danecek, P., Morelli, A. & Tromp, J., 2009. Imaging lateral heterogeneity in the northern Apennines from time reversal of reflected surface waves, *Geophys. J. Int.*, **177,** 543–554.

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, *J. geophys. Res.*, **108**(B3), 2170, doi:10.1029/2001JB001628.

Talagrand, O. & Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: theory, *Q. J. R. Meteorol. Soc.*, **113,** 1311–1328.

Tan, Y., 2006. Broadband Waveform Modeling Over a Dense Seismic Network, *PhD thesis*, California Institute of Technology, Pasadena, California, USA.

Tape, C., 2009. Seismic Tomography of Southern California Using Adjoint Methods, *PhD thesis*, California Institute of Technology, Pasadena, California, USA.

Tape, C., Liu, Q. & Tromp, J., 2007. Finite-frequency tomography using adjoint methods—methodology and examples using membrane surface waves, *Geophys. J. Int.*, **168,** 1105–1129.

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

Tape, C., Sambridge, M. & Tromp, J., 2010. Adjoint tomography based on source subspace projection, *Geophys. J. Int.*, in preparation.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**(8), 1259–1266.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia, USA.

ten Brink, U.S., Zhang, J., Brocher, T.M., Okaya, D.A., Klitgord, K.D. & Fuis, G.S., 2000. Geophysical evidence for the evolution of the California Inner Continental Borderland as a metamorphic core complex, *J. geophys. Res.*, **105**(B3), 5835–5857.

Thurber, C., Zhang, H., Waldhauser, F., Hardebeck, J., Michael, A. & Eberhart-Phillips, D., 2006. Three-dimensional compressional wavespeed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region, *Bull. seism. Soc. Am.*, **96**(4B), S38–S49.

Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal, and banana-doughnut kernels, *Geophys. J. Int.*, **160,** 195–216.

Virieux, J. & Operto, S., 2010. Full waveform inversion: theory, algorithms and applications, *Geophysics*, in press.

Wald, L.A., Hutton, L.K. & Given, D.D., 1995. The Southern California Network Bulletin: 1990–1993 summary, *Seism. Res. Lett.*, **66**(1), 9–19.

Wessel, P. & Smith, W.H.F., 1991. Free software helps map and display data, *EOS, Trans. Am. geophys. Un.*, **72**(41), 441.

Yan, Z. & Clayton, R.W., 2007. Regional mapping of the crustal structure in southern California from receiver functions, *J. geophys. Res.*, **112,** B05311, doi:10.1029/2006JB004622.

Zhou, Y., Dahlen, F.A. & Nolet, G., 2004. Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.*, **158,** 142–168.

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

Zhu, L. & Kanamori, H., 2000. Moho depth variation in southern California from teleseismic receiver functions, *J. geophys. Res.*, **105**(B2), 2969–2980.

APPENDIX A: UNCERTAINTY ESTIMATE FOR MEASUREMENTS

The cross-correlation function is given by

$$\Gamma(\tau) = \int s(t-\tau)\, d(t)\, \mathrm{d}t, \tag{A1}$$

where d denotes the observed seismogram and s the synthetic. Let $\tau = \Delta T$ denote the cross-correlation traveltime anomaly associated with the maximum of $\Gamma(\tau)$. Let $\Delta \ln A$ denote the amplitude anomaly (Dahlen & Baig 2002; Maggi *et al.* 2009, eq. 11):

$$\Delta \ln A = \ln\left(\frac{A^{\mathrm{obs}}}{A^{\mathrm{syn}}}\right) = 0.5\, \ln\left[\frac{\int d^2(t)\, \mathrm{d}t}{\int s^2(t)\, \mathrm{d}t}\right]. \tag{A2}$$

We seek to determine σ_{T} and σ_{A} associated with ΔT and $\Delta \ln A$. Therefore we write

$$d(t) = [\exp(\Delta \ln A) \pm \sigma_{\mathrm{A}}]\, s(t - \Delta T \pm \sigma_{\mathrm{T}})\ . \tag{A3}$$

