Inversion of massive surface wave data sets: Model construction and resolution assessment

Eric Debayle

Institut de Physique du Globe de Strasbourg, Ecole et Observatoire des Sciences de la Terre, Centre National de la Recherche Scientifique, and Université Louis Pasteur, Strasbourg, France

Malcolm Sambridge

Centre for Advance Data Inference, Research School of Earth Sciences, Australian National University, Canberra, Australia

Received 24 June 2003; revised 30 October 2003; accepted 28 November 2003; published 26 February 2004.

[1] A new scheme is proposed for the inversion of surface waves using a continuous formulation of the inverse problem and the least squares criterion. Like some earlier schemes a Gaussian a priori covariance function controls the horizontal degree of smoothing in the inverted model, which minimizes some artifacts observed with spherical harmonic parameterizations. Unlike earlier schemes the new approach incorporates some sophisticated geometrical algorithms which dramatically increase computational efficiency and render possible the inversion of several tens of thousands of seismograms in few hours on a typical workstation. The new algorithm is also highly suited to parallelization which makes practical the inversion of data sets with more than 50,000 ray paths. The constraint on structural and anisotropic parameters is assessed using a new geometric approach based on Voronoi diagrams, polygonal cells covering the Earth's surface. The size of the Voronoi cells is used to give an indication of the length scale of the structures that can be resolved, while their shape provides information on the variation of azimuthal resolution. The efficiency of the scheme is illustrated with realistic uneven ray path configurations. A preliminary global tomographic model has been built for *SV* wave heterogeneities and azimuthal variations through the inversion of 24,124 fundamental and higher-mode Rayleigh waveforms. Our results suggest that the use of relatively short paths (<10,000 km) in a global inversion should minimize multipathing, or focusing/defocusing effects and provide lateral resolution of a few hundred kilometers across the globe. *INDEX TERMS:* 7218 Seismology: Lithosphere and upper mantle; 7255 Seismology: Surface waves and free oscillations; 7260 Seismology: Theory and modeling; 8180 Tectonophysics: Tomography; *KEYWORDS:* surface wave, global tomography, resolution Voronoi

Citation: Debayle, E., and M. Sambridge (2004), Inversion of massive surface wave data sets: Model construction and resolution assessment, *J. Geophys. Res.*, *109*, B02316, doi:10.1029/2003JB002652.

1. Introduction

[2] The dramatic increase in the number of seismic stations over the past 20 years has led to modern global surface wave tomographic models being constrained by several tens of thousands of seismograms [e.g., *Su et al.*, 1994; *Trampert and Woodhouse*, 1995; *Ritsema et al.*, 1999]. With such large amounts of data the assessment of resolution can become problematic [*Nolet et al.*, 1999]. For massive data sets (i.e., >30,000 ray paths) the distribution of rays is also far from regular due to the uneven distribution of stations and seismic events [e.g., see *van Heijst and Woodhouse*, 1999]. An area of particularly poor coverage is the Southern Hemisphere where few permanent stations are available for global tomography. In such regions the number of independent data available can be much smaller than the number of independent parameters sought, and in these cases a sophisticated definition of the a priori information can help reduce artifacts in the inverted model.

[3] In most linear imaging problems involving large data sets, Earth structure is usually expanded into a set of global basis functions truncated at some arbitrary level, e.g., spherical harmonics. When the truncation is crude, side lobes in the averaging kernels can produce artifacts in the inverted models. In order to minimize such effects, *Whaler and Gubbins* [1981] proposed down-weighting of the coefficients as the degree of the spherical harmonic increased, i.e., a form of taper. An alternative approach is to use smooth local basis functions that behave as natural tapers. Such an approach was developed by *Montagner* [1986]. His "continuous regionalization" method was used to constrain velocity and azimuthal anisotropy from regional path average measurements and is derived from the continuous form of an inversion approach developed by *Tarantola and Valette* [1982]. In this case a smooth model is obtained by

Copyright 2004 by the American Geophysical Union.
0148-0227/04/2003JB002652$09.00

imposing correlation between neighboring points in the Earth through a Gaussian prior covariance function. In this way a natural taper is produced, which allows potential ringing phenomenon to be minimized. The use of a Gaussian covariance function also deals in a natural way with the effect of uneven data sampling. Indeed, the *Tarantola and Valette* [1982] least squares approach can be seen as a way of finding the model that gives the best fit to the data while keeping as "close" as possible to the a priori information. The smoothness of the inverted model in poorly sampled regions is therefore mostly constrained by the width of the Gaussian covariance function, while in regions with higher ray density the need for a satisfactory data fit results in a rougher model.

[4] Another attractive feature of this approach is that the prior information imposed on the inverse problem has a direct interpretation in terms of physical properties of the model. The Gaussian covariance function is defined by a prior variance $\sigma(r)$ and a horizontal correlation length L_{corr} (see equation (5)). These control the allowable perturbations of the velocity model in terms of amplitude and lateral length scale, respectively. The most reasonable way to chose L_{corr} is to make sure that the overlap between the surfaces of width $2L_{\mathrm{corr}}$ centered around each of the ray paths ensures a good coverage of the area under study and that L_{corr} is larger than the wavelength at the period considered [*Montagner*, 1986]. For large data sets it is expected that the first condition is fulfilled when L_{corr} corresponds to the wavelength of the considered period, and one might argue that the choice of L_{corr} is more physically based than choosing damping parameters in an iterative inversion scheme.

[5] These characteristics make the *Montagner* [1986] approach robust and well adapted for underdetermined problems where the lack of information in the data is compensated by the a priori information on the model. The main drawback of the continuous regionalization approach has been the computational cost required to determine the a priori model covariance function. Initially, it was designed for regional applications at the scale of a tectonic plate which involved relatively few data. By the end of the 1980s, applications typically included several hundred seismograms and could be performed on a single processor workstation. However, the computational cost of the original *Montagner* [1986] method scales with the square of the number of data N and can become impractical when the number of seismograms exceeds a few thousand. Modern global surface wave data sets often involve several tens of thousands of seismograms [*Ritsema et al.*, 1999]. To overcome this problem, *Montagner and Tanimoto* [1990] introduced several simplifications to the original method and were able to apply a continuous regionalization scheme to 2600 seismograms within a global-scale inversion. In their simplified approach the matrix to be inverted was approximated as quasi-diagonal, and the inverse was obtained using a standard series expansion. In terms of the prior model covariance function this approximation corresponds to assuming that the correlation in the Earth model between points belonging to different ray paths is null, except at points where the paths cross.

[6] In this paper, we present an efficient extension of the *Montagner* [1986] method that improves upon the N^2 run time dependence while preserving the complete use of prior information contained in the original scheme; that is, correlation between points on all pairs of ray paths is included, not just intersection points. The efficiency of the new scheme makes it practical to include several tens of thousands of seismograms in an inversion for local variations of both seismic velocity and azimuthal anisotropy. Calculations show that data sets of up to 20,000 paths can be inverted in a few hours on a single processor of a standard desktop workstation, and just a few minutes are required to process several thousand paths.

[7] The new approach is highly parallel and lends itself for use on multiprocessor machines. This will further increase the number of seismograms that can be considered routinely, thereby making the scheme well suited for construction and testing of tomographic models with modern data sets. To circumvent the difficulty in computing the resolution and a posteriori covariance of the resulting massive tomographic systems of equations [*Nolet et al.*, 1999], we also propose a procedure to assess the variability of model constraint. This is a geometric approach which makes use of Voronoi diagrams [*Okabe et al.*, 1992] to qualitatively estimate how well a given parameter can be resolved from the available ray coverage. The new approach for massive surface wave data sets is demonstrated using both synthetic experiments and a real data set of 24,124 Rayleigh wave seismograms. Software implementing the new scheme are available on request from the first author.

2. Continuous Regionalization Algorithm

[8] A classical approach in surface wave tomography is to build a three-dimensional (3-D) model of seismic velocities in two stages. The first involves a 2-D regionalization and the second involves an inversion for depth. The regionalization step is 2-D because local seismic parameters are retrieved from a set of "path average" measurements between an epicenter and a station. Phase or group velocity maps are built at different periods and then combined using depth kernels to produce fully 3-D S wave models [e.g., *Ritzwoller et al.*, 2001]. This regionalization approach can also be used to retrieve the local shear velocity at each depth from a set of 1-D depth-dependent path-average measurements [e.g., *Debayle and Lévêque*, 1997]. In this paper, we focus on the 2-D regionalization step and follow the formalism of *Tarantola and Valette* [1982] for a least squares inverse problem, where unknowns are a function of a continuous variable and where the theoretical relationship between the data and unknowns is assumed to be linear. This formalism has been applied to body waves by *Tarantola and Nercessian* [1984] and to surface waves by *Montagner* [1986]. In this section we first describe the *Montagner* [1986] approach for regional surface wave tomography before presenting our new efficient implementation.

2.1. Forward Problem

[9] In a medium where seismic parameters vary smoothly and velocity gradients are not strong, the incremental phase for a given surface wave is the integral of the phase slowness along the path, while the path itself is determined

by the lateral variations in phase velocity [*Woodhouse*, 1974]. For a source-receiver path i we have

$$1/C_i(T) = 1/L_i \int_i 1/C(T, \theta, \phi) ds, \tag{1}$$

where T is the period, θ and ϕ are the coordinates of the geographical points along the great circle, and L_i is the source receiver distance. $C_i(T)$ is the phase velocity measured along the ith path at period T, and $C(T, \theta, \phi)$ is the local phase velocity at geographical point (θ, ϕ) and period T.

[10] For longer-period waves an additional approximation is to neglect the deviations of the path from the great circle. There is a general agreement that the great circle approximation is reasonable for the fundamental mode and the first few higher modes of surface waves at periods larger than 40 s [e.g., *Kennett*, 1995; *Ritsema et al.*, 1999]. Using the great circle approximation, the inverse problem can be written in the form of a linear relationship between a data vector **d** (which contains the average slowness along each path $1/C_i(T)$) and a parameter vector **m** (which contains the local slowness $1/C(T, \theta, \phi)$ at each geographical point). Segmenting path i in segments δs, one can state in matrix form

$$\mathbf{d} = G\mathbf{m}, \tag{2}$$

where the matrix G contains the partial derivatives $\delta s/L_i$.

2.2. Inversion

[11] The simplest way to retrieve seismic heterogeneities from a set of path average surface wave measurements is to divide the Earth into tectonic provinces. This constitutes a "pure path" inversion which significantly reduces the size of the inverse problem, because the parameters at each depth correspond to only a few tectonic regions. The early regionalization techniques were based on this type of pure path inversion [*Dziewonski*, 1971; *Okal*, 1977; *Lévêque*, 1980]. The drawback of such approaches is that they require a priori the definition of boundaries between tectonic provinces, and deeper structure may not be simply related to surface tectonics in many regions of the Earth [*Montagner*, 1986; *Debayle and Lévêque*, 1997].

[12] A more flexible approach would be to avoid the a priori definition of tectonic boundaries and instead use a large number of smaller regular blocks [*Aki and Husebye*, 1977]. With a large number of blocks the inverse problem is likely to become underdetermined, and in order to stabilize the solution, a priori information on the model must be introduced. A classical least squares solution for discrete linear inverse problem is proposed by *Tarantola and Valette* [1982]

$$\hat{\mathbf{m}} = \mathbf{m}_0 + C_{m0} G^t S^{-1} (\mathbf{d} - G\mathbf{m}_0) \tag{3}$$

with

$$S = (G C_{m0} G^t + C_{d0}), \tag{4}$$

where t denotes the transpose, $\hat{\mathbf{m}}$ is the inverted model, $\mathbf{m}_0$ is the a priori model, C_{m0} represents an a priori covariance matrix on the model, and C_{d0} is an a priori covariance matrix on the data.

[13] The continuous regionalization consists then in decreasing indefinitely the size of the blocks while increasing their number toward infinity. With an "infinite" number of blocks the problem would clearly be underdetermined and require a priori constraints to be introduced to stabilize the solution. This can be done through the definition of an a priori model $\mathbf{m}_0$ and an a priori covariance function $C_{m0}(r, r')$. *Montagner* [1986] choose $C_{m0}(r, r')$ to be in the form

$$C_{m0}(r, r') = \sigma(r)\sigma(r') \exp\left(\frac{-\Delta^2_{r,r'}}{2L^2_{\text{corr}}}\right), \tag{5}$$

where r and r' are two geographical points separated by a distance Δ. The Gaussian defined through the exponential term has a standard deviation given by L_{corr}. This acts as a spatial filter, imposing correlation between points separated by distances of order L_{corr}. The function $\sigma(r)$ controls the amplitude of the model perturbation allowed at r.

