Title

A Bayesian approach to assess the importance of crustal corrections in global anisotropic surface wave tomography

Journal

Geophysical Journal International, 203(3)

ISSN

0956-540X

Authors

Xing, Z
Beghein, C

Publication Date

2015-12-01

DOI

10.1093/gji/ggv401

Peer reviewed

Geophysical Journal International

Geophys. J. Int. (2015) **203,** 1832–1846
GJI Seismology

doi: 10.1093/gji/ggv401

A Bayesian approach to assess the importance of crustal corrections in global anisotropic surface wave tomography

Z. Xing and C. Beghein*

Department of Earth, Planetary, and Space Sciences, University of California Los Angeles, 595 *Charles Young Drive East, Los Angeles, CA* 90095, *USA.*
E-mail: zxing1989@ucla.edu

Accepted 2015 September 17. Received 2015 September 15; in original form 2015 July 3

SUMMARY

Several recent studies have demonstrated the importance of crustal corrections when inverting surface wave data to model lateral variations in mantle radial anisotropy. It has also been shown that the choice of the prior crustal model to correct the data can strongly influence the anisotropy model and potentially lead to different geodynamic interpretations. In comparing tomographic models of radial anisotropy obtained from different crustal corrections, these studies did not, however, determine quantitative model uncertainties. Nevertheless, mantle models resulting from different prior crustal corrections are statistically different only if the posterior model errors stemming from the non-uniqueness of the inverse problem are smaller than the effect of the crustal correction itself. Here, we applied a model space search approach to global fundamental and higher mode Rayleigh and Love wave phase velocity maps to determine reliable, quantitative model uncertainties on seismic velocities and radial anisotropy. The technique employed enabled us to describe the model space with a posterior probability density function, and therefore to test whether models obtained from different crustal corrections are statistically different. We thus assessed the significance of the choice of the crustal model by comparing the posterior model errors to the differences in mantle structure resulting from different crustal corrections. We tested prior crustal models CRUST2.0, CRUST1.0 and 3SMAC. Our study shows that the use of prior crustal corrections from different crustal models yields significant discrepancies in mantle velocities around 50 km depth and in radial anisotropy down to 100 km. The impact of the crustal correction on radial anisotropy can extend down to 250 km in some locations. We found that choosing 3SMAC instead of the other crustal models has a stronger influence on the mantle model, but that CRUST1.0 and CRUST2.0 yield statistically identical anisotropy models at all depths, except at a few grid cells. Importantly, the effect of the crustal model is most significant in continental regions and not so much beneath oceans, which has important consequences for determining the depth of continental roots. Our results therefore suggest that improving constraints on crustal structure in continents is essential for our understanding of continent formation. Our work also demonstrates that the prior crustal model does not significantly affect radial anisotropy and velocities at depths greater than 100 km. This implies that if geodynamic interpretations of radial anisotropy below 100 km depth were to account for tomographic model uncertainties, they would not depend on the choice of the prior crustal model. It is therefore important for geodynamicists and seismologists to work in concert and to put effort into determining quantitative tomographic model uncertainties before interpreting the results. Our results also caution against the use of 3SMAC to correct surface wave data for studies of the continental lithosphere and suggest that the solid Earth community would benefit from putting some efforts towards building a revised 3SMAC. The discrepancies between mantle models built based on 3SMAC crustal corrections and those based on CRUST1.0 or CRUST2.0 should also help shed light on the validity of the geodynamical assumptions made in the construction of models like 3SMAC.

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

*UCLA.

1 INTRODUCTION

Seismic anisotropy, which is the directional dependence of seismic wave velocity, can be a signal of large-scale mantle deformation (Karato 1998), and is therefore a useful tool to constrain mantle flow patterns. It can be caused by the crystallographic or lattice preferred orientation (CPO or LPO) of intrinsically anisotropic minerals or by the shape preferred orientation (SPO) of isotropic material with contrasting elastic properties, such as layered structures or the alignment of cracks, melt tubules, or lenses (Nicolas & Christensen 1987; Karato 1989; Montagner 1994; Kendall & Silver 1996; Silver 1996; Karato 1998; Long & Becker 2010). In the upper mantle, LPO of olivine in mantle flow is generally considered to be the main source of the observed seismic anisotropy.

Anderson (1961) showed that the discrepancy between Rayleigh and Love wave dispersion curves can only be reconciled if radial anisotropy is present. Radial anisotropy is a particular case of anisotropy that quantifies the velocity difference between the vertical and horizontal directions. In this case, the medium in which the waves propagate has one symmetry axis oriented in the radial direction. Seismic anisotropy has also been observed in the azimuthal dependence of oceanic P_n velocity (Hess 1964) and in Rayleigh wave phase velocities (Forsyth 1975). Radial anisotropy was included in the top 220 km of the 1-D Preliminary Earth Reference Model (Dziewonski & Anderson 1981), thereafter referred to as PREM. Several global models of lateral variations in radial anisotropy have since been published using a variety of data (Nataf *et al.* 1984; Woodhouse & Dziewonski 1984; Nataf *et al.* 1986; Montagner & Tanimoto 1991; Ekström & Dziewonski 1998; Gung *et al.* 2003; Beghein & Trampert 2004; Panning & Romanowicz 2006; Zhou *et al.* 2006; Kustowski *et al.* 2008; Nettles & Dziewoński 2008; Visser *et al.* 2008b; Auer *et al.* 2014; Chang *et al.* 2014; French & Romanowicz 2014; Moulik & Ekström 2014). However, despite great progress in modelling Earth structure and increased computational capabilities, there are still large discrepancies between models of seismic anisotropy. Those inconsistencies arise from differences in data sets, inversion techniques (linear versus nonlinear, waveforms versus dispersion curves inversions), prior constraints, etc. Another source of discrepancies comes from the mapping of crustal structure into the mantle. There are multiple ways to account for crustal structure when modelling mantle velocities and anisotropy. Some researchers invert simultaneously for the Moho depth, mantle structure, and/or crustal velocities (Meier *et al.* 2007; Visser *et al.* 2008b; Burgos *et al.* 2014; Chang *et al.* 2014). Others choose to correct the data using a prior crustal model and invert the remaining signal for mantle velocities and anisotropy using linear (e.g. Ekström & Dziewonski 1998) or nonlinear (e.g. Boschi & Ekström 2002; Beghein & Trampert 2004; Panning & Romanowicz 2006; Marone & Romanowicz 2007; Kustowski *et al.* 2008; Nettles & Dziewoński 2008; Beghein 2010; Ferreira *et al.* 2010; Lekić *et al.* 2010; Panning *et al.* 2010; Auer *et al.* 2014; Moulik & Ekström 2014) crustal corrections.

Some authors have recently used an approach that does not involve any existing crustal model, but that uses a Markov Chain Monte Carlo method instead to construct a smooth crustal model that fits a global data set of fundamental mode dispersion data (Lekić & Romanowicz 2011; French *et al.* 2013; French & Romanowicz 2014). The new crustal model is then used directly in the computation of predicted waveforms. This latter technique has the advantage of not biasing the modelling towards specific crustal models, which could be a source of concern in other studies. All of the studies cited above do, indeed, include some prior information based on a pre-existing crustal model before inverting their data, which could influence the mantle model. For instance, CRUST5.1 (Mooney *et al.* 1998) was used as a prior crustal model by Ekström & Dziewonski (1998), Boschi & Ekström (2002), and Beghein & Trampert (2004), and its updated version, CRUST2.0 (Bassin *et al.* 2000), has been widely employed either to correct the data a priori or as a starting model (Panning & Romanowicz 2006; Zhou *et al.* 2006; Marone & Romanowicz 2007; Meier *et al.* 2007; Kustowski *et al.* 2008; Nettles & Dziewoński 2008; Visser *et al.* 2008b; Beghein 2010; Lekić *et al.* 2010; Panning *et al.* 2010; Auer *et al.* 2014; Burgos *et al.* 2014; Chang *et al.* 2014). Model 3SMAC (Nataf & Ricard 1996) has also been used in a few regional studies (e.g. Silveira *et al.* 2006; Debayle *et al.* 2001; Montagner 2002; Yoshizawa & Kennett 2004).

Ferreira *et al.* (2010) showed, however, that the choice of the prior crustal model used to correct surface wave dispersion data has an impact on the data fit comparable to or larger than that obtained by including or neglecting 3-D variations in radial anisotropy in the inversion. The authors also showed that while the crustal model does not strongly affect the global distribution of shear velocity anomalies in global models, it can strongly bias models of shear wave radial anisotropy, especially around 100 km depth, suggesting different geodynamical interpretations. Differences between mantle models obtained with different prior crustal corrections are, nevertheless, significant only if the affected features are well-resolved, that is, if these differences are larger than the intrinsic model uncertainties.

The models obtained by Ferreira *et al.* (2010) resulted from regularized inversions. Geophysical inverse problems are inherently non-unique. One usually chooses one 'best' or preferred model among many by imposing a subjective regularization and minimizing a cost function that compromises between model size and data fit. This reduces the number of possible solutions and helps avoiding unrealistically large models (Jackson 1979; Trampert 1998). Regularizing the problem is equivalent to adding prior information to the data and therefore affects the resulting models (Jackson 1979). The resolution of the inverted model can be tested by performing checkerboard tests and a resolution matrix can be calculated. A resolution matrix is, however, function of the regularization and parametrization applied and is thus not ideal to evaluate parameter trade-offs. In addition, the stability of tomographic models can be tested qualitatively by performing a few inversions with different dampings, different parametrizations, or different prior information, but it does not necessarily reflect the true model uncertainties. These can indeed be strongly underestimated in traditional inversions if the statistics are not Gaussian as usually assumed, or if the model null-space is large, which could result in a cost function with a larger valley (Trampert 1998).

The objective of this study was to assess the significance of the discrepancies between surface wave derived models of radial anisotropy obtained using different *a priori* crustal models. This can be done by determining quantitative mantle model uncertainties for models resulting from different prior crustal corrections, and comparing differences between the models to the intrinsic posterior model uncertainties. For this purpose, we applied the Neighbourhood Algorithm (NA; Sambridge 1999a,b), thereafter referred to as NA, to fundamental and higher mode surface wave phase velocities. The NA is a model space search approach that enabled us to forward model lateral variations in shear-wave velocities and radial anisotropy in the mantle and to obtain more reliable, quantitative mantle model uncertainties than one would with a traditional inverse approach.

