

Imaging upper mantle anisotropy with teleseismic *P*-wave delays: insights from tomographic reconstructions of subduction simulations

Brandon P. VanderBeek¹ and Manuele Faccenda

Dipartimento di Geoscienze, Università degli Studi di Padova, 35131 Padova PD, Italy. E-mail: brandon.p.vanderbeek@gmail.com

Accepted 2021 February 25. Received 2021 February 24; in original form 2020 July 31

SUMMARY

Despite the well-established anisotropic nature of Earth's upper mantle, the influence of elastic anisotropy on teleseismic *P*-wave imaging remains largely ignored. Unmodelled anisotropic heterogeneity can lead to substantial isotropic velocity artefacts that may be misinterpreted as compositional heterogeneities. Recent studies have demonstrated the possibility of inverting *P*-wave delay times for the strength and orientation of seismic anisotropy. However, the ability of *P*-wave delay times to constrain complex anisotropic patterns, such as those expected in subduction settings, remains unclear as synthetic testing has been restricted to the recovery of simplified block-like structures using ideal self-consistent data (i.e. data produced using the assumptions built into the tomography algorithm). Here, we present a modified parametrization for imaging arbitrarily oriented hexagonal anisotropy and test the method by reconstructing geodynamic simulations of subduction. Our inversion approach allows for isotropic starting models and includes approximate analytic finite-frequency sensitivity kernels for the simplified anisotropic parameters. Synthetic seismic data are created by propagating teleseismic waves through an elastically anisotropic subduction zone model created via petrologic-thermomechanical modelling. Delay times across a synthetic seismic array are measured using conventional cross-correlation techniques. We find that our imaging algorithm is capable of resolving large-scale features in subduction zone anisotropic structure (e.g. toroidal flow pattern and dipping fabrics associated with the descending slab). Allowing for arbitrarily oriented anisotropy also results in a more accurate reconstruction of isotropic slab structure. In comparison, models created assuming isotropy or only azimuthal anisotropy contain significant isotropic and anisotropic imaging artefacts that may lead to spurious interpretations. We conclude that teleseismic *P*-wave traveltimes are a useful observable for probing the 3-D distribution of upper mantle anisotropy and that anisotropic inversions should be explored to better understand the nature of isotropic velocity anomalies particularly in subduction settings.

Key words: Composition and structure of the mantle; Body waves; Seismic anisotropy; Seismic tomography.

1 INTRODUCTION

Seismic imaging is a primary means for constraining the physical state of Earth's mantle which is otherwise isolated from direct observation. Most commonly, tomographic studies yield maps of wave speeds which can be used to infer physical properties (e.g. density, temperature, melt fraction and volatile content). Observations from dense seismic deployments (e.g. Goes & van der Lee 2002; Wiens *et al.* 2008; Villagómez *et al.* 2014; Plank & Forsyth 2016; Byrnes *et al.* 2017; Eilon & Abers 2017; Lebedev *et al.* 2018;

Ruan *et al.* 2018) combined with experimental and theoretical work on the relationship between seismic properties and petrology (e.g. Karato & Jung 1998; Hammond & Humphreys 2000; Takei 2002; Cammarano *et al.* 2003; Jackson & Faul 2010) has allowed researchers to make more robust and detailed estimates of mantle thermal structure and composition. Micromechanical models and petrologic studies linking mantle flow to mineral and compositional fabric development (e.g. Ribe 1989; Blackman & Kendall 2002; Kaminski *et al.* 2004; Karato *et al.* 2008; Faccenda 2014; Skemer & Hansen 2016) yields detailed predictions regarding the

anisotropic structure of the Earth. These fabrics can create significant directional variations in seismic velocities allowing mantle dynamics to be investigated via the anisotropic properties of the seismic wavefield. While the anisotropic nature of the mantle has been well-established through decades of theoretical and observational research (e.g. Savage 1999; Karato *et al.* 2008; Long & Becker 2010), it remains commonplace to assume an isotropic Earth in the construction of tomographic models, particularly those derived from body waves. This is a poor assumption considering that *P* waves exhibit strong sensitivity to anisotropic fabrics and neglecting directional velocity variations can lead to substantial imaging artefacts and misleading interpretations (Kendall 1994; Blackman *et al.* 1996; Blackman & Kendall 1997; Sobolev *et al.* 1999; Lloyd & van der Lee 2008; Menke 2015; Bezada *et al.* 2016). Furthermore, tomographic models provide a metric for evaluating geodynamic simulations and give constraints on density anomalies that drive flow (e.g. Becker & Boschi 2002; Simmons *et al.* 2006; Faccenda & Becker 2010; Wang & Becker 2019). Therefore, creating accurate seismic images is fundamental to advancing our understanding of the physical state and dynamics of the mantle. This in turn requires accounting for the seismically anisotropic structure of the Earth.

The primary means for investigating upper mantle anisotropy is through the analysis of surface waves and shear wave splitting. Surface waves are particularly sensitive to depth variations in both azimuthal and radial anisotropic structure. However, the relatively long periods used for mantle imaging result in poor lateral resolution (hundreds of kilometres). In comparison, the higher frequency content of body waves used in shear wave splitting analyses allow for the investigation of anisotropic heterogeneity at finer lateral scales limited primarily by station spacing (50–100 km). However, the near vertical nature of the *S* phases used in these studies (most commonly SKS) results in poor depth resolution and limited sensitivity to the dip of anisotropic fabrics (Chevrot & van der Hilst 2003; Chevrot 2006; Beller & Chevrot 2020). The vast number of publications on surface wave tomography and shear wave splitting studies unambiguously demonstrate the seismic anisotropic nature of the upper mantle at both global and local scales (e.g. Vinnik *et al.* 1992; Becker *et al.* 2003; Conrad *et al.* 2007; Long & Becker 2010). Such anisotropic structure will impact the arrival of body waves, particularly *P* phases (Bokelmann 2002; Schulte-Pelkum & Blackman 2003; Sieminski *et al.* 2007). Despite these findings, seismic anisotropy remains largely ignored in the tomographic inversion of teleseismic body wave data. This is in part due to a general reluctance to include additional inversion parameters in an already underdetermined problem. However, it is also difficult to justify the wilful neglect of parameters known to have a strong influence on seismic observables.

Here we present and rigorously evaluate a new approach for the inversion of teleseismic *P*-wave delays for upper mantle isotropic velocity and arbitrarily oriented hexagonal anisotropy. Our method includes approximations for finite-frequency effects on anisotropic traveltimes. Realistic synthetic seismic data sets are created by performing waveform modelling through geodynamic simulations of subduction. We systematically evaluate the ability of isotropic, azimuthally anisotropic and arbitrarily oriented anisotropic inversions to accurately capture subduction zone structure. We find that isotropic and azimuthally anisotropic inversions can generate strong imaging artefacts that should be considered when interpreting seismic images. Anisotropic inversions that account for dipping fabrics yield the most accurate reconstruction of the synthetic subduction zone. Our results highlight the ability of teleseismic *P* waves to

constrain regional upper mantle anisotropy. Application of our method to existing seismic data sets holds promise for yielding new insights into upper mantle dynamics.

2 BACKGROUND

A variety of methods have been developed for extracting body wave anisotropic structure primarily from the analysis of *P* phases. Seismic birefringence complicates the analysis of shear phases. While methods that account for anisotropy when measuring *S*-wave arrivals have been proposed (e.g. Sieminski *et al.* 2007), we focus on *P* phases. *P*-wave residuals exhibit azimuthal variations that can be linked to underlying mantle fabrics (Bokelmann 2002). Babuška *et al.* (1993) identify homogeneous anisotropic domains by searching for the orientation of an assumed elastic tensor that can best explain directional variations in *P*-wave station delays. This approach has subsequently been applied across Europe (e.g. Babuška & Plomerová 2006; Plomerová *et al.* 2012) and combined with shear wave splitting observations (e.g. Šílený & Plomerová 1996; Eken *et al.* 2012) to map lateral variations in anisotropic fabrics. A similar strategy, combined with isotropic inversions, was used by Hammond & Toomey (2003) to identify anisotropic domains beneath the East Pacific Rise.

To better constrain the 3-D distribution of seismic velocity heterogeneity, anisotropic traveltimes inversion algorithms subject to various simplifications have been developed. Grésillaud & Cara (1996) present a linear inversion method for *P*-wave speed and strength of anisotropy assuming a simplified olivine elastic tensor with a prescribed orientation. A number of tomography algorithms have been developed for imaging azimuthal anisotropy using *P*-wave arrivals from active-sources (e.g. Dunn *et al.* 2005; Weekly *et al.* 2014) or local earthquakes (e.g. Hearn 1996; Eberhart-Phillips & Henderson 2004; Wang & Zhao 2008; Koulakov *et al.* 2009). These methods rely on approximations to the directional dependence of *P*-wave speeds in weakly hexagonally anisotropic media initially presented by Backus (1965). It is often sufficient to describe velocity in such media by sinusoidal functions with period π allowing anisotropy to be parametrized through the introduction of two additional terms that control the anisotropic strength and azimuth. Assuming hexagonal anisotropy with a vertically oriented symmetry axis, Wang & Zhao (2013) present a method for mapping radial anisotropic structure from body waves. Building on this work, Liu & Zhao (2017) simultaneously determine 3-D azimuthal and radial anisotropic structure by considering vertically oriented orthorhombic symmetry. These azimuthal and radial anisotropic imaging methodologies have subsequently been applied to modelling teleseismic data (see review by Zhao *et al.* 2016).

The aforementioned studies all assume that the anisotropic symmetry axis is horizontally or vertically oriented. However, this assumption is likely not valid for upper mantle fabrics. Backazimuthal variations in *P*-wave delays (e.g. Babuška *et al.* 1993; Babuška & Plomerová 2006; Plomerová *et al.* 2012) are indicative of inclined fabrics. Geodynamic models of plumes (e.g. Ito *et al.* 2014), mid-ocean ridges (e.g. Blackman & Kendall 2002) and subduction zones (e.g. Faccenda 2014) all predict diverse distributions of mineral fabrics that are not confined to the horizontal plane. To account for arbitrarily oriented fabrics, Munzarová *et al.* (2018a) devised a *P*-wave traveltimes inversion algorithm that solves for three spherical parameters characterizing a hexagonally anisotropic medium. While the anisotropic imaging method of Munzarová *et al.* (2018a) proved

successful in their synthetic tests, the strongly non-linear problem requires an anisotropic starting model that is sufficiently close to the true structure. Consequently, numerous anisotropic starting models must be explored to identify plausible solutions. Thus, it may be difficult to apply this approach in environments with complex anisotropic fabrics and limited prior knowledge.

Recently, Beller & Chevrot (2020) developed a full-waveform inversion (FWI) algorithm capable of solving for the 21 elastic coefficients leading to a more linear imaging problem. This approach does not require simplifying assumptions regarding the symmetry system or orientation and includes both *P* and *S* phases. Importantly, Beller & Chevrot (2020) found that the inclusion of *P* phases are paramount to the recovery of anisotropic fabric orientations. While anisotropic FWI will be required to generate the highest resolution images of upper mantle elastic structure, the applicability of FWI remains limited by the availability of computational resources and sufficiently dense high-quality seismic recordings.

In this study, we evaluate of the ability of teleseismic *P*-wave delay times to constrain realistic upper mantle anisotropic structure. The main contributions of this work are as follows. We present a new parametrization for imaging arbitrarily oriented hexagonal anisotropy that allows for isotropic starting models. An approximation to model finite-frequency traveltimes in simplified anisotropic media is presented and shown to yield more accurate traveltimes over ray theory. Our synthetic tests focus on the reconstruction of a synthetic subduction zone model created via petrologic-thermomechanical geodynamic simulations. While some of the aforementioned tomography methods have been applied to real data sets (e.g. Zhao *et al.* 2016; Munzarová *et al.* 2018b), their ability to resolve realistic anisotropic fabrics has not been extensively evaluated. Analysis of prior anisotropic imaging methods relies on the reconstruction of simple block-like structures. The synthetic data for such tests is generated via the same algorithms used for imaging leading to an idealized test. Therefore, it is difficult to glean the nature of biases and artefacts inherent in each method that may arise when applied to real data. Our synthetic seismic data sets are generated by simulating seismic wave propagation through an elastic geodynamic model resulting in delay times that are independent of the assumptions built into our tomography code. The resulting seismic images provide an accurate view of the structures and artefacts we can expect to observe in tomographic results from real-world subduction environments.

3 METHODS

3.1 Synthetic subduction zone model

We construct an anisotropic synthetic subduction zone using petrologic-thermomechanical modelling described in detail by Faccenda (2014). The subduction zone consists of a slab with a half-width of 1000 km that subducts freely in response to its negative buoyancy and stagnates over the 660 km discontinuity. Anisotropic fabric is simulated by micromechanical modelling of polymineralic aggregates that are passively advected through the mantle flow field. The rheological model accounts for brittle deformation according to a Drucker–Prager criterion, as well as for dislocation, diffusion and low-*T* Peierls creep mechanisms. Upper mantle aggregates have a harzburgitic composition (Ol:Ens = 70:30, <410 km) while a pyroclitic composition is used at greater depths (Wd:Grt = 60:40 in the upper transition zone, 410–520 km; Rw:Grt = 60:40 in the lower transition zone, 520–660 km; Bridgmanite:Mg = 80:20 for the lower

mantle, >660 km; (Mainprice 2007; Faccenda 2014). Our model does not contain a crustal layer or topography. At each time-step, the anisotropic fabric for each polycrystalline aggregate is calculated following Kaminski *et al.* (2004) with modifications to account for non-steady-state deformation and deformation history, which is particularly relevant when considering subduction systems (Faccenda & Capitanio 2013), as well as strain-induced LPO of mid-mantle aggregates. For each Lagrangian aggregate, the micromechanical modelling yields a full elastic tensor scaled by the local *P*–*T* conditions predicted by the thermomechanical model of Faccenda (2014). Elastic moduli at any point in the model are obtained via interpolation. In addition to flow-induced fabrics, pre-existing anisotropy is added to the slab to represent ‘frozen’ anisotropy created at the spreading centre and subsequently incorporated into the lithosphere. Finally, we note that the micro- and macroflow simulations are run in a Cartesian coordinate system. The resulting solution is distorted to replicate the Earth’s curvature.

While the resulting elastic tensors are fully anisotropic, the dominant symmetry is hexagonal. We have chosen to simplify the elastic model by approximating the tensors with a hexagonally symmetric tensor (Browaeys & Chevrot 2004) and disregard lower-order symmetry components. The isotropic velocity structure relative to a far-field 1-D reference velocity profile that resembles IASP91 Kennett & Engdahl (1991) without a crustal layer is shown in Fig. 1. Note that the geodynamic model extends to 1000 km depth but the tomographic model domain ends at 700 km (below which there is no significant structure). The only significant isotropic anomaly is the relatively cold subducting slab which is characterized by a 3–5 per cent increase in *P*-wave speed. The thin horizontal anomalies are due to differences in the fine-scale structure of seismic discontinuities between the 1-D reference profile and the 3-D model.

Selected cross-sections illustrating the anisotropic structure are shown in Fig. 2. The anisotropic structure can be characterized by four main domains. (1) There is a frozen anisotropic fabric within the slab defined by 4–6 per cent *P*-wave speed variations. (2) A similarly aligned fabric is produced by entrained mantle flow that follows the movement of the descending plate. (3) Surrounding the edges of the slab there is a clear toroidal flow pattern in the anisotropic fabric that is present throughout the upper ~300 km. (4) Finally, there is a region of trench-parallel anisotropy beneath the incoming plate. In addition to these four zones, a corner-flow type pattern can be observed in the wedge near the mid-plate. However, this geometry is less evident towards the slab edges where symmetry axes become more steeply dipping. Anisotropy is nearly extinguished beneath 400 km where the dominant minerals create weaker fabrics. All of these features are potential imaging targets that, if recovered, provide insights into the dynamics of the subduction system.

3.2 Synthetic seismic data

We perform anisotropic teleseismic wavefield simulations using the spectral element code SPECFEM 3-D (Komatitsch & Tromp 1999; Chen & Tromp 2007) with the AxiSEM grid-injection technique (Monteiller & Long 2013; Nissen-Meyer *et al.* 2014). Our 3-D anisotropic elastic model is imbedded in an otherwise 1-D radial earth model defined by IASP91 velocities (Kennett & Engdahl 1991). Each of the 16 events in Fig. 1 is modelled with a double couple source mechanism with a Gaussian moment function. The dominant period of the source is 15 s while the computation mesh is accurate to periods longer than ~10 s. The time step used in SPECFEM simulations is 100 ms. This relatively long period was

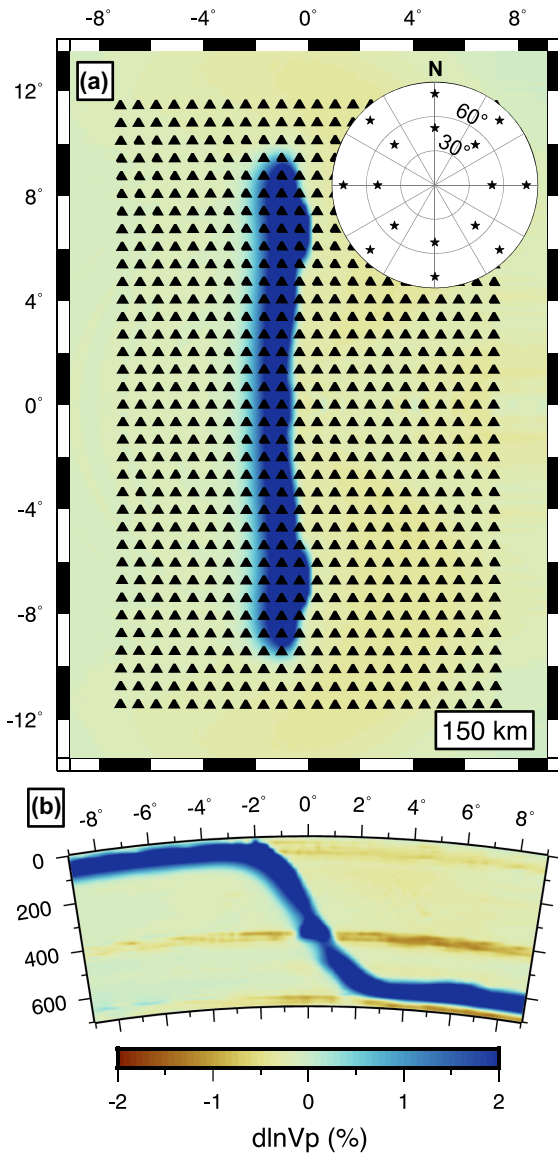


Figure 1. Isotropic structure, array geometry and distribution of teleseismic sources considered in the present study. (a) Seismic stations (black triangles) are uniformly spaced 75 km apart and plotted over isotropic velocity heterogeneity in the true model at 150 km depth. Inset shows location of teleseismic sources (stars) relative to the experiment centre. Sources are located at distances of 50° and 80° and evenly distributed in backazimuth. An east-west cross-section through the true isotropic model at 4°30'S is shown in (b). Note that the isotropic structure is symmetric about 0°N.

chosen for computational efficiency but we note that it is comparable to the observational period used for teleseismic body waves in noisier offshore environments (e.g. Byrnes *et al.* 2017; Bodmer *et al.* 2018). The modelled wavefields are recorded on 770 receivers uniformly spaced 75 km apart (Fig. 1), a density comparable to the USArray deployment. Both direct P and S phases are simulated as well as SKS waveforms from 8 additional events evenly distributed in backazimuth at a range of 120° so that we could assess splitting intensity through the subduction model. A second-order bandpass filter with corners at 15 and 40 s is applied to the seismograms and relative delays are measured via waveform cross-correlation (VanDecar & Crosson 1990) using an iterative method (Lou *et al.* 2013) to semi-automate the procedure. Estimated measurement errors are

~56 ms. However, based on the accuracy of our approximate finite-frequency traveltimes calculations discussed in Section 3.3.2, all data are assigned an error of 150 ms. For comparison, most modern regional teleseismic tomography models fit the observed data to 300–500 ms. The final set of delay times is completely independent of the simplifications made in our ray-based tomography algorithm.

