Adaptively parametrized surface wave tomography: methodology and a new model of the European upper mantle

J. F. Schaefer, L. Boschi and E. Kissling

Institute of Geophysics, ETH Zurich, Sonneggstr. 5, CH-8092 *Zurich, Switzerland. E-mail: schaefer@erdw.ethz.ch*

SUMMARY

In this study, we aim to close the gap between regional and global traveltime tomography in the context of surface wave tomography of the upper mantle implementing the principle of adaptive parametrization. Observations of seismic surface waves are a very powerful tool to constrain the 3-D structure of the Earth's upper mantle, including its anisotropy, because they sample this volume efficiently due to their sensitivity over a wide depth range along the ray path. On a global scale, surface wave tomography models are often parametrized uniformly, without accounting for inhomogeneities in data coverage and, as a result, in resolution, that are caused by effective under- or overparametrization in many areas. If the local resolving power of seismic data is not taken into account when parametrizing the model, features will be smeared and distorted in tomographic maps, with subsequent misinterpretation. Parametrization density has to change locally, for models to be robustly constrained without losing any accurate information available in the best sampled regions. We have implemented a new algorithm for upper mantle surface wave tomography, based on adaptive-voxel parametrization, with voxel size defined by both the 'hit count' (number of observations sampling the voxel) and 'azimuthal coverage' (how well different azimuths with respect to the voxel are covered by the source-station distribution). High image resolution is achieved in regions with dense data coverage, while lower image resolution is kept in regions where data coverage is poorer. This way, parametrization is everywhere tuned to optimal resolution, minimizing both the computational costs, and the non-uniqueness of the solution. The spacing of our global grid is locally as small as $\sim$50 km. We apply our method to identify a new global model of vertically and horizontally polarized shear velocity, with resolution particularly enhanced in the European lithosphere and upper mantle. We find our new model to resolve lithospheric thickness and radial anisotropy better than earlier results based on the same data. Robust features of our model include, for example, the Trans-European Suture Zone, the Panonnian Basin, thinned lithosphere in the Aegean and Western Mediterranean, possible small-scale mantle upwellings under Iberia and Massif Central, subduction under the Aegean arc and a very deep cratonic root underneath southern Finland.

Key words: Inverse theory; Tomography; Surface waves and free oscillations; Seismic anisotropy; Seismic tomography; Dynamics of lithosphere and mantle.

1 INTRODUCTION

In all aspects of upper mantle geophysics, seismic models play a fundamental role. They are essential for accurate localization of earthquakes, estimating chemical heterogeneities (e.g. Forte *et al.* 1994), mapping important discontinuities (lithosphere–asthenosphere boundary, upper-mantle phase transitions) and as a constraint for geodynamic modelling: Faccenna & Becker (2010) for instance use seismic models to estimate upper-mantle density and temperature. This allows them to estimate the pattern of mantle flow in the region, which in turn they use to predict plate motion and dynamic topography. In general it is still important, however, to significantly improve the quality of seismic models, if one hopes to define a model of the Earth's structure and dynamics that simultaneously explains data from different disciplines.

Surface wave measurements provide the most useful data set to obtain anisotropic models of the upper mantle (e.g. Nataf *et al.* 1984; Woodhouse & Dziewonski 1984; Montagner & Tanimoto 1991; Ekström & Dziewonski 1998; Khan *et al.* 2011). Depending on their type (Love or Rayleigh) and their frequency they show different depth sensitivity and thus resolve structure in different depth. Phase-velocity maps (e.g. Boschi *et al.* 2009) show that, for example, Love waves with a period of 35 s resolve crustal structure while with Rayleigh waves

with a period of 150 s deep cratonic structures can be identified. Assuming for a measurement all fundamental modes could be identified, a very good depth sampling along the great-circle path is achieved. If combined with a good source-station coverage, theoretically a data set can be obtained which provides a much more uniform and, hence, better sampling of the upper mantle than can be achieved by body waves. Different frequencies can be used to be inverted together for 3-D crustal and mantle structure. In practice however frequencies high enough to adequately resolve crustal structure cannot be measured and crustal corrections have to be applied. In future, studies of ambient noise tomography are expected to provide us with the measurements necessary to jointly invert for crust and upper-mantle structure.

A problem of significance and difficulty in tomography resides in the appropriate selection of model parametrization. In the ideal experimental setup sources and stations would be distributed in such a way that the target structure is sampled from all directions adequately and uniformly. In real applications of seismic tomography unfortunately this is impossible, as sources are mainly limited to plate boundaries, and stations are distributed inhomogeneously and predominantly located on continents. This leads to a quite heterogeneous data coverage (Fig. 1), with some areas, and some directions of wave propagation, sampled by much more observations than others. A uniform parametrization thus results in under- and/or overparametrization of at least some regions. If model resolution is not taken into account when parameterizing the model, real features might be deformed in the solution model, which leads to misinterpretation as shown, for example, by Kissling *et al.* (2001) for local body wave tomography. For the horizontally travelling surface waves, however, resolution in any layer is only depending on azimuthal distribution of rays or 'crossfiring' within each cell. In order for our parametrization to be everywhere adequate, we follow an adaptive-grid approach analogue to that of Bijwaard *et al.* (1998) who used the number of rays to estimate the resolution capability rather than using the formal resolution matrix. In practice, the size of each parametrization pixel depends on the local model resolution capability, which we estimate based on the density of seismic ray paths and their azimuthal coverage ('crossfiring'). This way, higher image resolution is possible in regions where observations permit. Lower image resolution is kept in regions with lower data coverage and lower crossfiring. Note that the theoretically smallest possible cell size, that is, highest resolution, does not only depend on number of rays and crossfiring but rather on shortest wave length in data. Optimal model resolution is then achieved everywhere in the region resolved by data, which maximizes the model's information content without over- and/or underparametrization. An important consequence is that computational costs are kept to a minimum, while employing a 50-km grid spacing globally would be impossible even on powerful computational platforms.

On a global scale, most surface wave models today are still parametrized uniformly (e.g. Kustowski *et al.* 2008) with few exceptions. Boschi *et al.* (2004, 2009) use a multigrid approach with a rough grid on a global scale and a fine grid over Europe, an area better covered by their data. Due to the sudden and strong change in pixel size at interface between the two grids, however, imaging artefacts like boundary
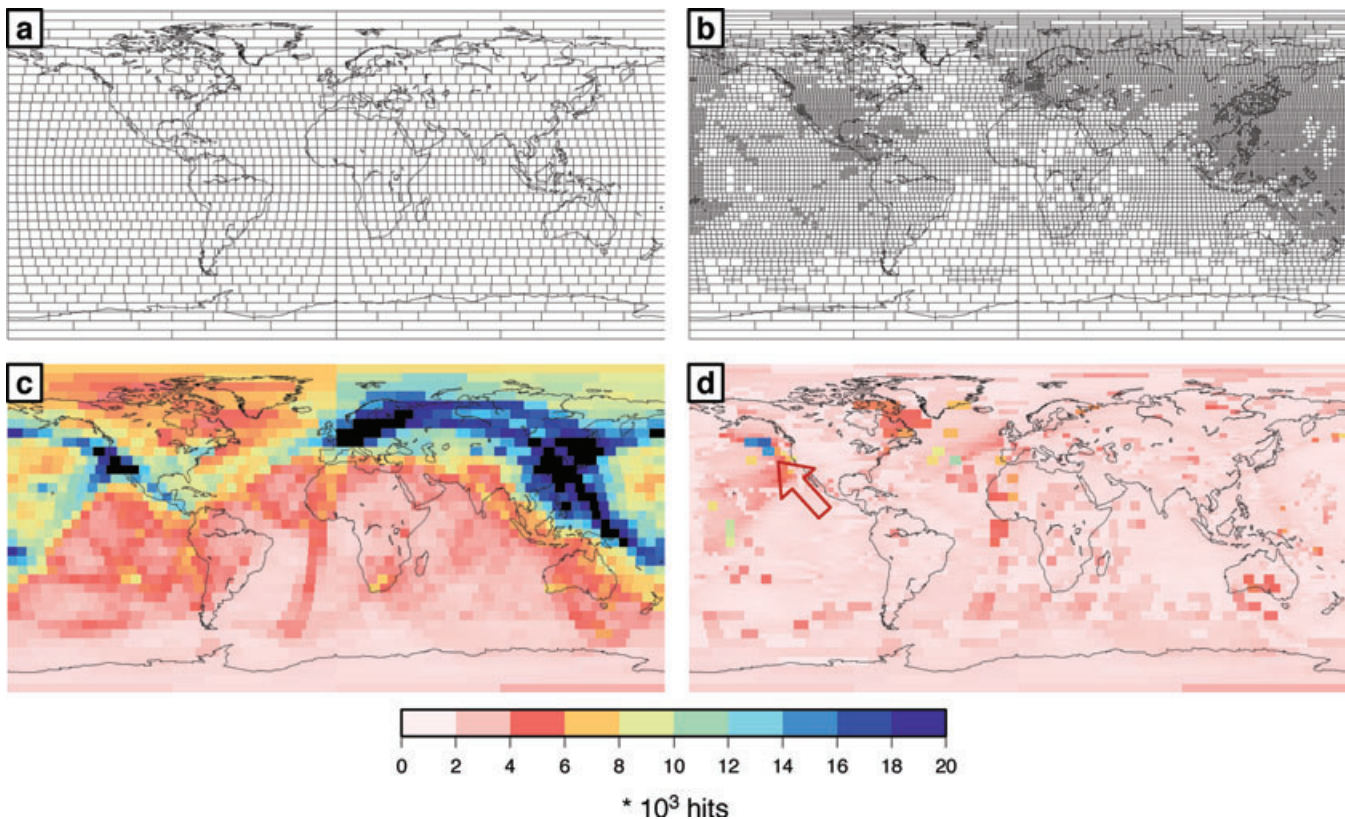
Figure 1. (a) Uniform, equal-area (5° × 5° at equator) pixel grid as in Boschi & Dziewonski (1999). (b) Non-uniform grid based on our algorithm applied to Boschi *et al.* (2009) 35 s Rayleigh wave database. (c) Number of hits per pixel, for the uniform grid of panel (a) and the 35 s Rayleigh wave database. (d) same as (c), grid as in (b). The arrow marks some pixel where the hit count is much higher. The resolution is not high enough to use a finer parametrization due to bad crossfiring.

effects might have somewhat polluted their model. Schivardi & Morelli (2009) use a similar approach, but first conduct a global inversion on a uniform grid to obtain a laterally heterogeneous starting model. In a second inversion they use a multigrid approach to invert a regional data set for Europe and the Mediterranean. Wang *et al.* (1998) use a multiresolution spherical-spline inversion to conduct global phase-anomaly maps. Experiments in finding data-adaptive grids have been conducted by, for example, Sambridge *et al.* (1995); Sambridge & Gudmundsson (1998); Curtis & Snieder (1997) by using Delauney tetrahedra and Voronoi polyhedra. Debayle & Sambridge (2004) employed voronoi cells, which also account for azimuthal coverage, to invert for vertically polarized shear velocity heterogeneities. The same technique has been used by Pilidou *et al.* (2004). Bodin & Sambridge (2009) use Voronoi cells with ambient-noise Rayleigh wave group velocity data. Data-adaptive inversion techniques for body wave tomography are used by, for example, Bijwaard *et al.* (1998) or Sambridge & Faletic (2003), who refine their model during inversion via tetrahedral cells. Another approach to parametrize the model is wavelet decomposition (e.g. Chiao & Kuo 2001; Loris *et al.* 2007; Tikhotsky & Achauer 2008). Some adaptive-grid approaches work on hit count (number of rays) while some use a more sophisticated method. The parametrization of, for example, Bodin & Sambridge (2009) is driven by data fit while Curtis & Snieder (1997) use a genetic algorithm to arrange the vertices of a triangular parametrization so as to minimize the condition number of the resulting tomographic system of equations. We define here our parametrization based on the hitcount and on crossfiring (resolution estimates). Taking the azimuthal coverage into account for defining the parametrization has the advantage that the size of parametrization pixels changes more smoothly with location than based on hitcount alone, suppressing boundary effects.

The aim of this paper is to present how adaptive pixel and voxel grids can be used in surface wave based phase-velocity and shear-velocity tomography and demonstrate the advantages to the resulting model. We first provide a summary of the theory of surface wave tomography (Boschi *et al.* 2009) and describe our new adaptive-grid implementation, including the definition and computation of an appropriate roughness damping operator. We show a set of adaptively parametrized phase-velocity maps corresponding to various Love and Rayleigh wave frequencies, and present a new, adaptively parametrized, 3-D, radially anisotropic model of the upper mantle, focused on Europe and the Mediterranean Basin, based on the global database of (Ekström *et al.* 1997; Boschi & Ekström 2002; Fry *et al.* 2008). To to quantify the improvements of the resulting model and the better depth resolution of the lithosphere we carefully compare our results with those of (Boschi *et al.* 2009), who inverted the same data on a simpler (multiple but not adaptive) parameterization grid.

2 THEORETICAL BACKGROUND

We summarize here the theory of surface wave tomography as described by Boschi & Dziewonski (1999), Boschi & Ekström (2002) and Boschi *et al.* (2009), which constitutes the starting point of our work.

2.1 Phase-velocity tomography

The linear relation between phase delay time $\delta t(\omega)$ for a certain frequency ω of a Love or Rayleigh wave and the according mode slowness $\delta p(\theta, \phi; \omega)$ is given by Fermat's principle as described by, for example, Boschi & Ekström (2002):

$$\delta t(\omega) = \omega \int_{\text{path}_m} \delta p(\theta(s), \phi(s); \omega) \, ds, \tag{1}$$

where θ and ϕ are colatitude and longitude, respectively, and s the ray path. Eq. (1) is valid for a smooth structure (Boschi & Ekström 2002). The inverse problem is linearized around the starting model PREM (Dziewonski & Anderson 1981). Since the starting model has no lateral heterogeneities the ray path coincides with the great-circle path. We can express the heterogeneity of phase slowness δp for each frequency through a set of horizontal basis functions $S_i(\theta, \phi)$ ($i = 1, \ldots, I$) and their unknown coefficients x_i

$$\delta p(\theta, \phi; \omega) = \sum_{i=1}^{I} x_i S_i(\theta, \phi). \tag{2}$$

After substituting (2) into (1) for each datum (denoted by an index *m*) with its associated ray path, a linear system can be found

$$\delta \boldsymbol{t}(\omega) = \mathbf{C} \cdot \mathbf{x}, \tag{3}$$

where $\mathbf{C}$ is given by

$$C_{mi} = \omega \int_{\text{path}_m} S_i(\theta(s), \phi(s)) \, ds. \tag{4}$$

The coefficients $\mathbf{x}$ are found via least-squares inversion of (3) and used to calculate the phase-slowness heterogeneities $\delta p(\theta, \phi; \omega)$ in (2).

