Geochemistry Geophysics Geosystems G3
AN ELECTRONIC JOURNAL OF THE EARTH SCIENCES
Published by AGU and the Geochemical Society

Technical Brief
Volume 4, Number 3
5 March 2003
1022, doi:10.1029/2001GC000213
ISSN: 1525-2027

Adaptive whole Earth tomography

Malcolm Sambridge
Research School of Earth Sciences, Australian National University, Canberra, ACT 0200, Australia (malcolm@rses.anu.edu.au)

Rado Faletič
Department of Physics, Faculty of Science, Australian National University, Canberra, ACT 0200, Australia (Rado.Faletic@anu.edu.au)

[1] An approach for seismic tomography is presented which allows the parameterization to be refined during the inversion. The objective is to use the data to refine the mesh and the velocity model together, and hence both are considered part of the solution. Some simple rules are used to identify the volumes of a three-dimensional model in need of refinement. The self-adaptive parameterization is applied to an initial mesh built from uniformly distributed spherical triangles and Delaunay tetrahedra. Application of the technique to a typical summary ray *P*-wave arrival time data set shows it to be both feasible and practical for large scale whole Earth tomography. A noticeable trend in the resulting models, as the parameterization is refined, is the thinning of the Farallon and Tethys subduction features imaged in the mid mantle, together with an increase in amplitude of the velocity perturbation.

Components: 6948 words, 15 figures, 1 table.

Keywords: Seismic tomography; adaptive parameterization; numerical techniques.

Index Terms: 8180 Tectonophysics: Evolution of the Earth: Tomography; 3260 Mathematical Geophysics: Inverse theory; 3230 Mathematical Geophysics: Numerical solutions; 7218 Seismology: Lithosphere and upper mantle.

Received 9 August 2001; **Revised** 14 February 2002; **Accepted** 3 December 2002; **Published** 5 March 2003.

Sambridge, M., and R. Faletič, Adaptive whole Earth tomography, *Geochem. Geophys. Geosyst.*, *4*(3), 1022, doi:10.1029/2001GC000213, 2003.

1. Introduction

[2] Tomographic imaging of mantle structure has been performed for more than 20 years. Comprehensive summaries are given by *Romanowicz* [1991], *Iyer and Hirahara* [1993], *Nolet et al.* [1994], *Ritzwoller and Lavely* [1995], and *Dziewonski* [1995]. In recent years the increased availability of high quality seismic arrival times and waveform data, together with improvements in methodology have led to detailed information on lateral heterogeneities, and more consistency between velocity models *Grand et al.* [1997]. However, it has long been known that tomographic images are limited by a number of factors. One is the uneven distribution of sources and recording stations across the globe. This is particularly true in the case of arrival time tomography where the distribution of seismic ray paths often leads to highly variable sampling within the mantle.

[3] It is self evident that uneven ray path sampling leads to limited resolution in regions of poor data coverage. In tomography the usual way of dealing

with ill-constrained parts of a model are to apply some spatial smoothing, norm damping, or simply to coarsen the parameterization, e.g., increase block sizes. Traditionally these forms of "regularization" have been applied uniformly across the entire model, which raises the possibility that, while the ill-constrained regions are being damped, the well constrained regions are being over-smoothed and hence information may be lost.

[4] Some applications of body wave tomography have used non-uniform sized rectangular blocks to account for uneven ray path sampling [*Inoue et al. 1990*; *Bijwaard et al.*, 1998; *Karason and van der Hilst*, 2001]. Surface wave inversions have also been performed using unevenly sized spherical triangles [*Wang et al.*, 1998]. In each case the cell sizes were constructed so that their density approximately matches that of the ray paths calculated in a 1-D reference model. In this way more unknowns are used in regions where the data density is higher, and less where it is lower. These studies appear to have been very successful. The next logical question to ask might be whether the parameterization could be made to adapt to the structural signal in the data during the inversion, i.e., not to impose an irregular grid a priori, but rather to only add detail where the data require it.

[5] The idea of a data adaptive grid in seismic tomography is not new. Several approaches have been proposed for 2-D problems. *Michelini* [1995] proposed an adaptive method using B-splines, and tested it on cross bore-hole tomography. *Curtis and Snieder* [1997] used a genetic algorithm to build a triangular parameterization while minimizing the condition number of the resulting tomographic system of equations. Very recently, *Chiao and Kuo* [2001] have proposed a multi-scale tomographic parameterization based on 2-D spherical wavelets, and used it to estimate of lateral shear wave heterogeneity in the D'' layer. In each case the procedures were quite successful and lend encouragement for much larger scale 3-D problems.

[6] In this paper we describe a data adaptive approach which is practical for whole Earth tomographic problems. Our method is an extension of the work of *Sambridge and Gudmundsson* [1998], where most of the geometric tools are described in detail. We illustrate the technique by applying it to a large summary ray *P*-wave data set. This also allows us to investigate the robustness of common structural features in tomographic models produced with earlier static parameterizations.

2. The Method

2.1. Parameterizing the Earth With Tetrahedral Cells

[7] The adaptive parameterization used in this study is based on a division of the Earth into a set of tetrahedra. We begin with a uniform tetrahedral mesh and locally subdivide the elements according to a prescribed set of rules at successive stages during the inversion. A convenient way of producing the initial uniform mesh is to first generate a series of 'shells', constructed from spherical triangles (see Figure 1), and then join these together to produce Delaunay tetrahedra. Here we use the *quickhull* algorithm of *Barber et al.* [1993] to construct Delaunay tetrahedra, and the method of *Wang and Dahlen* [1995] to construct spherical triangles.

[8] The initial regular mesh is constructed from sub-divisions of an icosahedron. Figure 1 shows the first four spherical triangle meshes, produced in this way. The average size of the tetrahedra in the uniform mesh is determined by the density of the spherical triangles in each layer (see Table 1). Here we use layers at 18 depths spanning the mantle (0, 100, 200, 300 410, 520, 660, 820, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600, 2750, 2889 km). The advantage of using spherical triangles is that a near uniform coverage is achieved across a sphere, and the distortion effects from the poles are much reduced compared to a mesh built on a lat-long grid. For other applications of spherical triangles see *Constable et al.* [1993].

[9] Three dimensional Delaunay tetrahedra can then be built from the vertices of the spherical triangles using any Delaunay algorithm (see *Sam-*

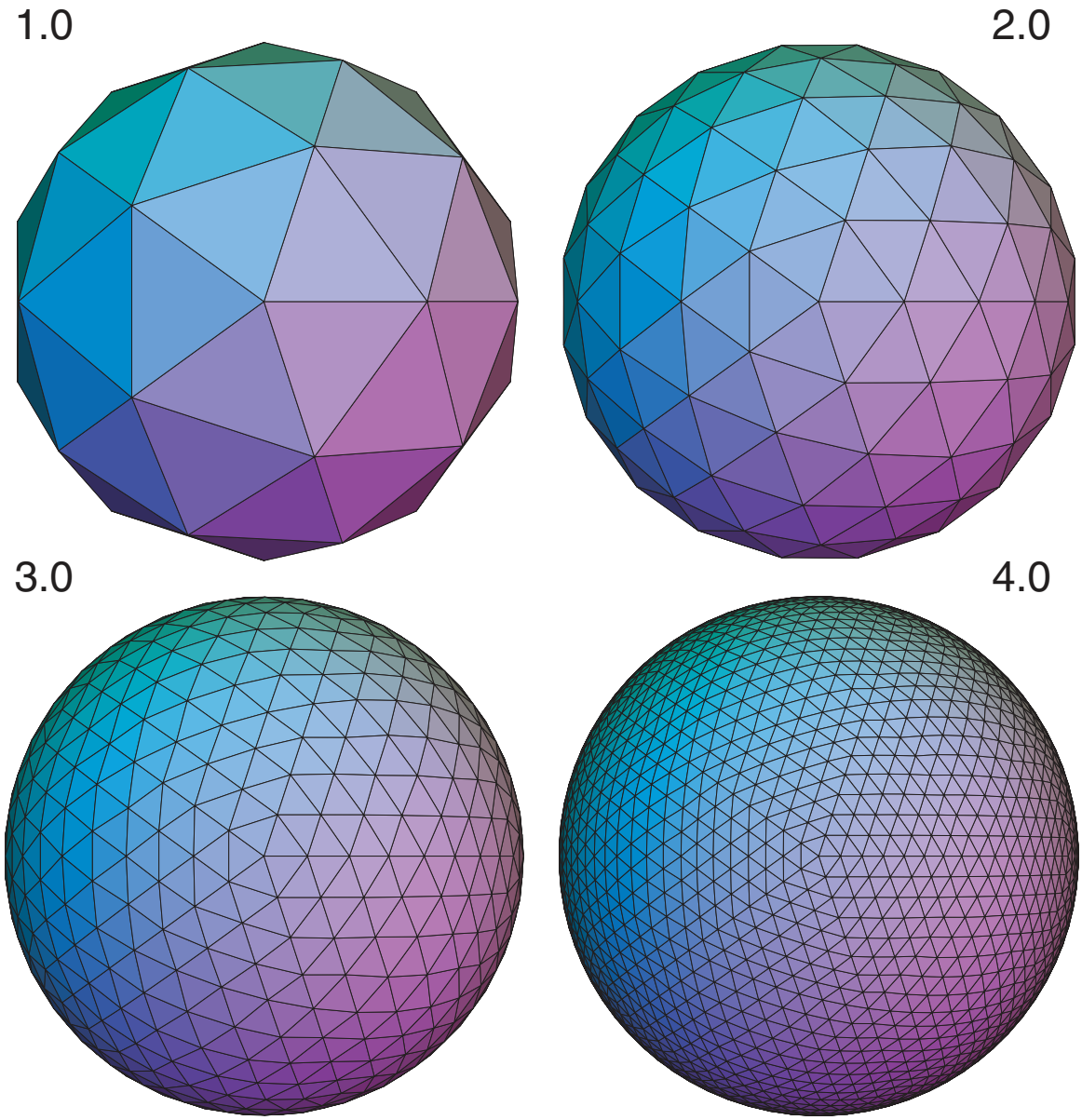

Figure 1. Four subdivisions of an icosahedron which define a near uniform triangular grid on the surface of the Earth. These are the surface representations of the four uniform 3-D meshes *param1.0–param4.0*.

bridge et al. [1995] for a discussion). Table 1 gives some statistics of the four uniform 3-D parameterizations built in this way. The simplest choice of inversion parameters are then the perturbations in slowness (reciprocal of velocity) from some reference model, within each tetrahedron.

