Gradient of image-space wave-equation tomography by the adjoint-state method

Claudio Guerra

ABSTRACT

Optimization with gradient-descent techniques requires computing the gradient of the objective function. The gradient can be determined by using the Frechét derivatives, but, for practical problems, this can be very expensive. The gradient can be more efficiently computed by the adjoint-state method, which does not require the use of the Frechét derivatives. Here, I derive the gradient of the image-space wave-equation tomography using the adjoint-state method. I also show its application with a numerical example using image-space phase-encoded gathers.

INTRODUCTION

Wave-equation tomography aims to solve for earth models that explain observed seismograms under some norm. There are two main categories, depending on the domain in which the objective function is computed. In one category, known as waveform inversion (Lines and Treitel, 1984; Tarantola, 1987; Woodward, 1992), the objective function is defined in the data space, where the modeled data are compared with the recorded seismograms. In the other category, here called image-space wave-equation tomography (ISWET), the objective function is minimized in the image space. Two variants of ISWET are wave-equation migration velocity analysis (WEMVA) (Sava and Biondi, 2004a,b) and differential semblance velocity analysis (DSVA) (Shen, 2004; Shen and Symes, 2008).

Like waveform inversion, ISWET is a computationally demanding process. This computational cost is commonly decreased by using generalized sources (Shen and Symes, 2008; Tang et al., 2008). Because wavefield propagation is a linear process, generalized sources are computed by linearly combining source wavefields and receiver wavefields, using phase-encoding techniques (Whitmore, 1995; Romero et al., 2000). In particular, Guerra et al. (2009) used phase-encoding modeling in the image space to synthesize generalized source functions that drastically decrease the cost of DSVA.

WEMVA and DVSA seek the optimal slowness by driving an image perturbation to a minimum. However, they differ in the way the image perturbation is computed and, consequently, in the numerical optimization scheme applied. As Biondi (2008) points out, WEMVA is not easily automated. The image perturbation is computed

by the linearized residual prestack-depth migration (Sava and Biondi, 2004a), which uses a manually picked residual-moveout parameter. Because the perturbed image computed with the linearized residual prestack-depth migration is consistent with the application of the forward wave-equation tomographic operator, the forward and adjoint ISWET operators can be used in conjugate-gradient methods to invert for the slowness perturbation. In DSVA, the perturbed image is computed by applying the fully automated differential-semblance operator (DSO) to the subsurface-offset gathers (ODCIGs) or angle gathers (ADCIGs). When applied to ODCIGs, DSVA minimizes the energy not focused at zero-offset. When applied to ADCIGs, DSVA minimizes energy departing from flatness of the reflectors. Although DSVA easily automates ISWET, it produces perturbed images that do not present the depth phase-shift introduced by the forward one-way ISWET operator, neither the DVSA amplitude behavior can be modeled by this operator. Therefore, the objective function computed with DSO is tipically minimized by quasi-Newton algorithms, which require computation of the gradient of the objective function.

The gradient of the objective function can be computed with the Frechét derivatives. However, even for 2D applications of ISWET this computation can be very expensive. An efficient way to compute the gradient without using Frechét derivatives is the adjoint-state method (Chavent and Jacewitz, 1995; Plessix, 2006). Plessix (2006) describes two methodologies for computing the gradient of the objective function using the adjoint-state method. One methodology uses the perturbation theory, which states that, at first order, a perturbation of the model parameters causes a perturbation of the objective function. The other uses the augmented Lagrangian. The augmented Lagrangian is formed by the objective function and the scalar product of the adjoint-state variables with general solutions of the forward modeling equations. The adjoint-state variables are, in turn, solutions of the adjoint-state equations. The adjoint-state equations are defined by equating to zero the derivatives of the augmented Lagrangian with respect to the state variables. For the linear case, the adjoint of the modeling operator applied to the adjoint-state variables gives the gradient of the objective function.

Although previous studies have computed the gradient of the ISWET objective function by the adjoint-state method (for example, Shen et al. (2003)), they have not provided a detailed derivation. Here, I show a detailed derivation the gradient of the ISWET objective function using the augmented Lagrangian methodology. The derivation is valid whether ISWET uses areal-shot migration or shot-profile migration. A complete description of the forward and adjoint of ISWET operators is given in Tang et al. (2008). A numerical example of slowness optimization on the Marmousi model illustrates the use of the gradient by a quasi-Newton algorithm.

Because I use image-space phase-encoded gathers in the numerical example, for completeness, I first briefly describe how to compute these phase-encoded gathers. A detailed treatment of the image-space phase-encoded gathers can be found in Biondi (2006, 2007) and Guerra and Biondi (2008a). Then I derive the ISWET gradient using the adjoint-state method, and show the numerical example .

IMAGE-SPACE PHASE-ENCODED GATHERS

Biondi (2006, 2007) introduced the concept of the prestack exploding-reflector as a generalization of the exploding-reflector method (Loewenthal et al., 1976). The prestack exploding-reflector modeling synthesizes areal data and the corresponding areal source function, having as an initial condition a prestack image computed with wave-equation migration. If the slowness is accurate and the energy is focused at zero-subsurface offset, the prestack exploding-reflector modeling reduces to the conventional exploding-reflector method. Basically, the prestack exploding-reflector method models one single reflection event from one single ODCIG by recursive upward continuation with the following one-way wave equations:

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2\hat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) d(\mathbf{x},\omega;x_m,y_m) = r_D(\mathbf{x},\mathbf{h};x_m,y_m) \\ d(x,y,z=z_{\max},\omega;x_m,y_m) = 0 \end{cases}, \tag{1}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2\hat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) u(\mathbf{x},\omega;x_m,y_m) = r_U(\mathbf{x},\mathbf{h};x_m,y_m) \\ u(x,y,z=z_{\max},\omega;x_m,y_m) = 0 \end{cases}, \tag{2}$$

where $r_D(\mathbf{x},\mathbf{h};x_m,y_m)$ and $r_U(\mathbf{x},\mathbf{h};x_m,y_m)$ are a single ODCIG at the horizontal location (x_m,y_m) with a single reflector, and are suitable initial conditions for modeling the source and receiver wavefields, respectively. They are obtained by rotating the original unfocused ODCIGs according to the apparent geological dip of the reflector. This operation maintains the velocity information needed for migration velocity analysis, especially for dipping reflectors (Biondi, 2007). Note that $d(x,y,z=z_c,\omega;x_m,y_m)$ is the areal source data and $u(x,y,z=z_c,\omega;x_m,y_m)$ is the areal receiver data for a single reflector and a single ODCIG located at (x_m,y_m); $z=z_c$ denotes that the wavefields can be collected at any depth level, z_c. This characteristic is important for accelerating ISWET, especially if z_c separates regions of sufficiently accurate slowness above and inaccurate slowness below. Therefore, the synthesized gathers are naturally datumized, and the wavefield propagations during ISWET can be restricted to the region where the slowness model must be updated. This feature allows the application of ISWET to be target-oriented.

