Shear Wave Tomography Beneath the United States Using a Joint Inversion of Surface and Body Waves

E. M. Golos1, H. Fang1,2, H. Yao2, H. Zhang2, S. Burdick3, F. Vernon4, A. Schaeffer5, S. Lebedev6, and R. D. van der Hilst1

1Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA, 2School of Earth and Space Sciences, University of Science and Technology of China, Hefei, China, 3Department of Geology, Wayne State University, Detroit, MI, USA, 4Institute of Geophysics and Planetary Physics, University of California, San Diego, La Jolla, CA, USA, 5Department of Earth and Environmental Sciences, University of Ottawa, Ottawa, Ontario, Canada, 6Geophysics Section, School of Cosmic Physics, Dublin Institute for Advanced Sciences, Dublin, Ireland

Abstract Resolving both crustal and shallow-mantle heterogeneity, which is needed to study processes in and fluxes between crust and mantle, is still a challenge for seismic tomography. Body wave data can constrain deep features but often produce vertical smearing in the crust and upper mantle; in contrast, surface wave data can provide good vertical resolution of lithospheric structure but may lack lateral resolution and are less sensitive to the deeper Earth. These two data types are usually treated and inverted separately, and tomographic models therefore do not, in general, benefit from the complementary nature of sampling by body and surface waves. As a pragmatic alternative to full waveform inversions, we formulate linear equations for teleseismic S wave traveltimes and surface wave phase velocities and solve them simultaneously for variations in shear wave speed anomalies in the crust and upper mantle. We apply this technique to data from USArray and permanent seismic networks and present a model of seismic shear wave speed anomalies beneath the continental United States. Our joint model fits the individual data sets almost as well as separate inversions but provides a better explanation of the combined data set. It is generally consistent with previous models but shows improvements over both body wave-only and surface wave-only tomography and can lead to refinements in interpretation of features on the scale of the lithosphere and mantle transition zone.

Plain Language Summary Variations in the speed at which seismic waves travel through the Earth reveal information about the structure and history of the planet. In this study, we investigate seismic velocity variations using two common types of data from seismograms: body waves, which travel through the deep Earth, and surface waves, which provide information about the shallower layers. Commonly, these two waves are studied separately, but we adopt the method of Fang et al. (2016, https://doi.org/10.1002/2015JB012702) to produce a model of the crust and mantle of the whole Earth by using both types of data. The goal of this paper is to validate the application of this technique on a large scale, using the continental United States as a test region. We perform qualitative and quantitative tests to show that this method improves upon models made with only body or surface waves while maintaining the best fits of the individual models. We conclude that this technique is a valuable and efficient tool to study the Earth’s interior at multiple scales.

1. Introduction

Seismic tomography has become an indispensable tool for mapping the Earth’s interior structure and (combined with other information) provides constraints on temperature, composition, mineralogy, history, and dynamic behavior of the planet at multiple scales. The United States has witnessed a dramatic increase in the availability of high-quality seismic data, due largely to the USArray Transportable Array. Many tomography studies have used these data (see, e.g., Bedle & van der Lee, 2009; Bensen et al., 2009; Burdick et al., 2017; Ekström, 2017; Porter et al., 2016; Schaeffer & Lebedev, 2014; Schmandt & Lin, 2014; Yuan et al., 2014), and there is growing consensus on the first-order seismic structures (Pavlis et al., 2012). Details are still debated, however, and many questions remain about North America’s complex tectonic history.
For instance, is magmatic activity in the Yellowstone region related to a mantle plume or interactions between plates and mantle (Fouch, 2012; Nelson & Grand, 2018; Pierce & Morgan, 2009)? What is the structure of the relic Farallon Plate (Pavlis et al., 2012; Sigloch, 2011; Sigloch & Mihalynuk, 2013), and how has its subduction influenced North America’s evolution?

To address these questions, which involve lithospheric and asthenospheric interactions, it is necessary to understand Earth’s interior structure across a range of lateral scales and from the shallow crust to the deep mantle. Seismic tomography provides us with several ways to do this. Full waveform inversions are gaining traction within the global seismology community (Fichtner et al., 2009; Tape et al., 2010; Tromp et al., 2005), but these methods are still very expensive in terms of computational resources, which prohibits model space searches, and they are not well set up to benefit from the tremendous increase in seismic data. Cognizant of its shortcomings, we apply a ray approximation to develop a pragmatic approach toward joint inversions of body and surface wave data.

Traveltime tomography using body waves can offer good 3-D resolution where many rays intersect, but seismic phases that provide deep coverage often propagate nearly vertically through the crust and shallow mantle, leading to vertical smearing. In contrast, the dispersive behavior of fundamental mode surface waves provides good vertical resolution in the crust and shallow mantle, but at the frequency ranges commonly used they are less sensitive to deeper heterogeneity. By combining the complementary sensitivities of both types of data, one can aim to produce an internally consistent model with good resolution from the crust to the deeper mantle. In an attempt to achieve this objective, some studies use the results from surface wave inversions as a starting model for the inversion of body wave data (e.g., Nunn et al., 2014; Schmandt & Lin, 2014), but this does not fully address trade-offs between the two data types.

