Localized imaging of the uppermost mantle with USArray Pn data

J. S. Buehler[1] and P. M. Shearer[1]

Received 6 May 2012; revised 1 August 2012; accepted 2 August 2012; published 14 September 2012.

[1] USArray has now provided several years of high-quality seismic data and improved ray coverage for much of the western United States. This allows increased resolution for regional studies of the lithosphere and deeper structure of the North American continent. In this study, we use Pn phases in the USArray data set to solve for velocity structure in the uppermost mantle in the western United States. This article focuses on localized imaging techniques that complement traditional Pn tomography analysis. We apply waveform cross-correlation to obtain inter-station travel times between the closely and uniformly spaced USArray stations. This allows us to use traces without phase picks and reduces errors associated with the picking. We obtain differential times that can directly be used to fit locally for slowness and, depending on the approach, for the direction and curvature of the incoming wavefront. The various measurements of incoming wavefronts at different sub-arrays provide constraints on azimuthal variations in velocity. The traditional tomography approach and the local fitting method reveal similar large-scale features. No regularization is applied with the local method, and the resulting velocity maps reveal smaller-scale structures than the tomographic images.

Citation: Buehler, J. S., and P. M. Shearer (2012), Localized imaging of the uppermost mantle with USArray Pn data, *J. Geophys. Res.*, *117*, B09305, doi:10.1029/2012JB009433.

1. Introduction

[2] With the implementation of USArray, data coverage in regions without local seismic networks has been much improved, which has led to better resolved crust [e.g., *Moschetti et al.*, 2010] and mantle structure [e.g., *Burdick et al.*, 2008; *Schmandt and Humphreys*, 2010] beneath the western United States. However, even with greater data coverage, traditional tomography approaches usually require regularization due to uneven grid sampling and ray tracing is often simplified. Alternative methods for surface- and body-wave data that profit from the large USArray station network and complement traditional tomography techniques have been proposed, as for example local surface-wave tomography [e.g., *Lin et al.*, 2009; *Pollitz and Snoke*, 2010; *Lin and Ritzwoller*, 2011] or waveform based analysis [*Sun and Helmberger*, 2011].

[3] Here we present a localized array approach to map regional velocity structure using horizontally traveling Pn waves. We focus mainly on the localization method, and the reader is referred to *Buehler and Shearer* [2010] for a discussion of the tectonic implications of our models. Because Pn is a Moho refracted wave that travels in the mantle lid, it provides constraints on seismic velocity and anisotropy within a very confined depth interval. These maps provide information on temperature and compositional variations in the upper mantle [e.g., *Goes and van der Lee*, 2002; *Perry et al.*, 2006]. Even with the dense USArray station coverage, traditional teleseismic body-wave tomography studies often do not have the vertical resolution to accurately map velocity structure in the uppermost mantle. Combining surface- and body-wave observations provides better depth resolution, and a recent USArray shear wave study adds short-period surface-wave data to increase resolution in the shallowest part of the model [*Obrebski et al.*, 2011]. Moho refracted Pn waves provide additional information in the uppermost mantle, with the capacity to further constrain surface-wave or other body-wave tomographies that average anomalies over larger depth intervals.

[4] We exploit the closely spaced transportable array stations to apply localized wavefront fitting techniques to Pn data to image regional velocity structures and anisotropy in the uppermost mantle below the western United States. These array-based methods provide velocity maps similar to traditional tomography, but have the potential to reveal anomalous regions with better resolution because they do not rely on regularization. These local approaches are fast and simple, can account for off-great circle-path arrivals, and help validate traditional Pn tomography results.

[5] Working with the Pn phase is challenging because the emergent character of the arrival makes accurate phase picking difficult. In section two we describe the data selection and the cross-correlation procedure we apply to improve the relative timing between stations. Using these times, we solve for the direction and apparent velocity of the incoming wavefronts using a local approach, as illustrated in section three. We use the apparent velocities to produce maps

[1]Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, USA.

Corresponding author: J. S. Buehler, Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92121–2460, USA. (jsbuehle@ucsd.edu)

0148-0227/12/2012JB009433

similar to traditional tomography. The directions of arrival are later used to investigate velocity changes with back-azimuth, which can be related to upper mantle anisotropy and Moho slope. In section four we show that traditional Pn travel time tomography produces images of large-scale velocity structure and fast anisotropy directions that largely agree with the local estimates.

2. Data Selection and Processing

[6] The USArray Array Network Facility (*ANF*) analysts routinely pick first (and on a number of traces later) arrivals in the USArray data set. These picks are available for download through the *ANF* website (http://anf.ucsd.edu/tools/events/download.php, last accessed December 2011), and provide data that can be used directly in traditional travel-time tomography [e.g., *Buehler and Shearer*, 2010; *Burdick et al.*, 2008, 2010, 2012; *Steck et al.*, 2011].

[7] Our goal is to improve the relative timing between the transportable array stations in an automated and objective procedure. Pn waveforms characteristically have weak amplitudes and/or emergent onsets. This makes waveform processing and precise timing challenging. Accurate inter-station differential times are important for local velocity measurements among nearby stations. To reduce timing errors and improve the consistency of the times, we cross-correlate the Pn waveforms of regional earthquakes. This has the further advantage of often providing times for waveforms with no analyst-assigned Pn picks. The inter-station differential times associated with the cross-correlation peaks can then be used to fit for the apparent slowness, azimuth, and curvature of the wavefront as it travels across a localized subset of the stations. In addition, we can use the differential times to realign the seismograms to obtain self-consistent data sets for traditional tomography [*Vandecar and Crosson*, 1990].

2.1. Stations and Events

[8] Pn is a refracted ray traveling in the mantle lid just below the Moho, and arrives first at stations at regional distances. The cross-over distance, where the first arrivals change from Pg to Pn, depends on crustal thickness. Pn typically arrives first at stations with an epicentral distance of about 150 km and larger on continents. Here we use stations that are between 180 km and 1450 km from the epicenter for receivers west of longitude 110°W, and between 240 km and 1450 km for receivers east of 110°W. We make this division to accommodate the increased thickness of the crust east of the Great Basin in order to avoid erroneous picks and/or cross-correlation alignments on the usually higher-amplitude crustal phase Pg without eliminating too much data in the west. In addition we discard events with hypocenter depths larger than 30 km.

[9] We collect all available ANF picks from May 2004 to August 2011 that satisfy the criteria above, recorded at stations west of 100°W. To identify the Pn picks we proceed similarly to *Hearn* [1996] and apply an iterative procedure to identify picks with a maximum residual of 6 s following a simple straight-line fit to time versus distance (requiring at least 10 picks per earthquake and 10 picks per station). There are ~371,000 P picks currently in the ANF database at epicentral distances between 180 and 1450 km, of which ~125,000 are labeled Pn. After the windowing procedure ~161,000 picks remain. This group of arrivals is associated with 6348 earthquakes and quarry blasts.

[10] For this set of events we download waveforms from the ANF with the Antelope database tools, including traces without an analyst pick. Figure 1 shows the resulting event and station locations. The stations are color coded by the number of months data from our time period was recorded at that receiver. We save the waveforms at 40 Hz sample rate (stations with lower sampling rates are re-sampled using spline interpolation) and associated meta-data in an event-based data archive for further processing.

2.2. Pn Waveform Cross-Correlation

[11] Pn waveforms are challenging to cross-correlate reliably because they often have marginal signal-to-noise and are typically followed by the higher-amplitude Pg phase. It is important to initially align the waveforms well to avoid misidentifying Pg as Pn. Fortunately the ANF analyst picks provide good reference times for many of the seismograms. For seismograms without picks we apply an automatic picking algorithm based on a short-time-average to long-time-average ratio (STA/LTA) function [*Earle and Shearer*, 1994]. This algorithm works well with emergent and low amplitude arrivals. We high-pass all the seismograms at 1 Hz for consistent automatic picking.

[12] In some cases the autopicker may only trigger on the generally higher-amplitude crustal phase. To avoid late alignments on Pg, we restrict the search for picks to a predefined Pn time window. We calculate a mean slowness and intercept point with straight-line fits using ANF picks that are 'Pn' labeled. If an event is associated with at least 10 Pn picks, we calculate the theoretical arrival times for all the seismograms with a straight-line fit through the single-event data points. Otherwise we use the slope and intersection values calculated for the whole data set. We then use this average velocity ±0.4 km/s to calculate the start and end times of the picking window. We require an STA/LTA value of at least 2.5 within the window in order for the pick to be recorded. Once the autopicker is applied to all the traces, we discard seismograms that have neither an analyst pick nor an automatic one in the defined Pn windows. Then we iterate again to remove events that are associated with less than ten picks, stations that record less than ten earthquakes, and arrivals with residuals larger than six seconds after a straight-line fit.

[13] For each event, we then cross-correlate every trace with every other trace for all the n first arrival times of an event. This produces $n(n-1)/2$ cross-correlations per event. We experimented with different correlation window lengths and obtained satisfactory results with widths two to three times the length of the dominant period centered around the pick. The seismograms are bandpass filtered between 0.8 and 2.5 Hz prior to cross-correlation to eliminate microseisms and uncorrelated high-frequency noise. We find that in many cases we can prevent cycle skipping if we keep the maximum allowed time shift around one second. Figure 2 shows examples of cross-correlations of various qualities. Following some experimentation, we settled on a uniform cross-correlation window length of 2.5 seconds for all the traces, starting one second before the pick.

