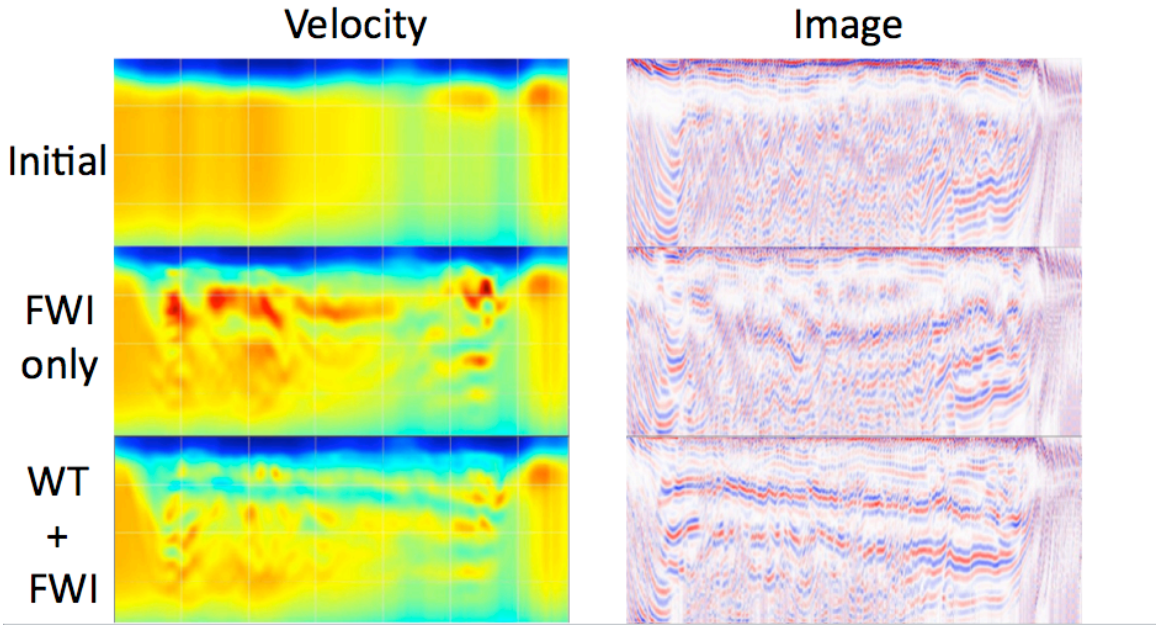
STANFORD EXPLORATION PROJECT

Ali Almomin, Gboyega Ayeni, Ran Bachrach, Ohad Barak, Biondo Biondi, Jon Claerbout, Robert Clapp, Joe Dellinger, Qiang Fu, Adam Halpert, Tim Keho, Chris Leader, Robert Ley, Yunyue Li, Yi Luo, Dave Nichols, Konstantin Osypov, Sjoerd de Ridder, Shuki Ronen, Xukai Shen, Yi Shen, Yaxun Tang, Kittinat Taweesintananon, Thierry Tonellot, Mandy Wong, and Yang Zhang

Report Number 143, June 2011



Preface

The electronic version of this report[1] makes the included programs and applications available to the reader. The markings [ER], [CR], and [NR] are promises by the author about the reproducibility of each figure result. Reproducibility is a way of organizing computational research that allows both the author and the reader of a publication to verify the reported results. Reproducibility facilitates the transfer of knowledge within SEP and between SEP and its sponsors.

ER denotes Easily Reproducible and are the results of processing described in the paper. The author claims that you can reproduce such a figure from the programs, parameters, and makefiles included in the electronic document. The data must either be included in the electronic distribution, be easily available to all researchers (e.g., SEG-EAGE data sets), or be available in the SEP data library[2]. We assume you have a UNIX workstation with Fortran, Fortran90, C, X-Windows system and the software downloadable from our website (SEP makerules, SEPlib, and the SEP latex package), or other free software such as SU. Before the publication of the electronic document, someone other than the author tests the author's claim by destroying and rebuilding all ER figures. Some ER figures may not be reproducible by outsiders because they depend on data sets that are too large to distribute, or data that we do not have permission to redistribute but are in the SEP data library.

CR denotes Conditional Reproducibility. The author certifies that the commands are in place to reproduce the figure if certain resources are available. The primary reasons for the CR designation is that the processing requires 20 minutes or more, or commercial packages such as Matlab or Mathematica.

NR denotes Non-Reproducible figures. SEP discourages authors from flagging their figures as NR except for figures that are used solely for motivation, comparison, or illustration of the theory, such as: artist drawings, scannings, or figures taken from SEP reports not by the authors or from non-SEP publications.

Our testing is currently limited to LINUX 2.6 (using the Intel Fortran90 compiler), but the code should be portable to other architectures. Reader's suggestions are welcome. More information on reproducing SEP's electronic documents is available online[3].

[1] http://sepwww.stanford.edu/private/docs/sep143
[2] http://sepwww.stanford.edu/public/docs/sepdatalib/toc_html
[3] http://sepwww.stanford.edu/research/redoc/

SEP143 — TABLE OF CONTENTS

Wave-equation Migration Velocity Analysis

Full Waveform Inversion & velocity

Linearized wave-equation inversion

Reverse time migration & Angle gathers

Velocity estimation and model building

Bidirectional deconvolution

Subsalt velocity analysis by target-oriented wavefield tomography: A 3-D field-data example

Yaxun Tang and Biondo Biondi

ABSTRACT

We apply target-oriented wavefield tomography to a 3-D field data set acquired from the Gulf of Mexico. Instead of using the original surface-recorded data set, we use a new data set synthesized specifically for velocity analysis to update subsalt velocities. The new data set is generated based on an initial unfocused target image and by a novel application of 3-D generalized Born wavefield modeling, which correctly preserves velocity kinematics by modeling non-zero subsurface-offset-domain images. We show that the target-oriented inversion strategy drastically reduces the data size and the computation domain for 3-D wavefield tomography, greatly improving its efficiency and flexibility. We apply differential semblance optimization (DSO) using the synthesized new data set to optimize subsalt velocities. The updated velocity model significantly improves the continuity of subsalt reflectors and yields flattened angle-domain common-image gathers.

INTRODUCTION

Accurate reflectivity imaging requires an accurate background velocity model. As seismic exploration moves towards structurally complex areas, wavefield-based tomography that better models band-limited wave phenomena becomes necessary for high-quality velocity model building. Wavefield-based tomography, however, is still expensive for industrial-scale applications (Biondi and Sava, 1999; Shen et al., 2005; Albertin et al., 2006; Fei et al., 2009), both because the method uses more expensive wavefield modeling engines, and because it lacks flexibility and usually requires the use of the whole recorded data set for velocity analysis.

To reduce the cost and increase the flexibility of wavefield-based tomography, Biondi (2006); Guerra (2010); Tang and Biondi (2010) originated the idea of target-oriented wavefield tomography. The idea is to synthesize a target-oriented new data set specifically for velocity analysis. The new data set is designed to be much smaller than the original surface-recorded data set, while still containing all necessary information for velocity updating. This strategy allows us to apply the powerful but expensive wavefield-based technique only in areas where it is necessary, such as in subsalt regions with complex overburdens, and leave areas with relatively simple geologies to be handled by conventional velocity-analysis methods, which are sufficient to produce accurate results.

By localizing the computation within a selected target zone, the target-oriented inversion strategy dramatically improves the efficiency and flexibility of wavefield-based tomography. Therefore, it can greatly shorten the cycle time from seismic processing to interpretation, enabling interpretation-driven interactive wavefield-based velocity analysis, where different

geological scenarios can be tested in almost real time (Halpert et al., 2008). The high efficiency may also make velocity uncertainty analysis feasible, which requires inverting velocity models multiple times to build the probability distribution (Tarantola, 2005).

Biondi (2006); Guerra et al. (2009); Guerra (2010) synthesize a new data set for local tomography using the concept of prestack exploding-reflector modeling (PERM). PERM, however, generates crosstalk when multiple image events (reflectors) are modeled simultaneously. This limits the number of reflectors it can model. Manual picking and stochastic encoding methods, such as random-phase encoding, are required to mitigate the impact of the crosstalk.

In contrast, Tang and Biondi (2010) formulate the problem under a seismic-data-mapping (SDM) framework (Hubral et al., 1996; Bleistein and Jaramillo, 2000) and use generalized Born wavefield modeling as the mapping operator to synthesize a new data set for velocity analysis. Generalized Born wavefield modeling is extended from conventional Born modeling (Stolt and Benson, 1986) to include modeling of the non-zero subsurface offset images. As shown by Tang and Biondi (2010), the inclusion of the subsurface offset in the modeling process preserves the correct velocity information and is crucial to the success of this method. One advantage of generalized Born wavefield modeling is that it does not require any picking, but picking can be incorporated if it is desired. Another advantage is that it can model arbitary number of reflectors simultaneously and is minimally affected by crosstalk artifacts.

In this paper, we follow the method of Tang and Biondi (2010) and use generalized Born wavefield modeling to generate a target-oriented data set for local velocity analysis. In the subsequent sections, we first review the theory of target-oriented tomography using synthesized Born data. We then apply the methodology to a 3-D field data set acquired from the Gulf of Mexico (GOM), where we update the subsalt velocities in a target-oriented fashion.

THEORY

We formulate our method under the framework of seismic data mapping (SDM) (Hubral et al., 1996; Bleistein and Jaramillo, 2000), where the idea is to transform the original observed seismic data from one acquisition configuration to another with a designed mapping operator. SDM can be summarized as two main steps: (1) apply the (pseudo) inverse of the designed mapping operator to the original data set to generate a model, and (2) apply the forward mapping operator to the model to generate a new data set with a different acquisition configuration than the original one.

In our case, we use generalized Born wavefield modeling to perform data mapping. With an initial velocity model, seismic prestack images can be obtained using the pseudo-inverse of the generalized Born modeling operator as follows:

$$\mathbf{m} = \mathbf{H}_0^{\dagger}\mathbf{L}_0^{*}\mathbf{d}_{\text{obs}}, \tag{1}$$

where * and † denote adjoint and pseudo-inverse, respectively; $\mathbf{m}$ is the seismic image; $\mathbf{L}_0$ is the generalized Born modeling operator computed using an initial velocity model $\mathbf{v}_0$, whose adjoint $\mathbf{L}_0^*$ is the well-known depth migration operator; $\mathbf{H}_0$ is the Hessian operator (Plessix and Mulder, 2004; Valenciano, 2008; Tang, 2009); and $\mathbf{d}_{\text{obs}}$ is the observed surface data.

It is important to note that the seismic image $\mathbf{m}$ must be parameterized as a function of both spatial location and some prestack parameter, such as the subsurface offset, reflection angle, etc., in order to preserve the velocity information for later velocity analysis (Tang and Biondi, 2010). In this paper, we use the subsurface offset as our prestack parameter. The significance of the Hessian operator in equation 1 is that its pseudo-inverse removes the influence of the original acquisition geometry in the least-squares sense, and the resulting image is independent from the original data. However, the full Hessian $\mathbf{H}_0$ is impossible to obtain in practice due to its size and computational cost; we therefore approximate it by a diagonal matrix : $\mathbf{H}_0 \approx \text{diag}\{\mathbf{H}_0\}$. We further reduce the cost of computing the diagonal of Hessian by using the phase encoding method (Tang, 2007, 2009; Tang and Lee, 2010).

We obtain a target image $\mathbf{m}_{\text{target}}$ by applying a selecting operator $\mathbf{S}$ to the initial image: $\mathbf{m}_{\text{target}} = \mathbf{S}\mathbf{m}$, where the selecting operator $\mathbf{S}$ can be simply a windowing operator. A new data set $\widetilde{\mathbf{d}}_{\text{obs}}$ can then be simulated as follows:

$$\widetilde{\mathbf{d}}_{\text{obs}} = \widetilde{\mathbf{L}}_0 \mathbf{m}_{\text{target}}, \tag{2}$$

where $\widetilde{\mathbf{L}}_0$ is the Born modeling operator computed using the same initial velocity $\mathbf{v}_0$, but with a different acquisition configuration. The wavefield propagation can be restricted to regions with inaccurate velocities, and the modeled data can be collected at the top of the target region. The target-oriented modeling strategy makes the new data set much smaller than the original one. The new data set can be imaged using the migration operator, i.e., the adjoint of $\widetilde{\mathbf{L}}$, with an arbitary velocity $\mathbf{v}$, as follows:

$$\widehat{\mathbf{m}}_{\text{target}} = \widetilde{\mathbf{L}}^* \widetilde{\mathbf{d}}_{\text{obs}}. \tag{3}$$

We pose the velocity estimation problem as an optimization problem that seeks an optimum velocity model by minimizing a user-defined image residual (or maximizing some measure of the image coherence). There are many ways of defining the objective functions. In this paper, we use the differential semblance optimization (DSO) (Symes and Carazzone, 1991) as the criterion to estimate the velocity. The DSO objective function in the subsurface-offset domain is (Shen, 2004; Shen and Symes, 2008)

$$F(\mathbf{v}) = \frac{1}{2} \sum_{\mathbf{x}} \sum_{\mathbf{h}} |\mathbf{h}|^2 \widehat{m}^2_{\text{target}}(\mathbf{x}, \mathbf{h}), \tag{4}$$

where $\mathbf{x} = (x, y, z)$ is the image point in the subsurface and $\mathbf{h} = (h_x, h_y, 0)$ is the subsurface offset. The physical interpretation of the subsurface-offset-domain DSO is that it optimizes the velocity model by penalizing energy at non-zero subsurface offset, taking advantage of the fact that seismic events should focus at zero-subsurface offset if migrated using an accurate velocity model (Shen, 2004). However, the gradient of the objective function defined by equation 4 is sensitive to the amplitude variation of the images due to uneven illumination (Vyas and Tang, 2010; Fei and Williamson, 2010). We propose to normalize the DSO objective function by the square of the root-mean-squared (RMS) image amplitudes to reduce the influence of image amplitude variations. The normalized DSO objective function is

$$J(\mathbf{v}) = \frac{1}{2} \sum_{x,y} \frac{\sum_{z,\mathbf{h}} |\mathbf{h}|^2 |\widehat{m}_{\text{target}}(\mathbf{x}, \mathbf{h})|^2}{\sum_{z,\mathbf{h}} |\widehat{m}_{\text{target}}(\mathbf{x}, \mathbf{h})|^2}. \tag{5}$$

We use nonlinear conjugate-gradient method to minimize J. The gradient is calculated using the adjoint-state method (with a one-way wave-equation formulation) without explicitly computing the Jacobian matrix (Shen and Symes, 2008; Sava and Vlad, 2008; Tang et al., 2008).

3-D FIELD-DATA EXAMPLES

We apply the target-oriented wavefield tomography to a data set acquired from the Gulf of Mexico (GOM). The data set was acquired using a narrow-azimuth towed streamer (NATS) acquisition system, and further rotated using azimuth moveout (AMO) (Biondi et al., 1998) into zero azimuth. The minimum and maximum inline offsets are 0.3 km and 8.2 km, respectively. The frequency content of the data set ranges between 5 Hz and 35 Hz. The initial 3-D velocity model is shown in Figure 1. Velocities above the target (outlined by a black box) and the salt interpretation are known to be accurate. The goal is to invert for subsalt velocities inside the target region. The initial velocities inside the box are set to be $v(z)$.

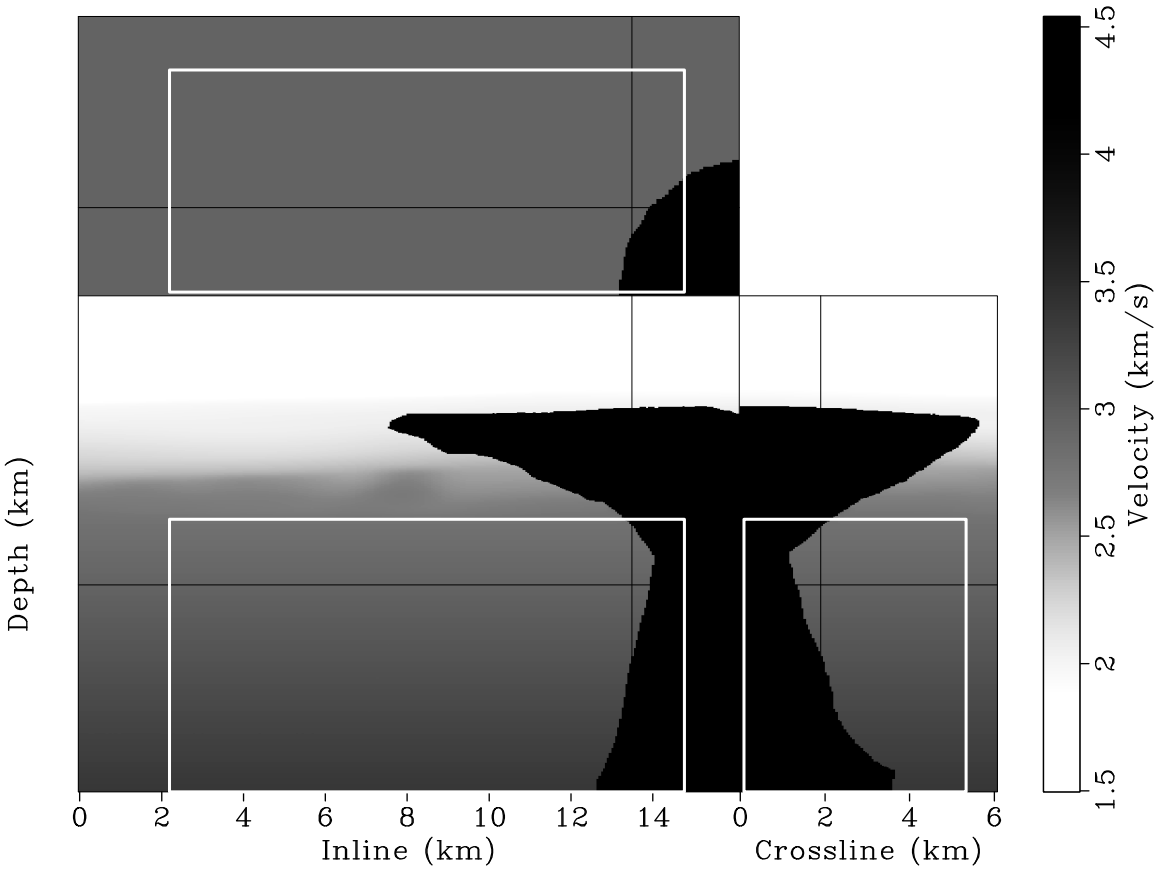


Figure 1: The initial 3-D velocity model. The black box outlined area is the target region for velocity analysis. Velocities above the target region and the salt interpretation are known to be accurate. [**ER**] yaxun1/. bpgom3d-bvel-full

Generating initial image gathers

We perform conical-wave migration (Whitmore, 1995; Duquet et al., 2001; Zhang et al., 2005; Liu et al., 2006) to generate the initial image gathers, where we synthesize 101 conical waves for each crossline and migrate 12625 conical waves in total. The minimum and maximum inline take-off angles at the surface for the conical waves are -30° and 30°, respectively. The maximum frequency used for the initial migration is 20 Hz. We also compute the diagonal of the Hessian in the conical-wave domain using the phase-encoding method (Tang, 2009), which encodes the source-side Green's functions using inline plane-wave phase-encoding functions and the receiver-side Green's functions using random-phase encoding functions. The simultaneous encoding dramatically reduces the computational cost of the Hessian (Tang and Biondi, 2011).

Figure 2 shows the zero-subsurface offset image for the target region obtained using the initial velocity model (Figure 1). The image has been normalized with the diagonal of the Hessian (not shown here). To more accurately preserve the velocity information, we compute both inline and crossline subsurface offsets (Figure 3). The crossline subsurface offset is included because wavefields can travel out of plane during propagation, and therefore they may image the subsurface with different azimuths than that on the surface. The 3-D subsurface-offset-domain common-image gathers (SODCIGs) shown in Figure 3 confirm this; the events in the top panels of Figures 3a and 3b are tilted, suggesting that they are not imaged by zero subsurface azimuth. Also note that the 3-D SODCIGs are not well focused at the zero subsurface offset in either the inline or crossline directions. The defocusing in the inline subsurface offset (h_x) is mainly caused by velocity inaccuracies, whereas the defocusing in the crossline subsurface offset (h_y) is mainly due to insufficient crossline data coverage (single surface azimuth data)(Tang, 2007).

3-D Born wavefield modeling

The next step is to model a new data set for velocity analysis. The initial image and gathers (Figures 2 and 3) have been preprocessed before the modeling step, where we simply mute events inside the salt body. More sophisticated preprocessing, such as noise attenuation, key reflection selection, etc., can also be applied. The preprocessing step allows human interaction in the velocity analysis workflow, adding desired flexibility and control to this method.

We use 3-D plane-wave modeling to generate a Born data set at the top of the target region. To determine the optimum parameters for the 3-D modeling, we carry out seismic visibility analysis (Jin and Xu, 2010), which can provide us a quantitative estimate of which part of the data contributes most to the target region for a given prestack acquisition geometry. Figure 4 shows the average source visibility map for the dominant frequency 10 Hz, as a function of both inline and crossline ray parameters. Note that most of the energy contributing to the target region falls within the range where $\widetilde{p}_{s_x} \in (-213, 213)$ μs/m, and $\widetilde{p}_{s_y} \in (-180, 180)$ μs/m.

Based on the result of visibility analysis, we modeled 315 plane waves at the top of the target region, with 21 plane waves inline and 15 plane waves crossline. The minimum and maximum dip and azimuth angles of these plane waves are approximately -47.7° and

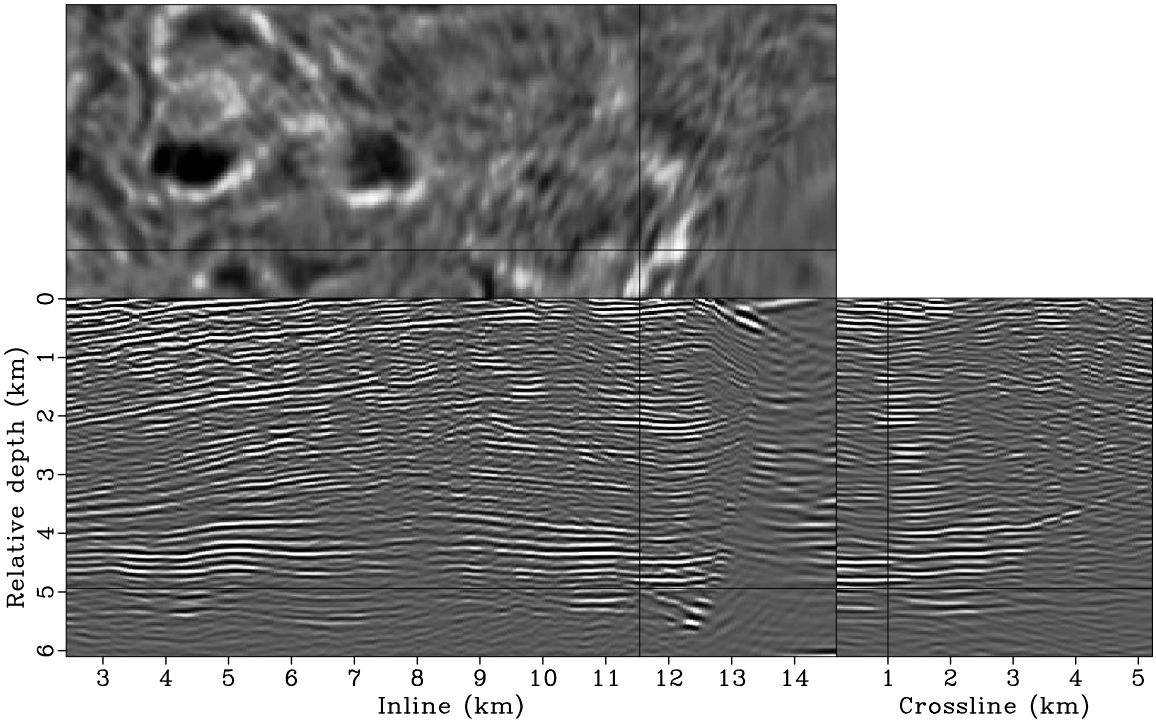


Figure 2: The zero-subsurface-offset-domain image migrated using the original data set and the initial velocity. The image has been normalized using the diagonal of the phase-encoded Hessian. [**CR**] yaxun1/. bpgom3d-bimg-cpst1-copy

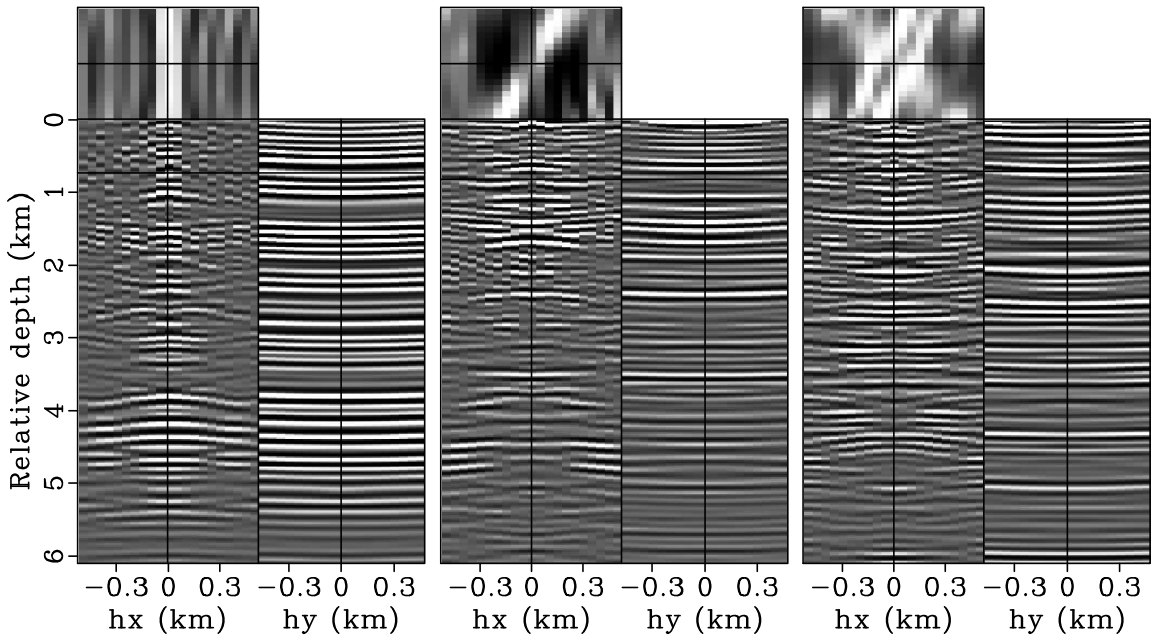


Figure 3: The Hessian-normalized 3-D SODCIGs obtained using the original data set and the initial velocity model. Panels (a), (b) and (c) are extracted at $(3.04, 1.34)$ km, $(6.09, 2.86)$ km and $(9.13, 4.38)$ km, respectively. [**CR**] yaxun1/. bpgom3d-cig3d-orig-cpst

47.7°, −40.2° and 40.2°, respectively. The dominant frequency used for the generalized Born wavefield modeling is 10 Hz. Figure 5 shows a 3-D plane-wave source function with both incident dip and azimuth angles equal to 10°, shooting from the top of the target region (Figure 1), and the corresponding 3-D plane-wave gather collected at the top of the target region.

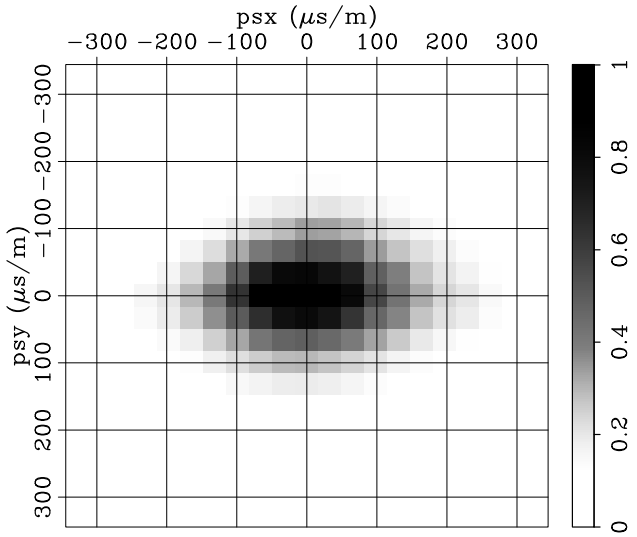


Figure 4: The average source visibility (data-domain illumination) for the target region using 3-D plane-wave acquisition geometry for the Born wavefield modeling. Note that most of the energy contributing to the target region falls within the range where $\widetilde{p}_{s_x} \in (-213, 213)$, and $\widetilde{p}_{s_y} \in (-180, 180)$. [**CR**]
yaxun1/. bpgom3d-born-visi

Tomographic inversion

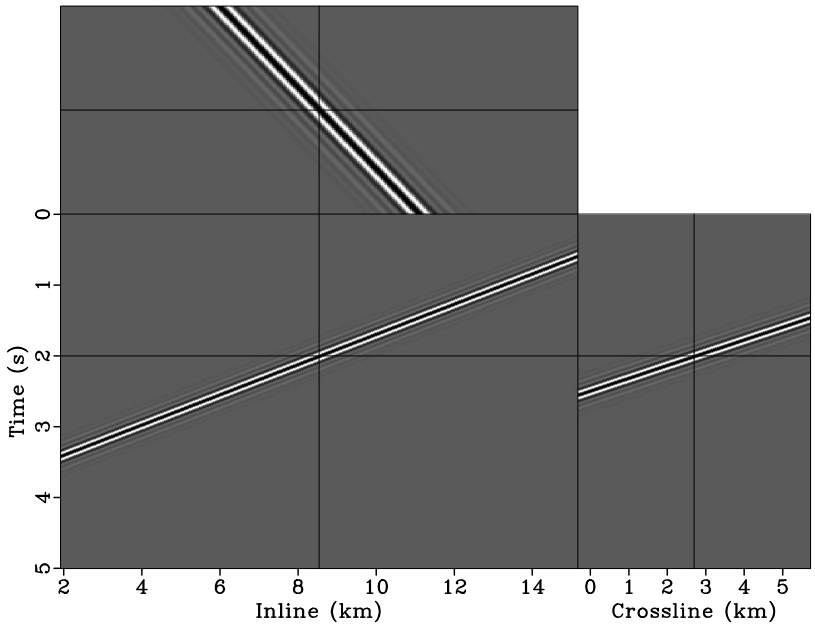
We use the normalized DSO (equation 5) to optimize the subsalt velocity. Although the 3-D Born data set is synthesized with both inline and crossline subsurface offsets, we use only inline subsurface offsets for velocity inversion due to the limited angular coverage in the crossline direction. We regularize the inversion by smoothing the gradient using a B-spline operator as follows:

$$\mathbf{g}_s = \mathbf{KBB}^*\mathbf{g}, \tag{6}$$

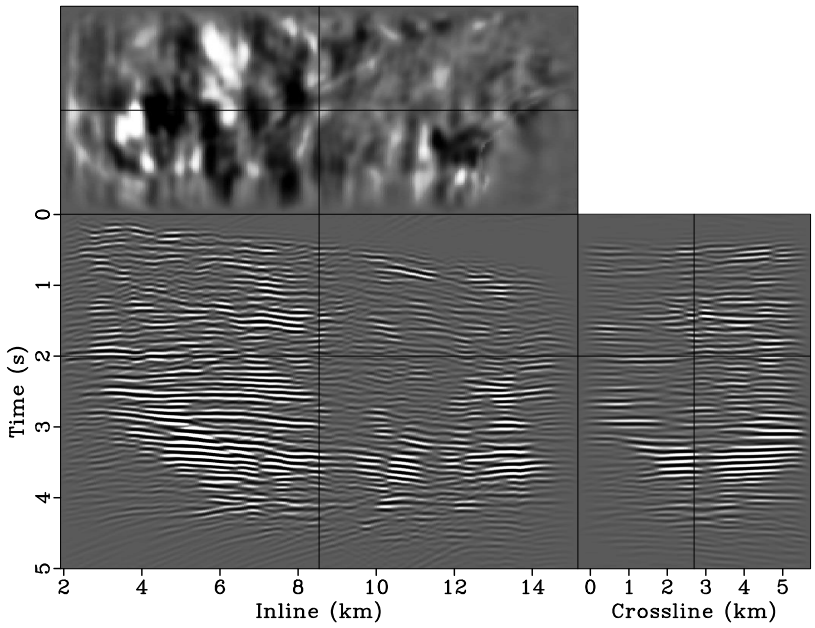
where $\mathbf{g}_s$ and $\mathbf{g}$ are the smoothed and raw gradient vectors, respectively; $\mathbf{B}$ is the B-spline projection operator. The spacing of the B-spline nodes controls the smoothness of the velocity updates. A mask operator $\mathbf{K}$ has been introduced to prevent updating the salt velocities (Figure 6).

Instead of using a fixed degree for gradient smoothing, we gradually decrease the smoothness of the gradient after every few iterations by decreasing the spacing of the B-spline nodes. We have found this strategy effective in finding an acceptable minimum, even when starting with a velocity model far from being accurate. Decreasing the smoothness of the gradient at later iterations also helps improve the resolution of the velocity model. This procedure is similar to the multi-scale inversion strategy, which has proven useful in practice (Bunks et al., 1995; Soubaras and Gratacos, 2007). Table 1 illustrates the spacings of the B-spline nodes for different iterations. Figures 7 presents the raw and smoothed gradients at different iterations.

We restart the nonlinear conjugate gradient solver every 10 iterations, and we terminate the inversion after 40 iterations when the objective function does not decrease significantly.



(a)



(b)

Figure 5: The synthesized 3-D plane-wave data. Panel (a) is the plane-wave source function with both incident dip and azimuth angles equal 10° at the top of the target region shown in Figure 1, and (b) is the corresponding plane-wave gather collected also at the top of the target region. [**CR**] yaxun1/. bpgom3d-born-tsou,bpgom3d-born-trec

Iterations	Node spacing in x	Node spacing in y	Node spacing in z
1 to 10	2.0 km	2.0 km	0.6 km
11 to 20	1.8 km	1.8 km	0.4 km
21 to 30	1.2 km	1.2 km	0.3 km
31 to 40	0.7 km	0.7 km	0.2 km

Table 1: The spacings of B-spline nodes for different iterations.

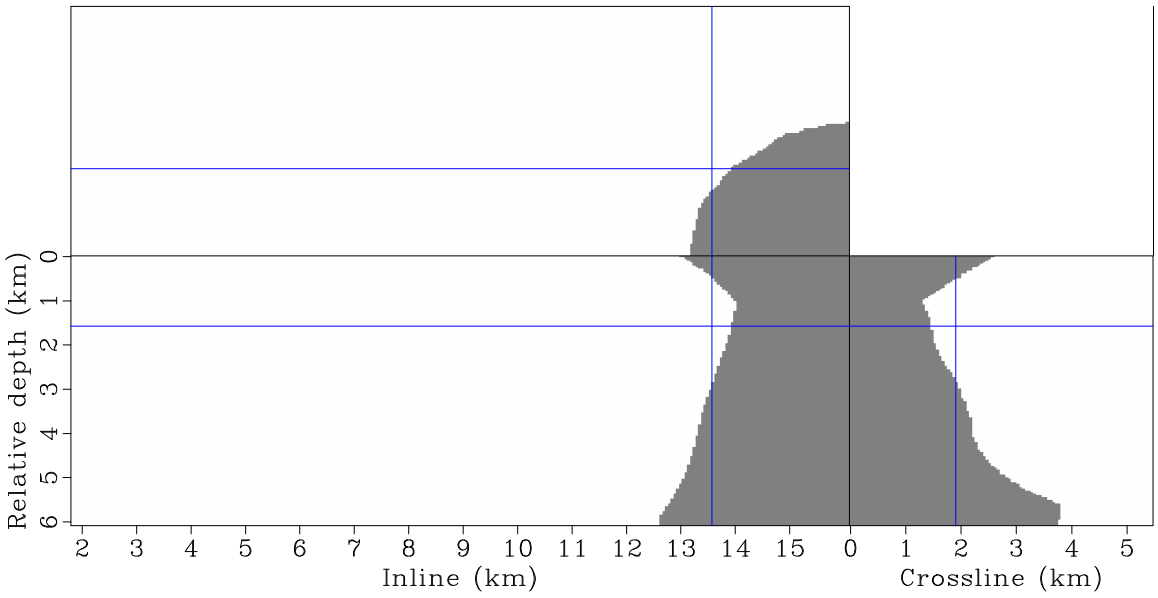


Figure 6: The mask operator applied to the gradient to prevent updating the velocity inside the salt. [**CR**] yaxun1/. bpgom3d-gmsk-target

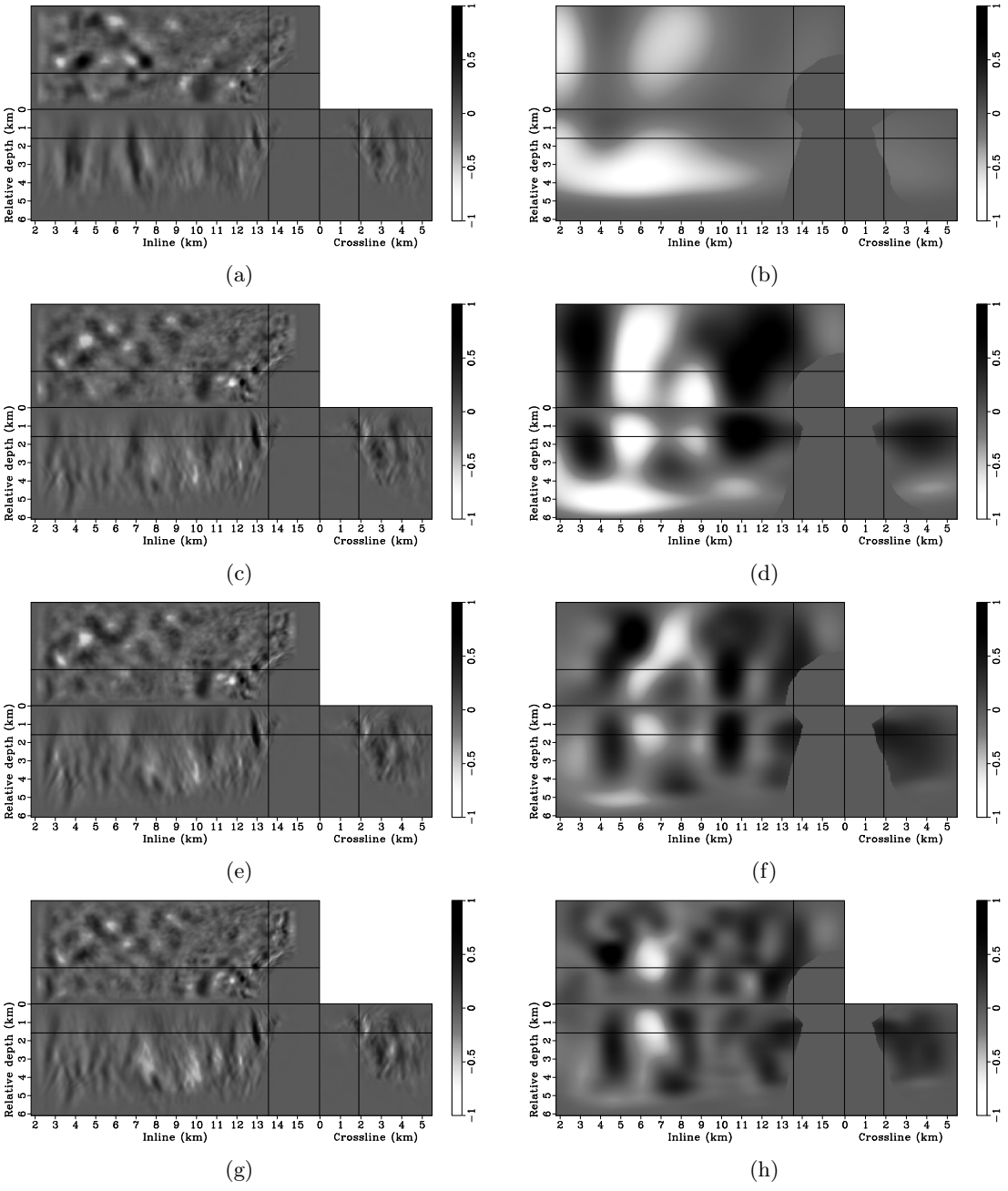


Figure 7: Panels (a), (c), (e), (g) are the raw gradients at iterations 1, 11, 21 and 31, respectively. Panels (b), (d), (f) and (h) are the corresponding gradients after applying the smoothing and mask operators. [**CR**] yaxun1/. r1,s1,r11,s11,r21,s21,r31,s31

Figure 8 shows how the objective function evolves over the first 40 iterations, whereas Figure 9 plots the step lengths as a function of iterations. It seems that the objective function converges very fast at the first several iterations, then the convergence slows down quickly, and that the velocity model gets the most significant updates at the first several iterations.

The final velocity model obtained by merging the inverted velocity model in the target region with the velocity model above the target is shown in Figure 10. It is interesting to note that the velocities beneath the salt body are slightly lower than the surrounding sediment velocities. The low velocities might indicate overpressure due to the compaction of the salt body.

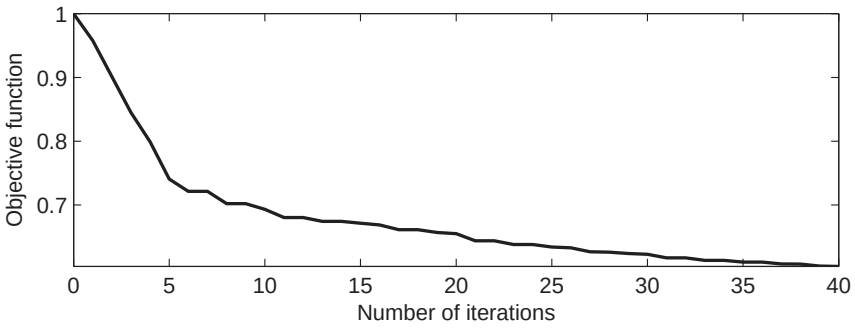


Figure 8: The evolution of the objective function over the first 40 DSO iterations. The objective function values have been normalized to 1. [**CR**] yaxun1/. fobj

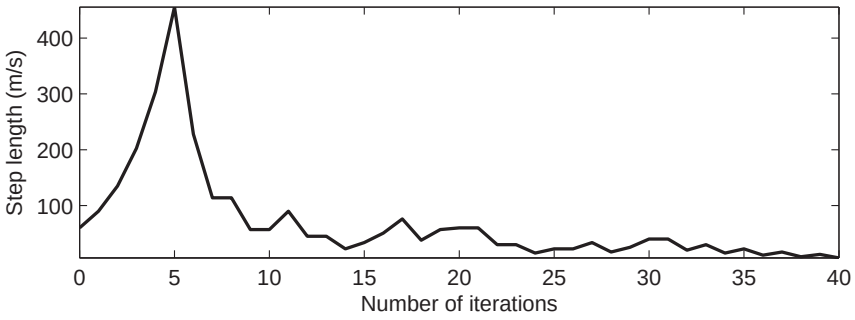


Figure 9: Step length versus the number of iterations. The initial step length is about 20 m/s. [**CR**] yaxun1/. step

Migrating the original data set using the updated velocity model

After the tomographic inversion step, we migrate the original data set using the updated velocity model. Once again, we perform 3-D conical wave migration and normalize the migrated image with the diagonal of the phase-encoded Hessian, also computed using the updated velocity model, to compensate for uneven subsurface illumination. The updated image (Figures 11(b), 12(b) and 13(b)) significantly improves the continuity of the reflectors in both the inline and crossline directions and is also more focused than the initial one (Figures 11(a), 12(a) and 13(a)). Figures 14, 15 and 16 show the corresponding angle-domain common-image gathers (ADCIGs) before and after updating the velocity model.

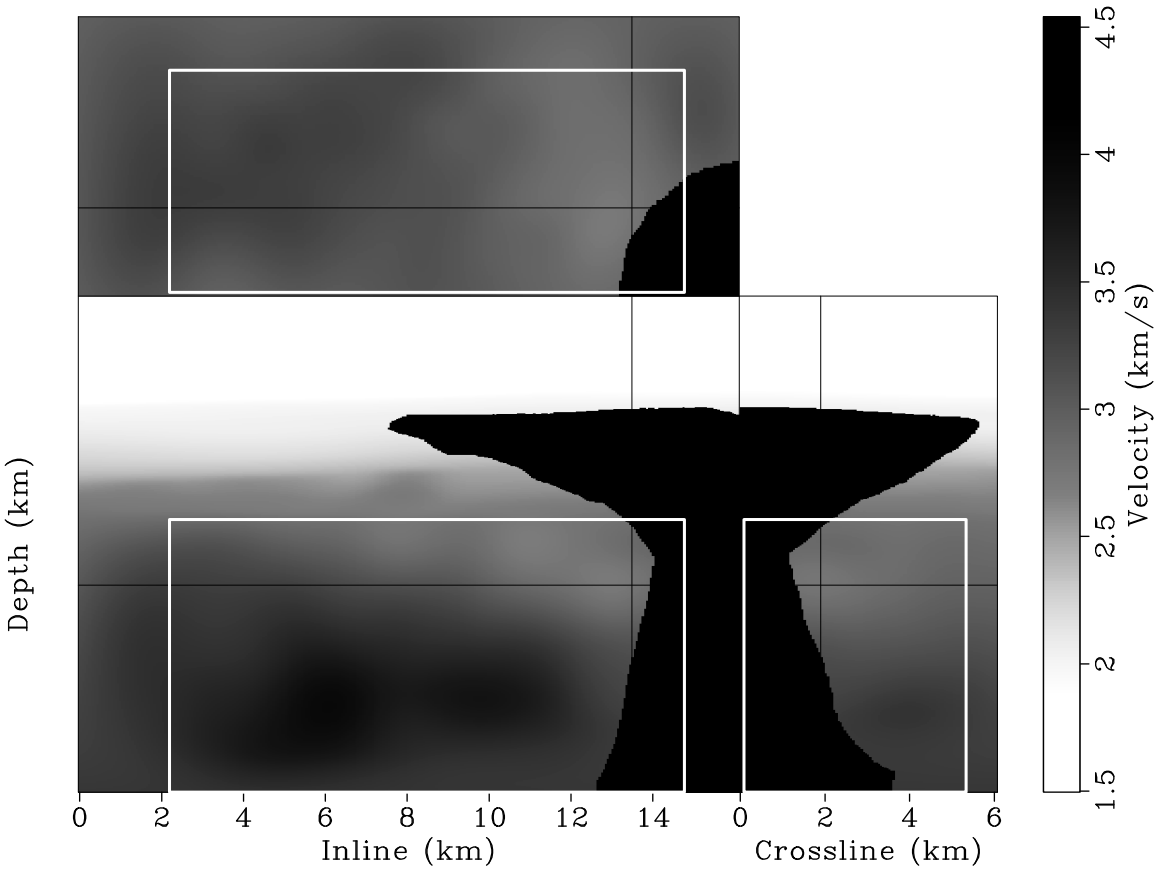


Figure 10: The final velocity model after merging the inverted velocity model in the target region with the velocity model above the target. [**CR**] yaxun1/. bpgom3d-vmod-full

The flatness and coherence of ADCIGs have been improved considerably after updating the subsalt velocities.

CONCLUSIONS

We have presented a methodology for performing image-domain wavefield tomography. Our method is extremely efficient, because we formulate the velocity estimation problem in a target-oriented fashion, and use a synthesized Born data set for velocity inversion. Our method is wavefield-based; therefore it is suitable for estimating velocities in complex geologies. Numerical examples on a 3-D field data set give excellent inversion results. The entire velocity analysis workflow can be fully automated, and no picking is necessary, although human interaction can be easily incorporated if it is desired. Thus our method provides a good tool for fast, accurate and flexible velocity model building in areas with complex geologies.

ACKNOWLEDGEMENTS

We thank BP and ExxonMobil for providing the field data set. We also thank the Stanford Center for Computational Earth and Environmental Science for providing computing resources.

REFERENCES

Albertin, U., P. Sava, J. Etgen, and M. Maharramov, 2006, Adjoint wave-equation velocity analysis: SEG Technical Program Expanded Abstracts, **25**, 3345–3349.

Biondi, B., 2006, Prestack exploding-reflectors modeling for migration velocity analysis: SEG Technical Program Expanded Abstracts, **25**, 3056–3060.

Biondi, B., S. Fomel, and N. Chemingui, 1998, Azimuth moveout for 3-d prestack imaging: Geophysics, **63**, 574–588.

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.

Bleistein, N. and H. Jaramillo, H., 2000, A platform for Kirchhoff data mapping in scalar models of data acquisition: Geophysical Prospecting, **48**, 135–161.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473.

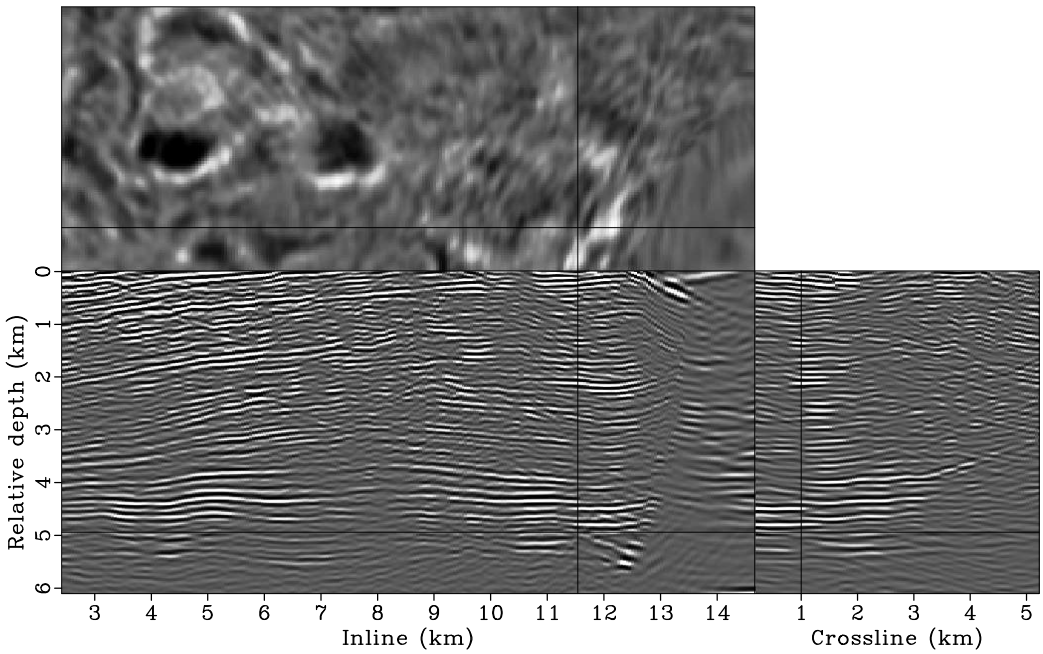
Duquet, B., P. Lailly, and A. Ehinger, 2001, 3d plane wave migration of streamer data: SEG Technical Program Expanded Abstracts, **20**, 1033–1036.

Fei, W. and P. Williamson, 2010, On the gradient artifacts in migration velocity analysis based on differential semblance optimization: SEG Technical Program Expanded Abstracts, **29**, 4071–4076.

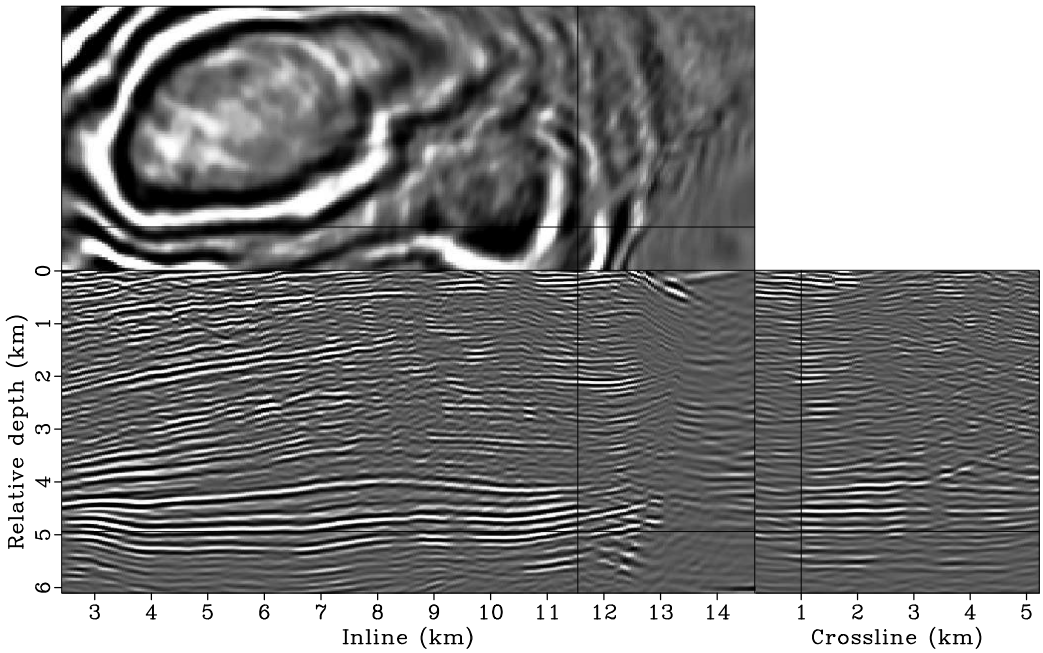
Fei, W., P. Williamson, and A. Khoury, 2009, 3-d common-azimuth wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **28**, 2283–2287.

Guerra, C., 2010, Migration-velocity Analysis Using Image-space Generalized Wavefields: PhD thesis, Stanford University.

Guerra, C., Y. Tang, and B. Biondi, 2009, Wave-equation tomography using image-space phase encoded data: SEG Technical Program Expanded Abstracts, **28**, 3964–3968.

