Anisotropic P-wave traveltime tomography implementing Thomsen's weak approximation in TOMO3D

Adrià Meléndez[1], Clara Estela Jiménez[1], Valentí Sallarès[1], César R. Ranero[2]

[1]Barcelona Center for Subsurface Imaging, Institut de Ciències del Mar (CSIC), Barcelona, 08003, Spain
[2]Barcelona Center for Subsurface Imaging, ICREA at Institut de Ciències del Mar (CSIC), Barcelona, 08003, Spain

Correspondence to: Adrià Meléndez (melendez@icm.csic.es)

Abstract. We present the implementation of Thomsen's weak anisotropy approximation for VTI media within TOMO3D, our code for 2-D and 3-D joint refraction and reflection traveltime tomographic inversion. In addition to the inversion of seismic P-wave velocity and reflector depth, the code can now retrieve models of the Thomsen's parameters δ and ε. Here we test this new implementation following four different strategies on a canonical synthetic experiment in ideal conditions with the purpose of estimating the maximum capabilities and potential weak points of our modeling tool and strategies. First, we study the sensitivity of traveltimes to the presence of a 25 % anomaly in each of the parameters. Next, we invert for two combinations of parameters, (v, δ, ε) and (v, δ, $v^\perp$), following two inversion strategies, simultaneous and sequential, and compare the results to study their performances and discuss their advantages and disadvantages. Simultaneous inversion is the preferred strategy and the parameter combination (v, δ, ε) produces the best overall results. The only advantage of the parameter combination (v, δ, $v^\perp$) is a better recovery of the magnitude of v. In each case we derive the fourth parameter from the equation relating ε, $v^\perp$ and v. Recovery of v, ε and $v^\perp$ is satisfactory whereas δ proves to be impossible to recover even in the most favorable scenario. However, this does not hinder the recovery of the other parameters, and we show that it is still possible to obtain a rough approximation of δ distribution in the medium by sampling a reasonable range of homogeneous initial δ models and averaging the final δ models that are satisfactory in terms of data fit.

1 Introduction

An isotropic velocity field is the rare exception in the Earth subsurface. Anisotropy is a multiscale phenomenon, and its causes are diverse. In the crust, it can be produced by the preferred orientation of mineral grains or their crystal axes (Schulte-Pelkum and Mahan, 2014; Almqvist and Mainprice, 2017), the alignment of cracks and fracture networks and the presence of fluids (Crampin, 1981; Maultzsch et al., 2003; Yousef and Angus, 2016), or the bedding of layers much thinner than the wavelength used to explore them (Backus, 1962; Johnston and Christensen, 1995; Sayers, 2005). In the mantle, anisotropy is related to the alignment of olivine crystals due to mantle flow (Nicolas and Christensen, 1987; Montagner et al., 2007), aligned melt inclusions (Holtzman et al., 2003; Kendall et al., 2005), large-scale deformation (Vinnik et al., 1992; Vauchez et al., 2000), and preexisting lithospheric fabric (Kendall et al., 2006), among others. Anisotropy has proven an

30   informative physical property in the understanding of the Earth's interior (Ismaïl and Mainprice, 1998; Long and Becker, 2010), most particularly in continental rifts (Eilon et al., 2016), mid-ocean ridges (Dunn et al., 2001), and subduction zones (Long and Silver, 2008).

The theory of anisotropic wave propagation has been described in several publications (Kraut, 1963; Babuska and Cara, 35   1991). Numerous formulations of varying complexity have been proposed to approximate anisotropy depending on its magnitude and the symmetry conditions of the medium (Nye, 1957). Twenty-one elastic stiffness parameters define the most general anisotropic medium with the lowest symmetry conditions, whereas for the highest symmetry not equivalent to isotropy, only five parameters are needed (Almqvist and Mainprice, 2017). Regarding the strength of anisotropy, in view of the overall success of isotropic methods in studying the Earth's subsurface, and of the experimental evidence and sample 40   measurements available, it is admitted that the anisotropy is generally weak (Thomsen, 1986). Specifically, anisotropy is considered weak when Thomsen's parameters are much smaller than 1, i.e. for a ~20 % or smaller velocity variation with angle. Precisely Thomsen (1986) presented the formulation for the transverse isotropy symmetry on weakly anisotropic media, which is the reference that we follow in this work. Thomsen's parameters are by far the most common and convenient combinations of stiffness tensor elements used in seismic anisotropy modeling (Tsvankin, 1996; Thomsen and Anderson, 45   2015). Applications of simpler approximations exist, as is the case of elliptical symmetry (Song et al., 1998; Giroux and Gloaguen, 2012), as well as others that assume the most general anisotropic model (Zhou and Greenhalgh, 2008).

The objectives of this work are (1) presenting the anisotropic version of TOMO3D (Meléndez et al., 2015) for the study of VTI weakly-anisotropic media in terms of the Thomsen's parameters δ and ε using P-wave arrival times, and (2) comparing 50   several parametrizations and inversion strategies under optimal and equal conditions for all parameters with the purpose of defining an upper limit to the code's capabilities, an ideal but generalizable estimation of the code's performance, as well as highlighting its potential weaknesses. Moreover, the development of this code is motivated by the need to combine wide-angle and near-vertical traveltime picks in field data applications, as we plan to do with the trench-parallel 2-D profile in Sallarès et al. (2013), which is affected by a ~15 % anisotropy judging from the mismatch in the interplate boundary 55   locations obtained separately from near-vertical and wide-angle data. The P-wave velocity, δ and ε models obtained from the modeling of field data would be useful and geologically informative by themselves, but they would also serve as initial models in anisotropic FWI.

In the following section, we describe the anisotropy formulation and the modifications implemented on our 3-D joint 60   refraction and reflection traveltime tomography code TOMO3D to incorporate the inversion of Thomsen's parameters. Next, in the third section, we present the synthetic tests performed and their results, including accuracy and sensitivity analyses and synthetic inversions. In section four, these results are discussed and interpreted in terms of the ability of the code to retrieve

both the velocity field and Thomsen's anisotropy parameters. Finally, in the last section we summarize the main conclusions of this work.

2 Modeling anisotropy

The first part of this section is a general overview of the treatment of anisotropy within the field of seismic inversion, while the second one describes the implementation of Thomsen's weak VTI anisotropy formulation in the 3-D joint refraction and reflection traveltime tormography code TOMO3D.

2.1 Anisotropy in seismic inversion methods

Anisotropy was first incorporated to seismic inversion methods in traveltime tomography with the development of the linearized perturbation theory (Cervený, 1982; Cervený and Jech, 1982; Jech and Psencik, 1989). Previously, the approach to deal with anisotropy was to approximately remove its estimated effect to then apply an isotropic method (e.g. McCann et al., 1989). Linearized perturbation theory was first implemented in anisotropic traveltime tomography by Chapman and Pratt (1992) and Pratt and Chapman (1992) assuming the weak anisotropy approximation, which allowed them to use isotropic ray tracing and approximate anisotropy effects as being caused by small perturbations of the isotropic system. The initial development of anisotropic ray tracing is attributed to Cervený (1972). Methods for anisotropic ray tracing and traveltime computation depend on the symmetry assumptions made regarding the medium. The most common of those is rotational symmetry around a vertical pole. This formulation is known as vertical transverse isotropy (VTI), also polar anisotropy (e.g. Rüger and Alkhalifah, 1996; Alkhalifah, 2002), and it is the simplest geologically applicable case: it reproduces the symmetry exhibited by minerals in sedimentary rocks, and that produced by parallel cracks or fine layering. Furthermore, it significantly simplifies the mathematical formulae since anisotropy is defined by only five parameters, which contributes to a greater computational efficiency. The generalization of VTI to a tilted symmetry axis is the so called tilted transverse isotropy (TTI). Some authors argue that it is not possible to distinguish TTI from VTI in real experimental cases without a priori information (Bakulin et al., 2009). Assuming the most general anisotropic media has also become rather usual, in particular with the improvement of computational resources, allowing for a more detailed and complex reconstruction of the subsurface physical properties (e.g. Zhou and Greenhalgh, 2005) although successful field data applications are yet to be achieved to the best of our knowledge. Regarding the inversion process, the main difficulty arises from the trade-off between velocity heterogeneity and anisotropy (e.g. Bezada et al., 2014). To the best of our knowledge, Stewart (1988) was the first to propose an inversion algorithm, specifically for the recovery of Thomsen's parameters in a weakly-anisotropic VTI medium. Other authors have produced inversion algorithms for different formulations such as azimuthal anisotropy (e.g. Eberhart-Phillips and Henderson, 2004; Dunn et al., 2005) or a 3-D TTI medium (e.g. Zhou and Greenhalgh, 2008). Concerning FWI, anisotropy in active data is typically modeled following Thomsen's parameters and the VTI and/or TTI approximation for the medium. The first anisotropic wave propagators appeared during the 80s and 90s (e.g. Helbig, 1983;

Alkhalifah, 1998) and new improvements on this matter continue today (e.g. Fowler et al., 2010; Duveneck and Bakker, 2011). When performing anisotropic full waveform inversion (FWI), both 2-D and 3-D, some authors choose to invert only for the velocity field, fixing the initial anisotropy models throughout the inversion because it simplifies the process (e.g. Prieux et al., 2011; Warner et al., 2013). However, other works have explored the feasibility of multiparameter inversions, that is, using different combinations of velocity and anisotropy parameters and of inversion strategies (e.g. Gholami et al., 2013a,b; Alkhalifah and Plessix, 2014).

2.2 Anisotropy in TOMO3D: Thomsen's weak VTI anisotropy formulation

We adapted TOMO3D (Meléndez et al., 2015) to perform anisotropic ray tracing and traveltime calculations, as well as inversion of Thomsen's parameters for P-wave data. In TOMO3D, the forward problem solver is parallelized to simultaneously trace rays for multiple sources and receivers, and it uses an hybrid ray tracing algorithm that combines the graph or shortest path method (Moser, 1991) and the bending refinement method (Moser et al., 1992). The inverse problem is solved sequentially using the LSQR algorithm (Paige and Saunders, 1982). Velocity models are discretized as 3-D orthogonal and vertically-sheared grids that can account for topography and/or bathymetry. Velocity values are assigned to the grid nodes, and the velocity field is built by trilinear interpolation within each cell. Apart from first-arrival traveltimes, the code allows for the inversion of reflection traveltimes to obtain the geometry of major geological boundaries associated to impedance contrasts that produce strong seismic energy reflections in the data recordings. Such reflecting interfaces are modeled as 2-D grids independent of the velocity grid. The code is also prepared to extract information from the water-layer multiple of refracted and reflected seismic phases (Meléndez et al., 2014). A detailed description of the code can be found in Meléndez (2014).

Our anisotropy formulation is based on Thomsen (1986) and specifically in the following weakly-anisotropic velocity equation for the P-wave velocity:

$$v^a(v, \delta, \varepsilon, \theta) = v \cdot \left(1 + \delta \cdot sin^2(\theta) \cdot cos^2(\theta) + \varepsilon \cdot sin^4(\theta)\right) \tag{1}$$

where v^a is the anisotropic velocity, v is the velocity along the symmetry axis (α_0 in Thomsen, 1986), θ is the angle with respect to the symmetry axis, and δ and ε are Thomsen's parameters controlling the anisotropic P-wave propagation. Studying the cases of $\theta=0$ and $\theta=\pi/2$ the meaning of ε becomes clear: it is the relative difference between the velocities along and across the symmetry axis, that we refer to as parallel and perpendicular velocities, respectively.

$$v^a(v, \theta = 0) = v$$

$$v^a\left(v, \theta = \frac{\pi}{2}\right) = v \cdot (1 + \varepsilon) \equiv v^{\perp}$$

125

$$\varepsilon = (v^{\perp} - v)/v \tag{2}$$

According to Thomsen (1986) the meaning of δ is far from intuitive but the author states that it is associated to the near-vertical anisotropic response and shows that it relates v and the normal move-out velocity (V_{NMO}). V_{NMO} models are a by-product of multichannel seismic reflection data processing, a mathematical construct that involves the assumptions of a stratified media with constant velocity layers and of small spread, i.e. near-vertical propagation. It does not seem wise to try estimating it by other means, less so if the data and modeling used do not necessarily fulfill the assumptions for the normal move-out correction that define V_{NMO}. Moreover, we want to combine traveltimes from as many types of seismic data sets as possible, notably from multichannel reflection (near-vertical propagation) and wide-angle (sub-horizontal propagation) experiments, in order to have the best polar coverage, and with that, the best recovery of v and ε (or $v^{\perp}$). Thus, we do not consider V_{NMO} to be a useful parameter in describing the general anisotropic VTI media for our modeling method, and we did not implement parametrizations (v, V_{NMO}, ε) and (v, V_{NMO}, $v^{\perp}$).We do think however that V_{NMO} models can help in the building of initial δ models, just as the comparison between near-vertical propagation and sub-horizontal data can provide an initial estimation of ε. In conclusion, we only considered Eq. (2), and we implemented two parametrizations of the medium: (v, δ, ε) and (v, δ, $v^{\perp}$). From here on, for simplicity, we will refer to them as P[ε] and P[$v^{\perp}$] respectively.