Expanding the second term to first order, we obtain

$$\begin{aligned} d(t) &\approx [\exp(\Delta \ln A) \pm \sigma_{\mathrm{A}}]\, [s(t-\Delta T) \pm \sigma_{\mathrm{T}}\, \dot{s}(t-\Delta T)] \\ &= [\exp(\Delta \ln A) \pm \sigma_{\mathrm{A}}]\, s(t-\Delta T) \pm \sigma_{\mathrm{T}}\, [\exp(\Delta \ln A) \pm \sigma_{\mathrm{A}}]\, \dot{s}(t-\Delta T) \\ &= \exp(\Delta \ln A)\, s(t-\Delta T) \pm \sigma_{\mathrm{A}}\, s(t-\Delta T) \pm \sigma_{\mathrm{T}}\, \exp(\Delta \ln A)\, \dot{s}(t-\Delta T) \pm \sigma_{\mathrm{T}}\, \sigma_{\mathrm{A}}\, \dot{s}(t-\Delta T). \end{aligned}$$

To first order in σ_{T} and σ_{A}, this may be written as

$$d(t) - \exp(\Delta \ln A)\, s(t-\Delta T) = \pm\sigma_{\mathrm{T}} \exp(\Delta \ln A)\, \dot{s}(t-\Delta T) \pm \sigma_{\mathrm{A}}\, s(t-\Delta T). \tag{A4}$$

Squaring (A4) and integrating over time, we obtain

$$\begin{aligned} \int [d(t) - \exp(\Delta \ln A)\, s(t-\Delta T)]^2\, \mathrm{d}t \ = \ & \sigma_{\mathrm{T}}^2 \int [\exp(\Delta \ln A)\, \dot{s}(t-\Delta T)]^2\, \mathrm{d}t \\ & + \sigma_{\mathrm{A}}^2 \int [s(t-\Delta T)]^2\, \mathrm{d}t \\ & \pm \sigma_{\mathrm{T}}\, \sigma_{\mathrm{A}} \int [\exp(\Delta \ln A)\, \dot{s}(t-\Delta T)\, s(t-\Delta T)]\, \mathrm{d}t. \end{aligned} \tag{A5}$$

If we assume the errors are uncorrelated, we find that

$$\sigma_{\mathrm{T}}^2 = \frac{\int [d(t) - \exp(\Delta \ln A)\, s(t-\Delta T)]^2\, \mathrm{d}t}{\int [\exp(\Delta \ln A)\, \dot{s}(t-\Delta T)]^2\, \mathrm{d}t}\ , \tag{A6}$$

$$\sigma_{\mathrm{A}}^2 = \frac{\int [d(t) - \exp(\Delta \ln A)\, s(t-\Delta T)]^2\, \mathrm{d}t}{\int [s(t-\Delta T)]^2\, \mathrm{d}t}\ . \tag{A7}$$

If one is using σ_{T} or σ_{A} in the denominator of the misfit function, then one must specify a non-zero water-level value for each. Otherwise, for a perfect cross-correlation measurement, $\sigma_{\mathrm{T}} = \sigma_{\mathrm{A}} = 0$, and the adjoint source (and therefore event kernel) is singular. As discussed in Section 3.4, we chose $\sigma_0 = 1.0$ s for the tomographic inversion. Note that amplitude measurements were not explicitly used in the tomographic inversion; however, they were used in computing σ_{T} (A6), which appears as σ_i in (3).

APPENDIX B: MODEL VECTOR, MODEL COVARIANCE MATRIX, AND GRADIENT

In the tomographic inversion, the entries of $\mathbf{m}$ contain log-normalized values of V_{S} and V_{B}, which we define as β and c, respectively:

$$\beta = \ln(V_{\mathrm{S}}/V_{\mathrm{S}_0}), \tag{B1}$$

$$c = \ln(V_{\mathrm{B}}/V_{\mathrm{B}_0}). \tag{B2}$$

The steepest ascent vector, $\mathbf{g}$, is related to the gradient, $\hat{\mathbf{g}}$, via (e.g. Tarantola 2005)