As initially formulated, if one models a single reflection from one ODCIG at a time, the prestack exploding-reflector method generates a dataset that can be orders of magnitude bigger than the original dataset. As discussed by Biondi (2006) and Guerra and Biondi (2008b,a), modeling several reflectors and several ODCIGs simultaneously and using random phase encoding generates a much smaller dataset that, when migrated, do not produce crosstalk. The randomly encoded areal source and areal receiver wavefields can be computed as follows:

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2\hat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widetilde{d}(\mathbf{x},\mathbf{p}_m,\omega) = \widetilde{r}_D(\mathbf{x},\mathbf{h},\mathbf{p}_m,\omega) \\ \widetilde{d}(x,y,z=z_{\max},\mathbf{p}_m,\omega) = 0 \end{cases}, \tag{3}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 \hat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widetilde{u}(\mathbf{x}, \mathbf{p}_m, \omega) = \widetilde{r}_U(\mathbf{x}, \mathbf{h}, \mathbf{p}_m, \omega) \\ \widetilde{u}(x, y, z = z_{\max}, \mathbf{p}_m, \omega) = 0 \end{cases}, \tag{4}$$

where $\widetilde{r}_D(\mathbf{x}, \mathbf{h}, \mathbf{p}_m, \omega)$ and $\widetilde{r}_U(\mathbf{x}, \mathbf{h}, \mathbf{p}_m, \omega)$ are the encoded ODCIGs after rotations. They are defined as follows:

$$\widetilde{r}_D(\mathbf{x}, \mathbf{h}, \mathbf{p}_m, \omega) = \sum_{x_m} \sum_{y_m} r_D(\mathbf{x}, \mathbf{h}, x_m, y_m) \beta(\mathbf{x}, x_m, y_m, \mathbf{p}_m, \omega), \tag{5}$$

$$\widetilde{r}_U(\mathbf{x}, \mathbf{h}, \mathbf{p}_m, \omega) = \sum_{x_m} \sum_{y_m} r_U(\mathbf{x}, \mathbf{h}, x_m, y_m) \beta(\mathbf{x}, x_m, y_m, \mathbf{p}_m, \omega), \tag{6}$$

where $\beta(\mathbf{x}, x_m, y_m, \mathbf{p}_m, \omega) = e^{i\gamma(\mathbf{x}, x_m, y_m, \mathbf{p}_m, \omega)}$ is a pseudo-random phase-encoding function, with $\gamma(\mathbf{x}, x_m, y_m, \mathbf{p}_m, \omega)$ being a uniformly distributed random sequence in $\mathbf{x}$, x_m, y_m and ω; the variable $\mathbf{p}_m$ is the index of different realizations of the random sequence. The recursive solution of equations 3 and 4 gives the encoded areal source data $\widetilde{d}(x, y, z = z_c, \mathbf{p}_m, \omega)$ and areal receiver data $\widetilde{u}(x, y, z = z_c, \mathbf{p}_m, \omega)$, at the depth level, z_c.

GRADIENT OF ISWET BY THE ADJOINT-STATE METHOD

Image-space wave-equation tomography aims to iteratively solve for the slowness model, $s = s(\mathbf{x})$, that minimizes the nonlinear objective function:

$$J(s) = \frac{1}{2} ||\Delta r(s)||^2 = \frac{1}{2} ||r(s) - \mathbf{M} r(s)||^2, \tag{7}$$

where $\Delta r = \Delta r(\mathbf{x}, \mathbf{h})$ is the image perturbation that measures the accuracy of the slowness model. To compute Δr, a differential residual-focusing operator $\mathbf{M}$ is applied to the image $r = r(\mathbf{x}, \mathbf{h})$ obtained with the current slowness (Biondi, 2008), using either differential residual prestack migration (Sava and Biondi, 2004a,b) or differential-semblance optimization (DSO) operators (Shen and Symes, 2008). In this paper, operators are represented by bold capital letters.

If the differential residual-focusing operator $\mathbf{M}$ is independent of the slowness, the gradient of this objective function evaluated at the current slowness $\widehat{s} = \widehat{s}(\mathbf{x})$ is

$$\nabla J(s) = \left(\frac{\partial r}{\partial s}\right)' \Bigg|_{s=\widehat{s}} (\mathbf{I} - \mathbf{M}') \Delta \widehat{r}, \tag{8}$$

where the prime denotes the adjoint, $\mathbf{I}$ is the identity operator, and $\Delta \widehat{r} = \Delta \widehat{r}(\mathbf{x}, \mathbf{h})$ is the perturbed image obtained with the current slowness model. The linear operator $\frac{\partial r}{\partial s}$ defines the mapping $\frac{\partial r}{\partial s} \Delta s = \Delta r$ between the slowness perturbation Δs and the image perturbation Δr; it is called the image-space wave-equation tomographic operator.

Because the image-space wave-equation tomographic operator is composed of different operators, it is difficult to envision from equation 8 which operations are performed to compute the gradient. Therefore, for a clear explanation of the operators involved, I use the adjoint-state method to derive the gradient of the objective function (equation 7).

