Waveform-based simulated annealing of crosshole transmission data: a semi-global method for estimating seismic anisotropy

Michael V. Afanasiev,[1,2] R. Gerhard Pratt,[1] Rie Kamei[1,3] and Glenn McDowell[4]

[1]*Department of Earth Sciences, University of Western Ontario, London, ON, Canada* N6A5B7
[2]*Institute of Geophysics, ETH Zürich, Zürich 8053, Switzerland. E-mail: michael.afanasiev@erdw.ethz.ch*
[3]*School of Earth and Environment, University of Western Australia, Crawley, WA 6009, Australia*
[4]*Vale Technology Development (Canada) Ltd., Sudbury, Ontario, Canada*

SUMMARY

We successfully apply the semi-global inverse method of simulated annealing to determine the best-fitting 1-D anisotropy model for use in acoustic frequency domain waveform tomography. Our forward problem is based on a numerical solution of the frequency domain acoustic wave equation, and we minimize wavefield phase residuals through random perturbations to a 1-D vertically varying anisotropy profile. Both real and synthetic examples are presented in order to demonstrate and validate the approach.

For the real data example, we processed and inverted a cross-borehole data set acquired by Vale Technology Development (Canada) Ltd. in the Eastern Deeps deposit, located in Voisey's Bay, Labrador, Canada. The inversion workflow comprises the full suite of acquisition, data processing, starting model building through traveltime tomography, simulated annealing and finally waveform tomography.

Waveform tomography is a high resolution method that requires an accurate starting model. A cycle-skipping issue observed in our initial starting model was hypothesized to be due to an erroneous anisotropy model from traveltime tomography. This motivated the use of simulated annealing as a semi-global method for anisotropy estimation. We initially tested the simulated annealing approach on a synthetic data set based on the Voisey's Bay environment; these tests were successful and led to the application of the simulated annealing approach to the real data set. Similar behaviour was observed in the anisotropy models obtained through traveltime tomography in both the real and synthetic data sets, where simulated annealing produced an anisotropy model which solved the cycle-skipping issue. In the real data example, simulated annealing led to a final model that compares well with the velocities independently estimated from borehole logs. By comparing the calculated ray paths and wave paths, we attributed the failure of anisotropic traveltime tomography to the breakdown of the ray-theoretical approximation in the vicinity of strong velocity discontinuities.

Key words: Inverse theory; Controlled source seismology; Seismic anisotropy; Seismic tomography; Computational seismology; Wave propagation.

1 INTRODUCTION

Seismic tomography is a versatile subsurface imaging method appropriate for use with transmitted seismic waveforms on a wide range of scales. Data are commonly processed using traveltime tomography, often based on a ray-theoretical approximation of the elastic or acoustic wave equation (Červeny 1972; Červeny & Jech 1982; Zelt & Barton 1998; Zelt *et al.* 2003). The goal of traveltime tomography is to estimate a set of elastic Earth parameters (usually velocities) that minimize the differences between a set of observed and predicted seismic traveltimes. Although traveltime tomography has been used widely and successfully, the spatial resolution of

ray-theoretical methods are limited, and structures smaller than the order of the width of the first Fresnel zone are effectively invisible (Williamson 1991; Williamson & Worthington 1993).

As the power of computers has increased, it has become feasible to move beyond the asymptotic, high-frequency ray approximation, and instead to base tomographic methods on numerical models of seismic wave propagation via the elastic or acoustic wave equations. Tomographic methods using these techniques are commonly referred to as 'waveform tomography' (Brenders & Pratt 2007a,b), and are based on adjoint methods for waveform inversion first introduced by Lailly (1983) and Tarantola (1984). As in traveltime tomography, in waveform tomography we seek to iteratively update

an earth model with the goal of producing an accurate image of subsurface elastic parameters that minimizes the misfit between observed and predicted data. Whereas in the traveltime method the data are reduced to a set of traveltime 'picks', the data in waveform tomography are edited versions of the recorded seismic waveforms themselves. Furthermore, because waveform tomography replaces ray methods for data prediction with numerical modelling of the wave equation, this naturally negates the need to approximate wave propagation as occurring along an infinitely thin ray path, and complex, heterogeneous media is accurately accounted for.

In implementation, waveform tomography redistributes waveform residuals over spatially broad 'wave paths' (Woodward 1992). This is a more correct model of wave propagation physics that leads to an improved resolution limit on the order of one half of the seismic wavelengths of the data (Wu & Toksoz 1987). Due to the cyclical nature of waveform data the waveform tomography problem is highly non-linear, and as a consequence of the local inverse methods commonly used the successful application of the technique requires a starting model which predicts first arrival waveforms to within one-half cycle of the lowest frequency present (Sirgue & Pratt 2004). This is known as the 'half-cycle criterion', and it is a condition that is often difficult to satisfy in real data examples, in which low frequency data may be difficult to acquire. Failure to satisfy this condition leads to 'cycle-skipping' during the inversions.

Crosshole seismic tomography has been used by Vale Technology Development (Canada) Ltd. (referred to as Vale hereafter) to provide geophysical images of massive sulphides in the Eastern Deeps deposit at Voisey's Bay, Labrador, Canada (Enescu *et al.* 2002; Perozzi *et al.* 2012). The purpose of these surveys is to obtain information on ore-body geometries, and to use this information, along with any relevant geological information, to place spatial constraints on the location and extent of economic ores. The goal of this particular study is to provide Vale with a high resolution seismic velocity map of the target region via waveform tomography, and to investigate the feasibility of waveform tomography in similar hardrock environments. The geological setting is highly heterogeneous, with strong velocity contrasts and variable levels of seismic anisotropy (Enescu *et al.* 2002; Cosma & Enescu 2003); Perozzi *et al.* (2012) performed traveltime tomography on the same data set presented in this paper, however they did not consider anisotropy to be a factor.

Seismic velocity anisotropy can be significant in hard rock settings (Pratt *et al.* 1993). Cosma & Enescu (2003) estimate the anisotropy within the host rock at Voisey's Bay at approximately 3 per cent, while laboratory measurements have reported on troctolites showing anisotropy as high as 13 per cent (Babuska 1968), with Iturrino *et al.* (1991) reporting values as high as 6 per cent. While these laboratory measurements are very accurate, their relevance to true, *in situ* anisotropy parameters at crosshole frequencies are uncertain, as *in situ* pressures may close any fracture planes which are present at atmospheric pressure (Tsvankin 2012). Regardless, when applying high-resolution seismic imaging techniques to this area, the proper quantification of anisotropy at the relevant scale length is essential.

This geological setting presents particular challenges, including significant velocity discontinuities. This leads to the failure of ray-based traveltime tomography for this case study to correctly extract a useful anisotropy model, initially preventing us from satisfying the half-cycle criterion in our starting models for waveform tomography. It is the purpose of this paper to show that semi-global inverse methods can overcome the cycle-skipping phenomena. To do this,

we reduce the dimensionality of the model space to a level which permits the time-efficient computation of many forward models.

Since an anisotropy model which fits the data to the degree required for waveform tomography could not be found through manual picking and traveltime tomography, a novel method was developed to discover a 'best-fitting' 1-D anisotropy model using waveform data rather than traveltime data. We present a semi-global method, waveform-based simulated annealing, as an intermediate step that allows us to extract a layer-by-layer estimate of the seismic anisotropy of sufficient accuracy to allow for successful waveform tomography. The traveltime tomography method for determining anisotropy was abandoned, and attention was given to reducing the waveform residuals through perturbations to the anisotropy model using simulated annealing (while preserving the horizontal velocity model obtained from the traveltimes).

This paper is organized as follows: We begin with a review of relevant inversion methods, including those used in anisotropic traveltime tomography and in seismic waveform tomography. After a brief discussion of the geological setting of the Voisey's Bay deposit, we describe the data acquisition, instrumentation and survey procedures for the seismic data set. We then present some sample data gathers and comment on data quality, and we illustrate the effects of pre-processing. Following this, we discuss initial velocity and anisotropy models found through anisotropic traveltime tomography, and the cycle-skip phenomenon is identified in these results. After hypothesizing that the results are due to an erroneous anisotropy model, we present a synthetic study in which results of a similar nature are reproduced via numerical waveform simulations. We demonstrate that traveltime tomography may indeed produce incorrect anisotropy models, and we then illustrate how we successfully overcome this issue with the waveform-based simulated annealing approach we have developed. This is followed by a successful application of the simulated annealing method to the original data set, along with a validation of our results by comparisons between the inverted velocities and independent borehole velocity measurements, and a comparison with core logs. We finish with a discussion of the physical limitations of ray theory and traveltime tomography in a strongly heterogeneous, anisotropic model in order to better understand the breakdown of the method in this example.

2 METHODS

2.1 Inverse methods

We begin with a quick review of inverse techniques, followed by an introduction to the specific techniques relevant to the work presented. In the most general sense, Tarantola (2005) defines the solution of the inverse problem as a probability distribution $P(\mathbf{m} \in \mathfrak{M})$ over the space of all models $\mathfrak{M}$, with the form

$$P(\mathbf{m} \in \mathfrak{M}) = \exp\left[-\frac{1}{2}E(\mathbf{m})\right]. \tag{1}$$

Here, $E(\mathbf{m})$, known as the 'objective' or 'misfit' function, is an attempt to objectively measure of the quality of the model $\mathbf{m}$. A common way to define this function is by the L_2 norm of the differences between observed data ($\mathbf{d}$) and estimated data $g(\mathbf{m})$, combined with appropriate *a priori* information about the model $\mathbf{m}_0$

$$E(\mathbf{m}) = [g(\mathbf{m}) - \mathbf{d}]^T \mathbf{C}_d^{-1} [g(\mathbf{m}) - \mathbf{d}] + (\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0)$$

$$\tag{2}$$

(after Tarantola 2005). The notation $g(\mathbf{m})$ is used in eq. (2) to symbolize forward modelling, which produces an estimated data set by applying a functional operator g to a set of model parameters. We have included a measure of model and data uncertainty through the incorporation of model and data covariance matrices, $\mathbf{C}_m$ and $\mathbf{C}_d$, respectively.

The probabilistic interpretation of inverse problems requires a full description of the behaviour of the objective function over the space of all possible models. A less ambitious but highly useful goal is to find a set of models that are most likely given the data, and given the prior model. To accomplish this, we try to explore the maximum of eq. (1), which corresponds to exploring the minimum of $E(\mathbf{m})$. There are a variety of approaches that can be used to minimize $E(\mathbf{m})$ in eq. (2). These may be classified into

(i) Local methods that search in the neighbourhood of a starting model for nearby minima, and

(ii) Global or semi-global methods that attempt a more general search of the model space to identify global minima, which may be unreachable with local methods.

2.1.1 Local inverse methods

In local inverse methods, updates to an initial model are usually found by using information on the local gradient of the misfit function (Pratt *et al.* 1998), and model updates are iteratively estimated using linearized systems of the form

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} - \alpha^{(k)} \nabla_m E(\mathbf{m}^{(k)}), \tag{3}$$

where k refers to the iteration number, and α is known as the 'step length', a quantity that must be estimated as a part of the procedure.

While there are many variations on eq. (3), all local inverse methods rely on the same concept: that of successive minimizations to the misfit function along the gradient direction, with the requirement that the initial model is strictly within the 'zone of attraction' of the global minimum of $E(\mathbf{m})$.

2.1.2 Global or semi-global inverse methods

An alternative to local inverse methods are global inverse methods. Global inverse methods instead try to find the best fit model by exploring the model space $\mathfrak{M}$ with little reference to local derivative information. The simplest example is a grid search, which tests discrete models $\mathbf{m}^{(k)}$ in $\mathfrak{M}$, where k refers to the iteration number, and which returns an ensemble of models which best fit the data (as evaluated by the misfit function). An important, point must be made here. In defining $P(\mathbf{m})$ in eq. (1), we interpreted the solution of the inverse problem as a probability distribution over $\mathfrak{M}$. In the local optimization routines discussed above we seek a single solution $\mathbf{m}^{(k)}$ which best fits the data. This is because of the deterministic nature of gradient algorithms, which in most cases sample $\mathfrak{M}$ much too narrowly to make any rigorous claims about the full distribution $P(\mathbf{m})$. Global optimization methods overcome this problem as they attempt to sample the entire model space, and their solutions lead to the natural interpretation of eq. (1) as a probability distribution.

Similar to global inverse methods are 'semi-global' inverse methods which attempt to search a large portion of the model space, but also exploit some information on the structure of the misfit function to direct the search. While extremely robust, global and semi-global inverse methods have not been widely adopted in

traveltime or waveform tomography, although examples do exist (Sen & Stoffa 1991; Varela *et al.* 1998; Sambridge & Mosegaard 2002; Kaufl *et al.* 2013). This is mainly due to the extreme size of the parameter space, and the significant amount of time needed to perform the forward modelling. A comprehensive review of global and semi-global inverse methods within a geophysical context is given by Sambridge & Mosegaard (2002).

2.1.3 Simulated annealing

Of specific relevance to this paper is the semi-global method of simulated annealing (Rothman 1985, 1986; Kirkpatrick *et al.* 1987; Mosegaard & Vestergaarad 1991; Sen & Stoffa 1991; Varela *et al.* 1998). The concepts behind simulated annealing are borrowed from statistical mechanics, specifically from the analysis of a solid cooling in a heat bath. The method attempts to sample a probability distribution of the form

$$P(\mathbf{m}) = \exp[-E(\mathbf{m})/T], \tag{4}$$

where the probability assigned to a certain model $P(\mathbf{m})$ is dependent on the current value of the misfit $E(\mathbf{m})$, and a control parameter T. There are several ways to generate samples of the distribution in eq. (4) (Sen & Stoffa 1995). The one we adopt is the Metropolis algorithm (Metropolis *et al.* 1953), which proceeds as follows: First, a random perturbation, $\Delta\mathbf{m}$ is made to a single, randomly chosen model parameter belonging to model $\mathbf{m}^{(k)}$, resulting in $\mathbf{m}^{(k+1)}$ where

$$\mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \triangle\mathbf{m}; \left\{ \mathbf{m}^{(k)}, \mathbf{m}^{(k+1)} \in \mathfrak{M} \right\}. \tag{5}$$

The misfit function $E(\mathbf{m})$ is evaluated for $\mathbf{m}^{(k)}$ and $\mathbf{m}^{(k+1)}$. Depending on the values of $E(\mathbf{m}^{(k)})$ and $E(\mathbf{m}^{(k+1)})$, $\mathbf{m}^{(k+1)}$ is accepted with a probability $\mathbb{A}(\mathbf{m}^{(k+1)}, T^{(k+1)})$ where

$$\mathbb{A}\left(\mathbf{m}^{(k+1)}, T^{(k+1)}\right)$$
$$= \begin{cases} \exp\left(-\frac{E\left(\mathbf{m}^{(k+1)}\right) - E\left(\mathbf{m}^{(k)}\right)}{T^{(k+1)}}\right) & \text{if } E\left(\mathbf{m}^{(k+1)}\right) > E\left(\mathbf{m}^{(k)}\right) \\ 1 & \text{if } E\left(\mathbf{m}^{(k+1)}\right) \leq E\left(\mathbf{m}^{(k)}\right) \end{cases} \tag{6}$$

This recipe for acceptance is known as the Metropolis criterion. We can write the probability $\mathbb{P}(\mathbf{m}_j)$, of accepting any model $\mathbf{m}_j \in \mathfrak{M}$ as

$$\mathbb{P}(\mathbf{m}_j, T^{(k)}) = \mathbb{G}(\mathbf{m}_j) \times \mathbb{A}(\mathbf{m}_j, T^{(k)})$$
$$\mathbb{G}(\mathbf{m}_j) = \frac{1}{M}, \tag{7}$$

where $\mathbb{G}(\mathbf{m}_j)$ is the generation probability of $\mathbf{m}_j$, depending inversely on the number of elements M in $\mathfrak{M}$. The subscript j is used here to represent a certain realization of model parameters, while the superscript k retains its role as an iteration counter. Note that here we are discretizing and assuming a finite size for $\mathfrak{M}$, and that the generation probability is uniform.

To use simulated annealing with the Metropolis algorithm, we begin with a high value for $T^{(0)}$, and slowly decrease $T^{(k)}$ over the course of the inversion. While $T^{(k)}$ is large, most model perturbations will be accepted. As $T^{(k)}$ decreases, only perturbations which decrease $E(\mathbf{m}^{(k)})$, or which only increase $E(\mathbf{m}^{(k)})$ by a small amount, are likely to be accepted. Eventually, once $T^{(k)}$ drops to a low enough value, it will be overwhelmingly unlikely that any model perturbations that increase the misfit will be accepted, we assume convergence.

