GEOPHYSICS, VOL. 62, NO. 3 (MAY-JUNE 1997); P. 869–883, 8 FIGS., 1 TABLE.

Improved inversion through use of the null space

Peter S. Rowbotham* and R. Gerhard Pratt‡

ABSTRACT

Standard least-squares traveltime inversion techniques tend to produce smoothed estimates of the velocity field. More complete results can be obtained by incorporating into the inversion scheme a priori information about the media to be imaged, derived from well logs, core data, and surface geology. A promising technique for achieving this involves projecting this information onto the null space model singular vectors of the inverse problem and including this projection with the non-null space contribution in order to produce a solution. The method, demonstrated with both field and synthetic crosshole traveltime data acquired through layered, anisotropic media, successfully produces improved inversion solutions with lower traveltime residuals, layers that are more homogeneous, sharper interfaces, and better correlated anisotropy parameters than solutions obtained with standard techniques.

INTRODUCTION

Traveltime tomography, whereby direct wave traveltimes are inverted, has become an established technique for estimating velocity images (see, e.g., reviews by Worthington, 1984; Nolet, 1987; Iyer and Hirahara, 1993). Tomography using primary reflected arrivals has also proved successful (e.g., Bishop et al., 1985). The resulting velocity models are of fundamental importance for detecting subtle changes in lithology, porosity and permeability, and for accurate migration of seismic data.

To overcome the problem of nonuniqueness for ill-posed problems, standard inversion techniques, such as least-squares Gaussian methods (e.g., the simultaneous iterative reconstruction technique (SIRT), Dines and Lytle, 1979), tend to smooth the resulting images compared to the actual geologic media (Gersztenkorn and Scales, 1988; Pratt and Sams, 1996). It is not the model parameters themselves that are estimated, but weighted spatial averages of the model parameters (Menke, 1984). The rather ad hoc Gaussian and smoothness assumptions are at variance with our knowledge of the actual geology (Walden and Hosken, 1986). We know a priori from well logs, core data, and surface geology analogs that geologic processes operate on different scales, creating a combination of sharp contrasts between physical parameters of the adjoining rock masses and subtle variations within such lithological units. Upon least-squares inversion, however, the interfaces become gradual gradients from one limit to the other, and the relatively homogeneous ("blocky") geologic units acquire false velocity textures.

Anisotropy is a further problem for velocity field determination. Although the nature of the deposition process would favor transverse isotropy with a vertical symmetry axis (TIV) in many sedimentary rocks, until recently inversion methods have standardly assumed isotropic velocity structures. Isotropic inversion of an anisotropic medium will result in distorted images (as noted by, e.g., Chiu and Stewart, 1987; McCann et al., 1989; Carrion et al., 1992; Pratt et al., 1993; Michelena et al., 1993; Williamson, 1993; Pratt and Sams, 1996). Even when anisotropic inversion is performed, it is found that the resolution of the individual anisotropy parameters is dissimilar at the same spatial locations (Williamson, 1993; Ory and Pratt, 1995), an effect often caused by limited ray-angle coverage. Using a limited number of "blocks" to parameterize the geology would overcome this problem (e.g., Michelena, 1992), though such a parameterization does not respect subtle velocity variations within lithological units and can lead to misinterpretation if an incorrect scale is used for the size of the blocks (Delprat-Jannaud and Lailly, 1992).

Since the effective anisotropy of a medium is often estimated from the local ratio of horizontal and vertical velocity parameters, it would be desirable to balance the spatial resolution of these parameters at all locations in a velocity section, but without an explicit blocky parameterization. The anisotropy parameters are usually geologically dependent on each other, i.e., where there is a discontinuity in one parameter (e.g., across an interface), we would expect a discontinuity in the others. In other words, all parameters are expected to respond to a change in lithology in a similar manner, if not to the same degree.

Manuscript received by the Editor August 14, 1995; revised manuscript received June 28, 1996.
*Formerly Department of Geology, Imperial College, Prince Consort Road, London SW7 2BP, United Kingdom; presently Elf Exploration UK, Geoscience Research Centre, 30 Buckingham Gate, London SW1E 6NN, United Kingdom.
‡Department of Geology, Imperial College, Prince Consort Road, London SW7 2BP, United Kingdom.

The incorporation of a priori information into inverse problems has been suggested by many authors (e.g., Tarantola and Nercessian, 1984; Gersztenkorn and Scales, 1988; Carrion, 1991; Ellis and Oldenburg, 1994). The most popular schemes, such as regularization, singular-value decomposition (SVD) truncation, and other damping schemes, implicitly incorporate a priori information. Cluster analysis, which assumes that lithological types have similar properties, has been described for seismic attribute mapping (Lendzionowski et al., 1990). Hyndman et al. (1994) describe lithological zonation methods for aquifer geometry structure. Maximum entropy methods have also been used (e.g., Zhang et al., 1994).

A scheme to encourage horizontally layered solutions, achieved by locally clustering the velocity cells, has been reported in Neri et al. (1993) and Carrion et al. (1993). Their method centers around maximizing the a posteriori probability field (MAP) using a Gibbs distribution to describe the model statistics. They used a neighborhood system of the eight adjacent cells, this being adequate to describe smoothly varying slownesses within geologic units, with sharp discontinuities at the boundaries. Connectivity orientation was introduced so that horizontal layering could be favored, and the minimum potential field would then occur where layers were homogeneous. This potential field was introduced in Neri et al. (1993) into the objective equation of the inverse problem. They sought the minimum potential field and the minimum traveltime residuals simultaneously.

Deal and Nolet (1993; 1996) recently described the concept of filtering the inversion solution (a posteriori) and projecting this filtered image onto the null-space model singular vectors to ascertain the permissible changes wrought by the filter such that the data fit is conserved. This projection onto the null space, termed the "nullspace shuttle," is a starting point for this paper. Deal and Nolet (1996) demonstrated the nullspace shuttle with two types of filtering process: clustering and an alpha-trimmed mean filter (a combination of mean and median filtering which preserves edges).

In this paper, we claim to improve on the basic least-squares tomographic approach by incorporating into our inversion scheme a priori information about the media to be imaged, derived from logs, core data, and geologic outcrops. In the case of our example data sets, we believe a priori that the media are homogeneous, TIV sedimentary units bounded by sharp interfaces, and with correlated anisotropy parameters. The improvement in imaging is achieved by projecting a *desirable* solution onto the null space of the inversion, and including this nullspace contribution with the standard non-null space inversion solution. The desirable solution is derived a posteriori from the solutions obtained after regularized inversion and after SVD truncation, and is consistent with our a priori attributes. This is a more efficient route than using the whole model space. By working solely in the null space of a linear problem, the traveltimes are guaranteed to be unaffected, and there is no need to reiterate at every step. However, for quasi-linear problems such as traveltime inversion, retracing of rays could be performed occasionally if desired.

We should note from the outset the similarities between our proposed method and that of initializing the tomographic model with the best independent measurements of the velocity field, i.e., our desirable solution. In this latter case, tomographic inversion would change only those parts of the assumed model that contradicted the data; we shall see that the method of projection of a desirable solution onto the null space works in a similar fashion. It is important to note that our choice of a desirable model to use in the projection is only one choice of many possible operations on the model. Deal and Nolet (1996) have discussed many of these, and we will touch on this subject later in this paper.

We use the method of singular-value decomposition (SVD) (Lanczos, 1961) to define the null space of the inversion and to elucidate the projection method. For large inverse problems where SVD is impractical, Deal and Nolet (1996) have shown how to achieve the projection onto the null space without SVD. For a general introduction to the inversion methodology used in this work, we refer readers to Bregman et al. (1989b) and Pratt and Chapman (1992). Here we shall first give mathematical treatments of the advantages of using the null space and the projection of a desirable a posteriori solution onto the null space of the inverse problem, before demonstrating the improvements achieved with field and synthetic crosshole traveltime data in an anisotropic environment. The results will show that projection of a filtered solution onto the null space is a practical method for significantly improving the match between velocity images and a priori information without affecting the data fit, and of correlating the spatial resolution of different anisotropy parameters.