In migration with generalized sources or shot-profile migration, the source and receiver wavefields are propagated independently, and the image $r_z = r_z(\mathbf{x}, \mathbf{h})$ at a depth level z, is computed by the crosscorrelation

$$r_z(\mathbf{x}, \mathbf{h}) = \sum_{\omega} d_z^*(\mathbf{x} - \mathbf{h}, \omega) u_z(\mathbf{x} + \mathbf{h}, \omega), \tag{9}$$

where $d_z(\mathbf{x}, \omega)$ is the source wavefield for a single frequency ω at horizontal coordinates $\mathbf{x} = (x, y)$; $u_z(\mathbf{x}, \omega)$ is the receiver wavefield, $\mathbf{h} = (h_x, h_y)$ is the subsurface half-offset, and "*" stands for the complex-conjugate. An additional summation over shots is required when migrating more than one shot. Hereafter, letters d and u stand for source and receiver wavefields, respectively, irrespective of the migration scheme.

In a more compact notation, not explicitly writing the dependencies on $\mathbf{x}$ and $\mathbf{h}$, equation 9 can be re-written as follows:

$$r_z = \mathbf{S}\mathbf{D}'_z(\omega) u_z(\omega) = \mathbf{S}\mathbf{U}_z(\omega) d_z^*(\omega), \tag{10}$$

where $\mathbf{D}$ and $\mathbf{U}$ are convolutional operators composed of (h_x, h_y)-shifted versions of $d_z(\mathbf{x}, \omega)$ and $u_z(\mathbf{x}, \omega)$, respectively. Operator $\mathbf{S}$ corresponds to the summation over frequency.

For subsequent depth levels, $d(\mathbf{x}, \omega)$ is computed by means of the recursive downward propagation

$$\begin{cases} d_{z+1}(\omega) = \mathbf{T}_z^{\downarrow}(\omega, s) d_z(\omega) \\ d_1(\omega) = q(\omega), \end{cases} \tag{11}$$

where $\mathbf{T}_z^{\downarrow}$ is the downward continuation operator, which is a function of the slowness s, and $q(\omega)$ is the source wavefield used as a boundary condition. In the case of conventional shot-profile migration, $q(\omega) = f_s(\omega)\delta(\mathbf{x} - \mathbf{x_s})$ is the source signature located at $\mathbf{x}_s = (x_s, y_s, 0)$. If using the generalized sources in the image space, $q(\omega)$ represents the image-space phase-encoded source wavefield of equation 3.

The downward continuation of the receiver wavefield is performed by

$$\begin{cases} u_{z+1}(\omega) = \mathbf{T}_z^{\downarrow}(\omega, s) u_z(\omega) \\ u_1(\omega) = w(\omega), \end{cases} \tag{12}$$

where $w(\omega)$ is the recorded data at the surface for shot-profile migration. If using generalized sources in the image space, $w(\omega)$ is the phase-encoded areal receiver wavefield of equation 4. In equations 11 and 12, I omitted the dependencies of the wavefield with respect to $\mathbf{x}$. The subscript *1* in equations 11 and 12 represents the surface for

the shot-profile migration and the "collection" depth level, z_c, for the image-space phase-encoded wavefields.

In the image-space wave-equation tomography problem, the perturbed source and receiver wavefields, and the image perturbation are used to compute the slowness perturbation that updates the current slowness model. From the perturbation theory, we have $d = \widehat{d}+\Delta d$, $u = \widehat{u}+\Delta u$, and, consequently, $r = \widehat{r}+\Delta r$ are physical realizations with $s = \widehat{s}+\Delta s$, where the *hat* refers to fields obtained with the background slowness. To the first order (Born approximation), these perturbed fields are given by

$$\Delta d_{z+1}(\omega) = \mathbf{T}_z^{\downarrow}(\omega,\widehat{s})\Delta d_z(\omega) + \widetilde{\widetilde{\mathbf{D}}}_z(\omega)\Delta s_z \tag{13}$$

and

$$\Delta u_{z+1}(\omega) = \mathbf{T}_z^{\downarrow}(\omega,\widehat{s})\Delta u_z(\omega) + \widetilde{\widetilde{\mathbf{U}}}_z(\omega)\Delta s_z. \tag{14}$$

The diagonal operators $\widetilde{\widetilde{\mathbf{D}}}_z$ and $\widetilde{\widetilde{\mathbf{U}}}_z$ have in the diagonal entries the scattered source and receiver wavefields, respectively. These wavefields are given by the action of the scattering operator $\Delta\mathbf{T}_z^{\downarrow}$ on the background wavefields:

$$\widetilde{\widetilde{\mathbf{D}}}_z(\omega) = \Delta\mathbf{T}_z^{\downarrow}(\omega,\widehat{s})\widehat{d}_z(\omega) = i\frac{\omega^2\widehat{s}}{\sqrt{\omega^2\widehat{s}^2 - |\mathbf{k}|^2}}dz\ \widehat{d}_z(\omega) \tag{15}$$

and

$$\widetilde{\widetilde{\mathbf{U}}}_z(\omega) = \Delta\mathbf{T}_z^{\downarrow}(\omega,\widehat{s})\widehat{u}_z(\omega) = -i\frac{\omega^2\widehat{s}}{\sqrt{\omega^2\widehat{s}^2 - |\mathbf{k}|^2}}dz\ \widehat{u}_z(\omega). \tag{16}$$

The perturbed image is given by

$$\Delta r_z = \mathbf{S}\left(\widehat{\mathbf{U}}_z(\omega)\Delta d_z^*(\omega) + \widehat{\mathbf{D}}'_z(\omega)\Delta u_z(\omega)\right). \tag{17}$$

The matrix representations of equations 13, 14, 17 are

$$\Delta\underline{\mathbf{d}} = \mathbf{T}^{\downarrow}\Delta\underline{\mathbf{d}} + \widetilde{\widetilde{\mathbf{P}}}\mathbf{S}'\Delta\mathbf{s}, \tag{18}$$

$$\Delta\underline{\mathbf{u}} = \mathbf{T}^{\downarrow}\Delta\underline{\mathbf{u}} + \widetilde{\widetilde{\mathbf{U}}}\mathbf{S}'\Delta\mathbf{s}, \tag{19}$$

and

$$\Delta\underline{\mathbf{r}} = \mathbf{S}\left(\widehat{\mathbf{U}}\Delta\underline{\mathbf{d}}^* + \widehat{\mathbf{D}}'\Delta\underline{\mathbf{u}}\right), \tag{20}$$

where $\mathbf{S}'$ is a spreading operator that replicates the slowness perturbation for every frequency.

