Title

A Bayesian method to quantify azimuthal anisotropy model uncertainties: application to global azimuthal anisotropy in the upper mantle and transition zone

Authors

Yuan, K
Beghein, C

A Bayesian method to quantify azimuthal anisotropy model uncertainties: application to global azimuthal anisotropy in the upper mantle and transition zone

K. Yuan* and C. Beghein

Department of Earth, Planetary, and Space Sciences, University of California Los Angeles, Los Angeles, CA 90095, USA. E-mail: cbeghein@ucla.edu

SUMMARY

Seismic anisotropy is a powerful tool to constrain mantle deformation, but its existence in the deep upper mantle and topmost lower mantle is still uncertain. Recent results from higher mode Rayleigh waves have, however, revealed the presence of 1 per cent azimuthal anisotropy between 300 and 800 km depth, and changes in azimuthal anisotropy across the mantle transition zone boundaries. This has important consequences for our understanding of mantle convection patterns and deformation of deep mantle material. Here, we propose a Bayesian method to model depth variations in azimuthal anisotropy and to obtain quantitative uncertainties on the fast seismic direction and anisotropy amplitude from phase velocity dispersion maps. We applied this new method to existing global fundamental and higher mode Rayleigh wave phase velocity maps to assess the likelihood of azimuthal anisotropy in the deep upper mantle and to determine whether previously detected changes in anisotropy at the transition zone boundaries are robustly constrained by those data. Our results confirm that deep upper-mantle azimuthal anisotropy is favoured and well constrained by the higher mode data employed. The fast seismic directions are in agreement with our previously published model. The data favour a model characterized, on average, by changes in azimuthal anisotropy at the top and bottom of the transition zone. However, this change in fast axes is not a global feature as there are regions of the model where the azimuthal anisotropy direction is unlikely to change across depths in the deep upper mantle. We were, however, unable to detect any clear pattern or connection with surface tectonics. Future studies will be needed to further improve the lateral resolution of this type of model at transition zone depths.

Key words: Inverse theory; Probability distributions; Seismic anisotropy; Seismic tomography; Statistical seismology; Surface waves and free oscillations.

1 INTRODUCTION

The directional and polarization dependence of seismic wave velocity, or seismic anisotropy, is a powerful tool to investigate mantle deformation and geodynamics (Montagner 1994; Karato 1998; Becker *et al.* 2003; Long 2013). The lattice-preferred orientation (LPO) of the crystallographic axes of elastically anisotropic material is generally assumed to be the cause of the seismic anisotropy detected in Earth's mantle, though it could alternatively be caused by the shape-preferred orientation (SPO) of isotropic structures with contrasting elastic properties such as cracks, layered structures, melt tubules, or lenses (Kendall & Silver 1996; Montagner 1994). In the mantle lithosphere, "frozen-in" seismic anisotropy is often attributed to olivine LPO related to past tectonic processes (Karato 1989;

Nicolas & Christensen 1987; Silver 1996). In the asthenosphere, olivine LPO associated with present-day mantle deformation is often invoked to explain observations of seismic anisotropy because the fast seismic direction generally aligns with the absolute plate motion (Nishimura & Forsyth 1988; Smith *et al.* 2004; Debayle *et al.* 2005; Marone & Romanowicz 2007; Beghein *et al.* 2014), and the preferred alignment of olivine can be used to determine the direction of mantle flow (Becker *et al.* 2003). In the lowermost mantle, both SPO through horizontal layering or aligned inclusions (Kendall & Silver 1996) and LPO of the post-perovskite phase (Oganov 2005) have been proposed to explain observations of anisotropy.

Most tomographic models of seismic anisotropy are obtained by regularized inversion of seismic data such as surface waves, free oscillations, and/or long-period body waves, and they all provide ample evidence for the presence of seismic anisotropy in the uppermost 250 km of the mantle. Radial anisotropy, which quantifies differences in seismic wave velocity between the vertical and horizontal directions, is required in the uppermost mantle to

* Formerly at: Department of Earth, Planetary, and Space Sciences, University of California Los Angeles, Los Angeles, CA 90095, USA.

simultaneously explain Love and Rayleigh wave dispersion data. It is included in the top 220 km of the 1-D Preliminary Reference Earth Model (PREM) of Dziewonski & Anderson (1981), and in several 3-D radially anisotropic global models of the uppermost mantle models (see Chang *et al.* 2014, for a recent review). Azimuthal anisotropy, that is, the dependence of seismic wave velocities with the azimuth of propagation, is also present in the uppermost mantle at the global scale (Montagner & Tanimoto 1991; Trampert & Woodhouse 2003; Debayle *et al.* 2005; Ekström 2011; Debayle & Ricard 2013; Yuan & Beghein 2013; Becker *et al.* 2014; Yuan & Beghein 2014; Schaeffer *et al.* 2016). Most global models of azimuthal anisotropy display common features at these depths, such as the alignment of the fast axes with the plate motion direction at asthenospheric depths beneath ocean basins and with the palaeospreading directions in the oceanic lithosphere. The 1-D models of radial anisotropy generally agree with one another, but there are discrepancies in models of lateral variations in radial anisotropy even at shallow depths (Chang *et al.* 2014).

The D'' layer is also known to be radially anisotropic: at the regional scale, radial anisotropy has been observed with shear wave splitting measurements (Kendall & Silver 1996), and azimuthal anisotropy has been detected with *S* and *S*diff waveform modeling (Maupin *et al.* 2005). A few global tomographic models suggest D'' radial anisotropy is present at the global scale as well, though the effect of the crustal correction (Panning *et al.* 2010), of prior scaling relationships between elastic parameters (Beghein & Trampert 2004a,b; Beghein *et al.* 2006; Beghein 2010), and trade-offs between isotropic and anisotropic structures (Kustowski *et al.* 2008; Chang *et al.* 2014) cast doubt on the global nature of radial anisotropy at these depths. To date, there is no global azimuthal anisotropy model of the lowermost mantle.

For years, the lack of evidence for seismic anisotropy below ~250 km depth was interpreted as the result of deformation by diffusion creep (Karato *et al.* 1995). Evidence for radial anisotropy in the deep upper mantle and uppermost lower mantle has, however, been accumulating over the past two decades. The first global model displaying radial anisotropy in the deep upper mantle was the 1-D model of Montagner & Kennett (1996), which was followed by multiple 1-D and 3-D global radial anisotropy models (Beghein & Trampert 2004b; Panning & Romanowicz 2004; Beghein *et al.* 2006; Panning & Romanowicz 2006; Kustowski *et al.* 2008; Visser *et al.* 2008a; Panning *et al.* 2010; Romanowicz & Lekić 2011; Auer *et al.* 2014; Chang *et al.* 2014; French & Romanowicz 2014; Moulik & Ekström 2014). Shear wave splitting studies have also suggested the presence of radial anisotropy near subduction zone in the mantle transition zone (MTZ) and top of the lower mantle (Fouch & Fischer 1996; Wookey & Barruol 2002; Chen & Brudzinski 2003; Foley & Long 2011; Lynner 2015; Nowacki *et al.* 2015). Azimuthal anisotropy may additionally be present at these depths. It has been shown to be compatible with higher mode Love waves (Trampert & van Heijst 2002) and coupled free oscillation data (Beghein *et al.* 2008; Hu *et al.* 2012). A study combining data from surface waves and shear wave splitting beneath North America also suggested that azimuthal anisotropy is needed at greater depths than commonly assumed (Marone & Romanowicz 2007), and similar conclusions were drawn by Kosarian *et al.* (2011) for California. At the global scale, while early 3-D models did not show any significant azimuthal anisotropy below 250 km depth (Montagner & Tanimoto 1991; Debayle *et al.* 2005), more recent studies present about 1 per cent anisotropy in the deep upper mantle at least down to 400 km (Debayle & Ricard 2013; Yuan & Beghein 2013, 2014; Schaeffer *et al.* 2016), and possibly even deeper down to the bottom of the

MTZ and top of the lower mantle (Yuan & Beghein 2013, 2014). There are still large discrepancies among models, but the increasing evidence for seismic anisotropy in the deep upper mantle challenges our understanding of mantle deformation (Fig. 1).

Our previously published global azimuthal anisotropy model (Yuan & Beghein 2013), hereafter referred to as YB13SVani, not only displayed a non-negligible amount of azimuthal anisotropy (1–3 per cent) below 250 km depth, but it also revealed changes in the seismic fast direction at the MTZ boundaries. The interpretation of these results is non-unique due to the paucity of mineral physics data on MTZ material anisotropy. Nevertheless, they have important consequences for our understanding of mantle convection and the anisotropy of deep upper-mantle material as it could imply changes in mantle flow direction at the MTZ, changes in volatile content, in slip system in MTZ material, etc. It is thus essential to determine which model features are robust.

In this paper, we present a Bayesian forward modeling method to quantify uncertainties and trade-offs of azimuthal anisotropy model parameters. Like most tomographic models, YB13SVani was obtained by regularized inversion of seismic data. In this particular case, the model was derived from fundamental and higher mode surface wave phase velocity maps. Estimating reliable model uncertainties from linear inversions is, however, not straightforward since most inversions yield a posterior model covariance smaller or equal to the prior covariance by construction (Tarantola 1987). If there is a large model null space, the posterior covariance can be strongly underestimated (Trampert 1998), making both the interpretation and the uncertainty assessment of tomographic models difficult (Beghein & Trampert 2003; Beghein 2010). Model space search approaches are generally better suited to determine posterior model uncertainties as they can explore a larger part of the model space, including the null space, and map the range of models that can fit the data reasonably well. In some cases, this type of method can even find solutions to the problem that could not be found with traditional inverse methods (Beghein & Trampert 2003). In this paper, we modeled azimuthal anisotropy in the upper mantle and topmost lower mantle and quantified parameter uncertainties and trade-offs using the Neighbourhood Algorithm (Sambridge 1999a,b), hereafter referred to as the NA. Among many other applications, this direct search technique has been used successfully to model inner core anisotropy (Beghein & Trampert 2003), regional and global mantle seismic velocities (Beghein *et al.* 2002; Snoke & Sambridge 2002) and radial anisotropy (Beghein & Trampert 2004a,b; Beghein *et al.* 2006; Visser *et al.* 2008a; Yao *et al.* 2008; Beghein 2010). Yao (2015) recently proposed a two-step method using the NA to model azimuthal anisotropy from fundamental-mode surface waves. However, as explained in Section 3, the author did not display or discuss the posterior uncertainties on the fast seismic direction and the anisotropy amplitude. The method we present here solves the linear problem that relates azimuthal anisotropy elastic parameters to phase velocities using laterally varying sensitivity kernels to account for variations in crustal structure, and quantifies model uncertainties for the azimuthal anisotropy amplitude and the fast axes directions.

2 PHASE VELOCITY DATA

2.1 Phase velocity anisotropy

In this study, we employed the same data set as we did to construct YB13SVani (Yuan & Beghein 2013). It consists in the fundamental
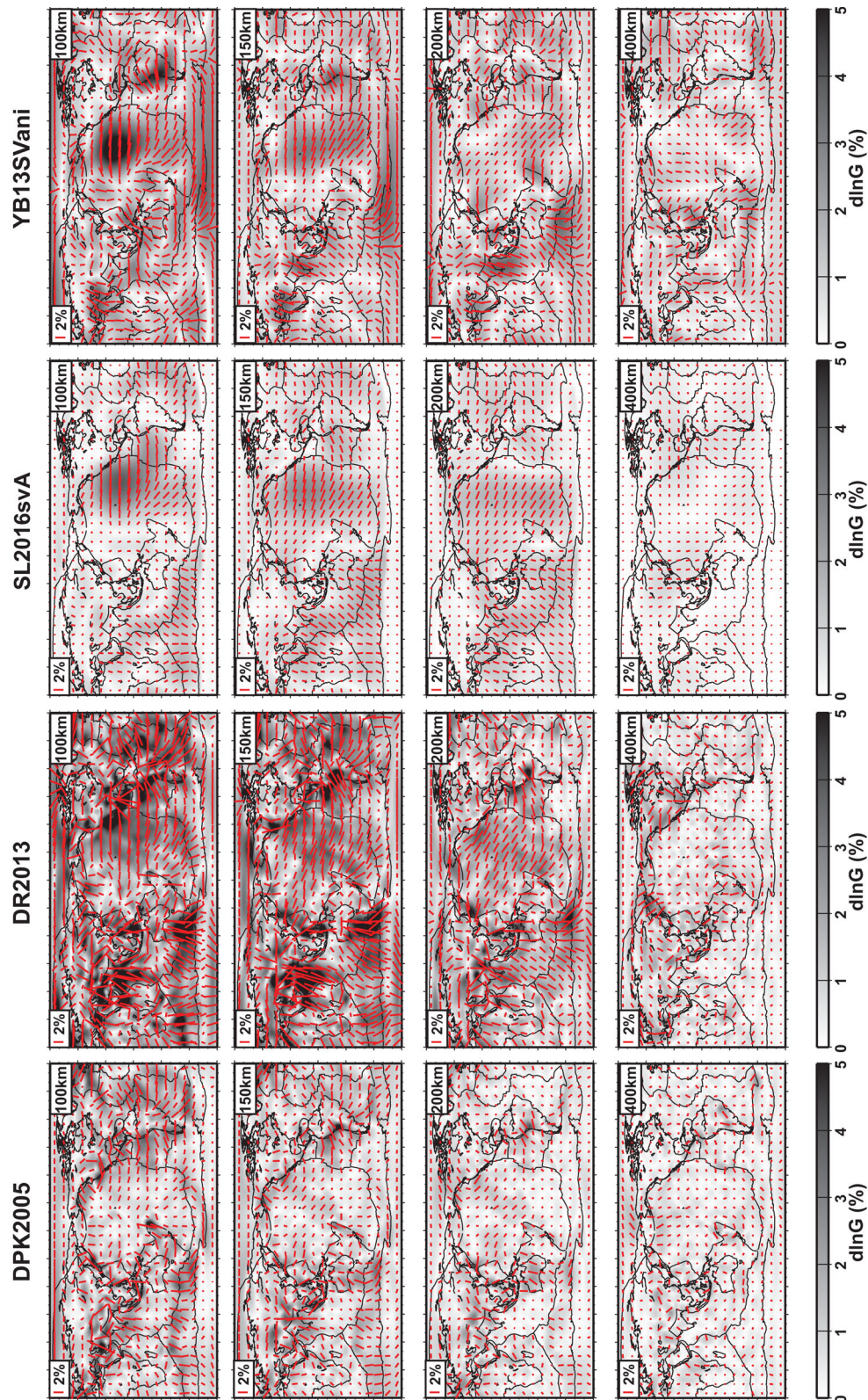
Figure 1. Azimuthal anisotropy models in the top 400 km of the mantle from previous studies: DPK2005 (Debayle *et al.* 2005), DR2013 (Debayle & Ricard 2013), SL2016svA (Schaeffer *et al.* 2016) and YB13SVani (Yuan & Beghein 2013).

and first six higher mode anisotropic Rayleigh wave phase velocity maps determined by Visser *et al.* (2008b). There were 16 fundamental modes between 35 and 175 s, 16 first overtones between 35 and 172 s, 15 second overtones between 35 and 150 s, 11 third overtones between 35 and 88 s, 8 fourth overtones between 35 and 62 s, 7 fifth overtones between 35 and 56 s, and 6 sixth overtones between 35 and 51 s. This type of seismic data are ideal to provide depth constraints on Earth's internal structure because of their dispersive properties (Fig. 2). In addition, combining fundamental and higher mode surface wave data significantly increases the depth resolution of tomographic models. Contrary to fundamental-mode surface waves, which can only resolve the top ∼200–300 km of the mantle, the set of higher modes employed here have sensitivity to azimuthal anisotropy well into the deep upper mantle and topmost lower mantle (Fig. 5).

