

Joint transmission and reflection traveltime tomography using the fast sweeping method and the adjoint-state technique

Jun-Wei Huang* and Gilles Bellefleur

Geological Survey of Canada, 615 Booth Street, Ottawa ON, K1A 0E9 Canada. E-mail: junwei.huang@vibrometric.com

Accepted 2011 October 17. Received 2011 October 17; in original form 2011 April 29

SUMMARY

We present a joint transmission and reflection traveltime tomography algorithm based on the Fast Sweeping Method and the adjoint-state technique. In contrast to classical ray based tomography, this algorithm utilizes a grid-based Eikonal equation solver to circumvent the non-linearity of conventional ray shooting and bending approaches in complex media. The adjoint-state technique is used to obtain the gradient of the objective function without the explicit estimation of the Fréchet derivative matrix, which is usually computationally prohibitive for large-scale problems. When combined with Huygens' Principle, the tomographic inversion can simultaneously use direct and reflected arrivals to optimize a final velocity model, further mitigate the ambiguity of the inverse problem and reveal deeper structures not visible to transmission tomography alone. In this paper, we describe the theoretical basis of our algorithm, evaluate its performance on synthetic models, and then apply it to a 20 km long 2-D seismic survey acquired in the Mackenzie Delta, Northwest Territories of Canada. The subsurface at that location is characterized by a thick permafrost (600 m) comprising high- and low-velocity areas associated with thermokarst lakes. Our results show the potential of the joint tomography in characterizing multi-scale heterogeneous velocity structures within the permafrost.

Key words: Numerical solutions; Tomography; Gas and hydrate systems; Arctic region.

1 INTRODUCTION

Seismic wave traveltimes have been used to image the Earth's interior at a variety of scales ranging from global and continental scales in earthquake seismology (e.g. Aki & Lee 1976; Sengupta & Toksöz 1976; Dziewonski *et al.* 1977) to local scales in seismic exploration (e.g. Stewart 1991). In general, deterministic traveltime tomography consists of two essential steps performed iteratively: (1) calculating traveltimes in a starting model, and (2) updating the model to reduce the discrepancy between the calculated and observed traveltimes. In the first step, ray-tracing is often used to find the trajectory of seismic waves through a medium with varying propagation velocity (Virieux & Farra 1991; Červený 2001). The calculated traveltimes of the seismic body waves are then combined with the traveltimes observed with seismic sensor to reconstruct the seismic propagation velocity (Tarantola 2005).

Conventional ray tracing can be problematic in a complex medium due to shadow zones and multipathing problems and is computationally expensive for large-scale seismic survey with many source–receiver pairs (Vidale 1988; Rawlinson *et al.* 2008). An alternative to ray tracing is to numerically solve the Eikonal equation for the traveltime field discretized on regular grids with the Fi-

nite Difference method (Vidale 1988) or on irregular grids with the Finite Element method (Cheng & Shu 2007). Such grid-based methods possess numerous advantages over conventional ray tracing (see Rawlinson *et al.* 2008). For example, grid-based Eikonal equation solver can avoid shadow zone problems in conventional ray tracing, since a traveltime is assigned to each grid point. In addition, grid-based methods are formulated as a forward problem which is highly stable for first arrival traveltime calculations.

One of the first grid-based solvers of Eikonal equation was proposed by Vidale (1988) and later improved by many authors including Podvin & Lecomte (1991), Hole & Zelt (1995), Sethian (1996), Kim & Cook (1999), Qian & Symes (2002), Zhao (2004) and Taillandier *et al.* (2009). Interested readers are referred to Rawlinson *et al.* (2008) for a comprehensive review. Currently, there are two competing algorithms used to solve the Eikonal equation: Fast Marching Method (FMM) and Fast Sweeping Method (FSM). FMM advances wave fronts in a monotonic manner, but must be conducted in a sequential order to ensure causality and cannot be easily modified to achieve high-order accuracy (Kim & Cook 1999). FSM advances wave fronts with iterative sweeping and updates traveltime at each grid point monotonically to ensure the causality (Tsai *et al.* 2003; Zhao 2004). The implementation of FSM is thus significantly simplified and its computational efficiency is superior to FMM even when velocity changes rapidly (Bak *et al.* 2010). In addition, FSM can be easily extended to high-order accuracy (Zhang *et al.* 2006; Xiong *et al.* 2010), achieve high computational performance on

*Now at: Vibrometric Canada Limited, 425 Adelaide St West, Suite 702, Toronto, ON, M5V 3C1, Canada.

shared or distributed memory systems (Zhao 2007) and can be implemented on unstructured elements as well (Qian *et al.* 2007). In this study, we employ FSM as the forward modelling algorithm of the traveltimes tomography.

One limitation common to grid-based methods is the complicated calculation of secondary arrivals such as reflection traveltimes. Podvin & Lecomte (1991), Matsuoka & Ezaka (1992) and Riahi & Juhlin (1994) computed reflection times by tracking first arrivals to the interfaces from both source and receiver and applied Fermat's principle of stationary time to locate reflection points. This scheme requires the discretization of the interfaces at depth. The computation load also increases significantly with increasing pairs of source and receiver and number of reflecting interfaces. Another option is to utilize local rectangular and triangular meshes to model interfaces with irregular topography. The incident traveltimes field from the source is calculated directly on nodes at the interface and the reflected traveltimes field can be obtained by reinitializing the traveltimes calculation from the interface node with minimum traveltimes (Li & Ulrych 1993) or from the entire interface nodes (Rawlinson & Sambridge 2004). In 3-D situations, this local regridding technique requires irregular tetrahedral meshes and the implementation is not easy (e.g. de Kool *et al.* 2006). In our study, we adopted the method in Hole & Zelt (1995) which consist in applying Snell's law in a local plane wave approximation in the neighbourhood of an interface within the modelling region. The direct traveltimes at all nodes immediately above the interface are replaced by the reflected traveltimes. According to Huygens' Principle, the reflected traveltimes field can be obtained by using the reflected traveltimes as secondary sources. For multiple reflections and multiple interfaces, a multi-stage FSM scheme similar to the multi-stage FMM approach used in Rawlinson & Sambridge (2004) is applied.

Classical tomography formulates the inverse problem as a sequence of linear problems, in which established algorithms such as LSQR (Paige & Saunders 1982) can be applied to efficiently solve the linear matrix system. In our study, we maintain the non-linear relationship between the traveltimes and the velocity and formulate the tomographic inversion as a non-linear optimization problem to find an optimum model that minimizes the discrepancies between the measured and the calculated traveltimes. The key step of the deterministic minimization algorithm is to calculate the derivative of the misfit with respect to the model parameters (Fréchet derivatives). The adjoint-state (or simply adjoint) technique allows us to obtain the gradient with cost equivalent to the forward modelling solution, regardless of the size of the model. The mathematical details on the adjoint method can be found in Fichtner *et al.* (2006) and a comprehensive review of geophysical applications is given by Plessix (2006).

With computational efficiency from FSM and the adjoint method, we are able to perform tomography using both direct and reflected traveltimes. The joint tomography is able to recover deep structure undetectable to first arrival tomography and to reduce the ambiguity in the inversion, particularly in regions where the wavefield of direct and reflected waves overlaps. In such case, the misfit function is defined as a weighted summation of the traveltimes discrepancies for both direct and reflected arrivals, and the final velocity update corresponds to the weighted summation of the velocity perturbations.

This paper is composed as follows. In Section 2, we present the essential ingredients of our joint traveltimes tomography algorithm, i.e. the FSM for transmission and reflection arrivals and the adjoint method. In Section 3, we demonstrate the performance of the transmission, reflection and joint traveltimes tomography on a syn-

thetic 2-D model and apply it to the 2-D seismic data acquired from Mackenzie delta, Northwest Territories of Canada. The results from synthetic data demonstrate its benefits and robustness in recovering both shallow and deep velocity structures. The joint tomography on the field data recovers both shallow and deep velocity in the presence of heterogeneous permafrost and complex thermokarst lakes. The image of small-scale low-velocity anomalies corresponding to geologically disturbed areas such as partially frozen sediments beneath lakes is also enhanced by the joint inversion of transmission and reflection traveltimes.

2 METHODOLOGY

The Eikonal equation belongs to the class of static Hamilton–Jacobi equations (e.g. Bento 1977). In seismology, it can be derived from the full elastic wave equation with the high-frequency approximation (Chapman 2004) as expressed in eq. (1):

$$|\nabla T(\mathbf{x})|^2 = \frac{1}{c^2(\mathbf{x})}, \quad \mathbf{x} \in \Omega \setminus \Gamma, \quad (1a)$$

with boundary condition

$$T(\mathbf{x}_s) = g(\mathbf{x}_s), \quad \mathbf{x}_s \in \Gamma, \quad (1b)$$

where $T(\mathbf{x})$ is the traveltimes at location \mathbf{x} , and $c(\mathbf{x})$ are the propagation velocity, respectively, $\Omega = R^n$ for an n -dimensional problem, and Γ is the boundary with known function of $g(\mathbf{x}_s)$. In the case of point source, the boundary values are defined at a point for first arrivals and defined along a reflecting interface for reflection arrivals.

There are generally two different methods used to solve a static Hamilton–Jacobi equation. One converts eq. (1) from a static to time-dependent state with the level set method (Osher 1993) or the paraxial formulation (Qian & Symes 2002). A variety of numerical methods can be applied to solve the time-dependent Hamilton–Jacobi equations (see Shu 2007 and references therein). The other method is to discretize directly eq. (1) as a stationary boundary value problem and tracks the wave front by following the causality of wave propagation. The FMM and FSM fall in this category. FMM was initially developed in the field of computational mathematics to solve interface evolution problems (Sethian 1996). Starting from the boundary, FMM stores all active nodes (the wave front) in a heap data structure named Narrow Band and uses an upwind finite difference method to calculate traveltimes at all the neighbouring nodes. To ensure the causality, only the grid point with minimum traveltimes is updated and the wave front is moved one step forward. The complexity of FMM is $O(N \log N)$ for an N grid model, where $O(\log N)$ is from the heap sorting algorithm. FSM (Tsai *et al.* 2003; Zhao 2004) uses Gauss–Seidel iterations and alternating sweeping orders such that each sweeping direction covers a class of wave propagation directions. During each sweep, the traveltimes can only monotonically increase to ensure the causality of wave propagation and the complexity drops to $O(N)$. In the next two sections, we provide a brief description on how FSM is used to solve both direct and reflected traveltimes from eq. (1). The gradient of misfit function of direct and reflected arrivals is then derived using the adjoint method.

