V_P and V_S structure of the Yellowstone hot spot from teleseismic tomography: Evidence for an upper mantle plume

Gregory P. Waite,[1,2] Robert B. Smith,[1] and Richard M. Allen[3]

Received 5 June 2005; revised 2 November 2005; accepted 28 December 2005; published 13 April 2006.

[1] The movement of the lithosphere over a stationary mantle magmatic source, often thought to be a mantle plume, explains key features of the 16 Ma Yellowstone–Snake River Plain volcanic system. However, the seismic signature of a Yellowstone plume has remained elusive because of the lack of adequate data. We employ new teleseismic *P* and *S* wave traveltime data to develop tomographic images of the Yellowstone hot spot upper mantle. The teleseismic data were recorded with two temporary seismograph arrays deployed in a 500 km by 600 km area centered on Yellowstone. Additional data from nearby regional seismic networks were incorporated into the data set. The V_P and V_S models reveal a strong low-velocity anomaly from ~50 to 200 km directly beneath the Yellowstone caldera and eastern Snake River Plain, as has been imaged in previous studies. Peak anomalies are −2.3% for V_P and −5.5% for V_S. A weaker, anomaly with a velocity perturbation of up to −1.0% V_P and −2.5% V_S continues to at least 400 km depth. This anomaly dips 30° from vertical, west-northwest to a location beneath the northern Rocky Mountains. We interpret the low-velocity body as a plume of upwelling hot, and possibly wet rock, from the mantle transition zone that promotes small-scale convection in the upper ~200 km of the mantle and long-lived volcanism. A high-velocity anomaly, 1.2% V_P and 1.9% V_S, is located at ~100 to 250 km depth southeast of Yellowstone and may represent a downwelling of colder, denser mantle material.

Citation: Waite, G. P., R. B. Smith, and R. M. Allen (2006), V_P and V_S structure of the Yellowstone hot spot from teleseismic tomography: Evidence for an upper mantle plume, *J. Geophys. Res.*, *111*, B04303, doi:10.1029/2005JB003867.

1. Introduction

[2] The Yellowstone Plateau volcanic field in northwestern Wyoming, a region associated with the extensive geysers and hot springs of Yellowstone National Park, is the youngest manifestation of the Yellowstone hot spot. The track of the hot spot extends 800 km across the northern basin-range province (Figure 1). This track of bimodal basaltic-rhyolitic volcanism is considered the result of southwest movement of the North America Plate across a mantle magma source. Yellowstone's mantle source has often been attributed to a mantle plume [e.g., *Morgan*, 1972], but this model has remained equivocal partly because there have not been adequate seismic data to resolve the volcanic system's mantle structure.

[3] The mantle heat source has produced three caldera-forming explosions at Yellowstone, as well as numerous lava flows that have erupted 6000 km^3 of lava in the past 2 million years [*Christiansen*, 2001]. Within the youngest 0.64 Ma, 3000 km^2 caldera, high heat flow (averaging more than 1700 mW/m^2 [*Blackwell*, 1969]), a ~60 mGal gravity low [*Lehman et al.*, 1982], and a low (−8% to −15%) V_P body in the upper crust beneath the caldera [*Benz and Smith*, 1984; *Miller and Smith*, 1999; *Husen et al.*, 2004] suggest an upper crustal magma body that fueled Yellowstone volcanism and drives the hydrothermal system.

[4] Yellowstone is also the most seismically active area of the 1300-km-long Intermountain Seismic Belt, which stretches from Montana to Arizona. Earthquake swarms [e.g., *Waite and Smith*, 2002] and episodes of crustal uplift and subsidence [*Pelton and Smith*, 1982; *Wicks et al.*, 1998; *Puskas et al.*, 2002] are common at Yellowstone. Seismicity at Yellowstone also includes the largest historical earthquake of the basin-range province, the M_S 7.5 1959 Hebgen Lake, Montana earthquake. The earthquakes and crustal deformation result from the interaction of regional tectonics with the magmatic system.

[5] Beginning from the youngest, 0.64 Ma caldera of the Yellowstone Plateau volcanic field, a line of progressively older silicic eruptive centers extends SW along the eastern Snake River Plain (ESRP) to the 16 Ma McDermitt volcanic field on the Oregon-Nevada border [*Christiansen and Yeats*, 1992]. Ashfall deposits analyzed by *Perkins and Nash* [2002] suggest there were 142 caldera-forming eruptions along the track of the hot spot. The rate and direction of the progression of the hot spot across the southwesterly moving North America plate are consistent with a persistent, relatively stationary, sublithospheric source.

[1] Department of Geology and Geophysics, University of Utah, Salt Lake City, Utah, USA.
[2] Now at U.S. Geological Survey, Menlo Park, California, USA.
[3] Seismological Laboratory, Department Earth and Planetary Science, University of California, Berkeley, California, USA.

Copyright 2006 by the American Geophysical Union.
0148-0227/06/2005JB003867$09.00

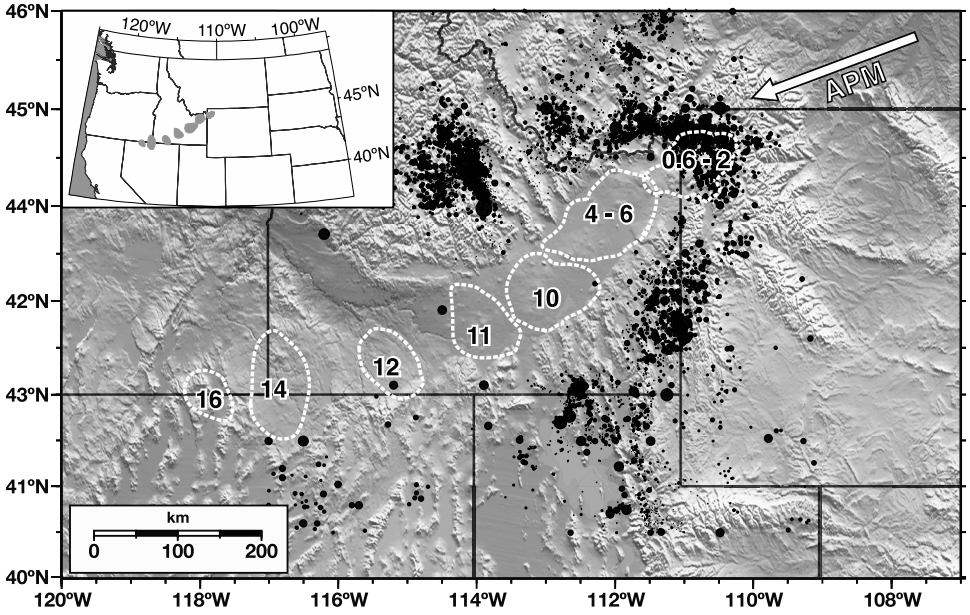

Figure 1. Yellowstone–eastern Snake River Plain volcanic system with earthquake epicenters (black circles) and topography to show the surrounding parabolic pattern of seismicity and high topography. Approximate ages of silicic volcanic centers are noted in Ma. The dashed white lines outline the locations of the eruptive centers. State boundaries are plotted for reference. The direction of absolute plate motion (APM) from *Gripp and Gordon* [2002] is shown with a white arrow. Inset shows the location of the study area in the western United States.

[6] Yellowstone has been viewed as the archetypical continental hot spot because of characteristics that suggest a mantle source: (1) the well-defined track of progressively older silicic volcanism in the direction of plate motion; (2) a parabolic pattern of high topography (>1000 m) and seismicity around the ESRP with its apex at the Yellowstone plateau [*Smith and Sbar*, 1974; *Anders et al.*, 1989; *Pierce and Morgan*, 1992; *Smith and Braile*, 1994]; (3) high $^3He/^4He$ ratios suggestive of an anomalous mantle source [e.g., *Craig et al.*, 1978]; and (4) a 10 to 12 m positive geoid anomaly with a ~500 km radius, comparable to that of Hawaii, centered at Yellowstone [*Smith and Braile*, 1994]. These observations have often been attributed to a mantle plume beneath Yellowstone [e.g., *Morgan*, 1972; *Smith and Sbar*, 1974; *Anders and Sleep*, 1992; *Bijwaard et al.*, 1998; *Pierce and Morgan*, 1992; *Steinberger*, 2000].

[7] Plumes were first defined as hot upwellings of relatively primordial material that rise from a thermal boundary layer, presumed to be the core-mantle boundary [*Morgan*, 1971]. Some researchers have suggested there is not definitive evidence for or against a deep-mantle plume beneath Yellowstone [e.g., *Humphreys et al.*, 2000; *Smith and Braile*, 1994], while others have argued against a plume source [e.g., *Hamilton and Myers*, 1966; *Favela and Anderson*, 2000; *Christiansen et al.*, 2002]. *Christiansen et al.* [2002] point to observations that they believe are not consistent with a deep-mantle plume source for Yellowstone. For example, they suggest that such a mantle plume does not explain the persistence of basaltic volcanism along the hot spot track hundreds of km from the present location of the hot spot. *Christiansen et al.* [2002] also note that preexisting lithospheric structures coincidentally parallel the ESRP and may explain the propagation of the hot spot. They also question the assumption of a deep, primordial source for ^3He to explain the high $^3He/^4He$ ratios in light of research that casts doubt on the assumption [e.g., *Anderson*, 2000; *Meibom et al.*, 2003].

[8] The northwestward progression of volcanism associated with the Newberry system in Oregon has been cited as evidence against a plume [e.g., *Christiansen et al.*, 2002]. While the Newberry volcanic track is not as distinct as the ESRP, it originates near the start of the Yellowstone hot spot track at about the same time. The age-progressive volcanism of the Newberry trend is not consistent with a stationary mantle source beneath a moving plate. However, some models suggest the Newberry trend is a result of spreading of a Yellowstone plume head [*B. T. Jordan et al.*, 2004; *Camp and Ross*, 2004], possibly aided by corner flow associated with the subducting Juan de Fuca slab [e.g., *Humphreys et al.*, 2000]. In these models, the Newberry trend does not contradict the notion of a plume source for Yellowstone.

[9] Perhaps the most compelling evidence against a mantle plume source for Yellowstone has been the lack of a clear seismic image of a plume beneath Yellowstone or the ESRP [*Iyer et al.*, 1981; *Evans*, 1982; *Dueker and Humphreys*, 1990; *Humphreys and Dueker*, 1994a, 1994b; *Saltzer and Humphreys*, 1997; *Christiansen et al.*, 2002; *Schutt and Humphreys*, 2004]. While these studies show a low-velocity anomaly to at least 200 km beneath Yellowstone and the ESRP, limitations of the regional data set prevent resolution of deeper anomalies.

[10] On the other hand, *Bijwaard et al.* [1998] and *Montelli et al.* [2004] image a low *P* wave velocity anomaly beneath Yellowstone and extending westward to at least 650 km depth in global seismic tomography models.

Bijwaard et al. [1998] interpret the anomaly as a plume. *Montelli et al.* [2004], seeing no evidence for continuation of the anomaly through the lower mantle to the core-mantle boundary do not interpret the Yellowstone anomaly as a plume. The regional and global seismic tomography studies agree that the Yellowstone hot spot has a shallow, <300 km, upper mantle low-velocity anomaly on the order of −2% to −5% for V_P and V_S. However, these studies have some uncertainty in the depth extent of the anomaly. In addition, there has not been a regional-scale mantle shear wave tomography study of Yellowstone.

[11] Since *Morgan* [1971], researchers have defined plumes variously based on seismic, chemical and thermal properties. The lack of consistency in defining plumes has led to confusion. Catalogs of proposed plume-induced hot spots vary depending on what criteria are used to define the plume [e.g., *Courtillot et al.*, 2003] and numerical modeling reveals a range of possible plume shapes and sizes [*Farnetani and Samuel*, 2005]. In order to avoid similar confusion in this paper we use a general plume definition that does not define the source depth, size, or chemical composition: a plume is a near-vertical, approximately axisymmetric, buoyant upwelling of hot or wet material. We are not able to resolve lower mantle seismic velocities with this study, so we cannot address the proposed core-mantle boundary source of a Yellowstone plume. Likewise, the term hot spot refers to a persistent mantle melting anomaly that produces concentrated volcanism, but does not presuppose a deep mantle origin.

[12] The continental setting of Yellowstone affords a unique opportunity to study a hot spot with a large-scale passive-source seismology experiment. We modeled teleseismic delay time data from a 500 km by 600 km array of 86 broadband three-component seismograph stations. The average station spacing is 50 km NW–SE and 75 km NE–SW. The array aperture and station spacing provide higher resolution in the upper mantle than global tomographic models and over a larger area and to greater depths than has been imaged in the previous regional studies. These data facilitate the imaging of a continuous low *P* and *S* wave anomaly that extends from Yellowstone Plateau to the top of the mantle transition zone [*M. Jordan et al.*, 2004; *Smith et al.*, 2003; *Waite et al.*, 2003; *Yuan and Dueker*, 2005].

2. Traveltime Data for the Yellowstone Hot Spot Experiment

2.1. Earthquake Data Selection and Processing

[13] Seismic data used in this project came from four networks: two temporary deployments of broadband seismographs deployed specifically for this Yellowstone hot spot study, broadband stations in the University of Utah Seismograph Stations (UUSS) permanent network in Yellowstone and Utah, and broadband seismic stations in the United States National Seismograph Network (USNSN). Data from six USNSN and four UUSS broadband stations were examined for time period 1999 to 2003. The temporary seismic arrays deployed for this study operated for 11 months each. The first array, near Billings, Montana, 200 km NE of the Yellowstone caldera (hereafter referred to as the Billings array), was deployed from August 1999 through August 2000 (Figure 2). The second array, abbreviated YISA for Yellowstone–Intermountain Seismograph Array, consisted of 47 stand-alone stations distributed over a 500 km by 600 km area that encompassed Yellowstone deployed from June 2000 through April 2001. To take advantage of the seismic ray paths from persistent earthquakes in the northwest direction (northern and western Pacific subduction zones) and to the southeast (Central and South America subduction zones) of Yellowstone, the seismograph stations were deployed in six parallel NW-SE oriented lines. Approximately 70 percent of the earthquakes recorded during the deployment were in these source regions (Figure 3). Most of the remaining data came from the Tonga region, SW of the array.

[14] Event gathers generated from all $M \geq 6.0$ earthquakes recorded on the arrays were inspected for clear *P*, *PKiKP*, *S*, and *SKS* arrivals. In addition, earthquakes with $5.6 \leq M < 6.0$ from regions not sampled by good quality $M > 6$ earthquakes were also used to identify and examine arrivals. For the *P* wave velocity model, 169 earthquakes with useable phases were recorded, but many events had roughly the same locations. This leads to redundant ray paths beneath the array that may bias the solution toward a particular portion of the model. For example, the large number of rays with NW backazimuths could bias anomalies along those ray paths with respect to other directions.

[15] We removed selected events from the data set to reduce the redundant data and produce a more evenly distributed set of rays. In particular, this effort was aimed at deriving an approximately even number of rays from both NW and SE back azimuths. Epicenters were sorted into 1° by 1° bins and events from bins with multiple events were reexamined to identify the event with the most clear, impulsive arrivals. We removed the other event(s) in the bins from the data set. In addition, earthquakes with fewer than 10 clear arrivals were removed. Most of the events that were removed were from NW back azimuths. This process was performed separately for the *P* and *S* wave data sets. Figure 3 shows the locations of all the epicenters used in the V_P and V_S inversions. In addition to removing redundant earthquakes, data from 14 selected stations that were part of the densely spaced Billings array NE of Yellowstone were removed to reduce the much higher ray density in that area.