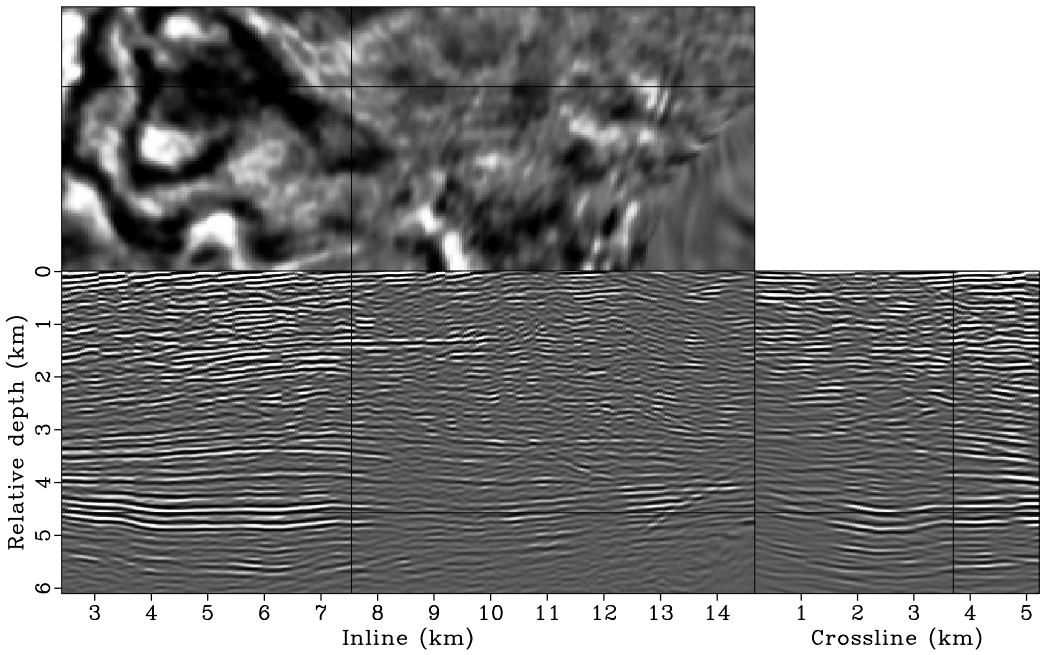


(a)

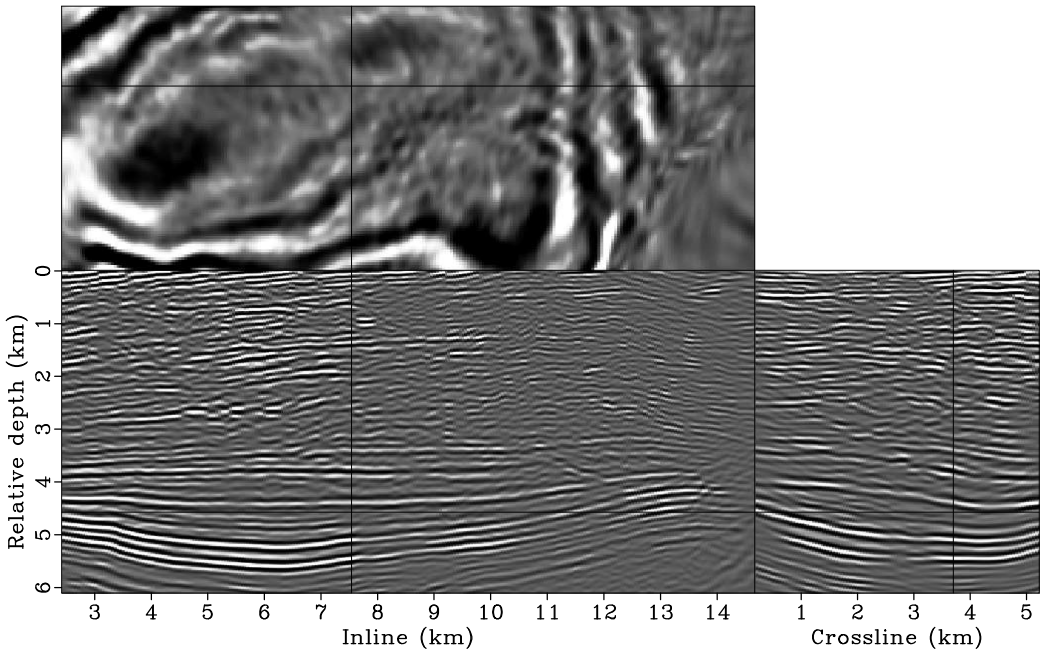


(b)

Figure 11: Comparison between the zero-subsurface-offset image (a) before and (b) after updating the velocities. Both results are obtained using the original surface-recorded data set. In both panels, the crosshair is taken at inline 11.55 km, crossline 1.00 km and relative depth 4.94 km. [**CR**] yaxun1/. bpgom3d-bimg-cpst1,bpgom3d-imag-cpst1

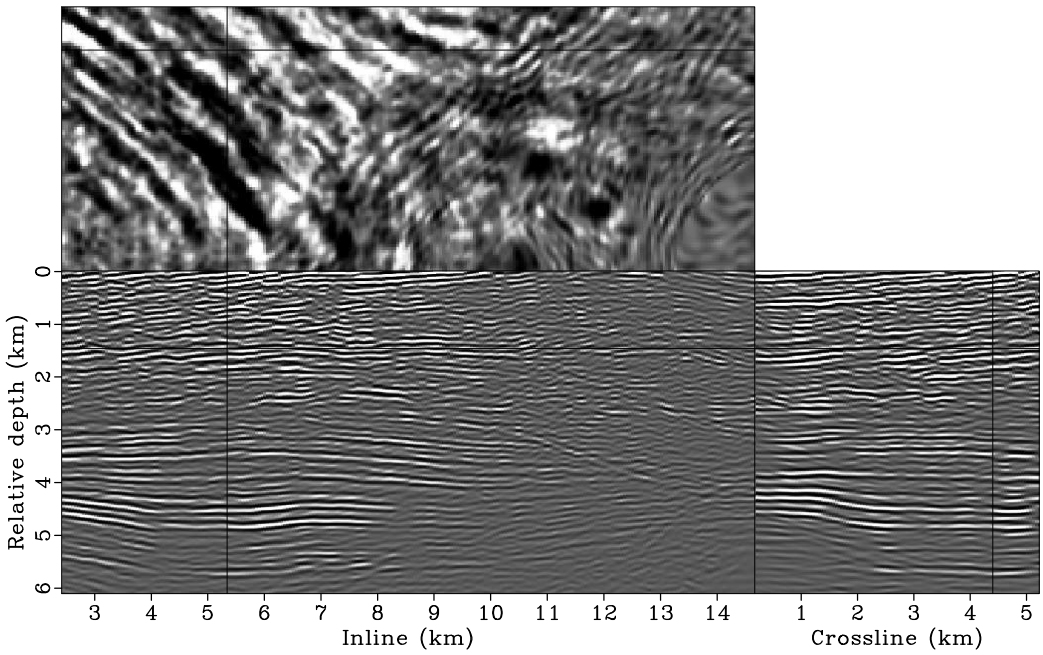


(a)

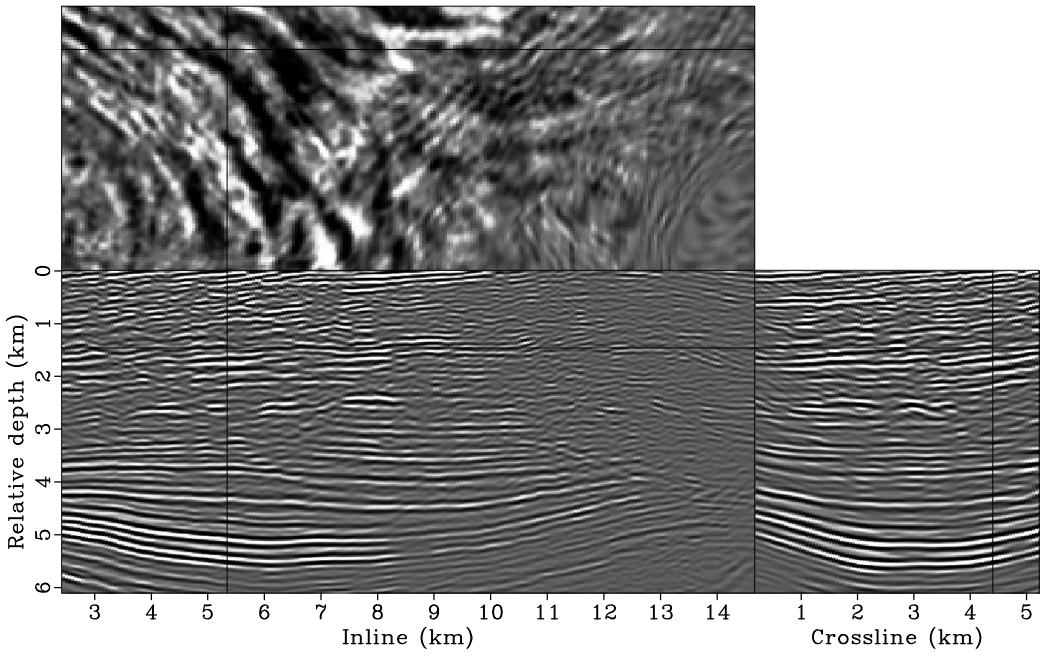


(b)

Figure 12: Comparison between the zero-subsurface-offset image (a) before and (b) after updating the velocities. Both results are obtained using the original surface-recorded data set. In both panels, the crosshair is taken at inline 7.55 km, crossline 3.70 km and relative depth 4.57 km. [**CR**] yaxun1/. bpgom3d-bimg-cpst2,bpgom3d-imag-cpst2



(a)



(b)

Figure 13: Comparison between the zero-subsurface-offset image (a) before and (b) after updating the velocities. Both results are obtained using the original surface-recorded data set. In both panels, the crosshair is taken as inline 5.35 km, crossline 4.40 km and relative depth 1.46 km. [**CR**] yaxun1/. bpgom3d-bimg-cpst3,bpgom3d-imag-cpst3

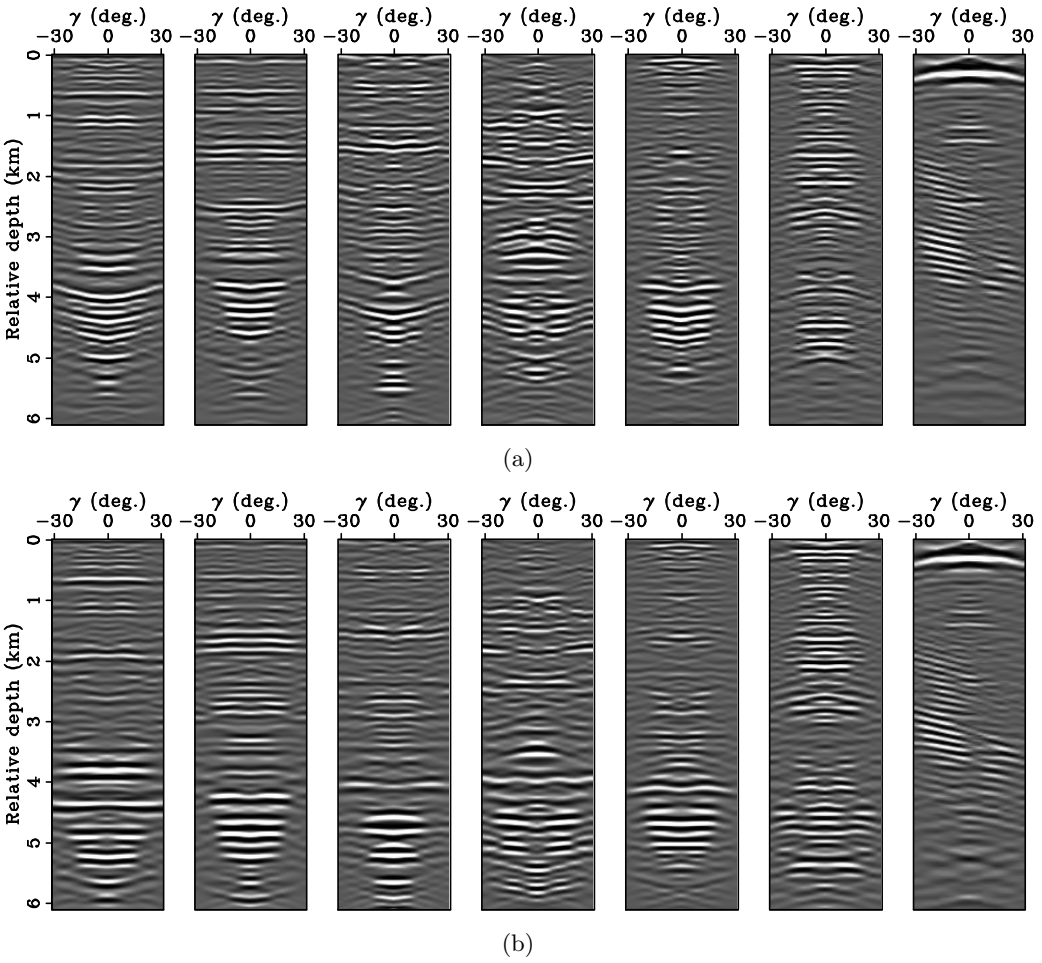


Figure 14: Comparisons between ADCIGs (a) before and (b) after updating the velocities. Both results are obtained using the original surface-recorded data set. All of the ADCIGs are extracted at the same crossline (1.00 km), but at inlines 1.52, 3.05, 4.57, 6.10, 7.62, 9.14 and 10.67 km, from left to right. [**CR**] yaxun1/. adcig2d-bvel-cpst1,adcig2d-invt-cpst1

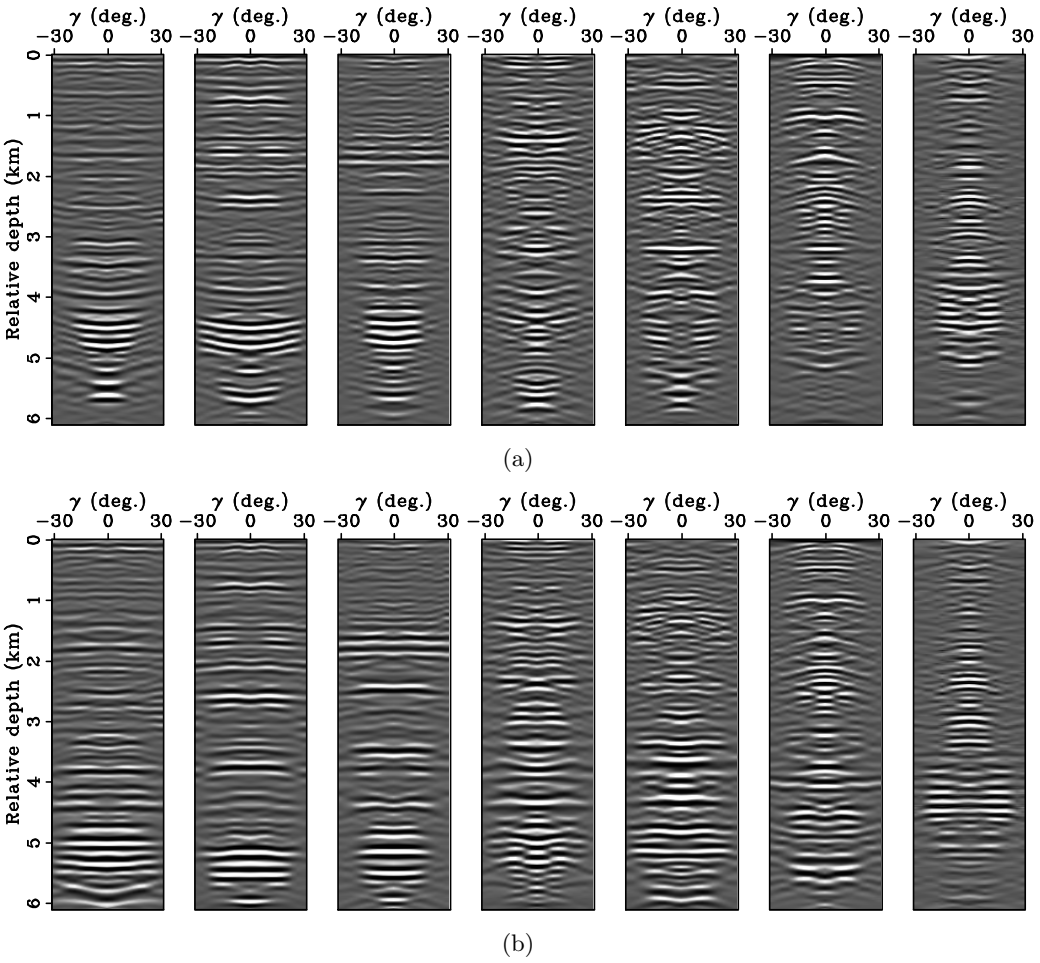


Figure 15: Comparisons between ADCIGs (a) before and (b) after updating the velocities. Both results are obtained using the original surface-recorded data set. All of the ADCIGs are extracted at the same crossline (3.70 km), but at inlines 1.52, 3.05, 4.57, 6.10, 7.62, 9.14 and 10.67 km, from left to right. [**CR**] yaxun1/. adcig2d-bvel-cpst2,adcig2d-invt-cpst2

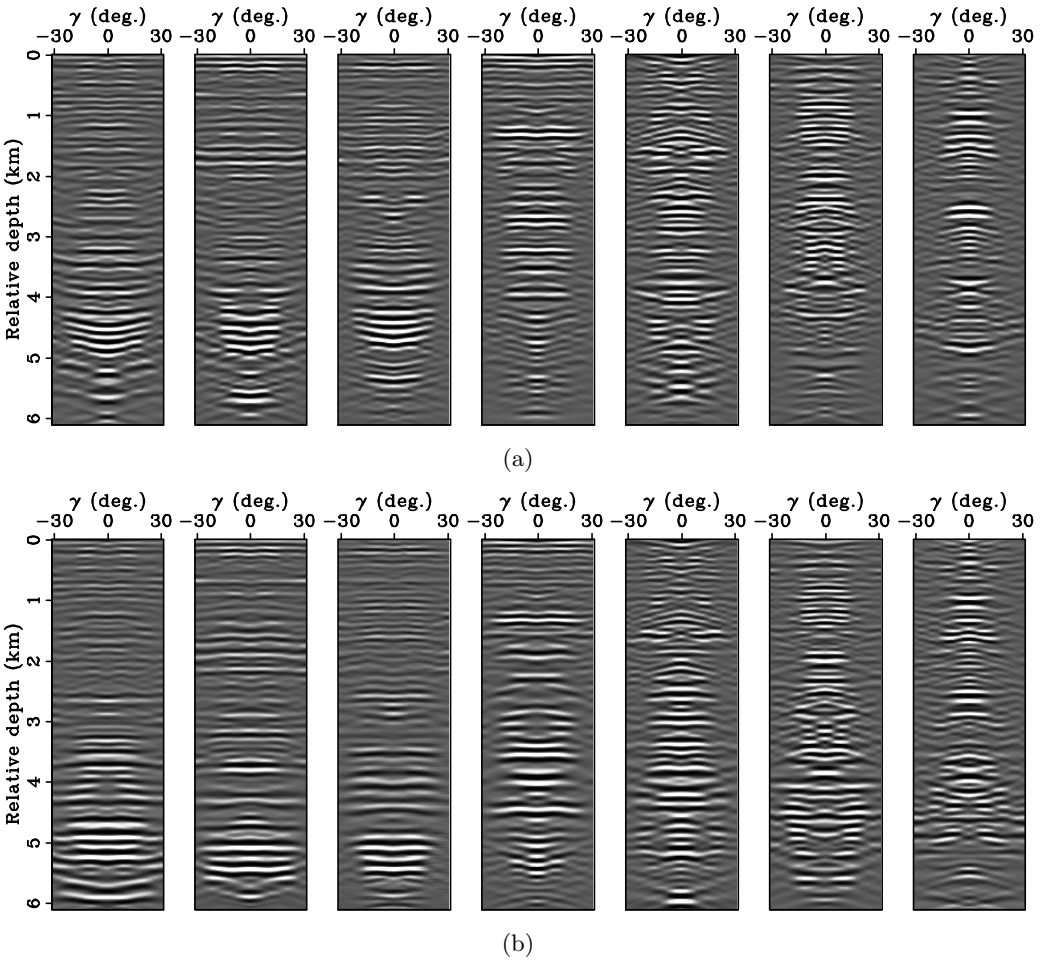


Figure 16: Comparisons between ADCIGs (a) before and (b) after updating the velocities. Both results are obtained using the original surface recorded data set. All of the ADCIGs are extracted at the same crossline (4.40 km), but at inlines 1.52, 3.05, 4.57, 6.10, 7.62, 9.14 and 10.67 km, from left to right. [**CR**] yaxun1/. adcig2d-bvel-cpst3,adcig2d-invt-cpst3

Halpert, A. D., R. G. Clapp, J. Lomask, and B. Biondi, 2008, Image segmentation for velocity model construction and updating: SEG Technical Program Expanded Abstracts, **27**, 3088–3092.

Hubral, P., J. Schleicher, and M. Tygel, 1996, A unified approach to 3-D seismic reflection imaging, part I: Basic concepts: Geophysics, **61**, 742–758.

Jin, S. and S. Xu, 2010, Visibility analysis for target-oriented reverse time migration and optimizing acquisition parameters: The Leading Edge, **29**, 1372–1377.

Liu, F., D. W. Hanson, N. D. Whitmore, R. S. Day, and R. H. Stolt, 2006, Toward a unified analysis for source plane-wave migration: Geophysics, **71**, no. 4, S129–S139.

Plessix, R.-E. and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophys. J. Int., **157**, 975–987.

Sava, P. and I. Vlad, 2008, Numeric implementation of wave-equation migration velocity analysis operators: Geophysics, **73**, VE145–VE159.

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, S. Morton, A. Hess, and H. Calandra, 2005, Differential semblance velocity analysis via shot profile migration: SEG Technical Program Expanded Abstracts, **24**, 2249–2252.

Soubaras, R. and B. Gratacos, 2007, Velocity model building by semblance maximization of modulated-shot gathers: Geophysics, **72**, U67–U73.

Stolt, R. H. and A. Benson, 1986, Seismic Migration: Theory and Practice: Geophysical Press.

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

Tang, Y., 2007, Selective stacking in the reflection-angle and azimuth domain: SEG Technical Program Expanded Abstracts, **26**, 2320–2324.

———, 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian: Geophysics, **74**, WCA95–WCA107.

Tang, Y. and B. Biondi, 2010, Target-oriented wavefield tomography using demigrated Born data: SEG Technical Program Expanded Abstracts, **29**, 4280–4285.

———, 2011, Subsalt imaging by target-oriented wavefield tomography: A 3-D field-data example: **SEP-143**, xxx–xxx.

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

Tang, Y. and S. Lee, 2010, Preconditioning full waveform inversion with phase-encoded hessian: SEG Technical Program Expanded Abstracts, **29**, 1034–1038.

Tarantola, A., 2005, Inverse problem theory: Methods for data fitting and model parameter estimation: Society for Industrial and Applied Mathematics.

Valenciano, A., 2008, Imaging by Wave-equation Inversion: PhD thesis, Stanford University.

Vyas, M. and Y. Tang, 2010, Gradients for wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **29**, 4077–4081.

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

Zhang, Y., J. Sun, C. Notfors, S. H. Gray, L. Chernis, and J. Young, 2005, Delayed-shot 3d depth migration: Geophysics, **70**, E21–E28.

Migration velocity analysis for anisotropic models

Yunyue (Elita) Li and Biondo Biondi

ABSTRACT

Anisotropic models are recognized as more realistic representations of the subsurface in complex geological environments. These models are widely needed by many kinds of migration and interpretation schemes. However, anisotropic model building is still a challenging problem in the industry. In this paper, we propose an approach to building anisotropic models from surface seismic data based on the theory of Wave-Equation Migration Velocity Analysis (WEMVA). Because of the ambiguity between depth and Thomsen parameter δ, we parametrize our model space using only NMO velocity (V_{nmo}) and the anellipticity parameter η. We tested the anisotropic WEMVA on a shallow part of the Hess synthetic VTI model. The results show that anisotropic WEMVA is effective in resolving some of the anisotropic perturbation. However, a unique solution to the inversion requires additional constraining information.

INTRODUCTION

Since first reported in exploration seismology in the 1930s (McCollum and Snell, 1932), anisotropy has become increasingly important in seismic imaging and exploration. Until now, the transverse isotropic (TI) model has been the most commonly used model in seismic imaging. Postma (1955), Helbig (1956) and Backus (1962) have shown that a sequence of isotropic layers on a scale much smaller than the wavelength leads to an anisotropic medium. If the layers are horizontal, the medium is defined as a vertical TI (VTI) medium. A VTI medium is commonly formed because of thin bedding during deposition. If the layers become dipping due to deformation, a tilted TI (TTI) medium is formed. Many authors (Shan, 2009; Fletcher et al., 2009; Zhang and Zhang, 2009; Fei and Liner, 2008) have developed migration and processing schemes for VTI and TTI media; however, the challenge of estimating the anisotropy model remains a bottleneck for the exploration workflow.

The existing anisotropic model-building schemes are mostly based on measuring the non-hyperbolic moveout along the traveltime curve to flatten the common image gathers (CIG) (Zhou et al., 2003, 2004; Yuan et al., 2006; Cai et al., 2009; Woodward et al., 2008). However, traveltime ray-based methods are prone to errors and unrealistic results when multi-pathing exists in areas of complex overburden. Hence, we propose to apply wave-equation tomography for anisotropic model building.

Wave-equation tomography has been widely studied in isotropic velocity building and can be implemented either in the data space, commonly known as Full-Waveform Inversion (FWI) (Tarantola, 1984; Woodward, 1992) or in the image space, commonly known as Wave-Equation Migration Velocity Analysis (WEMVA) (Sava and Biondi, 2004a,b; Shen, 2004; Shen and Symes, 2008; Guerra et al., 2009). Several advantages drive us to use the image-space wave-equation tomography instead of data-space wave-equation tomography: first,

WEMVA does not require as accurate an initial model to avoid the cycle-skipping problem as FWI requires. In fact, many studies (Guerra et al., 2009; Guerra and Biondi, 2010; Tang and Biondi, 2010) show that the resolution gap between ray-based tomography and FWI could be linked by the image-space WEMVA method; second, the objective function is directly related to the final image; third, the migrated image is often much cleaner than the recorded wavefields. Therefore, we choose to extend image-space WEMVA from isotropic velocity building to anisotropic model building.

In this paper, we first generalize the methodology of image-space WEMVA from an isotropic medium to an anisotropic medium and explain our parameterization. We show that the gradient of the tomographic objective functional for an isotropic medium can be modified to describe an anisotropic medium by simply adding a term for the additional parameter. Finally, we test our inversion scheme on a shallow part of the Hess anisotropic synthetic dataset.

MIGRATION VELOCITY ANALYSIS FOR ANISOTROPIC PARAMETERS

Anisotropic MVA is a non-linear inversion process that aims to find the background anisotropic model that minimizes the residual image $\Delta\mathbf{I}$. The residual image is derived from the background image $\mathbf{I}$, which is computed with the current background model. To form the image, both the source and receiver wavefields are downward continued using the one-way wave equations. Assuming that the shear velocity is much smaller than the P-wave velocity, one way of formulating up-going and down-going one-way acoustic wave equations for VTI is shown as follows (Shan, 2009):

$$\left(\frac{\partial}{\partial z} \mp i\Lambda\right) P = 0, \tag{1}$$

where $P = P(x, y, z, \omega)$ is the wavefield in the space-frequency domain and Λ describes the dispersion relationship in terms of P-wave vertical slowness s_0 and Thomsen parameters ϵ and δ (Thomsen, 1986):

$$\Lambda = \omega s_0 \sqrt{\frac{\omega^2 s_0^2 - (1+2\epsilon)|\mathbf{k}|^2}{\omega^2 s_0^2 - 2(\epsilon-\delta)|\mathbf{k}|^2}}, \tag{2}$$

where $\mathbf{k} = (k_x, k_y)$ is the spatial wavenumber vector.

Many authors (Tsvankin and Thomsen, 1994; Alkhalifah and Tsvankin, 1995) have shown that P-wave traveltime can be characterized by the NMO slowness, s_n, and the anellipticity parameter η. Therefore, the one-way wave-equation in terms of s_n, η and δ is:

$$\left(\frac{1}{\sqrt{1+2\delta}}\frac{\partial}{\partial z} \mp i\Lambda'\right) P = 0 \tag{3}$$

where

$$\Lambda' = \omega s_n \sqrt{1 - \frac{|\mathbf{k}|^2}{\omega^2 s_n^2 - 2\eta|\mathbf{k}|^2}}. \tag{4}$$

Notice that in the dispersion relationship in Equation 3, δ and the derivative in depth $\frac{\partial}{\partial z}$, are coupled with each other. This is a theoretical proof of the well-accepted observation that δ cannot be determined by the surface seismic data. To constrain this parameter, we

need well information (e.g. checkshots) to add the depth dimension into the inversion. Now, if we apply the change of variables

$$d\bar{z} = \sqrt{1+2\delta}dz \tag{5}$$

and neglect the derivatives of δ, Equation 3 becomes

$$\left(\frac{\partial}{\partial \bar{z}} \mp i\Lambda'\right) P = 0. \tag{6}$$

We can therefore formulate the image-space migration velocity analysis problem with NMO slowness s_n and anisotropic parameters η and δ, but we invert only for s_n and η assuming δ model is known from other source of information.

Notice that when $\eta = 0$, the dispersion relationship (equation 4) is the same as the isotropic dispersion relationship, and the corresponding one-way wave equation (equation 6) is almost the same as for the isotropic case, except for a depth stretch caused by δ. In other words, an elliptic anisotropic wavefield inversion is almost equivalent to an isotropic wavefield inversion. Plessix and Rynja (2010) reached the same conclusions for full-waveform inversion (FWI). Figure 1 compares the original NMO velocity to the stretched NMO velocity. Notice that the geological features are stretched downward for positive δ. Because we ignore δ in the inversion, we expect the inverted NMO velocity to have more similarity to the stretched NMO velocity than to the original one.

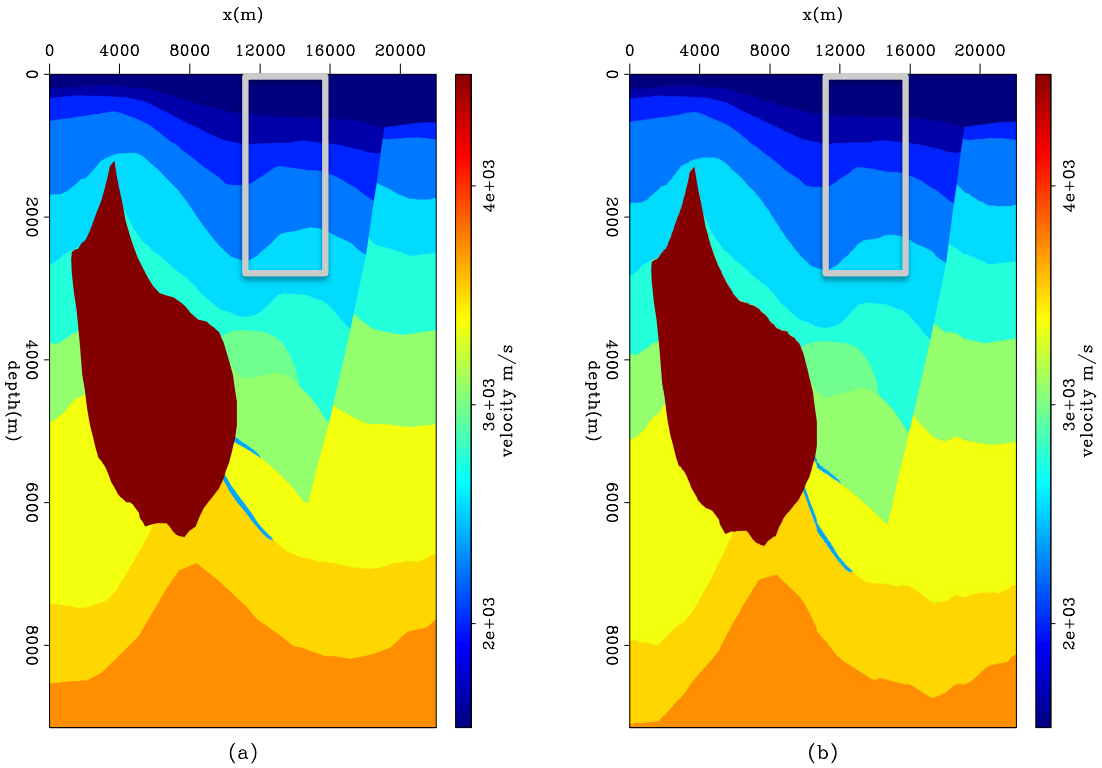


Figure 1: (a) Original NMO velocity for the anisotropic Hess model; (b) Stretched NMO velocity according to δ. The overlaid box denotes the part of model that we work on in the numerical test. [**CR**] elita1/. nmo

In general, the residual image is defined as (Biondi, 2008)

$$\Delta\mathbf{I} = \mathbf{I} - \mathbf{F}(\mathbf{I}), \tag{7}$$

where $\mathbf{F}$ is a focusing operator. In the least-square sense, the tomographic objective function can be written as follows:

$$J = \frac{1}{2}||\Delta\mathbf{I}||_2 = \frac{1}{2}||\mathbf{I} - \mathbf{F}(\mathbf{I})||^2. \tag{8}$$

To perform MVA for anisotropic parameters, we first need to extend the tomographic operator from the isotropic medium (Shen, 2004; Sava, 2004; Guerra et al., 2009) to the anisotropic medium. We define the wave-equation tomographic operator $\mathbf{T}$ for anisotropic models as follows:

$$\begin{aligned} \mathbf{T} &= \left.\frac{\partial \mathbf{I}}{\partial \mathbf{m}}\right|_{\mathbf{m}=\widehat{\mathbf{m}}} \\ &= \left.\frac{\partial \mathbf{I}}{\partial \mathbf{s_n}}\right|_{\mathbf{s_n}=\widehat{\mathbf{s_n}}} + \left.\frac{\partial \mathbf{I}}{\partial \eta}\right|_{\eta=\widehat{\eta}} \end{aligned} \tag{9}$$

where $\mathbf{m}$ is the anisotropy model, which in this case includes NMO slowness $\mathbf{s_n}$ and anellipticity parameter η; $\widehat{\mathbf{m}}$ is the background anisotropy model, consisting of the background NMO slowness $\widehat{\mathbf{s_n}}$ and background anellipticity $\widehat{\eta}$; and $\mathbf{I}$ is the image. This wave equation tomographic operator $\mathbf{T}$ is a linear operator that relates the model perturbation $\Delta\mathbf{m}$ to the image perturbation $\Delta\mathbf{I}$ as follows:

$$\Delta\mathbf{I} = \mathbf{T}\Delta\mathbf{m}. \tag{10}$$

In the shot-profile domain, both source and receiver wavefields are downward continued using the one-way wave equation (6):

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\Lambda'\right) D(\mathbf{x}, \mathbf{x}_s) = 0 \\ D(x, y, z = 0, \mathbf{x}_s) = f_s \delta(\mathbf{x} - \mathbf{x}_s) \end{cases}, \tag{11}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\Lambda'\right) U(\mathbf{x}, \mathbf{x}_s) = 0 \\ U(x, y, z = 0, \mathbf{x}_s) = Q(x, y, z = 0, \mathbf{x}_s) \end{cases}, \tag{12}$$

where $D(\mathbf{x}, \mathbf{x}_s)$ is the source wavefield at the image point $\mathbf{x} = (x, y, z)$ with the source located at $\mathbf{x}_s = (x_s, y_s, 0)$; $U(\mathbf{x}, \mathbf{x}_s)$ is the receiver wavefield at the image point $\mathbf{x}$ with the source located at $\mathbf{x}_s$; f_s is the source signature, and $f_s\delta(\mathbf{x} - \mathbf{x}_s)$ defines the point-source function at $\mathbf{x_s}$, which serves as the boundary condition of equation 11; and $Q(x, y, z = 0, \mathbf{x}_s)$ is the recorded shot gather for the shot located at $\mathbf{x}_s$, which serves as the boundary condition of equation 12.

The dispersion relationship in equation (4) can be approximated with a rational function by Taylor series and Padé expansion analysis (Shan, 2009):

$$\Lambda' = \omega s_n \left(1 - \frac{a|\mathbf{k}|^2}{\omega^2 s_n^2 - b|\mathbf{k}|^2}\right), \tag{13}$$

where, to the second order of the expansion, $a = 0.5, b = 0.25 + 2\eta$. Equation (13) using binomial expansion can be further expanded to polynomials:

$$\Lambda' = \omega s_n - \frac{a}{\omega^2 s_n^2}|\mathbf{k}|^2 - \frac{ab}{\omega^4 s_n^4}|\mathbf{k}|^4. \tag{14}$$

Now it is straightforward to take the derivative of Λ' with respect to s_n and η.

The background image is computed by applying the cross-correlation imaging condition:

$$I(\mathbf{x},\mathbf{h}) = \sum_{\mathbf{x}_s}\sum_{\omega} \overline{D(\mathbf{x}-\mathbf{h},\mathbf{x}_s)} U(\mathbf{x}+\mathbf{h},\mathbf{x}_s), \tag{15}$$

where the overline stands for complex conjugate, and $\mathbf{h} = (h_x, h_y, h_z)$ is the subsurface half-offset. Perturbing the wavefields in equation (15) and ignoring the higher-order term, we can get the perturbed image as follows:

$$\begin{aligned}\Delta I(\mathbf{x},\mathbf{h}) &= \sum_{\mathbf{x}_s}\sum_{\omega} \Big(\overline{\Delta D(\mathbf{x}-\mathbf{h},\mathbf{x}_s)} \widehat{U}(\mathbf{x}+\mathbf{h},\mathbf{x}_s) + \\ &\quad \overline{\widehat{D}(\mathbf{x}-\mathbf{h},\mathbf{x}_s)} \Delta U(\mathbf{x}+\mathbf{h},\mathbf{x}_s) \Big),\end{aligned} \tag{16}$$

where $\widehat{D}(\mathbf{x}-\mathbf{h},\mathbf{x}_s)$ and $\widehat{U}(\mathbf{x}+\mathbf{h},\mathbf{x}_s)$ are the background source and receiver wavefields computed with the background model $\widehat{m}(\mathbf{x})$; and $\Delta D(\mathbf{x}-\mathbf{h},\mathbf{x}_s)$ and $\Delta U(\mathbf{x}+\mathbf{h},\mathbf{x}_s)$ are the perturbed source wavefield and perturbed receiver wavefield, which are the results of the model perturbation $\Delta m(\mathbf{x})$.

To evaluate the adjoint tomographic operator $\mathbf{T}^*$, which maps from the image perturbation to the model perturbation, we first compute the wavefield perturbation from the image perturbation using the adjoint imaging condition:

$$\begin{aligned}\Delta D(\mathbf{x},\mathbf{x}_s) &= \sum_{\mathbf{h}} \Delta I(\mathbf{x},\mathbf{h}) \widehat{U}(\mathbf{x}+\mathbf{h},\mathbf{x}_s) \\ \Delta U(\mathbf{x},\mathbf{x}_s) &= \sum_{\mathbf{h}} \Delta I(\mathbf{x},\mathbf{h}) \widehat{D}(\mathbf{x}-\mathbf{h},\mathbf{x}_s).\end{aligned} \tag{17}$$

The perturbed source and receiver wavefields satisfy the following one-way wave equations, linearized with respect to NMO slowness and η:

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\Lambda'\right) \Delta D(\mathbf{x},\mathbf{x}_s) = \left(-i\frac{\partial \Lambda'}{\partial \mathbf{m}} \widehat{D}(\mathbf{x},\mathbf{x}_s)\right) \Delta \mathbf{m}^*(\mathbf{x}) \\ \Delta D(x,y,z=0,\mathbf{x}_s) = 0 \end{cases}, \tag{18}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\Lambda'\right) \Delta U(\mathbf{x},\mathbf{x}_s) = \left(-i\frac{\partial \Lambda'}{\partial \mathbf{m}} \widehat{U}(\mathbf{x},\mathbf{x}_s)\right) \Delta \mathbf{m}^*(\mathbf{x}) \\ \Delta U(x,y,z=0,\mathbf{x}_s) = 0 \end{cases}, \tag{19}$$

where $\mathbf{m}$ is the row vector $[\mathbf{s_n}\ \eta]$.

During the inversion, the model perturbation is unknown, and in fact must be estimated. Therefore, we obtain the image perturbation by applying a focusing operator (equation 7) to the current background image. Then the perturbed image is convolved with the background wavefields to get the perturbed wavefields (equation 17). The scattered wavefields are obtained by applying the adjoint of the one-way wave-equations (18) and (19). Finally, the model-space gradient is obtained by cross-correlating the upward propagated scattered wavefields with the modified background wavefields [the terms in the parentheses on the right-hand sides of equations (18) and (19)].

OBJECTIVE FUNCTION

As mentioned in the previous section, we estimate the optimum earth model by minimizing a user-defined image perturbation. There are many ways to define the objective function. Here we use the Differential Semblance Optimization (DSO) method (Symes and Carazzone, 1991; Shen, 2004) as the criterion:

$$\mathbf{F}(\mathbf{I}) = (\mathbf{1} - \mathbf{O})\mathbf{I}, \tag{20}$$

where $\mathbf{1}$ is the identity operator and $\mathbf{O}$ is the differential operator along the angle axes in the ADCIGs $\mathbf{I}$. In the subsurface-offset domain, the objective function (Equation 8) reads:

$$J = \frac{1}{2}||h\mathbf{I}(\mathbf{x}, \mathbf{h})||_2, \tag{21}$$

where h is the absolute value of subsurface-offset, and $\mathbf{I}(\mathbf{x}, \mathbf{h})$ is the image gather in the subsurface-offset domain. This operator is preferred by many researchers since it is a fully automated procedure, with no picking required. However, for isotropic migration velocity analysis, many authors (Vyas and Tang, 2010; Fei and Williamson, 2010) observe undesired artifacts generated by the DSO operator and suggest that a differential operator along h can help compensate for the phase shift caused by the velocity perturbation. Therefore, we use the modified DSO operator as follows:

$$J = \frac{1}{2}||h\mathbf{D}\mathbf{I}(\mathbf{x}, \mathbf{h})||_2, \tag{22}$$

where $\mathbf{D}$ is a differential operator in $\mathbf{h}$. Taking the derivative in the subsurface offset domain is equivalent to an α weighting in the angle domain. Therefore, the objective function (Equation 22) also emphasizes the contribution of the large angle information, which is crucial for velocity analysis.

To guarantee a smooth inversion, we choose a B-spline representation of the model space. The smoothed gradient in the original space is then represented as:

$$\hat{\mathbf{g}} = \mathbf{B}\mathbf{B}^*\mathbf{g}, \tag{23}$$

where $\mathbf{g}$ and $\hat{\mathbf{g}}$ are the original and the smoothed gradient on the original model grid; $\mathbf{B}$ is the B-spline projection operator. Then the number and spacing of the B-spline nodes control the smoothness of the model update. Practically, we can choose different B-spline parameters for velocity and η.

NUMERICAL TEST

We test our inversion scheme on the shallow part of the Hess synthetic anisotropic model, as denoted by the gray square in Figure 1. The initial model is a 1D gradient isotropic model from the seabed. Figure 2 compares the inversion results with the true models. The ratio of intial velocity and inverted velocity over the true stretched NMO velocity are shown in Figure 2(a) and 2(b), respectively. The error in the initial NMO velocity is up to 25%, which is far beyond the tolerance of FWI. The anisotropic WEMVA successfully reduces the error in velocity down to less than 5%. Notice that the error in velocity generally follows

the dip in the image. This suggests that we should use better smoothing operators such as dip (steering) filters (Hale, 2007; Clapp, 2000) to regularize the inversion.

On the other hand, the η update [Figure 2(d)] is in general larger than the true η model [Figure 2(c)]. A trade-off is observed below 1,600 m, where the inverted velocity is smaller but η is much larger than the true values. This result illustrates the null space of our inversion problem, since the reflector around 2,200 m is well-focused (although not perfectly focused) in the final image obtained with the inverted model [Figure 3(b)]. This problem can presumably be resolved by increasing the angle coverage at depth and allowing more iterations in the inversion.

Figure 3 compares the subsurface-offset images using the initial model (a), the updated model (b), and the true model (c). After the inversion, the reflectors are focused at zero subsurface-offset, and the depths of the reflectors are closer to the true depths. The focused image shows that we are dealing with a non-linear problem with a large null space. To reduce the size of the null space, and hence the uncertainty in the inverted model, other information such as checkshots or rock-physics prior knowledge is needed (Li et al., 2011a,b).

CONCLUSION AND DISCUSSION

We have presented a new methodology for performing image-domain migration velocity analysis in anisotropic media. Our method is a natural extension of isotropic MVA theory and retains the same properties as isotropic MVA. We demostrate our method on a 2-D synthetic data set. After inversion, we obtain better-focused subsurface-offset images and better-defined depths. By including the geological information and the wider-offset data, we should be able to eliminate the model error at depth.

Experience shows that the DSO operator has a layer-stripping effect during the iterations. One cause of this effect is the unbalanced amplitude for the reflectors in depth. Therefore, an illumination-corrected image is preferred to compensate for this effect. On the other hand, a residual-moveout-based objective function (Sava, 2004; Sava and Biondi, 2004a,b; Almomin, 2011; Zhang and Biondi, 2011) could avoid the problem.

Compared with ray-based image-space model-building methods, our wavefield-based image-space method is computationally more intensive. However, the wavefield method better approximates wave propagation in complex areas. We can also utilize the phase-encoded target-oriented image-space wavefield tomography (Guerra et al., 2009; Guerra and Biondi, 2010) technique to reduce the computational cost.

Finally, by introducing another parameter η into the MVA inversion, we now have a larger model space and hence a larger null space with respect to the same data. Therefore, the surface reflection seismic data is inadequate for resolving a unique earth model. Other information, such as borehole measurement, geological interpretation (Bakulin et al., 2010), or rock-physics prior knowledge (Li et al., 2011a,b), is necessary to obtain a consistent, unique and reliable earth model.

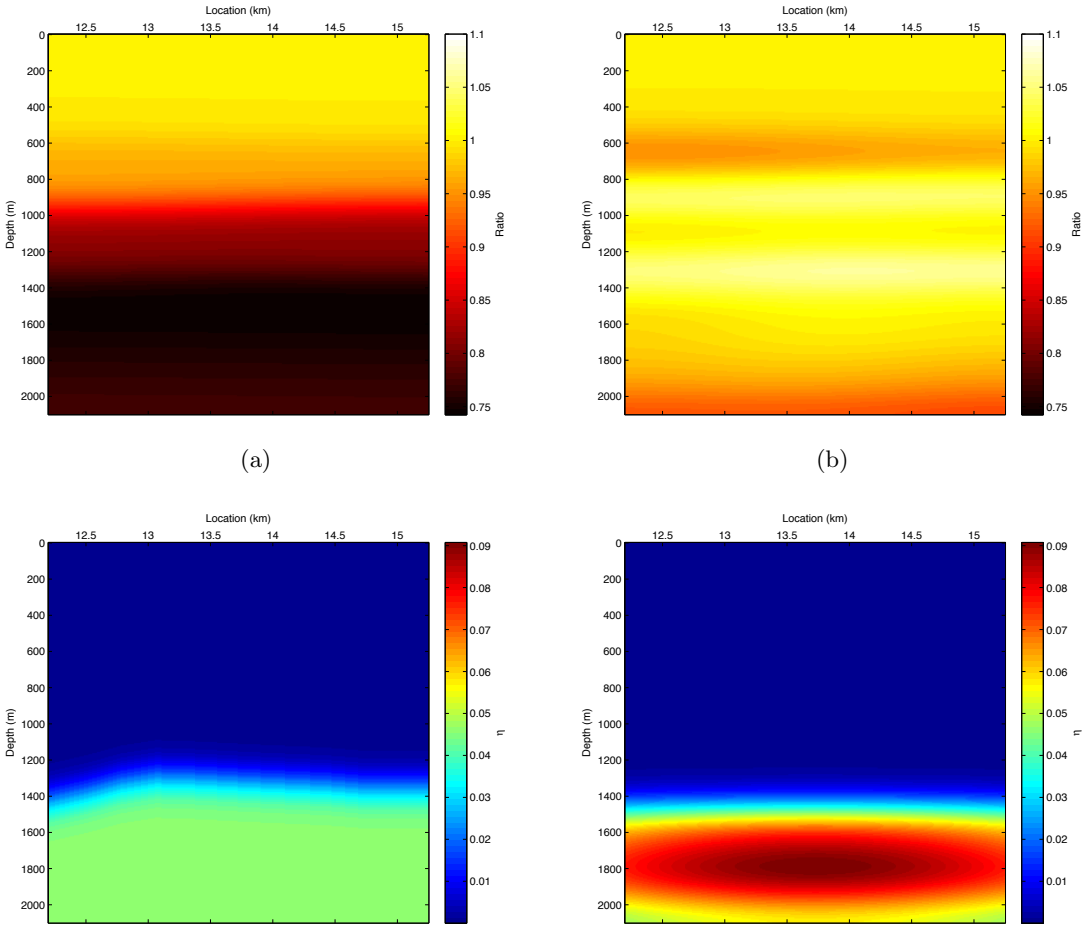


Figure 2: (a) Ratio of initial velocity over true velocity; (b) ratio of inverted velocity over true velocity; (c) true η model; (d) inverted η model. [**CR**] elita1/. init-over-true,inv-over-true,true-eta,inv-eta

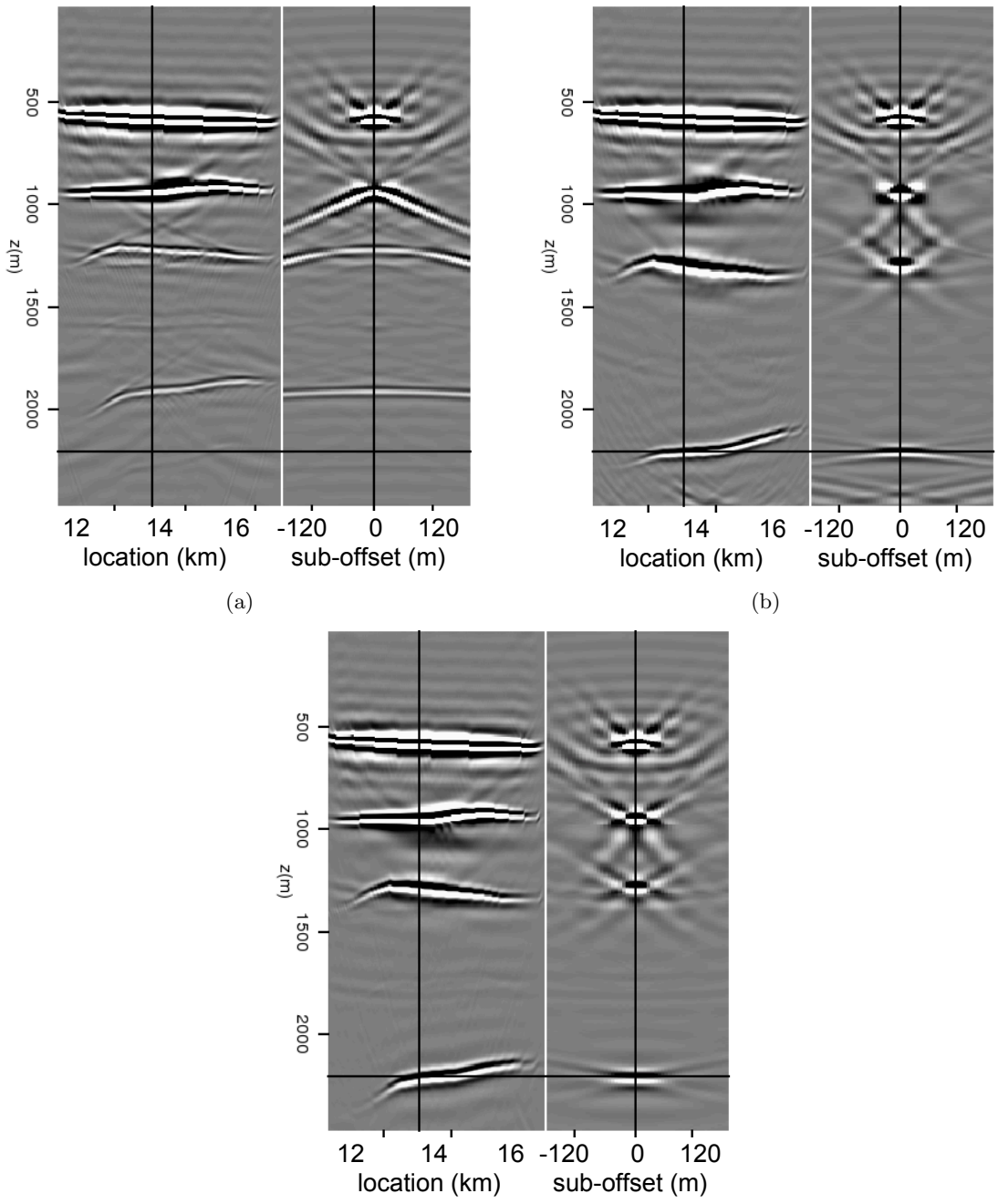


Figure 3: Subsurface offset images using the initial model (a), the updated model (b), and the true model (c). [**CR**] elita1/. init-image,final-image,true-image

ACKNOWLEDGMENT

We would like to thank James Berryman for fruitful discussions. We thank Hess for providing the Hess anisotropic synthetic dataset.

REFERENCES

Alkhalifah, T. and I. Tsvankin, 1995, Velocity analysis for transversely isotropic media: Geophysics, **60**, 1550–1566.

Almomin, A., 2011, Correlation-based wave-equation migration velocity analysis: SEP-Report, **143**, 33–42.

Backus, G. E., 1962, Long-wave elastic anisotropy introduced by horizontal layering: J. Geophys. Res., **67**, 4427–4440.

Bakulin, A., M. Woodward, Y. Liu, O. Zdraveva, D. Nichols, and K. Osypov, 2010, Application of steering filters to localized anisotropic tomography with well data: SEG Expanded Abstracts, **29**, 4286–4290.

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

Cai, J., Y. He, Z. Li, B. Wang, and M. Guo, 2009, TTI/VTI anisotropy parameters estimation by focusing analysis, Part I: theory: SEG Expanded Abstracts, **28**, 301–305.

Clapp, R., 2000, Geologically constrained migration velocity analysis: PhD thesis, Stanford University.

Fei, T. W. and C. L. Liner, 2008, Hybrid Fourier finite-difference 3D depth migration for anisotropic media: Geophysics, **73**, 27–34.

Fei, W. and P. Williamson, 2010, On the gradient artifacts in migration velocity analysis based on differential semblance optimization: SEG Expanded Abstracts, **29**, 4071–4076.

Fletcher, R., X. Du, and P. J. Fowler, 2009, Stabilizing acoustic reverse-time migration in TTI media: SEG Expanded Abstracts, **28**, 2985–2989.

Guerra, C. and B. Biondi, 2010, Fast 3D velocity updates using the pre-stack exploding reflector model: SEG Expanded Abstract, **29**, 4375–4379.

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

Hale, D., 2007, Local dip filtering with directional Laplacians: CWP Report 567.

Helbig, K., 1956, Die ausbreitung elastischer Wellen in anisotropen Medien: Geophys. Prosp., **04**, 70–81.

Li, Y., D. Nichols, K. Osypov, and R. Bachrach, 2011a, Anisotropic tomography using rock physics contraints: Accepted by the proceeding EAGE.

———, 2011b, Anisotropic tomography with rock physics constraints: SEP-Report, **143**, 249–258.

McCollum, B. and F. Snell, 1932, Asymmetry of sound velocity in stratified formations: Physics (Journal of Applied Physics), **2**, 174–185.

Plessix, R.-E. and H. Rynja, 2010, VTI full waveform inversion: a parameterization study with a narrow azimuth streamer data example: SEG Expanded Abstracts, **29**, 962–966.

Postma, G. W., 1955, Wave propagation in a stratified medium: Geophysics, **20**, 780–806.

Sava, P., 2004, Migration and velocity analysis by wavefield extrapolation: PhD thesis, Stanford University.

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.

Shan, G., 2009, Optimized implicit finite-difference and Fourier finite-difference migration for VTI media: Geophysics, WCA189–WCA197.

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.

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

Tang, Y. and B. Biondi, 2010, Target-oriented wavefield tomography using demigrated Born data: SEG Expanded Abstract, **29**, 4280–4285.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Thomsen, L., 1986, Weak elastic anisotropy: Geophysics, **51**, 1954–1966.

Tsvankin, I. and L. Thomsen, 1994, Nonhyperbolic reflection moveout in anisotropic media: Geophysics, **59**, 1290–1304.

Vyas, M. and Y. Tang, 2010, Gradients for wave-equation migration velocity analysis: SEG Expanded Abstracts, **29**, 4077–4081.

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

Woodward, M. J., D. Nichols, O. Zdraveva, P. Whitfield, and T. Johns, 2008, A decade of tomography: Geophysics, **73**, VE5–VE11.

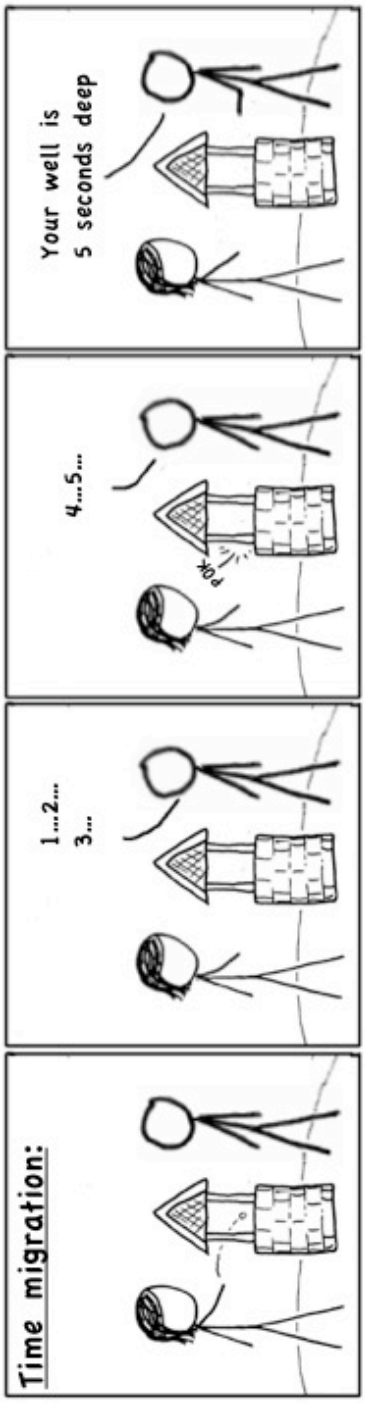
Yuan, J., X. Ma, S. Lin, and D. Lowrey, 2006, P-wave tomographic velocity updating in 3D inhomogeneous VTI media: SEG Expanded Abstracts, **25**, 3368–3372.

Zhang, Y. and B. Biondi, 2011, Moveout-based wave-equation migration velocity analysis: SEP-Report, **143**, 43–58.

Zhang, Y. and H. Zhang, 2009, A stable TTI reverse time migration and its implementation: SEG Expanded Abstracts, **28**, 2794–2798.

Zhou, H., D. Pham, and S. Gray, 2004, Tomographic velocity analysis in strongly anisotropic TTI media: SEG Expanded Abstracts, **23**, 2347–2350.

Zhou, H., D. Pham, S. Gray, and B. Wang, 2003, 3-D tomographic velocity analysis in transversely isotropic media: SEG Expanded Abstracts, **22**, 650–653.


Time migration:
1...2...
3...
4...5...
POK
Your well is
5 seconds deep

Correlation-based wave-equation migration velocity analysis

Ali Almomin

ABSTRACT

Wave-equation migration velocity analysis (WEMVA) is a family of techniques that aim to improve the subsurface velocity model by minimizing the residual in the image space. Since the true image is unknown, measuring the residual in the image space is a challenge for WEMVA techniques. In this paper, I present a new method of measuring the image perturbation that is based on the cross-correlation of the observed image with a reference image in reflection angle gathers. I derive the gradient of this technique and show some synthetic examples that compare it to the optimum WEMVA gradient. I then modify the gradient in order to handle multiple events and show that it becomes immune to the problem of cycle skipping. I finally show a synthetic example of the modified gradient and compare it to the optimum gradient.

INTRODUCTION

Seismic velocity-analysis methods can be divided into two major groups. First, there are techniques that aim to minimize the misfit in the data domain, such as full waveform inversion (Tarantola, 1984; Luo and Schuster, 1991). Second, there are other techniques that aim to improve the quality in the image domain such as migration velocity analysis (MVA)(Symes and Carazzone, 1991; Biondi and Sava, 1999; Shen, 2004). These techniques try to measure the quality of the image and then invert the estimated image perturbation using a linearized wave-equation operator.

There are several advantages to minimizing the residual in the image space, such as increasing the signal-to-noise ratio and decreasing the complexity of the data (Tang et al., 2008). However, the biggest challenge in WEMVA techniques is that the true image is unknown. Therefore, each technique uses a certain attribute of the background image and tries to estimate the residual using that attribute. The stack-power-maximization technique maximizes the angle stack, and differential semblance optimization (DSO) minimizes the difference between neighboring traces in angle gathers. However, these assumptions can cause some problems, such as cycle-skipping in stack power maximization and edge effects in DSO.

