GEOPHYSICS, VOL. 63, NO. 5 (SEPTEMBER-OCTOBER 1998): P. 1726–1737, 11 FIGS.

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Nonlinear refraction traveltime tomography

Jie Zhang* and M. Nafi Toksöz‡

ABSTRACT

A few important issues for performing nonlinear refraction traveltime tomography have been identified. They include the accuracy of the traveltime and raypath calculations for refraction, the physical information in the refraction traveltime curves, and the characteristics of the refraction traveltime errors. Consequently, we develop a shortest path ray-tracing method with an optimized node distribution that can calculate refraction traveltimes and raypaths accurately in any velocity model. We find that structure ambiguity caused by short and long rays in the seismic refraction method may influence the inversion solution significantly. Therefore, we pose a nonlinear inverse problem that explicitly minimizes the misfits of the average slownesses (ratios of traveltimes to the corresponding ray lengths) and the apparent slownesses (derivatives of traveltimes with respect to distance). As a result, we enhance the resolution as well as the convergence speed. To regularize our inverse problem, we use the Tikhonov method to avoid solving an ill-posed inverse problem. Errors in refraction traveltimes are characterized in terms of a common-shot error, a constant deviation for recordings from the same shot, and a relative traveltime-gradient error with zero mean with respect to the true gradient of the traveltime curve. Therefore, we measure the uncertainty of our tomography solution using a nonlinear Monte Carlo approach and compute the posterior model covariance associated with two different types of random data vectors and one random model vector. The nonlinear uncertainty analysis indicates that the resolution of a tomography solution may not correspond to the ray coverage. We apply this tomography technique to image the shallow velocity structure at a coastal site near Boston, Massachusetts. The results are consistent with a subsequent drilling survey.

INTRODUCTION

Using seismic refraction traveltime data for imaging the subsurface has long been a standard technique. It is appealing because of its low-cost field operation and easy data interpretation. However, conventional processing of refraction data falls short of showing the true strength of this technique because conventional processing uses oversimplified geometry and media, although it does attempt to develop unique physical concepts from refraction traveltimes. On the other hand, modern tomography methods seem to ignore the value of physical concepts that have been already established for the refraction problem. For example, the derivatives of traveltimes with respect to distance that are explicitly used in the conventional methods [e.g., the generalized reciprocal method (Palmer, 1980)] are not accounted for in refraction traveltime tomography. Existing refraction tomography methods simply match the absolute traveltimes using a least-squares criterion as do other tomography methods (e.g., reflection and crosshole tomography).

A number of refraction traveltime tomography methods have been developed. White (1989) described a refraction traveltime tomography method that applies a two point raytracing algorithm and solves a damped least-squares problem for both velocities and refractor depths. The inverse problem is regularized with a gradient smoothing operator in a creeping manner [for notation, see Scales et al. (1990)]. Zhu and McMechan (1989) performed refraction tomography using an analytic traveltime solution; their tomography approach is the same for the crosswell geometry (McMechan et al., 1987) and requires the initial model to have positive velocity gradients. Stefani (1995) showed a turning-ray tomography which is similar to White's (1989), but he inverted for velocities only. Using a finite-difference approach to solve the eikonal equation

Manuscript received by the Editor August 15, 1996; revised manuscript received February 9, 1998.
*Formerly Earth Resources Lab., Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02142; presently GeoTomo LLC, 225 South Boulder Road, Louisville, CO 80027. E-mail: jie@gmg.com.
‡Earth Resources Lab., Department of Earth, Atmospheric, and Planetary Sciences, E34-462 MIT, 42 Carleton Street, Cambridge, MA 02142. E-mail: nafi@erl.mit.edu.
© 1998 Society of Exploration Geophysicists. All rights reserved.

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

without rays involved, Ammon and Vidale (1993) developed a refraction traveltime tomography method that regularizes the inverse problem with second-order model derivatives in a jumping fashion [for notation, see Scales et al. (1990)]. Because Ammon and Vidale explicitly construct a sensitivity (Fréchet-derivative) matrix by repeating the forward calculation for each perturbed slowness cell, their approach is limited to a small number of model grids.

The existing techniques for refraction traveltime tomography are limited because of the ray-tracing approach used, or because they encounter ill-posed inverse problems or both. They also ignore the derivative information in the refraction traveltime data. In addition, some of these methods display raypaths to show the resolution of the solution; others use a resolution matrix to estimate the uncertainty associated with fixed raypaths. The nonlinear uncertainty of a refraction tomography solution has not been explored.

The use of the traditional two-point ray-tracing algorithms limits the accuracy of tomography inversion. The ray methods suffer from the problem of converging to a local minimum traveltime path and occasionally missing the global minimum (Moser, 1991). Comparison of several existing ray-tracing methods indicates that the finite-difference solution to the eikonal equation by Vidale (1988) is the fastest one (Matarese, 1993). However, Qin et al. (1992) showed that the finite-difference approach (Vidale, 1988) violates causality when the velocity contrast is larger than $\sqrt{2}$ because of expanding a square wavefront. Hole and Zelt (1995) applied a partial correction to Vidale's finite-difference solution on the basis of Fermat's principle. A complete correction to the finite-difference solution by expanding its virtual wavefront with Huygens's principle can guarantee the global minimum traveltimes in any complex model (Qin et al., 1992), but it loses the simplicity in the finite-difference method by Vidale (1988) and becomes more expensive.

An ill-posed inverse problem must be properly regularized in order to obtain a stable solution, which is independent of the model discretization (Delprat-Jannaud and Lailly, 1993; Zhang et al., 1996). Rodi (1989) and Zhang et al. (1996) showed that regularizing the model stepsize (creeping) rather than the model itself (jumping) cannot ultimately avoid solving an ill-posed inverse problem. Further, not all the regularization approaches perform well, and an appropriate criterion for refraction traveltime tomography must be defined.

In this study, we present a nonlinear refraction traveltime tomography method that consists of a new version of the shortest path ray-tracing approach (SPR), a regularized nonlinear inversion method that inverts "traveltime curves" rather than traveltimes alone, and a Monte Carlo method for nonlinear uncertainty analysis of the final solution. We demonstrate the application of this approach to real data from a small-scale refraction survey conducted at a coastal site near Boston, Massachusetts.

CALCULATING REFRACTION TRAVELTIMES AND RAYPATHS

To conduct a tomography study, we need a ray-tracing technique to calculate both traveltimes and raypaths of the first-arrival refractions or turning rays. Ray tracing is the most time-consuming step of nonlinear seismic traveltime tomography. Accuracy of both traveltimes and raypath loci is also important because the accuracy of the tomography inversion in part depends on the errors introduced by the forward calculation. In particular, the calculation of refraction traveltime has been difficult simply because of its highly variable raypath.

Recently, tremendous progress has been made in developing wavefront ray-tracing methods on a regular grid. These include three categories; solving an eikonal equation using finite-difference methods (e.g., Vidale, 1988; Qin et al., 1992, Hole and Zelt, 1995), applying an analytical solution to expand a wavefront (e.g., Vinje et al., 1993), and using graph theory to expand a wavefront by finding the shortest path (e.g., Moser, 1989, 1991; Saito, 1989, 1990). Wavefront methods are attractive because they can simulate wave propagation in the entire model and find diffracted raypaths, head waves, and raypaths to shadow zones. Among those methods, we chose the shortest path ray-tracing method (SPR) because it is more flexible and one can achieve desired accuracy by adjusting the graph template size. However, for highly accurate results, the methods of Saito (1989, 1990) and Moser (1989, 1991) require vast memory and intensive calculation. Recent improvements have been made by Fischer and Lees (1993) and Klimes and Kvasnicka (1993).

Seismic raypaths can be found by calculating the shortest traveltime paths through a network that represents the earth. The network consists of nodes, and the node connection is based on a graph template. The SPR method includes three steps: (1) timing nodes along an expanding wavefront from its original source or secondary source, (2) finding the minimum traveltime point along the wavefront and taking this point as a new secondary source, and (3) expanding the wavefront from this minimum time point. These three steps are repeated until the whole model is traced. With a small number of ray legs in a graph template, SPR usually yields zig-zag raypaths in homogeneous or smooth velocity zones in the model, and produces longer raypaths (Fischer and Lees, 1993). We improve the SPR method by sampling the wavefront with a uniform angle coverage and eliminating unnecessary nodes in the slowness network by analyzing the slowness model prior to ray tracing.

