Elastic wavefield inversion of reflection and transmission data

Peter Mora

ABSTRACT

Elastic inversion of seismic data can be performed by finding the earth parameters (P- and S-wave velocities and density) that minimize the square error between the wavefield computed using this model and the observed wavefield (Tarantola (1984) and Mora (1986a)). The resultant model is a maximum probability solution provided the assumptions implied using least squares, namely Gaussian probability distributions of the data and model parameters, are valid. This is not a bad assumption considering the elastic inversion algorithm assumes the elastic wave equation and therefore accounts for S-waves, mode conversions, head waves and rayleigh waves etc. that are normally considered as coherent noise. As the number of shot profiles used in an inversion is increased, the signal to noise ratio increases and there are more illumination angles of the seismic waves on the subsurface. More illumination angles implies a more complete picture of the subsurface can be obtained. Even using a single shot profile, Mora (1986a) showed with some synthetic studies that a reasonable resolution can be achieved (both spatially and between the P- and S-wave velocities). When many shots are used to invert surface seismic data (reflection data), the result of the inversion is comparable to the expected result a prestack elastic shot profile depth migration, namely, a high frequency image of P- and S-wave velocities and density. However, compared to migration, artifacts are smaller, the P and S-wave velocity images are better resolved from one another, the magnitudes of the velocity and density perturbations have significance in an absolute sense and there is a slight increase in the lower frequency components of the image making this result easier to interpret (see also Mora (1986a)). Synthetic studies show that when signal to noise ratios are high, that an inversion of just a few shots

profiles results in a *complete* image of the subsurface (whereas conventional processing (stack and migration) would have required at least an order of magnitude more data (Ronen (1985) comes to the same conclusion)). In comparison to the case of inversion of reflection data where only the high frequency part of the model can be resolved, inversion of transmission data yields a result containing *both high and low frequencies* (and hence blockiness in the model due to layering and other gross features). Transmission data (such as offset VSP and/or well to well data) contains direct P- and S-waves that are most affected by the low frequency blocky velocity perturbations which cause significant delays to these direct waves. If multiple shots are used and hence the direct waves have illuminated the subsurface at many different angles, then there is enough redundancy to resolve these gross features (this is comparable to elastic diffraction tomography, see Devaney (1984) for the case of linearized acoustic diffraction tomography). Synthetic studies show that even using offset VSP's from only two wells located on either side of a region of interest, it is possible to obtain an inversion result that is almost identical to the true solution.

INTRODUCTION

In a previous paper (Mora, 1986a) I motivated the use of the elastic wave equation for the treatment of seismic data and gave arguments for inversion rather than conventional processing. An algorithm was derived (based on the scheme of Tarantola (1984)) which performed inversion of shot profiles for P- and S-wave velocities and density (or impedances). The inversion attempted to find an earth model that minimized the difference between the synthetic data computed by performing elastic modeling using this earth model and the observed wavefield. Therefore, it is a complete wavefield inversion scheme and attempts to take into account all elastic wave events such as reflected waves, converted waves, shear events, head waves, multiples, rayleigh waves and so on (note that isotropy will be assumed in this paper so anisotropic effects, if present in field data, will be treated as coherent noise by this algorithm). The previous paper (Mora 1986a) demonstrated the use of the algorithm for the inversion of a single shot profile of reflection seismic data. It illustrated the spatial resolution and resolution between the different model parameters (P- and S-wave velocities and density) and that the low frequency components of the model (the blockiness of layers) only creeps in gradually as the iterations of the inversion proceed.

This paper concentrates on some synthetic studies using several shots and contrasts the results of inverting reflection and transmission data. Reflection data resolves mainly P- and S-wave impedance changes (because that is what causes reflections) while

transmission data resolves the large scale velocity variations such as blockiness. The reason is that these large scale (or low frequency) velocity variations have the biggest effect on the direct waves which typically constitute most of the energy in transmitted wavefields. In particular, the large scale (low frequency) velocity anomalies delay the direct waves and if the subsurface was illuminated with direct waves at many angles, these large scale anomalies will be resolved (see also Gauthier et al. (1985) for the acoustic case) (c.f. diffraction tomography (see Devaney for a linearized acoustic example)). Some approaches to obtain the low frequencies in the model when direct waves are not present in the data set are suggested by Mora (1986a) and Kolb (1986).

INVERSION ALGORITHM

Because of the size of the elastic inverse problem and slowness of doing elastic forward modeling, it is impossible to use the most general methods that avoid local minima in the objective function such as Monte Carlo methods. Therefore, a compromise is the use of the iterative least squares optimization methods that assume Gaussian distributed model and data spaces. In the case of elastic inversion of seismic data the model space is the P- and S-wave velocities and densities at every location in the zone of interest and the data space is the recorded seismic data. The Newton algorithm of least squares in infeasible in two (or three) dimensional elastic inversions using present day computers so gradient type iterative least squares algorithms must be relied on. The seismic problem is nonlinear and so the Polak conjugate gradient algorithm is preferred (see Powell (1981) or Luenberger (1984)). This reverts to the simple gradient algorithm when the nonlinearity is too strong and most conjugate gradient algorithms would make poorer choices of update directions. In the following description I will use $\mathbf{d}$ to denote the data vector (shot profiles, usually multi-component) and $\mathbf{m}$ for the model vector (P- and S-wave velocities and densities). It will be assumed that the joint posteriori probability function is given by

$$P(\mathbf{d},\mathbf{m}) \quad = \quad constant \ \exp - \frac{1}{2}\left(\Delta\mathbf{d}^* \, \mathbf{C}_\mathbf{d}^{-1} \Delta\mathbf{d} + \Delta\mathbf{m}^* \, \mathbf{C}_\mathbf{m}^{-1} \Delta\mathbf{m} \right) \ , \qquad (1)$$

where $\Delta\mathbf{d} = \mathbf{d} - \mathbf{d}_0 = \mathbf{d}(\mathbf{m}) - \mathbf{d}_0$ is the data error vector corresponding to the earth model parameters $\mathbf{m}$ and data observations $\mathbf{d}_0$, $\Delta\mathbf{m} = \mathbf{m} - \mathbf{m}_0$ the model perturbation vector measured relative to the a priori model $\mathbf{m}_0$ and $\mathbf{C}_\mathbf{d}$ and $\mathbf{C}_\mathbf{m}$ are the covariance matrices for data and model spaces respectively. Note that * indicates conjugate transpose, normally the data consists of real values so $^* \ = \ ^T$, unless the inversion is carried out in a complex space (such as the Fourier space). The maximum probability solution can be obtained by minimizing the square error functional

$$S(\mathbf{d},\mathbf{m}) \;=\; \left(\Delta\mathbf{d}^{*}\,\mathbf{C}_{\mathrm{d}}^{-1}\Delta\mathbf{d} \;+\; \Delta\mathbf{m}^{*}\,\mathbf{C}_{\mathrm{m}}^{-1}\Delta\mathbf{m}\right) \;. \tag{1b}$$

This may be achieved for nonlinear functions $\mathbf{d}(\mathbf{m})$ such as is the case for elastic waves by using the preconditioned conjugate gradient algorithm (with the Polak choice for the conjugate direction) as follows

$$\textit{for } n \;=\; 1 \;\textit{to}\; \infty \;\left\{\right.$$

$$\mathbf{d}_n \;=\; \mathbf{d}(\mathbf{m}_n) \qquad , \quad \textit{data calculation}$$

$$\Delta\mathbf{d}_n \;=\mathbf{d}_n - \mathbf{d}_0 \;,\; \Delta\mathbf{m}_n \;=\mathbf{m}_n - \mathbf{m}_0 \qquad , \quad \textit{compute residuals}$$

$$S(\mathbf{d},\mathbf{m}) = \Delta\mathbf{d}^{*}\,\mathbf{C}_{\mathrm{d}}^{-1}\Delta\mathbf{d} + \Delta\mathbf{m}_n^{*}\mathbf{C}_{\mathrm{m}}^{-1}\Delta\mathbf{m}_n \qquad , \quad \textit{square error functional}$$

$$\textit{exit if converged} \qquad , \quad \textit{convergence test}$$

$$\mathbf{g}_n \;= \mathbf{D}_n^{*}\mathbf{C}_{\mathrm{d}}^{-1}\Delta\mathbf{d} + \mathbf{C}_{\mathrm{m}}^{-1}\Delta\mathbf{m}_n \qquad , \quad \textit{gradient}$$

$$\mathbf{p}_n \;= P(\mathbf{g}_n) \;\approx\; \mathbf{C}_{\mathrm{m}}\mathbf{g}_n \qquad , \quad \textit{preconditioning}$$

$$\mathbf{c}_n \;= \mathbf{p}_n + \frac{\mathbf{p}_n^{*}\left(\mathbf{g}_n - \mathbf{g}_{n-1}\right)}{\mathbf{p}_n^{*}\mathbf{g}_n}\mathbf{c}_{n-1} \;,\; \mathbf{c}_1 = \mathbf{p}_1 \qquad , \quad \textit{conjugate direction}$$

$$\hat{\eta}_n \;= \frac{\mathbf{c}_n^{*}\mathbf{g}_n}{\mathbf{c}_n^{*}\mathbf{D}_n^{*}\mathbf{C}_{\mathrm{d}}^{-1}\mathbf{D}_n\,\mathbf{c}_n + \mathbf{c}_n^{*}\mathbf{C}_{\mathrm{m}}^{-1}\mathbf{c}_n} \qquad , \quad \begin{array}{l}\textit{calculate approximate steplength}\\ \textit{by assuming linearity}\end{array}$$

$$\eta_n \;= \min_{\eta_n} S(\mathbf{d}+\eta_n\,\mathbf{D}\mathbf{c}_n\,,\mathbf{m}) \qquad , \quad \begin{array}{l}\textit{optimize steplength via linesearch}\\ \textit{using } \hat{\eta}_n \textit{ as a starting point}\end{array}$$

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \eta_n\,\mathbf{c}_n \qquad , \quad \textit{update model}$$

$$\left.\right\} \;. \tag{2}$$

The reason the iteration number n starts at 1 is because I used subscripts 0 for the a priori model $\mathbf{m}_0$ and field data $\mathbf{d}_0$. Note that $\mathbf{g}$ defines the gradient vector, $\mathbf{D} = \partial\mathbf{d}/\partial\mathbf{m}$ is the Frechet derivative matrix, $\mathbf{p}$ is the modified gradient direction (the preconditioned direction), $\mathbf{c}$ is the Polak conjugate direction, $\hat{\eta}$ is the linearized steplength and η is the optimal steplength obtained by a linesearch (necessary because the $\hat{\eta}$ is calculated assuming linearity but the seismic problem is nonlinear, however, $\hat{\eta}$ gives a good value from which to commence the line search). This algorithm is described in detail in relation to

the elastic seismic inverse problem by Mora (1986a) (for general background on conjugate gradients see Luenberger (1984) and for least squares theory see Menke (1984) or Tarantola et al. (1982)). The Frechet derivative matrix is never actually required by this algorithm but only a knowledge of how to compute the operation of it on some model vector and the operation of its conjugate transpose on some data vector. The operation of $\mathbf{D}$ corresponds to the linearized forward problem and can be computed by any linear or nonlinear forward modeling scheme $\mathbf{d}(\mathbf{m})$ by observing that

$$\mathbf{d}(\mathbf{m}') = \mathbf{d}(\mathbf{m}) + \mathbf{D}(\mathbf{m}' - \mathbf{m}) = \mathbf{d}(\mathbf{m}) + \mathbf{D}\delta\mathbf{m} \quad , \tag{3}$$

for small $\delta\mathbf{m}$. After some algebra it becomes evident (see Mora (1986a) and Tarantola (1984)) that the adjoint operation (the operation of $\mathbf{D}^*$) required in order to compute the gradient direction can also be calculated using a forward simulation. From Mora (1986a) the equation for the gradient in terms of the P- and S-wave velocities and density, $\mathbf{g} = \left[\delta\hat{\hat{\alpha}} \, , \, \delta\hat{\hat{\beta}} \, , \, \delta\hat{\hat{\rho}}\right]^T$ is