Explicit joint inversion of both sets of data is not trivial because of the different ways the intrinsically nonlinear relationship between observables and model parameters is formulated. On the one hand, body wave traveltimes can be related directly to wave speed, for instance, in formulations used by Aki and Lee (1976), Spakman and Nolet (1998), and Grand (1994). On the other hand, traditional surface wave inversions are typically linearized in two steps: the construction of 2-D phase (or group) velocity maps at different frequencies, followed by a pointwise inversion for wave speed variation with depth in order to construct a 3-D model from the 1-D models thus obtained (Bensen et al., 2009; Boschi et al., 2009). Indeed, Obrebski et al. (2011) and Porritt et al. (2014) combine Rayleigh wave and body wave data but require an initial inversion for phase velocity structure. The partitioned waveform inversion formulated by Nolet (1990) avoids the intermediate step of phase velocity maps, but in its original form the use of 1-D path-average models as intermediate products is less amenable to including body wave traveltime data.

West et al. (2004) performed a one-step inversion of surface wave phase velocities and body wave traveltimes and used previous models to impose a priori constraints on crustal structures in their inversion. In order to determine 3-D variations in shear wave speed beneath the continental United States, we follow a similar strategy and linearize the sensitivity kernels of Rayleigh waves to invert directly for spatial structure (Fang et al., 2015, 2016). Unlike previous attempts at joint inversion, our model spans the entire globe, which reduces model domain artifacts introduced by anomalies outside of our study region. Similar to other joint inversions, we solve for isotropic Vs structure. Ignoring radial anisotropy may produce artifacts, because Rayleigh waves are sensitive to Vsv, whereas the body wave measurements in our data set are mostly sensitive to Vsh, but we assume (as in other studies) that radial anisotropy will not affect the first-order, large-scale recovery of seismic structure.

To evaluate the performance and efficacy of our approach, we generate and compare four models. First, we present results from body wave traveltime tomography using S arrivals only. Second, we invert Rayleigh wave phase velocity dispersion data, using the direct inversion method of Fang et al. (2015). Third, following Fang et al. (2016), we perform an explicit joint inversion of surface and body wave data. Finally, for comparison we also perform the body wave inversion with the surface wave model as an a priori constraint on the shallow structure. We use qualitative and quantitative tests to evaluate the impacts of the addition of different data sets, and we compare our results with other tomographic models in North America. The focus of this paper is to demonstrate the improvements in resolution offered by the joint inversion on large scales, and in-depth discussions of the features within our model are left to future studies.
2. Methods

This section summarizes the pertinent aspects of asymptotic inversion of body waves and surface waves and how they are combined into an explicit joint inversion. Although the theory for traditional traveltime and surface wave tomography is well known, we include some basic equations to provide context for the joint inversion. For details we refer to Fang et al. (2015, 2016).

2.1. Body Wave Traveltime Inversion

The body wave inversion is formulated as in Aki and Lee (1976), Li et al. (2008), and Spakman and Nolet (1998). The traveltime \( t \) is the integrated wave slowness along raypath \( L \), which upon linearization is calculated in a known reference model, and \( \delta t \) is the difference in traveltime due to slowness perturbations \( \delta S \) with respect to this model:

\[
\delta t = t - t_{\text{ref}} \approx \int \delta S \, dl.
\]  

(1)

We consider deviations from the global 1-D reference model AK135 (Kennett et al., 1995), but 3-D models could also be used. The model space is discretized with a three-dimensional grid of cells with slowness perturbations \( \Delta S \). We use the superscript BW for body waves and the subscript \( \beta \) to indicate that the model describes shear wave slowness:

\[
\delta t^{BW} = \sum_{k=1}^{K} \Delta S_{k,\beta} \delta l_k \quad \text{or, in matrix form,} \quad d^{BW} = G^{BW} m,
\]  

(2)

with \( G^{BW} \) the sensitivity matrix for the body wave data and \( d^{BW} \) the vector of traveltime residuals. To account for uneven data distribution, we use an adaptive grid with irregular spacing (Kárason & van der Hilst, 2001; Li et al., 2008; Spakman & Bijwaard, 2001).

For the body waves we minimize, using the least squares (LSQR) algorithm (Nolet, 1985; Paige & Saunders, 1982), the cost function \( \epsilon \):

\[
\epsilon = ||G^{BW} m - d||^2 + k_1 ||Lm||^2 + k_2 ||m||^2,
\]  

(3)

where \( L \) is a smoothing operator, \( ||m|| \) is an L2-norm damping term that controls the magnitude of anomalies, and \( k_1 \) and \( k_2 \) are parameters that can be tuned to balance the effects of damping (see Appendix A). \( G^{BW} \) and \( m \) also contain parameters to account for event mislocation. Hereinafter, the estimate of \( m \) thus obtained is referred to as \( \hat{m}^{BW} \).

We note that additional terms could be included in the cost function, for example, to account for 3-D crustal structure (Li et al., 2006). Section 2.4 explores this option for a prior inferred from surface wave tomography.

