Multiparameter Inversion and Energy Source Estimation
for a Reflection Seismic Experiment

Susan E. Minkoff

May 1995

TR95-12

RICE UNIVERSITY

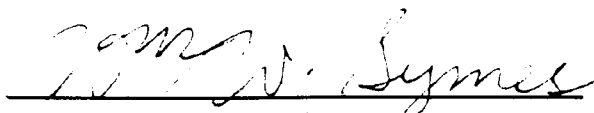# Multiparameter Inversion and Energy Source Estimation for a Reflection Seismic Experiment
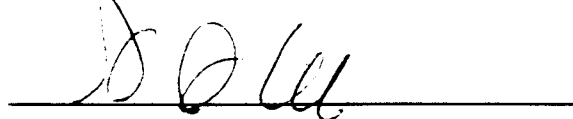
by

Susan E. Minkoff

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy
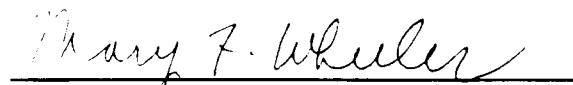
APPROVED, THESIS COMMITTEE:

William W. Symes, Chairman
Professor of Computational and Applied
Mathematics

Steven J. Cox
Associate Professor of Computational and
Applied Mathematics

Alan R. Levander
Professor of Geology and Geophysics

Mary F. Wheeler
Noah Harding Professor of Computational
and Applied Mathematics

Houston, Texas

May, 1995

Contents

4   Full Waveform Inversion of Marine Reflection Data in the Plane-Wave Domain   47

5   A Few More Interesting Marine Data Experiments   89

6   An Application of Model Resolution to Source-Reflectivity Inversions   123

Illustrations

Tables

Multiparameter Inversion and Energy Source Estimation for a Reflection Seismic Experiment

Susan E. Minkoff

Abstract

Reflection seismologists illuminate the subsurface by introducing energy into the ground. These propagating waves encounter heterogeneities in the subsurface material and are partly reflected back up to the surface where they are recorded as seismograms. The seismic energy source in most cases cannot be reliably measured in a laboratory but must be accurately estimated to allow one to extract the physical parameters which characterize the subsurface (such as velocity and density). The source and multiple earth parameters may be simultaneously successfully estimated by inversion.

When the seismogram model is the plane-wave convolutional model derived from the constant density, variable sound velocity acoustic wave equation, perturbations in the seismic data stably determine perturbations in the source and reflectivity (the high-frequency relative fluctuation in the velocity). The stability of this determination improves as the angular range over which the data is defined increases.

A more realistic model for wave propagation in the earth is the plane-wave convolutional model derived from the viscoelastic wave equation. Waveform inversion applied to field data from the Gulf of Mexico successfully estimates the long-wavelength compressional velocity, three elastic parameter reflectivities, and the anisotropic seismic source. The resulting reflectivities match measured well log data and agree with commonly-accepted lithological relationships. These inversion results predict 70% of the total seismic data and 90% of the data in an interval around the gas sand target.

The resolution matrix measures how close inversion-estimated reflectivities are to the true parameters which generated the data and is useful when independent information such as well logs is unavailable. However, computing the resolution matrix from the singular value decomposition of the forward map (the usual technique) is prohibitive for real seismic inverse problems. Instead we approximate the resolution

matrix from Lanczos estimates of the eigenvectors of the normal matrix. The resolution matrix indicates that our inversion-estimated source provides well resolved reflectivities in the depth interval of interest.

Chapter 1

Introduction

1.1 Overview

In exploration seismology, scientists illuminate the subsurface by generating a disturbance in the earth (for example by detonating an air gun in water or explosives on land). This man-made disturbance propagates down into the earth where heterogeneities in subsurface materials cause portions of the waves to be reflected back to the surface. Equipment set up on the surface captures this returning energy and creates a record of this activity as a function of time which is called a seismogram. This thesis deals with seismic inversion by which we mean one chooses a physical model to describe the propagation of waves through the medium and attempts to improve the fit of the model to the given data by successively updating the parameters in the model. As seismologists have long known, the seismic source or "system wavelet is one of the undesirable components of the raw seismic data", ([35] p. 164). The seismic source's signature (time history) and radiation pattern (direction-dependence) are of little use in themselves. However, the seismic source must be separated from the data in some way so that geophysicists can estimate and analyze the material properties of the subsurface which are of interest (for example, velocity and density).

Estimating and removing the source from the seismic data is a task which seismologists have been studying for at least the last forty years. Classic techniques related to communication theory and signal processing were adapted for geophysics by Robinson and Treitel in the 1950's and 1960's (see [42], [43], and [45]). Another group led by the scientist Ziolkowski advocates measuring the source directly in a laboratory or in the experimental environment (see [64], [65], and [66]). This idea works well in settings where one has control over the input source such as in electrical impedance imaging (for example see [26]). However, in the case of reflection seismology, direct measurement of the source wavelet is not only extremely unrealistic (especially for land seismology), but also very unreliable. Both linear and nonlinear inversion methods for estimating the air gun source array in marine seismic work

are described in [32]. Their methods rely, however, on having data which has been recorded on a ministreamer of receivers towed a few meters below the source array; in some sense they have measured the source.

In Chapter 4 we discuss how our source estimate is an "effectIve source" which is quite useful for estimating other physical parameters in the model but is clearly not the uncontaminated source one would get from measuring the air gun pressure pulse in a laboratory. The source is affected by the medium surrounding it (such as the rocks around the hole in which explosives are placed for land experiments). The source we obtain for the real field data discussed in Chapter 4 has absorbed into its description information about the receiver array and possibly an inadequacy in our model for wave propagation in this part of the subsurface. This interesting flexibility of the source still allows very stable estimates of the physical parameters.

This source is estimated directly from the data using least squares inversion at the same time that we are estimating the other earth parameters. The references [7], [33], [46] appear to be the first mathematical papers on the use of least squares inversion to estimate the source. This thesis is likely the only work in which the source for field data is estimated using inversion.

The results described in Chapter 4 have led to natural questions about how much trust we can place in the multiple physical parameters we estimate at the time the source is being estimated. These questions arise because of the high degree of interest over the last ten years in the topic of *Amplitude Versus Offset* (or AVO) analysis. As recently as 1983 geophysicists were satisfied with estimating only one elastic parameter. In [44] p.19, Robinson explains:

> Usually, only the P-waves are considered useful in seismic exploration. We try to avoid recording S-waves and surface waves; when they are recorded, they are relegated to the category of noise. Exceptions to this rule are now being made, however, and in the future S-waves may provide invaluable lithologic information that we cannot obtain from P-waves alone.

We define the amplitude of a normally incident seismic wave reflected from an interface between two materials as being governed by the reflection coefficient R where

$$R = \frac{\rho_2 v_2 - \rho_1 v_1}{\rho_2 v_2 + \rho_1 v_1}.$$

The densities ρ_1 and ρ_2 and velocities v_1 and v_2 correspond to the top and bottom sides of the interface respectively. When the incident compressional waves (P-waves)

hit an interface at non-normal angles, shear wave (S-wave) motion is induced causing the amplitudes of reflected P-waves to vary significantly with angle ([14] pp.357-358, 378).

Drilling wells is costly and detrimental to the environment, sd exploration seismologists need to be able to distinguish between changes in the reflection coefficient caused by hydrocarbons and changes due to other factors or materials. Ostrander had a significant impact on the current exploration seismology community when he showed that Poisson's ratio (an elastic parameter related to the ratio between P- and S-wave velocities) tends to be very low for high-porosity gas sands ([37] and [10]). Reflection seismologists are now concerned with accurately estimating more than just the P-wave velocity.

The source estimate affects one's ability to estimate more than one of these physical parameters accurately. As noted by the authors in [14] (p. 379),

> The primary data inputs to offset-amplitude analyses are amplitude values as functions of offset, but variation in the subsurface P and shear velocities and densities are not the sole parameters that can cause such changes. Other factors include source directivity and array effects, near surface velocity variations,... For land data, one of the variables that is least known but most important is the lateral variability in coupling between source and earth, and receiver and earth,... Some adequate accounting or control must be given for each of these effects in order to relate the remaining amplitude variation to the reflectivity change effects.

Although more work in this area needs to be done to make strong assertions concerning the success of using the inversion techniques described in this thesis to estimate multiple elastic parameters, we give one very successful example for real marine data from the Gulf of Mexico. The thesis is divided into three main sections. Chapter 3 describes a theoretical result for separate source-reflectivity estimation (where the *reflectivity* is the high-frequency relative fluctuation in the velocity in this case). Chapter 4 contains the real data inversion results. Chapter 6 describes a practical tool for analyzing how close the inversion results are to the true parameters which generated the data. The remaining chapters provide background for the mathematical model used and discuss some other experiments conducted on the real data (Chapters 2 and 5 respectively). The main points discussed in this thesis are summarized below.

1.2   The Convolutional Model for the Seismogram

The model used throughout this thesis is the convolutional model for the seismogram so called because it is the time convolution of the seismic source with the high-frequency relative fluctuation in the physical parameters in the model. Chapter 2 gives the derivation of this model for the acoustic wave equation with variable sound velocity and density.

There are three main assumptions on the model common to all work in this thesis. The first assumption is that we may model the earth as a *layered medium*. In other words, we assume the parameters in the wave equation model depend only on one spatial variable, depth z. This assumption allows us to Radon transform or integrate part of the solution of the wave equation along planes in the two remaining horizontal directions. The end result of this process is a family of 1-D wave equations parametrized by the slowness parameter p which is related to the plane-wave angle of incidence θ by

$$p = \frac{\sin \theta}{c(z)}$$

where $c(z)$ is the velocity.

The second assumption we make is that we can separate the parameters in the wave equation model into short and long wavelength pieces or linearize the wave equation. This linearized equation means we model only the seismic energy which has been reflected once off a subsurface layer (*primary reflections* as opposed to multiples). The long-wavelength portion of the model parameters determines the kinematics of wave propagation and we term these functions "background". The short-scale relative fluctuations in the model parameters determine the reflections and we call these functions the "reflectivities". Both the assumption that the earth is layered and that multiple reflections may be neglected are made fairly often although geophysicists feel these simplifications do not adequately describe some areas of the subsurface. Finally, to arrive at the convolutional model for the seismogram, we apply geometric optics which assumes the *source is high-frequency* (i.e., the peak frequency corresponds to a wavelength which is much shorter than that of a typical record, see [56] and [13] Ch.6).

The theoretical stability result in Chapter 3 applies to the primaries-only, plane-wave, layered-medium *constant density acoustic wave equation*. The model for the seismogram is

$$S(t,p) = f(t) * \tilde{r}(t,p)$$

$$\tilde{r}(t,p) \approx \int A(z,t,p)r(z)dz$$

The *isotropic source* is $f(t)$, and the relative fluctuation in the sound velocity is $r(z)$. $A(z,t,p)$ is the geometric optics reflectivity amplitude.

Similarly, the real data numerical experiments described in Chapters 4-6 assume a primaries-only, plane-wave, layered-medium, *viscoelastic wave equation* with an *anisotropic source*.

$$S(t,p) = f(t,p) * \tilde{r}(t,p)$$

$$\tilde{r}(t,p) \approx \int [A_P(z,t,p)r_P(z) + A_S(z,t,p)r_S(z) + A_d(z,t,p)r_d(z)]dz$$

Here r_P is the relative short-scale fluctuation in the P-wave velocity, r_S the fluctuation in the S-wave velocity and r_d the fluctuation in density. A_P, A_S, and A_d are the respective geometric optics reflectivity amplitudes.

1.3 Stable Determination of the Source and Reflectivity for the Acoustic Wave Equation

In Chapter 3 we extend work begun by R. M. Lewis in his PhD thesis ([33]). Lewis assumed that the long-wavelength background velocity in the acoustic model is constant. He then showed that one could stably determine perturbations in the source and reflectivity from perturbations in the data.

The extension we make in Chapter 3 drops the restriction that the background velocity be constant. Constant background velocities do not accurately describe the kinematics of wave propagation in the earth. Proving a direct result such as Lewis' was made difficult in this case by the change of variables from time to depth via velocity. We, therefore, show that the normal operator is continuous in the background velocity allowing us to extend the stability result at constant velocity to velocities close to constant. The stability depends on the frequency content of the source, how close to "white" (containing all frequencies) the reflectivity series is, and the range of slowness values (angles) in the data. Unfortunately, it is clear that the theoretical result described in Chapter 3 applies to a model which still does not accurately describe real data.

1.4 Full Waveform Inversion of Gulf of Mexico Field Data Using a Viscoelastic Model

We were interested to see how well we could invert for the seismic source which generated field data. This idea was the beginning of a large inversion study of a particular marine data set from the Gulf of Mexico. In fact all of the numerical experiments described in this thesis were performed on this real data or on synthetic data sets derived from this marine data. In Chapter 4 the model, techniques, and main results are described. Chapter 5 contains some additional experiments worth mentioning but which are somewhat less conclusive. We chose to invert for the P-wave background velocity, anisotropic seismic source, and three combinations of the elastic parameter reflectivities (P-wave velocity, S-wave velocity, and density) using a viscoelastic model for wave propagation.

An algorithm for successfully inverting real data eventually emerged from this study. The P-wave background velocity has a nonlinear (and therefore significant) influence on the data and is the most important parameter to estimate accurately. We inverted first for this parameter and the reflectivities. Once we were satisfied with our background velocity estimate, we alternated between re-estimating the reflectivities and estimating the seismic source. Each of these parameters separately has a linear influence on the data. We found that although the source has a major impact on how well the reflectivities are estimated, it has little effect on how well the background velocity is estimated. The other model parameters (attenuation factors, background S-wave velocity and density) were not estimated in the inversion.

Our best model estimates allowed us to predict all but 29% of the real seismic data, a rather remarkable result. This same data was studied in the PhD thesis of Igel ([25]). He used an elastic 2D modeler and did not invert for the seismic source. He was able to fit all but 45% of the data inverting for short-scale fluctuations (reflectivities) in the P-wave impedance and Poisson's ratio.

Data misfit is one mathematical indicator of accuracy for inversion results. Comparison to a priori geologic knowledge of the region of the subsurface is, however, a more common measure. We, therefore, also compared our reflectivity results to detrended well-log measurements of the high-frequency parts of the P-wave velocity and density, and we checked whether our results were in agreement with some common lithological laws governing the relation between different elastic parameters.

One of the most interesting aspects of this study involved understanding what role our source estimate played in the inversion. We found that the condition number (ratio of largest to smallest eigenvalues) of the source subproblem was orders of magnitude worse than the condition of the single-reflectivity subproblem (in which we only inverted for the P-wave impedance reflectivity). When we were inverting for all three reflectivities, however, the condition of the reflectivity subproblem was about the same as the condition of the source subproblem.

In truth, this data set did not agree with one of our modeling assumptions — namely, that the subsurface was layered. Even after pre-processing the data, there were obvious near-surface heterogeneities. Our source estimate, however, may have accounted for some of this modeling inaccuracy. Heiner Igel noted in his thesis ([25] p. 132):

> An initial inversion for P-impedance gave a reasonable fit for the part of the seismogram following the waterbottom reflection which was not well modeled. I currently do not understand the problems with the waterbottom. Most likely they are due to lack of knowledge of the source function and/or the structure of the waterbottom which may not well be modeled by an elastic modeling [sic] which does not take into account attenuation.

1.5 Resolution of the Model Parameters

Finally, we contrast the resolution as a function of depth of the elastic parameter reflectivity estimates for two different sources. There was no measured shear log for this area which could be used for comparison to inversion results. Moreover, it is conceivable that in many instances one would not be able to compare inversion results to logged wells. Chapter 6 provides a practical way of trying to quantify how close the inversion results are to the true parameters which generated the data.

In a now classic paper, Backus and Gilbert [4] describe how seismic inverse problems with non-unique solutions may have unique solutions in an average sense. They also describe how to determine the shortest length scale which the given data can resolve at a particular depth. (Related references include [17], [29], and [27].) The model resolution matrix is one way to quantify how close the model estimates are to the true model as a function of the independent variable (such as depth). Wiggins [62] defines this matrix in terms of the singular value decomposition (SVD) and applies this technique to a very small problem (less than ten unknowns). He comments in the abstract to the paper that "computation of parameter and information resolution is such a

simple extension of any inversion procedure based on perturbation parameters that such inversion studies are incomplete without considering resolution." Forming the resolution matrix does not seem to have become common practice, however, because it is computationally prohibitive to implement the idea directly oh large problems. Papers by Martinez and McMechan [35], Assous and Collino [3], and Bishop et al. [6] are attempts to use the SVD to compute the well and poorly-determined parameters. All of these attempts are restricted by the cost of computing the SVD. Martinez and McMechan use a viscoelastic simulator to invert for velocities, density, and quality factors. They examine model resolution for a simple geometry experiment of a target layer located between two homogeneous half-spaces. For field data experiments they do not compute resolution. Bishop et al. estimate reflector depth and seismic velocities from tomography data. They analyze the determination of components of a simplified model analytically and substantiate their conclusions with a small numerical experiment performed on real data. Finally, Assous and Collino determine resolution for a more complicated problem by storing the full normal matrix and calling an IMSL generalized eigenvalue problem solver.

In all cases these authors assume one had computed and stored the entire forward or normal matrices during the inversion. The in-core storage of these large matrices is unnecessary. We use a version of the conjugate gradient and Lanczos algorithms which allows us to estimate the eigenvalues and eigenvectors of the normal matrix with almost no additional cost over solving the inverse problem. The approximate resolution matrix we obtain gives preliminary evidence that the reflectivity estimates from the inversion-estimated source are slightly better resolved than are reflectivity estimates from an independently estimated air gun model source.

Chapter 2

Derivation of the Convolutional Model for the Acoustic Wave Equation

2.1 Introduction

In this chapter we explain the steps in the derivation of the convolutional model for the seismogram (extracted from the more complete treatment given in [56]). The convolutional model described here comes from assuming the earth is an acoustic layered fluid where the parameters in the equation (velocity and density) depend only on depth. The wave equation model we are assuming in this chapter is slightly more complicated than the model used in Chapter 3 (the *constant density* acoustic wave equation) and much less complicated than the model used in Chapters 4–6 (the *viscoelastic* wave equation). Nonetheless, the process one goes through to derive the convolutional model for the seismogram from these different wave equations is the same. Moreover, in Chapters 3 and 4 the convolutional model appropriate for those applications is defined explicitly. Fewer or more terms may be added to the "reflectivity" term, but the general model remains very similar.

Because the derivation is long, I will indicate here (and by section headings throughout the chapter) the basic steps which one must go through to arrive at the convolutional model. The first assumption one makes is that the parameters in the wave equation model (in this case sound velocity and density) depend on only one spatial variable, depth. The Radon transform can therefore be applied to reduce the 3-D equation to a family of 1-D equations parametrized by a quantity called slowness. Secondly, if one assumes that multiply reflected energy may be neglected, one is able to *linearize* the wave equation. Finally, one applies high-frequency asymptotics or geometric optics to this linearized equation to arrive at the final convolutional model.

2.2 The Model

In [47] this model is defined in more detail. The model we use for wave propagation in the earth is the 3-D acoustic wave equation:

$$\frac{1}{\rho(z)c^2(z)}\frac{\partial^2 P}{\partial t^2} - \nabla \cdot \frac{1}{\rho(z)}\nabla P = F(\vec{x}, t) \tag{2.1}$$

Here $P(t, \vec{x}) = P(t, x, y, z \geq 0)$ is the excess pressure; $\rho(z)$ and $c(z)$ are the depth-dependent density and sound velocity respectively. The time-dependent, anisotropic source is denoted F. We assume the source is causal, i.e., that

$$\left.\begin{array}{l} F(\vec{x}, t) \equiv 0 \\ P(\vec{x}, t) \equiv 0 \end{array}\right\} t << 0$$

Physical boundary conditions are imposed as well (for example, in marine geophysics, from the ocean surface). These are not discussed here.

The inverse problem may now be stated as given recordings of the pressure $P(x_r, t_r)$ at the receiver locations x_r and times t_r and for sources $F(\vec{x}, t)$, estimate the coefficients density $\rho(z)$ and sound velocity $c(z)$.

The geometry of the reflection seismology experiment allows us to assume the coefficients $\rho(z)$ and $c(z)$ are known for depths above z_0. In this setting depth z increases downward into the earth with the experimental surface denoted z_0. The source locations $\{x_s\}$ and receiver locations $\{x_r\}$ in the experiment are assumed to lie on the plane $z = z_0$. We also can realistically only measure the reflected signal over a finite time interval $0 \leq t \leq T$. The sampling rate is assumed to be the same for all receivers.

We assume the source has point support, i.e., that it has spatial extent much smaller than a seismic wavelength. In Chapter 3 we will make use of an *isotropic* source. For the real data experiments described in Chapters 4–6, the source is allowed to be *anisotropic*. We see in the experiments in Chapters 4–6 how important the anisotropy of the source is for fitting real seismic data. However, this complication has no impact on our current calculations and so will be avoided. Our sources will, therefore, have the form: $F(\vec{x}, t) = f(t)\delta(x - x_s)$ where x_s is the point source location and $f(t)$ is the time series for the source signal. One extremely important aspect of all sources discussed in this thesis is that they are band-limited, usually with frequency content in the range of 10–90 Hz. The frequency content of the source indicates to what scale inhomogeneities in the earth will be resolved.

For marine experiments, the measured seismic data is

$$P(t, x, y, z = 0) = S(t, x, y).$$

For land experiments, the measured data is the time derivative of the particle displacement.

2.3 Application of the Radon Transform

We first introduce the slowness parameter $p \in \Re$ which is related to the plane-wave angle of incidence θ by the relation

$$p = \frac{sin\theta}{c(z)}. \tag{2.2}$$

Then the Radon-transformed solution to the acoustic wave equation is given by

$$u(t, z, p) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(t + p \cdot x, \mathbf{x}, z) d\mathbf{x}$$

where here $\mathbf{x} \equiv (x, y)$. The Radon-transformed field $u(t, z, p)$ is a plane-wave component of P and satisfies the 1-D wave equation:

$$\frac{1}{\rho(z)} \left(\frac{1}{c^2(z)} - p^2 \right) \frac{\partial^2 u(t, z, p)}{\partial t^2} - \frac{\partial}{\partial z} \frac{1}{\rho(z)} \frac{\partial u(t, z, p)}{\partial z} = f(t)\delta(z - z_s) \tag{2.3}$$

With boundary conditions:

$$\left. \begin{array}{l} f(t) \equiv 0 \\ u(t, z, p) \equiv 0 \end{array} \right\} t << 0$$

And the seismic data is now:

$$u(t, z = 0, p) = S(t, p)$$

We will work with this set of simpler 1-D equations rather than with the original 3-D equation. We make the assumption of *pre-criticality* throughout this thesis, namely, that we only consider slowness values p such that $c_{max}(z)|p| < 1$. Then the apparent velocity

$$v(z, p) \equiv \frac{c(z)}{\sqrt{1 - c^2(z)p^2}}$$

is always real and equation 2.3 is hyperbolic.

2.4 Linearization of the Wave Equation

We consider the formal linearization of the wave equation 2.3 with respect to the parameters velocity and density. Substitute $\rho + \delta\rho$ in for the density and $c + \delta c$ for the sound velocity. The solution u becomes $u + \delta u$. In general terms, this formal linearization to the acoustic wave equation is a good approximation to the derivative of the forward map (model for the seismogram) if

1. The coefficients describing the background medium, ρ and c, are slowly-varying (smooth) relative to a typical data wavelength.
2. The perturbations, $\delta\rho$ and δc, are high-frequency or "rough".

The formal linearization has been studied for the 1-D problem in references [33], [55], and [57].

If in this modified version of equation 2.3 we write the perturbed density (and similarly the velocity) so that

$$\frac{1}{\rho + \delta\rho} = \frac{1}{\rho}\left(\frac{1}{1 + \delta\rho/\rho}\right)$$

then we can use the Binomial series ([2], p. 244):

$$\frac{1}{(1-x)^c} = \sum_{k=0}^{\infty}\left(\begin{array}{c} -c \\ k \end{array}\right)(-1)^k x^k, \qquad \text{if } -1 < x < 1.$$

The linearized equation for δu is found by simplifying the resulting equation and dropping terms of order greater than one in the perturbed quantities:

$$\frac{1}{\rho(z)}\left(\frac{1}{c^2(z)} - p^2\right)\frac{\partial^2 \delta u(p,z,t)}{\partial t^2} - \frac{\partial}{\partial z}\frac{1}{\rho(z)}\frac{\partial \delta u(p,z,t)}{\partial z} = \qquad (2.4)$$

$$\frac{1}{\rho(z)}\left(\left(\frac{1}{c^2(z)} - p^2\right)\frac{\delta\rho(z)}{\rho(z)}\right)\frac{\partial^2 u(p,z,t)}{\partial t^2} + \frac{2\delta c(z)}{\rho(z)c^3(z)}\frac{\partial^2 u(p,z,t)}{\partial t^2} - \frac{\partial}{\partial z}\left(\frac{\delta\rho(z)}{\rho^2(z)}\right)\frac{\partial u(p,z,t)}{\partial z}$$

2.5 Application of High Frequency Asymptotics to the Linearized Equation

2.5.1 Approximation of the Perturbed Green's Function

Rather than solving for δu directly, we note that we may solve instead for the perturbed Green's function δG and then write $\delta u = f *_t \delta G$. The process described in

this section allows us to derive an approximation containing the most significant portion of the perturbed Green's function. We assume that the source function is "high frequency", and therefore the most singular part of δG (corresponding to large high frequency Fourier coefficients) will give the greatest contribution to the convolution solution δu. One necessary assumption is that of *simple ray geometry*. This assumption basically states that the reflected ray cannot be a continuation of the incident ray, and we have ruled out the possibility of caustics. The calculations in this section are presented completely formally. For a more detailed version of this derivation, see [56].

We start by noting that

$$\delta G(z_s, z_r, t_r) = \int_{\Re} dz \int dt \delta G(z_s, z, t) \delta(z_r - z) \delta(t_r - t)$$

$$= \int_{\Re} dz \int dt \delta G(z_s, z, t) \left[\frac{1}{\rho(z)} \left(\frac{1}{c^2(z)} - p^2 \right) \frac{\partial^2 G}{\partial t^2}(z_r, z, t_r - t) - \frac{\partial}{\partial z} \frac{1}{\rho(z)} \frac{\partial G(z_r, z, t_r - t)}{\partial z} \right]$$

because G satisfies the wave equation with δ right hand side. We will now formally integrate these singular terms by parts.

$$= \int_{\Re} dz \int dt \left[\frac{1}{\rho(z)} \left(\frac{1}{c^2(z)} - p^2 \right) \frac{\partial^2 \delta G(z_s, z, t)}{\partial t^2} - \frac{\partial}{\partial z} \frac{1}{\rho(z)} \frac{\partial \delta G(z_s, z, t)}{\partial z} \right] G(z_r, z, t_r - t)$$

Note that this formal calculation is somewhat justified in that we have chosen as our second term $G(z_r, z, t_r - t)$ an "advanced fundamental solution" to keep the products of supports of the Green's functions bounded. We now substitute the right-hand side of the linearized model (equation 2.4) into the expression above to get

$$\delta G(z_s, z_r, t_r) = \int_{\Re} dz \int dt \left[\frac{1}{\rho(z)} \left(\left(\frac{1}{c^2(z)} - p^2 \right) \frac{\delta \rho(z)}{\rho(z)} \right) \frac{\partial^2 G(z_s, z, t)}{\partial t^2} \right. \tag{2.5}$$

$$\left. + \frac{2 \delta c(z)}{\rho(z) c^3(z)} \frac{\partial^2 G(z_s, z, t)}{\partial t^2} - \frac{\partial}{\partial z} \frac{\delta \rho(z)}{\rho^2(z)} \frac{\partial G(z_s, z, t)}{\partial z} \right] G(z_r, z, t_r - t)$$

We will integrate expression 2.5 by parts. We start by integrating the first term in the integrand with respect to t. The resulting expression is

$$\int_{\Re} dz \int dt \left[\frac{1}{\rho(z)} \left(\left(\frac{1}{c^2(z)} - p^2 \right) \frac{\delta \rho(z)}{\rho(z)} \right) \frac{\partial^2 G(z_s, z, t)}{\partial t^2} G(z_r, z, t_r - t) \right. \tag{2.6}$$

$$= \int_{\Re} dz \int dt \left[\frac{1}{\rho(z)} \left(\left(\frac{1}{c^2(z)} - p^2 \right) \frac{\delta \rho(z)}{\rho(z)} \right) \frac{\partial G(z_s, z, t)}{\partial t} \frac{\partial G(z_r, z, t_r - t)}{\partial t} \right.$$

Here we have assumed that the boundary term can be neglected since the Green's functions $G(z_s, z, t)$ and $G(z_r, z, t_r - t)$ have bounded support.

The second term in integral 2.5 is also integrated by parts with respect to t. In this case we get:

$$\int_{\mathfrak{R}} dz \int dt \frac{2\delta c(z)}{\rho(z)c^3(z)} \frac{\partial^2 G(z_s, z, t)}{\partial t^2} G(z_r, z, t_r - t) \qquad (2.7)$$

$$= \int_{\mathfrak{R}} dz \int dt \frac{2\delta c(z)}{\rho(z)c^3(z)} \frac{\partial G(z_s, z, t)}{\partial t} \frac{\partial G(z_r, z, t_r - t)}{\partial t}$$

Again, the boundary term goes to 0. Finally, the third term in the integral 2.5 is integrated by parts with respect to z.

$$- \int_{\mathfrak{R}} dz \int dt \frac{\partial}{\partial z} \frac{\delta\rho(z)}{\rho^2(z)} \frac{\partial G(z_s, z, t)}{\partial z} G(z_r, z, t_r - t) \qquad (2.8)$$

$$= - \int_{\mathfrak{R}} dz \int dt \left(\frac{\delta\rho(z)}{\rho^2(z)} \frac{\partial G(z_s, z, t)}{\partial z} \frac{\partial G(z_r, z, t_r - t)}{\partial z} \right)$$

We now define the Progressing Wave Expansion (for a thorough explanation see reference [13] Ch.6). We will approximate our Green's function solution by the singular series

$$G(z, t) = \sum_{n=0}^{N} a_n(z) S_n(t - \phi(z)) + R(z, t)$$

where $S_n'(t) = S_{n-1}(t)$ for $n = 0, 1, 2, \ldots$. Here, the prime denotes differentiation with respect to t. Thus, the most singular term is the first term in the series, and specifically, the remainder term R is smoother than the smoothest term in the sum, S_N. The phase is denoted $\phi(z)$ and the $a_n(z)$ are the transport coefficients. We will defer calculating the phase ϕ and first transport coefficient a_0 until later. For now we return to arriving at an expression for the Green's function itself. Since the fundamental solution of the 1-D acoustic wave equation in homogeneous medium is given by

$$G(x_s, x, t) = \rho H(ct - |x - x_s|)$$

We take as our progressing wave expansion for the Green's function the series

$$G(z, t) = a_0(z) H(t - \phi(z)) + R(z, t) \qquad (2.9)$$

where $H(t - \phi(z))$ is the Heaviside function defined so that

$$H = \begin{cases} 0 & \text{if } t < \phi(z) \\ 1 & \text{otherwise} \end{cases}$$

We substitute the progressing wave expansion expression 2.9 into the expression for δG gotten by combining 2.6, 2.7, 2.8.

$$\delta G(z_s, z_r, t_r) = \int_{\Re} dz \int dt \left\{ \left[\frac{1}{\rho(z)} \left(\frac{1}{c^2(z)} - p^2 \right) \frac{\delta \rho(z)}{\rho(z)} + \frac{2 \delta c(z)}{\rho(z) c^3(z)} \right] \right. \tag{2.10}$$

$$\frac{\partial}{\partial t}[a_0(z_s, z) H(t - \phi(z_s, z)) + R(z_s, z, t)] \frac{\partial}{\partial t}[a_0(z_r, z) H(t_r - t - \phi(z_r, z)) + R(z_r, z, t_r - t)]$$

$$- \frac{\delta \rho(z)}{\rho^2(z)} \frac{\partial}{\partial z}[a_0(z_s, z) H(t - \phi(z_s, z)) + R(z_s, z, t)] \frac{\partial}{\partial z}[a_0(z_r, z) H(t_r - t - \phi(z_r, z))$$

$$\left. + R(z_r, z, t_r - t)] \right\}$$

This expression is meaningful so long as the hypersurfaces defined by $0 = t_r - t - \phi(z_r, z)$ and $0 = t - \phi(z_s, z)$ intersect transversely (with normals not parallel at points of intersection). The transversality condition is violated when the assumption of simple ray geometry which we assumed earlier is also violated. (For further explanation see [56].)

If we carry the differentiations throughout the integrand now and drop all terms except the most singular we arrive at the expression:

$$\delta G(z_s, z_r, t_r) \approx \tag{2.11}$$

$$\int_{\Re} dz \int dt \left[-\frac{1}{\rho(z)} \left(\frac{1}{c^2(z)} - p^2 \right) \frac{\delta \rho(z)}{\rho(z)} a_0(z_s, z) \delta(t - \phi(z_s, z)) a_0(z_r, z) \delta(t_r - t - \phi(z_r, z)) \right.$$

$$\left. - \frac{2 \delta c(z)}{\rho(z) c^3(z)} a_0(z_s, z) \delta(t - \phi(z_s, z)) a_0(z_r, z) \delta(t_r - t - \phi(z_r, z)) \right]$$

$$- \frac{\delta \rho(z)}{\rho^2(z)} \left[a_0(z_s, z) \delta(t - \phi(z_s, z)) \frac{\partial \phi(z_s, z)}{\partial z} a_0(z_r, z) \delta(t_r - t - \phi(z_r, z)) \frac{\partial \phi(z_r, z)}{\partial z} \right]$$

We now need to use explicit expressions for the phase ϕ and first transport coefficient a_0. Rather than calculating those terms at this time, we jump ahead to the results themselves so we can finish the derivation of the convolutional model. We find in Subsection 2.5.2 that if we define the vertical velocity of the plane-wave of slowness p by

$$v(z, p) = \frac{c(z)}{\sqrt{1 - c^2(z) p^2}} \tag{2.12}$$

then

$$\phi'(z) = \frac{1}{v(z)} = \sqrt{\frac{1}{c^2(z)} - p^2}. \tag{2.13}$$

We also find that

$$a_0(z) = a_0(0)(\rho(z)v(z))^{1/2}. \tag{2.14}$$

If we substitute the expression for ϕ' into the last term in the integrand of expression 2.11, we have that

$$\delta G(z_s, z_r, t_r) \approx \int_{\Re} dz \int dt \left[-\frac{2\delta\rho(z)}{\rho^2(z)} \left(\frac{1}{c^2(z)} - p^2 \right) - \frac{2\delta c(z)}{\rho(z)c^3(z)} \right] \tag{2.15}$$

$$\times a_0(z_s, z)a_0(z_r, z)\delta(t - \phi(z_s, z))\delta(t_r - t - \phi(z_r, z))$$

To evaluate the integrals, change variables. Let $X = t - \phi(z_s, z)$ and $Y = t_r - t - \phi(z_r, z)$. Then

$$\frac{\partial(X, Y)}{\partial(t, z)} = \begin{pmatrix} 1 & -1/v(z) \\ -1 & -1/v(z) \end{pmatrix}$$

$$\frac{\partial(t, z)}{\partial(X, Y)} = -\frac{1}{2} \begin{pmatrix} -1 & 1 \\ v(z) & v(z) \end{pmatrix}$$

And,

$$\det \left| \frac{\partial(t, z)}{\partial(X, Y)} \right| = \frac{-v(z)}{2}$$

Applying this change of variables to expression 2.15 gives that

$$\delta G(z_s, z_r, t_r) \approx \int_{\Re} \int dX dY \left\{ \frac{-v}{2} \left[\frac{-2\delta\rho}{\rho^2} \left(\frac{1}{c^2} - p^2 \right) - \frac{2\delta c}{\rho c^3} \right] a_0^2 \delta(X)\delta(Y) \right\} \tag{2.16}$$

$$= \left[\frac{\delta\rho}{\rho^2} \left(\frac{1}{c^2} - p^2 \right)^{\frac{1}{2}} + \frac{\delta c}{\rho c^3} \left(\frac{1}{c^2} - p^2 \right)^{-\frac{1}{2}} \right] a_0^2 \Bigg|_{X=Y=0}$$

Now, $X = 0 \Rightarrow t = \phi(z_s, z)$ and $Y = 0 \Rightarrow z = \phi^{-1}(1/2t_r)$. From expressions 2.12 and 2.14 we have that

$$\delta G(z_s, z_r, t_r) \approx \left[\frac{\delta\rho}{\rho^2} v^{-1}(z) + \frac{\delta c}{\rho c^3} v(z) \right] \left[a_0(0)^2 \rho(z)v(z) \right] \Bigg|_{X=Y=0} \tag{2.17}$$

$$= \left[\frac{\delta\rho}{\rho} + \frac{\delta c}{c^3} v^2(z) \right] a_0(0)^2 \Big|_{X=Y=0}$$

The approximation to the perturbed Green's function is, therefore, given by

$$\delta G(z_s, z_r, t_r) \approx a_0^2(0) \left\{ \frac{\delta\rho}{\rho} + \frac{\delta c}{c} \left(\frac{1}{1 - p^2 c^2} \right) \right\} \left(\phi^{-1}(1/2t_r) \right) \tag{2.18}$$

Therefore, the perturbed solution $\delta u \approx f * \delta G$.

2.5.2 Calculation of the Phase Function and First Transport Coefficient

We first calculate the phase function ϕ. We substitute the progressing wave expansion

$$u(p, z, t) = \sum_{n=0}^{N} a_n(z) S_n(t - \phi(z)) + R(z, t) \tag{2.19}$$

into the acoustic wave equation

$$\frac{1}{\rho(z)} \left(\frac{1}{c^2(z)} - p^2 \right) \frac{\partial^2}{\partial t^2} u(p, z, t) - \frac{\partial}{\partial z} \frac{1}{\rho(z)} \frac{\partial}{\partial z} u(p, z, t) = 0 \tag{2.20}$$

(which hold for $z \neq z_s$). We chose the terms in the singular series S_n so that $S_n' = S_{n-1}$ so

$$\frac{\partial^2 u}{\partial t^2} = \sum_{n=0}^{N} a_n(z) S_{n-2}(t - \phi(z)) + \frac{\partial^2 R(z, t)}{\partial t^2} \tag{2.21}$$

And

$$\frac{\partial u}{\partial z} = \sum_{n=0}^{N} [a_n'(z) S_n(t - \phi(z)) - a_n(z) \phi'(z) S_{n-1}(t - \phi(z))] + \frac{\partial R(z, t)}{\partial z} \tag{2.22}$$

$$\frac{\partial^2 u}{\partial z^2} = \sum_{n=0}^{N} [a_n''(z) S_n(t - \phi(z)) - 2a_n'(z) \phi'(z) S_{n-1}(t - \phi(z)) \tag{2.23}$$

$$-a_n(z) \phi''(z) S_{n-1}(t - \phi(z)) + a_n(z)(\phi'(z))^2 S_{n-2}(t - \phi(z))] + \frac{\partial^2 R(z, t)}{\partial z^2}$$

Expressions 2.20, 2.21, 2.22, and 2.23 give that

$$\sum_{n=0}^{N} S_{n-2}(t - \phi(z)) \left[\frac{1}{\rho(z)} \left(\frac{1}{c^2(z)} - p^2 \right) a_n(z) - \frac{1}{\rho(z)} a_n(z)(\phi'(z))^2 \right] \tag{2.24}$$

$$+ \sum_{n=0}^{N} S_{n-1}(t - \phi(z)) \left[-\frac{\rho'(z)}{\rho^2(z)} a_n(z)\phi'(z) + \frac{2}{\rho(z)} a_n'(z)\phi'(z) + \frac{1}{\rho(z)} a_n\phi''(z) \right]$$

$$+ \sum_{n=0}^{N} S_n(t - \phi(z)) \left[\frac{\rho'(z)}{\rho^2(z)} a_n'(z) - \frac{1}{\rho(z)} a_n''(z) \right] \approx 0.$$

To find the phase function we solve for the coefficient of the S_{n-2} term (the most singular):

$$\left[\left(\frac{1}{c^2(z)} - p^2 \right) - (\phi'(z))^2 \right] \frac{1}{\rho(z)} a_0(z) = 0$$

This equation is the eikonal equation of geometric optics. The solution to this equation for heterogeneous media may be computed numerically via the method of characteristics (as described in [56]). We find that

$$\int_0^z \sqrt{\frac{1}{c^2(\xi)} - p^2} \, d\xi = \phi(z) \tag{2.25}$$

To recover the first transport coefficient, namely a_0, we must solve the equation

$$-\frac{\rho'(z)}{\rho(z)} \phi'(z) a_0(z) + 2\phi'(z) a_0'(z) + \phi''(z) a_0(z) = 0$$

We rewrite this equation as

$$a_0'(z) = \frac{1}{2} \left(\frac{\rho'(z)}{\rho(z)} - \frac{\phi''(z)}{\phi'(z)} \right) a_0(z)$$

And make use of the expression for the phase we just derived (expression 2.25). Then

$$a_0'(z) = \frac{1}{2} \left((\log \rho(z))' - \left(\log \frac{1}{v(z)} \right)' \right) a_0(z)$$

$$\Rightarrow \frac{a_0'(z)}{a_0(z)} = \frac{1}{2} \left((\log \rho(z))' + (\log v(z))' \right)$$

$$\Rightarrow a_0(z) = c \exp \left(\frac{1}{2} \int \left((\log \rho(z))' + (\log v(z))' \right) dz \right)$$

$$\Rightarrow a_0(z) = a_0(0)(\rho(z)v(z))^{1/2}$$

Chapter 3

A Stability Result for Estimation of the Source and Reflectivity

3.1   Introduction

In his 1954 thesis, Enders Robinson stated that "the seismic trace is the response of the system consisting of the earth and recording apparatus to the impulsive source, the explosion." Further, he said "in the final analysis one is interested in the various components of this total response; for example, one wishes to separate components of reflected energy from those of non-reflected energy" [42]. One of the most basic tasks in seismology (and the one which Robinson devoted much time to) is removal of the input energy source from the seismic data. In the last four decades, seismologists have devised numerous methods for estimating and removing the source from the seismic data.

Robinson, for example, suggested separating the "dynamic component" or wavelet shape from the "random components" (arrival times and strengths of the wavelets) in a method known today as predictive deconvolution. This method assumes the wavelet is minimum phase. Another technique for removing the energy source from the seismic data is homomorphic deconvolution. The seismic data is approximated by convolving the source wavelet with the impulse response. By Fourier transforming the data, the convolution becomes multiplication which is then replaced by addition when the log is taken. An assumed dichotomy of the frequency content between the source and reflectivity allows these quantities to be determined separately from the sum [61]. A third idea is to measure the direct wave and then try to figure the input energy source from this measurement [64].

In this chapter we suggest using inversion to estimate the energy source at the same time the earth parameters are being estimated. In seismic inversion one chooses a physical model to describe the propagation of waves through the medium and attempts to improve the fit of the model to the given data by successively updating the parameters which characterize the model. The problem of determining the velocity

and quasi-impulsive source via inversion is analyzed in papers by Bube et al. [7] and Sacks [46]. Lewis analyzed a similar problem. He assumed the earth is a constant density acoustic layered fluid with pressure measured at the earth's surface. The propagating waves are assumed to be plane waves and only primary reflections are recorded. He did not require the source to be impulsive but did require the background velocity in this linearized model to be constant. He showed that perturbations in the seismic data *stably* determine corresponding perturbations in the source and reflectivity [33].

Starting from the same model Lewis used, we sought to relax the unrealistic assumption of a constant background velocity medium. We give a second proof of the stability result for constant background velocities and prove that the normal operator is continuous with respect to velocity. Thus for background velocities which vary slowly and smoothly with depth (i.e., are close to constant), changes in the source and reflectivity must cause proportional changes in the seismic data. The ability to separately determine the two parameters improves with increasing slowness aperture.

Section 3.2 of this chapter details the mathematical model we consider. Section 3.3 provides the mathematical results, namely, a new proof of stability for the constant case and the theorem showing that the normal operator is continuous with respect to velocity. Finally, in Section 3.4 we apply inversion to $\tau - p$ transformed synthetic marine data generated with a variable background velocity model, estimating both the energy source and reflectivity functions. Four experiments are described. In each case the starting reflectivity is the same (zero) and the starting guess for the source is an isotropic Ricker wavelet with the correct peak frequency but incorrect temporal distribution. We use output least squares inversion to estimate the source and reflectivity. The four models differ in that the data for each successive experiment is defined over a smaller and smaller slowness aperture. In each case we are able to find parameters which allow the model to fit the data equally well. However, the original source and reflectivity parameters are only correctly recovered (including the time location of the source) when the data is defined over a sufficiently large slowness aperture.

3.2 Model and Problem Specifications

We model the earth as an acoustic fluid with constant density and depth-dependent variable sound velocity $c(z)$. The velocity is assumed to vary slowly on the scale

of a seismic wavelength and to determine the kinematics of wave propagation. The short-scale heterogeneities are modeled by the relative perturbation in the velocity, or the reflectivity, $r(z) = \delta c(z)/c(z)$. The source is assumed to be isotropic and to have point support. The layered medium assumption above (that the velocity depends only on depth) allows us to apply the Radon integral transform (or plane-wave decomposition) to the normal displacement gotten from the solution to the acoustic wave equation. Thus one can reduce the three-dimensional problem to a family of one-dimensional equations [60]. These equations are parametrized by slowness, p. By assuming a primaries only or single-scattering approximation and by using high-frequency asymptotics, we can write the convolutional equation for the seismogram,

$$S(t,p) = f(t) * \tilde{r}(t,p).$$

Here, "$*$" denotes convolution in time t, and f is the isotropic source. The expression for the reflectivity as a function of time, $\tilde{r}$ (or perturbation of the Green's function for the acoustic wave equation) is given by $\tilde{r}(t,p) \approx \int dz[A(z)r(z,p)]\delta(t - 2\tau)$ where A is the reflectivity amplitude from geometric optics. The vertical (plane-wave) velocity changes form when we Radon transform, becoming $v(z,p) = c(z)/\sqrt{1 - c^2(z)p^2}$. Similarly, the reflectivity $r(z,p)$ is now the relative perturbation in this transformed velocity. We may write the travel-time function

$$\tau(z,p) = \int_0^z d\zeta \frac{1}{v(\zeta,p)}.$$

We shall also need the inverse of the travel-time function, which we denote by $z(t,p)$; thus $\tau(z(t,p),p) = t$ and $z(\tau(\zeta,p),p) = \zeta$. This primaries-only, plane-wave, layered medium, constant density acoustic model is likely the simplest model of seismic wave propagation which one can use to describe real seismograms. For an example of its use in modeling seismic field data see the paper by Symes and Carazzone [58].

3.3 Mathematical Results

3.3.1 Background

In his thesis, Lewis analyzed the effect of small perturbations in the source and reflectivity on the seismogram, in the case that the background velocity is constant [33]. From the plane-wave convolutional model for the seismogram, we may write the perturbed seismogram as

$$\delta S(t,p) = \delta f(t) * \tilde{r}(t,p) + f(t) * \delta\tilde{r}(t,p). \tag{3.1}$$

Lewis showed that the perturbed seismogram determines both the perturbation in the source and the perturbation in the reflectivity uniquely. For instance, if the perturbation in the data is equal to zero, then the perturbations in both the source and reflectivity must also be identically zero. Obvious nonuniqueness due to scale ambiguities between the two parameters (source and reflectivity) is ruled out by an added constraint which fixes the scale of one of the two quantities.

Further, he showed that perturbations in the seismic data *stably* determine corresponding perturbations in the source and reflectivity. This determination for source perturbations is valid within the passband of the source itself. Similarly, determination of reflectivity perturbations is constrained to a corresponding spatial frequency passband for the reflectivity. The stability improves as the range of slowness values increases. Conversely, as the range is reduced to a single trace, the ability to separately determine the perturbations in the source and reflectivity simultaneously is lost. The numerical examples in Section 3.4 show that although it may be easy to find model parameters which explain a small aperture (or single trace) set of data, these parameters will not necessarily be the ones which generated the data originally.

As is realistic, in this work Lewis assumed that both the source and reflectivity have compact support in time. Thus, the source cannot mathematically be truly band-limited. It is reasonable, however, to assume that outside of some frequency band the source has small enough contribution to be considered negligible and ignored [52].

In an effort to generalize the above result to the case in which the background velocity need not be constant, we have arrived at a new proof of Lewis' result (Theorem 3.1 below). We show in Theorem 3.2 that the normal operator corresponding to the convolutional forward model for the seismogram is continuous in the velocity. Thus, having established stability at one point in velocity model space, namely for constant background velocity media, the continuity result implies that in a small neighborhood of this point (i.e., for background velocities which don't stray too far from constant), the seismic data stably determines corresponding perturbations in the source and reflectivity.

3.3.2 Notation

In the following theory section we make use of the symbols "$\hat{g}$", "$\tilde{g}$", and "$\bar{g}$" where g will generally be the source or reflectivity functions. The symbol "$\hat{g}$" is the Fourier transform of g with respect to either time or depth depending on the domain of the function. The symbol "$\tilde{g}$" implies a change of variables from depth to time has been applied, and "$\bar{g}$" is the conjugate of g.

3.3.3 Theory

Given the nature of the statistical deconvolution theory as described in the introduction, it is hardly surprising that a measure of "whiteness" turns up in the stability result to follow. For an interval $[\zeta_{min}, \zeta_{max}] \subset \mathbb{R}$ and $0 < \Delta\zeta < \zeta_{max} - \zeta_{min}$, define for $r \in L^2(\mathbb{R}), r \neq 0$

$$W(\zeta_{min}, \zeta_{max}, \Delta\zeta; r) = \tag{3.2}$$

$$\inf\left\{ \frac{\frac{1}{\zeta_+ - \zeta_-}\int_{\zeta_-}^{\zeta_+} |\hat{r}|^2}{\|r\|_{L^2(\mathbb{R})}^2} : \zeta_{min} \le \zeta_- < \zeta_+ \le \zeta_{max}, \frac{1}{2}\Delta\zeta \le \zeta_+ - \zeta_- \le \Delta\zeta \right\}$$

That is, W measures the uniformity of distribution of Fourier components averaged over frequency intervals of length roughly $\Delta\zeta$. Note that $W > 0$ for any r of compact support. Evidently if f (and perhaps r) are band-limited, it is only possible to determine band-limited information about $\delta f, \delta r$ from the perturbational relation 3.1. For $\eta > 0$, we define an η-passband for $f \in L^2(\mathbb{R})$ to be a symmetric frequency interval $[-\omega_{max}, -\omega_{min}] \cup [\omega_{min}, \omega_{max}]$ over which

$$|\omega \hat{f}(\omega)| \ge \eta \|f\|_{L^2(\mathbb{R})} \tag{3.3}$$

Theorem 3.1 Suppose that $0 \le p_{min} \le p_{max}$, $0 < \omega_{min} < \omega_{max}$, $\eta > 0$, and $\Omega_f = [-\omega_{max}, -\omega_{min}] \cup [\omega_{min}, \omega_{max}]$ is an η-passband for $f \in L^2[0,T]$. Fix the size of the source, i.e., $\|\hat{f}\|_{L^2(\Omega_f)} = 1$. Suppose also that $r \in L^2[0, Z]$. Then for any $\delta f \in L^2[0,T]$ which satisfies the linearization of the fixed size constraint above, namely that,

$$\int_{\Omega_f} d\omega \overline{\widehat{\delta f}(\omega)} \hat{f}(\omega) = 0 \tag{3.4}$$

and $\delta r \in L^2[0, Z]$,

$$\left\|\widehat{\delta f}\right\|_{L^2(\Omega_f)} + \left\|\widehat{\delta \tilde{r}}\right\|_{L^2(\Omega_f \times [p_{min}, p_{max}])} \leq K \left\|\delta S\right\|_{L^2([0,T] \times [p_{min}, p_{max}])}$$

where K depends on $p_{min}, p_{max}, c, \omega_{min}, \omega_{max}, \eta$, and

$$W := W\left(\frac{\omega_{min}}{v(p_{max})}, \frac{\omega_{max}}{v(p_{min})}, \frac{\omega_{min}}{2}\left(1 + \frac{v(p_{max})}{v(p_{min})}\right); r\right)$$

Proof Set $v(p) = v(z, p) = (c^{-2} - p^2)^{-\frac{1}{2}}$. Then $\hat{\tilde{r}}(\omega, p) = v(p)\hat{r}(\omega/v(p))$. Write $\zeta(\omega, p) = \omega/v(p)$. Then Fourier transformation of the expression for δS (expression 3.1) and a little algebra yield

$$\frac{\overline{\hat{r}}(\zeta)}{v(p)} \frac{\widehat{\delta S}(\omega, p)}{\hat{f}(\omega)} = \left[\frac{\widehat{\delta f}(\omega)}{\hat{f}(\omega)}\right] |\hat{r}(\zeta)|^2 + \overline{\hat{r}}(\zeta)\widehat{\delta r}(\zeta) \tag{3.5}$$

The map $p \to \zeta(\omega, p)$ is smooth and invertible. Viewed as an identity in ω and ζ, expression 3.5 holds over the ζ interval $[\zeta_{min}(\omega), \zeta_{max}(\omega)]$, where $\zeta_{min}(\omega) = \omega/v(p_{max})$, and $\zeta_{max}(\omega) = \omega/v(p_{min})$. A calculation shows that

$$[\zeta_{min}(\sigma\omega), \zeta_{max}(\omega)] \subset [\zeta_{min}(\omega), \zeta_{max}(\omega)] \cap [\zeta_{min}(\sigma\omega), \zeta_{max}(\sigma\omega)] \tag{3.6}$$

if $1 \leq \sigma \leq \frac{1}{2}\left(1 + \frac{v(p_{max})}{v(p_{min})}\right)$. Therefore we can subtract equation 3.5 at the point $\sigma\omega$ from the same expression at the point ω so long as ζ remains in the range $\zeta \in [\zeta_{min}(\sigma\omega), \zeta_{max}(\omega)]$. We obtain

$$\overline{\hat{r}}(\zeta)\left[\frac{\widehat{\delta S}(\omega, p(\omega, \zeta))}{\hat{f}(\omega)v(p(\omega, \zeta))} - \frac{\widehat{\delta S}(\sigma\omega, p(\sigma\omega, \zeta))}{\hat{f}(\sigma\omega)v(p(\sigma\omega, \zeta))}\right] = \left[\frac{\widehat{\delta f}(\omega)}{\hat{f}(\omega)} - \frac{\widehat{\delta f}(\sigma\omega)}{\hat{f}(\sigma\omega)}\right]|\hat{r}(\zeta)|^2 \tag{3.7}$$

Integrate both sides of 3.7 to get

$$\left|\frac{\widehat{\delta f}(\omega)}{\hat{f}(\omega)} - \frac{\widehat{\delta f}(\sigma\omega)}{\hat{f}(\sigma\omega)}\right|$$

$$= \left| \left(\int_{\zeta_{min}(\sigma\omega)}^{\zeta_{max}(\omega)} d\zeta \, |\hat{r}(\zeta)|^2 \right)^{-1} \int_{\zeta_{min}(\sigma\omega)}^{\zeta_{max}(\omega)} d\zeta \, \overline{\hat{r}}(\zeta) \left[\frac{\widehat{\delta S}(\omega, p(\omega, \zeta))}{\hat{f}(\omega) v(p(\omega, \zeta))} - \frac{\widehat{\delta S}(\sigma\omega, p(\sigma\omega, \zeta))}{\hat{f}(\sigma\omega) v(p(\sigma\omega, \zeta))} \right] \right|$$

$$\leq \frac{K}{\eta W \|r\|_{L^2[0,Z]}} \left[\left(\int_{p_{min}}^{p_{max}} dp |\widehat{\delta S}(\omega, p)|^2 \right)^{\frac{1}{2}} + \left(\int_{p_{min}}^{p_{max}} dp \, |\widehat{\delta S}(\sigma\omega, p)|^2 \right)^{\frac{1}{2}} \right] \quad (3.8)$$

The expression 3.8 is derived from the line above it through application of the "pass-band" and "whiteness" hypotheses 3.2 and 3.3 as well as the Cauchy-Schwarz and triangle inequalities. Thus, we arrive at the expression,

$$\left| \frac{\widehat{\delta f}(\omega)}{\hat{f}(\omega)} - \frac{\widehat{\delta f}(\sigma\omega)}{\hat{f}(\sigma\omega)} \right| \leq \frac{K}{\eta W \|r\|_{L^2[0,Z]}} \|\delta S\|_{L^2([0,T] \times [p_{min}, p_{max}])} \quad (3.9)$$

K stands for a quantity depending only on c, p_{min}, and p_{max}, which may vary from expression to expression. Denote by N the smallest integer greater than

$$\frac{\omega_{max} - \omega_{min}}{\frac{1}{2} \left(1 + \frac{v(p_{max})}{v(p_{min})} \right)}$$

Expression 3.9 holds for an ω in one of these N intervals. To generalize to any $\omega \in [\omega_{min}, \omega_{max}]$ we employ a telescoping sum

$$\left| \frac{\widehat{\delta f}(\omega)}{\hat{f}(\omega)} - \frac{\widehat{\delta f}(\omega_{min})}{\hat{f}(\omega_{min})} \right| = \left| \sum_{n=1}^{N} \frac{\widehat{\delta f}(\sigma^n \omega_{min})}{\hat{f}(\sigma^n \omega_{min})} - \frac{\widehat{\delta f}(\sigma^{n-1} \omega_{min})}{\hat{f}(\sigma^{n-1} \omega_{min})} \right|$$

$$\leq \frac{NK}{\eta W \|r\|_{L^2[0,Z]}} \|\delta S\|_{L^2([0,T]) \times [p_{min} p_{max}])} \quad (3.10)$$

Evidently the same inequality holds for the negative part of the passband Ω_f, with ω_{min} replaced by $-\omega_{min}$. Set $k = \widehat{\delta f}(\omega_{min})/\hat{f}(\omega_{min})$. Because $\widehat{\delta f}$ is orthogonal to $\hat{f}$ over Ω_f,

$$\left\| \widehat{\delta f} \right\|_{L^2(\Omega_f)}^2 = \int_{\Omega_f} d\omega \, \overline{\widehat{\delta f}}(\omega) \left(\widehat{\delta f}(\omega) - k \hat{f}(\omega) \right)$$

$$= \int_{\omega_{min}}^{\omega_{max}} d\omega \, \overline{\widehat{\delta f}}(\omega) \hat{f}(\omega) \left(\frac{\widehat{\delta f}(\omega)}{\hat{f}(\omega)} - \frac{\widehat{\delta f}(\omega_{min})}{\hat{f}(\omega_{min})} \right) + \int_{-\omega_{max}}^{-\omega_{min}} d\omega \, \overline{\widehat{\delta f}}(\omega) \hat{f}(\omega) \left(\frac{\widehat{\delta f}(\omega)}{\hat{f}(\omega)} - \frac{\widehat{\delta f}(\omega_{min})}{\hat{f}(\omega_{min})} \right)$$

$$\leq \left\|\widehat{\delta f}\right\|_{L^2(\Omega_f)} \left\|\hat{f}\right\|_{L^2(\Omega_f)} \frac{NK}{\eta W \left\|r\right\|_{L^2[0,Z]}} \left\|\delta S\right\|_{L^2([0,T]\times[p_{\min},p_{\max}])}$$

The last inequality above came from substitution of the expression 3.10. Thus we have established the required bound for $\widehat{\delta f}$. The corresponding bound for $\delta \tilde{r}$ follows from the definition of δS and the passband inequality 3.3.

