Nonlinear 3-D traveltime inversion of crosshole data with an application in the area of the Middle Ural Mountains

4 authors:

Pantelis Soupios
King Fahd University of Petroleum and Minerals
169 PUBLICATIONS **1,883** CITATIONS

SEE PROFILE

Christopher Juhlin
Uppsala University
395 PUBLICATIONS **5,247** CITATIONS

SEE PROFILE

C. B. Papazachos
Aristotle University of Thessaloniki
223 PUBLICATIONS **5,089** CITATIONS

SEE PROFILE

G.N. Tsokas
Aristotle University of Thessaloniki
152 PUBLICATIONS **1,992** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   H2020 RIA Smart Exploration View project

Project   Southern Aegean Subduction View project

Nonlinear 3-D traveltime inversion of crosshole data with an application in the area of the Middle Ural Mountains

Pantelis M. Soupios*, Constantinos B. Papazachos*, Christopher Juhlin‡, and Gregory N. Tsokas*

ABSTRACT

This paper deals with the problem of nonlinear seismic velocity estimation from first-arrival traveltimes obtained from crosshole and downhole experiments in three dimensions. A standard tomographic procedure is applied, based on the representation of the crosshole area into a number of cells which have an initial slowness assigned. For the forward modeling, the raypath matrix is computed using the revisited ray bending method, supplemented by an approximate computation of the first Fresnel zone at each point of the ray, hence using physical and not only mathematical rays. Since 3-D ray tracing is incorporated, the inversion technique is nonlinear. Velocity images are obtained by a constrained least-squares inversion scheme using both "damping" and "smooth-

ing" factors. The appropriate choice of these factors is defined by the use of appropriate criteria such as the L-curve.

The tomographic approach is improved by incorporating a priori information about the media to be imaged into our inversion scheme. This improvement in imaging is achieved by projecting a desirable solution onto the null space of the inversion, and including this null-space contribution with the standard non–null-space inversion solution. The efficiency of the inversion scheme is tested through a series of tests with synthetic data. Moreover, application in the area of the Ural Mountains using real data demonstrates that the proposed technique produces more realistic velocity models than those obtained by other standard approaches.

INTRODUCTION

The main target of seismic traveltime inversion in 2-D and 3-D heterogeneous media is to infer the slowness distribution of a medium, given a set of observed first-arrival traveltimes between sources and receivers of known locations within the medium. This problem is typical for crosshole seismic tomographic imaging of a 2-D region between vertical boreholes (Harris et al., 1995). Usually, the crosshole area is represented by a Cartesian grid of nodes with values of slowness assigned to the grid points. If $\mathbf{s}$ is the model slowness n-vector such that $\mathbf{s}^T = (s_1, \ldots, s_n)$ with s_j being the slowness of the jth node, then

$$\mathbf{Ms} = \mathbf{t} \qquad (1)$$

where $\mathbf{M}$ is an $m \times n$ matrix whose matrix element $\ell_{i,j}$ corresponds to the distance of the ith ray associated with the jth node, and $\mathbf{t}$ is the traveltime m-vector such that $\mathbf{t}^T = (t_1, \ldots, t_m)$ where t_i is the traveltime along the ith raypath.

To find $\mathbf{M}$, the forward problem (ray tracing) needs to be solved. In the present study, we examine possible improvements of the standard crosshole-downhole tomography using ray theory. Specifically, we explore the significance of 3-D ray tracing for tomographic purposes. For the determination of initial rays, we use a fast, approximate ray tracing method known as the ray initializer (Thurber and Ellsworth, 1980), which is equivalent to solving a 1-D approximation to the actual 3-D problem. This method provides an initial estimate of the minimum-time ray paths in heterogeneous media for the final ray tracing with a revised ray-bending technique (Moser et al., 1992).

Knowing, that the properties of the ray are influenced not only by the structure along the ray, but also by the structure in some "vicinity" of the ray, we incorporate physical rays by estimating the width of first Fresnel volume, which is frequency dependent (Červený and Soares, 1992).

It is generally assumed that the inversion problem [equation (1)] is overdetermined and, hence, can be solved in a least-squares sense by minimizing the data misfit. Since there is frequently a certain degree of nonuniqueness or ill-conditioning in the problem, the standard approach is to construct a "least-structure" solution (Franklin, 1970; Tarantola and Nercessian, 1984; Constable et al., 1987) by considering additional constraints and usually minimizing the model and/or one of the model derivative norms. The inverse problem is solved by either the Levenberg-Marquardt (L-M) method (also known as the damped least squares method) (Levenberg, 1944; Marquardt, 1963), or by using both a damping factor and a roughness term in the form of a second derivative smoothing filter to the slowness model. To choose the proper values of damping and smoothing factors, we also used two of the most well-known regularization tools, the Picard condition and the L-curve (Hansen, 1994), which generally provide a good estimation of the regularization parameters in combination with the standard maximum likelihood approach (Franklin, 1970). Moreover, we also use conservative filtering (Deal and Nolet, 1996), which allows a tomographic image to be "manipulated" without significantly changing the data fit. Hence, by restricting the changes made to the solution to lie in the nullspace, we can determine which a priori constraints or expectations are permitted by the data.

To show the applicability of the proposed method, we present two synthetic examples and the application of our method to a real VSP data set from the Middle Urals area, producing a velocity model which is in general agreement with previous studies (Juhlin et al., 1997).

MODEL CONFIGURATION AND FORWARD PROBLEM

In local inversion studies, a flat layered earth model is commonly used (Aki and Lee, 1976). A similar assumption was also adopted here because of the small scale of our study area and the convenience of the Cartesian coordinate system, which significantly speeds up all computations, especially those concerning 3-D ray tracing. Our model is composed of a 3-D grid of blocks of different sizes, with values of slowness assigned at the grid points. The slowness at each point which is not a slowness node is calculated by trilinear interpolation of the slowness values of the nodes of the grid rectangle where this point lies.

Given a set of observed traveltimes, $t_1, \ldots, t_m$, from m source-receiver pairs in a medium of slowness $s(\mathbf{x})$ and neglecting observational errors, we can write,

$$\int_{P_i} s(\mathbf{x}) d\ell^{P_i} = t_i, \qquad i = 1, \ldots, m. \qquad (2)$$

where P_i is the Fermat ray path connecting the ith source-receiver pair. Given a block model of slowness, let $\ell_{i,j}$ be the length of the ith ray path through the jth cell. If the model is divided in n cells, Equation (2) can then be rewritten as

$$\sum_{j=1}^{n} \ell_{ij} s_j = t_i, \qquad i = 1, \ldots, m. \qquad (3)$$

Note that for any given i, the raypath lengths ℓ_{ij} are zero for most cells j, as a given ray path will in general intersect only a few of the cells in the model.

