Dynamic upwelling beneath the Salton Trough imaged with teleseismic attenuation tomography

Joseph S. Byrnes[1*] and Maximiliano Bezada[1]

[1]*Department of Earth Sciences, University of Minnesota, 116 Church Street SE Minneapolis, MN 55455 USA*

To whom correspondence should be addressed; Email jsbyrnes@umn.edu

Key points:

- P waves recorded in the Salton Trough are strongly attenuated
- Qp likely requires the presence of melt in the asthenosphere
- Buoyancy forces drive a robust melting column

ABSTRACT

The Salton Trough is one of the few regions on Earth where rifting is sub-aerial instead of sub-marine. We use the relative attenuation of teleseismic P phases recorded by the Salton Trough Seismic Imaging Project to investigate lithospheric and asthenospheric structures that form during extension. Map-view analysis reveals stronger attenuation within the Salton Trough than in the adjacent provinces. We then construct tomographic models for variations in seismic attenuation with depth to discriminate between crustal and mantle signals with a damped least-squares approach and a Bayesian approach. Synthetic tests show that models from damped least-squares significantly under-estimate the strength of attenuation and cannot separate crustal and mantle signals even when the tomographic models are allowed to be discontinuous at the lithosphere-asthenosphere boundary. We show that a Bayesian approach overcomes these problems when inverting the same synthetic datasets, and that shallow and deep signals are more clearly separated when imposing a discontinuity. With greater than 95% confidence, the results reveal first, that attenuation occurs primarily beneath the LAB; second, that the width of the attenuative region is narrower than the rift at 120 km depth; and third, that the strength of attenuation requires that the attenuative feature represents a melting-column similar to those beneath mid-ocean ridges. The narrow width of the melting-column below the volatile-free solidus is inconsistent with models for passive upwelling, where flow is driven only by rifting. Instead, we attribute the generation of incipient oceanic crust to mantle upwelling focused by buoyancy into a narrow diapir.

1. Introduction

The Salton Trough is a low-lying transtensional basin in Southern California formed by motion between the Pacific and North America plates (Elders et al., 1972; Powell and Weldon, 1992; Stock and Hodges, 1989). The San Andreas and associated faults accommodate transform motion between the Pacific and North American plates to the north, while extension occurs in several segments to the south. As far south as the Cabo segment (Figure 1a), the plate boundary has fully transitioned into the East Pacific Rise, and the production of oceanic crust occurs at least as far north as the Guaymas segment (Lizarralde et al., 2007). Buoyancy-driven mantle upwelling occurs approximately every 250 km both north and south of the Guaymas segment at points not clearly associated with the spreading centers (Wang et al., 2009). Several lines of evidence show that mantle melting in response to such an upwelling occurs as far north as the Salton Trough, including Holocene eruptions (Schmitt et al., 2013; Wright et al., 2015) that exhume young basaltic xenoliths (Schmitt and Vazquez, 2006) and low velocity anomalies in the upper 200 km of the mantle (Humphreys et al., 1984; Schmandt and Humphreys, 2010). Geophysical evidence further supports a model where extension has rifted the mantle lithosphere beneath the Salton Trough, including high heat flow (Lachenbruch et al., 1985), crustal thinning (Zhu and Kanamori, 2000), and the shoaling of a seismic discontinuity inferred to be the lithosphere-asthenosphere boundary, or LAB (Lekić et al., 2011).

The Salton Trough Seismic Imaging Project (SSIP) (Rose et al., 2013) deployed an array of broad-band seismometers across the Salton Trough (Figure 1b). The western half of the array lies in the Peninsular Ranges, a batholith created by the subduction of the Farallon Slab (Dickinson, 2009 and references therein). The motion of the Pacific plate brought the Peninsular Ranges from approximately 200 km further south to the western edge of the Salton Trough (Darin and Dorsey, 2013). The array crosses from the Peninsular Ranges to the low-lying regions where sediments sourced from the Colorado River (Figure 1a, Dorsey, 2010) are typically 2-3 km thick (Han et al., 2016) and locally as thick as 7 km (Persaud et al., 2016). The depositional rate is great enough to delay the complete rupture of the crust and transition to sea-floor spreading (Han et al., 2016). The array lies just south of the volcanically active Salton Buttes (Figure 1b). Beneath this region, tomographic imaging reveals low shear-wave velocities below 40 km depth (Barak et al., 2015) where previous studies inferred that rifting has thinned the mantle lithosphere (Lekić et al., 2011). The mantle at the base of the moho has seismic velocities lower than the global average for cold continental lithosphere (Barak et al., 2015; Han et al., 2016) but typical of young oceanic lithosphere and higher than in the asthenosphere (e.g, Vanderbeek and Toomey, 2017). Moreover, while seismic anisotropy is thought to be weak below the lithosphere across much of Southern California (Monteiller and Chevrot, 2011), the splitting times of *SKS* waves increase where the array crosses into the trough (Barak and Klemperer, 2016). This confluence of low seismic velocities with strong seismic anisotropy suggests the presence of melt in the asthenosphere that has been organized by shear (Holtzman et al., 2012, 2003; Holtzman and Kendall, 2010; Kohlstedt and Holtzman, 2009; Barak and Klemperer, 2016).

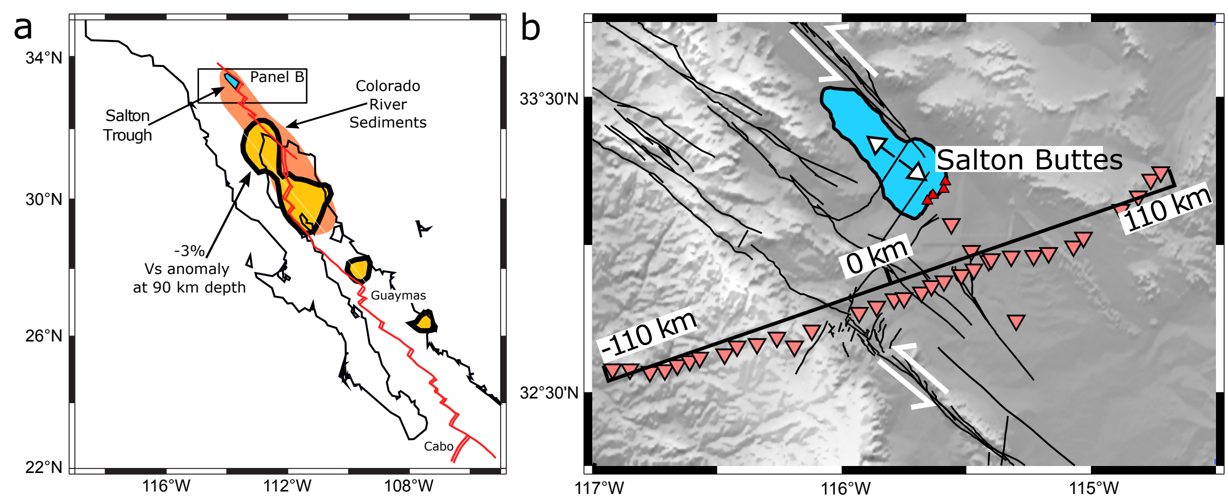

Figure 1. Tectonic setting of the Salton Trough. a) Thin black lines are coasts, blue is the Salton Sea, single red lines are transform faults, double red lines are spreading centers, and yellow regions are low seismic velocity anomalies from *Wang et al.,* (2009). b) Inverted pink triangles are seismic stations used in this study, thin black lines are faults from *Plesch et al.,* (2007), red triangles are the active volcanoes knows as the Salton Buttes, thick black line is the x-axis referred to in the main text, white arrows show relative motion along faults, and the shaded background is topography.

This study measures the attenuation of teleseismic *P* phases recorded by the SSIP broad-band array to better understand upwelling and mantle melting processes beneath the Salton Trough. Seismic attenuation, the loss of seismic energy during propagation, provides an observational constraint on the physical state of the Earth that is complementary to seismic velocity. The density of the SSIP array, in a region with geological and geophysical evidence for not only melt production but the presence of *in-situ* melt in the upper mantle, provides a unique opportunity to explore the relationship between melt production, mantle upwelling, and seismic attenuation. Significant unknowns remain regarding how seismic attenuation relates to the presence of melt in the upper mantle, and predictions for the relationship between melt and seismic attenuation range from no effect to strong effects (Faul et al., 2004; Hammond and

Humphreys, 2000; McCarthy and Takei, 2011). However, the presence of seismic attenuation stronger than can be explained by the current generation of models where attenuation depends solely on temperature and grain size provides evidence for the presence of melt as an additional attenuative mechanism (Abers et al., 2014, Eilon and Abers, 2017). Alternatively, a recently characterized reduction in seismic velocity at temperatures above 94% of the solidus (Takei, 2017; Yamauchi and Takei, 2016) may explain seismic anomalies without appealing to *in-situ* melt. Under either scenario, however, mapping seismic attenuation in the upper mantle can define the contours of the region of melt-production at those depths. Since mantle melting away from subduction zones is caused by decompression during upwelling, mapping the region of melt production also constrains the possible pattern of mantle flow in the upper mantle.

In this study we observe strong seismic attenuation at seismic stations within the Salton Trough. We apply two approaches for attenuation tomography to determine the distribution of the attenuation with depth: Damped least-squares and a transdimensional Monte Carlo search (Bodin and Sambridge, 2009; Green, 1995; Sambridge et al., 2006). We find that the damped least-squares approach cannot identify the depth at which the attenuation occurs nor the strength of attenuation due to the high levels of noise in the dataset and the limited distribution of ray-paths available in this study. Inversions of synthetic datasets confirm that the transdimensional approach can correctly identify the depth at which attenuation occurs and more accurately estimate the amplitude of anomalies, particularly when a discontinuity in the model is included. Inversion of the data reveals strong attenuation at depths below the LAB down to ~120 km. This is well-below the volatile-free solidus for fertile mantle and is interpreted in terms of a deep and narrow melting column due to the presence of volatiles. While the most probable solution to the model includes sufficiently strong attenuation to require that the presence of *in-situ* melt is the attenuative mechanism, the 95% confidence bound on the strength of attenuation in the mantle is plausibly consistent with a melt-free upper mantle that has reached the solidus according to the predictions of *Yamauchi and Takei* (2016). However, the results show that the width of the deep melting-column tapers to become narrower than the width of the rifted region. This tapering is consistent with geodynamic models of buoyancy-driven upwelling, which leads to downwelling on the flanks of a narrow diapir-like column. While the northward progression of rifting between the Pacific and North American plates has not wholly ruptured the continental crust due to the high rate of sedimentation (Han et al., 2016) and produced limited volcanism, the asthenospheric mantle behaves like a mid-ocean ridge.

2. Map view analysis

2.1 Measurements of Δt_P^*

We use the *P* phase from eight deep focus events recorded by the SSIP array (Table 1). Phases from deep events are preferred because of their higher frequency content relative to phases from shallow events. Apparent attenuation was measured with an approach that models the waveform in the time-domain (Bezada, 2017). An estimate of the source-time function is made by stacking the narrowest *P* waves in the dataset after applying a causal instrument response correction (Haney et al., 2012) and filtering between 0.01 and 1.5 Hz with a fourth-order Butterworth filter (Figure 2a). The degree of attenuation experienced by a *P* phase is described by the quantity t_P^*,

$$t_P^* = \int Q_p^{-1}(t)dt = \int Q_p^{-1}(r)V_p^{-1}(r)dr \qquad (1)$$

where *t* is the travel-time, $Q_P(t)$ is the quality factor encountered by the phase at a given time during propagation, *r* is distance along the ray path, and $V_P(r)$ and $Q_P(r)$ are the velocity and quality factor along the ray path. In this study, we only measure relative changes to t_P^*, which we refer to as Δt_P^*, which can be expressed as

$$\Delta t_P^* = \int \Delta Q_p^{-1}(r) \Delta V_p^{-1}(r) dr \quad (2)$$

where ΔQ_p^{-1} and ΔV_p^{-1} are deviations from an unknown reference model for quality factor and velocity for *P* phases. Variably attenuated versions of the estimate source were generated by convolution of the estimated source-time function with the impulse response for different values of Δt_P^* (Azimi et al., 1968), which were then fit to the *P* arrivals with a grid search over Δt_P^* in 0.01 s increments. Examples of attenuated synthetics compared with data are shown in Figure 2a, along with highlighted examples of *P* waves that have been attenuated to greater and lesser degrees. For more negative values of Δt_P^* the waveform is narrower in the time domain and contains more high frequency features than the reference trace. As Δt_P^* increases, the waveform broadens in response to the loss of high frequency energy and dispersion. Results for the measurement made in Figure 2a are shown alongside results with the spectral slope method (Figure 2b,c). As expected, the time-domain method returns less scattered measurements with less sensitivity to small changes in the maximum frequency considered (Bezada et al., 2019).

2.2 Mapping inversion

We first invert for a two-dimensional map of the Δt_P^* measurements (Bezada, 2017, Byrnes et al., 2019). Often, Δt_P^* are analyzed in map-view (e.g., Bezada, 2017; Cafferky and Schmandt, 2015; Dong and Menke, 2017; Eilon et al., 2018; Hwang et al., 2009) because of the small number of events suitable for Δt_P^* measurements and the near-vertical ray paths of teleseismic waves at depths shallower than 200 km. This approach averages measurements from different azimuths and epicentral distances and shows relative variations in the ray-path integrated Δt_P^*.

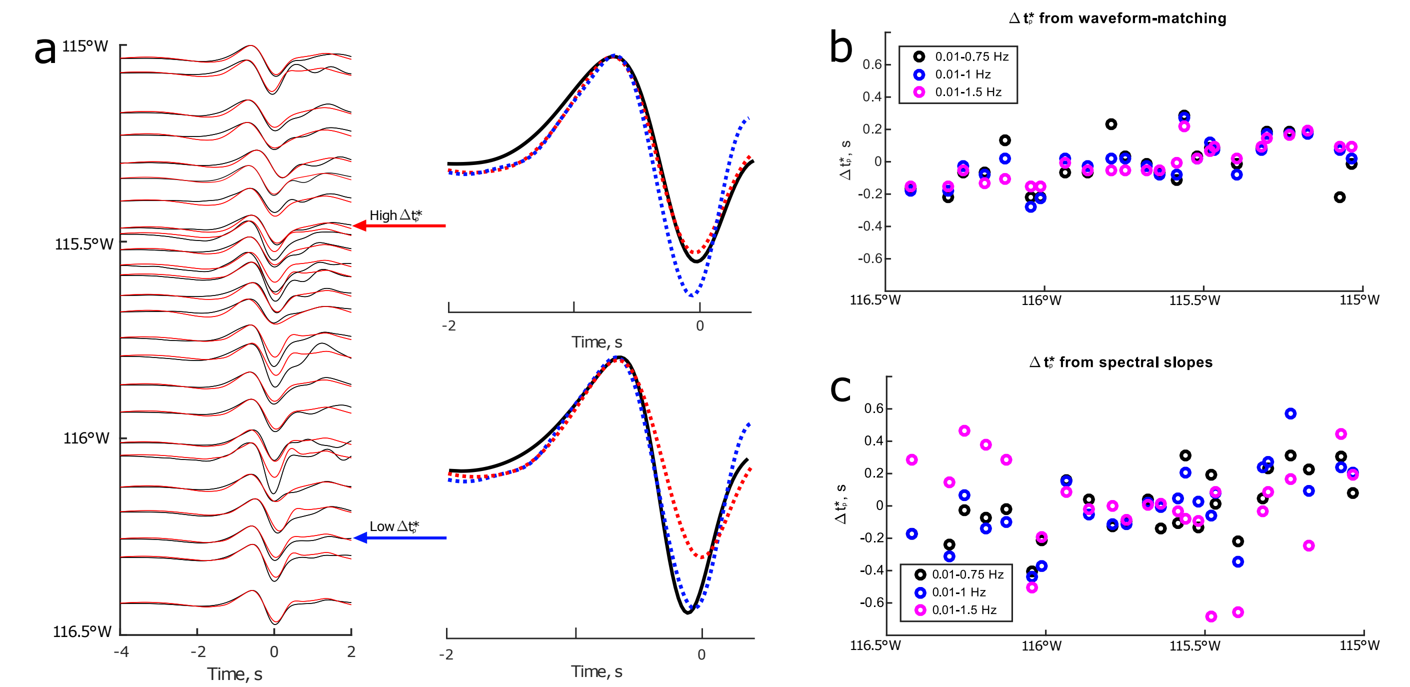

Figure 2. **Demonstration of the waveform matching approach.** a) First arriving *P* waves from a magnitude 5.9 event from the Izu-Bonin region recorded by the SSIP array in black, with the best fitting synthetic traces in red. Two traces are enlarged, the pair of which were measured to have a Δt_P^* of 0.15 s with the waveform-matching approach. The high and low Δt_P^* synthetics compared to the enlarged traces are in red and blue, respectively. b) Results for Δt_P^* with the waveform-matching approach for three different frequency bands. c) Results for Δt_P^* with the spectral-slope approach for three different frequency bands.