$\square$

Corollary 3.1 For any spatial frequency interval (ζ_-, ζ_+) contained in the interior of

$$\left[\frac{-\omega_{\max}}{v(p_{\min})}, \frac{-\omega_{\min}}{v(p_{\max})}\right] \quad \cdot \left[\frac{\omega_{\min}}{v(p_{\max})}, \frac{\omega_{\max}}{v(p_{\min})}\right] ,$$

there exists a constant $K(\zeta_-, \zeta_+)$ depending also on the quantities mentioned in the theorem, so that

$$\left\|\widehat{\delta r}\right\|_{L^2[\zeta_-,\zeta_+]} \leq K \left\|\delta S\right\|_{L^2([0,T]\times[p_{\min},p_{\max}])}$$

Note. The constant K in the corollary is not uniform in ζ_-, ζ_+ but depends on the size of the neighborhood of $[\zeta_-, \zeta_+]$ contained in the band defined above.

Theorem 3.2 The conclusion of Theorem 1 continues to hold, under the same hypotheses, for $c \in C^1[0, Z]$, provided that $\|dc/dz\|_{C^0[0,Z]}$ is sufficiently small. Specifically, there exist $\sigma, K > 0$ depending on the same quantities as in the statement of Theorem 1 so that

$$\left\|\frac{dc}{dz}\right\|_{C^0[0,Z]} \leq \sigma$$

implies that for arbitrary $\delta f \in L^2[0, T]$, $\delta r \in L^2[0, Z]$,

$$\left\|\widehat{\delta f}\right\|_{L^2(\Omega_f)} + \left\|\widehat{\delta \tilde{r}}\right\|_{L^2(\Omega_f \times [p_{\min},p_{\max}])} \leq K \left\|\delta S\right\|_{L^2(\mathbb{R}\times[p_{\min},p_{\max}])}$$

Proof Evidently, the assertion of the theorem boils down to a statement about

$$\|\delta S\|^2_{L^2(\mathbb{R}\times[p_{\min},p_{\max}])} = \left\langle \begin{pmatrix} \delta f \\ \delta r \end{pmatrix}, N \begin{pmatrix} \delta f \\ \delta r \end{pmatrix} \right\rangle_{L^2[0,T]\times L^2[0,Z]} \tag{3.11}$$

where N is the *normal operator*.

Explicitly,

$$N \begin{pmatrix} \delta f \\ \delta r \end{pmatrix} = \begin{pmatrix} N_{rr}\delta f + N_{fr}\delta r \\ N_{rf}\delta f + N_{ff}\delta r \end{pmatrix}$$

where

$$N_{rr}\delta f(t) = \int_{p_{min}}^{p_{max}} p\,dp \int ds_1 \int ds_2\, r(z(s_1,p))\, r(z(s_2,p))\, \delta f(t + s_1 - s_2)$$

$$N_{fr}\delta r(t) = \int_{p_{min}}^{p_{max}} p\,dp \int ds_1 \int ds_2\, f(s_1)\, r(z(s_2,p))\, \delta r(z(t + s_1 - s_2,p))$$

$$N_{rf}\delta f(z) = \int_{p_{min}}^{p_{max}} p\,dp\frac{1}{v(z,p)} \int ds_1 \int ds_2\, f(s_1)\, r(z(s_2,p))\, \delta f(\tau(z,p) + s_1 - s_2)$$

$$N_{ff}\delta r(z) = \int_{p_{min}}^{p_{max}} p\,dp\frac{1}{v(z,p)} \int ds_1 \int ds_2\, f(s_1)\, f(s_2)\, \delta r(z(\tau(z,p) + s_1 - s_2,p))$$

$N = N[c]$ is a self-adjoint operator on the Hilbert space $L^2[0,T] \times L^2[0,Z]$ for each positive smooth velocity profile $c(z)$. We will show that N is *continuous* in c, in an appropriate sense. Then the conclusion of the theorem will follow from the perturbation theory of bounded self-adjoint operators [28]. All four components of N may be treated the same way, so we show the calculations explicitly only for N_{rf}. Denote by $N_{rf}^0 = N_{rf}[c(0)]$ the operator N_{rf} for the constant velocity $c \equiv c(0)$. Then a change of variables of integration yields

$$\left(N_{rf}\delta f - N_{rf}^0\delta f\right)(z) = \int_{p_{min}}^{p_{max}} p\,dp \left\{ \frac{1}{v(z,p)} \int\int dz_1 ds\, r(z_1)\, \delta f(s)f(s + \tau(z_1,p) - \tau(z,p)) \right.$$

$$\left. - \frac{1}{v_0(p)} \int\int dz_1 ds\, r(z_1)\, \delta f(s)f(s + \tau_0(z_1,p) - \tau_0(z,p)) \right\}$$

where the zero subscripts denote quantities associated with the constant velocity $c(0)$. This last is

$$= \int_{p_{min}}^{p_{max}} p\,dp \left\{ \left(\frac{1}{v(z,p)} - \frac{1}{v_0(p)} \right) \int\int dz_1 ds\, r(z_1)\, \delta f(s)\, f(s + \tau(z_1,p) - \tau(z,p)) \right.$$

$$\left. + \frac{1}{v_0(p)} \int\int dz_1 ds\, r(z_1)\delta f(s)\, [f(s + \tau(z_1,p) - \tau(z,p)) - f(s + \tau_0(z_1,p) - \tau_0(z,p))] \right\}$$

Now

$$f(s + \tau(z_1, p) - \tau(z, p)) - f(s + \tau_0(z_1, p) - \tau_0(z, p))$$

$$= \int_s^{s+(\tau(z_1,p)-\tau(z,p))-(\tau_0(z_1,p)-\tau_0(z,p))} d\sigma f'(\sigma + \tau_0(z_1, p) - \tau_0(z, p))$$

$$= -\frac{1}{(z - z_1)v_0(p)p} \left\{ \frac{\partial}{\partial p} \int_s^{s+(\tau(z_1,p)-\tau(z,p))-(\tau_0(z_1,p)-\tau_0(z,p))} d\sigma f(\sigma + \tau_0(z_1, p) - \tau_0(z, p)) \right\}$$

$$+ \left(\int_z^{z_1} dz' p(v(z', p) - v_0(p)) \right) f(s + \tau(z_1, p) - \tau(z, p))$$

Substituting this last expression into the integral and integrating by parts with respect to p, we obtain a lengthy expansion for the kernel of the Hilbert-Schmidt operator N_{rf}, each term in which consists of

1. a factor of $r(z_1)$

2. either $f(\cdot)$ or $\frac{1}{z-z_1} \int_I f$, where the interval of integration I has length $= O(|z - z_1|)$

3. a coefficient function of z, z_1 and p which is uniformly bounded by a function of p_{min}, p_{max} and $\|c - c(0)\|_{C^1[0,Z]}$

hence in turn by $\|dc/dz\|_{C^0[0,Z]}$. Thus the Hilbert-Schmidt norm, or $N_{rf} - N_{rf}^0$ is bounded by

$$K \|r\|_{L^2[0,Z]} \|f\|_{L^2[0,T]} \left\| \frac{dc}{dz} \right\|_{C^0[0,Z]}$$

where K is a function of $Z, \|c\|_{C^1[0,Z]}$, p_{min}, and p_{max}. The other three components may be treated the same way, so

$$\left\| N - N^0 \right\| \leq K \|r\|_{L^2[0,Z]} \|f\|_{L^2[0,T]} \left\| \frac{dc}{dz} \right\|_{C^0[0,Z]}$$

To link this continuity result to Theorem 3.1, use expression 3.11 which ties the normal operator to the square of the norm of the perturbed seismogram. Recall the first part of the conclusion of Theorem 3.1 gives that $\left\| \widehat{\delta f} \right\|_{L^2(\Omega_f)} \leq K \|\delta S\|_{L^2([0,T] \times [p_{min}, p_{max}])}$ which may now be expressed as:

$$N^0 \geq \begin{pmatrix} \varepsilon E_{\Omega f} & 0 \\ 0 & 0 \end{pmatrix}$$

where $\varepsilon > 0$ and $E_{\Omega f}$ is the operator of Fourier transformation (of f), followed by multiplication by the bandpass filter (1 inside Ωf, 0 outside), followed by the inverse Fourier transform. The continuity estimate above now implies the same bound for N:

$$N \geq \begin{pmatrix} \varepsilon E_{\Omega f} & 0 \\ 0 & 0 \end{pmatrix}$$

where ε is uniform over $\|f\|_{L^2[0,T]} = 1$ and L^2-bounded sets of $r \in L^2[0, Z]$ subject to the "pseudowhiteness" constraint expressed in Theorem 3.1. Thus the result has been proved for the δf term, and the complete conclusion of Theorem 3.2 follows from the normal equations as before.

$\square$

To summarize, we have shown that perturbations in the source and reflectivity functions are both well-determined by the corresponding perturbations of the seismogram, within their passbands, so long as the rate of change of velocity with depth is sufficiently small, and so long as the reference reflectivity has a "whiteness" property. These conditions are merely sufficient for "well-posedness within the passband" of the linearized problem. Their necessity is far from obvious. Moreover, our results imply only that a reasonable misfit criterion, suitably regularized, will be convex near the global minimizer for noise-free data. In particular our result says little directly about the global behavior of misfit functions or algorithms to minimize them.

3.4 A Numerical Example

To test the theoretical results discussed above numerically, we performed experiments on $\tau - p$ transformed synthetic data generated from real data. The examples shown here were based on data provided by Exxon Production Research Company. All tests were done on a Sun Sparcstation 2 using the Differential Semblance Optimization package, a seismic simulation and inversion code under development at Rice University.

We began with one common-midpoint data gather taken from a marine seismic survey. The Radon transform was applied to yield 48 plane-wave traces with slowness values ranging from $p_{min} = .1158$ ms/m to $p_{max} = .36468$ ms/m. We were also given an estimate of the anisotropic air gun source in the form of a 31-term Legendre expansion in slowness. The coefficients in the series were functions of time only.

The procedure used to generate the target model for our experiment was the following. The target isotropic source is the first term of the 31-term source estimate

(which represents the isotropic component of the air gun source). This isotropic source has been filtered to have a peak frequency of roughly 15 Hz and a peak in time at 110 ms. To generate a realistic target reflectivity, we inverted the plane-wave field data for the reflectivity using an inversion-estimated background velocity. The output of the (least squares) inversion was an estimate of the reflectivity (the relative perturbation in the velocity in this case).

For the experiments reported here (in each case, determination of both the source and reflectivity) we chose the variable background velocity shown in Figure 3.1. We generated four new sets of data (each with a different slowness aperture) for our synthetic experiments using the reflectivity from the field data inversion as input. The seismic data was also filtered by the 15 Hz Ricker wavelet and consisted of 13 traces each of about 3 seconds time duration.

We successively updated the source and reflectivity estimates by minimizing the mean squared difference between the actual data and the data predicted by our model (Output Least Squares inversion). The method we used (coordinate search or alternation) requires the source parameters be fixed, and the reflectivities estimated by output least squares inversion. Then the reflectivities are updated and held fixed and the source parameters estimated by OLS inversion. The source is updated and this cycle is repeated until convergence.

Alternation, although notoriously inefficient, is attractive for initial experiments because it requires only successive solution of simple linear least-squares problems. Obviously, quasi-Newton methods could be applied to the problem and would likely reduce the number of iterations dramatically.

The initial estimate of the reflectivity used for the inversions was $r = 0$. For the initial source estimate, we chose a Ricker wavelet also with peak frequency of 15 Hz but which had its peak centered in time at 0 ms (Figure 3.2). The source location was of interest to us in this experiment. In practice, one might reasonably expect to be able to estimate the power spectrum of the source. However, one would not expect to be able to guess at the location of the peak in time (i.e., the phase). Each inversion round included an estimation of the source and an estimation of the reflectivity.

In experiment 1 we attempted to find source and reflectivity parameters to explain the data shown in Figure 3.3. This synthetic data was generated with the full slowness aperture range of the original plane wave data, namely from $p_{min} = .1158$ ms/m to $p_{max} = .36468$ ms/m. The $\tau - p$ transformed seismogram used in experiment 2 is shown in Figure 3.6. For this experiment, the slowness interval between consecutive

traces was cut in half, but the total number of traces remained the same . The resulting aperture range was from $p_{min} = .1158$ ms/m to $p_{max} = .2402$ ms/m. The slowness interval between neighboring traces was halved once more for experiment 3, yielding a slowness range from $p_{min} = .1158$ ms/m to $p_{max} = !1780$ ms/m. The corresponding data is displayed in Figure 3.9. Finally, a slowness interval of 0 ms/m was used, i.e., all thirteen traces in experiment 4 were the same. The data to be matched is the single-trace plane-wave seismogram of Figure 3.12.

Figures 3.4, 3.7, 3.10, and 3.13 display the target isotropic source (dashed line) and an inversion-estimated source (solid line) from the corresponding data (3.3, 3.6, 3.9, and 3.12 respectively). Similarly, Figures 3.5, 3.8, 3.11, and 3.14 display the target reflectivity (dashed line) and estimated reflectivity (solid line) for each experiment. Recall that using the convolutional model for the seismogram and determining both the source and reflectivity, a scale ambiguity exists (which is why we fix the size of the source function in the theoretical results). Thus, one may scale the source up by a constant α and the reflectivity down by the constant $1/\alpha$ and fit the data equally as well as one would without the scaling. Multiplication by the constant -1 is, of course, allowed as well. Thus, in each experiment, the graphs were designed to show the best results for that case (which may have included polarity reversal). In experiments 1 and 2, the source and reflectivity inversion results were scaled by a factor of -1, whereas in experiments 3 and 4 they were not. The inversion estimates plotted for the four different experiments correspond to the point at which the inversion-estimated source and reflectivity had reduced the root mean squared error to 7% of the data norm. Each time we narrowed the slowness aperture, the convergence rate for the alternation algorithm progressively worsened. For example, to reduce the error to 7% took only eighteen rounds of alternation with the full aperture range of data (experiment 1). The half aperture range (experiment 2) took forty rounds; and finally, in experiment 3, one-hundred and fifty rounds of alternation were required to reduce the rms error to 7% of the data norm. The degenerate case of data with only a single trace converged to less than 10% error in only two rounds of alternation. We note, however, that the experiments with a wider range of slowness values allow determination of the reflectivity and correctly locate the source peak at 110 ms. As the slowness range narrows, the inversion scheme is no longer able to move the initial source guess to its target location, although the fit to data is comparable across experiments. The ability to separately determine the two parameters has been lost.

3.5 Conclusion

In order to estimate the mechanical parameters which describe a section of the subsurface, seismologists must take into account the energy they introduce into the ground as an imaging device. Rather than removing the energy source from the data, one can invert for this source while simultaneously estimating the earth parameters. In this chapter we examine a simple seismogram model, namely the plane-wave convolutional model derived from the constant density, variable sound velocity acoustic wave equation. We invert for the energy source and a high-frequency perturbation of the velocity (or reflectivity). Theoretically we find that for slowly varying background velocities, the corresponding seismic data stably determines perturbations in the source and reflectivity. Numerically we describe four inversion experiments performed on $\tau - p$ transformed synthetic data generated from real marine data. The initial source guess was located at the time origin whereas the true source was peaked at 110ms. Although we are able to determine both parameters quite accurately for data defined over a wide range of slownesses, as this range narrows to the single trace limit, we can no longer separately recover the original source (and its correct time location) and the reflectivity.
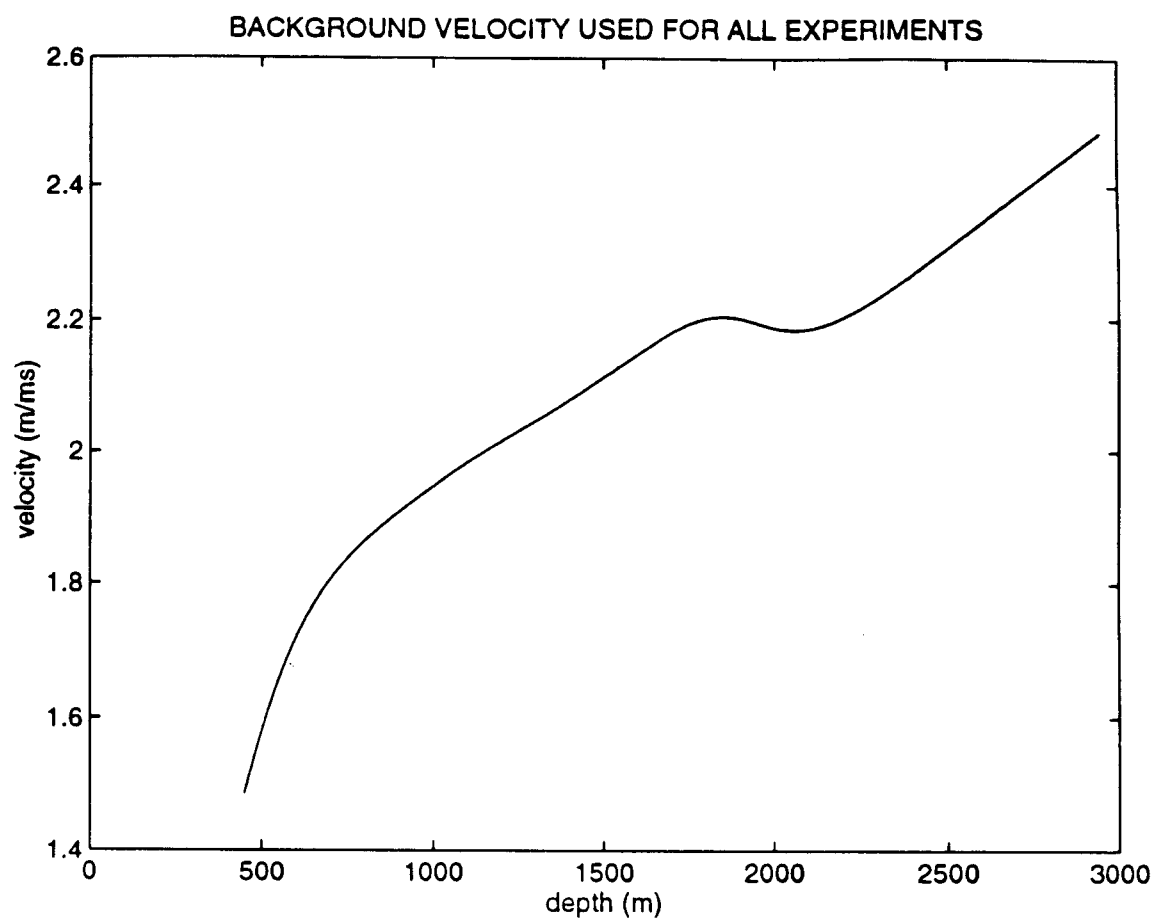
Figure 3.1 Variable background velocity model used to generate the seismic data shown in Figures 3.3, 3.6, 3.9, and 3.12.
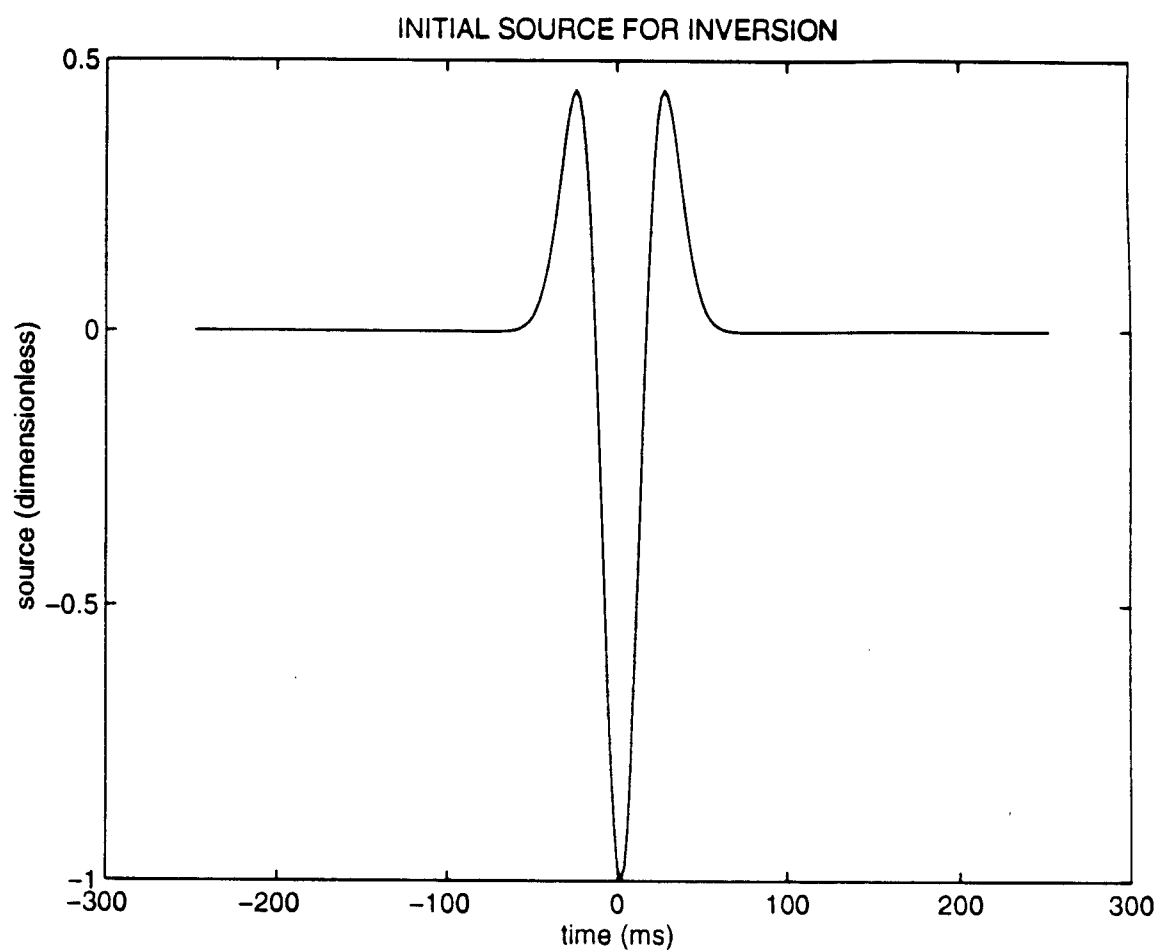
Figure 3.2 Initial guess for the source, an inverted
15 Hz Ricker wavelet centered at 0 ms.
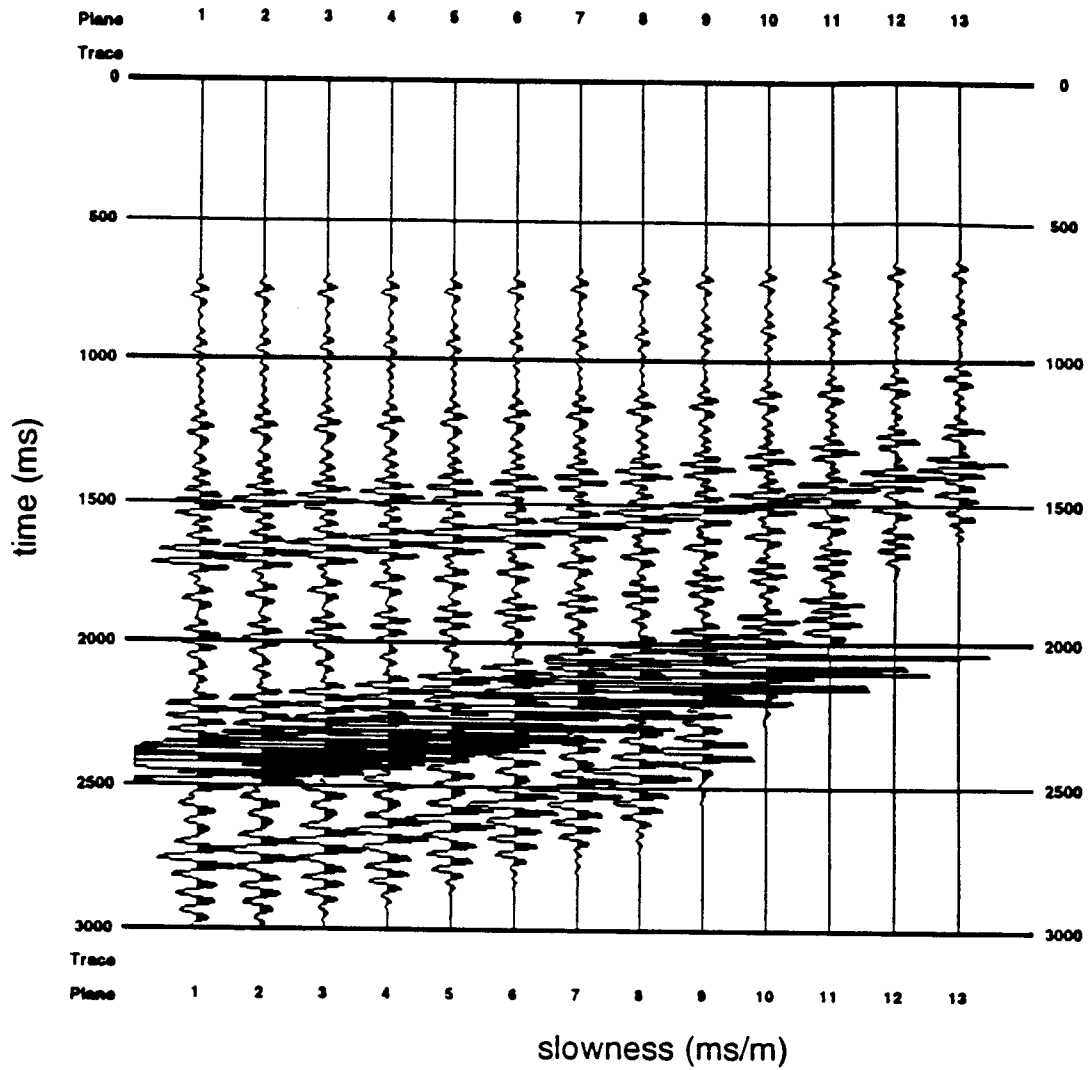
Figure 3.3 Synthetic $\tau - p$ transformed seismic data generated from the variable background velocity shown in Figure 3.1 and the isotropic source in Figure 3.4 (dashed line). The slowness aperture for the data is $p_{min} = .1158$ ms/m to $p_{max} = .36468$ ms/m.
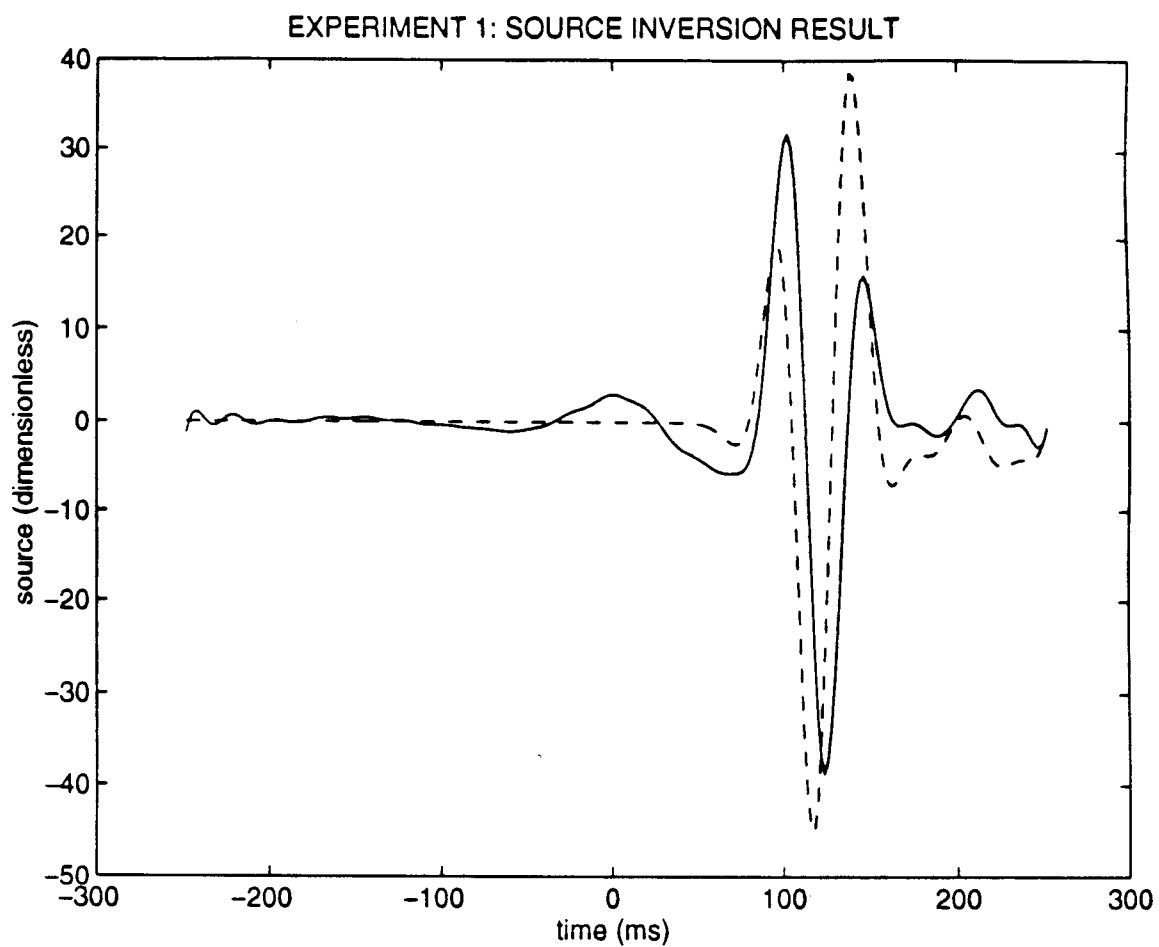
Figure 3.4 Source inversion result at alternation round 18 for the data shown in Figure 3.3. The rms error/data norm = 7.5%. Solid line: estimated source (scaled). Dashed line: target source.

EXPERIMENT 1: REFLECTIVITY INVERSION RESULT
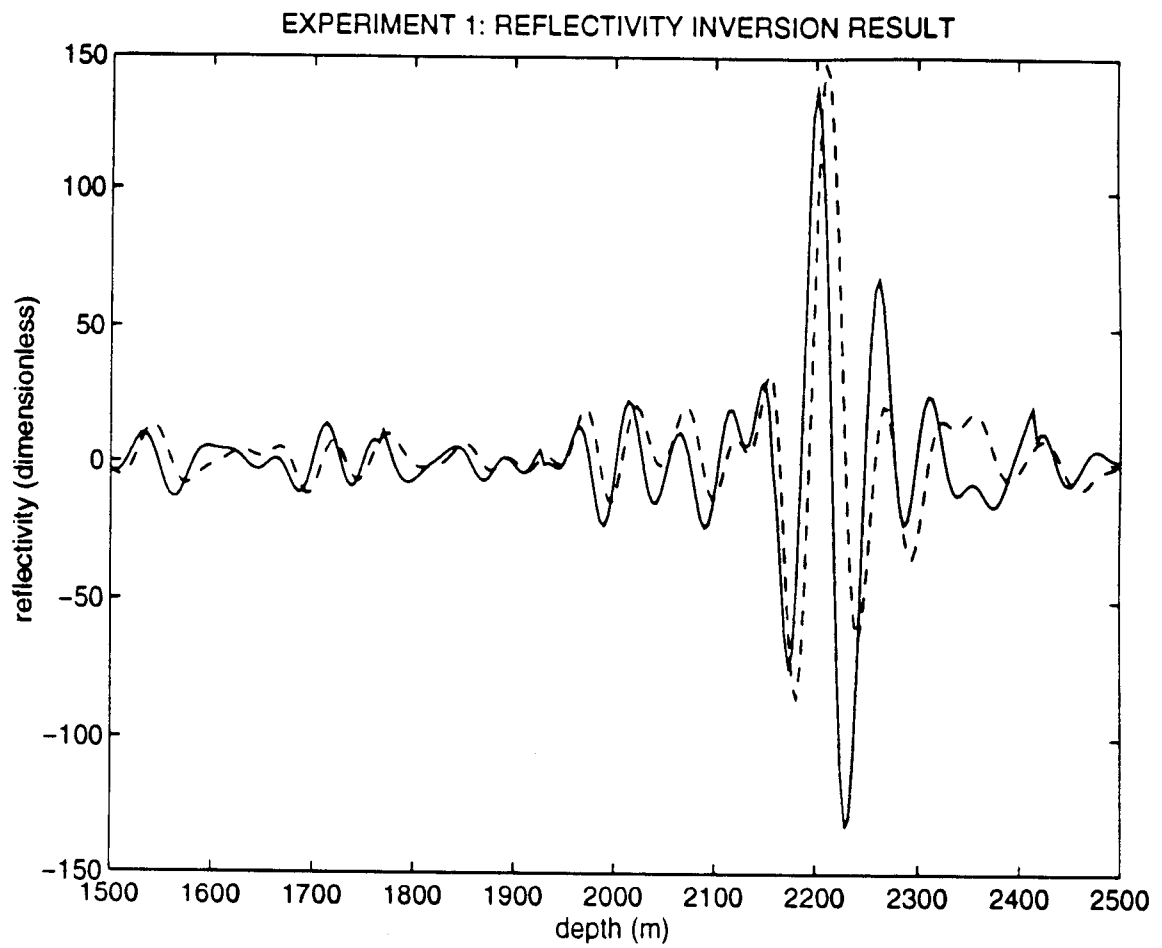
Figure 3.5 Reflectivity inversion result at alternation round 18 for the data shown in Figure 3.3. The rms error/data norm = 7%. Solid line: estimated reflectivity (scaled). Dashed line: target reflectivity.
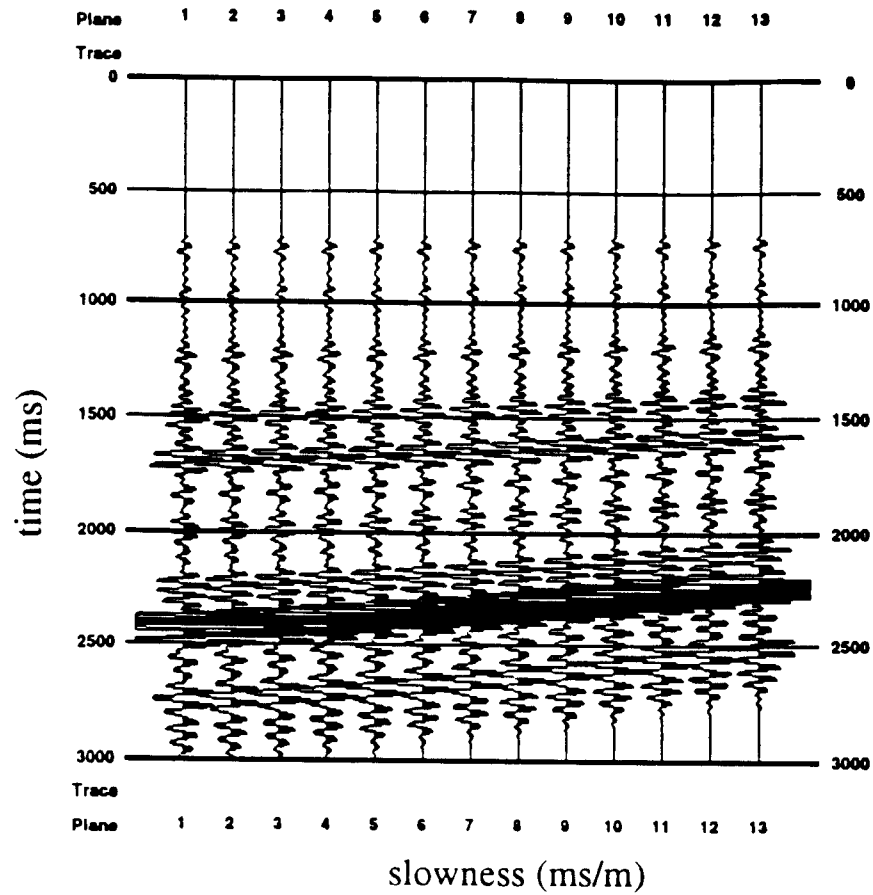
Figure 3.6 Synthetic $\tau - p$ transformed seismic data generated from the variable background velocity shown in Figure 3.1 and the isotropic source in Figure 3.7 (dashed line). The slowness aperture for the data is $p_{min} = .1158$ ms/m to $p_{max} = .2402$ ms/m.
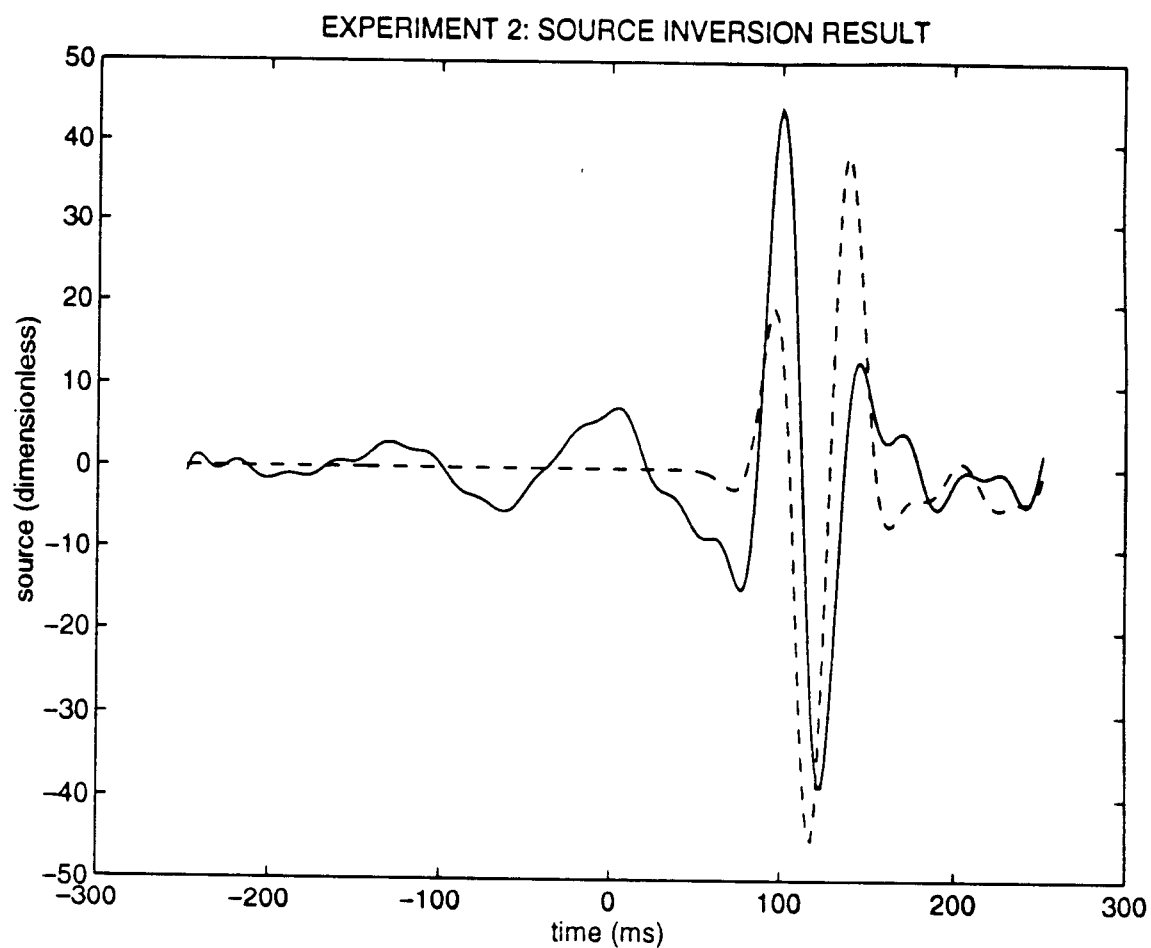
Figure 3.7 Source inversion result at alternation round 40 for the data shown in Figure 3.6. The rms error/data norm = 7%. Solid line: estimated source (scaled). Dashed line: target source.
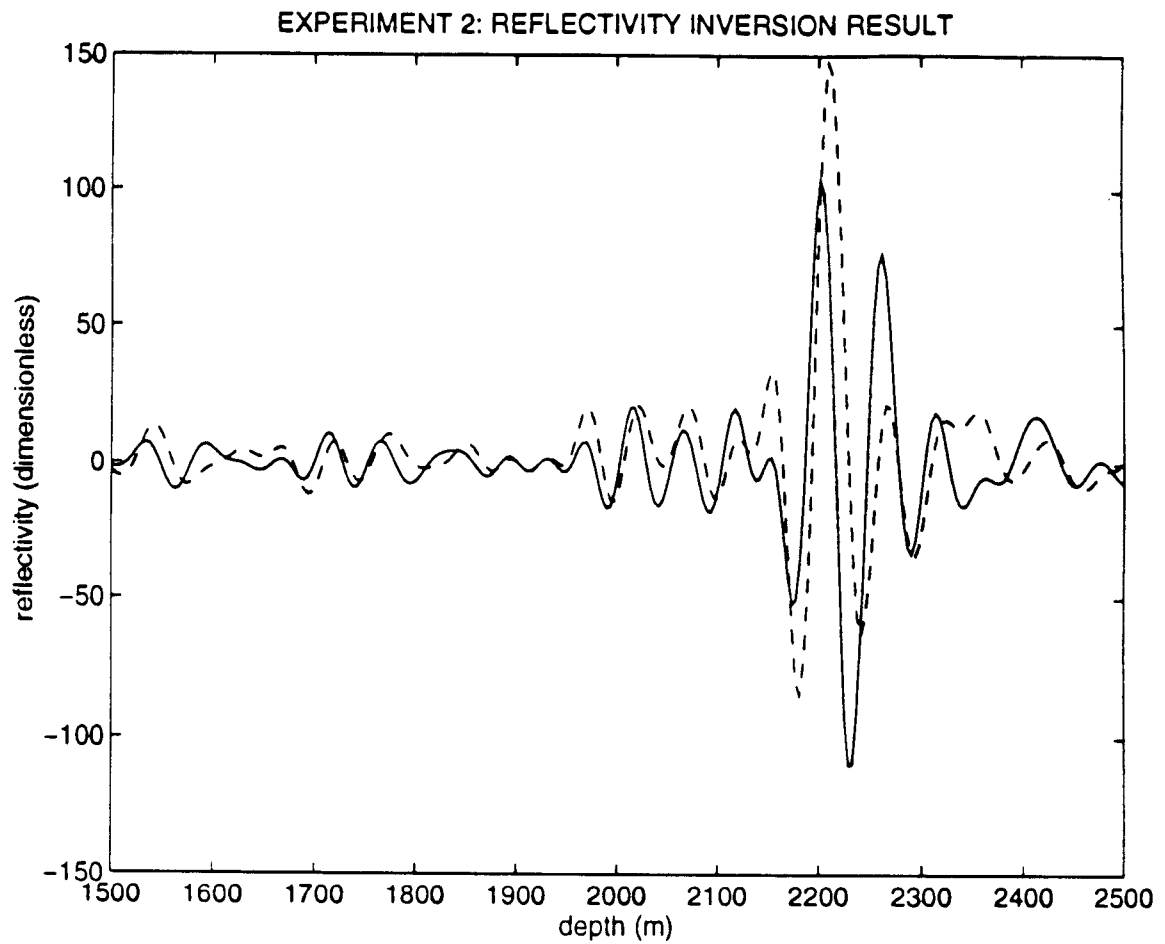
Figure 3.8 Reflectivity inversion result at alternation round 40 for the data shown in Figure 3.6. The rms error/data norm = 7%. Solid line: estimated reflectivity (scaled). Dashed line: target reflectivity.
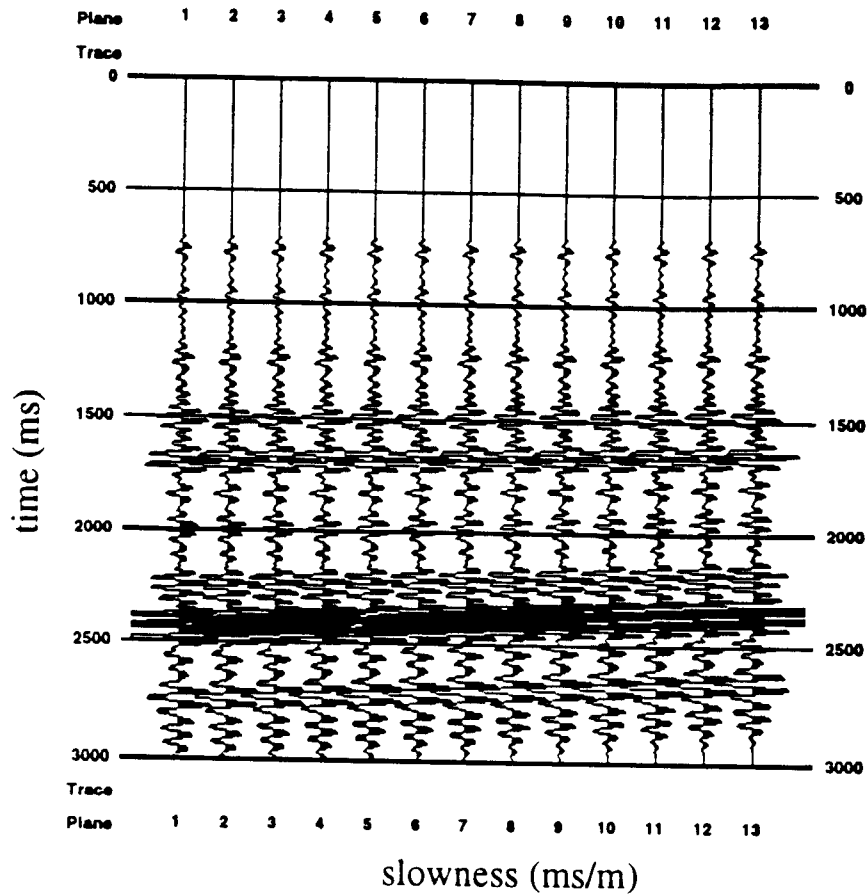
Figure 3.9 Synthetic $\tau - p$ transformed seismic data generated from the variable background velocity shown in Figure 3.1 and the isotropic source in Figure 3.10 (dashed line). The slowness aperture for the data is $p_{min} = .1158$ ms/m to $p_{max} = .1780$ ms/m.
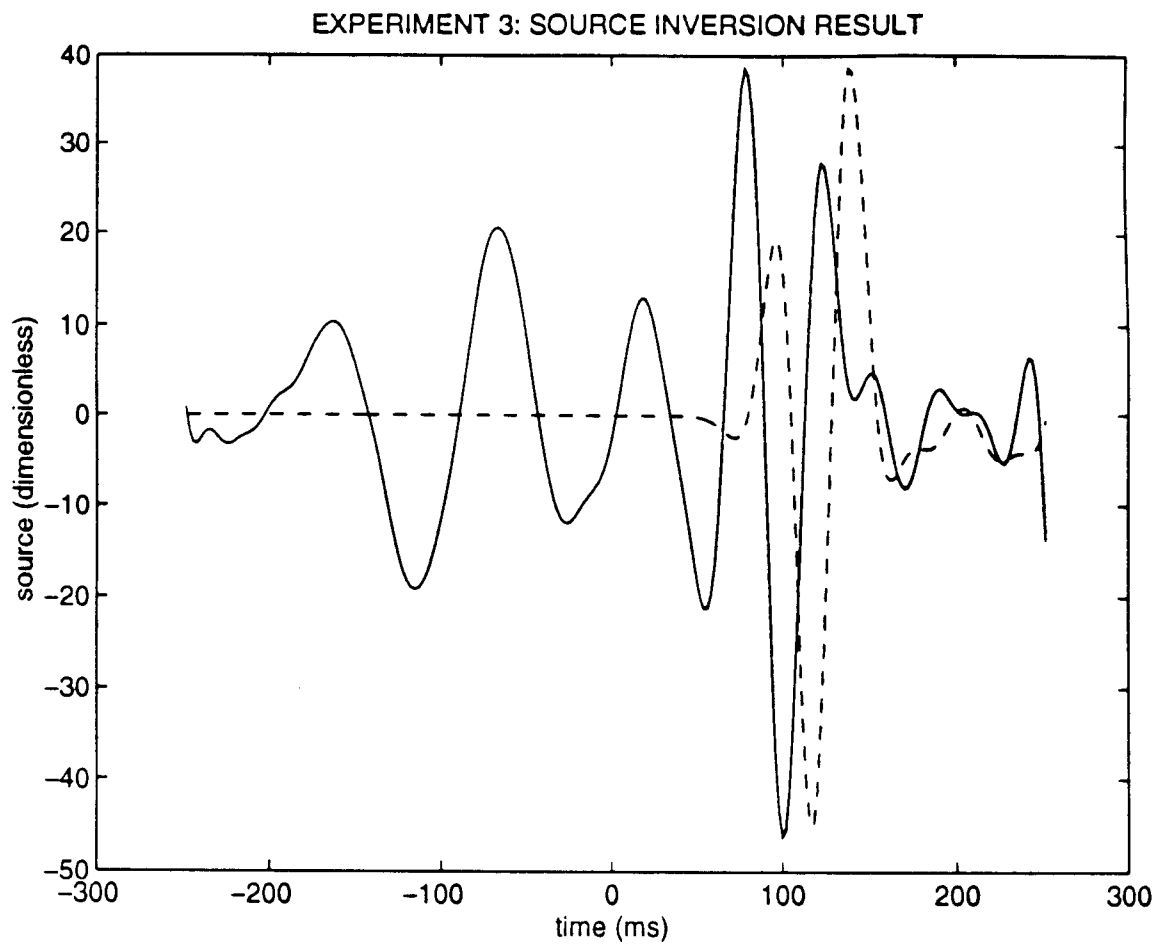
Figure 3.10 Source inversion result at alternation round 150 for the data shown in Figure 3.9. The rms error/data norm = 7%. Solid line: estimated source (scaled). Dashed line: target source.
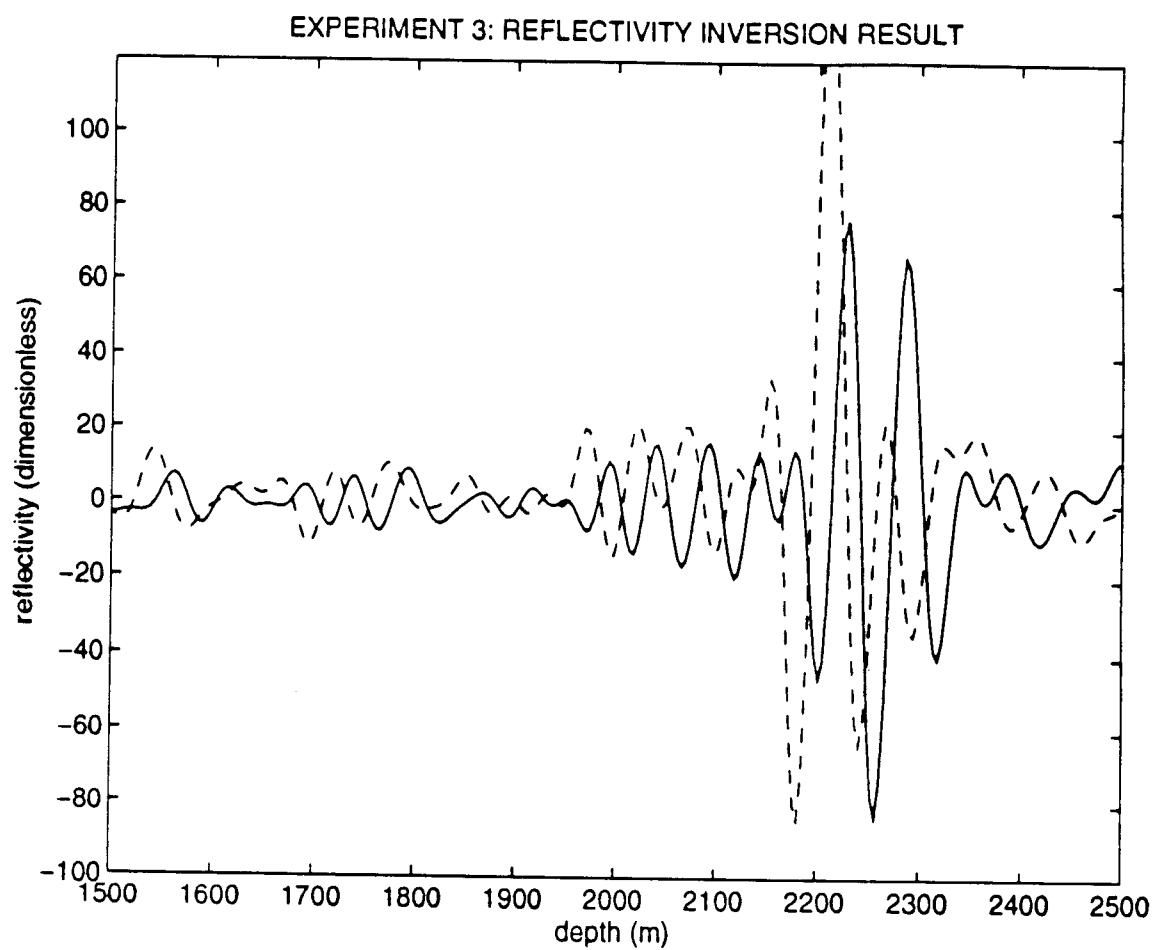
Figure 3.11 Reflectivity inversion result at alternation round 150 for the data shown in Figure 3.9. The rms error/data norm = 7%. Solid line: estimated reflectivity (scaled). Dashed line: target reflectivity.
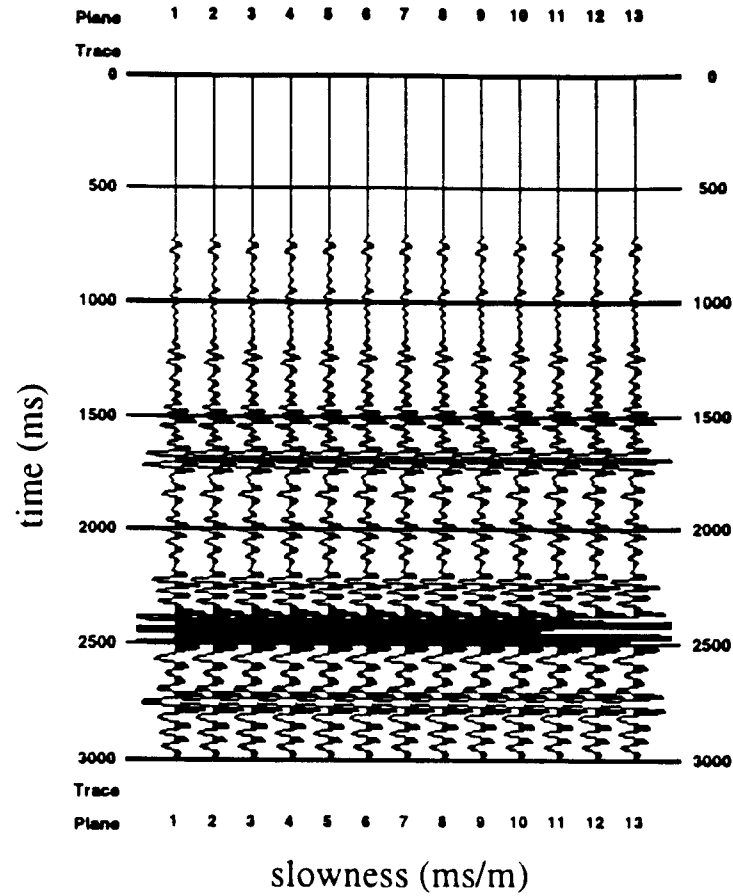
Figure 3.12 Synthetic $\tau - p$ transformed seismic data generated from the variable background velocity shown in Figure 3.1 and the isotropic source in Figure 3.13 (dashed line). The slowness aperture for the data is $p_{min} = p_{max} = .1158$ ms/m (i.e., all thirteen records are the same).
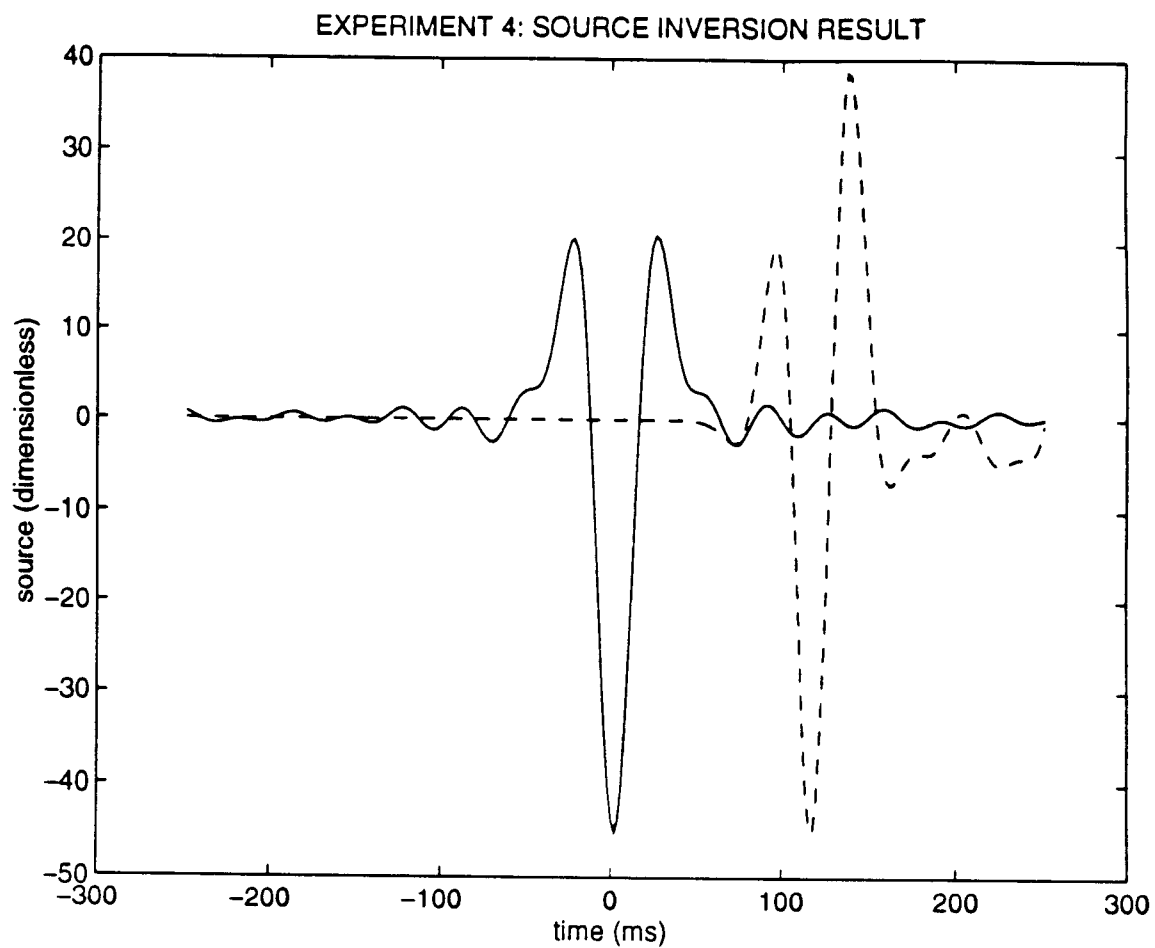
Figure 3.13 Source inversion result at alternation round 2 for the data shown in Figure 3.12. The rms error/data norm = 9%. Solid line: estimated source (scaled). Dashed line: target source.
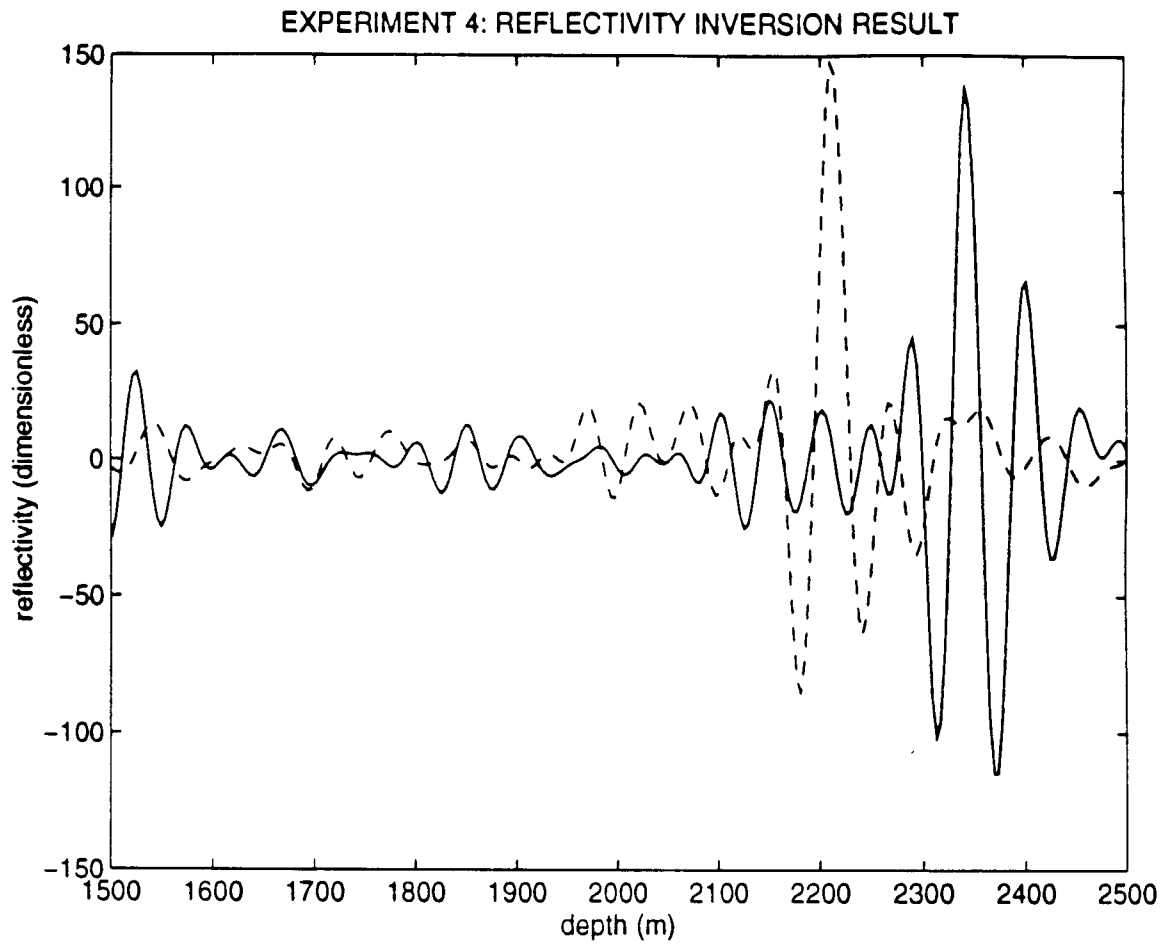
Figure 3.14 Reflectivity inversion result at alternation round 2 for the data shown in Figure 3.12. The rms error/data norm = 5%. Solid line: estimated reflectivity (scaled). Dashed line: target reflectivity.

