

Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies

Laurent Sirgue* and R. Gerhard Pratt†

ABSTRACT

Prestack migration and/or inversion may be implemented in either the time or the frequency domain. In the frequency domain, it is possible to discretize the frequencies with a much larger sampling interval than that dictated by the sampling theorem and still obtain an imaging result that does not suffer from aliasing (wrap around) in the depth domain. The selection of input frequencies can be reduced when a range of offsets is available; this creates a redundancy of information in the wavenumber coverage of the target. In order to optimize the use of this information, we define a new discretization strategy that depends on the maximum effective offset present in the surface seismic survey: the larger the range of offsets, the fewer frequencies are required.

The strategy, exact in a homogeneous 1D earth, selects frequencies by making use of the well-known effect of image stretch in normal-moveout (NMO) correction and in migration (usually considered detrimental for the imaging). The strategy is also useful in more general earth models: we apply it to the 2D Marmousi model and recover a continuous range of wavenumbers using only three input frequencies. The Marmousi inversion result accurately predicts all other data frequencies, demonstrating the redundancy of the data.

INTRODUCTION

This paper develops a strategy for waveform inversion, a tool that is related to seismic migration. The objective of prestack depth migration is to obtain an image of the subsurface reflectivity field; in most implementations, the image is not intended as a quantitative measure of the true subsurface impedances. The objective of waveform inversion is more complete: it is to estimate a quantitative model of the subsurface in a way that

minimizes the differences (residuals) between observed and calculated seismic data, by modeling the physics of the measurement. Because of the high computational cost of calculating synthetic seismic data, the full nonlinear problem of waveform inversion is usually formulated as an iterative “descent” method, in which the minimization of residuals is achieved through the repeated calculation of a local gradient. The gradient at each iteration provides the direction of minimization of the “objective” functional, usually the L_2 norm of the data residuals, possibly combined with some form of regularizing constraints on the model. Since 1983, it has been recognized that the calculation of the gradient is of the same computational order as the forward modeling task, and that the gradient image is closely related to seismic migration (Lailly, 1983; Tarantola, 1984a). Success depends upon the topography of the misfit function at the location of the starting model. In the presence of strong nonlinearities, the inversion may fail, and convergence into a local minimum may occur (Gauthier et al., 1986). The starting model for such schemes must therefore be located in the neighborhood of the global minimum, that is, a descent path leading to the global minimum must exist if the method is to succeed.

Waveform inversion may be implemented either in the time domain (Tarantola, 1986; Mora, 1987; Bunks et al., 1995; Shipp and Singh, 2002) or in the frequency domain (Pratt and Worthington, 1990; Liao and McMechan, 1996); implementations for both acoustic and elastic wave equations exist. The frequency-domain approach is equivalent to the time-domain approach when all frequencies are inverted simultaneously (Pratt et al., 1998). Similarly, the inversion of a single sinusoidal component of time-domain data is equivalent to the inversion of a single component of frequency-domain data.

The choice of a domain for inversion allows one to apply specific methodologies to precondition either the data residuals or the gradient in ways that may improve the convergence and/or the linearity of the inverse problem. For example, time windowing the residuals is often useful (Shipp and

Published on Geophysics Online November 26, 2003. Manuscript received by the Editor January 21, 2003; revised manuscript received April 11, 2003.

*Ecole Normale Supérieure de Paris, Laboratoire de Géologie, Paris, France, and Compagnie Générale de Géophysique, Research and Development, Massy, France. E-mail: lsirgue@free.fr.

†Queen's University, Department of Geological Sciences and Geological Engineering, Kingston, Ontario K7L 3N6, Canada. E-mail: pratt@geol.queensu.ca.

© 2004 Society of Exploration Geophysicists. All rights reserved.

Singh, 2002); this requires a time-domain representation of the data residuals. [It is nevertheless possible to devise strategies for time-domain conditioning of the data by making use of complex-valued frequencies, as in Mallick and Frazer (1987)]. Conversely, when using a frequency-domain approach it is straightforward to carry out the inversion adopting a strategy that proceeds sequentially from the low to the high frequencies (at no additional computational cost). Since low-frequency data are more linear with respect to the model perturbations than high-frequency data, such a strategy dramatically improves the chance for the inversion to locate the global minimum of the full, wide-band inverse problem. Such a methodology may of course also be applied in the time domain by low-pass filtering of the data (Bunks et al., 1995), however; the computational efficiency of the frequency-domain approach is far superior as the propagation of individual frequencies is far cheaper than full time-domain modeling (Marfurt, 1984).

One of the most important advantages of frequency-domain inversion is the ability to provide a unaliased image using a very limited number of frequencies, as observed by Pratt and Worthington (1988), Liao and McMechan (1996), and Forgues et al. (1998). This observation is consistent with the results of Wu and Toksöz (1987), who demonstrated that even a single frequency of a given seismic survey geometry yields information on a finite portion of the model in the wavenumber domain. Since the computational cost of frequency-domain inversion is directly proportional to the number of frequencies used, this is a significant advantage for waveform inversion of real, multishot seismic data.

The claim that a limited number of frequencies would suffice when inverting reflection seismic data was investigated by Freudenreich and Singh (2000), who pointed out that no strategy for selecting the appropriate frequencies had yet been developed. They concluded, incorrectly we believe, that frequency-domain inversion fails in the presence of a limited offset range.

In this paper, we show that frequency-domain inversion of reflection data using only a few frequencies (properly selected) can yield a result that is comparable to full-time domain inversion, for any available range of offsets. We develop a strategy that adapts the selection of imaging frequencies to the available offset range. The main idea is that the larger the offset, the fewer frequencies are required. This strategy takes advantage of the effect of an “image stretch”, which increases with increasing offset. This effect is commonly observed in prestack depth migration (Gardner et al., 1974). In a 1D earth, image stretch is equivalent to normal-moveout (NMO) stretch. Stretch is often considered to have a negative impact on seismic imaging, although Haldorsen and Farmer (1989) showed that the stretch effect can be used to compensate for the lack of low frequencies in the source and hence to improve the spectral content of stacked data. In our strategy, we select each successive frequency by evaluating the image stretch at the maximum offset for the previous frequency. Although we develop the strategy using a 1D assumption, we show in this paper that the same approach can also be applied effectively in 2D, heterogeneous media. Because of the kinematic equivalence between the first iteration gradient vector and prestack depth imaging for reflection data (Tarantola, 1986), the strategy also has important implications for frequency domain implementations of prestack

depth migration [such as that advocated by Schleicher et al. (1993)].

Our paper begins by restating the basic method of gradient-based waveform inversion in the frequency domain. We then derive a specific formula for the gradient using a plane-wave approximation in homogeneous background media that predicts analytically the wavenumber illumination of a given target. This allows us to define a strategy for choosing frequencies for continuous wavenumber illumination of 1D target. The key idea is to take advantage of the maximum offset data in the acquisition. We then validate this strategy on the 1D imaging experiment of Freudenreich and Singh (2000). We go on in the paper to show that the strategy can be successfully applied to data generated in the 2D heterogeneous Marmousi model.

A table of mathematical symbols is presented in Table 1.

METHODS

Waveform inversion

In this section, we briefly restate the principles of nonlinear waveform inversion using iterative gradient (steepest descent) methods, implemented specifically in the frequency domain (Pratt and Worthington, 1990). We use a constant-density acoustic-wave equation as our forward model, given by

$$\left(\nabla^2 + \frac{\omega^2}{c^2(\mathbf{x})}\right)\Psi(\mathbf{x}, \mathbf{s}, \omega) = -\delta(\mathbf{x} - \mathbf{s}), \quad (1)$$

which is nonlinear in the model parameter $m(\mathbf{x}) = 1/c^2(\mathbf{x})$, the inverse of the squared velocity. The forward solution, $\Psi(\mathbf{x}, \mathbf{s}, \omega)$ is the pressure field at the spatial location \mathbf{x} , arising from a disturbance at the source location \mathbf{s} . The dependence of the solutions on the temporal circular frequency ω will be implicit from now on. We define the data residuals for a single frequency component of a given source-receiver pair, with spatial coordinates \mathbf{s} and \mathbf{r} , as

$$\Delta\Psi(\mathbf{r}, \mathbf{s}) = \Psi_{calc}(\mathbf{r}, \mathbf{s}) - \Psi_{obs}(\mathbf{r}, \mathbf{s}), \quad (2)$$

where the complex quantities $\Psi_{calc}(\mathbf{r}, \mathbf{s})$ and $\Psi_{obs}(\mathbf{r}, \mathbf{s})$ are, respectively, the calculated and observed data for a particular frequency. Note that no assumption of a linear relation between data residuals and model parameters is made [i.e., we avoid the use of any Born approximation in the forward problem, as is the case for example in Clayton and Stolt (1981) and in Tarantola (1984b)]. The calculated data are always computed by solving the nonlinear forward problem of equation (1) (we use a finite-difference approach).

The inversion scheme is based on the minimization of the misfit, defined here as the L_2 norm of the data residuals

$$E = \frac{1}{2} \sum_s \sum_r \Delta\Psi^*(\mathbf{r}, \mathbf{s}) \Delta\Psi(\mathbf{r}, \mathbf{s}), \quad (3)$$

where the superscript $*$ denotes complex conjugation. Minimization of equation (3) is approached through the computation of the negative gradient, the direction of decrease of the misfit function with respect to the variation in model parameters. If the model is represented by $m(\mathbf{x})$, that is, as a function of spatial position \mathbf{x} , then the descent direction (the negative

gradient) is defined as

$$g(\mathbf{x}) = -\nabla_m E = -\frac{\partial E}{\partial m(\mathbf{x})}. \quad (4)$$

Since the descent vector is a function of model position \mathbf{x} , it may be considered to be an image of the model. The model is then updated at iteration ℓ in the direction of the descent direction with an appropriate step-length γ :

$$m(\mathbf{x})^{\ell+1} = m(\mathbf{x})^\ell + \gamma^\ell g(\mathbf{x})^\ell. \quad (5)$$

The step length is usually computed using a linear estimate (Tarantola, 1984a; Mora, 1987), although line search methods are also possible. Iterations are repeated until some chosen convergence criteria are met. The method of conjugate gradients (Concus et al., 1976) is usually applied to speed convergence.

The steps by which the gradient may be computed without explicitly calculating any partial derivatives of the data are well established in the literature (Lailly, 1983; Tarantola, 1987; Pratt and Worthington, 1990; Pratt et al., 1998). The fundamental (time-domain) result is that the gradient may be computed via a zero-lag correlation of the forward propagated wavefield (from the source point) and the back-propagated wavefield of the data residuals from the receiver positions. In the frequency domain, the descent direction is given by the multiplication of two wavefields:

$$g(\mathbf{x}) = -\omega^2 \sum_s \sum_r \text{Re}\{P_f^*(\mathbf{x}, \mathbf{s}) P_b(\mathbf{x}, \mathbf{r}, \mathbf{s})\} \quad (6)$$

[see Pratt et al. (1996), their equation (12)]. In equation (6) the forward propagated wavefield $P_f(\mathbf{x}, \mathbf{s})$, assuming an impulsive

point source of unit amplitude and zero phase is

$$P_f(\mathbf{x}, \mathbf{s}) = G_o(\mathbf{x}, \mathbf{s}), \quad (7)$$

and the back-propagated wavefield $P_b(\mathbf{x}, \mathbf{r}, \mathbf{s})$ from the receiver location is given by

$$P_b(\mathbf{x}, \mathbf{r}, \mathbf{s}) = G_o^*(\mathbf{x}, \mathbf{r}) \Delta\Psi(\mathbf{r}, \mathbf{s}), \quad (8)$$

where $G_o(\mathbf{x}, \mathbf{s})$ and $G_o(\mathbf{x}, \mathbf{r})$ are the Green's functions for an excitation at the source and receiver positions in the current model (or "reference medium"). The full expression of the gradient vector is then

$$g(\mathbf{x}) = -\omega^2 \sum_s \sum_r \text{Re}\{G_o^*(\mathbf{x}, \mathbf{s}) G_o^*(\mathbf{x}, \mathbf{r}) \Delta\Psi(\mathbf{r}, \mathbf{s})\}. \quad (9)$$

Equation (9) may also be considered as an operator applied to data residuals. This operator is the adjoint of the Fréchet partial derivative and has been termed a "wavepath" by Woodward (1992).

