2014

# Upper Mantle Structure of the Cascades from Full-Wave Ambient Noise Tomography: Evidence for 3D Mantle Upwelling in the Back-Arc

Haiying Gao

Yang Shen
*University of Rhode Island*, y.shen@icloud.com



This is a pre-publication author manuscript of the final, published article.



Citation/Publisher Attribution
Gao, H. & Shen, Y. (2014). Upper mantle structure of the Cascades from full-wave ambient noise tomography: Evidence for 3D mantle upwelling in the back-arc. *Earth and Planetary Science Letters, 390*(15), 222-233. http://dx.doi.org/10.1016/j.epsl.2014.01.012
Available at: http://dx.doi.org/10.1016/j.epsl.2014.01.012

# Upper mantle structure of the Cascades from full-wave ambient noise tomography: Evidence for 3D mantle upwelling in the back-arc

Haiying Gao[a,b,]*, Yang Shen[a]

[a] Graduate School of Oceanography, University of Rhode Island, Narragansett, RI 02882, USA

[b] now at Department of Geosciences, 250 Morrill Science Center, University of Massachusetts Amherst, MA 01003, USA, email: haiyinggao@geo.umass.edu, phone: 413-577-1250

## ABSTRACT

Melt generation and volcanism at subduction zones may result from several possible processes: hydration of the mantle wedge by fluid released from the slab, subduction-induced mantle upwelling beneath the back-arc, and heating of downgoing sediments/oceanic crust atop the slab. Each process predicts a distinctly different spatial pattern of melt generation and can thus be distinguished with high-resolution seismic imaging. Here we construct an upper mantle model of the Pacific Northwest using a full-wave ambient noise tomographic method. Normalized vertical components of continuous seismic records at station pairs are cross-correlated to extract empirical Green's functions at periods of 7-200 s. We simulate wave propagation within the 3D Earth structure using a finite-difference method and calculate sensitivity kernels of Rayleigh waves to

perturbations of Vp and Vs based on the Strain Green's Tensor database. Phase delays are extracted by cross-correlating the observed and synthetic waveforms at multiple frequency bands.

Our tomographic result reveals three separate low shear-wave velocity anomalies along the back-arc in the upper mantle ~200 km east of the Cascade volcanic arc, with the central one being the largest in size and lowest in velocity. These back-arc low-velocity anomalies are spatially correlated with the three arc-volcano clusters. The geometry of the low-velocity volumes relative to the slab and arc is consistent with the pattern of subduction-induced decompressional melting in the back-arc. Their along-strike variation suggests that the large-scale plate-motion-induced flow in the back-arc mantle wedge is modulated by small-scale convection, resulting in a highly 3D process that defines the segmentation of volcanism along the Cascade arc.

Keywords: full-wave ambient noise tomography, the Cascadia subduction zone, low-velocity anomaly, decompressional melting

## 1. Introduction

The mechanisms of melt generation in the upper mantle wedge (Figure 1) have been the focus of numerous studies, as they are fundamental to our understanding of arc volcanism along subduction zones. In general, melt production is positively correlated to water content in arc basalt (Kelley et al., 2006), which supports flux melting by fluid released from the subducting slab (van Keken, 2003). Subduction-induced mantle upwelling and decompressional melting explains existence of nearly anhydrous lavas, as well as low seismic velocities and high attenuation at the back-arc spreading center (Conder et al., 2002; Wiens et al., 2008). Melting of the oceanic crust/sedimentary layer atop the slab also contributes significantly to the trace element signatures at arc volcanoes and the thermal structure of the mantle wedge (Conder, 2005). The melting processes and their relative importance at various subduction zones may depend on the slab age, sediment thickness, subduction rate, and other factors.

Melt generation also varies along strike. At the Honshu subduction zone, northeast Japan, body-wave tomography reveals strong spatial correlation between the along-strike segments of arc volcanoes and the low-velocity anomalies at the back-arc in the upper mantle (named as 'hot fingers' by Tamura et al. (2002)). Numerical experiments with a low-viscosity mantle wedge, presumably due to water released from the slab by dehydration reaction, produce the finger-like small-scale convection (Honda and Yoshida, 2005). To our knowledge, the Honshu subduction zone is the only place where this distinctive hot-finger structure has been imaged and correlated with arc volcanism. Are the hot-finger structure and, by inference, small-scale mantle convection a general phenomenon at subduction zones or unique to those with a possibly hydrated mantle wedge in the back-arc produced by an old and cold slab subducting at a fast rate, such as the western Pacific plate?

With a relatively young and thin slab and presumably shallow dehydration (van Keken et al., 2011), the subduction of the Juan de Fuca plate beneath western North America represents an end member in the subduction zone system. Along strike, the subducting slab, the overriding plate and seismicity show clearly segmented signatures

(Tréhu et al., 1994; Brocher et al., 2003; Burdick et al., 2008). In particular, the Quaternary volcanoes are spatially clustered and different in the composition of primitive basalts (Figure 2, Schmidt et al., 2008). The arc basalts show evidence for both flux and dry melting, with the maximum water content lower than those found at other arcs (Elkins Tanton et al., 2001; Ruscitto et al., 2010). Unlike the subduction zones with well-developed back-arc spreading centers (e.g., Wiens et al., 2008), the Cascades have only volumetrically minor and sparse Quaternary volcanic activities behind the arc (Till et al., 2013). So the extent and form of subduction-induced mantle flow and decompressional melting in the back-arc of the Cascades and those of similar continental subduction zones remain enigmatic.

A well-defined crust and upper mantle velocity model is needed to understand the melting processes and the causes of along-strike segmentation of volcanism in the Cascadia subduction zone. Although there have been many velocity models for the Cascades (Lee and Crosson, 1990; Symons and Crosson, 1997; Parsons et al., 1999; Zhao et al., 2001; Brocher et al., 2001; Calvert et al., 2001; Shapiro and Ritzwoller, 2002; Tréhu et al., 2002; van Wagoner et al., 2002; Crosson et al., 2002; Graindorge et al., 2003; Ramachandran et al., 2004; Burdick et al., 2008, 2010; Roth et al., 2008; Yang et al., 2008; Abers et al., 2009; Audet et al., 2009; Moschetti et al., 2010; Schmandt and Humphreys, 2010; Calkins et al., 2011; Gao et al., 2011; Delorey and Vidale, 2011; Porritt et al., 2011; Wagner et al., 2012; Shen et al., 2013), none of these models covers the entire subduction zone, has high enough resolution at the depth of melt generation, and is adequately accurate to explain the observed waveforms as illustrated by Gao and Shen (2012) with full-wave simulation. In addition, the magnitudes of velocity perturbations in the existing models resolved from body-wave tomography in the Pacific Northwest (Roth et al., 2008; Burdick et al., 2008, 2010; Obrebski et al., 2010, 2011; Schmandt and Humphreys, 2010, 2011; James et al., 2011; Sigloch, 2011) vary within a wide range (Becker, 2012). Furthermore, the resolution gap in the upper mantle between surface-wave tomography and body-wave tomography limits geodynamic interpretations at the depth of melt generation.