Chapter 4

Full Waveform Inversion of Marine Reflection Data in the Plane-Wave Domain

4.1 Introduction

This chapter reports the full-waveform inversion of a small marine data set. The data used in this work are derived from a marine survey over a bright reflecting horizon embedded in essentially flat-lying layers. Little evidence of either multiply reflected energy or mode conversion is observable in the data. Accordingly, the inversion technique used a plane-wave viscoelastic model for P-wave propagation and primary reflections in a layered earth. A multistage iterative algorithm adjusted model components to minimize the mean-squared misfit between predicted plane-wave data gathers and those extracted by $p - \tau$ transform from field midpoint gathers. Inversion was used to estimate the compressional (or P-wave) background velocity, the anisotropic seismic source, and three elastic reflectivities (or short-scale relative fluctuations in combinations of the P-wave velocity, S-wave velocity, and density). This choice of inversion targets includes the background velocity which has a nonlinear effect on the data and thus can have a very large influence on the accuracy of the final answer. It also includes parameters which have a linear influence on the data (the elastic reflectivities and seismic source).

In order to gauge the accuracy of our results we used several tests. First, well logs provided an independent estimate of two of the three elastic parameters. The background velocity's ability to place the significant events in the reflectivities at the same depth independent of plane-wave slowness (p) provided us with a second, internal measure of consistency. This test is very similar to the use of "coherency" or "common image" panels in migration velocity analysis. Lastly, we used the relative misfit of the observed to predicted data to compare the inversion results. An initial estimate of the P-wave background velocity obtained by layer stripping in the $p - \tau$ domain proved inferior to the background velocity obtained from waveform inversion. Similarly, an initial estimate of the seismic source obtained from modeling the data

collection geometry and air gun apparatus was less successful in explaining the data than was the inversion result. Finally, we determined that a reasonably accurate quality factor estimate (Q_P and Q_S) is necessary to achieve an acceptable data fit. In particular, elastic inversion does not explain the dominant amplitude-versus-angle (AVA) trend observed in the data.

To rank earth models according to their ability to fit data (as we do here) is a radical proposition, not altogether conforming to contemporary geophysical practice, and so requiring justification. We compare estimates of small-scale features in the elastic parameters based on inversion for velocity model and source to estimates obtained from other velocity analysis and source parametrization techniques. We find that the inversion-based estimates conform more closely both to independent measurements (well logs) and to commonly-accepted lithological relationships.

Evidently, this justification of inversion is preliminary: it rests on analysis of a single-small data set. In fact most of our analysis is performed on a single midpoint gather. This small data set is extracted from a high-quality survey conducted in a region which conforms reasonably well to the modeling assumptions underlying the inversion. Multiparameter waveform inversion is a noise-sensitive process, and the validity of conclusions based on it could only become clear from analysis of considerably more evidence than is presented here. The present work shows that such analysis is possible, and that inversion can produce very reasonable results.

The prior work perhaps closest in spirit to our approach is that of Martinez and McMechan who use a viscoelastic simulator to forward model seismic data as the basis for an inversion algorithm [34] and [35]. Part 1 of their paper shows three synthetic examples which contrast elastic and viscoelastic modeling. The modeling technique is layer-based and produces $p - \tau$ (plane-wave) seismograms, as does the modeling technique used in this chapter.

In part 2 of their paper, Martinez and McMechan apply linearized inversion to estimate the quality factors, shear and compressional wave velocities, and density. They test the inversion on both synthetic and field data. Furthermore, eigenvalue analysis is performed on a simple single-layer model problem indicating the best and least well-determined of the estimated parameters. In contrast to the work reported here, however, Martinez and McMechan do not determine the energy source as part of their inversion.

The results presented in [25] are also very closely related to those described in this chapter. In fact, the experiments discussed in this chapter used data from the

same Gulf of Mexico survey. Igel employed a two-dimensional elastic finite difference simulator to invert for the P-wave impedance and Poisson's ratio. He did not invert for either the source or background compressional-wave velocity, relying on other techniques to provide these parameters. By inverting near offset'traces for the P-impedance and then using this result to invert all the traces for Poisson's ratio, he was able to obtain a final misfit error of 44.7%. An excellent reference list for viscoelastic modeling can be found in Igel's thesis.

Other recent references for elastic full waveform inversion include [63] and [38]. Wood estimates the shear velocity, compressional velocity, and density via inversion. Pan et al. include the layer thickness as the fourth parameter sought in the inversion. In all of the above inversion references, however, the source wavelet is either assumed known exactly ([63]) or is estimated via statistical techniques and initially extracted from the data ([34], [35], and [38]). Martinez and McMechan make mention of the fact that inversion could be used to estimate the energy source as well, but no examples are given.

Along with well-known source signature deconvolution techniques such as predictive deconvolution [43] and homomorphic deconvolution (see [61], [54], [8]), several authors have discussed estimating the energy source and other parameters via inversion. The references include a talk given by Canadas and Kolb [9], and mathematical justification for the idea of using inversion to estimate the energy source is given in [7], [46], and [33]. Lewis showed that band-limited sources and reflectivities could be estimated by simultaneous inversion. In Chapter 3 we extended Lewis' results to nonconstant background velocities and introduced the numerical technique used in this chapter. None of these references, however, combine viscoelastic modeling with energy source estimation.

In short, the work reported here differs from all prior work discussed in the references above on at least one of the following four counts:

- *use of waveform inversion to estimate velocities*

- *modeling of attenuation (viscoelastic wave propagation)*

- *estimation of energy source parameters via inversion*

- *correct treatment of three-dimensional amplitudes.*

Section 4.2 of this chapter explains the viscoelastic, plane-wave, primary-reflection model and the inversion methods based on it used in our work. Section 4.3 provides

details about the marine data set used in all the experiments shown in this chapter. The remainder of the chapter details the experiments we performed starting with the background or long-wavelength velocity inversions and continuing with the source and elastic parameter inversions enumerated in Table 4.1 and discussed in Section 4.4. Finally, some conclusions are drawn in Section 4.5.

4.2 The Method

4.2.1 Modeling

We model the earth as a viscoelastic medium. We do not give the full derivation of the convolutional solution approximation to the viscoelastic model here. The viscoelastic simulator we used was built on the solution approximation given in [1] pp.153–155. One of the few references which gives a detailed derivation of the viscoelastic equations from physical laws is [41]. See also references cited therein.

The mechanical parameters in the model include the density, ρ, the shear and compressional wave velocities, v_S and v_P, and the shear and compressional quality factors, q_S and q_P. The time-dependent, anisotropic source is assumed to have, approximately, point support.

Assume that the earth is a layered medium. Thus the parameters in our model vary only with depth, $z \equiv x_3$. By applying the Radon integral transform (or plane-wave decomposition) to the solution of the viscoelastic wave equation (and to common midpoint gathers of the data) we reduce the three-dimensional model to a family of one-dimensional models (see [60]). In effect, we have synthesized incident plane wave "shot" records parametrized by slowness, p, and by midpoint. Use of the plane-wave approach is justified by the fact that the reflection angles of interest for inversion were well sampled by the recording arrangement for the frequencies produced by the source.

Neither mode conversion nor multiply reflected energy appear to be important in the data set used for the experiments in this chapter (most likely because of the soft water-bottom materials in this part of the subsurface). Thus we were able to assume a primaries only, or single-scattering, approximation. The mechanical parameters are separated into the long-wavelength (smooth) background velocities and density v_P, v_S, and ρ, and short-wavelength relative perturbations of these parameters (reflectivities) $r_P = \delta v_P / v_P$, $r_S = \delta v_S / v_S$, and $r_d = \delta \rho / \rho$. High-frequency asymptotics, leads to the convolutional model prediction of the P-wave seismogram (see [5]),

$$S^{pred}(t,p) = f(t,p) * \tilde{r}(t,p).$$

In the above expression, S is the seismogram, f the source wavelet; p denotes slowness, and t time. The "$*$" symbol is convolution in time. The reflectivity $\tilde{r}$ may be written

$$\tilde{r}(t,p) \approx \int dz [A_P(z,t,p)r_P(z) + A_S(z,t,p)r_S(z) + A_d(z,t,p)r_d(z)].$$

The geometric optics reflectivity amplitude A_P is

$$A_P(z,t,p) = \int_{-\infty}^{\infty} d\omega F_P(z,p) \exp\left[i\omega(t - 2\tau(z,p)) - |\omega|\left(1 + \frac{2i}{\pi}\ln\left|\frac{\omega}{\omega_0}\right|\right)\alpha(z,p)\right].$$

The vertical travel time of plane waves with slowness, p, is

$$\tau(z,p) = \int_{z_0}^{z} dy \frac{\sqrt{1 - (pv_P(y))^2}}{v_P(y)},$$

and

$$\alpha(z,p) = \int_{z_0}^{z} dy \frac{1}{q_P(y)v_P(y)(1 - (pv_P(y))^2)^{1/2}}$$

is the corresponding attenuation factor. An algebraic combination of background parameters $v_P(z)$, $v_S(z)$, and $\rho(z)$, and slowness p gives $F_P(z,p)$. The temporal frequency variable is denoted ω; z_0 is the source depth, and q_P is the P-wave attenuation factor. Similar expressions define the geometric optics amplitudes A_S and A_d. The reference frequency ω_0 calibrates the (frequency-dependent) velocity. Waves at frequency ω_0 move with the P-wave velocity v_P.

The integrals described above are approximated by the trapezoidal rule in the modeling code. For use in the optimization, both the linearizations of the above expressions and their adjoints are required. These are computed by applying first order perturbation theory to the discretized integral transforms.

4.2.2 Inversion

The predicted seismogram S^{pred} is linear in each of the parameters f (the seismic source) and r_P, r_S, r_d (the elastic reflectivities). It is very nonlinear in the P-wave velocity v_p. This chapter discusses inversion for these five parameters. The predicted seismogram also depends on the (background) S-wave velocity v_S and density ρ, and on the quality factors Q_P and Q_S. We have assumed here that v_S and ρ are known with sufficient accuracy from logs and regional relationships which hold on the average over long scales. The quality factors were estimated by roughly matching the rate of energy decay in the data with predictions from log-derived synthetics.

The basic inversion principle embodied in our algorithms is *Output Least Squares* (or OLS). This method requires that we adjust the inversion parameters f, r_P, r_S, r_d and v_P to minimize the mean-squared error

$$J_{\text{OLS}} = \int_{t_{\min}}^{t_{\max}} dt \int_{p_{\min}}^{p_{\max}} dp\, \gamma(t,p) \left| S^{\text{pred}}(t,p) - S^{\text{obs}}(t,p) \right|^2$$

where $S^{\text{obs}}(t,p)$ is the "observed" $p - \tau$ data, and $\gamma(t,p)$ is a *conditioning weight* factor.

The production of the $p - \tau$ data set S^{obs} will be discussed in Section 4.3 (data characteristics and preparation). The conditioning factor $\gamma(t,p)$ enhances the resolution of deeper events. It is a model-based gain. The desired outcome of conditioning permits considerable freedom in the design of the weight. We have used

$$\gamma(t,p) = \exp(|\omega_0|\alpha(z,p))$$

where $t = \tau(z,p)$ in this formula. In the elastic limit ($Q_p \longrightarrow \infty$) no conditioning is necessary.

We have used several algorithms to invert for various combinations of parameters. These choices are outlined in the following paragraphs.

Elastic Reflectivities

The influence of the elastic reflectivities (r_P, r_S, and r_d) on S^{pred} is linear, so the mean-squared error is quadratic in these parameters. As the number of samples is moderately large (approximately 2000 total for the experiments described here), it is natural to use an iterative minimization method. We used the conjugate residual method which monotonically reduces the normal residual (see [15]). While the choice

of source wavelet influenced the rate of convergence, generally 20–50 iterations of the conjugate residual algorithm were required to reduce the normal residual to 1% of its starting value. In all cases, the initial estimates of the reflectivities were zero.

Elastic Reflectivities and Source Parameters

The source parameters alone have a linear influence on the data, so the same technique was applied to invert for them, namely, conjugate residual iteration. Typically more iterations are required to satisfy the convergence criterion (reduction of the normal residual) than is the case with the elastic reflectivities. This slower convergence appears to reflect the poorer numerical condition of the source inversion problem.

The elastic reflectivities and source parameters together have a nonlinear (bilinear) influence, so quadratic minimization algorithms cannot be used directly. Instead, we used a method known as coordinate search or alternation. This method of simultaneous inversion for source and reflectivities is introduced in Chapter 3. We will hereafter refer to a "round of alternation" to be one pass through the four steps of the following algorithm.

Repeat until convergence:

1. *Given the current source, f_c, and current reflectivity, r_c, invert for a new estimate of the reflectivity r_+.*

2. *Replace r_c by r_+.*

3. *Given the current source and reflectivity guesses, f_c, r_+, invert for a new estimate of the source f_+.*

4. *Replace f_c by f_+.*

Alternation, although notoriously inefficient, is attractive for initial experiments because it requires only successive solution of simple linear least-squares problems. Obviously, quasi-Newton methods could be applied to the problem and would likely reduce the number of iterations dramatically. While further algorithmic development for source-reflectivity inversion is definitely needed, we were able to obtain usable results in many cases with 2–3 rounds of alternation.

Elastic Reflectivities and P-Wave Velocity

The P-wave velocity has a strongly nonlinear influence on the data. Therefore the output least squares objective function J_{OLS} tends to be very complex. Gradient-based optimization algorithms typically have difficulty converging to useful estimates of velocity and reflectivity ([18], [59]).

A modification of the OLS approach appears to make it much more amenable to gradient-based optimization. The essential changes are these:

1. *Split* the optimization into *Stages*: i.e., for each velocity estimate, invert first for the elastic reflectivities in an *inner optimization*. Then treat the remaining error (*reduced objective*) as a function of velocity, and optimize over velocity.

2. *Relax* the model by permitting the reflectivity components r_P, r_S and r_d to depend on the slowness ("shot" parameter) as well. This modified inversion will produce a reflectivity "volume" quite analogous to the image volume of prestack depth migration.

3. *Penalize* the variation of reflectivity components with slowness by adding a *ferential semblance* term to J_{OLS}:

$$J_{DSO} := J_{OLS} + \sigma^2 \int \int dp\, dt \left\{ \left| \frac{\partial r_P}{\partial p} \right|^2 + \left| \frac{\partial r_S}{\partial p} \right|^2 + \left| \frac{\partial r_d}{\partial p} \right|^2 \right\}$$

The *semblance weight* σ^2 regulates the emphasis on semblance. As $\sigma \to \infty$, the objective functions J_{DSO} and J_{OLS} become equivalent:

$$\min_{(r_P,r_S,r_d)} J_{OLS} = \lim_{\sigma \to \infty} \min_{(r_P,r_S,r_d)} J_{DSO}$$

On the other hand, the opposite limit $\sigma \to 0$ completely decouples the reflectivity inversions for various slownesses. In effect, minimization of J_{DSO} for $\sigma \sim 0$ ("low DSO weight" below) produces independent inversions for each p. This amplitude corrected image volume is a useful diagnostic tool for velocity quality control, like its analogue in conventional migration velocity analysis.

The choice of the semblance weight σ^2 regulates the shape of the objective function J_{DSO}. This weight is best viewed as a Lagrange multiplier for a certain constrained formulation of inversion velocity analysis. In the experiments reported here σ was set by trial and error.

The reduction of the OLS objective function (step (1) above) already appears to eliminate most of the irregular behavior reported in the literature on output least squares inversion, and suggests that the use of very costly stochastic optimization methods, as has been advocated by a number of authors (for example [48], [50], [49]) can safely be avoided. The relaxation and differential semblance steps (2) and (3) tend to enhance considerably the convexity of the objective. For discussion and examples, see [58], [53], and, [19], and references cited therein. The algorithm just described has been given the name *differential semblance optimization* or DSO. This algorithm consists of two stages: (1) an *inner optimization* for elastic reflectivities r_P, r_S, r_d, which for the examples in this chapter was performed via conjugate residual iteration (2) an *outer optimization* over P-wave velocity v_P of a reduced objective produced by the inner optimization. We have used a version of (nonlinear) conjugate gradient optimization described in [16] to perform the outer optimization. See [30] for a complete description of the algorithmic details of DSO.

4.3 Seismic Field Data — Geometry of the Geophysical Experiment and Preprocessing

The data used in this work was derived from a marine survey in the Gulf of Mexico. The survey line consisted of 511 shots recorded with 301 hydrophone groups. The group interval was 15m with a minimum source-receiver separation of 148m. The shot interval was 22.5m. Each group contained 17 equally spaced and equally weighted hydrophones. The data was recorded without a low-cut filter. A 110 Hz high-cut filter was applied. The sampling rate was 2ms, and the total record length was 5s.

This area of the Gulf contains a strong gas-sand-related direct hydrocarbon indicator readily visible in the stack at about 2.3s (see Figure 4.1). The stack shows this target horizon to be embedded in a sequence of nearly horizontal strata, beginning at roughly 1.5s. Therefore, layered medium modeling appeared to be a plausible tool for target-oriented inversion. Very little evidence of multiply-reflected energy appears above or near the target event, suggesting the viability of a primaries-only approach. Similarly, the apparent absence of mode converted events justifies restriction of the propagation model to P-waves.

The data was Radon transformed, respecting 3D cylindrical symmetry, to yield 48 plane-wave traces per midpoint gather. Slowness values range from $p_{min}=.1158$ms/m to $p_{max}=.36468$ms/m. To remove diffraction artifacts originating in the shallow sub-

surface, the plane-wave data were time migrated in the midpoint dip domain, then modeled to $p_{mid} = 0$. This last step collapses diffractions while still preserving layered reflection amplitudes. The stack shown in Figure 4.1 is constructed from this migrated and remodeled $p - \tau$ data. While the diffracting structures are still clearly visible in the upper 1.5s, the diffraction tails are largely removed and no longer interfere with the lower, layered structure in the stack. The same is largely true of the prestack $p - \tau$ data (see Figures 4.3 and 4.7).

Techniques other than inversion supplied alternate estimates of several parameters. A layer stripping method produced a piecewise constant velocity by flattening $p - \tau$ migrated events in selected windows. This approach to velocity analysis is the $p - \tau$ analogue of a common method for estimation of interval velocities using prestack migration. The result of this layer-stripping velocity analysis is shown as the dashed line in Figure 4.4.

Also, air gun modeling software gave an a priori estimate of the source signature and radiation pattern. Figure 4.9 shows this air gun model source over the range of slowness values used in the data with every fourth trace displayed for clarity.

In order to lighten the computational burden of inversion, we performed most of our calculations on a low-pass filtered version of the $p - \tau$ data which resulted from convolving all the traces with a 15 Hz zero-phase Ricker wavelet. A single midpoint gather of the filtered data is shown in Figure 4.3. This same unfiltered gather is shown in Figure 4.7.

All of the numerical experiments described in this chapter were performed on one or more of eleven selected midpoint gathers. The midpoints were spaced 37.5m apart, so the total distance from the first to the last midpoint in this subset was 375m. We label these midpoint gathers CMP1–CMP11. A logged well is near CMP6, and most of our results concern this single midpoint gather.

For the well near CMP6, we obtained block sonic and density (gamma ray) logs in the 1.4–2.6s (two-way time) interval. The logs are shown in Figure 4.2 plotted as a function of depth and not frequency filtered. We have used these logs in assessing the accuracy of the inversions (see Figures 4.10, 4.11, 4.15, 4.16, 4.19, 4.20).

4.4 Numerical Experiments — Viscoelastic Inversion in the Plane-Wave Domain

4.4.1 Table of Experimental Inputs

	Numerical Inversion Experiments						
Experiment type	DSO		OLS				
Experiment no.	1	2	3	4	5	6	7
Model	viscoelastic		viscoelastic				elastic
Data gather no.	6	6	6	6	6	1,5,7,11	6
Data frequency filtered?	Y	N	Y	Y	Y	Y	Y
Data conditioned?	Y	Y	Y	Y	Y	Y	N
Velocity inversion?	Y		N				
Yes, initial velocity	Fig. 4.4 dash/dot line						
No, fixed velocity			Fig. 4.4 dark solid line				
Reflectivity inversion?	Y		Y				
Initial reflectivities	zeros		zeros				
Source inversion?	N	N	N	Y	Y	Y	Y
Yes, initial source				isotropic 15 Hz Ricker			
Yes, final source				isotropic	anisotropic		
No, fixed source	air gun		air gun				
DSO weight	10^{-4}		∞				
Damping weight	10^{-8}		10^{-8}				
Relative residual tolerance	5%		1%				
Alternation rounds	N/A		N/A	2	2	2	2
Relative rms error			55%	53%	29 %	30%	33%
Cost function reduction	35%	67%					

Table 4.1   DSO and OLS plane-wave inversion experiments performed on Gulf of Mexico field data.

4.4.2 Discussion of Results

1. The inversion-estimated P-wave background velocity places the significant events in the inverted reflectivities at the same depth location across different slowness values.

In Experiment 1, DSO inversion was used to estimate the P-wave background velocity shown in Figure 4.4 (dark solid line) from the frequency filtered data shown in Figure 4.3. The reflectivities, also estimated in this experiment, are allowed to depend on slowness, p (as described in Subsection 4.2.2). One measure of accuracy for the background velocity is its ability to place the significant events in the inverted reflectivities at the same depth location across traces. In order to gauge the background velocity's effectiveness at flattening these events, we set the differential semblance weight very low (10^{-9}). The differential semblance term tends to force consistent reflectivities from record to record. The main feature in Figure 4.5 (the P-wave impedance panel) is located between 2100m and 2200m and appears quite flat. For comparison, the background velocity derived via layer stripping in the $p - \tau$ domain is shown in Figure 4.4 (dashed line). The associated P-wave impedance gotten from performing an inversion with this background velocity and a very low semblance weight is shown in Figure 4.6. The visual "flatness" appears to be equally as good with the inverted background velocity as with the background velocity gotten from layer stripping.

2. The depth of the most significant event in the inverted reflectivity is close to the depth of this event on the sonic well log.

The gas-sand feature is located at about 2100m on the logs (Figure 4.2). The inverted background velocity places this event closer to 2150m (Figure 4.5); whereas the background velocity generated by layer stripping moves it towards 2200m (Figure 4.6).

3. DSO inversion performed on full bandwidth data still produces a background velocity which flattens significant events across p-dependent traces.

Using unfiltered CMP6 data shown in Figure 4.7, we ran a DSO inversion test for the P-wave background velocity and reflectivities (Experiment 2). Although we were not completely satisfied that we had pushed this inversion to its stopping point, the resulting velocity estimate can be seen in the Figure 4.4 (solid line) and is very close to the velocity derived from inverting the filtered data (dark solid line). Moreover, the migrated section with this background velocity estimate indicates that the significant events have been rendered relatively flat across traces. (see Figure 4.8).

4. Although we initially assumed the air gun model source estimate (see Figure 4.9) to be a good guess for the wavelet, we found that we were able to greatly improve the fit to data by deriving a new source via inversion.

Holding the P-wave background velocity fixed, we performed three OLS inversions for the elastic reflectivities. In Experiment 3, the energy source was not re-estimated in the process of inverting the reflectivities. The normalized misfit between actual and predicted data is 55% and the data residual is shown in Figure 4.12. The misfit graphs were all plotted on the same scale as the actual data being fit. It is clear from the graph of the residual that the reflectivities estimated using the air gun model source do not accurately predict the target event for large slowness values.

Experiments 4 and 5 are identical to Experiment 3 except that we estimated both the three elastic parameter reflectivities *and the energy source* using inversion. The energy source in Experiment 4 was, however, constrained to be isotropic. The initial guess for the isotropic wavelet is shown in Figure 4.13, and the final source wavelet estimated in this experiment is shown in Figure 4.14. In two rounds of the alternation algorithm applied to CMP6, the normalized misfit error was reduced to 53%. The residual graph makes manifest the fact that an isotropic source (even one gotten using inversion) and the corresponding reflectivities do an inadequate job of matching the target of the real data for large slowness values (see Figure 4.17).

Finally, in Experiment 5 we allowed the reflectivities and an *anisotropic* energy source to be estimated via inversion. In two rounds of alternation, we reduced the normalized misfit error to 29%. (In fact, in the time interval from 2200–2400ms, the *target interval,* these inversion results predicted all but 11% of the actual data.) The source estimate is shown in Figure 4.18, the residual in Figure 4.21. The residual graph now looks like random noise with the main target in the actual data well explained by data generated from the inversion-estimated parameters.

Similarly, we inverted the individual data gathers CMP1, CMP5, CMP7, and CMP11 in four separate experiments and achieved data fit errors of at most 30% after only two rounds of alternation (see the experiments described in Table 4.1 under "Experiment 6"). Thus we see that we can explain 25% more of the seismogram by allowing the earth parameters *and the source* to be estimated from the data. The air gun modeling package source had been estimated independently of the seismic data.

5. Of the three sets of estimated reflectivities (Experiments 3–5), the reflectivities from Experiment 5 which were generated by the inversion-estimated anisotropic energy source most closely match independent measurements.

The linearized inversion method assumes that a separation of frequency content exists between the background (low-frequency) density and velocities and the (high-frequency) reflectivities. To gauge the accuracy of our reflectivity estimates, we therefore found it necessary to detrend the well-logs. The logs had been detrended by subtracting a smooth average of the log from the original log and turned into (dimensionless) reflectivities by dividing this difference by the smooth average. Figure 4.10 compares the detrended sonic velocity log (dashed line) to the short-scale relative fluctuation in the P-wave velocity calculated from inversion (solid line) for Experiment 3. Figure 4.11 shows a similar comparison for the P-wave impedance. These reflectivity-log comparisons are all plotted as a function of two-way time and filtered using a zero-phase 8–32 Hz tapered Ormsby frequency filter (since the convolutional model constrains these parameters to lie in the passband of the source). For this experiment a shift of 20ms to the left was applied to the inversion result graphs in order that the log and inversion result could be overlaid and their shapes compared. The inversion result graphs were also scaled so that their peaks would be comparable to the well log measurement graphs.

Figures 4.15 and 4.16 show the log comparison to the P-wave velocity and P-wave impedance for Experiment 4. The inversion results in these graphs have been scaled to match the peak amplitude of the sonic log and a shift of 67ms to the left has been applied. A slightly larger time shift was needed here most likely because our inversion algorithm (alternation) was not powerful enough to change the initial (slightly erroneous) time location of the source by this small amount.

Finally, the parameter estimates versus logs for Experiment 5 are shown in Figures 4.19 and 4.20. Again, these inversion results were scaled and shifted 67ms to the left in order to be compared to the log measurements. Clearly, the comparisons to the logs improve with the addition of an inversion-estimated source and especially when this source is allowed to be anisotropic.

There was, in fact, no shear log measurement performed. Although agreement of the inversion-estimated parameters with the *measured* logs is generally good, none of the experiments described do an adequate job of matching in detail the shear log calculated from Castagna's relation. (For a reference to Castagna's relation see [11].) Since the target is a gas sand, however, we should not see any change in the shear modulus in this target range of 2200–2300ms. Figure 4.23 shows that the inversion results for the shear modulus from Experiments 4 and 5 do a better job of conforming to this expectation than does the modulus resulting from Experiment 3.

6. Inversion results for the anisotropic inversion-estimated source and reflectivities show that the V_p/V_s reflectivity is relatively constant and near zero except at the gas sand where it is more negative than the V_p reflectivity indicator.

[12] postulate that this V_p/V_s indicator is more dependable thaﾑ traditional AVO gas sand indicators such as the P-wave velocity. Their hypothesis comes from examining velocity and density well log data from 25 different regions of the world. The measurements were of adjacent shales, brine sands and gas sands. Our inversion results agree with their hypothesis. See Figure 4.22 which compares six elastic reflectivities from Experiment 5.

7. The viscoelastic model allows a better fit to this seismic data set than does an elastic modeler.

To test the importance of modeling attenuation for this data set, we inverted for a fully anisotropic energy source and the same three reflectivities using ungained data and an elastic modeler. The background velocity was the velocity in Figure 4.4 (dark solid line). Inversion for source and elastic reflectivities using two rounds of alternation reduced the root mean squared error to 33% of the data norm. However, one notes from the residual (Figure 4.24) that the parameters estimated via inversion do not generate data which fits the actual data well. For large slowness values, the target event can still be seen clearly in the residual.

8. The inversion-estimated P-wave background velocity and inversion-estimated anisotropic seismic source place the significant events in the inverted reflectivities at the same depth location across different slowness values.

As one final check of the validity of the inversion source estimates described in the previous subsection, we repeated the low DSO weight inversion experiments described in item 1 above with the source shown in Figure 4.18 as input. The seismic data was filtered common midpoint gather 6 (shown in Figure 4.3). The differential semblance penalty parameter was set at 10^{-9}. Figure 4.25 shows the P-wave impedance resulting from this inversion where the background velocity used was the inversion result from Experiment 1 shown in Figure 4.4 (dark solid line). Figure 4.26 only differs in that the background velocity used for the inversion is the layer stripping estimate shown in Figure 4.4 (dashed line). In both cases, the main events are quite flat.

