JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 105, NO. B12, PAGES 28,139–28,157, DECEMBER 10, 2000

Tomographic inversion of *S-SKS* times for shear velocity heterogeneity in D": Degree 12 and hybrid models

Ban-Yuan Kuo

Institute of Earth Sciences, Academia Sinica, Taipei, Taiwan

Edward J. Garnero

Department of Geology, Arizona State University, Tempe, Arizona

Thorne Lay

Institute of Tectonics, University of California, Santa Cruz, California

Abstract. Differential travel times between *S* (or diffracted *S*) and *SKS* were measured to study the global distribution of shear wave velocity heterogeneity in the lowermost 250 km of the mantle (the D" region). Commencing with ~3000 *S-SKS* times with variable qualities, we minimize uneven path coverage by thinning redundantly sampled regions (e.g., the Fiji-Tonga to North America corridor) and collect additional data for sparsely sampled areas, especially in the Southern Hemisphere. About 1500 paths were retained, distributed to reconcile both high-density and homogeneous sampling. We compare (1) spherical harmonic and (2) equal-area block parameterizations of D" shear velocity heterogeneity for identical minimum resolving lengths and mean model errors. We show that the two parameterizations result in indistinguishable patterns of heterogeneity and power distribution for resolution down to a block size of 4.5°×4.5° and maximum spherical harmonic degree of 40. We demonstrate synthetically that a high-degree (L) inversion followed by a lower-degree (L_s) spherical harmonic synthesis effectively circumvents model contamination from expansion truncation effects. Tomographic inversion of this data set yields a global distribution of D" heterogeneity robust up to degree 12. In our preferred model (L=40, L_s=12, rms ~1%), surface hotspots and estimated lower mantle plume roots are located in or at edges of low-velocity regions (δV_s<-2%), for both the Atlantic-Indian low-velocity corridor as well as the low velocities beneath the central Pacific. Although the above parameterizations entail no intrinsic difference in resolvability, higher resolution is regionally achieved by removing the derived degree 12 model from the data then inverting the corrected residuals using 4.5°×4.5° blocks. This hybrid method prevents power loss at intermediate to long wavelengths from otherwise severe damping and recovers smaller-scale structure where constraints are better than the global average, such as beneath the Pacific and Eurasia. A synthetic recovery test yields resolvability maps that reflect both the path geometry and the quality of data, which help to identify robust features in both the degree 12 and the hybrid D" models.

1. Introduction

The mantle's lower boundary layer, the D" region [*Bullen*, 1949], is heterogeneous on a wide range of scale lengths. Forward and inverse modeling of seismic travel times and waveforms has imaged both regional and global heterogeneity, providing a diverse range of shapes, sizes and strengths of lower mantle perturbations in a boundary layer up to 400 km thick. The pioneering tomographic inversions for mantle structure [e.g., *Dziewonski*, 1984; *Dziewonski and Woodhouse*, 1987; *Tanimoto*, 1990] yielded lower mantle *P* and/or *S* velocity variations with maximum degree of expansion (L) in spherical harmonics to degree 6 (L=6). These studies and others established the remarkable correlation between long-wavelength lower mantle heterogeneity patterns and surficial features, such as the lowermost mantle ring of circum-Pacific high velocities that underlies past and present subduction zones.

The past decade has seen a nearly twofold increase in global structure resolution, with inversions up to degree 12 or higher. Inversion methods have included use of arrival times reported in the International Seismological Centre (ISC) bulletin [e.g., *Vasco and Johnson*, 1998; *van der Hilst et al.*, 1997, 1998], normal mode coefficients [e.g., *Resovsky and Ritzwoller*, 1999], differential travel times along with long-period mantle waves [e.g., *Su et al.*, 1994; *Liu and Dziewonski*, 1998], normal modes and differential times [e.g., *Masters et al.*, 1996] and whole waveforms [*Li and Romanowicz*, 1996], as well as hand-picked travel times [e.g., *Grand et al.*, 1997].

Paper number 2000JB900334.
0148-0227/00/2000JB900334$09.00

Global inversions using a block parameterization, with sizes of a few degrees, rather than spherical harmonics have provided the most compelling suggestion of the connection between dynamical processes and velocity heterogeneity. For example, long tabular sheet-like high-velocity features beneath subduction zones are consistent with cold dense slabs penetrating into the lower mantle [e.g., *Grand et al.*, 1997; *van der Hilst et al.*, 1997, 1998]. A low-velocity conduit appears to extend from the core-mantle boundary to the Earth's surface beneath Iceland, suggesting a whole mantle plume [*Bijwaard and Spakman*, 1999; *Ritsema et al.*, 1999].

While it is desirable to resolve shorter-wavelength structure, it is important to recognize the difficult task of comparing results of one inversion to another, as well as understanding what parts of the structure are well resolved. Past studies invert different data sets, tolerate different model variances, and adopt different regularization philosophy and schemes. Thus degree L and block size (B) do not indicate the actual resolving scale length. D" models extracted from the above sources differ in heterogeneity pattern and power spectrum, much of which is due to a priori selection of model parameterization. A formidable aspect of inversion for three-dimensional (3-D) mantle structure is the particular difficulty in evaluating which structural features, especially those with shorter wavelengths, are actually constrained by the data.

Recently, global-scale single-layer (i.e., two-dimensional) maps of D" velocity variations have been determined by analysis of differential travel times sensitive to lower mantle structure. Some examples include maps of large-scale P velocity variation in the lowermost few hundred kilometers of the mantle from P_{diff}–PKP times [*Wysession*, 1996]. *Kuo and Wu* [1997] (hereinafter referred to as paper 1) inverted S_{diff}–SKS times for D" shear structure. Paper 1 adopted a conservative approach, inverting only up to spherical harmonic degree 6. Differential times between S and core-reflected ScS have also been useful for imaging patterns of deep mantle heterogeneity [e.g., *Woodward and Masters*, 1991; *Wysession et al.*, 1994].

Differential times have the advantage of minimizing possible contamination effects due to source mislocation and/or unmapped upper mantle heterogeneity; perturbations in these times tend to be most sensitive to Earth structure along the two paths where they are most distant. In the case of S–SKS differential times, sensitivity to velocity variations is greatest along both paths in the lower mantle and along the relatively long core-diffraction part of the S_{diff} path, for SKS bottoms in the outer core, which is likely homogenous due to rapid mixing [e.g., *Stevenson*, 1987]. S–SKS times are therefore potentially effective in situ probes of large-scale D" heterogeneity (Figure 1). In contrast, ScS-S times involve the S phase, which bottoms in the heterogeneous middle and lower mantle, and there is greater reliance on correction for mantle heterogeneity, since there is commonly a greater separation between S and ScS bottoming depths. While this is certainly not an insurmountable issue, in this paper we adopt the approach of limiting dependence of our results on correction for structure outside of D" by solely using S–SKS. Another advantage to the S–SKS data set is the fairly uniform global coverage that can be achieved. Inverting differential times for a single-layer model is a much more tractable inversion problem than typical whole mantle 3-D tomography, and the inversion itself can be examined empirically and the model easily compared to the original data. Of course, our 2-D model is inadequate in representing 3-D structures in the lower mantle.

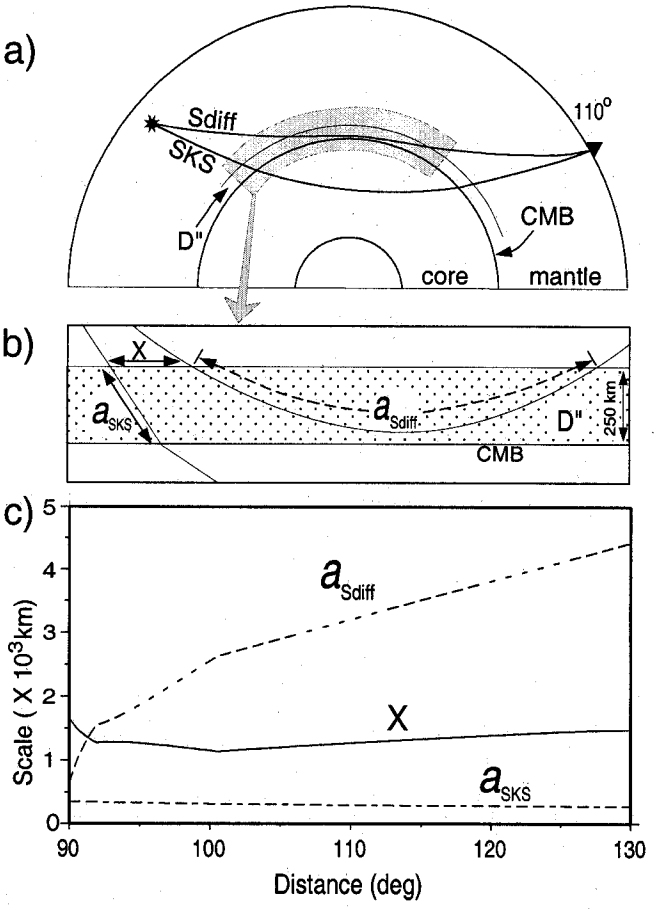

Figure 1. (a) Cross section of Earth showing ray paths of S_{diff} and SKS at an epicentral distance of 110° from a 600-km-deep event. The proximity of the two ray paths in the upper mantle reduces the effects of hypocentral mislocation and shallow lateral heterogeneity in S-SKS differential times. (b) Zoom in of the shaded part of Figure 1a, showing the definition of various length parameters: a_{Sdiff} and a_{SKS} are path lengths of S_{diff} and SKS traversing a 250-km-thick D" layer. X denotes the separation between the two phases upon entering the top of the layer. (c) The lengths defined in Figure 1b plotted against epicentral distance. The sensitivity of S overwhelms that of SKS to heterogeneity in D" due to its much longer sampling path. The nearly constant value of X roughly corresponds to the resolving length of degree 8, implying that a degree 12 mantle model for D" should be able to explain the S-SKS times.

In this paper, we adopt the term D" to represent the lowermost 250 km of the mantle. This distance is the upper limit of the wavelengths of our long-period (LP) analog and digital records and is roughly the width of the T/4 Fresnel zone for LP digital record S_{diff} above the core-mantle boundary (CMB) [*Kuo*, 1999]. The term D" was first designated to the base of the mantle by *Bullen* [1949] to recognize the decrease of the shear velocity gradient. In later 1-D reference models such as Preliminary Reference Earth Model (PREM) [*Dziewonski and Anderson*, 1981] and iasp91 [*Kennett and Engdahl*, 1991], this reduced gradient zone is ~150 km thick. Since the early 1980s when a velocity discontinuity some 250-300 km above the CMB had been mapped [*Lay and Helmberger*, 1983], the D" region has been routinely characterized as distinct from the mantle above. This velocity jump is often assumed to represent the top of the D" layer. In the geographically limited

samples available up to date, variable D" thicknesses from 100 to 400 km and from well defined to undetectable have been reported [e.g., *Wysession et al.*, 1998]. In paper 1 and this paper, we used the term D" to represent our target depth region, the lowermost 250 km of the mantle. Thus we do not attempt to address exactly where or what D" is, or if such can be defined at all. The effect of assuming a different target thickness is minor in both pattern and magnitude of heterogeneity, compared with that of other arbitrary choices in the inversion, as discussed in paper 1.

We extend the low-resolution modeling of paper 1 to more detailed inversion for D" heterogeneity. The *S-SKS* data set has been significantly expanded to densify, as homogeneously as possible, global coverage of D" with the goal of achieving a higher-resolution map of shear velocity variations. We first address a few issues regarding the inversion itself: spherical harmonic versus block parameterizations, aliasing or leakage effects from expansion truncation, and the strategy to mitigate such effects. We then demonstrate that the current data support a global structure robustly resolved up to degree 12. Maps that reveal variable resolution of the model are constructed through synthetic recovery analysis. This is then followed by a modeling in which blocks of velocity are superimposed on our a priori derived spherical harmonic aspherical reference structure to enhance regional resolution where such improvement is warranted. Interesting dynamic implications of this D" model are also addressed. In this paper, we select several whole mantle tomography models from the literature that are used to correct our data for non-D" heterogeneity and/or to compare with our results.

2. *S-SKS* Differential Travel Time Data

As mentioned above, differential travel times reduce contamination from possible errors in hypocentral location and origin time and from heterogeneity in the upper mantle beneath the source and receiver. Paper 1 utilized 340 Global Digital Seismic Network (GDSN) long-period (T=25-30 s) recordings and Global Seismic Network (GSN) broadband recordings of *S-SKS* filtered to the LP response of the GDSN data. Epicentral distances were restricted to be > 100°, and lower mantle ray paths were distributed over the globe as evenly as possible.