2.1 Fast sweeping method for direct arrivals

We first apply the fast sweeping algorithm to solve the Eikonal equation with a point source in a 3-D medium, i.e.

$$|\nabla T(\mathbf{x})|^2 = \frac{1}{c^2(\mathbf{x})}, \quad \mathbf{x} \in R^3 \setminus \mathbf{x}_s, \quad (2)$$

$$T(\mathbf{x}_s) = 0$$

where \mathbf{x}_s is the location of the point source and $c(\mathbf{x})$ is the propagation velocity.

Using the first-order Godunov upwind finite difference scheme (Rouy & Tourin 1993), we discretize eq. (2) as follow:

$$\left[\frac{(T(\mathbf{x}) - T(\mathbf{x})^{x \min})^+}{\Delta x} \right]^2 + \left[\frac{(T(\mathbf{x}) - T(\mathbf{x})^{y \min})^+}{\Delta y} \right]^2 + \left[\frac{(T(\mathbf{x}) - T(\mathbf{x})^{z \min})^+}{\Delta z} \right]^2 = \frac{1}{c^2(\mathbf{x})}, \quad (3a)$$

where $T(\mathbf{x})$ is the discrete approximation at the grid $\mathbf{x} = (i, j, k)$, the indices of i, j, k are running from 0 to $N_p - 1$ and $T(\mathbf{x})^{p \min}$ is the minimum traveltime value of two adjacent neighbours of $T(\mathbf{x})$ along the axis of p (p is x, y or z). For example, $T(\mathbf{x})^{x \min} = \min(T(x + \Delta x, y, z), T(x - \Delta x, y, z))$. The $\Delta x, \Delta y$ and Δz are the grid intervals along x, y, z directions, respectively, and the total number of grid points of the model is $N_x \times N_y \times N_z$. The superscript '+' is an operator taking the positive value only and expressed as:

$$(A)^+ = \begin{cases} A, & A > 0 \\ 0, & A \leq 0 \end{cases}, \quad (3b)$$

where A is an arbitrary value evaluated by the operator. Note that one-sided finite difference must be used when \mathbf{x} is at the edge of the modelling region (e.g. $\mathbf{x} = (0, j, k)$ or $\mathbf{x} = (N_x - 1, j, k)$). In this case, the first term in eq. (3) becomes

$$\left[\frac{(T(0, j, k) - T(1, j, k))^+}{\Delta x} \right]^2 \text{ or } \left[\frac{(T(N_x - 1, j, k) - T(N_x - 2, j, k))^+}{\Delta x} \right]^2, \quad (3c)$$

respectively. Here, we start the array index from zero.

The fast sweeping algorithm (referred to as Algorithm 1) can thus be implemented as follows.

1. Initialize the point source condition $T(\mathbf{x}_s) = 0$ and assign very large positive values to the rest of the grid points which will be updated.
2. Update grid points with Gauss-Seidel iterations with eight alternating sweeping orders. At each node, the solution from eq.

(3) denoted as T^* is compared with the old value (T^{old}) and the traveltime at this node T^{new} is updated with the smallest value between the old and calculated traveltime value, i.e. $\min(T^{\text{old}}, T^*)$. It is important to mention that the sweeping order is not necessarily applied in sequential order and can be performed in parallel on multiple processors.

3. Test the convergence by checking the criterion $\|T^{n+1} - T^n\|_{L1} \leq \varepsilon$ point-wise, for a given convergence criterion $\varepsilon > 0$.

During the sweeping of the whole domain (step 2), a unique solution must be obtained from eq. (3) to update node $T(i, j, k)$. The following algorithm (steps 2.1 to 2.3) is used to systematically find that unique solution:

2.1 Assign the values of $T(\mathbf{x})^{x \min}, T(\mathbf{x})^{y \min}$ and $T(\mathbf{x})^{z \min}$ to u, v and w such that $w \leq v \leq u$ and let $T^* = w + 1/c(\mathbf{x})$. For clarity of annotation in steps 2.2 and 2.3, u, v and w are intermediate variables used to store the values of $T(\mathbf{x})^{x \min}, T(\mathbf{x})^{y \min}$ and $T(\mathbf{x})^{z \min}$.

2.2 If T^* is less than v , then T^* is the solution. Otherwise, let $T^* = (v + w + \sqrt{-v^2 - w^2 + 2vw + 2/c^2})/2$.

2.3 If T^* is less than u , then T^* is the solution. Otherwise, the solution is $T^* = (2(u + v + w) + \sqrt{4(u + v + w)^2 - 12(u^2 + v^2 + w^2 - 1/c^2)})/6$.

Fig. 1 shows an example of seismic wave propagation in a 3-D layered media model comprising 10^6 grid points obtained with the fast sweeping algorithm. The iso-surface of different arrival times is rendered using different colours (Fig. 1a) and the first arrivals are obtained at each grid point of the volume (Fig. 1b). In this example, the source is located at the (5, 5, 10) grid point and a total of four iterations were completed on a single CPU in 7 s.

2.2 Fast sweeping method for reflection arrivals

To calculate the reflection times, we adopted the method of Hole & Zelt (1995). First, the FSM is applied to track the wave front from a point source down to the reflecting interface using eq. (2). Secondly, Snell's law is applied to update reflected traveltime at grid points immediately above the interface. Generally, there is no intrinsic limitation to low dip interface. As stated in Hole & Zelt (1995), however, the local incidence plane wave assumption can be challenged by highly irregular interfaces. Finally, the Eikonal equation with a different boundary condition is solved for the reflected

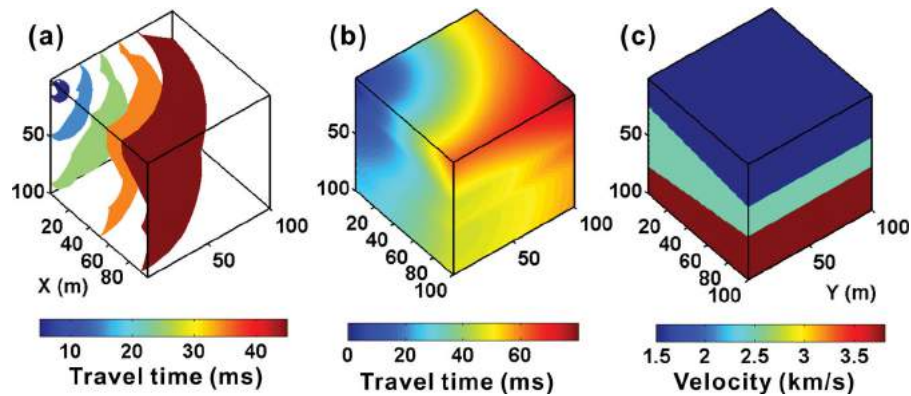


Figure 1. An example of seismic wave propagation in a 3-D model with 10^6 grid points. The traveltime is in millisecond and the distance is in metre. The traveltime field is calculated in 7 s. (a) The iso-surface of different arrival times is rendered using different colours. (b) A slice view of the traveltimes showing that each grid point of the volume has an assigned traveltime, and (c) a slice view of the 3-D velocity model used to calculate traveltimes in (a) and (b).

traveltime field:

$$|\nabla T_u(\mathbf{x})|^2 = \frac{1}{c^2(\mathbf{x})}, \quad \mathbf{x} \in \Omega' \setminus \Gamma \quad (4)$$

$$T_u(\mathbf{x} = \Gamma) = T_d(\Gamma)$$

where Ω' denotes the modelling region containing the source and bounded by the reflecting interface Γ , $T_d(\Gamma)$ is the traveltime of the down-going wave at the interface Γ , and $T_u(\mathbf{x})$ is the up-going reflection traveltimes in region Ω' . In a general situation, the calculation of T_u and T_d can be repeated an arbitrary number of times to retrieve traveltimes of multiples, similar to the ‘multi-stage’ approach described in Rawlinson & Sambridge (2004). This procedure is applied to many interfaces as demonstrated in Fig. 2, where two irregular reflectors and a variable surface topography are present. For display purpose, the wave front in the air is not plotted.

Appendix A, we demonstrate the accuracy of this ‘multi-stage’ FSM by comparing the calculated traveltimes with the analytical solution.

2.3 Adjoint tomography

The adjoint method was initially developed in the field of control theory (Lions 1971), first introduced into seismology by Bamberger *et al.* (1982), and later popularized by Tarantola (1984, 1988), Mora (1988), Pratt (1999) and Tromp *et al.* (2005). The adjoint method formulates the inverse problem as the minimization of the misfit function under the constraint of the Partial Differential Equation (PDE). To avoid the computation-intensive task of the Fréchet derivative matrix estimation, a new PDE termed adjoint PDE must be solved to obtain the gradient of the misfit function, which turns