2.2. Local Refinement

[10] The local refinement of the parameterization requires two problems to be solved. First we must decide which tetrahedra need to be sub-divided, and second how to divide them. Clearly many choices are possible. In the work of *Bijwaard et al.* [1998] a nested set of cubic cells were defined using the densities of ray paths, which remained fixed for the entire inversion. Here we wish to refine the parameterization during the inversion, in response to structural features detected. *Spakman and Bijwaard* [2001] have suggested that ray hit counts could also be used as the basis of an adaptive scheme, i.e., to drive the cell refinement process between iterations.

Table 1. Number of Points and Cells for the Four Spherical Triangle Meshes Shown in Figure 1 and the 3-D Tetrahedra Built From Them[a]

Label	$\mathcal{S}_s$	$\mathcal{T}_s$	N_t	$\mathcal{D}$
param0.0	12	20	1,315	8640 km
param1.0	42	80	4,056	4320 km
param2.0	162	320	16,189	2160 km
param3.0	642	1,280	64,973	1080 km
param4.0	2,562	5,120	259,418	540 km
2° × 2°	-	16,200	291,600	220 km

[a] *param0.0* refers to a simple icosahedron. $\mathcal{S}_s$ is the number of nodes and $\mathcal{T}_s$ is the number of spherical triangles on each surface, $\mathcal{D}$ is the edge length of the spherical triangle, and N_t is the total number of tetrahedra in each 3-D mesh. For a comparison, the last row shows the number of cells on the surface ($\mathcal{T}_s$) and total number of unknowns (N_t) in the 2° × 2° cubic cellular grid used by *van der Hilst et al.* [1997].

Such an approach would necessarily require 3-D ray tracing, and (in our view) it is questionable whether ray hit counts would vary sufficiently between iterations to drive a mesh refinement process. After experimenting with several alternatives we decided to use a criterion based on the maximum spatial gradients in seismic velocity perturbation measured across each tetrahedron face. The maximum gradient for the i-th tetrahedron is given by, g_i,

$$g_i = \max_j \frac{|V_{j,i} - v_i|}{\|\mathbf{b}_{j,i} - \mathbf{b}_i\|} \quad (1)$$

where v_i is the velocity perturbation in tetrahedron i, $V_{j,i}$ is the velocity perturbation in the j-th neighbour of tetrahedron i, and $\mathbf{b}$ represents the position of a tetrahedron's centroid. Each tetrahedron has a maximum of four neighbours, as illustrated in Figure 2, and so g_i is simply the largest of up to four alternatives.

[11] By ranking the tetrahedra according to g_i we may select any percentage of tetrahedra for refinement. To sub-divide a tetrahedron we add a node along each of its six edges, a method known as *bi-section*. The new nodes generated in this way form the input to the same Delaunay tetrahedralization algorithm used to construct the uniform meshes [*Barber et al.*, 1993]. Figure 3 shows an example of a single tetrahedron being sub-divided in this way. At least eight new tetrahedra are formed by a single sub-division, and hence the number of slowness parameters in the inversion is also increased by at least eight. Note that neighbouring cells are also likely to be subdivided giving a total of up to 24 new tetrahedra. The exact number will depend on the distribution of vertices of other tetrahedra surrounding the one that is being sub-divided. We choose to use Delaunay tetrahedra because they have the property of 'maximum-minimum' internal solid angle which makes them, on average, as least long and thin as possible. (For discussions of Delaunay

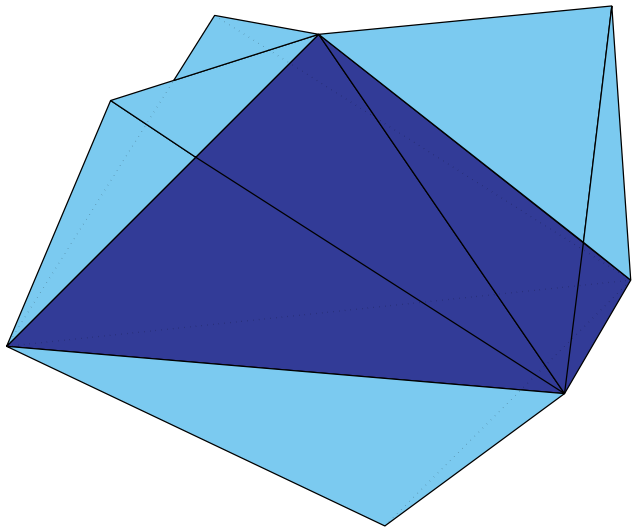

Figure 2. Gradients of the slowness field can be estimated using the difference between slownesses in neighbouring tetrahedra. Each tetrahedron has (at most) four neighbouring tetrahedra. The gradients across each face can also be used in the damping of the tomographic system of equations.

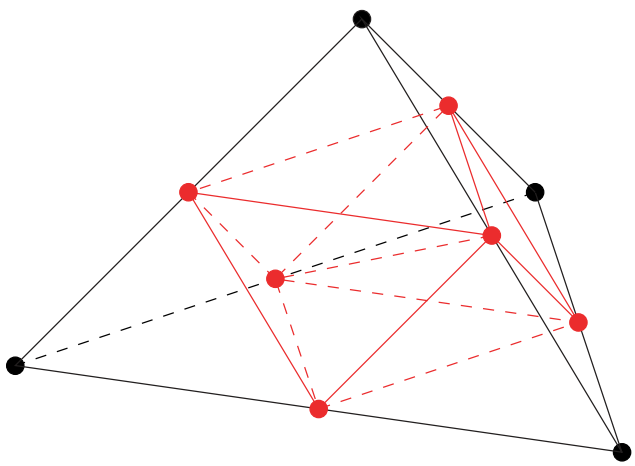

Figure 3. Bi-section of a tetrahedron by placing six (red) nodes at the centres of each edge.

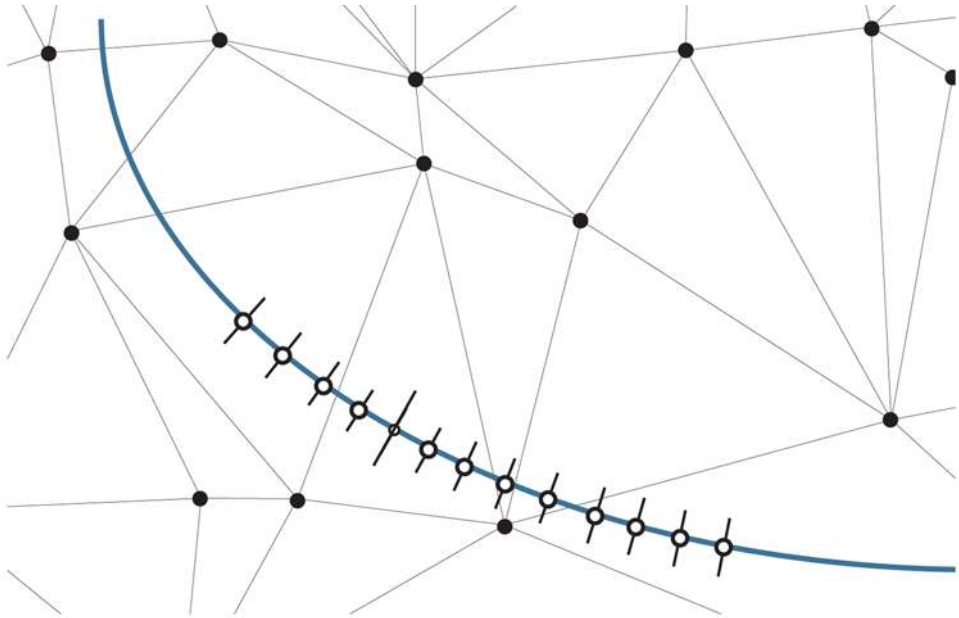

Figure 4. Calculation of ray lengths through irregular tetrahedra by stepping down the ray in equal steps of arc length.

tetrahedra and their properties see *Okabe et al.* [1992], *Watson* [1992], *Sambridge et al.* [1995], and *Gudmundsson and Sambridge*, [1998].) Note that alternatives to bi-section exist, e.g., placing a new node in the centre of the tetrahedron, but in our tests we found that this type of sub-division tended to create many distorted (long-thin) tetrahedra.

2.3. Calculating Ray Lengths in Irregular Tetrahedra

[12] An integral component of any tomographic study is the calculation of derivatives of travel times with respect to the inversion parameters. In our case this reduces to the calculation of ray lengths in a mesh of variably sized tetrahedra. The approach used here is simply to step along rays, traced in the 1-D reference model (*ak135* of *Kennett et al.* [1995]), and sum together the arc lengths contained in each tetrahedron encountered. Figure 4 shows an example in the 2-D case. After some experimentation the step length was chosen small enough to allow the ray to be approximated by a series of linear segments, but large enough to avoid excessive computation. (Here we use a constant arc length per ray chosen to give at least 10 segments per cell.)

[13] For each segment we also need to determine in which tetrahedra the end points lie. An efficient method was given by *Sambridge and Gudmundsson* [1998], which is used throughout the present study. Like in many large scale tomographic studies the computational effort required to calculate ray lengths is large compared with that needed to solve the resulting systems of equations. For the approximately 550000 summary rays used here, ray length calculations took approximately 12 hours on a uni-processor Sun ultra Sparc 20 with 128MB memory.