In this study, we invert for the upper mantle structure of the Pacific Northwest using an advanced full-wave tomographic method based on simulation of wave propagation within the 3D Earth structure (Zhang et al., 2012). Compared to previous studies, three factors significantly improve the model resolution: First, there has been a huge increase in broadband seismic data in the Pacific Northwest. We have processed ambient noise waveforms from ~1000 stations (see station distribution in Figure 2); Secondly, we have developed a new waveform normalization method (Shen et al., 2012) that improves the quality of surface waves extracted from ambient seismic noise and are able to obtain much more broadband and higher quality Rayleigh waves than in previous studies (Figure 3); and thirdly, the full-wave tomographic method is based on wave propagation simulation in 3D models, a more accurate theory that relates seismic data to the Earth structure (Zhao et al., 2005; Chen et al., 2007; Shen and Zhang, 2010). Unlike previous studies of the Pacific Northwest, we calculate the sensitivities of Rayleigh waves to perturbations of both Vp and Vs and jointly invert for the velocity model. The velocity model is then improved by iteratively reducing the misfit between the observed and synthetic waveforms.

## 2. Data and Methods

The procedure of the full-wave ambient noise tomography includes extraction of empirical Green's functions (EGFs) from continuous ambient noise waveform, finite-difference wave propagation simulation in the 3D Earth structure, measurement of phase delays between observed and synthetic waveforms, calculation of sensitivity kernels and inversion for velocity perturbations. The first three steps are fully described by Gao and Shen (2012), so we only briefly summarize here.

### 2.1. Extraction of empirical Green's functions

To retrieve Rayleigh-wave EGFs between station pairs, we process the vertical component of continuous seismic data recorded between 1995 and 2012 by about 1000 stations in an area extending from northern California to Vancouver Island, Canada (Figure 2). We include seismic stations from the EarthScope USArray Transportable Array (TA), the Canadian National Seismograph Network (CN), the Plate Boundary Observatory borehole seismic network (PB), the Portable Observatories for Lithospheric Analysis and Research Investigating Seismicity (PO), the University of Oregon regional network (UO), the Pacific Northwest regional seismic network (UW), the United States national seismic network (US), the Berkeley digital seismography network (BK), the Cascade chain volcano monitoring (CC), the Caltech regional seismic network (CI), and many other flexible arrays including High Lava Plains broadband seismic experiment (XC), Mendocino experiment (XQ), Cascadia arrays for EarthScope (XU), Flexarray along Cascadia experiment for segmentations (YW), and Wallowa broadband experiment (ZG). To our knowledge, this is the first time that all these networks are combined for surface-wave tomography, resulting in a dataset that provides the most comprehensive, dense coverage of the study area.

Prior to cross-correlating the vertical-component waveforms from station pairs, we remove instrument response, normalize ambient noise data with a frequency-time-normalization method (Shen et al., 2012), and eliminate time segments of large (M>5.5) earthquakes. To increase the signal-to-noise ratio, we stack daily cross-correlations for each station pair, producing high-quality Rayleigh waves at periods of 7-200 s (Figure 3). As this study focuses on structures from the mid-crust to upper mantle, for computational reasons we only use 15-200 s periods in the analysis described here. The EGFs are then recovered as the time derivative of the stacked cross-correlations (e.g., Sabra et al., 2005; Snieder, 2004). In addition, we obtain monthly stacks of cross-correlations, whose variations provide estimates of the uncertainties of EGFs and their travel times.

The conditions to equate EGFs with Green's functions of the Earth include a uniform distribution of noise sources around the seismic stations and zero attenuation (e.g., Wapenaar, 2004; Wapenaar and Fokkema, 2006). These conditions are usually not strictly satisfied in ambient noise seismic tomography (e.g., Yang and Ritzwoller, 2008) and the Pacific Northwest is no exception as most of the ambient noise comes from the Pacific Ocean. Following helioseismological practices in dealing with similar issues, Tromp et al. (2010) suggested construction of ensemble-averaged cross-correlation and corresponding ensemble-averaged sensitivity kernels. This method requires the power spectral distribution of the ambient noise sources, which is highly variable spatially and temporally (e.g., Uchiyama and McWilliams, 2008; Bromirski and Gerstoft, 2009). To construct ensemble cross-correlations, we must know the global power spectral distribution of ambient noise sources for the various overlapping recording periods of station pairs. This detailed knowledge of the global power spectral distribution of ambient noise sources is currently unavailable and requires substantial work that is beyond the scope of this study. On the other hand, it has been suggested that the non-uniformity of noise sources would significantly affect the surface-wave amplitude (Tsai and Moschetti, 2010) but not the velocity (Snieder, 2004). Numerical experiments show that the non-uniform distribution of noise sources leads to less than 0.5% error in travel times and phase velocity (Yang and Ritzwoller, 2008). This level of error is much less than the lateral velocity variations in the Cascades (e.g., Porritt et al., 2011; Gao et al., 2012). Furthermore, for the study area with an average travel time of ~150 s, a 0.5% error is equivalent to a ~0.75-s error in travel time, less than measurement errors and the RMS data misfit. In the following, we consider the effects of non-uniform noise source distribution on travel times secondary to those of the Earth structure, and EGF a close approximation to the Green's function of the Earth for velocity inversion.

### 2.2. Finite-difference wave simulation

We implement a nonstaggered-grid, finite-difference method to simulate wave propagation in the 3D spherical Earth structure (Zhang et al., 2012). Each seismic station is considered as a virtual source and all others as receivers. The regional 3D shear-wave velocity model by Gao et al. (2011) is chosen as the initial reference model, merged with CUB (Shapiro and Ritzwoller, 2002) at locations beyond the coverage of the original model. Deeper than 400 km, we use the AK135 model (Kennett et al., 1995). P-wave velocity is converted from shear-wave velocity with a Vp/Vs ratio of 1.74 in the crust

(Brocher, 2005) and the depth-dependent relationship of Vp and Vs of AK135 in the mantle (Kennett et al., 1995). Density is calculated as a function of Vp (Christensen and Mooney, 1995). Constraints on the Moho depth (Lowry and Pérez-Gussinyé, 2011) are added to the initial reference model. No anisotropy and attenuation are included in the simulation, though the effect of attenuation is considered in the interpretation of the tomographic results.

For computational reasons, we carry out two levels of finite-difference wave simulation, starting from a coarser grid for longer-period waves. The horizontal grid spacing is 10 km and 5 km for level 1 and level 2, respectively, along the geographic longitude and latitude. The vertical grid spacing is about one-third of the horizontal spacing near the surface and increases with depth to approximately the same as the horizontal spacing at ~100 km depth. Such grid sizes are sufficient to accurately simulate waves at periods greater than 40 s and 21 s for level 1 and level 2, respectively (Zhang et al., 2012). The total wave propagation time is 1000 s as the longest inter-station distance is ~3000 km. To maintain numerical stability, we use a time step of 0.5 s for level 1 and 0.25 s for level 2. To calculate Green's functions, we use a Gaussian pulse with a half width of 7.5 s and 4 s, respectively, as the source-time function of the vertical force applied at the virtual station. The wave simulations are executed on a Linux cluster with 17 nodes (each with 24 cpu-cores). It takes about 0.3 and 1 hour per simulation with two nodes for level 1 and level 2, respectively.

2.3. Cross-correlation of EGFs and synthetics

The phase delay times between the EGFs and synthetics are measured by cross-correlation at multiple overlapping period bands, with the central periods of 55 s, 75 s, 112.5 s, 150 s, and 200 s for level 1 and 22.5 s, 37.5 s, 56 s, 75 s, and 112.5 s for level 2. The corresponding Rayleigh waves have peak sensitivities to structures from the mid-crust to 250 km depth. In this paper, we focus our interpretations on the mantle structure (50-140 km). Before the delay measurement, the EGFs are convolved with the source-time function used in the calculation of Green's functions to account for the finite-frequency nature and initial time shift of the simulated Green's functions. To ensure high-quality signals, the signal-to-noise ratio of EGFs is required to be at least eight, the inter-station distance at least 1.5-wavelength, and the cross-correlation coefficient between the EGFs and synthetics greater than 0.90. The number of measured phase delays varies from ~3,000 to 20,000 within different frequency bands.