[14] In the continuous form, the unknowns are functions of a continuous variable and equation (3) becomes

$$\hat{m}(r) = m_0(r) + \sum_i \left[\int C_{m0}(r, r') G_i(r') dr' . \cdot \sum_j S^{-1}_{ij} \left(d_{0j} - \int G_j(r'') m_0(r'') dr'' \right) \right] \tag{6}$$

with

$$S_{ij} = C_{d0_{ij}} + \int_i \int_j G_i(r_1) C_{m0}(r_1, r_2) G_j(r_2) dr_1 dr_2. \tag{7}$$

For propagation along the great circle, equations (6) and (7) can be written

$$\hat{m}(r) = m_0(r) + \sum_i \left[\int_{\text{path i}} \frac{ds_k}{L_i} \sigma(r_k)\sigma(r) . \exp\left(\frac{-\Delta^2_{r_k,r}}{2L^2_{\text{corr}}}\right) \cdot \sum_j S^{-1}_{ij} \left(d_{0j} - \int_{\text{path j}} G_j(r'') m_0(r'') dr'' \right) \right] \tag{8}$$

$$S_{ij} = C_{d0_{ij}} + \int_{\text{path i}} \frac{ds_k}{L_i} \int_{\text{path j}} \frac{ds_l}{L_j} \sigma(r_k)\sigma(r_l) \exp\left(\frac{-\Delta^2_{r_k,r_l}}{2L^2_{\text{corr}}}\right). \tag{9}$$

Indices i and j correspond to the ith and jth surface wave paths while k and l refer to the kth and lth points along the ith and jth paths, respectively. The estimation of $\hat{m}(\mathbf{r})$ presents several practical difficulties that become increasingly severe as the number of rays increase. The first lies in the evaluation of the integral

$$A_i(r) = \int_{\text{path i}} \frac{ds_k}{L_i} \sigma(r_k)\sigma(r) \exp\left(\frac{-\Delta^2_{r_k,r}}{2L^2_{\text{corr}}}\right), \tag{10}$$

which in the discrete case corresponds to the computation of one element of the matrix $C_{m0}G^t$ (see equations (3) and (4)).

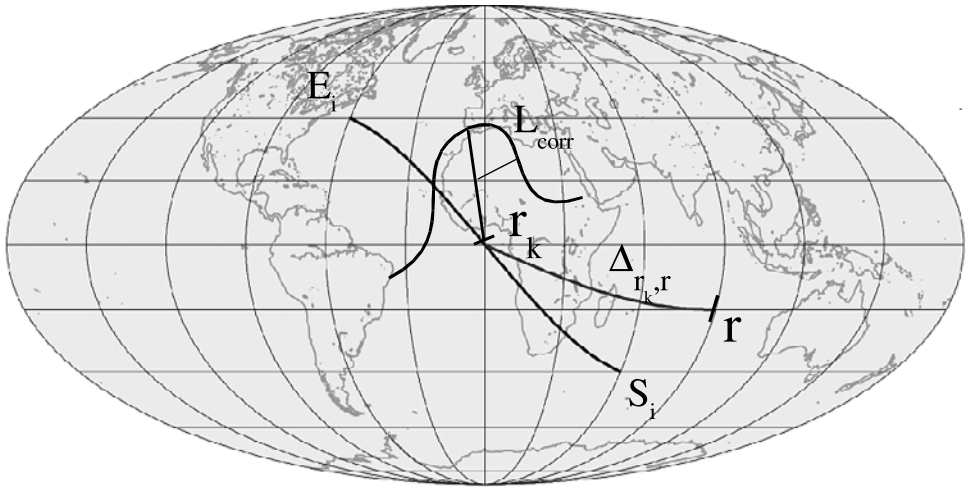

Figure 1. Estimation of $A_i(r)$ everywhere on the Earth requires the computation of $C_{m0}(r_k, r)$ between each point r_k of path i and each point r across the globe. When the distance $\Delta(r_k, r)$ between r_k and r becomes large compared to L_{corr}, the exponential term becomes negligible and its evaluation can be omitted. However, a computation of the distance $\Delta(r_k, r)$ and a comparison with L_{corr} are still required a large number of times, making the current inversion schemes impractical when the number of paths exceeds a few thousand.

Figure 1 shows how this integral can be visualized geometrically. To determine $A_i(r)$ at each geographical point r on the Earth, it is necessary to compute the correlation between the Earth model at r and the Earth model at each point r_k along the ith path. When the distance $\Delta_{r_k,r}$ is much larger than L_{corr}, the exponential term ensures that the contribution to the integral becomes negligible. Therefore, in the evaluation of $A_i(r)$ it makes sense to first compare $\Delta_{r_k,r}$ with L_{corr} so that the exponential term is only evaluated where the contribution is significant. However, even in this case, one must still evaluate $\Delta_{r_k,r}$ for all pairs of points (r, r_k) and for each path i, and this itself can become extremely time consuming for large numbers of paths. In section 3.1 we present an alternative approach that only requires the evaluation of $\Delta_{r_k,r}$ for geographical points (r, r_k) that contribute significantly to the integral. This is done by defining an "influence zone" about each path i, for which $\Delta_{r_k,r}$ is about $2.64 \times L_{\text{corr}}$, and then only collecting contributions to $A_i(r)$ from this zone. The factor of 2.64 has been chosen to truncate the Gaussian at -30 dB from its maximum following *Dziewonski et al.* [1969]. This choice allows capture of 99.18% of the surface below the Gaussian.

[15] A second practical difficulty is the computation of S_{ij} in equation (9) which requires the evaluation of the double integral

$$B_{ij} = \int_{\text{path i}} \frac{ds_k}{L_i} \int_{\text{path j}} \frac{ds_l}{L_j} \sigma(r_k)\sigma(r_l) \exp\left(\frac{-\Delta^2_{r_k,r_l}}{2L^2_{\text{corr}}}\right). \tag{11}$$

Figure 2 gives a geometrical interpretation of this integral. One sees that the model correlation between each geographical point belonging to the paths i and j is needed. However, only the points of path j located in the influence zone of path i contribute significantly to the double integral. In section 3.2 we again use the notion of the influence zone to present an approach that evaluates the double integral in equation (11) by only considering points in both influence zones of each pair of paths, i.e., again $\Delta(r_k, r_l)$ is not evaluated when r_k is outside of the influence zone of r_l. Numerical illustrations of the new procedures to evaluate $A_i(r)$ and B_{ij} are given in sections 3 and 5.

[16] A third practical problem for inverting massive surface wave data sets using equations (8) and (9) is the inversion and storage of the matrix S_{ij}. S is a square matrix whose dimension is the size of the data vector, i.e., the number of paths at a given period. With several tens of thousands of paths, inverting S becomes impractical with currently available computation. One solution is to approximate the determination of the double integral B_{ij} by considering only the case when $i = j$. In this case the resulting matrix is almost diagonal so that only its diagonal terms can be kept for the inversion. This type of approach was followed by *Montagner and Tanimoto* [1990] to successfully invert a few thousand seismograms. Here we determine S in full using a new implementation of the double integral calculation described below. We store S by utilizing a row-indexed compact sparse matrix storage technique [*Press et al.*, 1992] and avoid the need for directly calculating the inverse of S by using a conjugate gradient algorithm to determine $\boldsymbol{\alpha}_i = S^{-1}_{ij}(\mathbf{d}_{j0} - (G\mathbf{m}_0)_j)$ (i.e., we solve the system of equations $(\mathbf{d} - G\mathbf{m}_0) = S\boldsymbol{\alpha}$ for $\boldsymbol{\alpha}$).

2.3. A Posteriori Covariance and Resolution

[17] The discrete formulation of the a posteriori covariance matrix C_m and the resolution matrix R is given by *Tarantola and Valette* [1982]:

$$C_m = C_{m0} - C_{m0}G^t S^{-1} G C_{m0} \tag{12}$$

$$R = C_{m0}G^t S^{-1} G. \tag{13}$$

From equations (12) and (13) we see that evaluation of the matrices C_m and R requires the computation of the product

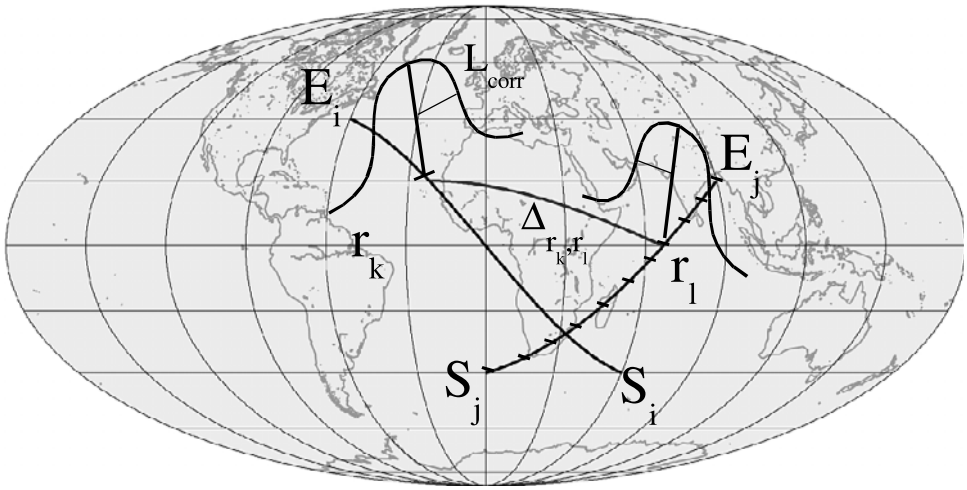

Figure 2. For each pair of ray paths (i, j) the calculation of B_{ij} involves a double integral of the cross-correlation function $C_{m0}(r_k, r_l)$ along paths i and j, where r_k is the kth point along path i and r_l is the lth point along path j. When the distance $\Delta(r_k, r_l)$ between r_k and r_l becomes large compared to L_{corr}, the exponential term in C_{m0} becomes negligible and its computation can be skipped. However, distance computations $\Delta(r_k, r_l)$ are still required for every pair of points along every pair of paths, making current inversion schemes impractical when the number of paths exceeds a few thousand.

$S^{-1}G$, with G being a $d \times m$ matrix, where d is the number of data and m is the number of model parameters. In terms of computational cost this is equivalent to solving the inverse problem m times. For example, in surface wave tomography inverting for only a single elastic parameter varying over a 1×1 degree global grid, the computation of C_m and R would cost about 64,800 times that of obtaining an inversion solution. With a large number of data this is clearly impossible. This problem is also well known in body wave tomography, and although some workable approximations for C_m and R have been proposed [e.g., *Nolet et al.*, 1999], no exact solution is currently available.

[18] In the continuous case the equivalent problem is the evaluation of the a posteriori covariance function $C_m(r, r')$,

$$C_m(r,r') = C_{m0}(r,r') - \sum_i \left[\int_{\text{path i}} \frac{ds_k}{L_i} C_{m0}(r,r_k) \sum_j \int_{\text{path j}} \cdot S_{ij}^{-1} C_{m0}(r_l,r') \frac{ds_l}{L_j} \right] \tag{14}$$

and the resolution kernel $R(r, r')$,

$$R(r,r') = \sum_i \sum_j \int_{\text{path i}} \frac{ds_k}{L_i} \frac{ds_l}{L_j} C_{m0}(r,r_k) S_{ij}^{-1} \tag{15}$$

(see *Tarantola and Nercessian* [1984] and *Montagner* [1986] for derivations). As with the discrete case, their evaluation also remains impractical for large data sets. In section 4 we propose a technique to circumvent this problem, which produces a qualitative estimation of model constraint as a function of geographical position.

3. Optimization

3.1. Optimized Computation of $A_i(r)$

[19] As described above, the computational cost of evaluating $A_i(r)$ for all paths (using equation (10)) is very large because the correlation function $C_{m0}(r_k, r)$ must be evaluated (using equation (5) for every pair of points (r_k, r), i.e., a "double loop" over every point in the model with every point along the ith path. Finally, this calculation has to be repeated for every ray path, i (see Figure 1).

[20] Our optimization strategy consists of restricting all calculations to only the pairs of points which will contribute significantly to $A_i(r)$. We define an influence zone about the ith great circle path, as containing all points r which satisfy

$$\Delta_{r_k,r} < 2.64 L_{\text{corr}}, \tag{16}$$

where r_k is any point along the ith great circle path. The influence zone defined in this way corresponds to the region where the Gaussian term in the integral (10) falls to about 3% of its maximum. Figure 3 shows an example. The complete influence zone for the ith great circle path can be thought of as the union of the influence zones of each point along the ith path, and it is this concept which forms the basis of the algorithm used to compute $A_i(r)$.