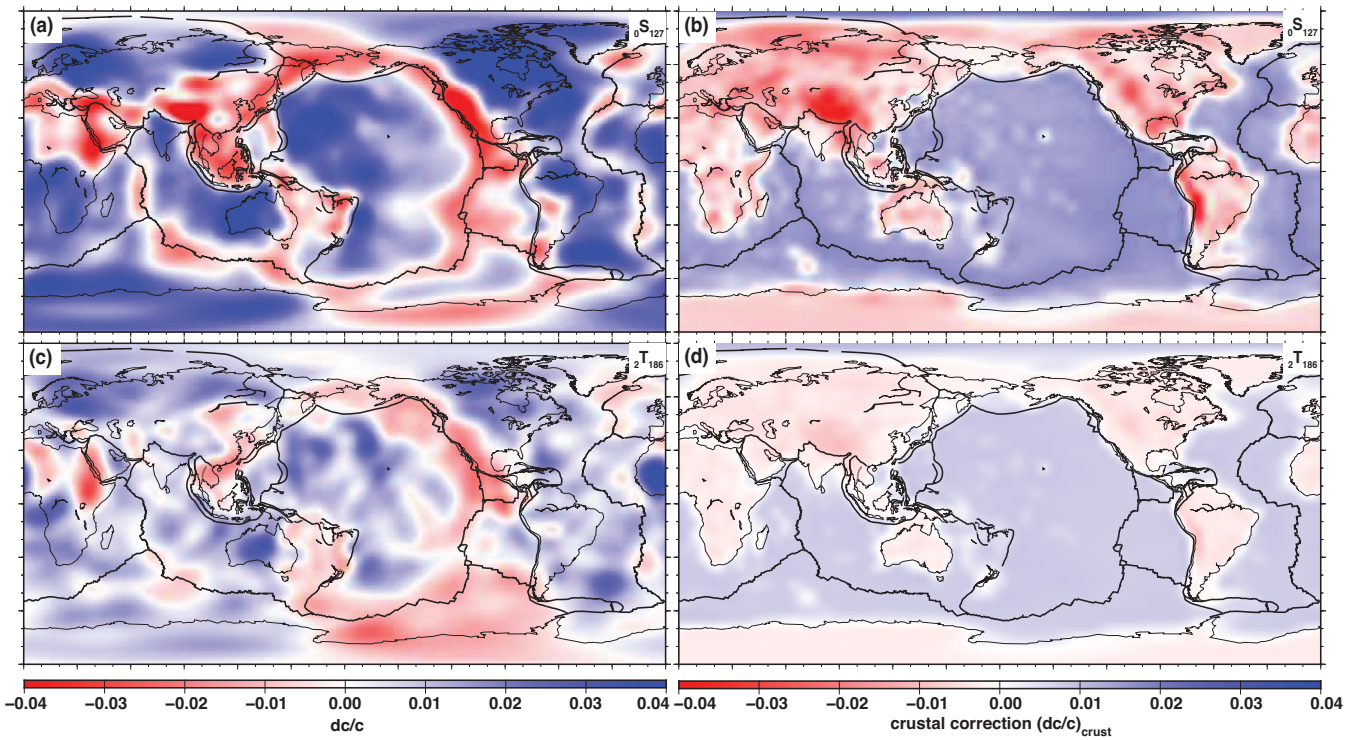

Figure 1. (a) Fundamental mode Rayleigh wave phase velocity maps at 78 s (Visser *et al.* 2008a) and (b) corresponding crustal correction using CRUST2.0 (Bassin *et al.* 2000); (c) second overtone Love wave phase velocity maps at 40 s and (d) corresponding crustal correction.

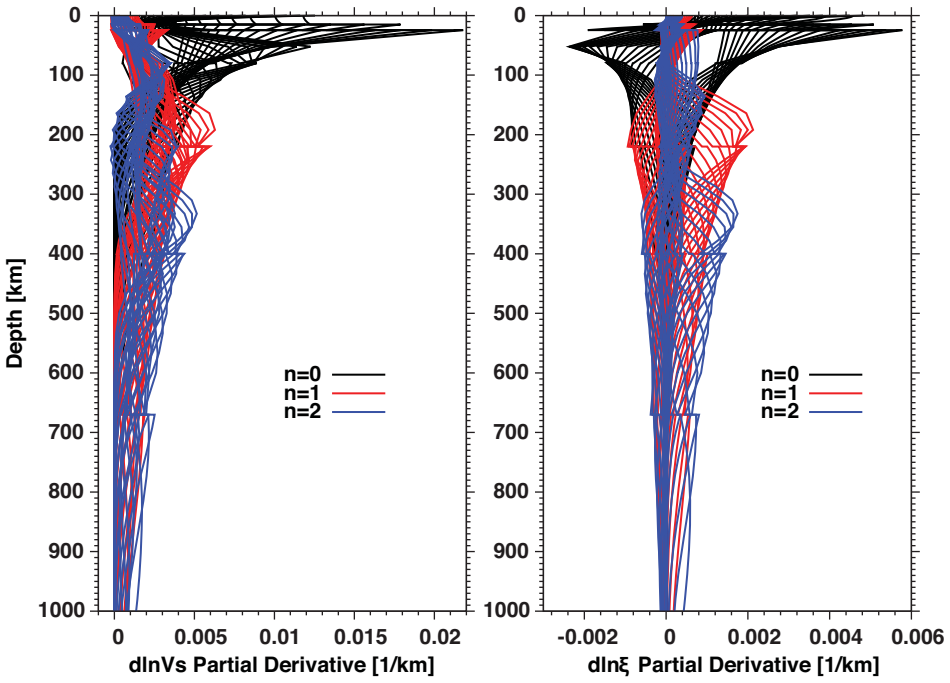

Figure 2. $d \ln V_s$ (left) and $d \ln \xi$ (right) sensitivity Kernels for the Rayleigh and Love Wave fundamental modes and first two overtones used in this study.

2 DATA

In this study, we used the isotropic part of the Rayleigh and Love wave phase velocity maps obtained by Visser *et al.* (2008a). Because we wanted to limit this study to the upper mantle, depths at which crustal corrections are most likely to impact the model, we used phase velocity maps for fundamental modes and the first two overtone only. Figs 1(a) and (c) show examples of these phase velocity maps for fundamental mode Rayleigh waves and the second Love wave overtone at a period of 78 s. Fig. 2 demonstrates that this data set has great sensitivity to Earth structure in the upper 400 km of the mantle, thereby improving model resolution in the uppermost mantle compared to studies that employ fundamental mode surface waves only (Beghein *et al.* 2014).

Higher mode surface waves are essential to model the deep upper mantle (below 150 km depth), but they can be difficult to measure and to separate from the fundamental modes (e.g. Visser 2008). This is especially true for Love waves in oceanic settings since the overtones and the fundamental modes have similar group velocities, which implies that they arrive almost simultaneously in the seismogram (Nettles & Dziewoński 2011). Visser *et al.* (2008a) measured dispersion curves for fundamental and higher mode surface wave phase velocities using the fully nonlinear waveform inversion method of Yoshizawa & Kennett (2002), which they combined with a model space search approach (Sambridge 1999a,b) to determine consistent and quantitative dispersion measurement uncertainties. This technique also allowed them to test whether the different modes were likely to be well separated from one another and to reject the ones that were not. This yielded dispersion curves for over 60 000 fundamental mode ($n = 0$) Rayleigh waves, over 50 000 first and second ($n = 1, 2$) Rayleigh wave overtones, over 45 000 Love wave fundamental modes, and between 30 000 and 35 000 Love wave higher modes for $n = 1, 2$.

Their path-averaged measurements were then inverted to obtain global azimuthally anisotropic phase velocity maps. Relative perturbations in surface wave phase velocities in a slightly anisotropic medium can be expressed as (Montagner & Nataf 1986):

$$\delta c/c_0(\omega, \Psi) = \alpha_0(\omega) + \alpha_1(\omega)\sin(2\Psi) + \alpha_2(\omega)\cos(2\Psi) + \alpha_3(\omega)\sin(4\Psi) + \alpha_4(\omega)\cos(4\Psi), \quad (1)$$

where ω and Ψ are the angular frequency and azimuth of propagation of the surface waves, and $\delta c/c_0$ is the relative phase velocity perturbation with respect to that predicted in a 1-D Earth reference model. α_0 is the average phase velocity perturbation over all azimuths and α_i ($i = 1, \ldots, 4$) are anisotropic terms that describe the dependence on the azimuth of the phase velocity. The data used in this study consist in the isotropic part (α_0) of the phase velocity maps. This term depends on elastic coefficients that characterize radial anisotropy at depth, as described in the Method section.

There are a number of factors that can affect the resolution of phase velocity maps, such as the data path azimuthal coverage, the applied regularization, and spectral leakage. Visser *et al.* (2008a) showed that their fundamental modes had generally good coverage which greatly minimized the trade-offs between isotropic and anisotropic terms of eq. (1). The authors also calculated resolution matrices and showed that little isotropic structure is mapped into anisotropic terms and vice versa. Though a decrease in the number of measurements was observed with increasing overtone number, the path coverage of the higher modes was very good as well, especially in the northern hemisphere. The damping chosen by the authors was based on the study of Trampert & Woodhouse (2003) and was such that the relative model uncertainty remains the same for all modes, resulting in phase velocity maps of decreasing resolution with increasing overtone number. They determined that on an average the isotropic term of the fundamental mode phase velocity maps have a resolution comparable to that of a spherical harmonic expansion of degree order 25, yielding a lateral resolution of about 1600 km in the uppermost mantle. The isotropic term of the higher modes has a resolution similar to that of degree and order 18, allowing for about 2200 km of lateral resolution in the deep upper mantle and mantle transition zone. In order to deal with spectral leakage, which is the mapping of small-scale structure into the low-degree structure due to uneven data coverage (Trampert & Snieder 1996), Visser *et al.* (2008a) applied a stronger damping to higher spherical harmonic degree order structure. This tends to reduce the resolvability of higher order structures, but since we focus on comparing the effects of different prior crustal models on the long wavelength features of the mantle, this and the varying resolution with depth should not affect our conclusions significantly.

3 METHOD

3.1 Forward problem and parametrization

In this work, we used the isotropic terms α_0 of the Love and Rayleigh wave phase velocity maps (eq. 1) to constrain 3-D radial anisotropy and wave velocity in the upper mantle. At any given point at the surface of the Earth, α_0 is the vertical average of the underlying elastic structure weighted by sensitivity kernels (Woodhouse & Dahlen 1978; Montagner & Nataf 1986):

$${}_k\alpha_0 = \int_0^a {}_k\mathbf{K}_m(r)\delta \ln \mathbf{m}(r)\, \mathrm{d}r. \quad (2)$$

Subscript k represents different modes and different frequencies ω. $\delta \ln \mathbf{m}(r)$ represents relative perturbations in parameter $\mathbf{m}$ with respect to a reference model, and a stands the radius of the Earth. ${}_kK_m(r)$ are the partial derivatives or sensitivity kernels for model parameter $\mathbf{m}$. They can be calculated for any 1-D Earth models using normal mode theory (Takeuchi & Saito 1972) and vary from mode to mode. Note that, as opposed to mode-based waveform inversion techniques, an implicit assumption of this method is that there is no mode coupling, which is an approximation since we know that lateral heterogeneities and anisotropy can cause coupling between some modes (e.g. Beghein *et al.* 2008).