SINGULAR-VALUE DECOMPOSITION AND THE NULL SPACE

Definition of the null space

Linear tomographic systems are characterized by an $M \times N$ Frechét matrix $\underset{\sim}{\mathbf{F}}$, which is a linear mapping from the N-dimensional parameter space to the M-dimensional data space:

$$\mathbf{t} = \underset{\sim}{\mathbf{F}}\mathbf{m} \qquad \text{or} \qquad \delta\mathbf{t} = \underset{\sim}{\mathbf{F}}\delta\mathbf{m}. \tag{1}$$

[The first case is appropriate only for inherently linear problems, such as fixed-raypath tomography involving only the slowness parameter. The second case, obtained by linearization of nonlinear systems, is more generally applicable. In this paper, we always linearize about an isotropic model, and we always use squared slowness parameters, so that we will use the second form of equation (1) from now on.] For isotropic tomography, N is the number of cells chosen to discretely represent the velocity field and M is the number of raypaths. For any matrix $\underset{\sim}{\mathbf{F}}$ there exists the SVD

$$\underset{\sim}{\mathbf{F}} = \underset{\sim}{\mathbf{U}}\,\underset{\sim}{\Lambda}\,\underset{\sim}{\mathbf{V}}^T. \tag{2}$$

For the linear tomographic problem, the matrix $\underset{\sim}{\mathbf{U}}$ is a matrix of singular vectors that span the data space, $\underset{\sim}{\mathbf{V}}$ is a matrix of singular vectors that span the parameter space, and $\underset{\sim}{\Lambda}$ is a matrix whose diagonal elements are known as the singular values of $\underset{\sim}{\mathbf{F}}$. Zero-valued singular values remove the contribution of the corresponding singular vectors, so that the matrices in the previous equation can be partitioned as

$$\underset{\sim}{\mathbf{F}} = \underset{\sim}{\mathbf{U}}_p\,\underset{\sim}{\Lambda}_p\,\underset{\sim}{\mathbf{V}}_p^T, \tag{3}$$

where the subscript p indicates that only the columns of $\underset{\sim}{\mathbf{U}}$ and $\underset{\sim}{\mathbf{V}}$ that are multiplied by nonzero singular values are retained. In other words, the SVD has been truncated, with the retained p singular vectors of $\underset{\sim}{\mathbf{V}}$ spanning the range of $\underset{\sim}{\mathbf{F}}$, and the rejected

columns of $\mathbf{V}$ spanning the null space of $\mathbf{F}$. We define the matrix $\mathbf{V}_{\text{null}}$ as the matrix of these columns of $\mathbf{V}$ thus excluded from $\mathbf{V}_p$. In fact, we use less than p singular vectors, i.e., we use an effective null space that is larger than the strict definition. Deal and Nolet (1996) have termed this the "generalized null space."

Why use the null space?

The central plank of the use of the null space is that any linear combination of the null space singular vectors will have no effect on the modeled traveltimes. Our estimate of the model perturbation parameters is

$$\delta\hat{\mathbf{m}} = \mathbf{F}^{\dagger}\delta\mathbf{t} = \mathbf{F}^{\dagger}\mathbf{F}\delta\mathbf{m} = \mathbf{V}_p\mathbf{V}_p^T\delta\mathbf{m}, \tag{4}$$

where the pseudo-inverse $\mathbf{F}^{\dagger} = \mathbf{V}_p\mathbf{\Lambda}_p^{-1}\mathbf{U}_p^T$ when using the "generalized inverse," or SVD truncation. $\mathbf{F}^{\dagger}\mathbf{F} = \mathbf{V}_p\mathbf{V}_p^T$ is the model resolution matrix (Menke, 1984). Therefore the predicted traveltime data are

$$\delta\hat{\mathbf{t}} = \mathbf{F}\delta\hat{\mathbf{m}} = \mathbf{F}\mathbf{F}^{\dagger}\delta\mathbf{t} = \mathbf{U}_p\mathbf{U}_p^T\delta\mathbf{t}, \tag{5}$$

where $\mathbf{F}\mathbf{F}^{\dagger} = \mathbf{U}_p\mathbf{U}_p^T$ is the data resolution matrix (Menke, 1984).

Now, if we add linear combinations of vectors from the null space,

$$\delta\hat{\mathbf{m}}_{\text{null}} = \sum_{i=p+1}^{N}\alpha_i\mathbf{v}_i = \sum_{\text{null}}\alpha_i\mathbf{v}_i, \tag{6}$$

to $\delta\hat{\mathbf{m}}$, where the limit *null* represents $i = p+1, N$, then the new predicted data $\delta\hat{\mathbf{t}}'$ are

$$\begin{aligned}\delta\hat{\mathbf{t}}' &= \delta\hat{\mathbf{t}} + \mathbf{F}\delta\hat{\mathbf{m}}_{\text{null}} \\ &= \delta\hat{\mathbf{t}} + \mathbf{U}_p\mathbf{\Lambda}_p\mathbf{V}_p^T\delta\hat{\mathbf{m}}_{\text{null}}.\end{aligned} \tag{7}$$

All scalar products $\mathbf{v}_i^T\delta\hat{\mathbf{m}}_{\text{null}}$ are zero, since the vectors $\mathbf{v}_i^T$ (the rows of $\mathbf{V}_p^T$) are orthogonal to all vectors in $\mathbf{V}_{\text{null}}$, and therefore

$$\delta\hat{\mathbf{t}}' = \delta\hat{\mathbf{t}} = \mathbf{U}_p\mathbf{U}_p^T\delta\mathbf{t}, \tag{8}$$

i.e., we do not change the data predictions.

Projection onto the null space

By performing a generalized inverse on the SVD truncated matrices [equation (4)], we obtain

$$\delta\hat{\mathbf{m}} = \mathbf{F}^{\dagger}\delta\mathbf{t} = \mathbf{V}_p\mathbf{\Lambda}_p^{-1}\mathbf{U}_p^T\delta\mathbf{t}, \tag{9}$$

where $\delta\hat{\mathbf{m}}$ is the estimated solution and $\mathbf{F}^{\dagger}$ represents the pseudo-inverse of the Frechét matrix. If we now project an a priori consistent desirable model perturbation $\delta\mathbf{m}_{\text{a priori}}$ onto the null space model singular vectors,

$$\delta\hat{\mathbf{m}} = \mathbf{V}_p\mathbf{\Lambda}_p^{-1}\mathbf{U}_p^T\delta\mathbf{t} + \mathbf{V}_{\text{null}}\mathbf{V}_{\text{null}}^T\delta\mathbf{m}_{\text{a priori}}. \tag{10}$$

The first term represents the contribution from the range of the SVD truncation, whereas the second term is the projection of the desirable solution onto the null space.

We propose the following scheme to exploit this (Figure 1a):

1) invert the traveltime data using SVD on the linearized forward problem to identify the singular vectors,

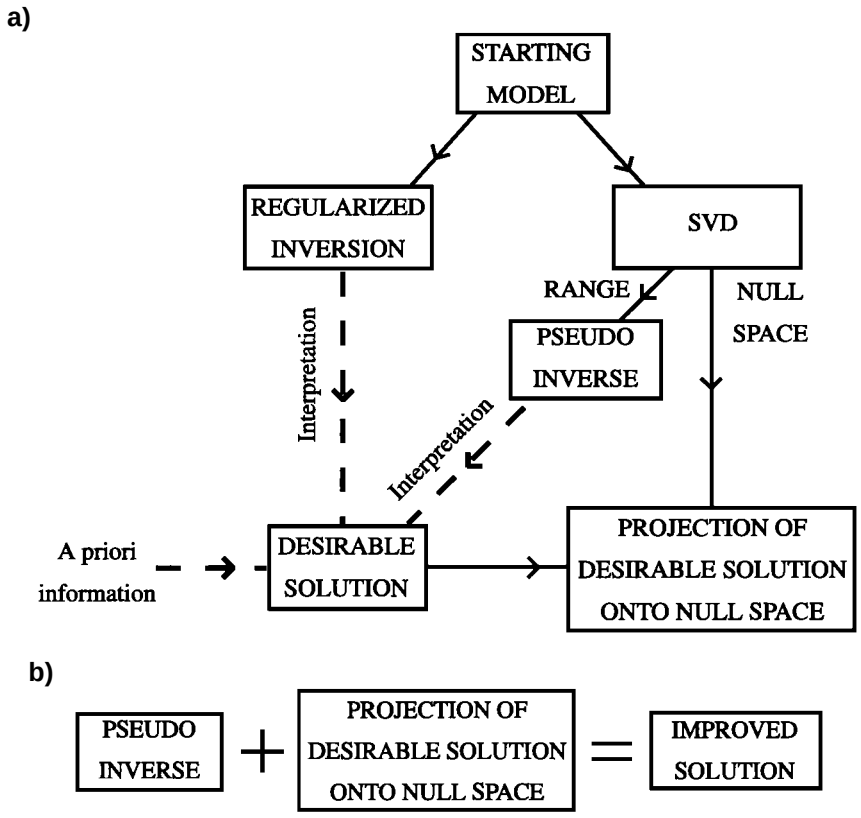

FIG. 1. Schematic representation of the processing scheme for inversion using null-space projection. (a) Projection of a desirable solution onto the null space. (b) Improved solution.

2) truncate the SVD to obtain a well-conditioned solution, and to define the null space of the problem,
3) project a desirable solution (i.e., one agreeing with the a priori information) onto the null space singular vectors, and include this (Figure 1b) with the truncated contribution from the range of $\underset{\sim}{\mathbf{F}}$ (pseudo-inverse) to obtain an improved final solution.

Since this scheme does not affect the traveltimes, the use of the null space is more efficient than using the whole model space, which would result in new traveltimes and thus an implied need for iteration.

Comment on truncation level

One problem with the above description of the SVD truncation method is that the singular values are not divided simply into nonzero and zero values. In practice, there is a drop off of singular values, and to damp the noise amplification effects of small but nonzero singular values, it is desirable to choose some threshold level, $k < p$, below which singular values and their corresponding singular vectors will be removed from the inversion (Bregman et al., 1989a). This threshold value is generally chosen so that the ratio of the maximum singular value to the cutoff value is similar to the ratio σ_d/σ_m, where σ_d is the standard deviation of the traveltime data perturbations and σ_m is the standard deviation of the model perturbations. Where estimates for these standard deviations are not readily available, the choice of truncation level is based instead on a trade off between model smoothness and data fit.