We can rewrite equation (3) in matrix notation by defining the $\mathbf{s}$ and $\mathbf{t}$ as column vectors containing s_j and t_i and $\mathbf{M}$ as the matrix containing the ℓ_{ij} values in the form of equation (1), which is the basic equation of forward modeling for ray theory analysis. In this paper, we examine the nonlinear inversion problem, where only $\mathbf{t}$ is available and we attempt to infer $\mathbf{s}$ and incidently $\mathbf{M}$. In this problem, the dependence of ray paths on the slowness distribution strongly influences the nonlinearity of the problem and the design of the inversion algorithm. Thus, an iterative algorithm is generally needed to find an approximate solution $\mathbf{s_b}$.

The calculation of the ray-path matrix is just the solution of the forward problem for $\mathbf{s_b}$. This part should not introduce any instability, since it can be performed, theoretically, as accurately as desired. On the other hand, the calculation of the model correction is known as the linear inversion step, imbedded in a nonlinear algorithm.

Ray tracing

The most computationaly "expensive" and important step in traveltime inversion is the forward modeling step associated with the ray tracing technique used. In the present study, we adopted a revised bending algorithm (Moser et al., 1992), which has been shown to have superior converging properties.

In an effort to alleviate problems such as convergence and execution speed associated with the bending method, we incorporated the approximate ray-tracing scheme of Thurber and Ellsworth (1980). In this approach, an equivalent 1-D structure is calculated from the 3-D model between each source-receiver pair, as shown in Figure 1. The velocities at the grid points that fall within the box shown in Figure 1 are harmonically averaged, which is equivalent to arithmetically averaging the slowness. These average values are used as layer velocities for the 1-D model, where the direct and all possible refracted
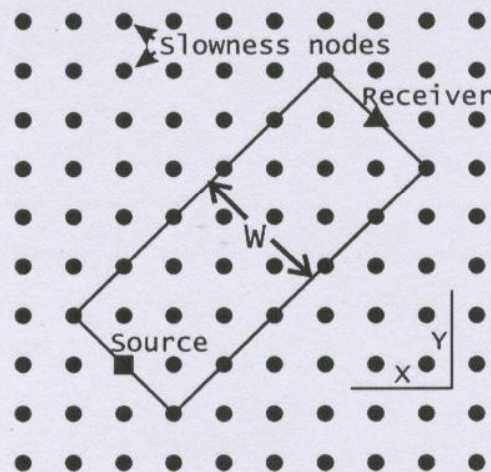

FIG. 1. Configuration of the initial approximate 1-D ray tracing used. A 1-D average velocity is constructed from the grid-point values lying within a rectangular volume, and standard 1-D rays are computed.

arrivals are considered. The width, W, of this box is given by

$$W = \frac{\Delta}{8} + GS \qquad (4)$$

where GS is the horizontal grid spacing. The direct ray is determined using a combination of the false position method and the secant method (Press et al., 1988), while refracted rays are handled using time terms (Officer, 1958).

The minimum traveltime path from this ray tracing routine is used as the initial raypath guess for our ray-bending tracing routine (Moser et al., 1992). In this technique, the ray is represented by a series of support points which describe the ray path with the use of Beta-splines (Pereyra, 1992). Using this configuration, the ray is split into a series of segments, each one of which is completely described by four succesive support points. The main advantage is that the interpolation between successive points of a ray is generally necessary to obtain sufficient precision in the quadrature, but it is also advantageous in terms of efficiency when Beta-splines are used. The travel time along the Beta-spline curve can be calculated with the trapezoidal rule:

$$T(\gamma) = \sum_{i=1}^{k} \sum_{j=1}^{m} \frac{1}{2}\left(\frac{1}{c_{i,j}} + \frac{1}{c_{i,j}-1}\right) \|Q_i(u_j) - Q_i(u_j - 1)\|, \qquad (5)$$

where $Q_i(u)$, $(i = 1, \ldots, k, 0 \leq u \leq 1)$ is a support point in the ith curve segment, m is the number of points $Q_i(u_j)$ to be integrated over on that segment, and c_{ij} is the seismic velocity at $Q_i(u_j)$. The usual and appropriate choice for u_j is $u_j = j/m$.

Finally, we defined the radius of Fresnel volume for each point (x) along the raypath from the source (a) to the receiver (b). Using the results of Červený and Soares (1992) for a homogeneous medium, we can define the following equation, which is a paraxial approximation of the exact Fresnel volume:

$$r = \sqrt{\frac{\left(\frac{\lambda}{2} + \ell_{ab}\right)^4 + \left(\ell_{xa}^2 - \ell_{xb}^2\right)^2 - 2\left(\frac{\lambda}{2} + \ell_{ab}\right)^2 \left(\ell_{xa}^2 + \ell_{xb}^2\right)}{4\left(\frac{\lambda}{2} + \ell_{ab}\right)^2}} \qquad (6)$$

where r is the radius of the Fresnel zone at point (x) along the ray, λ is the wavelength of the waveform, which depends on frequency and local velocity, ℓ_{ab} denotes the distance along the ray between source (a) and receiver (b) and ℓ_{xa}, ℓ_{xb} are the distances between the points $x - a$ and $x - b$. The effect of the first Fresnel zone in equation (1) merely consists of a simple algebraic manipulation of the elements of the raypath matrix, $\mathbf{M}$, by "distributing" the matrix elements corresponding to each ray point to its Fresnel volume. The wavelength, λ, is estimated at each point (x) using the velocity at this point and the dominant frequency of the seismic ray.

Equation (6) is an efficient approximation of the exact estimation of the width of the Fresnel volume. This is seen in Figure 2, where we test equation (6) against the example presented in the original paper of Červený and Soares (1992). In Figure 2a, the exact Fresnel volume is shown for a ray in a medium of constant velocity gradient ($v = 1 + 0.875z$). The ray

determined using our ray tracing and approximate Fresnel volume (Figure 2b) is almost identical. However, the application of the proposed equation (6) is much more simple for the calculation of the Fresnel volume, since we only need the raypath connecting the given source and receiver and the dominant frequency of our signal for the wavelength estimation.

INVERSE PROBLEM

The solution of equation (1), in the least-squares sense, is to find an s that minimizes the sum of squared differences between $\mathbf{t_b}$ (calculated traveltime) and $\mathbf{t}$ (observed traveltime). Hence,

$$\min \|\mathbf{t} - \mathbf{t_b}\|^2, \qquad (7)$$

In the present work, the solution of equation (1) was obtained by using a combination of "damping" and "smoothing" factor in the least-squares method. Both methods were implemented in this study by using singular value decomposition (SVD) (e.g., Pereyra, 1996) since the size of the studied tests allowed its practical use.