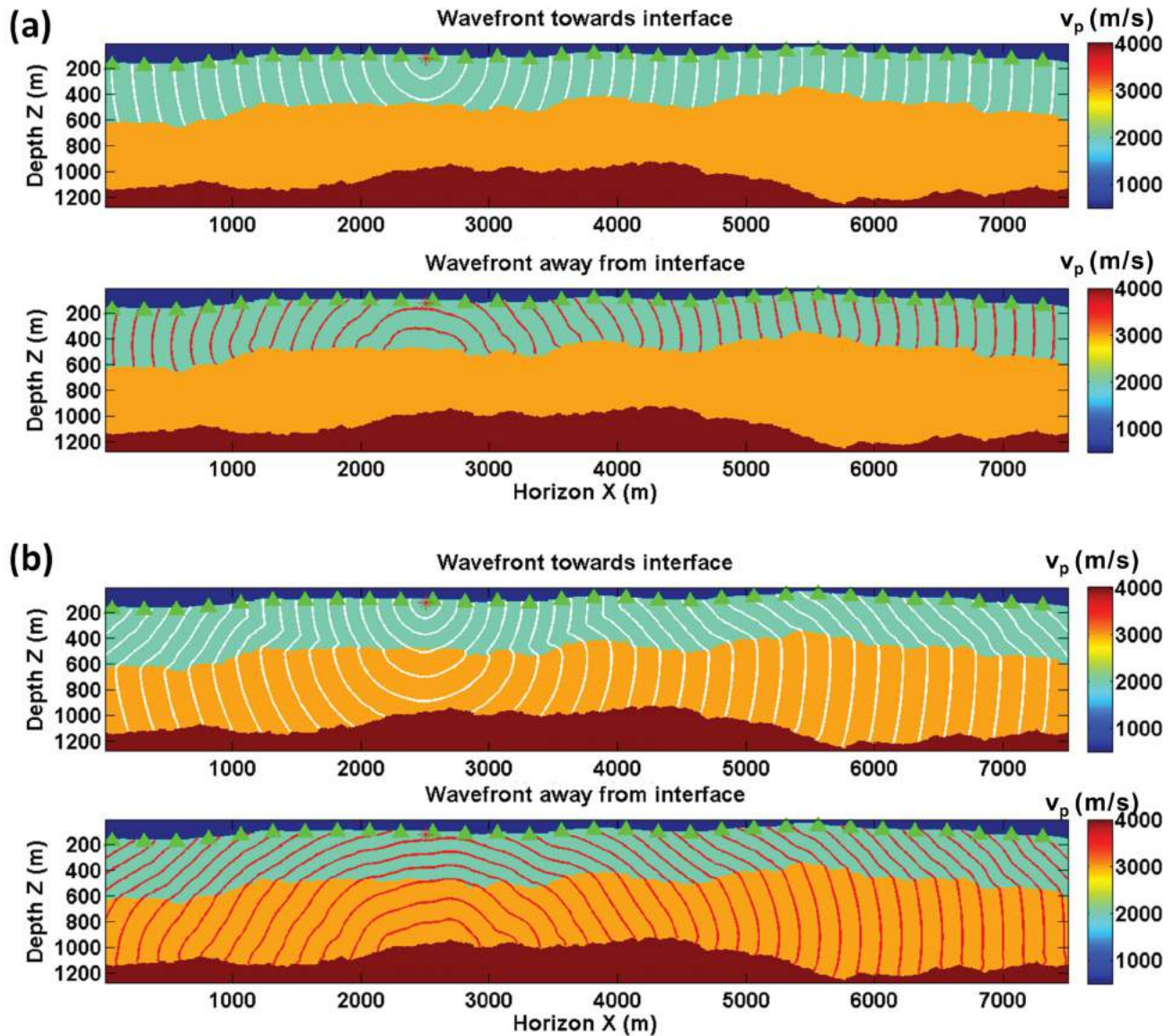


Figure 2. The ‘multi-stage’ procedure for tracking transmission and reflection wave front for all interfaces. The red star and the green triangles denote the shot and receiver locations, respectively. Velocity is increasing with depth but each layer has constant velocity. The propagating wave front is displayed as contours initiated from the shot for transmission and from the interface for reflection. Figures (a) and (b) show the wave front tracking for the shallow and deeper interfaces.

the inverse modelling problem into another forward modelling problem. In this section, we derive the adjoint PDE of the Eikonal equation and show that the model update can be efficiently calculated using the FSM method.

We first define the misfit function (energy) as the L_2 norm of the arrival time differences:

$$E(c) = \frac{1}{2} \int_{\partial\Omega} dr |T(c, \mathbf{x}) - T^*(\mathbf{x})|^2, \quad (5)$$

where c is the propagation velocity, a function of space, T is the calculated arrivals from the source, and T^* is the observed arrival time at the surface $\partial\Omega$, where the numerical sensors are deployed. In practice, the integral in eq. (5) should be regarded as a summation, since $T^*(\mathbf{x})$ is only defined at the position where sensors are deployed.

Secondly, we define the perturbation of the velocity as ∂c which causes a corresponding change in arrival time ∂T . Therefore the change of the energy can be expressed as

$$\partial E = \int_{\partial\Omega} dr (T - T^*) \partial T. \quad (6)$$

The perturbation in velocity is related to the perturbation of arrival time obtained by differentiating eq. (2):

$$T_x \partial T_x + T_y \partial T_y + T_z \partial T_z = -\frac{1}{c^3} \partial c, \quad (7)$$

where T_p denotes the partial derivative of T in the direction of p (p is x , y or z).

By multiplying eq. (7) by a Lagrange multiplier λ , integrating it over the domain Ω , and adding it to (6), we obtain a new expression for the change of energy:

$$\begin{aligned} \partial E &= \int_{\partial\Omega} d\mathbf{x} (T - T^*) \partial T \\ &+ \int_{\Omega} d\mathbf{x} \cdot \lambda (T_x \partial T_x + T_y \partial T_y + T_z \partial T_z + \partial c / c^3). \end{aligned} \quad (8a)$$

For reasons that will be clear in the following equations, multiplying eq. (7) by an arbitrary function λ eliminates ∂T , i.e. the variation of travel time with respect to the perturbation of velocity, whose evaluation is usually computationally prohibitive for large-scale problems.

The integration by parts of the second term on the right-hand side of (8a) yields:

$$\begin{aligned} \partial E &= \int_{\partial\Omega} d\mathbf{x} (T - T^*) \partial T + \int_{\partial\Omega} d\mathbf{x} (\mathbf{n} \cdot \nabla T) \lambda \partial T \\ &+ \int_{\Omega} d\mathbf{x} \nabla \cdot (-\lambda \nabla T) \partial T + \int_{\Omega} d\mathbf{x} \cdot \lambda \partial c / c^3. \end{aligned} \quad (8b)$$

By choosing a multiplier λ that satisfied the following PDE

$$\nabla \cdot (-\lambda \nabla T) = 0, \quad (\mathbf{x} \in \Omega), \quad (9a)$$

$$(\mathbf{n} \cdot \nabla T) \lambda = T^* - T, \quad (\mathbf{x} \in \partial\Omega), \quad (9b)$$

where \mathbf{n} is the normal vector perpendicular to the surface of sensors, the explicit knowledge of ∂T is not needed and the gradient of the objective function can be simply calculated by

$$\partial E = \int_{\Omega} d\mathbf{x} \cdot \lambda \partial c / c^3. \quad (10)$$

The multiplier λ and eq. (9) are also called the adjoint solution and the adjoint equation, respectively. To ensure that update of velocity reduces the misfit, we chose the velocity perturbation defined as

$\partial c = -\lambda / c^3$. We hereafter refer to ∂c as the velocity perturbation, a vector pointing towards the descent direction of the misfit function. In the gradient-based optimization framework, the n th iteration is usually performed as

$$c_{n+1} = c_n + \alpha_n \partial m, \quad (11)$$

where α_n is the step length, and ∂m is the model perturbation. In the case of the steepest descent method, the perturbation ∂m simply equals to ∂c .

We adapted the algorithm from Zhao (2006) (referred to as Algorithm 2) to solve the adjoint equation (eq. 9) with the fast sweeping algorithm:

1. Initialize λ with the boundary values using eq. (9b) at the boundary and assign zeros to the remaining grid points.
2. Sweep the whole domain alternatively as in step 2 of Algorithm 1 and update the values of λ at the non-boundary locations using eq. (9a).
3. Test the convergence by checking the criterion $\|\lambda^{n+1} - \lambda^n\|_{L1} \leq \varepsilon$ point-wise, for a given convergence criterion $\varepsilon > 0$.

The difference between Algorithm 1 and Algorithm 2 is that the values of λ are not required to change monotonically as traveltine in step 2 of Algorithm 1, since no causality is required for the adjoint equation.

For typical seismic survey with multiple sources (e.g. N_S) located at \mathbf{x}_s^i ($i = 1, 2, \dots, N_S$), the misfit function becomes

$$E(c) = \frac{1}{2} \sum_{i=1}^{N_S} \int_{\partial\Omega} d\mathbf{x} |T^i(c, \mathbf{x}) - T^{i*}(\mathbf{x})|^2, \quad (12a)$$

where sensors are distributed along the surface $\partial\Omega$. Following the same procedure, each shot provides a velocity perturbation ∂c^i and the final velocity perturbation is simply the summation of all individual perturbations:

$$\partial c = \sum_{i=1}^{N_S} \partial c^i. \quad (12b)$$

In the case of reflection tomography, the location of reflective interfaces is given *a priori* in model space coordinates, and eq. (12a) becomes

$$E_R(c) = \frac{1}{2} \sum_{r=1}^{N_R} \sum_{i=1}^{N_S} \int_{\partial\Omega} d\mathbf{x} |T_r^i(c, \mathbf{x}) - T_r^{i*}(\mathbf{x})|^2, \quad (13a)$$

with the total velocity perturbation

$$\partial c = \sum_{r=1}^{N_R} \sum_{i=1}^{N_S} \partial c_r^i, \quad (13b)$$

where N_R is the total number of reflecting interfaces and T_r^{i*} is the observed reflection time from the i th shot reflected by the r th interface.

For joint tomography, the misfit function can be defined as the weighted summation of (12a) and (13a):

$$\begin{aligned} E_{TR}(c) &= \frac{w_T}{2} \sum_{i=1}^{N_S} \int_{\partial\Omega} d\mathbf{x} |T^i(c, \mathbf{x}) - T^{i*}(\mathbf{x})|^2 \\ &+ \frac{w_R}{2} \sum_{r=1}^{N_R} \sum_{i=1}^{N_S} \int_{\partial\Omega} d\mathbf{x} |T_r^i(c, \mathbf{x}) - T_r^{i*}(\mathbf{x})|^2, \end{aligned} \quad (14a)$$

where the summation of weights equals one, i.e. $w_T + w_R = 1$ with $0 < (w_T \text{ and } w_R) < 1$, and the velocity perturbation becomes

$$\partial c = w_T \sum_{i=1}^{N_S} \partial c^i + w_R \sum_{r=1}^{N_R} \sum_{i=1}^{N_S} \partial c_r^i. \quad (14b)$$

Spatially varying weighting factors can also be employed to account for space-dependent reliability of direct and reflection arrivals. Then eq. (14b) becomes

$$\partial c = w_T(\mathbf{x}) \sum_{i=1}^{N_S} \partial c^i(\mathbf{x}) + w_R(\mathbf{x}) \sum_{r=1}^{N_R} \sum_{i=1}^{N_S} \partial c_r^i(\mathbf{x}), \quad (14c)$$

where the summation of weights still equals one, i.e. $w_T(\mathbf{x}) + w_R(\mathbf{x}) = 1$ with $0 < (w_T(\mathbf{x}) \text{ and } w_R(\mathbf{x})) < 1$, for all $\mathbf{x} \in R^3$. Empirically, direct arrival traveltimes are more reliable to define velocities of the shallow subsurface whereas reflected arrivals are more appropriate for the deeper structures. In the following synthetic example, a constant weighting function is used, whereas in the field data section, we use a weighting function $w_T(\mathbf{x})$ that gradually decreases from 1 at the surface to 0 at the reflective interface.