The traveltime errors in SPR result from space and angle discretization. These two errors are independent, that is, decreasing the grid size does not reduce the error due to a finite angle coverage (Moser, 1991). The error due to the angle discretization can be reduced by optimizing the node distribution in a graph template prior to the ray tracing. Because the angle error is associated with the largest angle that can be sampled by rays in a graph template, the angle differences for a given number of rays in a graph template should be minimized for accurate calculation. Figure 1 shows two different graph templates that contain two nodes on each grid boundary. The regular node distribution shown by the dash line was used previously with $a_1 = 0.25dx$, $a_2 = 0.5dx$, and $a_3 = 0.25dx$ (Nakanishi and Yamaguchi, 1986; Moser, 1991). The solid line shows the iso-angle node distribution that we use. When we choose $b_1 = 0.29289dx$, $b_2 = 0.41421dx$, and $b_3 = 0.29289dx$ (see Figure 1), the angle between two adjacent rays in each cell is equal (22.5°). If the number of nodes on each grid boundary is more than two, then the optimized propagation angles cannot be exactly equal, but they can be optimized to have a minimum difference. Further node optimization is made by analyzing the velocity model prior to the ray tracing. Instead of timing all the

network nodes with the same graph template, the areas in the model that are smooth enough (i.e., having velocity variations less than a predetermined value) are eliminated on the basis of a velocity-gradient lookup table. This results in a sparse node distribution in the model. The use of this sparse-node model avoids the zig-zag raypath problem and enhances numerical efficiency as well.

Figure 2 shows a comparison of traveltime errors caused by different node distributions used in a ray-tracing calculation (i.e., regular node, iso-angle node, and iso-angle and sparse node). The model consists of two layers and is gridded into 200 × 50 uniform cells (spacing interval 1.0 m). In this example, we show results using only two nodes on each cell boundary, as shown in Figure 1. In modeling real data, we usually use four or five nodes on each cell boundary for higher accuracy. Regular nodes produce large errors in some areas, including refraction in this case because of uneven angle sampling. Without additional computational effort, simply adjusting the node distribution according to Figure 1 reduces the error by nearly 50% for refraction traveltimes. Further optimizing the nodes by eliminating those in the smooth-velocity areas prior to ray tracing gives results with negligible errors.

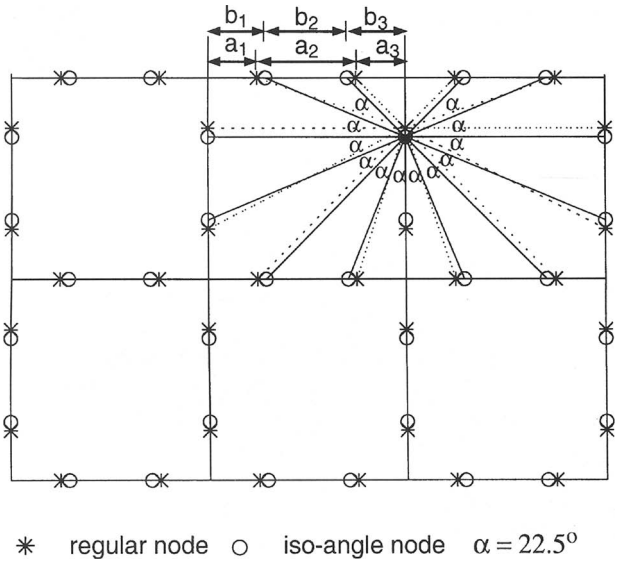

FIG. 1. Two graph templates for the shortest path ray-tracing method. A star denotes the regular-node template that samples the grid interface equally ($a_1 = 0.25$, $a_2 = 0.50$, and $a_3 = 0.25$); the open circle represents the iso-angle node ($b_1 = 0.29289dx$, $b_2 = 0.41421dx$, and $b_3 = 0.29289dx$). The iso-angle node distribution allows the angles between two adjacent rays in each cell to be approximately equal, reducing error due to the angle discretization.

MINIMIZING A PHYSICALLY MEANINGFUL OBJECTIVE FUNCTION

We solve a regularized nonlinear inverse problem. Starting from an initial model, we iteratively update traveltimes and raypaths without assuming any interfaces or velocity functionals as defined by few parameters in the model. The solution of the refraction traveltime inversion is nonunique. There are many solutions that can predict the observed data equally well. However, not all these solutions are physically meaningful. To obtain one optimal solution, our strategy is to establish quantitative criteria for a meaningful data fit and a meaningful model correlation when sufficient physical concepts are involved. Instead of inverting absolute traveltimes, we minimize the misfits of the average slownesses (traveltimes divided by ray lengths) and the apparent slownesses (traveltime derivatives with respect to distance). We apply the Tikhonov regularization (Tikhonov and Arsenin, 1977) to explicitly constrain the model roughness. Specifically, we minimize the following objective function,

$$\Phi(m) = (1-\omega)\|C_\ell(d - G(m))\|^2 + \omega\|D_x(d - G(m))\|^2 + \tau\|Rm\|^2 \quad (1)$$

$$= (1-\omega)\|\bar{d} - \bar{G}(m)\|^2 + \omega\|\hat{d} - \hat{G}(m)\|^2 + \tau\|Rm\|^2 \quad (2)$$

$$= (1-\omega)S_1 + \omega S_2 + \tau S_3, \quad (3)$$

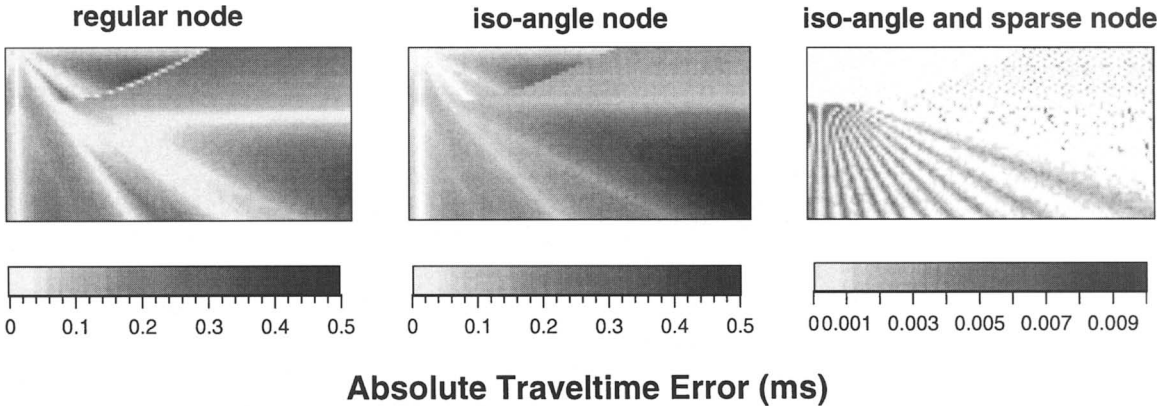

FIG. 2. Comparison of ray-tracing accuracy due to different graph templates used in the shortest path method. A two-layer model (200 × 50 m) is discretized with a grid size of 1.0 m. The upper layer velocity is 2500 m/s, the lower layer velocity is 4500 m/s. Using two regular nodes on each cell interface produces refraction traveltime errors up to 0.4 ms. Shifting these two nodes to achieve equal angles (22.5°) reduces the refraction traveltime error to 0.2 ms. Further eliminating nodes in the homogeneous areas reduces the error to a negligible level.

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

$$\bar{d} \overset{def}{=} \frac{d}{\ell}: \text{average-slowness data,} \tag{4}$$

$$\hat{d} \overset{def}{=} \frac{\partial d}{\partial x}: \text{apparent-slowness data,} \tag{5}$$

where d is the traveltime data; $G(m)$ is the calculated traveltime data for the current model m, and C_ℓ is an operator that scales a traveltime with the corresponding ray length ℓ, and returns an average slowness $\bar{d}$, along the raypath. Here, the ray length $\ell(m)$, is a variable parameter which is updated iteratively in the tomography inversion. Later, we will show that using the shot-receiver distance to replace ℓ in the inversion plays the same role as ℓ. D_x is a differential operator for traveltime with respect to distance, $\hat{d} = D_x d$ returns the gradients of the traveltime-distance curves (apparent slownesses), R is a regularization operator (e.g., derivative operator), τ is a smoothing trade-off parameter, and ω is a weighting factor between the average-slowness misfit norm and the apparent-slowness misfit norm. As one can see, the objective function (1) contains three terms all associated with slownesses in one way or another: the misfit of average slownesses, the misfit of apparent slownesses, and the roughness of model slowness.