$$\delta\hat{\hat{\alpha}} = 2\rho\alpha\delta\hat{\lambda} + \mathbf{C}_{\alpha\alpha}^{-1}(\alpha - \alpha_0) + \cdots \quad , \tag{4a}$$

$$\delta\hat{\hat{\beta}} = -4\rho\beta\delta\hat{\lambda} + 2\rho\beta\delta\hat{\mu} + \mathbf{C}_{\beta\beta}^{-1}(\beta - \beta_0) + \cdots \quad , \tag{4b}$$

$$\delta\hat{\hat{\rho}}_\alpha = (\alpha^2 - 2\beta^2)\delta\hat{\lambda} + \beta^2\delta\hat{\mu} + \delta\hat{\rho} + \mathbf{C}_{\rho\rho}^{-1}(\rho - \rho_0) + \cdots \quad . \tag{4c}$$

where $\delta\hat{\lambda}$, $\delta\hat{\mu}$ and $\delta\hat{\rho}$ are the adjoint operation with respect to the Lamé parameters and density given by

$$\delta\hat{\gamma} = \sum_S \int dt \, \mathbf{C}_{\mathbf{d}}^{-1}(t)\Big(\Omega^\gamma u_j(\mathbf{x},t)\Big)\Big(\Omega^\gamma \psi_j(\mathbf{x},t)\Big) \quad , \tag{5a}$$

where $\gamma = \lambda$, μ or ρ and Ω^γ is an operator to be defined shortly. The notation used here is consistent with the previous paper (Mora, 1986a) which gives a derivation of the gradient. Also, subscript α of $\delta\hat{\hat{\rho}}_\alpha$ in equation (4c) indicates that this term is the gradient with respect to density when the P- and S-wave velocities are used as the other model parameters. The $+ \cdots$ at the end of each line in equations (4) is used to allow for the possibility of covariances between the different model parameters such as $\mathbf{C}_{\alpha\beta}(\alpha - \alpha_0)$. In equation (5a), u_j represents the j-th component of displacement of a seismic wavefield, and ψ is the *back propagated residual wavefield* given by

$$\psi_j(\mathbf{x},t) = \sum_R \mathbf{C}_{\mathbf{d}}^{-1}(\mathbf{x}_R) G_{ij}(\mathbf{x}, -t \, ; \mathbf{x}_R \, , 0) * \Delta u_i(\mathbf{x}_R \, , t) \quad , \tag{5b}$$

where Δu_i is error in the displacement wavefield at the current iteration given by

$$\Delta u_i = u_i - u_{i_0}$$

and u_{i_0} is the observed displacement wavefield. Note that $\Omega^\gamma = \Omega^\gamma_{ijk}$ is an operator that depends on the model parameter γ and can be considered as a model parameter *unraveling* operator ($\gamma = \lambda$, μ, or ρ). Also, note that I use an unusual convention in this equation to avoid complicated notation with subscripts, namely that the implied summations within $\left(\right)$'s must be carried out first. From Mora (1986a), the unraveling operator Ω^γ is given by

$$\Omega^\rho = i\,\partial_t \quad , \quad \Omega^\lambda = i\,\partial_j \quad , \quad \Omega^\mu = \frac{i}{\sqrt{2}}\Big(\delta_{jk}\,\partial_i + \delta_{ji}\,\partial_k \Big) \quad . \tag{6}$$

The inversion formulae, equations (2), (4), (5) and (6), are derived and described by Mora (1986a) including a comparison with the the process of migration (see also Tarantola (1984) for the original derivation in terms of the Lamé parameters and density).

The above equations have assumed that the data covariance function can be represented as

$$\mathbf{C}_\mathbf{d}^{-1}(\mathbf{x}_R\,,t\,) \;=\; \mathbf{C}_\mathbf{d}^{-1}(\mathbf{x}_R\,)\mathbf{C}_\mathbf{d}^{-1}(t\,) \quad . \tag{7a}$$

In the examples in this paper I specifically used diagonal forms

$$\mathbf{C}_\mathbf{d}^{-1}(t\,) \;=\; \frac{t^{\,2p}}{\sigma_\mathbf{d}^2} \tag{7b}$$

and

$$\mathbf{C}_\mathbf{d}^{-1}(\mathbf{x}_R\,) \;=\; \mathit{taper}(\mathbf{x}_R\,) \quad . \tag{7c}$$

where $\mathit{taper}(\mathbf{x}_R\,)$ gradually tapers the data at the edges of the receiver array to zero in order to decrease artifacts caused by edge effects.

EXAMPLES

Introduction

The following are some inversions of synthetic reflection and transmission data sets to demonstrate what information the algorithm can gain about the deeper geology (non near surface geology) when applied to *reflection* and *transmission* seismic data. Therefore, considering rayleigh waves would resolve only the near surface they have been excluded from these calculations (i.e. an absorbing boundary condition was used rather than a free surface boundary condition at the earth's surface). Note that together with refracted waves, rayleigh waves should help resolve the near surface (and hence the solve statics problem). Therefore, it would be desirable to use a free surface condition when

an accurate inversion of the near surface is desired. However, because of the large range of velocity variation in the near surface (and therefore nonlinearity), convergence is expected to be slow and local minima may pose a problem. Another case when the free surface boundary condition should not be ignored is when there is a large range of angles (relative to the earth's surface) of the source to receiver raypaths at the source or receiver locations. This would mean that the free surface boundary condition plays an important role in the source directivity pattern and it is therefore necessary to obtain an accurate inversion. Often, near surface velocities are low so rays leaving the source are almost vertical and the effect of the free surface on the source (and receiver) directivity pattern is unimportant. Inversions of rayleigh and refracted waves will be the topic of a future paper. One further point is that the size of the velocity and density perturbations in these examples was deliberately chosen to be small (only a few percent) in order that the inverse problem be more linear and hence the convergence be more rapid. Considering the large CPU requirements of the algorithm and CPU limitations when carrying out this work, rapid convergence was crucial. Therefore, a demonstration of the algorithm when large velocity and density perturbations (away from the a priori model) are present will have to await a future paper. Even though velocity perturbations are small, the following examples illustrate effectively what can be hoped to be gained by using this inversion procedure. In particular, they show how it is possible to do inversions when the background velocity model is known accurately (to within a few percent) which would often be the case in a region where detailed velocity analyses were carried out. The results also illustrate that S-wave velocity can be obtained as well as P-wave velocity in inversions of data sets containing S-wave events such as mode conversions, direct S-waves and S-S reflections as well as the usual P-wave events.

The synthetic data was calculated from the model shown in Figure 1 with elastic finite differences using a similar (faster) version of the algorithm of Kosloff et al. (1984) (see also Cerjan et al. (1985)). The algorithm is fully described by Mora (1986b). This model represents a complex geologic situation consisting of a horst structure and limestone reef complex amidst a multitude of larger layers each with laterally variable random fine layer structure within them. It is parametrized by the three isotropic elastic parameters (P- and S-wave velocity and density) on a square grid with a grid spacing of 20 meters. Nine shot simulations were carried out with shots spaced every .5 km starting at .16 km using a highly band limited source wavelet (a 4-th derivative of a Gaussian curve with a fundamental frequency of 20 hertz). The surface geophone array was fixed and covered the 4.34 km region of the model. There were also two wells at .4 km and 3.92 km in which VSP data was recorded for all these shots. Geophone spacing on the

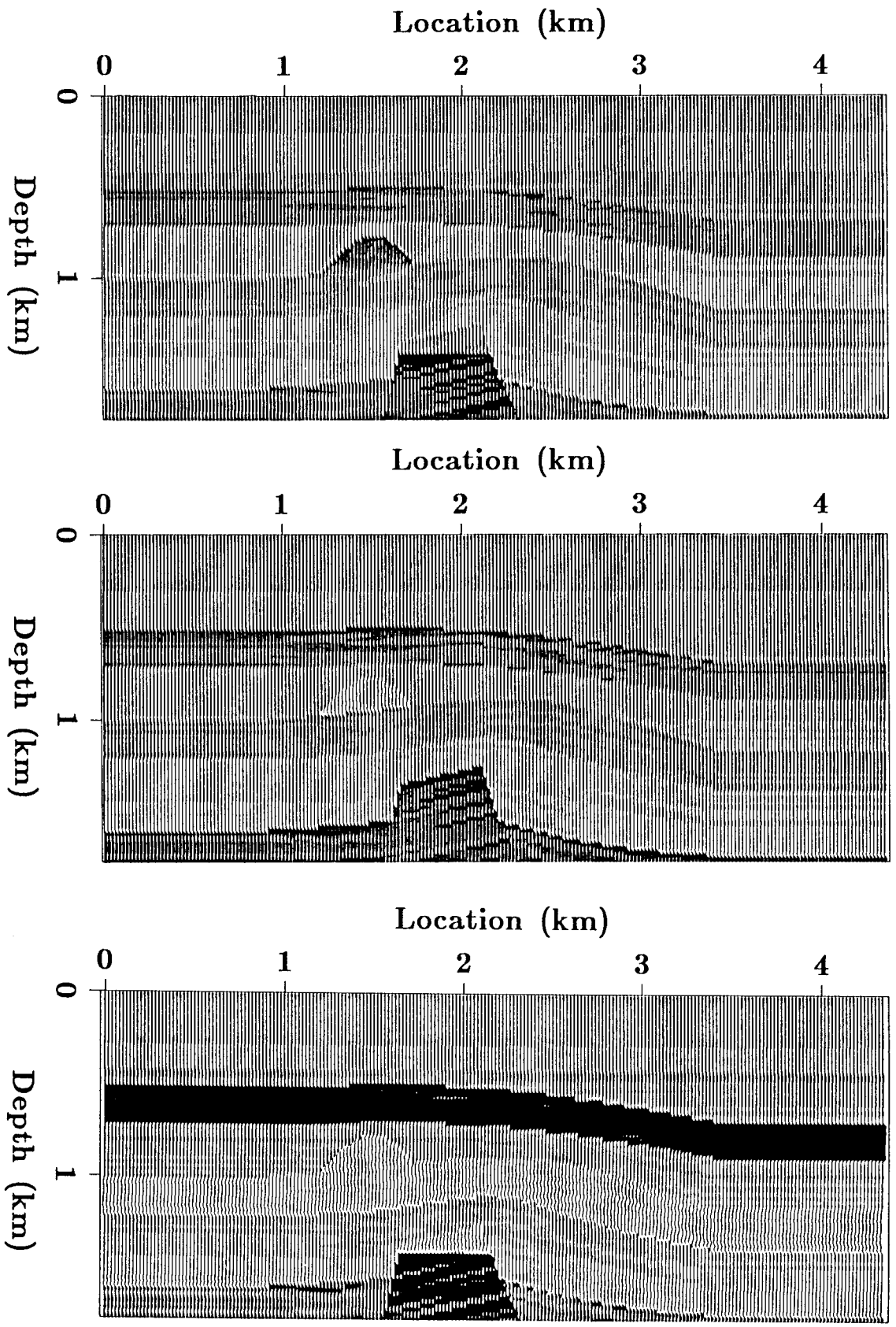

FIG. 1. Horst and reef model used to generate the synthetic data shown in Figures 2 and 3. (a) P-wave velocity, (b) S-wave velocity, and (c) density.