2.2. Surface Wave Inversion

Traditionally, 2-D phase velocity maps are formed by relating \( \delta t \) to phase velocity anomalies, \( \delta C(\omega) \):

\[
\delta t(\omega) = \sum_{k=1}^{K} v_k \Delta S_k \approx \sum_{k=1}^{K} \Delta l_k \frac{\delta C(\omega)}{C^2(\omega)},
\]  

(4)

where \( v_k \) describes the coefficients of sensitivity along the great-circle path between a source and receiver. As in the partitioned waveform inversion due to Nolet (1990), the frequency-specific sensitivity can be linked to variations in density \( (\rho) \), P wave speed \( (\alpha) \), and S wave speed \( (\beta) \) with depth \( (z) \):

\[
\delta C(\omega) = \sum_{j=1}^{J} \frac{\delta C(\omega)}{\delta \rho(z_j)} \Delta \rho(z_j) + \sum_{j=1}^{J} \frac{\delta C(\omega)}{\delta \alpha(z_j)} \Delta \alpha(z_j) + \sum_{j=1}^{J} \frac{\delta C(\omega)}{\delta \beta(z_j)} \Delta \beta(z_j).
\]  

(5)

Fang et al. (2015) estimate these linear sensitivity kernels from a reference model using the method of Thomson (1950) and Haskell (1953). Sensitivities to \( \rho \) and \( \alpha \) are related to \( \beta \) by the scaling relationships \( R_{\rho,\alpha}(z) \) and \( R_{\alpha,\beta}(z) \), derived from AK135. Equations (4) and (5) are combined to express traveltime anomaly \( \delta t^{SW} \) as a linear function of \( \Delta S_j \):

\[
\delta t^{SW} = \sum_{k=1}^{K} \left( -\frac{\Delta l_k}{C^2(\omega)} \right) \sum_{j=1}^{J} \left[ R_{\rho,\alpha}(z_j) \frac{\partial C(\omega)}{\partial \rho(z_j)} + R_{\alpha,\beta}(z_j) \frac{\partial C(\omega)}{\partial \alpha(z_j)} + \frac{\partial C(\omega)}{\partial \beta(z_j)} \right] \beta_k \Delta S_{j,k}.
\]  

(6)

or \( d^{SW} = G_{\beta}^{SW} m \).

(7)
We apply the same inversion scheme as for body waves (equation (3)). Hereinafter, the model derived from surface wave data only is referred to as \( \hat{m}^{SW} \). For simplicity we assume ray-like paths, but more realistic finite frequency behavior can be incorporated through approximate kernels (e.g., Lebedev & van der Hilst, 2008).

### 2.3. Joint Inversion

For an explicit inversion of the body and surface wave data, we combine equations (2) and (7):

\[
\begin{bmatrix}
W^{BW} x^{BW} \\
W^{SW} x^{SW}
\end{bmatrix} =
\begin{bmatrix}
W^{BW} g^{BW} \\
W^{SW} g^{SW}
\end{bmatrix} \begin{bmatrix}
\Delta S_p \\
\Delta S_p
\end{bmatrix}
\text{ or } d^{BW+SW} = g^{BW+SW} m.
\]  

(8)
We invert this system using equation (3), and the model of slowness (relative to AK135) that best fits our data is hereinafter referred to as $\mathbf{m}_{BW+SW}$. Figure 1 shows the irregular joint model grid within the United States. The choice of data weights ($w_{BW}$, $w_{SW}$) and regularization parameters is discussed in Appendix A1.

### 2.4. Body Wave Inversion With A Priori Constraint From Surface Waves

For comparison to the explicit joint inversion, we follow (Burdick et al., 2008; Li et al., 2006; Waldhauser et al., 2002) and perform a body wave traveltime inversion with a 3-D model from surface waves as an a priori constraint. For this purpose, an additional term is added to the cost function (equation (3)):

$$
\epsilon = ||G^{BW}\mathbf{m} - \mathbf{d}||^2 + k_1||\mathbf{L}\mathbf{m}||^2 + k_2||\mathbf{m}||^2 + k_3||\mathbf{m} - \mathbf{m}_{SW}||^2.
$$

The solution of this system of equations is referred to as $\mathbf{m}_{prior}$. $k_3$ is determined via L-curve analysis (see Appendix A).

### 3. Data

#### 3.1. Body Wave Data

Figure 2 shows the distribution of stations and events for all body wave data used. We use $\sim$610,000 direct $S$ phase arrival times from teleseismic earthquakes (epicentral distance range of 30°−90°) of magnitude 5.5 or more. About 140,000 picks are from the USArray TA network (accessibile via the IRIS Data Management Center); the picks were made by analysts at the Array National Facility on broadband or high broadband high-gain (BH or HH) horizontal seismogram components, for events between 2004 and 2016. The remainder of the $S$ arrivals are from the Engdahl-Hilst-Buland International Seismological Centre (EHB-ISC) global catalog (Engdahl et al., 1998; EHB Bulletin, 2015), for events between 1960 and 2008. All picks are manual and mostly from short-period records. After event clustering, approximately 370,000 rays are associated with our body wave data, each of which corresponds to a row in the sensitivity matrix. Ellipticity and elevation corrections are applied to traveltimes to account for deviations from a spherical Earth.