In an isotropic medium, the model parameter vector $\mathbf{m}$ contains density ρ and two elastic coefficients: the Lamé parameter λ and the shear modulus μ, or alternatively the bulk modulus κ and the shear modulus. In a general anisotropic medium, a total number of 21 independent parameters is needed to fully describe the fourth-order elastic stiffness tensor. Because seismic data cannot resolve all 21 parameters, we make assumptions regarding the symmetry of the anisotropic medium to reduce the number of unknowns. In the simple case of radial anisotropy, the medium has hexagonal symmetry with a vertical symmetry axis and can be described by five independent parameters: the velocity of vertically polarized shear waves V_{SV}, of horizontally polarized shear waves V_{SH}, of vertically and horizontally propagating P-waves V_{PV} and V_{PH}, respectively, and parameter η which describes wave propagation at an intermediate angle. A parametrization in terms of the Love elastic parameters A, C, N, L and F (Love 1927) is sometimes employed (Beghein 2010; Chang *et al.* 2014), with $A = \rho V_{PH}^2$, $C = \rho V_{PV}^2$, $N = \rho V_{SH}^2$, $L = \rho V_{SV}^2$, $F = (A - 2L)\eta$. Here, we chose instead to follow Panning & Romanowicz (2006) and used another equivalent parametrization:

$$V_P^2 = \left(V_{PV}^2 + 4V_{PH}^2\right)/5 \quad (3)$$

$$V_S^2 = \left(2V_{SV}^2 + V_{SH}^2\right)/3 \quad (4)$$

$$\phi = V_{PV}^2/V_{PH}^2 \quad (5)$$

$$\xi = V_{SH}^2/V_{SV}^2 \quad (6)$$

$$\eta = F/(A - 2L), \quad (7)$$

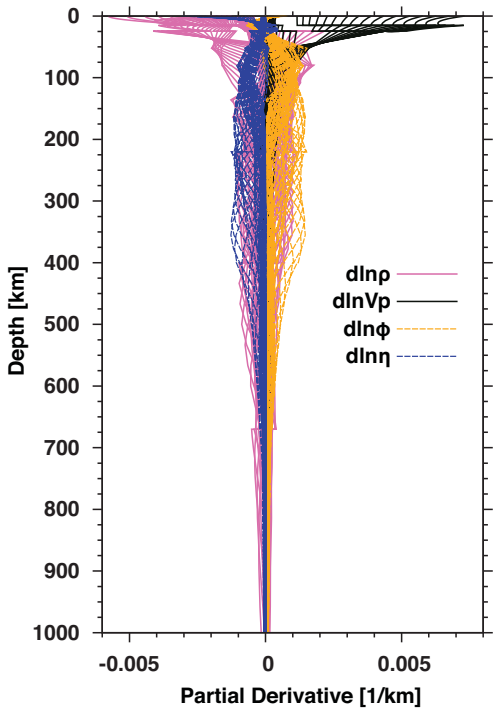

Figure 3. Sensitivity Kernels for Rayleigh and Love Wave fundamental modes and first two overtones calculated used in this study. The kernels were calculated for $d\ln V_P$, $d\ln\phi$, $d\ln\eta$ and $d\ln\rho$.

where V_P and V_S are Voigt average isotropic velocities representing upper bounds on the effective elastic moduli (Babuska & Cara 1991). ϕ and ξ represent P-wave and S-wave radial anisotropy, respectively. This parametrization enables us to directly invert for anisotropy, avoiding the possible roughness in anisotropic structure if it is calculated *a posteriori* from separate inversions of V_{SV} and V_{SH} (Nettles & Dziewoński 2008). Eq. (2) therefore becomes

$$
{}_k\left(\frac{\delta c}{c_0}\right) = \int_0^a \big[{}_kK_{V_P}(r)\delta\ln V_P(r) + {}_kK_{V_S}(r)\delta\ln V_S(r) + {}_kK_{\phi}(r)\delta\ln\phi(r) + {}_kK_{\xi}(r)\delta\ln\xi(r) + {}_kK_{\eta}(r)\delta\ln\eta(r) + {}_kK_{\rho}(r)\delta\ln\rho(r)\big]\,\mathrm{d}r. \tag{8}
$$

Not all of these six parameters are, however, well resolved by surface wave data because of reduced sensitivity to V_P, ϕ, η and ρ and because of parameter trade-offs (Fig. 3).

For this reason, we used the empirical relationships established by Montagner & Anderson (1989) from petrological considerations to scale perturbations in compressional wave velocity and density to changes in shear wave velocity, and to scale changes in compressional wave anisotropy and in η to perturbations in shear wave anisotropy:

$$
\frac{\delta\ln V_P}{\delta\ln V_S} = 0.5, \tag{9}
$$

$$
\frac{\delta\ln\rho}{\delta\ln V_S} = 0.33, \tag{10}
$$

$$
\frac{\delta\ln\phi}{\delta\ln\xi} = -2.5, \tag{11}
$$

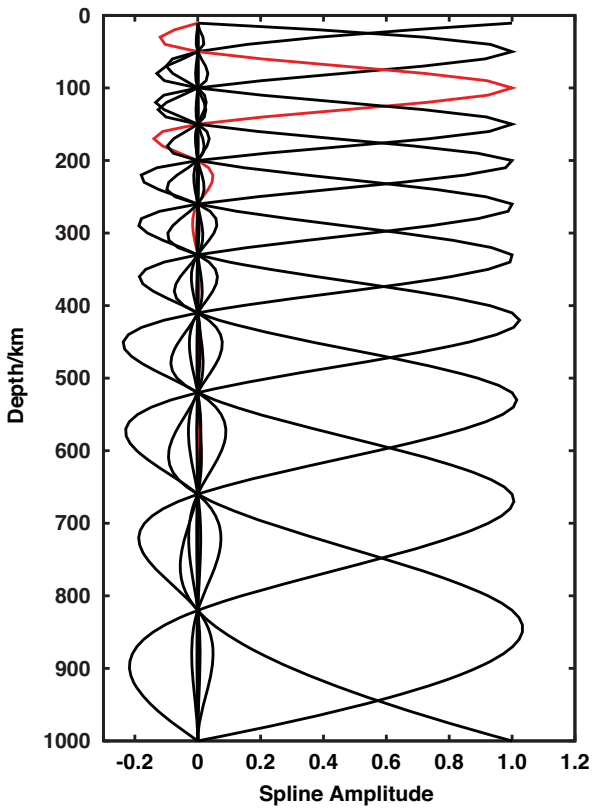

Figure 4. Spline functions used for the depth parametrization. The interval increases with depth to account for the decrease in sensitivity. The red curve highlights the spline that peaks at 100 km depth.

$$
\frac{\delta\ln\eta}{\delta\ln\xi} = -1.5. \tag{12}
$$

These relations were derived for the top 400 km of the mantle and were shown to not strongly affect the resulting mantle models (Beghein 2010). However, the data used here are sensitive to structure at greater depths where these scaling relationships may not be valid, as suggested by Beghein & Trampert (2004). Previous studies have also suggested that mantle models of radial anisotropy can be affected by these prior relationships in the lower transition zone and uppermost lower mantle (Panning & Romanowicz 2006). This should not, however, affect the conclusions of this paper since we focus on shallower structure, but should be kept in mind for future work involving deeper structure.

We parametrized the Earth horizontally by dividing its surface into $10^\circ \times 10^\circ$ cells, and conducted the inversions of eq. (8) for each cell separately. At every grid cell, we used a depth parametrization for $\delta\ln V_S(r)$ and $\delta\ln\xi(r)$ composed of 12 cubic spline function $S_i(r)$ ($i = 1,\ldots,12$) (Fig. 4) down to 1000 km:

$$
\delta\ln V_S(r) = \sum_{i=1}^{12} \delta\ln V_S^i S_i(r) \tag{13}
$$

$$
\delta\ln\xi(r) = \sum_{i=1}^{12} \delta\ln\xi^i S_i(r), \tag{14}
$$

where $\delta\ln V_S^i$ and $\delta\ln\xi^i$ are the coefficients of the ith spline. Note that the top of the splines is defined by the local Moho instead of Earth's surface, and they are thus laterally varying. The spacing between the splines ranges from 30 km in the uppermost mantle

to more than 100 km at larger depth to reflect the variable depth sensitivity of the data.

Eq. (8) can be written as:

$$\mathbf{Gm} = \mathbf{d}, \tag{15}$$

where vector **d** represents the data vector, and **m** is the vector containing the model parameters, that is, the $\delta \ln V_S^i$ and $\delta \ln \xi^i$ of eqs (13) and (14). Matrix **G** is the kernel matrix and its elements G_{ki} can be calculated using the prior constraints of eqs (9)–(12) and by integrating the sensitivity kernels projected onto the spline functions $S_i(r)$:

$$G_{ki} = \int_0^a [{}_kK_{V_S}(r) + 0.5\,{}_kK_{V_P}(r) + 0.33\,{}_kK_{\rho}(r) + {}_kK_{\xi}(r) - 2.5\,{}_kK_{\phi}(r) - 1.5\,{}_kK_{\eta}(r)]S_i(r)\,\mathrm{d}r. \tag{16}$$

At each grid cell, we have 47 Rayleigh and 45 Love wave phase velocities, and 24 model parameters. Matrix **G** has thus 92 rows and 24 columns.

3.2 Crustal corrections and prior crustal models

Surface wave phase velocities have high sensitivity to crustal structure and crustal thickness. It is thus of great importance to correctly account for the effect of the crust on the measurements to minimize the mapping of the crust into the mantle model as it could affect the interpretation of the results (Boschi & Ekström 2002; Panning *et al.* 2010). This is especially important when modelling radial anisotropy because Love waves are sensitive to shallower structure than Rayleigh waves at the same period, and incorrectly accounting for lateral variations in crustal structure could bias the anisotropy mantle model (Ferreira *et al.* 2010). It can be done by constructing a prior crustal model that fits global fundamental mode surface wave dispersion data, and using this model directly in the computation of predicted waveforms (Lekić & Romanowicz 2011; French *et al.* 2013; French & Romanowicz 2014). Other techniques involve inverting crustal thickness and/or structure simultaneously with mantle structure (Visser *et al.* 2008b), or correcting dispersion data with predictions from a prior crustal model before inverting those data (Boschi & Ekström 2002; Kustowski *et al.* 2007; Marone & Romanowicz 2007; Ferreira *et al.* 2010; Panning *et al.* 2010), or using a hybrid two-step method (Burgos *et al.* 2014; Chang *et al.* 2014).

In this work, we tested the effect of three different crustal models: CRUST2.0 (Bassin *et al.* 2000), CRUST1.0 (Laske *et al.* 2013) and 3SMAC (Nataf & Ricard 1996). CRUST1.0 is an updated, higher resolution (1° × 1°) version of CRUST2.0, which was defined on a 2° × 2° grid and is itself an update from the 5° × 5° resolution model CRUST5.1 (Mooney *et al.* 1998). In both CRUST2.0 and CRUST1.0, the crust is represented as a stack of layers (ice or water, sediments and crystalline crust) with assigned thicknesses as well as *P*- and *S*-wave velocities and density. Differences between those two models include the use of group velocity data in CRUST1.0 to better constrain crustal velocities, and new Moho depth estimates in locations where we now have constraints from recent active source seismic and receiver function studies. 3SMAC, which stands for '3-Dimensional Seismological Model A Priori Constrained', is a model of the upper mantle and the crust. It does not only incorporate seismic observations, but is also based on the geodynamical interpretation of near surface layers of the Earth by including measurements of surface heat flux, isostatic topography and hot spot distributions. Density and velocities in 3SMAC were calculated from models of temperature and pressure profiles with depth. 3SMAC was built, however, from a much lower amount of active source seismic data than CRUST2.0 and CRUST1.0.

Fig. 5 shows the Moho depth in each of the three crustal models as well as the difference between the Moho depth in each pair of crustal model. Even though these Moho depths may look similar at first glance, they can differ by as much as 10 km in some locations, mainly in continental regions. Since the Moho depth affects surface wave phase velocities, it is expected that the differences among those crustal models will result in differences in the mantle models. There are also discrepancies between crustal velocities in the different models as can be seen for instance in Fig. 6. They can affect the crustal corrections as well.