earths surface and down the wells was the same as the grid spacing. Two component receivers were used (P-SV type receivers) and a vertical force was applied. A representative shot profile due to a shot located at 2.16 km is shown in Figure 2 and a typical offset VSP recorded in the well at .4 km due to the shot at 2.66 km is shown in Figure 3. The inversion was carried out with the conjugate gradient algorithm (equations (2)) and using the gradient parametrized in P- and S-wave velocities and density denoted α, β and ρ. The gradient was calculated with equations (4) through (6). To concentrate on the resolvability of α, β and ρ, the source was assumed to be known and so was *not* varied but was fixed at its true value throughout the iterations of the inversions. The same method of elastic finite differences was used to compute *both* the forward modeled wavefield u_j *and* the back propagated residual wavefield ψ_j. Hence, the results will illustrate that the algorithm works in ideal circumstances but testing is certainly required with real seismic data to fully evaluate the method. There may be many factors in field situations that will effect the results of inversions and must be taken into account if the algorithm is to yield useful results (for instance, geophone couplings). Equations (7) were used to define the data covariance matrix where the value of p was .5 (this is equivalent to saying that a time varying gain of $t^{.5}$ was applied to allow for wave divergence in 2D and the edges of the data were tapered to reduce finite aperture artifacts). The data variance $\sigma_{\mathbf{d}}$ was set to a small value so the inversion was not heavily damped and hence the solution was not required to stay close to the a priori model $\mathbf{m}_0$ (i.e. little heed was paid in the c.g. algorithm (equations (2)) to $\Delta\mathbf{m}_n{}^*\mathbf{C}_{\mathbf{m}}^{-1}\Delta\mathbf{m}_n$, $\mathbf{C}_{\mathbf{m}}^{-1}\Delta\mathbf{m}$ or $\mathbf{c}_n{}^*\mathbf{C}_{\mathbf{m}}^{-1}\mathbf{c}_n$). A preconditioning of $\mathbf{C}_{\mathbf{m}}$ was applied in algorithm (2) where constant diagonal model covariance matrices were used, $\mathbf{C}_{\alpha\alpha} = \alpha_{ave}^2\mathbf{I}$, $\mathbf{C}_{\beta\beta} = \beta_{ave}^2\mathbf{I}$ and $\mathbf{C}_{\rho\rho} = \rho_{ave}^2\mathbf{I}$ and cross-covariances between different model parameters were assumed to be zero. This preconditioning has the effect of removing unwanted scale factors between different model parameters due to their different physical units (and perturbation magnitudes). A further preconditioning was applied that spatially equalized the energy in the gradient $\mathbf{g}$ (similarly to the method of AGC but applied over the space axes (using a long averaging window) rather than the time axis). This corresponds to the assumption that the standard deviation of the earth parameters from the a priori model $\mathbf{m}_0$ does not vary spatially (except rapidly because of geologic/stratigraphic effects) and so the gradual variations are assumed to be caused by differences in the amount of seismic energy that can reach different parts of the model. For example, less seismic energy reaches the edges of the model than the middle. If a preconditioning were not applied to remove this effect, the convergence would be slower.

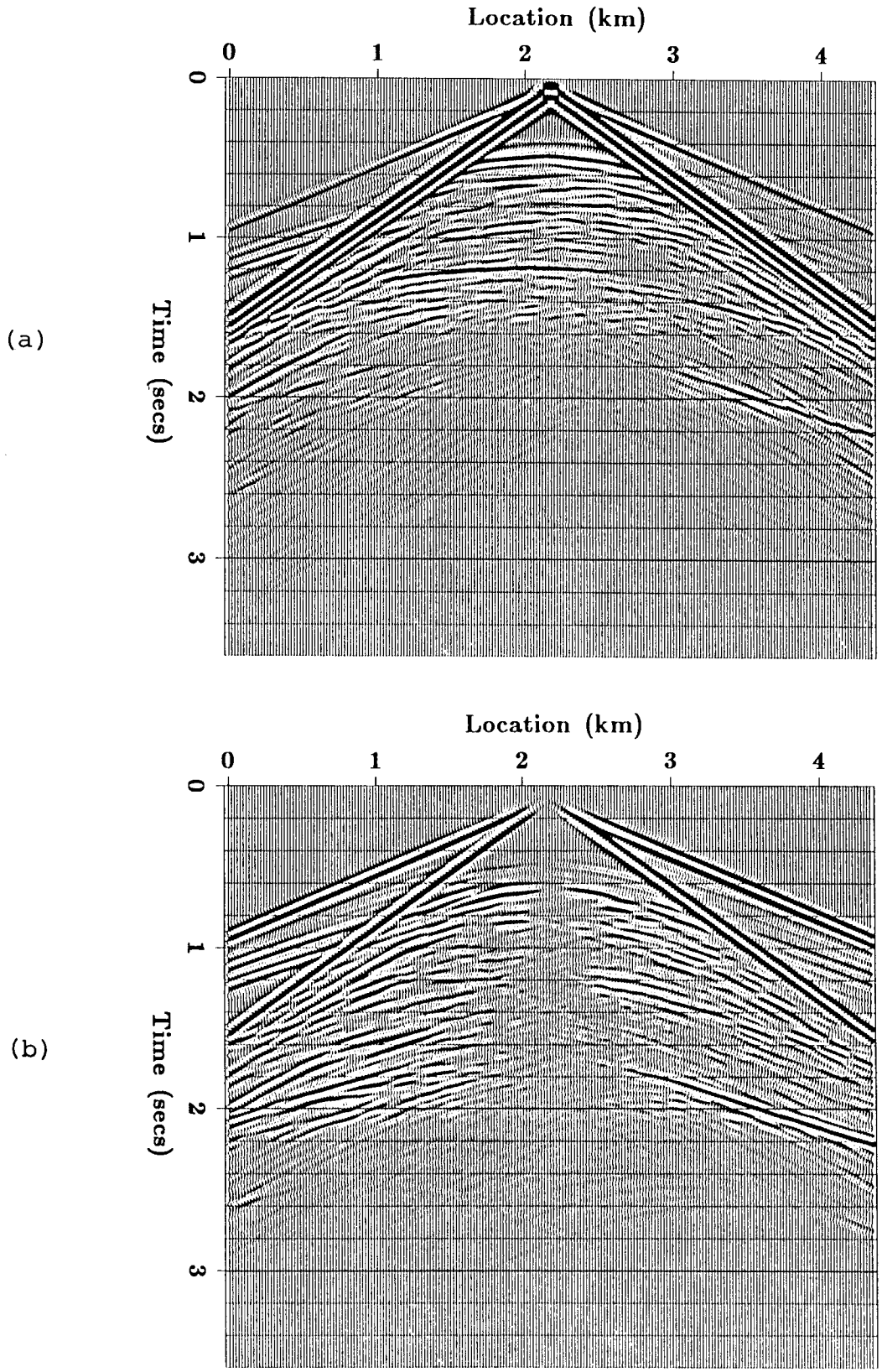

FIG. 2. One of the split spread shot profiles used in the inversion. It is due to a vertical source located at 2.16 km and was generated by elastic finite differences with the model shown in Figure 1. This synthetic data is noise free. (a) Vertical component, and (b) horizontal component.

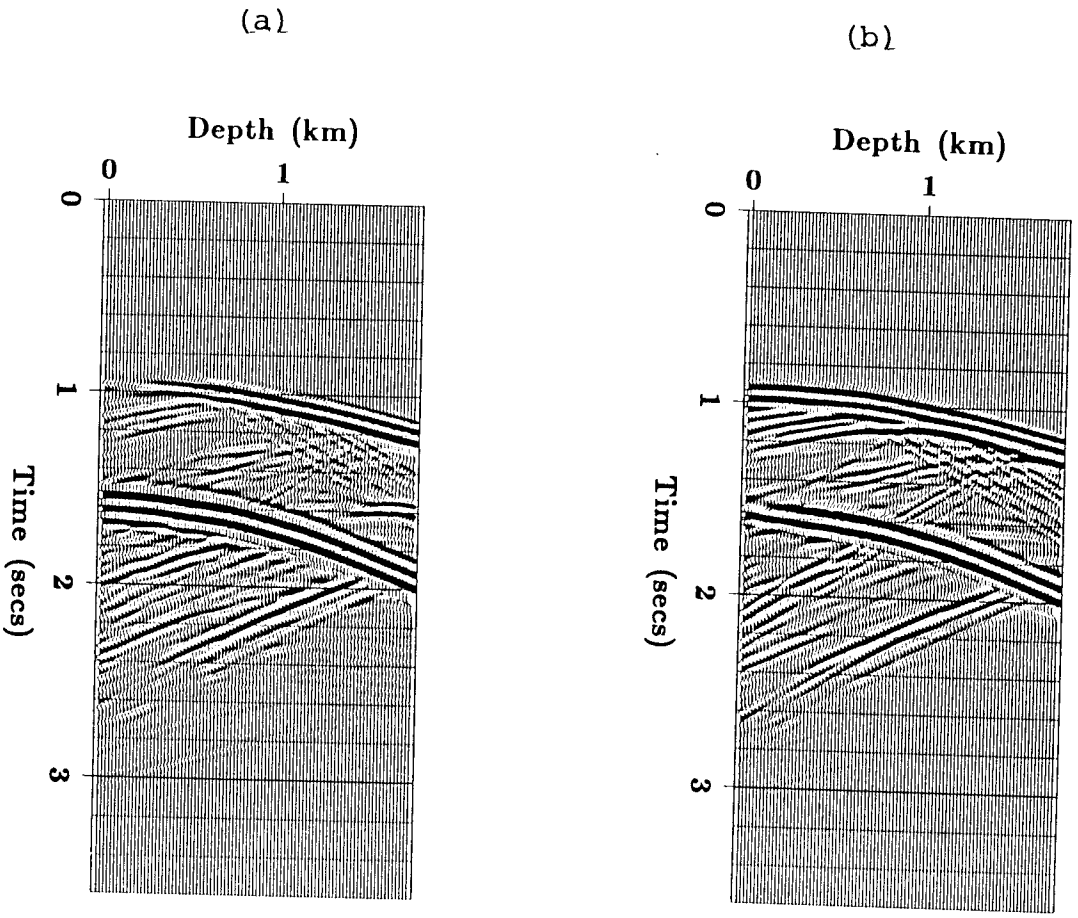

FIG. 3. One of the offset VSP's used in the inversion. It corresponds to a well located at .4 km and is due to the vertical source located at 2.66 km. It was generated by elastic finite differences with the model shown in Figure 1 and is noise free. (a) Vertical component, and (b) horizontal component.

Transmission example

All the data recorded for the nine shots was used simultaneously for an inversion. This included 9 two-component shot profiles, 9 two-component offset VSP's from the well at .4 km and 9 two-component offset VSP's from the well located at 3.92 km. This data set contains both reflected and transmitted waves. The reflections are produced by high frequency perturbations in the P- and S-wave velocities and density and so the reflections will help resolve the high frequency part of the model. By comparison, the transmitted waves (direct waves on the VSP's) are most affected by large scale P- and S-wave velocity anomalies that cause time delays and so they tend to help resolve these low frequency velocity anomalies. The true model is compared to the 1 and 10 iteration inversion result for the P-wave velocity model (Figure 4), S-wave velocity model (Figure 5) and density model (Figure 6). Notice that the 10 iteration results for the P- and S-wave velocities are close to the true model while the 10 iteration density result is only a high frequency approximation of the true density model. This is because the low frequency part of the P- and S-wave velocity model could be resolved by the traveltime delays in the direct waves while low frequency density fluctuations have little effect on the seismic waves and so cannot be resolved. A major effect of the iterations on the P- and S-wave velocity models is to make the solution more blocky and similar to the true model. Also, the S-wave velocity solution is much sharper after the 10 iterations. The reason is that initially the S-wave velocity is resolved mainly by the strong direct S-waves on the VSP data causing the solution to have too many low frequencies. After some iterations when the direct waves begin to be matched by the inversion algorithm, the reflections begin to have a greater effect and start to boost the high frequencies thereby sharpening the picture. Notice that just like in the true model, the flat spot at the top of the horst (gas water contact) can be seen on the P-wave velocity result and doesn't exist on the S-wave velocity result. This illustrates that there is a good resolution between P- and S-wave velocity. Also, observe that the reef can be seen in the inversion result and has the correct polarity. The magnitudes of the inversion results are better seen on the well logs shown in Figure 7. Here, the true model is shown at location 1.5 km (i.e. located on the reef) as well as the starting model, first iteration result and the 10 iteration result. Note that the magnitudes of the velocity perturbations have not quite built up to their true values after 10 iterations. Considering that the throughout the iterations, the magnitude of the velocity perturbation gradually increased (as the velocity perturbations become more blocky like the true model), it is expected that further iterations would cause the magnitude to build up to almost exactly the true value. However, considering that this is an inversion of ideal noise free synthetic data, it

