

## Resolution, uncertainty, and whole Earth tomography

D. W. Vasco<sup>1</sup> and Lane R. Johnson<sup>1</sup>

Center for Computational Seismology, Berkeley Laboratory, University of California, Berkeley, California, USA

Osni Marques

High Performance Computing Research Department, Berkeley Laboratory, University of California, Berkeley, California, USA

Received 1 February 2001; revised 1 July 2002; accepted 23 October 2002; published 16 January 2003.

[1] Using the arrival times of 10 distinct phases from the International Seismological Center catalogue ( $P$ ,  $S$ ,  $PP$ ,  $PcP$ ,  $SS$ ,  $ScS$ ,  $PKPab$ ,  $PKPbc$ ,  $PKPdf$ ,  $SKSac$ ) and two sets of differential times ( $SS - S_{410}S$  and  $SS - S_{660}S$ ), we estimate the  $P$  and  $S$  velocity structure of the whole Earth (mantle, outer core, and inner core), boundary topography of four interfaces within the Earth, hypocentral relocation parameters, and station corrections. The entire catalogue is relocated and reassociated using the radial velocity model ak135. Both  $P$  and  $S$  phases are used in the relocation of over 41,000 events. In depth, isotropic velocity variations within the Earth are parameterized by 22 layers. Lateral variations in shear and compressional velocity within the layers are represented by one of two grids. In the mantle, each layer is subdivided into 4560 equal-volume cells, of dimension  $3^\circ \times 3^\circ$  at the equator. However, in the outer and inner core, the cell size is increased to  $6^\circ \times 6^\circ$ . This change in grid spacing accounts for decreasing cell volumes with depth and variations in sampling between the mantle and core. Previous attempts using a fine grid throughout the Earth resulted in poor resolution of core structure. With the current grid, the resolution of velocity heterogeneity in the mantle and outer core is comparable. In the upper and midmantle, the well-resolved cells lie beneath Eurasia and a narrow zone surrounding the mantle beneath the Pacific basin. In the lowest mantle, the resolution decreases significantly. An examination of the averaging kernels associated with six locations near the base of the mantle reveals the significant lateral and vertical averaging inherent in estimates of compressional velocity. In well-sampled regions of the outer core, we are able to resolve compressional velocity variations in our  $6^\circ \times 6^\circ$  cells. However, velocity variations in the inner core are poorly constrained at the scale of our blocks. In well-sampled regions of the mantle and outer core model, parameter standard deviations remain below 0.25%. For much of the inner core, the standard deviations are quite large, exceeding 0.5%. Narrow, high-velocity anomalies are imaged in the upper mantle, associated with subduction zones. In a number of locations, such as beneath the Americas, these features extend well into the lower mantle. In the outer core, heterogeneity is generally lower in magnitude, of the order of 0.5%. The most notable feature in the outer core is a systematic shift to slightly higher velocity. In addition, we find a roughly symmetric distribution of heterogeneity at the base of the outer core. Compressional velocities are generally higher in the equatorial regions and lower at the poles.

*INDEX TERMS:* 7203 Seismology: Body wave propagation; 7207 Seismology: Core and mantle; 3260 Mathematical Geophysics: Inverse theory; 1734 History of Geophysics: Seismology

**Citation:** Vasco, D. W., L. R. Johnson, and O. Marques, Resolution, uncertainty, and whole Earth tomography, *J. Geophys. Res.*, 108(B1), 2002, doi:10.1029/2001JB000412, 2003.

### 1. Introduction

[2] Since the early studies of *Sengupta and Toksoz* [1976], *Dziewonski et al.* [1977], *Clayton and Comer* [1983], and

*Dziewonski* [1984] there has been a continual refinement in our three-dimensional images of the Earth's interior. The improvements have been brought about by both increasing amounts of high quality seismological data and advancements in methodology. Summaries of these past efforts can be found in review articles [*Romanowicz*, 1991; *Iyer and Hirahara*, 1993; *Nolet et al.*, 1994; *Ritzwoller and Lavelle*, 1995].

[3] Some of the more recent advances involve the use of finer scale models and the incorporation of such factors as

<sup>1</sup>Also at Berkeley Seismological Laboratory, University of California, Berkeley, California, USA.

This paper is not subject to U.S. copyright.  
Published in 2003 by the American Geophysical Union.

mantle anisotropy. A number of investigators have reduced the cell size used to represent mantle velocity structure to  $1^\circ$  to  $2^\circ$  [Zhou, 1996; van der Hilst *et al.*, 1997; Kennett *et al.*, 1998]. Using a variable sized grid Bijwaard *et al.* [1998] has allowed for heterogeneity on a scale as fine as  $0.6^\circ$ . On a related note, Trampert and Woodhouse [1996] have pushed the spherical harmonic representation up to degree 40. Such fine scale heterogeneity is suggested by surface geology, regional seismic studies and the scattering of seismic waves. Furthermore, there have been studies of the distribution of anisotropy in the upper mantle [Montagner and Tanimoto, 1991] as well as throughout the mantle [Boschi and Dziewonski, 2000]. The additional parameters associated with finer scale models and more complete representations will potentially reduce the chance of biasing our models by underparameterization. However, with so many model parameters we have the possibility of significant trade-offs between sets of unknowns and a corresponding lack of resolution.

[4] There have been many studies employing a single class of seismological data, such as surface waves, body waves, or normal mode data, to constrain Earth structure. Several recent studies have also integrated several types of data to better constrain Earth structure. Masters *et al.* [1996] have integrated a wide variety of data: long period absolute and differential times, surface wave phase velocities, and mode structure coefficients in order to infer mantle structure. Su and Dziewonski [1997] use a similar diversity of data types to constraint shear and bulk sound speed in the mantle.

[5] There are now numerous models of mantle velocity structure, each employing a particular representation of heterogeneity, either isotropic or anisotropic, and some combination of seismological data to constrain Earth structure. To a significant degree there has been a convergence in Earth models and many velocity anomalies are now accepted as features of our planet. However, it is desirable to compare the various models in more detail and to discern which features are reliably determined and which features might be influenced by factors such as the regularization used in the inversion. A prerequisite for quantitatively comparing models is some estimate of model parameter resolution and uncertainty. It does not make sense to compare models in those parts of the Earth which are poorly resolved nor to compare anomalies which are below their estimated uncertainties. Thus, a necessary step in comparing models is some form of model assessment.

[6] In this paper we describe the results of an inversion for mantle and core velocity heterogeneity. We also present estimates of resolution and uncertainty associated with our model. This work is an extension of a previous inversion for mantle and core heterogeneity [Vasco *et al.*, 1999] in which we represented lateral variations from a radial Earth model by a set of  $6^\circ \times 6^\circ$  cells. Here we utilize a two-scale grid with  $3^\circ \times 3^\circ$  cells in the mantle and  $6^\circ \times 6^\circ$  cells in the inner and outer core. We have reduced the cell size in the mantle so that we can begin to examine the resolution associated with finer scale models. As noted above, many such models have been published and it is not clear how well we can resolve features of the order of  $1^\circ$  or smaller. Also, the use of smaller cells in the mantle reduces the possibility of mapping finer scale mantle structure into core heterogeneity. Furthermore, more general parameteriza-

tions, such as radial anisotropy throughout the mantle [Boschi and Dziewonski, 2000] introduce the possibility of additional trade-offs between model parameters. We generalize the methodology for computing resolution and covariance to allow for nondiagonal forms of regularization, such as a roughness penalty. Our work should be viewed as a first step in formally assessing fine scale models of Earth structure. The methodology employed here should prove useful in estimating resolution for more complicated formulations of the inverse problem, such as those incorporating radial anisotropy. The techniques are also applicable to studies utilizing multiple data types, e.g. body waves, surface waves, and free oscillations.

[7] In order to reduce the length of this paper we will only discuss the results associated with compressional velocity estimates. That is, even though we also solve for shear velocity variations, topography on four boundaries, inner core anisotropy, hypocentral, and station corrections, we will not discuss our estimates of those quantities here.

## 2. Methodology

[8] In this section we outline the procedures involved in our estimation of Earth structure. The general approach is described in more detail in Vasco *et al.* [1994], Vasco *et al.* [1995], and Vasco and Johnson [1998]. The overall methodology is similar to that of Inoue *et al.* [1990] and Pulliam *et al.* [1993], and the more recent studies of van der Hilst *et al.* [1997] and Bijwaard *et al.* [1998]. We also touch upon model assessment using the singular value decomposition (SVD) [Noble and Daniel, 1977]. We describe enhancements to the methodology that allow us to consider more general, non-diagonal, regularization, such as roughness penalties. To our knowledge, the expressions presented here for resolution and covariance matrices for inverse problems with general regularization are original. Finally, we outline an approach for conducting a partial SVD using the Lanczos algorithm [Lanczos, 1950]. Full details of our particular implementation of this algorithm are presented by Vasco *et al.* [1999].

### 2.1. Travel Time Data Sets and Systems of Equations

[9] Our basic datum is the travel time residual, the observed arrival time minus the calculated arrival time. In the work of Vasco and Johnson [1998] we discuss the most significant contributions to a travel time residual: volumetric heterogeneity, event mislocations, stations corrections, boundary topography, anisotropy, and random data errors. We represent the travel time residual as a sum of these respective contributions

$$\begin{aligned} \delta t_{ijl} = & \sum_{k=1}^{N_{3D}} \left[ \frac{\partial t_{ijl}}{\partial v^k} \right] \delta v^k + \sum_{k=1}^4 \left[ \frac{\partial t_{ijl}}{\partial h^k} \right] \delta h^k + \sum_{k=1}^{N_s} \left[ \frac{\partial t_{ijl}}{\partial s^k} \right] \delta s^k \\ & + \sum_{k=1}^{N_b} \left[ \frac{\partial t_{ijl}}{\partial b^k} \right] \delta b^k + \sum_{k=1}^{N_a} \left[ \frac{\partial t_{ijl}}{\partial a^k} \right] \delta a^k + \epsilon \end{aligned} \quad (1)$$

where quantities in parenthesis are sensitivity coefficients relating changes in model parameters to changes in the travel time residual. Computation of these sensitivity coefficients is discussed by Morelli and Dziewonski [1987], Nolet [1987], Inoue *et al.* [1990], Pulliam *et al.*

[1993], and *Vasco and Johnson* [1998]. The velocity variations are represented by  $N_{3D}$  expansion coefficients  $\delta v_i^k$ . These are the coefficients in a representation of the velocity variations by an orthogonal set of basis functions, in our case the constant velocity cells mentioned by *Vasco and Johnson* [1998]. In our previous work we parameterized the volumetric heterogeneity using volume elements of dimension  $6^\circ \times 6^\circ$  (at the equator). Currently, we adopt a two level parameterization in which  $3^\circ \times 3^\circ$  cells are used in the mantle and  $6^\circ \times 6^\circ$  cells are used in the core. Larger cells are used in the inner and outer core to account for the decreasing surface area with decreasing radius. There are four relocation parameters,  $\delta h_j^k$ , for each summary event: shifts in origin time, depth, latitude, and longitude. Associated with each station there are 2 station corrections,  $\delta s_{il}^k$ : one for compressional waves and one for shear phases. The boundary variations,  $\delta b^k$ , are parameterized by  $N_b$  equal-area pixels of dimension  $3^\circ \times 3^\circ$  (at the equator). That is, each boundary of the four boundaries we consider, [the 410 km, the 660 km, the core–mantle and inner core–outer core boundaries] is decomposed into 4544 surface elements. Currently, anisotropy is only included in the inner core. The coefficients  $\delta a^k$  allow for inner core anisotropy that is a function of the angle from the rotation axis [*Vasco and Johnson*, 1998]. The final contribution,  $\epsilon$ , represents a random component. Note that, due to a preponderance of outliers in the arrival time data, this error is not expected to be governed by a normal distribution [*Buland*, 1986].

[10] Our image of the Earth is based upon a large collection of arrival times, 12 phases in all [*Vasco and Johnson*, 1998], representing energy which has propagated along various paths through the Earth. Ten phases are derived from a reprocessing of ISC observations, as detailed by *Vasco and Johnson* [1998]. The other two sets of measurements are  $SS - S_{410S}$  and  $SS - S_{660S}$  differential times obtained from long-period waveform data [*Shearer*, 1993]. More details on the incorporation of these differential times may be found in the work of *Vasco et al.* [1995]. The suite of travel time residuals (observed-calculated travel time) from  $N_e$  events recorded by different subsets of  $N_s$  stations may be represented as a vector  $\delta t$ . Furthermore, the unknown parameters may also be collected into vectors for each class of unknowns (velocity heterogeneity, hypocentral relocation, station parameters, boundary deviations, and anisotropy). Then the set of equations (1) may be written in the form of a matrix equation

$$\delta t = \mathbf{V}\delta v + \mathbf{H}\delta h + \mathbf{S}\delta s + \mathbf{B}\delta b + \mathbf{A}\delta a + \epsilon. \quad (2)$$

The model is parameterized by  $2 \times N_{3D}$  velocity unknowns ( $P$  and  $S$  for each of the  $N_{3D}$  cells),  $4 \times N_e$  hypocentral relocation parameters,  $2 \times N_s$  station corrections,  $N_b$  boundary perturbations, and  $N_a$  anisotropy terms.  $\mathbf{V}$ ,  $\mathbf{H}$ ,  $\mathbf{S}$ ,  $\mathbf{B}$ , and  $\mathbf{A}$  represent matrices containing the partial derivatives of the travel time with respect to each class of model parameter, as in equation (1). The above matrix equation represents the basic constraints used to infer Earth structure.

## 2.2. Robust Inversion With Regularization

[11] Due to the presence of numerous outliers, errors associated with our arrival time data do not obey a Gaussian distribution [*Vasco et al.*, 1994]. Using a least squares algorithm in the presence of significant outliers will result

in parameters values which are not maximum likelihood estimates. One could simply apply a residual cutoff such that the outliers are removed. However, useful data may be found beyond the cutoff and it is difficult to decide exactly where the cutoff should lie. Alternatively, one could choose the regularization, the damping and smoothing, to minimize the influence of large residuals. Unfortunately, this may result in a heavily damped or smoothed model that significantly underestimates the level of heterogeneity. We prefer to tailor the inversion algorithm to be compatible to the approximate distribution of the residual errors. Specifically, rather than minimize the sum of the squares of the misfit we minimize

$$|\mathbf{r}|_p = \left( \sum_i |r_i|^p \right)^{1/p} \quad (3)$$

where  $\mathbf{r}$  is the misfit vector

$$\mathbf{r} = \delta t - \mathbf{V}\delta v - \mathbf{H}\delta h - \mathbf{S}\delta s - \mathbf{B}\delta b - \mathbf{A}\delta a. \quad (4)$$

As described by *Vasco et al.* [1994], the value of  $p$  is chosen to produce approximately maximum likelihood estimates. The algorithm for estimating  $p$  is based upon an analysis of Monte Carlo simulations by *Gonin and Money* [1985]. In their study it was determined that  $p$  depends on the forth moment of the residual distribution. In particular, it is given by the empirical formula,  $p = 9/K^2 + 1$ , where  $K$  is the kurtosis of the residual distribution, the forth moment divided by the standard deviation to the fourth power.

[12] As in all mixed-determined inverse problems we must regularize the inversion algorithm in some fashion. This ensures that errors in the data are not greatly magnified in poorly constrained regions, leading to artifacts. We incorporate model norm and model roughness penalty terms in the inversion to effect the regularization. That is, for parameters which are poorly constrained, we bias the values toward the background model. Similarly, if no data constraints are active, the roughness penalty biases the estimates such that they vary smoothly within a layer. Note that we do not apply depth smoothing in our inversion. The details are given by *Pulliam et al.* [1993], *Vasco et al.* [1994, 1995], and *Vasco and Johnson* [1998].

[13] The solution of the inverse problem entails finding an Earth model which simultaneously minimizes the misfit to the data, equation (3), and the penalty measures. An algorithm for general  $p$  residual norm minimization is presented by *Scales et al.* [1988]. In their algorithm the  $p$  residual norm is minimized by solving a series of reweighted linear equations of the form

$$\mathbf{N}^{1/2} \mathbf{M} \mathbf{d} = \mathbf{N}^{1/2} \delta \mathbf{T}. \quad (5)$$

The scaling matrix  $\mathbf{N}$  is a diagonal matrix consisting of  $|r_i|^{p-2}$  on that portion of the diagonal corresponding to the data constraints and ones on the remainder of the diagonal. That is, we only reweight the data constraints, the regularization constraints are not adjusted. In our implementation,  $\mathbf{M}$  is an augmented matrix containing both data and penalty coefficients. The exact form of  $\mathbf{M}$  is given in Appendix A. The matrix  $\mathbf{M}$  has dimension  $1,433,102 \times 307,134$  but is quite sparse, consisting of 48,317,209 nonzero elements. The algorithm proceeds iteratively, first solving for the parameters  $\delta \mathbf{x}$ , using these values to calculate the weights ( $r_i$ ) and then solving the reweighted linear

system (5). The whole procedure is cycled through repeatedly until convergence.

### 2.3. Model Assessment

#### 2.3.1. Regularization and the Singular Value Decomposition (SVD)

[14] When presenting images of Earth structure it is important to provide estimates of model parameter resolution and uncertainty. In most geophysical inverse problems numerical measures of these quantities take the form of resolution and covariance matrices. Strictly speaking, resolution and covariance matrices are designed for least squares estimates, problems in which the errors are normally distributed. Our  $\ell^p$  inversion algorithm involves a series of reweighted least squares problems in which the weighting is designed to transform the data such that, in the transformed domain, the errors are normally distributed. Thus, we apply the assessment techniques to the last set of iterations of our algorithm, using the weights from the final cycle to calculate the matrices. In this subsection we define the reweighted coefficient matrix

$$\mathbf{G} = \mathbf{N}^{1/2} \mathbf{M} \quad (6)$$

and reweighted data vector

$$\mathbf{T} = \mathbf{N}^{1/2} \delta \mathbf{T} \quad (7)$$

and equation (5) then reads

$$\mathbf{T} = \mathbf{G} \delta \mathbf{x}. \quad (8)$$

The least squares solution of (8) is the solution of the linear system

$$[\mathbf{G}^T \mathbf{G}] \delta \mathbf{x} = \mathbf{G}^T \mathbf{T}. \quad (9)$$

Because of the structure of  $\mathbf{M}$  and  $\delta \mathbf{T}$  (see Appendix A), we may write both  $\mathbf{G}$  and  $\mathbf{T}$  in the partitioned forms

$$\mathbf{G} = \begin{bmatrix} \Gamma \\ \mathbf{L} \end{bmatrix} \quad (10)$$

and

$$\mathbf{T} = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix}. \quad (11)$$

Equation (9) then takes the form

$$[\Gamma^T \Gamma + \mathbf{L}^T \mathbf{L}] \delta \mathbf{x} = \Gamma^T \mathbf{d}. \quad (12)$$

The matrix  $\Gamma$  is associated with the scaled rows corresponding to data constraints while the matrix  $\mathbf{L}$  corresponds to the scaled regularization rows. Similarly,  $\mathbf{T}$  is partitioned into scaled data rows  $\mathbf{d}$  and zero rows.