[21] The "outer loop" of the algorithm is the same as previous implementations, i.e., we consider each point r_k along the ith great circle path and evaluate its contribution to the integral for $A_i(r)$ in equation (10). In practice the entire Earth's surface is divided into a fine grid of regular cells the centers of which form the set of model points r. The integral in equation (10) becomes a summation of the exponential term $\sigma(r_k)\sigma(r)\exp(-\Delta^2_{r_k,r}/2L^2_{\text{corr}})$ for all point pairs (r, r_k). However, rather than testing each point in the entire model grid to determine which are within the influence zone (which is the approach adopted in previous schemes) we use a more sophisticated geometric search technique based on a "last in first out (LIFO) stack". LIFO stacks can be used to "map out" arbitrarily shaped continuous regions, such as our influence zone, by starting from any point within the zone and stepping between neighbors in a systematic manner. Figure 4 shows a flow diagram for

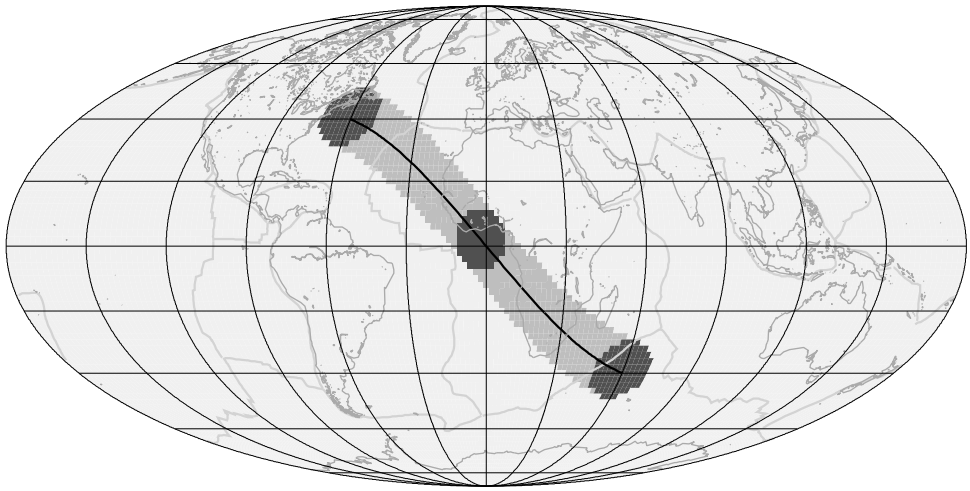

Figure 3. For each point r_k of the great circle i the algorithm in Figure 4 locates the current cell and explores the model in a region located within $2.64L_{\mathrm{corr}}$ of r_k. The "influence zone of the complete path" (in light gray) is the union of the "influence zones of each point along the great circle path" (shown in dark gray for the first, middle, and last points). In the new approach only the model contributions located within the influence zone of the rays are considered in the estimation of $A_i(r)$. Because of the use of the "mapping out" algorithm no distance evaluations, $\Delta(r_k, r_l)$, are required outside the influence zone.

the method. It can be summarized using the following six steps:

[22] 1. Let r be the center of the cell containing r_k.

[23] 2. Put r on the stack.

[24] 3. If the stack is empty, then we finish.

[25] 4. Get the current point r from top of the stack.

[26] 5. If r is in the influence zone of r_k, then (1) add contribution of pair (r, r_k) to $A_i(r)$, and (2) if the neighboring cells of r (e.g., N, S, E, W) are not already on the stack, then put them on it.

[27] 6. Go to step 3.

[28] The important point about this type of technique is that "nearly all" of the points r are in the influence zone of r_k. As soon as r falls outside of the zone, it makes no contribution to $A_i(r)$, but more importantly its neighboring cells are not placed on the stack (at step 5). This means that when the stack is empty (and the algorithm finished), distance evaluations, $\Delta_{r_k,r}$, have only been necessary for cells whose centers are either within the zone or in the immediate "ring" surrounding the zone. It turns out that this approach will work regardless of the size of the regular grid of cells and the shape of the influence zone (as long as it is continuously connected).

[29] This type of "mapping out" algorithm is common in computational geometry problems and was previously used by *Sambridge et al.* [1995] to solve similar geometric problems arising in the construction of complex 3-D parameterizations for seismic tomography. By restricting all calculations to only the influence zone of the paths, the total cost of the calculation becomes proportional to NKI, where N is the total number of paths, K is the average number of steps per path, and I is the average number of cells in the influence zone of the points along the paths, r_k. The computational cost can be compared to that of the classical *Montagner* [1986] approach which sweeps the entire Earth's surface for each point along every ray path. In that case, run time is proportional to NKX, where X is the number of cells spanning the globe. Since L_{corr} will typically be on the order of 500 km, we have $I \ll X$, and hence the new scheme will be considerably more efficient than the original (see Figure 3).

3.2. Optimized Computation of B_{ij}

[30] The evaluation of B_{ij} is essentially a double integral along every pair of paths but with an integrand which is negligible when the paths are further than a specified distance apart, which we call their common influence zone. Note that the distance between paths is itself a variable along each path, and so following a similar reasoning to previously, we need to restrict all calculations to point pairs (r_k, r_l) on the i and j paths, respectively, which are less than certain distance apart, i.e., $\Delta_{r_k,r_l} < 2.64\ L_{\mathrm{corr}}$.

[31] The new scheme begins with a preprocessing step, which consists of dividing the Earth's surface into a number of cells and determining which rays pass through each cell. This can be achieved by simply stepping down each ray j and locating the cell C containing each point r_l. The step length along the rays can be chosen independently of the cell size. A step length much smaller than the cell allows a better accuracy in sampling the cells crossed by the rays but is more expensive. However, our experience shows that a value of 2° (i.e., similar to the actual cell size) is both efficient and produces indistinguishable tomographic maps from much finer discretizations. In addition to storing the list of rays j in each cell C, we also record the list of points r_l along ray j which lie in cell C. We denote these as r_l ($l = C_{j1}, \ldots, C_{j2}$). Clearly, the computational cost of the preprocessing step scales with the number of rays and the average number of steps along each ray, i.e., NK.

[32] The main part of the algorithm consists of an outer loop over all rays, i ($i = 1, \ldots, N$). For each ray i we step along the path points, r_k ($k = 1, \ldots, K$) and use the mapping out algorithm described in Figure 4 to sample all of the cells C within the influence zone about r_k. For each cell C in the

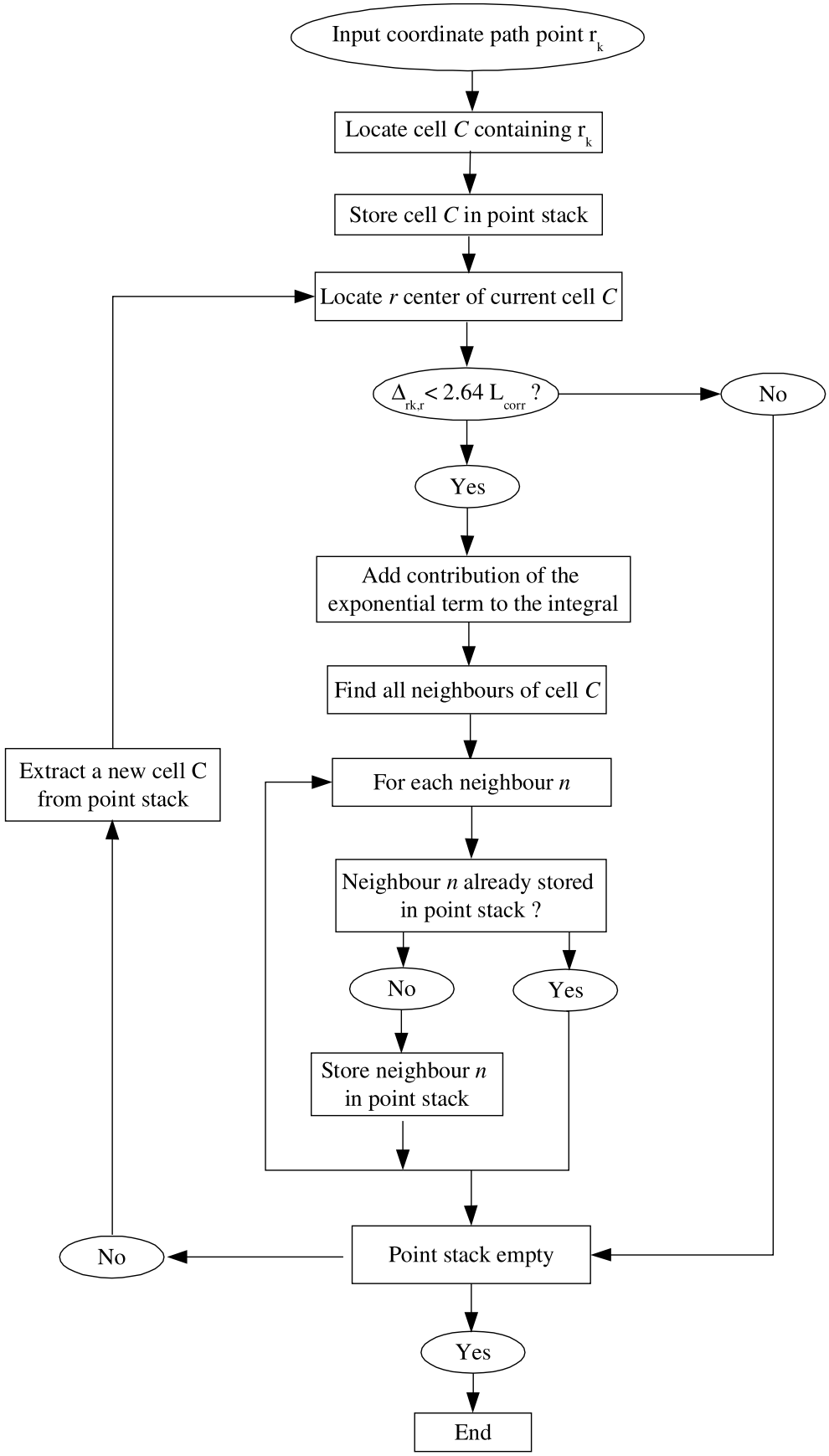

Figure 4. Flowchart of the "mapping out" algorithm used to determine the influence zone about a geographical point r_k and to calculate its contribution to $A_i(r)$.

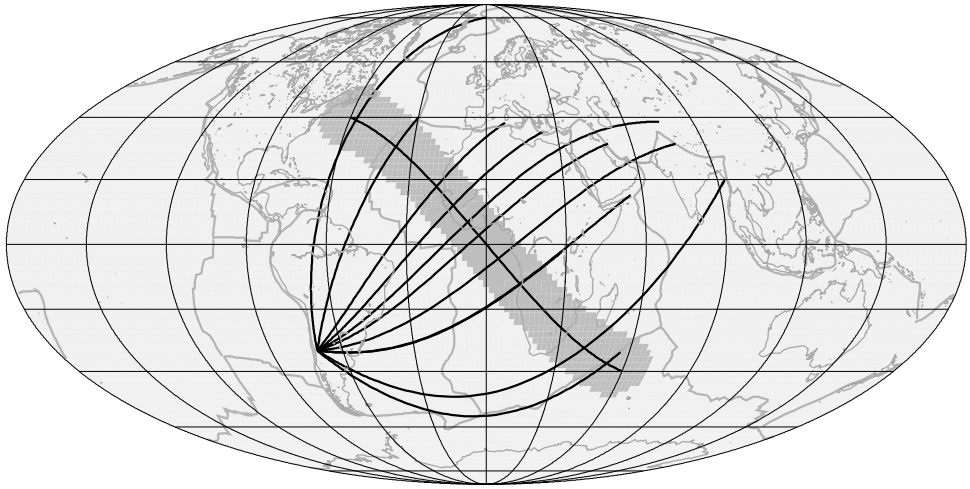

Figure 5. For each point r_k of the great circle i the algorithm in Figure 4 explores the Earth's surface within a distance $2.64L_{\mathrm{corr}}$ of r_k (region in grey). With the new scheme, only the part of paths j that are located within the influence zone of path i are considered in the estimation of B_{ij}. No computation is made outside of the influence zone.

influence zone we already know which other rays j also pass through that cell, together with their path points r_l ($l = C_{j1}, \ldots, C_{j2}$). Therefore, for every ray j and each cell C we can calculate the quantity

$$b_{ij} = \sum_{l=C_{j1}}^{l=C_{j2}} \frac{ds_k}{L_i}\frac{ds_l}{L_j}\sigma(r_k)\sigma(r_l)\exp\left(\frac{-\Delta^2_{r_k,r_l}}{2L^2_{\mathrm{corr}}}\right). \qquad (17)$$

Note that the values b_{ij} form a matrix of the same size as the matrix B formed by the values of B_{ij}. By summing the values b_{ij} for all cells C of the influence zone about r_k and then for all r_k along path i we obtain the ith row of the required matrix B. After repeating for all rays i we obtain the complete matrix B.

[33] As before, the key point is that by exploiting the mapping out algorithm, all calculations (including distance evaluations Δ_{r_k,r_l}) are restricted to the part of paths j that are located within the influence zone of path i (Figure 5).

[34] Assuming that n is the average number of rays per cell and p is the average number of steps along a ray within a cell, the computational cost for computing B_{ij} becomes proportional to $NK + NKInp$. In comparison, the Montagner original scheme involves a double loop over each point of each ray to compute the double integral in equation (11) and has computation time proportional to $(NK)^2$. Note that we always have $n \ll N$ and $p \ll K$. If N is very small and the path coverage very poor, it can be more efficient to evaluate B_{ij} using the classical Montagner approach because a large number of cells in the influence zone may not be crossed by any path. In this case, although the computation needed to estimate B_{ij} is increased, the more efficient calculation of $A_i(r)$ counteracts the effect, and overall, the new approach still remains more efficient than the classical *Montagner* [1986] scheme. As the number of data increases, say larger than a few tens of paths, we show in section 2.5 that the new algorithm for calculating B_{ij} becomes significantly more efficient.

3.3. Performance of the New Scheme

[35] We made a series of tests to compare the performance of our optimized scheme with the original one developed by *Montagner* [1986]. In all cases we generate a realistic global synthetic coverage based on permanent stations taken from the IRIS catalog and events from the Harvard centroid moment tensor catalog [*Dziewonski et al.*, 1981]. For all paths the minimum epicenter-station distance is taken as 1200 km, which is typical of surface waveform tomography. The synthetic path average measurements are generated by using the 3SMAC a priori Earth model [*Nataf and Ricard*, 1995]. The inversion is performed under the same conditions and using the same a priori values for both the optimized scheme and the *Montagner* [1986] approach.