3.3 Forward problem

We use the TauP Toolkit (Crotwell *et al.* 1999) to predict ray paths and traveltimes through a 1-D radial earth model. The use of 1-D rays is appropriate given the relatively long period data and the magnitude and dimensions of heterogeneity considered in this study. The effect of isotropic and anisotropic heterogeneity within the imaging volume on the 1-D ray theoretical traveltimes is computed using approximate finite-frequency slowness kernels. We present the details of this approximation below.

3.3.1 Anisotropy parametrization

It is well established that anisotropy in Earth's upper mantle can be approximated by a hexagonally symmetric media (Becker *et al.* 2006). For hexagonal anisotropy, the directional dependence of P -wave speeds is well approximated by a periodic function of 2α and 4α terms where α is the angle between the hexagonal symmetry axis and the P -wave propagation direction (Backus 1965; Thomsen 1986). However, the 4α terms for mantle fabrics are generally an order of magnitude smaller than the 2α oscillations leading to the following simplified and widely used expression for P -wave speed anisotropy,

$$v = \bar{v} [1 \pm f \cos(2\alpha)], \quad (1)$$

where \bar{v} is the mean or isotropic velocity and f is the fractional magnitude of the velocity variations. A positive or negative fraction indicates a seismically fast (e.g. olivine A-type fabric; Karato *et al.* 2008) or slow (e.g. aligned fractures) symmetry axis, respectively. The value of this sign term is fixed throughout the inversion. Noting that $\cos(\alpha)$ is equivalent to the dot product of the ray directional vector with the anisotropic symmetry axis vector, we can expand eq. (1) as,

$$v = \bar{v} [1 \pm f (2[\cos(\theta)\cos(\gamma)\cos(\phi - \psi) + \sin(\theta)\sin(\gamma)]^2 - 1)], \quad (2)$$

where θ and γ are the ray and symmetry axis elevation and ϕ and ψ are the ray and symmetry axis azimuth, respectively. See Fig. S1 for coordinate system angle conventions. The three anisotropic parameters (f , ψ , γ) and mean slowness ($\bar{u} = 1/\bar{v}$) are defined at every node in the forward grid (uniformly spaced at 10 km intervals) and eq. (2) is used to estimate wave speeds for anisotropic traveltimes calculations. In this study, the sign of the anisotropic fraction is assumed positive everywhere—a common assumption made in modelling upper mantle seismic anisotropy based the dominant forms of olivine lattice preferred orientation (Karato *et al.* 2008). For generality, we retain the '±' notation throughout the paper.

3.3.2 Approximate frequency-dependent traveltimes

Owing to the finite-frequency content of the seismic wavefield, phase arrival times reflect structure in a volume around the geometric ray path. The sensitivity of traveltimes to velocity heterogeneity

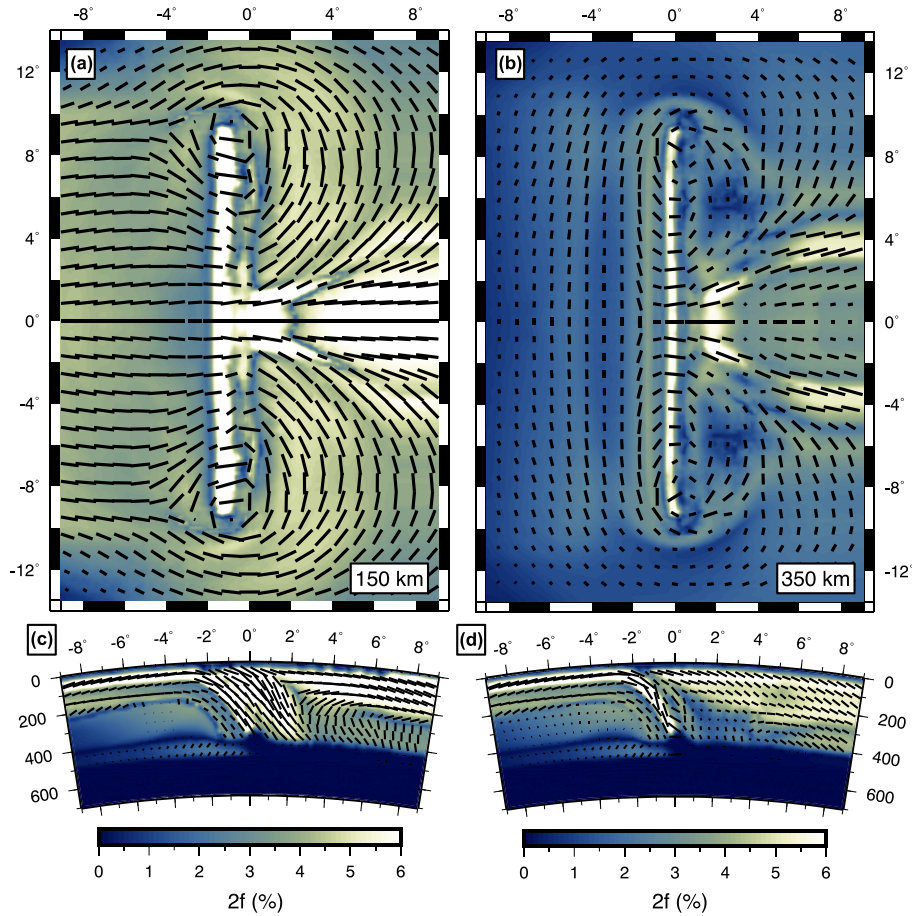


Figure 2. Cross-section through the true anisotropic model. Hexagonal symmetry axis vectors are plotted over the peak-to-peak magnitude of P -wave velocity anisotropy at (a) 150 km and (b) 350 km depth. East–west cross-section are shown at (c) 0°N and (d) $4^\circ 30'\text{S}$. Symmetry axis vectors are scaled by the anisotropic magnitude and projected onto the cross-section plane. Note that the anisotropic structure is symmetric about 0°N .

in this volume, commonly referred to as the banana–doughnut kernel, is described by Marquering *et al.* (1999) for the particular case of arrivals measured via waveform cross-correlation. Dahlen *et al.* (2000) and Hung *et al.* (2000) present methods for computing these velocity kernels in spherically symmetric media and these approaches have seen wide-spread use in teleseismic imaging studies. However, computation of the velocity kernels can become expensive because they require the calculation of forward and backward traveltimes for each station–receiver pair. Additionally, the velocity kernel formulas presented by Dahlen *et al.* (2000) and Hung *et al.* (2000) are only valid for isotropic heterogeneity. While anisotropic kernels have since been developed, they either assume a particular anisotropic geometry (e.g. vertically transverse isotropy; Zhao & Jordan 1998; Calvet *et al.* 2006) or use adjoint methods with spectral element wavefield modelling to compute the sensitivities (Sieminski *et al.* 2007, 2009). Here, we use a simple analytic approximation to the velocity kernel geometry originally suggested by Schmandt & Humphreys (2010) and show that it can be extended to anisotropic models.

We can write the frequency-dependent traveltimes as,

$$t = t' + \int_V (u - u') K dV, \quad (3)$$

where t' is the traveltimes predicted through the 1-D reference slowness (i.e. the inverse of velocity) model defined by u' ; u is the true 3-D slowness model and K is the Born slowness sensitivity kernel that maps a change in slowness to a delay time determined via cross-correlation with a reference waveform (Dahlen *et al.* 2000). Schmandt & Humphreys (2010) noted that the cross-sectional shape of the Born slowness sensitivity kernel within the first Fresnel zone can be approximated as,

$$K(x, r) = \frac{Q}{\pi R_f^2(x)} \sin\left(\pi \frac{r^2}{R_f^2(x)}\right), \quad (4)$$

where x is the along-ray distance; r is the ray-normal distance and R_f is the radius of the first Fresnel volume. The sine term is scaled such that the integral of a slowness kernel segment is proportional to the length of the corresponding ray segment and Q is a dimensionless scalar chosen such that the integral of the full slowness kernel equals the length of the ray path within the imaging volume. In other words, the integrated sensitivity of the Born slowness kernel equals the ray theoretical sensitivity (Hung *et al.* 2000). In practice, K is smoothed in the ray-normal distance to account for uncertainties in ray location (Schmandt & Humphreys 2010). We refer to this approximation as the heuristic finite-frequency slowness kernel (HFFK) because it represents a practical alternative description of the true Born slowness kernel that is sufficient for traveltimes tomography. The

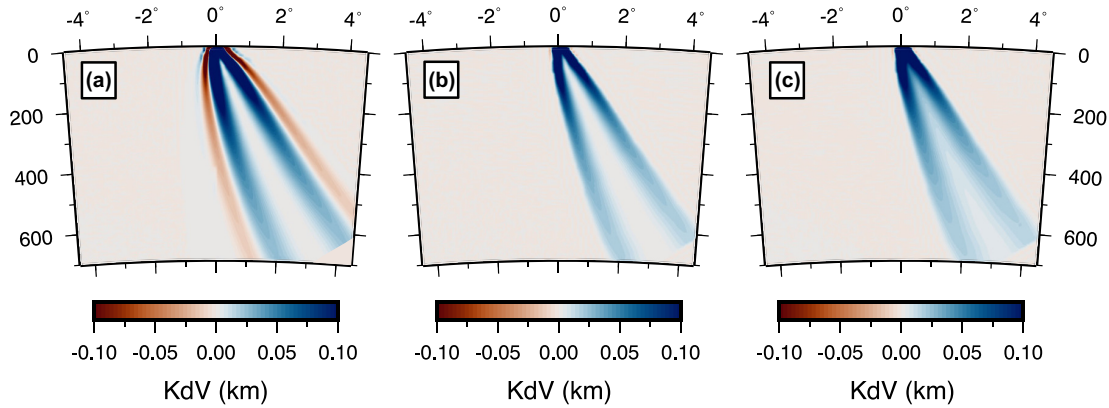


Figure 3. Comparison of traveltime sensitivity kernels for slowness perturbations. The (a) Born slowness sensitivity kernel computed following Hung *et al.* (2000) is compared to our approximate slowness kernels (b) without and (c) with ray-normal smoothing. Kernels are computed for a P wave with a period of 7.5 s recorded at a range of 81° . Cross-section is in the vertical plane containing the ray path.

structure of a Born slowness kernel and a HFFK are compared in Fig. 3. The sensitivities are qualitatively very similar, the most notable difference being that the Born slowness kernel has a small negative side-lobe.

For a direct teleseismic phase, a sufficient approximation for R_f is,

$$R_f = \left[\frac{Tx(L-x)}{Lu(x)} \right]^{1/2}, \quad (5)$$

where T is the dominant period of the seismic phase and L is the total ray length. To extend eqs (3)–(4) to anisotropic models, we simply define u using the apparent slowness in the direction of ray propagation (i.e. inverse of eq. 2). For this calculation, each node in K maps to a particular ray segment.

The absolute velocity kernel presented by Dahlen *et al.* (2000) is applicable for delays measured by cross-correlating an observed waveform with a reference waveform computed through a 1-D model. However, in local- or regional-scale imaging studies relative teleseismic delays are often derived from differential delays measured between stations recording the same event (i.e. multi-channel cross-correlation; VandeCar & Crosson, 1990). Maupin & Kolstrup (2015) demonstrate that the velocity kernel for a relative delay is the absolute kernel minus the average of all absolute kernels for a particular event. The removal of the mean kernel does not significantly alter sensitivities within the array footprint but does extinguish sensitivity outside the array where all kernels converge. This is precisely why relative delays are used—far field structure can be ignored. A relative HFFK is easily constructed by subtracting the mean of all HFFKs from each individual kernel for a given event.

While the HFFKs have been used in a number of studies for isotropic imaging (e.g. Bezada *et al.* 2010; Schmandt & Humphreys 2010; Villagómez *et al.* 2014; Byrnes *et al.* 2017; Bodmer *et al.* 2018; Portner *et al.* 2020), they have never been validated. By way of comparison, we show that this approximation yields more accurate relative traveltimes compared to ray theory for both isotropic and anisotropic models considered in this study. To quantify the error in our finite-frequency approximation, we compared cross-correlated relative delay times measured for the isotropic and anisotropic subduction zone models to those computed using ray theory, the HFFKs, and the paraxial Born slowness kernels. Mean station delays for each method are shown in Figs S2 and S3 for the isotropic and

anisotropic models, respectively. The distribution of traveltime errors predicted by each method are shown in Fig. S4. As expected, ray theory overpredicts the magnitude of delays and yields sharper gradients in delay time variations. The root-mean-squared error (RMSE) of the ray theoretical times with respect to SPEC-FEM is 254 and 325 ms for the isotropic and anisotropic models, respectively. The HFFKs provide a substantial improvement over ray theory yielding a RMSE of 148 and 165 ms for the isotropic and anisotropic cases. The paraxial slowness kernels provide a similar fit for the isotropic data with a RMSE of 147 ms. The similarity between the HFFK and the paraxial Born slowness kernel predictions suggests that traveltime errors are largely the result of discretization, assuming a single dominant period and differences between the true and paraxial Born slowness kernel structure arising from 3-D heterogeneity. For comparison, we estimate that the RMS error in the SPEC-FEM modelled delays is ~ 56 ms. This estimate was obtained by computing delays through the IASP91 radial Earth model following the same workflow used to create our synthetic subduction zone data set (Section 3.2). In the absence of any numerical or measurement errors, the delay times computed through the 1-D reference model should be identically zero. However, the RMS delay we measured was 56 ms. In all cases, the errors in our forward calculation are smaller than the data fit generally achieved by teleseismic tomographic models (300–500 ms).

Though not theoretically founded, our approximation for anisotropic finite-frequency slowness kernels has proved successful in yielding more accurate traveltimes relative to ray theory. The approximate slowness kernels also provide the main benefits of their more accurate counterparts at reduced computational expense. Most importantly, they provide a framework for the inversion of broadband traveltime measurements. They also act as physically based smoothing criteria resulting in a better posed inverse problem that requires less stringent regularization. Finally, we note that the true slowness kernel structure will depend on the true heterogeneity. However, in practice Born kernels are commonly computed using 1-D radial earth models under the assumption that small perturbations in slowness do not significantly influence kernel structure. Because the change in slowness caused by an anisotropic perturbation is expected to be similar in magnitude to an isotropic perturbation (i.e. a few percent), this assumption likely holds true and is supported by the aforementioned comparison between ray-theoretical and HFFK anisotropic traveltimes (Figs S2–S4).

3.4 Inverse problem

3.4.1 Anisotropic partial derivatives

A linear relationship between the traveltimes and anisotropic parameters is required for the inversion. This can be obtained by directly differentiating eq. (2) with respect to the spherical anisotropic parameters (i.e. f , ψ and γ). Such an approach was implemented by Munzarová *et al.* (2018a). However, the resulting partials are strongly dependent on the anisotropic starting model which may be poorly known. Additionally, regularization of the inversion via smoothing constraints can be problematic for angular quantities given their values are modulo 2π . We suggest an alternative parametrization for arbitrarily oriented hexagonal anisotropy in terms of coefficients that are combinations of the symmetry axis vector components. This parametrization is similar to those used in previous azimuthally anisotropic tomography studies (e.g., Eberhart-Phillips & Henderson 2004) but without making any assumptions regarding the dip of the symmetry axis. While our proposed parametrization is still non-linear, it has the advantage of separating variables that control the orientation and magnitude of azimuthal anisotropy from the variable that controls dip. This is beneficial because sensitivity to the azimuthal parameters are defined for isotropic models. This allows the inversion to determine the best-fitting azimuthally anisotropic solution without any prior assumptions regarding anisotropic structure. Importantly, we find that this initial azimuthal solution provides a sufficient starting point for subsequent iterations that include dip (which has zero-valued sensitivity when the model is isotropic). In contrast, parametrization in terms of f , ψ and γ requires an anisotropic starting model in which the fabric orientations are sufficiently close to the true solution (Munzarová *et al.* 2018a)—a criteria met by the azimuthal solution implicitly obtained through our parametrization.

We can write eq. (2) as a function of the symmetry axis vector components. Upon this substitution, and converting velocity to slowness (u)—the parameter solved for in the inversion—eq. (2) becomes,

$$u = \bar{u}[1 \pm ((n_1^2 - n_2^2) \cos(2\phi) + 2n_1n_2 \sin(2\phi)) \cos^2(\theta) - (n_1^2 + n_2^2) \sin^2(\theta) - n_3^2 \cos(2\theta) + 2n_3[n_1 \cos(\phi) + n_2 \sin(\phi)] \sin(2\theta)]^{-1}. \quad (6)$$

where the symmetry axis vector is defined as,

$$n = \begin{bmatrix} \sqrt{f} \cos(\psi) \cos(\gamma) \\ \sqrt{f} \sin(\psi) \cos(\gamma) \\ \sqrt{f} \sin(\gamma) \end{bmatrix}. \quad (7)$$

To invert directly for the components of n would require an anisotropic starting model such that the partial derivatives are defined. To avoid this requirement, we suggest the following reparametrization of eq. (6),

$$u = \bar{u}[1 \pm ([A \cos(2\phi) + B \sin(2\phi)] \cos^2(\theta) - G \sin^2(\theta) - C^2 \cos(2\theta) + \sqrt{2}C[s_1(G + A)]^{1/2} \cos(\phi) + s_2(G - A)^{1/2} \sin(\phi)] \sin(2\theta)]^{-1} \quad (8)$$

$$A = (n_1^2 - n_2^2) \quad (9)$$

$$B = 2n_1n_2 \quad (10)$$

$$C = n_3. \quad (11)$$

Introduction of the A , B and C terms results in a potential sign ambiguity that is accounted for by defining $s_1 = \text{sign}(n_1)$ and $s_2 = \text{sign}(n_2)$ in eq. (8) using values from the current model. We also define $G = \sqrt{A^2 + B^2}$ for notational convenience. The anisotropic parameters in eq. (2) can be recovered as $f = G + C^2$, $\psi = \arctan[B/(G + A)]$ and $\gamma = \arctan[C/G^{1/2}]$.