2.2 3-D anisotropic tomography

We follow here the formulation of Boschi & Ekström (2002) who invert phase-delay data for 3-D structure directly, without the need of determining preliminary phase-velocity maps. Their formulation is based on local normal-mode theory (Tromp & Dahlen 1992). The phase slowness heterogeneity δp depends only on the local radial structure (e.g. Jordan 1978) which in our case amounts to vertically and horizontally polarized shear velocity heterogeneities δv_{SV} and δv_{SH} (relative to PREM (Dziewonski & Anderson 1981)). The dependence of δp on δv_{SV},

δv_{SH} is described by an integral along the radius (r) of the Earth, weighted by kernel function K_{SV} and K_{SH} (Anderson & Dziewonski 1982). K_{SV} and K_{SH} are defined on the basis of normal mode theory (e.g. Takeuchi & Saito 1972) and depend on the starting model PREM (Dziewonski & Anderson 1981) and the assumed crustal model Crust5.1 (Mooney *et al.* 1998). An important aspect of the local normal-mode approach is that K_{SH} and K_{SV} change as a function of θ, ϕ.

$$\delta p(\theta, \phi; \omega) = \int_0^a K_{SV}(r, \theta, \phi; \omega)\delta v_{SV}(r, \theta, \phi) + K_{SH}(r, \theta, \phi; \omega)\delta v_{SH}(r, \theta, \phi) \, dr. \tag{5}$$

The velocity heterogeneities can be described through the horizontal basis functions $S_i(\theta, \phi)$, a set of vertical basis functions $R_j(r)$ ($j = 1,$ ..., J) and associated unknown coefficients x_{ij}^l
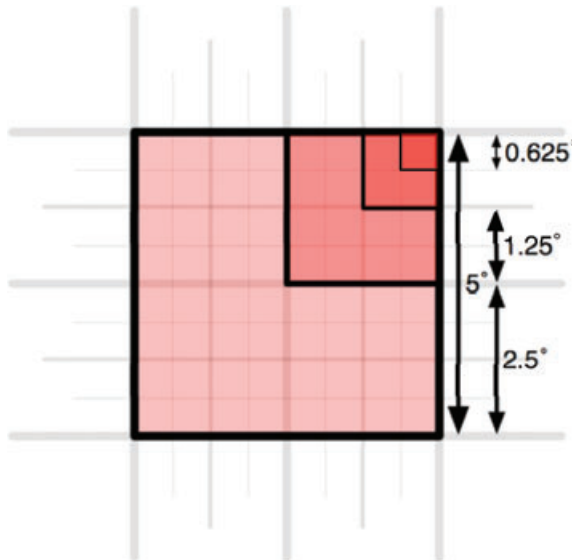
$$\delta v_l(r, \theta, \phi) = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^l \, S_i(\theta, \phi) R_j(r). \tag{6}$$

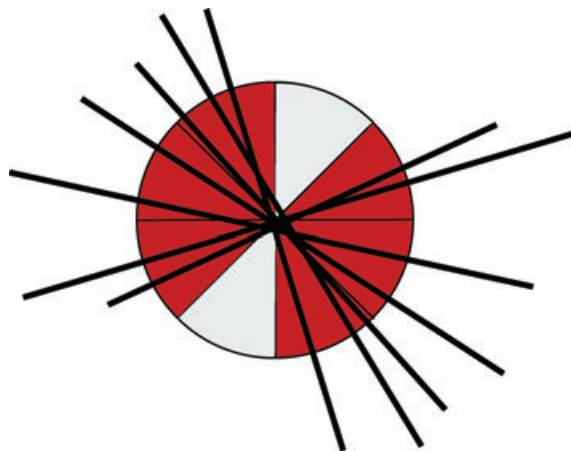Substituting (5) and (6) into (1) we are left with

$$\delta t_m(\omega) = \omega \sum_{l=1}^L \sum_{i=1}^I \sum_{j=1}^J x_{ij}^l \int_0^a \int_{\text{path}_m} K_l(r(s), \theta(s), \phi(s); \omega) \times S_i(\theta(s), \phi(s)) \, R_j(r(s)) \, dr \, ds.$$

Which can be described by a linear system of the same form as (3), with **C** given by

$$C_{ij}^l = \omega \int_0^a \int_{\text{ray path}_m} K_l(r(s), \theta(s), \phi(s); \omega) \, S_i(\theta(s), \varphi(s)) \, R_j(r(s)) \, dr \, ds. \tag{7}$$

Figure 2. The grid is parametrized by four different pixel sizes, where small pixels are chosen to be a simple fraction of bigger ones.

Figure 3. A cell is crossed by rays from various directions. We distinguish four quadrants within the 180° range of possible azimuths and classify the directional resolution as good when at least three quadrants are covered by a predefined minimum number of rays. The cell may be further split into four subcells, as long as this threshold for directional resolution is met.

Whatever is the definition of **C** the unknown coefficients **x**, which describe the model, can be obtained solving the linear system (3) in a least squares sense

$$\mathbf{x} = (\mathbf{C}^T \cdot \mathbf{C})^{-1} \cdot \mathbf{C}^T \cdot \delta\mathrm{t} \tag{8}$$

(e.g. Trefethen & Bau 1997; Boschi & Dziewonski 1999).

2.3 Crustal correction

Particularly at shorter periods, surface waves are not only sensitive to the mantle but also strongly influenced by the crust, as we shall show in detail in Section 5.3.1. Unfortunately, the crust is much more heterogeneous than the mantle, and our data set does not include enough high frequency observations to obtain appropriate resolution at shallow depths. Following Boschi & Ekström (2002), Boschi *et al.* (2004, 2009) and Marone & Romanowicz (2007), we correct *a priori* the data to account for crustal structure. We use the crustal model Crust5.1 (Mooney *et al.* 1998) and correct the data accounting for the effect of each crossed pixel according to the kernels K_{SH} and K_{SV}. Both Crust5.1 and the more recent Crust2.0 are low-resolution crustal models compared to resolution that we aim to achieve. We choose Crust5.1 because its 5° × 5° parametrization overlays perfectly our base grid (Section 3). This should be seen as a temporary solution, before a reference crustal model for the whole European-Mediterranean is defined. Discrepancies between recently proposed models (Tesauro *et al.* 2008; Grad *et al.* 2009) suggest that more work is still needed to reach a consensus (Molinari & Morelli 2011).
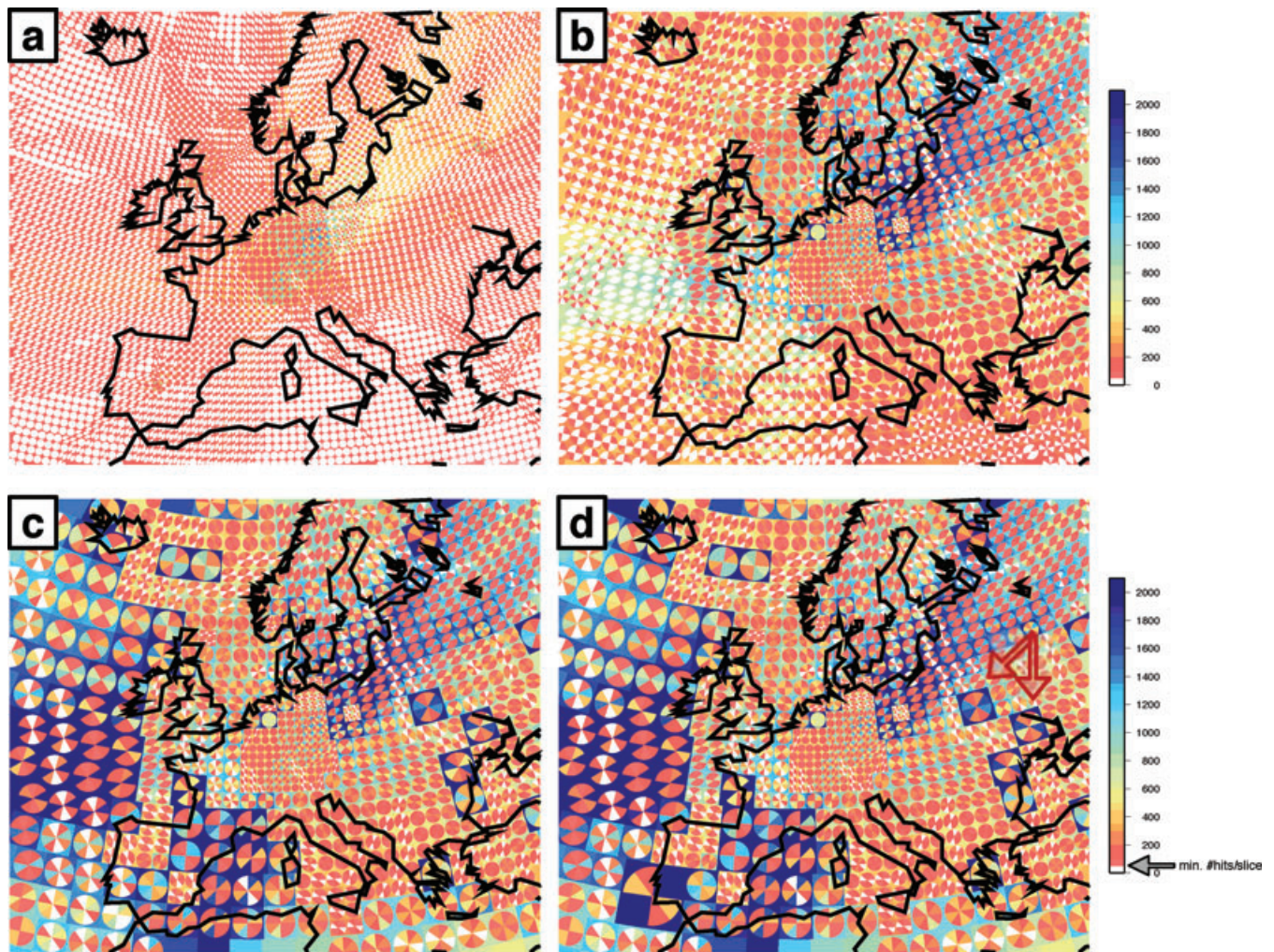
Figure 4. Hit count of Rayleigh waves with a period of 35 s, and four different grids: (a) uniform 0.625°; (b) adaptive with 0.625°, 1.25° pixels; (c) 0.625°, 1.25° and 2.5°; (d) final adaptive grid with 0.625°, 1.25°, 2.5° and 5°. For each pixel the total hit count is shown in the background and the azimuthal coverage by the piediagram. White slices are by definition not adequately resolved due to their hit count. Pixels which have not at least three slices adequately resolved are combined to a bigger pixel in the final grid. The arrows mark cells where the total hit count is much higher than in surrounding ones. The resolution is due to bad crossfiring not high enough to use a finer parametrization.

3 ADAPTIVE PARAMETRIZATION

3.1 Phase velocity (2-D)

Defining the parametrization grid is a stepwise process. In our initial uniform grids pixels are defined to have approximately equal area at all latitudinal zones and thus vary in their longitudinal size (Fig. 1a). We first pick the coarsest pixel size (lateral extent of equal-area pixels at the equator), for example, the 5° uniform pixel grid in Fig. 1(a). Smaller pixels are then defined by iteratively subdividing the coarser ones in four equal parts (Fig. 2).

We consider a pixel to be sampled well if it is (i) sampled by a relatively large number of ray paths associated with the data and (ii) if the azimuths of those ray paths at the pixels are distributed relatively uniformly. In practice, we split the 180° range of possible azimuths in a number of slices (e.g. four 45° slices in the example presented below: see Fig. 3). For each pixel, we count how many of the ray paths that cross it fall in each slice. If a minimum number of paths cross a cell in at least a subset of the slices (three out of four in the example below), we name the cell well-sampled, and allow for it to be split in four smaller cells provided the same resolution 'crossfiring' threshold is still met by each subcell. This minimum number is chosen so that the final parametrization is sufficiently dense to achieve the highest possible resolution in those places where data coverage is maximum without the inverse problem becoming too expensive. After all cells have been so processed, the procedure is repeated on all those smaller cells that have been obtained. This is iterated until no cell can be further split up, or until the *a priori* minimum cell size (0.625° here) is reached. This is defined by the physical resolution limit as defined by the wave length of the shortest wave, that is, the diagonal of the smallest cell is about half the wavelength of the shortest wave in our data set. Fig. 4 shows the hit count for different grids in combination with pieplots describing the angle dependent hit count. White slices are not adequately sampled according to the quality criterion. It can be observed that for a uniform fine parametrization (Fig. 4a) most of the cells do not fulfil our defined requirements of azimuthal coverage. The adaptive parametrization (Fig. 4b) shows fine cells wherever possible and bigger ones where necessary, resulting in improved resolution. Obviously, in certain regions an adaptive grid with four cell sizes (Fig. 4d) is necessary to achieve an almost uniformly adequate resolution with our data set according to the defined requirements to consider a pixel as well sampled when a certain hit count and azimuthal coverage is achieved. Since adequate resolution is defined depending on crossfiring it is possible that the total hit count varies strongly between neighbouring pixel as seen, for example in Fig. 4(d) marked by the red arrow.

To obtain the inverse solution according to eq. (8), $\mathbf{C}$ is first determined for the uniform grid associated with the smallest pixels (0.625°). We then obtain our actual tomographic $\mathbf{C}$ by simply summing the entries associated with insufficiently resolved (according to the minimum number of hits per azimuthal slice) pixels.

3.2 Shear velocity and radial anisotropy (3-D)

The sensitivity of surface waves to 3-D Earth structure is completely different in the vertical and horizontal directions (Fig. 5). Vertical resolution in surface wave tomography may only be achieved by simultaneously inverting data from various wave periods (Fig. 6). Hence, we define our vertical parametrization independently from the horizontal one.