We invert the data for a smooth map of Δt_P^* with a Bayesian approach (Byrnes et al., 2019) that, heuristically, solves for a solution only as complex as is required to explain the data. The inversion iterates a Markov Chain that treats both the dimensions of the model and the uncertainty of the data as unknowns referred to as ‘hyperparameters’ (Malinverno and Briggs, 2004). Independent chains are initialized with a model describing Δt_P^* via nearest-neighbor interpolation from a set of nodes. The number of nodes, location of nodes within the model domain, and Δt_P^* associated with each node are randomly drawn from uniform distributions. Here, the bounds are 5 to 30 nodes, each restricted to within a region shown in Figure 3, and with Δt_P^* values within -0.5 to 0.5 s. On each iteration of each chain, a new model is proposed by perturbing the current model, either by a “birth” step (the introduction of a new node), a “death” step (the removal of a node), a "move" step (moving a node to a new location), a "change" step (changing the Δt_P^* value of a node), or a change to the uncertainty of the data. A set of subjectively chosen parameters describe how the search proceeds, but do not influence the shape of the final model so long as a sufficient number of iterations are performed (MacKay, 2003). The Δt_P^* of a node, the location of a node, and the value of the uncertainty may be changed on an iteration of a chain by drawing a perturbation to the current value from a normal distribution with a standard deviation of 0.01 s, 2.2 km, and 0.01 s, respectively. We use the acceptance criteria of *Bodin et al*., (2009; 2012b), which are derived from Bayes' theorem, to decide whether or not a perturbed model is accepted as the new position of the chain. We initialize 48 independent chains and iterate each chain 1e5 times. Models generated during the first 5e4 iterations are discarded, after which every 500th model is saved.

The result of this approach is not a single model subject to *a priori* assumptions but an ensemble of models that describe the probability density functions for model parameters in light of the data. The individual models are Voronoi diagrams, but the average of a large ensemble of models will be appropriately smooth. This removes the need to tune subjective regularization parameters so as not to overfit the data with models that are too rough, or to exclude real features by over-damping the inversion, even when the uncertainty of the data is not known ahead of time (Bodin et al., 2012a).

2.3 Results

The median of the ensemble of models from the search shows distinct variations in the attenuation of *P* phases that correlate with tectonic features (Figure 3a). The resolution of the models will be determined by the station spacing and, since no smoothing length is assumed, both gradational and sharp features will be allowed if required by the data (Byrnes et al., 2019).

From west to east, Δt_P^* increases rapidly at approximately 116°W, where the SSIP array crosses over to a region of lowered topography and active faulting (Plesch et al., 2007). Values of Δt_P^* peak at 0.22 s within the Salton Trough, and are lower within the Basin and Range province to the east. The mean value of the uncertainty of the data is 0.10 s, which is nearly half the range of estimate for Δt_P^* across the region. We conclude that while the model in Figure 3 is well determined, individual measurements of Δt_P^* contain little information because of the large standard deviation describing the error of the measurements.

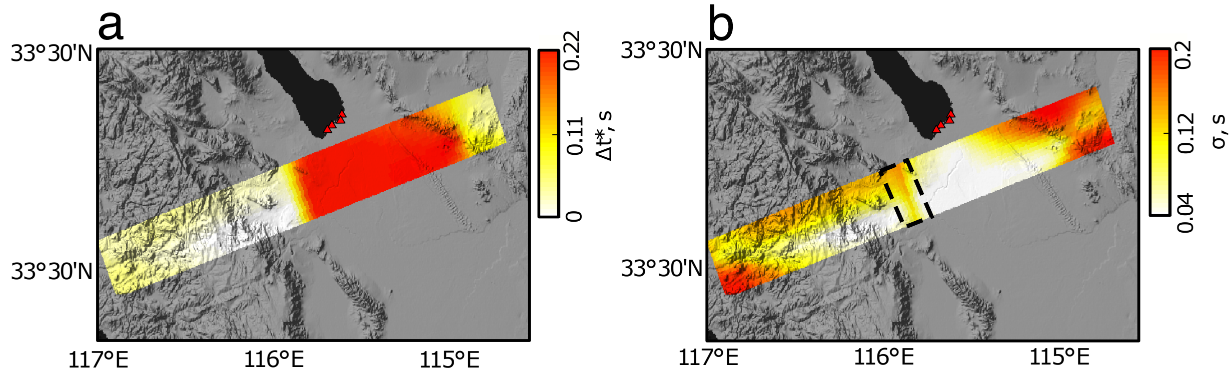

Figure 3. Results of the mapping inversion. The Salton Buttes are approximately marked by red triangles. a) Median Δt_P^* from the ensemble of models. b) Standard deviation of Δt_P^* from the ensemble of models. The location of the maximum in uncertainty referred to in the main text is highlighted by the dashed black box.

The standard deviation of the ensemble is shown in Figure 3b, which provides an estimate of the uncertainty of the solution. This uncertainty is for the models and is distinct from the uncertainty of the individual measurements. Within the entire Salton Trough and the Peninsular Ranges nearest to 116°W the uncertainty is approximately 0.04 s. The lower values of $\boldsymbol{\Delta t_P^*}$ on the eastern side of the Salton Trough and near the coast on the western side of the array are less well constrained with uncertainties as high as 0.08 s. Where there is no station coverage, uncertainties increase to 0.2 s, which is roughly the standard deviation of values drawn from the uniform distribution used to initialize the chains.

The change in $\boldsymbol{\Delta t_P^*}$ at the rapid onset of high attenuation near 116°W spans the full range of the model. The width of this transition is constrained by the local maximum in uncertainty near 116°W (Fig. 3b), where the uncertainty reaches 0.12 s. This maximum reflects the uncertainty in the location of the boundary between the regions of high and low attenuation. The width of this peak in uncertainty is only ~20 km, and hence we infer a rapid transition into the region of high attenuation. However, this analysis does not constrain the depth at which the attenuation occurs. Since the Salton Trough features thick sediments that typically have low quality factors (Hauksson and Shearer, 2006) and high-heat flow consistent with a warm crust (Lachenbruch et al., 1985), this analysis cannot determine if the attenuation occurs above or below the mantle lithosphere.

3. Tomographic inversions

In this section, we describe two approaches to teleseismic attenuation tomography, a damped least-squares approach and non-linear Bayesian approach. We show via the inversions of synthetic datasets that damped least-squares is inadequate for the inversion of the dataset in this study. We first define the parameter we are inverting for as

$$\boldsymbol{\zeta} = \frac{\mathbf{1000}}{\boldsymbol{Q_p}} \quad (3)$$

for convenient expression and refer to relative changes in our models in subsequent sections as $\Delta\zeta$. The small station spacing of the SSIP array means that teleseismic ray-paths cross at depths as shallow as 20 km and may therefore be useful for constraining the depth-distribution of attenuation, even if the number of useable teleseismic events is small (Heath et al., 2015). We show the ray paths for the events in our study in Figure 4a along the x-axis shown in Figure 1; all tomographic inversions are in this plane. We use the iasp91 velocity model (Kennet and Engdahl, 1991) and hold ray paths fixed during all inversions. We show the hit-count for a grid with 5 km wide square elements in Figure 4b; hit-count is defined as the number of rays that intersect with each element in the grid. The hit-count is low beneath the eastern-most portion of the array, particularly at depths below 50 km, and we expect this portion of the tomographic models to be the least constrained. By integrating along the ray-path in equation (2), we are using ray-theory and neglecting finite-frequency effects, as kernels for t^* have not been derived. Since we are working exclusively with events having energy above 1 Hz, ray-theory should provide a reasonable approximation.

A key difference between velocity and attenuation tomography is the size of perturbations. The velocity tomography study of *Schmandt and Lin* (2014) report a model from damped least-squares of velocity anomalies across the Continental United States that features perturbations to the starting model of less than ±5% at depths of 75 km. While non-linear methods recover perturbations with larger amplitudes (Burdick and Lekić, 2017), the linear approach is functional. The perturbations in attenuation tomography, in contrast, may be of significantly greater amplitude, and perturbations within a tomographic model may span over an order of magnitude. For example, $\boldsymbol{Q_P}$ values of 1350, 180, and 50 (typical values for the lithosphere (Dziewonski and Anderson, 1981), asthenosphere (Dziewonski and Anderson, 1981), and melting-producing regions (Abers et al., 2014)) have ζ of 0.7, 5.5, and 20. Moreover, assuming a reference ζ typical of the asthenosphere, 5.5 (Dziewonski and Anderson, 1981), means the largest negative perturbation possible is -5.5, while perturbations will be unbounded in the positive direction. A second difference is the low-quality of the observations. We estimated in Section 2.3 that the uncertainty for an observation of $\boldsymbol{\Delta t_P^*}$ was approximately half the range of the smoothed map of $\boldsymbol{\Delta t_P^*}$, and so tomographic inversions of $\boldsymbol{\Delta t_P^*}$ are likely to be more uncertain than analogous inversions of teleseismic travel times, which are themselves associated with significant uncertainty (Burdick and Lekić, 2017).

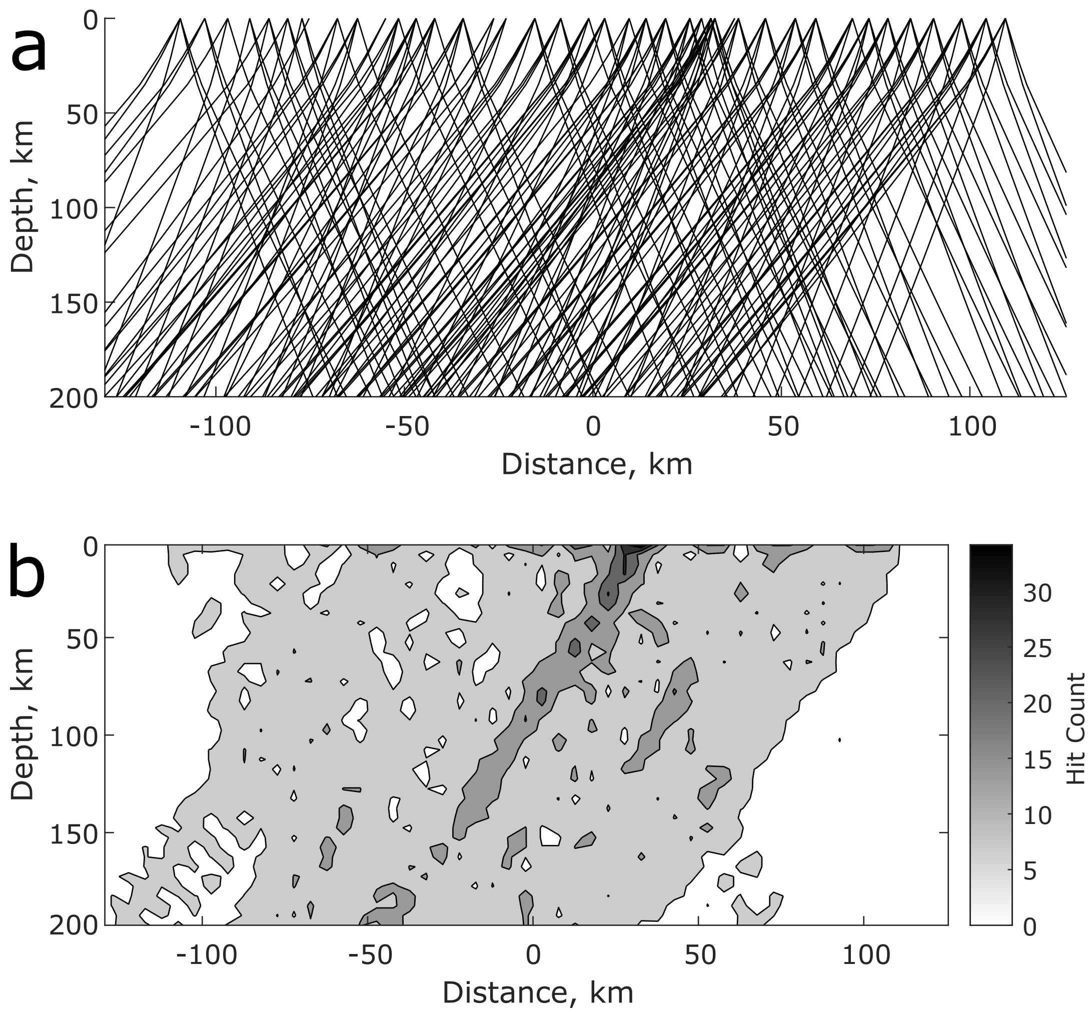

Figure 4. Ray path coverage. A) Ray-paths for all station event pairs available in this study plotted along the plane defined by the axis defined in Figure 1b. B) Ray-path hit-count for 5 km wide square regions.

We first test if a damped least-squares approach can accurately reconstruct synthetic models of ζ containing both small, symmetrically distributed anomalies – analogous to the case of velocity tomography – and large, asymmetrically distributed anomalies. We test inversions with smooth models and models that are discontinuous at the depths of the LAB as reported by *Lekić et al.,* (2011). Both surface-wave tomography and active-source experiments confirm that seismic velocity is high within the mantle lithosphere beneath the Salton Trough relative to typical values of the asthenosphere (Barak et al., 2015; Han et al., 2016), and *Qp* will be higher at lithospheric than asthenospheric conditions since anelastic mechanisms are weak or even inactive at temperatures typical of the lithosphere (Faul and Jackson, 2005; Jackson and Faul, 2010). Hence, we consider anomalies that are continuous across the LAB of *Lekić et al.,* (2011) to be improbable.