The relationship between a change in traveltime and a perturbation in mean slowness is derived via the chain rule,

$$\frac{\partial t}{\partial \bar{u}} = \frac{\partial t}{\partial u} \frac{\partial u}{\partial \bar{u}} = - \int_V K \frac{u^2}{\bar{u}} dV, \quad (12)$$

Similarly, we can write the partial derivative of the traveltime with respect to the three anisotropic variables (eqs 9–11) as,

$$\begin{aligned} \frac{\partial t}{\partial A} = \mp \int_V K \frac{u^2}{\bar{u}} & \left(\cos(2\phi) \cos^2(\theta) - \frac{\partial G}{\partial A} \sin^2(\theta) \right. \\ & + \frac{C}{\sqrt{2}} \left[s_1(G + A)^{-1/2} \left(\frac{\partial G}{\partial A} + 1 \right) \cos(\phi) \right. \\ & \left. \left. + s_2(G - A)^{-1/2} \left(\frac{\partial G}{\partial A} - 1 \right) \sin(\phi) \right] \sin(2\theta) \right) dV, \quad (13) \end{aligned}$$

$$\begin{aligned} \frac{\partial t}{\partial B} = \mp \int_V K \frac{u^2}{\bar{u}} & \left(\sin(2\phi) \cos^2(\theta) - \frac{\partial G}{\partial B} \sin^2(\theta) \right. \\ & + \frac{C}{\sqrt{2}} \left[s_1(G + A)^{-1/2} \cos(\phi) \right. \\ & \left. \left. + s_2(G - A)^{-1/2} \sin(\phi) \right] \frac{\partial G}{\partial B} \sin(2\theta) \right) dV, \quad (14) \end{aligned}$$

$$\begin{aligned} \frac{\partial t}{\partial C} = \mp \int_V K \frac{u^2}{\bar{u}} & (-2C \cos(2\theta) \\ & + \sqrt{2}[s_1(G + A)^{-1/2} \cos(\phi) \\ & + s_2(G - A)^{-1/2} \sin(\phi)] \sin(2\theta)) dV. \quad (15) \end{aligned}$$

We find this parametrization particularly convenient given that eq. (8) and the partials reduce to those widely used for azimuthal anisotropy (i.e. $C = n_3 = 0$) inversions where the model is isotropic (e.g. Eberhart-Phillips & Henderson 2004). Thus, isotropic starting models may be used in which case the first iteration will solve for the A and B parameters. Provided that ray paths are not exactly vertical where the model is also isotropic, sensitivity to anisotropic variables will be defined under this parametrization. Note that terms involving G assume null values when G goes to zero. Due to the non-linearity of the imaging problem, multiple iterations must be performed to converge to a solution. Because we assume 1-D rays, the purely isotropic inversions are linear and require only a single iteration. The AB anisotropic inversions converge after two iterations while the ABC anisotropic inversions converge after four iterations (Fig. 4). For the anisotropic inversions that require multiple iterations, the HFFK kernel is re-evaluated using the updated model parameters but the ray paths and associated kernel geometry remain the same.

The inversion scheme proposed here emphasizes the recovery of anisotropy in areas with strong (relative to the data uncertainty) 2ϕ signals. This because we sought a parametrization that would facilitate isotropic starting models such that no assumptions regarding the fabric orientation and strength are required. In this case, only the azimuthal parameters (A and B) have initially non-zero partials and this initial azimuthal solution provides a sufficient starting point for subsequent iterations that solve for fabric dip. We did explore inversions parametrized as a function of f , ψ and γ using a homogeneous anisotropic starting model as suggested by Munzarová *et al.* (2018a). Munzarová *et al.* (2018b) demonstrate

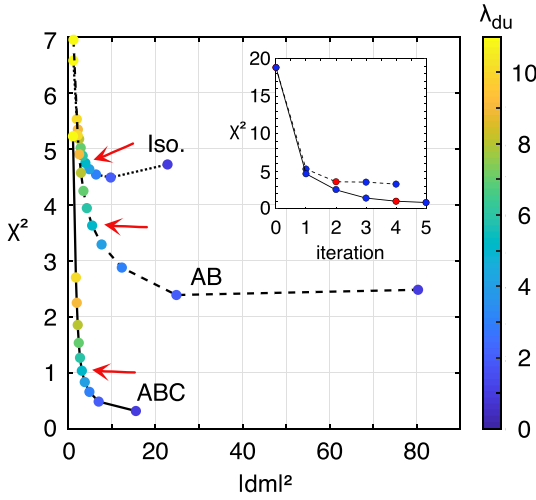


Figure 4. Selected trade-off curves for model selection. The squared norm of the delay time residuals normalized by the estimated data uncertainty (χ^2) is plotted as a function of the squared model norm ($|dm|^2$) for different values of λ_{du} (colour scale). The model norm is the length of the fractional slowness and anisotropic perturbational vectors. Trade-off curves for purely isotropic (dotted line), AB-anisotropic (dashed-line) and ABC-anisotropic (solid line) inversions are shown. All curves are constructed using $\lambda_s/\lambda_d = 10/1$. The ratio $\lambda_{da}/\lambda_{du}$ is 1/1 and 1/3 for the AB and ABC inversions, respectively. Red arrows point to the preferred solutions presented in Section 5 which all correspond to $\lambda_{du} = 5$. Inset shows the convergence of the preferred AB and ABC anisotropic inversions with our chosen solutions in red.

that the starting anisotropic fabric must be oriented sufficiently near the true structure to recover the input model. However, due to the diversity of anisotropy orientations in the subduction zone, we could not find a sufficient starting model and the resulting solutions always contained significant biases where the true and starting symmetry axes were at high angles to one-another. Note that spatially variable starting models based on, for example, SKS splits could be used but such models may provide a poor estimate of P -wave anisotropic structure due to their lack of depth-dependent information (Bezada *et al.* 2016). One could also generate an anisotropic starting model by first solving for azimuthal parameters. However, as we will show, purely azimuthal inversions in environments containing dipping fabrics can generate significant isotropic and anisotropic artefacts that could persist depending on how the inversion is regularized. Undoubtedly, other parametrizations of eq. (6) can be imagined that may be more or less suitable given the geometry of a particular imaging problem and prior structural constraints.

Inspection of eq. (8) yields a few useful insights into how teleseismic data ‘see’ regional anisotropic structure. It is clear that azimuthal variations in traveltimes are sensitive to all three anisotropic parameters. This is important considering the limited sampling of incidence angles by teleseismic rays within the imaging volume. Also apparent from eq. (8) is that delays accumulated beneath the local array will contain a significant 1ϕ azimuthal component for dipping symmetry axes superimposed on the expected 2ϕ pattern. When θ is poorly sampled, the recovered \bar{u} will represent the average slowness in a cross-section through the true slowness surface. Finally, it is important to note that the equations presented thus far describe absolute traveltimes. However, relative traveltimes (or delays if a 1-D prediction has been removed) are most often used for regional teleseismic imaging. While the equations are still appropriate, predicted data are demeaned in the same manner as the

input data. We discuss the implications of demeaning observations in Section 4.

3.4.2 Discretization, regularization and solution

Relative traveltimes are inverted for perturbations to \bar{u} , A , B and C . For convenience, we refer to azimuthal anisotropic inversions as ‘AB’ anisotropy and inversions that include symmetry axis elevation as ‘ABC’ anisotropy. The perturbational models are defined more coarsely than those used for the forward calculation with nodes spaced 50 km in each dimension. The discretization is sufficiently fine to capture the structure of the approximate finite-frequency slowness kernels which are mapped to the inversion grid via linear interpolation. While separate discretizations may be used for isotropic and anisotropic parameters, we use the same grid for both fields in the present analysis. To further constrain the anisotropic parameters, perturbations to A , B and C were limited to the upper 500 km of the mantle where anisotropy in the Earth is expected to be most significant. We also include event statics in the inversion that allow for mean shifts in the relative traveltimes associated with a given source. In practice, event statics accommodate hypocentral errors and inconsistencies arising from uneven sampling of the imaging volume. For all inversions, the starting model is isotropic and 1-D. The initial velocity profile is taken from the geodynamic model and closely follows IASP91 (Kennett & Engdahl 1991) but without a crustal layer.

After discretization, the linear system defined by eqs (12)–(15) is solved using the LSQR algorithm (Paige & Saunders 1982) subject to smoothing and damping constraints that are required to regularize the otherwise ill-posed inverse problem. We seek a solution that minimizes a function of the form,

$$y^2 = \Delta t' C_d \Delta t + \lambda_d \Delta m' C_m \Delta m + \lambda_s \Delta m' C_s \Delta m, \quad (16)$$

where Δt is the vector of data residuals; C_d is the data covariance matrix which we assume is diagonal and composed of the squared data uncertainties; Δm is the model perturbation vector; C_m is the model covariance matrix which is again assumed to be diagonal and composed of user-defined relative parameter uncertainties; C_s applies Laplacian smoothing to the model perturbations; λ_d and λ_s are weighting parameters that control the length and roughness of the solution vector relative to the length of the data residual vector. The damping and smoothing weights are defined separately for the isotropic (λ_{du} , λ_{su}) and anisotropic parameters (λ_{da} , λ_{sa}).

For purely azimuthal anisotropy, the A and B terms reduce to $f_{cos}(2\psi)$ and $f_{sin}(2\psi)$ and it is appropriate to directly apply the damping and smoothing operators to these parameters. However, when considering the elevation in addition to the azimuth the relationship between the A , B and C parameters to the phase and magnitude of anisotropy is more complex. Additionally, the magnitude of the A and B terms will generally be smaller than the C term since the former involve products of the symmetry axis components. If this discrepancy is not accounted for, then the inversion will preferentially perturb the A and B terms by virtue of their smaller values. When inverting for A , B and C parameters the relevant anisotropic quantities to damp and smooth are the vector components of the anisotropic symmetry axis (eq. 7). If L_m is the damping or smoothing equation for n_i symmetry axis component at the m th model node, then the regularization weights are obtained by differentiating L_m with respect to the A , B and C parameters.

Table 1. Inversion summary table. The type of data (isotropic/anisotropic), regularization weights, iteration at convergence, solution vector norms, delay time fit and relevant figures for each inversion presented in Section 5 are listed.

Parameter	Isotropic model 1	Isotropic model 2	AB model	ABC model
Data	Isotropic	Anisotropic	Anisotropic	Anisotropic
λ_s/λ_d	10	10	10	10
λ_{du}	5	5	5	5
λ_{da}	—	—	5	5/3
Iteration	1	1	2	4
$ du $	1.04	1.98	1.78	1.02
$ da $	0	0	1.46	1.44
χ^2	1.1	4.75	3.63	1.04
Figure	5	6	9, 10	11, 12

The choice of regularization weights has perhaps the greatest effect on the final image and is the most subjective part of the inversion. Our preferred values of λ_d and λ_s were chosen through inspection of L-curves (e.g. Aster *et al.* 2005) where the squared data residual norm is plotted against the squared model norm. We use a normalized data residual defined as, $\chi^2 = |dd^2/\epsilon^2|$, where dd is the delay time residual and ϵ is the uncertainty in the delay time measurement. Ideal solutions are considered those near the corner of the curve where further reduction of the regularization does not substantially improve the data fit. First, we systematically explored values of λ_{du} at different fixed ratios of $\lambda_{su}/\lambda_{du}$ in isotropic inversions of the isotropic data set. After identifying an appropriate smoothing-to-damping ratio, we ran a series of anisotropic inversions in which we varied λ_{du} over a range of $\lambda_{da}/\lambda_{du}$ ratios adopting the preferred smoothing-to-damping ratio previously identified for both the isotropic and anisotropic parameters. Through these analyses it became apparent that smoothing of the final image was largely controlled by the HFFKs and the choice of λ_s had only a minor effect on the image. Selected L-curves from this analysis using our preferred smoothing-to-damping and slowness-to-anisotropic parameter damping ratios are shown in Fig. 4; see Table 1 for list of inversion parameters. Note that in Fig. 4 the model norm is the total length of the fractional slowness perturbations (slowness perturbation weighted by starting model value) and the fractional anisotropic perturbations yielding consistent units between the different parameter sets.

3.4.3 Resolution

Assessing the resolution of tomographic models remains a challenge. The computational costs for obtaining resolution matrices via methods such as singular value decomposition (e.g. Zhang & Thurber 2007) are often prohibitively expensive for problems containing a large number of observations and parameters (such as those considered here). Consequently, we rely on traditional more qualitative measures of model fidelity. Specifically, the derivative weight sum (DWS) provides a reasonable proxy for resolution (e.g. Toomey & Foulger 1989; Zhang & Thurber 2007). Selected views of the DWS are shown in Fig. S5 and regions with $DWS < 150$ are masked in the presentation of our isotropic results. To assess the quality of anisotropic solutions, we calculate the mean resultant length (MRL; Fisher 1995; Zhang *et al.* 2009) at each model node. If a node is directionally well-sampled, the MRL tends toward zero. Conversely, ray sampling with a strong directional bias will have a MRL near 1. Of course, teleseismic data have a strong bias toward

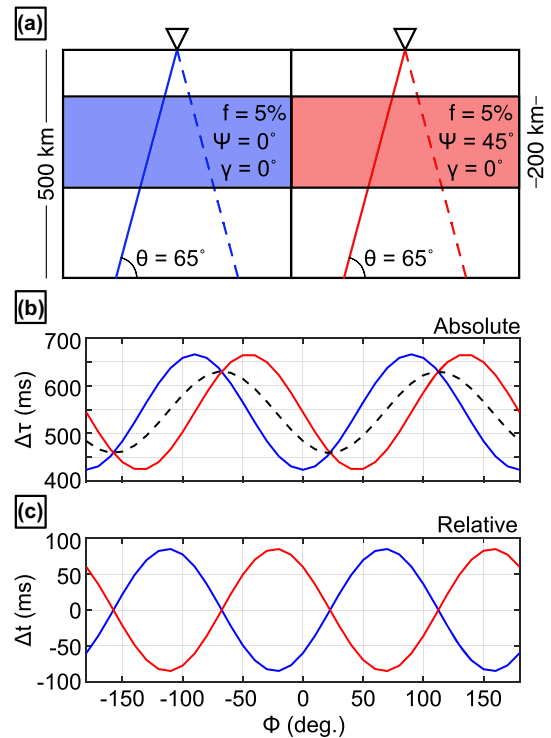


Figure 5. Example of how relative delay times can distort an anisotropic signal. (a) Illustration of simple azimuthally anisotropic model composed of two 200-km-thick anisotropic layers embedded in an otherwise homogeneous isotropic material. Two stations (inverted triangles) located above each layer record the same events evenly distributed in backazimuth at a constant range. Idealized ray paths for an event located to the west (solid lines) and east (dashed lines) are drawn. (b) Absolute traveltime residuals are plotted as a function of azimuth for the western (blue curve) and eastern (red curve) stations. Each azimuth represents a separate event. The dashed black curve is the mean traveltime residual for each event/azimuth. (c) Same as (b) but after removing the mean traveltime residual for each event.

the vertical axis and variations in directional sampling are better highlighted by considering the length of the azimuthal component of the MRL vector. Maps of the azimuthal MRL are also shown in Fig. S5. Directional sampling of our synthetic data set is rather uniform given the uniform distribution of events. Poorer azimuthal coverage is observed near the edge of the model and at depth where ray crossing is more limited. Anisotropic solutions presented below are masked in areas where the azimuthal MRL > 0.5 . The construction of ray density tensors (Kissling 1988; Munzarová *et al.* 2018a) provides another measure of directional sampling that better illustrates the 3-D distribution of ray paths and is useful when considering more heterogeneous data distributions.