In this study, we expand the data set of paper 1, adding new measurements from high-quality broadband digital GSN and portable Program for Array Seismic Studies of the Continental Lithosphere (PASSCAL) experiment data, as well as analog LP World-Wide Standardized Seismograph Network (WWSSN) data. Some of the broadband records are low-pass filtered to 10 s and onset picked, and their differential travel times are practically no different from those for LP records measured by first-swing matching (explained below). We have also included WWSSN measurements from two earlier regional studies with data sampling beneath the Pacific [*Garnero and Helmberger*, 1993] and beneath the North Pacific/Alaska region [*Young and Lay*, 1990]. These records have a dominant period at 15 s. The minimum epicentral distance for the present paper has been reduced to 95°. At this distance, standard radial reference structures predict the *S* wave bottoming depth to be ~100 km above the core-mantle boundary, with a path length in a 250-km-thick lowermost mantle shell ~30° long (~2000 km, Figure 1c). Strongly negative velocity gradients in D" (as predicted by *Ritsema et al.* [1997] for D" beneath the Pacific) can significantly change bottoming depths of core-grazing and diffracted phases. However, for our present purpose of deriving a 2-D map of D" heterogeneity (i.e., not considering radial dependency of shear wave perturbations), this should not significantly affect our results.

Our differential travel time measurement method is based on matching the onset or the first swing of *SKS* on the radial (*SV*) component and *S* on the transverse (*SH*) component of motion. First-swing alignment minimizes the finite frequency effect of waveform on determination of travel times, as elaborated in section 3.1, and as a result, it minimizes the discrepancy in travel times for records of different frequency contents as well. For events shallower than 50 km, differential times are measured by aligning first swings of data waveforms and reflectivity synthetic seismograms [e.g., *Kennett*, 1985; *Müller*, 1985] to account for possible depth phase interference and/or to rule out mispicks involving the surface reflections *sSKS* and sS_{diff}. This was unnecessary for the archive analog data, which are from events deeper than 50 km. First-swing alignment also minimizes the difference in travel times measured from different period of records. Data were classified as A, B, or C quality, with measurement uncertainty of 1, 2, and 3 s, respectively. These bounds were somewhat conservatively applied, resulting in 437 A, 603 B, and 456 C classifications; 52% of the data from events with source depth < 50 km are C quality (311 out of 595 measurements).

As described below, from an initial data set of roughly 3000 carefully measured *S-SKS* times, we retain about 1500 times chosen for best data quality and path coverage distribution. The *S* and *SKS* paths within a 250 km thick D" layer (i.e., a_{SKS} and a_{Sdiff} in Figure 1b) of the final data set, calculated for the PREM 1-D reference structure, are shown in Figure 2a. Decent global coverage with a relatively uniform distribution is achieved. This was facilitated by collecting data from shallow events that enhance coverage in the Southern Hemisphere, as well as through addition of data from portable broadband arrays.

Contributions to the differential times from mantle heterogeneity above the D" layer are estimated using model SKS12WM13 [*Dziewonski et al.*, 1997; *Liu and Dziewonski*, 1998] and removed from the raw times, resulting in a 22% variance reduction of the residuals. Later we discuss the dependence of our results on choice of model for this correction. Other corrections include receiver-side azimuthal anisotropy effects using the compilation of *Silver* [1996] and *Wolfe and Silver* [1998] and ellipticity correction [*Kennett and Gudmundsson*, 1996]. Synthetic anisotropic waveform experiments like that used in paper 1 indicate that the anisotropy correction (δt) can be parameterized as $\delta t=\tau\cos(2\theta)$, where θ is the angle between the polarization of *SKS* and the fast direction of anisotropy and τ is the delay time of the station. This is a realistic approximation considering the uncertainty of the anisotropy and the complexity in the waveform calculation. Travel time errors associated with event mislocation are ~0.5 s for 15 km epicenter error and ~0.2 s for a 10 km depth error. Errors in anisotropy and asphericity (above D") corrections are estimated to be up to 0.5 and 1.25 s, respectively; these corrections are assigned to each residual along with the measurement error. The resulting corrected *S-SKS* residuals as referenced to PREM are plotted at the midpoint of the *S* (and S_{diff}) paths in Figure 2b. The fully corrected residuals range roughly from -20 to 22 s, with a mean of ~1 s and a standard deviation of 5.0 s; the histogram of these data is shown in

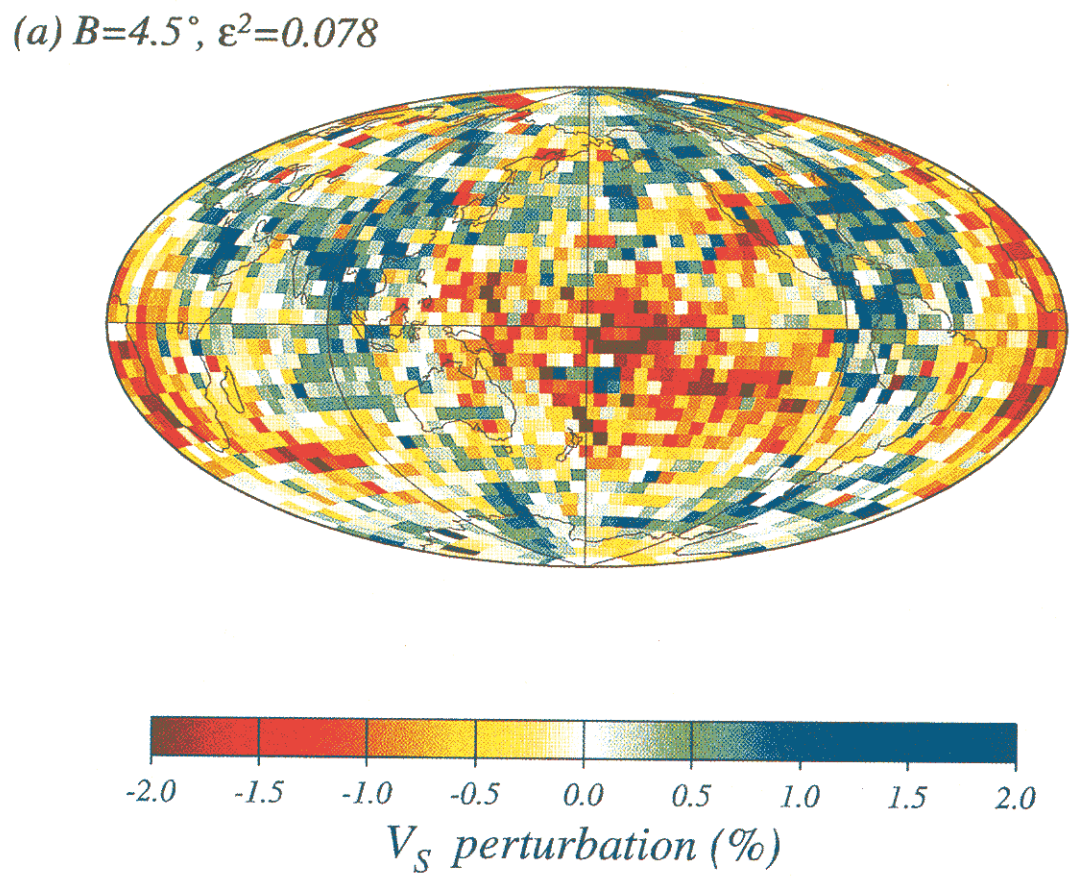

Plate 1. Heterogeneity images for models with (a) B=4.5° and (b) L=40, extracted from the box in Figure 3a. The images are very similar to the shortest resolving length, leading to the conclusion that these two types of parameterizations are not intrinsically different given our data geometry, as long as they are compared for comparable model variance.

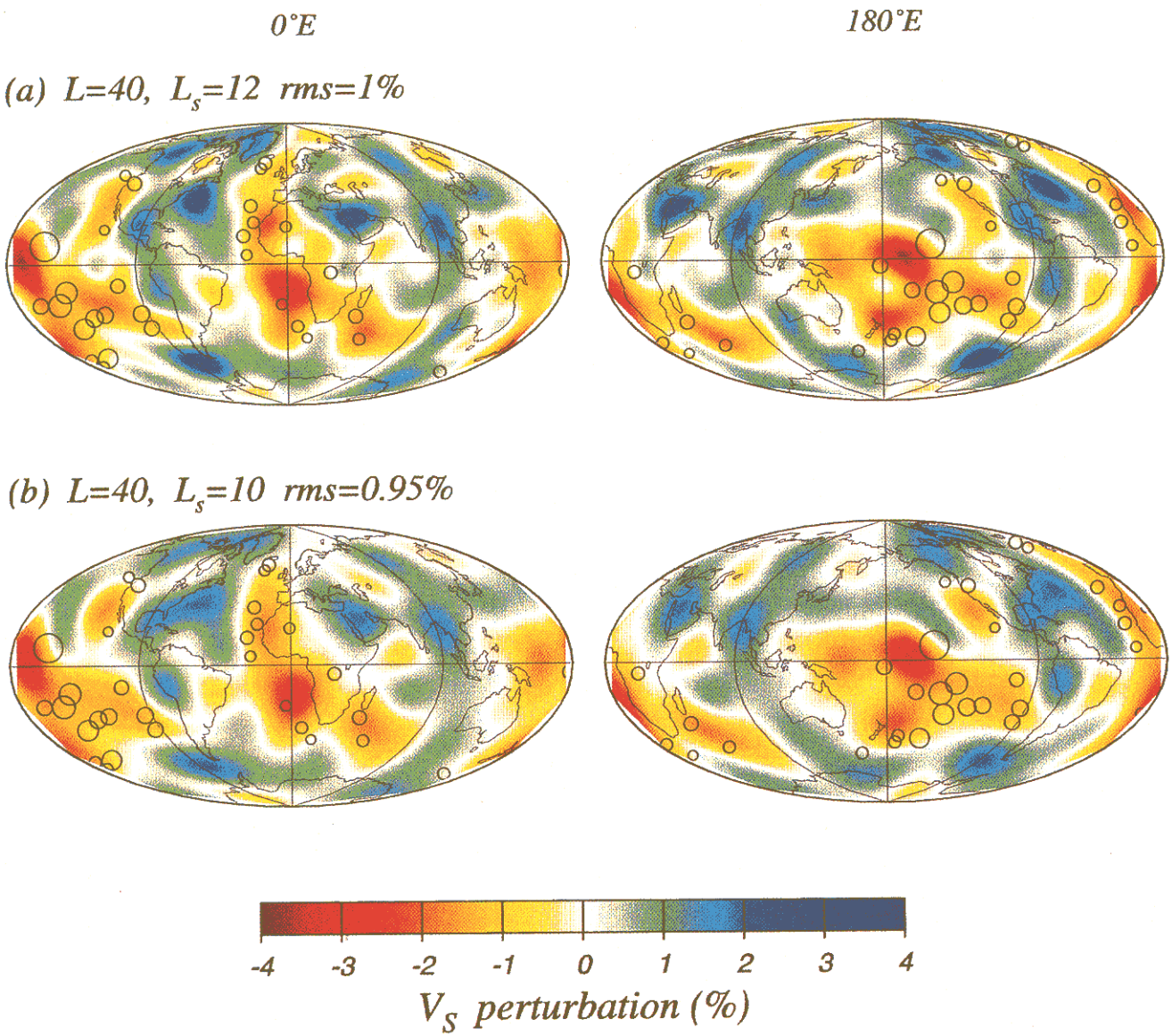

Plate 2. (a) D" shear velocity model with L=40, L_s=12, rms=1% and (b) D" model with L=40, L_s=10, rms=0.95% that is extracted from that in Plate 2a. (left) Maps centered at 0°E; (right) maps centered at 180°E. The degraded resolution in Plate 2b is compensated by higher robustness as indicated in Figure 7. Circles mark the positions of the plume roots at the CMB calculated by *Steinberger* [2000]. Note the possible clustering of roots with zones of low velocities in the Atlantic (left) and the distribution of roots around other low-velocity anomalies.