Gradient vector and wavenumber illumination

Up to this point, our formulation is applicable to arbitrary reference media: equation (9) may be used in cases where the background media is inhomogeneous, provided the Green's functions are calculated appropriately. In order to proceed, we now introduce the following assumptions with respect to the Green's functions:

- 1) Amplitude effects may be ignored.
- 2) The reference medium is homogeneous, with a velocity c_o .

Table 1. Symbols.

Symbol	Type	Unit	Description
$\Psi_{obs/calc}(\mathbf{r}, \mathbf{s}, \omega)$	Complex	-	Observed/calculated wavefield for a given source-receiver pair
$\Delta\Psi(\mathbf{r}, \mathbf{s}, \omega)$	Complex	-	Data residuals for a given source-receiver pair
\mathbf{x}	Real	m	Position vector in the model
\mathbf{r}, \mathbf{s}	Real	m	Receiver and source location vectors
E	Real	-	Misfit function
$c(\mathbf{x})$	Real	m/s	Velocity field
$c_o(\mathbf{x})$	Real	m/s	Reference velocity field
$m(\mathbf{x}) = \frac{1}{c^2(\mathbf{x})}$	Real	s ² /m ²	Squared slowness model parameter
$m(\mathbf{x}) = \frac{1}{c^2(\mathbf{x})} - \frac{1}{c_o^2(\mathbf{x})}$	Real	s ² /m ²	Perturbation to squared slowness
$g(\mathbf{x}) = -\nabla_m E$	Real	-	Descent vector
γ	Real	-	Step length of the model update
$G_o(\mathbf{x}, \mathbf{s}), G_o(\mathbf{x}, \mathbf{r})$	Complex	-	Source and receiver Green's functions
$P_b(\mathbf{x}, \mathbf{r}, \mathbf{s})$	Complex	-	Backpropagated field for a given source-receiver pair
$P_f(\mathbf{x}, \mathbf{s})$	Complex	-	Forward propagated wavefield from the source
f	Real	Hz	Temporal frequency
Δf_n	Real	Hz	n th frequency interval in the discretization strategy
Δf_{st}	Real	Hz	Sampling theorem frequency interval
$\omega = 2\pi f$	Real	rad/s	Circular frequency
$k_o = \omega/c_o$	Real	rad/m	Background wavenumber
k_x, k_z	Real	rad/m	Components in the wavenumber domain for the image and model.
$\hat{\mathbf{s}}, \hat{\mathbf{r}}$	Real	-	Source-to-scatterer and receiver-to-scatterer unit vectors
ϕ/θ	Real	rad	Incident/scattered angle with respect to vertical
h, z	Real	m	Half offset, target depth
R	Real	-	Half offset-to-depth ratio
$\alpha = \cos \phi$	Real	-	Incidence/scattering aperture
Δk_z	Real	rad/m	Wavenumber coverage
t, τ	Real	s	Time and zero-offset time of a reflection (used in the Appendix)

- 3) We are in the far field, so that we may replace the Green's functions with plane-wave approximations.

Note that these assumptions are introduced only in order to proceed to a frequency selection strategy; none are required for the actual inversions. Under the three assumptions, both Green's functions, $G_o(\mathbf{x}, \mathbf{s})$ and $G_o(\mathbf{x}, \mathbf{r})$, may be approximated by incident and scattered plane waves:

$$\begin{aligned} G_o(\mathbf{x}, \mathbf{s}) &\approx \exp(ik_o \hat{\mathbf{s}} \cdot \mathbf{x}), \\ G_o(\mathbf{x}, \mathbf{r}) &\approx \exp(ik_o \hat{\mathbf{r}} \cdot \mathbf{x}), \end{aligned} \quad (10)$$

where $k_o = \omega/c_o$ is the wavenumber of the incident and scattered waves in the homogeneous reference medium, and $\hat{\mathbf{s}}$ and $\hat{\mathbf{r}}$ are, respectively, unit vectors in the incident propagation direction (source to scatterer) and the inverse scattering direction (receiver to scatterer). Inserting equation (10) into equation (9) yields

$$\begin{aligned} g(\mathbf{x}) &= -\omega^2 \sum_s \sum_r \text{Re}\{\exp(-ik_o \hat{\mathbf{s}} \cdot \mathbf{x}) \\ &\quad \times \exp(-ik_o \hat{\mathbf{r}} \cdot \mathbf{x}) \Delta\Psi(\mathbf{r}, \mathbf{s})\} \\ &= -\omega^2 \sum_s \sum_r \text{Re}\{\exp(-ik_o (\hat{\mathbf{s}} + \hat{\mathbf{r}}) \cdot \mathbf{x}) \Delta\Psi(\mathbf{r}, \mathbf{s})\}. \end{aligned} \quad (11)$$

The only space-dependent term in equation (11) is the complex exponential, which, for a given source-receiver pair, oscillates at a single wavenumber given by the vector $k_o(\hat{\mathbf{s}} + \hat{\mathbf{r}})$. Equation (11) is therefore an inverse Fourier summation in which the weight function for each source/receiver pair plane-wave combination is determined by the data residuals $\Delta\Psi(\mathbf{r}, \mathbf{s})$.

A single-frequency, single-source-receiver pair in the imaging algorithm described by equation (11) would therefore yield only a single wavenumber in the image, $g(\mathbf{x})$ (see Figure 1). In order to reconstruct useful images, we need to provide additional wavenumbers. There are two ways to recover a range of wavenumbers: (1) we can use a wider band of frequencies (as we do if we use a time-domain algorithm), or (2) we can use a range of different source-receiver pairs that sample the same image point from different directions, $\hat{\mathbf{r}}$ and $\hat{\mathbf{s}}$. A combination of both approaches seems sensible. Almost all seismic imaging strategies take advantage of the first approach: the wavenumber bandwidth of the image is provided by temporal bandwidth in the experiment. Our suggestion is that some of the required image bandwidth can in fact be provided by making optimal use of different offsets in the reflection experiment to create a variety of plane-wave imaging directions. Temporal frequencies may then be selected to avoid or reduce redundancy of the wavenumber information.

Analysis of the gradient within the Born approximation

Let us now examine the gradient image of equation (11) by estimating the data residuals $\Delta\Psi(\mathbf{r}, \mathbf{s})$ in the context of the Born approximation, through which we may write

$$\Delta\Psi(\mathbf{r}, \mathbf{s}) \approx -\omega^2 \int d\mathbf{x} G_o(\mathbf{r}, \mathbf{x}) G_o(\mathbf{x}, \mathbf{s}) \delta m(\mathbf{x}), \quad (12)$$

where $\delta m(\mathbf{x})$ is the true perturbation in the parameter $m(\mathbf{x}) = 1/c^2(\mathbf{x})$ [see, for example, Miller et al. (1987), their equation (8)]. Equation (12) approximates the data residuals by establishing a linear relation with the true model perturbation δm .

Although in our inversion scheme, the data residuals are computed using the nonlinear forward problem, this linear approximation will allow us to establish an insightful link between the gradient and the true model perturbation. Note that the Born approximation is valid if the starting model is close enough to the true model. Substituting the plane-wave approximations of equation (10) into equation (12) we obtain

$$\Delta\Psi(\mathbf{r}, \mathbf{s}) \approx -\omega^2 \int d\mathbf{x} \delta m(\mathbf{x}) \exp(+ik_o (\hat{\mathbf{s}} + \hat{\mathbf{r}}) \cdot \mathbf{x}), \quad (13)$$

which is in the form of a forward Fourier transform, so that we may write

$$\Delta\Psi(\mathbf{r}, \mathbf{s}) = -\omega^2 \tilde{M}(k_o(\hat{\mathbf{s}} + \hat{\mathbf{r}})), \quad (14)$$

where $\tilde{M}(\mathbf{k})$ is the Fourier transform of the function $\delta m(\mathbf{x})$. The fact that the monochromatic plane-wave scattering response is sensitive only to a single Fourier component of the target is a well-established result that goes back at least to Ewald (1921). It is the basis of the reciprocal lattice method in X-ray crystallography (Cullity, 1978). A restatement of this basic principle may be found in Wu and Toksöz (1987).

Substituting the Born estimate of the data residuals, equation (14), into the plane-wave gradient-image equation (11), we obtain

$$g(\mathbf{x}) = \omega^4 \sum_s \sum_r \text{Re}\{\exp(-ik_o (\hat{\mathbf{s}} + \hat{\mathbf{r}}) \cdot \mathbf{x}) \tilde{M}(k_o(\hat{\mathbf{s}} + \hat{\mathbf{r}}))\}, \quad (15)$$

which, as stated above in reference to equation (11), is an inverse Fourier summation (apart from the ω^4 factor, which we comment on below). We have now written this summation in a

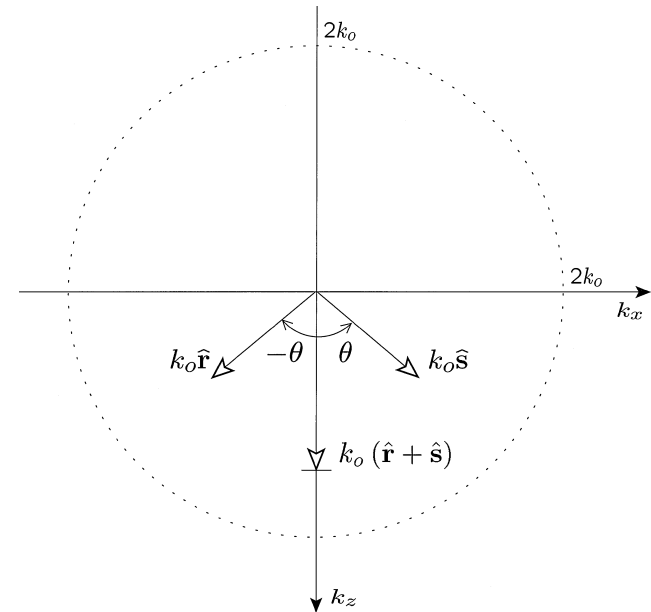


FIG. 1. Wavenumber illumination. A single frequency component from a single source-receiver pair generates a single wave vector in the spectrum of the gradient image. The wave vector illuminated is given by the sum of the source-to-scatterer wave vector, $k_o \hat{\mathbf{s}}$ and the receiver-to-scatterer wave vector $k_o \hat{\mathbf{r}}$. We illustrate here a particular case in which magnitudes of the incident/reflected angle are equal, so that only the vertical components are imaged.

form that makes it clear that the weights in the summation are given by the Fourier components of the model. Thus, the summation will yield an image of the true model, distorted however due to the limited availability of scattering directions. Let us emphasize this fundamental observation:

One source-receiver pair with an incidence direction $\hat{\mathbf{s}}$ and an inverse scattering direction $\hat{\mathbf{r}}$ yields information on only a single wavenumber in the spectrum of the model, given by $k_o(\hat{\mathbf{s}} + \hat{\mathbf{r}})$. Multiple source-receiver pairs yield information on a finite region in wavenumber space, even if only a single frequency component is used. This basic relationship is illustrated in Figure 1. This result is similar to demonstrations made in the context of diffraction tomography (Devaney, 1981; Wu and Toksöz, 1987) and linearized inversion in the (ω, k) domain e.g., Clayton and Stolt, 1981; Ikelle et al., 1986).

A comment on the gradient in the linear case

Equation (15) is in the form of an inverse Fourier summation: if we have a complete and uniform sampling in the source and receiver scattering directions (an impossibility for most geophysical scenarios), then this inverse Fourier summation will undo the forward Fourier transform in $\tilde{M}(\mathbf{k})$, i.e.,

$$g(\mathbf{x}) \rightarrow \omega^4 \delta m(\mathbf{x}), \quad (16)$$

in which case we see that the gradient will recover a scaled image of the original model. Equation (16) illustrates the limitation of the gradient as an imaging operator: the factor ω^2 in the forward problem is not canceled by the gradient operator, instead it is reused. A correct inverse modeling operator would cancel this term. In diffraction tomography and linearized inversion, the inverse operator contains a filter term inversely proportional to ω^2 , thus ensuring the fidelity of the reconstruction (Wu and Toksöz, 1987). In descent schemes, an iterative approach should ensure that the distortions in the gradient are eventually removed.