Equations 18, 19, and 20 are the forward modeling equations of the image-space wave-equation tomography problem using the generalized sources or shot-profile schemes. They depend on the state variables $\Delta\underline{\mathbf{d}}$, $\Delta\underline{\mathbf{u}}$, and $\Delta\underline{\mathbf{r}}$. Plessix (2006) describes how to compute the adjoint states using the augmented functional methodology. By introducing the adjoint-state variables $\underline{\lambda}_d$, $\underline{\lambda}_u$, and $\underline{\lambda}_r$, the augmented Lagrangian reads

$$\begin{aligned}\mathcal{L}(\Delta\underline{\mathbf{d}},\Delta\underline{\mathbf{u}},\Delta\underline{\mathbf{r}},\underline{\lambda}_d,\underline{\lambda}_u,\underline{\lambda}_r;\Delta\mathbf{s}) = \mathcal{R}\left[\tfrac{1}{2}\,||\Delta\underline{\mathbf{r}}||^2 - \right.\\ \left\langle \underline{\lambda}_d, \left(\mathbf{I}-\mathbf{T}^{\downarrow}\right)\Delta\underline{\mathbf{d}} - \widetilde{\widetilde{\mathbf{D}}}\mathbf{S}'\Delta\mathbf{s}\right\rangle - \\ \left\langle \underline{\lambda}_u, \left(\mathbf{I}-\mathbf{T}^{\downarrow}\right)\Delta\underline{\mathbf{u}} - \widetilde{\widetilde{\mathbf{U}}}\mathbf{S}'\Delta\mathbf{s}\right\rangle - \\ \left.\left\langle \underline{\lambda}_r, \Delta\underline{\mathbf{r}} - \mathbf{S}\left(\widehat{\mathbf{U}}\Delta\underline{\mathbf{d}}^{*} + \widehat{\mathbf{D}}'\Delta\underline{\mathbf{u}}\right)\right\rangle\right].\end{aligned}$$

The adjoint-state variables are computed by taking the derivative of $\mathcal{L}$ with respect to the state variables and equating to zero, which gives

$$\left(\mathbf{I}-\mathbf{T}^{\downarrow}\right)'\underline{\lambda}_d = \widehat{\mathbf{U}}\underline{\lambda}_r, \tag{21a}$$

$$\left(\mathbf{I}-\mathbf{T}^{\downarrow}\right)'\underline{\lambda}_u = \widehat{\mathbf{D}}\underline{\lambda}_r, \tag{21b}$$

$$\underline{\lambda}_r = \Delta\underline{\mathbf{r}}. \tag{21c}$$

Notice that

$$\left(\mathbf{I}-\mathbf{T}^{\downarrow}\right)' = \left(\mathbf{I}-\mathbf{T}^{\downarrow'}\right) = \left(\mathbf{I}-\mathbf{T}^{\uparrow}\right) \tag{22}$$

corresponds to the recursive upward propagation operator. Therefore, equations 21a and 21b can be written as

$$\underline{\lambda}_p = \mathbf{T}^{\uparrow}\underline{\lambda}_d + \widehat{\mathbf{U}}\underline{\lambda}_r, \tag{23a}$$

$$\underline{\lambda}_u = \mathbf{T}^{\uparrow}\underline{\lambda}_u + \widehat{\mathbf{D}}\underline{\lambda}_r, \tag{23b}$$

which correspond to the recursive upward propagation of the perturbed wavefields resulting from the convolution of the wavefields computed with the current slowness and the perturbed image.

Finally, the gradient of J is

$$\nabla_s J(\mathbf{s}) = \mathbf{S}\left(\widetilde{\widetilde{\mathbf{D}}}'\,\underline{\lambda}_d + \widetilde{\widetilde{\mathbf{U}}}'\,\underline{\lambda}_u\right). \tag{24}$$

To compute the gradient, the adjoint-state wavefields, $\underline{\lambda}_d$ and $\underline{\lambda}_u$, are upward propagated and cross-correlated in time with the scattered wavefields.

NUMERICAL EXAMPLE

The gradient of the ISWET objective function computed in the previous section can be used in a quasi-Newton optimization scheme. I use the L-BFGS-B bound

constrained optimization algorithm (Nocedal and Wright, 2000) to invert for slowness using the Marmousi model. A B-spline smoothing, with nodes spaced at 240 m in x and 16 m in z, is applied to the gradient to prevent problems caused by poor illumination and noise. I model 375 two-way shots with 24 m spacing using the original Marmousi model of Figure 1a. The maximum offset is 6600 m.

The one-way shot profile image with the correct slowness model can be seen in Figure 1b. The modeled data were also migrated using the background slowness of Figure 2a to compute the background image of Figure 2b. The background slowness differs from the true slowness in a limited region, as can be seen in Figure 3, which shows the ratio between true and background slowness. When comparing the two images, there is a clear pull-up effect at the center-bottom of the background image caused by migrating with a velocity which is too slow.

Following Guerra et al. (2009), representative reflectors were selected from the background image (Figure 4) to synthesize eleven image-space phase-encoded source and receiver wavefields using the background slowness. Prior to modeling, the selected reflectors were subjected to rotation according to the apparent geological dip as previously mentioned. I estimate that, by using only eleven areal shots, the computational cost decreased in 95% considering if the slowness optimization was carried out with the original 376 shot-profiles.

The optimization stopped after 4 iterations comprising 41 function and gradient evaluations. The optimized slowness model can be seen in Figure 5a. The optimized slowness decreased as expected, but did not recover the details of the true slowness. The smoothness of the optimized slowness is partially because of the B-spline filtering of the gradient but also because image-space wave-equation tomography cannot solve for the short wavelengths of the slowness model, as pointed out by Symes (2008).

The accuracy of the optimized slowness model can be evaluated by comparing the computed images with true, background, and optimized slowness. The shot-profile migration of the original shots using the optimized slowness is shown in Figure 5b. Notice that the pull-up effect, which occurs in the background image, is greatly mitigated. In addition, reflectors are more focused in the image computed with the optimized slowness than in the image computed with the background slowness. Another slowness accuracy indicator is flatness of the reflectors in the ADCIGs. The ADCIGs resulting from shot-profile migration of the original shots using, from top to bottom, the true slowness, the background slowness and the optimized slowness are shown in Figure 6. The reflectors computed with the optimized slowness are flatter than those with the background slowness, demonstrating the greater accuracy of the optimized slowness.

