

Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle

Frederik J. Simons,^{1,*} Rob D. van der Hilst,¹ Jean-Paul Montagner² and Alet Zielhuis¹

¹Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

E-mail: fjsimons@alum.mit.edu

²Département de Sismologie, Institut de Physique du Globe, 75005 Paris, France

Accepted 2002 June 10. Received 2002 May 8; in original form 2001 January 26

SUMMARY

We present an azimuthally anisotropic 3-D shear-wave speed model of the Australian upper mantle obtained from the dispersion of fundamental and higher modes of Rayleigh waves. We compare two tomographic techniques to map path-average earth models into a 3-D model for heterogeneity and azimuthal anisotropy. Method I uses a rectangular surface cell parametrization and depth basis functions that represent independently constrained estimates of radial earth structure. It performs an iterative inversion with norm damping and gradient regularization. Method II uses a direct inversion of individual depth layers constrained by Bayesian assumptions about the model covariance. We recall that Bayesian inversions and discrete regularization approaches are theoretically equivalent, and with a synthetic example we show that they can give similar results. The model we present here uses the discrete regularized inversion of independent path constraints of Method I, on an equal-area grid. With the exception of westernmost Australia, we can retrieve structure on length scales of about 250 km laterally and 50 km in the radial direction, to within 0.8 per cent for the velocity, 20 per cent for the anisotropic magnitude and 20° for its direction. On length scales of 1000 km and longer, down to about 200 km, there is a good correlation between velocity heterogeneity and geologic age. At shorter length scales and at depths below 200 km, however, this relationship breaks down. The observed magnitude and direction of maximum anisotropy do not, in general, appear to be correlated to surface geology. The pattern of anisotropy appears to be rather complex in the upper 150 km, whereas a smoother pattern of fast axes is obtained at larger depth. If some of the deeper directions of anisotropy are aligned with the approximately N–S direction of absolute plate motion, this correspondence is not everywhere obvious, despite the fast (7 cm yr⁻¹) northward motion of the Australian plate. More research is needed to interpret our observations in terms of continental deformation. Predictions of *SKS* splitting times and directions, an integrated measure of anisotropy, are poorly matched by observations of shear-wave birefringence.

Key words: Australia, azimuthal anisotropy, heterogeneity, inversion, surface-waves, tomography.

1 INTRODUCTION

The Earth is heterogeneous and anisotropic on many length scales. Polarization anisotropy (modelled by transverse isotropy with a vertical symmetry axis) affects fundamental and higher-mode surface-wave propagation. Body- and surface-wave velocities are dependent on propagation azimuth. Shear-wave birefringence reflects transverse isotropy with a horizontal symmetry axis. Since anisotropy

requires both the presence of anisotropic crystals and their (strain-induced) preferred orientation on seismic length scales, anisotropy measurements contain information about the mineralogy and dynamics of the mantle (see, e.g. Silver 1996; Plomerová *et al.* 1998; Savage 1999; Kendall 2000).

Knowledge of continental velocity heterogeneity and anisotropy may answer a number of outstanding questions. These include the very definition of the base of the continental lithosphere, the relation between the depth extent of the lithosphere and the age of the overlying crust, and the deformation of the continental lithospheric mantle and the underlying asthenosphere.

*Now at: Princeton University, 321B Guyot Hall, Princeton, NJ 08544

It has been argued that continental seismic anisotropy is best explained by a combination of ‘frozen’ (subcrustal) anisotropy, developed during major orogenic events, underlain by an actively deforming asthenosphere which displays anisotropy generated by the subhorizontal alignment of olivine crystals by present-day mantle flow (Silver 1996). However, considerable ambiguity exists regarding the amount and location of anisotropy inferred from body waves.

Surface-wave waveform tomography allows the delineation of lateral heterogeneities in the upper mantle with better radial resolution than body-wave traveltimes tomography. A variety of methods has been developed to extract the variation of wave speed with depth from surface waves, (e.g. Nolet *et al.* 1986; Cara & L  v  que 1987), and a number of tomographic methods to invert for aspherical structure exist (e.g. Montagner 1986; Nolet 1990). Inversions of surface waves for heterogeneity have traditionally been restricted to fundamental modes but more recent models have incorporated higher-mode measurements, thereby enhancing the resolution down to depths of about 400 km. Inversions for both heterogeneity and azimuthal anisotropy have been made using fundamental modes (e.g. Montagner & Jobert 1988; Nishimura & Forsyth 1989) and multimode waves (L  v  que *et al.* 1998; Debayle & Kennett 2000a).

Local measurements of the seismic structure of the Australian upper mantle have indicated the presence of a high-velocity lithospheric lid down to ~200 km in Northern Australia (e.g. Dey *et al.* 1993; Kennett *et al.* 1994), and to ~300 km in Central Australia (Gaherty & Jordan 1995; Gaherty *et al.* 1999). Seismic anisotropy has been proposed for the upper mantle under Central Australia down to 200–300 km (Gaherty *et al.* 1999), and between 210 and 410 km depth under Northern Australia (Tong *et al.* 1994). Girardin & Farra (1998) report evidence for two-layer anisotropy in southeast Australia. The conflicting results of different shear-wave birefringence studies (Clitheroe & van der Hilst 1998;   zalaybey & Chen 1999) indicate that the anisotropy of the Australian continental mantle is complex and not yet well understood.

The SKIPPY experiment (van der Hilst *et al.* 1994) has produced a unique broad-band data set of seismic waves propagating all across the Australian lithosphere, resulting in the path coverage used in the

present study (Fig. 1). Waveform inversions of vertical-component fundamental and higher-mode Rayleigh-wave data have yielded detailed information on the 3-D isotropic *S*-wave speed variations of the continent and the adjacent oceanic areas (Zielhuis & van der Hilst 1996; Simons *et al.* 1999; Debayle & Kennett 2000a).

The azimuthal anisotropy measured by Debayle (1999) and Debayle & Kennett (2000a) changes from a regime with fast-changing directions in the upper 150 km to smoother North-South patterns below 150 km, and an analysis of Love and Rayleigh waves by Debayle & Kennett (2000b) shows polarization anisotropy extending down to about 200–250 km.

In this paper, we present a new model for the Australian upper mantle that takes into account both wave-speed heterogeneity and azimuthal anisotropy. We quantify resolution and uncertainty by means of synthetic tests, trade-off estimates, and an exact calculation of the horizontal resolution matrix. While there is a general agreement between the long-wavelength structures imaged in our present model and the one by Debayle & Kennett (2000a) (hereafter: DK2000), there are substantial differences in heterogeneity and anisotropic structure at wavelengths smaller than about 500 km even in regions for which both groups claim good resolution.

Like DK2000, in this paper we attempt to invert for 3-D structure starting from path-average models. The validity of the approximation of independent propagation of the individual modes of the seismogram has been explored theoretically (Kennett & Nolet 1990; Kennett 1995; Meier & Malischewsky 2000) and been found to be subject to certain frequency limitations, to which we adhere. As the frequency bounds for uncoupled mode propagation presume a scale length of lateral heterogeneity, DK2000 conclude one should be wary of comparing differences in our and their model on lateral length scales shorter than 300 km. DK2000 report noticeable differences from great-circle propagation through their model at frequencies >25 mHz. In the present paper, the higher-mode window of the waveforms is fit up to 50 mHz where possible. Ideally, the capability of our models to explain the observed waveforms needs to be tested with 3-D wave propagation techniques (e.g. Komatitsch & Tromp 1999).

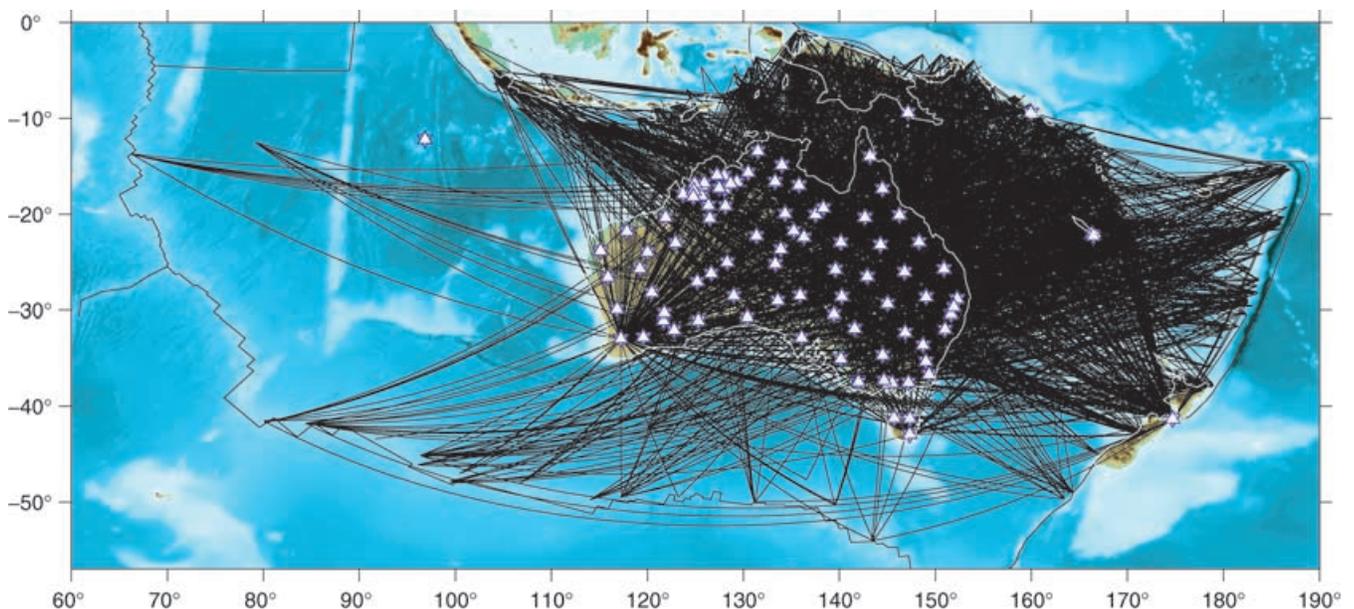


Figure 1. Path coverage. About 2250 vertical-component seismograms provided waveforms which were fitted to obtain the model presented in this paper. The stations belong to the SKIPPY, KIMBA, AGSO, IRIS and GEOSCOPE arrays. We used earthquake locations from Engdahl *et al.* (1998).

In addition to differences in the waveform inversion technique, the tomographic inversion method used to obtain the model presented in this paper differs from the one used by DK2000. They use a form of the regionalized inversion method due to Montagner (1986), while we use the discretely parametrized, regularized inversion method introduced by Zielhuis & Nolet (1994). We compare these tomographic techniques in theory and in their application to data.

2 INVERSION PARTITIONING AND DATA

Our method and the one of DK2000 both involve an inversion for aspherical earth structure that is partitioned into an initial waveform inversion that seeks to determine the average structure under the great-circle path between source and receiver and a subsequent tomographic inversion for 3-D Earth structure (which will be discussed in Section 4).

Our waveform inversion is due to Nolet *et al.* (1986). It produces a minimal set of independently constrained estimates of radial earth structure (Zielhuis 1992) that are used as data in the 3-D tomographic inversion. The independence of the linear constraints and the decorrelation of their errors is achieved by matrix diagonalization. This is the central distinguishing feature of Nolet's (1990) method. Furthermore, with this method the waveform analysis can be applied to separate analysis windows in the group velocity versus frequency space (see Simons *et al.* 1999, their Fig. 3), in order to separate fundamental from higher modes, where possible. The possibility of using two data windows prevents the highly energetic fundamental modes from dominating the waveform inversion. For their waveform inversion, DK2000 rely on the method by Cara & L ev eque (1987), which linearizes the waveform inversion by using the envelope of modal cross-correlation functions and fits the fundamental mode and the overtones simultaneously (Debayle 1999).

Both approaches rely on a path-average approximation and neglect mode coupling. As a result, some effects of complex source mechanisms and mantle heterogeneity are not adequately modelled. However, when applied within the frequency bounds derived for both fundamental and higher modes (Kennett & Nolet 1990; Kennett 1995) and for a large number of crossing paths, the path-average method leads to an acceptable approximation of the 3-D upper-mantle heterogeneity (Marquering & Snieder 1996).

For the present study we have expanded the data set of 1600 paths used by Simons *et al.* (1999) with additional data from the SKIPPY project and from permanent observations operated by IRIS, GEOSCOPE and AGSO, and we applied the waveform inversion to a total of ~ 2250 paths. In selecting the events we have avoided regions with complex source and mantle structure by limiting the data to events that originate mostly along or within the boundaries of the Indo-Australian plate (see Fig. 1).

Where data quality allows we fit the fundamental mode between 5–25 mHz (40–200 s period) or a narrower frequency band within this range, while the higher modes are fitted between 8–50 mHz (20–125 s) or less. With these frequency bounds the fundamental modes can provide sensitivity to structure down to at least 400 km, and the higher modes yield additional detail. For more details about the application to SKIPPY data, see Zielhuis & van der Hilst (1996).

At high frequencies, Rayleigh modes are sensitive to variations in crustal thickness and velocity. We use continental crustal thickness determinations from broad-band receiver functions at each of the SKIPPY and permanent network stations (Shibutani *et al.* 1996;

Clitheroe *et al.* 2000a,b). This is arguably the most accurate crustal model to date of the Australian continent.