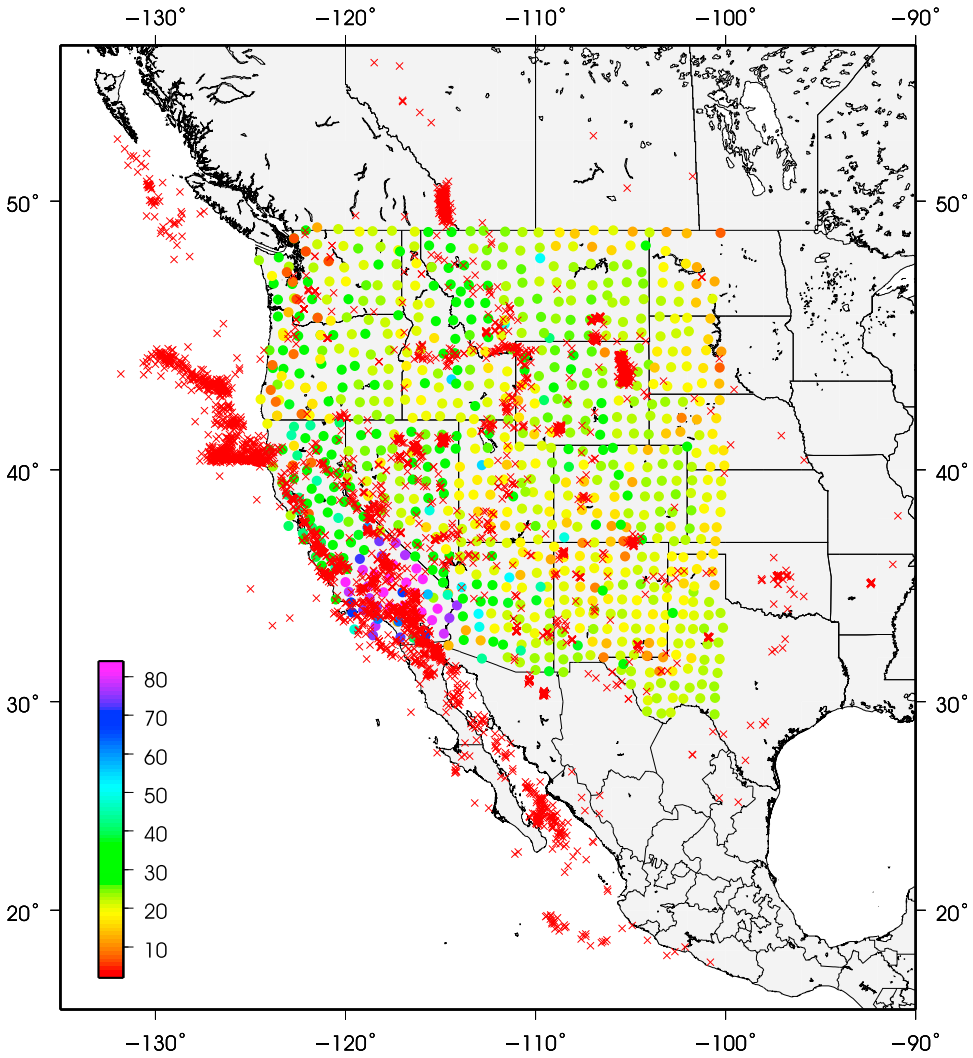

Figure 1. Overview map showing the study region, the final set of events (red crosses) used in the travel time tomography as well as the localization approach, and the station locations (colored circles). The stations are color coded by the number of months data from our time period was recorded at that receiver.

2.3. Waveform Re-Alignment

[14] Differential times from cross-correlations for a set of receivers can directly be used for subsequent wavefront fitting. However, the relative travel-time measurements are almost never consistent with a single set of absolute times, i.e., $\tau_{12} + \tau_{23} \neq \tau_{13}$, where τ_{12} is the differential time between station one and two, etc. *Vandecar and Crosson* [1990] used the measured differential times τ_{ij} between stations in a linear system to solve for a set of realigned absolute arrival times t for an event, minimizing the penalty function $F = \sum(\tau_{ij} - (t_i - t_j))^2$. This provides a self-consistent set of arrival times for every earthquake or quarry blast. The adjusted arrival times can be used both in traditional travel-time tomography and localization approaches. We proceed in similar fashion as *Vandecar and Crosson* [1990], but find that for the Pn differential times a robust regression method provides better results compared to ordinary least squares, due to occasional large errors in the differential times (e.g., cycle-skipping in the cross-correlation peaks). We applied the *robustfit* routine from MATLAB that implements the Iteratively Reweighted Least-Squares (IRLS) method.

[15] To further reduce errors associated with cycle skips we adapted a procedure based on the approach by *Vandecar and Crosson* [1990]: First, the system of equations is solved with a data vector d that contains the inter-station time shifts associated with the peak of the correlation-function. Next we find the data residuals >0.25 s and test if any other shifts associated with peaks in the cross-correlation function with values higher than 0.6 provide a better fit. We exchange the elements in the data vector accordingly and proceed iteratively to obtain the final set of realigned arrival times (Figure 2).

[16] We expect neighboring stations to have the same polarity and more similar waveform characteristics than receivers further apart. Hence the differential time observations are weighted according to inter-station distance, assigning a weight of 1 to all station pairs that are less than 250 km apart, and decaying weights with larger distances.

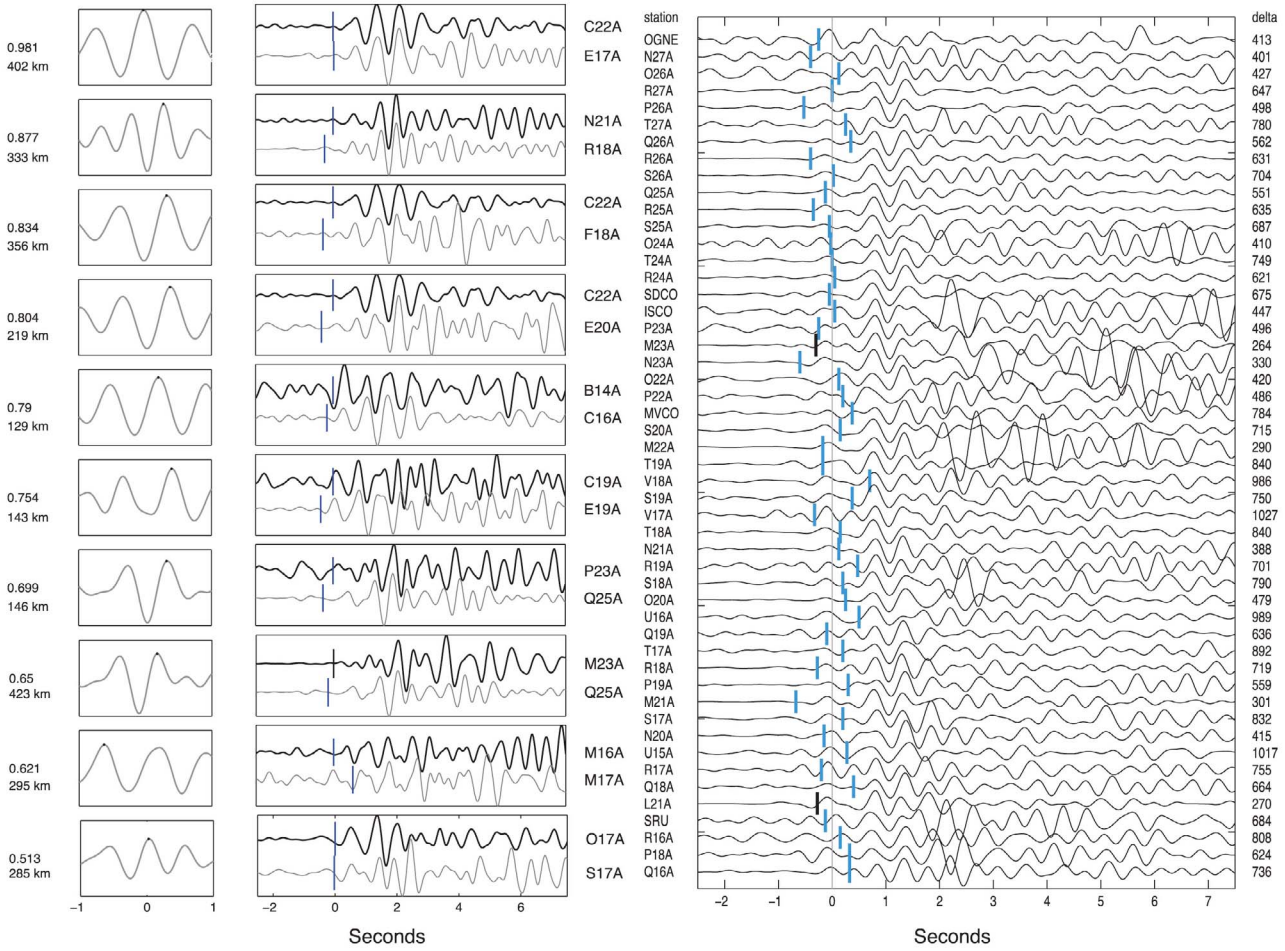

Figure 2. Sample cross-correlations. The left two columns contains the correlation functions and the corresponding cross-correlated and shifted seismograms. The correlation coefficients and inter-station distances are noted in the far left. The tick marks indicate the picks. The right column shows a selection of the realigned seismograms. The realigned times are obtained by fitting the measured differential times from cross-correlation. The amplitudes are scaled according to energy in the correlation window for better comparison. In this figure all but two of the picks are from the auto-picker (blue ticks).

We discard stations that correlate with a mean correlation coefficient of less than 0.6. This approach produces about 107,000 re-aligned picks (73,000 ANF picks and 34,000 auto-picks). The differential times obtained from cross-correlation and also the realigned traces can directly be used in the subsequent wavefront fitting analysis.

3. Localization

[17] The closely spaced stations of USArray allow us to experiment with local array methods to regionally image velocity structure in the uppermost mantle. We use a variety of approaches based on local plane- and circular-wavefront fitting techniques [e.g., *Almendros et al.*, 1999; *Pezzo and Giudicepietro*, 2002] that we apply to station subsets of USArray. These local methods do not require the regularization (i.e., damping and smoothing) used to stabilize the typically ill-posed inversion problem of traditional travel-time tomography, and thus should provide better lateral resolution and velocity anomaly amplitude retrieval.