In multifrequency or time-domain inversions, the factor ω^4 implies a double application of the second time-derivative operator. For this reason, time-domain schemes are not very good at recovering the very low wavenumbers, which are suppressed by the derivative operators. Iterations may eventually recover some of the low wavenumbers but, without preconditioning in some form, the convergence rate will be slow. Furthermore, since the low frequencies have a more linear relation with the low-wavenumber components of the model, the chance of locating the global minimum will be diminished if this information is not used at an early stage of the inversion. This effect can be compensated for by low-pass filtering of the data in time (Bunks et al., 1995), and progressively including higher frequencies at a later stage. However, if the inversion only involves a few frequencies, it is always more efficient to carry this out in the frequency domain (Marfurt, 1984; Jo et al., 1996). In single frequency inversions the factor ω^4 is of no consequence, as the step length used in equation (5) will compensate for this factor correctly by rescaling the gradient.

The 1D Case

Having reviewed the basic relationship between the wavenumber spectrum of the image and the incident and scattered wave directions, we now wish to use this relationship

to develop a strategy for the optimal selection of frequencies in frequency-domain imaging of seismic reflection data in one dimension. Our fundamental assumption here will be that a general strategy is developed by examining the illumination of a 1D model (i.e., one in which the velocity varies only as a function of depth), for which we need only to recover the vertical wavenumbers from a range of source-receiver offsets. At a later stage, we test the strategy on a model with significant 2D structure in order to evaluate the generality of the approach.

Wavenumber illumination for the 1D case

As shown previously using a plane-wave approximation, the contribution to the gradient image of a single source-receiver pair has only a single wavenumber component, given by $k_o(\hat{\mathbf{s}} + \hat{\mathbf{r}})$. If the model is 1D, then all reflection points in the subsurface are midpoint reflection points. In other words, we may consider a common-shot gather to be equivalent to a common-midpoint (CMP) gather. The basic configuration is illustrated in Figure 2: each source-receiver pair records data from a given midpoint with a particular combination of symmetric plane waves (the equivalent wavenumber diagram for this configuration appears in Figure 1). For a 1D earth, the incident and scattering angles are symmetric, so that

$$k_o \hat{\mathbf{s}} = (k_o \sin \theta, k_o \cos \theta) \quad \text{and} \quad k_o \hat{\mathbf{r}} = (-k_o \sin \theta, k_o \cos \theta). \quad (17)$$

In equation (17), the angles θ and $-\theta$ are the propagation directions of the source and receiver wave vectors, defined with respect to the vertical axis. From Figure 2, we observe that

$$\begin{aligned} \cos \theta &= \frac{z}{\sqrt{h^2 + z^2}} \\ \sin \theta &= \frac{h}{\sqrt{h^2 + z^2}}, \end{aligned} \quad (18)$$

where h is the half offset and z is the depth of the scattering layer. Substituting expression (18) into equation (17), we find that the components of the vector $k_o(\hat{\mathbf{s}} + \hat{\mathbf{r}})$ are

$$\begin{aligned} k_x &= 0 \\ k_z &= 2k_o \alpha, \end{aligned} \quad (19)$$

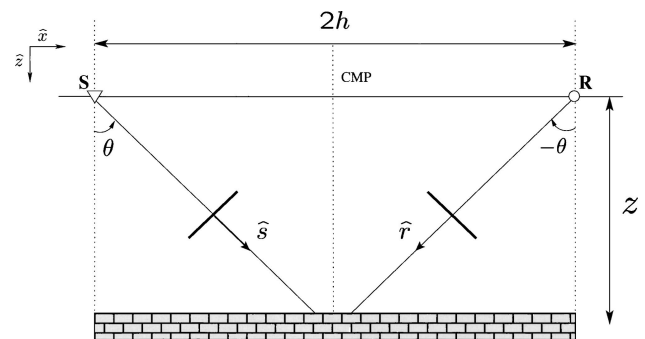


FIG. 2. The 1D basic scattering experiment. The incident plane wave is reflected from a 1D thin layer at the midpoint between source and receiver. For the 1D case, the incident and reflected angles, θ and $-\theta$ are equal.

with

$$\begin{aligned}\alpha &= \cos \theta \\ &= \frac{z}{\sqrt{h^2 + z^2}} \\ &= \frac{1}{\sqrt{1 + R^2}},\end{aligned}\quad (20)$$

where $R = h/z$ is the half offset-to-depth ratio.

Equation (19) defines the wavenumber illumination of a given 1D layer for a given source-receiver pair. The illumination varies with the wavenumber $k_o = \omega/c_o$ and with the factor α , which in turn depends on the half offset-to-depth ratio R of the source-receiver pair relative to the target. The horizontal wavenumber of the gradient is always zero, as we expect for a 1D model. Each source-receiver pair contributes a different vertical wavenumber to the gradient image; the larger the offset-to-depth ratio, the smaller is the vertical wavenumber of the contribution. Thus, for a given frequency and offset, the vertical wavenumber recovered will increase with depth.

STRATEGY FOR CHOOSING FREQUENCIES FOR 1D IMAGING

In the preceding sections, we clarified the wavenumber spectral coverage for a given source-receiver pair at a single frequency over a 1D earth. Using equations (19) and (20), but now for a range of offsets, we find that for a given surface seismic acquisition characterized by an offset range $[0, x_{max}]$, the vertical wavenumber coverage k_z of a 1D thin layer for a given frequency is limited to the range $[k_{zmin}, k_{zmax}]$, where

$$\begin{aligned}k_{zmin} &= 2k_o\alpha_{min}, \\ k_{zmax} &= 2k_o,\end{aligned}\quad (21)$$

with

$$\alpha_{min} = \frac{1}{\sqrt{1 + R_{max}^2}},\quad (22)$$

where $R_{max} = h_{max}/z$ is the half offset-to-depth ratio obtained at the maximum half offset h_{max} , and z is the depth of the target layer. The minimum and maximum wavenumbers k_{zmin} and k_{zmax} are produced by the furthest and nearest offsets, respectively. Expressing equation (21) in terms of frequency, we have

$$\begin{aligned}k_{zmin} &= 4\pi f\alpha_{min}/c_o, \\ k_{zmax} &= 4\pi f/c_o,\end{aligned}\quad (23)$$

where f is the frequency expressed in hertz and c_o is the velocity in the background medium. From relation (23), we may define the wavenumber coverage of a multioffset acquisition as

$$\begin{aligned}\Delta k_z &\equiv |k_{zmax} - k_{zmin}| \\ &= 4\pi(1 - \alpha_{min})f/c_o,\end{aligned}\quad (24)$$

while the bandwidth

$$\frac{k_{zmax}}{k_{zmin}} = \frac{1}{\alpha_{min}} = \sqrt{1 + R_{max}^2}.\quad (25)$$

The wavenumber coverage increases linearly with frequency and is therefore greater for high than for low frequencies (Wu

and Toksöz, 1987), while the wavenumber bandwidth is a function only of the offset-to-depth ratio.

Our strategy for choosing frequencies is determined as follows: each frequency has a limited, finite-band contribution to the image spectrum. In order to recover the target accurately over a broad range of wavenumbers, the continuity of the coverage of the object in the wavenumber domain must be preserved as the imaging frequencies are selected. We choose

$$k_{zmin}(f_{n+1}) = k_{zmax}(f_n),\quad (26)$$

where f_{n+1} is the next frequency to be chosen following the frequency f_n . The principle (illustrated in Figure 3) is that the maximum wavenumber of the smaller frequency must equal the minimum wavenumber of the larger frequency. This strategy thus relies on using the full range of offsets available.

Using the condition defined in equation (26) and substituting equation (23), we arrive at the relation

$$f_{n+1} = \frac{f_n}{\alpha_{min}}.\quad (27)$$

This leads us to a discretization law in which the frequency increment Δf_{n+1} is given by

$$\begin{aligned}\Delta f_{n+1} &= f_{n+1} - f_n \\ &= \left(\frac{1 - \alpha_{min}}{\alpha_{min}}\right)f_n \\ &= (1 - \alpha_{min})f_{n+1}.\end{aligned}\quad (28)$$

Equation (28) shows that the optimum frequency increment is not constant and increases linearly with frequency. This is an interesting result as the commonly used frequency-domain sampling theorem sets the frequency increment to be constant and equal to

$$\Delta f_{st} = \frac{1}{T_{max}},\quad (29)$$

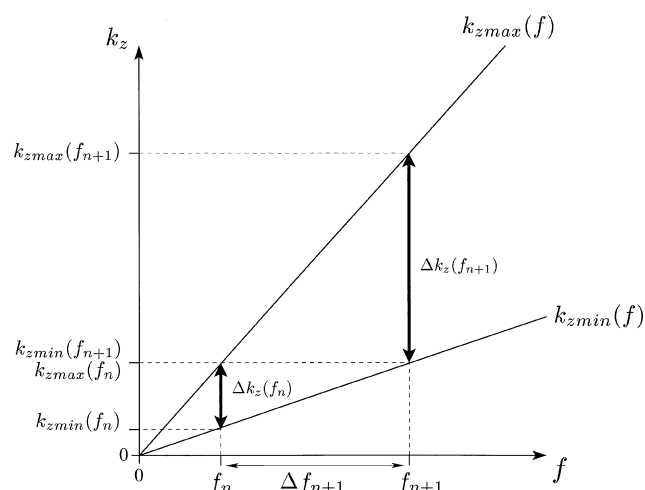


FIG. 3. Illustration of the frequency discretization strategy. A single discrete frequency yields a range of vertical wavenumbers in the image (governed by the maximum and minimum offsets in the survey). The next frequency is selected in such a way as to obtain a continuous coverage in vertical wavenumbers.

where T_{max} is the maximum recorded time. Equation (29) prevents “time-aliasing” (i.e., wrap-around in the time domain due to insufficient frequency sampling). Since Δf_n in equation (28) is typically bigger than Δf_{st} in equation (29), the strategy we are proposing is a discretization that would produce time aliasing of a single seismic trace in time, but will yield an unaliased image of the target in depth from a set of seismic traces with a range of offsets.

NUMERICAL TEST I: 1D MODELS

The 1D experiment we use in this section (see Figure 4) was originally used by Freudenreich and Singh (2000), who applied a time-domain waveform inversion method (Shipp et al., 1997) to evaluate and compare the time- and frequency-domain inversion approaches. (Freudenreich and Singh simulated a frequency-domain approach by inverting a single sinusoidal component of the time-domain data.) They observed that when they included far offsets (out to 10 km), their frequency-domain approach using three discrete frequencies did indeed provide an accurate imaging result equivalent to their time-domain inversion. However, when attempting waveform inversion for the same three frequencies, but using only a near-offset range (limited to 3 km), their inversion result suffered from oscillatory artifacts due to the limited set of input frequency components. They concluded that when only near offsets were available, the frequency-domain approach was less robust.

Freudenreich and Singh raised the important question of exactly how the inversion frequencies should be chosen. In this section, we use our strategy on the same experiment and thus provide an answer to this question.

The synthetic data

The data for the 1D experiment (Figure 4) consist of a single shot gather comprising 201 receivers, with a receiver interval of 50 m. The maximum offset range available is [0, 10] km. The target of the imaging experiment is the thin layer located at 2-km depth. The velocity model is exactly 1D, and has

thus only wavenumber components along the vertical k_z axis. For all inversion results we present below, the starting model was identical to this model, but without the thin layer. All the modeling and inversion tests were performed with absorbing boundary at the top of the model, and using a point source with a Ricker wavelet centered on 4 Hz. The time-domain seismograms shown in Figure 4 are synthesized from the frequency-domain finite-difference results, using 50 frequencies within a range [0.2–10] Hz (thus representing a maximum recording time of 5 s).

Validation of the plane-wave predictions

In the previous section, we quantified the wavenumber coverage k_z of a single source/receiver pair within the plane-wave assumption for a 1D model [equation (19)]. In this section, we generate the gradient (descent direction) using equation (9) with point-source Green's functions and test the results against equation (19).