3 AZIMUTHAL ANISOTROPY

Following Montagner & Nataf (1986), the perturbation of the phase speed of Rayleigh waves (δc_R) in a slightly anisotropic Earth can be modelled by a transversely isotropic perturbation (δL) and two combinations of elastic constants (G_c and G_s) modulated by, respectively, the cosine and sine of twice the azimuth of the ray (Ψ):

$$\delta c_R = \frac{\partial c_R}{\partial L} (\delta L + G_c \cos 2\Psi + G_s \sin 2\Psi). \quad (1)$$

In simplified index notation (C_{ij}) for the elastic tensor Γ_{ijkl} (Montagner 1996) the parameters of eq. (1) are:

$$L = (C_{44} + C_{55})/2, \quad (2)$$

$$G_c = (C_{55} - C_{44})/2, \quad (3)$$

$$G_s = C_{54}. \quad (4)$$

Like L ev eque *et al.* (1998) we define an effective shear-wave speed

$$\hat{\beta}_V = \left(\frac{L + G_c \cos 2\Psi + G_s \sin 2\Psi}{\rho} \right)^{\frac{1}{2}}. \quad (5)$$

If no azimuthal anisotropy is present or if it is effectively averaged out over all azimuths, $\rho \hat{\beta}_V^2 = L$ and $G_c = G_s = 0$ so that eq. (5) reduces to β_V , i.e. the wave speed of a vertically polarized shear-wave in a transversely isotropic medium with a vertical symmetry axis (Takeuchi & Saito 1972). Assuming $G_c/L \ll 1$ and $G_s/L \ll 1$ we can write:

$$\hat{\beta}_V = \beta_V + \frac{G_c}{2\rho\beta_V} \cos 2\Psi + \frac{G_s}{2\rho\beta_V} \sin 2\Psi. \quad (6)$$

In the tomographic inversion we will take as a starting point the particular radial velocity averages that result from the waveform inversion (see Section 2), and solve for heterogeneity (β_V) and the anisotropic parameters ($A_1 = G_c/(2\rho\beta_V)$ and $A_2 = G_s/(2\rho\beta_V)$). We calculate the anisotropic magnitude as $(G_c^2 + G_s^2)^{1/2}$ and the direction of maximum velocity as $\tan^{-1}(A_2/A_1)/2$.

The ubiquitous use of the parametrization represented by eqs (1) and (6) is based on the dominance of the $\partial c_R/\partial L$ partial for fundamental-mode surface-wave propagation (Montagner & Nataf 1986). Since $\partial c_R/\partial L$ is also dominant for higher modes, both for isotropic and realistically anisotropic media (to verify this, we computed sensitivity kernels for the model by Gaherty & Jordan 1995), a parametrization in terms of β_V , G_c and G_s is adequate when higher modes are included.

To fit seismograms, we used the $\partial c_R/\partial L$ partial derivative with an isotropic reference model. The resulting 1-D path model of β_V is an adequate representation of the Earth when no further azimuthally anisotropic information is available, because the sensitivity kernels of Rayleigh wave phase velocities to perturbations in S -velocity in isotropic and transversely isotropic earth models are very similar (Anderson & Dziewonski 1982), although a small error is probably made (L ev eque & Cara 1985).

4 THE TOMOGRAPHIC PROBLEM

4.1 Construction of a linear matrix problem

With the exception of a few computational details, our method of S -wave velocity tomography differs little from the procedure described by Zielhuis & Nolet (1994). The model vector \mathbf{m} containing 3-D shear-wave speed perturbations $\delta\beta(\mathbf{r})$ from a reference model $\beta(r)$ is related to the data vector \mathbf{d} of independent linear constraints by a sensitivity matrix \mathbf{G} . The data errors are uncorrelated by construction, and the sensitivity is weighted by the inverse of the data error. We seek the model vector \mathbf{m} that solves the linear equation $\mathbf{d} = \mathbf{G} \cdot \mathbf{m}$; we incorporate the effect of azimuthal anisotropy by modifying the shear-wave speed anomaly to include a transversely isotropic and an azimuthally anisotropic part:

$$\delta\beta(\mathbf{r}) = \delta\beta_V(\mathbf{r}) + A_1(\mathbf{r}) \cos 2\Psi(\mathbf{r}) + A_2(\mathbf{r}) \sin 2\Psi(\mathbf{r}), \quad (7)$$

i.e. by augmenting the sensitivity matrix \mathbf{G} to include the cosine and the sine of twice the azimuth Ψ of the ray in each surface cell. We try to explain the data with both heterogeneity and anisotropy. This leads to a threefold increase in the number of unknowns with respect

to isotropic inversions. In Section 5.5 we evaluate the resolution of heterogeneity and anisotropic parameters as well as the importance of the trade-off of between them.

4.2 Model parametrization

The model presented in this paper is based on a $2^\circ \times 2^\circ$ equal-area surface parametrization of the kind shown in Figs 2(a) and (b). An inversion on a $3^\circ \times 3^\circ$ grid showed essentially the same features, but an inversion for heterogeneity and anisotropy on a $1^\circ \times 1^\circ$ grid proved less stable. To further test the robustness of the structures retrieved, we performed inversions on a resolution-dependent grid (Fig. 2c). In order to determine the extent to which the Australian continental structure at varying depths bears a relationship to the geology observed at the surface, we also performed inversions on a pure-path (tectonic) grid (Fig. 2d).

Due to the uneven sampling density the resolution is not uniform across the model domain, but the grid spacing can be adapted to the locally expected resolution in order to reduce the number of free parameters. Our approach to adaptive grids is akin to that of Abers & Roecker (1991). We used a fine ($1^\circ \times 1^\circ$) Cartesian grid upon

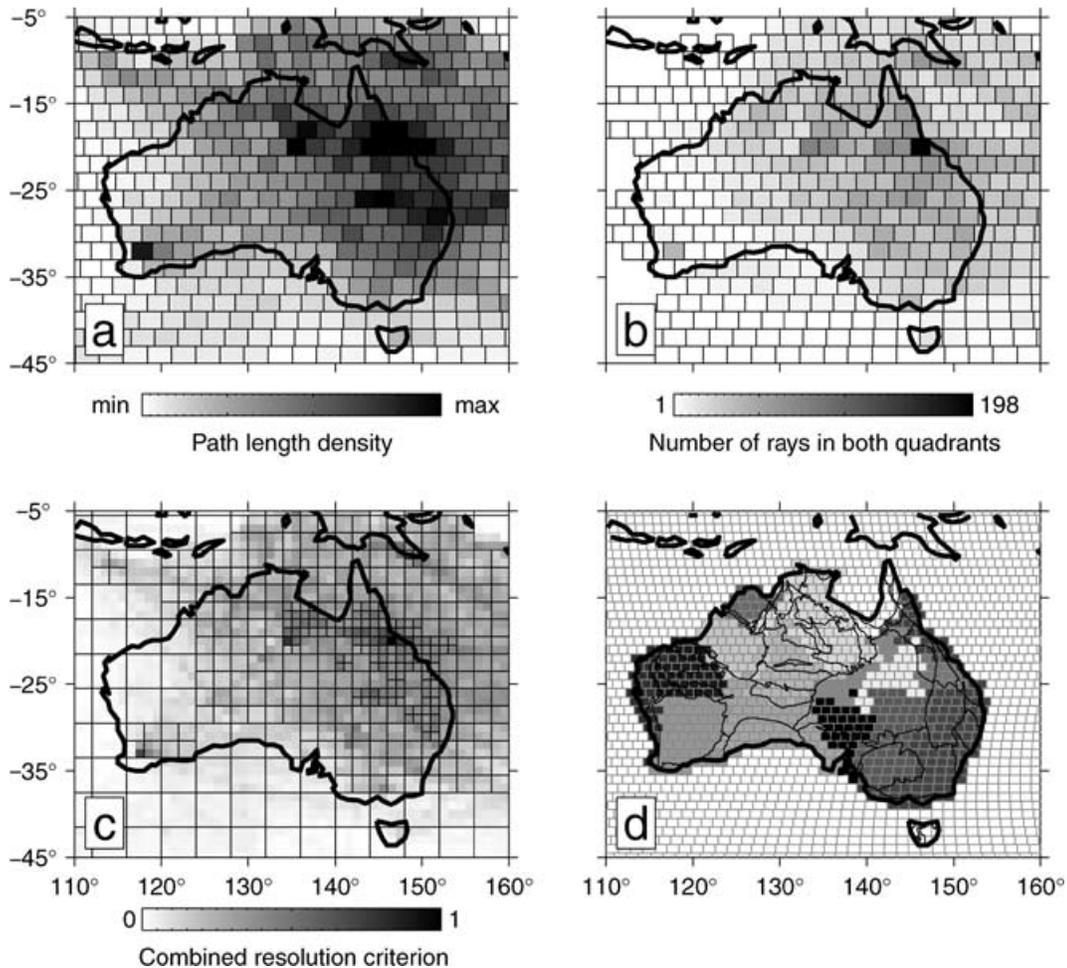


Figure 2. (a) Path density and (b) azimuthal coverage for the paths shown in Fig. 1, calculated on a $2^\circ \times 2^\circ$ equal-area grid. As a measure of azimuthal coverage, we took the number of pairs of rays in the cell coming into two different quadrants. (c) Design of a grid adapted to expected model resolution. Path density and azimuthal coverage calculated on a $1^\circ \times 1^\circ$ square grid were normalized and added to provide a measure of data quality. The cell size was adapted to allow a finer parametrization in areas where the resolution is expected to be better. (d) Tectonic subdivision. Based on an underlying equal-area grid, similarly shaded regions indicate approximate tectonic domains which are used as a single parameter in the pure-path inversion. Lines indicate the location of major sedimentary basins. For the definition of tectonic units, see Simons *et al.* (1999).

which we define a coordinate transform that attempts to equalize the sampling density by clustering cells together. Starting from a maximum cell size of $8^\circ \times 8^\circ$ we allow a further subdivision, into cells sized $4^\circ \times 4^\circ$, $2^\circ \times 2^\circ$ and $1^\circ \times 1^\circ$, as long as this brings the resulting mean of the sampling density over these new cells closer to the global mean. As a measure of resolution we use the sum of path density (see Fig. 2a) and azimuthal variability (Fig. 2b) in the surface cells. Although this is not necessarily an optimal *a priori* estimate of the spatial resolution of the model the similarity of the features of this measure with an exactly computed resolution matrix (see Section 5.2) is sufficient to warrant their use in the construction of an irregular grid. An example of an irregular grid that results from the clustering of the sampling density is plotted in Fig. 2(c).

Pure-path methods are used to invert for the average velocity and anisotropy structure of predefined tectonic regions (Fig. 2d) (Nishimura & Forsyth 1985). Similarly to the construction of a resolution-dependent irregular grid we achieve this tectonic parametrization by a coordinate transform, such that all surface cells belonging to a tectonic domain are treated as a single parameter in the inversion.

4.3 Regularization or *a priori* information?

The discrete surface parametrization of our method leads to classic regularization approaches of systems of linear equations (Section 4.3.1). On the other hand, the continuous regionalization approach adopted by DK2000 draws solutions from a class of functions defined by their *a priori* covariance structure (Section 4.3.2).

4.3.1 Discrete regularization

We construct a penalty function based on a weighted Euclidean norm of the solution error and a weighted norm in model parameter space as follows:

$$\Phi = (\mathbf{d} - \mathbf{G} \cdot \mathbf{m}) \cdot \mathbf{A}^{-1} \cdot (\mathbf{d} - \mathbf{G} \cdot \mathbf{m}) + \mathbf{m} \cdot \mathbf{B}^{-1} \cdot \mathbf{m}. \quad (8)$$

\mathbf{A}^{-1} and \mathbf{B}^{-1} are weighting matrices. Minimization of eq. (8) in a least-squares sense is equivalent to solving the following modified set of matrix equations:

$$\begin{bmatrix} \mathbf{A}^{-1/2} \cdot \mathbf{G} \\ \mathbf{B}^{-1/2} \end{bmatrix} \cdot \mathbf{m} = \begin{bmatrix} \mathbf{A}^{-1/2} \cdot \mathbf{d} \\ 0 \end{bmatrix}. \quad (9)$$

The generalized inverse,

$$\mathbf{m} = (\mathbf{B}^{-1} + \mathbf{G}^T \cdot \mathbf{A}^{-1} \cdot \mathbf{G})^{-1} \cdot \mathbf{G}^T \cdot \mathbf{A}^{-1} \cdot \mathbf{d}, \quad (10)$$

can be approximated by the iterative LSQR algorithm (Paige & Saunders 1982).

The optimal weighting matrix \mathbf{A}^{-1} is given by the inverse of the data covariance matrix \mathbf{C}_d^{-1} , so that the solution has minimum variance. A covariance matrix, \mathbf{C}_d is positive definite (hence invertible) and symmetric; in our implementation it is also diagonal because the individual parameters resulting from the waveform fitting procedure are independent by construction (Nolet 1990; Zielhuis & Nolet 1994).

If we choose $\mathbf{B}^{-1} = \mathbf{I}$, the identity matrix, we are imposing a minimum norm constraint on the solution by biasing the solution to zero in a least-squares sense. This is appropriate if \mathbf{m} represents perturbations to a reference model which is sufficiently close to the real solution so that linearity is preserved. Most inversion algorithms for the generalized inverse minimize the model norm implicitly (van der Sluis & van der Vorst 1987). In particular, the LSQR algorithm

provides an intrinsic damping of the norm $\mathbf{m} \cdot \mathbf{m}$ when the iterations are truncated at some small finite value (Nolet 1985; van der Sluis & van der Vorst 1987).

Alternatively, we may propose $\mathbf{B}^{-1} = \mathbf{D}_1^2$, the square of the first finite difference matrix given by

$$\mathbf{D}_1 = \begin{pmatrix} \dots & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & \dots \end{pmatrix}. \quad (11)$$

The effect of this “roughening” operator is to constrain the solution to be locally smooth (Nolet 1987). Eq. (11) is the Toeplitz matrix of a high-pass filter in model space (Strang & Nguyen 1997). Hence, we are imposing that the high-passed model vector be minimal, which, after inversion, results in a smooth model.