[18] The localization approach assumes that the Pn wavefront is locally coherent across station subarrays. If we knew the exact hypocenter location and assume that lateral velocity variations are very small, we could simply compute the great circle path between stations and events and use the differential times to solve directly for the apparent velocity of the passing wavefront. More likely, velocity variations between source and receiver will bend the raypaths, in which case more accurate results will be obtained by simultaneously solving for the direction and apparent velocity of the incoming wavefronts. Assuming a flat Moho and a uniform crust, the measured apparent velocity would still correspond to the true medium velocity in the uppermost mantle. Again, however, this is probably not a reasonable assumption, and just as in the case of traditional Pn tomography, station time terms should be applied to take into account the effects of crustal velocity and thickness variations. In addition, azimuthal anisotropy in the upper mantle may lead to varying velocity observations for events from different directions. For regions with good azimuthal coverage, we can estimate this anisotropy by fitting the azimuthal distribution of

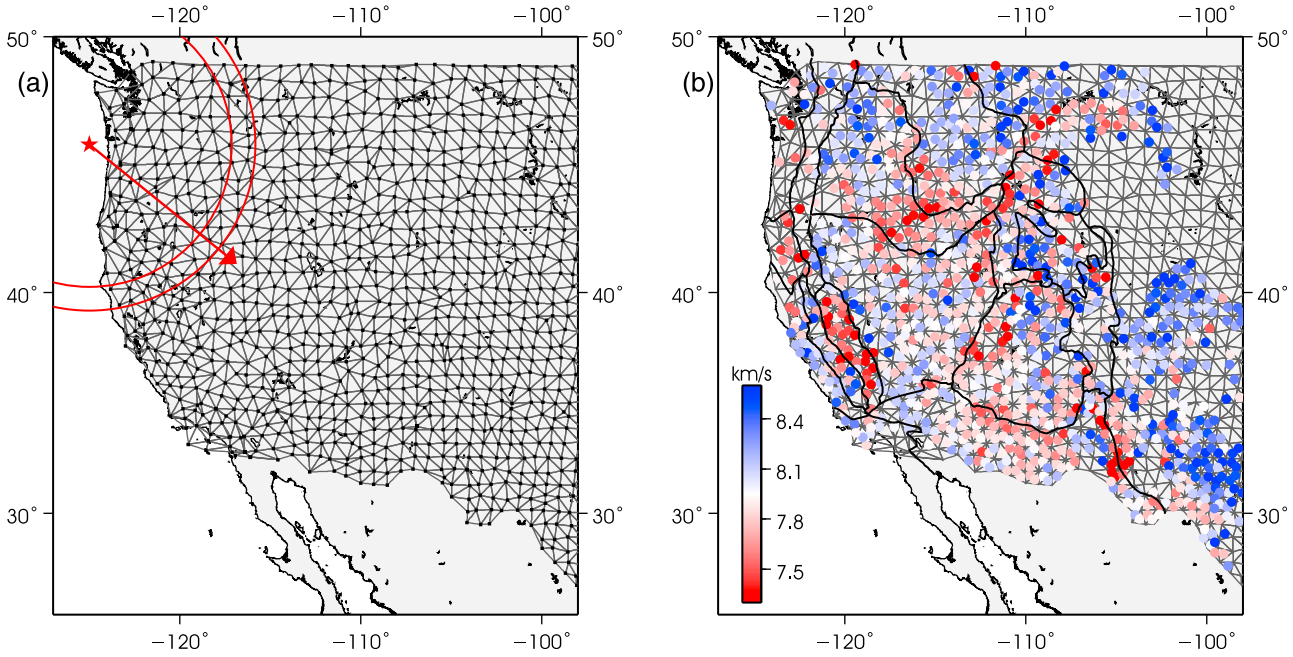

Figure 3. (a) USArray stations divided into triangles with a Delaunay tesselation (black lines). At each triangle, we solve for the direction and velocity of incoming wavefronts. (b) The colored circles show the mean measured velocity at each triangle. For some triangles, there are not enough good differential times to obtain a velocity measurement, especially in regions east of 110°W. No circle is plotted in triangles with fewer than five velocity measurements.

apparent velocities with cos 2θ curves. In the sections that follow, we will start with simple wavefront fitting, and then add in the complexities of station terms, Moho slope estimates, and anisotropy.

3.1. Station Triangularization and Wavefront Fitting

[19] The smallest array subset to estimate directions and velocities of incoming plane waves contains three stations. To obtain an even mesh of station triangles, we apply a Delaunay triangularization to all the USArray station locations (Figure 3a). We then solve for the direction and apparent slowness of incoming plane wavefronts for each earthquake and station triangle directly using inter-station differential times. To estimate the parameters for a plane wave, we solve $\tau_{ij} = \mathbf{p} \cdot \mathbf{x}_{ij}$, where τ_{ij} is the data vector with the differential times, $\mathbf{x}$ contains the station spacing in x and y (east and north), and $\mathbf{p}$ the horizontal slowness components.

[20] Inter-station differential times can be directly obtained from the cross-correlations (section 2.2), or can be calculated from the set of realigned arrival times (section 2.3). The advantage of using realigned times is their self-consistency, and in ensuring that large timing errors in individual cross-correlation measurements (as for example from cycle skipping) do not influence the wavefront fitting. In addition, the same set of realigned times can be used in traditional time-term tomography, which allows for detailed comparisons between the two methods. Alternatively, one can simply use the differential times from cross-correlation directly, and for good correlations these raw inter-station times may be the most exact. Hence we start with the raw differential times from cross-correlation (section 2.2), since this is the most straightforward approach to apply the localization method to a new data set.

[21] The aperture of a USArray station triangle (mean station separation of ~70 km) spans ~30 degrees for arrivals close to the cross-over distance of 180 km, and a plane wave simplification is not valid for these arrivals. We begin by fitting for plane waves if the distance between the earthquake and triangle center is larger than 500 km, and for circular wavefronts using the catalog epicenter location for closer ranges (the wavefront looks approximately like a circle at the surface of the Earth as it sweeps across the array of stations. In this case $\mathbf{x}$ simply contains the differences in epicentral distance, and we solve for the apparent slowness). We use the set of three differential times only if the mean correlation coefficient is greater than 0.7. For each station triangle we obtain several velocity measurements from different earthquakes at various back-azimuths. Figure 3b shows the mean velocity for each tripartite array, plotted at the center of each triangle. No circle is plotted in triangles associated with fewer than five velocity measurements.

[22] Because good cross-correlation results are not obtained for every station-event pair, a fixed triangularization will discard some data because usable traces are often scattered across subsets of the array. This creates the gaps seen in Figure 3 where some triangles do not show a measurement. To maximize usable data and to obtain a more continuous mesh of velocity measurements, we next apply the triangularization separately to all the recording stations of an earthquake, allowing the largest triangle leg to be 150 km. We then again solve for the incoming wavefronts for all the triangles whose differential times satisfy the