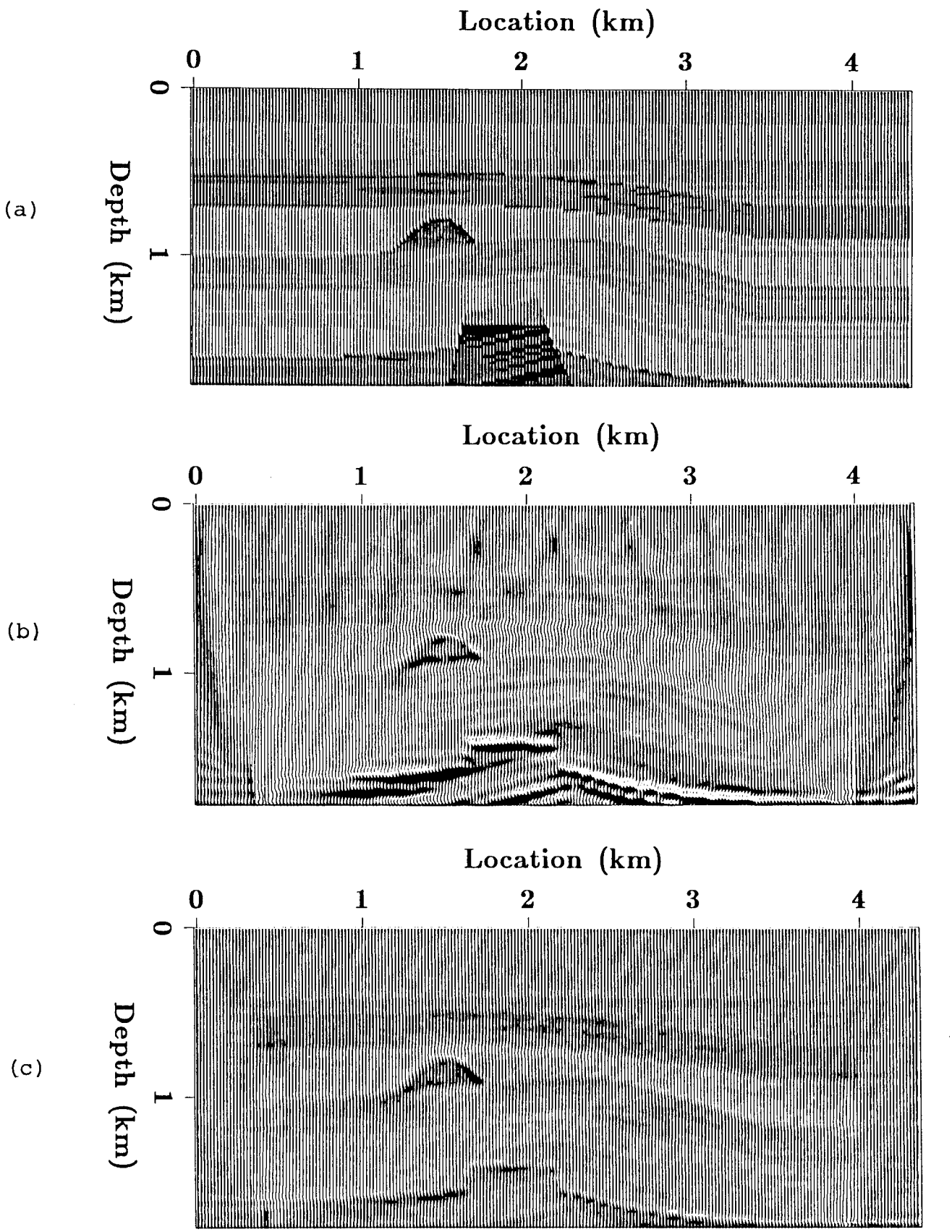

FIG. 4. Comparison of the true P-wave velocity model and inversion results at various iterations when using both reflection and transmission data (i.e. shot profiles and VSP's). Note that the plots are not at the same scale. (a) True P-wave velocity model, (b) P-wave velocity inversion result after 1 iteration, and (c) P-wave velocity inversion result after 10 iterations.

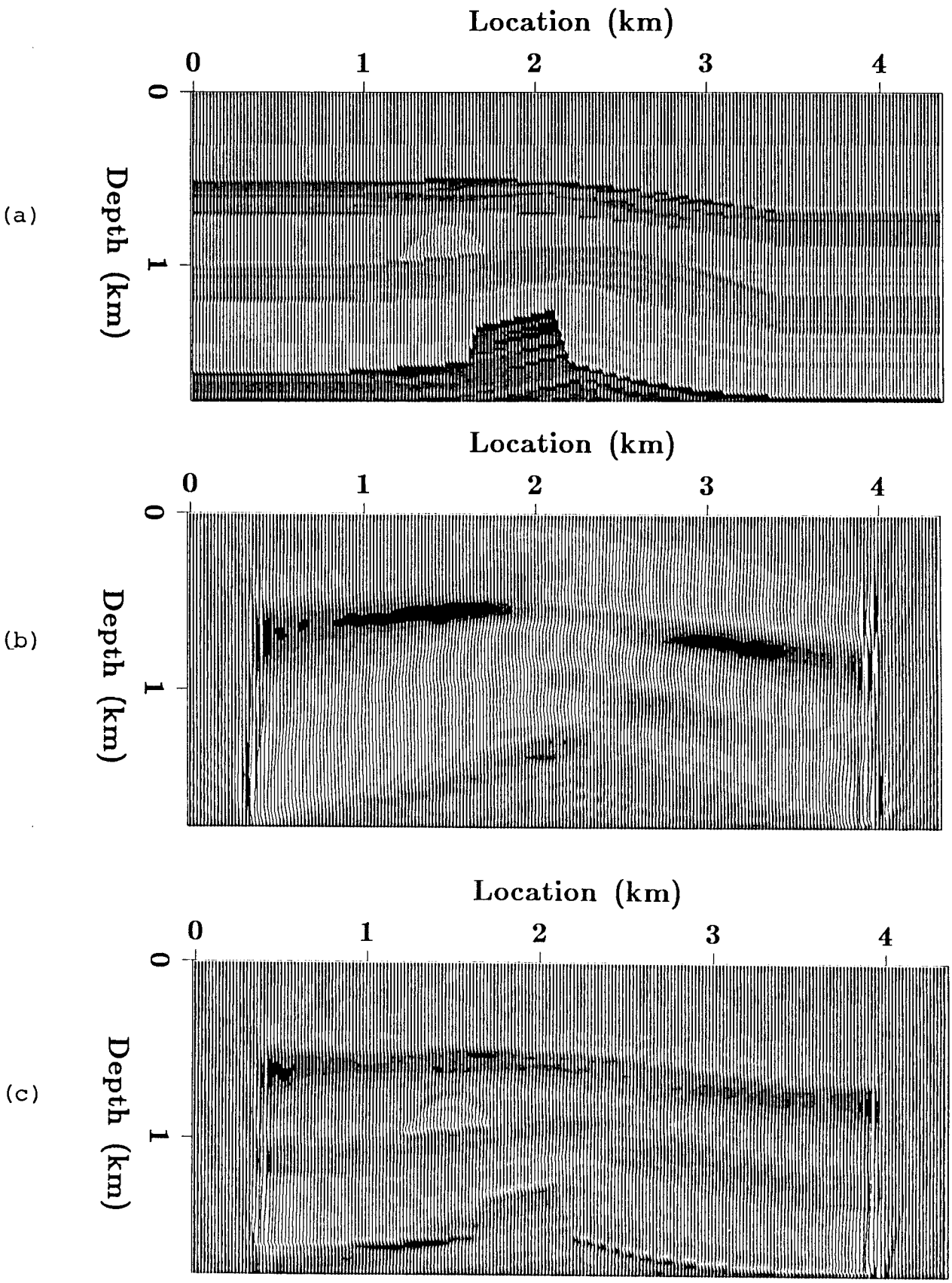

FIG. 5. Comparison of the true S-wave velocity model and inversion results at various iterations when using both reflection and transmission data (i.e. shot profiles and VSP's). Note that the plots are not at the same scale. (a) True S-wave velocity model, (b) S-wave velocity inversion result after 1 iteration, and (c) S-wave velocity inversion result after 10 iterations.

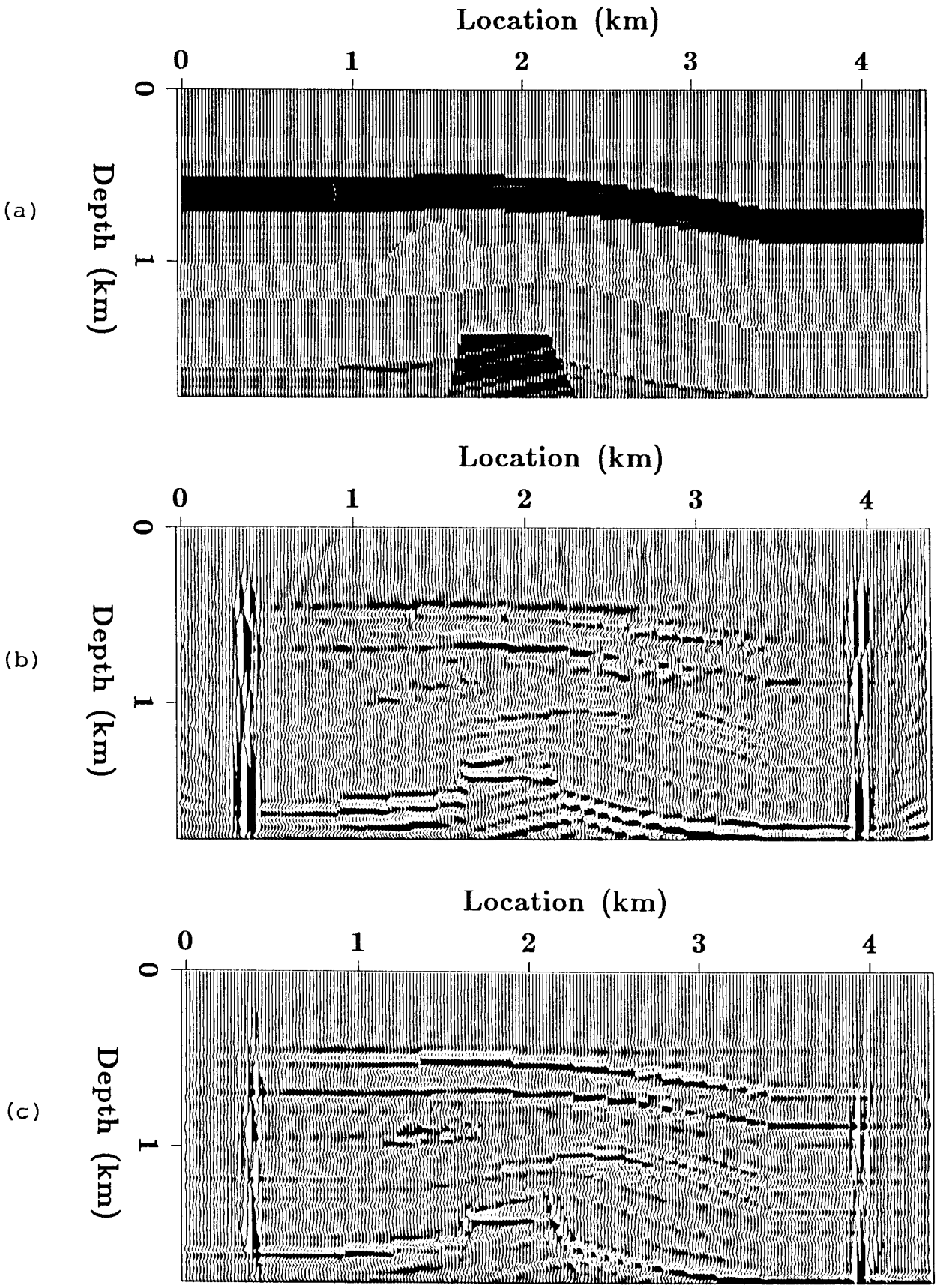

FIG. 6. Comparison of the true density model and inversion results at various iterations when using both reflection and transmission data (i.e. shot profiles and VSP's). Note that the plots are not at the same scale. (a) True density model, (b) density inversion result after 1 iteration, and (c) density inversion result after 10 iterations.