2.4. Sensitivity kernels and inversion method

In previous surface-wave tomographic studies (e.g., Yang et al., 2008; Gao et al., 2011; Porritt et al., 2011; Wagner et al., 2012), the inversion of phase velocity to shear-wave velocity is carried out under the assumption that Rayleigh waves are not affected by P-wave speed. This is not accurate, especially at shallow depths (Figure S1). We represent the Rayleigh wave phase delay time $\delta\mathbf{t}$ as a joint Vp and Vs inverse problem,

$$\delta\mathbf{t} = \int [K_\alpha(\mathbf{m}_0, x)\Delta\mathbf{m}_\alpha + K_\beta(\mathbf{m}_0, x)\Delta\mathbf{m}_\beta]dV \qquad (1)$$

where $\mathbf{m}_0$, $\Delta\mathbf{m}_\alpha$ and $\Delta\mathbf{m}_\beta$ are the 3D reference model, Vp and Vs perturbations, and $K_\alpha$ and $K_\beta$ the Rayleigh-wave sensitivity kernels to Vp and Vs, respectively. The integration is for the 3D volume of the model. Although the velocity structure in the shallow crust is not well constrained in this study due to the intermediate- to long-period data used, the inclusion of Vp in inversion provides additional degrees of freedom that minimize the extent to which Vp anomalies in the shallow crust are mapped into the deep crust and upper mantle. The effect of density is not explicitly expressed in the equation, but density is recalculated based on Vp (Christensen and Mooney, 1995) after each iterative model update. The sensitivity kernels are calculated with the strain-Green-tensor-based, scattering-integral method (Zhao et al., 2005; Chen et al., 2007; Zhang et al., 2007). The inverse problem is solved with damping and smoothness constraints. The best-fit damping and smoothing parameters, which are gradually reduced with iterative inversions, are chosen from the tradeoff of the normalized chi-squared value and the model variance (Gao and Shen, 2012).

We start wave propagation simulation from level 1 to construct a large-scale framework, which provides the reference model for level 2. The model is then iteratively updated by alternating the two-level full-wave tomographic imaging. The solution does not change significantly after 3-4 iterations. In total, we run five iterations for each level. Compared to the initial reference model, the updated model yields synthetic waveforms that match the observed EGFs much better (Figure 4), with the standard deviation of phase delays decreasing from 2.5 s to 0.7 s. The phase delays of the updated model as a function of inter-station distance are less scattered and centered around zero (Figure 5). We observe that, on average, our model has much stronger velocity perturbations compared to previous models (Shapiro and Ritzwoller, 2002; Yang et al., 2008; Gao et al., 2011). It appears that more data and/or a more accurate methodology result in a stronger contrast of the velocity anomalies. This has also been observed in the western

U.S. among various body-wave tomographic studies (as compared by Becker (2012)), Iceland (Hung et al., 2004), and the Lau Basin (Wiens et al., 2006, 2012).

## 3. Seismic Results and Discussion

We focus our discussion on the mantle structure, in the depth range that is best constrained by the ambient noise data and most relevant to melt generation (40-140 km). Because the Rayleigh-wave sensitivity to Vp concentrates primarily in the shallow depth (Figure S1), Vp in the mantle is not well constrained and thus not interpreted. Our shear-wave velocity model, as shown in Figure 6, has features that are similar to those in the previous tomographic models. For example, we image the low-velocity Yellowstone hotspot at all the depths (e.g., Moschetti et al., 2010; Gao et al., 2011; Wagner et al., 2012), an observation that is consistent with the active magmatism and high surface heat flow (Lowry and Pérez-Gussinyé, 2011). The geometry of the subducting slab is better resolved in this study at depths greater than 80 km than in previous surface-wave studies (e.g., Yang et al., 2008; Gao et al., 2011; Porritt et al., 2011; Wagner et al., 2012). The seismic velocity of the subducting slab is heterogeneous along strike. At depth of ~110 km, we image a slab hole (or weak slab) in northern Oregon, which spatially correlates with a similar gap or weak slab in the body-wave tomographic models at depths greater than 160 km (Roth et al., 2008; Burdick et al., 2008; Schmandt and Humphreys, 2010).

The most striking features of our Cascade model are three low shear-wave velocity volumes, with velocities as low as ~3.6 km/s, in the upper mantle along the back-arc (Figures 6 and 7). The back-arc low-velocity anomalies are about 200 km away from the arc, segmented along strike and correlate spatially with the three volcano clusters along the Cascades (Figure 2). The inter-spacing of these anomalies (center to center) is ~300 km. Large-scale, separate low-velocity anomalies have also been imaged in the upper mantle wedge along the Izu-Bonin-Mariana arc (where the spacing of the anomalies is ~500 km), and are interpreted as heterogeneous along-strike mantle flow (Isse et al., 2009). Among the three anomalies, the central one at the Oregon back-arc is the largest in size and lowest in velocity and corresponds to the arc segment with the largest Quaternary eruption volume (Sherrod and Smith, 1990).

Analysis of model resolution in full-wave inversion is complicated by several factors, including the non-linear relationship between the data and model and the multiple model iterations carried out to reach the final model. The prohibitive cost of forward wave propagation simulation makes the probabilistic approaches commonly used to deal with nonlinear inverse problems impractical. Fichtner and Trampert (2011) proposed a method based on the Fréchet derivatives of the misfit function. Their method, however, depends on the condition that the model is in the vicinity of an optimal Earth model and the global minimum of the misfit function has been found, an assumption that is difficult to verify in large-scale, nonlinear full-wave inversion.

Synthetic inversion of various input models is a common practice in tomographic resolution analysis. This approach has limitations and can be misleading in the sense that synthetic inversion explores only a limited model subspace (Lévêque et al., 1993). Nevertheless, it is useful if the limitations are understood and interpretations are restricted to the model subspaces explored. With this caveat in mind, we run multiple resolution tests. We first use our preferred model as the input (Figure 6). The synthetic phase delay times are calculated with the sensitivity kernels (Eq. 1). The uncertainties of individual observed phase measurements are estimated from the variations of monthly-stacked EGFs and added to the synthetic phase delays. The velocity perturbations of the input model are well reconstructed in inversion at the depth of our interest (Figures 9 and 10). We then run the 3D checkerboard resolution tests with a maximum of ±10% velocity perturbations for both Vp and Vs (Figures S2-S4). The velocity variation within each checkerboard cell is a cosine function. The dimensions of the checkerboard cells vary from 100-200 km along the geographic longitude and latitude, and from 90-150 km vertically. For the small checkerboard, although the pattern of the velocity perturbation can be fairly well reconstructed, the magnitude is underestimated. For the larger checkerboards, both the pattern and smooth variation of the magnitude can be well recovered. The sizes of our observed back-arc low-velocity volumes are comparable to the largest checkerboard. In the above tests, the recovered structures are obtained in a single model iteration. A fully non-linear inversion with multiple iterations, as for the observed data in this study, will further minimize the residual, resulting in a sharper reconstruction of the model. The results in Figures 9 and 10 can thus be considered conservative. Taken together, Figures 9 and 10 and the checkerboard tests indicate that an Earth model with a structure that resembles the inferred back-arc low-velocity anomalies is well resolved.