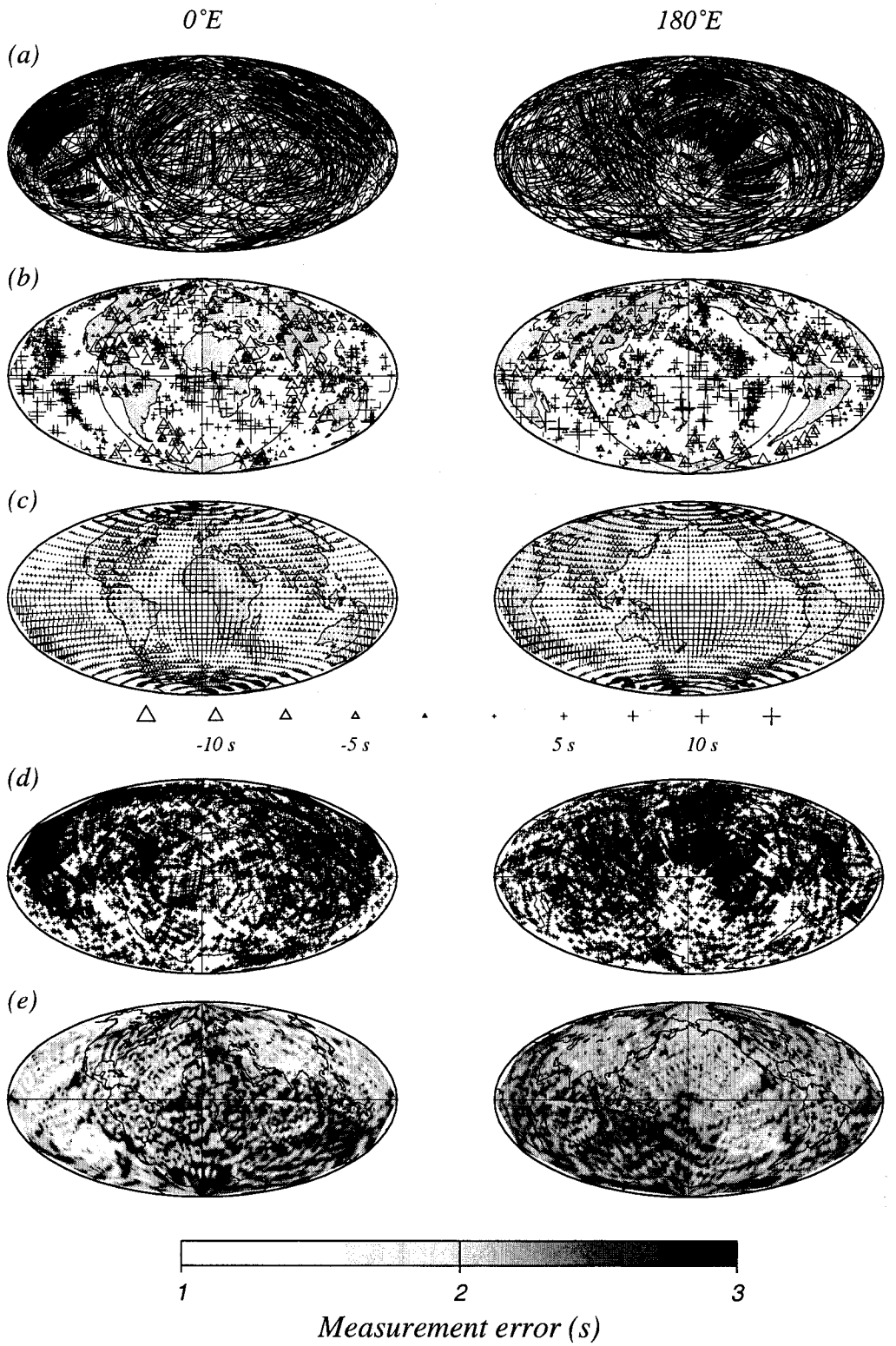

Figure 2. Maps displaying attributes of our data set. Figures on the left (right) column are centered at 0°E (180°E), and continents are outlined in all figures and lightly shaded in Figures 2b and 2c. (a) Surface projections of the paths of *S* and *SKS* within a 250-km-thick D" layer for the final data set. (b) Fully corrected *S-SKS* residuals plotted at the midpoints of the S segments in Figure 2a. (c) The 5° cap-averaged residuals showing long-wavelength variations in the midpoint anomaly distribution. (d) Locations where *S* segments in Figure 2a intersect at an angle >10°, representing ray crossings that provide lateral constraints and minimize systematic errors associated with mantle asphericity and mislocation. This map indicates that the Fiji-Tonga basin, southeastern Pacific, southern Atlantic, and southern Africa are relatively poorly constrained in terms of crossing coverage. (e) Measurement error distribution. Quality of data is relatively low in the Southern Hemisphere except the southern central Pacific, where D" is sampled through using Fiji-Tonga deep events. Figures 2d and 2e together determine the variable resolution provided by our data set.

Figure 3a. The large negative residuals are found in regions beneath the circum-Pacific; the large positive residuals are present beneath the central Pacific, the southern Atlantic, and Africa (Figure 2b). There are coherent large-scale patterns in the data, as emphasized by the 5° cap-averaged plot in Figure 2c, as well as small-scale irregularities.

The total number of data in this type of study can be increased fairly readily. A wealth of good quality measurements can be obtained from a few favorable source-receiver geometries, such as between Fiji-Tonga earthquakes and North American stations, providing dense sampling of D" beneath the central Pacific. However, adding redundant paths is undesirable for an inversion problem. First, although repeated sampling suppresses random errors, regional oversampling increases the size of the inversion and biases global statistics such as the mean model error and total variance reduction. This may lead to misjudgment in choosing maximum truncation degree or damping factor. Second, severely inhomogeneous data distribution induces spectral leakage in a truncated parameterization of spherical harmonic functions and thus biases model estimation [*Trampert and Snieder*, 1996], especially for components near the truncation end. We will examine this problem in detail below. Third, redundant sampling for similar azimuths can accumulate nonrandom errors, like those from systematic regional mislocation and mantle asphericity miscorrection, which can then be incorrectly mapped into D" solution structure. To minimize redundant sampling, we sort the rays on a nearly equal-area 5°×5° grid. *S* waves having D" segments with the exact same set of cells represent nearly identical D" paths; this redundancy is eliminated in a hierarchical fashion by discarding data of (1) lower quality, (2) shallower event depth, and (3) analog rather than digital recording. The starting data set contained regions with wide spread redundant sampling (e.g., more than five identical rays per path, such as in the central Pacific). In the densely sampled Pacific we limit ray path redundancy to two, and in more sparsely sampled regions the limit is set at three. This ultimately reduces our data set by half, resulting in the ~1500 samples displayed in Figure 2. This procedure is equivalent to the summary ray approach often taken in global tomography to reduce a huge data set, but we prefer to remove low-quality data if high-quality data exist along the same path, as well as those longer upper mantle paths from shallower sources.

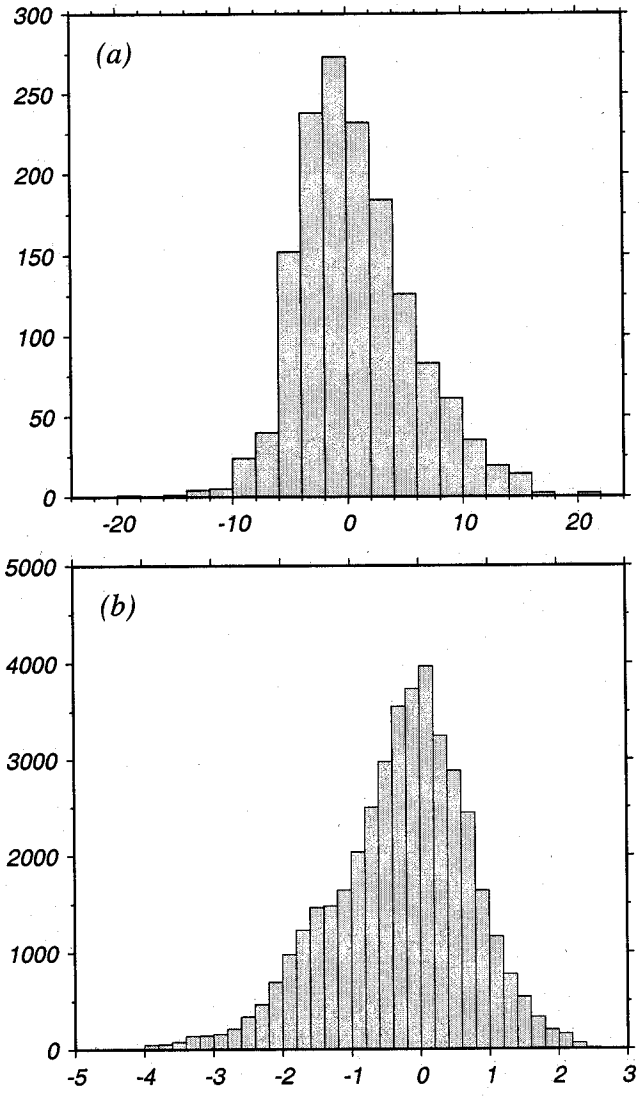

Figure 3. Histograms (a) of the fully corrected *S-SKS* residuals and (b) of velocity perturbation in percentage of our degree 12 model (in Plate 2a). The skewness of Figure 3b is -0.49, which favors basal heating as an indispensable source of convection in the mantle.

In addition to reduction of uneven sampling by the above procedure, we improved upon the coverage in past efforts (e.g., paper 1) by specifically searching for data with paths in undersampled regions, especially in the Southern Hemisphere where coverage is typically much less than that in the Northern Hemisphere. New digital stations from a variety of networks have only recently made possible a significant step forward in path coverage for this type of deep Earth investigation. The PASSCAL experiment of *Wiens et al.* [1995] that deployed stations in the Fiji-Tonga basin received several waveforms from South American events, which approximately reverse the frequently used Tonga to South America path, and we find consistent times from both. As shown in Figure 1, the *S* segments in D" can be relatively long, and constraint on intermediate-scale D" structure comes from the extent of crossing ray path coverage. Crossing path coverage is displayed for our final *S-SKS* data set in Figure 2d. Locations at which *S* rays intersect with angles larger than 10° are indicated by crosses. The crossover map reveals the spatial variation of resolution associated with the geometry of the paths alone. Despite our effort to enhance azimuthal coverage of samples, the crossing distribution is uneven, with particularly limited crossings in the southwest Pacific, southeast Pacific, and southern Atlantic. Figure 2e shows a map of measurement error distribution. Low-quality data are mainly in regions which are covered by paths from shallow events only, especially in Southern Hemisphere, and they are usually associated with few crossings. More data of acceptable quality are needed in these regions, and sustained efforts to improve data coverage will be important. Figures 2 demonstrates the generally acceptable overall coverage provided by our data and allows a qualitative assessment of model resolution, which will be further discussed in the following sections.

3. Inversion Methodology and Results

3.1 Data Kernel: Ray Theory Approximation

As in paper 1, the data kernel of inversion is constructed based on ray theory in the sense that the *S-SKS* times are sensitive to structures along the paths shown in Figure 2a. The

kernel is exact only for infinite-frequency waves and has been used as an approximation for band-limited waves in many global tomographic studies [e.g., *Grand et al.*, 1997; *Van der Hilst et al.*, 1997]. The common practice to validate the approximation is to pick the onset of the waveform as the travel time of the phase, and this is usually attainable for impulsive arrivals on relatively high frequency recordings. The approximation breaks down when measurement of travel time is dominated by the finite-frequency effect on the whole waveform.

Recently, 3-D delay time (residual) kernels for band-limited body waves have revealed a complex pattern of sensitivity to perturbations in the mantle [e.g., *Hung et al.*, 2000; *Zhao et al.*, 2000]. The most remarkable feature of the 3-D kernel is the paradoxical zero sensitivity along the geometrical ray path and maximum sensitivity off the path. The physics behind the "hollow banana" type of sensitivity for teleseismic phases is the integrated effect of diffraction and scattering from off-path heterogeneity and wavefront healing associated with finite-frequency waves. However, it should be noted that the 3-D kernels presented in above references are formulated for "cross-correlation" delay times that consider the accumulated travel time difference over usually the entire period of waveform. Later arrivals from off-path scattering or diffraction from the boundary of an on-path perturbation alter most strongly the waveform after the onset and therefore dominate the cross-correlogram between the perturbed and unperturbed waveforms. Instead, delay times of the onset or first break are least biased by off-path contribution and better approximate the sensitivity of the ray. For the emergent pulses on which exact onset is hard to identify, as the S_{diff} in our long-period records, matching the first swings between observed and synthetic waveforms is a good replacement of onset picking. For band-limited waves, first-swing delay time is less prone to the wavefront healing effect, providing greater sensitivity to small scale perturbation along the raypath than for the cross-correlation delay time.

"First-swing alignment" is technically tenable for band-limited waveforms of our data. We remeasure some of the data in paper 1 to assure that the residuals reflect the shift in the very beginning of the wave packet (see Figure 2 of paper 1 for waveforms). It is probably risky measuring S_{diff}-SKS by waveform cross-correlation because it will be biased by dispersion of the S_{diff} waveform due to attenuation in D". As the first-swing delay times are less "blind" to small-scale anomalies than cross-correlation times, we will restrict the minimum resolving scale in our models to only 3 times of the maximum wavelength of our data, or roughly 600 km, to validate the use of ray theory kernels. In section 3.4 we present degree 12 models whose minimum scale of variation is 900-1000 km. For the hybrid model in section 3.5, strong smoothing with minimum gradient is applied to meet the above criterion. The along-ray mantle asphericity correction performed in previous section is valid because the correction models employed are of long wavelength or highly smoothed.