In this paper, I present a new method of measuring the image residual that is based on the cross-correlation of the observed image with a reference image in reflection angle gathers. The reference image can be any image with the desired kinematics, i.e. flat angle gathers. Therefore, it is possible to choose an angle stack as the reference image. However, angle stacks do not take into account the limited acquisition, which can result in anomalies in the gradient if the angle gathers are not muted properly. In order to take acquision into account, I create my reference image by computing Born-modeling data with the background slowness and a reference reflectivity. This reference reflectivity could come

from the angle-stack image or from non-seismic data such as well-logs or geologic models. Therefore, modeling and migrating a dataset gives us more flexibility than just using the angle stack. The derivation of this method is based on traveltime inversion by Luo and Schuster (1991) but in the image domain instead of the data domain. After deriving the objective function and the gradient of this method, I provide some synthetic examples and compare the gradient to the optimum WEMVA gradient.

This technique is similar to differential residual migration (DRM) (Sava, 2004) in the sense that it uses the kinematics of a reference image. However, there are a few advantages in using correlation over DRM. First, the correlation method gives us more flexibility in choosing the reference image. Second, picking correlation lags could be automated more easily than picking DRM panels. Finally, the objective function of the correlation method could include the full correlation function as opposed to just maximum lags, which will eliminate picking and fully automate the inversion.

METHOD

The first step in evaluating a tomographic operator is to linearize the image $\mathbf{I}$ around the background slowness $\mathbf{s}_0$, as follows:

$$\mathbf{I} = \mathbf{I}_0 + \frac{\partial \mathbf{I}}{\partial \mathbf{s}}|_{s_0}(\mathbf{s} - \mathbf{s}_0) + ..., \tag{1}$$

where $\mathbf{I}_0$ is the background image and $\mathbf{s}$ is the slowness model. By neglecting the higher order terms in the image series, we can define the tomographic operator as follows:

$$\Delta \mathbf{I} = \frac{\partial \mathbf{I}}{\partial \mathbf{s}}|_{s_0}\Delta \mathbf{s} = \mathbf{T}\Delta \mathbf{s}, \tag{2}$$

where $\mathbf{T}$ is the tomographic operator. Now, we use the conventional imaging condition as follows:

$$I(\mathbf{x}, \mathbf{h}) = \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r} G^*(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) G^*(\mathbf{x} + \mathbf{h}, \mathbf{x}_r, \omega) d(\mathbf{x}_r, \mathbf{x}_s, \omega), \tag{3}$$

where I is the image, G is the Green's function, d is the surface data, ω is frequency, $\mathbf{x}_s$ and $\mathbf{x}_s$ are the source and receiver coordinates, and $\mathbf{h}$ is the subsurface offset. To evaluate the tomographic operator $\mathbf{T}$, I take the derivative of the imaging condition as follows:

$$\begin{aligned}
\Delta I(\mathbf{x}, \mathbf{h}) &= \sum_{\mathbf{y}} \frac{\partial I(\mathbf{x}, \mathbf{h})}{\partial s(\mathbf{y})} \Delta s(\mathbf{y}) \\
&= \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r, \mathbf{y}} \left\{ -2\omega^2 s_0(\mathbf{y}) G_0^*(\mathbf{y}, \mathbf{x}_s, \omega) G_0^*(\mathbf{x} - \mathbf{h}, \mathbf{y}, \omega) \right\} G_0^*(\mathbf{x} + \mathbf{h}, \mathbf{x}_r, \omega) d(\mathbf{x}_r, \mathbf{x}_s, \omega) \Delta s(\mathbf{y}) \\
&+ \sum_{\omega, \mathbf{x}_s, \mathbf{x}_r, \mathbf{y}} \left\{ -2\omega^2 s_0(\mathbf{y}) G_0^*(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) G_0^*(\mathbf{x} + \mathbf{h}, \mathbf{y}, \omega) \right\} G_0^*(\mathbf{y}, \mathbf{x}_r, \omega) d(\mathbf{x}_r, \mathbf{x}_s, \omega) \Delta s(\mathbf{y}),
\end{aligned} \tag{4}$$

where y is the slowness coordinate. The full derivation of the tomographic operator is presented in Almomin and Tang (2010).

After defining the tomographic operator, I use a cross-correlation function to estimate image perturbations:

$$f(\zeta, \gamma; \mathbf{x}) = \sum_z I_{\text{obs}}(z, \gamma; \mathbf{x}) I_{\text{cal}}(z + \zeta, \gamma; \mathbf{x}), \tag{5}$$

where ζ is the lag, γ is the reflection angle, $\mathbf{x}$ is the surface coordinates, z is depth, I_{obs} is the angle-domain image using the observed data, and I_{cal} is the angle-domain image using the calculated data, which is modeled with the background slowness. I_{cal} is always going to be flat, since I create Born-modeled data using a reference model as the reflectivity and the background slowness and then migrate that data using the same background slowness. Next, I define ξ to be the lag that maximizes the correlation function. Therefore, the derivative of the correlation function vanishes at that lag, as follows:

$$\mathbf{g} = \frac{\partial \mathbf{f}}{\partial \zeta}|_{\xi} = \sum_z \frac{\partial}{\partial z} I_{\text{obs}}(z, \gamma; \mathbf{x}) I_{\text{cal}}(z + \xi, \gamma; \mathbf{x}) = 0. \tag{6}$$

We can now use ξ as our measure of the residual to minimize, casted as the following objective function:

$$J(\mathbf{s}) = \frac{1}{2} \|\xi(\gamma, \mathbf{x})\|^2. \tag{7}$$

Then, we take the derivative of the objective function with respect to slowness as follows:

$$\nabla J = \left(\frac{\partial \xi}{\partial \mathbf{s}}\right)^* \xi, \tag{8}$$

where * indicates an adjoint. By using the chain rule of differentiation, I relate the derivative of the maximum lag ξ with respect to $\mathbf{s}$ to the derivative of the correlation function with respect to $\mathbf{s}$ as follows:

$$\left(\frac{\partial \xi}{\partial \mathbf{s}}\right)^* = \left(\frac{\partial \mathbf{g}}{\partial \mathbf{s}} \frac{\partial \xi}{\partial \mathbf{g}}\right)^*. \tag{9}$$

The second partial derivative in equation (9) is just a scalar that balances the energy between surface locations. The first partial derivative with respect to slowness can be calculated using equation (6) as follows:

$$\begin{aligned} \left(\frac{\partial \mathbf{g}}{\partial \mathbf{s}}\right)^* = & \sum_z \left(\frac{\partial I_{\text{obs}}(z, \gamma; \mathbf{x})}{\partial \mathbf{s}}\right)^* \frac{\partial}{\partial z} I_{\text{cal}}(z + \xi, \gamma; \mathbf{x}) \\ & + \sum_z \left(\frac{\partial I_{\text{cal}}(z + \xi, \gamma; \mathbf{x})}{\partial \mathbf{s}}\right)^* \frac{\partial}{\partial z} I_{\text{obs}}(z, \gamma; \mathbf{x}). \end{aligned} \tag{10}$$

The first tomographic operator in equation (10) can be computed as I described in equation (4). However, the second tomographic operator depends on how I_{cal} is computed. If a fixed-reflectivity model is used, such as well data, and only the background slowness is updated, then this derivative will be very small and could be ignored, since changing the slowness updates does not change the reflectivity estimate. On the other hand, if we allow the modeling reflectivity to change location, i.e. we update the reflectivity model as we iterate, then this operator could have a significant component. However, evaluating this operator is much more expensive than the first tomographic operator, since it is a cascade of three operators. Therefore, I will assume that the first tomographic operator is sufficient and ignore the second term.

SYNTHETIC EXAMPLES

The background model has a constant velocity of 2500 m/s. The model is 4000 m wide and 2800 m deep. The spatial sampling is 20 m, and the temporal sampling is 3 ms. A

Ricker wavelet with a fundamental frequency of 15 Hz is used to model the data. The receiver spacing is 20 m, and the shot spacing is 80 m. The reflector is at a depth of 2200 m. Born modeling was used in both the observed and the calculated data. There are three anomalies that I will estimate: first, a negative Gaussian anomaly at a depth of 1300 m with a maximum velocity of 800 m/s, second, a constant velocity decrease of 250 m/s, and third, a constant velocity increase of 250 m/s. First, I will compute the optimum WEMVA gradient by applying the forward and adjoint of these slowness perturbations. The results are shown in figures 1(a), 1(c) and 1(e). Then, I compute the angle-domain common-image gathers using the background velocity of the data with the three anomalies, which are shown in figures 1(b), 1(d) and 1(f).

Next I show the results of applying the first term of the tomographic operator in equation (10). First, I use the true reflector depth to create the reference image and use it in our method to generate the gradients for the three anomalies, as shown in figures 2(a), 2(c) and 2(e). Finally, I use the apparent reflector depth, i.e. the depth extracted from the zero-subsurface-offset image, to create the reference image and generate the gradients shown in figures 2(b), 2(d) and 2(f). The correlation lags were picked automatically by choosing the maximum value of the cross-correlation function.

Modified correlation-based WEMVA

There were two disadvantages to the original correlation-based WEMVA gradient. First, it assumes a single correlation lag per trace. Second, it is sensitive to the actual values of the picked correlation lags. If these lags are not correct, the gradient might have a cycle-skipping problem. In order to overcome these disadvantages, I approximate the shifted flat image by the observed image to get the following modified gradient:

$$\nabla J = \sum_z \left(\frac{\partial I_{\text{obs}}(z,\gamma;\mathbf{x})}{\partial \mathbf{s}} \right)^* \frac{\partial}{\partial z} I_{\text{obs}}(z,\gamma;\mathbf{x}) \frac{\xi}{E} \tag{11}$$

This modified gradient is immune to cycle-skipping since the residual and operator match by design. In addition, it is not as sensitive to the values of the picked lags. Finally, multiple picked lags could be used in this formulation.

Figures 3(a) and 3(b) show the true and background velocities of a salt model. ADCIGs using these two velocities on the observed data are shown in figures 4(a) and 4(b). Next, I used a sliding Gaussian window to compute and pick local cross-correlation panels between the observed image and the stacked image. Then, I picked the maximum correlation lag at each depth level. The picked lags are shown in figure 5.

As in the previous example, I computed the optimum gradient using the true velocity anomaly and the modified correlation-based WEMVA gradient using the picked correlation lags. These two gradients are shown in figure 6(a) and 6(b). Although we used the apparent depth of the stacked image, the modified gradient shows very good results that are very similar to the optimum gradient. In this case, the velocity information in large reflection angles was more dominant than those in zero reflection angle.

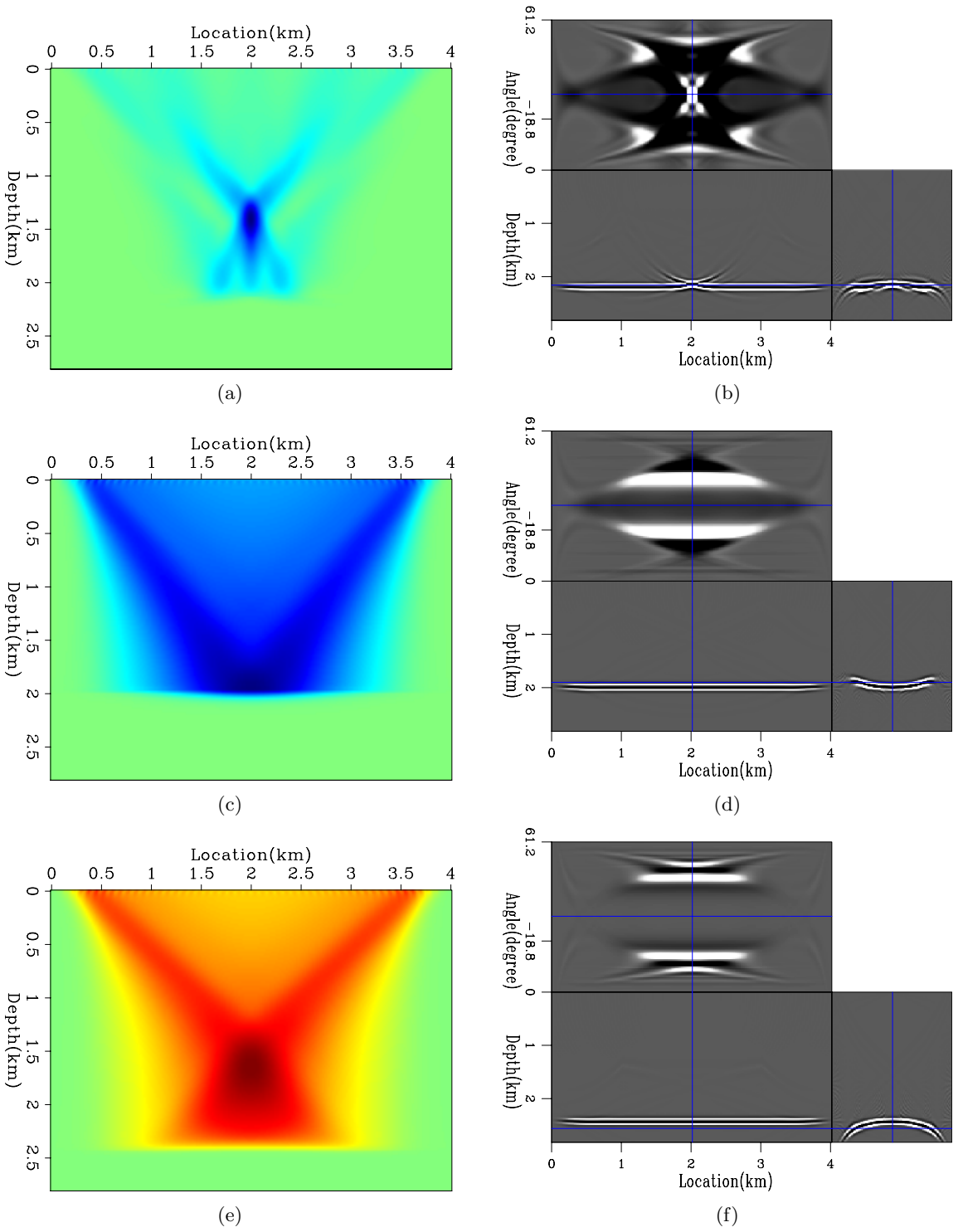


Figure 1: The left column shows the optimum WEMVA gradient of (a) a negative Gaussian anomaly, (c) a negative bulk shift and (e) a positive bulk shift. The right column shows the corresponding ADCIGs. [**CR**] ali1/. dS1,adcig1,dS2,adcig2,dS3,adcig3

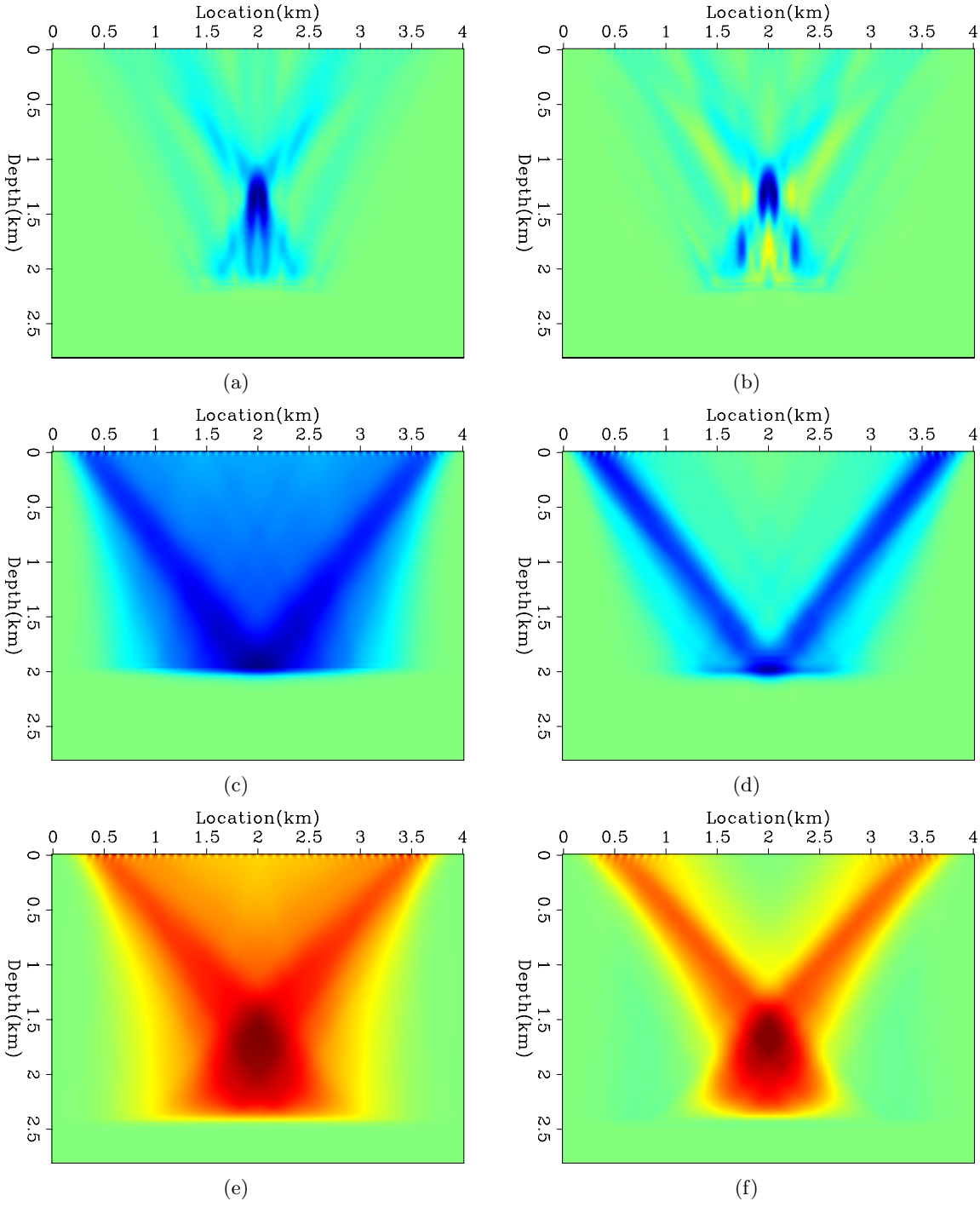
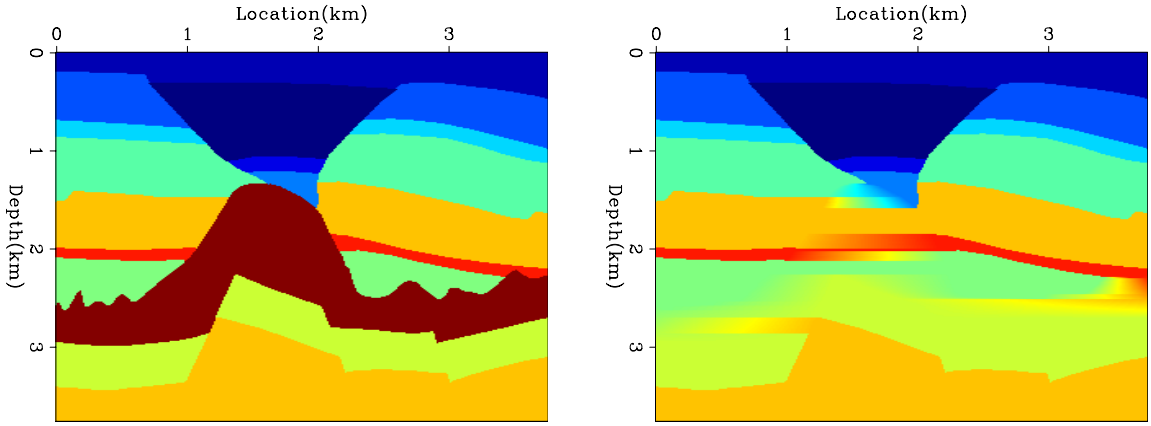
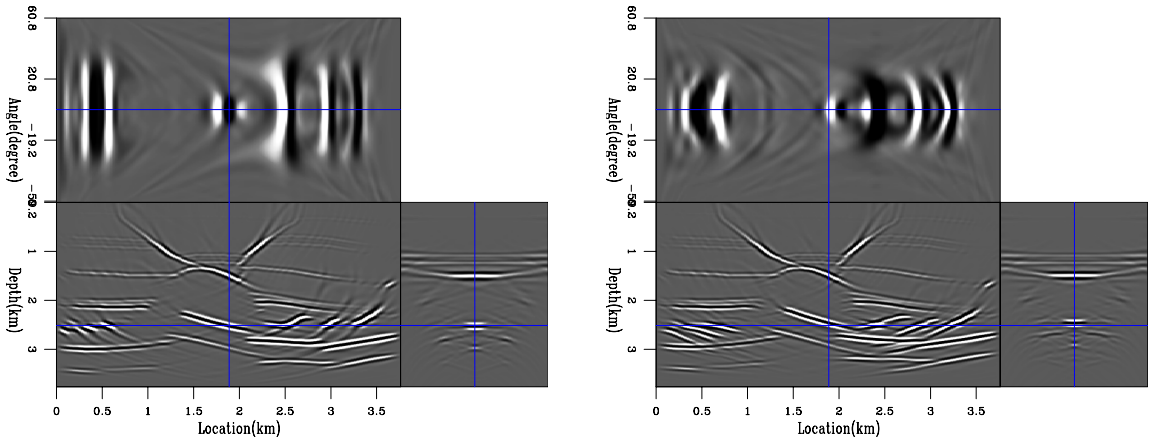


Figure 2: The left column shows the correlation-based WEMVA gradients using the true depth for the reference image. The right column shows the correlation-based WEMVA gradients using the apparent depth for the reference image. [**CR**] ali1/. dS221,dS321,dS222,dS322,dS223,dS323



(a) (b)

Figure 3: The true velocity salt model (a) and the background velocity model (b). [**ER**] ali1/. wemva22truev,wemva22bgv



(a) (b)

Figure 4: ADCIGs using the true velocity salt model (a) and using the background salt model (b). [**CR**] ali1/. wemva22adcigtv,wemva22adcigbgv

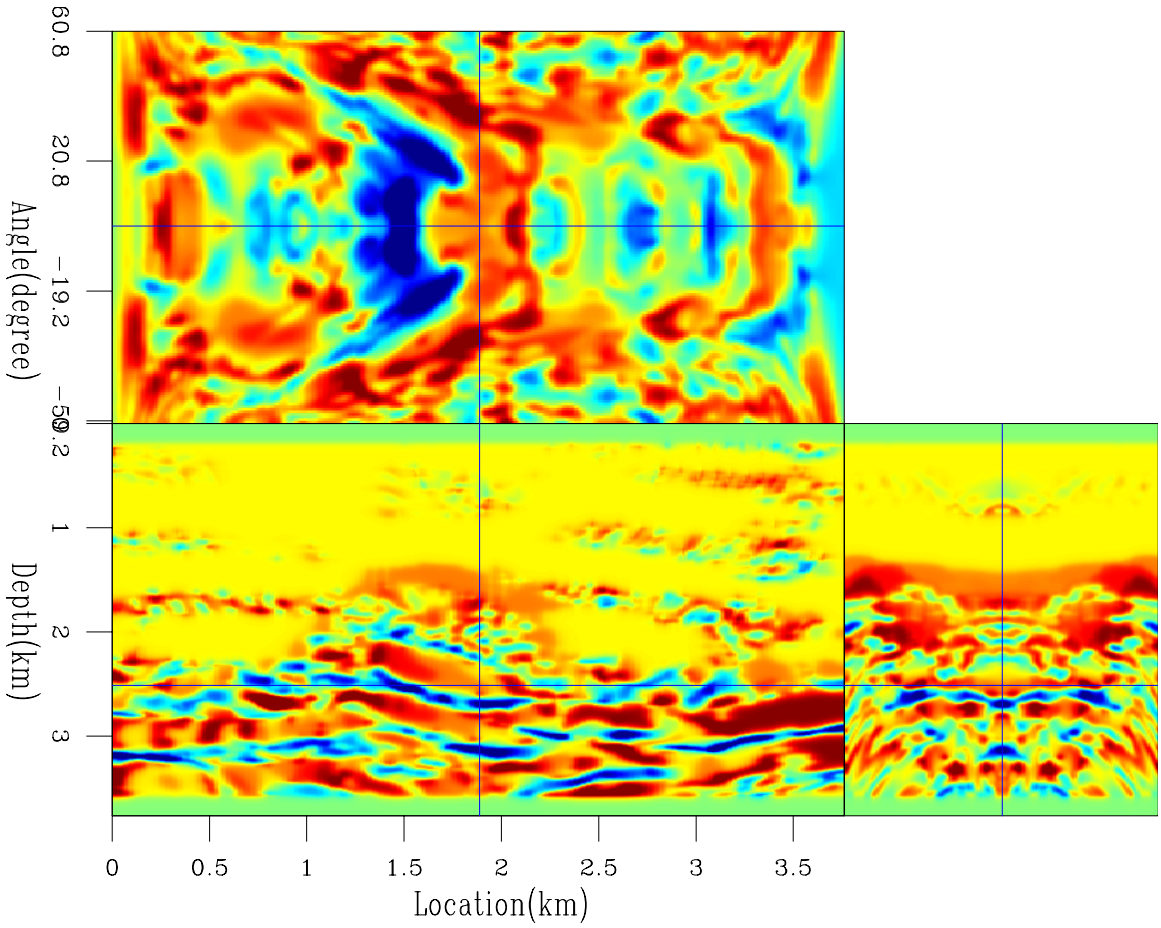


Figure 5: Lag estimation of the maximum local correlation. [**CR**] ali1/. wemva22lags3dc

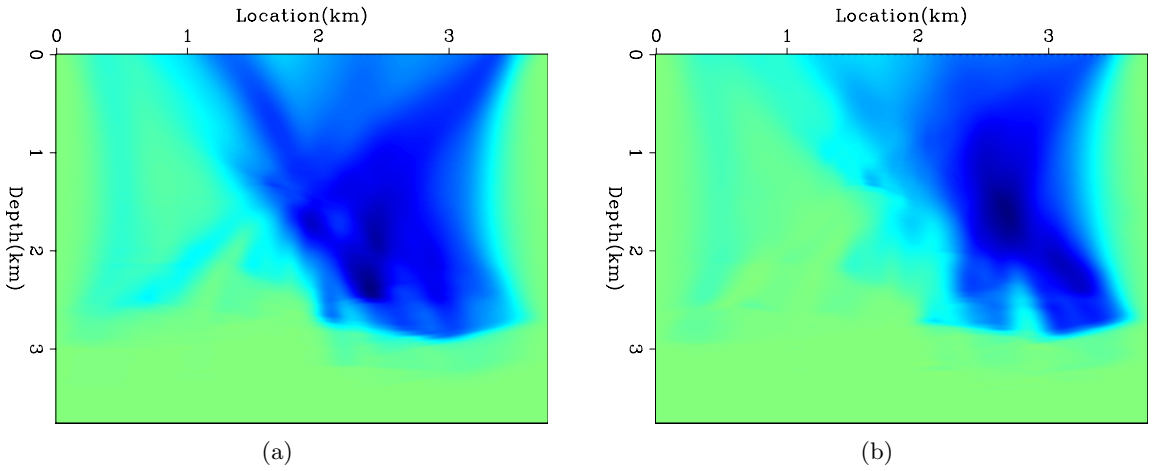


Figure 6: The optimum WEMVA gradient of the salt model (a) and the modified correlation-based WEMVA gradient of the salt model (b). [**CR**] ali1/. wemva22amacoopt,wemva22amacoz

DISCUSSION AND CONCLUSIONS

By examining the results, we see that, when provided with the true depth, the correlation-based WEMVA produces excellent results with no cycle-skipping or edge effects. When the apparent depth was provided, the results start to have some errors. However, these errors are to be expected due to the velocity-depth ambiguity. Therefore, the gradients will flatten the gathers but not necessarily resolve the correct depth (Zhang and Biondi, 2011). However, the results are still satisfactory, since the gradient is pointing toward the right direction.

Using the modified gradient has several advantages, such as handling multiple events. However, I need to further test the approximations made in the modified gradient. Leeuwen (2010) showed that it is possible to further improve the objective function by multiplying with the proper weights and minimizing the correlation function instead of the lags, which will eliminate any picking. Finally, using correlation functions over extended images (Yang and Sava, 2009) could help provide better results.

REFERENCES

Almomin, A. and Y. Tang, 2010, Migration velocity analysis based on linearization of the two-way wave equation: SEP-Report, **142**, 13–24.

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.

Leeuwen, T. V., 2010, Correlation-based seismic velocity inversion: PhD thesis, Delft University of Technology.

Luo, Y. and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

Sava, P., 2004, Migration and velocity analysis by wavefield extrapolation: PhD thesis, Stanford University.

Shen, P., 2004, Wave equation migration velocity analysis by differential semblance optimization: PhD thesis, Rice University.

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

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

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Yang, T. and P. Sava, 2009, Wave-equation migration velocity analysis using extended images: SEG Expanded Abstracts, **28**, 3715–3719.

Zhang, Y. and B. Biondi, 2011, Moveout-based wave-equation migration velocity analysis: SEP-Report, **143**, 43–58.

SEP 2011 ...
We have as many sunny days in England as you do rainy days in Saudi
I feel the inverse of your pain
Perhaps we should convolve together...

Moveout-based wave-equation migration velocity analysis

Yang Zhang and Biondo Biondi

ABSTRACT

We propose a new method to perform wave-equation migration velocity analysis using angle-domain common image gathers. Instead of maximizing the image-stack-power objective function directly with respect to the slowness, we link the objective function to the slowness indirectly through an intermediate moveout parameter. Since this approach is robust against the cycle-skipping problem, it produces more reasonable gradients. Also the proposed method does not require explicit picking of the moveout parameters. Our data examples shows the great potential of this method.

INTRODUCTION

Wave-equation migration velocity analysis (WEMVA) methods aim to utilize velocity information from the migrated images to improve the velocity model. Using the wave equation is potentially more accurate than ray-based methods becauses it better describes wave-propagation physics and will give a more physically realistic senstivity kernel for the velocity update. Evaluating the flatness of the angle-dommain common image gathers (ADCIGs) (Biondi and Symes, 2004) is so far the most favored choice when forming WEMVA optimization problems.

Several WEMVA methods have been proposed, but there is no consensus on the best method. The maximum-stack power method (Chavent and Jacewitz, 1995) directly maximizes the angle stack of the ADCIGs, but similar to the Full Waveform Inversion (FWI) (Tarantola, 1984) method, it is prone to cycle-skipping when the velocity error is too large. The differential-semblance optimization (DSO) (Symes and Carazzone, 1991; Shen et al., 2005; Shen and Symes, 2008) penalizes the first derivative along the angle axis on the ADCIGs. This objective function is easy to implement but will falsely over-penalize an already flat angle gather with variant amplitudes; and the differential operator significantly amplifies the noise in the image, thus generating unwanted artifacts in the velocity upgrade. Sava (2004) uses prestack Stolt residual migration to help construct the image perturbation. The cycle-skipping problem is avoided this way, however the user is required to pick a ρ parameter at each model point, and the picking is not trivial. Furthermore the Stolt migration can only migrate images using constant velocity for the entire velocity model. There is some question whether the ρ parameters picked with these image-ρ cubes always represent the correct trend of the velocity update.

In this report, we propose a new method which extends from the theory in Biondi (2010, 2011). This method extracts the velocity-focusing information in the angle domain and tries to maximize the angle-stack power of ADCIGs as well. To tackle the cycle-skipping issue, we present a new way to construct the image perturbation by introducing an intermediate moveout parameter that describes the kinematics change of the ADCIGs caused by the

velocity change; and that kinematic change then links to the change of objective function. The rest of the paper is divided into two parts: first the theoretical framework is explained; then we demonstrate the effectiveness of our method with several synthetic examples.

THEORY

Practical experience indicates that slowness is a better parameterization than velocity. For clarity, the model space is parameterized as slowness for the rest of the derivation in this paper. To make it simple, the derivation is presented in 2D; however it is straightforward to extend it to three dimensions.

from slowness perturbation to the change in image kinematics

In this section we present the formula that links the slowness perturbation to the shifts of the ADCIGs. Starting from the initial slowness model $s_0(z,x)$, we first define the pre-stack common-image gather in the angle domain as $I(z,\gamma,x;s_0)$, where γ is the reflection angle. If we choose a different slowness $s(z,x)$, the new image $I(z,\gamma,x;s)$ will be different from $I(z,\gamma,x;s_0)$ in terms of both kinematics and amplitude. If, as is commonly done, we focus on the kinematic change, then a way to characterize this kinematic change is to define a shift parameter b at each image location, $b(\gamma,z,x)$, such that if we apply this shift parameter to the initial image, the resulting image $I(z+b,\gamma,x;s_0)$ will agree with the new image $I(z,\gamma,x;s)$ in terms of kinematics. This is indicated by the maximum point of the auxilary objective function:

$$J_{aux}(b) = \int_{-L/2}^{L/2} dz_w \int d\gamma\, I(z+z_w+b,\gamma;z,x,s_0) I(z+z_w,\gamma;z,x,s) \text{ for each x,z.} \quad (1)$$

Note that in order to handle multiple events, we use a local window of length L along the depth axis. For the rest of the paper, the integration bound for variable z_w is always $[-L/2, L/2]$, and each J_{aux} is defined within that window around image point (x,z). $I(z+z_w,\gamma;z,x) = I(z+z_w,\gamma,x)$ represents a windowed version of the entire image.

This methodology is borrowed from Luo and Schuster (1991) who tried to find the relation of the travel-time perturbation to the slowness change. Then $\frac{\partial b}{\partial s}$ can be found using the rule of partial derivatvies for implicit functions (please refer to the appendix A):

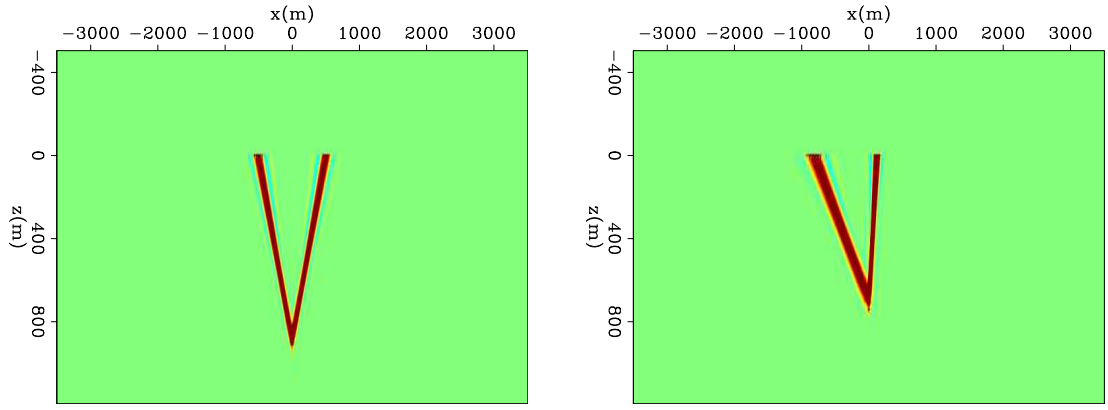
$$\frac{\partial b}{\partial s} = -\frac{\int dz_w\, \dot{I}(z+z_w+b,\gamma;z,x,s_0)\frac{\partial I(z+z_w,\gamma;z,x,s)}{\partial s}}{E(\gamma,z,x)}, \quad (2)$$

in which $E(\gamma,z,x) = \int dz_w \ddot{I}(z+z_w+b,\gamma;z,x,s_0) I(z+z_w,\gamma;x,s)$. $\dot{I}$ and $\ddot{I}$ indicate the first and second derivatives in z (depth). In practice, eq. (2) will be greatly simplified if we evaluate this expression at $s=s_0$, in other words $b=0$. In fact, this will always be the case if we update the intial slowness s_0 after each iteration. The simplified relation becomes

$$\frac{\partial b}{\partial s}|_{s=s_0} = -\frac{\int dz_w\, \dot{I}(z+z_w,\gamma;z,x,s_0)\frac{\partial I(z+z_w,\gamma;z,x,s)}{\partial s}}{E(\gamma,z,x)}, \quad (3)$$

and the $\frac{\partial I(z+z_w,\gamma;z,x,s)}{\partial s}$ term is indeed the wave-equation image-space tomographic operator. Each part in eq. (3) has clear physical implications: the E term acts as an energy term

to normalize the amplitude of the back-projected image; the back-projected image, $\dot{I}(z + z_w, \gamma; z, x, s_0)$ is built based on the initial image; it also has a first-order z derivative that introduces a proper 90° phase shift, ensuring a well behaved slowness update from the tomographic operator.



(a) (b)

Figure 1: Slowness sensitivity kernel at incident angle $\gamma = 30°$ for a flat reflector (a) and a dipping reflector (b). [**ER**] yang1/. sensKer1w.zref,sensKer1w.Tilt.zref

For a simple illustration of eq. (3), the slowness sensitivity kernel is calculated, by back-projecting a shift perturbation $\Delta b(\gamma, x)$ that has one single spike at $\gamma = 30°, x = 0$. A uniform background velocity of 2000 m/s is used. Figure 1(a) shows the sensitivity kernel if the reflector is flat, and figure 1(b) shows the sensitivity kernel with a dipping reflector (dip angle = 20°). As is clearly shown in these two plots, this operator will project the slowness perturbation along the corresponding wave path based on the location and reflection angle of the image shift.

Introducing the residual moveout parameter

The previous section showed how the slowness change can be associated with the kinematic change of the image. The remaining task is to determine how the image's kinematics should change by evaluating the objective function. Ideally we would like to know the true depth of the reflector, so that we can guarantee that the correct shift direction of the image is extracted. Unfortunately, we usually do not have the true depth of the reflector, and have to find the shift direction by relying only on the flatness criterion. A first attempt that directly maximizes the angle stack after shifting the initial image would be

$$\mathbf{J}(s) = \frac{1}{2} \sum_x \sum_z \int dz_w \left[\int d\gamma \, I(z + z_w + b, \gamma, z, x; s_0) \right]^2 . \tag{4}$$

The corresponding derivative over b for a fixed (γ, z, x) is

$$\frac{\partial J}{\partial b(\gamma, z, x)}|_{b=0} = \frac{1}{2} \sum_x \sum_z \int dz_w \Big\{ \left[\int d\gamma \, I(z + z_w, \gamma, z, x; s_0) \right] \dot{I}(z + z_w, \gamma, z, x; s_0) \Big\}. \tag{5}$$

Apparently, this objective function is very susceptible to the cycle-skipping problem: for a fixed (γ, z, x), if the image $I(z + z_w, \gamma, z, x; s_0)$ and the angle-averaged image $\int d\gamma \, I(z +$

$z_w, \gamma, z, x; s_0)$ become out of phase, the derivative of eq. (5) will point to the wrong shift direction.

To prevent cycle-skipping, we need a way to detect the global shape of the ADCIGs. Almomin (2011) propose to measure the relative shift the traces at each angle with respect to some reference trace by picking cross-correlation peak. Here we use the residual moveout (RMO) parameters so that the objective function knows whether the angle gather is curving up or curving down.

As shown in Biondi (2003) Chap 11, in the case of constant velocity error, the residual moveout of an ADCIG gather is

$$\theta(\gamma) = z_{\rho 0} \frac{\rho - 1}{\cos \alpha_x} \frac{\sin^2 \gamma}{\cos^2 \alpha_x - \sin^2 \gamma},$$

where $\rho = s/s0$, α_x and γ are the dip angle and reflection angle respectively, and z is the true reflector depth, $z_{\rho 0} = z/\rho$. If we assume the dip is small, then the expression above can be further simplified to

$$\theta(\gamma) = z_{\rho 0}(\rho - 1) \tan^2 \gamma.$$

Therefore we introduce the moveout parameter α and the moveout function $g(\gamma) = \tan^2 \gamma$. The objective function we want to maximize is the angle stack-power of the inital image after applying the residual moveout:

$$\mathbf{J} = \frac{1}{2} \sum_x \sum_z \int dz_w \left[\int d\gamma \, I(z + z_w + \alpha g(\gamma), \gamma, z, x; s_0) \right]^2.$$

The derivative is

$$\frac{\partial J}{\partial s}|_{s=s_0} = \sum_x \sum_z \int dz_w \left\{ \left(\int d\gamma I(z + z_w, \gamma, z, x; s_0) \right) \left[\int d\gamma \, \dot{I}(z + z_w, \gamma, z, x; s_0) g(\gamma) \frac{\partial \alpha}{\partial s} \right) \right] \right\}.$$

Define

$$\begin{aligned} A(z_w; z, x, s_0) = &\quad \int d\gamma \, I(z + z_w, \gamma; z, x, s_0), \\ B(z_w; z, x, s_0) = &\quad \int d\gamma \, \dot{I}(z + z_w, \gamma; z, x, s_0) g(\gamma), \end{aligned}$$

then

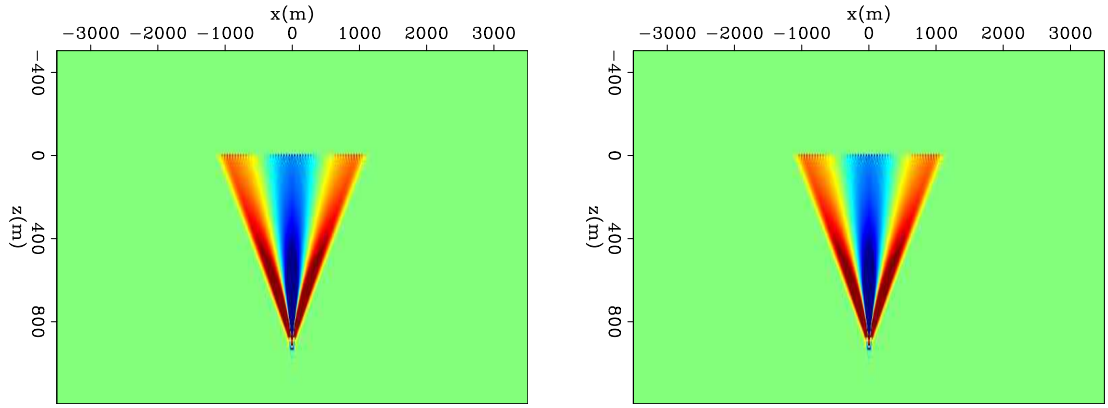
$$\begin{aligned} \frac{\partial J}{\partial s} &= \sum_x \sum_z \frac{\partial J}{\partial \alpha} \frac{\partial \alpha}{\partial s} \\ &= \sum_x \sum_z \{ \int dz_w A(z_w; z, x, s_0) B(z_w; z, x, s_0) \} \frac{\partial \alpha}{\partial s}. \end{aligned} \tag{6}$$

We find there are two ways to derive the $\frac{\partial \alpha}{\partial s}$ relation (see appendix B):

1. We can link Δs to $\Delta \alpha$ by defining an auxilary objective function; we call this the *direct operator*.

2. we can convert the perturbation of α to the shift parameter b perturbation at each angle, through a weighted least-squares fitting formula; thus $\Delta \alpha \rightarrow \Delta b \rightarrow \Delta s$; and as shown previously, we know how to calculate $\frac{\partial b}{\partial s}$. We call this the *indirect operator*.

The sensitivity kernel $\frac{\partial \alpha}{\partial s}$ calculated using the direct operator and the indirect operator are shown in figure 2, as with the Toldi operator (Toldi, 1985), the characteristic shape of such a sensitivity kernel is a center lobe, with two side lobes with oppositite polarity, which reaffirms the well known fact that velocity perturbation at the center and the side-end lateral position will change the curvature of ADCIGs towoard opposite directions. Yet the overall average is positive, which would give the correct update in case of a bulk shift slowness error.



(a) (b)

Figure 2: Sensitivity kernel $\partial\alpha/\partial s$ calulated using the direct operator (a) and the indirect operator (b), using constant background velocity and a flat reflector. [**CR**] yang1/. sensKer1w.bf2,sensKer1w.bf1m

Now if we review this method on eq. (6), the success of this method simply relies on the proper behavior of the two components in eq. (6): $\frac{\partial J}{\partial \alpha}$ needs to correctly detect the curvature of the ADCIGs so that the inversion will choose a moveout direction that properly flatten the gathers; $\frac{\partial \alpha}{\partial s}$ needs to properly convert the curvature perturbation to the update in slowness space.

In cases where velocity error is big, the actual curvature of the gather may be poorly represented the $\frac{\partial J}{\partial \alpha}$ term. To further improve the robustness and precondition the gradient, the analytic expression of $\frac{\partial J}{\partial \alpha}$ in eq. (6) is replaced by a numerical approach. First a semblance panel of $J(\alpha)$ will be calculated. To ensure that the derivative at $\alpha = 0$ can determine the correct curvature that maximizes the semblance value, a Gaussian derivative rather than a simple $(-1, 1)$ finite-difference derivative is applied. The width of the Gaussian can be reduced in later iterations.

RESULTS

We create several synthetic examples to show the effectiveness of our method. The model is 16 km in x and 1.2 km in z. The grid sampling is 20 m in x and 10 m in z. The survey geometry follows the marine acquisition convention with 4km cable length, receiver spacing is 20 m and a total of 150 shots are simulated from -6 km to +6 km with a spacing of 80 m. A total of 106 frequencies are calculated, ranging from 5 Hz to 40 Hz. Unless explicitly mentioned, there is one flat reflector at a depth around 800 m and a constant background

velocity of 2000 m/s is used. A one-way propagator is used for computational efficiency. For comparison, the subsurface-offset DSO and straight maximum-stack-power method are also implemented.

The first example is a true velocity with a 1 km width Gaussian anomaly at the center, with peak value 4000 m/s. Figure 3(a) shows the anomaly (in slowness) and Fig 3(b) shows the migrated zero subsurface-offset image, The velocity error is so large that the center part of the reflector is pulled up significantly.

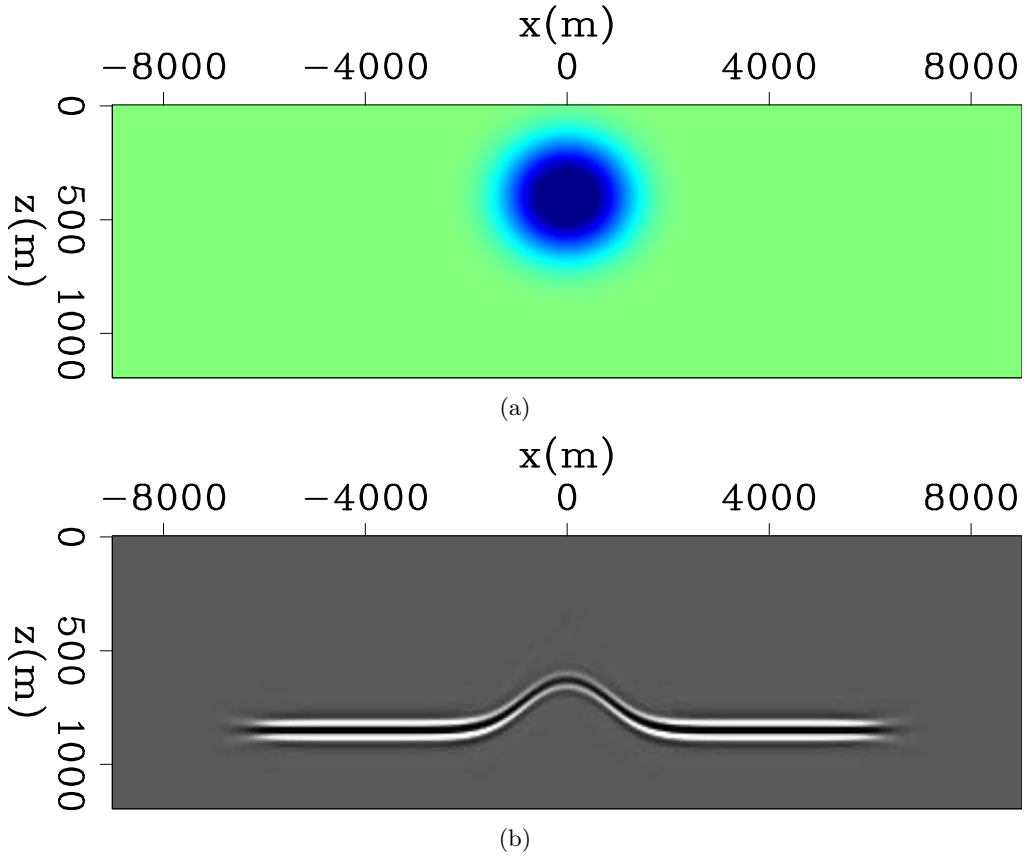


Figure 3: (a) The slowness model with a Gaussian anomaly. [**ER**](b) The migrated image using constant background slowness. [**CR**] yang1/. vel2.a5,img21w.a5.0off

Figure 4 shows the first slowness update using the three methods. Our method presents the most pleasing update among the three; the location of the anomaly is correctly located, and the target region is more uniformly updated, except for the small holes within the "W" shape due to the poorer wave path coverage caused by the curved reflector. The DSO result shows the typical strip-shaped artifacts, and the amplitude of the target region is weak compared to the edge artifacts. The maximum-stack-power method suffers from cycle skipping and is not able to locate the target area.

In the second example the true velocity has a horizontal gradient, where velocity increase linearly from 1500 m/s to 2700 m/s. Figure 5(a) shows the true model (in slowness) and figure 5(b) shows the migrated zero subsurface-offset image. The originally flat reflector is tilted in the image because of the horizontal slowness gradient.

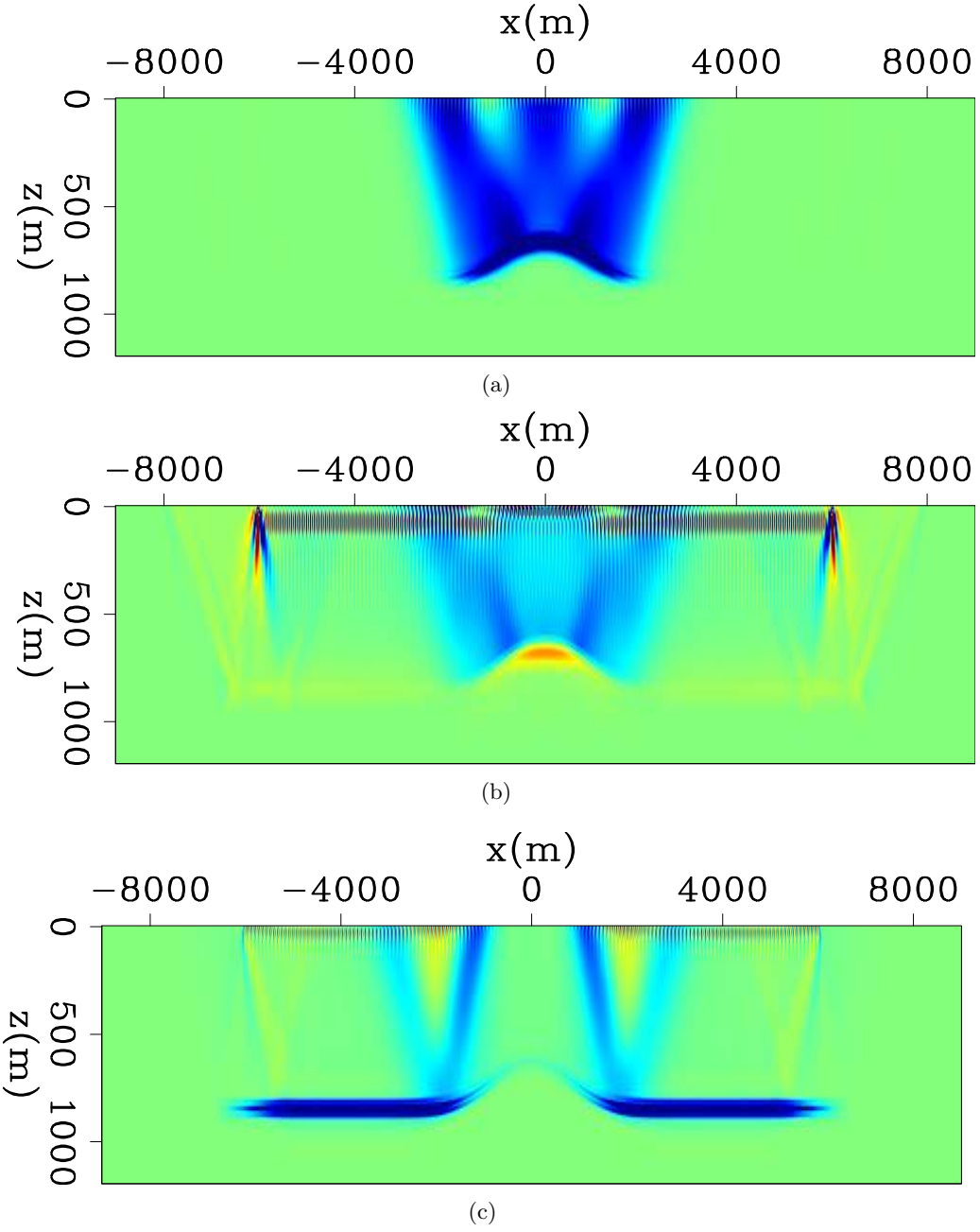


Figure 4: The first slowness update direction of our method (a), the subsurface-offset domain DSO (b), and the direct stack-power-maximization (c), true slowness refers to figure 3(a). [**CR**] yang1/. dS21w.a5.bf2g,dS21w.a5.dso,dS21w.a5.1

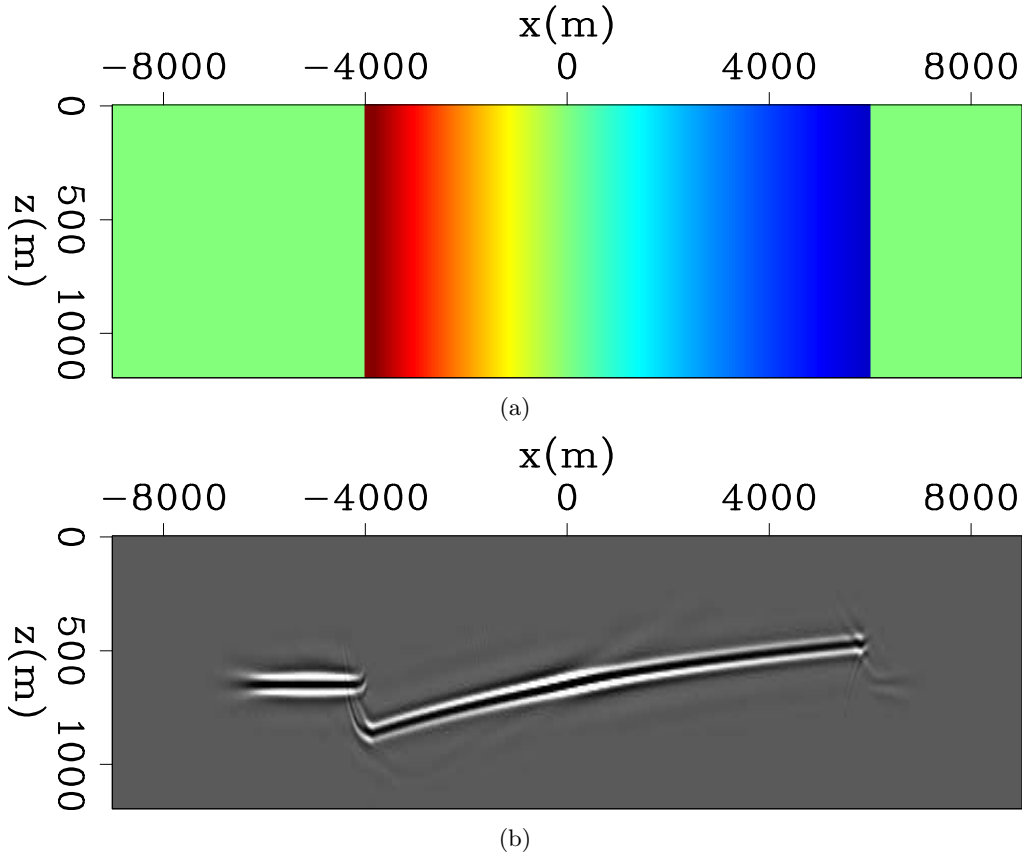


Figure 5: (a) The slowness model with a horizontal gradient from 1500m/s to 2700m/s [**ER**];(b) The migrated image using constant background slowness. [**CR**] yang1/. vel2.hg,img21w.hg.0off

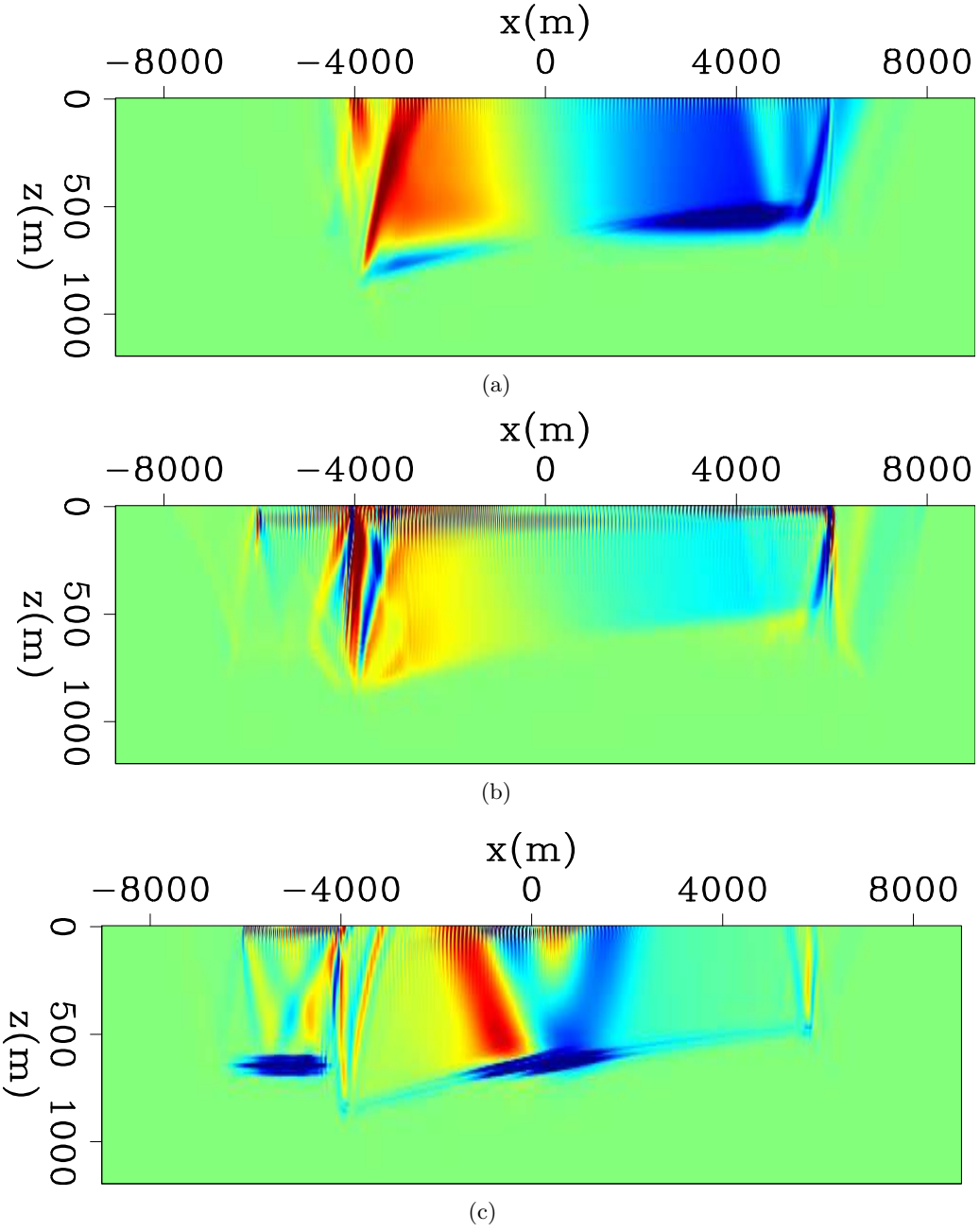


Figure 6: The first slowness update direction of our method (a), the subsurface-offset domain DSO (b), and direct stack power maximization (c). The true slowness refers to figure 5(a). [**CR**] yang1/. dS21w.hg.bf2g,dS21w.hg.dso,dS21w.hg.1

Figure 6 shows the first slowness update using the three methods. Again our method presents the most pleasing update among the three. The amplitude of the search direction is proportional to the actual magnitude of the slowness error. A large strip-shaped artifact is introduced at $x = -4000\,m$ in DSO update. this is commonly seen from DSO results when there are reflector discontinuities; although it does decently predicted the horizontal gradient shape of the slowness error. The maximum-stack-power method is able to predict correctly when the slowness error is small (around $x = 0\,m$). As the slowness error reaches a certain threshold, cycle-skipping happens and the result deteriorates.