3.1 Framework for the damped least-squares approach

To invert for ζ anomalies with damped least-squares, we seek a vector of model parameters $\vec{\boldsymbol{m}}$ that minimizes an objective function *E,* defined as

$$\boldsymbol{E} = \left\|\boldsymbol{G}\cdot\vec{\boldsymbol{m}} - \vec{\boldsymbol{d}}\right\| + \boldsymbol{\varepsilon}\|\vec{\boldsymbol{m}}\| + \boldsymbol{\lambda}\|\boldsymbol{L}\vec{\boldsymbol{m}}\| \qquad (4)$$

where $\boldsymbol{G}$ is a matrix of partial derivatives that relates model parameters to the data, $\vec{\boldsymbol{m}}$ is a vector of *m* model parameters, $\vec{\boldsymbol{d}}$ is a vector of the *n* $\boldsymbol{\Delta t_p^*}$ observations, *L* is a roughness operator, and $\boldsymbol{\varepsilon}$ and $\boldsymbol{\lambda}$ are regularization parameters. Model parameters consist of ζ values on a grid with spacing *dg* defining the model space and event terms, $\vec{\boldsymbol{e}}$. The event terms accounts for possible variations in the mean value of $\boldsymbol{t_p^*}$ between different events due to variations in station coverage. As an example, if $\boldsymbol{\Delta t_p^*}$ for one event were measured only with stations within the central Salton Trough, all the $\boldsymbol{\Delta t_p^*}$ values will be near zero since the mean would be removed. The event statics are designed to correct these measurements back to high values during the inversion using information from other events. The partial derivate of $\boldsymbol{\Delta t_p^*}$ with ζ parameters is *dg*/(1000**Vp*) if a ray crosses an element of grid, and zero otherwise. The partial derivative with an event term is 1 for $\boldsymbol{\Delta t_p^*}$ measurements of a given event and zero otherwise. The hit count then defines the derivate-weight-sum for a model parameter scaled by 1000 following (3). The roughness operator L applied to model parameter i defines a matrix of size m x m with the value at index j defined as

$$\boldsymbol{L_{ij}} = \begin{cases} \boldsymbol{-1 \; if \; i = j} \\ \mathbf{exp}\left(-(\frac{\|\overline{r_{ij}}\|}{r_\lambda})^2\right) \boldsymbol{if \; i \neq j} \end{cases} \quad (5)$$

where $\overline{\boldsymbol{r_{ij}}}$ is the line between the parameters i and j, and $\boldsymbol{r_\lambda}$ is the length over which roughness is considered. The length of the vector defined by the operator is then normalized to 1 along index i. When enforcing roughness across a discontinuity defined by the vector $\overline{\boldsymbol{d}}$, we define a new operator ${}^{d}\boldsymbol{L_{ij}}$ as

$${}^{d}\boldsymbol{L_{ij}} = \begin{cases} \boldsymbol{-L_{ij} \; if \; \overline{r_{ij}} \; intersects \; \overline{d}} \\ \boldsymbol{L_{ij} \; otherwise} \end{cases} \quad (6)$$

where $\boldsymbol{L_{ij}}$ is the normalized operator defined in (4). Then we solve for $\overrightarrow{\boldsymbol{m}}$ with LSQR (Paige and Saunders, 1982) from the constrained linear set of equations

$$\begin{bmatrix} \boldsymbol{G} \\ \boldsymbol{\varepsilon I} \\ \boldsymbol{\lambda L} \end{bmatrix} \overrightarrow{\boldsymbol{m}} = \begin{bmatrix} \overrightarrow{\boldsymbol{d}} \\ \boldsymbol{O} \\ \boldsymbol{O} \end{bmatrix} \quad (7)$$

where I is the identity matrix of size m x m and O is the zero vector of size m x 1. The constraint equations balance model misfit with the magnitude and roughness of the model, with greater values of $\boldsymbol{\varepsilon}$ reducing $||\overrightarrow{\boldsymbol{m}}||$ and greater values of $\boldsymbol{\lambda}$ reducing $||\boldsymbol{L}\overrightarrow{\boldsymbol{m}}||$. When enforcing a discontinuity, ${}^{d}\boldsymbol{L}$ replaces $\boldsymbol{L}$ and the inversion will favor a rapid contrast in parameters across the discontinuity.

3.2 Framework for the Bayesian approach

We next present a Bayesian approach that does not impose fixed, *a priori* regularization on the models (Green 1995, Malinverno 2002, Sambridge *et al.* 2006). Similar approaches have been widely applied to geophysical datasets in the past decade (e.g., Bodin and Sambridge, 2009;

Agostinetti and Malinverno, 2010; Dettmer et al., 2010; Bodin et al., 2012a, 2012b; Kolb and Lekić, 2014; Burdick and Lekić, 2017; Olugboji et al., 2017; Eilon et al., 2018; and many others). The final product is a probability density function (PDF) that describes a set of independent models at each point within the imaging domain. The uncertainty of the measurements, σ_{t^*} , is also treated as a "hierarchical" parameter that can be solved for (Malinverno and Briggs, 2004) as in Section 2.2. The formalism described here ensures that the PDFs are consistent with Bayes theorem and obviates the need to impose fixed assumptions regarding the complexity of the models or the quality of the data.

The set of models that form the PDF are generated by iterating a Markov chain. Individuals models are Voronoi diagrams constructed by near-neighbor interpolation from a set of nodes. When enforcing a discontinuity, Voronoi cells are not allowed to extend across the discontinuity and models must have at least one node both above and below the discontinuity; the algorithm is otherwise unchanged from what is described below. A chain is initialized by generating a random starting model, $\overrightarrow{m}$, defined by selecting a random number of nodes, k, which each have a ζ value randomly selected from a prior distribution. The ζ model is interpolated with nearest-neighbor interpolation to the ray paths in increments of 5 km in the vertical direction before line integration to find Δt_m^* via (2). We demean Δt_m^* for each event, and find the fit of the starting model to the data with an L2 norm defined by

$$\chi^2(\overrightarrow{m}) = \sum^n (\Delta t_m^* - (\Delta t_p^* + e))^2/\sigma_{t^*}^2 \quad (8)$$

where Δt_m^* is the Δt_p^* predicted for the starting model and e is the relevant event term for Δt_p^*. Event terms account for differences in the set of stations that record each event and minimize the RMS of $(\Delta t_m^* - \Delta t_p^*)$. We find $\vec{e}$ with LSQR via the equations

$$\begin{bmatrix} \boldsymbol{G} \\ \boldsymbol{I} \end{bmatrix} \vec{\boldsymbol{e}} = \begin{bmatrix} (\boldsymbol{\Delta t_m^*} - \boldsymbol{\Delta t_p^*}) \\ \boldsymbol{0} \end{bmatrix} \quad (9)$$

where the matrix *G* is defined only for event terms as an Sec. 3.1, and *I* is the identity matrix of the length of the vector $\vec{e}$. The additional constraint equation enforces a zero mean on the event terms.

The probability (to a constant term) of the initial model can be found once the fit is known, and this probability will be used to determine if later models are accepted. The acceptance probability is found with Bayes theorem,

$$P(\vec{m}|\vec{d}) \propto P(\vec{m})P(\vec{d}|\vec{m}) \quad (10)$$

where $P(\vec{m}|\vec{d})$ describes the probability of model vector $\vec{m}$ given the data vector $\vec{d}$, $P(\vec{m})$ describes the prior probability of drawing the vector $\vec{m}$, and $P(\vec{d}|\vec{m})$ describes the "likelihood" of drawing the data $\vec{d}$ given the model $\vec{m}$. We assume a uniform distribution for the locations of the nodes and for the uncertainty of the data. We use a log-uniform distribution for *k*, which is the relevant Jeffery's prior (as noted by *Kolb and Lekić,* (2014)), given by

$$P(k) = (k(\ln\left(\frac{k_{max}}{k_{min}}\right)))^{-1} \quad (11)$$

where k_{max} and k_{min} are the maximum and minimum number of nodes allowed, respectively.

For ζ, we assume a Gaussian distribution for the prior defined by

$$P(\vec{m}) = \frac{1}{\sqrt{2\pi}\sigma_\zeta} \exp\left(-\frac{\zeta^2}{2\sigma_\zeta^2}\right) \quad (12)$$

where σ_ζ is a standard deviation for ζ values. This prior better describes a model of anomalies than the uniform prior, since a model of deviations from the reference model should be dominated by near-zero values. Moreover, since our observations do not constrain the mean value of t_P^*, we require a prior that centers the models at zero. Assuming a uniform prior allows for models to be populated by large anomalies that vary little across the model so as to only effect the mean of t_P^* and therefore be invisible to Δt_P^* constraints.

Once the likelihood is known for the initial model, the chain is advanced by perturbing the initial model to produce a new model, $\overrightarrow{m'}$. One of five possible types of perturbations can occur on each step of the chain: A) a “birth” step introduces a new node, increasing the complexity of the model, B) a “death” step removes a node, reducing the complexity of the model, C) a “move” step changes the location of a node, D), a “change” step changes the ζ of a node, and finally E) an “error” step changes σ_{t^*} . The probability that the perturbed model is accepted, α, is found with the transdimensional form of the Metropolis-Hastings algorithm as given by *Green* (1995),

$$\alpha(m'|m) = \min\left[1, \frac{p(\overrightarrow{m\prime})}{p(\vec{m})}\frac{p(\vec{d}|\overrightarrow{m\prime})}{p(\vec{d}|\vec{m})}\frac{q(\vec{m}|\overrightarrow{m\prime})}{q(\overrightarrow{m\prime}|\vec{m})}|J|\right] \quad (13)$$

where $q\left(\overrightarrow{m'}|\vec{m}\right)$ is the probability of transitioning from model $\vec{m}$ to $\overrightarrow{m'}$, $q\left(\vec{m}|\overrightarrow{m'}\right)$ is the probability of transitioning from model $\overrightarrow{m'}$ to $\vec{m}$ (that is, the probability of the reverse of the perturbation made on the current step), and the $|J|$ is the determinant of the Jacobian matrix (see Green (1995) for a discussion; the term is always 1 for the problem considered here). The key to implementing this inversion is to evaluate (13) for each type of perturbation. We follow the derivations in *Bodin and Sambridge* (2009) and *Bodin et al.*, (2012b) for each step while assuming our priors described above, and arrive at the acceptance functions

$$\alpha_{move}(\overrightarrow{m'}|\vec{m}) = \min\left[1, e^{\frac{\chi^2(\vec{m}) - \chi^2\left(\overrightarrow{m'}\right)}{2}}\right] \quad (14)$$

$$\alpha_{change}(\overrightarrow{m'}|\vec{m}) = \min\left[1, e^{\left(\frac{\zeta o^2}{2\sigma_\zeta^2} - \frac{\zeta'^2}{2\sigma_\zeta^2} + \frac{\chi^2(\vec{m}) - \chi^2\left(\overrightarrow{m'}\right)}{2}\right)}\right] \quad (15)$$

$$\alpha_{birth}(\overrightarrow{m'}|\overrightarrow{m}) = \min\left[1, \frac{k}{k+1}\frac{\delta_\zeta}{\sigma_\zeta}e^{\left(\left(-\frac{\zeta'^2}{2\sigma_\zeta^2}\right)+\frac{(\zeta'-\zeta_o)^2}{2\delta_\zeta^2}+\frac{\chi^2(\overrightarrow{m})-\chi^2(\overrightarrow{m'})}{2}\right)}\right] \quad (16)$$

$$\alpha_{death}(\overrightarrow{m'}|\overrightarrow{m}) = \min\left[1, \frac{k}{k-1}\frac{\sigma_\zeta}{\delta_\zeta}e^{\left(\left(\frac{\zeta_o^2}{2\sigma_\zeta^2}\right)-\frac{(\zeta'-\zeta_o)^2}{2\delta_\zeta^2}+\frac{\chi^2(\overrightarrow{m})-\chi^2(\overrightarrow{m'})}{2}\right)}\right] \quad (17)$$

$$\alpha_{error}(\overrightarrow{m'}|\overrightarrow{m}) = \min\left[1, \left(\frac{\sigma'_{t*}}{\sigma_{t*o}}\right)^n e^{\left(\frac{\chi^2(\overrightarrow{m})-\chi'^2(\overrightarrow{m'})}{2}\right)}\right] \quad (18)$$

where the apostrophe indicates a value from the proposed model, and a subscript *o* indicates a value from the current model, and δ_ζ is the standard deviation of perturbations to the ζ parameters on "change" and "birth" steps. A step-size is also defined for the location of the nodes, δ_r, and for the error terms, δ_σ, but neither step-sizes appear in an acceptance function because we have assumed uniform priors for the coordinates of the nodes and σ_{t*}. A subjective sparsity constraint on the model is not imposed, but the acceptance functions for "birth" and "death" naturally balance the need to fit the data with a bias towards lower dimensional models when $\frac{\delta_\zeta}{\sigma_\zeta} < 1$, that is, the step-size is smaller than the width of the prior. The leading terms $\frac{k}{k+1}$ and $\frac{k}{k-1}$ appear in (16) and (17) because of the assumption of a log-uniform prior for *k*. We validate (14)-(18) by performing a test where χ^2 is set to 1 for all models (Bodin and Sambridge, 2009). The posterior will be the prior in this case, and (14)-(18) correctly return a Gaussian distribution with a standard deviation σ_ζ for ζ at all points in model space, a log-uniform distribution for *k*, and a uniform distribution for σ_{t*}.

Once α is found, the proposed model is accepted if $r < \alpha$, where *r* is a randomly selected value between 0 and 1. If a model is accepted, then the chain continues with the newly accepted model as the current model. If a model is rejected, the proposed model is discarded and a new perturbation to the current model is drawn on the next iteration. Chains continue to a maximum number of iterations, which is a parameter chosen ahead of time. Models are saved at a regular interval of iterations once a "burn-in" number of iterations has been reached. Once all the chains have been iterated a maximum value, the PDF at a given location given by a normalized histogram of ζ at each point in the model domain. The mean value of ζ from the interpolated models is removed before constructing the PDFs. The values of parameters that shape the search are given in Table 1, unless otherwise given in the text for specific inversions. However, we stress that these parameters only define the contours of the search and not the shape of the final PDF so long as a sufficient number of models have been explored (MacKay, 2003).

4. Tests of the inverse approaches

4.1 Synthetic datasets

In this section, we invert three datasets for Δt_p^* predicted with the ray geometry in Figure 4. The models are shown in Figure 5. The first model (Figure 5a with a symmetric color scale), contains two anomalies from 60 to 90 km depth with $\Delta\zeta$ of -2 and +2. The anomalies are separated by a 20 km wide gap. We note that relative to PREM, a $\Delta\zeta$ of 2 is a 36% percent perturbation, and so is large relative to the anomalies typically imaged in teleseismic body-wave tomography. However, the Δt_p^* for this model have a range of ±0.01 s, which is more than an order of magnitude smaller than the range of Δt_p^* in the map-view inversion for the Salton Trough (Figure 3). Before inverting the dataset, we add white Gaussian noise with a standard deviation of 0.003 s (30% of the range of the predicted Δt_p^* values) which obscures the signal when the dataset is plotted against distance (Figure 5b).