We calculated nonlinear crustal corrections following a procedure similar to that of Boschi & Ekström (2002). We first constructed a 3-D reference model composed of the PREM (Dziewonski & Anderson 1981) mantle and a 3-D crustal model. We then used the local 1-D model at each grid cell j with computer program MINEOS to calculate local, theoretical phase velocities ${}_kc_j$ and local eigenfunctions, which were then used to determine local sensitivity kernels ${}_kK_{m,j}$. In this notation, $k = (n, T)$ stands for the mode with overtone number n and period T and m represents a model parameter. At each grid cell j, the difference between the phase velocity ${}_kc_j$ predicted by the 3-D model and that calculated from PREM $c_0(k)$ is the contribution of the crustal model to the measured phase velocity perturbation ${}_kdc_{\mathrm{crust},j}$, that is, it is the crustal correction. After removing this correction from the measured phase velocity anomaly ${}_kdc_j$, we are left with a corrected signal ${}_k(\frac{dc_c}{c})_j$ that represents the phase velocity anomaly due to a local mantle perturbation with respect to the reference model at grid cell j. This remaining signal can thus be inverted to infer local perturbations in mantle structure with respect to PREM using the local sensitivity kernels:

$${}_k\left(\frac{\delta c_c}{c}\right)_j = \int_0^{\mathrm{Moho}} \sum_i \left[{}_kK_{m_i,j}(r)\delta \ln m_{i,j})(r)\right] \mathrm{d}r, \tag{17}$$

where the sum over i is carried out over all model parameters.

Note that this method is only strictly valid if there is no mode coupling since it assumes that the effect of crustal structure can be computed independently at each location on the Earth's surface. Examples of crustal corrections using CRUST2.0 Bassin *et al.* (2000) for fundamental mode Rayleigh waves and the second Love wave overtone at a period of 78 s are shown in Fig. 1, and demonstrate that the crust contributes significantly to the measurements.

3.3 Modelling

As explained above, regularized inversions do not, however, always provide reliable quantitative model uncertainties. An alternative approach can be found in direct search techniques. Without needing to implicitly introduce strong *a priori* information, this type of method is able to explore the model space, including null-space, giving us a better description of the range of possible models instead of just one 'best'. Here we utilized a two-part global optimization technique, namely the NA (Sambridge 1999a,b).

The first part of NA is a direct search of the model space. First, a number of samples are randomly generated in the model space and the model space is divided into Voronoi cells using this initial sampling. Each cell is centred on one of the models by construction. An approximate misfit surface is determined by calculating the misfit of each of these models. At each subsequent iteration, n_s

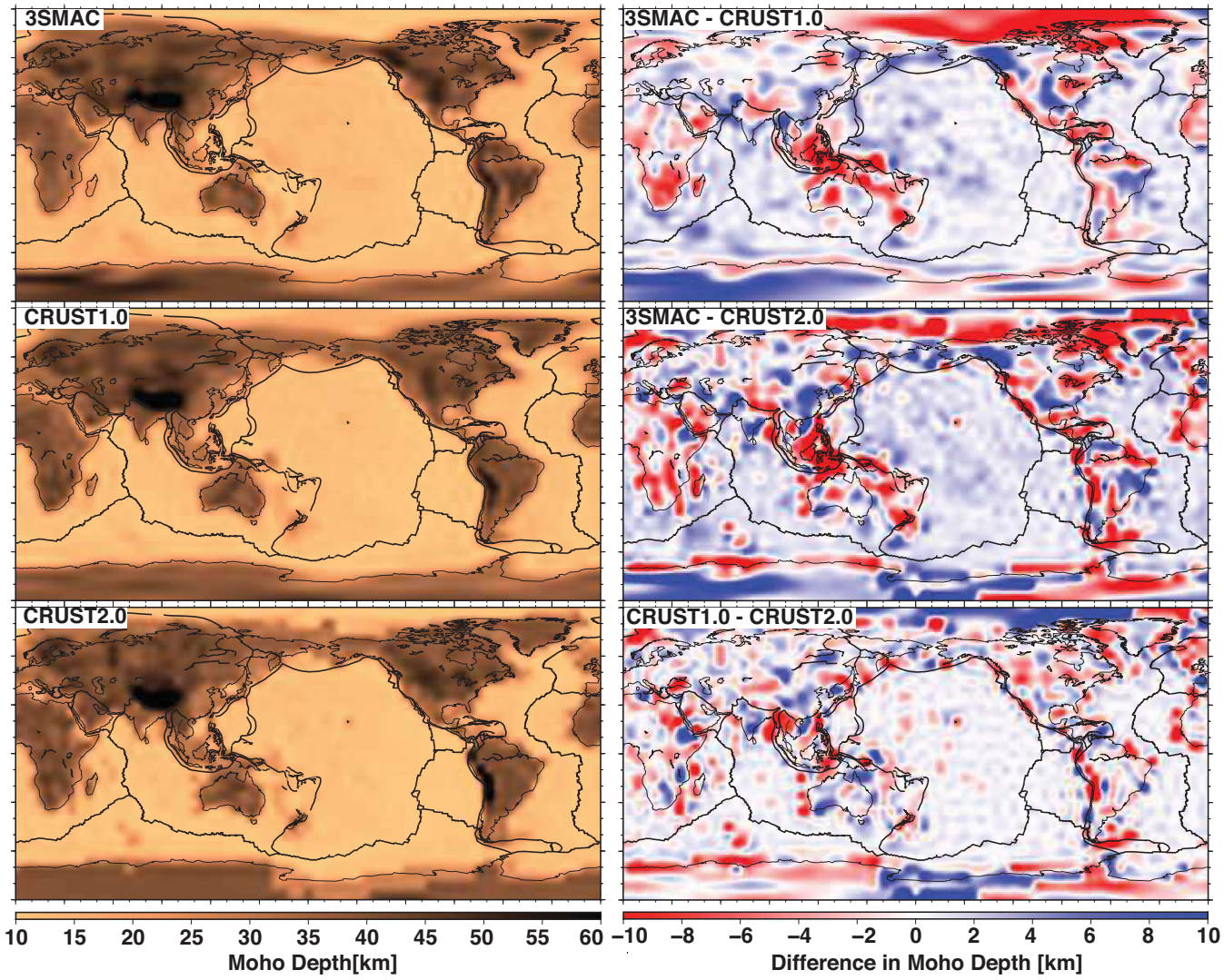

Figure 5. Moho depths (left) for 3SMAC, CRUST1.0 and CRUST2.0 and differences (right) between each pair of models.

new samples are then generated by performing a uniform random walk with a Gibbs sampler in the Voronoi cell of each of the n_r best-fitting models. The Voronoi cells are updated to accommodate the newly generated models, and misfits are calculated for those new models, determining a new misfit surface. Misfits are then ranked among all the existing models to determine the next set of n_r best-fitted models. At each iteration, the sampling density therefore increases in the neighbourhoods of the better data-fitting models, and the NA is able to use the information contained in the previous models to adapt the sampling. The choice of n_r and n_s is generally decided by trial and error to control the convergence speed and the sampling quality. A small n_s/n_r ratio implies a slower convergence of the algorithm but helps perform a sampling of the model space as thorough as possible to avoid getting trapped in a local minimum. The iteration is stopped by the user when the misfit does not show any clear decrease with each iteration.

In this work, we first carried out regularized inversions of the surface wave data using eq. (15) and the singular value decomposition (SVD) method of Matsu'ura & Hirata (1982). This technique is described in details in Yuan & Beghein (2014). We then ran the first part of the NA using a uniform prior model distribution around the results of the regularized inversion: each of the 24 model parameters were allowed to vary by 5 per cent with respect to the reference model starting from the inversion results. Searching the model space around the inversion results has the advantage of accelerating the sampling convergence if the model distribution is close to a Gaussian while still allowing the NA to find good data fitting models away from the inversion results.

In the second part of the NA, a Bayesian appraisal of the ensemble of models obtained from sampling the model space 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. The low and high misfits of the models are converted to high and low likelihoods, respectively. Assuming Gaussian-distributed errors on the measurements, the likelihood function is defined as:

$$P(\mathbf{d}|\mathbf{m}) = \exp\left[-\frac{1}{2}(\mathbf{d}-\mathbf{g}(\mathbf{m}))^T\mathbf{C}_\mathrm{D}^{-1}(\mathbf{d}-\mathbf{g}(\mathbf{m}))\right], \quad (18)$$

where $\mathbf{g}(\mathbf{m})$ are phase velocity predictions calculated from model $\mathbf{m}$ and eq. (15), and $\mathbf{C}_\mathrm{D}$ is the data covariance matrix. Here, the data covariance matrix is a diagonal matrix whose elements are given by the data variance estimated by Visser *et al.* (2008a) for

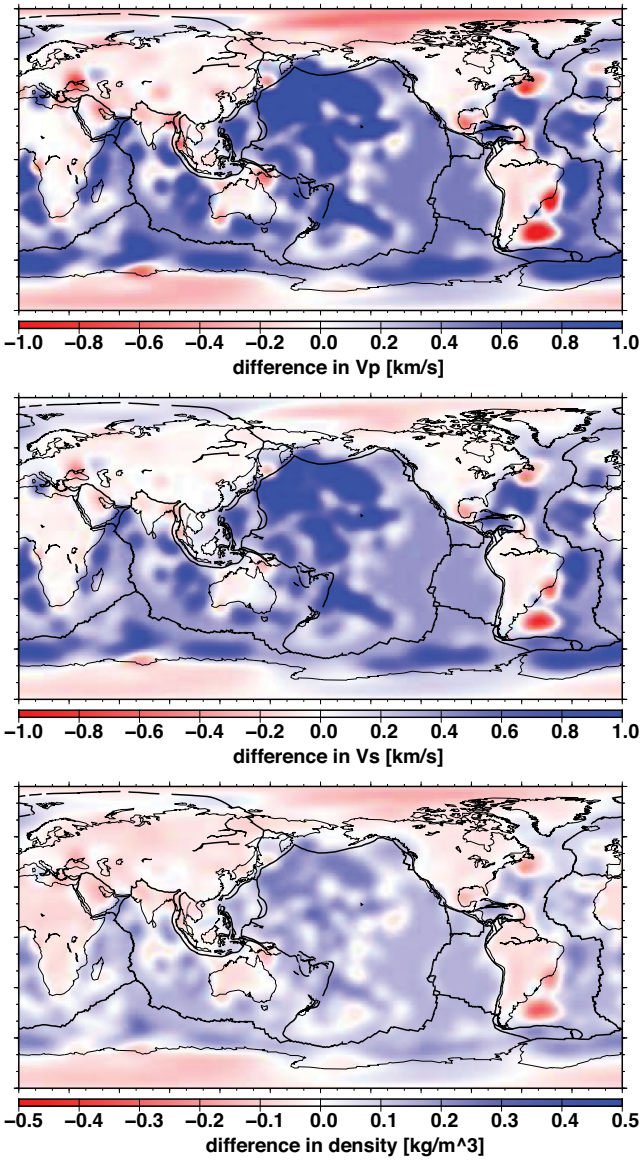

Figure 6. Example of difference in velocities and density between CRUST1.0 and CRUST2.0 at 7 km depth.

each fundamental and higher mode. The likelihood function $P(\mathbf{d}|\mathbf{m})$ is then used to obtain the posterior probability density functions (PPDFs) given by Bayes' theorem (Bayes & Price 1763):

$$P(\mathbf{m}|\mathbf{d}) \propto \rho(\mathbf{m})P(\mathbf{d}|\mathbf{m}), \tag{19}$$

where $P(\mathbf{m}|\mathbf{d})$ is a likelihood function representing the fit to the data, and $\rho(\mathbf{m})$ is the prior probability density distribution, defined here as uniform distributions around the inversion results. These PPDFs can be used to assess the robustness of the model parameters as they can be used to calculate the mean values and uncertainties of the model parameters, as well as covariance matrix and resolution matrix.