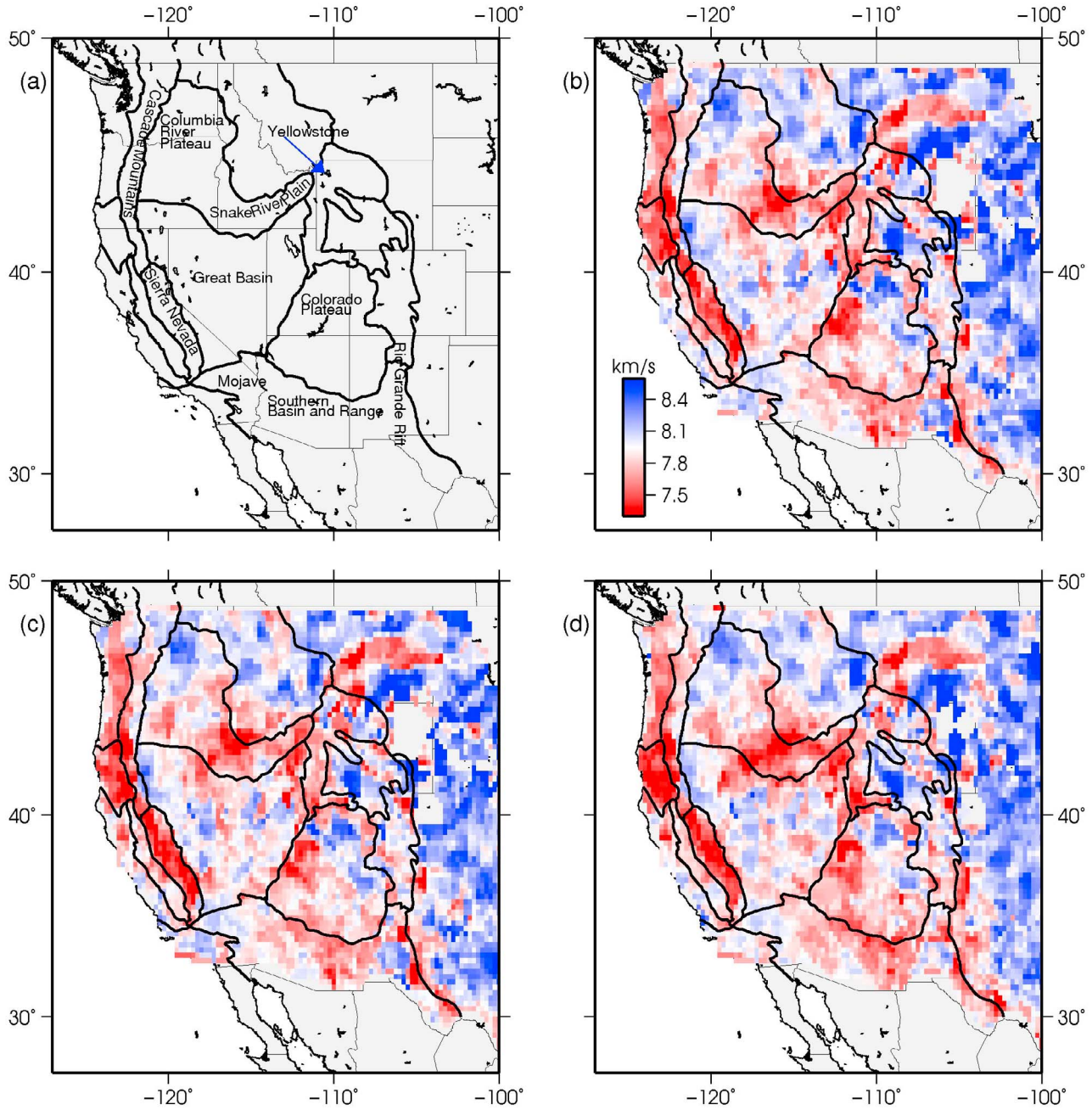

Figure 4. Localization approach with station triangles. Delaunay triangularization is applied separately to every event data set, and parameter estimates are stored in an underlying two-dimensional 0.25° × 0.25° grid. Plotted are the average values at each location. (a) Overview map with geological provinces. (b) Localization result assuming a planar wavefront, except at ranges smaller than 500 km where we fit a circular wavefront from a fixed epicenter. (c) Accounting for curvature with circular wavefront fits. (d) Same as Figure 4c but with self-consistent realigned relative arrival times.

minimum correlation criteria. We put a two-dimensional 0.25° × 0.25° grid over the model area to accommodate the results of the various triangularizations and we assign the measured velocity for a tripartite array to all the regular grid points that lay within the triangle. The mantle piercing point of the head wave is not located directly below the station triangle, but in practice we have found that this offset introduces little bias. Figure 4b shows the mean value per grid point after we solved for the velocity parameters of the triangle-event combinations.

[23] To avoid errors from the plane wave approximation and fixed hypocenter locations, we can account for the curvature of the wavefront [e.g., *Almendros et al.*, 1999]). Using the predetermined hypocenter locations in the catalog keeps the problem simple and linear. But ideally, as the hypocenter locations might be inaccurate or lateral velocity

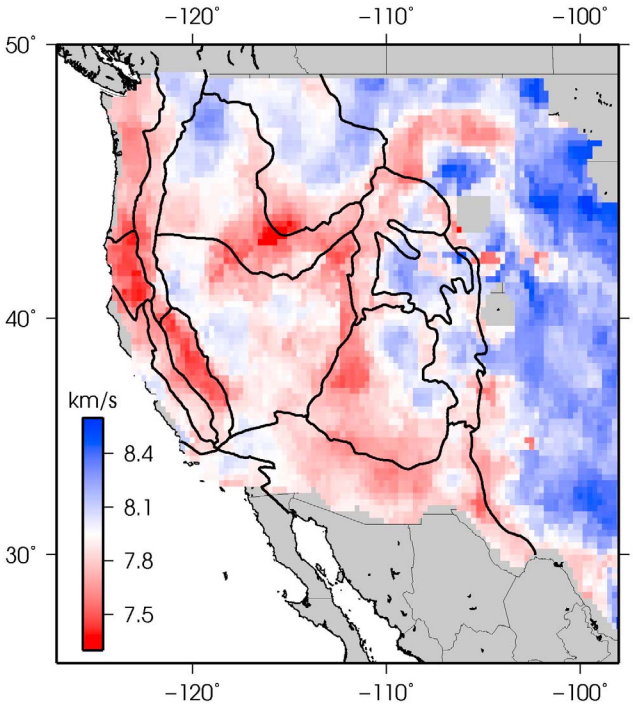

Figure 5. Local velocity estimates from larger sub-array polygons that contain at least three station triangles. This provides a smoother image with similar large-scale features as obtained with tripartite arrays.

variations may cause the direction of the incoming plane wave to deviate from the great circle path, we want to simultaneously solve for the curvature, apparent velocity, and azimuth of the passing wavefront. To take this into account, we solve for apparent slowness S and origin coordinates lat_o and lon_o of the non-linear equation $\tau_{ij} = S(\delta_i - \delta j)$ where δ_i is the great circle distance between station and event.

[24] Although we measure three differential times for each tripartite array, only two of them are independent. However, our problem now consists of three unknowns, generating an underdetermined system with non-unique solutions. To stabilize the approach, we search for a solution close to the reported catalog epicenter location. As described in the next section, we also experiment with larger station subsets. However, in general, these more complicated approaches provide very similar results to the simple combined plane- and spherical wavefront method (Figure 4c). Thus, it does not appear that epicentral errors or ray bending effects due to lateral velocity variations are very important in our local velocity estimates.

[25] As mentioned in the starting paragraph of this section, another approach is to analyze the self-consistent realigned times obtained over all stations for each event (described in section 2.3), rather than the raw differential times from the cross-correlation. This is our preferred approach to compare maps from localization imaging to traditional travel-time tomography as the same data set can be used for both. In addition, for a generally low signal-to-noise data set, or for locations with only few arrivals, this likely provides the more accurate localization map since individual correlation measurements are error-prone. In this case, we use the three absolute times t_k of a station triangle to solve for the arrival time at the origin (center of the triangle) t_0, and the horizontal slowness components: $t_k = t_0 + \mathbf{p} \cdot \mathbf{x}_k$. This provides slightly smoother varying velocity maps (Figure 4d) compared to the ones with cross-correlated differential times but the difference is small and all the images show consistent velocity structure. We will later compare this image to that obtained using the same set of self-consistent times in traditional Pn tomography.

3.2. Larger Sub-Arrays and Station Time Terms

[26] Here we experiment with using larger station subsets, which should provide more stable, albeit lower resolution, velocity estimates. Residuals of this now overdetermined system allow us to introduce station time terms as described in the next paragraph. To obtain station polygons, we first identify all triangles associated with each station. If a station is part of at least three triangles (i.e., sub-arrays of at least five stations), and the mean correlation coefficient for the differential times of this group of receivers is larger than 0.7, we proceed to fit for the incoming wavefront. The obtained slowness and azimuth information is assigned to all the grid points within the polygon spanned by the three or more triangles. Figure 5 shows the resulting mean measured apparent velocity at each location, which, as expected, resembles a smoothed version of Figure 4.

[27] To this point we have not considered changes in station elevation or underlying crustal structure. Crustal velocity structure and crustal thickness differences below individual stations will contribute to the measured differential times. To account for delays introduced by these effects, we can include station time terms just as in the classic Pn time-term approach [*Hearn*, 1996]. We define the travel-time difference between two stations as $\tau_{ij} = S(\delta_i - \delta j) + (\xi_i - \xi_j)$, where ξ_i denote the station corrections and δ_i the great circle distances. The station terms are initially set to zero. We then solve for the slowness of incoming wavefronts for all the events and sub-arrays, keeping track of the residuals after each fit $res_{ij} = \delta\xi_{ij} = \tau_{ij} - S(\delta_i - \delta j)$, where res_{ij} represents the residual for the differential time between station i and event j. After one run through all the events, the residuals are put into a linear system to solve for individual station statics ξ using least squares by minimizing $F = \sum(\delta\xi_{ij} - (\xi_i - \xi_j))^2$. The mean station term is set to zero.

[28] In subsequent iterations, the difference in station time terms is subtracted from the observed differential times, and wavefronts are fitted to the corrected τ_{ij}. After about ten iterations, the change in time terms is small enough to consider the system converged. For stability, we start with the fixed epicenter method, solve for the apparent slowness only, and iterate for the station terms. These station statics are then used as starting values for the more sophisticated method with more model parameters where the wavefront curvature is allowed to vary.

[29] The station time terms and velocity estimates with corrected times are illustrated in Figure 6. Large station term values (e.g., stations in the Sierra Nevada) can reduce the size of the associated velocity anomaly. Time terms in the Snake River Plain are close to zero, not influencing the mapped velocity structure. However, the station delays are large at receivers above the current location of the Yellowstone hot spot, and the mean apparent velocity is increased with the site-term corrected differential times.

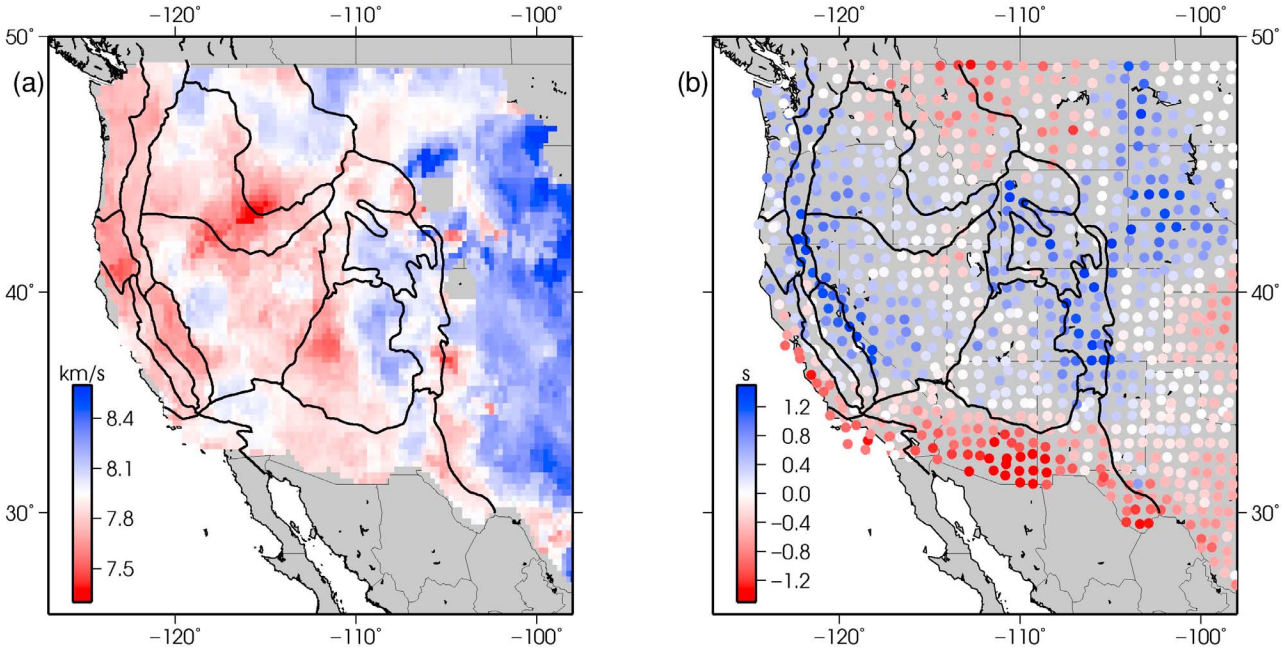

Figure 6. (a) Local velocity estimates from larger sub-array polygons that contain at least three station triangles with differential times that are corrected for station delays. (b) Station time terms in seconds. Station terms might absorb crustal velocity anomalies below stations and changes in crustal thickness. Compare these terms to the station time terms obtained from traditional Pn tomography (Figure 11b).