Horizontally an adaptive grid can be defined with the method of Section 3.1. We now select 1800 as the minimum number of hits per azimuthal slice. Fig. 7 shows the horizontal parametrization we use together with the hit count and pieplots describing horizontal crossfiring resolution. We shell show in section 5.4.1 that the regions with high image resolution actually coincide with the regions of high sensitivity of our data set. Since the depth-dependence of sensitivity is constant along each ray path, the geographic distribution of data does not change with depth, and the horizontal parametrization is the same for all depth layers. Because the sensitivity of our data is maximum at shallower depths, layer thickness increases with increasing depth (Fig. 6) and the number of layers has to be chosen carefully. The 3-D inverse problem is
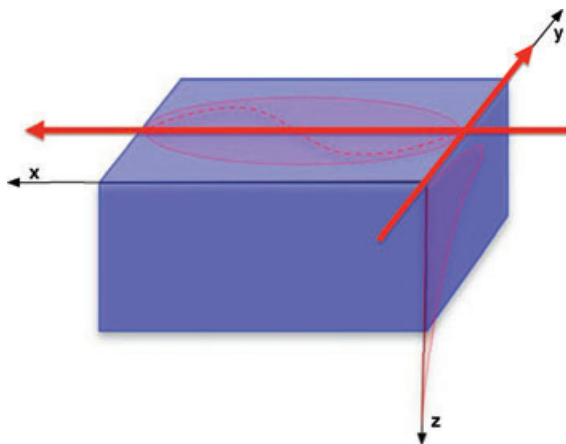

Figure 5. The sensitivity of a surface wave to 3-D velocity is different in horizontal and vertical directions, requiring a decoupling of horizontal and vertical parametrization.
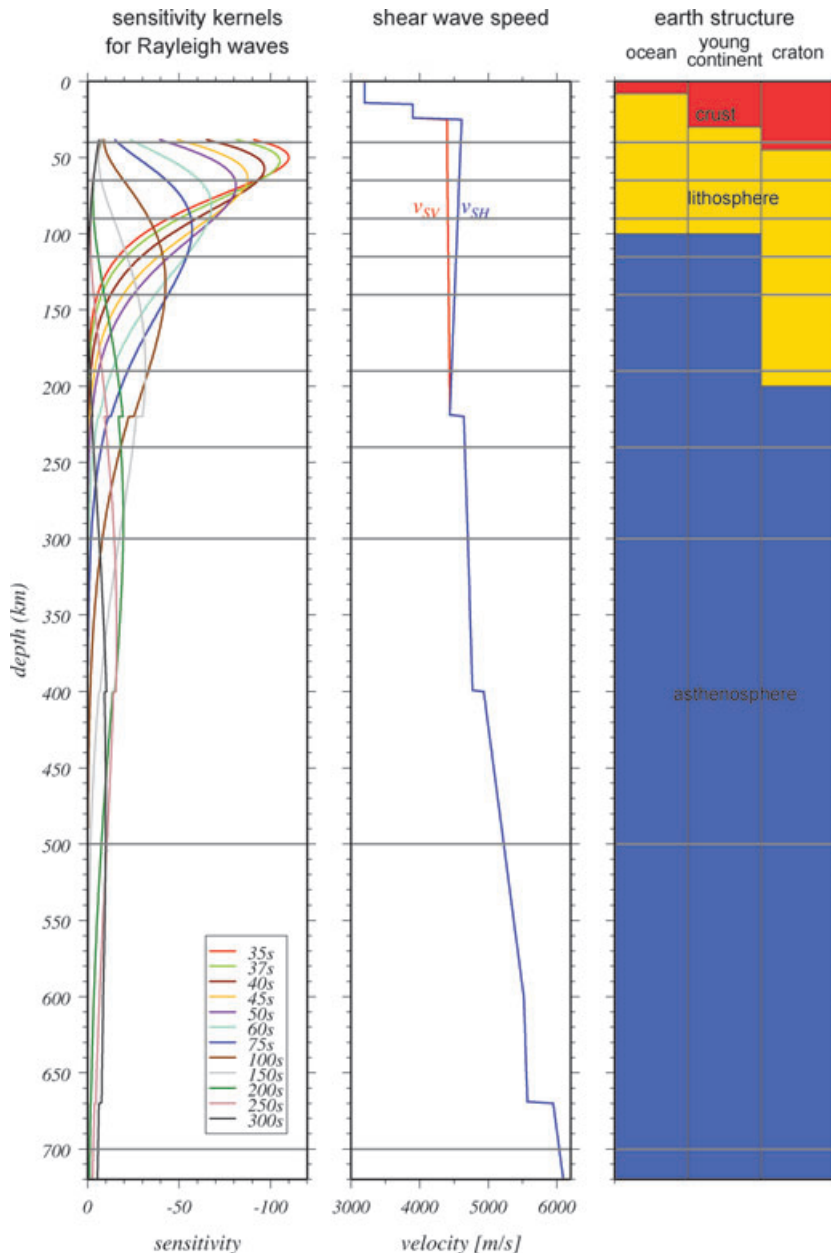
Figure 6. Sensitivity kernels K_{SV} (left-hand side) are calculated for sensitivity to v_{SV} of Rayleigh waves in a 3-D model consisting of PREM combined with the crustal model Crust5.1. We apply *a priori* crustal correction and, therefore, we do not invert for the crust which is equivalent to setting the kernel equal to zero in the corresponding depth range at each location. We show here an example for a pixel to the west of Sardinia. In the middle column shear wave velocity structure according to PREM is shown and on the right average vertical structure for different tectonic regions. The grey horizontal lines mark boundaries between layers. The thickness of the vertical layers is chosen to increase with depth since the overall sensitivity of our database decreases with depth.

naturally much larger than the 2-D one, with a higher computational cost. To partly compensate for this effect while still being able to employ a fine vertical parametrization, we reduce the number of pixels by limiting the adaptively parametrized region to Europe, while keeping a uniform, equal-area ($5° \times 5°$ at the equator) parametrization elsewhere (Fig. 7). Inverting also for heterogeneities outside Europe is important since we use teleseismic data. Otherwise we might map heterogeneities from outside Europe into it.

The sensitivity of surface waves to Earth structure depends on frequency and overtone number. With decreasing frequency and increasing overtone number, surface waves tend to be more structure-sensitive at larger depth. In this study, we only use fundamental-mode surface waves, which have maximum sensitivity in the lower crust and uppermost mantle. In comparison, their sensitivity to transition-zone and lower-mantle structure is negligible. Sensitivity kernels in Fig. 6 equivalent to those of Anderson & Dziewonski (1982) describe the sensitivity of the data used here. As shown by Boschi & Ekström (2002) and Marone & Romanowicz (2007) the sensitivity to mantle structure depends on the local crustal structure. The example shown here is for a pixel in the Mediterranean effectively west of Sardinia. We herein do not invert
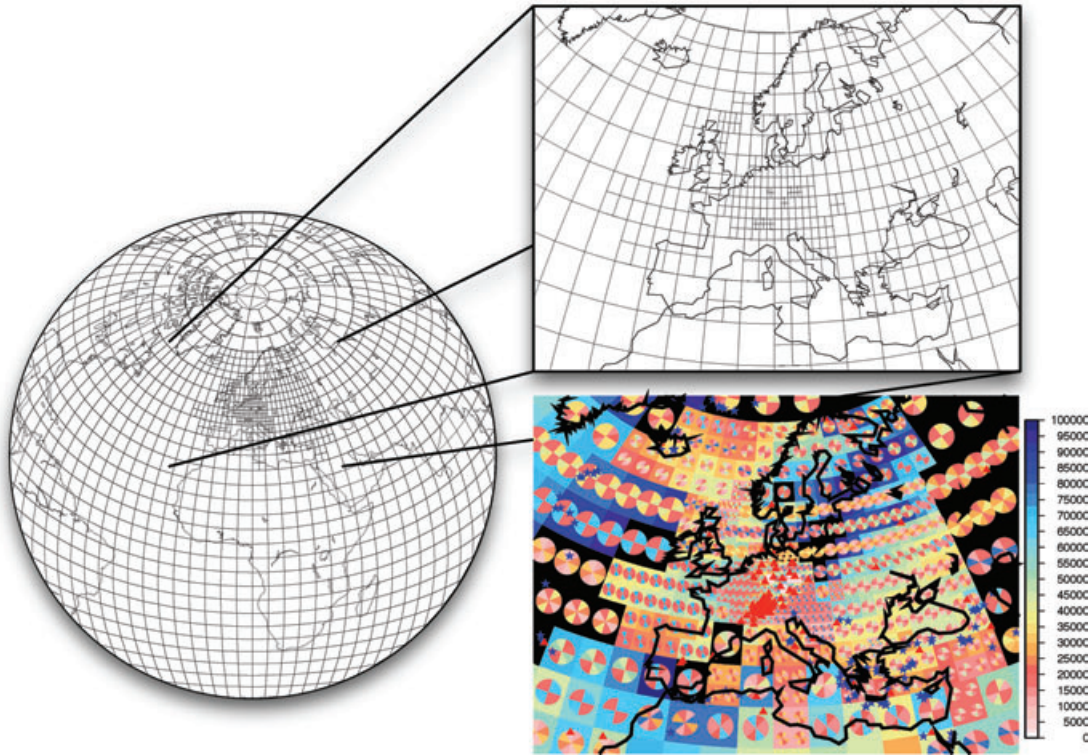
Figure 7. For 3-D inversions the number of free parameters increases drastically. To limit the computational cost of our inversions, we only apply the adaptive parametrization to Europe, the region we are focusing on and where data coverage is best. On the lower right the hit count (colours) is displayed in combination with pieplots describing horizontal crossfiring resolution by angle dependent hit count. Red triangles indicate stations and blue stars earthquake locations.

for crust which amounts to setting K_{SV}, K_{SH} to zero throughout the crust. We use eleven layers down to a depth of 700 km, as illustrated by Fig. 6.

4 ROUGHNESS DAMPING

4.1 2-D horizontal roughness damping: From spherical harmonic to cell domain

Our inverse problem is defined by a coefficient matrix $\mathbf{C}$, model parameters $\mathbf{x}$ and traveltime heterogeneities $\delta\mathbf{t}$ (eq. 3). In seismic tomography the inverse problem is commonly mixed-determined due to non-uniform data coverage and hence eq. (3) can only be solved in least-squares sense (e.g. Trefethen & Bau 1997). Additionally, seismic data are, at least to some extent, noisy and the matrix $\mathbf{C}^T\mathbf{C}$ is singular. 'Regularization' (e.g. Kissling 1988; Inoue *et al.* 1990) must be applied (e.g. Boschi & Dziewonski 1999): by taking *a priori* information about the resulting model into account, the inverse problem can be augmented with some additional rows, which make the matrix invertible. A common method is roughness damping, which requires the solution to be smooth: a meaningful assumption considering that strong small-scale heterogeneities would not be properly accounted for in a linearized (ray theory, but also first-order scattering) regime. Inoue *et al.* (1990) and Boschi & Dziewonski (1999) suggest that in surface wave tomography roughness damping is generally preferable to norm damping.

Boschi & Dziewonski (1999) show how a simple roughness-damping operator is defined for a uniform, approximately equal area pixel parametrization. Our adaptive-grid approach requires a more complicated procedure. We start by noting that the roughness operator can take a particularly simple form if the model is parametrized in terms of spherical harmonics Y_{lm}, that is,

$$\delta p(\theta, \phi) = \sum_{l,m} \delta p_{lm} Y_{lm}, \tag{9}$$

(obtained from eq. (2) after replacing the generic horizontal basis function S_i with Y_{lm}) with l harmonic degree and m order. If the Laplacian, integrated over the whole globe, is used as a measure of model roughness, we can use the fact that spherical harmonics are eigenfunctions of the Laplacian with eigenvalues $l(l+1)$ (e.g. Dahlen & Tromp 1998), and as a result the roughness operator

$$R(\delta p) = \sum_{l,m} l(l+1)\delta p_{lm}^2. \tag{10}$$

We seek the model δp_{lm} such that R is minimum, or

$$\frac{\partial R}{\partial \delta p_{lm}} = 2l(l+1)\delta p_{lm} = 0. \tag{11}$$

The horizontal roughness damping operator defined by eq. (11) in the spherical harmonic domain is a diagonal matrix with entries $l(l+1)$. Regularization amounts to replacing eq. (3) of section 2.1 with

$$\begin{pmatrix} \mathbf{C} \\ \lambda \cdot \mathbf{D_h} \end{pmatrix} \cdot \mathbf{x} = \begin{pmatrix} \delta \mathbf{t} \\ \mathbf{0} \end{pmatrix}, \qquad (12)$$

with λ a 'damping parameter', weighting the regularization constraint with respect to the original inverse problem (e.g. Boschi & Dziewonski 1999).
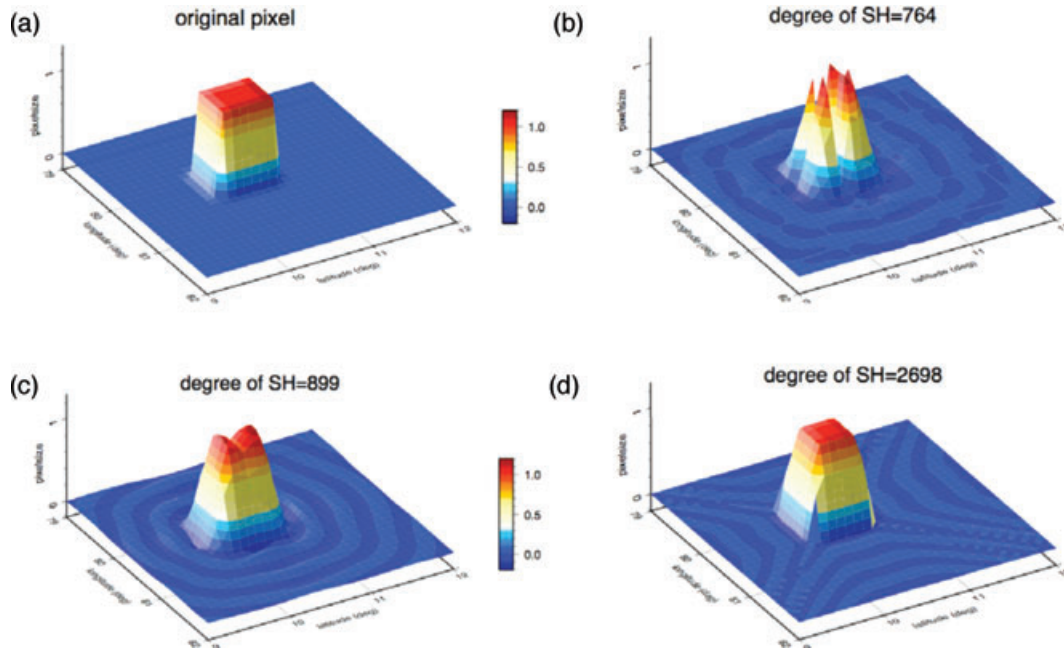
Figure 8. We determine the degree of spherical harmonic coefficients needed to represent an finest pixel in an adequate way. Panel (a) shows the original pixel while panels (b)–(d) show a representation of this pixel by using spherical harmonic coefficients up to a certain degree, as indicated.
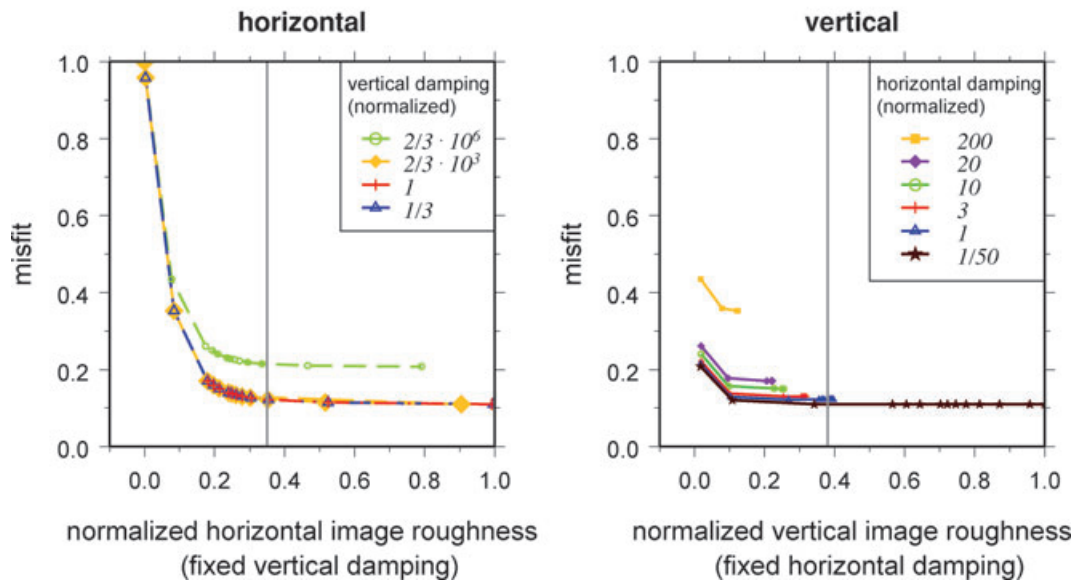
Figure 9. We choose the coefficients λ and μ for horizontal and vertical roughness damping after analysing how the misfit of the resulting models changes as a function of image roughness normalized to its maximum value. In the left panel we show how horizontal image roughness varies with λ, keeping μ constant with the value indicated in the legend. These values are normalized to the finally chosen value since the actual number is depending on the exact construction of the operator. The right panel shows the result of varying μ, with λ fixed as indicated. Grey vertical lines indicate the preferred model.