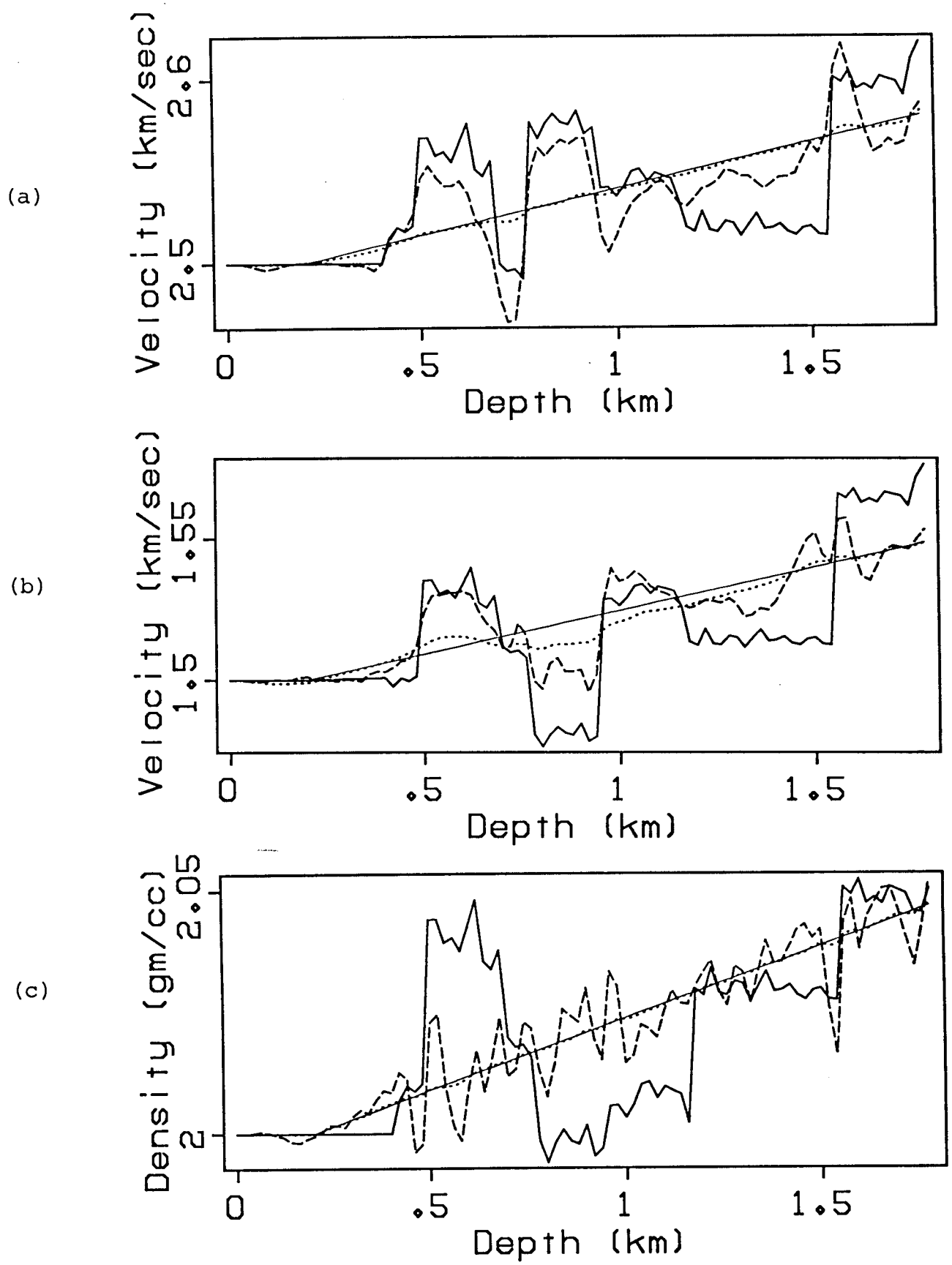

FIG. 7. Comparison of the true model and inversion results located at 1.5 km (i.e. well logs through the reef) when using both reflection and transmission data (i.e. shot profiles and VSP's). The true model is shown in the bold solid line, the starting model in the fine solid line, the result after one iteration is in the dotted line and the result after 10 iterations is in the broken line. (a) P-wave velocity, (b) S-wave velocity, and (c) density.

would be unrealistic to expect such perfect results when applying the technique to field data. Figures 8 and 9 show the data residual (unmatched part of the wavefield) after 10 iterations for the gather and VSP respectively shown in Figures 2 and 3. A graph of the square error sum as a function of iteration is shown in Figure 10 indicating that the amount of mismatched energy in the seismic wavefield steadily decreases as the least squares iterations proceed.

Reflection example

Figures 11 through 13 show the filtered true model and the 1 and 10 iteration inversion results when reflection data is used but *not* transmission (VSP) data. The true model was filtered with the seismic wavelet to make the comparison easier here where the inversion method could not resolve the low frequency (blocky) velocity variations. The high frequencies velocity and density models are well resolved by the the inversion and in particular, the P- and S-wave velocity models are well resolved between one another (for example, observe that the inversion correctly obtained that there is a gas-water contact (flat spot) at the top of the horst on the P-wave velocity result but not on the S-wave velocity result. The main effects of iterations are: (i) to decrease crosstalk between P- and S-wave velocity (notice that the flat spot is present in the 1 iteration S-wave velocity result but not in the 10 iteration result). Therefore, the iterations help improve resolution between the different model parameters, (ii) to decrease the level of noise on the S-wave velocity result (this is greater than the level of noise on the P-wave velocity result because the S-wave velocity is mainly resolved using high angle events such as mode converted waves and S-S reflections, and (iii) to decrease the strength of elliptical smiles (i.e. finite aperture artifacts similar to migration edge truncation artifacts). Note that the density is not well resolved from P-wave velocity as seen by fact that the reef is almost invisible in the density result (i.e. the true P-wave velocity perturbation in the reef was positive while the density perturbation was negative and since these both affected the density result some cancellation occurred). See also Mora (1986a) for more examples of poor resolution between P-wave velocity and density using the velocity-density parametrization of the gradient (equations (4) through (6)). cancellation). The inversion results can be seen in more detail on the well logs shown in Figure 14. Here, the true model is shown at location 1.5 km (i.e. located on the reef) as well as the starting model, first iteration result and the 10 iteration result. Note that the magnitudes of the velocity perturbations after 10 iterations are much smaller than the true velocity perturbations. Also, compared to the inversion results when VSP data was used shown in Figure 7, the magnitudes are smaller. This is because the reflection data alone

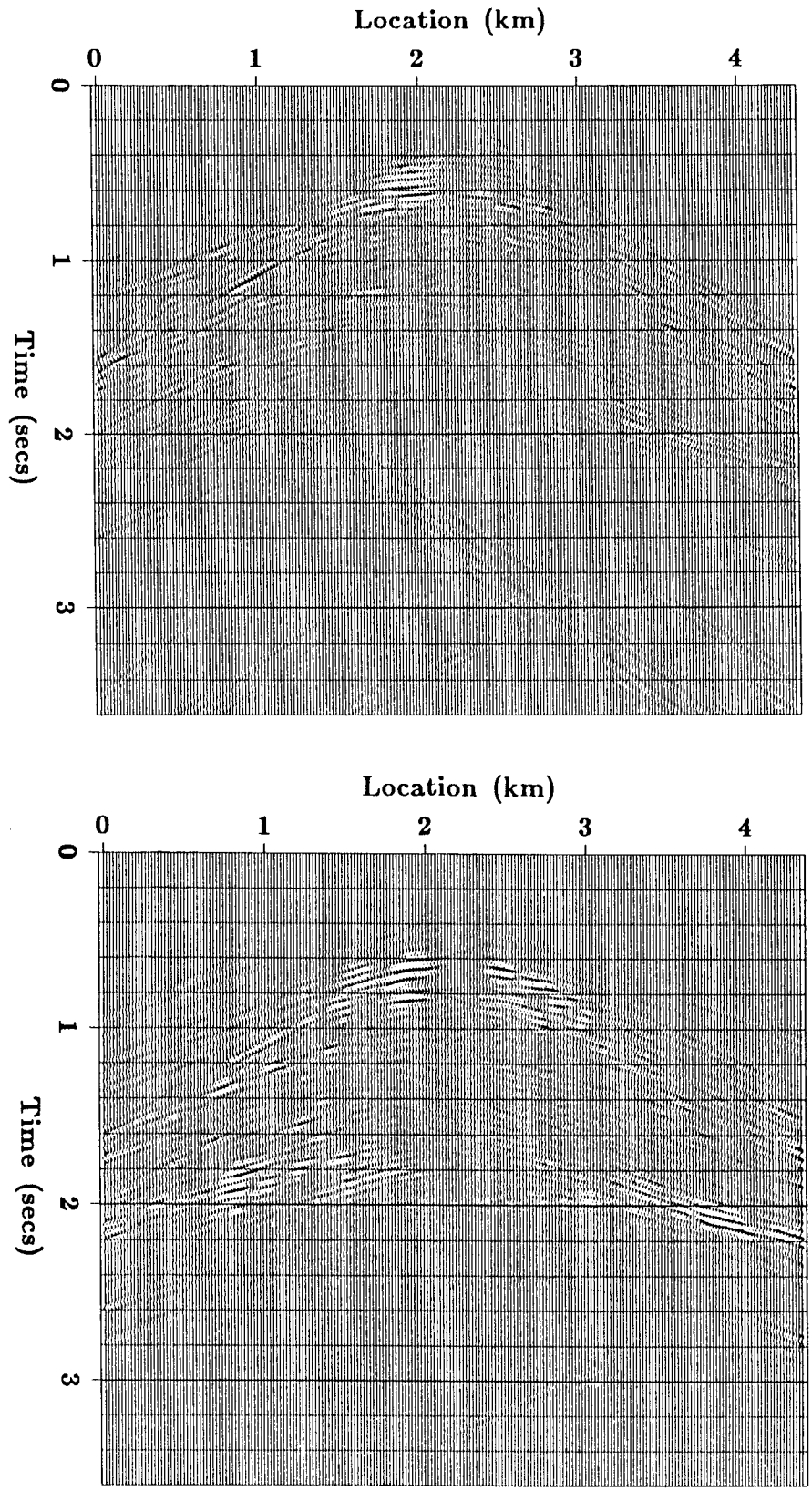

FIG. 8. Unmatched part of the shot profile of Figure 2 (i.e. the residual) after 10 iterations of the inversion algorithm when using both reflection and transmission data in the inversion (i.e. using both shot profiles and VSP's). It is plotted at the same scale as Figure 2. (a) Vertical component, and (b) horizontal component.

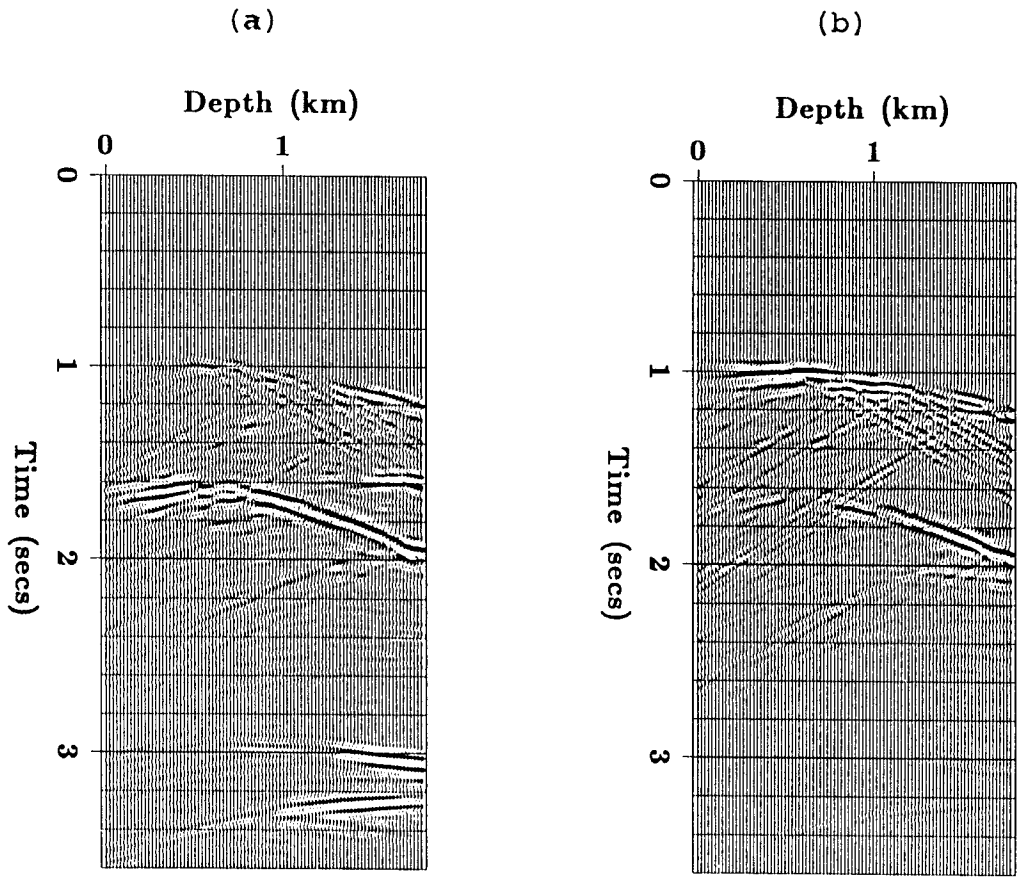

FIG. 9. Unmatched part of the VSP of Figure 3 (i.e. the residual) after 10 iterations of the inversion algorithm when using both reflection and transmission data in the inversion (i.e. shot profiles and VSP's). It is plotted at the same scale as Figure 3. (a) Vertical component, and (b) horizontal component.