At any given point at Earth's surface, perturbations dc in surface wave phase velocity with respect to predictions from a reference Earth model can be expressed as a function of the azimuth of propagation Ψ as follows (Montagner & Nataf 1986):

$$dc(T, \Psi) = dc_0(T) + dc_1(T)\cos(2\Psi) + dc_2(T)\sin(2\Psi)$$
$$+ dc_3(T)\cos(4\Psi) + dc_4(T)\sin(4\Psi) \qquad (1)$$

where T is the period of the wave. dc_0 represents the phase velocity anomaly averaged over all azimuths and the dc_i ($i = 1, ..., 4$) terms represent the azimuthal dependence of the phase velocity. The 2Ψ terms can help constrain depth variations in elastic parameters that relate to the azimuthal anisotropy of vertically polarized shear (*SV*) wave, as explained below with eqs (2) and (3), and the 4Ψ terms can help determine the depth dependence of horizontally polarized shear (*SH*) wave azimuthal anisotropy [see Yuan and Beghein (Yuan & Beghein 2014) for details]. Eq. (1) is valid for fundamental and higher mode surface waves.

The relation between 2Ψ phase velocity anisotropy and azimuthal anisotropy at depth is given by the following set of equations (Montagner & Nataf 1986):

$$dc_1(T) = \int [G_c(r)K_G(T,r) + B_c(r)K_B(T,r) + H_c(r)K_H(T,r)]dr$$
$$(2)$$

$$dc_2(T) = \int [G_s(r)K_G(T,r) + B_s(r)K_B(T,r) + H_s(r)K_H(T,r)]dr$$
$$(3)$$

where elastic parameters $G_c(r)$ and $G_s(r)$ relate to V_{SV} azimuthal anisotropy, and $B_c(r)$ and $B_s(r)$ relate to *P*-wave azimuthal anisotropy. $H_s(r)$ and $H_c(r)$ do not control body wave azimuthal anisotropy and only appear in surface waves (Montagner & Nataf 1986) and in normal modes (Beghein *et al.* 2008). $K_G(r, T)$, $K_B(r, T)$, and $K_H(r, T)$ are the local partial derivatives, or sensitivity kernels, for Rayleigh wave at period T and radius r, which can be calculated for a reference model using normal-mode theory (Takeuchi 1972). The fast azimuth of propagation Θ and the anisotropy amplitude G of vertically polarized shear waves are given by:

$$\Theta = \frac{1}{2}\arctan(G_s/G_c) \qquad (4)$$

and

$$G = \sqrt{G_s^2 + G_c^2} \qquad (5)$$

Similar relations exist for $B_{c,s}$ and $H_{c,s}$. Examples of kernels calculated using model PREM (Dziewonski & Anderson 1981) are

shown in Figs 2 and 5, and demonstrate that including higher modes in the data set significantly increases and extends the sensitivity to anisotropy into the deep upper mantle.

Fundamental-mode Rayleigh waves typically are not expected to have a strong 4Ψ dependence in comparison to the 2Ψ terms, as demonstrated by Montagner & Tanimoto (1991) for realistic petrological models. The same may not, however, be true for higher modes since they are sensitive to deeper structure (Fig. 5), and indeed Visser *et al.* (2008b) determined that a 4Ψ dependence significantly improved the fit of their Rayleigh wave fundamental and higher mode phase velocity path-averaged measurements (Visser *et al.* 2008b). Nevertheless, because the sensitivity of fundamental and higher mode Rayleigh waves to *SH* anisotropy is very small, here we only used the 2Ψ terms of eq. (1) to build a 3-D model of *SV* azimuthal anisotropy in the top 1000 km of the mantle.

2.2 Phase velocity resolution

As explained by Visser *et al.* (2008b), the lateral resolution of their phase velocity models generally decreases with increasing overtone number because the quality of the path azimuthal coverage (and thus the number of modes measured reliably) decreases with the overtone number. Ray coverage was very good everywhere for the fundamental modes, and in most continental regions and the northwestern Pacific for the higher modes, but it was poorer for the third through sixth higher modes in the southeastern Pacific, southern Indian Ocean and southern Atlantic. Another factor that affected the lateral resolution of the phase velocity maps was the choice of the damping made by the authors who opted for maintaining a constant relative model uncertainty for all modes. This too resulted in phase velocity maps of decreasing resolution with increasing overtone number. Visser *et al.* (2008b) estimated that the fundamental-mode 2Ψ terms are resolved up to spherical harmonic degrees 8 and 5 for the higher modes, which corresponds to a resolving power of about 4500 km near the surface, decreasing to 6500 km near the MTZ. Because the inferences made in this paper focus on large-scale anisotropy, using data of varying resolution should not strongly affect our results. Trade-offs between the different terms of eq. (1) constitute another source of uncertainty when constructing anisotropic phase velocity maps from path-averaged measurements. One cannot completely separate the different terms because data coverage is imperfect owing to the uneven distribution of earthquakes and seismic stations over the globe. The resolution matrices calculated by Visser *et al.* (2008b) showed that these trade-offs were minimal and that there was therefore little mapping of lateral heterogeneities or topography at discontinuities into the anisotropic terms, though one should of course always keep in mind that trade-offs are not completely inexistent.

3 METHOD

3.1 Parametrization

We divided Earth's surface into $10° × 10°$ cells and the data were inverted by applying the NA to eqs (2) and (3) at each grid cell. The reader should note, however, that $10°$ does not correspond to the lateral resolution of our models since it is directly controlled by the resolution of the phase velocity maps and is limited to larger wavelengths in the deep upper mantle than in the shallow mantle, as discussed in Section 2. It is also important to note that the quantitative uncertainty analysis performed in this study does not account
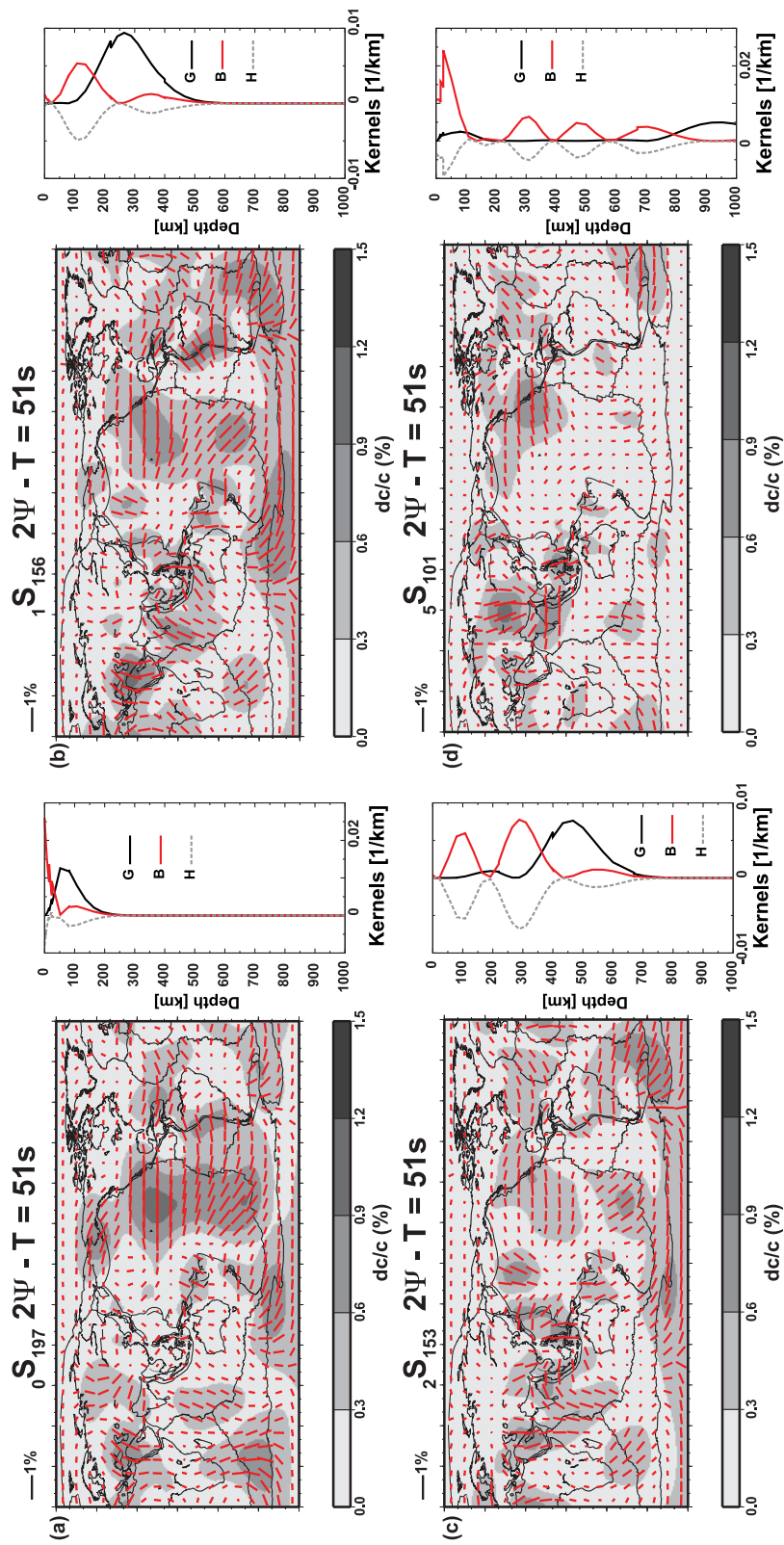
Figure 2. 2Ψ azimuthally anisotropic Rayleigh wave phase velocity maps (Visser *et al.* 2008b) at 51 s period for the fundamental mode (a), the first (b), second (c) and fifth (d) overtone, and associated partial derivatives. The sensitivity kernels were calculated using reference model PREM (Dziewonski & Anderson 1981) for relative perturbations in parameters *G*, *B* and *H*.
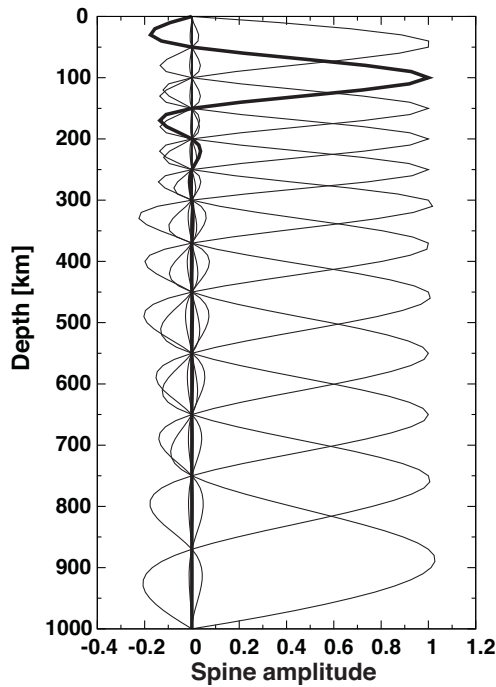
Figure 3. Splines functions employed to parametrize $G_c(r)$ and $G_s(r)$.

for uncertainties stemming from the non-uniform ray path coverage, but is solely focused on the model parameter resolution for a given set of dispersion curves and estimated data uncertainties. While it would be interesting to investigate the effect of the non-uniform coverage, in particular the variations from one mode to another and one frequency to another, it is beyond the scope of this paper.

At every grid cell we parameterized $G_c(r)$ and $G_s(r)$ vertically using 12 cubic spline functions $S_i(r)$ ($i = 1,...,12$) of varying depth spacing (Fig. 3):

$$G_c(r) = \sum_{i=1}^{12} G_c^i S_i(r) \qquad (6)$$

$$G_s(r) = \sum_{i=1}^{12} G_s^i S_i(r) \qquad (7)$$

Parameters $B_c(r)$, $B_s(r)$, $H_c(r)$, and $H_s(r)$ are poorly resolved due to the similarity of their partial derivatives (Fig. 2), and we therefore opted to neglect them and invert for $G_{c,s}$ only. Most previous authors have either neglected these parameters in surface wave inversions (Marone & Romanowicz 2007) or assumed to be proportional to $G_{c,s}$ (Yao 2015). Such assumptions also enable us to run the NA more efficiently because increasing the number of unknowns quickly raises the computation cost of a model space search. Fig. S1 in the Supporting Information shows that neglecting the *P*-wave related parameters does not strongly affect the results for $G_c(r)$ or $G_s(r)$. Under this assumption, eqs (2) and (3) become:

$$dc_1(T) = \sum_{i=1}^{12} G_c^i I_i(T) \qquad (8)$$

$$dc_2(T) = \sum_{i=1}^{12} G_s^i I_i(T) \qquad (9)$$

where

$$I_i(T) = \int S_i(r) K_G(T, r) \, dr \qquad (10)$$

In this work, we used a parametrization in terms of relative perturbations $d\ln G_{c,s} = G_{c,s}/L$, where L is one of the so-called Love elastic parameter (Love 1927), which determines the wave speed of vertically polarized shear waves ($V_{SV} = \sqrt{L/\rho}$). Perturbations are expressed with respect to a local reference model composed of CRUST2.0 (Bassin *et al.* 2000) and PREM (Dziewonski & Anderson 1981) at each grid cell, as explained in Section 3.2.