To assess the influence of solution regularization and potential trade-offs between isotropic and anisotropic heterogeneity we performed a series of simple block recovery tests (Rawlinson & Spakman 2016). We considered five synthetic models each containing a single cylindrical anomaly with a radius of 150 km that extends from 100 to 400 km depth and is placed at the centre of the experiment. Two purely isotropic anomalies were considered defined by a +4 per cent and -4 per cent change in velocity. Three purely anisotropic anomalies are defined with $f = 5$ per cent, $\psi = 60^\circ$ and $\gamma = 0^\circ$, 30° or 60° . Isotropic, AB-anisotropic, and ABC-anisotropic inversions to reconstruct each synthetic model were performed using our preferred inversion parameters (Table 1)

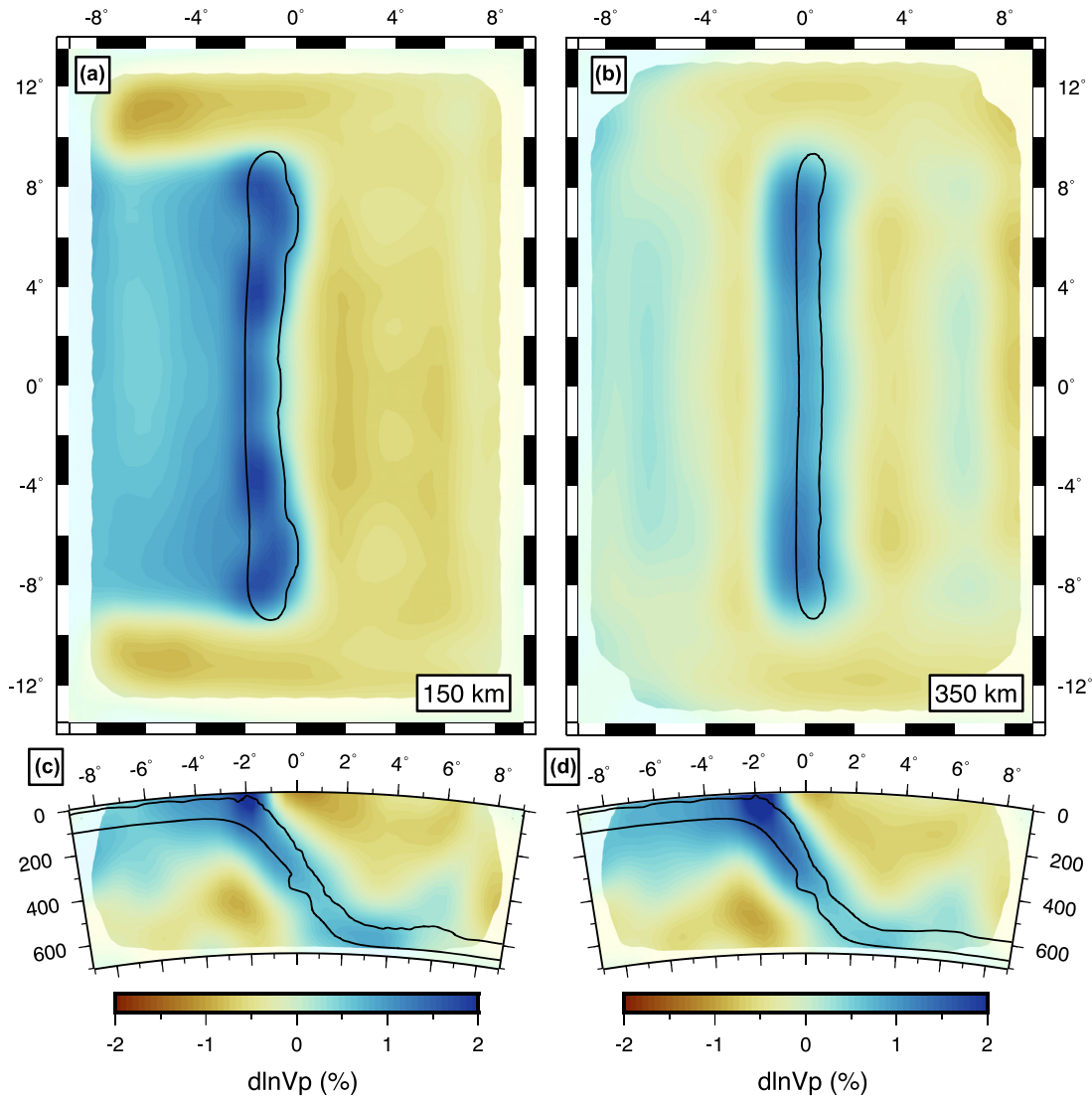


Figure 6. Isotropic inversion of isotropic data. Velocity perturbations with respect to the 1-D starting velocity model are plotted at (a) 150 km and (b) 350 km depth. East–west cross-sections are plotted at (c) 0°N and (d) 4°30'S. The 1 per cent velocity contour in the true model is drawn in black and outlines the true slab geometry. Areas with poor data coverage are masked.

with the exception that anisotropic perturbations are not depth limited. The results highlight potential biases implicit in the inversion strategies.

The reconstruction of the high-velocity and low-velocity isotropic cylinders are shown in Figs S6 and S7, respectively. All three inversion strategies accurately recover the true anomaly. However, perturbations are smeared vertically by ~ 100 km owing to the steep incidence of the ray paths. Importantly, minimal anisotropic structure is recovered in the anisotropic inversions. This suggests that the risk of mapping truly isotropic heterogeneity into anisotropic parameters is minimal in a directionally well-sampled model.

Isotropic structure recovered from inversions of the different anisotropic data sets are shown in Fig. S8. Despite the fact that the anomaly is directionally well-sampled, purely isotropic inversions generate significant artefacts ($d\ln V > 1$ per cent). Dipping anisotropic fabrics tend to create clover shaped patterns of high and low-velocity artefacts. The magnitude of these artefacts is reduced as a more complete description of anisotropy is considered.

Anisotropic structure recovered from inversions of the different anisotropic data sets are shown in Figs S9 and S10. All inversions accurately recover ψ except the AB-anisotropic inversion when γ is 60° (Fig. S9). In cross-section (Fig. S10), we see that the AB-anisotropic inversion creates significant anisotropic artefacts surrounding the true anomaly when the symmetry axes are inclined from horizontal. These spurious anisotropic anomalies tend to be oriented orthogonal to the true ψ . This highlights a potentially significant issue for interpreting azimuthally anisotropic seismic models in environments that may contain dipping fabrics. This issue is further discussed when we present the azimuthally anisotropic inversions for subduction zone structure.

These artefacts are effectively absent from the ABC-anisotropic solutions which generally constrain γ to within 10–15°. However, there is a tendency for the recovered anisotropic structure to elongate in the direction of the symmetry axis especially for steeply dipping fabrics. Based on this set of synthetic tests, the ABC-anisotropic inversion seems well-suited for imaging purely isotropic as well as arbitrarily oriented anisotropic fabrics. We note that anisotropy will

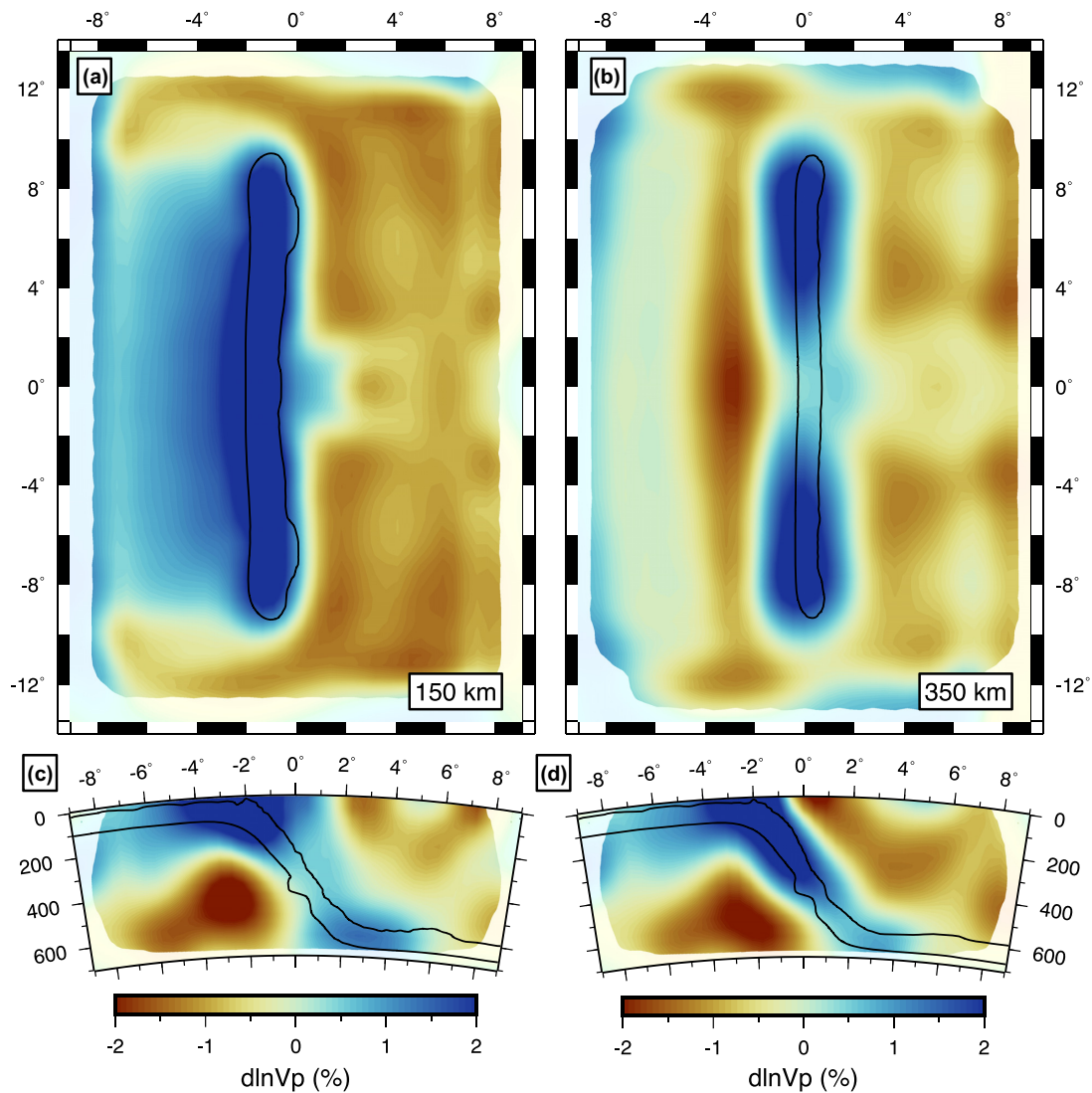


Figure 7. Isotropic inversion of anisotropic data. Velocity perturbations are plotted as in Fig. 6.

become progressively more difficult to image as the symmetry axis approaches vertical and the azimuthal signal in the delay times is reduced.

4 ANISOTROPY IN RELATIVE TRAVELTIMES

Before exploring synthetic tomographic models, it is worth considering how anisotropy is seen by the regional teleseismic imaging problem. The inversion is driven not by absolute traveltime residuals but relative traveltime residuals measured with respect to some unknown absolute time (Aki *et al.* 1977). Consequently, the true mean of the resulting model perturbations is not constrained (Aki *et al.* 1977; Lévêque & Masson 1999). While the implications of this for the interpretation of isotropic anomalies are generally well-understood, the consequences for anisotropic images are not. Here we consider how anisotropy manifests in relative residuals and its significance for constraining anisotropic heterogeneity.

Regional teleseismic tomography is built upon the assumption that heterogeneity outside the imaging volume equally affects all

traveltimes for a specific event which is reasonable considering the similarity in ray paths and relatively large Fresnel volumes outside the array (Masson & Romanowicz 2017). With this assumption, by rendering the traveltimes relative through the demeaning of all arrivals associated with a specific event, arrival time perturbations associated with paths outside the imaging volume cancel along with any error in origin time. The resulting residuals reflect traveltime time differences accumulated within the tomographic model subject to some mean offset. Neglecting any errors in the source hypocentre, we can define the relative traveltime residual as,

$$\Delta t_{ij} = \tau_{ij} - t_{ij} - \frac{1}{N} \sum_{j=1}^N (\tau_{ij} - t_{ij}), \quad (17)$$

where τ_{ij} is the observed/true traveltime through the imaging volume for the i th-event and j th-station; t_{ij} is the predicted time and N is the total number of stations recording the i th-event.

Now we examine how anisotropy is mapped into this residual. For simplicity, we assume that local teleseismic ray paths for a particular event can be approximated with a single line segment characterized

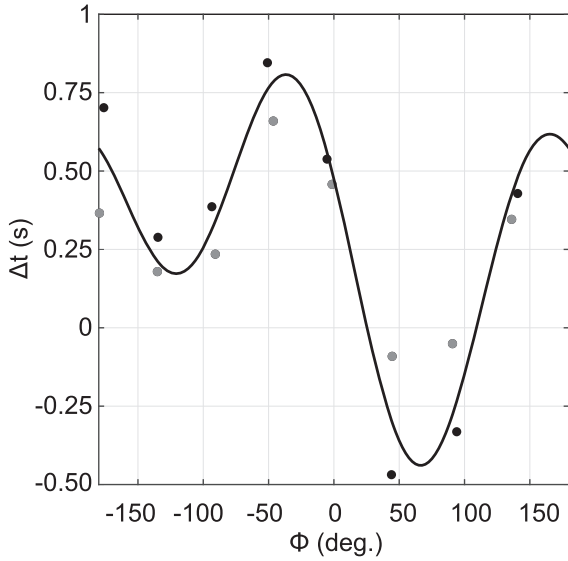


Figure 8. Example of azimuthal variations in relative P -wave delays observed for a station located at $5^{\circ}23'S$, $4^{\circ}24'E$. Grey and black points correspond to events with an epicentral distances of 80° and 50° , respectively. Black curve is the result of fitting eq. (18) to the black and grey points and is plotted for $\theta = 55^{\circ}$ (i.e. the mean ray elevation predicted by TauP). Note the strong 1ϕ trend in addition to the 2ϕ oscillations.

by a constant θ (Fig. 5a). Using the expression for anisotropic slowness in eq. (6) and assuming a weak anisotropic magnitude (i.e. $[1 \pm f]^{-1} \approx [1 \mp f]$), we can write the average slowness perturbation along a ray path (i.e. $\Delta t_{ij}/L_{ij}$) within the imaging volume as the following harmonic series,

$$\begin{aligned} \frac{\Delta t_{ij}}{L_{ij}} \approx & (U_j - \bar{U}) \cos(2\phi_{ij}) \cos^2(\theta_{ij}) + (V_j - \bar{V}) \sin(2\phi_{ij}) \cos^2(\theta_{ij}) \\ & + (W_j - \bar{W}) \cos(\phi_{ij}) \sin(2\theta_{ij}) + (X_j - \bar{X}) \sin(\phi_{ij}) \sin(2\theta_{ij}) \\ & + (Y_j - \bar{Y}) \sin^2(\theta_{ij}) + (Z_j - \bar{Z}), \end{aligned} \quad (18)$$

where L_{ij} is the ray length within the imaging volume; the coefficients U_j, \dots, Z_j are non-linear combinations of \bar{u} and n averaged across all paths taken to the j th-station; the coefficients \bar{U}, \dots, \bar{Z} are similarly defined but represent averages over the entire array introduced through demeaning. eq. (18) effectively describes the P -wave residual sphere discussed by Plomerová *et al.* (1996). Clearly, subtraction of \bar{U}, \dots, \bar{Z} will distort the directional variations in residuals that constrain the anisotropic parameters potentially yielding misleading results. Considering that teleseismic ray azimuths and elevations do not vary significantly across a regional array for a particular event, this distortion amounts to removing some model-averaged azimuthal trend. Note that if the imaging volume is well-sampled by each event and is anisotropically heterogeneous, then the \bar{U}, \dots, \bar{X} terms will tend toward zero.

Eq. (18) can be fit to relative P -wave delays as a first-order indication of anisotropy. The fit can be simplified by assuming a constant θ —a reasonable approximation for a regional teleseismic study—allowing the Y and Z terms to be combined into a single parameter. This curve fitting is analogous to a shear wave splitting intensity analysis (Chevrot 2000). As with splitting intensity, the magnitude of the 1ϕ -component relates to the dip of the symmetry axis but the nonlinear relationships among the harmonic coefficients precludes a meaningful definition of this angle. However, we can

define the apparent anisotropy azimuth as,

$$\psi_{app} = \arctan \left(\bar{V}_j \left[\bar{U}_j + \sqrt{\bar{U}_j^2 + \bar{V}_j^2} \right]^{-1} \right), \quad (19)$$

where $\bar{U}_j, \dots, \bar{Z}_j = (U_j - \bar{U}), \dots, (Z_j - \bar{Z})$. The mean station delay would be $\bar{Y}_j + \bar{Z}_j$. This analysis is applied to our synthetic data in Section 5.1 and could easily be applied to real P -wave data as an initial assessment of potential anisotropic structure. However, limited azimuthal coverage and/or strongly biased recording of events across an array may limit the applicability. Additionally, the relative nature of the delays may make direct interpretation of the results difficult as we demonstrate below.

In Fig. 5, we illustrate one potential pitfall of using relative travel-times to image anisotropic structure. We consider a model containing two uniform anisotropic layers defined by horizontal symmetry axes with $\psi = 0^{\circ}$ and $\psi = 45^{\circ}$. The model is uniformly sampled in azimuth by rays with constant incidence angle $\theta = 65^{\circ}$ (each azimuth represents a different teleseismic event). Azimuthal variations in absolute traveltime residuals accurately reflect the underlying structure. However, demeaned residuals are phase-shifted by 45° with respect to the true signal and have a reduced magnitude. The inversion of these relative traveltime residuals would obviously lead to spurious anisotropic results.

The consequences of demeaning on the anisotropic image will of course depend on the particular data geometry under consideration and the nature of anisotropic heterogeneity sampled. However, we can make some general comments on these consequences. (1) If the medium is sufficiently heterogeneous in terms of anisotropic structure or contains a significant volume of isotropic material and is well sampled by each event, the $\bar{V} - \bar{Y}$ terms in eq. (18) will tend toward zero and the true directional variations in the travel-times will be preserved. (2) A dominant anisotropic fabric in the imaging volume could bias the recovered anisotropic parameters. This is because demeaning will remove this trend from the data and modify the phase and magnitude of the relative traveltime variations (Fig. 5). This effect could be mitigated by using accurate anisotropic starting models and/or including data sensitive to absolute velocity structure. (3) While anisotropic checkerboard tests may provide a proxy for resolution within the imaging volume, they may be misleading in light of comment (2). The $\bar{V} - \bar{Y}$ coefficients will be near zero for alternating anisotropic orientations that are at large angles to one another. This may give the false sense that a large domain of coherent anisotropy in a particular solution is well resolved though it may be biased for the reasoning in comment (2). (4) Finally, a homogeneous anisotropic volume is invisible to the imaging method (i.e. $V_j, \dots, Y_j \approx \bar{V}, \dots, \bar{Y}$). Despite these potential limitations, we will show through the following synthetic examples that including anisotropy in regional teleseismic inversions can yield a more faithful representation of Earth structure that better fits the observations.

5 TOMOGRAPHIC RESULTS

In the following sections, we progress through a series of isotropic, AB-anisotropic and ABC-anisotropic inversions highlighting the ability of each imaging strategy to accurately capture subduction zone structure. We conclude with a quantitative comparison of the models in terms of their ability to fit the relative traveltime data and true model structure. All inversions are summarized in Table 1.