2.4. Formulating and Solving Tomographic Systems of Equations

[14] Since the construction and solution of the tomographic systems of equations is independent of the parameterization, the procedure used here is similar to many previous studies [*Iyer and Hirahara*, 1993; *van der Hilst et al.*, 1997]. The linear system of equations relating perturbations in slowness to travel time residuals with respect to a 1-D reference model can be written,

$$A\mathbf{s} = \delta\mathbf{d} \qquad (2)$$

where **s** is the vector of slowness perturbations, $\delta\mathbf{d}$ is the vector of data residuals, and A is the matrix of ray lengths in each tetrahedron. (For simplicity we will assume that the data covariance matrix is diagonal, and merely re-scales each equation in (2).) We expect this system to be under-determined (or mixed-determined) and require regularization (damping). A popular approach is to minimize a combination of slowness perturbation (minimum norm damping) and slowness gradients (often called flattening), at the same time as solving (2) in a least squares sense. This leads to the minimization of, $\psi(\mathbf{s})$,

$$\psi(\mathbf{s}) = (A\mathbf{s} - \delta\mathbf{d})^T(A\mathbf{s} - \delta\mathbf{d})) + \lambda_1^2\mathbf{s}^T\mathbf{s} + \lambda_2^2\sum_{i=1}^{N_t}\sum_{j=1}^{n_i}(s_i - S_{i,j})^2 \tag{3}$$

where s_i is the slowness perturbation of tetrahedron i, n_i is its number of neighbours ($n_i \leq 4$), $S_{i,j}$ is the slowness perturbation in the j-th neighbour of tetrahedron i, N_t are the total number of tetrahedra, and (λ_1, λ_2) are empirically determined scaling constants. The quadratic expression (3) may be rewritten as,

$$\psi(\mathbf{s}) = (M\mathbf{s} - \mathbf{q})^T(M\mathbf{s} - \mathbf{q}) \tag{4}$$

where

$$M = \begin{pmatrix} A \\ \lambda_1 I_{N_t} \\ \lambda_2 D \end{pmatrix}, \qquad \text{and} \qquad \mathbf{q} = \begin{pmatrix} \delta\mathbf{d} \\ 0 \\ 0 \end{pmatrix}. \tag{5}$$

Here I_{Nt} is an $N_t \times N_t$ identity matrix and D is a non-square matrix of size ($\Sigma\, n_i$) $\times N_t$ given by,

$$D_{ij} = \begin{cases} 1 & j = i, \\ -1 & j = \text{neighbour of } i, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

The multiplication of D and **s** produces a vector whose entries are the differences in slowness of each parameter with one of its neighbours, i.e., ($s_i - S_{i,j}$). Clearly (4) corresponds to the least squares solution of a larger (augmented) linear system of equations,

$$M\mathbf{s} = \mathbf{q} \tag{7}$$

One can verify that combining (4), (5) and (6) leads to (3). In some large scale tomographic studies [e.g., *Nolet*, 1987, 1993; *Widiyantoro and van der Hilst*, 1997], the linear system (7) is not actually solved directly. Instead the matrix D is replaced by the smaller ($N_t \times N_t$) matrix D^TD and hence the gradient damping equations correspond to

$$D^TD\mathbf{s} = 0 \tag{8}$$

These are the normal equations corresponding to minimization of the gradient damping term in (3). The reader will be able to verify that the square matrix D^TD corresponds to an approximate second derivative operator. It is worthwhile noting that if the system (8) were exactly satisfied while solving (7), then the gradient term in (3) would also be minimized. However, in general, the least squares solution of (7), with the second derivative operator, D^TD, replacing D in (5), does not correspond to the minimization of (3). Here we follow previous authors and make use of (8) in (7), and use the iterative equation solver *LSQR* of *Nolet* [1987].

2.5. Damping, Smoothing and Mesh Refinement

[15] It is well known that tomography is an ill-posed inverse problem, and as such requires "regularization" in order to construct a model. This means that since ray paths are approximated as zero width curves it is not possible to uniquely constrain a 3-D velocity field, and hence the range and character of allowable models needs to be restricted. The particular form of regularization used is a subjective choice and represents the biases imposed on the inversion. In this work we use the standard approach of gradient and minimum norm damping (represented by the corresponding terms in (3)), Since the parameterization restricts the range of allowable models it is also a form of regularization, and here we allow the mesh to vary both spatially and between iterations. We might therefore ask how these different forms of regularization (damping, smoothing and mesh) interact.

[16] It is clear that there is a complex inter-dependency between damping, smoothing and the adaptive parameterization. Note that the damping and smoothing terms in equation (3) depend only on the grid, and not on the physical gradients, and hence their influence on the inversion will directly depend to the particular mesh being used. Therefore we always observe the combined effect

of the damping, smoothing and mesh in regularizing the inversion. (It may not be possible to separate them in any meaningful way.) Here the tetrahedral mesh is able to evolve during the course of the inversion, and locally adapt in response to model gradients imaged in the earlier stages. In this way the adaptive mesh is a mechanism for enhancing the positive role of the damping and smoothing processes, i.e., by allowing them to vary in strength across the model, and during the inversion. For ''adaptive parameterization'' we might equally well read ''adaptive regularization,'' and all results should be viewed in these terms, i.e., it is the overall effect of mesh, damping and smoothing terms which is responsible for the results presented here. It is worthwhile noting that spatially varying regularization schemes have previously been applied to tomographic problems with relatively few unknowns [*Chou and Booker*, 1979; *Tarantola and Nercessian*, 1984].

[17] With an adaptive mesh the regularization process is also being influenced by the observed data, and so we might expect the problem to become more non-linear. In this work the numerical examples are restricted to the linearized regime (i.e., no 3-D ray-tracing) and so the non-linearity is not considered. Another downside of increased flexibility is the potential for numerical instabilities, i.e., where mesh refinement might become too aggressive and result in locally under-determined tetrahedra with too few rays. In this case the tuning of the gradient and smoothing processes, i.e., choice of λ_1, λ_2 in equation. (3), will be driven by the need to locally damp out such solution instabilities. It is clear that a balance must be achieved between the choice of λ_1, λ_2 (which will damp out changes to the model) and the mesh refinement process (which will allow ever smaller scale, and possibly higher amplitude, changes). Ultimately the models generated in the series of numerical tests described below are a reflection of the, necessarily subjective, tuning processes used here.

3. Results

[18] To illustrate the adaptive algorithm we perform a series of tomographic inversions using the summary ray data set of *van der Hilst et al.* [1997], [see also *Widiyantoro*, 1997; *Widiyantoro and van der Hilst*, 1997], built on the relocated global catalogue of *Engdahl et al.* [1998], which was made available to us by the authors. This data set involves approximately 550,000 summary rays derived from 6×10^6 P and *pwP* phases from more than 77,000 events recorded at 3750 stations worldwide. Rays with absolute P-residuals greater than 5 seconds are treated as outliers and removed. The averaging of many similar source/receiver pairs into summary rays helps reduce the influence of the noise processes, e.g., earthquake mislocation, small scale heterogeneity beneath stations, and other picking and timing errors common to similar source receiver paths etc. For full details of the data processing see *Widiyantoro and van der Hilst*, [1997].

[19] Even though data processing is aimed at reducing the influence of noise and other unmodelled effects (e.g., boundary perturbations anisotropy etc), errors will still be present in the summary rays, and these will inevitably contaminate the velocity models. What's more the noise reduction processes themselves may well remove signal from the data, especially on localized heterogeneity near the surface. *Widiyantoro and van der Hilst* [1997] and *van der Hilst et al.* [1997] studied the resolving power of the data set by means of synthetic tests on a regular parameterization, and their results may be taken as a guide for the present study.

[20] Each of the four regular tetrahedral meshes in Figure 1 are used as the starting point of separate inversions, and in each case we 'update' the tetrahedral mesh four times, by sub-dividing the 5% of tetrahedra with the largest g_i value given by (1). The bi-section algorithm is used for subdivision. We therefore perform 16 separate tomographic inversions in total. In each case the starting velocity model is the 1-D reference model *ak135* of *Kennett et al.* [1995], but the parameterization is the one obtained in the previous inversion in the series. For example, solution of the tomographic system using the regular mesh *param4.0* produces P-velocity model *AP4.0*, and the irregular parameterization *param4.1*. In the next stage *param4.1* is

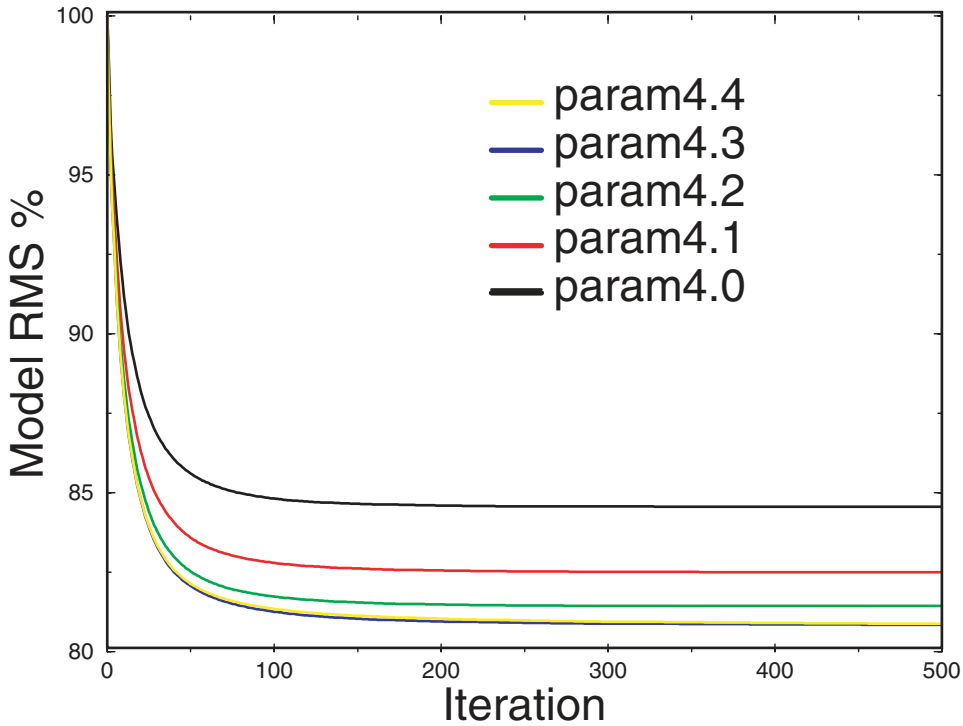

Figure 5. RMS deviation on the slowness field against iteration of the conjugate gradient solver for the *param4.** series.

used to produce model *AP4.1* and the more refined mesh *param4.2*, and so on. (Here we use the prefix "*AP*" to denote a P-wavespeed model produced with the adaptive scheme, and "*param*" to indicate a parameterization.)

[21] At each stage the parameterization is inherited from the previous step but the velocity model is not. In this way we avoid the difficult task of translating a velocity model from one irregular parameterization to another. Note that because each irregular mesh is determined from the previous stage, the inversion procedure is being used to constrain both the parameterization and the structure simultaneously.

[22] Our results are an extension of the earlier work of *Faletič* [1997], where full details of algorithm design, testing and implementation can be found. [This material is available at the URL given at the end of this paper.] Statistics for the 16 inversion results are summarized in Figures 5, 6 and 7.