Now that a simple roughness operator is defined in the spherical harmonic domain, we map it to the adaptive pixel domain. Let us choose as basis function $S_i(\theta, \phi)$ the function whose value is 1 within the ith pixel and 0 elsewhere, and let $S_{i,lm}$ be its harmonic coefficients, such that

$$S_i(\theta, \phi) = \sum_{lm} S_{i,lm} Y_{lm}(\theta, \phi).$$ (13)

Let us rewrite eq. (2) after selecting S_i as our horizontal basis functions

$$\delta p(\theta, \phi) = \sum_i x_i S_i(\theta, \phi).$$ (14)

We next replace $S_i(\theta, \phi)$ in eq. (14) with its spherical harmonic expansion (13), and

$$\delta p(\theta, \phi) = \sum_i x_i \sum_{lm} S_{i,lm} Y_{lm}(\theta, \phi),$$ (15)

from which it follows that the harmonic and pixel coefficients of $\delta p(\theta, \phi)$ are related by

$$\delta p_{lm} = \sum_i x_i S_{i,lm}.$$ (16)

Replacing eq. (16) into (11), we find

$$2l(l+1)\delta p_{lm} = \sum_i [2l(l+1)S_{i,lm}]x_i.$$ (17)

Let us now introduce a linear operator $\mathbf{D_h}$, defined in the pixel domain, acting on the vector x_i ($i = 1, 2, 3, \ldots$) and whose entries are defined by the term in square brackets in eq. (17). This is the roughness damping operator in the pixel domain, used to regularize our inversions. The matrix associated with $\mathbf{D_h}$ has as many rows as there are spherical harmonics in our spherical harmonic domain, and as many columns as there are pixels in our grid. There is no requirement, of course, for the pixel grid to have uniform density. In practice, we need to take into account harmonic degrees sufficiently high to properly represent the finest of our pixels.

A similar approach is used by Soldati *et al.* (2011) to map a spherical-harmonic description of mantle flow into a voxel tomography parametrization.
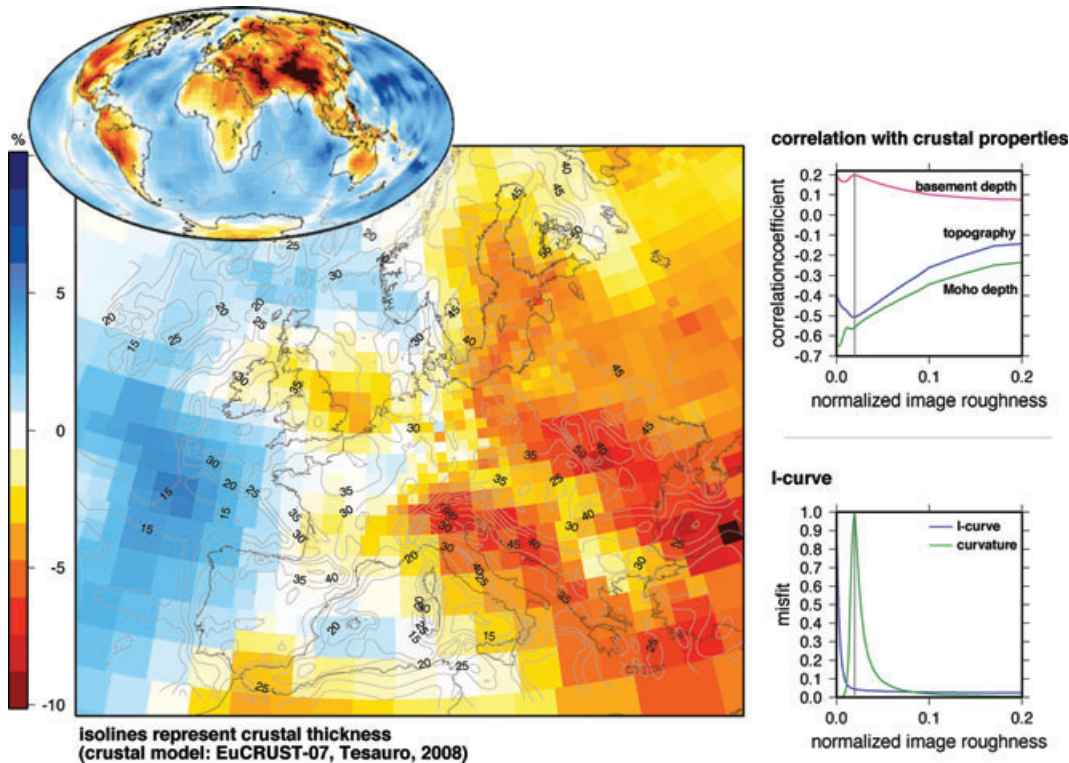
Figure 10. Phase-velocity heterogeneity relative to PREM for Love waves with a period of 35 s. The grey lines indicate isolines representing crustal thickness according to EuCRUST-07 (Tesauro *et al.* 2008). A high correlation to crustal parameters can be observed, which is also shown in the top right panel. The correlation coefficient is calculated for all models and is highest for the model favoured according to the L-curve (bottom right, roughness normalized to the highest value). Values for the shown model are indicated by grey vertical lines.

4.2 2-D horizontal roughness damping: Implementation

To find an appropriate roughness damping operator, we must calculate the spherical harmonic coefficients of each pixel. We use the freely available software archive SHTOOLS (available at http://www. ipgp.fr/wieczor/SHTOOLS/SHTOOLS.html) for this. After a number of preliminary tests we determined that the maximum harmonic degree $l = 899$ was sufficient to resolve even our finest pixels (Fig. 8). Expansion to higher l increases the computational cost excessively. $l = 899$ results in $\sim 10^6$ spherical harmonic coefficients, which is the leading dimension of the damping matrix $\mathbf{D}_h$. The large size of $\mathbf{D}_h$ increases drastically our computer memory requirements. For solving the inverse problem in a least squares sense it is necessary to determine $\mathbf{D}_h{}^T\mathbf{D}_h$ (eq. 8). Due to the size of the matrix this step in our work has the highest computational cost, but use of parallel operations is possible.

4.3 3-D roughness damping

In the 3-D case we invert for radial anisotropy (horizontal and vertical shear wave velocity, where horizontal shear wave velocity is azimuthally averaged) for n layers. The parameter vector $\mathbf{x}$ is defined

$$\mathbf{x} = \begin{pmatrix} \delta \mathbf{v}_{sh}^{layer1} \\ \vdots \\ \delta \mathbf{v}_{sh}^{layern} \\ \delta \mathbf{v}_{sv}^{layer1} \\ \vdots \\ \delta \mathbf{v}_{sv}^{layern} \end{pmatrix}. \tag{18}$$

Regularization of the inverse problem is now given by horizontal ($\mathbf{D}_h^{3D}$) and vertical ($\mathbf{D}_v^{3D}$) roughness damping and anisotropy damping $\mathbf{D}_a^{3D}$

$$\begin{pmatrix} \mathbf{C} \\ \lambda \cdot \mathbf{D}_h^{3D} \\ \mu \cdot \mathbf{D}_v^{3D} \\ \nu \cdot \mathbf{D}_a^{3D} \end{pmatrix} \cdot \mathbf{x} = \begin{pmatrix} \delta \mathbf{t} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{19}$$

where in the following we shell refer to the matrix at the left-hand side as $\mathbf{G}$.
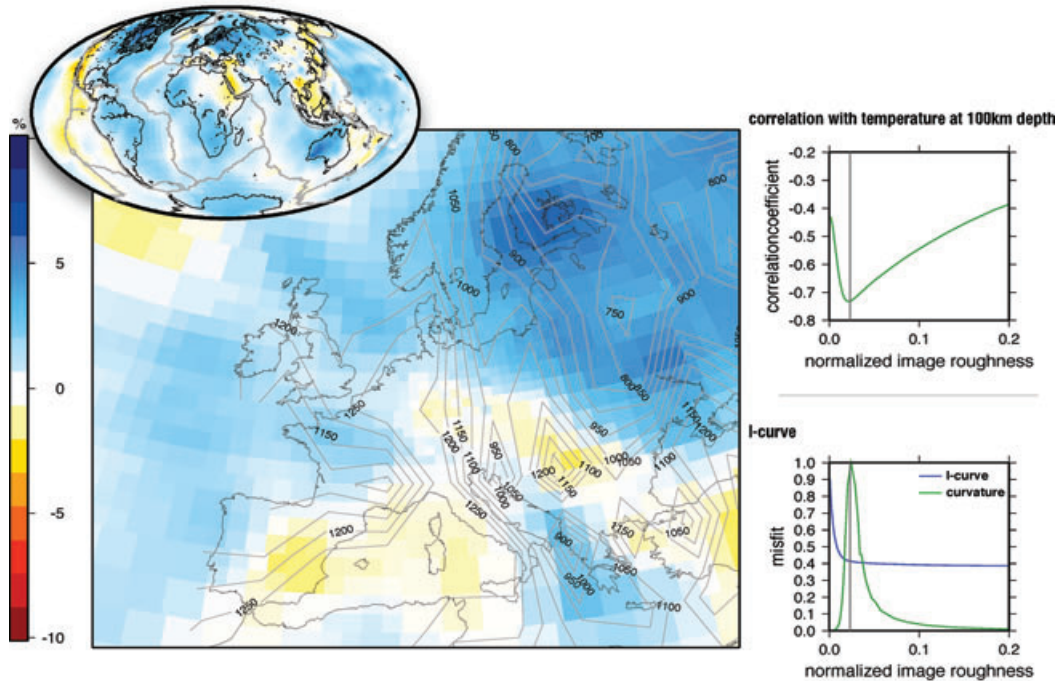
Figure 11. Phase-velocity heterogeneity relative to PREM for Rayleigh waves with a period of 75 s. The grey lines indicate isolines representing temperature at 100 km depth according to the geothermal data set TC1 (Artemieva 2006). Temperature in this depth is an indicator for the thickness of the lithosphere, that is, thin lithosphere can be identified by high temperature. Temperature and velocity anomalies are strongly anticorrelated (top right-hand panel), and the anticorrelation is highest for our preferred model based on the L-curve (bottom right, roughness normalized to the highest value). The preferred model is marked by a grey vertical line.

Anisotropy can be damped by minimizing the difference between δv_{SV} and δv_{SH}. Our preliminary tests show that also with no damping ($\nu = 0$) the v_{SV}, v_{SH} pattern in globally best resolved areas (e.g. Pacific) is similar to that found in previous studies (Ekström & Dziewonski 1998). We therefore chose to apply no anisotropy damping.

4.3.1 3-D: Horizontal roughness damping

The operator $\mathbf{D_h}$ defined in Section 4.1 by eq. (17) can naturally be applied to any scalar field defined on the sphere. In the following we shall apply it to velocity heterogeneities δv_{SH} and δv_{SV} each for all n layers resulting in a diagonal matrix with $2n$ times $\mathbf{D_h}$ on the diagonal

$$\mathbf{D_h^{3D}} = \begin{pmatrix} \mathbf{D_h} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D_h} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{D_h} \end{pmatrix} \qquad (20)$$

4.3.2 3-D: vertical roughness damping

Vertically adjacent voxels have the same latitudinal-longitudinal extent (note that, as a result, deeper voxels have smaller volume) and should be about equally resolved if the geographic coverage of the various periods is about the same. This allows us to choose for simplicity a much
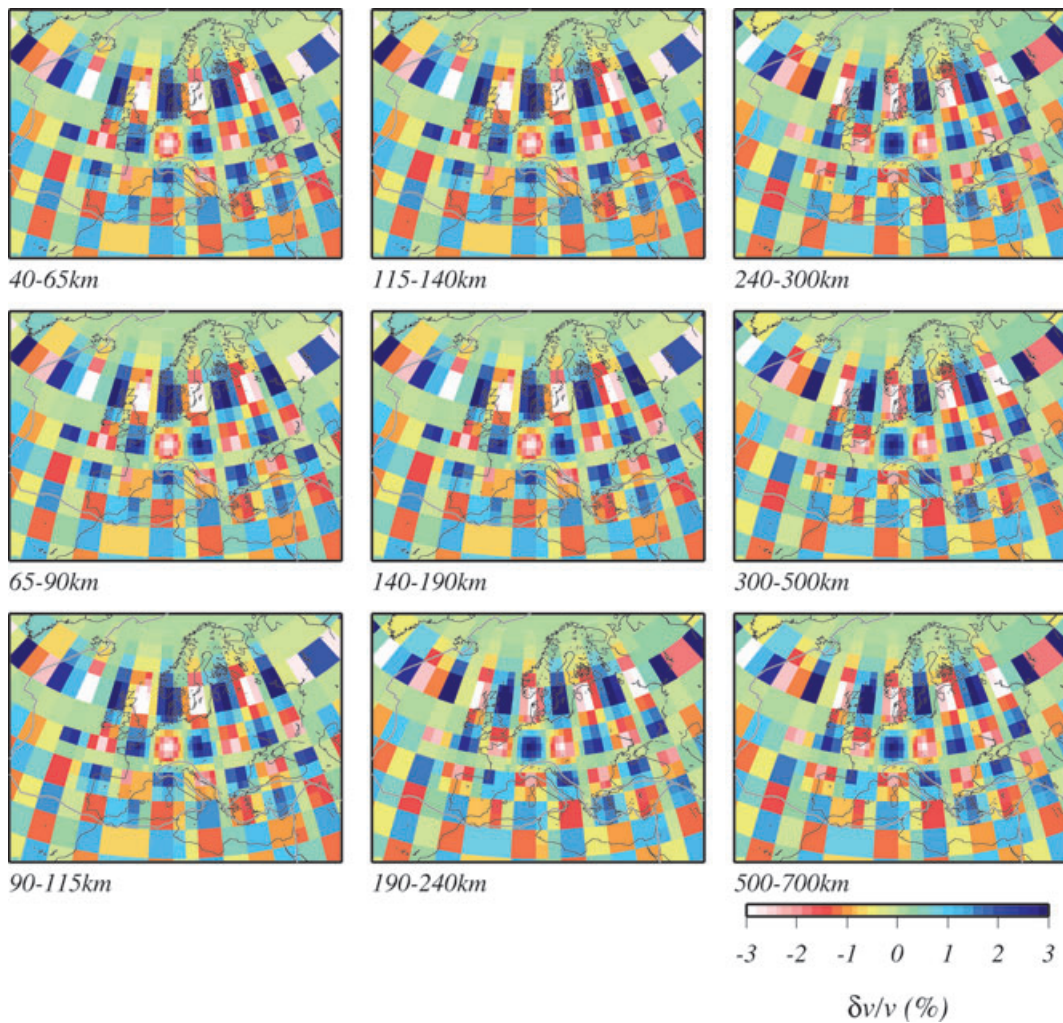
Figure 12. Isotropic input model for the checkerboard test. The pattern coincides with a spherical harmonic degree 42 and order of 21, projected on our adaptive grid. The nine panels correspond to layers with the indicated depth range in our model.