It should be noted that the spline functions used in this study differ slightly from the ones employed to obtain model YB13SVani (Yuan & Beghein 2013). We therefore cannot fairly compare YB13SVani with our new model resulting from the NA. Thus, in this paper, in addition to presenting the results of a model space search approach (see Section 3.3), we display an updated 3-D model obtained using the new splines described above together with the same data set and singular value decomposition (SVD) method as in Yuan & Beghein (2013). This new model, hereafter referred to as YB17SVaniSVD, is almost identical and displays the same features as YB13SVani. Fig. 4 shows that the two models are well correlated with one another and they present similar anisotropy amplitudes. Both models display peaks and minima in the root mean square (rms) amplitude and peaks in the gradient of the fast axes at the same depths.

3.2 Effect of the crust

An important aspect of modeling lateral heterogeneities or anisotropy in the mantle relates to crustal structure. Many first generation 3-D velocity and anisotropy models were obtained using sensitivity kernels calculated based on the 1-D reference mantle model PREM (Dziewonski & Anderson 1981). However, crustal thickness, velocities and density vary laterally, and neglecting these variations can bias the model due to the mapping of crustal structure into the mantle (Boschi & Ekström 2002; Marone & Romanowicz 2007; Kustowski *et al.* 2007; Bozdağ & Trampert 2010).

When inverting surface wave data for mantle velocities or anisotropy, it is essential to account for the effect of lateral crustal variations on the sensitivity kernels (Boschi & Ekström 2002; Marone & Romanowicz 2007) and to either correct the data with an *a priori* crustal model (Boschi & Ekström 2002) or invert the data simultaneously for the Moho depth, crustal structure and mantle structure (Meier *et al.* 2007; Visser *et al.* 2008a; Chang *et al.* 2014). It should, however, be noted that in this last case data uncertainties need to be small to resolve the Moho depth due to trade-offs with velocities (Lebedev *et al.* 2013). Similarly, when inverting the azimuthally anisotropic part of phase velocity data, one would ideally be able to correct the data for azimuthal anisotropy in the crust or invert the data simultaneously for crust and mantle azimuthal anisotropy. However, to this day there exists no global azimuthal anisotropy model of the crust that we can use to correct the data *a priori*, and because the data used here have little sensitivity to crustal depths, they are likely not sufficient to resolve azimuthal anisotropy in the crust (Fig. 5).

We thus used a depth parametrization that averages azimuthal anisotropy in the crust and uppermost part of the mantle (Fig. 3), and we accounted for the effect of crustal structure and variations in Moho depth on the $d\ln G_{c,s}$ partial derivatives following Yuan & Beghein (2013). More specifically, we generated a local 1-D reference model composed of the PREM mantle to which we superimposed crustal model CRUST2.0 (Bassin *et al.* 2000) at each grid cell, and calculated the corresponding partial derivatives (Takeuchi 1972). We refer to Yuan & Beghein (2013) for examples of laterally varying sensitivity kernels. The approach taken here is slightly
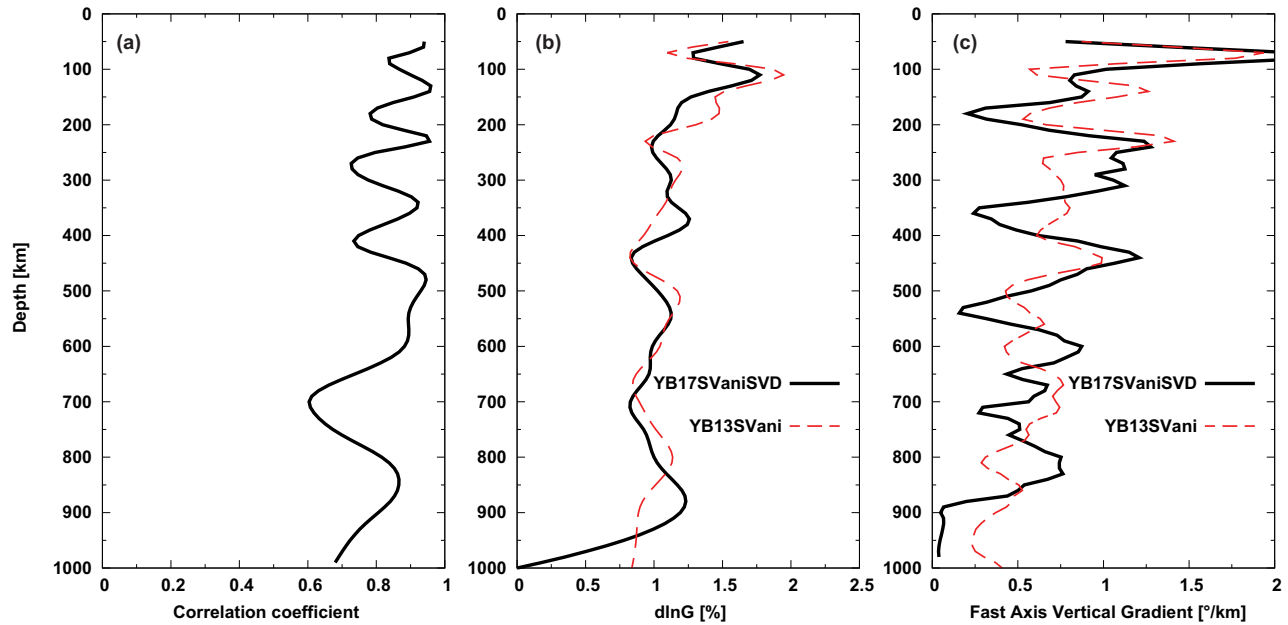
Figure 4. (a) Correlation coefficient between new model YB17SVaniSVD and model YB13SVani (Yuan & Beghein 2013); (b) root mean square amplitude and (c) gradient of the fast axes direction of models YB17SVaniSVD and YB13SVani.
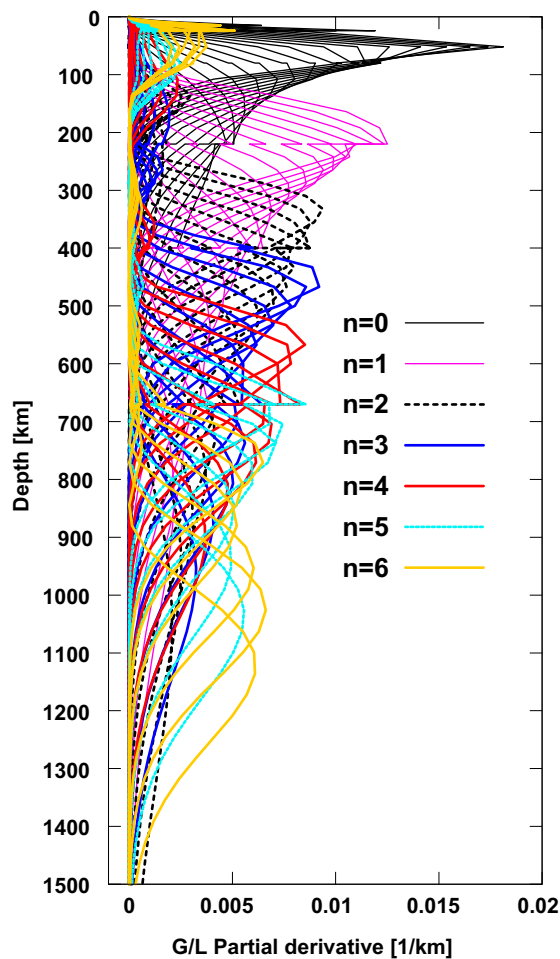
Figure 5. Phase velocity partial derivatives for relative perturbations in vertically polarized shear wave azimuthal anisotropy. These sensitivity kernels were calculated using model PREM for the fundamental modes and first six higher modes employed in this study.

different from that of Yao (2015) who used the isotropic part of the phase velocity maps together with the NA to generate a new local 1-D mantle model. However, we do not expect this to strongly influence our results since it was demonstrated that accounting for lateral variations in mantle structure to calculate the sensitivity kernels does not yield any significant difference in the 3-D azimuthal anisotropy model (Yuan & Beghein 2013).

3.3 The Neighbourhood Algorithm

Model space search techniques are most often applied to non-linear problems, which can be highly non-unique and can have a non-Gaussian cost function with multiple minima. In that case, the solution obtained by traditional inverse techniques is strongly dependent on prior assumptions and regularization. Forward modeling methods offer a more robust way to solve non-linear problems. They are also useful to solve linear problems, since these do not necessarily have Gaussian model parameter distributions (Beghein 2010).

The NA (Sambridge 1999a,b) is a guided Monte Carlo search technique that identifies regions of relatively low and relatively high misfit, associated with high and low likelihoods, respectively. For a given parametrization and cost function, if the boundaries of the model space are wide enough, it allows us to map a larger part of the model space (within these selected boundaries) than a damped inversion. In inverse theory, one usually assumes that the prior information on both model and data covariances follow a Gaussian, which implies a Gaussian distribution of the posterior model covariance (Tarantola 1987). Model space searches, however, enable the user to map the model null space and to obtain information on the model space approximate topology without having to introduce explicit regularization on the model parameters (e.g. assuming Gaussian prior model distributions) other than the imposed parametrization and the chosen boundaries of the model space being explored. One should also keep in mind that because the imposed range within which we search the parameters is a form

of prior information, if this range is very small, it is equivalent to imposing a strong damping. This type of method requires therefore a compromise between efficiency of the model space search and thoroughness of the model space search. In cases like this one, where linearized perturbation theory lies behind the equation employed, one also has to be careful to not sample too wide of a model space which could break the conditions of application of the theory.

The NA is composed of two stages. During the first stage, the model space is sampled randomly, and a cost function is calculated to determine how well each model explains the data. At each iteration, the number of models generated increases in the vicinity of the best-fitting regions of the model space. The first stage of the NA differs from many other Monte Carlo techniques in that its objective is not to locate a single optimal model, but to obtain an overview of the model space. It therefore keeps track of all the models generated instead of discarding the worse data-fitting models at each iteration.

The cost function $\phi(m)$ employed in this study to drive the sampling is defined as:

$$\phi(m) = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{d_i - (Am)_i}{\sigma_i} \right)^2 \tag{11}$$

where N is the total number of data, m_i is the ith component of the model vector $\mathbf{m}$ generated by the NA, d_i is the ith component of the data vector $\mathbf{d}$ and $(Am)_i$ is the ith component of the vector $\mathbf{Am}$ containing the data predictions calculated using eqs (2) or (3) and the sensitivity kernels projected onto the spline functions. σ_i is the error in the phase velocity maps estimated by Visser *et al.* (2008b). Note that eq. (11) assumes that data uncertainties follow Gaussian distributions. In addition, it is good to remind the reader that these standard deviations result from global inversions of path-averaged phase velocity measurements and are therefore, in fact, posterior errors on the phase velocity dispersion curves that we use here to build a prior data covariance matrix.

In this work, we solved the problem for dc_1 and dc_2 separately at each grid cell, that is, we ran the NA $36 \times 18 = 648$ times for dc_1 and 648 times for dc_2, searching the model space for 12 spline parameters each time. Each model parameter ($d\ln G_{c,s}^i$ for $i = 1, \ldots, 12$) was allowed to vary uniformly between -0.03 and $+0.03$ around model YB17SVaniSVD, which resulted from a regularized inversion as explained in Section 3.1. This range was selected to allow most parameters to change sign if required by the data. We tested that running the NA around PREM (for which all $d\ln G_{c,s}^i$ are zero) does not affect the outcome of the model space search provided convergence is achieved and a broad enough model space search is performed (Figs S2 and S3, Supporting Information). Running the model exploration around YB17SVaniSVD was, however, more computationally efficient because the sampling started from a reasonably good data-fitting model, enabling faster convergence in cases where the model space topography was approximately Gaussian. Nevertheless, doing so did not prevent the NA to find other solutions, away from the starting model, because we insured the model space search was broad and thorough using multiple tests and settings in the NA algorithm.

In a second stage, a Bayesian appraisal of all the models is performed. Unlike other statistical techniques, such as importance sampling, that draw inferences on the models using only a subset of the ensemble of models generated, the NA makes use of all the models, good and bad, generated during the first stage. As pointed out by Sambridge (1999b), in some cases one can learn from the models that fit the data poorly as much as from those that fit the data

well. The entire family of models obtained in the first stage is thus converted into posterior probability density functions (PPDFs) by associating the relatively low- and high-misfit values to high and low likelihoods, respectively. These PPDFs can be used to assess the robustness of the model parameters.

For a PPDF denoted by $P(\mathbf{m})$, where $\mathbf{m}$ is a point in the model space, the posterior mean model for the ith parameter is given by the following integral performed over the model space (Sambridge 1999b):

$$< m_i > = \int m_i P(\mathbf{m}) \, d\mathbf{m} \tag{12}$$

The posterior variances of the model parameters can be obtained from the diagonals of the posterior model covariance matrix given by:

$$C_{i,j} = \int m_i m_j P(\mathbf{m}) \, d(\mathbf{m}) - < m_i > < m_j > \tag{13}$$

Because the model space, including the null space, was sampled, the model uncertainties inferred are more accurate than those resulting from regularized inversions. Those result from the local curvature of the cost function around a model chosen with an explicit regularization and assuming Gaussian statistics. However, if the underlying statistics are not Gaussian or if the cost function has a wide valley (e.g. if the null space is large), error estimates from regularized inversions underestimate the posterior model variance (Trampert 1998; Beghein & Trampert 2003; Beghein 2010). An example of model uncertainties estimated with the NA compared to those resulting from an inversion by SVD (Menke 2012) is shown in Fig. S4 in the Supporting Information.