Revisiting the discussion on the pitfalls of local search methods, note the main advantage of the simulated annealing method: as long as the initial temperature is high enough the final model does not depend on the initial (or seed) model. In the high temperature limit, all randomly proposed models are accepted, and we are truly working with a global inverse method. However, the algorithm comes with a major limitation: if any finite-time cooling schedule is used, there is a chance that the global minimum of $E(\mathbf{m})$ will not be found. To mitigate this possibility, we should set the cooling schedule to be as slow as possible. However, practical requirements dictate that the algorithm finish in a reasonable amount of time. As a compromise, the cooling schedule is usually set so that simulated annealing will complete after a certain amount of computer time, or after a certain number of iterations (Sen & Stoffa 1995). By doing this, we trade a guarantee of convergence for a high likelihood of convergence, although unless the full model space is searched one cannot say for certain whether the final model obtained is the model which produces the minimum misfit.

Unfortunately, setting cooling schedules in this manner is largely *ad hoc* and problem dependent (Sen & Stoffa 1995; Tarantola 2005), although Nulton & Salamon (1988) and Mosegaard & Vestergaarad (1991) develop implementations which set the cooling schedule according to statistical information gained during the model search. Another difficulty is the estimation of the initial temperature $T^{(0)}$. We must ensure that at this temperature the Metropolis criterion will be passed for a large variety of models; otherwise, the initial model will have an impact on the final result. This choice of the initial temperature is also problem dependent, although a strategy is suggested in a later section.

It should be noted that the use of the term 'simulated annealing' does not automatically imply that models are perturbed only one parameter at a time, or that the generation probability is uniform and temperature independent. Indeed, Szu & Hartley (1987) generate perturbations with a temperature dependent probability. The behaviour of the algorithm depends on how the perturbations are treated. In our formulation, where a single model parameter is perturbed at each iteration, not all points in model space are accessible at one instant: we may only 'crawl' along one axis at a time. This is in contrast to the approach of perturbing all model parameters, in which case a multidimensional 'jump' through model space is possible. The choice of a 'crawling' or 'jumping' formulation is likely problem dependent, although for problems where certain model parameters are better constrained than others the choice of a crawling formulation may speed up convergence. This is the case we encountered in our field data set, and the benefits of adopting the crawling approach will be expanded upon in section 5.

2.2 Traveltime tomography in 2-D, weakly anisotropic media

Our anisotropic traveltime tomography approach is based on the work of Chapman & Pratt (1992) and Pratt & Chapman (1992); the reader is referred to these papers for a detailed description of the methodology. Here we will quickly review some critical aspects which are especially relevant to this paper.

Pratt & Chapman (1992) advocated a system for controlling the *a priori* model variance in which the variance of the prior (isotropic) velocity can be set independently from the variance of the parameters controlling anisotropy. This two-tiered system is required due to the fact that the parameters controlling anisotropy are often much more poorly determined than those controlling isotropic velocity

(Tsvankin 2012). A single parameter is used to quantify the relative variances between isotropic and anisotropic parameters. If this parameter is small, this allows the introduction of anisotropy into the model without significant increases in the objective function $E(\mathbf{m})$. In contrast, if this parameter is large, this is equivalent to assuming the variance of the anisotropy parameters is much smaller than the *a priori* (isotropic) model variance, and perturbations to any anisotropic parameters will thereby be penalized due to their significant influence on the value of $E(\mathbf{m})$.

Transversely isotropic media with either a tilted symmetry axis (TTI) or a symmetry axis aligned with the vertical (VTI) are investigated in this paper. The parameters we use for describing TTI or VTI systems are those proposed by Thomsen (1986): V_p (the velocity in the fast direction), ε (the fractional difference between fast and slow velocity) and δ (the parameter describing the anellipticity). Note that although this approach allows for general, 2-D TTI anisotropy, ultimately our acoustic waveform tomography scheme (described below) is limited to VTI systems with elliptical ($\delta = \varepsilon$) anisotropy that varies with depth alone.

The degrees of freedom available for the inversion scheme necessitate a proper strategy for choosing values for the required regularization parameters. Pratt & Chapman (1992) suggested that the regularization parameters be chosen according to the appearance of a suite of tomograms, and relaxing the constraints until a model fits the data to within an acceptable margin of error while remaining geologically plausible. Deciding what an 'acceptable margin of error' is can be somewhat ambiguous, and Pratt & Chapman's (1992) suggestion is to try and fit the data to within the estimated picking error. As discussed below, if the goal of traveltime tomography is the generation of a starting model for waveform tomography, we have the additional requirement that the error in the traveltime residuals be less than one-half of the period of the lowest frequency signal present in the survey data (Sirgue & Pratt 2004).

2.3 Acoustic waveform tomography in 2-D, weakly anisotropic media

As mentioned in the introduction, the purpose of waveform tomography is to discover a subsurface model that minimizes the difference between observed and calculated waveforms. A full overview of the theory for isotropic media is given in Pratt *et al.* (1998), Shin & Cha (2009) and Kamei & Pratt (2013).

Following the development of Kamei *et al.* (2012), we first transform the recorded time-domain wavefield $u(\mathbf{x}, t)$ into the Laplace–Fourier domain via a (discrete version) of the complex-valued Laplace Transform:

$$u(\mathbf{x}, s) = \int_0^\infty u(\mathbf{x}, t) \exp[-st] \, dt, \tag{8}$$

where

$$s = \sigma + i\omega, \tag{9}$$

with a real-valued Laplace constant σ and a real-valued angular frequency ω. Eq. (8) can be expanded to yield

$$u(\mathbf{x}, s) = \int_0^\infty u(\mathbf{x}, t) \exp[-\sigma t] \exp[-i\omega t] \, dt. \tag{10}$$

We interpret the parameter σ as the inverse of a characteristic decay time

$$\tau = \frac{1}{\sigma}, \tag{11}$$

so that eq. (10) represents the Fourier transform of each time-damped trace

$$u(\mathbf{x},\, s) = \int_0^\infty \left\{ u(\mathbf{x},\, t) \exp\left[-\frac{t}{\tau}\right] \right\} \exp[-i\omega t]\, dt. \tag{12}$$

The purpose of formulating waveform tomography in the Laplace–Fourier domain, instead of just the Fourier domain, is to use the parameter τ as a data pre-conditioner (Sirgue & Pratt 2004; Brenders & Pratt 2007a; Kamei *et al.* 2013), which preferentially weights data corresponding to early arrivals. We can further rewrite eq. (12) as

$$u(\mathbf{x},\, s) = \int_{-\infty}^\infty u(\mathbf{x},\, t) \exp[-i\Omega t]\, dt \tag{13}$$

$$= u(\mathbf{x},\, \Omega), \tag{14}$$

where $\Omega = \omega - i/\tau$ is a complex-valued frequency. In this development we see that the complex Laplace and Fourier transforms are equivalent, and as such we refer to the transformed wavefield as the 'Laplace–Fourier domain wavefield' (Shin & Cha 2009).

We employ a 'logarithmic residual', or

$$\delta d_j = \ln\left(\frac{u_j}{d_j}\right) \tag{15}$$

(Shin & Min 2006), where u_j and d_j refer to the modelled and recorded wavefield, respectively, at receiver j. Expressing the complex-valued data in terms of their amplitude (A) and phase components (θ), we can write

$$u_j = A_{u_j} \exp[i\theta_{u_j}], \tag{16}$$

$$d_j = A_{d_j} \exp[i\theta_{d_j}], \tag{17}$$

so eq. (15) becomes

$$\delta d_j = \ln\left(\frac{A_{u_j}}{A_{d_j}}\right) + i\left(\theta_{u_j} - \theta_{d_j}\right). \tag{18}$$

This allows for a trivial separation of amplitude and phase in the misfit function in the real and imaginary parts of the data residuals.

Note that when using this development, we assume that the phases of u_j and d_j have been unwrapped to the same degree, meaning $|\theta_{u_j} - \theta_{d_j}| \leq \pi$. This is critical, as a perturbation to θ_{u_j} by $2n\pi$, where n is any integer, will leave the misfit unchanged (due to the periodicity of the phase). This leads to the realization that if the predicted and observed wavefields are more than π (one half-cycle) out of phase in the initial model, any local optimization method is at risk of convergence towards the wrong cycle. This is what is meant when we refer to the cycle-skipping phenomenon; the half-cycle criterion is the condition on the initial model that is required to strictly avoid this phenomenon.

2.3.1 Weak anisotropy

In order to make use of the waveform tomographic approach described above where the media is anisotropic, a fully anisotropic forward and inverse approach would be required. Although such an approach has been proposed (Alkhalifah 2000; Brossier *et al.* 2010; Plessix & Perkins 2010; Warner *et al.* 2013), we follow here the approach developed and applied by Pratt *et al.* (2004), in which the isotropic approach is modified in a simple way that allows us to incorporate a 1-D background model consisting of VTI, elliptical anisotropy that varies on a layer-by-layer basis. Under these conditions the kinematic effects of the anisotropy may be accounted for by implementing a layer-by-layer geometrical stretch proportional to the amount of anisotropy in each layer, an approach originally suggested by Dellinger (1991). This limitation requires that we attempt to represent an-elliptic anisotropy with an elliptic model; as such the parameter ε has the added responsibility of approximating the effect of δ over the range of angles in the survey. The anisotropy is assumed to be fully described in the starting model, and is not updated during the waveform tomography stages.

3 ACQUISITION AND INITIAL PROCESSING

The main purpose of the experimental crosshole study was to investigate the continuity of a mineralized massive sulphide zone located in Voisey's Bay, Labrador, Canada. We used seismic waveform data from Vale's property obtained from a crosshole seismic survey designed to map the continuity of the sulphide mineralization. The data set included the seismic waveform data, the source and receiver geometry, and acoustic borehole logging results from the source and receiver boreholes.

3.1 Geological and geophysical setting

The Voisey's Bay deposit, located on the East Coast of Canada in Labrador, contains significant economic quantities of sulphide-hosted nickel, copper and cobalt ores (Evans-Lamswood *et al.* 2000). The mineralized deposits are part of the Voisey's Bay Intrusive Suite, which consists primarily of layered troctolite-gabbros intruding into Archean quartz-feldspar-biotite gneisses of the Nain province, and Proterozoic sulfidic garnetiferous paragneisses of the Churchill province (Ryan *et al.* 1995). There are two magma chambers associated with the Voisey's Bay deposit, known as the Eastern Deeps and Western Deeps chambers. This project deals solely with the Eastern Deeps magmatic system, consisting of a feeder tube (the Eastern Deeps Feeder), which leads into the large Easter Deeps Chamber. The main control on the location of mineralization is structural, which implies that deposition is controlled by morphological changes in the conduit geometry, and not by gravitational settling within the chamber itself (Evans-Lamswood *et al.* 2000). Apart from these areas of massive mineralization, the chamber is dominated by troctolite-gabbros, containing a variable amount of sulphide fragments, and granite-syenite.

The specific boreholes used in our crosshole study were drilled in 2000 by Boart Longyear (Inco 2000), and they were designed to intersect a zone of massive sulphide mineralization at depth, referred to as the 'target'. A mock-up of the target area is schematically illustrated in Fig. 1. As expected from *a priori* geological information, the borehole paths mainly intersected troctolites containing variable amounts of sulphide in addition to successfully intersecting the target: an approximately 60-m-thick mineralized zone, corresponding to the location of the Eastern Deeps Feeder at a depth of approximately 720 m. After drilling, the borehole geometries were surveyed with the use of a north-seeking gyroscope, which recorded the azimuthal orientation and dip of the holes at 30 m intervals.

Borehole sonic logs, also presented in Fig. 1, give a high resolution measurement of the vertical seismic wave velocity in the immediate vicinity of the borehole (at frequencies on the order of 20 kHz). We observe a significant and rapid decrease in sonic velocity associated with the mineralized target zone. This is confirmation
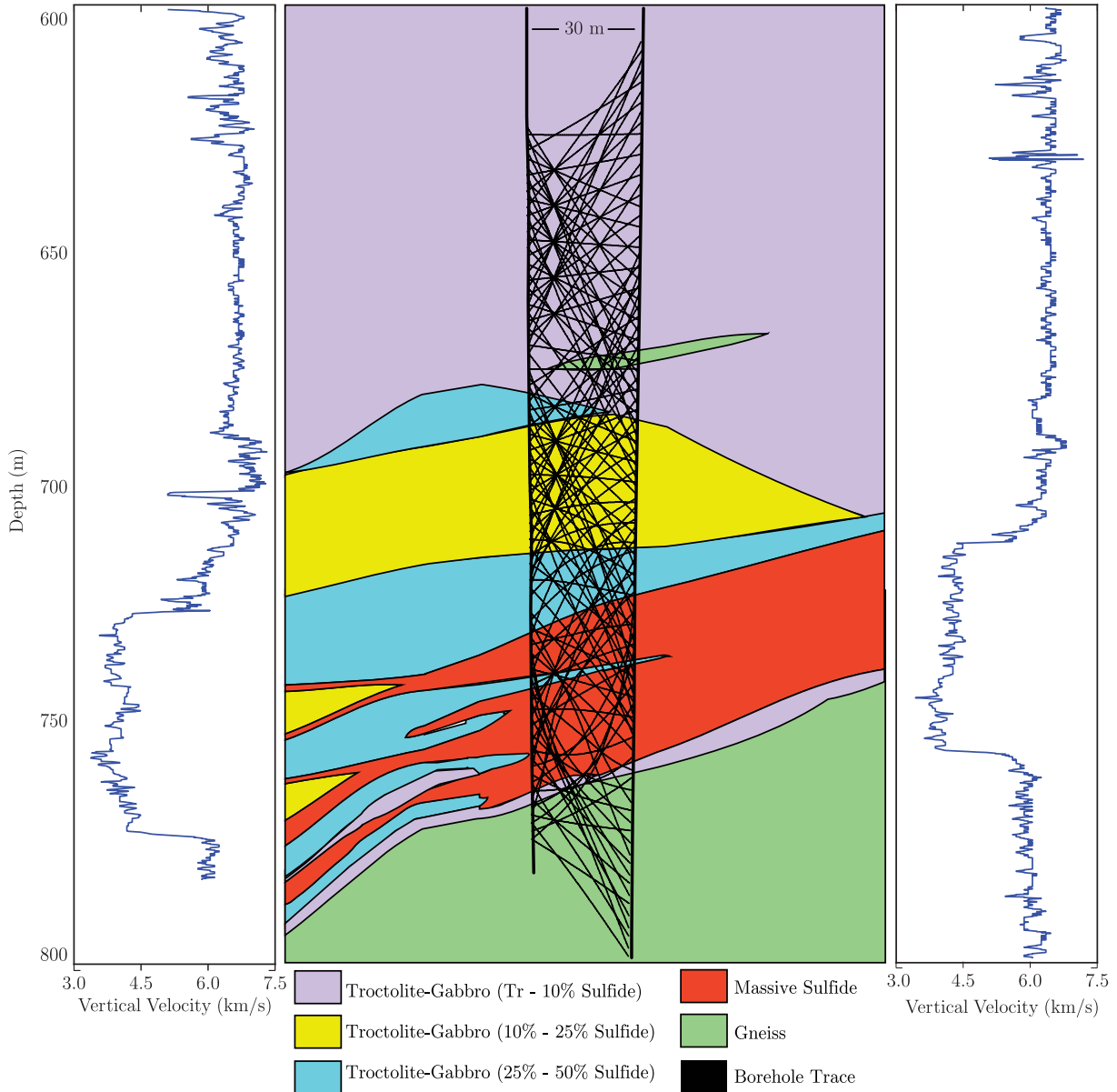
Figure 1. An idealized mock-up of the geological cross-section illustrating (schematically) the expected geology at survey depths [derived from Evans-Lamswood *et al.* (2000)]. The majority of the country rock is troctolite-gabbro, which surrounds a mineralized massive sulphide zone. The two boreholes used in the survey, and a subset of the seismic raypaths are overlain. Also present are the actual sonic velocity logs in the receiver (left-hand panel) and source (right-hand panel) holes. The low velocities seen in the sonic logs are associated with the mineralized zone.

that the seismic velocity of the massive sulphide body is in major contrast to the troctolites of the host rock, an observation supported by laboratory measurements (Eaton *et al.* 2010). Since seismic tomography attempts to produce a map of subsurface velocity, this contrast motivates the use of the method in the Voisey's Bay area, as we expect to see a strong velocity anomaly associated with the target zone in our results.