CONCLUSION

I showed a detailed derivation of the gradient of the objective function for the image-space wave-equation tomography by means of the adjoint-state method. I

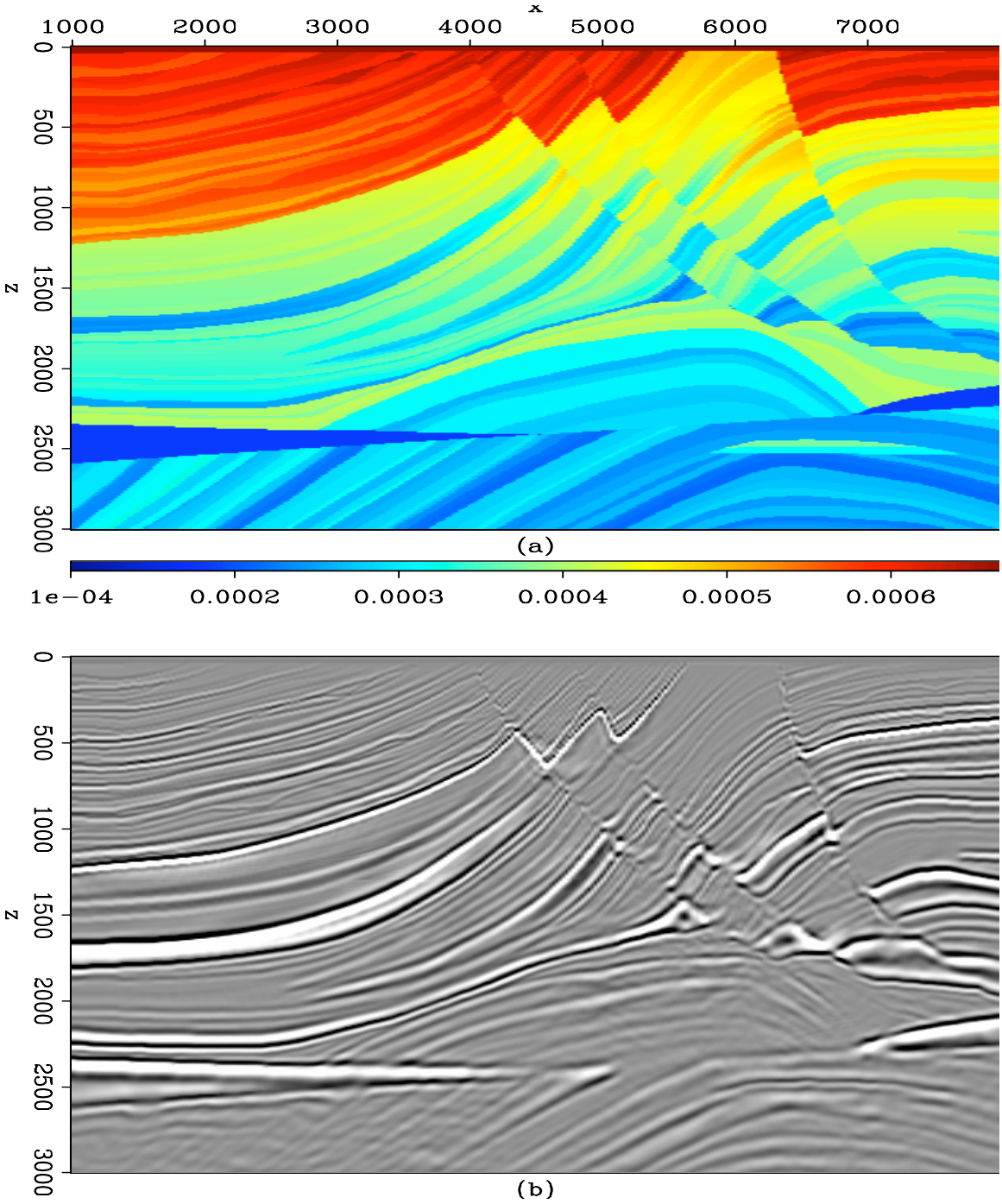

Figure 1: (a) True slowness of the Marmousi model. (b) Zero-subsurface-offset section of the one-way shot-profile migration with the true slowness model.[**CR**]