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

$$\langle m_i \rangle = \int m_i P(\mathbf{m}|\mathbf{d})\, \mathrm{d}\mathbf{m}. \tag{20}$$

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

$$C_{i,j} = \int m_i m_j P(\mathbf{m}|\mathbf{d})\mathrm{d}(\mathbf{m}) - \langle m_i \rangle \langle m_j \rangle. \tag{21}$$

Note that the concepts of covariance matrix and the resolution kernels are linearized concepts, and are most useful if the PPDF has a single dominant peak, for example, if the distribution is Gaussian. They are used here to get quantitative uncertainties on the model parameters, but one should keep in mind the limitations of these definitions. Because the null-space is included in the model space sampling, the model uncertainties inferred are more accurate than those resulting from regularized inversions which tend to underestimate posterior variance, especially in the presence of a large model the null-space (Trampert 1998; Beghein & Trampert 2003; Beghein 2010).

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

$$M(m_i) = \int \ldots \int P(\mathbf{m}|\mathbf{d}) \prod_{k=1,k\neq i}^{d} \mathrm{d}m_k, \tag{22}$$

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 $\langle m_i \rangle$ coincides with the peak of the distribution, that is, the most likely value. Information about parameter trade-offs can be obtained from the off-diagonal terms of the posterior covariance matrix, and from the 2-D marginal distributions, which are calculated by integrating $P(\mathbf{m}|\mathbf{d})$ over all but two parameters. The 2-D marginal PPDF for the ith and jth parameters is given by:

$$M(m_i, m_j) = \int \ldots \int P(\mathbf{m}|\mathbf{d}) \prod_{k=1,k\neq i,k\neq j}^{d} \mathrm{d}m_k \tag{23}$$

Examples of 1-D and 2-D marginals are given in Fig. 7.

We note that some parameters displayed in Fig. 7 have their posterior truncated at the edge of the model space. One might therefore wonder whether searching a wider parameter space than the 5 per cent we imposed would help better characterize the model space and yield different 'best' models and posterior variances. Our choice of choosing a prior uniform distribution between −5 per cent and +5 per cent of the starting model was motivated by two main factors: computational costs (exploring a larger model space is more time-consuming), and the need to remain within the framework of perturbation theory, which would be violated if we increased the range of the search substantially. Besides, as explained by Beghein & Trampert (2002), one has to keep in mind that there are parameter trade-offs. While increasing the sampling range for one or multiple parameters may have the effect of reducing the minimum misfit somewhat, the trade-offs among the model parameters will cause the global minimum of other parameters to move as well. And they may, in turn, be directed towards the edge. Therefore, as long as there are trade-offs between model parameters, the model space cannot be surveyed completely guaranteeing that no solution is on the edge of the model space. This is, fortunately, not a major problem because we can quantify these trade-offs.

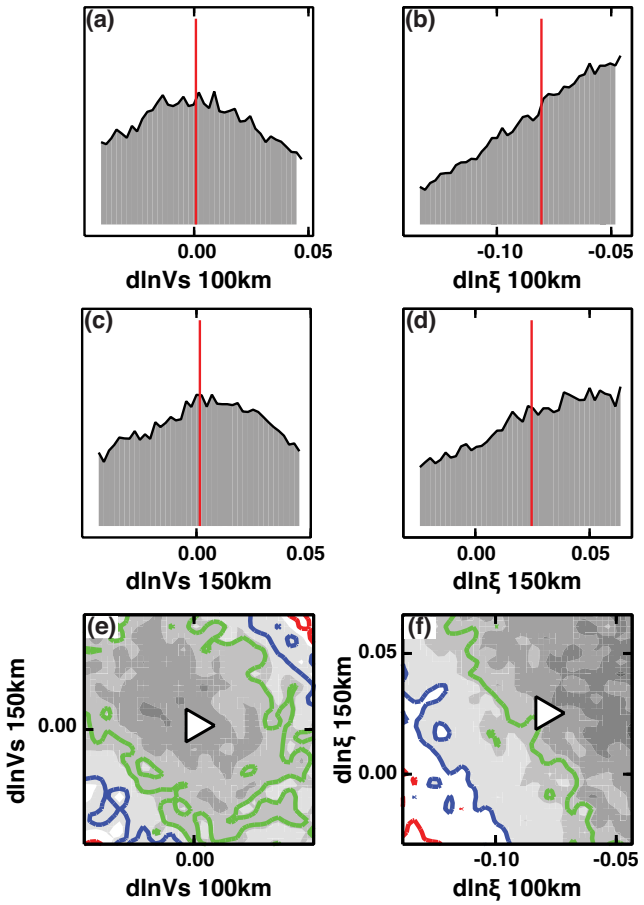

Figure 7. Examples of 1-D and 2-D marginal distributions at a grid cell located in the older part of the Pacific ocean at 191° longitude and −5° latitude. The vertical red lines in the 1-D marginal plots and the triangles in the 2-D marginals represent the mean model (eq. 20). The confidence levels are for 68 per cent (green), 95 per cent (blue), and 99.7 per cent (red). Panels (a) and (c) represent the PPDFs for the $d\ln V_S$ spline coefficients of the splines that peak at 100 and 150 km depth, respectively. Panels (b) and (d) are for relative perturbations in anisotropy. Panels (e) and (f) represent the 2-D marginals for two pairs of parameters. They show large uncertainties and clear trade-offs are also visible between $d\ln\xi$ at 100 km and $d\ln\xi$ at 150 km.

4 RESULTS AND DISCUSSION

Figs 8 and 9 show the mean shear velocity and mean radial anisotropy models, respectively, obtained from the NA (eq. 20) using CRUST1.0, CRUST2.0 and 3SMAC to correct the data. There is a good agreement between the structure of the different mean models for $d\ln V_S$ at all depths, and they are consistent with previously published models (Panning & Romanowicz 2006; Zhou *et al.* 2006; Kustowski *et al.* 2008; Nettles & Dziewoński 2008; Ferreira *et al.* 2010; Auer *et al.* 2014; Chang *et al.* 2014; French & Romanowicz 2014; Moulik & Ekström 2014). The amplitude of $d\ln V_S$ anomalies is lower in the deep upper mantle than at shallow depths. In all three velocity models, we observe low shear velocities associated with the location of mid-ocean ridges in the uppermost mantle. Positive velocity anomalies are associated with continents and negative anomalies with oceanic regions in the shallow mantle. This large contrast in $d\ln V_S$ between oceans and continents disappears around 250 km. This may coincide with the depth of continental roots if we use the change in sign of the velocity anomalies as a proxy, and is consistent with previous estimates from anisotropy (Gung *et al.* 2003; Yuan & Romanowicz 2010; Yuan & Beghein 2014).

The three radial anisotropy models are also generally consistent with one another below 100 km (Fig. 9). Strong differences are, however, observed at 50 km depth and, to some extent, at 100 km. The model obtained with 3SMAC displays $V_{SH} > V_{SV}$ ($\xi > 1$) at 50 km in the central Pacific and $V_{SV} > V_{SH}$ ($\xi < 1$) at ocean ridges and subduction regions where we expect vertical flow (though one should keep in mind that short scale vertical layering could also cause this signal (Bodin *et al.* 2015)). This is fundamentally different from the results of Kustowski *et al.* (2008) or Nettles & Dziewoński (2008) who found a clear $V_{SV} > V_{SH}$ signal in the middle of the Pacific surrounded by $V_{SH} > V_{SV}$ at that depth. Our 3SMAC-derived model at 50 km is also different from the models of Ferreira *et al.* (2010) for all the crustal models they tested, and most resembles the models of Panning & Romanowicz (2006) and of Visser *et al.* (2008b). Fig. 9 shows that CRUST1.0 also yields $\xi > 1$ at 50 km depth in the middle of the Pacific, but the model is more heterogeneous than with 3SMAC. The model resulting from crustal corrections made with CRUST2.0 shows even more lateral variations. The anisotropy found in continental areas at 50 km is also very dependent on the crustal model. CRUST1.0 and CRUST2.0 yield $\xi > 1$ in those regions, but 3SMAC gives $\xi < 1$.

At 100 km, while we observe a dominant global $V_{SH} > V_{SV}$ signal for all three models, there are also clear regional differences. This can be seen beneath Tibet, North America, along the South American subduction region, Eurasia, and Iceland. The model resulting from crustal corrections calculated using 3SMAC displays a $V_{SV} > V_{SH}$ signal at these locations, but those obtained with CRUST1.0 or CRUST2.0 show $V_{SH} > V_{SV}$. We observe a similar discrepancy in the older parts of the Pacific ocean with $V_{SV} > V_{SH}$ in models obtained using 3SMAC, and a weak $V_{SH} > V_{SV}$ in the other models, which appears to support the findings of Ferreira *et al.* (2010). These discrepancies disappear at 150 km where all three models show a $V_{SH} > V_{SV}$ signal.

A difference between the models presented here and models from several other studies is that the contrast in ξ between oceans and continents between 100 and 150 km is much weaker here (Panning & Romanowicz 2006; Zhou *et al.* 2006; Kustowski *et al.* 2008; Nettles & Dziewoński 2008; Ferreira *et al.* 2010; Auer *et al.* 2014; Chang *et al.* 2014). In that regard, our mean models are closer to the models of Moulik & Ekström (2014) and French & Romanowicz (2014). Our mean models are also very similar to the results of Visser *et al.* (2008b) even though we used only a subset of their data set and made different modelling assumptions. At 250 km, our models display $V_{SV} > V_{SH}$ near subduction zone regions, possibly reflecting vertical flow. This has been seen in most other models before (Montagner & Tanimoto 1991; Panning & Romanowicz 2006; Zhou *et al.* 2006; Nettles & Dziewoński 2008; Visser *et al.* 2008b; Ferreira *et al.* 2010; Chang *et al.* 2014; French & Romanowicz 2014; Moulik & Ekström 2014). At 400 km the models all display $V_{SV} > V_{SH}$ in most regions, except in central Pacific where $\xi > 0$, as seen by Visser *et al.* (2008b) and Auer *et al.* (2014).

