GEOPHYSICS, VOL. 78, NO. 4 (JULY-AUGUST 2013); P. S179–S192, 16 FIGS.
10.1190/GEO2012-0064.1

Automatic velocity analysis with reverse-time migration

Wiktor Waldemar Weibull[1] and Børge Arntsen[1]

ABSTRACT

We apply a method to automatically estimate the background velocities using reverse-time migration. The method uses a combination of differential semblance and similarity-index (a.k.a., "semblance" or "stacking-power") to measure the focusing error in imaging and a nonlinear optimization procedure to obtain the background velocities. A challenge in this procedure is that, for media consisting of complex and strongly refracting velocities, artifacts in the reverse-time migrated image (low-frequency noise) can cause the velocity analysis to diverge. We successfully overcome this issue by applying a simple vertical derivative filter to the image that is input to velocity analysis. The resultant velocity analysis method is tested in two 2D synthetic examples and one 2D field data example. Due to the assumptions inherent to prestack depth migration, the data that are input to velocity analysis must be singly scattered. To apply the method to multiple-rich data, we propose an image-based demultiple method. The method consists of muting events in the subsurface offset common image point gathers constructed with reverse-time migration, and remodeling the data using a kinematic demigration. A field data example shows how the image-based demultiple of the data helps to improve the velocity analysis in the presence of multiple scattering.

INTRODUCTION

An accurate estimate of the distribution of the subsurface seismic velocities is an indispensable component for obtaining an accurate image of Earth's reflectivity by prestack depth migration. Methods for velocity estimation in routine use today are often based on ray tomography and, although usually semiautomatic, require time-consuming picking of gathers and quality control. Due to the shortcomings of ray-theoretical depth-migration approaches in areas with complex geology (Arntsen et al., 2009), one-way — and more recently, two-way — wave equation imaging methods have become popular. These approaches are often combined with velocity estimation based on ray theory. For consistency and improved resolution, velocity model building and seismic imaging should preferably be based on wave equation methods.

Wave equation migration velocity analysis (WEMVA) is based on focusing of seismic reflection data in the image domain and uses an automatic optimization procedure to estimate the velocity field, avoiding manual picking.

The approach formulates an objective function that measures the extent to which subsurface offset- or angle-gathers are either focused or flattened, and then that function is minimized with respect to the velocity field.

Chavent and Jacewitz (1995) implement WEMVA by using a similarity-index and reverse-time migration (RTM) to compute the velocity field. The procedure uses the complete wavefield and requires no picking. Biondi and Sava (1999) use one-way migration operators and image perturbations for computing corrections to the initial wavefield and Sava and Biondi (2004) extend this approach to a fully nonlinear iterative scheme. The numerical implementation is described in detail by Sava and Vlad (2008).

Shen et al. (2003) use the double square root approach to depth migration and an objective function based on differential semblance Symes and Carazzone (1991) to estimate the velocity field. The approach was extended to shot-profile migration based on one-way migration operators, through an objective function consisting of the difference between differential semblance and similarity-index (Shen and Symes, 2008). Mulder (2008) used depth migration based on the two-way wave equation in the frequency-domain and an objective function related to the differential semblance cost function divided by the similarity-index cost function to implement a nonlinear scheme for computing the velocity field. Gao and Symes (2009) propose to use a differential semblance cost function and RTM to solve the velocity estimation problem, and also give an initial theoretical framework.

We implement an objective function for WEMVA using differential semblance, similarity-index, and RTM. To minimize

Manuscript received by the Editor 21 February 2012; revised manuscript received 9 January 2013; published online 3 June 2013.
[1]Norwegian University of Science and Technology, Institute for Petroleum Technology and Applied Geophysics, Trondheim, Norway. E-mail: wiktor.weibull@ntnu.no; borge.arntsen@ntnu.no.

problems related to amplitude-sensitivity caused by the two-way wave equation, we choose to modify the image by a spatial differentiation operator, similar to but simpler than the filter proposed by Mulder (2008). As we will show, this leads to an objective function with improved convergence properties. We also give complete expressions for computing the gradient of the cost function with respect to the velocity field, and show how the gradient can be used in a full nonlinear optimization scheme illustrated with synthetic and real data examples.

One challenge in applying WEMVA to field data is the presence of free-surface multiples (van Leeuwen and Mulder, 2008; Mulder and van Leeuwen, 2008). To overcome this problem, Mulder and van Leeuwen (2008) modify their objective function with an asymmetric weighting function. We extend this idea and develop a method of multiple attenuation based on RTM, muting, and data reconstruction, and show how this method helps in improving the results of velocity analysis in the case of multiple-rich data.

REVERSE-TIME MIGRATION AND VELOCITY ANALYSIS

Migration

In RTM, a common image-point gather (CIG), $\mathcal{R}$, can be produced by crosscorrelating a forward modeled source wavefield (D) with a reverse-time modeled scattered wavefield (U)

$$\mathcal{R}(\mathbf{x},\mathbf{h}) = \sum_s \int dt\, U_s(\mathbf{x}+\mathbf{h},t) D_s(\mathbf{x}-\mathbf{h},t), \qquad (1)$$

where $\mathbf{x} = (x_1, x_2, x_3)$ is the subsurface midpoint coordinate, with x_3 being the vertical depth axis coordinate, $\mathbf{h} = (h_1, h_2, 0)$ is the subsurface horizontal half-offset, t is the time, and s is the source index.

This imaging condition is an extension of Claerbout's principle (Claerbout, 1971; Rickett and Sava, 2002). According to this principle, given an accurate estimate of the material velocities, the crosscorrelation of the reconstructed source and receiver wavefields will have a maximum at zero lag in time and space. By parameterizing the image with an additional lag parameter, we can capture the deviation of the maximum in crosscorrelation from zero lag, and use this to quantify the error in the estimates of the velocities. In this paper we only consider horizontal spatial lags in the crosscorrelation, however vertical spatial lags, and even temporal lags can be incorporated in the image (Biondi and Shan, 2002; Biondi and Symes, 2004; Sava and Fomel, 2006). The U_s and D_s wavefields are obtained through

$$D_s(\mathbf{x},t) = \int d\mathbf{x}' G(\mathbf{x},t;\mathbf{x}',0) * \sum_{\text{sou}=1}^{N\text{sou}} \delta(\mathbf{x}'-\mathbf{x}_{\text{sou}}) S_s(\mathbf{x}',t), \qquad (2)$$

and

$$U_s(\mathbf{x},t) = \int d\mathbf{x}' G(\mathbf{x},0;\mathbf{x}',t) * \sum_{\text{rec}=1}^{N\text{rec}} \delta(\mathbf{x}'-\mathbf{x}_{\text{rec}}) P_s(\mathbf{x}',t), \qquad (3)$$

where G is the Green's function for the constant-density two-way acoustic wave equation, and * means time convolution. Here, P is the recorded reflection data and S is the source data, while $\mathbf{x}_{\text{sou}}$ are source datum coordinates for one shot, and $\mathbf{x}_{\text{rec}}$ are the receiver datum coordinates for one shot.

Velocity analysis

The velocity analysis is based on the optimization of the following objective function:

$$\mathcal{J} = \mathcal{DS} - \mathcal{SI} + \mathcal{J}_{\text{reg}}. \qquad (4)$$

The objective function is composed of three parts; the differential semblance misfit ($\mathcal{DS}$), the similarity-index ($\mathcal{SI}$) and a regularization term ($\mathcal{J}_{\text{reg}}$).

An important assumption in WEMVA is that when the velocity model is optimum, the CIGs are maximally focused at zero subsurface offset. Thus, any deviation from perfect focus can be taken as an indication that the background velocity model must be improved. The differential semblance misfit represents a simple and direct way of quantifying the deviation from focus of CIGs,

$$\mathcal{DS} = \frac{1}{2}\left\| \mathbf{h}\frac{\partial \mathcal{R}}{\partial x_3}(\mathbf{x},\mathbf{h})\right\|^2 = \frac{1}{2}\int d\mathbf{x}\int d\mathbf{h}\,\mathbf{h}^2\left[\frac{\partial \mathcal{R}}{\partial x_3}(\mathbf{x},\mathbf{h})\right]^2. \qquad (5)$$

This differential semblance formulation is similar to that in Shen and Symes (2008). However, the fact that we use RTM to construct the image makes a fundamental difference. The solution of the one-way wave equation neglects any scattering from vertical contrasts in the velocities, whereas the same is not true for the solution of the two-way wave equation. The scattering during wavefield reconstruction produces undesired artifacts in the final RTM image. In the context of velocity analysis, it is desirable to remove these artifacts because they are coherent events and particularly sensitive to changes in the velocities, thus affecting the results of the optimization. Due to the low-wavenumber character and the predominantly vertical orientation of this effect, a simple vertical derivative filter acting over the image is sufficient to remove it (Guitton et al., 2007). A similar procedure, albeit involving multiplication in the wavenumber domain, has been applied by Mulder (2008) in his implementation of WEMVA using the two-way wave equation.

The differential semblance measure is based solely on kinematic considerations. To exploit the dynamic effect that the velocities can have through the improvement of the stack quality, the objective function can be augmented with the similarity-index (a.k.a., "stacking-power" or "semblance") (Chavent and Jacewitz, 1995)

$$\mathcal{SI} = \frac{\gamma}{2}\left\| \frac{\partial \mathcal{R}}{\partial x_3}(\mathbf{x},0)\right\|^2 = \frac{\gamma}{2}\int d\mathbf{x}\left[\frac{\partial \mathcal{R}}{\partial x_3}(\mathbf{x},0)\right]^2, \qquad (6)$$

where γ is a constant that balances the weight of $\mathcal{SI}$ over $\mathcal{DS}$. Ideally, the weight should be chosen such that the similarity-index only acts as a regularization.

Finally, to further improve the well posedness of the velocity analysis we also apply regularization (Tikhonov and Arsenin, 1977),

$$\mathcal{J}_{\text{reg}} = \frac{\alpha}{2}\left\| \frac{\partial v}{\partial \mathbf{x}}(\mathbf{x})\right\|^2 + \frac{\beta(\mathbf{x})}{2}\|v(\mathbf{x}) - v_{\text{prior}}(\mathbf{x})\|^2, \qquad (7)$$

where α is a scalar constant, β is a vector of constant weights, and v_{prior} is the vector containing a priori known values of velocity.

The optimization of equation 4 is performed using a gradient-based nonlinear method (Byrd et al., 1995; Nocedal and Wright, 2000). The gradient of the misfit function with respect to velocity is then required.