3.3. Anisotropy and Moho Slope

[30] Station terms can correct for mean crustal thickness variations beneath stations, but cannot account for the azimuthally-dependent time delays produced by a dipping Moho under a station or by azimuthal anisotropy in the uppermost mantle (Figure 7). Because Moho topography and mantle anisotropy vary across the western United States, it is important to consider these effects. We do this by fitting the azimuthal distribution of apparent velocities of triangles with good azimuthal coverage to the equation $Vapp = A + [B \sin\phi + C \cos\phi] + [D \sin 2\phi + E \cos 2\phi]$, where A is the isotropic velocity, the sin φ and cos ϕ terms approximate the

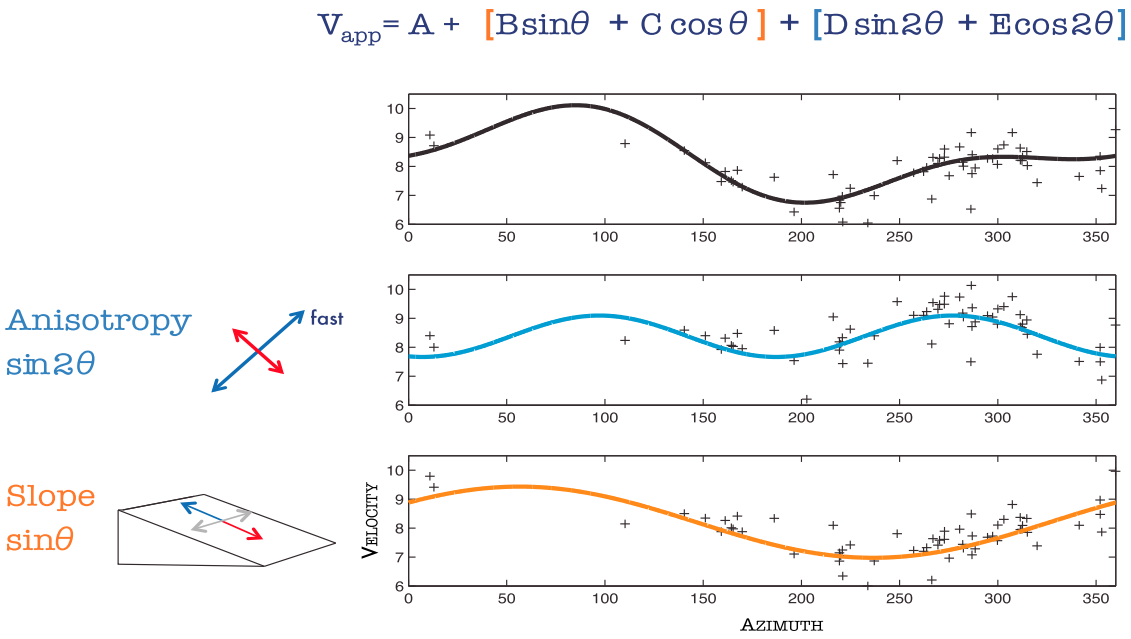

Figure 7. Example of using azimuthal velocity variations to solve for Moho slope and azimuthal anisotropy. Moho slope and anisotropy influence the observed apparent velocity. We fit curves described by the equation on the top to the measured velocity distribution with azimuth to take these effects into account.

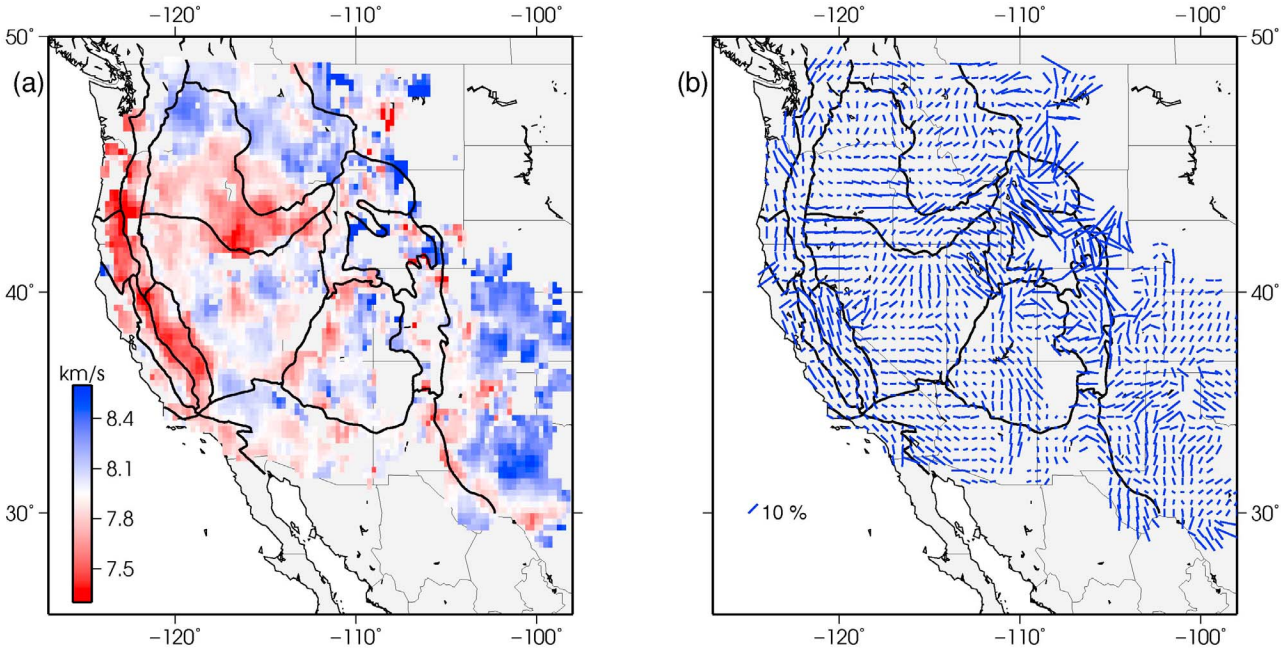

Figure 8. Moho slope and anisotropy influence the observed apparent velocity. We fit curves to the measured velocity distribution with azimuth to take these effects into account. (a) The isotropic velocities after accounting for Moho slope and anisotropy, and (b) the fast directions and strength of anisotropy (scaling with line length). Velocities and azimuths are measured using triangular station subsets, and without the addition of station terms.

effects of Moho slope, and the sin 2ϕ and cos 2ϕ terms describe the azimuthal anisotropy of Pn velocity. We do not include 4ϕ terms since observations and models of azimuthal anisotropy in the uppermost mantle have shown a dominant 2ϕ variation [*Raitt et al.*, 1969; *Christensen*, 1984].

[31] For the curve fitting we again apply an iterative re-weighted least squares approach since the velocity measurements typically show large scatter. To obtain stable results, we only consider grid locations with good azimuthal event coverage, and require that the largest gap in direction of the arrival is smaller than 100°. For each grid point, we use the velocity and azimuth measurements stored at that point, as well as the estimates stored for the surrounding eight nodes. Figure 8 shows the estimated fast directions and magnitude of the anisotropy ($(V_{fast} - V_{slow})/V_{iso}$) and corresponding isotropic velocities for the triangle fits. Figure 9 shows the same parameters for the larger sub-arrays measurements with station-term-corrected differential times. For mapping clarity, we plot the average fast direction in 0.5 degree cells. We observe mostly fault-parallel Pn fast axes in central California, and fairly homogeneous northeast fast axes in the northwestern Basin and Range that rotate to north–south in eastern Nevada. Note how the inclusion of anisotropic parameters seems to show some of the velocity anomalies more clearly, as for example the south–west to north–east striking low velocity anomaly in eastern New Mexico, east of the Rio Grande Rift, near the axis of the Jemez Lineament [e.g., *Karlstrom and Humphreys*, 1998].

[32] Moho topography can be calculated from the station time terms and compared to the localized slope estimates. We find that these results only agree in regions with likely large topography and good azimuthal coverage, as for example below the Sierra Nevada. However, changes in crustal velocity below a station can produce a similar signal as Moho slope. In the analysis for Figure 8, where we used station triangles, we did not include station time terms, but for Figure 9, with station pentagons, we solved for time terms prior to the curve fitting. In the first case, it is possible that Moho slope terms absorb delays introduced by varying crustal velocities. In the latter case, the time terms already partially account for changes in crustal thickness. It is difficult to resolve the trade-offs between the various parameters, and some experimentation with varying model complexity seems necessary.