3.2. Spherical Harmonic Versus Block Parameterizations

We compare two types of parameterization commonly used in tomography: one with continuous spherical harmonic functions and the other with discretized blocks. A comparison of parameterizations requires that both have the same minimum resolving scales. We therefore design models having 4.5°×4.5° and 9°×9° nearly equal-area blocks, which are to be compared with spherical harmonic models L=40 and 20, respectively. Among 2038 B=4.5° blocks, only two in the Southern Hemisphere are missed by either phase, and all of the 508 B=9° blocks are sampled. Two-dimensional inversion for D" shear velocity heterogeneity using spherical harmonics as basis functions is addressed in paper 1. Here we weight the data kernel by the data covariance matrix, which is diagonal but variable, in contrast to the constant data variance assumed in paper 1. As in paper 1, we adopt damped least squares as the main regularization scheme in which the reciprocal of the singular value, $1/\lambda$, is weighted as $\lambda/(\lambda^2+\Lambda^2\varepsilon^2)$, where Λ is the maximum singular value [*Aki and Richards*, 1980]. Trade-off curves between the mean data misfit and the mean model variance are then constructed with different levels of damping ε^2 (Figure 4a). A fair comparison requires that models have the same model error; we therefore conduct our comparison for values in the boxed region in Figure 4a. The box is arbi-

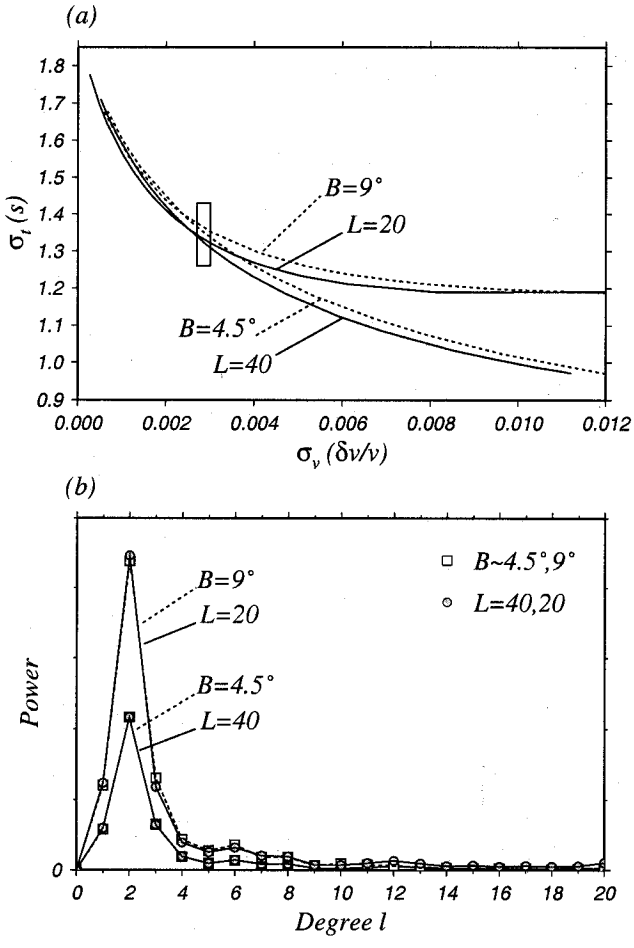

Figure 4. (a) Trade-off curves between data misfit and mean model variance (as in paper 1) for inversions with spherical harmonic functions (solid) and blocks (dotted). Each set, B=9°/L=20 and B=4.5°/L=40, has the same basic resolving length. The models within the box have the same model variance and are compared in Figure 4b for power distribution and in Plate 1 for heterogeneity image. (b) The power spectra for the models extracted from the box in Figure 4a. Comparable resolving length parameterizations yield nearly identical power spectra, but low-degree structure is affected by the choice of resolving length.

trarily chosen for models which have the same overall error but have variable misfits to data.

The L=40 spherical harmonic model and the B=4.5° block model are highly similar in both power spectrum (Figure 4b) and distribution of heterogeneity (Plate 1). This similarity also exists for the L=20 and B=9° models, as well as for other positions on the trade-off curve in Figure 4a, except for those with large model variance. The relatively increased power for the L=20/B=9° models (Figure 4b) results partly from the fact that damping for L=20 is slightly smaller than for L=40 and mainly from forcing energy from degrees $21 < L < 40$ into the first 20 degrees. While the energy (not the power) is conserved against L, strong smoothing causes it to be skewed to low degrees and the redistribution of energy would not be as obvious. The L=40/B=4.5° models are a better choice to avoid aliasing or energy leakage, as we discuss below. For each pair of models in Figure 4, the block and spherical harmonic models depart significantly in regions where large data gaps are present. It is well documented that continuous basis functions can spread undesired energy to data gaps (or where constraints are weak) in order to enhance total data fitting [e.g., *Pulliam and Stark*, 1993]. In contrast, block or pixel-like parameterizations do not have global aliasing effects; instead, unsampled blocks have zero singular values for which local model properties will be determined by the regularization alone. Only if coverage gaps are smaller than the scale length of interest do the two parameterizations yield similar results.

Although the scale of 4.5° is too small to be within the assumed validity range of the ray theory kernel, these results serve to demonstrate that spherical harmonic and block inversions retrieve consistent model power spectra. The differences that exist in published models arise from differences in data coverage and/or sampling, the shortest resolving scales contained in the model, or the regularization scheme. For example, by comparing models L=40 and B=9° in Figure 4b, one might erroneously conclude that spherical harmonics retrieve long-wavelength heterogeneity more effectively than blocks. Recognizing no intrinsic difference in resolvability between the two, we employ spherical harmonics in relatively low resolution modeling for convenience and smoothness.

3.3. Truncation Effects

Bias due to finite parameterization has long been a concern for inversions [e.g., *Parker*, 1977]. *Trampert and Snieder* [1996] showed that a truncated expansion introduces artifacts to the solution model for unevenly distributed data, and they provide a remedy to suppress leakage from the neglected basis functions. While we have reduced redundant sampling as described above, this effect remains unpredictable because the real Earth structure is unknown. *Trampert and Snieder* [1996] demonstrate that components near the truncation point are the most affected. Following the approach in paper 1, we employ a high-degree (L) expansion for inversion and a low-degree (L_s) spherical harmonic synthesis to obtain a robust solution. This is a natural antileakage scheme in that the model is least biased if $L_s << L$.

To illustrate the leakage due to truncation and to demonstrate our strategy, we perform a checkerboard experiment in which a degree (l) 12, order (m) 6 checkerboard pattern of heterogeneity, with spatially alternating positive-negative anomalies (in %) having a root-mean-square (rms) magnitude of 1%, is assumed as an input test model. We generate artificial "data" with the ray geometry in Figure 2a for this input structure by randomly adding to the ith model prediction zero-mean Gaussian noise with the standard deviation σ_i equal to the assigned error of the ith observation. There are infinite sets of random errors, but they cause little difference in what we conclude below. We perform inversions for the artificial data, with L=12 and 11, employing a damping that reproduces the same 1% rms magnitude in the L=12 inversion. For both models we synthesize the components to L_s (spherical harmonic summation from 0 to L_s), for L_s=1 to L, and show the magnitude spectra versus L_s in Figure 5a. If the model fully parameterizes the "real" earth, as in this case L=12, the true structure at L=12 dominates the cumulative errors below L. However, if the parameterization is truncated before L=12, the energy of the real structure will not be completely truncated but "leaks" into the model [*Trampert and Snieder*, 1996]. The

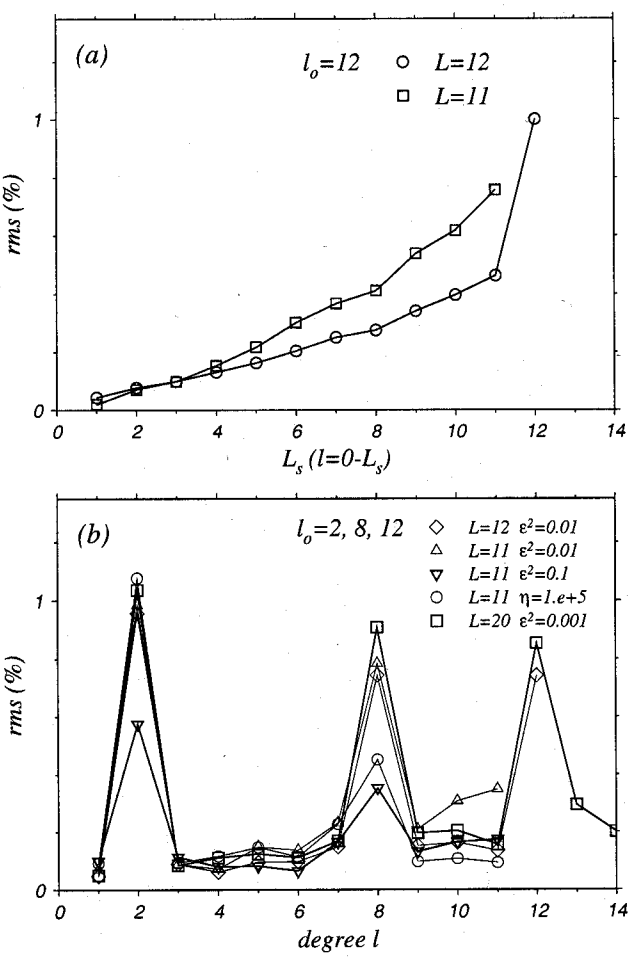

Figure 5. Model rms spectra for synthetic recovery tests. (a) The rms versus L_s for a degree 12 input structure. An L=12 inversion retrieves this signal well, while an L=11 inversion retrieves a significant, but erroneous component at degree 11, caused by leakage of the truncated signal. The leakage effect decays with decreasing L_s, a fact that provides the basis for our antileakage scheme. (b) The rms versus degree l for an input structure with components at l=2, 8, 12. High damping (ε^2=0.1 for damped least squares, inverse triangle; and η=10^5 for minimum curvatures, circles) suppresses leakage at degree 11 but overdamp all other signals. Inversions with L>=12 recover all three components with low damping. Using L=11 causes significant leakage at degrees 11 and 10.

leakage is most serious from those degrees that have been truncated, i.e., when L=11 inversion fails to adequately parameterize the degree 12 signals. Figure 5a implies that the L=11 model has a significant degree 11 feature that is prominent enough to be misinterpreted as a real structure. However, the artifacts due to truncation vanish toward lower degrees, so that, without knowing the true spectrum of the Earth, models synthesized to L_s remain little biased as long as L_s<<L. Note the errors from the two inversions converge and become negligible for L_s<=3.

To further examine how damping affects the leakage, we experiment with a multi-component input model having variations at (1) l=12, m=6; (2) l=8, m=4; and (3) l=2, m=1, each with an rms magnitude of 1%. Artificial "data" are generated the same way as above. We perform inversions this time on the artificial data with L=20, 12, and 11, employing high damping (ε^2=0.1) and low damping (ε^2=0.01 and 0.001). Smoothing with weighted minimum curvature is sometimes employed in tomography [e.g., *Li and Romanowicz*, 1996] and is also conducted here for comparison. The weighted minimum curvature involves a smoothing proportional to $1/[l(l+1)]^2$, with a weighting factor η. For the relatively low-degree models, which are already "smoothed" by truncation in synthesis, we do not prefer imposing any predefined decay rate but rather taper each component according to the singular values of the data kernel. Figure 5b shows the recovered heterogeneity spectra in this experiment versus each degree (l) rather than versus L_s as in Figure 5a. Similar to the above experiment, errors in the data are mapped to all degrees, but the signals at degrees 12, 8, and 2 still dominate the noise at other degrees for L>=12. Inversion with L=11 and low damping (ε^2=0.01) leaves a considerable artificial feature at degree 11, which results from leakage of the unparameterized degree 12 signal (compare to the L=12, ε^2=0.01 curve). For uniform damping severe enough to suppress the leakage (e.g., ε^2=0.1), real signals are substantially suppressed too, as manifested in the reduced magnitudes at L=2 and 8. A similar effect occurs for inversion with minimum curvature; in order to suppress the leakage at degree 11 (η =10^5) the signal at degree 8 is overdamped. The decay of leakage with l can also be observed from degree 11 to 9 in Figure 5b.

Although these two experiments provide little quantitative insight (as we lack knowledge of Earth's true spectrum), we infer that the L>> L_s strategy naturally filters out much of the truncation contamination due to possible model expansion failure to account for real, short scale D" structure. This approach also avoids undesirable overdamping. Fortunately, the 2-D problem that we address enables us to use a much higher L than would be the case for 3-D mantle tomography due to limits on computation, although we do not know how high the L should go. We perform inversion truncation with L=40, which, as shown below, is sufficiently higher than the L_s that we adopt for a robust model.

3.4. Degree 12 Model, 1% rms Heterogeneity

The proper choice of L_s depends on the global resolution offered by the geometry and quality of data and on how heavily the high-degree components are damped. To determine L_s quantitatively, we perform a series of synthetic recovery tests for checkerboard structures with degree and order l=l_o, m=l_o/2, for various l_o, each of which has a 1% rms magnitude. Inversion for each structure is performed with L= l_o to exclude leakage effects. The damping for each inversion is chosen so that the output model reproduces the 1% rms magnitude, which usually leads to modest to low damping (ε^2<0.1). To measure how well the inversion recovers the structure, the rms difference between input and output is calculated. Each rms difference value in Figure 6 is the average of more than five calculations corresponding to different sets of random errors used in generating the artificial data, which we believe gives results relatively free of bias. The model recovery is considered unsuccessful if this difference exceeds half (0.5) of the magnitude of the input model. This criterion is relatively strict because it is equivalent to requiring roughly 75% variance reduction for a model to be acceptable. If the synthetic data are calculated without random errors, the ray geometry alone (Figures 2a) renders a recovery successful up to L=40. With the presence of error as defined here, a degree 12 variation is probably the finest we can retrieve reliably in the global average sense (rms difference of 0.48). Degree 12 is thus the mean global resolution and L_s is set at 12 in this study.