[36] Figure 6 and Table 1 summarize the results. In all cases the velocity maps obtained using the two schemes are identical. The largest number of paths for which use of the original scheme was practical was 3090, and this took 9 hours 24 min 49 s to produce a result on a single 2.4-GHz Pentium 4 processor with 2 Gb of RAM. The optimized scheme achieved the same result in 3 min 6 s. (As seen in Table 1, decrease in computation time by more than 98% was achieved in all cases.) With the optimized scheme we have been able to invert 21,184 seismograms in 3 hours 58 min on a standard workstation. The main limitation is the amount of memory required. Inverting 21,184 seismograms requires about 1.9 Gb of memory in this case. With increasing speed and memory becoming available, much larger data sets should be soon become practical on a single workstation. Using about 4 Gb of memory on a single processor of an IBM Power 4 machine, we have been able to invert 37,320 path-average measurements in 22 hours 43 min. The input and recovered models for this case are shown in Figure 7.

[37] An important point to note about the algorithms for $A_i(r)$ and B_{ij} is that they are ideally suited to parallelization. In the calculation of both $A_i(r)$ and B_{ij} each ray path is processed independently and hence calculations may be carried out simultaneously on different processors.

[38] Current efforts are directed to implementing a parallelized version of the algorithms described in section 3. Preliminary results show that using parallelization, it is possible to improve efficiency still further. As an example,

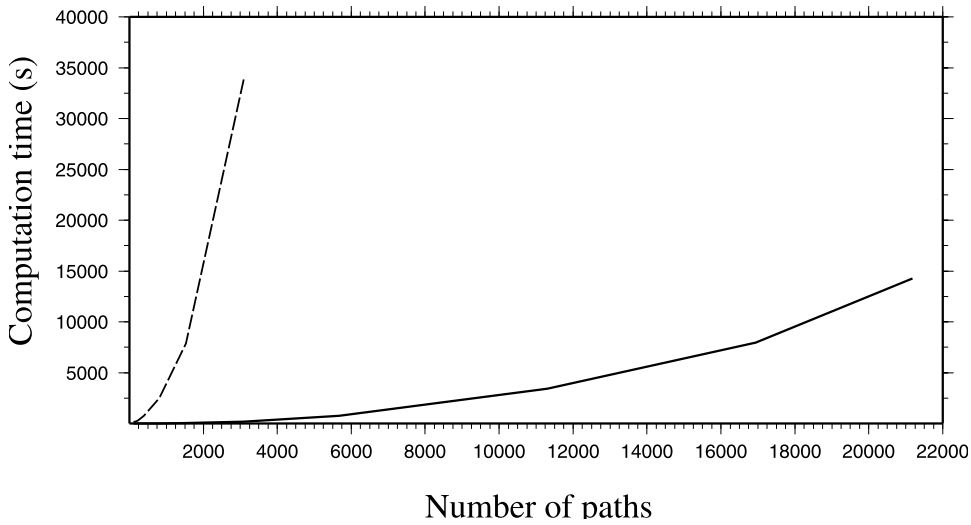

Figure 6. Performance of the optimized scheme (solid line) against the original Montagner approach (dashed line). All tests were carried out on a single 2.4-GHz Pentium 4 processor with 2 Gb of RAM. Here azimuthal anisotropy is present in the inverted model, and the Earth model was discretized into cells of $2° \times 2°$.

we have been able to invert 25,460 path average measurements in 3 hours 16 min using 16 processors of the IBM power 4 parallel machine at Institute for development and Resources in Intensive Scientific Computing. On the basis of these results we expect that data sets exceeding 50,000 paths will soon become practical on multiprocessor machines.

4. Using Voronoi Diagrams to Assess Model Constraint

[39] To circumvent the problems in assessing the resolution and a posteriori covariance when massive data sets are involved, we propose a new procedure to estimate the variation of constraint provided by the ray path coverage. The approach makes use of computational geometry techniques, which have seen several applications to geophysical problems in recent times. *Sambridge et al.* [1995] presented a technique for interpolating a property of the Earth specified at a series of reference points, while *Gudmundsson and Sambridge* [1998] and *Sambridge and Faletič* [2003] make use of Voronoi polyhedra and Delaunay tetrahedra as a parameterization for 3-D travel time tomography.

[40] In all cases, algorithms from the field of computational geometry are used to partition a medium into either Delaunay triangles (tetrahedra in three dimensions) or Voronoi cells (polyhedra in three dimensions) constructed around a set of irregularly spaced references nodes. In two dimensions a Voronoi diagram divides the plane into a set of polygons, one about each node, such that all points in a particular cell are closer to its defining node than any other node. Figure 8 shows an example on the Earth's surface.

[41] Surface wave regionalization is a 2-D problem in which the degree of constraint placed on seismic structure at a given location is highly dependent on the distribution of rays, which is always irregular. To produce a qualitative measure of constraint we aim to build an "optimized" parameterization of the model in which each geographical point belongs to the smallest cell for which an appropriate "quality criterion" is satisfied. If the quality criterion is chosen to measure the degree of constraint imposed locally on a particular type of inversion parameter, then the overall

Table 1. Performance of the Optimized Scheme Against the Original *Montagner* [1986] Approach[a]

Number of Paths	Original Scheme Computation Time, s	Optimized Scheme Computation Time, s	Improved Efficiency, %
60	50	1	98.0
111	116	2	98.3
223	288	5	98.3
395	762	8	98.9
805	2,499	21	99.2
1,532	7,862	53	99.3
3,090	33,889	186	99.4
5,672		763	
11,303		3,422	
16,936		7,956	
21,184		14,282	

[a]All calculations were performed on a 2.4-GHz single processor with 2 Gb of memory. These values correspond to the curve displayed in Figure 6.

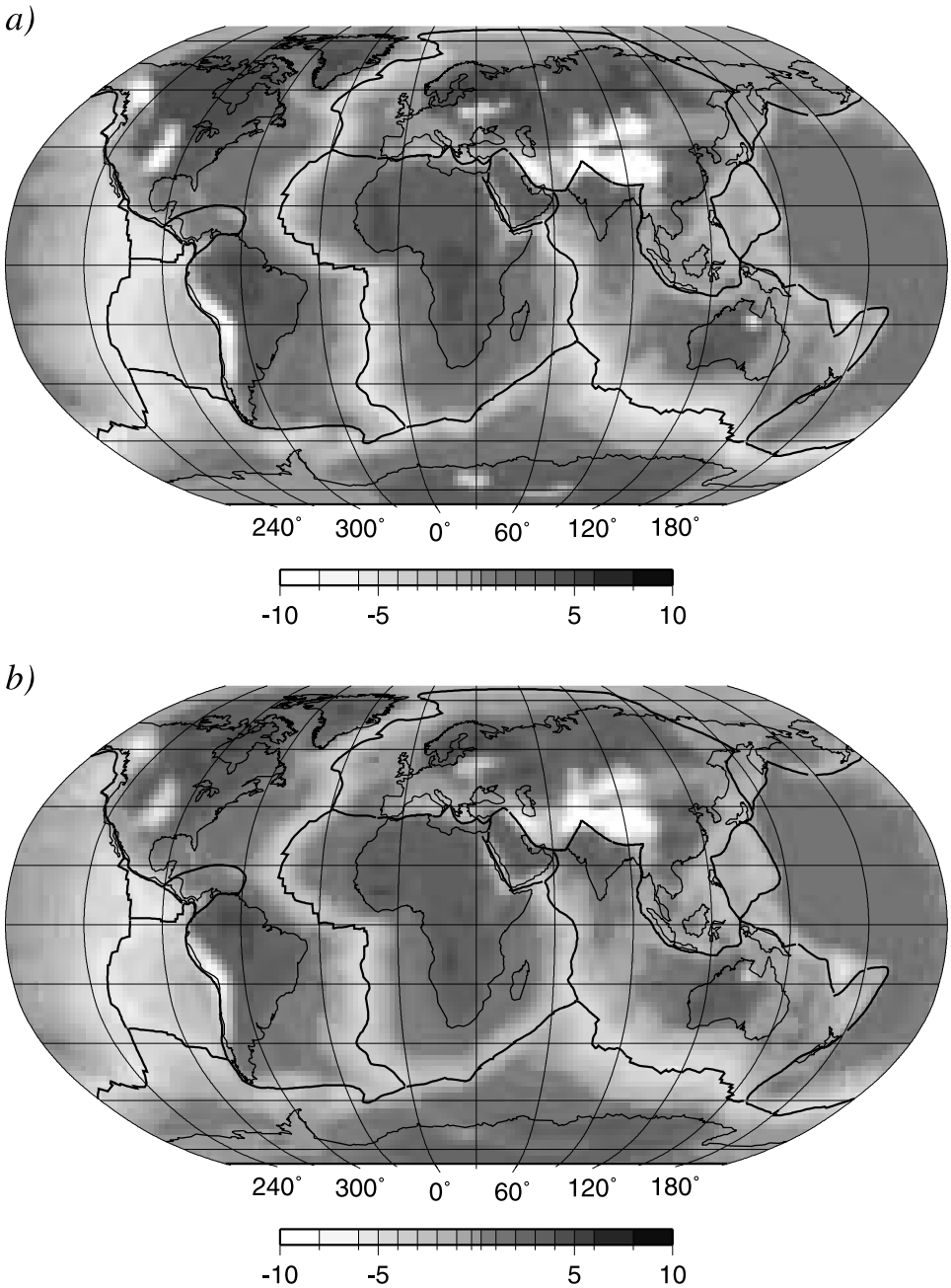

Figure 7. (a) Input *SV* velocity distribution used in the synthetic tests (from the 3SMAC model of *Nataf and Ricard* [1995]). (b) Final model after the inversion of 37,320 synthetic path average measurements. The corresponding ray density is shown in Figure 15. Recovery is excellent in most parts of the globe.

pattern of the optimized parameterization will reflect variations in model constraint.

4.1. Constructing Optimized Voronoi Meshes

[42] As a starting point, we build an "initial Voronoi diagram" from a uniform set of nodes, e.g., falling on a 2° × 2° grid. Using a 2° × 2° grid, our initial Voronoi diagram has therefore only 16,200 Voronoi cells across the globe and can be handled easily on a single workstation.

[43] From this starting Voronoi diagram we proceed by building a new Voronoi diagram in such a way that a particular quality criterion (which depends on the ray distribution) is satisfied for each cell. (For the moment we omit the definition of the quality criterion and simply assume that for any given cell on the Earth's surface it is possible to determine whether it is satisfied or not.) The process of generating the new "optimized" Voronoi diagram is simply a matter of deleting nodes which do not match the criterion and then recalculating the Voronoi cells about the remaining nodes. Note that when nodes are deleted, the neighboring Voronoi cells inherit the area and ray paths previously contained in the deleted Voronoi cell (see Figure 8). This makes the new cells larger on average and more likely to satisfy the quality criterion.

Figure 8. An example of a Voronoi diagram on the Earth's surface. The boundaries of each cell are great circles equidistant from the defining nodes, which are plotted as filled circles. Each Voronoi cell contains that part of the Earth's surface closest to the defining node. In this case the nodes are randomly positioned. The Voronoi diagram is used to build qualitative measures of model constraint is section 4.

[44] Imposing a $2° \times 2°$ lower limit on the initial Voronoi diagram is not a disadvantage when dealing with surface wave measurements in the period range of 40–300 s because the shortest distance scales resolvable will be of the order of a few hundred kilometers anyway.

[45] A complication in the approach described thus far is that every time a node is removed, one should, in principle, recalculate the positions of all remaining Voronoi cells and update their quality criteria. This is because the removal of any one node can affect the value of the quality criterion in its neighbors, and so all need to be recalculated before deciding on which node to delete next. The resulting optimized Voronoi diagram therefore depends on the order in which the nodes are removed. In practice, we iterate the procedure and at each stage randomly remove a fixed fraction of the nodes whose cells do not satisfy the quality criterion. Experience shows that so long as the fraction removed is small ($\sim$10%) the order is not important but the cost of the procedure is considerably reduced. Figure 9 shows a flowchart of the deletion process. By deleting a small proportion of nodes at each iteration in this way we can build an optimized Voronoi diagram very close to the ideal one that would contain the largest possible number of cells for which the corresponding quality criterion is satisfied.

4.2. Quality Criterion

[46] Here we propose a particular quality criterion based on the azimuthal distribution of surface wave paths within each cell. The azimuthal variation of surface waves is well known since the work of *Smith and Dahlen* [1973], who established the following relation for surface waves phase and group velocities in a slightly anisotropic medium:

$$C(T) = C_0(T) + A_1(T)\cos(2\theta) + A_2(T)\sin(2\theta) + A_3(T)\cos(4\theta) + A_4(T)\sin(4\theta), \quad (18)$$

$C_0(T)$ is the isotropic term representing the local value of the phase and group velocities and $A_1(T)$, $A_2(T)$, $A_3(T)$ and $A_4(T)$ are anisotropic coefficients.