3.2 Data acquisition

An array of 29 hydrophones was used in the receiver hole, with a hydrophone interval of 2 m. This array was deployed at three separate depths, spaced 60 m apart. This resulted in a total of 87 unique

receiver locations. The sources were deployed at 1 m intervals in the source hole, and 121 shots were recorded for each receiver layout. The source was based on the Swept Impact Seismic Technique (SIST; Park *et al.* 1996). This allows an acceptable signal-to-noise ratio to be achieved without using high energy impulsive sources, which can damage, and sometimes collapse, the host boreholes. Both the source and the hydrophones were manufactured by Vibrometric Ltd.

We refer to the combination of a static receiver array depth and the corresponding shots as a 'panel'; in total three panels were shot. As there was some overlap in the shot location after a layout change, the total number of unique shot locations was 236. The shots thus extended for a total of 236 m, beginning approximately 100 m above the top of the massive sulphide zone, and ending
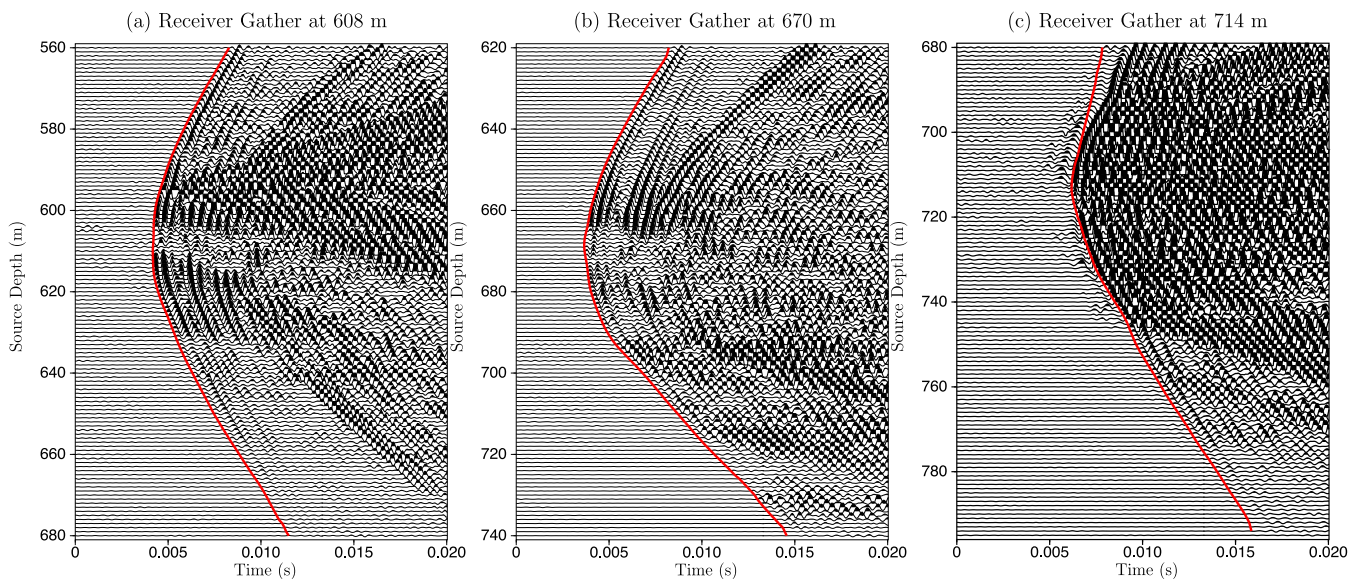
Figure 2. Selected raw receiver gathers from the Voisey's Bay data; (a) [608 m] and (b) [670 m] are representative examples from the high velocity troctolites (as seen in the sonic logs in Fig. 1); (c) [714 m] shows a receiver located within the low velocity massive sulphide zone. Red lines on the figures depict the first arrival picks.

approximately 50 m below the bottom. Due to the dip of the target zone, the first receiver was also approximately 100 m above the top of the massive sulphide zone. Unfortunately, due to the limited depth of the receiver borehole, the final receiver was located only a few metres below the target zone, limiting angular coverage for the deepest part of the target zone. The hole-to-hole offset at the survey depth was approximately 30 m, although this narrowed to approximately 26 m towards the bottom of the holes.

3.3 Data description

The signal quality from the crosshole survey was in general good, with clear first arrival waveforms and a low level of noise. A small number of receiver gathers exhibited corruption, and these were identified and removed from the data set without further processing. Examples of single receiver gathers from the Voisey's Bay survey, each showing the received waveforms from 121 unique source positions, are given in Fig. 2. For these data, the first arrival times were picked by hand using ProMAX. Where present in Fig. 2, first arrival times are highlighted in red. Not all traces were picked; where picking was difficult due to noise levels no arrival time was assigned. An absence of an arrival time pick resulted in the rejection of that particular trace in subsequent waveform tomography. Figs 2(a) and (b) show receivers located within the high velocity troctolite zone. Here, the waveforms were well behaved, and the first arrival times were relatively easy to pick on the waveform data. Fig. 2(c) shows a receiver gather within the low velocity zone. The waveforms here were significantly more complicated than those in the troctolite zone, due mainly to the influence of the strong velocity contrast at the troctolite-sulphide boundary, and arrival time picking was quite difficult.

A 'precursor' component of the waveforms may be observed in Fig. 2, visible as oscillations appearing prior to the first arrival picks. These precursors, which likely arise from an acausal low cut filter applied either during or shortly after the data acquisition, made picking even more difficult. The filter appears to have been used to remove most of the energy in the waveforms below 1000 Hz.

Also evident in Fig. 2 (especially in Fig. 2c) are high amplitude arrivals which appear to have a linear, steep slope representing waves propagating in the source borehole with a low and constant propagation velocity. These are caused by the presence of tube waves, which are waves propagating vertically within the water-filled boreholes. These waves are the result of body wave energy being injected into the water column of the source hole, where they propagate with a constant velocity of approximately $1500 \, \mathrm{m \, s^{-1}}$. When these tube waves encounter geological discontinuities, such as those at the edges of the massive sulphide zone, their energy is converted to body waves, which eventually reach the receiver hole. As our waveform tomography routine does not model tube waves, their presence in the data can lead to artifacts in the final image. However, their removal is challenging (as described below).

3.4 Traveltime tomography

In order to generate an appropriate starting model for waveform tomography, traveltime tomography was applied to the data following traveltime picking. As described previously, our traveltime tomography routine is based on the work of Chapman & Pratt (1992) and Pratt & Chapman (1992). In addition to 2-D images of the horizontal *P*-wave velocity, the traveltime inversions include 2-D images of Thomsen's epsilon (ε) and delta (δ), along with estimates of the orientation of symmetry axes under the assumption of 2-D TTI media. The initial results of traveltime tomography, consisting of images of the isotropic velocity and the magnitude of ε, are given in Fig. 3. These results were produced via five outer (nonlinear) iterations, between each of which new raypaths were calculated by the ray-bending method of Um & Thurber (1987). Within each nonlinear iteration, the raypaths were kept constant, and the traveltime residuals were minimized via the LSQR algorithm (Paige & Saunders 1982). The use of a tilted symmetry axis, instead of a vertical axis (as is required by our waveform tomography routine), was necessary. Assuming a vertical axis of symmetry led to degradations in the plausibility of the model; in particular resulting in horizontal velocities at the edges of the mineralized zone that were
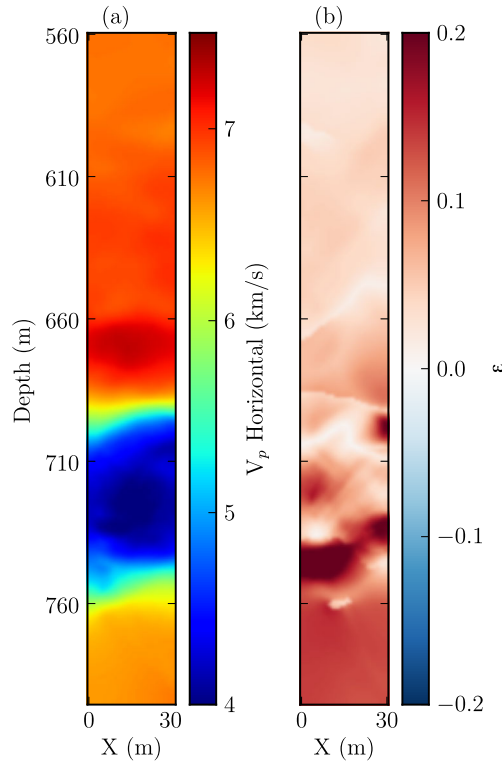
Figure 3. Traveltime velocity and anisotropy models. (a) Horizontal velocity; (b) Thomsen's ε.

in places less than $2\,\mathrm{km\,s^{-1}}$. The constraint applied to perturbations to parameters controlling anisotropy was the same as that applied to perturbations to isotropic velocity. A heavier constraint on the parameters controlling anisotropy led to an increase in traveltime residuals, along with velocity models which were geologically unrealistic. The rms traveltime residual after five non-linear iterations was 0.049 ms. Anisotropy in the high velocity troctolites (Fig. 3b; $\varepsilon = 4$–5 per cent) coincides well with the previous studies cited in the introduction (i.e. Cosma & Enescu 2003). Around the edges of the low velocity zone, a strange lobe-like structure is observed in the anisotropy model. We believe this is an inversion artefact, and a discussion of its origin will be returned to in Section 5.3.1.

We should also note that although we assume a transversely isotropic model (with an arbitrary axis of symmetry) in traveltime tomography, this is likely a simplification of reality. The borehole core logs (Inco 2000) report on highly fractured rock, with multiple fracture symmetry axes, overprinted on local flow and shear fabrics. As well, post emplacement serpentine metamorphism (Inco 2000) is known to be highly anisotropic (Kern 1993), and may add additional anisotropic properties along some cracks. To completely characterize the anisotropy would likely require the use of orthorhombic, monoclinic or even triclinic symmetry systems.

3.5 Data preprocessing

Prior to waveform tomography, waveform data must be preprocessed in an attempt to attenuate undesirable signals. These signals include records of tube waves, ambient noise, and elastic effects. During the preprocessing stage, it was observed that the data had been high-pass filtered to above 1000 Hz at acquisition.

3.5.1 f-k domain and attempts to remove tube waves

Since tube waves appear at a near constant slope, it is common to remove them from the data through frequency-wavenumber filtering (*f–k* filtering) (Chen *et al.* 1990). Unfortunately, the high-pass filtering at acquisition resulted in the tube waves being aliased in the usable frequency band (above 1000 Hz). Due to this aliasing, the tube wave energy could not be removed cleanly from the data using simple *f–k* filtering, and the method was not applied. We might therefore expect this to result in artefacts in our final waveform tomography images without further processing.

3.5.2 Data pre-conditioning: windowing, bandpass filtering, editing and normalization

The waveforms were pre-conditioned for waveform tomography via high-cut filtering to less than 2000 Hz (to avoid the necessity of modelling high-frequency data), windowing in the time domain, editing by removal of anomalous source and/or receiver gathers and normalization of the data so that each source gather, and each receiver gather contain consistent amplitudes. The windowing served to mute any noise present before the first arrivals, and also to exclude (as much as possible) late-arriving tube waves, as well as any shear waves in the data (which are not modelled by the acoustic wave equation). The window began 0.5 ms before the first arrival picks, and extended to 3 ms after.

Images of selected frequency domain data files for several frequencies, after bandpass filtering, time windowing, editing and normalization are given in Fig. 4. In these images, sources increase from 1 to 236 along the horizontal axis, and receivers increase from 1 to 87 along the vertical axis. Three distinct blocks are visible, and each block corresponds to one survey panel. Note that some sources are present in more than one panel, illustrating the repetition of some source positions. Due to the geometry of the survey, the source–receiver pairs with the closest offsets appear approximately along the main diagonal in the images, and source–receiver pairs with approximately equal offsets appear along subdiagonals parallel to the main diagonal.

While time-windowing of the data introduces a low frequency component due to spectral leakage, it is obvious from Fig. 4 that most of these low frequency data are corrupted and noisy. While the waveforms at 500 Hz show some coherency (Fig. 4a), those between 600 and 1000 Hz are very incoherent (a representative example at 800 Hz is given in Fig. 4b). This echoes our observation that the data were filtered at acquisition to remove data below 1000 Hz. We expect the lack of low frequency data to increase the chances of cycle-skipping.

Above 1000 Hz, the processed data are of good quality and coherent. Representative examples at 1300 and 1600 Hz can be seen in Figs 4(c) and (d), respectively.

Other relevant information can be visually extracted from Fig. 4. Over most of the survey (i.e. in the high velocity troctolite host rocks), the coherent patterns in these displays are subparallel to the main diagonal, and cycle in phase perpendicular to the main diagonal. However, for those sources and receivers in the low velocity mineralized zone (sources 130–180, receivers 58–87), we see that these coherent patterns in the data change in slope. Within this zone the coherency is no longer parallel to the the main diagonal. This is a sign that the data in this region are spatially aliased, which strongly suggests a corresponding decrease in seismic velocity (shortening the wavelengths to below the Nyquist criterion).
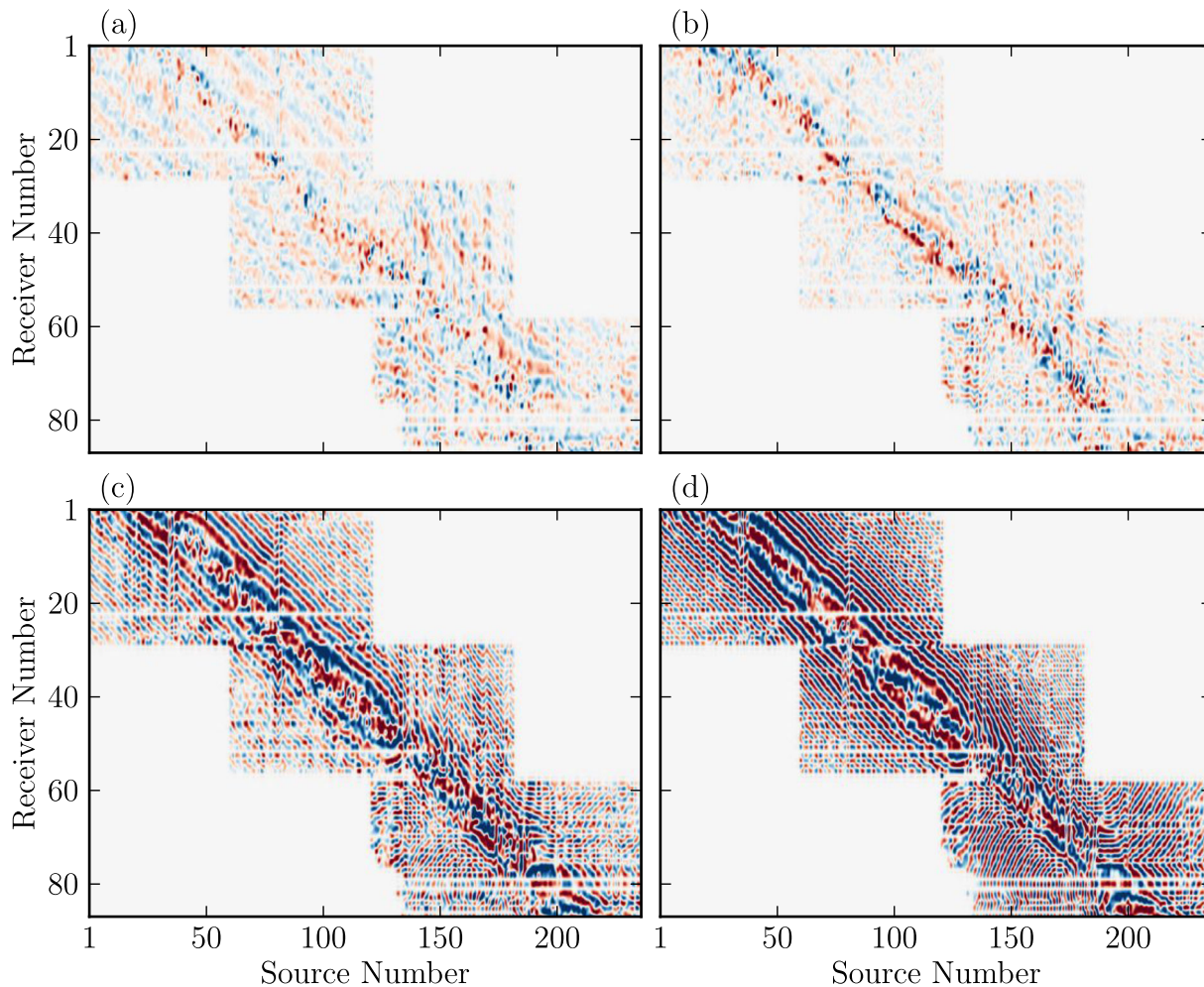
Figure 4. Frequency domain displays of the Voisey's Bay data following bandpass filtering, time windowing and amplitude normalization over the source and receiver gathers. (a) 500 Hz; (b) 800 Hz; (c) 1300 Hz; (d) 1600 Hz. We see ample evidence of coherent patterns parallel to the main diagonal, especially on panels (c) and (d). As the distances between sources and receivers are increased (i.e. perpendicular to the diagonals) the signal may be observed to cycle in phase [evident as a change from red to blue caused by the change in the polarity of Re(u)].