simpler form of the damping operator $\mathbf{D}_v^{3D}$ with respect to $\mathbf{D}_h^{3D}$. We define on the diagonal of $\mathbf{D}_v^{3D}$ two submatrices for the two parameters. For each matrix the ith row has only two non-zero entries: the diagonal one, equal to $+1$, and the $(n+i)$th one, equal to -1, with n the number of pixels per layer:

$$
\mathbf{D}_v^{3D} = \begin{pmatrix}
1 & 0 & \cdots & 0 & -1 & 0 & \cdots & 0 & & & & & & & & \\
0 & \ddots & 0 & \cdots & 0 & \ddots & \ddots & \vdots & & & & & & & & \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & & & 0 & & & & & \\
0 & \cdots & 0 & 1 & 0 & \cdots & 0 & -1 & & & & & & & & \\
& & & & & & & & 1 & 0 & \cdots & 0 & -1 & 0 & \cdots & 0 \\
& & & & & & & & 0 & \ddots & 0 & \cdots & 0 & \ddots & \ddots & \vdots \\
& & 0 & & & & & & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\
& & & & & & & & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & -1
\end{pmatrix}.
$$
(21)

Eq. (19) results in the context of the least-squares problem (eq. 8) to

$$
\mathbf{x} = \left(\mathbf{C}^T \cdot \mathbf{C} + \lambda^2 \mathbf{D}_h^{3D\,T} \cdot \mathbf{D}_h^{3D} + \mu^2 \mathbf{D}_v^{3D\,T} \cdot \mathbf{D}_v^{3D} \right)^{-1} \cdot \mathbf{C}^T \cdot \delta t.
$$
(22)

Figure 13. Output model from the checkerboard test, v_{SV} heterogeneity. Good sensitivity is down to ~115 km, fair down to ~500 km. We note a resolution problem achieved in depth range 190–240 km resulting from to thin layering.

4.4 Choice of the damping coefficients

To estimate optimal damping values we apply the L-curve approach proposed by Hansen (1992) and similar to Eberhart-Phillips (1986), Kissling (1988) or Boschi (2006). Figs 10 and 11 show applications of this approach for the choice of adequate damping to derive phase-velocity models. Comparing resulting models obtained with different values for λ when correlated to geological features to which they are supposed to be mostly sensitive, we observe that our criterion for choosing the damping parameter is sound.

In 3-D, we define image roughness

$$\sum_{j=1}^{2 \times n \times (l+1)^2} \left(\sum_{i=1}^{N_{par}} D_{h,ji}^{3D} \cdot x_i \right)^2 \tag{23}$$

and data misfit

$$\frac{\sum_{j=1}^{N_{obs}} \left(\sum_{i=1}^{N_{par}} C_{ji} \cdot x_i - \delta t_j \right)^2}{\sum_{j=1}^{N_{obs}} \delta t_j^2}, \tag{24}$$

and monitor their variations for different values of λ and μ. An acceptable damping should result in a model that is not exceedingly rough but fits the data sufficiently well. Fig. 9 shows for the 3-D case horizontal and vertical image roughness each plotted against the data misfit in analogue manner as for body waves proposed by Eberhart-Phillips (1986) who choose a damping value which greatly reduced the data variance with a moderate increase in the model roughness. We keep the vertical damping parameter constant while varying the horizontal one, and vice versa. This analysis is repeated for different values of the fixed parameter. Due to difference in horizontal versus vertical resolution capability of surface waves the resulting damping is different. As to be expected, different values for the vertical damping do not significantly influence horizontal roughness and vice versa.
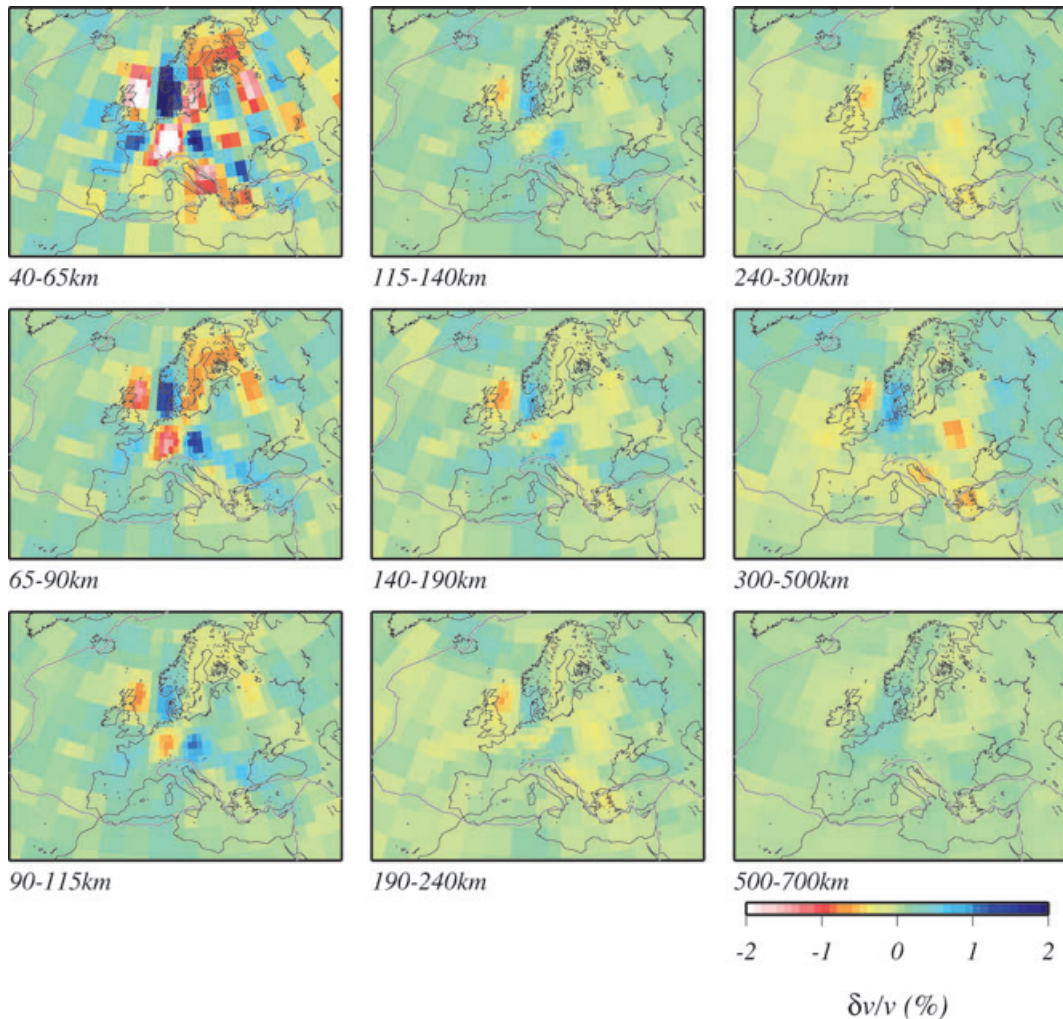

Figure 14. Output model from the checkerboard test, v_{SH} heterogeneity. Fair sensitivity is limited within the lithosphere, that is, top three layers. Below 115 km depth the layers in this test are obviously to thin: note the somewhat better resolution capability in the thick layer 300–500 km.

We choose for the 2-D inversions a roughness damping operator λ of 0.011 in the case of Love-waves with a period of 35 s and of 0.017 in the case of Rayleigh-waves with a period of 75 s. For 3-D inversions we choose as a coefficient for horizontal roughness damping $\lambda = 500$ and for the vertical roughness damping $\mu = 1500$.

5 APPLICATION TO EUROPE

5.1 Data set

We use a database of global teleseismic phase-velocity measurements of Love and Rayleigh waves with a period range between 35 and 300 s compiled and described by Fry *et al.* (2008). This database, focused on Europe, includes the global data of Ekström *et al.* (1997) and a later update of these measurements by Boschi & Ekström (2002). The dense coverage of Europe was achieved by Fry *et al.* (2008) including data from the European networks MidSEA (van der Lee 2001), SDSNet (Baer *et al.* 2000), TomoCH (Fry *et al.* 2008) and GRSN (Henger *et al.* 2002). Overall the database consists of $\sim 10^6$ measurements of phase anomaly.

5.2 Computational issues

We have to invert eq. (19) in least-squares sense (eq. 22). In the 3-D case the inversion problem is defined by $\sim 10^6$ observations and 5.1×10^4 parameters. The inversion can be implemented in various ways (e.g. Boschi & Dziewonski 1999) and the inversion algorithm to be used should be selected based on the properties of the matrix that has to be inverted. In our case the matrix $\mathbf{G}$ (eq. 19) to be inverted is relatively dense, because (i) a surface wave samples, at each location along its path, a large portion of the upper mantle (hence ~ 10 layers, while a body wave typically samples a small depth range at each location), and (ii) the regularization matrix $\mathbf{D}_h^{2D}$ is full, resulting in a dense 3-D regularization operator. The matrix $\mathbf{G}^T \cdot \mathbf{G}$, needed to solve the inverse problem in a least-squares sense is positive-definite by construction, is also dense as a result. The most efficient algorithm to invert a dense positive-definite matrix is Cholesky factorization (e.g. Trefethen & Bau 1997). Cholesky factorization is also preferable to conjugate-gradients (CG) type algorithms because it reduces RAM-requirements: CG requires the storage of $\mathbf{G}$, a rectangular $(N_{obs} + N_{D_h^{3D}} + N_{D_v^{3D}}) \times N_{par}$, with N_{obs} the number of inverted data, $N_{D_h^{3D}}$ the number of rows in D_h^{3D}, $N_{D_v^{3D}}$ the number of rows in D_v^{3D}, N_{par} twice the number of model voxels (since we invert for v_{SV} and v_{SH}), and typically $(N_{obs} + N_{D_h^{3D}} + N_{D_v^{3D}}) \gg N_{par}$; Cholesky factorization operates directly on $\mathbf{G}^T \cdot \mathbf{G}$, whose dimension is only $N_{par} \times N_{par}$ (needed RAM is further reduced of a factor 2 because $\mathbf{G}^T \cdot \mathbf{G}$ is symmetric). This is significant due to the high number of observations and parameters in our case.

The density of $\mathbf{G}^T \cdot \mathbf{G}$ also makes it hard to effectively parallelize the Cholesky factorization process. Using the Intel® Math Kernel Library (MKL; http://software.intel.com/en-us/articles/intel-mkl/) on a Linux cluster, with the benefit of automatic parallelization in a multithreading context, we achieve a speed-up of 10 on 16 processors compared to a single processor.

5.3 2-D tomography: phase-velocity maps

We next use our method to map phase-velocity heterogeneities relative to PREM for Love waves with a period of 35 s, and Rayleigh waves with a period of 75 s. These periods are chosen owing to the quite different depth-dependence of their sensitivities. While the first one is mainly sensitive to the crust and uppermost mantle lithosphere, the latter has maximum sensitivity at 100 km depth (Fig. 6) and we compare the heterogeneities to the depth of Moho and lithosphere asthenosphere boundary, respectively. We conduct a number of inversions varying
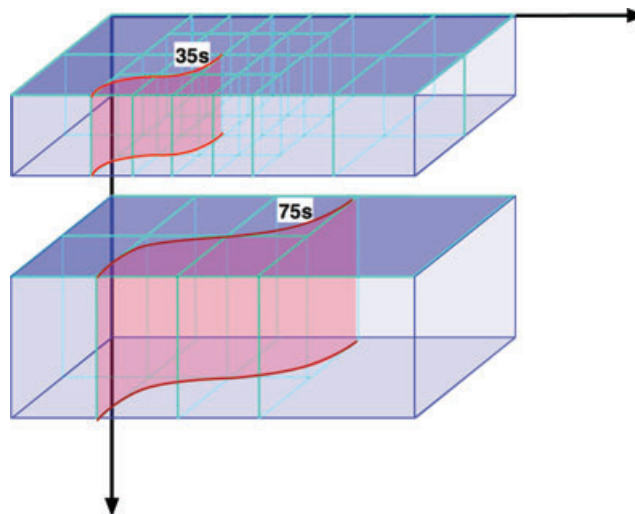
Figure 15. Deeper layers are mainly resolved by surface waves with larger wavelength, thus the horizontal physical resolution limit is lower and the minimal pixel size should be chosen bigger.

the roughness damping parameter λ, and choose the model corresponding to the point of maximum curvature of the L-curve as the preferred model (e.g. Boschi 2006).