For the third example, the true velocity is a constant $1900\,m/s$ plus a high veloicty anomaly at $x = 2000\,m/s$ with peak value $2850\,m/s$. We designed a hump-shaped reflector to test this method's ability to handle mild dips. Figure 7(a) shows the true model (in slowness) and Fig 7(b) shows the migrated zero subsurface-offset image. Figure 7(c) and 7(d) shows the first slowness update using our method and the DSO method. The effect of the high velocity anomaly is partially cancelled by the lower constant part ($1900\,m/s$), nontheless, the overall result still requires a negative update at the anomaly's location. In the presence of a mild dipping reflector, our method still yields a satisfying result.

The effect of the high velocity anomaly is partially cancelled by the lower constant part ($1900\,m/s$), nontheless, the overall result still requires a negative update at the anomaly's location. In the presence of a mild dipping reflector, our method still yields a satisfying result.

Based on the gradient calculation, we implement an non-linear slowness inversion framework. The example we tested is the marmousi velocity model. The model is $6\,km$ in x and $1.6\,km$ in z. The spatial sampling is $20\,m$. The survey geometry follows the land acquisition convention with receiver at every surface location on the top and a total of 51 shots are simulated covering the whole lateral span on the top with a spacing of $120\,m$.

Figure 8(a) shows the true model (in velocity) and figure 8(b) shows the starting model, which has a vertical gradient increasing from $1600\,m/s$ to $3200\,m/s$. Figure 8(c) and 8(d) show the inverted velocity model using our method and subsurface-offset DSO method. We run 20 nonlinear iterations for each inversion, and the same extent of gradient smoothing is applied. Although both inversion result improves the focusing of the subsurface-offset image, the outline of the true velocity structure is much better captured with our method.

CONCLUSION

We present a new method to perform wave-equation migration velocity analysis. It properly captures the kinematic update from the ADCIGs and projects that update into the slowness space. We demonstrated this promising approach using several examples. As shown previously, this method does not suffer from cycle-skipping, does not require moveout parameters picking, and can robustly improve the flatness of the angle gathers. We plan to further develop this method and apply it to more challenging data.

ACKNOWLEDGEMENT

We thank Robert Clapp for his insights and suggestions that help us solve quite a few practical issues during this project. The first author would like to thank Yaxun Tang

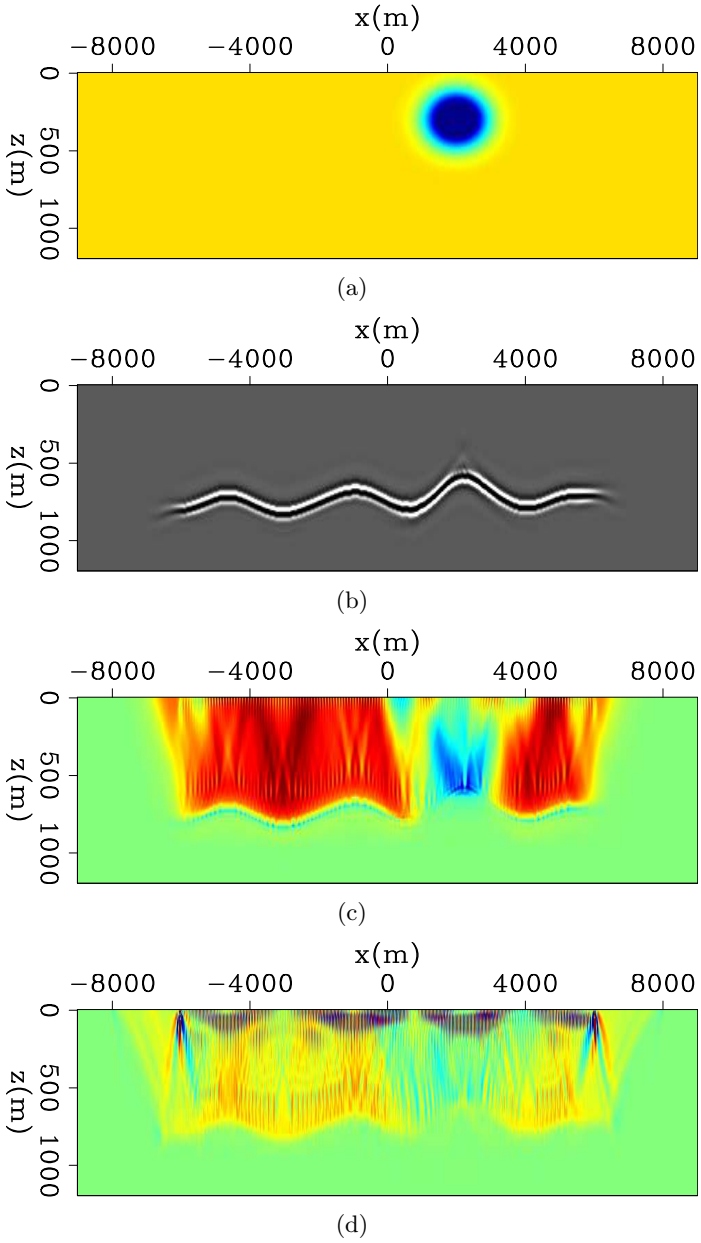


Figure 7: (a) The slowness model with a constant slower velocity $1900\,m/s$, there is also a high velocity anomaly at $x = 2000\,m$, with peak value $2850\,m/s$; [**ER**](b) The migrated image using constant background velocity ($2000\,m/s$). [**CR**](c) The first slowness update direction using our method; [**CR**](d) The first slowness update using the subsurface-offset domain DSO method. [**CR**] yang1/. vel2.bump,img21w.bump.0off,dS21w.bump.bf2g,dS21w.bump.dso

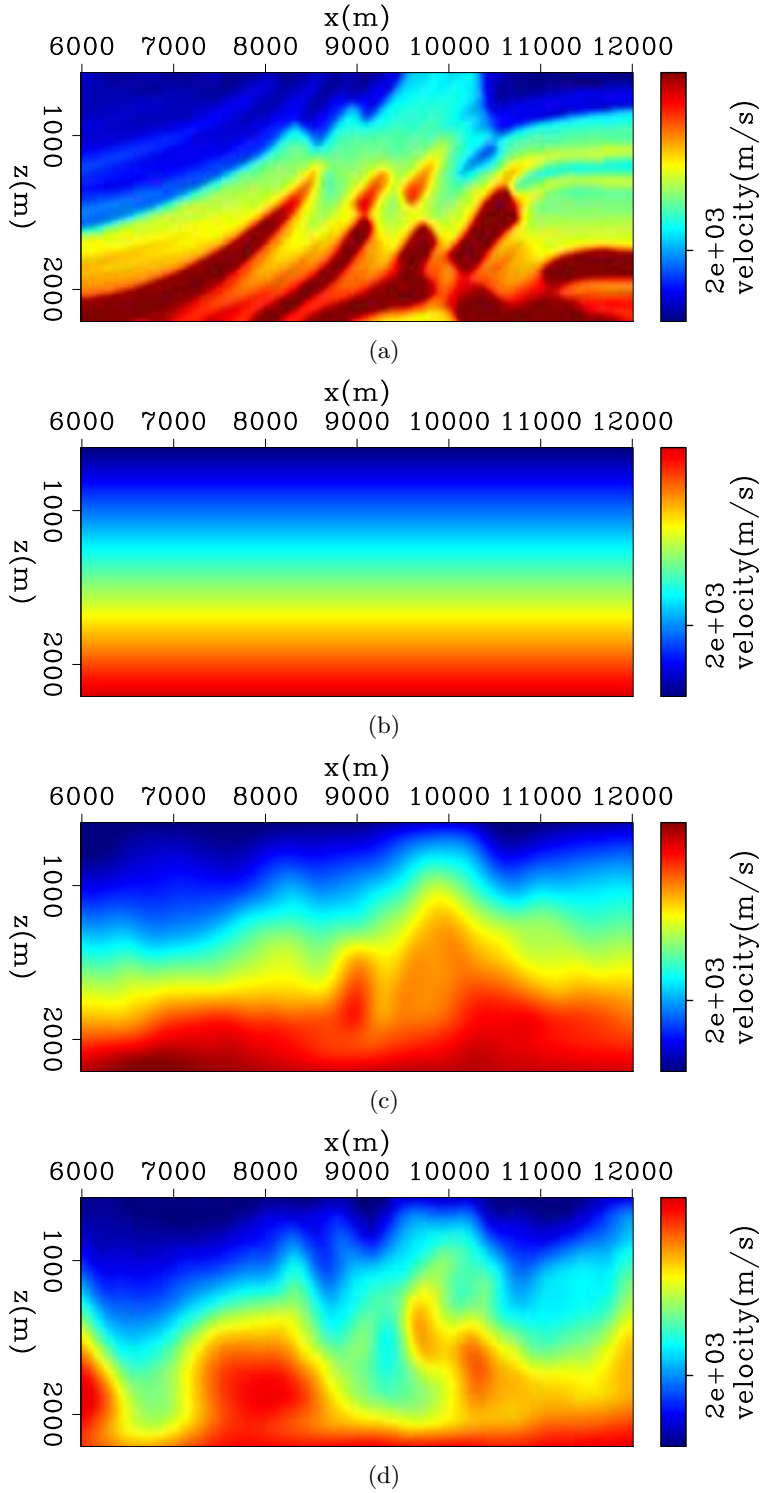


Figure 8: The true velocity model of Marmousi (a) (slightly smoothed) [**ER**]; the starting velocity model (b) [**ER**]; the inverted velocity model using our method (c) and using the subsurface-offset DSO (d). [**CR**]

yang1/. v.sm.marm,v.zg.marm,vinv.zg.bf2g.marm,vinv.zg.dso.marm

specially for many helpful discussions and the one-way wave-equation propagator used in this report.

REFERENCES

Almomin, A., 2011, Correlation-based wave-equation migration velocity analysis: SEP-Report, **143**, 33–42.

Biondi, B., 2003, 3-D Seismic Imaging: SEG.

———, 2010, Wave-equation migration velocity analysis by residual moveout fitting: SEP-Report, **142**.

———, 2011, Migration velocity analysis by one-parameter residual moveout fitting in presence of strong lateral velocity anomalies: SEP-Report, **143**, 59–66.

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

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

Luo, Y. and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

Sava, P., 2004, Migration and Velocity Analysis by Wavefield Extrapolation: PhD thesis, Stanford University.

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

Shen, P., W. W. Symes, S. Morton, A. Hess, and H. Calandra, 2005, Differential semblance velocity analysis via shot profile migration: SEG Technical Program Expanded Abstracts, **24**, 2249–2252.

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

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Toldi, J., 1985, Velocity analysis without picking: PhD thesis, Stanford University.

APPENDIX-A

Details of the $\frac{\partial b}{\partial s}$ sensitivity kernel calculation

This section illustrates how to calculate the sensitivity kernel of the image shift parameter $b(\gamma, x)$: $\frac{\partial b}{\partial s}$. Since b maximizes the auxillary objective function,

$$J_{aux}(b) = \int dz_w \int d\gamma \, I(z + z_w + b, \gamma; z, x, s_0) I(z + z_w, \gamma; z, x, s) \text{ for each z,x}, \tag{7}$$

we have

$$\frac{\partial J_{aux}}{\partial b} = 0. \tag{8}$$

To find the relation between b and $s(x)$, we differentiate equation (8) with respect to b and s, which yields

$$\frac{\partial^2 J_{aux}}{\partial b^2} \frac{\partial b}{\partial s} = -\frac{\partial J_{aux}}{\partial b \partial s}, \tag{9}$$

in which we can find

$$\frac{\partial J_{aux}}{\partial b} = \int dz_w \dot{I}(z + z_w + b, \gamma; z, x, s_0) I(z + z_w, \gamma; z, x, s),$$

$\dot{I}, \ddot{I}$ indicate the first and second derivatives in z (depth). Let

$$\frac{\partial^2 J_{aux}}{\partial b^2} = \int dz_w \, \ddot{I}(z + z_w + b, \gamma; x, z, s_0) I(z + z_w, \gamma; x, z, s) = E(x, z).$$

Then substituting the above two equations into eq. (9) leads to

$$\frac{\partial b}{\partial s} = -\frac{\int dz_w \, \dot{I}(z + z_w + b, \gamma; z, x, s_0) \frac{\partial I(z+z_w, \gamma; z, x, s)}{\partial s}}{E(\gamma, z, x)}, \tag{10}$$

which is eq. (2).

APPENDIX-B

Details of $\frac{\partial \alpha}{\partial s}$ sensitivity kernel calculation

This section provides the derivation of the *direct* $\frac{\partial \alpha}{\partial s}$ operator. The approach is to define the auxilary function that links moveout parameter $\alpha(z, x)$ and the slowness s.

$$J_{aux} = \int dz_w \int d\gamma \, I(z + z_w + \theta(\alpha, \beta, \gamma), \gamma; z, x, s_0) I(z, \gamma; z, x, s) \text{ for each z,x}, \tag{11}$$

in which

$$\theta(\alpha, \beta, \gamma) = \alpha \tan^2 \gamma + \beta = \alpha g(\gamma) + \beta h(\gamma).$$

Notice that here the moveout between the initial image $I(s_0)$ and the new image $I(s)$ is characterized by *both curvature α and constant shift β*. The β parameter does **not** affect the gather flatness; therefore there is no need to put it into the objective function in () , however β is essential to describe the change of image's kinematics.

Since α, β maximize eq (11),

$$\begin{cases} \frac{\partial J_{aux}}{\partial \alpha} = 0 \\ \\ \frac{\partial J_{aux}}{\partial \beta} = 0 \end{cases}. \tag{12}$$

We differentiate the equation (12) with respect to α, β and s, which gives

$$\begin{bmatrix} \frac{\partial^2 J_{aux}}{\partial \alpha^2} & \frac{\partial^2 J_{aux}}{\partial \alpha \partial \beta} \\ \\ \frac{\partial^2 J_{aux}}{\partial \alpha \partial \beta} & \frac{\partial^2 J_{aux}}{\partial \beta^2} \end{bmatrix} \begin{bmatrix} \frac{\partial \alpha}{\partial s} \\ \\ \frac{\partial \beta}{\partial s} \end{bmatrix} = - \begin{bmatrix} \frac{\partial J_{aux}}{\partial \alpha \partial s} \\ \\ \frac{\partial J_{aux}}{\partial \beta \partial s} \end{bmatrix}. \tag{13}$$

Now we need to invert a Jacobian to get $d\alpha/ds$. We define the following:

$$\begin{aligned}
\frac{\partial^2 J_{aux}}{\partial \alpha^2} &= \int dz_w \int d\gamma\, \ddot{I}(z+z_w+\theta,\gamma;z,x,s_0) g^2(\gamma) I(z+z_w,\gamma;z,x,s) = \quad E_{11}(z,x) \\
\frac{\partial^2 J_{aux}}{\partial \alpha \partial \beta} &= \int dz_w \int d\gamma\, \ddot{I}(z+z_w+\theta,\gamma;z,x,s_0) g(\gamma) h(\gamma) I(z+z_w,\gamma;z,x,s) = \quad E_{12}(z,x) \\
\frac{\partial^2 J_{aux}}{\partial \beta^2} &= \int dz_w \int d\gamma\, \ddot{I}(z+z_w+\theta,\gamma;z,x,s_0) h^2(\gamma) I(z+z_w,\gamma;z,x,s) = \quad E_{22}(z,x)
\end{aligned} \tag{14}$$

Let the inverse of matrix E to be matrix F:

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{12} & F_{22} \end{bmatrix} = \begin{bmatrix} E_{11} & E_{12} \\ E_{12} & E_{22} \end{bmatrix}^{-1}$$

Then

$$\frac{\partial \alpha}{\partial s}|_{s=s_0} = -\int dz_w \int d\gamma (F_{11} g(\gamma) + F_{12} h(\gamma)) \dot{I}(z+z_w,\gamma;z,x,s_0) \frac{\partial I(z+z_w,\gamma;z,x,s)}{\partial s}. \tag{15}$$

There is another way to define the relation between α and s, leading to the *indidirect* $\frac{\partial \alpha}{\partial s}$ operator using a weighted least-squares fitting formula. Suppose we have the locations of one event in the ADCIGs at location (γ_i, z_i), and we introduce a moveout formula $\theta(\gamma) = z_0 + \alpha \tan^2\gamma$. Now we define the best fitted intercept and curvature (z_0 and α) values as follows:

$$(z_0, \alpha) = \arg\min_{z_0,\alpha} \sum_i \{(z_i - \alpha \tan^2\gamma_i - z_0)^2 w_i^2\}, \tag{16}$$

where w_i is the energy of the event at angle γ_i, serving as a weight for the least-squares fitting. We denote $\overline{x} = \sum_i x_i w_i^2$ (the weighted average of quantity x), then

$$\alpha = \frac{\overline{1}\,\overline{(z\tan^2\gamma)} - \overline{\tan^2\gamma}\,\overline{z}}{\overline{1}\,\overline{\tan^4\gamma} - (\overline{\tan^2\gamma})^2}.$$

It is easy to find $\Delta\alpha$ if there is a pertubation of b_i on z_i:

$$\frac{\partial \alpha}{\partial b_i} = \frac{w_i^2(\overline{1}\,\tan^2\gamma_i - \overline{\tan^2\gamma})}{\overline{1}\,\overline{\tan^4\gamma} - (\overline{\tan^2\gamma})^2}\,. \tag{17}$$

Finally,

$$\frac{\partial \alpha}{\partial s} = \sum_i \frac{\partial \alpha}{\partial b_i} \frac{\partial b_i}{\partial s}, \tag{18}$$

where $\frac{\partial \alpha}{\partial b_i}$, $\frac{\partial b_i}{\partial s}$ are defined respectively in equations (17) and (10).

The real reason for the reproducibility problems in SEP:
I THINK WE SHOULD GIVE IT ANOTHER SHOT.
WE SHOULD BREAK UP, AND I CAN PROVE IT.
OUR RELATIONSHIP
HUH.
MAYBE YOU'RE RIGHT.
I KNEW DATA WOULD CONVINCE YOU.
NO, I JUST THINK I CAN DO BETTER THAN SOMEONE WHO DOESN'T LABEL HER AXES.

Residual-moveout analysis in presence of strong lateral velocity anomalies

Biondo Biondi

ABSTRACT

The analysis of a simple synthetic data set recorded above a strong velocity anomaly and a flat reflector illustrates the challenges that can be encountered when performing residual-moveout analysis using a family of curves described by a single parameter. Overcoming these challenges is important if we want to use automatic velocity analysis methods that rely on the derivative of the stack power with respect to the residual-moveout parameter to compute velocity gradients. My analysis shows how, at some reflector locations, the stack-power may have a poorly defined peak because the residual moveout is more complex than the one-parameter model assumes. At other reflector locations, the peak of the stack-power is sharp but it is too far from the value of the parameter corresponding to no residual moveout. Consequently, the derivatives are unreliable, and possibly have even the wrong sign. More robust information could be provided by migrating data with lower frequencies, when available. A more general solution is smoothing the stack power along the residual-moveout parameter before evaluating its derivatives.

INTRODUCTION

Methods to perform wave-equation migration velocity analysis without requiring velocity-spectra picking are attractive and have been the focus of substantial effort at SEP (Biondi, 2008, 2010; Zhang and Biondi, 2011). Robustness should be an important characteristic of these methods. Therefore, we have focused on algorithms that extract velocity information from migrated angle-domain common image gathers (ADCIG) using a one-parameter residual-moveout analysis. One-parameter moveout analysis has the advantage of being more robust to noise, imaging artifacts, and cycle skipping than alternative methods for measuring residual moveout. It has been extensively used for ray-based migration velocity analysis where it has proven to provide useful information for velocity updating when the moveout parameter is picked from stack-power, or semblance, scans. However, it has not been extensively tested when the velocity information is extracted by computing the derivative of the focusing measure (e.g. semblance or stack power) around the origin of the moveout-parameter axis, as is required by the automatic methods we have been developing.

A one-parameter moveout cannot accurately describe the actual moveout of the migrated gathers in some important cases, such as in the presence of strong lateral velocity anomalies and anisotropy. When the velocity errors are large and the migrated gathers display a significant (i.e. larger than the dominant wavelength in the image) moveout at wide angles, numerical differentiation of stack-power scans can be prone to errors caused by cycle skipping.

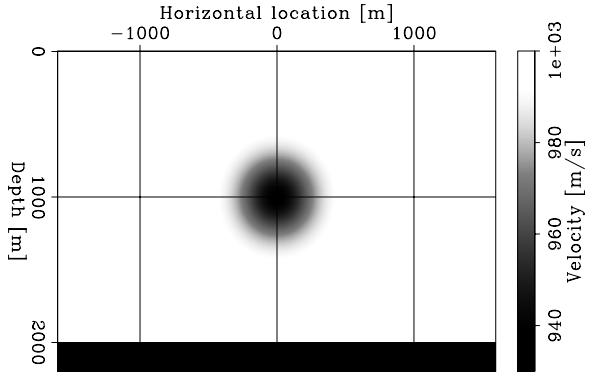


Figure 1: Central part of the velocity model used to model the data. [**ER**] biondo1/. Vel-overn

To test the robustness of one-parameter moveout analysis for automatic wave-equation migration velocity analysis, I performed numerical experiments on a simple synthetic data set. I started from the angle-domain image generated from data that were modeled assuming a strong velocity anomaly, but migrated with a constant background velocity. In the migrated image there are areas that illustrate both the challenges described above (complex residual moveout and cycle skipping caused by large velocity errors). At the original frequency band of the data (25 Hz dominant frequency) the straightforward computation of the gradients would likely result in poor convergence. After I applied a low-pass filter to the data (high cut at 8 Hz) the gradient becomes better behaved. However, seismic data are not currently recorded with sufficient signal-to-noise ratio at arbitrarily low frequencies. Smoothing the velocity spectra along the moveout parameter axis is a simple remedy that does not require low-frequency data. This smoothing is sufficient to overcome the problems identified from the test data set. Numerical differentiation of smoothed stack-power scans provides useful information to be used for a tomographic update even when the data are migrated at full bandwidth.

TEST DATA AND IMAGE

I performed my test on a simple synthetic data set. The data were modeled assuming a strong, and fairly localized, velocity anomaly above a flat reflector. Figure 1 shows the central part of the model. I modeled 400 split-spread shot gathers with offsets ranging from -2 km to 2 km. The source function was a Ricker wavelet with central frequency of 25 Hz.

The localized velocity anomaly caused clear non-hyperbolicity in the data common-midpoint (CMP) gathers. Figure 2 shows two typical CMP gathers in the data set. Figure 2a shows the CMP gather at horizontal location X=0 km; that is, in the middle of the model, and Figure 2b shows the CMP gather at X=.55 km. The non-hyperbolicity in the CMP gathers makes this a challenging data set for velocity methods that characterize moveout with one-parameter curves.

Migrated image with background velocity

I migrated the data set introduced above by using reverse time migration and assuming a constant background velocity of 1 km/s. Figure 3 shows the angle-domain image produced by this migration. Figure 3a shows the stacked section (i.e. zero subsurface-offset section).

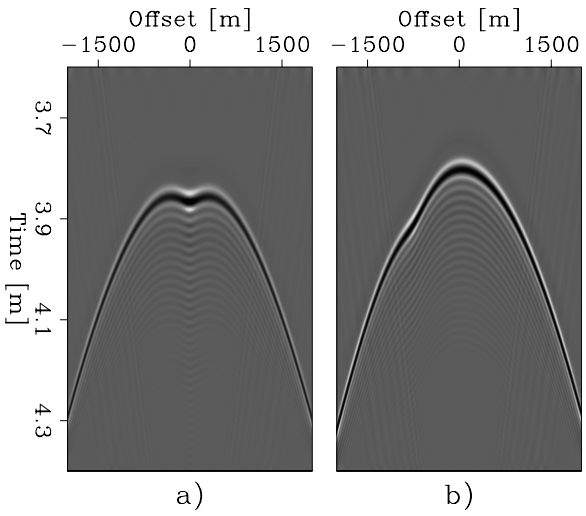


Figure 2: Two typical CMP gathers in the data set: a) X=0 km, and b) X=.55 km. [**CR**] biondo1/. Data-all-overn

The strong residual moveout prevents a coherent stack in the central area of the reflector, approximately between X=-.8 km and X=.8 km. The three panels on the right of the stacked section display the ADCIGs taken at three midpoint locations, identified by the vertical lines superimposed onto the stacked section; that is, at: b) X=0 km, c) X=.55 km, and d) X=1.1 km.

The one-parameter residual-moveout analysis conducted on this image is based on approximating the vertical shifts as being directly proportional to the square of the tangent of the aperture angle (Biondi and Symes, 2004). The ADCIG shown in Figure 3b obviously does not fulfill this approximation because the velocity error is quickly changing along the horizontal direction. Furthermore, because the cumulative kinematic error caused by the anomaly is large, the moveout at wide angles in the ADCIG shown in Figure 3c is sufficiently large to cause cycle-skip problems even when using a supposedly robust one-parameter moveout analysis.

Figure 4 shows panels equivalent to the ones shown in Figure 3 after the data were drastically low passed before migration. The peak frequency of the data was reduced to 5 Hz, from the original 25 Hz of the full-bandwidth data. Because of the lower frequency, the wide angles in the ADCIG at .55 km (Figure 4c) do not suffer from cycle-skip problems. Consequently, the stack (Figure 4a) is now coherent over the majority of the reflector, except in the very central part approximately located between X=-.2 km and X=.2 km.

RESIDUAL MOVEOUT ANALYSIS

Starting from the images shown in the previous section, I performed a conventional residual moveout analysis by applying the following angle-domain moveout

$$\Delta z = (1 - \rho) \tan^2 \gamma, \tag{1}$$

over a range of values for ρ, and then computing the stack power from the moved-out ADCIGs. For constant velocity errors in the half space above the reflector, the parameter ρ is approximately related to the ratio between the current migration slowness s_{mig} and the true slowness s; that is, $\rho \approx s_{\text{mig}}/s$ (Biondi and Symes, 2004).

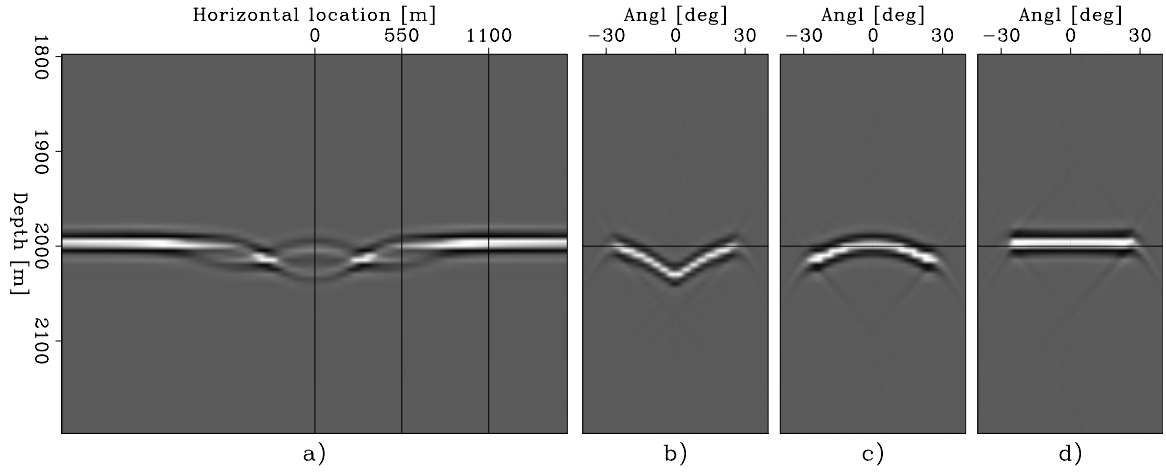


Figure 3: Angle-domain prestack image obtained from the full-bandwidth data set: (peak frequency at 25 Hz) a) stacked section, b) ADCIG at X=0 km, c) ADCIG at X=.55 km, and d) ADCIG at X=1.1 km. [**CR**] biondo1/. Mig-0-stack-X-overn

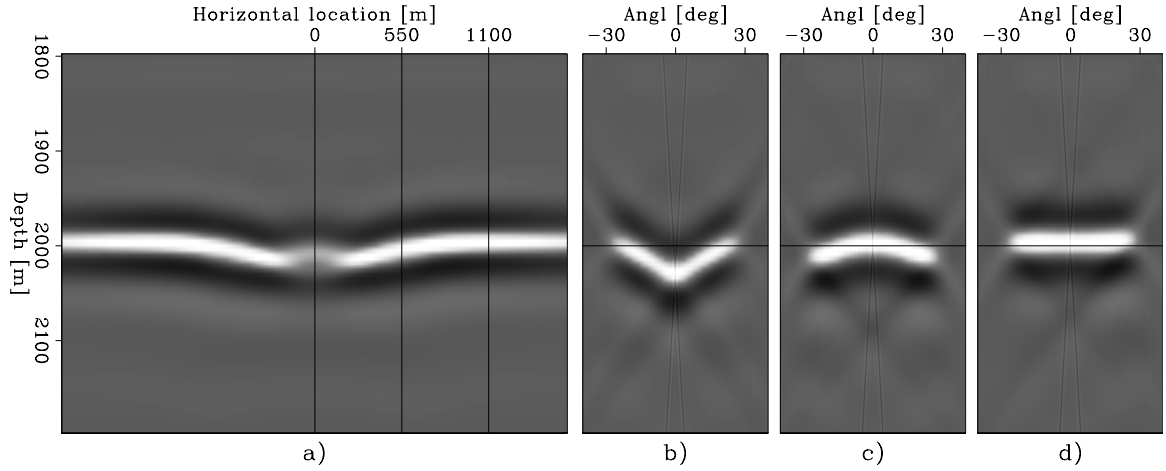


Figure 4: Angle-domain prestack image obtained from the low-passed data set (peak frequency at 5 Hz): a) stacked section, b) ADCIG at X=0 km, c) ADCIG at X=.55 km, and d) ADCIG at X=1.1 km. [**CR**] biondo1/. Mig-0-VLowFreq-stack-X-overn

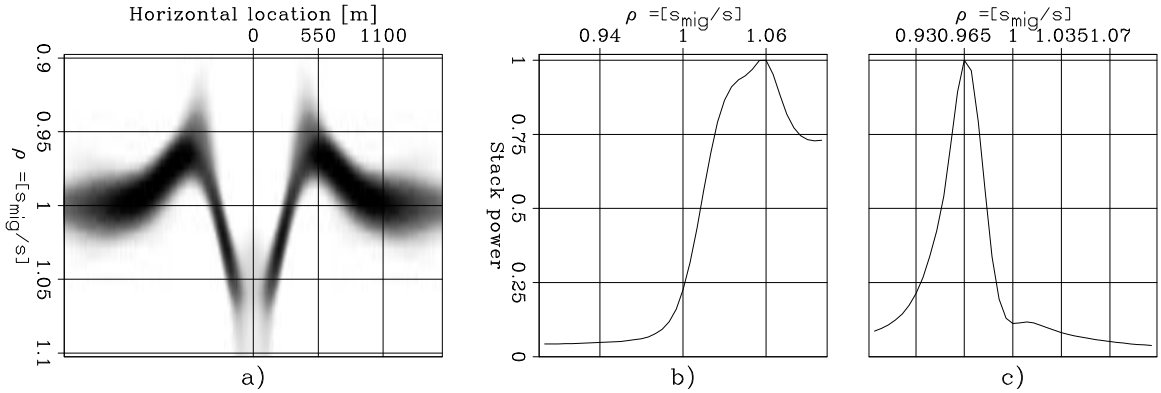


Figure 5: (a) Stack power as a function of horizontal location X and the moveout parameter ρ, corresponding to the full-bandwidth image shown in Figure 3. Graphs of the function in panel a) at (b) X=0 km, and (c) X=.55 km. [**CR**] biondo1/. Avg-Power-all-0-overn

Figure 5a shows the stack power as a function of horizontal location X and moveout parameter ρ, averaged over the depth interval of the reflector. The panels in Figure 5b and 5c show the graphs of this function at (b) X= 0 km and (c) X=.55 km.

In the middle of the reflector the residual moveout is not well described by a one parameter curve, and thus in Figure 5b the stack power peak is broad and not well defined. Figure 6 shows the central ADCIG before (a) and after (b) residual moveout with ρ=1.06. Whereas the power of the stack is maximum for ρ=1.06 (see Figure 5b), the gather shown in Figure 6b is far from being flat.

In contrast, at X=.55 km, the residual moveout is well described by a one-parameter curve and the stack power peak is sharp and well defined in Figure 5c. However, at $\rho = 1$ the stack power curve is almost flat. If we relied on the numerical derivative of this curve to compute the velocity gradient, we might be relying on the wrong information. The power of the stack is maximum for ρ=.965 (see Figure 5c) and indeed the ADCIG moved-out with this value of ρ is flat, as shown in Figure 7b.

A simple solution to the problems identified above could be to image only the low frequency component of the data. Figure 8 shows the stack-power function when computed from the low-frequency image shown in Figure 4. In this case the stack-power peaks are well defined at both X=0 km and X=.55 km, and they are sufficiently broad that the derivative of the stack-power with respect to ρ, evaluated at $\rho = 1$, would provide useful information for the computation of the velocity gradient.

However, seismic data are not always available with sufficient signal-to-noise ratio at low frequencies. In these cases, the challenge can be tackled by smoothing the stack-power function along the moveout parameter before evaluating the derivatives. Figure 9 shows the stack-power function when computed from the full-bandwidth image and then smoothed along the ρ axis. This function has many similarities to the low-frequency one shown in Figure 8, but does not require data with good signal-to-noise ratio at low frequencies.

Finally, Figure 10 shows the derivatives of the stack-power functions shown in the previous three figures, evaluated numerically at $\rho = 1$. These functions would be the starting

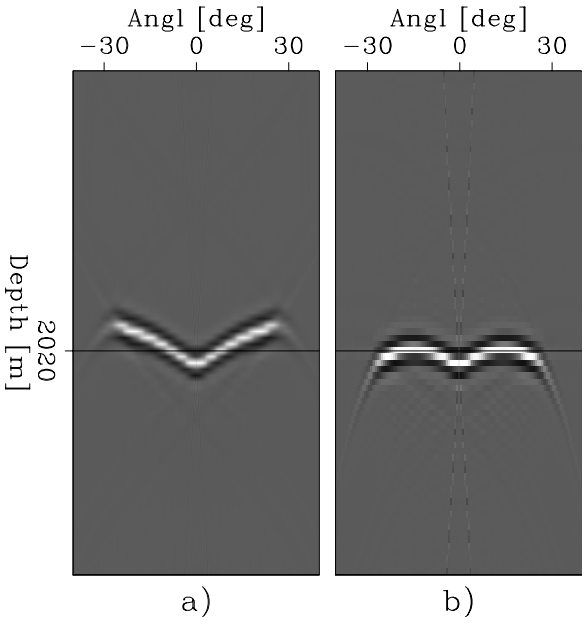


Figure 6: ADCIGs at X=0 km before (a) and after (b) residual moveout with ρ=1.06. [**CR**] biondo1/. Rmo-all-X0-overn

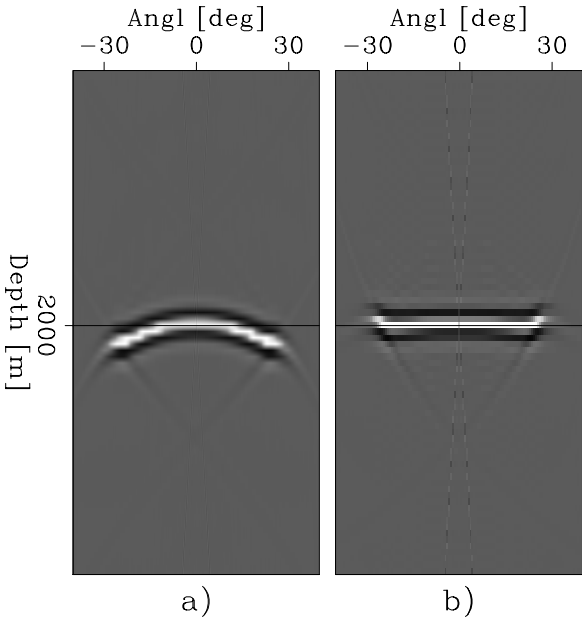


Figure 7: ADCIGs at X=.55 km before (a) and after (b) residual moveout with ρ=.965. [**CR**] biondo1/. Rmo-all-X550-overn

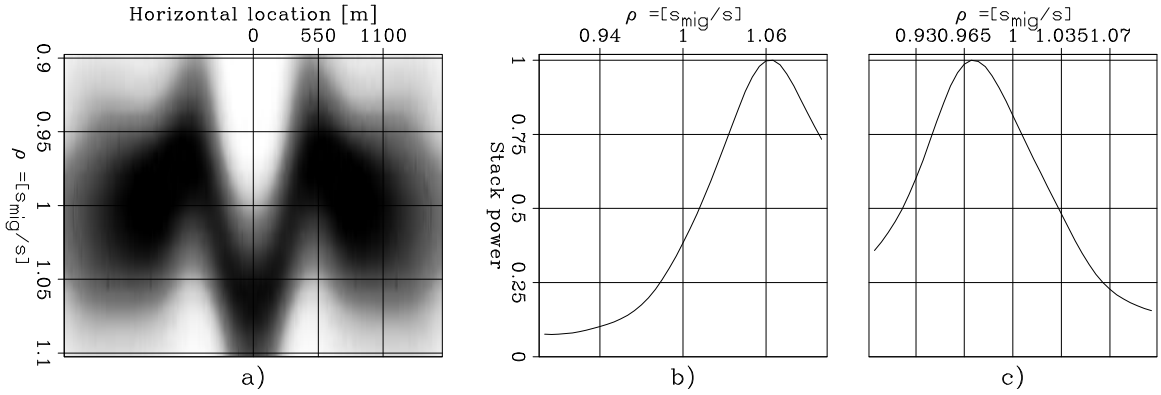


Figure 8: (a) Stack power as a function of horizontal location and moveout parameter ρ corresponding to the low-passed image shown in Figure 4. Graphs of the function in panel a) at (b) X=0 km, and (c) X=.55 km. [**CR**] biondo1/. Avg-Power-all-0-VLowFreq-overn

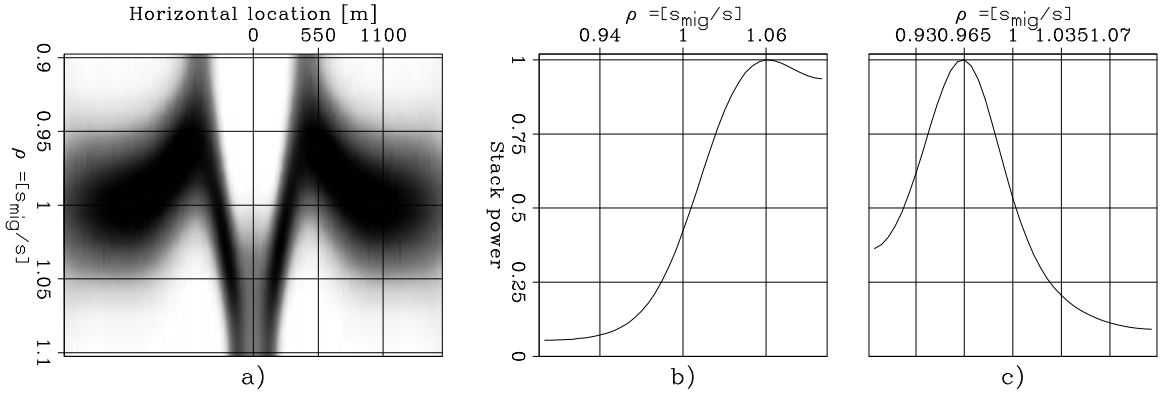


Figure 9: (a) Stack-power function resulting from smoothing along the ρ axis the function shown in Figure 5. Graphs of the function in panel a) at (b) X=0 km, and (c) X=.55 km. [**CR**] biondo1/. Smooth-Power-all-0-overn

data from which the velocity gradient is computed in a wave-equation migration velocity analysis method (Biondi, 2008, 2010; Zhang and Biondi, 2011). The solid line, which corresponds to the full-bandwidth data without smoothing, would provide misleading information and possibly would prevent proper convergence of the velocity estimation algorithm. On the contrary, both the curve computed from the low-frequency data (dotted line) and the one obtained by smoothing the stack-power along ρ (dashed line) would provide useful information for the computation of the gradient.

CONCLUSIONS

Strong lateral velocity anomalies are challenging for one-parameter residual-moveout analysis. The migrated ADCIGs may display complex moveouts that are not accurately described by a a one-parameter family of curves. Furthermore, when the residual velocity error is large enough, the moveout at far angles may be larger than the dominant wavelength, and thus

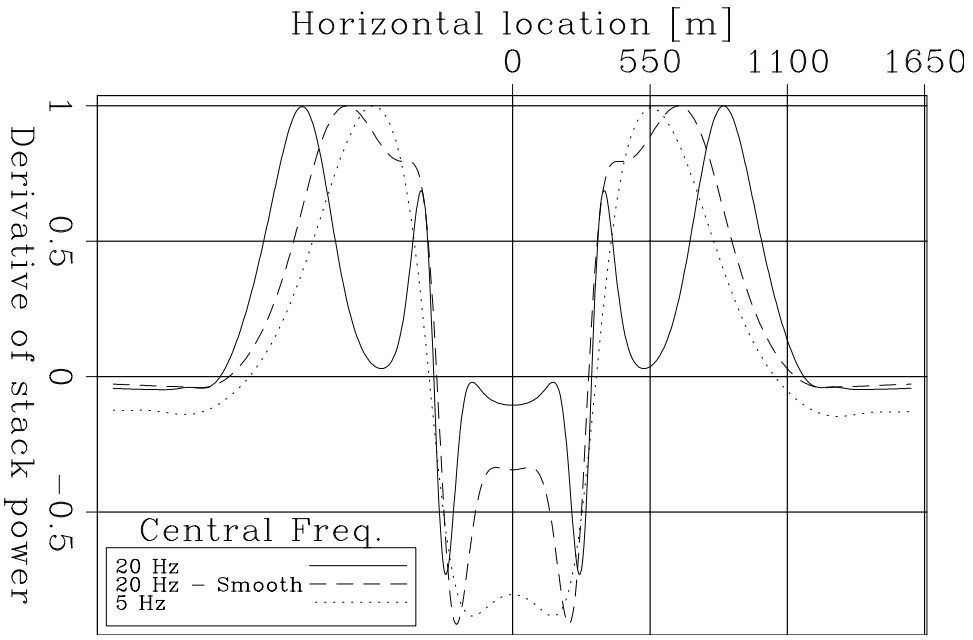


Figure 10: Derivatives of the stack-power functions evaluated numerically at $\rho = 1$: solid line - full-bandwidth data (Figure 5), dashed line - full-bandwidth data with smoothing (Figure 9), dotted line - low-passed data (Figure 8). [**CR**] biondo1/. Der-Power-all-overn

stack-power spectra may display local maxima. The negative effects of these phenomena can be avoided by imaging only the low-frequency components of the data, when they are available. A more general solution is to smooth the stack-power function along the residual-moveout parameter axis.

REFERENCES

Biondi, B., 2008, Automatic wave-equation migration velocity analysis: 2008, **134**, 65–77.

———, 2010, Wave-equation migration velocity analysis by residual moveout fitting: SEP-Report, **142**, 25–32.

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

Zhang, Y. and B. Biondi, 2011, Moveout-based wave-equation migration velocity analysis: SEP-Report, **143**, 43–58.

A new waveform inversion workflow: Application to near-surface velocity estimation in Saudi Arabia

Xukai Shen, Thierry Tonellot, Yi Luo, Tim Keho and Robert Ley

ABSTRACT

Waveform inversion is a more accurate near-surface velocity estimation tool than ray-based methods. It is able to solve complex near-surface velocity structure where conventional ray-based methods fail. If errors in the initial model are too large, waveform inversion will fail to converge to the correct model. This convergence problem is particularly obvious for large near-surface velocity contrasts and velocity inversions. We propose to address this issue with a new inversion workflow that adds a wave-equation traveltime inversion step prior to waveform inversion. The performance of our approach is evaluated on both synthetic and field data.

INTRODUCTION

Typically, conventional ray-based methods (Hampson and Russell, 1984; Olson, 1984; White, 1989) are used to derive near-surface velocity. Such smooth solutions may be adequate for areas with simple near-surface velocity distributions, but in geologically complex areas, smooth velocities are not accurate enough for imaging deeper reflectors (Marsden, 1993; Bevc, 1995; Hindriks and Verschuur, 2001). In such cases, waveform inversion (Tarantola, 1984; Pratt et al., 1998; Mora, 1987) tends to give more accurate results (Ravaut et al., 2004; Sheng et al., 2006; Sirgue et al., 2009) by simulating seismic wave propagation.

With more and more long offset data being acquired, particularly on land, it is not uncommon to see shingling in data (Figure 1). Shingling is defined here as where first-

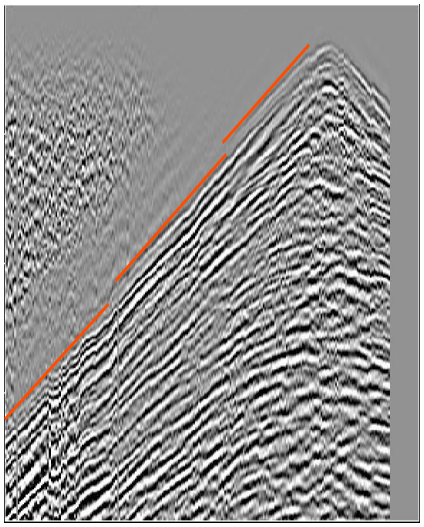


Figure 1: Shot gather of 2D land dataset showing shingling phenomenon, red lines denote approximate first-break picks, notice how they become discontinuous as offset increases.[**NR**] xukai1/. realshot

break traveltimes are only piecewise continuous when we look at data in shot gathers. This phenomenon is likely to be caused by a near-surface low velocity layer (Figure 2). Shingling poses a serious problem for ray-based inversion methods. Note that at the shot gather in Figure 1, the first-break amplitudes decay so much that that at a certain offset, the only reliable pick on a trace becomes the later refraction arrival in time. The discontinuity is caused by tracking different events as we move further away from source location. Since ray-based methods assume the same event is being picked, they tend to fail in these types of situations.

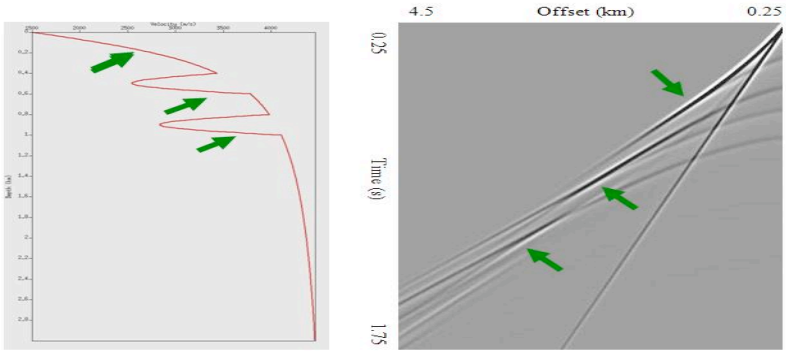


Figure 2: A synthetic data case showing a) a 1D velocity model with low velocity layers in the near-surface and b) the resulting shot gather. Green arrows denote decaying refractions on the right and where they occur in near-surface model on the left. [**NR**] xukai1/. shingling

Waveform inversion relies on the wave equation and is able to solve the velocity structure quite well even in the presence of low velocity layers. Results from waveform inversion are highly dependent on the accuracy of the starting model and on the frequency content of the data. To relax the dependence of the waveform inversion on the initial model, we propose the work flow described in the next section.

WAVEFORM INVERSION WORKFLOW

Our waveform inversion workflow consists of three successive steps:

```
1) First arrival traveltime tomography.
2) Wave-equation traveltime inversion.
3) Early arrival acoustic waveform inversion.
```

The first step relies on ray theory to estimate an initial low frequency velocity model to explain the first arrival picks. In the second step, we use the output of the conventional traveltime tomography as input to wave-equation traveltime inversion (Luo and Schuster, 1991). This result is then used as input for the third step of full waveform inversion. Since both the traveltime and waveform inversion are derived from the same wave-equation, they can both be described in a common framework where the objective function can be written as

$$f\left(\mathbf{d_{obs}}, \mathbf{D}(\mathbf{m})\right) \approx \mathbf{0}, \tag{1}$$

where f is a function of $\mathbf{d_{obs}}$ and $\mathbf{D(m)}$, the observed data and forward modeled synthetic data from the velocity model $\mathbf{m}$ respectively. Observed data can be either in frequency domain or in time domain, depending on the actual form of $\mathbf{f}$. For example if we take $\mathbf{f}$ as the L2 norm of $(\mathbf{d_{obs}} - \mathbf{D(m)})$, we obtain the objective function of conventional waveform inversion (Tarantola, 1984; Pratt et al., 1998); if we take $\mathbf{f}$ as the L2 norm of the time lag difference of the cross-correlation of $\mathbf{D(m)}$ and $\mathbf{d_{obs}}$, we obtain an objective function for wave-equation traveltime inversion. (Luo and Schuster, 1991). This second formulation is more robust than the first one in the presence of large velocity contrasts or to inaccuracies in the initial model. However, wave-equation traveltime inversion provides lower model resolution compared to conventional full waveform inversion. near-surface low velocity layer and resulting shingling data lead to inaccurate velocity estimates using ray-based methods. The workflow adopted in this paper tries to compensate for these issues by adding an intermediate wave-equation traveltime inversion to the conventional workflow.

SYNTHETIC DATA APPLICATION

To evaluate the performance of the new workflow, a synthetic dataset was constructed where the true velocity model has two low velocity layers with a background velocity that smoothly increases with depth. The starting model used in the inversion is the background velocity without the two low velocity layers. The velocity is laterally invariant. The maximum offset of the survey is 8 km. The synthetic data are generated using constant density acoustic modeling engine. A typical shot gather (Figure 3) shows shingling as we move further away from source location. In the inversion, the source wavelet was known. We employed the workflow described above and ran the inversion using data with a peak frequency of 10 Hz. Inversion results are shown in Figure 4. The blue and red curves denote the true velocity model and the starting velocity model respectively, and the fuchsia curve is the inversion result. Notice that we were able to successfully recover both low velocity layers even when using relatively high frequencies. We also tried direct waveform inversion without wave-equation traveltime inversion using the same data. In this case, convergence to the right solution occurred only with peak frequencies of 7 Hz or lower.

FIELD DATA APPLICATION

Our new workflow was applied to a 2D land dataset acquired in Saudi Arabia. The line geometry and the stacked section of the 2D data are shown in Figure 5. The section of the line used for waveform inversion is between the two vertical red lines. The acquisition geometry was not strictly 2D, that is it was not in a straight line. The waveform inversion algorithm that was used is a 2D implementation of the time-domain approach described in Shen (2010). The line geometry was hence converted to 2D by a simple projection of the line onto the x-axis. The starting model was obtained by ray-based method using picked first-break traveltimes. Starting source wavelets were obtained by stacking part of the moved-out refractions, and vary from shot to shot. The source wavelets were then updated at each iteration of the inversion. We used a total of 110 shots, with 180m shot spacing and 30 m receiver spacing. Offset used for inversion ranged from -4000 to -400 m for each shot. The lowest frequency in the data was 10 Hz, which makes direct application of waveform-based inversion for the low frequencies of velocity difficult. Two waveform inversion runs were

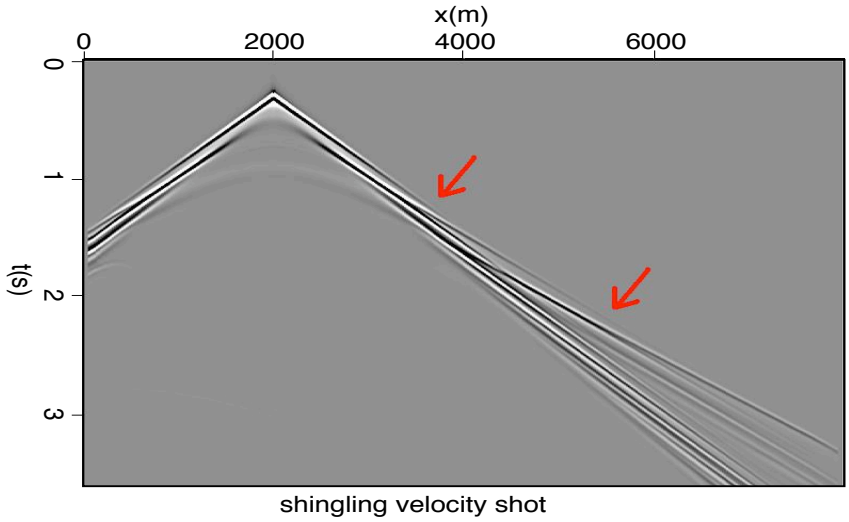


Figure 3: Synthetic shot gather by acoustic modeling using true velocity model. [**NR**] xukai1/. synshot

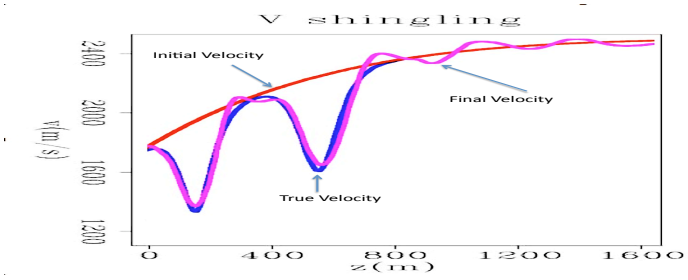


Figure 4: Inversion result with blue being true model, red being starting model and fuchsia being inverted model. [**NR**] xukai1/. syninv

performed. The first run started directly from the first arrival traveltimes tomography model and the second run included an intermediate step of wave-equation traveltime inversion applied to the full bandwidth data.

The starting model and final results are shown in Figure 6. The inversion result after the wave-equation traveltime inversion shows a continuous low velocity event in the upper middle portion of the section

The reverse time migrated (RTM) images in Figure 7 clearly show that wave-equation traveltime inversion has improved the final image. Indeed, with better velocities, we get more spatial coherence, particularly in the neighborhood of the low velocity layer where reflections are seen both in the RTM result and on the stacked section in Figure 5 (indicated by the red ellipse).

Figure 8 shows ray tracing through the final velocity model within the data offsets used for inversion. This display verifies the depth of the valid velocity model updates, which in this case confirmed that the low velocity layer is indeed the result of inversion rather than an artifact. Since we are matching waveforms, it is also important to compare the modeled data with the input refraction data (Figure 9). It can be seen that waveforms and traveltimes match quite well despite differences in absolute amplitude. However, these difference are not a problem for the waveform-based waveform inversion objective function proposed by Shen (2010). Based on these results, the wave-equation traveltime inversion step seems to considerably improve the final results. This improvement could also be explained by the fact that the ray-based model was not totally consistent with the waveform inversion algorithm used.