Inversion algorithm

We want to avoid explicitly dealing with the sensitivity matrix in the inversion, that is, one should implicitly account for the effects of sensitivity matrices rather than explicitly constructing or storing them for the inversion (Scales, 1987; Mackie and Madden, 1993; Zhang et al., 1995). For example, using the Gauss-Newton (GN) method to linearize the stationarity equation associated with minimizing the objective function (1) and then applying a conjugate gradient (CG) technique to solve the inversion for each iteration (Scales, 1987) is one of these approaches. Directly minimizing the objective function using a CG method (Nolet, 1987) is another. Both techniques avoid the storage of any matrices, but only store the results of a matrix multiplying a vector or its transpose multiplying a vector. We take the first approach to demonstrate our concepts in inverting refraction traveltime data. This leads to the following equation,

$$\begin{aligned}&\left((1-\omega)A_k^T A_k + \omega B_k^T B_k + \tau R^T R + \epsilon_k I\right)\Delta m_k \\ &= (1-\omega)A_k^T(\bar{d} - \bar{G}(m_k)) + \omega B_k^T(\hat{d} - \hat{G}(m_k)) \\ &\quad - \tau R^T R m_k,\end{aligned} \tag{6}$$

$$A \overset{def}{=} \frac{\partial \bar{G}}{\partial m} = \frac{1}{\ell}\cdot\frac{\partial G}{\partial m}, \tag{7}$$

$$B \overset{def}{=} \frac{\partial \hat{G}}{\partial m} = \frac{\partial^2 G}{\partial m \partial x}, \tag{8}$$

$$m_{k+1} = m_k + \Delta m_k, \qquad k = 1, 2, 3, \ldots, N, \tag{9}$$

where each nonzero term in A_k (average-slowness sensitivity matrix) contains the length of a local ray across a model cell divided by the entire ray length, and each nonzero term in B_k (apparent-slowness sensitivity matrix) represents the differences of the ray lengths in a cell divided by the receiver spacing when the locations of two adjacent receivers are used to calculate derivatives.

Linearizing a nonlinear inverse problem in an iterative fashion may be highly efficient for numerical computation. Linearization does not necessarily lead to a "poor" solution or miss the "global" solution. In fact, the suitability of linearization depends on how a nonlinear function is linearized and for what scale the linearization is applied, rather than linearization itself. We account for nonlinearity by adding a variable damping parameter in the GN method as suggested by Morgan (1981). Because we explicitly minimize model roughness as described by equation (1), large nonlinearity occurs only in the early inversion stage due to a poor starting model, and the objective function approaches a quadratic form when the data misfits [S_1 and S_2 in equation (1)] are small and the model roughness (S_3) becomes dominant. Following a procedure described by Morgan (1981), we add a variable damping term $\epsilon_k I$, to the left-hand side of equation (6) with $\epsilon_k = \alpha \times rhs$, where α is an empirical parameter (about 0.1) and *rhs* is the rms misfit norm of the right-hand side of equation (6). If the objective function is not minimized well and remains quite nonlinear, then *rhs* is large and a large damping ϵ_k is automatically applied, resulting in only small model updates. With the inversion proceeding further and *rhs* decreasing, a smaller ϵ_k drives the convergence speed faster. Several authors have applied this approach to solve nonlinear electrical tomography problems (e.g., Mackie and Madden, 1993; Zhang et al., 1995).

Average-slowness and apparent-slowness data

We chose to minimize the misfits of average-slowness and apparent-slowness data instead of absolute traveltimes (integrated slownesses). This leads to the inversion of traveltime curves rather than traveltimes alone. Inverting average slownesses tends to reconstruct the shallow earth structure, whereas inverting apparent slownesses helps recover the deep structure. Hence, jointly inverting both average-slowness and apparent-slowness data with a variable weighting factor allows us to isolate data information and invert for the slownesses progressively from the shallow to the deep structure.

Figure 3 shows several possibilities for calculated traveltimes that give the same rms traveltime misfit but may correspond to different models. The calculated traveltime at the receiver R_0 is assumed to be fixed. If the next receiver is placed at

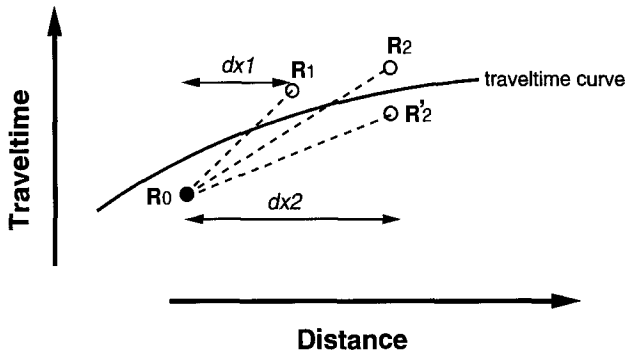

FIG. 3. Schematic traveltime fit. The calculated traveltimes at R_1, R_2, and R_2' all have the same rms misfit, but their traveltime gradients relative to R_0 fit the data gradients differently.