The gradient of equation 4 is given by

$$\nabla_v \mathcal{J}(\mathbf{x}) = \nabla_v \mathcal{DS}(\mathbf{x}) - \nabla_v \mathcal{SI}(\mathbf{x}) + \nabla_v \mathcal{J}_{\text{reg}}(\mathbf{x}). \quad (8)$$

The adjoint state method (Lions and Magenes, 1972; Chavent and Lemonnier, 1974; Chavent, 2009) offers an exact and efficient way to compute the gradients of the differential semblance misfit and the similarity-index functions. In Appendix A, we show how to derive the gradient of equations 5 and 6 with respect to the acoustic velocities by the adjoint state method. In this case the gradient is given by

$$\nabla_v(\mathcal{DS} - \mathcal{SI})(\mathbf{x}) = -\sum_s \int dt \frac{2}{v^3(\mathbf{x})} \frac{\partial^2 D_s}{\partial t^2}(\mathbf{x},t) D'_s(\mathbf{x},t) - \sum_s \int dt \frac{2}{v^3(\mathbf{x})} \frac{\partial^2 U_s}{\partial t^2}(\mathbf{x},t) U'_s(\mathbf{x},t). \quad (9)$$

The wavefields D'_s and U'_s are adjoint states associated with the constraints that the direct states (D_s and U_s) satisfy the constant density acoustic wave equation. These wavefields can be computed by the following adjoint modelings:

$$D'_s(\mathbf{x},t) = \int \mathbf{dx}' G(\mathbf{x},0;\mathbf{x}',t) * \int d\mathbf{h}(\mathbf{h}^2 - \gamma\delta(\mathbf{h})) \frac{\partial^2 \mathcal{R}}{\partial x'^2_3}(\mathbf{x}'+\mathbf{h},\mathbf{h}) U_s(\mathbf{x}'+2\mathbf{h},t), \quad (10)$$

and

$$U'_s(\mathbf{x},t) = \int \mathbf{dx}' G(\mathbf{x},t;\mathbf{x}',0) * \int d\mathbf{h}(\mathbf{h}^2 - \gamma\delta(\mathbf{h})) \frac{\partial^2 \mathcal{R}}{\partial x'^2_3}(\mathbf{x}'-\mathbf{h},\mathbf{h}) D_s(\mathbf{x}'-2\mathbf{h},t), \quad (11)$$

where δ is the Kronecker delta.

The cost of computing the gradient in this way is approximately the same as that of evaluating the misfit function; however, it requires the state variables (D and U) to be stored for each shot, which can be expensive. In the discussion, we suggest some measures to reduce this cost.

To complete the gradient of $\mathcal{J}$ we need to compute the gradient with respect to the regularization term

$$\nabla_v \mathcal{J}_{\text{reg}}(\mathbf{x}) = \beta(\mathbf{x})(v(\mathbf{x}) - v_{\text{prior}}(\mathbf{x})) - \alpha \frac{\partial}{\partial \mathbf{x}} \left[\frac{\partial v}{\partial \mathbf{x}}(\mathbf{x}) \right], \quad (12)$$

where $\partial v/\partial \mathbf{x}$ is taken to be zero at the boundaries.

Velocity preconditioning

The solution of differential semblance optimization is notoriously rough. Fei and Williamson (2010) indicate that the updated velocities can have artificial roughness features (vertical stripes). To ensure a smooth solution to the velocity analysis, we precondition the velocity model. In addition to speeding up the convergence, preconditioning also helps to make the velocity analysis well posed, as it reduces the space of possible solutions and the number of parameters to be estimated. A popular choice of representing a velocity model is given by the cubic B-spline representation (Dierckx, 1993)

$$v(\mathbf{x}) = \sum_{\mathbf{i}} c_{\mathbf{i}} B_{\mathbf{i}}(\mathbf{x}), \quad (13)$$

where $B_{\mathbf{i}}$ are cubic splines defined at predetermined control points and $c_{\mathbf{i}}$ are coefficients to be determined by the velocity analysis. One advantage of using this representation is that it ensures continuous second derivatives, which is important for the derivative regularization implementation. At the same time, due to the local support of the cubic spline functions, this representation is also well-suited to describe the spatial variations necessary for velocity analysis. When using the B-spline representation for optimization, the gradient must be transformed from the Cartesian to the spline basis,

$$\nabla_c \mathcal{J}(\mathbf{i}) = \int d\mathbf{x} B_{\mathbf{i}}(\mathbf{x}) \nabla_v \mathcal{J}(\mathbf{x}). \quad (14)$$

RESULTS

Synthetic data examples

We present the results of optimizing two 2D synthetic data sets. Both synthetic data sets were generated using 2D finite difference modeling, with synthetic density and acoustic velocity models. The geometry simulates a typical marine acquisition. We use a monopole point source and a Ricker wavelet with a peak frequency of 25 Hz. The modeling was carried out with an algorithm that is eighth-order accurate in space and second-order accurate in time (Virieux, 1986). The modeling aperture is taken as twice the cable length. To avoid reflections from the boundaries of the models, PML absorbing boundary conditions were implemented at all sides (Qin et al., 2009). This way, the resulting data sets are devoid of free-surface multiples. Although interbed multiples are still present in the data, they are significantly weaker than the primaries and are neglected in these examples. Preprocessing of the data sets consisted in muting the direct wave and the refracted waves at the receiver level.

Optimization is carried out with an L-BFGS method (Byrd et al., 1995). In both examples, the regularization parameter α was spatially invariable, and its value was chosen such that the initial derivative regularization error value was 1% of the initial $\mathcal{DS}$ value. The value of parameter β was chosen in a similar way as α, but it was set to zero outside of the topmost layer.

Shallow gas accumulation and leakage model

The first example consists of a 4-km long and 1-km deep model shown in Figure 1. This model simulates a dipping layered sediment

succession with a small reef-like structure in the center. The density varies between 1800 and 2400 kg/m^3. The velocity model has the same structure as the density model, but with a localized low-velocity lens (Gaussian with a peak of −500 m/s) under the reef structure.

This model simulates a scenario of shallow gas accumulation and leakage, a typical situation that, if not accurately predicted by the velocity model, can produce significant distortion on depth images.

The acquisition geometry simulates a 2D marine acquisition with 311 sources separated by 20 m. The cable length is 3.2 km with 10-m channel interval, with minimum offset of 500 m. Recording time length is 2 s.

The initial model used for migration consists of a linear 1D velocity profile varying from 1.8 to 2.4 km/s as shown in Figure 2a. For optimization, a bicubic B-spline spline with control points every 40 m in the x- and z-directions was then fitted to this initial model. The parameter γ was chosen so that the initial $\mathcal{SI}$ value was 20% of the initial $\mathcal{DS}$ value.

The CIGs are produced according to equation 1 with the half-offset axis varying between ±400 m (81 offset samples).

The quality of our initial estimate of the velocity can be quantified by looking at the initial image shown in Figure 2b and the CIGs output by RTM displayed in Figure 2c. The image is obtained from the CIGs by taking only the zero subsurface offset component (i.e., zero lag crosscorrelation), and shows significant distortion, especially below the low-velocity lens. At the same time, the energy in CIGs is significantly spread across the offset axis.

To justify the use of the spatial derivative in our modified differential semblance misfit function, we make two attempts at estimating velocities from this data. One without the spatial

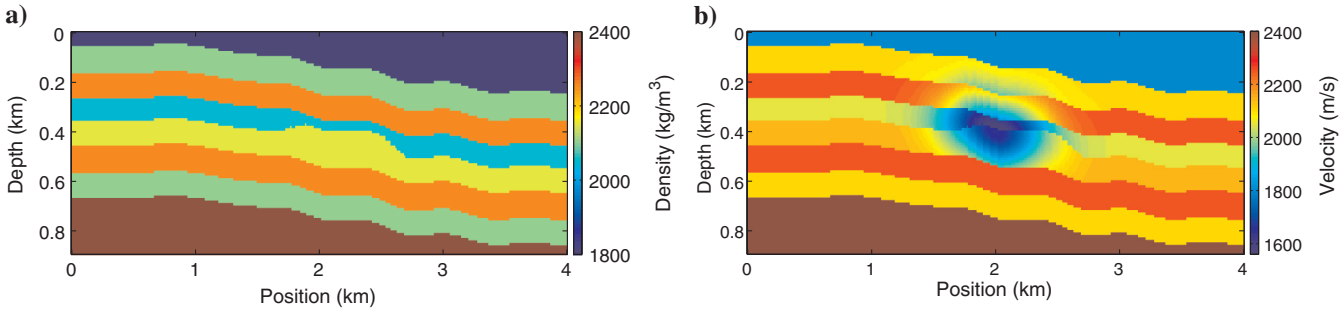

Figure 1. Shallow gas accumulation and leakage synthetic model: (a) density model (kg/m^3) and (b) acoustic velocity model (m/s).

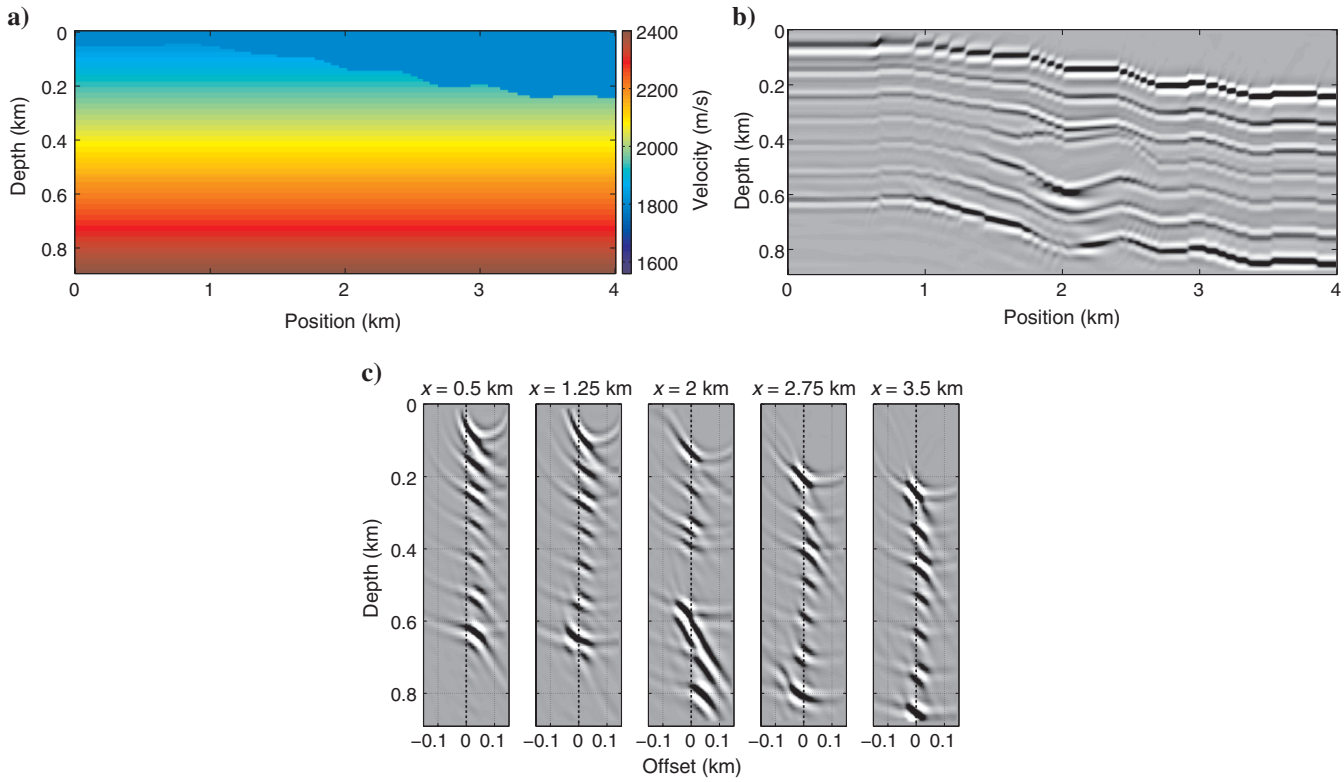

Figure 2. (a) Initial velocity model (m/s). (b) Initial image. (c) Initial subsurface offset CIGs.

derivative, and other with our modified differential semblance objective.