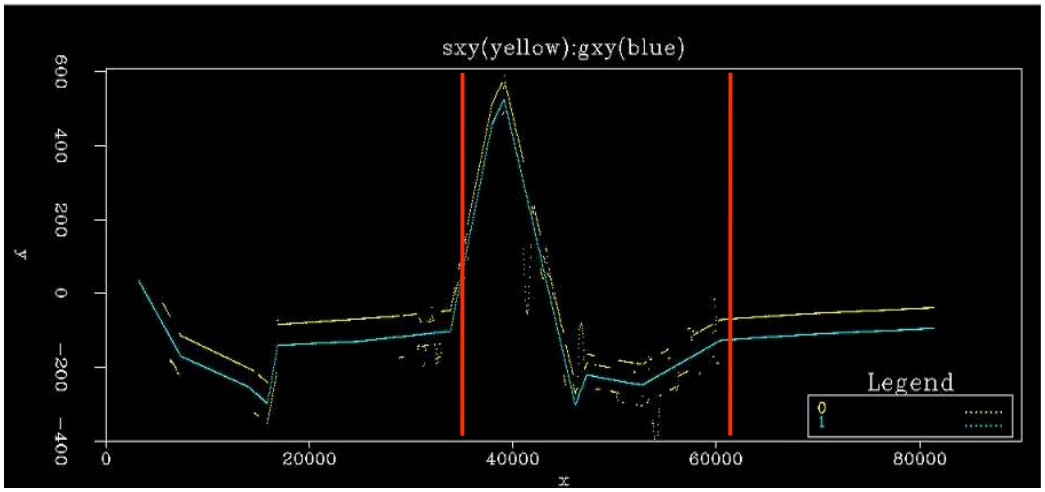
CONCLUSION

We showed a new three stage waveform inversion workflow for near-surface velocity estimation:

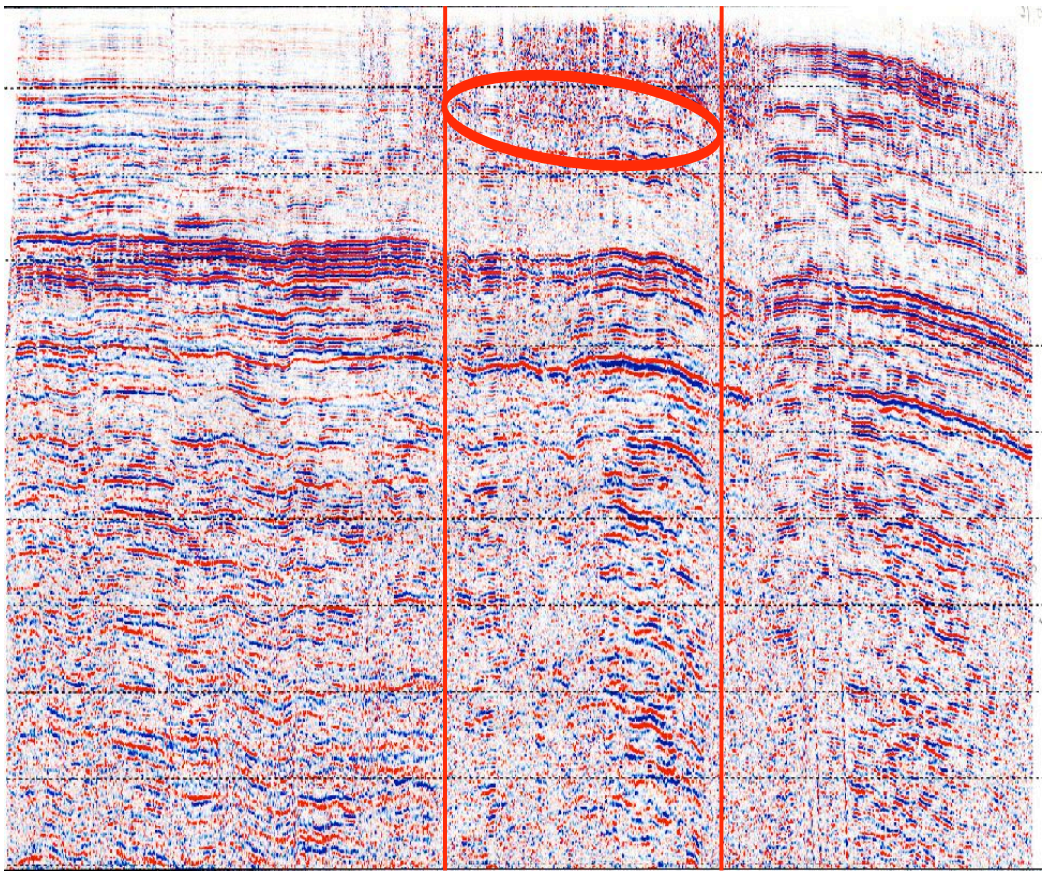
```
1) First arrival traveltime tomography.
2) Wave-equation traveltime inversion.
3) Early arrival acoustic waveform inversion.
```

The wave-equation traveltime inversion step helps correct inaccuracies in the conventional first arrival traveltime inversion result due to the inability of ray based methods to properly handle low velocity zones and large velocity contrasts.

This new workflow was evaluated using both a synthetic data example and a 2D field dataset from Saudi Arabia. Our new inversion workflow produced a superior result on the field data compared to results obtained using a more conventional strategy that did not include the wave-equation traveltime inversion step. Since the 2D field data was not acquired in a straight line, it is possible that the improvement we observed when using wave-equation traveltime inversion was partly due to the correction of the non-2D geometry. We plan to further investigate the wave-equation traveltime inversion, using a full 3D parametrization for the modeling and inversion.



a)



b)

Figure 5: A land 2D data case showing a) the x,y source receiver geometry, and b) the stacked. The area for waveform inversion is between the vertical red lines. Blue indicates source, and yellow indicates receiver locations. [**NR**] xukai1/. geostack

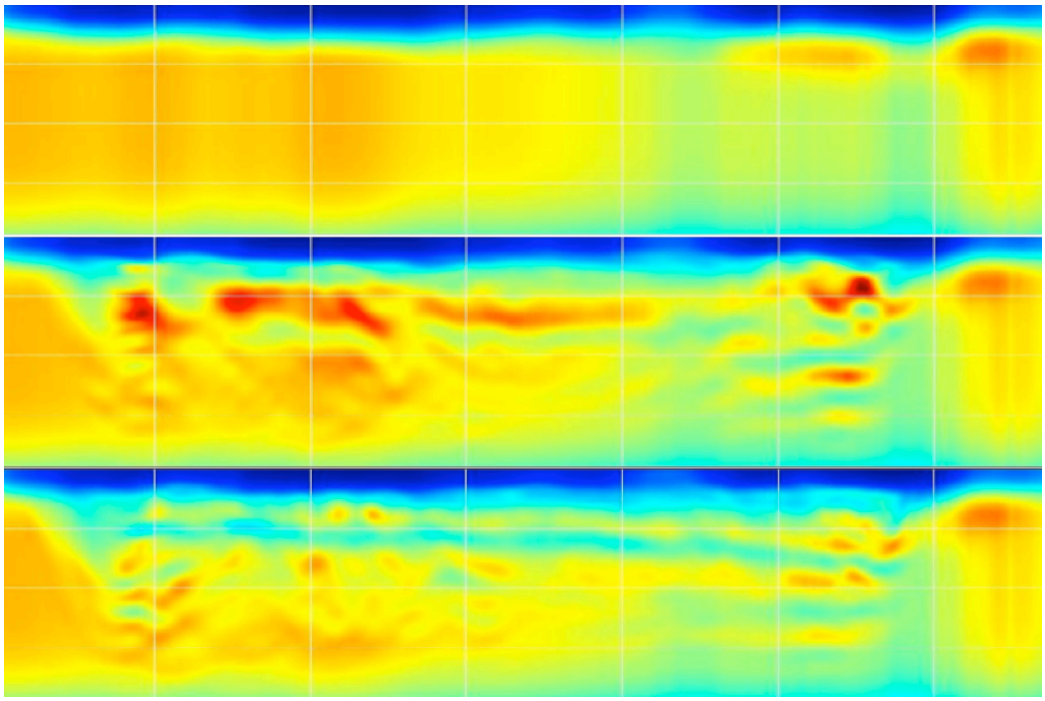


Figure 6: Starting depth velocity model using ray-based tomography (top), and waveform inversion results without (middle) and with (bottom) wave-equation traveltime inversion. [**NR**] xukai1/. threevel

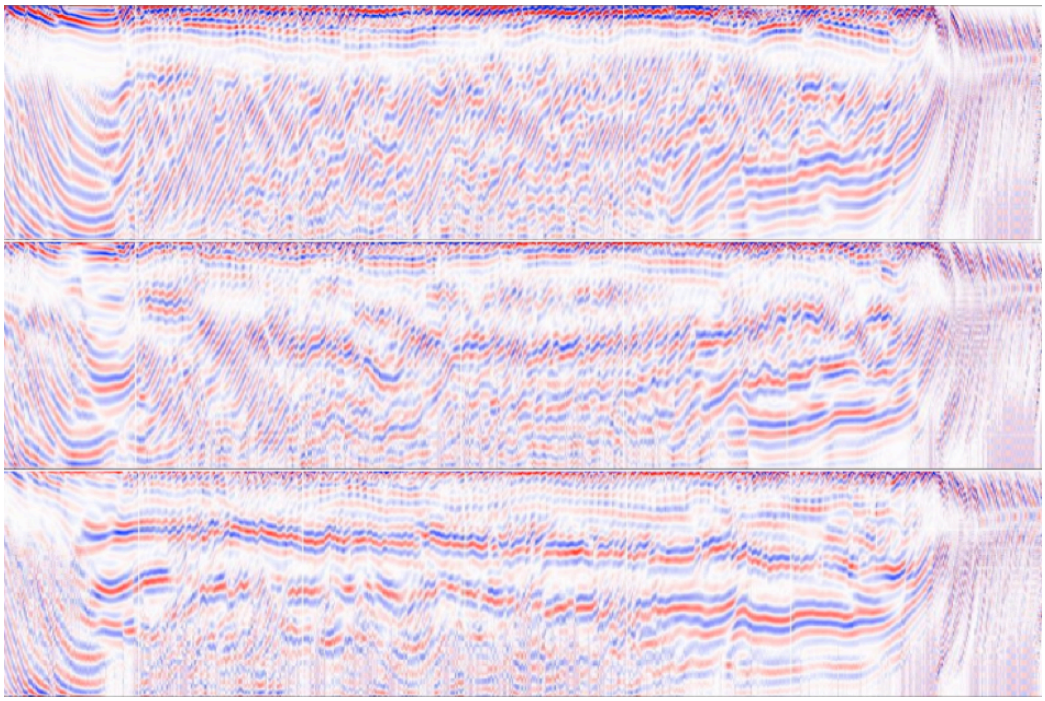


Figure 7: RTM images corresponding to the starting depth velocity model using ray-based tomography (top), and to the waveform inversion results without (middle) and with (bottom) wave-equation traveltime inversion. [**NR**] xukai1/. threeimg

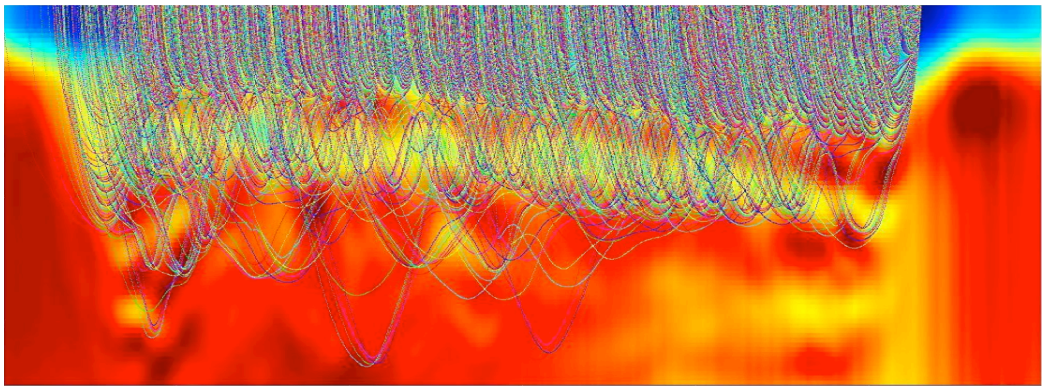


Figure 8: Ray tracing through the final velocity model. This shows that low velocity layer is indeed from inversion instead of being an artifact. [**NR**] xukai1/. realray

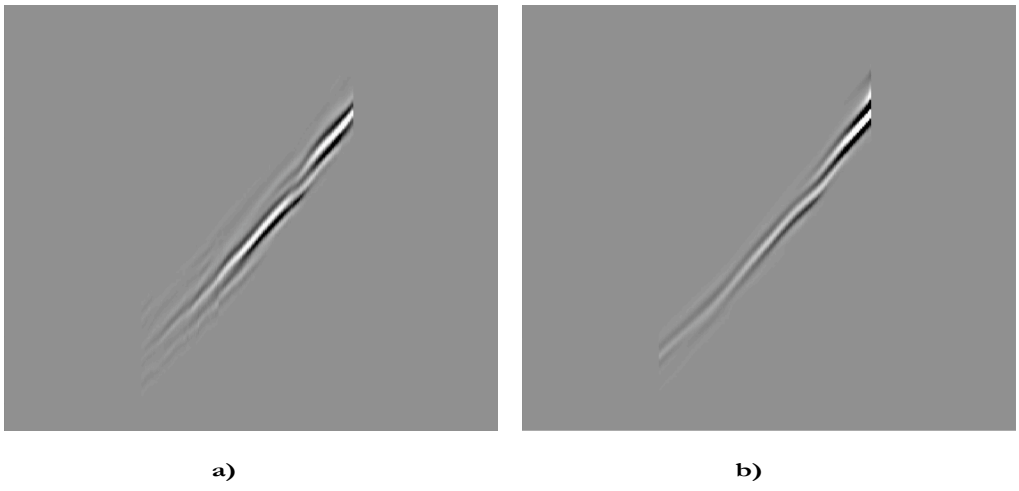


Figure 9: Comparison of a) input and b) modeled refractions. Note the similar kinematics despite minor differences in absolute amplitudes. [**NR**] xukai1/. realcomp

ACKNOWLEDGMENTS

Xukai would like to thank the Geophysical Technology Team of the EXPEC Advanced Research Center, Saudi Aramco for providing the data and the financial support for this research. We would like to thank Saudi Aramco for allowing us to publish the result.

REFERENCES

Bevc, D., 1995, Imaging under rugged topography and complex velocity structure: SEP Ph.D Thesis.

Hampson, D. and B. Russell, 1984, First-break interpretation using generalized linear inversion: Journal of CSEG, **20**, 40–54.

Hindriks, C. and D. Verschuur, 2001, Common focus point approach to complex near surface effects: SEG Expanded Abstracts.

Luo, Y. and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

Marsden, D., 1993, Statics corrections-a review: The Leading Edge.

Mora, P., 1987, Elastic wavefield inversion: SEP Ph.D Thesis.

Olson, K. B., 1984, A stable and flexible procedure for the inverse modelling of seismic first arrivals: Geophysical Prospecting, **37**, 455–465.

Pratt, R. G., C. Shin, and G. Hicks, 1998, Gauss-Newton and full Newton methods in frequency domain seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Ravaut, C., S. Operto, L. Improta, J. Virieux, A. Herrero, and P. Dell'Aversana, 2004, Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt: Geophysical Journal International, **159**, 1032–1056.

Shen, X., 2010, Near-surface velocity estimation by weighted early-arrival waveform inversion: SEG Expanded Abstracts.

Sheng, J., A. Leeds, M. Buddensiek, and G. T. Schuster, 2006, Early arrival waveform tomography on near-surface refraction data: Geophysics, **71**, U47–U57.

Sirgue, L., O.I.Barkved, J. V. Gestel, O. Askim, and R. Kommedal, 2009, 3d waveform inversion on valhall wide-azimuth obc: EAGE 71th Conference.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

White, D. J., 1989, Two-dimensional seismic refraction tomogaphy: Geophysical Journal International, **97**, 223–245.

Arghl This Makefile is really complicated!
Yeah... and this is only 2D modeling
That's what happens when you go elastic.
Just be careful.....
There is a Y axis coming.
OHAD

Combining forward-scattered and back-scattered wavefields in velocity analysis

Ali Almomin, Biondo Biondi, and Robert Clapp

ABSTRACT

Wave-Equation Migration Velocity Analysis (WEMVA) is a family of techniques that aims to improve the subsurface velocity model by minimizing the residual in the image space. This process is performed iteratively by linearizing the imaging operator in order to relate image perturbations to model updates. However, WEMVA techniques only utilize the kinematics in the forward-scattered wavefield. Ignoring the back-scattered wavefield results in model updates that have low vertical resolution. I present a method that combines the forward-scattered wavefield information from WEMVA with the back-scattered wavefield information from full-waveform inversion (FWI). This can be done by first decomposing the FWI gradient into forward- and back-scattered gradients, and then applying the proper weights to combine the back-scattered FWI gradient with the WEMVA gradient. These weights aim to enhance the components of the FWI gradients that overlap with the WEMVA gradient. Preliminary results show that the combined gradient converges faster and to a better solution.

INTRODUCTION

Seismic velocity-analysis methods can be divided into two major groups. First, there are techniques that aim at minimizing the misfit in the data domain, such as full-waveform inversion (Tarantola, 1984; Luo and Schuster, 1991). Second, there are other techniques that aim at improving the quality in the image domain, such as migration velocity analysis (MVA) (Symes and Carazzone, 1991; Biondi and Sava, 1999; Shen, 2004; Biondi, 2010). These techniques try to measure the quality of the image in several ways and then invert the estimated image perturbation using a linearized wave-equation operator.

There are several advantages to minimizing the residual in the image-space, such as increasing the signal-to-noise ratio and decreasing the complexity of the data (Tang et al., 2008). However, a common drawback in doing velocity analysis in the image domain is that only forward-scattered wavefields are used. This results in a loss of vertical resolution in the estimated model updates.

On the other hand, full-waveform inversion (FWI) does not have that problem, since it utilizes the information from both the forward-scattered and back-scattered wavefields, i.e. both the kinematics and the dynamics. This results in higher resolution in the model estimates. However, FWI has the disadvantage of being highly nonlinear, which requires the starting model to be very close to the true model to avoid converging to local minima.

A straightforward solution is to first invert for the velocity model using MVA techniques and then use the output as the initial model for FWI. However, this practice might not work

if the results of MVA are not accurate enough for FWI to start. This could be a result of the larger null space that forward-scattered wavefields do not constrain. Moreover, the convergence rate of the MVA techniques is going to be sub-optimal, since they do not use all of the information in the data.

In this paper, I try to improve the MVA results and convergence rate by supplying the back-scattered information to the gradient. First, I decompose the FWI gradient into two components: the forward-scattering gradient and the back-scattering gradient. The decomposition is done in Fourier domain based on the direction of wave propagation. Second, I develop a new gradient that combines the MVA gradients with the back-scattered component of the FWI gradient using the proper weighting function. This weighting function aims to emphasize the components of the back scattered FWI gradients that overlap with the MVA gradient. Finally, I apply the combined gradient in a synthetic example.

METHOD

When linearized over slowness, the FWI gradient can be written as follows (Tarantola, 1984):

$$g(\mathbf{x}) = \sum_{\omega,\mathbf{x}_s,\mathbf{x}_r} -2\omega^2 s_0(\mathbf{x}) G^*(\mathbf{x},\mathbf{x}_s,\omega) G^*(\mathbf{x},\mathbf{x}_r,\omega) \Delta d(\mathbf{x}_r,\mathbf{x}_s,\omega), \tag{1}$$

where G is the Green's function, $\mathbf{s}_0$ is the background slowness, Δd is the data residual, $\mathbf{x}_s$ and $\mathbf{x}_r$ are the source and receiver coordinates, $\mathbf{x}$ is the Green's functions' coordinate, and ω is frequency. FWI updates contain two components: a tomographic component and a migration component. To better understand these components, we will look at the migration and tomographic operators.

The operator relating the slowness model and the image is nonlinear. Therefore, the first step in evaluating a tomographic operator is to linearize the image $\mathbf{I}$ around the background slowness $\mathbf{s}_0$, as follows:

$$\mathbf{I} = \mathbf{I}_0 + \frac{\partial \mathbf{I}}{\partial \mathbf{s}}|_{s_0}(\mathbf{s}-\mathbf{s}_0) + ..., \tag{2}$$

where $\mathbf{I}_0$ is the background image. Under the Born approximation, the background image can be obtained as follows:

$$\mathbf{I}_0 = \mathbf{L}^*\mathbf{d}, \tag{3}$$

where $\mathbf{L}$ is the Born modeling operator, and its adjoint $\mathbf{L}^*$ is the migration operator. By neglecting the higher-order terms in the image series, we can define the tomographic operator as follows:

$$\Delta\mathbf{I} = \frac{\partial \mathbf{I}}{\partial \mathbf{s}}|_{s_0}\Delta\mathbf{s} = \mathbf{T}\Delta\mathbf{s}, \tag{4}$$

where $\mathbf{T}$ is the tomographic operator. Now, we use the conventional imaging condition as follows:

$$I(\mathbf{x},\mathbf{h}) = \sum_{\omega,\mathbf{x}_s,\mathbf{x}_r} G^*(\mathbf{x}-\mathbf{h},\mathbf{x}_s,\omega) G^*(\mathbf{x}+\mathbf{h},\mathbf{x}_r,\omega) d(\mathbf{x}_r,\mathbf{x}_s,\omega), \tag{5}$$

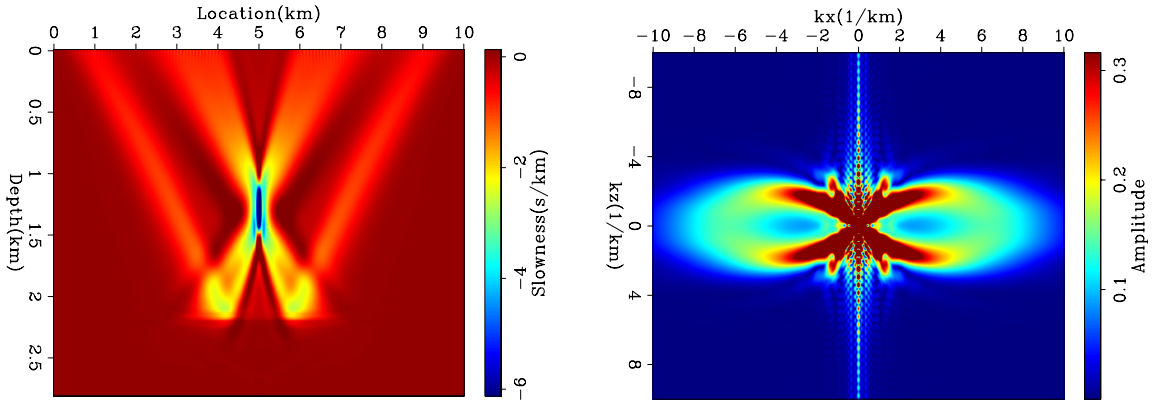
where I is the image, d is the surface data, and $\mathbf{h}$ is the subsurface offset. The background image I_0 can be written as follows:

$$I_0(\mathbf{x},\mathbf{h}) = \sum_{\omega,\mathbf{x}_s,\mathbf{x}_r} G_0^*(\mathbf{x}-\mathbf{h},\mathbf{x}_s,\omega) G_0^*(\mathbf{x}+\mathbf{h},\mathbf{x}_r,\omega) d(\mathbf{x}_r,\mathbf{x}_s,\omega), \tag{6}$$

where G_0 is the background Green's function. After derivation, as shown in Almomin and Tang (2010), we obtain the forward tomographic operator as follows:

$$\begin{aligned}\Delta I(\mathbf{x},\mathbf{h}) &= \sum_{\mathbf{y}} \frac{\partial I(\mathbf{x},\mathbf{h})}{\partial s(\mathbf{y})}\Delta s(\mathbf{y}) \\ &= \sum_{\omega,\mathbf{x}_s,\mathbf{x}_r,\mathbf{y}} \left\{-2\omega^2 s_0(\mathbf{y}) G_0^*(\mathbf{y},\mathbf{x}_s,\omega) G_0^*(\mathbf{x}-\mathbf{h},\mathbf{y},\omega)\right\} G_0^*(\mathbf{x}+\mathbf{h},\mathbf{x}_r,\omega) d(\mathbf{x}_r,\mathbf{x}_s,\omega)\Delta s(\mathbf{y}) \\ &+ \sum_{\omega,\mathbf{x}_s,\mathbf{x}_r,\mathbf{y}} \left\{-2\omega^2 s_0(\mathbf{y}) G_0^*(\mathbf{x}-\mathbf{h},\mathbf{x}_s,\omega) G_0^*(\mathbf{x}+\mathbf{h},\mathbf{y},\omega)\right\} G_0^*(\mathbf{y},\mathbf{x}_r,\omega) d(\mathbf{x}_r,\mathbf{x}_s,\omega)\Delta s(\mathbf{y}),\end{aligned} \tag{7}$$

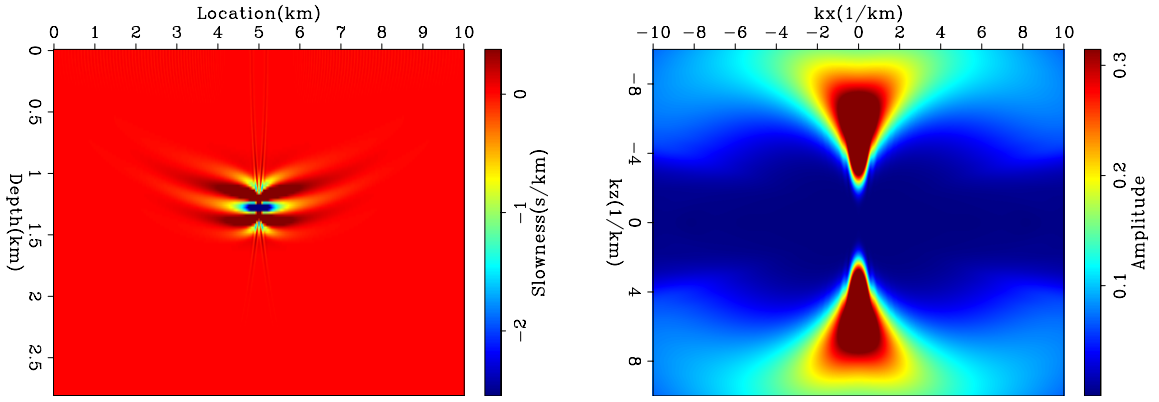
where $\mathbf{y}$ is the slowness coordinate. Notice that all the Green's functions in the migration and tomographic operators are background Green's functions, which means that they do not contain any perturbations. Thus, the migration operator correlates two background wavefields, whereas the tomographic operator correlates a background and a perturbed wavefield. Therefore, a migration component can be generalized as the result of correlating wavefields moving in the opposite direction, whereas a tomographic component is the result of correlating wavefields moving in the same direction. In FWI, the background and perturbed wavefields are summed together in one wavefield. Nonetheless, these components are correlated together to compute the gradient.



(a) (b)

Figure 1: The tomographic operator response (a) and its amplitude spectrum (b). [**CR**] ali2/. wemva26deltaS,wemva26deltaSfk

I now compute the operators' responses in a surface acquisition. The receiver spacing is 20 m, the source spacing is 80 m, and the temporal sampling is 2 ms. Figure 1(a) shows the tomographic operator response by applying its forward and adjoint on a spike. This operator response represents one column of the Hessian matrix of the tomographic operator. Figure 1(b) shows the amplitude spectrum of the tomographic operator response, which shows that the tomographic updates illuminate only the very small vertical wavenumbers. Similarly, figure 2(a) shows the migration operator response of a spike, and figure 2(b) shows its amplitude spectrum. Unlike the tomographic component, the migration component illuminates larger vertical wavenumbers. The illuminated wavenumbers in both operators depend on the reflection angles as well as the frequency content of the data. In this case, the data was created with a Ricker wavelet that has a dominant frequency of 15 Hz.



(a) (b)

Figure 2: The migration operator response (a) and its amplitude spectrum (b). [**CR**] ali2/. wemva27imaged,wemva27imagedfk

Now that I have defined the components in FWI, we can take a closer look at their error sensitivity. The migration operator is linear with respect to model updates. Hence, the migration component in FWI also behaves in a linear sense. This can be better explained by observing that wavefields moving in opposite directions are always going to result in the correct sign of update, as long as the source wavelet is correct. Background velocity changes, or errors, will mostly affect the position at which the wavefields correlate. On the other hand, tomographic components correlate wavefields that are moving in the same direction. For the correlation to have the correct direction, the two wavefields should be within the wavelength of one another. This strict requirement causes the nonlinear behavior of FWI. Since WEMVA techniques do not have such a requirement, they tend to have more stable behavior when the velocity error is large. Our goal is to combine the robust components of FWI and WEMVA.

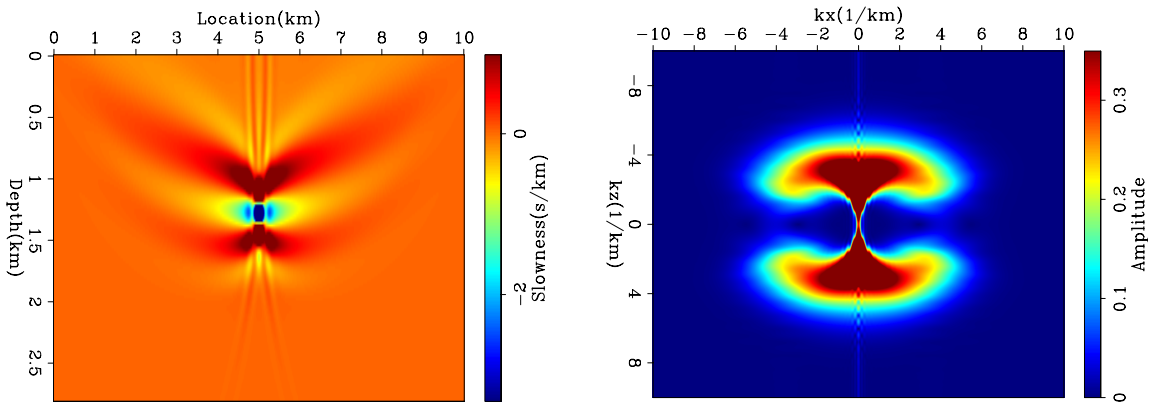
One way to decompose the FWI gradient into its tomographic and migration components is to separate the wavefields based on their direction of propagation, e.g. up or down, which can be done in Fourier domain (Hu and McMechan, 1987; Liu et al., 2007; Taweesintananon, 2011). The decomposed FWI gradient can be written as follows:

$$\begin{aligned} g(\mathbf{x}) = &- \sum_{\omega,\mathbf{x}_s,\mathbf{x}_r} 2\omega^2 s_0(\mathbf{x}) G^{*+}(\mathbf{x},\mathbf{x}_s,\omega) G^{*-}(\mathbf{x},\mathbf{x}_r,\omega) \Delta d(\mathbf{x}_r,\mathbf{x}_s,\omega) \\ &- \sum_{\omega,\mathbf{x}_s,\mathbf{x}_r} 2\omega^2 s_0(\mathbf{x}) G^{*-}(\mathbf{x},\mathbf{x}_s,\omega) G^{*+}(\mathbf{x},\mathbf{x}_r,\omega) \Delta d(\mathbf{x}_r,\mathbf{x}_s,\omega) \\ &- \sum_{\omega,\mathbf{x}_s,\mathbf{x}_r} 2\omega^2 s_0(\mathbf{x}) G^{*+}(\mathbf{x},\mathbf{x}_s,\omega) G^{*+}(\mathbf{x},\mathbf{x}_r,\omega) \Delta d(\mathbf{x}_r,\mathbf{x}_s,\omega) \\ &- \sum_{\omega,\mathbf{x}_s,\mathbf{x}_r} 2\omega^2 s_0(\mathbf{x}) G^{*-}(\mathbf{x},\mathbf{x}_s,\omega) G^{*-}(\mathbf{x},\mathbf{x}_r,\omega) \Delta d(\mathbf{x}_r,\mathbf{x}_s,\omega), \end{aligned} \tag{8}$$

where the positive and negative superscript signs of G indicate the direction of propagation of the wavefield.

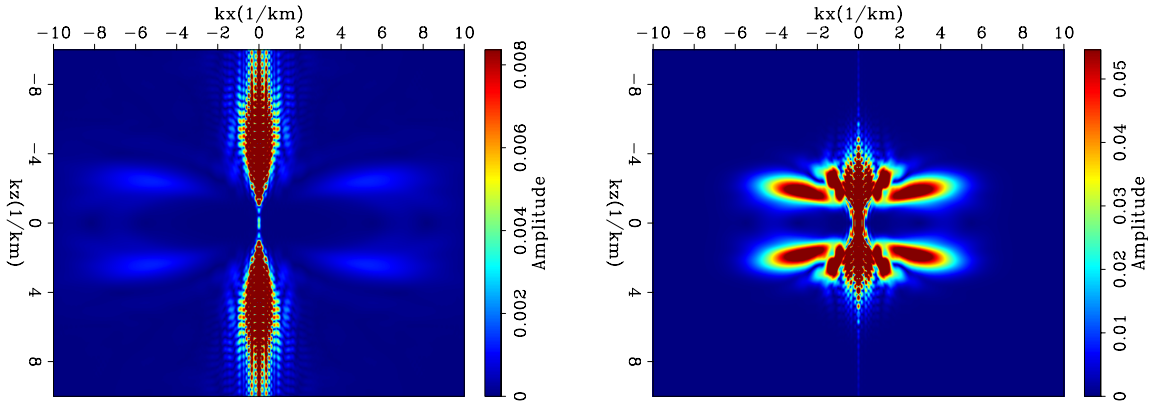
In order for the migration component of FWI gradient to improve the convergence and results of WEMVA gradient, the two need to have some overlap of their illuminated

wavenumbers. Figure 4(a) shows the multiplication of the amplitude spectra in 1(b) and 2(b). The overlap seems to be only at very low horizontal wavenumbers, which have some truncation artifacts. However, the operator's response is a function of frequency. Therefore, by taking lower frequencies when computing the migration component, we may increase the overlap region of the two gradients.



(a) (b)

Figure 3: The low-frequency migration operator response (a) and its amplitude spectrum (b). [**CR**] ali2/. wemva27imagedlf,wemva27imagedlffk



(a) (b)

Figure 4: The spectra product of the tomographic operator with (a) the migration operator and (b) the low-frequency migration operator. [**CR**] ali2/. wemva27deltaSmfk0,wemva27deltaSmfk

Figure 3(a) shows the migration operator response of a spike with a 5 Hz Ricker wavelet, and figure 3(b) shows its amplitude spectrum. With lower frequencies, the migration operator response shifts toward smaller vertical wavenumbers. Figure 4(b) shows the multiplication of the amplitude spectra in 1(b) and 3(b), which shows a much larger and stronger overlap compared to figure 4(a). Figure 5 shows the vertical average migration response at 15 Hz and 5 Hz. It is clear that at low frequencies, the migration update starts to have small vertical-wavenumber contribution.

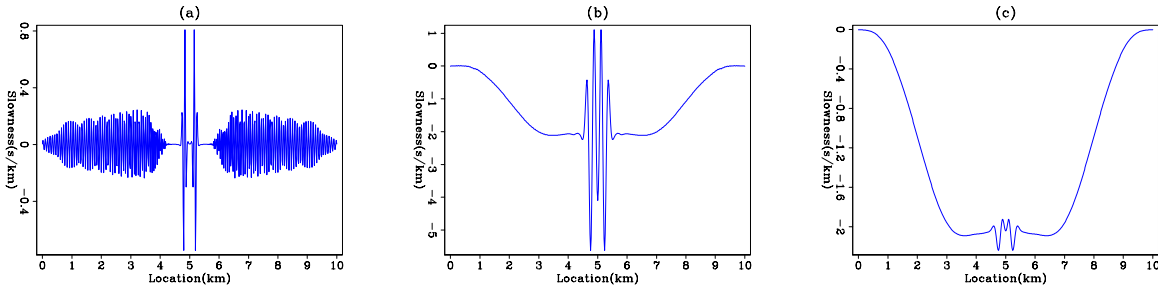


Figure 5: The vertical stack of (a) migration operator response, (b) low-frequency migration operator response and (c) laterally smooth low-frequency migration operator response. [**CR**] ali2/. wemva27stacks

SYNTHETIC EXAMPLES

In this synthetic example, we will compare three inversion results after 20 iterations: low-frequency back-scattered FWI, correlation-based WEMVA (Almomin, 2011), and the combined gradient of WEMVA and the back-scattered FWI weighted by a scalar. The receiver spacing is 20 m, the source spacing is 80 m, and the temporal sampling is 2 ms. Figure 6(a) shows the true slowness anomaly we are inverting for. There is one reflector at the bottom of the slowness model. Figure 6(b) shows the results of low-frequency back-scattered FWI. Although it has the correct anomaly shape, the result has strong side-lobes and the amplitude is still very weak, because convergence is very slow.

Figure 7(a) shows the results of correlation-based WEMVA. As shown earlier, the gradient has low vertical resolution. Also, the anomaly has negative side-lobes. However, the amplitude of the anomaly is much stronger than with back-scattered FWI, and the kinematics of the reflector below converges much faster. Figure 7(b) shows the results of the combined gradient. The amplitude of the anomaly is even stronger and is closer to the true anomaly than is the WEMVA inversion. Furthermore, the side-lobes are much weaker, and the model has much better vertical resolution. Notice that figures 6(b), 7(a) and 7(b) have different color scales.

DISCUSSION AND CONCLUSIONS

I showed that the forward-scattered and back-scattered wavefields have overlap in their estimated slowness models. The synthetic example showed that supplying the back-scattered information can improve the tomographic inversion results in terms of convergence rate and accuracy. The overlap is strongest in the low-frequency part of the data. Therefore, the data must contain low frequencies for the combined gradient to give the best results. Moreover, I used a scalar weight that does not vary with iterations to calculate the combined gradient. Varying the weights with iterations could improve the results even more.

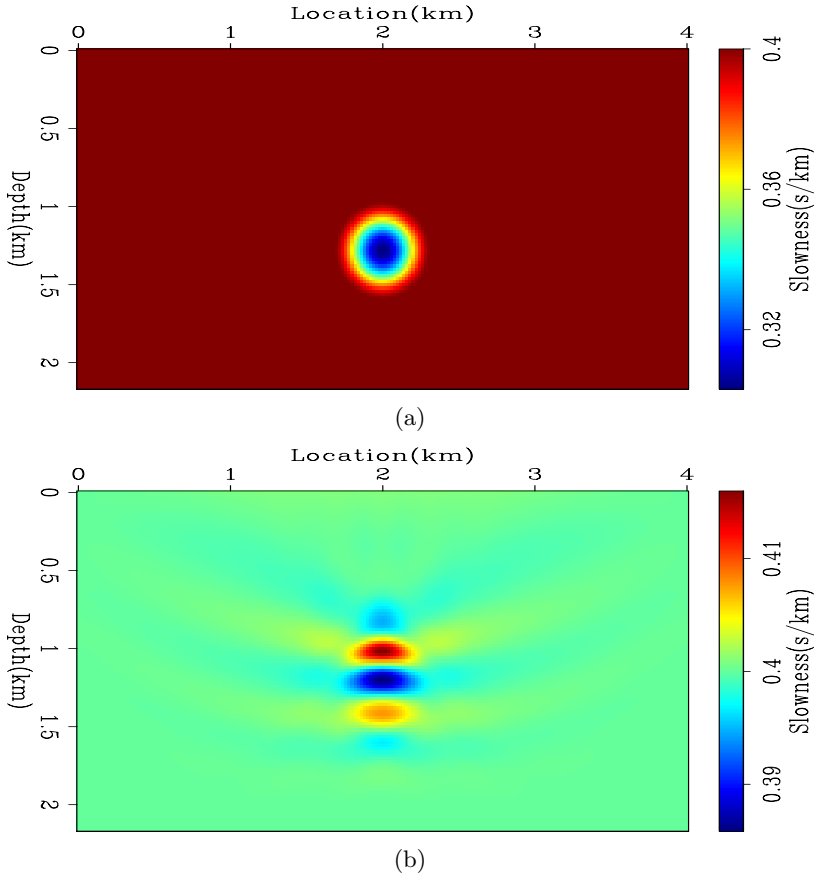


Figure 6: The true slowness model with a cosine perturbation (a) and the inversion results using low-frequency back-scattered FWI (b). [**ER**] [**CR**] ali2/. wemva30tv,wemva30bp

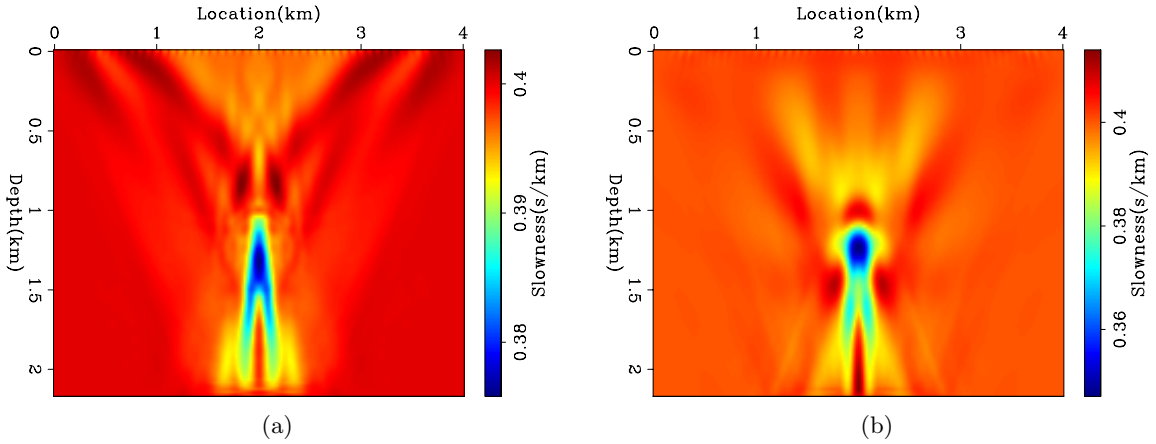


Figure 7: Slowness inversion results using (a) correlation-based WEMVA inversion and (b) the combined gradient. [**CR**] ali2/. wemva30fp,wemva30fbp

REFERENCES

Almomin, A., 2011, Correlation-based wave-equation migration velocity analysis: SEP-Report, **143**, 33–42.

Almomin, A. and Y. Tang, 2010, Migration velocity analysis based on linearization of the two-way wave equation: SEP-Report, **142**, 13–24.

Biondi, B., 2010, Wave-equation migration velocity analysis by residual moveout fitting: SEP-Report, **142**, 25–32.

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.

Hu, L.-Z. and G. A. McMechan, 1987, Wave-field transformations of vertical seismic profiles: Geophysics, **52**, 307–321.

Liu, F., G. Zhang, S. A. Morton, and J. P. Leveille, 2007, Reverse-time migration using one-way wavefield imaging condition: SEG Technical Program Expanded Abstracts, **26**, 2170–2174.

Luo, Y. and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

Shen, P., 2004, Wave equation migration velocity analysis by differential semblance optimization: PhD thesis, Rice University.

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

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

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Taweesintananon, K., 2011, Reverse-time migration using wavefield decomposition: SEP-Report, **143**, 171–190.

Random boundary condition for low-frequency wave propagation

Xukai Shen and Robert G. Clapp

ABSTRACT

In this paper we present a random boundary that is effective at both high-frequencies (Reverse Time Migration) and low-frequencies (waveform inversion). This boundary condition uses larger, irregularly-shaped randomized velocity grains that are effective in introducing incoherency in wavefronts at a large range of wavelengths. We use source functions with a range of peak frequencies to compare this boundary condition to alternative implementations.

INTRODUCTION

A random boundary condition (Clapp, 2009) is attractive in Reverse Time Migration (RTM) imaging, since it generates time-reversible wavefields that scatter incoherently from boundaries. The advantage of the random boundary condition is that the receiver wavefield can be first propagated from max-time to 0, then simultaneously propagated with the source wavefield from 0, to max-time. This eliminates the need to either store the entire wavefield on disk, or use checkpointing schemes at the cost of an extra propagation. Such reduction in memory requirement is particularly suitable for implementation on novel hardwares such as GPUs (Micikevicius, 2008). The incoherent nature of the reflections off the random boundary creates random correlations when the imaging condition is applied, resulting in little to no degradation of the final image.

In theory, this random boundary condition can be easily extended to wavefields at all temporal frequency ranges. However, due to memory and computational requirements, overly large boundary regions are unfavorable. Therefore, problems arise when we move to low frequencies. In that case, increased spatial wavelengths mandate larger boundary regions if we are to use the same random boundary scheme. For low-frequency applications such as waveform inversion, if we can keep the size of the random boundary approximately the same as that used in RTM, computational efficiency will be greatly improved, especially in 3D by avoiding extra computation of waves propagating in the boundary domain. In this paper, we introduce a scheme of perturbing the grain shape of the random boundary condition to keep the boundary region small but still effective with low frequency wave propagation. The same random boundary is also effective for high-frequency wave propagation. We briefly discuss the theory and illustrate the method using synthetic examples.

BOUNDARY DESIGN

Randomness in a velocity boundary has different effects at different temporal frequencies. Because lower-frequency signals have longer wavelengths, the averaging effect within the

dominant wavelength makes the same random boundary appear less random than at higher frequencies. This can be easily understood by taking the extreme case; zero frequency has infinite wavelength, so it sees a random boundary as a constant velocity, no matter what velocity distribution is used. For a random velocity field to appear random to a low-frequency signal, a coarser "grain" of random velocity anomalies is necessary. If grains are bigger than a single cell, two adjustable parameters become important in determining the effectiveness of such random boundaries.

The first parameter is the size of grains. A simple way to determine the grain size is:

$$l_{\text{grain}} \approx \frac{f_{\text{RTM}}}{f_{\text{lowfreq}}} \tag{1}$$

where l_{grain} is the effective length of the random velocity grain, f_{RTM} is the dominant frequency in the RTM that uses the random boundary condition, and $f_{lowfreq}$ is the dominant low frequency used in modeling. The above equation holds because of the inverse relationship between frequency and wavelength at a given velocity. The second parameter is the shape of grains, the easiest to implement of which is cubic, with side lengths equal to the effective length l_{grain}. Although this works much better than a single-cell random velocity anomaly, its effectiveness is diminished by its regular shape. To further increase the randomness of reflected and scattered wavefields, we propose randomly shaped grains in place of cubic grains. We generate randomly shaped grains by perturbing cubic grains of certain lengths. In this case, the effective length of such a randomly shaped grain is equal the the side length of the cubic grain being perturbed. The perturbed grains will have similar volume or grain size to the cubic grains, but have random shapes that more effectively scatter coherent wavefields. It will be shown next that randomly shaped grains also work well with higher-frequency signals, because the irregular, small-scale features at grain boundaries scatter shorter wavelengths effectively.

EXAMPLES

To illustrate the idea, we show examples of acoustic modeling using constant velocity fields with random boundaries of different "grain" size and shape. Each example is compared with a constant boundary condition, where wavefields propagate through boundary regions that have the same velocity values as the inner region. In addition, we use zero-value boundary conditions when wavefields reach the end of each boundary region. The entire velocity field, including the boundary regions, is 6000 meters in x,y and z, with the same 20-meter spacing used in all directions. Boundary regions are 1000 meters (50 gridpoints) in each direction, and either the velocity is constant throughout the region, or velocity values are randomly assigned different grain sizes. In the inner region of the velocity fields, the velocity is constant at 3 km/s; in the boundary regions the velocity is at least 1.5 km/s and no greater than 4.2 km/s. Two types of grains are used in random boundaries, shown in Figure 1: cubic grains with two different side lengths: 20 m (1 grid point)and 200 m (10 grid point), and randomly shaped grains with effective length of 200 m (10 grid points). 3-D modeling is used for all examples, but for display purposes, we only show 2D slices of 3D velocity and wavefields.

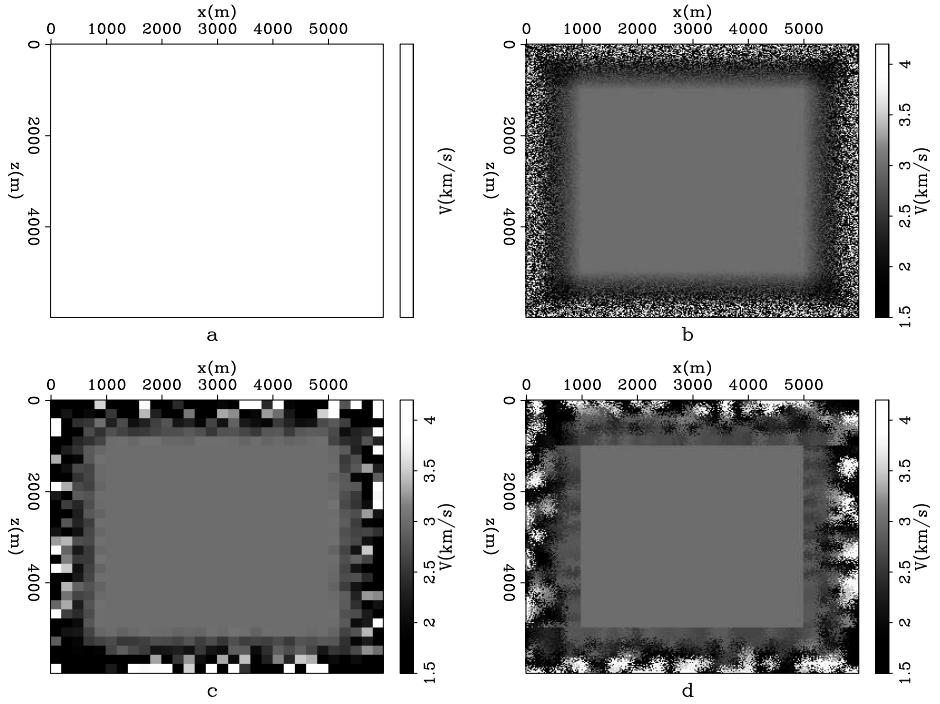


Figure 1: Velocity with different boundary regions: a) constant velocity ; b) cubic grains with 20 m side length ; c) cubic grains with 200 m side length; d) randomly shaped grains with 200 m effective length. [**ER**] xukai2/. velview

Low-frequency modeling

Figure 2 and 3 shows wavefields at 3.5 seconds using a point source with a peak frequency of 6 Hz in velocity fields with the different boundary regions mentioned in Figure 1. The top row of figure 2 shows 2D slices of the wavefield using one realization of each velocity. For a 6 Hz peak frequency (which is typical in waveform inversion), a random boundary with small grain size is only slightly better than constant velocity. In this case, the dominant wavelength is 500 m, and only when grain size becomes comparable to the dominant wavelength does the wavefield start to appear random. However, there are some low-frequency residuals remaining in the images using big cubic grains. The bottom row of figure 2 shows a stacked wavefield of the same source and record time using 16 different realizations of random boundaries. It is now obvious that the low-frequency residuals are evidence of poor performance from using big cubic grains. The regular shape of these cubic grains cannot scatter certain incident wavefronts, thus leaving coherent energy at these angles. Randomly shaped grains, on the other hand, scatter those low-frequency components quite well. This can be seen in the wavenumber domain amplitude spectrum of these wavefields (Figure 3), where the randomly shaped grains reduce lower-wavenumber components much more effectively than cubic grains.

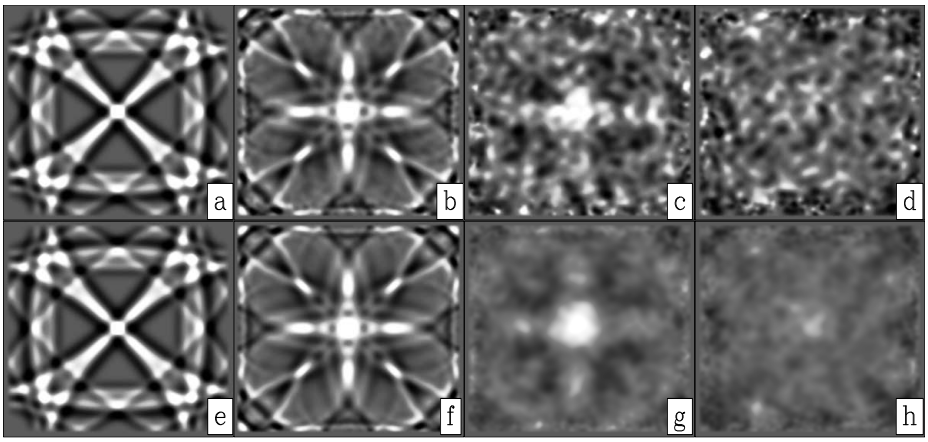


Figure 2: Time-domain wavefield snapshots for a point source with 6 Hz peak frequency in velocity fields with different random boundary conditions. The top row shows one realization of velocity with a) constant velocity, b) cubic grains with 20 m side length, c) cubic grains with 200 m side length, and d) randomly shaped grains with effective length 200 m. The bottom row shows average wavefields using 16 realizations of velocity with e) constant velocity, f) cubic grains with 20 m side length, g) cubic grains with 200 m side length, and h) randomly shaped grains with 200 m effective length. [**CR**] xukai2/. wvmvsixhoriz

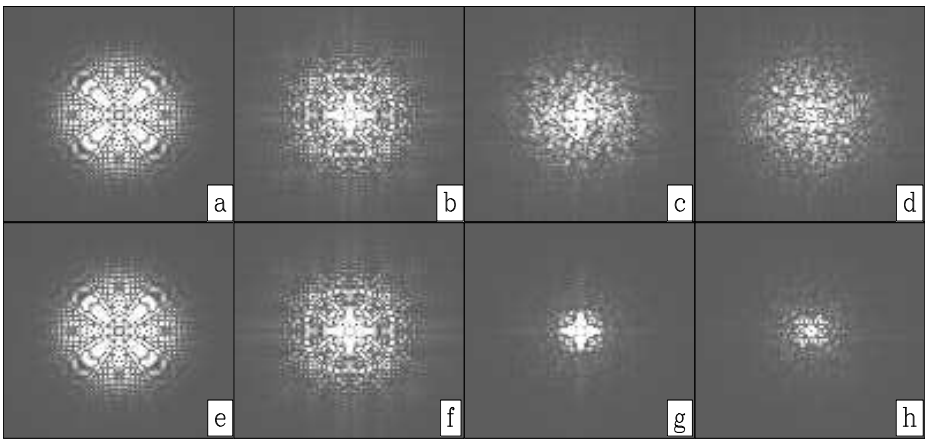


Figure 3: Wave-number-domain amplitude spectrum of wavefield snapshots for a point source with 6 Hz peak frequency in velocity fields with different random boundary conditions. The top row shows one realization of velocity with a) constant velocity, b) cubic grains with 20 m side length, c) cubic grains with 200 m side length, and d) randomly shaped grains with effective length 200 m. The bottom row shows average wavefields using 16 realizations of velocity with e) constant velocity, f) cubic grains with 20 m side length, g) cubic grains with 200 m side length, and h) randomly shaped grains with 200 m effective length. [**CR**] xukai2/. wvmvsixkabshoriz

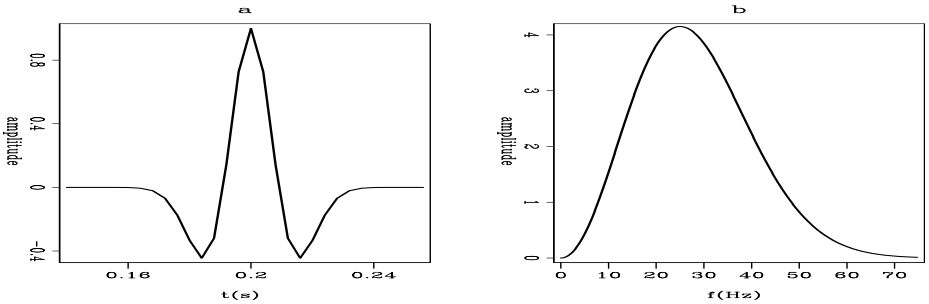


Figure 4: Time-domain and frequency-domain broadband source with a peak frequency of 25 Hz: a) Time-domain wavelet; b) Frequency-domain amplitude spectrum. [**ER**] xukai2/. srctwentyfive

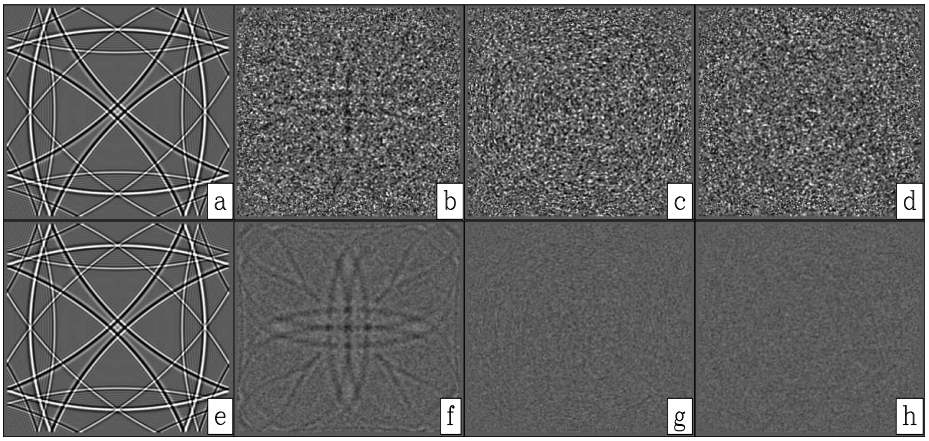


Figure 5: Time domain wavefield snapshots for a broadband point source with a peak frequency of 25 Hz, in velocity fields with different random boundary conditions. The top row shows one realization of velocity with a) constant velocity, b) cubic grains with 20 m side length, c) cubic grains with 200 m side length, and d) randomly shaped grains with effective length 200 m. The bottom row shows average wavefields using 16 realizations of velocity with e) constant velocity, f) cubic grains with 20 m side length, g) cubic grains with 200 m length, and h) randomly shaped grains with effective length 200 m. [**CR**] xukai2/. wvmvtwentyfivehoriz

Broadband modeling application

In this section, we test a broadband point source with a peak frequency of 25 Hz, using the same boundary as in the previous example. Figure 4 shows the amplitude spectrum of the point source used, which contains non-trivial high- and low-frequency components. Figure 5 (top row) shows 2D slices of the wavefield using one realization of each velocity. In this case, the proposed random boundaries still work quite well. A random boundary with a small grain size is effective at high frequency, but has difficulty eliminating low-frequency components of the source. This is more obvious in the bottom row of Figure 5, which is the stacked wavefield of the same source and record time, using 16 different realizations of random boundaries.