In our inversion, we combine both approaches and minimize the penalty function

$$\Phi = (\mathbf{d} - \mathbf{G} \cdot \mathbf{m}) \cdot \mathbf{C}_d^{-1} \cdot (\mathbf{d} - \mathbf{G} \cdot \mathbf{m}) + \alpha \mathbf{m} \cdot \mathbf{I} \cdot \mathbf{m} + \beta \mathbf{m} \cdot \mathbf{D}_1^2 \cdot \mathbf{m}, \quad (12)$$

where α and β are weights which can be adjusted arbitrarily. The solution

$$\mathbf{m} = (\alpha \mathbf{I} + \beta \mathbf{D}_1^2 + \mathbf{G}^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{G})^{-1} \cdot \mathbf{G}^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{d} \quad (13)$$

is calculated by LSQR.

4.3.2 Continuous regionalization

In the “Bayesian” approach (Tarantola & Valette 1982a,b; Montagner 1986) the model parameters \mathbf{m} are distributed according to a Gaussian probability density function

$$\rho(\mathbf{m}) \propto \exp\left(-\frac{1}{2} \mathbf{m} \cdot \mathbf{C}_m^{-1} \cdot \mathbf{m}\right). \quad (14)$$

The solution is obtained by maximizing the joint distribution of data and model parameters, and is given by

$$\mathbf{m} = \mathbf{C}_m \cdot \mathbf{G}^T \cdot (\mathbf{C}_d + \mathbf{G} \cdot \mathbf{C}_m \cdot \mathbf{G}^T)^{-1} \cdot \mathbf{d}. \quad (15)$$

The choice of the *a priori* model covariance matrix is important. DK2000 use

$$\mathbf{C}_m(\mathbf{r}_1, \mathbf{r}_2) = \sigma^2 \exp\left(-\frac{|\mathbf{r}_1, \mathbf{r}_2|^2}{2L^2}\right), \quad (16)$$

where L denotes a correlation length, the determination of which is, like the level of regularization, essentially at the discretion of the seismologist. It is possible, perhaps even sensible, to use different correlation lengths for velocity heterogeneity and anisotropy. In addition, the *a priori* standard deviations σ can be chosen independently.

Using the trivial matrix identity

$$\mathbf{G}^T \cdot \mathbf{C}_d^{-1} \cdot (\mathbf{G} \cdot \mathbf{C}_m \cdot \mathbf{G}^T + \mathbf{C}_d) = (\mathbf{G}^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{G} + \mathbf{C}_m^{-1}) \cdot \mathbf{C}_m \cdot \mathbf{G}^T, \quad (17)$$

we can rewrite eq. (15) as

$$\mathbf{m} = (\mathbf{C}_m^{-1} + \mathbf{G}^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{G})^{-1} \cdot \mathbf{G}^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{d}. \quad (18)$$

4.3.3 Equivalence of regularization and Bayesian approaches

Regularization of a discretely parametrized model and stochastic extensions using an *a priori* model covariance structure have been presented as distinct alternatives (see, e.g. Scales & Snieder 1997) but the approaches are equivalent in many ways (e.g. Ho-Liu *et al.* 1989).

Damping criteria to stabilize inverse solutions induce a particular form of a model covariance matrix. Comparison of eqs (18) and (13) reveals that choosing norm and gradient regularization amounts to identifying

$$\mathbf{C}_m^{-1} = \alpha \mathbf{I} + \beta \mathbf{D}_1^2 \quad (19)$$

as a subjective approximation to the inverse of the *a priori* model covariance matrix. Yanovskaya & Ditmar (1990) show that for a Gaussian covariance function of the form of eq. (16), the explicit equivalence holds:

$$\mathbf{m} \cdot \mathbf{C}_m^{-1} \cdot \mathbf{m} = \frac{(\sigma L)^{-2}}{2\pi} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{L^2}{2} \right)^n \nabla^n \mathbf{m} \cdot \nabla^n \mathbf{m}. \quad (20)$$

With this choice of covariance function one obtains smooth, infinitely differentiable functions. Eq. (20) corresponds to a model norm weighted by an infinite number of finite difference matrices, starting with the zeroth difference, i.e. the traditional model norm. Neglecting the higher-order smoothing terms makes this approach equivalent to eq. (19).

To illustrate this equivalence we conducted a synthetic experiment to compare the discrete regularization approach with the continuous regionalization for one particular depth. For a fabricated path coverage (Fig. 3a), (noiseless) data were synthesized from a model of shear-wave speed heterogeneity (Fig. 3b), azimuthal anisotropy (Fig. 3c), or both. Fig. 4 shows the inversion results obtained with the discrete method (left column) and the continuous method (right column). Figs 4(a) and (b) show the result for an input model with only heterogeneity (i.e. Fig. 3b). Figs 4(c) and (d) show the solution for a model with only anisotropy (i.e. Fig. 3c), and Figs 4(e) and (f) give the results for a model which has both heterogeneity and anisotropy. This test demonstrates that a set of inversion parameters (correlation lengths, *a priori* standard deviations and regularization parameters) can be found with which both methods yield very similar results. The exact values are unimportant given the synthetic nature of this example. Differences in amplitude recovery are due to the different implicit weighting of the data constraints versus the *a priori* information in both methods. The differences in smoothness are due to the chosen horizontal correlation length (chosen to maximize variance reduction; in this example we chose a correlation length of 175 km for both the velocity heterogeneity and the anisotropy) and the implicit correlation length provided by the damping.

4.4 Differences between Method I and Method II

While conceptually similar there are, however, theoretical and practical differences that set our method (Method I) apart from the method used by DK2000 (Method II).

In the previous section we have shown the theoretical equivalence of the Bayesian inversion approach with the regularization method, and that inversion parameters can be found so that both methods produce similar results (Figs 3 and 4). One should realize, however, that we could find such parameters only because the input model was known. In inversions for unknown earth structure this tuning is not possible. The challenge is to find *a priori* horizontal correla-

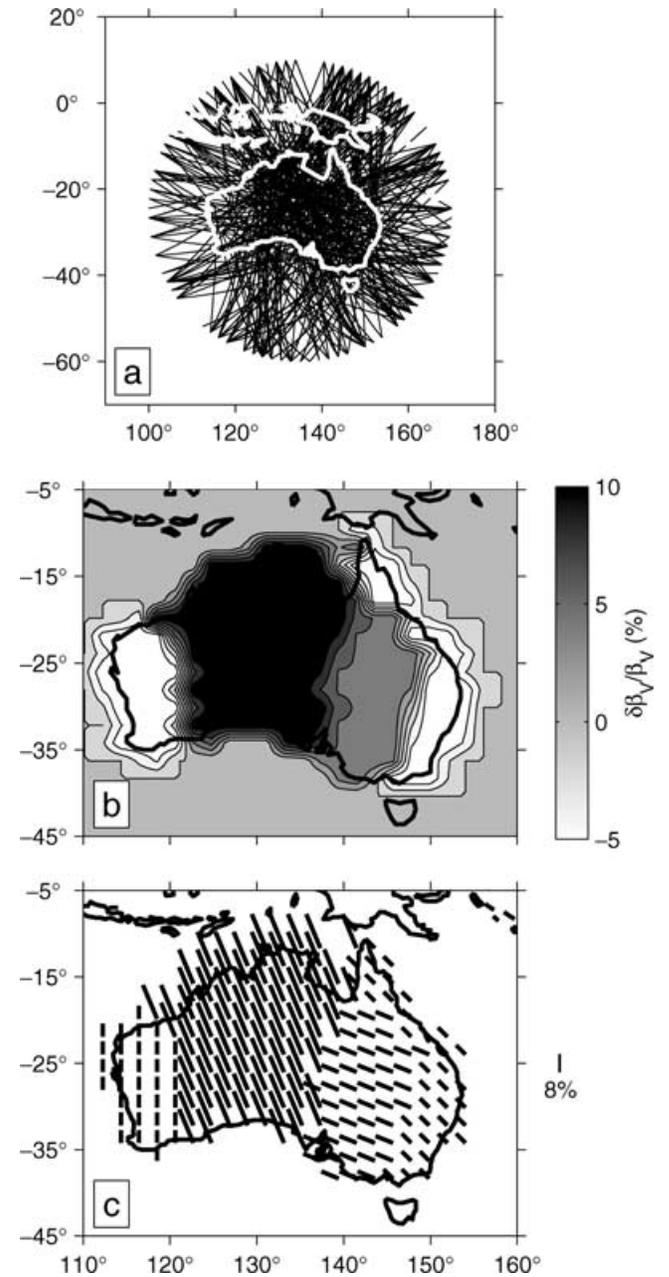


Figure 3. Synthetic experiment. (a) Distribution of 450 events and stations randomly distributed on a circle around Australia. (b) Input velocity heterogeneity, in per cent. (c) Input azimuthal anisotropy, direction of maximum velocity.

tion lengths for both the velocity heterogeneity and the anisotropic parameters, and *a priori* standard deviations of the anomalies (σ_β for $\delta\beta_V$ and σ_A for A_1 and A_2). The correlation length affects the smoothness of the resulting model and the *a priori* uncertainties regulate the amplitude. Moreover, the ratio of σ_β and σ_A controls the trade-off between these quantities. The choice of σ does not affect the location but only the amplitude of the anomalies. Various authors using this technique have made different choices. L ev eque *et al.* (1998), for example, choose $\sigma_\beta = 100 \text{ m s}^{-1}$ and $\sigma_A = 5 \text{ m s}^{-1}$. Griot *et al.* (1998) have $\sigma_\beta = 200 \text{ m s}^{-1}$ and $\sigma_A = 200 \text{ m s}^{-1}$, and Silveira *et al.* (1998) take $\sigma_\beta = 200 \text{ m s}^{-1}$ and $\sigma_A = 130 \text{ m s}^{-1}$. DK2000 take $\sigma_\beta = 50 \text{ m s}^{-1}$ and $\sigma_A = 3 \text{ m s}^{-1}$.

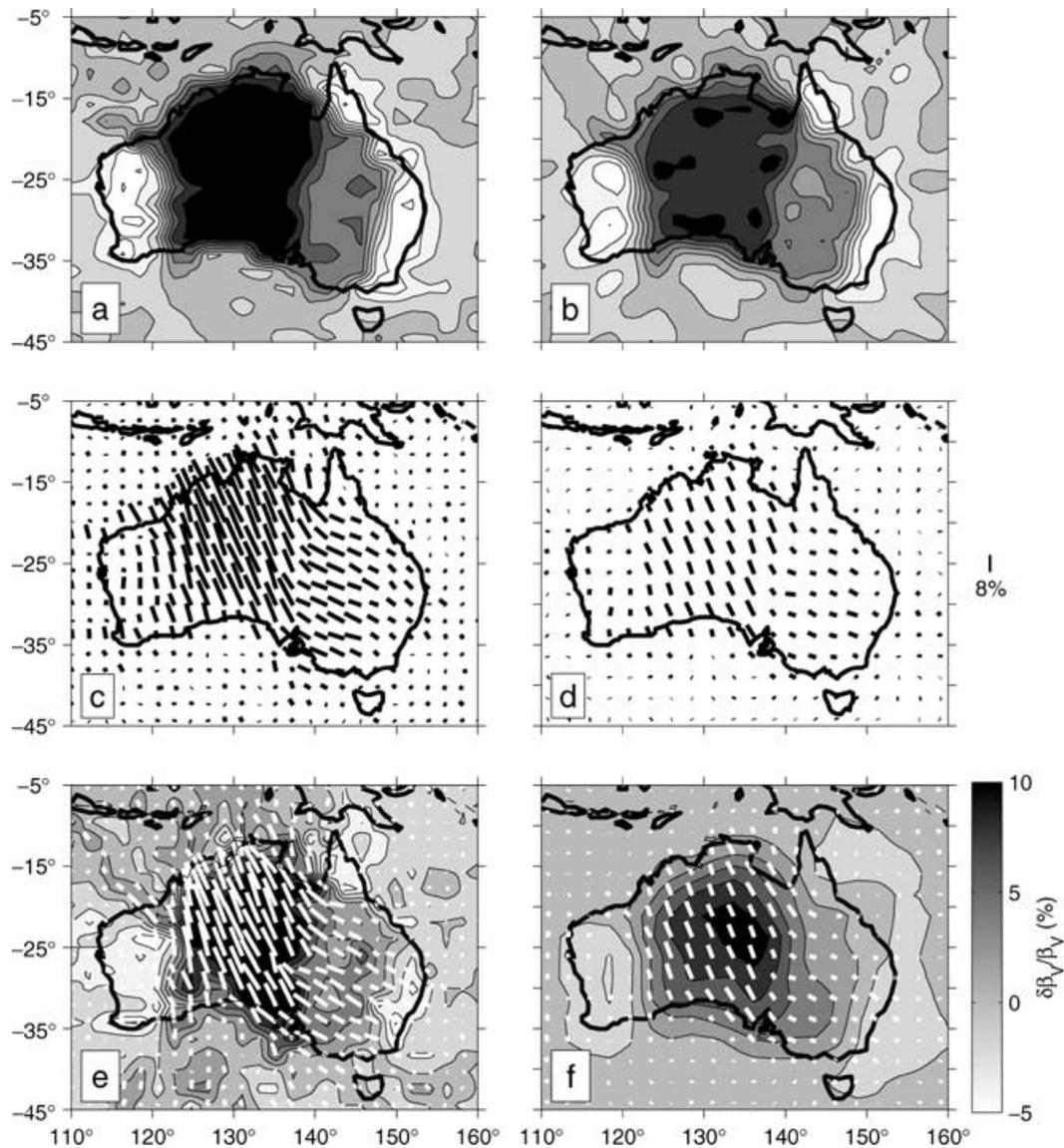


Figure 4. Comparison of the discrete iterative approach (left column) with the continuous regionalization inversion (right column). (a, b) Isotropic inversion result with the path distribution of Fig. 3(a) and input heterogeneity of Fig. 3(b). (c, d) Anisotropy-only inversion with input anisotropy of Fig. 3(c). (d, e) Joint inversion for velocity heterogeneity and anisotropy with input anomalies as in Figs 3(b) and (c).