In the first attempt, we did not apply the spatial derivative. Optimization was stopped after 30 iterations because, despite the decrease in value of the objective function, the quality of the migrated image was getting worse with increasing number of iterations. This was taken as an indication that velocity analysis was converging to a nonoptimal minimum. The updated velocities are shown in Figure 3a, the updated image is shown in Figure 3b and the updated subsurface CIGs are shown in Figure 3c. The optimization fails here because we allow overly strong spatial variations in the updated velocities. These introduce artifacts in the RTM image that are more sensitive to the velocities than the kinematic errors that we want to correct for.

In the second attempt, we introduce the spatial derivative in the image. Now optimization converges after 17 iterations. The optimized velocity model is shown in Figure 4a, and it is clear that the low-velocity lens and the layered structure of the velocity model are successfully detected.

Figure 4b shows the results of the optimization on the image, whereas Figure 4c shows the optimized CIGs. The image migrated with the updated velocities is much better focused and the reflectors are now well positioned. At the same time, the CIGs are well-focused at zero subsurface offset, indicating that the velocities are adequate to describe the kinematics of the data.

Figure 5a shows a comparison of the initial, updated, and true traces of velocity at different spatial positions, while Figure 5b does the same for the reflectivity. The velocity traces show that, locally, the updated velocities can deviate significantly from the true velocities, though without compromising the reflectivity fit. This demonstrates the nonuniqueness inherent to the solution of this type of problem. From the reflectivity, we can see that the method is able to correct mispositioning errors that are, at times, larger than half-wavelength (e.g., Figure 5b at $x = 2$ km).

In yet another experiment with this data set, we compare the results of velocity analysis using the differential semblance misfit alone, the similarity index alone and their combination. For this test, a bicubic B-spline with control points every 100 m in the x-direction and 50 m in the z-direction was fitted to the initial model. In the combination result, the parameter γ is chosen such that the initial $\mathcal{SI}$ value is equal to 50% of the initial $\mathcal{DS}$ value.

The updated velocities obtained with the three different objective functions are shown in Figure 6, while the updated images are displayed in Figure 7. The results show that, in this example, all three objective functions converge to models that improve the quality of the initial migrated image. As can be seen in Figure 6a and 6b, the velocities obtained from velocity analysis based on $\mathcal{SI}$ and $\mathcal{DS}$ can be quite different. This reflects, in part, the fact that these objective functions have different sensitivities to the initial model. However, it is also a consequence of using different sources of information to constrain the velocity model.

Combining both objective functions seems to have the effect of averaging out artifacts and strengthening similarities between the different velocity models (Figure 6c). This helps to improve the quality of the final migrated image, as can be seen in Figure 7c.

Gullfaks model

The second data set is generated from a 2D synthetic model of the Gullfaks oil field, located in the Norwegian margin of the North

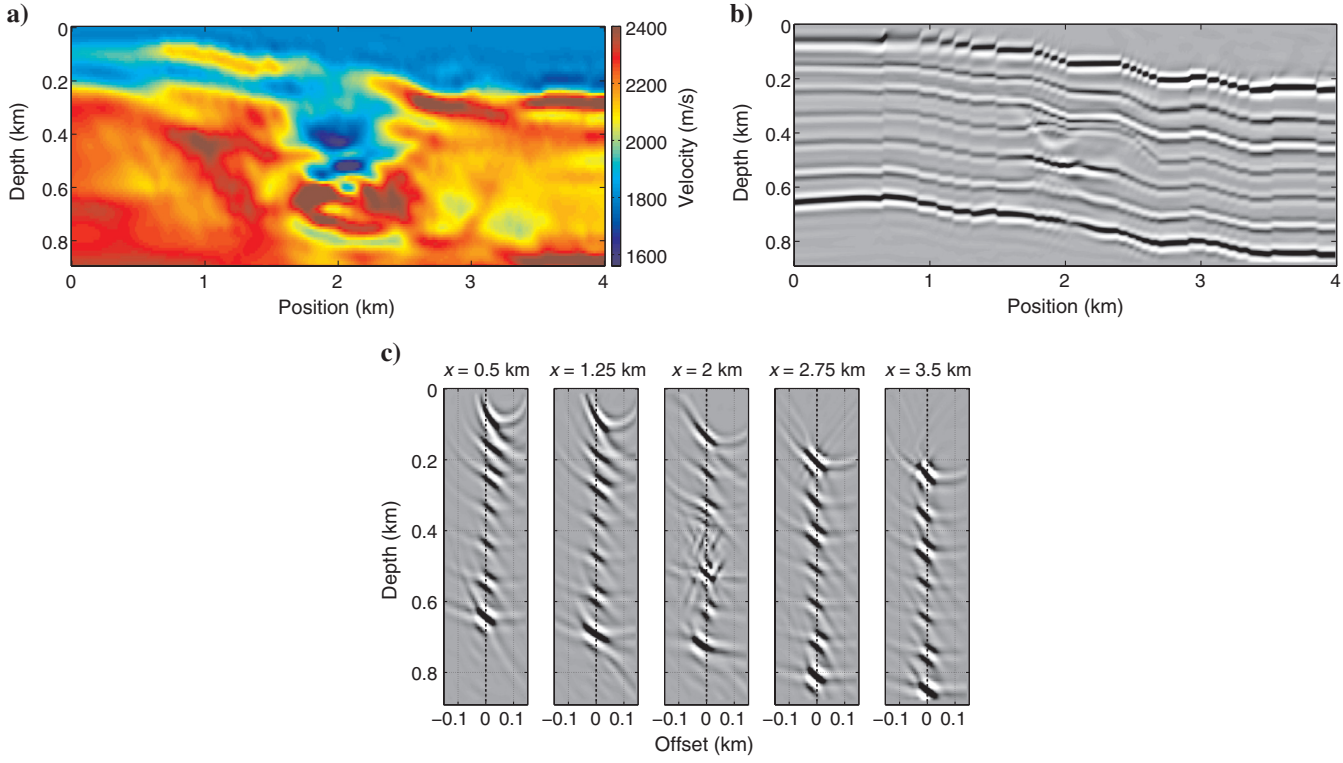

Figure 3. Updated velocity model (a), updated image (b), and updated subsurface offset CIGs (c) after 30 iterations of velocity analysis using the objective function without the spatial derivative.

Sea. The model is 3 km deep and 4 km across. At the top of the model, there is a 200-m deep water layer. Beneath the water layer and down to about 1.8 km, the model consists largely of a layered sediment succession. The layered overburden unconformably overlies a set of rotated fault blocks. The density varies between 1200 and 2400 kg/m^3 (Figure 8a), while the velocity varies between 1.48 and 3.4 km/s (Figure 8b).

The geometry of the data in this example is a half-spread marine setup with offsets ranging from 0 to 6 km, and spaced every 12.5 m. The recording time length is 4 s.

The initial model used for migration consists of a 1D velocity profile linearly varying from 1.48 to 3.4 km/s (Figure 9a). This model carries large deviations from the true model, and these stretch over several hundred meters. This results in significant traveltime

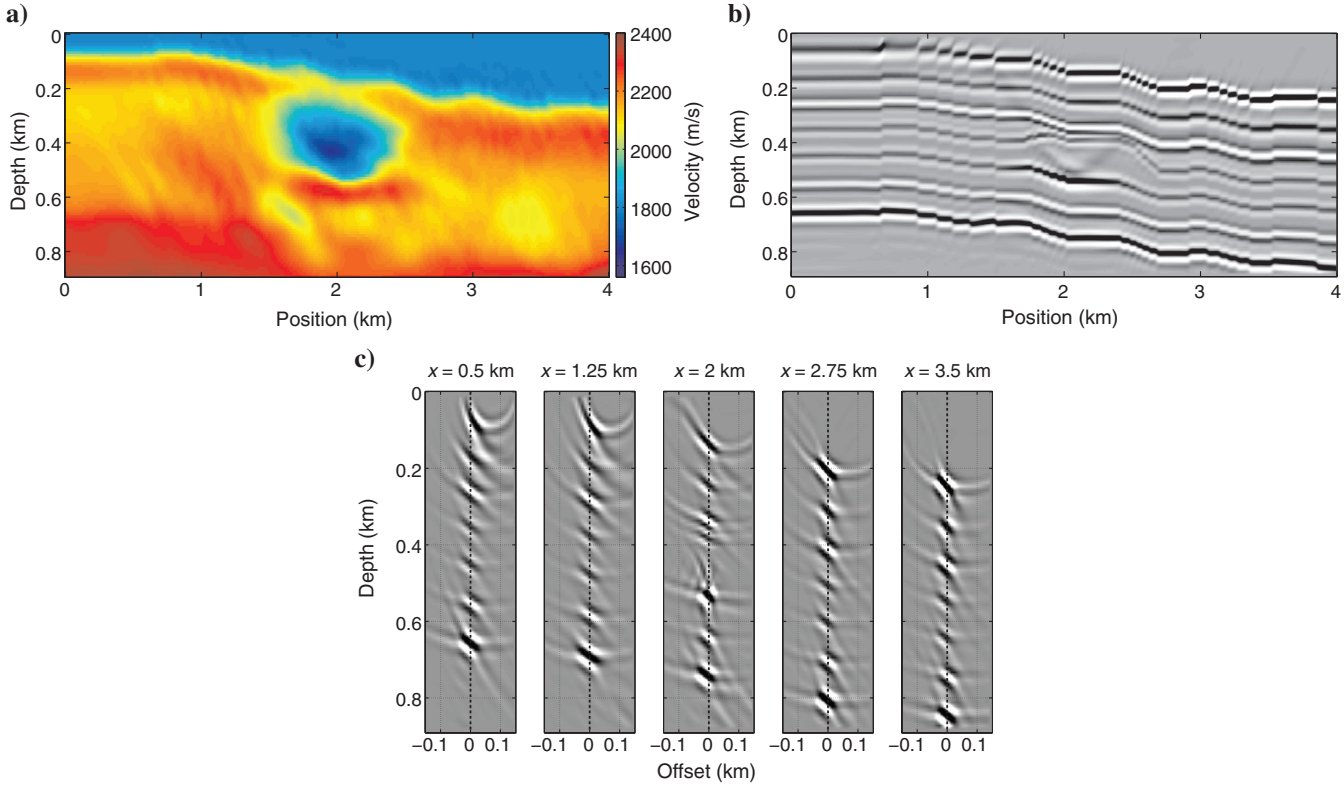

Figure 4. Updated velocity model (a), updated image (b), and updated subsurface CIGs (c) after 17 iterations of velocity analysis with the spatial derivative.