[23] The performance of the iterative linear equation solver is illustrated in Figure 5 for the inversion series beginning from *param4.0*. In each case the *LSQR* algorithm is run for 500 iterations and the scaling constants λ_1 and λ_2 are adjusted empirically to achieved a balance between data fitting and explicit model damping. This was done by searching through pairs of values for (λ_1, λ_2) and solving the linear system of equations in each case. Our preference was to choose the smallest pair of values consistent with a stable solution, as determined by inspection of the model perturbations and data variance reduction. (As noted above the overall regularization of the inversion is determined by the combination of explicit damping and smoothing, i.e., choice of (λ_1, λ_2), together with the mesh refinement algorithm). Inevitably these choices are subjective, which is undesirable, but common to many (if not all) tomographic studies involving large linear systems. Figure 5 clearly shows that the *LSQR* solver has converged, in terms of producing no further changes to the velocity model.

[24] Figure 6 shows the variance reduction in each case, (color coded by the four separate regular starting models). One immediately notices that

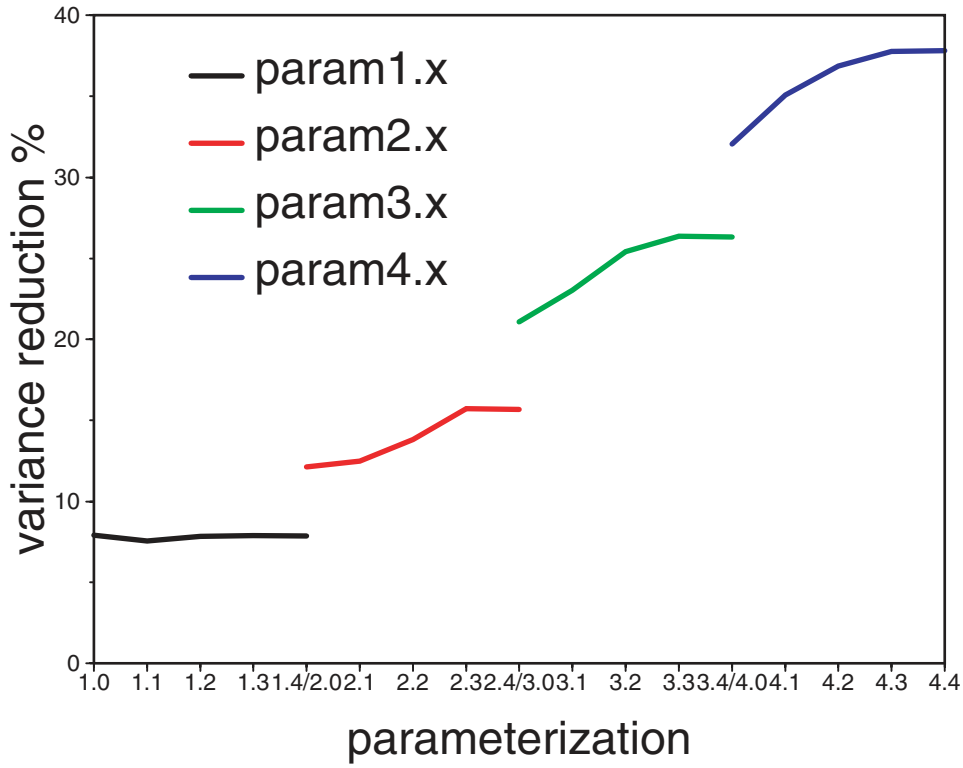

Figure 6. Data variance reduction for the adaptive parameterization algorithm starting from the four uniform meshes in Figure 1.

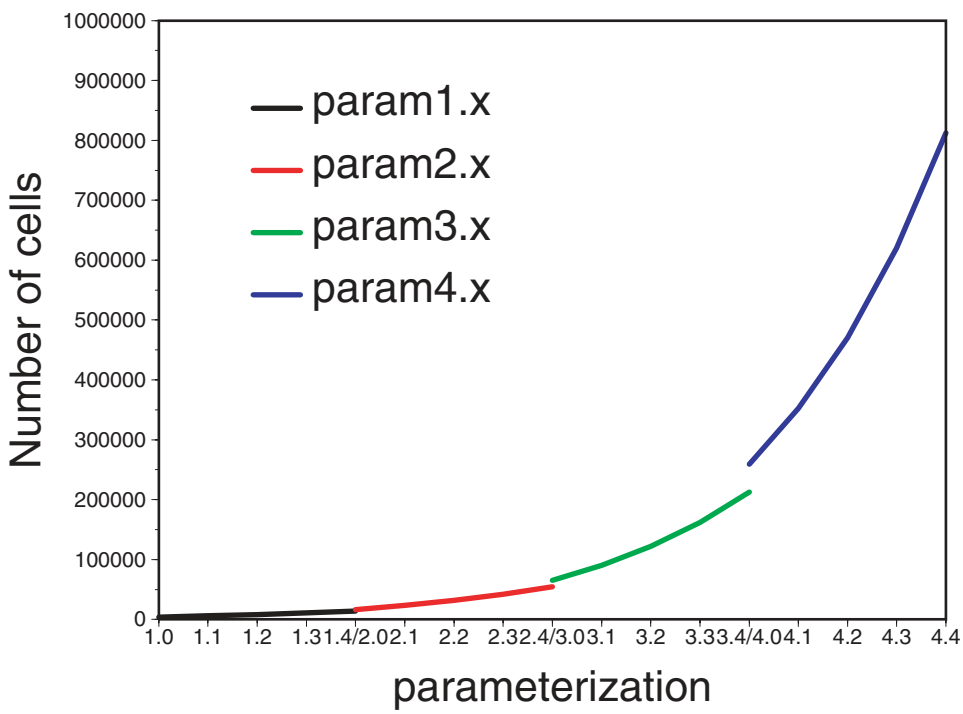

Figure 7. Number of tetrahedra introduced by the adaptive algorithm for all four inversion series.

the curves do not overlap. In each case a better data fit is more easily achieved with the next finer scale regular mesh, rather than with the most detailed irregular mesh produced by the previous series, i.e., models *AP2.0*, *AP3.0* and *AP4.0* have variance reductions of 12%, 21% and 32%) respectively, while models *AP1.4*, *AP2.4* and *AP3.4* have variance reductions of 8%, 16% and 26% respectively.

[25] This result is consistent with Figure 7, which shows the number of unknowns (tetrahedra) produced by the adaptive algorithm in each case. Again there is no overlap and hence the four separate series of inversions produce a virtual continuous spectrum of mesh sizes from the most coarse, at 4056 tetrahedra (*param1.0*), to the most fine, 812686 (*param4.4*). In terms of numbers of unknowns this spans the entire range from the earliest whole Earth spherical harmonic inversions [*Dziewonski*, 1984] to the most recent "high resolution" cellular parameterizations [*Bijwaard et al.*, 1998; *Karason and van der Hilst*, 2001].

[26] One clear feature of Figure 6 is the lack of any improvement in data fit in the final stage of each series, e.g., model *AP4.4* (37.8% variance reduction) gave virtually no improvement in data fit over model *AP4.3* (37.7% variance reduction), even though considerable numbers of extra tetrahedra are introduced (see Figure 7). This suggests that little could be achieved by further refinement in each stage. Indeed it does not appear possible (with the current mechanism) to produce a single adaptive refinement which could span the full range of length scales, e.g., from *param1.0* to say *param3.4*. It is not clear if this effect is simply an artifact of the current method, or whether it reflects the information content of the data, i.e., it may suggest that the shortest length scales of heterogeneity in the Earth (to which the data fit is sensitive) are rather broadly distributed. [For a discussion of spatial scales of heterogeneity see *Gudmundsson et al.*, 1990; *Davies et al.*, 1992.]

[27] Another point to note is that previous authors [*van der Hilst et al.*, 1997; *Widiyantoro*, 1997] obtained a 47% variance reduction with the same data set, where as we have a maximum of around 38%. The differing levels of data fit achieved will be influenced by differences in both parameterization and inevitably damping, however, the earlier work also included almost 30,000 cluster event relocation parameters which, may well be responsible for this difference.

[28] Apart from examining the trends in data fit and number of unknowns, we can also assess performance by viewing the velocity models and adaptive meshes directly. Since there are 16 3-D models and tetrahedral meshes in total we only present a subset of the results here to illustrate the main features of the adaptive algorithm. (Some animations and a facility for interactive visualization of whole mantle *P* wavespeed model *AP4.3* are presented in the appendix.)

[29] It might be argued that interpreting the results of tomography using an irregular mesh is more difficult than with a regular mesh. Usually one might be suspicious when trends in the velocity model are similar to those in the density of the mesh. However, with adaptive tomography the mesh is not imposed a priori but instead its deformation is part of the inversion process. We therefore regard the parameterization as part of the solution. Trends in mesh density should correlate with velocity perturbations, because they are both being driven by the signal in the data.

[30] Figures 8 and 9 show slices, at a depth of 1300 km, through the meshes *param4.0* to *param4.3*. Figure 10 shows the same slice through the wavespeed models *AP4.0* and *AP4.3*. It is evident that the adaptive algorithm has introduced more tetrahedra in regions where the velocity gradients are highest. In particular, notice, the finer detail in regions occupied by the Farallon and Tethys subductions, as well as the "great African plume." These are clearly identified in *param4.1* and subsequently enhanced by further refinement in meshes *param4.2* and *param4.3*. By the final stage *param4.3* many other smaller scale features have been selected and the irregularity of the mesh has spread across the globe. Clearly the adaptive nature of the parameterization has successfully identified regions of rapid velocity change, and progressively refined the 3-D structure in these regions.