The linearized inverse problem matrix equation including anisotropy parameters for a refraction-only case is as follows

$$
\begin{pmatrix} \Delta t^0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
=
\begin{pmatrix}
\boldsymbol{G}^{u0} & \boldsymbol{G}^{\delta0} & \boldsymbol{G}^{\varepsilon0} \\
\lambda_u \boldsymbol{L}^{uX} & 0 & 0 \\
\lambda_u \boldsymbol{L}^{uY} & 0 & 0 \\
\lambda_u \boldsymbol{L}^{uZ} & 0 & 0 \\
0 & \lambda_\delta \boldsymbol{L}^{\delta X} & 0 \\
0 & \lambda_\delta \boldsymbol{L}^{\delta Y} & 0 \\
0 & \lambda_\delta \boldsymbol{L}^{\delta Z} & 0 \\
0 & 0 & \lambda_\varepsilon \boldsymbol{L}^{\varepsilon X} \\
0 & 0 & \lambda_\varepsilon \boldsymbol{L}^{\varepsilon Y} \\
0 & 0 & \lambda_\varepsilon \boldsymbol{L}^{\varepsilon Z} \\
\alpha_u \boldsymbol{D}^{u} & 0 & 0 \\
0 & \alpha_\delta \boldsymbol{D}^{\delta} & 0 \\
0 & 0 & \alpha_\varepsilon \boldsymbol{D}^{\varepsilon}
\end{pmatrix}
\begin{pmatrix} \Delta u \\ \Delta \delta \\ \Delta \varepsilon \end{pmatrix}
\tag{3}
$$

145

Smoothing (**L**) and damping (**D**) constraints for δ and ε parameters follow the same formulation described in Meléndez (2014) for velocity parameters. The kernels (**G**) have been modified to account for anisotropy. The linearized and discretized equation that relates the traveltime residual of the n-th refracted pick to changes in the model parameters is written as

$$\Delta t_n^0 = \sum_i^I \sum_{m=1}^8 r_m^u \cdot \frac{\partial t}{\partial u^a} \cdot \frac{\partial u^a}{\partial u} \cdot \Delta u_i + \sum_j^J \sum_{m=1}^8 r_m^\delta \cdot \frac{\partial t}{\partial u^a} \cdot \frac{\partial u^a}{\partial \delta} \cdot \Delta \delta_j + \sum_k^K \sum_{m=1}^8 r_m^\varepsilon \cdot \frac{\partial t}{\partial u^a} \cdot \frac{\partial u^a}{\partial \varepsilon} \cdot \Delta \varepsilon_k \qquad (4)$$

In each of the three terms, the first summation corresponds to the model cells that are illuminated by the n-th ray path. A cell is considered illuminated if it contains a ray path segment. The second summation is over the eight nodes of each of those cells. In the third term, ε is replaced by $v^\perp$ when using this alternative parametrization. Δt_n^0 is the traveltime residual for the n-th refracted pick, $u^a = 1/v^a$ is the anisotropic slowness, $u=1/v$ is the along-axis slowness, Δu_i, $\Delta \delta_j$, and $\Delta \varepsilon_k$ (or $\Delta v_k^\perp$) are the parameter perturbations for each illuminated cell in their respective grids, and the r_m factors are the weights that distribute these perturbations among the eight nodes of each illuminated cell according to the trilinear interpolation used to define the four fields (u, δ, ε, and $v^\perp$).

In order to build the kernel matrices we need to compute two partial derivatives for each model parameter. The first order partial derivative of traveltime with respect to the anisotropic slowness is the ray path segment s_i within each cell that is covered in a given traveltime at a given slowness

$$t = \sum_{i=1}^N u_i^a \cdot s_i \qquad (5)$$

$$\frac{\partial t}{\partial u_i^a} = s_i \qquad (6)$$

From Eq. (1) the first order partial derivatives of the anisotropic slowness with respect to the model parameters (u, δ, ε, and $v^\perp$) are as follows

$$\frac{\partial u^a}{\partial u} = \frac{1}{1+\delta \cdot sin^2(\theta) \cdot cos^2(\theta) + \varepsilon \cdot sin^4(\theta)} \qquad (7)$$

$$\frac{\partial u^a}{\partial \delta} = \frac{-u \cdot sin^2(\theta) \cdot cos^2(\theta)}{\left(1+\delta \cdot sin^2(\theta) \cdot cos^2(\theta) + \varepsilon \cdot sin^4(\theta)\right)^2} \qquad (8)$$

$$\frac{\partial u^a}{\partial \varepsilon} = \frac{-u \cdot sin^4(\theta)}{\left(1+\delta \cdot sin^2(\theta) \cdot cos^2(\theta) + \varepsilon \cdot sin^4(\theta)\right)^2} \qquad (9)$$

$$\frac{\partial u^a}{\partial v^\perp} = \frac{-(u)^2 \cdot \sin^4(\theta)}{\left(1 + \delta \cdot \sin^2(\theta) \cdot \cos^2(\theta) + \varepsilon \cdot \sin^4(\theta)\right)^2} \qquad (10)$$

3. Synthetic tests

We have performed a number of tests using canonical synthetic models made of an anomaly centered in a uniform
background with two main objectives: (1) checking that the newly implemented anisotropic traveltime tomography method
works properly, and (2) providing a quantitative measure of the potential recovery of anisotropy based on P-wave traveltimes
alone. First, we run a sensitivity test to assess the effect that a variation in each model parameter has in the synthetic
traveltimes, and we calibrated the code by comparing the synthetic data that it generates to analytically calculated data. Next,
we performed a number of synthetic inversion tests considering both possible parametrizations and inversion strategies.
These tests are conducted under ideal and equal conditions for all parameters with the purpose of obtaining an upper-limit
but widely-applicable estimation of the code's performance, and detecting any potential weak points.

The models in all these tests are cubes of 5-km-long edge. The background model of all four parameters is set to a constant
value, i.e. v, δ, ε and $v^\perp$ background models are homogeneous. Note that the z axis positive direction points downwards. Grid
spacing is 0.125 km for the four parameters in all three dimensions so that differences in model discretization do not
influence the test results. The volume of the anomaly is determined by the 3σ region of a 3-D Gaussian function centered in
the cube setting $3\sigma = 0.5$ km. The values of v, δ, ε and $v^\perp$ within this volume are homogeneously increased resulting in a
discretized representation of a spherical anomalous body.

3.1 Sensitivity

We define the sensitivity of a parameter as the difference between the first arrival traveltimes with and without the anomaly
in that parameter. Figure 2 shows synthetic and analytic sensitivities for two selected meridians. We express sensitivity both
as normalized and as relative traveltime difference. In its normalized form, traveltime differences are divided by the greatest
of these differences among all parameters, i.e. normalized to 1, whereas in its relative form, these differences are given with
respect to the traveltime without the anomaly and multiplied by 100. Note that the analytic response does not change
between meridians, i.e. given the symmetry of the models and of the anisotropic formulation, sensitivity is independent of
the azimuth angle.

The acquisition configuration consists of 482 diametrically-opposed source–receiver pairs. Each receiver records exclusively
the first arrival traveltime from its corresponding source for a total of 482 traveltimes (Fig. 1). Sources and receivers are
located at 2.5-km of the center of the model, at the locus defined by the surface of the sphere inscribed in the cube, and
placed at the crossing points of 32 meridians with 15 parallels, and at each pole.

According to the definition of sensitivity, alternately for v, δ, and ε, the said anomaly was added at the center of the cube representing a 25 % increase on the background value, while the models for the rest of parameters in the parametrization remained homogeneous. Table 1 summarizes the background and anomaly values for all parameters in each sensitivity test, along with their equivalence in the alternative parametrization. Since v is related to ε and $v^\perp$ through Eq. (2) its sensitivity pattern changes depending on the parametrization used (Fig. 2). Indeed, a 25 % increase in v with respect to equivalent background models in P[ε] and P[$v^\perp$] (Table 1) yields different sensitivities because of the different parameters involved (ε or $v^\perp$) in the representation of the medium. Contrarily, δ sensitivity pattern is independent of the parametrization used. ε and $v^\perp$ sensitivities are only defined in their respective parametrizations. However, a 25 % increase in $v^\perp$ yields an equivalent ε of 0.45 which is greater than the ~0.2 limit for weak anisotropy approximation. Thus, instead of establishing the comparison with $v^\perp$ sensitivity based on a proportional anomaly increment and measuring its effect on traveltimes, we based it on an equal traveltime change, i.e. the same change in traveltime requires a change of 25 % in ε but only a ~3.4 % change in $v^\perp$ (Table 1), indicating that data is ~7 times more sensitive to $v^\perp$ than to ε changes.

v sensitivity in P[ε] is the highest for all angles, 4.5 % to 5 % (Fig. 2), and it can be shown that expressed in its relative form, the analytic solution is a constant 5 % (see mathematical proof in supplementary material). In its normalized form it follows the same sinusoidal pattern as in P[$v^\perp$] (Fig. 2e), and both have equal maxima in the directions parallel to the symmetry axis. In both cases minima are found in the directions perpendicular to the symmetry axis, but in P[$v^\perp$] they reach down to 0 whereas in P[ε] the value is ~0.85. As expected from Eq. (1), ε sensitivity goes to 0 % in the directions parallel to the symmetry axis and has its maxima (~0.8 % or ~0.15) for the polar angles perpendicular to it. $v^\perp$ sensitivity would follow the same angular dependence but with maxima of the order of magnitude of v sensitivities. Finally, δ sensitivity is, at its maxima, more than one order of magnitude smaller than v sensitivity, around 0.25 % or 0.05. These sensitivity results indicate that we can generally expect similar recoveries for v and $v^\perp$, better than for ε, and that retrieving δ might prove complicated. Keep in mind that these sensitivities for v, δ, and ε are produced by an anomaly that represents a 25 % increase with respect to the background value, and that an anomaly in $v^\perp$ would produce the same sensitivity pattern as ε with only a ~3.4 % increment.

The differences in synthetic sensitivities between meridians arise from the discretization of the model space in a Cartesian system of coordinates. Such approximation inevitably defines privileged directions for ray tracing, and consequently produces differences in synthetic traveltimes. The mismatch between synthetic and analytic sensitivities (Fig. 2) occurs because the discretization used cannot represent the surface of a perfect sphere. These effects are most notable in the v sensitivity, in P[ε] and to a lesser extent in P[$v^\perp$], precisely because it is the most sensitive parameter, and thus the errors in the representation of a sphere and the existence of privileged directions have a much larger influence on the calculated traveltimes. Figure S1 shows how refining the v model reduces the relative traveltime error (Fig. S1a), and generates a more

accurate sensitivity pattern (Fig. S1b). In a real case study one can always refine the grid spacing of a particular parameter to achieve better accuracy, but here we wish to test the performance of the code in the modeling and recovery of each parameter under the same conditions, i.e. equivalent anomalies and identical model discretization.

3.2 Accuracy

245   For the four simulations in the sensitivity analysis, we compared the synthetic traveltimes obtained with our code to the analytic solution to quantify the accuracy of the code's performance. The comparison along the two selected meridians at 0 rad and $\pi/4$ rad azimuths is displayed in Fig. 3 expressed as relative traveltime error, i.e. the difference between synthetic and analytic traveltimes relative to the latter in percentage. Table 2 contains the mean of the relative traveltime errors and their respective mean deviations for these four tests and along the two selected meridians, as well as the overall values for each of

250   them.

Comparison of Fig. 2a,b with the values in Table 2 and Fig. 3 indicates that the forward calculation of traveltimes is accurate enough with respect to the traveltime residuals expected for the selected anomalies. Sensitivity is at least 5 times and up to two orders of magnitude greater than traveltime accuracy errors depending on the parameters, with the exception of angles

255   for which sensitivity tends to zero. Figure S2 illustrates that the code is able to reproduce nearly identical accuracies using both parametrizations. Furthermore, given that we are using the same synthetic traveltimes that we used for the sensitivity analysis, this also implies that the sensitivity patterns obtained with the alternative parametrization would be virtually equal to those in Fig. 2.

3.3. Inversion results

260   For the inversion tests, we considered a synthetic medium defined by the anomaly models of all four parameters. Here we refer to these models as target models, and the goal of the inversion is to retrieve the heterogeneity in each of them. These tests are conducted for the two parametrizations of the anisotropic medium described in section 2: $P[\varepsilon]$ and $P[v^{\perp}]$. Note that, in order to perform the inversion tests on equivalent cases for both parametrizations, the heterogeneity in $v^{\perp}$ is calculated with Eq. (2) considering the 25 % anomalies in v and ε, which yields a ~29.3 % anomaly in $v^{\perp}$ (Table 3). If not indicated

265   otherwise, we use background models as initial models. Finally, we study the potential recovery of δ because inverting this particular parameter proves notoriously difficult due to its low sensitivity (Fig. 2).