Figure 5 depicts the first stage in the numerical computation of the gradient at 5 Hz using equation (9) for two individual source-receiver pairs (at 3-km and 10-km offsets). The 5-Hz Green's functions were calculated by the method of frequency-domain finite differences (Pratt and Worthington, 1990; Jo et al., 1996). The background model is as described above (i.e., as in Figure 4a), without the thin layer. The figure shows the real parts of the forward and back propagated wavefields and the gradient, the product $Re\{P_f^*(\mathbf{x}, \mathbf{s})P_b(\mathbf{x}, \mathbf{r}, \mathbf{s})\}$ [see equation (6)]. Note that to compute the back-propagated wavefield $P_b(\mathbf{x}, \mathbf{r}, \mathbf{s})$, we first computed the data residuals $\Delta\Psi(\mathbf{r}, \mathbf{s})$ for the experiment by subtraction of the data generated with and without the thin layer.

The fundamental properties of the image in 1D media (noted previously) may also be noted on Figure 5 by examining the wavepaths at the midpoint locations:

- 1) At the midpoint location, the gradient image contains only vertical wavenumbers.

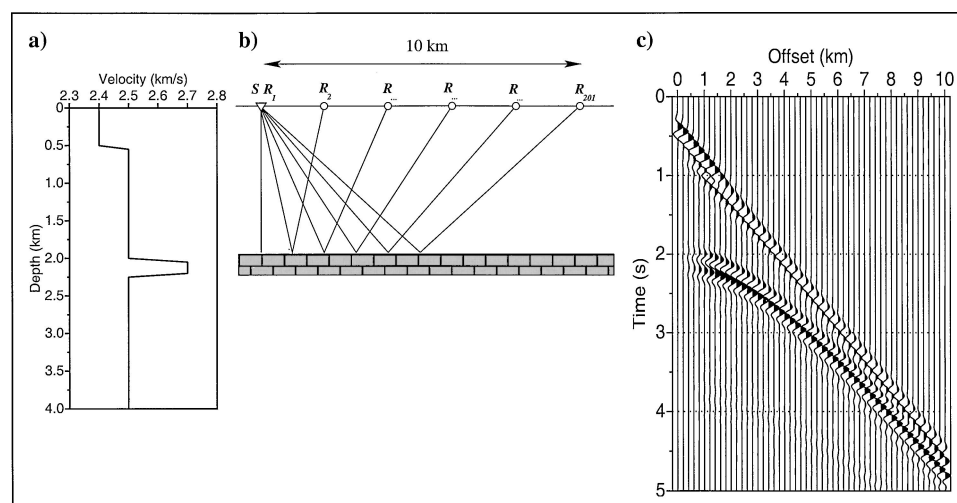


FIG. 4. Velocity model after Freudenreich and Singh (2000). (a) Velocity profile of the 1D model. The target is located at 2 km depth. (b) Acquisition geometry of the 1D basic imaging experiment with a single source and 201 receivers. (c) Synthetic, time-domain data from the model, using a Ricker wavelet with a peak frequency at 4 Hz and with 400 ms of time delay.

- 2) Large offsets generate low vertical wavenumbers and near offsets generate high wavenumbers.
- 3) The wavenumber of the reconstruction increases with decreasing offset-to-depth ratio (i.e., the deeper in the model we go, the higher the wavenumber for a given offset).

A full reconstruction is provided by combining gradient images such as those in Figure 5 for all source-receiver pairs (i.e., all offsets). The reconstructed images for 2 Hz and for 5 Hz are illustrated in Figure 6a. Because of the equivalence between shot gathers and midpoint gathers for 1D models, the images may be considered to be equivalent to a set of individual 1D images (or common-image-point gathers), each obtained from a midpoint gather with a different half-offset h . Each result is a 2D image which manifests the expected decrease in vertical wavenumber as the offset increases.

Figure 6b shows the Fourier transform of the gradient images in depth, with an overlay representing the predicted wavenumber coverage from equation (19). The comparison confirms that the plane-wave prediction is an adequate representation

of the wavenumber coverage in the far field and with a homogeneous background velocity model. This shows that the analytic equations developed within the plane-wave assumption may be safely used in the development of the frequency-selection strategy.

A remark on “image stretch”

Equation (19) predicts a decrease in vertical wavenumber content of the gradient image with offset; this effect is evident on Figure 6. We refer to this shift towards low wavenumbers with offset as “image stretch.” The notion that the seismic image from large offsets is lower in resolution has long been observed (Buchholtz, 1972; Dunkin and Levin, 1973), this effect is usually termed “NMO stretch” (see, for example Yilmaz, 1987). As we show in the Appendix, the stretch factor $1/\alpha$ in equation (19) is identical to that caused by NMO stretch. If the processes of NMO correction, stacking, and depth conversion are applied, NMO stretch will, inevitably, add low-wavenumber content to the image in a manner identical to that observed in Figure 6. If these processes are replaced with

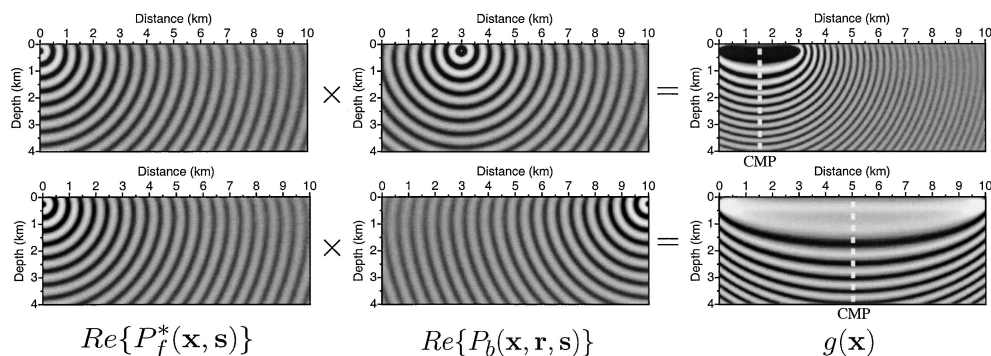


FIG. 5. Illustration of the computation of the 5-Hz gradient [equation (6)] for the model in Figure 4, using two source-receiver pairs at offsets of 3 km (top row) and 10 km (bottom row). These patterns have been referred to as wavepaths by Woodward (1992). The midpoint for each offset pair is marked by vertical dashed line.

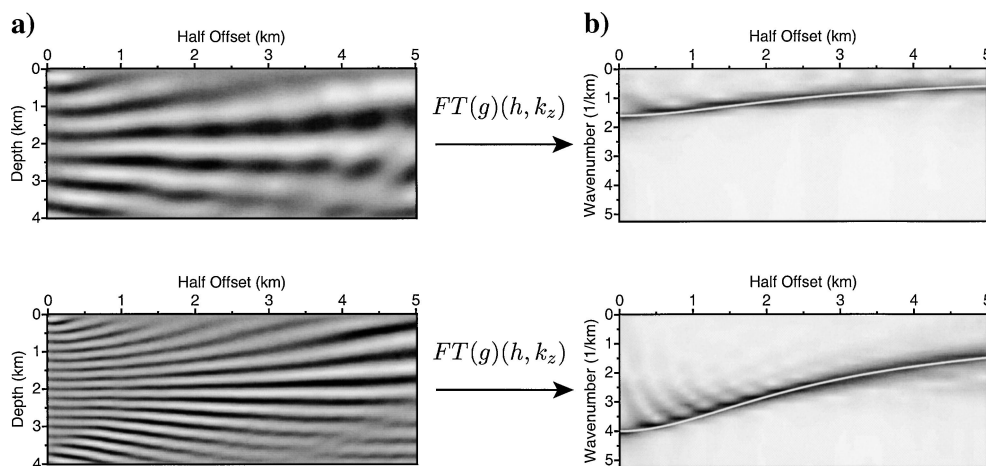


FIG. 6. (a) A set of 1D amplitude normalized gradient images at 2 Hz (top) and 5 Hz (bottom) as a function of halfoffset h for a midpoint gather. The 1-D model used in this test is shown in Figure 4a; the location of the expected anomaly is at 2 km depth. The 1D gradient image at a single frequency for a single source-receiver pair contains only a single vertical wavenumber, and hence is oscillatory. (b) The vertical Fourier transform of the images in (a). The white line is the vertical wavenumber predicted from plane-wave theory in equation (19). The vertical wavenumber decreases with increasing half offset in a manner exactly equivalent to that for NMO stretch.

prestack depth migration, the effect remains (Gardner et al., 1974; Levin, 1998). Usually this is considered detrimental, and data with a large offset-to-depth ratio are muted in standard data processing. An exception to this is the work of Haldorsen and Farmer (1989), who showed that NMO stretch may be used to compensate for a lack of low-frequency content in the data spectrum in order to improve the spectral content of a stack section.

1D waveform inversion

We now describe the application of the frequency discretization strategy developed above to the 1D basic imaging experiment in Figure 4, using two survey geometries: (1) A near-offset survey with a maximum offset of 3 km, and (2) a far-offset survey with a maximum offset of 10 km. Given the depth of the bottom of the scatterer $z = 2.2$ km, equation (27) gives the following sequence of frequencies for the two offset ranges: (1) for the near-offset survey [0, 3] km, $f = 2, 2.4, 2.9, 3.5, 4.3, 5.2, 6.3, 7.6, 8$ Hz; (2) for the far-offset survey [0, 10] km, $f = 2, 5, 8$ Hz. (The last frequency 8 Hz is chosen arbitrarily so that both experiments use identical frequency ranges within the spectrum of the source signature.) The two discretization sequences are illustrated in Figure 7, in which the continuity of the wavenumber coverage is evident. The near-offset survey requires nine frequencies, whereas the far-offset survey only three frequencies. It is implicit in the strategy that the larger the offset range is, the fewer frequencies are required.

The inversion results for each of the surveys are shown Figure 8. In order to complete the imaging of the midpoints, we stacked the full set of common-image-point gradient images (shown in Figure 6) to form a single, 1D image at an equivalent midpoint location for each frequency. In selecting frequencies, we followed the sequence above one frequency at a time (i.e., “sequentially”), iterating 20 times at each discrete frequency. Figure 8 shows both the final results (top row) and the individual contributions for each frequency to the image (bottom row). Our selection of frequencies, although discretized far below the frequency-domain sampling theorem, yields a

continuous wavenumber coverage of the target without aliasing (see the wavenumber domain representations in Figure 8). Each imaging frequency for the far-offset survey clearly yields a wider band of vertical wavenumbers than does the equivalent frequency for the near-offset survey.

The far-offset survey is also more successful in recovering the magnitude of the velocity perturbation. This is due to the fact that the inversion of the first frequency (2 Hz) from the far-offset survey yields information on much lower wavenumbers than does the 2 Hz result from the near-offset survey. This is fundamental, and it is due to the image stretching remarked on above. The lowest wavenumbers are particularly important in recovering the true velocity perturbation.

Time-domain waveform inversion can be simulated using our frequency-domain approach by inverting from all frequency components within the data simultaneously, rather than sequentially. By “all” frequencies, we mean with a complete sampling according to the sampling theorem (29). The gradients for this simultaneous, “time like” inversion procedure is the sum of the single-frequency gradients at each iteration. Figure 9 shows the results of the inversion using the same offset ranges using time-like inversion (i.e., inverting for 50 frequencies simultaneously). The far-offset survey, once more, does a better job of recovering the magnitude of the velocity perturbation (as in frequency-domain inversion, and for the same reasons). Comparing the sequential frequency-domain inversions and the time-like inversions, we note that performing the waveform inversion with a limited number of frequencies according to our strategy yields an acceptable result at a much smaller computational cost, i.e., a factor of 5 fewer frequencies for the near-offset range, and a factor of 16 fewer frequencies for the far-offset range.