[15] We assume that, due to the regularization, the matrix product  $\mathbf{G}^T \mathbf{G}$  in equation (9) is invertible (nonsingular). We may then write the estimate of  $\delta \mathbf{x}$  as

$$\delta \hat{\mathbf{x}} = [\mathbf{G}^T \mathbf{G}]^{-1} \mathbf{G}^T \mathbf{T}. \quad (13)$$

Our estimates will be based upon a decomposition of  $\mathbf{G}$  as the product

$$\mathbf{G} = \mathbf{U} \Lambda \mathbf{V}^T \quad (14)$$

where  $\Lambda$  is a diagonal matrix of singular values and the matrices  $\mathbf{U}$  and  $\mathbf{V}$  satisfy the orthogonality relations  $\mathbf{U}^T \mathbf{U} = \mathbf{I}$  and  $\mathbf{V}^T \mathbf{V} = \mathbf{I}$ .

[16] The generalized inverse follows from a truncated representation of  $\mathbf{G}$

$$\mathbf{G} = \mathbf{U}_\kappa \Lambda_\kappa \mathbf{V}_\kappa^T$$

where the  $\kappa$  signifies that only  $\kappa$  columns of  $\mathbf{V}$ ,  $\mathbf{U}$  and  $\Lambda$  are retained in the decomposition. Then, the generalized inverse has the truncated representation

$$\mathbf{G}^\dagger = \mathbf{V}_\kappa \Lambda_\kappa^{-1} \mathbf{U}_\kappa^T, \quad (15)$$

resulting in the model parameter estimates

$$\delta \hat{\mathbf{x}} = \mathbf{V}_\kappa \Lambda_\kappa^{-1} \mathbf{U}_\kappa^T \mathbf{T}. \quad (16)$$

Note that we have formed the generalized inverse of the augmented matrix  $\mathbf{G}$  in (10). The presence of the additional equations, such as those associated with the roughness penalty will act to shape or filter the spectrum of the original matrix  $\Gamma$ . As noted by *Pratt and Chapman* [1992], in the presence of smoothing, our estimates may contain contributions from the singular vectors of the null-space of the unconstrained problem.

#### 2.3.2. Model Parameter Resolution

[17] Model parameter resolution is a numerical measure of averaging or filtering inherent in our estimation procedure. As shown in many texts on inverse methods [*Menke*, 1984; *Parker*, 1994], both the resolution and covariance matrices may be derived directly from the singular value decomposition of the matrix  $\mathbf{G}$ . There are alternatives to the SVD for computing resolution and covariance matrices. For example, an LU decomposition technique, based upon the approach of *Tarantola* [1987], was applied to mantle tomography by *Vasco et al.* [1993]. Recently, an approximation based upon a one-step backprojection approach was introduced [*Nolet et al.*, 1999]. For inverse problems in which regularization takes the form of a singular value cutoff, in which only the  $\kappa$  largest values are retained, the resolution matrix takes the form

$$\mathbf{R} = \mathbf{V}_\kappa \mathbf{V}_\kappa^T \quad (17)$$

[*Menke*, 1984; *Parker*, 1994]. However, for more general, nondiagonal regularization such as roughness penalties, the simple formulation is no longer valid [*Pratt and Chapman*, 1992]. In Appendix B we derive the modifications required to compute model parameter resolution in the presence of nondiagonal regularization measures. The approach we take is somewhat different from *Pratt and Chapman* [1992] because we decompose  $\mathbf{G}$  rather than  $\Gamma$ . Our representation of the resolution matrix is

$$\mathbf{R} = \mathbf{V}_\kappa \mathbf{\Upsilon} \mathbf{V}_\kappa^T \quad (18)$$

where

$$\Upsilon = \mathbf{I} - \Lambda_{\kappa}^{-1} (\mathbf{U}_2)_{\kappa}^T (\mathbf{U}_2)_{\kappa} \Lambda_{\kappa} \quad (19)$$

and the matrix  $\mathbf{U}_2$  is associated with the rows of a partitioned form of  $\mathbf{U}$ ,

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix}. \quad (20)$$

The rows of  $\mathbf{U}_2$  correspond to the constraints provided by the regularization.

### 2.3.3. Model Parameter Covariance

[18] The uncertainties associated with our model parameter estimates are described by the model parameter covariance matrix [Menke, 1984; Parker, 1994]. Because the model parameter estimates are linearly related to the data, as indicated in equation (16), there is a direct mapping from data covariance to model parameter covariance [Menke, 1984]. As in our resolution calculations, we may base our model parameter covariance estimates upon a singular value decomposition of the augmented matrix  $\mathbf{G}$ . In the presence of nondiagonal regularization, such a roughness penalties, we must modify the usual formula for model parameter covariances. The necessary modifications are outlined in Appendix B. In essence, we write the covariance matrix  $\mathbf{C}_m$  as

$$\mathbf{C}_m = \mathbf{V}_{\kappa} \Omega \mathbf{V}_{\kappa}^T \quad (21)$$

where

$$\Omega = \Lambda_{\kappa}^{-2} - \Lambda_{\kappa}^{-1} (\mathbf{U}_2)_{\kappa}^T (\mathbf{U}_2)_{\kappa} \Lambda_{\kappa}^{-1}. \quad (22)$$

As indicated in equation (20), the matrix  $(\mathbf{U}_2)$  is composed of those rows of  $\mathbf{U}$  associated with the regularization.

### 2.3.4. Iterative Singular Value Decomposition Using the Lanczos Algorithm

[19] In this subsection we provide an outline of our algorithm for obtaining a partial SVD. The details of our approach may be found in the work of Vasco et al. [1999]. The Lanczos recursion is an algorithm for transforming a symmetric  $L \times L$  matrix  $\mathbf{A}$  into an  $l \times l$  tridiagonal matrix  $\Pi$  using an orthogonal transformation  $\mathbf{Q}$

$$\Pi = \mathbf{Q}^T \mathbf{A} \mathbf{Q} \quad (23)$$

[Lanczos, 1950]. This equation is equivalent to the recursion

$$\beta_{i+1} \mathbf{q}_{i+1} = \mathbf{A} \mathbf{q}_i - \beta_i \mathbf{q}_{i-1} - \alpha_i \mathbf{q}_i \quad i = 1, 2, \dots, l \quad (24)$$

where  $\mathbf{q}_0 = \mathbf{0}$  and the orthonormality property of the Lanczos vectors  $\mathbf{Q}$  is used to choose  $\beta_i$ . The sequence of equations (24) is started with an arbitrary nonzero vector  $\mathbf{q}_1$  of unit length.

[20] Upon completion of  $l$  steps we have the  $l \times l$  tridiagonal matrix  $\Pi$  and the  $L \times l$  matrix  $\mathbf{Q}$  from which the eigenvalues and eigenvectors of  $\mathbf{A}$  may be estimated. The full SVD of  $\mathbf{A}$  follows from an eigen-decomposition of the matrix  $\Pi$  as detailed by Vasco et al. [1999]. When  $l$  equals the order of the matrix  $\mathbf{A}$ , the eigen-decomposition is identical to a conventional SVD of the matrix. Typically

many fewer steps are actually needed for the convergence of eigenvalues and eigenvectors. A rough rule of thumb is that the first  $l/2$  eigenvalues have converged to those which would be produced by a conventional SVD [Berry, 1992].

[21] Convergence of the Lanczos estimates of model parameter resolution and covariance is something of an issue. In particular, it has been pointed out [Deal and Nolet, 1996] that if very few Lanczos iterations are taken,  $l$  is very much less than  $L$  and the resolution estimates will be significantly different from those produced by a conventional QR based SVD calculation. Some of this criticism is based upon a particular implementation of the Lanczos recursion in which orthogonality is not enforced between the vectors  $\mathbf{q}_i$ . This differs from the approach we adopt in which a measure of nonorthogonality is retained throughout the iteration [Berry, 1992]. When the loss of orthogonality is indicated, we reorthogonalize the set of vectors. Our experience with the approach advocated by Cullum and Willoughby [1985], in which orthogonality is not enforced, indicates that it is not a viable method for efficiently computing a partial SVD. Once orthogonality is lost their algorithm simply generates too many repeated singular values and vectors to be practical for constructing a large suite of pairs. Rather, we have found that the selective reorthogonalization scheme of Simon [1984] to be quite effective. In this approach a recursion relates the dot product of the new vector  $\mathbf{q}_{i+1}$  and all columns of  $\mathbf{Q}_i$ , that is  $\mathbf{Q}_i^T \mathbf{q}_{i+1}$  to  $\mathbf{Q}_i^T \mathbf{q}_i$  and  $\mathbf{Q}_i^T \mathbf{q}_{i-1}$ . Only when orthogonality is lost do we reorthogonalize the set of vectors. Using this methodology, we can generate large sets of vectors in a stable fashion. The validity of the Lanczos resolution estimates has been defended by Zhang and McMechan [1995] and Yao et al. [1999]. Our view is that the Lanczos resolution calculations represent conservative estimates of model parameter resolution, lower bounds on conventional SVD-based resolution measures [Vasco et al., 1999]. Furthermore, we conduct enough Lanczos iterations [20,000 for the results described below] so that our estimates have converged in some sense. We measure convergence by calculating the sum of the diagonal of the resolution matrix as a function of the number of iterations. Finally, we compare our resolution and covariance estimates with previous computations, based upon a direct method [Vasco et al., 1993].

[22] For an efficient implementation, we apply the Lanczos algorithm to the symmetric matrix

$$\mathbf{A} = [\mathbf{G}^T \mathbf{G}]. \quad (25)$$

In doing so we obtain the singular values, the diagonal elements of  $\Lambda$ , and the singular vectors, contained in the columns of  $\mathbf{V}$ . Using the relationship

$$\lambda_i \mathbf{u}_i = \mathbf{G} \mathbf{v}_i \quad (26)$$

we obtain the necessary columns of  $\mathbf{U}$ . In this fashion we calculate all the quantities required for our partial SVD.

## 3. Results

### 3.1. Data Analysis

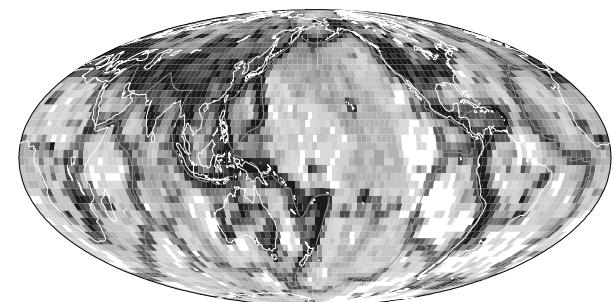
[23] The data reduction necessary to obtain useful travel time residuals has been described in previous publications

[Pulliam et al., 1993; Vasco et al., 1994; Vasco et al., 1995; Vasco and Johnson, 1998; Vasco et al., 1999]. The basic procedure is a complete relocation of some 41,000 events in the ISC catalogue for the years 1964–1987. The one-dimensional model ak135 [Kennett et al., 1995] is used for the relocation. Both  $P$  and  $S$  arrival times are utilized in relocating the earthquakes and stringent criteria are employed to ensure that the events may be reliably located. All arrival times are reassigned, based upon the ak135 travel time tables. The procedures we followed are essentially the same as those described by Engdahl et al. [1998]. However, we did not incorporate depth phases in the relocation. It was felt that such phases might be influenced significantly by regional heterogeneity in the uppermost mantle and crust. We used a more stringent requirement that only events with 50 or more arrivals be used in the inversion.

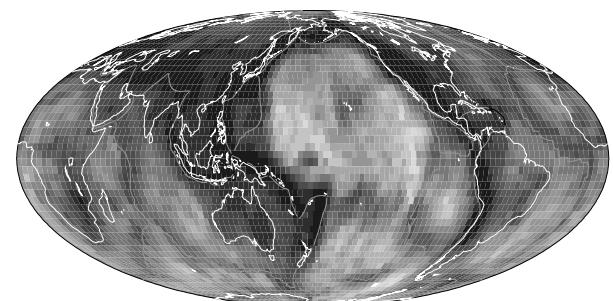
[24] Ten phases ( $P$ ,  $S$ ,  $PP$ ,  $SS$ ,  $PKPab$ ,  $PKPbc$ ,  $PKPdf$ ,  $PcP$ ,  $ScS$ , and  $SKSac$ ), which are sensitive to structure in the mantle, outer core, and inner core were used in the inversion. Two sets of differential times ( $SS - S_{410}S$  and  $SS - S_{660}S$ ), derived from long-period waveforms [Shearer, 1993], also constrain the Earth model. Travel time residuals are formed using the calculated arrival times from the ak135 tables. Arrival times are corrected for the ellipticity of the Earth [Dziewonski and Gilbert, 1976]. Internal reflections such as  $PP$  are corrected for bounce point topography. Summary residuals are formed by averaging travel time residuals for nearly identical travel paths. Rays are traced for all summary residuals and the ray lengths and velocities in each cell are used to form the coefficients for the inversion. For a given layer in our model the cells are roughly of equal volume. In the mantle the surface area is equivalent to  $3^\circ \times 3^\circ$  at the equator. In the core the cells are larger,  $6^\circ \times 6^\circ$  equivalents. The Earth is divided into 22 layers, 12 layers in the mantle, 6 layers in the outer core and 4 layers in the inner core. Layer thickness varies from 35 to 400 km, averaging around 200 km thick. For exact layer boundaries, consult Vasco and Johnson [1998]. Ray path sampling by  $P$  waves, the number of rays intersecting the cells of the model, are shown in Figure 1 for three layers within the mantle. Note the great variation in sampling within the mantle, the scale is logarithmic. At shallow depths within the mantle (35–200 km) the subduction zone source regions are well sampled. Also, the volume beneath the continents of the Northern Hemisphere is well sampled. Just below the transition zone (660–870 km) the best sampled region is a ring extending around the Pacific Ocean and eastward into Europe. At the base of the mantle (2670 km–CMB) the sampling is more uniform and the peak sampling is less than for overlying regions. One would expect the variations in sampling to influence our ability to resolve variations in Earth structure. Model parameter resolution is discussed in the next subsection.

[25] Initially we plotted our travel time residuals as bottoming point averages [Vasco and Johnson, 1998]. That is, for each phase we calculated the median travel time residual of all rays which bottomed in the same cell in a particular depth interval. The cells were approximately equal area, equivalent to  $6^\circ \times 6^\circ$  at the equator. In the work of Vasco and Johnson [1998] the bottoming point medians were plotted as maps, displaying their geographical variation in a particular depth interval. For phases which

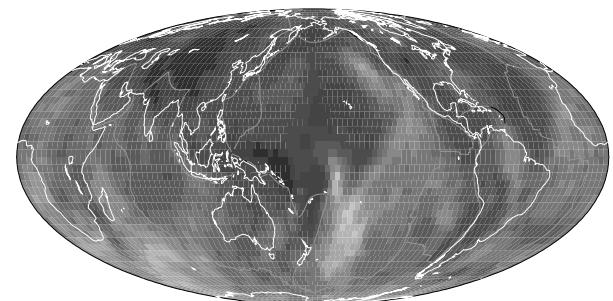
### COMPRESSIVE (35–200 KM)



### COMPRESSIVE (660–870 KM)



### COMPRESSIVE (2670 KM–CMB)



0.00 LOG (RAY DENSITY) 8.00

**Figure 1.** Ray path sampling by compressional waves in the mantle. The logarithmic scale indicates the number of rays intersecting each cell in the three layers shown. White cells indicate unsampled blocks.

interacted with the core we note a systematic variation with latitude. Such patterns have been documented previously by Shearer et al. [1988], Gudmundsson [1989], Souriau and Poupinet [1990], and Roudil and Souriau [1993]. In order to emphasize this trend we averaged the bottoming point medians at each latitude. In Figure 2 the medians for  $P$  in

the lowest mantle, for core-reflected  $PcP$ , and for core transmitted  $SKSac$ ,  $PKPab$ ,  $PKPbc$ , and  $PKPdf$  are shown. Only latitudes for which there were 25 or more longitudinal bins are shown in this figure and the errors on the median are computed from the variation within the bins. The error bars denote one standard deviation associated with the median estimates. Note the positive travel time residuals in the polar regions and the negative residuals near the equator for transmitted core phases ( $SKSac$ ,  $PKPab$ ,  $PKPbc$ , and  $PKPdf$ ). The latitudinal symmetry for transmitted core phases is weak but does appear to exceed the associated errors on the medians. This trend is not seen for P travel times in the lowest mantle, nor is it particularly strong for core reflected  $PcP$ . We varied the depth interval over which we averaged the lower mantle P residuals from about 1,000 km in thickness to that shown in Figure 2, and the resulting pattern was always similar. It appears that there is a systematic variation in core residuals which is not seen in phases which only interact with the mantle. Differences between mantle and core residuals were also noted by *Boschi and Dziewonski* [2000] and *Karason and van der Hilst* [2001].

[26] The methodology used to infer Earth structure, the iterative  $\ell^p$  inversion algorithm, has been outlined above. Our value of  $p$  is based upon the analysis of *Vasco and Johnson* [1998]. They concluded that  $p$  equal to 1.25 was compatible with the composite travel time data set. Three cycles of iterations are conducted, with 15, 30, and 200 iterations in each cycle. The matrix coefficients used in the final set of outer iterations were subsequently used in the Lanczos recursion. The Lanczos algorithm for estimating resolution, equation (18), was implemented on the massively parallel T3E computer at NERSC as described by *Vasco et al.* [1999]. A total of 20,000 iterations were conducted in constructing the Lanczos resolution estimates. The entire procedure took approximately 20 hours on the T3E. The resulting spectral estimates are shown in Figure 3. As noted above, roughly half of these values are expected to have actually converged to 'true' singular values. However, the other estimated values are useful in calculating resolution and covariance. These values represent vectors which have converged to almost degenerate subspaces. That is, subsets of singular vectors associated with nearly equal singular values. From the flattening of the spectrum in Figure 3 one observes how the singular values approach each other in magnitude. As noted by *Vasco et al.* [1999], as the singular values become closer in value the subspace spanned by the associated singular vectors becomes degenerate. To some extent, any linear combination of these vectors may be used to represent the subspace. The Lanczos algorithm converges to such a subset and it is this set that is used to represent the entire subspace.