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

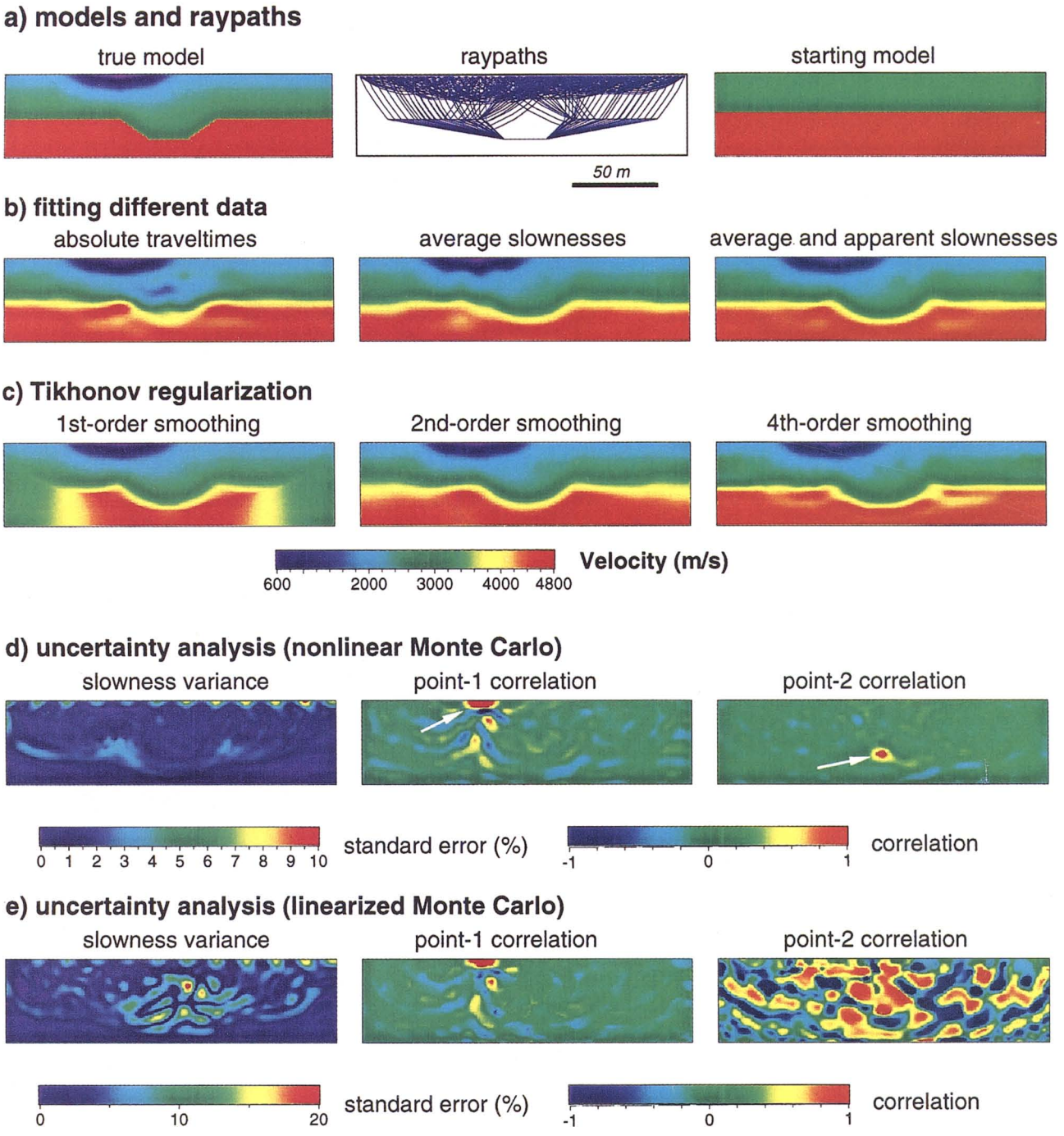

FIG. 4. Numerical experiment for resolving a graben model containing a near-surface low-velocity zone. (a) "True" velocity model (200 × 50 m) gridded into 200 × 50 cells, raypaths for 12 sources and 48 receivers on the surface, and an estimated starting model for inversion. (b) With the third-order derivative regularization, inversion results from inverting traveltime data only, average slownesses only (ratios of traveltimes to the ray lengths), or both average slownesses and apparent slownesses (traveltime gradients). The last one shows the best performance for resolving the graben shape and velocities. (c) Inverting average slowness and apparent slowness data, tomography results using first-order derivative operator (over-smoothed), second- and fourth-order derivative operators for regularization. (d) Uncertainty analysis using nonlinear Monte Carlo inversions. It shows percent standard error in the reconstructed slowness values as a function of position, and model correlation at two selected points. (e) Uncertainty analysis using linearized Monte Carlo inversions. It correctly estimates the variance of the first point, but fails for the second point where rays do not pass.

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

R_1 or R_2 and the rms traveltime misfit remains constant, then the apparent-slowness misfits (traveltime gradient) for R_1 and R_2 are different, implying two different models. One must account for the receiver spacing when measuring the traveltime misfit. For the same receiver location (R_2 and R_2'), the same rms traveltime misfit may also correspond to two different traveltime gradients. Jointly inverting average-slowness data and apparent-slowness data, one can reconstruct a model that fits the traveltime curves rather than data points alone.

Figure 4a shows a numerical model consisting of 200×50 cells with a uniform grid spacing of 1.0 m. Its velocities range from 600 to 4500 m/s, and the structure consists of a near-surface low-velocity zone and a deep graben refractor. A total of 12 shots are assumed with 48 surface receivers at a 4.0 m spacing. Figure 4a also illustrates the first-arrival raypaths for this survey. We first estimate the upper-bound velocities using a GRM-type velocity analysis function (Palmer, 1980) and construct a two-layer starting model. Figure 4b demonstrates tomography results obtained by inverting different types of data: integrated slownesses, average slownesses, and jointly average slownesses and apparent slownesses. A third-order derivative operator [R in equation (1), further discussion will follow] is applied in all three inversions for regularization, and appropriate parameters are selected so that all these inversions converge to the same rms traveltime misfit (0.1 ms). It appears that inverting the absolute traveltimes tends to overestimate the features in the deep structure. Given an initial guess of the model, a larger misfit almost always occurs for the longer range traveltimes. Therefore, inverting the absolute traveltimes attempts to produce deep features before the shallow structure is well reconstructed and quickly reduces the misfit variance mostly associated with long rays. On the other hand, inverting average slownesses constrains the shallow velocities, but cannot precisely reproduce the deep structure due to a larger misfit variance allowed for the longer rays. Finally, an optimal solution can be obtained only by jointly inverting the average slownesses and the apparent slownesses.

Figure 5 illustrates the progress of the inversion in terms of rms average-slowness misfit, rms apparent-slowness misfit, and the objective function. Two experiments are conducted. One, inverts average-slowness data only, the other, inverts average-slowness data during the first six iterations and jointly inverts apparent-slowness data from the seventh iteration when the misfit of the average-slowness data is sufficiently small. It appears that the misfit of traveltime gradients rapidly decreases after the inversion starts to explicitly minimize the misfit of apparent slowness (see Figure 5b). An appropriate weighting factor [ω in equation (1)] is selected so that the rms misfit of the average slowness and the apparent slowness converge to a similar variance, corresponding to the compatible resolving capability. Figure 5a shows that one can also invert pseudo–average slownesses (traveltimes divided by shot-receiver distances) instead of "true" average slownesses (traveltimes divided by ray lengths). In terms of misfit measurements, they are nearly identical, but inverting pseudo–average slownesses avoids tracking exact ray lengths and thus saves computation time.

Figure 6 shows the traveltime fit for the noise-free data at iteration 4 (inverting average slowness only) and at iteration 25 (jointly inverting average slowness and apparent slowness from the seventh iteration). At an early inversion stage (iteration 4), traveltimes associated with short rays are preferentially fitted. With the inversion proceeding, traveltimes associated with long rays are also fitted well due to inverting the traveltime gradients (apparent slowness) until the misfit of the traveltime curves becomes negligible (iteration 25).

Traveltime calculation is the most time-consuming step of nonlinear traveltime tomography. As we described earlier, our traveltime calculation method defines a node distribution on the basis of the model complexity. Therefore, the CPU time for each inversion iteration is nonuniform. Figure 7 presents CPU time for the tomography experiment inverting both average-slowness data and apparent-slowness data. Given a simple two-layer starting model, the inversion is fast at the beginning, but becomes slow when the detailed features appear in the model, and eventually stabilizes when the whole model is reconstructed. This calculation was conducted on a DEC 3000 workstation.

Tikhonov regularization

When an inverse problem is ill-posed, regardless of how sophisticated the optimization approach is, there can be no definitive "solution" solely on the basis of fitting the data. In other words, minimizing only the data misfit (even including the traveltime gradient data) does not have one "global solution." Of course, one could increase the grid size up to the

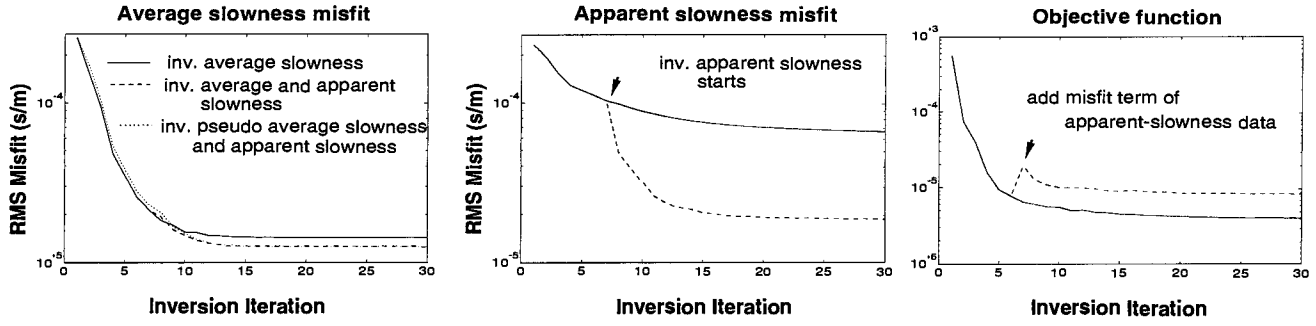

FIG. 5. Two inversions using the third-order derivative operator for regularization but minimizing different data misfits are conducted for the model shown in Figure 4a. One inversion (solid curve) inverts average-slowness data only, the other (dash curve) inverts average-slowness data for the first six iterations, then jointly inverts average-slowness and apparent-slowness data. The apparent-slowness misfit is always calculated and displayed for both inversions even when it is not minimized. Note a rapid decrease of the rms misfit of the apparent slownesses when it is explicitly minimized. The objective function becomes slightly larger because of the sudden inclusion of an ℓ_2 norm of the traveltime gradient misfit.