[47] From a set of path-average phase or group velocity curves, it is possible to build some tomographic maps for the lateral variations and azimuthal anisotropy in group or phase velocity. A recent global study by *Trampert and Woodhouse* [2003] suggests that both 2θ and 4θ terms are required to explain Rayleigh phase velocities while Love waves do not require 2θ terms. *Montagner and Nataf* [1986] have shown that the local values of $C_0(T)$, $A_1(T)$, $A_2(T)$, $A_3(T)$, and $A_4(T)$ can be inverted for an isotropic term associated with the local shear velocities and for anisotropic terms corresponding to linear combinations of the elastic coefficients. The inversion at depth is straightforward because the partial derivatives needed are those of a transversely isotropic medium with a vertical axis of symmetry [*Montagner and Nataf*, 1986]. In the case of Rayleigh waves the largest partial derivatives for the anisotropic terms are associated with the $\cos(2\theta)$ and $\sin(2\theta)$ terms and the azimuthal variation for a long-period *SV* wave propagating horizontally with velocity β_v at a given depth z can be approximated with the following expression involving only the 2θ variation:

$$\beta_v(z) = \beta_{v0}(z) + (G_c/2\rho\beta_v)\cos(2\theta) + (G_s/2\rho\beta_v)\sin(2\theta), \quad (19)$$

where $\beta_{v0}(z)$ is the isotropic shear velocity at depth z and G_c and G_s are some combinations of the elastic parameters described by *Montagner and Nataf* [1986].

[48] In the case of Love waves the dominant anisotropic terms for the inversion are those associated with the $\cos(4\theta)$ and $\sin(4\theta)$ terms and the following relation yields for a long-period *SH* wave propagating horizontally with velocity β_h:

$$\beta_h(z) = \beta_{h0}(z) + (C_c/2\rho\beta_h)\cos(4\theta) + (C_s/2\rho\beta_h)\sin(4\theta). \quad (20)$$

The combinations of the elastic parameters C_c and C_s are also given by *Montagner and Nataf* [1986]. *Lévêque et al.* [1998] demonstrate that equations (19) and (20) are those that govern, in the long-period approximation, the velocities of long-period *SV* and *SH* waves propagating horizontally in the azimuth θ. Therefore, even in fully anisotropic structure, the velocity of horizontally propagating long-period *SV* and *SH* waves, in a depth range controlled by the rank and the period of the mode, is the most influential factor in the first-order variation of surface wave phase velocities. Finally, note that modern surface wave tomographic procedures often involve the building of path average 1-D shear velocity models before inverting them to retrieve the local structure [e.g., *Zielhuis and van der Hilst*, 1996; *Debayle and Lévêque*, 1997]. *Lévêque et al.* [1998] show that equations (19) and (20) can be used to retrieve the local distribution of shear velocity and anisotropy from a set of

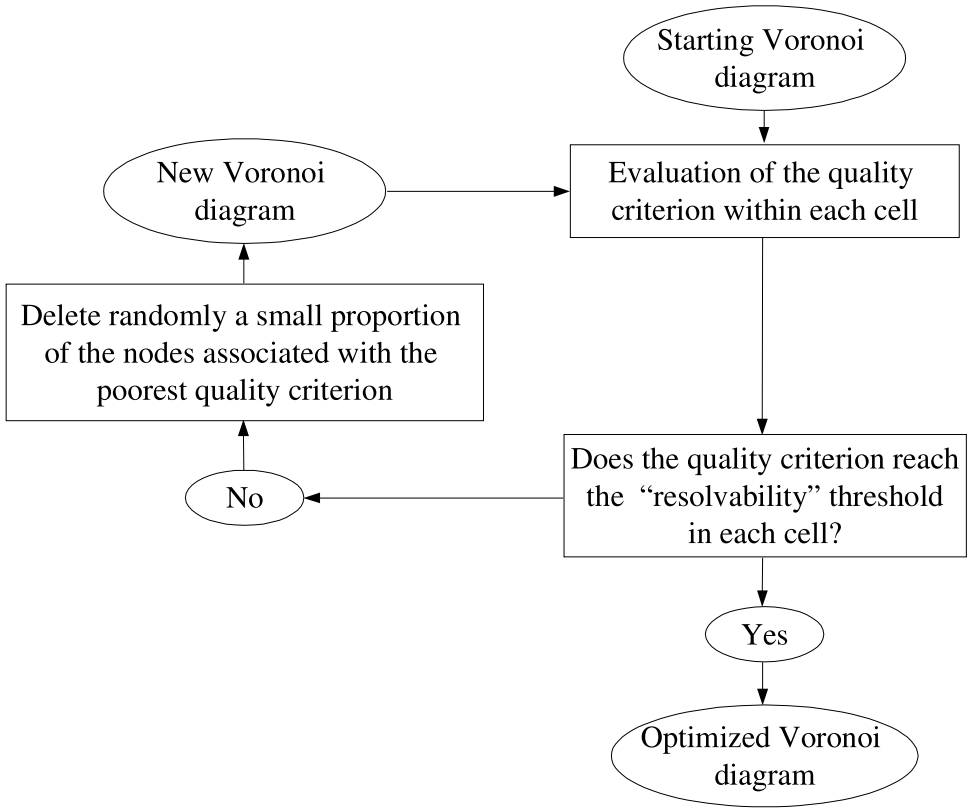

Figure 9. Flowchart of the iterative procedure used to generate the optimized Voronoi diagram with respect to any quality criterion. At each iteration the cells neighboring those deleted increase in size and capture more rays, which makes them more likely to satisfy the quality criterion.

path average *SV* and *SH* models obtained after applying a waveform inversion technique.

[49] The Rayleigh waves cos(2θ) sin(2θ) variation are periodic in π and require at least three paths with different azimuths to be properly retrieved without aliasing artifacts (the Nyquist criterion). Figure 10 illustrates a quality criterion which ensures that the azimuthal coverage is dense enough in each cell to resolve a 2θ azimuthal variation. Here the quality criterion is based on a subdivision of the 180° azimuth range into five bins of 36° each. A score of 1 is given if at least one path samples the corresponding azimuth range in a given cell, otherwise it is set to 0. With five azimuthal bins, there are 2^5 possible combinations of scores for any cell, and we divide these into the eight possible classes shown. Each class represents an approximately equivalent level of azimuthal coverage.

[50] For any given ray path coverage an optimized Voronoi diagram can then be generated using the nodal deletion technique in Figure 9 by requiring that the final cells all belong to class 1, i.e., at least one path in each 36° azimuth bin. This ensures that a minimum of three different azimuths are sampled in each cell, so that the azimuthal variation of Rayleigh waves can be resolved. Note that a similar quality criterion can be developed for Love waves by using nine azimuthal bins of 20° each. This would guarantee that a minimum of five directions are sampled, and hence that the cos(4θ) sin(4θ) azimuthal variation, which is the most important for resolving Love waves and long-period horizontally propagating *SH* waves.

4.3. Examples

[51] We present here two examples illustrating applications of the automated procedure for building optimized Voronoi diagrams. In each case the quality criteria represented by Figure 10 for Rayleigh waves are used. At each iteration of the process, N nodes corresponding to a percentage P of the initial number of nodes are deleted. The N nodes are taken in the "worst" class, according to the classification of Figure 10. If N is larger than the number of nodes N_i in class i, N_i nodes are removed in class i, while the remaining $N - N_i$ nodes are taken from class $i - 1$. This process is applied until nodes remain only in classes 1 and 2. At this stage the N_2 remaining nodes are removed by halving, over two iterations.

[52] In Figure 11 we plot three examples of optimized Voronoi diagrams obtained by applying our procedure with P = 10%, P = 5%, and P = 1% to an actual database of 24,124 Rayleigh wave seismograms (Figure 12) previously analyzed in Australia [*Debayle and Kennett*, 2003], northeastern Africa [*Debayle et al.*, 2001], Asia [*Priestley and Debayle*, 2003], and South America [*Heintz et al.*, 2000]. The three optimized Voronoi diagrams are extremely similar in the well-resolved areas where small Voronoi cells are obtained. The most significant differences occur in the poorly resolved regions where the shape of some of the cells can be affected by the random process used to delete

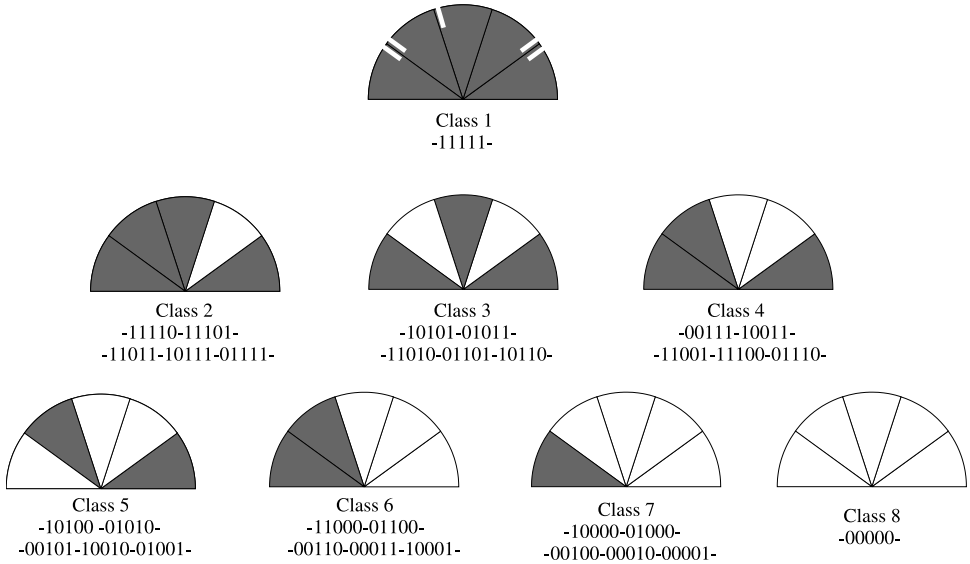

Figure 10. Definition of a quality criterion measuring azimuthal path coverage. The 180° azimuth range is divided into five "boxes" of 36°. The quality criterion is based on a sequence of five numbers, one for each box. Each number is either set to 1 if at least one path samples the azimuth range represented by the box, otherwise it is set to 0. With five boxes, 2^5 possible values exist for the quality criterion which can be grouped into the eight classes shown. These have equivalent azimuthal coverage. Each cell of the optimized Voronoi diagram is designed to be in class 1 (at least one path in each 36° box), thereby ensuring that azimuthal variation of Rayleigh waves is well resolved. White bars simulate the worst azimuthal ray sampling possible in class 1, where only three different azimuths are sampled.

the nodes. Large Voronoi cells fill up regions with no seismic rays such as the Pacific Ocean. Only the edges of these large cells are sampled, generally by small portions of rays with different azimuths. They tell us that if we aim to determine anisotropic directions on the whole Earth with our ray coverage, in some regions it is not possible to do a better job than using the crossing rays on the edge of the cells to constrain the azimuthal direction before extrapolating the measurement over a large surface.

[53] Note that the inversion procedure is designed to deal with this kind of underdetermined problem through the definition of the a priori model covariance. In regions not sampled by seismic rays the inverted model stays close to the a priori model. This can be seen in Figure 13, where no azimuthal anisotropy is observed in the regions poorly sampled by the data. The size and shape of the Voronoi cells are more interesting in the regions sampled by seismic rays. In the well-sampled regions, such as the eastern and north central part of Australia, most of Siberia, the northeastern part of Africa and Saudi Arabia and most of South America the initial 2° × 2° cells are preserved in the optimized Voronoi diagram. This means that restricting the inversion to those parameters that describe the 2θ azimuthal anisotropy of Rayleigh waves, our ray distribution allows resolution of azimuthal anisotropy in these regions on a 2° × 2° grid. In practice, the data often require variations of both isotropic and anisotropic parameters and this decreases the horizontal resolution.

[54] With the quality criterion of Figure 10 the optimized Voronoi diagram is a proxy for resolution of the 2θ azimuthal parameters but does not incorporate any a priori information on data and model and does not take into account the decrease in resolution which would result when other parameter types are introduced. At first glance, our optimized Voronoi diagram is far from a rigorous estimation of resolution which should incorporate both effects. However, in practice, we can argue that the information provided is quite reasonable. This is because the interpretation of the optimized Voronoi diagram can be guided by the horizontal degree of smoothing imposed a priori. Note that in order to minimize trade-off between isotropic and anisotropic parameters the horizontal degree of smoothing should be chosen in such a way that both parameters can be resolved on the smallest surface over which the model is required to be smooth (in our case this surface is circular and has a diameter close to L_{corr}). Therefore when the size of the Voronoi cells is comparable to the horizontal degree of smoothing imposed a priori, the azimuthal anisotropy is guaranteed to be resolved. Conversely, where the size of the Voronoi cells are larger, data constraint on local Rayleigh wave azimuthal variations is less, and so if the Voronoi cell size exceeds the horizontal degree of smoothing imposed a priori, then azimuthal anisotropy should not be interpreted.