The synthetic data set is made of 114 sources each recorded at 113 receivers for a total 12,882 first arrival traveltimes. For the acquisition geometry, again sources and receivers are located at the surface defined by the sphere inscribed in the cube

270   (Fig. 1). The 114 positions at the surface of this sphere are shared by sources and receivers, and each receiver records all sources, except for the one source located at its same position.

For both parametrizations, we compared two inversion strategies: simultaneously inverting for all parameters and a two-step sequential inversion. First, in Figs 5 and 6 we show the best results for the simultaneous inversion strategy. For each parametrization we derived the fourth parameter applying Eq. (2).

Table 4 shows several statistical measures to quantify the quality of these inversion results. As a measure of data fit improvement, we provide the RMS of traveltime residuals for the first and last iterations. As a measure of model recovery or fit, for each parameter, we calculate the mean relative difference for the background area between the inverted model and either the target or the initial one, as they are identical in this area, as well as for the anomaly area comparing the inverted model to both the target and the initial models. In the case of a perfect recovery, mean relative difference for the background area would be 0 %, whereas for the anomaly area, it would be 0 % when using the target model as a reference, and 25 % (~29.3 % for $v^{\perp}$) when comparing to the initial model.

In an attempt to improve the recovery of δ, we repeated these two tests for different values of the smoothing constraints, but it proved impossible. Correlation lengths tested for all four parameters include 0.25 km (twice the grid spacing), 0.5 km, and 1 km. The weights of the smoothing submatrices for each parameter, λ in Eq. (3), were varied between 1 and 100, with intermediate values of 2, 5, 10, 20, 30, and 60. For successful inversions, very similar results for the other parameters were obtained regardless of the final δ model. In other words, the low sensitivity of δ makes it extremely hard, if not impossible, to recover this parameter from traveltime data, but for this same reason it has little or no influence on the recovery of v and ε or $v^{\perp}$.

The two-step sequential inversion strategy was also tested for both parametrizations P[ε] and P[$v^{\perp}$]. For the first step, we tested two options: (a) inverting for v while fixing δ and ε or $v^{\perp}$ and (b) fixing only δ. In the second step we used the inverted models from step 1 as initial models and tested three options: (c) inverting for all three parameters, (d) fixing only δ, when following option (a) in step 1, and (e) fixing v and/or ε or $v^{\perp}$, when following option (b) in step 1. Figure 4 shows a flowchart describing all the sequential inversion combinations tested. Again smoothing constraints were varied for similar correlation lengths and submatrix weights as detailed for the simultaneous inversion case.

As indicated in Fig. 4, in the case of P[ε], the best result (Fig. 7) was obtained inverting for v and ε while fixing δ in the first step, and fixing only ε in the second step, whereas for P[$v^{\perp}$], the best combination for the two-step inversion (Fig. 8) was fixing only δ in step 1, and inverting for the three parameters in step 2. Tables 5 and 6 summarize the statistical quantification of data and model fit for each parametrization.

3.4 Modelling δ

305   Observing that good results for v and ε or $v^{\perp}$ are achieved regardless of the result in δ, and knowing that the sensitivity of δ is notably smaller than that of the other parameters, we explored a strategy to have an estimate of this parameter. First, as a reference, we considered an unrealistically optimal scenario in which the real v and ε or $v^{\perp}$ models are known to us. Figs 9 and 10 show the resulting δ achieved by repeating inversions in Figs 5 and 6 with v and ε or $v^{\perp}$ target models as initial models. Table 7 summarizes the traveltime residuals RMS and the mean relative difference for each parameter in these

310   inversions. Again, these two tests were repeated for ranges of smoothing constraints in all four parameters, as described for the cases in Figs 5 and 6. Table 7 and Figs 9 and 10 correspond to the best results obtained, which indicate that the recovery of δ is, at best, extremely complicated due to the limited sensitivity of traveltime data to changes in this parameter.

  Next, we decided to try neglecting δ in Eq. (1), and we repeated a number of inversions, such as the ones displayed in Figs 5

315   and 6, following the equation

$$v^{a}(v,\varepsilon,\theta)=v \cdot \left(1 + \varepsilon \cdot sin^{4}(\theta)\right) \tag{11}$$

  The purpose of these tests was checking whether it was possible to invert v and ε or $v^{\perp}$ with data generated following Eq. (1)

320   using the approximation in Eq. (11) given that the influence of δ on the results for other parameters is rather small, that δ cannot be accurately retrieved from traveltime alone, and that it has the smallest sensitivity. To do so, a homogeneous model of $\delta=0$ was fixed throughout the inversions. These tests were unsuccessful, with noticeably poorer results than when considering a dependence on δ (Table 8). However, they were useful in proving that even if a detailed δ model is not necessary to successfully retrieve the other parameters, at least a rough approximation of the δ field is needed to recover the

325   other parameters, e.g. the background δ model that we used as initial model in inversions displayed in Figs 5 and 6.

  Finally, we tested whether it would be possible to obtain at least this rough approximation of δ in the medium, valid in the sense that it allows for the successful recovery of the rest of parameters, using any a priori information available such as compilations of anisotropy measurements (e.g. Thomsen, 1986; Almqvist and Mainprice, 2017). Once again we repeated

330   inversions from Figs 5 and 6 (initial $\delta = 0.16$), now using different homogeneous initial models for δ within a range of possible values from 0.1 to 0.24. Table 9 contains the initial and final RMS of traveltime residuals, as well as δ mean values for the inverted model along with the corresponding mean deviations. It is straightforward to note that, for a central subrange of the tested initial δ values, final RMS values are an order of magnitude smaller, a few tenths of millisecond compared to the few milliseconds outside this subrange. Specifically, very similar results to those in Figs 5 and 6 in terms of traveltime

335   residuals RMS are produced by initial δ values between 0.13 and 0.22 for $P[\varepsilon]$, and between 0.12 and 0.22 for $P[v^{\perp}]$. The

narrowing of the initial δ distribution to a smaller subrange of mean δ values for the inverted models is indicative of a good general convergence trend.

The rough estimate of the δ field could be built, for instance, as the average of the mean δ values for the inverted models in the central subranges defined by the change in magnitude of the final RMS of traveltime residuals. One final inversion could be run using a homogeneous initial δ model with this average value, with the additional option of fixing it and inverting only for the other two parameters. As mentioned in subsection 2.2, potentially more detailed initial δ models could be obtained from the normal move-out correction of near-vertical reflection seismic data.

4. Discussion

We have tested two parametrizations of the VTI anisotropic media, P[ε] (v, δ, ε) and P[$v^\perp$] (v, δ, $v^\perp$), and two inversion strategies, simultaneously inverting for all three parameters and a two-step sequential process fixing some of the parameters in each step. We consider three criteria for evaluating and comparing the quality of the inversion results obtained following the four possible combinations of strategies and parametrizations: visual inspection of the results, as well as traveltime data and model fits. For both inversion strategies and both parametrizations, the recovery of the parameters is positively correlated with their respective sensitivities; the more sensitive parameters are systematically better recovered.

For the simultaneous inversion, both parametrizations were able to produce acceptable final results (Figs 5 and 6). According to our tests, P[ε] provides the best outcome, specifically because data and model fits (Table 4) are better for this option but, more importantly, because of the difference in the quality of the recovery of ε. However, visual comparison of the recovery of v as well as the v model fit in Table 4 indicate that P[$v^\perp$] yields a slightly better result regarding the magnitude of the anomaly for this parameter, as is particularly evident at its center. In P[ε], given the disparity in sensitivities between v and ε, the former might be prone to accounting for data misfit that actually corresponds to the latter, whilst in P[$v^\perp$] this disparity is less acute. The boundaries of the anomalies for both velocities are still better retrieved with P[ε]. Also, the recovery of δ, even though it is far from acceptable, is significantly better in the case of P[ε]. This might be explained because the disparity of δ sensitivity is greater with respect to the other two parameters in P[$v^\perp$] than in P[ε].

In general, sequential inversion is a more complex process that requires more human intervention and fine tuning in each step. In addition, fixing some of the parameters in the first step may result in the inverted parameters artificially accounting for part of the data misfit that is actually related to the fixed ones. This can more easily lead convergence into a local minimum, and it might be impossible to correct this tendency in the second step. For this inversion strategy, it is also P[ε] that produces the best results (Figs 7 and 8). Data fit and the model fit of ε are slightly better than for the simultaneous inversion of this parametrization (Tables 4 and 5), whereas model fits and the visual aspect of both velocities are almost

identical to those obtained by simultaneously inverting all three parameters (Figs 5 and 7). Visually, it is difficult to decide whether the recovery of ε is better or not than for the simultaneous inversion (Figs 5 and 7). As for δ, recovery is unsuccessful and artifacts appear in the background area of the model but, according to both its model fit and its visual aspect, it is notably better than for the simultaneous inversion. As in the case of simultaneous inversion, the only advantage of using P[$v^\perp$] instead of P[ε] is that it yields a better recovery of the anomaly magnitude of v (Tables 4 and 6). The results for both velocities are virtually identical to those obtained by simultaneous inversion of this parametrization (Figs 6 and 8). δ and ε are not properly retrieved, but the results are significantly better than for the simultaneous inversion of this same parametrization.

δ has been shown to be by far the most complicated parameter to retrieve because of the low sensitivity of traveltime to its variation (Fig. 2). Even when excellent v and ε or $v^\perp$ models are available, i.e. the target models for these parameters in our synthetic tests, the recovery of δ is limited at best (Figs 9 and 10 and Table 7). However, and for the same reason, poor recoveries of δ do not affect the recovery of the other two parameters, meaning that a detailed δ model is not necessary to satisfactorily retrieve v and ε or $v^\perp$ (Figs 5-10). Still, our inversion tests also proved that neglecting δ in Eqs (1) and (3) is not an option, the accuracy in the recovery of the other parameters resulting severely affected (Table 8). Thus, even if a detailed inversion of δ is, at the very least, hard to achieve, and it is not needed for a successful result in the other parameters, some sort of simple, even homogeneous, initial δ model with a value or values about the average δ in the medium, is necessary for a good recovery of the other parameters.

We showed that given some a priori information on the range of possible δ values in the medium, it should be possible to create the necessary initial δ model. In order to illustrate this, we chose a range of δ values for the initial model and rerun the inversions in Figs 5 and 6. The results indicate that for any homogeneous initial δ model in a certain subrange close to the actual average δ value of the medium, the results for v and ε or $v^\perp$ are satisfactory and virtually identical to those of the original inversions (Table 9). This subrange is easily defined by looking at the final RMS of traveltime residuals, which experiences a notorious change of one order of magnitude. Any model within this subrange works similarly well as initial δ model. Alternatively, a possibly more robust selection of the constant value for a homogeneous initial δ model might be the mean (or also the median or the mode) of the mean δ values for the inverted models in this subrange. It is worth noting that, whereas for the purpose of this work we used the same discretization for all parameters, in a real case study it would probably be recommendable to use a coarser discretization for δ than for the other parameters, and in general a finer discretization for the more sensitive parameters (Fig. S1). Indeed, a heterogeneity of a given spatial scale and relative variation will produce a greater effect on data for a parameter of greater sensitivity. Thus, for a parameter of greater sensitivity, it will be easier for the code to identify smaller heterogeneities both in scale and variation, which will require a finer grid.

5 Conclusions

We have successfully implemented and tested a new anisotropic traveltime tomography code. For this implementation we had to modify both the forward problem and the inversion algorithms of the TOMO3D code (Meléndez et al., 2015). The forward problem was adapted to compute the velocities observed by rays considering Eq. (1) for the weak VTI anisotropy formulation in Thomsen (1986). The inversion solver was extended to include the δ, ε and $v^\perp$ kernels in the linearized forward problem matrix equation, as well as smoothing and damping matrices for these parameters defined following the same scheme as for velocity in the isotropic code (Eqs 3-10).

Regarding the synthetic tests, we first determined the sensitivity of traveltime data to changes in each of the parameters defining anisotropy in the medium (v, δ, ε, $v^\perp$) (Fig. 2), and we checked the proper performance of the code by comparing the synthetic traveltimes produced with their respective analytic solutions (Fig. 3). Next, we performed canonical inversion tests to compare two possible media parametrization, P[ε] and P[$v^\perp$], and two possible inversion strategies, simultaneous and sequential. According to our tests, both parametrizations have their strengths: P[ε] produces the best overall result in the sense that all parameters are acceptably recovered, with the exception of δ, and trade-off between parameters is lower, but P[$v^\perp$] yields the best result for the magnitude of the anomaly in v. Regarding the inversion strategy, simultaneous inversion is more straightforward and involves less human intervention, and given that both strategies yield similar results, it would be our first choice. Sequential inversion is always a more complex process that can be shown to work in a synthetic case because the target models are available, but in field data applications the complexity would most likely be unmanageable. These tests were conducted under ideal conditions, and thus the conclusions provided by their results are an upper-limit but generalizable estimation of the strengths and weaknesses in the code's performance.