4. Traditional Travel Time Tomography

[33] For the localization approach, we typically require a minimum correlation coefficient for the differential times among the selected stations. Each measurement is carried out independently, therefore we just use all the available data that pass the cut-off. To use the globally adjusted times in traditional travel-time tomography, we ideally want a well-distributed subset of events to provide uniform data coverage over the study area. Uneven ray coverage can introduce artifacts into velocity maps, as for example smearing along prominent raypaths. To provide more uniform coverage, we apply a modified version of the composite event method of *Lin et al.* [2007], which has the further advantage of reducing the size of the inverse problem.

[34] We first window the realigned picks to remove large outliers after straight-line fits and then remove any events that are recorded at fewer than ten stations. Next, we perform the following steps, following the general approach described in

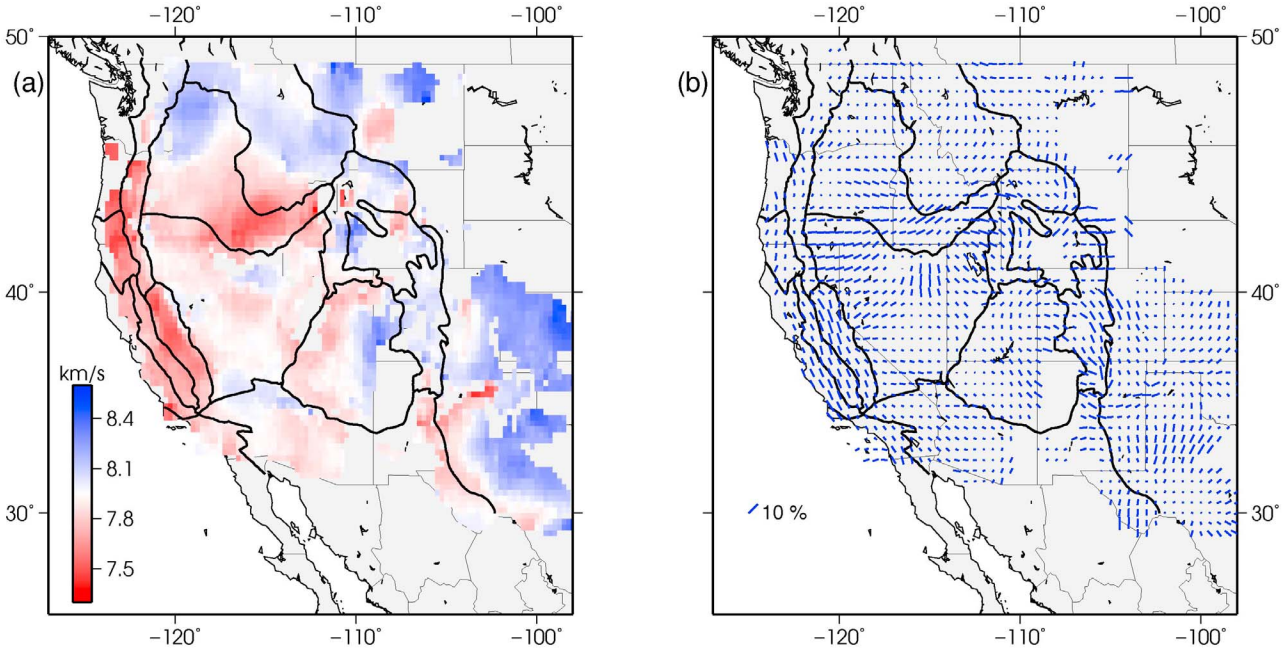

Figure 9. (a) The isotropic velocities after accounting for Moho slope and anisotropy, and (b) the fast directions and strength of anisotropy (scaling with line length). Velocities and azimuths are measured using station pentagons with five or more receivers, and the differential times were corrected for station delays.

Lin et al. [2007]: We find the event with the most picks and keep track of the recording stations. We then find all the events within a radius R of this reference event (here we use $R = 20$ km), and sort these neighboring events in descending order by their number of picks. Because USArray stations do not stay in place but move approximately every two years, two events at the same location might be recorded by different sets of stations. For this reason, we check the number of newly activated stations for each subsequent event within radius R of the target event, starting with the second best recorded (by number of picks) earthquake within the group. We keep an event within the same radius if the arrival set includes more than 30% new stations. Then we move on to the text target event and so on. Contrary to the composite event method of *Lin et al.* [2007] we do not average the arrival-time residuals, but try to select the best-recorded events in each region. The final set of events contains about 20% of the original earthquakes, but is much more geographically uniform (Figure 10).

[35] Next, we invert these times following the tomographic approach of *Hearn* [1996] and further described in *Buehler and Shearer* [2010], which applies a modified time-term method to the large set of residuals. Using a two-dimensional grid along the Moho, the travel-time residuals are parameterized as $\delta t_{es} = \delta\tau_e + \delta\tau_s + \sum\Delta_{esk}(\delta S_k + A_k \cos 2\phi_{esk} + B_k \sin 2\phi_{esk})$, where Δ_{esk} is the distance the ray travels in the cell k, δS_k is the isotropic slowness perturbation in cell k, A_k and B_k are the anisotropic parameters for cell k, ϕ is the back azimuth, and $\delta\tau_e$ and $\delta\tau_s$ are the event and station time terms. As in the localization approach, we do not include any 4ϕ velocity variations. Here the event time term $\delta\tau_e$ absorbs not only the time the ray spends in the crust, but also errors in hypocenter locations and absolute timing (since we only have accurate relative times). Variance reduction is about 70% compared to a simple straight-line fit (uniform Pn velocity, no station or event terms), and 30% compared to a one-dimensional time-term tomography (uniform Pn velocity with station and event terms). We find that variance reduction is not much influenced by the use of the realigned times versus the raw analyst and auto-picker picks. Hence the cross-correlation likely does not improve the timing accuracy much for this particular generally low signal-to-noise data set (or the model parameterization is not good enough to capture these small time changes), even though the waveforms appear visually more aligned. Still, we prefer to work with the re-aligned times in the tomography, as this provides us with a consistent and completely objective data set.

[36] The resulting isotropic velocity variations are shown in Figure 11 and largely agree with the locally imaged velocity structure, especially when comparing them with the localization maps from the larger sub-arrays. Prominent features include the distinct slow velocity anomalies below the Sierra Nevada and the Snake River Plain. The Yellowstone hot spot track is bounded to the north by a large positive velocity anomaly spanning eastern Washington, northern Idaho and western Montana. The localization and tomography provide consistent fast directions (Figure 12); the largest difference is observed in the magnitude of the anisotropy. Synthetic tests in *Buehler and Shearer* [2010] showed that the anisotropy magnitude is generally less well retrieved than the fast axis with the traditional travel-time tomography. In California we consistently observe mostly fault-parallel Pn fast axes, indicating that the anisotropy is caused by shearing along the plate boundary. Both methods show a rotation of the fast axis around the great

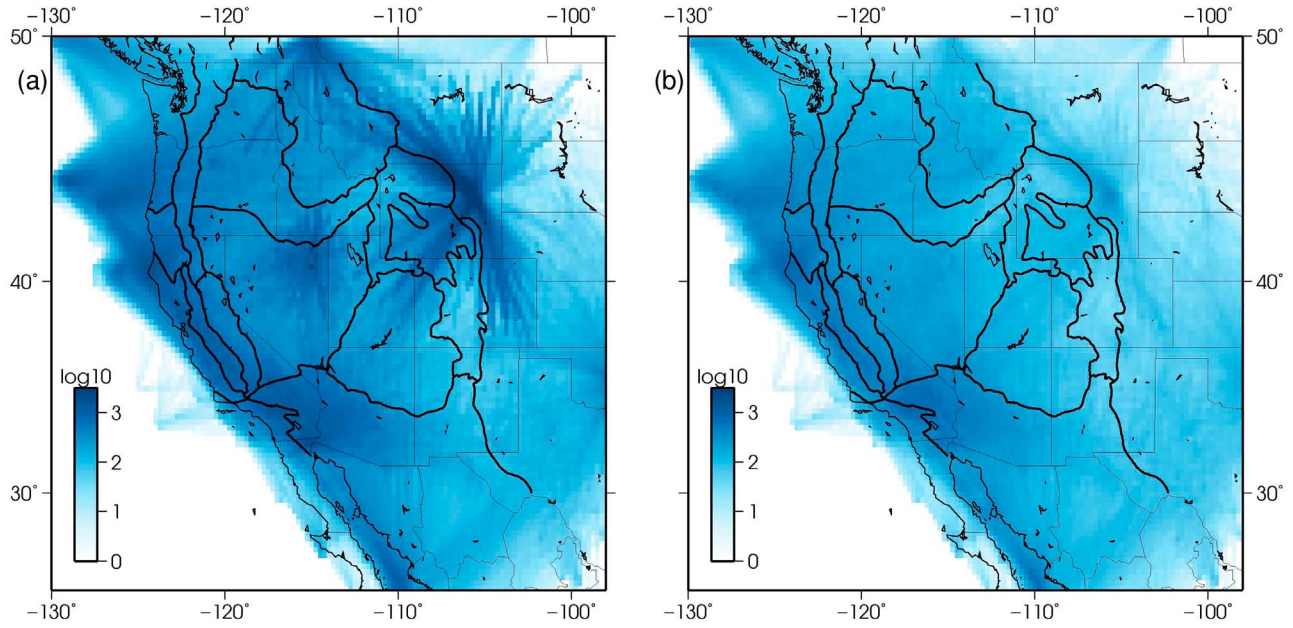

Figure 10. Ray count per cell for the traditional Pn tomography: (a) with the whole data set and (b) after removing data on much traveled paths.

Basin, taking a sharp corner in northeastern Nevada at the southern edge of the Snake River Plain.