In contrast, in Method I we incorporate the sensitivity to heterogeneity and anisotropy with an equal weighting. The relative amplitudes of anisotropy and heterogeneity are controlled by the weights of the damping terms of their norm and not well constrained by the data. In our model, the norm damping for the anisotropic parameters was twice as large as for the velocity heterogeneity. We evaluate the amount of trade-off *a posteriori* in various experiments (see Section 5.5).

Another difference is the treatment of the third dimension in the 3-D inversion. In Method I the third dimension is an integral part of the minimization problem, entering the equations through the depth-dependent linear constraints on the velocity and anisotropic structure (see Zielhuis & Nolet 1994). In this method, the path-averaged constraints entering the tomographic inversion are depth-dependent functions representing the particular depth-average of structure that can be resolved independently from all other constraints (Nolet 1990; Zielhuis 1992). With the continuous regularization technique of Montagner (1986) such a treatment of the

depth-dependent resolvability of Earth structure can be achieved by retaining the full 3-D covariance matrices $C(\mathbf{r}, \mathbf{r}')$ in the development. However, DK2000 invert the shear-wave speed path-averages layer by layer. The data used in their tomographic inversion are given by the shear-wave speed profiles obtained from the waveform inversion, sampled with an interval of 25 km, and individual depths are treated separately. The results are then stacked into a 3-D model. The variation of the resolution matrix with depth is given by the errors on the shear-wave speed at the particular depths. However, judging from the kernel functions that represent how surface waves sample the Earth, it is not possible to determine shear-wave speeds at such finely spaced depth independently from one another. As a consequence, with this method, it is impossible to obtain direct estimates of the variation of lateral resolution with depth and the extent of radial interdependence in the 3-D model.

Finally, in Method I, we use the inversion algorithm LSQR (Paige & Saunders 1982), which speeds up the computations by several

orders of magnitude, but no full resolution matrix can be computed directly. In contrast, Method II performs the inversion directly, which is computationally much more expensive but enables the explicit computation of the resolution matrix, a capability we will use in the next section.

5 RESOLUTION AND ROBUSTNESS ANALYSIS

5.1 Data coverage

The best sampled regions are the easternmost two-thirds of the Australian continent (Fig. 1), with a maximum path density on the eastern rim of the continent (Fig. 2). Towards the west, the path density degrades. Fig. 2(b) depicts the number of pairs of rays that come into the cell in a different azimuthal quadrant. This score is more uniform across the continent, but it also diminishes in the western half. The northeastern portion of the continent has a maximum of about 200 such pairs of crossing rays in one cell. In Fig. 2(c) we combined both resolution criteria into a single value ranging from 0 to 1.

5.2 Exact calculation of the resolution matrix

The resolution kernel R is given by the reduction of the *a posteriori* covariance C_p with respect to the *a priori* covariance structure:

$$R(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') - \frac{C_p(\mathbf{r}, \mathbf{r}')}{C_m(\mathbf{r}, \mathbf{r}')}. \quad (21)$$

Using the direct Bayesian inversion we calculate the exact resolution matrix both for the actual (Fig. 1) and the synthesized uniform (Fig. 3a) data coverage and compare this to the qualitative resolution estimates from checkerboard tests. We input the path-average velocity anomaly at a single depth into the Bayesian inversion code, with a correlation length of 175 km for both velocity and anisotropy, and *a priori* model uncertainties of 120 m s^{-1} for $\delta\beta_V$, A_1 and A_2 (see eq. 7). The value of these choices only affect the amplitude of the result; equating σ_{β_V} and σ_A corresponds to an equal weighting of velocity heterogeneity and anisotropy.

The results for the synthetic uniform path coverage of Fig. 3(a) are plotted in Fig. 5(a) (resolution of $\delta\beta_V$), Fig. 5(b) (resolution of A_1), and for the actual coverage of Fig. 1 in Figs 5(c) and (d). The resolution for A_2 is virtually identical to that of A_1 and is therefore omitted. The comparison of the synthetic resolution with the

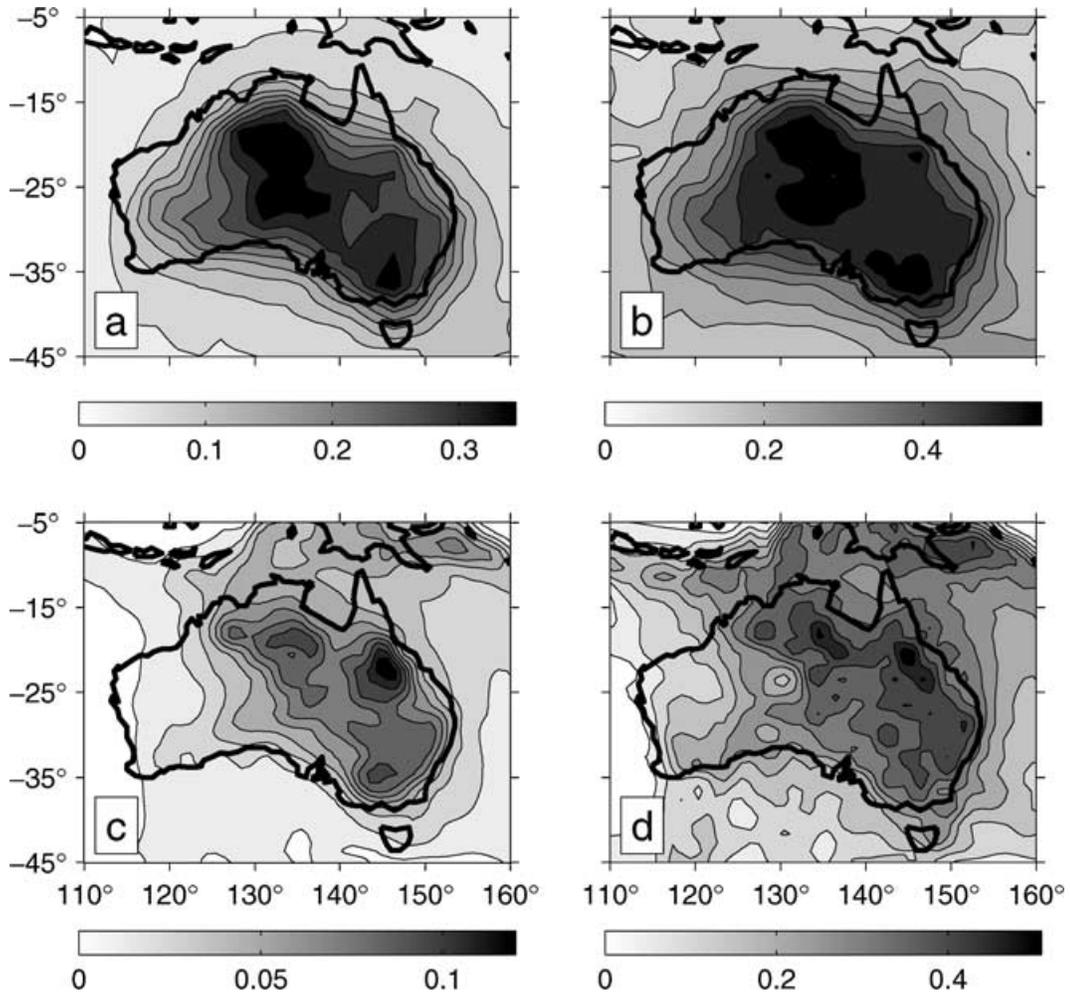


Figure 5. Resolution tests. Exact horizontal resolution computation for the synthetic path density experiment (a–b) and the real data coverage (c–d). Plotted is the diagonal value of $R(\mathbf{r}, \mathbf{r}')$ for the isotropic velocity solution $\delta\beta_V$ (a, c), and the cosine term A_1 (b, d). The resolution for the sine term A_2 is virtually indistinguishable from that of the cosine term.

resolution from the actual data gives us a feel for the values we can expect where the model is well resolved. As anticipated, the solution in the eastern two-thirds of the continent is markedly better than in the west.

5.3 Checkerboard tests

A popular way of assessing the resolution in tomographic models is to calculate the recovery of a synthetic (e.g. checkerboard) input pattern. By themselves checkerboard tests are difficult to interpret. It is often hard to guess the exact pattern of the anomalies without knowledge of the input function, which strongly degrades their diagnostic value. Also, the damping parameters can be chosen to reproduce the input pattern optimally, a luxury which is not available when choosing the damping needed to model the actual data for unknown Earth structure. These and other, intrinsic, limitations of checkerboard tests have been pointed out by several authors (e.g. L ev eque *et al.* 1993). Hence, we will use the results of checkerboard tests in a qualitative way only, and compare them to the exact resolution computed in Section 5.2.

Since the sensitivity matrix maps the input anomalies onto the synthetic data according to the independently constrained depth-dependent kernel functions, the solution at one particular depth is related to the solution at adjacent depths in Method I. Therefore, we can use checkerboard tests to get a sense of the radial resolution and vertical smearing of the tomographic model (Fig. 6) as well as the horizontal resolution and its depth variation (Fig. 7).

We have created two staggered checkerboard patterns, one for heterogeneity and one for azimuthal anisotropy (Figs 6a and b). Synthetic data were generated for a model with this input pattern introduced at 80 km depth. The inversion result is shown at this same depth (80 km: Figs 6c and d) as well as at the next level of the solution (140 km: Figs 6e and f). Except for the diminished amplitudes the velocity input pattern is reasonably well recovered,

including the sharp gradients between the cells. This and other tests suggest that large wave speed contrasts occurring over as little as 250 km can be resolved by our data. On the other hand this test does not tell us the amplitude of the minimum wave speed variation that can be detected at that same wavelength.

The N–S and E–W pattern of azimuthal anisotropy is fairly well recovered in the central and eastern part of the continent but the angular variability is at least 20  (Fig. 6d). Indeed, in large parts of our study region it is not obvious from Fig. 6(d) alone that the input pattern was a regular checkerboard pattern, which hampers the interpretation of the results of such tests significantly.

Radial smearing affects our results, but a comparison of the amplitudes of the solutions at both depths reveals that the level of contamination is small.

In Fig. 7 we illustrate the variation of horizontal resolution with depth. We have plotted the output-to-input ratio (in per cent) of an isotropic checkerboard test with patches of about 550 km each (5  imes 5 , performed at 80 (Fig. 7a) and 200 km (Fig. 7b)). At 80 km, we observe that up to 80 per cent of the input anomaly is recovered in the best resolved regions; at 200 km this number has dropped to about 60 per cent, but in most of the study region the amplitude recovery is less than 50 per cent. Figs 7(c) and 7(d) plot the difference in angle between the synthetic input pattern of Fig. 6(b) in the presence of the velocity heterogeneity of Fig. 6(a) and the recovery of the anisotropy from a joint inversion for heterogeneity and anisotropy, for an input pattern and model retrieval at 80 (also shown in Fig. 6d) and 200 km depth (also shown in Fig. 6e).

We conclude that, except perhaps in western Australia, our final model will give a fair representation of the velocity and anisotropy structure of the upper mantle under Australia.

5.4 Model robustness

In addition to the parametrization strategies described in Section 4.2 on the grids shown in Fig. 2, we evaluate the robustness of the

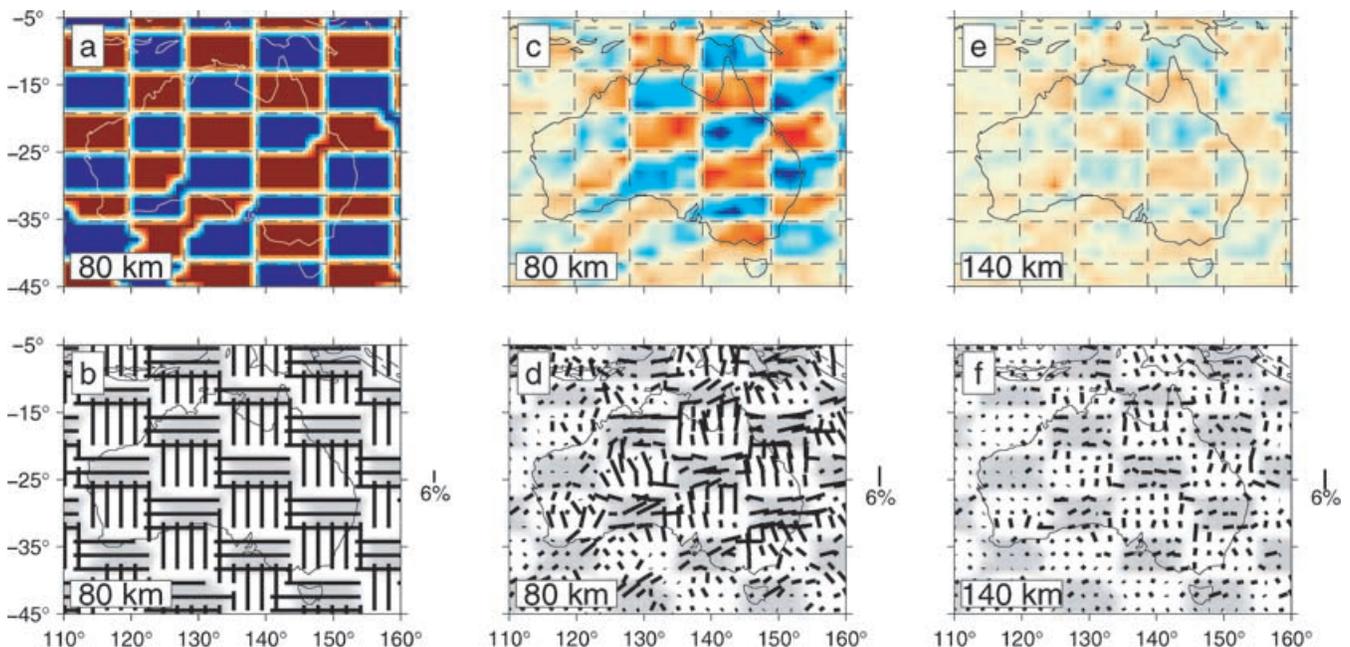


Figure 6. Checkerboard test for vertical smearing. An input pattern of velocity heterogeneity at 80 km (a) superimposed on a pattern of azimuthal anisotropy (b; shaded for clarity) offset by half a cell was used to generate synthetic data. The inversion for both wave speed anomalies and anisotropy yields a recovery at 80 km given in (c) and (d). Due to the radial averaging properties of the surface-wave kernels and the smearing in the inversion, some of the input structure at 80 km is also mapped into the next layer at 140 km (e and f). Panels a, c, and e are plotted on the same color scale.