[55] In Australia it is possible to compare the pattern of Voronoi cells observed in Figure 11 with the a posteriori error maps published at 100 km by *Debayle and Kennett* [2000] or with the resolution maps published by *Simons et al.* [2002]. At 100 km depth the structure is constrained by most of the paths used by *Debayle and Kennett* [2000], with a ray coverage very similar to the one available here. A good agreement exist between the pattern of Voronoi cells observed in Australia and the a posteriori error maps. However, if the Voronoi diagram provides us with a somewhat more qualitative information than the absolute value of the error, the information regarding the size and the shape of the region where a given parameter can be resolved

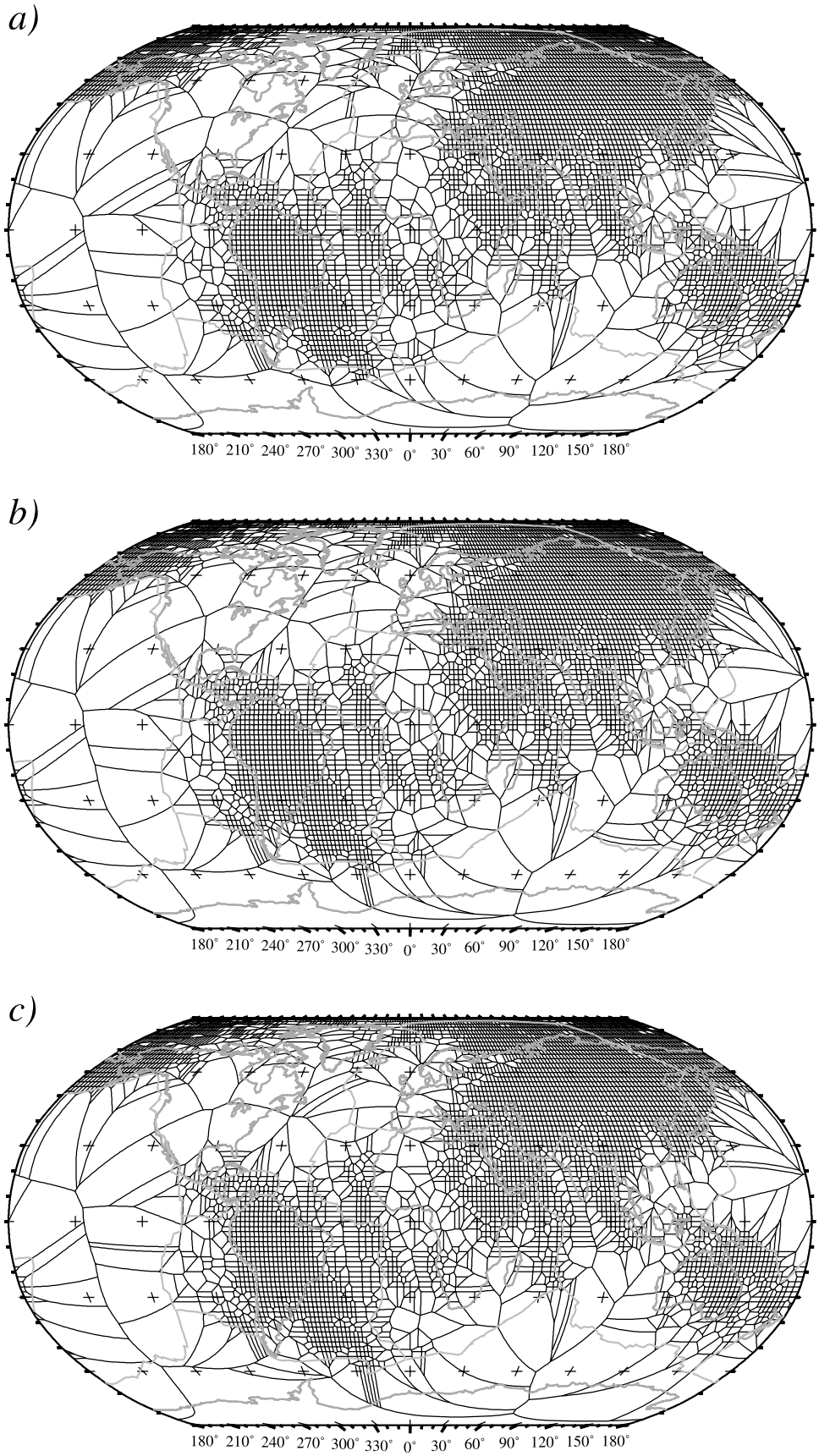

Figure 11. Optimized Voronoi diagram for (a) $P = 10\%$, (b) $P = 5\%$, and (c) $P = 1\%$. The corresponding ray path coverage is shown in Figure 12.

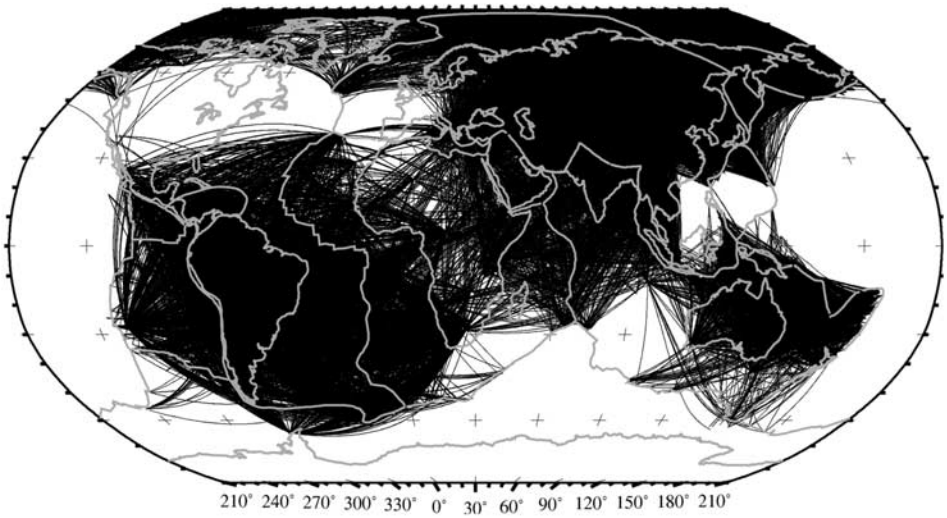

Figure 12. Ray coverage corresponding to the 24,124 path average measurements compiled from four regional tomographic studies by *Debayle et al.* [2001] (eastern Africa, northern Indian Ocean, and southern Asia), *Debayle and Kennett* [2003] (Australia and surrounding oceans), *Heintz et al.* [2000] (South America and surrounding oceans), and *Priestley and Debayle* [2003] (Asia).

is more complete. A set of Voronoi cells elongated in the east-west directions, as observed on the western coast of Indian or in the southern Atlantic Ocean eastward of the Brazilian coast, suggests that changes in anisotropic directions are more easy to resolve in the north-south directions than in the east-west direction.

[56] Figure 14 shows a second example where the optimized Voronoi diagram is built using P = 10% for a synthetic coverage of 37,320 paths corresponding to a realistic distribution of stations and events shown in Figure 15. The optimized Voronoi diagrams for P = 1% and P = 5% are virtually indistinguishable from Figure 14 and are not shown. In this case the process took 12 iterations (54 min 29 s) for P = 1%, 3 iterations (15 min 04 s) for P = 5%, and 2 iterations (14 min 14 s) for P = 10% (using the same hardware as previously). The optimized Voronoi diagram of Figure 14 mostly consists of a regular $2° \times 2°$ grid. Larger Voronoi cells are only needed in some parts of the Southern Hemisphere because of the lack of seismic stations. This confirms that on the basis of ray distribution alone, the azimuthal variations of Rayleigh waves could, in principle, be resolved almost everywhere on the Earth at this scale. However, as the inversion is likely to be unstable when both isotropic and anisotropic parameters are involved, a priori information on these parameters needs to be introduced and will therefore reduce the resolution. Global surface wave tomography has been until now restricted to the period range between 40 and 300 s, where surface waves have Fresnel zones with maximum widths exceeding a few hundred kilometers. This further limits the horizontal resolution attainable since under the assumption of ray theory, such surface wave data cannot in principle resolve structures smaller than the Fresnel zone width [e.g., *Spetzler and Snieder*, 2001]. The horizontal correlation length L_{corr} introduced in the inversion (see equation (5)) is chosen to mimic the Fresnel zone of long-period surface waves and forces the inverted model to be smooth, and hence it is not possible to resolve structure with wavelengths shorter than the Fresnel zone. As a result, a somewhat smoother pattern of anisotropy is obtained in Figure 13 compared to the size of the Voronoi cells in Figure 11. It will always be the case that the actual horizontal resolution achieved in surface wave tomography will be a combination of the geometrical constraint given the azimuthal distribution of rays and the limitations imposed by the underlying physics of surface waves.

[57] With the quality criterion used in this paper our aim has been to produce a simple, efficient, and practical guide on the resolution attainable for azimuthal parameters. In principle, it is possible to design a more sophisticated Voronoi diagram that would tell us more about the resolvability of a given parameter at a given depth or period. For example, an optimized Voronoi diagram could be built separately for the fundamental and the higher modes coverage. In the case of the regionalization of 1-D depth-dependent path average shear velocity model, the "a posteriori" error at each depth on the path average models could be used to discriminate between paths that constrain different depth intervals. It may also be possible to design other quality criteria that would be closer to the inverse operator and hence obtain a better estimate of resolution. These remain areas for further study.

5. Application to Real Data

[58] We have applied the algorithms described above to the tomographic inversion of 24,124 *Sv* velocity path average measurements obtained using the automated version [*Debayle*, 1999] of the *Cara and Lévêque* [1987] waveform inversion technique. The corresponding ray coverage is shown in Figure 12, and two maps of the lateral variations in shear wave velocity and azimuthal anisotropy at a depth of 100 and 150 km are presented in Figure 13. The tomographic inversion took approximately 5 hours per layer on a single 2.4-GHz Pentium 4 processor, using 2 Gb of memory.

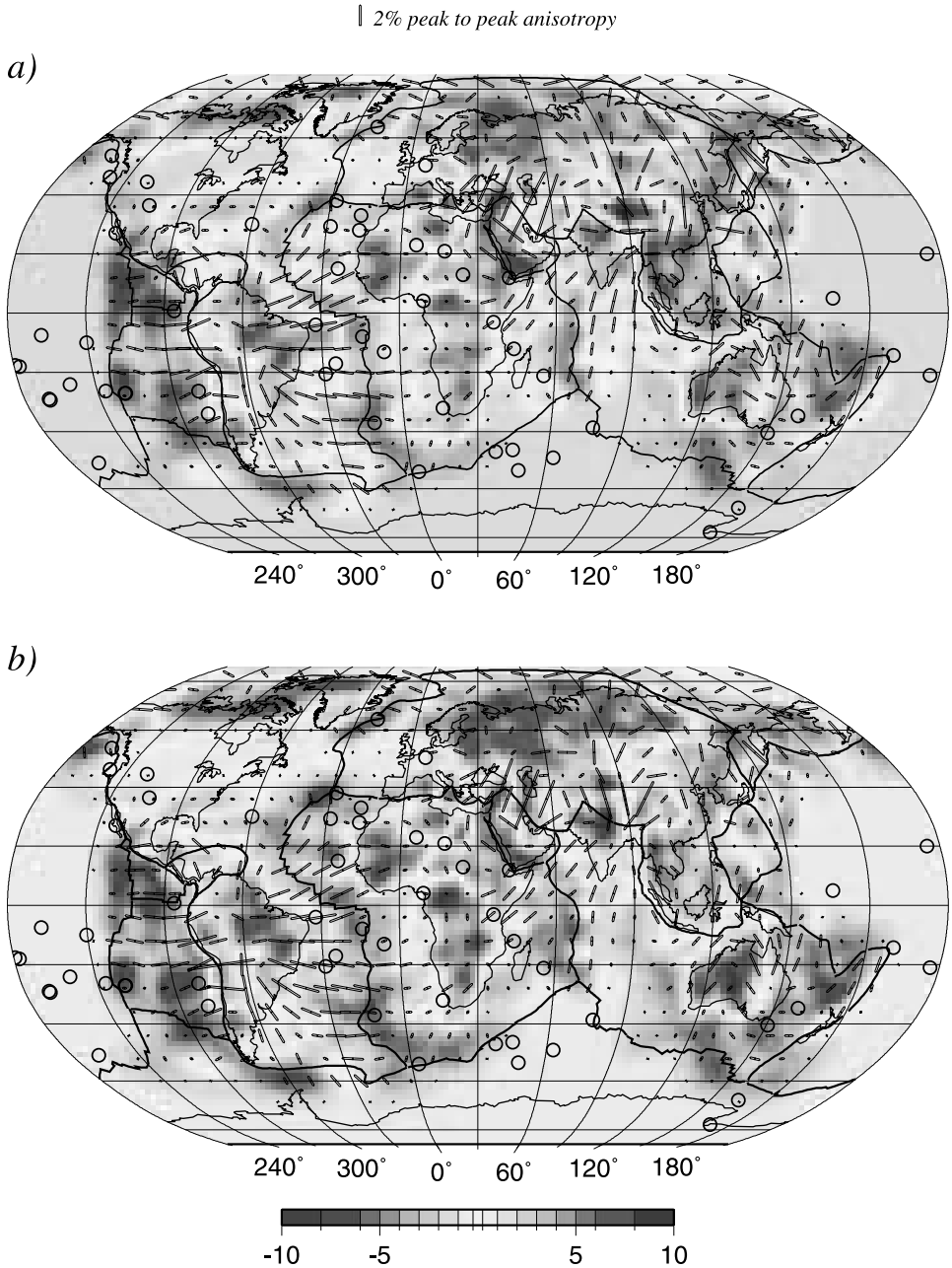

Figure 13. *SV* wave heterogeneities and azimuthal anisotropy (black bars) at (a) 100 km and (b) 150 km depth obtained from the inversion of 24,124 path average measurements of Figure 12. Hot spots locations are indicated with green circles. The length of the black bars is proportional to the maximum amplitude of azimuthal anisotropy. The reference velocity for the *Sv* wave heterogeneities is PREM. See color version of this figure at back of this issue.