Often, the level of detail increases as the magnitude of the singular value decreases; thus the low-wavenumber features of the solution are provided by the large singular values, and high-wavenumber features by the small singular values (Zhang and McMechan, 1995). However, readers should note that this is not always the case (Bregman et al., 1989a; Scales et al., 1990), especially when parameters that can trade off with another are involved, e.g., velocity and depth. By assigning small singular values to the null space, we are allowing the projection method freedom to improve on the poorly resolved features of an inverse problem.

Considering the effect of an expanded generalized null space on our data predictions, equation (6) will become

$$\delta\hat{\mathbf{m}}_{\text{gen-null}} = \sum_{i=k+1}^{N} \alpha_i \mathbf{v}_i = \sum_{\text{gen-null}} \alpha_i \mathbf{v}_i, \qquad (11)$$

where the limit gen-null represents $i = k+1, N$ ($k < p$), and equation (7) for the predicted data $\delta\hat{\mathbf{t}}'$ becomes

$$\begin{aligned}\delta\hat{\mathbf{t}}' &= \delta\hat{\mathbf{t}} + \underset{\sim}{\mathbf{F}}\delta\hat{\mathbf{m}}_{\text{gen-null}} \\ &= \delta\hat{\mathbf{t}} + \underset{\sim}{\mathbf{U}}_p\underset{\sim}{\Lambda}_p\underset{\sim}{\mathbf{V}}_p^T \sum_{\text{gen-null}} \alpha_i \mathbf{v}_i \\ &= \delta\hat{\mathbf{t}} + \underset{\sim}{\mathbf{U}}_p\underset{\sim}{\Lambda}_p(0, 0, \ldots, \alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_p)T. \qquad (12)\end{aligned}$$

There now exist $p - k$ nonzero component in the final vector. The choice of truncation level k can now be seen to be a trade off between the magnitude of the second term in equation (12) (the data fit) and the effective damping required to minimize the variance of the model solution. The important difference between using a generalized null space rather than a true null space is that the data misfit will be increased. There is some art involved in the choice of the truncation level, which will act as a balancing factor between the data misfit and the variance of the solution.

Comment on choice of solution to project

The choice of a desirable solution to project onto the null space can be made in several ways. Deal and Nolet (1996) used synthetic data to demonstrate two methods of producing a desirable solution a posteriori. First, by imposing a bimodal velocity distribution on the initial inversion solution (c.f., Hyndman et al., 1994), the a priori assumption is that the region is composed of only two distinct rock types. The second, more generally applicable method uses an alpha-trimmed mean filter (a combination of arithmetic mean and median filters after Gersztenkorn and Scales, 1988), which smooths the velocity solution a posteriori while trying to preserve edges.

To add to the above suggestions, we have tested several more a posteriori strategies for producing a desirable solution: (1) multiparameter modal clustering, (2) median filtering with correlation between the separate anisotropy parameters at the same spatial locations, (3) use of a regularized solution, and (4) construction by the geophysicist of a desirable solution based on regularized and SVD truncated solutions (Figure 1a). Here we have chosen to discuss only this last scheme that has produced the most promising results to date. Although there is no doubt that this scheme will incorporate user bias into the desirable solution, the projection onto the null space supplies a safety net by showing which aspects of this bias are justified by the data. The user bias can be considered as additional information, or data, that should be incorporated when a meaningful solution is required.

APPLICATION TO SYNTHETIC AND FIELD DATA

To illustrate the method, we have used crosshole traveltime data. However, it should be stressed that these inversion methods are not restricted to the crosshole traveltime case, nor even to seismics, but are generally applicable to many inverse problems.

Acquisition and previous analysis

In 1988 the British Geological Survey (BGS) performed a series of crosshole experiments at Purton in the UK (McCann et al., 1989). The boreholes penetrated the Oxford Clay, an Upper Jurassic formation outcropping in the south of England. At the location of the survey, the clay is about 100 m thick and overlies a limestone layer. The Oxford Clay is largely homogeneous and plane layered, but there are stringers of variable sand content that influence its elastic properties significantly. One such high-velocity sandy layer with a thickness of about 5 m has been noted at 30 m depth on natural gamma logs and on the velocity tomograms once a simple elliptical anisotropy assumption was made. The Oxford Clay has been determined by a number of experiments (reviewed in McCann et al., 1989) to exhibit transverse isotropy with near-vertical symmetry axes, with an anisotropy factor on the order of 10%. Our a priori geologic information about the media then is that they are largely homogeneous, plane layered, and with TIV of about 10%.

For the purposes of testing our methods, we have followed Pratt and Chapman (1992) in using crosshole compressional

traveltime data from a survey comprising 15 source and 14 receiver positions at 4 m spacing in boreholes that were 20 m apart (Figure 2). Media velocities were in the range 1.4–2.2 km/s. McCann et al. (1989) found that their initial tomograms, formed without taking anisotropy into account, bore little relationship to the known stratigraphy. Using a crude approximation for anisotropy, a more plausible image was obtained. Pratt and Chapman (1992) successfully inverted these 210 traveltimes using parameters for an isotropic velocity plus five anisotropic velocities (i.e., five independent parameters) (Chapman and Pratt, 1992). Regularization parameters were chosen so as to minimize the traveltime residuals and the roughness of the solution. Here we have used just three anisotropy parameters to characterize the TIV medium, which is an important consideration because of the size limitations of the SVD matrix computations.

Isotropic inversion

The model was parameterized using a grid 8 cells wide and 24 cells high, yielding a total of 192 cells of size 2.5 m square. Initially, SIRT isotropic inversion was performed from a homogeneous starting model with straight rays. Although there were more traveltime data than parameters, anisotropy has not yet been accounted for and the solution contains artifacts and does not agree with the known stratigraphy (Figure 3a).

The traveltime residual mapping for this isotropic solution is included in Figure 3a. Each pixel in this figure is shaded according to the traveltime residual (computed observed traveltime) for a single source-receiver pair. The source depths increase from left to right along the horizontal axis, whereas the receiver depths increase from top to bottom down the vertical axis. In this format, horizontal ray times corresponding to equal source-receiver depths plot down the leading diagonal, with higher angle raypaths plotting in the top right and bottom left-hand corners. The importance of these residual time surfaces was emphasized in Pratt et al. (1993). A low rms residual error alone is insufficient support for an inversion solution, since a good inversion solution should also result in removing all systematic trends from the residuals on such displays. A correlation between ray angle and data residuals is usually an indication that the solutions are insufficiently anisotropic.

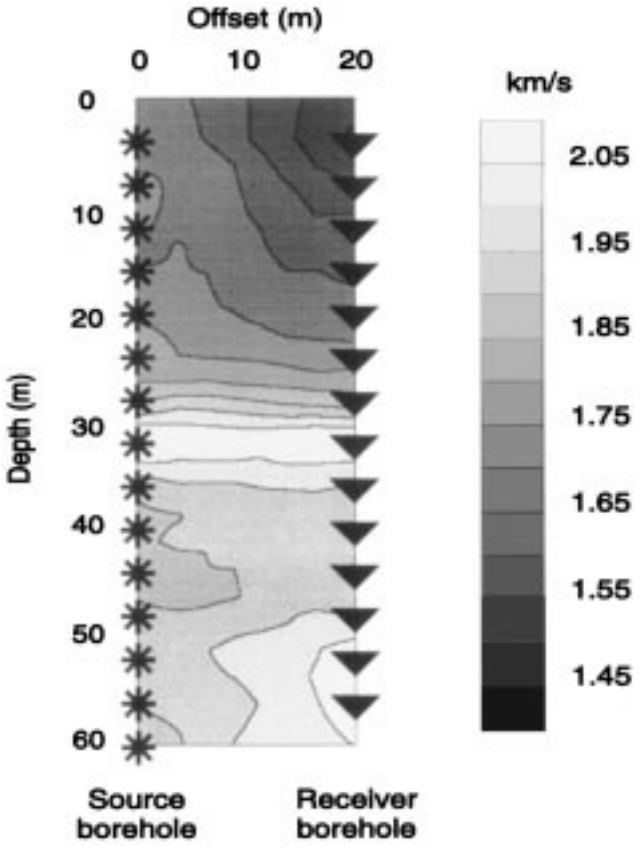

FIG. 2. Crosshole acquisition geometry and average velocity of the regularized solution shown in Figure 3b.

Anisotropic parameterization and inversion

TIV anisotropy can be parameterized using the three anisotropy parameters q_1, q_3, and q_5 (Chapman and Pratt, 1992), specifying group velocities according to

$$v^2(\theta) = q_1 \sin^4\theta + q_3 \sin^2\theta \cos^2\theta + q_5 \cos^4\theta, \qquad (13)$$

where θ is the ray angle with respect to the vertical. In equation (13), q_1 is the square of the horizontal velocity, q_5 is the square of the vertical velocity, and q_3 affects velocities at intermediate angles. These parameters are directly interchangeable with Thomsen's (1986) parameters [$q_1 = \alpha_0^2(1+2\epsilon)$, $q_3 \approx 2\alpha_0^2(1+\delta)$, $q_5 = \alpha_0^2$].

(In Figures 3, 5, 6, 7, and 8, we will display the anisotropic solutions both in terms of the three parameters q_1, $q_3/2$, and q_5, and in terms of ϵ and δ. We find it instructive to plot the results in this fashion, since different features are apparent in the two parameterizations. The choice of $q_3/2$ was made so as to be able to plot the three q parameters on the same scale, and hence the perturbations of the $q_3/2$ plot will appear to be of half the magnitude of those for the horizontal q_1 and vertical q_5 solutions.)