[16] All the arrivals were picked by hand and traveltime residuals were computed relative to the IASP91 velocity models [*Kennett and Engdahl*, 1991]. We optimized the hand-picked data with the cross-correlation method of *VanDecar and Crosson* [1990]. The median pick uncertainties estimated from the cross correlation are 0.019 s for *P* and *PKiKP*, and 0.024 s for *S* and *SKS* picks. These uncertainties may be too optimistic, however, as estimates of pick errors from the initial picking are almost an order of magnitude larger. The uncertainty has a strong inverse correlation with the cross-correlation coefficient. Picks with low cross-correlation coefficients, <0.65, were reexamined and were either removed or assigned a larger error than that estimated by the cross-correlation algorithm. In most cases, the waveforms are similar and cross-correlation coefficients are 0.8 and above. A total of 3399 *P* phases and 380 *PKiKP* phases from 115 earthquakes were used in the V_P inversion. The *S* wave velocity model used 1705 *S* phases and 459 *SKS* phases from 79 earthquakes.

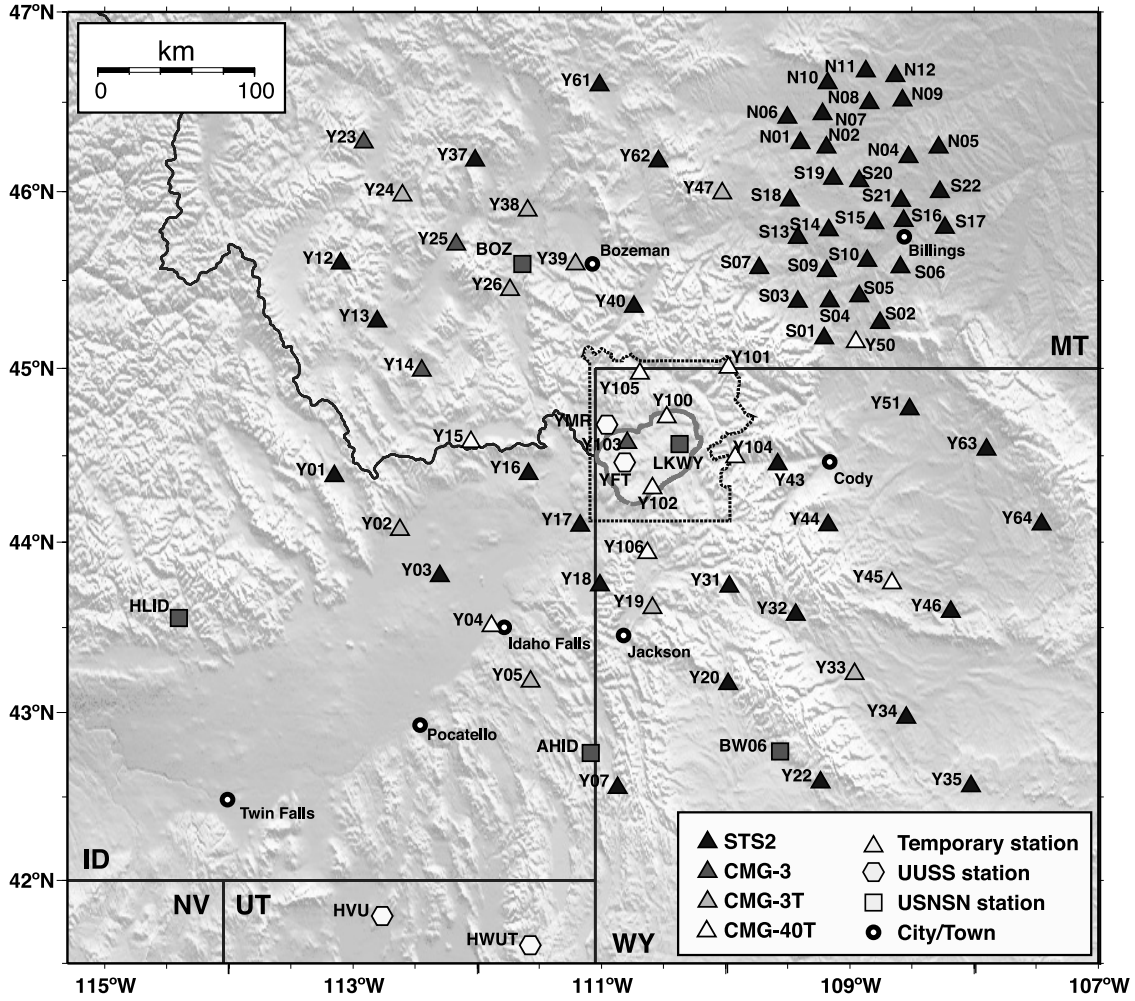

Figure 2. Seismographs used in the study. Billings array stations are the tight array in the NE. The Yellowstone–Intermountain Seismic Array (YISA) array consists of five lines of stations oriented NW-SE. Additional stations are part of USNSN and UUSS permanent networks. Symbols indicate station owner and are shaded by sensor type. State boundaries, Yellowstone National Park boundary, selected cities, and the 0.6 Ma caldera are shown for reference.

2.2. Limitations of Ray-Theoretical Tomography

[17] We expect that there is little difference between results obtained with our 1-D ray tracing and 3-D ray tracing [e.g., *Saltzer and Humphreys*, 1997; *Yuan and Dueker*, 2005]. However, the use of ray theory will underestimate the true amplitude of seismic anomalies by ignoring the Fresnel volume. The width of the Fresnel volume depends on the total distance between the source and receiver, L, the distance from the source, d, and the wavelength, λ. The variable f, is given by *Spetzler and Snieder* [2004] as

$$f = 2\left[\frac{\lambda d(L-d)}{L}\right]^{\frac{1}{2}}. \quad (1)$$

The wavelengths of teleseismic P and S waves used in this study are about 20 km given periods of 2 and 4 s, respectively, and upper mantle velocities. For a ray path length of 10^4 km, the maximum Fresnel width (at the midpoint) is 450 km. This is a factor of 22 times larger than the wavelength. For smaller (or larger) d, the Fresnel width is smaller. For example, at the 410 km discontinuity, a ray with an incidence angle of 40° from vertical will be about d = 550 km away from the receiver. This gives $f \approx 200$ km for the same λ and L as above, so anomalies much smaller than about 200 km wide may not expected to be well resolved at the base of the upper mantle (410 km depth). At 200 km depth, $f \approx$ 140 km.

[18] The effects of wavefront healing, the diffraction of a wavefront around a low-velocity anomaly [*Wielandt*, 1987], are also unaccounted for with ray theory. *Nolet and Dahlen* [2000] found that anomalies can be resolved with high-frequency rays when $l/h \ll \pi h/\lambda$, where l is the distance from the anomaly, h is the half width of the anomaly, λ is the wavelength of the seismic wave and the source is at infinity. The amplitude of the recovered anomaly will be decreased as d increases. The wavefront will "heal" (i.e., the traveltimes will not be delayed) when $l/h \gg \pi h/\lambda$ For teleseismic wavelengths on the order of 20 km, an anomaly with 100 km half width at 400 km depth (l = 550 km for a

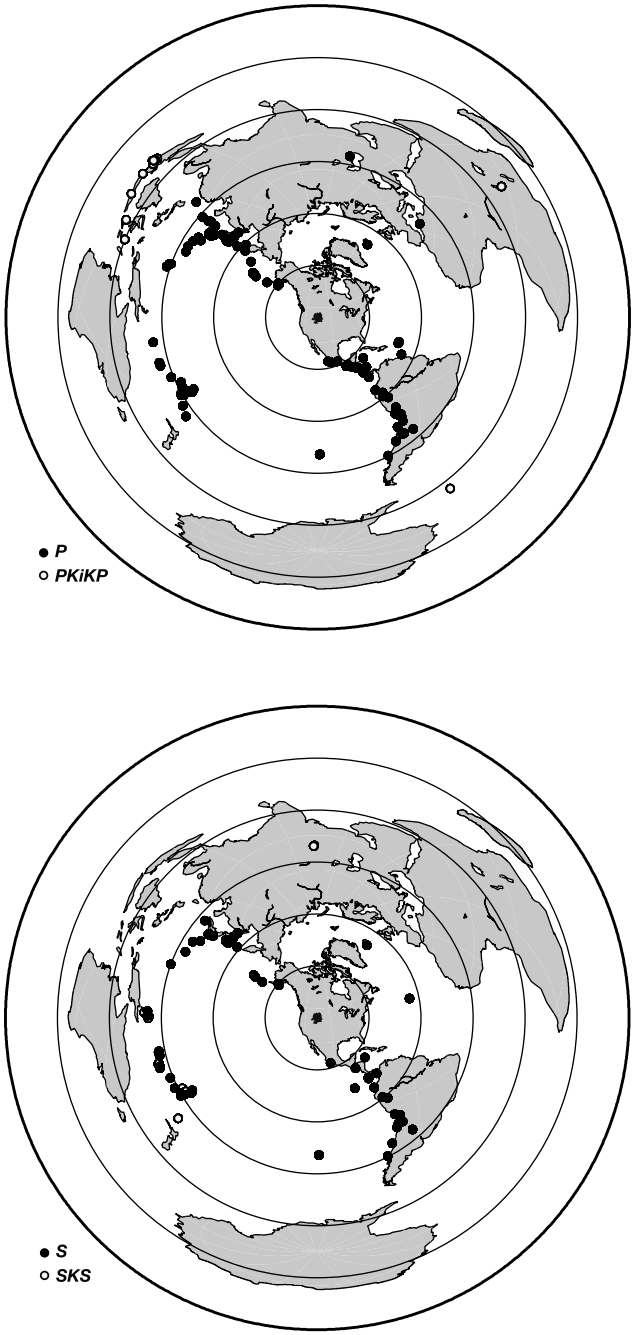

Figure 3. Distribution of (top) 115 earthquakes used in the V_P inversion and (bottom) 79 earthquakes used in the V_S inversion. Events are clustered at NW, SE, and SW azimuths. The study area is at the center of the plots.

ray with 40° incidence angle) gives $l/h = 5.5$ and $\pi h/\lambda = 15.7$. Therefore most of the traveltime delay due to the anomaly should be preserved. The amplitude of smaller volume or deeper anomalies will be underestimated, however.

[19] Given the Fresnel width and wavefront healing considerations, anomalies in the bottom of the upper mantle (300–400 km) will not be well resolved unless they have a width of about 200 km. Anomalies at least 140 km wide should be resolvable at 200 km depth. The amplitudes of the anomalies that are imaged will be underestimated because wave effects are neglected.

2.3. Crustal Structure Corrections

[20] Synthetic modeling has shown that crustal anomalies can be mapped into the uppermost mantle in teleseismic tomography studies because ray paths through the crust are near vertical [*Waldhauser et al.*, 2002]. Variations in the depth of the Moho velocity discontinuity represent one source of this type of error. In the Yellowstone study area, the Moho is as shallow as 30 km in the basin-range at southwest end of the model and almost 50 km deep in the Archean cratonic NE corner. A 20 km difference in the Moho depth corresponds to ~0.6 s difference in the traveltime of a vertical *P* wave and a difference of nearly 0.9 s for a vertical *S* wave.

[21] Traveltime residuals were corrected using the global CRUST2.0 model (the model can be found online at http://mahi.ucsd.edu/Gabi/rem.html) [*Bassin et al.*, 2000]. This global model of crustal structure, topography, and bathymetry has a 2° grid spacing which only accounts for large-scale variations, but the Moho depth in the CRUST2.0 model is estimated to be accurate to ±5 km. We linearly interpolated the model to 0.25° to smooth the Moho. Individual ray parameters and velocities were used to compute the ray paths and traveltimes through the CRUST2.0 and IASP91 models to the station elevations. The Moho and elevation corrections reduced the initial *P* wave traveltime RMS by 7% from 0.49 s to 0.45 s, and the initial *S* wave traveltime RMS was reduced 3% from 1.50 s to 1.45 s.

[22] While large velocity anomalies in Yellowstone have been found using local earthquake tomography [*Benz and Smith*, 1984; *Lynch*, 1999; *Miller and Smith*, 1999; *Husen et al.*, 2004], the anomalies are well resolved to a depth of only ~10 km and cover a relatively small volume beneath the Yellowstone Plateau volcanic field. *Lynch's* [1999] V_P model extends from the Intermountain region into the ESRP, but covers only a third of the teleseismic array area. Since the teleseismic rays are essentially vertical through the upper 15 km of the model, all the rays to a given station follow the same path and have the same delay due to shallow crust anomalies. Instead of using an incomplete upper crust model to correct for these anomalies, station correction parameters were included in the inversion scheme. Synthetic tests with 30-km-thick, synthetic anomalies having 10% velocity perturbations demonstrate that the station correction parameters effectively account for these delays.

2.4. Delay Times

[23] Delay times across the network and across azimuths at individual stations give an indication of the distribution of velocity anomalies beneath Yellowstone. The polar plots in Figure 4 show *P* phase and *S* phase relative delays at representative stations after the Moho and elevation corrections have been applied. The largest positive delays of 1.9 s for *P* and 9.5 s for *S* are at station LKWY within the Yellowstone caldera. Stations on the ESRP also have delays of over 1 s for *P* and 2–4 s for *S*. Stations NW of the ESRP and Yellowstone generally have positive delays for arrivals from the SE (<1 s for *P* and *S*), but negative delays (<−1 s

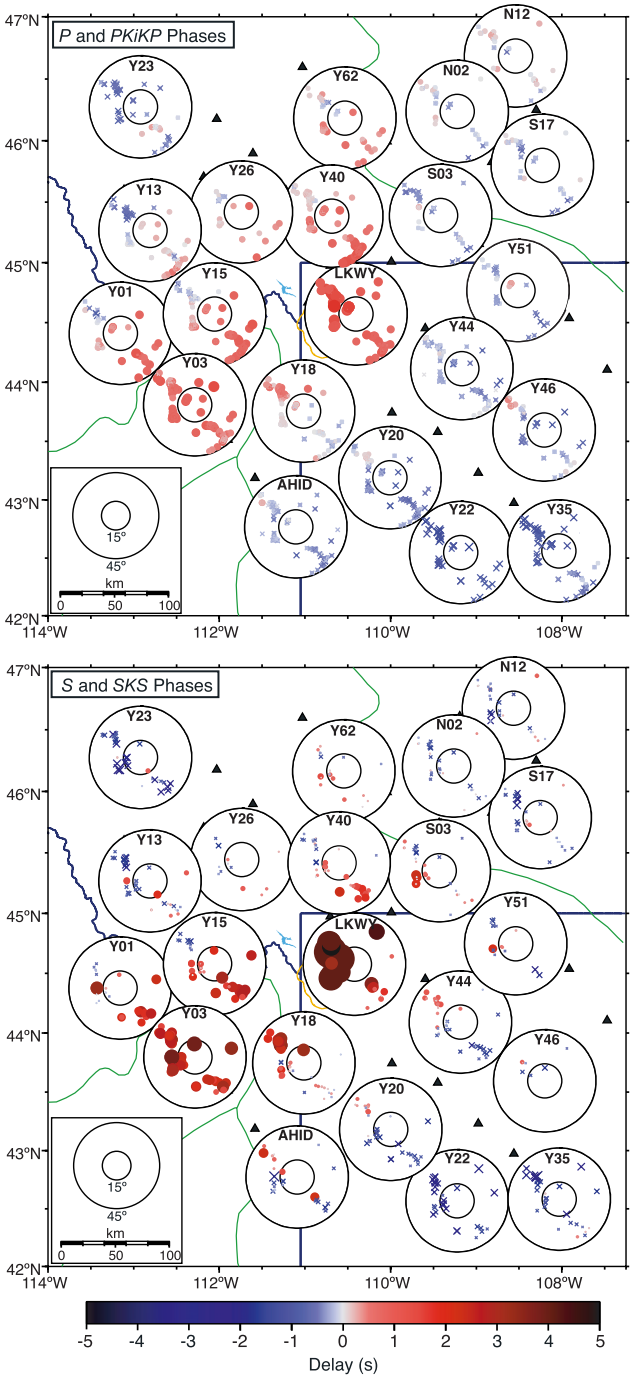

Figure 4. *P* and *S* wave delays for representative stations indicating low velocities beneath Yellowstone and the eastern Snake River Plain. Delays are plotted by back azimuth and incidence angle at 200 km. Positive delays are plotted with red circles, and negative delays are blue crosses. The color scale applies to both (top) *P* and (bottom) *S* delays.

for *P* and *S*) for arrivals from other back azimuths. Similarly stations SE of Yellowstone have positive delays for events from the NW, but negative delays for events from the SE. Positive delays for near-vertical arrivals at stations NW of Yellowstone (i.e., Y26, Y38, Y40, Y62) indicate low velocities below this area as well. While arrivals from the SE have positive delays, steep incidence arrivals also have positive delays indicating low velocities directly below the stations. Stations in the SE part of the array have the largest negative delays of up to −1.8 s for *P* and −4 s for *S*. Correcting for the 50 km deep Moho in the NE part of the array reduces the negative delays there significantly.