4.5 Conclusion

In this chapter we have shown that we were able to estimate quite accurately parameters that affect real marine $\tau - p$ transformed seismic data in both linear and nonlinear ways by a completely deterministic approach. In a sequence of steps we inverted for the P-wave background velocity, the seismic energy source, and three elastic parameter reflectivities. The viscoelastic modeler used assumed a layered medium and single scattering. The results were conclusively in favor of using the data directly (via inversion) to estimate as much as possible about the model. We were able to match the most significant event in the P-wave velocity well log accurately; the background velocity estimate placed the events in the reflectivity panels at the same depth location across traces. By inverting for the energy source as well we were able to fit 70% of the seismic data. In fact, in a small window around the target, we were able to fit 90% of the seismic data in a single CMP data gather. We obtained these results despite a modeler which did not completely describe the real data. Although the data was pre-processed, near surface heterogeneities remained which did not conform to the assumptions of a layered model.
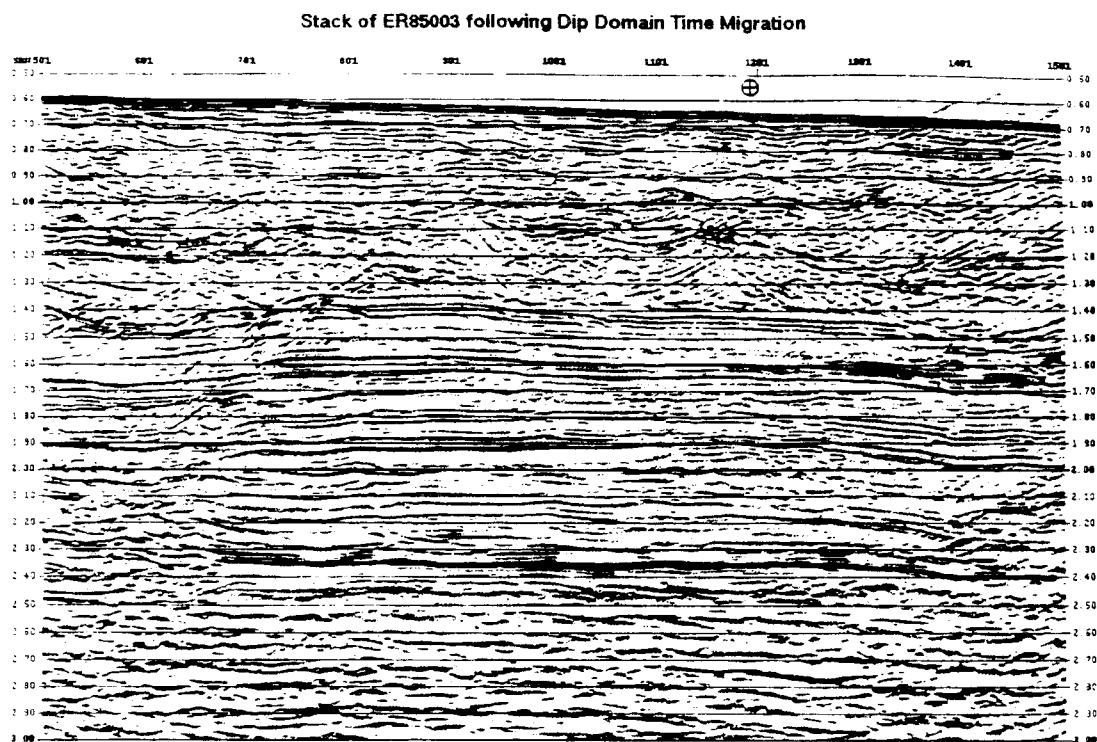
Figure 4.1 The stacked section of marine data. The diffracted energy originating in the shallow subsurface was suppressed by prestack time migration in the offset-midpoint slowness domain. Modeling recreated the data with zero midpoint slowness. The location of the logged well referenced in the text is marked. Note that the bright reflecting horizon, a gas sand at 2.3s, is embedded in a flat-lying sequence.
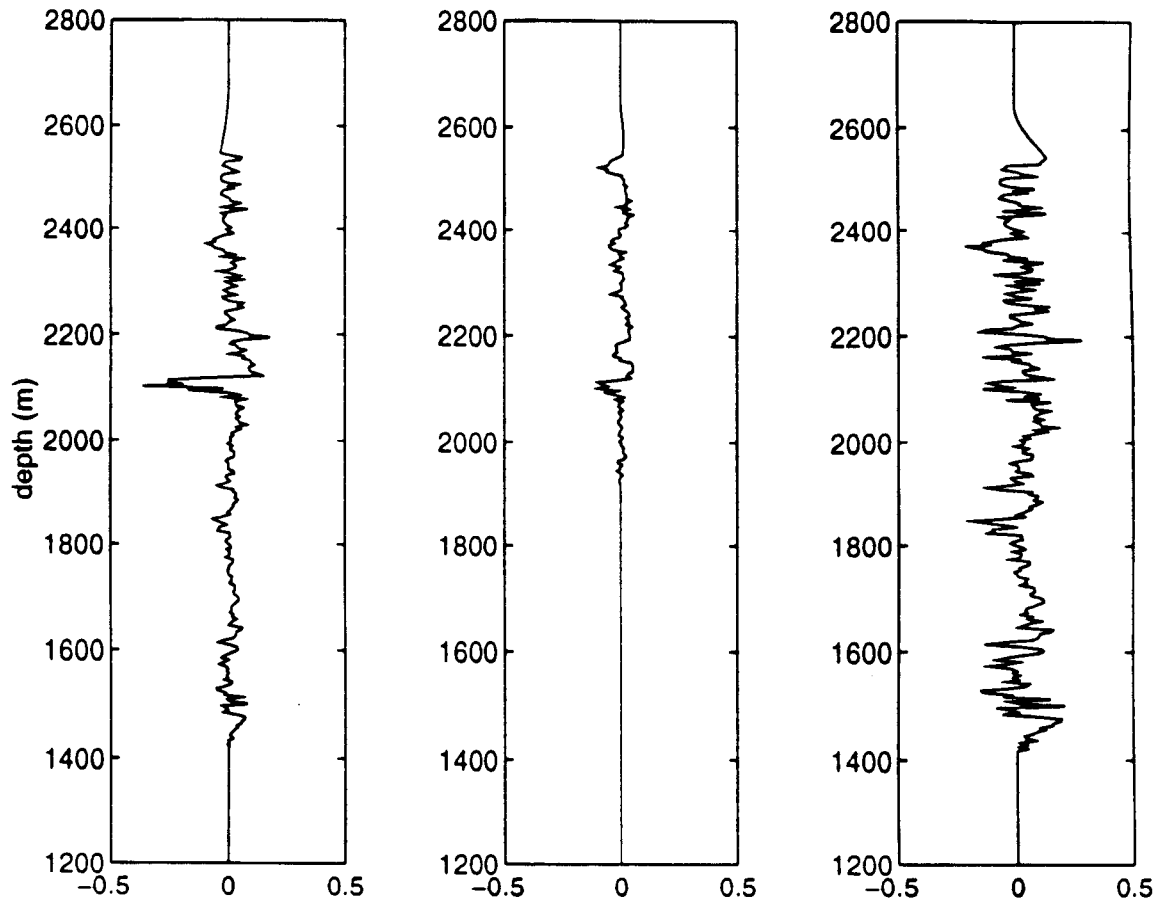
Figure 4.2 The measured and extracted logs from a well near CMP6. Note that unlike the comparisons to the inversion results shown in Figures 4.10, 4.11, 4.15, 4.16, 4.19, 4.20, these logs are plotted as a function of depth and not frequency filtered. The left panel shows the detrended measured P-wave velocity log. The middle panel shows the detrended measured density log. The right panel shows the shear wave velocity log estimated using Castagna's relation (see reference [11]) which did not prove to be reliable. There was no S-wave velocity measurement made of this area.
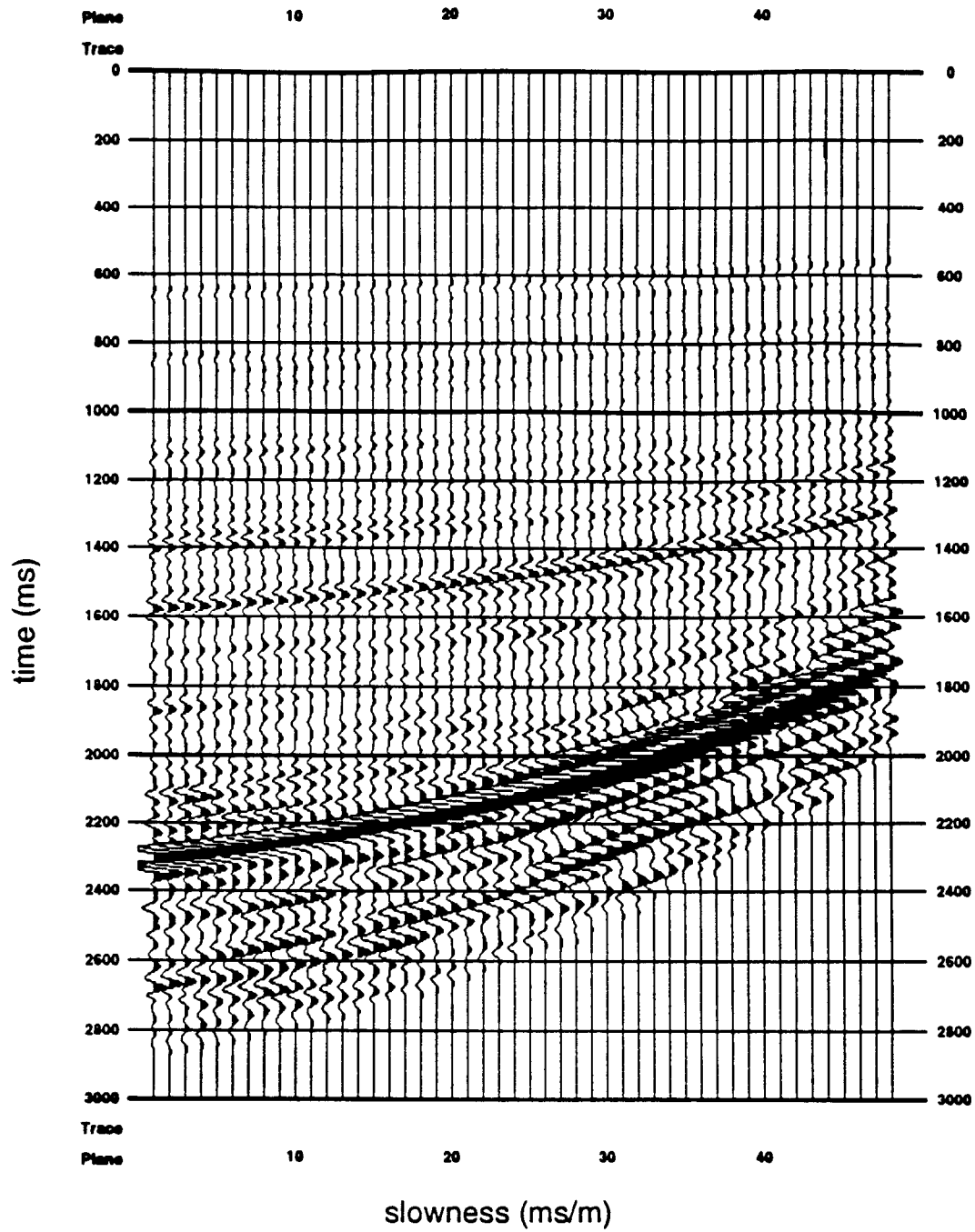
Figure 4.3 The $\tau - p$ transformed seismic data from common midpoint gather 6 which has had a 15 Hz Ricker filter applied.
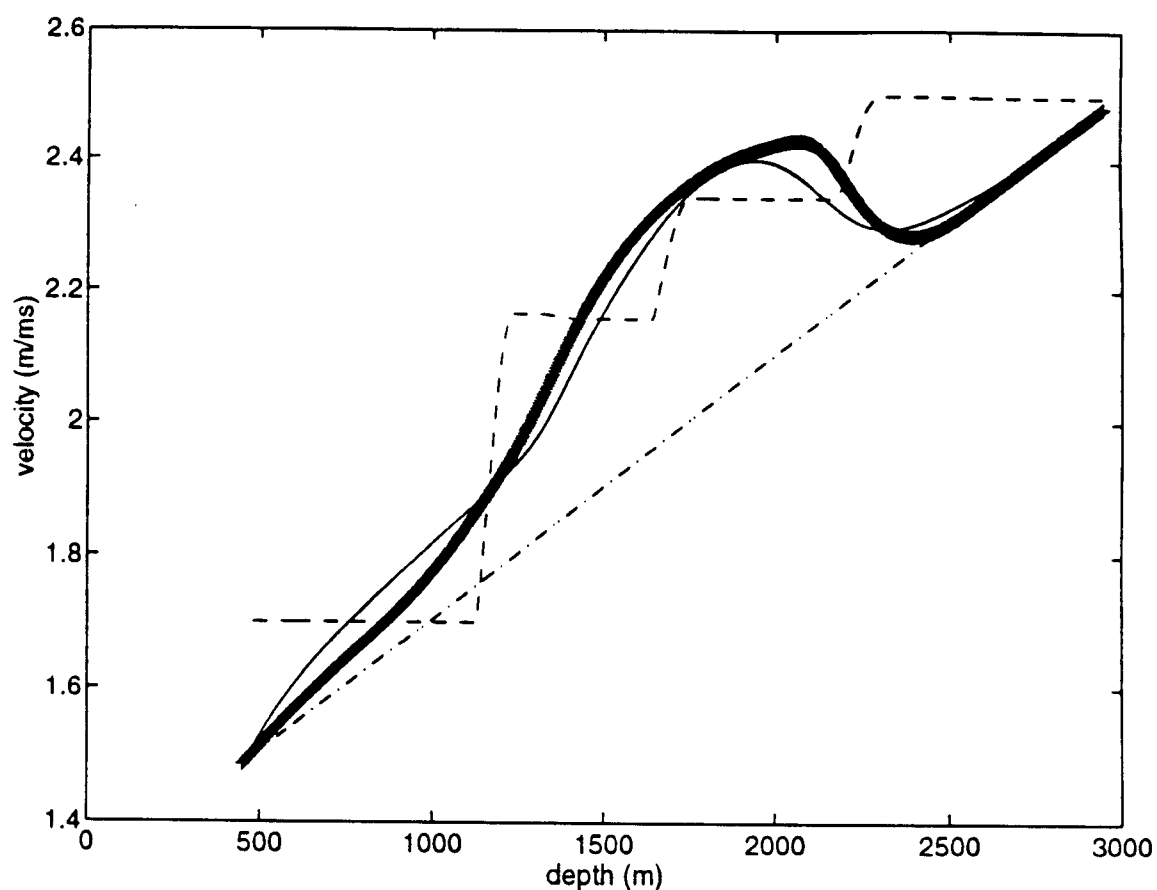
Figure 4.4 Dashed/dotted line: Initial estimate for P-wave background velocity used for inversion of common midpoint data gather 6 (Experiments 1 and 2). Dashed line: Background velocity predicted for the same data set using layer stripping in the $p - \tau$ domain. Solid line: P-wave background velocity determined via inversion for *unfiltered* CMP6 (Experiment 2). Dark solid line: Background velocity estimate gotten from performing the same inversion on *filtered* CMP6 data (Experiment 1).
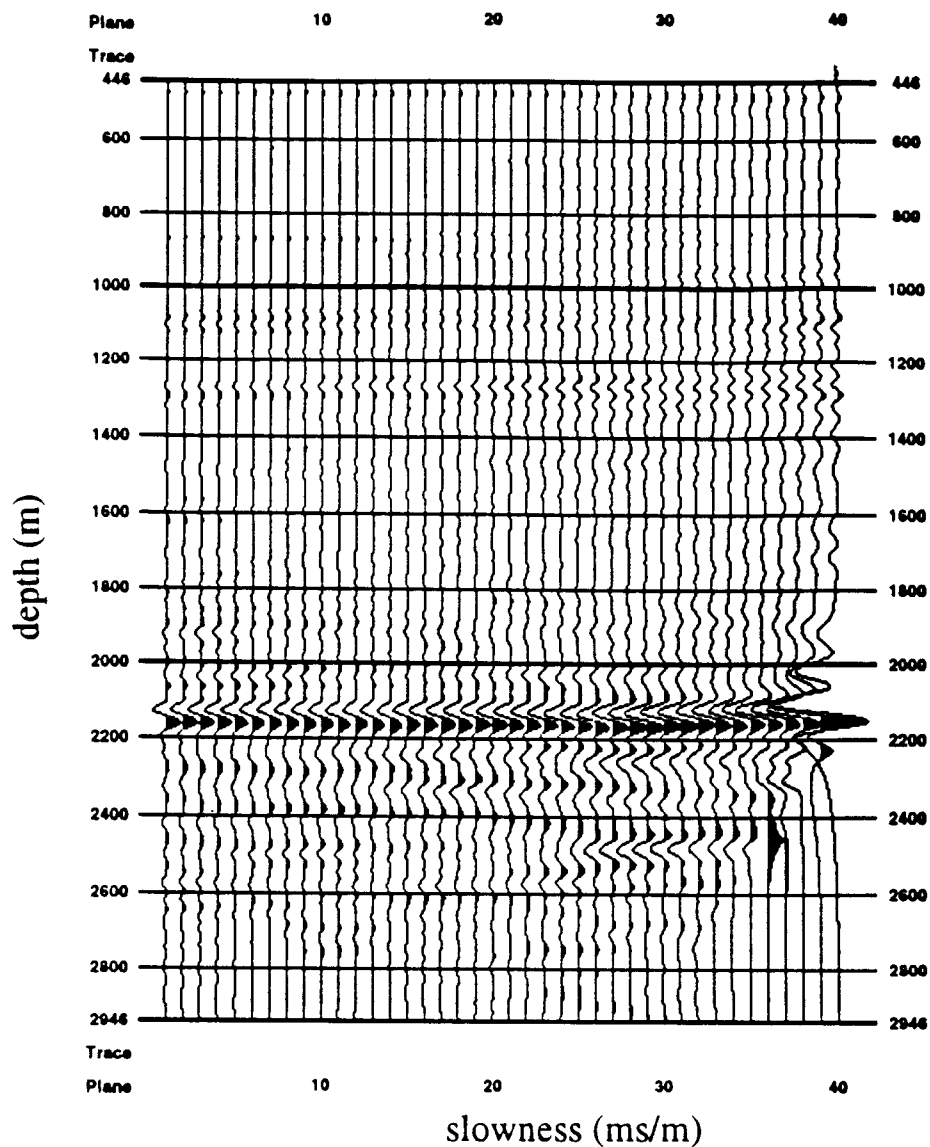
Figure 4.5 The P-wave impedance resulting from a linear inversion with the background velocity shown in Figure 4.4 (dark solid line), fixed (air gun model) source, and low DSO weight.
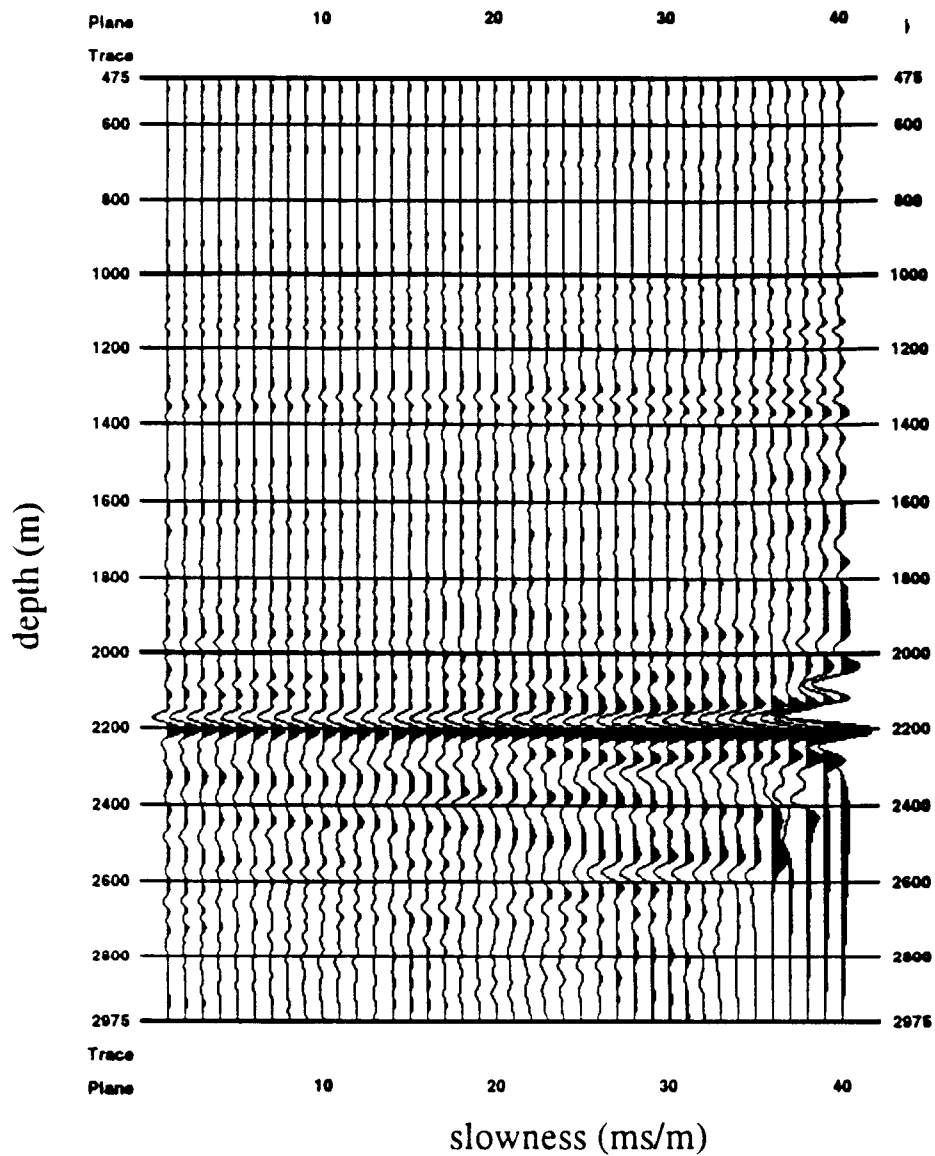
Figure 4.6   The P-wave impedance resulting from a DSO inversion with the layer-stripping background velocity shown in Figure 4.4 (dashed line).
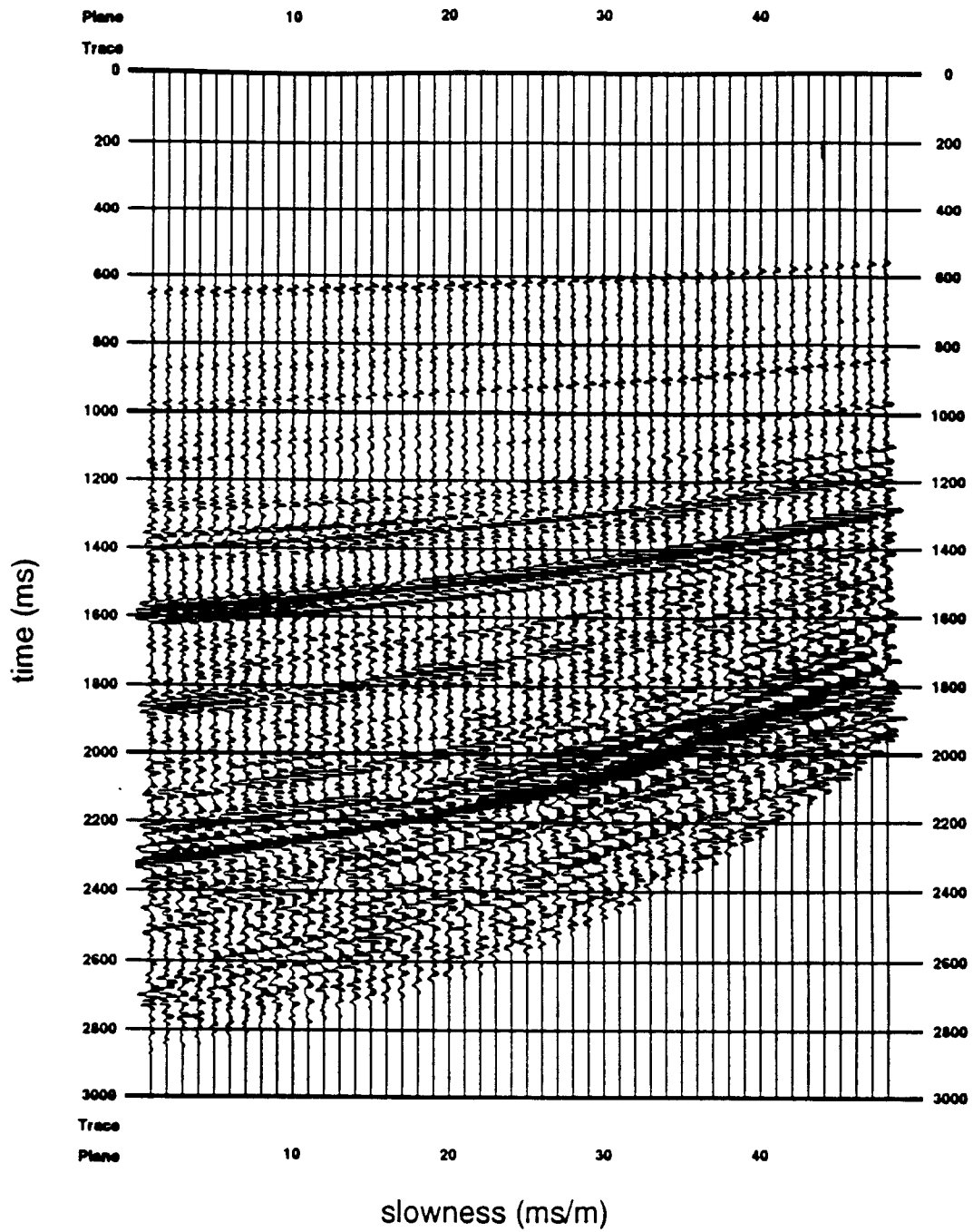
Figure 4.7 The unfiltered seismic data for common midpoint data gather 6 (used in Experiment 2).
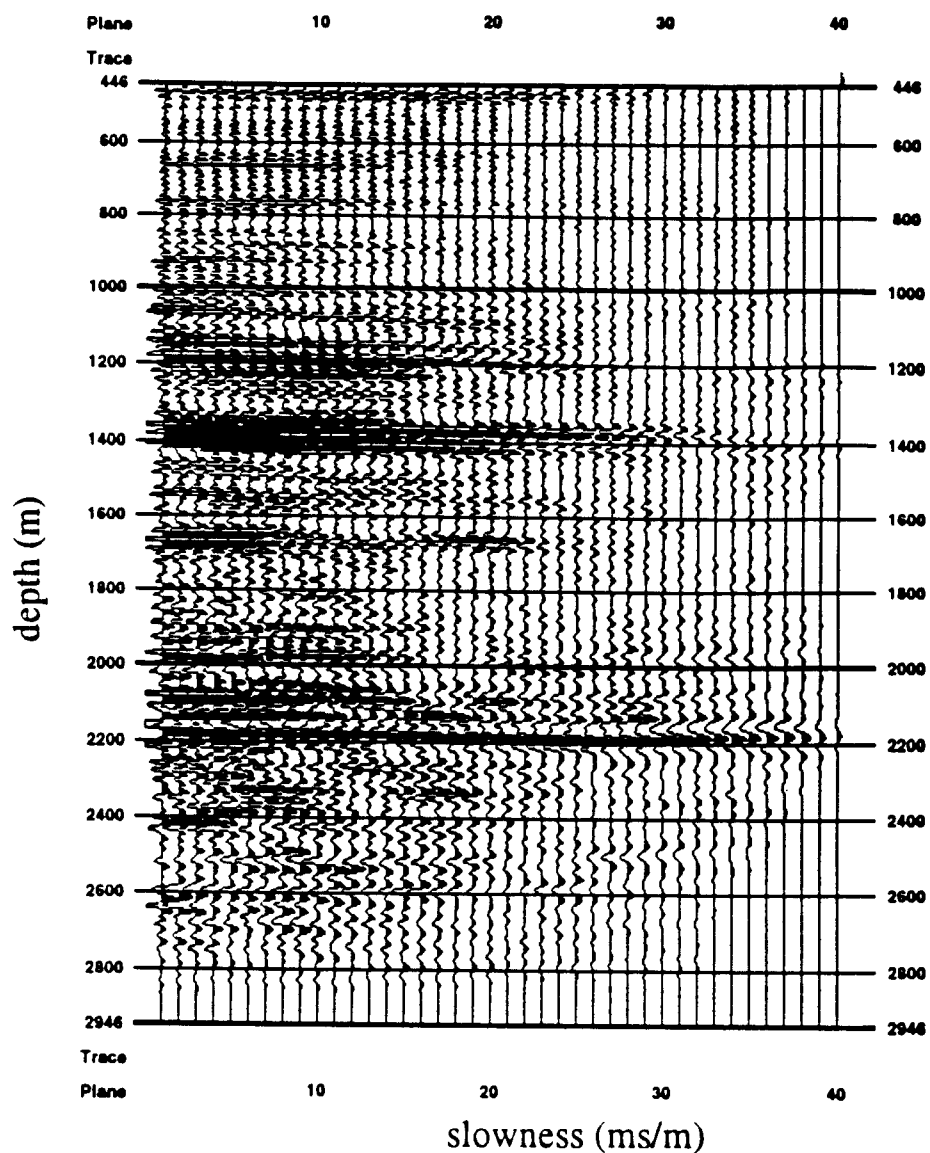
Figure 4.8 The P-wave impedance resulting from a migration performed on unfiltered CMP6 data with the background velocity shown in **Figure 4.4** (solid line) and an unfiltered air gun model source estimate.
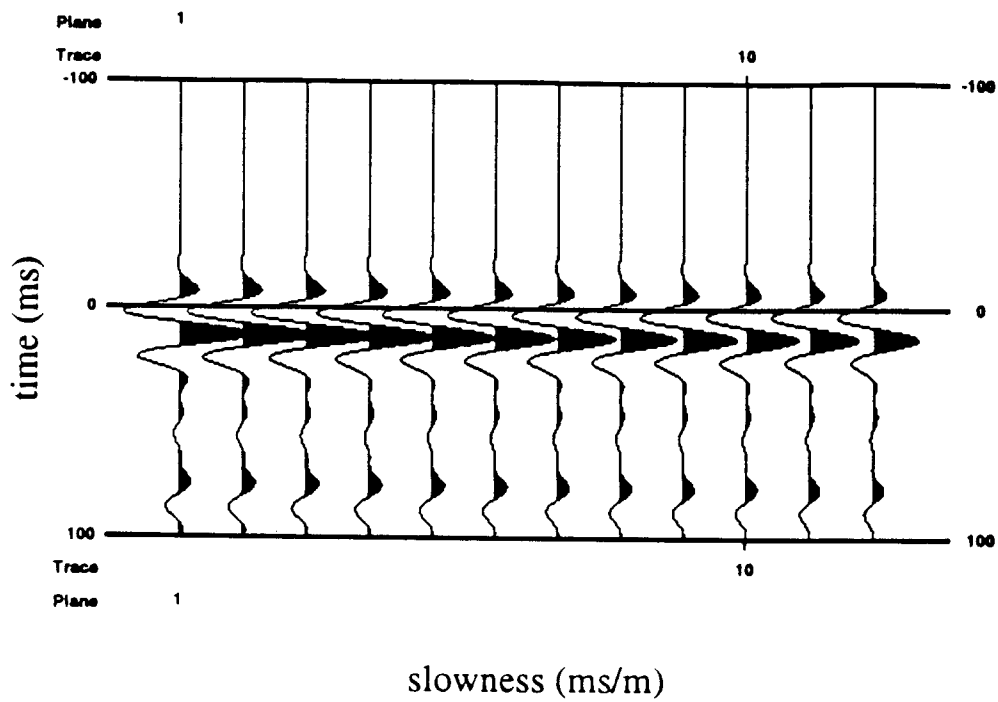
Figure 4.9 The air gun model source estimate with every fourth trace shown.
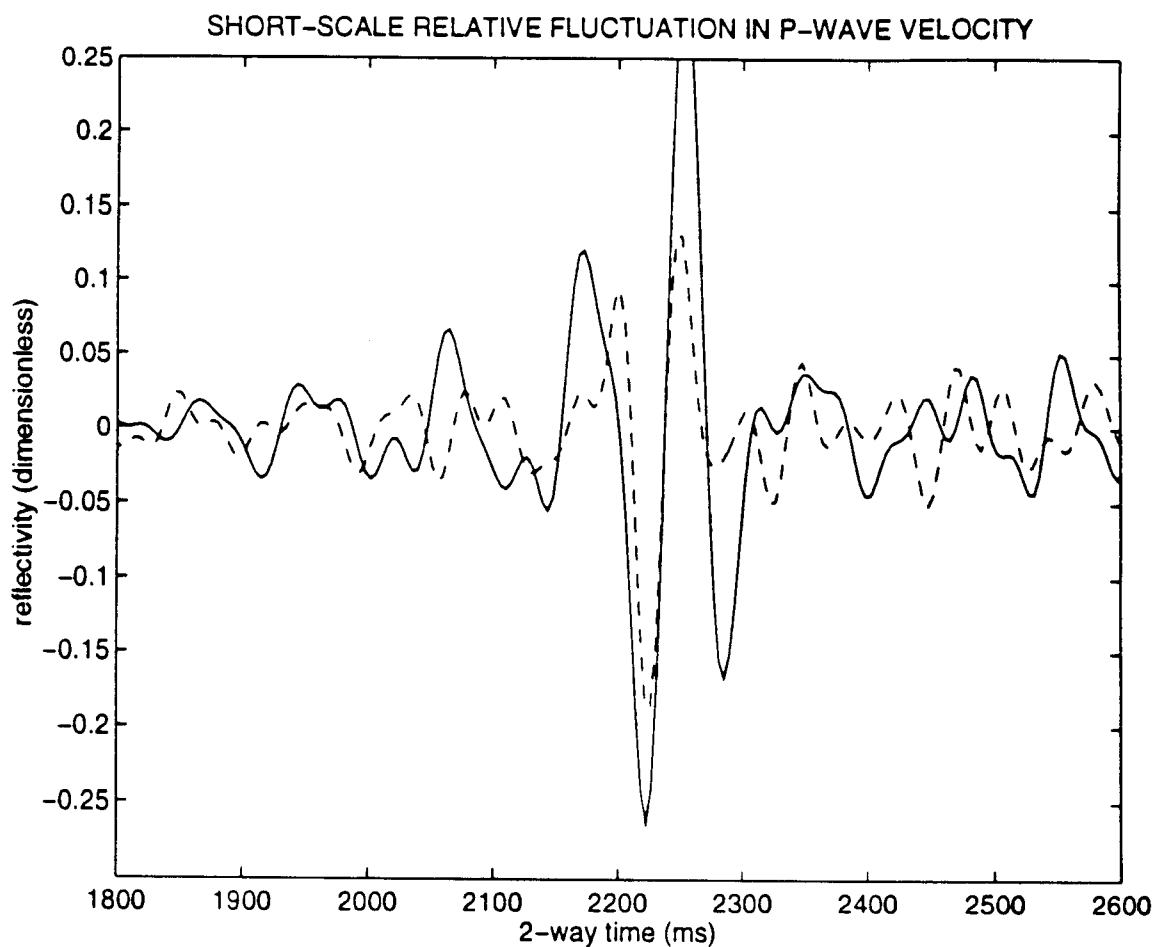
Figure 4.10 Comparison of the independent well-log measurement of the relative short-scale fluctuation in the P-wave velocity with the result of inversion done on CMP6 using the air gun model source estimate (Experiment 3). The solid line shows the inversion result (scaled and shifted left 20ms). The dashed line shows the detrended well log. Both graphs have been plotted as a function of two-way time and filtered to match the frequency content of the source.
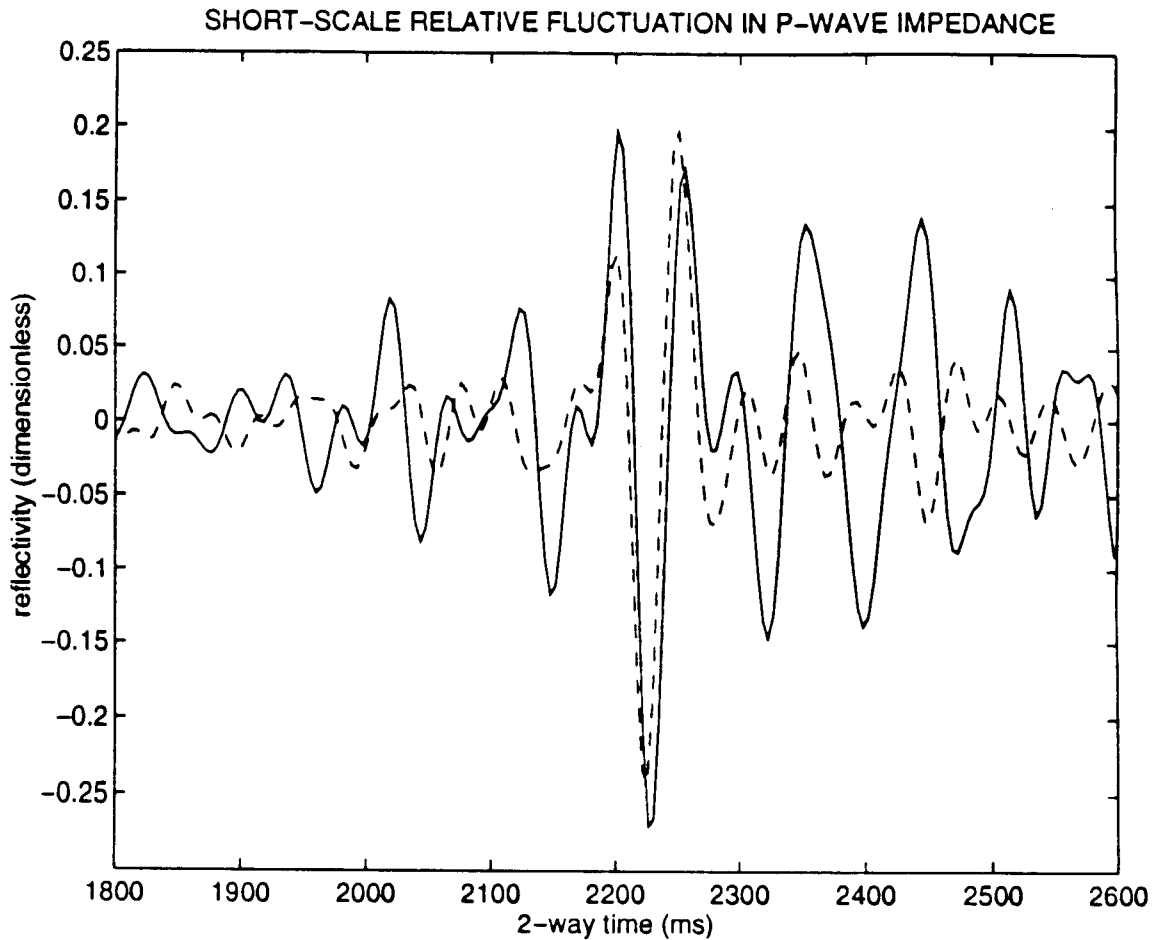
Figure 4.11 Comparison of the independent well-log measurement of the relative short-scale fluctuation in the P-wave impedance with the result of inversion done on CMP6 using the air gun model source estimate (Experiment 3). The solid line shows the inversion result (scaled and shifted left 20ms). The dashed line shows the detrended well log. Both graphs have been plotted as a function of two-way time and filtered to match the frequency content of the source.

Figure 4.12 Difference between the actual data (CMP6) and the data predicted by parameters gotten from inversion using the air gun source (fixed background velocity) (Experiment 3). The data misfit is plotted on the same scale as the CMP6 data.
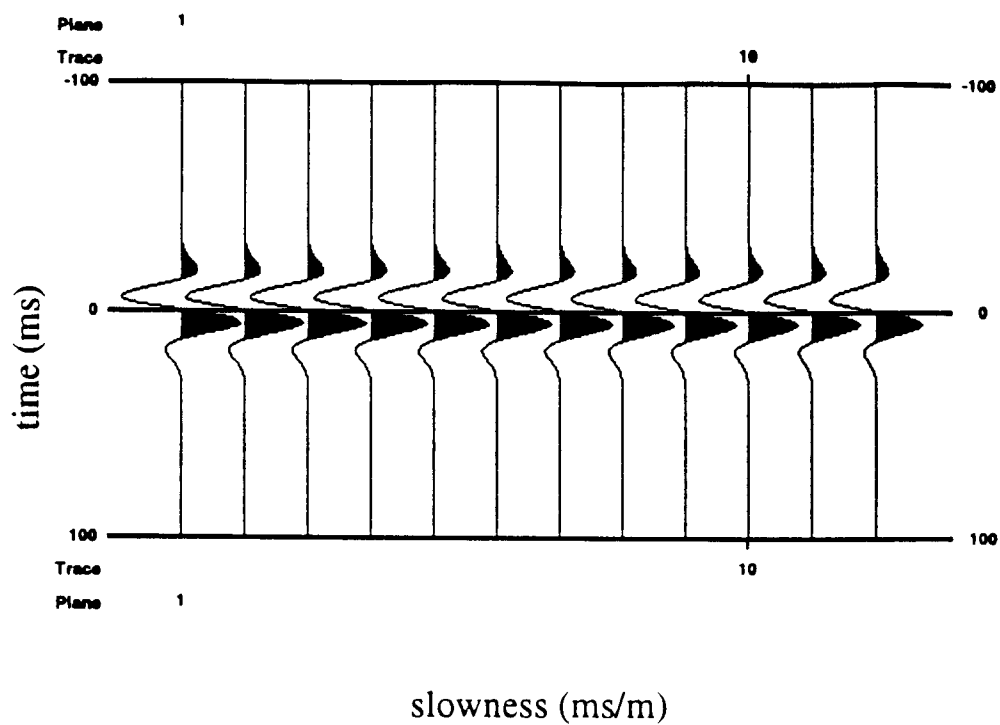
Figure 4.13 The initial source guess for the inversions described in Experiments 4 and 5 (derivative of an isotropic Ricker wavelet) where only every fourth trace is shown for clarity.
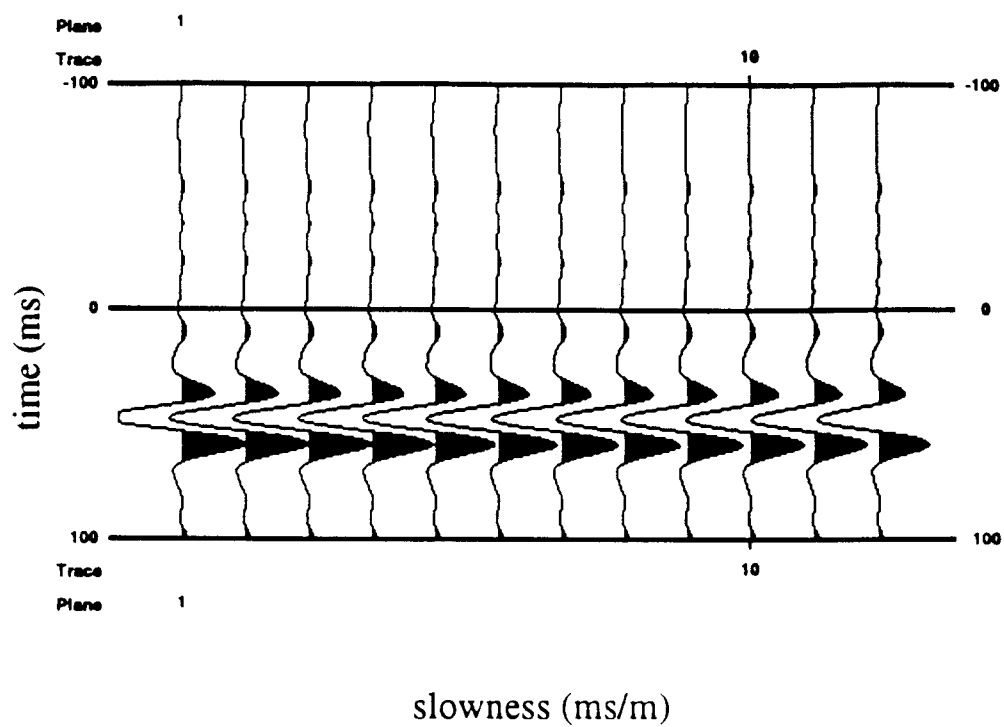
Figure 4.14 The isotropic inversion-estimated source from a linear source-reflectivity inversion job with every fourth trace shown (Experiment 4).
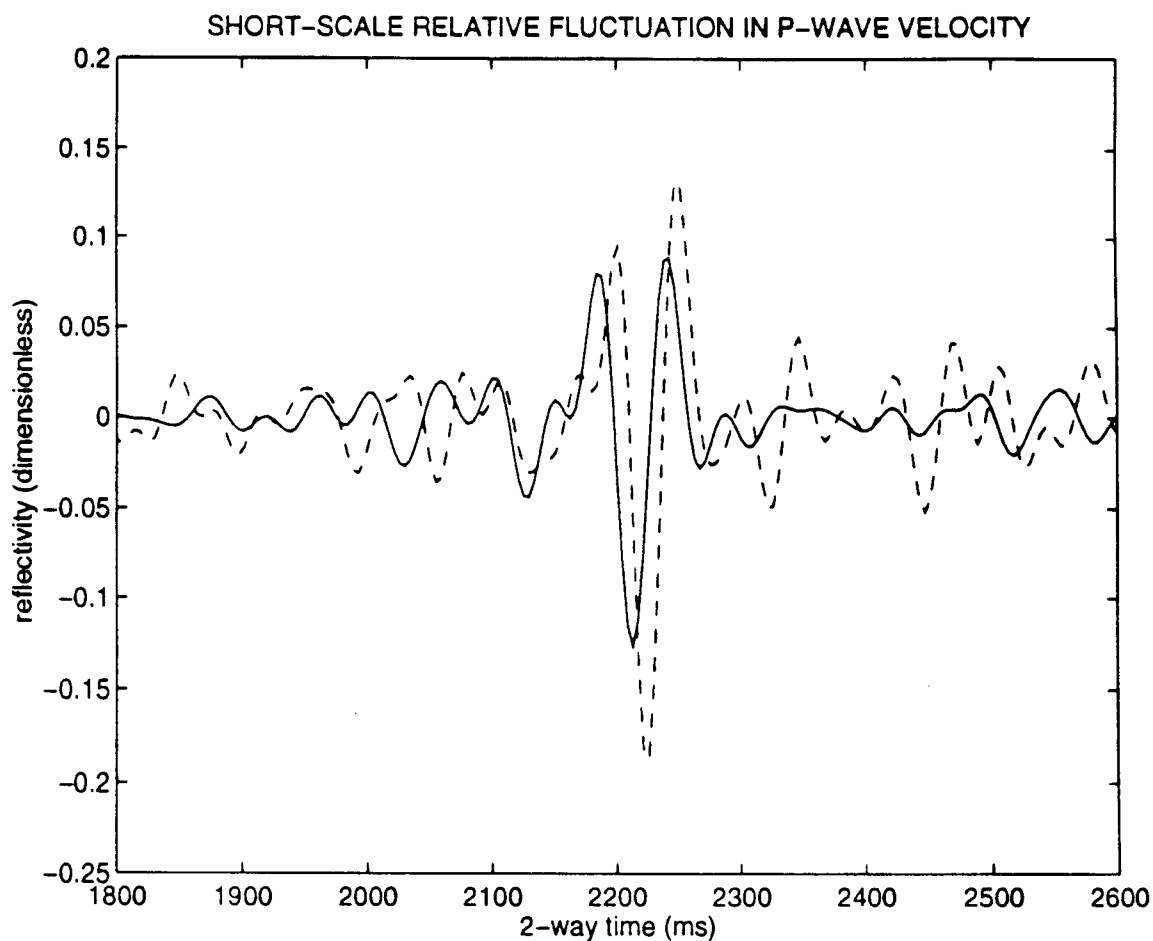
Figure 4.15 Comparison of the independent well-log measurement of the relative short-scale fluctuation in the P-wave velocity with the result of inversion done on CMP6 using an isotropic source estimate from inversion (Experiment 4). The solid line shows the inversion result (scaled and shifted left 67ms). The dashed line shows the detrended well log. Both graphs have been plotted as a function of two-way time and filtered to match the frequency content of the source.
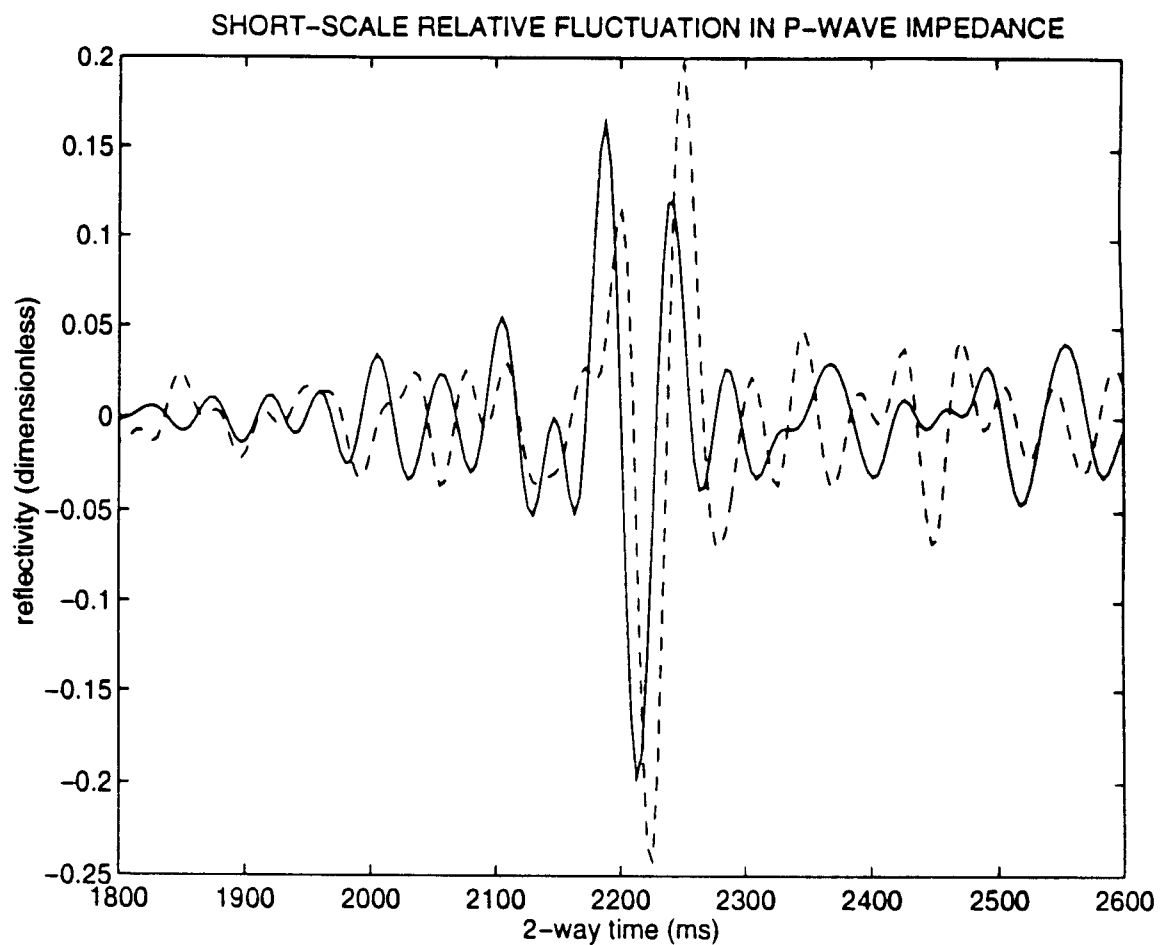
Figure 4.16 Comparison of the independent well-log measurement of the relative short-scale fluctuation in the P-wave impedance with the result of inversion done on CMP6 using an isotropic source estimate from inversion (Experiment 4). The solid line shows the inversion result (scaled and shifted left 67ms). The dashed line shows the detrended well log. Both graphs have been plotted as a function of two-way time and filtered to match the frequency content of the source.
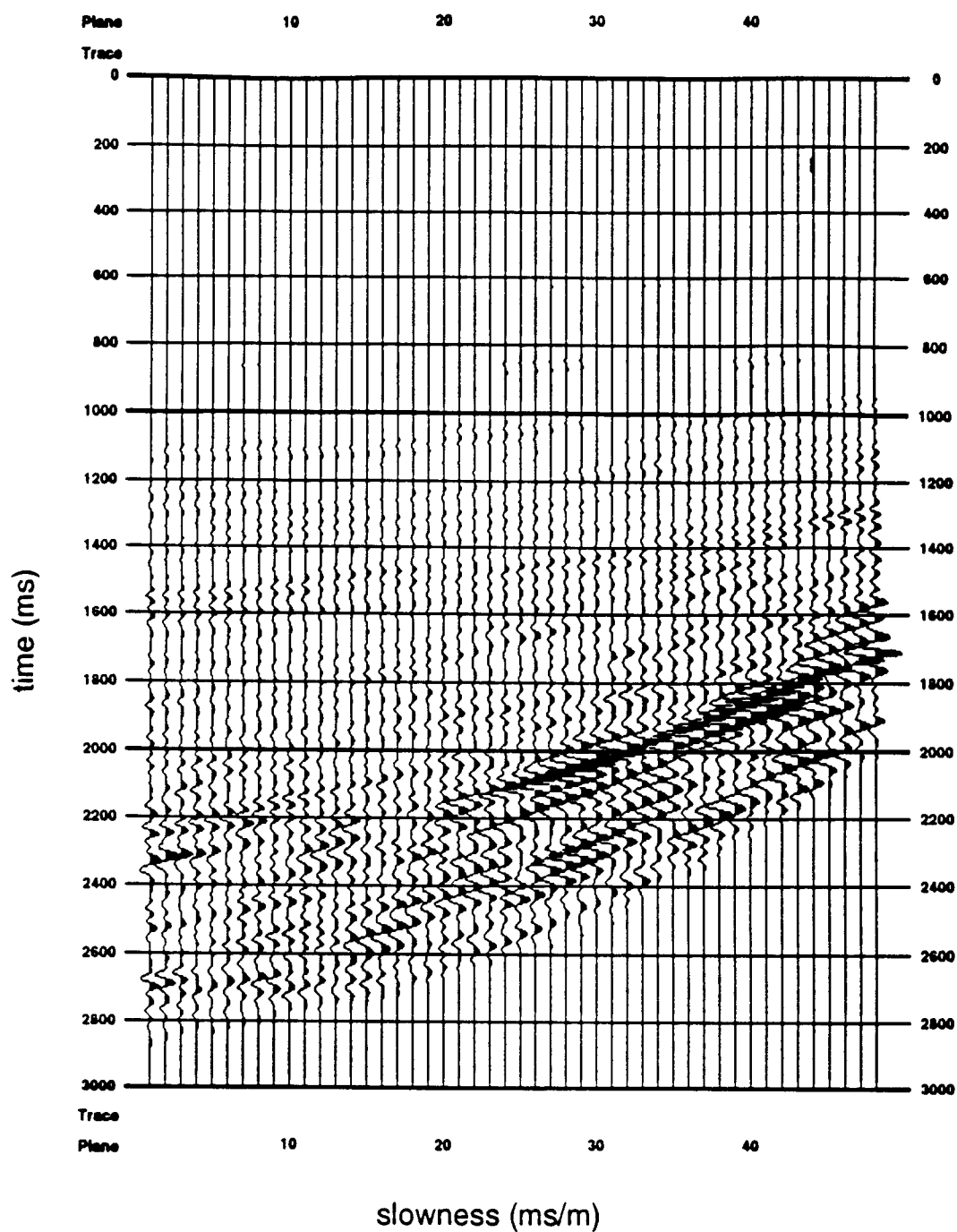
Figure 4.17 Difference between the actual data (CMP6) and the data predicted by parameters gotten from inversion using an isotropic inversion-estimated source (fixed background velocity) (Experiment 4). The data misfit is plotted on the same scale as the CMP6 data.
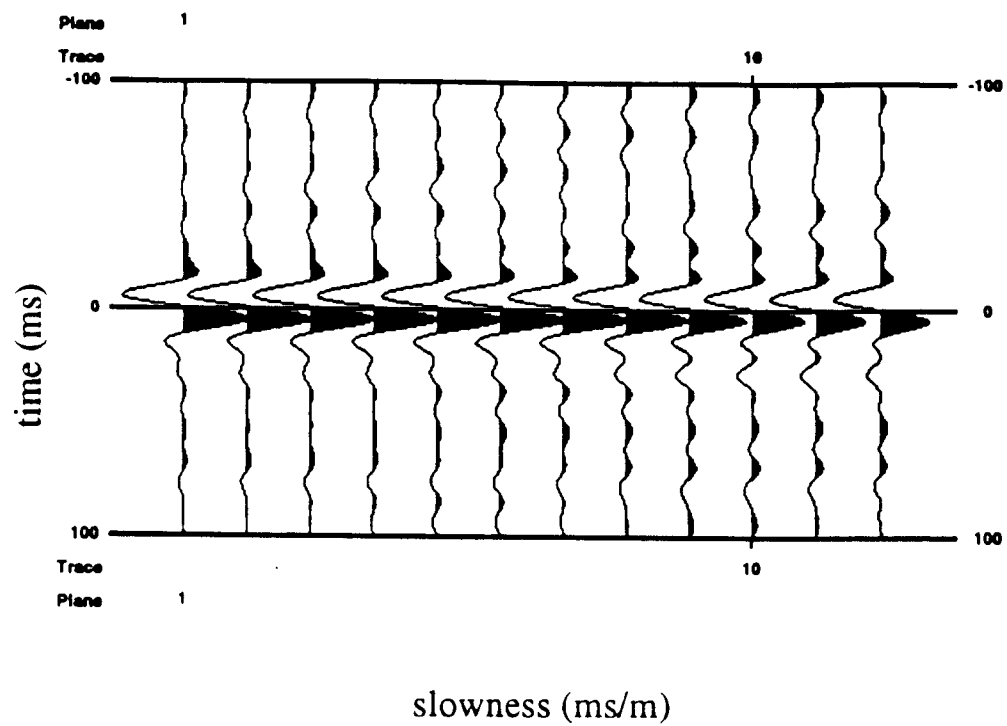
Figure 4.18 The estimated anisotropic source from the linear source-reflectivity inversions where only every fourth trace is shown for clarity (Experiment 5).
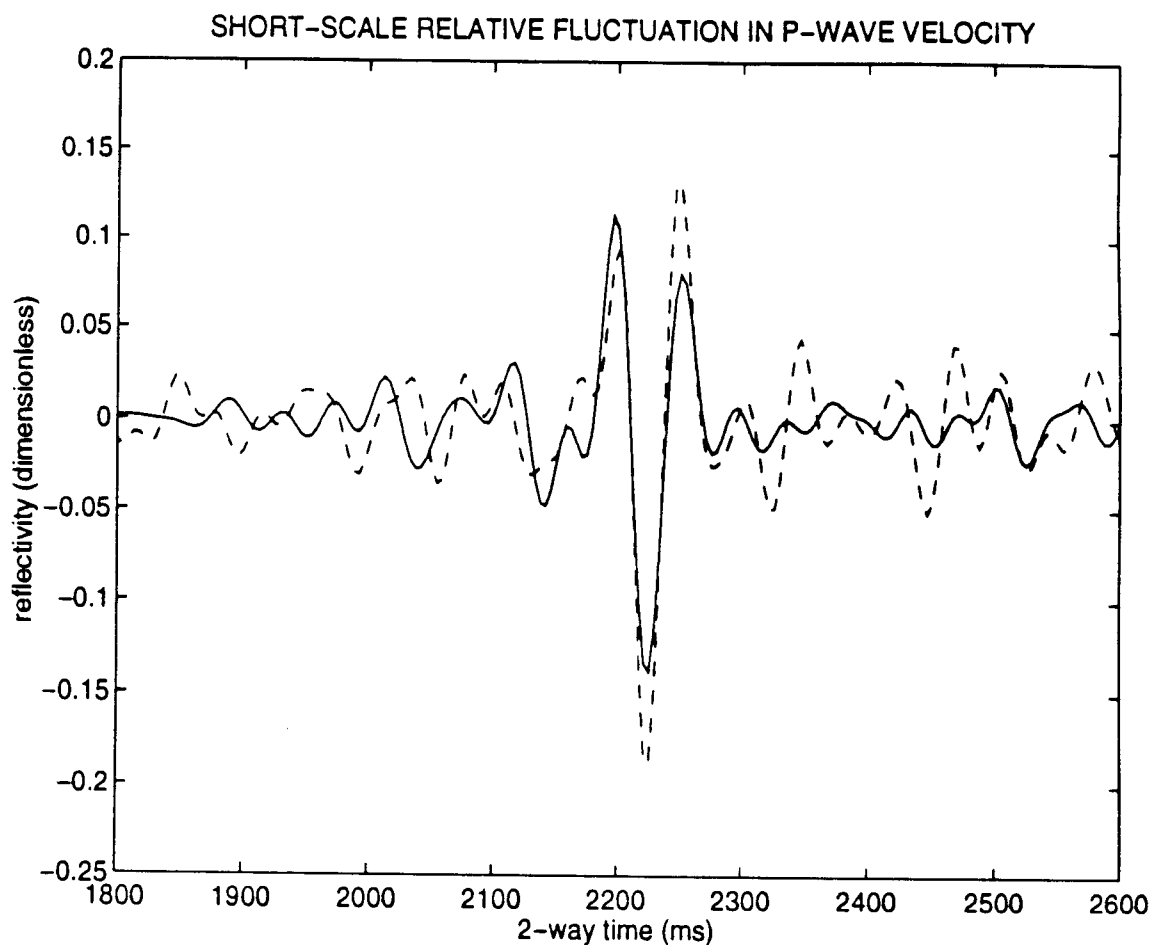
Figure 4.19   Comparison of the independent well-log measurement of the relative short-scale fluctuation in the P-wave velocity with the result of inversion done on CMP6 using an anisotropic source estimate from inversion (Experiment 5). The solid line shows the inversion result (scaled and shifted left 67ms). The dashed line shows the detrended well log. Both graphs have been plotted as a function of two-way time and filtered to match the frequency content of the source.
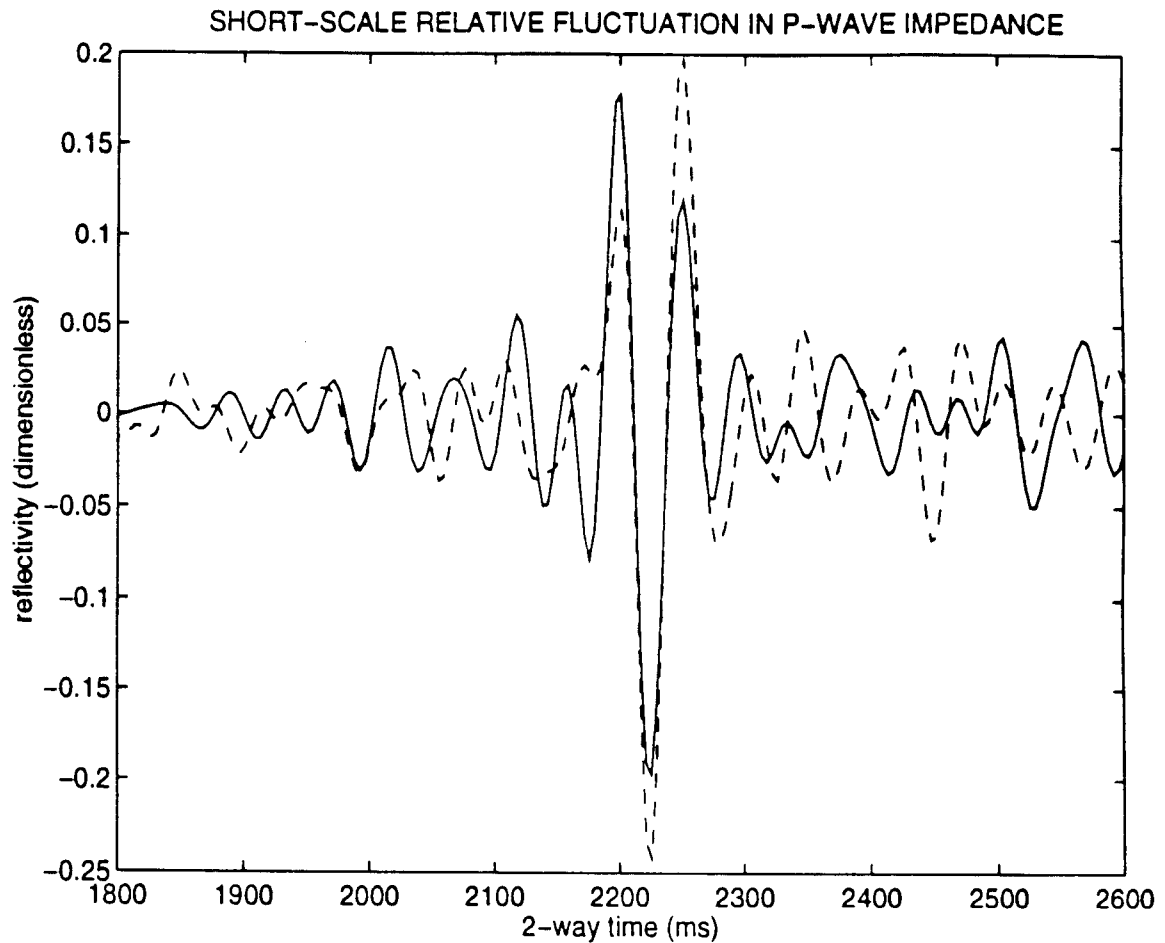
Figure 4.20   Comparison of the independent well-log measurement of the relative short-scale fluctuation in the P-wave impedance with the result of inversion done on CMP6 using an anisotropic source estimate from inversion (Experiment 5). The solid line shows the inversion result (scaled and shifted left 67ms). The dashed line shows the detrended well log. Both graphs have been plotted as a function of two-way time and filtered to match the frequency content of the source.
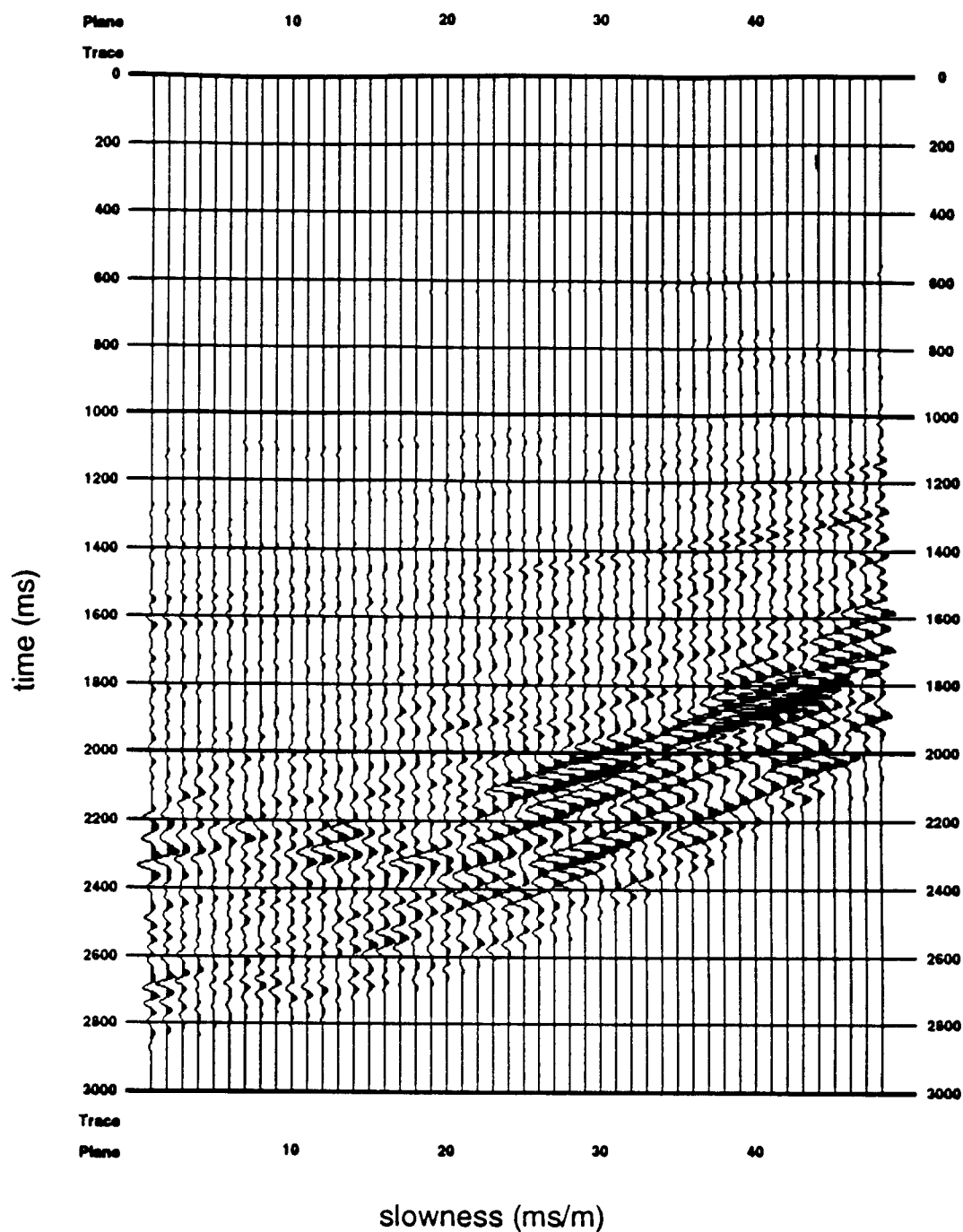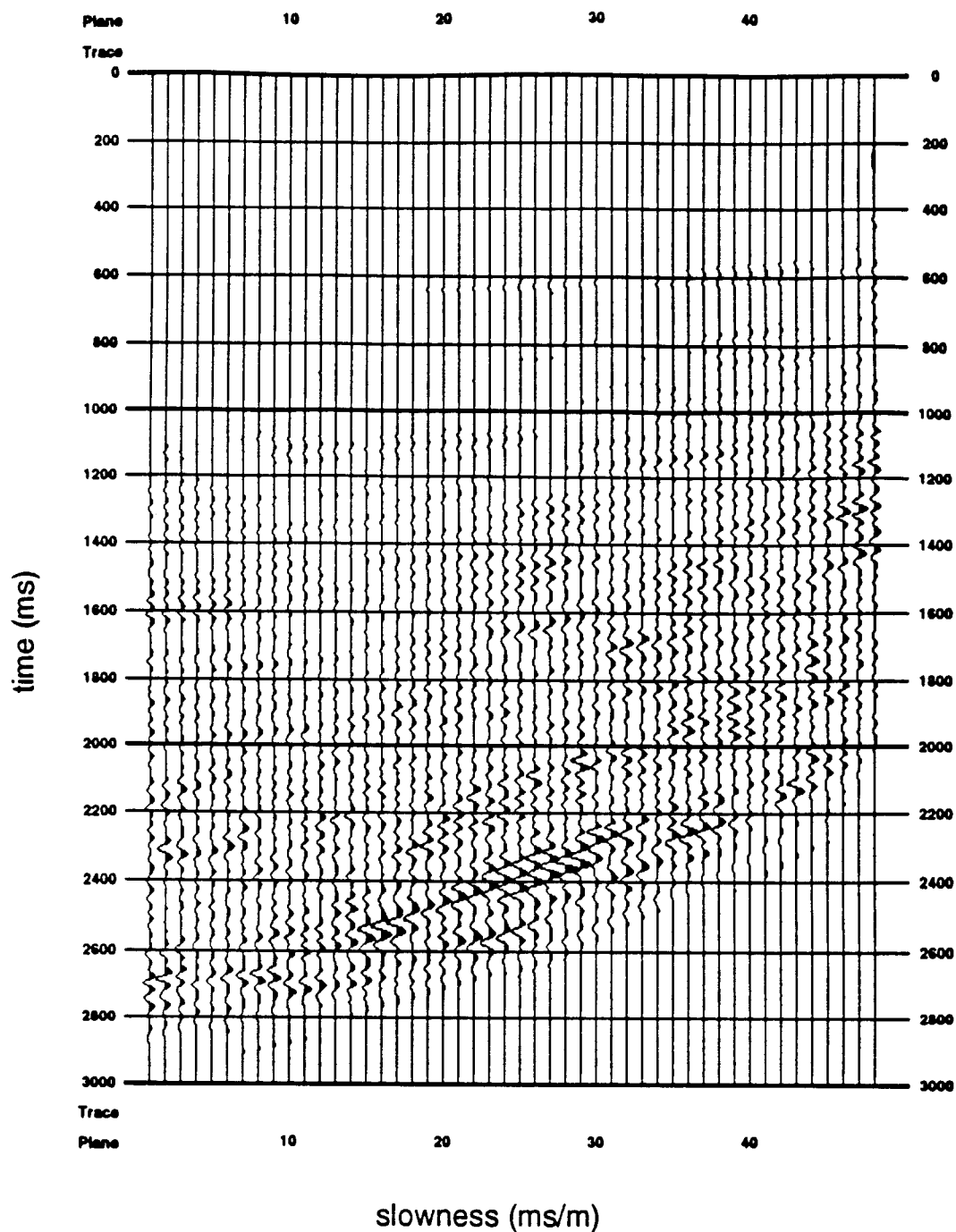
Figure 4.21 Difference between the actual data (CMP6) and the data predicted by parameters gotten from inversion using an anisotropic inversion-estimated source (fixed background velocity) (Experiment 5). The data misfit is plotted on the same scale as the CMP6 data.
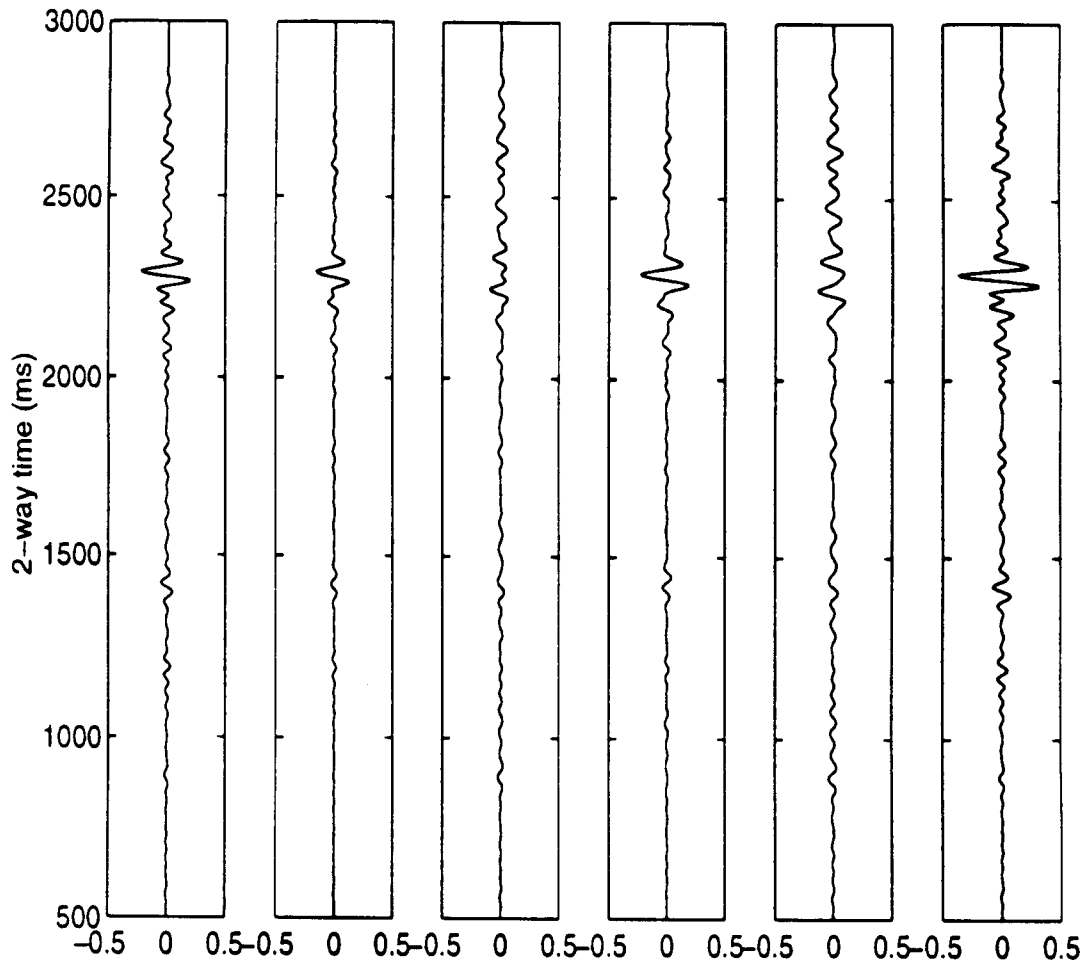
Figure 4.22 Comparison of six estimated reflectivities from the linear inversion job in which the reflectivities and an anisotropic source are estimated (Experiment 5). The reflectivity estimates shown above are (from left to right) the relative short-scale fluctuation in the P-wave impedance, the P-wave velocity, the S-wave velocity, the ratio of P-wave velocity to S-wave velocity, the shear modulus, and the combination of Lamé constants $\lambda + 2\mu$.
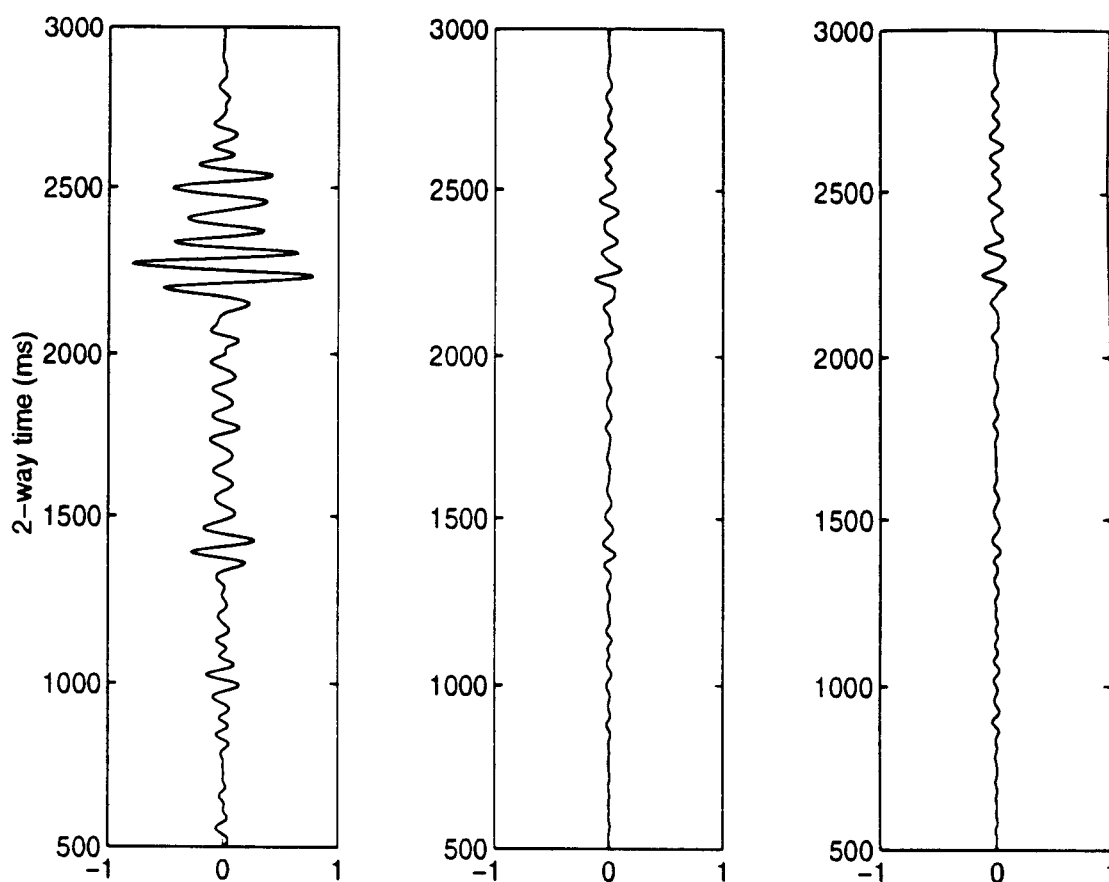
Figure 4.23 From left to right: comparison of the relative short-scale fluctuation in the shear modulus gotten from performing an inversion with the air gun model source (Experiment 3), an isotropic inversion-estimated source (Experiment 4), and an anisotropic inversion-estimated source (Experiment 5). The shear moduli have all been shifted and scaled along with the other parameters gotten via inversion so the P-wave impedance inversion result matches the log.
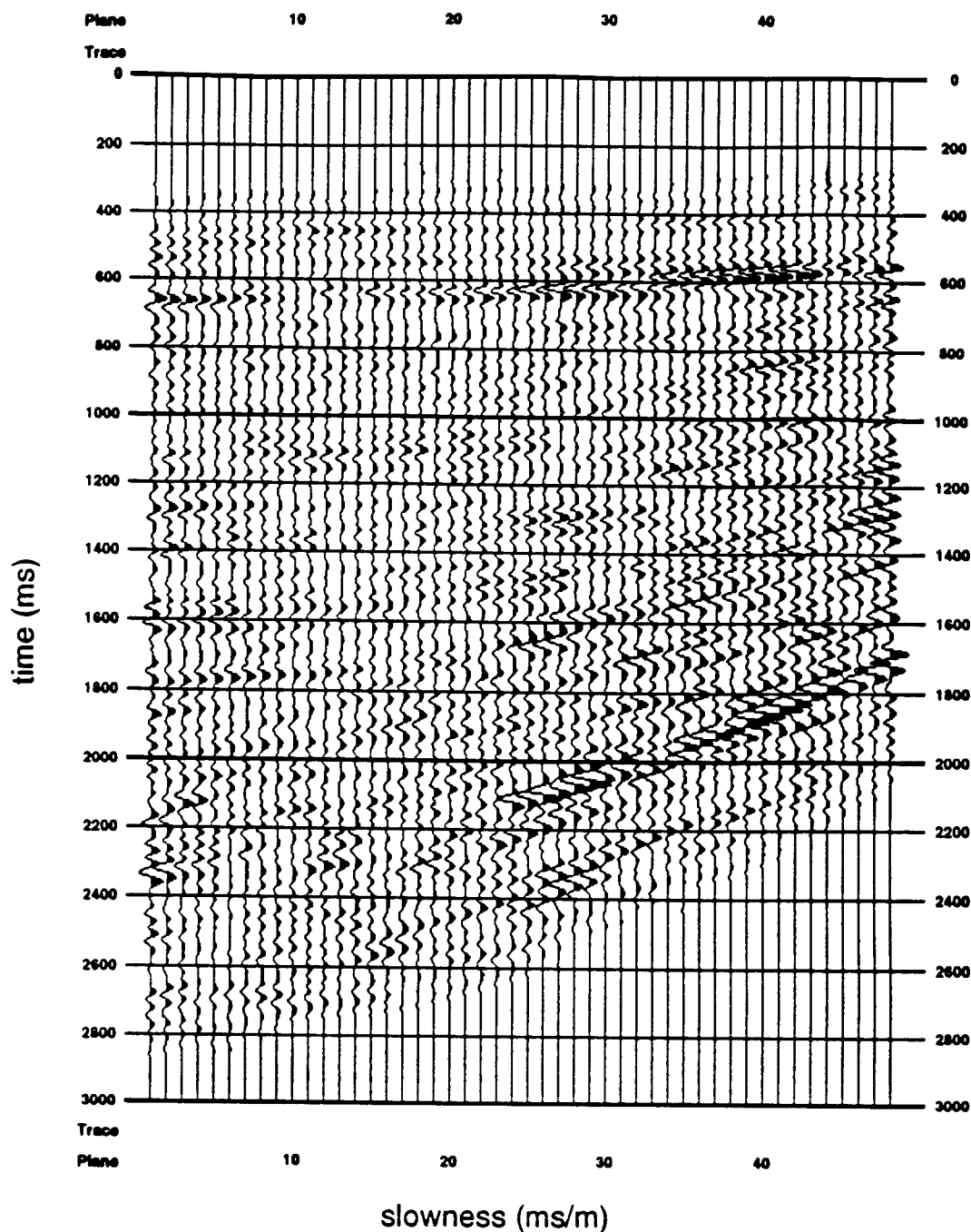
Figure 4.24 The difference between the filtered, *ungained* CMP6 data and the data predicted by an *elastic* model. The data misfit is plotted on the same scale as the ungained CMP6 data. The fixed background velocity is shown in Figure 4.4 (dark solid line). The reflectivities and energy source were gotten from inversion (Experiment 7).
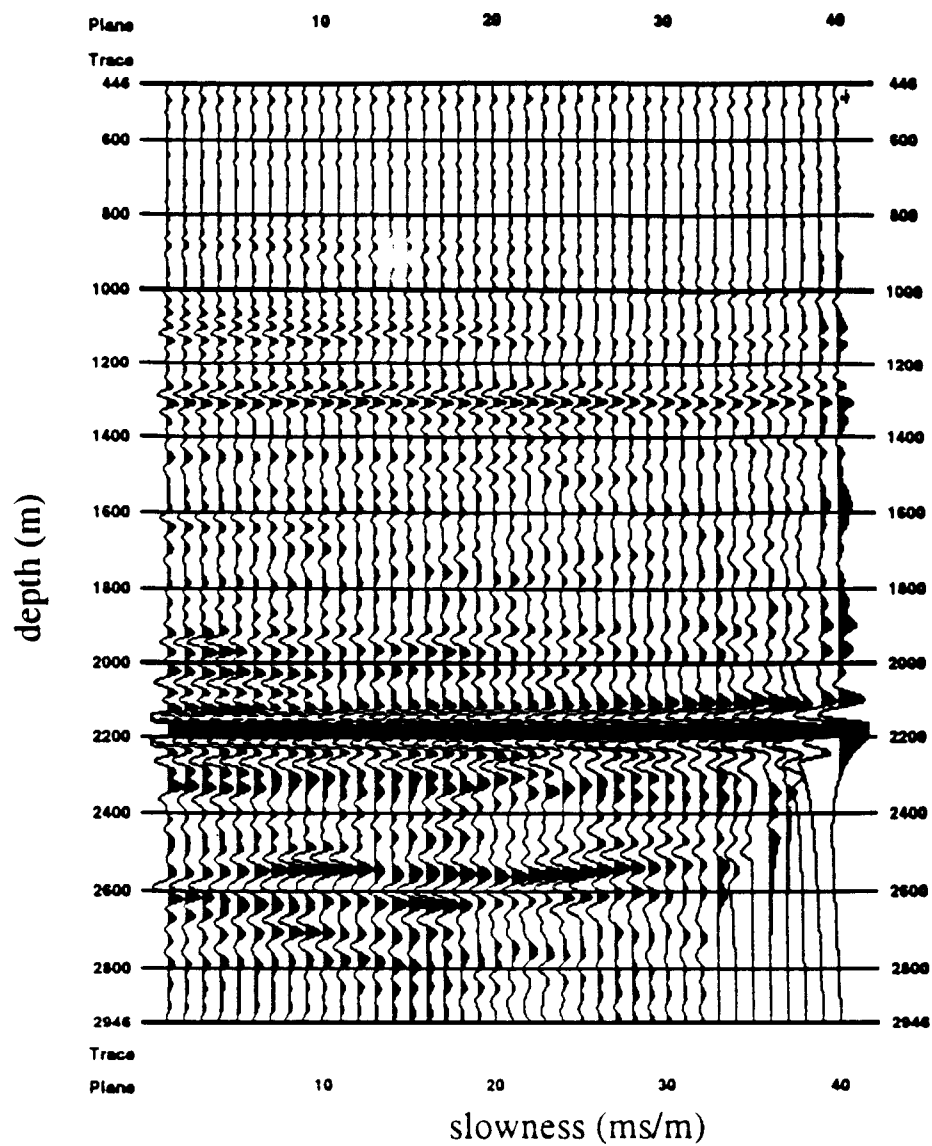
Figure 4.25 The P-wave impedance resulting from an inversion performed on CMP6 with the background velocity shown in Figure 4.4 (dark solid line) and the source estimate shown in Figure 4.18.