In most cases the derivative matrix has small singular values and is typically highly ill-conditioned. One approach in dealing with this is to use the Lagrangian multiplier (LM) method and augment the system in equation (1) by a set of additional constraints in the following manner:

$$\begin{bmatrix} \mathbf{M} \\ \varepsilon \mathbf{I} \end{bmatrix} \mathbf{s} = \begin{bmatrix} \mathbf{t} \\ \mathbf{0} \end{bmatrix} \qquad (8)$$

where $\varepsilon \mathbf{I}$ is an $m \times m$ identity matrix multiplied by the constant ε. This leads to a damped least-squares solution, which balances the size of the residual and the solution norm.
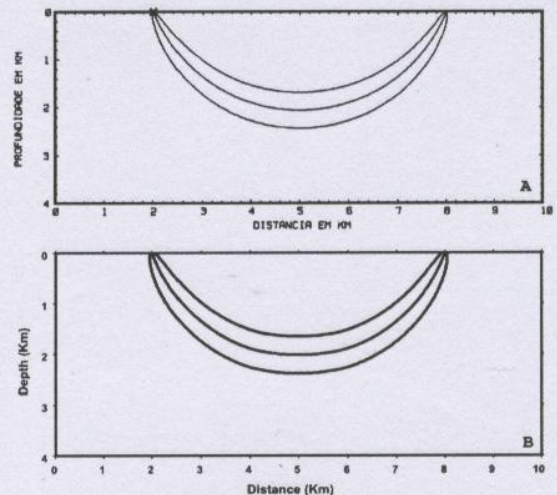

FIG. 2. Comparison between Cerveny's (Cerveny and Soares, 1992) and our approximation of the Fresnel volume using the same configuration, velocity model, and dominant frequency. (a) is an original copy from Cerveny's paper, and (b) is our Fresnel volume calculation.

Using equations (8), the solution can be written as

$$\delta s(\varepsilon) = (M^T M + \varepsilon^2 I)^{-1} M^T (t - t_b) \qquad (9)$$

where δs is the correction that should be applied to our pre-liminary solution. The controlling effect of ε on the inversion results has been discussed by various investigators (Jupp and Vozoff, 1975; Crosson, 1976; Hawley et al., 1981). The choice of ε depends upon the problem considered. An optimal ε, which should be as small as possible to achieve the maximum reso-lution and convergence rate yet large enough to stabilize the solution, also varies from iteration to iteration. If the obser-vational data errors are Gaussian with a variance of σ_t^2 and the model perturbations, δs, also follow a normal distribution with variance σ_s^2, then Franklin (1970) showed that ε^2 should be equal to σ_t^2 / σ_s^2.

If we assume that the slowness field for which we are invert-ing is a discrete version of a continuously varying slowness, we may want our solutions to have a certain measure of smooth-ness. In our case, smoothness constraints can be imposed by minimizing a measure of the roughness of the model, using a second differential operator for a 2-D earth. In vector notation, this is just the Laplacian operator, ∇^2. We can then augment the system (1) by replacing εI with λL in equation (8), where L is the 2-D, finite difference Laplacian operator applied in our slowness model. In our algorithm, we applied the roughness term, L, as proposed in Sasaki (1992). Using this augmented system, we will minimize the following functional:

$$\|Ms - t\|^2 + \lambda^2 \|Ls\|^2. \qquad (10)$$

The parameter λ is used to adjust the relative weight of roughness versus misfit reduction as a tradeoff parameter. A discussion of the use of differential operators as roughening constraints is given by Constable et al. (1987). Finally, the so-lution is given by the following equation:

$$\delta s = (M^T M + \lambda^2 L^T L)^{-1} M^T (t - t_b), \qquad (11)$$

where $L^T L$ is the smoothing matrix.

Smoothing is often applied to reduce inversion artifacts with short correlation lengths. However, an a posteriori smoothing generally affects the data fit. For specific nonlinear filters this effect can be significant. Thus, we applied the method proposed in Deal and Nolet (1996), that allows us to filter an image a pos-teriori, while retaining the fit to the data. The filter is restricted to operate only on components of the solution that do not affect the data fit.

NUMERICAL TESTS

Numerical tests on synthetic data were undertaken to study the proposed algorithms. Results from such a synthetic cross-hole and crosshole-downhole test are presented. In these tests, we select a specified velocity distribution for the cross-hole area and solve the forward problem (ray tracing). Then, all the minimum traveltimes were calculated and inverted back to a velocity structure that is compared to the original one. Our synthetic data are error-free crosshole data, which means that the ideal solution should be recovered. Hence, comparison of the original velocity model and what is recovered by our al-gorithm permits us to analyse the efficiency of the inversion scheme and of the forward modeling. This way, we not only

study the impact of incomplete ray coverage in certain regions, but also observe how different velocity zones are mapped by our algorithm. For both tests described in this section, a large set of values (0–50) was used as the damping factor in order to evaluate the trade-off between resolution and error. From this analysis, a value of $\varepsilon = 30$ was selected for the final solution, since larger values resulted in strongly "damped" models. Also, a similar range of values (5–35) was used for the smoothness factor. We will present velocity models for the first and third iteration, applying both constraints (damping-smoothing).

In both tests, we use a 1-D background velocity model, where velocities increase linearly with depth according to the follow-ing formula:

$$V = 0.7 + 0.01z. \qquad (12)$$

Our 3-D grid of nodes had 11 nodes in the x direction, 3 in the y direction, and 11 in the z direction (depth), resulting in a total of 363 nodes. A 2-D velocity model was used to produce the synthetics traveltimes, which was identical to the 1-D model [equation (12)], except for layers 3, 4, and 5, which were set to 15% above the background velocity in the center of the model, and layers 7, 8, and 9, which were set to 10% below background velocity near the lower left part of our model (Figure 3).

Test 1. Crosswell model

For this test, we used two ideal boreholes (A and B) located 25 m apart, with 51 sources in borehole A and 51 geophones in borehole B (Figure 3). This configuration results in 2601 rays for the 363 unknown variables (grid velocities).
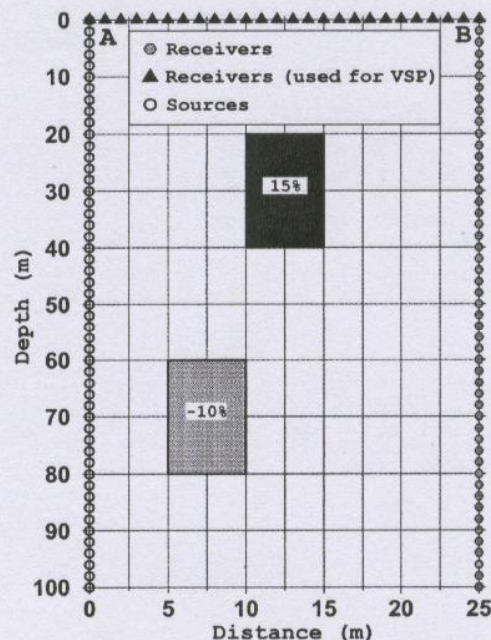

FIG. 3. The crosswell and VSP experimental geometry and the velocity distribution used for the synthetic tests.