Due to the sparse sampling in field survey and spikes in boundary values caused by an irregular topography (see eq. 9b), the velocity perturbation (∂c) is usually not smooth and may cause unstable inversion results or render models trapped in local minima. To eliminate spikes and avoid local minima, we regularize the model perturbation by a low pass filter expressed in the wavenumber domain as

$$f(k_x, k_y, k_z) = \frac{1}{(v_x k_x)^2 + (v_y k_y)^2 + (v_z k_z)^2}, \quad (15)$$

where v_x, v_y, v_z are filter parameters controlling the smoothness (or roughness) of the model in three dimensions. When $v_x = v_y = v_z = v$, the filter in eq. (15) is identical to the Laplacian operator in Leung & Qian (2006), which guarantees the well-posedness of the inverse problem.

As an example, we show some regularized velocity perturbations from the transmission, reflection and joint tomography in Fig. 3. The velocity perturbations are calculated during the first iteration with an initial velocity model defined with an increasing vertical velocity gradient. The velocity perturbations suggest some velocity modifications consistent with the known anomaly located in the middle of the model (Fig. 3a). Figs 3(b) and (c) illustrate the lateral and vertical smearing effect due to the raypaths of direct and reflected arrivals. The joint tomography increases the resolution in both directions by combining the two velocity perturbations (see Fig. 3d).

3 JOINT TOMOGRAPHY EXAMPLES

3.1 A simple synthetic case

A 2-D numerical seismic survey is simulated over a 7500 m long by 1280 m deep model comprising two layers (Fig. 4). A total of 56 sources and 600 receivers are deployed along the topographic surface with source and receiver intervals of 125 and 12.5 m, respectively. The velocity in the top layer increases with depth and becomes homogeneous in the bottom layer. As seen in Fig. 4(a), the background velocity in each layer is perturbed by a velocity anomaly (Fig. 4c). For the tomography, we choose the background velocity without any anomalies (Fig. 4b) as the starting model and used direct arrivals for transmission tomography and the primary reflection from the reflector located at the base of the model for

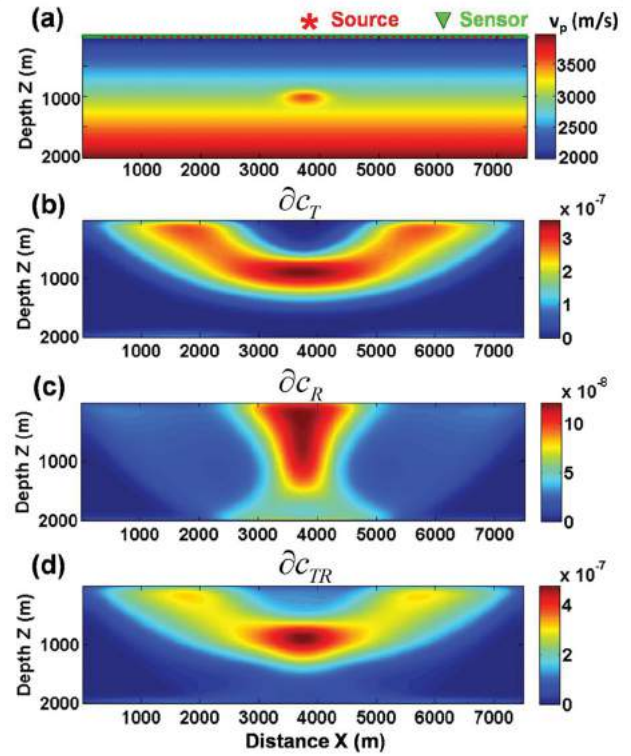


Figure 3. The velocity perturbations of transmitted (b) and reflected waves (c) from the first iteration using an initial velocity model defined with an increasing vertical gradient. This starting model is similar to the model shown in (a) but without the anomaly. The source and receivers are deployed evenly along the flat surface. Both velocity perturbations detect the central anomaly but with different spatial resolutions. The joint tomography in (c) reduces the spatial smearing effect and thus increases the resolution by combining the information of both direct and reflected arrivals. Additional iterations are required to converge to final solutions.

reflection tomography. We defined that the tomographic model has converged when the total energy (eq. 12a, 13a, or 14a) decreased by less than 0.1 per cent. A constant weighting function ($w_T = w_R = 0.5$) was adequate to recover the anomalies. As shown in Fig. 5, the transmission tomographic model recovers the shallow high-velocity anomaly but is unable to sample the low-velocity anomaly in the second layer due to the lack of turning rays in this area. The reflection tomography, on the other hand, samples both anomalies, but slightly stretches the deeper anomaly along the vertical axis. This is due to the limited vertical resolution of reflection tomography as reflected waves mainly travel vertically. The joint tomography combines the merits of both and provides the best image of the subsurface anomalies. The rms traveltme residual:

$$\Delta \bar{T} = \sqrt{\frac{2E}{N_T}}, \quad (16)$$

where E is the misfit function in eq. (12a, 13a, 14a) and N_T is the total number of arrivals, is plotted as a function of iteration number in Fig. 6. After convergence, the joint tomographic model reduces the residual traveltme to 0.13 ms. In practice, a higher plateau is expected due to the presence of noise which cannot be further reduced by iterative velocity perturbations. The robustness of our tomography algorithm in the presence of random noise is briefly demonstrated with checkerboard tests in Appendix B.

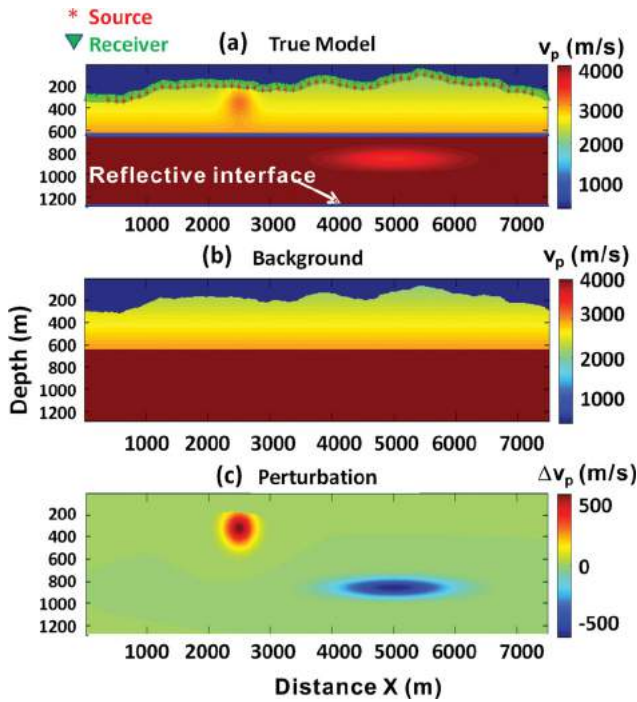


Figure 4. (a) The layered synthetic model consisting of a vertically heterogeneous layer located on top of a homogeneous layer. Each layer comprises a localized velocity anomaly. The source and receiver locations, denoted by star and triangles respectively, are deployed along the irregular topography. Two interfaces inside the model are marked by the blue lines, one separating the two layers and the other at the lower boundary of the model (at a constant depth of 1200 m). Note that in this example, the reflection tomography only used the reflected arrivals from the lower boundary. (b) The background velocity is used as the initial model. (c) Two elliptical shaped anomalies to be recovered by traveltimes tomography.

3.2 Field data from permafrost region

In the winter of 2001, four 2-D seismic profiles were acquired by oil and gas industry on Richards Island in the Mackenzie Delta, where complex thermokarst lakes and heterogeneous permafrost exist in the near surface (Bellefleur *et al.* 2008). In the central Mackenzie Delta, permafrost may be less than 80 m thick whereas in northern Richards Island it may be more than 700 m thick (Judge *et al.* 1987). In particular lakes which cover between 20 and 50 per cent of the landscape of the Mackenzie Delta area have played an important role in conditioning ground temperatures. Positive mean annual temperatures beneath deep lakes, if imposed for a significant amount of time, can thaw the permafrost and affect the proportion of water versus ice within the sediment matrix. Such variations modify the physical properties of sediments, including the velocity of *P* waves which can vary from 1.9 km s^{-1} in unfrozen sediments up to 4 km s^{-1} in fully frozen sediments (Dallimore *et al.* 1999). We applied the joint inversion tomography to one of the 2-D profile (line 120) to provide a better understanding of velocity distribution in the permafrost. This seismic profile was designed to avoid major lakes in the area but still intersects some water bodies (marked by arrows in Fig. 7a). The permafrost near line 120 is approximately 600 m thick and confirmed by well-logging data acquired in oil and gas industry wells. One of these wells is the Mallik 2L-38 gas hydrate research well (Dallimore *et al.* 1999) located at about 4 km from line 120 (Fig. 7a). Borehole logging in Mallik 2L-38 defines

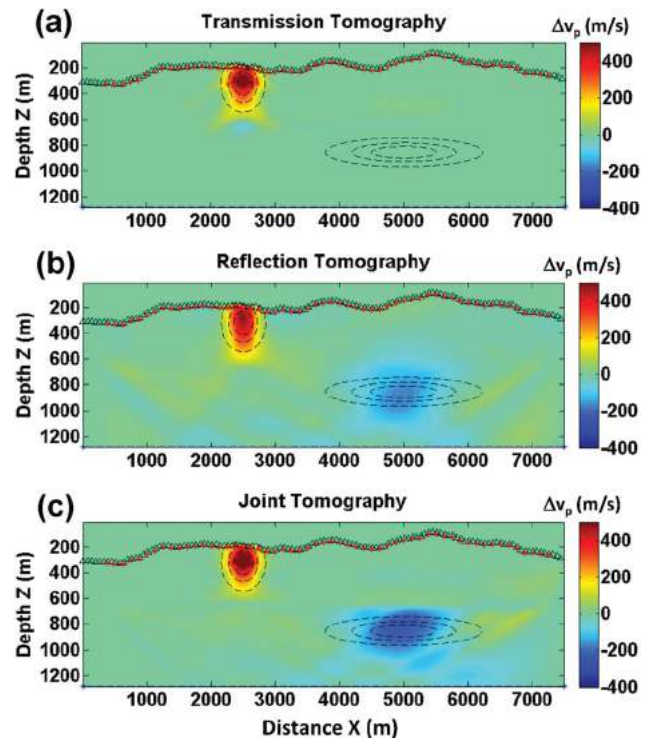


Figure 5. The tomographic models from transmission (a), reflection (b) and joint transmission-reflection (c). The sources and receivers show the acquisition geometry and the dashed contours mark the location and shape of the anomalies as defined in Fig. 4(c). (a) Transmission tomography only recovers the shallow anomaly, due to the lack of turning rays from the second layer. (b) Reflection tomography recovers both anomalies but smears the anomalies along the vertical direction, particularly for the deep anomaly (due to vertical raypath). (c) Joint tomography recovers both anomalies with relatively higher resolution. The vertically elongated anomaly recovered with reflection-only tomography is more focused with the transmission tomography.