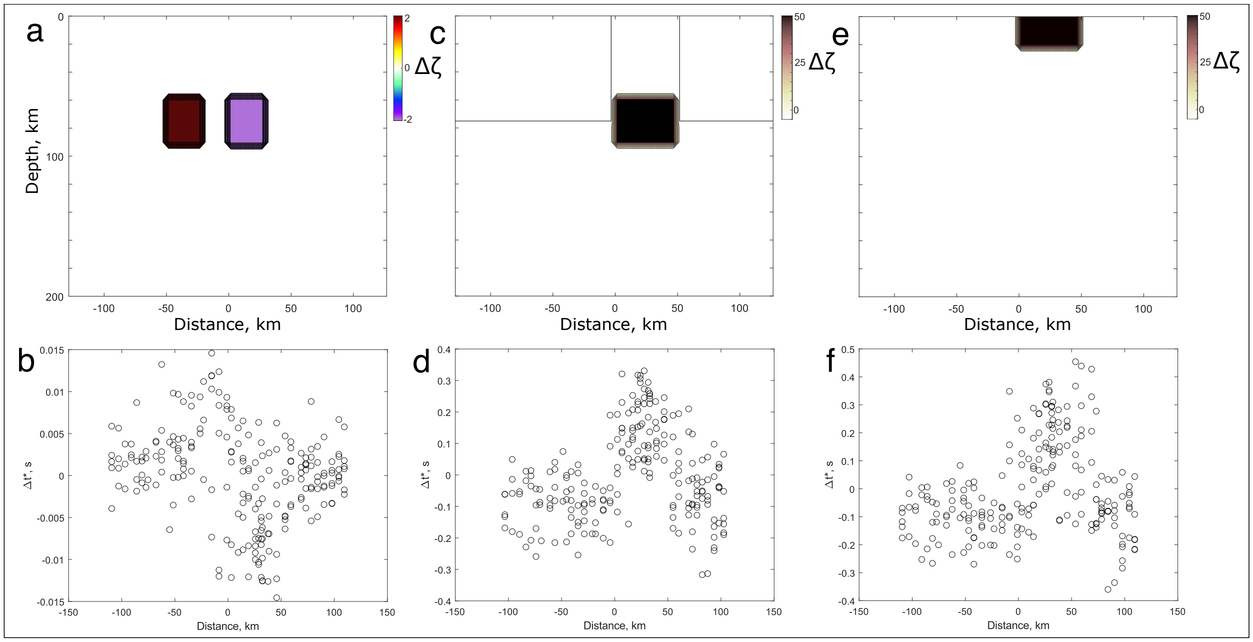

Figure 5. Synthetic models used for testing the inversion frameworks. Top row: Synthetic models discussed in the text. Bottom row: Synthetic data predicted with the event-station distribution in this study for the model in the above panel with Gaussian white noise added.

We also invert models containing asymmetrically distributed large anomalies. The first model (Figure 5c with a unipolar color scale) contains a 30 km thick and 50 km wide box with a $\Delta\zeta$ of +50 at 60 km depth. To test the effect of lithospheric structure, two regions of $\Delta\zeta$ with -5 bound the anomaly. The second model (Figure 5e with a unipolar color scale) contains a 20 km thick anomaly with a $\Delta\zeta$ of +57 at the surface. The amplitude and dimensions of the anomalies were chosen to give a peak Δt_p^* of approximately 0.2 s (as in Figure 3a). These two models test whether our inversion can resolve the difference between the hypotheses that attenuation occurs in the upper crust or in the asthenosphere. White Gaussian noise with a standard deviation of 0.1 s was added to the data to mimic the quality of the real dataset (see Section 2). We note that the standard deviation of the noise is half the range of the clean Δt_P^* values, and thus simulates remarkably poor imaging conditions (Figure 5d,f).

4.2 Damped least-squares inversions of synthetic datasets

We first invert the "symmetric" dataset (Figure 5b) with the damped least-squares approach. The damped least-squares inversion requires two regularization parameters, ε and λ, which damp the magnitude and roughness of the model, respectively. To choose these values, we construct a misfit surface of $ln(\overline{\chi^2})$ (the natural log of χ^2 scaled to the length of the data vector to given a mean misfit) as a function of $ln(\varepsilon)$ and $ln(\lambda)$ and take the point of maximum (Gaussian) curvature as the "ideal" model (Figure 6a). This is analogous to the traditional L-curve approach when considering a single model parameter (Hansen, 1992). The preferred model features an accurate representation of the input model but with vertical exaggeration of the anomalies (Figure 6b). The peak-to-peak amplitudes of $\Delta\zeta$ are ±0.3, which is 15% of the range in the input model (Figure 5a). The lateral extent of the anomalies is approximately recovered, but the vertical extent of both anomalies is greatly exaggerated.

In Figure 6c and d, we show inversions of the asymmetric models with deep (Figure 5c) and shallow (Figure 5e) anomalies, respectively. The regularization parameters were chosen to give peak curvature to the $ln(\overline{\chi^2})$ surface as was done for the previous test. The recovered anomalies span the entire vertical dimension of the model space. Peak-to-peak $\Delta\zeta$ in the recovered models are 10 and 8, which are 16% and 17% recoveries, for the deep and shallow anomalies, respectively. While there is a difference in the depth of the peak anomaly between the two tests, one could not confidently infer whether the input anomaly was at mantle or crustal depths from these inversions alone. An attempt to discriminate between a mantle or crustal origin of the anomaly is made by applying the discontinuity of *Lekić et al.,* (2011) according to the discontinuous roughness constraint in (6) (Figures 6e, 6f). Enforcing the discontinuity does not reduce the amplitude of the artifacts near the surface in Figure 6e and increases the peak amplitude of the deeper artifact in Figure 6f. We conclude that the enforcement of the discontinuity does not aid in discerning the original depth of the anomaly and more generally that the damped least-squares approach is inadequate for imaging large anomalies in attenuation with the dataset available in this study.

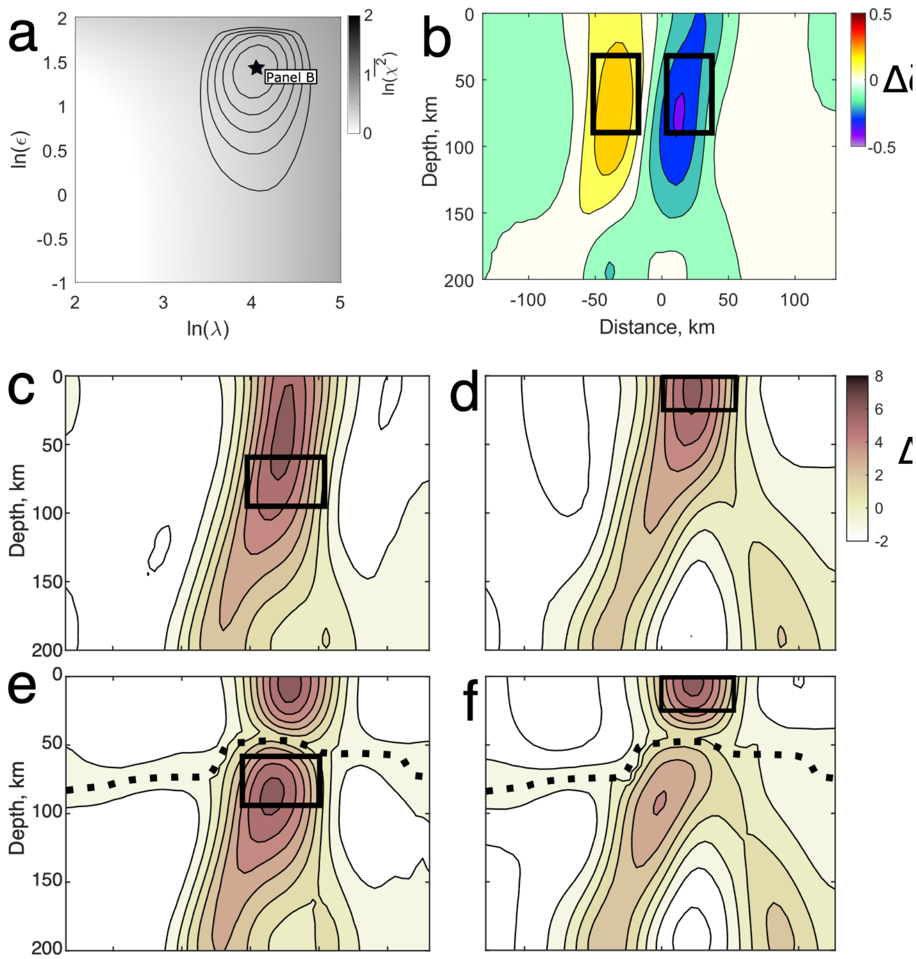

Figure 6. Inversions of the synthetic datasets with damped least-squares. Black rectangles indicate the locations of input anomalies (cf. Figure 5). a) Gray shows the natural log of the scaled misfit against the natural log of the regularization parameters. Contours show the Gaussian curvature of the surface with the star marking the maximum curvature. b) Inversion of the synthetic dataset shown in Figure 5b. Cool and warm colors contours negative and positive values of $\Delta\zeta$. Note the restricted color scale relative to Figure 5a. d, d) Inversion of the synthetic dataset show in Figure 5d and 5f, respectively. Darker colors contour more positive values of $\Delta\zeta$. e,f) Identical with panels c and d, respectively, but with a discontinuity, marked by the dashed black line, imposed on the inversion.

4.3 Bayesian inversions of synthetic datasets

We first test the Bayesian approach on the dataset for the symmetric anomalies. Figure 7a shows the median values of $\Delta\zeta$ across model space using a prior with a σ_ζ of 3. The resulting model is a better match to the input model than the model recovered by damped least-squares.

Peak $\Delta\zeta$ are ±1.7, which is 85% of the amplitude of the anomalies in Figure 5a. Vertical smearing is restricted to approximately 10 km away from the input anomalies, showing that the vertical smearing in Figure 6 was due to the inverse procedure and not to the limitations of the

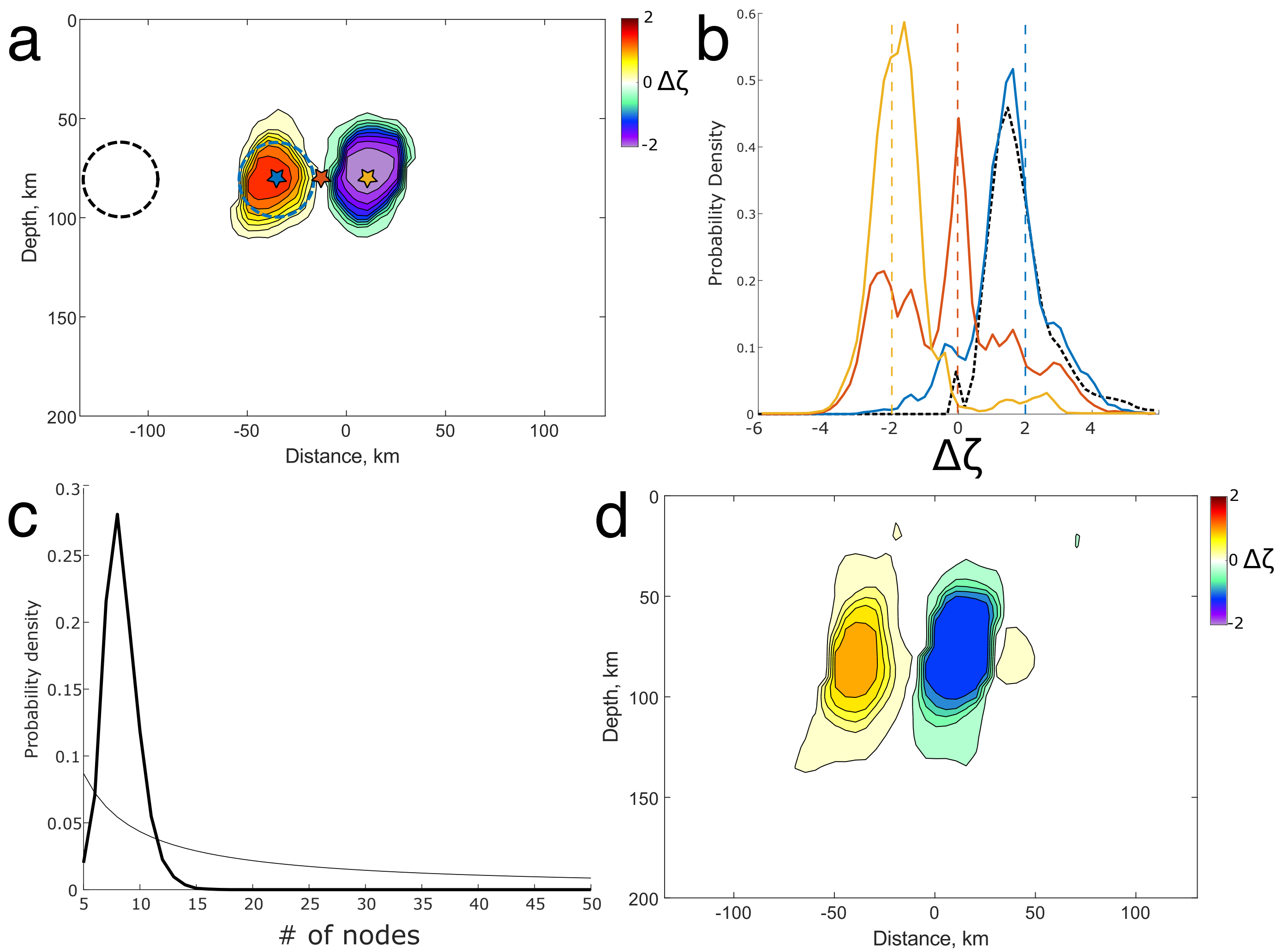

Figure 7. Inversions of a synthetic dataset with the Bayesian approach. a) Median of the ensemble of solutions from the inversion of the synthetic dataset in Figure 5b. Cool and warm colors contour negative and positive values of $\Delta\zeta$. Note that the color scale is the same as in Figure 5a and compare with Figure 6b. The dashed circles mark regions used for a relative measurement in panel b, and colored stars show the locations corresponding to the PDFs in panel b. b) Solid lines show PDFs for $\Delta\zeta$ at points in the model marked by stars of matching color in panel a. Vertical dashed lines show the value in the input model for the locations marked in panel a by stars. Black dashed lines show the PDFs of values within the blue circle minus the values in the black circle in panel a. c) Thin and thick black lines show the prior and posterior probability distributions for the number of nodes describing the individual models. d) Same as panel a but with the width of the prior restricted (see text for discussion).

dataset. No spurious anomalies are produced near the edges of model space. The error on the input data is estimated to be 2.8e-3 s, with a standard deviation below 1e-4 s. This result implies that the inversion slightly over-fits the data, and that all models in the ensemble have similar fits.

While the median of the posterior distribution closely matches the input model, the uncertainty associated with this result is large relative to the amplitude of the anomalies. We show PDFs at three points in Figure 7b. At points in the center of the anomalies, the peaks of the PDFs are well defined and close to the true values, but the tails of the distribution are broad and asymmetric. Outlying values deviate greatly from the input anomalies. At the center of the negative anomaly, a set of low-probability solutions have amplitudes larger than the input anomaly but with the opposite sign, despite the apparent accuracy of the median solution. At the gap between the input anomalies (red star in Figure 7a), the modal value is accurate but there are heavy tails that reflect solutions without any gap. The most common number of nodes defining the individual models is 8 (Figure 7c), which confirms that the inversion finds appropriately simple models even when models with up to 50 nodes are allowed.

The wide PDFs in Figure 7b are strictly for point-estimates of $\Delta\zeta$, and so the width represents uncertainty in both the shape, location, and amplitude of anomalies. To focus on the uncertainty of the amplitude and neglect the fine-scale shape of the anomaly, we show the PDFs for the difference between the maximum value within 20 km of the center of the positive anomaly (blue, dashed circle in Figure 7a) and the minimum value within a reference location (black, dashed circle in Figure 7a). This allows for an estimate of $\Delta\zeta$ that is less sensitive to uncertainties in the shape and location of the anomaly, and the negative tails of the anomaly no longer appear (dashed PDF in Figure 7b). The lower bound amplitude for this anomaly at 95% confidence is 0.45, which is higher than the peak amplitude recovered by the damped least-squares inversion. A small peak at a $\Delta\zeta$ of 0, which mean that no anomaly is recovered where the input model features a positive anomaly, contains less than 1% of the cumulative probability density.