#### 3.2. Surface Wave Data

We incorporate about 1 million fundamental mode phase velocity measurements of Rayleigh waves. From the data set of Ekström (2017), which were measured from interstation cross correlations of ambient noise recorded by USArray, we use $\sim$490,000 measurements at periods of 5−40 s. The second data set, from Schaeffer and Lebedev (2014), is determined...
by waveform fitting (Lebedev & van der Hilst, 2008; Lebedev et al., 2005) of earthquake seismograms. Raw data were obtained from USArray and the Canadian National Seismograph Network. We select \( \sim 610,000 \) measurements at 40–290 s with good coverage of the contiguous United States.

Depth sensitivity kernels of Rayleigh wave phase velocity to Vs (specifically Vsv, but we assume isotropy) are calculated according to the 1-D AK135 model and shown at a selection of periods in Figure 3. Source-receiver distributions are shown in Figure 4.

4. Results

4.1. Body Wave Inversion

The body wave model, or \( \hat{m}_{BW} \), is depicted in map view in Figure 5, and in vertical cross sections in Figures 6a and 7a. In this model, low (seismic) wave speeds below the western United States (WUS) and high wave speeds beneath the central United States (CUS) persist into the transition zone, with no clear distinction between the lithosphere and asthenosphere. In fact, the structure in map view is nearly identical between 20 km and 200 km depth, and similar even at 400 km. As shown below, this is most likely an effect of smearing along the relatively steep upper mantle raypaths of teleseismic waves.

The lateral variations near 100 km depth generally agree with previous studies. There is a distinct pattern of slow anomalies beneath the Basin and Range Province and the Columbia Plateau, with higher wave speeds near the Colorado Plateau. In the eastern United States (EUS), low wave speeds are observed below the Appalachian Mountains, the northernmost extent of which may be associated with the North Appalachian Anomaly highlighted by Porter et al. (2016) and Levin et al. (2017).

4.2. Surface Wave Inversion

The surface wave model, or \( \hat{m}_{SW} \), is depicted in Figures 8, 6b, and 7b. As expected from the sensitivity kernels (Figure 3), anomalies below \( \sim 400 \) km are faint and smooth.

Wave propagation is fast in the east to a depth of \( \sim 200 \) km, which may be interpreted as the depth extent of the cratonic lithosphere. In the west, a thinner crustal layer is characterized by higher wave speeds and is underlain by high-amplitude slow material (Figure 8). Vertical heterogeneity is more pronounced in \( \hat{m}_{SW} \) than in \( \hat{m}_{BW} \); for instance, slow anomalies along the Gulf Coast are no longer imaged as a (vertically) continuous feature as in \( \hat{m}_{BW} \) but rather as separate anomalies in the crust (possibly due to thick sediments), and near 100 km. Meanwhile, the low wave speeds in the southern and central Appalachians are restricted to the 40–60 km depth range, whereas the NAA is a distinct mantle feature, first appearing at 60 km and extending below 100 km.

4.3. Joint Inversion

Results from the full joint inversion, \( \hat{m}_{BW+SW} \), are shown in Figures 9, 6c, and 7c. While the first-order pattern of fast wave speeds in the EUS and slow in the WUS resembles \( \hat{m}_{BW} \), it is here confined to shallower depths than in \( \hat{m}_{BW} \). At larger depths wave speeds are generally lower than AK135, but fast features reappear in the mantle transition zone (seen best at B-B'). Furthermore, features that were earlier obscured by effects of smearing emerge. For instance, the high wave speed anomaly beneath the WUS at 400 km is stronger in \( \hat{m}_{BW+SW} \).
Figure 5. Map views of wave speed anomalies according to the inversion of body wave data only, $\mathbf{m}^{BW}$, at (a) 20, (b) 40, (c) 60, (d) 100, (e) 200, (f) 400, (g) 600, and (h) 1,000 km. The amplitude of anomalies decreases from $\pm5\%$ in the crust to $\pm2\%$ in the lower mantle. Major geologic boundaries are superimposed on slices: CP = Columbia Plateau; BNR = Basin and Range Province; RM = Rocky Mountain range; CUS = central US; O = Ozark Plateau; ME = Mississippi Embayment; A = Appalachian Mountain range; NAA = North Appalachian Anomaly. Boundaries are adapted from https://water.usgs.gov/GIS/metadata/usgswrd/XML/physio.xml. The gray lines A-A$'$ and B-B$'$ indicate the locations of vertical profiles for Figures 6, 7, 11, and 13. All perturbations are relative to AK135.
Figure 6. Vertical cross sections of wave speed anomalies along line A-A′ (see Figure 5a) relative to AK135 for (a) the body wave only inversion, (b) the surface wave only inversion, and (c) the explicit joint inversion. The slices extend from the surface to 1,000-km depth, and exaggerated topography is shown (gray and blue indicate above and below sea level, respectively). Geologic provinces are the same as in Figure 5, with the addition of the Yellowstone hot spot in red (Y). Anomaly magnitudes are ±3%.

than in \( \hat{m}_{BW} \) or \( \hat{m}_{SW} \). In addition, lower-crustal heterogeneities in the Midcontinent Rift and the North Appalachian Anomaly areas visible in \( \hat{m}_{SW} \).