The three distinct, segmented low-velocity volumes along the Cascade back-arc in the upper mantle are resolved for the first time. We attribute the resolution to the dense data coverage, an EGF dataset with a broad frequency band well suited for imaging the crust and upper 200 km mantle, and an advanced full-wave tomographic method. This allows us to gain new insight into the dynamic processes of the Cascadia mantle wedge.

### 3.1. What controls the reduction of shear-wave velocity?

The seismic velocity can be affected by a few factors, including temperature, water content and presence of partial melt. As shown in Figure 8, the shear-wave velocity of the back-arc anomalies is 3.6-4.0 km/s within the depths of 80-120 km, which are deeper and lower compared to where melting is inferred for the 0-4 Ma Pacific mantle (Nishimura and Forsyth, 1989), the Lau back-arc basin (Wiens et al., 2006, 2008), and the Izu-Bonin-Mariana arc (Isse et al., 2009). This strongly suggests presence of melt beneath the Cascade back-arc, although temperature, water content, and grain size can also contribute to the reduction of seismic velocity. The young, subducting Juan de Fuca plate leads to shallow dehydration and less subduction-related water input into the back-arc mantle wedge to affect seismic velocities. This inference is supported by the observation that the primitive basalts erupted on the back-arc side of the Cascades are nominally dry (Ruscitto et al., 2010). The back-arc lithosphere inferred from our model is relatively thin (Figure 7), consistent with the receiver function images of the lithosphere-asthenosphere boundary (Kumar et al., 2012; Hopper et al., 2013) and the high surface heat flow (Blackwell et al., 1990; Ingebritsen and Mariner, 2010). However, mantle temperature alone cannot explain the observed back-arc low-velocity anomalies. Using the method of Jackson and Faul (2010) for the geothermal profiles estimated from the Cascade heat flow (based on Currie et al. (2004), Figure 8c), we find that the predicted shear-wave velocities are all much higher than the observed (Figure 8b). The lowest velocity calculated with the geotherms is ~4.2 km/s within the depth range of 80-110 km, which is more than 0.4 km/s higher than the observed velocities beneath the back-arc low-velocity anomalies (blue lines in Figure 8b). Correction for attenuation ignored in forward wave simulation assuming a low Q value of 50 (Dalton et al., 2008) reduces the velocity mismatch by ~0.1 km/s to 0.3 km/s. The additional velocity reduction needed to match the observed is indicative of the presence of partial melt, which can drastically reduce shear-wave velocities. The electromagnetic study in the region (Egbert, 2012) also supports the possible presence of melt in the back-arc.

Using S-to-P converted phases (Sp), Hopper et al. (2013) map the lithosphere-asthenosphere boundary (LAB) beneath the Pacific Northwest. The depth of their LAB is consistent with the base of the imaged high-velocity mantle lid of the upper plate (Figure 7), including the deepening of the LAB from the back-arc in south and central Oregon to the back-arc of Washington, where the latest magmatism occurred more than 15 Ma ago. Hopper et al. (2013) attribute the consistent Sp observed beneath the Cascades back-arc to the negative LAB velocity gradient created by a layer of partial melt ponding beneath a solidus-defined boundary. Interestingly, strong Sp phases - in other words a large and/or sharp negative velocity gradient at the LAB - are clustered in three areas in the back-arc that roughly overlap with the three low-velocity volumes (Figure 7). A simple explanation of the spatial overlap between the strong Sp phase and low-velocity volumes is that the low-velocity volumes represent the regions of partial melt production. Melts migrate upwards to collect at the base of the lithosphere, causing a relatively large and/or sharp negative velocity gradient to generate the strong Sp phase.

3.2. What processes contribute to the pattern of the back-arc anomalies?

The geometry and magnitude of the velocity anomalies in the mantle wedge and their spatial correlation with the arc volcanoes provide constraints on the mechanism of melt generation at the Cascades. Compositionally buoyant small-scale diapirs (Figure 1) triggered by fluid released at the slab arise from the top of the slab with a more or less vertical geometry beneath the arc (Hasenclever et al., 2011). These clearly do not match the observed back-arc low-velocity volumes. Nevertheless, we cannot exclude the possibilities of melting related to small-scale diapirs (Hall and Kincaid, 2001) because of the difficulties in tomographically imaging such small-scale features. Beneath the Oregon Cascade arc, where a hydrated uppermost mantle wedge has been suggested previously (Bostock et al., 2002), the low-velocity zone atop and approximately parallel to the slab extending from depth of ~100 km upward to ~60 km near the Cascade arc (Figure 7c) is consistent with flux melting (Wiens et al., 2008; Zhao et al., 1997). Note that the plate interface terminates at depth of 100 km and is poorly defined at greater depths (McCrory et al., 2004, 2012). The lack of such a low-velocity zone along the entire arc (Figures 7b, 7c and7d), however, indicates a variable strength of flux melting along the arc.

The pattern of the back-arc low-velocity anomalies (Figures 7b, 7c and 7d) is most consistent with subduction-induced mantle upwelling and decompressional melting (Figure 1). The asthenospheric flow from beneath the old and thick North America continental lithosphere towards the mantle wedge must undergo decompression. The fact that the lowest velocities are at 80-110 km depth suggests that melting may involve damp (50-200 p.p.m $H_2O$) peridotite and/or carbonated peridotite (Dasgupta et al., 2013). Bifurcation of the Yellowstone plume driven by subduction-induced mantle flow (Kincaid et al., 2013) may contribute to an excess mantle temperature and the stronger anomaly beneath the Oregon back-arc, though bifurcation of the plume cannot explain all three back-arc low-velocity volumes. We suggest that the decompressional melts in the back-arc are not only responsible for the volcanism in the back-arc (Till et al., 2013), but also a likely source of the low-water-content magmas at the arc (Elkins Tanton et al., 2001; Ruscitto et al., 2010) ~200 km away. Melts may migrate upslope along a dipping decompaction channel near the base of the lithosphere (Sparks and Parmentier, 1991), though the exact mechanism of melt migration remains unclear.

Extension of the northern Basin and Range may contribute to the velocity reduction in the southern Cascades (Ingebritsen and Mariner, 2010; Wang et al., 2002). However, this mechanism cannot explain the back-arc low-velocity anomalies to the north. Furthermore, the Basin and Range goes beyond the study area, while the southern low-velocity anomaly extends only to near the southern end of the slab (Figure 2). The spatial mismatch with the Basin and Range and the close correlation with the slab and the arc volcanoes suggest that the low-velocity anomaly beneath the southern Cascade back-arc reflects primarily the subduction processes.

Laboratory experiments with slab rollback and small back-arc extension that mimic the Cascadia subduction zone indicate that upwelling above the slab in the mantle wedge is influenced by the large-scale plate motion (Druken et al., 2011). The laboratory experiments show the strongest upwelling in the center of the slab, which matches the largest low-velocity anomaly in our seismic imaging. However, the magnitude of upwelling varies gradually along strike from the center to the edge in the laboratory experiments, different from the segmented pattern of our observed low-velocity anomalies. Thus, besides the plate-motion-controlled processes other mechanisms must also be at work in Cascadia. One likely mechanism is small-scale convection due to buoyancy associated with melting, differences in the temperature of the upwelling mantle, negative buoyancy from the cooling of the lithosphere above, and pre-existing lithosphere structure. Between the central and northern low-velocity volumes, the SKS-splitting anisotropy pattern is complex and deviates from the plate motion direction (Yuan and Romanowicz, 2010), possibly the consequence of the disturbance of lattice-preferred orientation by small-scale convection in the mantle wedge (Morishige and Honda, 2011).