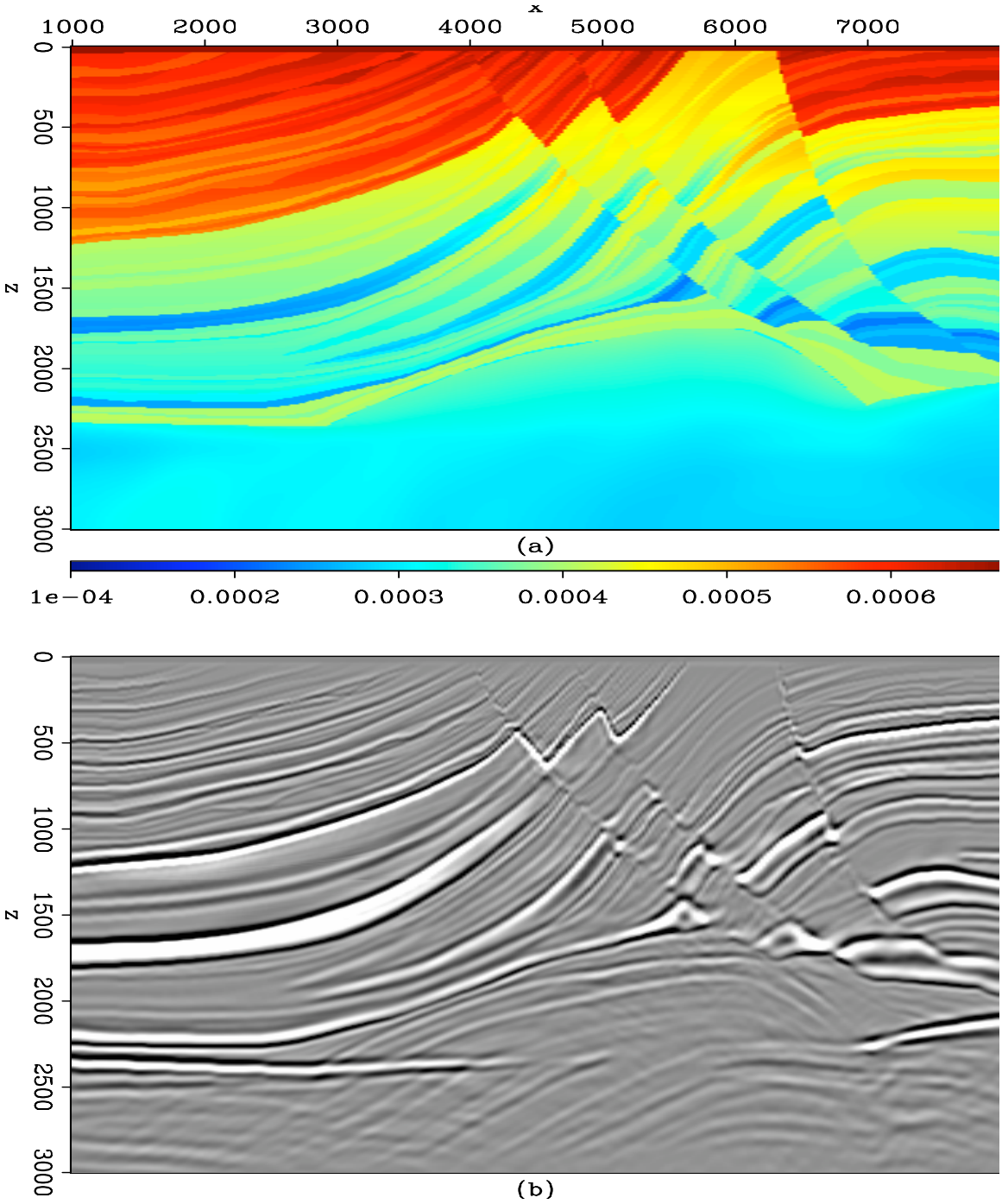

Figure 2: (a) Background slowness. (b) Zero-subsurface-offset section of the one-way shot-profile migration with the background slowness model.[**CR**]

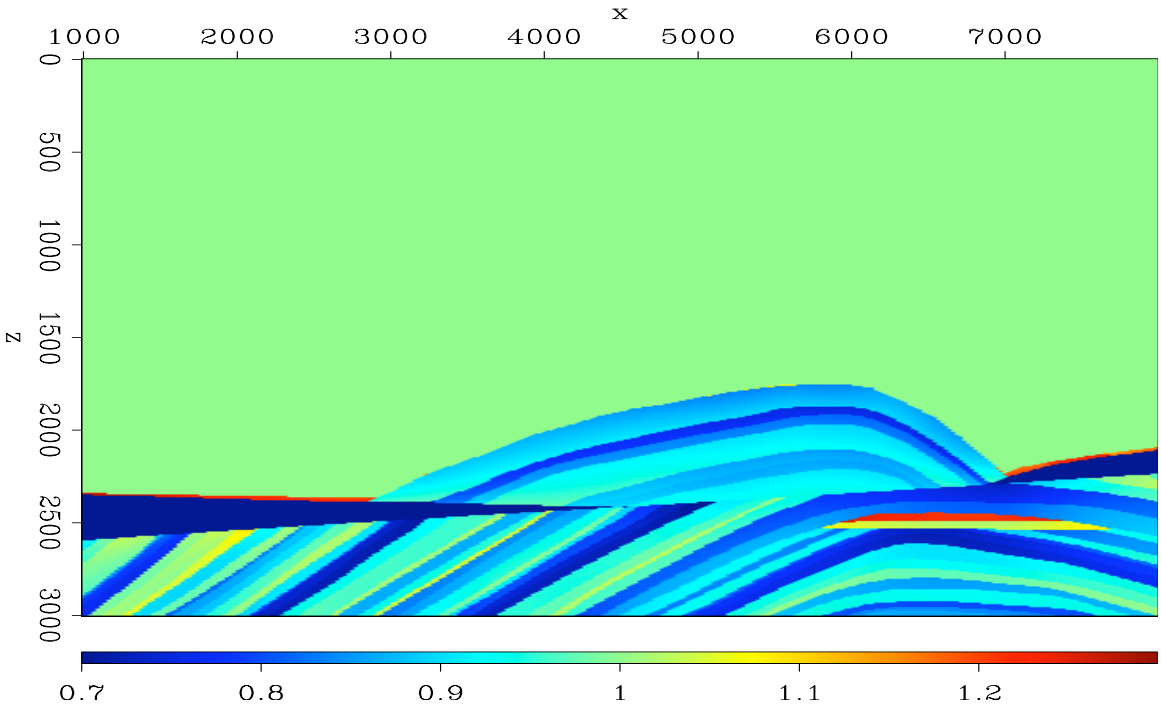

Figure 3: Ratio between the true and the background slowness.[**ER**]