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

point where the number of grids is equal to the rank of the least-squares matrix. But such a coarse model would not be sufficient to produce accurate forward solutions. For refraction traveltime tomography, most methods attempt to apply ad hoc constraints to keep the inversion stable. These include filling areas between rays with nonzero sensitivity (e.g., Cai and Qin, 1994), applying model gradient derivatives to damp the model step size in a creeping manner (e.g., White, 1989), or using a posterior low-pass filter to smooth the model itself after each iteration (e.g., Zhu and McMechan, 1989). We chose to solve an inverse problem that explicitly minimizes data misfit as well as model roughness using the Tikhonov regularization that mitigates an ill-posed inverse problem and provides parameter constraints for an infinite number of unknowns (Delprat-Jannaud and Lailly, 1993; Zhang et al., 1996).

Applying Tikhonov regularization involves the use of model derivative operators, and may produce a "smooth" solution. However, when we say "model smoothness" or "model roughness," we must ask, "by what criterion?" In fact, we found that not all smoothness criteria of the Tikhonov method perform equally well for refraction traveltime problems. We investigate the following criteria:

first-order smoothing: $R = \nabla$,

$$S_3 = \int (\nabla m(\mathrm{x}))^2 \, d\mathrm{x}, \tag{10}$$

second-order smoothing: $R = \nabla^2$,

$$S_3 = \int (\nabla^2 m(\mathrm{x}))^2 \, d\mathrm{x}, \tag{11}$$

third-order smoothing: $R = \nabla^3$,

$$S_3 = \int (\nabla^3 m(\mathrm{x}))^2 \, d\mathrm{x}, \tag{12}$$

fourth-order smoothing: $R = \nabla^4$,

$$S_3 = \int (\nabla^4 m(\mathrm{x}))^2 \, d\mathrm{x}. \tag{13}$$

Each of these derivatives yields one criterion of smoothness. One can also apply a higher order derivative operator and define a different smoothness criterion for inversion. Figure 4c shows a comparison of the inversion results for smoothness of three different criteria; the first-order smoothing, the second-order smoothing, and the fourth-order smoothing. The inversion result with the third-order operator is already shown in Figure 4b. All these numerical experiments are performed by inverting both average-slowness and apparent-slowness data. We select appropriate parameters to allow them to converge to the same misfit level.

The results show that the use of the first-order smoothness criterion for smoothing (gradient smoothing) produces over-smoothed images, whereas the higher order derivative operators allow us to recover the nonlinear variations in slowness. Although various theoretical studies (e.g., Rodi, 1989; Delprat-Jannaud and Laily, 1993) were conducted to find the roles of the regularization operators in geophysical inversion, the use of the first-order derivative operator is numerically equivalent to the application of a linear interpolation method in the model space, and the minimization of the second or higher order model derivatives is equivalent to the application of nonlinear interpolation methods. In particular, the use of the second-order derivative operator corresponds to the cubic spline interpolation. If the slowness variations in the earth are linear, the first-order smoothness criterion should be sufficient

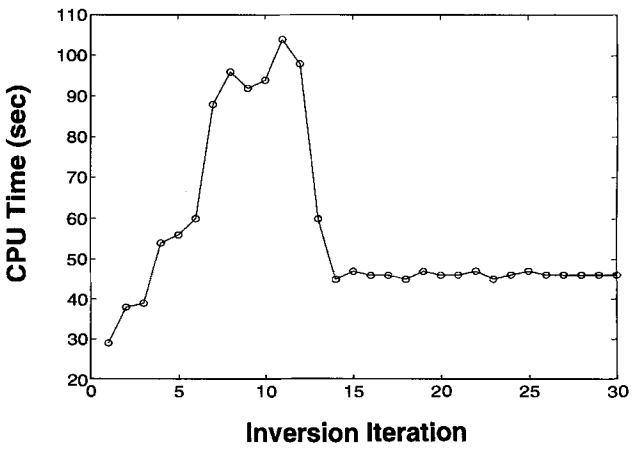

FIG. 7. CPU run times for each inversion iteration. Note they are not equal because our forward traveltime calculation depends on model heterogeneity and because inversion of the apparent slowness data starts at the seventh iteration.

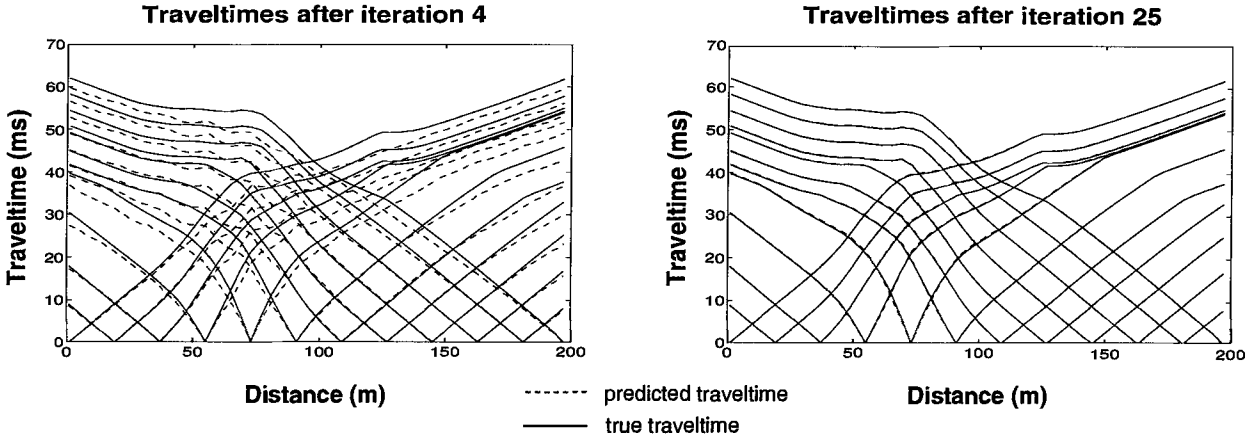

FIG. 6. Traveltime data fit at iteration 4 and iteration 25. It demonstrates that tomography tends to fit the near-shot traveltimes more accurately in the early inversion stage, and fits the long-offset traveltimes later in the inversion.

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

for regularization. However, in the real earth, we often encounter nonlinear variations in slowness. Therefore, nonlinear interpolation schemes must be applied.

From these experiments, one can also observe that the resolution increases with the solution uncertainty (see Figure 4c). Applying higher order smoothness criteria tends to cause the slowness (or velocity) to oscillate wildly near the structure interfaces where the deep rays pass. To reduce oscillation for the fourth-order smoothing operator, we actually had to force the highest velocity in the model to be less than 4800 m/s, which was estimated using the GRM method (the highest true velocity is 4500 m/s). Similarly to the performance of spline interpolation/extrapolation (Press et al., 1992), we found that minimizing the second- or third-order model derivatives is more stable and has less possibility of nonphysical oscillations.

MEASURING NONLINEAR UNCERTAINTY

A complete solution to any inverse problem must address issues of uncertainty. Two important approaches in geophysics are the Backus-Gilbert method (Backus and Gilbert, 1968, 1970), measuring uncertainty in terms of spread, and Bayesian inference (Tarantola, 1987), characterizing uncertainty completely through the posterior covariance operator. We apply Monte Carlo methods to measure variance and resolution from the posterior model covariance associated with our regularized inverse problem. Monte Carlo methods require inverting data vectors with random errors starting from models with random perturbation for many realizations until the posterior model covariance can be well approximated. We shall compare the performance of nonlinear and linearized Monte Carlo methods in dealing with the refraction traveltime problem.

Unlike in the waveform problems, noise influences traveltime data implicitly. Although numerical algorithms can be applied to pick traveltimes automatically, the user often has to determine where the first-break energy starts on the basis of the traveltime moveout. Errors in the refraction traveltime data are not arbitrarily "random" and generally do not have a zero mean for the "true traveltimes" from a statistical point of view. From our experiences with numerous field experiments, we found it often difficult to determine an accurate zero shooting time even with a modern recording system for the large crustal imaging projects or an electronic trigger system for the small-scale surveys. For instance, in the latter case, a constant trigger threshold under different near-surface geological conditions may lead to different trigger delays. Therefore, we define two types of errors: (1) common-shot error, a random shift for traveltimes at all the receivers from the same shot, and (2) a traveltime-gradient error, which is caused by the traveltime-picking uncertainty.