Fig. 10 shows the global correlation between the three pairs of models obtained using three different crust models. The correlation coefficients were calculated following Becker *et al.* (2007) after expanding the models in spherical harmonics up to degree 20. The velocity models obtained with CRUST1.0 and CRUST2.0 display the largest correlation at any given depth, except around 350 km. The velocity model obtained using 3SMAC displays an especially low correlation with the other two models between 50 and 100 km, and at 200 km. It is also low at 700 km, but this may not

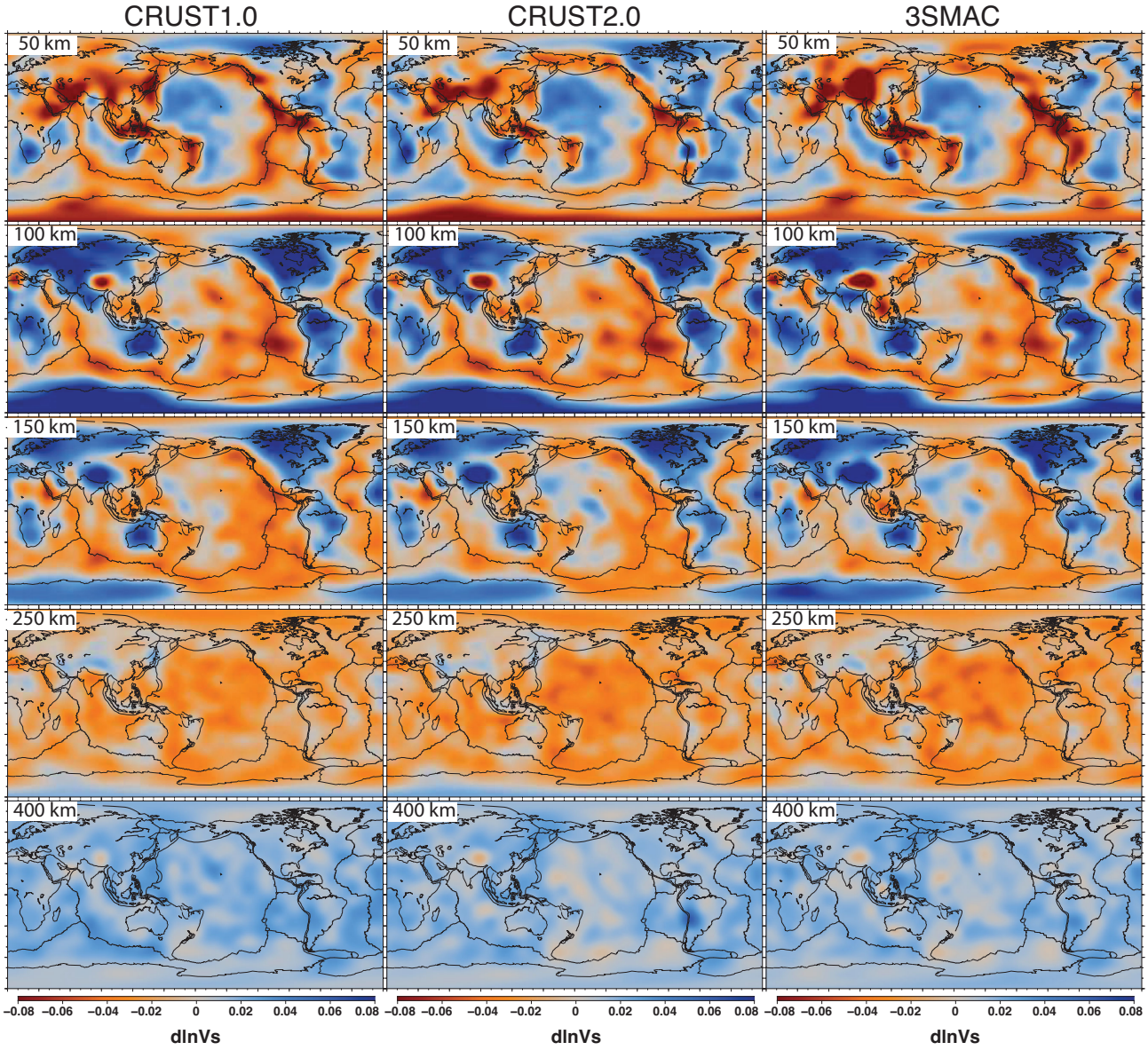

Figure 8. Shear velocity perturbations with respect to PREM at 50, 100, 150, 250 and 400 km obtained using, from left to right, CRUST1.0, CRUST2.0 and 3SMAC.

be significant since the data sensitivity is low at these depths. For radial anisotropy, the lowest correlation can be found at 150 km between the models obtained with CRUST1.0 and CRUST2.0 and with CRUST1.0 and 3SMAC. The correlation between the CRUST2.0- and 3SMAC-derived radial anisotropy models is good at that depth. Below 150 km, the radial anisotropy models correlate well with one another.

From the discussion of Figs 8 and 9, it appears that the choice of the crustal model does not strongly affect the velocity model, but strongly influences the radial anisotropy in the upper 100 km of the mantle. Ferreira *et al.* (2010) had found a similar effect at 100 km, but not at shallower depths. The purpose of this study was, however, to test the significance of these model discrepancies, which was done by comparing the differences between pairs of models to the size of their uncertainties. With m_1 and m_2 being the mean values of model 1 and model 2, and σ_1 and σ_2 their respective standard deviations, we calculated the difference between the two mean models $m_1 - m_2$ and compared it to the standard deviation of the distribution of $m_1 - m_2$, given by $\sqrt{\sigma_1^2 + \sigma_2^2}$ assuming independent Gaussian distributions. Figs 11 and 12 represent the difference $m_1 - m_2$ between two velocity and anisotropy models and the ratio between $m_1 - m_2$ and the standard deviation $\sqrt{\sigma_1^2 + \sigma_2^2}$. Values smaller than 1 correspond to discrepancies between models that are smaller than the uncertainty on the model difference, and are therefore not significant. Values larger than 1 correspond to discrepancies between models that are larger than the uncertainty on the model difference. These are significant and should be carefully analysed as they could yield different interpretations of the models.

Figs 11 and 12 demonstrate that differences in amplitude and short wavelength features exist between the models. We find differences in velocities even when comparing results from CRUST1.0 and CRUST2.0 (Fig. 11a), and they can reach 4 per cent with respect to PREM locally at 50 km depth and about 2 per cent at 150 km. Even stronger discrepancies can be seen between $d\ln V_S$ models resulting from using 3SMAC and CRUST1.0 (Fig. 12a). However,

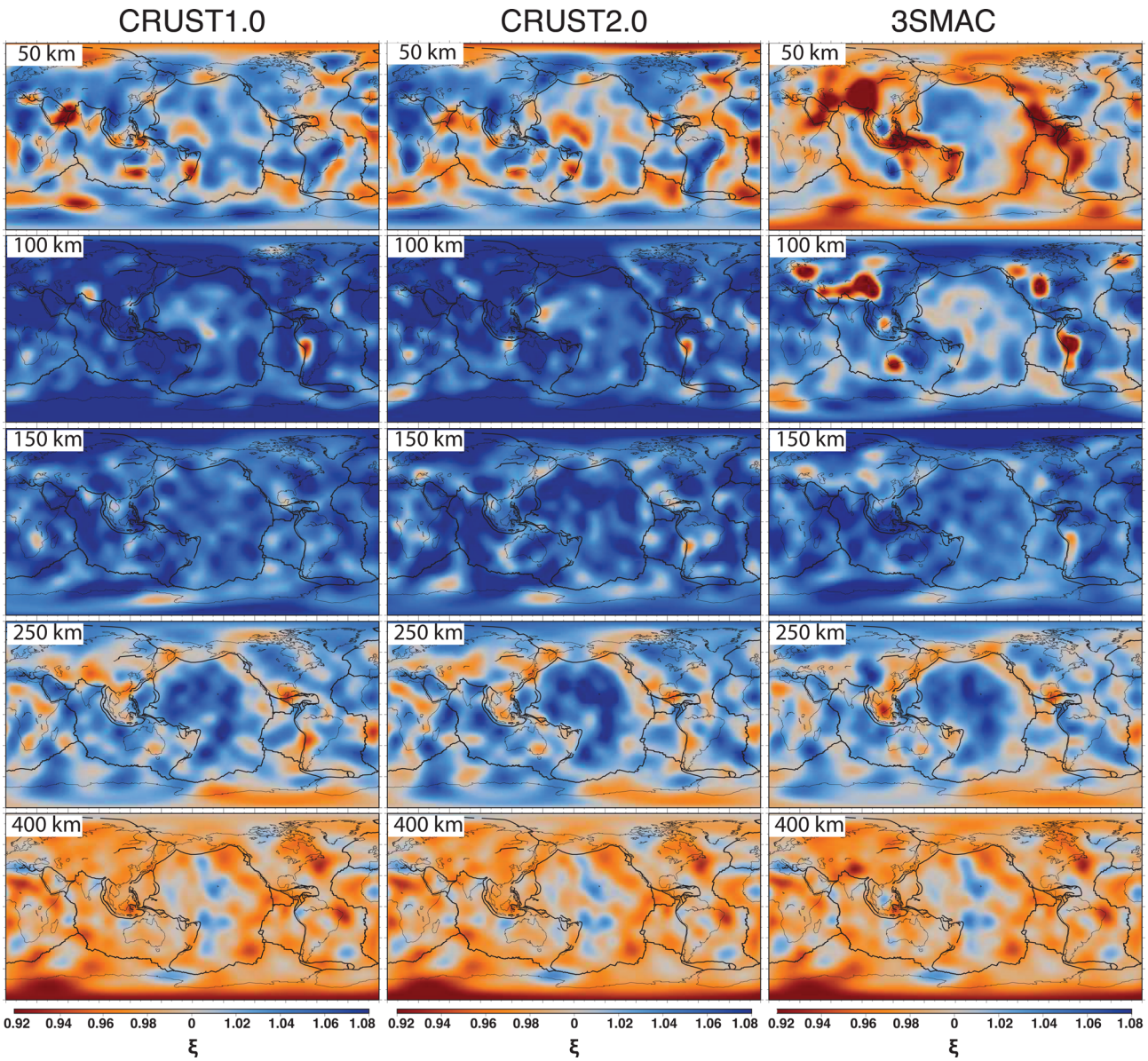

Figure 9. Radial anisotropy at 50, 100, 150, 250 and 400 km obtained using, from left to right, CRUST1.0, CRUST2.0 and 3SMAC. $\xi = 1$ corresponds to an isotropic mantle model. $\xi > 0$ corresponds to $V_{SH} > V_{SV}$.

if we compare the $d \ln V_S$ model difference with the uncertainty (Figs 11b and 12b), we find that generally the model discrepancies are not significant except in a few continental areas (North America, Northern Africa, and Eastern Asia) at 50 km when comparing CRUST1.0 and CRUST2.0 (Fig. 11b). Differences between the CRUST1.0- and 3SMAC-derived velocity models are not only stronger than when comparing CRUST1.0- and CRUST2.0-derived models, but they are also significant over much broader regions and to a greater depth range (down to 150 km depth beneath Tibet and part of the Andes). Discrepancies in mantle radial anisotropy were also found at most depths whether we compare models obtained using CRUST1.0 and CRUST2.0 or results from 3SMAC and CRUST1.0 (Figs 11c and 12c). The differences between the anisotropy models obtained using CRUST1.0 and CRUST2.0 are, however, generally smaller than the model uncertainties at most locations and are therefore not significant (Fig. 11d). On the contrary, when choosing between 3SMAC and CRUST1.0 to correct the dispersion data, one can affect the anisotropy model significantly in several continental areas in the top 100 km as shown by the large $d \ln \xi$ discrepancies with respect to the models standard deviations (Fig. 12d). In some locations such as near Tibet, the affected areas can extend as deep as 250 km.