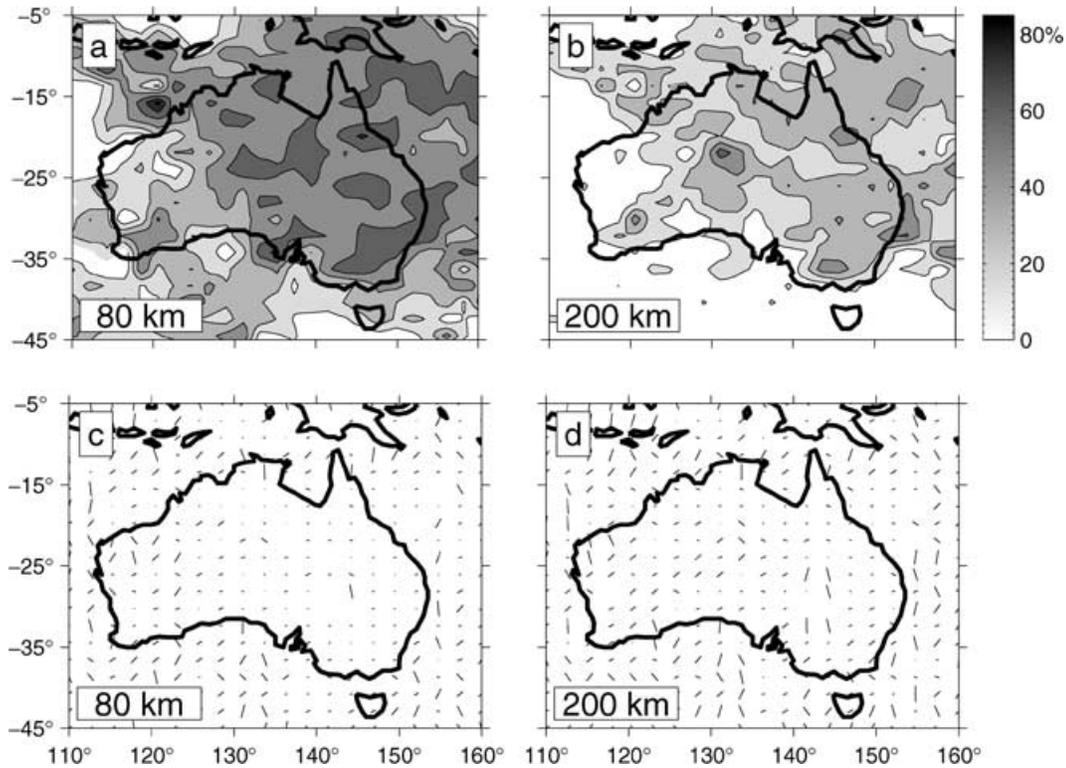


Figure 7. Checkerboard test for depth-varying horizontal resolution. (a, b) Percentage of input anomaly recovered of an isotropic checkerboard pattern at 80 and 200 km depth. (c, d) Difference in angle between input and output anisotropy in a joint inversion for a pattern of heterogeneity and anisotropy as in Figs 6(a)–(b) at 80 and 200 km depth. Bars scale with the magnitude of the difference such that excellent recovery is depicted by a dot.

models by an extensive series of tests. We performed inversions with up to 40 per cent of the data randomly removed. We also calculated the mean model obtained from 500 inversions with 5 per cent of the data randomly removed, as well as the mean of 500 inversions performed on a data set to which noise was added (with signal-to-noise ratios up to 15 dB). The average models obtained in this way were visually indistinguishable from our final model. For this class of mixed-determined inversion problems it is not possible to use inversion experiments like these to interpret the variance of the model parameters quantitatively everywhere. Areas with low resolution are always damped towards the reference model, and this perfectly known state of total uncertainty has therefore almost zero variance in such inversions. However, for the well resolved areas in the east, the standard deviation obtained is everywhere less than 0.8 per cent, which is much smaller than the magnitude of the anomalies we obtain in the final model. This bound holds for all depths of our final model.

5.5 Trade-off of heterogeneity and anisotropy

Several previous studies (e.g. Vinnik *et al.* 1992; Gaherty & Jordan 1995; Clitheroe & van der Hilst 1998; Girardin & Farra 1998; Debayle & Kennett 2000a,b) have suggested that the upper mantle beneath Australia is anisotropic. However, we need to show that the inclusion of the azimuthal anisotropy into the inversion problem is required by our data.

The increase in variance reduction of isotropic-only (82.6 per cent in a separate inversion with all parameters unchanged) versus azimuthally anisotropic models (86.1 per cent for our model) is an indication of the improvement of data fit when anisotropy is

included, but its statistical significance is hard to assess. Methods to evaluate the improvement of data fit with an increasing number of parameters, such as Fischer tests, are only applicable to overdetermined problems (Menke 1989) and can thus not be used here. One indication of improvement of data fit is by comparing the increase in variance reduction from an isotropic model ($\delta\beta_V$) parametrized in 7059 surface cells of a certain area with an anisotropic inversion ($\delta\beta_V, A_1, A_2$) of 2353 cells of an area three times as large. The number of inversion parameters is the same in either case, but the variance reduction increases from 83.5 to 86.1 per cent. Thus, the inclusion of anisotropy results in an increase of variance reduction even when the number of free parameters is left unchanged.

We have studied the incremental improvement in variance reduction in our final model due to the inclusion of progressively deeper layers of heterogeneity, anisotropy, or both. For the velocity heterogeneity alone, the depth intervals between 30 and 400 km depth contribute most to explain the variance in the data, reaching a plateau of 76.9 per cent for the model between 0–400 km and a maximum of 77.5 per cent for the entire depth range (0–670 km). Anisotropy alone between 0–400 km only explains 20.1 per cent of the variance in the data, terminating at 20.5 per cent for the entire model between 0–670 km. The improvement in variance reduction by the addition of successive layers of anisotropy to a model with heterogeneity between 0–670 km is from 77.5 per cent to 85.9 per cent at 400 km, terminating at 86.1 per cent for the entire model.

We explore the significance of the azimuthal anisotropy with another synthetic experiment designed to test to which extent uneven data coverage and heterogeneity can combine to produce an anisotropic bias. We represent our full solution by $(\delta\beta_V, A_1, A_2)$. In one test we inverted data synthesized from the heterogeneity part only, (i.e. $(\delta\beta_V, 0, 0)$) for both heterogeneity and anisotropy. From

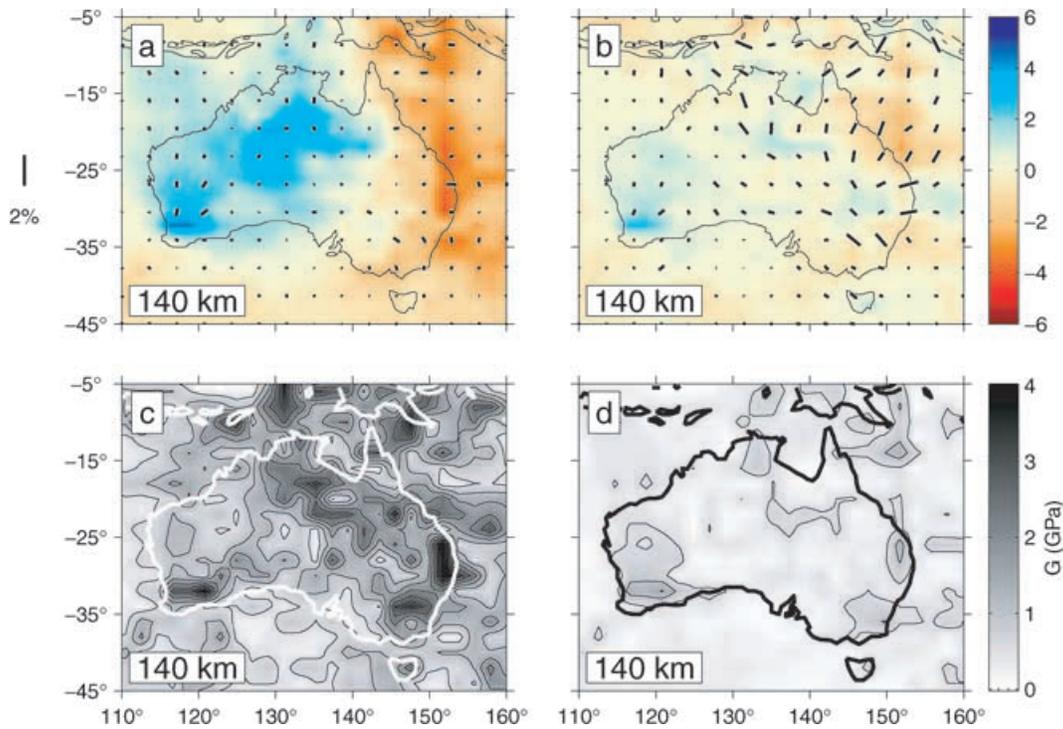


Figure 8. Trade-off of heterogeneity with anisotropy at 140 km. (a) Inversion for heterogeneity and anisotropy of synthetic data calculated from the isotropic part of our model. With perfect resolution and control of the trade-off, the anisotropy retrieved should be zero everywhere. (b) Inversion of synthetic data calculated from the anisotropic part of the model for heterogeneity and anisotropy. In this case perfect resolution should have resulted in zero heterogeneity. (c) Magnitude of anisotropy (in GPa) of the actual model (see Fig. 9). (d) Magnitude of the erroneous anisotropy of Fig. 8(a). Contour lines plotted at 1 GPa interval for both panels.

the result we can then infer how much anisotropy is produced by uneven data coverage and heterogeneity. In a second test we invert synthetic data from $(0, A_1, A_2)$ for heterogeneity and anisotropy to assess the projection of artificial anisotropy into heterogeneity.

Fig. 8 summarizes the test results for a depth of 140 km. Figs 8(a) and 8(b) show the inversion of models which should have zero residual anisotropy and zero residual velocity heterogeneity, respectively. The magnitude of the anisotropy in the actual model plotted in Fig. 8(c), can be compared to the magnitude of the “erroneous” anisotropy (Fig. 8d). For all depths considered, the average anisotropic bias is less than 20 per cent of the magnitude required to satisfy the data in the actual model.

In regions of uneven data coverage some artificial anisotropy is clearly produced. The direction of the artificially induced anisotropy is especially sensitive to uneven sampling, but, in general, the magnitude of the bias is small. In Western Australia, however, much of the anisotropy present in the actual model may well consist of bias due to the scarcity of crossing paths in that region. This makes it impossible to distinguish the fast velocity anomaly, which is clearly detectable in the waveforms, from a possible alignment of the fast axis of anisotropy.

6 RESULTS

In Figs 9 and 10 we present models for heterogeneity and azimuthal anisotropy of Australia, using Method I. The inversions were performed on the regular (equal-area) grid shown in Figs 2(a) and (b).

The azimuthally isotropic component $\delta\beta_V$ shows many of the characteristics that were present in our earlier models (Zielhuis & van der Hilst 1996; Simons *et al.* 1999). Low wave speed anomalies

mark the eastern Phanerozoic rim of the continent. They are associated with thermal anomalies, recent volcanism and high attenuation (Mitchell *et al.* 1998). Wave speeds in the central and western region are high, but the correlation between the high velocities and the age of the overlying crust is scale-dependent (Simons *et al.* 1999).

In most regions the anisotropy suggested by the data exceeds the anisotropic bias (see also Fig. 8), and the lateral variability in directions is much larger than the nominal uncertainty of 20° . There is no obvious relationship between wave speeds and anisotropy, neither in magnitude nor in direction, nor is there a clear correlation between the anisotropy and the large-scale geological structure (Fig. 11).

The cross-sections shown in Fig. 10 suggest a depth to the base of the continental lithosphere of up to 250–300 km (see also Simons *et al.* 1999; Fischer & van der Hilst 1999), and Fig. 11 shows that the wave speed difference between the large-scale tectonic provinces is confined to the shallowest 300 km. The central-Australian region is underlain by the thickest high-velocity lid, which may be an effect of preferential preservation (Simons *et al.* 1999). Even with the uncertainty estimates given in Section 5.5, the magnitude of the anisotropy does not fall off with depth as rapidly as the magnitude of the velocity heterogeneities. The average anisotropic magnitudes reach a maximum around 100 km (Fig. 11) and diminish at larger depth, in accordance with observations elsewhere (Babuška *et al.* 1998). For comparison, the magnitudes of the anisotropy from laboratory estimates of upper-mantle materials range from 2 to 7 GPa (Peselnick & Nicolas 1978; Estey & Douglas 1986).