Sensitivity to noise

The strategy that we developed suppresses the redundancy in the wavenumber coverage of the 1D target by reducing the number of frequencies given by the sampling theorem. However, some of this redundancy is most desired when noise is present in the seismic data to assure the destructive summation of incoherent signal. In order to investigate the sensitivity to

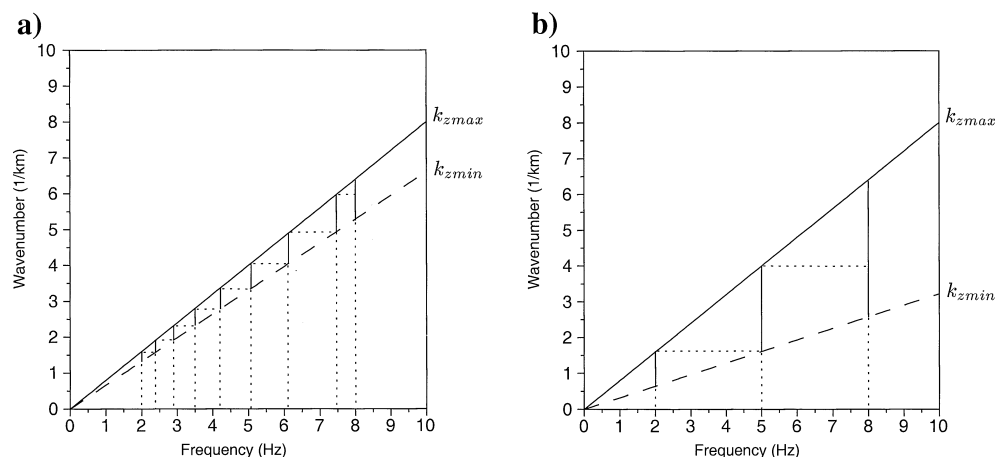


FIG. 7. Frequency sequence generated by equation (27), for (a) the near-offset survey (0–3km) and (b) the far-offset survey (0–10km). The near-offset survey requires nine frequencies, whereas the far-offset survey requires only three frequencies.

noise of our frequency selection strategy, we carry out the 1D inversion after having added white noise to the data using the program *suaddnoise* of the Seismic Un*x package (Stockwell, 1997). The reflection hyperbola of the 1D target with noise corresponding to a signal-to-noise ratio of 1 is shown in Figure 10b and can be compared with the original reflection in Figure 10a. The 1D inversion of the near-offset data (0–3 km) using the nine frequencies dictated by our strategy is shown in Figure 10c. The reconstructed target contains more noise than the equivalent time domain inversion using 50 frequencies (Figure 10e). The noise can nevertheless be attenuated by increasing some of the redundancy using 16 frequencies in the inversion, as shown Figure 10d. The sequence of 16 frequencies was generated using a maximum offset of 2 km in the calculation of α_{min} [equation (22)].

Despite the high level of noise present in the data, the inversion using nine frequencies does yield a good image. The frequency decimation suppresses some of the redundancy, but the stack of each offset contribution to the image does nevertheless allow destructive summation of some of the random noise.

NUMERICAL TEST II: 2D MODELS

Our strategy explicitly ensures the continuity of the wavenumber coverage only along the k_z axis, and only for a

1D target in a homogeneous medium. The applicability of the strategy to 2D, heterogeneous velocity models is therefore uncertain. In the 2D case, other directions in the wavenumber spectrum of the model must be recovered. In this section, we test the strategy using a more demanding 2D application.

Marmousi velocity model

In order to evaluate the applicability of the strategy to 2D models, we test the approach with synthetic seismic data from the 2D Marmousi velocity model (Versteeg, 1994). Waveform inversion has already been applied to data from the Marmousi model using both a time-domain approach (Bunks et al., 1995) and a frequency-domain approach (Forgues et al., 1998). Forgues et al. (1998) performed frequency-domain waveform inversion with a starting velocity model which was a smoothed version of the true model. They also used an arbitrary selection of imaging frequencies. We avoid these two oversimplifications in the test described below.

Synthetic data

A key conclusion of our analysis is that large offsets are important if we want to recover the lowest wavenumbers in the model (which, in turn allow us to recover the true magnitude

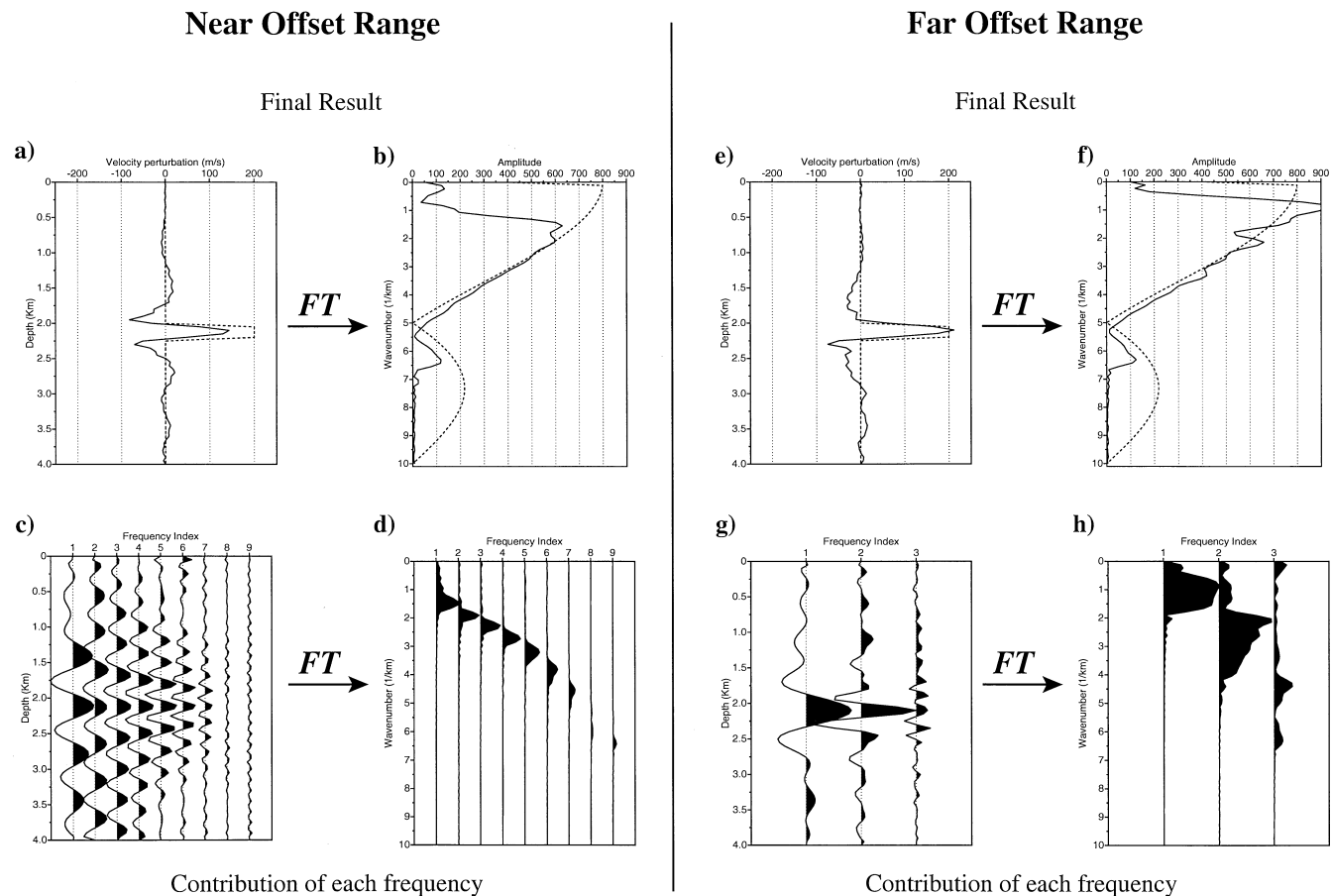


FIG. 8. Sequential frequency domain inversion results for the near-offset survey (a–d) and the far-offset survey (e–h). The true perturbation is shown as a dotted line. The top row shows the final results in both the depth domain (a, e) and the wavenumber domain (b, f), whereas the bottom row shows the contributions at each discrete frequency, both in the depth domain (c, g) and in the wavenumber domain (d, h).

of the velocity anomalies). The original Marmousi experiment was limited in offset to only 3 km. We therefore modified and remodeled the seismic survey in the Marmousi model to create new wide-angle data. The velocity model was restricted to 2 km in depth to assure the effective illumination of the model at large offset-to-depth ratios by the diving and refracted waves (Figure 11). These arrivals will be needed for the determination of the starting model of the waveform inversion (see below). The new synthetic survey comprises 96 sources (every 96 m) and 384 receivers (every 24 m), with a maximum offset of 9.2 km [we effectively simulated the geometry of an ocean-bottom cable (OBC)]. The synthetic data shown in Figure 11 were created using frequency-domain finite differences; time-domain data were extracted from 122 frequencies in the range [0.12–15] Hz by Fourier synthesis. A Ricker source wavelet with a spectrum centered on 5.5 Hz was used in the simulation. The forward modeling and the inversions were all performed with a free surface boundary at the top of the model; our simulations therefore include the effect of free surface multiples.

Determination of the starting model by traveltimes inversion

In order to apply waveform inversion to the wide-angle Marmousi data set, a starting velocity model is required that

must be located in the neighborhood of the global minimum of the objective function for the starting frequency (see our comments in the Introduction). A suitable starting model may be obtained by combining waveform inversion with ray-based traveltimes inversion (Pratt and Goulet, 1991). We used travel time inversion of the hand-picked first arrivals, using the program FAST (Zelt and Barton, 1998). The FAST software uses a finite-difference solver for the eikonal equation to forward model the arrival times (Vidale, 1990; Podvin and Lecomte, 1991). We proceeded by first carrying out an inversion to estimate the 1D medium that best fit the data. We followed this by solving for a strongly smoothed 2D model, and then progressively less smoothed models. Our final traveltimes inversion result is a relatively smooth, 2D velocity structure that fits the picked time with an rms residual of 27 ms (Figure 12a) reasonable level considering the 200-ms period of the central frequency of the source. Finite-difference modeling in this result produces seismic traces that closely match the first arrival waveforms, but contain virtually no reflected waves (Figure 12c).

Waveform inversion

We used the 2D, smooth traveltimes inversion velocity model in Figure 12a, as a starting model for the subsequent waveform

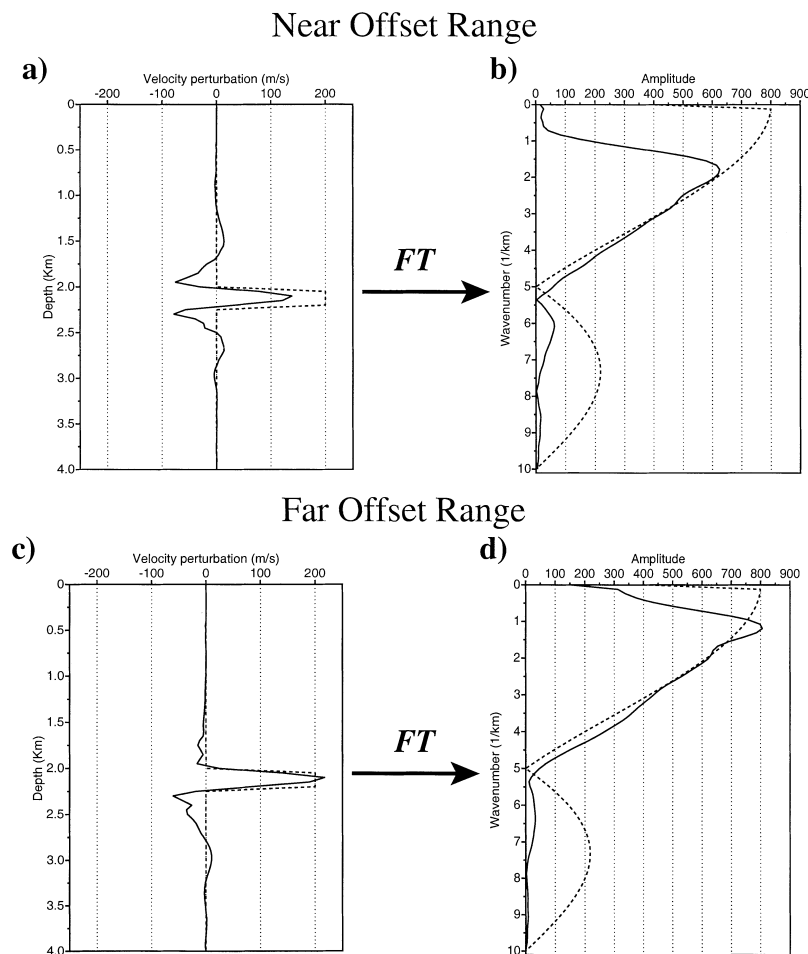


FIG. 9. Timelike waveform inversion of the near-offset range (a, b) and the far-offset range (c, d). The true perturbation shown as a dotted line. Fifty frequencies were inverted simultaneously in order to simulate time-domain waveform inversion. Results are shown in both depth domain (a, c) and wavenumber domain (b, d).