Discontinuities parallel to the receiver axis (the vertical axis) are also apparent at several locations (for example, around source 75). As the horizontal axis represents source location, these discontinuities may be related to source-borehole coupling issues, as well as to possible source mis-location issues. Finally, white zones parallel to the source axis are also present. These represent deleted receivers, which were excluded due to excessive noise. Specifically, the 23rd receiver in the array was consistently corrupted, and is therefore absent from each panel on Fig. 4.

3.6 Assessment—waveform modelling of traveltime tomography result

In order to assess the quality of the traveltime tomography results for the purposes of waveform tomography, we simulated waveform data numerically in the traveltime velocity model (Fig. 3a) and compared them with the real (pre-processed) data. To prepare the model for waveform modelling, we created a 1-D, vertically varying anisotropy profile by averaging the ε model in Fig. 3(b) over horizontal layers. The resultant profile can be seen in Fig. 5. Note that this approach to waveform tomography assumes that the 2-D anisotropy varies only with depth, that the anisotropy is elliptical

($\varepsilon = \delta$), and that the symmetry axes are close to vertical. Examining Fig. 3(b), and noting that the image was created by allowing for ($\varepsilon \neq \delta$) and allowing a tilted symmetry axis, we see that these assumptions are unrealistic. Nevertheless, a realistic traveltime tomography velocity image could only be found by assuming tilted transverse isotropy. We therefore interpret Fig. 5 as a measure of the average magnitude of ε at each depth level—regardless of symmetry direction. The velocity model was interpolated onto a 96×500 cell finite difference grid, with the cell edge length set at 0.5 m.

Time domain waveforms were generated through this anisotropic velocity model. The main purpose of performing time-domain modelling was to determine if the half-cycle criterion was satisfied. The resulting predicted receiver gathers from the traveltime model (Fig. 6b) show a mismatch in the first arrival times at larger offsets of more than a half cycle. Specifically, we see this for shots deeper than 730 m. At near offsets, the first arrivals were consistently predicted to within one half wavelength of the dominant frequency. The pattern of a good fit at near offsets, and of cycle-skipping at higher offsets, was observed everywhere within or near the low velocity zone. Since an increase in offset corresponds to an increase in angle away from the horizontal, this strongly suggests that the averaged, elliptical, VTI anisotropy model found through traveltime tomography is insufficient for use in waveform tomography.
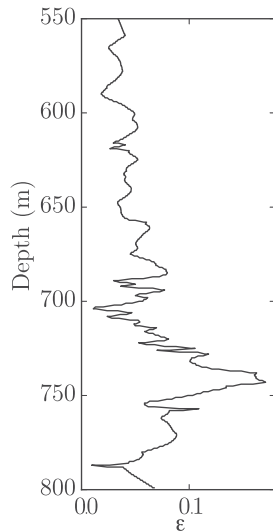
Figure 5. 1-D anisotropy profile obtained by averaging the magnitude of ε in the 2-D model shown in Fig. 3(b) over horizontal layers.

3.7 Necessity of a semi-global approach

A usable anisotropy model could not be found through manual picking and traveltime tomography. Instead, generating such a model required a novel method; such a method would have to respect the waveform nature of the data. Since cycle-skipping was present, it was clear that a local inverse method for waveform tomography would not be sufficient. Our solution to this was to develop a waveform-based simulated annealing approach that would find the appropriate anisotropy model for subsequent waveform tomography. To determine the viability of the simulated annealing approach as a global method for anisotropy inversions, a test of the method was initially performed on synthetically generated crosshole data, for which the true models would be available for comparison purposes. Synthetic models based on the Voisey's Bay geological setting were created, and seismic data were calculated numerically within these models, fully respecting the anisotropy distributions. The next section follows the development of simulated annealing strategies and methodologies with reference to the synthetic examples, and we will return to the real Voisey's Bay data in Section.

4 SIMULATED ANNEALING—DEVELOPMENT AND APPLICATION TO SYNTHETIC CROSSHOLE DATA

The synthetic velocity and anisotropy models we use in this section to test the effectiveness of simulated annealing were designed to simulate the essential elements of the actual Voisey's Bay field data. In order to do this, we simulated the forward data with a full elastic, anisotropic method yielding realistic synthetic data that include both qP and qS waves. Elastic time domain VTI finite difference modelling was performed using the freely available open source software package Madagascar (Yan & Sava 2011). Madagascar was chosen so that a realistic, anelliptic 2-D anisotropy model could be used, and as well to avoid the pitfalls of performing an 'inverse crime' (Colton & Kress 1998; Wirgin 2004). The anisotropy we introduced matched that observed in the traveltime inversions above with the real data, with one important distinction: the synthetic anisotropy models assumed a vertical axis of symmetry (VTI). Since the primary goal of these experiments was to investigate the applicability

of simulated annealing to the anisotropy problem, we consider this sufficient. The anisotropy was elliptic in nature ($\varepsilon = \delta$).

The true models for V_p and ε are given in Fig. 7. The source-receiver geometry was the same as that for the field data, with boreholes approximately 30 m apart, sources spaced 1 m apart, and receivers spaced 2 m apart. The background velocity was set to be 7.5 km s^{-1}. A dipping, low velocity layer was introduced, simulating the actual location and orientation of the massive sulphide zone, as estimated from the borehole sonic logs (see Fig. 1). As well, a layer of slightly higher velocity was placed just above the low velocity layer, corresponding to the slight increase in borehole sonic velocities seen above the massive sulphide zone. Random heterogeneities were then added to the model using the methods of Ikelle *et al.* (1993) and Kamei *et al.* (2005); the magnitudes of these heterogeneities were estimated from observed fluctuations in the borehole logs. In the high velocity zone, the heterogeneities had a standard deviation of 300 m s^{-1}, while in the low velocity layer the heterogeneities had a standard deviation of 210 m s^{-1}. A vertical to horizontal ratio for the correlation lengths of 1:4 was used for these heterogeneities, to investigate the impact of horizontal layering on the inverted anisotropy models. The long axis of the heterogeneities added to the low velocity zone was rotated to coincide with the dip of the zone itself. Density values for the high and low velocity zones were chosen to correspond to the average density of troctolite (2.7 g cc^{-1}) and pyrrhotite (4.6 g cc^{-1}), respectively. The *P*- to *S*-wave velocity ratio was set everywhere equal to 1.87. In the high velocity zone, ε was set to a constant 0.15, while in the low velocity layer it dropped to 0.05.

For both models, an explosive, Küpper-wavelet source was modelled, with a peak frequency of 666 Hz and a maximum frequency of approximately 2000 Hz, and the scalar curl-free (pressure) component of the wavefield was extracted at each receiver position.

4.1 Traveltime tomography

To prepare the synthetic data for traveltime tomography, little effort was required. Due to the noise-free nature of the synthetic data, the majority of the first break picking was very easy, and could mostly be carried out semi-automatically. This being said, there were still significant difficulties at the edges of the low velocity zone, where some manual intervention was required.

Traveltime tomography was performed following the same philosophy given in Section 3.4. Five nonlinear inversions were carried out, during which the constraint on model roughness was gradually relaxed. Traveltime velocity and anisotropy results for the elliptical synthetic data are depicted in Fig. 8. The values of ε in the high velocity host rock zone are overestimated by approximately 5 per cent. Lobes of strong anisotropy around the low velocity zone, similar to those observed in Fig. 3, are also present. The failures of the traveltime tomography method is attributed to the fine scale of the heterogeneities that lie outside the resolution capability of ray tomography; a discussion of this effect follows in Section 5.3.

4.2 Waveform tomography

4.2.1 Frequency domain waveform data

Since the synthetic data were noise free and not corrupted by tube waves, it was not necessary to apply time damping as was done with the real data, and thus the damping parameter τ in eq. (10) was set to infinity. This results in eq. (10) reducing to the classic Fourier
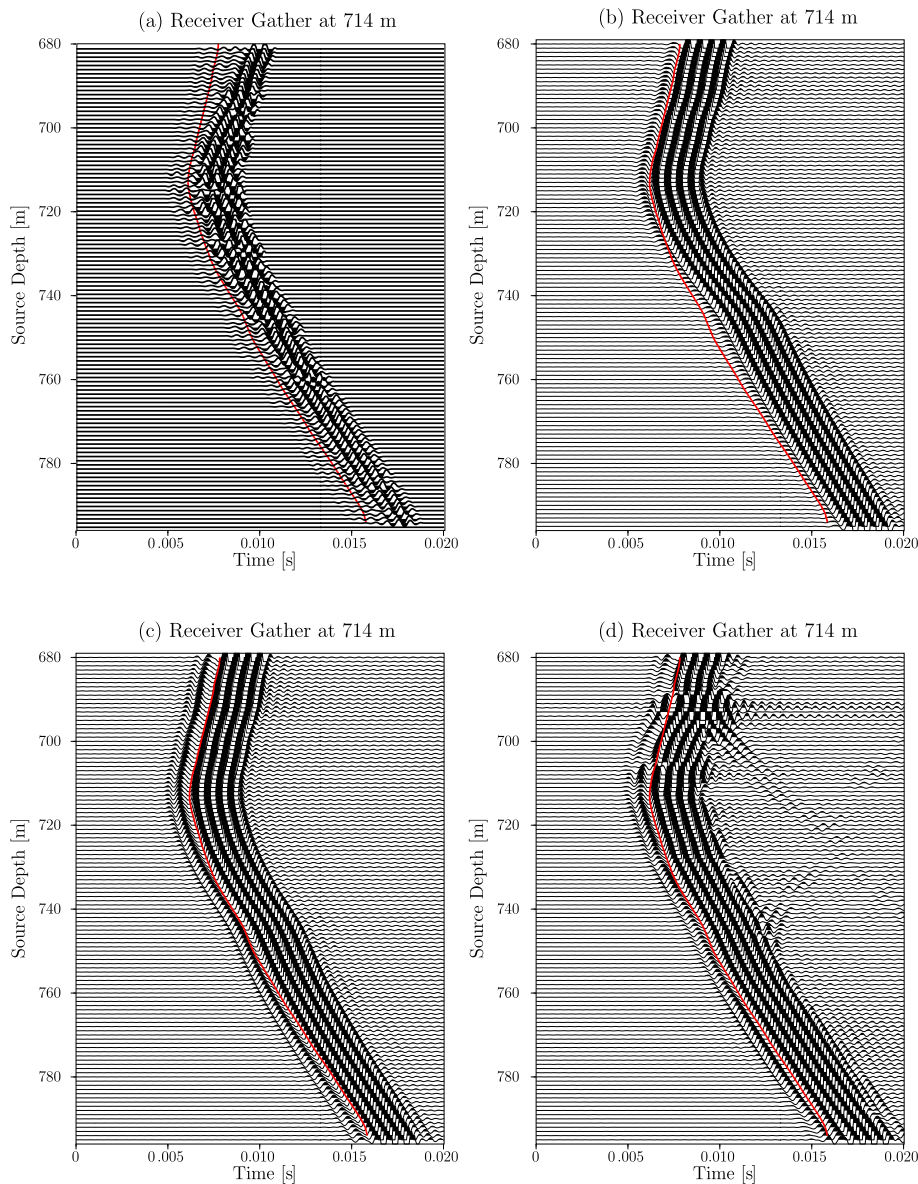
Figure 6. Receiver gathers at 714 m. (a) Real, windowed, bandpass filtered and time-damped ($\tau = 0.01$) data; (b) predicted data, modelled using traveltime (TT) velocity model and TT anisotropy; (c) predicted data, modelled using TT velocity model and simulated annealing (SA) anisotropy; (d) predicted data, modelled using waveform tomography (WT) velocity model and simulated annealing (SA) anisotropy. For reference, and to visually indicate cycle-skipping, the red lines on each plot represent the same, picked arrival times.

transform. The noise free nature of the synthetic data also resulted in minimal pre-processing requirements: the waveforms were simply windowed in the time domain (to remove late-arriving shear modes) and bandpass filtered (to below 2000 Hz) before being transformed into the frequency domain. An example of a frequency domain wavefield at 1300 Hz is given in Fig. 9(a). We can see significant waveform complexity in the low velocity zone (sources 120–180; receivers 59–70). Fig. 10(a) shows the original, or 'true' 1600 Hz wavefield, zoomed into the vicinity of the low velocity zone.

4.2.2 Anisotropy estimation by waveform fitting using simulated annealing

Since simulated annealing, by its semi-global nature, relies on the computation of many forward models, care must be taken to design

the model space in a computationally efficient way, while still allowing a wide variety of candidate models to be tested. We defined our model space as the space of all epsilon values [$\varepsilon(z)$] between -0.1 and 1.0, and discretized the model space in intervals of $\triangle \varepsilon = 0.001$. To reduce the dimensionality of the problem to a computationally tractable level, we enforced smoothness on $\varepsilon(z)$ via a cubic spline representation with tie points spaced every 25 m; this resulted in a total of 10 model parameters. A seed model of $\varepsilon(z) = 0.1$ was used to initialize the algorithm. The finite-difference grid size and spacing were identical to that used for the real data, with 96 cells in the horizontal direction, 500 cells in the vertical direction, and a grid size of 0.5 m $\times$ 0.5 m.

At each iteration, a random model parameter, defined as the value of ε at a particular cubic spline tie point, was perturbed to a random value in model space. Synthetic waveforms at 1000, 1300 and 1600 Hz were then calculated, with the frequency values chosen
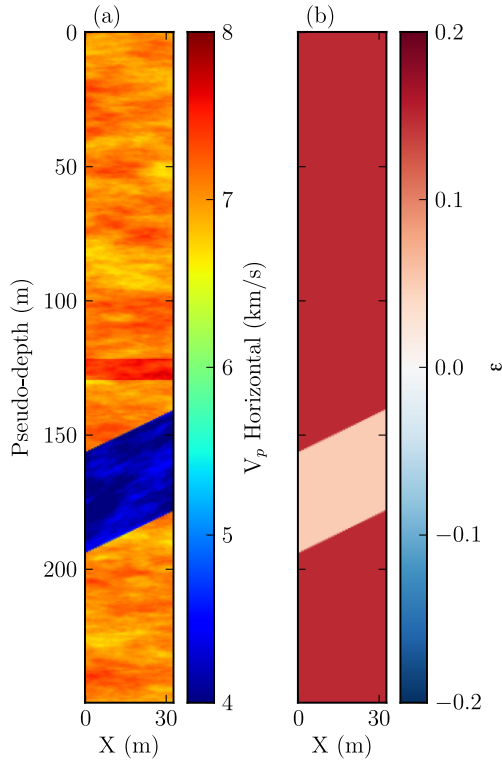
Figure 7. True models used in the synthetic tests: (a) V_p horizontal and (b) Thomsen's ε. The anisotropy symmetry system was VTI. As the model was elliptic, $\delta = \varepsilon$.
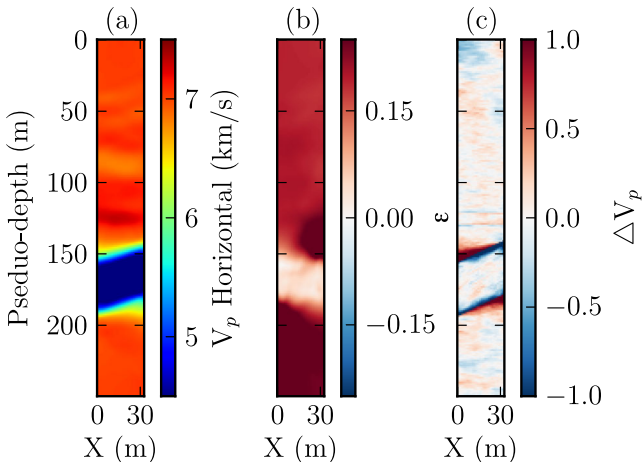

Figure 8. Traveltime tomography results for the synthetic model. (a) velocity model found through traveltime tomography; (b) ε model found through traveltime tomography; (c) velocity error ($\triangle V_p = V_{\text{true}} - V_{\text{est}}$).

to cover the frequency range of the Voisey's Bay data. The L_2 norm of the phase residuals, which refers to the imaginary part of eq. (18), was compared to the L_2 norm of the phase residuals in the previously accepted model. If the residuals in the current model were less than those in the previous model, the new model was automatically accepted. If the residuals were greater than those in the previous model, the probability of this new model being accepted (P_{accept}) was then calculated from the Metropolis criterion, given in eq. (6). A random number R was then drawn on the interval [0, 1]. If P_{accept} was greater than R, the Metropolis criterion was satisfied, and the new model was accepted. If $P_{\text{accept}} < R$ the new model was discarded. A flowchart illustrating this procedure is given in Fig. 11.