The 1-D marginal distribution of a given model parameter m_i can be obtained by integrating $P(\mathbf{m})$ numerically over all other parameters (Sambridge 1999b):

$$M(m_i) = \int \ldots \int P(\mathbf{m}) \prod_{k=1, k \neq i}^{d} dm_k \tag{14}$$

where d is the total number of model parameters. The shape and width of these 1-D marginals provide useful information on how well constrained a given parameter is and whether the model distribution is Gaussian, in which case the mean $<m_i>$ coincides with the peak of the distribution, that is, the most likely value. Information about parameter trade-offs can be obtained from the diagonal terms of the posterior covariance matrix, and from the 2-D marginal distributions, which are calculated by integrating $P(\mathbf{m})$ over all but two parameters. The 2-D marginal PPDF for the ith and jth parameters is given by:

$$M(m_i, m_j) = \int \ldots \int P(\mathbf{m}) \prod_{k=1, k \neq i, k \neq j}^{d} dm_k \tag{15}$$

Fig. 6 displays a representative example of 1-D and 2-D marginals. No trade-off is visible between the spline parameters displayed, and we checked that this was the case for other pairs of G_c and G_s parameters at several grid cells. We note that the posterior model parameter distributions are not necessarily Gaussian. Posterior model distributions result from the product of the *a priori* probability density in the model space and the probability density describing the result of the measurements (Tarantola 1987):

$$\sigma_M(\mathbf{m}) = k\rho_M(\mathbf{m})L(\mathbf{m}) \tag{16}$$

where k is a normalization constant, $\sigma_M(\mathbf{m})$ is the PPDF for model $\mathbf{m}$ from which one may calculate the probability for a model to sat-
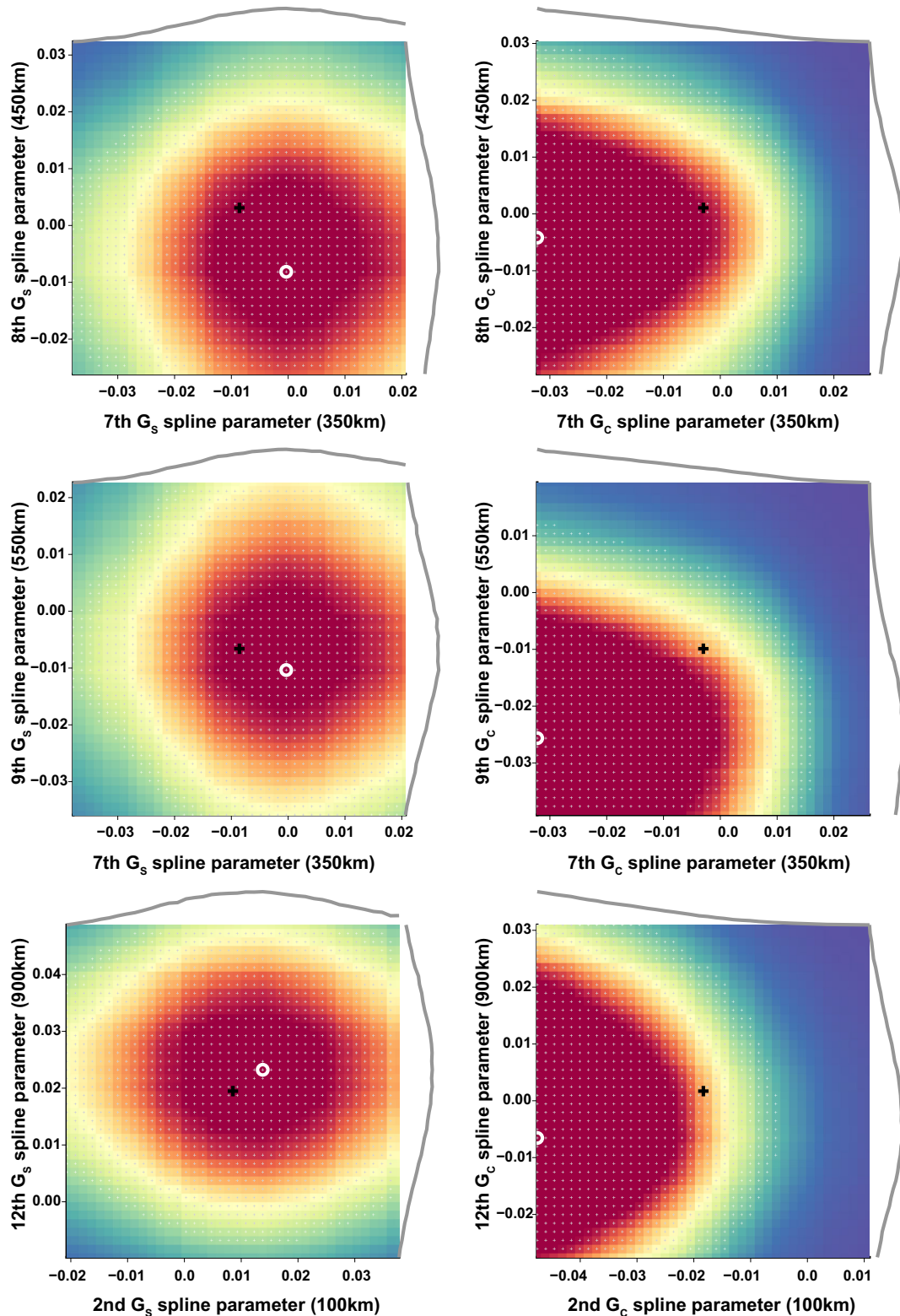
Figure 6. Examples of 1-D and 2-D marginal distributions at the grid cell located at $-55°$ longitude and $-60°$ latitude. The 1-D marginals are represented by the thick grey curves. The black cross indicates the location of the regularized inversion result around which the model space search was performed. The white circle is for the peak of the 2-D PPDF.

isfy some characteristic and $\rho_M(\mathbf{m})$ is the prior model distribution. $L(\mathbf{m}) = \rho_D(\mathbf{g}(\mathbf{m}))$ is the likelihood function that measures how well a model $\mathbf{m}$ explains the data, with ρ_D representing the (prior) information on the data and $\mathbf{g}(\mathbf{m})$ is the forward operator that represents the mathematical model of the physical system under study. When $\mathbf{g}(\mathbf{m})$ is linear ($\mathbf{g}(\mathbf{m}) = \mathbf{Gm}$) and both the prior on the model parameters and the measurements are normal distributed, the posterior model distribution follows a Gaussian as well. However, in the case
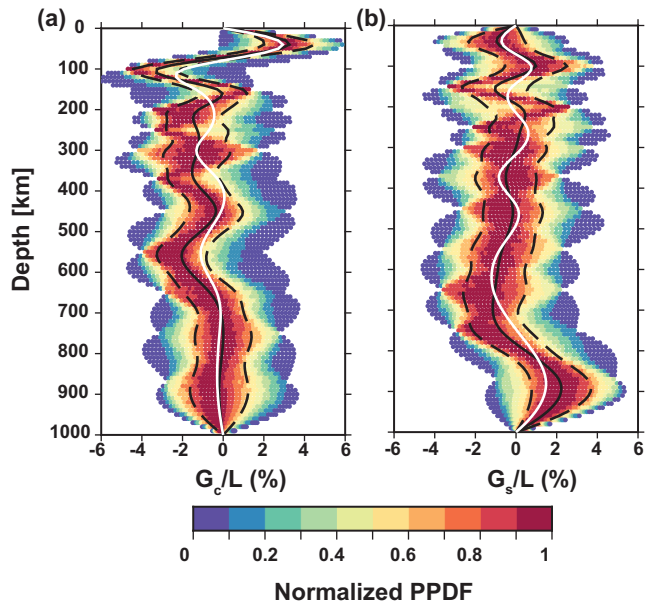
Figure 7. Example of normalized PPDFs for (a) G_c/L, and (b) G_s/L at $-55°$ longitude and $-60°$ latitude. L is the Love (1927) elastic coefficient that controls the speed of vertically polarized shear waves. The solid black lines represent the mean values of the distributions, the dashed black lines represent one standard deviation and the solid white line is from model YB17SVaniSVD at the same location.

presented in this paper, while $\mathbf{g}(\mathbf{m})$ is linear and the data are assumed to be Gaussian-distributed, the prior model parameter density function is uniform across the model space search range. The posterior model PPDFs obtained are not necessarily Gaussian and reflect the state of information we possess on the model parameters.

From the PPDFs of these $G_{c,s}^i$ spline parameters, we can reconstruct probability distributions for G_c and G_s as a function of depth by:

(i) Drawing 10 000 random values for each of the $G_{c,s}^i$ ($i = 1, \ldots,$ 12) coefficients from their posterior 1-D marginal distributions;

(ii) For each set of $G_{c,s}^i$ values, calculate the $G_{c,s}(r)$ profile [eqs 6 for $G_c(r)$ and 7 for $G_s(r)$], which results in 10 000 $G_{c,s}(r)$ models.

This yields distributions of $d\ln G_c(r)$ and $d\ln G_s(r)$ models drawn directly from the G_c^i and G_s^i PPDFs at each grid cell. An example is shown in Fig. 7. We note that the mean models as identified by the NA do not necessarily coincide with the inversion results, and that the data generally favour larger amplitudes than obtained from a regularized inversion. It should also be pointed out that the mean model does not necessarily correspond to the best-fitting model since not all PPDFs are Gaussian. This is why it is important to not discuss the mean model alone, but to account for its uncertainties. Finally, the reader should note that the existence of discontinuities in seismic velocities in the local reference models was shown to not be responsible for changes in G_c and G_s by Yuan & Beghein (2013).

3.4 Error propagation

To evaluate the robustness of the features observed in a tomographic model such as YB13SVani (Yuan & Beghein 2013) or YB17SVaniSVD, we ideally need to determine the mean amplitude $d\ln G$ and mean fast axis direction Θ at each depth and grid cell together with their respective uncertainties. However, the for-

mulation of the linearized forward problem described in Section 2 does not allow us to model $d\ln G$ and Θ directly. Previous attempts at estimating model uncertainties on azimuthal anisotropy with the NA (Yao 2015) have focused on the G_c and G_s uncertainties, and only discussed the azimuthal anisotropy ($d\ln G$ and Θ) model resulting from the most likely or mean G_c and G_s only. Because we adopted a Bayesian approach in this study, posterior uncertainties on $d\ln G$ and Θ can, however, be transmitted from the $d\ln G_c$ and $d\ln G_s$ model distributions, as explained below.

One approach consists in calculating $d\ln G$ and Θ distributions by (1) drawing models from the PPDFs of the individual $d\ln G_c(r)$ and $d\ln G_s(r)$ profiles (Fig. 7), and (2) calculating $d\ln G(r)$ and $\Theta(r)$ for each pair of $d\ln G_c(r)$ and $d\ln G_s(r)$ models drawn. This would yield distributions of $d\ln G(r)$ and $\Theta(r)$ models drawn directly from the $d\ln G_c$ and $d\ln G_s$ PPDFs. A similar method was taken by Beghein & Trampert (2004a) and Visser *et al.* (2008a) for radial anisotropy. One can then derive a mean and standard deviation from the reconstructed PPDFs, though we point out that they might not be Gaussian. In our case, however, this technique yields a mean model that has little to do with the model that results from the best-fitting $d\ln G_c$ and $d\ln G_s$ because of the highly non-Gaussian nature of the resulting $d\ln G$ and Θ PPDFs. The interpretation of the model and its error bars is thus extremely challenging as demonstrated in Figs S5–S7 in the Supporting Information.

We opted for another approach instead, involving the propagation of the errors obtained from the individual $d\ln G_c(r)$ and $d\ln G_s(r)$ PPDFs. This approach assumes the PPDFs are normally distributed, which is clearly an approximation for some parameters (Figs 6 and 7), but it enables us to avoid possible artefacts such as those seen in the synthetic examples. Let us take a function f that depends on parameters x and y that are assumed to be Gaussian with standard deviations σ_x and σ_y, respectively. If we further assume that the x and y variables have no covariance, the variance σ_f^2 of function f depends on the variances σ_x^2 and σ_y^2 of x and y as follows (Clifford 1973):

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2 \qquad (17)$$

Therefore, for $f = G = \sqrt{G_s^2 + G_c^2}$, we can determine that :

$$\sigma_G^2 = \frac{G_s^2 \sigma_{G_s}^2 + G_c^2 \sigma_{G_c}^2}{G_s^2 + G_c^2} \qquad (18)$$

And for $f = \Theta = \frac{1}{2} \arctan(G_s/G_c)$:

$$\sigma_\Theta^2 = \frac{1}{4} \frac{G_c^2 \sigma_{G_s}^2 + G_s^2 \sigma_{G_c}^2}{(G_s^2 + G_c^2)^2} \qquad (19)$$

Here, we used $G_c = G_{c,\,\text{mean}}$ and $G_s = G_{s,\,\text{mean}}$ as determined from eq. (12) and the variances $\sigma_{G_c}^2$ and $\sigma_{G_s}^2$ result from the off-diagonals of the posterior covariance matrix (eq. 13). The uncertainty maps displayed in Section 4 were determined from these error propagation calculations. We assumed no covariance between G_s and G_c, which is a reasonable approximation since Visser *et al.* (2008b) showed there was little covariance between the dc_1 and dc_2 terms of eq. (1).