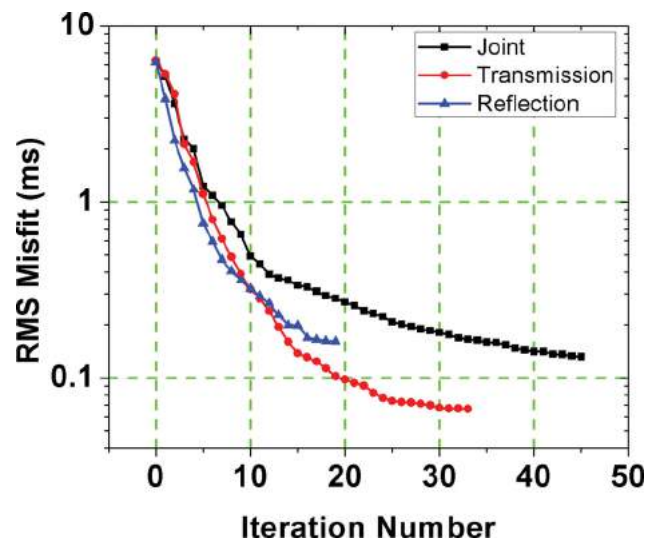


Figure 6. The rms traveltimes residual for transmission, reflection and joint tomography decreases at each iteration until the convergence criterion is satisfied, i.e. the change of the misfit is less than 0.1 per cent in this case. The vertical axis is logarithmic.

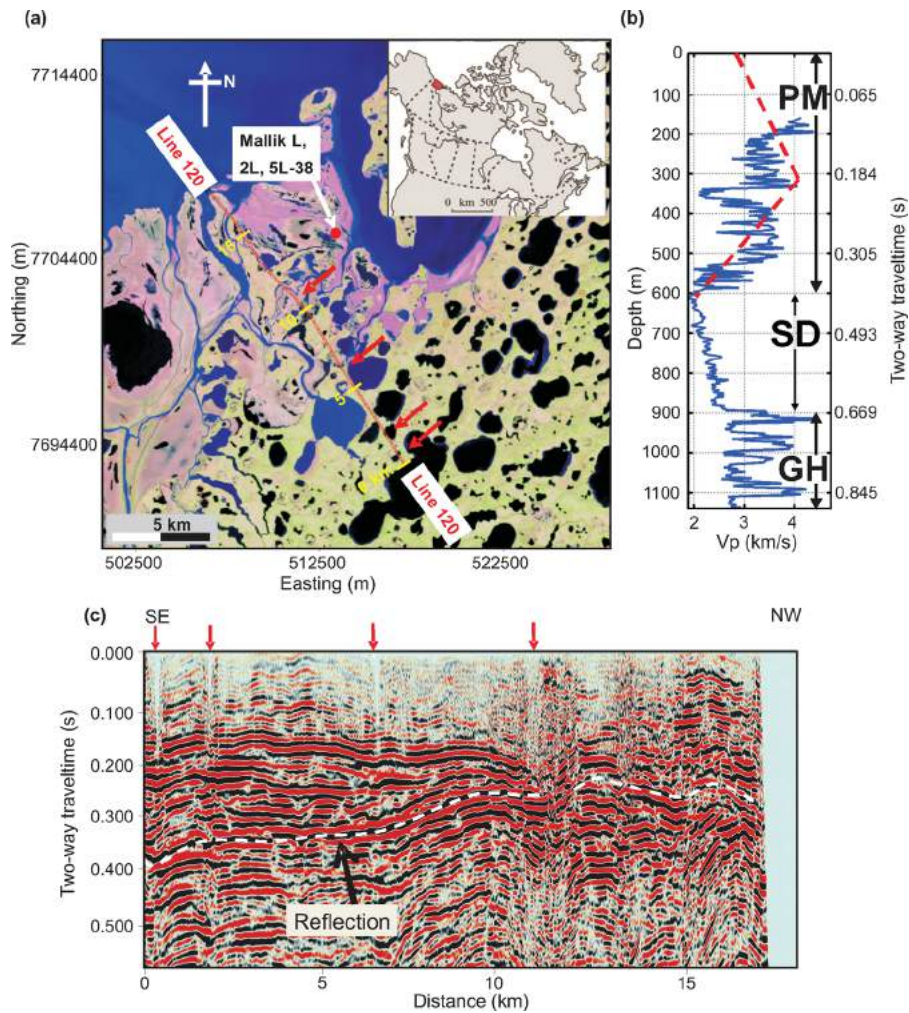


Figure 7. (a) The 2-D seismic survey and well locations on Richards Island, Mackenzie Delta, Northwest Territories of Canada. Line 120 was acquired parallel to the axis of a regional anticline structure and thus is characterized by relatively smooth and shallowly dipping reflection events. Red arrows show water bodies intersected by the profile. (b) P -wave sonic logs measured at Mallik 2L-38 allow the identification of three zones (PM, SD and GH) representing permafrost, fluid saturated sediment and gas hydrate bearing sediment, respectively. We observe the significant velocity variations in the permafrost. Velocities are generally increasing to a maximum value (at approximately 300 m depth) and gradually decrease to values of water-saturated sediments at the base of permafrost. (c) The pre-stack time migrated seismic image of Line 120. The dashed line marks the reflection used for the joint inversion. The water bodies intersected by the seismic line smear and distort reflections (arrows).

three zones: permafrost, sediment saturated by fluid and sediment saturated by gas hydrate (Fig. 7b). Velocity in permafrost usually increases to a maximum value before gradually decreasing to velocities of unfrozen sediments. The turning point of velocity shall be referred to as the peak velocity zone. The effect of lakes (marked in Fig. 7a) on the seismic section is observable in Fig. 7(c), and corresponds to areas where near surface reflection events lose lateral continuity (marked by arrows). Line 120 is parallel to the axis of an anticline structure and possesses relatively continuous and shallowly dipping reflections, especially on the South-East side of the profile (Fig. 7c).

Line 120 comprises 386 shot gathers acquired with 30 m shot and receiver intervals. The vibroseis source generated frequencies ranging from 6 to 96 Hz. Over its total length (19 km), line 120 has one very minimal deviation located approximately in the middle of the profile. A total of 77 220 first arrivals with offset up to 4 km were automatically picked, inspected and manually edited when appropriate. Since traveltimes were restricted to a maximum offset

of 4 km, we assumed that most of the traveltimes picks followed 2-D geometry. A clear reflection was identified from a prestack-time migrated section (dashed line in Fig. 7c) and used in the model after time-to-depth conversion using interval velocities obtained from pre-stack time migration velocity analysis. This reflection was also used to define pilot traces to guide the picking of reflection traveltimes on unprocessed common midpoint (CMP) gathers. Reflection arrivals at large offset are weak and usually contaminated by noise and thus were not included in the inversion process. A total of 14 862 picks with offset up to 0.5 km were used as input data for reflection traveltimes. The limited offsets and much smaller number of reflected arrivals result in sparser wavefield coverage in the lower part the model. In regions with limited transmitted waves, reflection arrivals tend to smear the velocity model vertically. To diminish the effect of smearing and balance the contribution between the first and the reflected arrivals, we employed a space varying weighting function, $w_T(x)$, that equals to 1 at the surface and gradually decreases to 0 at the reflective interface.

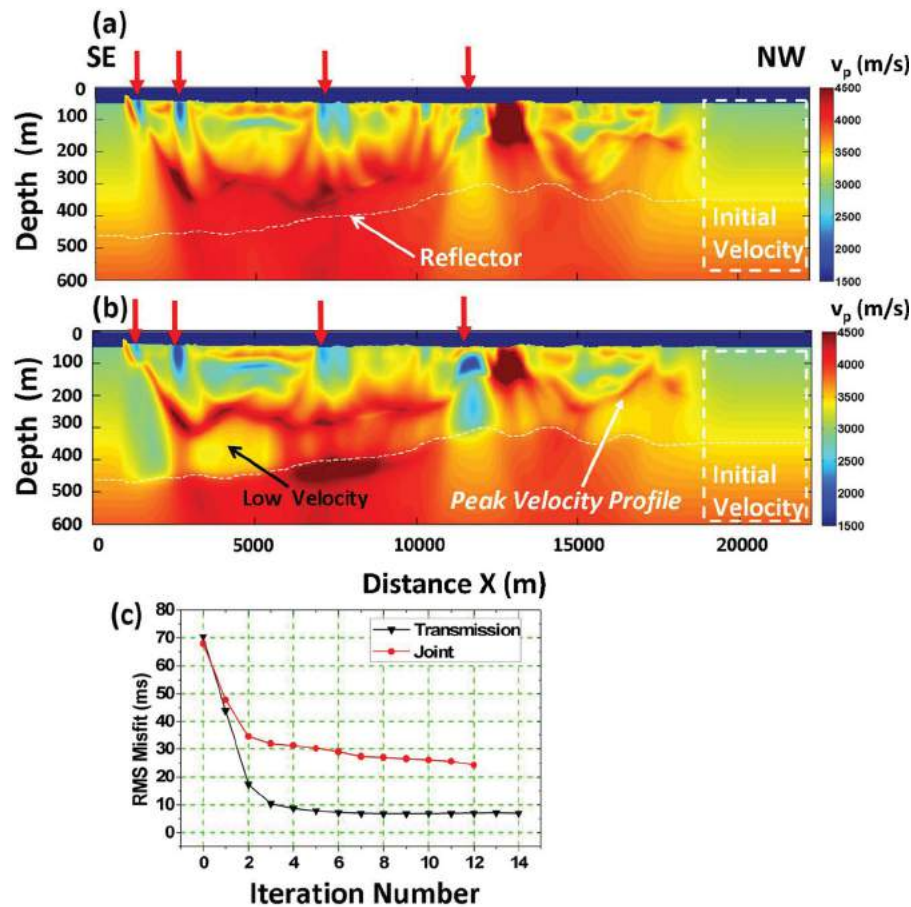


Figure 8. The final velocity image from (a) first-arrival turning-ray tomography and (b) joint tomography using first and reflected arrivals from Line 120. The starting model consisted of monotonically increasing velocity with depth and is still observed on the NW side of the profile. The low-velocity water bodies (arrows) and peak velocity profiles absent in the starting model are well recovered in both tomographic models. The tomographic model from field data shows complicated velocity distribution characterized by various scales and corresponding to water bodies observed on the satellite image shown in Fig. 7(a). (c) The rms misfit for the joint and transmission tomography as a function of iteration number is reduced to 21 and 7 ms, respectively.