4.2.3 Determining the starting temperature

The choice of a starting temperature in simulated annealing has traditionally been problem dependent (Sen & Stoffa 1995). A useful suggestion is to choose a temperature that allows 30–50 per cent of randomly proposed models to be accepted (Tarantola 2005). To determine our starting temperature $T^{(0)}$, a 'warm up' phase was added to the simulated annealing algorithm. Throughout this phase, all perturbations to the model are accepted. Assuming that a random perturbation would be accepted 50 per cent of the time [i.e. $\mathbb{A}(\mathbf{m}^{(k+1)}, \ T^{(k+1)})$ from eq. (6) was set to 0.5], the Metropolis criterion is used to determine the temperature of an equivalent system. This procedure is time-invariant, i.e. the absolute values of $\triangle E(\mathbf{m})$ were used. After 100 iterations, the average equivalent temperature of all these perturbations is calculated, and this is set as $T^{(0)}$. This warm up phase also ensures that the starting model used for the subsequent simulated annealing varies randomly from run to run. Longer warm up periods were tested, although they were found to have a negligible effect on the starting temperature.

4.2.4 Determining the cooling schedule

The cooling schedule is set to reduce the temperature to 95 per cent of its current value when the number of perturbations which either (i) minimized the misfit or (ii) passed the Metropolis criterion, summed to 50. This value was determined by trial and error. Longer cooling schedules were also tested, although this resulted in no significant change to the final models, and vastly increased the time needed for convergence. Shorter schedules resulted in the inversions sometimes 'freezing' at local minima, and converging to unrealistic results. An 'iteration wall', where the temperature is automatically decreased regardless of the number of acceptances, is not included. Not including an iteration wall results in the temperature approaching a non-zero value as the number of iterations increases. We believe this strategy more accurately samples the posterior probability distribution, and renders the final result less susceptible to modelization error and noise.

4.2.5 Results of simulated annealing

The results of using simulated annealing to find the best-fitting 1-D elliptical anisotropy profile for the synthetic data set is shown in Fig. 12. The velocity model used is given in Fig. 8(a). The black line is the 1-D elliptical anisotropy profile found through simulated annealing. We see a good agreement between the (horizontally averaged) actual values of ε (15 per cent in the high velocity host rock zone, and 5 per cent in the low velocity mineralized zone) and those found from simulated annealing. There are some discrepancies across the boundary, although this is expected due to the 2-D nature of the anisotropy in this location.

Fig. 12(b) shows the behaviour of the value of the misfit function over the course of the inversion, as a percent of the misfit evaluated for the seed model. Also shown, in red, is the value of the misfit function obtained by using the average ε model from traveltime tomography. The horizontal axis represents the index number for all accepted models: those which either pass the Metropolis criterion, or decrease $E(\mathbf{m})$. The behaviour of $E(\mathbf{m})$ is as expected: early on the value of the objective function varies wildly with subsequent iterations. As the number of iterations increases, and the temperature $T^{(k)}$ decreases, the variations become more stable, and the objective function decreases more-or-less monotonically as it converges to its final value.
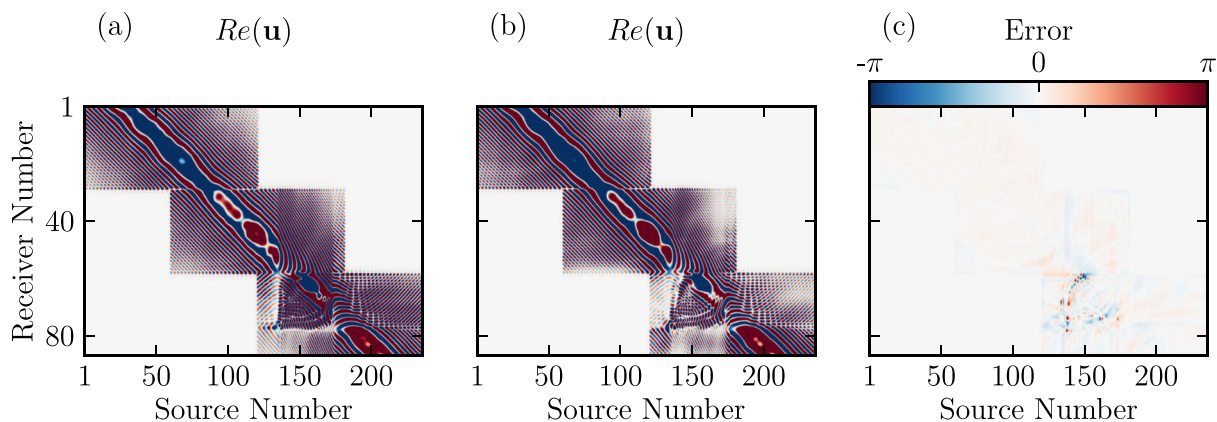
Figure 9. Frequency domain wavefield at 1300 Hz for the synthetic elastic data. (a) Original, simulated elastic, anisotropic wavefield; (b) estimated wavefield after waveform tomography, using the anisotropy profile found through simulated annealing; (c) wavefield phase residuals after waveform tomography, using the anisotropy profile found through simulated annealing.

The 1-D anisotropy profile from simulated annealing was extracted by first examining the convergence via Fig. 12(b). The dip in the objective function at model 1200 is interpreted as the point where the algorithm discovers the global minimum, and we refer to this as the 'phase change' (as it is where the model 'freezes' close to its final state). An ensemble of all models accepted after this phase change was then taken. Those models for which the misfit value lay within one standard deviation of the model with the lowest misfit were then averaged to give the model subsequently used for waveform inversion.

Figs 10(b) and (c) compare the estimated frequency domain wavefields in the vicinity of the low velocity zone before and after simulated annealing, respectively. When comparing with the true data in Fig. 10(a), we see that cycle-skipping has been eliminated at sources 120–140, between receivers 78 and 87. As well, a discontinuity at receiver 137 has been partially remedied by the new anisotropy model. Note finally the dramatic change in waveform phase around source 130 and receiver 50.

4.2.6 Velocity estimation using waveform tomography

Following the extraction of the ε profile from simulated annealing, acoustic frequency domain waveform tomography, as described in Section 2.3, was performed on the synthetic elliptical VTI data set. A sequential frequency strategy was used (Sirgue & Pratt 2004; Kamei *et al.* 2012), with each block of five iterations using four frequencies, spaced at 100 Hz intervals. After five iterations, a new frequency was added, and the lowest frequency from the previous block was dropped. The lowest frequency used was 500 Hz, and the highest was 1600 Hz. At each iteration, the amplitude and phase of the estimated source signature was updated (Pratt *et al.* 1998).

Velocity models from traveltime tomography, and from waveform tomography, along with the true model, are depicted in Fig. 13. It is apparent that the waveform tomography result using the traveltime anisotropy model (Fig. 13c) is significantly fast. In contrast, the final model from waveform tomography using the anisotropy model from simulated annealing (Fig. 13d) matches the true model (Fig. 13a) very well. 1-D velocity profiles extracted from the middle of the velocity models shown in Fig. 14, further support these observations.

Our explanation of the increased velocities obtained when the traveltime anisotropy profile is used is as follows: The geometry of the survey results in the majority of the data representing wave propagation at non-horizontal angles. Since our goal in waveform tomography is the minimization of a misfit function which is based on waveform differences, it is immediately apparent that the value of the misfit function will be significantly affected by anisotropy (as the majority of the estimated wavefield is susceptible to anisotropy). As such, the horizontal velocity plays a minimal role. If the estimated anisotropy is too strong, and if as a result the non-horizontal velocities are too slow, this can be compensated for by increasing the horizontal velocity. Since truly horizontally propagating waves contribute negligibly to the misfit function, there is not much penalty in this incorrect adjustment. We believe this is why the horizontal velocity is significantly faster when the traveltime anisotropy profile is used. Even when the anisotropy is estimated from simulated annealing, the horizontal velocities from waveform tomography are slightly too fast on average (Fig. 14). This is likely due to the fact that the anisotropy profile, while good, is not perfect, and in fact is slightly too strong. Nevertheless, the velocity model created using simulated annealing is much more accurate than the inversion using traveltime anisotropy.

Fig. 9(b) shows the final frequency domain wavefield, modelled using the results from waveform tomography. By comparing this to Fig. 9(a), we see a good match almost everywhere, that is at most sources, receivers, and offsets. Fig. 9(c) shows the residual phase differences between the modelled and true wavefield. Again, the phase differences throughout most of the model are negligible. However, for sources and receivers near the edges of the low velocity zone, we see a significant increase in wavefield phase residuals. This is expected, as the velocity heterogeneity in this area is extremely strong, and elastic mode conversions, which are not modelled in the acoustic inverse routine, are likely affecting the data significantly. As a result, it is also in this area where the estimated anisotropy is furthest from its true value.

5 SIMULATED ANNEALING—APPLICATION TO REAL CROSSHOLE DATA

5.1 Anisotropy estimation by simulated annealing of waveform data

Following the development of the simulated annealing workflow described above, we now describe the application to the original real data from Voisey's Bay. The resulting best fit 1-D elliptical
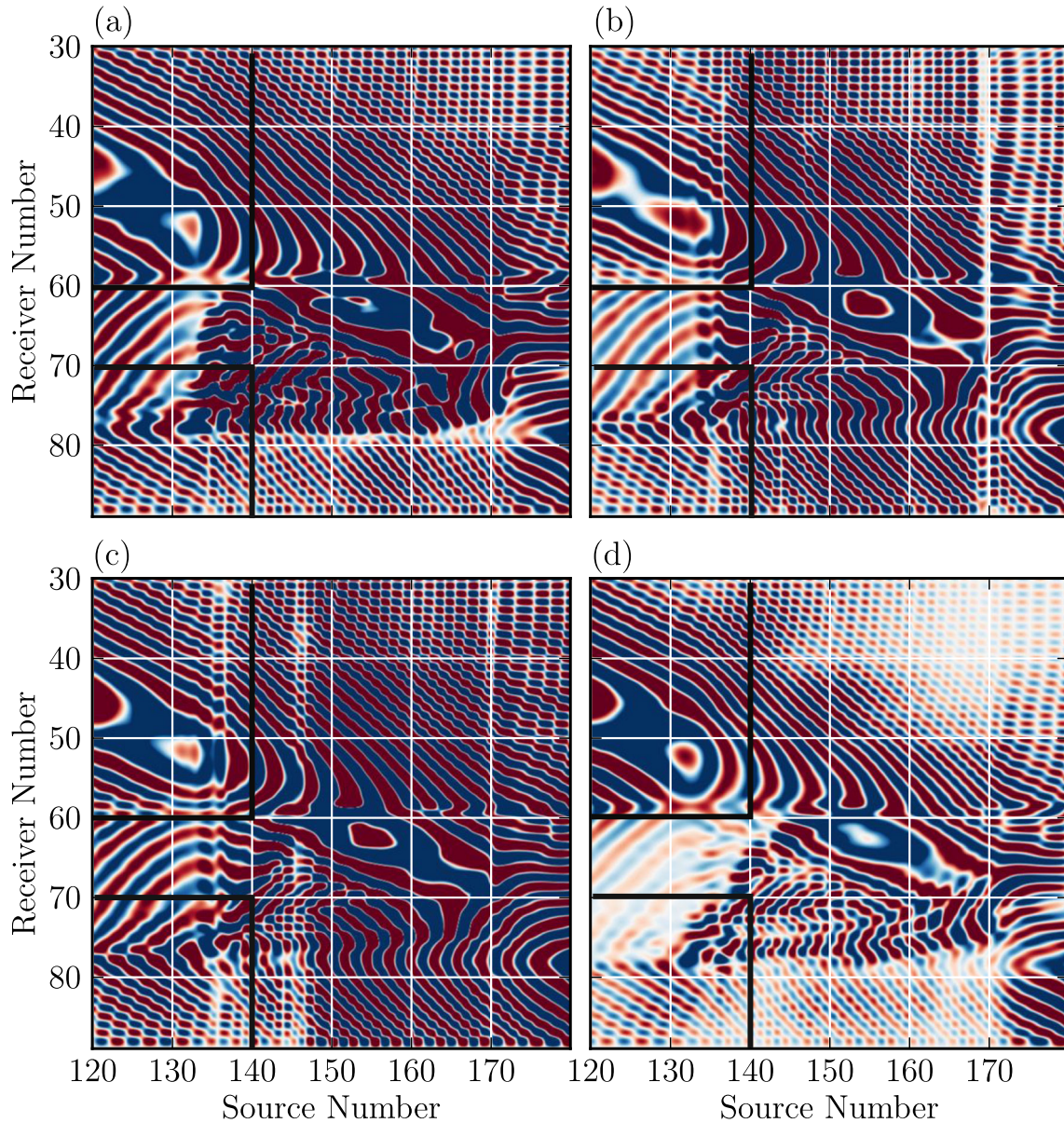
Figure 10. A subset of the frequency domain wavefield [$Re(\mathbf{u})$] at 1600 Hz for the synthetic elastic data. The subset represents a 'zoomed' view in the area surrounding the low velocity 'mineralized' zone. (a) Original simulated elastic, anisotropic wavefield; (b) estimated wavefield using the traveltime velocity model, with the anisotropy model found from traveltime tomography; (c) estimated wavefield using the traveltime velocity model, with the anisotropy model found through simulated annealing; (d) estimated wavefield after waveform tomography, with the anisotropy model found through simulated annealing. To identify cycle-skipping in these images, compare the values of $Re(\mathbf{u})$ from one panel to the next. For example, if you follow lines of red or blue in the top left corner of panel (a), you will see that they are continuous. If the same area is examined in panel (b), a discontinuity can be seen. This discontinuity may be a sign of a cycle-skip. The areas within the black boxes show the most significant improvement.

anisotropy profile found via simulated annealing of the Laplace–Fourier waveforms from the real data is shown in Fig. 15. The parameters and frequencies used for simulated annealing algorithm were identical to those used during the synthetic tests, with the exception of τ which was set to 0.01. The velocity model used is given in Fig. 3(a).

In Fig. 15, the black line is the 1-D elliptical anisotropy profile found through simulated annealing. As described in the previous section, the grey area represents those anisotropy models which lay within one standard deviation of the absolute best fit model. Fig. 15(b) shows the behaviour of the value of the misfit function over the course of the inversion, as a percentage of the misfit evaluated for the seed model. Also shown, in red, is the value of

the misfit function obtained by using the average ε model from traveltime tomography. The horizontal axis represents all accepted models: those which either pass the Metropolis criterion, or decrease the misfit function. The behaviour of misfit function is as expected from the synthetic examples described above. As stated before, since it is unlikely that the true anisotropy in the target region is elliptical, the values of ε obtained should be interpreted as a geometric correction factor for anisotropy, and not as a robust estimate of the true anisotropy.