Finally, we test the effect of the width of the prior on the results by repeating the inversions with $\sigma_\zeta = 0.5$, which is a factor of 4 smaller than the peak-to-peak values in the input model. The resulting model (Figure 7d) is more similar to the result from damped least-squares than the result in Figure 7a, which assumed $\sigma_\zeta = 3$. We therefore attribute the smearing from damped least-squares to the regularization and not to the limitations of the dataset. The peak-to-peak values of $\Delta\zeta$ are ±1.1, which is a 55% recovery. Small side-lobes have also appeared next to the recovered anomalies with opposite sign that were not present in the input model. However, despite this under-recovery and appearance of minor artifacts, the mean and standard deviation of σ_{t*} are identical to the inversion that assumed $\sigma_\zeta = 3$. As a note of caution, this means that models fit the data as well as when the prior was more appropriately chosen. The only indication that the prior was inappropriately chosen is that the peak anomalies are near $\pm 2\sigma_\zeta$, while the inversion assuming $\sigma_\zeta = 3$ features peak anomalies near $\pm 0.6\sigma_\zeta$.

We next show inversions of the large, asymmetrically distributed anomalies. We show inversions for two datasets with large anomalies in Figure 8. We assume $\sigma_\zeta = 30$. In all four cases, the estimated error on the input data is within two standard deviations of the true value of 0.1 s. While there is error in the peak amplitudes relative to the synthetic models in Figure 5c and 5e, the amplitudes are better estimated than by damped least-squares and there is no intrinsic bias towards underestimation. The $\Delta\zeta$=*-5* anomalies on the flanks of Figure 5c are not recovered, suggesting that lithospheric structure cannot be recovered alongside strong positive anomalies. While a comparison of Figure 8a and 8b suggests that enforcing the discontinuity may result in the worse reconstruction of anomalies near the discontinuity, a comparison of Figure 8c and 8d show that enforcing the discontinuity leads to a clean separation of shallow and deep structure. This separation could not be achieved with damped least-squares (Figure 6e-f). Finally, we note that the relative attenuation across the discontinuity is poorly estimated, which we attribute to the insensitivity of relative teleseismic observations to lateral structure that spans the whole model space.

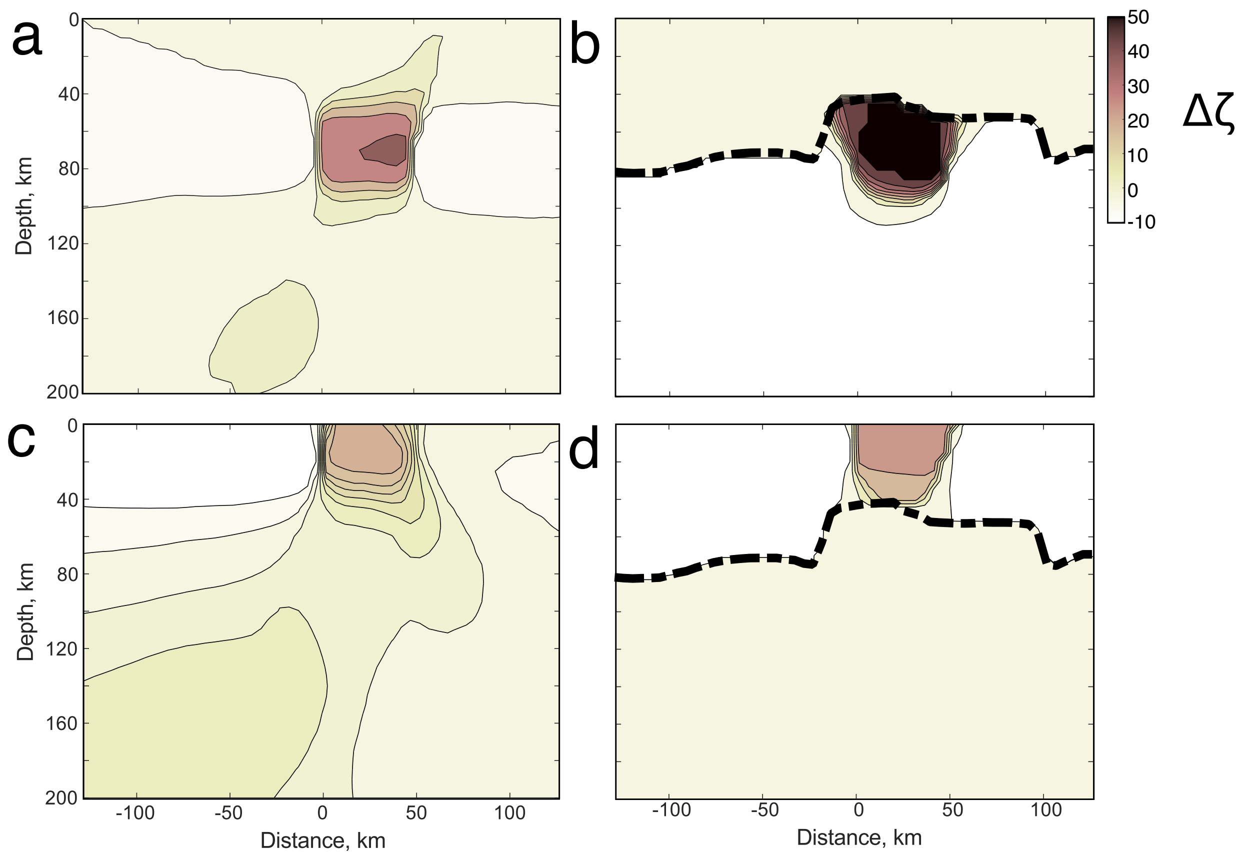

Figure 8. Inversions of synthetic datasets with the Bayesian approach. a,c) Median of the ensemble of solutions from the inversion of the synthetic dataset in Figure 5d and 5f, respectively. Darker colors show more positive values of $\Delta\zeta$. b,d) Same as panels A and C, but with the discontinuity of *Lekić et al.,* (2011), marked by the dashed black line, imposed.

5. Results

5.1 Tomographic inversion of the data

Our preferred approach to inverting the data is the Bayesian method with the discontinuity of *Lekić et al.,* (2011) enforced (Figure 9). In supporting information S1, we also show inversions of the data with the damped least-squares approach with and without the discontinuity enforced. We note that our approach does not enforce the condition that attenuation be weak above the discontinuity, and in synthetic tests the occurrence of strong attenuation above the discontinuity was correctly identified. In supporting information S2, we also show results with the Bayesian approach without the discontinuity enforced. Our choice of preferred approach is based on prior geophysical constraints on the study area and the results of the inversions of synthetic datasets. In Figure 9a, we have plotted a rectangular sub-section of the model space that contains required features (see supporting information S2).

The median solution (Figure 9a) features attenuation in both the crust and upper mantle beneath the Salton Trough, and lateral variations in attenuation are greater in the upper mantle than in the crust (Figure 9b). At depths greater than 80 km, $\Delta\zeta$ reaches a minimum below the Peninsular Ranges in the west and increases eastwards by 11 where the discontinuity begins to shoal. Below the central Salton Trough, a triangular-shaped anomaly with $\Delta\zeta$ of 34 becomes narrower with increasing depth and does not extend below a depth of 150 km. The Salton Buttes lie above the triangular anomaly and are slightly west (negative direction on the x-axis) of the deepest point (Figure 9a).

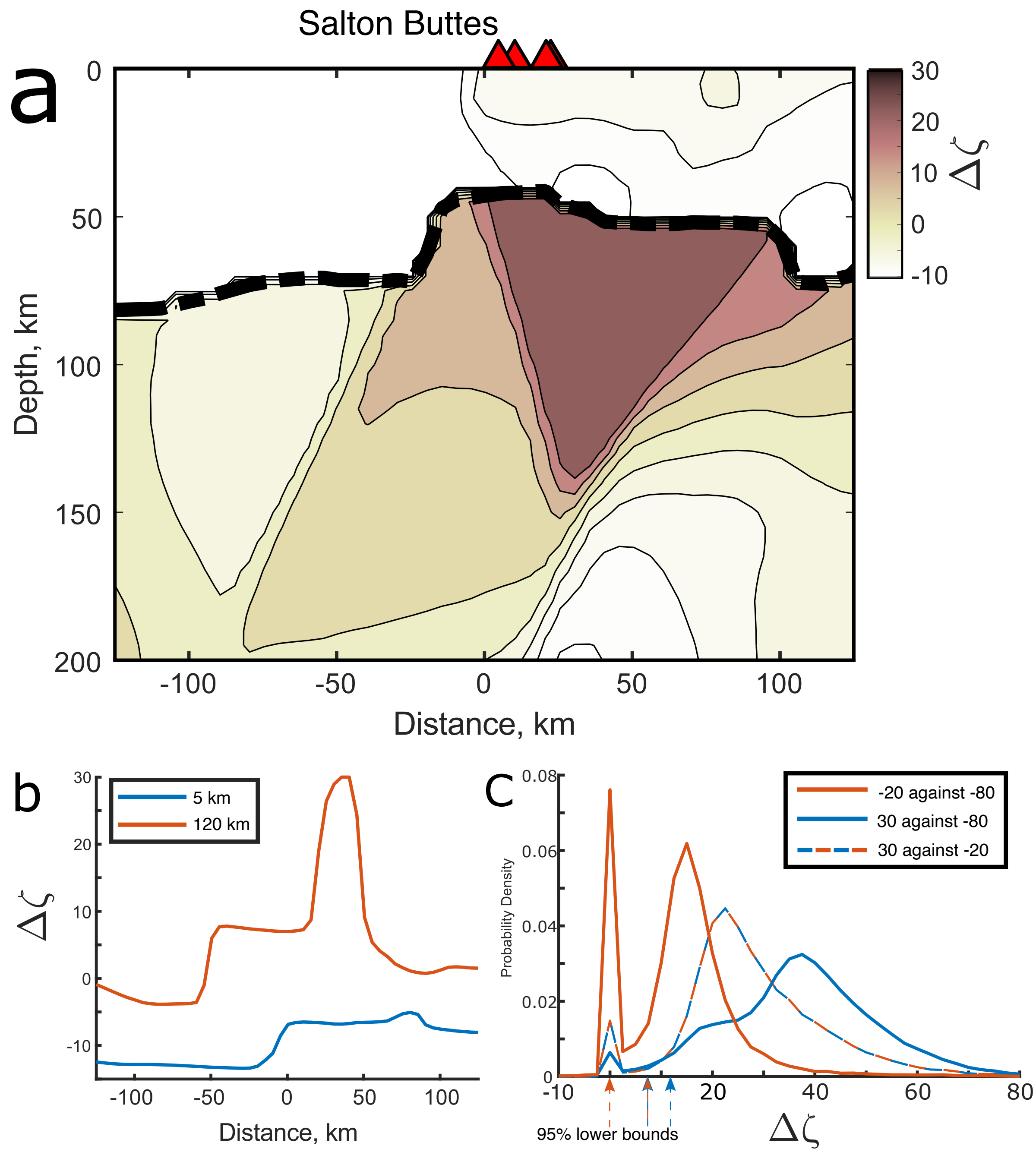

Figure 9. Preferred inversion of the data. A) Median of the ensemble is contoured, with darker colors indicating more positive values of *Δζ*. In order to emphasize features over a wide range of *Δζ*, the contour interval scales with the square of *Δζ*, with 10 contours spaced from 0.5 at a *Δζ* of -10 and 8.3 at a *Δζ* of 30. The discontinuity of *Lekić et al*., (2011) is shown by the dashed black line, and the location Salton Buttes are shown by red triangles on the top. B) Slices thorough the median model in panel A at 5 and 120 km depth. C) Probability density functions for *Δζ* relative to reference regions (cf. Fig 7b, black dashed line). The regions considered have radii of 20 km with a central depth of 120 km, and the legend gives the location of the center of the regions along the x-axis. The 95% confidence lower bound is marked by arrows below the x-axis.

Lower bounds for $\Delta\zeta$ are found between circular regions with radii of 20 km as in Section 3.4 (Figure 9c). The peaks at $\Delta\zeta = 0$ reflect models where the same Voronoi cell covers both the test and reference circles. The lower bound at 95% confidence in the deep portion of the triangular anomaly is 11.5 (blue line in Figure 9c). In contrast, the PDF for the region in between the Peninsular Range and the Trough is bimodal (orange line in Figure 9c). The probability that the minimum in attenuation beneath the Peninsular Range extends further eastward than in the median of the ensemble (that is, $\Delta\zeta = 0$ in Figure 9c) is approximately 20%. Thus, while a strong contrast in attenuation at 120 km depth in the western half of our study area is likely, this contrast is not robust at the 95% confidence interval. We further construct a PDF for the difference between the two test sites, with the intermediate region acting as the reference against the deep triangular anomaly. This difference is robust with a minimum $\Delta\zeta$ of 7 with 95% confidence. There is no evidence for a significant contrast within the triangular anomaly itself, and the eastern edge of the triangular anomaly is not well constrained by the tomographic approach (see supporting text Section 2), likely due to the small number of crossing ray-paths in this portion of model space (Figure 3).

5.2 Comparison with previous geophysical observations

Our preferred model for seismic attenuation correlates well with results from previous geophysical studies. First, the location of the high attenuation beneath the center of the Salton Trough lies beneath the region of low velocities at 50 km depth in the surface-wave tomography model of *Barak et al.* (2015) (Figure 10a). These low velocities lie beneath a high-velocity lid in the upper-most mantle, supporting the interpretation that the attenuation in our preferred model occurs beneath the LAB. The model of *Barak et al.* (2015) does not extend below the discontinuity of *Lekić et al.* (2011) beneath the Peninsular Ranges, and so we cannot determine if the minimum in attenuation beneath the Peninsular Ranges has elevated or typical seismic velocity beneath the discontinuity. At crustal depths, the rifted crust has generally low *Vs* except for a maximum at ~12 km depth. This high velocity anomaly lies beneath the Salton Buttes and above the deepest point of the high-attenuation anomaly in the asthenosphere (Figure 10a).