The results from the transmission and joint tomography are shown in Fig. 8. In general, the upper part of the joint tomographic model is mostly controlled by transmission arrivals whereas deeper part is updated by reflected arrivals. In particular, the peak velocity zone is well defined indicating that long wavelength variations within permafrost are well resolved. Some low-velocity anomalies are also well-defined between the depth of 0 and 200 m, such as the low-velocity areas beneath water bodies (marked by arrows in Figs 8a and b). The lowest velocity anomaly recovered for this profile at 10.5 km corresponds to a highly disturbed area on the 2-D seismic profile and confirms imaging problem due to significant lateral velocity variations near the surface. Velocities beneath this low-velocity anomaly (i.e. beneath the reflector) were not significantly updated during the inversion process, and remained very close to the velocity of the initial model. In general, no useful velocity information can be obtained below the fixed reflector which is located in the lower part of the permafrost. As a result, no velocity information could be recovered close to the base of the permafrost. The high velocity imaged below the reflector at approximately 7.5 km is believed to be artefact. By the same token, velocity structures at the edge of the model area maintained their initial values due to the lack of wave coverage. In addition, the effective bandwidth of the seismic data (between 20 and 80 Hz) suggests that structures with scales larger than 150 m are most reliable and Ap-

pendix B confirms that scales of 150 m are also resolvable by our tomographic algorithm. The heterogeneous structures with smaller scales are likely the result of sparse wave coverage or wave path footprint. Since most recovered features are well above 150 m, the tomographic model in Fig. 8 is sufficiently smooth to the primary wavelength of the data and valid for the application of the Eikonal equation.

Overall, results from the joint inversion provide additional information that could not be recovered with transmission tomography alone, especially for the deeper parts of the permafrost. However, results obtained with this method depend strongly on the accurate positioning of the reflector. Inversion results will tend to underestimate the velocity if such reflector is too shallow or overestimate the velocity if it is too deep. Thus, we assumed that this reflector was well-imaged during the pre-stack time migration process and that interval velocities appropriately positioned the reflector at depth.

4 CONCLUSION

In this study, we present a joint traveltimes tomography method based on FSM, Huygens' Principle and adjoint-state technique. The FSM solves the upwind finite difference formulated Eikonal equation while maintaining the causality of the wave propagation. When combined with Huygens' Principle, the FSM provides traveltimes

of both direct and reflection arrivals. The adjoint method turns the inverse problem to two forward modelling problems and thus circumvents the computation intensive task of the Fréchet derivative matrix estimation.

In a synthetic case example, the joint tomography recovered both shallow and deep anomalies and provided the best image among the three tomographic approaches tested (i.e. transmission only, reflection only and joint approach). The results from 2-D seismic data from the Mackenzie Delta demonstrate some benefits and robustness in recovering both shallow and deep velocity structures in the presence of heterogeneous permafrost and low-velocity zones associated with thermokarst lakes. The permafrost peak velocity zone not present in the starting models is recovered in the field data indicating that large-scale variations within permafrost are well-resolved. The image of small-scale low-velocity anomalies corresponding to geologically disturbed areas such as partially frozen lakes is also enhanced by the joint inversion of transmission and reflection traveltimes. Although the joint tomography provides additional information that could not be recovered with transmission tomography alone, especially for the deeper parts of the permafrost, our method depends strongly on the accurate positioning of the reflector.

ACKNOWLEDGMENT

The authors acknowledge the advices from Dr. Shingyu Leung and Dr. Jianliang Qian on the FSM and adjoint method for first arrival tomography. We greatly appreciate the insightful comments from one anonymous reviewer and Dr. Andreas Fichtner, which have improved the paper considerably. BP Canada Energy Company and Chevron Canada Limited kindly provided access to the 2-D seismic data from Richards Island (line 120). Geological Survey of Canada contribution number is 20110139.

REFERENCES

- Aki, K. & Lee, H.K., 1976. Determination of three-dimensional anomalies under a seismic array using first P arrival times from local earthquakes: A homogeneous initial model. *J. geophys. Res.*, **81**, 4381–4399.
- Bak, S., McLaughlin, J. & Renzi, D., 2010. Some improvements for the fast sweeping method, *SIAM J. Sci. Comput.*, **32**(5), 2853–2874
- Bamberger, A., Chavent, G., Hemons, Ch. & Lailly, P., 1982. Inversion of normal incidence seismograms, *Geophysics*, **47**(5), 757–770.
- Bento, S.H., 1977. The Hamilton–Jacobi Equation: A Global Approach, Mathematics in Science and Engineering Vol. 131, Academic Press, London, 158pp.
- Bellefleur, G., Ramachandran, K., Riedel, M., Brent, T. & Dallimore, S.R., 2008. Recent advances in mapping permafrost and gas hydrate occurrences using industry seismic data from Mallik, Richards Island, Northwest Territories, Canada, in *Proceedings of the Ninth International Conference on Permafrost*, 2008 June 29–July 2, University of Alaska, Fairbanks, p. 6.
- Červený, V. 2001. *Seismic Ray Theory*, 1st edn., Cambridge University Press, Cambridge, 722pp.
- Chapman, C. 2004. *Fundamentals of Seismic Wave Propagation*. Cambridge University Press, Cambridge, 172pp.
- Cheng, Y. & Shu, C.-W., 2007. A discontinuous Galerkin finite element method for directly solving the Hamilton–Jacobi equations, *J. Comput. Phys.*, **223**, 398–415.
- Dallimore, S.R., Collett, T.S. & Uchida, T., 1999. Overview of science program, in *Scientific Results from JAPEX/JNOC/GSC Mallik 2L-38 Gas Hydrate Research Well, Mackenzie Delta, Northwest Territories, Canada*, Geol. Surv. Canada Bull. 544, p. 11–17, eds Dallimore, S.R., Uchida, T. & Collett, T.S., Geological Survey of Canada, Ottawa.
- de Kool, M., Rawlinson, N. & Sambridge, M., 2006. A practical grid based method for tracking multiple refraction and reflection phases in 3-D heterogeneous media, *Geophys. J. Int.*, **167**, 253–270.
- Dziewonski, A.M., Hager, B.H. & O’Connell, R.J., 1977. Large-scale heterogeneities in the lower mantle. *J. Geophys. Res.*, **82**, 239–255.
- Fichtner, A., Bunge, H.-P. & Igel, H., 2006. The adjoint method in seismology: I. Theory, *Phys. Earth planet. Inter.*, **157**, 86–104.
- Judge, A.S., Pelletier, B.R. & Norquay, I., 1987. Permafrost base and distribution of gas hydrates, in *Marine Sciences Atlas of the Beaufort Sea—Geology and Geophysics*, Geol. Surv. Canada Misc. Rep. 40, p. 39, ed. Pelletier, B.R., Geological Survey of Canada, Ottawa.
- Hole, J.A. & Zelt, B.C., 1995. 3-D finite-difference reflection traveltimes, *Geophys. J. Int.*, **121**, 427–434.
- Kim, S. & Cook, R., 1999. 3-D traveltime computation using second-order ENO scheme, *Geophysics*, **64**(6), 1867–1876.
- Leung, S. & Qian, J. 2006. An adjoint-state method for three-dimensional transmission traveltime tomography using first-arrivals, *Commun. Math. Sci.*, **4**, 249–266.
- Lions, J.L., 1971. *Optimal Control of Systems Governed by Partial Differential Equations* (translated by S. K. Mitter), Springer-Verlag, New York, NY, 396pp.
- Li, X. & Ulrych, T.J., 1993. Traveltime computation in discrete heterogeneous layered media, *J. Seismic Explor.*, **2**(4): 305–318
- Matsuoka, T. & Ezaka, T., 1992. Ray tracing using reciprocity, *Geophysics*, **57**, 3276–333.
- Mora, P., 1988. Elastic wave-field inversion of reflection and transmission data, *Geophysics*, **53**, 750–759.
- Osher, S., 1993. A level set formulation for the solution of the Dirichlet problem for Hamilton–Jacobi equations, *SIAM J. Math. Anal.*, **24**, 1145–1152.
- Paige, C. & Saunders, M.A., 1982. LSQR: Sparse linear equations and least squares problems, Part I and Part II, *ACM Trans. Math Softw.*, **8**, 43–71.
- Plessix, R.-E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, *Geophys. J. Int.*, **167**, 495–503, doi: 10.1111/j.1365-246X.2006.02978.x.
- Pratt, R. G., 1999. Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model, *Geophysics*, **64**, 888–901.
- Podvin, P. & Lecomte, I., 1991. Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools, *Geophys. J. Int.*, **105**, 271–284.
- Qian, J. & Symes, W.W., 2002. Finite-difference quasi-P traveltimes for anisotropic media, *Geophysics* **67**, 147–155.
- Qian, J., Zhang, Y.-T. & Zhao, H.-K., 2007. Fast sweeping methods for Eikonal equations on triangular meshes, *SIAM J. Numer. Anal.*, **45**, 83–107.
- Rawlinson, N., Hauser, J. & Sambridge, M., 2008. Seismic ray tracing and wavefront tracking in laterally heterogeneous media, *Adv. Geophys.*, **49**, 203–273.
- Rawlinson, N. & Sambridge, M., 2004. Multiple reflection and transmission phases in complex layered media using a multistage fast marching method, *Geophysics*, **69**, 1338–1350.
- Riahi M.A. & Juhlin, C., 1994. 3-D interpretation of reflected arrivals times by finite-difference techniques, *Geophysics*, **59**, 844–849.
- Rouy, E. & Tourin, A., 1993. A viscosity solutions approach to shape-from-shading, *SIAM J. Numer. Anal.*, **29**, 867–884.
- Sengupta, M.K. & Toksöz, M.N., 1976. Three-dimensional model of seismic velocity variation in the Earth’s mantle, *Geophys. Res. Lett.*, **3**, 84–86.
- Sethian, J.A. 1996. A fast marching level set method for monotonically advancing fronts, *Proc. Natl. Acad. Sci.*, **93**, 1591–1595.
- Shu, C.-W., 2007. High order numerical methods for time dependent Hamilton–Jacobi equations, in *Mathematics and Computation in Imaging Science and Information Processing*, IMS Lecture Notes Series Vol. 11, eds Goh, S.S., Ron, A. & Shen, Z., World Scientific Publishing, Singapore, 262pp.
- Stewart, R.H., 1991. *Exploration Seismic Tomography: Fundamentals*, Society of Exploration Geophysicists, Houston, TX, 190pp.