For a given degree of final model synthesis (L_s), L is chosen such that L >> L_s. A parameter left undetermined is the damping factor ε^2, which is explicitly involved in constructing the model. The choice of ε^2 has an effect of smoothing and dominates the magnitude of heterogeneity strength of the model, and the two effects trade off with each other. We choose the damping to achieve a desired level of heterogeneity perturbation for given L and L_s. Although this choice is arbitrary, it is directly related to the nature of the output model and therefore we can choose parameters appropriate for com-

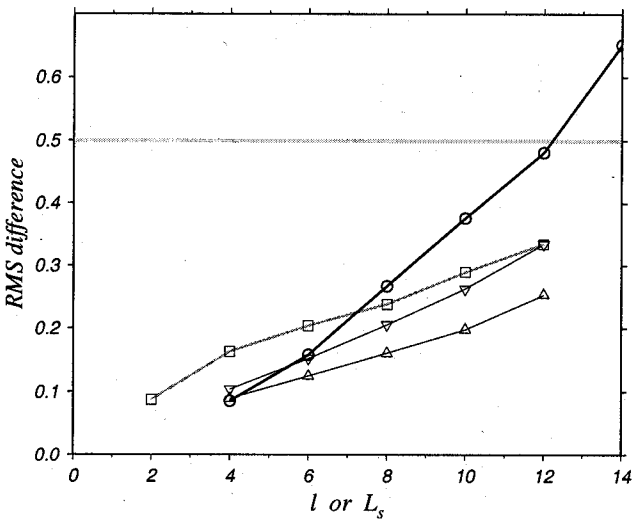

Figure 6. The rms difference between models as a function of each degree (l) or the maximum degree to which the models are synthesized (L_s). Circles and thick line are rms difference between input and output model versus l for the recovery experiment to quantify the global mean resolution. The highest rms that is <0.5, the threshold of successful recovery (horizontal shaded bar), occurs at degree 12, which is therefore considered the maximum global mean resolution. Triangles are rms difference between L=20 and L=12 models as synthesized to various degrees L_s. Inverse triangles are the same as for the triangle but between L=40 and L=12. Squares and thick shaded lines are rms difference, as a function of L_s, between models with non-D" paths corrected using SKS12WM13 and SAW12D.

parison to other models in the literature. The rms magnitude of shear velocity perturbation in our model is set at 1% in order to compare with the following models. D" heterogeneity maps were extracted from SKS12WM13 (L=12) [*Dziewonski et al.*, 1997], SAW12D (L=12) [*Li and Romanowicz*, 1996], and S16B30 (L=16) [*Masters et al.*, 1996] by averaging 6 layers with 50 km intervals in the lowermost 250 km, and from the bottom layer of *Grand et al.* [1997] (~2.5°×2.5° blocks, 240 km thick). The rms magnitudes of these models are ~1%, 0.92%, 0.90%, and ~1%, respectively. We therefore construct degree 12 models by performing inversions with L=40, damped with values of ε^2 such that when synthesized to L_s=12, the rms magnitude is 1% (Plate 2a). The map is demeaned by -0.28%, or -0.021 km/s, a close value to that obtained in paper 1 (-0.022 km/s). We also extract a degree 10 model (L_s=10) from the degree 12 model for comparison (Plate 2b). This model has an rms magnitude of 0.95%.

To test how an L_s=12 model depends on L, inversions with L=20 and L=12 are also performed and the results are compared with the preferred L=40 inversion. The rms differences between resulting models are considered quite modest: 0.33 between L=40 and L=12 and 0.26 between L=40 and L=20 (Figure 6). The differences likely arise from truncation effects of lower expansion, and the differences decrease and converge for smaller values of L_s. The D" maps of SKS12WM13 and SAW12D are displayed in Plate 3. The rms differences between our preferred model (L=40, L_s=12) and SKS12WM13 and SAW12D are significant (>0.9%), and they decay slowly with lower L_s (Figure 6). The difference between SKS12WM13 and SAW12D is also as high as 0.8% to their full synthesis. As described in section 1, different data sets, regularization schemes, and formulations result in such significantly different D" images.

3.5. A Hybrid Model

The degree 12 model presented above (Plate 2a) is a global aspherical image, which somewhat overparameterizes regions that are less constrained and sacrifices high resolution in densely sampled regions. The resolution of any tomographic model is spatially variable. To illustrate the spatial recovery capability for our data set, we calculate the local rms difference at each point over a suite (more than 5) of degree 12 and degree 10 synthetic input data sets with random errors, yielding the rms differences in Figure 7. The rms's of the two maps are virtually the same values as those plotted in Figure 6, i.e., 0.48 and 0.38, respectively, and the same threshold for a successful recovery, i.e., rms < 0.5, holds for the map too. These maps indicate nonuniform resolvability of the model associated with both density and quality of samples. In general, the recovery in the Northern Hemisphere is better than in the Southern Hemisphere and that in the Pacific is better than both sides of the Atlantic. Note that some regions with rms >

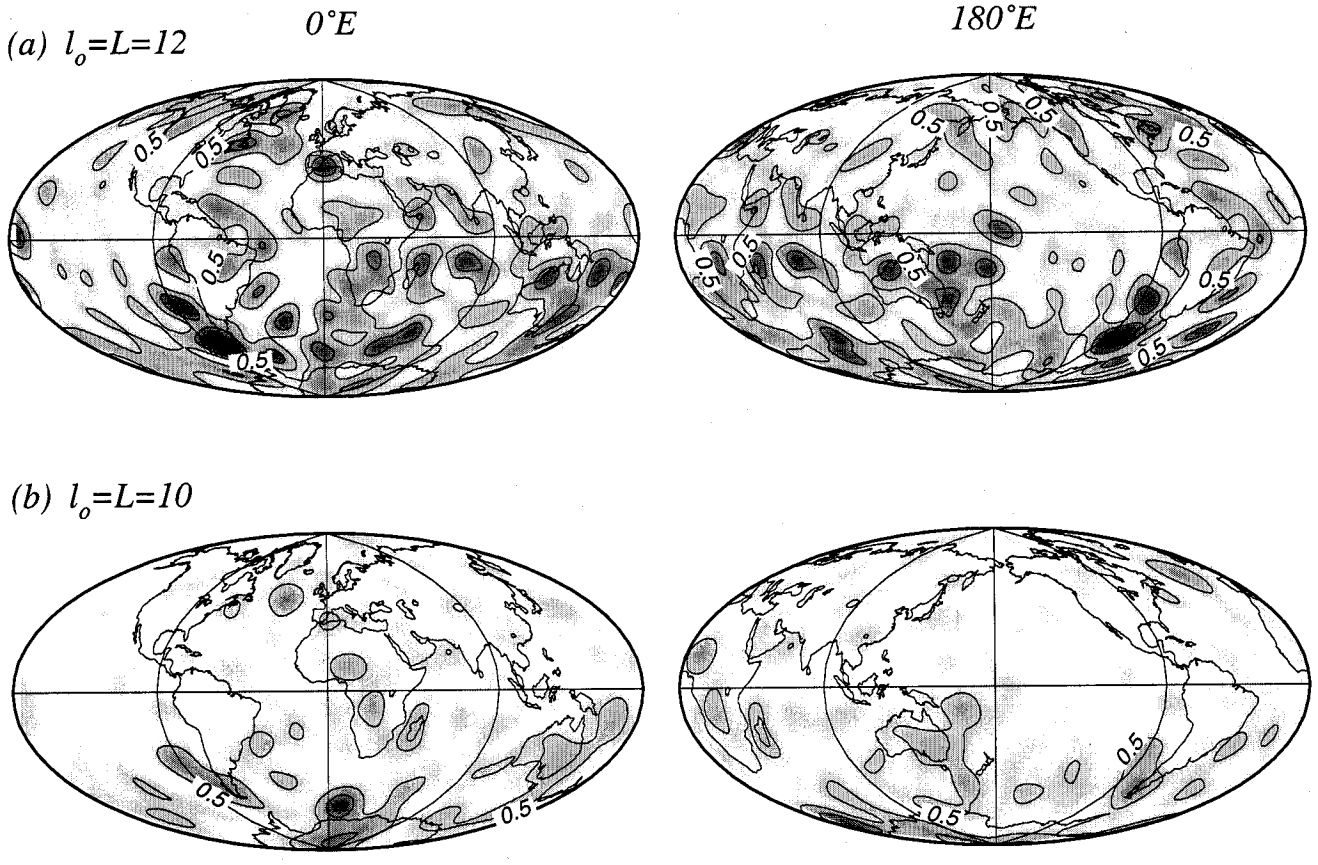

Figure 7. Maps of rms difference (see text) in the recovery test for input structures of (a) l=12 and (b) l=10. Regions with rms<0.5 are better resolved than the global average degree 12 set in this study. The rms of each map is equivalent to that plotted in Figure 6 (circles). In Figure 7a, regions below the Pacific, Eurasia, and North America can be imaged with resolution beyond degree 12. Note that for some regions, recovery is acceptable (rms<0.5) only for degree 10 variation. These maps can be compared with those in Plate 2 in order to identify well-resolved features.

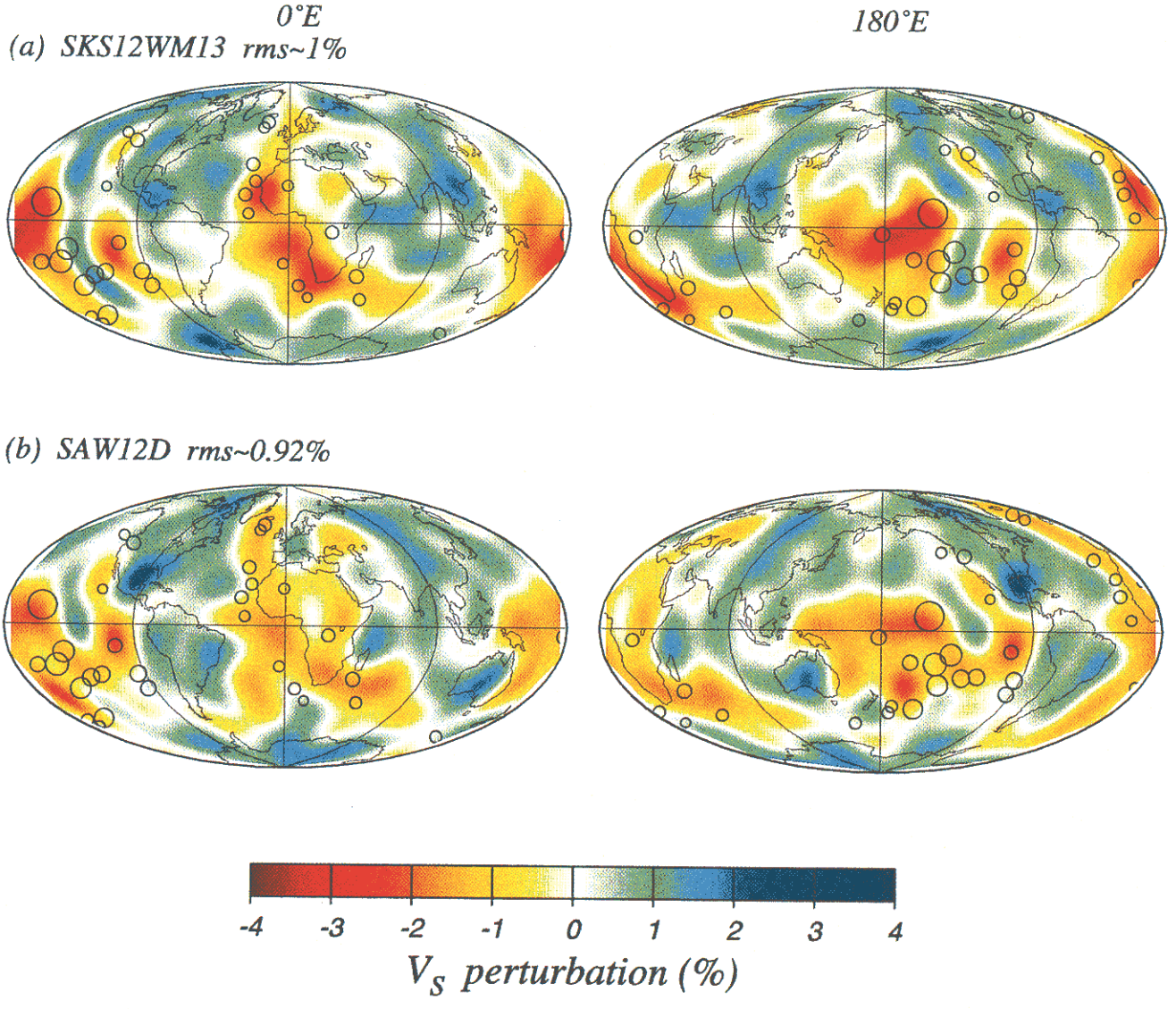

Plate 3. The D" models extracted from (a) SKS12WM13 and (b) SAW12D. (left) Maps centered at 0°E; (right) maps centered at 180°E. The differences between our degree 12 model in Plate 2a and these two are significant. Note that the proposed plume roots are less clustered with patches of negative anomalies beneath the Atlantic and visually less correlated with the negative anomalies in these two models than in our preferred model in Plate 2a.