With three TIV parameters and 192 cells, there are 576 parameters in the anisotropic inversion, and the problem is underdetermined. The SIRT isotropic solution (Figure 3a) was initially used as the background model to which the anisotropic perturbations were applied. As the solution improves, better estimates of the isotropic model can be made so as to minimize the perturbations necessary (Figure 1). Straight rays were used for computing the Frechét partial derivatives matrix $\mathbf{F}$ and the data differences $\delta\mathbf{t}$.

We acknowledge the limitation of straight rays (for example, horizontal rays will have less curvature than nearer vertical ones, and straight rays will be slower than the correct curved ones, introducing an apparent velocity anisotropy). For the instructive purposes of this paper, though, we have chosen to use only straight rays, since the use of bent rays tends to produce a variation of ray coverage, with subsequent effects on the solution and on the null space. However, the methods discussed here are equally applicable to curved ray tomography with iterative corrections, with the proviso that cells may become unsampled by the curved rays, and, in the absence of regularization, the associated parameters would be free to take on any value. The use of curved rays would therefore lead to an even greater need for the regularization and null-space techniques advocated in this paper.

For regularized inversion, a first differences gradient regularization on the model solution was used in both horizontal and

vertical directions. The regularization augments the Frechét matrix and creates a new problem of full rank. The optimum regularization parameters were obtained by producing trade off curves of the data residuals and the solution roughness, and the knee point selected (Pratt and Chapman, 1992). The anisotropic solution and the "average" velocity (effectively an average over ray angles of the velocities, Pratt and Chapman, 1992) derived from this anisotropic solution are shown in Figure 3b and Figure 2, respectively. We can proceed using this average velocity as the isotropic background, since it is a truer representation of the stratigraphy. The perturbations are then almost purely anisotropic.

An SVD of the Frechét matrix was performed and the singular values were truncated after the 150th largest singular value (Figure 4). The sharp cut off after the 210th singular value is because 210 data are being used to invert for 516 parameters. Although we would have been inclined to choose a lower threshold value, say after 190–200 singular values, it was

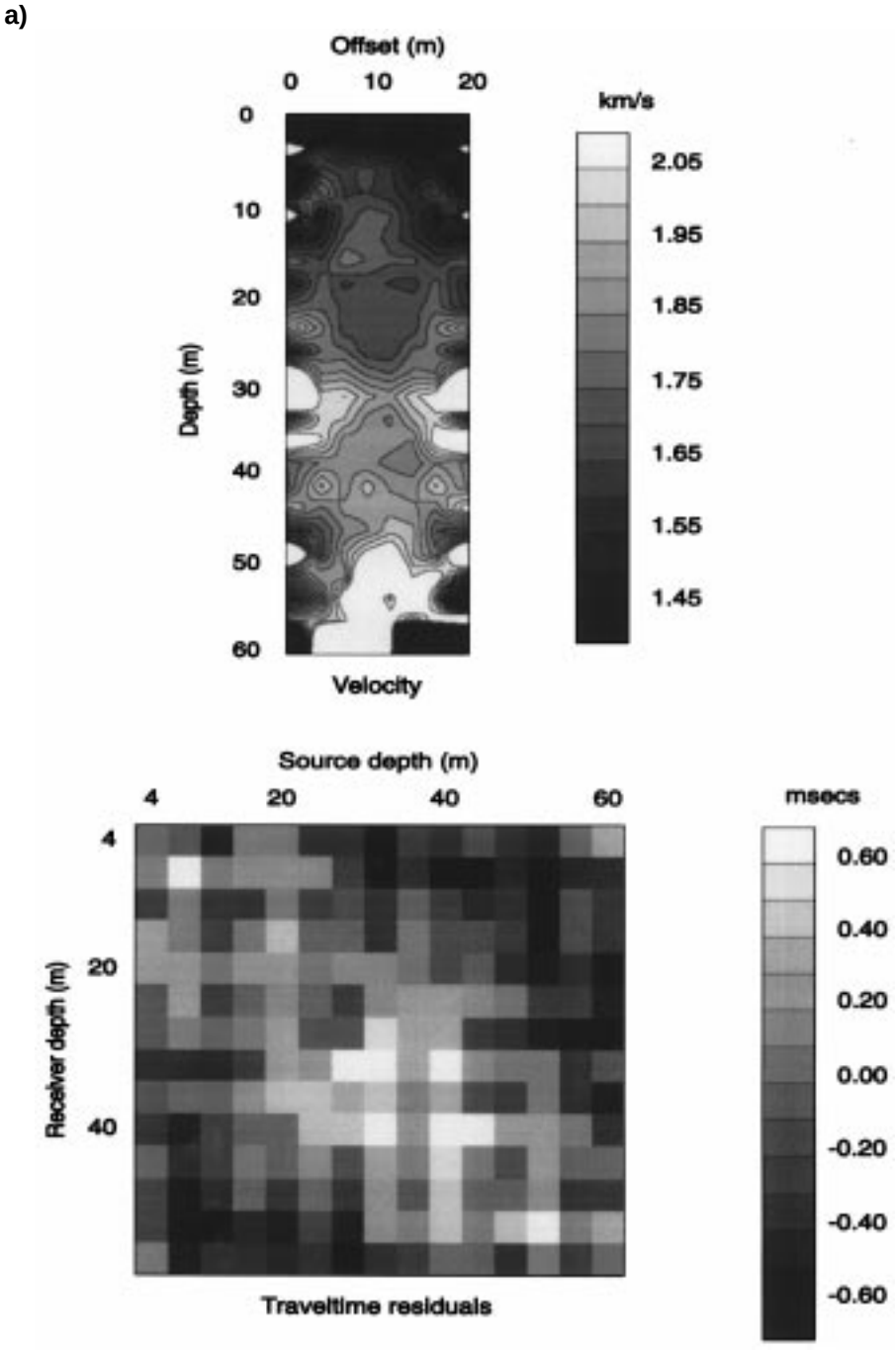

FIG. 3. (a) Field data. Isotropic inversion solution for the Purton traveltime data.

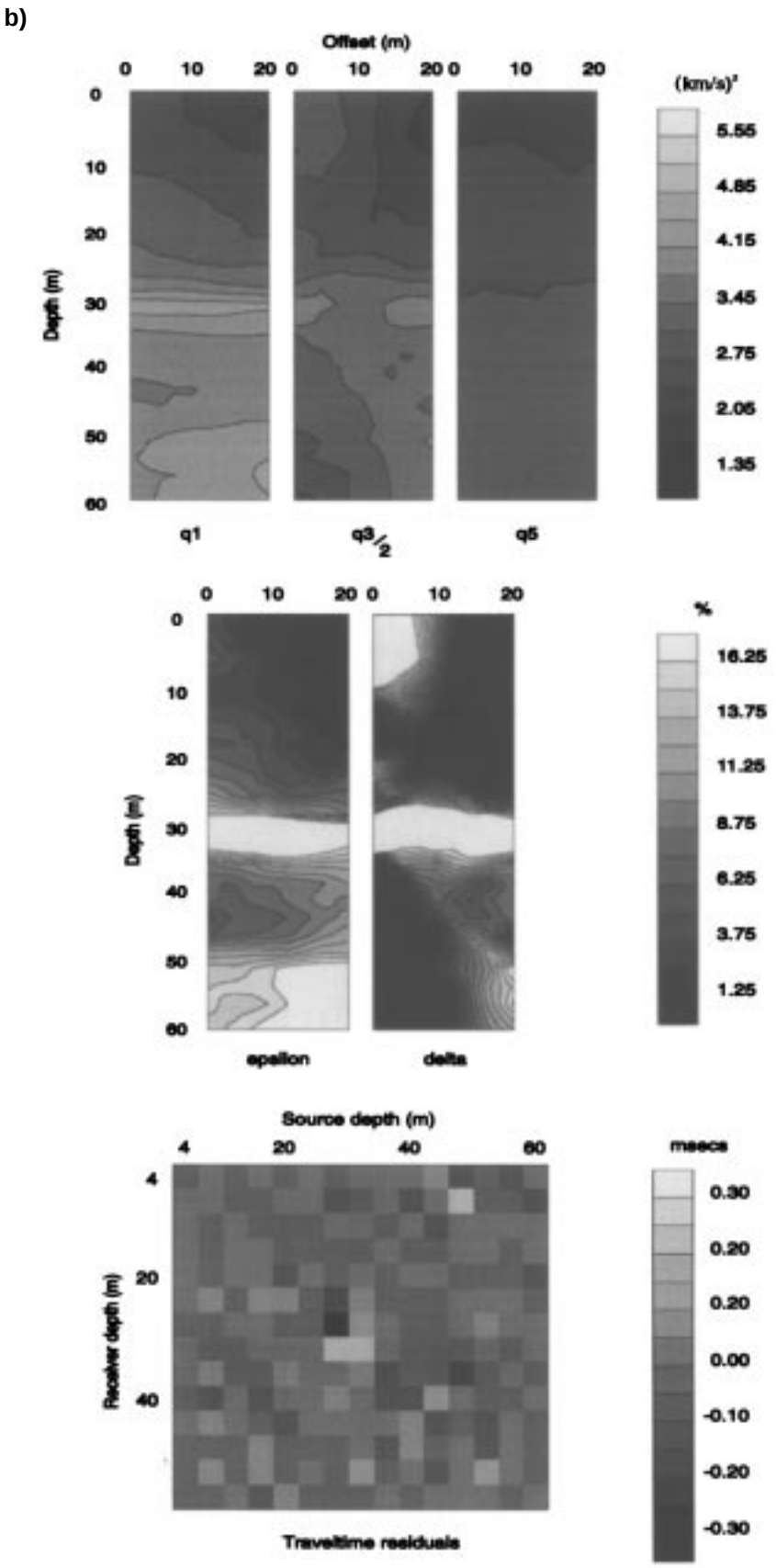

FIG. 3. (b) Field data. Solution for regularized anisotropic inversion.