88
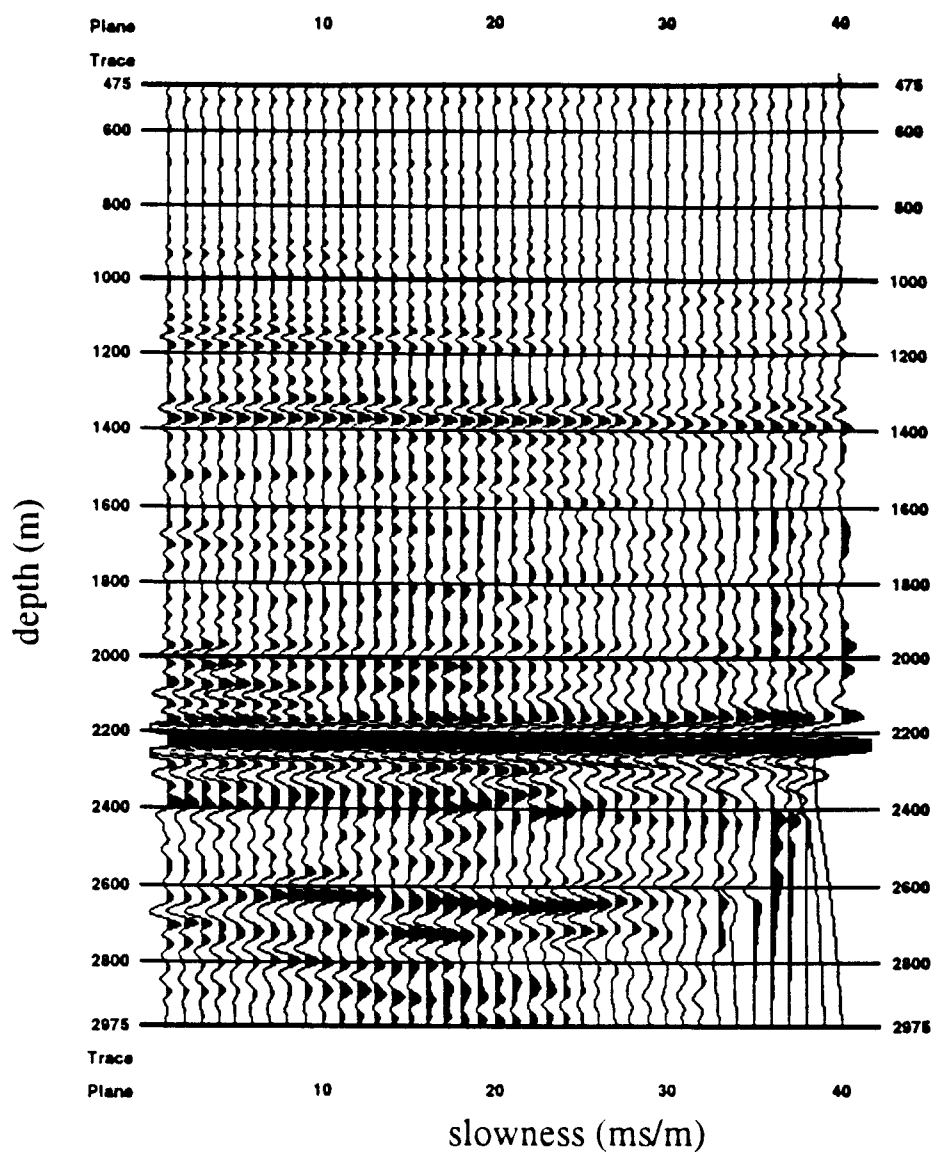
Figure 4.26 The p-wave impedance resulting from an inversion performed on CMP6 with the background velocity shown in Figure 4.4 (dashed line) and the source estimate shown in Figure 4.18.

Chapter 5

A Few More Interesting Marine Data Experiments

5.1 Introduction

Many numerical experiments were carried out to produce the real data inversion picture painted in Chapter 4. In this chapter I will describe some of the other experimental results which helped us to analyze this data set and which allow us to propose the process for doing inversion on real data outlined in the last chapter. The mathematical model, seismic marine experiment (data set), and software package are the same as those described in Chapter 4. Rather than repeat that information in this chapter, I will immediately focus on the experimental results themselves.

In this chapter we start by showing a series of experiments which makes clear that the energy source which generated this (and likely most) real data is anisotropic. Although the idea that real seismic sources are anisotropic should not be a surprise to exploration seismologists, it is worthwhile to highlight the large amount of misfit which results when one assumes an isotropic source because many scientists still make the assumption that an isotropic source will suffice. (Some recent references include [51], [22], [35].) We also performed an inversion using an isotropic source estimated via predictive deconvolution. Predictive deconvolution is a traditional signal processing technique for estimating the source in a seismic experiment. In Section 5.3 we compare the misfit graph gotten from this experiment to a misfit graph from the series described in Section 5.2.

The series of experiments detailed in Section 5.2 also includes multi- and single-parameter reflectivity inversions. We would like to compare this series of seven experiments where in each subsequent experiment more "degrees of freedom" are added; i.e., the source is allowed to be increasingly anisotropic and the reflectivity term contains either one or three parts to be estimated. We would like to ascertain whether misfit is enough for one to conclude that multiparameter inversion really gives ad-

ditional information over single parameter inversion especially when one allows the source to be anisotropic.

In the last chapter we showed that inverting single gathers CMP1, 5, 7, and 11 for the three reflectivities and source we were able to fit the data with'about 30% misfit (see Table 4.1 Experiment 6). In this chapter we invert the data set CMP1 for the reflectivities only with the source and background velocity fixed at the CMP6 results. In the course of performing the inversion experiments, we found that one assumption we had made, that the data is well-described by a layered model, did not hold. Even for this small subset of the total data set (where the distance between gathers CMP1 and CMP11 is 375m) one notes in Figure 5.19 that the difference between gathers CMP1 and CMP6 (offset 187m) is nearly as large as the data itself. As discussed in Chapter 4, some initial processing was applied in an effort to mitigate the nonlayered effects, namely, time migration in the midpoint dip domain to collapse diffractions. Nonetheless, lateral heterogeneities in the near surface layer still influence the data used in the inversions. Re-estimation of the source for each gather is necessary to compensate for this modeling deficiency.

Finally, we discuss briefly the importance of the initial source guess for this type of inversion.

5.2 Comparison of a Series of Experiments for Source and Reflectivity With Increasing Model Complexity

This section addresses two issues. First, we indicate how important an *anisotropic* source is for fitting real data. Second, we examine the misfit between actual and predicted data for single and multi-reflectivity inversions and show how difficult it is to use only this crude measure for analyzing the accuracy of the parameter estimates. Of the experiments described in this section, five are single-reflectivity inversions and two are multi-reflectivity. In each experiment all inputs are the same except for the number of reflectivities estimated and the number of terms allowed in the Legendre series describing the source. (An isotropic source is described by a one-term Legendre series.) The experiments were all performed on the common midpoint gather 6, and the background velocity is the inversion result for CMP6 data given in Figure 4.4, the dark solid line. Figures 5.1-5.7 show the difference (misfit) between the actual and predicted data gotten from inverting for the reflectivities and source as specified in Table 5.1. Figures 5.8-5.14 show the misfit between actual and predicted data in

a zone around the target event. For the close to normal incidence record ($p = .1158$ ms/m), this interval is 2200-2400ms. As slowness (p) increases, this muted zone follows the main event upward in time, keeping the width of this envelope approximately that of the original 200ms window. *Note that all the data misfit pictures shown are plotted on the same scale as the actual data.* The percent data misfit for the whole recording interval and for the mute zone are given in Table 5.1.

Experiment number	1	2	3	4	5	6	7
No. of reflectivities estimated	1	1	1	1	1	3	3
No. of source components	1	3	10	31	40	1	31
Percent misfit	71%	55%	41%	33%	30%	51%	27%
Percent misfit in target zone	64%	46%	26%	13%	10%	42%	11%

Table 5.1 Misfit comparisons for seven experiments in which single or multi- reflectivities and sources are estimated.

Discussion of Results:

1. Large slowness-value records are not well matched on the $\tau - p$ seismogram by sources which are isotropic or mildly anisotropic (having $\leq$ 10 Legendre components).

2. From the misfit pictures and percentages alone it is extremely difficult to ascertain how reliable the information derived from multiparameter inversion is over single parameter inversion, especially with an anisotropic source. However, Table 5.1 indicates that the inversion-estimated parameters in Experiment 7 (*three reflectivities* and a 31-component source) give distinctly better fit to data than do the parameters estimated by inversion in Experiment 4 (a *single reflectivity* and 31-component source).

3. The muted misfit pictures indicate that the single-reflectivity, 40-component source in Experiment 5 does a good job of fitting the data in the target zone. However, over the whole recording time interval, the multi-reflectivity, 31-component source (Experiment 7) does a slightly better job of fitting the data. The difference between the two misfits is subtle. The fact that the misfit for the multi-reflectivity, 31-component source (Experiment 7) is *overall* smaller than for the single-reflectivity, 40-component source (Experiment 5) implies

more of the data in the zone *above* the target is being fit in Experiment 7. For instance, the first noticeable event in the data (see Figure 4.3) is at about 1600ms. The multi-reflectivity, 31-component source result does a better job of completely predicting this event than does the single-reflectivity, 4b-component source. This observation is a small clue that in fact there is information about the second and third reflectivities (not just P-wave impedance) in the data.

5.3 Comparison of a Source Generated by Traditional Techniques with the Inversion-Estimated Isotropic Source

We also used the commercial package, ProMAX, to estimate an isotropic *minimum phase* (with energy concentrated at the front end of the pulse) seismic wavelet (source) for this data. The technique ProMAX implements is called least squares Wiener filtering (or predictive deconvolution). The general ideas are described in detail in [45]. We define briefly the method and parameters for this particular use of the package below.

For this data which is defined over a time interval from 0–3000ms with samples taken every 2ms, let z_t, x_t, f_t, y_t, $t = 0, 2, 4, \ldots, 3000$, be discrete time series. The desired output (in this case a seismic trace filtered so the estimated source has been removed) is represented by z_t. The actual seismic trace is x_t, and f_t is a filter to be determined. The cross-correlation of x_t and f_t is y_t, or

$$y_\tau = \sum_k f_k x_{k+\tau}.$$

Given the cross-correlation of the actual and desired traces

$$\phi_{xz}(\tau) = \sum_k x_k z_{k+\tau}$$

and the autocorrelation of the actual trace with itself

$$\phi_{xx}(\tau) = \sum_k x_k x_{k+\tau}$$

we can determine the filter f_t that minimizes the function $\|z_t - y_t\|^2$ by solving the normal equations

$$\sum_t f_t \phi_{xx}(\tau - t) = \phi_{xz}(\tau), \qquad \tau = 0, 1, 2, \ldots.$$

Once ProMAX finds the appropriate filter to deconvolve the source from the data, the source itself is estimated by finding an inverse to the first filter. For our application, the final output wavelet was defined over an interval of 1000ms. The filter function f_t was defined over an interval of 120ms. Finally, the discrete convolutions were implemented so that the time gate for summation was $k = 500, \ldots, 2000$ms.

The inversion for the three reflectivities using this predictive deconvolution source allowed a fit to data with 57% rms error. The source estimate is shown in Figure 5.15. The associated data misfit is shown in Figure 5.16. For comparison, the isotropic inversion-estimated source we derived as well as the associated data misfit for that experiment (see Table 5.1 Experiment 6) are redisplayed in Figures 5.17 and 5.18. This inversion-estimated source is clearly not minimum phase. The time gates in the process described above determine the width of the estimated source pulse. The inversion-estimated source appears to be higher frequency than the predictive deconvolution estimate we obtained.

5.4 Comparison of Source-Reflectivity Inversions of Two Different Common Midpoint Data Gathers

In this section we contrast three experiments. The first two experiments are inversions for the source and three reflectivities for common midpoint data gathers 6 and 1 respectively. The third experiment is an inversion for the reflectivities which correspond to CMP1 data. The source is not estimated. The source estimate from the first experiment (CMP6 data) is used. The two sources used are shown in Figures 5.20 and 5.21.

Figures 5.22, 5.23, and 5.24 show the inverted reflectivities for the three experiments described in the paragraph above. One sees again that the modeling assumption of a layered earth cannot be correct as the results for the two different gathers are so different. Also, a source estimated from one gather is not sufficient for one to estimate reflectivities from another gather. Interestingly, the P-wave impedance reflectivity is identical at the target for the three experiments. The shear wave velocity and P-wave velocity divided by density reflectivities are not so consistent across experiments.