Using the background model as the starting model, equation (3) was inverted. Figure 4 shows, in a x-z cross-section presentation, the retrieved velocity anomalies when 1-D ray tracing (first iteration) and 3-D ray tracing (third iteration) are employed.

From our velocity models, we can observe that both the high- and low-velocity anomalies are quite well recovered. It was quite interesting that in the area of interest (velocity anomalies), both the amplitude and the shape of the velocity anomalies are significantly improved by the nonlinear inversion. Some noise and artifacts in our velocity model are caused as a result of insufficient ray coverage in some regions, especially at the top and at the bottom part of the model. This is an important feature of the crosshole experiment where a small variation in velocity can lead to a deflection of the ray and the appearence of a shadow zone (low-velocity anomaly). Better tomographic images are produced when smoothing constraints are applied to the data.

Examination of the travel-time misfit (rms) for the first three iterations is shown in Figure 5. After the third iteration, we observed that we have almost no misfit reduction, implying that matrix **M** has no more information to provide and, therefore, we stopped the inversion. The initial rms misfit was 1.05 s and reduced to 0.14 s after one iteration (1-D ray tracing) and to 0.08 s after 3-D ray tracing. However, the amount of misfit reduction after each iteration is not accurate since this misfit is calculated from the linear approximation and not from the real nonlinear problem. Between succesive steps (iterations), we recalculate the derivative matrix **M** by performing 3-D ray tracing and estimated the misfit for the solution derived from the previous step. Hence, at these points, we can compare the linear and nonlinear misfits. Figure 5 shows that at the end

of the first iteration we are underestimating the linear misfit by at least 60% because we are using a derivative matrix **M**, which was calculated for the 1-D model. Thus, as the model is updated and deviates from the starting solution, the search directions determined by the inversion of **M** are inadequate,
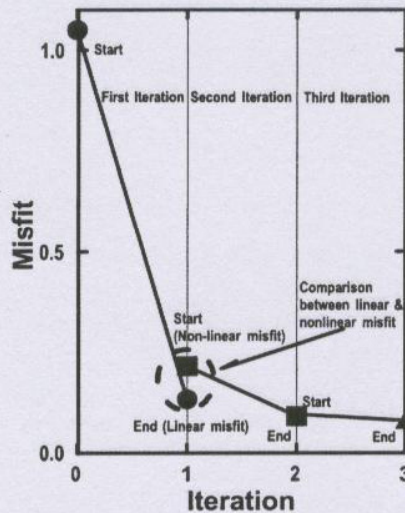

FIG. 5. Plot of the traveltime misfit versus each iteration of the inversion. The transition between the first and second iteration is marked by the dashed circle. At the end of the first iteration, the linear misfit underestimates the true (nonlinear) misfit by approximately 60%.
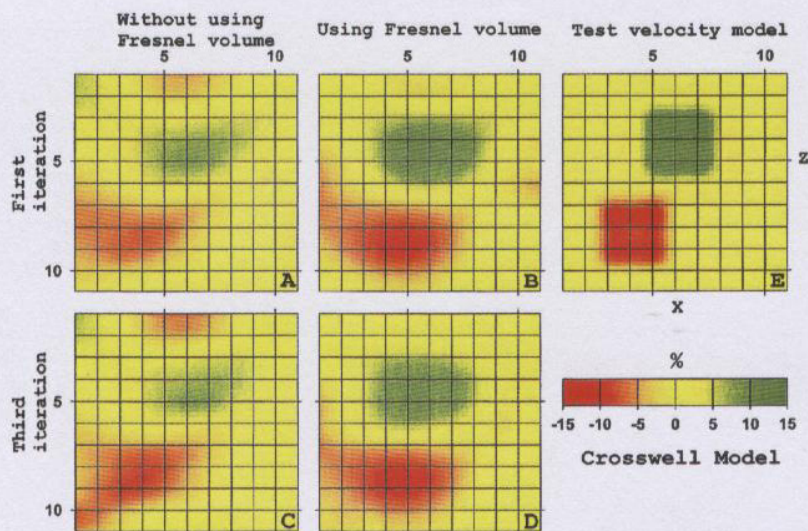

FIG. 4. Vertical cross-section of the study area resulting from the inversion of synthetic crosswell data. Solutions have been determing using both damping ($\varepsilon = 5.$) and smoothing ($\lambda = 20.$) (a, b) 1-D ray tracing (first iteration) is used with and without Fresnel volume calculation. (c, d) 3-D ray tracing (third iteration) is also given with and without Fresnel calculation. The determined velocity anomalies are significantly improved when 3-D ray tracing is used. The reference "real" velocity model is also shown (e).

and the linear approximation residuals give us a false image of convergence of the nonlinear problem. This behavior, also observed in other cases (Sambridge, 1990; Papazachos and Nolet, 1997), verifies the original suggestion of Tarantola (1987) that, when solving a nonlinear problem with a local linear approximation, all resolution and error estimates should incorporate the nonlinear residuals and not their linearized approximation.

Test 2. Crosswell + VSP model

For this test, we keep the same configuration for boreholes and sources but use 76 geophones both in borehole B and on the surface (Figure 3), resulting in 3876 rays and 363 unknown variables.

Figure 6 presents (in a cross-section) the tomographic images for the first and third iteration. The results show that both velocity anomalies are well reconstructed. Artifacts have been eliminated by applying apropriate constraints. The algorithm recovers the true model in this case much better than in the crosshole case (Figure 4), due to the better ray coverage. We observed that by using 3-D ray tracing, we recover quite well the amplitude of the ideal anomaly in our synthetic model, and also we avoided the introduction of bias near the boreholes, a typical characteristic of 1-D ray tracing (linear inversion).