Although not all individual directions of anisotropy agree with the model by DK2000, we concur with these authors that the pattern of anisotropy changes with increasing depth. The solutions at 80 and 140 km depth show a complex pattern of azimuthal

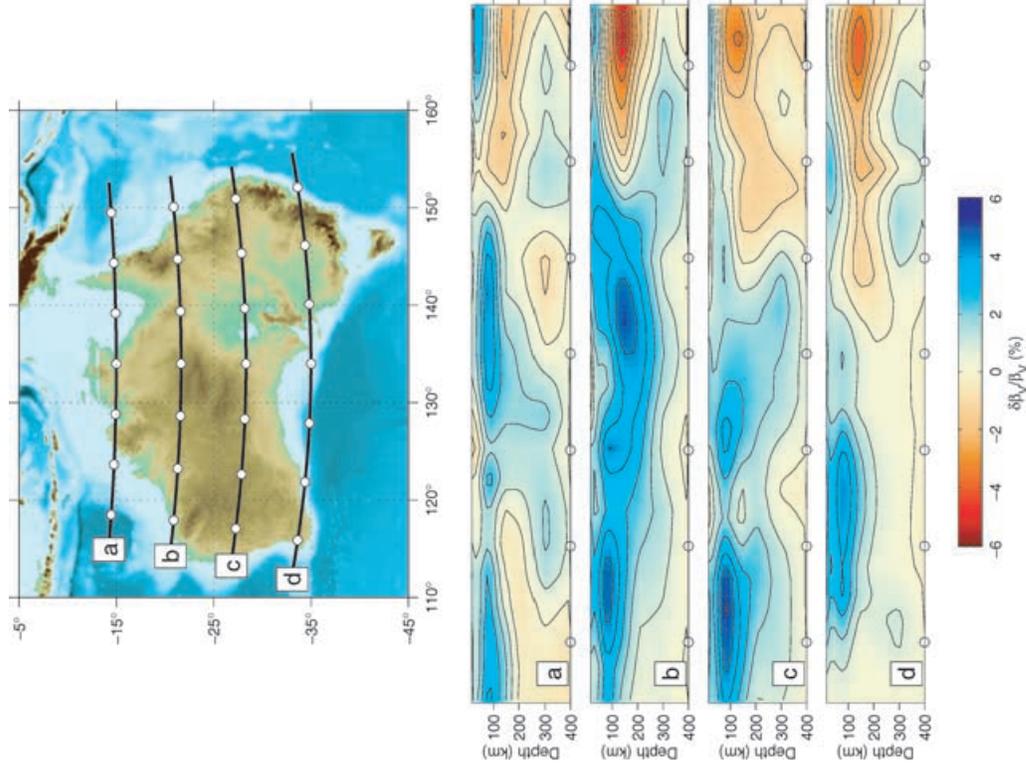


Figure 10. Cross-sections at various latitudes. Fifteen contour lines are plotted between -6 and 6 per cent.

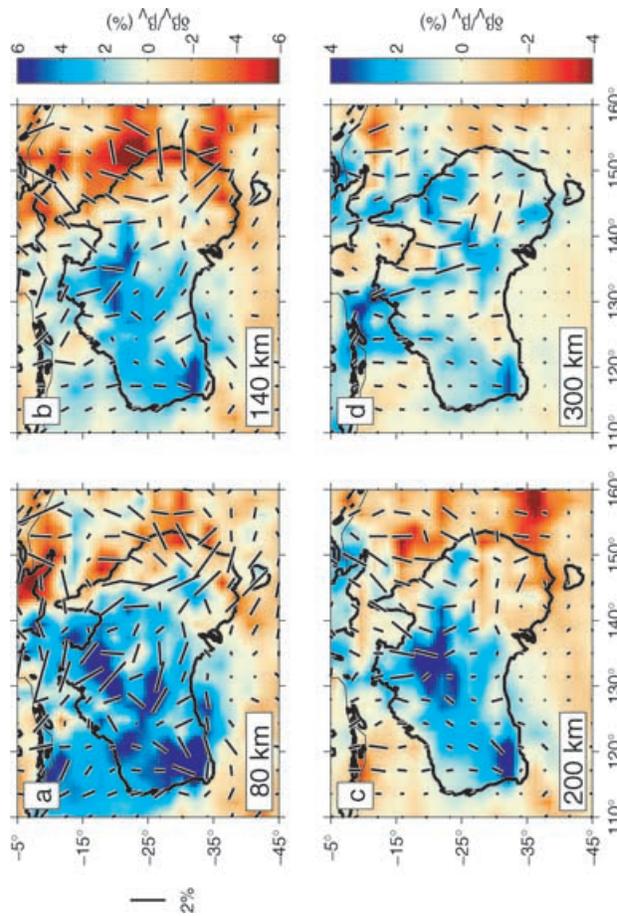


Figure 9. Inversion results at different depths. Percentages from reference model with background shear-wave velocities of 4500 m s^{-1} at 80 and 140 km depth; 4513.5 m s^{-1} at 200 km ; 4635 m s^{-1} at 300 km .

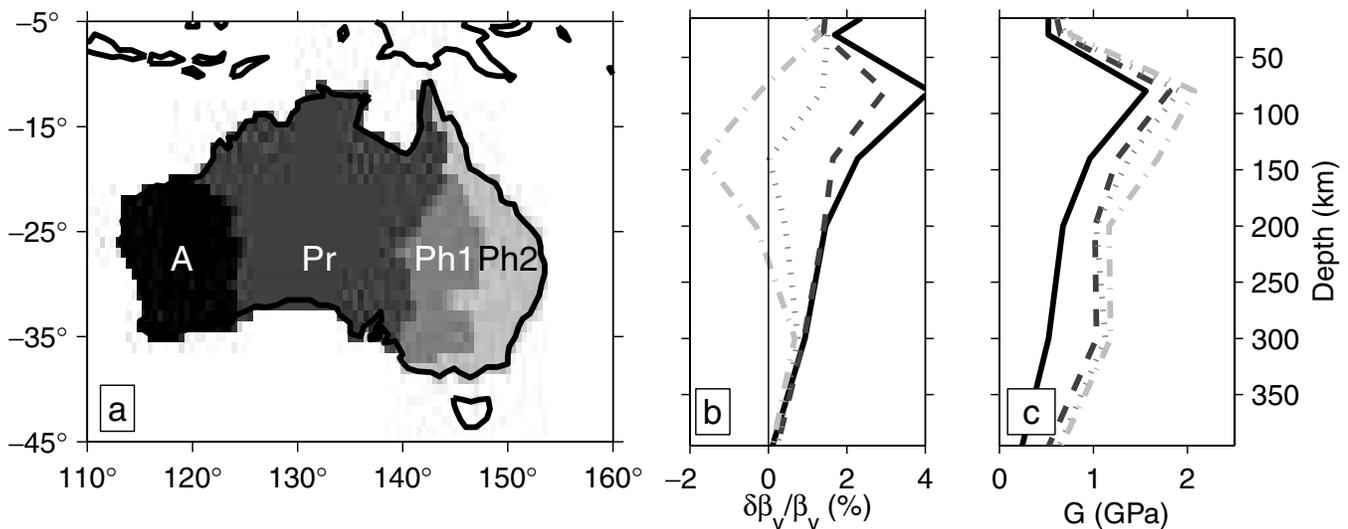


Figure 11. Average heterogeneity and anisotropy in function of crustal age. A coarse regionalization of the Australian continent consists of a predominantly Archaean domain (A), the central Proterozoic region (Pr) and domains of Phanerozoic age (Ph1 and Ph2). While wave speed anomalies become progressively more positive with increasing crustal age (see also Simons *et al.* 1999), no defining relationship is found in the magnitude of the anisotropy.

anisotropy (short correlation wavelengths), whereas in the layers at 200 and 300 km a simpler, smoother pattern (long correlation wavelengths) is observed (Fig. 9). The character of the anisotropy and the length scale over which it varies seem to change rapidly across the 150 and 200 km depth interval.

The velocity heterogeneity revealed in Fig. 9 was reproduced satisfactorily with the irregular parametrization (Fig. 2c), albeit with somewhat reduced amplitudes. In the western regions, where the path coverage is demonstrably worse, our fine equal-area grid overparametrizes the structure but this does not introduce noticeable artifacts. Within the uncertainty of 20° , the different parametrizations yield similar directions of fast axes of anisotropy. The pure-path average solutions (on the grid shown in Fig. 2d) at 80, 140 and 200 km were very similar to the averages of the wave speeds imaged on the equal-area grid. Differences appear at larger depth, suggesting again that below 200 km most correspondence between wave speed structure and surface geology is lost. The central and northern region of Australia have stable directions of maximum anisotropy, both under the equal-area and the tectonic parametrization, but the magnitudes are much reduced in the latter. In the eastern portion of the continent a tectonic parametrization reveals little anisotropy, at any depth. This could indicate that the anisotropy is small, i.e. at or below the level of the anisotropic bias, or that through much of the depth range the anisotropy varies on length scales smaller than the main tectonic units of the Australian continent.

7 DISCUSSION

7.1 Thickness, age-relations and deformation of the Australian lithosphere

Many global models of Rayleigh wave phase velocities (e.g., Trampert & Woodhouse 1996; Ekström *et al.* 1997) have measured faster than average wave speeds for the Archaean and Proterozoic portions of the Australian continent (which correspond roughly to the western two-thirds of the continent). This pattern is corroborated by our regional model, but significant variations in heterogeneity and anisotropy exist at smaller scales than is apparent

from global models. In particular, the depth extent of the lithosphere is determined in much more detail than possible with global data sets. Some global models suggest that the high wave speeds associated with ancient continents extend to depths exceeding 400 km (e.g. Su *et al.* 1994; Masters *et al.* 1996). Our detailed 3-D model verifies instead the local measurements of lithospheric thickness of ~ 200 km quoted for Northern Australia (Dey *et al.* 1993; Kennett *et al.* 1994; Tong *et al.* 1994), and ~ 300 km under Central Australia (Gaherty & Jordan 1995; Gaherty *et al.* 1999). A ~ 200 km thick lithosphere in North Australia is compatible with our observations (see, e.g. Fig. 10a), but clearly the 3-D structure of the lithosphere–asthenosphere boundary is complex, and the thickness of the lithosphere probably reaches at least 300 km in certain parts of Central Australia (Fig. 10b). Eastern Australia, which is the most recently tectonized region of the continent, is characterized by a pronounced low velocity zone between 100–200 km depth underlain by fast anomalies at a depth of about 300 km (see Fig. 10b). The origin of the latter is uncertain and it may not be related to the overlying lithosphere. We conclude that the thickness of the Australian lithosphere is about 225 ± 50 km. This is in agreement with recent observations for the Canadian Shield (Frederiksen *et al.* 2001) and South-Africa (James *et al.* 2001). Similar lithospheric thicknesses are obtained for other cratonic regions (for a review, see Artemieva & Mooney 2001). Elsewhere, Simons & van der Hilst (2002) report the full variation of lithospheric thickness based on contours of the velocity anomalies in our present model in more detail.

The wave speed heterogeneity of the Australian continent seems to correlate with crustal evolution and lithospheric age, at least to depths of about 200 km (Fig. 11) (see also Simons *et al.* 1999). At depths greater than 250 km there is no unambiguous relationship between crustal age and wave speed. The anisotropy does not seem to change on the length scale of the major geological domains at any depth, although its magnitude remains higher at larger depth beneath the regions affected by the most recent tectonic activity (the easternmost parts of the continent). In contrast, Babuška *et al.* (1998) argued that there are systematic variations in radial anisotropy which are related to the age of continental provinces in North America and Eurasia.

Global and regional studies (e.g. Montagner & Tanimoto 1991; L ev eque *et al.* 1998) have been interpreted to show evidence of a shallow regime of rapidly varying anisotropy underlain by a region with smoother patterns of anisotropy. DK2000 observe this change at 150 km depth. There is uncertainty on the magnitude and directions of the azimuthal anisotropy, but qualitatively, the wave speed and anisotropy maps presented in Fig. 9 argue for a change in regime between the first two depths (80 and 140) and the deeper parts of the solution (200 and 300 km). The first regime is characterized by a complex pattern of strong anisotropy, whereas the second presents smoother variations of weaker anisotropy and a tendency towards N–S fast axes increasing with depth.

Gaherty *et al.* (1999) and Debayle & Kennett (2000b) found that polarization anisotropy in the Australian continent is concentrated above 250 km. Our Rayleigh wave model does not contain information about polarization anisotropy. However, most of the anisotropic signal is present in the depth range above 200 km depth (Fig. 11), and the improvement in variance reduction due to the inclusion of azimuthal anisotropy is largely explained by the layers above 300 km depth. While we believe we see evidence of azimuthal anisotropy until a depth of about 300 km, its magnitude is not incompatible with the statement by Gaherty *et al.* (1999) that polarization anisotropy below 250 km depth must be small (<1 per cent). We interpret our model to show evidence for fossil anisotropy above 200 km and deeper anisotropy which can be reconciled with the overall direction of absolute plate motion. The precise interpretation of this model in terms of continental deformation will be the subject of forthcoming research.

7.2 Comparison with results from shear-wave splitting

From the analysis of refracted *S*-waves bottoming under North Australia, Tong *et al.* (1994) obtain shear-wave splitting times exceeding 4 s in certain places. Vinnik *et al.* (1989) and Barruol & Hoffman (1999) report no perceptible splitting for *SKS* arrivals at station CAN in southeastern Australia, but Girardin & Farra (1998) interpret *P*-to-*S* converted phases to be indicative of a two-layer anisotropic system under that station with nearly perpendicular fast axes distributed over comparable depth intervals so the contributions to splitting cancel out. Clitheroe & van der Hilst (1998) present evidence for complex, frequency-dependent splitting from *SKS* and *SKKS* phases, with split times generally less than 1 s in the Precambrian central part of Australia. Their fast axis directions do not align with the direction of absolute plate motion, which is roughly SSW–NNE, but rather suggest alignment with crustal fabric or a complex lattice preferred orientation of olivine in the subcrustal lithosphere. However,  zalaybey & Chen (1999) find no convincing evidence for splitting at all. They argue that the analysis of a small amount of data over a limited azimuthal range and complex multipathing effects due to heterogeneities at depth may contribute to erroneous splitting measurements, especially at the frequencies higher than 0.3 Hz considered by Clitheroe & van der Hilst (1998).