An acceptable recovery of δ turned out to be impossible due to the small sensitivity of traveltimes to this parameter, but we verified that it cannot simply be neglected in the equations. Whereas the recovery of the other parameters is not significantly affected by that of δ, a rough estimate of the average δ value in the medium is necessary and sufficient to generate a homogeneous initial model that allows for satisfactory inversion results in these other parameters. We also proved that it is possible to obtain it, provided that some a priori knowledge on δ values in the medium is available to define a range of plausible values, such as field or laboratory measurements.

Code availability

The anisotropic version of TOMO3D is available only for academic purposes on our group website at www.barcelona-csi.cmima.csic.es/software-development.

Data availability

All data and files used in the synthetic tests presented in this article are available at the repository Digital.CSIC. The corresponding DOI is https://dx.doi.org/10.20350/digitalCSIC/8957.

Author contribution

435 The formulation of the overarching research goals of this work is a product of discussion among the four co-authors. AM was in charge of software development and data curation, analysis, visualization, and validation. AM also prepared the manuscript. The methodology for the synthetic tests was designed by AM, CEJ, and VS. CR was responsible for the acquisition of financial support.

Competing interests

440 The authors declare that they have no conflict of interest.

Acknowledgements

AM and CLJ are funded by Respol through the SOUND collaboration project with CSIC, and the work in this paper was conducted at the Grup de Recerca de la Generalitat de Catalunya 2009SGR146: Barcelona Center for Subsurface Imaging (B-CSI). We thank all our fellows at the B-CSI for their contribution to this work. We also wish to thank Dr. Marko Riedel
445 for his constructive comments at the 18th edition of the biennial International Symposium on Deep Seismic Profiling of the Continents and their Margins (SEISMIX 2018). The comments of the anonymous referees have also contributed significantly to the improvement of our work.

References

Alkhalifah, T.: Acoustic approximations for processing in transversely isotropic media, Geophysics, 63, 623-631,
450 https://doi.org/10.1190/1.1444361, 1998.

Alkhalifah, T.: Travel Time computation with the linearized eikonal equation for anisotropic media, Geophys. Prospect., 50, 373-382, https://doi.org/10.1046/j.1365-2478.2002.00322.x, 2002.

Alkhalifah, T. and Plessix, R. É.: A recipe for practical full-waveform inversion in anisotropic media: An analytical parameter resolution study, Geophysics, 79, R91-R101, https://doi.org/10.1190/geo2013-0366.1, 2014.
455 Almqvist B. S. G. and Mainprice, D.: Seismic properties and anisotropy of the continental crust: Predictions based on mineral texture and rock microstructure, Rev. Geophys., 55, 367-433, https://doi.org/10.1002/2016RG000552, 2017.

Babuska, V. and Cara, M.: Seismic Anisotropy in the Earth, 1st edition, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.

Backus G. E.: Long-wave elastic anisotropy produced by horizontal layering. J. Geophys. Res., 67, 4427-4440, https://doi.org/10.1029/JZ067i011p04427, 1962.

Bakulin, A., Woodward, M., Osypov, K., Nichols, D., and Zdraveva, O.: Can we distinguish TTI and VTI media?, in: SEG Technical Program Expanded Abstracts 2009 of the 79th SEG Annual Meeting, Houston, USA, 25-30 October 2009, 226-230, https://doi.org/10.1190/1.3255313, 2009.

Bamford, D. and Crampin, S.: Seismic anisotropy – the state of the art, Geophys. J. R. astr. Soc., 49, 1-8, https://doi.org/10.1111/j.1365-246X.1977.tb03697.x, 1977.

Bezada, M. J., Faccenda, M., Toomey, D. R., and Humphreys, E.: Why ignoring anisotropy when imaging subduction zones could be a bad idea, in: AGU Fall Meeting Abstracts of the 47th AGU Fall Meeting, San Francisco, USA, 15-19 December 2014, S41D-04, 2014.

Bishop, T. N., Bube, K. P., Cutler, R. T., Langan, R. T., Love, P. L., Resnick, J. R., Shuey, R. T., Spindler, D. A., and Wyld, H. W.: Tomographic determination of velocity and depth in laterally varying media, Geophysics, 50, 903-923, https://doi.org/10.1190/1.1441970, 1985.

Crampin, S.: A review of wave motion in anisotropic and cracked elastic-media, Wave Motion, 3, 343-391, https://doi.org/10.1016/0165-2125(81)90026-3, 1981.

Cervený, V.: Seismic rays and ray intensities in inhomogeneous anisotropic media, Geophys. J. R. Astr. Soc., 29, 1-13, https://doi.org/10.1111/j.1365-246X.1972.tb06147.x, 1972.

Cervený, V.: Direct and inverse kinematic problems for inhomogeneous anisotropic media – linearized approach, Contr. Geophys. Inst. Slov. Aca. Sci., 13, 127-133, 1982.

Cervený, V. and Jech, J.: Linearized solutions of kinematic problems of seismic body waves in inhomogeneous slightly anisotropic media, J. Geophys., 51, 96-104, 1982.

Chapman C. H. and Pratt, R. G.: Traveltime tomography in anisotropic media – I. Theory, Geophys. J. Int., 109, 1-19, https://doi.org/10.1111/j.1365-246X.1992.tb00075.x, 1992.

Dunn, R. A., Toomey, D. R., Detrick, R. S., and Wilcock, W. S. D.: Continuous Mantle Melt Supply Beneath an Overlapping Spreading Center on the East Pacific Rise, Science, 291, 1955-1958, https://doi.org/10.1126/science.1057683, 2001.

Dunn, R. A., Lekic, V., Detrick, R. S., and Toomey, D. R.: Three-dimensional seismic structure of the Mid-Atlantic Ridge (35ºN): Evidence for focused melt supply and lower crustal dike injection, J. Geophys. Res., 110, https://doi.org/10.1029/2004JB003473, 2005.

Duveneck, E. and Bakker, P. M.: Stable P-wave modeling for reverse-time migration in tilted TI media, Geophysics, 76, S65–S75, https://doi.org/10.1190/1.3533964, 2011.

490 Eberhart-Phillips, D. and Henderson, M.: Including anisotropy in 3-D velocity inversion and application to Marlborough, New Zealand, Geophys. J. Int., 156, 237-254, https://doi.org/10.1111/j.1365-246X.2003.02044.x, 2004.

Eilon, Z., Abers, G. A., and Gaherty, J. B.: A joint inversion for shear velocity and anisotropy: the Woodlark Rift, Papua New Guinea, Geophys. J. Int., 206(1), 807-824, https://doi.org/10.1093/gji/ggw177, 2016.

Fowler, P.J., Du, X., and Fletcher, R.P.: Coupled equations for reverse-time migration in transversely isotropic media,
495 Geophysics, 75, S11–S22, https://doi.org/10.1190/1.3294572, 2010.

Giroux, B. and Gloaguen, E.: Geostatistical traveltime tomography in elliptically anisotropic media, Geophys. Prospect., 60(6), 1133-1149, https://doi.org/10.1111/j.1365-2478.2011.01047.x, 2012.

Gholami, Y., Brossier, R., Operto, S., Ribodetti, A., and Virieux, J.: Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 1: Sensitivity and trade-off analysis, Geophysics, 78, R81-R105,
500 https://doi.org/10.1190/geo2012-0204.1, 2013.

Gholami, Y., Brossier, R., Operto, S., Prieux, V., Ribodetti, A., and Virieux, J.: Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 2: Synthetic and real data case studies from Valhall. Geophysics, 78, R107-R124, https://doi.org/10.1190/geo2012-0203.1, 2013.

Helbig, K.: Elliptical anisotropy—Its significance and meaning, Geophysics, 48, 825-832, https://doi.org/10.1190/1.1441514,
505 1983.

Holtzman, B. K., Kohlstedt, D. L., Zimmerman, M. E., Heidelbach, F., Hiraga, T., and Hustoft, J.: Melt segregation and strain partitioning: implications for seismic anisotropy and mantle flow, Science, 301, 1227-1230, https://doi.org/10.1126/science.1087132, 2003.

Ismaïl, W. B. and Mainprice, D.: An olivine fabric database: an overview of upper mantle fabrics and seismic anisotropy,
510 Tectonophysics, 296, 145-157, https://doi.org/10.1016/S0040-1951(98)00141-3, 1998.

Jech, J. and Psencik, I.: First-order perturbation method for anisotropic media, Geophys. J. Int., 99, 369-376, https://doi.org/10.1111/j.1365-246X.1989.tb01694.x, 1989.

Johnston, J. E. and Christensen, N. I.: Seismic anisotropy of shales, J. Geophys. Res., 100, 5991-6003, https://doi.org/10.1029/95JB00031, 1995.

515 Kendall, J. M., Stuart, G. W., Ebinger, C. J., Bastow, I. D., and Keir, D.:. Magma-assisted rifting in Ethiopia, Nature, 433, 146–148, https://doi.org/10.1038/nature03161, 2005.

Kendall, J. M., Pilidou, S., Keir, D., Bastow, I. D., Stuart, G. W., and Ayele, A.: Mantle upwellings, melt migration and the rifting of Africa: insights from seismic anisotropy, Geol. Soc. Lond. Spec. Publ., 259(1), 55–72, https://doi.org/10.1144/GSL.SP.2006.259.01.06, 2006.

520 Kraut, E. A.: Advances in the Theory of Anisotropic Elastic Wave Propagation, Rev. Geophys., 1(3), 401-448, https://doi.org/10.1029/RG001i003p00401, 1963.

Long, M. D. and Silver, P. G.: The Subduction Zone Flow Field from Seismic Anisotropy: A Global View, Science, 319, 315-318, https://doi.org/10.1126/science.1150809, 2008.

Long, M. D. and Becker, T. W.: Mantle dynamics and seismic anisotropy, Earth Planet. Sci. Lett., 297, 341-354, https://doi.org/10.1016/j.epsl.2010.06.036, 2010.

Maultzsch, S., Chapman, M., Liu, E., and Li, X. Y.: Modelling frequency-dependent seismic anisotropy in fluid-saturated rock with aligned fractures: implication of fracture size estimation from anisotropic measurements, Geophys. Prospect., 51(5), 381-392, https://doi.org/10.1046/j.1365-2478.2003.00386.x, 2003.

McCann, C., Assefa, S., Sothcott, J., McCann, D. M., and Jackson, P. D.: In-situ borehole measurements of compressional and shear wave attenuation in Oxford clay, Sci. Drilling, 1, 11-20, 1989.

Meléndez, A.: Development of a New Parallel Code for 3-D Joint Refraction and Reflection Travel-Time Tomography of Wide-Angle Seismic Data – Synthetic and Real Data Applications to the Study of Subduction Zones, PhD thesis, Universitat de Barcelona, Barcelona, Spain, pp. 171, http://hdl.handle.net/2445/65200, 2014.

Meléndez, A., Sallarès, V., Ranero, C. R. and Kormann, J.: Origin of water layer multiple phases with anomalously high amplitude in near-seafloor wide-angle seismic recordings, Geophys. J. Int., 196, 243-252, https://doi.org/10.1093/gji/ggt391, 2014.

Meléndez, A., Korenaga, J., Sallarès, V., Miniussi, A., and Ranero, C. R.: TOMO3D: 3-D joint refraction and reflection traveltime tomography parallel code for active-source seismic data – synthetic test, Geophys. J. Int., 203, 158-174, https://doi.org/10.1093/gji/ggv292, 2015.

Montagner, J. P., Marty, B., and Stutzmann, E.: Mantle upwellings and convective instabilities revealed by seismic tomography and helium isotope geochemistry beneath eastern Africa, Geophys. Res. Lett., 34, L21303, https://doi.org/10.1029/2007GL031098, 2007.

Moser, T. J.: Shortest path calculation of seismic rays, Geophysics, 56, 59-67, https://doi.org/10.1190/1.1442958, 1991.

Moser, T. J., Nolet, G., Snieder, R.: Ray bending revisited, B. Seismol. Soc. Am., 82, 259-288, 1992.

Nicolas, A. and Christensen, N. I.: Formation of anisotropy in upper mantle peridotites-A review. Composition, structure and dynamics of the lithosphere-asthenosphere system, 16, 111-123, https://doi.org/10.1029/GD016p0111, 1987.

Nye, J. F.: Physical Properties of Crystals: their representation by tensors and matrices, 1st edition, Oxford University Press, Oxford, England, 1957.

Pratt, R. G. and Chapman C. H.: Traveltime tomography in anisotropic media – II. Application, Geophys. J. Int., 109, 20-37, https://doi.org/10.1111/j.1365-246X.1992.tb00076.x, 1992.

Prieux, V., Brossier, R., Gholami, Y., Operto, S., Virieux, J., Barkved, O. I., and Kommedal, J. H.: On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study. Geophysical Journal International, 187, 1495-1515, https://doi.org/10.1111/j.1365-246X.2011.05209.x, 2011.