[59] The optimized Voronoi diagrams of Figure 11 confirm that the ray coverage is dense enough to constrain the *Sv* wave azimuthal anisotropy beneath most of Asia, northeastern Africa, central and eastern Australia, and South America. In western and southern Africa, in the southeast Atlantic Ocean, and in western Australia the larger Voronoi cells indicate a degradation of resolution, but their size remains comparable to the horizontal correlation length of 400 km used for the inversion. This suggests that anisotropy can still be interpreted. In other regions (e.g., North America) the size of the Voronoi cells generally exceeds the horizontal correlation length indicating a lack of resolution. Here the inverted seismic parameters generally remain close to the a priori values (i.e., no anisotropy and a *Sv* wave perturbation close to 0%) in Figure 13 and should not be interpreted.

[60] This preliminary joint inversion of the data associated with four regional tomographic models underlines the ability of the continuous regionalization approach to process heterogeneous data sets. The *Tarantola and Valette* [1982] inversion scheme naturally finds a compromise between the smoothness required by the Gaussian a priori covariance

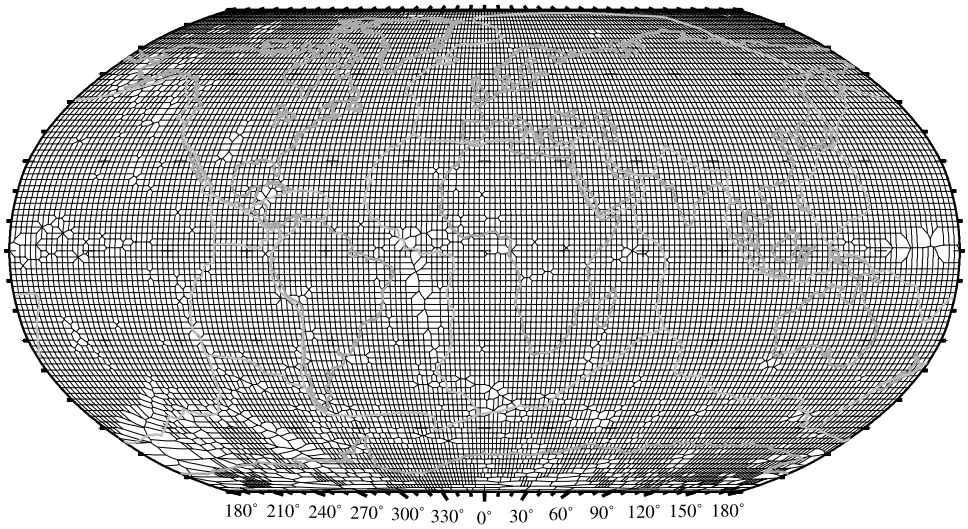

Figure 14. Optimized Voronoi diagram for P = 10%. The corresponding ray density is shown in Figure 15.

function and the level of heterogeneities required by the data. In regions where no coverage is available, such as the Pacific, southern Indian Ocean, or Antarctica, the inverted model displays no lateral heterogeneities and stays at the uniform a priori model. In poorly covered regions such as the northern Canadian shield, long-wavelength lateral variations imposed by the a priori covariance function characterize the model, while much more structural detail is present beneath the well-covered continents.

[61] The geodynamical results associated with each regional tomographic model have been discussed in detail by *Heintz et al.* [2000], *Debayle et al.* [2001], *Priestley and Debayle* [2003], and *Debayle and Kennett* [2003]. We note here that the joint inversion allows us to improve the coverage in the regions where the different regional studies overlap, mostly in southern Asia and central Africa (Figure 12). It also allows a direct comparison of results of different regional tomographic studies. For example, the joint inversion reveals that significant heterogeneities exist in the amplitude of seismic anisotropy in regions where good resolution is achieved. This is especially clear beneath the slowly moving African plate, which is associated with a weak azimuthal anisotropy, compared to Australia, South America, and Asia. We also note that the correlation between seismic heterogeneities and surface tectonics is extremely good. At 100 km depth, high seismic velocities are associated with the main cratons and the bottom of old oceanic basins where surface waves are still sensitive to the thickest oceanic lithosphere. At 150 km depth, high seismic velocities are observed beneath cratons and some tecton-

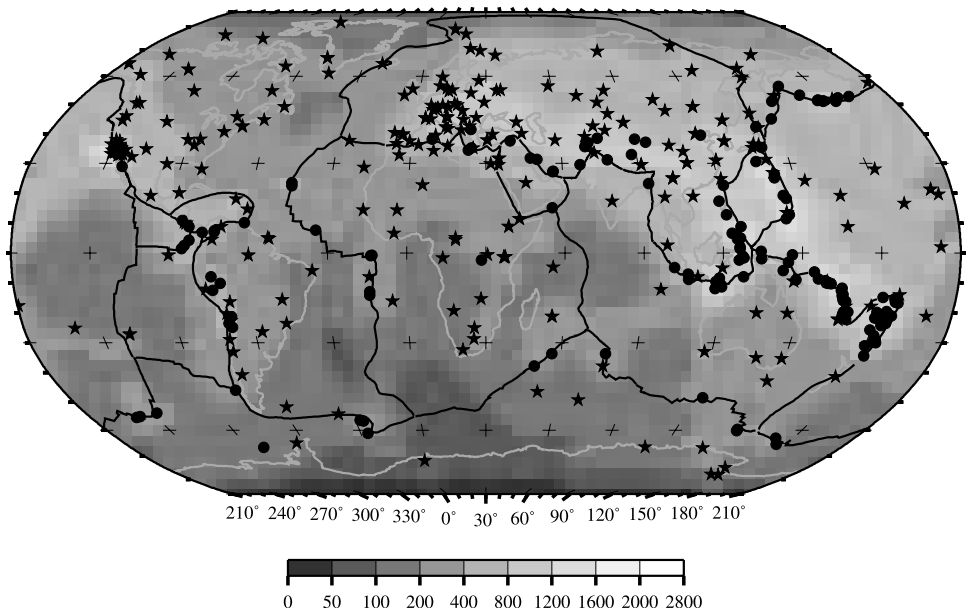

Figure 15. Ray density averaged over 4 × 4 degree cells, distribution of events (circles) and stations (stars).

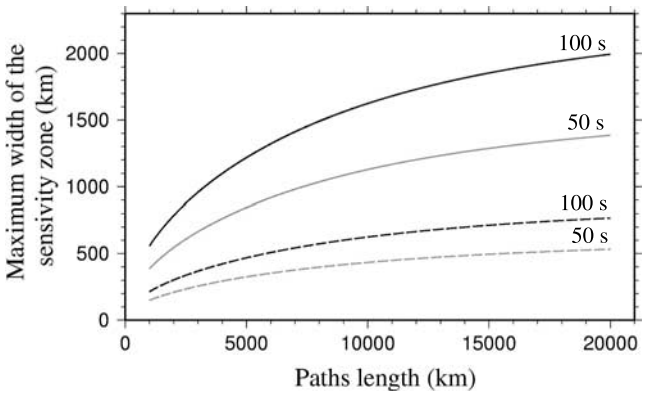

Figure 16. Maximum width of the sensitivity zone (as a function of path length) required to adequately account for scattering effects according to *Yoshizawa and Kennett* [2002] (dashed line) and *Spetzler et al.* [2002] (continuous line). The light gray curves show the case for 50 s period, and the dark curves show the case for 100 s. In each case this "sensitivity zone" corresponds to only a subregion of the first Fresnel zone.

ically active regions where the lithosphere is thickened by subduction (e.g., Tibet). Slower seismic velocities are located in the oceanic asthenosphere, and beneath the thin Phanerozoic lithosphere of eastern Australia and Asia.

[62] Finally, we note a reasonable correlation between the pattern of slow velocities and the location of the main hot spots according to *Müller et al.* [1993]. This is, for example, relevant at a depth of 100 km in the southern Atlantic Ocean and in Africa. Our ability to retrieve heterogeneities with wavelengths shorter than 1000 km is enhanced by the large number of short paths used in this preliminary inversion.

[63] Under the assumption of ray theory, surface waves cannot in principle resolve structures smaller than the Fresnel zone width [e.g., *Spetzler and Snieder*, 2001]. It is, however, a matter of debate as to which subregion of the first Fresnel zone should be considered to properly model the scattering sensitivity of surface waves. *Yoshizawa and Kennett* [2002] show that a primary sensitivity zone spanning only about one third of the width of the first Fresnel zone represents the region over which surface waves are coherent in phase. They argue that the crossover region between the primary sensitivity zone of two physical rays provides an upper bound for the lateral resolution that can be achieved in surface wave tomography. This crossover region is slightly larger than the width of the sensitivity zones which are less than 500 km for paths shorter than 6000 km and periods smaller than 100 s (Figure 16). On the other hand, *Spetzler and Snieder* [2001] found that a broader sensitivity zone which almost entirely encompasses the width of the first Fresnel zone has to be considered in order to adequately model scattering under the Born/Rytov approximation. For path lengths typical of global surface wave tomography (8000–10,000 km), Figure 16 shows that the upper bound for horizontal resolution is on the order of 1000–1500 km between 50 and 100 s periods. Note that the maximum width increases with path length, irrespective of the choice of subregion for the first Fresnel zone. It is clear that short paths are associated with thinner Fresnel zones and allow improved lateral resolution using ray theory. Short paths also have the advantage of being less sensitive to spurious effects such as multipathing or focusing/defocusing compared to the longer R1 and R2 paths commonly used in global tomography. Therefore, in regions where short paths allow sufficient coverage it is expected that structures with smaller horizontal wavelengths can be resolved.

[64] A histogram of path lengths for the 24,124 path average measurements presented in Figure 17 has 73% of the paths shorter than 6000 km. Our preliminary results suggest that structure with wavelengths smaller than 1000 km can be resolved in the uppermost 300 km of the mantle. This is consistent with the prediction of *Yoshizawa and Kennett* [2002] but suggests that the estimation by *Spetzler et al.* [2002] is too pessimistic. Note also that these short-wavelength heterogeneities have not previously been resolved in global tomography where the lateral resolution is usually larger than 1000 km. The recent S20RTS global *Sv* wave model of *Ritsema et al.* [1999] has a horizontal resolution of the order of 2000 km. In the upper mantle and transition zone, S20RTS is primarily constraint by fundamental and higher-mode surface waves, while long-period body waves have also been incorporated to provide additional resolution down to the lower mantle. In S20RTS the fundamental and higher-mode surface wave data set comes mostly from the work of *van Heijst and Woodhouse* [1999], who did not process paths shorter than 3500 km and favored large earthquakes with well-excited overtones at periods greater than 100 s. Long-period overtones contribute significantly to resolve heterogeneities in the transition zone because of their deep penetration, but they limit lateral resolution. In this study we favor shorter paths (33% of our paths are shorter than 3500 km), and to achieve a good horizontal coverage, we need to incorporate a larger proportion of intermediate size earthquakes. The use of shorter paths allows us to improve the lateral resolution in the uppermost mantle; however, the use of intermediate size earthquakes limits the number of events that can be analyzed at periods greater than 100 s, especially for the overtones and therefore reduces the sensitivity to structure deeper than 500 km compared to *Ritsema et al.* [1999].

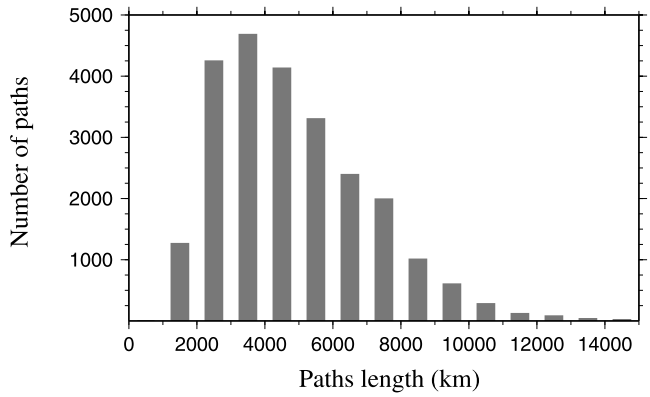

Figure 17. Distribution of lengths for the 24,124 paths of Figure 12.

[65] We are currently working on building a new global tomographic model, primarily constrained by a massive data set of epicenter-station paths shorter than 10,000 km. The new model will benefit from a large proportion of short epicenter-station paths related to intermediate size earthquakes, typical of regional tomography at the scale of a tectonic plate, but will also include larger events to improve the sampling of the deep structure using long-period fundamental modes and overtones. In this way, we aim to image the upper mantle structure at a global scale with a lateral resolution of a few hundred kilometers, typical of what can be currently achieved in regional surface wave tomography.

6. Conclusions

[66] We have presented an approach that allows surface wave tomography to be applied to the massive volumes of data currently available at the global scale. The new scheme is an extension of the continuous regionalization method of *Montagner* [1986] in that it preserves its use of a priori information on the model but considerably improves efficiency. All of the advantages of the *Montagner* [1986] continuous regionalization approach are retained, including its robustness and use of a continuous parameterization instead of expansion in terms of arbitrarily truncated basis functions (a well-known source of artifacts in tomographic models). The use of a Gaussian a priori covariance function on the model is particularly useful with large or mixed regional and global data sets because it naturally handles the inevitable variations in the ray path density.

[67] Experiments show that the scheme allows the regionalization of up to 25,000 path-average measurements in a few hours on a standard workstation. It is also ideally suited to parallelization, which in the near future should make routine the inversion of data sets exceeding 50,000 paths.