4 RESULTS AND DISCUSSION

4.1 Goodness of fit

Fig. 8 compares some of the azimuthally anisotropic phase velocity maps measured by Visser *et al.* (2008b) with predictions from the model resulting from our regularized inversion (YB17SVaniSVD)
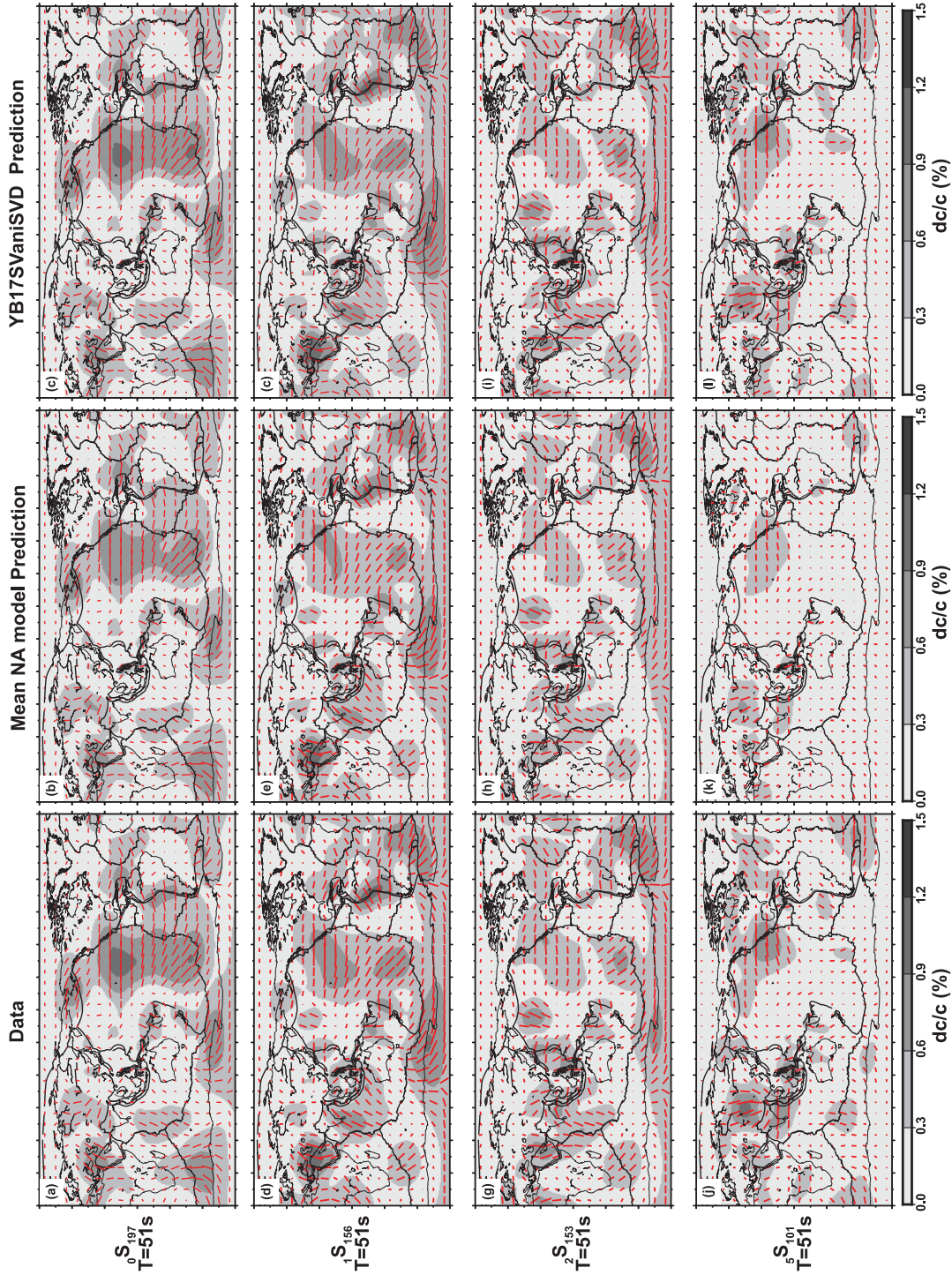
Figure 8. Phase velocity maps measured by Visser *et al.* (2008b) (left) compared to predictions from the mean NA model (middle) and from the model obtained by singular value decomposition (right) for the 51s Rayleigh wave fundamental mode (a)–(c), first overtone (d)–(f), second overtone (g)–(i) and fifth (j)–(l) overtone.

Table 1. Average χ^2 misfit for the model obtained by regularized inversion and for the mean NA model.

Model	$n = 0$–6	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
YB17SVaniSVD	0.04	0.02	0.023	0.04	0.07	0.04	0.07	0.07
Mean NA model	0.17	0.56	0.18	0.20	0.19	0.14	0.24	0.33

Table 2. Average variance reduction for the model obtained by regularized inversion and for the mean NA model.

Model	$n = 0$–6	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
YB17SVaniSVD	0.94	0.95	0.93	0.89	0.80	0.77	0.66	0.45
Mean NA model	0.79	0.88	0.68	0.71	0.67	0.50	0.40	0.20

and from the mean NA model, that is, the model corresponding to the mean of the $G_c(r)$ and $G_s(r)$ distributions. We see that both models can generally reproduce the data well and that the discrepancies are mostly in the amplitudes and less so in the fast axes directions. This figure also shows that an NA inversion for G only can yield a model that fits the data as well as a model obtained by regularized inversion for parameters B, G, and H.

Following Yuan & Beghein (2014), we calculated an average χ^2 misfit by averaging the χ^2 from the G_c model (χ_c^2) and from the G_s model (χ_s^2) over all grid cells:

$$\chi_{c,s}^2 = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{d_i - (Am)_i}{\sigma_i} \right)^2 \tag{20}$$

$$\chi^2 = \frac{1}{N_c} \sum_{i=1}^{N_c} \left(\chi_{c,i}^2 + \chi_{s,i}^2 \right) \tag{21}$$

where N is the total number of data, d_i is the ith component of the data vector **d**, σ_i is the ith component of the vector containing data errors and $(Am)_i$ is the ith component of the data prediction vector **Am** calculated using eqs (8) and (9). N_c is the total number of grid cells. Table 1 gives the average χ^2 misfit for each model, and confirms that the two models can explain the data within uncertainties.

Table 2 compares the average variance reduction for the two models, using the following definition of the variance reduction:

$$\text{VR} = 1 - \frac{\sum_{i=1}^{N} (d_i - (Am)_i)^2}{\sum_{i=1}^{N} d_i^2} \tag{22}$$

It shows that model YB17SVaniSVD explains 94 per cent of the data, and the mean NA model explains 79 per cent of the data. We attribute the better data fit of YB17SVaniSVD compared to the mean NA model to the fact that the mean NA model does not correspond exactly to one of the best data-fitting models due to the non-Gaussian nature of the posterior model 1-D distributions. We note that the low χ^2 values do not mean the model overfits the data, as shown by the values calculated for the variance reduction. The low χ^2 is due, instead, to the large data uncertainties.

One more point of importance when discussing data fit regards the depth extent of the anisotropy. Our Bayesian approach was applied to the parameter estimation problem but not to explore the model selection problem. A transdimensional method (Bodin *et al.* 2012) would have been able to determine the optimum depth parametrization in addition to the model parameter distribution at each depths, but with the NA employed here we had to fix the depth parametrization *a priori*, which can cause bias in the model. It is thus important to test whether the data require anisotropy to depths as great as the transition zone or whether they could be explained equivalently well with anisotropy restricted to the upper 410 km of the mantle for instance. Yuan & Beghein (2013) demonstrated using

statistical *F*-tests that the fit to the Rayleigh wave data employed was significantly improved by allowing azimuthal anisotropy below 410 km depth, which is what guided or choice of spline functions (Fig. 3). We performed identical tests for this study and forced the anisotropy to exist only in the upper 400 km of the mantle at a few locations where our model displays strong MTZ anisotropy. The *F*-tests performed on the mean G_c and G_s determined that including anisotropy below 400 km depth significantly decreases the data misfit.

4.2 Global averages

Some of the most interesting features detected in YB13SVani (Yuan & Beghein 2013) were the changes in the average azimuthal anisotropy fast directions associated with amplitude minima at about 220 km depth, and near the MTZ boundaries. While the interpretation of these results is non-unique because too few mineral physics data are available on the anisotropy of MTZ material, these results provide new constraints on deep upper-mantle circulation, and the observed changes in anisotropy at the MTZ boundaries could be the signature of changes in mantle flow direction. To determine whether our new results confirm these findings, we determined the vertical gradient of the fast axes ($d\Theta/dr$) and the relative anisotropy amplitude ($d\ln G$) at each grid cell every 10 km depth with a 20 km window, after which we calculated their rms as a function of depth, following Yuan & Beghein (2013) Fig. 9 represents the rms of $d\ln G$ and of $d\Theta/dr$ calculated for the mean NA model and for YB17SVaniSVD.

We see that even though the 1-D average of the mean NA model presents a few more oscillations below 300 km depth than the model obtained by regularized inversion, the two models display similar features. We observe 1.5–2 per cent anisotropy in the top 200 km and about 1 per cent below, down to at least the bottom of the MTZ. We also detect amplitude minima between 50 and 100 km, around 220 km and 250 km, and near the boundaries of the MTZ. These minima are associated with higher gradients in the fast axes direction, as observed in YB13SVani.

Note that in Yuan & Beghein (2013), we demonstrated that these changes in anisotropy are not artefacts due to the presence of discontinuities in seismic velocities in the local reference models, and that they were stable with respect to regularization and with respect to the presence of lateral heterogeneities in the mantle. We also previously demonstrated (Yuan & Beghein 2013, 2014) that the model does not depend on the choice of the spline functions, the position of their peaks, or their spacing. The robustness of these features is of course better tested with quantitative model uncertainties, but these are difficult to display for rms($d\ln G$) and rms($d\Theta/dr$). We decided to focus on the uncertainties of the 3-D model instead (see below).
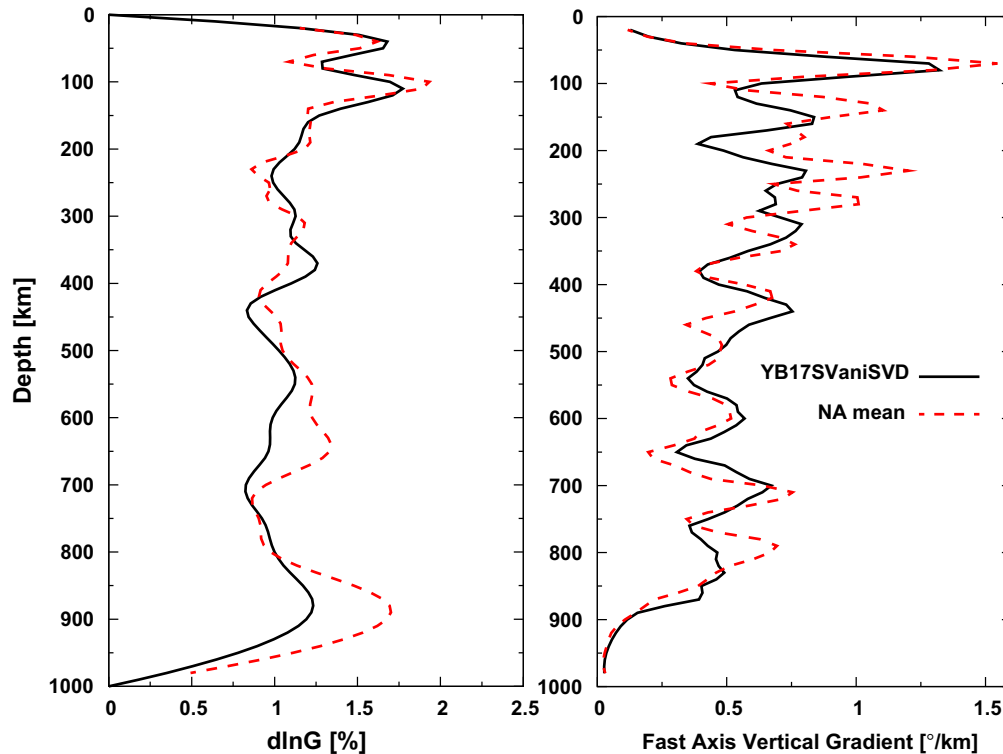
Figure 9. 1-D average model amplitude (left) and gradient of the fast axis direction (right) calculated from the mean G_s and mean G_c distributions. Model YB17SVaniSVD is shown by the black curves for comparison.
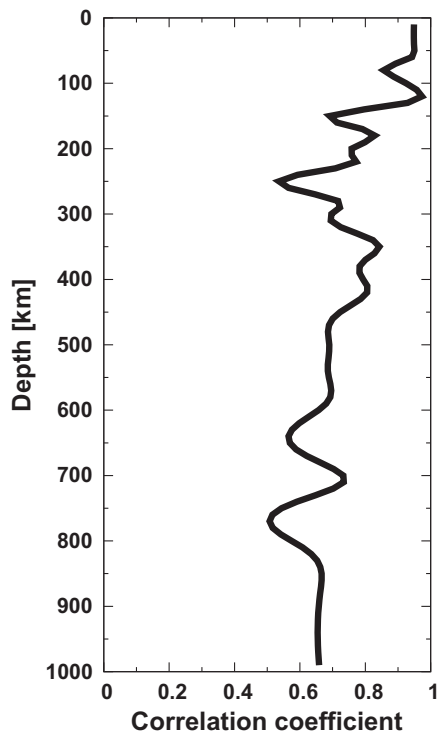
Figure 10. Correlation between YB17SVaniSVD and the mean NA model obtained from the mean $G_s(r)$ and mean $G_c(r)$ distributions.

4.3 3-D models

In this section, we discuss the 3-D models obtained with the NA and compare them with YB17SVaniSVD. Fig. 10 shows the correlation coefficient between the two models as a function of depth. It was

calculated after expansion of each model map in generalized spherical harmonics up to degree 20, following Yuan & Beghein (2014). At all depths, the correlation is well above the 95 per cent significance level as calculated by Becker *et al.* (2007), demonstrating that the two models are overall consistent with one another.
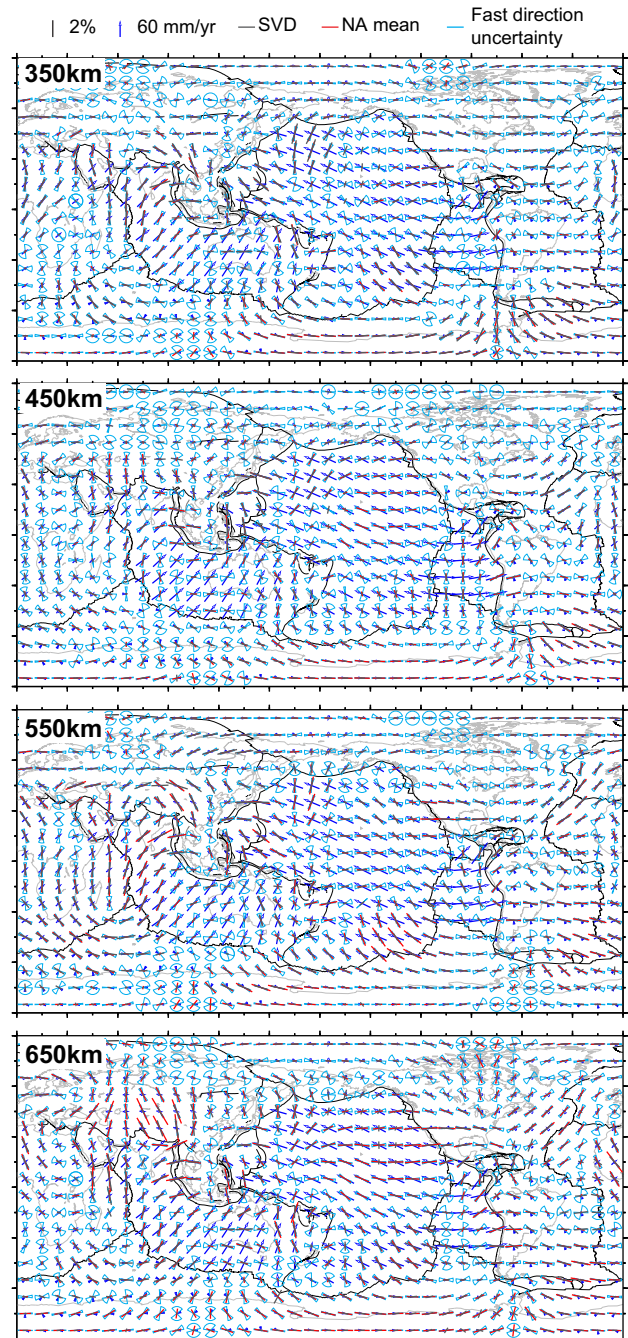
Figs 11–14 are maps that represent model YB17SVaniSVD and the NA results at different depths. In Figs 11 and 12, both the mean NA model and the fast axes standard deviation are displayed. The fast axes standard deviation was estimated at each grid cell with eq. (19). Figs 13 and 14 focus on the anisotropy amplitude and its standard deviation (eq. 18). The two models show very similar fast axes directions at most depths and comparable amplitudes. They are also consistent with previous studies in the top 200 km (Nishimura & Forsyth 1989; Montagner & Tanimoto 1991; Debayle *et al.* 2005). Differences in model amplitudes are generally within the model uncertainties [e.g. differences in the anisotropy pattern in the western Pacific at 100 km between YB17SVaniSVD and DR2013 (Debayle & Ricard 2013) or SL2016svA (Schaeffer *et al.* 2016)]. We note, however, that the large amplitudes seen in the Debayle & Ricard (2013) model in the uppermost mantle are difficult to reconcile with our results, even accounting for the posterior model variance, except near the South American subduction zone and at the Eurasia–Africa boundary. Similarly, we note that Schaeffer *et al.* (2016) published SL2016svAr (not represented in Fig. 1 for clarity), which is a rougher model than SL2016svA. It was built with the same data set and data fit as SL2016svA, but a different parametrization resulting in larger amplitudes and higher resolution in well-sampled areas.