SKS splitting times represent an integrated effect over a large depth range h due to the subvertical propagation of the phases. It is possible to relate our observations of surface-wave anisotropy to shear-wave splitting directly (Montagner *et al.* 2000). For simple media with a horizontal symmetry axis, the maximum splitting time is given by

$$\delta t = \sqrt{\left(\int_0^h \sqrt{\frac{\rho}{L}} \frac{G_c(z)}{L(z)} dz\right)^2 + \left(\int_0^h \sqrt{\frac{\rho}{L}} \frac{G_s(z)}{L(z)} dz\right)^2}, \quad (22)$$

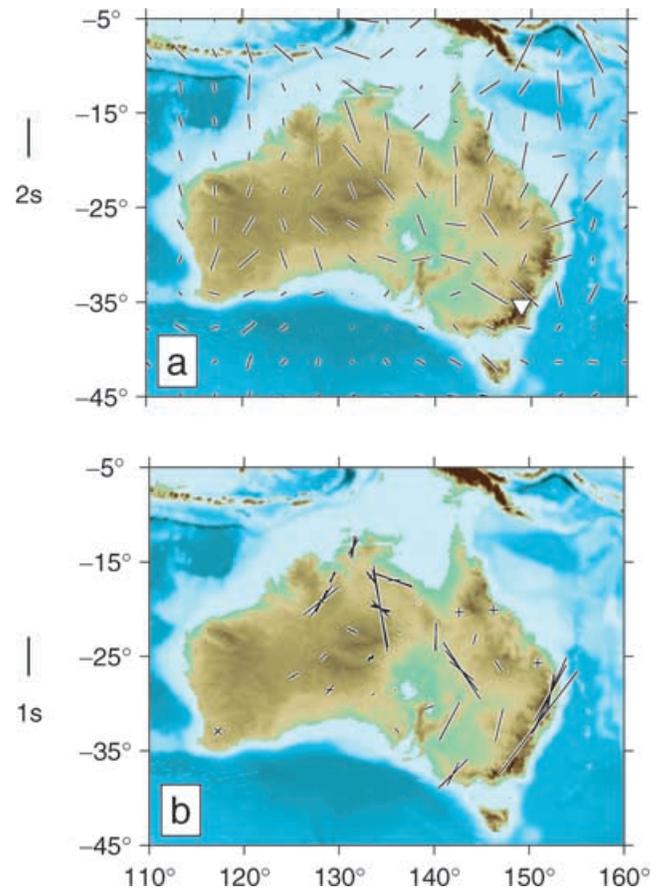


Figure 12. (a) *SKS* splitting directions predicted from the solution given in Fig. 9. Station CAN is represented by the white triangle. (b) Shear-wave splitting observations (from Idaka, in prep.).

and the splitting direction Ψ is calculated by

$$\tan 2\Psi = \int_0^h \sqrt{\frac{\rho}{L}} \frac{G_s(z)}{L(z)} dz / \int_0^h \sqrt{\frac{\rho}{L}} \frac{G_c(z)}{L(z)} dz. \quad (23)$$

Applying eqs (22) and (23) to our model (depicted in Fig. 9) yields considerable splitting times throughout the continent (Fig. 12a). The results of the surface-wave inversion presented here suggest higher levels of azimuthal anisotropy than inferred from *SKS* by either Clitheroe & van der Hilst (1998) or  zalaybey & Chen (1999). For comparison, we have plotted a recent data set of splitting times and directions in Fig. 12(b) (from Idaka, in prep.).

In addition to calculating *SKS* times due the entire depth range of anisotropy, we have investigated the incremental effect on the splitting times from successively deeper levels in our model. The transverse isotropy (with horizontal symmetry axis) between 140 and 400 km depth contributes most to the calculated *SKS* split times. This is in good agreement with Tong *et al.* (1994), who argue that, in Northern Australia, the most likely source region of anisotropy is the asthenospheric zone between 210–410 km.

The fast axis directions in the eastern portions of the continent follow the rough outline of the high wave speeds associated with the Australian Precambrian (Fig. 9a). Clitheroe & van der Hilst (1998) have suggested this might be related to the crystal alignment by asthenospheric flow around the irregularly shaped eastern edge of the deep continent (see also Fouch *et al.* 2000). This is supported by the fact that the split times originate mainly at a depth near 140 km, in

a zone of anomalously low velocities and, presumably, low viscosities (Figs 10 and 11). In the center of the continent, the predicted splitting times are somewhat smaller, which may in part be due to a change from E–W in the shallower part to N–S in the deeper continental lithosphere (see, e.g. Silver & Savage 1994; Saltzer *et al.* 2000).

A mismatch between surface-wave and body-wave anisotropy has been observed on other continents as well (Saltzer *et al.* 2000; Freybourger *et al.* 2001). We should point out that the assumption of a horizontal orientation of the fast axis is not unlikely to be violated, especially in the shallow parts of the continental lithosphere. Further analysis will require taking into account the depth variation of fast axis directions, the possibility of dipping axes and frequency-dependent effects.

Our observations point to a complex geometry of the continental anisotropy at various length scales. We cannot rule out that strong anisotropy in the top 80–140 km exists if the direction of fast axes varies rapidly on a scale that is smaller than can be resolved by our data. With the magnitude of anisotropy increasing and the lateral correlation lengths decreasing towards the surface, it is conceivable that strong, rapidly changing azimuthal anisotropy in the top 140 km diminishes the effective birefringence of vertically propagating body waves without similarly influencing the long-period surface-wave data used in our study (Saltzer *et al.* 2000).

8 CONCLUSIONS

Manual data selection and processing of stations operated by AGSO, IRIS and GEOSCOPE and data from the portable SKIPPY experiments have produced a data coverage that is superior to surface-wave studies on any other continent in terms of density and azimuthal distribution.

We have compared two tomographic approaches for the inversion of path-average earth models for 3-D models of *S*-heterogeneity and azimuthal anisotropy. Method I (Zielhuis & van der Hilst 1996; Simons *et al.* 1999) performs a 3-D inversion for heterogeneity and anisotropy based on a discrete parametrization at the surface with depth basis functions representing independently constrained estimates of radial earth structure. Method II, used by Debayle & Kennett (2000a), uses a continuous regionalization algorithm which performs a separate inversion for each depth layer. Discretely parametrized regularization methods and Bayesian continuous regionalization methods can lead to similar results. The differences between our results and the model by Debayle & Kennett (2000a) can be due to (1) the independence of the constraints from the waveform inversion used in Method I and the decorrelation of their errors; (2) the treatment of the third dimension (3-D inversion in Method I but a stack of layers in Method II); (3) somewhat different distribution of earthquake sources; (4) different data selection criteria (e.g., the frequency bands used); (5) balancing of fundamental-mode and higher-mode information through the use of one (Method II) or more (Method I) data windows; and (6) the crustal models used to apply shallow corrections.

The advantage of Method II is that a formal resolution matrix can be computed, but this comes at a substantial computational expense. We have shown that the information contained therein is qualitatively similar to *ad hoc* checkerboard response tests. Furthermore, the formal resolution is 2-D since the full 3-D covariance matrices of data and model are not used in Method II. As a result, variations in lateral resolution with depth and radial smearing can not be investigated efficiently with Method II. In contrast, Method I allows

us to evaluate the uncertainty as well as the radial smearing in all parts of the 3-D model.

We have obtained a new shear-wave speed model for the Australian upper mantle which includes azimuthal anisotropy. The inversion presented here was based on 30 per cent more data than used by Simons *et al.* (1999), but we confirm their conclusions regarding the aspherical variation of isotropic wave speed in the Australian upper mantle. In particular, the conclusion that the high correlation between the westward increase in wave speed and lithospheric age holds only for depths less than 200 km is confirmed.

We have determined the spatial variation of azimuthal anisotropy (transverse anisotropy with a horizontal axis of symmetry) and investigated the trade-off between heterogeneity and anisotropy. We conclude that with the exception of westernmost Australia, where data coverage is worst, the level of anisotropy that is produced by a combination of uneven data coverage and heterogeneity is at most 20 per cent of the magnitude of anisotropy required by the data. In general the uncertainty in the angle of fast polarization is about 20°; locally the directions are more tightly constrained, but there are also regions where the uncertainty is much larger, in particular in western Australia. In the top 150 km the inversions yield fairly strong anisotropy, and the orientation of the fast axis changes in a complex manner. In contrast to the wave speed heterogeneity, there is no obvious relationship between anisotropy and the large scale geological domains that we considered, suggesting that the length scale for variation is much smaller than the crustal elements identified by Shaw *et al.* (1995). The horizontal length scale over which anisotropy changes increases between 150 and 200 km depth.

In the northern central region we observe a predominance of N–S directions throughout the depth range ≥ 140 km. Within the substantial uncertainty on their direction, these N–S oriented domains are subparallel to the direction of absolute plate motion (DeMets *et al.* 1990). If shearing due to plate motion is the dominant process generating seismic anisotropy, then it is perhaps surprising that so many of the strong N–S anisotropic directions are confined to central Australia and start at depths as shallow as 140 km or even less, i.e., well within what we interpret as continental lithosphere. Furthermore, complex flow around stable continental interiors could yield anisotropic fabrics in relationships that are not straightforward to correlate with a single motion vector of the entire Indo-Australian plate. If fossil anisotropy is the dominant process, then one needs to explain the change in length scale and behavior between 150 and 200 km depth for the regions outside of the Central Australian domain. Most likely, an interpretation will involve a combination of fossil and present-day anisotropy, but modelling this will require further study.

We have used the approach suggested by Montagner *et al.* (2000) to predict *SKS* splitting times and directions from our 3-D anisotropic model. Much as the original inversion result for depths above 200 km, the predicted *SKS* splitting map does not reveal a predominance of directions parallel to plate motion. The first order difference between our result and observations of *SKS* and *SKKS* splitting is that the surface waves seem to prefer a higher level of horizontal transverse anisotropy than the body waves. This is true, in particular, for the Precambrian shields where shear-wave birefringence yielded split times of less than 1 s whereas the surface-wave inversion predicts times that are at least twice that large. The amplitude of anisotropy relative to the velocity anomalies, to which splitting delay times are sensitive, is not well resolved, but even with correct delay times, there remains a significant mismatch between the predicted and the observed splitting directions. In addition, frequency-dependent effects or the presence of dipping axes of

symmetry might be responsible. We argue that, at least conceptually, the observed shear-wave splitting could be reconciled with the inferences from surface-wave propagation with a model of anisotropy in which the correlation wavenumbers and magnitude of anisotropy increase towards the surface.

ACKNOWLEDGMENTS

Funding for the instruments and field campaigns involved in the SKIPPY and KIMBA experiments was provided by the Australian National University. We thank Eric Debayle, Don Forsyth, Hrafnkell Káráson, Brian Kennett, Sergei Lebedev and Albert Tarantola for stimulating discussions, Takashi Iidaka for providing his shear-wave splitting analyses and Mark Leonard and the IRIS Data Management Center for data access. Guust Nolet, Steve Ward, and an anonymous reviewer provided comments which improved the manuscript. RvdH and FJS thank the Institut de Physique du Globe in Paris, where parts of this paper were written. This research was supported by NSF (grant EAR-0001136) and the David and Lucile Packard Foundation, and it is IGP contribution # 1793.