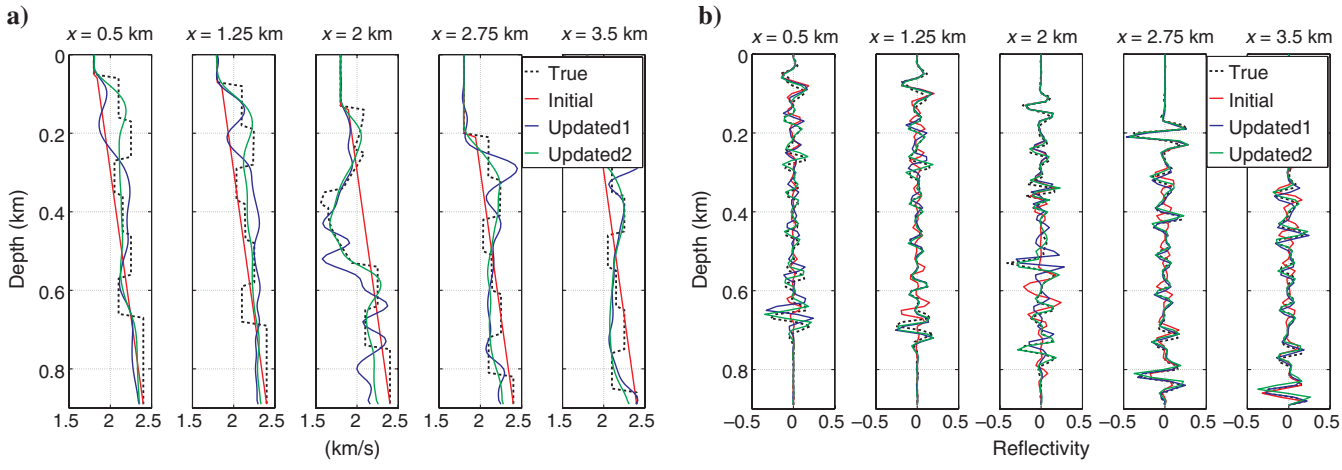

Figure 5. (a) Comparison of velocity traces. (b) Comparison of reflectivity traces. Updated1 corresponds to the updated model and reflectivity from the optimization of the objective function without the vertical derivative, while updated2 comes from the optimization of the objective function with the vertical derivative.

errors (more than one wavelength at the considered frequencies). For optimization, a bicubic B-spline with control point spacings of 150 m in the x-direction and 50 m in the z-direction was fitted to this model. The parameter γ was chosen so that the initial $\mathcal{SI}$ value was 20% of the initial $\mathcal{DS}$ value. Migration and optimization are carried out in a similar way as in the first example.

The results of migration with the initial model are shown in Figure 9b and 9c.

Despite the simplicity of the Gullfaks model, consisting mostly of flat layers with only mild variations in velocity, the convergence rate for this example is relatively slow. We interpret this as a consequence of the fact that there are not many reflectors in the upper

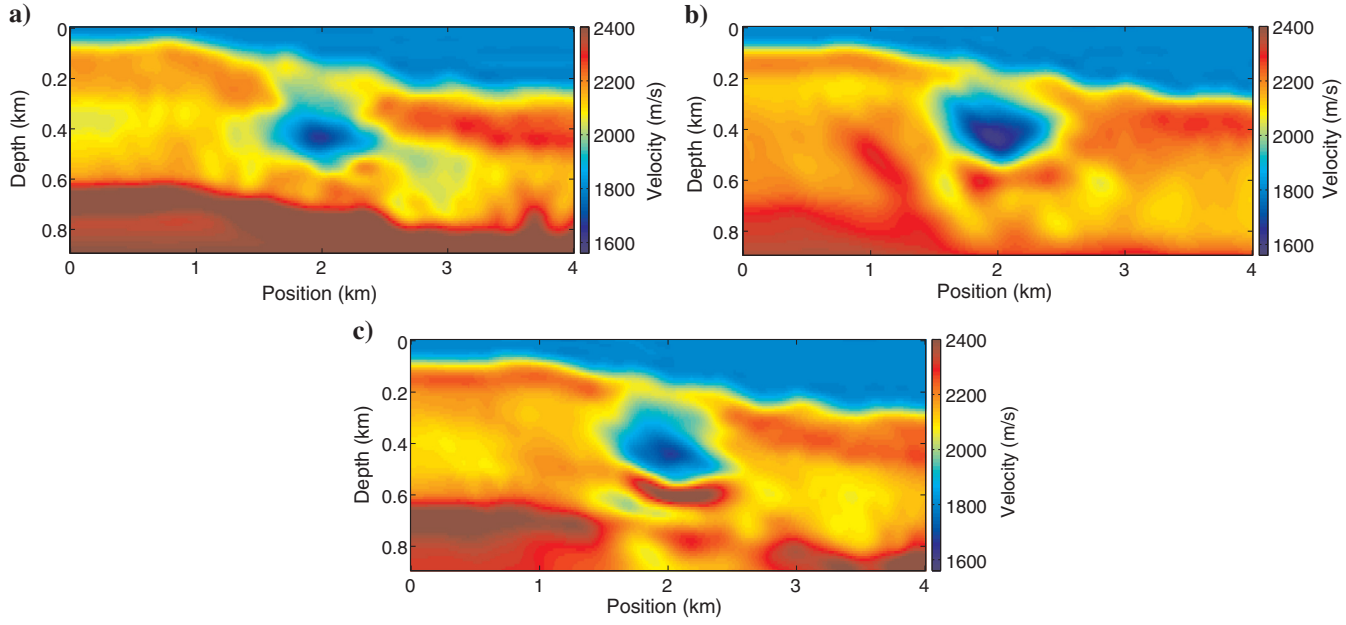

Figure 6. Updated velocity model using: (a) similarity-index alone; (b) differential semblance alone; (c) differential semblance and similarity-index combined.

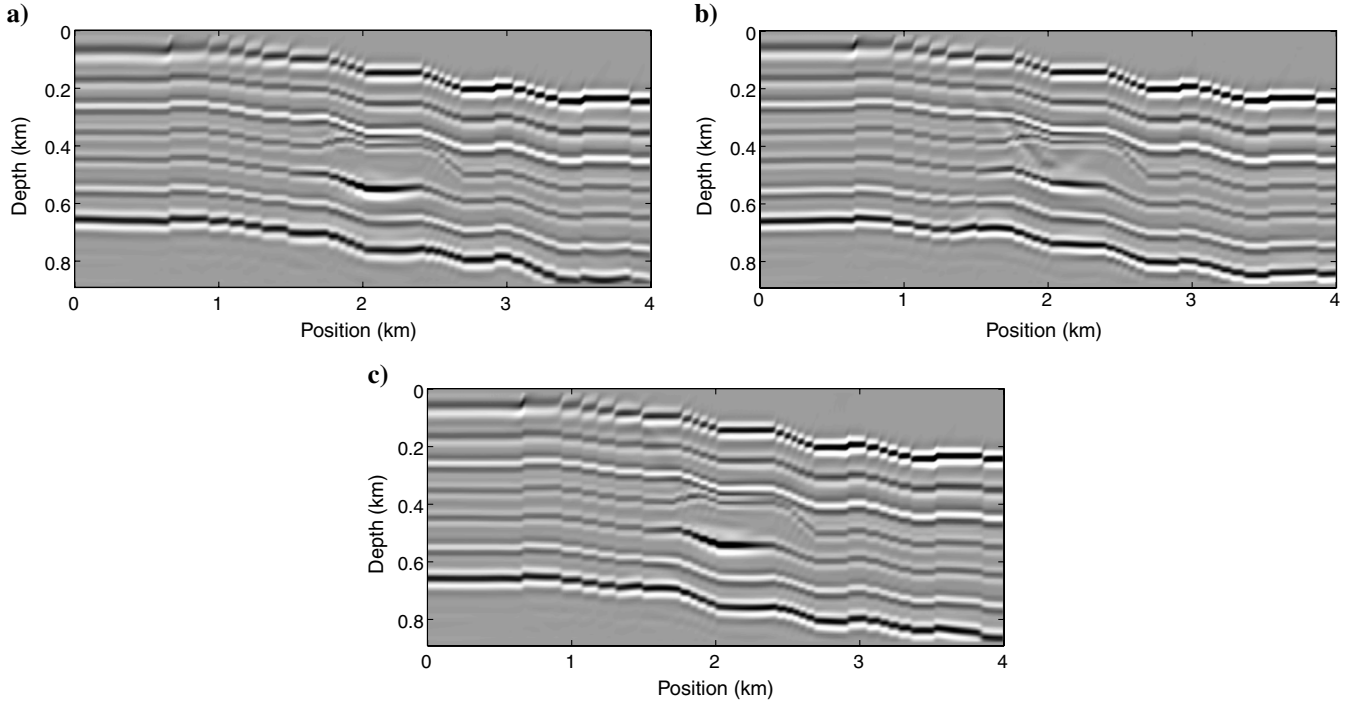

Figure 7. Updated image using: (a) similarity-index alone; (b) differential semblance alone; (c) differential semblance and similarity-index combined.

1.8 km of the model, which makes the inversion poorly constrained. Nevertheless, the optimization is stopped after 50 iterations, at which point the CIGs were deemed sufficiently focused.

The optimized velocity model is shown in Figure 10a. The results of optimization show that, again, in this case the method is capable of improving the kinematics of the velocity model. This can be clearly seen if we compare the initial images, shown in Figure 9b and 9c, to the updated images, shown in Figure 10b and 10c. At the same time, Figure 11a and 11b shows that the updated velocities and reflectivities are close to their true value (from the synthetic model).

Field data example

The method is tested on a field data set taken off the Norwegian North Sea. The data are originally a 3D data set, from which we extracted a 2D line. The geometry of the data consists of a line with minimum offset of 150 m and maximum offset of 5 km. The original receiver interval is 25 m and the original shot interval is 18.7 m. The data processing included multiple removal, and muting of direct wave, wide-angle reflections and refractions. To increase the contribution from far offsets and deeper events, a power of two time gain (t^2) is applied to the data. The maximum frequency

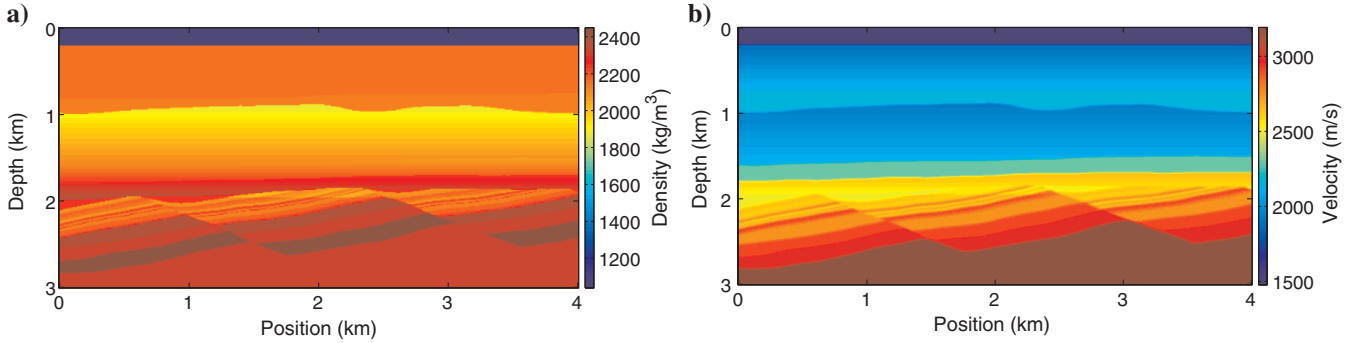

Figure 8. Gullfaks synthetic model: (a) density model (kg/m^3) and (b) acoustic velocity model (m/s).