The tight spatial correlation between the low-velocity volumes in the Cascade back-arc and the volcano clusters is similar to the hot-finger structure at the Honshu subduction zone (Tamura et al., 2002), which may be a consequence of small-scale convection within a low-viscosity wedge in the Japan back-arc (Honda and Yoshida, 2005). However, there is a notable difference between the Cascade back-arc anomalies and the hot-finger structure beneath Honshu: The spacing of the volcano clusters and the finger-like feature in Honshu is 50-100 km, much smaller than at the Cascades. Numerical simulations show that the formation and dimension of small-scale convection depend on the slab age, subducting rate and water flux into the mantle released from the slab, which affects mantle viscosity (Honda and Yoshida, 2005). Furthermore, long-wavelength rolls dominate in the early development of small-scale convection and/or for cases with a small subduction speed, while short-wavelength rolls take over in late stages and/or for cases with a large subduction speed (Honda, 2011). Compared to Honshu, the Cascadia subduction zone is younger (subduction started about 48 Ma versus over 130 Ma) and has a lower subduction speed (3.5 cm/yr versus 9 cm/yr). Thus, the link between the wavelength of small-scale 3D convection and clustering of arc volcanoes at different subduction zones may reflect the different subduction histories and the relatively dry or wet nature of the mantle wedge.

## 4. Conclusions

A new upper mantle shear-wave velocity model in the Pacific Northwest has been constructed in this study using an advanced full-wave ambient noise tomographic method. We have imaged three segmented low-velocity (~3.6-4.0 km/s) volumes along the Cascade back-arc within the depth range of 80-115 km, which provides new insights into the melting processes at the Cascadia subduction zone. Such low-velocity anomalies require presence of partial melt within a hot upper mantle wedge. The correlation of the three back-arc low-velocity anomalies in the mantle wedge with the volcanic arc and the subducting slab is consistent with the pattern predicted by subduction-induced decompressional melting. Furthermore, the along-strike variation suggests existence of small-scale convection with a scale of ~300 km. Whether and how the melts generated within the back-arc low-velocity volumes supply the low-water-content magmas at the arc (Elkins Tanton et al., 2001; Ruscitto et al., 2010) remain unknown.

Whether small-scale mantle convection and 3D decompressional melting are ubiquitous in the back-arc of subduction zones requires further understanding of the connections between large-scale plate-driven processes and small-scale convection. Although our shear-wave velocity model provides important constraints on the melting processes in Cascadia, several other lines of work should be carried to further constrain the complex 3D processes at the subduction zone. Our interpretation of the existence of back-arc decompressional melting is based mainly on the shear-wave velocity model. Poisson's ratio is in fact more sensitive to melt and fluid than Vp and Vs. Attenuation also reflects temperature, water content, and grain size. Therefore, a Vp/Vs model and attenuation structure (e.g., Lawrence and Prieto, 2011) will help distinguish the contributions of melt, fluid, temperature and chemical composition, and understand the melting generation processes (e.g., Wiens et al., 2008).

Acknowledgements

All the seismic waveforms used in this study are obtained from the IRIS Data Management Center and the Canadian National Seismograph Network. We thank Katherine Kelley, Mark Parmentier, Kelsey Druken and Chris Kincaid for helpful discussions. We thank Ulrich Faul for providing the code to calculate shear-wave velocities. This research is supported by the National Science Foundation (grant number 1144771) and Air Force Research Laboratory (grant number FA9453-10-C-0217).

References

Abers, G. A., L. S. MacKenzie, S. Rondenay, Z. Zhang, A. G. Wech, and K. C. Creager (2009), Imaging the source region of Cascadia tremor and intermediate-depth earthquakes, *Geology, 37,* 1119-1122.

Audet, P., M. G. Bostock, N. I. Christensen, and S. M. Peacock (2009), Seismic evidence for overpressured subducted oceanic crust and mega-thrust fault sealing, *Nature, 457,* 76-78.

Becker, T. W. (2012), On recent seismic tomography for the western United States, *Geochem. Geophys. Geosyst., 13,* Q01W10, doi:10.1029/2011GC003977.

Blackwell, D. D., J. L. Steele, M. K. Frohme, C. F. Murphey, G. R. Priest, and G. L. Black (1990), Heat flow in the Oregon Cascade Range and its correlation with regional gravity, Curie point depths, and geology, *J. Geophys. Res., 95,* 19,475-19,493.

Bostock, M. G., R. D. Hyndman, S. Rondenay, and S. M. Peacock (2002), An inverted continental Moho and serpentinization of the forearc mantle, *Nature, 47,* 536-538.

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

Brocher, T. M., T. Parsons, A. M. Trehu, C. M. Snelson, and M. A. Fisher (2003), Seismic evidence for widespread serpentinized forearc upper mantle along the Cascadia margin, *Geology, 31,* 267-270.

Brocher, T. M., T. Parsons, R. A. Blakely, N. I. Christensen, M. A. Fisher, R. E. Wells, and the SHIPS Working Group (2001), Upper crustal structure in Puget Lowland, Washington: Results from the 1998 Seismic Hazards Investigation in Puget Sound, *J. Geophys. Res., 106,* 13,541-13,564.

Bromirski, P. D., and P. Gerstoft (2009), Dominant source regions of the Earth's "hum" are coastal, *Geophys. Res. Lett., 36*, L13303, doi:10.1029/2009GL038903.

Burdick, S., C. Li, V. Martynov, T. Cox, J. Eakins, T. Mulder, L. Astiz, F. L. Vernon, G.

L. Pavlis, and R. D. van der Hilst (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., R. D. van der Hilst, F. L. Vernon, V. Martynov, T. Cox, J. Eakins, L. Astiz, and G. L. Pavlis (2010), Model Update January 2010: Upper mantle heterogeneity beneath North America from travel time tomography with global and USArray Transportable Array data, *Seismol. Res. Lett., 81,* 689–693.

Calkins, J. A., G. A. Abers, G. Ekström, K. C. Creager, and S. Rondenary (2011), Shallow structure of the Cascadia subduction zone beneath western Washington from spectral ambient noise correlation, *J. Geophys. Res., 116,* B07302, doi:10.1029/2010JB007657.

Calvert, A. J., M. A. Fisher, and SHIPS Working Group (2001), Imaging the Seattle fault zone with high-resolution seismic tomography, *Geophys. Res. Lett., 28,* 2337-2340.

Chen, P., T. H. Jordan, and L. Zhao (2007), Full three-dimensional tomography: a comparison between the scattering-integral and adjoint-wavefield methods, *Geophys. J. Int., 170,* 175-181.

Christensen, N. I., and W. D. Mooney (1995), Seismic velocity structure and composition of the continental crust: A global view, *J. Geophys. Res*., *100,* 9761-9788.

Conder, J. A., (2005), A case for hot slab surface temperatures in numerical viscous flow models of subduction zones with an improved fault zone parameterization, *Phys. Earth Planet. Inter., 149,* 155–164.

Conder, J. A., D. A. Wiens, and J. Morris (2002), On the decompression melting structure at volcanic arcs and back-arc spreading centers, *Geophys. Res. Lett., 29,* doi:10.1029/2002GL015390.

Crosson, R. S., N. P. Symons, K. C. Creager, L. A. Preston, T. van Wagoner, T. M. Brocher, M. Fisher, and the SHIPS Working Group (2002), Three-dimensional structure of the Cascadia forearc region from SHIPS active experiment and earthquake observations: tomographic inversion provides a high-resolution view into the core of the Cascadia forearc complex, U.S. Geological Survey Open-File Report 02-328 and Geological Survey of Canada Open File 4350, 33-34.