REFERENCES

- Abers, G.A. & Roecker, S.W., 1991. Deep structure of an arc-continent collision: Earthquake relocation and inversion for upper mantle *P* and *S* wave velocities beneath Papua New Guinea, *J. geophys. Res.*, **96**, 6379–6401.
- Anderson, D.L. & Dziewonski, A.M., 1982. Upper mantle anisotropy: Evidence from free oscillations, *Geophys. J. R. astr. Soc.*, **69**, 383–404.
- Artemieva, I.M. & Mooney, W.D., 2001. Thermal thickness and evolution of Precambrian lithosphere: A global study, *J. geophys. Res.*, **106**, 16 387–16 414.
- Babuška, V., Montagner, J.-P., Plomerová, J. & Girardin, N., 1998. Age-dependent large-scale fabric of the mantle lithosphere as derived from surface-wave velocity anisotropy, *Pure appl. Geophys.*, **151**, 257–280.
- Barrool, G. & Hoffman, R., 1999. Upper mantle anisotropy beneath the GEOSCOPE stations, *J. geophys. Res.*, **104**, 10 757–10 773.
- Cara, M. & Lévêque, J.-J., 1987. Waveform inversion using secondary observables, *Geophys. Res. Lett.*, **14**, 1046–1049.
- Clitheroe, G. & van der Hilst, R.D., 1998. Complex anisotropy in the Australian lithosphere from shear-wave splitting in broad-band *SKS*-records, in *Structure and Evolution of the Australian Continent, Geodynamics Series*, vol. 26, eds Braun, J., Dooley, J.C., Goleby, B., van der Hilst, R.D. & Klootwijk, C., pp. 73–78, AGU, Washington, DC.
- Clitheroe, G., Guðmundsson, Ó. & Kennett, B.L.N., 2000. The crustal thickness of Australia, *J. geophys. Res.*, **105**, 13 697–13 713.
- Clitheroe, G., Guðmundsson, Ó. & Kennett, B.L.N., 2000. Sedimentary and upper crustal structure of Australia from receiver functions, *Aust. J. Earth Sci.*, **47**, 209–216.
- 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. & Kennett, B.L.N., 2000. The Australian continental upper mantle: Structure and deformation inferred from surface waves, *J. geophys. Res.*, **105**, 25 423–25 450.
- Debayle, E. & Kennett, B.L.N., 2000. Anisotropy in the Australasian upper mantle from Love and Rayleigh waveform inversion, *Earth Planet. Sci. Lett.*, **184**, 339–351.
- DeMets, C., Gordon, R.G., Argus, D.F. & Stein, S., 1990. Current plate motions, *Geophys. J. Int.*, **101**, 425–478.
- Dey, S.C., Kennett, B.L.N., Bowman, J.R. & Goody, A., 1993. Variations in upper mantle structure under northern Australia, *Geophys. J. Int.*, **114**, 304–310.
- Eckström, G., Tromp, J. & Larson, E.W.F., 1997. Measurements and global models of surface wave propagation, *J. geophys. Res.*, **102**, 8137–8157.
- Engdahl, E., van der Hilst, R.D. & Buland, R., 1998. Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seism. Soc. Am.*, **88**, 722–743.
- Estey, L.H. & Douglas, B.J., 1986. Upper mantle anisotropy: A preliminary model, *J. geophys. Res.*, **91**, 11 393–11 406.
- Fischer, K.M. & van der Hilst, R.D., 1999. A seismic look under the continents, *Science*, **285**, 1365–1366.
- Fouch, M.J., Fischer, K.M., Parmentier, E.M., Wyssession, M.E. & Clarke, T.J., 2000. Shear wave splitting, continental keels, and patterns of mantle flow, *J. geophys. Res.*, **105**, 6255–6275.
- Frederiksen, A.W., Bostock, M.G. & Cassidy, J.F., 2001. *S*-wave velocity structure of the Canadian upper mantle, *Phys. Earth Planet. Inter.*, **124**, 175–191.
- Freybourger, M., Gaherty, J.B. & Jordan, T.H., 2001. Structure of the Kaapvaal craton from surface waves, *Geophys. Res. Lett.*, **28**, 2489–2492.
- Gaherty, J.B. & Jordan, T.H., 1995. Lehmann discontinuity as the base of an anisotropic layer beneath continents, *Science*, **268**, 1468–1471.
- Gaherty, J.B., Kato, M. & Jordan, T.H., 1999. Seismological structure of the upper mantle: A regional comparison of seismic layering, *Phys. Earth Planet. Inter.*, **110**, 21–41.
- Girardin, N. & Farra, V., 1998. Azimuthal anisotropy in the upper mantle from observations of *P*-to-*S* converted phases: Application to southeast Australia, *Geophys. J. Int.*, **133**, 615–629.
- Griot, D.-A., Montagner, J.-P. & Tapponnier, P., 1998. Phase velocity structure from Rayleigh and Love waves in Tibet and its neighboring regions, *J. geophys. Res.*, **103**, 21 215–21 232.
- Ho-Liu, P., Montagner, J.-P. & Kanamori, H., 1989. Comparison of iterative back-projection inversion and generalized inversion without blocks: Case studies in attenuation tomography, *Geophys. J. Int.*, **97**, 19–29.
- James, D.E., Fouch, M.J., VanDecar, J.C., van der Lee, S. & Kaapvaal Seismic Group, 2001. Tectospheric structure beneath southern Africa, *Geophys. Res. Lett.*, **28**, 2485–2488.
- Kendall, J.-M., 2000. Seismic anisotropy in the boundary layers of the mantle, in *Earth's Deep Interior: Mineral Physics and Tomography from the Atomic to the Global Scale, Geophysical Monograph*, vol. 117, eds Karato, S., Forte, A., Liebermann, R., Masters, G. & Stixrude, L., pp. 133–159, AGU, Washington, DC.
- Kennett, B.L.N., 1995. Approximations for surface-wave propagation in laterally varying media, *Geophys. J. Int.*, **122**, 470–478.
- Kennett, B.L.N. & Nolet, G., 1990. The interaction of the *S*-wavefield with upper mantle heterogeneity, *Geophys. J. Int.*, **101**, 751–762.
- Kennett, B.L.N., Guðmundsson, Ó. & Tong, C., 1994. The upper-mantle *S*-velocity and *P*-velocity structure beneath Northern Australia from broad-band observations, *Phys. Earth Planet. Inter.*, **86**, 85–98.
- Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophys. J. Int.*, **139**, 806–822.
- Lévêque, J.-J. & Cara, M., 1985. Inversion of multimode surface wave data: Evidence for sub-lithospheric anisotropy, *Geophys. J. R. astr. Soc.*, **83**, 753–773.
- Lévêque, J.-J., Rivera, L. & Wittlinger, G., 1993. On the use of the checkerboard test to assess the resolution of tomographic inversions, *Geophys. J. Int.*, **115**, 313–318.
- Lévêque, J.-J., Debayle, E. & Maupin, V., 1998. Anisotropy in the Indian Ocean upper mantle from Rayleigh- and Love-waveform inversion, *Geophys. J. Int.*, **133**, 529–540.
- Marquering, H. & Snieder, R., 1996. Shear-wave velocity structure beneath Europe, the Northeastern Atlantic and Western Asia from waveform inversions including surface-wave mode coupling, *Geophys. J. Int.*, **127**, 283–304.
- Masters, G., Johnson, S., Laske, G. & Bolton, H., 1996. A shear-velocity model of the mantle, *Phil. Trans. R. Soc. London, Ser. A*, **354**, 1385–1411.
- Meier, T. & Malischewsky, P.G., 2000. Approximation of surface wave mode conversion at a passive continental margin by a mode-matching technique, *Geophys. J. Int.*, **141**, 12–24.
- Menke, W., 1989. *Geophysical Data Analysis: Discrete Inverse Theory*, 2nd edn, Academic Press, San Diego.

- Mitchell, B.J., Baqer, S., Akinci, A. & Cong, L., 1998. L_g coda Q in Australia and its relation to crustal structure and evolution, *Pure appl. Geophys.*, **153**, 639–653.
- Montagner, J.-P., 1986. Regional three-dimensional structures using long-period surface waves, *Ann. Geophys.*, **4**, 283–294.
- Montagner, J.-P., 1996. Surface waves on a global scale, Influence of anisotropy and anelasticity, in *Seismic Modelling of Earth Structure*, eds, Boschi, E., Ekström, G. & Morelli, A., pp. 81–148, Editrice Compositori, Bologna.
- Montagner, J.-P. & Jobert, N., 1988. Vectorial tomography—II. Application to the Indian Ocean, *Geophys. J. Int.*, **94**, 309–344.
- Montagner, J.-P. & Nataf, H.-C., 1986. A simple method for inverting the azimuthal anisotropy of surface waves, *J. geophys. Res.*, **91**, 511–520.
- Montagner, J.-P. & Tanimoto, T., 1991. Global upper mantle tomography of seismic velocities and anisotropies, *J. geophys. Res.*, **96**, 20 337–20 351.
- Montagner, J.-P., Griot-Pommeroy, D.-A. & Lavé, J., 2000. How to relate body wave and surface wave anisotropy?, *J. geophys. Res.*, **105**, 19 015–19 027.
- Nishimura, C.E. & Forsyth, D.W., 1985. Anomalous Love-wave phase velocities in the Pacific: Sequential pure-path and spherical harmonic inversion, *Geophys. J. R. astr. Soc.*, **81**, 389–407.
- Nishimura, C.E. & Forsyth, D.W., 1989. The anisotropic structure of the upper mantle in the Pacific, *Geophys. J. Int.*, **96**, 203–229.
- Nolet, G., 1985. Solving or resolving inadequate and noisy tomographic systems, *J. Comput. Phys.*, **61**, 463–482.
- Nolet, G., ed., 1987. *Seismic Tomography*, Reidel, Hingham.
- Nolet, G., 1990. Partitioned waveform inversion and two-dimensional structure under the Network of Autonomously Recording Seismographs, *J. geophys. Res.*, **95**, 8499–8512.
- Nolet, G., van Trier, J. & Huisman, R., 1986. A formalism for nonlinear inversion of seismic surface waves, *Geophys. Res. Lett.*, **13**, 26–29.
- Özalaybey, S. & Chen, W.P., 1999. Frequency-dependent analysis of SKS-SKKS waveforms observed in Australia: Evidence for null birefringence, *Phys. Earth Planet. Inter.*, **114**, 197–210.
- Paige, C.C. & Saunders, M.A., 1982. LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Softw.*, **8**, 43–71.
- Peselnick, L. & Nicolas, A., 1978. Seismic anisotropy in an ophiolite peridotite. Application to oceanic upper mantle, *J. geophys. Res.*, **83**, 1227–1235.
- Plomerová, J., Liebermann, R.C. & Babuška, V., eds., 1998. *Geodynamics of Lithosphere and Earth's Mantle*, vol. 151 of Pure appl. Geophys., Birkhäuser, Boston, Mass.
- Saltzer, R.L., Gaherty, J.B. & Jordan, T.H., 2000. How are vertical shear-wave splitting measurements affected by variations in the orientation of azimuthal anisotropy with depth?, *Geophys. J. Int.*, **141**, 374–390.
- Savage, M.K., 1999. Seismic anisotropy and mantle deformation: What have we learned from shear wave splitting?, *Rev. Geophys.*, **37**, 65–106.
- Scales, J.A. & Snieder, R., 1997. To Bayes or not to Bayes?, *Geophysics*, **62**, 1045–1046.
- Shaw, R.D., Wellman, P., Gunn, P., Whitaker, A.J., Tarlowski, C. & Morse, M.P., 1995. Australian crustal elements map, *AGSO Res. Newslett.*, **23**, 1–3.
- Shibutani, T., Sambridge, M. & Kennett, B.L.N., 1996. Genetic algorithm inversion for receiver functions with application to crust and uppermost mantle structure beneath eastern Australia, *Geophys. Res. Lett.*, **23**, 1829–1832.
- Silveira, G., Stutzmann, E., Griot, D.-A., Montagner, J.-P. & Victor, L.M., 1998. Anisotropic tomography of the Atlantic Ocean from Rayleigh surface waves, *Phys. Earth Planet. Inter.*, **106**, 257–273.
- Silver, P.G., 1996. Seismic anisotropy beneath the continents: Probing the depths of geology, *Annu. Rev. Earth Planet. Sci.*, **24**, 385–432.
- Silver, P.G. & Savage, M.K., 1994. The interpretation of shear-wave splitting parameters in the presence of two anisotropic layers, *Geophys. J. Int.*, **119**, 949–963.
- Simons, F.J. & van der Hilst, R.D., Age-dependent seismic thickness and mechanical strength of the Australian lithosphere, *Geophys. Res. Lett.*, in **29**, 1529, DOI: 10.1029/2002GLO14962.
- Simons, F.J., Zielhuis, A. & van der Hilst, R.D., 1999. The deep structure of the Australian continent from surface-wave tomography, *Lithos*, **48**, 17–43.
- Strang, G. & Nguyen, T., 1997. *Wavelets and Filter Banks*, 2nd edn, Wellesley-Cambridge Press, Wellesley.
- Su, W.-j., Woodward, R. & Diewonski, A.M., 1994. Degree-12 model of shear velocity heterogeneity in the mantle, *J. geophys. Res.*, **99**, 6945–6980.
- Takeuchi, H. & Saito, M., 1972. Seismic surface waves, in *Seismology: Surface Waves and Earth Oscillations*, ed., Bolt, B.A., pp. 217–295, Academic Press, San Diego.
- Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys.*, **20**, 219–232.
- Tarantola, A. & Valette, B., 1982. Inverse problems = Quest for information, *J. Geoph.*, **50**, 159–170.
- Tong, C., Guðmundsson, Ó. & Kennett, B. L. N., 1994. Shear wave splitting in refracted waves returned from the upper mantle transition zone beneath northern Australia, *J. geophys. Res.*, **99**, 15 783–15 797.
- Trampert, J. & Woodhouse, J.H., 1996. High resolution global phase velocity distributions, *Geophys. Res. Lett.*, **23**, 21–24.
- van der Hilst, R.D., Kennett, B.L.N., Christie, D. & Grant, J., 1994. Project SKIPPY explores the lithosphere and mantle beneath Australia, *Eos Trans. AGU*, **75**, 177–181.
- van der Sluis, A. & van der Vorst, H.A., 1987. Numerical solution of large sparse linear algebraic systems arising from tomographic problems, in *Seismic Tomography*, ed., Nolet, G., Chap. 3, pp. 49–83, Reidel, Hingham, Mass.
- Vinnik, L.P., Farra, V. & Romanowicz, B., 1989. Azimuthal anisotropy in the Earth from observations of SKS at GEOSCOPE and NARS broad-band stations, *Geophys. J. Int.*, **79**, 1542–1558.
- Vinnik, L.P., Makeyeva, L.I., Milev, A. & Usenko, A.Y., 1992. Global patterns of azimuthal anisotropy and deformations in the continental mantle, *Geophys. J. Int.*, **111**, 433–447.
- Yanovskaya, T.B. & Ditmar, P.G., 1990. Smoothness criteria in surface wave tomography, *Geophys. J. Int.*, **102**, 63–72.
- Zielhuis, A., 1992. *S-Wave Velocity Below Europe from Delay-Time and Waveform Inversions*, Ph.D. thesis, Utrecht Univ., Utrecht.
- Zielhuis, A. & Nolet, G., 1994. Shear-wave velocity variations in the upper mantle beneath central Europe, *Geophys. J. Int.*, **117**, 695–715.
- Zielhuis, A. & van der Hilst, R.D., 1996. Upper-mantle shear velocity beneath eastern Australia from inversion of waveforms from SKIPPY portable arrays, *Geophys. J. Int.*, **127**, 1–16.