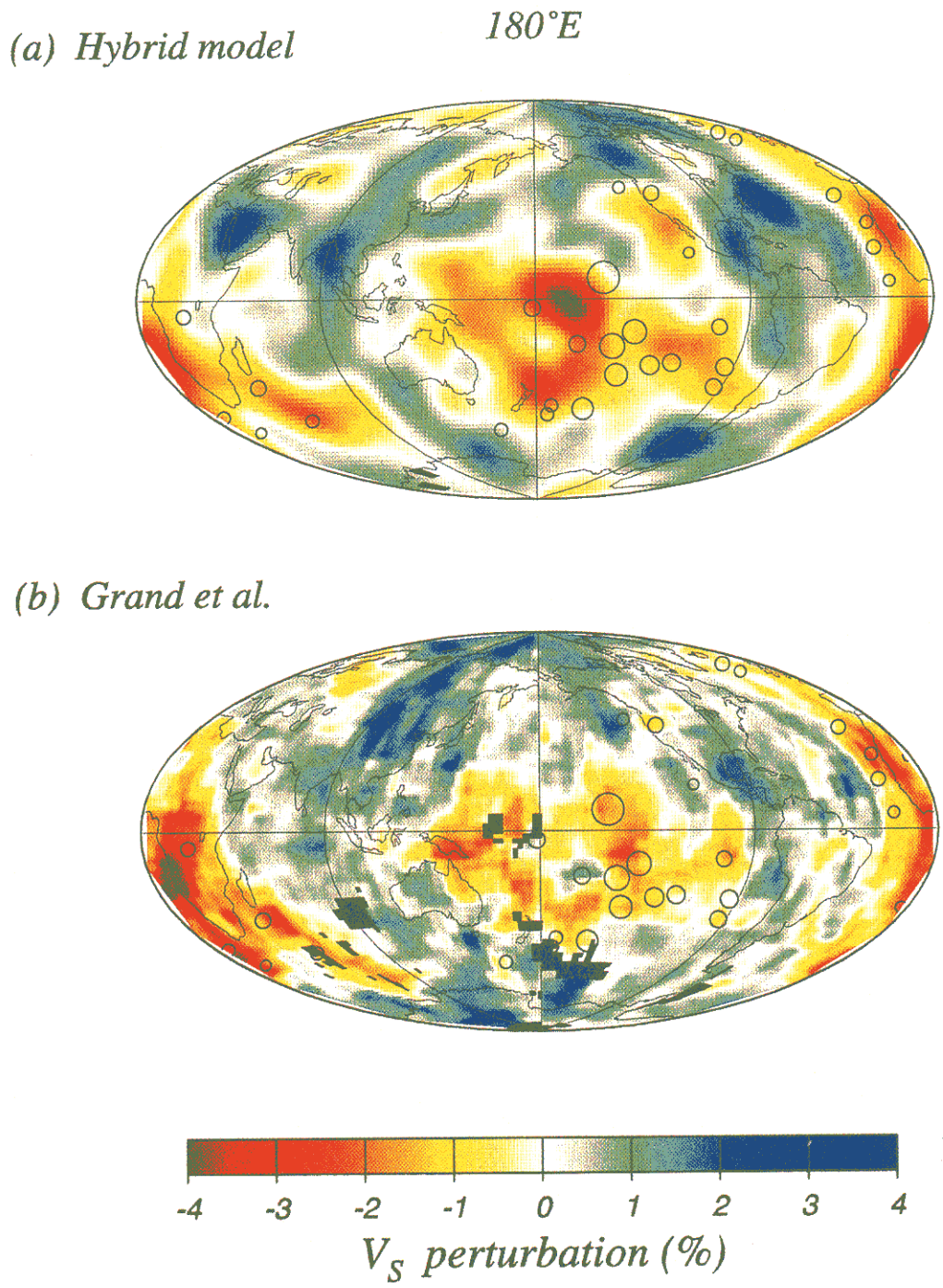

Plate 4. (a) The hybrid model of this study and (b) the block model of *Grand et al.* [1997]. The rms map in Figure 7a indicates that the hybrid model is valid in the central Pacific, Eurasia, and North America. The two images deviate substantially in the Pacific where our hybrid model is probably best resolved.

0.5 for degree 12 are resolved satisfactorily for degree 10, e.g., southern Africa and the Indian Ocean.

The failure to retrieve accurate degree 12 structure in regions like south Africa and the southeastern and southwestern Pacific is a consequence of both few samples and low quality of measurements (Figure 2e); most of these regions can only be accessed through using quality C shallow events from relatively inactive zones. Similar calculations demonstrate that in the Pacific, Eurasia, and North America where rms < 0.5 for degree 12 (Figure 7a), the resolution can be pushed to degree 14-16, under the same criterion for success. These regions warrant further investigation for smaller-scale details. To accomplish this, we employ a second-stage block parameterization, since blocks are more flexible for imaging regional, small-scale perturbations. The *S-SKS* times predicted for the degree 12 aspherical reference structure in Plate 2a are removed from actual data, then the corrected residuals are inverted using a 4.5°×4.5° block parameterization, with a regularization of minimum gradient to suppress overparameterization (explained below). The results of the second step inversion are added back to our degree 12 reference structure, and the final hybrid model is shown in Plate 4a in a style with block geometry removed. Regions with sufficient resolution to warrant such a two-step procedure are identified by using the rms map in Figure 7a. Regions below the Pacific, Eurasia, and North America are suitable for the two-step hybrid method.

The reason for adopting this hybrid approach, and not solely a B=4.5° block inversion, is twofold. First, it is convenient and natural to explore model details with the large-scale background removed. Second, a pure B=4.5° model, which is equivalent to degree 40 in maximum resolution, necessarily requires extreme regularization to penalize the small wavelength components for a robust solution, if simple form of regularization is used. As shown in Figure 5, overdamping and smoothing suppresses all scales of signals and distorts the spectrum of the model, unless a combination of regularization forms is employed. The minimum leakage, degree 12 model retrieves the low to intermediate-degree signals without unnecessarily overdamping rms magnitudes.

The 4.5° block parameterizations must be smoothed to represent the actual maximum regional resolution of about degree 16 (or 700 km), which is roughly of the same order of the 600 km limit set for the validity of the ray theory kernel. To find the characteristic scale of a model, we establish the velocity difference, Δv, between any two points of the model as a function of the distance range within which the calculation is made. The rms deviation of Δv, which is a measure of the scatter of velocity, usually increases with distance range until it approaches the global characteristic value, at the distance corresponding to the scale of the heterogeneity. To smooth the step 2 model to a characteristic scale of at least 700 km, the regularization of minimum gradient has to be at least 10 times more strongly weighted than the least squares fitting in the inversion. Different combinations of minimum gradient and damped least squares may yield similar results. Since the first step of the hybrid method preserves the low-degree power, power loss by strong smoothing or damping is not a major concern in step 2. The block parameterization also entails leakage problems, but the strong smoothing effectively suppresses possible leakage. The signal power resolved in the second step is roughly 2-3 orders of magnitude lower than the long-wavelength inversion of step 1. The rms magnitude of the hybrid model is 1.1%, which is slightly increased from the degree 12 reference (of 1%) but with greater detail revealed in some regions. The Pacific side of the hybrid model is compared with the map of *Grand et al.* [1997] which use strongly damped, variable but small size blocks for inversion, with some subsequent smoothing (Plate 4b). The differences between these two maps are prominent in the central Pacific and Eurasia where our hybrid model is particularly warranted.

4. Discussion

The success of using S_{diff}-*SKS* differential times rests upon the enhanced sensitivity to D" by crossing ray paths in this layer and randomizing ray paths elsewhere in the mantle. The maps in Figure 2d demonstrate this point. The current geometry and quality of data preclude global resolution higher than degree 12, which is similar in resolution to other recent global studies. If the errors are estimated correctly, better ray path coverage is needed to extend the mean resolution above this value. In future modeling efforts, filling data coverage gaps, guided by the regions of poor recovery in Figure 7a, is as important as advancing the inversion technique. Given that the poorly resolved regions tend to have limited sampling by numerous, high-quality data (Figures 2d and 2e), filling gaps will be hard to achieve in the near future. In this circumstance, analyzing the variation in resolvability is critical for a tomographic inversion problem. In this paper, we pursued a model that has the maximum viable mean resolution and which, in theory, contains the fewest inversion artifacts. We then extended the resolution to shorter wavelengths in regions that are considered viable for further parameterization. The rms difference maps in Figure 7 help to identify the geographical distribution where the solution is robust.

Variable polarization anisotropy has been documented for S_{diff} in some extensively explored regions (see review of *Lay et al.* [1998]). Beneath Alaska, the northern Pacific, and the central America, velocity of the *SH* wave is faster than that of the *SV* wave by 1-2%, whereas the sense of anisotropy in the central Pacific is not uniform. *Garnero and Lay* [1998] indicate that a 2% transverse isotropy in a 300-km-thick D" layer generates <1 s anomaly in S_{diff}-*SKS* at distances 100° or greater. This is because of the short path and the subhorizontal polarization of the *SKS* phase in the lowermost mantle. We have conducted experiments to resolve *SV* and *SH* velocities simultaneously. As expected, the resulting *SV* model is not required statistically to explain the data, even for very low degree expansions. Thus the models presented in this study are most appropriately characterizing *SH* shear velocity. In addition to attempting retrieval of *SV/SH* anisotropy we also experimented with resolving a multilayer velocity model, as well as a model with variable linear vertical velocity gradient. However, the lack of vertical resolvability of the current data set precluded progress in this direction. Recent works by *Ritsema et al.* [1997] and *Kuo* [1999] demonstrate that the amplitude of S_{diff} may help constrain the gradient in D".

Yanagisawa and Hamano [1999] suggest that the skewness of the velocity perturbation could be an indicator of relative importance between different sources that drive thermal convection. The skewness is a measure of asymmetry of a random variable about its mean; it is zero for symmetrical distribution and nonzero for skewed distribution. They found in their simple simulation that the skewness of the temperature

field is positive (skewed to positive) averaged over the lower mantle if basal heating dominates internal heating as driving forces and negative if the former is turned off. The skewness as a function of depth for various global tomographic models supports the notion of basal heating, assuming that all velocity perturbations are thermal in origin. Our degree 12, 1% rms model is characterized by a skewness of -0.5, which is qualitatively consistent with other models. However, if chemical heterogeneity is present and perturbs shear velocity and drives flow in the lower mantle, the above skewness implications may, of course, not be applicable.

On the basis of forward modeling of a lower mantle triplication arrival between *S* and *ScS* phases, D" shear velocity models with a discontinuity of 2-3% have been proposed for various regions [*Lay and Helmberger*, 1983] (see review by *Wysession et al.*, [1998]). For a tomographic model derived using a single-layer structure (explicitly ignoring any vertical velocity gradients in the lowermost 250 km of the mantle), the aspherical structure is manifested as a ~2-3% anomaly in regions of large positive velocity perturbation (e.g., see paper 1). Regions where evidence for a D" shear velocity discontinuity exists, such as beneath the Caribbean [e.g., *Lay and Helmberger*, 1983; *Kendall and Nangini*, 1996; *Ding and Helmberger*, 1997], Alaska [*Young and Lay*, 1990; *Matzel et al.*, 1996], India [*Young and Lay*, 1987], Eurasia [e.g., *Gaherty and Lay*, 1992; *Weber*, 1993], and off the southern tip of South America [*Olivieri et al.*, 1997], do indeed correspond to high-velocity regions in our final model shown in Plate 4a (similar to the finding of paper 1 for degree 6). Regions where there is little evidence for a D" shear velocity discontinuity or evidence for intermittent layering, such as beneath the central Pacific [*Garnero et al.*, 1993], the southwest Pacific, and Australia [*Kendall and Shearer*, 1995], show negative or variable velocity structure in our final model. This indicates that the integrated volumetric heterogeneity of the D" layer is related to the characteristic of any discontinuity in the radial structure.

We have corrected the non-D" portions of our wave paths for mantle asphericity predicted by various mantle models. These models, however, may contain signal from the D" region smeared to above D" as we have defined it here (250 km above the CMB). Some recent global mantle model inversions have incorporated *S-SKS* residuals to place better constraint on the deepest mantle structure. SKS12WM13 utilizes more than 6000 *S-SKS* times, and its non-D" part provides a 22% variance reduction in our data set. SAW12D, without the radial resolution provided by *S-SKS* times, provides 12% variance reduction, despite using a more sophisticated formalism. The ability of these models to predict trends seen in the data is a strong function of geographic locale. For instance, SKS12WM13 explains 53% of the variance for our 254 Pacific paths from Tonga and South American events. This is much higher than the global mean, implying that Pacific data dominate the determination of lower mantle asphericity in SKS12WM13. For the same Pacific paths, SAW12D yields 30% variance reduction. The long-wavelength components of both models show less disparate abilities to explain our *S-SKS* residuals. The first six degrees (L_s=6) of SKS12WM13 and SAW12D (non-D") contribute 17% ad 13% variance reductions, respectively, for all paths and 30% and 25% for the Pacific paths.