5.5 The Impact of the Initial Guess on the Inversions

Finally, a source-reflectivity inversion starting from the anisotropic air gun model source is contrasted with a source-reflectivity inversion starting from an isotropic Ricker wavelet (Table 4.1 Experiment 5). Unfortunately, this experiment points out some of the problems with using the alternation algorithm. OLS inversion coupled with alternation is a somewhat inefficient local technique so it is not surprising that the results are affected by the starting source guess. The air gun model used as a starting source did not allow the alternation method to converge. The normal residual for the final source alternation experiment was 17% when the job terminated (reached the iteration limit). We noted in Chapter 4 how well the inversion results matched the well logs when we started the source estimation process with an isotropic Ricker wavelet (an extremely reasonable initial source guess). We find that although the data misfit for this second experiment is 31% after only two rounds of the alternation algorithm described in Chapter 4, the final source looks very similar to the initial air gun source and the final reflectivities deviate quite dramatically from the reflectivities estimated in Experiment 5 Chapter 4 (and, therefore, from the well logs). The initial source (air gun model) is shown in Figure 5.25 and the final source estimate for this inversion may be seen in Figure 5.26. Figures 5.27 and 5.28 show the P-wave impedance and P-wave velocity divided by density comparison between the two experiments described in this paragraph (which differ by the initial source guess). All reflectivities were shifted and scaled so the P-wave impedance results for the two experiments would match.
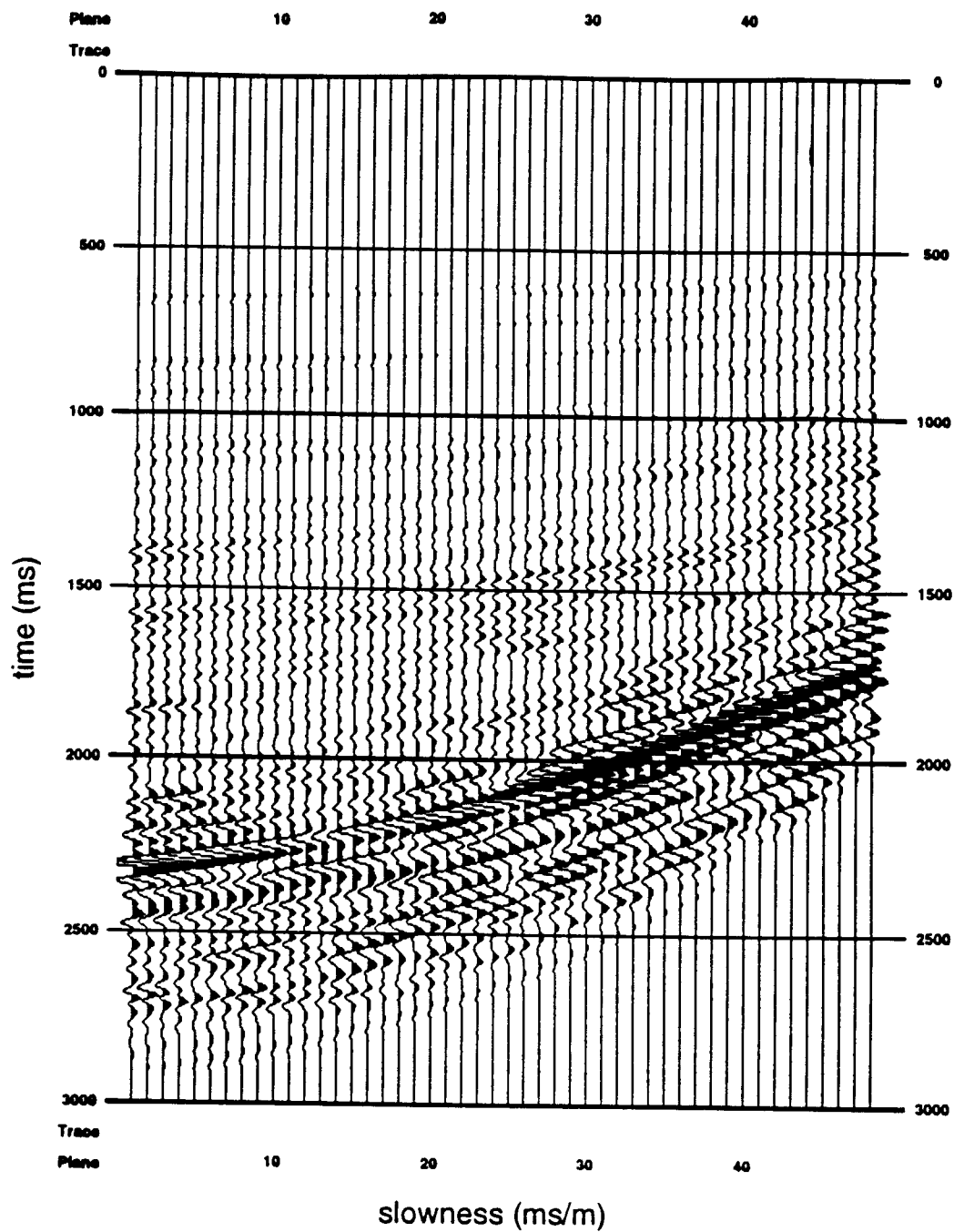
Figure 5.1 Table 5.1 Experiment 1 misfit
between actual and predicted data.
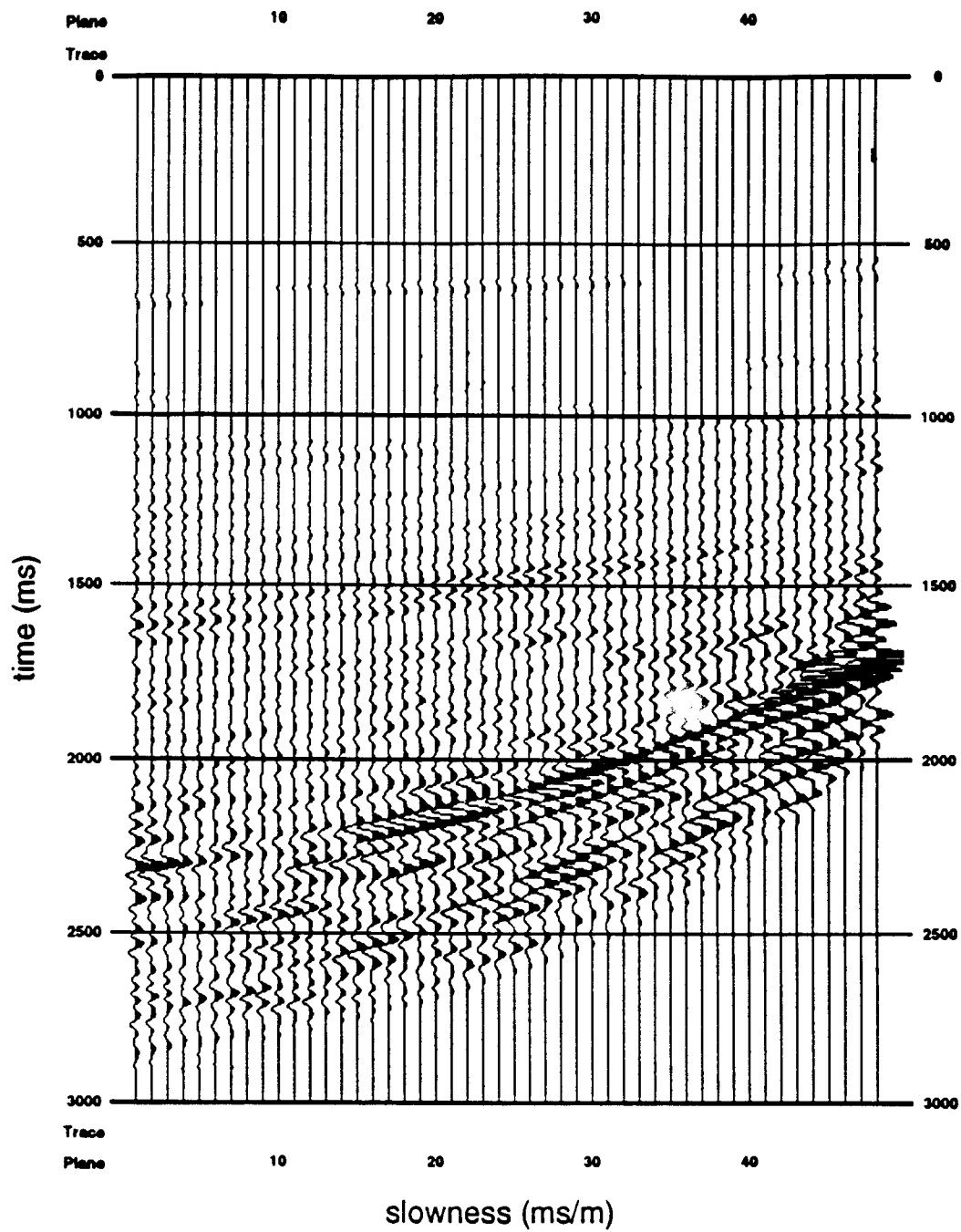
Figure 5.2   Table 5.1 Experiment 2 misfit between actual and predicted data.
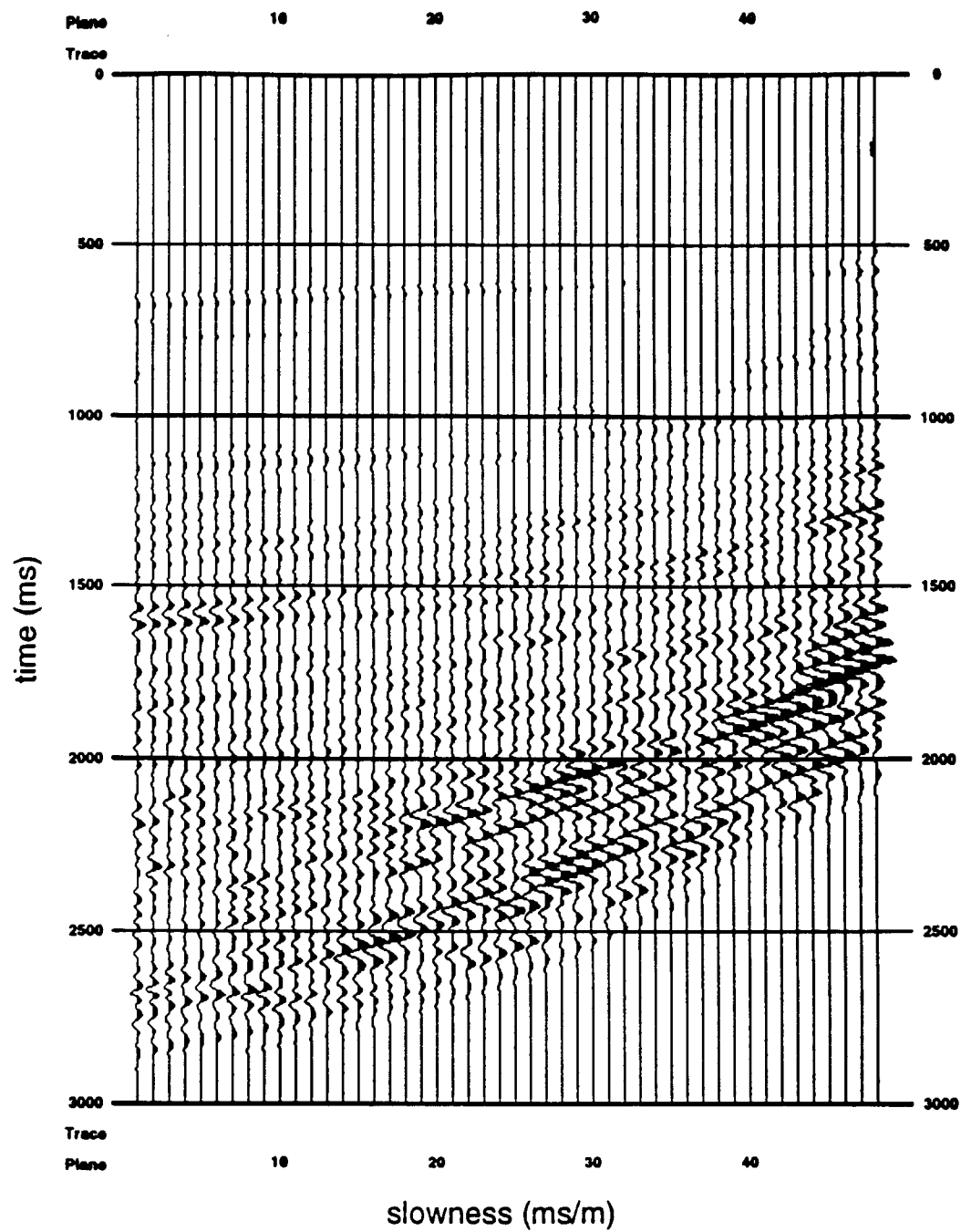
Figure 5.3 Table 5.1 Experiment 3 misfit between actual and predicted data.
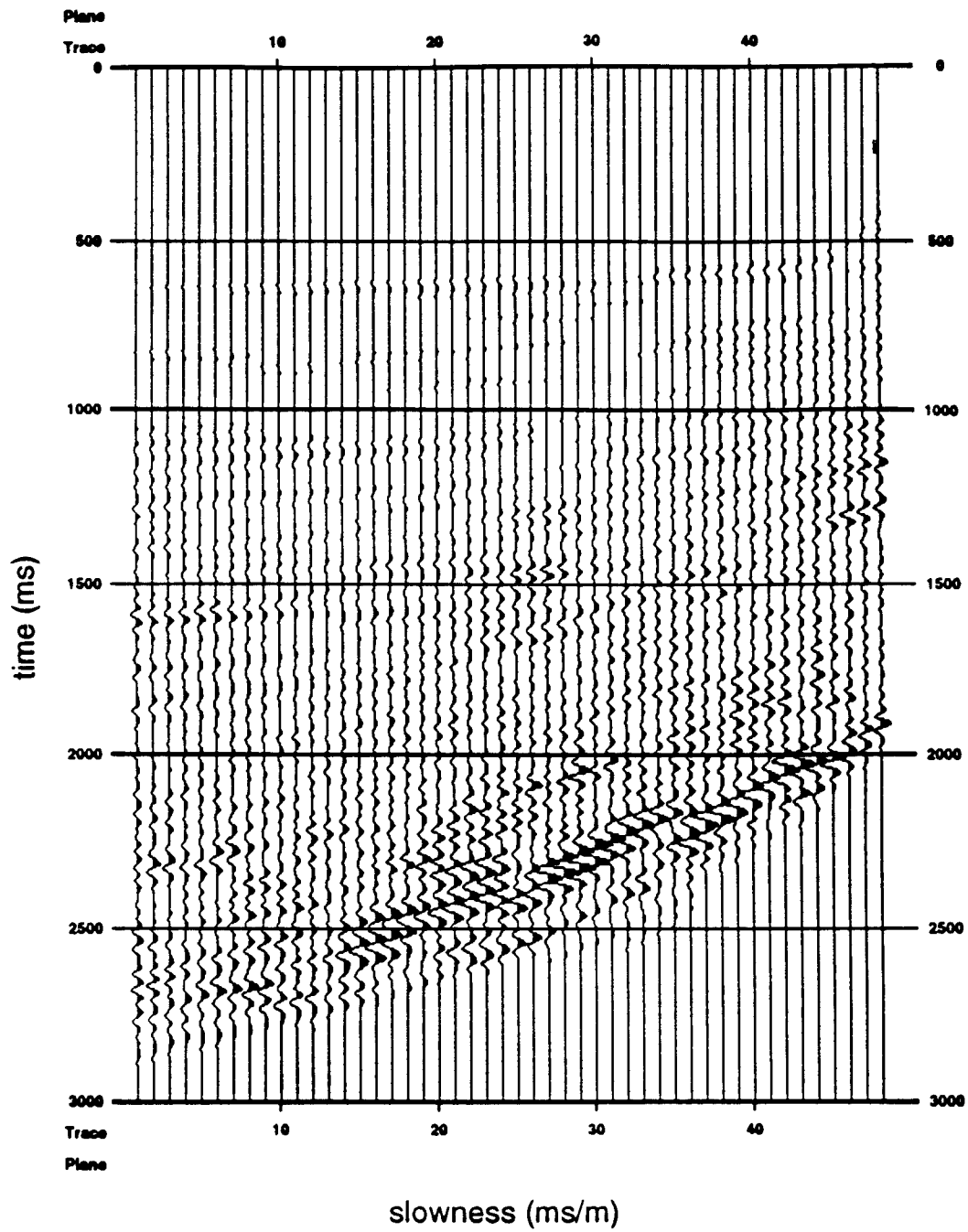
Figure 5.4 Table 5.1 Experiment 4 misfit
between actual and predicted data.
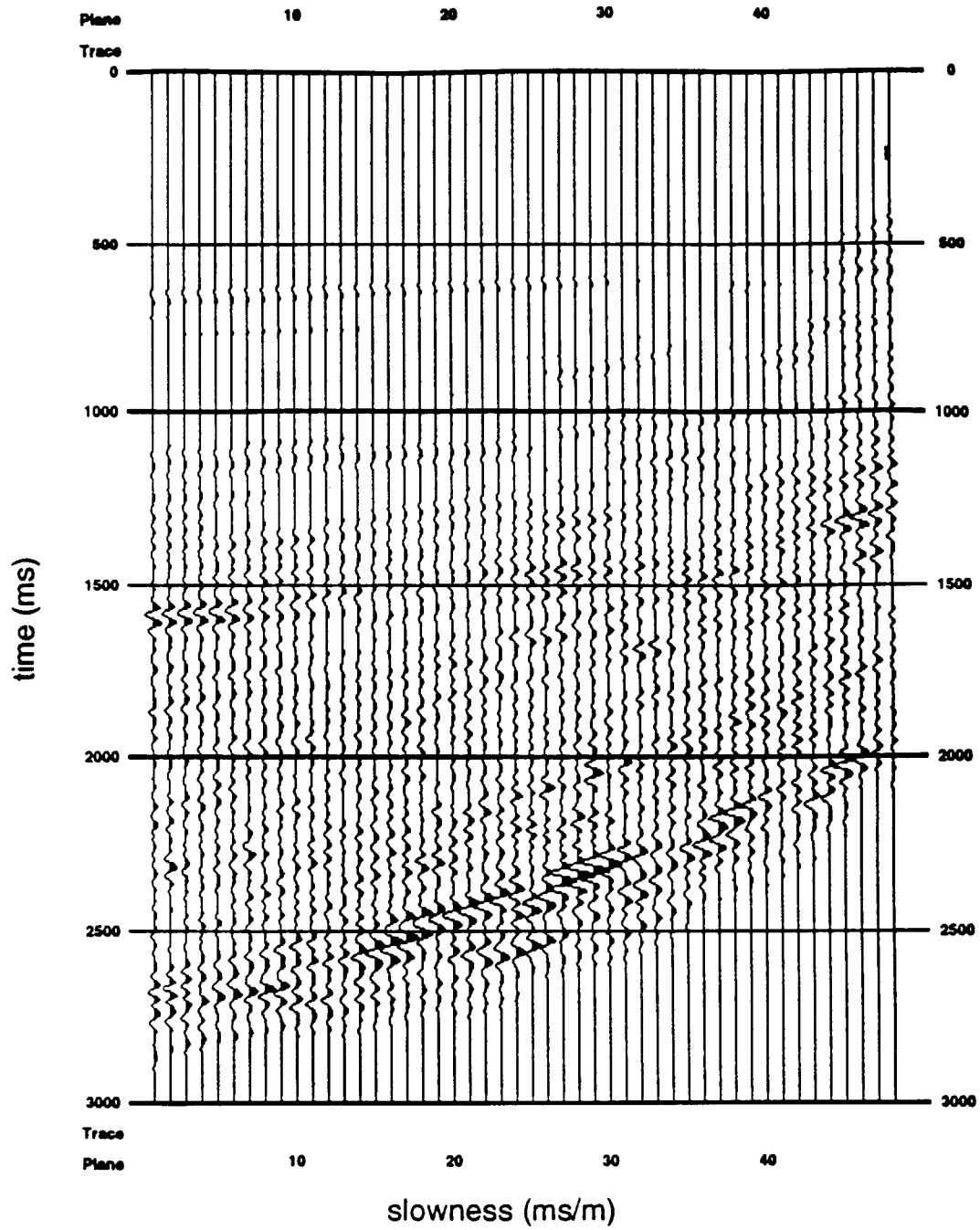
Figure 5.5 Table 5.1 Experiment 5 misfit
between actual and predicted data.
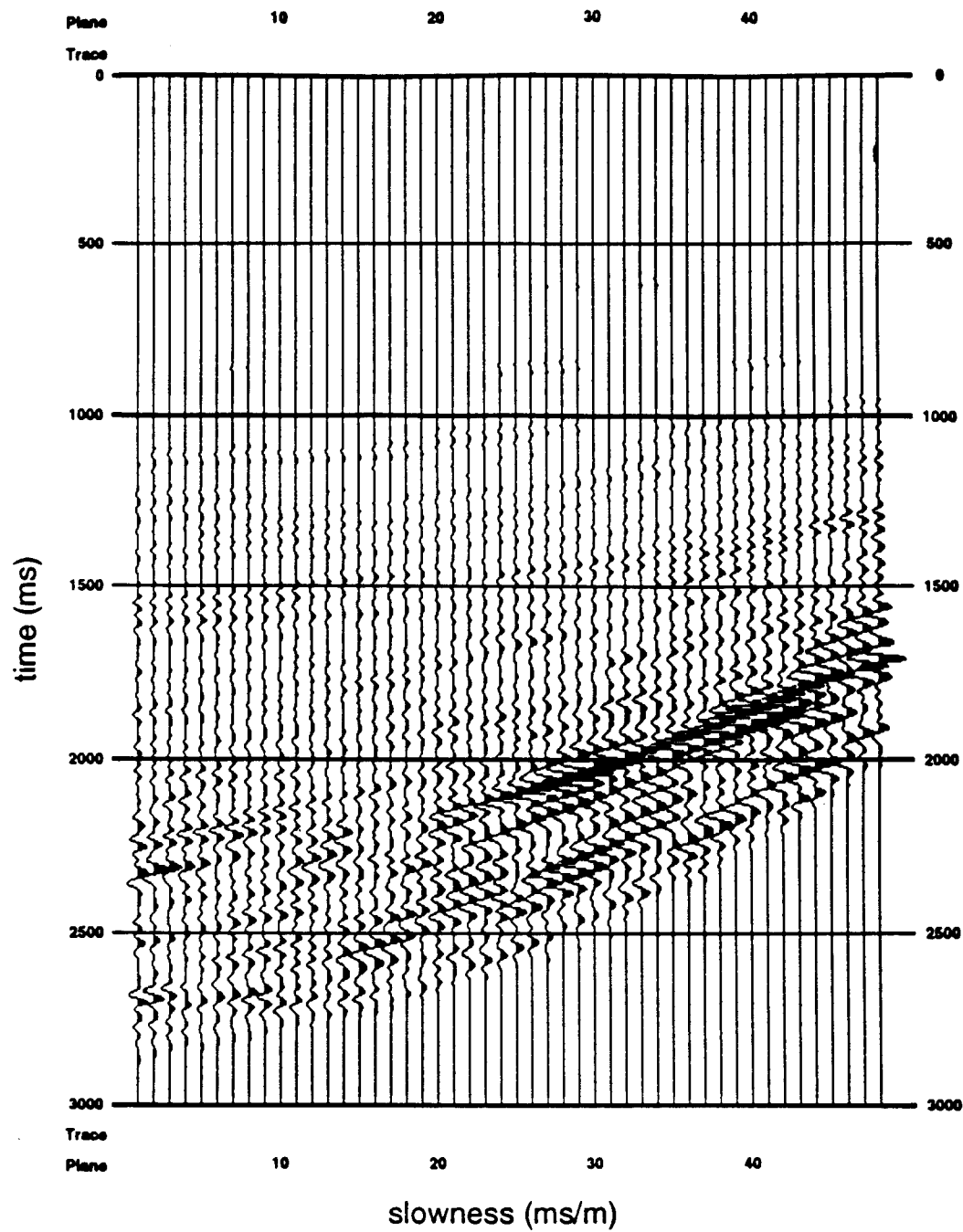
Figure 5.6 Table 5.1 Experiment 6 misfit
between actual and predicted data.
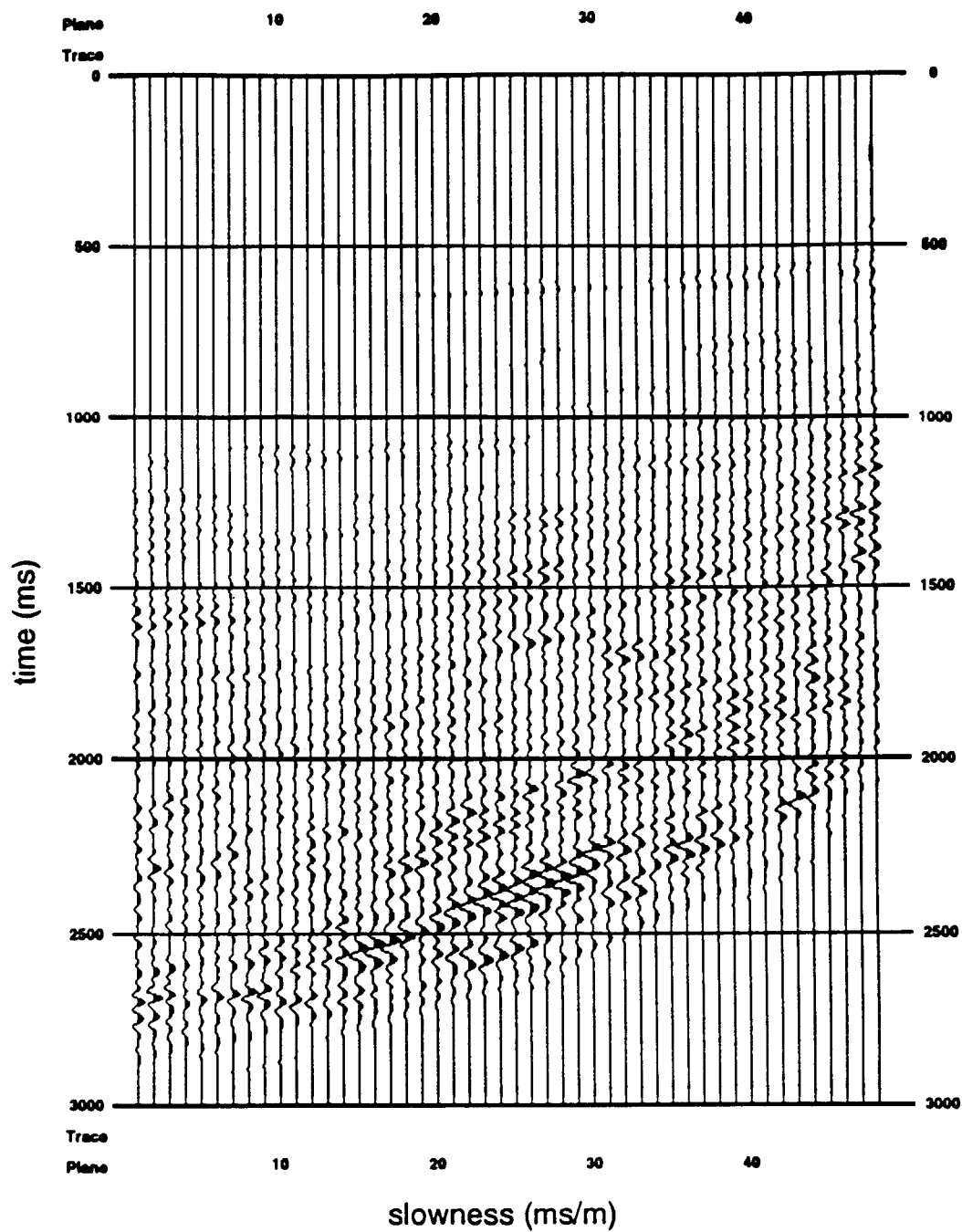
Figure 5.7   Table 5.1 Experiment 7 misfit
between actual and predicted data.
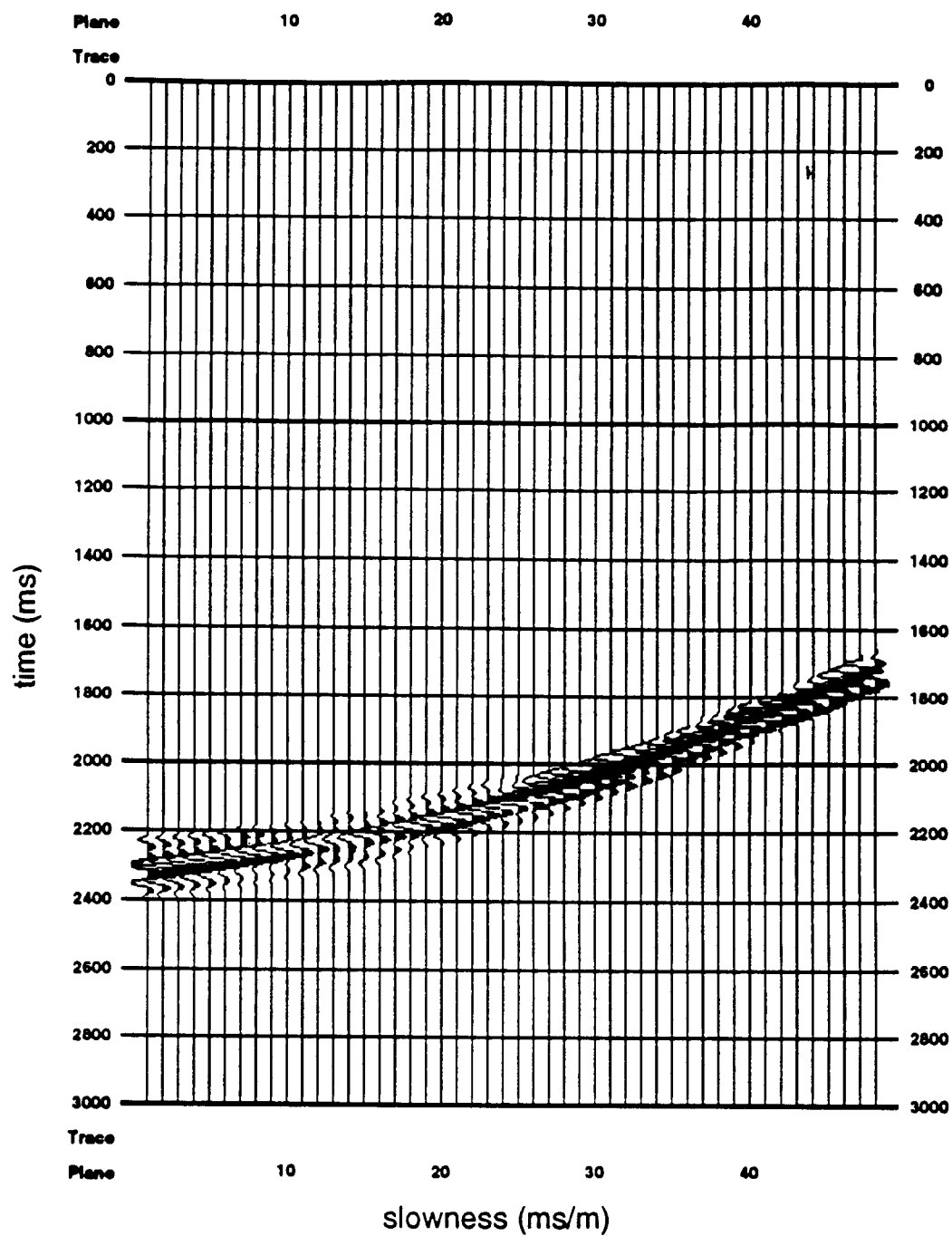
Figure 5.8 Table 5.1 Experiment 1 muted data misfit
(misfit in target range from 2200-2400ms).
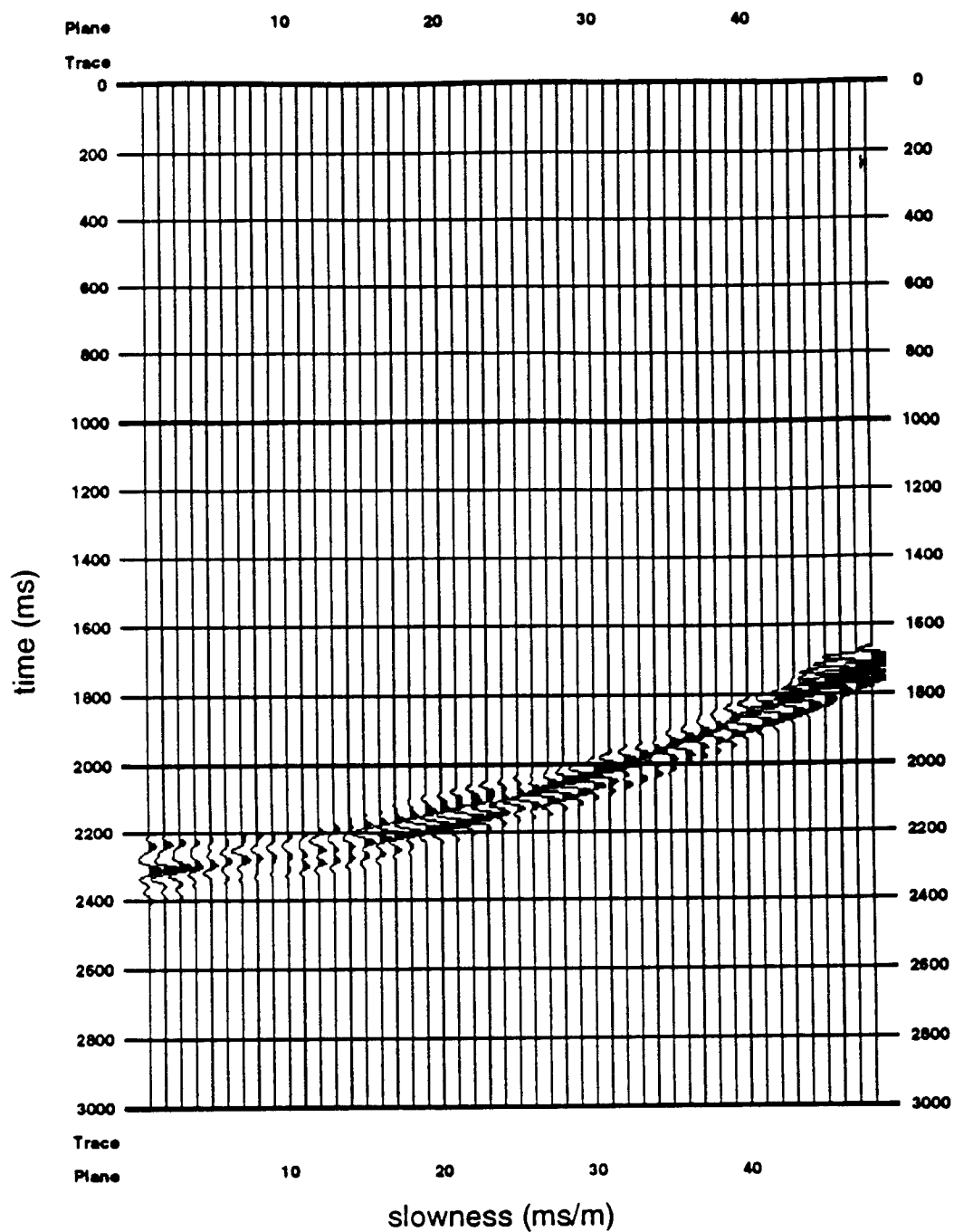
Figure 5.9 Table 5.1 Experiment 2 muted data misfit (misfit in range from 2200-2400ms).
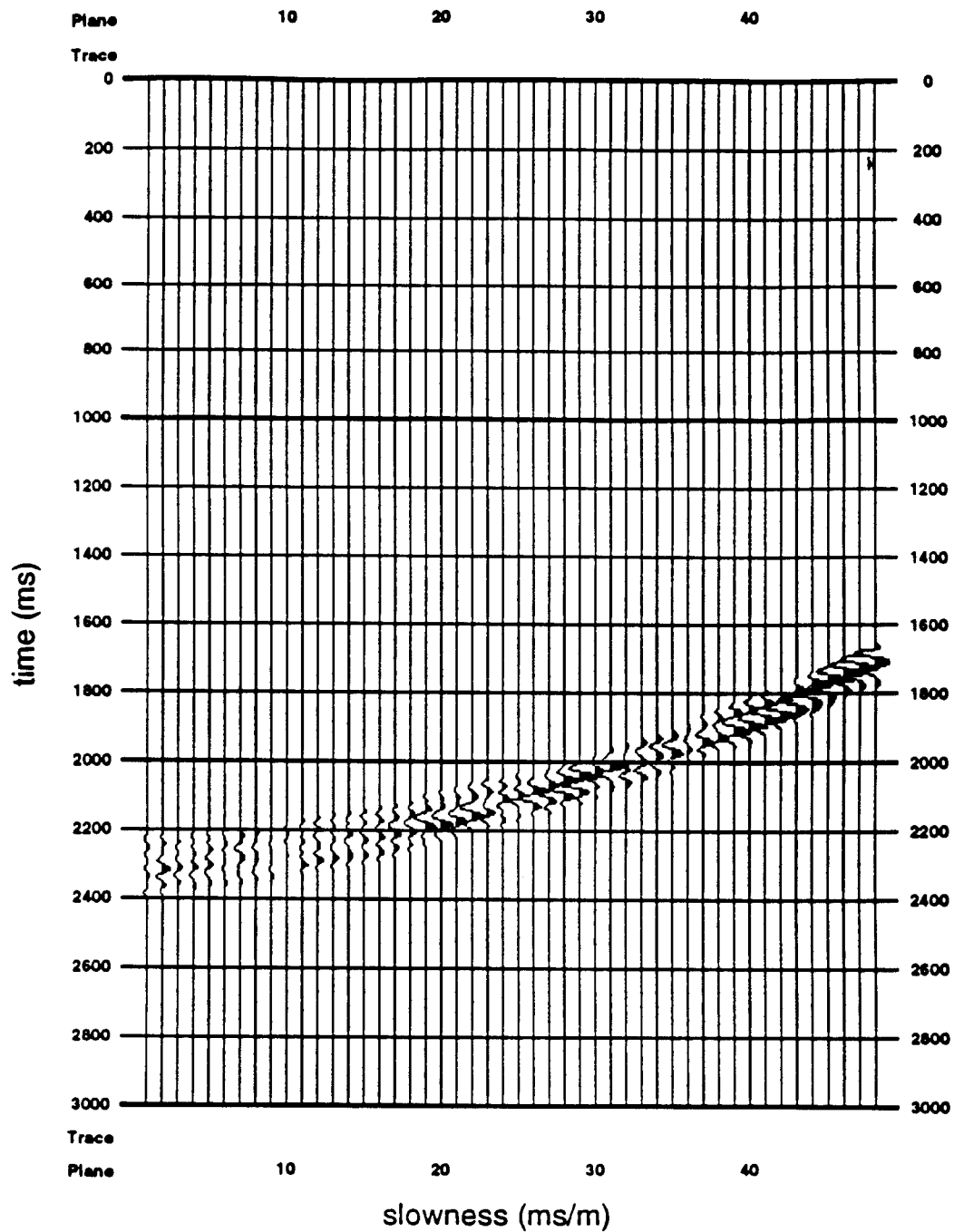
Figure 5.10   Table 5.1 Experiment 3 muted data misfit (misfit in range from 2200-2400ms).
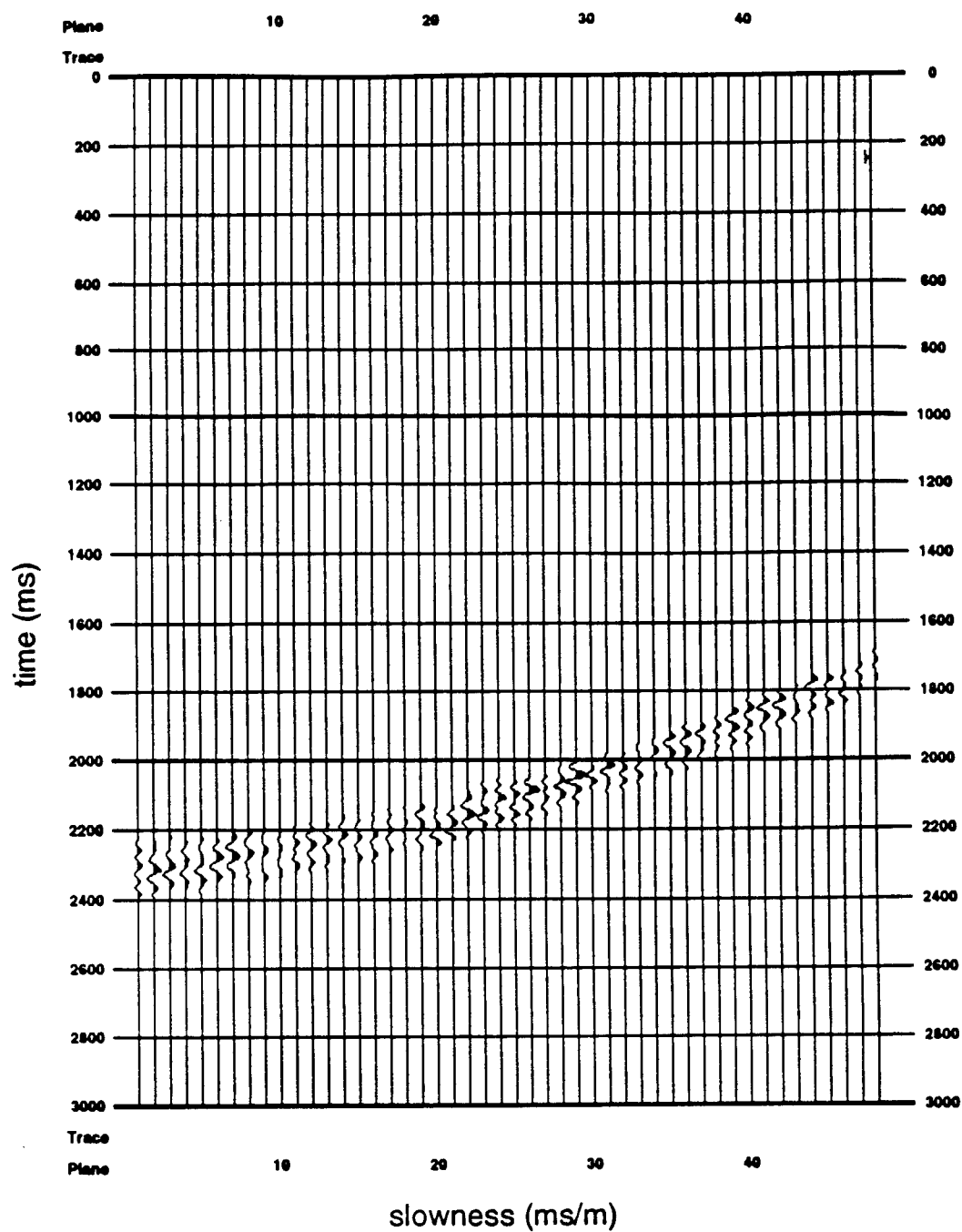
Figure 5.11 Table 5.1 Experiment 4 muted data
misfit (misfit in range from 2200-2400ms).
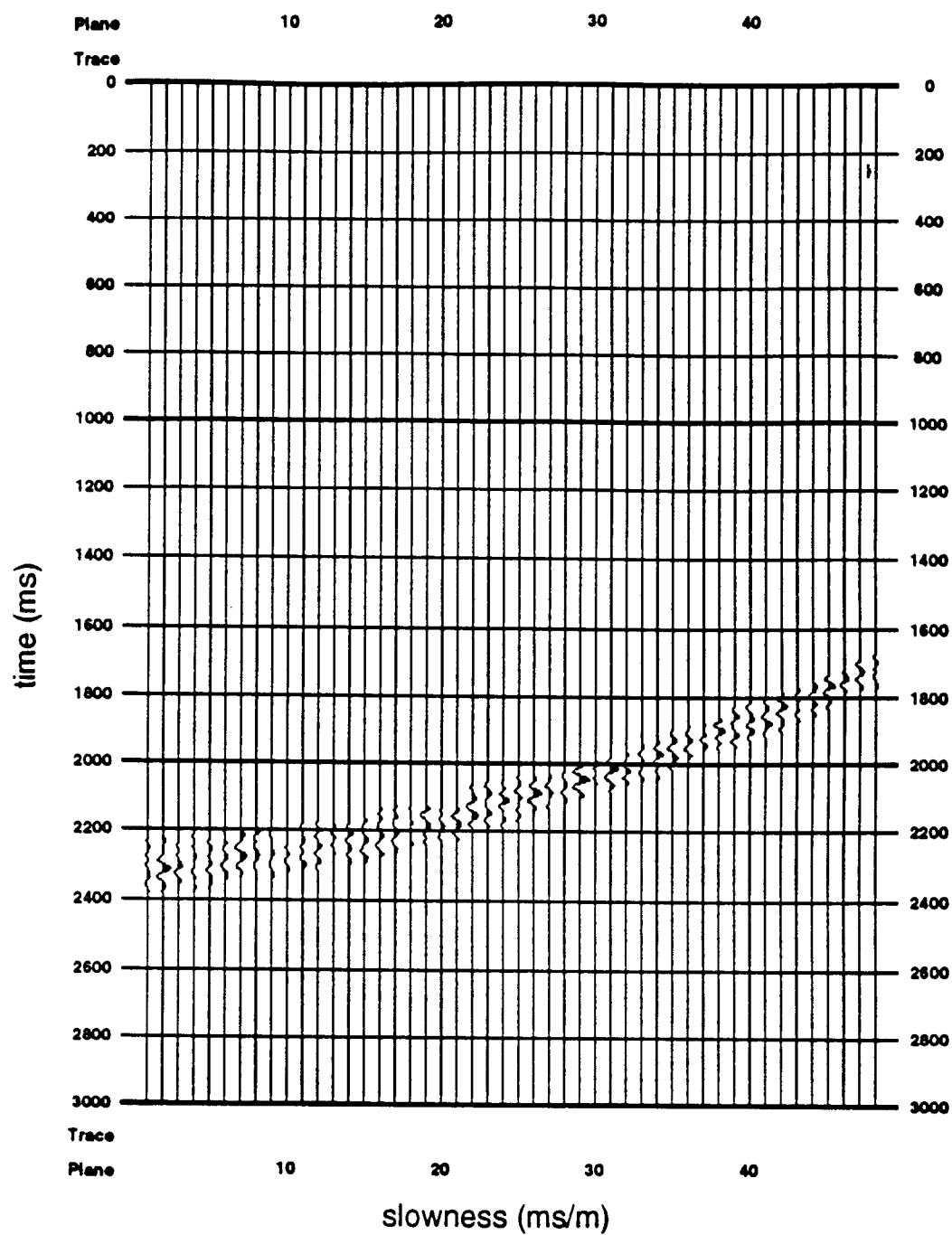
Figure 5.12 Table 5.1 Experiment 5 muted data misfit (misfit in range from 2200-2400ms).
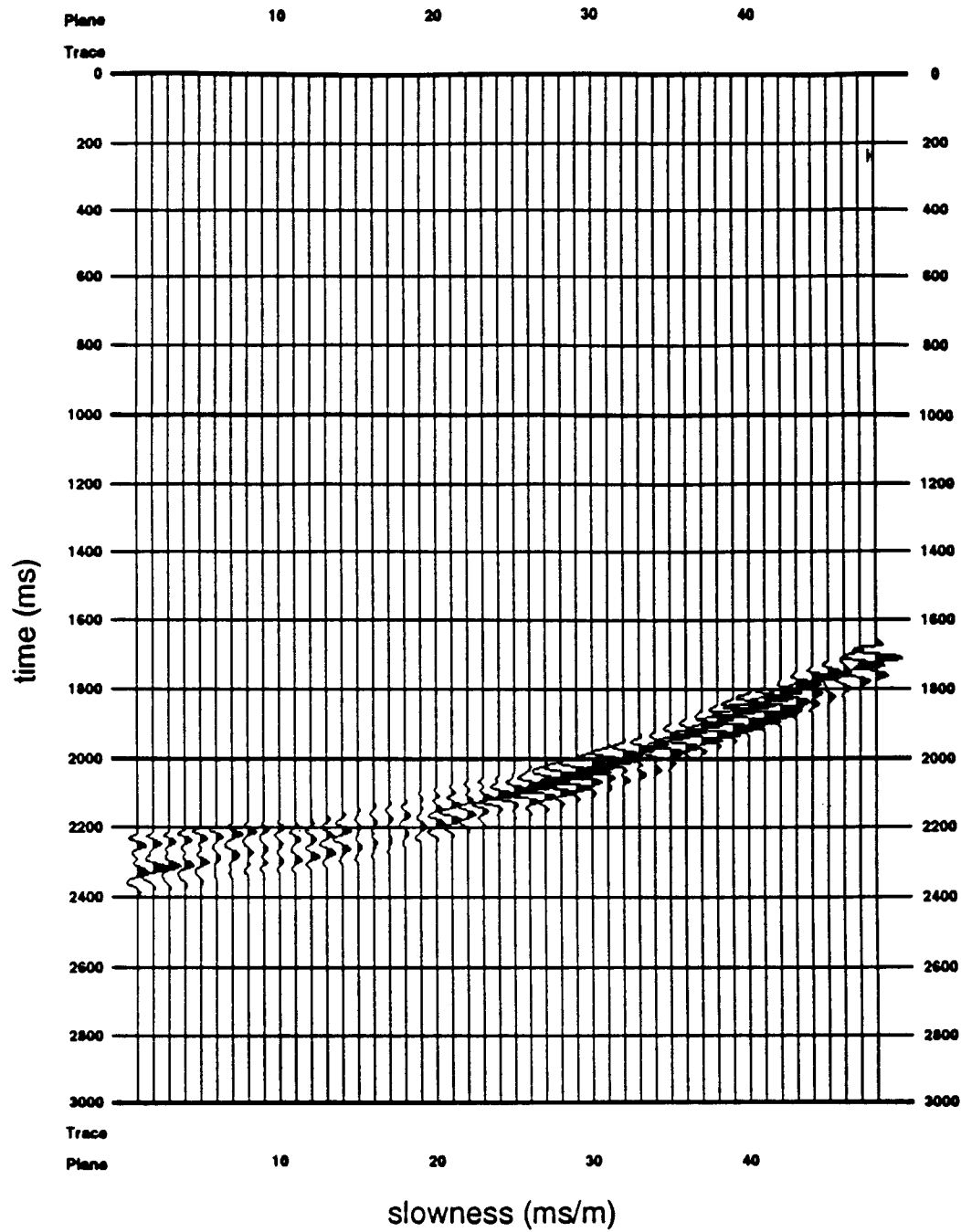
Figure 5.13   Table 5.1 Experiment 6 muted data
misfit (misfit in range from 2200-2400ms).
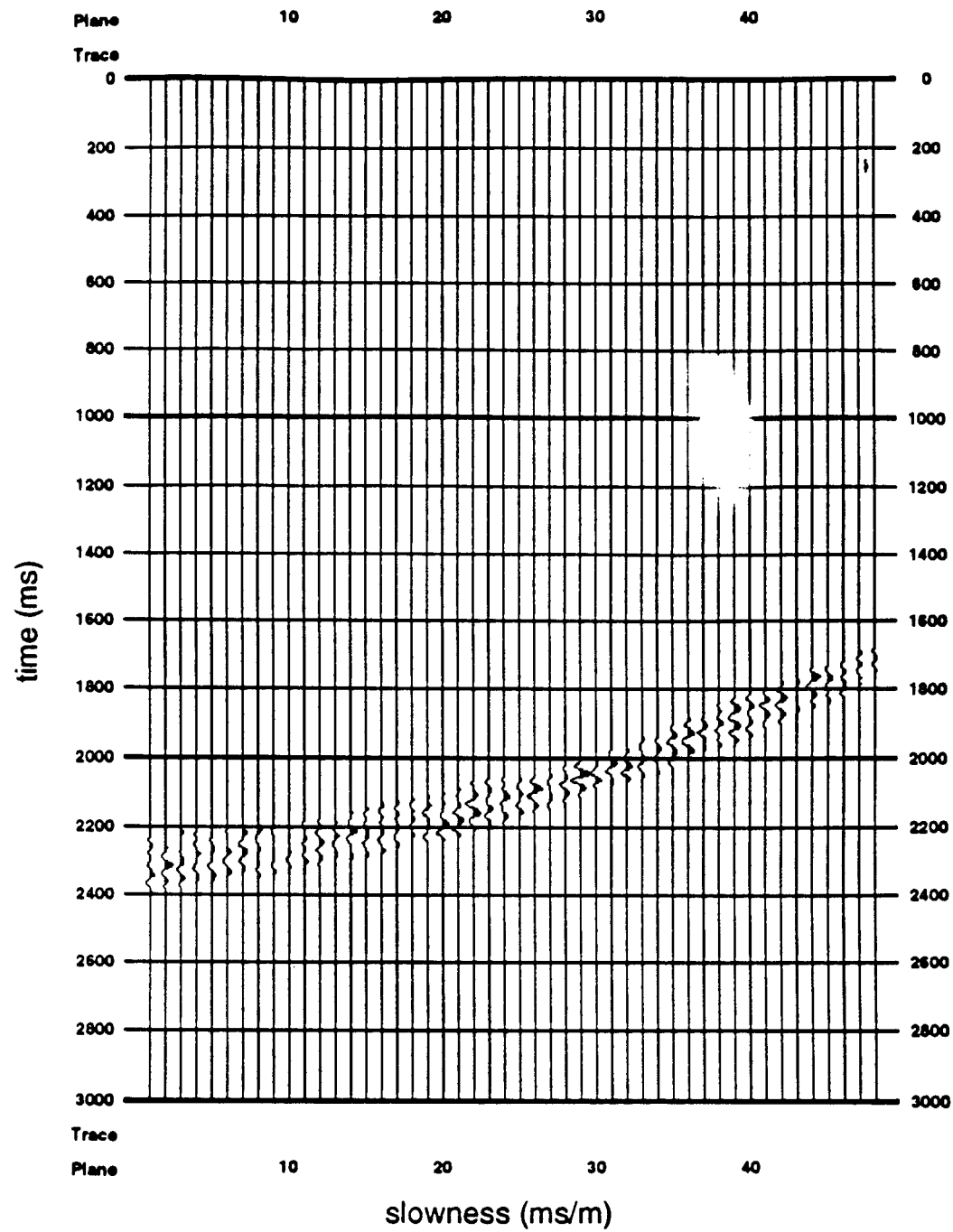
Figur 5.14 Table 5.1 Experiment 7 muted data
r t (misfit in range from 2200-2400ms).
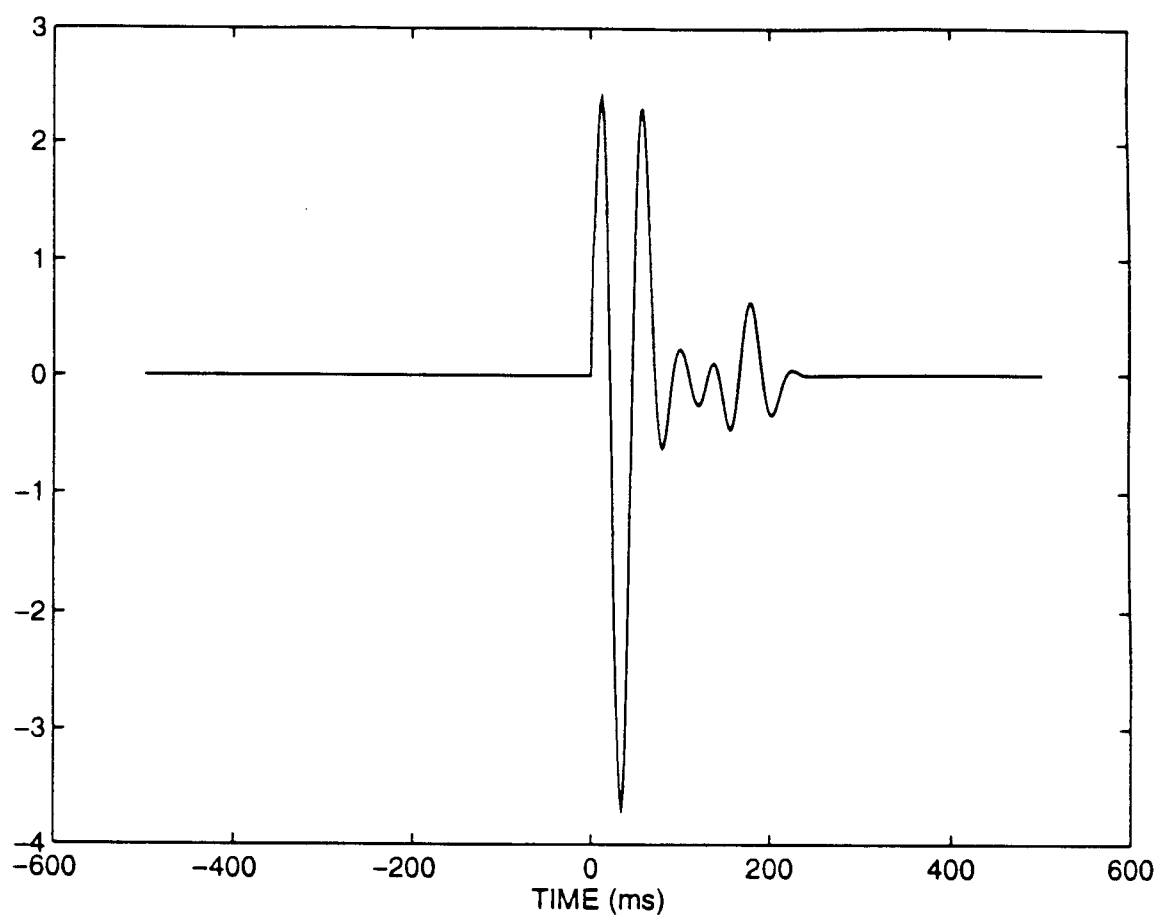
Figure 5.15 Source estimated using predictive deconvolution.
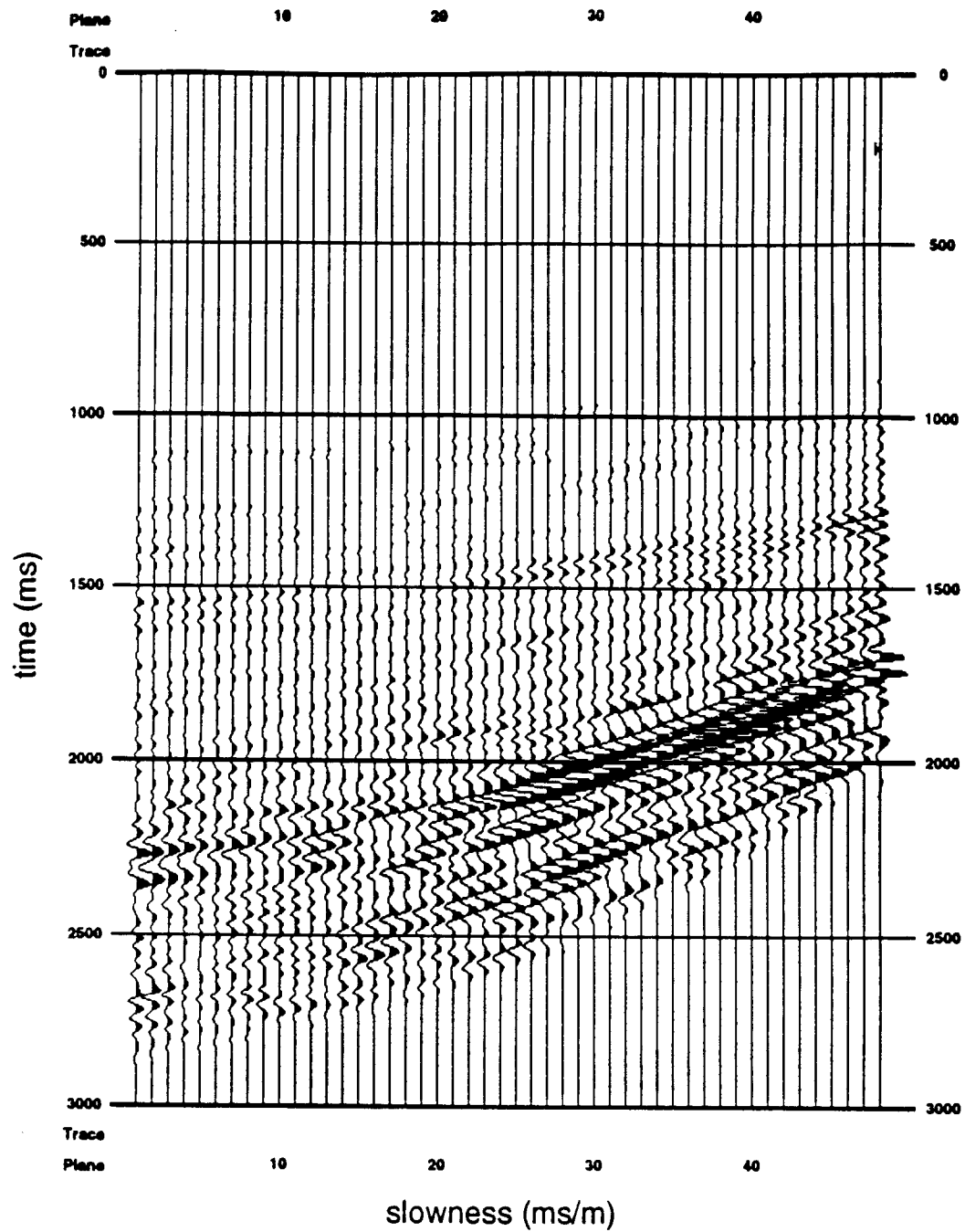
Figure 5.16   Data misfit for isotropic predictive deconvolution source and associated inversion-estimated reflectivities.
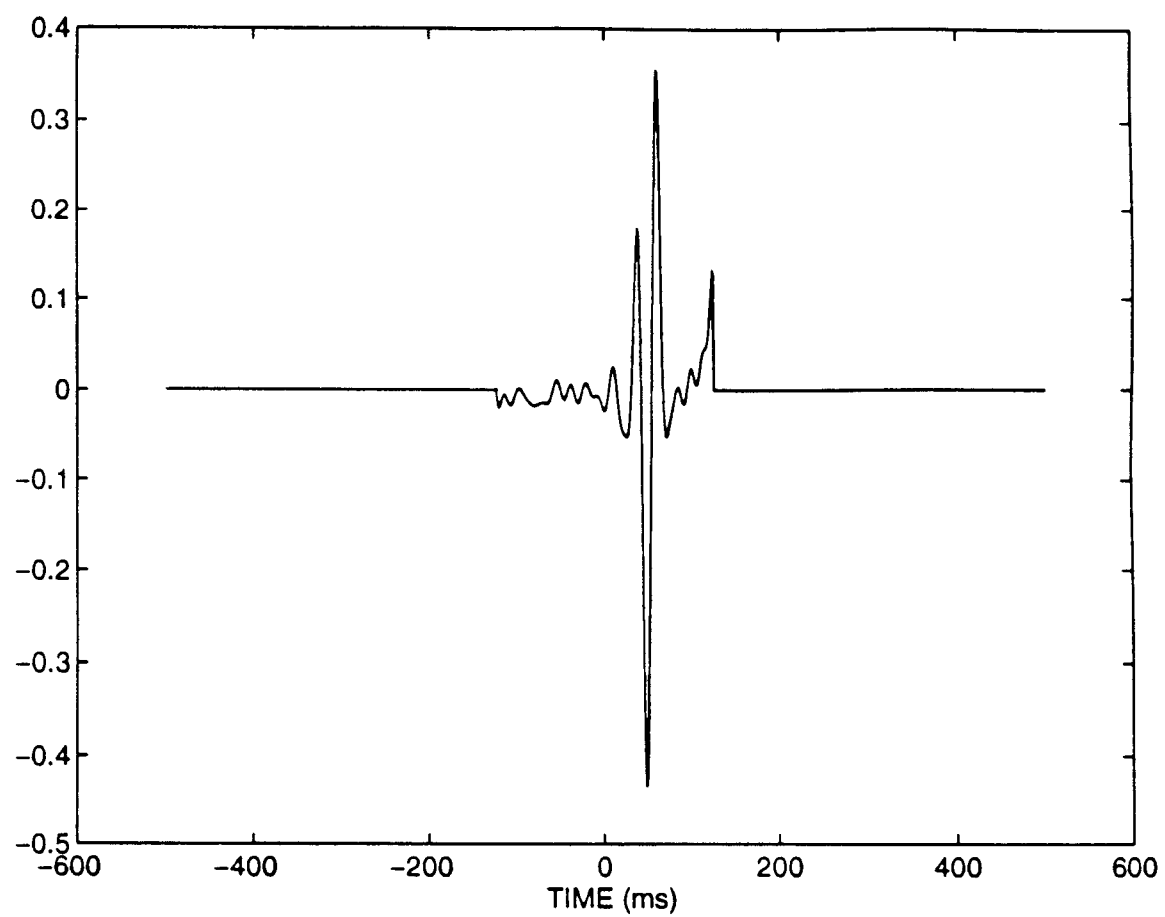
Figure 5.17 Isotropic source estimated from inversion.
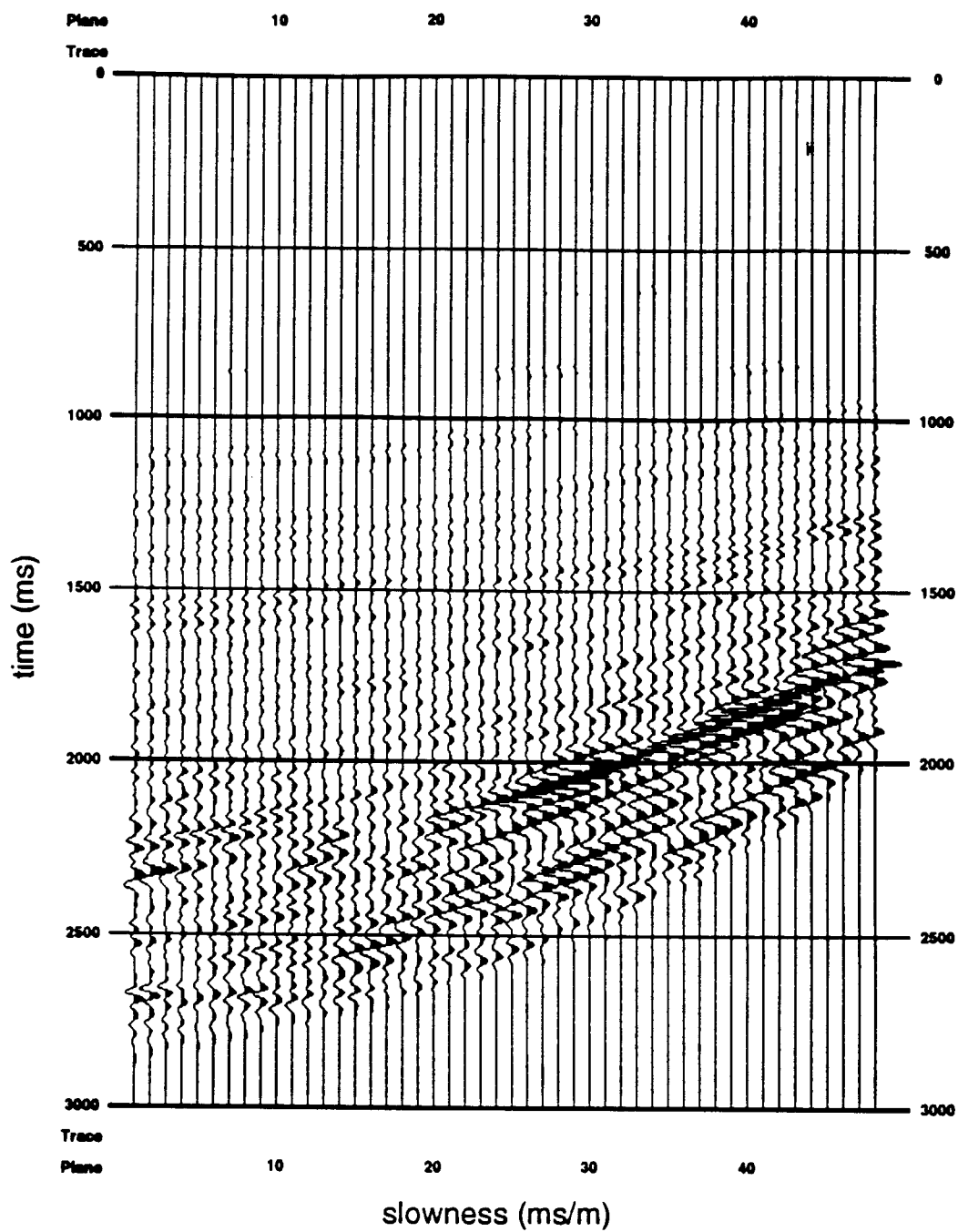
Figure 5.18 Data misfit for isotropic inversion-estimated source and associated inversion-estimated reflectivities.
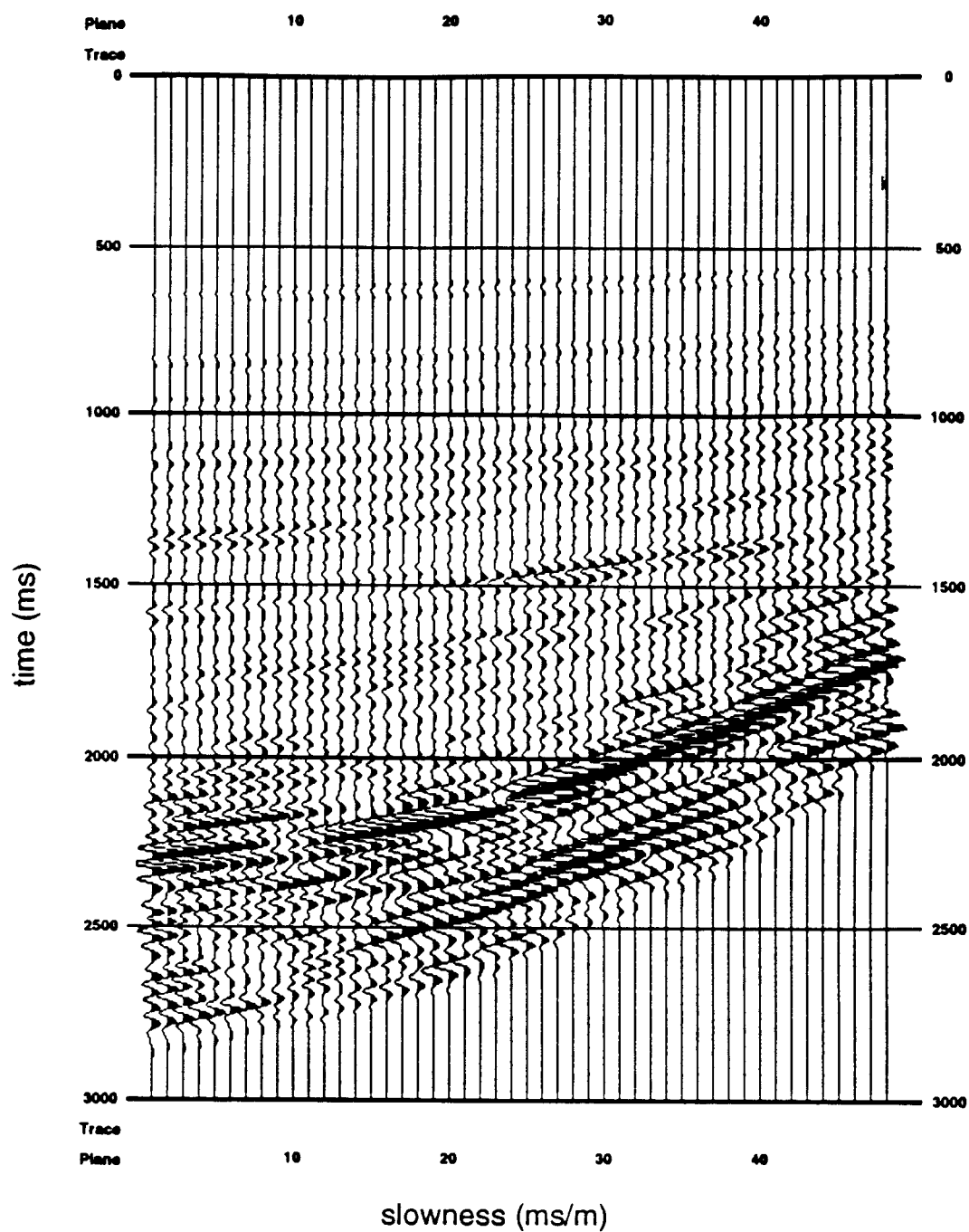
Figure 5.19 The data difference between common midpoint gathers 1 and 6 (separated by 187m). The data difference is plotted on the same scale as the CMP6 data.

Plane   1

Trace   10

-100   -100

time (ms)