APPLICATION TO REAL VSP DATA

Following the experiments on synthetic data, we also tested the proposed technique on data sets acquired in the area of SG-4 borehole in the Middle Ural Mountains, Russia. The SG4 borehole is located in the western part of the Tagil synform in the Middle Urals, about 15 km east of the Main Uralian fault. A large number of seismic experiments in the area show the upper crust to be highly complicated and reflective (Bliznetsov and Juhlin, 1994; Juhlin et al., 1995, 1997; Steer et al., 1995) due to the complex geological situation at the SG4 borehole. The island arc volcanics and their derivatives form a highly heterogeneous rock mass of varying composition where there are numerous interfaces which may cause velocity anomalies. In addition, there are at least two sets of fault zones of opposing dip, which may also produce velocity anomalies. We applied tomography in this complex region because we wanted to find out the velocity model for this area, and we also intented to test the algorithm in real noisy data, not only in synthetic models.

Two data sets were used for the inversion. The first data set used was recorded in March 1997 in the SG-4 borehole; the second one was recorded by a surface receiver array. A total of 15 shot points, following an approximate north-south linear configuration was used (Figure 7) from which high-resolution (1 ms) data were collected by three component geophones located for the first data set at depths between 1120 and 3920 m at 20 m intervals (Ayarza et al., 1998). For the second experiment, 6 of the 15 previous shots (SP1, 2, 3, 4, 6, and 7) were used and recorded by 48 surface channels (Figure 7). The horizontal geophone spacing was 50 m. Recordings at only 28 geophones were finally used due to the relatively poor signal quality. Topographic correction were applied in order to reduce all arrivals to the altitude of the head of the borehole. The final data set consisted of 1103 first breaks of P-waves hand picked from the VSP waveforms and 115 first breaks of P-waves from the surface data set.

The tomographic model consisted of 656 velocity nodes (16 nodes in south-north direction and 41 nodes in the vertical direction). The cells in the 2-D cross section in the borehole area were chosen to be squares 100 m on a side in order to get a quite accurate image of the study area. Using the surface data, we applied the plus/minus refraction method to retrieve
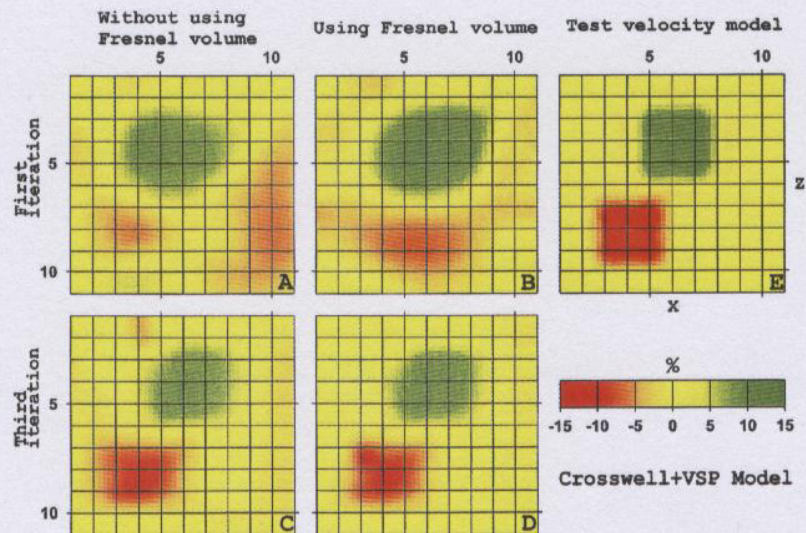

FIG. 6. Same as Figure 4 with both crosswell and VSP data inverted. In the final velocity model, both the amplitudes and the shape of the anomalies are very well resolved. A further improvement in the reconstruction of the anomalies is observed when using Fresnel volume calculations.

an initial guess of the structure and velocity model of the shallow layers. The results show a two-layer model with a contact dipping from south to north, with mean velocities 3200 and 4700 m/s, respectively (Figure 8). Using these velocities to constrain the first three node layers of the model, and a positive gradient with depth for the rest of grid nodes, we produced an initial smooth 1-D velocity model.

Since the experiment was not specifically designed for tomographic studies (subparallel rays) and the configuration control of the experiment was quite poor, a high smoothing factor was necessary. In order to choose the best values for regularization parameters, we used two different methods. The optimum damping factor, ε, was determined as follow. The variance of observational traveltime errors was of the order of 2 ms, while the mean velocity perturbation of the formations in the area is expected to be of the order of 1 km/s. Hence, according to Franklin (1970), the resulting damping factor, ε, should be of the order of 70. We also checked the effect of the smoothing factor by applying several values for λ. Finally, we concluded

that smoothing factors of the order of 10–500 gave reasonable velocity models. In order to define an "optimum" solution, we checked our estimations for the regularization parameters using an appropriate mathematical approach. For this reason, we used two criteria based on the L-curve and the Picard condition (Hansen, 1994). The first one presents the variation of the minimum norm regularized solution versus the corresponding residual norm. In this way, the L-curve clearly displays the compromise between minimization of these two quantities, which is the heart of any regularization problem. On the other hand, the Picard condition examines w_i, $|\mathbf{u}_i^{\mathrm{T}}\mathbf{b}|$ and $|\mathbf{u}_i^{\mathrm{T}}\mathbf{b}|/w_i$, where w_i and $\mathbf{u}_i$ are the singular value and the left singular vector of the raypath matrix, respectively (Hansen, 1994). In almost all applications, the data vector $\mathbf{b}$ is contaminated by errors (measurements, approximation, and rounding errors). Under this assumption, the first two quantities can reach the order of the numerical precision of the estimation. If the solution coefficient $|\mathbf{u}_i^{\mathrm{T}}\mathbf{b}|/w_i$ starts to increase (and eventually "explode") as we move to smallest w_i, the eigenvalue before
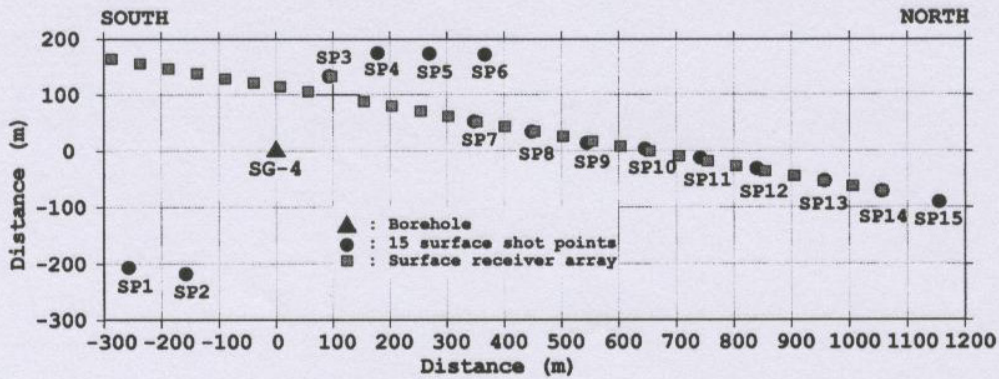