[24] Delays from phases of similar incidence angle and back azimuth are generally consistent at each station across the network. The *S* phase picks are less consistent than *P* phase picks, however. Some of the inconsistencies are due to the subset of stations that recorded a given event. For example, the large delays at station LKWY inside the caldera (Figure 4) influence the mean delay for an event. Since the mean is removed, the same event recorded across the network with or without LKWY, will have slightly different traveltime residuals. These differences are accounted for in the inversion by using event correction terms.

3. Traveltime Inversion

[25] The isotropic velocity perturbations are solved relative to the 1-D IASP91 models [*Kennett and Engdahl*, 1991] using the linearized inversion described by *Nolet* [1993], *Allen et al.* [2002], and *Waite* [2004]. The model grids for the V_P and V_S inversions extend ~1000 km in all three dimensions with either 30 or 50 km grid node spacing. The specific model parameterizations are detailed below. The model space is larger than the volume in which structures will be resolved to ensure anomalies are not compressed into the model.

[26] We solve the system of equations

$$\mathbf{Ax} = \mathbf{d}, \tag{2}$$

where **d** is the vector of traveltime residuals, *dt*, **A** is the matrix of ray path data, and **x** is the vector of model updates. A smoothing matrix, **S**, with weights decreasing linearly in a spherical volume of some radius, *r*, is incorporated to require that the model is smooth: $\mathbf{x} = \mathbf{Sy}$. Substituting into equation (2), we have $\mathbf{ASy} = \mathbf{By} = \mathbf{d}$, where **B** is the smooth matrix used for inversion. Following inversion, the smooth model is reconstructed from $\mathbf{x} = \mathbf{Sy}$.

[27] Arrival time picks' uncertainty estimates were calculated with a cross-correlation algorithm [see *Allen et al.*, 2002] and are included in the diagonal data covariance matrix, $\mathbf{C_d}$. The addition of station and event correction parameters to the model vector requires a diagonal model covariance matrix, $\mathbf{C_m}$, of a priori model parameter weights to scale the matrix **B**, so that the magnitudes of the free parameters are similar in the inversion. These values represent expected relative variations in the velocity perturbations and corrections.

[28] The covariance matrices are applied giving

$$\mathbf{C_d B C_m z} = \mathbf{C_d d}, \tag{3}$$

making the final model vector, **z**, nondimensional. The model covariance matrix is applied to the model after inversion to obtain the true values, $\mathbf{y} = \mathbf{C_m z}$. Starting with $\mathbf{Bz} = \mathbf{d}$, where the covariance matrices have been applied,

the system of equations is solved by minimizing the least squares misfit function, with the LSQR algorithm [*Paige and Saunders*, 1982] after modification to include damping, λ:

$$\| \mathbf{Bz} - \mathbf{d} \|^2 + \lambda \| \mathbf{z} \|^2 . \quad (4)$$

Several combinations of model grid spacing (30, 40, 50, 60 km), smoothing lengths (0 up to 90 km) and damping were tested to explore the sensitivity of the inverted model velocity perturbations to model parameterization. Synthetic data and real data were used in these tests. The principal features of the model are persistent in every solution, although the amplitudes vary by up to a few percent in the V_S models. Many small features are inconsistent and are not interpreted. As expected, smoothing tends to reduce the amplitude of small volume anomalies, but spreads them out over a larger volume.

[29] An alternative type of smoothing uses grid offset and averaging [*Evans and Zucca*, 1988]. Two grids are used in this procedure: a coarse grid for inversion, and a second fine grid with spacing some fraction of the inversion grid. The latter grid is used to shift the inversion grid. The grids are not shifted vertically. The procedure is as follows: inversion is performed using a coarse model grid; the grid is then shifted horizontally 10 km (e.g., to the east) and the inversion is performed with this new coarse model grid. The shifting and inversion is continued until a coarse grid node has occupied the each node of the fine grid. Finally, the average value of each node in the fine grid is computed from the value of the velocity at each of the 25 fine grid nodes that surround it in a 5 node by 5 node square.

[30] We present results of inversion with both linear smoothing and multimodel average for comparison. The models with linear smoothing imposed in the inversion have grid node spacing of 30 by 30 by 30 km and 70 km smoothing in both the horizontal and vertical directions. The smoothed V_P and V_S models are designated S30VP and S30VS, respectively. The offset-and-average models have a coarse grid spacing of 50 km by 50 km in horizontal and the fine grid node spacing is 10 km by 10 km in horizontal. Both grids have 50 km spacing in the vertical as discussed above. The offset-and-average V_P and V_S models are called OSA50VP and OSA50VS. Including two models for both the V_P and V_S inversions is useful for interpreting the results. For example, higher confidence is afforded to anomalies that are consistent between the models.

[31] Higher damping was used for the OSA models than for the corresponding S30 models because there were fewer grid nodes in the OSA models. Similarly, the model covariance values were chosen so that the station and event corrections computed with the two methods would be equivalent. The station corrections computed in the inversions vary from -0.8 to 0.7 s (-1.2 to 1.5 s) in the V_P (V_S) models. Event corrections vary from -0.2 to 0.2 s (-0.5 to 0.8 s) in the V_P (V_S) models. The final RMS and total variance reduction of the corresponding OSA and S30 models for V_P and V_S are equivalent (see results below). While the damping and weighting of station and event corrections contributes to the variations in the structure and amplitude of the seismic anomalies in the models, we attribute most of the differences between the models to the different types of smoothing employed.

4. Results of the *P* and *S* Wave Tomographic Inversion

[32] The P and S wave velocity models were solved independently as described above. The V_P and V_S models show strong low-velocity anomalies in the upper 200 km beneath Yellowstone. In addition, the V_P and V_S models have a smaller-amplitude low-velocity anomaly extending from 250 km depth to the top of the midmantle transition zone $\sim$100 km NNW of the caldera. The locations of the lower V_P and V_S anomalies are slightly different.

4.1. *P* Wave Velocity Structure

[33] The V_P models are constructed from 3779 P and *PKiKP* rays and traveltimes. The initial RMS residual is 0.45 s and the final RMS residual is 0.17 s for both the S30VP and OSA50VP models. This is an order of magnitude higher than the pick uncertainty estimate after cross correlation and roughly equal to the estimated median uncertainty in the handpicked data. The data variance reductions are 77% for S30VP and 78% for OSA50VP.

[34] Plots of the ray density through slices of the S30VP model are shown in Figure 5. These plots give a rough estimate of the model resolution since they do not take into account the orientation of the rays, but they provide a way to quickly estimate areas of good and poor data coverage. For example, note the high density of rays beneath the Yellowstone caldera and the Billings array. The predominance of rays arriving from NW and SE back azimuths is demonstrated by the volumes of high ray density to the NW and SE of the caldera. In the 90 km depth slice, there is a high density under the NW-SE lines of stations, but low density in between the lines. This results in less resolution in the NE-SW direction at shallow depths, but the effect is smaller in deeper parts of the model.

[35] The ray density plots do not demonstrate the vertical resolution problem inherent in this type of regional teleseismic tomography study [see, e.g., *Keller et al.*, 2000; *Wolfe et al.*, 2002a, 2002b]. The angles of the incoming rays in the middle upper mantle ($\sim$200 km) are between $\sim$45° and vertical. The ray paths through the model to a given station define a cone that opens with depth. There are crossing rays in the middle of the model down to at least the 410 km discontinuity, so reasonable resolution is expected to about that depth. Smearing is expected to be strongest at the sides and bottom of the model where the rays are parallel. Resolution information obtained from the synthetic tests described below is more useful.

[36] The P wave models (Figure 6) are dominated by a tilted low V_P anomaly that extends from directly beneath Yellowstone through the upper mantle to the 410 km discontinuity 100 km WNW of Yellowstone. The anomaly has peak amplitudes of -2.0% (S30VP) and -2.3% (OSA50VP) above 200 km and -1.0% (S30VP and OSA50VP) from 250 to 400 km depth. The shallow portion of the anomaly continues down the ESRP to the SW but decreases in amplitude. It is roughly the width of the ESRP.

[37] *Schutt and Humphreys* [2004] used a similar teleseismic tomography technique to image the upper mantle

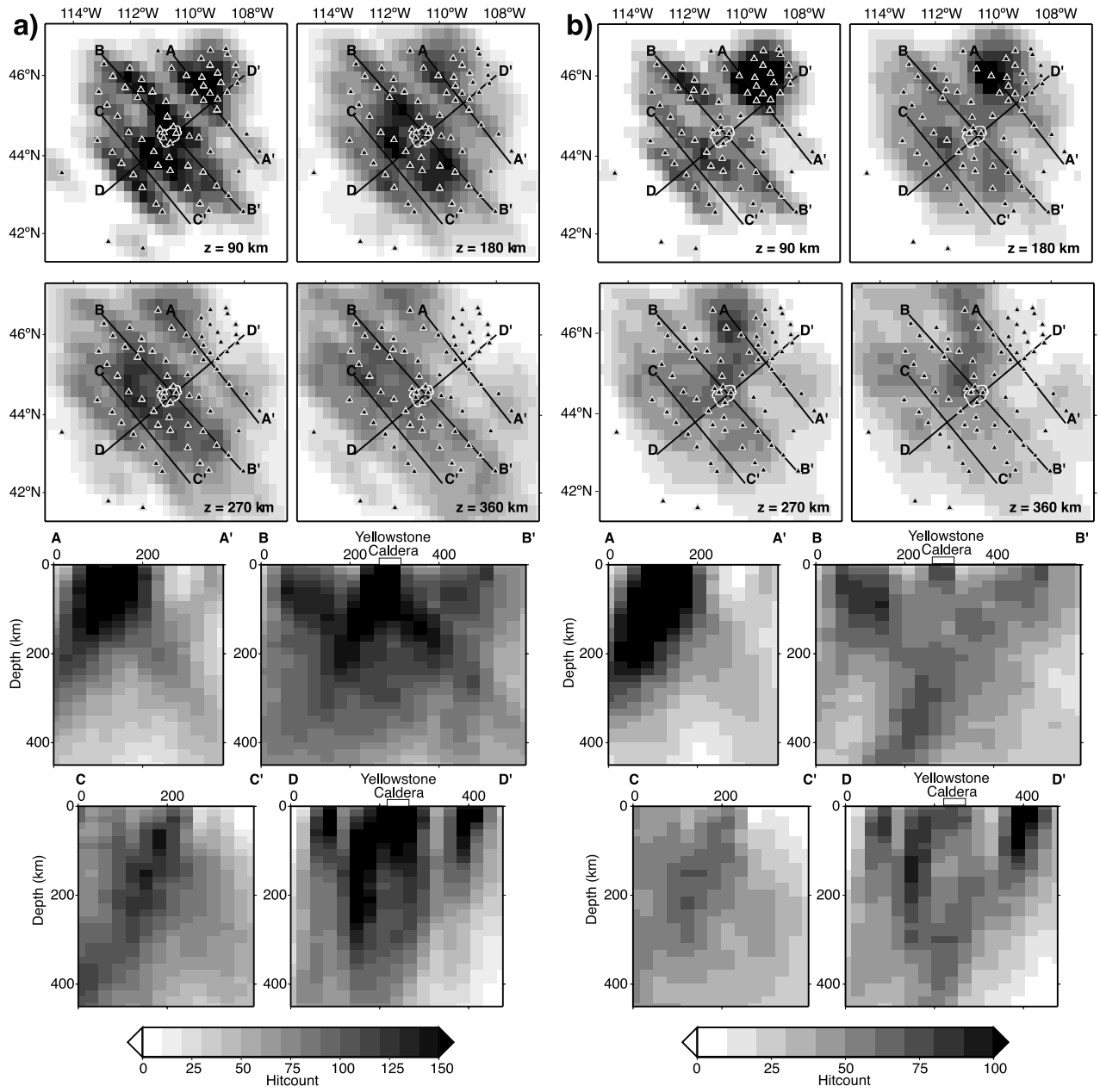

Figure 5. Ray density plots for the (a) S30VP and (b) S30VS models showing four vertical and four horizontal slices through the model. Horizontal slices are shown at 90, 180, 270, and 360 km. The locations of the vertical slices are noted on the horizontal slices. The Yellowstone caldera is shown with a white line in the horizontal slices and a white box at the top of the B-B′ and D-D′ cross sections. Seismograph stations are shown on the horizontal cross sections as triangles.

under the ESRP, ~100 km SW of the YISA array. They found a low-velocity anomaly in the upper 200 km directly beneath the ESRP, and high-velocity anomalies on the flanks of the ESRP. While the low-velocity anomaly is continuous up to the Yellowstone caldera, the high-velocity features are not. There is a 1.1% (S30VP) and 1.5% (OSA50VP) high V_P anomaly in the SE part of the models, but only discontinuous high-velocity anomalies NW of the ESRP.

[38] Our results reliably image a continuous *P* wave velocity anomaly through the upper mantle, which previous studies had not done [*Iyer et al.*, 1981; *Dueker and Humphreys*, 1990; *Humphreys and Dueker*, 1994a, 1994b; *Christiansen et al.*, 2002]. An important difference between our study and the previous studies is our consistent data set collected with a wide aperture array of broadband stations. The previous studies all used some variation of data collected by *Iyer et al.* [1981] in the late 1970s. We tested this data set with our methodology and found results similar to those of previous studies; the tomographic inversion did not clearly reveal structure below about 200 km. Data from the 500 by 600 km array of digital broadband seismographs used in this study made the imaging of the deeper part of the low-velocity anomaly possible. We note that *Yuan and Dueker*