Clearly, the choosing 3SMAC to correct our surface wave data instead of CRUST2.0 or CRUST1.0 has significant influence on the models, mostly in continental areas at depths less than 100 km. It is thus important to keep in mind that 3SMAC was built with strong geodynamical assumptions. In particular, it included surface heat flow data that were sparse, especially in continents. It also assumed that the temperature distribution under continents can be estimated using a plate model, and employed a crude regionalization for crustal ages. The distribution of points where crustal thickness was known was uneven, it included a much lower amount of active source seismic data than CRUST2.0 and CRUST1.0, and constraints on crustal *S*-wave velocities were sparse. In addition, 3SMAC did not include constraints from surface wave observations, and was shown to not be able to reproduce fundamental mode Love and Rayleigh wave phase

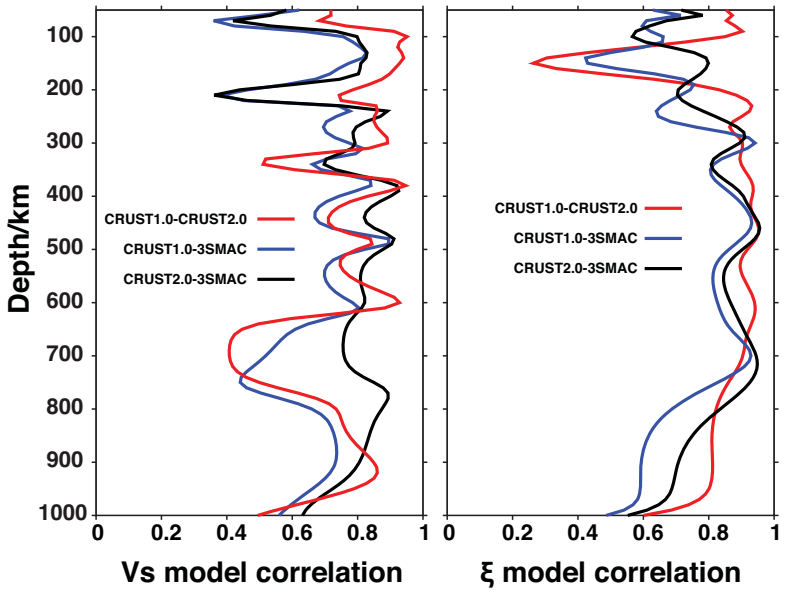

Figure 10. Correlation between the mean $d \ln V_S$ models obtained with different crustal models (left) and between the mean $d \ln \xi$ models (right).

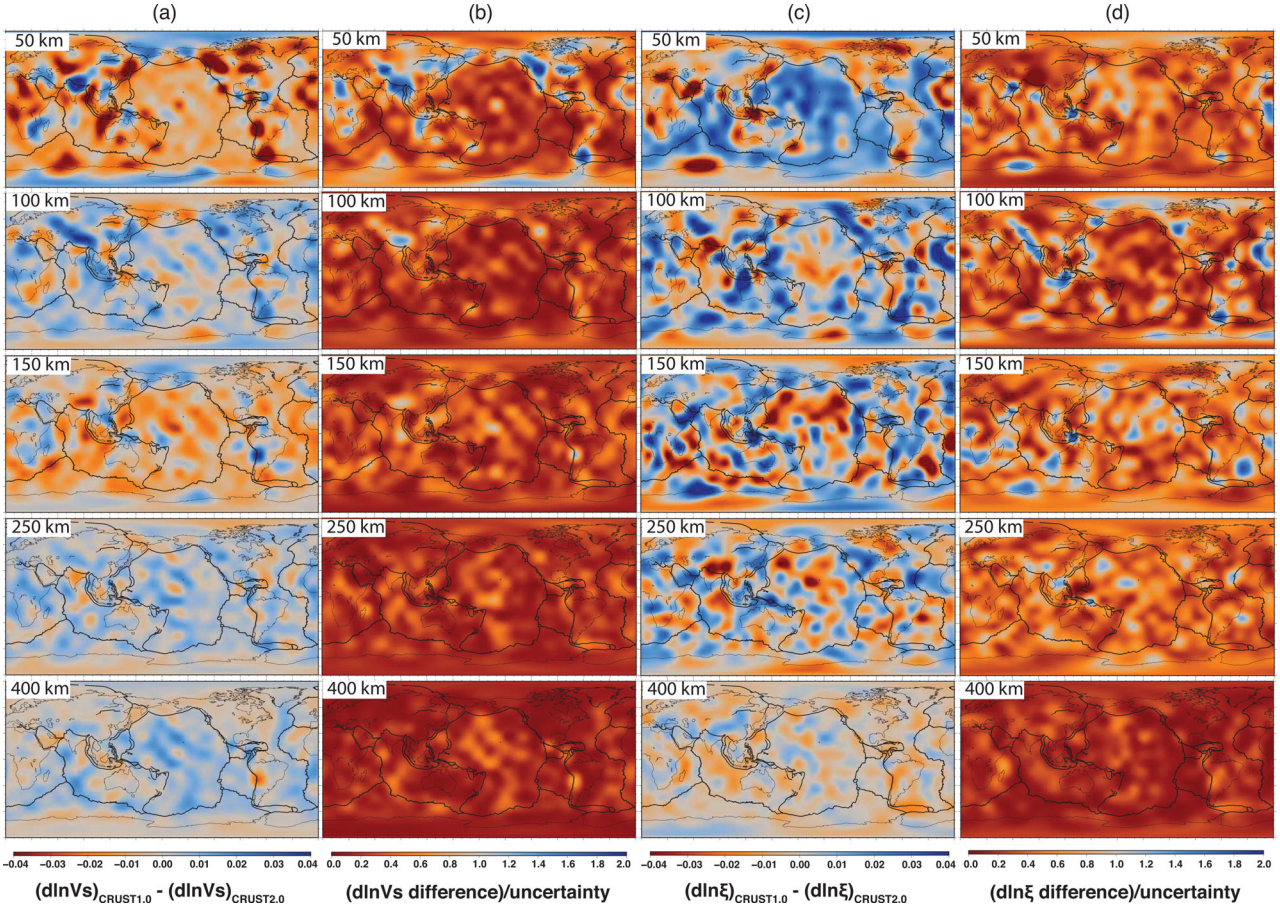

Figure 11. (a) Differences between the mean shear-wave velocity models obtained using CRUST1.0 and CRUST2.0; (b) model differences shown in panel (a) divided by the corresponding uncertainties; (c) differences between the mean radial anisotropy models obtained using CRUST1.0 and CRUST2.0; (d) model differences shown in panel (c) divided by the corresponding uncertainties.

velocities between 70 and 250 s under the Red Sea, Southeast China, and under Greenland (Ricard *et al.* 1996). In contrast, CRUST1.0 was not only built from a larger active source data set but also included constraints from group velocities, and therefore has better constraints on crustal velocities and Moho depth. Caution should therefore be taken in future work using prior crustal corrections based on 3SMAC in continental regions such as Eurasia, Southern Africa, North America and the South American subduction area.

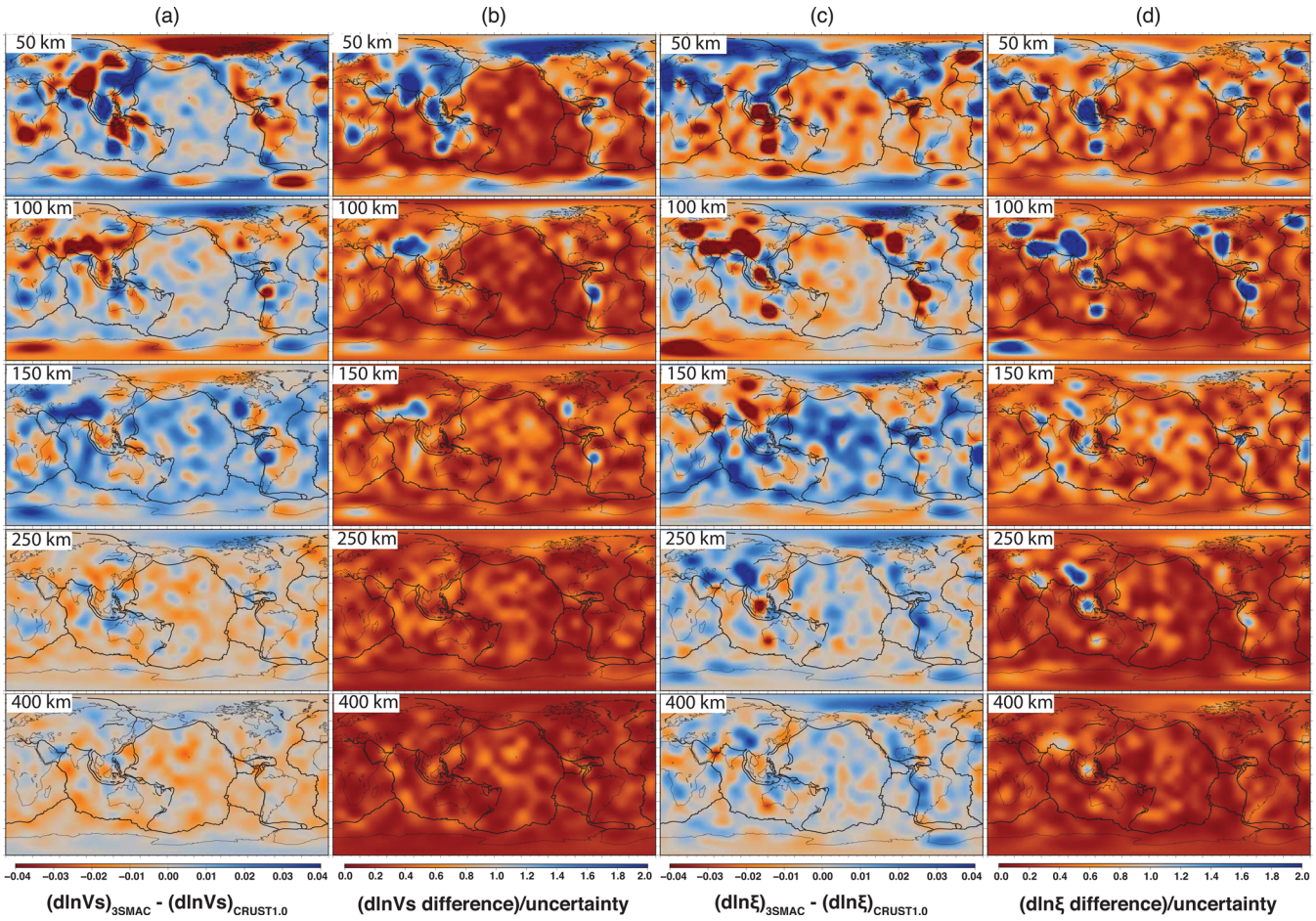

Figure 12. Same as in Fig. 11 but for 3SMAC and CRUST1.0.

The reader should note, of course, that in a Bayesian formulation such as the one employed here the level of data uncertainty directly determines posterior model uncertainties. Changing the data covariance matrix $\mathbf{C}_D$ would directly change the model uncertainties and possibly affect the conclusions of this study. It is therefore important to keep that caveat in mind since the phase velocity uncertainties determined by Visser *et al.* (2008a) by might not be fully characterized (Yuan & Beghein 2013), and because measurements of higher modes and estimates of their uncertainties will likely be refined in future studies. Our conclusions are, however, valid considering the current level of uncertainty in the published higher mode phase velocities employed here.

5 CONCLUSIONS

In this study we tested whether the choice of a prior crustal model to perform nonlinear corrections of phase velocity data affects mantle models of seismic velocities and radial anisotropy significantly. For this purpose, we compared the amplitude of the discrepancies between tomographic models obtained using three crustal models, namely CRUST1.0, CRUST2.0 and 3SMAC, with the size of the model uncertainties that result from the inherent non-uniqueness of the problem. While other studies have addressed this topic before (e.g. Ferreira *et al.* 2010), they did not estimate quantitative model uncertainties. Determining which model features are resolved is, however, an important part of modelling mantle structure and should be done whenever possible before discussing the effect of a chosen prior information, such as the crustal model, and before making any geodynamic interpretation of the mantle model.

Here, we determined reliable quantitative model uncertainties using a model space search approach. It enabled us to obtain a PPDF to describe the model parameters instead of choosing one particular model among many possible solutions as one does with a regularized inverse technique. This likelihood function could then be used to obtain uncertainties on the model parameters, that is, shear wave velocity and radial anisotropy, which were then used to assess the significance of the changes in mantle radial anisotropy observed when testing different crustal models.