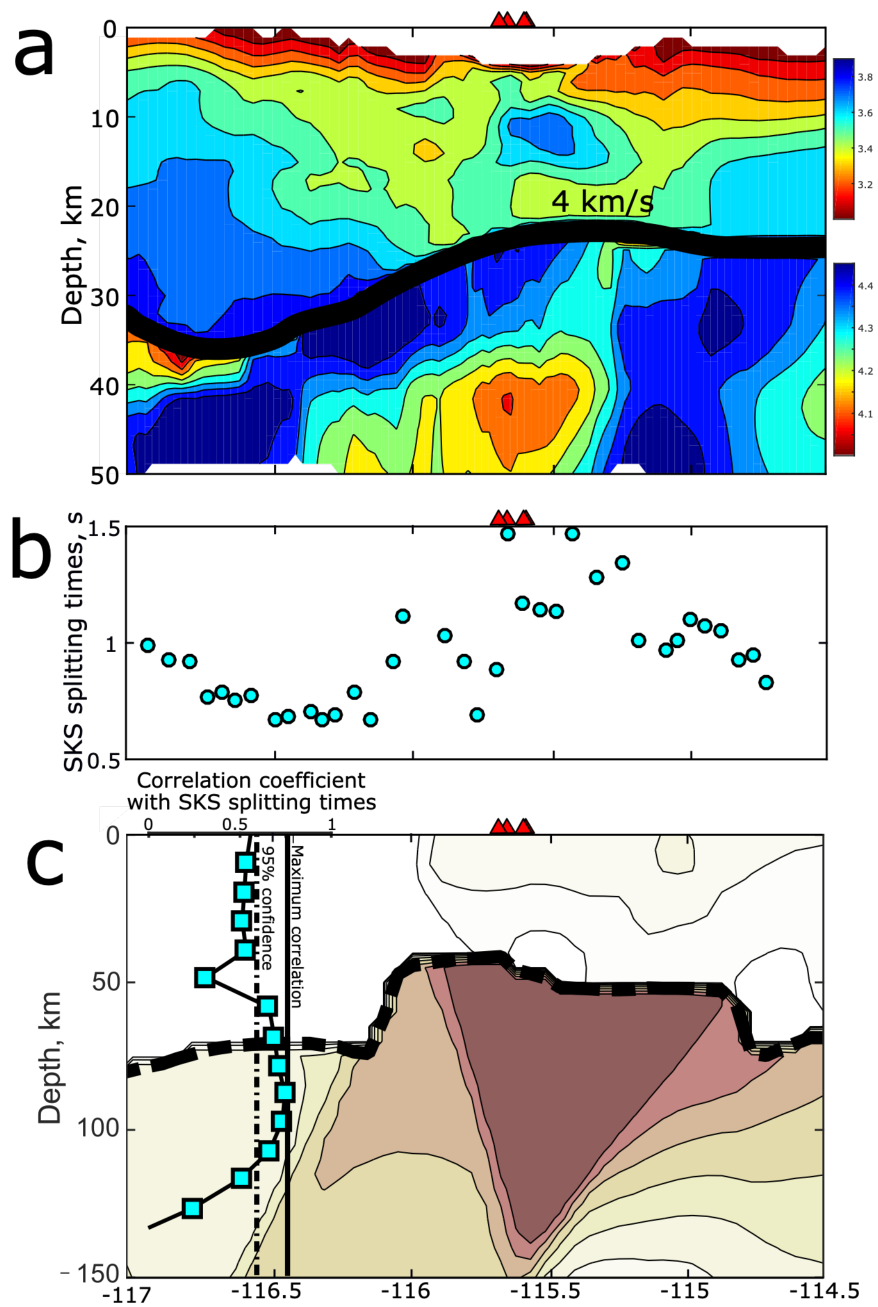

Figure 10. Correlation between seismic attenuation and previous geophysical results. The location of the Salton Buttes is approximately marked by red triangles above each panel. A) Seismic velocity from *Barak et al.* (2015). The moho is approximated by the 4 km/s contour, and the crust and mantle are separately contoured from 3 to 4 km/s in increments of 0.1 km/s and 4 to 4.5 km/s in increments of 0.05 km/s, respectively. B) Splitting times of *SKS* phases from *Barak and Klemperer* (2016) at SSIP stations. C) Model of ζ from Figure 9a is contoured, the dashed black line is the discontinuity of *Lekić et al.*, (2011), and teal squares show the correlation coefficients between the attenuation model at different depths and the *SKS* splitting times in Figure 10b. The solid black line shows the maximum correlation coefficient reached, and the dashed-dotted black line shows the 95% confidence bound.

Our results also correlate with the *SKS* splitting results of *Barak and Klemperer* (2016). The splitting times are smaller in the Peninsular Range and Basin and Range provinces outside of the Salton Trough, and increase by nearly 1 s over distances of less than 100 km (Figure 10b). We show the correlation coefficient between the splitting times and the median solution along different depths on top of the median solution in Figure 10c to test the consistency between these two datasets. The two datasets are positively correlated down to depths of approximately 120 km, with greater correlation at mantle than crustal depths. The correlations at depths between 60 and 110 km are within twice the standard error of the maximum correlation at 90 km, and the correlation at crustal depths lies outside of the 95% confidence interval. The correlation coefficient between Δt_P^* from the mapping inversion (Figure 3a) and the *SKS* splitting times is 0.68 and therefore lies outside of the confidence interval for the maximum at 90 km depth. This estimate of the error does not account for the uncertainty on the tomographic solution, but this analysis show that the structure we recover below the discontinuity is consistent with an independent dataset and improves upon map-view analysis.

6. Discussion

6.1 The physical origin of the high attenuation

Before comparing our results to previous constraints on attenuation, we must first estimate what absolute values of *Qp* are consistent with our relative constraints. For a reference when computing lower bounds, we used the minimum in attenuation beneath the Peninsular Ranges. We cannot reject the possibility of a cold root that extends below the discontinuity, and so *Qp* could be high or even infinite (Dalton et al., 2017; Ma et al., 2020). The scale between our $\Delta\zeta$ constraints and absolute ζ approaches zero as *Qp* approaches infinity. The minimum ζ beneath the Salton Trough would then be 11, which corresponds to a *Qp* of 91. We consider 91 the maximum possible *Qp* (with 95% confidence) in the asthenosphere beneath the Salton Trough. While we cannot strictly reject the possibility of a cold root with nearly infinite quality factor at 120 km below the Peninsular Ranges, a more likely scenario is that this region represents nearly ambient asthenosphere. *Yang and Forsyth* (2008) estimated an average *Qs* of 60 below 100 km depth in Southern California. Assuming a *Qp/Qs* value of 2.25 (Karato and Spetzler, 1990), the value of *Yang and Forsyth* (2008) implies a ζ of 18.5, or a *Qp* of 54, at 120 km depth below the Salton Trough. We note that this is the lower bound with 95% confidence for ζ. Assuming the reference value from *Yang and Forsyth* (2008), there is a 95% chance that *Qp* in the anomaly we image is less than 54. Doing the same exercise but adopting the modal value of $\Delta\zeta$ from the PDF in Figure 9c for the base of the triangular anomaly gives a *Qp* of 22. We consider the ζ within the triangular anomaly to be effectively unbounded in the positive direction.

To place our results in a broader context, we compare the seismic velocities and quality factors from the Salton Trough to a selection of regions with published *Vs* and quality factors and to laboratory predictions in Figure 11. Results from the Juan de Fuca Ridge (Bell et al., 2016; Eilon and Abers, 2017), Lau Basin (Abers et al., 2014; Wei and Wiens, 2018, 2020), and the Mariana Spreading Center (Barklage et al., 2015; Pozgay et al., 2009) have similar values of ζ. Lower ζ values were reported for a range of *Vs* values from Southern California (Yang and Forsyth, 2008), the East Pacific Rise (Yang et al., 2007), and Nicaragua (Harmon et al., 2013; Rychert et al., 2008). Only the isotropic component of velocity is reported in Figure 11. We have assumed a *Qp/Qs* ratio of 2.25 where only *Qs* is available (Karato and Spetzler, 1990), we note that the effects of variable *Qp/Qs* ratios are smaller than the ranges of ζ shown in Figure 11, and we refer the reader to the original studies for detailed considerations of the uncertainties.

Focusing, scattering, and frequency-dependence effects are neglected for the purpose of this comparison (Hwang and Ritsema, 2011; Lekić et al., 2009; Shito et al., 2004).

To explore the physical factors that could explain these observations, we compare these results to experimental predictions for seismic velocity and quality factor. We first consider predictions based on pure olivine that only consider variation in temperature and grain size (Jackson and Faul, 2010). In this case, *Vs* is not predicted to reach values below 4.1 km/s (Jackson and Faul, 2010) and ζ is predicted to be less than 4 (Abers et al., 2014), assuming a *Qp/Qs* ratio of 2.25. Based on the comparison in Figure 11 and these experimental predictions, we reject the hypothesis that elevated temperature or reduced grain sizes alone are sufficient to explain the high attenuation beneath the Salton Trough and other volcanic regions. Next, we consider predictions for seismic velocity and quality factor based on measurements of an organic crystalline analog (Yamauchi and Takei, 2016) parameterized by viscosity. This model predicts that seismic attenuation is enhanced in the asthenosphere by a Debye peak that appears at seismic frequencies when the temperature is near the solidus due to grain-boundary disordering before melting initiates. For values shown in Figure 11, we use a potential temperature of 1350°C, a frequency of 1 Hz, a homologous temperature of 1, and a depth of 75 km over a range of sample viscosities. The predictions from the premelting effect are able to approximately explain the highest values of *Qp* allowed beneath the Salton Trough for reasonable estimates of mantle viscosity of ~10^{18} Pa s (Hirth and Kohlstedt, 2004), and so a premelting effect is not ruled out by our observations as a possible explanation for the attenuation beneath the Salton Trough. In this scenario, no *in-situ* melt fraction is required at mantle depths, though the mantle must be at the solidus beneath the central portion of the rifted region at 120 km depth and below the solidus in the adjacent regions.

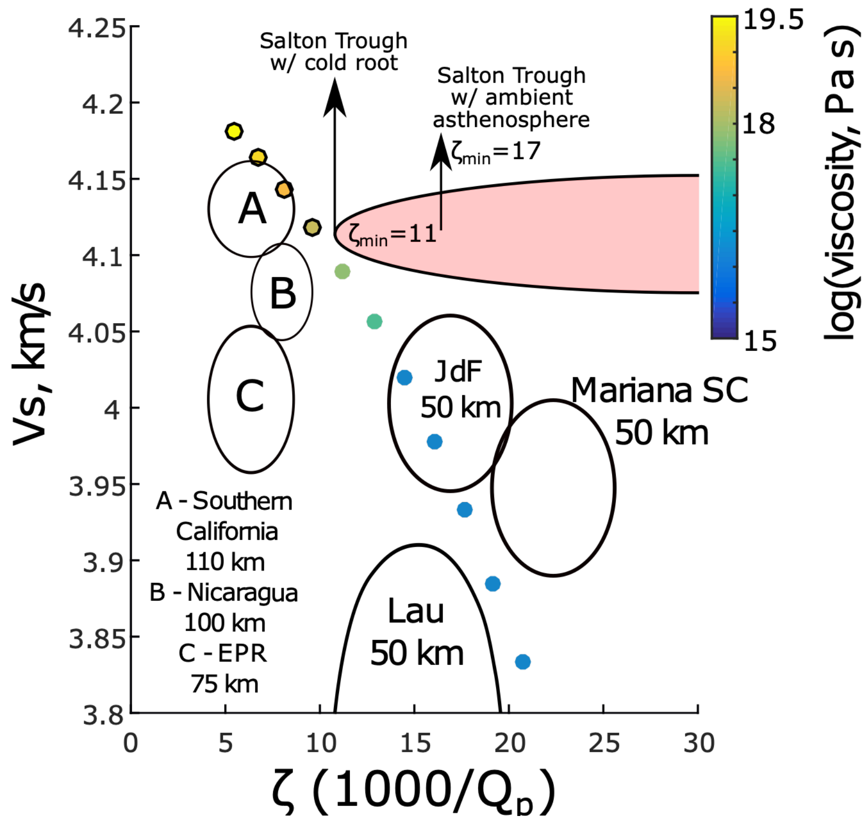

Figure 11. Compilation of *Vs* and ζ ($1000Q_P^{-1}$) values from this and previous studies against an experimental prediction. JdF is the Juan de Fuca Ridge, Mariana SC is the Mariana Spreading Center, Lau is the Lau Basin, and EPR stands for the East Pacific Rise. Circles are predicted *Vs* and ζ values from measurements with an organic crystalline analog (Takei, 2017; Yamauchi and Takei, 2016); circles for viscosities above and below a plausible lower limit of 10^{18} Pa s (Hirth and Kohlstedt, 2004) are outlined and bare, respectively. The size of the ovals approximately shows uncertainties from the respective studies, without any reevaluation here. The red oval uses attenuation measurements from this study.

Since the premelting effect can only explain the lower bounds of our results, we consider two other hypotheses for larger values of attenuation. First, the presence of *in-situ* melt has been proposed to cause attenuation at seismic frequencies by both experimental (Faul et al., 2004) and seismological (Abers et al., 2014; Eilon and Abers, 2017) studies. The presence of melt could cause attenuation by enhancing diffusionally assisted grain boundary sliding (Faul et al., 2004; Eilon and Abers, 2017) or by the displacement of the fluid phase (the "melt squirt" mechanism). Numerical studies have suggested that basaltic melt in the mantle causes attenuation via the melt squirt mechanism only at frequencies above the seismic band (Hammond and Humphreys, 2000), but this result depends on the poorly known geometry of the melt at the grain scale (Garapić et al., 2013). The distribution of attenuation at mantle depths spatially correlates with the splitting times of *SKS* phases, and *Barak and Klemperer* (2016) attributed the large splitting times to the ear-driven segregation of melt in the upper mantle beneath the Salton Trough. The organization of melt into networks connected over longer distances than considered in the model of *Hammond and Humphreys* (2000) could theoretically give rise to strong attenuation at seismic frequencies or to bulk attenuation, which is strong in the Lau Basin (Wei and Wiens, 2020), though we are unaware of a quantitative estimate of this effect.

Another possible mechanism is elastically accommodated grain boundary sliding (EAGBS, Jackson and Faul, 2010), but the strength and central frequency of the peak associated with this mechanism must vary within the study area to influence relative constraints on attenuation. *Olugboji et al.*, (2013) estimated that variations in the water content of the mantle could produce *Qp* values of 67, assuming a *Qp/Qp* ratio of 2.25, and since the strength of the peak is poorly constrained, we consider this mechanism possible. To explain our relative constraints on *Qp*, the central frequency or strength of the peak would need to change within the seismic band (Ma et al., 2020), which could be explained by dehydration during mantle upwelling (Olugboji et al., 2013).

In summary, the presence of melt in the mantle beneath the Salton Trough likely causes the strong attenuation. The organization of the melt into high-porosity bands can potentially explain the most probable values of attenuation, the correlation between *SKS* splitting times with attenuation, and how pockets of melt could lead to attenuation in light of the conclusion of *Hammond and Humphreys* (2000) that non-aligned melt pockets does not cause attenuation. However, both a “premelting” effect due to the onset of melting or the activation of EAGBS due to dehydration can explain the minimum allowed attenuation. Any combination of these three mechanisms is also possible. In contrast, the hypothesis that only temperature and grain size variations can explain the results can be rejected based on the results of both experimental predictions and comparisons with previous observations.

6.2. Evidence for dynamic upwelling

While the mechanism causing the high attenuation cannot be isolated, each mechanism we considered in Section 6.1 considers the attenuative feature at the base of the triangular anomaly as upwelling mantle. A synthetic test of a strong anomaly restricted to depths above 100 km depth did not produce a similar triangular anomaly (Figure 8a, b) and so we do not attribute the maximum depth extent of the anomaly to the limited resolving power of the dataset. Moreover, we confirmed that attenuation is stronger beneath the center of the rifted region than beneath the western flank of the rift. Taken together, we conclude that the attenuative region represents a melting column that is narrower than the rifted region at 120 km depth.

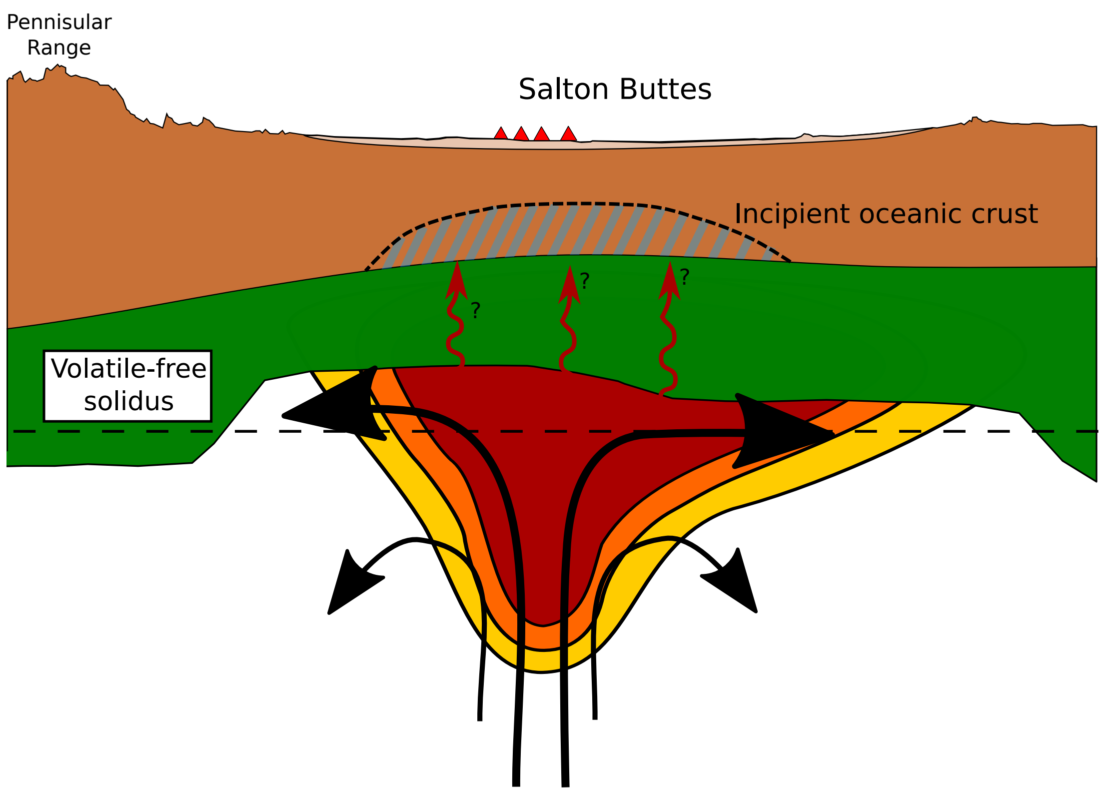

Figure 12. Illustration of mantle and crustal structure in this region. Dark brown shows the crust with topography, light brown shows sediments, green shows the mantle lithosphere, blue striping indicates a gabbroic composition in the crust, warmer colors indicate regions of mantle melting, black lines trace mantle flow, red lines are inferred pathways for melt, and the deep dashed black line is the volatile-free solidus of the ambient asthenosphere. The vertical axis is not to scale.