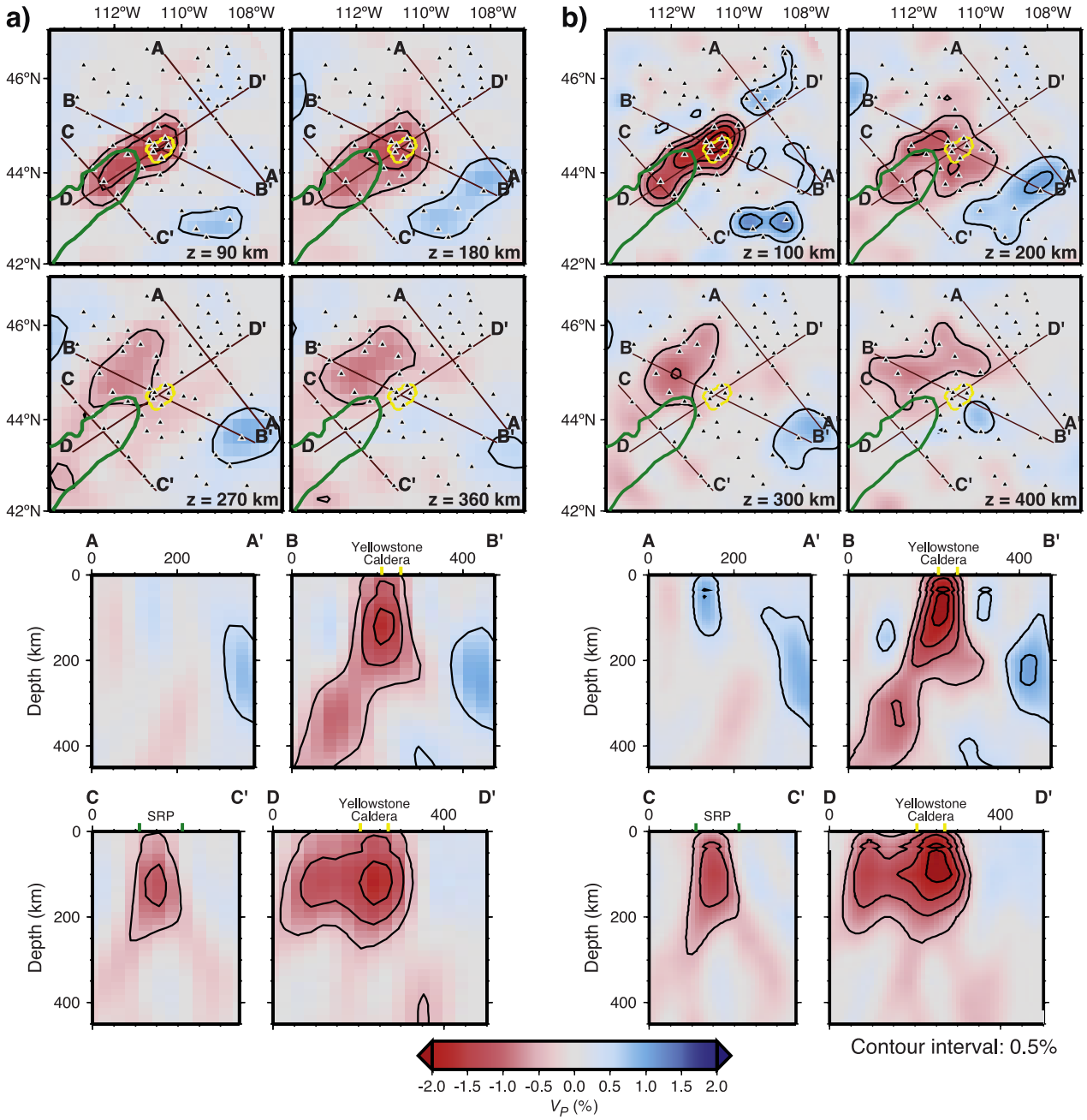

Figure 6. Slices through the V_P models (a) S30VP and (b) OSA50VP plotted along with station locations, the outline of the ESRP, and the outline of the 0.6 Ma Yellowstone caldera. Note that depth slices are slightly different for each model because of different grid node locations. The cross sections are taken through similar locations to those in Figure 5, except B-B′, which highlights the WNE plunging low-velocity anomaly.

[2005] found a similar low V_P anomaly using essentially the same data we used.

4.2. *S* Wave Velocity Structure

[39] The V_S models result from the inversion of 2164 *S* and *SKS* traveltimes. The initial RMS residual is 1.45 s, a factor of three higher than the initial V_P model RMS. The final RMS residuals are 0.64 s for model S30VS and 0.62 s for model OSA30VS. Data variance was reduced 70% (S30VS) and 73% (OSA50VS). Figure 5b shows ray density through slices of the S30VS model. This plot demonstrates the high density of rays beneath the Billings array in the NE part of the model. A gap in density between the NW-SE lines of stations is evident at shallow depths. Both of these characteristics are similar to the S30VP model ray density shown in Figure 5a. The high density of rays directly beneath the caldera in Figure 5a, however, is noticeably absent in Figure 5b. This is a result of the difficulty in picking *S* and *SKS* waveforms at stations inside the caldera. Many arrivals at stations LKWY, Y100, Y102, Y103, and YMR, which are inside, or adjacent to, the

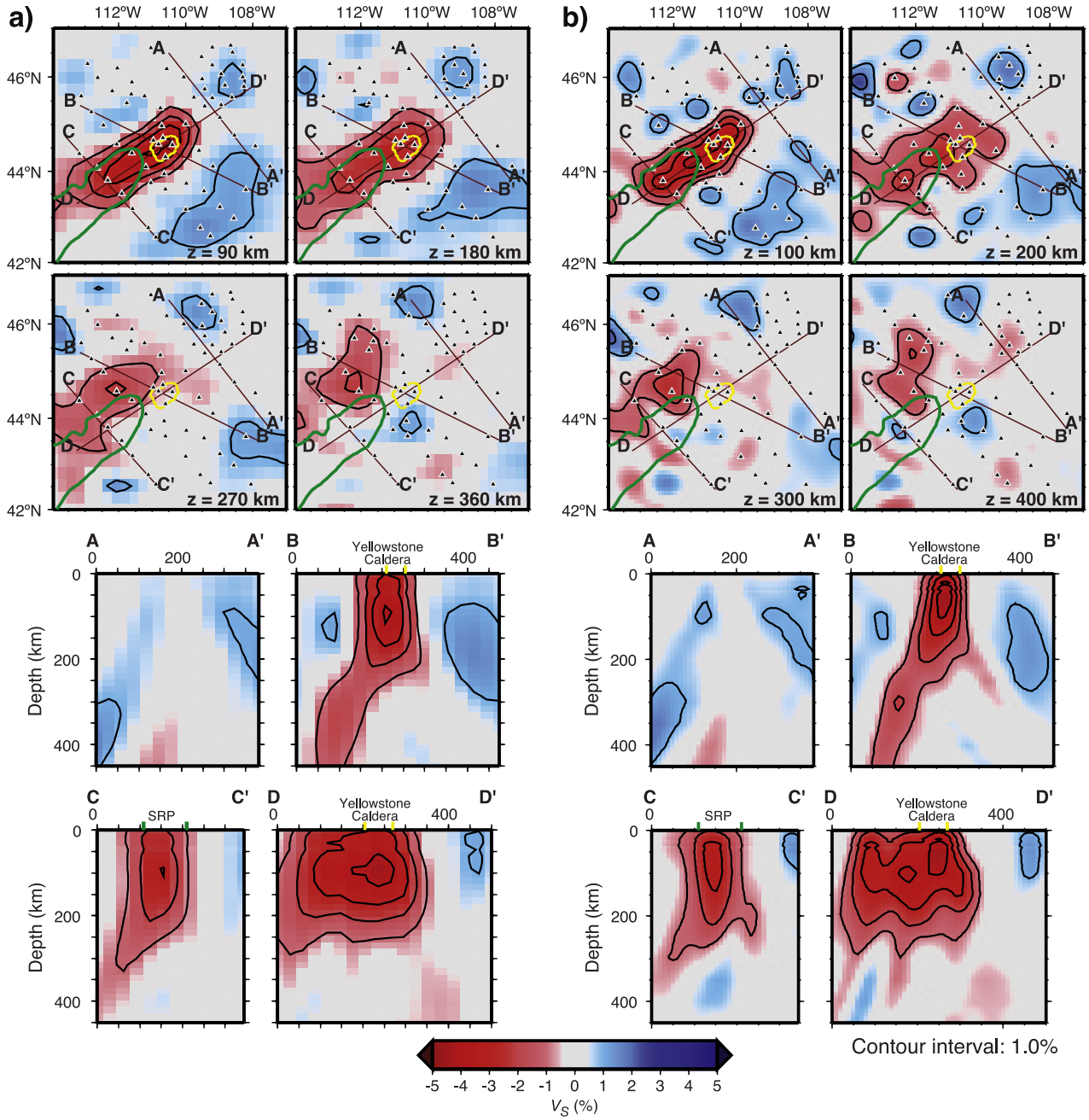

Figure 7. Slices through the V_S models (a) S30VS and (b) OSA50VS in the same locations as in Figure 6.

caldera, have broad, distorted waveforms that do not correlate with arrivals at stations outside the caldera. A combination of scattering by small-scale heterogeneities and attenuation directly beneath the caldera affects the waveforms recorded there.

[40] The geometries of the anomalies in the V_S models (Figure 7) are similar to those in the V_P models, but they have larger amplitudes. The peak negative V_S anomalies of −4.5% (S30VS) and −5.5% (OSA50VS) are directly beneath the caldera. The deeper part of the anomaly to the WNW of the caldera has peak V_S anomalies of −2.2% (S30VS) and −2.5% (OSA50VS). The small volume low-velocity anomaly beneath the caldera is clearly smoothed in the S30VS model, but the amplitude of the deeper anomaly, as well as the geometry, is almost the same in both S30VS and OSA50VS. A high-velocity anomaly with peaks of 1.6% (S30VS) and 1.9% (OSA50VS) is in a similar location to the high-velocity anomaly in the V_P model, although the V_S anomaly is larger in volume and amplitude. Smaller-volume positive anomalies in the lower half of the model are outside the array, where resolution is poor, and are not interpreted.

4.3. Resolution Tests

[41] We conducted several resolution tests to assess the reliability of the tomographic solutions. In some of these tests, an anomaly, or set of anomalies, was used to calculate synthetic traveltimes with the ray set used in the inversion. Normally distributed, random noise with standard deviations of 0.1 s for *P* picks and 0.15 s for *S* picks was added to

the residuals to approximate the uncertainties in the real data and the data were inverted with the same model parameterization used for the real data. These tests measure the ability of the data to resolve synthetic structures of various size, strength, and shape at points in the model space using the data, or subsets of the data. The LSQR algorithm does not explicitly solve the generalized inverse needed to construct the commonly used resolution matrix. Instead, various synthetic models are tested and examined to determine where anomalies are recovered, where leakage strongly affects the model nodes, and where there is no resolution.

4.3.1. Checkerboard Sensitivity Tests

[42] These tests use alternating anomalies of high and low velocity evenly spaced throughout the model in a three-dimensional checkerboard pattern. Synthetic models with anomalies from 1 to 4 nodes wide were tested. The single-node anomalies are poorly resolved even in the middle of the model where the resolution is expected to be the best. The larger checkerboard anomalies are well resolved in the middle of the model, from ~100 to 500 km. Figure 8 shows results using the S30VP and S30VS parameterization for three-node anomalies. There is a large degree of leakage at the edges of the array where all the rays are parallel. Leakage also occurs between adjacent layers as shown in the 150 km depth slice and the cross sections. The size and shape of the anomalies are preserved fairly well in the middle of the model, although the ability to resolve small anomalies does not necessarily demonstrate the ability to resolve large volume anomalies [e.g., *Leveque et al.*, 1993].

4.3.2. Realistic Anomaly Recovery Tests

[43] A second type of test uses a synthetic model that contains low-velocity anomalies with shapes and velocity contrasts similar to those found with the real data. The test explores the smearing that occurs between the shallow and deeper anomalies as well as the percentage of synthetic amplitude recovered. The use of station and event correction terms, damping, and smoothing, as well as smearing and incomplete fitting of the synthetic data in the inversion, results in a reduction of the true amplitudes. In addition, the limitations of high-frequency ray theory, which does not properly account for wavefront healing or the true sensitivity volume for each measurement, can result in up to 70% reduction of true amplitudes in small volume anomalies [e.g., *Allen et al.*, 1999, 2002].

[44] Low-velocity anomalies were placed in the upper mantle beneath the Yellowstone caldera and 100 km NW in the approximate locations of the largest low-velocity anomalies in both models. The shallow anomaly extends from 50 to 200 km, has a 50 km radius, and peak amplitude of -4% V_P (-6% V_S). The deeper anomaly extends from 330 to 390 km, has a 75 km radius and peak amplitude of -2% V_P (-3% V_S). The inverted models show how the anomalies have been smeared together slightly (Figures S1 and S2 in the auxiliary material[1]). In addition, the lower anomaly is tilted toward the center of model and the shallower anomaly. In the V_P model with the same parameterization as in S30VP, the recovered anomalies have peak amplitudes of -1.4% and -1.2% for the shallow and deeper anomalies, respectively. The reduced amplitude recovery in the shallower model is partly a result of the smaller volume of the anomaly.

[45] When the lower anomaly is moved to a position 100 km SE of the caldera, approximately the same percentage of the anomalies are recovered and the smearing is similar. On the basis of these results, 60 to 70% of true anomalies are expected to be recovered in S30VP. Similarly, for the synthetic model with the same parameterization as S30VS, ~60% of the true peak shallow and deep anomalies is recovered. Inversions with the OSA50VP and OSA50VS parameterizations recover 60 to 65% of the shallow anomaly and ~50% of the deeper anomaly. The consistency between the recovery of the deeper anomalies placed NW and SE is evidence that the deep anomaly imaged with the real data is not an artifact of the ray set.

[46] Another test illustrates the degree to which vertical leakage may affect the tomographic model. In this test, a single anomaly is placed directly beneath the Yellowstone caldera between 50 and 200 km depth and the depth of the 410 km discontinuity is depressed to 430 km WNW of Yellowstone to coincide with the discontinuity imaging of *Fee and Dueker* [2004]. The modeled anomaly is elongated vertically, especially in the shallowest part of the model; however there is very little vertical leakage below ~250 km depth and none of the discontinuity topography is modeled as shallower structure. We conclude that unmodeled topography on the 410 km discontinuity does not significantly influence the tomography results. Further testing was done with vertical plume-like anomalies to test the ability to resolve such structures. The anomalies tested were confined to the upper mantle between 50 and 400 km depth. These vertical structures are easy to resolve given the geometry of the teleseismic ray set. Even small, 30 km diameter, -2% V_P anomalies were recovered in these tests indicating such a feature is not likely to be directly beneath Yellowstone.

4.3.3. Squeezing Tests

[47] Squeezing tests are also used to investigate the degree of vertical leakage. These tests are performed using a two-stage inversion approach. In the first step, a portion of the model is overdamped while the rest of the model is regularized with the same damping value used in the whole model inversion. This forces traveltime residuals to be resolved only in a certain portion of the model. When this first stage is complete, the overdamping is removed and the remaining data residuals are inverted in a second stage. If the first stage model sufficiently explains the data, the second stage model, will have no anomalies. When the second stage is complete, the sum of the anomalies from both stages forms the final model.

[48] Squeezing tests using both the *P* and *S* data sets were performed using several depth ranges for the top of the overdamped part of the model from 150 to 300 km. In each test, some of the traveltime residual was mapped into a deep low-velocity anomaly northwest of the caldera in the final model. The results of some of these tests are shown in Figure 9. While squeezing tests are not definitive, the results imply the Yellowstone anomaly extends below 300 km depth.

4.3.4. Inversion With Data Subsets

[49] Additional tests were performed to investigate the sensitivity of certain anomalies to the ray geometry. In these tests, a subset of the data is removed and the inversion is

[1] Auxiliary material is available at ftp://ftp.agu.org/apend/jb/2005jb003867.