Rüger, A. and Alkhalifah, T.: Efficient two-dimensional anisotropic ray tracing, in: Seismic Anisotropy, edited by: Erling, F., Rune, M. H., and Jaswant, S. R., Soc. of Expl. Geophys., 556-600, https://doi.org/10.1190/1.9781560802693.ch18, 1996.

Sallarès, V., Meléndez, A., Prada, M., Ranero, C. R., McIntosh, K., and Grevemeyer, I.: Overriding plate structure of the Nicaragua convergent margin: Relationship to the seismogenic zone of the 1992 tsunami earthquake, Geochem. Geophys. Geosyst., 14, 3436–3461, https://doi.org/10.1002/ggge.20214, 2013.

Sayers, C. M.: Seismic anisotropy of shales. Geophys. Prospect., 53, 667-676, https://doi.org/10.1111/j.1365-2478.2005.00495.x, 2005.

Schulte-Pelkum, V. and Mahan, K. H.: A method for mapping crustal deformation and anisotropy with receiver functions and first results from USArray. Earth Planet. Sci. Lett., 402, 221-233, https://doi.org/10.1016/j.epsl.2014.01.050, 2014.

Song, L., Zhang, S., Liu, H., Chun, S., and Song, Z.: A formalism for acoustical traveltime tomography in heterogenous anisotropic media, in: Review of Progress in Quantitative Nondestructive Evaluation, edited by: Thompson, D. O. and Chimenti, D. E., Springer, Boston, USA, 1529-1536, https://doi.org/10.1007/978-1-4615-5339-7_198, 1998.

Stewart, R. R.: An algebraic reconstruction technique for weakly anisotropic velocity, Geophysics, 53, 1613-1615, https://doi.org/10.1190/1.1442445, 1988.

Thomsen, L.: Weak elastic anisotropy, Geophysics, 51, 1954-1966, https://doi.org/10.1190/1.1442051, 1986.

Thomsen, L. and Anderson, D. L.: Weak elastic anisotropy in global seismology. Geol. Soc. Spec. Pap., 514, 39-50, https://doi.org/10.1130/2015.2514(04), 2015.

Tsvankin, I.: P-wave signatures and notation for transversely isotropic media: An overview, Geophysics, 61(2), 467-483, https://doi.org/10.1190/1.1443974, 1996.

Vauchez, A., Tommasi, A., Barruol, G., and Maumus, J.: Upper mantle deformation and seismic anisotropy in continental rifts, Phys. Chem. Earth A: Solid Earth Geod., 25(2), 111–117, https://doi.org/10.1016/S1464-1895(00)00019-3, 2000.

Vinnik, L. P., Makeyeva, L. I., Milev, A., and Usenko, A. Y.: Global patterns of azimuthal anisotropy and deformations in the continental mantle. Geophys. J. Int., 111(3), 433-447, https://doi.org/10.1111/j.1365-246X.1992.tb02102.x, 1992.

Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Štekl, I., Guasch, L., Win, C., Conroy, G., and Bertrand, A.: Anisotropic 3D full-waveform inversion. Geophysics, 78(2), R59-R80, https://doi.org/10.1190/geo2012-0338.1, 2013.

Yousef B. M. and Angus, D. A.: When do fractured media become seismically anisotropic? Some implications on quantifying fracture properties, Earth Planet. Sci. Lett., 444, 150-159, https://doi.org/10.1016/j.epsl.2016.03.040, 2016.

Zhou, B. and Greenhalgh, S.: 'Shortest path' ray tracing for most general 2D/3D anisotropic media, J. Geophys. Eng., 2, 54-63, https://doi.org/10.1088/1742-2132/2/1/008, 2005.

Zhou, B. and Greenhalgh, S.: Non-linear traveltime inversion for 3-D seismic tomography in strongly anisotropic media, Geophys. J. Int., 172, 383-394, https://doi.org/10.1111/j.1365-246X.2007.03649.x, 2008.

Tables

Sensitivity test for	Model area	P[ε]			P[$v^\perp$]		
		v	δ	ε	v	δ	$v^\perp$
v in P[ε]	Anomaly	2.5	0.16	0.16	2.5	0.16	2.9
	Background	2	0.16	0.16	2	0.16	2.32
v in P[$v^\perp$]	Anomaly	2.5	0.16	-0.072	2.5	0.16	2.32
	Background	2	0.16	0.16	2	0.16	2.32
δ	Anomaly	2	0.2	0.16	2	0.2	2.32
	Background	2	0.16	0.16	2	0.16	2.32
ε (and $v^\perp$)	Anomaly	2	0.16	0.2	2	0.16	2.4
	Background	2	0.16	0.16	2	0.16	2.32

Table 1: Values for anomaly and background areas of the models used in each of the four sensitivity tests (white cells), and their equivalences in the alternative parametrization (gray cells) (Fig. 2). Accuracy tests in Fig. 3 are conducted for these same four cases (white cells). A comparison with the accuracies achieved using the alternative parametrizations can be seen in Fig. S2.

Anomaly in	v in P[ε]		v in P[$v^\perp$]		δ		ε	
Azimuth	0 rad	π/4 rad	0 rad	π/4 rad	0 rad	π/4 rad	0 rad	π/4 rad
Mean relative error ± mean deviation (%)	0.8 ± 0.1	0.6 ± 0.1	0.6 ± 0.2	0.5 ± 0.2	0.01 ± 0.01	0.01 ± 0.01	0.04 ± 0.05	0.03 ± 0.04
	Overall values							
	0.7 ± 0.1		0.5 ± 0.2		0.017 ± 0.009		0.04 ± 0.03	

Table 2: Mean relative traveltime errors in percentage and their mean deviations for the two selected meridians in Fig. 3, and for the entire set of 482 source – receiver pairs. Compared to Fig. 2a,b, these average traveltime error values indicate that the code is sufficiently accurate to model the traveltime residuals arising from the inclusion of the selected anomalies.

	v (km/s)	δ	ε	$v^{\perp}$ (km/s)
Background value	2	0.16	0.16	2.32
Anomaly value	2.5	0.2	0.2	3

Table 3: For inversion tests, background and anomaly values of all four parameters for the initial/background model and for the anomaly in the target model. The model is a cube of edge 5 km. The anomaly is a discretized sphere of 1 km in diameter at the center of the cube.

	Residuals RMS (ms)	Mean relative differences (%)											
		v			δ			ε			$v^{\perp}$		
		BG	AI	AT	BG	AI	AT	BG	AI	AT	BG	AI	AT
P[ε]	30 – 0.4	0.5	21.0	3.3	4.8	22.3	15.2	1.6	11.1	11.2	0.5	22.8	5.0
P[$v^{\perp}$]	30 – 0.5	0.8	25.9	1.9	5.0	20.8	29.1	5.8	26.7	41.0	0.6	21.3	6.2

Table 4: Quantification of the quality of recovery in terms of data and model fit for the simultaneous inversion of both parametrizations (Figs 5 and 6). Initial and final RMS values for traveltime residuals to quantify data fit. For each parameter, as a measure of model fit, we computed the mean relative differences between the background areas of the inverted and target models (BG), the anomaly areas of the inverted and initial models (AI), and the anomaly areas of the inverted and target models (AT). Since the initial model is equal to the target model in the background area, it is not necessary to calculate the difference between inverted and initial models for this area. The ideal difference value for BG is 0 %. AT and AI indicate the resemblance between true and recovered anomalies, and their ideal values are 0 % and 25 % (~29.3 % in the case of $v^{\perp}$) respectively. Recovery is consistent with sensitivity (Fig. 2): v and $v^{\perp}$ are well retrieved, ε is only partially recovered with P[ε], whilst inversion of δ is unsuccessful in both cases.

P[ε]	Residuals RMS (ms)	Mean relative differences (%)											
		v			δ			ε			$v^{\perp}$		
		BG	AI	AT	BG	AI	AT	BG	AI	AT	BG	AI	AT
Step 1	30 – 0.5	0.5	21.4	2.9	-	-	-	1.3	8.2	13.5	0.5	22.8	5.1
Step 2	0.5 – 0.3	0.5	21.4	3.0	0.9	5.3	19.4	-	-	-	0.5	22.8	5.1

Table 5: Same as Table 4 but for the two-step sequential inversion of P[ε] (Fig. 7). Recoveries in terms of model fit are virtually identical to those for the simultaneous inversion of this parametrization, with ε recovery being just slightly better. Final data fit is also better than the one achieved by simultaneous inversion of P[ε].

P[$v^\perp$]	Residuals RMS (ms)	Mean relative differences (%)											
		v			δ			ε			$v^\perp$		
		BG	AI	AT	BG	AI	AT	BG	AI	AT	BG	AI	AT
Step 1	30 – 0.7	0.7	26.7	2.7	-	-	-	5.4	33.0	45.7	0.5	21.0	6.4
Step 2	0.7 – 0.5	0.7	26.2	2.5	1.8	17.4	14.9	5.7	29.8	41.8	0.5	21.5	6.1

Table 6: Same as Table 5 but for the two-step sequential inversion of P[$v^\perp$] (Fig. 8). Model fits are similar to those achieved by the simultaneous inversion of this parametrization: v and $v^\perp$ are well retrieved, while recovery of δ and ε is unsuccessful. Final data fit is identical to the one obtained by simultaneously inverting for P[$v^\perp$].

	Residuals RMS (ms)	Mean relative differences (%)											
		v			δ			ε			$v^\perp$		
		BG	AI	AT	BG	AI	AT	BG	AI	AT	BG	AI	AT
P[ε]	19 – 0.1	0.1	25.2	0.3	0.8	15.4	7.7	0.3	24.4	0.5	0.1	29.4	0.3
P[$v^\perp$]	21 – 0.4	0.2	24.2	1.1	1.4	30.6	9.9	1.2	31.9	7.4	0.1	29.7	0.5

Table 7: Same as Table 4 but for the simultaneous inversion of P[ε] and P[$v^\perp$] using v and ε or $v^\perp$ target models as initial models (Figs 9 and 10). Model fits for v and ε or $v^\perp$ are close to perfect as expected, but even so the recovery of δ is partial at most. Data misfit is smaller than for the original inversions in Figs 5 and 6. P[ε] yields a better result according to all indicators, and as for all previous tests, the recovery of $v^\perp$ from P[ε] is notably better than that of ε from P[$v^\perp$].

	Residuals RMS (ms)	Mean relative differences (%)											
		v			δ			ε			$v^\perp$		
		BG	AI	AT	BG	AI	AT	BG	AI	AT	BG	AI	AT
P[ε]	61 – 0.8	3.3	22.0	3.1	-	-	-	14.6	28.6	5.2	2.6	26.8	3.2
P[$v^\perp$]	61 – 0.8	3.5	21.1	3.3	-	-	-	24.7	40.7	15.2	2.0	27.8	2.4

Table 8: Same as Table 4 but for the simultaneous inversion of P[ε] and P[$v^\perp$] following Eq. (11), i.e. neglecting δ in Eqs (1) and (3). Model differences for v, ε, and $v^\perp$, as well as data misfit are all significantly worse than for any of the previous tests.

δ value for initial model	P[ε]		P[$v^\perp$]	
	Initial and final traveltime residuals RMS (ms)	Mean ± mean deviation of inverted δ model	Initial and final traveltime residuals RMS (ms)	Mean ± mean deviation of inverted δ model
0.1	39 – 11	0.13 ± 0.05	39 – 5	0.13 ± 0.04
0.11	37 – 8	0.13 ± 0.04	37 – 4	0.14 ± 0.03
0.12	35– 4	0.15 ± 0.03	35 – 0.5	0.14 ± 0.02
0.13	34 – 0.5	0.15 ± 0.01	34 – 0.5	0.14 ± 0.02
0.14	32 – 0.4	0.15 ± 0.01	32 – 0.5	0.15 ± 0.01
0.15	31 – 0.4	0.16 ± 0.01	31 – 0.5	0.16 ± 0.01
0.16	30 – 0.4	0.160 ± 0.008	30 – 0.5	0.160 ± 0.008
0.17	28 – 0.4	0.16 ± 0.01	29 – 0.6	0.17 ± 0.01
0.18	27 – 0.4	0.17 ± 0.01	28 – 0.5	0.17 ± 0.01
0.19	27 – 0.5	0.17 ± 0.01	27 – 0.5	0.18 ± 0.01
0.2	27 – 0.5	0.18 ± 0.02	27 – 0.5	0.18 ± 0.02
0.21	27 – 0.5	0.18 ± 0.02	27 – 0.5	0.19 ± 0.02
0.22	27 – 0.5	0.19 ± 0.02	27 – 0.5	0.19 ± 0.02
0.23	27 – 5	0.20 ± 0.04	27 – 2	0.20 ± 0.03
0.24	28 – 11	0.22 ± 0.06	28 – 5	0.20 ± 0.04

Table 9: Results of the procedure to approximate an initial δ model. The RMS of the final traveltime residuals shows a clear change in order of magnitude in the subranges (0.13,0.22) and (0.12,0.22) depending on the parameterization. Results for initial δ = 0.16 correspond to the examples in Figs 5 and 6.