Comparison of the asphericity corrections along the 1500 paths calculated from various models with those from SKS12WM13 is shown in Figure 8. The scatter between SKS12WM13 and SAW12D is much more pronounced than for other pairs (Figure 8a), though the correlation is high if both models are synthesized only to L_s=6 (Figure 8b), indicating that the two models deviate substantially at short wavelengths. The asphericity corrections calculated from *Grand et al.* [1997] and S16B30 are better correlated with those from SKS12WM13 (Figures 8c and 8d). Note that the model of *Grand et al.* [1997] is heavily damped, yielding a small range of corrections (Figure 8c). We also invert for a rms=1% D" model using data corrected by SAW12D, and we compare it with our preferred model. The rms difference between the two D" models is 0.33, which is considered to be modest (Figure 6). If our data coverage were everywhere azimuthally well sampled, the differences in asphericity corrections between models would behave geometrically as "random noise" and would be suppressed. The fact that our degree 12 model is still slightly sensitive to this correction implies that at least some regions are dominated by less diversified raypaths, along which differences in corrections (or "errors") can accumulate and be projected into the D" model. However, through our effort to enhance azimuthal diversity of the crossing raypaths, the overall mantle dependence is not very significant. Since the scatter in Figures 8c and 8d is less than in Figure 8a, corrections using the model of *Grand et al.* [1997] or S16B30 are expected to cause even less deviation in final models.

For quite some time, it has been supposed that hotspots may be due to plumes that originate at the CMB [e.g., *Stacey and Loper*, 1983; *Williams et al.*, 1998]. The identification of plumes beneath a few hotspots at lithospheric or asthenospheric depths has been documented [e.g., *Nataf and VanDecar*, 1993; *VanDecar et al.*, 1995], and recently evidence for a plume across the whole mantle beneath Iceland has been reported [*Bijwaard and Spakman*, 1999]. In forward modeling efforts, *Helmberger et al.* [1998] and *Russell et al.* [1998] infer the presence of the root of the Icelandic and Hawaiian plumes, respectively, at the base of the mantle. If whole mantle plumes indeed originate from the D" boundary layer, then D" shear velocity maps may have characteristic patterns that relate to hotspot distribution. In paper 1, positions of major hotspots were shown to colocate favorably with depressed D" shear velocities. We revisit this issue here with our improved model and utilize the distribution of the roots of plumes from *Steinberger* [2000], who calculated the effects of mantle plume deflection due to large-scale mantle flow. The buoyancy-driven flow was estimated by converting shear velocity perturbations of various tomographic models to temperature anomalies. The estimated locations of plume roots for 40 major hotspots (mass flux > 0.3 in the *Sleep* [1990] catalogue) calculated for SKS12WM13 are shown with all the models considered in this study (Plates 2, 3, and 4). Nearly all of the estimated root locations fall in low-velocity regions, close to transitions from average to regional minima. We test the hypothesis that plumes randomly occur in fast or slow regions with equal probability. The binomial test rejects this at the 99% confidence level in favor of the hypothesis that plumes originate from lowermost mantle slow regions. Furthermore, the colocation of paleoslab locations [from *Wen and Anderson*, 1995] and D" high velocities is statistically significant at the 95% confidence level.

The correlation between D" velocities and proposed hotspot root locations is stronger than between D" velocities and

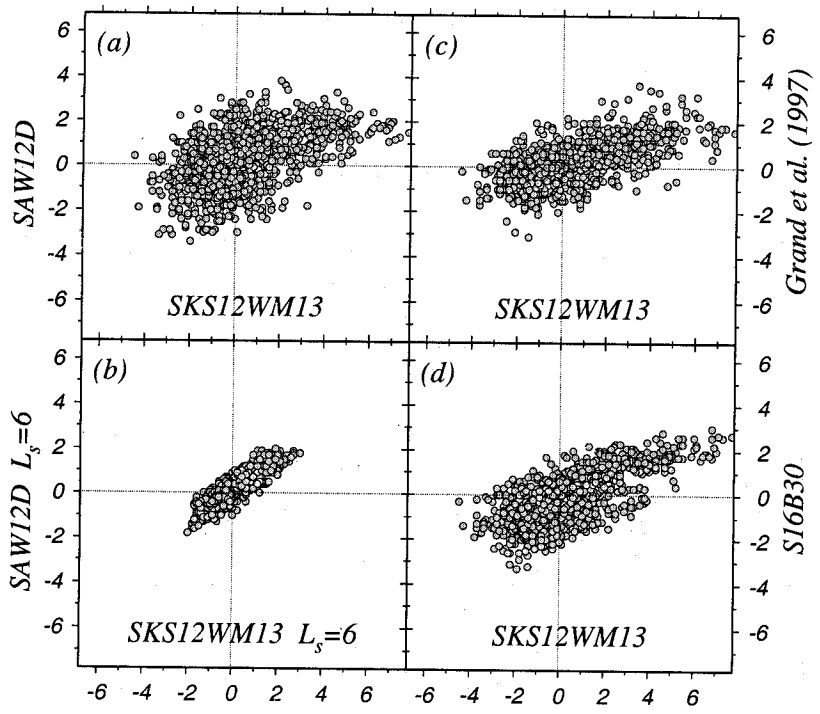

Figure 8. Mantle asphericity corrections for non-D" portions of the *S-SKS* differential times for our ~1500 data calculated for various models, plotted against those for SKS12 WM13 to show their correlations. (a) SAW12D; (b) SKS12WM13, L_s=6; (c) *Grand et al.* [1997]; and (d) S16B30. The most scattered distribution is in Figure 8a, which causes 0.33% rms difference in D" models when inversions are performed using the two reference models for corrections. Corrections using other models are expected to yield more similar models than in that case.

hotspot surface locations. This is not surprising, since plumes, in *Steinberger*'s [2000] calculation, are swept along the CMB toward buoyant sources in the driving model SKS12WM13, which does not greatly differ from our model at very long wavelength. The distribution of plumes on the hybrid model show similar geographic characteristics with respect to the low-velocity zones (Plate 4). Hotspot roots are clustered more systematically around the smaller-scale low-velocity features in our model than any other structures. For example, the Atlantic-Indian low-velocity corridor has average perturbations of ~-1%; three strongly negative patches (δV_s < -2%) along this trend seem to be closely affiliated with proposed plume root locations that belong to three separate groups of hotspots. This affiliation is considerably stronger in our map than in SKS12WM13 or SAW12D. Given that the rms map in Figure 7a implies insufficient resolvability in southern Africa for degree 12 to firmly establish this pattern; we turn to the degree 10 model (Plate 2b), in which the recovery is acceptable in the same region (Figure 7b). The degree 10 model shows slightly reduced variation in the vicinity of Africa, but the relationship with the hotspot roots still exists. It is possible that the instability of the thermal boundary layer may be related to the small- to intermediate-scale low-velocity anomalies (e.g., ~1000-3000 km). High-resolution studies [e.g., *Bréger and Romanowicz*, 1998; *Ritsema et al.*, 1999] will help to place tighter constraints on the shapes and strengths of these features.

Smaller-scale fine structure, such as thin localized ultra-low-velocity zones (ULVZ) [e.g., *Garnero and Helmberger*, 1996; *Wen and Helmberger*, 1998; *Helmberger et al.*, 1998; *Vidale and Hedlin*, 1998] are unresolvable with the *S-SKS* data set. However, if such structures are caused by partial melt [*Williams and Garnero*, 1996], then one would certainly expect an intimate relationship between ULVZ distribution and reduced lower mantle velocities and thus locations of plume genesis [*Williams et al.*, 1998; *Garnero et al.*, 1998]. The low D" velocities beneath the Pacific host nearly all of the proposed plume roots for that region. This correlation, as with the Atlantic and Indian Oceans, is stronger in our model than for others we tested. Hawaii, the largest hotspot on Earth's surface in terms of estimated mass flux [*Sleep*, 1990], is located above the margin of the lowest-velocity anomaly (<-3%) in our model. As in other models we considered, Hawaii's surface location and proposed plume root location are not directly over the minimum of the low-velocity anomaly. The shape of the lowest-velocity structure beneath the central Pacific is resolved differently by our model (Plate 2a and 4a) and others (Plates 3 and 4b). The rms maps in Figure 7 reveal that this feature may be well resolved only up to degree 10, for which the feature is less isolated and prominent than it appears in the degree 12 model, although the general geometry remains unchanged. The structure to the southwest of this low anomaly is, however, less reliable in both degrees 12 and 10, and therefore the western stretches of the Pacific low-velocity corridor are generally unresolved.

As with the case of the Hawaiian plume, it is interesting to note that hardly any of the proposed plume root locations of *Steinberger* [2000] directly coincide with the maxima in low-velocity features and instead appear to surround them. The velocity depressions are most commonly presumed to be thermal anomalies related to whole mantle flow. Hence it is perplexing that we see no evidence for large flux plumes

originating from the central zone of a large thermal upwelling. This may be a dynamical result because upwelling attracts lateral flow that extends only a limited distance away from the center of the buoyant source, making a plume finally "rooted" on the boundary of a negative anomaly. On the other hand, future work must consider scenarios involving a possible chemical component to D" heterogeneity and refined models of possible plume root locations, as well as better resolution models of D" heterogeneity from improved wave path coverage.

SKKS-SKS times have been used previously to map D" lateral variations [e.g., *Sylvander and Souriau*, 1996]. We tested this approach by adding a small data set of ~300 *SKKS-SKS* times, which were dominated by paths across the Pacific. For model SKS12WM13 the variance reduction for the *SKKS-SKS* data used in constructing that model is only 25%, much lower than the 75% achieved for the corresponding *S-SKS* data. This is because *SKKS* and *SKS* separate spatially far less than *S* and *SKS* and a degree 12 model lacks the lateral resolution to account for their relative behavior. The parameter *X* in Figure 1 for *SKKS* and *SKS* is nearly half of that for *S* and *SKS*, and that is shorter than the degree 12 resolution in the lower mantle. This is the disadvantage of using these times. However, the proximity of the paths for *SKKS* and *SKS* reduces the effect of mantle heterogeneity better than for *S* and *SKS*. Another slight advantage of *SKKS-SKS* times is that they are largely free from error associated with azimuthal anisotropy in the upper mantle because both phases are on the radial component. The reasons we ultimately did not incorporate *SKKS-SKS* residuals in our model are twofold. First, the degree 12 model is dominated by *S-SKS* data, and the damping suitable for *S-SKS* is found to be too high to explain *SKKS-SKS*, leaving large errors that are always associated primarily with the *SKKS-SKS* samples. Second, despite the enhanced resolution provided by the step 2, block parameterization, our *SKKS-SKS* data set has limited coverage and path crossings. As emphasized above, without crossings, errors from inadequate corrections for passage through the mantle will be translated into false velocity perturbations in D". Block inversion will be most sensitive to this disadvantage if most of the blocks are undersampled azimuthally from a small data set.

5. Conclusions

Approximately 1500 *S-SKS* residuals, selected to maintain as uniform coverage over the globe as possible, provide an opportunity to illuminate the inversion problems and models of D" structure (or the lowermost 250 km of the mantle). Ray theory approximation along with first-swing alignment for travel time measurement is used to simplify the construction of the data kernel for inversion. Experiments on this data set confirm that inversions with continuous functions or discretized blocks provide no intrinsic difference in characterizing the spectrum of Earth as long as both parameterizations possess the same basic resolution and are compared for the same model variance. In the presence of error, coverage of rays in our data set enables retrieval of structure robust overall up to degree 12, given a stringent criterion. We employ a simple and effective antialiasing strategy in constructing the degree 12 image of D" shear velocity structure to minimize the leakage due to expansion truncation. This global mean resolution, however, underestimates that in the Pacific and Eurasia where dense crossings and high quality of data ensure a sharper imaging than elsewhere. For such regions a hybrid model with an additional block parameterization extracts small-scale variations superimposed on the global degree 12 aspherical structure. Our final hybrid model has spatially variable detail added to the long-wavelength image of D". An experimental recovery map that realistically takes into account the errors and distribution of the data is found helpful for identifying resolvable features in the hybrid model. Circum-Pacific high-velocity features are found as in earlier studies, with a general correlation between regions observed to have D" shear velocity discontinuities and volumetric high velocities in D". The predicted roots of plumes show strong affinity to the low-velocity zones in our model; however, the roots tend to be located on the margins of the lowest-velocity areas, not directly centered on them. In pursuit of higher-degree resolution for a global model in the future, this study suggests that simply accumulating huge data sets does not mandate better results: what is more important is to enhance crossing path coverage along with always assessing the variable resolution of any model.

Acknowledgments. We thank the many researchers who kindly supplied us with data from portable experiments or regional arrays: H. Benz, M. Bostock, X.-M. Ding, M. Fouch, S. Grand, D. Kee, V. Levin, H. Patton, J. Ritsema, E. Sandoval, K. Stammler, and D. Wiens. K. C. Tsai has been helpful in expanding the data set. Support for this work was partially covered by NSF grants EAR-9896047, EAR9896048 (E.J.G), and EAR9418643 (T.L.). BYK is supported by National Science Council of Taiwan, Republic of China, under grant NSC-89-2116-M-001-014.

References

Aki, K., and P. G. Richards, *Quantitative Seismology*, W. H. Freeman, New York, 1980.

Bijwaard, H., and W. Spakman, Tomographic evidence for a narrow whole mantle plume below Iceland, *Earth Planet. Sci. Lett., 166*, 121-126, 1999.

Bréger, L., and B. Romanowicz, Three-dimensional structure at the base of the mantle beneath the central Pacific, *Science, 282*, 718-720, 1998.