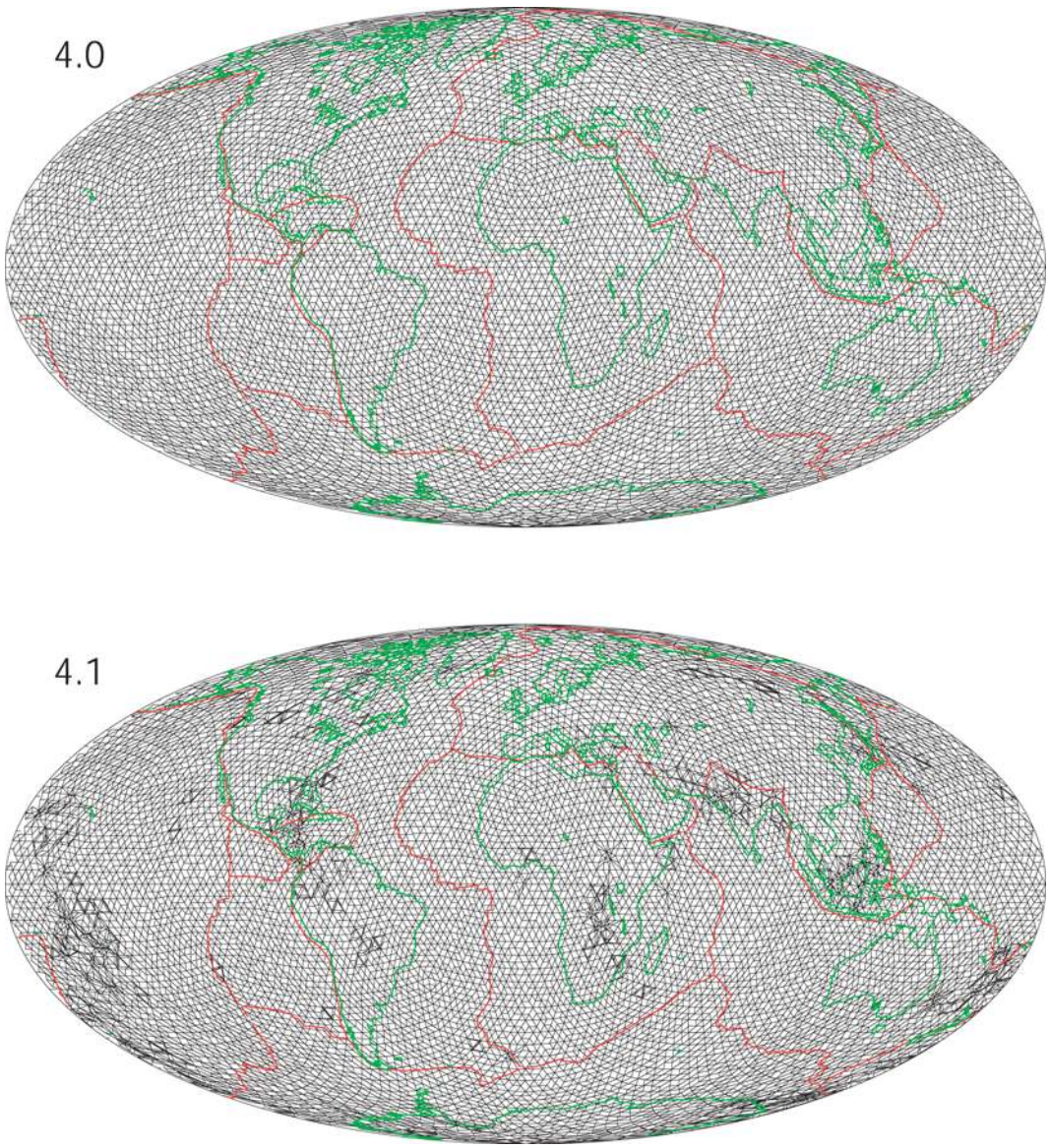

Figure 8. A slice at 1300 km depth through the tetrahedral meshes produced by the adaptive parameterization. The upper panel shows the uniformly distributed tetrahedra of *param4.0*. The lower panel shows the same slice for the mesh *param4.1*.

[31] Figure 10 shows the initial (*AP4.0*) and final (*AP4.3*) preferred velocity model for the 4-th series. (We reject model *AP4.4* because it contains considerably larger perturbations which do not significantly improve the data fit.) At a depth of 1300 km model *AP4.0* bears a strong similarity to the whole mantle *P*-wavespeed model of *van der Hilst et al.* [1997], especially away from the poles. This is unsurprising since the underlying data set is almost identical. The appearance of tomographic images constructed from irregular tetrahedral meshes will differ from the now familiar images generated with cubic blocks, simply because of the nature of the complex 3-D mesh. We must therefore be careful not to mis-interpret parameterization effects in terms of structural information.

[32] Comparing models *AP4.0* to *AP4.3*, we notice that the smooth features of model *AP4.0* are consistent with the finer detail in model *AP4.3*. Model *AP4.0* looks like a low pass version of *AP4.3*.

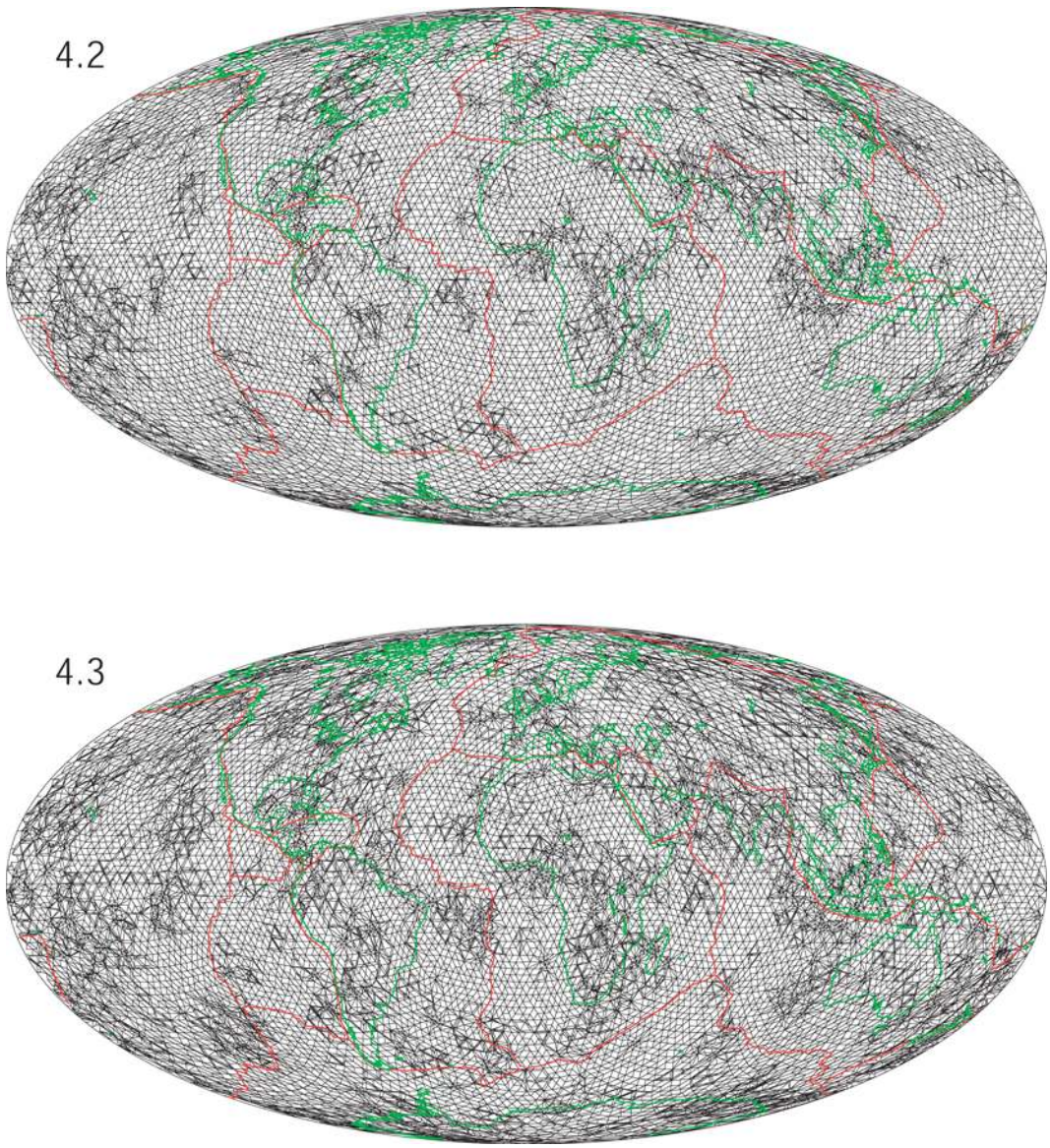

Figure 9. The upper panel shows a slice at 1300 km depth through the the mesh *param4.2*, and the lower panel shows the next irregular mesh *param4.3*.

Comparing *P*-wavespeed models *AP4.0* to *AP4.3*, the general trend appears to be that as the parameterization is refined both globally and locally, the Tethys and Farallon subduction features become progressively narrower, with increasing lateral velocity gradients. The same is also true of the more "plume" like slower anomaly in southern Africa. Moreover the broad patterns of heterogeneity observed in model *AP4.0* become more broken up with 'spots' of higher amplitude perturbation in model *AP4.3*.

[33] The adaptive algorithm does not appear to have been successful everywhere. The broad low velocity regions in the mid pacific at 1300 km depth, would appear to be artifacts, since the ray sampling there is relatively low. The general pattern of anomalies in the models *AP4.0* and *AP4.3* are similar to those found in previous models. Notice, from Table 1, that *param4.0* has a similar overall number of unknowns to the regular cubic parameterization of *van der Hilst et al.* [1997]. *param4.4* has many more unknowns, which are