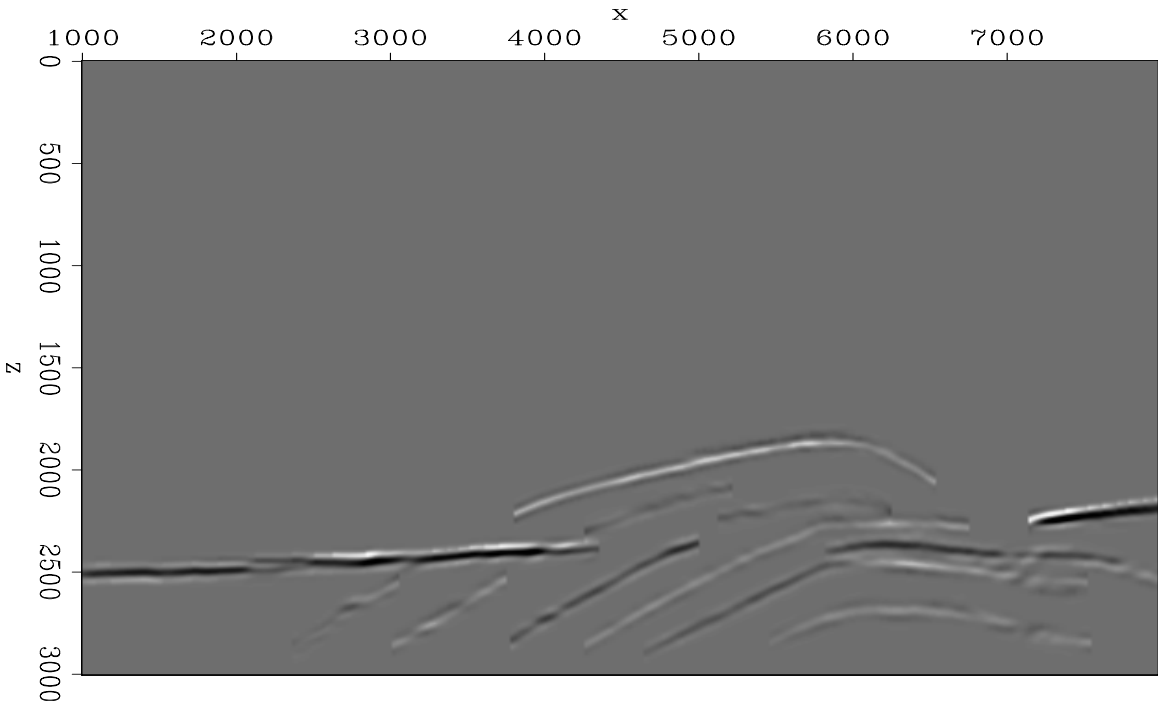

Figure 4: Zero-subsurface-offset section showing the reflectors selected on the one-way shot-profile migration with the background slowness model.[**CR**]

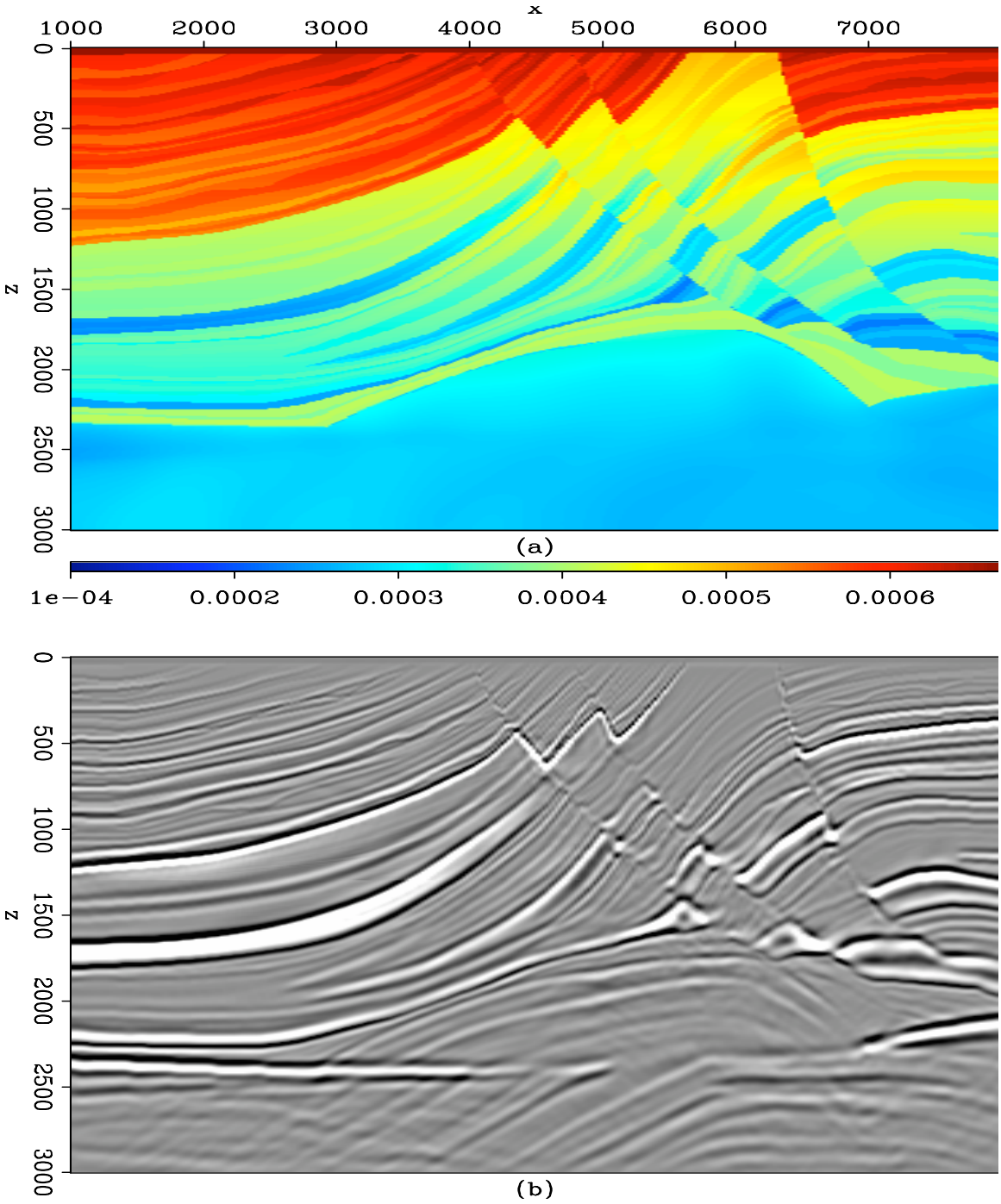

Figure 5: (a) Optimized slowness after 4 iterations with 41 function and gradient evaluations. (b) Zero-subsurface offset section of the one-way shot-profile migration with the optimized slowness model.[**CR**]