$$\mathbf{g} = \mathbf{C}_{\mathrm{M}}\, \hat{\mathbf{g}}, \tag{B3}$$

$$\hat{\mathbf{g}} = \mathbf{C}_{\mathrm{M}}^{-1}\, \mathbf{g}. \tag{B4}$$

A gradient-based minimization algorithm requires taking norms of $\mathbf{m}$ and $\mathbf{g}$. The L2-norm operation requires a covariance matrix, $\mathbf{C}_{\mathrm{M}}$:

$$\|\mathbf{m}\|_2^2 = \mathbf{m}^T\, \mathbf{C}_{\mathrm{M}}^{-1}\, \mathbf{m}, \tag{B5}$$

$$\|\mathbf{g}\|_2^2 = \mathbf{g}^T\, \mathbf{C}_{\mathrm{M}}^{-1}\, \mathbf{g}, \tag{B6}$$

$$\left\|\hat{\mathbf{g}}\right\|_2^2 = \hat{\mathbf{g}}^T\, \mathbf{C}_{\mathrm{M}}\, \hat{\mathbf{g}}. \tag{B7}$$

In theory, $\mathbf{C}_{\mathrm{M}}$ should contain all prior information about the distribution of allowable models. In practice, much of the prior information is specified in choices of smoothing and regularization (Section 4.1). The diagonal entries of $\mathbf{C}_{\mathrm{M}}$ must be chosen according to the units of the entries m_j, as well as the 'weight' of each m_j that is based on the model parametrization.

Considering the two inversion variables V_{S} and V_{B}, we organize the entries of $\mathbf{m}$, $\hat{\mathbf{g}}$, and the diagonal of $\mathbf{C}_{\mathrm{M}}$ as follows:

$$\begin{aligned}
\mathbf{m} &= [\beta_1 \quad \cdots \quad \beta_j \quad \cdots \quad \beta_G \quad c_1 \quad \cdots \quad c_j \quad \cdots \quad c_G]^T, \\
\hat{\mathbf{g}} &= [K_1^{\beta} V_1 \quad \cdots \quad K_j^{\beta} V_j \quad \cdots \quad K_G^{\beta} V_G \quad K_1^{c} V_1 \quad \cdots \quad K_j^{c} V_j \quad \cdots \quad K_G^{c} V_G]^T, \\
\mathbf{C}_{\mathrm{M}_{jj}} &= [\sigma_\beta^2 V/V_1 \cdots \sigma_\beta^2 V/V_j \cdots \sigma_\beta^2 V/V_G \; \sigma_c^2 V/V_1 \cdots \sigma_c^2 V/V_j \cdots \sigma_c^2 V/V_G]^T,
\end{aligned} \tag{B8}$$

where K^{β} and K^{c} are the event kernels corresponding to V_{S} and V_{B}, V_j is the volume corresponding to the jth gridpoint in the mesh, V is the simulation volume, such that

$$V = \sum_{j=1}^{G} V_j, \tag{B9}$$

and σ_β and σ_c are the uncertainty estimates for β and c, which we assume to be fixed. Because the gridpoints in a spectral-element mesh are not uniformly spaced, it is important to consider the V_j factors.

APPENDIX C: VARIANCE REDUCTION

The variance reduction, VR, corresponding to the reduction in misfit in going from model $\mathbf{m}_1$ to $\mathbf{m}_2$, is given by

$$\mathrm{VR} = \frac{F(\mathbf{m}_1) - F(\mathbf{m}_2)}{F(\mathbf{m}_1)}\,, \tag{C1}$$

where $F(\mathbf{m}) > 0$ is a least-squares misfit function, which we assume is positive, that is, there is some water-level minimum misfit value. The formula for VR is a first-order approximation for a logarithmic difference:

$$\mathrm{VRL} = \ln\left[\frac{F(\mathbf{m}_1)}{F(\mathbf{m}_2)}\right]. \tag{C2}$$

The relationship between VR and VRL is given by

$$\mathrm{VRL} = \ln\left(\frac{1}{1-\mathrm{VR}}\right), \tag{C3}$$

$$\mathrm{VR} = 1 - e^{-\mathrm{VRL}}. \tag{C4}$$