FIG. 7. Sketch of the VSP97 experiment. Fifteen shot points (SP) and 141 receivers inside the SG4 borehole (triangle) and 28 surface receivers are used.
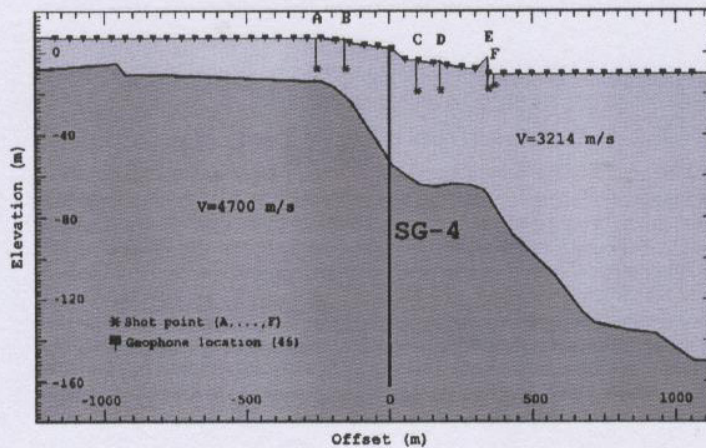

FIG. 8. Seismic velocity image as obtained from the traveltime data recorded at surface receivers, obtained by standard refraction techniques.

this "explosion" is the smallest eigenvalue that we are able to reconstruct. Of course, this also gives us an idea of the acceptable eigenvalue threshold. Generally, the purpose of both methods is to damp or filter out the contributions to the solution corresponding to the small generalized singular values. Finally, using the L-curve, we can define the best damping factor for a given smoothing parameter (Figure 9), which is also in agreement with the Picard's condition plot (Figure 10). Using these criteria, we chose the best pair of these regularization parameters. The values of $\varepsilon = 70$ (damping) was found to correspond (Figure 10) to a smoothing factor of $\lambda = 100$. These values were selected for the final solution.

Estimation of the traveltime misfit reduction for the three iterations is shown in Figure 11. After the third iteration, we observed that we have almost no misfit reduction and, therefore, we decided to stop the inverstion. The initial rms misfit was 19.1313 ms and reduced to 3.1187 ms at the end of the first iteration (1-D ray tracing) and to 2.9471 ms after 3-D ray tracing. As noted previously, the amount of misfit reduction after each iteration is not accurate because this misfit is calculated from the linear approximation and not from the real nonlinear problem. Figure 11 shows that at the end of the first iteration, we are understimating the true (nonlinear) misfit by at least 200%. This "jump" is caused by the fact that our velocity model is very complex, which means that 1-D ray tracing is inadequate to describe correctly the traveltime field.

Computationally, on a Silicon Graphics (O2-MIPS R10000 150 MHz) workstation, the model used in our case (656 nodes) for 1500 ray paths required an average CPU time of 2 minutes for the initial ray path calculations when a 1-D starting velocity model was used and less than 8 minutes when 3-D ray tracing was incorporated.
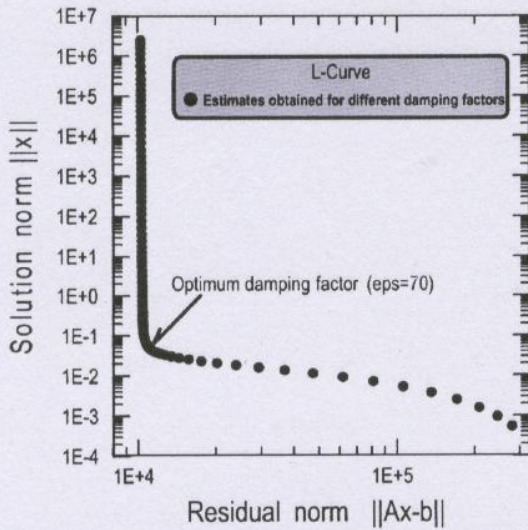

FIG. 10. Picard condition: a linear-log plot of index number versus w_i, $|\mathbf{u}_i^T\mathbf{b}|$, and $|\mathbf{u}_i^T\mathbf{b}|/w_i$. The value of $w_i = 70$ is the smallest singular value that we were able to reconstruct. This value is also the damping factor chosen for a smoothing factor equal to 100.

FIG. 9. L-curve: a log-log plot of residual norm versus solution norm which has the characteristic L-shaped appeerence. The corner of the curve gives an approximation of the optimal regularization parameter ($\varepsilon = 70$ for smoothing factor equal to 100).
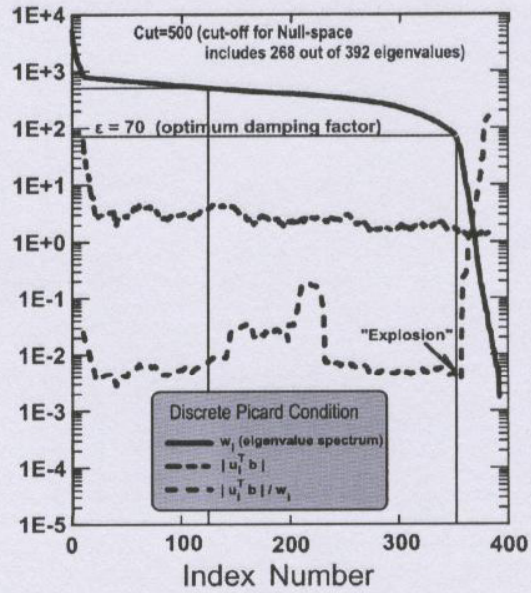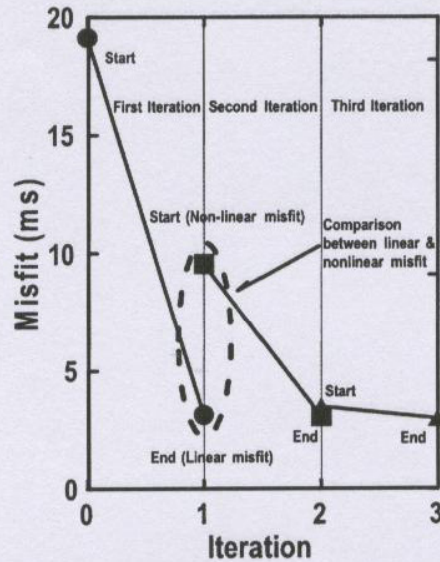

FIG. 11. Same as Figure 5, but the transition between the first and second iteration (marked by the dashed ellipse) is enlarged. This misfit "jump" confirms the nonlinearity of the problem. We can observed that at the end of the first iteration, the linear misfit underestimates the true (nonlinear) misfit by approximately 200%.