Currie, C. A., K. Wang, R. D. Hyndman, and J. He (2004), The thermal effects of steady-state slab-driven mantle flow above a subducting plate: the Cascadia subduction zone and backarc, *Earth Planet. Sci. Lett*., *223,* doi:10.1016/j.epsl.2004.04.020.

Dahlen, F. A., and Y. Zhou (2006), Surface-wave group-delay and attenuation kernels, *Geophys. J. Int., 165,* 545-554.

Dalton, C. A., G. Ekström, and A. M. Dziewoński (2008), The global attenuation structure of the upper mantle, *J. Geophys. Res., 113*, B09303, doi:10.1029/2007JB005429.

Dasgupta, R., A. Mallik, K. Tsuno, A. C. Withers, G. Hirth, and M. M. Hirschmann (2013), Carbon-dioxide-rich silicate melt in the Earth's upper mantle, *Nature, 493,* 211-216.

Delorey, A. A., and J. E. Vidale (2011), Basin shear-wave velocities beneath Seattle, Washington, from noise-correlation Rayleigh waves, *Bull. Seism. Soc. Am., 101,* 2162-2175, doi:10.1785/0120100260.

Druken, K. A., M. D. Long, and C. Kincaid (2011), Patterns in seismic anisotropy driven by rollback subduction beneath the High Lava Plains, *Geophys. Res. Lett., 38,* L13310, doi:10.1029/2011GL047541.

Egbert, G. D. (2012), Three-dimensional inversion of EarthScope magnetotelluric data: crustal and mantle conductivity beneath the NW USA. *IRIS Webinars*, http://www.iris.edu/hq/webinar (November 27, 2012).

Elkins Tanton, L. T., T. L. Grove, and J. Donnelly-Nolan (2001), Hot, shallow mantle melting under the Cascades volcanic arc, *Geology, 29,* 631-634.

Fichtner, A., and J. Trampert (2011), Resolution analysis in full waveform inversion, *Geophys. J. Int*., *187,* 1604-1624.

Gao, H., and Y. Shen (2012), Validation of Shear-wave velocity models of the Pacific Northwest, *Bull. Seismol. Soc. Am., 102(6),* 2611-2621.

Gao, H., E. D. Humphreys, H. Yao, and R. D. van der Hilst (2011), Crust and lithosphere structure of the Northwestern U.S. with ambient noise tomography: Terrane accretion and Cascade arc development, *Earth Planet. Sci. Lett*., *304,* doi:10.1016/j.epsl.2011.01.033.

Graindorge, D., G. Spence, P. Charvis, J. Y. Collot, R. Hyndman, and A. M. Tréhu (2003), Crustal structure beneath the Strait of Juan de Fuca and southern Vancouver Island from seismic and gravity analyses, *J. Geophys. Res., 108 (B10),* 2484, doi:10.1029/2002JB001823.

Hall, P. S., and C. Kincaid (2001), Diapiric flow at subduction zones: A recipe for rapid transport, *Science, 292,* 2472-2475.

Hasenclever, J., J. P. Morgan, M. Hort, and L. H. Rüpke (2011), 2D and 3D numerical models on compositionally buoyant diapirs in the mantle wedge, *Earth Planet. Sci. Lett*., *311,* 53-68, doi:10.1016/j.epsl.2011.08.043.

Honda, S. (2011), Planform of small-scale convection under the island arc, *Geochem. Geophys. Geosyst., 12,* Q11005.

Honda, S., and T. Yoshida (2005), Application of the model of small-scale convection under the island arc to the NE Honshu subduction zone, *Geochem. Geophys. Geosyst., 6,* Q01002, doi:10.1029/2004GC000785.

Hopper, E., H. A. Ford, K. M. Fischer, V. Lekic, and M. J. Fouch (2013), The lithosphere-asthenosphere boundary and the tectonic and magmatic history of the northwestern United States, in review.

Hung, S.-H., Y. Shen, and L.-Y. Chiao (2004), Imaging seismic velocity structure beneath the Iceland hotspot - A finite-frequency approach, *J. Geophys. Res.*, *109*, B08305, doi:10.1029/2003JB002889.

Ingebritsen, S. E., and R. H. Mariner (2010), Hydrothermal heat discharge in the Cascade Range, northwestern United States, *J. Volcanol. Geotherm. Res., 196*, 208-218, doi:10.1016/j.jvolgeores.2010.07.023.

Isse, T. *et al.* (2009), Seismic structure of the upper mantle beneath the Philippine Sea from seafloor and land observation: Implications for mantle convection and magma genesis in the Izu-Bonin-Mariana subduction zone, *Earth Planet. Sci. Lett., 278,* 107-117.

Jackson, I., and U. H. Faul (2010), Grainsize-sensitive viscoelastic relaxation in olivine: Towards a robust laboratory-based model for seismological application, *Phys. Earth Planet. Inter., 183,* 151-163.

James, D. E., M. J. Fouch, R. W. Carlson, and J. B. Roth (2011), Slab fragmentation, edge flow and the origin of the Yellowstone hotspot track, *Earth Planet. Sci. Lett.,* *311,* 124–135.

Kelley, K. A., T. Plank, T. L. Grove, E. M. Stolper, S. Newman, and E. Hauri (2006), Mantle melting as a function of water content beneath back-arc basins, *J. Geophys. Res., 111,* B09208, doi:10.1029/2005JB003732.

Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int*., *122,* 108-124.

Kincaid, C., K. A. Dreken, R. W. Griffiths, and D. R. Stegman (2013), Bifurcation of the Yellowstone plume driven by subduction-induced mantle flow, *Nature Geoscience, 6,* doi:10.1038/NGEO1774.

Kumar, P., R. Kind, X. Yuan, and J. Mechie (2012), USArray receiver function images of the lithosphere-asthenosphere boundary, *Seis. Res. Lett., 83,* doi:10.1785/gssrl.83.3.486.

Lawrence, J. F., and G. A. Prieto (2011), Attenuation tomography of the western United States from ambient seismic noise, *J. Geophys. Res., 116*, B06302, doi:10.1029/2010JB007836.

Lee, J. M., and R. S. Crosson (1990), Tomographic imaging of local earthquake delay times for three-dimensional velocity variation in western Washington, *J. Geophys. Res., 95,* 4763– 4776.

Lévêque, J. J., L. Rivera, and G. Wittlinger (1993), On the use of the checkerboard test to assess the resolution of tomographic inversions, *Geophys. J. Int., 115*, 313-318.

Lowry, A. R., and M. Pérez-Gussinyé (2011), The role of crustal quartz in controlling Cordilleran deformation, *Nature*, *471*, 353-357, doi:10.1038/nature09912.

McCrory, P. A., J. L. Blair, D. H. Oppenheimer, and S. R. Walter (2004), Depth to the Juan de Fuca slab beneath the Cascadia subduction margin: A 3-D model for sorting earthquakes, *U.S. Geol. Surv. Data Ser. DS-91*.

McCrory, P. A., J. L. Blair, F. Waldhauser, and D. H. Oppenheimer (2012), Juan de Fuca slab geometry and its relation to Wadati-Benioff zone seismicity, *J. Geophys. Res., 117*, B09306, doi:10.1029/2012JB009407.

Morishige, M., and S. Honda (2011), Three-dimensional structure of P-wave anisotropy in the presence of small-scale convection in the mantle wedge, *Geochem. Geophys. Geosyst., 12,* Q12010, doi:10.1029/2011GC003866.

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.