The benefits of the 'crawling' approach to model perturbations, that is perturbing *one* model parameter per iteration instead of *all*, becomes apparent when viewing a time-lapse of the accepted models. In areas where the model is well constrained (within the
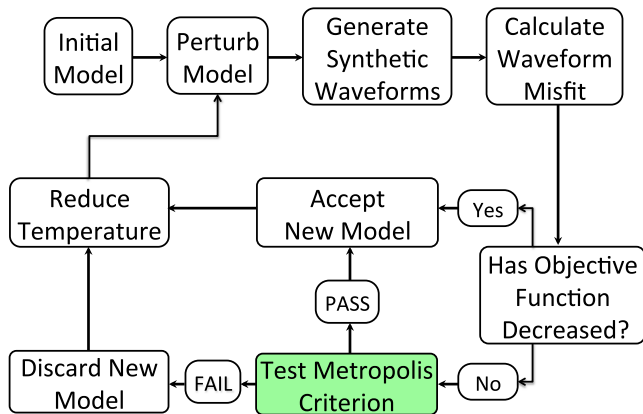
Figure 11. Flowchart illustrating the simulated annealing algorithm. The initial 1-D anisotropy model is randomly generated during the warm up phase, and perturbations to this model are again random. Synthetic waveforms are generated using the acoustic wave equation, and the waveform misfit is defined as the L_2 phase residuals at 1000, 1300 and 1600 Hz. Depending on the value of the misfit, the perturbed model is either accepted, or subjected to the Metropolis criterion (eq. 6). When the sum of perturbations which either (i) minimizes the misfit or (ii) passes the Metropolis criterion, reaches 50, the temperature is decreased to 95 per cent of its current value.

high-velocity troctolites, for example), the local anisotropy values converge, or 'freeze' close to their final states very quickly. The algorithm is then free to explore the subtleties of the misfit function in the vicinity of the less constrained parts of the model (i.e. in the low velocity target zone), while retaining information on those well constrained portions of the model. If we had chosen to perturb all model parameters at each iteration, this information would not be preserved. While adopting the latter approach would result in a less-biased sampling of the misfit function, the crawling approach allows us to independently perturb the model parameters which are most uncertain. It should further be noted that the warm up phase of

our approach ensures that the starting model is completely random, so no footprint of the seed model is present in these results.

Fig. 16(c) illustrates the effect of the new anisotropy model derived from simulated annealing on the modelled time domain waveforms. Much of the misfit at larger offsets has now been explained; most importantly, cycle-skipping is no longer apparent on the time domain wavefields. These same improvements can also be seen in the frequency domain wavefields: Fig. 16(c) shows a zoomed in subset of the estimated frequency domain wavefield at 1200 Hz after simulated annealing. Comparing this with the real data in Fig. 16(a), and that modelled with the traveltime anisotropy model in Fig. 16(b), we see some significant improvements. First, at near offsets (between sources 120 and 130, and receivers 40 and 50) we see that the wavefield post-simulated annealing matches the real data more closely than it did with the traveltime anisotropy model. As well, just above the low velocity zone (shots 120–140, and receivers 30–40) the phase in Fig. 16(c) is closer to the real data, and does not contain the rapid variations seen in Fig. 16(b).

5.2 Waveform tomography with anisotropy from simulated annealing

Once the initial model satisfied the half cycle criterion, we were able to proceed confidently with waveform tomography with the Voisey's Bay data. For waveform tomography, a sequential frequency selection strategy was employed (Sirgue & Pratt 2004; Kamei *et al.* 2012), in which blocks of four frequencies were inverted at a time, beginning with 1000 Hz (the lowest, clean frequency), and increasing in 100 Hz intervals. After five iterations, a new block of frequencies was inverted, retaining the highest three frequencies from the preceding block. This was continued until the final block containing 1600 Hz data, the highest frequency used in the inversions, was inverted. At each iteration, the phase and amplitude of the averaged source wavelet was updated. No gradient preconditioning techniques (such as smoothing, or wavenumber filtering) were
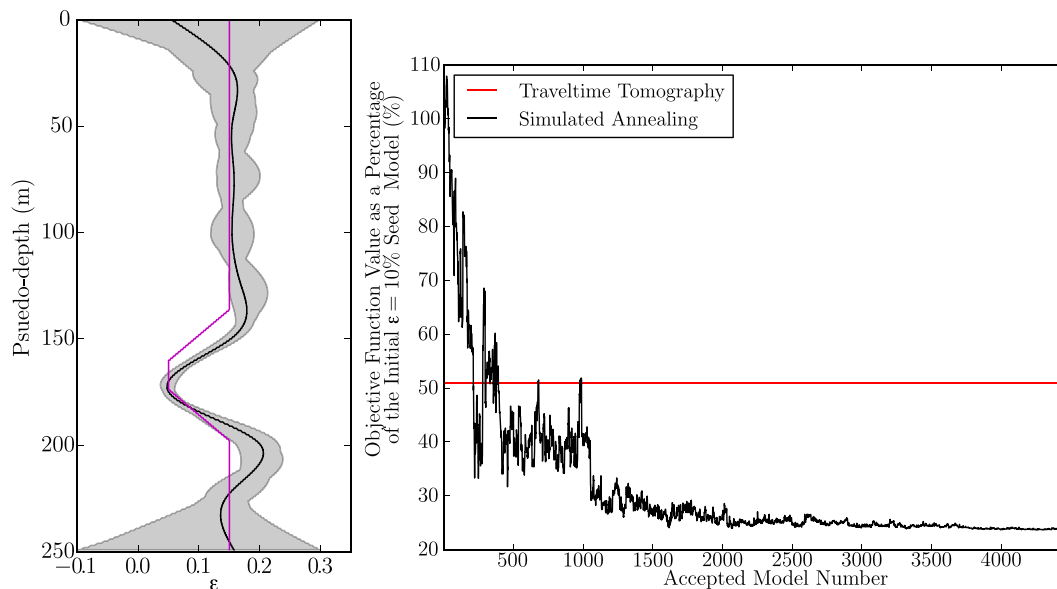

Figure 12. Left-hand panel: best-fitting ε profile obtained from simulated annealing (black) for the synthetic data, using data frequencies 1000, 1300 and 1600 Hz. The shaded area represents the range of anisotropy profiles that lay within 1 SD of the absolute best-fitting profile. The true, horizontally averaged ε profile is shown in purple. Right-hand panel: objective function behaviour over the course of simulated annealing (black), compared to the value obtained in the traveltime tomography model (red). The best-fitting profile from simulated annealing was constructed by considering all models after the 'phase change' around iteration 1000.
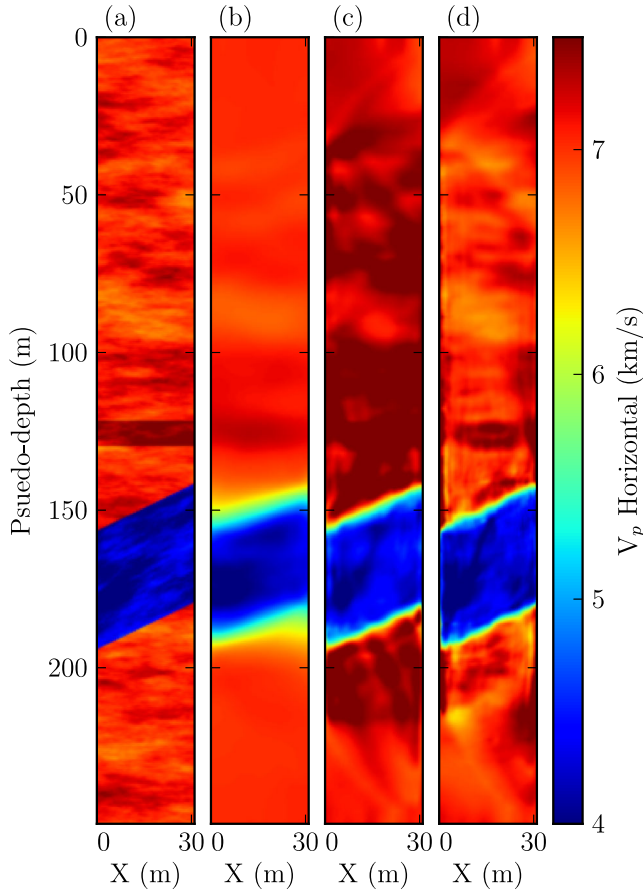
Figure 13. Velocities in the synthetic model: (a) True, horizontal velocity (as in Fig. 7); (b) velocity model found through anisotropic traveltime tomography; (c) velocity model found through waveform tomography, using the anisotropy profile from traveltime tomography; (d) velocity model found through waveform tomography, using the anisotropy profile from simulated annealing.

used. The resultant velocity model from waveform tomography can be seen in Fig. 17, along with vertical velocity profiles extracted from near the source and receiver positions, and a comparison to the sonic well log velocities. A predicted receiver gather computed using this velocity model can be seen in Fig. 6, and the corresponding 1200 Hz frequency domain data are shown in Fig. 16(d).

A comparison of the real and estimated wavefields, as well as the final wavefield residuals calculated at 1200 Hz using the final velocity model from waveform tomography, is given in Fig. 18. Through layouts 1 and 2, the waveform residuals are acceptable and the phase residuals are largely much less than π as we would hope, although a significant phase error is present along the main diagonal, representing the near offsets. These are attributed to the presence of tubes waves in the near-offset data, which are not modelled in the inversions (at near offsets, the tubes waves are consistently mixed in with the first arrivals, but as the offsets increase, the slow-moving tube waves are excluded from the data by the time window). Also apparent in Fig. 18(c) are some vertical stripes, which are attributed to misfits introduced by source-borehole coupling problems, or possible source mislocation errors. Finally, in these two layouts, a trend in phase mismatch is present. For source receiver pairs where the source is located below the receiver, the phase error is dominantly negative, while for pairs where the receiver is below the source, the error is dominantly positive.

This suggests that the error is dependent on angle, and that our anisotropy model may not be entirely correct. This is expected due to the simplified 1-D elliptical approximation with vertical symmetry.

In the third (deeper) layout (largely within the low velocity mineralized zone), the results are less encouraging. Semi-circular rings of constant phase mismatch are apparent, and these are indicative of cycle-skipping at 1200 Hz. This is not entirely unexpected, as these particular data (shots 130–180, shooting to receivers 60–87) are the most difficult; they arise from wave paths which spend the longest time within the dipping low velocity zone. Since this zone is where the 1-D anisotropy assumption is most likely to be invalid, and since this is the region most corrupted by tube waves, we expect these data to be the most difficult to fit. Interestingly, a similar residual pattern is also seen in the noise-free synthetic data (Fig. 9c), suggesting that these residuals are also partially related to the challenges of imaging the strong velocity discontinuity. Fig. 16d shows a zoomed-in view of the estimated 1200 Hz frequency domain wavefield in the low velocity zone, calculated using the velocity model from waveform tomography.

Borehole sonic velocities were available from each of the boreholes, and are depicted along with the equivalent vertical velocities extracted from the waveform tomography velocity model in Fig. 17. The sonic velocities measure the seismic velocity along the near vertical boreholes, so the tomographic velocities must be corrected for anisotropy before these data can be compared to the horizontal velocities in Fig. 17. Since we are assuming elliptical anisotropy, an estimate of the vertical velocity can be obtained by dividing the horizontal velocities in each layer by $\sqrt{1 + 2\varepsilon} \approx (1 + \varepsilon)$ (for weak anisotropy). These equivalent vertical velocities from the waveform tomography are plotted in blue Fig. 17, together with the sonic velocities plotted in green. The match is good, although not exact throughout the model, however this is expected: borehole sonic velocities are usually measured at approximately 20 kHz, while our inversions were for much lower frequencies. As well, the 1-D elliptic VTI assumption is likely invalidated by the complex shearing and fracturing present. This may lead to not only image distortion, but also to errors in calculating the equivalent vertical velocities (as the symmetry axis may be at some other, arbitrary angle). In the source hole, an apparent static shift in vertical velocities is apparent, where the inverted velocities are consistently faster than the sonic velocities, although they show the same major features. This may be a result of the true, 2-D anisotropy structure.

Despite these caveats on the potential differences between waveform tomography velocities and sonic velocities, some promising observations emerge. In the receiver hole, a sharp decrease is seen in the sonic velocity at approximately 675 m. This is echoed closely by a corresponding decrease in the inverted velocity. By examining the borehole core logs, it was possible to tie these velocity anomalies to a 1-m-thick massive sulphide veinlet. Further examination of the image allows us to trace the continuity of this zone of relatively low velocity: it traverses the panel, and appears to split near the source hole. This is also in agreement with the core logs (Inco 2000), on which it was observed that brecciated host rock in this zone give way to local irregular massive sulphide veins. The high velocities on either side of the vein at the receiver well exhibit a good match between the sonic and inverted velocities, and appear to be geologically reasonable structures, which however do not appear to extend to the source well (as was expected from the sonic velocities). Just below 650 m, a zone of relatively low velocity can be traced across the panel. While this agrees well with the sonic velocities, there is no corresponding geologic signature noted in the core logs.
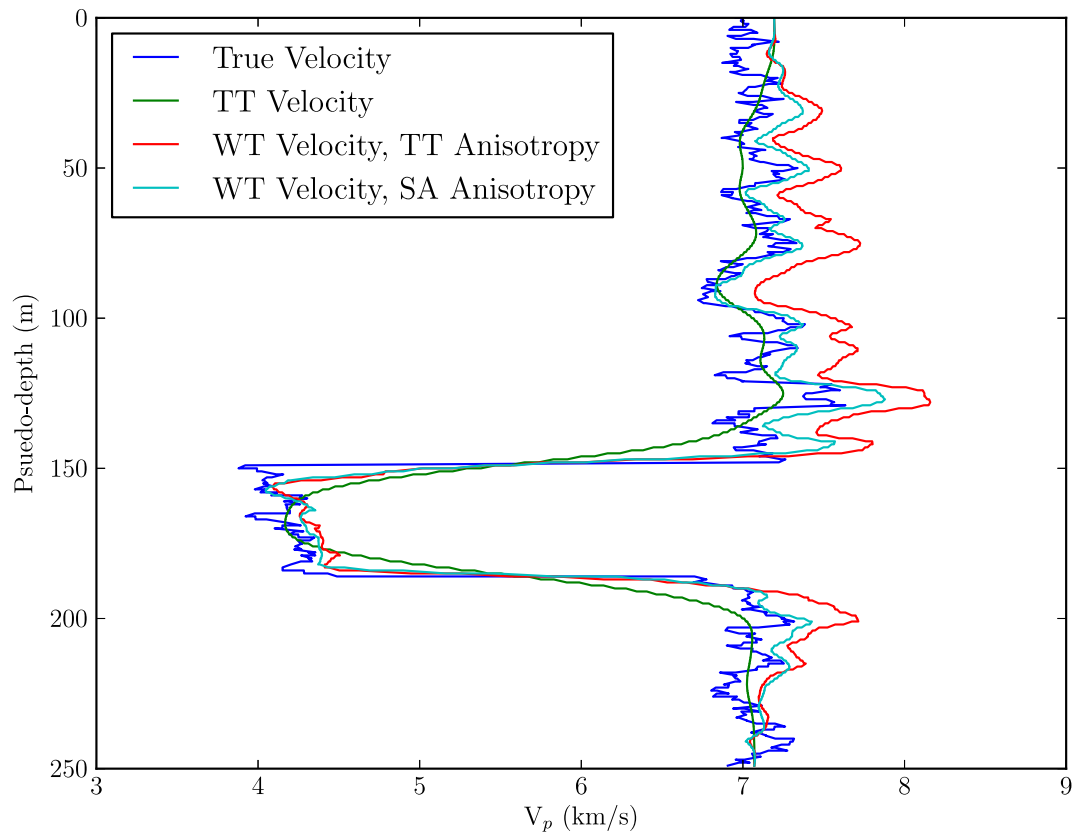
Figure 14. Vertical profiles taken through the middle of the horizontal velocity models shown in Fig. 13. TT refers to traveltime tomography, WT to waveform tomography, and SA to simulated annealing.
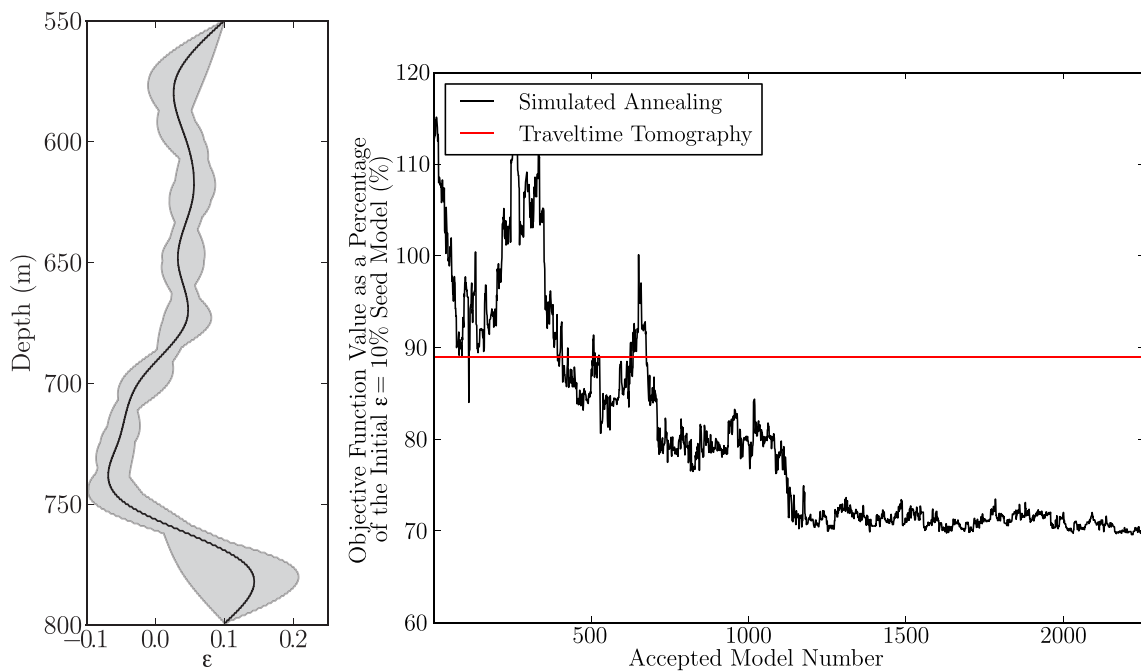