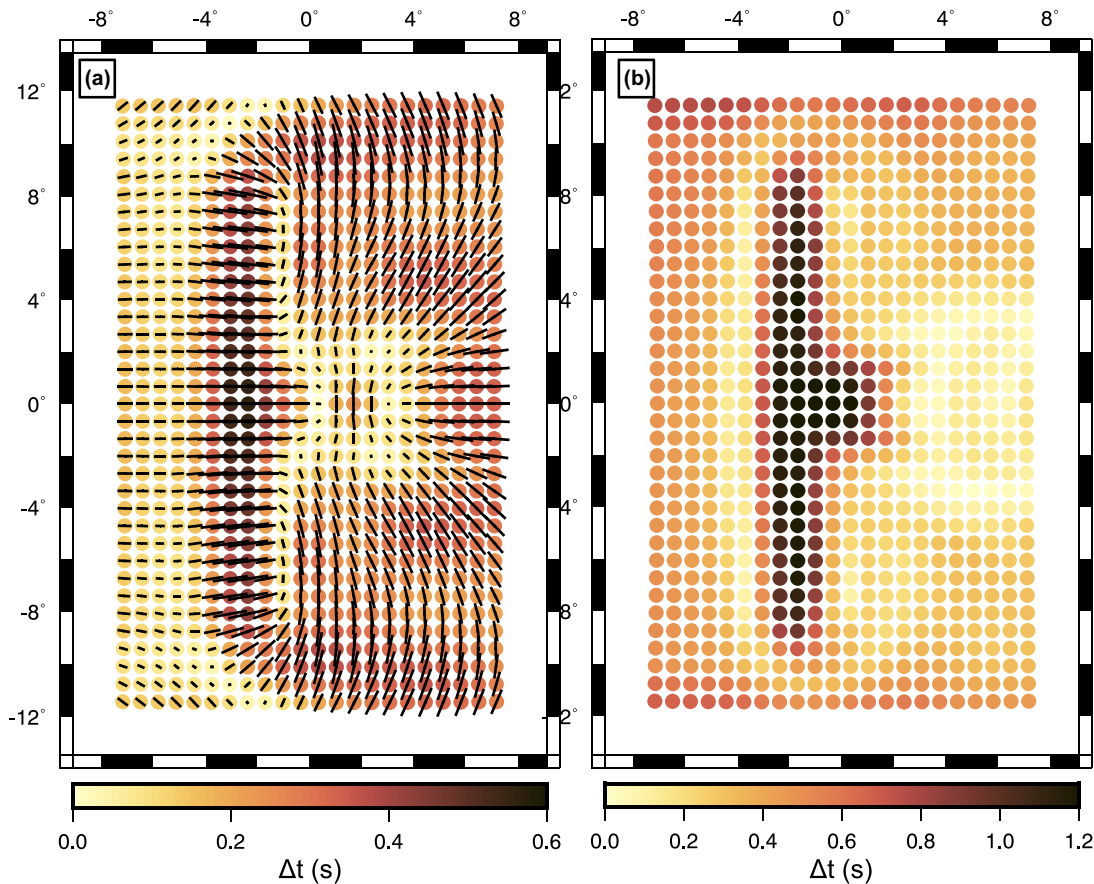


Figure 9. Apparent anisotropy from relative P -wave delays. (a) The apparent fast-azimuth (quivers; eq. 19) is plotted over the peak-to-peak magnitude of 2ϕ delay time variations predicted by fitting eq. (18) to the relative P -wave delays at each station. In (b), the peak-to-peak magnitude of the best-fitting 1ϕ component is plotted.

5.1 Isotropic solutions

Before considering the effect of anisotropy on the tomographic image, it is useful to establish an appropriate reference model. To do this, we first inverted a set of delays computed through a purely isotropic model. This solution (Fig. 6) represents the ideal recovery of isotropic structure, given the constraints of the imaging problem (e.g. finite-frequency approximation, imperfect data coverage), if the anisotropic component of the delay times was perfectly accounted for. The general geometry of the slab is reconstructed but with a significant amount of smoothing, which is to be expected given the relatively long period data inverted (15 s). Due to the relative nature of the data, apparent low-velocity zones (LVZs) are imaged above and below the slab. However, their magnitude is generally small and they are spatially distributed. A similar isotropic solution was obtained when we inverted the anisotropic delays using a starting model containing the true anisotropic structure (Fig. S11). This test indicates that our forward approximation to the anisotropic structure is accurate and does not generate any appreciable imaging artefacts.

In comparison, a purely isotropic solution to the anisotropic delays contains a number of significant artefacts (Fig. 7). The magnitude of LVZs increases throughout the model and are particularly concentrated beneath the slab. The apparent velocity within and around the descending slab is increased as a consequence of the steeply dipping anisotropy found in this region. More horizontal

symmetry axes, particularly westward of the synthetic trench, act to increase the apparent velocity contrast between the slab and surrounding mantle leading to stronger LVZs in the slab region. The nature of the LVZ artefacts are in many ways similar to those identified in the synthetic tomography study of Bezada *et al.* (2016). However, the Bezada *et al.* (2016) results represent the high frequency expression of anisotropic artefacts since their synthetic data was computed using ray theory. Unlike Bezada *et al.* (2016), we also find significant distortion in slab geometry with along-strike changes in LVZ amplitude accompanied by along-strike changes in LVZ amplitude. This difference may be due to the fact that Bezada *et al.* (2016) only modelled half of the subduction zone. Any practitioner of tomography can easily imagine a number of intriguing interpretations to which such velocity structure could be, in this case, falsely attributed.

A clear anisotropic signal can be identified in the data prior to performing anisotropic inversions by fitting eq. (18) to azimuthal variations in station delays. The results of this curve fitting are shown for a station at $5^{\circ}23'S$, $4^{\circ}24'E$ in Fig. 8 and summarized for all stations in Fig. 9; the average θ_{ij} predicted by TauP for each ray was used for the curve fitting. The azimuth of anisotropy has a pattern very similar to that constrained by SKS splitting intensity (Fig. S12) and clearly captures the toroidal flow pattern around the slab. The magnitude of the 1ϕ -component correctly identifies the area around the slab as having steeper dipping fabrics. Also

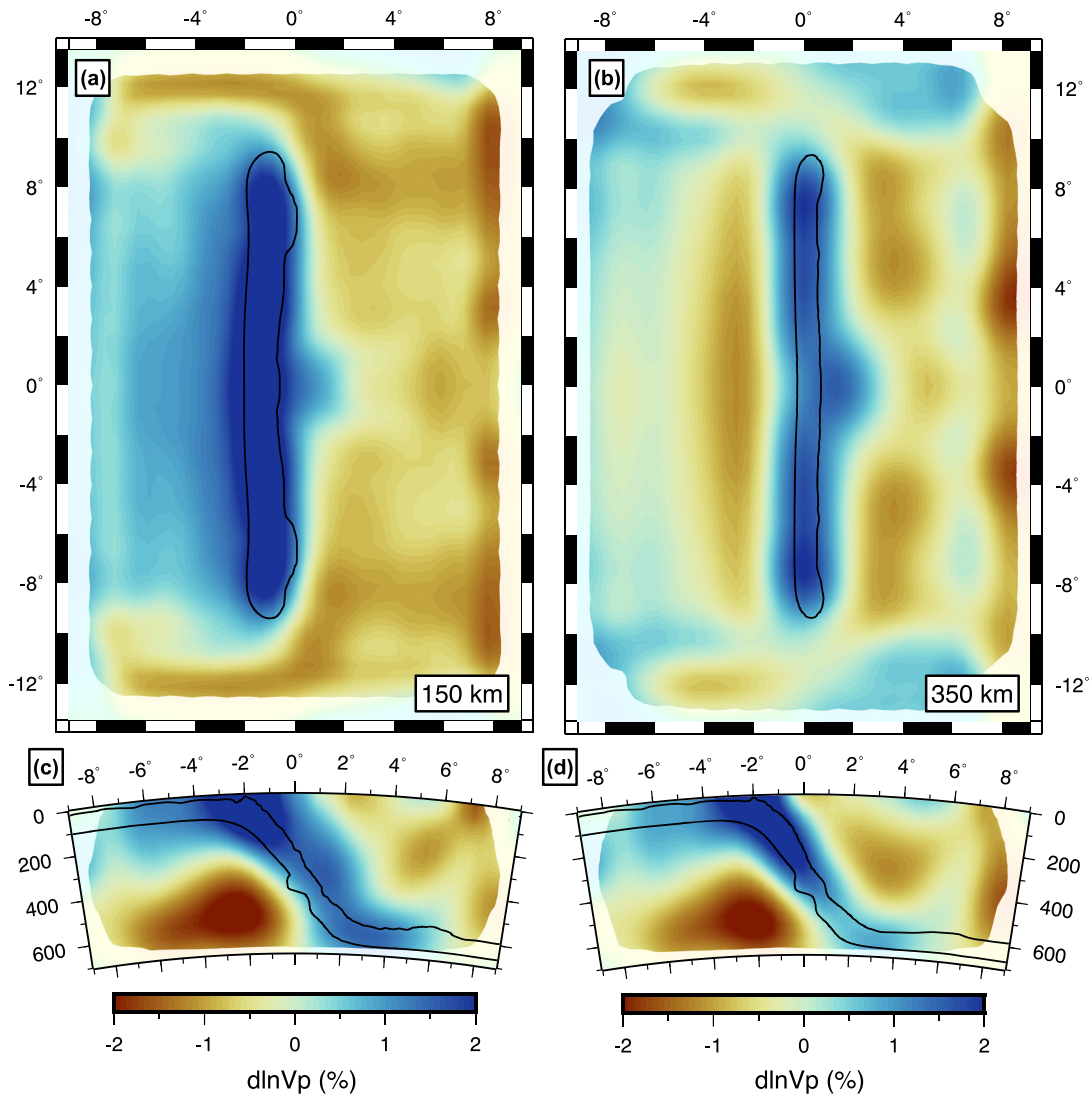


Figure 10. Isotropic structure recovered from azimuthally anisotropic inversion. Velocity perturbations are plotted as in Fig. 6.

note that magnitude of the 1ϕ - and 2ϕ -components are as strong as the station-averaged delays (Fig. S4a) attesting to the influence of anisotropy on the data.

5.2 AB anisotropic solutions

Next we consider the most common type of anisotropic imaging—the inversion for azimuthal anisotropic parameters. It is clear from the recovered isotropic structure (Fig. 10) that inclusion of the azimuthal anisotropic parameters does not significantly reduce the presence of the LVZ artefacts. This is perhaps not surprising considering that changes in the symmetry axis dip have the greatest influence on the velocity perceived by P waves with steep-incidence angles. However, the magnitude of the LVZs is modestly reduced and the slab amplitude appears more uniform relative to the purely isotropic solution. There is an apparent thickening of the slab towards 0°N because of the thicker layer of entrained flow results in strong anisotropy with steeply dipping fast axes (Fig. 2).

The recovered anisotropic structure is shown in Fig. 11. While the azimuthally anisotropic model generally captures the toroidal

flow pattern around the slab, the azimuthal assumption introduces a number of artefacts. The most egregious of these artefacts is the apparent transition in ψ from a trench-normal to trench-parallel orientation beneath the incoming plate. While this structure occurs in the true model (Fig. 2), synthetic tests (Section 3.4.3; Fig. S10) demonstrate that it is an artefact arising from the inversion attempting to fit the 1ϕ signals in the data with parameters that only control 2ϕ oscillations. This behaviour is clearly observed in the resolution tests where AB anisotropic inversions are performed for an anisotropic block containing steeply dipping fabrics (Fig. S10). There is a tendency to construct zones of anisotropy with ψ orthogonal to the true orientation above and below the true structure. A similar pattern is seen about the slab in Fig. 11(c). As an additional test, we also inverted synthetic data created from a subduction zone model where we removed the zone of trench-parallel anisotropy. The recovered model was nearly identical to Fig. 11. This is a surprising result that calls into question the accuracy of azimuthally anisotropic models in environments where we may expect dipping fabrics.

There are other notable departures from the true model in addition to the erroneous sub-slab structure. At shallower depths

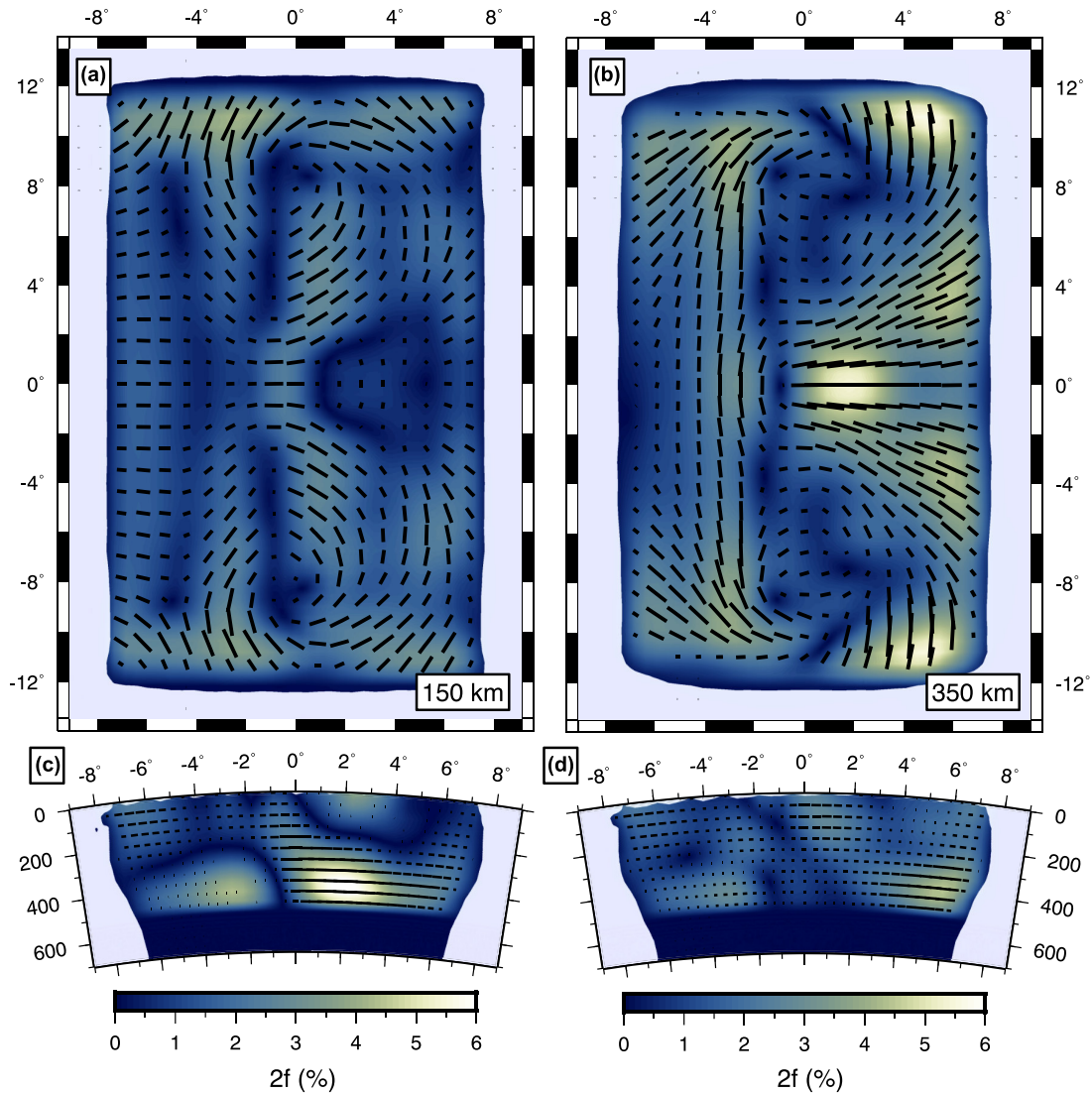


Figure 11. Azimuthal anisotropic structure recovered from inversion for \bar{u} , A and B terms. Seismic anisotropy is plotted as in Fig. 2. Areas with poor directional sampling are masked. Note that the trench-parallel orientations are actually an imaging artefact introduced by the azimuthal assumption; see Sections 3.4.3 and 5.2.

(<150 km) within the wedge near the centre of the plate the true model contains some of the strongest anisotropy but it is imaged as a weakly anisotropic zone where symmetry axis azimuths may be 90° from the true orientation. We suspect that this is due to the rapid changes in symmetry axis elevations with depth (Fig. 2c). The anisotropic magnitude is generally underestimated. This is primarily due to the steep incidence of the teleseismic rays resulting in a reduced 2ϕ -signal in the delay times. We conclude that the azimuthal simplification leads to both isotropic and anisotropic imaging artefacts in environments containing dipping anisotropic fabrics and the resulting models should be cautiously interpreted.

5.3 ABC anisotropic solutions

Inverting for all three anisotropic parameters has the most significant influence on the recovered isotropic model (Fig. 12). The LVZ artefacts are substantially reduced and the solution more closely

resembles the ideal recovery of isotropic structure (i.e. Fig. 6). There remains some distortion of the slab at 0°N where the true anisotropic fabric is strongest. Here, the geometry of the high-velocity anomaly gives the impression of a slightly steeper descending plate.

The anisotropic structure recovered from the ABC inversion (Fig. 13) more clearly captures the toroidal flow pattern without many of the artefacts observed in the azimuthal solution. Specifically, the trench-parallel fabrics westward of the subducting plate are not present. This area is largely imaged as isotropic. This makes sense considering it is characterized by two stacked layers with orthogonal symmetry axis orientations. Such structure would require more diverse sampling to resolve. Similar to the true model, we recover strong steeply dipping anisotropy within and around the descending plate and this zone weakens and thins toward the edges of the slab. Owing to strong vertical gradients in the symmetry axis elevation, weak anisotropic fabrics are recovered within the wedge near 0°N . Away from the mid-plate, the generally more inclined symmetry axes within the wedge are reconstructed. The

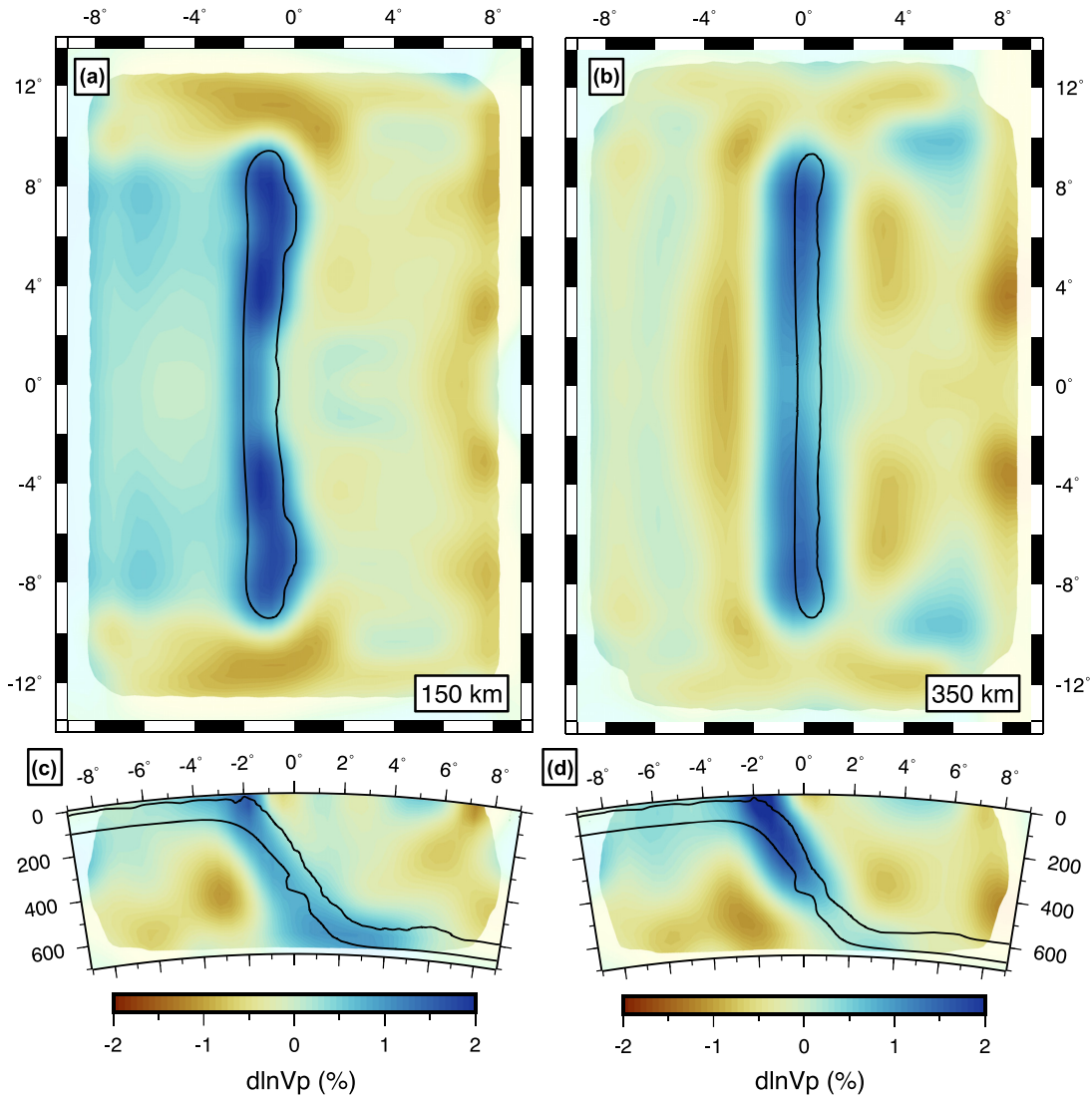


Figure 12. Isotropic structure recovered from anisotropic inversion that includes symmetry axis azimuth and elevation. Velocity perturbations are plotted as in Fig. 6.

anisotropic magnitude is generally underestimated throughout the model. In summary, the ABC anisotropic model captures the large-scale trends in ψ and γ and yields insights into the nature of mantle flow in the synthetic subduction zone.