Spectra of random boundaries

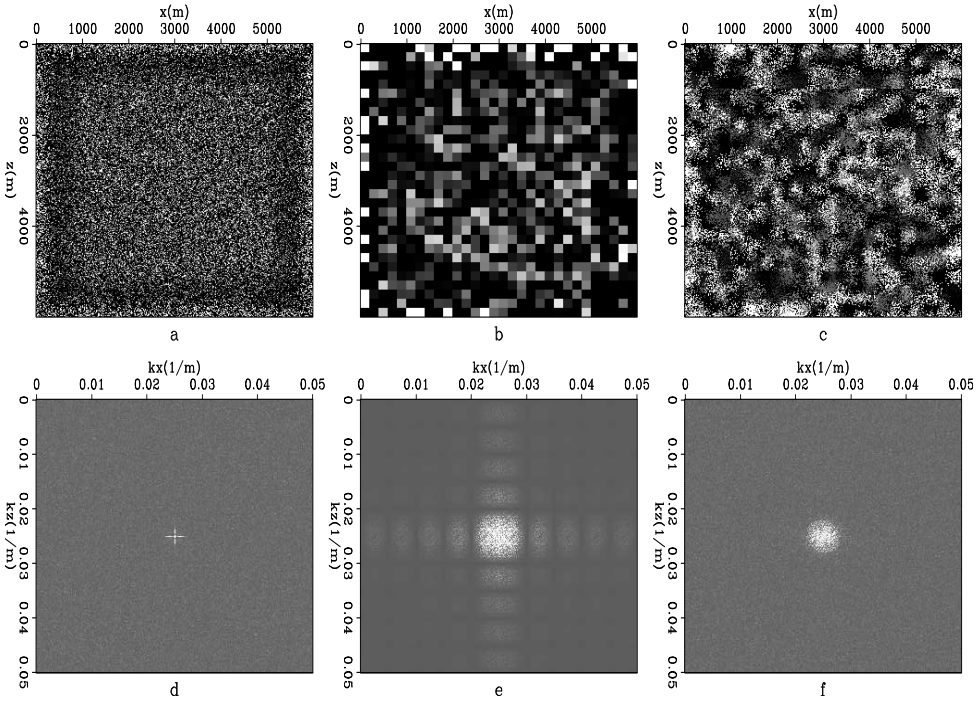


Figure 6: Velocity field with different random boundaries: a) cubic grains with 20 m length; b)cubic grains with 200 m length; and c) randomly shaped grains with 200 m effective length. The bottom row shows their corresponding k spectra: d) cubic grains with 20 m side length, e) cubic grains with 200 m side length, and f) randomly shaped grains with 200 m effective length. [**ER**] xukai2/. velfullrndbnd

Another way to analyze the previous two examples is to consider the spectra of different random boundaries. Figure 6 shows a velocity field filled with different random boundaries. We can look at their k spectra by Fourier transform along both the x and z axes and take the absolute amplitudes (Figure 6). It can be seen that randomly shaped grains at high k values have an amplitude similar to that of single-cell random boundary, and at low k values have an amplitude similar to that of a large cubic-grain random boundary. The large cubic-grain random boundary, on the other hand, has a far lower high-k component, and

there are even notches at certain k values. These features become obvious in stacked k_z and k_x spectra of each random field (Figure 7). Large amplitude at high k values means more detailed randomness, which is useful for scattering high-frequency waves. Large amplitude at low k values means many coarse "grains", which is useful for scattering low-frequency waves. Random fields that have both can effectively scatter broadband signals, and this is the case for randomly shaped grains. The large volume of the grains effectively scatters low-frequency signals, while the edges of grains effectively scatter high-frequency signals. On the other hand, regular notches in the spectrum mean that the random field (cubic grains in this case) displays certain patterns that will render it unable to deal with certain wavefronts; this is analogous to the null space in inversion.

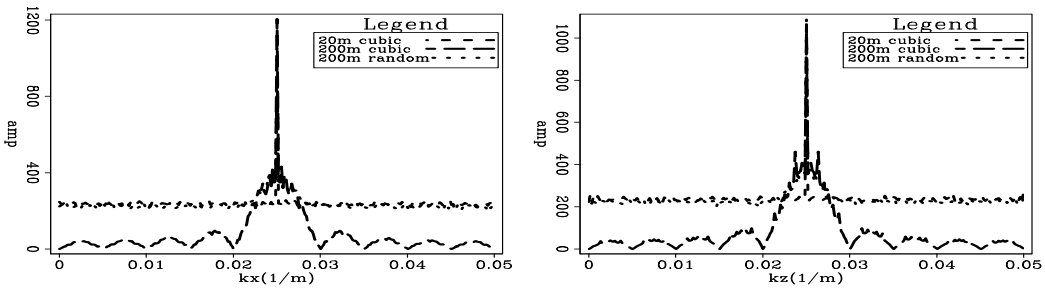


Figure 7: Stacked k_x and k_z spectra of each random field obtained by stacking the amplitude k spectrum along the x and z directions. On the left is the k_x spectrum; on the right is the k_z spectrum. The blue is from cubic grains of side length 20 m, red is from cubic grains of side length 200 m, and pink is from randomly shaped grains of length 200 m. [**ER**] xukai2/. vkabsstk

CONCLUSIONS

Large grain size in a random boundary condition is effective for low-frequency data. Randomly shaped grains work better than regularly shaped grains, and are also very effective in dealing with broadband sources. Of key importance is matching the grain size in the random boundary to the spectrum of the wavefield used in modeling. Applying such frequency-matched boundary conditions extends the utility of the random boundary condition beyond RTM, with possible applications including waveform inversion.

ACKNOWLEDGMENTS

We would like to thank Stanford Exploration Project for the financial support of this research.

REFERENCES

Clapp, R. G., 2009, Reverse time migration with random boundaries: SEG Expanded Abstracts, **28**, 2809–2813.

Micikevicius, P., 2008, 3d finite difference computation on GPUs using CUDA: 2nd Workshop on General Purpose Processing on Graphics Processing Units,Expanded Abstracts.

Subsalt imaging by target-oriented wavefield least-squares migration: A 3-D field-data example

Yaxun Tang and Biondo Biondi

ABSTRACT

We pose the reflectivity-imaging problem as a linear inversion problem and solve it in the image domain in a target-oriented fashion. The most computationally intensive part of the image-domain inversion is the explicit computation of the Hessian matrix. We show how we can overcome the cost issue by using the phase-encoding technique in the 3-D conical-wave domain. We apply the inversion-based imaging methodology to a 3-D field data set acquired from the Gulf of Mexico (GOM), and we precondition the inversion with non-stationary dip filters, which naturally incorporate interpreted geological information. Numerical examples demonstrate that imaging by regularized inversion successfully recovers the reflectivities from the effects of uneven illumination, yielding images with more balanced amplitudes and higher spatial resolution.

INTRODUCTION

Successful geological interpretation requires accurate reflectivity images of the subsurface. Reflectivity images obtained by prestack depth migration, however, are often distorted by uneven subsurface illumination. This is because the migration operator is only the adjoint of the forward Born modeling operator (Lailly, 1983), which is non-unitary due to the limited acquisition geometry, complex overburden and band-limited wavefields. The distorted image, as exemplified by biased amplitudes and the shadow zone effect, presents significant difficulties for accurate geological interpretation.

To correct the effects of uneven illumination, the reflectivity imaging problem can be posed as a linear inverse problem, which, instead of using the adjoint operator, uses the pseudo-inverse of the Born modeling operator to optimally reconstruct the reflectivity. This inversion-based imaging method is also widely known as least-squares migration (Nemeth et al., 1999; Kühl and Sacchi, 2003; Clapp, 2005; Valenciano, 2008).

Least-squares migration can be implemented in either the data domain or the image domain. In this paper, we focus on the image-domain inversion scheme because it can be implemented in a target-oriented fashion and hence is more suitable for large-scale 3-D applications. The target-oriented image-domain formulation allows us to invert only areas of particular interest, enabling accurate imaging at the reservoir level.

As shown by Valenciano (2008), image-domain least-squares migration contains three main steps: (1) compute the migrated image for a chosen target area, (2) compute the Hessian, the normal operator of the Born modeling operator, for the same target area, and (3) deblur the migrated image using the Hessian with an iterative solver. Among the three steps, the explicit computation of the Hessian is the most computationally intensive part,

because it requires either storing a huge number of Green's functions for reuse (Valenciano, 2008), or performing a large number of wavefield propagations. (Each receiver-side Green's function has to be recomputed for each shot, if the Green's functions are not stored) (Tang and Lee, 2010). Fortunately, the computational cost can be drastically reduced by using the phase-encoding method, which does not require storing any Green's functions and significantly reduces the required number of wavefield propagations (Tang, 2009). In this paper, we extend phase-encoding theory to 3-D and show how the cost of Hessian can be drastically reduced by using a simultaneous phase-encoding scheme in the 3-D conical-wave domain.

Regularization is a crucial component of solving an ill-posed inverse problem. One important advantage of the image-domain inversion scheme is that solving the linear inversion problem (step 3) is very fast, which involves only sparse-matrix and vector multiplications. Therefore, different regularization parameters or schemes can be tried at very low cost. The high efficiency of this method also makes interpretation-driven interactive reflectivity imaging possible, where we can repeat the inversion with regularizations that incorporate different geological scenarios and obtain the results in almost real time. In this paper, we precondition the inversion problem with non-stationary dip filters (Clapp, 2003; Hale, 2007; Claerbout, 2008), which impose smoothness on the reflectivities along given dip directions. We show that the dip filter naturally incorporates prior knowledge based on interpreter's geological interpretation into the inversion, and it helps the inversion converge to a geologically meaningful result.

This paper is organized as follows: we first review the theory of image-domain least-squares migration. Then we show how the phase-encoded Hessian can be efficiently computed in the 3-D conical-wave domain. Following that, we discuss how to precondition the inversion with dip filters. Finally, we apply the method to a 3-D field data set acquired from the Gulf of Mexico (GOM), where we invert subsalt reflectivities in a target-oriented fashion.

THEORY

Image-domain least-squares migration

Image-domain least-squares migration (Valenciano, 2008; Tang, 2009) optimizes the reflectivity model by minimizing an objective function defined in the image domain as follows:

$$J(\mathbf{m}) = \frac{1}{2}||\mathbf{H}\mathbf{m} - \mathbf{m}_{\text{mig}}||^2 + \epsilon\mathcal{R}(\mathbf{m}), \tag{1}$$

where $\mathbf{m}$ is the reflectivity model, and $\mathbf{m}_{\text{mig}}$ is the migrated image

$$\mathbf{m}_{\text{mig}} = \mathbf{L}^*\mathbf{d}_{\text{obs}}, \tag{2}$$

where * denotes taking the adjoint, $\mathbf{d}_{\text{obs}}$ is the vector of observed primaries, and $\mathbf{L}$ is the Born modeling operator, which models only singly scattered waves. In equation 1, $\mathbf{H} = \mathbf{L}^*\mathbf{L}$ is the Hessian operator, which contains all necessary information, including information of acquisition geometry, velocity model and frequency content of seismic waves, for correcting the effects of distorted illumination. The second term $\mathcal{R}(\mathbf{m})$ in equation 1 is a regularization term that incorporates user-defined model covariance into the inversion, and ϵ determines the strength of the regularization. Objective function J can be minimized

with any iterative solver, such as the conjugate gradient method. The most important components in minimizing J are the explicit calculation of the Hessian operator $\mathbf{H}$ and the definition of the regularization term $\mathcal{R}(\mathbf{m})$. In the subsequent subsections, we first demonstrate how to calculate the Hessian $\mathbf{H}$ efficiently in 3-D. Then we discuss how to incorporate dip constraints into the inversion and solve it as a preconditioning problem.

The 3-D phase-encoded Hessian

3-D conical-wave migration (Whitmore, 1995; Duquet et al., 2001; Zhang et al., 2005; Liu et al., 2006) has been widely used to migrate marine streamer data sets. In this section, we demonstrate how the Hessian can be efficiently computed in this domain using simultaneous phase encoding, which encodes both source- and receiver-side Green's functions.

As derived in Appendix A, each component of the the 3-D conical-wave domain Hessian reads

$$\begin{aligned} H(\mathbf{x},\mathbf{x}') &= \sum_{\omega} |\omega|^5 \sum_{y_s} \sum_{p_{s_x}} \sum_{\mathbf{x}_r} G(\mathbf{x},\mathbf{x}_r,\omega) G^*(\mathbf{x}',\mathbf{x}_r,\omega) \\ &\quad \sum_{x_s} W(\mathbf{x}_r,x_s,y_s) f_s(\omega) G(\mathbf{x},x_s,y_s,\omega) \mathrm{e}^{i\omega p_{s_x} x_s} \\ &\quad \sum_{x'_s} W(\mathbf{x}_r,x'_s,y_s) f^*_s(\omega) G^*(\mathbf{x}',x'_s,y_s,\omega) \mathrm{e}^{-i\omega p_{s_x} x'_s}, \end{aligned} \tag{3}$$

where $f_s(\omega)$ is the source signature at frequency ω; x_s and y_s are the source locations in the inline and crossline directions, respectively; $\mathbf{x}_r = (x_r, y_r, 0)$ is the receiver location; p_{s_x} is the horizontal component of the source ray parameter; $W(\mathbf{x}_r, x_s, y_s)$ is the acquisition mask operator, which contains ones where we record data, and zeros where we do not; and $G(\mathbf{x}, x_s, y_s, \omega)$ and $G(\mathbf{x}, \mathbf{x}_r, \omega)$ are the Green's functions connecting the source and receiver positions to the image point $\mathbf{x} = (x, y, z)$, respectively.

The diagonal part of the Hessian (when $\mathbf{x} = \mathbf{x}'$), which contains autocorrelations of both source and receiver-side Green's functions, can be interpreted as a subsurface illumination map with contributions from both sources and receivers. The rows of the Hessian (for fixed $\mathbf{x}$'s and varying $\mathbf{x}'$), which contains crosscorrelations of both source and receiver-side Green's functions, can be interpreted as resolution functions (Lecomte, 2008; Tang, 2009). They measure how much smearing an image can have due to a given acquisition setup.

The exact Hessian defined by equation 3, however, is nontrivial and very expensive to implement. It requires either storing a huge number of Green's functions for reuse, or performing a large number of wavefield propagations to repeatedly calculate the Green's functions, resulting in a computational cost proportional to $N_{y_s} N_{p_{s_x}} N_{x_r} N_{y_r}$, with N_{y_s}, $N_{p_{s_x}}$, N_{x_r} and N_{y_r} being the number of crossline shots, inline conical waves, inline receivers and crossline receivers, respectively.

In order to reduce the computational cost, we use the simultaneous phase-encoding technique to efficiently calculate an approximate version of equation 3. The simultaneous phase-encoding, however, is only strictly valid when the acquisition mask operator is independent along the encoding axes (Tang, 2009). For the 3-D conical-wave-domain Hessian, the encoding axes are the inline source axis x_s and the receiver axis $\mathbf{x}_r = (x_r, y_r)$,

respectively. Ocean-bottom cable (OBC) and land acquisition geometries, where receivers are fixed for all sources, obviously satisfy this condition. But marine-streamer acquisition geometry, where the receiver cable moves with sources, apparently does not. To make the theory applicable to the marine-streamer data case, we assume that the receiver location $\mathbf{x}_r$ depends only on the crossline source position y_s, but is independent of the inline source position x_s. This implicitly assumes that for a fixed crossline y_s, all inline shots share the same receiver array. Therefore, we can express the acquisition mask operator as a product of two separate functions:

$$W(\mathbf{x}_r, x_s, y_s) \approx W_r(\mathbf{x}_r, y_s) W_s(x_s, y_s), \tag{4}$$

where W_r and W_s define the distributions of receiver position $\mathbf{x}_r$ and the inline source position x_s, respectively, for a given crossline source position y_s.

Substituting equation 4 into equation 3 yields

$$\begin{aligned} H(\mathbf{x}, \mathbf{x}') &= \sum_{\omega} |\omega|^5 \sum_{y_s} \sum_{p_{s_x}} \sum_{\mathbf{x}_r} W_r(\mathbf{x}_r, y_s) G(\mathbf{x}, \mathbf{x}_r, \omega) G^*(\mathbf{x}', \mathbf{x}_r, \omega) \\ &\quad \sum_{x_s} W_s(x_s, y_s) f_s(\omega) G(\mathbf{x}, x_s, y_s, \omega) \mathrm{e}^{i\omega p_{s_x} x_s} \\ &\quad \sum_{x'_s} W_s(x'_s, y_s) f_s^*(\omega) G^*(\mathbf{x}', x'_s, y_s, \omega) \mathrm{e}^{-i\omega p_{s_x} x'_s}. \end{aligned} \tag{5}$$

With further encoding on the receiver-side Green's functions, we obtain the simultaneously phase-encoded Hessian as follows:

$$\begin{aligned} \widetilde{\widetilde{H}}(\mathbf{x}, \mathbf{x}', \mathbf{p}_r) &= \sum_{\omega} |\omega|^5 \sum_{y_s} \sum_{p_{s_x}} \\ &\quad \times \sum_{\mathbf{x}_r} W_r(\mathbf{x}_r, y_s) G(\mathbf{x}, \mathbf{x}_r, \omega) \alpha(\mathbf{x}_r, \mathbf{p}_r, \omega) \\ &\quad \times \sum_{\mathbf{x}'_r} W_r(\mathbf{x}'_r, y_s) G^*(\mathbf{x}', \mathbf{x}'_r, \omega) \alpha(\mathbf{x}'_r, \mathbf{p}_r, \omega) \\ &\quad \times \sum_{x_s} W_s(x_s, y_s) f_s(\omega) G(\mathbf{x}, x_s, y_s, \omega) \mathrm{e}^{i\omega p_{s_x} x_s} \\ &\quad \times \sum_{x'_s} W_s(x'_s, y_s) f_s^*(\omega) G^*(\mathbf{x}', x'_s, y_s, \omega) \mathrm{e}^{-i\omega p_{s_x} x'_s}, \end{aligned} \tag{6}$$

where α is the receiver-side encoding function, to be specified later. Equation 6 can be greatly simplified as follows:

$$\begin{aligned} \widetilde{\widetilde{H}}(\mathbf{x}, \mathbf{x}', \mathbf{p}_r) &= \sum_{\omega} |\omega|^5 \sum_{y_s} \sum_{p_{s_x}} S(\mathbf{x}, p_{s_x}, y_s, \omega) S^*(\mathbf{x}', p_{s_x}, y_s, \omega) \\ &\quad R(\mathbf{x}, p_{s_x}, y_s, \mathbf{p}_r, \omega) R^*(\mathbf{x}', p_{s_x}, y_s, \mathbf{p}_r, \omega), \end{aligned} \tag{7}$$

if we define

$$S(\mathbf{x}, p_{s_x}, y_s, \omega) = \sum_{x_s} W_s(x_s, y_s) f_s(\omega) G(\mathbf{x}, x_s, y_s, \omega) \mathrm{e}^{i\omega p_{s_x} x_s}, \tag{8}$$

and

$$R(\mathbf{x}, p_{s_x}, y_s, \mathbf{p}_r, \omega) = \sum_{\mathbf{x}_r} W_r(\mathbf{x}_r, y_s) G(\mathbf{x}, \mathbf{x}_r, \omega) \alpha(\mathbf{x}_r, \mathbf{p}_r, \omega). \tag{9}$$

For one-way wave-equation-based applications, S and R can be obtained by solving the following one-way wave equations:

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\frac{\omega^2}{v^2(\mathbf{x})} + \nabla^2}\right) S(\mathbf{x}, p_{s_x}, y_s, \omega) = 0 \\ S(x, y, z=0, p_{s_x}, y_s, \omega) = \sum_{x_s} W_s(x_s, y_s)\delta(x - x_s, y - y_s) f_s(\omega) \mathrm{e}^{i\omega p_{s_x} x_s} \end{cases}. \tag{10}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\frac{\omega^2}{v^2(\mathbf{x})} + \nabla^2}\right) R(\mathbf{x}, p_{s_x}, y_s, \mathbf{p}_r, \omega) = 0 \\ R(x, y, z=0, p_{s_x}, y_s, \mathbf{p}_r, \omega) = \sum_{\mathbf{x}_r} W_r(\mathbf{x}_r, y_s)\delta(x - x_r, y - y_r)\alpha(\mathbf{x}_r, \mathbf{p}_r, \omega) \end{cases}. \tag{11}$$

In both equations 10 and 11, $v(\mathbf{x})$ is the velocity at image point $\mathbf{x}$, $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the Laplacian operator, and $\delta(\cdot)$ is the Dirac delta function. Therefore, S is the wavefield generated by propagating the conical-wave source $\sum_{x_s} W_s(x_s, y_s)\delta(x - x_s, y - y_s) f_s(\omega) \mathrm{e}^{i\omega p_{s_x} x_s}$, whereas R is the wavefield generated by propagating the encoded-area source $\sum_{\mathbf{x}_r} W_r(\mathbf{x}_r, y_s)\delta(x - x_r, y - y_r)\alpha(\mathbf{x}_r, \mathbf{p}_r, \omega)$. It is quite obvious that the computational cost of equation 7 is independent of the number of receivers, as opposed to equation 3, for which the cost is proportional to the number of receivers.

However, the phase-encoded Hessian brings unwanted crosstalk. This becomes clear by comparing equations 6 and 3. The crosstalk can be suppressed by carefully choosing the phase-encoding function α (Tang, 2009). In this paper, we choose α to be a random phase-encoding function; thus $\mathbf{p}_r$ denotes the realization index of the random phase-encoding function. It would be very easy to verify that, with this choice of encoding functions, the expectation of the crosstalk becomes zero. Therefore, equation 6 converges to equation 5 by stacking over $\mathbf{p}_r$, according to the law of large numbers (Gray and Davisson, 2003):

$$\widetilde{\widetilde{H}}(\mathbf{x}, \mathbf{x}') = \sum_{\mathbf{p}_r} \widetilde{\widetilde{H}}(\mathbf{x}, \mathbf{x}', \mathbf{p}_r) \approx H(\mathbf{x}, \mathbf{x}'). \tag{12}$$

For most practical applications where the number of shots is big, the randomly phase-encoded Hessian with one realization seems to be sufficient (Tang, 2009).

Regularizations

Regularization helps to stabilize the inversion; it can shape the null space and remove unwanted features in the inverted result by introducing user-defined model-covariance operators. In this paper, we choose to use the following regularization term:

$$\mathcal{R}(\mathbf{m}) = \frac{1}{2}||\mathbf{D}^*\mathbf{D}\mathbf{m}||^2, \tag{13}$$

where operator $\mathbf{D}$ contains wavekill filters (Claerbout, 2008), which annihilate local planar-events with given dips. The operator imposes continuity of reflectors along its dipping direction. This idea has also been explored by Clapp (2005) and Ayeni et al. (2009), who

use similar filters (Clapp, 2003; Hale, 2007) to regularize the data-domain least-squares migration.

Instead of solving the inversion problem as a regularization problem, we solve it as a preconditioning problem by making change of variables as follows:

$$\mathbf{m} = \mathbf{S}\mathbf{n}, \tag{14}$$

where $\mathbf{n}$ is the vector of preconditioned variables and $\mathbf{S}$ is the preconditioning operator, which is defined to be an approximate inverse of the regularization operator $\mathbf{D}^*\mathbf{D}$. To find the inverse of $\mathbf{D}^*\mathbf{D}$, we factorize it into minimum-phase filters $\mathbf{A}$ such that $\mathbf{D}^*\mathbf{D} \approx \mathbf{A}^*\mathbf{A}$. We use the Wilson-Burg factorization (Claerbout, 1992; Fomel et al., 2003) and apply it on the helix (Claerbout, 1998, 2008). Since minimum-phase filters have stable inverses, we can define the preconditioning operator as follows:

$$\mathbf{S} = \mathbf{A}^{-1} \left(\mathbf{A}^*\right)^{-1}. \tag{15}$$

Unlike $\mathbf{D}^*\mathbf{D}$, operator $\mathbf{S}$ contains dip filters, which smooth along given dip directions. Substituting equations 13, 14 and 15 into 1 yields

$$J_p(\mathbf{n}) = \frac{1}{2}||\mathbf{HSn} - \mathbf{m}_{\text{mig}}||^2 + \epsilon||\mathbf{n}||^2. \tag{16}$$

Objective function 16 is often solved by setting $\epsilon = 0$ and iterating until an acceptable result is obtained (Claerbout, 2008). Solving it in this way implicitly assumes that we are starting with a model that has all the user-defined covariance, and that the more iterations we run, the more we honor the data. Once a solution vector $\mathbf{n}_{\text{sol}}$ has been found, the final model is obtained by computing $\mathbf{m}_{\text{sol}} = \mathbf{S}\mathbf{n}_{\text{sol}}$.

3-D FIELD-DATA EXAMPLES

We apply the target-oriented inversion method to a data set acquired from the Gulf of Mexico (GOM). The data set was acquired using a narrow-azimuth towed-streamer (NATS) acquisition system, and further rotated using azimuth moveout (AMO) (Biondi et al., 1998) into zero azimuth. The minimum and maximum inline offsets are 0.3 km and 8.2 km, respectively. The frequency content of the data set ranges between 5 Hz and 35 Hz. The velocity model used for migration and Hessian computation is shown in Figure 1, which is obtained using target-oriented wavefield tomography (Tang and Biondi, 2011).

We compute the migrated image using the 3-D conincal-wave migration operator, where we synthesize 101 conical waves for each crossline and migrate 12625 conical waves in total. The minimum and maximum inline take-off angles at the surface for the conical waves are -30° and 30°, respectively. The maximum frequency used for migration is 20 Hz. The image obtained for the target area is shown in Figure 2. Note that the amplitudes of the sediment reflectors are biased; also notice the illumination shadows below the salt due to the non-unitary characteristic of the Born modeling operator.

The 3-D Hessian matrix

We compute the phase-encoded Hessian using equations 6, 10 and 11. Since the number of shots is big, we compute only one random realization of the phase-encoded Hessian. The

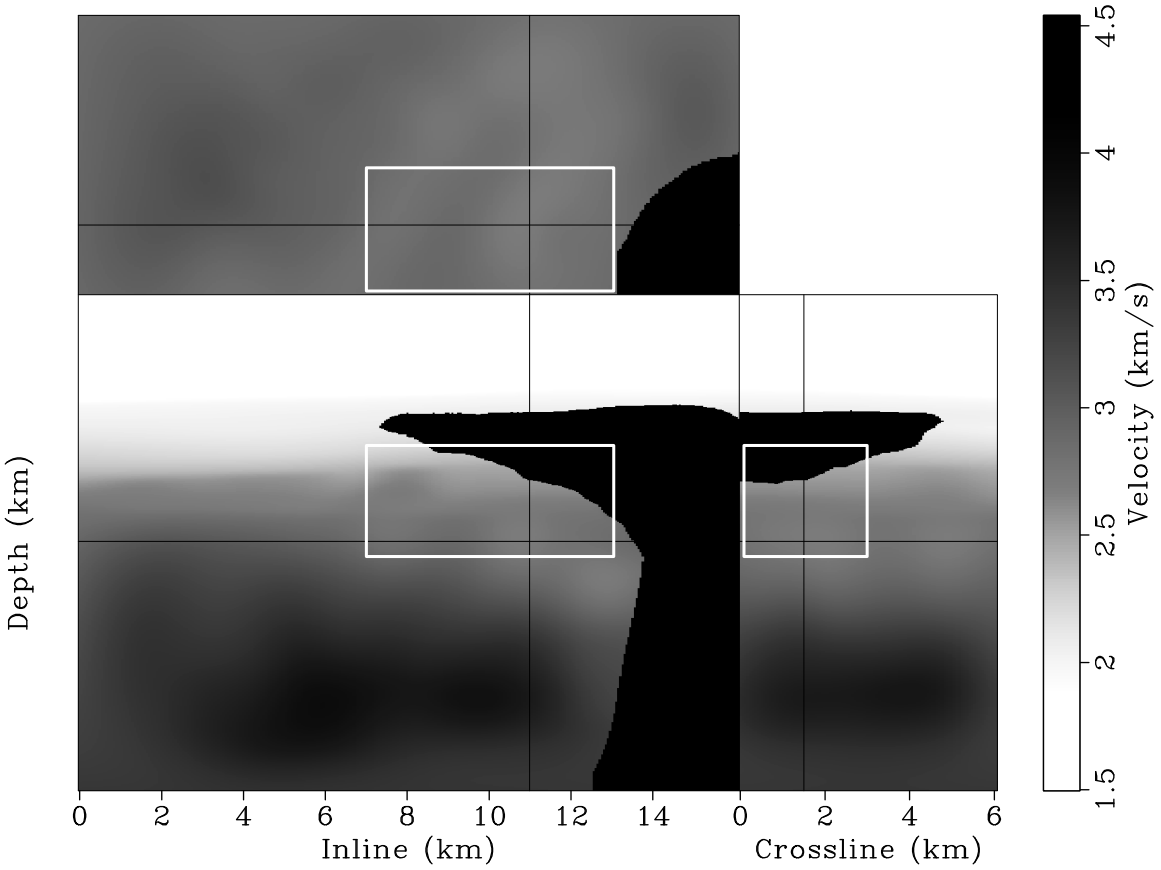


Figure 1: Target area selected (outlined by a box) for wavefield least-squares migration. [**CR**] yaxun2/. lsm3d-target

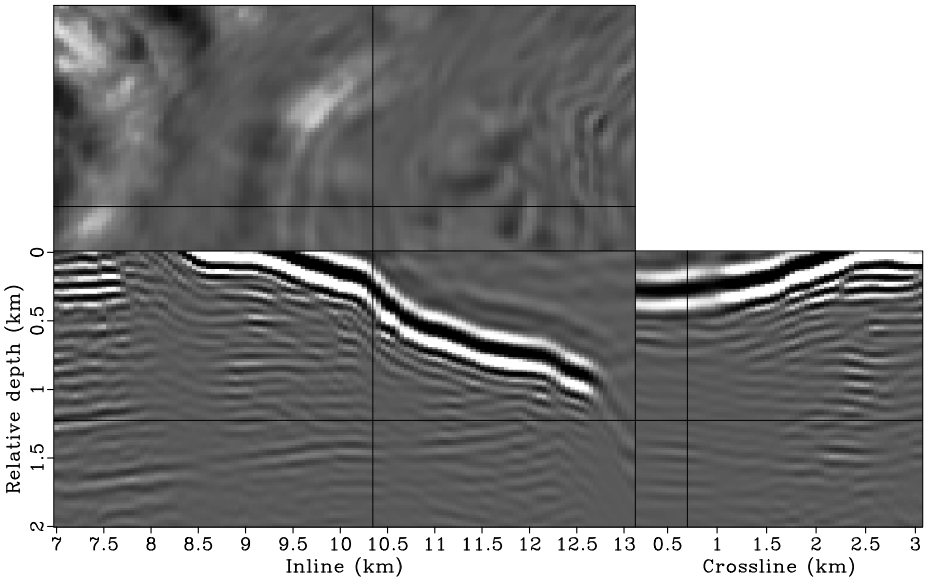


Figure 2: Migrated image for the selected target region. Note the illumination shadows below the salt. [**CR**] yaxun2/. lsm3d-imag

target region contains 784740 points, with 123 samples inline, 58 crossline and 110 in depth. The number of elements computed per row for the Hessian is 5155 (11 in x, 15 in y and 31 in z). Therefore, the widths of the local filter for each image point are 0.25 km, 0.35 km and 0.27 km in x, y and z directions, respectively. The top panels in Figures 3 and 4 present the diagonal components of the Hessian matrix at two different slices. Note that the values of the diagonal of the Hessian are far from uniform, and the left-side values are much higher than those elsewhere in the target region. This is because the salt body, which has relatively high velocities, prevents most of the energy from penetrating itself. The unevenness of the diagonal components also suggests that the Hessian matrix is highly nonstationary (each row is substantially different than the others). This is further illustrated by the bottom panels in Figures 3 and 4, which show the off-diagonal components of the Hessian matrix at two different image points. For an image point that is well illuminated (bottom panels in Figure 3), the off-diagonals have relatively wide spectrum coverage and are more focused around the diagonal. On the other hand, for an image point that is poorly illuminated (bottom panels in Figure 4), the off-diagonals have relatively narrow spectrum coverage and are more spread in the space domain.

To further appreciate the nonstationarities of the 3-D Hessian matrix, we apply the Hessian to a reflectivity model containing a collection of point scatterers (Figure 5). The result can also be considered as the filter response of the Hessian (where each row can be seen as a filter) to point scatterers. Note how the shape and strength of the filters change across the space. Also note that the filter is more elongated in the crossline direction than in the inline direction. This is a result of the single-azimuth acquisition geometry. Figure 6 shows the Hessian filter response for four horizontal reflectors. Note the imprint of shadow zones on the reflectors. The characteristics of the shadow zones very closely match those in the migrated image (Figure 2), indicating that the computed Hessian matrix, albeit with some approximations, accurately captures the effects of uneven subsurface illumination due to limited acquisition geometry, band-limited wave phenomena and complex overburden. In the subsequent section, I demonstrate how the effects of uneven illumination can be optimally removed by inverting the Hessian matrix through regularized linear inversion.

Inversion result

Since the goal is to invert sediment reflectivities, we incorporate a mask operator (Figure 7) into the inversion to prevent updating the reflectivities of the salt boundary. We first run the inversion without applying any regularizations; the inverted image after 100 iterations is shown in Figure 9. Compared to migration (Figure 8), inversion significantly improves the spatial resolution of the image; the amplitudes are more balanced, and the illumination shadows are filled in. However, the inverted image is more noisy, and the continuity of the reflectors seems to be degraded. The increased noise level reveals the ill-posedness of the inversion problem due to the narrow bandwidth of the Hessian filter (bottom panels of Figures 3 and 4). Another reason might be that the approximations used to compute the Green's function (acoustic one-way wave equation) and the migration do not fully match the way seismic waves propagate through the earth. The inconsistency of the operator and the data may further increase the ill-posedness of the inversion problem. Therefore, regularization becomes necessary.

In order to regularize the inversion with a reasonably accurate model covariance, we

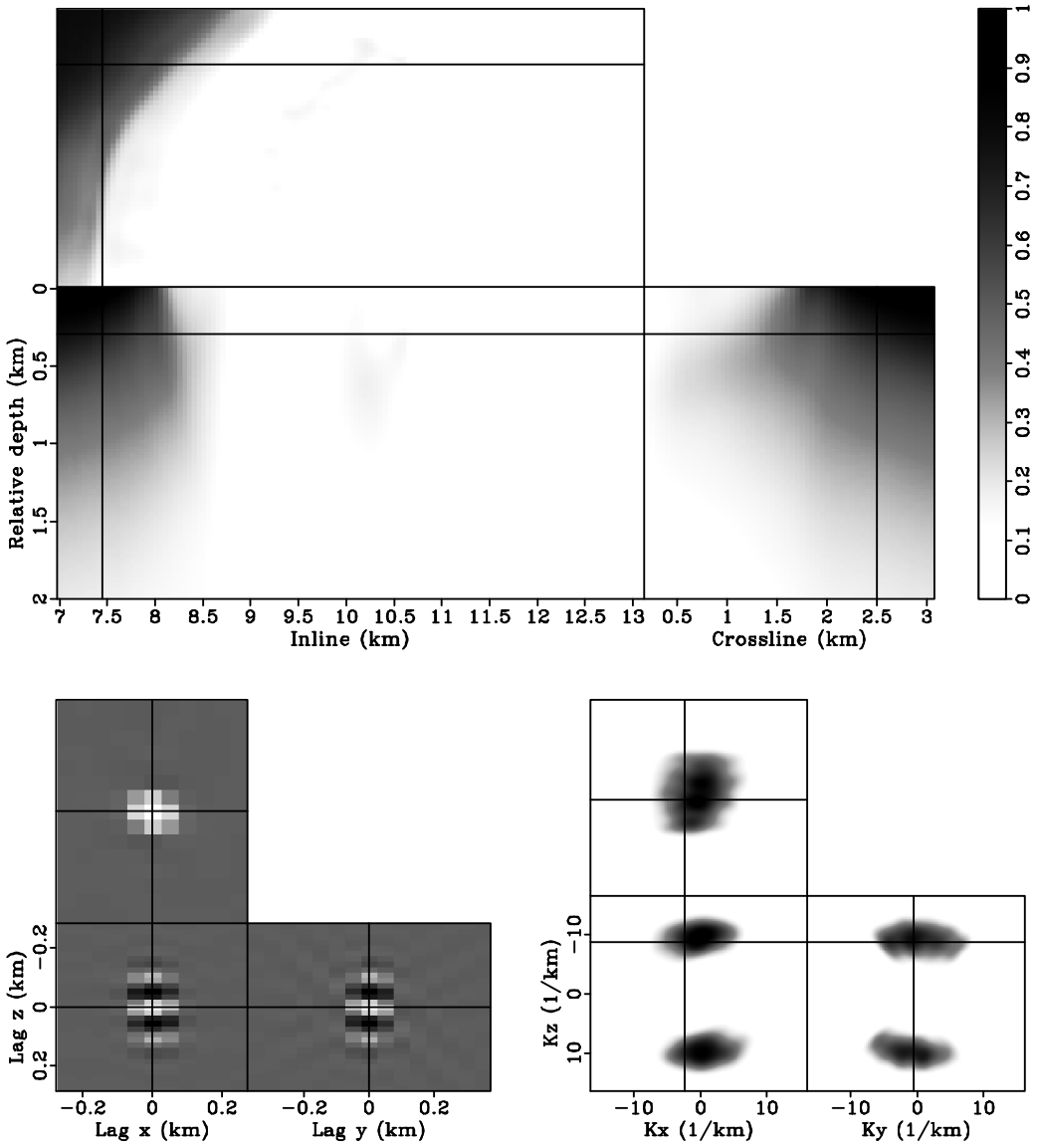


Figure 3: The Hessian matrix for the target region. The top panel shows the diagonal components of the matrix; the bottom left panel shows the off-diagonals of the matrix taken from the image point at inline 7.45 km, crossline 2.50 km and relative depth 0.30 km (the intersection of the crosshairs in the top panel); the bottom right panel shows the amplitude spectrum of the off-diagonal components. [**CR**] yaxun2/. lsm3d-hess-filter1

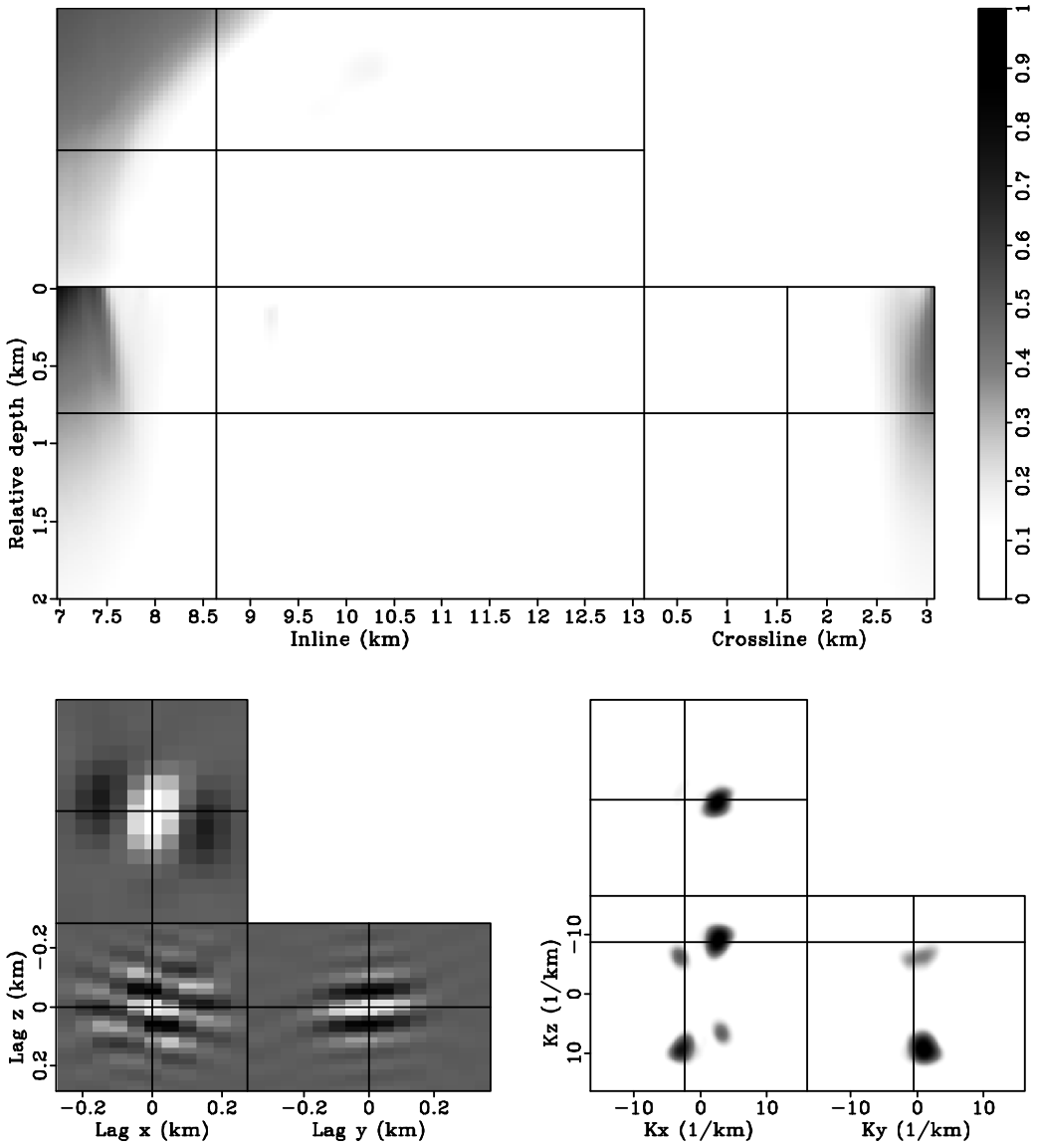
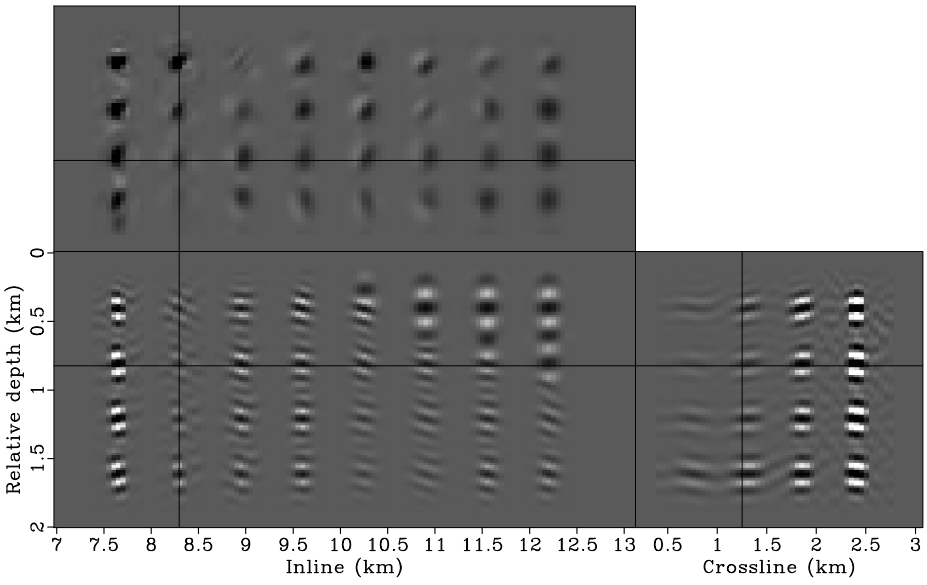
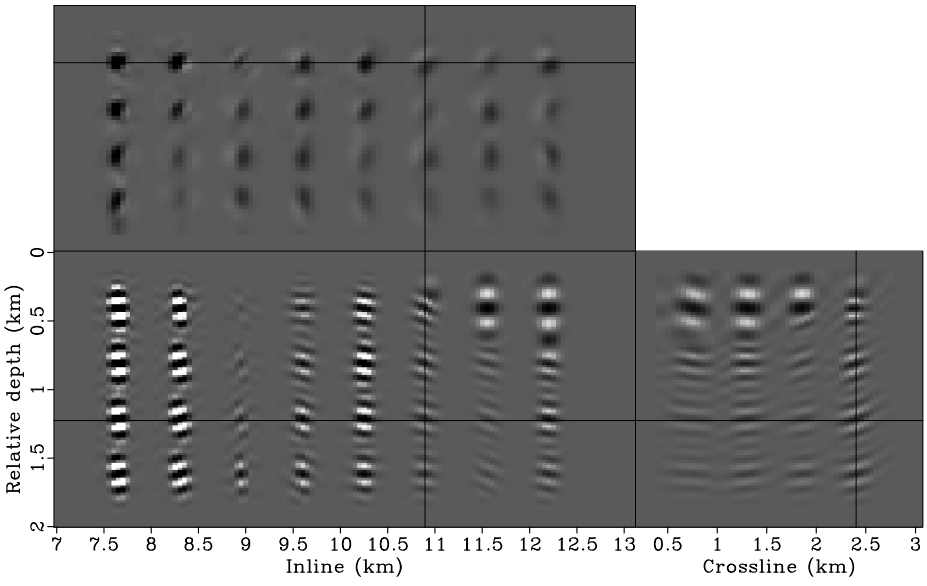


Figure 4: The Hessian matrix for the target region. The top panel shows the diagonal components of the matrix; the bottom left panel shows the off-diagonals of the matrix taken from the image point at inline 8.65 km, crossline 1.60 km and relative depth 0.80 km (the intersection of the crosshairs in the top panel); the bottom right panel shows the amplitude spectrum of the off-diagonal components. [**CR**] yaxun2/. lsm3d-hess-filter2

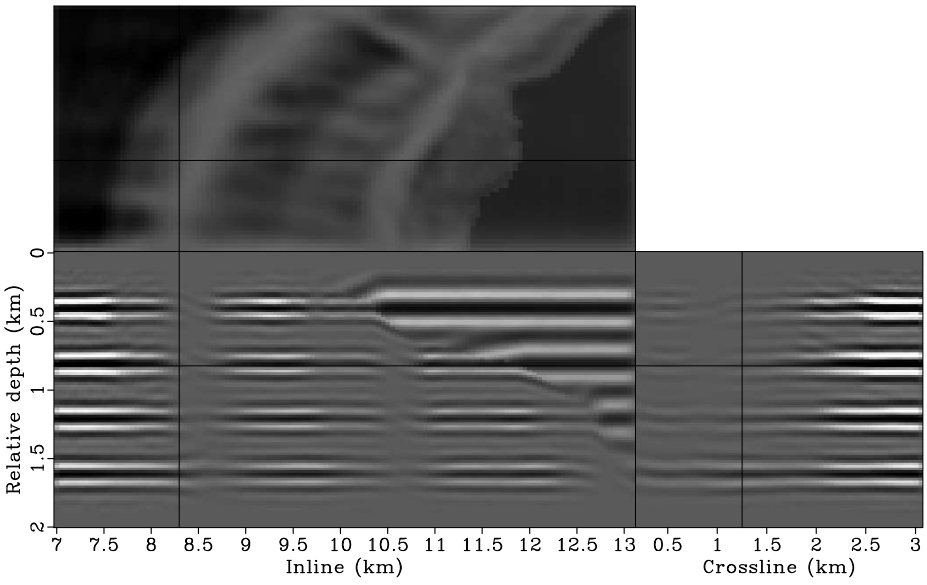


(a)

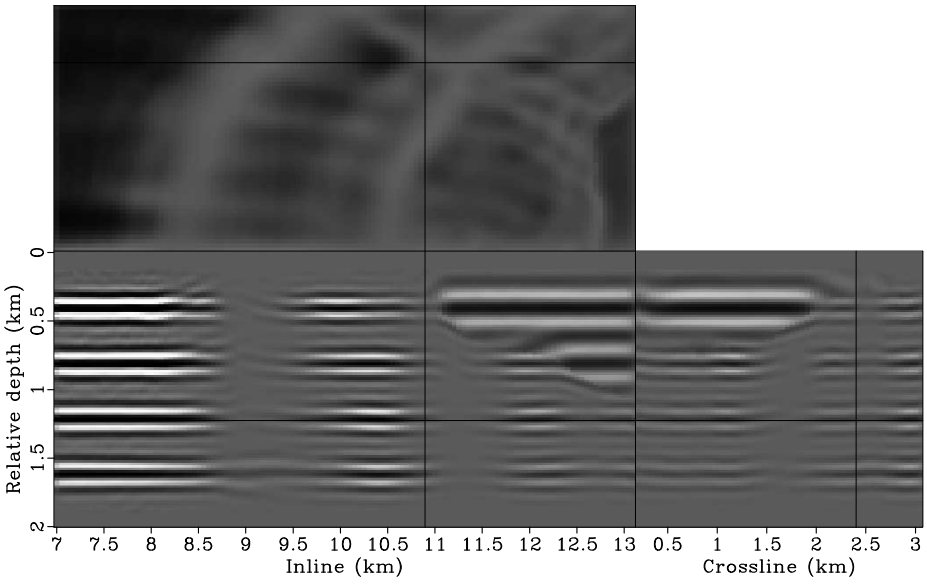


(b)

Figure 5: Hessian filter response for point scatterers. Panels (a) and (b) show different slices of the same 3-D cube. Note the nonstationarity of the filters. [**CR**] yaxun2/. lsm3d-imag-spike1,lsm3d-imag-spike2



(a)



(b)

Figure 6: Hessian filter response for horizontal reflectors. Panels (a) and (b) show different slices of the same 3-D cube. Note the imprint of shadow zones on the reflectors. [**CR**] yaxun2/. lsm3d-imag-flat1,lsm3d-imag-flat2

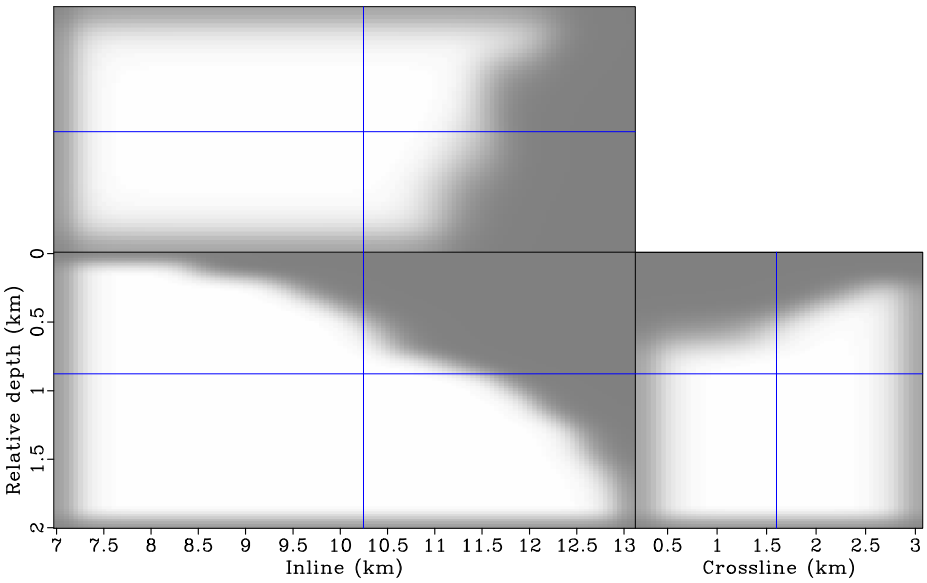


Figure 7: A mask operator used during inversion. [**CR**] yaxun2/. lsm3d-mask

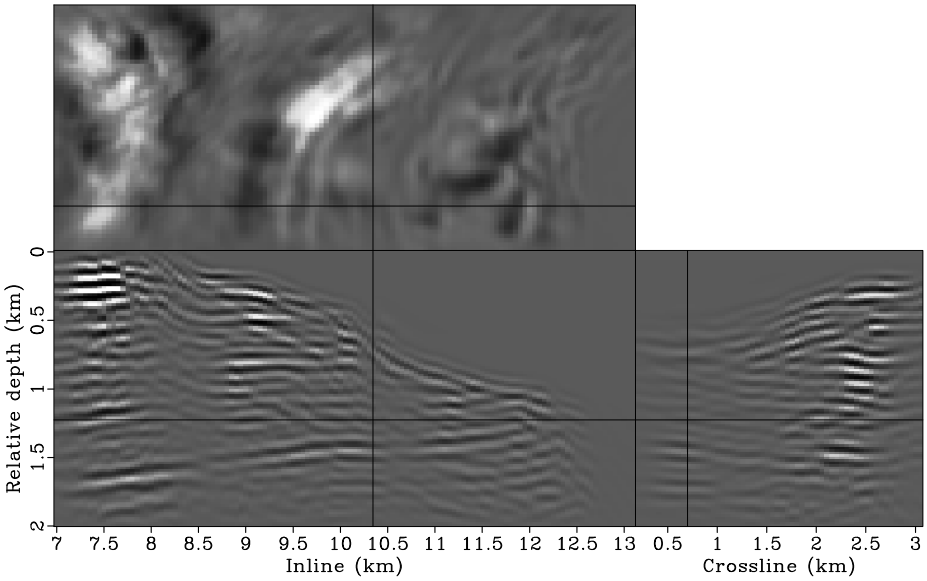
interpret the migrated image and manually pick several key reflectors (Figure 10). We estimate dip fields based on the interpreted reflectors using the structure tensor method (van Vliet and Verbeek, 1995; Hale, 2007). The estimated dips have been smoothed with a triangle filter and are shown in Figure 11. The dip field is then used to build a bank of dip filters for preconditioning. Figures 12, 13 and 14 show impulse responses of the dip filters as smoothing strength increases. Note that the stronger the smoothing effect, the longer the filter response. These dip filters also vary spatially, and each dip filter smooths along a direction conformal to the corresponding dip values.

Figure 15 presents the inverted image after 100 iterations when preconditioned by the dip filter with weak smoothing (Figure 12). The result is significantly improved over the one obtained without any regularizations (Figure 9). The reflectors are much more coherent and they extend further into the shadow zone, filling in the illumination gap almost completely. The regularized image is much easier to interpret geologically.

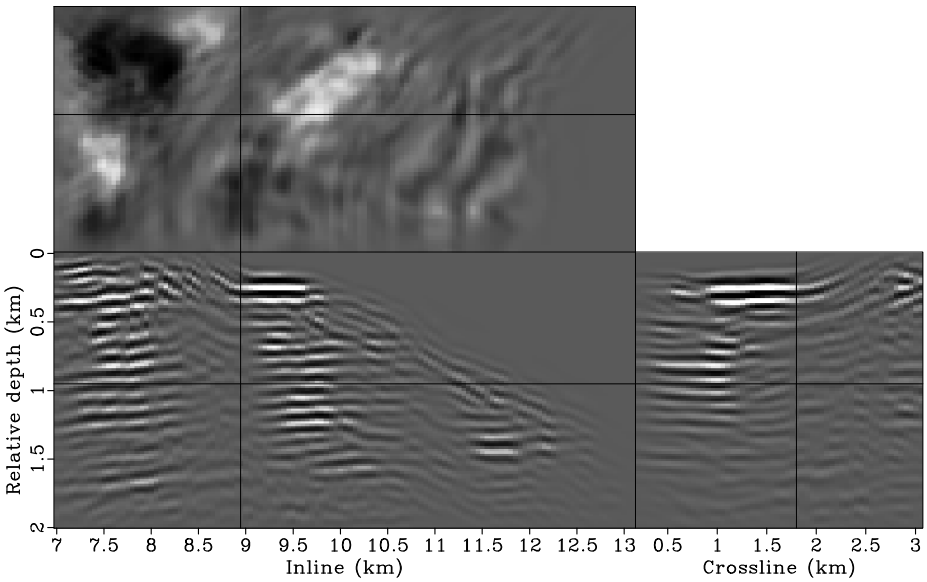
To see the effects of regularization (or preconditioning), we increase the smoothing strength of the dip filters. Figure 16 shows the inversion result after 100 iterations when we use dip filters with moderate smoothing (Figure 12) for preconditioning. The result further enhances the coherence and continuity of the reflectors and the inverted image looks even cleaner. However, the spatial resolution of the image seems somewhat degraded by the smoothing effect of the preconditioner (This becomes clear by comparing the depth slices of Figures 9, 15 and 16).

Figure 17 shows the inversion result when we further increase the smoothing strength of the dip filters (Figure 14). In this extreme case, inversion is dominated by preconditioning. The inverted image honors the user-supplied model-covariance, but not necessarily the data. (The data fitting plays little role in this case.)