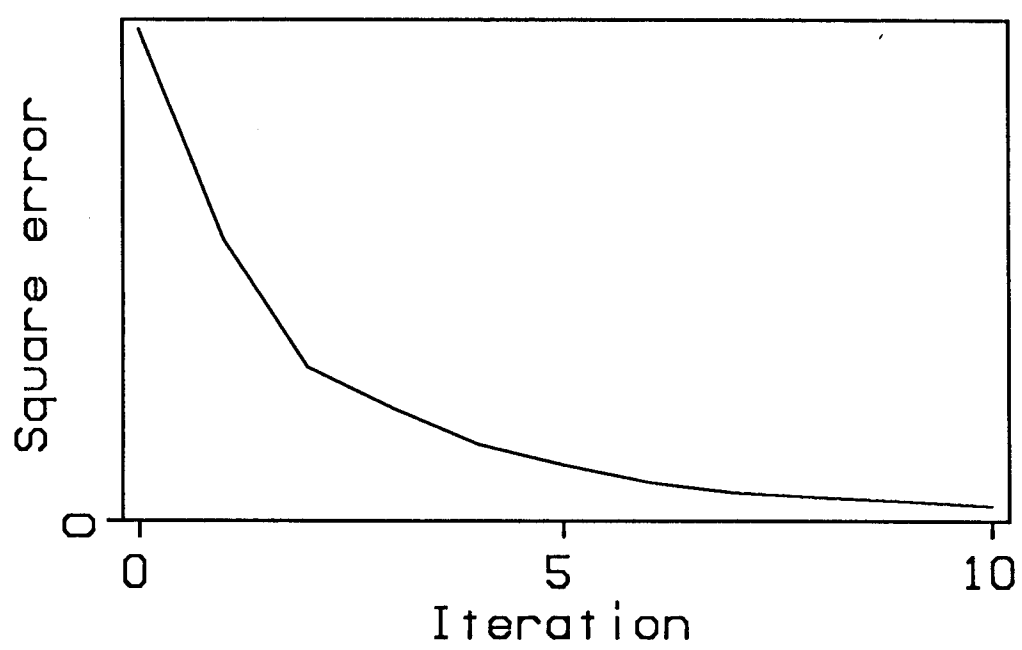

FIG. 10. The sum of square error as a function of iteration when both reflection and transmission data are used in the inversion (i.e. shot gathers and VSP's).

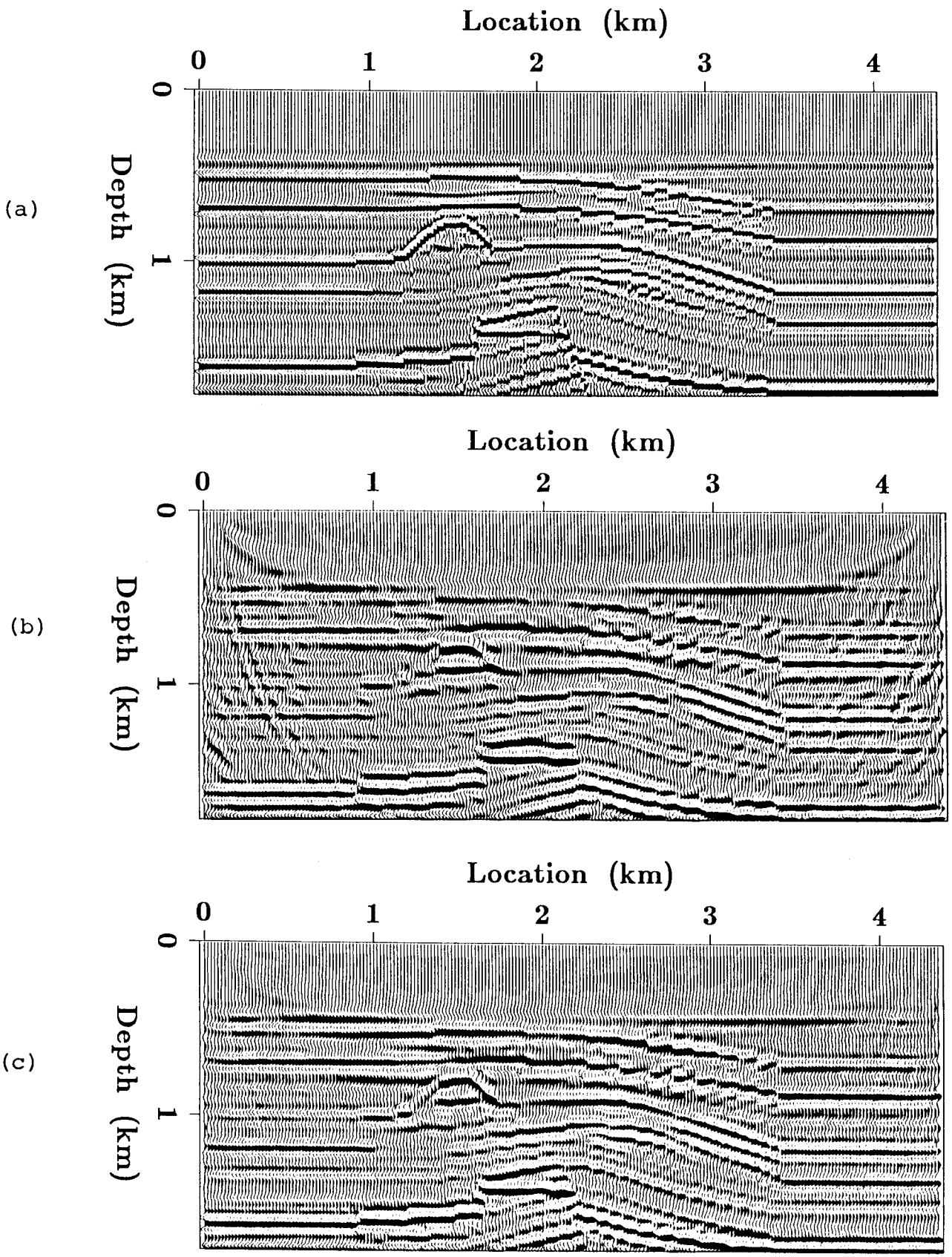

FIG. 11. Comparison of the true P-wave velocity model and inversion results at various iterations when using only reflection data (i.e. shot profiles only). Note that the plots are not at the same scale. (a) True P-wave velocity model, (b) P-wave velocity inversion result after 1 iteration, and (c) P-wave velocity inversion result after 10 iterations.

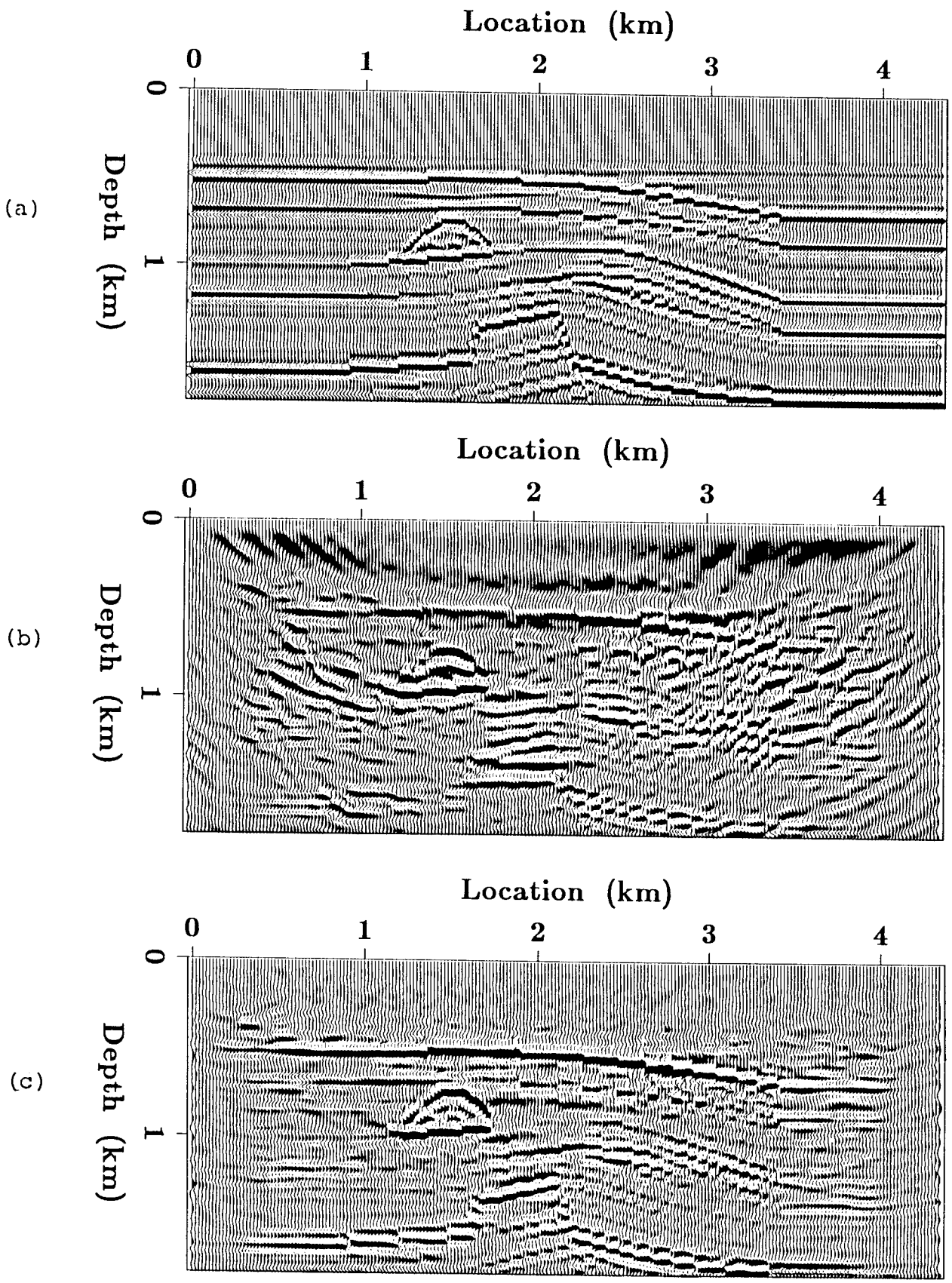

FIG. 12. Comparison of the true S-wave velocity model and inversion results at various iterations when using only reflection data (i.e. shot profiles only). Note that the plots are not at the same scale. (a) True S-wave velocity model, (b) S-wave velocity inversion result after 1 iteration, and (c) S-wave velocity inversion result after 10 iterations.