It is important to acknowledge that maximum depth extent of anisotropic structure is poorly constrained in either the ABC or AB models. This is illustrated in Fig. S13 where we plot a solution obtained without imposing a depth limitation on anisotropic perturbations. In this case, anisotropy is distributed throughout the vertical extent of the model. However, lateral variations in anisotropic and isotropic structure remain unchanged. Caution should be exercised when interpreting deep anisotropy constrained solely by teleseismic P -wave delays.

5.4 Comparison of solutions

We have demonstrated that teleseismic P -wave delays can constrain many anisotropic features in subduction zones predicted by geodynamic models and reduce isotropic artefacts. However, assuming azimuthal anisotropy is a problematic simplification for environments

containing dipping fabrics that allows isotropic imaging artefacts to persist and introduces artefacts in the anisotropic solution. The inclusion of anisotropy does provide a better fit to the data (Figs 4 and 14; Table 1). While this is to be expected given the increase in the number of free parameters, it cannot solely be attributed to increased model complexity. From Fig. 4, it is apparent that for a given perturbational vector length, anisotropic models consistently yield better data predictions. As evident from Table 1 and illustrated in Fig. 14, as a more complete description of anisotropy is implemented in the inversion the norm of the slowness perturbations is reduced implying a less complex isotropic model nearer to the true solution. These trends suggest anisotropy is a necessary parameter for modelling the P -wave delays.

To assess how well the anisotropic solutions capture the true anisotropic structure, we evaluate individually the recovery of f , ψ and γ . We focus the comparison on the upper 400 km and only include nodes that are directionally well-sampled (Fig. S5). In Fig. 15, we present histograms of the error in f , ψ and γ in the AB and ABC solutions. Similar to isotropic anomalies, the anisotropic magnitude is on average underestimated by 1 per cent. The ABC solution

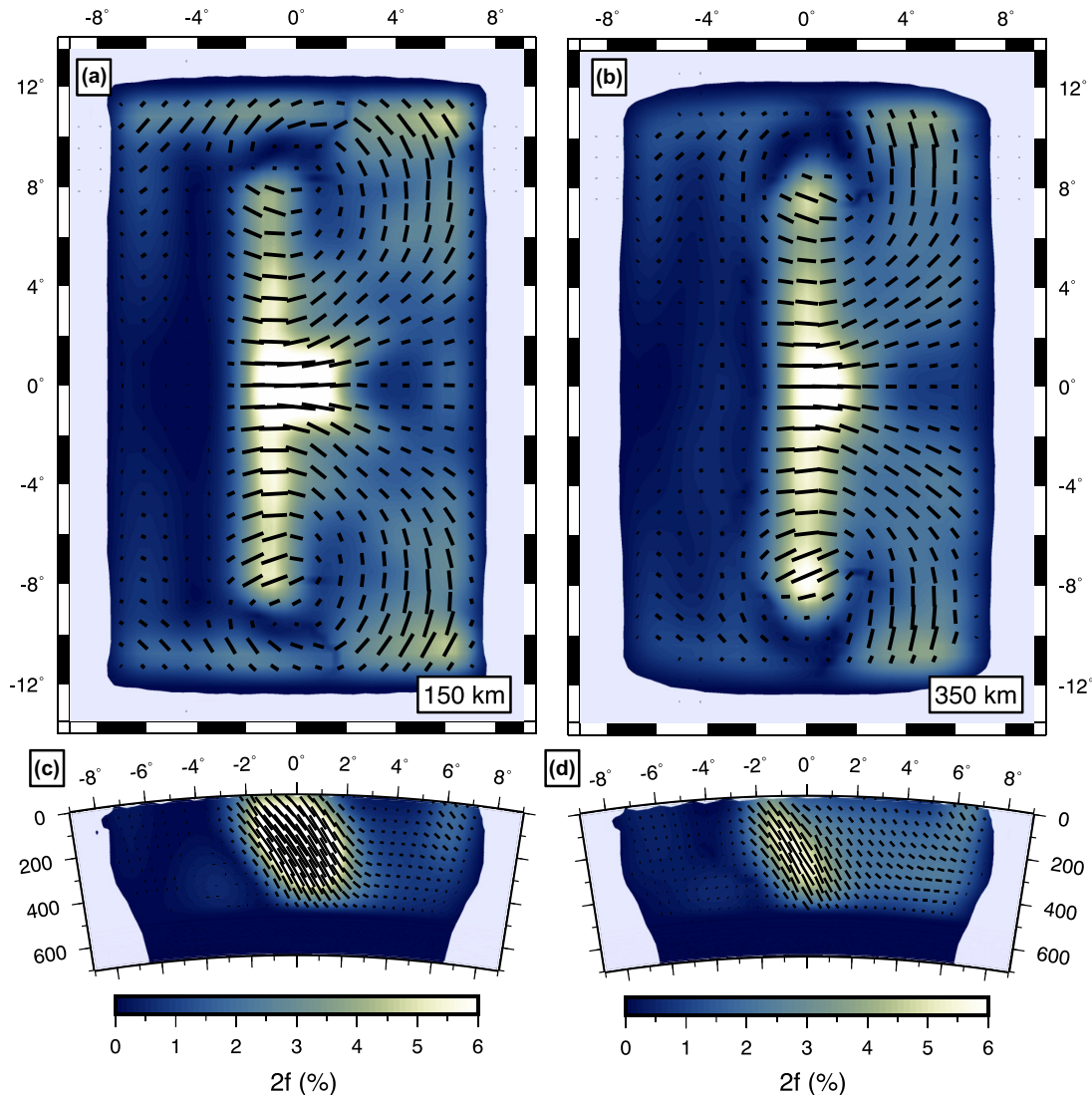


Figure 13. Anisotropic structure recovered from inversion for $\bar{\mu}$, A , B and C terms. Seismic anisotropy is plotted as in Fig. 2. Areas with poor directional sampling are masked.

provides a better average estimate of both the anisotropic azimuth and elevation. The weighted mean error in ψ for the AB and ABC models is 23° and 16° , respectively. Considering that the angular misfit loses significance as the anisotropic magnitude approaches zero, the weights for the mean calculation are defined as $w_k = \sqrt{f_k f'_k}$ where f_k and f'_k are the recovered and true anisotropic fractions at the k th model node. These misfits are comparable to but generally less than those measured between mantle azimuthal anisotropy constrained by surface waves and predicted from global circulation models (Becker *et al.* 2003, 2014). In this study, we can unequivocally attribute this error to limitations of the imaging method as opposed to unmodelled mantle dynamics. The weighted mean error in γ is 18° for the ABC solution compared to 26° if symmetry axes are restricted to the horizontal plane. For a more granular view of the symmetry axis orientation errors, see Figs S14 and S15 which show maps of the angular error in ψ and γ .

In Fig. 16, the weighted mean error in ψ is plotted as a function of depth and compared to orientations derived from SKS splitting intensity (Fig. S12) and azimuthal variations in P -wave delays (Fig. 9a) assuming the latter two measurements remain constant

with depth. Measurements of ψ derived from SKS splitting intensity and P -wave delays provide an accurate view of anisotropy within the upper ~ 200 km. Here the azimuthal component of the toroidal flow pattern is relatively constant and westward of the trench frozen-in anisotropy and that due to entrainment are east-west oriented. Splitting intensity gives a more accurate characterization of the subsurface relative to the P delays likely because splitting intensity is an absolute rather than a relative measurement and less sensitive to isotropic structure. Below ~ 200 km the toroidal flow pattern evolves and there is a 90° change in ψ beneath the incoming plate (Fig. 2) causing a step-like increase in the angular misfit.

The weighted mean errors in ψ as a function of depth for the ABC and AB tomographic solutions are also shown in Fig. 16. The ABC solution provides an accurate view of ψ below ~ 100 km depth with an average error of 10° – 20° . Average errors in the AB model are notably larger, between 20° and 35° . The errors in ψ for both the ABC and AB solutions increase toward the surface where ray paths become more vertical and, consequently, less sensitive to anisotropic structure. Despite the fact that the SKS splitting intensity vectors

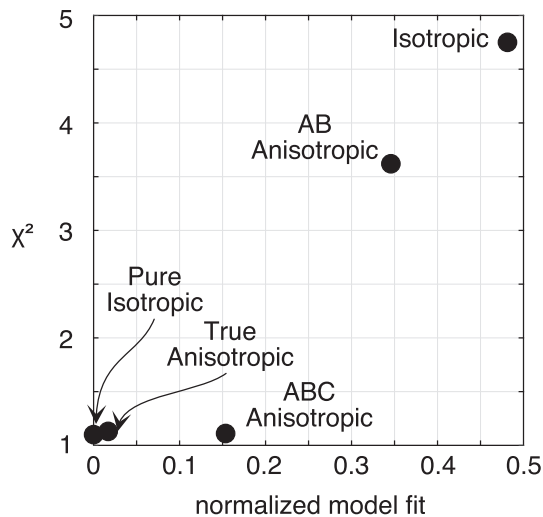


Figure 14. Comparison of isotropic solutions. The squared data residual norm (normalized by the data uncertainty) is plotted against the normalized fit to the isotropic model component for various inversions. The model fit is defined as, $|\text{dln}V - \text{dln}V'|^2 / |\text{dln}V_0 - \text{dln}V'|^2 - 1$, where $\text{dln}V$ and $\text{dln}V'$ are the recovered true fractional velocity perturbations and $\text{dln}V_0$ is a reference solution which we chose as the ‘Pure Isotropic’ model. The ‘Pure Isotropic’ result refers to the solution in Fig. 6; the ‘True Anisotropic’ result refers to the model obtained when the true anisotropic structure is included in the starting model (Fig. S11).

provide a depth integrated view of mantle anisotropy, they yield the most accurate view of ψ in the upper 200 km of the model with an average error of $\sim 10^\circ$. Of course, these measurements alone offer limited information on fabric dip and 3-D variations in anisotropic heterogeneity which can be illuminated by the P -wave inversions.

6 DISCUSSION AND CONCLUSIONS

Our synthetic tests demonstrate that teleseismic P waves alone can constrain 3-D upper mantle anisotropic structure. Moreover, constraining the dip in addition to the azimuth of anisotropic fabrics is critical to accurately imaging both isotropic and anisotropic heterogeneity in subduction zones. Purely isotropic or azimuthal approximations result in substantial low-velocity artefacts ($\text{dln}V > -1.5$ per cent) particularly behind the slab. Such anomalies could be interpreted as thermal anomalies of $\sim 120^\circ\text{K}$ or attributed to ~ 1 per cent partial melt (Bezada *et al.* 2016). The azimuthal simplification tends to generate spurious anisotropic structure flanking areas where the true fabric is inclined from horizontal. When the dip of anisotropic structure is included in our inversions, such artefacts are nearly eliminated. Notably, the concentration of low-velocity artefacts behind the slab is strikingly similar to those imaged in a number of real subduction settings (e.g. Li & van der Hilst 2010; Hawley *et al.* 2016; Portner *et al.* 2017; Bodmer *et al.* 2018; Ma *et al.* 2019). While such structures are often interpreted in a thermal context, this study and prior work by Bezada *et al.* (2016) stress that an equally likely cause may be unmodelled anisotropic structure. As demonstrated by Bezada *et al.* (2016), we can expect anisotropy-induced artefacts to be commonplace considering that their occurrence is largely independent of a given event distribution. However, the detailed geometry and amplitude of these features will depend on the particular experiment geometry and true mantle structure. The role of anisotropy on teleseismic tomography models

should be carefully assessed to yield more robust interpretations. There now exists a number of frameworks for this purpose. These include the inspection of azimuthal variations in residuals (Babuška & Plomerová 1992; Bokelmann 2002; Plomerová *et al.* 2012; Eken *et al.* 2012), forward modelling of anticipated anisotropic structure (Hammond & Toomey 2003; Bezada *et al.* 2016; Confal *et al.* 2020) and inversion for anisotropic parameters presented here and elsewhere (e.g. Zhao *et al.* 2016; Munzarová *et al.* 2018a; Beller & Chevrot 2020).

Well-constrained azimuthal anisotropy in a tomographic solution does not necessarily imply that isotropic heterogeneity is more robustly determined. While azimuthal inversions can improve model fit and image some true anisotropic features (e.g. general toroidal flow pattern), they are prone to artefacts and they do not substantially reduce anomalous LVZs. These artefacts are largely caused by changes in symmetry axis dip and are not unique to teleseismic imaging but may also exist in local imaging studies where absolute arrival times are known. This may explain why attempts to correct images based on shear wave splitting observations generally yield only modest changes in anomaly amplitudes (Lloyd & van der Lee 2008; O’Driscoll *et al.* 2011; Mohanty *et al.* 2016). Corrections based on dipping anisotropic domains (e.g. Eken *et al.* 2012) can be more significant than their azimuthal counterparts but unmodelled depth-dependent structure may still cause spurious anomalies. Bezada *et al.* (2016) found that attempts to estimate azimuthal anisotropic structure from SKS splits and incorporate this information into the starting tomographic model tended to exacerbate many imaging artefacts. These results all highlight that accounting for dipping anisotropic fabrics is crucial to obtaining accurate tomographic models.

While our synthetic tests demonstrate the utility of teleseismic P waves in constraining realistic upper mantle anisotropy, the proposed imaging strategy has important limitations. Below, we discuss some of these limitations and possible means by which they may be overcome.

As with most seismic imaging problems, the resulting tomographic models are non-unique—a problem that is exacerbated by introducing additional parameters required to describe anisotropy. This non-uniqueness is further complicated by the relative nature of teleseismic data recorded by a local or regional array (see Section 4). Therefore, it is important that anisotropic models are validated against independent data sets such as shear wave splitting observations. Even more beneficial to combating this non-uniqueness would be to devise inversion strategies that incorporate observables from other seismic phases. For example, absolute traveltimes from local earthquakes can be jointly inverted with teleseismic phases (e.g. Wang & Zhao 2013) to further constrain anisotropy in the upper 100–200 km. Imaging deeper anisotropy could also be improved by including splitting intensity observations in the inversion of P delays. While there are applications of 3-D splitting intensity tomography using exceptionally dense seismic arrays (e.g. Monteiller & Chevrot 2011; Lin *et al.* 2014; Mondal & Long 2020), the method has been largely limited by the near vertical incidence of the SKS phases typically used in such analyses (e.g. Chevrot & van der Hilst 2003; Chevrot 2006). The more diverse ray geometries of teleseismic P phases combined with their sensitivity to both the azimuth and dip of anisotropy would provide important and complimentary information. Ultimately, to fully characterize upper mantle anisotropy will require FWI of both compressional and shear waves (Beller & Chevrot 2020). However, such analyses remain limited by computational resources and availability of sufficiently dense seismic arrays. Considering the limited application of

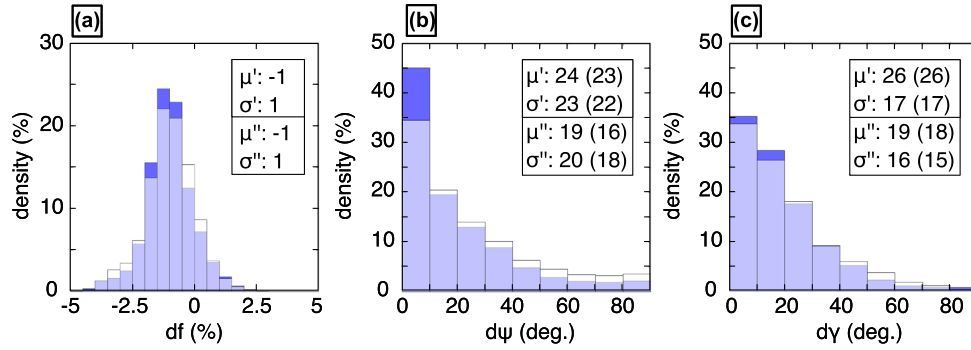


Figure 15. Errors in recovered anisotropic parameters. Distributions of the errors in the recovered anisotropic (a) fraction, (b) azimuth and (c) elevation are shown for the ABC (blue bars) and AB (transparent white bars) anisotropic inversions. The mean (μ) and standard deviation (σ) are provided for each distribution. Single primes refer to the AB solution while double primes refer to the ABC solution. Values in parentheses are anisotropic magnitude-weighted estimates of the distribution statistics.