The resulting velocity model for the area is given in Figure 12. Figure 12a shows the inversion results. Information is shown only for the areas where the model has been inverted and corresponds to a damping factor ε of 70 and smoothing factor λ of 100. The tomographic images show two large anomalies (layers) that intersect the borehole at 400 and 1000 m. The first anomaly is a low-velocity layer which dips at an angle of ~45° from south to north, and the second is a high-velocity anomaly dipping in the same direction. No veritical seismic profile (VSP) reflection images are available from these depths. However, the subsurface image produced by traveltime tomography is consistent with results from surface seismic reflection (Juhlin et al., 1997). Moreover, images produced by refraction processing of the first arrivals from the surface line exhibit a similar behavior because the layers which have been identified are dipping from south to north (Figure 8). Two more velocity anomalies at depths 2300 m (negative) and 3200 m (positive) are indicated. These anomalies may be related to reflectors observed on surface seismic data from the area (Juhlin et al., 1997).

In order to test the reliability of our results, we calculated the covariance and resolution matrix for the final solution of Figure 12a. Figure 13a shows the distribution of the square root of the diagonal of the covariance matrix. Errors in estimation of velocity model do not exceed 250 m/s and are typically less than 150 m/s. This error is acceptable if we consider that variation of the velocities in the study area is of the order of 1000–2000 km/sec. Figure 13b shows the distribution of the spread function, which is defined as (Toomey and Foulger, 1989)

$$S_j = \frac{\sum_i D_{ij} R_{ij}^2}{\sum_i R_{ij}^2}, \qquad (13)$$

where D_{ij} is the distance between grid nodes i and j, R_{ij} is the element of the resolution matrix of the ith parameter (node), which correspond to the jth parameter (node). The spread function is an attempt to present the resolution matrix in a physical way, such that it displays how smoothed the solution is at each node. Areas where S is larger than 4 grid nodes are not shown in Figure 13b.

In Figure 11a, we observed that the size of the identified anomalies is larger than 4 grid nodes, whereas the corresponding S values range from 2 to 4. Areas with S larger than 4 have also been "masked" in Figure 12.

The nullspace shuttle method (Deal and Nolet, 1996) was used in an attempt to improve a posteriori the tomographic image and to check if we could get information about predefined anomalies from the nullspace. Thus, we performed two types of a posteriori filtering based on a priori geological knowledge. For the first case, we considered that the maximum expected velocity variation for the study area ranges from 4.5 to 7.5 km/sec. Thus, we imposed a specific velocity distribution on the model
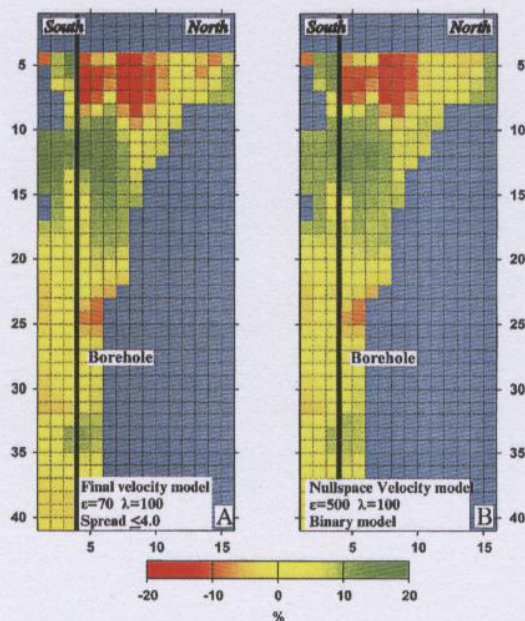

FIG. 12. Final velocity model for study area. (a) Minimum norm solution as obtained after applying damping =70 and smoothing =100. Velocities are plotted only at nodes with spread function $S \leq 4$. We can identify two distinct bodies of anomalous velocities at the depth of 400 and 1200 m. (b) Corresponding conservative nullspace solution starting from a homogeneous background velocity model.
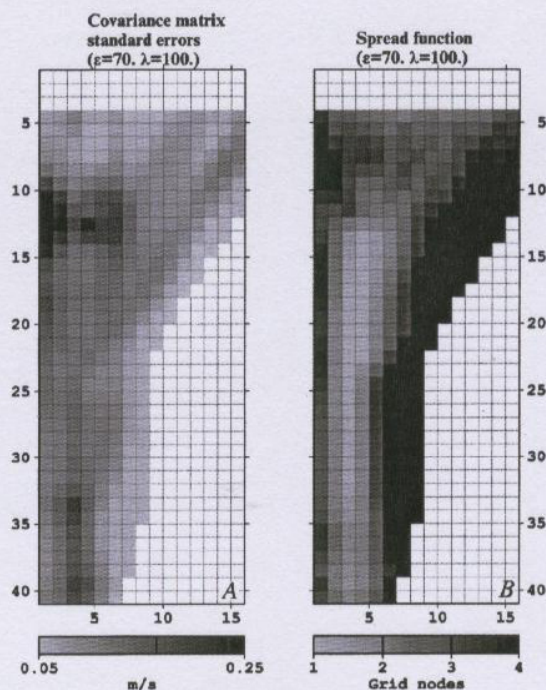

FIG. 13. Plots of covariance matrix and spread function. (a) Obtained velocity standard errors (root of diagonals of the covariance matrix) for the study area. Errors generally do not exceed 200 m/s, which is an acceptable error for the velocity estimation. (b) Plot of spread function for values ≤ 4. Colored (i.e., nonwhite) nodes are quite well resolved.

by employing a binary filter. The simple binary filter we used
compares the velocity of each node with two threshold veloc-
ities (5 and 6.5 km/sec). If the velocity exceeded these limits,
then the velocity of the node was fixed to this a priori value
of the upper/lower velocity; otherwise, it was not changed. The
second filter we applied attempts to use null-space information
to reconstruct the 1-D background velocity model. The reason
we used this velocity model was to check if there was any 1-D
model able to fit the data and explain most of the identified
traveltimes. In order to enlarge the nullspace, we considered
that all singular values less than 500 (268 of 392 eigenvalues, see
Figure 10) belonged in the nullspace. The difference between
these two filtered images and the least-squares solution was
projected onto the nullspace of the model and, therefore, the
resulting components more or less "lie" in the nullspace and
do not violate the misfit of the data. Finally, we added these
components to the minimum norm solution and we obtained
the conservatively filtered images. The corresponding error of
the solution was increased by 10% compared to the model of
Figure 12a. We observe that although less than one third of the
eigenvalues are used, the reconstructed model is practically
the same with our final model. Similar results were obtained
for the 1-D filtered model. These results, together with the re-
sults of Figure 13 confirm the robustness of the final solution.
The a posteriori "improved" tomographic image based on the
binary filter is shown in Figure 12b.