The strongest model amplitudes (of at least 2–3 per cent anisotropy) in the top 150 km of our model are well resolved and can be found in the youngest parts of the Pacific plate, at the Africa–Eurasia plate boundary, and around the South American subduction zone. Lower amplitudes are seen in the western Pacific at these
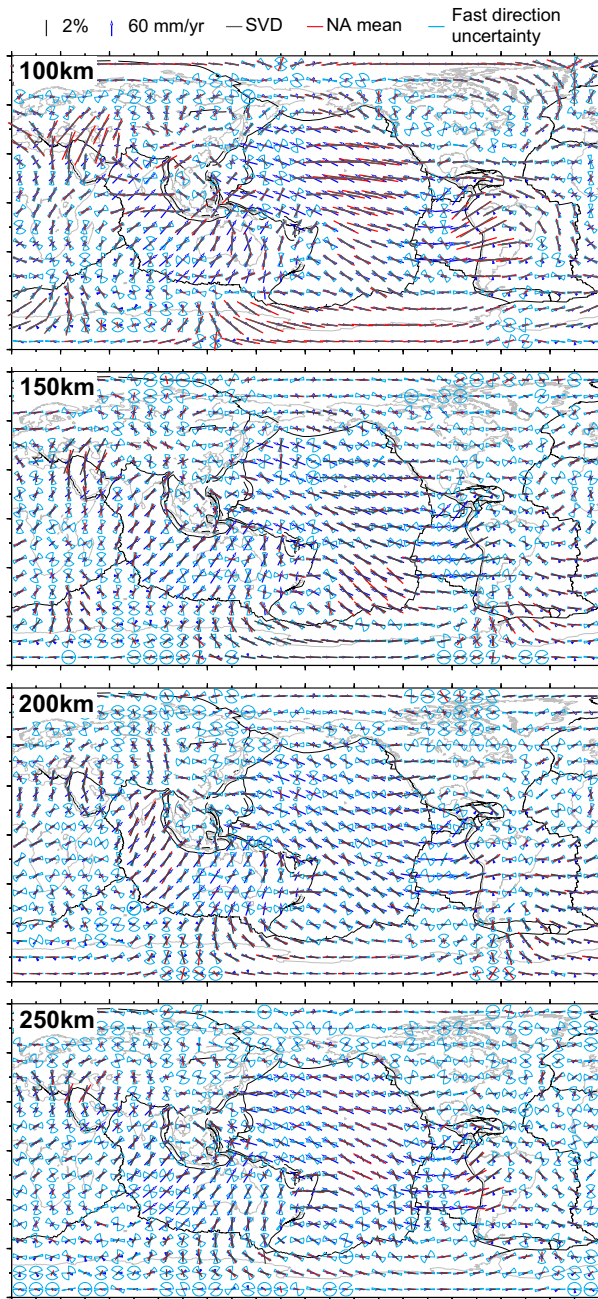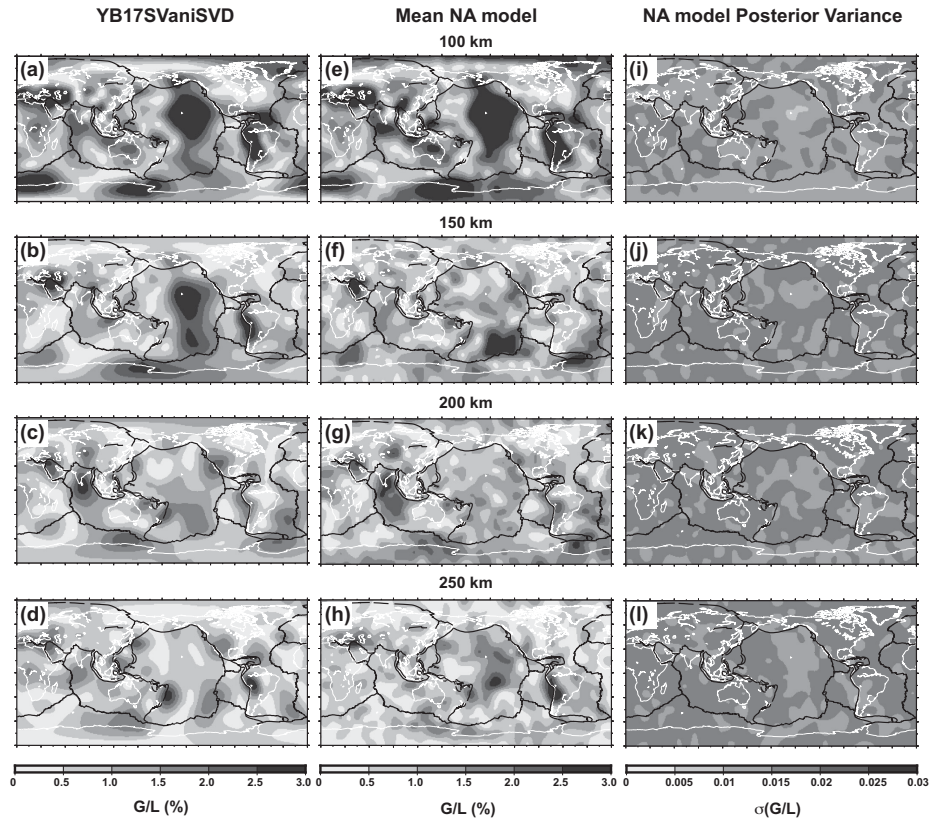
Figure 11. Model YB17SVaniSVD (black bars) obtained by regularized inversion and mean mantle model (red bars) obtained using the NA for the uppermost mantle. The standard deviation on the fast axes direction as obtained from the NA is shown in light blue. The anisotropy amplitude is proportional to the length of the red and black bars. Model NNR-NUVEL 1A is represented by the darker blue arrows. Plate boundaries are shown by thin black lines, and continents are delimited by thin grey lines.
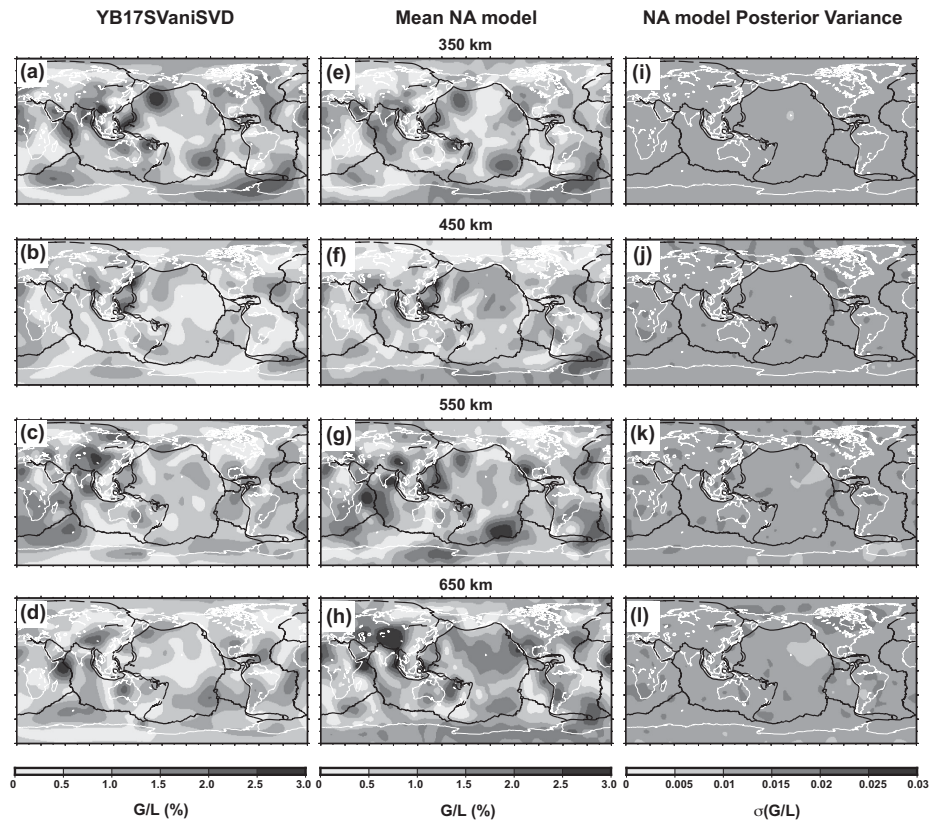

Figure 12. Model YB17SVaniSVD (black bars) obtained by regularized inversion and mean mantle model (red bars) obtained using the NA for the deep upper mantle. See Fig. 11 for more details.

depths. These low amplitudes were first detected by Nishimura & Forsyth (1989) who related them to changes in the horizontal direction of anisotropic fabric with depth rather than being due to a decrease of *in situ* anisotropy. In Yuan & Beghein (2014), we showed, however, that the lower *SV* anisotropy amplitude in the western Pacific is close to the average amplitudes of other oceanic plates and is therefore not anomalously low with respect to other plates. We also note that while seismic anisotropy amplitudes are anomalously high in the shallow mantle in the middle of the Pacific

plate, it is one of the places where the amplitude is the weakest at greater depths, suggesting a relatively shallow origin for this signal, such as asthenospheric mantle flow. This was previously suggested for radial anisotropy models (Ekström & Dziewonski 1998; Gaboret *et al.* 2003). We also find a relatively strong signal of about 3 per cent anisotropy at 100 km depth near the India–Eurasia convergence zone and in the Indonesian subduction region. Amplitude uncertainties are closer to the mean model amplitudes at greater depths, except in a few locations between 200 and 350 km such as the Western part of the Pacific where subduction occurs, near the Arabian plate, and India. Below 350 km depth, a stronger,

Figure 13. (a)–(d) Amplitude of model YB17SVaniSVD, (e)–(h) of the mean mantle model obtained using the NA for the uppermost mantle and (i)–(l) mean model amplitude standard deviation. Plate boundaries are shown by black lines, and continents are delimited by thin white lines.

Figure 14. (a)–(d) Amplitude of model YB17SVaniSVD, (e)–(h) of the mean mantle model obtained using the NA for the deep upper mantle and (i)–(l) mean model amplitude standard deviation. Plate boundaries are shown by black lines, and continents are delimited by thin white lines.
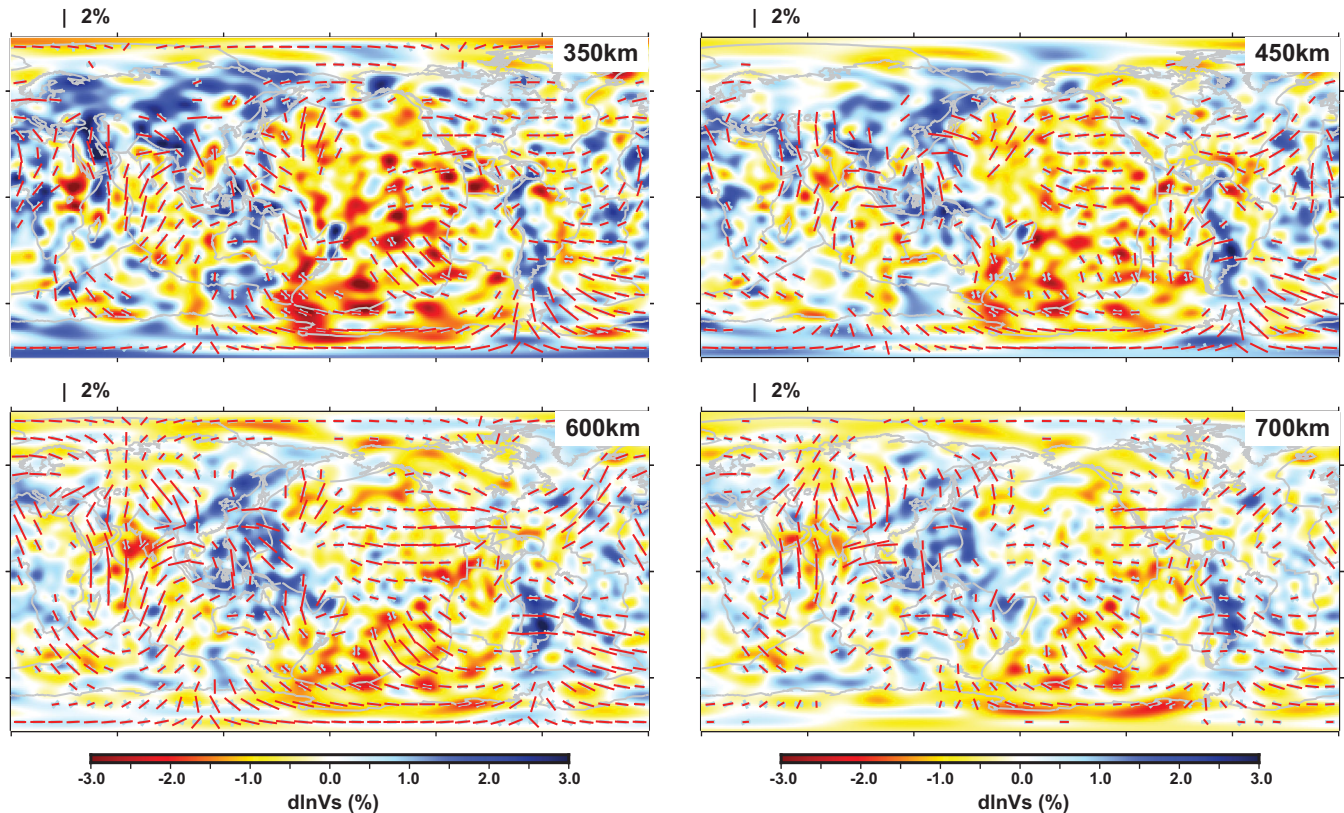
Figure 15. Fast seismic direction in locations where the uncertainty is lower than 45°. The colour background represents isotropic velocity model SEMUCB (French & Romanowicz 2014).

well-resolved signal appears in the northwestern part of the Pacific and Asia.