5.3.1 Love waves, 35 s

In Fig. 10, we show phase-velocity heterogeneities of Love waves with a period of 35 s. This surface wave mode is sensitive to very shallow structures with still relatively high sensitivity in the crust. It was thus interesting to calculate the correlation between observed heterogeneities (including all models obtained from various values of λ) and crustal parameters like Moho depth, basement depth and topography from EuCrust-07 (Tesauro *et al.* 2008). A significant correlation to crustal parameters can be observed, suggesting that variations in crustal structure are the main cause of velocity heterogeneities. The correlation between crustal thickness (isolines in Fig. 10) and δp is particularly high: continents are generally slower than oceans, with an especially low velocity in regions of thick crust, for example, Alps, Himalayas and Andes. This effect is expected based, for example, on Fig. 6. In regions of thin crust, the speed of surface wave propagation (which is eventually a weighted average of shear velocity over depth) is mostly controlled by fast mantle material, while crustal elastic properties are dominant in regions of thick crust. A comparison with the L-curve shows that correlation is highest for our preferred model (maximum of L-curve curvature, $\lambda = 0.011$).

5.3.2 Rayleigh Waves, 75 s

We show in Fig. 11 phase-velocity heterogeneities for Rayleigh waves with a period of 75 s. As can be seen from the sensitivity kernels (Fig. 6) this surface wave mode has its main sensitivity at a depth of 100 km.
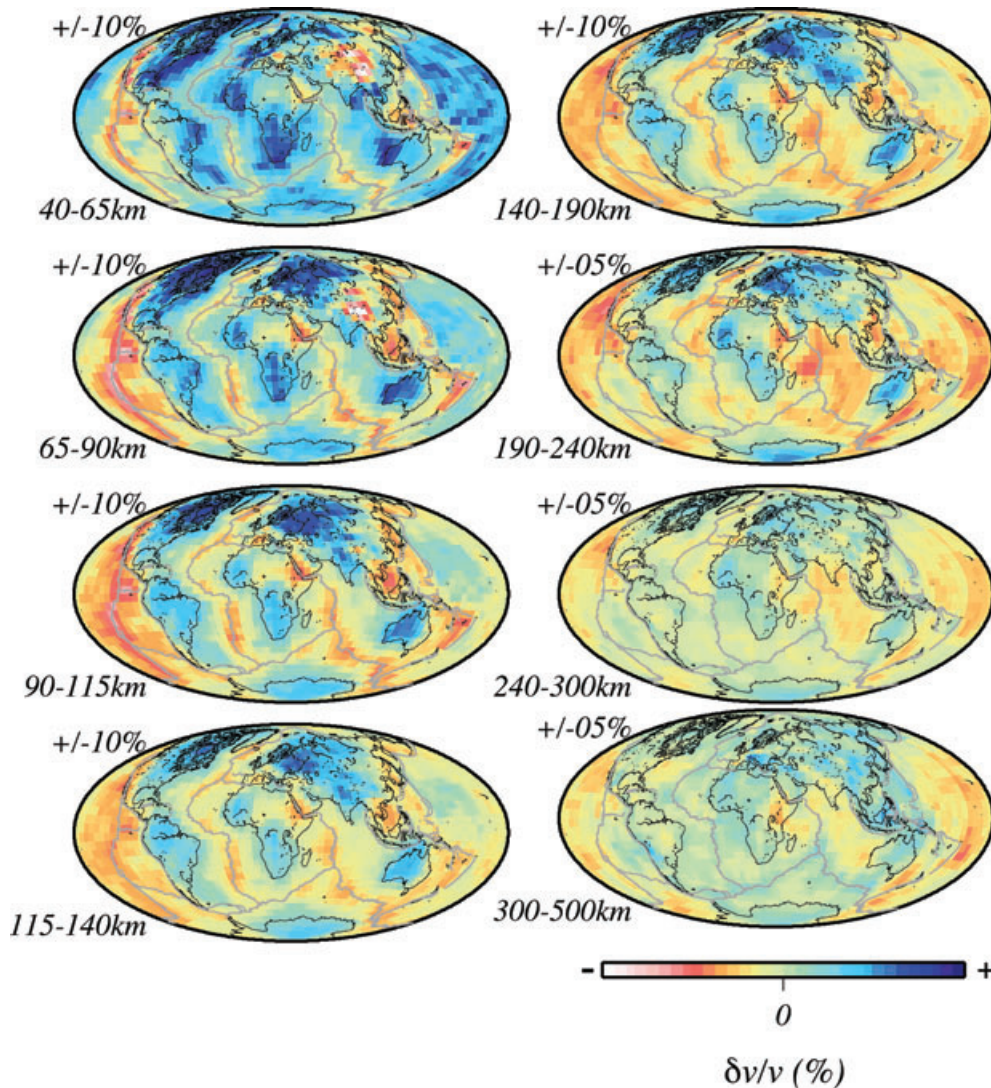
Figure 16. Heterogeneities of v_{SV} in FMADVOXEU, relative to PREM. The colourscale varies with depth as indicated in upper left corner of each panel corresponding to a different depth layer, as indicated in lower left corner. Details on the model are described in the text.

Young continental lithosphere and old oceanic lithosphere have both a thickness of about 100 km. The temperature in this depth gives a good estimate on the thickness of the lithosphere, that is, thin lithosphere can be observed due to higher temperature caused by more shallow mantle material. Velocity variations due to differences in the thickness of the lithosphere should therefore correlate with temperature variations in that depth range. We calculated the correlation between our velocity heterogeneities and temperature at 100 km depth according to the geothermal model TC1 (Artemieva 2006). The correlation coefficient reaches a value of −0.73 showing a significant correlation between high velocities and low temperature. As in the case of Love waves with a period of 35 s the highest correlation is reached for our preferred model according to the L-curve ($\lambda = 0.017$). The anticorrelation is largely explained by the anomalously large thickness of the lithosphere under cratons, resulting in low temperatures and high velocities (russian platform in Fig. 11). The same phenomenon takes place in subduction zones (Aegean arc in Fig. 11), while areas of thin lithosphere like the western Mediterranean, the Pannonian Basin and the Aegean are hot and slow.

5.4 3-D tomography: new anisotropic shear velocity model for Europe

5.4.1 Sensitivity tests

To estimate the sensitivity of our data set in context of our adaptive model parametrization we carried out a checkerboard test. Fig. 12 shows the input model characterized by heterogeneities with an amplitude of 3 per cent. The model is based on the spherical harmonic function of degree 42 and order 21, projected on the adaptive grid used here. In the vertical direction, a sharp change of sign is imposed across the 190 km parametrization boundary. From this model we calculate synthetic data through the forward implementation of eq. (3) with **x** the input model. We subsequently generate random Gaussian noise with a standard deviation of 10 s and sum it to the data. The value is a rough estimate of
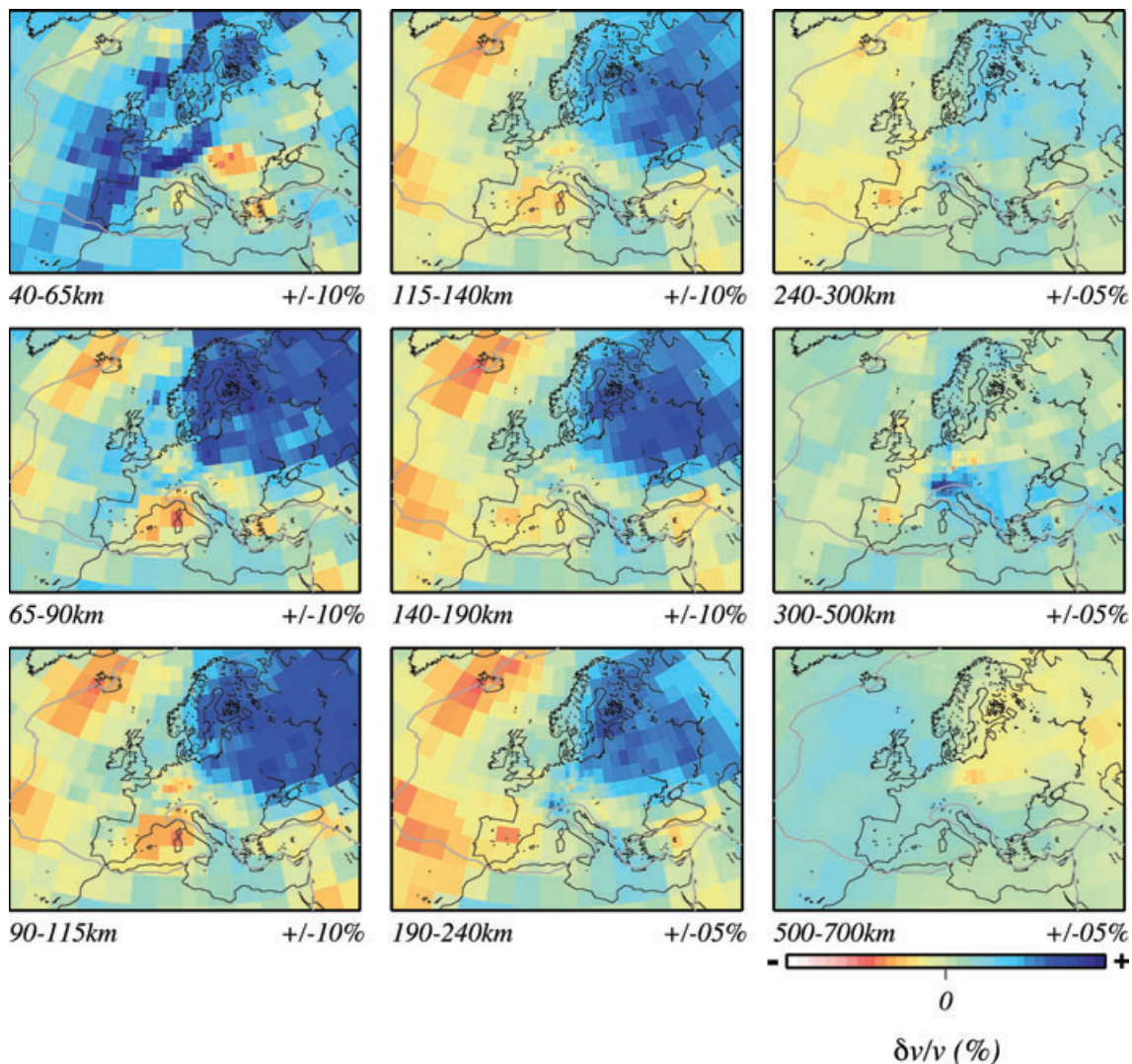
Figure 17. Detail of Fig. 16: Heterogeneities of v_{SV} in Europe and the Mediterranean Basin. The colourscale varies with depth as indicated in lower left corner of each panel corresponding to a different depth layer, as indicated in lower right corner.

data error based on Ekström *et al.* (1997) and Visser *et al.* (2008a,b). Fig. 13 shows sensitivity to v_{SV}, mainly given by Rayleigh waves, from the inversion of the synthetic noised data. Sensitivity is good down to ∼115 km and fair to a depth of ∼500 km. The apparent decrease in sensitivity around 190 km depth is caused by smearing of the sharp change in sign in the input model in vertical direction. This smearing effect points out that layer thickness is chosen to thin in this depth considering the sensitivity of the data (compare also sensitivity kernels in Fig. 6) to distinguish between these two layers. Future models might consider this by joining two layers (e.g. 190–240 km and 240–300 km) to obtain one well-resolved layer. Fig. 14 shows the v_{SH} component of the output model. Resolution capability decreases drastically with depth since v_{SH} is mainly constrained by Love waves, which have a more limited depth resolution capability down to ∼115 km only.

This study is limited to fundamental modes, which are sensitive to a relatively small depth range in the uppermost mantle. In future applications when mapping larger depths, we will hence to account for the depth-dependence of the surface wave resolution capability limit. Essentially, larger depths are sampled by larger periods, hence larger wavelength and a lower resolution capability (Fig. 15). We consider this methodological improvement to go beyond the scope of this paper and leave it for future work.

5.4.2 Results

In Figs 16–21, we illustrate our new anisotropic shear velocity model FMADVOXEU (Fundamental Mode ADaptive VOxel model of EUrope). We show v_{SV} (vertically polarized shear velocity) and v_{SH} (horizontally polarized) heterogeneities relative to PREM on a global scale (Figs 16 and 18) and with increased image resolution for Europe (Figs 17 and 19) as well as absolute values for anisotropy parameter $\xi = v_{SH}^2/v_{SV}^2$ (Figs 20 and 21). According to tests (Section 5.4.1) our data is sensitive to v_{SV} down to ∼115 km and fair down to ∼500 km; v_{SH}, and thus anisotropy, only down to ∼115 km.

We can clearly distinguish the first three layers which results in a better resolution of the lithosphere than has been achieved before, for example Boschi *et al.* (2009) have just one layer down to 100 km depth. Below vertical smearing can be observed as seen in earlier studies. The amplitudes decrease like the sensitivity of the surface waves does. As described in Section 5.4.1 layer thickness is chosen to thin.

Compared to earlier global studies (e.g. Debayle *et al.* 2005; Lebedev & van der Hilst 2008) as well as regional studies (e.g. Pilidou *et al.* 2004; Boschi *et al.* 2009) we improve in the correlation with lithospheric structure. All of these models are able to identify the Transeuropean Suture Zone, marking the transition from relatively young continental western Europe to the thicker and seismically faster lithosphere of the eastern and northern European Craton. This feature is prominently visible in layers below 90 km depth. We are now able to identify more local variations in lithospheric thickness. An example for this is the narrow cratonic 'keel' underneath southern Finland (Sandoval *et al.* 2004), clearly visible down to ∼240 km. The very deep lithosphere-asthenosphere boundary in this area has been found by the teleseismic body wave study of Sandoval *et al.* (2004), and it is vaguely visible in the study of Pilidou *et al.* (2004) but so far it was not resolved in global surface wave studies.

Compared to earlier studies like Boschi *et al.* (2009), who use the same data set, the model shows a global improvement especially in v_{SH}. Their model shows strong vertical smearing, for example, in regions of the mid-oceanic ridges down to 200 km depth while in our results for v_{SH} the low velocity anomalies are significant only in the first layer. Also the model can distinguish between young continental and old oceanic lithosphere, which have both a thickness of about 100 km, but due to vertical resolution problems this pattern is smeared to about 300 km. Our model shows a fast decrease in the amplitude of the mid-oceanic ridge anomaly and is able to distinguish between continental and oceanic lithosphere as can be observed, for example, for Africa and northern Europe, without strong vertical smearing. The increase in the amplitude for the lowest layers is obviously an artefact. Another improvement is the mapping of the former Tethyan convergence belt, especially in the second layer. Generally, our model is less patchy, showing only features, which are expected in the geological context. The first layer is still strongly influenced by crustal effects. Thus the only low velocity feature is the Panonnian Basin due to the upwelling and the accordingly thin lithosphere. Our results for the deeper layers are in general agreement with large-scale features which were observed in earlier studies. Several geological features can be identified for vertical and horizontal shear velocity heterogeneities which are in agreement with Boschi *et al.* (2009) and thus, as shown for that model, also with Piromallo & Morelli (2003), Boschi *et al.* (2004), Schmid *et al.* (2008) and Peter *et al.* (2008). The high velocity underneath the Alps and the Aegean Arc (stronger in v_{SV} than in v_{SH}) should be caused by the slab of the subducting plate.