[27] Because of the size of our inverse problem we cannot compare our resolution and covariance estimates against more complete measures, such as a QR-based SVD, as a way to gauge convergence. Rather, we will follow a procedure described by *Vasco et al.* [1999] to assess the convergence of our resolution estimates. That is, we compare successive traces of the resolution matrix as the number of Lanczos iterations increases. The trace of the resolution matrix is a numerical invariant which is relatively easy to compute. In Figure 4 the ratio of successive diagonal

sums are plotted against the number of iterations. After about 16,000 iterations the change in the diagonal sum appears to be leveling off, having obtained a value of 0.91. This behavior was also noted by *Vasco et al.* [1999], though the transition occurred around 3,000 iterations for their smaller problem. We also monitored convergence by comparing the resolution estimates themselves as the algorithm progressed. We noted little change in the maps of model parameter resolution during the last 3,000 iterations of the algorithm. However, to some degree our estimates will underestimate the resolution obtained by a QR-based algorithm. Thus, our estimates serve as a conservative measure of resolution.

[28] As is evident from equations (15), (16), (18), and (21), our model parameter estimates and assessment are based upon a truncated representation of the weighted coefficient matrix  $\mathbf{G}$ . That is,  $\kappa$  singular values and singular vectors are used in constructing both the inverse and the resolution and covariance estimates. Our choice of  $\kappa$  is based upon a singular value cutoff in which only singular values greater than  $0.001 \times \lambda_{max}$  are retained, where  $\lambda_{max}$  is the largest singular value. Based upon this strategy we utilized 19,890 singular values and vectors in our computations. This number should be compared with the size of the matrix  $\mathbf{G}$  which contained 1,433,102 rows and 307,134 columns. Thus, the effective rank of the matrix is more than an order of magnitude smaller than the number of columns.

### 3.2. Lanczos Resolution Estimates

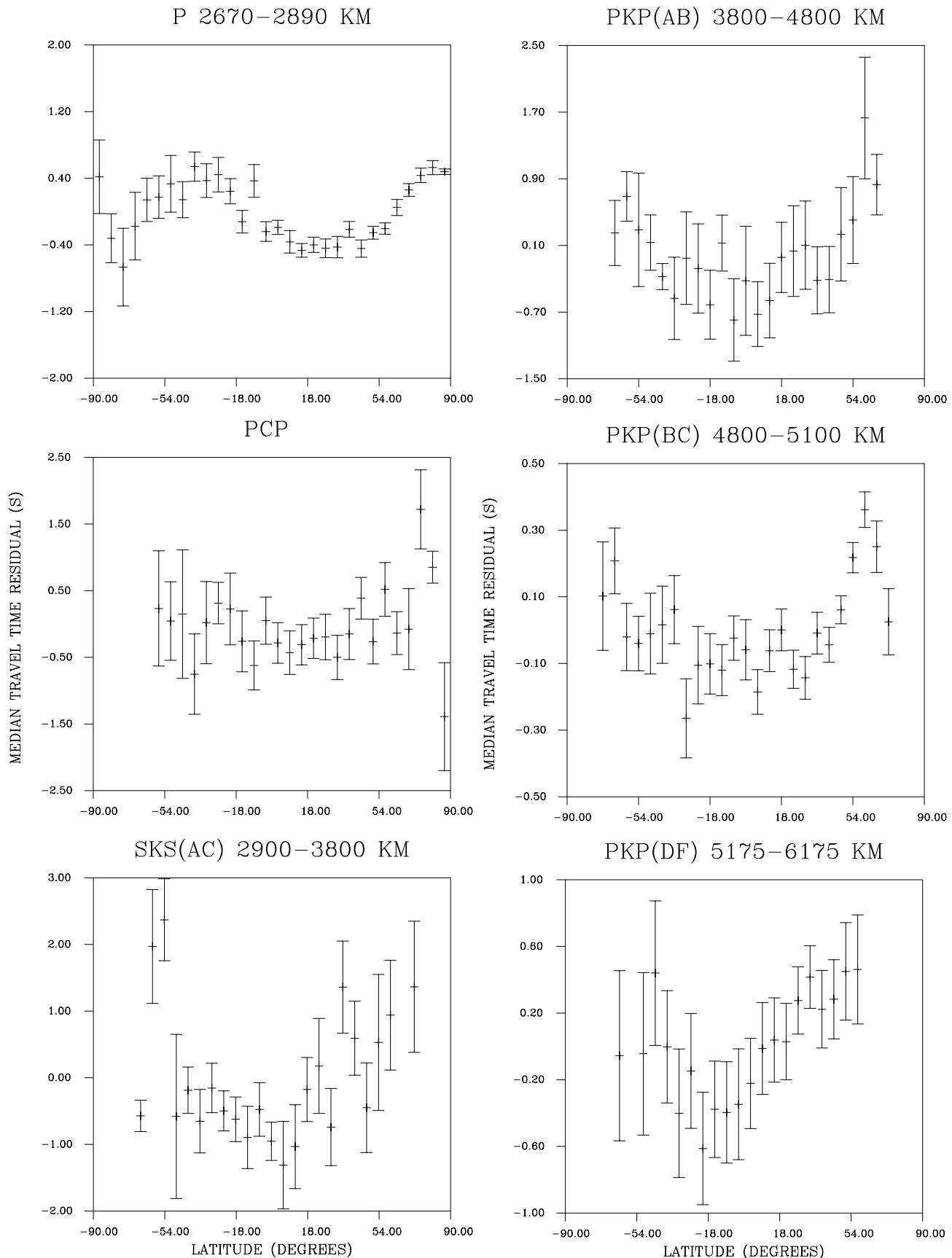
[29] Figure 5 displays a selected portion of the diagonal of the Lanczos resolution matrix, those elements associated with our estimates of compressional velocity. We consider an estimate of compressional velocity to be well resolved if the associated diagonal element is close to unity. Conversely, poorly resolve estimates correspond to diagonal elements which are close to zero. Perhaps a more informative measure of spatial resolution is provided by the resolving width [Jackson, 1979],  $s_k$ , of the  $k$  block

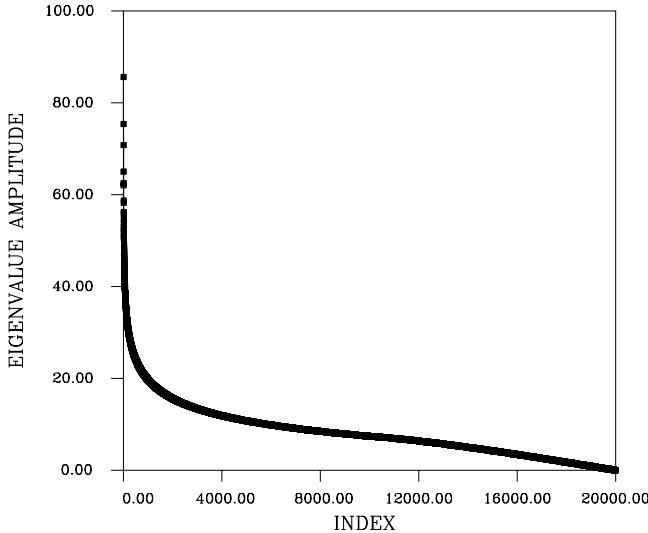
$$s_k = \sum_i \sum_j R_{ki} P_{ij} R_{kj}$$

where  $P_{ij}$  is a matrix that makes  $s_k$  large if  $R_{ki}$  is not close to a delta function. This measure was utilized by *Vasco et al.* [1993] in assessing tomographic compressional velocity estimates for the mantle. Unfortunately, calculating  $s_k$  for all cells in our model is computationally prohibitive.

[30] We should note that the resolution computations are based upon constraints provided by several travel time phases. The errors associated with the various phases are quite different. For example, the errors associated with first arriving P phases are typically much less than errors associated with core phases. Thus, simply adding various phases in the inversion does not guarantee enhanced resolution of Earth structure. We account for variable errors by weighting the rows of the data constraint matrix by the standard errors of the associated data. Thus, phases with large uncertainties will be downweighted with respect to phases with smaller uncertainties.

[31] The overall pattern of the diagonal elements of the resolution matrix is similar to the maps of coverage (Figure

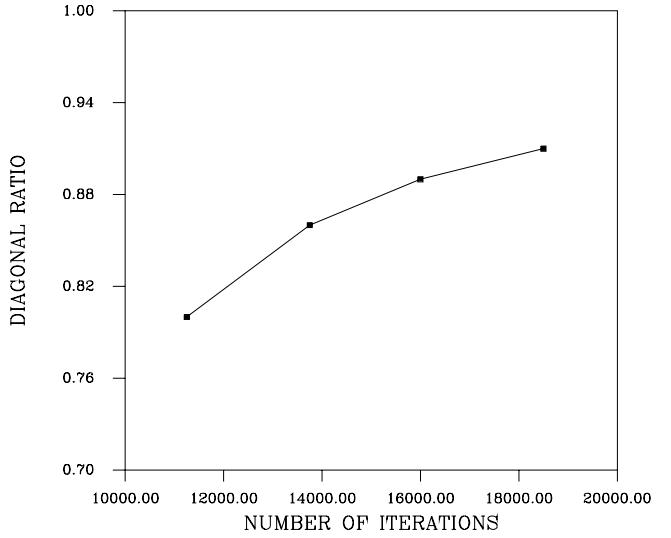




**Figure 3.** Spectral amplitude estimates, based upon 20,000 iterations of the Lanczos algorithm.

1). Thus, there is a clear correlation between ray density and model parameter resolution. One implication of this is that the slabs might appear as narrow, well-resolved features simply because they are sampled by a narrow corridor of rays. The correspondence between ray density and model parameter resolution is not exact however, because of the significant anisotropy in the direction of rays through cells in the model. In some cells, particularly those beneath the central Pacific, the ray directions are highly biased along specific directions. Thus, the cell may have a high ray density but is poorly resolved due to trade-offs along the direction of the ray bundle.

[32] In the uppermost mantle, the Lanczos resolution is greatest beneath the major subduction zones encircling the Pacific and under the continents of the northern hemisphere. The coincidence of seismic events and seismographic stations in the Pacific subduction zones and tectonically active continental regions is primarily responsible for the well-resolved circum-Pacific velocity heterogeneity. Beneath oceanic ridges, another seismically active province, the mantle structure is moderately to poorly resolved. The lack of seismographic stations in the world's ocean basins is primarily responsible for the generally poor Lanczos resolution beneath the Pacific, Atlantic, and Indian ocean basins. Another contributing factor are the relatively few large earthquakes at oceanic ridges. There are many fewer seismographic stations in the southern hemisphere and correspondingly the resolution in the southern hemisphere is poor for the most part with exceptions under Australia, southern Africa, and western South America. The pattern of well-resolved cells beneath the continents of the northern hemisphere and in a broad circum-Pacific swath continues downward to mid and



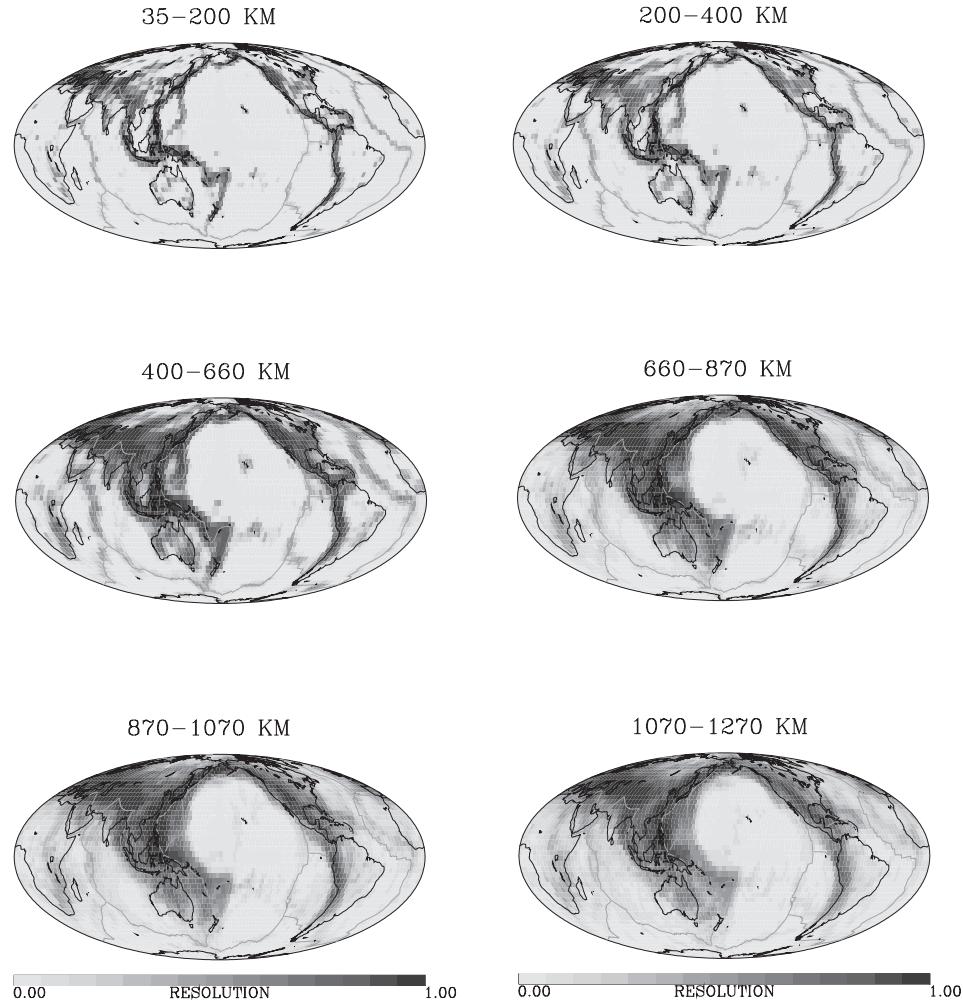
**Figure 4.** Ratio of the diagonal sum of the resolution matrix for successive sets of iterations.

lower mantle depths (Figure 5). As might be expected, the resolution of velocity in our  $3^\circ \times 3^\circ$  representation of the mantle is generally lower than our resolution of compressional velocity in  $6^\circ \times 6^\circ$  cells [Vasco et al., 1999]. There are narrow regions, primarily associated with circum-Pacific subduction zones, where the resolution of velocity heterogeneity is quite high, approaching unity. For the most part, these narrow, highly resolved zones lie above 1470 km.

[33] In the lowermost mantle, below about 1870 km, the peak Lanczos resolution decreases significantly. Such a decrease was also observed in the LU decomposition calculations of Vasco et al. [1993] and in the checkerboard resolution estimates of Vasco and Johnson [1998]. It is most likely due to ray path geometry, mantle P waves bottom in these layers and only core reflected and core transmitted phases pass completely through the layers. In order to investigate this loss of resolution further, we examine individual rows of the resolution matrix. The  $i$ th row of  $\mathbf{R}$  contains the averaging kernel associated with the  $i$ th model parameter. In Figure 6 filled rectangles denote six cells located in the bottom layer of the mantle. For each of cells we computed the corresponding row of the resolution matrix associated with our estimate of compressional velocity (Figures 7a–7f). For cell 46,435 (Figure 7a), located at the base of the mantle just to the northwest of the Azores, we note substantial lateral averaging elongated in a north-easterly to southwesterly direction. There is a contribution from cells in the overlying layer as well. Cell 46,661 (Figure 7b) is situated in a moderately resolved region of the lowest mantle (see Figure 5). Still, there is substantial lateral averaging inherent in our estimate of the compres-

**Figure 2.** (opposite) Bottoming point averages for various phases which interact with the core and mantle. The residual averages are computed by averaging all summary residuals associated with rays bottoming in a particular block. We then compute the residual median based upon all blocks in a particular latitude band in the depth range of interest. The error bar indicates the standard deviation of the mean for a given latitude band.

## Mantle - Compressional



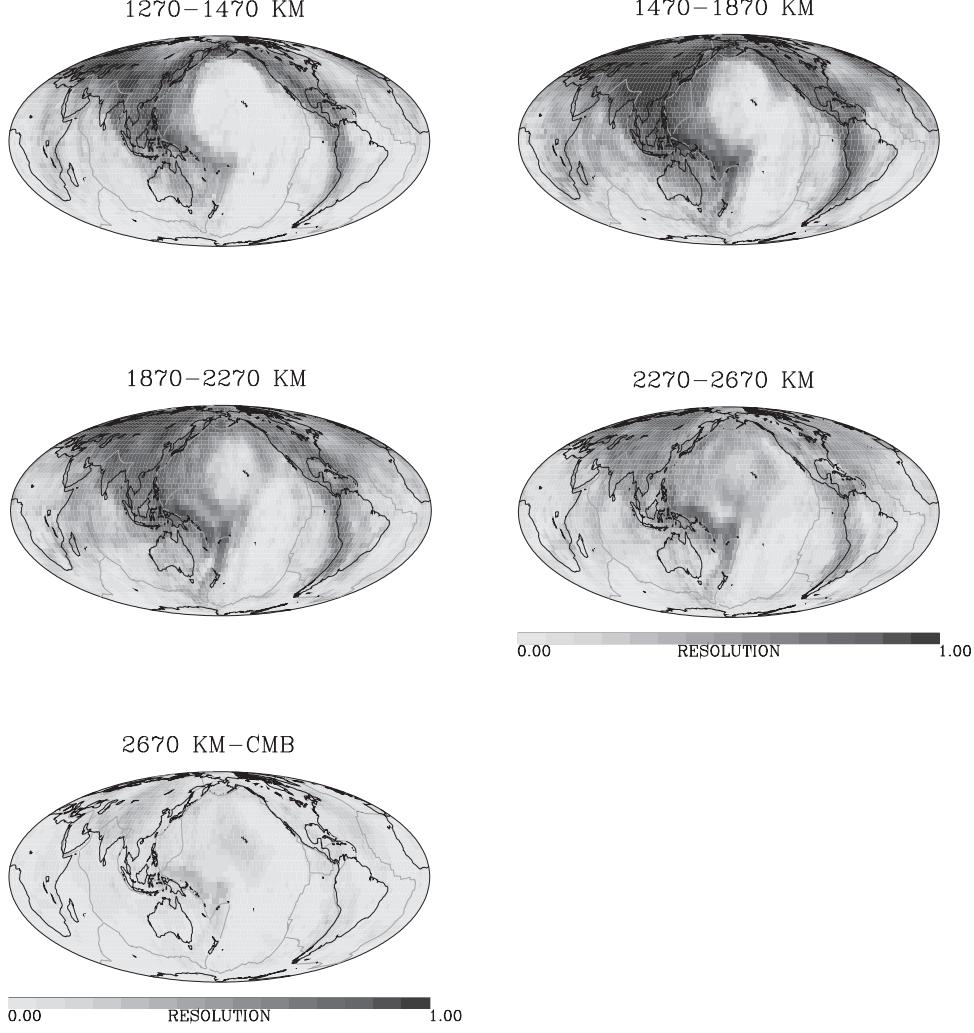
**Figure 5.** Lanczos resolution estimates for 11 layers in the mantle. The resolution variations are associated with estimates of mantle compressional wave velocity.

sional velocity within the block. The averaging is elongated in a southeasterly to northwesterly direction, highlighting the contribution of travel times associated with events from the southwestern Pacific recorded by stations in Europe. Block 47,817 is to the north of the Solomon Islands (Figure 7c), a relatively poorly resolved area at the base of the mantle. From the averaging kernel we observe a substantial contribution from rays trending roughly north-south. There is also a small trade-off with cells in the overlying layer. Substantial lateral and vertical averaging is associated with our compressional velocity estimate for block 47,834 (Figure 7d), located to the north of the Marquesas Islands. The averaging volume is highly elongated in a northeast to southwest direction, spanning several thousand kilometers in extent. The averaging kernel for block 48,353, in the South Atlantic (Figure 7e), highlights contributions from ray paths spanning at least three major orientations. The primary contribution are from north-south trending ray paths with secondary contributions from east-west and northwesterly

ray paths. Cell 48,594, beneath southern Africa (Figure 7f), is also poorly resolved. However, in this case we observe vertical, rather than lateral, averaging. It may be that this cell is primarily constrained by  $PcP$  and core phases. Thus, there may be trade-offs with CMB topography and core structure as well. The nature of these averaging kernels has implications for those attempting to constrain compressional velocity variations in the lowest mantle. For the most part, we can only estimate large scale features in this region. The ray geometry simply precludes imaging small-scale features, even when including a substantial number of compressional phases. Furthermore, obtaining a reliable estimate of the anisotropy of compressional wave velocity in the lowest mantle may prove to be difficult. Because of the preferential sampling along specific directions near the base of the mantle, we cannot constrain the azimuthal variation of compressional velocity.