Strong discrepancies were found between model YB13SVani (Yuan & Beghein 2013) and the uppermost mantle model of Marone & Romanowicz (2007) under North America. We had attributed this disagreement to differences in the horizontal resolution of the models (Yuan & Beghein 2013). Here, we see that the uncertainties in the fast axes directions at 100 km are strong beneath this region, which would reconcile the differences between the models. Uncertainties on the fast axes are also slightly stronger toward the western part of the Pacific. As we go deeper, more regions display larger standard deviations in the fast axes direction, but a few features appear well constrained. For instance, the fast seismic direction is close to the present-day plate motion, which was calculated with model NNR-NUVEL 1A (DeMets *et al.* 1994), beneath the young and mid-Pacific plate down to about 150 km depth (Fig. 11), though model YB17SVaniSVD appears to reflect the plate motion slightly better than the mean NA model (see, for instance, in the younger parts of the Pacific plate). This is attributed to the fact that the mean model is not necessarily the best-fitting model due to the non-Gaussian topology of the model space, as explained above. We do not expect other reference frames to yield a better alignment with the anisotropy since our new models are very similar to YB13SVani, which had been tested against other plate motion models (Yuan & Beghein 2013). In Fig. S8 in the Supporting Information, we additionally compared our models with the more recent plate motion model of Becker *et al.* (2015), which was not published at the time of the Yuan & Beghein (2013) study, and which optimizes the match between absolute plate motions and spreading orientations. We did not find strong differences between the match of our models with

the Becker *et al.* (2015) Absolute Plate Motion (APM) model and with NNR-NUVEL 1A. Interestingly, the alignment of the fast seismic direction and present-day plate motion appears to continue to depths as great as 350 km (Fig. 12) in a few locations such as the eastern Indian Ocean and central Africa, though whether this is the manifestation of deep upper-mantle physical processes or artefacts in the model needs to be investigated more thoroughly.

A question that arises from Fig. 9 is whether the changes in azimuthal anisotropy at the MTZ boundaries are global or appear only at a few locations. If they occur globally, they might be caused purely by the effect of pressure on MTZ material anisotropy. If they occur only in a few regions, compositional effects might come into play. To try to answer this question, one can make a simple visual comparison of the model maps at different depths. It is, however, important to keep in mind when comparing maps that not all grid cells have well-resolved fast directions and that the phase velocity maps may have been affected by small trade-offs between isotropic and anisotropic anomalies. We thus took advantage of the fact that the forward modeling method used here yielded quantitative posterior model uncertainties and plotted the anisotropy only at locations where the fast direction was best resolved. This is displayed in Fig. 15 for depths of 350, 450, 600, and 700 km. They represent the fast direction at grid cells where the error on Θ was less than 45°. This is a subjective cut-off value, but looking at smaller cut-off values (e.g. 35°) did not change our conclusions. Visual inspection of these maps shows that there is little variations of the fast seismic direction across depths in the general area where the Pacific plate subducts underneath the North American plate and under the Philippine plate. The same is also true, for instance, around the South American subduction zone, between the Arabian and
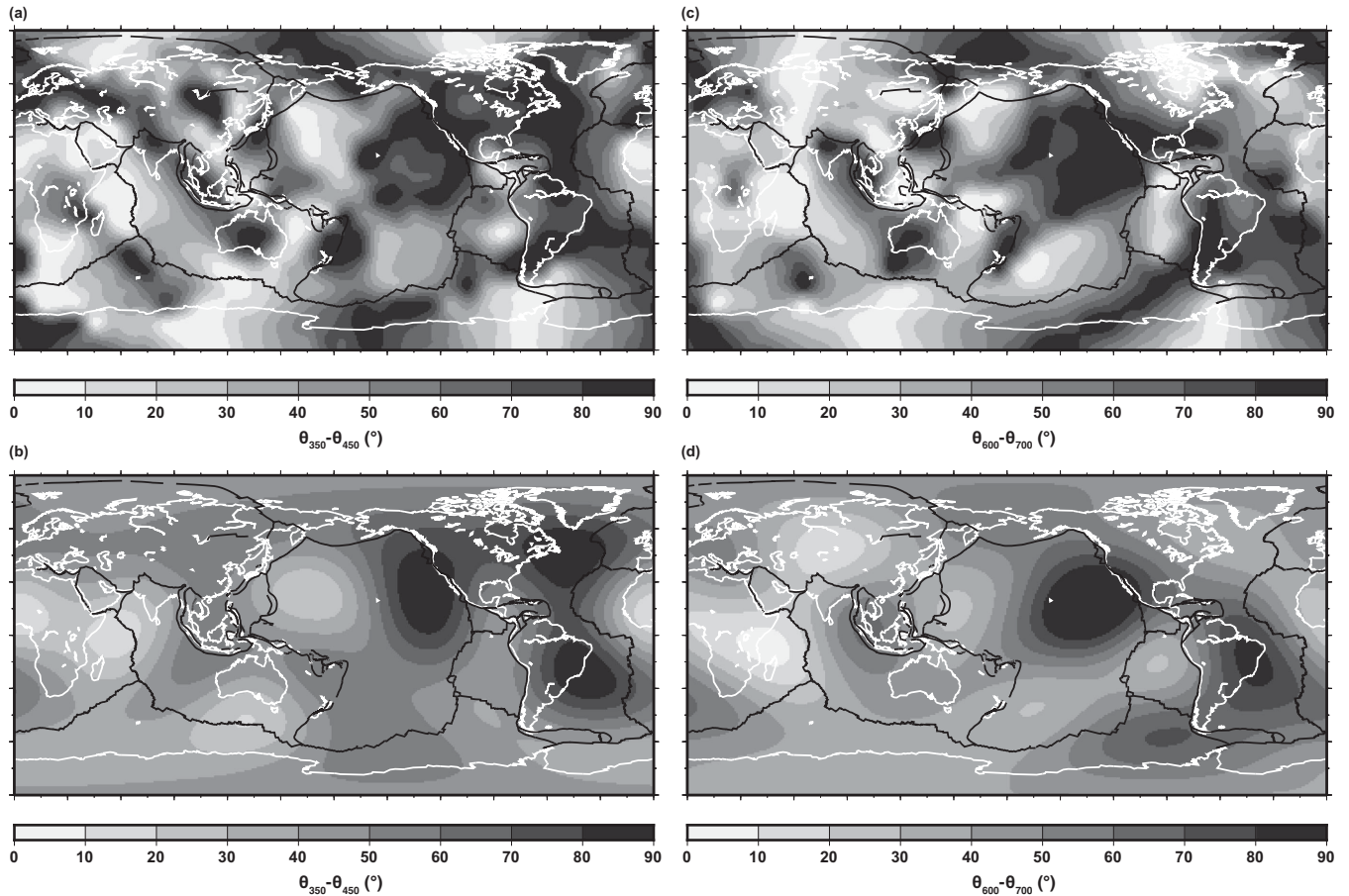
Figure 16. Difference between mean fast axis Θ_m across the 410 and 670 discontinuities. (a) and (b) represent the difference between the fast direction at 350 and 450 km depth, and (c) and (d) is the difference between fast directions at 600 and 700 km. (a) and (c) were determined at every grid cell we used to parametrize Earth's surface. (b) and (d) were obtained by filtering (a) and (c) up to spherical harmonic degree 5. Plate boundaries are shown by black lines, and continents are delimited by white lines.

Eurasian plates, in the Southeastern Pacific, and where the Indian, African and Indo-Australian plates meet.

Another way of looking at this is by calculating the difference between the fast direction at depths above and below the MTZ boundaries. This is what is represented in Fig. 16. It shows the difference between the mean fast directions at 350 and 450 km and between depths of 600 and 700 km at all grid cells and for spherical harmonic degrees 1–5, which is the estimated lateral resolution for azimuthal anisotropy at these depths (see Section 2). While one might be tempted to conclude from Fig. 15 that subduction zones are characterized by the same fast direction across the MTZ, there is no clear pattern relating to surface tectonics visible in Fig. 16. Nevertheless, from the degree 5 maps, one can conclude that regions such as Africa, Asia and the northwestern Pacific are characterized by similar azimuthal anisotropy above, below and inside the MTZ. Thus, even though the lateral resolution of the higher modes and posterior model uncertainties do not allow us to determine with high precision where the fast axes do and do not change at MTZ depths, our results suggest that the change in fast direction across the MTZ boundaries is not likely to occur globally and is thus not solely due to pressure effects.

5 CONCLUSIONS

The goal of this research was to present a new method to model and obtain quantitative uncertainties on 3-D azimuthal seismic

anisotropy. It was applied to global higher mode surface wave phase velocity data to assess the likelihood of azimuthal anisotropy in the deep upper mantle. For this, we employed the Neighbourhood Algorithm developed by Sambridge (1999a,b), a model space search approach that enables searching a broader part of the model space than a damped inversion, including the null space. Even though the linearized formulation of the problem relating phase velocities to azimuthal anisotropy at depth does not allow us to directly obtain uncertainties on the anisotropy amplitude and fast axes direction, we showed that that they can be determined *a posteriori*.

The PPDFs of the resulting models yielded a mean model that was overall consistent (correlation above the 95 per cent significance level) with models obtained by regularized inversion with the same data set and parametrization, but with somewhat larger amplitudes. The posterior model variance was also larger than estimates from regularized inversions, which is to be expected in the presence of a large model null space. We confirm our previously published results showing that azimuthal anisotropy of 1–2 per cent is present in the MTZ and that, on average, the anisotropy changes across the MTZ boundaries. This change is therefore required by the higher mode data utilized, and did not result from inversion artefact or parameter trade-offs that could have affected our previous model, YB13SVani. We showed, however, that the anisotropy change across the MTZ boundaries is likely not a global feature, but further studies will be required to improve the lateral resolution of the models at

those depths and determine whether there is any relation between the change (or lack thereof) of anisotropy across the 410 and 670 discontinuities.

REFERENCES

Auer, L., Boschi, L., Becker, T., Nissen-Meyer, T. & Giardini, D., 2014. Savani: a variable resolution whole-mantle model of anisotropic shear velocity variations based on multiple data sets, *J. geophys. Res.,* **119**(4), 3006–3034.

Bassin, C., Laske, G. & Masters, G., 2000. The current limits of resolution for surface wave tomography in North America, *Eos Trans. AGU,* **81**(48) Fall Meet. Suppl., Abstract S12A-03.

Becker, T.W., Kellogg, J.B., Ekström, G. & O'Connell, R.J., 2003. Comparison of azimuthal seismic anisotropy from surface waves and finite strain from global mantle-circulation models, *Geophys. J. Int.,* **155**(2), 696–714.

Becker, T.W., Ekström, G., Boschi, L. & Woodhouse, J.D., 2007. Length scales, patterns and origin of azimuthal seismic anisotropy in the upper mantle as mapped by Rayleigh waves, *Geophys. J. Int.,* **171**(1), 451–462.

Becker, T.W., Conrad, C.P., Schaeffer, A.J. & Lebedev, S., 2014. Origin of azimuthal seismic anisotropy in oceanic plates and mantle, *Earth planet. Sci. Lett.,* **401**(1), 236–250.

Becker, T.W., Schaeffer, A.J., Lebedev, S. & Conrad, C.P., 2015. Toward a generalized plate motion reference frame, *Geophys. Res. Lett.,* **42**, 3188–3196.

Beghein, C., 2010. Radial anisotropy and prior petrological constraints: a comparative study, *J. geophys. Res.,* **115**, B03303, doi:10.1029/2008JB005842.

Beghein, C. & Trampert, J., 2003. Robust normal mode constraints on inner-core anisotropy from model space search, *Science,* **299**, 552–555.

Beghein, C. & Trampert, J., 2004a. Probability density functions for radial anisotropy from fundamental mode surface wave data and the neighbourhood algorithm, *Geophys. J. Int.,* **157**, 1163–1174.

Beghein, C. & Trampert, J., 2004b. Probability density functions for radial anisotropy: implications for the upper 1200 km of the mantle, *Earth planet. Sci. Lett.,* **217**, 151–162.

Beghein, C., Resovsky, J. & Trampert, J., 2002. P and S tomography using normal mode and surface wave data with a Neighbourhood Algorithm, *Geophys. J. Int.,* **149**(3), 646–658.

Beghein, C., Trampert, J. & van Heijst, H.J., 2006. Radial anisotropy in seismic reference models of the mantle, *J. geophys. Res.,* **111**, B02303, doi:10.1029/2005JB003728.

Beghein, C., Resovsky, J. & van der Hilst, R.D., 2008. The signal of mantle anisotropy in the coupling of normal modes, *Geophys. J. Int.,* **175**(3), 1209–1234.

Beghein, C., Yuan, K., Schmerr, N. & Xing, Z., 2014. Changes in seismic anisotropy shed light on the nature of the Gutenberg discontinuity, *Science,* **343**, 1237–1240.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K. & Rawlinson, N., 2012. Transdimensional inversion of receiver functions and surface wave dispersion, *J. geophys. Res.,* **117**, doi:10.1029/2011JB008560.