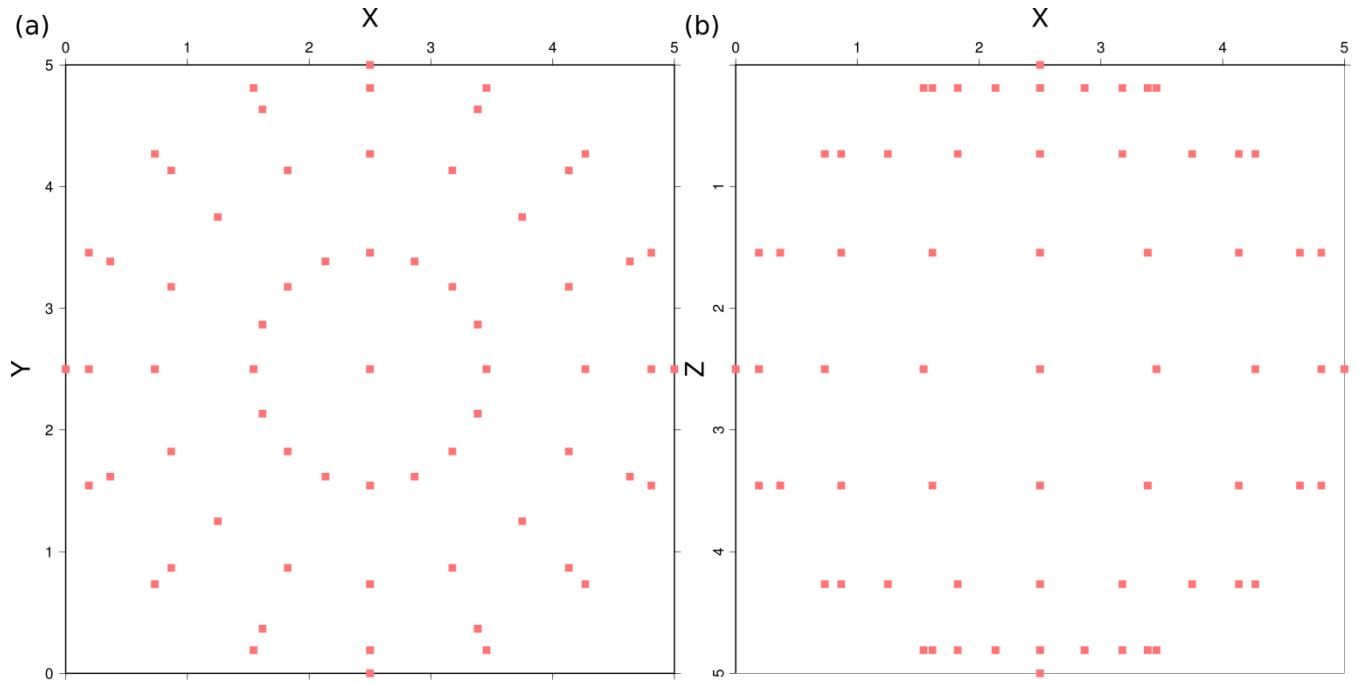
625

Figures

Figure 1: (a) Horizontal and (b) vertical views of the acquisition geometry for the inversion tests. 114 sources and 114 receivers (red boxes) are located at 2.5 km from the center of the model, at the locus defined by the surface of the sphere inscribed in the cube, and placed at the crossing points of 16 meridians with 7 parallels, and at each pole. Thus, in the inversion tests we used 12,882 traveltimes from 114 sources each recorded at 113 receivers, i.e. all receivers record all sources, except for the one coinciding in location. The acquisition geometry for the accuracy and sensitivity tests is similar, only in this case for sources and receivers at the crossing points of 32 meridians and 15 parallels, plus one of each at the two poles, and using just 482 traveltimes from diametrically-opposed source–receiver pairs arranged, i.e. each receiver exclusively records the first arrival traveltime from its paired source.

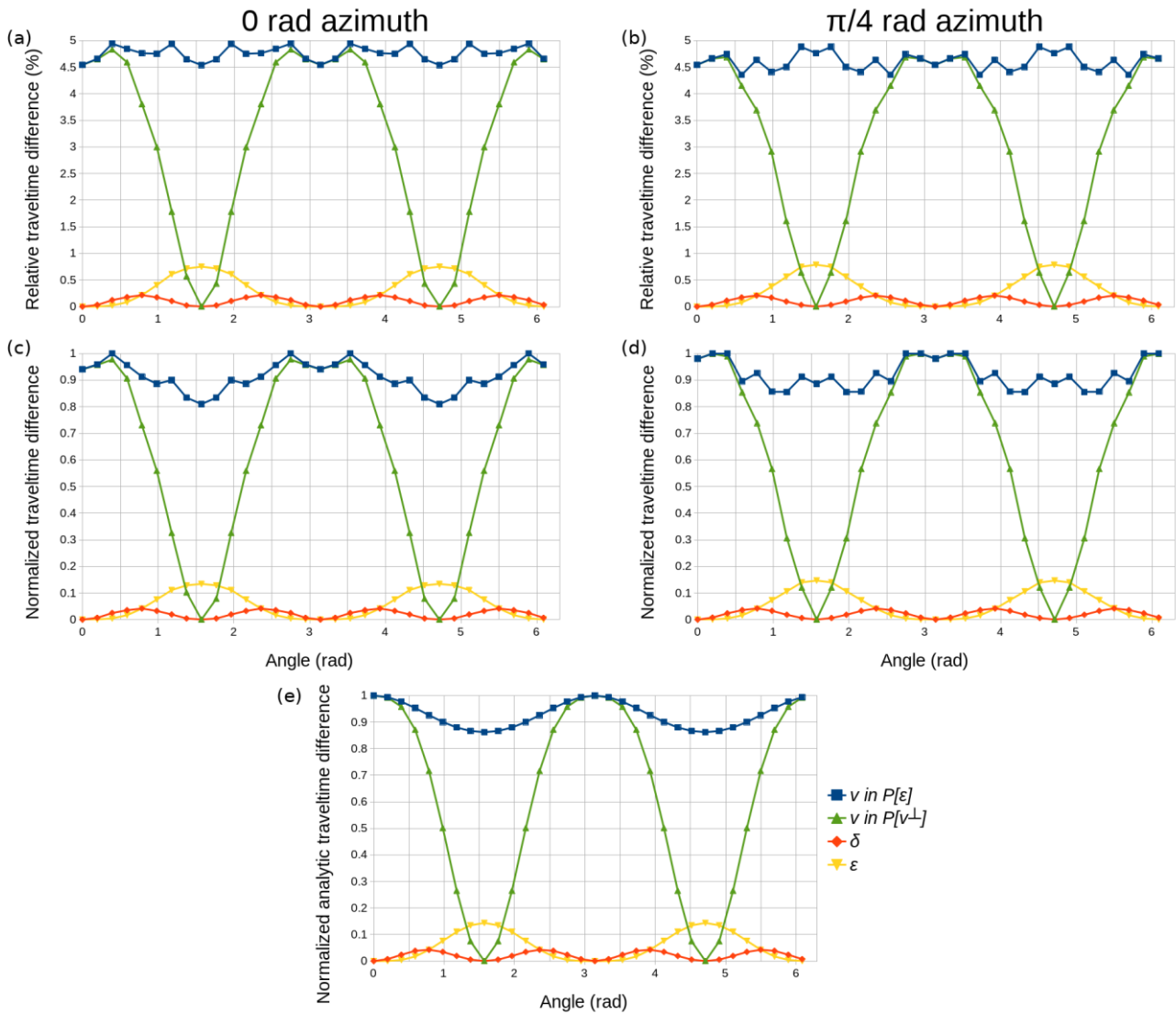



Figure 2: Sensitivities for the meridians at azimuths 0 rad (left) and π/4 rad (right) as a function of the polar angle (origin in the downward vertical axis). (a) and (b) Synthetic relative sensitivities in percentage, (c) and (d) synthetic normalized sensitivities, and (e) normalized analytic sensitivities. Sensitivity values displayed correspond to all source – receiver pairs along the selected meridians of the acquisition configuration.

640

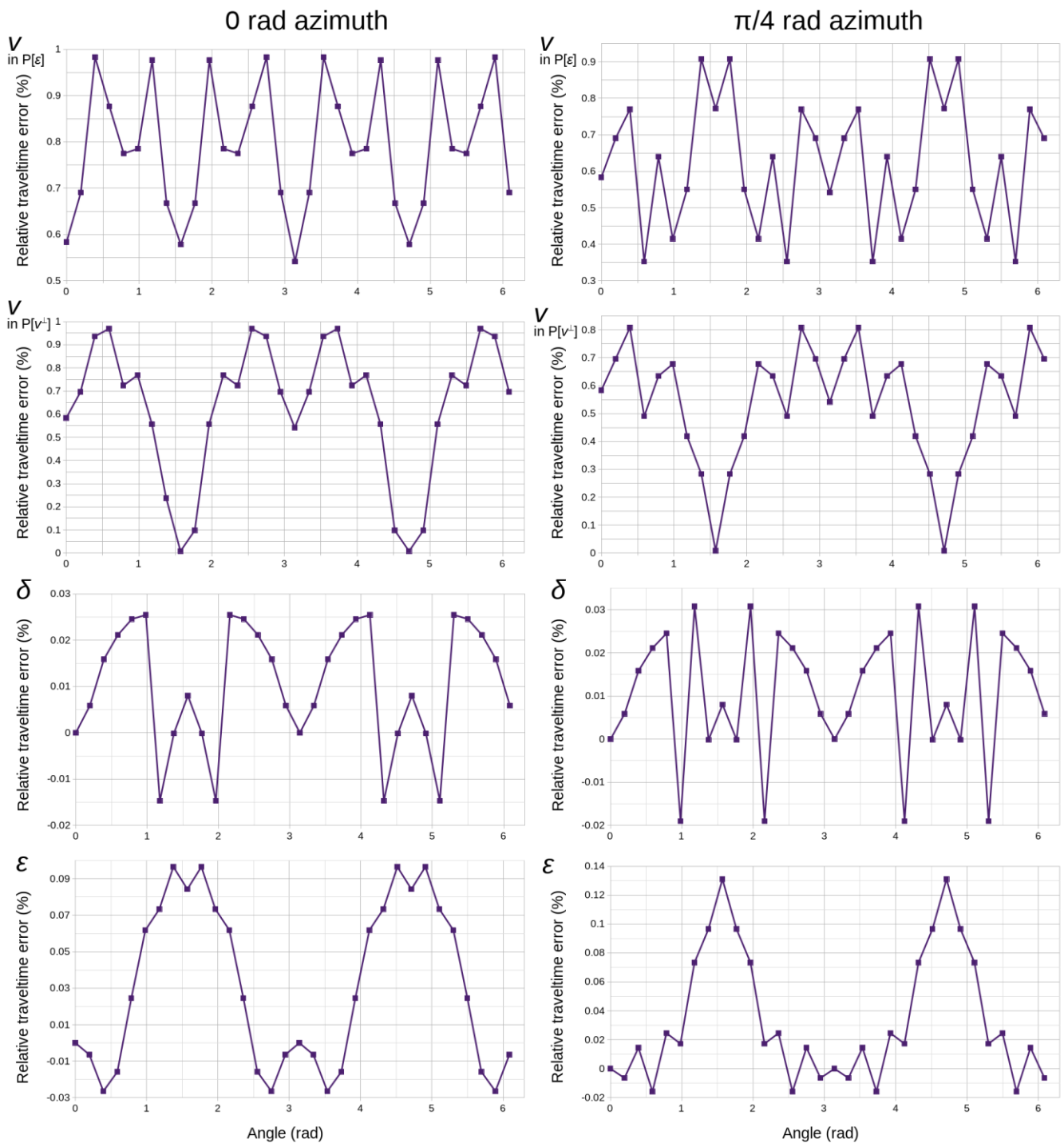



Figure 3: For both selected meridians, 0 rad and π/4 rad azimuths, relative traveltime errors in percentage with respect to the analytic value for each of the four simulations used in the sensitivity analysis. Polar angle origin is in the downward vertical axis. Mean values and deviations are shown in Table 2.

Figure 4: Flowchart describing the two steps for all the sequential inversion options tested. The best options for the sequential inversion strategy in P[ε] (Fig. 7) and in P[$v^\perp$] (Fig. 8) are marked in green and red respectively.