Figure 15. Left-hand panel: best-fitting ε profile obtained from applying simulated annealing to the Voisey's Bay data set, using data frequencies 1000, 1300 and 1600 Hz. The shaded area represents the range of anisotropy profiles that lay within one standard deviation of the absolute best-fitting profile. Right-hand panel: objective function behaviour over the course of simulated annealing (black), compared to the value obtained from traveltime tomography (red). The best-fitting profile from simulated annealing was constructed by considering all models after the 'phase change' around iteration 1000.
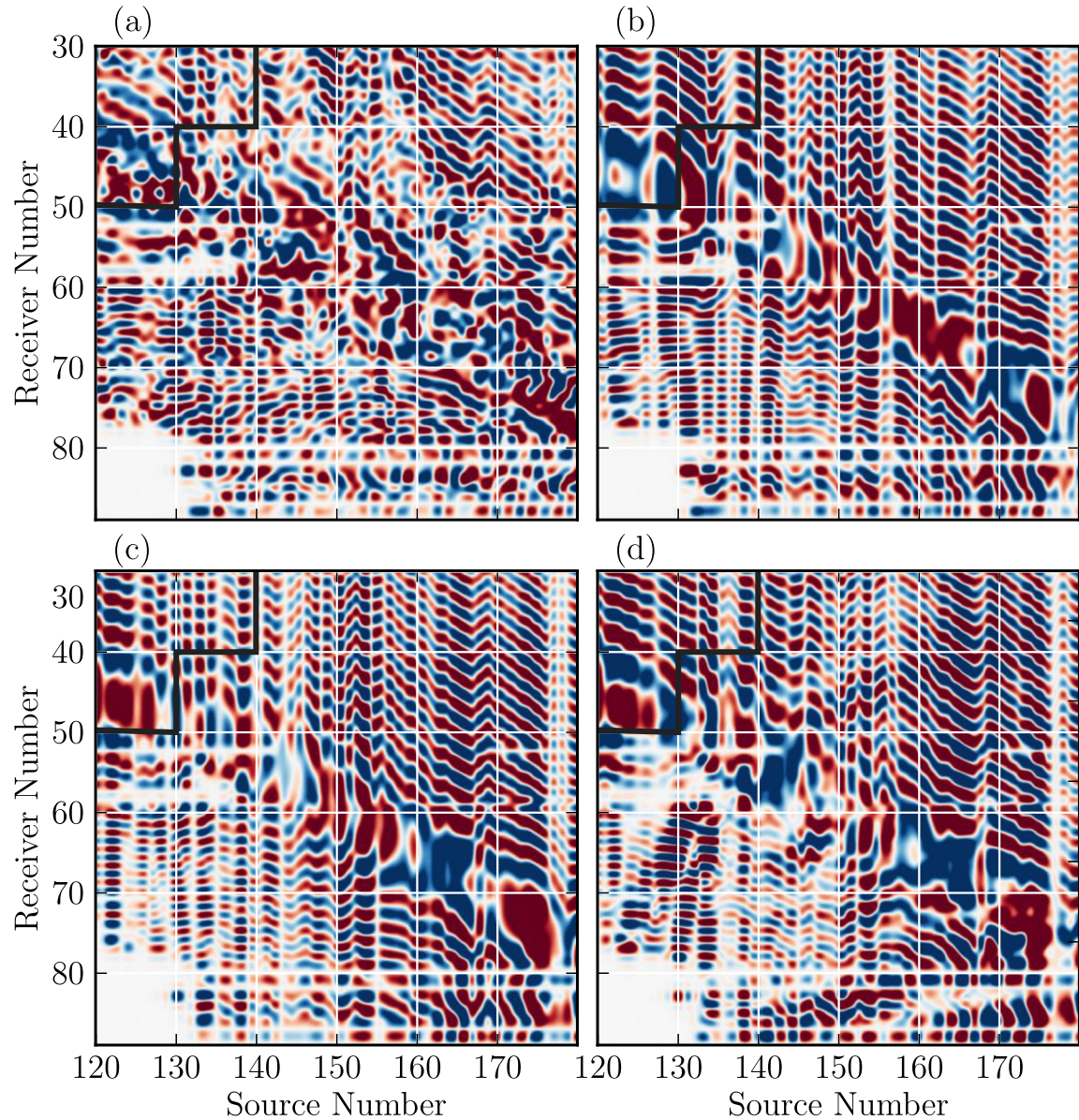
Figure 16. A subset of the frequency domain wavefields at 1200 Hz for the real crosshole data, zoomed in on the area surrounding the low velocity zone. (a) Observed wavefield; (b) estimated wavefield using the traveltime velocity model, and the anisotropy model found from traveltime tomography; (c) estimated wavefield using the traveltime velocity model, and the anisotropy model found through simulated annealing; (d) estimated wavefield after waveform tomography, using the anisotropy model found through simulated annealing. The areas within the black boxes show the most significant improvement.

Finally, at the top edge of the low velocity zone, it is very apparent that the velocity gradient in the source well is much stronger than that in the receiver well. This matches with the sonic velocities, and also has geologic significance. In the receiver well, significant quantities (40–75 per cent) of massive sulphide are present in the troctolites beginning at 689 m, which results in a decrease in velocity. Also, between 697 and 700 m, the existence of a granite intrusion is noted, and this may be represented as a slight levelling off of the velocity gradient in both the sonic and inverted velocities. While the receiver well intersects a zone where the massive sulphide content increases gradationally, the source well experiences a much sharper geological contrast. Between 683 and 686 m, the sulphide content increases to 40 per cent, with the massive sulphide zone beginning directly below.

The match between the sonic velocities and inverted velocities within the massive sulphide zone is slightly less promising than

the match outside this zone. While the overall structure and velocities match well, there are some discrepancies when examining fine structure. We attribute this partially to the breakdown of the 1-D anisotropy assumption, when combined with the 2-D massive sulphide zone. The borehole logs report on significant off-axis fracture systems (30°, 50° and 60° to the core axis) which also likely leads to the failure of the VTI elliptic approximation. We also expect artifacts due to the presence of tube waves in the data, which are especially apparent within the low velocity zone.

The comparison of the inverted velocities and borehole logs allow us to comment on the resolution of our final image. Theoretically, we expect the resolution of waveform tomography to be on the order of a half-wavelength of the highest frequency used in the inversion. Our maximum resolution is therefore expected to occur in the low velocity massive sulphides, during the inversion for 1600 Hz. These conditions should result in a maximum local resolution length of
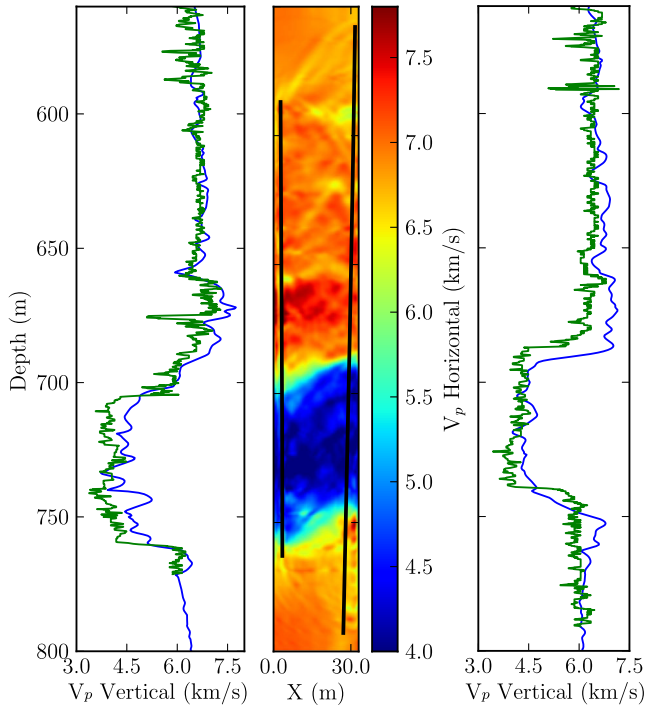
Figure 17. Final velocity model obtained from waveform tomography, using the anisotropy profile found from simulated annealing. Comparisons between the inverted vertical velocities (blue) and the vertical sonic velocities obtained from the borehole logs (green) are shown for the source hole (left-hand side) and receiver hole (right-hand side).

approximately 1 m. The appearance of the 1-m-thick massive sulphide veinlet in the final image from waveform tomography is then in accordance with this estimate of resolution. Throughout the model the resolution will vary depending on the local wave speed and data coverage.

5.3 Discussion

It is worth investigating why it is that waveform tomography, combined with simulated annealing, succeeds in generating appropriate anisotropy models, while anisotropic traveltime tomography was less successful for these examples. To answer this question, let us

look more closely at the synthetic data set for the case of elliptical VTI anisotropy.

5.3.1 Anisotropic lobes

Changing the constraint on perturbations to anisotropy produced completely different ε models, but the resultant velocity models were quite similar. This suggests that some of the anisotropy parameters are within the null space of the traveltime tomography problem (a well understood notion), and do not have much of an effect on the traveltime fit. Nevertheless, the slight reduction in traveltime residuals, and the consistent manner in which the anisotropy models behave in the presence of a strong velocity contrast, lead us to believe that certain situations squeeze the normally vastly underdetermined anisotropy parameters into the range of the traveltime tomography inverse problem. Pratt & Chapman (1992) achieve this by enforcing smoothness constraints, but here we make the case that the fundamental physical assumptions of classic traveltime tomography make the anisotropy parameters very difficult to estimate under the conditions encountered with the real data within the Eastern Deeps deposit.

Fig. 19 illustrates the reason for the success of waveform based simulated annealing. As discussed in the introduction, waveforms are influenced not only by velocities along the ray path, but by velocities within the spatially broader wavepath. The underlying red-white-blue image in Fig. 19 is the $V_{\rm true} - V_{\rm est}$ from Fig. 8(c). This image is overlain with a single, curved seismic ray path used in traveltime tomography (yellow); the corresponding wave path from 500 Hz waveform tomography is also superimposed using a greyscale. Note the magnitude and sign of the velocity error sampled by either the ray path or wave paths: Between this specific source-receiver pair, the ray is only sensitive to velocities which are on average too fast. This is a result of the unresolvability of the sharp velocity boundary by the ray-theoretical approximation, which stems from the fundamental resolution limit of ray-based inversion. As well, we have placed constraints on the roughness of the model, which serves to further blur this boundary. These two effects result in the velocity residual pattern seen in Fig. 19, as the velocity discontinuity is smoothed out over the boundary.

Infinitely thin rays, traced within erroneous velocity models, will experience a corresponding increase or decrease in traveltime. In most of the model, we see that this is not a major issue, as positive and negative errors tend to average out by the time the ray
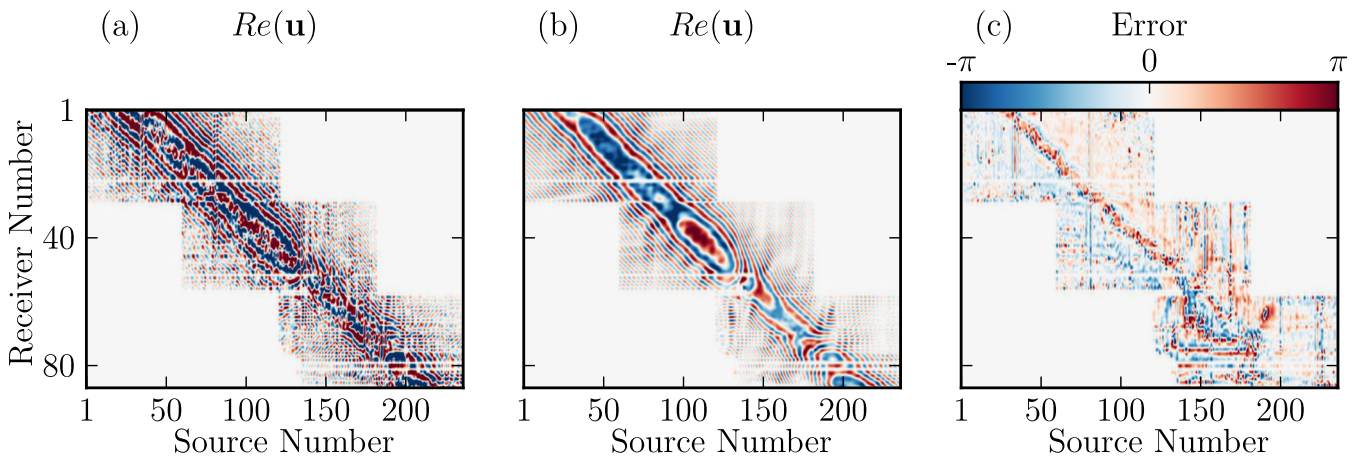

Figure 18. Frequency domain wavefields [$Re(\mathbf{u})$] for the Voisey's Bay data at 1200 Hz (a) Real data after data pre-conditioning (see Fig. 4); (b) predicted data calculated using the final velocity model from waveform tomography; (c) phase residuals in the final model.
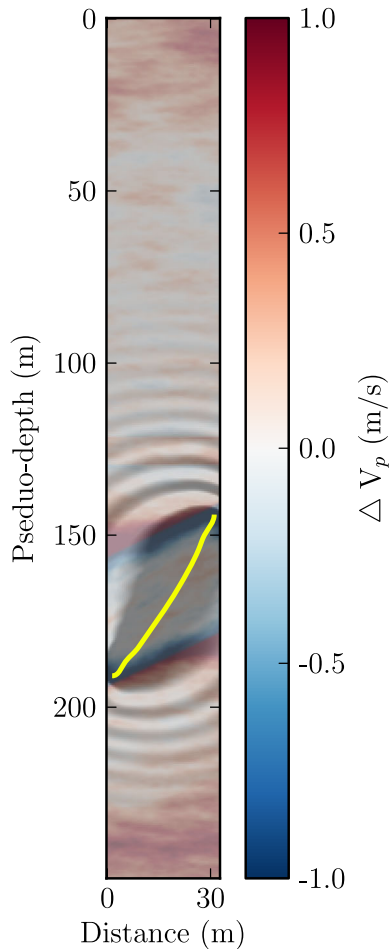
increased anisotropy appears in the low velocity zone, corresponding with those highly angled ray paths. In contrast, if there is *no* constraint on anisotropy, the lobe anisotropy is allowed to grow arbitrarily large, and the traveltime errors can be accounted for without invoking anisotropy in the low velocity zone.