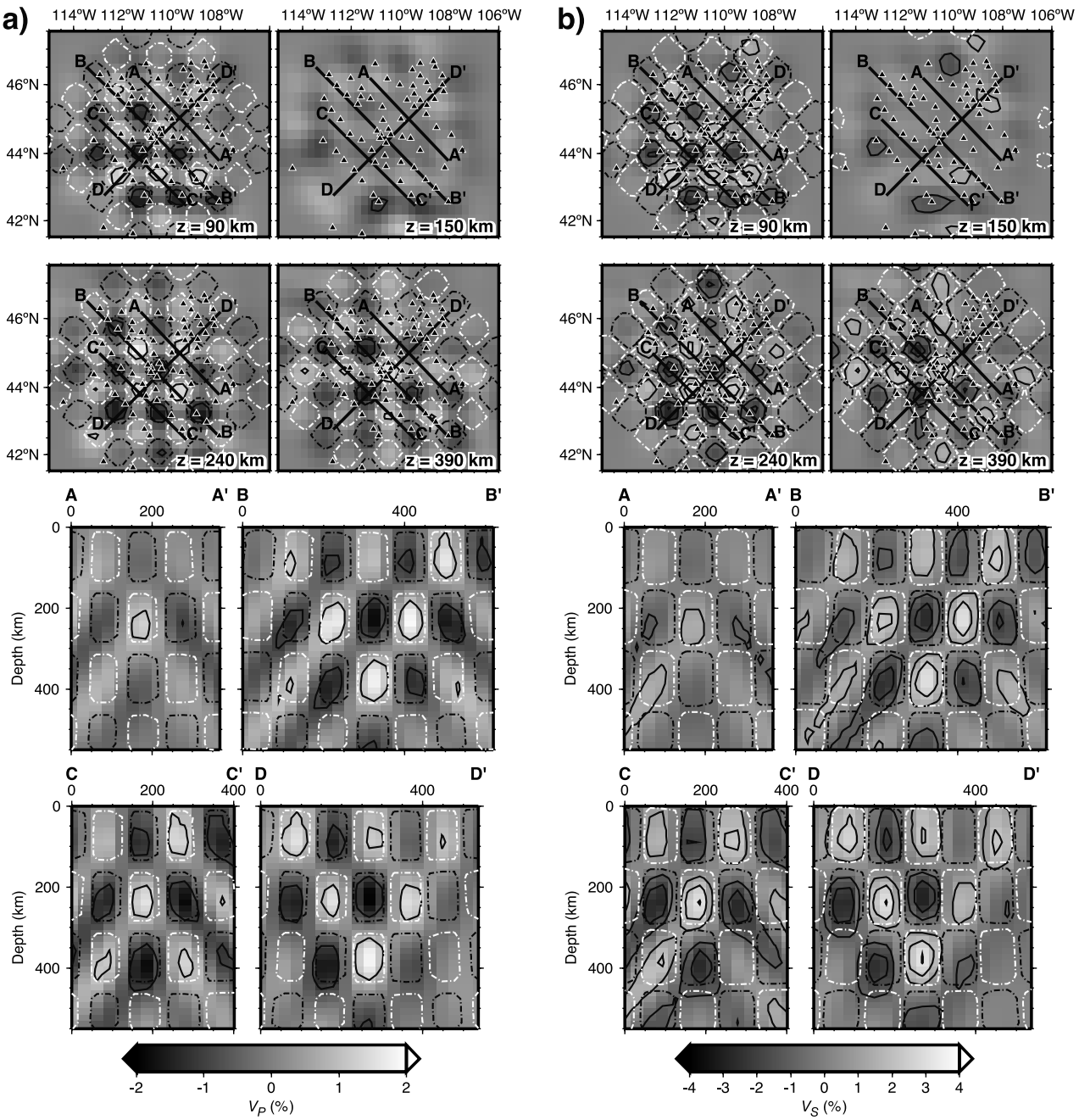

Figure 8. Slices through recovered velocity anomalies in checkerboard sensitivity tests plotted in similar locations as previous plots. Dashed black and white lines outline the locations of input low- and high-velocity anomalies. Input anomalies have peak values of (a) ±2% for V_P and (b) ±4% for V_S. Three horizontal sections are chosen to correspond to the middle of input anomalies (90, 240, and 390 km) and one (150 km) corresponds to a depth with no input anomalies.

performed using a smaller data set. Two of these tests, described below, address the sensitivity of deeper anomalies to the data. Inversion of these data subsets was performed using the same model parameterization used for the full data set.

[50] The upper crustal velocity structure is complicated at Yellowstone, with a −8% low V_P body extending from ~5 km to at least 10 km depth [*Husen et al.*, 2004]. In addition, the tomography results from this study indicate a large-amplitude low-velocity anomaly at the top of both the V_P and V_S mantle models beneath the caldera. These anomalies may smear to greater depths. In the first test, all the traveltime residuals from the stations LKWY, Y100, Y102, Y103, and YFT, which are inside the Yellowstone caldera above the strongest part of the upper mantle anomaly, were removed. Inversion of the reduced data set yields a model with anomalies of essentially the same geometry as inversion with full data set (Figures S3a and S4a in the auxiliary material). While there is a significant reduction in the amplitude of the shallow, <100 km, low-velocity mantle anomaly directly beneath

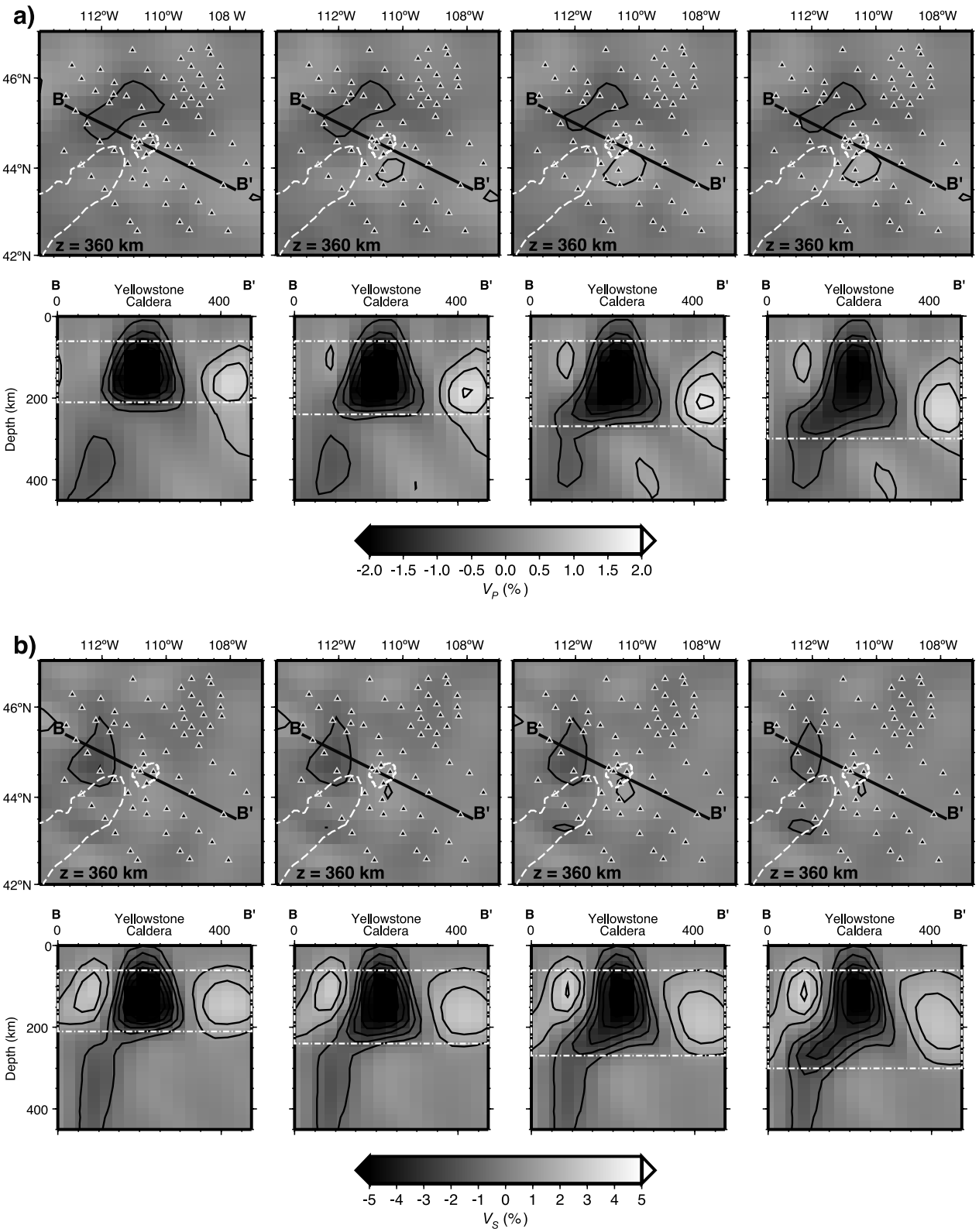

Figure 9. Results of several "squeezing" tests for (a) S30VP and (b) S30VS model parameterizations. Horizontal slices at 360 km and vertical slices in the same place as the B-B′ section in Figures 6 and 7 are shown. The white dash-dotted boxes in the vertical cross sections show the depth range where the model was permitted to change in the first step of the inversion. See text for definition of the "squeezing" test.

the caldera, this is expected since there are few rays remaining in this part of the model. The deeper anomaly, however, is not significantly affected by the smaller data set.

[51] In the other test, data from earthquakes to the NW (i.e., earthquakes from azimuths between 300° and 360° from the caldera) were removed from the data set. While the anomalies have different shapes than those imaged with the full data set, they are generally in the same positions (Figures S3b and S4b in the auxiliary material). The deep, 300–400 km depth anomaly is clear although it is nearly separated from the shallow anomaly. This is an important difference from the inversion of the entire data set that shows a continuous low-velocity feature from near the surface to ~400 km depth. Taken together, these tests provide confidence that the deeper, 250–400 km portion of the low-velocity anomaly imaged with the full inversion is not a result of smearing of shallow anomalies.

5. Discussion of the Yellowstone Hot Spot V_P and V_S models

[52] The interpretation of seismic tomography requires knowledge of the effects of temperature, anisotropy, and composition including the presence of water or partial melt. Forward modeling of seismic velocity for a large number of upper mantle thermal and compositional parameters shows that variations in temperature have the largest effect [e.g., *Goes et al.*, 2000; *Goes and van der Lee*, 2002]. Exceptions may include regions where plumes or small-scale convection may produce volumes of different composition through melting, hydration and dehydration. For example, *Schutt and Humphreys* [2004] interpret velocity variations across the ESRP, ~100 km SW of the YISA array, primarily in terms compositional heterogeneity. The low-velocity anomaly beneath the ESRP is attributed to up to 1% partial melt. The high-velocity bodies on the flanks of the ESRP are interpreted to be only 80 K cooler, but 5% depleted in basaltic component.

[53] Seismic anisotropy has largely been ignored in velocity tomography studies although it can effect on the ability to resolve velocities [e.g., *Levin et al.*, 1996]. The anisotropic contribution to the traveltime delay depends on the amplitude of the anisotropy, direction of propagation and polarization, and thickness of the anisotropic medium. *Schutt and Humphreys* [2004] used a correction for the anisotropy beneath the ESRP [*Schutt and Humphreys*, 2001] to remove the effect of anisotropy in their study. The simple anisotropic structure of the upper mantle beneath the ESRP, with roughly parallel directions of fast anisotropy everywhere, allowed corrections to be made with reasonable assumptions about the mean direction of fast anisotropy and thickness of the anisotropic layer. However, they found little difference between their tomography results, which include correction for anisotropy and those that did not. *Keyser et al.* [2002] found no first-order effect of S wave anisotropy in their shear wave tomography of the Eifel hot spot, despite a complex pattern of shear wave splitting fast directions [*Walker*, 2004]. Since the distribution of fast S wave polarization directions at Yellowstone is comparable or simpler than at Eifel, we do not expect that accounting for anisotropy will have a significant effect on the tomography results.

[54] Some additional items should be considered when interpreting the seismic anomalies in terms of thermal and compositional variations. First, recognizing that not all of the true anomaly amplitude is recovered with the inversion, the modeled anomalies should be considered as minimums. The relative seismic anomalies contribute the primary source of uncertainty in the interpretation. Second, the excess temperature estimates are relative to a mantle that is warmer than average. *Goes and van der Lee* [2002] estimate a temperature anomaly of 200 K to at least 250 km depth beneath the active basin-range province. Third, no significant chemical anomalies are interpreted to be in this region [*Godey et al.*, 2004]; however, anomalies on the scale of a narrow upwelling may not appear in the surface wave tomography used by *Godey et al.* [2004] to estimate temperatures and chemical variations. Finally, as revealed by the synthetic testing, some vertical leakage of seismic anomalies occurs in the inversion. In particular, the depths of the velocity anomalies may be overestimated.

[55] To aid in interpreting the anomalies, three-dimensional, perspective view plots of the low velocities in models OSA50VP and OSA50VS are shown in Figures 10 and 11. Surfaces of equal velocity perturbation from the starting one-dimensional models are shown from −1.5% to −0.75% for V_P and −4.5% to −1.5% for V_S. These plots clearly show that the strongest velocity anomaly is in the uppermost mantle beneath Yellowstone and the ESRP. At smaller velocity contrasts, the anomaly stretches from the crust to the top of the transition zone ~100 km WNW of Yellowstone. The main seismic velocity anomalies derived in this study are summarized in Table 1.

5.1. Effect of Temperature on Seismic Velocities in a Dry Mantle

[56] We begin by interpreting the velocity anomalies in a dry, chemically homogeneous mantle model and show that very high attenuation is required to explain the observations. The significance of the temperature effect on anelasticity is well documented [e.g., *Karato*, 1993; *Goes et al.*, 2000; *Cammarano et al.*, 2003]. Preliminary work by *Adams and Humphreys* [2003] found high S wave attenuation in the upper mantle beneath Yellowstone and the ESRP. Beneath Yellowstone caldera, the high attenuation region extends to the base of the crust.

[57] We compute partial derivatives, $\partial \ln V_P/\partial T$ and $\partial \ln V_S/\partial T$ following the work of *Goes et al.* [2000] for an average continental garnet lherzolite [*Jordan*, 1979] along a 1550K adiabat at 100 and 300 km depth. Values for the mantle mineral parameters are taken from laboratory measurements (see *Schutt and Lesher* [2006] for a summary). For a Q model such as that used by *Cammarano et al.* [2003], the shallow low V_P anomaly corresponds to a temperature anomaly of ~150 K, but the low V_S anomaly requires a much higher temperature anomaly of at least 300 K. At greater depths, the V_P and V_S models predict a similar temperature anomaly ($\Delta T \approx 170$ K), but the V_S perturbations still favor a larger temperature anomaly than the V_P. We take advantage of the nonlinear dependence of Q on temperature to find a single temperature anomaly that can explain the P and S anomalies given a relatively constant ratio Q_P/Q_S.

[58] Variations in Q can range over 2 orders of magnitude in some regional seismic studies [e.g., *Sato*, 1992; *Umino*

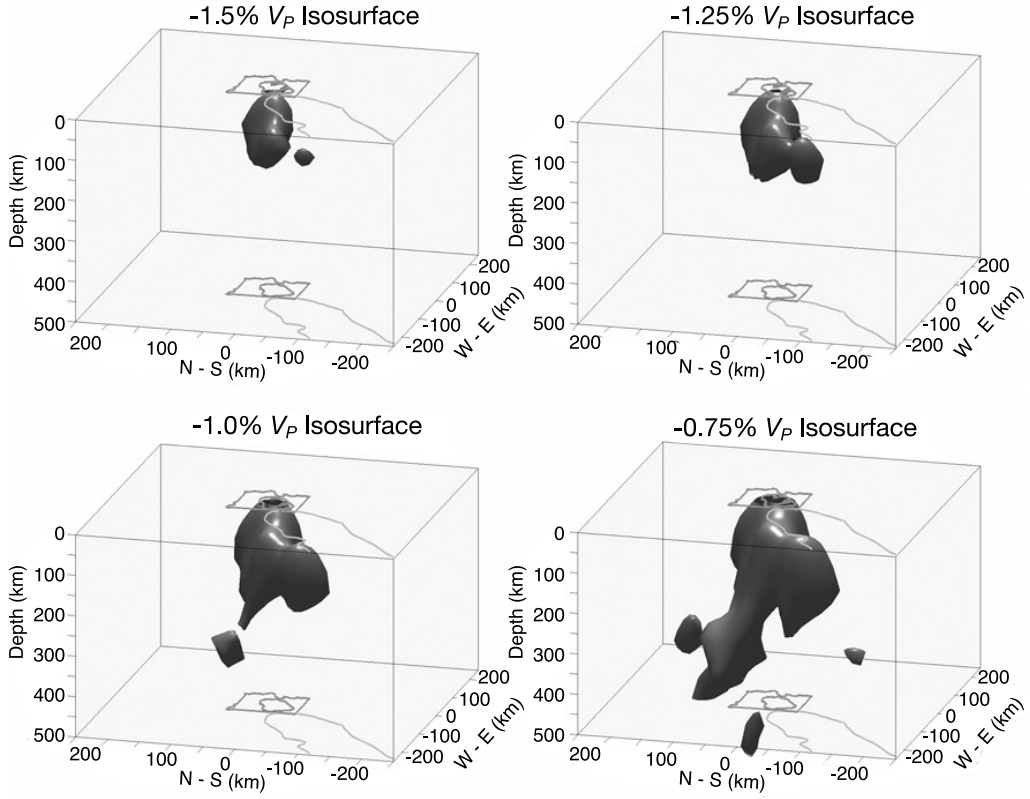

Figure 10. Perspective view of *P* wave isovelocity perturbations from the starting model for model OSA50VP looking NE. The outline of the Yellowstone caldera, ESRP, and Yellowstone National Park are plotted on the top and the bottom of the 500 km by 500 km by 500 km cubes for reference.

and Hasegawa, 1984]. Similarly, experimental models predict large variations in Q with depth and temperature [e.g., *Berckhemer et al.*, 1982]. An infinite number of combinations of Q and T fit our V_P or V_S data. For example, a 2.3% reduction in V_P is predicted from very high attenuation ($Q_P = 25$) and small temperature changes ($\Delta T = 65$ K) or from negligible attenuation (Q_P = 1000) and large temperature increase ($\Delta T = 375$ K). However, with knowledge of both V_P and V_S perturbations, the range of likely Q and T values can be narrowed considerably.