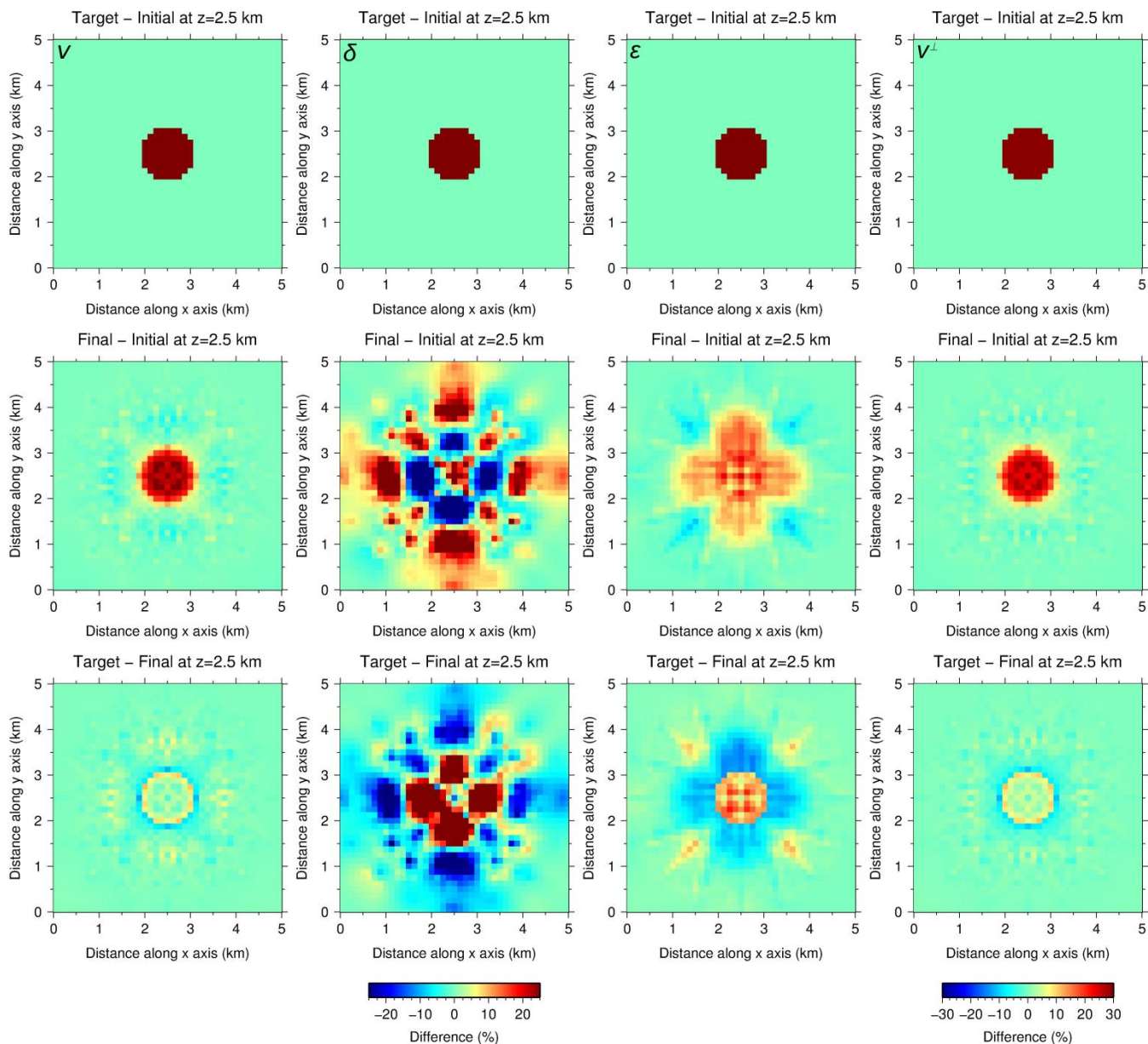



Figure 5: Simultaneous inversion with P[ε]. Horizontal slices of the relative differences between target and initial (first row), final and initial (second row), and target and final (third row) models at 2.5-km depth for the four parameters. $v^\perp$ is derived from Eq. (2). The range of the color scale for $v^\perp$ is wider than for the rest of parameters because the heterogeneity is calculated considering the 25 % anomalies in v and ε, which yields a ~29.3 % anomaly in $v^\perp$. First and second row would be identical if the inversion were perfect, whereas the third row would display a homogeneous value of 0 %. The quality of the recovery of each parameter is in correlation with their sensitivities (Fig. 2). Recovery of v is satisfactory with anomaly values close to the target and well-defined anomaly boundaries. ε recovery is partial, the anomaly is centered but its magnitude and shape are not as accurate as in the case of both velocities; even so it allows for a successful recovery of $v^\perp$ through Eq. (2), both in anomaly magnitude and shape. As for δ, recovery is unsuccessful.

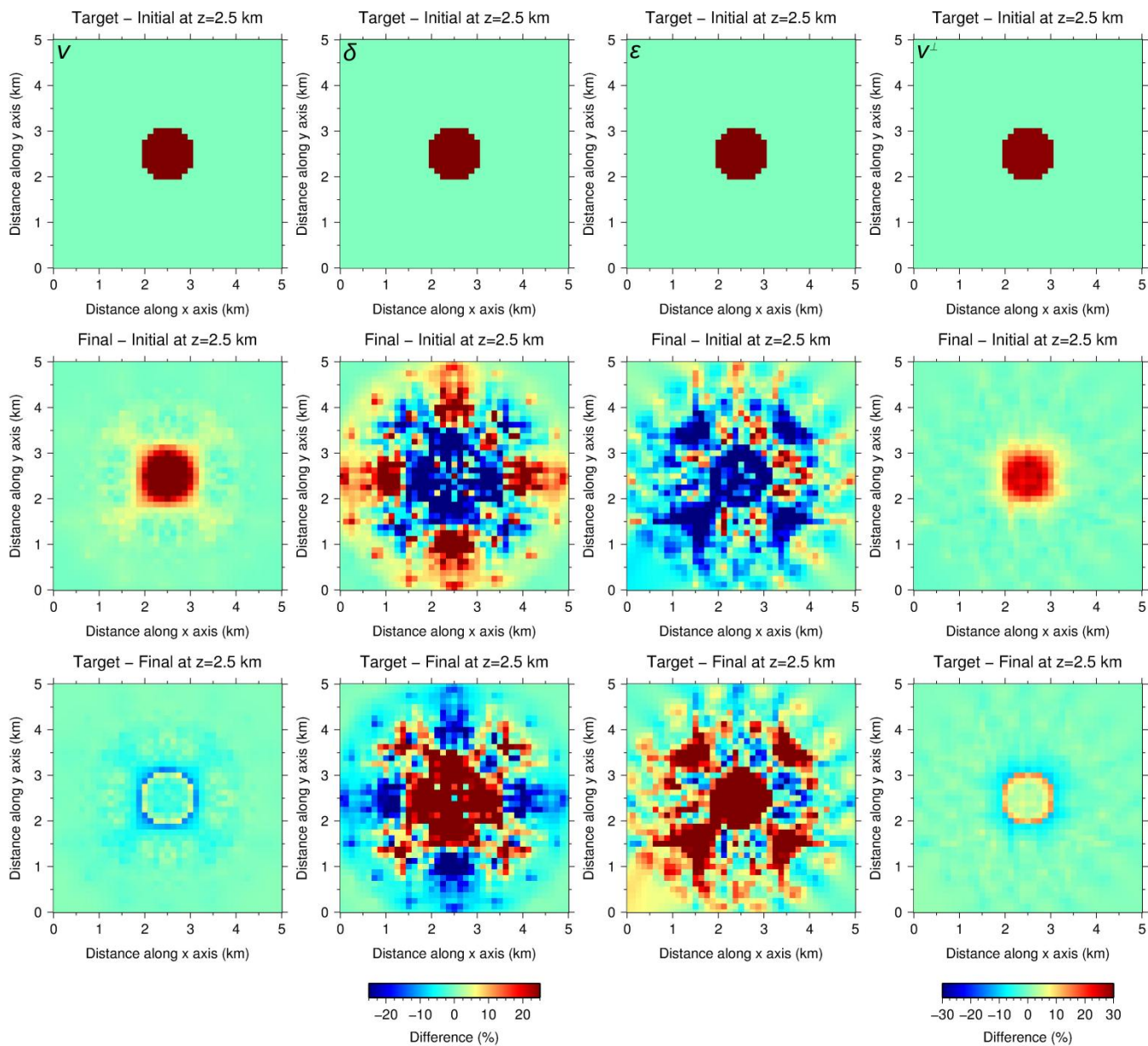



Figure 6: Same as Fig. 5 but with P[$v^\perp$]. ε is derived from Eq. (2). The quality of the recovery is in correlation with sensitivity (Fig. 2). Both velocities are satisfactorily recovered. The magnitude of the anomaly in v is better recovered than for P[ε], whereas the opposite occurs for $v^\perp$. Anomaly boundaries for both velocities are not as well determined as for P[ε]. ε and δ are not recovered.

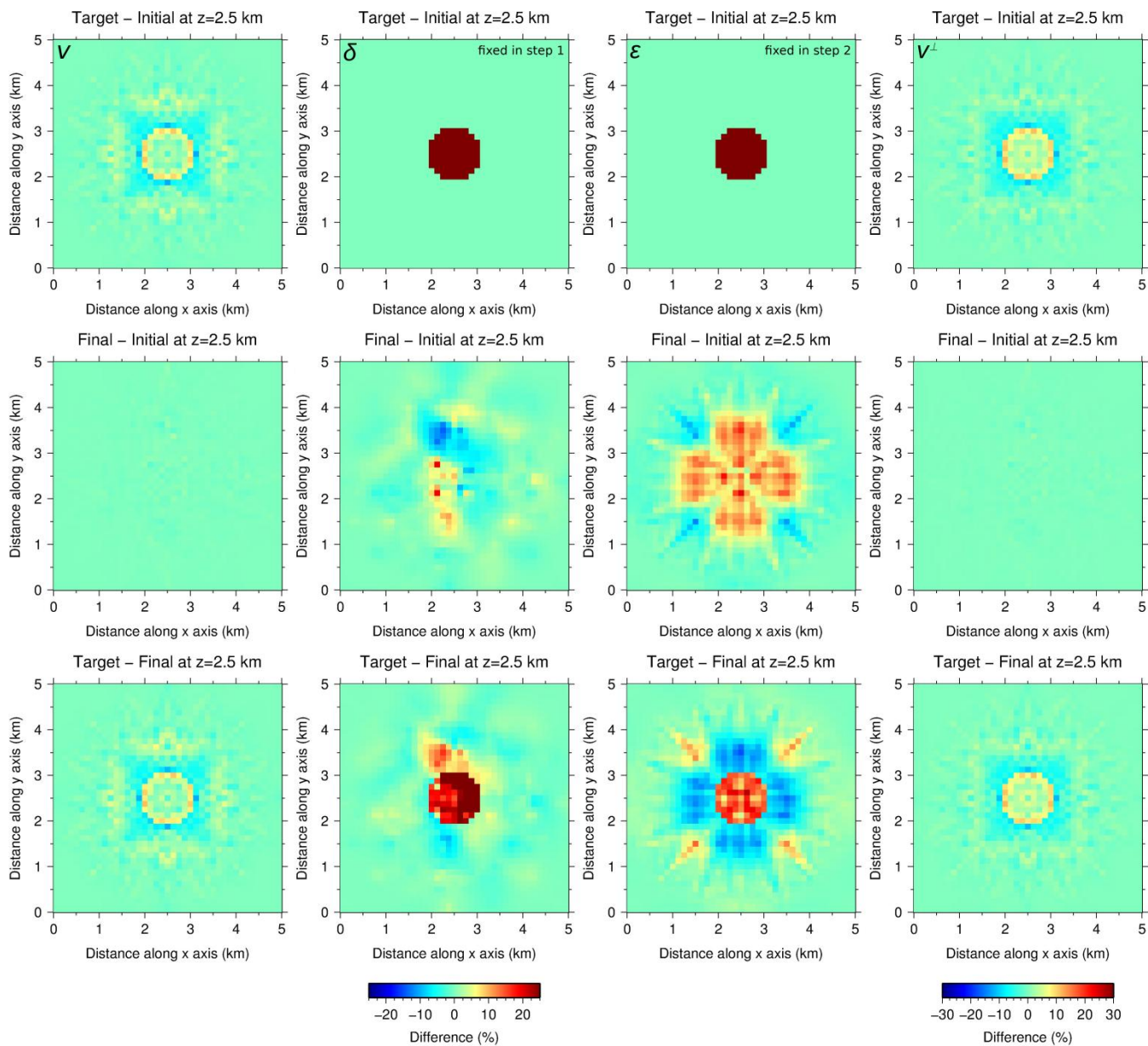
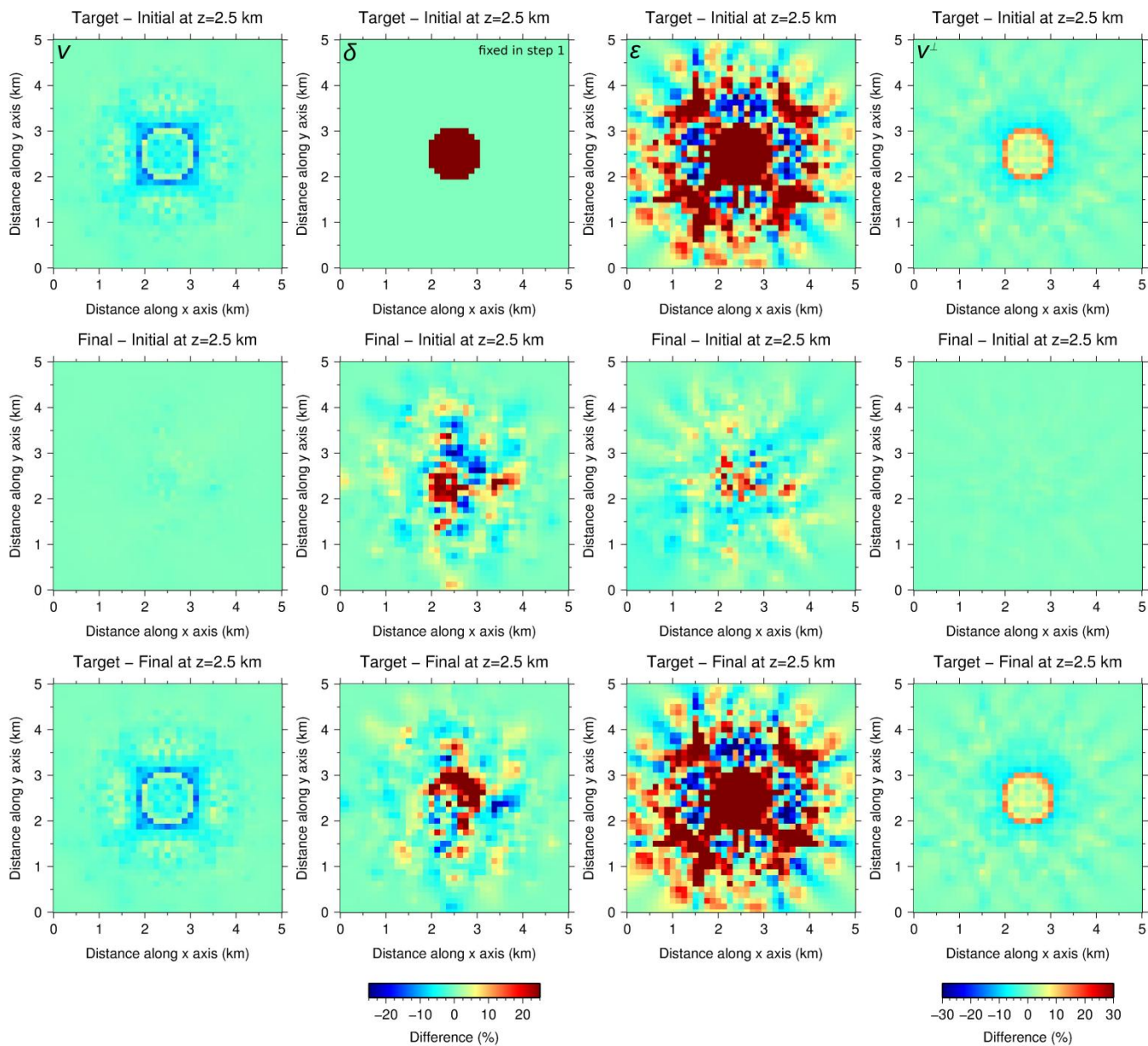