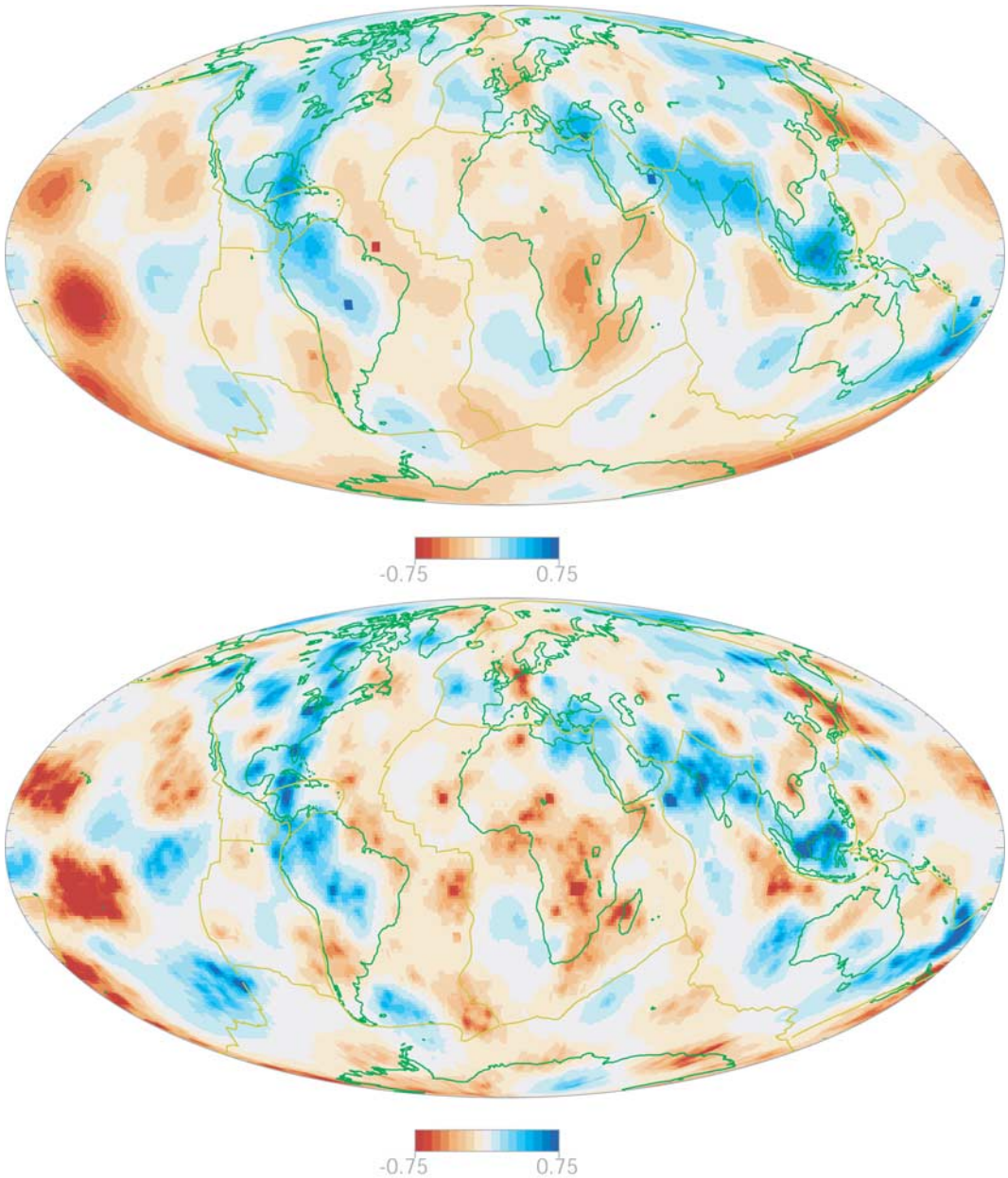

Figure 10. The upper panel shows a slice at 1300 km depth through the velocity model *AP4.0*, and the lower panel through the model *AP4.3*. Both velocity models have been smoothed over length scales of 200 km laterally and 50 km in depth to aid visual interpretation.

concentrated in the heterogeneous regions identified by the algorithm.

[34] Figures 11, 12 and 13 show enlargements of the meshes *param4.0*, *param4.4* and models *AP3.0*, *AP3.4*, *AP4.0* and *AP4.3*, at 1300 km depth in the region of the subducted Tethys ocean. These images show more clearly the progression towards thinner subduction zones as the parameterization is refined. Also the higher amplitudes and the irregularity of the fine scale structure become more apparent. Since the refined parameterization is automatically chosen by the algorithm we would conclude that the thinner subduction zones are necessary to fit the data, and in this sense they are well resolved.

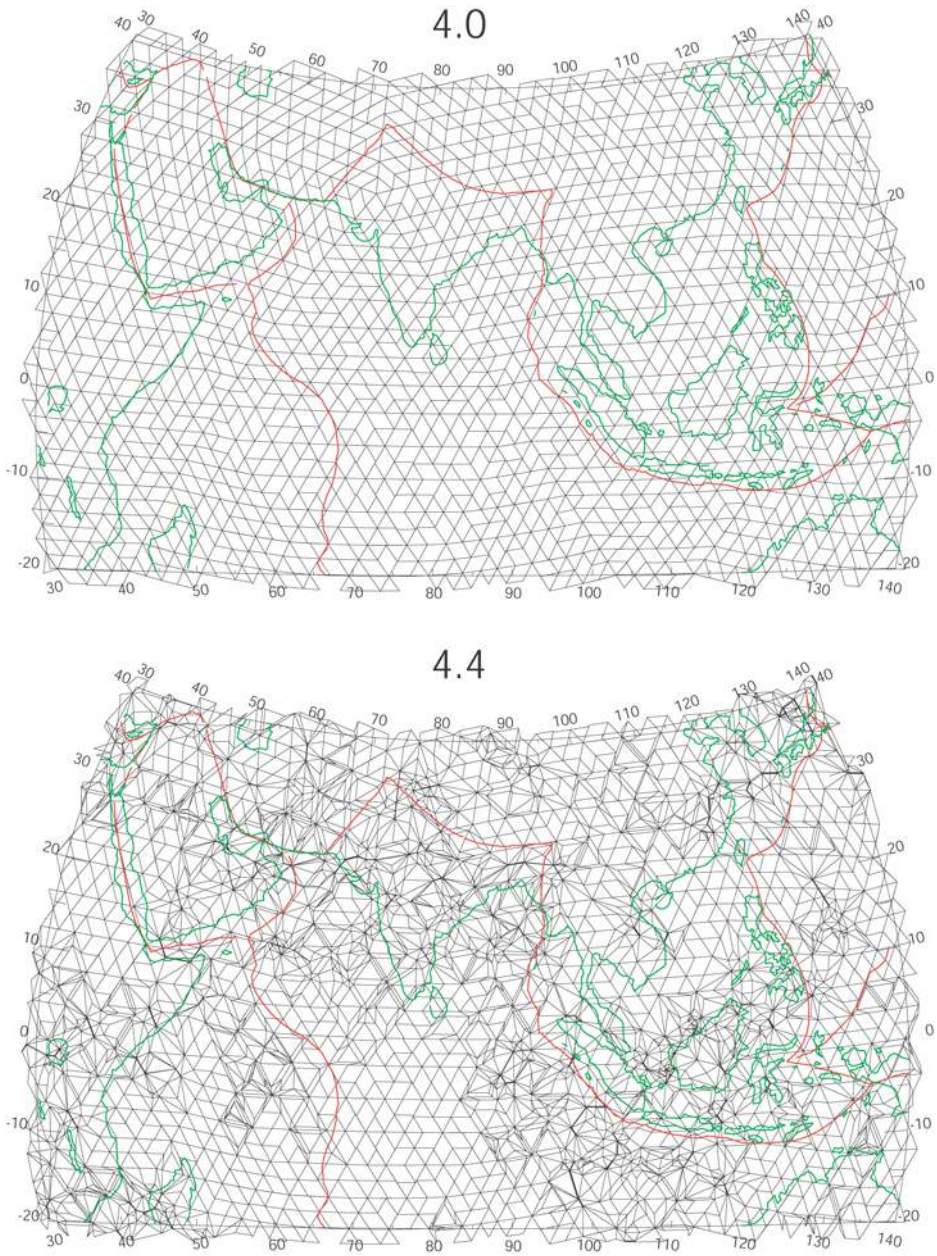

Figure 11. The upper panel shows a slice through mesh *param4.0*, and lower panel through *param4.3* for the region containing the Tethys ocean, both at 1300 km depth.

[35] Inevitably as one introduces a locally fine scale parameterization the variation in ray path density will influence the results. This may be a factor in the appearance of the localized "spot like" anomalies seen in Figures 12 and 13, which presumably result from local instabilities due to lack of smoothing in these areas. Nevertheless our results lend support to the resolvability of the narrow subduction like structures seen in the other models obtained with static parameterizations [e.g., *Fukao et al.*, 1992; *Grand*, 1994; *van der Hilst et al.*, 1997; *Bijwaard et al.*, 1998; *Gorbatov et al.*, 2000].

4. Discussion

[36] We have presented a new approach to tomography which makes use of a self-adaptive parameterization, and illustrated it with application to a whole Earth summary ray data set. Through a series of numerical tests we illustrate what can be

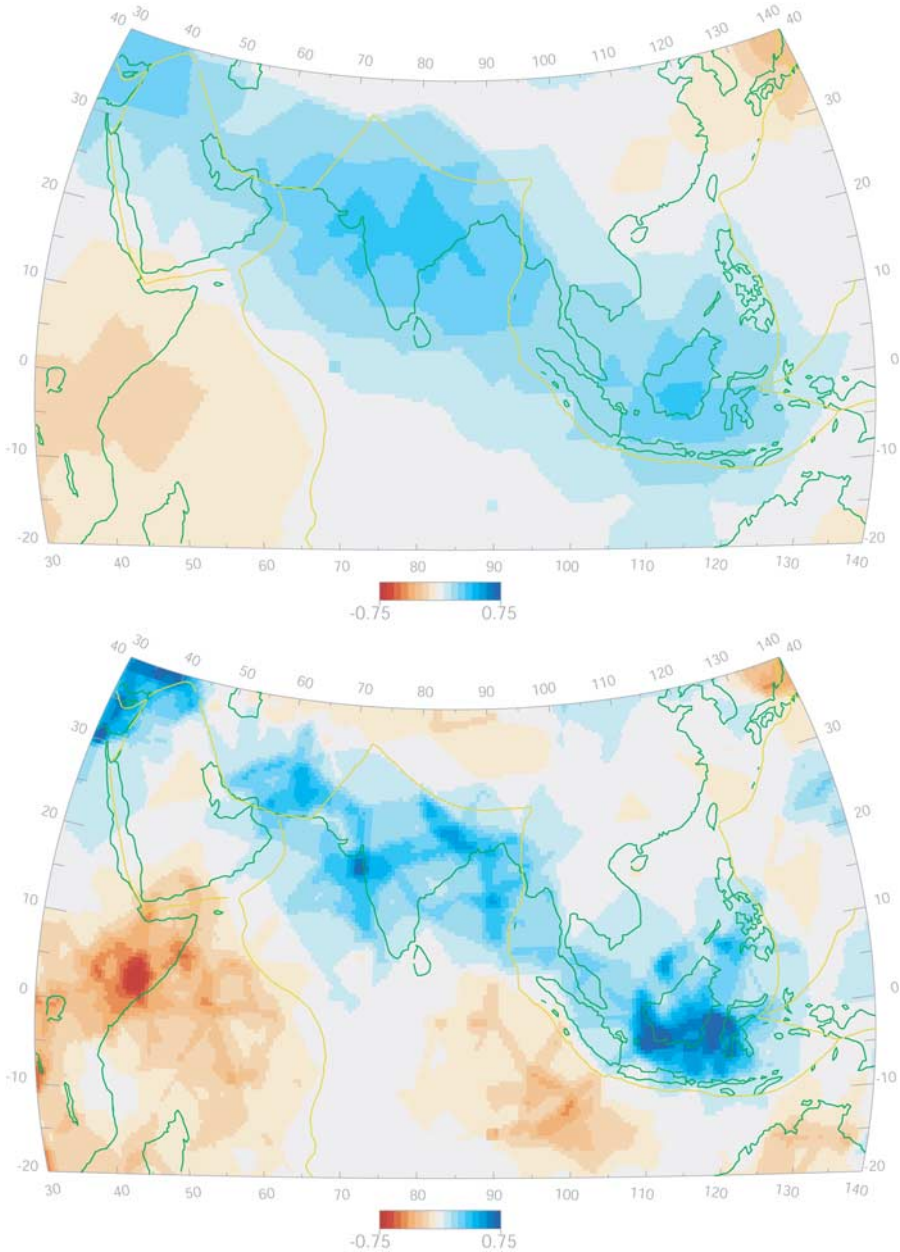

Figure 12. The upper panel shows a slice through the velocity model *AP3.0* obtained with mesh *param3.0*, and lower panel through model *AP3.4* for the Tethys ocean region, both at 1300 km depth.

achieved with such an approach, while highlighting the potential pitfalls.