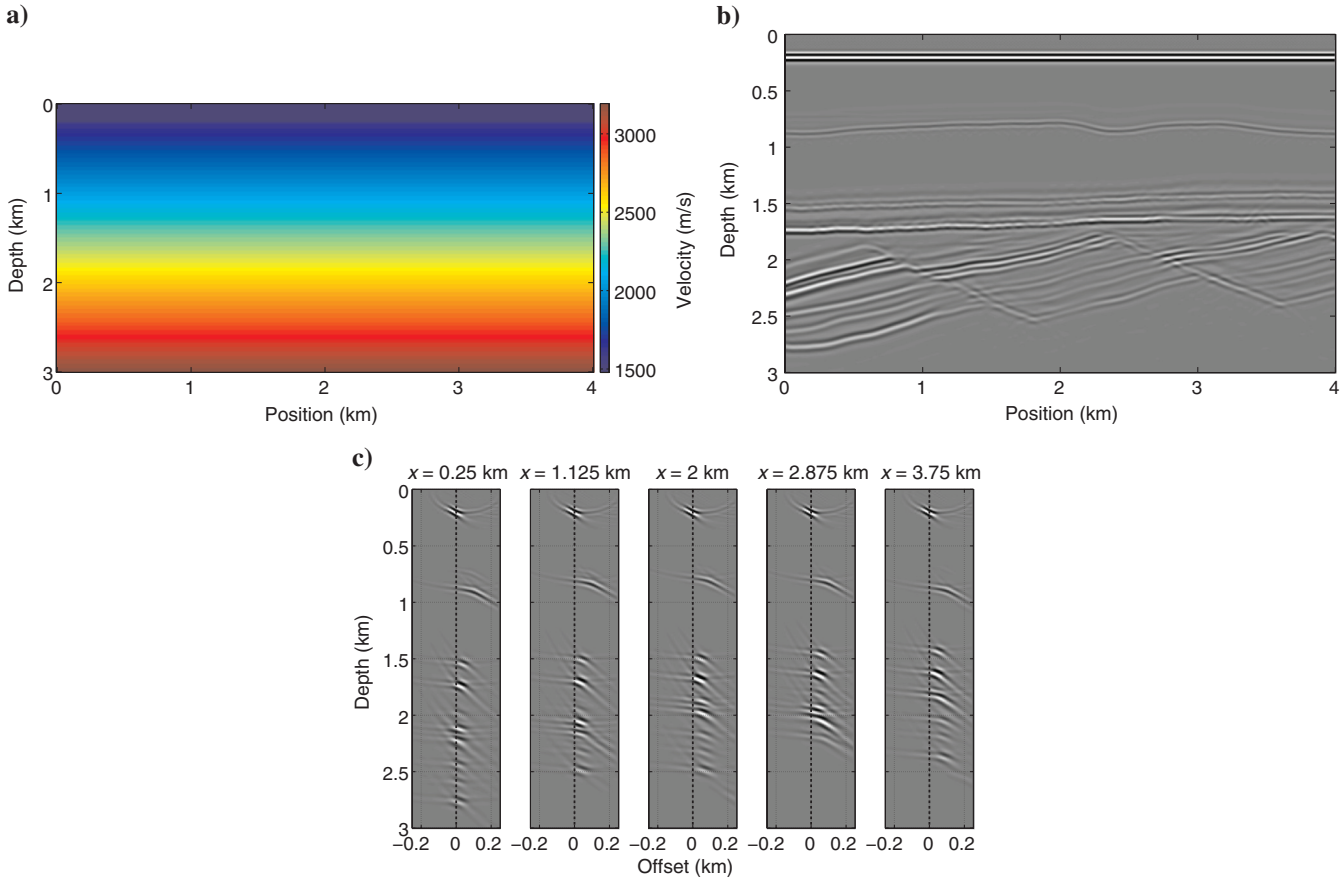

Figure 9. (a) Initial velocity model (m/s). (b) Initial image. (c) Initial subsurface offset CIGs.

of the data was filtered down to 30 Hz, so that a grid of 20 × 20 m could be used for modeling and migration. However, for display purposes, the updated images are migrated on a finer grid of 10 × 10 m using frequencies up to 80 Hz. The choice of regularization parameters (α, β and γ) for optimization follow the same guidelines as in the synthetic examples.

The starting point for the velocity analysis is a 1D velocity model shown in Figure 12a. The model is constructed from a smoothed well log of P-wave velocities. For optimization, this initial model was fitted to a B-spline representation using a grid of control points spaced 600 m in the x-direction and 100 m in the z-direction.

The initial image is shown in Figure 12b. Due to the approximately plane-layered overburden, the initial image shows relatively well-focused reflectors. Indeed, the biggest challenge for velocity analysis in this data is the presence of resilient free-surface multiples that were not properly attenuated in preprocessing.

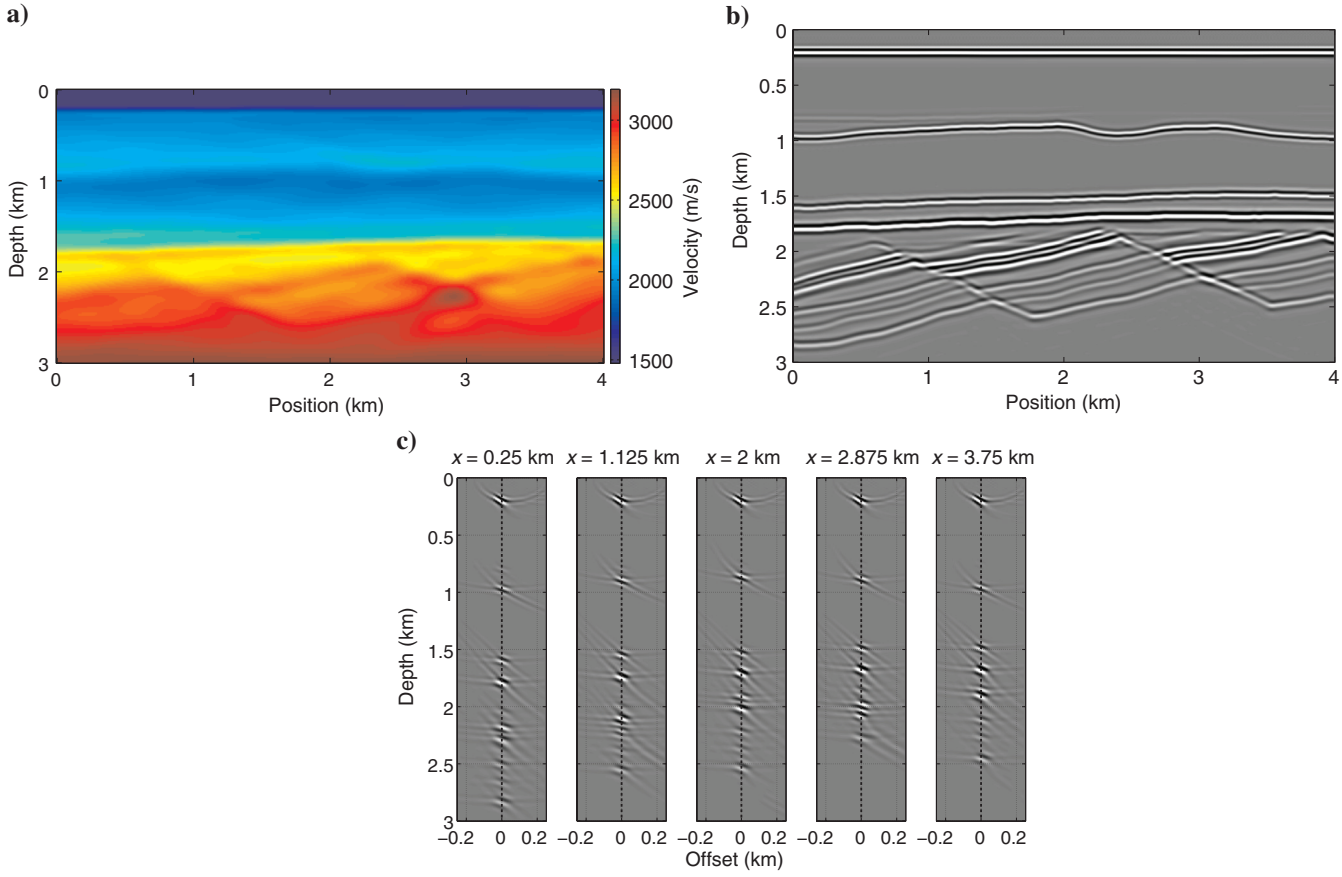

Figure 10. Updated velocity model (a), updated image (b), and updated subsurface CIGs (c) after 50 iterations of velocity analysis.

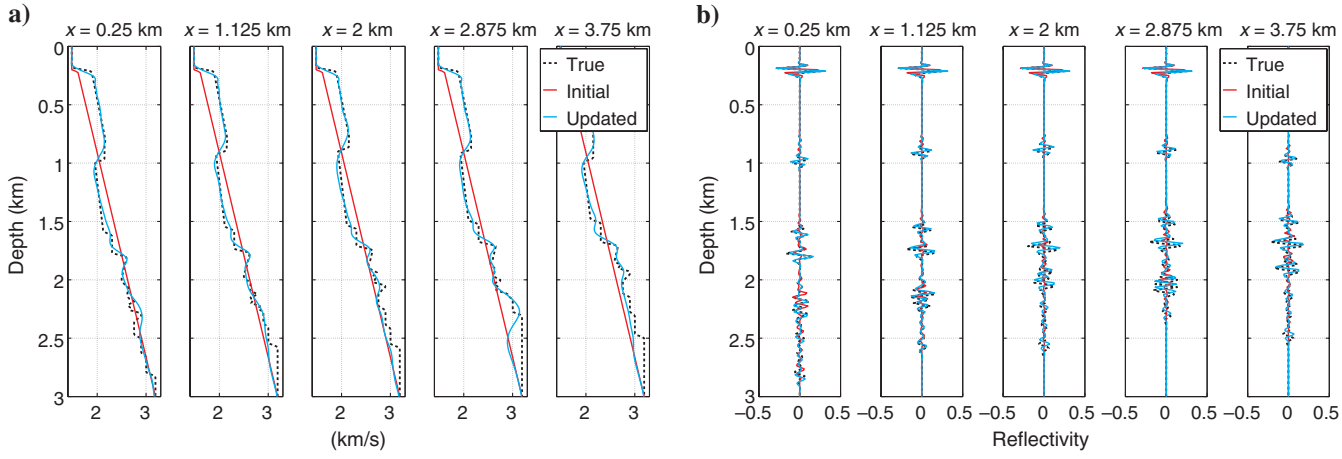

Figure 11. (a) Comparison of velocity traces. (b) Comparison of reflectivity traces.

To overcome this problem, we developed a method based on identifying and muting the multiples directly on the subsurface offset CIGs. The muted CIGs are used in a demigration procedure that kinematically reconstructs the reflection data. The resultant demultipled data can then be used as input for the velocity analysis.