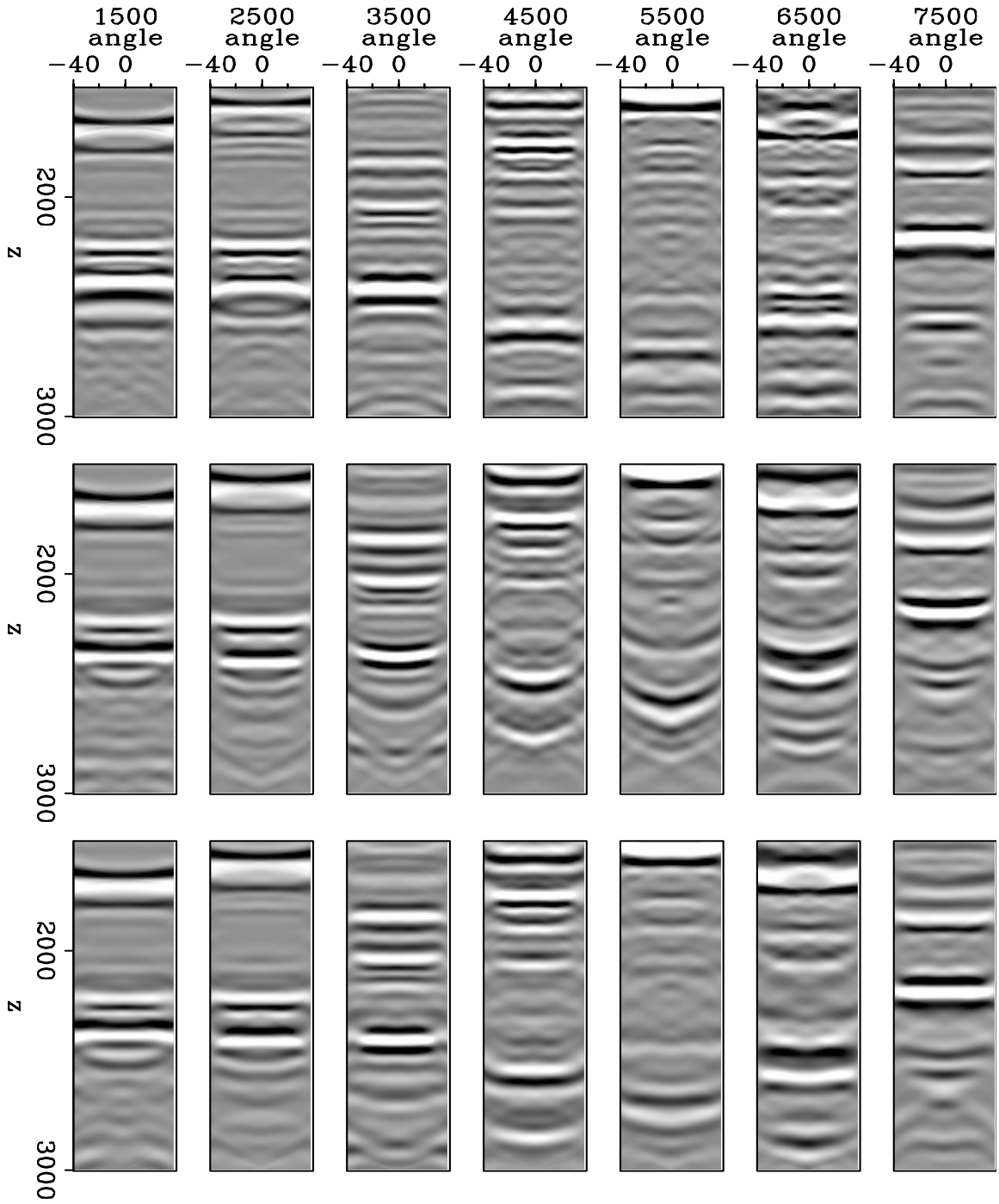

Figure 6: ADCIGs obtained with true slowness (top), background slowness (center), and optimized slowness (bottom).[**CR**]

used the augmented Lagrangian methodology, which clarifies the action of the different operators that compose the image-space wave-equation tomography operator. The derivation is valid whether one considers migration of shot-profiles or generalized sources. Image-space phase-encoded gathers significantly accelerate image-space wave-equation tomography. In addition to decreasing the computational cost, the numerical example shows that image-space phase-encoded gathers can be used to generate sufficiently accurate optimized slowness.

REFERENCES

Biondi, B., 2006, Prestack exploding-reflectors modeling for migration velocity analysis: 76th Ann. Internat. Mtg., Expanded Abstracts, 3056–3060, Soc. of Expl. Geophys.

———, 2007, Prestack modeling of image events for migration velocity analysis: **SEP-131**, 101–118.

———, 2008, Automatic wave-equation migration velocity analysis: **SEP-134**, 65–78.

Chavent, G. and C. A. Jacewitz, 1995, Determination of background velocities by multiple migration fitting: Geophysics, **60**, 476–490.

Guerra, C. and B. Biondi, 2008a, Phase-encoding with Gold codes for wave-equation migration: **SEP-136**.

———, 2008b, Prestack exploding reflector modeling: The crosstalk problem: **SEP-134**, 79–92.

Guerra, C., Y. Tang, and B. Biondi, 2009, Wave-equation tomography using image-space phase-encoded data: **SEP-138**, 95–116.

Lines, L. and S. Treitel, 1984, A review of least-squares inversion and its application to geophysical problems: Geophys. Prospecting, **32**, 159–186.

Loewenthal, D., L. Lu, R. Roberson, and J. Sherwood, 1976, The wave equation applied to migration: Geophys.Prosp., **24**, 380–399.

Nocedal, J. and S. Wright, 2000, Numerical optimization: Springer Verlag,New York.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophys. J. Int., **167**, 495–503.

Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, 426–436.

Sava, P. and B. Biondi, 2004a, Wave-equation migration velocity analysis-I: Theory: Geophysical Prospecting, **52**, 593–606.

———, 2004b, Wave-equation migration velocity analysis-II: Examples: Geophysical Prospecting, **52**, 607–623.

Shen, P., 2004, Wave-equation Migration Velocity Analysis by Differential Semblance Optimization: PhD thesis, Rice University.

Shen, P. and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.

Shen, P., W. W. Symes, and C. C. Stolk, 2003, Differential semblance velocity analysis by wave-equation migration: SEG Technical Program Expanded Abstracts, **22**, 2132–2135.

Symes, B., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790.

Tang, Y., C. Guerra, and B. Biondi, 2008, Image-space wave-equation tomography in the generalized source domain: **SEP-136**, 1–22.

Tarantola, A., 1987, Inverse problem theory: Methods for data fitting and model parameter estimation: Elsevier.

Whitmore, N. D., 1995, An Imaging Hierarchy for Common Angle Plane Wave Seismogram: PhD thesis, University of Tulsa.

Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, 15–26.