We further demonstrate the error pattern in the refraction traveltimes with synthetics. Figure 8 displays noise-free synthetic traveltimes (solid curve) from the numerical experiments that we described, and "picked" traveltimes with some assumed error patterns (dash curves). It seems unlikely that one would pick such rough traveltimes for inversion as those shown in Figure 8a. Accordingly, assuming Gaussian errors for the absolute traveltimes is not realistic, and inclusion of additional random shooting-time errors is not realistic either (see Figure 8b). In fact, in most cases, one would pick traveltimes by identifying the moveout of the first-break waves rather than picking a time solely based on a single trace. This leads to the possible picked traveltime curves that one may determine as shown in Figure 8c, which contains common-shot error and traveltime-gradient error. Therefore, for uncertainty analysis with the Monte Carlo methods, we need to construct two random vectors for data errors and one random vector for model perturbation with uncorrelated normal Gaussian distribution:

$$\hat{d}_k = \hat{d} + e_k^{(1)}, \qquad e_k^{(1)} \sim N(0, \sigma_1^2), \tag{14}$$

$$d_k = (d + e_k^{(2)}) + \hat{d}_k \Delta x, \qquad e_k^{(2)} \sim N(0, \sigma_2^2), \tag{15}$$

$$m_k = m_\infty + e_k^{(3)}, \qquad e_k^{(3)} \sim N(0, \sigma_3^2), \tag{16}$$

$$k = 1, 2, \ldots, K,$$

where d is a data vector and $\hat{d}$ is a data-gradient vector, $e_k^{(1)}$ is the traveltime-gradient error, $\hat{d}_k$ is the data-gradient vector perturbed by error $e_k^{(1)}$, $e_k^{(2)}$ is the shooting-time error, d_k is a data vector perturbed by both the shooting-time error and the traveltime-gradient error, Δx is the receiver spacing for which we consider the traveltime gradient, m_∞ is the solution of our nonlinear tomography inversion, and m_k is a model vector perturbed by error $e_k^{(3)}$. Because of our objective function

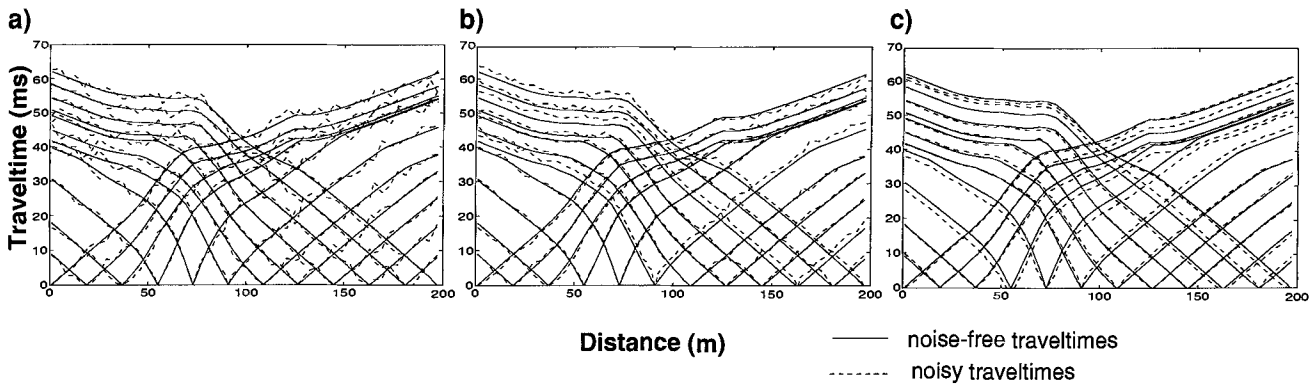

FIG. 8. Noise-free synthetic traveltimes and "picked" traveltimes with an assumed noise pattern. (a) Adding random Gaussian noise to the absolute traveltimes. (b) Adding random Gaussian noise to the absolute traveltimes and to each shot record. (c) Adding random Gaussian noise to the traveltime gradients and also adding random Gaussian noise to each shot record. Pattern (c) is realistic for the refraction problem.

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

definition, we can directly take the rms apparent-slowness misfit to be the variance of the traveltime-gradient error σ_1^2, and use the rms average-slowness misfit for the variance of the model perturbation σ_3^2. The variance of shooting times can be estimated using the near-shot traveltimes or simply half of the period of the refracted wave to account for the variance of trigger responses.

After solving K realizations, we compute the posterior covariance matrix (Matarese, 1993),

$$C_m = \frac{1}{K}\sum_{k=1}^{K}(\hat{m}_k - m_\infty)(\hat{m}_k - m_\infty)^T, \qquad (17)$$

where $\hat{m}_k$ is the solution of the kth realization. In practice, we estimate a convergence criterion S_0 for the objective function and only accept those realizations with the value of the objective function less than S_0.

We measure variance and resolution of our regularized nonlinear solution from the matrix C_m. The square roots of the diagonal elements in C_m can be interpreted as "error bars" describing the uncertainties on the posterior values of the model parameters. More importantly, model correlation tells how well each point is resolved (Tarantola, 1987). For a point i, its correlation with the entire model is given by

$$\rho_i(n) = \frac{C_m(i,n)}{(C_m(i,i))^{1/2}(C_m(n,n))^{1/2}}, \quad n = 1, 2, \ldots N, \qquad (18)$$

where N is the total number of model parameters. If the posterior correlation is close to zero, the posterior uncertainties are uncorrelated. If the correlation is close to +1, the uncertainties are highly correlated, indicating that the parameters have not been independently resolved by the data set, and are only linear combination of some values. If the correlation is close to −1, the uncertainties are highly anticorrelated.

For the numerical experiment that we performed, we show uncertainty measurements using a nonlinear Monte Carlo approach (250 realizations), that is allowing the raypath to be perturbed in each realization (see Figure 4d). The solution is obtained by minimizing the misfit of the average-slowness and apparent-slowness data, and the inversion is regularized by the third-order Tikhonov method. The results show that common-shot error primarily influences the near-surface structure, particularly near the shot locations. The traveltime-gradient error mainly produces uncertainty along the deep rays. Model correlation results from the nonlinear realizations show that two particular points that we select are fairly well resolved.

It is interesting to observe that a better resolution is shown for the deep point right above the graben bottom where rays actually do not pass. This occurs because nonlinear traveltime tomography would not allow the velocities above the graben refractor to be arbitrary. If the velocities are too high, the deep rays would move up and reduce the velocities. If they are too low after the random model perturbation, regularized inversion would remove the low-velocity features which are not required for fitting the data. Therefore, areas where rays do not pass may not be resolved poorly. We also performed linearized Monte Carlo inversions to measure uncertainty by fixing the raypaths for each realization (see Figure 4e), which are much faster than the nonlinear approach. For the same model point in the shallow low-velocity zone, applying linearized Monte Carlo method yields a resolution estimate similar to the result obtained by using the nonlinear Monte Carlo inversions. This is because many rays pass through the area, and the fixed raypaths are sufficient to constrain the whole area in the model. On the other hand, linearized Monte Carlo inversions fail to delineate the resolution of the second model point where rays do not pass. Since insufficient ray coverage often occurs in the refraction traveltime problem, applying nonlinear Monte Carlo inversions is a necessary step for measuring uncertainty.

INVERTING FIELD DATA

Our nonlinear refraction traveltime tomography method has been applied to several field data at different geometry scales. We shall demonstrate its application to a small-scale survey at a coastal site near Boston, Massachusetts. The goal of the survey was to locate those areas where bedrock is deep so that construction of a new storm-drainage system may proceed without costly blasting. The environment at the working site was quite unusual because the survey area was covered by sea water during the high-tide period, and exposed for only 1–2 hours during low tide each day. Further details of the project are reported by Kutrubes et al. (1996).