The shape of the melting column inferred in this study is predicted by models of dynamic upwelling. If the viscosity of the mantle and the spreading-rate of the rift are sufficiently low that buoyancy forces can contribute to upwelling, then upwelling takes the form of a narrow column with downwelling on the flanks (Buck and Su, 1989; Scott and Stevenson, 1989; Su and Buck, 1993). The complementary downwelling sharply demarcates edges of the region of both melt production and melt retention (Katz, 2010), which conforms with our tomographic image. Some geodynamic simulations specifically predict an upwelling column that is narrower than 50 km wide at 120 km depth that widens toward the surface due to the progressive dehydration of the upwelling mantle (Braun et al., 2000; Choblet and Parmentier, 2001), which is a scenario in excellent agreement with our results. In geodynamic simulations of passive upwelling, where only viscous coupling with the spreading lithosphere drives mantle flow, the upwelling region becomes wider with increasing depth (Shen and Forsyth, 1992; Ligi et al., 2008), which is inconsistent with our tomographic results. Consistent with this interpretation, several localized low-velocity anomalies are distributed beneath the Gulf of California that are not centered beneath spreading centers (Figure 1a). *Wang et al.,* (2009) attributed these anomalies to dynamic upwelling, since plate-spreading forces cannot center upwelling away from the ridge-axis. Finally, we note that the deep attenuative anomaly is better constrained on the western than eastern side. However, an eastward extension of the attenuative anomaly would not contradict our interpretation as it would imply a highly asymmetric melting column. Dynamic upwelling causes asymmetric upwelling in the presence of modest compositional or thermal anomalies in the mantle (Katz, 2010). In contrast, asymmetric upwelling is not predicted under passive-upwelling conditions without temperature or pressure anomalies in the mantle that are likely unrealistic for this region (Conder et al., 2002; Toomey et al., 2002).

We illustrate a first-approximation scenario for mantle and crustal dynamics in the study area in Figure 12. A narrow column of upwelling mantle forms near 120 km, where melting can be initiated due to the presence of H_2O and CO_2 (Dasgupta et al., 2007). Holocene volcanic activity at the Salton Buttes (Schmitt et al., 2013; Wright et al., 2015) that exhume young basaltic xenoliths (Schmitt and Vazquez, 2006) occurs approximately at the longitude of deepest attenuation and approximately 35 km north of the main SSIP line (Figure 1b). Moreover, results from seismic refraction support a gabbroic lower crust (Han et al., 2016). In between these signatures of incipient oceanic crust is a high velocity layer in the shallow mantle with seismic velocity characteristic of young oceanic lithosphere and a sharp lower boundary (Lekić et al., 2011; Barak et al., 2015; Han et al., 2016). Taken together, we illustrate the upwelling mantle as bounded at the base of this remaining lithospheric layer, with pathways for basaltic melts to migrate towards the surface inferred. The deep melting column is sharply bounded below the lithosphere by downwelling in response to buoyancy-driven flow.

7. Conclusions

This study constrains seismic attenuation beneath the Salton Trough with broad-band seismic data collected by the Salton Trough Seismic Imaging Project. The relative attenuation of first-arriving *P* phases from deep-focus teleseismic events reveals stronger attenuation within the rifted regions of the study area than within the surrounding regions. Synthetic tests show that a Bayesian approach to the inverse problem reconstructs both the amplitude and depth of anomalies more accurately than a damped least-squares approach in the presence of severe noise. We find that attenuation beneath the study area mainly occurs in the asthenosphere with some crustal contribution. Inversions of synthetic datasets give confidence that these different signals can be isolated. Below the lithosphere-asthenosphere boundary, attenuation is stronger than can be explained by variations in the temperature and grain size of the mantle, and we ultimately attribute the high attenuation to melting driven by upwelling via several possible mechanisms. Of the possible mechanisms, we prefer shear-driven segregation of the melt into high-porosity bands. The highly attenuative region at 120 km depth is narrower than the rifted lithosphere, consistent with geodynamic models of buoyancy-driven upwelling.

ACKNOWLEDGMENTS

Data from the Salton Trough Seismic Imaging Project (SSIP) are available from the IRIS DMC under the network code XD. Support for this project was provided by the University of Minnesota and the NSF Earthscope program under grant EAR-1827277. All $\boldsymbol{\Delta t_P^*}$ values, the ensemble of models, and necessary MATLAB 2016a scripts for plotting are available from the Data Repository for the University of Minnesota (http://hdl.handle.net/11299/166578)[1].

[1] Note to reviewers (see acknowledgements): Data will be uploaded at the time of publication so as to be linked with the manuscript's DOI. Until then, this information will be available from the lead author's github account at https://github.com/jsbyrnes/Salton-Trough.git

Tables

Table 1. Events used with the SSIP array

Region	Latitude, °	Longitude, °	Origin Time	Magnitude	Depth, km	Backazimuth, °
Izu-Bonin	29.057	139.251	2012-10-23 08:53:38.240	5.9	436	302
Russia	49.8	145.064	2012-08-14 02:59:38.460	7.7	583	317
Argentina	-22.059	-63.555	2012-06-02 07:52:53.990	5.9	527	130
Tonga	-18.012	-178.436	2012-02-10 01:47:33.110	6	577	240
Tonga	-21.593	-179.324	2011-09-15 19:31:03.160	7.3	629	237
Argentina	-28.413	-63.136	2011-09-02 13:47:10.700	6.7	592	135
Tonga	-17.606	-178.533	2011-04-03 14:07:08.980	6.4	547	240
Tonga	-26.043	178.476	2011-02-21 10:57:51.760	6.5	551	235

Table 2. Default parameters used during the inversions

Variable (if used)	Meaning	Value
k_{max}	Maximum number of nodes	50
k_{min}	Minimum number of nodes	5
σ_ζ	Standard deviation of the prior for ζ	30
δ_ζ	Step size for ζ	3
δ_r	Step size for the location of a node along both axes	10% of the range
δ_σ,	Step size for σ_{t*}	0.01 s
	Maximum σ_{t*} allowed	1 s
	Maximum number of iterations for an individual chain	5e6
	Burn-in interval	2.5e6
	Interval at which to save a model once the burn-in has been reached	1e4
	Number of independent chains	96
r_λ	Length over which roughness is considered (damped least squares only)	35 km
dg	Grid spacing (damped least squares only)	5 km

References

Abers, G. A., Fischer, K. M., Hirth, G., Wiens, D. A., Plank, T., Holtzman, B. K., McCarthy, M., and Gazel, E. (2014), Reconciling mantle attenuation-temperature relationships from seismology, petrology, and laboratory measurements. *Geochem., Geophys., Geosys.*, *15*, 3521–3542. doi:10.1002/2014GC005444

Agostinetti, N. P., and Malinverno, A. (2010),), Receiver function inversion by trans-dimensional Monte Carlo sampling. *Geophys. J. Int.*, *181*, 858–872, doi:10.1111/j.1365-246X.2010.04530.x

Azimi, S. A., Kalinin, A. V., Kalinin, V. V., and Pivovarov, B. L. (1968), Impulse and transient characteristics of media with linear and quadratic absorption laws. *Izv. Phys. Solid Earth*, *2*(2), 88–93.

Barak, S., and Klemperer, S. L. (2016), Rapid variation in upper-mantle rheology across the San Andreas fault system and Salton Trough, southernmost California, USA. *Geology*, *44*, 575–578, doi:10.1130/G37847.1

Barak, S., Klemperer, S. L., and Lawrence, J. F. (2015), San Andreas Fault dip, Peninsular Ranges mafic lower crust and partial melt in the Salton Trough, Southern California, from ambient-noise tomography. *Geochem., Geophys., Geosys.*, *16*(11), 3946–3972 doi:10.1002/2015GC005970

Barklage, M., Wiens, D. A., Conder, J. A., Pozgay, S., Shiobara, H., and Sugioka, H. (2015), P and S velocity tomography of the Mariana subduction system from a combined land-sea seismic deployment. *Geochem. Geophys. Geosys.*, *16*(3), 681–704, doi:10.1002/2014GC005627

Bell, S., Ruan, Y., and Forsyth, D. W. (2016), Ridge asymmetry and deep aqueous alteration at the trench observed from Rayleigh wave tomography of the Juan de Fuca plate. *J. Geophys. Res.*, *121*, 2016JB012990, doi:10.1002/2016JB012990

Bezada, M. J., Byrnes, J., and Eilon, Z. (2019), On the robustness of attenuation measurements on teleseismic P waves: insights from micro-array analysis of the 2017 North Korean nuclear test. *Geophys. J. Int.*, *218*, 573–585, doi:10.1093/gji/ggz169

Bezada, Maximiliano J. (2017), Insights into the lithospheric architecture of Iberia and Morocco from teleseismic body-wave attenuation. *Earth Planet. Sci. Lett.*, *478*, 14–26, doi:10.1016/j.epsl.2017.08.029

Bodin, T., and Sambridge, M. (2009), Seismic tomography with the reversible jump algorithm. *Geophys. J. Int.*, *178*(3), 1411–1436, doi:10.1111/j.1365-246X.2009.04226.x

Bodin, T., Sambridge, M., Rawlinson, N., and Arroucau, P. (2012a), Transdimensional tomography with unknown data noise. *Geophys. J. Int.*, *189*, 1536–1556, doi:10.1111/j.1365-246X.2012.05414.x

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., and Rawlinson, N. (2012b), Transdimensional inversion of receiver functions and surface wave dispersion. *J. Geophys. Res.*, *117*, B02301, doi:10.1029/2011JB008560

Braun, M. G., Hirth, G., and Parmentier, E. M. (2000), The effects of deep damp melting on mantle flow and melt generation beneath mid-ocean ridges. *Earth Planet. Sci. Lett.*, *176,* 339–356. doi:10.1016/S0012-821X(00)00015-7

Buck, W. R., and Su, W. (1989), Focused mantle upwelling below mid-ocean ridges due to feedback between viscosity and melting. *Geophys. Res. Lett.*, *16*, 641–644, 10.1029/GL016i007p00641

Burdick, S., and Lekić, V. (2017), Velocity variations and uncertainty from transdimensional P-wave tomography of North America. *Geophys. J. Int.*, *209*, 1337–1351, 10.1093/gji/ggx091

Byrnes, J. S., Bezada, M., Long, M. D., and Benoit, M. H. (2019), Thin lithosphere beneath the central Appalachian Mountains: Constraints from seismic attenuation beneath the MAGIC array. *Earth Planet. Sci. Lett.*, *519*, 297–307, doi:0.1016/j.epsl.2019.04.045

Cafferky, S., and Schmandt, B. (2015), Teleseismic P wave spectra from USArray and implications for upper mantle attenuation and scattering. *Geochem., Geophys., Geosys.*, *16*, 3343–3361, doi:0.1002/2015GC005993

Choblet, G., and Parmentier, E. M. (2001), Mantle upwelling and melting beneath slow spreading centers: effects of variable rheology and melt productivity. *Earth Planet. Sci. Lett.*, *184*, 589–604, doi:10.1016/S0012-821X(00)00330-7

Conder, J. A., Forsyth, D. W., and Parmentier, E. M. (2002), Asthenospheric flow and asymmetry of the East Pacific Rise, MELT area. *J. Geophys. Res.*, *107*, 2344, doi:10.1029/2001JB000807

Dalton, C. A., Bao, X., and Ma, Z. (2017), The thermal structure of cratonic lithosphere from global Rayleigh wave attenuation. *Earth Planet. Sci. Lett.*, *457*, 250–262, doi:10.1016/j.epsl.2016.10.014

Darin, M. H., and Dorsey, R. J. (2013), Reconciling disparate estimates of total offset on the southern San Andreas fault. *Geology*, *41*, 975–978, doi:10.1130/G34276.1

Dettmer, J., Dosso, S. E., and Holland, C. W. (2010), Trans-dimensional geoacoustic inversion. *J. Acoust. Soc. Am.*, *128*, 3393–3405, doi:10.1121/1.3500674

Dickinson, W. R. (2009), Anatomy and global context of the North American Cordillera, in *Backbone of the Americas: Shallow Subduction, Plateau Uplift, and Ridge and Terrane Collision*, vol. 204, edited by S. M. Kay et al., pp. 1–29, Geological Society of America Memoir, Denver, USA, doi:10.1130/2009.1204.

Dong, M. T., and Menke, W. H. (2017), Seismic High Attenuation Region Observed Beneath Southern New England From Teleseismic Body Wave Spectra: Evidence for High Asthenospheric Temperature Without Melt, *Geophys. Res. Lett.*, *44*, 2017GL074953. doi:0.1002/2017GL074953

Dorsey, R. J. (2010), Sedimentation and crustal recycling along an active oblique-rift margin: Salton Trough and northern Gulf of California. *Geology*, *38*, 443–446. doi:10.1130/G30698.1

Dziewonski, A. M., and Anderson, D. L. (1981), Preliminary reference Earth model. *Phys. Earth Planet. Int.*, *25*, 297–356.