found that the resulting solutions were implausible since they contained negative values for velocity. Furthermore, assigning more singular values to the null space allows greater flexibility in obtaining improved inversion results, though at the expense of data residuals. Using the pseudo-inverse [equation (9)] to produce the SVD truncated solution in Figure 5, we have chosen to remove from the matrix computation those columns of zeroes corresponding to cells that are not crossed by any ray. These were eight cells along the top of the model grid and three in the bottom right-hand corner, and so the number of singular values was reduced to 543. These cells were assumed to have no anisotropic perturbation and were assigned the isotropic background value in the solution.

From this SVD truncated solution (Figure 5) and from the regularized solution (Figure 3b), we have interpreted a plane layered *desirable* solution, with variable anisotropy in each layer (Figure 6a and Table 1). The degree of anisotropic behavior is seen to increase with depth.

For validation and comparison purposes, a similar set of synthetic traveltimes were obtained by straight ray tracing through this model, and processed in an identical scheme to the real field data (Figures 6b, 6c), the only difference being that 16 source and 16 receiver positions from 0 to 60 m depth at 4 m spacing were used, and thus all cells were crossed. To make a direct comparison between the field and synthetic results, we have kept the same regularization parameters to produce Figure 6b and level of truncation for Figure 6c that were designed for the field data, even though the noise-free synthetic data would afford the possibility of less severe regularization and truncation. The plane layered model (Figure 6a) was then projected onto the null-space model singular vectors (Figure 1) and added to the non-null space SVD truncated solutions for the synthetic data (Figure 7) and the field data (Figure 8).

Table 1. Parameters of the plane-layered model.

Layer depth (m)	α_0 (km/s)	ϵ	δ	q_1 $(km/s)^2$	q_3 $(km/s)^2$	q_5 $(km/s)^2$
0–14	1.60	.05	.02	2.816	5.220	2.56
14–30	1.75	.095	.05	3.644	6.418	3.0625
30–34	2.00	.11	.075	4.880	8.568	4.00
34–48	1.80	.14	.10	4.147	7.078	3.24
48–60	2.00	.17	.12	5.360	8.884	4.00

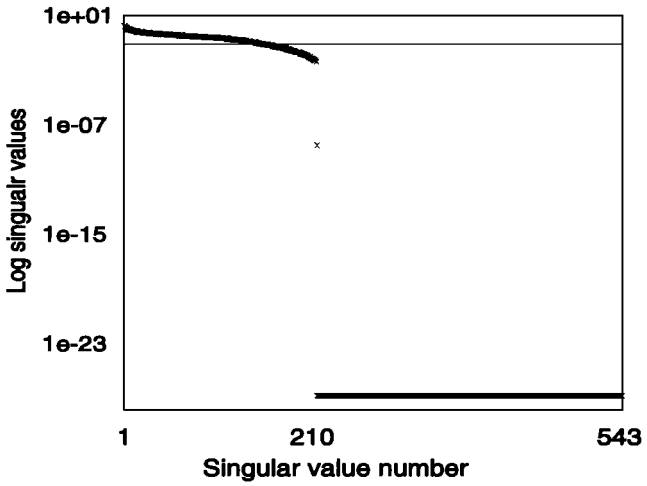

FIG. 4. Singular values for the field data example, showing the truncation level after 150 singular values used to produce the SVD truncated solution shown in Figure 5 and to define the null space of the problem.

RESULTS

Considering first the synthetic results, the regularized inversion has failed to image the high-velocity layer on the vertical solution q_5, only poorly images the layer on the q_3 solution (Figure 6b), and the interfaces are smeared. This is shown dramatically on the ϵ and δ plots, where the thin layer acquires very large values (though had the Thomsen parameters been regularized, rather than the δq parameters, the ϵ and δ plots would have been the more reasonable). This is corroborated also by the traveltime map, where the largest residuals (of both polarities) are associated with source and receiver positions around the central high-velocity layer, and near layer interfaces. Even though the horizontal velocity solution is seen to have been better imaged than the vertical, the largest residuals lie along the leading diagonal of the traveltime surface, i.e., associated with near-horizontal raypaths. This is because these horizontal raypaths will spend relatively more of their length in an incorrect velocity field than will more vertical raypaths.

The problems of poor ray coverage and poor angular coverage are apparent in the strongly varying patterns at the sides of the SVD truncated solution and in the lower regions of the section (Figure 6c). This observation is borne out in the traveltime map, where patterns are apparent for both deep horizontal raypaths along the leading diagonal and for higher angle raypaths. Regularization combats the problem of poor ray coverage near the borehole locations (Figure 6b), but at the expense of resolution, especially of the vertical velocity solution.

Projection of the plane-layered model onto the null space (Figure 7) overcomes the problem of ray coverage without resorting to smoothing. The projection result recovers the plane-layered model perfectly, as one would expect from the manner in which this test was carried out. The traveltime residuals betray slight numerical error (a factor of 1000 less than variations in Figures 6b and 6c). As desired, the different anisotropy parameters are now of a comparable vertical resolution, resulting in improved ϵ and δ solutions, and the interfaces and high-velocity layer are well resolved.

The isotropic solution for the field data set suffers from poor ray coverage at the sides of the image (Figure 3a). The traveltime residuals have a systematic trend of high positive polarity residuals for horizontal raypaths and high negative residuals for steeper angle raypaths, indicative of the need to account for anisotropy (note the scale change for this residual plot compared to Figures 3b, 5, and 8, and note also that the figures for the field data case and for the synthetic case have different scales). As with the inversion of the synthetic data, the regularized inversion (Figure 3b) again loses the high-velocity layer for the q_3 and q_5 solutions, leading to false values for ϵ and δ, and the largest residuals reflect the incorrect imaging of this layer. The SVD truncated results (Figure 5) are rough, but show the general lithological units and interfaces under interrogation. Null-space projection results (Figure 8) again show dramatically improved resolution of the high-velocity layer and other interfaces, improved homogeneity of the layers, and better correlation between the δq parameters. Further a posteriori median filtering of this solution followed by repeated

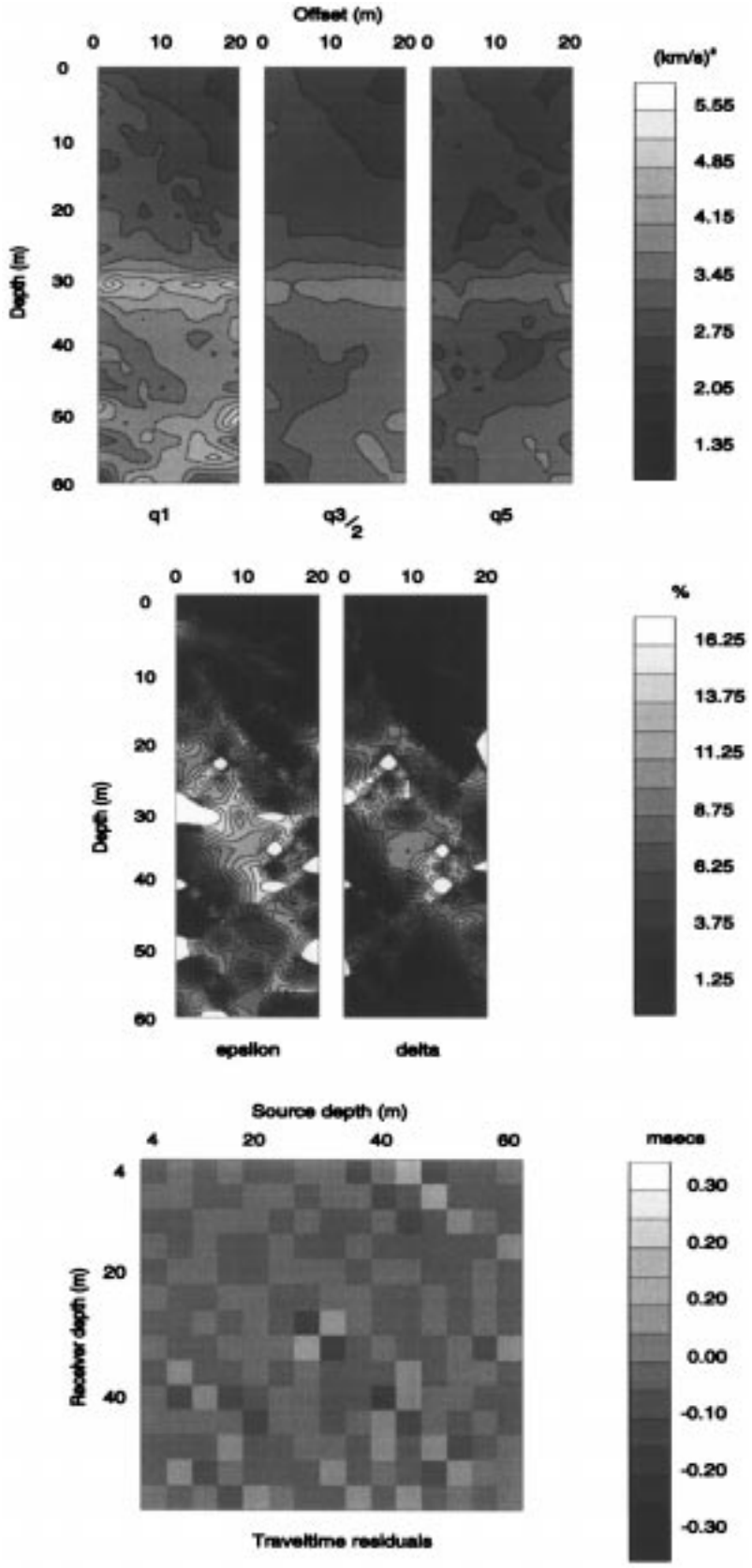

FIG. 5. Field data. Solution for anisotropic inversion following SVD truncation.