[37] Although the application here is to linearized tomography, the tetrahedral mesh is equally suited to non-linear tomography [*Widiyantoro et al.*, 2000; *Bijwaard and Spakman*, 2000]. Ray tracing through 3-D models built from Delaunay tetrahedra with constant velocities is, in principle, no more difficult than through cubic cells. The basic geometric problem to be solved in each case is to find the cell/tetrahedron containing a given point (r, θ, φ). Efficient tools described by *Sambridge and Gudmundsson* [1998], are available for this purpose, and so numerical ray tracing and ray length calculations can be performed in a straightforward manner.

[38] An area where the current approach might be improved is in the criteria for mesh refinement. Here we simply use the gradient measure (1) which is

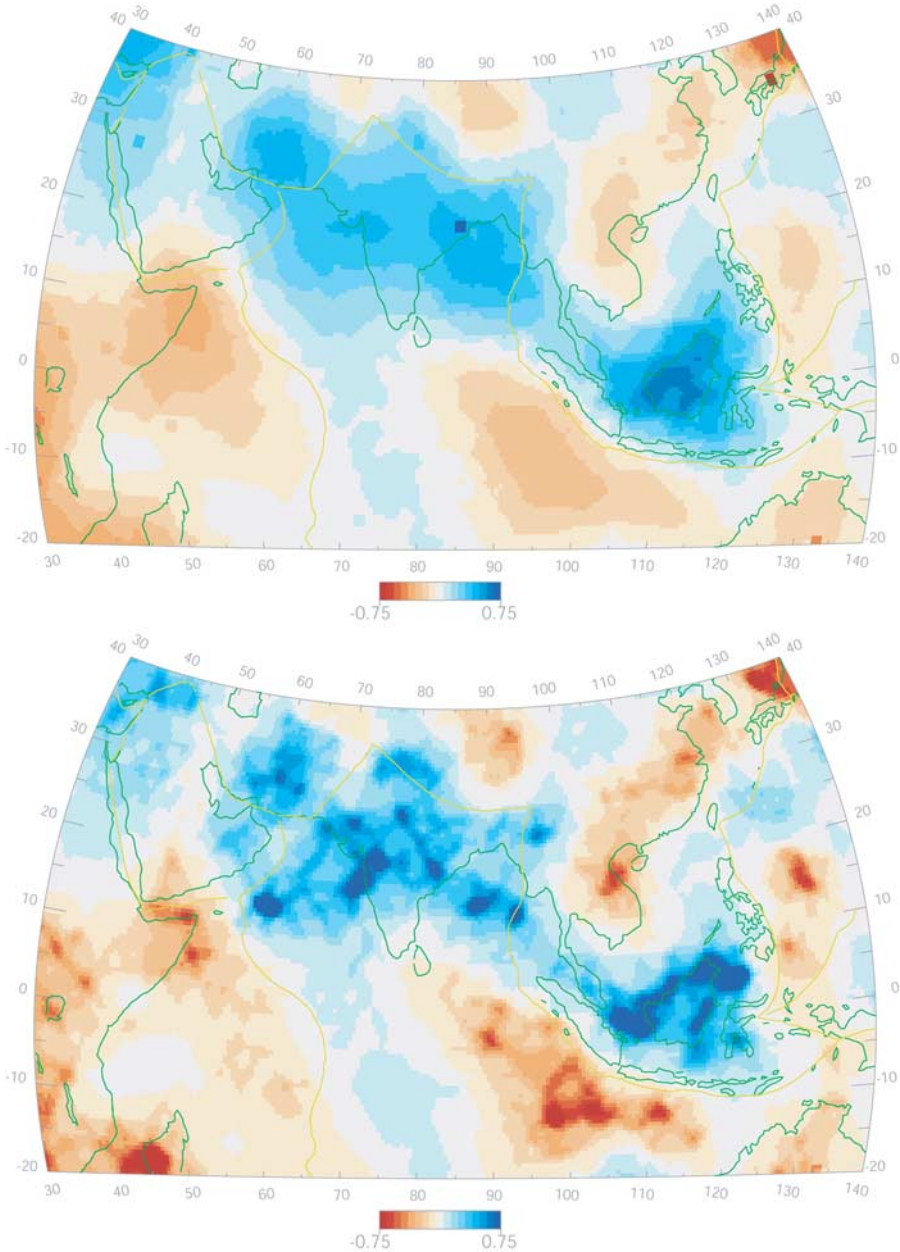

Figure 13. The upper panel shows a slice through the velocity model *AP4.0* obtained with mesh *param4.0*, and lower panel through model *AP4.3* for the Tethys ocean region, both at 1300 km depth.

straightforward to calculate, and clearly identifies the regions where the model is changing most. However the regions of high velocity gradient do not necessarily correspond to regions of good ray coverage. It may be worthwhile to consider some form of model resolution or covariance measure, to determine the regions for local refinement. Procedures for calculating these quantities have recently been proposed for large scale tomographic problems [*Nolet et al.*, 1999; *Yao et al.*, 1999]. Although it is not yet clear whether they are efficient and accurate enough for use in data adaptive tomography [*Nolet et al.*, 2001; *Yao et al.*, 2001].

[39] A major criticism we would have about our algorithm is that it is possibly too "enthusiastic" in that, on average, one unknown is replaced by 8 more in each sub-division of a tetrahedron. This appears to limit the range of distance scales over which the tetrahedral mesh can be locally refined

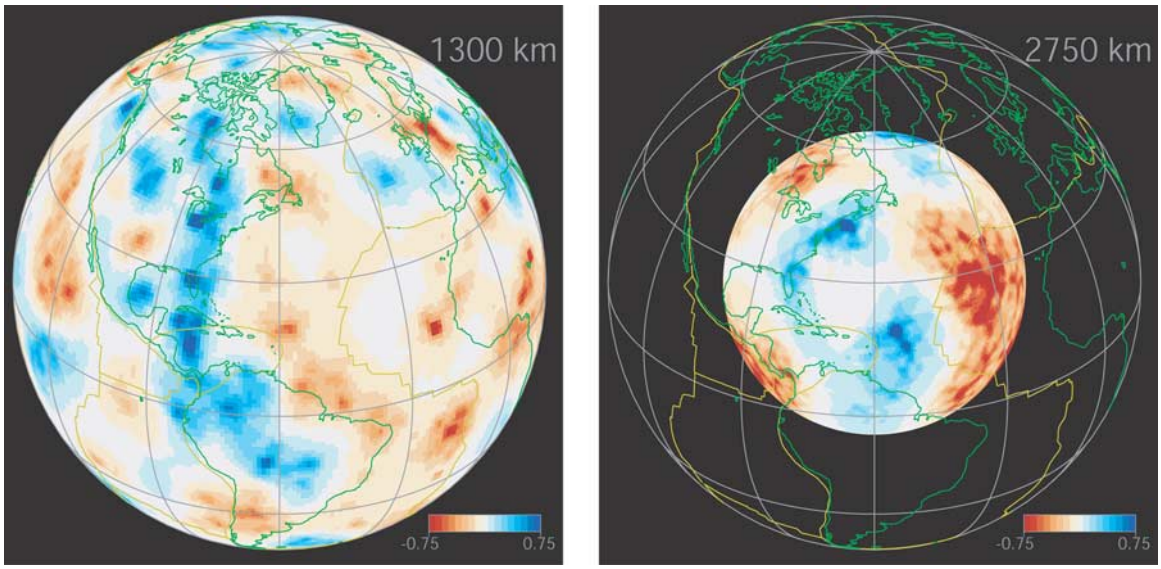

Figure A1. Two images from the gif animation files showing depth slices at 1300 km and 2750 km through model *AP4.3*.

while reducing the globally sensitive data variance measure. A less aggressive algorithm, (i.e., one introducing less unknowns per step) may allow a greater range of distance scales to be covered in a single application. This remains a direction for further study.

[40] As with all large scale tomographic studies this work lacks any formal (linearized) estimates of uncertainty. The usefulness or otherwise of checkerboard tests of "resolution" is not a debate we enter into here. We have opted to examine the performance of the images by looking at the combined results of models and meshes. Application of the adaptive algorithm to a global data set has shown that it is able to concentrate detail in sub-regions identified by the data itself. Prominent features of the mid-mantle seen in previous studies using static meshes are also seen here. We observe a clear improvement in the definition of structural features as the parameterization evolves, together with an increase in amplitude of perturbation away from a reference model.

[41] Tomography using static irregular parameterizations based on a priori information, e.g., ray density, are receiving increasing attention [*Bijwaard et al.*, 1998; *Karason and van der Hilst*, 2001; *Spakman and Bijwaard*, 2001]. A self adaptive approach which responds directly to the tomographic image obtained at each iteration might be viewed as a natural progression, but clearly more work will be required before all of its nuances are understood. The inclusion of 3-D ray tracing with a self-adaptive parameterization may be an interesting direction for further research.