[34] Preliminary Lanczos resolution estimates indicated that we could not reliably estimate velocity variations for

## Mantle - Compressional



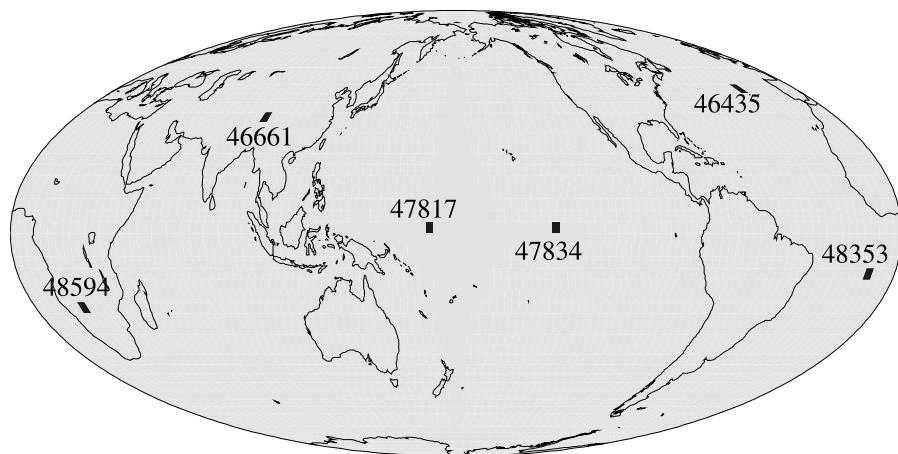
**Figure 5.** (continued)

individual  $3^\circ \times 3^\circ$  cells in either the inner and outer core. In this depth interval, the volume of such cells is so small that their associated velocity cannot be resolved by available seismic arrival time data. The results of Vasco *et al.* [1999] indicated that much of the core is moderately resolvable on the scale of  $6^\circ \times 6^\circ$  cells. In order to obtain well constrained velocity estimates in both regions, we conducted an inversion with two different cell sizes:  $3^\circ \times 3^\circ$  in the mantle and  $6^\circ \times 6^\circ$  in the core. With this parameterization we obtain moderate Lanczos resolution within much of the outer core (Figure 8). If we ignore the narrow features associated with subduction zone source regions, the level of resolution within the outer core is similar to that of the midmantle. Our pattern of resolution within the core is very similar to that presented by Vasco *et al.* [1999]. In particular, regions which are moderately well-resolved in the outer core are correspondingly moderately well-resolved in the midmantle (Figure 5). The similarity reflects the fact that most core phases are traveling at very high angles through the mantle.

Thus, in many respects the pattern in the outer core is a downward mapping of midmantle resolution. There is a notable change in resolution near the base of the outer core, below about 4471 km in-depth, where many  $PKPbc$  rays are turning. For the most part, at the base of the outer core we can no longer resolve velocity variations within individual  $6^\circ \times 6^\circ$  cells. However, as emphasized by the checkerboard tests of Vasco and Johnson [1998], we can still recover anomalies at a scale larger than a single cell. The checkerboard tests of Vasco and Johnson [1998] also indicated some trade-off between velocity heterogeneity at the base of the mantle and outer core structure. This trade-off was also noted by Karason and van der Hilst [2001] in a series of resolution tests in the lower mantle and core. Such a trade-off may explain the somewhat lower resolution of outer core structure. In spite of such a trade-off, many regions of the outer core are moderately well-resolved.

[35] In the inner core Lanczos resolution is poor overall (Figure 9). At least two factors may influence our ability

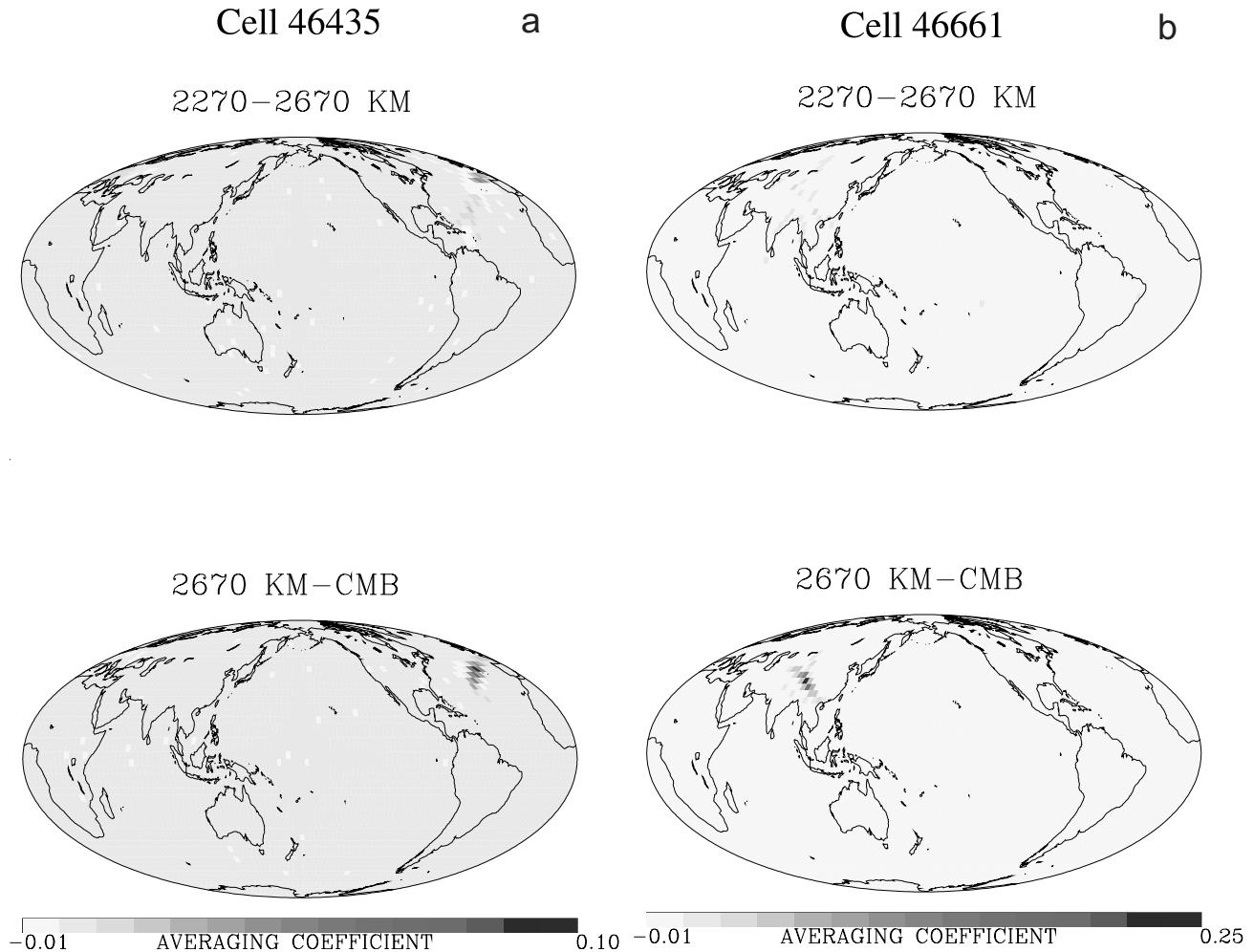
## 2670 KM-CMB



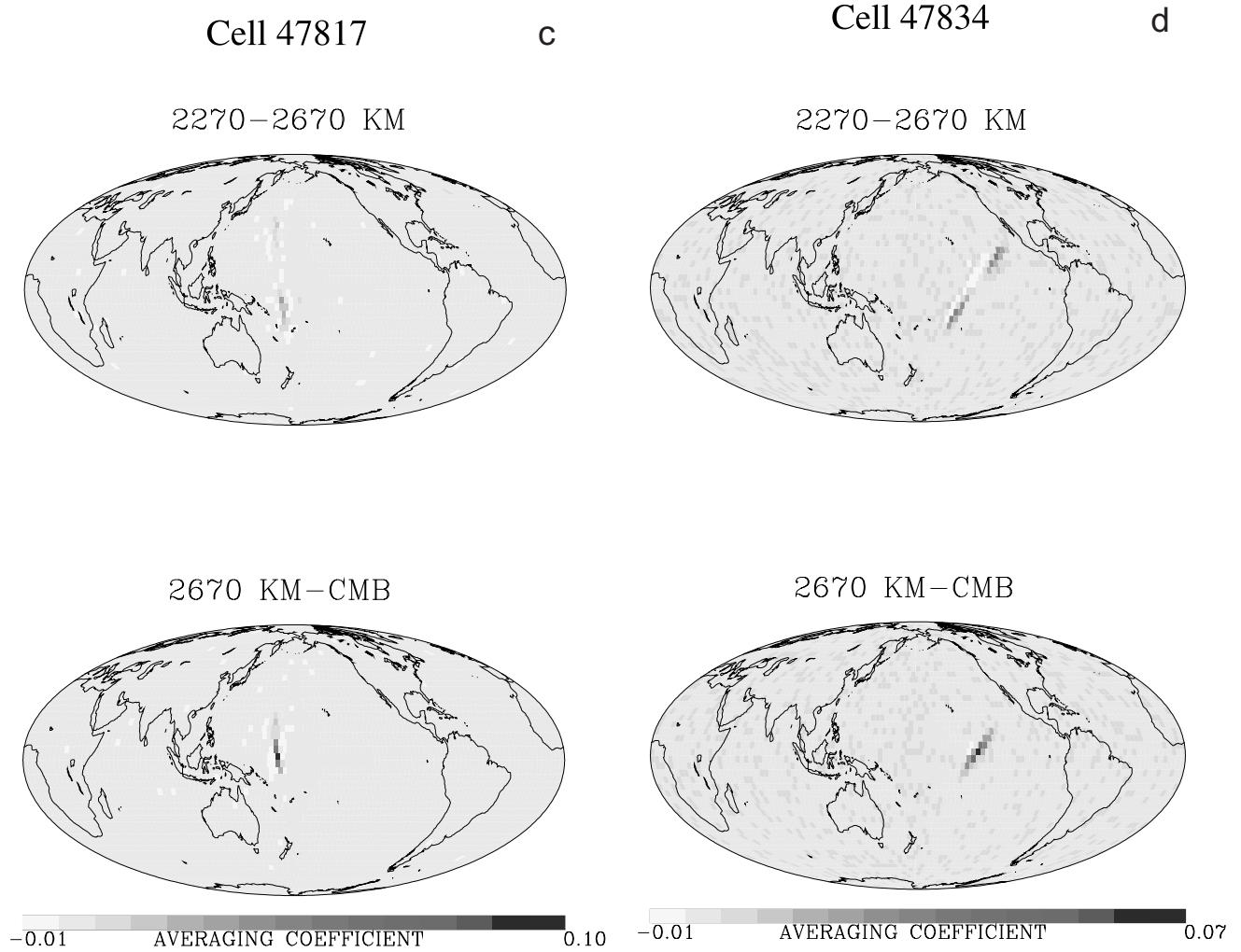
**Figure 6.** Location map indicating cells for which averaging kernels were computed. The filled cells indicate the outline of the block. Adjacent to each cell is the block number.

to recover compressional velocity anomalies here. First, the cell volumes are again fairly small in the inner core, even for  $6^\circ \times 6^\circ$  blocks. Second, in the inner core isotropic compressional velocity heterogeneity may trade-

off with laterally varying anisotropy, further reducing our Lanczos resolution. Our constraints on heterogeneity may be quite poor in the inner core as we allow for spatially varying anisotropy. Poor resolution in the inner core is



**Figure 7.** Averaging kernels for six blocks within the layer just above the core-mantle boundary. The cell locations are given in Figure 6.

**Figure 7.** (continued)

indicated by the large scale checkerboard tests of *Vasco and Johnson* [1998].

### 3.3. Lanczos Covariance Estimates

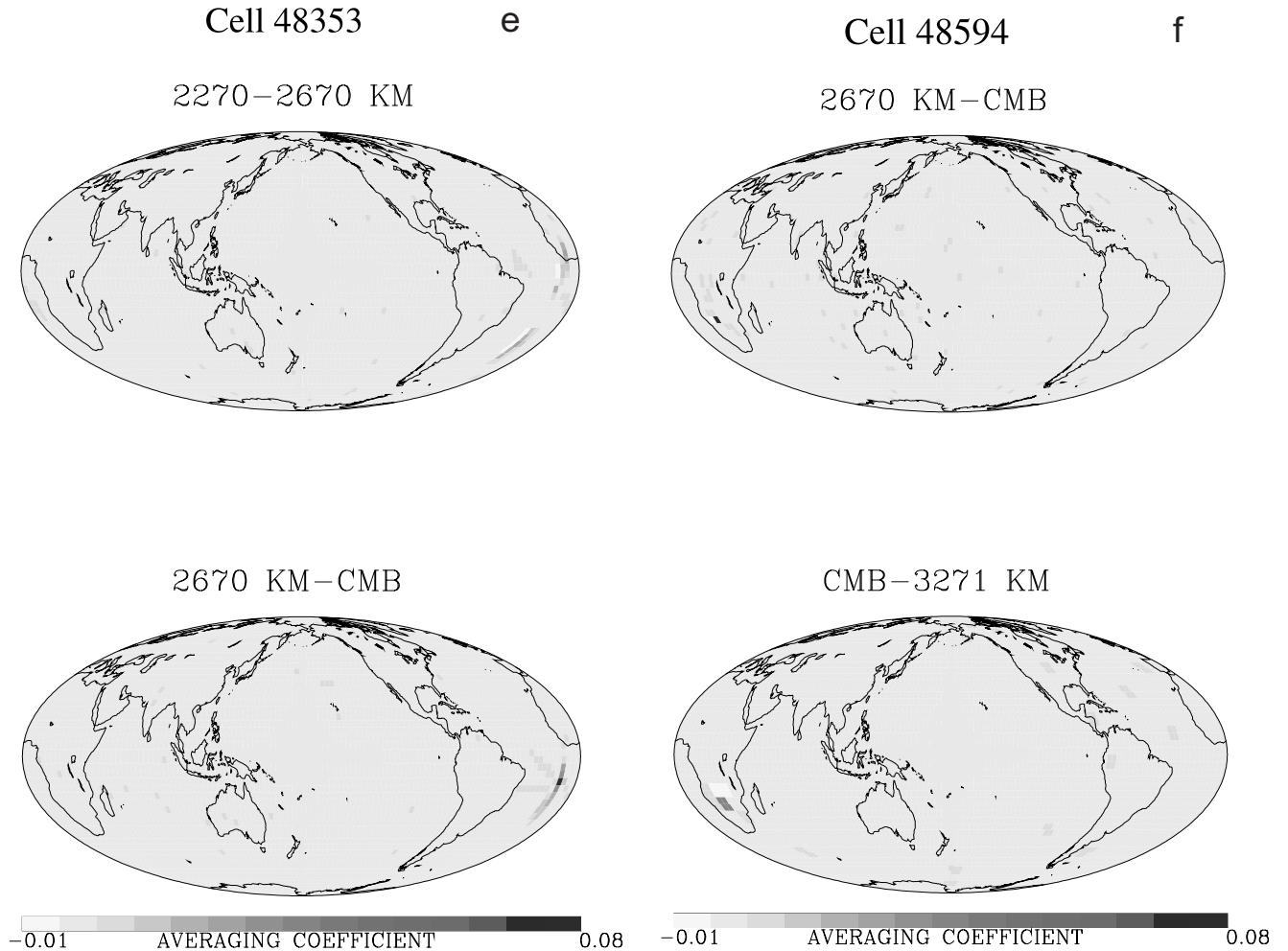
[36] Our covariance estimates are computed using formula (21). This equation holds for data that have been normalized by their standard deviations (see Appendix B). Thus, a necessary first step involves scaling the equations by their sample standard deviations. As discussed by *Vasco et al.* [1999], we estimate the standard deviations for each phase, as a function of epicentral distance. The process of estimating the sample standard deviations is described by *Vasco et al.* [1994, 1995]. The standard deviations associated with compressional wave velocity variations in the mantle are shown in Figure 10. Note, in order to facilitate comparisons with resolution, the scale has been reversed in this and in subsequent plots of standard deviations. In Figures 10, 11, and 12, darker shades indicate smaller values. In general, regions of low resolution correlate with larger standard deviations. This correspondence is to be expected because the posterior covariance matrix is given by

$$\mathbf{C}_m = (\mathbf{I} - \mathbf{R})\mathbf{C}_p$$

where  $\mathbf{C}_p$  is the prior covariance matrix [*Tarantola*, 1987]. This association between model parameter resolution and

uncertainty contrasts with model parameter error estimates calculated from inversions of random deviates. In such inversions, the lowest noise is found in regions with the poorest coverage. This counter-intuitive result points to the fact that the mapping of random deviations cannot be interpreted as model parameter uncertainty. This problem with the mapping of noise into model parameter errors has been pointed out by *Vasco et al.* [1993, 1999].