Figures 18 and 19 compare the residuals at the last iteration and the convergence of

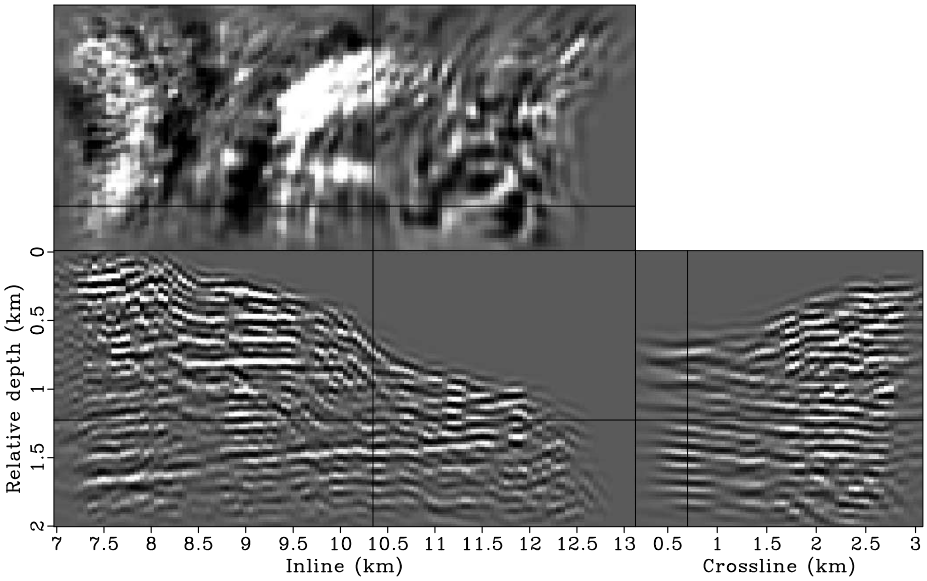


(a)

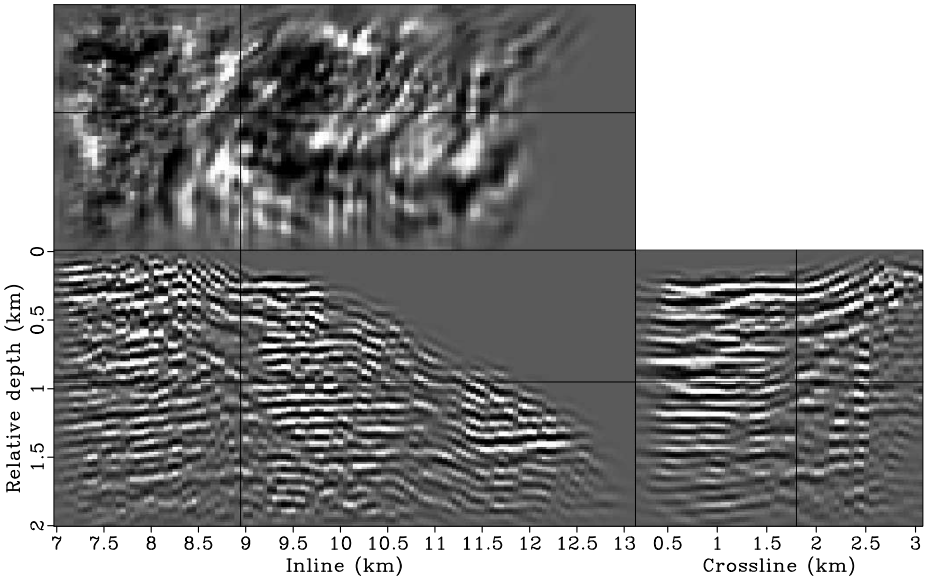


(b)

Figure 8: The migrated image of the 3-D GOM data set. Panels (a) and (b) show different slices of the same 3-D cube. The image has been masked using the mask operator shown in Figure 7 to focus on comparing sediment reflectivities. [**CR**] yaxun2/. lsm3d-imag-mask1,lsm3d-imag-mask2



(a)



(b)

Figure 9: The inverted image without applying any regularization. Panels (a) and (b) show different slices of the same 3-D cube. [**CR**] yaxun2/. lsm3d-noreg-invt1,lsm3d-noreg-invt2

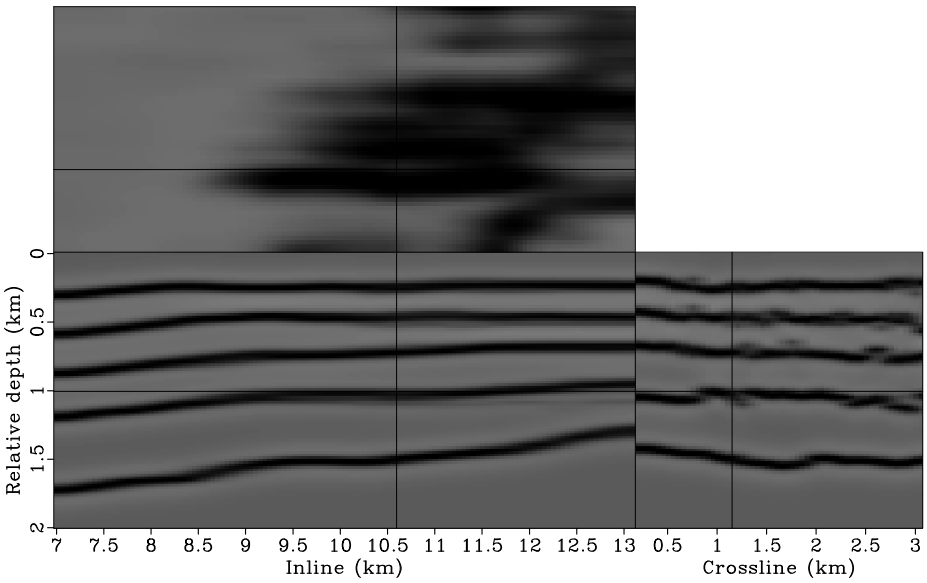
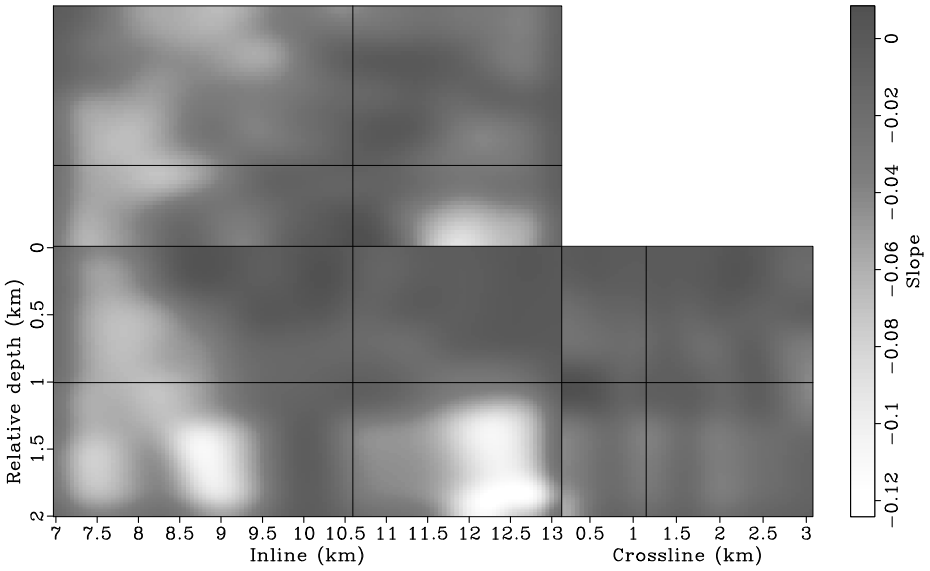


Figure 10: Interpreted horizons from the migrated image. The horizons are then used to build the dip field for dip filtering. [**CR**] yaxun2/. lsm3d-refl-pick

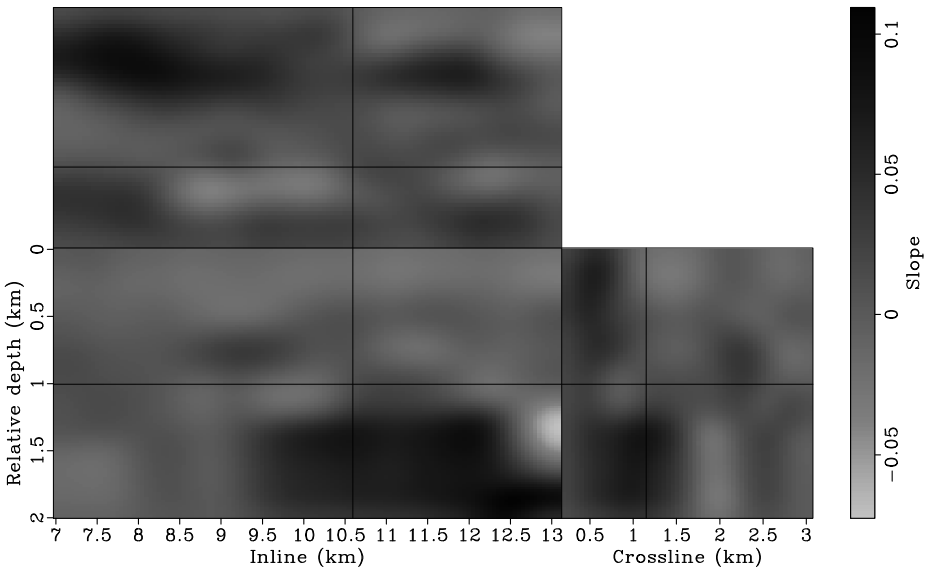
objective functions for different methods. As expected, inversion without any regularization, which fits the data most closely, has the smallest residual, whereas inversion preconditioned with strong dip filters, which fits the data least closely, has the biggest residual.

CONCLUSIONS

Solving wavefield least-squares migration in the image domain makes possible target-oriented application of this method, allowing reflectivity inversion at the reservoir level. The 3-D examples demonstrate that simultaneous phase-encoding in the conical-wave domain drastically reduces the computational cost of the 3-D Hessian matrix. The phase-encoded Hessian, albeit with some approximations, accurately quantifies the illumination effects on the migrated image. Since inverting the Hessian is very fast, different regularization parameters or schemes can be tried at very low cost. For the 3-D example shown in this chapter, it takes only about 6 minutes to run 100 iterations using 34 CPUs (17 nodes with 2 cores on each). This is a very important advantage over the conventional data-domain implementation, which requires full-domain modeling and migration at each iteration. The high efficiency of this method also makes interactive reflectivity imaging possible, where we can repeat the inversion with regularizations that incorporate different geological scenarios and obtain the results in almost real time. The 3-D reflectivity inversion results illustrate that inversion preconditioned with dip filters successfully recovers the reflectivity from the effects of uneven illumination, yielding more balanced amplitudes and higher spatial resolution in the inverted image.



(a)



(b)

Figure 11: The estimated dip field from the interpreted reflectors. Panels (a) and (b) are the slopes in the inline and crossline directions, respectively. [**CR**] yaxun2/. lsm3d-dipx-pick,lsm3d-dipy-pick

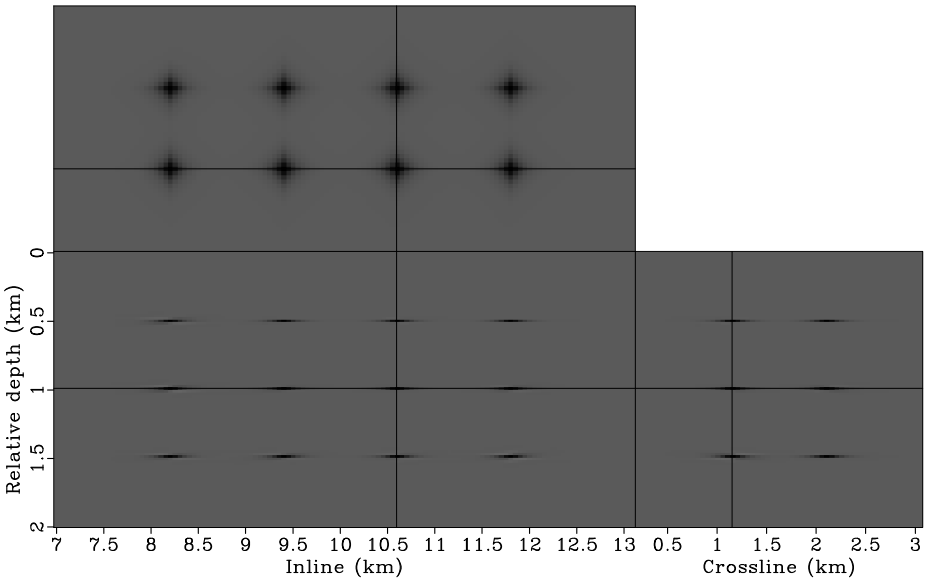


Figure 12: Impulse responses of the dip filters with weak smoothing. [**CR**] yaxun2/. lsm3d-dipflt3

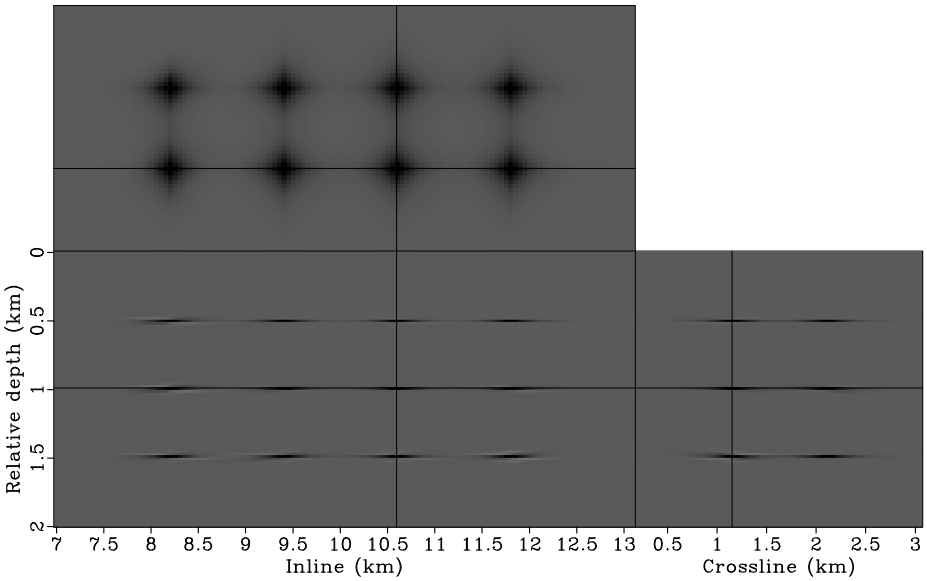


Figure 13: Impulse responses of the dip filters with moderate smoothing. [**CR**] yaxun2/. lsm3d-dipflt1

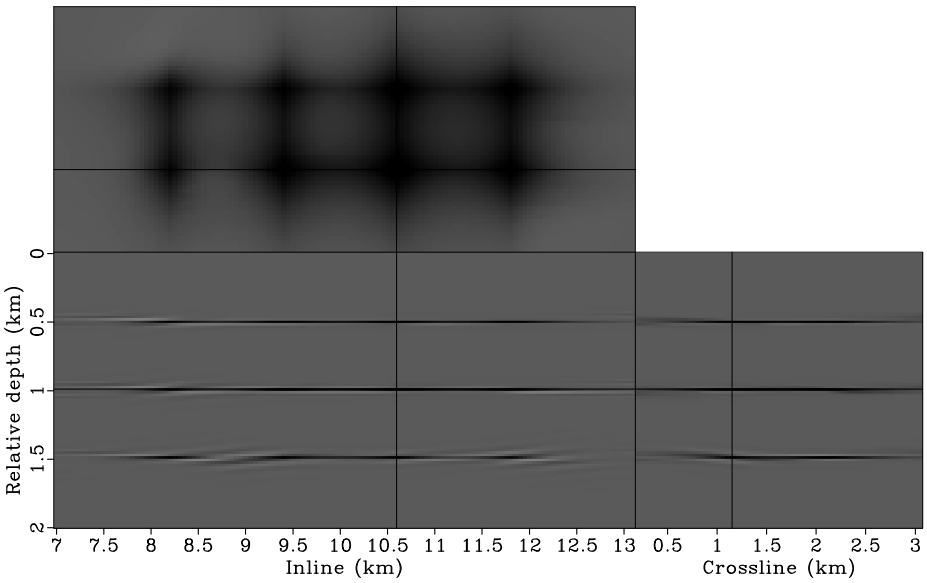


Figure 14: Impulse responses of the dip filters with strong smoothing. [**CR**] yaxun2/. lsm3d-dipflt6

ACKNOWLEDGEMENTS

We thank Gboyega Ayeni for insightful discussion about dip filtering. We acknowledge BP and ExxonMobil for providing the field data set. We also thank the Stanford Center for Computational Earth and Environmental Science for providing computing resources.

REFERENCES

Ayeni, G., Y. Tang, and B. Biondi, 2009, Joint preconditioned least-squares inversion of simultaneous source time-lapse seismic data sets: SEG Technical Program Expanded Abstracts, **28**, 3914–3918.

Biondi, B., S. Fomel, and N. Chemingui, 1998, Azimuth moveout for 3-d prestack imaging: Geophysics, **63**, 574–588.

Claerbout, J., 1998, Multidimensional recursive filters via a helix: Geophysics, **63**, 1532–1541.

Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Scientific Publication.

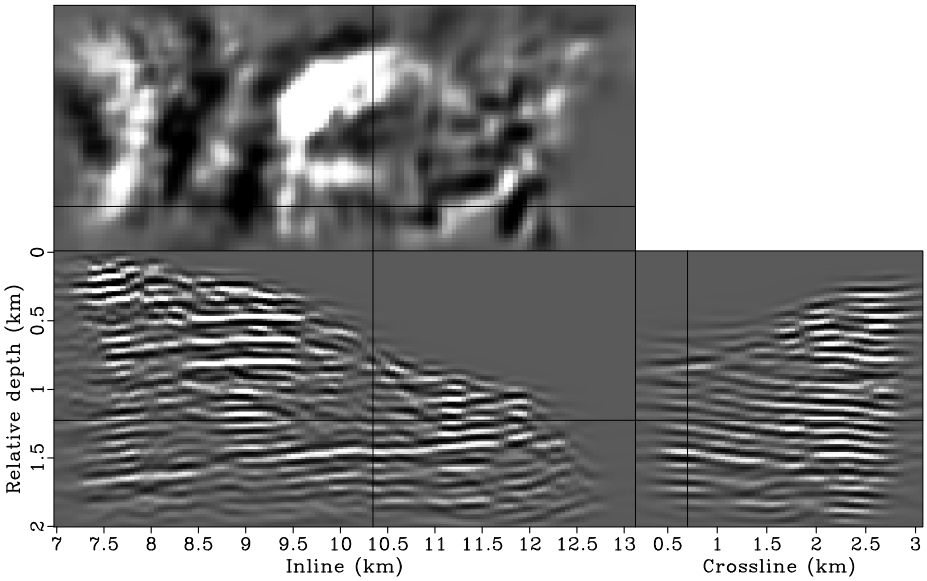
———, 1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific Publication.

———, 2008, Image Estimation By Example: Geophysical Sounding Image Construction: Multidimensional Autoregression: Stanford University.

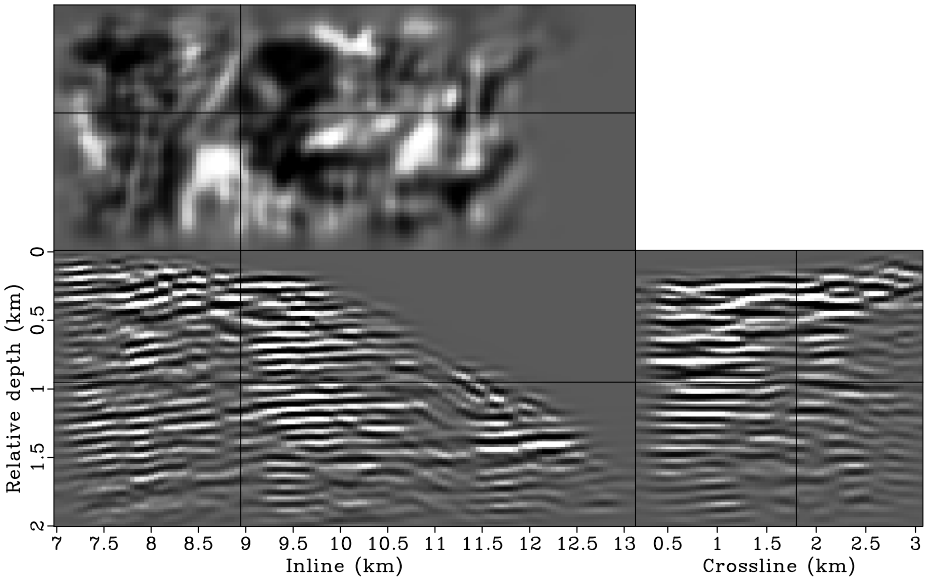
Clapp, M. L., 2005, Imaging Under Salt: Illumination Compensation by Regularized Inversion: PhD thesis, Stanford University.

Clapp, R., 2003, Geologically Constrained Migration Velocity Analysis: PhD thesis, Stanford University.

Duquet, B., P. Lailly, and A. Ehinger, 2001, 3d plane wave migration of streamer data: SEG Technical Program Expanded Abstracts, **20**, 1033–1036.

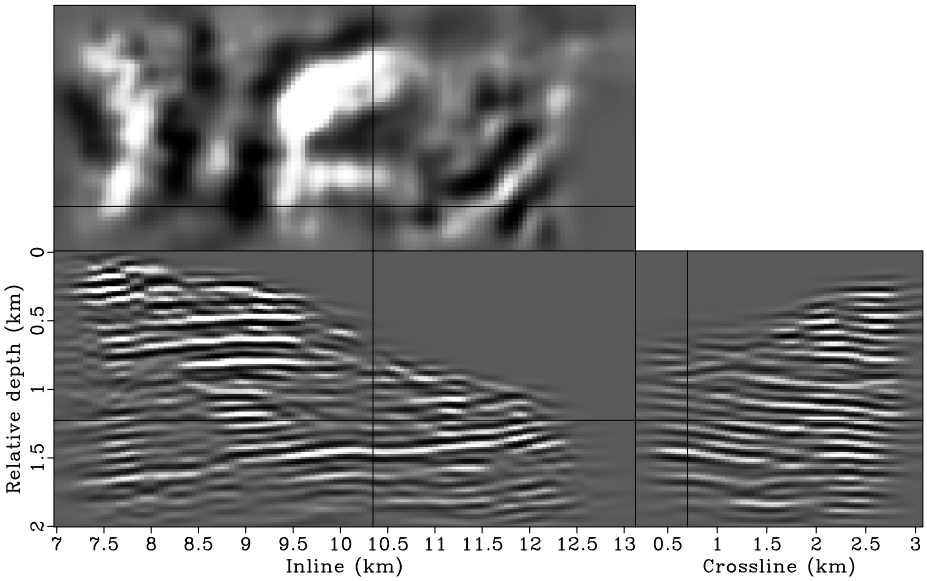


(a)

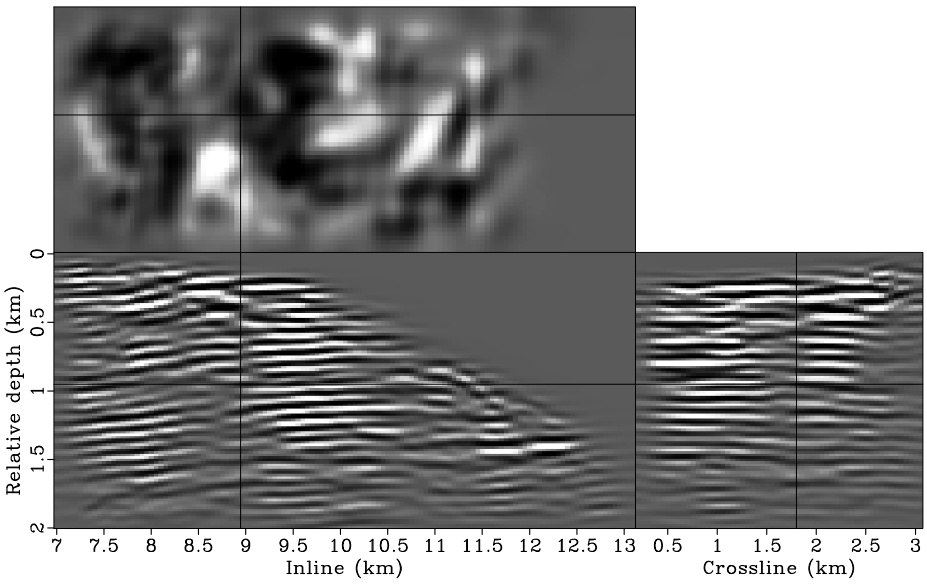


(b)

Figure 15: The inverted image when preconditioned using dip filters with weak smoothing (Figure 12). [**CR**] yaxun2/. lsm3d-a-invt1,lsm3d-a-invt2

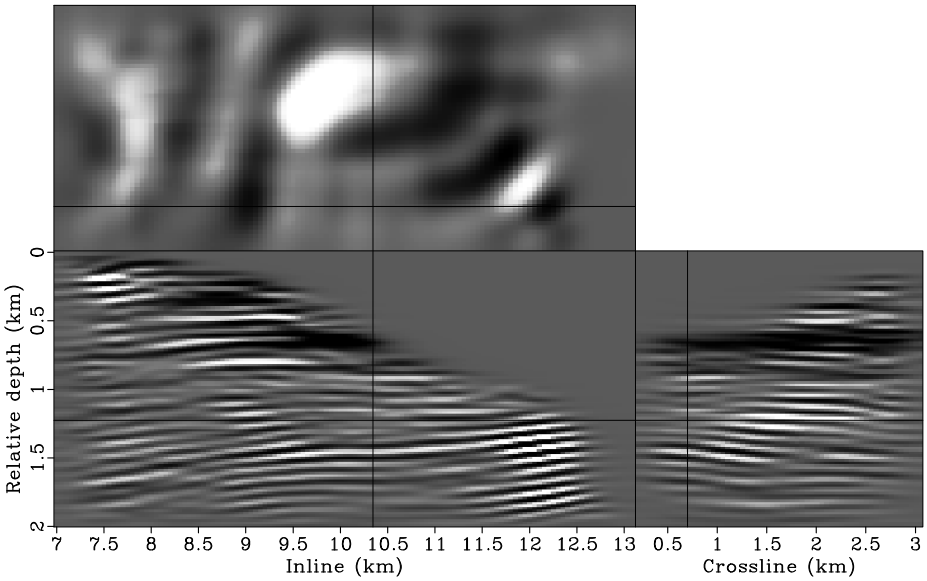


(a)

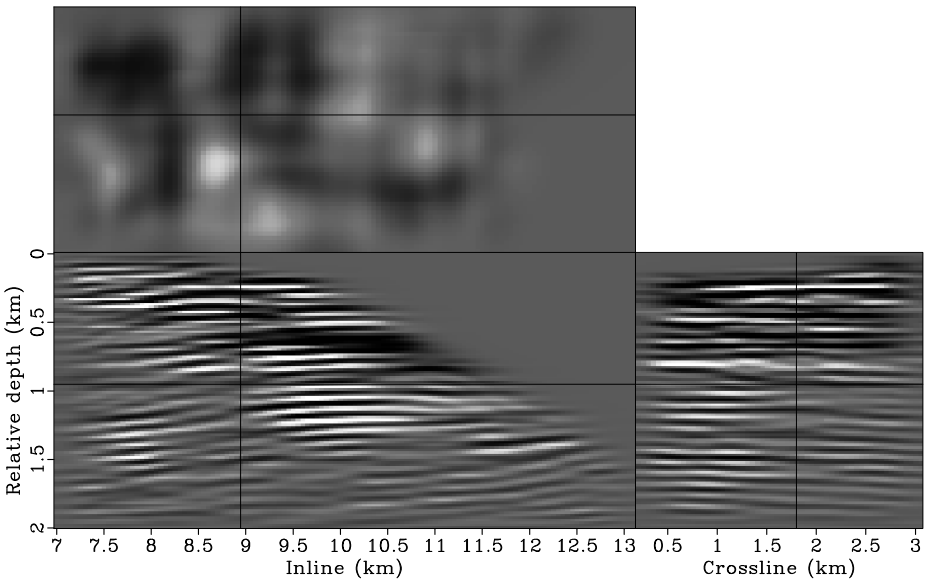


(b)

Figure 16: The inverted image when preconditioned using dip filters with moderate smoothing (Figure 13). [**CR**] yaxun2/. lsm3d-b-invt1,lsm3d-b-invt2



(a)



(b)

Figure 17: The inverted image when preconditioned using dip filters with strong smoothing (Figure 14). [**CR**] yaxun2/. lsm3d-c-invt1,lsm3d-c-invt2

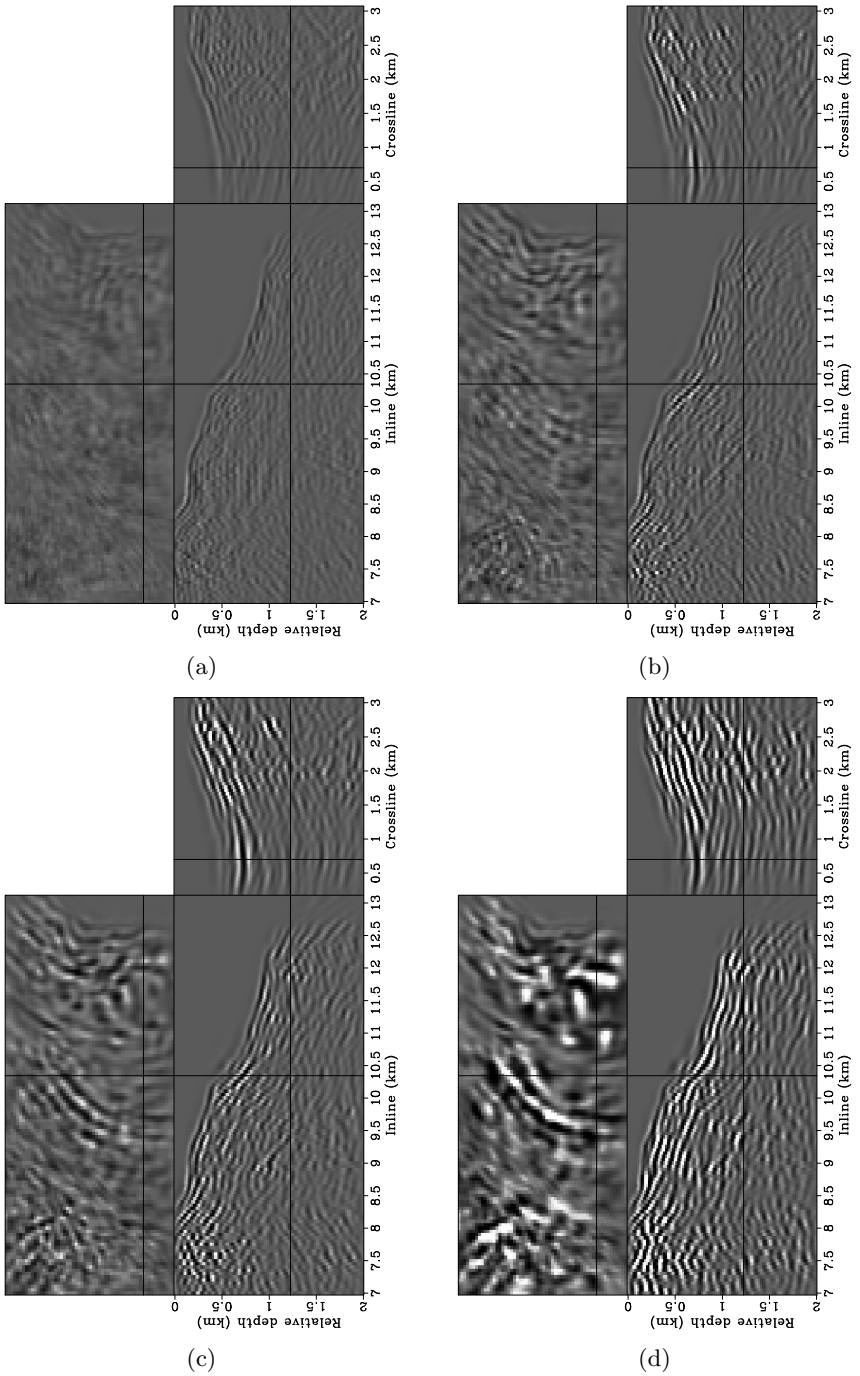


Figure 18: Residuals at the last iteration for different methods. Panel (a) is obtained using inversion without regularization. Panels (b), (c) and (d) are obtained using inversion preconditioned with weak, moderate and strong dip filtering, respectively. All panels are clipped to the same value. [**CR**] yaxun2/. lsm3d-noreg-resd,lsm3d-a-resd,lsm3d-b-resd,lsm3d-c-resd

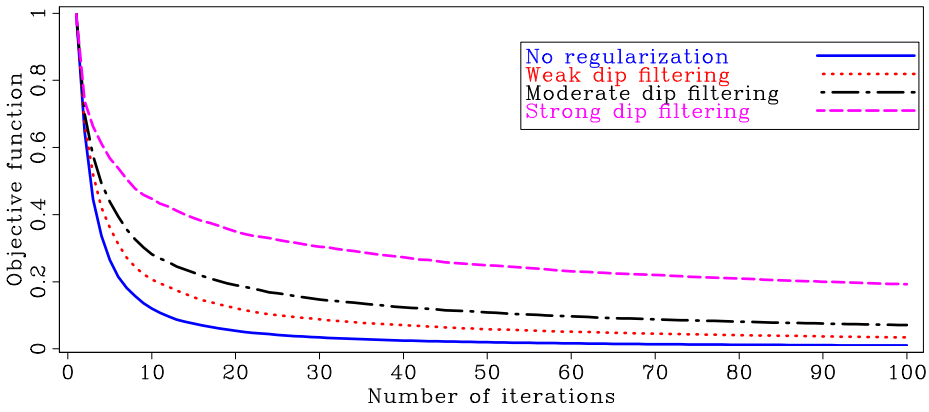


Figure 19: Evolution of objective functions for different inversion methods. [**CR**] yaxun2/. lsm3d-fobj

Fomel, S., P. Sava, J. Rickett, and J. F. Claerbout, 2003, The wilsonburg method of spectral factorization with application to helical filtering: Geophysical Prospecting, **51**, 409–420.

Gray, R. and L. D. Davisson, 2003, Introduction to Statistical Signal Processing: Cambridge University Press.

Hale, D., 2007, Local dip filtering with directional laplacians: CWP Project Review, **567**, 91–102.

Kühl, H. and M. D. Sacchi, 2003, Least-squares wave-equation migration for avp/ava inversion: Geophysics, **68**, 262–273.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migration: Proc. Conf. on Inverse Scattering, Theory and Applications, Expanded Abstracts, Philadelphia, SIAM.

Lecomte, I., 2008, Resolution and illumination analyses in PSDM: A ray-based approach: The Leading Edge, **27**, no. 5, 650–663.

Liu, F., D. W. Hanson, N. D. Whitmore, R. S. Day, and R. H. Stolt, 2006, Toward a unified analysis for source plane-wave migration: Geophysics, **71**, no. 4, S129–S139.

Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.

Stolt, R. H. and A. Benson, 1986, Seismic Migration: Theory and Practice: Geophysical Press.

Tang, Y., 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian: Geophysics, **74**, WCA95–WCA107.

Tang, Y. and B. Biondi, 2011, Subsalt velocity analysis by target-oriented wavefield tomography: A 3-D field-data example: **SEP-143**, xxx–xxx.

Tang, Y. and S. Lee, 2010, Preconditioning full waveform inversion with phase-encoded hessian: SEG Technical Program Expanded Abstracts, **29**, 1034–1038.

Valenciano, A., 2008, Imaging by Wave-equation Inversion: PhD thesis, Stanford University.

van Vliet, L. J. and P. W. Verbeek, 1995, Estimators for orientation and anisotropy in digitized images: ASCI95, Proc. First Annual Conference of the Advanced School for Computing and Imaging (Heijen, NL, May 16-18), ASCI, 442–450.

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

Zhang, Y., J. Sun, C. Notfors, S. H. Gray, L. Chernis, and J. Young, 2005, Delayed-shot 3d depth migration: Geophysics, **70**, E21–E28.

APPENDIX A

3-D CONICAL-WAVE DOMAIN HESSIAN

In general, a 3-D surface seismic data set can be represented by a 5-D object $d(\mathbf{x}_r, \mathbf{x}_s, \omega)$, with $\mathbf{x}_r = (x_r, y_r, z_r = 0)$ and $\mathbf{x}_s = (x_s, y_s, z_s = 0)$ being the receiver and source position, respectively, and ω being the angular frequency. Under the Born approximation (Stolt and Benson, 1986), the data can be modeled by a linear operator as follows:

$$d(\mathbf{x}_r, \mathbf{x}_s, \omega) = \sum_{\mathbf{x}} \omega^2 f_s(\omega) G(\mathbf{x}, \mathbf{x}_s, \omega) G(\mathbf{x}, \mathbf{x}_r, \omega) m(\mathbf{x}), \tag{A-1}$$

where $f_s(\omega)$ is the source function; $G(\mathbf{x}, \mathbf{x}_s, \omega)$ and $G(\mathbf{x}, \mathbf{x}_r, \omega)$ are the Green's functions connecting the source and receiver position to the image point $\mathbf{x} = (x, y, z)$, respectively. We can transform data into the conical-wave domain by slant-stacking along the inline source axis x_s as follows:

$$d(\mathbf{x}_r, p_{s_x}, y_s, \omega) = \sum_{x_s} W(\mathbf{x}_r, x_s, y_s) d(\mathbf{x}_r, x_s, y_s, \omega) \mathrm{e}^{i\omega p_{s_x} x_s}, \tag{A-2}$$

where $W(\mathbf{x}_r, x_s, y_s)$ is the acquisition mask operator, which contains ones where we record data, and zeros where we do not; p_{s_x} is the surface ray parameter in the inline direction. The inverse transform is

$$W(\mathbf{x}_r, x_s, y_s) d(\mathbf{x}_r, x_s, y_s, \omega) = |\omega| \sum_{p_{sx}} d(\mathbf{x}_r, p_{s_x}, y_s, \omega) \mathrm{e}^{-i\omega p_{s_x} x_s}, \tag{A-3}$$

where $|\omega|$ on the right hand side of the equation is also known as the "rho" filter (Claerbout, 1985).

To find a reflectivity model $\mathbf{m}$ that best fits the observed data for a given background velocity, we can minimize a data-misfit function that measures the differences between the observed data and the synthesized data in a least-squares sense. In the point-source case, the data-misfit function is

$$F(\mathbf{m}) = \frac{1}{2} \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} |W(\mathbf{x}_r, \mathbf{x}_s)[d(\mathbf{x}_r, \mathbf{x}_s, \omega) - d_{\mathrm{obs}}(\mathbf{x}_r, \mathbf{x}_s, \omega)]|^2, \tag{A-4}$$

where d_{obs} is the observed data. Substituting equation A-3 into A-4 yields

$$\begin{aligned} F(\mathbf{m}) \;=\; & \frac{1}{2} \sum_{\omega} \sum_{y_s} \sum_{\mathbf{x}_r} \sum_{p_{s_x}} \sum_{p'_{s_x}} |\omega|^2 \\ & \times [d(\mathbf{x}_r, p_{s_x}, y_s, \omega) - d_{\mathrm{obs}}(\mathbf{x}_r, p_{s_x}, y_s, \omega)]^* \\ & \times [d(\mathbf{x}_r, p'_{s_x}, y_s, \omega) - d_{\mathrm{obs}}(\mathbf{x}_r, p'_{s_x}, y_s, \omega)] \sum_{x_s} e^{-i\omega(p'_{s_x} - p_{s_x})x_s}. \end{aligned} \tag{A-5}$$

If the inline source axis x_s is reasonably well sampled, we have $\sum_{x_s} e^{-i\omega(p'_{s_x}-p_{s_x})x_s} \approx \frac{1}{|\omega|}\delta(p'_{s_x}-p_{s_x})$, where $\delta(\cdot)$ is the Dirac delta function. Therefore, an objective function equivalent to equation A-4 in the 3-D conical-wave domain takes the following form:

$$F(\mathbf{m}) \approx \frac{1}{2}\sum_{\omega}|\omega|\sum_{y_s}\sum_{p_{s_x}}\sum_{\mathbf{x}_r}|d(\mathbf{x}_r,p_{s_x},y_s,\omega)-d_{\text{obs}}(\mathbf{x}_r,p_{s_x},y_s,\omega)|^2. \tag{A-6}$$

The Hessian operator in the 3-D conical-wave domain can be obtained by taking the second-order derivatives of $F(\mathbf{m})$ (equation A-6) with respect to the model parameters:

$$H(\mathbf{x},\mathbf{x}') = \sum_{\omega}|\omega|^5\sum_{y_s}\sum_{p_{s_x}}\sum_{\mathbf{x}_r}\left(\frac{\partial d(\mathbf{x}_r,p_{s_x},y_s,\omega)}{\partial m(\mathbf{x})}\right)\left(\frac{\partial d(\mathbf{x}_r,p_{s_x},y_s,\omega)}{\partial m(\mathbf{x}')}\right)^*. \tag{A-7}$$

When $\mathbf{x}=\mathbf{x}'$, we obtain the diagonal components of the Hessian, which are also known as the subsurface illumination; otherwise, we obtain the off-diagonal components of the Hessian, which are also known as the resolution function for a given acquisition setup.

With equations A-1 and A-2, we obtain the expression of the derivative of d with respect to m as follows:

$$\begin{aligned}\frac{\partial d(\mathbf{x}_r,p_{s_x},y_s,\omega)}{\partial m(\mathbf{x})} &= \sum_{x_s}\omega^2 W(\mathbf{x}_r,x_s,y_s)f_s(\omega)G(\mathbf{x},x_s,y_s,\omega)\\ &\quad\times G(\mathbf{x},\mathbf{x}_r,\omega)\mathrm{e}^{i\omega p_{s_x}x_s}.\end{aligned} \tag{A-8}$$

Substituting equation A-8 into equation A-7 yields the expression for each component of the Hessian matrix in the 3-D conical-wave domain:

$$\begin{aligned}H(\mathbf{x},\mathbf{x}') &= \sum_{\omega}|\omega|^5\sum_{y_s}\sum_{p_{s_x}}\sum_{\mathbf{x}_r}G(\mathbf{x},\mathbf{x}_r,\omega)G^*(\mathbf{x}',\mathbf{x}_r,\omega)\\ &\quad\sum_{x_s}W(\mathbf{x}_r,x_s,y_s)f_s(\omega)G(\mathbf{x},x_s,y_s,\omega)\mathrm{e}^{i\omega p_{s_x}x_s}\\ &\quad\sum_{x'_s}W(\mathbf{x}_r,x'_s,y_s)f^*_s(\omega)G^*(\mathbf{x}',x'_s,y_s,\omega)\mathrm{e}^{-i\omega p_{s_x}x'_s}.\end{aligned} \tag{A-9}$$

Wave-equation inversion of time-lapse seismic data sets

Gboyega Ayeni and Biondo Biondi

ABSTRACT

We propose a linearized wave-equation inversion formulation for time-lapse seismic data sets. Our method poses time-lapse imaging as a joint least-squares problem that utilizes target-oriented approximations to the Hessian of the objective function. Because our method accounts for illumination mismatches —caused by differences in acquisition geometries— and for band-limited wave-propagation effects, it provides better estimates of production-related changes in reservoir acoustic properties than conventional time-lapse processing methods. Using data sets from a North Sea field, we demonstrate how our method can be used to image differences between time-lapse data sets. Furthermore, we show that obstruction artifacts may be attenuated by wave-equation inversion.

INTRODUCTION

Reservoir rock and fluid property changes can be obtained from seismic amplitude and/or travel-time changes (Lumley, 1995; Calvert, 2005). Several successful applications of time-lapse seismic imaging to reservoir monitoring have been published by previous authors (Lefeuvre et al., 2003; Whitcombe et al., 2004; Ebaid et al., 2009).

However, in practice, production-related changes in time-lapse seismic images can be masked by non-repeatability artifacts (e.g., changes in geometry, ambient noise). To correctly interpret time-lapse seismic differences, such artifacts must be attenuated —a prerequisite conventionally achieved through image cross-equalization methods (Rickett and Lumley, 2001). Although cross-equalization methods are well developed and provide reliable results in many practical applications, they are inadequate where large inconsistencies exist between time-lapse data sets or where the reservoir overburden is complex. Where these conventional methods fail, wave-equation inversion provides a way to attenuate unwanted artifacts in time-lapse images, thereby enhancing production-related changes.

The proposed method is based on linear least-squares migration/inversion of seismic data sets (Nemeth et al., 1999; Kühl and Sacchi, 2003; Clapp, 2005). Because each iteration of a data space implementation of least-squares migration/inversion is approximately twice the migration cost, this approach is expensive. However, by posing this problem in the image space, it can be efficiently solved in a target-oriented way (Valenciano et al., 2006; Tang, 2008). For the time-lapse imaging problem, we can either invert for the complete baseline and monitor images or we can invert for a static baseline and time-lapse images between surveys. The input vectors in the resulting formulations contain the migrated images (or combinations thereof) and the output vector contains the inverted images. The operators are a concatenations of target-oriented approximations to the Hessian of the least-squares objective function (Ayeni and Biondi, 2010). We regularize the inversion using spatial (dip) and temporal (difference) constraints.

Because we assume that the data contain only primaries, robust multiple/noise attenuation is required prior to inversion. Furthermore, we assume compaction and velocity changes between surveys are small, therefore these can be removed by warping of the monitor images to the baseline.

First, we discuss linearized wave-equation inversion of time-lapse data sets. Then, we apply this method to a North Sea field time-lapse data set. We show that even with the presence of a gap (caused by a simulated obstruction) in the monitor data-set, wave-equation inversion give improved results over conventional methods.

METHODOLOGY

Given a linearized modeling operator $\mathbf{L}$, the seismic data $\mathbf{d}$ for survey i due to a reflectivity model $\mathbf{m}$ is

$$\mathbf{d}_i = \mathbf{L}_i \mathbf{m}_i. \tag{1}$$

Assuming we have two data sets (baseline $\mathbf{d}_0$ and monitor $\mathbf{d}_1$) acquired at different times over an evolving reservoir, joint least-squares migration/inversion involves solving the regression

$$\left[\begin{array}{cc} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_1 \\ \hline \epsilon_0\mathbf{R}_0 & \mathbf{0} \\ \mathbf{0} & \epsilon_1\mathbf{R}_1 \\ \hline -\zeta_0\Lambda_0 & \zeta_1\Lambda_1 \end{array}\right] \left[\begin{array}{c} \mathbf{m}_0 \\ \mathbf{m}_1 \end{array}\right] \approx \left[\begin{array}{c} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \hline \mathbf{0} \\ \mathbf{0} \\ \hline \mathbf{0} \end{array}\right], \tag{2}$$

where $\mathbf{R}_i$ and Λ_i are the spatial and temporal regularization operators respectively, and ϵ_i and ζ_i are the corresponding regularization parameters. Although we can directly solve equation 2 by minimizing the quadratic-norm of the regression (Ajo-Franklin et al., 2005), we choose to transform it to an image space problem of the form (Appendix A)

$$\left[\begin{array}{cc} \mathbf{H_0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H_1} \\ \hline \mathbf{R_{00}} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_{11}} \\ \hline \Lambda_{00} & -\Lambda_{01} \\ -\Lambda_{10} & \Lambda_{11} \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{array}\right] \approx \left[\begin{array}{c} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1 \\ \hline \mathbf{0} \\ \mathbf{0} \\ \hline \mathbf{0} \\ \mathbf{0} \end{array}\right], \tag{3}$$

where $\mathbf{H}_i = \mathbf{L}_i^T \mathbf{L}_i$ is the wave-equation Hessian, and $\mathbf{R}_{ii} = \epsilon_i^2 \mathbf{R}_i^T \mathbf{R}_i$ and $\Lambda_{ij} = \zeta_i {\Lambda_i}^T \zeta_j \Lambda_j$ are the spatial and temporal constraints. Note that $\mathbf{L}^T$, the migration operator, is the adjoint of the modeling operator $\mathbf{L}$. The inverted time-lapse image $\Delta\hat{\mathbf{m}}$ is then the difference between the inverted baseline and monitor images ($\hat{\mathbf{m}}_0$ and $\hat{\mathbf{m}}_1$). Alternatively, we can re-write equation 3 to invert directly for the time-lapse image and a static baseline image (Appendix A). Furthermore, equation 3 can be extended to multiple seismic data sets (Ayeni and Biondi, 2010).

As shown in Appendix A, if there are physical movements of reflectors and velocity changes (e.g., due to reservoir depletion and compaction) between surveys, the baseline and monitor images will not be aligned. Such misalignments must be accounted for before or during inversion. As is the case in many practical time-lapse monitoring problems,

we assume that the monitor data are migrated with the baseline velocity. The updated inversion problem is then given by (Appendix A)

$$\left[\begin{array}{cc} \mathbf{H}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1^b \\ \hline \mathbf{R}_{00} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{11}^b \\ \hline \mathbf{\Lambda}_{00} & -\mathbf{\Lambda}_{01}^b \\ -\mathbf{\Lambda}_{10}^b & \mathbf{\Lambda}_{11}^b \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1^b \end{array}\right] \approx \left[\begin{array}{c} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1^b \\ \hline \mathbf{0} \\ \mathbf{0} \\ \hline \mathbf{0} \\ \mathbf{0} \end{array}\right], \tag{4}$$

where $\tilde{\mathbf{m}}_1^b$ and $\hat{\mathbf{m}}_1^b$ are respectively the migrated and inverted monitor images repositioned (warped) to the baseline image. The superscript ^{b} on the operators denotes that they are referenced to the baseline image. For example, $\mathbf{H}_1^b$ is the Hessian matrix with the monitor geometry but with the baseline velocity. Note that whereas the conventional time-lapse image $\Delta\tilde{\mathbf{m}}^b$ estimated at the baseline position is given by

$$\Delta\tilde{\mathbf{m}}^b = \tilde{\mathbf{m}}_1^b - \tilde{\mathbf{m}}_0, \tag{5}$$

the inverted time-lapse image $\Delta\hat{\mathbf{m}}^b$ is given by

$$\Delta\hat{\mathbf{m}}^b = \hat{\mathbf{m}}_1^b - \hat{\mathbf{m}}_0. \tag{6}$$

The wave-equation Hessian at image point $\mathbf{x}$ is defined as (Plessix and Mulder, 2004; Valenciano et al., 2006)

$$H\left(\mathbf{x},\mathbf{x}'\right) = \sum_{\omega}\omega^4 \sum_{\mathbf{x}_s} |f_s\left(\omega\right)|^2 G(\mathbf{x}_s,\mathbf{x},\omega)\bar{G}(\mathbf{x}_s,\mathbf{x}',\omega) \sum_{\mathbf{x}_r} G(\mathbf{x},\mathbf{x}_r,\omega)\bar{G}(\mathbf{x}',\mathbf{x}_r,\omega), \tag{7}$$

where $\mathbf{x}'$ denotes all image points, $\bar{G}$ is the complex conjugate of Green's function G at frequency ω, f_s is the source function, and $\mathbf{x}_s$ and $\mathbf{x}_r$ are the source and receiver positions, respectively. For any practical application, it is infeasible (and unnecessary) to compute the full Hessian matrix. Because the problem is posed in the image space, we only need to compute the Hessian for a target region of interest around the reservoir. In addition, we only compute off-diagonal elements sufficient to capture the dominant structure of the Hessian. This target-oriented approximation of the Hessian is given by (Valenciano et al., 2006)

$$H\left(\mathbf{x_T},\mathbf{x_{T+a_x}}\right) = \sum_{\omega}\omega^4 \sum_{\mathbf{x_s}} |f\left(s\right)|^2 G(\mathbf{x_s},\mathbf{x_T},\omega)\bar{G}(\mathbf{x_s},\mathbf{x_{T+a_x}},\omega) \sum_{\mathbf{x_r}} G(\mathbf{x_T},\mathbf{x_r},\omega)\bar{G}(\mathbf{x_{T+a_x}},\mathbf{x_r},\omega), \tag{8}$$

where $\mathbf{x_T}$ is an image point within the target area, and $\mathbf{x_{T+a_x}}$ represents points within a *small* region around $\mathbf{x_T}$. For any image point, elements of $\mathbf{H}\left(\mathbf{x_T},\mathbf{x_{T+a_x}}\right)$ represents a row of a sparse Hessian matrix $\mathbf{H}$ whose non-zero components are defined by $\mathbf{a_x}$. Therefore, $\mathbf{a_x}$ defines the how many off-diagonal elements of the Hessian are computed — which represents

the size of the point spread function (PSF) around each image point. Note that because of symmetry, only one half of the approximate Hessian is required. The target-oriented Hessian and computational savings are discussed in further detail by Valenciano et al. (2006) and Tang (2008).

In this paper, the spatial regularization operators are non-stationary dip filters. First, we estimate the local dips on the migrated baseline image using a plane-wave destruction method (Fomel, 2002). Then we construct the operator based on factorization of directional Laplacian representations of the local dip filters (Hale, 2007). The temporal constraint is implemented as a difference between the aligned images. To attenuate multiples and other unwanted artifacts in the data, we perform a high-resolution Radon demultiple on the data. In order to align the baseline and monitor images prior to inversion, we perform pre-stack warping using a cyclic local cross-correlation method (Ayeni, 2010).

In the next section, we apply the proposed method to a field time-lapse data set. We show applications to complete and incomplete monitor data sets.

EXAMPLE

We apply the proposed method to a North Sea time-lapse data set with two years of production between the baseline and the monitor. The shots and receivers are spaced at 25 m, and the maximum offset is 3200 m. First, we show how wave-equation inversion can improve time-lapse amplitude information on complete data sets. In this example, the baseline and monitor data sets have similar geometries. Secondly, we simulate an obstruction in the monitor data with a gap 400 m wide. In this example, except at the obstruction where no sources or receivers are present in the monitor, the baseline and monitor data sets have similar geometries. Figure 1 shows the full migrated baseline stacked section. After preliminary processing of the data, we identified the reservoir location around which the target area for inversion is defined. Henceforth, we show the results only for the target area. All inversion results are obtained using the RJMI formulation.

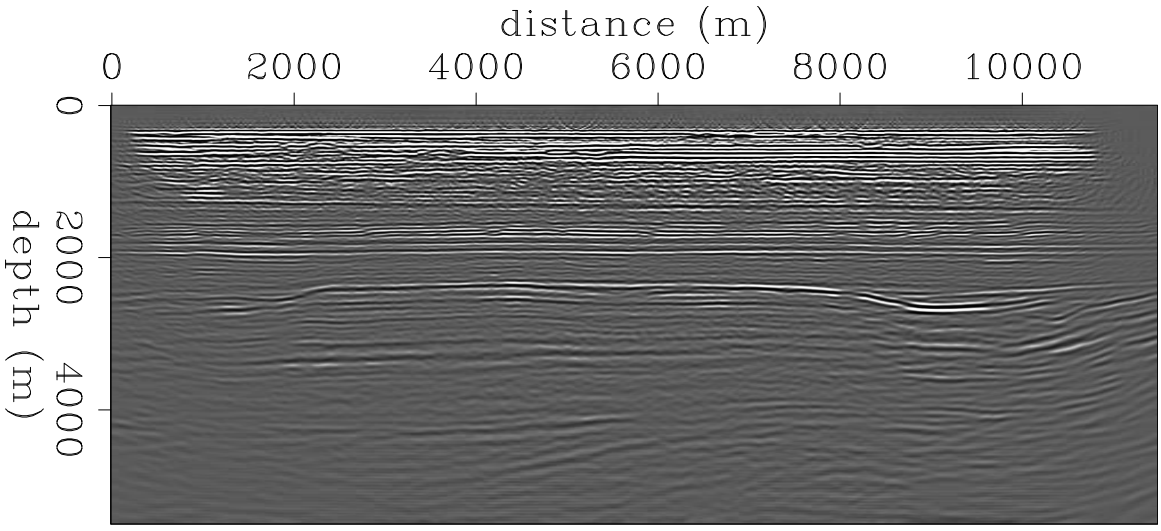


Figure 1: Raw baseline stacked image. [**CR**] gayeni1/. raw-2759-full-g-06

Example 1: Complete baseline and monitor data

Figure 2 shows the migrated pre-stack image obtained from the raw baseline data set. Note in Figure 2 that multiples show significant curvature relative to the correctly-migrated flat primaries. The same image after multiple attenuation is shown in Figure 3. In order to remove small velocities errors, we applied a residual moveout correction to the image gathers. The pre-processed pre-stack monitor image is shown in Figure 4. To obtain this image, in addition to the pre-processing applied to the baseline, we also applied a bulk amplitude correction to the monitor data. We estimate the amplitude correction as the ratio between the root-mean-square energy in the baseline versus the monitor. In addition, we perform pre-stack warping of the monitor to the baseline image. Apparent vertical displacements between the baseline and monitor are shown in Figure 5. Figure 6 shows the Hessian diagonal for the target area. Local dip estimates obtained from the stacked baseline image (not shown) are shown in Figure 7. Figure 8 shows the time-lapse images after different processing steps. Note the incremental improvements in the time-lapse image quality after processing and after inversion.