Nishimura, C. E., and D. W. Forsyth (1989), The anisotropic structures of the upper mantle in the Pacific, *Geophys. J., 96,* 203-229, doi:10.1111/j.1365-246X.1989.tb04446.x.

Obrebski, M., R. M. Allen, M. Xue, and S.-H. Hung (2010), Slab-plume interaction beneath the Pacific Northwest, *Geophys. Res. Lett., 37,* L14305, doi:10.1029/2010GL043489.

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.

Parsons T., R. E. Wells, M. A. Fisher, E. Flueh, and ten Brink U. S. (1999), Three-dimensional velocity structure of Siletzia and other accreted terranes in the Cascadia forearc of Washington, *J. Geophy. Res., 104,* 18,015–18,039, doi: 10.1029/1999JB900106.

Porritt, R. W., R. M. Allen, D. C. Boyarko, and M. R. Brudzinski (2011), Investigation of Cascadia segmentation with ambient noise tomography, *Earth Planet Sci. Lett., 309,* 67-76, doi:10.1016/j.epsl.2011.06.026.

Ramachandran, K., S. E. Dosso, C. Zelt, G. D. Spence, R. D. Hyndman, and T. M. Brocher (2004), Upper crustal structure of Southwestern British Columbia from the 1998 Seismic Hazards Investigation in Puget Sound, *J. Geophy. Res., 109,* B9, B09303, doi: 10.1029/2004JB003092.

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.

Ruscitto, D. M., P. J. Wallace, E. R. Johnson, A. J. R. Kent, and I. N. Bindeman (2010), Volatile contents of mafic magmas from cinder cones in the Central Oregon High Cascades: implications for magma formation and mantle conditions in a hot arc, *Earth Planet. Sci. Lett., 298,* 153–161, doi:10.1016/j.epsl.2010.07.037.

Sabra, K. G., P. Roux, and W. A. Kuperman (2005), Emergence rate of the time-domain Green's function from the ambient noise cross-correlation function, *J. Acoust. Soc. Am., 118,* 3524–3531.

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, doi:10.1016/j.epsl.2010.06.047.

Schmandt, B., and E. Humphreys (2011), Seismically imaged relict slab from the 55 Ma

Siletzia accretion to the northwest United States, *Geology, 39,* 175–178.

Schmidt, M. E., A. L. Grunder, and M. C. Rowe (2008), Segmentation of the Cascade Arc as indicated by Sr and Nd isotopic variation among diverse primitive basalts. *Earth Planet. Sci. Lett., 266,* 166-181.

Shapiro, N. M., and M. H. Ritzwoller (2002), Monte-Carlo inversion for a global shear velocity model of the crust and upper mantle, *Geophys. J. Int., 151,* 88-105.

Shen W., M. H. Ritzwoller, and V. Schulte-Pelkum (2013), A 3-D model of the crust and uppermost mantle beneath the Central and Western US by joint inversion of receiver functions and surface wave dispersion, *J. Geophys. Res. Solid Earth, 118,* doi:10.1029/2012JB009602.

Shen, Y., Y. Ren, H. Gao, and B. Savage (2012), An improved method to extract very broadband empirical Green's functions from ambient seismic noise, *Bull. Seism. Soc. Am., 102 (4).*

Shen, Y., and W. Zhang (2010), Full-wave ambient noise tomography of the northern Cascadia, SSA meeting (abstract), *Seismological Research Letters*, *81*, 300.

Sherrod, D. R., and J. G. Smith (1990), Quaternary extrusion rates of the Cascade Range, Northwestern United States and Southern British Columbia, *J. Geophys. Res., 95,* 19,465–19,474.

Sigloch, K. (2011), Mantle provinces under North America from multi-frequency P wave tomography, *Geochem. Geophys. Geosyst., 12,* Q02W08, doi:10.1029/2010GC003421.

Snieder, R. (2004), Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev. E., 69,* 046610.

Sparks, D. W., and E. M. Parmentier (1991), Melt extraction from the mantle beneath spreading centers, *Earth Planet. Sci. Lett., 105,* 368-377.

Symons, N. P., and R. S. Crosson (1997), Seismic velocity structure of the Puget Sound region from 3-D non-linear tomography, *Geophys. Res. Lett., 24,* 2593–2596.

Tamura, Y., Y. Tatsumi, D. Zhao, Y. Kido, and H. Shukuno (2002), Hot fingers in the mantle wedge: New insights into magma genesis in subduction zones, *Earth Planet. Sci. Lett*., *197,* 105-116.

Till, C. B., T. L. Grove, R. W. Carlson, J. M. Donnelly-Nolan, M. J. Fouch, L. S. Wagner, and W. K. Hart (2013), Depths and temperatures of <10.5Ma mantle melting and the lithosphere-asthenosphere boundary below southern Oregon and northern California, *Geochem. Geophys. Geosyst., 14,* 864–879, doi:10.1002/ggge.20070.

Tréhu, A. M., I. Asudeh, T. M. Brocher, J. H. Luetgert, W. D. Mooney, J. L. Nabelek, and Y. Nakamura (1994), Crustal architecture of the Cascadia forearc, *Science, 266,* 237-243, doi:10.1126/science.266.5183.237.

Tréhu, A. M., T. M. Brocher, K. C. Creager, M. A. Fisher, L. Preston, and G. D. Spence (2002), Geometry of the subducting Juan de Fuca Plate: New constraints from SHIPS98, in The Cascadia Subduction Zone and Related Subduction Systems: Seismic Structure, Intraslab Earthquakes and Processes, and Earthquake Hazards, edited by S. Kirby, K. Wang, and S. Dunlop, U.S. Geol. Surv. Open File Rep., 02328, 25–32. (Available at http://geopubs.wr.usgs.gov/open-file/of02-328/).

Tromp, J., Y. Luo, S. Hanasoge, and D. Peter (2010), Noise cross-correlation sensitivity kernels, *Geophys. J. Int.*, doi:10.1111/j.1365-246X.2010.04721.x.

Tsai, V. C., and M. P. Moschetti (2010), An explicit relationship between time-domain noise correlation and spatial autocorrelation (SPAC) results, *Geophys. J. Int., 182(1),* 454–460, doi:10.1111/j.1365-246X.2010.04633.x.

Uchiyama, Y., and J. C. McWilliams (2008), Infragravity waves in the deep ocean: Generation, propagation, and seismic hum excitation, *J. Geophys. Res., 113*, C07029, doi:10.1029/2007JC004562.

van Keken, P. E., (2003), The structure and dynamics of the mantle wedge, *Earth Planet. Sci. Lett., 215,* 323–338.

van Keken, P. E., B. R. Hacker, E. M. Syracuse, and G. A. Abers (2011), Subduction factory: 4. Depth-dependent flux of H2O from subducting slabs worldwide, *J. Geophys. Res., 116,* B01401, doi:10.1029/2010JB007922.

van Wagoner, T. M., R. S. Crosson, K. C. Creager, G. F. Medema, L. A. Preston, N. P. Symons, and T. M. Brocher (2002), Crustal structure and relocated earthquakes in the Puget Lowland, Washington from high resolution seismic tomography, *J. Geophys. Res., 107,* B12, 10.10129/2001JB000710.

Wagner, L. S., M. J. Fouch, D. James, S. Hanson-Hedgecock (2012), Crust and upper mantle structure beneath the Pacific Northwest from joint inversions of ambient noise and earthquake data, *Geochem., Geophys., Geosys.*, 13, doi:10.1029/2012GC04353.