CONCLUSIONS

We have presented an algorithm for nonlinear traveltime
tomography which incorporates an efficient 3-D ray tracing
technique in conjunction with a combination of inversion sta-
bilization approaches. Incorporation of 3-D ray tracing in to-
mography is traditionally prohibited by computational ineffi-
ciency. However, the proposed method and algorithm are quite
efficient since they require a relative small computation time.
The use of supercomputing and more efficient programming
could make the application of our algorithm even more feasi-
ble. We have also demonstrated that our algorithm is capable of
producing improved tomographic images using both synthetic
and real data. The performed tests suggest that 3-D ray tracing
and Fresnel volume calculation cannot be considered a small
second-order effect in small-scale tomography, especially as far
as the amplitude of the recovered anomaly is concerned.

REFERENCES

Aki, K., and Lee, W. H. K., 1976, Determination of three-dimensional
   velocity anomalies under a seismic array using first p arrival times
   from local earthquakes, 1, A homogeneous initial model: J. Geophys.
   Res., **81**, 4381–4399.

Ayarza, P., Juhlin, C., Brown, D., Pechning, R., Beckholmen, Ayala,
   C., Kimbell, G., Rybalka, A., and Glushkov, A., 1998, Origin of the
   shallow reflectivity in the SG4 borehole area from reflection seismic,
   gravity and borehole data: Constrains on the upper crustal structure
   of the Tagil Synform (Middle Urals), Submitted to Journal of Geo-
   physical Research.
Bliznetsov, M., and Juhlin, C., 1994, Analyses of the wavefields by the
   Common Excitation Array (CEA) method: J. Appl. Geophys., **32**,
   245–256.
Červený, V., and Soares E. P., 1992, Fresnel volume ray tracing: Geo-
   physics, **57**, 902–915.
Constable, S. C., Parker, R. L., and Constable, C., 1987, Occam's in-
   version: A practical algorithm for generating smooth models from
   electromagnetic sounding data: Geophysics, **52**, 289–300.
Crosson, R. W., 1976, Crustal structure modeling of earthquake data, 1.
   Simultaneous least-squares estimation of hypocenters and velocity
   parameters: J. Geophys. Res., **81**, 3036–3046.
Deal, M. M., and Nolet, G., 1996, Nullspace shuttles: Geophys. J. In-
   ternat., **124**, 372–380.
Franklin, J. N., 1970, Well-posed stochastic extension of ill-posed linear
   problems: J. Math. Anal. Appl., **31**, 682–716.
Hansen, P. C., 1994, REGULARIZATION TOOLS: A Matlab pack-
   age for analysis and solution of discrete ill-posed problems: Numer-
   ical Algorithms, **6**, 1–35.
Harris, J. M., Nolen-Hoeksema, R. C., Langan, R. T., Van Schaack, M.,
   Lazaratos, S. K., Rector III, J. W., 1995, High-resolution crosswell
   imaging of west Texas carbonate reservoir: Part 1—Project summary
   and interpretation: Geophysics, **60**, 667–681.
Hawley, B. W., Zandt, G., and Smith, R. B., 1981, Simultaneous in-
   version for hypocenters and lateral velocity variations: An iterative
   solution with layered model: J. Geophys. Res., **86**, 7073–7080.
Juhlin, C., Kashubin, S., Knapp, J. H., Makovsky, V., and Ryberg, T.,
   1995, Project conducts seismic reflection profiling in the Ural Moun-
   tains: EOS, **76**, 193–198.
Juhlin, C., Bliznetsov, M., Pevzner, L., Hismatulin, T., Rybalka, A.,
   and Glushkov, A., 1997, Seismic imaging of reflectors in the SG4
   borehole, Middle Urals. Russia: Tectonophysics, **276**, 1–18.
Jupp, D. L. B., and Vozoff, K., 1975, Stable iterative methods for the
   inversion of geophysical data: Geophys. J. Roy. Astr. Soc., **42**, 957–
   976.
Levenberg, K., 1944, A method for the solution of certain nonlinear
   problems in least-squares: Quart. Appl. Math., **2**, 164–168.
Marquardt, D. W., 1963, An algorithm for least-squares estimation of
   nonlinear parameters: SIAM J. Appl. Math., **11**, 431–441.
Moser, T. J., Nolet, G., and Snieder, R., 1992, Ray bending revisited:
   Bull. Seis. Soc. Am., **82**, 259–288.
Officer, C. B., 1958, Introduction to the theory of sound transmission:
   McGraw-Hill.
Papazachos, C. B., and Nolet, G., 1997, Non-linear arrival time tomog-
   raphy: Annali di Geofisica, **40**, 85–97.
Pereyra, V., 1992, Two-point ray tracing in general 3D media: Geophys.
   Prosp., **40**, 267–287.
Pereyra, V., 1996, Modeling, ray tracing and block nonlinear travel-
   time inversion in 3D: Pageoph, **148**, 345–386.
Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T.,
   1988, Numerical recipes in Fortran—The art of scientific computing:
   Cambridge Univ. Press, **292**, 305–309.
Sambridge, M. S., 1990, Non-linear arrival time inversion: constraining
   velocity anomalies by seeking smooth models in 3D, Geophys. J.
   Internat., **102**, 653–677.
Sasaki, Y., 1992, Resolution of resistivity tomography inferred from
   numerical simulation: Geophys. Prosp., **40**, 453–464.
Steer, D. N., Knapp, J. H., Brown, L. D., Rybalka, A. V., and Sokolov,
   V. B., 1995, Crustal structure of the Middle Urals based on repro-
   cessing of Russian seismic reflection data: Geophys. J. Internat., **123**,
   673–682.
Tarantola, A., 1987, Inverse problem theory: Elsevier.
Tarantola, A., and Nercessian, A., 1984, Three-dimensional inversion
   without blocks: Geophys. J. Roy. Astr. Soc., **79**, 299–306.
Thurber, C. H., and Ellsworth, W. L., 1980, Rapid solution of ray tracing
   problems in heterogeneous media: Bull. Seis. Soc. Am., **70**, 1137–
   1148.
Toomey, D. R., and Foulger, G. R., 1989, Tomographic inversion of local
   data from the Hengil-Grensdalur central volcano complex, Iceland:
   J. Geophys. Res., **94**, No. B12, 17497–17510.
Wessel, P., and Smith, W. H. F., 1991, Free software helps map and
   display data: EOS, **72**, 441, 445–446.