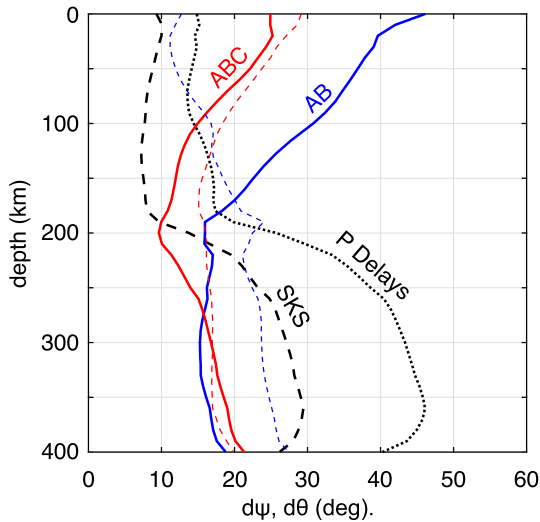


Figure 16. Depth profiles of the angular misfit in symmetry axis orientation. The weighted mean error in ψ is plotted for the ABC (red curve) and AB (blue curve) solutions at each depth in the tomographic model. Errors in ψ are also shown for SKS splitting intensity (dashed black curve; Fig. S12) and *P*-wave delays (dotted black curve; Fig. 9a) assuming vertically invariant structure. Thin red and blue dashed curves define the weighted mean angular misfit in γ for the ABC and AB models, respectively.

anisotropic imaging in teleseismic body wave tomography, we believe even the simplified treatment proposed here has the potential to yield new and geodynamically relevant insights into upper mantle structure.

While the tomographic solutions are non-unique, teleseismic anisotropic imaging methods are robust in the sense that truly isotropic structure tends to not generate erroneous anisotropic heterogeneity. Conversely, unaccounted for anisotropic structure does create significant isotropic artefacts. These points are clear from our own synthetic tests (Figs S6–S8) as well as number of other studies investigating the role of anisotropy in regional teleseismic imaging (Huang *et al.* 2015; Bezada *et al.* 2016; Munzarová *et al.* 2018a; Beller & Chevrot 2020). Thus, the risk of mapping relevant isotropic structure into anisotropic parameters is low and tomographic solutions that include anisotropy should generally yield more accurate representations of isotropic structure. Therefore, it seems appropriate to recommend the more broad use of anisotropic imaging methods especially in locations where anisotropic heterogeneity is expected to be significant.

Our linearized inversion scheme is not well-suited for characterizing uncertainty in the obtained models. Considering this limitation and the non-linear relationship between anisotropic parameters and the *P*-wave delays, non-linear inverse methods capable of systematically exploring the model space (e.g. Bodin & Sambridge 2009) are ideal. Such approaches have proved successful in constraining arbitrarily oriented hexagonal anisotropy (e.g. Mondal & Long 2019). However, these probabilistic methods are limited to a relatively small number of free parameters. In this regard, the simplified description of *P*-wave anisotropy considered here is beneficial and model space searches could help elucidate a smaller number of discrete anisotropic domains with robust error estimates.

Our synthetic tests demonstrate that the maximum depth extent of anisotropy is poorly constrained. Recent anisotropic tomography using teleseismic body waves in the Western Pacific (Ma *et al.* 2019) recovered ~ 2 per cent *P*-wave anisotropy throughout the upper 1000 km of the mantle. Such deep (> 500 km) anisotropy is surprising considering other seismic observables (Kustowski *et al.* 2008; Lebedev & van der Hilst 2008), geodynamic predictions (Sturgeon *et al.* 2019), and mineral physics experiments (Karato *et al.* 2008) that generally support decreasing anisotropic strength with depth. Our synthetic tests highlight the tendency of teleseismic *P* waves to distribute anisotropic perturbations along their high-angle paths (Fig. S13). We suggest that squeezing tests, where the depth interval over which anisotropic perturbations are allowed is systematically varied, should be performed to identify the anisotropic region required by the data.

The primary goal of this study was to evaluate the ability of relative teleseismic *P*-wave delays to constrain realistic upper mantle anisotropy patterns. We have shown that this is possible with an ideal backazimuthal event distribution despite limited sampling of ray incidence angles. However, we did not explore the effects of uneven data geometries. On the basis of checkerboard tests, events spanning a 60° backazimuthal range appear sufficient to capture at least the anisotropic azimuth (Huang *et al.* 2015). Such angular coverage can be achieved by many arrays particularly those operated over extended periods of time (e.g. USArray). Heavily biased sampling of certain directions and non-uniform event recording across an array may also adversely effect the tomographic model. Such biases may be overcome via ray averaging schemes that act to homogenize data coverage (e.g. Babuška *et al.* 1984; Ruan *et al.* 2019). Ongoing research is aimed at applying our imaging strategy to the Western United States, specifically targeting the Cascadia subduction system, in which we will fully investigate these potential issues.

We conclude by emphasizing that all seismic waveforms (SAC format), delay and splitting measurements, and elastic models generated during this study are available via the online repository figshare (see Supplementary Information and references VanderBeek & Faccenda 2020a–f for dataset DOIs). We hope that other researchers will find this material useful in developing and evaluating their own imaging methods. We find that the mantle imaging community in general lacks a comprehensive synthetic model that may be used for benchmarking imaging algorithms. In comparison, the seismic exploration community has extensively used the Marmousi sedimentary basin model (Versteeg 1994; Martin *et al.* 2006) for this purpose. Such reference data is critical to facilitate informed comparisons of seismic models generated via different techniques and to assess errors and biases in a particular method. While checkerboard and other block-like structures commonly used in synthetic tests are easy to reproduce, they are poor analogues for realistic earth structure. Additionally, synthetic tests often rely on ideal self-consistent data (i.e. data that conforms to assumptions built into a particular method). While we do not claim that our synthetic data is the perfect benchmark, we hope that it serves as a start to this larger goal and that other researchers will find it useful for evaluating tomographic methods and models.

ACKNOWLEDGEMENTS

This research was performed as part of the NEWTON project and supported by the European Research Council (grant 758199). Waveform modelling was performed using computational resources on the GALILEO supercomputer maintained by the CINECA HPC facility. Synthetic models and seismic data generated as a result of this study are available via the online repository figshare (VanderBeek & Faccenda 2020a–f). Figures were created using Generic Mapping Tools (Wessel *et al.* 2019) with perceptually uniform colour maps developed by Crameri (2018a, b).

REFERENCES

Aki, K., Christofferson, A. & Husebye, E.S., 1977. Determination of the three-dimensional seismic structure of the lithosphere, *J. geophys. Res.*, **82**(2), 277–296.

Aster, R., Borchers, B. & Thurber, C., 2005. *Parameter Estimation and Inverse Problems*, Academic Press.

Babuška, V. & Plomerová, J., 1992. The lithosphere in central Europe—seismological and petrological aspects, *Tectonophysics*, **207**(1–2), 141–163.

Babuška, V. & Plomerová, J., 2006. European mantle lithosphere assembled from rigid microplates with inherited seismic anisotropy, *Phys. Earth planet. Inter.*, **158**(2–4), 264–280.

Babuška, V., Plomerová, J. & Šílený, J., 1984. Spatial variations of P residuals and deep structure of the European lithosphere, *J. geophys. Int.*, **79**(1), 363–383.

Babuška, V., Plomerová, J. & Šílený, J., 1993. Models of seismic anisotropy in the deep continental lithosphere, *Phys. Earth planet. Inter.*, **78**(3–4), 167–191.

Backus, G.E., 1965. Possible forms of seismic anisotropy of the uppermost mantle under oceans, *J. geophys. Res.*, **70**(14), 3429–3439.

Becker, T.W. & Boschi, L., 2002. A comparison of tomographic and geodynamic mantle models, *Geochem. Geophys. Geosyst.*, **3**(1), doi:10.1029/2001GC000168.

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, *J. geophys. Int.*, **155**(2), 696–714.

Becker, T.W., Chevrot, S., Schulte-Pelkum, V. & Blackman, D. K., 2006. Statistical properties of seismic anisotropy predicted by upper mantle geodynamic models, *J. geophys. Res.*, **111**(B8), doi:10.1029/2005JB004095.

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**, 236–250.

Beller, S. & Chevrot, S., 2020. Probing depth and lateral variations of upper-mantle seismic anisotropy from full-waveform inversion of teleseismic body-waves, *J. geophys. Int.*, **222**(1), 352–387.

Bezada, M.J., Levander, A. & Schmandt, B., 2010. Subduction in the southern Caribbean: images from finite-frequency P wave tomography, *J. geophys. Res.*, **115**(B12), doi:10.1029/2010JB007682.

Bezada, M.J., Faccenda, M. & Toomey, D.R., 2016. Representing anisotropic subduction zones with isotropic velocity models: a characterization of the problem and some steps on a possible path forward, *Geochem. Geophys. Geosyst.*, **17**(8), 3164–3189.

Blackman, D.K. & Kendall, J.M., 1997. Sensitivity of teleseismic body waves to mineral texture and melt in the mantle beneath a mid-ocean ridge, *Phil. Trans. R. Soc. Lond., A.*, **355**(1723), 217–231.

Blackman, D.K. & Kendall, J.M., 2002. Seismic anisotropy in the upper mantle 2. Predictions for current plate boundary flow models, *Geochem. Geophys. Geosyst.*, **3**(9), doi:10.1029/2001GC000247.

Blackman, D.K., Kendall, J.M., Dawson, P.R., Wenk, H.R., Boyce, D. & Morgan, J.P., 1996. Teleseismic imaging of subaxial flow at mid-ocean ridges: traveltimes effects of anisotropic mineral texture in the mantle, *J. geophys. Int.*, **127**(2), 415–426.

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

Bodmer, M., Toomey, D.R., Hooft, E.E. & Schmandt, B., 2018. Buoyant asthenosphere beneath Cascadia influences megathrust segmentation, *Geophys. Res. Lett.*, **45**(14), 6954–6962.

Bokelmann, G.H., 2002. Convection-driven motion of the North American craton: evidence from P-wave anisotropy, *J. geophys. Int.*, **148**(2), 278–287.

Browaey, J.T. & Chevrot, S., 2004. Decomposition of the elastic tensor and geophysical applications, *J. geophys. Int.*, **159**(2), 667–678.

Byrnes, J.S., Toomey, D.R., Hooft, E.E., Nábělek, J. & Braunmiller, J., 2017. Mantle dynamics beneath the discrete and diffuse plate boundaries of the Juan de Fuca plate: results from Cascadia initiative body wave tomography, *Geochem. Geophys. Geosyst.*, **18**(8), 2906–2929.

Calvet, M., Chevrot, S. & Souriau, A., 2006. P-wave propagation in transversely isotropic media: I. Finite-frequency theory, *Phys. Earth planet. Inter.*, **156**(1–2), 12–20.

Cammarano, F., Goes, S., Vacher, P. & Giardini, D., 2003. Inferring upper-mantle temperatures from seismic velocities, *Phys. Earth planet. Inter.*, **138**(3–4), 197–222.

Chen, M. & Tromp, J., 2007. Theoretical and numerical investigations of global and regional seismic wave propagation in weakly anisotropic earth models, *J. geophys. Int.*, **168**(3), 1130–1152.

Chevrot, S., 2000. Multichannel analysis of shear wave splitting, *J. geophys. Res.*, **105**(B9), 21 579–21 590.

Chevrot, S., 2006. Finite-frequency vectorial tomography: a new method for high-resolution imaging of upper mantle anisotropy, *J. geophys. Int.*, **165**(2), 641–657.

Chevrot, S. & van der Hilst, R.D., 2003. On the effects of a dipping axis of symmetry on shear wave splitting measurements in a transversely isotropic medium, *J. geophys. Int.*, **152**(2), 497–505.

Confal, J.M., Bezada, M.J., Eken, T., Faccenda, M., Saygin, E. & Taymaz, T., 2020. Influence of upper mantle anisotropy on isotropic P-wave tomography images obtained in the Eastern Mediterranean Region, *J. geophys. Res.*, **125**, doi:10.1029/2019JB018559.

Conrad, C.P., Behn, M.D. & Silver, P.G., 2007. Global mantle flow and the development of seismic anisotropy: differences between the oceanic and continental upper mantle, *J. geophys. Res.*, **112**(B7), doi:10.1029/2006JB004608.

Crameri, F., 2018a. Scientific colour-maps, Zenodo. <http://doi.org/10.5281/zenodo.1243862>.

- Cramer, F., 2018b. Geodynamic diagnostics, scientific visualisation and StagLab 3.0, *Geosci. Model Dev.*, **11**, 2541–2562.
- Crotwell, H.P., Owens, T.J. & Ritsema, J., 1999. The TauP Toolkit: flexible seismic traveltimes and ray-path utilities, *Seismol. Res. Lett.*, **70**(2), 154–160.
- Dahlen, F.A., Hung, S.H. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes—I. Theory, *J. geophys. Int.*, **141**(1), 157–174.
- Dunn, R.A., Lekić, V., Detrick, R.S. & Toomey, D.R., 2005. Three-dimensional seismic structure of the Mid-Atlantic Ridge (35 N): evidence for focused melt supply and lower crustal dike injection, *J. geophys. Res.*, **110**(B9), doi:10.1029/2004JB003473.
- Eberhart-Phillips, D. & Mark Henderson, C., 2004. Including anisotropy in 3-D velocity inversion and application to Marlborough, New Zealand, *Geophys. J. Int.*, **156**(2), 237–254.
- Eilon, Z.C. & Abers, G.A., 2017. High seismic attenuation at a mid-ocean ridge reveals the distribution of deep melt, *Sci. Adv.*, **3**(5), doi:10.1126/sciadv.1602829 Article.
- Eken, T., Plomerová, J., Vecsey, L., Babuška, V., Roberts, R., Shomali, H. & Bodvarsson, R., 2012. Effects of seismic anisotropy on P-velocity tomography of the Baltic Shield, *J. geophys. Int.*, **188**(2), 600–612.
- Faccenda, M., 2014. Mid mantle seismic anisotropy around subduction zones, *Phys. Earth planet. Inter.*, **227**, 1–19.
- Faccenda, M. & Capitanio, F.A., 2013. Seismic anisotropy around subduction zones: Insights from three-dimensional modeling of upper mantle deformation and SKS splitting calculations, *Geochem. Geophys. Geosyst.*, **14**(1), 243–262.
- Faccenna, C. & Becker, T.W., 2010. Shaping mobile belts by small-scale convection, *Nature*, **465**(7298), 602–605.
- Fisher, N.I., 1995. *Statistical Analysis of Circular Data*, Cambridge Univ. Press.
- Goes, S. & van der Lee, S., 2002. Thermal structure of the North American uppermost mantle inferred from seismic tomography, *J. geophys. Res.*, **107**(B3), ETG–2.
- Grésillaud, A. & Cara, M., 1996. Anisotropy and P-wave tomography: a new approach for inverting teleseismic data from a dense array of stations, *J. geophys. Int.*, **126**(1), 77–91.
- Hammond, W.C. & Humphreys, E.D., 2000. Upper mantle seismic wave velocity: effects of realistic partial melt geometries, *J. geophys. Res.*, **105**(B5), 10 975–10 986.
- Hammond, W.C. & Toomey, D.R., 2003. Seismic velocity anisotropy and heterogeneity beneath the Mantle Electromagnetic and Tomography Experiment (MELT) region of the East Pacific Rise from analysis of P and S body waves, *J. geophys. Res.*, **108**(B4), doi:10.1029/2002JB001789.
- Hawley, W.B., Allen, R.M. & Richards, M.A., 2016. Tomography reveals buoyant asthenosphere accumulating beneath the Juan de Fuca plate, *Science*, **353**(6306), 1406–1408.
- Hearn, T.M., 1996. Anisotropic Pn tomography in the western United States, *J. geophys. Res.*, **101**(B4), 8403–8414.
- Huang, Z., Zhao, D. & Liu, X., 2015. On the trade-off between seismic anisotropy and heterogeneity: numerical simulations and application to Northeast Japan, *J. geophys. Res.*, **120**(5), 3255–3277.
- Hung, S.H., Dahlen, F.A. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes - II. Examples, *J. geophys. Int.*, **141**(1), 175–203.
- Ito, G., Dunn, R., Li, A., Wolfe, C.J., Gallego, A. & Fu, Y., 2014. Seismic anisotropy and shear wave splitting associated with mantle plume-plate interaction, *J. geophys. Res.*, **119**(6), 4923–4937.
- Jackson, I. & Faul, U.H., 2010. Grain-size-sensitive viscoelastic relaxation in olivine: towards a robust laboratory-based model for seismological application, *Phys. Earth planet. Inter.*, **183**(1–2), 151–163.
- Kaminski, E., Ribe, N.M. & Browaeys, J.T., 2004. D-Rex, a program for calculation of seismic anisotropy due to crystal lattice preferred orientation in the convective upper mantle, *J. geophys. Int.*, **158**(2), 744–752.
- Karato, S.I. & Jung, H., 1998. Water, partial melting and the origin of the seismic low velocity and high attenuation zone in the upper mantle, *Earth planet. Sci. Lett.*, **157**(3–4), 193–207.
- Karato, S.I., Jung, H., Katayama, I. & Skemer, P., 2008. Geodynamic significance of seismic anisotropy of the upper mantle: new insights from laboratory studies, *Annu. Rev. Earth Planet. Sci.*, **36**, 59–95.
- Kendall, J.M., 1994. Teleseismic arrivals at a mid-ocean ridge: effects of mantle melt and anisotropy, *Geophys. Res. Lett.*, **21**(4), 301–304.
- Kennett, B.L.N. & Engdahl, E.R., 1991. Traveltimes for global earthquake location and phase identification, *J. geophys. Int.*, **105**(2), 429–465.
- Kissling, E., 1988. Geotomography with local earthquake data, *Rev. Geophys.*, **26**(4), 659–698.
- Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophys. J. Int.*, **139**(3), 806–822.
- Kustowski, B., Ekström, G. & Dziewoński, A.M., 2008. Anisotropic shear-wave velocity structure of the Earth's mantle: a global model, *J. geophys. Res.*, **113**(B6), doi:10.1029/2007JB005169.
- Koulakov, I., Jakovlev, A. & Luehr, B.G., 2009. Anisotropic structure beneath central Java from local earthquake tomography, *Geochem. Geophys. Geosyst.*, **10**(2), doi:10.1029/2008GC002109.
- Lebedev, S. & van der Hilst, R.D., 2008. Global upper-mantle tomography with the automated multimode inversion of surface and S-wave forms, *J. geophys. Int.*, **173**(2), 505–518.
- Lebedev, S., Schaeffer, A.J., Fullaer, J. & Pease, V., 2018. Seismic tomography of the Arctic region: inferences for the thermal structure and evolution of the lithosphere, *Geol. Soc., Lond., Spec. Publ.*, **460**(1), 419–440.
- Lévesque, J.J. & Masson, F., 1999. From ACH tomographic models to absolute velocity models, *J. geophys. Int.*, **137**(3), 621–629.
- Li, C. & van der Hilst, R.D., 2010. Structure of the upper mantle and transition zone beneath Southeast Asia from traveltimes tomography, *J. geophys. Res.*, **115**(B7), doi:10.1029/2009JB006882.
- Lin, Y.P., Zhao, L. & Hung, S.H., 2014. Full-wave multiscale anisotropy tomography in Southern California, *Geophys. Res. Lett.*, **41**(24), 8809–8817.
- Liu, X. & Zhao, D., 2017. P-wave anisotropy, mantle wedge flow and olivine fabrics beneath Japan, *J. geophys. Int.*, **210**(3), 1410–1431.
- Lloyd, S.M. & van der Lee, S., 2008. Influence of observed mantle anisotropy on isotropic tomographic models, *Geochem. Geophys. Geosyst.*, **9**(7), doi:10.1029/2019JB018559.
- Long, M.D. & Becker, T.W., 2010. Mantle dynamics and seismic anisotropy, *Earth planet. Sci. Lett.*, **297**(3–4), 341–354.
- Lou, X., van der Lee, S. & Lloyd, S., 2013. AIMBAT: a python/matplotlib tool for measuring teleseismic arrival times, *Seismol. Res. Lett.*, **84**(1), 85–93.
- Ma, J., Tian, Y., Zhao, D., Liu, C. & Liu, T., 2019. Mantle dynamics of western Pacific and East Asia: new insights from P wave anisotropic tomography, *Geochem. Geophys. Geosyst.*, **20**(7), 3628–3658.
- Mainprice, D., 2007. Seismic anisotropy of the deep Earth from a mineral and rock physics perspective, *Treatise Geophys.*, **2**, 437–491.
- Marquering, H., Dahlen, F.A. & Nolet, G., 1999. Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox, *J. geophys. Int.*, **137**(3), 805–815.
- Martin, G.S., Wiley, R. & Marfurt, K.J., 2006. Marmousi2: an elastic upgrade for Marmousi, *Leading Edge*, **25**(2), 156–166.
- Masson, Y. & Romanowicz, B., 2017. Box tomography: localized imaging of remote targets buried in an unknown medium, a step forward for understanding key structures in the deep Earth, *J. geophys. Int.*, **211**(1), 141–163.
- Maupin, V. & Kolstrup, M.L., 2015. Insights in P- and S-wave relative traveltimes tomography from analysing finite-frequency Fréchet kernels, *J. geophys. Int.*, **202**(3), 1581–1598.
- Menke, W., 2015. Equivalent heterogeneity analysis as a tool for understanding the resolving power of anisotropic travel-time tomography, *Bull. seism. Soc. Am.*, **105**(2A), 719–733.
- Mohanty, D.D., Singh, A., O'Driscoll, L.J., Ravi Kumar, M., Srinagesh, D. & Humphreys, E.D., 2016. P wave velocity structure below India and Tibet incorporating anisotropic delay time effects, *Geochem. Geophys. Geosyst.*, **17**(3), 725–738.
- Mondal, P. & Long, M.D., 2019. A model space search approach to finite-frequency SKS splitting intensity tomography in a reduced parameter space, *J. geophys. Int.*, **217**(1), 238–256.