4.4. Body Wave Inversion With Prior Model Constraint Term

For comparison, the results from the inversion of body wave data using \( \hat{m}_{SW} \) as a prior model, \( \hat{m}_{prior} \), are shown in Figures 10 and 11. The contrast between the fast EUS and slow WUS still extends into the transition zone as in \( \hat{m}_{BW} \), but some vertical variation is seen around 200 km. In addition, the amplitudes of anomalies in the lithosphere are closer to those of \( \hat{m}_{SW} \) than \( \hat{m}_{BW} \). This model thus contains elements from both the body wave and surface wave inversions, but some artifacts, such as vertical smearing, are still apparent.

5. Discussion

5.1. Visual Model Analysis

A first analysis of the behavior of the four inversions, and how the resulting models differ on continental scales, is based on visual inspection of the results. As expected, \( \hat{m}_{BW-SW} \) contains elements of \( \hat{m}_{BW} \) and \( \hat{m}_{SW} \). Deeper features resemble \( \hat{m}_{BW} \), whereas shallower regions exhibit similarities to \( \hat{m}_{SW} \). At 100 km depth (Figures 5d, 8d, 9d), lateral resolution is adequate for both data sets and the inferred lateral variations are similar. For \( \hat{m}_{BW} \) the shallow structure tends to be smeared to larger depths, but upon joint inversion this effect is mitigated by surface waves. At 600 and 1,000 km, the joint inversion model resembles (though is not identical to) \( \hat{m}_{BW} \). Thus, body waves and surface waves each dominate the portions of our model space where their constraints are most robust.
Figure 7. Vertical cross sections of wave speed anomalies along line B-B′ (see Figure 5a) relative to AK135 for (a) the body wave only inversion, (b) the surface wave only inversion, and (c) the explicit joint inversion. The slices extend from the surface to 1,000-km depth, and exaggerated topography is shown (gray and blue indicate above and below sea level, respectively). Geologic provinces are the same as in Figure 5. Anomaly magnitudes are ±3%.

\( \hat{m}_{\text{prior}} \) shows higher-magnitude wave speed variations at lithospheric depths than \( \hat{m}_{\text{BW}} \), in agreement with \( \hat{m}_{\text{SW}} \), but the slow anomalies in the WUS extend into the mantle transition zone, suggesting that vertical smearing is still a problem in this type of inversion. In addition, the fast anomaly in the EUS diminishes abruptly below 200 km across much of the study region (Figure 11a), and the variations in depth of this feature that are visible in \( \hat{m}_{\text{SW}} \) are lost in \( \hat{m}_{\text{prior}} \). These (and other) visual differences between \( \hat{m}_{\text{BW}}+\hat{m}_{\text{SW}} \) and \( \hat{m}_{\text{prior}} \) lead us to conclude that our explicit inversion is more effective in mitigating smearing than body wave data, and in capturing features from both data sets consistently between the EUS and WUS.

The Yellowstone anomaly, a significant slow feature, highlights the differences between the four models. In vertical section A-A′ (Figure 6) in \( \hat{m}_{\text{BW}}+\hat{m}_{\text{SW}} \) a strongly slow feature extends below the lithosphere to nearly 400 km and may connect to a deeper slow anomaly. These observations are consistent with Nelson and Grand (2018). In contrast, in \( \hat{m}_{\text{BW}} \) and \( \hat{m}_{\text{prior}} \) it is hard to distinguish small-scale features from the slow background wave speed; in \( \hat{m}_{\text{SW}} \) resolution is insufficient for a confident analysis of the geometry of the anomaly below about 200 km.

Another region that illustrates the effects of different inversion strategies is the Appalachian Mountain Range, which in the upper mantle is characterized by wave speeds that are lower than those of its surroundings. In \( \hat{m}_{\text{BW}} \) and \( \hat{m}_{\text{prior}} \) the slow central and northern anomalies are hard to distinguish from one another and are imaged from 20 to 400 km depths. In \( \hat{m}_{\text{SW}} \) and \( \hat{m}_{\text{BW}}+\hat{m}_{\text{SW}} \) we detect a shallower feature (40–60 km) in the central and southern Appalachians. Distinct from this, the NAA below New England is visible from 60 to possibly 400 km, which is consistent with the interpretation of Levin et al. (2017) that this anomaly represents a small-scale mantle upwelling.
Figure 8. Map views of wave speed anomalies relative to AK135 according to the inversion of surface wave data only, $m^\text{SW}$. Due to the frequencies used for fundamental mode surface waves (Figure 3), the model is less robust below 300 km. At depths of (a) 20, (b) 40, (c) 60, (d) 100, (e) 200, (f) 400, (g) 600, and (h) 1,000 km.
Figure 9. Map views of wave speed anomalies relative to AK135 according to the explicit joint inversion, $\hat{m}_{BW+SW}$. At depths of (a) 20, (b) 40, (c) 60, (d) 100, (e) 200, (f) 400, (g) 600, and (h) 1,000 km.
Figure 10. Map views of wave speed anomalies relative to AK135 according to the inversion of body wave data with the surface wave model as an a priori constraint, $m_{\text{prior}}$. At depths of (a) 20, (b) 40, (c) 60, (d) 100, (e) 200, (f) 400, (g) 600, and (h) 1,000 km.
We also note improved resolution at the scale of the entire lithosphere. In the eastern half of section B-B’ in $\hat{m}^{BW-SW}$ (Figure 7c) the data resolve a separation between fast material near the surface and multiple anomalies in the transition zone and top of the lower mantle. We interpret the former as thick continental lithosphere; the average depth extent is approximately 200 km, which is roughly consistent with estimates derived from the heat flow data of Jaupart and Mareschal (2015) and the heat flow-lithospheric thickness relationships in Mareschal and Jaupart (2004).