Appendix A: Visualization of Whole Earth Tomography Model *AP4.3*

[42] To aid visualization of the whole mantle model *AP4.3* we have produced a set of four GIF animations, two "snapshots" of which are shown in Figure A1. These are attached to this paper and may also be downloaded directly from http://rses.anu.edu.au/seismology/projects/tireg. Each animation shows a sequence of contoured slices which cycle through either depth or longitude. An orthographic projection is used with the central point either in the northern or southern hemisphere. In addition we also include model *AP4.3* in virtual reality modelling language (VRML) format, which may also be accessed through the URL above. Figure A2 shows a snapshot example. The VRML format allows interactive viewing of the model as a function of latitude, longitude and depth. (VRML viewers are available for a range of computer platforms.) In this case the

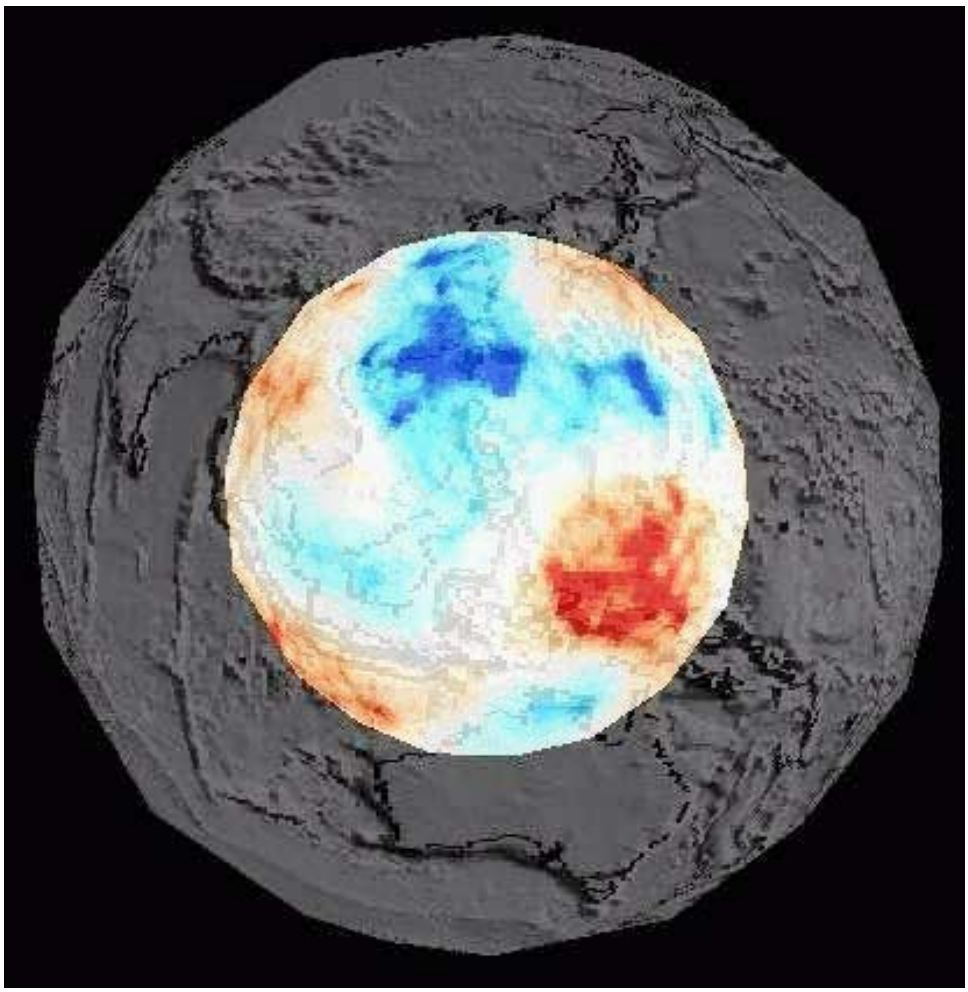

Figure A2. A snapshot of a VRML view of model *AP4.3*. This format allows interactive examination of the model in three dimensions.

model has been re-sampled on to a $1^\circ \times 1^\circ$ grid on a series of 18 constant depth slices throughout the mantle. This re-sampling lowers the resolution of the model but greatly assists speed of down load and interactivity in viewing of the 3-D tomographic model.

Acknowledgments

[43] Discussions with B. L. N. Kennett and A. Gorbatov have been helpful in shaping this study. S. Widiyantoro and R. van der Hilst kindly provided the summary ray data set used. J. Braun wrote the software which produced VRML files for interactive examination of tomographic models. Constructive criticisms on an earlier version of this manuscript were received from B. L. N. Kennett and N. Rawlinson, W. Spakman and two anonymous reviewers. Their input is much appreciated. Additional material from this study, including animations and R. Faletič's honours thesis, may be obtain via http://rses.anu.edu.au/seismology/projects/tireg. Software implementing the quickhull algorithm is acknowledged, and available from the Geometry Centre at http://www.geom.umn.edu/software/qhull.

References

Barber, B., D. P. Dobkin, and H. Huhdanpaa, The Quickhull Algorithm for Convex Hull, *Tech. Rep. GCG53*, The Geometry Cent., Univ. of Minnesota, Minneapolis, 1993.

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.

Chiao, L. Y., and B. Y. Kuo, Multiscale seismic tomography, *Geophys. J. Int.*, *145*, 517–527, 2001.

Chou, C. W., and J. R. Booker, A Backus-Gilbert approach to inversion of travel time data for three-dimensional velocity structure, *Geophys. J. R. Astron. Soc.*, *59*, 325–344, 1979.

Constable, C. G., R. L. Parker, and P. B. Stark, Geomagnetic field models incorporating frozen flux constraints, *Geophys. J. Int.*, *113*, 419–433, 1993.

Curtis, A., and R. Snieder, Reconditioning inverse problems using the genetic algorithm and revised parameterization, *Geophysics*, *62*, 1524–1532, 1997.

Davies, J. H., O. Gudmundsson, and R. W. Clayton, Spectra of Mantle shear wave velocity structure, *Geophys. J. Int.*, *108*, 865–882, 1992.

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., Global seismic tomography of the mantle, *Rev. Geophys.*, *33*, 419–423, 1995.

Engdahl, E. R., R. D. van der Hilst, and R. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Am.*, *88*, 722–743, 1998.

Faletič, R., Seismic Tomography with a self-adaptive parameterisation, Honours thesis, Australian National Univ., 1997.

Fukao, Y., M. Obayashi, H. Inoue, and M. Nenbai, Subducting slabs stagnant in the mantle transition zone, *J. Geophys. Res.*, *97*, 4809–4822, 1992.

Gorbatov, A., S. Widiyantoro, Y. Fukao, and E. Gordeev, Signature of remnant slabs in the North Pacific from *P*-wave tomography, *Geophys. J. Int.*, *142*, 27–36, 2000.

Grand, S. P., Mantle shear structure beneath the Americas and surrounding oceans, *J. Geophys. Res.*, *99*, 11,591–11,621, 1994.

Grand, S. P., R. D. van der Hilst, and S. Widiyantoro, Global seismic tomography: A snapshot of convection in the Earth, *GSA Today*, *7*, 1–7, 1997.

Gudmundsson, O., and M. Sambridge, A regionalized upper mantle (RUM) seismic model, *J. Geophys. Res.*, *103*, 7121–7136, 1998.

Gudmundsson, O., J. H. Davies, and R. W. Clayton, Stochastic-analysis of global travel time data: Mantle heterogeneity and random errors in the ISC data, *J. Geophys. Res.*, *102*, 25–43, 1990.

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 (Eds.), *Seismic Tomography—Theory and Practice*, pp. 9–22, Chapman and Hall, New York, 1993.

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-diff), *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 travel times, *Geophys. J. Int.*, *122*, 108–124, 1995.

Michelini, A., An adaptive-grid formalism for travel time tomography, *Geophys. J. Int.*, *121*, 489–510, 1995.

Nolet, G., Seismic wave propagation and seismic tomography, in *Seismic Tomography: With Application in Global Seismology and Exploration Geophysics*, edited by G. Nolet, pp. 1–23, D. Reidel, Norwell, Mass., 1987.

Nolet, G., Solving large linearized tomographic problems, in *Seismic Tomography: Theory and Practice*, edited by H. M. Iyer and K. Hirahara, chap. 9, pp. 227–247, Chapman and Hall, New York, 1993.

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.

Nolet, G., R. Montelli, and J. Virieux, Reply to comment by Z. S. Yao, R. G. Roberts and A. Tryggvason on "Explicit, approximate expressions for the resolution and a posteriori covariance of massive tomographic systems", *Geophys. J. Int.*, *145*, 315, 2001.

Okabe, A., B. Boots, and K. Sugihara, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, John Wiley, New York, 1992.

Ritzwoller, M. H., and E. M. Lavely, Three-dimensional seismic models of the Earth's mantle, *Rev. Geophys.*, *33*, 1–66, 1995.

Romanowicz, B., Seismic tomography of the Earths mantle, *Ann. Rev. Earth Planet. Sci.*, *19*, 77–99, 1991.

Sambridge, M., and O. Gudmundsson, Tomographic systems of equations with irregular cells, *J. Geophys. Res.*, *103*, 773–781, 1998.

Sambridge, M., J. Braun, and H. McQueen, Geophysical parametrization and interpolation of irregular data using natural neighbours, *Geophys. J. Int.*, *122*, 837–857, 1995.

Spakman, W., and H. Bijwaard, Optimization of cell parameterizations for tomographic inverse problems, *Pure Appl. Geophys.*, *158*, 1401–1423, 2001.

Tarantola, A., and A. Nercessian, Three-dimensional inversion without blocks, *Geophys. J. R. Astron. Soc.*, *76*, 299–306, 1984.

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

Wang, Z., and F. A. Dahlen, Spherical-spline parameterization of three dimensional earth models, *Geophys. Res. Lett.*, *22*, 3099–3102, 1995.

Wang, Z., J. Tromp, and G. Ekstrom, Global and regional surface-wave inversions: A spherical-spline parameterization, *Geophys. Res. Lett.*, *25*, 207–210, 1998.

Watson, D. F., *Contouring: A Guide to the Analysis and Display of Spatial Data*, Pergamon, New York, 1992.

Widiyantoro, S., Studies of seismic tomography on regional and global scale, Ph.D. thesis, Australian National Univ., 1997.

Widiyantoro, S., and R. D. van der Hilst, Mantle structure beneath Indonesia inferred from high-resolution tomographic imaging, *Geophys. J. Int.*, *130*, 167–182, 1997.

Widiyantoro, S., A. Gorbatov, B. L. N. Kennett, and Y. Fukao, Improving global shear wave traveltime tomography using

three-dimensional ray tracing and iterative inversion, *Geophys. J. Int.*, *141*, 747–758, 2000.

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.

Yao, Z. S., R. G. Roberts, and A. Tryggvason, Comment on "Explicit, approximate expressions for the resolution and a posteriori covariance of massive tomographic systems" by G. Nolet, R. Montelli and J. Virieux, *Geophys. J. Int.*, *145*, 307–314, 2001.