- Mondal, P. & Long, M.D., 2020. Strong seismic anisotropy in the deep upper mantle beneath the Cascadia backarc: constraints from probabilistic finite-frequency SKS splitting intensity tomography, *Earth planet. Sci. Lett.*, **539**, 116172.
- Monteiller, V. & Chevrot, S., 2011. High-resolution imaging of the deep anisotropic structure of the San Andreas Fault system beneath southern California, *J. geophys. Int.*, **186**(2), 418–446.
- Monteiller, V., Chevrot, S., Komatitsch, D. & Fuji, N., 2013. A hybrid method to compute short-period synthetic seismograms of teleseismic body waves in a 3-D regional model, *J. geophys. Int.*, **192**(1), 230–247.
- Munzarová, H., Plomerová, J. & Kissling, E., 2018a. Novel anisotropic teleseismic body-wave tomography code AniTomo to illuminate heterogeneous anisotropic upper mantle: Part I—theory and inversion tuning with realistic synthetic data, *J. geophys. Int.*, **215**(1), 524–545.
- Munzarová, H., Plomerová, J., Kissling, E., Vecsey, J. & Babuška, V., 2018b. Novel anisotropic teleseismic body-wave tomography code AniTomo to illuminate heterogeneous anisotropic upper mantle: Part II – application to data of passive seismic experiment LAPNET in northern Fennoscandia, *J. geophys. Int.*, **215**(2), 1388–1409.
- Nissen-Meyer, T., Driel, M.V., Stähler, S., Hosseini, K., Hempel, S., Auer, L., Colombi, A. & Fournier, A., 2014. AxisSEM: broadband 3-D seismic wavefields in axisymmetric media, *Solid Earth*, **5**(1), 425–445.
- O'Driscoll, L.J., Humphreys, E.D. & Schmandt, B., 2011. Time corrections to teleseismic P delays derived from SKS splitting parameters and implications for western US P-wave tomography, *Geophys. Res. Lett.*, **38**(19), doi:10.1029/2011GL049031.
- Paige, C.C. & Saunders, M.A., 1982. LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Softw. (TOMS)*, **8**(1), 43–71.
- Plank, T. & Forsyth, D.W., 2016. Thermal structure and melting conditions in the mantle beneath the Basin and Range province from seismology and petrology, *Geochem. Geophys. Geosyst.*, **17**(4), 1312–1338.
- Plomerová, J., Šílený, J. & Babuška, V., 1996. Joint interpretation of upper-mantle anisotropy based on teleseismic P-travel time delays and inversion of shear-wave splitting parameters, *Phys. Earth planet. Inter.*, **95**(3–4), 293–309.
- Plomerová, J., Vecsey, L. & Babuška, V., 2012. Mapping seismic anisotropy of the lithospheric mantle beneath the northern and eastern Bohemian Massif (central Europe), *Tectonophysics*, **564**, 38–53.
- Portner, D.E., Beck, S., Zandt, G. & Scire, A., 2017. The nature of slab slow velocity anomalies beneath South America, *Geophys. Res. Lett.*, **44**(10), 4747–4755.
- Portner, D.E. *et al.*, 2020. Detailed structure of the subducted Nazca slab into the lower mantle derived from continent-scale teleseismic P wave tomography, *J. geophys. Res.*, **125**(5), doi:10.1029/2019JB017884.
- Rawlinson, N. & Spakman, W., 2016. On the use of sensitivity tests in seismic tomography, *J. geophys. Int.*, **205**(2), 1221–1243.
- Ribe, N.M., 1989. Seismic anisotropy and mantle flow, *J. geophys. Res.*, **94**(B4), 4213–4223.
- Ruan, Y., Forsyth, D.W. & Bell, S.W., 2018. Shear attenuation beneath the Juan de Fuca plate: implications for mantle flow and dehydration, *Earth planet. Sci. Lett.*, **496**, 189–197.
- Ruan, Y., Lei, W., Modrak, R., Örsveran, R., Bozdağ, E. & Tromp, J., 2019. Balancing unevenly distributed data in seismic tomography: a global adjoint tomography example, *J. geophys. Int.*, **219**(2), 1225–1236.
- Savage, M.K., 1999. Seismic anisotropy and mantle deformation: what have we learned from shear wave splitting?, *Rev. Geophys.*, **37**(1), 65–106.
- Schmandt, B. & Humphreys, E., 2010. Seismic heterogeneity and small-scale convection in the southern California upper mantle, *Geochem. Geophys. Geosyst.*, **11**(5), doi:10.1029/2010GC003042.
- Schulte-Pelkum, V. & Blackman, D.K., 2003. A synthesis of seismic P and S anisotropy, *J. geophys. Int.*, **154**(1), 166–178.
- Sieminski, A., Liu, Q., Trampert, J. & Tromp, J., 2007. Finite-frequency sensitivity of body waves to anisotropy based upon adjoint methods, *J. geophys. Int.*, **171**(1), 368–389.
- Sieminski, A., Trampert, J. & Tromp, J., 2009. Principal component analysis of anisotropic finite-frequency sensitivity kernels, *J. geophys. Int.*, **179**(2), 1186–1198.
- Simmons, N.A., Forte, A.M. & Grand, S.P., 2006. Constraining mantle flow with seismic and geodynamic data: a joint approach, *Earth planet. Sci. Lett.*, **246**(1–2), 109–124.
- Šílený, J. & Plomerová, J., 1996. Inversion of shear-wave splitting parameters to retrieve three-dimensional orientation of anisotropy in continental lithosphere, *Phys. Earth planet. Inter.*, **95**(3–4), 277–292.
- Skemer, P. & Hansen, L.N., 2016. Inferring upper-mantle flow from seismic anisotropy: an experimental perspective, *Tectonophysics*, **668**, 1–14.
- Sobolev, S.V., Grésillaud, A. & Cara, M., 1999. How robust is isotropic delay time tomography for anisotropic mantle?, *Geophys. Res. Lett.*, **26**(4), 509–512.
- Sturgeon, W., Ferreira, A.M., Faccenda, M., Chang, S.J. & Scharndong, L., 2019. On the origin of radial anisotropy near subducted slabs in the midmantle, *Geochem. Geophys. Geosyst.*, **20**(11), 5105–5125.
- Takei, Y., 2002. Effect of pore geometry on Vp/Vs: from equilibrium geometry to crack, *J. geophys. Res.*, **107**(B2), ECV–6.
- Thomsen, L., 1986. Weak elastic anisotropy, *Geophysics*, **51**(10), 1954–1966.
- Toomey, D.R. & Foulger, G.R., 1989. Tomographic inversion of local earthquake data from the Hengill-Grensdalur central volcano complex, Iceland, *J. geophys. Res.*, **94**(B12), 17497–17510.
- VanderBeek, B. & Faccenda, M., 2020a. Dataset S1. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.12745151.v1>.
- VanderBeek, B. & Faccenda, M., 2020b. Dataset S2. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.12745169.v1>.
- VanderBeek, B. & Faccenda, M., 2020c. Dataset S3. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.12745178.v1>.
- VanderBeek, B. & Faccenda, M., 2020d. Dataset S4. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.12745229.v1>.
- VanderBeek, B. & Faccenda, M., 2020e. Dataset S5. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.12745265.v1>.
- VanderBeek, B. & Faccenda, M., 2020f. Dataset S6. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.12745370.v1>.
- VanDecar, J.C. & Crosson, R.S., 1990. Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. seism. Soc. Am.*, **80**(1), 150–169.
- Versteeg, R., 1994. The Marmousi experience: Velocity model determination on a synthetic complex data set, *Leading Edge*, **13**(9), 927–936.
- Villagómez, D.R., Toomey, D.R., Geist, D.J., Hooft, E.E. & Solomon, S.C., 2014. Mantle flow and multistage melting beneath the Galápagos hotspot revealed by seismic imaging, *Nature Geoscience*, **7**(2), 151–156.
- Vinnik, L.P., Makeyeva, L.I., Milev, A. & Usenko, A.Y., 1992. Global patterns of azimuthal anisotropy and deformations in the continental mantle, *J. geophys. Int.*, **111**(3), 433–447.
- Wang, W. & Becker, T.W., 2019. Upper mantle seismic anisotropy as a constraint for mantle flow and continental dynamics of the North American plate, *Earth planet. Sci. Lett.*, **514**, 143–155.
- Wang, J. & Zhao, D., 2008. P-wave anisotropic tomography beneath Northeast Japan, *Phys. Earth planet. Inter.*, **170**(1–2), 115–133.
- Wang, J. & Zhao, D., 2013. P-wave tomography for 3-D radial and azimuthal anisotropy of Tohoku and Kyushu subduction zones, *J. geophys. Int.*, **193**(3), 1166–1181.
- Weekly, R.T., Wilcock, W.S., Toomey, D.R., Hooft, E.E. & Kim, E., 2014. Upper crustal seismic structure of the Endeavour segment, Juan de Fuca Ridge from traveltome tomography: Implications for oceanic crustal accretion, *Geochem. Geophys. Geosyst.*, **15**(4), 1296–1315.
- Wessel, P., Luis, J.F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W.H.F. & Tian, D., 2019. The generic mapping tools version 6, *Geochem. Geophys. Geosyst.*, **20**(11), 5556–5564.
- Wiens, D.A., Conder, J.A. & Faul, U.H., 2008. The seismic structure and dynamics of the mantle wedge, *Annu. Rev. Earth Planet. Sci.*, **36**, 421–455.
- Zhang, H. & Thurber, C.H., 2007. Estimating the model resolution matrix for large seismic tomography problems based on Lanczos bidiagonalization with partial reorthogonalization, *J. geophys. Int.*, **170**(1), 337–345.
- Zhang, H., Thurber, C. & Bedrosian, P., 2009. Joint inversion for Vp, Vs, and Vp/Vs at SAFOD, Parkfield, California, *Geochem. Geophys. Geosyst.*, **10**(11), doi:10.1029/2009GC002709.

Zhao, L. & Jordan, T.H., 1998. Sensitivity of frequency-dependent travel-times to laterally heterogeneous, anisotropic Earth structure, *J. geophys. Int.*, **133**, 683–704.

Zhao, D., Yu, S. & Liu, X., 2016. Seismic anisotropy tomography: new insight into subduction dynamics, *Gondwana Res.*, **33**, 24–43.

SUPPORTING INFORMATION

Supplementary data are available at [GJI](https://doi.org/10.1093/gji/gjz001) online.

Figure S1. Illustration of coordinate system and angle convention. The red vector represents a ray segment unit vector (\mathbf{r}) or anisotropic symmetry axis vector (\mathbf{n}). The azimuth of the ray (ϕ) or symmetry axis vector (ψ) is measured counter-clockwise-positive from east. The elevation of the ray (θ) or symmetry axis vector (γ) is measured counter-clockwise-positive from the horizontal east–north plane.

Figure S2. Mean station delays through the isotropic subduction model (a) measured from simulated waveforms are compared to those predicted using (b) ray theory, (c) HFFKs and (d) Born sensitivity kernels.

Figure S3. Mean station delays through the anisotropic subduction model (a) measured from simulated waveforms are compared to those predicted using (b) ray theory and (c) HFFKs. Anisotropic delays are not estimated using the isotropic Born sensitivity kernels.

Figure S4. Distributions of delay times errors predicted using (a, d) ray theory, (b, e) HFFKs and (c) Born sensitivity kernels for the (a–c) isotropic and (d–e) anisotropic subduction zone models. The delay time errors are with respect to delays measured from the simulated waveforms. Note the long positive tail in the ray theory distributions resulting from overestimating the traveltime anomaly associated with the slab. Anisotropic delays are not estimated using the isotropic Born sensitivity kernels.

Figure S5. Maps of the (a) isotropic derivative weight sum (DWS) and (b) azimuthal mean resultant length (AMRL) are shown at 150 km depth. East–west cross-sections at 0°N through the (c) DWS and (d) AMRL fields are shown immediately below. Well-sampled nodes are considered those with DWS > 150 or AMRL < 0.5.

Figure S6. Recovery of a homogeneous isotropic high-velocity cylinder (radius is 150 km) characterized by a 4xd per cent velocity increase. True anomaly is outlined in black. Top row shows the tomographic solution at 250 km depth (mid-plane of anomaly). Bottom row shows SW-to-NE cross-sections (dashed line in map views) through the tomographic models. Each column corresponds to isotropic, AB-anisotropic and ABC-anisotropic solutions. The symmetry axis orientations recovered from the anisotropic solutions are projected onto the cross section planes and scaled by the anisotropic magnitude. Note that the peak-to-peak anisotropic magnitude is very weak (everywhere <0.5 per cent).

Figure S7. Same as Fig. S6 but for a low-velocity cylinder characterized by a 4 per cent velocity reduction. Again, anisotropic perturbations are weak (everywhere <0.5 per cent).

Figure S8. Isotropic artefacts from purely anisotropic structure. Each row corresponds to isotropic, AB-anisotropic and ABC-anisotropic inversions of delays predicted through a model

containing a single purely anisotropic cylindrical anomaly (outlined in black; radius is 150 km). The symmetry axis azimuth in each block is 60° and the peak-to-peak anisotropic magnitude is 5 per cent. Each column corresponds to solutions obtained by inverting delays predicted through models with different symmetry axis elevations (0°, 30° or 60°). The true symmetry axis orientation in each block is shown by the black quiver. Each cross-section passes through the centre of the anomaly and is oriented parallel to the symmetry axis azimuth where isotropic artefacts are most prevalent; see Fig. S6 for cross-section location.

Figure S9. Recovery of anisotropic cylinder models at 250 km depth (mid-plane of true anomaly). Each row corresponds to the AB- and ABC-anisotropic solutions. Each column shows the recovery of anisotropic blocks with different symmetry axis elevations. The true symmetry axis azimuth in each block is 60° and the peak-to-peak anisotropic magnitude is 5 per cent. The true anomaly is outlined in white and the white quiver in the bottom right corner shows the true symmetry axis orientation in the cross-section plane. Anisotropy quivers are projected onto cross-section plane and scaled by the anisotropic magnitude.

Figure S10. Cross-sections through the anisotropic models presented in Fig. S9. Cross-section location is shown in Fig. S6 and parallels the true symmetry axis azimuth. Isotropic structure from these inversions is shown in the bottom two rows of Fig. S8. The true anomaly is outlined in white and the white quiver in the bottom right corner shows the true symmetry axis orientation in the cross-section plane. Anisotropy quivers are projected onto cross-section plane and scaled by the anisotropic magnitude.

Figure S11. Isotropic inversion of anisotropic data with true anisotropic structure in starting model. Velocity perturbations are plotted as in Fig. 5.

Figure S12. Splitting intensity vectors measured from SKS waveforms propagated through the anisotropic subduction zone model. Quivers are scaled and color-coded by split time.

Figure S13. Anisotropic structure recovered from inversion for \bar{u} , A , B and C parameters without imposing a depth limit on anisotropic perturbations; compare to Fig. 12. Seismic anisotropy is plotted as in Fig. 2. Areas with poor directional sampling are masked.

Figure S14. Angular errors in the AB-anisotropic solution (Fig. 10). Difference between the recovered and true symmetry axis azimuth is shown at (a) 150 km and (b) 350 km depth. The difference between the recovered and true symmetry axis elevation is shown in east–west cross-sections at (c) 0°N and (d) 4°30'S. Note that the symmetry axis elevation in the AB solution is everywhere 0°.

Figure S15. Angular errors in the ABC-anisotropic solution (Fig. 12) are plotted as in Fig. S14 (depth sections correspond to ϕ and cross-section correspond to γ).

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.