This lithosphere is underlain by laterally heterogeneous slower anomalies (e.g., Figure 7c), but near 600 km we again detect fast wave speeds. In $\hat{m}^{BW}$ and previous studies by others, fast features at these depths are difficult to distinguish from the lithospheric signature, but in $\hat{m}^{BW-SW}$ they are clearly separate from shallower structures. Fast anomalies near the mantle transition zone are a robust feature of many other tomographic models, including those which account for anisotropy (French & Romanowicz, 2014; Porritt et al., 2014; Ritsema et al., 2011). These anomalies have been associated with fragments of the Farallon plate (Forte et al., 2007; Sigloch, 2011), remnants of other subducted material (Schmandt & Lin, 2014; Sigloch & Mihalynuk, 2013), or foundering lithosphere (Biryol et al., 2016). A full interpretation of these features will be the subject of future work.

5.2. Synthetic Tests

For a qualitative assessment of the ability of data to resolve structural features we perform a checkerboard test. As a quantitative test this has many shortcomings (e.g., Lévêque et al., 1993; van der Hilst et al., 1993), but it is an effective tool to assess the influence of different inversion strategies. We invert synthetic data with the matrices used to construct $\hat{m}^{BW}$, $\hat{m}^{SW}$, and $\hat{m}^{BW-SW}$. We show results for a $5^\circ \times 5^\circ$ horizontal, 20-km thick, checkerboard pattern (Figure 12). This analysis was not performed for $\hat{m}^{prior}$, because the surface wave data play no role in the inversion.

The body wave inversion constrains anomaly amplitudes poorly in the upper 100 km, but lateral resolution is good. At 400 and 1,000 km amplitude recovery improves, but some checkers are distorted. The surface wave inversion recovers both shape and amplitude well in the shallower portion of the model; quality degrades with depth, however, and at 400 km only faint and smeared patterns are resolved, suggesting that for these depths we cannot interpret the results of $\hat{m}^{SW}$ with confidence.
The joint inversion improves the recovery of checkerboard patterns at all depths. The recovered structures resemble the results of the body wave inversion at 400 km and deeper, and the surface wave inversion at smaller depths. Amplitudes are more uniformly strong at 400 and 1,000 km than for $m_{BW}$. Some of the improvement may be attributed to the increase in data quantity in the joint inversion.

We also perform a test in which we invert body wave data that are calculated from the joint inversion model $m_{BW+SW}$. As expected, the body wave data cannot resolve the structural features introduced by the surface waves, and the results of this inversion (Figure 13b) more closely resemble $m_{BW}$ than $m_{BW+SW}$. In particular, as in $m_{BW}$, high wave speeds continue across the transition zone and the lithospheric signature...
is not distinguishable from the deep features, even if they are clearly separate in $\hat{m}_{BW+SW}$. This suggests that body waves alone are insufficient for recovering the thickness of cratonic lithosphere and demonstrates that the appearance of the high wave speeds continuing from lithospheric depths across the transition zone in the BNR (as in Figures 13c) is, indeed, an artifact due to smearing along steep teleseismic raypaths.

### 5.3. Variance Reduction Comparison

Variance reduction (VR) provides a quantitative measure of the improvement in data fit upon tomographic inversions. The term data residual, or $\hat{r}$, refers to the difference between observed traveltime anomalies and those predicted from a particular model, that is, $\hat{r} = \delta t - G \hat{m}$. We calculate $\hat{r}$ for the nine combinations produced by three data sets (body wave data only, surface wave data only, and body and surface wave data combined) and three inversion models ($\hat{m}_{BW}$, $\hat{m}_{SW}$, and $\hat{m}_{BW+SW}$).

Assuming a normal distribution of each set of $\hat{r}$, we compute VR:

$$VR = 1 - \frac{\text{Var}(\hat{r})}{\text{Var}(\delta t)}. \quad (10)$$

These values are displayed in Figure 14. A higher VR implies a better fit to the data. As expected, the fit of body waves and surface waves for $\hat{m}_{BW+SW}$ is slightly less than that of the separate inversions ($\hat{m}_{BW}$, $\hat{m}_{SW}$), but the overall data VR is substantially greater after the joint inversion. Thus, the joint inversion preserves the best fits from the body and surface wave inversions while improving the fit of the combined data set. In Appendix A2 we discuss the error in these VR measurements.
5.4. Assumptions

As mentioned in section 1, we ignore anisotropy and attenuation, and we use an asymptotic (ray theoretical) formulation for wave propagation.