Eilon, Z., and Abers, G. (2017), High seismic attenuation at a mid-ocean ridge reveals the distribution of deep melt, *Sci. Adv.*, *3*, e1602829, doi:10.1126/sciadv.1602829

Eilon, Z., Fischer, K. M., and Dalton, C. A. (2018), An adaptive Bayesian inversion for upper-mantle structure using surface waves and scattered body waves. *Geophys. J. Int.*, *214*, 232–253, doi:0.1093/gji/ggy137

Elders, W. A., Rex, R. W., Robinson, P. T., Biehler, S., and Meidav, T. (1972), Crustal Spreading in Southern California: The Imperial Valley and the Gulf of California formed by the rifting apart of a continental plate. *Sci.*, *178*, 15–24. doi:0.1126/science.178.4056.15

Faul, U. H., and Jackson, I. (2005), The seismological signature of temperature and grain size variations in the upper mantle. *Earth Planet. Sci. Lett.*, *234*., 119–134, doi:10.1016/j.epsl.2005.02.008

Faul, U. H., Fitz Gerald, J. D., and Jackson, I. (2004), Shear wave attenuation and dispersion in melt-bearing olivine polycrystals: 2. Microstructural interpretation and seismological implications. *J. Geophys. Res.*, *109*, B06202. doi:10.1029/2003JB002407

Garapić, G., Faul, U. H., and Brisson, E. (2013), High-resolution imaging of the melt distribution in partially molten upper mantle rocks: evidence for wetted two-grain boundaries. *Geochem., Geophys., Geosys.*, *14*, 556–566, doi:10.1029/2012GC004547

Green, P. J. (1995), Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, *82*, 711–732, doi:10.1093/biomet/82.4.711

Hammond, W. C., and Humphreys, E. D. (2000), Upper mantle seismic wave attenuation: Effects of realistic partial melt distribution. *J. Geophys. Res.*, *105*, 10987–10999. doi:10.1029/2000JB900042

Han, L., Hole, J. A., Stock, J. M., Fuis, G. S., Kell, A., Driscoll, N. W., Kent, G. M., Harding, A. J., Rymer, M. J., Gonzáles-Fernández, A., Lázaro-Mancilla, O. (2016), Continental rupture and the creation of new crust in the Salton Trough rift, Southern California and northern Mexico: Results from the Salton Seismic Imaging Project: Continental Rupture in the Salton Trough. *J. Geophys. Res.*, *121*, 7469–7489, doi:10.1002/2016JB013139

Haney, M. M., Power, J., West, M., and Michaels, P. (2012), Causal Instrument Corrections for Short-Period and Broadband Seismometers. *Seis. Res. Lett.*, *83*, 834–845. doi:10.1785/0220120031

Hansen, P. C. (1992), Analysis of Discrete Ill-Posed Problems by Means of the L-Curve. *SIAM Review*, *34*, 561–580, doi:10.1137/1034115

Harmon, N., Cruz, M. S. D. L., Rychert, C. A., Abers, G., and Fischer, K. (2013), Crustal and mantle shear velocity structure of Costa Rica and Nicaragua from ambient noise and teleseismic Rayleigh wave tomography. *Geophys. J. Int.*, *195*, 1300–1313. doi:10.1093/gji/ggt309

Hauksson, E., and Shearer, P. M. (2006), Attenuation models (Qp and Qs) in three dimensions of the southern California crust: Inferred fluid saturation at seismogenic depths. *J. Geophys. Res.*, *111*, doi:10.1029/2005JB003947

Heath, B. A., Hooft, E. E. E., Toomey, D. R., and Bezada, M. J. (2015), Imaging the magmatic system of Newberry Volcano using joint active source and teleseismic tomography. *Geochem., Geophys., Geosys.*, *16*, 4433–4448, doi:10.1002/2015GC006129

Hirth, G., and D. Kohlstedt (2003), Rheology of the upper mantle and the mantle wedge: A view from the experimentalists, in Inside the Subduction Factory, *Geophys. Monogr. Ser.*, vol. 138, edited by J. Eiler, pp. 83 – 105, AGU, Washington, D. C.

Holtzman, B. K., Groebner, N. J., Zimmerman, M. E., Ginsberg, S. B., and Kohlstedt, D. L. (2003), Stress-driven melt segregation in partially molten rocks. *Geochem., Geophys., Geosys.*, *4*, 8607, doi:10.1029/2001GC000258

Holtzman, B. K., and Kendall, J.-M. (2010), Organized melt, seismic anisotropy, and plate boundary lubrication. *Geochem., Geophy., Geosys.*, *11*, Q0AB06. doi:10.1029/2010GC003296

Holtzman, B. K., King, D. S. H., and Kohlstedt, D. L. (2012), Effects of stress-driven melt segregation on the viscosity of rocks. *Earth Planet. Sci. Lett.*, *359*, 184–193. doi: 10.1016/j.epsl.2012.09.030

Humphreys, E., Clayton, R. W., and Hager, B. H. (1984), A tomographic image of mantle structure beneath Southern California. *Geophys. Res. Lett.*, *11*, 625–627, doi:10.1029/GL011i007p00625

Hwang, Y. K., and Ritsema, J. (2011), Radial Qμ structure of the lower mantle from teleseismic body-wave spectra. *Earth Planet. Sci. Lett.*, *303*, 369–375, doi:10.1016/j.epsl.2011.01.023

Hwang, Y. K., Ritsema, J., and Goes, S. (2009), Spatial variations of P wave attenuation in the mantle beneath North America. *J. Geophys. Res.*, *114*, B06312, doi:10.1029/2008JB006091

Jackson, I., and Faul, U. H. (2010), Grainsize-sensitive viscoelastic relaxation in olivine: Towards a robust laboratory-based model for seismological application, *Phys. Earth Planet. Int.*, *183*, 151–163, doi:10.1016/j.pepi.2010.09.005

Karato, S., and Spetzler, H. A. (1990), Defect microdynamics in minerals and solid-state mechanisms of seismic wave attenuation and velocity dispersion in the mantle. *Rev. Geophys.*, *28*, 399–421, doi:10.1029/RG028i004p00399

Katz, R. F. (2010), Porosity-driven convection and asymmetry beneath mid-ocean ridges. *Geochem., Geophys., Geosys.*, *11*, Q0AC07, doi:10.1029/2010GC003282

Kennett, B. L. N., and Engdahl, E. R. (1991) Travel times for global earthquake location and phase association, *Geophys. J. Int.*, *105*, 429-465, doi:10.17611/DP/9991809

Kohlstedt, D. L., and Holtzman, B. K. (2009), Shearing Melt Out of the Earth: An Experimentalist's Perspective on the Influence of Deformation on Melt Extraction. *Ann. Rev. Earth Planet. Sci.*, *37*, 561–593, doi:10.1146/annurev.earth.031208.100104

Kolb, J. M., and Lekić, V. (2014), Receiver function deconvolution using transdimensional hierarchical Bayesian inference. *Geophys. J. Int.*, *197*, 1719–1735, doi:10.1093/gji/ggu079

Lachenbruch, A. H., Sass, J. H., and Galanis, S. P. (1985), Heat flow in southernmost California and the origin of the Salton Trough. *J. Geophys. Res.*, *90*, 6709, doi:10.1029/JB090iB08p06709

Lekić, V., Matas, J., Panning, M., and Romanowicz, B. (2009), Measurement and implications of frequency dependence of attenuation. *Earth Planet. Sci. Lett.*, *282*, 285–293, doi:10.1016/j.epsl.2009.03.030

Lekić, V., French, S. W., and Fischer, K. M. (2011), Lithospheric Thinning Beneath Rifted Regions of Southern California. *Sci.*, *334*, 780–783, doi:10.1126/science.1211906

Ligi, M., Cuffaro, M., Chierici, F., and Calafato, A. (2008), Three-dimensional passive mantle flow beneath mid-ocean ridges: An analytical approach. *Geophys. J. Int.*, *175*, 783–805. doi:10.1111/j.1365-246X.2008.03931.x

Lizarralde, D., Axen, G. J., Brown, H. E., Fletcher, J. M., González-Fernández, A., Harding, A. J., Holbrook, W. S., Kent, G. M., Paramo, P., Sutherland, F., and Umhoefer, P. J. (2007), Variation in styles of rifting in the Gulf of California. *Nature*, *448*, 466–469, doi:10.1038/nature06035

Ma, Z., Dalton, C. A., Russell, J. B., Gaherty, J. B., Hirth, G., and Forsyth, D. W. (2020), Shear attenuation and anelastic mechanisms in the central Pacific upper mantle. *Earth Planet. Sci. Lett.*, *536*, 116148, doi:10.1016/j.epsl.2020.116148

MacKay, D. (2003), Information Theory, Inference, and Learning Algorithms, Cambridge Univ. Press, New York.

Malinverno, A. (2002), Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. *Geophys. J. Int.*, *151*, 675–688. doi:10.1046/j.1365-246X.2002.01847.x

Malinverno, A., and Briggs, V. A. (2004), Expanded uncertainty quantification in inverse problems: Hierarchical Bayes and empirical Bayes. *Geophys.*, *69*, 1005–1016. doi:0.1190/1.1778243

McCarthy, C., and Takei, Y. (2011), Anelasticity and viscosity of partially molten rock analogue: Toward seismic detection of small quantities of melt. *Geophys. Res. Lett.*, *38*, L18306, doi:10.1029/2011GL048776

Monteiller, V., and Chevrot, S. (2011), High-resolution imaging of the deep anisotropic structure of the San Andreas Fault system beneath southern California. *Geophys. J. Int.*, *186*, 418–446. doi:10.1111/j.1365-246X.2011.05082.x

Olugboji, T. M., Karato, S., and Park, J. (2013), Structures of the oceanic lithosphere-asthenosphere boundary: Mineral-physics modeling and seismological signatures: Structure of the oceanic LAB *Geochem., Geophys., Geosys.*, *14*(4), 880–901, doi:10.1002/ggge.20086

Olugboji, T. M., Lekic, V., and McDonough, W. (2017), A statistical assessment of seismic models of the U.S. continental crust using Bayesian inversion of ambient noise surface wave dispersion data. *Tectonics*, *36*, 1232–1253, doi:10.1002/2017TC004468

Paige, C. C., and Saunders, M. A. (1982), LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares. *ACM Trans. Math. Soft.*, *8*, 43–71, doi:10.1145/355984.355989

Persaud, P., Ma, Y., Stock, J. M., Hole, J. A., Fuis, G. S., and Han, L. (2016), Fault zone characteristics and basin complexity in the southern Salton Trough, California. *Geology*, *44*, 747–750, doi:10.1130/G38033.1

Plesch, A., Shaw, J. H., Benson, C., Bryant, W. A., Carena, S., Cooke, M., et al. (2007), Community Fault Model (CFM) for Southern California. *Bull. Seis. Soc. Am.*, *97*, 1793–1802, doi:10.1785/0120050211

Powell, R. E., and Weldon, R. J. (1992), Evolution of the San Andreas Fault. *Ann. Rev. Earth Planet. Sci.*, *20*, 431–468, doi:10.1146/annurev.ea.20.050192.002243

Pozgay, S. H., Wiens, D. A., Conder, J. A., Shiobara, H., and Sugioka, H. (2009), Seismic attenuation tomography of the Mariana subduction system: Implications for thermal structure, volatile distribution, and slow spreading dynamics. *Geochem., Geophys., Geosys.*, *10*, Q04X05, doi:10.1029/2008GC002313

Rychert, C. A., Fischer, K. M., Abers, G. A., Plank, T., Syracuse, E., Protti, J. M., et al. (2008), Strong along-arc variations in attenuation in the mantle wedge beneath Costa Rica and Nicaragua. *Geochem., Geophys., Geosys.*, *9*, Q10S10, doi:10.1029/2008GC002040

Sambridge, M., Gallagher, K., Jackson, A., and Rickwood, P. (2006), Trans-dimensional inverse problems, model comparison and the evidence. *Geophys. J. Int.*, *167*, 528–542, doi:10.1111/j.1365-246X.2006.03155.x

Schmandt, B., and Humphreys, E. (2010), Seismic heterogeneity and small-scale convection in the southern California upper mantle. *Geochem., Geophys., Geosys.*, *11*, Q05004, doi:10.1029/2010GC003042

Schmandt, B., and Lin, F.-C. (2014), P and S wave tomography of the mantle beneath the United States. *Geophys. Res. Lett.*, *41*, 2014GL061231, doi:10.1002/2014GL061231

Schmitt, A. K., and Vazquez, J. A. (2006), Alteration and remelting of nascent oceanic crust during continental rupture: Evidence from zircon geochemistry of rhyolites and xenoliths from the Salton Trough, California. *Earth Planet. Sci. Lett.*, *252*(3), 260–274, doi:10.1016/j.epsl.2006.09.041

Schmitt, Axel K., Martín, A., Stockli, D. F., Farley, K. A., and Lovera, O. M. (2013), (U-Th)/He zircon and archaeological ages for a late prehistoric eruption in the Salton Trough (California, USA), *Geology*, *41*, 7–10, doi:10.1130/G33634.1

Scott, D. R., and Stevenson, D. J. (1989), A self-consistent model of melting, magma migration and buoyancy-driven circulation beneath mid-ocean ridges. *J. Geophys. Res. 94*(B3), 2973–2988, doi:10.1029/JB094iB03p02973

Shen, Y., and Forsyth, D. W. (1992), The effects of temperature- and pressure-dependent viscosity on three-dimensional passive flow of the mantle beneath a ridge-transform System. *J. Geophys. Res. 97*, 19717–19728, doi:10.1029/92JB01467

Shito, A., Karato, S., and Park, J. (2004), Frequency dependence of Q in Earth's upper mantle inferred from continuous spectra of body waves. *Geophys. Res. Lett.*, *31*(12), doi:10.1029/2004GL019582

Stock, J. M., and Hodges, K. V. (1989), Pre-Pliocene Extension around the Gulf of California and the transfer of Baja California to the Pacific Plate. *Tectonics*, *8*(1), 99–115, doi:10.1029/TC008i001p00099

Su, W., and Buck, W. R. (1993), Buoyancy effects on mantle flow under Mid-Ocean Ridges. *J. Geophys. Res. 98*(B7), 12191–12205, doi:10.1029/93JB00994

Takei, Y. (2017), Effects of Partial Melting on Seismic Velocity and Attenuation: A New Insight from Experiments. *Ann. Rev. Earth Planet. Sci.*, *45*, 447–470, doi:10.1146/annurev-earth-063016-015820

Toomey, D. R., Wilcock, W. S. D., Conder, J. A., Forsyth, D. W., Blundy, J. D., Parmentier, E. M., and Hammond, W. C. (2002), Asymmetric mantle dynamics in the MELT region of the East Pacific Rise. *Earth Planet. Sci. Lett.*, *200*, 287–295, doi:10.1016/S0012-821X(02)00655-6

Wang, Y., Forsyth, D. W., and Savage, B. (2009), Convective upwelling in the mantle beneath the Gulf of California. *Nature*, *462*, 499–501, doi:10.1038/nature08552

Wei, S. S., and Wiens, D. A. (2018), P-wave attenuation structure of the Lau back-arc basin and implications for mantle wedge processes. *Earth Planet. Sci. Lett.*, *502*, 187–199, doi:10.1016/j.epsl.2018.09.005

Wei, S. S., and Wiens, D. A. (2020), High Bulk and Shear Attenuation Due to Partial Melt in the Tonga-Lau Back-arc Mantle. *J. Geophys. Res. 125*, e2019JB017527, doi:10.1029/2019JB017527

Wright, H. M., Vazquez, J. A., Champion, D. E., Calvert, A. T., Mangan, M. T., Stelten, M., Cooper, K. M., Herzig, C., Schriener Jr., A. (2015), Episodic Holocene eruption of the Salton Buttes rhyolites, California, from paleomagnetic, U-Th, and Ar/Ar dating. *Geochem., Geophys., Geosys.*, *16*(4), 1198–1210, doi:10.1002/2015GC005714

Yamauchi, H., and Takei, Y. (2016), Polycrystal anelasticity at near-solidus temperatures. *J. Geophys. Res. 121*, 2016JB013316, doi:10.1002/2016JB013316

Yang, Y., and Forsyth, D. W. (2008), Attenuation in the upper mantle beneath Southern California: Physical state of the lithosphere and asthenosphere. *J. Geophys. Res. 113*, B03308, doi:10.1029/2007JB005118

Yang, Y., Forsyth, D. W., and Weeraratne, D. S. (2007), Seismic attenuation near the East Pacific Rise and the origin of the low-velocity zone. *Earth Planet. Sci. Lett.*, *258*, 260–268, doi:10.1016/j.epsl.2007.03.040

Zhu, L., and Kanamori, H. (2000), Moho depth variation in southern California from teleseismic receiver functions. *J. Geophys. Res. 105*, 2969–2980, doi:10.1029/1999JB900322