[37] Overall, standard deviations in the mantle are somewhat larger for this finer parameterization, which is based upon  $3^\circ \times 3^\circ$  cells in the mantle. In well constrained regions the associated standard deviation is of the order of 0.2% (Figure 10). This is roughly twice the standard deviation computed for velocity estimates in a  $6^\circ \times 6^\circ$  mantle model [*Vasco et al.*, 1999]. In poorly constrained cells the estimated errors exceed several percent. As noted above, there is a good correspondence between well-resolved regions and areas with the lowest standard deviations. In particular, low standard deviations are associated with the volume of the mantle lying beneath the continents of the northern hemisphere. In addition, we find a ring-like pattern of lower standard deviations, surrounding the mantle beneath the Pacific Ocean basin (Figure 10). Unlike model parameter resolution, which decreases markedly at the base of the mantle, the standard deviations



**Figure 7.** (continued)

in well-sampled portions of the deepest mantle remain moderate, around 0.2 to 0.3%. Even so, much of the lowest mantle is poorly sampled and the standard deviations exceed 0.5%. This is particularly true in the southern hemisphere of our model.

[38] In the outer core the distribution of model parameter standard deviation is similar to that in the mantle, both in pattern and magnitude (Figure 11). In well constrained regions the standard deviation varies between 0.15 and 0.25%. In poorly resolved blocks the uncertainties exceed 0.5%. For example, in the outer core below the mid-Pacific the standard deviations may be of the order of several percent. In the uppermost outer core the smallest standard deviations are associated with blocks lying beneath the major seismic source regions, the subduction zones encircling the Pacific basin. In general, this pattern persists throughout the outer core though, with depth, it becomes somewhat diffuse. This is due to the bottoming of core refracted phases, such as  $PKPab$  and  $PKPbc$ . At the base of the outer core standard deviations of approximately 0.25% are associated with blocks in the northern hemisphere of our model. Uncertainties in the southernmost portion of this layer (4871 km-ICB) approach 0.5%.

[39] The model parameter uncertainty increases dramatically with depth in the inner core (Figure 12). The standard

deviation exceeds 0.25% for all cells within the inner core. In the bottom two layers (5671–5971 km and 5971–6371 km) most standard deviations are greater than 0.5%. As we shall see in the next section, such uncertainties are larger than the estimated inner core compressional velocity variations in our model.

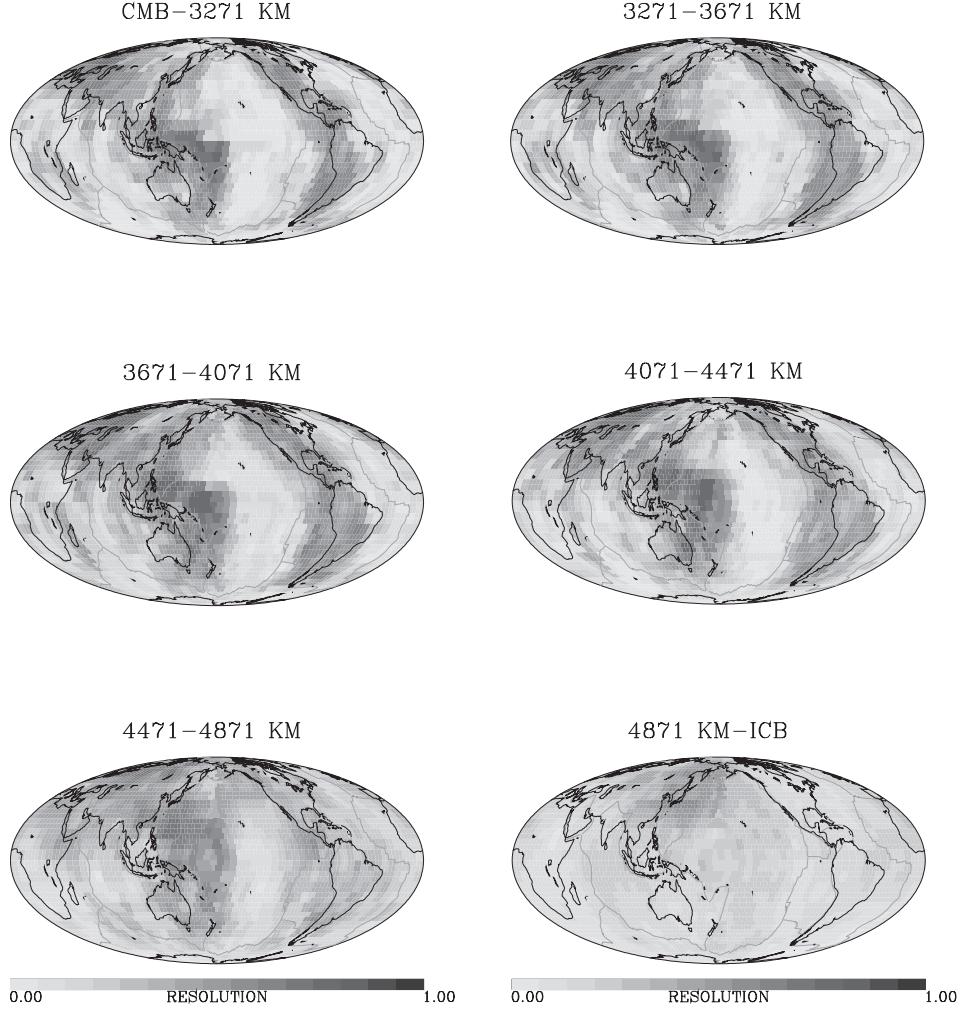
### 3.4. Whole Earth Structure

[40] Estimates of compressional velocity variations within the mantle, outer core, and inner core are presented in Figures 13, 14, and 15. In these figures we have utilized our Lanczos covariance estimates in order to delete anomalies in poorly constrained regions of our model. Specifically, blocks in which the model parameter uncertainties exceeded the velocity anomaly magnitudes are set to zero. This resolution-weighted approach to presenting global tomographic images was used previously by Pulliam *et al.* [1993] and Vasco *et al.* [1993].

#### 3.4.1. Compressional Velocities Within the Mantle

[41] In the depth range 35–200 km (Figure 13) the  $P$  velocity heterogeneity is dominated by surface tectonic elements: the continental cratons, subduction zones, and associated back arc basins, with peak anomalies exceeding 2%. Several back arc basin low-velocity anomalies extend to the depth interval 200–400 km. Also, many

## Outer Core



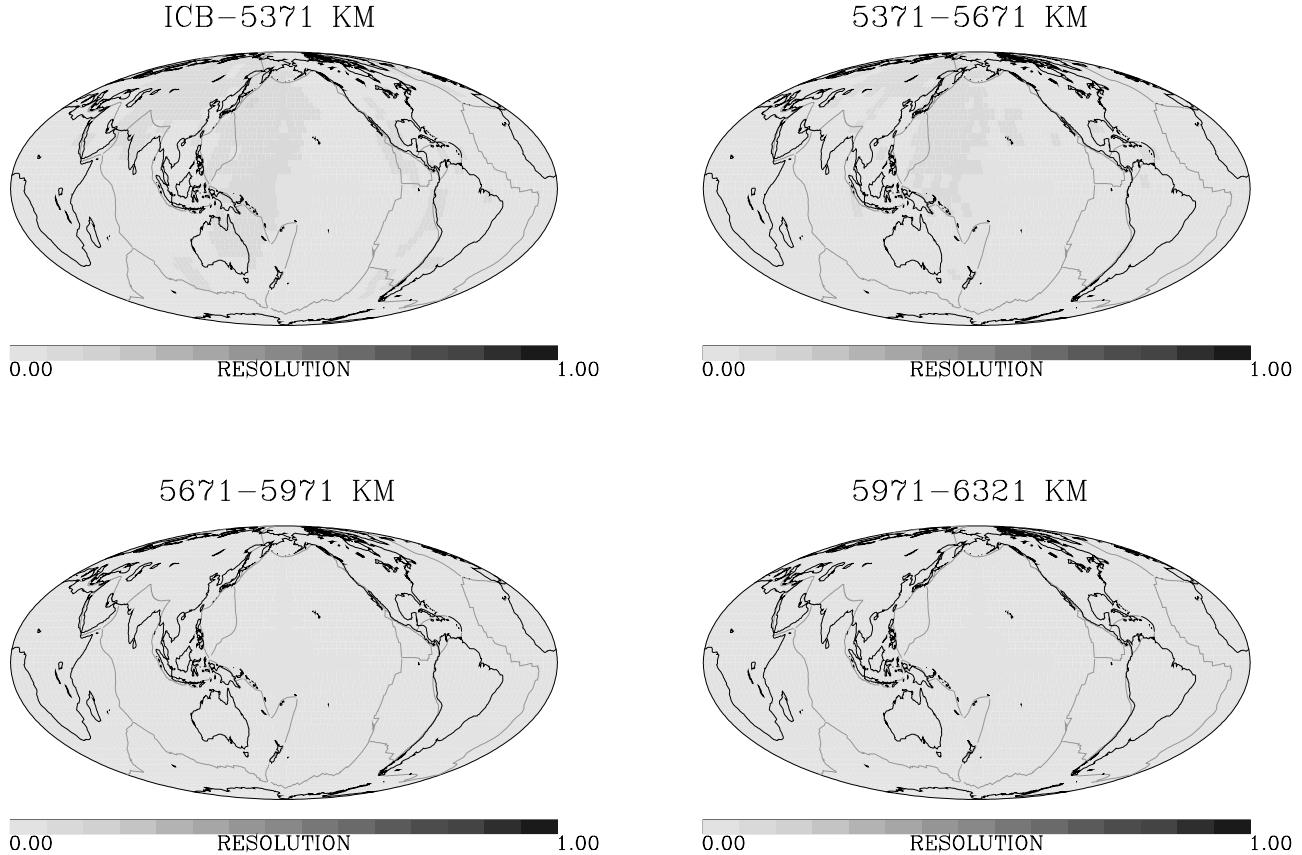
**Figure 8.** Diagonal elements of the resolution matrix associated with compressional velocity heterogeneity in the Earth's outer core.

continental cratons are still notable between 200 and 400 km, though as much smaller anomalies. The subduction zones are imaged as narrow, high-velocity features extending throughout the upper and lower mantle. As in previous studies [Fukao *et al.*, 1992; *van der Hilst et al.*, 1997; *Bijwaard et al.*, 1998], the high-velocity slabs are somewhat better defined in the current  $3^\circ \times 3^\circ$  model than in our previous  $6^\circ \times 6^\circ$  model [Vasco and Johnson, 1998]. We should note however, that the high-velocity slab-like features often coincide with narrow zones of high resolution in the upper mantle (Figure 5). It is not clear how such sharp variations in resolution in the upper mantle influence the pattern of velocity heterogeneity.

[42] Below the transition zone, a nearly continuous narrow ribbon of fast velocity encircles much of the mantle below the Pacific Ocean basin, between 660 and 870 km in-depth. This arc of high velocity extends intermittently down to a depth of about 1500 km. *van der Hilst et al.* [1997] have suggested that a few of these

anomalies are continuous down to the core mantle boundary. Between 660 and 1470 km a trend of higher velocities extends from India to New Guinea (Figure 13). As noted in other mantle tomography models [Romanowicz, 1991], the amplitude of lateral heterogeneity increases in the lowest mantle, and a high-velocity ring underlies the Pacific Ocean basin. Relative to the anomalies in the mid to upper mantle, the features in the lowermost mantle are more smoothly varying. It has been suggested that there is a change in the character of compressional velocity heterogeneity somewhere between 1500 and 1800 km in-depth [*van der Hilst and Karáson*, 1999]. Above this transition, the anomalies associated with subducting slabs are sharp and well-defined. Below 1800 km the high-velocity ring around the Pacific basin is more diffuse and of a much longer wavelength. While such a transition may indeed exist, it should be pointed out that our resolution of Earth structure changes significantly in this depth range (Figure 5). That is, above 1470 km the best resolution

## Inner Core



**Figure 9.** Diagonal elements of the resolution matrix associated with compressional velocity heterogeneity in the Earth's inner core.

is concentrated in fairly narrow zones surrounding the Pacific. Below 1870 km in-depth, the resolution is lower overall and more smoothly varying. As indicated by the averaging kernels associated with the lowest mantle (Figures 7a–7f), there is significant lateral and depth averaging involved in constructing our velocity estimates. Thus, more work is required before we can definitively conclude that there is indeed a change in the nature of mantle heterogeneity somewhere between 1500 and 1800 km in-depth.

[43] The major features of the mantle model, shown in Figure 13, are contained in our previous  $6^\circ \times 6^\circ$  compressional velocity model [Vasco and Johnson, 1998]. We observe somewhat more detailed variations in the current model. But these details should be weighed against the large variations in model parameter resolution in the upper mantle. The major features of our  $P$  velocity model also agree with previous studies such as *van der Hilst et al.* [1997] and *Bijwaard et al.* [1998]. That is, as in those studies we observe high-velocity anomalies associated with subduction zones which extend to depths greater than 1200 km. Furthermore, low velocities are prominent in the back arc basins and tectonically active regions in the upper 200 km. Finally, the pattern of heterogeneity in the lowermost

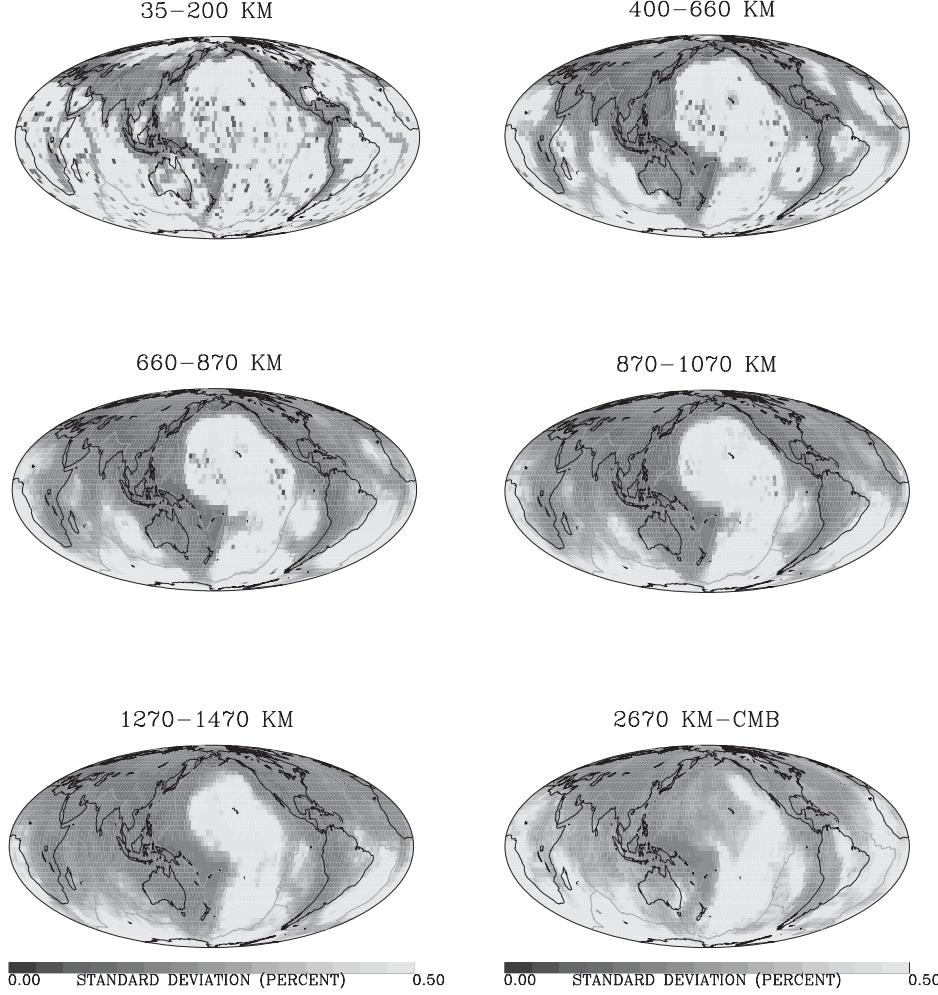
mantle is similar to that of *Bijwaard et al.* [1998] though, as noted above, the lateral resolution is poor there.

### 3.5. Compressional Velocities Within the Core

[44] The velocity variations in the outermost core are smaller in amplitude than our mantle variations: the RMS and peak variations are about 30% of those in the mantle (Figure 14). However, the size of the heterogeneity increases as the inner core is approached. The depth distribution of velocity heterogeneity in our model may be influenced by a number of factors. Different phases constrain various depth intervals of the outer core. In general, our depth resolution is poor in the core, and it is difficult to distinguish between heterogeneity in adjacent layers. The topmost and middle parts of the outer core are primarily constrained by *SKSac* and *PKPab* phases which are more emergent and difficult to pick, while the lowermost outer core is where the larger-amplitude *PKPbc* rays turn. The topmost core structure trades off with CMB topography and overlying mantle structure which are constrained by significantly more data.

[45] In our model of outer core heterogeneity there is a significant shift toward higher compressional velocity overall. That is, we note a significant positive mean shift

## Mantle - Compressional



**Figure 10.** Model parameter standard deviations corresponding to compressional velocity estimates for six layers in the mantle. Note that the scale is inverted in this figure such that darker shades indicate smaller standard deviations.