Twenty-four geophones with 10-ft (3.05-m) spacing were used. Surveys along two adjacent short lines (each 230 ft or 70.1 m) were conducted with 12 shots for each line. Figure 9a shows the seismic waveforms recorded from a forward shot and a reverse shot on line 1. Data from both shots show delayed first arrivals between receivers 7 and 14, although the topography along line 1 is flat. Moreover, the amplitudes of these delayed first arrivals are relatively small. This evidence suggests that a low-velocity zone with strong seismic attenuation exists beneath receivers 7 to 14. It became obvious when we placed sources (hammer and Betsy Seisgun) at locations between receivers 7 and 14, the seismic energy was observed to be significantly attenuated across all of the receivers. Although survey line 2 is adjacent to line 1, a different structural influence was found. Figure 9b displays the forward- and reverse-shot records on line 2. They show a relatively high-velocity anomaly in the shallow structure. In fact, weathered sandstone is found to outcrop at the surface in the central area.

To map bedrock topography precisely, we conducted tomography studies with a model consisting of 250×100 cells for each survey area (71.0×30.5 m). The results and predicted raypaths are presented in Figures 10a and 10b. The two tomograms clearly are quite consistent in terms of velocity range, though they are reconstructed independently. They both show three types of media: saturated sand and gravel, weathered sandstone, and sedimentary rock. The shallow low-velocity zone in the line-1 profile was found to be caused by organic deposits verified by drilling. To understand the uncertainty of our numerical solutions, we apply a nonlinear Monte Carlo method to estimate the posterior covariance. Figure 10c shows the posterior correlation for one particular point in each profile. The point in the first profile is selected right on the bedrock that we resolved. The solutions all show small variance with a similar correlation size. Because the geological evidence on the surface is fairly consistent with the near-surface image as shown in the line-2 tomogram, we should have the same good confidence for the bedrock image on line 1.

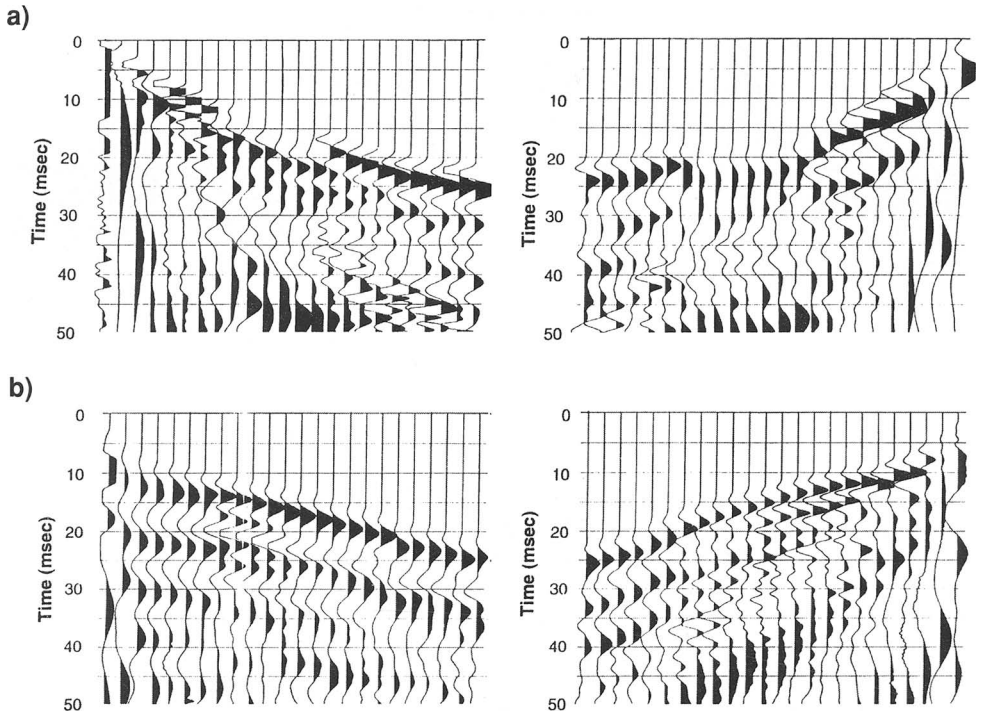

FIG. 9. Waveform data from forward and reverse shots along line 1 (a) and forward and reverse shots along line 2 (b). Note the arrival-time influences caused by a shallow low-velocity zone between traces 7 and 14 along line 1, and due to a shallow high-velocity zone along line 2.

Figure 11 shows the traveltime data and calculated traveltimes corresponding to the final solutions. Apparently, the line-1 data are not fit as well as the line-2 data, primarily because of the large traveltime uncertainty over the low-velocity zone. Accurate first-break times are difficult to pick from the attenuated signals. However, by inverting average slownesses and apparent slownesses and applying Tikhonov regularization, our tomography attempts to match traveltimes with meaningful data correlation and model correlation. Our tomography results were presented one day after the field work. Drilling from two nearby locations confirmed that our estimates of the bedrock depths were accurate (Kutrubes et al., 1996).

CONCLUSIONS

In this study, we develop a nonlinear refraction traveltime tomography method. This development includes three contributions: the forward ray-tracing method, the inversion approach that fits traveltime curves, and a nonlinear method for uncertainty analysis. These three issues in the tomography problem are of equal importance. A combination of these approaches presents a high-performance tomography product. The tomography method proved effective in both a numerical experiment and a real case.

ACKNOWLEDGMENTS

We are grateful to Doria Kutrubes for her help in the field operation. Comments and suggestions from the following people helped improve this work: Joe Matarese (MIT), Matthijs Haartsen (MIT), Chuck Doll (MIT), Dale Morgan (MIT), Roger Turpening (MIT), Bill Rodi (MIT), Huawei Zhou (University of Houston), Joe Stefani (Chevron), Bob Langan (Chevron), Doria Kutrubes (Hager GeoScience), and Chantal Chauvelier (Texas A&M). We also thank reviewers Carlos A. Cunha Filho, Tamas Nemeth, and Colin Zelt. During this project, J. Zhang was supported by a fellowship from the U.S. Environmental Protection Agency.

REFERENCES

Ammon, C. J., and Vidale, J. E., 1993, Tomography without rays: Bull. Seis. Soc. Am., **83**, 509–528.
Backus, G., and Gilbert, F., 1968, The resolving power of gross earth data: Geophys. J. Roy. Astr. Soc., **16**, 169–205.
——— 1970, Uniqueness in the inversion of inaccurate gross earth data: Phil. Trans. Roy. Soc. Lond., **A266**, 123–192.
Cai, W., and Qin, F., 1994, Three-dimensional refraction imaging: 63rd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstract, 629–632.
Delprat-Jannaud, F., and Lailly, P., 1993, Ill-posed and well-posed formulations of the reflection travel time tomography problem: J. Geophys. Res., **98**, 6589–6605.
Fischer, R., and Less, J. M., 1993, Shortest path ray tracing with sparse graphs: Geophysics, **58**, 987–996.
Hole, J. A., and Zelt, B. C., 1995, 3-D finite-difference reflection traveltimes: Geophys. J. Internat., **121**, 427–434.
Klimes, L., and Kvasnicka, M., 1993, 3-D network ray tracing: Geophys. J. Internat., **116**, 726–738.
Kutrubes, D. L., Zhang, J., and Hager, J., 1996, Comparison of conventional processing techniques and nonlinear refraction traveltime tomography for surveys at eastern Massachusetts coastal site: Proc., SAGEEP, Environ. Eng. Geophys. Soc., 135–138.
Mackie, R. L., and Madden, T. R., 1993, Three-dimensional magnetotelluric inversion using conjugate gradients: Geophys. J. Internat. **115**, 215–229.
Matarese, J. R., 1993, Nonlinear traveltime tomography: Ph.D. thesis, Mass. Inst. Tech.
McMechan, G. A., Harris, J. M., and Anderson, L. M., 1987, Cross-hole tomography for strongly variable media with applications to scale model data: Bull. Seis. Soc. Am., **77**, 1945–1960.