null-space projection did not produce dramatic improvements. Although the systematic overestimation of the ϵ and δ parameters within the high-velocity layer after regularization of the q parameters (Figure 3b) is no longer present, the projected ϵ and δ solutions (Figure 8) have little interpretive value. To glean maximum insight into the results of anisotropic inversion, we would suggest inverting with alternative parameterizations, and we certainly recommend simultaneous viewing of the results in different parameterizations together with the traveltime residual mappings.

DISCUSSION

One concern with the null-space projection method is its conservative nature with respect to the traveltime residuals. Recall that any linear combination of the null-space singular vectors will have no effect on the modeled traveltimes. This transpires to be both a help and a hindrance; although improving efficiency by saving the need to reiterate at every step, we believe this to be too harsh a constraint, since the method is restricted to the particular realization of the traveltime residuals obtained from SVD truncation. Since this realization is no better than any other with similar rms data residual and incoherent pattern on the traveltime mapping, a preferable method to null-space projection would be restricted to keeping the same rms data error, but would allow flexibility of the modeled traveltimes to vary. The use of a larger generalized null space alleviates the constraint problem somewhat, by granting extra freedom as compared to the strictly defined null space.

Defining the null space of an inverse problem is simple when the problem is solved using SVD because the singular vectors that span the null space are explicitly generated. However, SVD becomes computationally unfeasible for larger tomographic problems than the ones used in this paper. We would refer the interested reader to Deal and Nolet (1996) who have discussed how to use the conjugate gradient method for applying the projection method to large-scale tomographic problems with acceptable computational effort.

text continues on p. 883

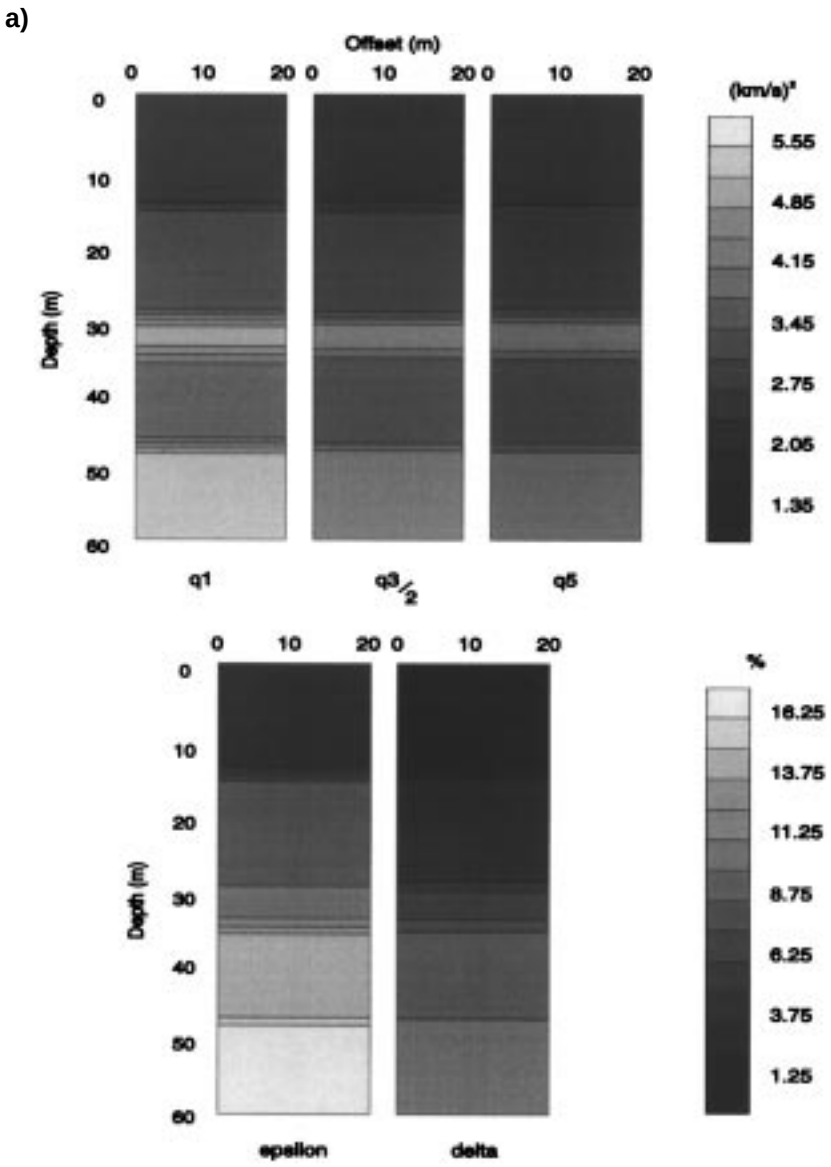

FIG. 6. (a) TIV plane-layered model constructed following regularized inversion and SVD truncation (Figures 3b, 5) of the field data.

b)

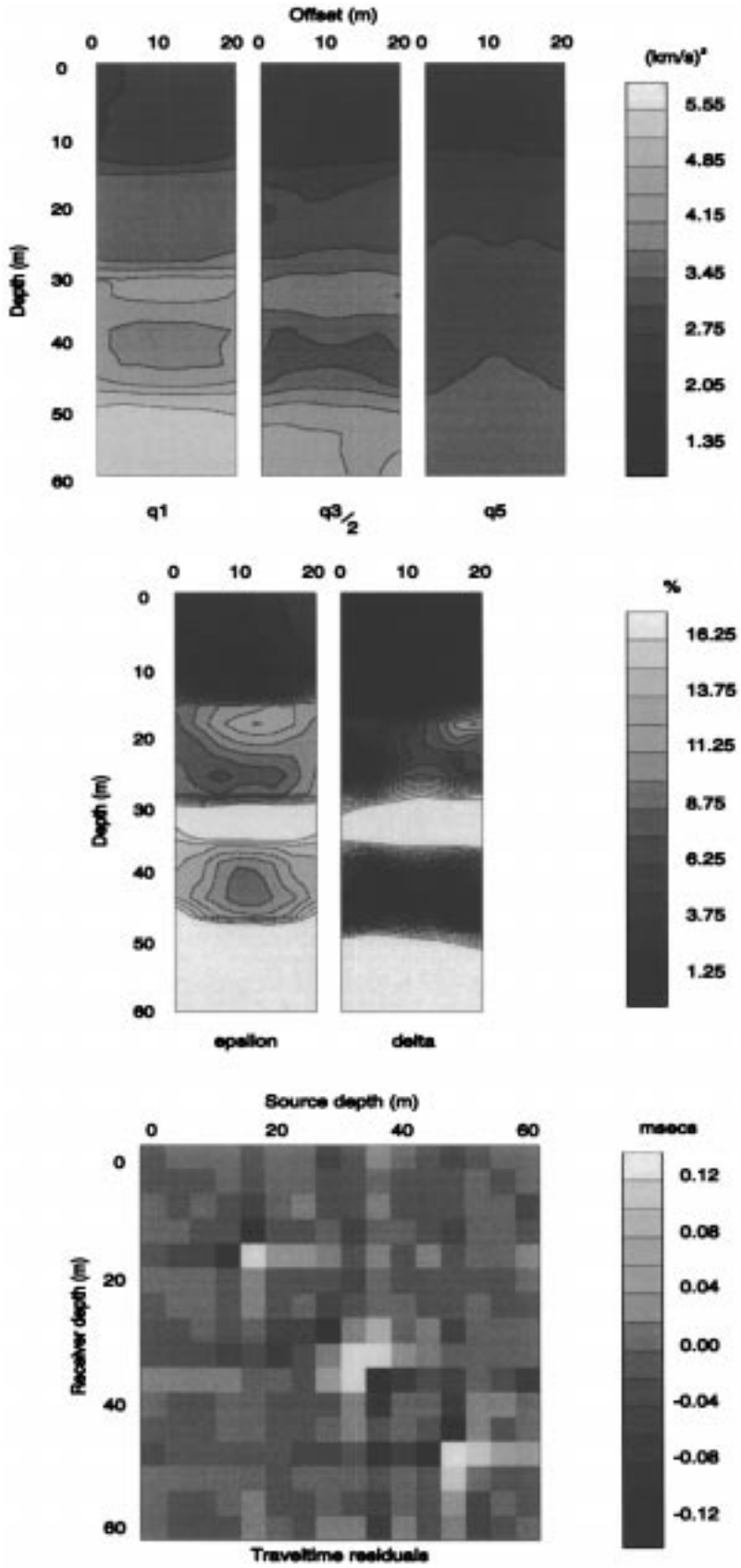

FIG. 6. (b) Synthetic data. Solution for regularized anisotropic inversion.