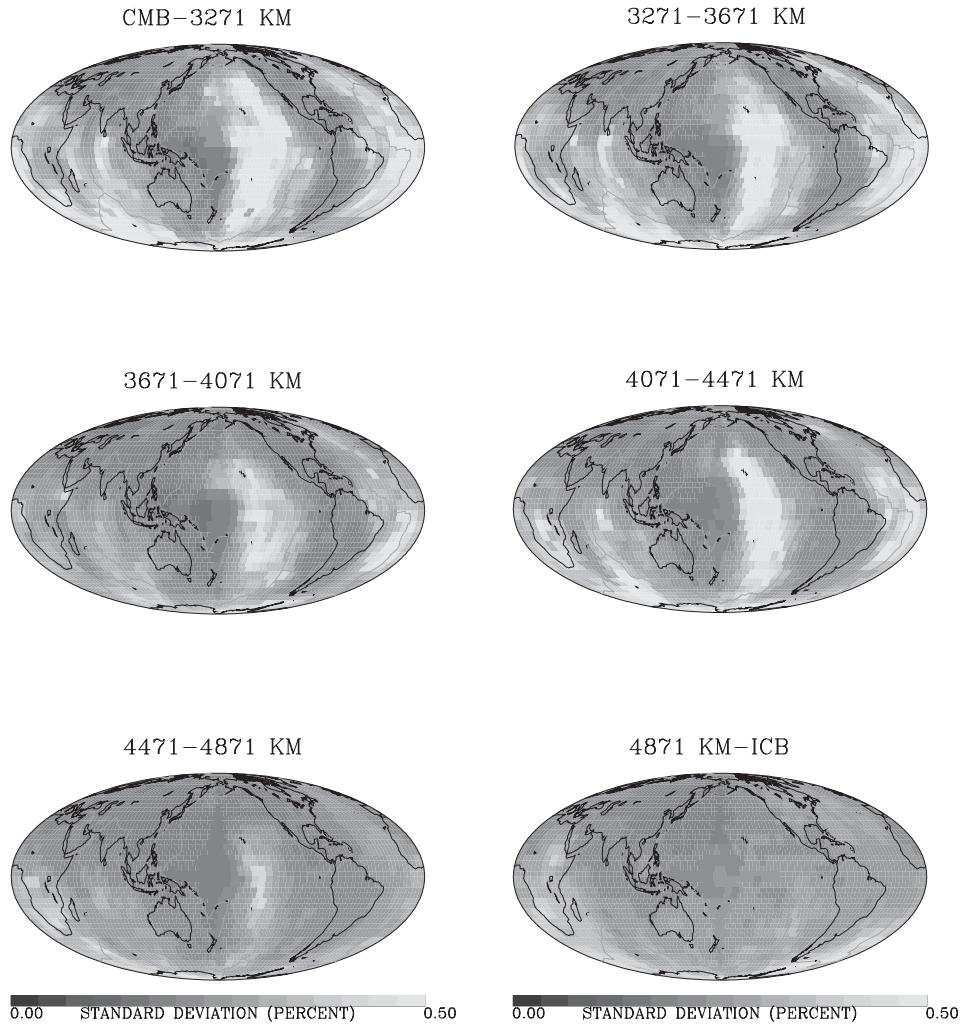
in our velocity perturbations, particularly at mid to low latitudes. This shift in average velocity is suggested by our residuals plots in Figure 2. On average, the travel time residuals are negative at the mid to low latitudes, suggesting that the velocity in our starting model is too low. In previous work [Vasco and Johnson, 1998; Vasco *et al.*, 1999] we removed the mean shift before plotting the results. However, because the mean shift is a significant component of our outer core velocity model we have included it in Figure 14. Another large scale feature at the base of the outer core is a rough symmetry in the velocity perturbations. In particular, lower velocities are concentrated near the poles and higher velocities are observed at the lower latitudes (Figure 14). This symmetry is also indicated by the variation of bottoming point residuals as a function of latitude (Figure 2). We observe negative residuals for rays which bottom at low latitudes and positive residuals for rays bottoming near the poles. Unfortunately, our resolution of the southernmost portion

of the base of the outer core is rather poor. More data is required to better constrain the heterogeneity in this region. As is clear from our Lanczos resolution and covariance estimates (Figures 9 and 12), inner core structure is poorly determined at the scale of our model ( $6^\circ \times 6^\circ$ ). Thus, most of the anomalies in blocks within the inner core have been set to zero.

## 4. Discussion and Conclusions

[46] Imaging the three-dimensional structure of the Earth has been an area of active research for over twenty years. We are now able to construct images of the entire planet from crust to core. In the past there have been major improvements in methodology and data quantity and quality. These improvements are likely to continue in the future. In particular, utilization of the complete seismic waveform and broadband data promise to better constrain Earth structure. The maps of velocity produced by seis-

## Outer Core

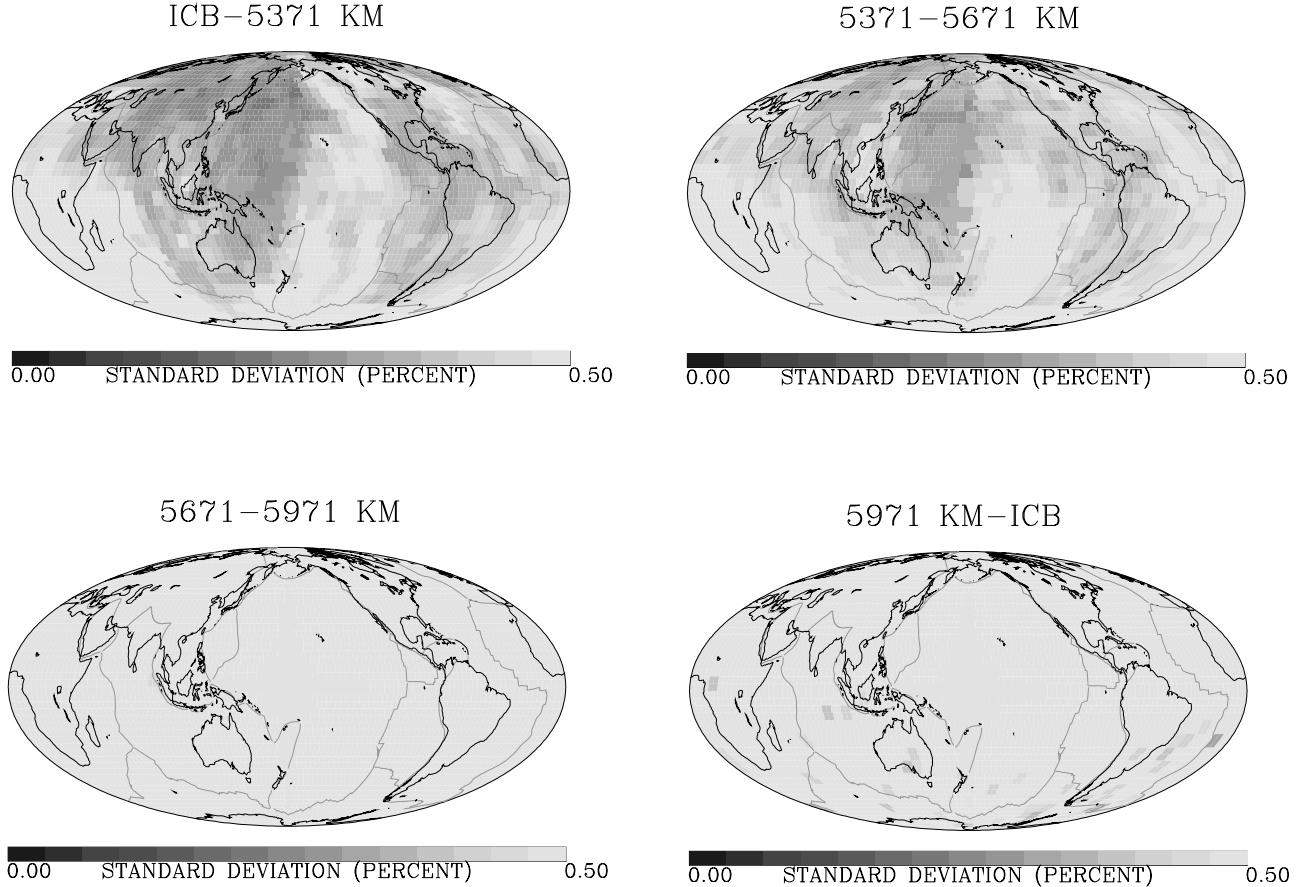


**Figure 11.** Model parameter standard deviations corresponding to compressional velocity estimates in the outer core.

mologists have many interesting features and are subject to geological interpretation. However, some caution is warranted before these results are taken as fact. The images themselves are based upon a highly biased collection of sources and receivers resulting in quite variable resolution. This points to the need for a rather complete model assessment in addition to estimates of velocity heterogeneity. In our Lanczos resolution computations we find that, at the scale of our  $3^\circ \times 3^\circ$  model, well-resolved cells are confined to narrow zones underlying Eurasia, North America, and circum-Pacific arc-trench regions. Thus, there is a notable degradation from our resolution of heterogeneity within larger  $6^\circ \times 6^\circ$  cells [Vasco *et al.*, 1999]. As we approach the core–mantle boundary, resolution decreases markedly and we observed significant lateral and vertical averaging associated with our estimates of compressional velocity (Figures 7a–7f). Such variations in resolution must be accounted for when interpreting our estimates of velocity heterogeneity in terms of mantle processes and dynamics.

[47] As noted in the introduction, there are now a wealth of Earth models. It is worthwhile to compare the various estimates of heterogeneity within the Earth. Work in this direction has already been initiated. However, there are some issues that should be addressed in examining how well, or how poorly, a set of models are correlated. First, the various inversions for Earth structure may be based upon very diverse sets of data. The situation can become somewhat complicated in that many models are based upon quite similar data sets, such as ISC derived first arrival time data, while other sets of models are based upon other observations such as surface wave data and long period travel times. This suggests that the resulting resolution of Earth structure can be very different from one model to the next. Furthermore, model parameterization may be quite different. For example, an inversion may account for mantle anisotropy or boundary topography in addition to lateral variations in the isotropic component of velocity. Differences in parameterization may introduce different trade-offs between sets of param-

## Inner Core



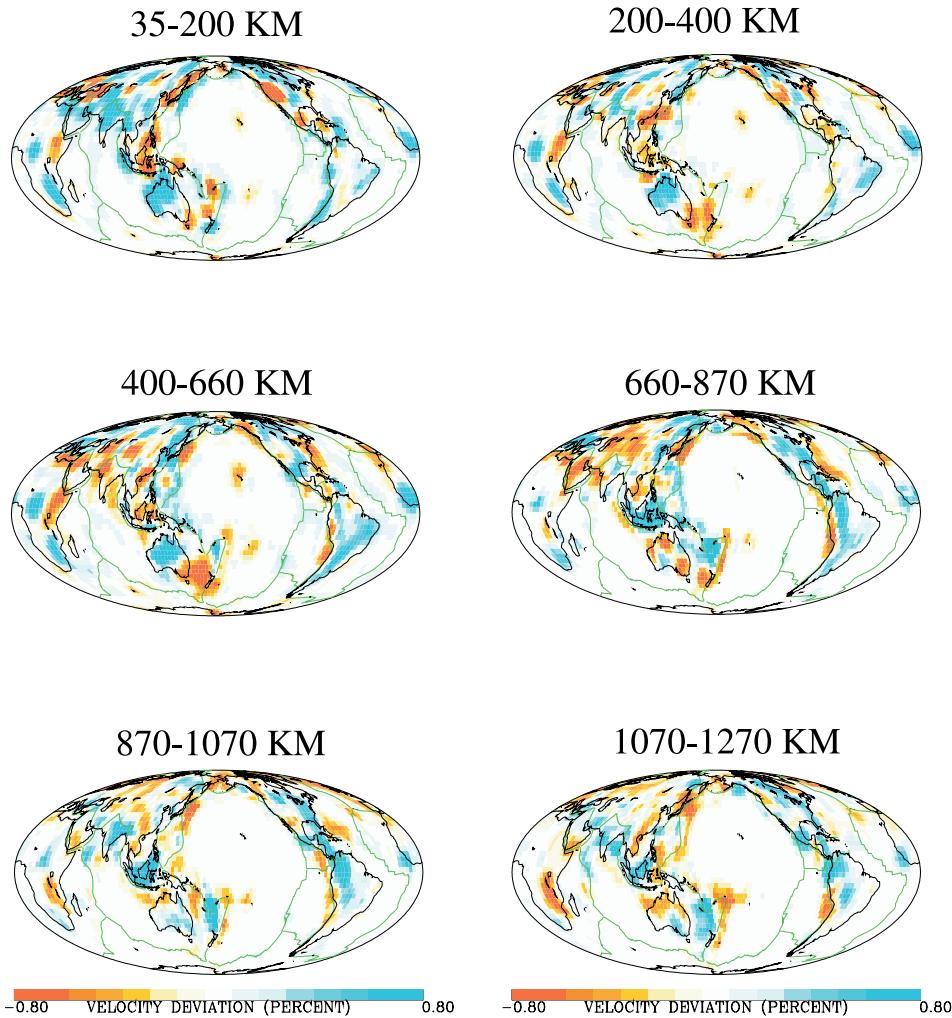
**Figure 12.** Model parameter standard deviations corresponding to compressional velocity estimates in the inner core.

eters. This will also influence our resolution of velocity heterogeneity within the Earth. Finally, differences in the regularization used to stabilize the inverse problem can significantly influence model parameter estimates, particularly in poorly constrained volumes within the model. In such regions, where the data constraints are weak, the structure can be dominated by the details of the regularization. For example, an emphasis on damping might result in very small perturbations while giving greater weight to a roughness penalty could result in significant though smoothly varying heterogeneity. All of these factors suggest that, at the very least, we should account for the variable resolution of velocity heterogeneity when comparing Earth models. In doing so, we can take advantage of our explicit calculation of Lanczos resolution. The underlying idea is that we should only compare well-resolved structure in our models. Thus, before correlating Earth models, we remove velocity estimates associated with cells which are poorly constrained.

[48] We have begun to incorporate such resolution filtering in the correlation of Earth models. For example, in Figure 16 we compare  $P$  variations from our model (WE2000) to  $P$  velocity variations derived from the

Scripps shear and bulk sound speed model SB10L18 [Masters et al., 2000], a  $10^\circ \times 10^\circ$  block-based representation. Because we do not have formal resolution estimates for model SB10L18, the resolution estimates shown in Figure 5 were used to eliminate cells which are poorly determined in our model. Ideally, resolution estimates for both models would be used to filter out poorly resolved anomalies in each. The map of resolution for the depth range 1070–1270 km is shown in Figure 16. A cutoff of 0.6 was used to eliminate cells which are not resolved in our model. Similarly, cells in model SB10L108 which are, on average, not well resolved are not used in the correlation. The original estimates and the filtered estimates are shown in Figure 16. It is clear from Figure 16 that significant anomalies in SB10L108 lie in poorly resolved regions of our model, particularly under the ocean basins. It should be noted that the anomalies may be resolved by the data used in constructing SB10L108. The important point is that they lie in a volume of the mantle which is poorly resolved in our model. Thus, we should not include such regions in a correlation of the two models. The Spearman rank-order correlation coefficient was computed for each depth range,

## Mantle - Compressional



**Figure 13.** Estimates of compressional velocity deviations within the mantle. The deviations are percent variations from the background radial  $P$  model ak135.

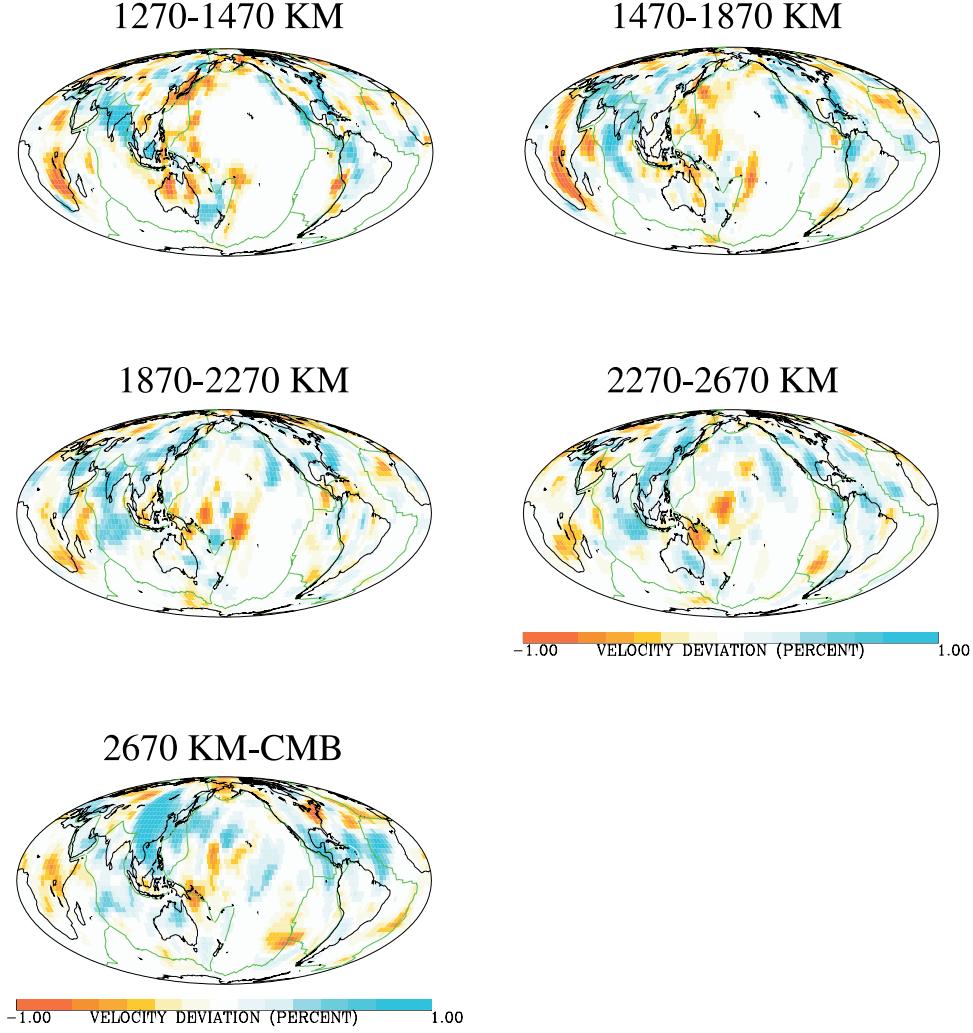
a nonparametric correlation coefficient based upon the rank ordering of anomalies [Lehmann, 1975]. The results are shown in Figure 17 for layers in the lower mantle where the resolution of the two models is likely to be similar. In the upper mantle the resolution associated with the two models is likely to be extremely different because surface wave phase velocities are utilized in the construction of SB10L108. In comparing the correlations between  $P$  velocity variations in both models we note a significant difference if poorly resolve cells are deleted. The correlation coefficient is uniformly higher, and can change by over 100% in many depth intervals.