Bullen, K. E., Compressibility-pressure hypothesis and Earth's interior, *Mon. Not. R. Astron. Soc. Geophys. Suppl., 5*, 355-368, 1949.

Ding, X., and D. V. Helmberger, Modeling D" structure beneath Central America with broadband seismic data, *Phys. Earth Planet. Inter., 101*, 245-270, 1997.

Dziewonski, A. M., Mapping the lower mantle: Determination of lateral heterogeneity in P velocity up to degree and order 6, *J. Geophys. Res. 89*, 5929-5952, 1984.

Dziewonski, A. M., and D. L. Anderson, Preliminary Reference Earth Model, *Phys. Earth Planet. Inter. 25*, 297-356, 1981.

Dziewonski, A. M., and J. H. Woodhouse, Global images of the Earth's interior, *Science, 236*, 37-48, 1987.

Dziewonski, A. M., X. F. Liu, and W. J. Su, Lateral heterogeneity in the lowermost mantle, in *Earth's Deep Interior: The Doornbos Memorial Volume*, edited by D. J. Crossley, pp. 11-49, Gordon and Breach Sci., Newark, N. J., 1997.

Gaherty, J. B., and T. Lay, Investigation of laterally heterogeneous shear velocity structure in D" beneath Eurasia, *J. Geophys. Res., 97*, 417-435, 1992.

Garnero, E. J., and T. Lay, Effects of D" anisotropy on seismic velocity models of the outermost core, *Geophys. Res. Lett., 25*, 2341-2344, 1998.

Garnero, E. J., J. Revenaugh, Q. Williams, T. Lay, and L. H. Kellogg, Ultralow velocity zone at the core-mantle boundary, in *The Core-Mantle Boundary Region, Geodyn. Ser., vol. 28*, edited by M. Gurnis et al., pp. 319-334, AGU, Washington, D. C., 1998.

Garnero, E. J., and D. V. Helmberger, Travel times of *S* and *SKS*: Implications for three-dimensional lower mantle structure beneath the central Pacific, *J. Geophys. Res., 98*, 8225-8241, 1993.

Garnero, E. J., and D. V. Helmberger, Seismic detection of a thin laterally varying boundary layer at the base of the mantle beneath the central Pacific, *Geophys. Res. Lett., 23*, 977-980, 1996.

Garnero, E. J., D. V. Helmberger, and S. Grand, Preliminary evidence for a lower mantle shear wave velocity discontinuity beneath the central Pacific, *Phys. Earth Planet. Inter., 79*, 335-347, 1993.

Grand, S. P., R. D. van der Hilst, and S. Widiyantoro, High resolution global tomography: A snapshot of convection in the Earth, *GSA Today, 7*, 1-7, 1997.

Helmberger, D. V., L. Wen, and X. Ding, Seismic evidence that the source of the Iceland hotspot lies at the core-mantle boundary, *Nature, 396*, 251-255, 1998.

Hung, S. H., F. A. Dahlen, and G. Nolet, Fréchet kernels for finite-frequency traveltimes, II, Examples, *Geophys. J. Int., 141*, 175-203, 2000.

Kendall, J. M., and C. Nangini, Lateral variations in D" below the Caribbean, *Geophys. Res. Lett., 23*, 399-402, 1996.

Kendall, J. M., and P. M. Shearer, On the structure of the lowermost mantle beneath the southwest Pacific, southeast Asia and Australasia, *Phys. Earth Planet. Inter., 92*, 85-98, 1995.

Kennett, B. L. N., *Seismic Wave Propagatiion in Stratified Media*, 342, pp., Cambridge Univ. Press, New York, 1985.

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

Kennett, B. L. N., and O. Gudmundsson, Ellipticity corrections for seismic phases, *Geophys. J. Int., 127*, 40-48, 1996.

Kuo, B. Y., Amplitude of S_{diff} across Asia: effects of velocity gradient and Qs in the D" region and the asphericity of the mantle, *Earth Planet. Sci. Lett., 173*, 101-112, 1999.

Kuo, B. Y., and K. Y. Wu, Global shear velocity heterogeneities in the D" layer: Inversion from *Sd-SKS* differential travel times, *J. Geophys. Res., 102*, 11,775-11,788, 1997.

Lay, T., and D. V. Helmberger, A lower mantle *S*-wave triplication and the shear velocity structure of D", *Geophys. J. R. Astr. Soc., 75*, 799-838, 1983.

Lay, T., Q. Williams, E. J. Garnero, L. Kellog, and M. E. Wysession, Seismic wave anisotropy in the D" region and its implications, in *The Core-Mantle Boundary Region, Geodyn. Ser., vol. 28*, edited by M. Gurnis et al., pp. 299-318, AGU, Washington, D. C., 1998.

Li, X. D., and B. Romanowicz, Global mantle shear-velocity model developed using nonlinear asymptotic coupling theory, *J. Geophys. Res., 101*, 22,245-22,272, 1996.

Liu, X. F., and A. M. Dziewonski, Global analysis of shear wave velocity anomalies in the lower-most mantle, in *The Core-Mantle Boundary Region, Geodyn. Ser., vol. 28*, pp. 21-36, edited by M. Gurnis et al., AGU, Washington, D. C., 1998.

Masters, G., S. Johnson, G. Laske, and H. Bolton, A shear-velocity model of the mantle, *Philos. Trans. R. Soc. London, Ser. A, 354*, 1385-1411, 1996.

Matzel, E, M. K. Sen, and S. P. Grand, Evidence for anisotropy in the deep mantle beneath Alaska, *Geophys. Res. Lett., 23*, 2417-2420, 1996.

Müller, G., The reflectivity method: A tutorial, *J. Geophys., 58*, 153-174, 1985.

Nataf, H. C., and J. VanDecar, Seismological detection of a mantle plume?, *Nature, 364*, 115-120, 1993.

Olivieri, M., N. A. Pino, and A. Morelli, Evidence for an S-velocity discontinuity in the lowermost mantle beneath the South Eastern Pacific Basin, *Geophys. Res. Lett., 24*, 2617-2620, 1997.

Parker, R. L., Understanding inverse theory, *Annu. Rev. Earth Planet Sci., 5*, 35-64, 1977.

Pulliam, R. J., and P. B. Stark, Bumps on the core-mantle boundary: Are they facts or artifacts?, *J. Geophys. Res., 98*, 1943-1955, 1993.

Resovsky, J. S., and M. H. Ritzwoller, A degree 8 mantle shear velocity model from normal mode observations below 3 mHz, *J. Geophys. Res., 104*, 993-1014, 1999.

Ritsema, J. E., E. J. Garnero, and T. Lay, A strongly negative shear velocity gradient and lateral variability in the lowermost mantle beneath the Pacific, *J. Geophys. Res., 102*, 20,395-20,411, 1997.

Ritsema, J., H. J. van Heijst, and J. H. Woodhouse, Complex shear wave velocity structure imaged beneath Africa and Iceland, *Science, 286*, 1925-1928, 1999.

Russell, S. A., T. Lay, and E. J. Garnero, Seismic evidence for small-scale dynamics in the lowermost mantle at the root of the Hawaiian hotspot, *Nature, 396*, 255-258, 1998.

Silver, P. G., Seismic anisotropy beneath the continents: Probing the depths of geology, *Annu. Rev. Earth Planet. Sci., 24*, 385-432, 1996.

Sleep, N. H., Hotspots and mantle plumes: Some phenomenology, *J. Geophys. Res., 95*, 6715-6736, 1990.

Stacey, F. D., and D. E. Loper, The thermal boundary layer interpretation of D" and its role as a plume source, *Phys. Earth Planet. Inter., 33*, 45-55, 1983.

Steinberger, B., Plumes in a convecting mantle: Models and observations for individual hotspots, *Geophys. J. Int.*, in press, 2000.

Stevenson, D. J., Limits on lateral density and velocity variations in the Earth's outer core, *Geophys. J. R. Astron. Soc., 88*, 311-319, 1987.

Su, W. J., R. L. Woodward, and A. M. Dziewonski, Degree 12 model of shear velocity heterogeneity in the mantle, *J. Geophys. Res., 99*, 6945-6980, 1994.

Sylvander, M., and A. Souriau, Mapping *S*-velocity heterogeneities in the D" region, from *SmKS* differential travel times, *Phys. Earth Planet. Inter., 94*, 1-21, 1996.

Tanimoto, T., Long-wavelength S-wave velocity structure throughout the mantle, *Geophys. J. Int., 100*, 327-336, 1990.

Trampert, J., and R. Snieder, Model estimations biased by truncated expansions: Possible artifacts in seismic tomography, *Science, 271*, 1257-1260, 1996.

VanDecar, J. C., D. E. James, and M. Assumpcao, Seismic evidence for a fossil mantle plume beneath South America and implications for plate driving forces, *Nature, 378*, 25-31, 1995.

van der Hilst, R. D., S. Widiyantoro, and E. R. Engdahl, Evidence for deep mantle circulation from global tomography, *Nature, 386*, 578-584, 1997.

van der Hilst, R. D., S. Widiyantoro, K. Creager, and T. J. McSweeney, Deep subduction and aspherical variations in P-wavespeed at the base of Earth's mantle, in *The Core-Mantle Boundary Region, Geodyn. Ser. Vol. 28*, edited by Gurnis et al., pp. 5-20, AGU, Washington, D. C., 1998.

Vasco, D. W., and L. R. Johnson, Whole Earth structure estimated from seismic arrival times, *J. Geophys. Res., 103*, 2633-2671, 1998.

Vidale, J. E., and M. A. H. Hedlin, Evidence for partial melt at the core-mantle boundary north of Tonga from the strong scattering of seismic waves, *Nature, 391*, 682-685, 1998.

Weber, M., *P*- and *S*-wave reflections from anomalies in the lowermost mantle, *Geophys. J. Int., 115*, 183-210, 1993.

Wen, L., and D. L. Anderson, The fate of slabs inferred from seismic tomograpy and 130 millions years of subduction, *Earth Planet. Sci. Lett., 133*, 185-198, 1995.

Wen, L., and D. V. Helmberger, Ultra-low velocity zones near the core-mantle boundary from broadband *PKP* precursors, *Science, 279*, 1701-1703, 1998.

Wiens, D. A., P. J. Shore, J. McGuire, E. Roth, M. G. Bevis, and K. Draunidalo, The southwest Pacific seismic experiment, *IRIS Newsl., 14*, 1-4, 1995.

Williams, Q., and E. J. Garnero, Seismic evidence for partial melt at the base of Earth's mantle, *Science, 273*, 1528-1530, 1996.

Williams, Q., J. Revenaugh, and E. J. Garnero, A correlation between ultra-low basal velocities in the mantle and hot spots, *Science, 281*, 546-549, 1998.

Wolfe, C. J., and P. G. Silver, Seismic anisotropy of oceanic upper mantle: Shear wave splitting methodologies and observations, *J. Geophys. Res., 103*, 749-771, 1998.

Woodward, R. L., and G. Masters, Lower mantle structure from *ScS-S* differential travel times, *Nature, 352*, 231-233, 1991.

Wysession, M. E., Large-scale structure at the core-mantle boundary from diffracted waves, *Nature, 382*, 244-248, 1996.

Wysession, M. E., L. Bartko, and J. B. Wilson, Mapping the lowermost mantle using core-reflected shear waves, *J. Geophys. Res., 99*, 13,667-13,684, 1994. (Correction to "Mapping the lowermost mantle using core-reflected shear waves" by M. E. Wysession, L. Bartko, and J. B. Wilson, *J. Geophys. Res., 100*, 8351, 1995.)

Wysession, M. E., T. Lay, J. Revenaugh, Q. Williams, E. J. Garnero, R. Jeanloz, and L. H. Kellogg, The D" discontinuity and its implications, in *The Core-Mantle Boundary Region, Geodyn. Ser. Vol. 28*, edited by Gurnis et al., pp. 273-297, AGU, Washington, D. C., 1998.

Yanagisawa, T., and Y. Hamano, "Skewness" of S-wave velocity in the mantle, *Geophys. Res. Lett., 26*, 791-794, 1999.

Young, C. J., and T. Lay, Evidence for a shear velocity discontinuity in the lower mantle beneath India and the Indian Ocean, *Phys. Earth Planet. Inter., 49*, 37-53, 1987.

Young, C. J., and T. Lay, Multiple phase analysis of the shear velocity structure in the D" region beneath Alaska, *J. Geophys. Res., 95*, 17,385-17,402, 1990.

Zhao, L., T. H. Jordan, and C. H. Chapman, Three-dimensional Fréchet differential kernels for seismic delay times, *Geophys. J. Int., 141*, 558-576, 2000.

E. J. Garnero, Department of Geology, Arizona State University, Box 871404, Tempe, AZ 85287-1404. (garnero@asu.edu)

B.Y. Kuo, Institute of Earth Sciences, Academia Sinica, POB 1-55 Nankang, Taipei, Taiwan. (byk@earth.sinica.edu.tw)

T. Lay, Institute of Tectonics, University of California, Earth and Marine Science Bldg., Santa Cruz, CA 95064. (thorne @emerald.ucsc.edu

(Received August 12, 1999; revised April 17, 2000;
accepted September 1, 2000)