Boschi, L. & Ekström, G., 2002. New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies, *J. geophys. Res.,* **107**, B4, doi:10.1029/2000JB000059.

Bozdağ, E. & Trampert, J., 2010. Assessment of tomographic mantle models using spectral element seismograms, *Geophys. J. Int.,* **180**(3), 1187–1199

Chang, S., Ferreira, A., Ritsema, J., van Heijst, H. & Woodhouse, J.H., 2014. Global radially anisotropic mantle structure from multiple datasets: a review, current challenges, and outlook, *Tectonophysics,* **617**, 1–19.

Chen, W.P. & Brudzinski, M.R., 2003. Seismic anisotropy in the mantle transition zone beneath Fiji-Tonga, *Geophys. Res. Lett.,* **30**(13), 1682, doi:10.1029/2002GL016330.

Clifford, A.A., 1973. *Multivariate Error Analysis: A Handbook of Error Propagation and Calculation in Many-Parameter Systems,* Wiley, New York.

Debayle, E. & Ricard, Y., 2013. Seismic observations of large-scale deformation at the bottom of fast-moving plates, *Earth planet. Sci. Lett.,* **376**, 165–177.

Debayle, E., Kennett, B. & Priestley, K., 2005. Global azimuthal seismic anisotropy and the unique plate-motion deformation of Australia, *Nature,* **433**(7025), 509–512.

DeMets, C., Gordon, R.G., Argus, D.F. & Stein, S., 1994. Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions, *Geophys. Res. Lett.,* **21**, 2191–2194.

Dziewonski, A. M. & Anderson, D. L., 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.,* **25**(4), 297–356.

Ekström, G., 2011. A global model of Love and Rayleigh surface wave dispersion and anisotropy, 25-250 s, *Geophys. J. Int.,* **187**(3), 1668–1686.

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

Foley, B.J. & Long, M.D., 2011. Upper and mid-mantle anisotropy beneath the Tonga slab, *Geophys. Res. Lett.,* **38**(2), L02303, doi:10.1029/2010GL046021.

Fouch, M.J. & Fischer, K.M., 1996. Mantle anisotropy beneath northwest Pacific subduction zones, *J. geophys. Res.,* **101**(B7), 15987–16002.

French, S.W. & Romanowicz, B., 2014. Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography, *Geophys. J. Int.,* **199**(3), 1303–1327.

Gaboret, C., Forte, A.M. & Montagner, J.-P., 2003. The unique dynamics of the Pacific Hemisphere mantle and its signature on seismic anisotropy, *Earth planet. Sci. Lett.,* **208**(3–4), 219–233.

Hu, X.-G., Xue, X.-X., Liu, L.-T. & Sun, H.-P., 2012. Normal mode coupling due to azimuthal anisotropy in the transition zone: an example from Taiwan Island, *Geophys. J. Int.,* **190**(1), 323–334.

Karato, S., 1998. Seismic anisotropy in the deep mantle, boundary layers and the geometry of mantle convection, *Pure appl. Geophys.,* **151**(2–4), 565–587.

Karato, S., Zhang, S. & Wenk, H.-R., 1995. Superplasticity in Earth's lower mantle – evidence from seismic anisotropy and rock physics, *Science,* **270**(5235), 458–461.

Karato, S.-i., 1989. Seismic anisotropy: mechanisms and tectonic implications, in *Rheology of Solids and of the Earth,* pp. 393–422, eds Karato, S.-I. & Toriumi, M., Oxford University Press, New York.

Kendall, J.-M. & Silver, P.G., 1996. Constraints from seismic anisotropy on the nature of the lowermost mantle, *Nature,* **381**(6581), 409–412.

Kosarian, M., Davis, P.M., Tanimoto, T. & Clayton, R.W., 2011. The relationship between upper mantle anisotropic structures beneath California, transpression, and absolute plate motions, *J. geophys. Res.,* **116**(B8), B08307, doi:10.1029/2010JB007742.

Kustowski, B., Dziewonski, A. & Ekström, G., 2007. Nonlinear crustal corrections for normal-mode seismograms, *Bull. seism. Soc. Am.,* **97**(5), 1756–1762.

Kustowski, B., Ekström, G. & Dziewonski, A., 2008. Anisotropic shear-wave velocity structure of the Earth's mantle: a global model, *J. geophys. Res.,* **113**(B6), B06306, doi:10.1029/2007JB005169.

Lebedev, S., Adam, J.M.C. & Meier, T., 2008. Mapping the Moho with seismic surface waves: a review, resolution analysis, and recommended inversion strategies, *Tectonophysics,* **609**(C), 377–394.

Lynner, C. & Long, M.D., 2015. Heterogeneous seismic anisotropy in the transition zone and uppermost lower mantle beneath Japan, Izu-Bonin, and South America, *Geophys. J. Int.,* **201,** 1545–1552.

Long, M.D., 2013. Constraints on subduction geodynamics from seismic anisotropy, *Rev. Geophys.,* **51**(1), 76–112.

Love, A.E.H., 1927. *A Treatise on the Mathematical Theory of Elasticity.* Cambridge University Press, Cambridge.

Marone, F. & Romanowicz, B., 2007. The depth distribution of azimuthal anisotropy in the continental upper mantle, *Nature,* **447**(7141), 198–201.

Maupin, V., Garnero, E.J., Lay, T. & Fouch, M.J., 2005. Azimuthal anisotropy in the D" layer beneath the Caribbean, *J. geophys. Res.,* **110,** B08301, doi:10.1029/2004jb003506.

Meier, U., Curtis, A. & Trampert, J., 2007. Global crustal thickness from neural network inversion of surface wave data, *Geophys. J. Int.,* **169**(2), B706–B722.

Menke, W., 2012. *Geophysical Data Analysis: Discrete Inverse Theory,* 3rd edn, Elsevier, 330 pp.

Montagner, J.-P., 1994. Can seismology tell us anything about convection in the mantle? *Rev. Geophys.,* **32**(2), 115–137.

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**(B12), 20 337–20 351.

Montagner, J.-P. & Kennett, B.L.N., 1996. How to reconcile body-wave and normal-mode reference Earth models, *Geophys. J. Int.,* **125**(1), 229–248.

Moulik, P. & Ekström, G., 2014. An anisotropic shear velocity model of the Earth's mantle using normal modes, body waves, surface waves and long-period waveforms, *Geophys. J. Int.,* **199**(3), 1713–1738.

Nicolas, A. & Christensen, N., 1987. Formation of anisotropy in upper mantle peridotites – a review, in *Composition, Structure, and Dynamics of the Lithosphere–Asthenosphere System,* eds Fuchs, A. & Froideveaux, C., American Geophysical Union, Washington, DC. doi: 10.1029/GD016p0111.

Nishimura, C.E. & Forsyth, D.W., 1988. Rayleigh wave phase velocities in the Pacific with implications for azimuthal anisotropy and lateral heterogeneities, *Geophys. J. Int.,* **94**(3), 479–501.

Nishimura, C.E. & Forsyth, D.W., 1989. The anisotropic structure of the upper mantle in the Pacific, *Geophys. J.,* **96**(2), 203–229.

Nowacki, A., Kendall, J.-M., Wookey, J. & Pemberton, A., 2015. Mid-mantle anisotropy in subduction zones and deep water transport, *Geochem. Geophys. Geosyst.,* **16,** 764–784.

Oganov, A., 2005. Anisotropy of Earth's D" layer and stacking faults in the $MgSiO_3$ post-perovskite phase, *Nature,* **438,** 1142–1144.

Panning, M. & Romanowicz, B., 2004. Inferences on flow at the base of Earth's mantle based on seismic anisotropy, *Science,* **303**(351), 351–353.

Panning, M. & Romanowicz, B., 2006. A three-dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophys. J. Int.,* **167**(1), 361–379.

Panning, M.P., Lekić, V. & Romanowicz, B., 2010. Importance of crustal corrections in the development of a new global model of radial anisotropy, *J. geophys. Res.,* **115**(B12), B12325, doi:10.1029/2010JB007520.

Romanowicz, B. & Lekić, V., 2011. Inferring upper-mantle structure by full waveform tomography with the spectral element method, *Geophys. J. Int.,* **185**(2), 799–831.

Sambridge, M., 1999a. Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space, *Geophys. J. Int.,* **138**(2), 479–494.

Sambridge, M., 1999b. Geophysical inversion with a neighbourhood algorithm—II. Appraising the ensemble, *Geophys. J. Int.,* **138**(3), 727–746.

Schaeffer, A.J., Lebedev, S. & Becker, T.W., 2016. Azimuthal seismic anisotropy in the Earth's upper mantle and the thickness of tectonic plates, *Geophys. J. Int.,* **207**(2), 901–933.

Silver, P.G., 1996. Seismic anisotropy beneath the continents: probing the depths of Ggology, *Annu. Rev. Earth Planet. Sci.,* **24**(1), 385–432.

Smith, D.B., Ritzwoller, M.H. & Shapiro, N.M., 2004. Stratification of anisotropy in the Pacific upper mantle, *J. geophys. Res.,* **109**(B11), doi:10.1029/2004JB003200.

Snoke, J.A. & Sambridge, M., 2002. Constraints on the S wave velocity structure in a continental shield from surface wave data: comparing linearized least squares inversion and the direct search Neighbourhood Algorithm, *J. geophys. Res.,* **107**(B5), doi:10.1029/2001JB000498.

Takeuchi, H. & Saito, M., 1972. Seismic surface waves, in *Methods in Computational Physics,* pp. 217–295, ed. Bolt, B., Academic Press New York.

Tarantola, A., 1987. *Inverse Problem Theory and Methods for Model Parameter Estimation.* Elsevier, New York.

Trampert, J., 1998. Global seismic tomography: the inverse problem and beyond, *Inverse Probl.,* **14,** 371–385.

Trampert, J. & van Heijst, H.J., 2002. Global azimuthal anisotropy in the transition zone, *Science,* **296**(5571), 1297–1299.

Trampert, J. & Woodhouse, J.H., 2003. Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s, *Geophys. J. Int.,* **154**(1), 154–165.

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

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

Wookey, J. & Barruol, G., 2002. Mid-mantle deformation inferred from seismic anisotropy, *Nature,* **415,** 777–780.

Yao, H., 2015. A method for inversion of layered shear wavespeed azimuthal anisotropy from Rayleigh wave dispersion using the Neighborhood Algorithm, *Earthq. Sci.* **28**(1), 59–69.

Yao, H., Beghein, C. & van der Hilst, R.D., 2008. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis: II. - Crust and upper mantle structure, *Geophys. J. Int.,* **173**(1), 205–219.

Yuan, K. & Beghein, C., 2013. Seismic anisotropy changes across upper mantle phase transitions, *Earth planet. Sci. Lett.,* **374,** 132–144.

Yuan, K. & Beghein, C., 2014. Three-dimensional variations in Love and Rayleigh wave azimuthal anisotropy for the upper 800 km of the mantle, *J. geophys. Res.,* **119**(4), 3232–3255.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. Inversion using the NA of the Visser *et al.* (2008b) dataset at a grid cell located at −25° lat and 305° lon. (A) is for a model space search for $d$lnGc, $d$lnBc, and $d$lnHc using 12 cubic spline functions for $d$lnGc. $d$lnHc and $d$lnBc were parameterized each with 6 cubic spline functions. We chose to use a coarser parameterization for these other parameters because are poorly resolved due to the similarity of their partial derivatives. (B) is for a model space search for $d$lnGc only using the same 12 splines as in (A). In each case, the model space search was performed around model YB17SVaniSVD, indicated by the red cross, allowing for perturbations between −0.03 and 0.03. The spline parameters are displayed as a function of the number of models generated. We labeled them as G_i (1 = 1,. . . ,12) instead of $d$lnGc,i for simplicity.

Figure S2. Synthetic tests comparing NA results when the model is sampled around the SVD inversion results (case 1) and around PREM (case 2, zero azimuthal anisotropy). Note that the axes labels in case 1 and case 2 are different. The model employed to calculate the synthetic data was model YB17SVaniSVD, denoted by the red bar.

Figure S3. NA results using the Visser *et al.* (2008b) dataset and searching the model space for G values (neglecting B and H). Case 1 corresponds to searching the mode space around YB17SVaniSVD (represented by the red line) and case 2 corresponds to searching around PREM (zero azimuthal anisotropy). Note that the axes labels in case 1 and case 2 are different.

Figure S4. Comparison of uncertainties calculated using the standard deviation (dashed blue lines) of the PPDF obtained from NA and using the covariance matrix of the model obtained by regularized inversion (dashed red line). The solid blue line corresponds to the mean model. The data point corresponding to this PPDF was located at 175° longitude and −85° latitude.

Figure S5. Synthetic examples of resampled $d\ln Gc$ and $d\ln Gs$ PPDFs to calculate $d\ln G$ and Theta distributions. Each of the six panels corresponds to different $d\ln Gc$ and $d\ln Gs$ distributions, with different means and variances as indicated in the legends. In each panel, the black vertical line in the $d\ln Gc$ and $d\ln Gs$ indicates the mean of the $d\ln Gc$ and $d\ln Gs$ PPDFs. The black vertical line in the reconstructed $d\ln G$ and Θ distributions indicates the value of $d\ln G$ and Θ calculated from the mean of the $d\ln Gc$ and $d\ln Gs$ distributions.

Figure S6. (A) Map of mean fast direction calculated from the mean Gc and Gs and (B) map of the mean model and its standard deviation obtained by drawing random samples from the Gc and Gs PPDFs. Both were calculated at 150 km depth. Plate boundaries are represented by thin black lines and continents are in grey.

Figure S7. (A) and (B) are PPDFs for $d\ln Gc$ and $d\ln Gs$ obtained from the NA using the Visser *et al.* (2008b) data and the NA. (C) and (D) are the reconstructed $d\ln G$ and Θ PPDFs after drawing random samples from the $d\ln Gc$ and $d\ln Gs$ PPDFs.

Figure S8. Comparison between our models (black bars for YB17SVaniSVD and red bars for the mean NA model) and the new APM model of Becker *et al.* (2015) shown in dark blue. The standard deviation on the fast axes direction as obtained from the NA is shown in light blue. The anisotropy amplitude is proportional to the length of the red and black bars. The plate boundaries are shown by thin black lines, and continents are delimited by thin grey lines.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.