[49] It is important to keep model parameter resolution in mind when correlating or comparing models. Some of the difficulties of correlating models in the presence of nonuniform sampling were noted by Pulliam and Stark [1993]. In poorly resolved regions of the Earth regularization can dominate an inversion result. Thus, if many investigators are using only a smoothing penalty and no

damping, for example, they may all produce the same anomaly in a poorly resolved region. That does not mean that the heterogeneity in that region is either smoothly varying or of that particular pattern. It simply means that, under the assumptions implicit in the use of the regularization, the smooth anomaly is compatible with the observations. However, many other patterns of heterogeneity may still be possible. The variations in our resolution of Earth structure highlights the need for continual improvement in the quality and distribution of seismographic stations. In particular, it is necessary to incorporate as many sources of information as possible such as body waves, free oscillations, and surface waves.

[50] Our Earth model is based upon an inversion of body wave arrival times. This work should be viewed as a first step in the estimation and assessment of whole Earth structure. In the future we hope to improve our resolution of Earth structure by incorporating surface wave and free-oscillation data. It is of particular interest

## Mantle - Compressional



**Figure 13.** (continued)

to examine how the resolution changes as each new class of data is included in the inverse problem. We can use the methodology developed here to examine several possible sources of nonuniqueness and trade-off. In particular, because we solve for topography at the core–mantle boundary, we can quantify the trade-off between core–mantle boundary topography and velocity structure. Similarly, we can estimate the trade-off between isotropic velocity heterogeneity and anisotropy. We can also estimate the uncertainty associated with ratios of  $S$  to  $P$  velocity heterogeneity within the mantle. Finally, all the results in this paper are based upon a linearization about the radial model ak135. That is, we are not conducting an iterative nonlinear inversion in which rays are recomputed at each iteration. Such an approach might prove useful, though initial indications [Bijwaard and Spakman, 2000] are that successive iterations do not drastically change the Earth model.

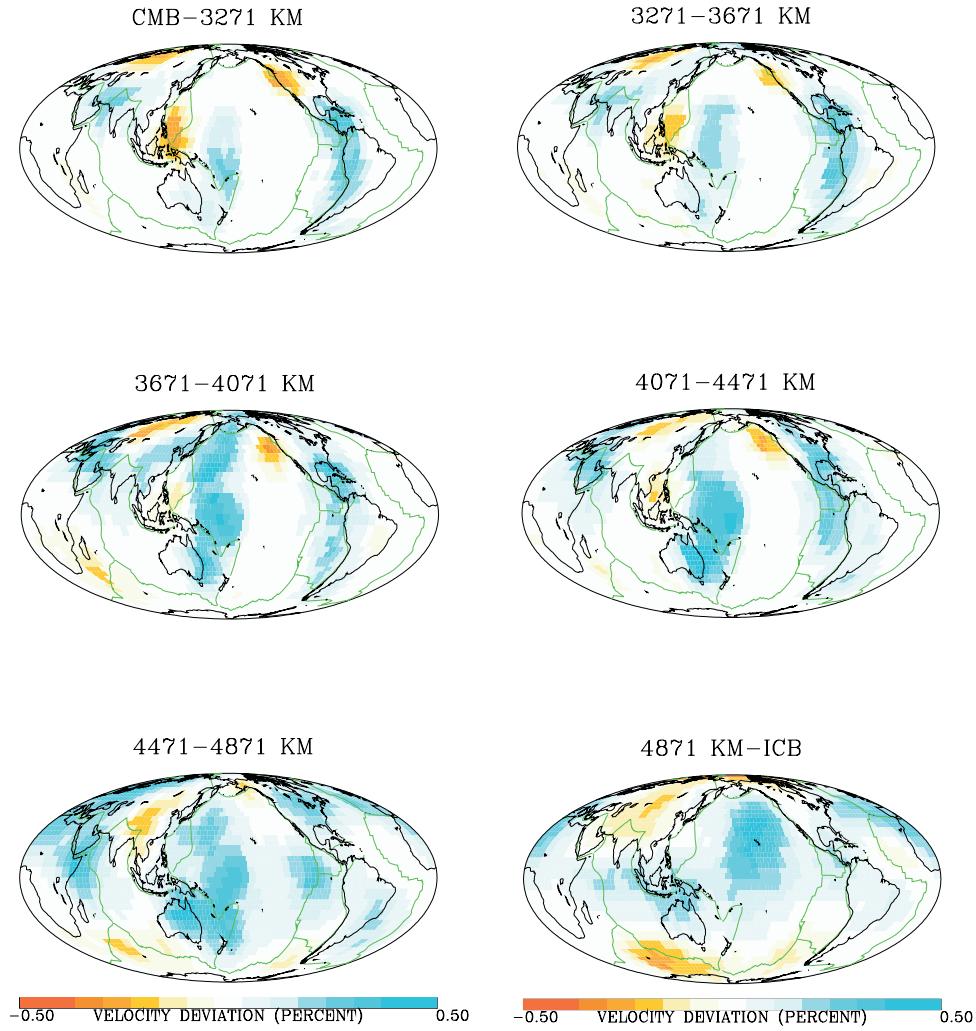
[51] We end this paper with a discussion on the computational costs of our approach. We have implemented the Lanczos algorithm on both a scalar workstation and a

massively parallel T3E computer. The results discussed here were produced by a run on 16 nodes of the T3E computer which took of the order of twenty hours. The main constraint in tackling large problems is computer memory. If the entire set of nonzeros in the coefficient matrix can be held in core memory the performance of the algorithm improves by over an order of magnitude. It is also helpful during reorthogonalizations to store the entire set of vectors  $\mathbf{q}_i$  in core. The algorithm, as implemented here, computed the columns of the matrix  $\mathbf{V}$  and thus scales as the number of model parameters. In addition, as more vectors  $\mathbf{q}_i$  are generated, the algorithm requires more storage and computation. Still, computer memory has increased dramatically in the past few years as the price per megabyte has fallen. Thus, it is entirely feasible to attack large problems using clusters, workstations, or even single PCs.

## Appendix A

[52] In this appendix we give the exact form of the matrices used in our inversion and assessment. The data

## Outer Core



**Figure 14.** Estimates of compressional velocity deviations within the outer core.

constraints are provided by the system of equations (4). Combining the data terms and the regularization produces a penalized misfit function, the  $\ell^p$  norm of the data misfit and penalty terms [Vasco *et al.*, 1994, 1995; Vasco and Johnson, 1998]:

$$\begin{aligned} P(\delta\mathbf{v}, \delta\mathbf{h}, \delta\mathbf{s}, \delta\mathbf{b}, \delta\mathbf{a}) = & |\mathbf{W}_D \mathbf{r}|_p + |\mathbf{W}_R^v \mathbf{D} \delta\mathbf{v}|_p + |\mathbf{W}_A^v \delta\mathbf{v}|_p + |\mathbf{W}_R^b \mathbf{D} \delta\mathbf{b}|_p \\ & + |\mathbf{W}_A^b \delta\mathbf{b}|_p + |\mathbf{W}_R^a \mathbf{D} \delta\mathbf{a}|_p + |\mathbf{W}_A^a \delta\mathbf{a}|_p \end{aligned} \quad (\text{A1})$$

where  $\mathbf{W}_D$  is a diagonal matrix containing the weights for the data components,  $\mathbf{W}_R^v$  and  $\mathbf{W}_A^v$  are the weight matrices for the velocity roughness and norm components,  $\mathbf{W}_R^b$  and  $\mathbf{W}_A^b$  are the weight matrices for the boundary roughness and norm components, and  $\mathbf{W}_R^a$  and  $\mathbf{W}_A^a$  are the weight matrices for the anisotropy roughness and norm components. The matrix  $\mathbf{D}$  is a differencing operator which is used in the computation of lateral roughness. For our application, we weight the data rows by the inverse of the

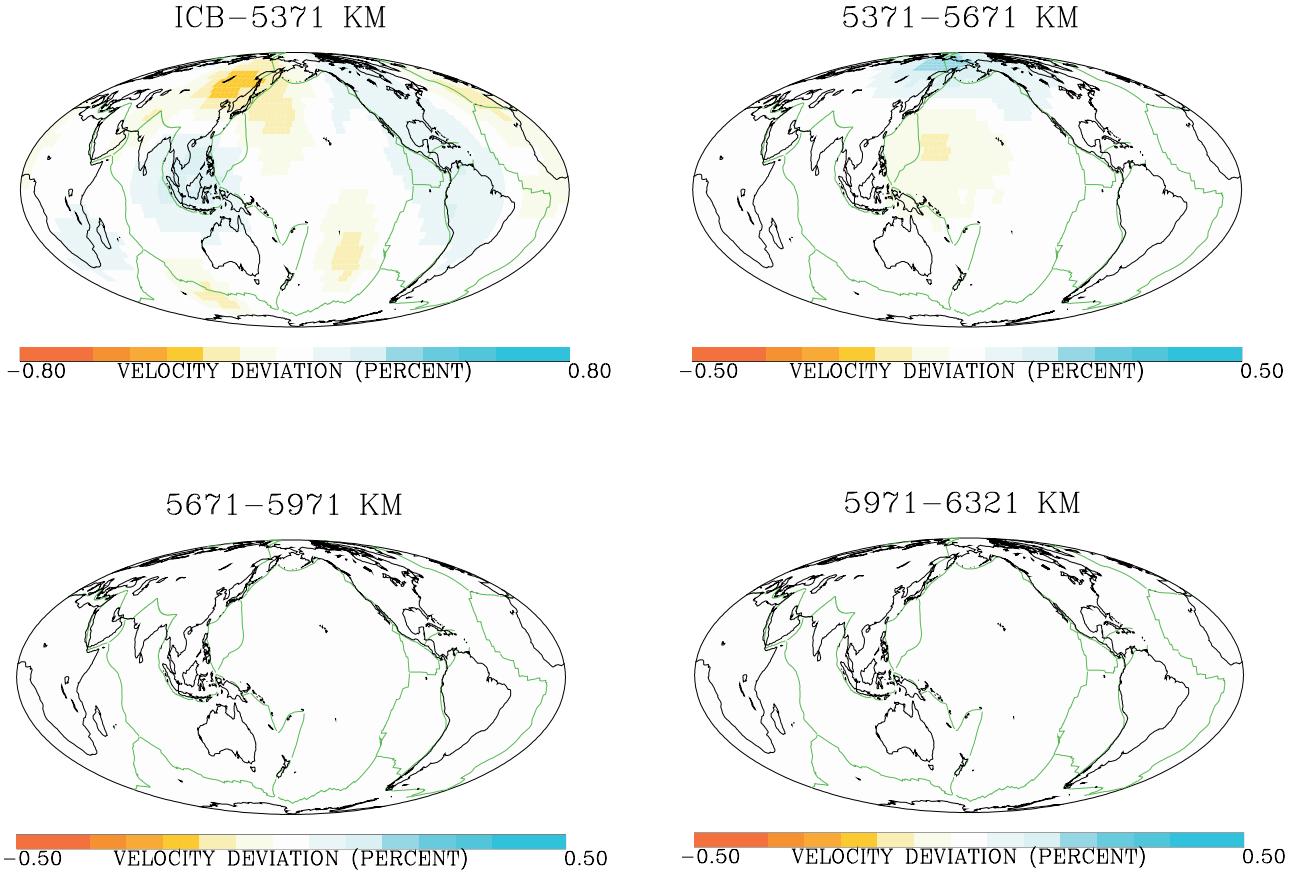
standard error of the respective datum. Thus, our constraint matrix is given by

$$\mathbf{M} = \begin{bmatrix} \mathbf{W}_D \mathbf{V} & \mathbf{W}_D \mathbf{H} & \mathbf{W}_D \mathbf{S} & \mathbf{W}_D \mathbf{B} & \mathbf{W}_D \mathbf{A} \\ \mathbf{W}_R^v \mathbf{D} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{W}_A^v & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{W}_R^b \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{W}_A^b & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{W}_R^a \mathbf{D} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{W}_A^a \end{bmatrix},$$

a very large but sparse matrix. The vector of unknown parameters is given by

$$\delta\mathbf{x} = \begin{bmatrix} \delta\mathbf{v} \\ \delta\mathbf{h} \\ \delta\mathbf{s} \\ \delta\mathbf{b} \\ \delta\mathbf{a} \end{bmatrix} \quad (\text{A2})$$

## Inner Core



**Figure 15.** Estimates of compressional velocity deviations within the inner core.

while the right-hand-side is

$$\delta\mathbf{T} = \begin{bmatrix} \mathbf{W}_D \mathbf{d} \\ \mathbf{0} \end{bmatrix}$$

### Appendix B

[53] As noted by *Pratt and Chapman* [1992], simple formulas for SVD-based estimates of resolution and covariance only hold for diagonal regularization, such as regularization associated with norm penalty functions. Non-diagonal penalties, such as a roughness penalty, require modifications, as derived in this appendix.

#### B1. Resolution Matrix

[54] The formal definition of model parameter resolution results when we relate our model parameter estimates  $\hat{\mathbf{x}}$ , as given by equation (16),

$$\hat{\mathbf{x}} = \mathbf{V}_\kappa \Lambda_\kappa^{-1} \mathbf{U}_\kappa^T \mathbf{T} \quad (\text{B1})$$

to a hypothetical ‘true’ model,  $\delta\mathbf{x}$ . The ‘true’ model satisfies  $\mathbf{d} = \Gamma\delta\mathbf{x}$  and is not subject to any roughness or norm

constraints, hence  $\mathbf{L} = 0$ , as is clear from equation (12). From equation (16) we have

$$\hat{\mathbf{x}} = \mathbf{V}_\kappa \Lambda_\kappa^{-1} \mathbf{U}_\kappa^T \begin{bmatrix} \Gamma \\ \mathbf{0} \end{bmatrix} \delta\mathbf{x}. \quad (\text{B2})$$

The resolution matrix is formally defined as

$$\mathbf{R} = \mathbf{V}_\kappa \Lambda_\kappa^{-1} \mathbf{U}_\kappa^T \begin{bmatrix} \Gamma \\ \mathbf{0} \end{bmatrix}. \quad (\text{B3})$$

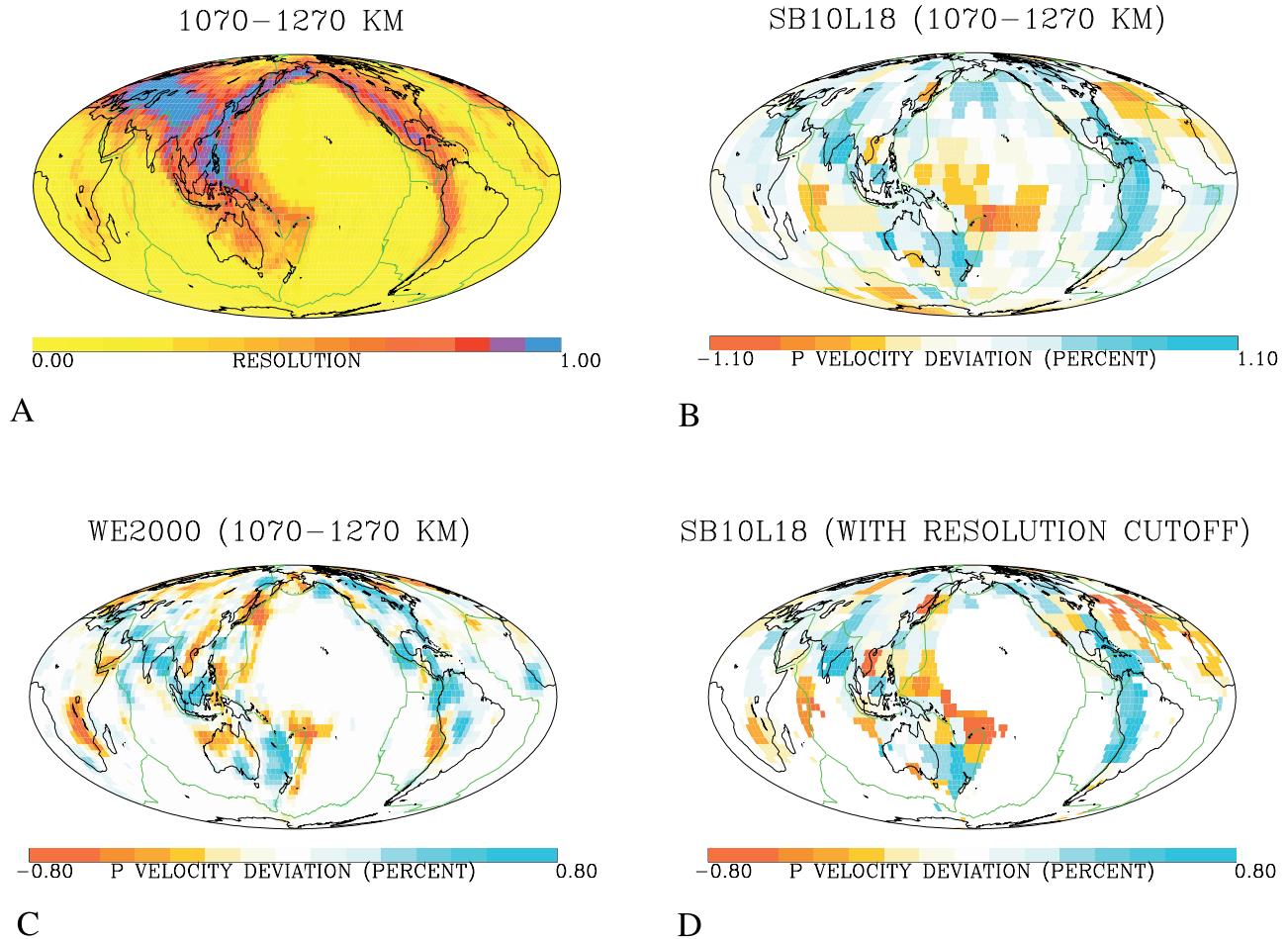
If we consider the SVD of  $\mathbf{G}$

$$\mathbf{G} = \begin{bmatrix} \Gamma \\ \mathbf{L} \end{bmatrix} = \mathbf{U} \Lambda \mathbf{V}^T \quad (\text{B4})$$

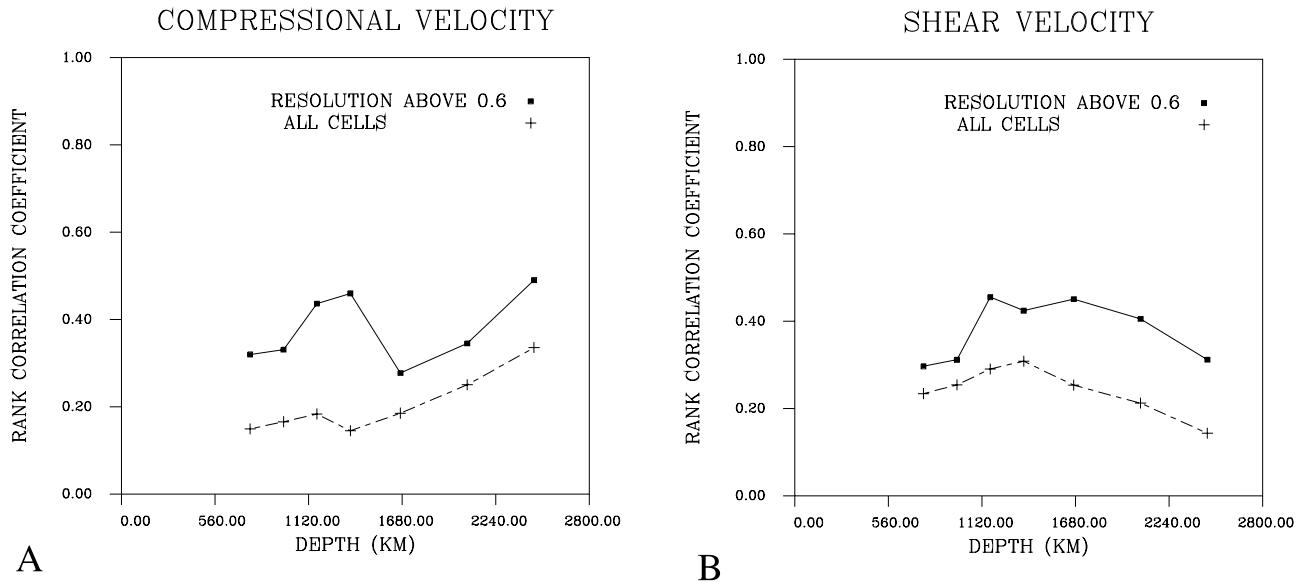
in a partitioned form, where the matrix  $\mathbf{U}$  is written

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} \quad (\text{B5})$$

we can simplify our treatment. The matrix  $\mathbf{U}_1$  corresponds to the  $N_d$  data rows (associated with  $\Gamma$ ), while  $\mathbf{U}_2$



**Figure 16.** (A) Model parameter resolution associated with our  $P$  velocity model. (B) Model SB10L18, the  $P$  velocity variations in the depth range 1070–1270. (C) The model discussed in this paper (WE2000) with anomalies in poorly resolved regions set to zero. (D) Model SB10L18 with anomalies in poorly resolved regions set to zero.