Wang, K., T. Plank, J. D. Walker, and E. I. Smith (2002), A mantle melting profile across the Basin and Range, SW USA, *J. Geophys. Res., 107,* doi:10.1029/2001JB000209.

Wapenaar, K. (2004), Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation, *Phys. Rev. Lett., 93,* 254301.

Wapenaar, K., and J. Fokkema (2006), Green's function representations for seismic interferometry, *Geophysics, 71,* S133-S146.

Wiens, D. A., K. A. Kelley, and T. Plank (2006), Mantle temperature variations beneath back-arc spreading centers inferred from seismology, petrology, and bathymetry, *Earth Planet. Sci. Lett*., *248,* 30-42, doi:10.1016/j.epsl.2006.04.011.

Wiens, D. A., J. A. Conder, and U. H. Faul (2008), The seismic structure and dynamics of the mantle wedge, *Annu. Rev. Earth Planet. Sci., 36,* 421-455, doi:10.1146/annurev.earth.33.092203.122633.

Wiens, D. A., S. S. Wei, H. N. Relyea, M. Barklage, S. Pozgay, and J. A. Conder (2012), Imaging mantle melting beneath island arcs and backarc spreading centers, Eos Trans. AGU, V21C-03.

Yang, Y., and M. H. Ritzwoller (2008), Characteristics of ambient seismic noise as a source for surface wave tomography, *Geochem. Geophys. Geosys., 9*, Q02008, doi:10.1029/2007GC001814.

Yang, Y., M. H. Ritzwoller, F.-C. Lin, M. P. Moschetti, and N. M. Shapiro (2008), Structure of the crust and uppermost mantle beneath the western United States revealed by ambient noise and earthquake tomography, *J. Geophys. Res., 113,* B12310, doi:10.1029/2008JB005833.

Yuan, H., and B. Romanowicz (2010), Lithospheric layering in the North American craton, *Nature, 466,* doi:10.1038/nature09332.

Zhang, Z., Y. Shen, and L. Zhao (2007), Finite-frequency sensitivity kernels for head waves, *Geophys. J. Int., 171,* 847-856.

Zhang, W., Y. Shen, and L. Zhao (2012), Three-dimensional anisotropic seismic wave modeling in spherical coordinate by a collocated-grid finite difference method, *Geophys. J. Int*., *188,* 1359-1381, doi:10.1111/j.365-246X.2011.05331.x.

Zhao, D., Y. Xu, D. A. Wiens, L. Dorman, J. Hildebrand, and S. Webb (1997), Depth extent of the Lau back-arc spreading center and its relation to subduction processes, *Science, 278,* 254-257.

Zhao, L., T. H. Jordan, K. B. Olsen, and P. Chen (2005), Fréchet kernels for imaging regional earth structure based on three-dimensional reference models, *Bull. Seism. Soc. Am., 95,* 2066-2080.

Zhao, D. P., K. Wang, G. C. Rogers, and S. M. Peacock (2001), Tomographic image of low P velocity anomalies above slab in northern Cascadia subduction zone, *Earth, Planets and Space, 53,* 285-293.

Zhou, Y. (2009), Surface-wave sensitivity to 3-D anelasticity, *Geophys. J. Int., 178,* 1403-1410.

Figure Captions

Figure 1. Schematic illustration of possible melt generation and migration processes at general subduction zones, which may depend on the slab age, sediment thickness, subduction rate, and other factors. Thin lines with open arrows indicate melt paths, while those with solid arrows represent mantle flow lines.  VF stands for the volcanic front.

Figure 2. Segmentation of the Cascade arc volcanoes (red triangles), defined by geochemical signatures (Schmidt et al., 2008), and distribution of seismic stations used in

this study (white dots). The depth contours of the Juan de Fuca plate interface at 20-100 km are from the model of McCrory et al. (2004). The time-progressive Newberry (NB) and Yellowstone (YS) rhyolite eruptive progression across Snake River Plain (SRP) and High Lava Plains (HLP) are shown as thin brown lines (in Ma). CB stands for the Columbia basin.

Figure 3. Example of Empirical Green's Functions. (a) The lines connect the receivers (blue triangles) to the "virtual source" (red dot). (b-c) EGFs from the "source" to all the other stations derived from ambient noise cross-correlation of vertical-to-vertical components are plotted by the inter-station distance, filtered at 100-200 s and 10-25 s, respectively.

Figure 4. Observed EGFs vs. synthetics, filtered at periods of 10-25 s. (a) The lines connect the "source" (red triangle) to the receivers (blue triangles). (b-c) Comparison of observed waveforms (black lines) and synthetics (red lines), sorted by the distance between the "source" and each receiver. Synthetic waveforms are generated from the initial reference model (b) and our full-wave tomographic model (c), respectively.

Figure 5. Comparison of phase delay time between observed and synthetic waveforms from the initial reference model (black) and our full-wave ambient noise tomography (red). (a-b) The phase delay time versus inter-station distance at periods of 50-100 s and 25-50 s, respectively. (c-d) The histogram of phase delay time at corresponding period ranges. It shows that the improvement in fitting the data with our updated model.

Figure 6. Shear-wave velocity structure (in km/s) at multiple depths. All the panels share the same velocity scale, as denoted by the color bar. Other symbols are the same as in Figure 2.

Figure 7. Segmented low-velocity anomalies along the Cascade back-arc. (a) Horizontal slice at depth of 94 km (Vs in km/s). The black dashed lines outline the amplitude of largest negative Sp phase from receiver functions in the back-arc (Hopper et al., 2013). The magenta lines mark the profile locations in (b), (c), (d) and (e), respectively. The three white dots mark the point locations in Figure 8. All the panels share the same color bar. (b-d) W-E profiles across the back-arc anomalies. The y-axis has the approximate same length scale as the x-axis. The triangles mark the volcano centers. The Juan de Fuca plate interface at depths of 20-100 km from the model of McCrory et al. (2004) is projected. At greater depth, the plate interface is poorly defined. (e) S-N profile along the back-arc low-velocity anomalies, which spatially correlate with the three volcano clusters as in Figure 1. The length scale of y-axis is exaggerated two times of the x-axis.

Figure 8. Point velocity versus depth. (a) Comparison of velocities for points above the back-arc anomalies (blue) with the velocities of the Lau Basin (red) and the 0-4 Ma Pacific upper mantle (black) (based on Wiens et al. (2008)), and the calculated velocity for the Cascade geotherm (gray, the same as the black line in (b)). (b) Comparison of the observed and calculated shear-wave velocities. The blue lines are the observed back-arc velocities, the same as in (a). The other four lines are calculated (with the code from

Jackson and Faul (2010)) for the four geothermal profiles in (c) based on Figure 4 of Currie et al. (2004). In the legend of (c), the first number indicates the surface heat flow (in $mW/m^2$) and the second one is the potential temperature of the mantle (in $^oC$). The difference between the calculated velocity and the observed is suggestive of the presence of partial melt.

Figure 9. Resolution test for the low-velocity anomalies along the back-arc imaged in this study. (Upper panels) Input model is our preferred tomographic model; (Lower panels) Recovered velocities at the corresponding depths. Note that the amplitude of the velocity perturbation is not fully recovered partly because of the damping and smoothing factors used in the inversion.

Figure 10. Vertical resolution test along three W-E profiles that cross the back-arc low velocity anomalies. (Upper panels) Input model is our preferred tomographic model, same as in Figure 9. The three profiles correspond to the cross-sections in Figures 7b-7d, respectively. The velocity perturbation varies within a range of ±10%. (Lower panels) The recovered velocity models along the corresponding vertical profiles.