- Taillandier, C., Noble, M., Chauris, H. & Calandra, H., 2009. First-arrival travel time tomography based on the adjoint state method, *Geophysics*, **74**(6), WCB1–WCB10.
- Tarantola, A., 1984. Linearized inversion of seismic-reflection data, *Geophys. Prospect.*, **32**, 998–1015.
- Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation, *Pure appl. Geophys.*, **128**, 365–399.
- Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 352pp.
- Tromp, J., Tape, C. & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and bananadoughnut kernels, *Geophys. J. Int.*, **160**, 195–216.
- Tsai, Y-H. R., Cheng, L-T., Osher, S. & Zhao, H-K., 2003. Fast sweeping algorithms for a class of Hamilton-Jacob Equations, *SIAM, J. Numer. Anal.*, **53**, 673–694.
- Virieux, J. & Farra, F. 1991. Ray tracing in 3-D complex isotropic media: An analysis of the problem, *Geophysics*, **56**, 2057–2069.
- Vidale, J.E. 1988. Finite-difference calculations of traveltimes, *Bull. seism. Soc. Am.*, **78**, 2062–2076.
- Xiong, T., Zhang, M., Zhang, Y.-T. & Shu C.-W., 2010. Fast sweeping fifth order WENO scheme for static Hamilton-Jacobi equations with accurate boundary treatment, *J. Sci. Comput.*, **45**, 514–536, doi:10.1007/s10915-010-9345-6.
- Zhang, Y.-T., Zhao, H.-K. & Qian, J., 2006. High order fast sweeping methods for static Hamilton-Jacobi equations, *J. Sci. Comput.*, **29**, 25–56.
- Zhao, H-K., 2004. A fast sweeping method for Eikonal equations, *Math. Comput.*, **74**, 603–627.
- Zhao, H-K., 2007. Parallel implementation of the fast sweeping method, *J. Comput. Math.*, **25**(4), 421–429.

APPENDIX A: ACCURACY ANALYSIS OF THE FAST SWEEPING METHOD

In Section 2.1, we proposed to use FSM to efficiently solve the Eikonal equation (eq. 1). In this section, we compare traveltome of both direct and reflected arrivals calculated using FSM with the analytical solution of a simple 2-D layered model with one dipping reflector (Fig. A1). The analytical solution is obtained by directly calculating the incidence and reflection wave travelling distance normalized by the propagation speed in the first layer. The match between FSM arrival times and the analytical results (see Fig. A2a) confirms that the grid-based Eikonal equation solver can provide an accurate traveltome for the transmitted and reflected body waves. To eliminate the effect of the numerical errors on the final tomographic model, the grid interval with numerical errors less than the error in field data traveltome picking should be used.

APPENDIX B: RESOLUTION AND SENSITIVITY ANALYSIS

In Section 3, the joint tomography is applied to synthetic data and to a 2-D seismic profile from Mackenzie Delta, Northwest Territories of Canada. To analyse the resolution of the tomographic model, we applied the joint tomography algorithm to the synthetic traveltome simulated from checkerboard models (Fig. B1c). These models result from the superposition of checkerboard anomaly (Fig. B1b) and a background velocity with vertical gradient (Fig. B1a). The 2-D synthetic acquisition geometry is based on seismic line 120 and has an identical surface topography, and the same distribution

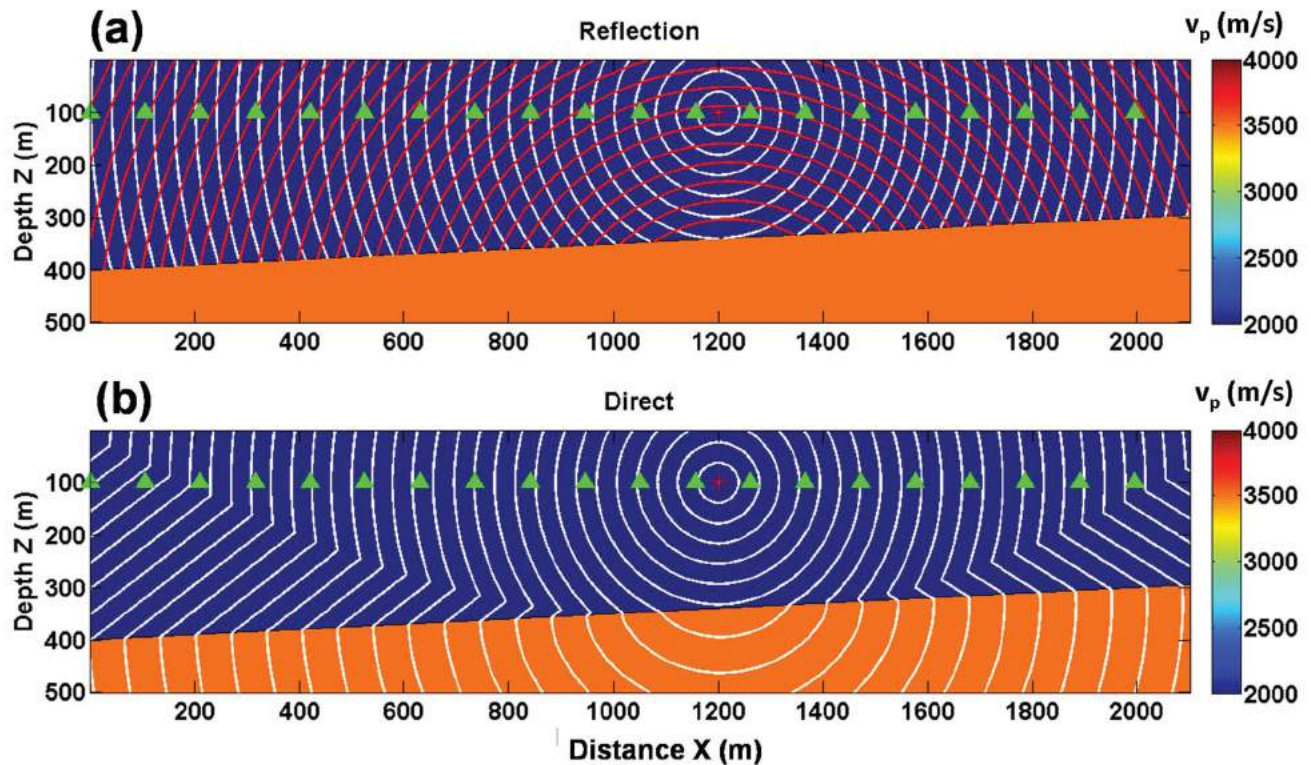


Figure A-1. A 2-D model ($500 \text{ m} \times 2100 \text{ m}$) with a dipping layer is used to solve traveltimes with the FSM. The star and triangles denote the location of sources and receivers. The velocity above and below the reflective interface is 2000 m s^{-1} and 3500 m s^{-1} , respectively. (a) The traveltome of down-going (incident wave) and up-going (reflected wave) solved with the FSM. (b) The traveltome of direct arrivals. The contours show the propagation of the wave fronts.

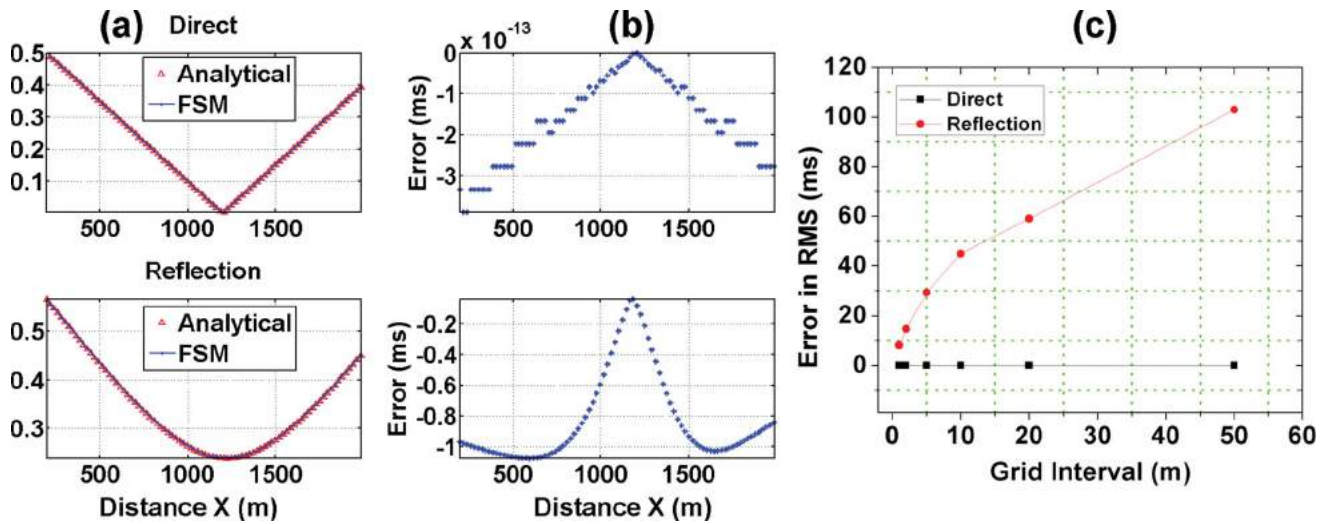


Figure A-2. The analytical direct and reflected arrivals are compared with the numerical results from the 2-D model discretized with different grid intervals, i.e., 1 m, 2 m, 5 m, 10 m, 20 m and 50 m. (a) The analytical traveltimes are superimposed with the numerical results from the 2-D model with 1 m grid interval. (b) The error between the analytical and numerical results (1 m grid interval) generally increases with offset. (c) The rms error is shown as a function of grid interval. The rms error of the direct traveltimes remains negligible for all test grid intervals due to the simple wave path, whereas the rms error of the reflection traveltimes increases with grid interval. Thus the grid interval should be used that reduces the numerical error to the level lower than the errors in the field data traveltimes picking.