c)

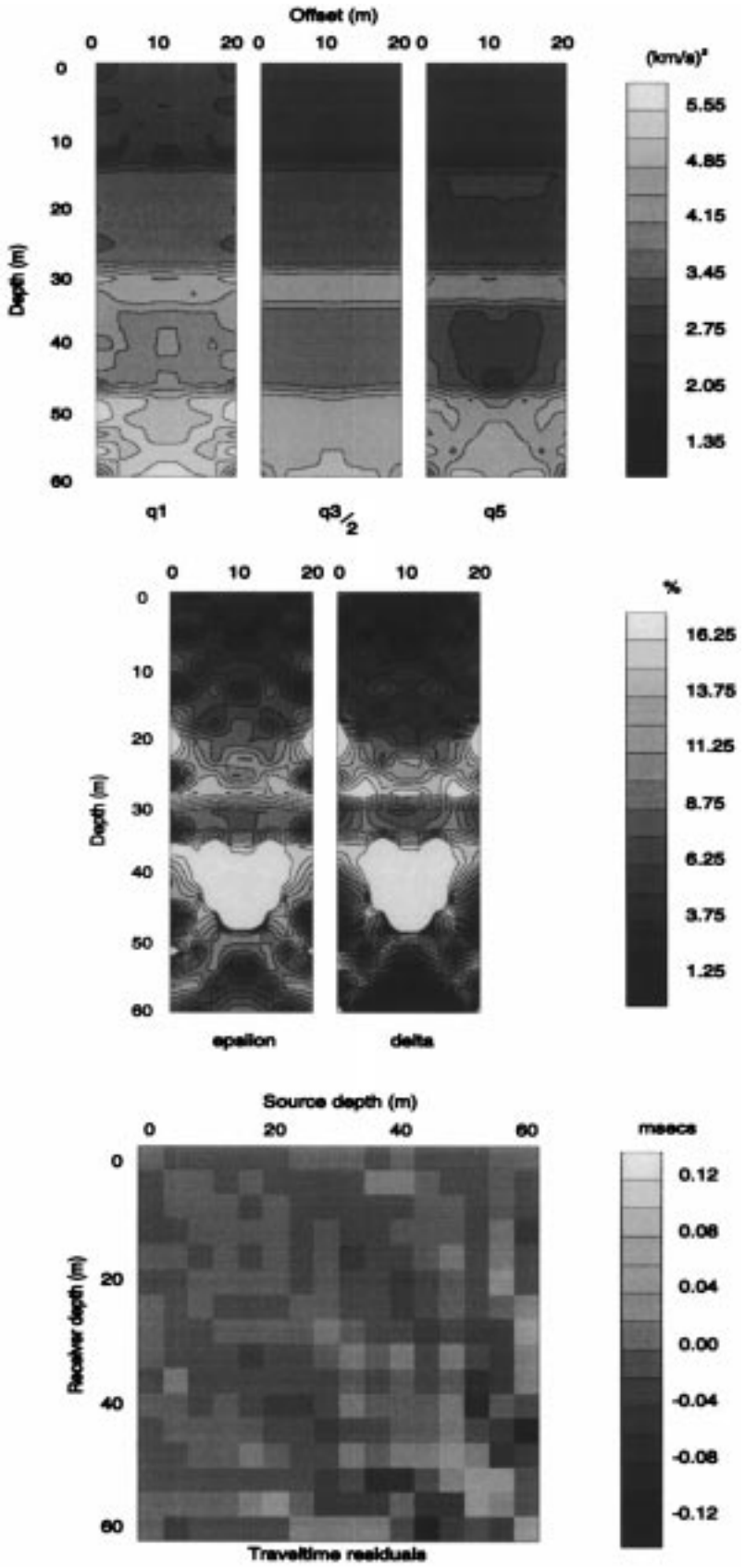

FIG. 6. (c) Synthetic data. Solution for anisotropic inversion following SVD truncation.

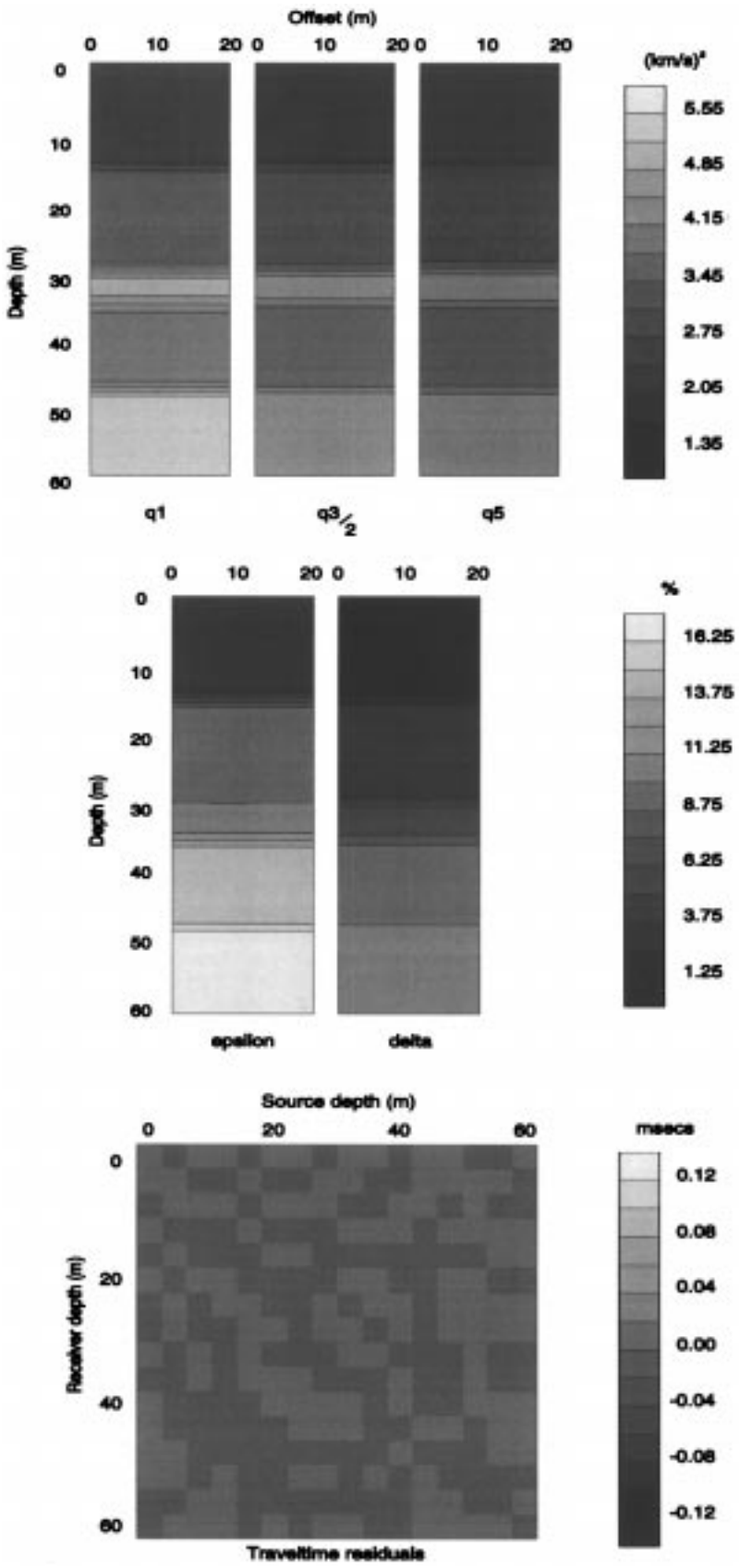

FIG. 7. Synthetic data. Improved solution following null-space projection of the TIV plane-layered model shown in Figure 6a.