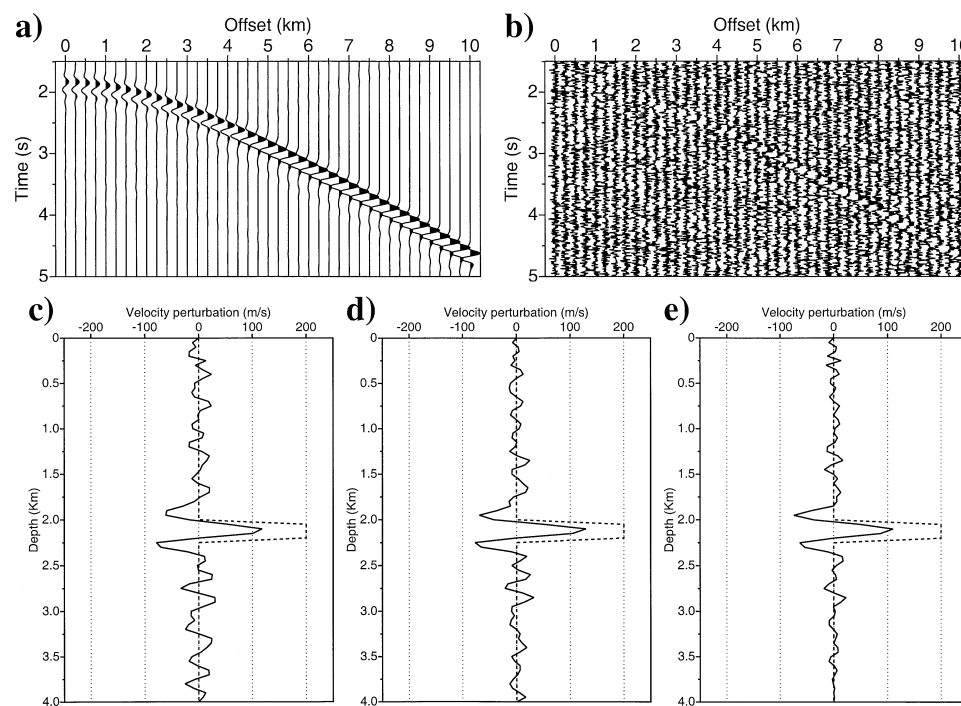


FIG. 10. Impact of noise on the reconstruction of the 1D target for an inversion of the near-offset data (0–3 km), as in Figure 8 and 9. (a) Original reflection hyperbola on the 1D target. (b) Reflection hyperbola with white noise (signal-to-noise ratio=1). (c) Result using 9 frequencies according to our strategy. (d) Result using 16 frequencies obtained with an effective maximum offset of 2 km. (e) Timelike inversion using 50 frequencies.

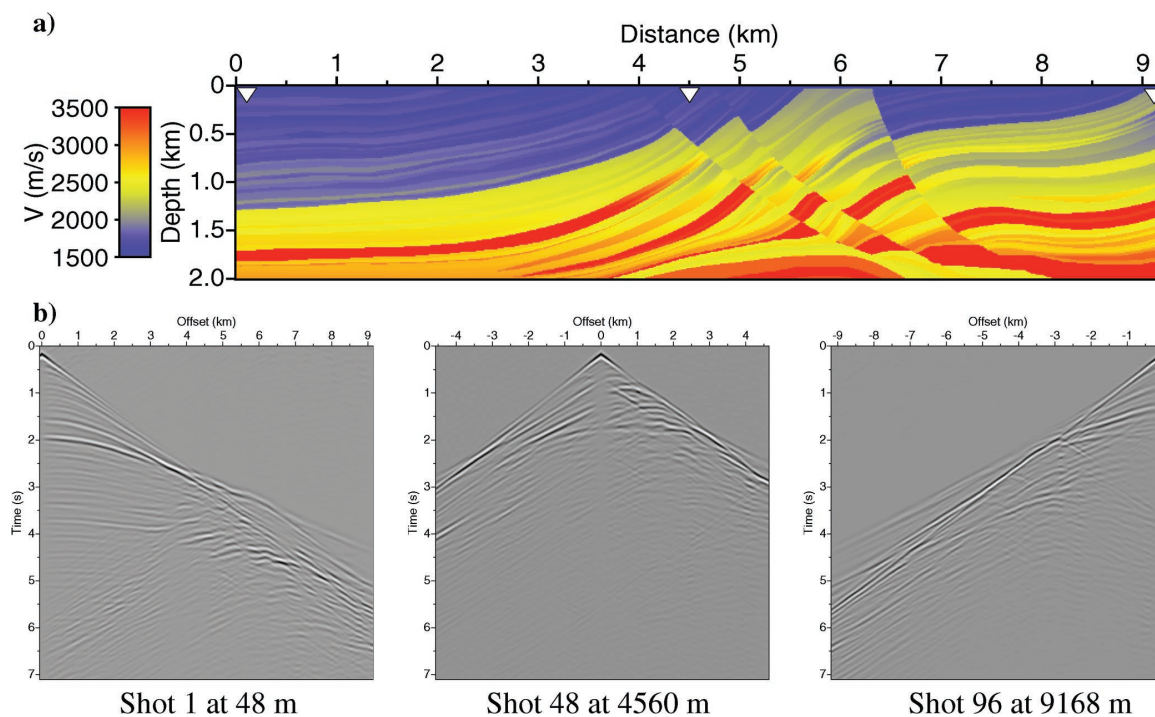


FIG. 11. Marmousi wide-angle synthetic data. (a) The velocity model, unchanged from the original Marmousi model but truncated at 2 km depth. (b) Three representative wide angle shot gathers. The acquisition was expanded to wide offsets by recording data for all shots at all 384 receiver points along the surface (simulating an OBC recording geometry).

inversion. The frequency sequence in the inversion was calculated using equation (27), with a target depth of $z = 2$ km (i.e. the maximum depth in the model) and a maximum offset $x_{max} = 4$ km. Although the maximum offset is in reality 9.2 km, the sequence was calculated with this “effective” offset range because the far offsets are underrepresented in the OBC shooting geometry, and hence these will carry a disproportionately small weight in the inversion. Furthermore, the edges of the model are not illuminated by the widest offsets, especially at depth. Using these parameters, the frequency selection strategy yields three frequencies within the bandwidth of the data: 5, 7, and 10 Hz. Twenty iterations were carried out at each imaging frequency sequentially from low to high frequency. The minimum frequency (5 Hz) was chosen so that the resolution limit imposed by traveltime tomography was compatible with the linearity requirement of the waveform inversion. This frequency is unrealistically low for real data [although there have been recent moves toward lower frequency reflection surveys (Ziolkowski et al., 2001)]. Using a higher starting frequency would introduce difficulties due to nonlinearities that we do not address within this paper.

Figure 13 shows the contribution of each frequency; the inversion results are shown in Figure 14. As expected, the res-

olution of the image improves as higher frequencies are used in the inversion. The final velocity model (following the use of 10-Hz data) is very close to the true model. Except at the very edges of the model where the coverage is incomplete, the full complexity of the 2D structure is recovered. Figure 15 shows three representative velocity profiles within the model, and compares these to the true model and the starting model from traveltime inversion, from which it may be concluded that not only are the structures correctly imaged, but also that the magnitudes of the velocity perturbations are accurately estimated. In spite of the oscillation of the individual contributions to the image at intermediate frequencies (Figure 13), there is no evidence of any oscillatory artifacts (“ringing”) in these results.

An essential aspect of any inversion is an illustration of the degree to which the result actually predicts the observations. Figure 16 shows the data residuals for the same shot gathers showed in Figure 11 and 12 computed in the starting model and in the final velocity model (after 10 Hz). Although only three frequencies were used for the waveform inversion, the waveforms in the time domain (created using all 122 frequencies) show an excellent fit with the true waveforms. Some misfit remains at the edge of the model (near offset of shots 1 and 96) due to the lack of coverage of the edge of the model.

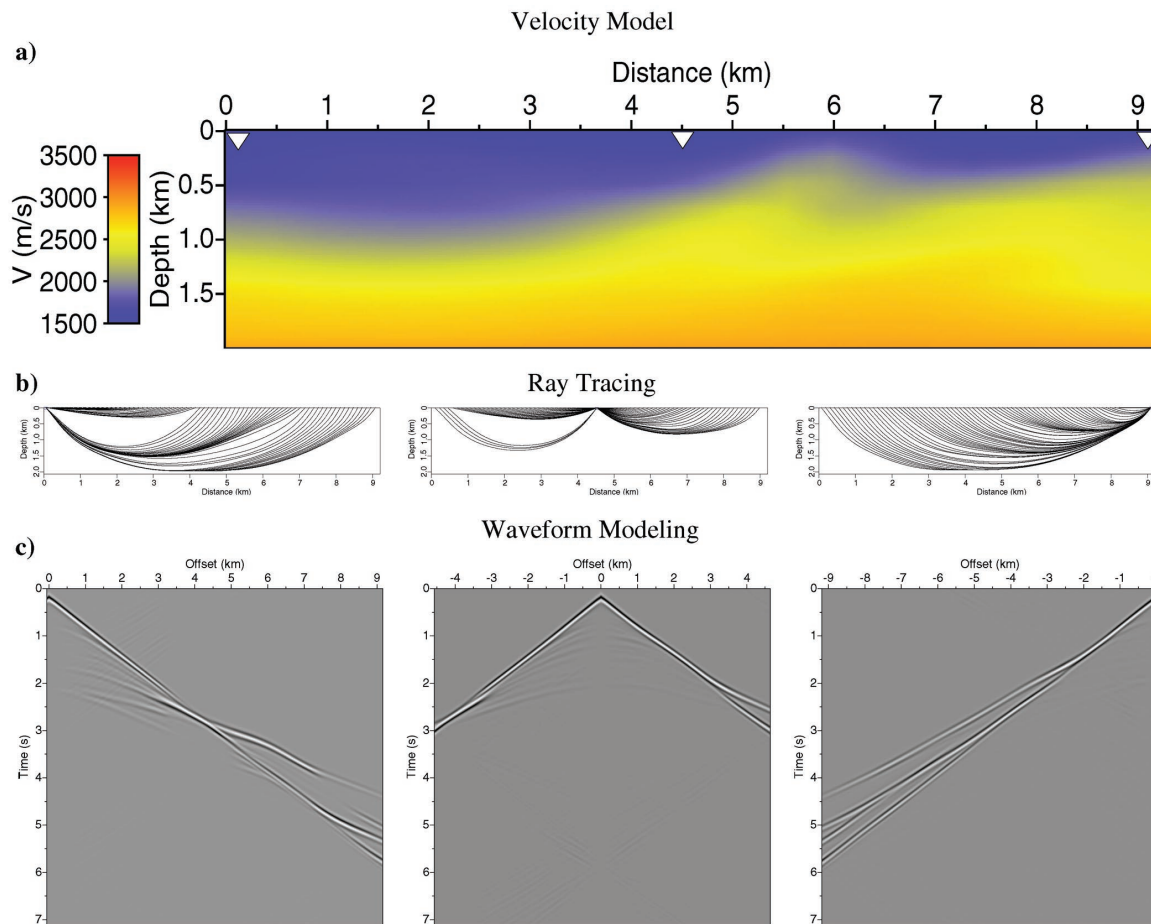


FIG. 12. First arrival traveltime inversion showing (a) the final traveltime inversion velocity model, and (b) examples of ray tracing in the final model for shots 1, 48, and 96. This velocity model fits the first arrival picks to an rms level of 27 ms. (c) Finite-difference seismic traces (normalized) in this velocity model for the three shot positions.