Since Rayleigh waves are sensitive to Vsv structure and the steeply arriving teleseismic body waves mostly to Vsh, the neglect of radial anisotropy could introduce artifacts. First, we note that this would similarly affect other joint inversions (e.g., Obrebski et al., 2011) or body wave inversions that use a surface wave model as a starting model (e.g., Schmandt & Lin, 2014). Second, while important for absolute quantitative accuracy, ignoring radial anisotropy is not likely to be a main source of uncertainty for the results presented here. Porritt et al. (2014) showed that Vsv and Vsh in model DNA13, which are derived from body wave data, are nearly identical at all depths, including within the crust. This suggests that body waves generally do not introduce strong artifacts from the Vsh structure into inversions at the continental scale. Furthermore, the average travel-time anomaly accrued by passing a vertical ray through a 250-km thick layer with wave speeds equivalent to the Vsh and Vsv models from French and Romanowicz (2014) is about 1 s, which is comparable to the average magnitude of misfit for our body wave model (2.3 s).

Azimuthal anisotropy has been inferred to be significant in the North American craton (e.g., Yuan et al., 2014) but is not considered in the illustration of the different inversion strategies presented here. Azimuthal anisotropy mostly concerns surface waves, and several studies have demonstrated that with adequate data coverage the trade-offs between isotropic and (azimuthally) anisotropic structures are small (Simons et al., 2002; Yao et al., 2010). For example, Simons et al. (2002) reported an increase in variance reduction of 3.5% in the Australian craton. We thus assume that the effects of azimuthal anisotropy are secondary to the effects we examine in this study.

Finite frequency effects can be accounted for within the framework of our joint inversion, but to illustrate the effects of different (joint) inversion strategies, we use ray theory. Several studies have shown that the use of ray theory or simple approximations to finite frequency theory produces similar results (e.g., Maceira et al., 2015; Montelli et al., 2004; Ritzwoller et al., 2002; van der Hilst & de Hoop, 2005). Full wave theory is, of course, superior, but in view of its formidable demands on computational resources there is still value in the development of simple but effective inversion schemes. We note that in our approach, the ray-like character of body wave sensitivity kernels for short-period data and the path-average character of the fundamental mode surface waves are preserved.

Another possible effect of seismic wave propagation on our model is frequency-dependent attenuation, which influences the low-frequency surface waves and high-frequency body waves in our model differently. Attenuation can alter the observed phase velocities for surface waves (Lekić et al., 2009). In Rayleigh waves at periods of 60–150 s anelastic effects may account for approximately 20% of observed delays (Ruan & Zhou, 2010), but this effect is smaller over continental regions (Ruan & Zhou, 2010) and the temperature dependence of elastic and anelastic variations are correlated. Therefore, attenuation is not expected to have a significant influence on the 3-D heterogeneity inferred here.

Figure 15 displays horizontal slices near 100, 400, and 600 km for different models: MITS_BWSW_18 (this study), SL2013NA (Schaeffer & Lebedev, 2014), SEMUCB-WM1 (French & Romanowicz, 2014), and US-SL-2014 (Schmandt & Lin, 2014). This selection represents a range of methods and scopes: SL2013NA and SEMUCB-WM1 are waveform-based inversions which account for anisotropy, while US-SL-2014 and our model are (partly) based on traveltime tomography. All models exhibit similar large-scale behavior around 100 km. Below this depth, SL2013NA and SEMUCB-WM1 are smoother and cannot distinguish features such as the two fast anomalies below the CUS at 600 km. US-SL-2014 does not include crustal depths and shows differences, for example, in high wave speeds near 400 km beneath CUS. This feature is also present in \( \hat{m} \text{BW} \), and we suspect it to be an artifact due to vertical smearing (see above and Figure 13).
6. Conclusion

We have adapted the joint inversion algorithm of Fang et al. (2016) to a continental scale to generate a shear wave speed model of the United States using both surface wave dispersion data and teleseismic body wave traveltimes. This paper validates the application of this method on such scales, within the limits, of course, of a ray-based isotropic inversion. Furthermore, we demonstrate that such a method performs better quantitatively and qualitatively than traditional ray-based inversions: the explicit joint inversion model fits surface and body wave data nearly as well as their respective individual models and greatly improves the overall data fit. The resultant model, $m^{\text{BW}+\text{SW}}$, resembles previous shear wave mantle models derived from body waves but has better resolution in the crust and upper mantle due to the incorporation of surface waves.

To a first order, our new model contains many features described in previous studies, including a slow Basin and Range, faster wave speeds beneath the Colorado Plateau, and slow shallow anomalies along the Mississippi Embayment. The southern and eastern boundaries of the cratonic region approximately follow the margins of Precambrian rift systems, as observed by Schmandt and Lin (2014). A notable effect of the joint inversion is the suppression of vertical smearing along steep raypaths resulting in the separate resolution of fast anomalies in the lithosphere and the deep transition zone beneath the CUS (instead of a continuous structure across the transition zone as seen in most body wave models). The anomalies in our favored model correspond well with features in the isotropic Vs portion of models that account for anisotropy and finite frequency effects.