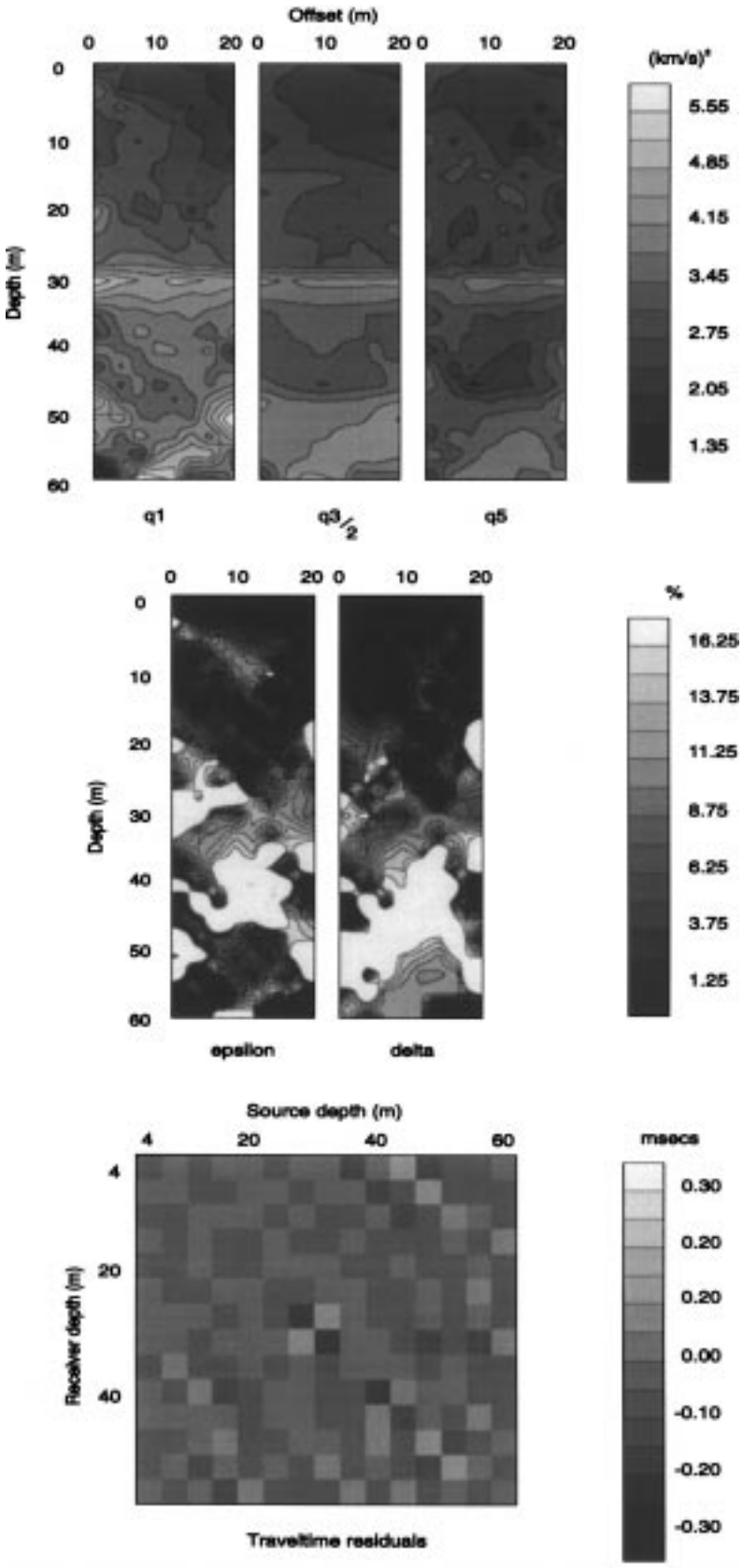

FIG. 8. Field data. Improved solution following null-space projection of the TIV plane-layered model shown in Figure 6a.

In the Introduction to this paper, we commented that the method proposed here bears similarities to that of initializing the tomographic model with the desirable solution. The latter scheme could function iteratively, with one iteration's desirable solution, produced after a posteriori filtering, becoming the starting model for the next. In this manner, the traveltime residuals would be guaranteed to decrease, and were an iterative process being used anyway, such as for curved ray tomography, this would be a more sensible approach than null-space projection. What we see as the main advantage of the null-space projection method is that of speed, since reiteration is not required. It is a trivial matter to null-space project any number of possible solutions to ascertain the allowable features in the solution, whereas considerable computer effort would be required to run an inversion with each as a starting model.

Although in this paper we have only discussed the null space in terms of projection of solutions, we note that more powerful uses of the null space are possible, such as global search mechanisms over the contributions from each null-space model singular vector for optimizing defined solution characteristics. That lies beyond the scope of the present paper.

CONCLUSIONS

Methods that rely on least-squares regularization are inadequate for high-resolution traveltime inversion, since they provide smoothed estimates of the velocity structure. We have demonstrated the use of the null space following singular-value decomposition and the incorporation of a priori information and a posteriori filtering for better resolving sharp discontinuities. We have shown the necessity of introducing anisotropy into inversion methods and have shown how we may achieve a compatible spatial resolution of each anisotropy parameter.

ACKNOWLEDGMENTS

This work was supported by the European Commission in the framework of the JOULE programme (JOU2-CT93-0321). We would like to thank the reviewers and associate editor for their clear and incisive comments.

REFERENCES

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

Bregman, N. D., Bailey, R. C., and Chapman, C. H., 1989a, Ghosts in tomography: the effects of poor angular coverage in 2-D seismic traveltime inversion: Can. J. of Expl. Geophys., **25**, 7–27.

——— 1989b, Crosshole seismic tomography: Geophysics, **54**, 200–215.

Carrion, P., 1991, Dual tomography for imaging complex structures: Geophysics, **56**, 1395–1404.

Carrion, P., Costa, J., Pinheiro, J. E. F., and Schoenberg, M., 1992, Cross-borehole tomography in anisotropic media: Geophysics, **57**, 1194–1198.

Carrion, P., Jacovitti, G., and Neri, A., 1993, Gaussian and non-Gaussian tomographic modelling via simulated annealing: J. Seis. Expl., **2**, 189–204.

Chapman, C. H., and Pratt, R. G., 1992, Traveltime tomography in anisotropic media–I: Theory: Geophys. J. Internat., **109**, 1–19.

Chiu, S. K. L., and Stewart, R. R., 1987, Tomographic determination of 3-D seismic velocity structure using well logs, vertical seismic profiles, and surface seismic data: Geophysics, **52**, 1085–1098.

Deal, M. M., and Nolet, G., 1993, Conservative filtering of tomographic images: Am. Geophys. Union Fall Meeting, 419.

——— 1996, Null space shuttles: Geophys. J. Internat., **124**, 372–380.

Delprat-Jannaud, F., and Lailly, P., 1992, What information on the earth model do reflection travel times provide?: J. Geophys. Res., **97**, 19827–19844.

Dines, K. A., and Lytle, R. J., 1979, Computerized geophysical tomography: Proc. Inst. Electr. Electron. Eng., **67**, 1065–1073.

Ellis, R. G., and Oldenburg, D. W., 1994, Applied geophysical inversion: Geophys. J. Internat., **116**, 5–11.

Gersztenkorn, A., and Scales, J. A., 1988, Smoothing seismic tomograms with alpha-trimmed means: Geophys. J. Internat., **92**, 67–72.

Hyndman, D. W., Harris, J. M., and Gorelick, S. M., 1994, Coupled seismic and tracer test inversion for aquifer property characterization: Water Resources Research, **30**, 1965–1977.

Iyer, H. M., and Hirahara, K., 1993, Seismic tomography: Theory and practice: Chapman and Hall.

Lanczos, C., 1961, Linear differential operators: Van Nostrand.

Lendzionowski, V., Walden, A. T., and White, R. E., 1990, Seismic character mapping over reservoir intervals: Geophys. Prosp., **38**, 951–969.

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

Menke, W., 1984, Geophysical data analysis: Discrete inverse theory: Academic Press Inc.

Michelena, R. J., 1992, Traveltime tomography in azimuthally anisotropic media: 62nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 757–761.

Michelena, R. J., Muir, F., and Harris, J. M., 1993, Anisotropic traveltime tomography: Geophys. Prosp., **41**, 381–412.

Neri, A., Carrion, P., Jacovitti, G., and Vesnaver, A., 1993, Tomographic reconstruction from incomplete data set with deterministic and stochastic constraints: SPIE, **2033**, 22–33.

Nolet, G., 1987, Seismic tomography with application to global seismology and exploration geophysics: D. Reidel Publ. Co.

Ory, J., and Pratt, R. G., 1995, Are our parameter estimators biased? The significance of finite-difference regularization operators: Inverse Problems, **11**, 397–424.

Pratt, R. G., and Chapman, C. H., 1992, Traveltime tomography in anisotropic media–II. Application: Geophys. J. Internat., **109**, 20–37.

Pratt, R. G., McGaughey, W. J., and Chapman, C. H., 1993, Anisotropic velocity tomography: A case study in a near surface rockmass: Geophysics, **58**, 1748–1763.

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

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.

Tarantola, A., and Nercessian, A., 1984, Three-dimensional inversion without blocks: Geophys. J. Roy. Astr. Soc., **76**, 299–306.

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

Walden, A. T., and Hosken, J. W. J., 1986, The nature of the non-Gaussianity of primary reflection coefficients and its significance for deconvolution: Geophys. Prosp., **34**, 1038–1066.

Williamson, P. R., 1993, Anisotropic crosshole tomography in layered media. Part II: Applications, results and conclusions: J. Seis. Expl., **2**, 223–238.

Worthington, M. H., 1984, An introduction to geophysical tomography: First Break, **2**, No. 11, 20–26.

Zhang, J., and McMechan, G. A., 1995, Estimation of resolution and covariance for large matrix inversions: Geophys. J. Internat., **121**, 409–426.

Zhang, Y. C., Shen, L. C., and Liu, C., 1994, Inversion of induction logs based on maximum flatness, maximum oil, and minimum oil algorithms: Geophysics, **59**, 1320–1326.