5.3.2 Bulk shifts in anisotropy

There is another interesting feature in the traveltime anisotropy models, and that is the difference in anisotropy in the high velocity zone between Fig. 8(b) and the true value. This is likely due to the presence of horizontal layering in the true velocity models, which is not resolvable at the resolutions of traveltime tomography. Backus (1962) showed that a horizontally layered medium is indistinguishable from a homogenous anisotropic medium in the long wavelength limit. Since the horizontal layering in the true model is beyond the resolution of traveltime tomography, the inverted velocity model is relatively homogeneous compared to the true model, and thus exhibits *extrinsic*, or additional, anisotropy beyond the *intrinsic* value of $\varepsilon = 15$ per cent. Pratt & Sams (1996) showed the effects of extrinsic anisotropy on traveltime tomography, and they concluded that fine layering beyond the resolution limit may indeed appear as extrinsically anisotropic when inverted using the ray-theoretical approach.

5.3.3 Wave paths versus ray paths

The question remains as to why simulated annealing, which uses a finite-difference implementation of the acoustic wave equation, succeeds in determining the correct (1-D) anisotropy model, whereas traveltime tomography does not. On Fig. 19, the wave paths for a 500 Hz wave is depicted in greyscale. Comparing this to the equivalent ray path, we see a significant difference: the multipath nature of the wave equation allows velocities external to the ray path to be sampled. In this particular case, the wave propagates outside the low velocity zone, experiencing areas of both positive and negative velocity errors. By the time the wave arrives at the receiver, the cumulative error along the wave paths results in the predicted phase to be roughly equivalent to the observed phase, thus negating the need to perturb any model parameters. As well, in the high velocity zone where the anisotropy was overestimated by traveltime tomography, information contained in the waveforms supports the notion that the medium is finely layered, instead of being extrinsically anisotropic.

5.3.4 Computational cost

The additional computational cost associated with simulated annealing was not extreme. The routine was parallelized over frequencies and sources, and each iteration took approximately 4–5 s when run on 16 Intel Xeon processors with a clock speed of 2.83 GHz. To generate Figs 12 and 15, approximately 50 000 trial models were generated, resulting in a total compute time on the order of 50 hr.

6 CONCLUSIONS

In this paper, we demonstrated the application of waveform-based simulated annealing to the problem of estimating seismic anisotropy both in a real 2-D data example from a mining environment, and in a 2-D synthetic model. To our knowledge this is the first time such a method has been successfully used. The global nature of the simulated annealing routine allowed us to bypass the time-intensive

Figure 19. Ray path (yellow) and 500 Hz wave path (greyscale), for a single source–receiver pair, superimposed on the traveltime inversion velocity error ($\triangle V_p = V_{\text{true}} - V_{\text{est}}$) from Fig. 8(c).

arrives at the receiver. The ray in Fig. 19 however, is part of a small subset of rays which travel mostly within areas in which the cumulative velocity error along the ray path is positive, and which as a result carry large, positive traveltime residuals. To remedy this mismatch in traveltime error, either the velocity or the anisotropy must be perturbed. Since the inverted velocity in the majority of the low velocity zone is roughly correct, perturbing velocities along these troublesome ray paths will not solve the problem, as it will increase the residuals between other source-receiver pairs. Instead, it appears that the traveltime tomographic algorithm resolves these errors through perturbations to the anisotropy model.

To see these effects, let us re-examine the inverted anisotropy model for the synthetic tests closely (Fig. 8b). While the correct value of anisotropy is resolved for the the majority of the low velocity zone, highly anisotropic lobes appear at the velocity discontinuity. In other tests where perturbations to anisotropy were heavily constrained, the magnitude of anisotropy *within* the low velocity zone increased, away from the correct value, but the size and magnitude of the lobe-like structures decreased.

These phenomenon arise as follows: If the anisotropy is heavily constrained, there is a significant penalty on any anisotropic model parameters which deviate from zero (this is equivalent to assuming *a priori* that the region is isotropic, and heavily weighting this information). As such, the magnitude of the lobes does not increase to the values seen in for the unconstrained case (Fig. 8b). Instead,

job of building a starting anisotropy model through traveltime tomography. This negated the need to choose the relative weighting of regularization parameters governing perturbations to velocity and anisotropy, an often tedious and subjective process. In fact, an analysis of the behaviour of seismic rays in areas of strong velocity contrast led to the realization that in certain situations the correct estimation of anisotropy may be impossible. This is due to the imperfect resolution of traveltime tomography and the limitations of ray theory. In such situations, the use of the wave equation in forward modelling becomes a necessity. This paper also served to investigate the applicability of acoustic waveform tomography to the highly heterogeneous hardrock environments frequently encountered in mineral exploration. The successful correlation of the final velocity model with independently collected borehole sonic logs is exciting, and motivates future work in this field.

Waveform tomography represents a key component of the state-of-the-art of seismic tomography across the scales. With current computational power and prowess, we are now able to accurately simulate the passage of seismic waves through a strongly heterogeneous and anisotropic Earth. Global inverse methods treat the seismic inverse problem as it really is, underdetermined and non-unique, while mitigating the pitfalls associated with incorporating often subjective *a priori* information. While their applicability to higher-dimensional parameter spaces is a topic for future work, we hope that this paper represents a step forward in uniting waveform tomography with global inverse methods, leading to a global, objective description of our state of knowledge of Earth's internal structure.

REFERENCES

Alkhalifah, T., 2000. An acoustic wave equation for anisotropic media, *Geophysics*, **65**(4), 1239–1250.

Babuska, V., 1968. Elastic anisotropy of igneous and metamorphic rocks, *Stud. Geophys. Geod.*, **12**(3), 291–303.

Backus, G.E., 1962. Long-wave elastic anisotropy produced by horizontal layering, *J. geophys. Res.*, **67**(11), 4427–4440.

Brenders, A.J. & Pratt, R.G., 2007a. Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model, *Geophys. J. Int.*, **168**, 133–151.

Brenders, A.J. & Pratt, R.G., 2007b. Efficient waveform tomography for lithospheric imaging: implications for realistic, two-dimensional acquisition geometries and low-frequency data, *Geophys. J. Int.*, **168**, 152–170.

Brossier, R., Gholami, Y., Virieux, J. & Operto, S., 2010. 2D frequency-domain seismic wave modeling in VTI media based on a Hp-adaptive discontinuous Galerkin method, in *Proceedings of the 72nd EAGE Conference & Exhibition*, Extended abstract, SPE, EAGE.

Červeny, V., 1972. Seismic rays and ray intensities in inhomogeneous anisotropic media, *Geophys. J. R. astr. Soc.*, **29**, 1–13.

Červeny, V. & Jech, J., 1982. Linearized solutions of kinematic problems of seismic body waves in inhomogeneous slightly anisotropic media, *J. Geophys.*, **51**, 96–104.

Chapman, C.H. & Pratt, R.G., 1992. Traveltime tomography in anisotropic media I. Theory, *Geophys. J. Int.*, **109**(1), 1–19.

Chen, S.T., Zimmerman, L.J. & Tugnait, J.K., 1990. Subsurface imaging using reversed vertical seismic profiling and crosshole tomographic methods, *Geophysics*, **55**(11), 1478–1487.

Colton, D. & Kress, R., 1998. *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer.

Cosma, C. & Enescu, N., 2003. Ore delineation by crosshole seismics, transmission and reflection imaging at Voisey's Bay, Canada, in *Proceedings of the 7th International Symposium on Recent Advances in Exploration Geophysics*, Extended Abstracts.

Dellinger, J.A., 1991. Anisotropic seismic wave propagation, *PhD thesis*, Stanford University.

Eaton, D.W., Adam, E., Milkereit, B., Salisbury, M.H., Roberts, B., White, D. & Wright, J., 2010. Enhancing base-metal exploration with seismic imaging, *Can. J. Earth Sci.*, **47**(5), 741–760.

Enescu, N., McDowell, G., Cosma, C. & Bell, C., 2002. Crosshole seismic investigations at Voisey's Bay, Canada, in *Proceedings of the 72nd Annueal SEG Meeting*, Extended Abstracts.

Evans-Lamswood, D.M., Butt, D.P., Jackson, R.S., Lee, D.V., Muggridge, M.G. & Wheeler, R.I., 2000. Physical controls associated with the distribution of sulfides in the Voisey's Bay Ni-Cu-Co deposit, Labrador, *Econ. Geol.*, **95**, 749–769.

Ikelle, L., Yung, S. & Daube, F., 1993. 2-D random media with ellipsoidal autocorrelation functions, *Geophysics*, **58**(9), 1359–1372.

Inco, 2000. INCO Technical Services Limited: Borehole core log.

Iturrino, G.J., Christensen, N.I., Kirby, S. & Salisbury, M.H., 1991. 11. Seismic velocities and elastic properties of oceanic gabbroic rocks from hole 735B, *Proc. Ocean Drilling Prog., Scient. Results*, **118**, 227–244.

Kamei, R. & Pratt, R.G., 2013. Inversion strategies for visco-acoustic waveform inversion, *Geophys. J. Int.*, pp. 859–884.

Kamei, R., Hato, M. & Matsuoka, T., 2005. Random heterogeneous model with bimodal velocity distribution for Methane Hydrate exploration, *Explor. Geophys.*, **58**(1), 41–49.

Kamei, R., Pratt, R.G. & Tsuji, T., 2012. Waveform tomography imaging of a megasplay fault system in the seismogenic Nankai subduction zone, *Earth planet. Sci. Lett.*, **317-318**, 343–353.

Kamei, R., Pratt, R.G. & Tsuji, T., 2013. On acoustic waveform tomography of wide-angle OBS data—strategies for pre-conditioning and inversion, *Geophys. J. Int.*, **194**(3), 1250–1280.

Kaufl, P., Fichtner, A. & Igel, H., 2013. Probabilistic full waveform inversion based on tectonic regionalization—development and application to the Australian upper mantle, *Geophys. J. Int.*, **193**(1), 437–451.

Kern, H., 1993. P- and S-wave anisotropy and shear-wave splitting at pressure and temperature in possible mantle rocks and their relation to rock fabric, *Phys. Earth planet. Inter.*, **78**, 245–256.

Kirkpatrick, S., Gelatt, C.D. & Vecchi, M.P., 1987. Optimization by simulated annealing, *Science*, **220**(4598), 671–680.

Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migrations, in *Proceedings of the Conference on Inverse Scattering: Theory and Application*, pp. 206–220, SIAM Philadelphia.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. & E., T., 1953. Equation of state calculations by fast computing machines, *J. Chem. Phys.*, **21**, 1087–1092.

Mosegaard, K. & Vestergaarad, P.D., 1991. A simulated annealing approach to seismic model optimization with sparse prior information, *Geophys. Prospect.*, **39**, 599–611.

Nulton, J.D. & Salamon, P., 1988. Statistical mechanics of combinatorial optimization, *Phys. Rev. A*, **37**(4), doi:10.1103/PhysRevA.37.1351.

Paige, C.C. & Saunders, M.A., 1982. LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Software*, **8**(1), 43–71.

Park, C.B., Miller, R.D., Steeples, D.W. & Black, R.A., 1996. Swept impact seismic technique (SIST), *Geophysics*, **61**(6), 1789–1803.

Perozzi, L., Gloaguen, E., Rondenay, S. & McDowell, G.M., 2012. Using stochastic crosshole seismic velocity tomography and Bayesian simulation to estimate Ni grades: case study from Voisey's Bay, Canada, *J. appl. Geophys.*, **78**, 85–93.

Plessix, R.-E. & Perkins, C., 2010. Full waveform inversion of a deep water ocean bottom seismometer dataset, *First Break*, **28**, 71–78.

Pratt, R.G. & Chapman, C.H., 1992. Traveltime tomography in anisotropic media II. Application, *Geophys. J. Int.*, **109**(1), 20–37.

Pratt, R.G. & Sams, M.S., 1996. Reconciliation of crosshole seismic velocities with well information in a layered sedimentary environment, *Geophysics*, **61**(2), 549–560.

Pratt, R.G., McGaughey, W.J. & Chapman, C.H., 1993. Anisotropic velocity tomography: a case study in near-surface rock mass, *Geophysics*, **58**(12), 1748–1764.

Pratt, R.G., Shin, C. & Hicks, G.J., 1998. Gauss Newton and full Newton methods in frequency space seismic waveform inversion, *Geophys. J. Int.*, **133**, 341–362.

Pratt, R.G., Hou, F., Bauer, K. & Weber, M.H., 2004. Waveform tomography images of velocity and inelastic attenuation from the Malik 2002 Crosshole Seismic Surveys, *Tech. rep.*, Malik 2002 Gas Hydrate Production Research Programme.

Rothman, D.H., 1985. Nonlinear inversion, statistical mechanics, and residual statics estimation, *Geophysics*, **50**(12), 2784–2796.

Rothman, D.H., 1986. Automatic estimation of large residual statics corrections, *Geophysics*, **51**(2), 332–346.

Ryan, B., Wardle, R.J., Gower, C.F. & Nunn, G.A.G., 1995. Nickel-copper sulphide mineralization in Labrador: the Voisey's Bay discovery and its exploration implications, Report 95-1, Newfoundland Department of Natural Resources.

Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.*, **40**(3), doi:10.1029/2000RG000089.

Sen, M. & Stoffa, P.L., 1995. *Global Optimization Methods in Geophysical Inversion*, Elsevier.

Sen, M.K. & Stoffa, P.L., 1991. Nonlinear one-dimensional seismic waveform inversion using simulated annealing, *Geophysics*, **56**(10), 1624–1638.

Shin, C. & Cha, Y.J., 2009. Waveform inversion in the Laplace-Fourier domains, *Geophys. J. Int.*, **177**(3), 1067–1079.

Shin, C. & Min, D.-J., 2006. Waveform inversion using a logarithmic wavefield, *Geophysics*, **71**(3), R31–R42.

Sirgue, L. & Pratt, R.G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics*, **69**(1), 231–248.

Szu, H. & Hartley, R., 1987. Fast simulated annealing, *Phys. Lett. A*, **122**(3), doi:10.1016/0375-9601(87)90796-1.

Tarantola, A., 1984. Inversion of seismic reflection data in the acousitc approximation, *Geophysics*, **49**(8), 1259–1266.

Tarantola, A., 2005. *Inverse Problem Theory*, Society for Industrial and Applied Mathematics.

Thomsen, L., 1986. Weak elastic anisotropy, *Geophysics*, **51**, 1954–1966.

Tsvankin, I., 2012. *Seismic Signatures and Analysis of Reflection Data in Anisotropic Media*, 3rd edn, Society of Exploration Geophysicists.

Um, J. & Thurber, C., 1987. A fast algorithm for two-point seismic ray tracing, *Bull. seism. Soc. Am.*, **77**(3), 972–986.

Varela, C.L., Stoffa, P.L. & Sen, M.K., 1998. Background velocity estimation using non-linear optimization for reflection tomography and migration misfit, *Geophys. Prospect*, **46**(1), 51–78.

Warner, M., Ratcliffe, A., Nangoo, T. & Morgan, J., 2013. Anisotropic 3D full-waveform inversion, *Geophysics*, **78**(2), R59–R80.

Williamson, P.R., 1991. A guide to the limits of resolution imposed by scattering in ray tomography, *Geophysics*, **56**(2), 202–207.

Williamson, P.R. & Worthington, M.H., 1993. Resolution limits in ray tomography due to wave behavior: numerical experiments, *Geophysics*, **58**(3), 727–735.

Wirgin, A., 2004. The inverse crime, arXiv:math-ph/0401050, 1–10.

Woodward, M.J., 1992. Wave-equation tomography, *Geophysics*, **57**(1), 15–26.

Wu, R.S. & Toksoz, M.N., 1987. Diffraction tomography and multisource holography applied to seismic imaging, *Geophysics*, **52**, 11–25.

Yan, J. & Sava, P., 2011. Elastic wave-mode separation for TTI media, *Geophysics*, **76**(4), 65–78.

Zelt, C.A. & Barton, P.J., 1998. Three-dimensional seismic refraction tomography: a comparison of two methods applied to data from the Faeroe Basin, *J. geophys. Res.*, **103**(B4), 7187–7210.

Zelt, C.A., Sain, K., Naumenko, J.V. & Sawyer, D.S., 2003. Assessment of crustal velocity models using seismic refraction and reflection tomography, *Geophys. J. Int.*, **153**(3), 609–626.