[59] We examined a range of values for the shear modulus quality factor, Q_μ, from 5 to ∞, and computed the corresponding $\partial \ln V_P/\partial T$ and $\partial \ln V_S/\partial T$ for a range of temperatures. We account for the frequency dependence of Q_μ with ω^a following the attenuation model of *Minster and Anderson* [1981] where $a = 0.15$ [*Sobolev et al.*, 1996]. The ratio Q_P/Q_S is computed from:

$$Q_P/Q_S = (3/4)(V_P/V_S)^2(\omega_P/\omega_S)^a, \quad (5)$$

which has been modified from the work of *Anderson and Given* [1982], where we assume anelasticity related to the bulk modulus is negligible ($Q_K^{-1} = 0$) and the ratio of *P* to *S* wave frequencies, $\omega_P/\omega_S = 2$. Assuming uncertainties in the model's resolution might incorrectly estimate V_P/V_S (see below), we compute Q_P/Q_S for a range of reasonable values (V_P/V_S = 1.75 to 1.90). The range of Q_P/Q_S, 2.5 to 3.0, is relatively small, permitting the observed V_P and V_S perturbations to be explained by the same temperature anomaly and $Q_P \approx 2.75 Q_S$. If we assume the anelasticity effect associated with the bulk modulus is finite, the ratio Q_P/Q_S is reduced. For Q_K = 1000, Q_P/Q_S varies from about 2.7 for Q_S = 10 to 2.4 for Q_S = 100.

[60] The shallowest upper mantle anomaly beneath the Yellowstone caldera has a peak amplitude of −2.3% in the OSA50VP model and −5.5% in the OSA50VS model (Table 1). For a low Q_S = 10, a 5.5% reduction in V_S can be obtained with a temperature anomaly of only ∼70 K. This corresponds to $Q_P \approx 27$ which yields a V_P reduction of 2.3%. For larger values of Q_S and ΔT that fit the V_S reduction, the corresponding Q_P is not large enough to explain the model V_P reduction. For a temperature anomaly of 380 K, $\Delta V_S = -5.5\%$ for $Q_S = 100$. However, there has to be very little *P* wave attenuation ($Q_P > 1000$), or the V_P reduction is too large at $\Delta T = 380$ K.

[61] The S30VP and S30VS results also suggest a very high degree of anelasticity in the shallow mantle and relatively small temperature anomaly. If the observed seismic anomalies in the upper 200 km are purely due to thermal (and attenuation) effects, both the smoothed and offset-and-averaged model parameterizations predict very low Q and relatively small temperature anomalies 40 to 70 K. However, such high attenuation may be unlikely without higher temperatures, melt or water. The smaller amplitude, deeper anomalies can be explained by a wider range of Q (Q_S: 10 to 50; Q_P: 27 to 150) and temperature anomalies of

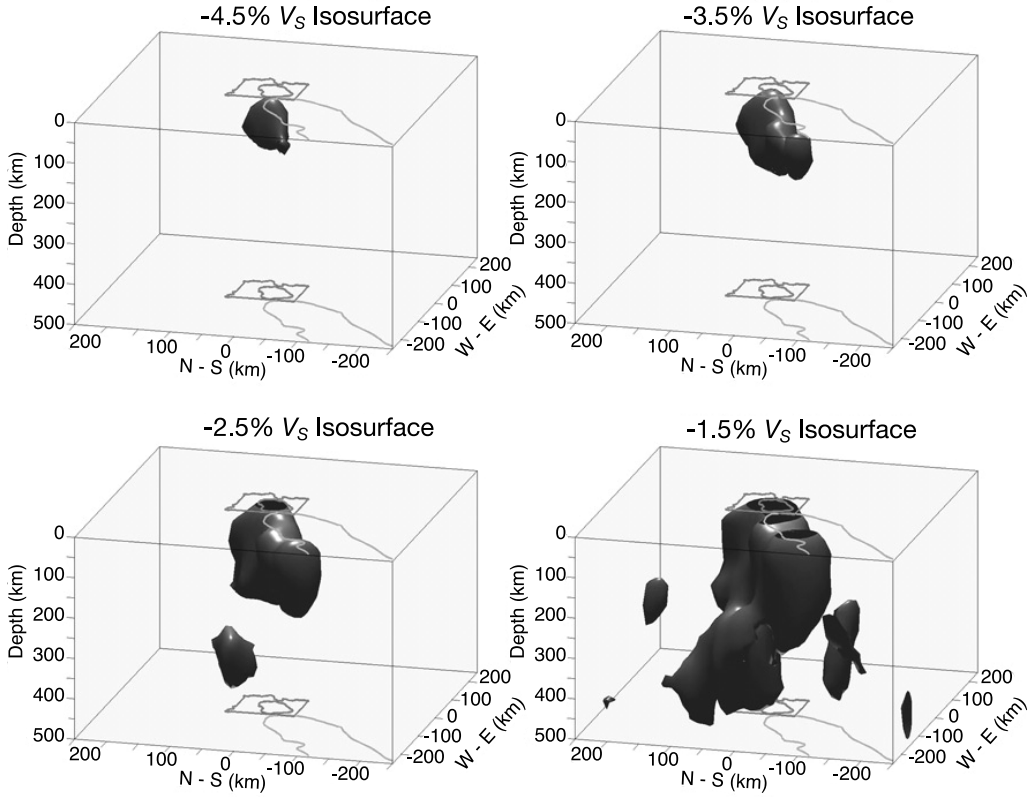

Figure 11. Perspective view of *S* wave isovelocity perturbations from the starting model for model OSA50VS looking NE as in Figure 10.

30 to 120 K. The upper ends of these ranges may be realistic for the lower part of the upper mantle indicating the deeper (300 km) seismic velocity anomalies may not require compositional anomalies to explain them.

5.2. Effects of Compositional Variations on Seismic Velocities

[62] The presence of melt can reduce seismic velocities, but the strong dependence on melt geometry makes predicting melt percent from seismic velocity perturbations difficult [*Goes et al.*, 2000]. In addition to reducing seismic velocities, the orientation of these lenses can cause strong anisotropy [*Kendall*, 1994], further complicating the interpretation. *Faul et al.* [1994] found the addition of 1% partial melt distributed in ellipsoidal lenses can lower V_P by 1.8% and V_S by 3.3%. *Hammond and Humphreys* [2000] calculated a larger reduction of 3.6% and 7.9% in V_P and V_S, respectively, per 1% partial melt distributed in geometries inferred from laboratory experiments. If this is correct, velocity reductions due to less than 1% partial melt could explain all of the observed shallow anomaly in the models. Importantly, because a small amount of partial melt primarily affects the shear modulus, the percent reduction in V_S is much greater than the reduction in V_P. This is consistent with our models where the percent reduction in V_S is more than a factor of two greater than the percent reduction in V_P.

[63] If water is present, the melting temperature may be 700 K lower than the melting temperature of a dry mantle at 200 km depth [*Thompson*, 1992]. Water may be present in hydrated minerals or possibly as free water [*Kawamoto and Holloway*, 1997] and has been shown to reduce seismic velocities through enhanced anelasticity [*Karato and Jung*, 1998]. As with partial melt, the reduction of seismic velocities due to the presence of water is greater for V_S than V_P. Water may be transported up through the upper mantle within an upwelling of hotter, buoyant material. Transition zone minerals can dissolve 2 to 3% water [*Kohlstedt et al.*, 1996]. Therefore the upwelling material may have a higher concentration of water than the surrounding mantle. However, in addition to lowering the seismic velocities, the water could lower the solidus enough to produce melt to depths of 250 km [*Kawamoto and Holloway*, 1997]. Since water is preferentially removed by melting, anomalies in the upper 250 km of the mantle are more likely due to partial melt than water.

[64] Mineralogical heterogeneity can also produce seismic velocity variations. Melt depletion (i.e., preferential removal of iron-rich olivines with respect to magnesium-rich olivines) was predicted to increase seismic velocities [*Jordan*, 1979] and has been used to explain high seismic velocities in some areas. However, recent work on the effect of melt depletion on velocities based on new laboratory

Table 1. Peak Seismic Velocity Anomalies

Depth of Anomaly, km	S30VP	OSA50VP	S30VS	OSA50VS
50–200	−2.0%	−2.3%	−4.5%	−5.5%
250–400	−1.0%	−1.0%	−2.2%	−2.5%
100–250	1.1%	1.5%	1.6%	1.9%

observations, suggests earlier estimates may have been too high [*Schutt and Lesher*, 2006]. Finally, *Faul and Jackson* [2005] demonstrate the correlation of grain size with seismic velocity and Q and suggest grain size increases with depth in the upper mantle. While these results suggest intriguing new models to test, we assume constant grain size in our interpretations.

5.3. Interpretation of Velocity Perturbations at Yellowstone

[65] The modeled low-velocity anomalies are likely due to a combination of temperature and compositional anomalies. The shallower, 50 to 200 km part of the low-velocity anomaly can be interpreted to consist of low Q material ($Q_S <$ 100), with a temperature anomaly of <100 K, and less than 1% partial melt. The Q estimates are comparable to those associated with the Eifel plume [*Keyser et al.*, 2002], but larger than values inferred for the shallow upper mantle beneath ESRP ($Q_S \leq 20$) [*Schutt and Humphreys*, 2004]. The deeper, 250 to 400 km, part of the anomaly can be explained by low Q material with a temperature anomaly of up to 120 K. If water is present between 250 and 400 km, such low values for Q are not necessary to explain the seismic anomalies. A combination of slightly higher temperatures, concentrations of water, and/or melt is likely in both depth ranges since the true amplitude of the anomalies is underestimated by the inversion.

[66] At 200 to 250 km depth beneath Yellowstone, the anomaly is slightly weaker. The decrease in the amplitude of the anomalies may be due to melting. *Karato and Jung* [1998] suggest that partial melting will increase seismic velocities through the removal of water. A similar decrease in the amplitude of the Eifel plume was found near 200 km depth [*Keyser et al.*, 2002]. The following scenario, modified from the work of *Keyser et al.* [2002] may explain the zone of weaker anomaly in the Yellowstone plume: upwelling material moving from the transition zone could have a higher concentration of water than the surrounding upper mantle; with enough water in the upwelling, melting could initiate at 200 to 250 km depth [*Kawamoto and Holloway*, 1997]; as the melt rises buoyantly, the seismic velocity of the material left behind will increase. The melt may pond at shallower depths reducing the velocities and increasing the attenuation there.

[67] Unlike the two-dimensional tomography results of *Schutt and Humphreys* [2004] across the ESRP 200 km southwest of Yellowstone, our Yellowstone tomography results do not reveal symmetric high-velocity volumes on both sides of the low-velocity anomaly. The V_P and V_S models do show strong high velocities ~200 km SE of the Yellowstone caldera from 100 to 300 km depth. The depth extent of this anomaly is overestimated because it is at the edge of the seismograph array. The anomaly is outside the seismic and topographic parabola and is probably too far from the caldera to be interpreted as buoyant melt residuum. It may represent a downwelling of colder mantle, possibly to accommodate new material that has moved up through the plume [*Yuan and Dueker*, 2005].

5.4. Comparison With Other Hot Spots

[68] Regional tomographic imaging studies conducted at other hot spots have revealed structures that are unlike the low-velocity anomaly under Yellowstone. For example, the Eifel hot spot upper mantle has a narrow, vertical, low-velocity anomaly from the crust to the transition zone. The Iceland upper mantle also has a vertical low-velocity anomaly through the upper mantle, but it is slightly elongated N–S parallel to the mid-Atlantic spreading center [*Foulger et al.*, 2001; *Allen et al.*, 2002]. In addition, *Allen et al.* [2002] combine surface wave and body wave data and find a relatively high-velocity anomaly in the uppermost mantle interpreted as the result of melt extraction. Unfortunately, the upper mantle beneath Hawaii, which is perhaps the best known midplate hot spot, has not been imaged with the resolution of Iceland, Eifel or Yellowstone due to the difficulty of instrumenting the ocean floor [e.g., *Wolfe et al.*, 2002a, 2000b].

[69] Plots of −1% V_P isovelocity surfaces of the upper mantle beneath Iceland (from ICEMAN-HP from *Allen et al.* [2002]), Eifel [*Ritter et al.*, 2001; M. Jordan, personal communication, 2003] and Yellowstone (OSA50VP) are shown in Figure 12 for comparison with the Yellowstone results. We do not interpret small-scale differences between the three models because different model parameterizations were used to calculate them. However, large-scale (tens to hundreds of kilometers) variations are also evident and are not likely to be the result of parameterization.

[70] The differences between the upper mantle structures beneath Iceland, Eifel, and Yellowstone are not surprising given their different tectonic settings. Iceland is an ocean island atop a spreading center. Regardless of whether the upwelling begins at the core-mantle boundary or the transition zone, the ascent of material must be influenced by the rift and associated upwelling. Eifel is a continental hot spot like Yellowstone, but sits above the relatively stable Eurasia plate while Yellowstone is at the edge of the extending basin-range province and Archean craton. These differences make direct comparisons of the amount of erupted material relative to the upper mantle anomalies difficult. The differences in the upper mantle structures, however, suggest that the different upper mantle settings may have an important influence on the surface expression of the hot spots.

5.5. A Yellowstone Plume?

[71] Debate about the existence of plumes in the mantle has persisted since they were first proposed [e.g., *DePaolo and Manga*, 2003; *Foulger and Natland*, 2003]. Several alternatives to the plume hypothesis have been suggested to explain Yellowstone hot spot volcanism, that include, but are not limited to: an eastward propagating rift [*Hamilton and Myers*, 1966; *Smith*, 1977]; volcanism along a preexisting crustal weakness [*Eaton et al.*, 1975]; flow around a melt residuum body [*Humphreys et al.*, 2000]; a thermal feedback between extension and shear melting to generate a self-sustaining melting anomaly that is guided by an ancient structural zone [*Christiansen et al.*, 2002]; and propagating asthenospheric ''Richter rolls'' [*Hernlund and Tackley*, 2003; *Tackley and Stevenson*, 1993].