5. Discussion

[37] Consistent large-scale features are seen in the localized images as well as the traditional Pn tomography. These include the slow Pn velocities below the Snake River Plain, the Sierra Nevada, and the western edge of the Colorado Plateau. As expected, the localization method produces more detailed images of the uppermost mantle velocity structure, especially when the fitting is performed with small triangular sub-arrays. For example, we observe velocities larger than

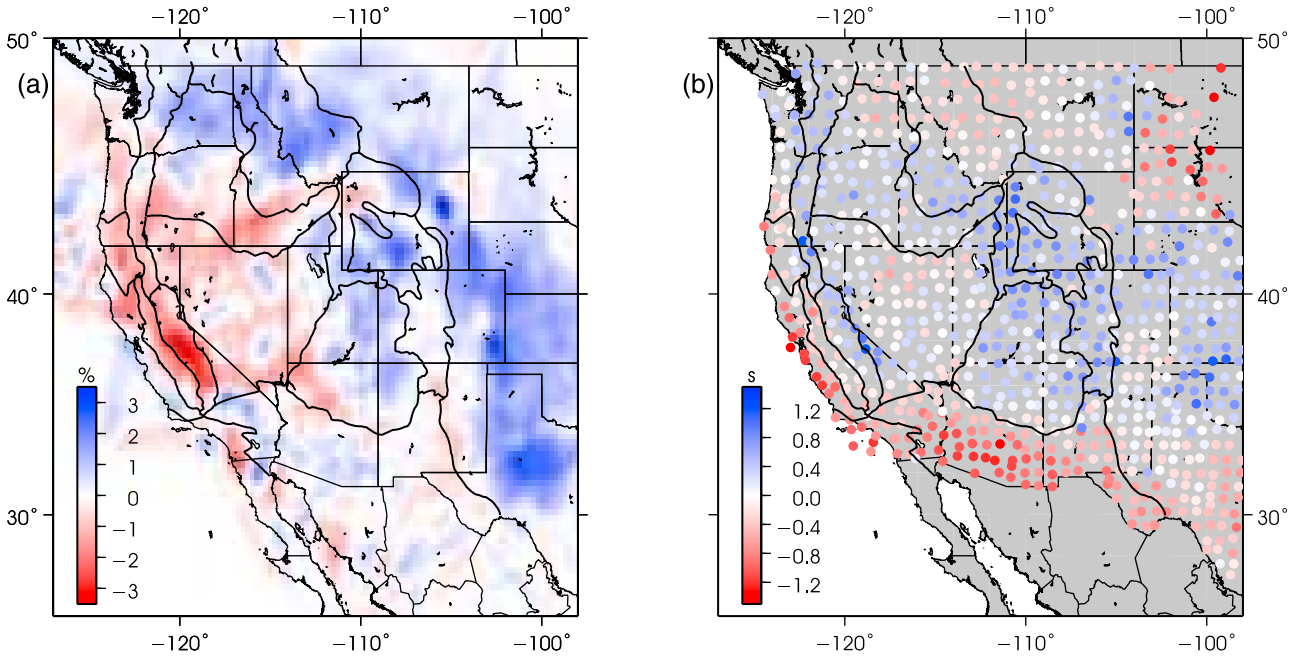

Figure 11. (a) Isotropic velocities obtained from two dimensional Pn travel time tomography using realigned picks. The average velocity is 7.91 km/s. (b) Station time terms in seconds. The station terms should account for the time it takes the ray to travel from the mantle pierce point to the receiver. With the assumption of a crustal velocity, Moho depths can be estimated. Positive station terms indicate a thicker crust, negative terms a thinner crust.

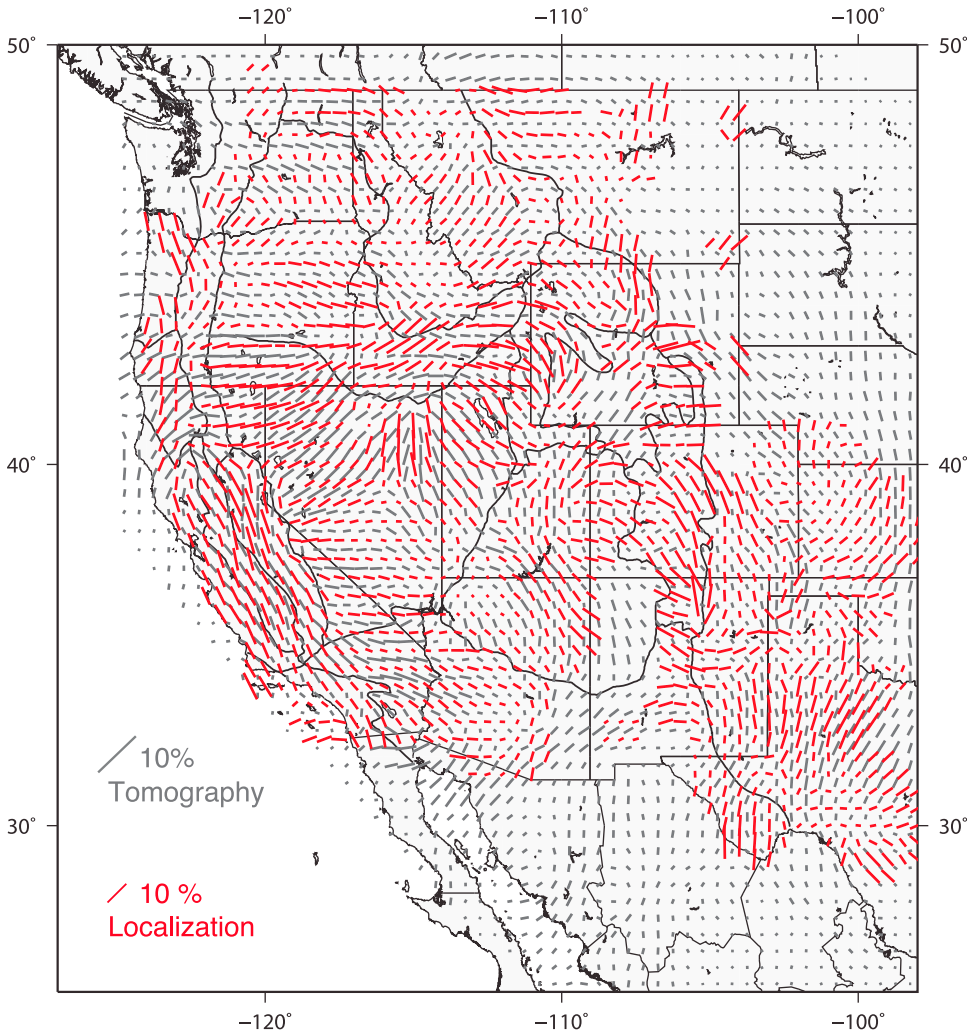

Figure 12. Pn fast axis obtained from localization with station pentagons (red) and traditional travel time tomography (gray). The traditional travel-time tomography results generally show weaker anisotropy. For visibility, the fast axes from the tomography are multiplied by a factor of two.

8 km/s just east of the Sierra Nevada in Figure 4, compared to the more muted changes in the tomography image (Figure 11). The localization images show overall higher-amplitude anomalies. It is likely that the travel-time tomography underestimates the amplitude of velocity structures because of damping in the inversion.

[38] Plane and circular wave-front methods show very similar results, the small changes in individual measurements do not appear in maps with averaged velocities because of generally large scatter in the measured apparent velocities. We found that for many traces in our data set the Pn arrivals are hard to pick and cross-correlate due to their low signal-to-noise ratios. The inclusion of wavefront curvature may have a larger effect in different regions, perhaps in areas with more impulsive Pn arrivals that would lead to more accurate timing. The addition of terms that account for Moho topography and anisotropy have a larger influence on the mapped velocity anomalies. As expected, localization images resulting from larger sub-arrays that average over larger regions agree better with the smoothed traditional tomography images.

[39] Consistent with the older Pn tomography study by *Hearn* [1996] and our newer version with realigned USArray data, the localization results show a large low-velocity region below the Snake River Plain, but the anomaly is strongest west of the current location of the Yellowstone hot spot. Contrary to new P-velocity map slices in the lithosphere [*Schmandt and Humphreys*, 2010; *Obrebski et al.*, 2011; *Tian and Zhao*, 2012], the slow velocities are not concentrated around the hot spot location, but spread over the Snake River Plain. This is observed even in our localization maps from triangular sub-arrays. The traditional tomography approach shows an anomaly that fades somewhat toward the northeast. However this could be an artifact of reduced ray coverage below Yellowstone, or off-great circle-path arrivals, as the localization maps show consistent slow velocities. It would be interesting to apply 2D ray tracing to model the deviations of the directions of the incoming wavefronts due to changes in velocity structure, information that could be compared to the ray angles obtained from the localization method and possibly also

included to improve ray tracing in travel-time tomography, but we defer this to planned future work.

[40] We observe consistent very low mean Pn velocities below the Sierra Nevada and azimuthal anisotropy with fault-parallel fast directions. As discussed in *Hearn* [1999], velocities much lower than 7.8 km/s may be associated with mantle material that contains a small amount of partial melt. Several factors, however, can influence the measured apparent velocities, for example an increased crustal thickness because of a mantle root, a dipping Moho, cross-correlation alignments on refractions from a shallower layer, multipathing, or a combination of these effects. It seems inevitable that there are trade-offs between the various parameters, and it is difficult to distinguish between geometrical, compositional, and thermal effects.