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

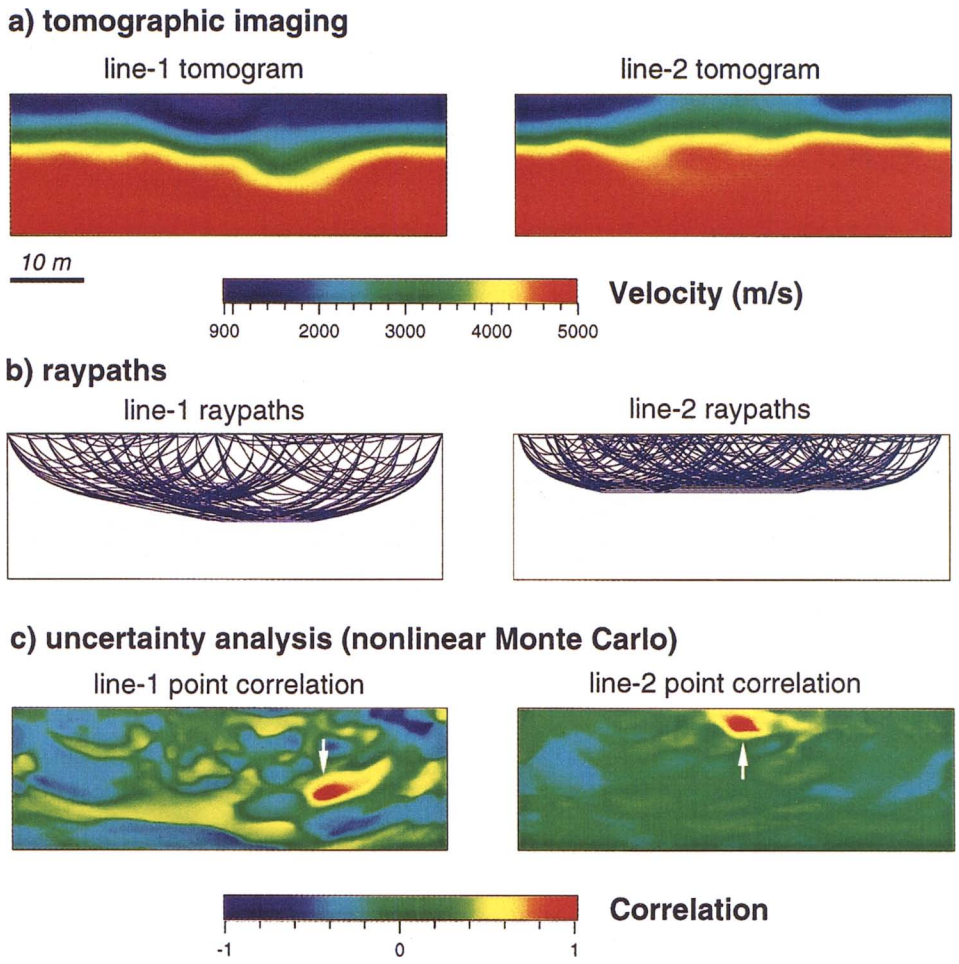

FIG. 10. (a) Tomographic images for the line-1 and line-2 profiles. Note the shallow low-velocity zone on line 1 and the intermediate-velocity zone outcropping at the surface on line 2. (b) Raypaths corresponding to the above models. (c) Uncertainty analysis using nonlinear Monte Carlo inversions. The posterior model correlations at two selected points are small, indicating that they are well resolved.

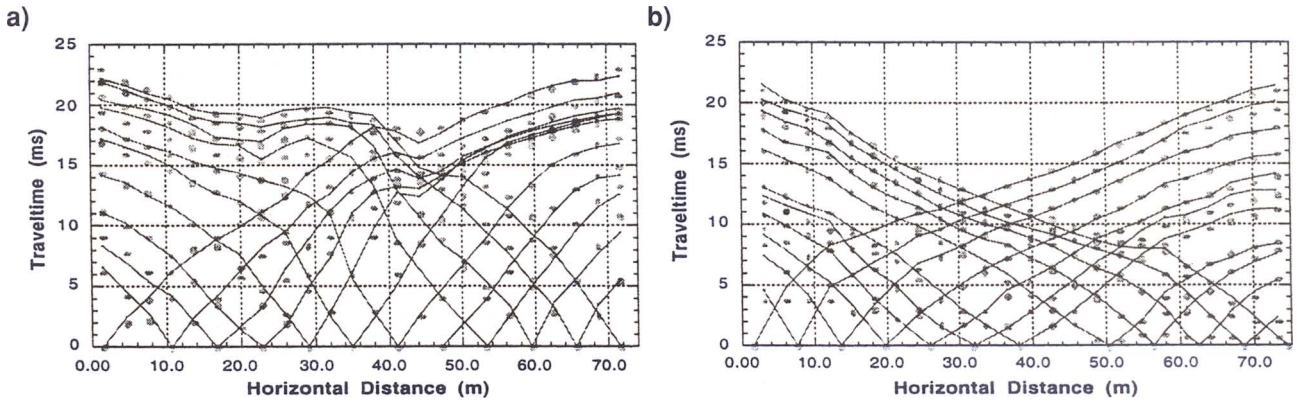

FIG. 11. Field traveltime data (curves) from line 1 (a) and line 2 (b) and the calculated traveltimes (gray dots) for the resolved models shown in Figure 10a.

Morgan, F. D. O., 1981, Electronics of sufide minerals: implications for induced polarization: Ph.D. thesis, Mass. Inst. Tech.
Moser, T. J., 1989, Efficient seismic ray tracing using graph theory: 59th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstract, 1106–1031.
——— 1991, Shortest path calculation of seismic rays: Geophysics, **56**, 59–67.
Nakanishi, I., and Yamaguchi, K., 1986, A numerical experiment on nonlinear image reconstruction from first-arrival times for two-dimensional island arc structure: J. Phys. Earth, **34**, 195–201.
Nolet, G., 1987, Waveform tomography, *in* G. Nolet, Ed., Seismic tomography: D. Reidel Publ. Co., 301–323.
Palmer, D., 1980, The generalized reciprocal method of seismic refraction interpretation: Soc. Expl. Geophys.
Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992, Numerical recipes in Fortran: the art of scientific computing, 2nd edition: Cambridge Univ. Press.
Qin, F., Luo, Y., Olsen, K. B., Cai, W., and Schuster, G. T., 1992, Finite-difference solution of the eikonal equation along expanding wavefronts: Geophysics, **57**, 478–487.
Rodi, W. L., 1989, Regularization and Backus-Gilbert estimation in nonlinear inverse problems: application to magnetotellurics and surface waves: Ph.D. thesis, Pennsylvania State Univ.
Saito, H., 1989, Traveltimes and raypaths of first arrival seismic waves: Computation method based on Huygens' Principle: 59th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstract, 244–247.
——— 1990, 3-D ray-tracing method based on Huygens' principle: 60th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstract, 1024–1027.
Scales, J. A., 1987, Tomographic inversion via the conjugate gradient method: Geophysics, **52**, 179–185.
Scales, J. A., Docherty, P., and Gersztenkorn, A., 1990, Regularisation of nonlinear inverse problems: imaging the near-surface weathering layer: Inverse Problems, **6**, 115–131.
Stefani, J. P., 1995, Turning-ray tomography: Geophysics, **60**, 1917–1929.
Tarantola, A., 1987, Inverse problem theory: Elsevier.
Tikhonov, A. N., and Arsenin, V. Y., 1977. Solutions of ill-posed problems: W. H. Winston and Sons.
Vidale, J. E., 1988, Finite-difference traveltime calculation: *Bull. Seis. Soc. Am.*, **78**, 2062–2076.
Vinje, V., Iversen, E., and Gjoystdal, H., 1993, Traveltime and amplitude estimation using wavefront construction: Geophysics, **58**, 1157–1166.
White, D. J., 1989, Two-dimensional seismic refraction tomography: Geophy. J., **97**, 223–245.
Zhang, J., Mackie, R. L., and Madden, T. R., 1995, 3-D resistivity forward modeling and inversion using conjugate gradients: Geophysics, **60**, 1313–25.
Zhang, J., Rodi, W., Mackie, R. L., and Shi, W., 1996, Regularization in 3-D dc resistivity tomography: Proc. SAGEEP, Environ. Eng. Geophy. Soc., 130–135.
Zhu, X., and McMechan, G. A., 1989, Estimation of two-dimensional seismic compressional-wave velocity distribution by iterative tomographic imaging: Internat. J. Imag. Sys. Tech., **1**, 13–17.

Downloaded 02/22/18 to 160.91.216.167. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/