0   0

100   100

Trace   10

Plane   1

slowness (ms/m)
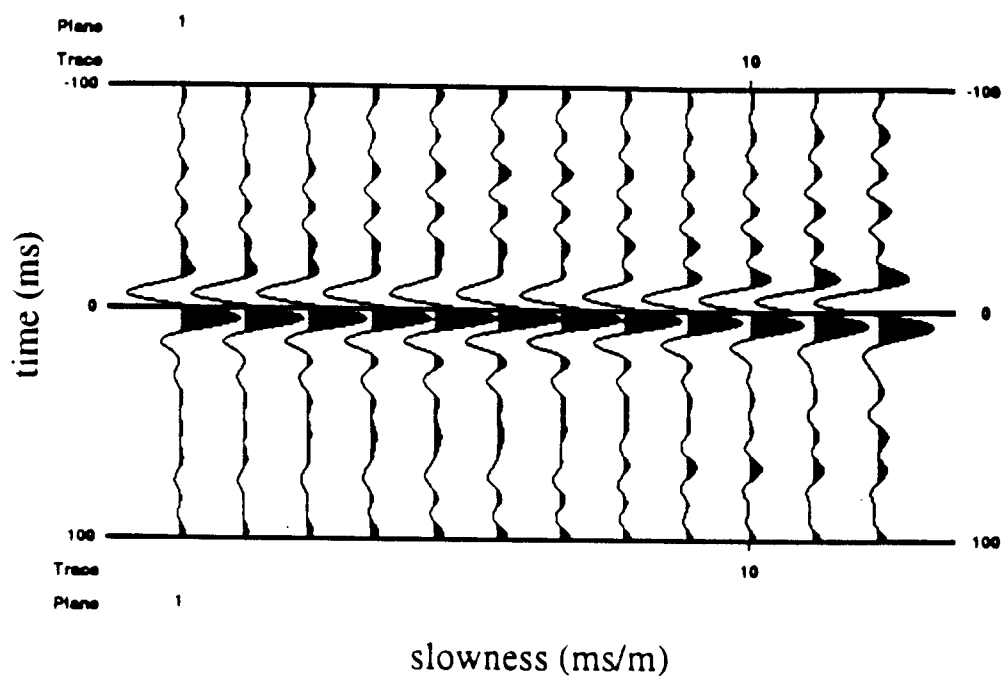
Figure 5.20   CMP1p anisotropic inversion-estimated source.
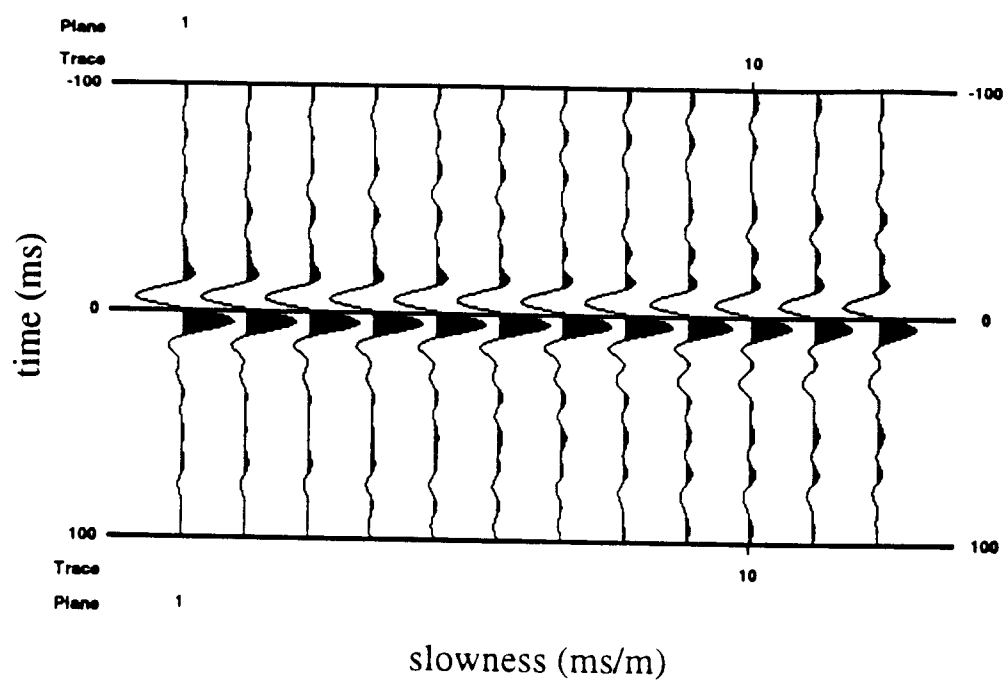
Figure 5.21 CMP6p anisotropic inversion-estimated source.
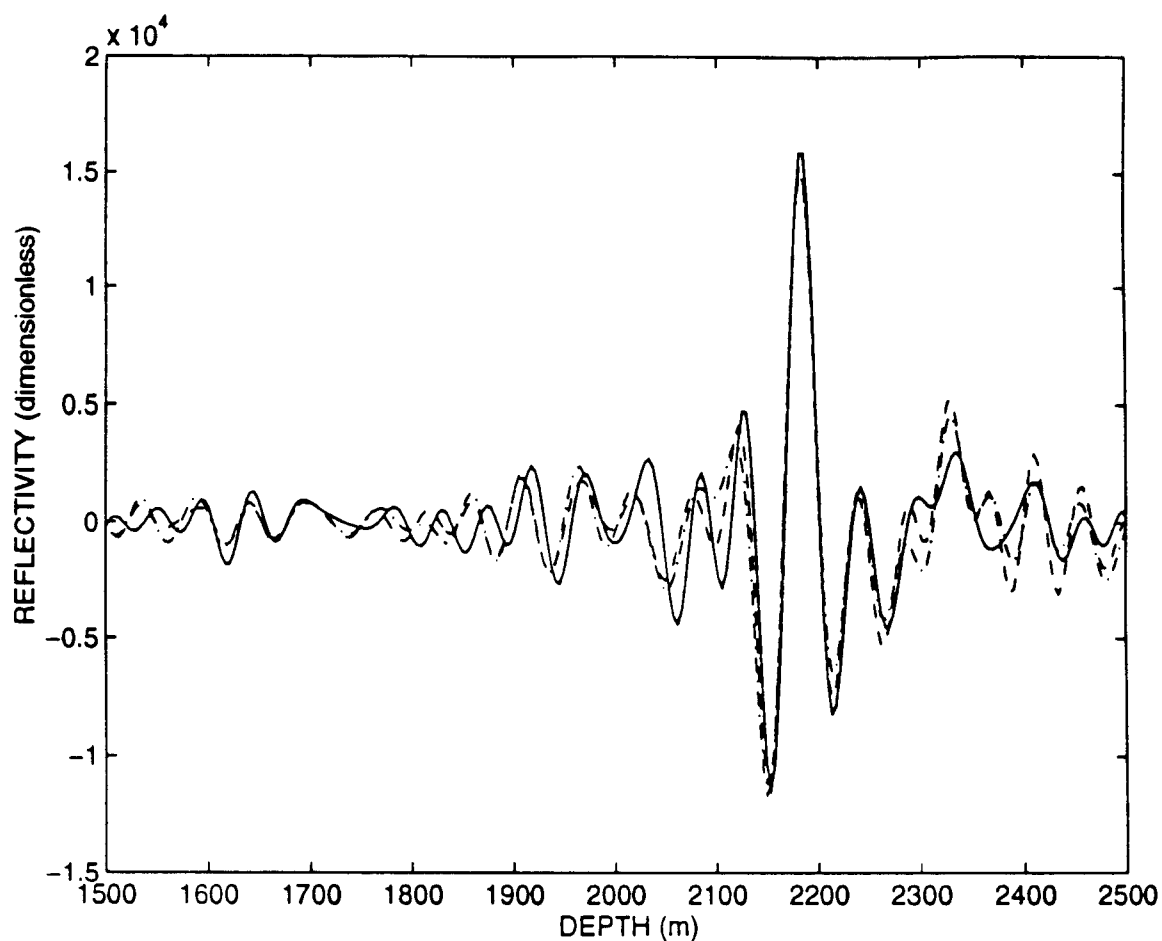
Figure 5.22 Comparison of three estimates for the short-scale relative fluctuation in the P-wave impedance (reflectivity). The solid line comes from inverting CMP6 data for the source and reflectivities. The dashed line comes from inverting CMP1 data for the source and reflectivities. The dot/dashed line comes from inverting CMP1 data for the reflectivities only. The source is fixed at the CMP6 estimate.
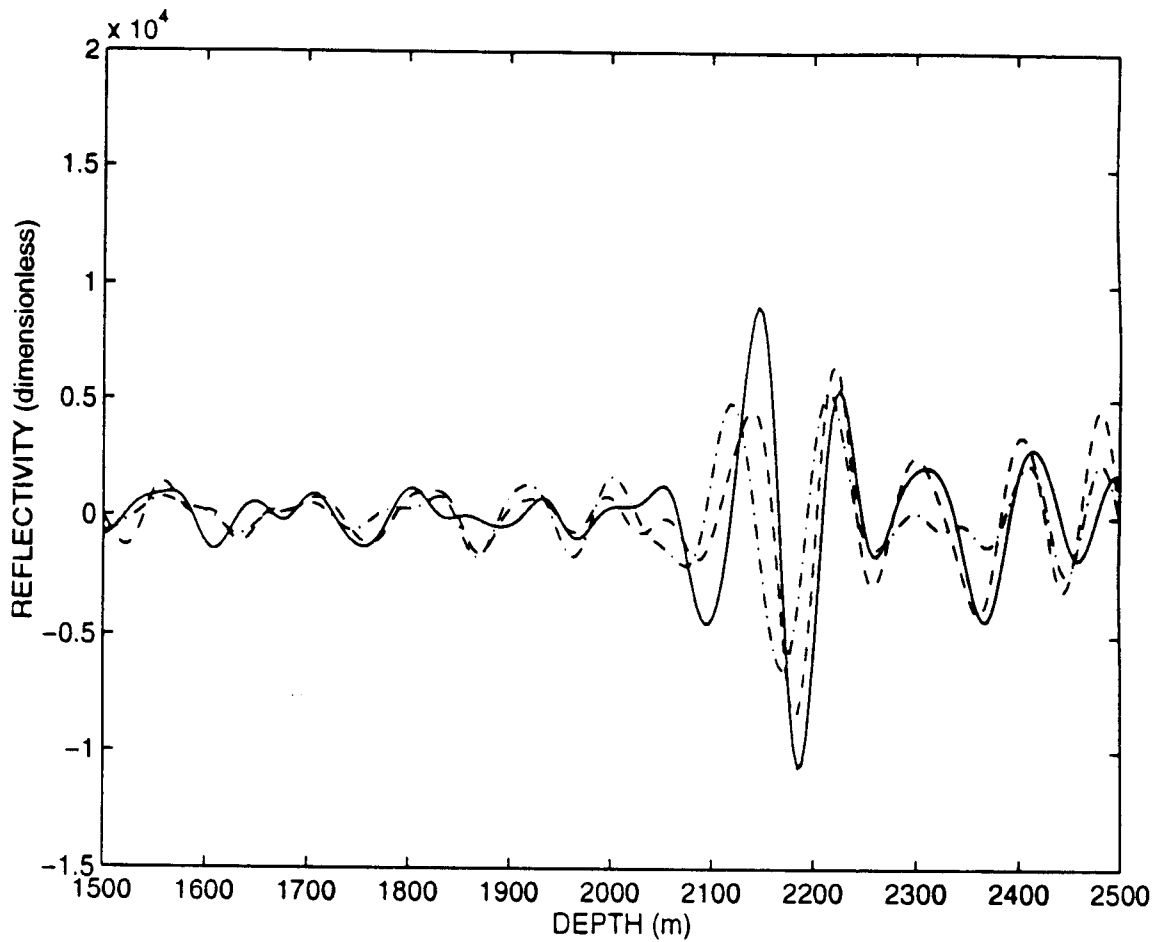
Figure 5.23 Comparison of three estimates for the short-scale relative fluctuation in the S-wave velocity (reflectivity). The solid line comes from inverting CMP6 data for the source and reflectivities. The dashed line comes from inverting CMP1 data for the source and reflectivities. The dot/dashed line comes from inverting CMP1 data for the reflectivities only. The source is fixed at the CMP6 estimate.
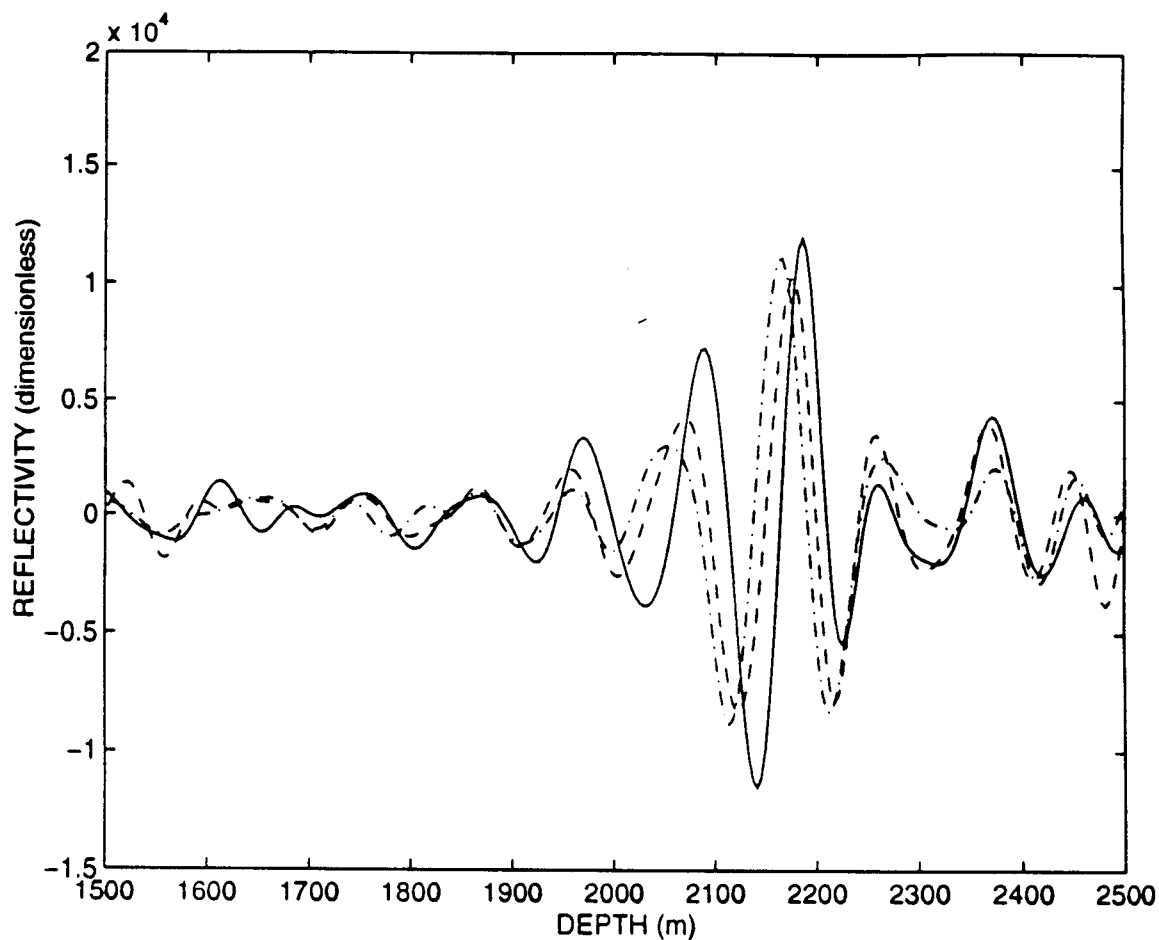
Figure 5.24 Comparison of three estimates for the short-scale relative
fluctuation in the P-wave velocity divided by density (reflectivity). The solid
line comes from inverting CMP6 data for the source and reflectivities. The
dashed line comes from inverting CMP1 data for the source and reflectivities.
The dot/dashed line comes from inverting CMP1 data for the reflectivities
only. The source is fixed at the CMP6 estimate.
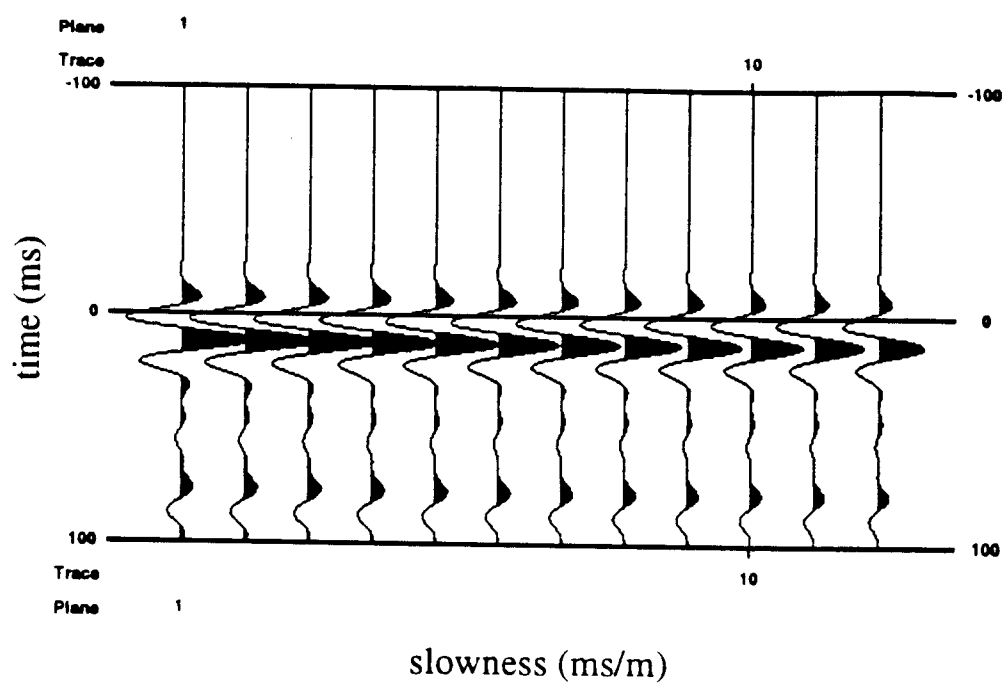
Figure 5.25 Air gun model anisotropic source estimate.
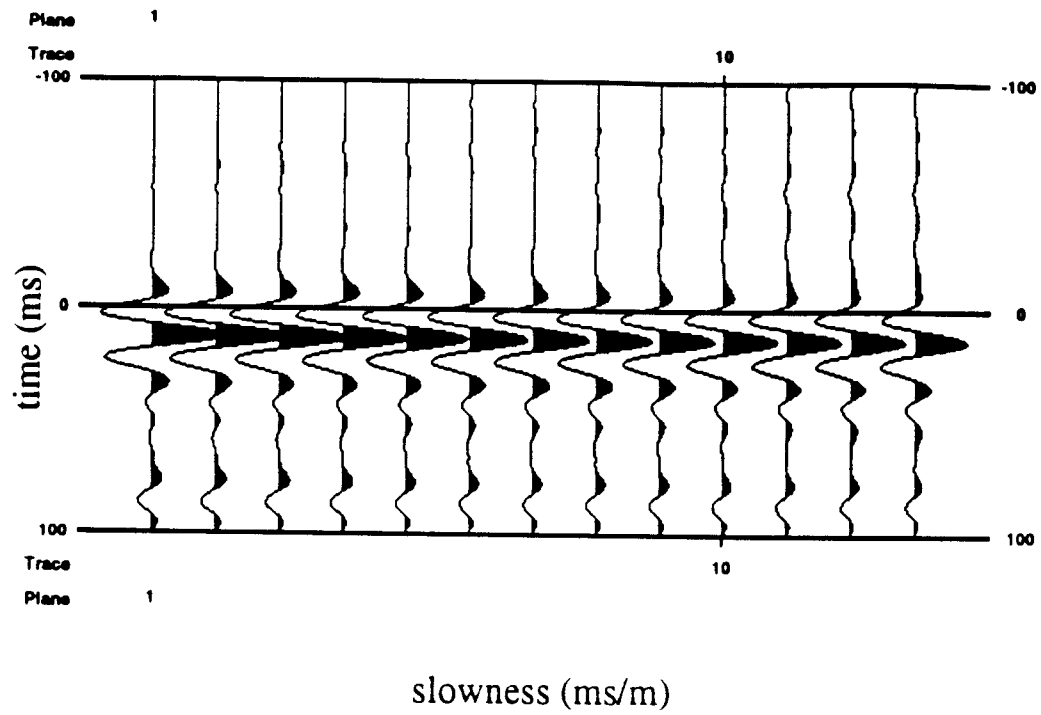
Figure 5.26 Final source estimate from source-reflectivity inversion experiment with air gun model source as initial guess.
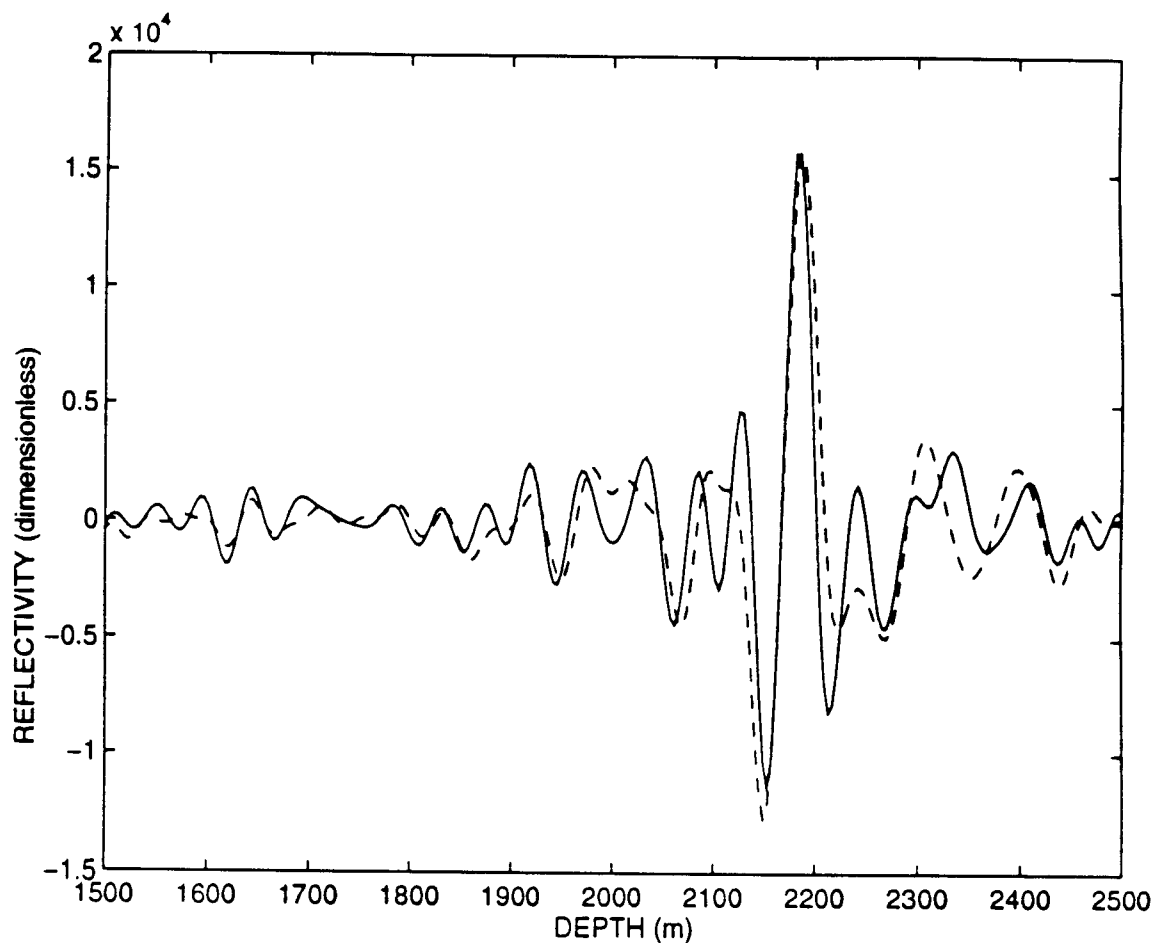
Figure 5.27   Comparison of short-scale fluctuation in P-wave impedance for two source-reflectivity experiments which are identical except for the starting source guess. The peak of the dashed graph (initial source guess was the anisotropic air gun model) was scaled to match the peak amplitude of the solid graph (initial source an isotropic Ricker wavelet). The dashed graph was also shifted 100m to the left.
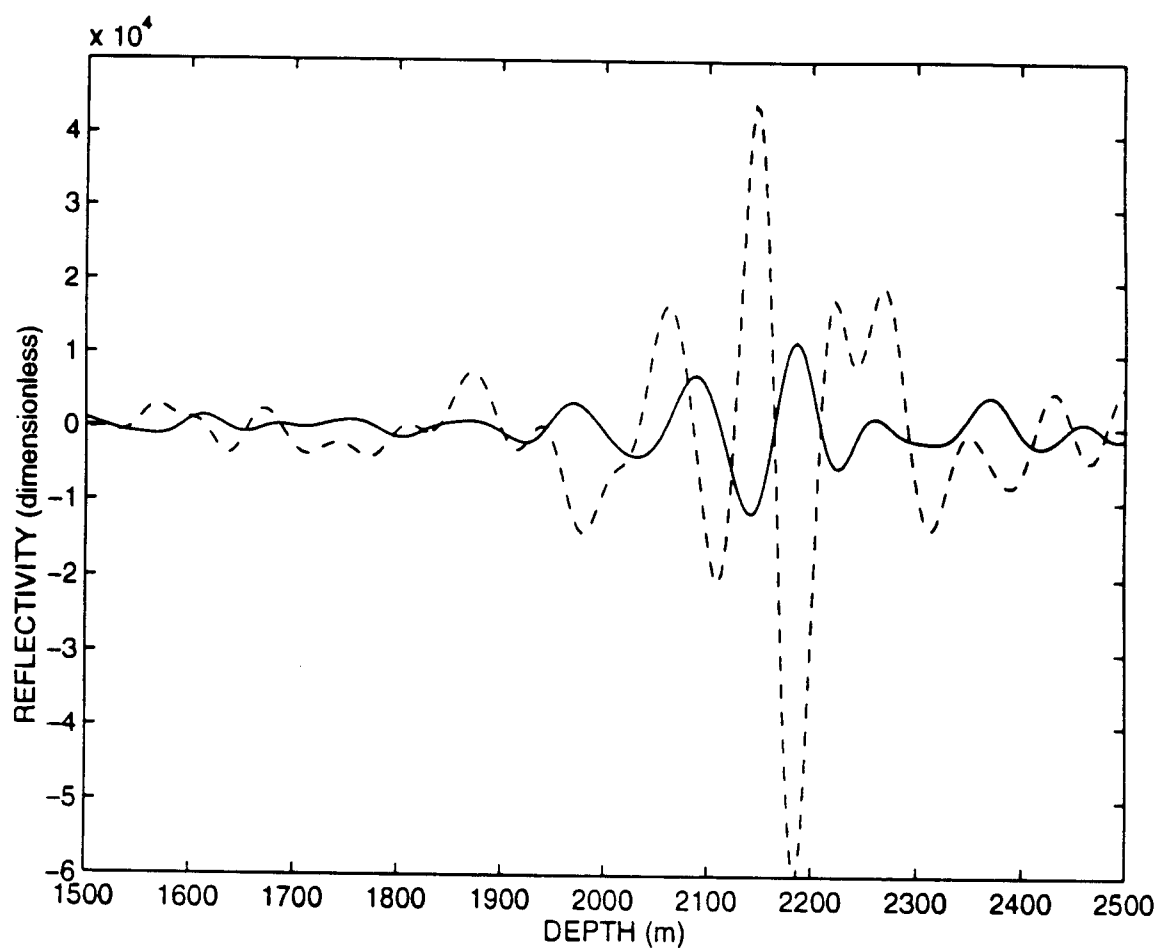
Figure 5.28 Comparison of short-scale fluctuation in P-wave velocity divided by density for two source-reflectivity experiments which are identical except for the starting source guess. Note that the solid graph was shown to be a good estimate relative to the well logs in Chapter 4.

Chapter 6

An Application of Model Resolution to Source-Reflectivity Inversions

6.1 Introduction

It is well known that seismic inverse problems have non-unique solutions. When we solve a linear inverse problem, we cannot rely on the model estimates to be accurate at each point. Least squares inversion, for example, gives a well-determined solution in an average sense. This solution minimizes the distance from the data to the column space of the forward operator.

The goal of this chapter is to quantify which parts of the model parameters are well determined insofar as one can regard that part of the model estimate as being very close to the true model. We first discuss the concept of model resolution for the general linear inverse problem and then describe the tools which we used to estimate this resolution. An explanation of the general linear inverse problem and model resolution may be found in Menke's book [36] and in the paper by Backus and Gilbert [4]. For another complete description of the problem and concepts of model resolution see [62] where these ideas are applied to surface waves and free oscillation observations. The generalized inverse and model resolution matrix for this least squares problem can be defined in terms of the singular value decomposition. Instead we estimate the model solution and eigenvalues and eigenvectors of the normal matrix through the conjugate gradient and Lanczos algorithms. These estimated eigenvectors of the normal matrix give us a practical (although clearly approximate) way to compute the model resolution matrix.

6.2 The Model Resolution Idea

We will now describe the normal equations for general data $d \in H^r$ and model $m \in H^s$ where H^r and H^s are the Sobolev spaces of order r and s respectively. We start by defining the Sobolev space H^s (see [40] Ch.2). Let $\lambda^s(\xi) = (1 + |\xi|^2)^{s/2}$ for $s \in \Re$ and

$\xi \in \Re^n$. Then we say that $u \in H^s$ (the *Sobolev space of exponent* s) if $u \in S'$ and $\lambda^s \hat{u} \in L^2$. In other words, $u \in H^s$ if $\hat{u}$ is a function satisfying

$$\|u\|_s^2 = (2\pi)^{-n} \int (1 + |\xi|^2)^s |\hat{u}(\xi)|^2 d\xi < \infty. \qquad ,$$

$\| \cdot \|_s$ is the sth Sobolev norm, and $\langle \cdot, \cdot \rangle_s$ its associated inner product. We note that $H^0 \equiv L^2$. Here S' is the space of *temperate distributions*. The distribution $u \in S'$ if u is a semi-linear form $\phi \mapsto (u, \phi)$ on S with two constants $C \in \Re$ and $N \in Z_+$ such that

$$|(u, \phi)| \le C |\phi|_N, \qquad \text{for } \phi \in \text{ }$$

Here, the *Schwartz Space* S is the space of C^∞ functions that are rapidly decreasing at infinity. The norm on the S is defined by

$$|\phi|_k = \sup\{|x^\alpha \partial^\beta \phi(x)|; x \in \Re^n \text{ and } |\alpha + \beta| \le k\}.$$

Our forward operator $G : H^s \mapsto H^r$. It is logical to assume throughout that $s \ge r \ge 0$. We generally asume that given data we can find a model such that $Gm = d$. Suppose, however, that we cannot solve this equation exactly. Then one approximate solution (the least squares solution) requires that we seek model parameters m such that we have minimized the objective function

$$J_{OLS} = \|Gm - d\|_r^2 = \langle Gm - d, Gm - d \rangle_r. \qquad (6.1)$$

To find the normal equations for these general data and model spaces we require that at the minimizer, m, the gradient of J_{OLS} vanish. Starting from the variational formulation,

$$\langle G(m + \delta m) - d, G(m + \delta m) - d \rangle_r - \langle Gm - d, Gm - d \rangle_r$$

and dropping terms in the perturbed quantity δm of order greater than 1, we get that in order that our model minimize the misfit,

$$\langle Gm, G\delta m \rangle_r = \langle G\delta m, d \rangle_r$$

$$\langle \Lambda^r Gm, \Lambda^r G\delta m \rangle = \langle \Lambda^r G\delta m, \Lambda^r d \rangle$$

$$\langle (\Lambda^r G)^\dagger \Lambda^r Gm, \delta m \rangle_s = \langle (\Lambda^r G)^\dagger \Lambda^r d, \delta m \rangle_s$$

$$\langle \Lambda^s (\Lambda^r G)^\dagger \Lambda^r Gm, \Lambda^s \delta m \rangle = \langle \Lambda^s (\Lambda^r G)^\dagger \Lambda^r d, \Lambda^s \delta m \rangle$$

$$\Lambda^s(\Lambda^r G)^\dagger \Lambda^r Gm = \Lambda^s(\Lambda^r G)^\dagger \Lambda^r d \qquad (6.2)$$

Where $(\Lambda^r G)^\dagger : L^2 \mapsto H^s$. We now need to determine a way to write $(\Lambda^r G)^\dagger$ in terms of its L^2 adjoint (the transpose). Let $u \in L^2$ and $v \in H^s$. Again, the operator $G : H^s \mapsto H^r$. Also, recall that we assumed at the start that $s \geq r \geq 0$. Thus, $H^s \subset L^2$.

$$\langle u, \Lambda^r Gv \rangle = \langle (\Lambda^r G)^\dagger u, v \rangle,$$

$$\langle (\Lambda^r G)^t u, v \rangle = \langle \Lambda^{2s}(\Lambda^r G)^\dagger u, v \rangle$$

$$(\Lambda^r G)^t = \Lambda^{2s}(\Lambda^r G)^\dagger$$

Or

$$\Lambda^{-2s}(\Lambda^r G)^t = (\Lambda^r G)^\dagger$$

Equation 6.2, therefore, becomes

$$\Lambda^{-s} G^t \Lambda^{2r} Gm = \Lambda^{-s} G^t \Lambda^{2r} d. \qquad (6.3)$$

$$m_{est} = [\Lambda^{-s} G^t \Lambda^{2r} G]^{-1} \Lambda^{-s} G^t \Lambda^{2r} d \equiv G^{-g} d. \qquad (6.4)$$

where $G^{-g}d$ is the *least squares generalized inverse*.

Most geophysical inverse problems are of "mixed determined" type; i.e., these problems are ones in which some model parameters are well determined and others are not. When we perform an inversion, we should not think of the resulting model parameters as being accurate at each discretized point. It is more reasonable to think of the solution vector m as an *average* of the true model parameters $\langle m \rangle = a^t m$. The easiest such models to interpret are ones which have an averaging vector a which is only nonzero over a small interval in the domain of the model parameters. Then, if the model parameters are discretely parametrized in a physically meaningful way, one can interpret this average. In the experiments I have described throughout the previous chapters, the model parameters depend on one spatial coordinate only, depth z, and the averaging vectors correspond to averages over depths.

We wish to quantify our ability to resolve the individual model parameters. In other words, assuming that there is an estimate of the model parameters m^{true} such that $Gm^{true} = d^{obs}$, we would like to be able to decide how close our model estimate m^{est} is to these (unknown) true model parameters. Then

$$m^{est} = G^{-g}d^{obs} = G^{-g}Gm^{true} \equiv Rm^{true} \tag{6.5}$$

Where $R \in \Re^{n \times n}$ is the *model resolution matrix*. If $R = I$ then $m^{true} = m^{est}$ and the model parameters are perfectly resolved in the inversion. In general, $R \neq I$, and then the model estimates m^{est} are weighted averages of the true parameters $m^{est} = a^t m^{true}$.

One way to compute the model resolution matrix is to use the singular value decomposition. The singular value decomposition is defined in [20], page 71, by the following theorem:

Theorem 6.1

If $A \in \Re^{m \times n}$ matrix, then there exist orthogonal matrices $U = [u_1, \ldots, u_m] \in \Re^{m \times m}$ and $V = [v_1, \ldots, v_n] \in \Re^{n \times n}$ such that $U^t AV = diag(\sigma_1, \ldots, \sigma_p) \in \Re^{m \times n}$, $p = \min\{m, n\}$, where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0$.

The columns of U are eigenvectors of AA^t (the left singular vectors), and the columns of V are the eigenvectors of $A^t A$ (the right singular vectors). Finally, the p singular values on the diagonal of $\Sigma = diag(\sigma_1, \ldots, \sigma_p)$ are the square roots of the nonzero eigenvalues of AA^t and $A^t A$ (the singular values).

In this part of the discussion we assume for simplicity that the forward operator $G : L^2 \mapsto L^2$. This assumption makes the representation of the resolution matrix in terms of the SVD simpler. *In our implementation, we do not explicitly compute the SVD.* Instead we calculate the eigenvalues and eigenvectors of the normal operator through the Lanczos process directly (see Section 6.4). When the Lanczos process is used to calculate these eigenvectors, the forward operator has been transformed for computational purposes so that $G : L^2 \mapsto L^2$ rather than as originally defined $G : H^s \mapsto H^r$.

In our example, the forward operator G, can therefore be described in terms of the singular value decomposition by $G = U\Sigma V^t$. In other words, the matrix U has columns which are the eigenvectors spanning the data space D, and V's columns are the eigenvectors spanning the model space M. The singular value matrix Σ can be partitioned into a square submatrix Σ_p of nonzero singular values and three zero matrices:

$$\begin{pmatrix} \Sigma_p & 0 \\ 0 & 0 \end{pmatrix}$$

We write $G = U\Sigma V^t = U_p \Sigma_p V_p^t$ (the truncated SVD) where U_p and V_p^t consist of the first p columns of U and V respectively. We note that although $V_p^t V_p = I$, in general since the p right singular vectors do not span the whole space, $V_p V_p^t \neq I$.

The least squares generalized inverse G^{-g} may now be expressed using the singular value decomposition as

$$G^{-g} = V_p \Sigma_p^{-1} U_p^t$$

and the model resolution matrix (from expression 6.5) is simply

$$R = G^{-g}G = (V_p \Sigma_p^{-1} U_p^t)(U_p \Sigma_p V_p^t) = V_p V_p^t \tag{6.6}$$

6.3 Computation of the Resolution Matrix Using the Conjugate Gradient and Lanczos Algorithms

One aspect of the Differential Semblance Optimization approach is that the model parameters sought in the inversion are grouped according to the influence they have on the data. The elements which have a nonlinear effect on the data are grouped separately from those elements which have a linear influence on the data. For example, at each step in velocity model space (the most common nonlinear parameter) an inversion must be carried out to estimate the parameters which linearly influence the data (such as the short-scale relative fluctuations in the elastic parameters or reflectivities). This inversion involves computing and solving the normal equations. (For a discussion of the algorithmic construction of DSO see [30])

Due to the large size of the normal matrices which are typical for seismic inverse problems, we prefer to use iterative methods to solve these linear systems. Since the normal operator is symmetric and positive definite, an obvious technique to use to solve this system is the conjugate gradient algorithm. Moreover, the conjugate gradient algorithm automatically generates the parameters needed to construct the tridiagonal matrix which arises in the Lanczos iteration. The extreme eigenvalues of this tridiagonal matrix approximate those of the original matrix (in our case, the normal matrix). The corresponding eigenvectors of the normal matrix can also be found from the eigenvectors of the tridiagonal matrix by multiplying by the Lanczos matrix (changing bases). The eigenvalues provide information about the condition of the normal matrix. The eigenvectors of the normal matrix (right singular vectors of the forward operator) allow us to measure model resolution.

The next section describes the conjugate gradient and Lanczos algorithms in general terms. The classic references for these two algorithms are [31], [24], and [23]. The paper [21] gives a bibliographical history of the two methods. The following general description of the two algorithms may be found in expanded form in [20].

6.4 Discussion of the Conjugate Gradient and Lanczos Algorithms

6.4.1 The Conjugate Gradient Idea

The Hestenes-Stiefel conjugate gradient algorithm may be understood in the context of minimizing the function $\phi(x)$ defined by

$$\phi(x) = \frac{1}{2}x^t A x - x^t b \tag{6.7}$$

where $b \in \Re^n$, and $A \in \Re^{n \times n}$ is assumed to be positive definite and symmetric. The minimizer of ϕ is $x = A^{-1}b$. So, minimizing the function ϕ and solving the linear system $Ax = b$ are seen to be equivalent problems.

One obvious choice for decreasing the function ϕ is to travel in the negative gradient direction $-\nabla\phi(x_c) = b - Ax_c$ from the current point x_c. One notices that the negative gradient direction is the residual direction r_c of the system at the current point. Unfortunately, as is well known, this method (steepest descent) may converge extremely slowly if the condition of the system (or ratio of largest to smallest eigenvalues) is large. The conjugate gradient algorithm, therefore, chooses to minimize ϕ in a set of directions $\{p_1, p_2, ...\}$ which do not necessarily correspond to the residual directions. One approach with obvious benefits is to choose linearly independent directions p_i so that each x_k solves

$$\min_{x \in \text{span}\{p_1, \dots, p_k\}} \phi(x) \tag{6.8}$$

This choice of search directions ensures finite termination of the algorithm in at most n steps. We would like a vector p_k such that when we solve the one-dimensional minimization problem

$$\min_\alpha \phi(x_{k-1} + \alpha p_k) \tag{6.9}$$

we also solve the k-dimensional problem 6.8. Luckily, such a solution is possible if we require the directions p_k to be A-*conjugate* to the previous directions $p_1, ..., p_{k-1}$. The vectors $p_1, ..., p_k$ are A-conjugate if $P_{k-1}^t A p_k = 0$. These requirements can be satisfied and an algorithmic implementation is described in Subsection 6.4.3 below.

6.4.2 The Lanczos Idea and Connection to the Conjugate Gradient Algorithm

Estimates of the eigenvalues and eigenvectors of the normal operator could be very useful for analyzing the inversion results we obtain from DSO. The Lanczos algorithm when applied to a symmetric matrix $A \in \Re^{n \times n}$, generates a sequence of tridiagonal matrices $T_j \in \Re^{j \times j}$ with extreme eigenvalues which are progressively better estimates of the extreme eigenvalues of A.

One way to motivate the Lanczos idea is to recall the Rayleigh quotient which can be used to approximate the eigenvalues of a matrix A. Let λ_1 be the largest eigenvalue of A and λ_n the smallest. For $Q_j = [q_1, ..., q_j]$ a matrix in $\Re^{n \times j}$ with orthonormal columns, we define the scalars M_j and m_j by

$$M_j = \max_{y \neq 0} \frac{y^t (Q_j^t A Q_j) y}{y^t y} \leq \lambda_1(A) \tag{6.10}$$

$$m_j = \min_{y \neq 0} \frac{y^t (Q_j^t A Q_j) y}{y^t y} \geq \lambda_n(A) \tag{6.11}$$

The Lanczos algorithm provides a way to compute the q_j so that the scalars M_j and m_j are better and better estimates of $\lambda_1(A)$ and $\lambda_n(A)$. Let $x = Q_j y$. Then the Rayleigh quotient changes most rapidly in the direction of its gradient which is a vector contained in $span\{x, Ax\}$. For this reason, the Lanczos vectors $\{q_i\}_1^j$ are chosen to be an orthonormal basis for the Krylov subspace

$$\kappa(A, q_1, j) \equiv span\{q_1, Aq_1, ..., A^{j-1}q_1\} = span\{q_1, ..., q_j\} \tag{6.12}$$

At the jth iteration of the Lanczos algorithm we have a matrix Q (the Lanczos matrix) whose columns are the normalized residuals resulting from the conjugate gradient algorithm (which can be shown to be orthonormal) and a symmetric, tridiagonal

matrix $T \in \Re^{j \times j}$. In fact, the Lanczos matrix "tridiagonalizes" the matrix A up to an error matrix.

$$AQ_j = Q_j T_j + r_j e_j^t. \qquad\qquad \text{(6.13)}$$

The entries in T are combinations of the parameters generated in the conjugate gradient iteration (for details see the algorithm next section).

6.4.3   Algorithm

We present here a pseudocode version of the two algorithms (conjugate gradient and Lanczos) which have been implemented in DSO.

Variables Used

A :	normal operator
b:	data
r:	residual
x_0:	starting solution
x:	approximate solution
p:	conjugate gradient direction
β:	parameter used in computation of new direction p
α_c	step length in current direction p
α_p	step length in previous direction
rtr_c:	inner product of current residual with itself
rtr_p:	inner product of previous residual with itself
tol:	relative residual tolerance used for determining algorithm convergence
Q:	Lanczos matrix
T	tridiagonal matrix resulting from Lanczos process with eigenvalues approximating those of A
Z	matrix of eigenvectors of the tridiagonal matrix T
X	matrix of approximate eigenvectors of the original matrix A

Algorithm: (Conjugate Gradient/Lanczos) If $A \in \Re^{n \times n}$ is symmetric and positive definite and $b \in \Re^n$ then the following algorithm computes $x \in \Re^n$ so that $Ax = b$. The algorithm also optionally approximates some of the eigenvalues and eigenvectors of the matrix A.

initialize:

$r = b - Ax_0$

$x = x_0$

for $k = 1$:iteration limit

    if eigenvector flag = true

        $Q(:, k) = r/\|r\|$

    end if

    if $k = 1$

        $\beta = 0$

        $p = r$

        $rtr_c = \langle r, r \rangle$

    else

        $\beta = rtr_c/rtr_p$

        $p = r + \beta p$

    end if

    $ap = Ap$

    $ptap = \langle p, ap \rangle$

    if $ptap < tol$

        break

    end if

    $\alpha_c = rtr_c/ptap$

    $x = x + \alpha_c p$

    $r = r - \alpha_c ap$

    if eigenvalue flag = true

        if $k = 1$

            $T(k, k) = 1/\alpha_c$

            $T(k, k - 1) = 0$

            $T(k - 1, k) = T(k, k - 1)$

        else

            $T(k, k) = rtr_c/(rtr_p \alpha_p) + 1/\alpha_c$

            $T(k, k - 1) = -\sqrt{rtr_c/rtr_p}/\alpha_p$

            $T(k - 1, k) = T(k, k - 1)$

        end if

        Call LAPACK routine SSTEQR to get eigenvalues/vectors (Z) of T.

        Compute error in approximate eigenvalue for normal operator.

```
end if
rtr_p = rtr_c
rtr_c = ⟨r, r⟩
α_p = α_c
if √(rtr_c) < tol
   break
end if
if eigenvector flag = true
   QZ = X
end if
end
```

6.5 Application of Resolution to the Marine Data Experiments

In this section we apply the conjugate gradient and Lanczos algorithms just described to the linear inverse problem of estimating the three elastic parameter reflectivities described in Chapter 4. We will compare the resolution obtained for the reflectivities in Experiments 3 and 5 described in Table 4.1. These experiments were performed on a real marine common midpoint data gather close to a logged well. The data is shown in Figure 6.1 . The background velocity (shown in Figure 6.2) was the same for both experiments and was not updated. In Experiment 3 the air gun model anisotropic source was used to estimate the three reflectivities (short-scale relative fluctuation in the P-wave impedance, S-wave velocity and P-wave velocity divided by density). In Experiment 5 an anisotropic inversion-estimated source was used. Here I reran these experiments using the conjugate gradient and Lanczos procedures and L^2 norm.

Of course the Lanczos process only *estimates* the eigenvalues and eigenvectors. The error in these estimates is given by the theorem repeated below from [20], p. 479:

Theorem 6.2 Suppose that j steps of the Lanczos algorithm have been performed and that $S_j^t T_j S_j = diag(\theta_1, \ldots, \theta_j)$ is the Schur decomposition

of the tridiagonal matrix T_j. If $Y_j = [y_1, \ldots, y_j] = Q_j S_j \in \Re^{j \times j}$ then for $i = 1 : j$ we have $\|Ay_i - \theta_i y_i\|_{L^2} = |\beta_j| |s_{ji}|$ where $S_j = (s_{pq})$.

In the theorem above, the matrix S contains the eigenvectors of the tridiagonal matrix T which comes out of the Lanczos procedure. Q is the Lanczos matrix. Y is the matrix of eigenvectors of the original matrix A; θ_i are the eigenvalues of T and approximate eigenvalues of A. Finally, β_j is the last off-diagonal entry in the tridiagonal matrix T_j.

To compute our analogue of the SVD resolution matrix $R = VV^t$ for each of the three experiments, we could not, therefore, rely on all the computed eigenvectors of the normal operator (columns of V). Figures 6.4, and 6.10 show the approximate eigenvalues graphed with the associated errors. Figures 6.5, and 6.11 show the relative error in the eigenvalues plotted against eigenvalue number (where "1" is always the smallest computed eigenvalue.)

The matrix V was then composed of the eigenvectors corresponding to the largest eigenvalues (which were the best determined in these experiments) with approximation error $\leq 30\%$.

A second well-known source of error in the Lanczos process comes from roundoff and cancellation. These sources of error cause the Lanczos vectors to lose orthogonality. Error analysis done on this problem has been the motivation behind newer Lanczos procedures which attempt to minimize this loss of orthogonality (see [39]). The approach we implemented does not take advantage of these newer methods and is, therefore, subject to the problems of loss of orthogonality of the Lanczos vectors after a large number of iterations of the conjugate gradient procedure have been performed. We ran numerous experiments, therefore, to try to maximize the accuracy of the eigenvalue approximations while minimizing the loss of orthogonality of the Lanczos vectors. Trial and error indicated that about 30 iterations of the conjugate gradient algorithm is optimal for this particular problem. If we let $Q \in \Re^{n \times j}$ be the Lanczos matrix gotten after j iterations of the conjugate gradient algorithm, then ideally we should have that $Q^t * Q = I$ with I the $j \times j$ identity matrix. Plots of $Q^t * Q$ for the two Experiments are shown in Figures 6.6 and 6.12.

One of the main difficulties one has with trying to understand the resolution matrix R is processing the information. The model resolution matrix $R \in \Re^{n \times n}$ may be quite large. The closer this matrix is to the identity $I \in \Re^{n \times n}$, the better the resolution. Various techniques have been devised for selecting interesting columns of

R to examine. Three columns of the parts of the resolution matrix corresponding to the P-wave impedance reflectivity for the two experiments are shown in Figure 6.7. The last pair of graphs corresponds to an area near the target for each experiment. However, the top two pairs were chosen at random. If the model'parameter is perfectly resolved in the inversion at a certain depth, the plotted column should have a unit spike at the depth corresponding to that column of the resolution matrix and be zero elsewhere. For the three column comparisons shown, the top graph of the pairs (corresponding to Experiment 5) is closer to a spike than the bottom graph (Experiment 3). However, unsurprisingly, none of the columns shown is very close to a delta pulse. We are after all only estimating 30 eigenvalues out of the total of 626 possible in these inversion experiments. Wiggins [62] advocates examining a more intelligent choice of columns of the resolution matrix, namely those with the largest diagonal elements in the resolution matrix (which he calls "delta vectors"). These vectors should indicate the locations of well-determined components of the model.

Rather than plotting arbitrary columns of the large resolution matrix, one could calculate a function of the matrix termed the resolution spread (see [4]). One example is the Backus-Gilbert spread function which returns a single number for each resolution matrix:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} (i - j)^2 R_{ij}^2.$$

Unfortunately, this function is attempting to convey a large amount of information in a single number and is easily corrupted by lack of information in some parts of the model domain or noise in the data. Another choice is the vector-valued spread function:

$$Sp_i = \frac{\sum_{j=1}^{n} (i - j)^2 R_{ij}^2}{\sum_{j=1}^{n} R_{ij}^2}. \tag{6.14}$$

The idea is to weight more heavily the parts of the model estimate derived from averaging the true model over wide intervals. The points on the resolution spread curve closest to zero are the best resolved depths. More work needs to be done to understand the meaning of the size of the elements in the spread vector for these experiments.

We show various comparisons of resolution spread. Figure 6.8 overplots the resolution spread for each of the three reflectivities estimated in Experiment 3. These curves give little information (increase rapidly) at the water bottom (where there is no information in the data) and after the target located in the interval of 2100-2200m

(due to noise). Nonetheless, one can use this graph to compare relative resolution for the three different parameters estimated in one inversion experiment. The normalization of the spread function prevents the P-wave impedance resolution (the part of the resolution matrix with biggest norm) from overpowering the other two parameters' resolution information. Figure 6.13 plots the same spread functions for the three reflectivities estimated in Experiment 5. Figure 6.14 compares the resolution spread for the P-wave impedance reflectivity for Experiment 3 (dashed line) and Experiment 5 (solid line). Figure 6.15 is a graph of the same resolution comparison for the S-wave velocity reflectivity. Finally, 6.16 compares the resolution spread for the two experiments for the P-wave velocity divided by density reflectivity. One notes that in all three of these graphs, the interval of interest (1000-2300m depth) is better resolved in Experiment 5 (the reflectivity estimation job using the anisotropic inversion-estimated source) than in Experiment 3 (the air gun source experiment). At the target, however, the air gun source does slightly better than the anisotropic inversion-estimated source. These resolution spread pictures agree with the well log comparisons given in Chapter 4, Figures 4.10, 4.11, 4.15, 4.16, 4.19, 4.20. The inversion-estimated sources tend to do a better job of agreeing with the well logs over the depth domain of interest. However, at the target, the air gun source does a good job of matching the well log.

6.6   Conclusion

Model resolution for seismic inverse problems was first described in classic papers by Backus and Gilbert 25 years ago. The limited data obtained about the subsurface due to measuring tools and computational discretization prevents the inversion process from being able to return model estimates that are the true parameters which generated the data. The singular value decomposition of the forward (seismogram) model and generalized inverse allow us to quantify to what extent the model estimates are localized averages of the true model parameters. The smaller the averaging interval (for example in depth) of the model estimates, the better the resolution. A number of papers were written on this subject in the late 1960's and early 1970's when it seemed a promising tool. However, computing the resolution matrix for inverse problems is still not common practice because it is prohibitive to calculate the SVD for realistic experiments. We present here an inexpensive way to approximate the resolution matrix. We solve the least squares inverse problem for the reflectivities using

the conjugate gradient algorithm and simultaneously estimate some of the eigenvalues and eigenvectors of the normal matrix by the Lanczos procedure. The estimated eigenvectors are used to form the resolution matrix for the reflectivities. We illustrate this process on two reflectivity experiments with different sources. The resulting resolution spread curves agree with the well log comparisons shown in Chapter 4 and imply that resolution defined in this way could be useful for gauging the accuracy of inversion results when well log information is not available.
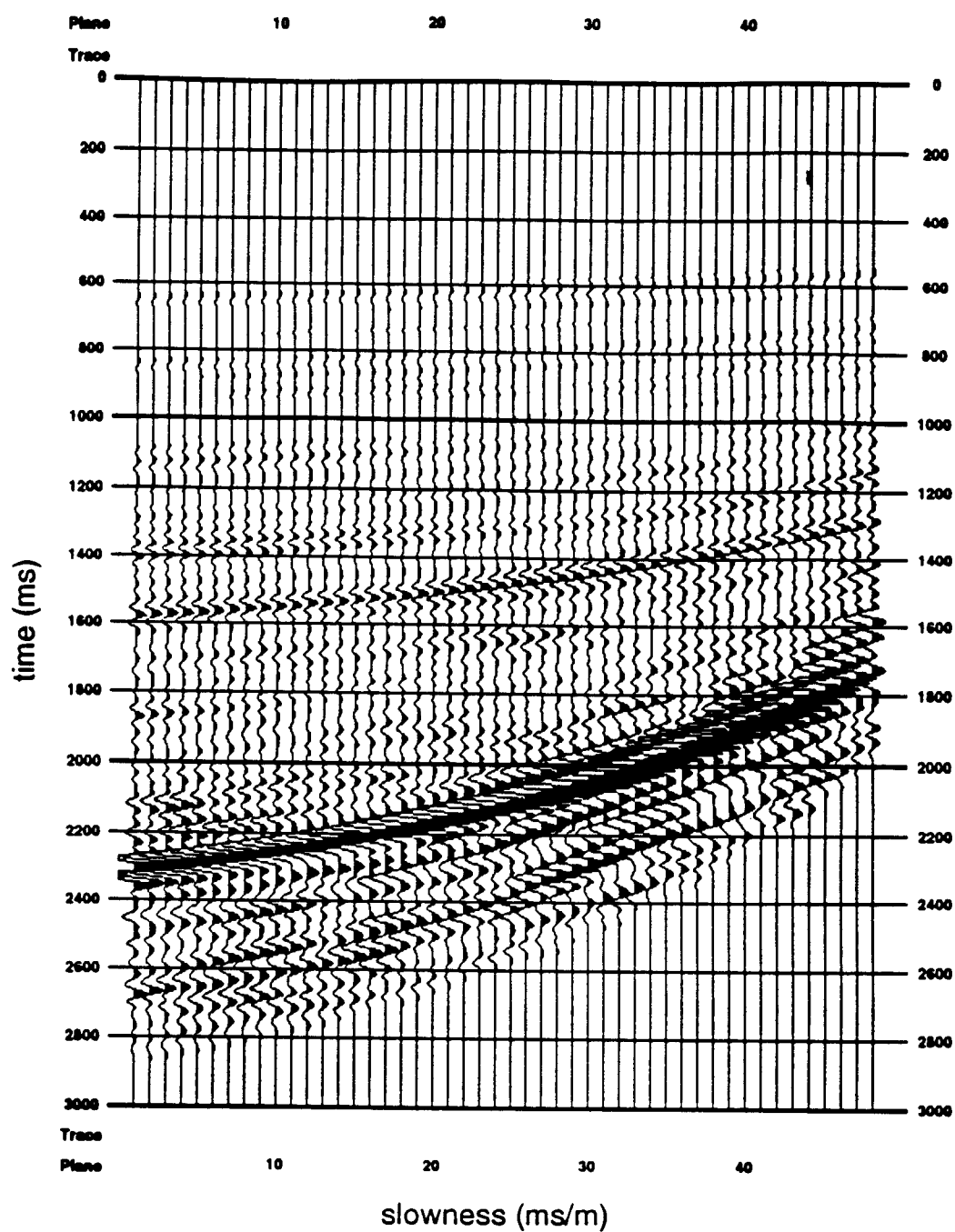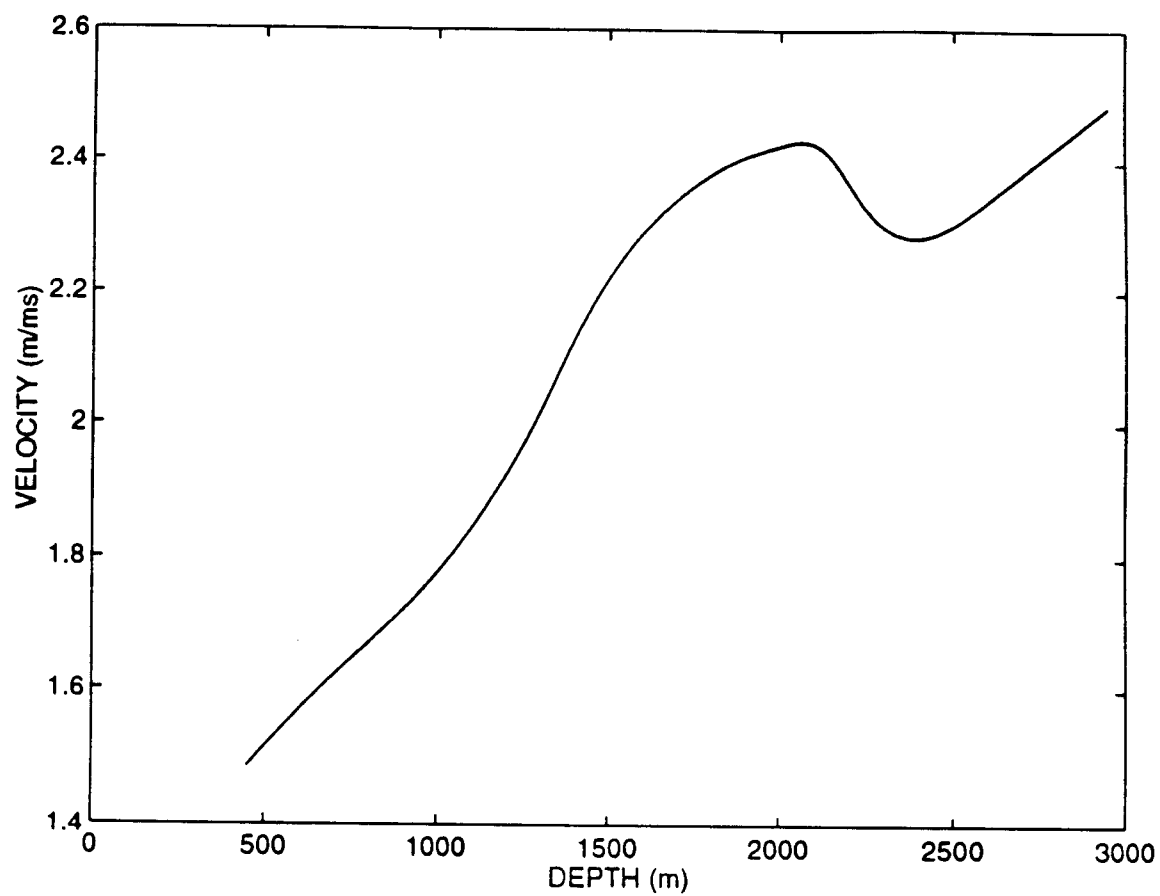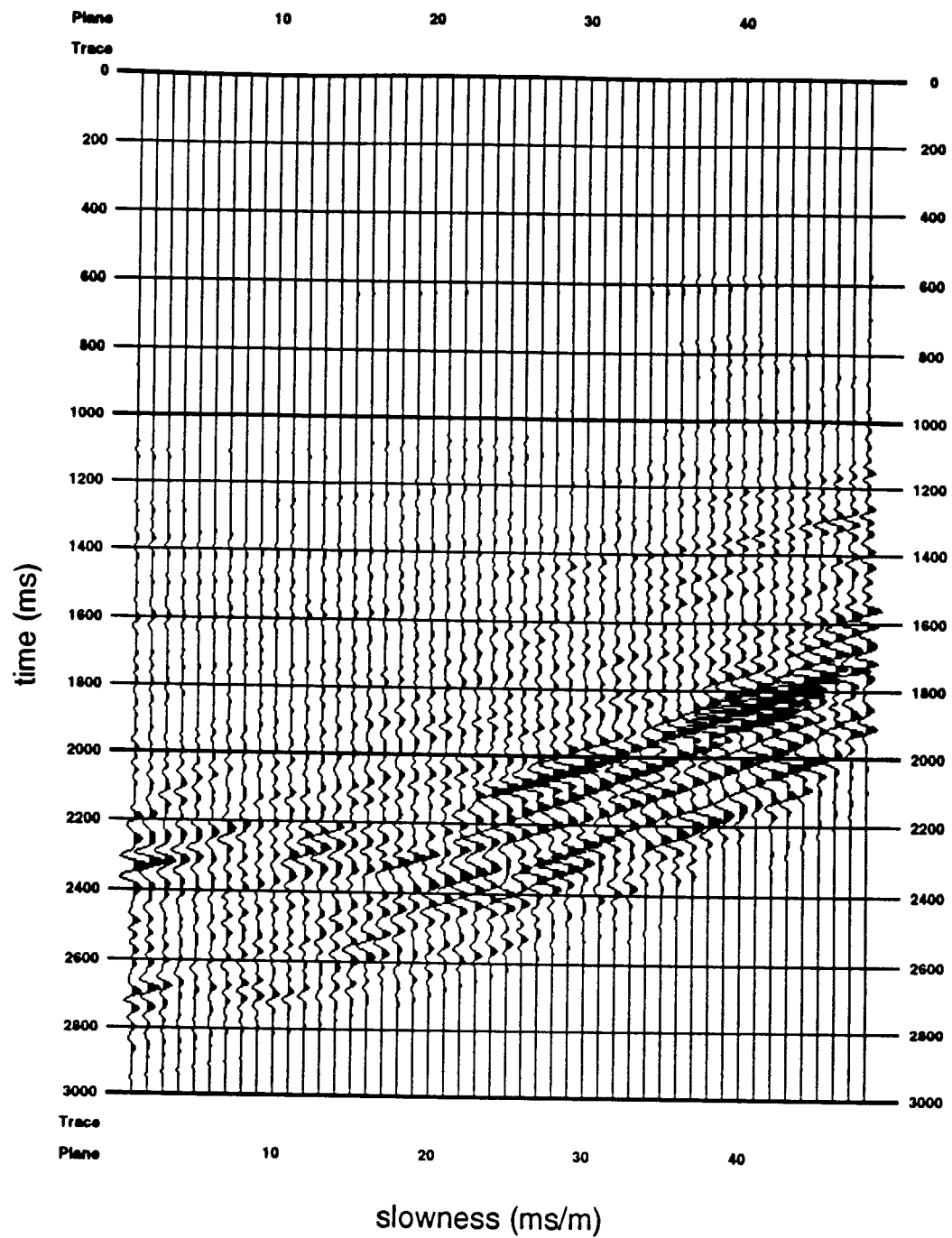
Figure 6.1 The common midpoint data gather (CMP6) used for the three experiments contrasted in this chapter.

Figure 6.2 The P-wave background velocity estimate used for both inversion experiments. This background velocity was estimated from a DSO inversion described in chapter 4.