Comprehensive interpretation of this model will be left to future papers, but we mention several interesting features. In contrast to previous studies, we now begin to distinguish lithospheric from asthenospheric anomalies in the EUS. We identify a fast layer near the surface as the lithosphere, which is separate from fast anomalies near the mantle transition zone. Relying on complementary data from multiple parts of the seismic waveform is essential to understanding these features and the processes of which they are brief snapshots. This is best done through rigorous full waveform inversion, but in view of its formidable computational cost simplified approaches such as the joint inversion presented here may present attractive alternatives for expedient determination of first-order features of lithospheric and mantle heterogeneity.
Appendix A

A1. Data Weighting and Inversion Parameter Selection

A major challenge when jointly interpreting multiple data sets is determining the relative weighting of data and inversion parameters; here we introduce a workflow for visually determining these parameters. We use a Pareto front to balance the fit of body wave and surface wave data. We perform 15 inversions, varying the relative contribution, $p$, of the surface wave data from 0 to 1; the body wave contribution decreases accordingly from 1 to 0. Figure A1 displays the size of the normalized data error vector $(\mathbf{Gm} - \mathbf{d})$ for surface waves against that of body waves.

The weights for surface waves are $w_{SW,i} = a^p N_i$, and for body waves, $w_{BW,i} = a^{1-p} N_i$, where $N_i$ is the number of measurements in data set $i$ and $a$ is a scaling constant which brings both $w_i$ to a reasonable order of magnitude.

We assign objective functions $\Theta$ based on the misfit for each data set:

$$\Theta_i = \sqrt{\sum_{j=1}^{J} G_{i,j} m_{i,j} - d_{i,j}} / N_i.$$  (A1)

We choose $p = 0.4$ as the utopia point, where both objective functions are balanced (Dal Moro & Pipan, 2007; Stadler, 1979). Points with $0.3 \leq p \leq 0.6$ plot close to one another, suggesting that the trade-off between the two types of data is fairly stable.

Next, we use an L curve to find the values of regularisation parameters $k_1$ and $k_2$ (equation (3)), and $k_3$ for $\hat{m}_{prior}$. Because shallower cells see more ray passes, even accounting for the increased fineness, we prescribe stronger smoothing within the United States above 300 km; lateral and radial smoothing are equal. We visually select the value at which data misfit and smoothness are balanced (Figure A2). The same technique is employed to select $k_2$. Since these values optimize the model globally, we may adjust $k_1$ and $k_2$ according to other visual tests for the United States (section 5.2).

A2. Variance Reduction Error and Uncertainty

The efficiency of our inversion algorithm allows us to estimate error and uncertainty using statistical bootstrapping methods. This is often done by jackknifing (Gung & Romanowicz, 2004; Su & Dziewonski, 1997). Because our resolution and parameterization are tied to data coverage, instead of resampling, we perturb the observed traveltimes of the full data set. This is akin to Humphreys and Clayton (1988), but we derive our perturbations from the misfit distribution of our inversion.

We define a vector $\mathbf{\hat{r}} = \delta \mathbf{t} - \mathbf{Gm}$, and its random permutation, $\mathbf{\hat{r}}^*$. We add this permuted vector to the observed traveltimes to obtain a perturbed data set $\mathbf{d}^*$.

$$\mathbf{d}^* = \delta \mathbf{t} + \mathbf{\hat{r}}^*.$$  (A2)

We perform such a test for three models: $\hat{m}_{BW}$, $\hat{m}_{SW}$, and $\hat{m}_{BW+SW}$. Fifty inversions of independently perturbed data are performed for each, from which we can estimate error and uncertainty for statistical properties.

We analyze the error in the variance reduction calculations shown in the main text. Figure A3 displays the mean VR, $\overline{VR}$, from these experiments, using the set of 50 bootstrapped models. The size of each point is proportional to the standard deviation of all $\overline{VR}$, that is, the standard error associated with this value. Body wave models tend to fit each type of data with a consistent VR, but higher standard errors, while surface wave models return lower errors but fit the body wave data poorly. The joint models do have higher error than the surface wave models, but this error...
is consistent between the data sets. Furthermore, as we observed in section 5.3, the body wave and over-all data have a higher VR from the joint inversion without sacrificing the high VR of surface wave data. We conclude that the joint inversion improves the overall fit of data, even with added noise, in a robust and consistent manner.

Figure A3. Mean variance reduction (VR) for each model and data type for the bootstrap test. The size of point is proportional to standard error of the VR, that is, standard deviation of the VR mean for all trials.

Acknowledgments
We thank Luciana Astiz and the team at the Array National Facility at UCSD for their tireless help in assembling the data set of S wave arrivals. Thanks also to Raj Moulik for his help in navigating the EHB catalog and to an anonymous reviewer for their helpful feedback, which has improved this manuscript. All body wave data used in this study can be accessed via the EarthScope ANF archives (anf.ucsd.edu/tools/events/) or EHB database (http://www.isc.ac.uk/). The surface wave measurements from the USANT15 data set can be accessed at http://www.ldeo.columbia.edu/~ekstrom/Projects/ANT/USANT15/data.html. For more information regarding the Earthscope USArray project, please see http://www.earthscope.org/; for more information regarding the Canadian National Seismograph Network, see http://www.earthquakescanada.nrcan.gc.ca/stndon/CNSN-RNSC/index-en.php. Station and event maps were generated using the Generic Mapping Tools (Wessel et al., 2013).

References