First, we found that the choice of the crustal model can significantly influence the mantle models in continental regions at shallow depths, which has important consequences for determining the depth of continental roots and understanding continent formation. Using 3SMAC instead of CRUST1.0 has a significant influence on the radial anisotropy model in the top 100 km. That effect can even extend down to 250 km in some locations. It should be noted, however, that CRUST1.0 and CRUST2.0 yield statistically identical radial anisotropy models at all depths, except at a few grid cells. This means that mantle models obtained using CRUST1.0 can fit the data corrected using CRUST2.0 almost equally well, within 1σ uncertainty.

Second, we also observed that the choice of the crustal model does not only affect mantle radial anisotropy but also seismic velocities especially in the top 50 km, and it can extend to 150 km depth locally

when using 3SMAC. Interestingly, we found that radial anisotropy is not significantly affected by the crustal model in oceanic regions. This contradicts the conclusions drawn by Ferreira *et al.* (2010) who found that the crustal models mostly influenced radial anisotropy in the Pacific at 100 km depth. We note, however, that we did find discrepancies between models in the Pacific as well, but our analysis of model non-uniqueness demonstrated that these differences are smaller than the model uncertainties.

Finally, we found that the prior crustal model does not significantly affect the radial anisotropy and velocity models at depths greater than 100 km. The mantle models are thus statistically indistinguishable below that depth. This implies that if geodynamic models of radial anisotropy below 100 km depth were to account for tomographic model uncertainties, they would not depend on the choice of the prior crustal model. Our results also suggest that 3SMAC as it stands today should not be used for crustal corrections in seismotectonic studies of Eurasia, Southern Africa, North and America, and that the geoscience community would benefit from putting some efforts towards building a revised 3SMAC. The discrepancies between mantle models built based on 3SMAC crustal corrections and those based on CRUST1.0 or CRUST2.0 should also help shed light on the validity of the geodynamical assumptions made in the construction of models like 3SMAC.

ACKNOWLEDGEMENTS

The authors thank Malcolm Sambridge for making his Neighbourhood Algorithm available and Jeannot Trampert and Karin Visser for sharing their higher mode data online at http://www.geo.uu.nl/~jeannot/My_web_pages/Downloads.html. Reviews by Barbara Romanowicz and Thomas Bodin helped improve this manuscript.

REFERENCES

Anderson, D.L., 1961. Elastic wave propagation in layered anisotropic media, *J. geophys. Res.,* **66**(9), 2953–2963.

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.

Babuska, V. & Cara, M., 1991. *Seismic Anisotropy in the Earth*, vol. 10 of Modern Approaches in Geophysics, Springer.

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

Bayes, M. & Price, M., 1763. An essay towards solving a problem in the doctrine of chances. By the Late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFRS, *Phil. Trans. R. Soc. Lond., A.,* **53,** 370–418.

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

Beghein, C., 2010. Radial anisotropy and prior petrological constraints: a comparative study, *J. geophys. Res.,* **115**(B3), doi:10.1029/2008JB005842.

Beghein, C. & Trampert, J., 2002. P and S tomography using normal-mode and surface waves data with a neighbourhood algorithm, *Geophys. J. Int.,* **149**(3), 646–658.

Beghein, C. & Trampert, J., 2003. Robust normal mode constraints on inner-core anisotropy from model space search, *Science,* **299**(5606), 552–555.

Beghein, C. & Trampert, J., 2004. Probability density functions for radial anisotropy: implications for the upper 1200 km of the mantle, *Earth planet. Sci. Lett.,* **217**(1), 151–162.

Beghein, C., Resovsky, J. & van der Hilst, R., 2008. The signal of mantle anisotropy in the coupling of normal modes, *Geophys. J. int.,* **175,** 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**(6176), 1237–1240.

Bodin, T., Capdeville, Y., Romanowicz, B. & Montagner, J.-P., 2015. Interpreting radial anisotropy in global and regional tomographic models, in *The Earth's Heterogeneous Mantle*, pp. 105–144, eds Khan, A. & Deschamps, F., Springer International Publishing.

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:10.1785/0120070041.

Burgos, G., Montagner, J.-P., Beucler, E., Capdeville, Y., Mocquet, A. & Drilleau, M., 2014. Oceanic lithosphere-asthenosphere boundary from surface wave dispersion data, *J. geophys. Res.,* **119**(2), 1079–1093.

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

Debayle, E., Lévêque, J.J. & Cara, M., 2001. Seismic evidence for a deeply rooted low-velocity anomaly in the upper mantle beneath the northeastern Afro/Arabian continent, *Earth Planet. Sci. Lett.,* **193**(3–4), 423–436.

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

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

Ferreira, a. M.G., Woodhouse, J.H., Visser, K. & Trampert, J., 2010. On the robustness of global radially anisotropic surface wave tomography, *J. geophys. Res.,* **115**(B4), 1–16.

Forsyth, D.W., 1975. The early structural evolution and anisotropy of the oceanic upper mantle, *Geophys. J. Int.,* **43**(1), 103–162.

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.

French, S.W., Lekić, V. & Romanowicz, B., 2013. Waveform tomography reveals channeled flow at the base of the oceanic asthenosphere, *Science,* **342**(6155), 227–230.

Gung, Y., Panning, M. & Romanowicz, B., 2003. Global anisotropy and the thickness of continents, *Nature,* **422**(6933), 707–711.

Hess, H.H., 1964. Seismic anisotropy of the uppermost mantle under oceans, *Nature,* **203,** 629–631.

Jackson, D.D., 1979. The use of a priori data to resolve non-uniqueness in linear inversion, *Geophys. J. Int.,* **57**(1), 137–157.

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

Karato, S.-I., 1998. Some remarks on the origin of seismic anisotropy in the D″ layer, *Earth, Planets Space,* **50**(11–12), 1019–1028.

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

Kustowski, B., Dziewonski, A.M. & Ekström, G., 2007. Nonlinear Crustal Corrections for Normal-Mode Seismograms, *Bull. seism. Soc. Am.,* **97**(5), 1756–1762.

Kustowski, B., Ekström, G. & Dziewonski, A.M., 2008. Anisotropic shear-wave velocity structure of the Earth's mantle: a global model, *J. geophys. Res.,* **113**(B6), 1–23.

Laske, G., Masters, G., Ma, Z. & Pasyanos, M., 2013. Update on CRUST1.0—a 1-degree global model of Earth's crust, in *EGU General Assembly Conference Abstracts*, vol. 15, 2658.

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

Lekić, V., Panning, M. & Romanowicz, B., 2010. A simple method for improving crustal corrections in waveform tomography, *Geophys. J. Int.,* **182**(1), 265–278.

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

Love, A.E.H., 1927. *A Treatise on the Mathematical Theory of Elasticity*, Dover.

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

Matsu'ura, M. & Hirata, N., 1982. Generalized least-squares solutions to quasi-linear inverse problems with a priori information, *J. Phys. Earth,* **30**(6), 451–468.

Meier, U., Curtis, A. & Trampert, J., 2007. Fully nonlinear inversion of fundamental mode surface waves for a global crustal model, *Geophys. Res. Lett.,* **34**(16), doi:10.1029/2007GL030989.

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

Montagner, J.-P., 2002. Upper mantle low anisotropy channels below the Pacific Plate, *Earth planet. Sci. Lett.,* **202**(2), 263–274.

Montagner, J.-P. & Anderson, D.L., 1989. Petrological constraints on seismic anisotropy, *Phys. Earth planet. Inter.,* **54**(1-2), 82–105.

Montagner, J.-P. & Nataf, H.-C., 1986. A simple method for inverting the azimuthal anisotropy of surface waves, *J. geophys. Res.,* **91**(B1), 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.

Mooney, W., Laske, G. & Masters, G., 1998. CRUST 5.1: A global crustal model at 5° × 5°, *J. geophys. Res.,* **103,** 727–747.

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.

Nataf, H.-C. & Ricard, Y., 1996. 3SMAC: an a priori tomographic model of the upper mantle based on geophysical modeling, *Phys. Earth planet. Inter.,* **95**(1), 101–122.

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

Nataf, H.-C., Nakanishi, I. & Anderson, D.L., 1986. Measurements of mantle wave velocities and inversion for lateral heterogeneities and anisotropy: 3. Inversion, *J. geophys. Res.,* **91**(B7), 7261–7307.

Nettles, M. & Dziewoński, A.M., 2008. Radially anisotropic shear velocity structure of the upper mantle globally and beneath north America, *J. geophys. Res.,* **113**(B2), doi:10.1029/2006JB004819.

Nettles, M. & Dziewoński, A.M., 2011. Effect of higher-mode interference on measurements and models of fundamental-mode surface-wave dispersion, *Bull. seism. Soc. Am.,* **101**(5), 2270–2280.

Nicolas, A. & Christensen, N.I., 1987. Formation of anisotropy in upper mantle peridotites–a review, in *Composition, Structure and Dynamics of the Lithosphere-Asthenosphere System*, pp. 111–123, eds Fuchs, K. & Froidevaux, C., American Geophysical Union.

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.

Ricard, Y., Nataf, H.-C. & Montagner, J.-P., 1996. The three-dimensional seismological model a priori constrained: confrontation with seismic data, *J. geophys. Res.,* **101**(B4), 8457–8472.

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.

Silveira, G., Stutzmann, E., Griot, D.A. & Montagner, J.-P., 1998. Anisotropic tomography of the Atlantic Ocean from Rayleigh surface waves, *Phys. Earth planet. Inter.,* **106**(3–4), 257–273.

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

Takeuchi, H. & Saito, M., 1972. Seismic surface waves, in *Methods in Computational Physics*, vol. 11, pp. 217–295, ed. Bolt, B., Academic Press.

Trampert, J., 1998. Global seismic tomography: the inverse problem and beyond, *Inverse Probl.,* **14**(3), 371–385.

Trampert, J. & Snieder, R., 1996. Model estimations biased by truncated expansions: possible artifacts in seismic tomography, *Science,* **271**(5253), 1257–1260.

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., 2008. Monte Carlo search techniques applied to the measurement of higher mode phase velocities and anisotropic surface wave tomography, *PhD thesis*, Utrecht University.

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

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

Woodhouse, J.H. & Dahlen, F.A., 1978. The effect of a general aspherical perturbation on the free oscillations of the Earth, *Geophys. J. Int.,* **53**(2), 335–354.

Woodhouse, J.H. & Dziewonski, A.M., 1984. Mapping the upper mantle: three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. geophys. Res.,* **89**(B7), 5953–5986.

Yoshizawa, K. & Kennett, B., 2002. Non-linear waveform inversion for surface waves with a neighbourhood algorithm, application to multimode dispersion measurements, *Geophys. J. Int.,* **149**(1), 118–133.

Yoshizawa, K. & Kennett, B., 2004. Multimode surface wave tomography for the Australian region using a three-stage approach incorporating finite frequency effects, *J. geophys. Res.,* **109**(B02310), 1978–2012.

Yuan, H. & Romanowicz, B., 2010. Lithospheric layering in the North American craton, *Nature,* **466**(7310), 1063–1069.

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.

Zhou, Y., Nolet, G., Dahlen, F. & Laske, G., 2006. Global upper-mantle structure from finite-frequency surface-wave tomography, *J. geophys. Res.,* **111**(B4), doi:10.1029/2005JB003677.