**Figure 17.** Nonparametric rank correlation coefficients for depths in the lower mantle. The crosses denote correlations based upon velocities in all cells of the models. Filled squares denote correlations in the models after eliminating cells with resolution below the 0.6 cutoff. (A) Correlation coefficients for compressional ( $P$ ) velocity. (B) Correlation coefficients for shear ( $S$ ) velocity.

corresponds to the  $N_r$  regularization rows (associated with  $\mathbf{L}$ ). Our representation of  $\mathbf{\Gamma}$  reduces to

$$\mathbf{\Gamma} = \mathbf{U}_1 \mathbf{\Lambda} \mathbf{V}^T \quad (\text{B6})$$

and the resolution matrix takes the form

$$\mathbf{R} = \mathbf{V}_\kappa \mathbf{\Lambda}_\kappa^{-1} (\mathbf{U}_1)_\kappa^T (\mathbf{U}_1)_\kappa \mathbf{\Lambda}_\kappa \mathbf{V}_\kappa^T. \quad (\text{B7})$$

Because the  $i$ th column of  $\mathbf{U}_1$  comprises only a part of the singular vector  $\mathbf{u}_i$ , we cannot expect orthogonality. However, the orthogonality of the singular vectors  $\mathbf{u}_i$  ensures that

$$\mathbf{U}^T \mathbf{U} = [\mathbf{U}_1^T \mathbf{U}_2^T] \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} = \mathbf{I}. \quad (\text{B8})$$

Thus, we may write the product  $\mathbf{U}_1^T \mathbf{U}_1$  in terms of  $\mathbf{U}_2$  and its transpose

$$\mathbf{U}_1^T \mathbf{U}_1 = \mathbf{I} - \mathbf{U}_2^T \mathbf{U}_2. \quad (\text{B9})$$

The advantage of equation (B9) is in the relative sizes of the matrices  $\mathbf{U}_1$  and  $\mathbf{U}_2$ . In particular, for tomographic problems we usually have many more data than unknowns, making  $\mathbf{U}_1$  much larger than  $\mathbf{U}_2$ . Using (B9), our representation of the resolution matrix takes the form

$$\mathbf{R} = \mathbf{V}_\kappa \mathbf{\Upsilon} \mathbf{V}_\kappa^T \quad (\text{B10})$$

where

$$\mathbf{\Upsilon} = \mathbf{I} - \mathbf{\Lambda}_\kappa^{-1} (\mathbf{U}_2)_\kappa^T (\mathbf{U}_2)_\kappa \mathbf{\Lambda}_\kappa. \quad (\text{B11})$$

Note that when no regularization is present,  $\mathbf{\Upsilon} = \mathbf{I}$ , and the expression reduces to the conventional SVD estimate

$$\mathbf{R} = \mathbf{V}_\kappa \mathbf{V}_\kappa^T \quad (\text{B12})$$

[Menke, 1984; Parker, 1994].

## B2. Covariance Matrix

[55] In the derivation of model parameter covariances we use the fact that the model parameter estimates,  $\delta\hat{\mathbf{x}}$  are linearly related to the data  $\mathbf{d}$ ,

$$\delta\hat{\mathbf{x}} = \mathbf{V}_\kappa \mathbf{\Lambda}_\kappa^{-1} \mathbf{U}_\kappa^T \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix}. \quad (\text{B13})$$

We use the partitioned form of  $\mathbf{U}$ , see equation (B5), to write

$$\hat{\mathbf{x}} = \mathbf{V}_\kappa \mathbf{\Lambda}_\kappa^{-1} (\mathbf{U}_1)_\kappa^T \mathbf{d} = \mathbf{G}^\dagger \mathbf{d} \quad (\text{B14})$$

where  $\mathbf{G}^\dagger$  is the generalized inverse. Based upon this linear relationship we can map the covariance associated with the data,  $\mathbf{C}_d$  into model parameter covariances,

$$\mathbf{C}_m = \mathbf{G}^\dagger \mathbf{C}_d (\mathbf{G}^\dagger)^T \quad (\text{B15})$$

[Menke, 1984]. Because the problem has been scaled, and every row has been normalized by the standard error of the associated datum,  $\mathbf{C}_d$  takes the form  $\mathbf{I}$ . Thus,

$$\mathbf{C}_m = \mathbf{G}^\dagger (\mathbf{G}^\dagger)^T, \quad (\text{B16})$$

or, in terms of the SVD of  $\mathbf{G}$

$$\mathbf{C}_m = \mathbf{V}_\kappa \mathbf{\Lambda}_\kappa^{-1} (\mathbf{U}_1)_\kappa^T (\mathbf{U}_1)_\kappa \mathbf{\Lambda}_\kappa^{-1} \mathbf{V}_\kappa^T. \quad (\text{B17})$$

Using (B9), we write  $\mathbf{C}_m$  as

$$\mathbf{C}_m = \mathbf{V}_\kappa \mathbf{\Omega} \mathbf{V}_\kappa^T \quad (\text{B18})$$

where

$$\mathbf{\Omega} = \mathbf{\Lambda}_\kappa^{-2} - \mathbf{\Lambda}_\kappa^{-1} (\mathbf{U}_2)_\kappa^T (\mathbf{U}_2)_\kappa \mathbf{\Lambda}_\kappa^{-1}. \quad (\text{B19})$$

[56] **Acknowledgments.** We would like to thank Gabi Laske and Guy Masters for providing a copy of their model SB10L18. This work was supported by a Laboratory Directed Research and Development grant, Office of Energy Research, Division of Basic Energy Sciences, Engineering, and Geosciences, of the U.S. Department of Energy under contract DE-AC03-76SF00098. All computations were carried out at the Center for Computational Seismology and the National Energy Research Scientific Computing (NERSC) Center of the Berkeley Laboratory.

## References

- Berry, M., scale singular value computations, *Int. J. Supercomput Appl.*, 6, 13–49, 1992.
- Bijwaard, H., and W. Spakman, Non-linear global P-wave tomography by iterated linearized inversion, *Geophys. J. Int.*, 141, 71–82, 2000.
- Bijwaard, H., W. Spakman, and E. R. Engdahl, Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, 103, 30,055–30,078, 1998.
- Boschi, L., and A. M. Dziewonski, Whole Earth tomography from delay times of  $P$ ,  $PcP$ , and  $PKP$  phases: Lateral heterogeneities in the outer core or radial anisotropy in the mantle?, *J. Geophys. Res.*, 105, 13,675–13,696, 2000.
- Buland, R., Uniform reduction error analysis, *Bull. Seismol. Soc. Am.*, 76, 217–230, 1986.
- Clayton, R. W., and R. P. Comer, A tomographic analysis of mantle heterogeneities from body wave travel times, *Eos Trans. AGU*, 64, 776, 1983.
- Cullum, J. K., and R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations*, vol. 1, *Theory*, Birkhäuser, Boston, 1985.
- Deal, M. M., and G. Nolet, Comment on “Estimation of resolution and covariance for large matrix inversions” by Zhang and McMechan, *Geophys. J. Int.*, 127, 245–250, 1996.
- 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 F. Gilbert, The effect of small, aspherical perturbations on travel times and a re-examination of the corrections for ellipticity, *Geophys. J. R. Astron. Soc.*, 44, 7–17, 1976.
- Dziewonski, A. M., B. H. Hager, and R. J. O’Connell, Large-scale heterogeneities in the lower mantle, *J. Geophys. Res.*, 82, 239–255, 1977.
- Engdahl, E. R., R. D. van der Hilst, and R. P. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Am.*, 88, 722–743, 1998.
- Fukao, Y., M. Obayashi, H. Inoue, and M. Nenbai, Subducting slabs stagnant in the transition zone, *J. Geophys. Res.*, 99, 2809–2822, 1992.
- Gonin, R., and A. H. Money, Non-linear  $L_p$ -norm estimation, I, On the choice of the exponent,  $p$ , where the errors are additive, *Commun. Stat., Theory Methods*, 14, 827–840, 1985.
- Gudmundsson, O., Some problems in global tomography: Modeling the core–mantle boundary and statistical analysis of travel-time data, Ph.D. thesis, Calif. Inst. of Technol., Pasadena, 1989.
- Inoue, H., Y. Fukao, K. Tanabe, and Y. Ogata, Whole mantle P-wave travel time tomography, *Phys. Earth Planet. Inter.*, 59, 294–328, 1990.

- Iyer, H. M., and K. Hirahara, *Seismic Tomography: Theory and Practice*, Chapman and Hall, New York, 1993.
- Jackson, D. D., The use of *a priori* data to resolve non-uniqueness in linear inversion, *Geophys. J. R. Astron. Soc.*, 57, 137–157, 1979.
- Karason, H., and R. D. van der Hilst, Tomographic imaging of the lowermost mantle with differential times of refracted and diffracted core phases ( $PKP, P_{dig}$ ), *J. Geophys. Res.*, 106, 6569–6587, 2001.
- Kennett, B. L. N., E. R. Engdahl, and R. Buland, Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, 122, 108–124, 1995.
- Kennett, B. L. N., S. Widjiantoro, and R. van der Hilst, Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle, *J. Geophys. Res.*, 103, 12,469–12,493, 1998.
- Lanczos, C., An iterative method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Natl. Bur. Stand.*, 45, 255–282, 1950.
- Lehmann, E. L., *Nonparametrics: Statistical Methods Based on Ranks*, Holden-Day, San Francisco, 1975.
- 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.
- Masters, G., H. Bolton, G. Laske, and A. Dziewonski, The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: Implications for chemical and thermal structure, in *Earth's Deep Interior: Mineral Physics and Tomography From Atomic to the Global Scale*, *Geophys. Monogr. Ser.*, vol. 117, edited by Karato, S., et al., pp. 63–87, AGU, Washington, D. C., 2000.
- Menke, W., *Geophysical Data Analysis: Discrete Inverse Theory*, Academic, San Diego, Calif., 1984.
- Montagner, J. P., and T. Tanimoto, Global upper mantle tomography of seismic velocities and anisotropies, *J. Geophys. Res.*, 96, 20,337–20,351, 1991.
- Morelli, A., and A. M. Dziewonski, Topography of the core–mantle boundary and lateral homogeneity of the liquid core, *Nature*, 325, 678–683, 1987.
- Noble, B., and J. W. Daniel, *Applied Linear Algebra*, Prentice-Hall, Old Tappan, N. J., 1977.
- Nolet, G., Seismic wave propagation and seismic tomography, in *Seismic Tomography*, edited by Nolet, G., pp. 1–23, D. Reidel, Norwell, Mass., 1987.
- Nolet, G., S. P. Grand, and B. L. N. Kennett, Seismic heterogeneity in the upper mantle, *J. Geophys. Res.*, 99, 23,753–23,766, 1994.
- Nolet, G., R. Montelli, and J. Virieux, Explicit, approximate expressions for the resolution and a posteriori covariance of massive tomographic systems, *Geophys. J. Int.*, 138, 36–44, 1999.
- Paige, C. C., and M. A. Saunders, LSQR: An algorithm for sparse linear equations and sparse linear systems, *ACM Trans. Math. Software*, 8, 195–209, 1982.
- Parker, R. L., *Geophysical Inverse Theory*, Princeton Univ. Press, Princeton, N. J., 1994.
- Pratt, R. G., and C. H. Chapman, Travelttime tomography in anisotropic media, 2, Application, *Geophys. J. Int.*, 109, 20–37, 1992.
- 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.
- Pulliam, R. J., D. W. Vasco, and L. R. Johnson, Tomographic inversions for mantle P wave velocity structure based on the minimization of  $\ell^2$  and  $\ell^1$  norms of International Seismological Centre travel time residuals, *J. Geophys. Res.*, 98, 699–734, 1993.
- Ritzwoller, M. H., and E. M. Lavelle, Three-dimensional seismic models of the Earth's mantle, *Rev. Geophys.*, 33, 1–66, 1995.
- Romanowicz, B., Seismic tomography of the Earth's mantle, *Annu. Rev. Earth Planet. Sci.*, 19, 77–99, 1991.
- Roudil, P., and A. Souriau, Liquid core structure and  $PKP$  station anomalies derived from  $PKP(BC)$  propagation times, *Phys. Earth Planet. Inter.*, 77, 225–236, 1993.
- Scales, J. A., A. Gerszenkorn, and S. Treitel, Fast  $\ell^p$  solution of large, sparse, linear systems: Application to seismic travel time tomography, *J. Comput. Phys.*, 75, 314–333, 1988.
- Sengupta, M. K., and M. N. Toksoz, Three-dimensional model of seismic velocity variation in the Earth's mantle, *Geophys. Res. Lett.*, 3, 84–86, 1976.
- Shearer, P. M., Global mapping of upper mantle reflectors from long-period SS precursors, *Geophys. J. Int.*, 115, 878–904, 1993.
- Shearer, P. M., K. M. Toy, and J. A. Orcutt, Axi-symmetric Earth models and inner-core anisotropy, *Nature*, 333, 228–232, 1988.
- Simon, H., Analysis of the symmetric Lanczos algorithm with reorthogonalization methods, *Linear Algebra Appl.*, 61, 101–131, 1984.
- Souriau, A., and G. Poupinet, A latitudinal pattern in the structure of the outermost liquid core, revealed by the travel times of SKKS-SKS seismic phases, *Geophys. Res. Lett.*, 17, 2005–2007, 1990.
- Su, W.-J., and A. M. Dziewonski, Simultaneous inversions for 3-D variations in shear and bulk velocity in the mantle, *Phys. Earth Planet. Inter.*, 100, 135–156, 1997.
- Tarantola, A., *Inverse Problem Theory: Methods for Fitting and Model Parameter Estimation*, Elsevier Sci., New York, 1987.
- Trampert, J., and J. H. Woodhouse, High resolution global phase velocity distribution, *Geophys. Res. Lett.*, 23, 21–24, 1996.
- van der Hilst, R. D., and H. Karason, Compositional heterogeneity in the bottom 1000 kilometers of the Earth's mantle: Toward a hybrid convection model, *Science*, 283, 1885–1888, 1999.
- van der Hilst, R. D., S. Widjiantoro, and E. R. Engdahl, Evidence for deep mantle circulation from global tomography, *Nature*, 386, 578–584, 1997.
- Vasco, D. W., and L. R. Johnson, Whole Earth structure estimated from seismic arrival times, *J. Geophys. Res.*, 103, 2633–2671, 1998.
- Vasco, D. W., R. J. Pulliam, and L. R. Johnson, Formal inversion of ISC arrival times for mantle P-velocity structure, *Geophys. J. Int.*, 113, 586–606, 1993.
- Vasco, D. W., R. J. Pulliam, L. R. Johnson, and P. S. Earle, Robust inversion of IASP91 travel time residuals for mantle P and S velocity structure, earthquake mislocations, and station corrections, *J. Geophys. Res.*, 99, 13,727–13,755, 1994.
- Vasco, D. W., L. R. Johnson, and R. J. Pulliam, Lateral variations in mantle velocity structure and discontinuities determined from P, PP, S, SS, and SS-SdS travel time residuals, *J. Geophys. Res.*, 100, 24,037–24,059, 1995.
- Vasco, D. W., L. R. Johnson, and O. Marques, Global Earth structure: Inference and assessment, *Geophys. J. Int.*, 137, 381–407, 1999.
- Yao, Z. S., R. G. Roberts, and A. Tryggvason, Calculating resolution and covariance matrices for seismic tomography with the LSQR method, *Geophys. J. Int.*, 138, 886–894, 1999.
- Zhang, J., and G. A. McMechan, Estimation of resolution and covariance for large matrix inversions, *Geophys. J. Int.*, 121, 409–426, 1995.
- Zhou, H.-W., A high resolution P wave model for the top 1200 km of the mantle, *J. Geophys. Res.*, 101, 27,791–27,810, 1996.

L. R. Johnson and D. W. Vasco, Earth Sciences Division, Center for Computational Seismology, Berkeley Laboratory, University of California, 1 Cyclotron Road, Berkeley, CA 94720, USA. (lrj@ccs.lbl.gov; dwvasco@lbl.gov)

O. Marques, High Performance Computing Research Department, Berkeley Laboratory, University of California, Berkeley, CA 94720, USA.