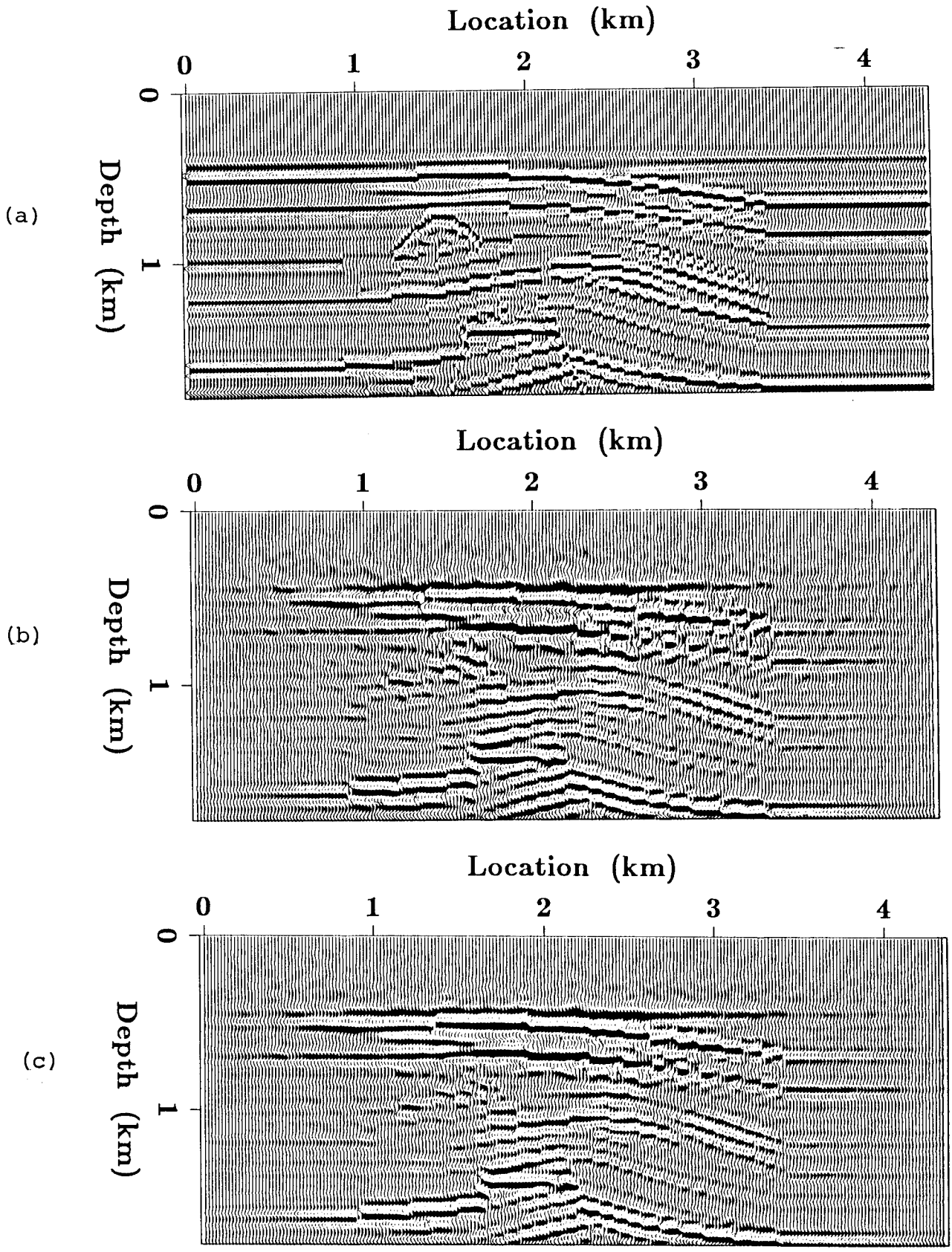

FIG. 13. Comparison of the true density model and inversion results at various iterations when using only reflection data (i.e. shot profiles only). Note that the plots are not at the same scale. (a) True density model, (b) density inversion result after 1 iteration, and (c) density inversion result after 10 iterations.

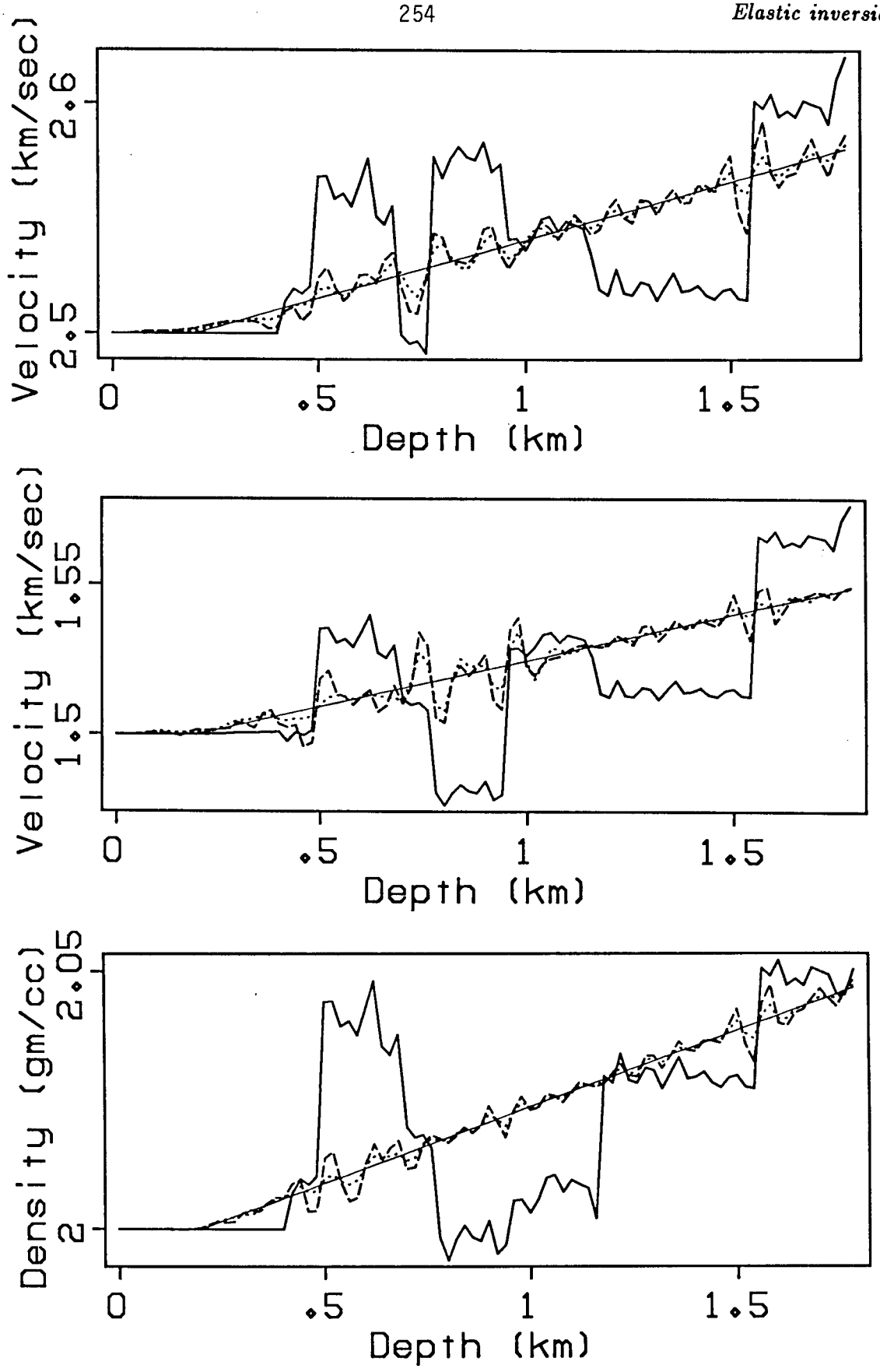

FIG. 14. Comparison of the true model and inversion results located at 1.5 km (i.e. well logs through the reef) when using only reflection data (i.e. shot profiles only). The true model is shown in the bold solid line, the starting model in the [illegible]

did not well resolve the low frequency P- and S-wave velocity models and hence kinematical errors exist even after 10 iterations. Therefore, it is impossible to perfectly match the data by simply increasing the magnitude high frequency model perturbations so the algorithm makes a compromise in order to decrease the sum of square error and finds smaller magnitude high frequency perturbations. Figure 15 shows the data residual (unmatched part of the wavefield) after 10 iterations for the gather shown in Figure 2. The residual is slightly larger than for the case where VSP data was also used in the inversion (see Figure 7). A graph of the square error sum as a function of iteration is shown in Figure 16 indicating that the amount of mismatched energy in the seismic wavefield steadily decreases as the least squares iterations proceed. However, it does not decrease by as much as the case when VSP data was also used in the inversion. This is because the VSP data helped resolve the blocky part of the model and hence obtained a more accurate kinematical model (than the starting guess) thereby allowing the error to be reduced by a greater degree. Note that as shown by Mora (1986a), the low frequencies slowly build up as iterations proceed even when only reflection data is used. This is illustrated in Figure 17 which shows the true amplitude spectrum of a P-wave velocity well log and the 1 and 10 iteration spectra. However, considering that there is only a minimal increase in the low frequency content of the spectra even after 10 iterations, it would be inefficient to try to obtain all the low frequencies by iterating further. Also, considering that the residuals shown in Figure 15 are small, it is expected that when noise is present not all the low frequencies would be well resolved no matter how many iterations were carried out. The problem is that the choice of the objective function leads to a particular way that the model solution is built up with iteration (i.e. the gradient direction depends on the objective function). When only reflection data is used in the inversion, the high frequency parts of the model are much better resolved than the low frequency parts and so appear rapidly in the early iterations. A way to solve this problem is by introducing the correct a priori information on the model parameters (namely that the velocity model is blocky) thereby leading to different objective functions and hence procedures that converge equally rapidly on all frequencies components of the model (see Mora (1986a) and Kolb et al. (1986)).