DISCUSSION

Efficiency of the strategy for 2D heterogeneous media

We have shown in this paper that our frequency selection strategy yields accurate results in both 1D and 2D heterogeneous media, despite the fact that the strategy was developed using only the simplest case of a 1D target embedded in an homogeneous model. Our strategy assumes 1D illumination by straight rays only, in which case the incident and scattering angles yield information only on the vertical component of the target. In a more realistic 2D, heterogeneous medium, other wavenumber directions must be recovered. The success of the strategy when applied to more complex models (such as the Marmousi model in the previous section) may be explained by the nature of the background velocity model. A more complex model can produce additional coverage due to two distinct effects:

- 1) Non symmetric illumination: In complex, 2D models, the incident and scattering angles are different. Energy reaches the receivers from nonhorizontal structures (containing nonvertical wavenumber components). Hence, the inversion will recover additional wavenumbers away from the vertical axis.
- 2) Ray bending: In models with an increase in velocity with depth, the incident and scattering angles are wider than

those predicted in a homogeneous model due to the progressive curvature of ray paths toward the horizontal axis with depth. Wider scattering angles provide lower wavenumber information than that predicted in a homogeneous medium; at very wide angles, diving waves will exist, which are scattered from the very low wavenumber components of the velocity model (i.e., they tell us about the background velocities). The application of the frequency selection strategy can thus produce some redundancy in wavenumber coverage, since the wavenumber coverage of a single frequency is actually wider than our prediction.

These effects together seem to permit the effective 2D coverage of the model observed with the Marmousi experiment, even with a limited number of frequencies selected according to a strategy based on 1D, homogeneous models.

The equivalence between gradient images and migration

Many authors have identified the kinematic equivalence between the first iteration gradient image and prestack migration (e.g., Tarantola, 1986; Mora, 1989). Migration maps the data (after muting of the first arrivals) to “isochrones” in the model space (Miller et al., 1987), whereas the gradient maps the data residuals (which may contain transmitted as well as reflected

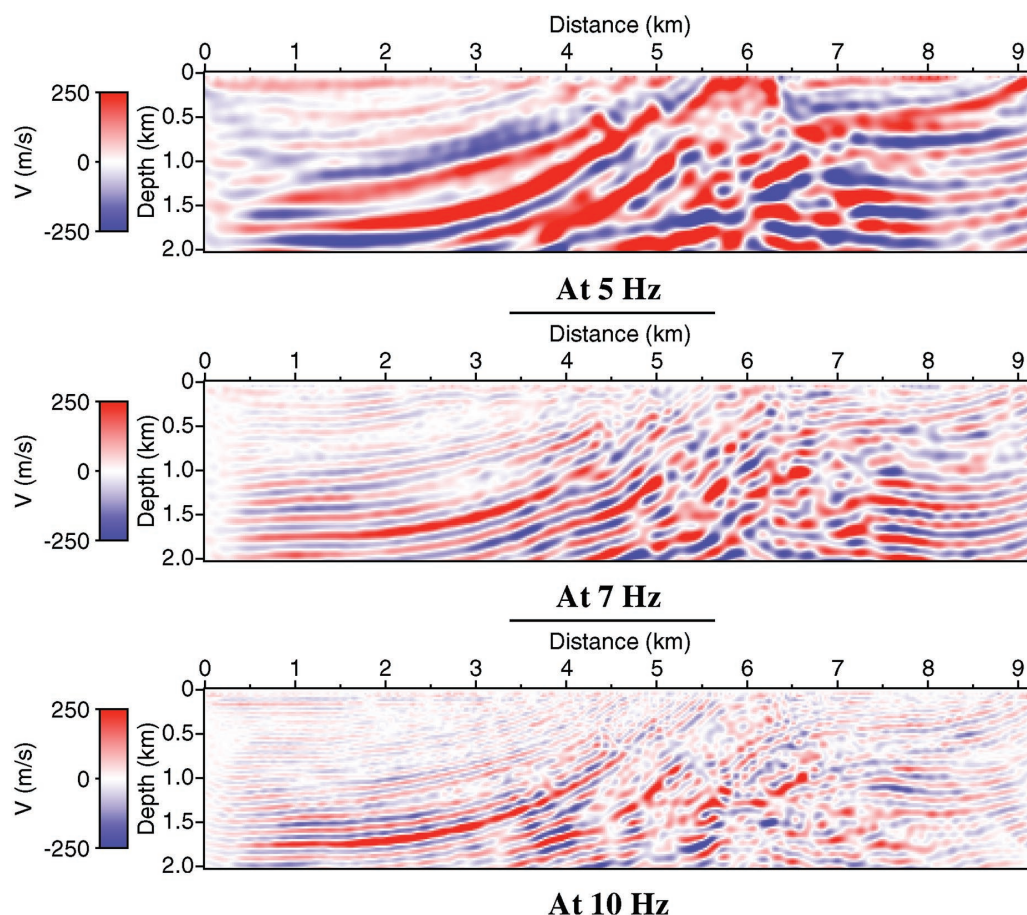


FIG. 13. Waveform inversion in Marmousi: velocity perturbation reconstructed by each frequency.

arrivals) to the wavepath (see Figure 5). Transmitted events map within the first Fresnel zones of the wavepath, while reflected events map to the higher order Fresnel zones, which are elliptical zones that corresponds to the isochrones used in migration (Woodward, 1992).

In the first iteration of a waveform inversion scheme, the starting model is normally a highly smoothed model, as in our numerical test of the Marmousi model above. Such a model will generate accurate transmitted arrivals but will generate no reflected energy (as in Figure 12). As a result, the first iteration data residuals will be dominated by reflections. In our numerical test using the 1D model above, the initial data residuals consisted only of reflected energy from the target at 2-km depth. If the data residuals consist only of reflected energy, the first iteration image is kinematically equivalent to a migration of the data. The gradient is not a particularly well-designed migration operator, since there are no focusing terms or amplitude terms to act as preconditioners. Nevertheless, the correspondence is significant, and frequency-domain migration algorithms could potentially make use of the strategy developed in this paper.

This correspondence between gradient images and migration operators implies that the wavenumber coverage of the

gradient image noted in this paper is also relevant in migration. Multioffset migration also contains an “image stretch” effect (Gardner et al., 1974; Tygel et al., 1994) that is normally considered detrimental (Brown, 1994; Levin, 1998). The meaning commonly given to migration is the recovery of discontinuities in the subsurface, so that the suppression of low wavenumbers may even be considered desirable. We have shown in this paper that the image stretch, when properly handled, will in fact contribute the low wavenumbers to the imaged velocity structure. Furthermore, large offsets are not required in the acquisition to take advantage of the effect of stretch. Implementation of the strategy to standard reflection profiles with offsets limited to less than 3 km still allow a significant decimation of the number of required frequencies.

However, the successful application of our strategy to prestack migration would depend on the accurate combination of multioffset images. This implies that the contribution of each offset to the image must carry true-amplitude information. The result of a simple stack of the multioffset images will fail to recover the true velocity perturbations if the images contain inaccurate relative amplitudes. This is also true for the simple gradient in waveform inversion, but the iterative

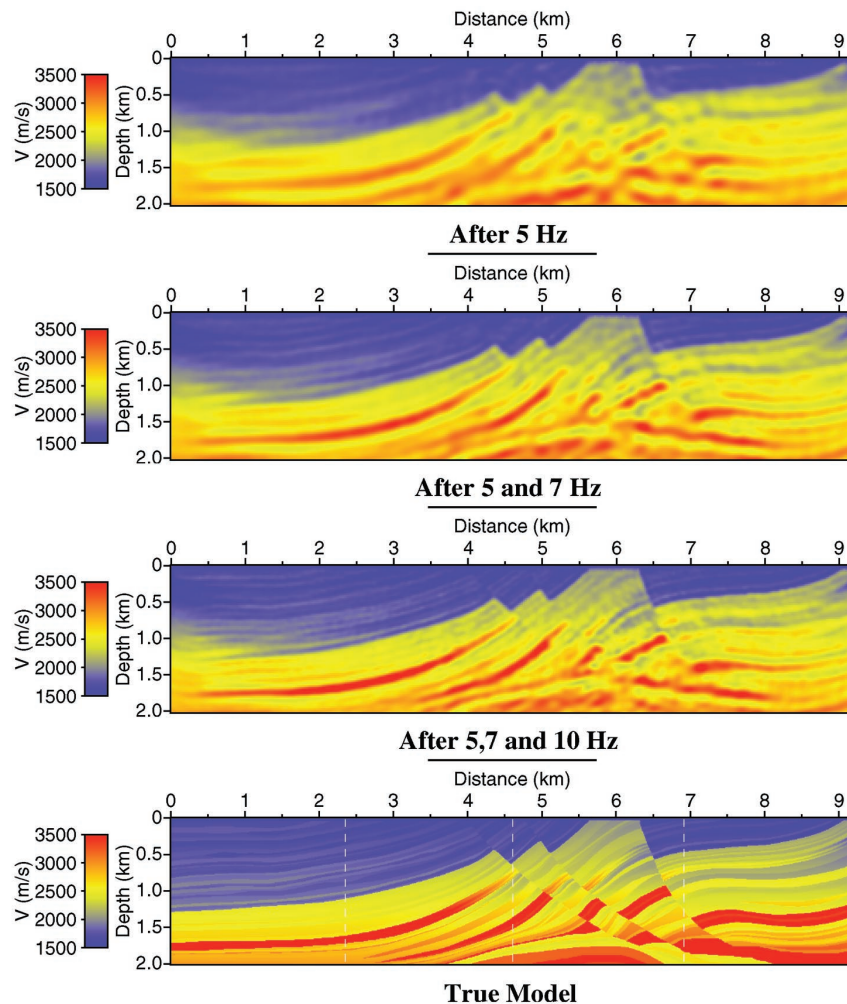


FIG. 14. Waveform inversion results in the Marmousi model after accumulating contributions from progressively higher frequencies. White lines indicate velocity profiles illustrated in Figure 15.

implementation of the descent methods corrects for the inaccuracies in the stack. Migration is usually formulated as a one-step process, without iteration, in which case effects such as spherical divergence and the number of input traces contributing to each offset image must be explicitly compensated for. In order to use our approach to frequency selection for prestack depth migration, the required migration operators need to at least preserve amplitudes during the combination of multiple offset images (e.g., Schleicher et al., 1993).

As an alternative, since the implementation of our strategy implies a significant reduction of the computational cost of the imaging, iterative migration/inversion techniques may become more feasible.

CONCLUSION

We have defined a strategy for selecting temporal frequencies for efficient prestack imaging. The selection scheme defines a much sparser set of frequencies than does the frequency sampling theorem. This is possible because our approach makes use of the multioffset coverage of each reflection point and optimizes the use of information contributed by each source-receiver pair. By taking advantage of the image stretch occurring with increasing offset, the strategy adapts the selection of frequencies to the maximum offset present in the data. The idea is simple: the larger this maximum offset is, the fewer frequencies are needed. In deriving the strategy, we made the approximations of a 1D target in an homogeneous background media, illuminated by plane waves.

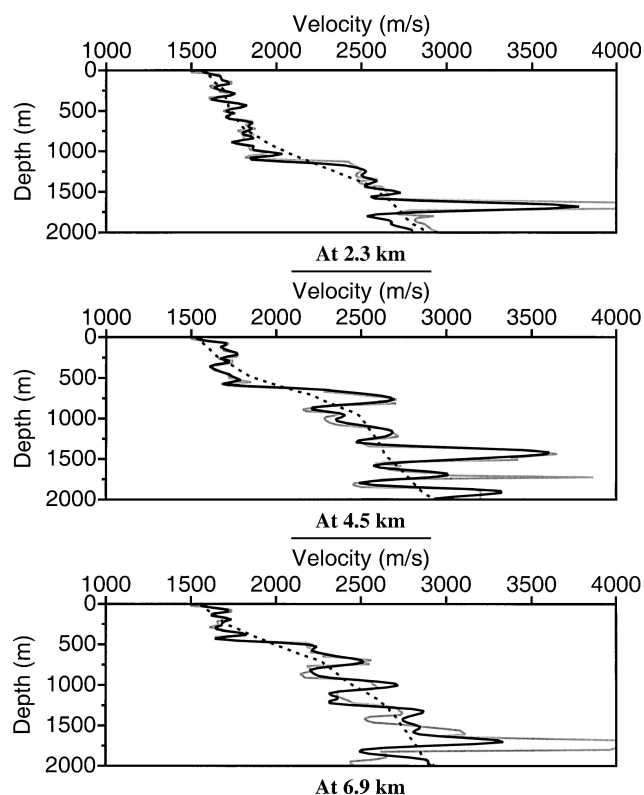


FIG. 15. Velocity profiles at three locations in the Marmousi model (see Figure 14) showing the true model (gray), the starting model (dotted), and the result of the waveform inversion (solid).