[68] To circumvent the difficulty of assessing resolution and a posteriori covariance when such large data sets are involved, we have proposed a new technique, based on the construction of optimized Voronoi diagrams. In this approach the size of the Voronoi cells gives an indication of the scale length of the structure that can be resolved, while their shape provides information on resolvability of structural variation along different azimuths.

[69] The first application of the complete approach to a real data set of 24124 path average *Sv* wave measurements suggests that in most parts of the Earth, upper mantle heterogeneities with wavelength smaller than 1000 km can be resolved. With the new approach it will be possible to combine into a single global inversion, the multiple data sets of fundamental and higher-mode surface waves that are more typically associated with regional tomography.

[70] **Acknowledgments.** Part of this work was carried out while E. Debayle was a visiting fellow in the Seismology and Geomagnetism group at the Research School of Earth Sciences, Australian National University. This work is supported by the INSU program "Intérieur de la Terre", and through a Discovery project (DP0208039) funded by the Australian Research Council. Supercomputer facilities were provided by the CNRS through the Institut du Développement et des Resources en Informatique Scientifique and the French ministry of research through the Centre Informatique National de l Enseignement Supérieur (CINES). We have benefitted from constructive reviews by J. Trampert, J. Ritsema, and an anonymous reviewer. Special acknowledgment is addressed to Nicole Audiffren and Alain Mango from CINES for their help in developing a parallel version of the regionalization approach. Finally, E. Debayle thanks J. J. Lévêque, Luis Rivera, and Michel Cara for numerous discussions about this work. Software implementing the algorithms presented here are available on request from the first author.

References

Aki, K., and E. Husebye (1977), Determination of the three-dimensional seismic structure of the lithosphere, *J. Geophys. Res.*, *82*, 277–296.

Cara, M., and J. Lévêque (1987), Waveform inversion using secondary observables, *Geophys. Res. Lett.*, *14*, 1046–1049.

Debayle, E. (1999), *SV*-wave azimuthal anisotropy in the Australian upper-mantle: Preliminary results from automated Rayleigh waveform inversion, *Geophys. J. Int.*, *137*, 747–754.

Debayle, E., and B. Kennett (2000), The Australian continental upper mantle: Structure and deformation inferred from surface waves, *J. Geophys. Res.*, *105*, 25,423–25,450.

Debayle, E., and B. Kennett (2003), Surface waves studies of the Australian region, in *Evolution and Dynamics of the Australian Plate*, edited by R. R. Hillis and R. D. Müller, *Spec. Pap. Geol. Soc. Am.*, 372, 19–34.

Debayle, E., and J. Lévêque (1997), Upper mantle heterogeneities in the Indian Ocean from waveform inversion, *Geophys. Res. Lett.*, *24*, 245–248.

Debayle, E., J. Lévêque, and M. Cara (2001), Seismic evidence for a deeply rooted low velocity anomaly in the upper mantle beneath the northeastern Afro/Arabian continent, *Earth Planet. Sci. Lett.*, *193*, 369–382.

Dziewonski, A. (1971), Upper mantle models from "pure-path" dispersion data, *J. Geophys. Res.*, *76*, 2601–2687.

Dziewonski, A., S. Bloch, and M. Landisman (1969), A technique for the analysis of transcient seismic signals, *Bull. Seismol. Soc. Am.*, *59*, 427–444.

Dziewonski, A., T. Chou, and J. Woodhouse (1981), Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.*, *86*, 2825–2852.

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

Heintz, M., E. Debayle, A. Vauchez, and M. Assumpcao (2000), Seismic anisotropy and surface wave tomography of South America, *Eos Trans. AGU*, *81*(48), Fall Meet. Suppl., Abstract S52B-22.

Kennett, B. (1995), Approximations for surface-wave propagation in laterally varying media, *Geophys. J. Int.*, *122*, 470–478.

Lévêque, J. J. (1980), Regional upper-mantle S-velocity models from phase velocities of great Rayleigh waves, *Geophys. J. R. Astron. Soc.*, *63*, 23–43.

Lévêque, J., E. Debayle, and V. Maupin (1998), Anisotropy in the Indian Ocean upper mantle from Rayleigh- and Love-waveform inversion, *Geophys. J. Int.*, *133*, 529–540.

Montagner, J. (1986), Regional three-dimensional structures using long-period surface waves, *Ann. Geophys.*, *4*, 283–294.

Montagner, J., and H. Nataf (1986), A simple method for inverting the azimuthal anisotropy of surface waves, *J. Geophys. Res.*, *91*, 511–520.

Montagner, J., and T. Tanimoto (1990), Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *J. Geophys. Res.*, *95*, 4797–4819.

Müller, R., J. Royer, and L. Lawver (1993), Revised plate motions relative to the hotspots from combined Atlantic and Indian Ocean hotspot tracks, *Geology*, *21*, 275–278.

Nataf, H., and Y. Ricard (1995), 3SMAC: An a priori tomographic model of the upper mantle based on geophysical modeling, *Phys. Earth Planet. Inter.*, *95*, 101–122.

Nolet, G., R. Montelli, and J. Virieux (1999), Explicit, approximate expressions for the resolution and a posteriori covariance of massive tomographic systems, *Geophys. J. Int.*, *138*, 36–44.

Okabe, A., B. Boots, and K. Sugihara (1992), *Spatial Tessellations Concepts and Applications of Voronoi Diagrams*, John Wiley, Hoboken, N. J.

Okal, E. (1977), The effect of intrinsic ocean upper mantle heterogeneity on the regionalization of long period Rayleigh wave phase velocities, *Geophys. J. R. Astron. Soc.*, *49*, 357–370.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery (1992), *Numerical Recipies in FORTRAN 77: The Art of Scientific Programming*, 2nd ed., Cambridge Univ. Press, New York.

Priestley, K., and E. Debayle (2003), Seismic evidence for a moderately thick lithosphere beneath the Siberian platform, *Geophys. Res. Lett.*, *30*, 1118, doi:10.1029/2002GL015931.

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

Ritzwoller, M., N. Shapiro, A. Levshin, and G. Leahy (2001), Crustal and upper mantle structure beneath Antarctica and surrounding oceans, *J. Geophys. Res.*, *106*, 30,645–30,670.

Sambridge, M., and R. Faletič (2003), Adaptive whole earth tomography, *Geochem. Geophys. Geosyst.*, *4*, 1022, doi:10.1029/2001GC000213.
Sambridge, M., J. Braun, and H. McQueen (1995), Geophysical parametrization and interpolation of irregular data using natural neighbours, *Geophys. J. Int.*, *122*, 837–857.
Simons, F., R. van der Hilst, J. Montagner, and A. Zielhuis (2002), Multimode rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle, *Geophys. J. Int.*, *151*, 738–754.
Smith, M., and F. Dahlen (1973), The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium, *J. Geophys. Res.*, *78*, 3321–3333.
Spetzler, J., and R. Snieder (2001), The effect of small-scale heterogeneity on the arrival time of waves, *Geophys. J. Int.*, *145*, 786–796.
Spetzler, J., J. Trampert, and R. Snieder (2002), The effect of scattering in surface wave tomography, *Geophys. J. Int.*, *149*, 755–767.
Su, W., R. Woodward, and A. Dziewonski (1994), Degree 12 model of shear velocity heterogeneity in the mantle, *J. Geophys. Res.*, *99*, 6945–6980.
Tarantola, A., and A. Nercessian (1984), Three-dimensional inversion without blocks, *Geophys. J. R. Astron. Soc.*, *76*, 299–306.
Tarantola, A., and B. Valette (1982), Generalised nonlinear inverse problems solved using the least square criterion, *Rev. Geophys.*, *20*, 219–232.
Trampert, J., and J. Woodhouse (1995), Global phase velocity maps of Love and Rayleigh waves between 40 and 150 seconds, *Geophys. J. Int.*, *122*, 675–690.
Trampert, J., and J. Woodhouse (2003), Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s, *Geophys. J. Int.*, *154*, 154–165.
van Heijst, H., and J. Woodhouse (1999), Global high-resolution phase velocity distributions of overtone and fundamental-mode surface waves determined by mode branch stripping, *Geophys. J. Int.*, *137*, 601–620.
Whaler, K., and D. Gubbins (1981), Spherical harmonic analysis of the geomagnetic field: An example of a linear inverse problem, *Geophys. J. R. Astron. Soc.*, *65*, 645–693.
Woodhouse, J. (1974), Surface waves in a laterally varying layered structure, *Geophys. J. R. Astron. Soc.*, *37*, 461–490.
Yoshizawa, K., and B. Kennett (2002), Determination of the influence zone for surface wave paths, *Geophys. J. Int.*, *149*, 441–454.
Zielhuis, A., and R. van der Hilst (1996), Upper-mantle shear velocity beneath eastern Australia from inversion of waveforms from Skippy portable arrays, *Geophys. J. Int.*, *127*, 1–16.

E. Debayle, IPGS-Ecole et Observatoire des Sciences de la Terre, CNRS and Université Louis Pasteur, 5 rue René Descartes, F-67084 Strasbourg, France. (eric.debayle@eost.u-strasbg.fr)

M. Sambridge, Centre for Advance Data Inference, Research School of Earth Sciences, Australian National University, Canberra ACT 0200, Australia. (malcolm.sambridge@anu.edu.au)

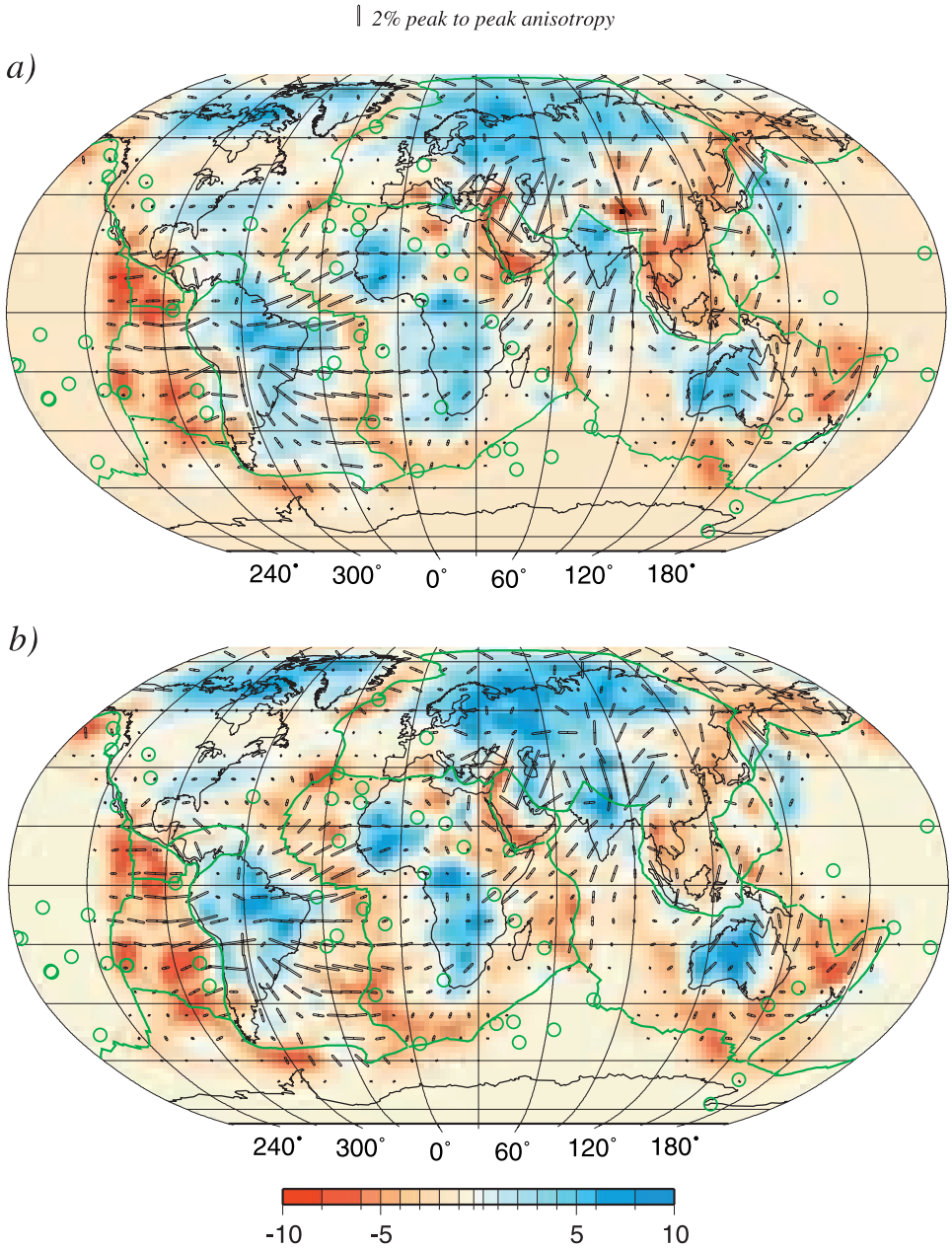

Figure 13. *SV* wave heterogeneities and azimuthal anisotropy (black bars) at (a) 100 km and (b) 150 km depth obtained from the inversion of 24,124 path average measurements of Figure 12. Hot spots locations are indicated with green circles. The length of the black bars is proportional to the maximum amplitude of azimuthal anisotropy. The reference velocity for the *Sv* wave heterogeneities is PREM.