Low-velocity heterogeneities are observable in both components for the following features: from ∼65 km downwards along the Mid Atlantic Ridge and under Iceland, connected to upwelling of warmer mantle material; under the Western Mediterranean and Aegean Sea, due to hot asthenosphere beneath thin crust (compare isolines in Fig. 10 for crustal thickness and Fig. 11 for temperature in 100 km depth); under Sardinia and Corsica. Some features are more prominent in v_{SV} than in v_{SH}: under the Pannonian Basin in the upper ∼150 km, the mantle is relatively hot resulting in low velocity; a low velocity feature in Germany might be connected to upwelling at the Rhine Graben and/or the Eiffel Plume; possible local upwellings under Iberia (Boschi *et al.* 2009) can be recognized in the 140–500 km depth range.

At the global scale, the anisotropy we map is consistent with the general pattern proposed in other studies. Fig. 20 shows that v_{SH} is overall higher than v_{SV} in the top 200 km of the mantle, as to be expected and consistent with the study of Chang *et al.* (2010) for the Tethyan margin. The anisotropy we find under the region of the Pacific Ocean is similar in pattern to that observed in previous studies (e.g. Ekström & Dziewonski 1998).

For Europe (Fig. 21) we can observe $v_{SH} > v_{SV}$ for almost the whole frozen-in anisotropy in the lithosphere (top 80–100 km). Below the lithosphere the pattern of ξ changes drastically. Unfortunately, according to the tests in Section 5.4.1 our data are not anymore very sensitive to v_{SH} and thus ξ from about ∼140 km depth down. Thus tomographic results and their interpretation should be regarded as preliminary.
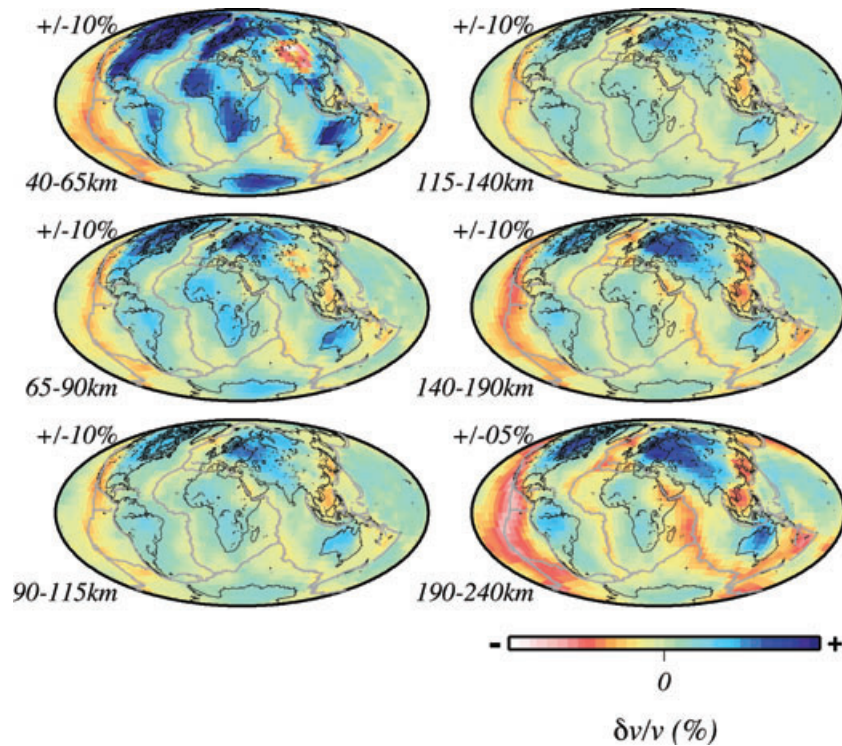
Figure 18. Heterogeneities of v_{SH} in FMADVOXEU, relative to PREM. The colour scale varies with depth as indicated in upper left-hand corner of each panel corresponding to a different depth layer, as indicated in lower left corner. Details on the model are described in the text.

Lattice preferred orientation (LPO) of olivine crystals is probably the main cause of anisotropy in the mantle (Long & Becker 2010). For A- and E-type LPO, which are most likely the dominant fabrics at asthenosphere depth, the fast axes of olivine crystals tend to align in the direction of shear and thus in the direction of mantle flow (Long & Becker 2010). Thus subduction processes accompanied by downwelling as well as upwelling are expected to be observations of $v_{SV} > v_{SH}$. In our anisotropy model of Europe (Fig. 21) this form of anisotropy is found
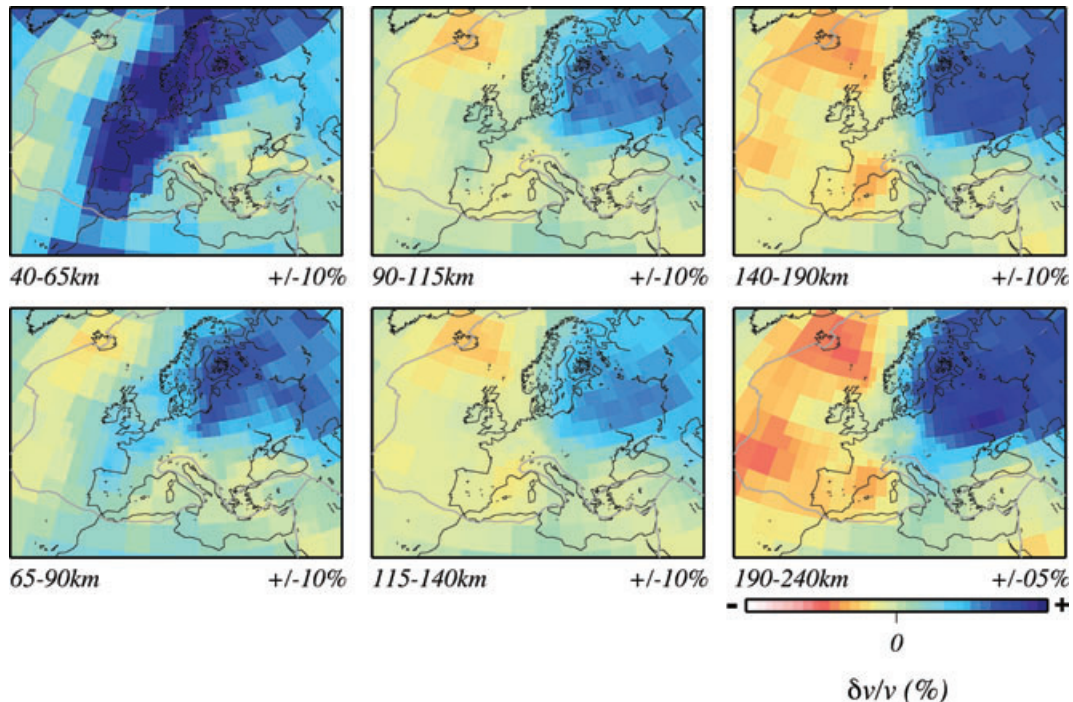
Figure 19. Detail of Fig. 18: Heterogeneities of v_{SH} in Europe and the Mediterranean Basin. The colour scale varies with depth as indicated in lower left-hand corner of each panel corresponding to a different depth layer, as indicated in lower right-hand corner.
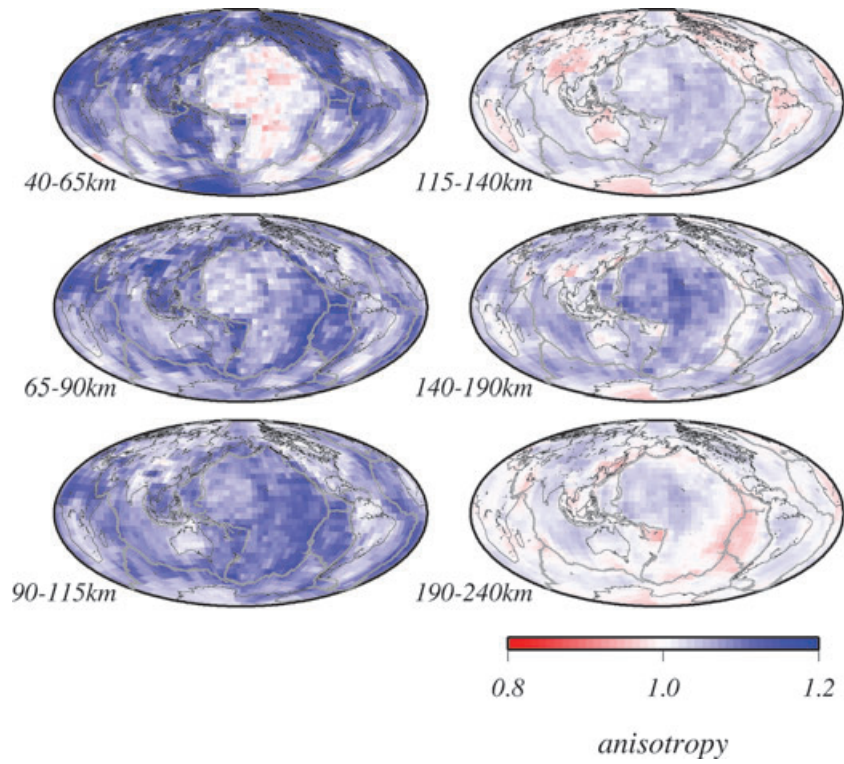
Figure 20. Anisotropy parameter $\xi = v_{SH}^2/v_{SV}^2$ according to model FMADVOXEU. The depth range of each layer is given in the lower left-hand corner of each panel. Details on the model are described in the text.
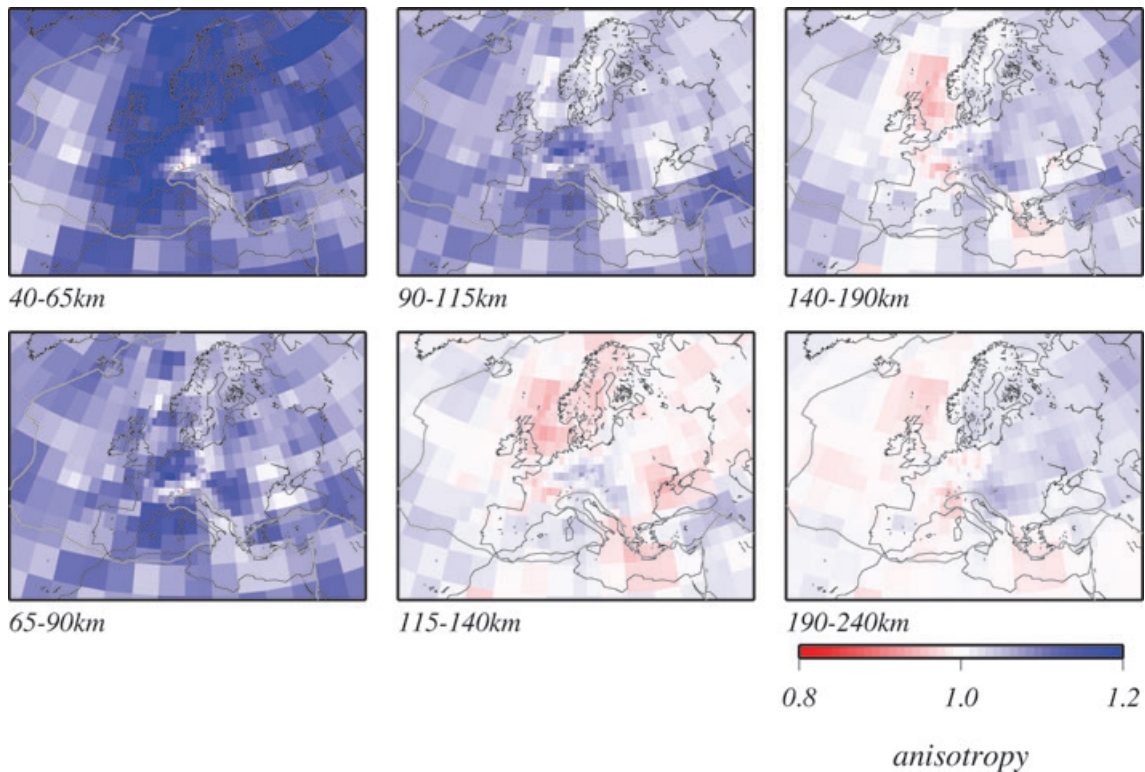
Figure 21. Detail of Fig. 20: ξ from model FMADVOXEU in Europe. The depth range of each layer is given in the lower left-hand corner of each panel.

at the Massif Central, presumably as a consequence of local upwelling (115 km downwards), and quite strongly between Great Britain and Scandinavia (115–240 km) and in the north of the Black Sea (115–140 km), which might be caused by downwelling at the edge of the craton. Like Chang *et al.* (2010) we find $v_{SH} > v_{SV}$ at shallow depths along the Hellenic Arc and $v_{SV} > v_{SH}$ from 115 km depth downwards (150 km in their study). This effect might be associated to frozen-in fossil anisotropy with faster v_{SH} in the subducting oceanic lithosphere. This anisotropy seems to be preserved in the slab which causes the fast axes to be aligned quasi-horizontally in regions with slight dip, resulting in high v_{SH}, and quasi-vertically in regions with steeper dip resulting in higher v_{SV} (Chang *et al.* 2010, and references therein). This effect might also cause $v_{SH} > v_{SV}$ for the Appenines associated to gentle dipping (Chang *et al.* 2010). Further $v_{SH} > v_{SV}$ for the Dinarides and the Tyrrhenian is consistent with Chang *et al.* (2010). While Chang *et al.* (2010) also observe $v_{SV} > v_{SH}$ for the East European platform we observe $v_{SH} > v_{SV}$ for most layers. The robustness of our observations of radial anisotropy in terms of mantle flow is analysed in Schaefer *et al.* (2011).

6 CONCLUSIONS AND OUTLOOK

We successfully applied the adaptive grid approach to surface wave tomography. We are now able to automatically adapt local cell size to local resolution capability estimated from the local ray density and azimuthal coverage rather than the formal resolution matrix. Resulting models document local model resolution capability directly in the model, thanks to the local cell size. Furthermore the adaptive grid provides an overview about the quality of data coverage and allows to identify regions for which additional data should be collected. Horizontal resolution in these models is optimized for the used wavelengths with a minimum resolvable horizontal cell size of around 50 km × 50 km. For smaller cells data with higher frequency would have to be added. For the moment, surface wave tomography is still limited by the amount and number of data and the prime aim should be to increase data coverage to be able to parametrize larger areas with the smallest cells.

Vertical resolution depends on the number of different wavelengths. To check the limits of resolution we overparametrized in vertical direction and conducted sensitivity tests. We can observe that the data set can resolve well the three top mantle layers with 25 km thickness. Below, layer thickness has to be increased to 50–100 km thickness. Fine vertical resolution also depends on data quality for each wavelength data set, layer thickness must be defined for each area specifically.