[72] Analysis of teleseismic shear wave splitting provides little evidence for plume-related radial flow below the Yellowstone lithosphere [*Waite et al.*, 2005]. Upon reaching the base of the lithosphere, buoyant plume material is expected to flow outward since not all of the material can penetrate the more viscous plate. The combination of radial

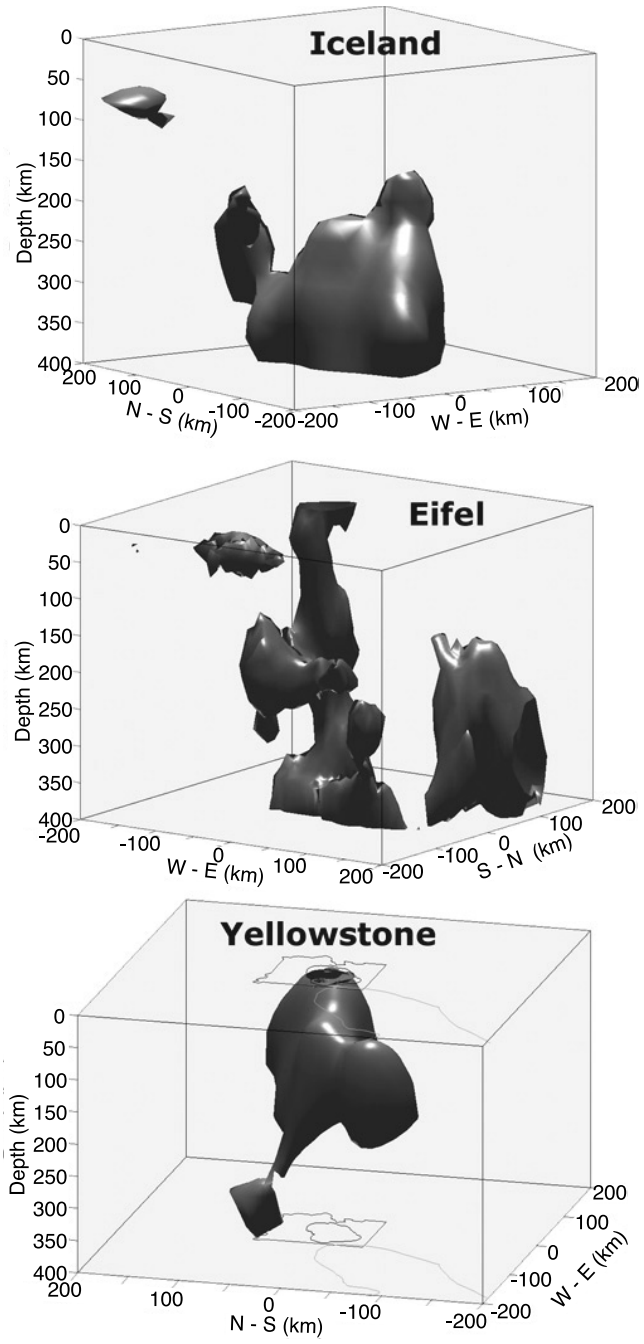

Figure 12. Perspective plots of −1% V_P perturbation surfaces of models of the upper mantle beneath of other hot spots: Iceland, Eifel, and Yellowstone. The scale of each 400 km by 400 km by 400 km cube is the same. While different model parameterizations used to construct the models may affect the amplitude of the anomalies, and therefore the shape of the isovelocity surface, the plots clearly show differences between the upper mantle structures beneath these three hot spots.

spreading with plate motion will produce a parabolic flow pattern in the asthenosphere and a similar pattern of anisotropy. However, split *S* waves show little to no sign of contribution from radially spreading plume material, indicating the contribution of gravitationally spreading plume material beneath Yellowstone is undetectably small with respect to the plate motion velocity.

[73] The phase changes that are primarily responsible for the 410 and 660 km discontinuities have opposite Clapeyron (dP/dT) slopes so thermal anomalies that cross the transition zone should have opposite effects on the discontinuity topography [*Bina and Helffrich*, 1994]. The 15 km increase in the depth of the 410 km discontinuity observed ∼100 km WNW of Yellowstone implies a positive thermal anomaly of ∼200 K at that depth [*Fee and Dueker*, 2004]. However, the 660 km discontinuity topography is not correlated with the deep 410 km discontinuity in that area. The thermal anomaly may not continue downward through the transition zone to 660 km depth, or the 660 km discontinuity is more complex than *Bina and Helffrich*'s [1994] estimate and involves multiple phase transitions [*Vacher et al.*, 1998; *Simmons and Gurrola*, 2000]. The topography on the 410 and 660 km discontinuities elsewhere in the western U.S. varies by 20–30 km and is also uncorrelated in general [*Gilbert et al.*, 2003].

[74] While the shear wave anisotropy pattern does not favor a buoyant plume beneath Yellowstone the discontinuity imaging is consistent with our tomography results that image a continuous low-velocity feature through the upper mantle. This anomaly is a plume by our general definition. In order to reconcile the seismic tomography with the observations cited as against a mantle plume we employ a plume-fed upper mantle small-scale convection model. This model follows the work of *Saltzer and Humphreys* [1997], *Humphreys et al.* [2000], and *Hernlund and Tackley* [2003], which suggests that a plume may fuel small-scale upper mantle convection.

[75] Numerical modeling demonstrates that longitudinal, small-scale, convection cells can develop spontaneously in the upper mantle where there is available partial melt from, for example, an upwelling plume [*Hernlund and Tackley*, 2003; *Tackley and Stevenson*, 1993]. Density differences between buoyant mantle containing partial melt and denser mantle with no melt, initiates convection. Decompression of ascending mantle results in more melting. This causes a larger density contrast and the result is a positive feedback. Melt residuum, which is also lower in density than normal mantle, accumulates on the sides of the convection cells. These convection cells could be aligned by the moving plate to mimic linear hot spot trends. The accumulation of melt residuum at the sides of the cells would eventually halt convection, but the addition of hot and/or wet material from a plume could sustain the melting anomaly. In addition, basin-range extension above this type of system would thin the upper mantle and encourage upwelling and melt production [*Saltzer and Humphreys*, 1997].

[76] It is plausible that the tilt of the plume may be due to upper mantle convection. For example, if a Yellowstone plume is advected in the eastward mantle flow [*Bunge and Grand*, 2000; *Steinberger*, 2000], it should be plunging to the west, similar to the WNW plunge of the low-velocity feature imaged in the tomography models. When combined with plate motion, *Steinberger*'s [2000; personal communication, 2003] models predict a Yellowstone hot spot track north of the ESRP. The location of the hot spot track, however, is also likely to be influenced by the linear lithospheric anomaly it seems to follow [e.g., *Eaton et al.*,

1975; *Smith*, 1977]. *Dueker et al.* [2001] imply the NE-SW Proterozoic Madison mylonite zone, interpreted as a deep, ancient shear zone [*Erslev and Sutter*, 1990], just NE of Yellowstone may provide a favorable guide for small-scale convection. Magnetic and gravity anomalies associated with this shear zone suggest it is a continuous, deep lithospheric structure [*Lemieux et al.*, 2000]. It is reasonable that as the hot spot encountered thicker lithosphere on its NE progression, the plume found the path of least resistance to the surface.

[77] We favor a combination plume-fed upper mantle convection model to reconcile the geologic as well as the seismic observations. While the plume is capable of transporting material up from the transition zone, the volume may not be large enough to sustain the energetic volcanism at Yellowstone alone. A lineation of weak lithospheric structure may be important in guiding the hot spot by allowing melt to penetrate into the crust more easily. The persistence of magmatism along the ESRP may be attributed to continued convection millions of years after the plate has passed the plume. The complex upper mantle flow field expected for longitudinal rolls can explain why evidence for a parabolic flow pattern is not seen in the shear wave splitting data.

6. Concluding Remarks

[78] Tomographic inversions of traveltime delays across the Yellowstone region provide an image of a low V_P and V_S anomaly at the bottom of the upper mantle and the unusual finding of a low-velocity body tilted $\sim 30^\circ$ from vertical and extending laterally more than 100 km northwest of Yellowstone. We interpret this structure as an upper mantle plume. In addition, the modeling reveals a low V_P and V_S anomaly directly beneath the Yellowstone caldera extending to 200–250 km depth. This shallow feature is continuous, with a smaller amplitude, to the SW beneath the ESRP to the edge of the model.

[79] Yellowstone has a plume source, although it is not necessarily deep plume that originates at the core-mantle boundary. In fact, there is no evidence to show that the low-velocity anomaly continues through the transition zone to the lower mantle. As such, it may be strictly an upper mantle feature. The coincidence of Yellowstone with the boundary of the Archean craton and basin-range as well as structural trends that parallel the hot spot track indicate lithosphere features may be important in guiding the hot spot.

[80] Upper mantle convection models are not contradicted by a plume model. Rather, convection, lithosphere extension, and upwelling from below likely work together at Yellowstone. Small-scale convection helps explain the strong low-velocity anomaly beneath Yellowstone and the Snake River Plain to $\sim$200 km depth. The high topography on both sides of the ESRP may be supported by melt residuum that has been pushed away from the upwelling zone under the ESRP. The possible eastward migration of the basin-range extensional regime is a partly a consequence of the active system moving in the direction opposite plate motion. Without all three mechanisms, Yellowstone volcanism may not have persisted for $\sim$16 million years.

[81] **Acknowledgments.** This project was part of the collaborative NSF Continental Dynamics project: Geodynamics of the Yellowstone Hot spot from Seismic and GPS Imaging. Collaborators included Gene Humphreys, Jason Crosswhite of the University of Oregon; Paul Tackley and John Hernlund of UCLA; and Ken Dueker and Derek Schutt of the University of Wyoming. We are especially appreciative of the dedicated field support teams that advised, installed, and maintained the instruments including J. Crosswhite, Dave Drobeck, K. Dueker, D. Schutt, and Brian Zurek. The seismographs were provided by the PASSCAL facility of IRIS through the PASSCAL Instrument Center. Data collected from this experiment are available at the IRIS Data Management Center. The IRIS Consortium is supported by the NSF. Additional data were acquired from the USGS National Seismograph Network and University of Utah Seismograph Network. John Evans provided the teleseismic picks from earlier USGS field experiments in Yellowstone. We appreciate discussions with Uli Achauer, Thorsten Becker, Robert Christiansen, Gillian Foulger, Michael Jordan, Rafaella Montelli, Richard O'Connell, Christine Puskas, and Bernard Steinberger. The NSF Continental Dynamics Program grants EAR-CD-9725431 and 0314237 provided support. The University of Utah supported computational aspects of the project.

References

Adams, D. C., and E. D. Humphreys (2003), *Q* tomography in the Yellowstone region: Effects of mantle water content, *Eos Trans. AGU*, *84*, Fall Meet. Suppl., Abstract S21A-05.

Allen, R. M., et al. (1999), The thin hot plume beneath Iceland, *Geophys. J. Int.*, *137*, 51–63.

Allen, R. M., et al. (2002), Imaging the mantle beneath Iceland using integrated seismological techniques, *J. Geophys. Res.*, *107*(B12), 2325, doi:10.1029/2001JB000595.

Anders, M. H., and N. H. Sleep (1992), Magmatism and extension: The thermal and mechanical effects of the Yellowstone hotspot, *J. Geophys. Res.*, *97*, 15,379–15,393.

Anders, M. H., J. W. Geissman, L. A. Piety, and J. T. Sullivan (1989), Parabolic distribution of circumeastern Snake River Plain seismicity and latest Quaternary faulting, migratory pattern and association with the Yellowstone hotspot, *J. Geophys. Res.*, *94*, 1589–1621.

Anderson, D. L. (2000), The statistics and distribution of helium in the mantle, *Int. Geol. Rev.*, *42*, 289–311.

Anderson, D. L. (2001), Top down tectonics?, *Science*, *293*, 2016–2018.

Anderson, D. L., and J. W. Given (1982), Absorption band *Q* model for the Earth, *J. Geophys. Res.*, *87*, 3893–3904.

Bassin, C., G. Laske, and G. Masters (2000), The current limits of resolution for surface wave tomography in North America, *Eos Trans. AGU*, *81*(48), Fall Meet. Suppl., Abstract S12A-03.

Benz, H. M., and R. B. Smith (1984), Simultaneous inversion for lateral velocity variations and hypocenters in the Yellowstone region using earthquake and refraction data, *J. Geophys. Res.*, *89*, 1208–1220.

Berckhemer, H., W. Kampfman, E. Aulbach, and H. Schmeling (1982), Shear modulus and Q of forsterite and dunite near partial melting from forced oscillation experiments, *Phys. Earth Planet. Inter.*, *29*, 30–41.

Bijwaard, H., W. Spakman, and E. R. Engdahl (1998), Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, *103*, 30,055–30,078.

Bina, C. R., and G. Helffrich (1994), Phase transition Clapeyron slopes and transition zone seismic discontinuity topography, *J. Geophys. Res.*, *99*, 15,853–15,860.

Blackwell, D. D. (1969), Heat-flow determinations in the northwestern United States, *J. Geophys. Res.*, *74*, 992–1007.

Bunge, H.-P., and S. P. Grand (2000), Mesozoic plate-motion history below the northeast Pacific Ocean from seismic images of the subducted Farallon slab, *Nature*, *405*, 337–340.

Cammarano, F., S. Goes, P. Vacher, and D. Giardini (2003), Inferring upper-mantle temperatures from seismic velocities, *Phys. Earth Planet. Inter.*, *138*, 197–222.

Camp, V. E., and M. E. Ross (2004), Mantle dynamics and genesis of mafic magmatism in the intermontane Pacific Northwest, *J. Geophys. Res.*, *109*, B08204, doi:10.1029/2003JB002838.

Christiansen, R. L. (2001), The Quaternary and Pliocene Yellowstone Plateau volcanic field of Wyoming, Idaho, and Montana, *U.S. Geol. Surv. Prof. Pap., 729-G*.

Christiansen, R. L., and R. S. Yeats (1992), Post-Laramide geology of the U.S. Cordilleran region, in *The Geology of North America*, vol. G3, *The Cordilleran Orogen: Conterminous U.S.*, edited by B. C. Burchfiel, P. W. Lipman, and M. L. Zoback, pp. 261–406, Geol. Soc. of Am., Boulder, Colo.

Christiansen, R. L., G. R. Foulger, and J. R. Evans (2002), Upper-mantle origin of the Yellowstone hotspot, *Geol. Soc. Am. Bull.*, *114*, 1245–1256.

Courtillot, V., A. Davaille, J. Besse, and J. Stock (2003), Three distinct types of hotspots in the Earth's mantle, *Earth Planet. Sci. Lett.*, *205*, 295–308.

Craig, H., J. E. Lupton, J. A. Welhan, and R. Poreda (1978), Helium isotope ratios in Yellowstone and Lassen Park volcanic gases, *Geophys. Res. Lett.*, *5*, 897–900.

DePaolo, D. J., and M. Manga (2003), Deep origin of hotspots: The mantle plume model, *Science*, *300*, 920–921.

Dueker, K. G., and E. D. Humphreys (1990), Upper-mantle velocity structure of the Great Basin, *Geophys. Res. Lett.*, *17*, 1327–1330.

Dueker, K., H. Yuan, and B. Zurek (2001), Thick-structured Proterozoic lithosphere of the Rocky Mountain region, *GSA Today*, *11*, 4–9.

Eaton, G. P., R. L. Christiansen, H. M. Iyer, A. M. Pitt, D. R. Mabey, H. R. Blank Jr., I. Zietz, and M. E. Gettings (1975), Magma beneath Yellowstone National Park, *Science*, *188*, 787–796.

Erslev, E. A., and J. F. Sutter (1990), Evidence for Proterozoic mylonitization in the northwestern Wyoming province, *Geol. Soc. Am. Bull.*, *102*, 1681–1694.

Evans, J. R. (1982), Compressional wave velocity structure of the upper 350 km under the eastern Snake River Plain near Rexburg, Idaho, *J. Geophys. Res.*, *87*, 2654–2670.

Evans, J. R., and J. J. Zucca (1988), Active high-resolution seismic tomography of compressional wave velocity and attenuation structure at Medicine Lake Volcano, northern California Cascade Range, *J. Geophys. Res.*, *93*, 15,016–15,036.

Farnetani, C. G., and H. Samuel (2005), Beyond the thermal plume paradigm, *Geophys. Res. Lett.*, *32*, L07311, doi:10.1029/2005GL022360.

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

Faul, U. H., D. R. Toomey, and H. S. Waff (1994), Intergranular basaltic melt is distributed in thin, elongated inclusions, *Geophys. Res. Lett.*, *21*, 29–32.

Favela, J., and D. L. Anderson (2000), Extensional tectonics and global volcanism, in *Problems in Geophysics for the New Millennium*, edited by E. Boschi, G. Ekström, and A. Morelli, pp. 463–498, Compositori, Bologna, Italy.