Waveform inversion was used in order to illustrate numerically the accuracy of the method. We showed that, for a 1D target in a homogeneous medium, using only the frequencies defined by our strategy, we could produce a result comparable with equivalent, wide-band, time-domain inversion (using the full set of frequencies as defined by the sampling theorem).

We also used the 2D Marmousi velocity model to test the strategy in a more realistic heterogeneous medium. A traveltime tomographic inversion of the first arrivals was used in order to create a smooth starting velocity model for our waveform inversion. This experiment yielded very good results for the recovery of both the background velocities and the discontinuities of the velocity model, demonstrating the imaging potential of parameterizing the full wavefield using only a limited set of frequencies.

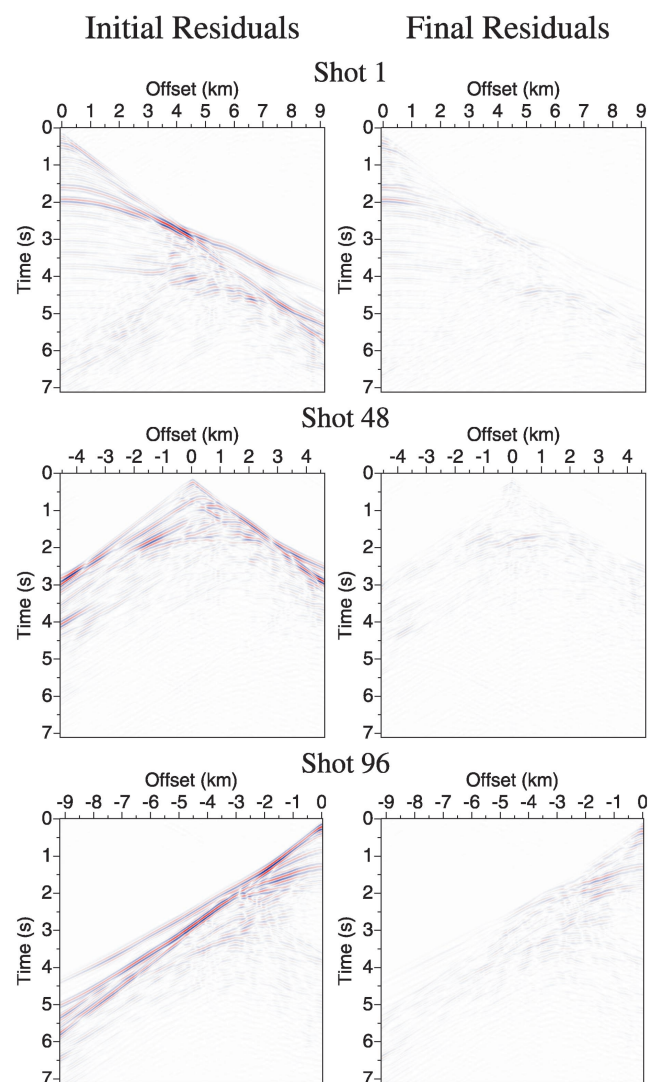


FIG. 16. Data residuals in the Marmousi model in the starting model (left column) and in the final model of the waveform inversion (right column) for shots 1, 48, and 96. Despite using only three frequencies in the waveform inversion, the modeling using 122 frequencies shows a good fit between the real and estimated data. This is a true-amplitude display, the color scale of initial and final residuals are identical.

Our strategy may easily be implemented in practice to prestack depth migration techniques. However, further research should be carried out in order to investigate the efficacy of the selection scheme to migration.

ACKNOWLEDGMENTS

We thank CGG France for sponsoring this work. We also to thank Gilles Lambaré for his useful comments on an early draft of the paper, as well as the reviewers and associate editor for their constructive criticisms.

REFERENCES

- Brown, R., 1994, Image quality depends on your point of view: The Leading Edge, **13**, 669–673.
- Buchholtz, H., 1972, A note on signal distortion due to dynamic (NMO) correction: Geophysical Prospecting, **20**, 395–402.
- Bunks, C., Saleck, F. M., Zaleski, S., and Chavent, G., 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473.
- Clayton, R. W., and Stolt, R. H., 1981, A Born-WBKJ inversion method for acoustic reflection data: Geophysics, **46**, 1559–1567.
- Concus, P., Golub, G., and O'Leary, D., 1976, A generalized conjugate gradient method for the numerical solution of elliptic pde's, in: Bunch, J., and Rose, D., Eds. Sparse matrix computations: Academic Press.
- Cullity, B. D., 1978, Elements of X-ray diffraction: Addison-Wesley.
- Devaney, A. J., 1981, Inverse-scattering theory within the Rytov approximation: Optics Letters, **6**, 374–376.
- Dunkin, J. W., and Levin, F. K., 1973, Effect of normal moveout on a seismic pulse: Geophysics, **38**, 635–642.
- Ewald, P. P., 1921, Das "reziproke Gitter" in der struktur theorie: Zeitschrift für Kristallographie, **56**, 129–156.
- Forgues, E., Scala, E., and Pratt, R. G., 1998, High resolution velocity model estimation from refraction and reflection data: 68th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 1211–1214.
- Freudenreich, Y., and Singh, S., 2000, Full waveform inversion for seismic data-frequency versus time domain: 62nd Meeting, European Association of Geoscientists and Engineers, C0054.
- Gardner, G. H. F., French, W. S., and Matzuk, T., 1974, Elements of migration and velocity analysis: Geophysics, **39**, 811–825.
- Gauthier, O., Virieux, J., and Tarantola, A., 1986, Two-dimensional nonlinear inversion of seismic waveforms—Numerical results: Geophysics, **51**, 1387–1403.
- Haldorsen, J. B. U., and Farmer, P. A., 1989, Resolution and NMO-stretch: Imaging by stacking: Geophysical Prospecting, **37**, 479–492.
- Ikelle, L. T., Diet, J. P., and Tarantola, A., 1986, Linearized inversion of multioffset seismic reflection data in the frequency-wavenumber domain: Geophysics, **51**, 1266–1276.
- Jo, C.-H., Shin, C., and Suh, J. H., 1996, An optimal 9-point, finite-difference, frequency-space, 2-D scalar wave extrapolator: Geophysics, **61**, 529–537.
- Lailly, P., 1983, The seismic inverse problems as a sequence of before stack migration, in: Bednar, J. B., Redner, R., Robinson, E., and Weglein, A., Eds., Conference on inverse scattering: theory and application: Society of Industrial and Applied Mathematics.
- Levin, S. A., 1998, Resolution in seismic imaging: Is it all matter of perspective: Geophysics, **63**, 743–749.
- Liao, Q., and McMechan, G. A., 1996, Multifrequency viscoacoustic modeling and inversion: Geophysics, **61**, 1371–1378.
- Mallick, S., and Frazer, N. L., 1987, Practical aspects of reflectivity modeling: Geophysics, **52**, 1355–1364.
- Marfurt, K. J., 1984, Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave-equations: Geophysics, **49**, 533–549.
- Miller, D. E., Oristaglio, M., and Beylkin, G., 1987, A new slant on seismic imaging: Migration and integral geometry: Geophysics, **52**, 943–964.
- Mora, P. R., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228.
- 1989, Inversion = migration + tomography: Geophysics, **54**, 1575–1586.
- Podvin, P., and Lecomte, I., 1991, Finite difference computation of traveltimes in very contrasted velocity models: A massively parallel approach and its associated tools: Geophysical Journal International, **105**, 271–284.
- Pratt, R. G., and Goulty, N. R., 1991, Combining wave-equation imaging with traveltome tomography to form high-resolution images from crosshole data: Geophysics, **56**, 208–224.
- Pratt, R. G., Shin, C., and Hicks, G. J., 1998, Gauss-Newton and full Newton methods in frequency space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.
- Pratt, R. G., Song, Z.-M., Williamson, P., and Warner, M., 1996, Two-dimensional velocity models from wide-angle seismic data by wavefield inversion: Geophysical Journal International, **124**, 323–340.
- Pratt, R. G., and Worthington, M. H., 1988, The application of diffraction tomography to cross-hole seismic data: Geophysics, **53**, 1284–1294.
- 1990, Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave-equation method: Geophysical Prospecting, **38**, 287–310.
- Schleicher, J., Tygel, M., and Hubral, P., 1993, 3-D true-amplitude finite-offset migration: Geophysics, **58**, 1112–1126.
- Shipp, R. M., and Singh, S. C., 2002, Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data: Geophysical Journal International, **151**, 325–344.
- Shipp, R., Singh, S., and Barton, P., 1997, Sub-basalt imaging using full wavefield inversion: 67th Annual International Meeting, Society of Exploration Geophysicists, Expanded Abstracts, 1563–1566.
- Stockwell, J. W., 1997, Free software in education: A case study of CWP/SU: Seismic Un*x: The Leading Edge, **16**, 1045–1049.
- Tarantola, A., 1984a, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.
- 1984b, Linearized inversion of seismic reflection data: Geophysical Prospecting, **32**, 998–1015.
- 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903.
- 1987, Inverse problem theory: Methods for data fitting and parameter estimation: Elsevier Science Publishers.
- Tygel, M., Schleicher, J., and Hubral, P., 1994, Pulse distortion in depth migration: Geophysics, **59**, 1561–1569.
- Versteeg, R., 1994, The Marmousi experience: Velocity model determination on a synthetic complex data set: The Leading Edge, **13**, 927–936.
- Vidale, J. E., 1990, Finite-difference calculation of traveltimes in three dimensions: Geophysics, **55**, 521–526.
- Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, 15–26.
- Wu, R. S., and Toksöz, M. N., 1987, Diffraction tomography and multisource holography applied to seismic imaging: Geophysics, **52**, 11–25.
- Yilmaz, Ö., 1987, Seismic data processing: Society of Exploration Geophysicists.
- Zelt, C. A., and Barton, P. J., 1998, 3D seismic refraction tomography: A comparison of two methods applied to data from the Faeroe Basin: Journal of Geophysical Research, **103**, 7187–7210.
- Ziolkowski, A., Hanssen, P., Gatliff, R., Li, X., and Jakubowicz, H., 2001, The use of low frequencies for sub-basalt imaging: 71st Annual International Meeting, Society of Exploration Geophysicists, 74–77.

APPENDIX

IMAGE STRETCH AND NMO STRETCH

In a homogeneous medium, the zero-offset two-way traveltime for a single horizontal reflection is given by

$$\tau = \frac{2z}{c}, \quad (\text{A-1})$$

where z is the depth of the reflector and c is the velocity. At finite offsets, the two-way traveltime is

$$\begin{aligned} t &= \frac{2}{c} \sqrt{z^2 + h^2} \\ &= \left(\tau^2 + \left(\frac{2h}{c} \right)^2 \right)^{\frac{1}{2}}, \end{aligned} \quad (\text{A-2})$$

where h is the half offset.

The NMO correction seeks to move all reflections to their zero-offset times. NMO stretch occurs because reflections are associated with a wavelet of finite (nonzero) duration. Therefore, the start of the reflection is moved to the correct zero-offset time, whereas the end of the wavelet is moved to a different zero-offset time. If we characterize the width of the wavelet as $\Delta\tau$ and the width of the wavelet after NMO correction as $\Delta\tau'$ the NMO stretch is

$$\frac{\Delta\tau'}{\Delta\tau}, \quad (\text{A-3})$$

which varies as a function of depth and offset. Both these dependencies are captured by considering the NMO stretch to be a function of two-way time, t . The instantaneous value of the

NMO stretch is therefore

$$\lim_{\Delta t \rightarrow 0} \left(\frac{\delta\tau}{\Delta t} \right) = \frac{d\tau}{dt}. \quad (\text{A-4})$$

We may evaluate this derivative with the use of equation (A-2), from which we obtain

$$\frac{d\tau}{dt} = \sqrt{1 + R^2} = \frac{1}{\alpha}, \quad (\text{A-5})$$

where $R = h/z$ is the offset-to-depth ratio and $\alpha = \cos \theta$ as in equation (19).

We therefore conclude that the shift of information toward the low wavenumber with offset predicted by equation (19) is exactly the same as that caused by NMO stretch after depth correction.