To identify the free-surface multiples in the subsurface offset CIGs, we first note that these multiples focus at lower velocities than the primaries (Mulder and van Leeuwen, 2008). Thus, when migrating a data with free-surface multiples using an initial velocity model that is closer to the primaries, the multiples will, in general, appear as "smiles." However, in the particular case of single-spread acquisition, events migrated with nonoptimal velocities are also asymmetrically shifted relative to the zero subsurface offset. In the case of a source to the left of the streamer, the events that require lower velocity to focus will be shifted to the left, while those requiring higher velocity will be shifted to the right. Whereas Mulder and van Leeuwen (2008) introduce a bias toward higher velocities in the penalization of CIGs, we propose to mute the multiples and reconstruct the data as a preprocessing step. This is similar to the multiple attenuation procedure proposed by Sava and Guitton (2005), where multiples are attenuated by a dip filter in the subsurface angle gathers. The main differences here are that the multiple attenuation is carried out by a mute in the subsurface offset domain, and that, after the CIGs are muted, the data are reconstructed by "demigrating" the muted subsurface offset CIGs. The data can be reconstructed from the muted CIGs according to

$$U_s(\mathbf{x_{rec}}, t) = \int \mathbf{dx}' G(\mathbf{x_{rec}}, t; \mathbf{x}', 0) * \int d\mathbf{h} \frac{\partial^2 \mathcal{R}_m}{\partial x_3'^2}(\mathbf{x}' - \mathbf{h}, \mathbf{h}) D_s(\mathbf{x}' - 2\mathbf{h}, t), \qquad (15)$$

where $U_s(\mathbf{x_{rec}}, t)$ is the demigrated demultipled data at the receiver positions $\mathbf{x_{rec}}$, and $\mathcal{R}_m$ are the muted subsurface offset CIGs.

The data reconstruction avoids possible instabilities associated with the energy in the CIGs being moved in and out of the mute during the optimization procedure, which can occur in the case of a "static" weighting applied directly in the objective function (Mulder and ten Kroode, 2002; Mulder and van Leeuwen, 2008).

Note that the demultiple procedure is done only once, as a preprocessing step, and requires no modification of the objective function. The cost of the demultiple is equivalent to the cost of one and a half RTMs, because the source wavefields, D_s, can be stored during the migration step (equation 1), and accessed later during the data reconstruction (equation 15).

The demultiple procedure requires picking a mute that separates multiple from primary events in the CIGs. One way to pick this mute is to perform an initial velocity analysis, where the goal is to converge to a velocity in between primaries and multiples. The primaries and multiples could then be separated in the resultant CIGs by a vertical mute at zero subsurface offset (Li and Symes, 2007).

Here, we follow a more subjective approach, and pick a mute based on our own interpretation. As an example of the application of the demultiple procedure, Figure 13a shows a set of subsurface offset CIGs generated using the initial velocity shown in Figure 12a. The CIGs are separated by red vertical lines, while the black vertical lines show the position of the zero subsurface offset, and the stipple-dotted red lines show the position of picked mutes. The mutes are picked to the left of the strongest events in the CIGs, which are interpreted to be primaries. However, our experience with this data suggests that some strong events, in particular those between 2.5 and 3 km depth, are multiples and are, therefore, also included in the mute. Figure 13b shows the same CIGs after the mute is applied. Whereas Figure 14a and 14b shows the data before and after the demultiple procedure. Note how low-velocity events, indicated by steep moveouts, have been attenuated while the kinematics of primary events are preserved. This example shows that, although subjective, picking the mute based on interpretation allows great flexibility, and can be used to remove multiples and any other events that can be prejudicial to the velocity analysis.

To demonstrate the efficiency of the demultiple method in reducing the sensitivity of WEMVA to free-surface multiples, we compare the results of velocity analysis using the original field data with those obtained using the demultipled data. The initial model used in both cases is the one shown in Figure 12a. For better comparison of the results, all updated images are constructed using the original data.

The updated velocities after 19 iterations of optimization using the original data are shown in Figure 15a, while the updated image is shown in Figure 15b. The results show that, as expected, the velocities updated by WEMVA represent a compromise, as the method attempts to simultaneously focus primaries and multiples. As a consequence, the final migrated image is locally distorted, and some reflectors that were well-focused in the initial image (Figure 12b) are now clearly mispositioned.

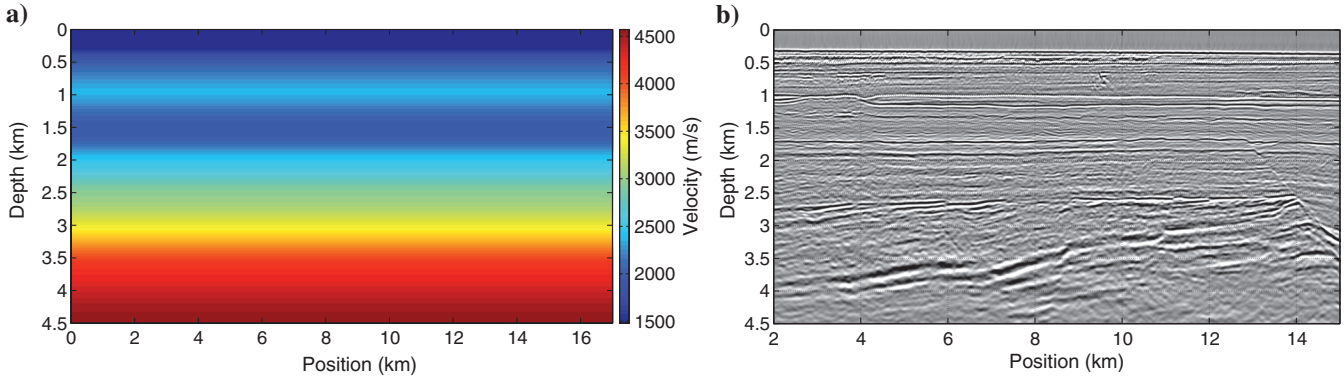

Figure 12. (a) Initial velocity model (m/s). (b) Image constructed using initial velocity model.

In comparison, the velocities obtained after 26 iterations of velocity analysis using the demultipled data are shown in Figure 16a. The updated image (Figure 16b) now shows a slight improvement over the initial image (Figure 12b), as we would expect in this case.

DISCUSSION

RTM-based WEMVA provides an automatic way of improving the quality of depth migrated images. However, artifacts in RTM which occur in the presence of strong and sharp velocity contrasts can cause the method to diverge. In the first synthetic example, we show how modifying the image with a vertical derivative operator improves the stability of the velocity analysis. In this example, the modification was necessary to ensure convergence and an adequate result. In general, we predict that this modification of the objective function is most significant when the velocity field has large velocity contrasts and/or is locally characterized by strong refracting mediums. In other cases, where one can accurately describe the

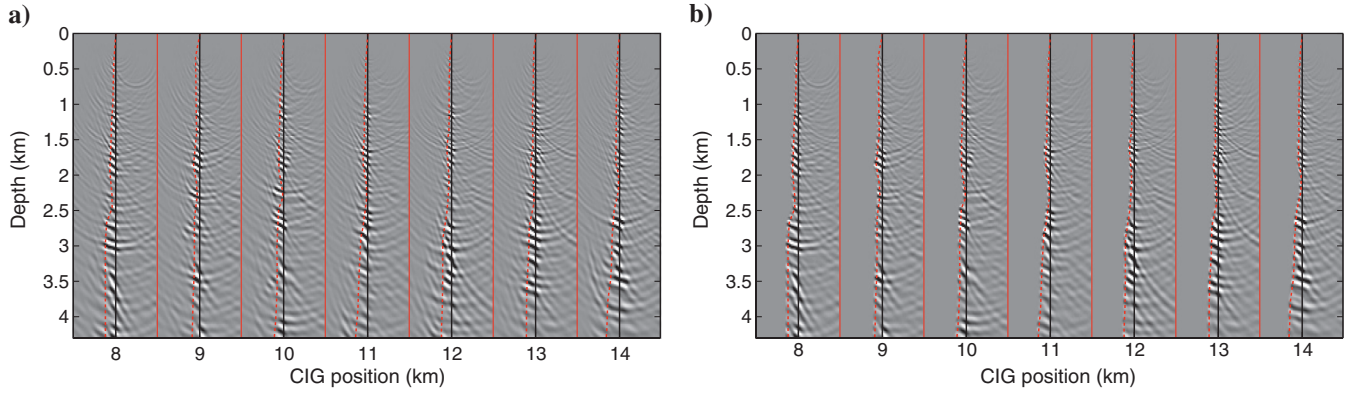

Figure 13. (a) Subsurface offset CIGs constructed using the initial velocity model in Figure 12a. The CIGs are separated by red vertical lines, while the stipple-dotted red lines mark the mute picks, and the black vertical lines mark the zero subsurface offset. (b) Subsurface offset CIGs after mute.

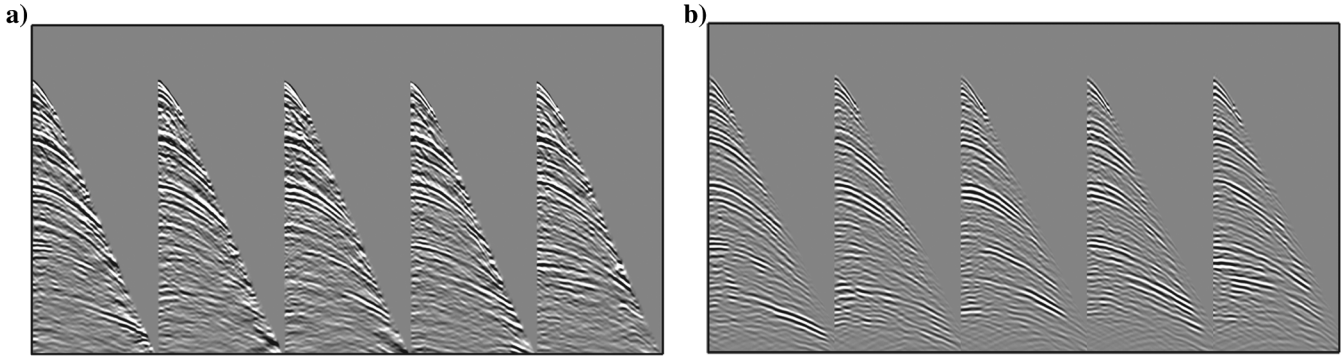

Figure 14. (a) Five shots taken from the original data. (b) Five shots taken from the data obtained from demigration of muted subsurface offset CIGs.

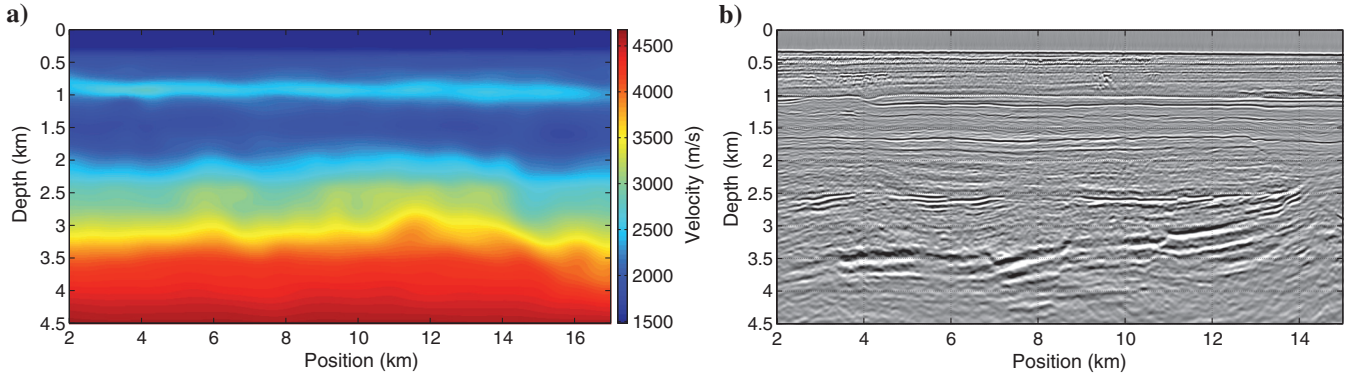

Figure 15. (a) Updated velocities from original data (m/s). (b) Image constructed using updated velocities in (a).

kinematics of the acoustic medium with a smoother velocity model (i.e., with only small acoustic impedance contrasts), such as in the second synthetic example, this modification may not significantly change the results.