Fee, D., and K. Dueker (2004), Mantle transition zone topography and structure beneath the Yellowstone hotspot, *Geophys. Res. Lett.*, *31*, L18603, doi:10.1029/2004GL020636.

Foulger, G. R., and J. H. Natland (2003), Is "hotspot" volcanism a consequence of plate tectonics?, *Science*, *300*, 921–922.

Foulger, G. R., et al. (2001), Seismic tomography shows that upwelling beneath Iceland is confined to the upper mantle, *Geophys. J. Int.*, *146*, 504–530.

Gilbert, H. J., A. F. Sheehan, K. G. Dueker, and P. Molnar (2003), Receiver functions in the western United States, with implications for upper mantle structure and dynamics, *J. Geophys. Res.*, *108*(B5), 2229, doi:10.1029/2001JB001194.

Godey, S., F. Deschamps, and J. Trampert (2004), Thermal and compositional anomalies beneath the North American continent, *J. Geophys. Res.*, *109*, B01308, doi:10.1029/2002JB002263.

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.

Goes, S., R. Govers, and P. Vacher (2000), Shallow mantle temperatures under Europe from *P* and *S* wave tomography, *J. Geophys. Res.*, *105*, 11,153–11,169.

Gripp, A. E., and R. G. Gordon (2002), Young tracks of hotspots and current plate velocities, *Geophys. J. Int.*, *150*, 321–361.

Hamilton, W. B., and W. B. Myers (1966), Cenozoic tectonics of the western United States, *Rev. Geophys.*, *4*, 509–549.

Hammond, W. C., and E. D. Humphreys (2000), Upper mantle seismic wave velocity: Effects of realistic partial melt geometries, *J. Geophys. Res.*, *105*, 10,975–10,986.

Hernlund, J. W., and P. J. Tackley (2003), Post-Laramide volcanism and upper mantle dynamics in the western US: Role of small-scale convection, *Eos Trans. AGU*, *84*(46), Fall Meet. Suppl., Abstract V31F-07.

Humphreys, E. D., and K. G. Dueker (1994a), Physical state of the western U.S. upper mantle, *J. Geophys. Res.*, *99*, 9635–9650.

Humphreys, E. D., and K. G. Dueker (1994b), Western U.S. upper mantle structure, *J. Geophys. Res.*, *99*, 9615–9634.

Humphreys, E. D., K. G. Dueker, D. L. Schutt, and R. B. Smith (2000), Beneath Yellowstone; evaluating plume and nonplume models using teleseismic images of the upper mantle, *GSA Today*, *10*, 1–7.

Husen, S., R. B. Smith, and G. P. Waite (2004), Evidence for gas and magmatic sources beneath the Yellowstone Volcanic Field from seismic tomographic imaging, *J. Volcanol. Geotherm. Res.*, *131*, 397–410, doi:10.1016/S0377-0273(03)00416-5.

Iyer, H. M., J. R. Evans, G. Zandt, R. M. Stewart, J. M. Coakley, and J. N. Roloff (1981), A deep low-velocity body under the Yellowstone caldera, Wyoming: Delineation using teleseismic *P*-wave residuals and tectonic interpretation: Summary, *Geol. Soc. Am. Bull.*, *92*, 792–798, 1471–1646.

Jordan, B. T., A. L. Grunder, R. A. Duncan, and A. L. Deino (2004), Geochronology of age-progressive volcanism of the Oregon High Lava Plains: Implications for the plume interpretation of Yellowstone, *J. Geophys. Res.*, *109*, B10202, doi:10.1029/2003JB002776.

Jordan, M., R. B. Smith, and G. P. Waite (2004), Tomographic images of the Yellowstone hotspot structure, *Eos Trans. AGU*, *85*, Fall Meet. Suppl., Abstract V51B-0556.

Jordan, T. H. (1979), Mineralogies, densities and seismic velocities of garnet lherzolites and their geophysical implications, in *The Mantle Sample: Inclusions in Kimberlites and other Volcanics*, edited by F. R. Boyd and H. O. A. Meyer, pp. 1–14, AGU, Washington, D. C.

Karato, S.-I. (1993), Importance of anelasticity in the interpretation of seismic tomography, *Geophys. Res. Lett.*, *20*, 1623–1626.

Karato, S.-I., and H. Jung (1998), Water, partial melting and the origin of the seismic low velocity and high attenuation zone in the upper mantle, *Earth Planet. Sci. Lett.*, *157*, 193–207.

Kawamoto, T., and J. R. Holloway (1997), Melting temperature and partial melt chemistry of H_2O-saturated mantle peridotite to 11 gigapascals, *Science*, *276*, 240–243.

Keller, W. R., D. L. Anderson, and R. W. Clayton (2000), Resolution of tomographic models of the mantle beneath Iceland, *Geophys. Res. Lett.*, *27*, 3993–3996.

Kendall, J. M. (1994), Teleseismic arrivals at a mid-ocean ridge; effects of mantle melt and anisotropy, *Geophys. Res. Lett.*, *21*, 301–304.

Kennett, B. L. N., and E. R. Engdahl (1991), Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.*, *105*, 429–465.

Keyser, M., J. R. R. Ritter, and M. Jordan (2002), 3D shear-wave velocity structure of the Eifel Plume, Germany, *Earth Planet. Sci. Lett.*, *203*, 59–82.

Kohlstedt, D. L., H. Keppler, and D. C. Rubie (1996), Solubility of water in the α, β and γ phases of $(Mg,Fe)_2SiO_4$, *Contrib. Mineral. Petrol.*, *123*, 345–357.

Lehman, J. A., R. B. Smith, M. M. Schilly, and L. W. Braile (1982), Upper crustal structure of the Yellowstone caldera from delay time analyses and gravity correlations, *J. Geophys. Res.*, *87*, 2713–2730.

Lemieux, S., G. Ross, and F. A. Cook (2000), Crustal geometry and tectonic evolution of the Archean crystalline basement beneath the southern Alberta Plains, from new seismic reflection and potential-field studies, *Can. J. Earth Sci.*, *37*, 1473–1491.

Leveque, 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.

Levin, V., W. Menke, and A. Lerner-Lam (1996), Seismic anisotropy in the north-eastern US as a source of significant teleseismic *P* traveltime anomalies, *Geophys. J. Int.*, *126*, 593–603.

Lynch, D. P. (1999), Three-dimensional finite difference tomography of the Basin and Range-Colorado Plateau-Rocky Mountain transition using earthquake and controlled source data, M.S. thesis, Univ. of Utah, Salt Lake City.

Meibom, A., D. L. Anderson, N. H. Sleep, R. Frei, C. P. Chamberlain, M. T. Hren, and J. L. Wooden (2003), Are high $^3He/^4He$ ratios in oceanic basalts an indicator of deep-mantle plume components?, *Earth Planet. Sci. Lett.*, *208*, 197–204, doi:10.1016/S0012-821X(03)00038-4.

Miller, D. S., and R. B. Smith (1999), *P* and *S* velocity structure of the Yellowstone volcanic field from local earthquake and controlled-source tomography, *J. Geophys. Res.*, *104*, 15,105–15,121.

Minster, J. B., and D. L. Anderson (1981), A model of dislocation-controlled rheology for the mantle, *Philos. Trans. R. Soc. London*, *299*, 319–356.

Montelli, R., G. Nolet, F. A. Dahlen, G. Masters, E. R. Engdahl, and S.-H. Hung (2004), Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, *303*, 338–343, doi:10.1126/science.1092485.

Morgan, W. J. (1971), Convection plumes in the lower mantle, *Nature*, *230*, 42–43.

Morgan, W. J. (1972), Deep mantle convection plumes and plate motions, *Am. Assoc. Pet. Geol. Bull.*, *56*, 203–213.

Nolet, G. (1993), Solving large linearized tomographic problems, in *Seismic Tomography: Theory and Practice*, edited by H. M. Iyer and K. Hirahara, CRC Press, Boca Raton, Fla.

Nolet, G., and F. A. Dahlen (2000), Wavefront healing and the evolution of seismic delay times, *J. Geophys. Res.*, *105*, 19,043–19,054.

Paige, C. C., and M. A. Saunders (1982), LSQR: An algorithm for sparse linear equations and sparse least squares, *Trans. Math. Software*, *8*, 43–71.

Pelton, J. R., and R. B. Smith (1982), Contemporary vertical surface displacements in Yellowstone National Park, *J. Geophys. Res.*, *87*, 2745–2761.

Perkins, M. E., and B. P. Nash (2002), Explosive silicic volcanism of the Yellowstone hotspot: The ashfall tuff record, *Geol. Soc. Am. Bull.*, *114*, 367–381.

Pierce, K. L., and L. A. Morgan (1992), The track of the Yellowstone hot spot; volcanism, faulting, and uplift, in *Regional Geology of Eastern Idaho and Western Wyoming*, edited by P. K. Link, M. A. Kuntz, and L. B. Platt, *Mem. Geol. Soc. Am.*, *179*, 1–53.

Puskas, C. M., R. B. Smith, and C. M. Meertens (2002), GPS-derived models of intraplate deformation of the Yellowstone Hotspot, *Eos Trans. AGU*, *83*(47), Fall Meet. Suppl., Abstract F1037.

Ritter, J. R. R., M. Jordan, U. R. Christensen, and U. Achauer (2001), A mantle plume below the Eifel volcanic fields, Germany, *Earth Planet. Sci. Lett.*, *186*, 7–14.

Saltzer, R. L., and E. D. Humphreys (1997), Upper mantle *P* wave velocity structure of the eastern Snake River Plain and its relationship to geodynamic models of the region, *J. Geophys. Res.*, *102*, 11,829–11,841.

Sato, H. (1992), Thermal structure of the mantle wedge beneath northeastern Japan: Magmatism in an island arc from the combined data of seismic anelasticity and velocity and heat flow, *J. Volcanol. Geotherm. Res.*, *51*, 237–252.

Schutt, D. L., and E. D. Humphreys (2001), Evidence for a deep asthenosphere beneath North America from western United States *SKS* splits, *Geology*, *29*, 291–294.

Schutt, D. L., and E. D. Humphreys (2004), *P* and *S* wave velocity and V_P/V_S in the wake of the Yellowstone hot spot, *J. Geophys. Res.*, *109*, B01305, doi:10.1029/2003JB002442.

Schutt, D. L., and C. E. Lesher (2006), Effects of melt depletion on the density and seismic velocity of garnet and spinel lherzolite, *J. Geophys. Res.*, doi:10.1029/2003JB002950, in press.

Simmons, N. A., and H. Gurrola (2000), Multiple seismic discontinuities near the base of the transition zone of the Earth's mantle, *Nature*, *405*, 559–562, doi:10.1038/35014589.

Smith, R. B. (1977), Intraplate tectonics of the western North American Plate, *Tectonophysics*, *37*, 323–336.

Smith, R. B., and L. W. Braile (1994), The Yellowstone hotspot, *J. Volcanol. Geotherm. Res.*, *61*, 121–187.

Smith, R. B., and M. L. Sbar (1974), Contemporary tectonics and seismicity of the western United States with emphasis on the Intermountain Seismic Belt, *Geol. Soc. Am. Bull.*, *85*, 1205–1218.

Smith, R. B., G. P. Waite, C. M. Puskas, D. L. Schutt, and E. D. Humphreys (2003), Dynamic and kinematic models of the Yellowstone Hotspot constrained by seismic anisotropy, GPS measurements and fault slip rates, *Eos Trans. AGU*, *84*(46), Fall Meet Suppl., Abstract T51G-05.

Sobolev, S. V., H. Zeyen, G. Stoll, F. Werling, R. Altherr, and K. Fuchs (1996), Upper mantle temperatures from teleseismic tomography of French Massif Central including effects of composition, mineral reactions, anharmonicity, anelasticity and partial melt, *Earth Planet. Sci. Lett.*, *139*, 147–163.

Spetzler, J., and R. Snieder (2004), The Fresnel volume and transmitted waves, *Geophysics*, *69*, 653–663, doi:10.1190/1.1759451.

Steinberger, B. (2000), Plumes in a convecting mantle; models and observations for individual hotspots, *J. Geophys. Res.*, *105*, 11,127–11,152.

Tackley, P. J., and D. J. Stevenson (1993), A mechanism for spontaneous self-perpetuating volcanism on the terrestrial planets, in *Flow and Creep in the Solar System: Observations, Modeling and Theory*, edited by D. B. Stone and S. K. Runcorn, pp. 307–322, Springer, New York.

Thompson, A. B. (1992), Water in the Earth's upper mantle, *Nature*, *358*, 295–302.

Umino, N., and A. Hasegawa (1984), Three-dimensional Q_S structure in the northeastern Japan arc, *J. Seismol. Soc. Jpn.*, *37*, 217–228.

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*, 150–159.

Vacher, P., A. Mocquet, and C. Sotin (1998), Computation of seismic profiles from mineral physics: The importance of the non-olivine components for explaining the 660 km depth discontinuity, *Phys. Earth Planet. Inter.*, *106*, 275–298.

Waite, G. P. (2004), Upper mantle structure of the Yellowstone hotspot from teleseismic body-wave velocity tomography and shear-wave anisotropy, Ph.D. thesis, Univ. of Utah, Salt Lake City.

Waite, G. P., and R. B. Smith (2002), Seismic evidence for fluid migration accompanying subsidence of the Yellowstone caldera, *J. Geophys. Res.*, *107*(B9), 2177, doi:10.1029/2001JB000586.

Waite, G. P., R. B. Smith, C. M. Puskas, D. L. Schutt, and R. M. Allen (2003), Kinematics of the Yellowstone hotspot derived from seismic anisotropy, tomography, and GPS, *Geophys. Res. Abstr.*, *5*, 00670.

Waite, G. P., D. L. Schutt, and R. B. Smith (2005), Models of lithosphere and asthenosphere anisotropic structure of the Yellowstone hotspot from shear wave splitting, *J. Geophys. Res.*, *110*, B11304, doi:10.1029/2004JB003501.

Waldhauser, F., R. Lippitsch, E. Kissling, and J. Ansorge (2002), High-resolution teleseismic tomography of upper-mantle structure using an a priori three-dimensional crustal model, *Geophys. J. Int.*, *150*, 403–414.

Walker, K. T. (2004), Exploring Problems in Tectonics and Geodynamics with Seismology, Ph.D. thesis, Stanford Univ., Palo Alto, Calif.

Wicks, C. W. J., W. R. Thatcher, and D. Dzurisin (1998), Migration of fluids beneath Yellowstone caldera inferred from satellite radar interferometry, *Science*, *282*, 458–462.

Wielandt, E. (1987), On the validity of the ray approximation for interpreting delay times, in *Seismic Tomography*, edited by G. Nolet, pp. 85–98, Springer, New York.

Wolfe, C. J., I. T. Bjarnason, J. C. VanDecar, and S. C. Solomon (2002), Assessing the depth resolution of tomographic models of upper mantle structure beneath Iceland, *Geophys. Res. Lett.*, *29*(2), 1015, doi:10.1029/2001GL013657.

Wolfe, C. J., S. C. Solomon, P. G. Silver, J. C. VanDecar, and R. M. Russo (2002), Inversion of body-wave delay times for mantle structure beneath the Hawaiian islands: Results from the PELENET experiment, *Earth Planet. Sci. Lett.*, *198*, 129–145.

Yuan, H., and K. Dueker (2005), Teleseismic *P*-wave tomogram of the Yellowstone plume, *Geophys. Res. Lett.*, *32*, L07304, doi:10.1029/2004GL022056.

R. B. Smith, Department of Geology and Geophysics, University of Utah, 135 S. 1460 E., Salt Lake City, UT 84112, USA.

G. P. Waite, U.S. Geological Survey, 345 Middlefield Road, MS-910, Menlo Park, CA 94025, USA. (gwaite@usgs.gov)

R. M. Allen, Seismological Laboratory, Department of Earth and Planetary Science, University of California, 215 McCone Hall, Berkeley, CA 94720, USA.