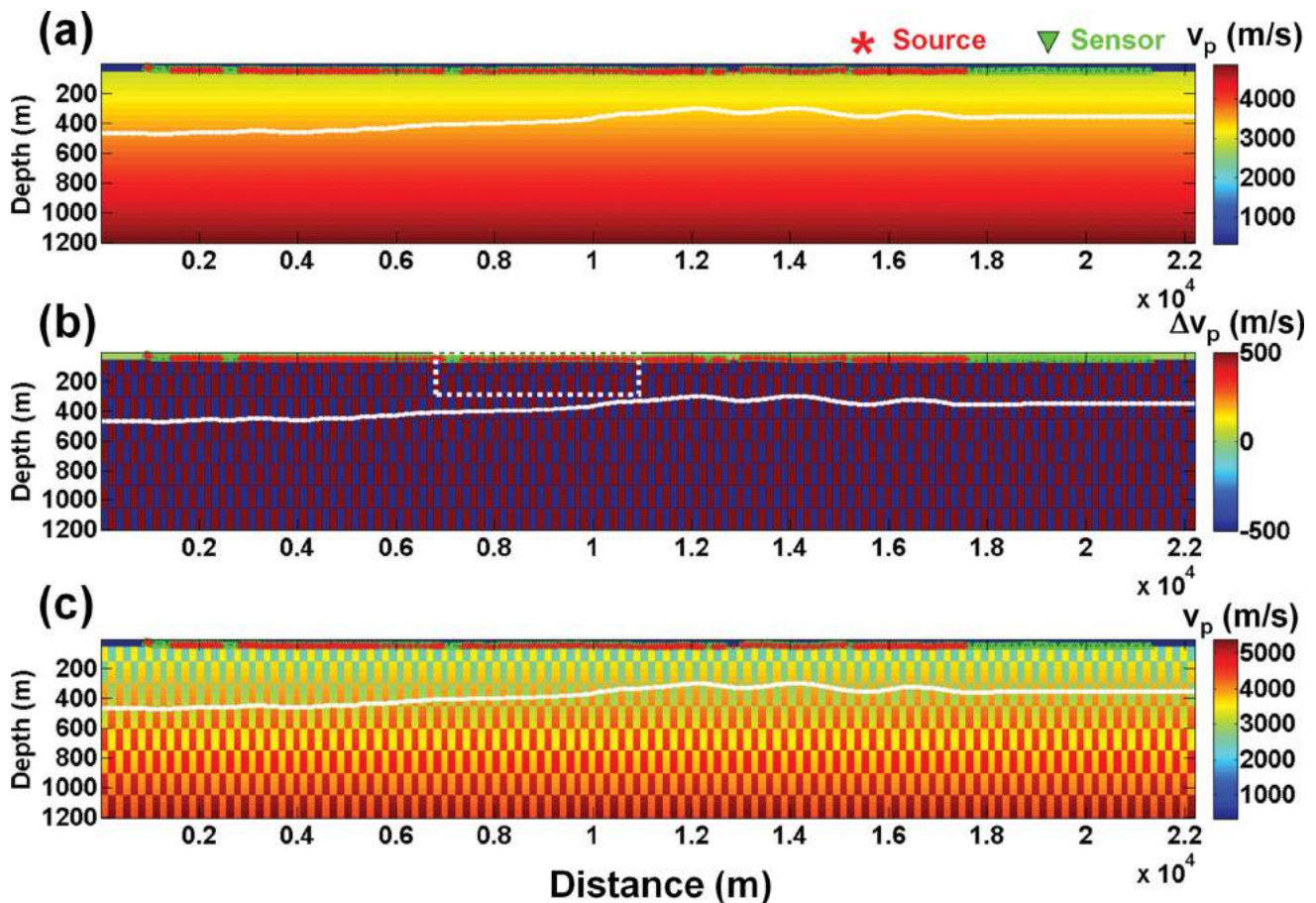


Figure B-1. (a) The background velocity increases with depth. (b) The velocity perturbations with checkerboard pattern. (c) The true model is composed of the background velocity in (a) and the checkerboard style perturbations in (b). The white line denotes the reflector used for reflection tomography. A representative region (dashed rectangle) is selected for display purpose in the following figures. The source (star) and receivers (triangle) are distributed along the irregular surface topography.

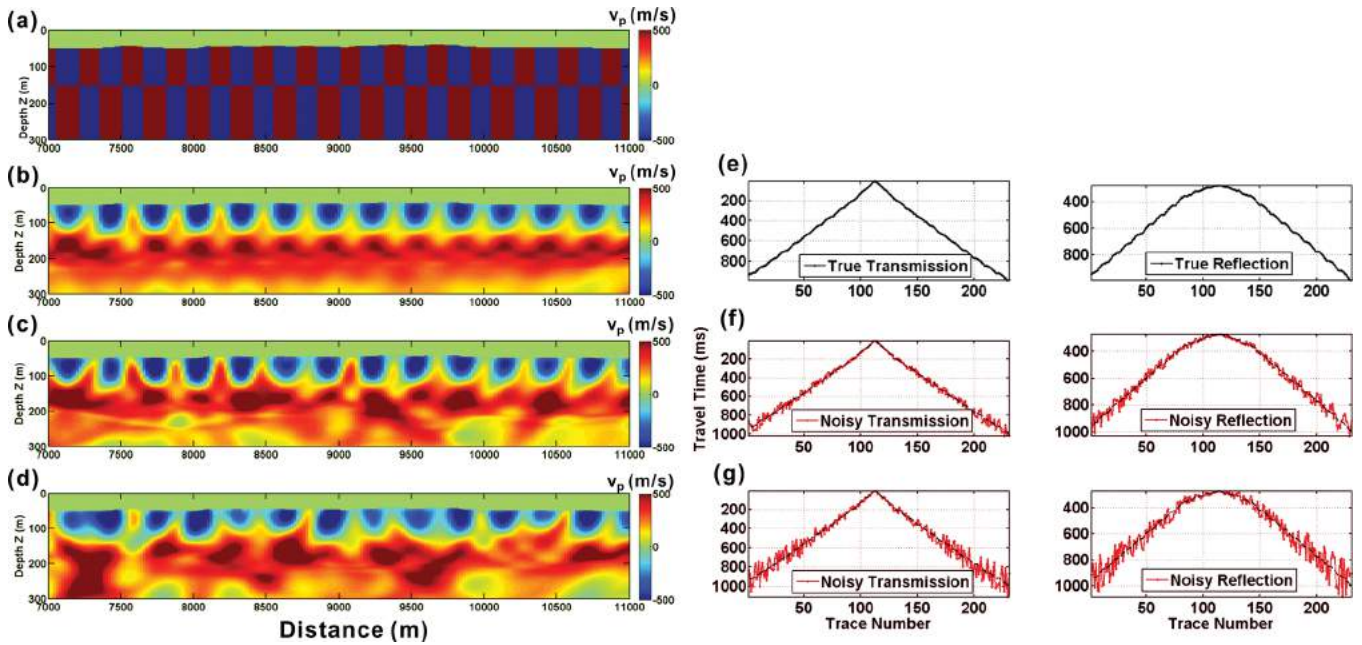


Figure B-2. The joint tomographic images from the checkerboard model in Fig. B-1. The size of each square tiles is 150 m by 150 m. Only the part from the representative region is shown for display purpose. (a) The true model. (b) Joint tomographic results without noise. (c) and (d) Joint tomographic results with different magnitude of noise. (e) Samples of direct and reflected arrivals without noise from one shot. (f) and (g) Samples of direct and reflected arrivals with different magnitude of noise. The magnitude of the noise is increasing with offset to mimic lower signal-to-noise ratio often observed at large offset in practical situations. The joint tomography recovers the top 200 m of the checkerboard perturbations.

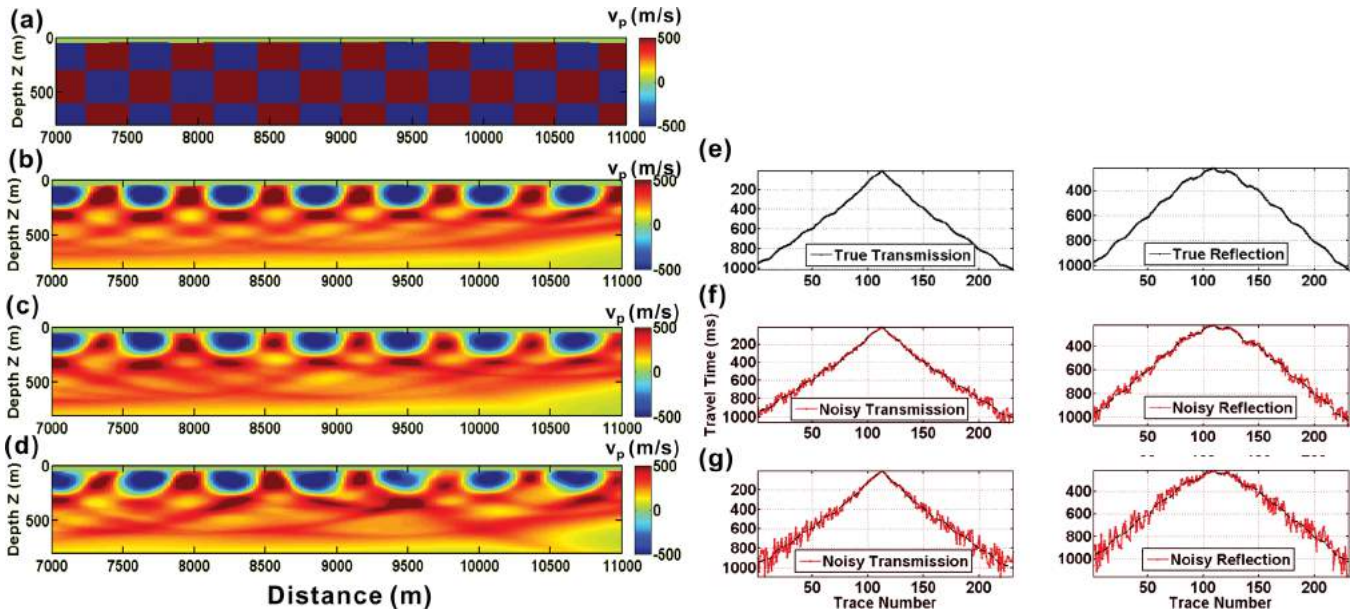


Figure B-3. The joint tomographic images similar to the ones in Fig. B-2. The size of the square tiles is now 300 m by 300 m. (a) is the true model. (b) shows the joint tomographic model without noise, whereas (c) and (d) show the joint tomographic model with different magnitude of noise. (e), (f) and (g) show samples of direct and reflected arrivals without and with different magnitude of noise. The joint tomography recovers the top 400 m of the checkerboard perturbations.

of sources and receivers. The synthetic traveltome picks used in the joint inversion are also representative of picks made on real data to ensure that input data are not unrealistically oversampled. Following this, the input traveltome used the same number of transmission and reflection arrivals that could be picked on real data from line 120 (77 220 and 14 862, respectively). The joint tomography uses the background velocity as the initial model and

is applied to the synthetic traveltome with and without contamination of random noise which increases with offset. Two different sizes of checkerboard are used, 150 m (Fig. B2) and 300 m (Fig. B3). It is shown that the joint tomography is efficient and robust in recovering the top 200 and 400 m velocity characteristics of the checkerboard perturbation even in the presence of random noise.