The well posedness of the velocity analysis is strongly dependent on the velocity preconditioning, and the regularization. To guarantee a stable solution, the number of parameters to be estimated needs to be adjusted to the quantity of information present in the image. In our numerical examples, we controlled the number of optimized parameters by choosing different grid spacings for the B-spline coefficients. In the first example, where we have well-illuminated, evenly distributed reflectors, the B-spline grid spacing could be made quite small (40×40 m). Whereas in the second example, with a scarcer number of reflectors to constrain the upper 2 km of the velocity field, the grid was purposely chosen more sparse (150×50 m). In the field data example, where there are large differences in the continuity and illumination of the reflectors a very sparse B-spline grid was adopted (600×100 m). At the same time, the derivative regularization was used to prevent large artificial spatial velocity variations. These can appear during optimization due to the fact that the gradient of the differential semblance misfit and the similarity-index are scaled by the reflection coefficients of the image. However, it tends to slow down optimization. Therefore, the regularization parameter should be chosen as small as possible. The choice of 1% of the initial differential semblance misfit value seemed to be satisfactory for all the examples.

Free-surface multiples have long been a problem for automatic velocity analysis methods (Mulder and ten Kroode, 2002; Li and Symes, 2007; Mulder and van Leeuwen, 2008; van Leeuwen and Mulder, 2008). Even if multiple attenuation is used as a part of preprocessing, some multiple energy might still remain and bias the automatic velocity analysis (Li and Symes, 2007). Different methods have been proposed to reduce the sensitivity of WEMVA to the multiples. These methods are either based on including a filter in the objective function that bias the optimization toward higher velocities (Mulder and ten Kroode, 2002; Mulder and van Leeuwen, 2008), or iteratively modeling and subtracting the multiples as part of the velocity analysis (van Leeuwen and Mulder, 2008). We developed and tested a new preprocessing method where we mute the free-surface multiples in the subsurface offset CIGs. Instead of modifying the objective function, we use the muted CIGs to construct demultipled data, that is then used in the velocity analysis. The distinction between multiples and primaries is done based on their focusing velocities, and might require some level of interpretation. On the other hand, the procedure is very flexible and can also be used to remove other events from the data, such as steep dips and "fast" interbed multiples, which are known to cause problems for WEMVA (Biondi and Shan, 2002).

In all examples, the initial models used as starting point for optimization, although 1D, were approximating well the true background trends of the actual velocities. These models could be found, in a first instance, by preconditioning the velocity model to the space of 1D velocity models or to the space of linear 1D models. This approach works here because, in the examples shown, the geology consisted approximately of plane layered overburdens. In more complicated geological environments, with strong and sharp contrast velocity variations, more refined initial models are likely to be required to avoid converging to nonoptimal minima.

The high computational cost, in terms of computation and storage, is currently limiting the method to 2D and low-frequency data sets. There are, however, several measures that can be used to reduce these costs. First, in terms of computational cost, we suggest some strategies to improve the runtime of the method on large data sets.

The cost of RTM can be reduced significantly if shots can be combined and migrated together. If only two shots are combined the cost of migration is already halved. The speedup does not come for free, though, because this approach introduces crosstalk in the resulting image. The crosstalk can be attenuated by using some sort of source encoding (Romero et al., 2000). This approach has been widely experimented in RTM and FWI (Ben-Hadj-Ali et al., 2011). It is, however, not clear how the crosstalk artifacts could affect the results of WEMVA. Alternatively, the number of shots can be reduced through subsampling (Diaz and Guitton, 2011). In this case, only a subsample of the shots is taken and used for updating the velocity model at each iteration. By updating the subsampling along the iterative procedure, all shots are eventually used. A drawback of this procedure for WEMVA is that it may introduce aliasing in the subsurface offset CIGs, which may deteriorate the results.

In RTM, a typical problem has been the need to model separately the incident wavefield and the scattered wavefield, which means that one of the two wavefields must be stored and accessed later for the imaging step (checkpointing). In case of WEMVA, both fields need to be stored and accessed during gradient computation. Therefore, the cost of storage is double of that of RTM. In this case,

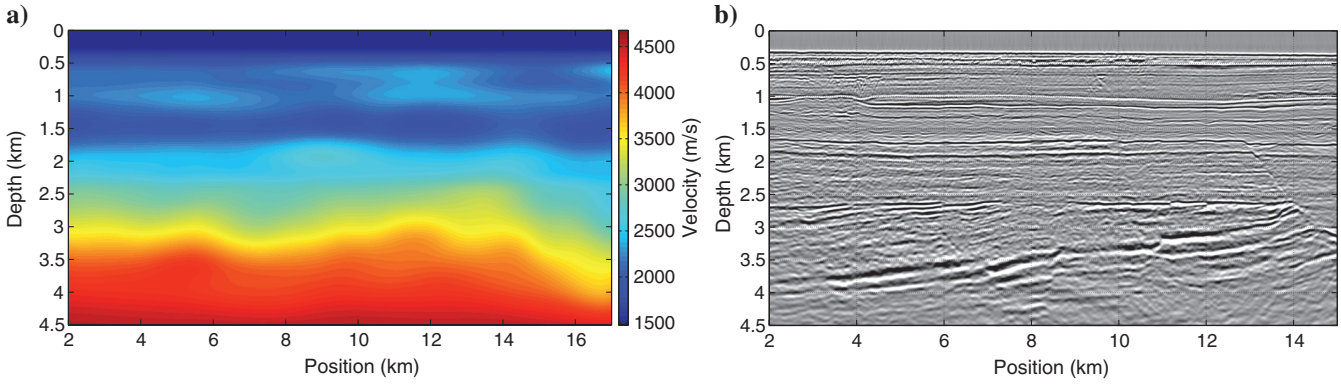

Figure 16. (a) Updated velocities from image-based demultipled data (m/s). (b) Image constructed using updated velocities in (a).

the cost of storing these fields can be reduced by optimal checkpointing, as suggested in Symes (2007).

CONCLUSION

We implemented a WEMVA method that can be used to estimate migration velocity fields from prestack seismic reflection data. The method estimates the velocities by minimizing an objective function based on differential semblance and similarity index of subsurface offset CIGs constructed by RTM. Artifacts of RTM (low-frequency noise) can be very sensitive to changes in the velocities and cause WEMVA to diverge. We showed that by modifying the image with a simple spatial differentiation operator helps to stabilize the velocity analysis in the presence of strong velocity contrasts. We showed that multiples can be attenuated in a procedure based on RTM, muting, and demigration, with optimal application for WEMVA.

ACKNOWLEDGMENTS

We are grateful to T. Nemeth, M. Sacchi, J. Schleicher, and five anonymous reviewers who greatly helped to improve the quality of this manuscript. The authors thank Statoil Petroleum AS and the sponsors of the ROSE consortium for financial support of this work. We also acknowledge the partners in the Snorre license, Statoil Petroleum AS, Petoro AS, ExxonMobil Exploration and Production Norway AS, Idemitsu Petroleum Norge AS, RWE Dea Norge AS, and Core Energy AS for permission to publish the results.

APPENDIX A

GRADIENT COMPUTATION

We now present the main steps required to derive the gradient of equations 5 and 6 by the adjoint state method (Lions and Magenes, 1972; Chavent and Lemonnier, 1974; Chavent, 2009).

The objective function is given by

$$\mathcal{J}(\mathbf{U},\mathbf{D},\mathbf{v}) = \frac{1}{2}\int \mathbf{dh}\int \mathbf{dx}\hat{\mathbf{h}}\left[\frac{\partial \mathcal{R}}{\partial x_3}(\mathbf{x},\mathbf{h})\right]^2, \tag{A-1}$$

where $\mathcal{R}(\mathbf{x},\mathbf{h}) = \sum_s \int dt U_s(\mathbf{x}+\mathbf{h},t) D_s(\mathbf{x}-\mathbf{h},t)$, and $\hat{\mathbf{h}} = \mathbf{h}^2 - \gamma\delta(\mathbf{h})$.

A Lagrangian function associated with the problem of minimizing equation A-1 with respect to $\mathbf{v}$ can be written as

$$\mathcal{L}(\mathbf{U},\mathbf{D},\mathbf{U}',\mathbf{D}',\mathbf{v}) = \mathcal{J}(\mathbf{U},\mathbf{D},\mathbf{v}) + \sum_s \left\langle \mathbf{U}'_\mathbf{s}, H^T\mathbf{U}_\mathbf{s} - \mathbf{P}_\mathbf{s}\right\rangle_{\mathbf{x},t} + \sum_s \left\langle \mathbf{D}'_\mathbf{s}, H\mathbf{D}_\mathbf{s} - \mathbf{S}_\mathbf{s}\right\rangle_{\mathbf{x},t}, \tag{A-2}$$

where $U'_s(\mathbf{x},t)$ and $D'_s(\mathbf{x},t)$ are Lagrange multipliers (adjoint states). The operator $H(\mathbf{v}) = (\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2)$ is the acoustic wave-equation forward-time marching operator. The transpose, $H^T(\mathbf{v})$, leads to a backward time marching scheme.

Here $\mathcal{L}$ is related to $\mathcal{J}$ by

$$\mathcal{J} = \mathcal{L}(\mathbf{U}_\mathbf{v},\mathbf{D}_\mathbf{v},\mathbf{U}',\mathbf{D}',\mathbf{v}), \tag{A-3}$$

where $\mathbf{U}_\mathbf{v}$, $\mathbf{D}_\mathbf{v}$ denote one realization of the direct states for a particular shot with a velocity vector $\mathbf{v}$.