Figure 7: Same as Fig. 5 but for the two-step sequential inversion strategy with P[ε]. In the first step only δ was fixed. In the second step, only ε was fixed. Final models from step 1 were used as initial model for step 2. v (1st column) is well recovered in step 1 (top panel), and it is barely modified by the second step (middle and bottom panels). δ is fixed to the initial homogeneous model in step 1, and its recovery is unsuccessful in step 2. Recovery of ε is limited compared to v, but significantly better than that of δ. Nonetheless, it proves to be good enough to provide a satisfactory recovery of $v^{\perp}$ using Eq. (2).

670   **Figure 8: Same as Fig. 7 but for the two-step sequential inversion strategy with P[$v^\perp$]. In the first step only δ is fixed, whereas in the second one all parameters are inverted. v (1st column) and $v^\perp$ (4th column) are well recovered in step 1 (top panel), and they are barely modified by the second step (middle and bottom panels). δ and ε are not properly recovered; in both cases some sort of irregular perturbations approximately centered in the cube are retrieved but bearing no resemblance to the target anomalies.**

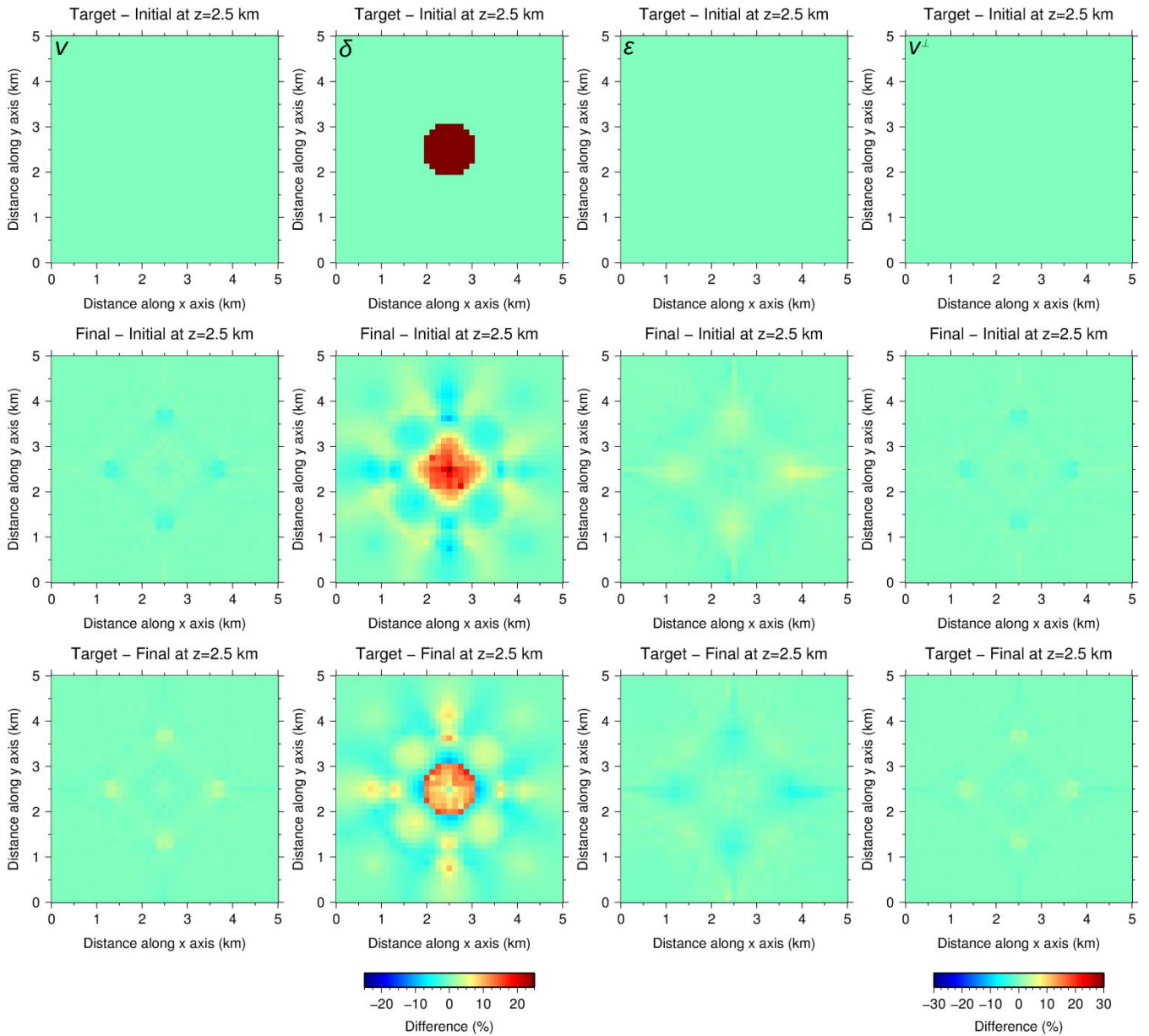



Figure 9: Same as Fig. 5 but using target models as initial models for all parameters in P[ε] but δ. This test was conducted to study the recovery of δ under unrealistically optimal circumstances, and even with these perfect initial conditions, recovery is, at best, extremely complicated due to the small sensitivity (Fig. 2); magnitude and shape are only partially recovered. In the first row, differences for v and ε are 0 % since we use target models as initial ones. For this same reason, the second and third rows show that differences between target and inverted for these three parameters are hardly observable, indicating that inversion is not modifying v and ε even though they are not fixed. Consequently, the resulting recovery of $v^{\perp}$ through Eq. (2) is almost perfect as well.

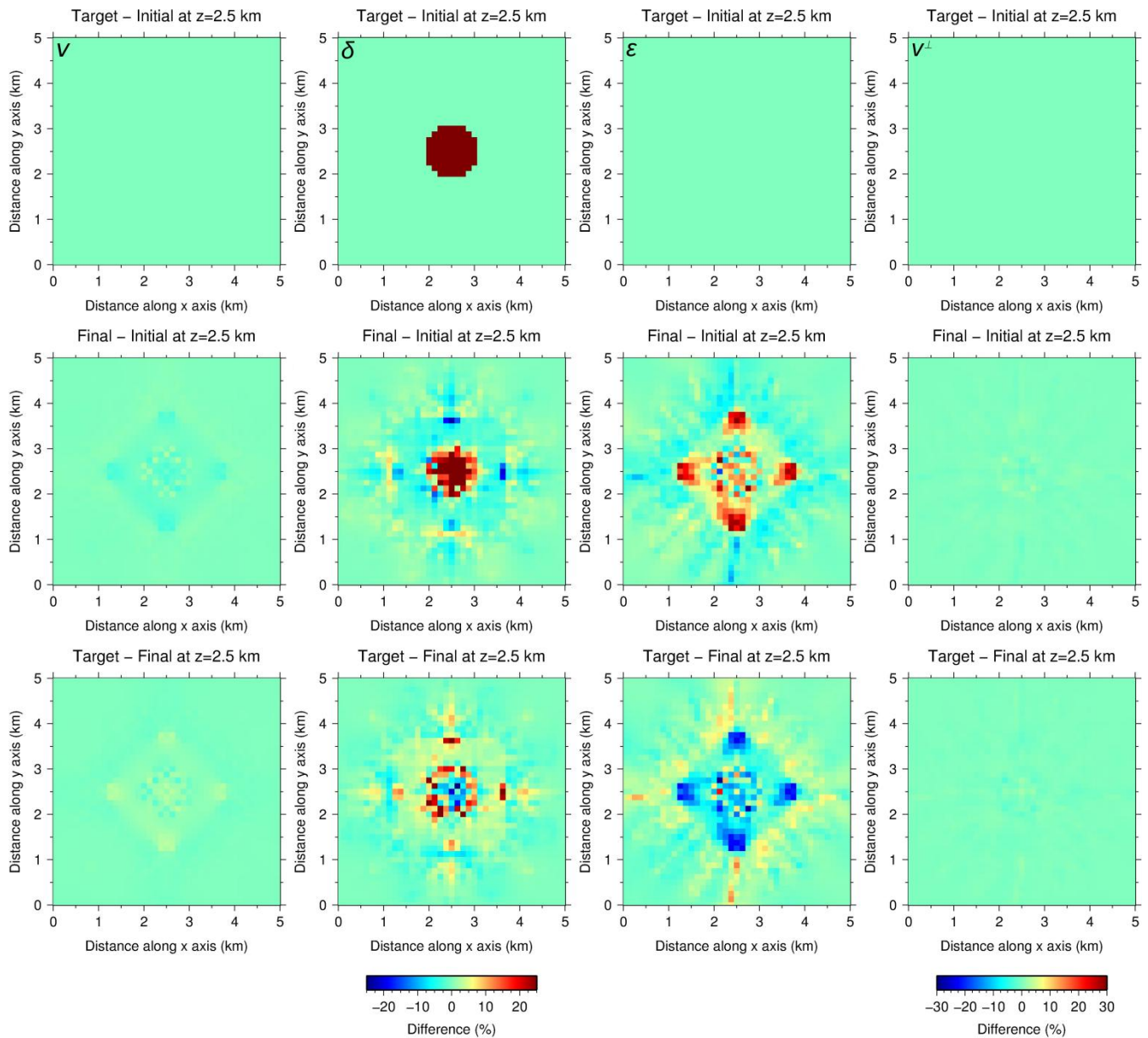



Figure 10: Same as Fig. 9 but using target models as initial models for all parameters in P[$v^{\perp}$] except for δ. The magnitude of the δ anomaly is better recovered than for P[ε], but the shape is not as well retrieved, and artifacts appear in the background area. Differences between inverted and target v and $v^{\perp}$ models are still hardly observable but not as much as for v and ε in the case of P[ε], and thus the recovery of ε through Eq. (2) is also not as good.