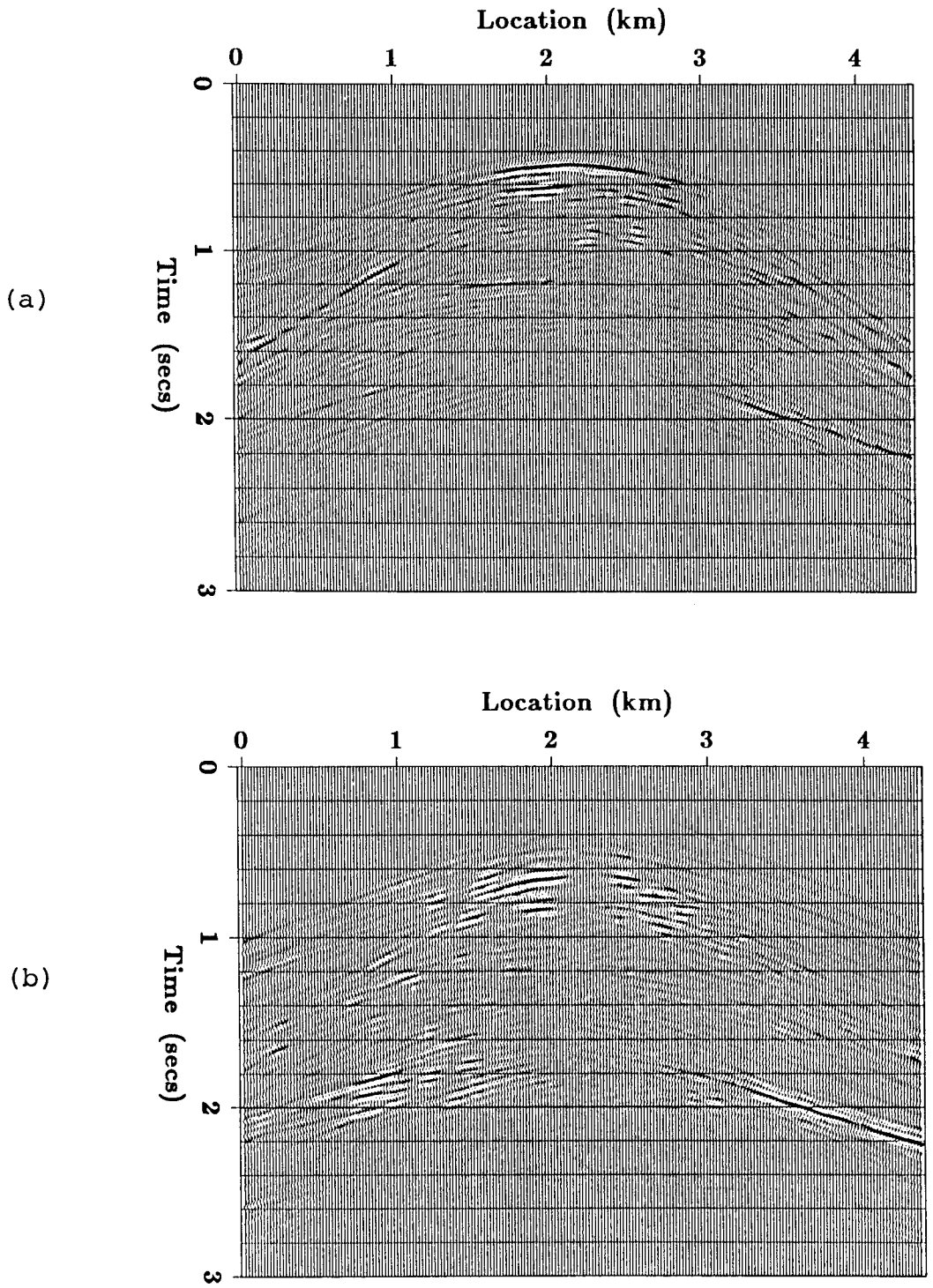

FIG. 15. Unmatched part of the shot profile of Figure 2 (i.e. the residual) after 10 iterations of the inversion algorithm when using only reflection data in the inversion (i.e. shot profiles only). It is plotted at the same scale as Figure 2. (a) Vertical component, and (b) horizontal component.

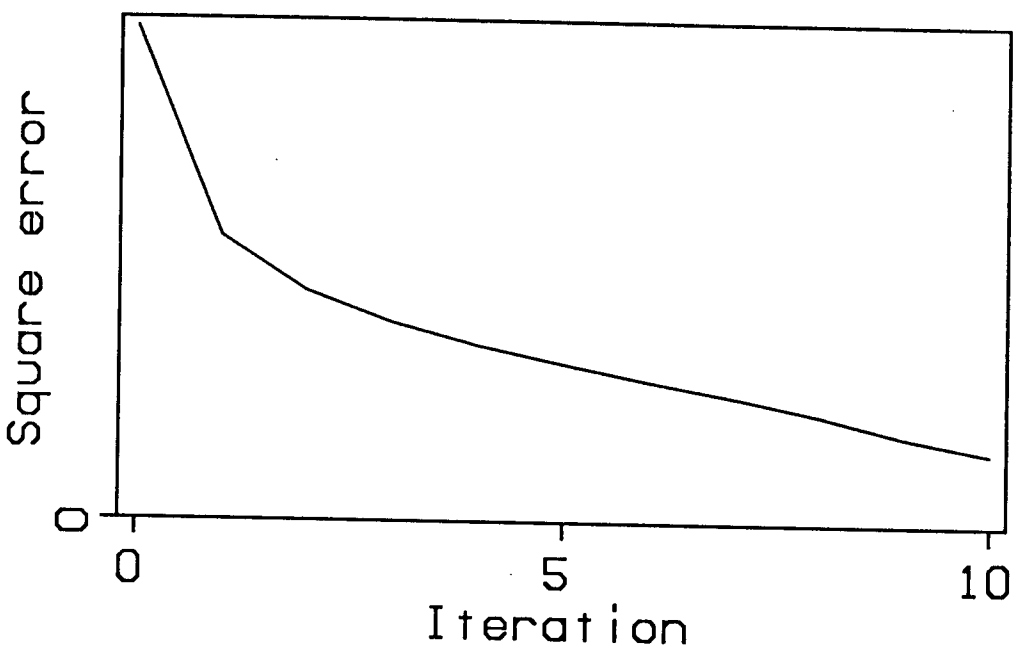

FIG. 16. The sum of square error as a function of iteration when only reflection data is used in the inversion (i.e. shot gathers only).

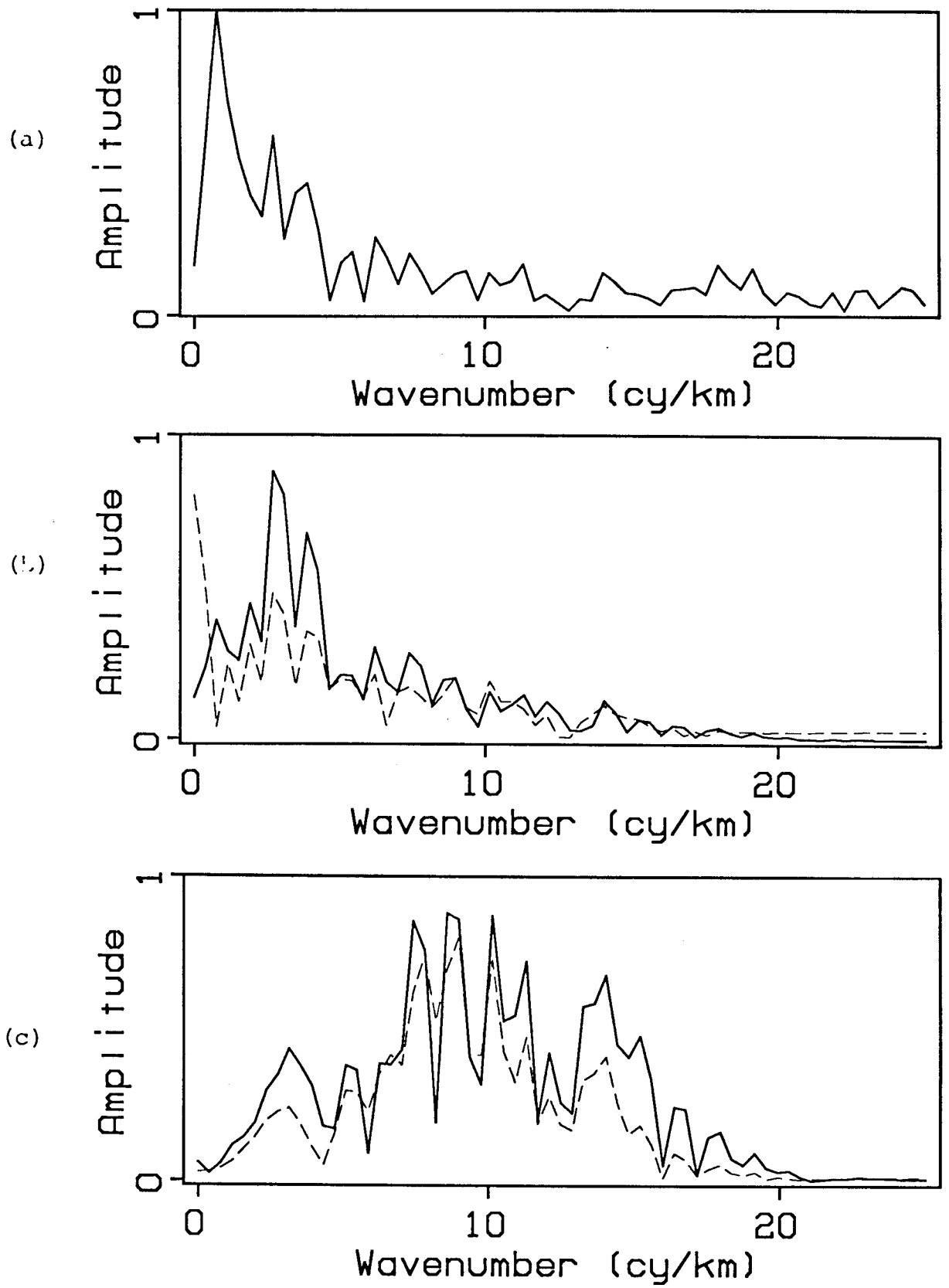

FIG. 17. Vertical wavenumber spectra of the P-wave velocity model located at 1.5 km (i.e. the spectrum of the P-wave velocity log at 1.5 km). (a) The spectrum of the true P-wave velocity log, (b) spectra of inversion results when both transmission and reflection data are used after 1 iteration (broken line) and 10 iterations (solid line), and (c) spectra of inversion results when only reflection data is used after 1 iteration (broken line) and 10 iterations (solid line).

CONCLUSIONS

Two-dimensional elastic inversion of reflection and transmission seismic data has been described and with some speedups it should be feasible though costly on todays super computers. It is based on matching the observed wavefield to a wavefield computed with an elastic modeling scheme. The advantage of this method over the conventional processing approach is that it treats all elastic wave events as signal and uses them in the inversion to obtain three elastic parameters, P- and S-wave velocity and density. This is especially important for wide-offset seismic experiments and multi-component surveys where both P and S waves are present. Synthetic data tests show that the elastic inversion works on realistic sized problems with excellent resolution between the P and S-wave velocities. For the inversion of reflection data alone (shot profiles), only the high frequency components of the elastic parameters were resolved. In the case where transmission seismic data (i.e. data containing direct waves such as VSP data) is used in the inversion in addition to reflection seismic data, the low frequency components are well resolved and the inversion result looks almost exactly like the true model in synthetic tests. Future work is required to test the algorithm more thoroughly under different circumstances such as when significant multiples (and hence nonlinearities) are present, for the case of inversions of rayleigh waves and refractions for the near surface and most importantly, for the case of field seismic data.

ACKNOWLEDGMENTS.

I acknowledge the support of the sponsors of the Stanford Exploration Project (SEP) and Jon Claerbout. Also, I acknowledge the CNRS (C_2VR) who supplied some CRAY time. Thanks to Albert Tarantola for many stimulating discussions and making possible an 8 month stay at the Institut de Physique du Globe where this work was initiated and the early stages of development carried out. Particular thanks to Alexandre Nercessian for priceless computer aid. Thanks to Odile Gauthier and Antonio Pica who are working on the acoustic equivalent of the algorithm presented and helped me to understand many aspects. Thanks to Jean Remy for computer assistance and to Jean Virieux for help in linking to the CRAY. Finally, thanks to my fellow students at Stanford, particularly, John Toldi, Dan Rothman, Kamal Al-Yahya and Jos Van Trier, for provocative and interesting discussion.

REFERENCES

Cerjan, C., Kosloff D., Kosloff R., and Resheff M., 1985, Short note: A nonreflecting boundary condition for discrete acoustic and elastic wave equations: Geophysics, **50**, 705-708.

Devaney, A. J., 1984, Geophysical diffraction tomography: Inst. Electr. and Electron. Eng., Trans., Geosci. and Remote Sensing, **GE-22**, 3-13.

Gauthier O., and Tarantola A., 1985, Nonlinear inversion of two-dimensional seismic wavefields: Preliminary results: Geophysics, submitted for publication.

Kolb P., and Canadas G., 1986, Least-squares inversion of prestack data: simultaneous identification of density and velocity: presented at the 16-th conference on mathematical geophysics, June 22 - 28, Oosterbeek, Holland.

Kosloff D., Reshef M., and Loewenthal D., 1984, Elastic wave calculations by the Fourier method: Bulletin of the Seismological Society of America, **74**, 875-891.

Luenberger D., 1984, Linear and nonlinear programming, second edition: Adison-Wesley.

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

Mora P., 1986a, Nonlinear 2D elastic inversion of multi-offset seismic data, SEP-48.

Mora P., 1986b, Elastic finite differences with convolutional operators, SEP-48.

Powell, M.J.D., 1981, Approximation theory and methods: Cambridge university press, Cambridge.

Ronen, J. M., 1985, Multichannel inversion in reflection seismology: Ph.D. thesis, Stanford Univ..

Tarantola A., and Valette, B., 1982, Inverse problems = quest for information: Journal of geophysics, **50**, 159-170.

Tarantola A., 1984, The seismic reflection inverse problem, in: Inverse problems of acoustic and elastic waves, edited by: F. Santosa, Y.H. Pao, W. Symes, and Ch. Holland, SIAM, Philadelphia.