[41] *Jones et al.* [1994] found slow Pn velocities (7.6–7.65 km/s) below the southern Sierra Nevada, and discuss the possibility of a two-layer structure that includes an additional refractor with a P-wave velocity greater than about 7.2 km/s. *Savage et al.* [1994] found further evidence for this intermediate layer. The subtraction of station delays from the differential times increases the apparent velocity below the Sierra Nevada in our maps (compare Figures 5 and 6) to about 7.6 km/s as well. Station terms might not only absorb crustal velocity anomalies below stations, but also changes in crustal thickness. Both the delay terms from traditional tomography and the localization site terms are largest at stations located at the eastern border of the mountain range. Further analysis that could include Pn and Sn amplitude measurements to locally resolve attenuation structure (but working with these amplitudes is difficult, see *Bakir and Nowack* [2011] for the complex behavior of Pn amplitudes), as well as combined Pn/Sn travel time inversions to estimate Vp/Vs ratios, might help constrain material and temperature properties.

[42] Localization anisotropy measurements largely agree with fast-direction estimates from traditional Pn tomography, in particular the mostly fault-parallel Pn fast axis in central California, indicating that the anisotropy is caused by shearing along the plate boundary. As discussed in greater detail in *Buehler and Shearer* [2010], these fast directions are complementary to fast polarization directions obtained in recent USArray shear wave splitting studies [*Liu*, 2009; *West et al.*, 2009]. The SKS fast polarization directions show a distinct cylindrical pattern around small splitting times in the central Great Basin. In addition, USArray P-wave tomographic models show an almost vertical high velocity anomaly in the region with small splitting times [*Roth et al.*, 2008] ranging from a depth of about 75 km to at least 500 km that can be interpreted as a result of a lithospheric drip [*West et al.*, 2009]. Figures 8 and 9 of this study suggest low velocities in central Nevada in the uppermost mantle, bounding the vertical high velocity anomaly imaged by *Roth et al.* [2008]. Pn fast axes from both tomography and localization show a less distinct rotation, and anisotropy is stronger in northern than southern Nevada. This observation suggests that the orientation of azimuthal anisotropy changes in the upper mantle as shear wave splitting measurements provide a vertically integrated measure of anisotropy in the upper mantle. We plan to apply similar methods to Sn-SV and Sn-SH to solve for radial and azimuthal anisotropic parameters. This would allow us to compare our observations to seismic anisotropy predicted by aligned olivine models that are believed to be the dominant cause of upper-mantle anisotropy.

[43] In conclusion, we demonstrate an alternative method to traditional Pn travel-time tomography to regionally image uppermost-mantle velocity structure and anisotropy. The dense station network of USArray provides a chance to apply simple array tools to regional scale investigations. The localization approach does not depend on regularization, and therefore has the capacity to map velocity structure with increased resolution and with more reliable amplitudes. Our localization maps generally show stronger anomalies compared to travel-time tomography images, which is important to consider, for example, in mantle temperature calculations. Similar localized approaches can be applied to Sn, and critically or nearly critically refracted crustal phases. This analysis should provide a more complete picture of crust and uppermost mantle structure at well-defined depths.

[44] **Acknowledgments.** We thank the USArray Array Network Facility for making their picking information available. Luciana Astiz assisted with many ANF database related questions. Many of the figures were produced with the GMT software [*Wessel and Smith*, 1995]. We thank Tom Hearn and an anonymous referee for constructive reviews of this manuscript. This research was supported by grant EAR-0950391 from the National Science Foundation.

References

Almendros, J., J. M. Ibanez, G. Alguacil, and E. Del Pezzo (1999), Array analysis using circular-wave-front geometry: An application to locate the nearby seismo-volcanic source, *Geophys. J. Int.*, *136*, 159–170.

Bakir, A. C., and R. L. Nowack (2011), Modeling seismic attributes of Pn waves using the spectral-element method, *Pure Appl. Geophys.*, *169*(9), 1539–1556.

Buehler, J. S., and P. M. Shearer (2010), Pn tomography of the western United States using USArray, *J. Geophys. Res.*, *115*, B09315, doi:10.1029/2009JB006874.

Burdick, S., et al. (2008), Upper mantle heterogeneity beneath North America from travel time tomography with Global and USArray transportable array data, *Seismol. Res. Lett.*, *79*, 384–392.

Burdick, S., et al. (2010), Model update January 2010: Upper mantle heterogeneity beneath North America from traveltime tomography with global and USArray transportable array data, *Seismol. Res. Lett.*, *81*, 689–693.

Burdick, S., et al. (2012), Model update March 2011: Upper mantle heterogeneity beneath North America from traveltime tomography with global and USArray transportable array data, *Seismol. Res. Lett.*, *83*, 23–28.

Christensen, N. I. (1984), The magnitude, symmetry and origin of upper mantle anisotropy based on fabric analyses of ultramafic tectonites, *Geophys. J. R. Astron. Soc.*, *76*, 89–111.

Earle, P. S., and P. M. Shearer (1994), Characterization of global seismograms using an automatic-picking algorithm, *Bull. Seismol. Soc. Am.*, *84*, 366–376.

Goes, S., and S. van der Lee (2002), Thermal structure of the North American uppermost mantle inferred from seismic tomography, *J. Geophys. Res.*, *107*(B3), 2050, doi:10.1029/2000JB000049.

Hearn, T. M. (1996), Anisotropic Pn tomography in the western United States, *J. Geophys. Res.*, *101*, 8403–8414.

Hearn, T. M. (1999), Uppermost mantle velocities and anisotropy beneath Europe, *J. Geophys. Res.*, *104*, 15,123–15,139.

Jones, C. H., H. Kanamori, and S. W. Roecker (1994), Missing roots and mantle and drips: Regional Pn and teleseismic arrival times in the southern Sierra Nevada and vicinity, California, *J. Geophys. Res.*, *99*, 4567–4601.

Karlstrom, K. E., and E. D. Humphreys (1998), Persistent influence of Proterozoic accretionary boundaries in the tectonic evolution of southwestern North America, *Rocky Mt. Geol.*, *33*, 161–179.

Lin, F.-C., and M. H. Ritzwoller (2011), Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure, *Geophys. J. Int.*, *186*, 1104–1120.

Lin, F.-C., M. H. Ritzwoller, and R. Snieder (2009), Eikonal tomography: Surface wave tomography by phase front tracking across a regional broad-band seismic array, *Geophys. J. Int.*, *177*, 1091–1110.

Lin, G., P. M. Shearer, E. Hauksson, and C. H. Thurber (2007), A three-dimensional crustal seismic velocity model for southern California from

a composite event method, *J. Geophys. Res.*, *112*, B11306, doi:10.1029/2007JB004977.

Liu, K. H. (2009), NA-SWS-1.1: A uniform database of teleseismic shear wave splitting measurements for North America, *Geochem. Geophys. Geosyst.*, *10*, Q05011, doi:10.1029/2009GC002440.

Moschetti, M. P., M. H. Ritzwoller, F.-C. Lin, and Y. Yang (2010), Crustal shear wave velocity structure of the western United States inferred from ambient seismic noise and earthquake data, *J. Geophys. Res.*, *115*, B10306, doi:10.1029/2010JB007448.

Obrebski, M., R. M. Allen, F. Pollitz, and S.-H. Hung (2011), Lithosphere-asthenosphere interaction beneath the western United States from the joint inversion of body-wave traveltimes and surface-wave phase velocities, *Geophys. J. Int.*, *185*, 1003–1021.

Perry, H. K. C., C. Jaupart, J.-C. Mareschal, and N. M. Shapiro (2006), Upper mantle velocity-temperature conversion and composition determined from seismic refraction and heat flow, *J. Geophys. Res.*, *111*, B07301, doi:10.1029/2005JB003921.

Pezzo, E. D., and F. Giudicepietro (2002), Plane wave fitting method for a plane, small aperture, short period seismic array: A MATHCAD program, *Comput. Geosci.*, *28*, 59–64.

Pollitz, F. F., and J. A. Snoke (2010), Rayleigh-wave phase-velocity maps and three-dimensional shear velocity structure of the western US from local non-plane surface wave tomography, *Geophys. J. Int.*, *180*, 1153–1169.

Raitt, R. W., G. G. Shor Jr., T. J. G. Francis, and G. B. Morris (1969), Anisotropy of the Pacific upper mantle, *J. Geophys. Res.*, *74*(12), 3095–3109, doi:10.1029/JB074i012p03095.

Roth, J. B., M. J. Fouch, D. E. James, and R. W. Carlson (2008), Three-dimensional seismic velocity structure of the northwestern United States, *Geophys. Res. Lett.*, *35*, L15304, doi:10.1029/2008GL034669.

Savage, M. K., L. Li, J. P. Eaton, C. H. Jones, and J. N. Brune (1994), Earthquake refraction profiles of the root of the Sierra Nevada, *Tectonics*, *13*, 803–817.

Schmandt, B., and E. Humphreys (2010), Complex subduction and small-scale convection revealed by body-wave tomography of the western United States upper mantle, *Earth Planet. Sci. Lett.*, *297*, 435–445.

Steck, L. K., M. L. Begnaud, S. Phillips, and R. Stead (2011), Tomography of crustal *P* and *S* travel times across the western United States, *J. Geophys. Res.*, *116*, B11304, doi:10.1029/2011JB008260.

Sun, D. Y., and D. Helmberger (2011), Upper-mantle structures beneath USArray derived from waveform complexity, *Geophys. J. Int.*, *184*, 416–438.

Tian, Y., and D. Zhao (2012), *P*-wave tomography of the western United States: Insight into the Yellowstone hotspot and the Juan de Fuca slab, *Phys. Earth Planet. Inter.*, *200–201*, 72–84.

Vandecar, J. C., and R. S. Crosson (1990), Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least-squares, *Bull. Seismol. Soc. Am.*, *80*, 150–169.

Wessel, P., and W. H. F. Smith (1995), New version of the generic mapping tools, *Eos Trans. AGU*, *76*, 329–329.

West, J. D., M. J. Fouch, J. B. Roth, and L. T. Elkins-Tanton (2009), Vertical mantle flow associated with a lithospheric drip beneath the Great Basin, *Nat. Geosci.*, *2*, 439–444.