Implicit differentiation of the above equation with respect to $\mathbf{v}$ gives

$$\delta\mathcal{J} = \frac{\partial\mathcal{L}}{\partial\mathbf{U}}(\mathbf{U}_\mathbf{v},\mathbf{D}_\mathbf{v},\mathbf{U}',\mathbf{D}',\mathbf{v})\cdot\delta\mathbf{U} + \frac{\partial\mathcal{L}}{\partial\mathbf{D}}(\mathbf{U}_\mathbf{v},\mathbf{D}_\mathbf{v},\mathbf{U}',\mathbf{D}',\mathbf{v}) \cdot\delta\mathbf{D} + \frac{\partial\mathcal{L}}{\partial\mathbf{v}}(\mathbf{U}_\mathbf{v},\mathbf{D}_\mathbf{v},\mathbf{U}',\mathbf{D}',\mathbf{v})\cdot\delta\mathbf{v}. \tag{A-4}$$

Now, if we assume $\mathbf{U}'$ and $\mathbf{D}'$ to satisfy

$$\frac{\partial\mathcal{L}}{\partial\mathbf{U}}(\mathbf{U}_\mathbf{v},\mathbf{D}_\mathbf{v},\mathbf{U}',\mathbf{D}',\mathbf{v})\cdot\delta\mathbf{U} = 0, \tag{A-5}$$

and

$$\frac{\partial\mathcal{L}}{\partial\mathbf{D}}(\mathbf{U}_\mathbf{v},\mathbf{D}_\mathbf{v},\mathbf{U}',\mathbf{D}',\mathbf{v})\cdot\delta\mathbf{D} = 0, \tag{A-6}$$

for all $\delta\mathbf{U}$ and $\delta\mathbf{D}$, then equation A-4 reduces to

$$\delta\mathcal{J} = \frac{\partial\mathcal{L}}{\partial\mathbf{v}}(\mathbf{U}_\mathbf{v},\mathbf{D}_\mathbf{v},\mathbf{U}',\mathbf{D}',\mathbf{v})\cdot\delta\mathbf{v}. \tag{A-7}$$

The problem now is to solve equations A-5 and A-6 for the adjoint states for each shot. Starting with equation A-6. Noting that $(\partial/\partial x_3)^T = -(\partial/\partial x_3)$, and that $\int \mathbf{dh}\mathcal{R}(\mathbf{x})U(\mathbf{x}+\mathbf{h})\delta D(\mathbf{x}-\mathbf{h}) = \int \mathbf{dh}\mathcal{R}(\mathbf{x}+\mathbf{h})U(\mathbf{x}+2\mathbf{h})\delta D(\mathbf{x})$, the stationary points of the Lagrangian $\partial\mathcal{L}/\partial\mathbf{D} = 0$ lead to the reverse-time problem

$$H^T(\mathbf{v})D'_s(\mathbf{x},t) = \int \mathbf{dh}\hat{\mathbf{h}}\frac{\partial^2\mathcal{R}}{\partial x_3^2}(\mathbf{x}+\mathbf{h},\mathbf{h})U_s(\mathbf{x}+2\mathbf{h},t), \tag{A-8}$$

which can be solved through a backward time marching scheme, starting from a final condition of rest, i.e., $D'_s(\mathbf{x},T) = 0$.

While equation $\partial\mathcal{L}/\partial\mathbf{U} = 0$ leads to the forward-time problem

$$H(\mathbf{v})U'_s(\mathbf{x},t) = \int \mathbf{dh}\hat{\mathbf{h}}\frac{\partial^2\mathcal{R}}{\partial x_3^2}(\mathbf{x}-\mathbf{h},\mathbf{h})D_s(\mathbf{x}-2\mathbf{h},t), \tag{A-9}$$

which can be solved through a forward time marching scheme, starting from an initial condition of rest, i.e., $U'_s(\mathbf{x},0) = 0$.

The solutions to equations A-8 and A-9 can be expressed in the form of Green's functions as

$$D'_s(\mathbf{x},t) = \int \mathbf{dx}'G(\mathbf{x},0;\mathbf{x}',t) * \int \mathbf{dh}\hat{\mathbf{h}}\frac{\partial^2\mathcal{R}}{\partial x'^2_3}(\mathbf{x}',\mathbf{h})U_s(\mathbf{x}'+2\mathbf{h},t), \tag{A-10}$$

and

$$U'_s(\mathbf{x},t) = \int \mathbf{dx}' G(\mathbf{x},t;\mathbf{x}',0) * \int \mathbf{dh}\hat{\mathbf{h}} \frac{\partial^2 \mathcal{R}}{\partial x'^2_3}(\mathbf{x}'-\mathbf{h},\mathbf{h}) D_s(\mathbf{x}'-2\mathbf{h},t), \tag{A-11}$$

where * means time convolution.

We finally turn to the problem of finding the derivative of $\mathcal{J}$ with respect to velocity ($\mathbf{v}$). Differentiating equation A-2 with respect to $\mathbf{v}$, and noting that $\langle \mathbf{D}', (\partial H/\partial \mathbf{v})\mathbf{D}\rangle_t = \langle \mathbf{D}', (\partial H/\partial \mathbf{v})^T \mathbf{D}\rangle_t = \langle \mathbf{D}', -(2/\mathbf{v}^3)(\partial^2 \mathbf{D}/\partial t^2)\rangle_t$, yields

$$\delta\mathcal{J} = -\sum_s \left\langle \mathbf{D}'_\mathbf{s}, \frac{2\delta\mathbf{v}}{\mathbf{v}^3}\frac{\partial^2 \mathbf{D}_\mathbf{s}}{\partial t^2}\right\rangle_{x,t} - \sum_s \left\langle \mathbf{U}'_\mathbf{s}, \frac{2\delta\mathbf{v}}{\mathbf{v}^3}\frac{\partial^2 \mathbf{U}_\mathbf{s}}{\partial t^2}\right\rangle_{x,t}, \tag{A-12}$$

from which we obtain the gradient by picking the coefficients of $\delta\mathbf{v}$

$$\nabla_\mathbf{v}\mathcal{J}(\mathbf{x}) = -\sum_s \int dt \frac{2}{v^3(\mathbf{x})}\frac{\partial^2 D_s}{\partial t^2}(\mathbf{x},t) D'_s(\mathbf{x},t) - \sum_s \int dt \frac{2}{v^3(\mathbf{x})}\frac{\partial^2 U_s}{\partial t^2}(\mathbf{x},t) U'_s(\mathbf{x},t). \tag{A-13}$$

REFERENCES

Arntsen, B., C. Gerea, and T. Rosten, 2009, Imaging salt bodies using explicit migration operators offshore Norway: Geophysics, **74**, no. 2, S25–S32, doi: 10.1190/1.3063660.

Ben-Hadj-Ali, H., S. Operto, and J. Virieux, 2011, An efficient frequency-domain full waveform inversion method using simultaneous encoded sources: Geophysics, **76**, no. 4, R109–R124, doi: 10.1190/1.3581357.

Biondi, B., and P. Sava, 1999, Wave-equation migration velocity analysis: 69th Annual International Meeting, SEG, Expanded Abstracts, 1723–1726.

Biondi, B., and G. Shan, 2002, Prestack imaging of overturned reflections by reverse time migration: 72nd Annual International Meeting, SEG Expanded Abstracts, 1284–1287.

Biondi, B., and W. W. Symes, 2004, Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging: Geophysics, **69**, 1283–1298, doi: 10.1190/1.1801945.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu, 1995, A limited memory algorithm for bound constrained optimization: SIAM Journal on Scientific Computing, **16**, 1190–1208, doi: 10.1137/0916069.

Chavent, G., 2009, Non-linear least squares for inverse problems: Theoretical foundations and step-by-step guide for applications: Springer.

Chavent, G., and C. A. Jacewitz, 1995, Determination of background velocities by multiple migration fitting: Geophysics, **60**, 476–490, doi: 10.1190/1.1443785.

Chavent, G., and P. Lemonnier, 1974, Identification de la non-linéarité d'une e´quation parabolique quasiline´aire: Applied mathematics and Optimization, **1**, 121–162, doi: 10.1007/BF01449027.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481, doi: 10.1190/1.1440185.

Diaz, E., and A. Guitton, 2011, Fast full waveform inversion with random shot decimation: 81st Annual International Meeting, SEG, Expanded Abstracts, 2804–2808.

Dierckx, P., 1993, Curve and surface fitting with splines: Oxford University Press.

Fei, W., and P. Williamson, 2010, On the gradient artifacts in migration velocity analysis based on differential semblance optimization: 80th Annual International Meeting, SEG, Expanded Abstracts, 4071–4076.

Gao, F., and W. Symes, 2009, Differential semblance velocity analysis by reverse time migration: Image gathers and theory: 79th Annual International Meeting, SEG, Expanded Abstracts, 2317–2321.

Guitton, A., B. Kaelin, and B. Biondi, 2007, Least-squares attenuation of reverse-time-migration artifacts: Geophysics, **72**, no. 1, S19–S23, doi: 10.1190/1.2399367.

Li, J., and W. W. Symes, 2007, Interval velocity estimation via NMO-based differential semblance: Geophysics, **72**, no. 6, U75–U88, doi: 10.1190/1.2767401.

Lions, J. L., and E. Magenes, 1972, Nonhomogeneous boundary value problems and applications: Springer Verlag.

Mulder, W. A., 2008, Automatic velocity analysis with the two-way wave equation: Presented at the 70th Annual International Conference and Exhibition, EAGE.

Mulder, W. A., and A. P. E. ten Kroode, 2002, Automatic velocity analysis by differential semblance optimization: Geophysics, **67**, 1184–1191, doi: 10.1190/1.1500380.

Mulder, W. A., and T. van Leeuwen, 2008, Automatic migration velocity analysis and multiples: 78th Annual International Meeting, SEG Expanded Abstracts, 3128–3132.

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

Qin, Z., M. Lu, X. Zheng, Y. Yao, C. Zhang, and J. Song, 2009, The implementation of an improved NPML absorbing boundary condition in elastic wave modeling: Applied Geophysics, **6**, 113–121, doi: 10.1007/s11770-009-0012-3.

Rickett, J. E., and P. C. Sava, 2002, Offset and angle-domain common image-point gathers for shot-profile migration: Geophysics, **67**, 883–889, doi: 10.1190/1.1484531.

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, doi: 10.1190/1.1444737.

Sava, P., and B. Biondi, 2004, Wave-equation migration velocity analysis. I. Theory: Geophysical Prospecting, **52**, 593–606, doi: 10.1111/j.1365-2478.2004.00447.x.

Sava, P., and S. Fomel, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, no. 6, S209–S217, doi: 10.1190/1.2338824.

Sava, P., and L. Vlad, 2008, Numeric implementation of wave-equation migration velocity analysis operators: Geophysics, **73**, no. 5, VE145–VE159, doi: 10.1190/1.2953337.

Sava, P. C., and A. Guitton, 2005, Multiple attenuation in the image space: Geophysics, **70**, no. 1, V10–V20, doi: 10.1190/1.1852789.

Shen, P., and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, no. 5, VE49–VE59, doi: 10.1190/1.2972021.

Shen, P., W. W. Symes, and C. C. Stolk, 2003, Differential semblance velocity analysis by wave equation migration: 73rd Annual International Meeting, SEG, Expanded Abstracts, 2132–2135.

Symes, W. W., 2007, Reverse time migration with optimal checkpointing: Geophysics, **72**, no. 5, SM213–SM221, doi: 10.1190/1.2742686.

Symes, W. W., and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663, doi: 10.1190/1.1443082.

Tikhonov, A. N., and V. Y. Arsenin, 1977, Solution of ill-posed problems: W. H. Winston and Sons.

van Leeuwen, T., and W. A. Mulder, 2008, Velocity analysis with multiples – NMO modeling for layered velocity structures: 78th Annual International Meeting, SEG, Expanded Abstracts, 1925–1929.

Virieux, J., 1986, P-SV wave propagation in heterogeneous media; velocity-stress finite-difference method: Geophysics, **51**, 889–901, doi: 10.1190/1.1442147.