**Figure 6.3** Difference between actual and predicted data gotten from inverting the reflectivities with the air gun source fixed (L2 norm). The misfit is plotted on the same scale as the actual data.
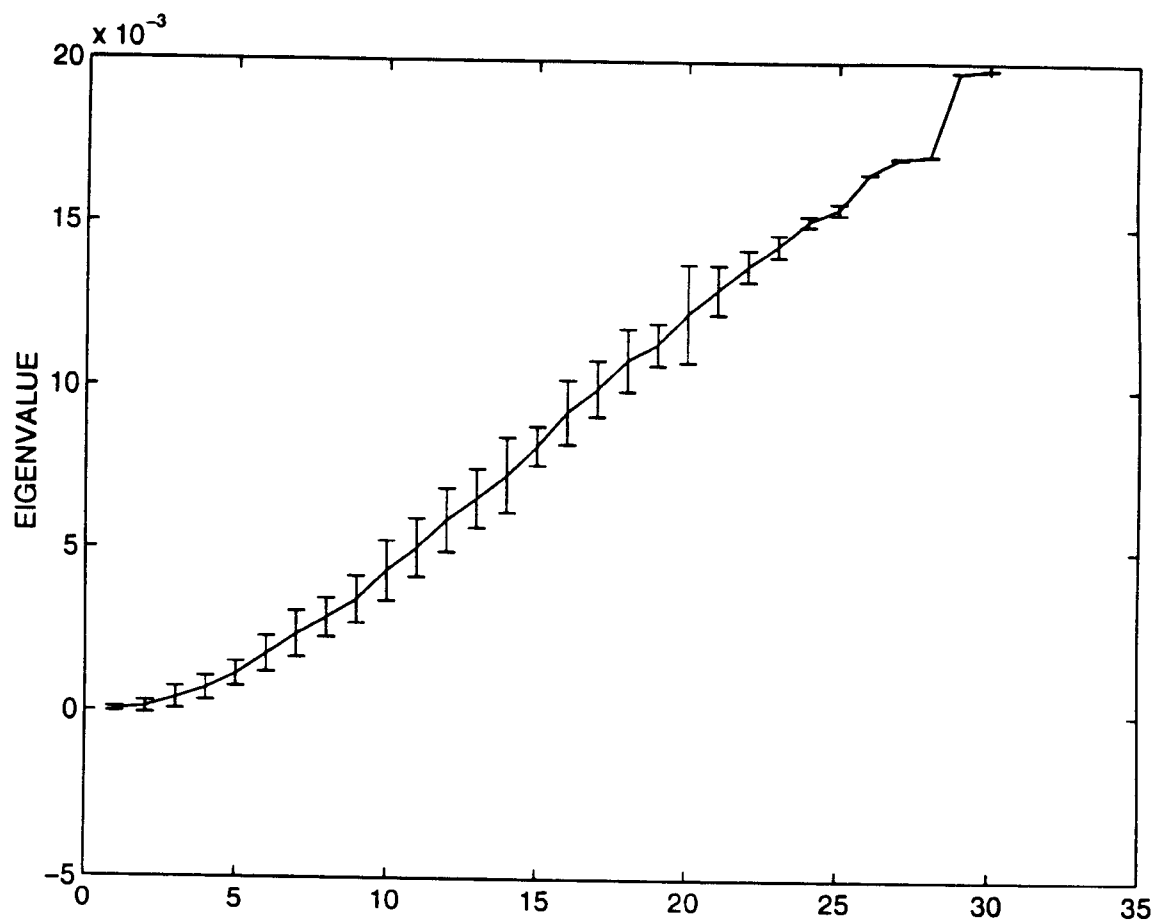
Figure 6.4 The approximate eigenvalues of the normal operator for the reflectivity inversion with a fixed air gun model source (Experiment 3 Table 4.1). The errors in the eigenvalues are shown as errorbars on the graph.
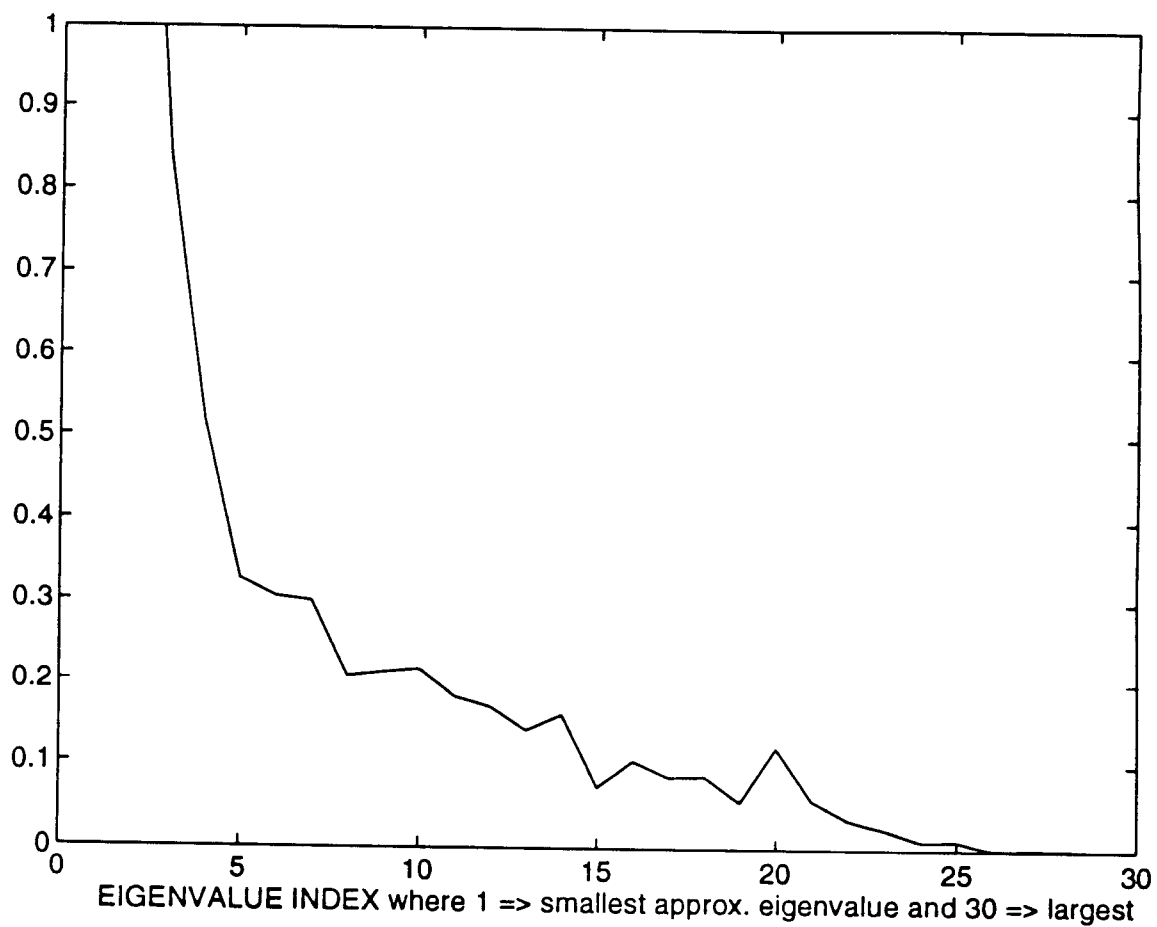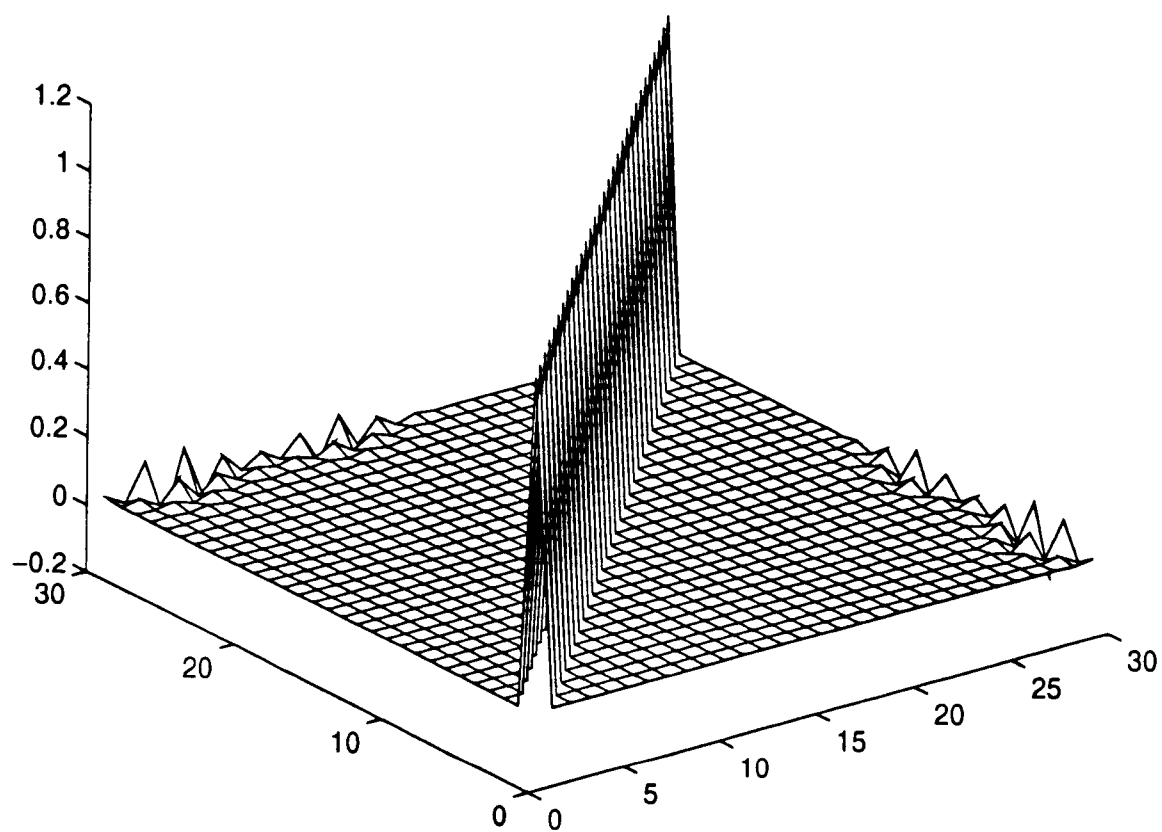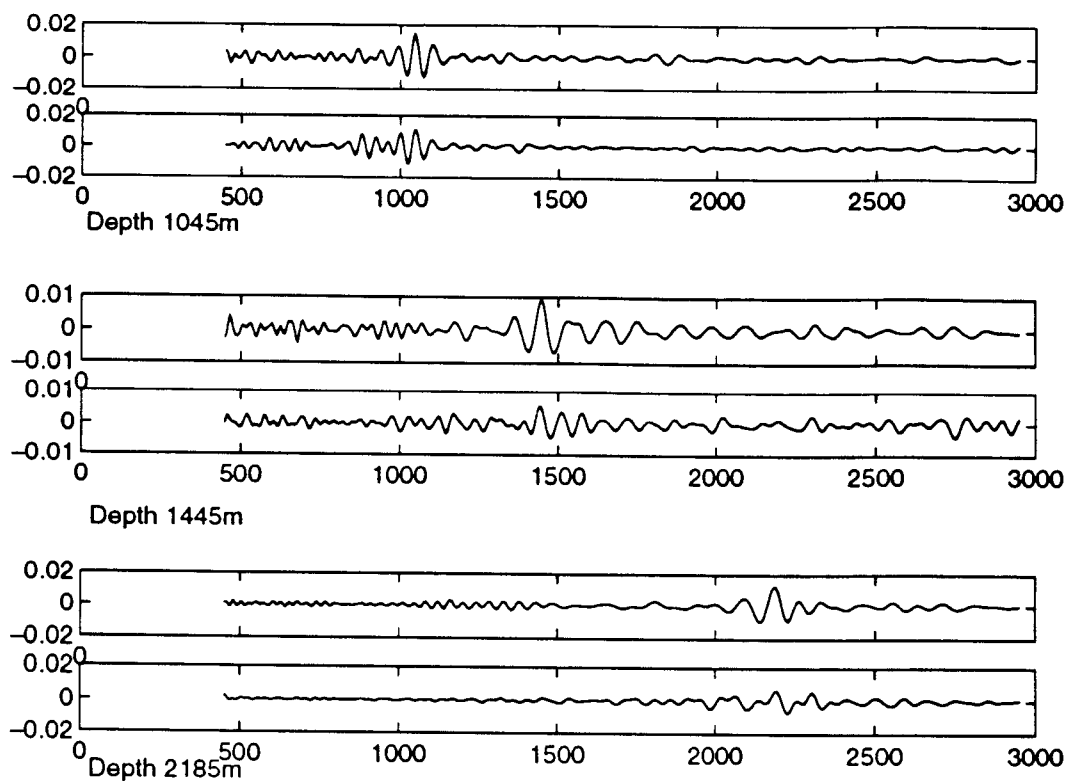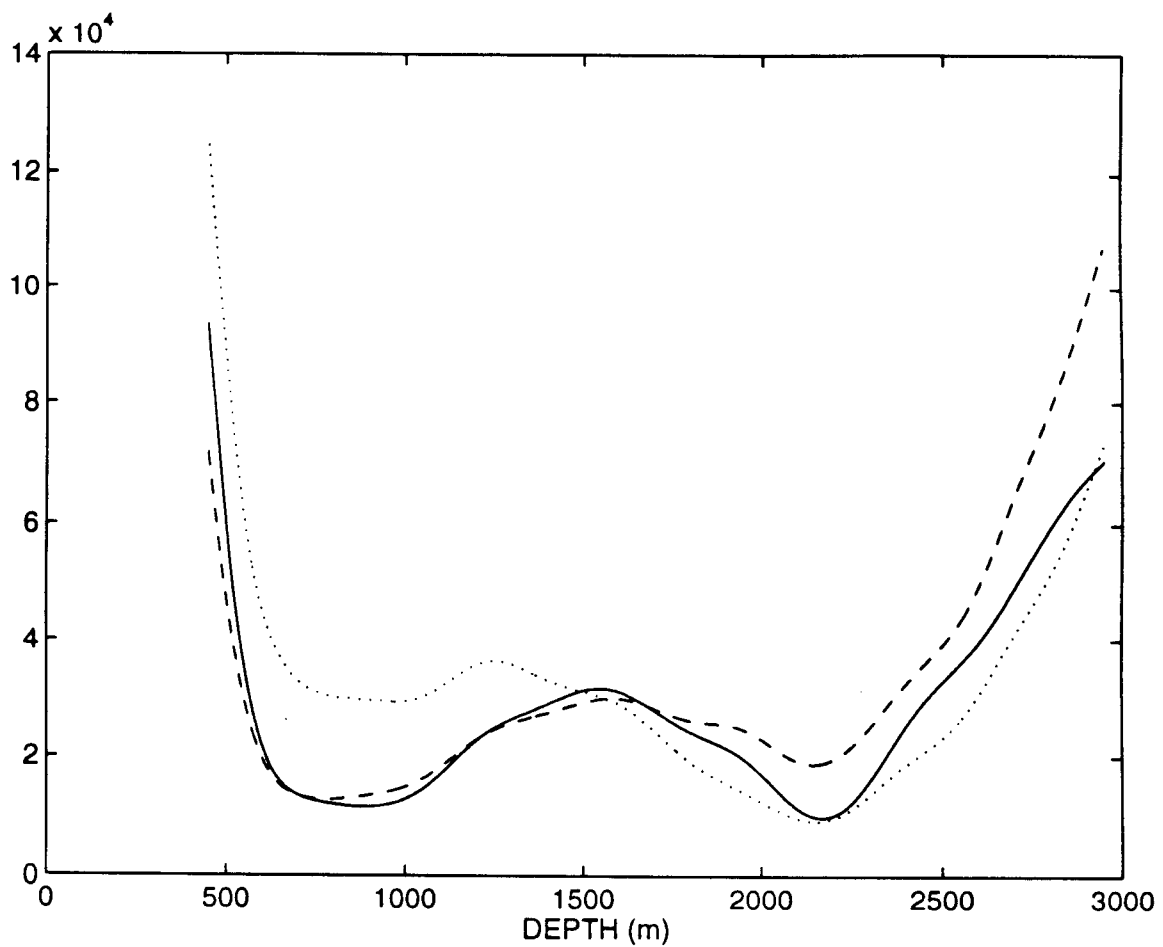
Figure 6.5   The relative error in the approximate eigenvalues of the normal operator for the reflectivity inversion with a fixed air gun model source (Experiment 3 Table 4.1).

Figure 6.6 The dot product of Lanczos matrix with itself for Experiment 3 Table 4.1 after 30 iterations of the conjugate gradient algorithm.

Figure 6.7   Three selected columns of the resolution matrices for Experiments 3 and 5. The top two graphs correspond to depth 1045m. The middle two graphs are the column of the resolution matrices corresponding to 1445m. The bottom two graphs correspond to depth 2185m (a region between where the two experiments place the target). In each pair, the top graph corresponds to Experiment 5 (reflectivity estimation with the anisotropic inversion-estimated source). The bottom graph corresponds to Experiment 3 (reflectivity estimation with the air gun model source).

Figure 6.8 Graph of the spread of the resolution matrix for Experiment 3. The solid line corresponds to the spread for the P-wave impedance reflectivity estimate. The dashed line describes the spread for the S-wave velocity reflectivity. The dotted line is the spread for the P-wave velocity divided by density reflectivity.
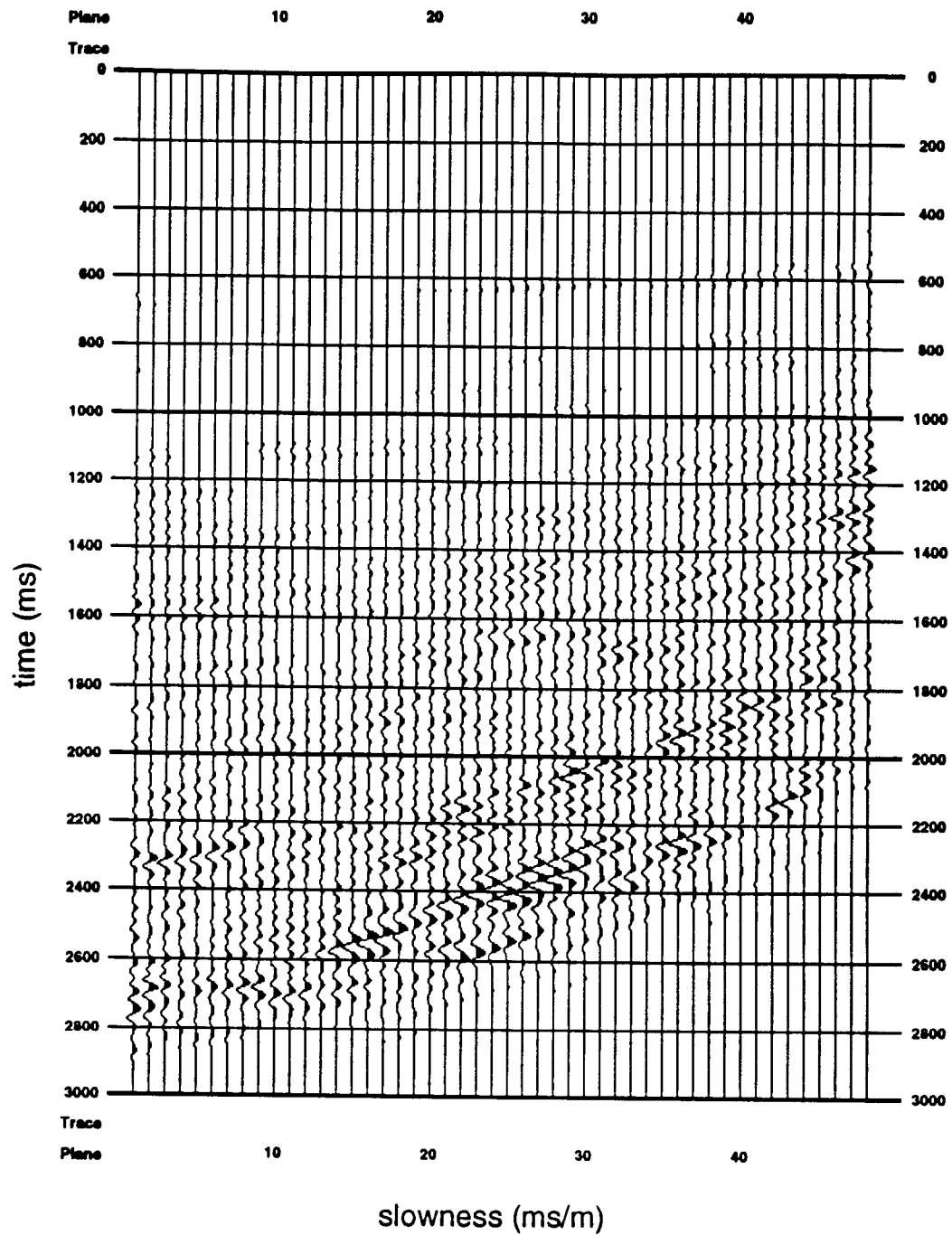
Figure 6.9 Difference between actual and predicted data gotten from inverting the reflectivities and an anisotropic source (L2 norm). The misfit is plotted on the same scale as the actual data.
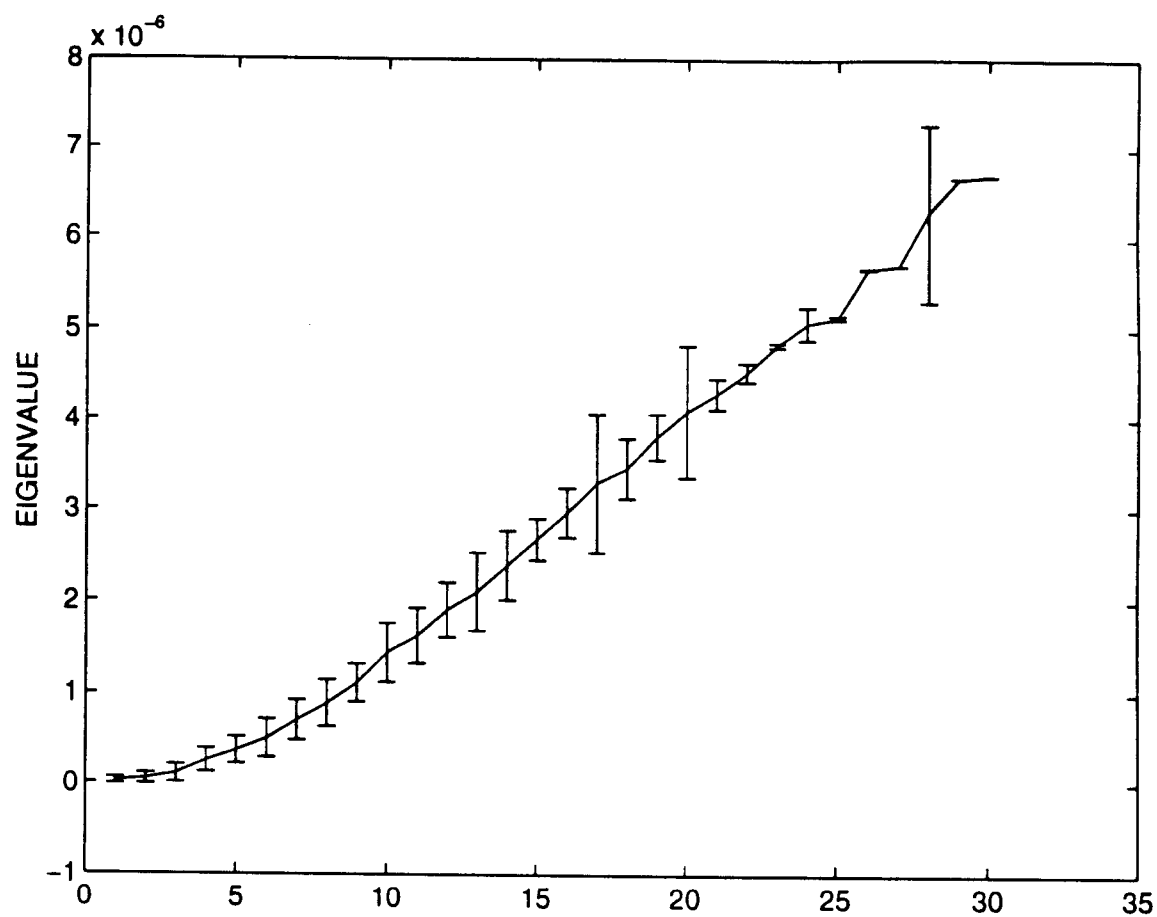
Figure 6.10 The approximate eigenvalues of the normal operator for the reflectivity inversion with an inversion-estimated anisotropic source (Experiment 5 Table 4.1). The errors in the eigenvalues are shown as errorbars.

Figure 6.11 The relative error in the approximate eigenvalues of the normal operator for the reflectivity inversion with an anisotropic inversion-estimated source (Experiment 5 Table 4.1).

Figure 6.12   The dot product of Lanczos matrix with itself for Experiment 5 Table 4.1 after 30 iterations of the conjugate gradient algorithm.
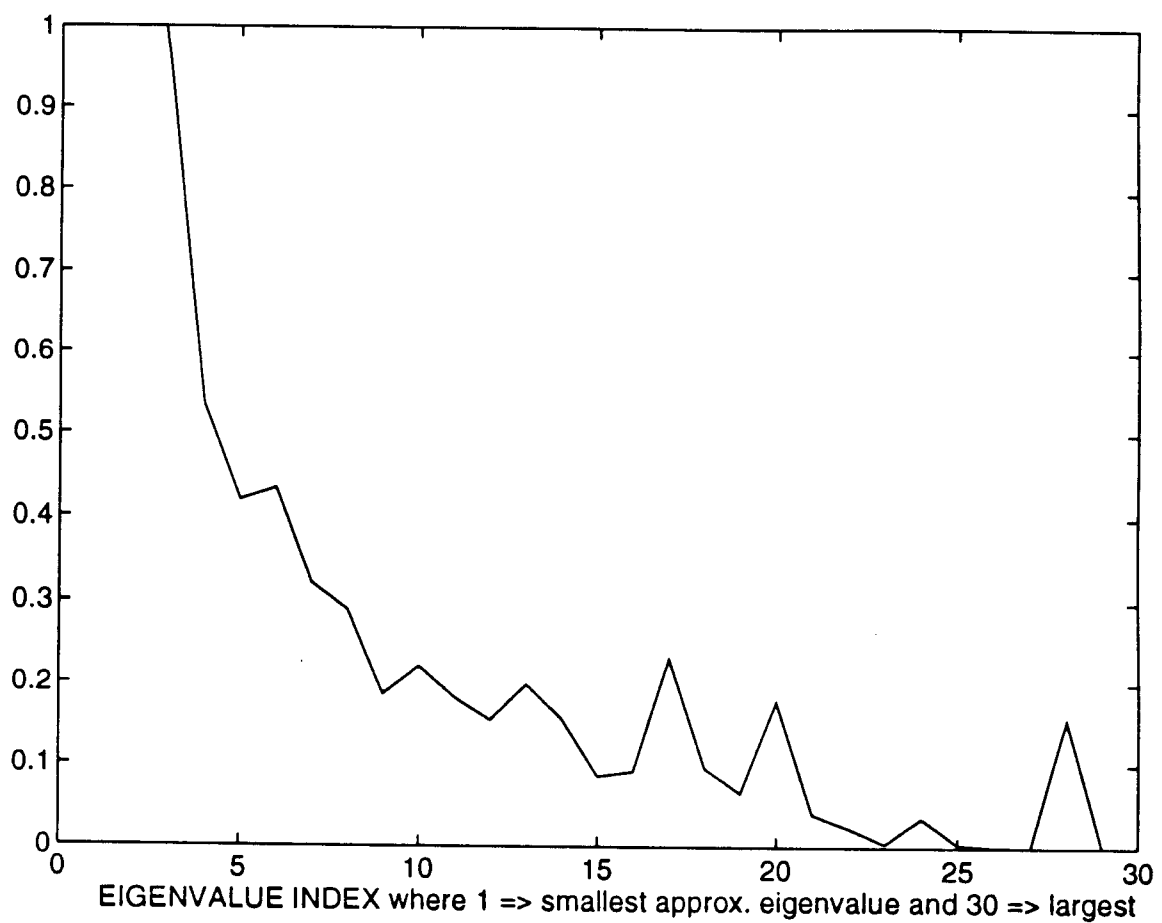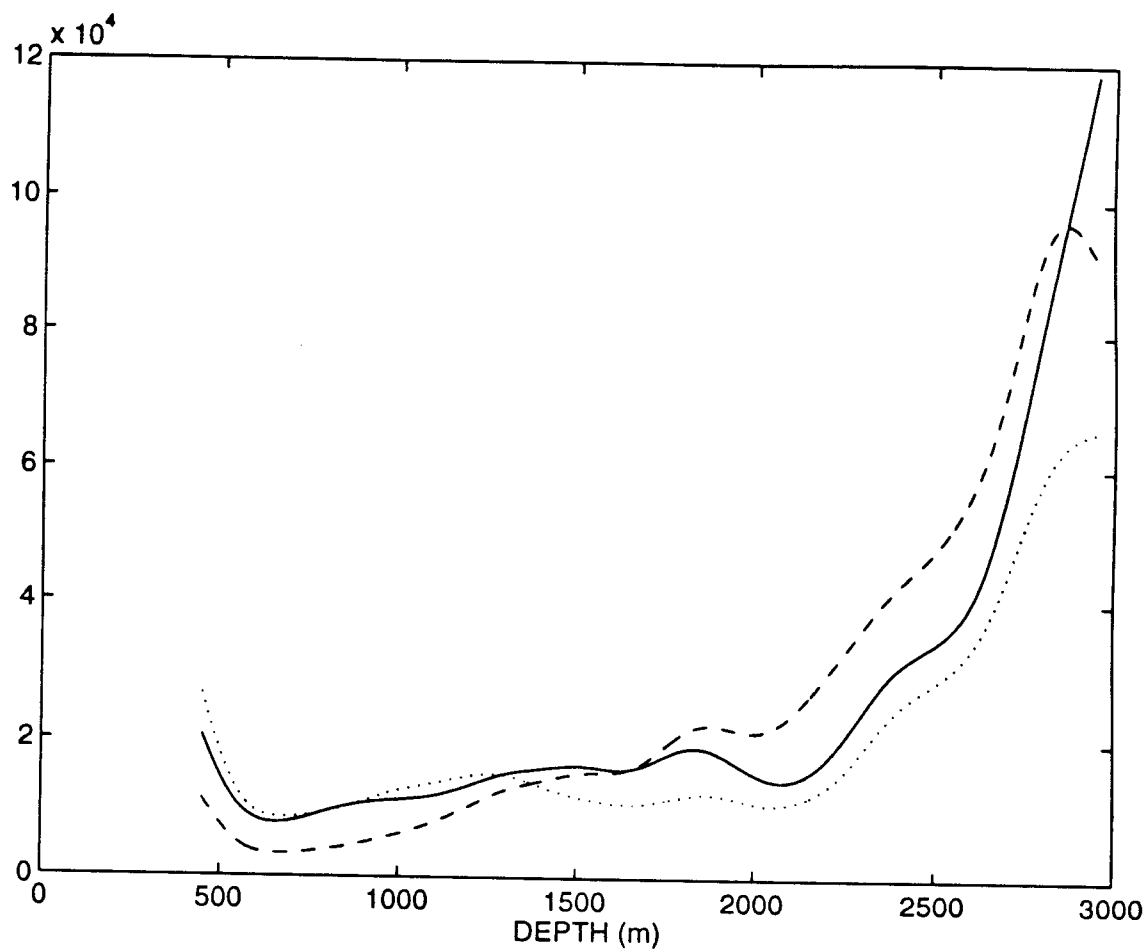
Figure 6.13 Graph of the spread of the resolution matrix for Experiment 5. The solid line corresponds to the spread for the P-wave impedance reflectivity estimate. The dashed line describes the spread for the S-wave velocity reflectivity. The dotted line is the spread for the P-wave velocity divided by density reflectivity.
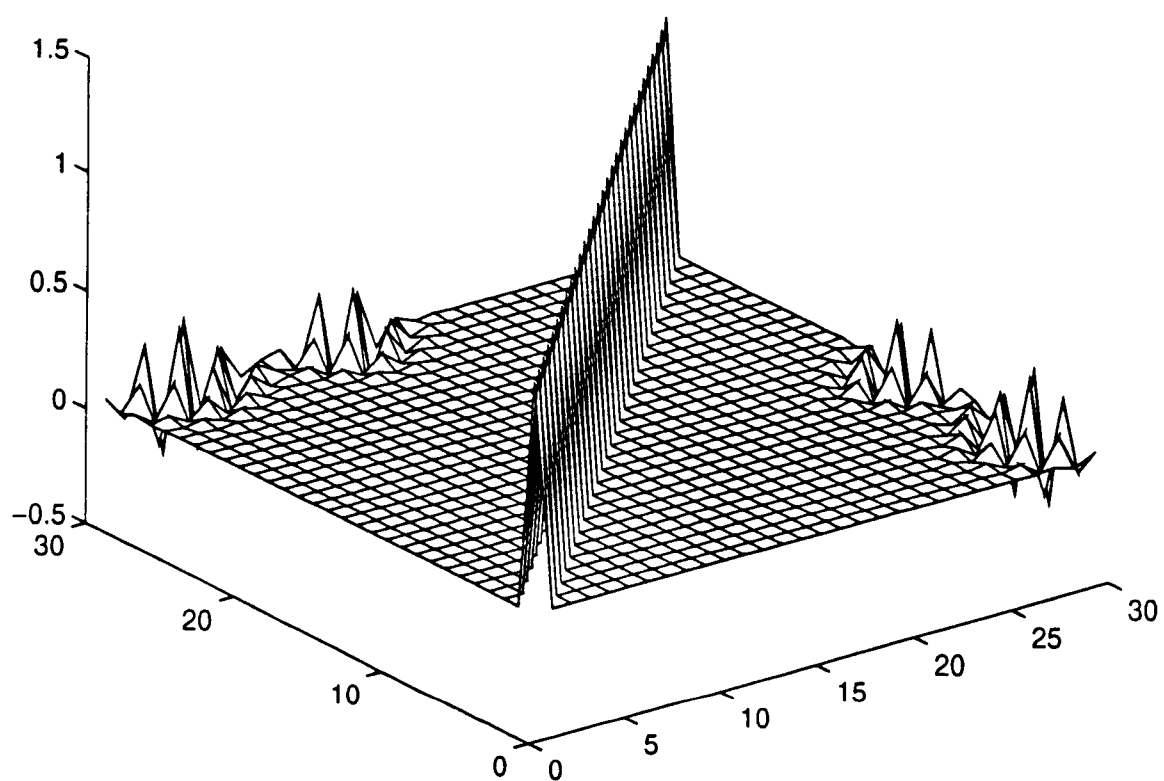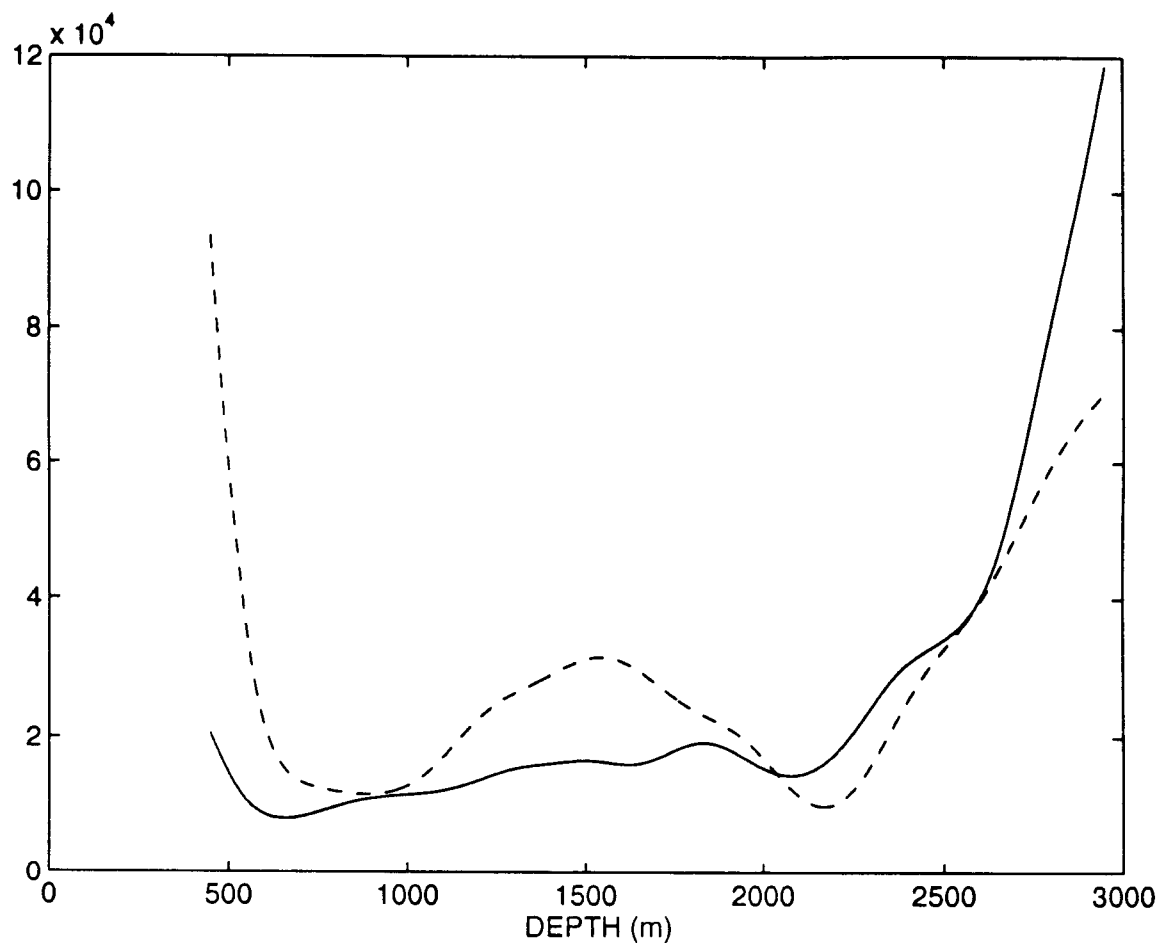
Figure 6.14 The graph of the resolution spread for two estimates of the P-wave impedance reflectivity. The solid line corresponds to Experiment 5; the dashed line to Experiment 3.
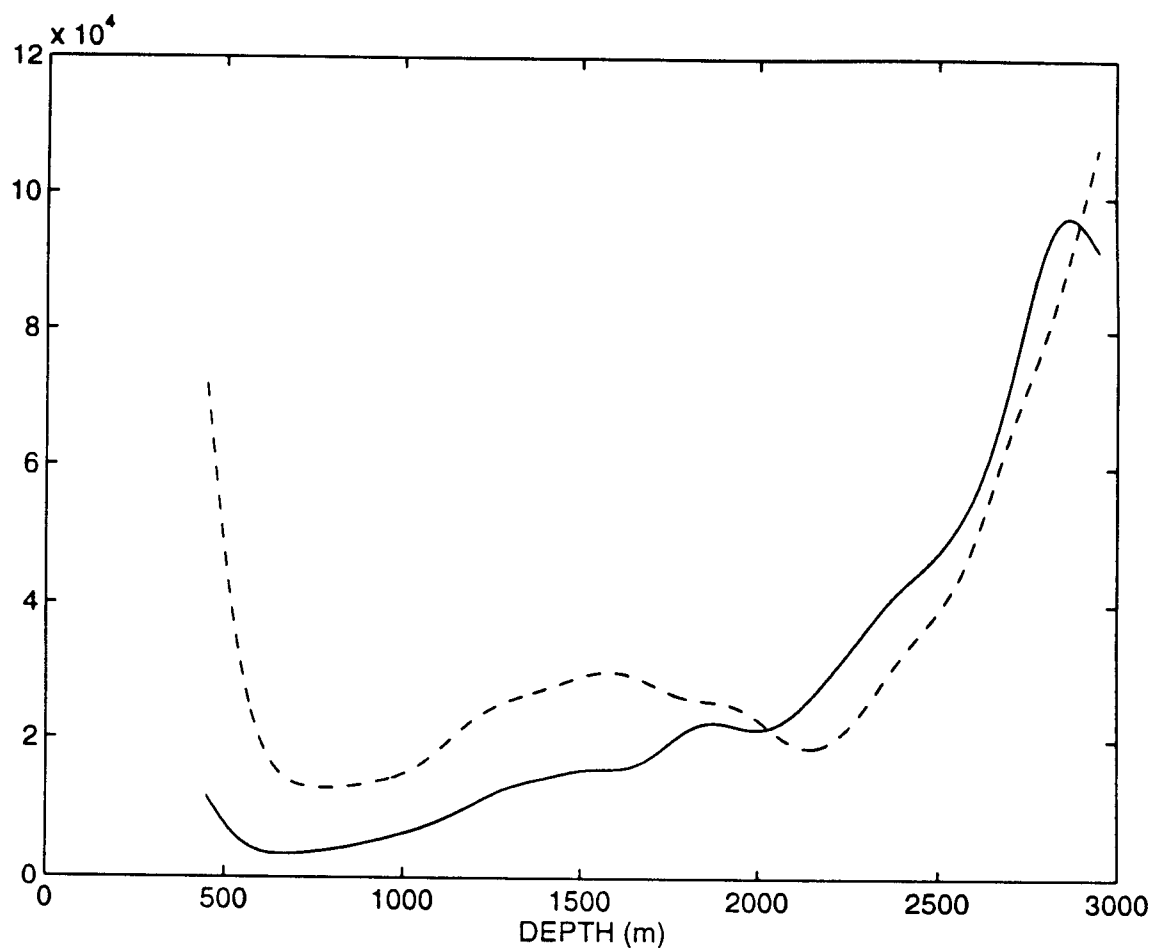
Figure 6.15 The graph of the resolution spread for two estimates of the S-wave velocity reflectivity. The solid line corresponds to Experiment 5; the dashed line to Experiment 3.
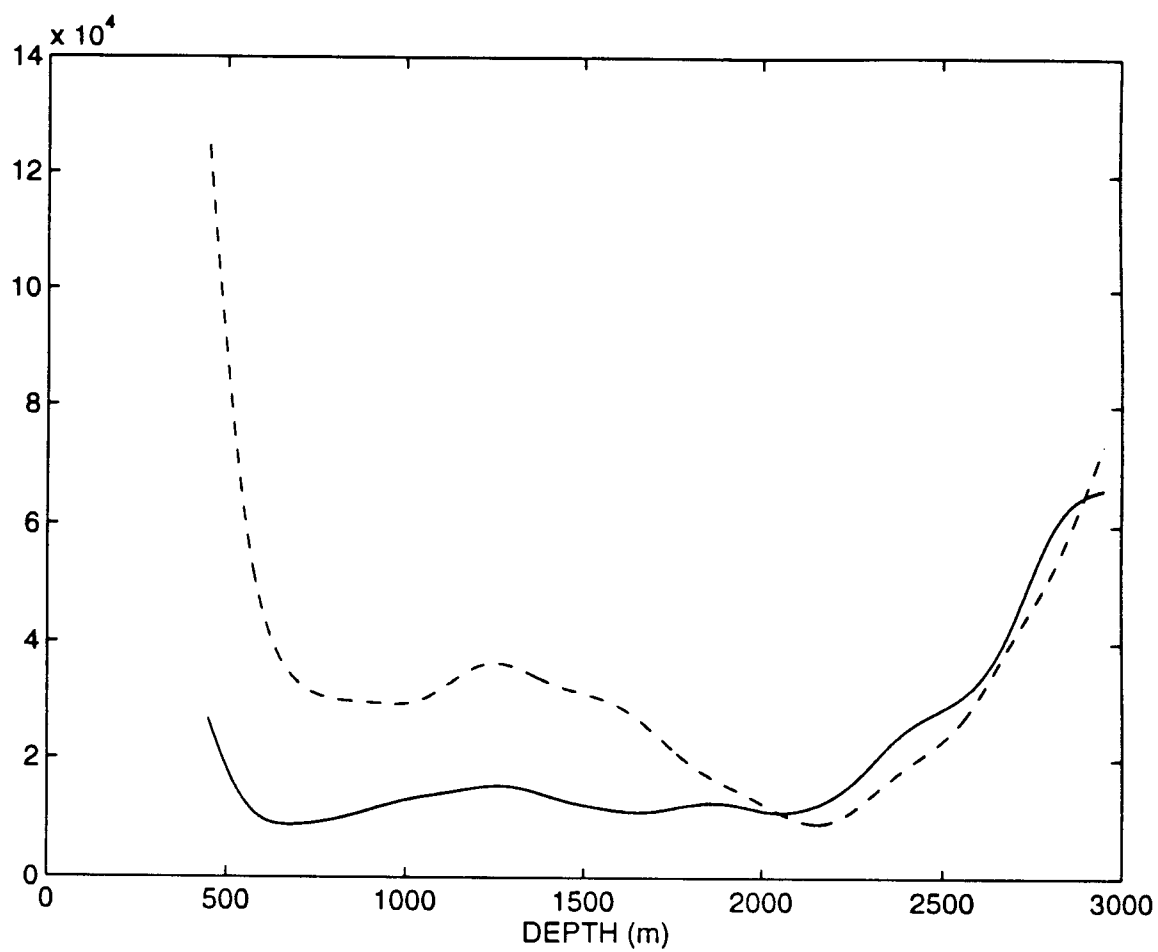
Figure 6.16 The graph of the resolution spread for two estimates of the P-wave velocity divided by density reflectivity. The solid line corresponds to Experiment 5; the dashed line to Experiment 3.

Bibliography

[1] K. AKI and P. RICHARDS. *Quantitative Seismology: Theory and Methods.* Freeman, San Francisco, 1980.

[2] T. M. APOSTOL. *Mathematical Analysis.* Addison-Wesley Publishing Company, Reading, Massachusetts, 2 edition, 1974.

[3] F. ASSOUS and F. COLLINO. A numerical method for the exploration of sensitivity: the case of the identification of the 2d stratified elastic medium. *Inverse Problems*, 6:487–513, 1990.

[4] G. BACKUS and F. GILBERT. The resolving power of gross earth data. *Geophys. J. R. Astr. Soc.*, 16:169–205, 1968.

[5] G. BEYLKIN. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized radon transform. *J. Math. Phys.*, 26:99–108, 1985.

[6] T.N. BISHOP, K.P. BUBE, R.T. CUTLER, R.T. LANGAN, P.L. LOVE, J.R. RESNICK, R.T. SHUEY, D.A. SPINDLER, and H.W. WYLD. Tomographic determination of velocity and depth in laterally varying media. *Geophysics*, 50:903–923, 1985.

[7] K. BUBE, P. LAILLY, P. SACKS, F. SANTOSA, and W.W. SYMES. Simultaneous determination of source wavelet and velocity profile using impulsive point-source data from a layered fluid. *Geophys. J.*, 95:449–462, 1988.

[8] P. BUHL, P.L. STOFFA, and G.M. BRYAN. The application of homomorphic deconvolution to shallow-water marine seismology – part ii: real data. *Geophysics*, 39:417–426, 1974.

[9] G. CANADAS and P. KOLB. Least-squares inversion of prestack data: simultaneous identification of density and velocity of stratified media. In *Proc. 56th*

Annual International Meeting, pages 604–607, Houston, 1986. Society of Exploration Geophysicists. Expanded abstract.

[10] J.P. CASTAGNA and M.M BACKUS, editors. *Offset-Dependent Reflectivity – Theory and Practice of AVO Analysis*. SEG, Tulsa, 1993.

[11] J.P. CASTAGNA, M. BATZLE, and R. EASTWOOD. Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks. *Geophysics*, 50:571–581, 1985.

[12] J.P. CASTAGNA and S.W. SMITH. Comparison of avo indicators: A modeling study. *Geophysics*, 59:1849–1855, 1994.

[13] R. COURANT and D. HILBERT. *Methods of Mathematical Physics*, volume II. Wiley-Interscience, New York, 1962.

[14] M.B. DOBRIN and C.H SAVIT. *Introduction to Geophysical Prospecting*. McGraw-Hill Book Company, New York, fourth edition, 1988.

[15] S.C. EISENSTAT, H.C. ELMAN, and M.H. SCHULTZ. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM J. Numer. Anal.*, 20:345–357, 1983.

[16] R. FLETCHER. *Practical methods of optimization, I:Unconstrained Optimization*. Wiley & Sons, New York, 1980.

[17] J.N. FRANKLIN. Well-posed stochastic extensions of ill-posed linear problems. *Journal of Mathematical Analysis and Applications*, 31:682–716, 1970.

[18] O. GAUTHIER, A. TARANTOLA, and J. VIRIEUX. Two-dimensional nonlinear inversion of seismic waveforms. *Geophysics*, 51:1387–1403, 1986.

[19] M.S. GOCKENBACH, W.W. SYMES, and R.A. TAPIA. The dual regularization approach to seismic velocity inversion. *Inverse Problems*, in press, 1995.

[20] G. GOLUB and C. V. LOAN. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 1989.

[21] G.H. GOLUB and D.P. O'LEARY. Some history of the conjugate gradient and lanzos algorithms: 1948-1976. *SIAM Review*, 31:50–102, 1989.

[22] J HELGESEN. *Inversion of seismograms using finite difference modelling techniques.* PhD thesis, Institutt for Teleteknikk, Trondheim, Norway, 1990.

[23] M.R. HESTENES. *Conjugate Direction Methods in Optimization.* Springer-Verlag, Berlin, 1980.

[24] M.R. HESTENES and E. STIEFEL. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Stand.*, 49:409–436, 1952.

[25] H. IGEL. *Seismic Modeling and Inversion.* PhD thesis, Institut de Physique du Globe de Paris, Paris, France, 1993.

[26] D. ISAACSON. Distinguishability of conductivities by electric current computed tomography. *IEEE Transactions on Medical Imaging*, 5:91–95, 1986.

[27] D.D. JACKSON. The use of *a priori* data to resolve non-uniqueness in linear inversion. *Geophys. J. R. astr. Soc.*, 57:137–157, 1979.

[28] T. KATO. *Perturbation Theory for Linear Operators.* Springer, New York, 1966.

[29] B. KENNETT and G. NOLET. Resolution analysis for discrete systems. *Geophys. J. R. astr. Soc.*, 53:413–425, 1978.

[30] M. KERN and W.W. SYMES. Inversion of reflection seismograms by differential semblance analysis: algorithm structure and synthetic examples. *Geophysical Prospecting*, 42:565–614, 1994.

[31] C. LANCZOS. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bur. Stand.*, 45:255–282, 1950.

[32] M. LANDRO, J. LANGHAMMER, R. SOLLIE, L. AMUNDSEN, and E. BERG. Source signature determination from ministreamer data. *Geophysics*, 59:1261–1269, 1994.

[33] R.M. LEWIS. *Source-Velocity Identification for a Layered Model of Reflection Seismology.* PhD thesis, Department of Mathematical Sciences, Rice University, Houston, Texas, U.S.A, 1989.

[34] R.D. MARTINEZ and G.A. MCMECHAN. $\tau - p$ seismic data for viscoelastic media – part 1: modelling. *Geophysical Prospecting*, 39:141–156, 1991.

[35] R.D. MARTINEZ and G.A. MCMECHAN. $\tau - p$ seismic data for viscoelastic media – part 2: linearized inversion. *Geophysical Prospecting*, 39:157–181, 1991.

[36] W. MENKE. *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, Inc., San Diego, 1989.

[37] W.J. OSTRANDER. Plane-wave reflection coefficients for gas sands at nonnormal angles of incidence. *Geophysics*, 49:1637–1648, 1984.

[38] G.S. PAN, C.Y. YOUNG, and J.P. CASTAGNA. Net pay delineation of gas sand using integrated target-oriented prestack elastic waveform inversion. In *63rd Annual International SEG Meeting*, pages 431–434, Washington, D.C., USA, September 1993. Society of Exploration Geophysicists.

[39] B.N. PARLETT. *The Symmetric Eigenvalue Problem*. Prentice-Hall, 1980.

[40] J. RAUCH. *Partial Differential Equations*. Springer-Verlag, New York, 1991.

[41] J.O.A. ROBERTSSON, J.O. BLANCH, and W.W. SYMES. Viscoelastic finite-difference modeling. *Geophysics*, 59:1444–1456, 1994.

[42] E.A. ROBINSON. Predictive decomposition of seismic traces. *Geophysics*, 22:767–778, 1957.

[43] E.A. ROBINSON. Predictive decomposition of time series with application to seismic exploration. *Geophysics*, 32:418–484, 1967.

[44] E.A. ROBINSON. *Seismic velocity analysis and the convolutional model*. International Human Resources Development Corporation, Boston, 1983.

[45] E.A. ROBINSON and S. TREITEL. Principles of digital wiener filtering. *Geophysical Prospecting*, 15:311–333, 1967.

[46] P.E. SACKS. A velocity inversion problem involving an unknown source. *SIAM J. Appl. Math.*, 50(3):931–941, 1990.

[47] F. SANTOSA and W. W. SYMES. The determination of a layered acoustic medium via multiple impedance profile inversions from plane wave responses. *Geophysical Journal of the Royal Astronomical Society*, 81:175–195, 1985.

[48] J.A. SCALES, M.L. SMITH, and T.L. FISCHER. Global optimization methods for highly nonlinear inverse problems. In G. Cohen, L. Halpern, and P. Joly, editors, *Mathematical and Numerical Aspects of Wave Propagation Phenomena*, pages 434–444. Society for Industrial and Applied Mathematics, Philadelphia, 1991.

[49] M.K. SEN and P. STOFFA. Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane wave seismograms. *Geophysics*, 56:1794–1810, 1991.

[50] M.K. SEN and P. STOFFA. Nonlinear one-dimensional seismic waveform inversion using simulated annealing. *Geophysics*, 56:1624–1636, 1991.

[51] J.L. SIMMONS. *Practical seismic inversion*. PhD thesis, The University of Texas, Austin, Texas, U.S.A, 1994.

[52] D. SLEPIAN. Some comments on fourier analysis, uncertainty and modeling. *SIAM Review*, 25(3):379–393, 1983.

[53] H. SONG and W.W. SYMES. On a transmission inverse problem. *Inverse Problems*, in press, 1995.

[54] P.L. STOFFA, P. BUHL, and G.M. BRYAN. The application of homomorphic deconvolution to shallow-water mari ne seismology – part i: models. *Geophysics*, 39:401–416, 1974.

[55] T. SUZUKI. Stability of an inverse hyperbolic problem. *Inverse Problems*, 4:273–290, 1988.

[56] W. W. SYMES. Mathematical foundations of reflected wave imaging. Technical Report TR90-2, Department of Mathematical Sciences, Rice University, Houston, Texas, U.S.A., 1990.

[57] W.W. SYMES. On the relation between coefficient and boundary values for solutions of webster's horn equation. *SIAM J. Math. Anal.*, 17:1400–1420, 1986.

[58] W.W. SYMES and J. CARAZZONE. Velocity inversion by differential semblance optimization. *Geophysics*, 56(5):654–663, 1991.

[59] W.W. SYMES and J.J. CARAZZONE. Velocity inversion by coherency optimization. In J.B. Bednar, editor, *Geophysical Inversion*, pages 59–89. Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[60] S. TREITEL, P.R. GUTOWSKI, and D.E. WAGNER. Plane-wave decomposition of seismograms. *Geophysics*, 47:1375–1401, 1982.

[61] T.J. ULRYCH. Application of homomorphic deconvolution to seismology. *Geophysics*, 36:650–660, 1971.

[62] R. A. WIGGINS. The general linear inverse problem: implication of surface waves and free oscillations for earth structure. *Reviews of Geophysics and Space Physics*, 10(1):251–285, 1972.

[63] W.T. WOOD. Linearized least-squares inversion of field seismic data for an elastic 1-d earth. In *63rd Annual International SEG Meeting*, pages 661–664, Washington, D.C., USA, September 1993. Society of Exploration Geophysicists.

[64] A. ZIOLKOWSKI. *Deconvolution.* International Human Resources Development Corporation, Boston, 1984.

[65] A. ZIOLKOWSKI. Determination of the signature of a dynamite source using source scaling, part 1: Theory. *Geophysics*, 58:1174–1182, 1993.

[66] A. ZIOLKOWSKI. Determination of the signature of a dynamite source using source scaling, part 2: Experiment. *Geophysics*, 58:1183–1194, 1993.