We have redefined the voxel roughness damping operator in the adaptive-grid domain. The regularization parameter is selected via the L-curve criterion as implemented by Boschi (2006). We have verified that our phase-velocity models correlate very well with geological features. In the examples considered here, the model we pick based on the L-curve criterion achieves the maximum correlation with known independent geological features. This suggests that our method to pick a model is reasonable.

As a result of our study we propose here a new 3-D anisotropic model of the European upper mantle (FMADVOXEU) determined from a joint inversion of various phases. Compared to previous surface wave tomography studies (e.g. Boschi *et al.* 2009) we are increasing the reliability of our models by making sure that horizontal model parametrization is optimal. This results in higher resolution of the lithosphere and its anisotropy heterogeneities. In most areas our new model shows similar large scale features compared to the results of European tomography reviewed by Boschi *et al.* (2009). Yet, some possibly important improvements do emerge: (i) The better correlation with cratonic lithosphere which allows to detect depth variations like the cratonic keel under southern Finland, not imaged by previous global surface wave studies; (ii) an apparently more robust regional map of radial anisotropy, whose geodynamic significance we are evaluating in an independent study (Schaefer *et al.* 2011). In a future step, we will augment our database with measurements of surface wave overtones. Further we will adopt an optimized layering according to sensitivity tests. This will allow us to resolve the entire upper mantle down to the transition zone. We are also combining our results with dynamic modelling and mineral physics observations, to explain our anisotropy maps in terms of mantle flow. Improvement of our model can be achieved as soon as a better crustal model is available (e.g. Molinari & Morelli 2011) by improving crustal corrections. Crustal corrections are especially important to constrain anisotropy (e.g. Ferreira *et al.* 2010; Panning *et al.* 2010). Alternatively, short/intermediate-period surface wave measurements from ambient noise tomography might be used to jointly invert for crust and upper-mantle structure. The future growth of computational infrastructure will provide us with larger RAM capacities, allowing us to extend the region of high resolution, and to invert a continuously growing database of global and regional seismic observations.

ACKNOWLEDGMENTS

REFERENCES

Anderson, D.L. & Dziewonski, A.M., 1982. Upper Mantle Anisotropy - Evidence from Free Oscillations, *Geophys. J. R. astr. Soc.,* **69**(2), 383–404.

Artemieva, I.M., 2006. Global 1 degrees x 1 degrees thermal model TC1 for the continental lithosphere: implications for lithosphere secular evolution, *Tectonophysics,* **416**(1–4), 245–277.

Baer, M., Zweifel, P. & Giardini, D., 2000. The Swiss digital seismic network (SDSNet), *Orfeus Newsl.,* **2**(2), 1–5.

Bijwaard, H., Spakman, W. & Engdahl, E.R., 1998. Closing the gap between regional and global travel time tomography, *J. geophys. Res.,* **103**(B12), 30 055–30 078.

Bodin, T. & Sambridge, M., 2009. Seismic tomography with the reversible jump algorithm, *Geophys. J. Int.,* **178**(3), 1411–1436.

Boschi, L., 2006. Global multiresolution models of surface wave propagation: comparing equivalently regularized Born and ray theoretical solutions, *Geophys. J. Int.,* **167**(1), 238–252.

Boschi, L. & Dziewonski, A.M., 1999. High- and low-resolution images of the Earth's mantle: implications of different approaches to tomographic modeling, *J. geophys. Res.,* **104**(B11), 25 567–25 594.

Boschi, L. & Ekström, R., 2002. New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies, *J. geophys. Res.,* **107**(B4), doi:10.1029/2000JB0059.

Boschi, L., Ekström, G. & Kustowski, B., 2004. Multiple resolution surface wave tomography: the Mediterranean basin, *Geophys. J. Int.,* **157**(1), 293–304.

Boschi, L., Fry, B., Ekström, G. & Giardini, D., 2009. The European upper mantle as seen by surface waves, *Surv. Geophys.,* **30**(4-5), 463–501.

Chang, S.J., van der Lee, S., Matzel, E. & Bedle, H., 2010. Radial anisotropy along the Tethyan margin, *Geophys. J. Int.,* **182**(2), 1013–1024.

Chiao, L. & Kuo, B., 2001. Multiscale seismic tomography, *Geophys. J. Int.,* **145**(2), 517–527.

Curtis, A. & Snieder, R., 1997. Reconditioning inverse problems using the genetic algorithm and revised parameterization, *Geophysics,* **62**(5), 1524–1532.

Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology,* Princeton University Press, Princeton, NJ.

Debayle, E. & Sambridge, M., 2004. Inversion of massive surface wave data sets: model construction and resolution assessment, *J. geophys. Res.,* **109**(B2), doi:10.1029/2003JB002652.

Debayle, E., Kennett, B. & Priestley, K., 2005. Global azimuthal seismic anisotropy and the unique plate-motion deformation of Australia, *Nature,* **433**, 509–512.

Dziewonski, A. & Anderson, D., 1981. Preliminary reference earth model, *Phys. Earth planet. Inter.,* **25**, 297–356.

Eberhart-Phillips, D., 1986. 3-Dimensional velocity structure in Northern California coast ranges from inversion of local earthquake arrival times, *Bull. seism. Soc. Am.,* **76**(4), 1025–1052.

Ekström, G. & Dziewonski, A.M., 1998. The unique anisotropy of the Pacific upper mantle, *Nature,* **394**(6689), 168–172.

Ekström, G., Tromp, J. & Larson, E.W.F., 1997. Measurements and global models of surface wave propagation, *J. geophys. Res.,* **102**(B4), 8137–8157.

Faccenna, C. & Becker, T.W., 2010. Shaping mobile belts by small-scale convection, *Nature,* **465**(7298), 602–605.

Ferreira, A.M.G., Woodhouse, J.H., Visser, K. & Trampert, J., 2010. On the robustness of global radially anisotropic surface wave tomography, *J. geophys. Res.,* **115**, doi:10.1029/2009JB006716.

Forte, A.M., Woodward, R.L. & Dziewonski, A.M., 1994. Joint inversions of seismic and geodynamic data for models of 3-dimensional mantle heterogeneity, *J. geophys. Res.,* **99**(B11), 21 857–21 877.

Fry, B., Boschi, L., Ekström, G. & Giardini, D., 2008. Europe-Mediterranean tomography: high correlation between new seismic data and independent geophysical observables, *Geophys. Res. Lett.,* **35**(4), doi:10.1029/2007GL031519.

Grad, M., Tiira, T. & Grp, E.W., 2009. The Moho depth map of the european plate, *Geophys. J. Int.,* **176**(1), 279–292.

Hansen, P.C., 1992. Analysis of Discrete Ill-Posed Problems by Means of the L-Curve, *Siam Review,* **34**(4), 561–580.

Henger, M., Berckhemer, H. & Seidl, D., 2002. The history of the development of the German Regional Seismic Network, in *Ten Years of German Regional Seismic Network (GRSN),* Report 25 of the Senate Commission for Geoscience (DFG), pp. 1–8, ed. Korn, M., Wiley-VCH, Weinheim.

Inoue, H., Fukao, Y., Tanabe, K. & Ogata, Y., 1990. Whaole mantle P-wave travel time tomography, *Phys. Earth planet. Inter.,* **59**(4), 294–328.

Jordan, T.H., 1978. Procedure for estimating lateral variations from low-frequency eigen-spectra data, *Geophys. J. R. astr. Soc.,* **52**(3), 441–455.

Khan, A., Boschi, L. & Connolly, J.A.D., 2011. Mapping the Earth's thermochemical and anisotropic structure using global surface wave data, *J. geophys. Res.,* **116**, doi:10.1029/2010JB007828.

Kissling, E., 1988. Geotomography with local earthquake data, *Rev. Geophys.,* **26**(4), 659–698.

Kissling, E., Husen, S. & Haslinger, F., 2001. Model parametrization in seismic tomography: a choice of consequence for the solution quality, *Phys. Earth planet. Inter.,* **123**(2-4), 89–101.

Kustowski, B., Ekström, G. & Dziewonski, A.M., 2008. Anisotropic shear-wave velocity structure of the Earth's mantle: a global model, *J. geophys. Res.,* **113**(B6), doi:10.1029/2007JB005169.

Lebedev, S. & van der Hilst, R.D., 2008. Global upper-mantle tomography with the automated multimode inversion of surface and S-wave forms, *Geophys. J. Int.,* **173**, 505–518.

Long, M.D. & Becker, T.W., 2010. Mantle dynamics and seismic anisotropy, *Earth planet. Sci. Lett.,* **297**(3-4), 341–354.

Loris, I., Nolet, G., Daubechies, I. & Dahlen, F., 2007. Tomographic inversion using l_1-norm regularization of wavelet coefficients, *Geophys. J. Int.,* **170**(1), 359–370, doi:10.1111/j.1365-246X.2007.03409.x.

Marone, F. & Romanowicz, B., 2007. Non-linear crustal corrections in high-resolution regional waveform seismic tomography, *Geophys. J. Int.,* **170**(1), 460–467.

Molinari, I. & Morelli, A., 2011. EPcrust: a reference crustal model for the european plate, *Geophys. J. Int.,* **185**, 352–364, doi:10.1111/j.1365-246X.2011.04940.x.

Montagner, J.P. & Tanimoto, T., 1991. Global upper mantle tomography of seismic velocities and anisotropies, *J. geophys. Res.,* **96**(B12), 20 337–20 351.

Mooney, W.D., Laske, G. & Masters, T.G., 1998. CRUST 5.1: a global crustal model at 5 degrees x 5 degrees, *J. geophys. Res.,* **103**(B1), 727–747.

Nataf, H.C., Nakanishi, I. & Anderson, D.L., 1984. Anisotropy and shear-velocity heterogeneities in the upper mantle, *Geophys. Res. Let.,* **11**(2), 109–112.

Panning, M.P., Lekic, V. & Romanowicz, B.A., 2010. Importance of crustal corrections in the development of a new global model of radial anisotropy, *J. geophys. Res.,* **115**, doi:10.1029/2010JB007520.

Peter, D., Boschi, L., Deschamps, F., Fry, B., Ekström, G. & Giardini, D., 2008. A new finite-frequency shear-velocity model of the European-Mediterranean region, *Geophys. Res. Lett.,* **35**(16), doi:10.1029/2008GL034769.

Pilidou, S., Priestley, K., Gudmundsson, O. & Debayle, E., 2004. Upper mantle S-wave speed heterogeneity and anisotropy beneath the North Atlantic from regional surface wave tomography: the Iceland and Azores plumes, *Geophys. J. Int.,* **159**(3), 1057–1076.

Piromallo, C. & Morelli, A., 2003. P wave tomography of the mantle under the Alpine-Mediterranean area, *J. geophys. Res.,* **108**(B2), doi:10.1029/2002JB001757.

Sambridge, M. & Faletic, R., 2003. Adaptive whole Earth tomography, *Geochem. Geophys. Geosyst.,* **4**, doi:10.1029/2001GC000213.

Sambridge, M. & Gudmundsson, O., 1998. Tomographic systems of equations with irregular cells, *J. geophys. Res.,* **103**(B1), 773–781.

Sambridge, M., Braun, J. & Mcqueen, H., 1995. Geophysical parametrization and interpolation of irregular data using natural neighbors, *Geophys. J. Int.,* **122**(3), 837–857.

Sandoval, S., Kissling, E., Ansorge, J. & Work, S.S.T., 2004. High-resolution body wave tomography beneath the SVEKALAPKO array – II. Anomalous upper mantle structure beneath the central Baltic Shield, *Geophys. J. Int.,* **157**(1), 200–214.

Schaefer, J.F., Boschi, L., Becker, T. & Kissling, E., 2011. Radial anisotropy in the European mantle: tomographic studies explored in terms of mantle flow, *Geophys. Res. Lett.,* submitted.

Schivardi, R. & Morelli, A., 2009. Surface wave tomography in the European and Mediterranean region, *Geophys. J. Int.,* **177**(3), 1050–1066.

Schmid, C., van der Lee, S., VanDecar, J.C., Engdahl, E.R. & Giardini, D., 2008. Three-dimensional S velocity of the mantle in the Africa-Eurasia plate boundary region from phase arrival times and regional waveforms, *J. geophys. Res.,* **113**(B3), doi:10.1029/2005JB004193.

Soldati, G., Boschi, L. & Forte, A.M., 2011. Geodynamically self-consistent tomography of core-mantle boundary and lowermost mantle, *Geophys. J. Int.,* submitted.

Takeuchi, M. & Saito, M., 1972. Seismic surface waves, in *Seismology: Surface Waves and Earth Oscillations,* Methods in Computational Physics Vol. 11, pp. 217–295, eds Alder, B., Fernbach, S. & Rotenberg, M., Academic Press, New York, NY.

Tesauro, M., Kaban, M.K. & Cloetingh, S., 2008. EuCRUST-07: a new reference model for the European crust, *Geophys. Res. Lett.,* **35**(5), doi:10.1029/2007GL032244.

Tikhotsky, S. & Achauer, U., 2008. Inversion of controlled-source seismic tomography and gravity data with the self-adaptive wavelet parametrization of velocities and interfaces, *Geophys. J. Int.,* **172**(2), 619–630.

Trefethen, L. & Bau, D.I., 1997. *Numerical Linear Algebra,* Society for Industrial and Applied Mathematics, Philadelphia, PA.

Tromp, J. & Dahlen, F.A., 1992. Variational-principles for surface-wave propagation on a laterally heterogeneous Earth—II. Frequency-domain JWKB theory, *Geophys. J. Int.,* **109**(3), 599–619.

van der Lee, S. *et al.*, 2001. Eurasia-Africa plate boundary region yields new seismographic data, *EOS, Trans. Am. geophys. Un.,* **82**(51), 637.

Visser, K., Trampert, J. & Kennett, B.L.N., 2008a. Global anisotropic phase velocity maps for higher mode Love and Rayleigh waves, *Geophys. J Int.,* **172**(3), 1016–1032.

Visser, K., Trampert, J., Lebedev, S. & Kennett, B.L.N., 2008b. Probability of radial anisotropy in the deep mantle, *Earth planet. Sci. Lett.,* **270**(3-4), 241–250.

Wang, Z., Tromp, J. & Ekstrom, G., 1998. Global and regional surface-wave inversions: a spherical-spline parameterization, *Geophys. Res. Lett.,* **25**(2), 207–210.

Wessel, P. & Smith, W. H.F., 1991. Free software helps map and display data, *EOS, Trans. Am. geophys. Un.,* **72**(441).

Woodhouse, J.H. & Dziewonski, A.M., 1984. Mapping the upper mantle—3-dimensional modeling of earth structure by inversion of seismic waveforms, *J. geophys. Res.,* **89**(Nb7), 5953–5986.