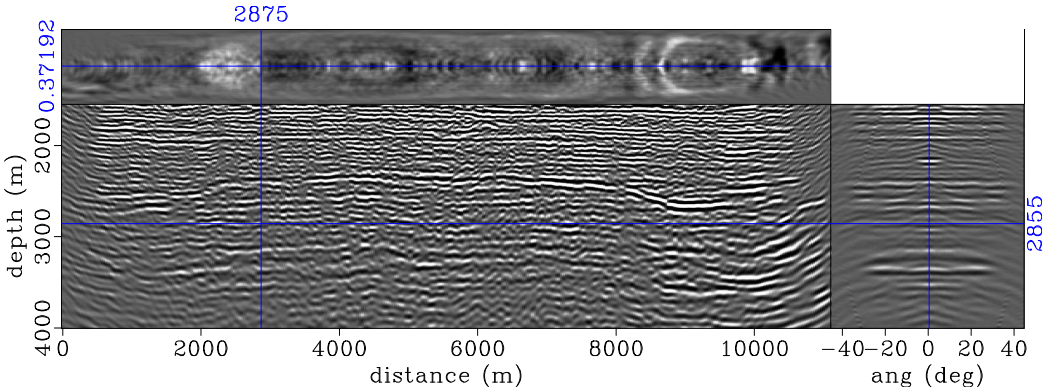


Figure 2: Raw pre-stack baseline image for the target area. Note the presence of several multiples and other undesired artifacts. [**CR**] gayeni1/. raw-2759-06

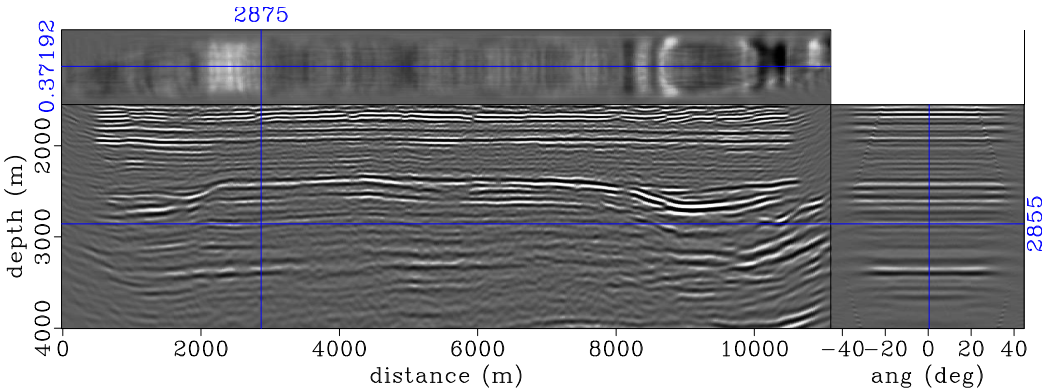


Figure 3: Preprocessed pre-stack baseline image of the target area. Note that the artifacts in the raw image (Figure 2) have been attenuated. [**CR**] gayeni1/. flat-2759-06-s

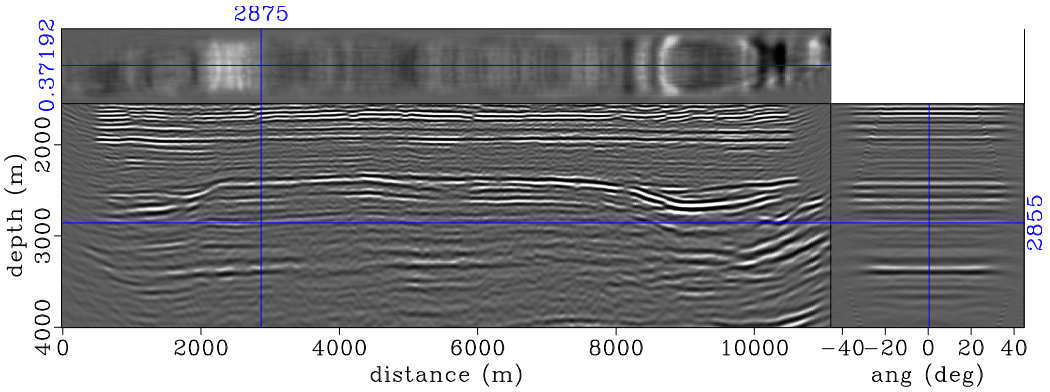


Figure 4: Preprocessed pre-stack monitor image of the target area obtained using the same parameters as in the baseline (Figure 3). Note that this image has been warped (using apparent displacements in Figures 5) to the baseline image. [**CR**] gayeni1/. warp-2759-08

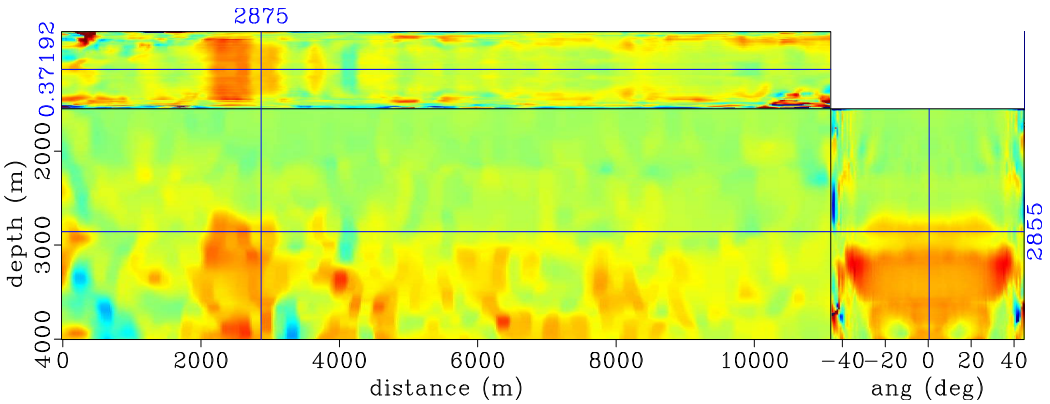


Figure 5: Apparent vertical displacements between the baseline and monitor images (Figures 3 and 4). Red, blue and green denote positive, negative and zero displacements respectively. Note that the maximum apparent displacements correspond to the reservoir location between position 2000 m to 3000 m. Also, note that the apparent displacement varies with opening angle. [**CR**] gayeni1/. warp-2759-08-rflat-ts

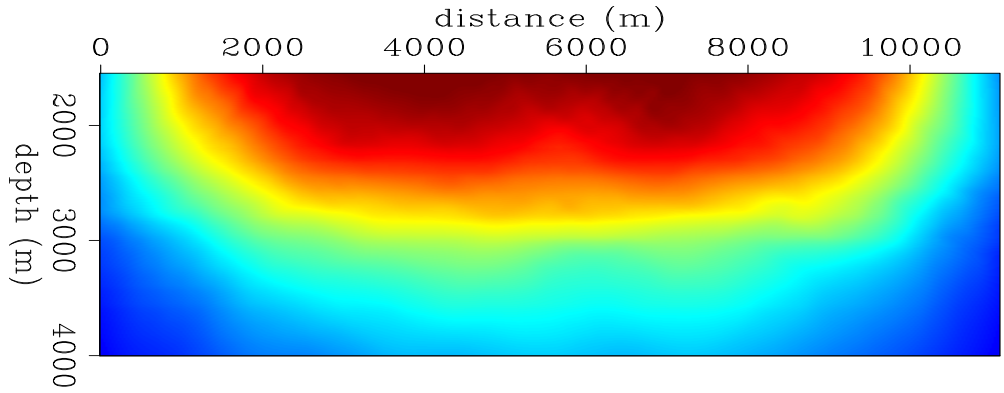


Figure 6: Illumination (Hessian diagonal) for the target area. Red indicates high illumination whereas blue indicates low illumination. [**CR**] gayeni1/. hs-2759-06

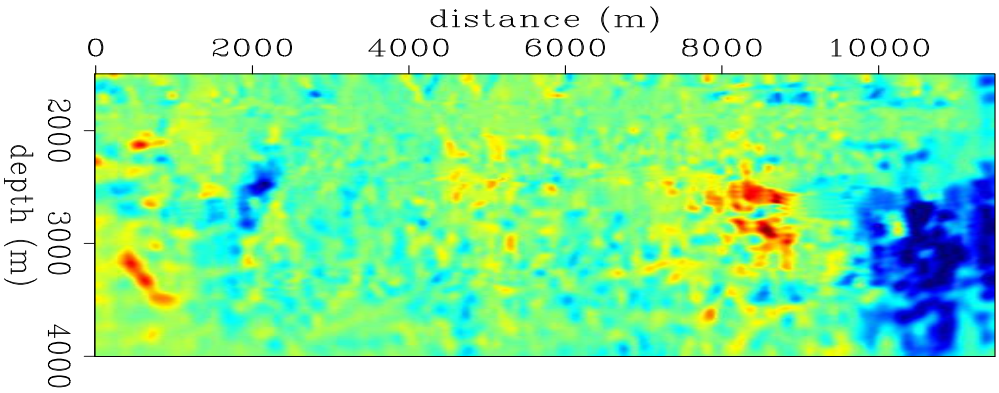


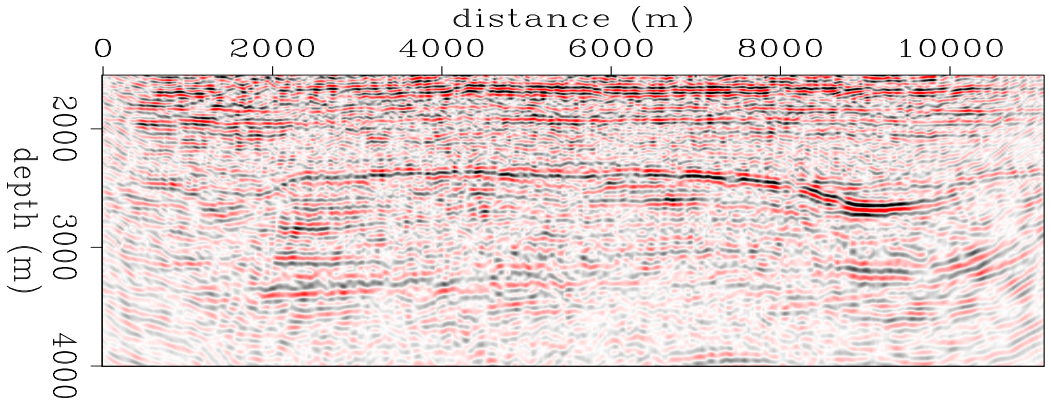
Figure 7: Dip estimates obtained from the stacked baseline image. Red indicates positive dips whereas blue indicates negative dips. These dip estimates are used to construct the spatial regularization operator. [**CR**] gayeni1/. dm-dip-2759

Example 2: Complete baseline and Incomplete monitor data

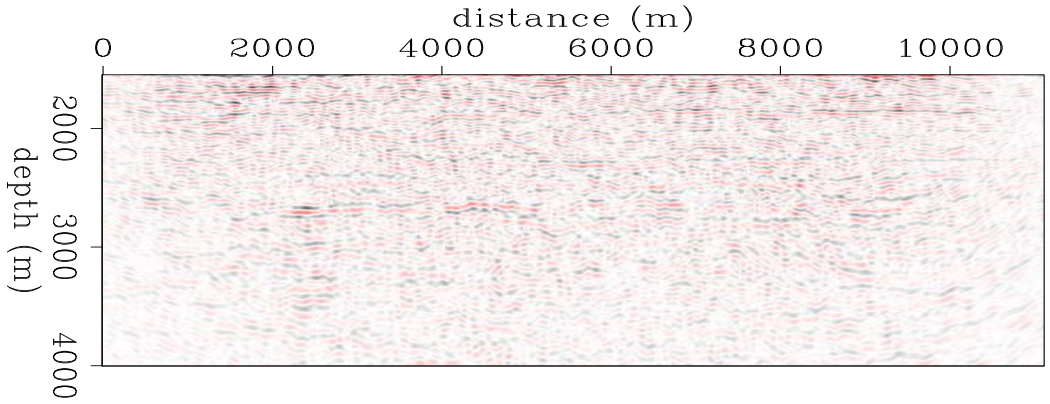
Figure 9 shows that gapped monitor data. Here, we simulate an obstruction in the geometry such that neither sources nor receivers are present between 2500 and 2900 m. This gap in the the data coincides with the reservoir location. We image gapped monitor data using the same parameters and in the previous example. However, in order to warp this incomplete monitor to the complete baseline, we first perform Radon interpolation to fill in the monitor gap. The interpolated monitor data are then imaged and processed using the same parameters as in the complete monitor case. The resulting pre-stack image is then used to estimate the apparent displacement vectors (Figure 10) used to warp the incomplete monitor image to the baseline. The ratio of the Hessian diagonal (illumination ratio) between the baseline and monitor is shown Figure 11. Note that the region with low illumination ratios correspond to the location of the simulated obstruction. Figures 12 and 13 show the spatial and wavenumber domain point-spread-functions (PSFs) at point $x = 2600$ m and $z = 2600$ m, respectively. Note that there is significant differences in the PSFs away from the diagonal (center of the spatial PSF) and at various wavenumbers. Figure 14 shows the time-lapse images after different processing steps. Note the incremental improvements in the time-lapse image quality after processing and after inversion. Furthermore, note that compared to the complete data example (Figure 8(b)), the time-lapse image from conventional processing (Figure 14(b)) is of poorer quality.

DISCUSSION

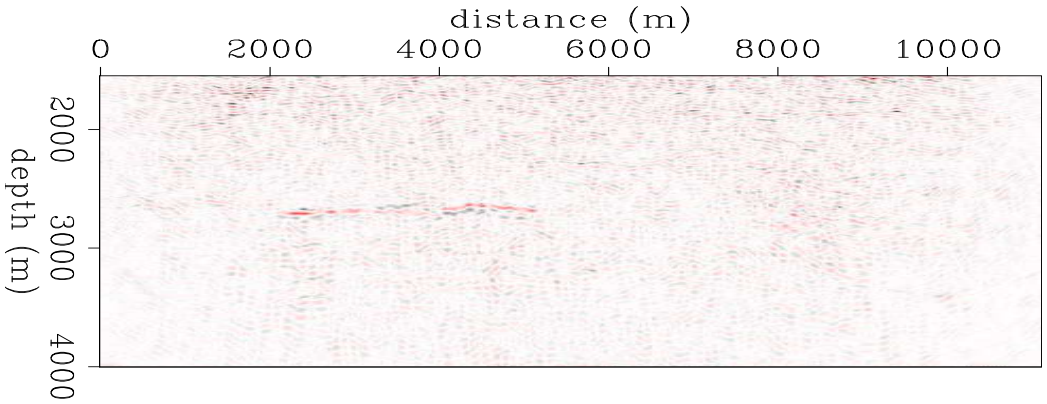
Data preprocessing sufficiently attenuates multiples and other artifacts in the data (Figures 2 and 3). This is required to ensure that the data satisfy sufficiently the primaries-only assumption in our inversion formulation. By warping the images before stacking, we ensure that defocusing effects due to velocity and compaction effects are minimized. Because the overburden geology along the studied section is fairly simple, the monitor image (Figure 4) is sufficiently aligned to the baseline using only vertical components of the prestack apparent displacement vectors (Figure 5). In practice, even with good repeatability between surveys,



(a)



(b)



(c)

Figure 8: Time-lapse images (a) from the raw data, (b) after time-lapse processing and (c) after wave-equation inversion. Note the incremental improvements in the time-lapse image from (a) to (c). [**CR**] gayeni1/. raw-2759-4d,rwarp-2759-4d,inv-2759-4d

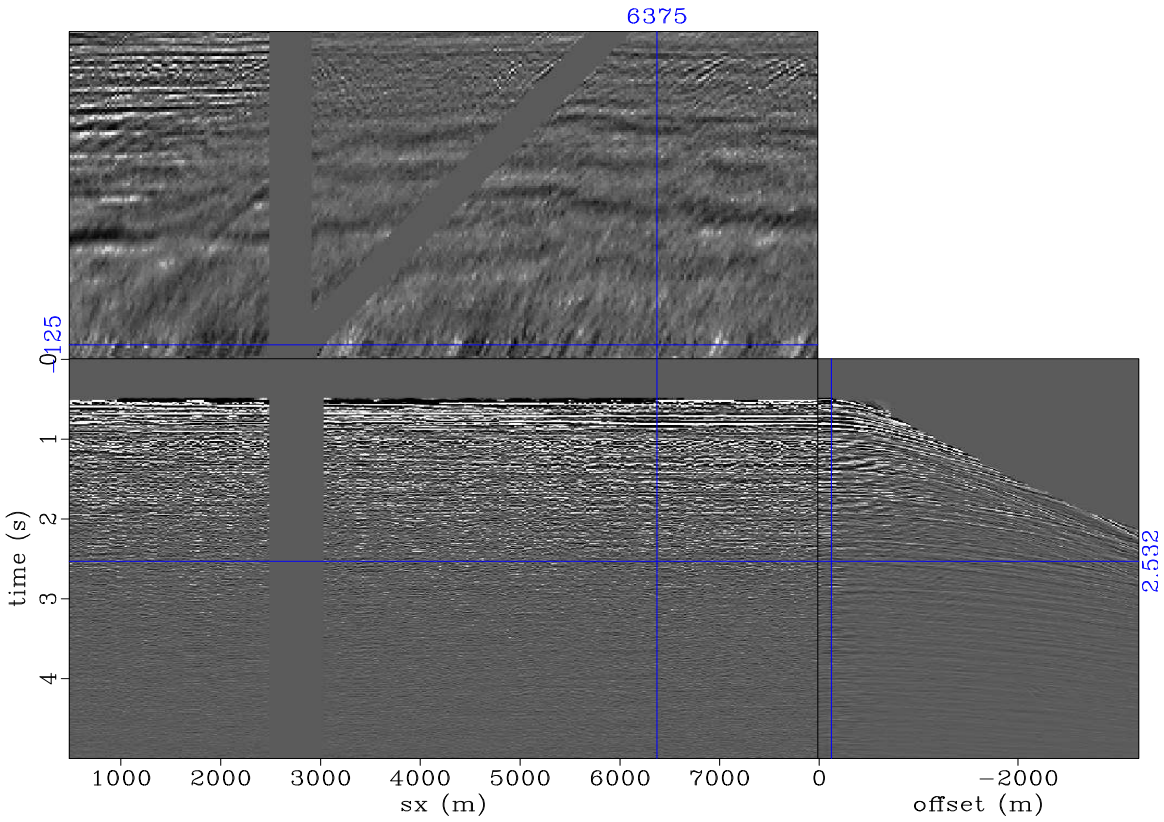


Figure 9: Gapped monitor data. Note that we have sources and receivers are missing around within the simulated obstruction. [**CR**] gayeni1/. dmhl-dat-2759-b4-hole

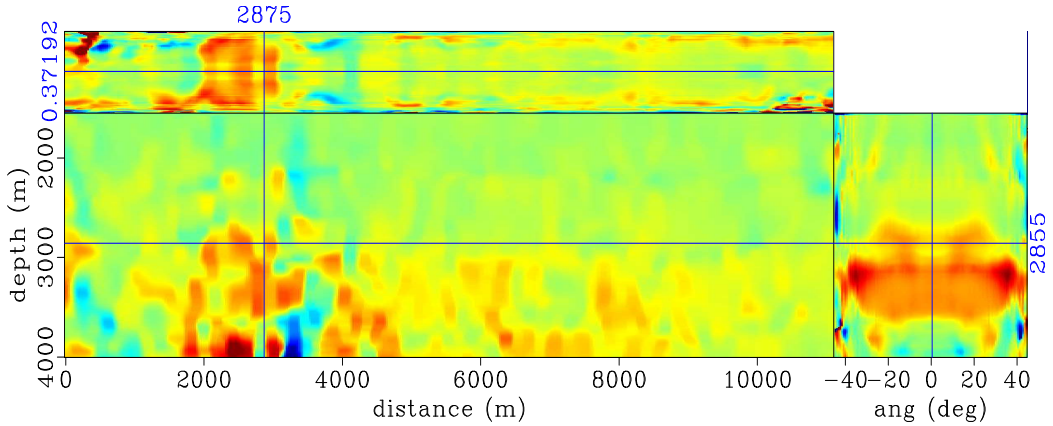


Figure 10: Apparent vertical displacements between images from the baseline and interpolated monitor data sets. Comparing this to Figure 5, note that estimates of the apparent displacements are similar to those from the complete data case. [**CR**] gayeni1/. dmhl-warp-2759-08-rflat-ts

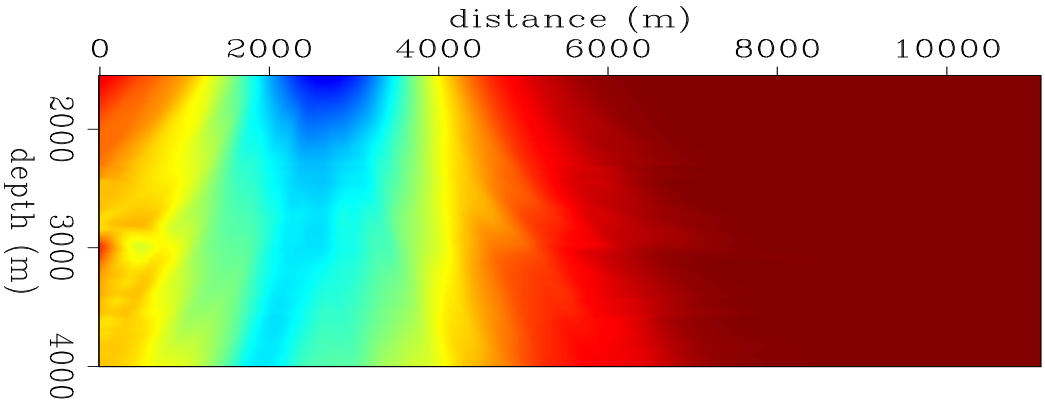


Figure 11: Illumination ratio between the baseline and monitor. Red indicates regions with equal illumination (i.e., ratio equals unity) whereas blue indicates unequal illumination (i.e., ratio less than unity). [**CR**] gayeni1/. hs-2759-rat

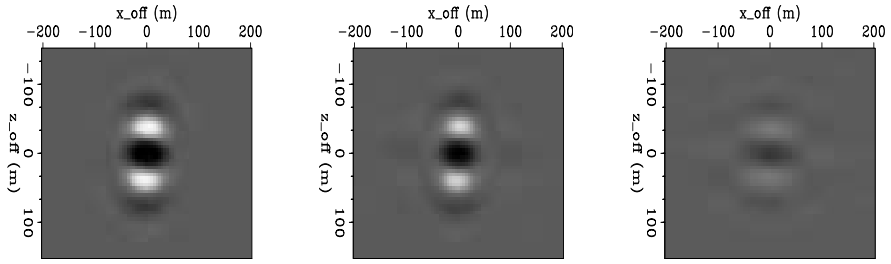


Figure 12: Point-spread-functions at point $x = 2600$ m and $z = 2600$ m for the (a) baseline and (b) monitor. Note in the difference PSF (c), that there is significant energy away from the center of the PSF (i.e. in the off-diagonal parts of the Hessian). [**CR**] gayeni1/. hs-2759-offd-06,hs-2759-offd-08,hs-2759-offd-diff

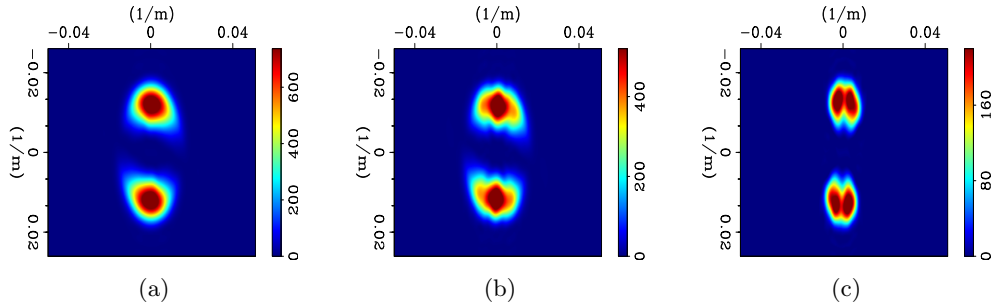
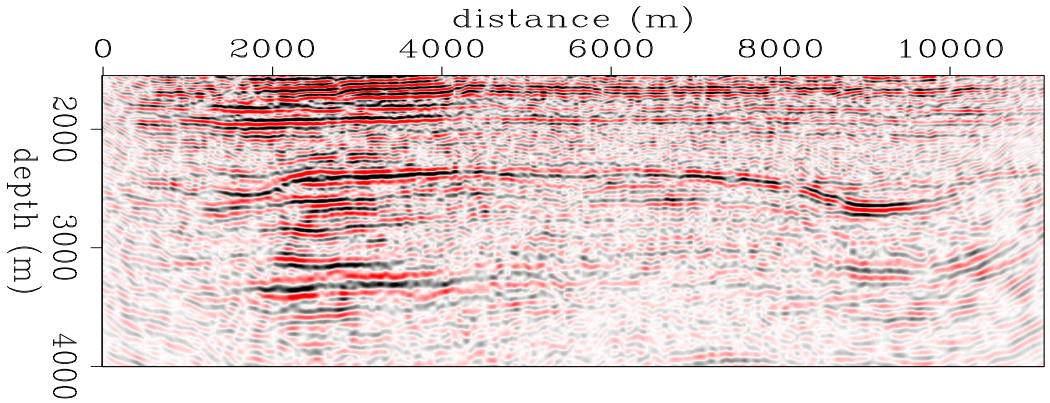
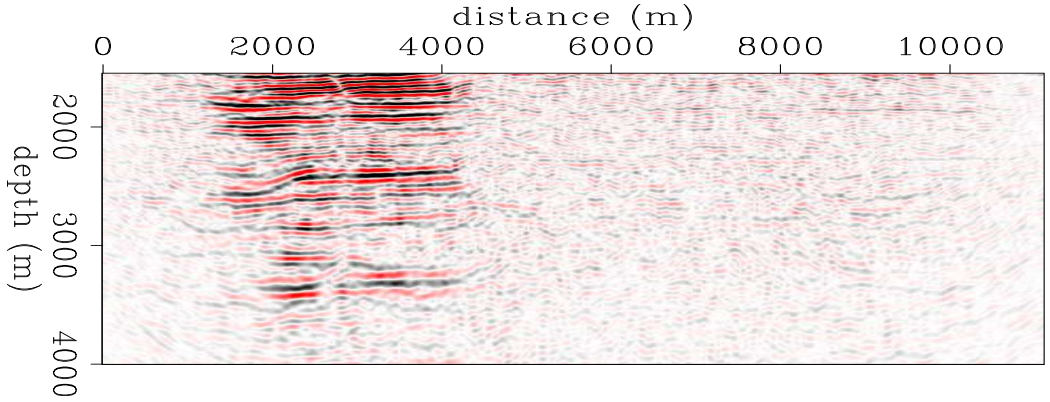


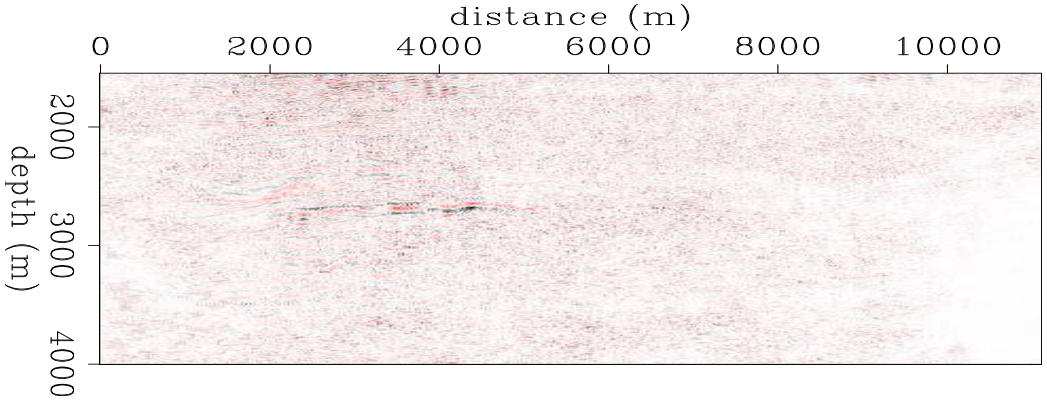
Figure 13: Wavenumber domain point-spread-functions at point $x = 2600$ m and $z = 2600$ m for the (a) baseline and (b) monitor. Note in the difference PSF (c), that there is significant illumination differences at various wavenumbers. [**CR**] gayeni1/. hs-2759-fft-06,hs-2759-fft-08,hs-2759-fft-diff



(a)



(b)



(c)

Figure 14: Time-lapse images (a) from the raw data, (b) after time-lapse processing and (c) after wave-equation inversion. Also, because of artifacts introduced by the incomplete monitor data, conventional methods fail to provide results of comparable quality to the complete data example (Figure 8(b)). Note that inversion provides satisfactory results. [**CR**] gayeni1/. dmhl-raw-2759-4d,dmhl-rwarp-2759-4d,dmhl-inv-2759-4d

it is difficult to interpret an unprocessed time-lapse image (Figure 8(a)). However, after careful processing, it is possible to make meaningful interpretation of amplitude information in the time-lapse image (Figure8(b)). Although in many cases, results from a conventional processing workflow may suffice, the quality of the time-lapse image can be improved by wave-equation inversion (Figure 8(c)). In the second example, because the monitor data is incomplete, the effective geometries differ for the two surveys, thereby leading to illumination mismatch (i.e., illumination ratios not equal to unity) in parts of the target area (Figure 11). However, the Hessian diagonal gives only a partial measure of the illumination mismatch between time-lapse surveys (Figure 11). Large geometry differences (e.g., an obstruction in the monitor acquisition) can cause large differences in the off-diagonal terms of the Hessian (Figure 12). Such geometry differences lead to differences in wavenumber illumination between surveys (Figure 13). Therefore, a point-by-point amplitude compensation using only the Hessian diagonal is inadequate. Where there is significant geometry difference between surveys, conventional time-lapse processing is insufficient (Figure 14(b)). In this case, wave-equation inversion provides a significant improvement to the time-lapse image (Figure 14(c)). Although the amplitude information derivable from the interpolated monitor data (not shown) are of poor quality the kinematics are similar to those of the complete baseline image. Therefore these provide adequate estimates of the warping parameters (Figure 10) comparable to those from the complete data case (Figure 5).

CONCLUSIONS

We have proposed a wave-equation inversion scheme for time-lapse seismic data sets. Applying the proposed method to a North sea field data set shows that it can improve production-related time-lapse seismic responses. This method can improve time-lapse amplitude information and can be used to attenuate artifacts caused by geometry/illumination differences between surveys.

ACKNOWLEDGMENTS

We thank Statoil and partners (ENI and Petoro) for donating the field data set used in this study. We thank the Stanford Center for Computational Earth & Environmental Science (CEES) for providing the computer resources used in this study.

APPENDIX A

Assuming we have two data sets (baseline $\mathbf{d}_0$ and monitor $\mathbf{d}_1$) acquired at different times, the following sections shows derivations of the the joint inversion formulations for image differences and for multiple images.

Joint Inversion for Image Differences (JID)

We can formulate baseline and monitor data modeling as follows:

$$\begin{bmatrix} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{L}_1 & \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{\Delta m} \end{bmatrix} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix} , \tag{A-1}$$

which can be divided into the following two parts:

$$\mathbf{L}_0\mathbf{m}_0 = \mathbf{d}_0, \tag{A-2}$$

$$\mathbf{L}_1\mathbf{m}_0 + \mathbf{L}_1\mathbf{\Delta m} = \mathbf{d}_1, \tag{A-3}$$

where the time-lapse reflectivity image $\mathbf{\Delta m}$ is given by

$$\mathbf{\Delta m} = \mathbf{m}_1 - \mathbf{m}_0. \tag{A-4}$$

Note that Equation A-1 assumes that both $\mathbf{m}_0$ and $\mathbf{m}_1$ are collocated. This means that there is no physical movement of the reflector between the baseline and the monitor images. In addition, equation A-1 assumes that there are no overburden velocity changes. If stress changes cause any physical movement of a point from baseline position (x_0, y_0, z_0) in $\mathbf{m}_0$ to monitor position (x_1, y_1, z_1) in $\mathbf{m}_1$, we can update equation A-3 such that the point in $\mathbf{m}_0$ is repositioned at (x_1, y_1, z_1). The updated modeling equation for the monitor data then becomes

$$\mathbf{L}_1\mathbf{S}^{m-}\mathbf{m}_0 + (\mathbf{L}_1\mathbf{m}_1 - \mathbf{L}_1\mathbf{S}^{m-}\mathbf{m}_0) = \mathbf{L}_1\mathbf{S}^{m-}\mathbf{m}_0 + \mathbf{L}_1\mathbf{\Delta m}^m = \mathbf{d}_1, \tag{A-5}$$

where $\mathbf{S}^{m-}$ is an orthogonal warping operator that aligns $\mathbf{m}_0$ to $\mathbf{m}_1$, and

$$\mathbf{\Delta m}^m = \mathbf{m}_1 - \mathbf{S}^{m-}\mathbf{m}_0 \tag{A-6}$$

is the time-lapse image estimated at the monitor position (x_1, y_1, z_1). And the combined modeling equation becomes

$$\begin{bmatrix} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{L}_1\mathbf{S}^{m-} & \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{\Delta m}^m \end{bmatrix} \approx \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix}. \tag{A-7}$$

However, note that equation A-7 requires that we know the true reflector position in the monitor which we may obtain from a geomechanical model. Furthermore, any regularization on the time-lapse image must be applied at the monitor position, or by first repositioning the time-lapse image to the baseline position as follows:

$$\mathbf{\Delta m} = \mathbf{\Delta m}^b = \mathbf{S}^{m+}\mathbf{\Delta m}^m, \tag{A-8}$$

where $\mathbf{\Delta m}^b$ is the time-lapse image at the baseline position and $\mathbf{S}^{m+} = (\mathbf{S}^{m-})^{-1}$ is an operator that repositions events from the monitor position to the baseline position.

Assuming we migrate the monitor data with the true monitor velocity, we arrive at the image-space inversion problem by minimizing the quadratic-norm

$$\left\| \begin{bmatrix} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{L}_1\mathbf{S}^{m-} & \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{\Delta m}^m \end{bmatrix} - \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix} \right\| \approx 0 \tag{A-9}$$

for which the solutions $\hat{\mathbf{m}}_0$ and $\mathbf{\Delta}\hat{\mathbf{m}}^m$ satisfy the solution

$$\begin{bmatrix} \mathbf{L}_0^T\mathbf{L}_0 + \mathbf{S}^{m+}\mathbf{L}_1^T\mathbf{L}_1\mathbf{S}^{m-} & \mathbf{S}^{m+}\mathbf{L}_1^T\mathbf{L}_1 \\ \mathbf{L}_1^T\mathbf{L}_1\mathbf{S}^{m-} & \mathbf{L}_1^T\mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \mathbf{\Delta}\hat{\mathbf{m}}^m \end{bmatrix} \approx \begin{bmatrix} \mathbf{L}_0^T & \mathbf{S}^{m+}\mathbf{L}_1^T \\ \mathbf{0} & \mathbf{L}_1^T \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix}, \tag{A-10}$$

or simply

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{S}^{m+}\mathbf{H}_1\mathbf{S}^{m-} & \mathbf{S}^{m+}\mathbf{H}_1 \\ \mathbf{H}_1\mathbf{S}^{m-} & \mathbf{H}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \mathbf{\Delta}\hat{\mathbf{m}}^m \end{bmatrix} \approx \begin{bmatrix} \tilde{\mathbf{m}}_0 + \mathbf{S}^{m+}\tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_1 \end{bmatrix}. \tag{A-11}$$

Note that the time-lapse image we obtain is $\mathbf{\Delta\hat{m}}^m$ at the monitor position and not $\mathbf{\Delta\hat{m}} = \mathbf{\Delta\hat{m}}^b$ at the baseline position. Although what is most interesting is $\mathbf{\Delta\hat{m}}^m$, as shown later in this section, we may choose to re-write the formulation as a function of $\mathbf{\Delta\hat{m}}^b$.

Assuming we migrate the monitor data with the wrong (e.g. baseline) velocity, then equation A-10 becomes

$$\begin{bmatrix} \mathbf{L}_0^T\mathbf{L}_0 + \mathbf{S}^{\alpha+}\check{\mathbf{L}}_1^T\mathbf{L}_1\mathbf{S}^{m-} & \mathbf{S}^{m+}\check{\mathbf{L}}_1^T\mathbf{L}_1 \\ \check{\mathbf{L}}_1^T\mathbf{L}_1\mathbf{S}^{m-} & \check{\mathbf{L}}_1^T\mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \mathbf{\Delta\hat{m}}^m \end{bmatrix} \approx \begin{bmatrix} \mathbf{L}_0^T & \mathbf{S}^{\alpha+}\check{\mathbf{L}}_1^T \\ \mathbf{0} & \check{\mathbf{L}}_1^T \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix}, \quad \text{(A-12)}$$

where $\check{\mathbf{L}}_1^T$, the migration operator with the monitor geometry but with baseline velocity migrates the monitor data to apparent position (x_1', y_1', z_1'), and $\mathbf{S}^{\alpha+}$ repositions the migrated data from (x_1', y_1', z_1') to (x_0, y_0, z_0). However, because the operator $\mathbf{L}_1$ is a function of the true monitor velocity, if the true monitor velocity is known, we should solve equation A-11 instead of equation A-12. Note that in the case where the monitor migration velocity is the correct one, equation A-12 becomes equation A-11. If we have neither the true monitor velocity nor a geomechanical model, we may modify the Hessian in equation A-12 using the apparent displacements between (x_1', y_1', z_1') and (x_0, y_0, z_0) so we can approximate equation A-12 as

$$\begin{bmatrix} \mathbf{L}_0^T\mathbf{L}_0 + \mathbf{S}^{\alpha+}\check{\mathbf{L}}_1^T\check{\mathbf{L}}_1\mathbf{S}^{\alpha-} & \mathbf{S}^{\alpha+}\check{\mathbf{L}}_1^T\check{\mathbf{L}}_1 \\ \check{\mathbf{L}}_1^T\check{\mathbf{L}}_1\mathbf{S}^{\alpha-} & \check{\mathbf{L}}_1^T\check{\mathbf{L}}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \mathbf{\Delta\hat{m}}^a \end{bmatrix} \approx \begin{bmatrix} \mathbf{L}_0^T & \mathbf{S}^{\alpha+}\check{\mathbf{L}}_1^T \\ \mathbf{0} & \check{\mathbf{L}}_1^T \end{bmatrix} \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix}, \quad \text{(A-13)}$$

Then equation A-11 becomes

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{S}^{\alpha+}\check{\mathbf{H}}_1\mathbf{S}^{\alpha-} & \mathbf{S}^{\alpha+}\check{\mathbf{H}}_1 \\ \check{\mathbf{H}}_1\mathbf{S}^{\alpha-} & \check{\mathbf{H}}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \mathbf{\Delta\hat{m}}^a \end{bmatrix} \approx \begin{bmatrix} \tilde{\mathbf{m}}_0 + \mathbf{S}^{\alpha+}\tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_1^a \end{bmatrix}, \quad \text{(A-14)}$$

where $\check{\mathbf{H}}_1$ is the modified Hessian in which we account for the mis-positioning due to compaction and velocity change. To account for such mis-positioning, we compute the updated Hessian using perturbed Green's functions:

$$\check{H}\left(\mathbf{y_T}, \mathbf{y_{T+a_x}}\right) = \sum_{\omega}\omega^4 \sum_{\mathbf{y_s}} |f\left(s\right)|^2 G_\alpha(\mathbf{y_s}, \mathbf{y_T}, \omega)\bar{G}_\alpha(\mathbf{y_s}, \mathbf{y_{T+a_y}}, \omega) \sum_{\mathbf{y_r}} G_\alpha(\mathbf{y_T}, \mathbf{y_r}, \omega)\bar{G}_\alpha(\mathbf{y_{T+a_y}}, \mathbf{y_r}, \omega), \quad \text{(A-15)}$$

where $\mathbf{y}$ denotes an apparent point in the monitor image that corresponds to baseline point $\mathbf{x}$. The modified Green's function G_α is given by

$$G_\alpha = G(\mathbf{x}) \exp^{-i\omega\Delta t_\alpha} \approx G(\mathbf{x}) \exp^{-i\omega\frac{|\mathbf{x}_\alpha|}{v_o}}, \quad \text{(A-16)}$$

where Δt_α is the time-delay corresponding to the absolute apparent displacement $|\mathbf{x}_\alpha|$ and v_o is the baseline velocity.

Instead of inverting for the time-lapse image at the monitor position, another approach is to directly invert for $\mathbf{\Delta\hat{m}}^b = \mathbf{\Delta\hat{m}}$ at the baseline position by making the substitution

$$\mathbf{\Delta m}^m = \mathbf{S}^{m-}\mathbf{\Delta m} \quad \text{(A-17)}$$

into equation A-9 to obtain

$$\begin{bmatrix} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{L}_1\mathbf{S}^{m-} & \mathbf{L}_1\mathbf{S}^{m-} \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{\Delta m} \end{bmatrix} \approx \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix} \tag{A-18}$$

which leads to the image-space problem

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{S}^{m+}\mathbf{H}_1\mathbf{S}^{m-} & \mathbf{S}^{m+}\mathbf{H}_1\mathbf{S}^{m-} \\ \mathbf{S}^{m+}\mathbf{H}_1\mathbf{S}^{m-} & \mathbf{S}^{m+}\mathbf{H}_1\mathbf{S}^{m-} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \mathbf{\Delta}\hat{\mathbf{m}} \end{bmatrix} \approx \begin{bmatrix} \tilde{\mathbf{m}}_0 + \mathbf{S}^{m+}\tilde{\mathbf{m}}_1 \\ \mathbf{S}^{m+}\tilde{\mathbf{m}}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1^b \\ \tilde{\mathbf{m}}_1^b \end{bmatrix}. \tag{A-19}$$

where $\tilde{\mathbf{m}}_1^b$, the migrated monitor image repositioned to the baseline position (x_0, y_0, z_0), is defined as

$$\tilde{\mathbf{m}}_1^b = \mathbf{S}^{m+}\tilde{\mathbf{m}}_1. \tag{A-20}$$

If we migrate the monitor data with the baseline velocity, equation A-19 becomes

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{S}^{\alpha+}\check{\mathbf{H}}_1\mathbf{S}^{\alpha-} & \mathbf{S}^{\alpha+}\check{\mathbf{H}}_1\mathbf{S}^{\alpha-} \\ \mathbf{S}^{\alpha+}\check{\mathbf{H}}_1\mathbf{S}^{\alpha-} & \mathbf{S}^{\alpha+}\check{\mathbf{H}}_1\mathbf{S}^{\alpha-} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \mathbf{\Delta}\hat{\mathbf{m}} \end{bmatrix} \approx \begin{bmatrix} \tilde{\mathbf{m}}_0 + \mathbf{S}^{\alpha+}\tilde{\mathbf{m}}_1^a \\ \mathbf{S}^{\alpha+}\tilde{\mathbf{m}}_1^a \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1^b \\ \tilde{\mathbf{m}}_1^b \end{bmatrix}, \tag{A-21}$$

where,

$$\mathbf{S}^{\alpha+}\tilde{\mathbf{m}}_1^a \approx \tilde{\mathbf{m}}_1^b = \mathbf{S}^{m+}\tilde{\mathbf{m}}_1. \tag{A-22}$$

However, provided the velocity change is isotropic, compaction effects are small, differences in kinematics are small, and the velocity change is small, we can make the following approximation:

$$\mathbf{L}_1\mathbf{S}^{m-}\mathbf{m}_1^b \approx \mathbf{U}^{m-}\mathbf{L}_1^b\mathbf{m}_1^b = \mathbf{d}_1, \tag{A-23}$$

where the operator $\mathbf{L}_1$ is a function of both the monitor velocity and geometry, whereas $\mathbf{L}_1^b$ is a function of the baseline velocity but the monitor geometry. $\mathbf{U}^{m-}$ is an orthogonal operator that translates a data due to a reflectivity spike at baseline position (x_0, y_0, z_0) and baseline background velocity v_0, to data due to a spike at (x_1, y_1, z_1) and monitor background velocity v_1. Note that in equation A-23, we have made the following approximation

$$\mathbf{U}^{m-} \approx \mathbf{L}_1\mathbf{S}^{m-}\left[\left[(\mathbf{L}_1^b)^T\mathbf{L}_1^b\right]^{-1}(\mathbf{L}_1^b)^T\right]. \tag{A-24}$$

Provided equation A-23 holds, we can write

$$(\mathbf{L}_1^b)^T(\mathbf{U}^{m-})^T\mathbf{U}^{m-}\mathbf{L}_1^b\mathbf{m}_1^b \approx (\mathbf{L}_1^b)^T(\mathbf{U}^{m-})^T\mathbf{d}_1, \tag{A-25}$$

where,

$$\begin{gathered}(\mathbf{U}^{m-})^T\mathbf{U}^{m-} = (\mathbf{U}^{m-})^{-1}\mathbf{U}^{m-} = \mathbf{U}^{m+}\mathbf{U}^{m-} = \mathbf{I}, \\ (\mathbf{L}_1^b)^T(\mathbf{U}^{m-})^T = \mathbf{S}^{m+}\mathbf{L}_1^T.\end{gathered} \tag{A-26}$$

Therefore, we can write

$$(\mathbf{L}_1^b)^T\mathbf{L}_1^b\mathbf{m}_1^b \approx \mathbf{S}^{m+}\mathbf{L}_1^T\mathbf{L}_1\mathbf{S}^{m-}\mathbf{m}_1^b \approx \mathbf{S}^{m+}\mathbf{L}_1^T\mathbf{d}_1 \approx \tilde{\mathbf{m}}_1^b, \tag{A-27}$$

where, $(\mathbf{L}_1^b)^T\mathbf{L}_1^b = \mathbf{H}_1^b$ is the Hessian computed using the baseline velocity but with the monitor geometry. Making these substitutions into equation A-19, we have

$$\begin{bmatrix} \mathbf{H}_0 + \mathbf{H}_1^b & \mathbf{H}_1^b \\ \mathbf{H}_1^b & \mathbf{H}_1^b \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \mathbf{\Delta}\hat{\mathbf{m}} \end{bmatrix} \approx \begin{bmatrix} \tilde{\mathbf{m}}_0 + \tilde{\mathbf{m}}_1^b \\ \tilde{\mathbf{m}}_1^b \end{bmatrix}. \tag{A-28}$$

An important advantage of the formulation in equation A-28 is that it allows us to readily regularize the time-lapse image. However, it may be desirable to invert directly for the individual seismic images, as shown in the following sections.

Joint Inversion of Multiple Images (JMI)

We can pose the problem as an inversion for the individual (baseline and monitor) images. Then, the modeling equation becomes

$$\begin{bmatrix} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{m}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_0 \\ \mathbf{d}_1 \end{bmatrix} . \tag{A-29}$$

Using the same assumptions and procedure as in the previous section, we arrive at the following formulation:

$$\begin{bmatrix} \mathbf{H}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} \approx \begin{bmatrix} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1 \end{bmatrix} , \tag{A-30}$$

where the baseline and monitor velocities are known. Where the monitor has been aligned with baseline, we have

$$\begin{bmatrix} \mathbf{H}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1^b \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1^b \end{bmatrix} \approx \begin{bmatrix} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1^b \end{bmatrix} , \tag{A-31}$$

which holds approximately whether or not we migrate the monitor data with the baseline or monitor velocity.

Regularized Joint inversion of Multiple Images (RJMI)

Because in the JMI formulation, the models are completely decoupled, they can be regularized by minimizing the norm

$$\left\| \begin{bmatrix} \epsilon_0\mathbf{R}_0 & \mathbf{0} \\ \mathbf{0} & \epsilon_1\mathbf{R}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} \right\| \approx 0 , \tag{A-32}$$

where $\mathbf{R}_i$ is the spatial regularization operator and ϵ_i the spatial regularization parameter for survey i. To add any temporal regularization, we need to warp the inverted monitor images to the baseline and then apply temporal constraints or we can regularize the time-lapse image directly by minimizing the norm:

$$\left\| \begin{bmatrix} -\zeta_0\mathbf{\Lambda_0} & \zeta_1\mathbf{\Lambda}_1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} \right\| \approx 0 , \tag{A-33}$$

where Λ_i is the temporal regularization operator and ζ_i is the regularization parameter. Therefore the full regularized inversion requires a minimization of the norm:

$$\left\| \left[\begin{array}{cc} \mathbf{L}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_1 \\ \hline \epsilon_0\mathbf{R}_0 & \mathbf{0} \\ \mathbf{0} & \epsilon_1\mathbf{R}_1 \\ \hline -\zeta_0\mathbf{\Lambda}_0 & \zeta_1\mathbf{\Lambda}_1 \end{array}\right] \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} - \left[\begin{array}{c} \mathbf{d}_0 \\ \mathbf{d}_1 \\ \hline \mathbf{0} \\ \mathbf{0} \\ \hline \mathbf{0} \end{array}\right] \right\| \approx 0 , \tag{A-34}$$

which leads to the image-space problem

$$\left[\begin{array}{cc} \mathbf{H}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1 \\ \hline \mathbf{R}_{00} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{11} \\ \hline \mathbf{\Lambda}_{00} & -\mathbf{\Lambda}_{01} \\ -\mathbf{\Lambda}_{10} & \mathbf{\Lambda}_{11} \end{array}\right] \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} \approx \left[\begin{array}{c} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1 \\ \hline \mathbf{0} \\ \mathbf{0} \\ \hline \mathbf{0} \\ \mathbf{0} \end{array}\right] , \tag{A-35}$$

where $\mathbf{R}_{ii} = \epsilon_i^2 \mathbf{R}_i^T \mathbf{R}_i$ and $\mathbf{\Lambda}_{ij} = \zeta_i \mathbf{\Lambda}_i^T \zeta_j \mathbf{\Lambda}_j$ are the spatial and temporal constraints, respectively.

If the monitor has been aligned to the baseline, then we can impose the spatial regularization by minimizing

$$\left\| \begin{bmatrix} \epsilon_0 \mathbf{R}_0 & \mathbf{0} \\ \mathbf{0} & \epsilon_1 \mathbf{R}_1^b \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1^b \end{bmatrix} \right\| \approx 0 \ , \tag{A-36}$$

and the temporal regularization by minimizing

$$\left\| \begin{bmatrix} -\zeta_0 \mathbf{\Lambda_0} & \zeta_1 \mathbf{\Lambda}_1^b \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1^b \end{bmatrix} \right\| \approx 0 \ , \tag{A-37}$$

where $\mathbf{R}_1^b$ and $\mathbf{\Lambda}_1^b$ are defined with respect to the baseline-aligned monitor image. If the time-lapse image at the baseline position, the regularized image-space inversion problem is given by

$$\left[\begin{array}{cc} \mathbf{H}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_1^b \\ \hline \mathbf{R}_{00} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{11}^b \\ \hline \mathbf{\Lambda}_{00} & -\mathbf{\Lambda}_{01}^b \\ -\mathbf{\Lambda}_{10}^b & \mathbf{\Lambda}_{11}^b \end{array} \right] \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1^b \end{bmatrix} \approx \left[\begin{array}{c} \tilde{\mathbf{m}}_0 \\ \tilde{\mathbf{m}}_1^b \\ \hline \mathbf{0} \\ \mathbf{0} \\ \hline \mathbf{0} \\ \mathbf{0} \end{array} \right] , \tag{A-38}$$

where the superscript $(^b)$ denotes that the operators and images are referenced to the baseline position. Note that in the simplest case, where the temporal regularization is a difference operator equation A-33 becomes

$$\left\| \zeta \begin{bmatrix} -\mathbf{I} & \mathbf{S}^{m+} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1 \end{bmatrix} \right\| \approx 0 \ , \tag{A-39}$$

and for the baseline-aligned images, the temporal constraint in equation A-37 becomes

$$\left\| \zeta \begin{bmatrix} -\mathbf{I} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{m}}_0 \\ \hat{\mathbf{m}}_1^b \end{bmatrix} \right\| \approx 0 \ . \tag{A-40}$$

REFERENCES

Ajo-Franklin, J. B., J. Urban, and J. M. Harris, 2005, Temporal integration of seismic traveltime tomography: SEG Technical Program Expanded Abstracts, **24**, 2468–2471.

Ayeni, G., 2010, Cyclic 1d matching of time-lapse seismic data sets: A case study of the norne field: SEP Report, **140**.

Ayeni, G. and B. Biondi, 2010, Target-oriented joint least-squares migration/inversion of time-lapse seismic data sets: Geophysics, **75**, R61–R73.

Calvert, R., 2005, Insights and methods for 4D reservoir monitoring and characterization: SEG/EAGE DISC (Distinguished Instructor Lecture Course).

Clapp, M. L., 2005, Imaging under salt: illumination compensation by regularized inversion: PhD thesis, Stanford University.

Ebaid, H., M. Nasser, P. Hatchell, and D. Stanley, 2009, Time-lapse seismic makes a significant business impact at Holstein: SEG, Expanded Abstracts, **28**, 3810–3814.

Fomel, S., 2002, Applications of plane-wave destruction filters: Geophysics, **67**, 1946–1960.

Hale, D., 2007, Local dip filtering with directional Laplacians: CWP Project Reiew, **567**, 91–102.

Kühl, H. and M. D. Sacchi, 2003, Least-squares wave-equation migration for avp/ava inversion: Geophysics, **68**, 262–273.

Lefeuvre, F., Y. Kerdraon, J. Peliganga, S. Medina, P. Charrier, R. L'Houtellier, and D. Dubucq, 2003, Improved reservoir understanding through rapid and effective 4D: Girassol field, Angola, West Africa: SEG Technical Program Expanded Abstracts, **22**, 1334–1337.

Lumley, D. E., 1995, Seismic time-lapse monitoring of subsurface fluid flow: PhD thesis, Stanford University, http://sepwww.stanford.edu/public/docs/sep91/.

Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.

Plessix, R.-E. and W. Mulder, 2004, Frequency-domain finite-frequency amplitude-preserving migration: Geophysical Journal International, **157**, 975–985.

Rickett, J. E. and D. E. Lumley, 2001, Cross-equalization data processing for time-lapse seismic reservoir monitoring: A case study from the Gulf of Mexico: Geophysics, **66**, 1015–1025.

Tang, Y., 2008, Wave-equation Hessian by phase encoding: SEG Technical Program Expanded Abstracts, **27**, 2201–2205.

Valenciano, A. A., B. Biondi, and A. Guitton, 2006, Target-oriented wave-equation inversion: Geophysics, **71**, A35–A38.

Whitcombe, D. N., J. M. Marsh, P. J. Clifford, M. Dyce, C. J. S. McKenzie, S. Campbell, A. J. Hill, R. S. Parr, C. Pearse, T. A. Ricketts, C. P. Slater, and O. L. Barkved, 2004, The systematic application of 4D in BP's North-West Europe operations — 5 years on: SEG Technical Program Expanded Abstracts, **23**, 2251–2254.

Least-squares reverse time migration for the Cascadia ocean-bottom dataset

Mandy Wong, Shuki Ronen, and Biondo Biondi

ABSTRACT

We present a method based on least-squares reverse time migration (LSRTM) for imaging ocean-bottom data. We show that by using LSRTM, we not only enhance the resolution of the image but also suppress the migration artifacts. Furthermore, LSRTM also raise the relative amplitude of true reflectors in the subsurface. This method can also be extended to imaging with higher-order multiples. In certain geometries, LSRTM with multiples can further improve the image with a larger aperture. We demonstrate the concept and methodology in 2D and apply our proposed scheme to an ocean-bottom field survey located in the northern Cascadia continental margin.

INTRODUCTION

Ocean-bottom seismic (OBS) acquisition is an established technology in which seismometers are placed on the sea bottom and shots are fired at the sea surface. In areas congested by platforms or other obstacles, ocean-bottom seismic is advantageous, because it is operated by small boats without cumbersome towed streamers.

There are several different processing schemes for ocean-bottom data. The traditional method, inherited from surface seismic processing, is to remove all free-surface multiples and to migrate only with the primary signal (Wang et al., 2009). Therefore, initial work on OBS data processing has been dedicated to the removal of free-surface multiples.

While multiples are often treated as noise, they are formed by the same source signal as primaries, but travel along different paths in the medium. The receiver ghost, also known as the mirror signal, is the next order of reflection beyond the primaries, with an additional reflection off the sea surface. For a deep water OBS survey, the subsurface reflection point of the receiver ghost is located at a greater distance from the receiver station than the primary (Figure 1). Therefore, the mirror signal can provide wider subsurface illumination than the primaries. Several authors have used the mirror signal in migrating OBS data (Godfrey et al., 1998; Ronen et al., 2005; Grion et al., 2007; Dash et al., 2009).

While mirror imaging can greatly expand the illumination aperture of an OBS survey, a similar concept can be applied to image with even higher-order surface-related multiples. Figure 1 shows that the subsurface reflection point of the next order of surface-related multiples, denoted as the double-mirror signal, is at an even greater distance from the receiver station than the mirror signal. This translates to an even wider subsurface illumination for the entire survey.

In OBS acquisition, Muijs et al. (2007) made an early attempt to image primary and free-surface multiples together. This requires the data to be decomposed into up-going

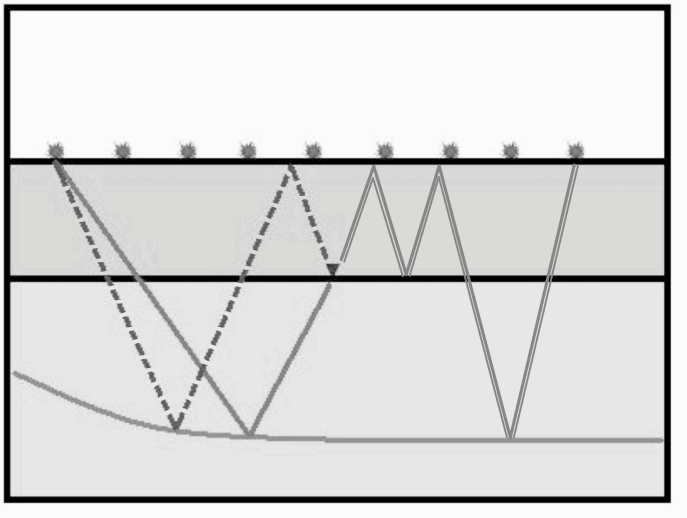


Figure 1: The subsurface reflection point of the double-mirror signal (in green double-line), the mirror signal (in black dashed-line), and the primary signal (in red solid-line). Note the double mirror image point is located at a greater distance from the receiver station than the mirror signal and the primary. [**NR**] mandy1/. illum2

and down-going constituents, followed by downward extrapolation and a 2D deconvolution-based imaging condition. While this technique is computationally efficient, its image contains crosstalk artifacts caused by interference from up-going and down-going waves not associated with the same subsurface reflector. In contrast to existing methods, we use least-squares reverse time migration to optimally combine structural information provided by each order of multiple reflection into images that are free from crosstalk.

It is the objective of the data processor to put everything together and not overload the interpreter with alternative images. In our earlier work (Wong et al., 2010), we have shown how the information in the primary and mirror signal can be combined to give one coherent image. In this paper, we focus on the improvement from performing LSRTM with the mirror signal and the subsequent higher-order multiples. Ultimately, we wish to perform joint inversion of the primaries and all the down-going surface related multiples. The rest of this paper is organized as follows. We first discuss the theory of linear least-squares inversion, we then apply the inversion scheme to a field survey located on the northern Cascadia continental margin.

THEORY

In general, pre-stack imaging of OBS acquisition is done in the common-receiver domain. For processing, the idea of reciprocity is applied, in which these common-receiver gather (CRG) are injected at the source location while the source wavelet is injected at the receiver location. We will first discuss how the migration operators are constructed. Afterwards, we will discuss how the least-squares migration problem is defined.

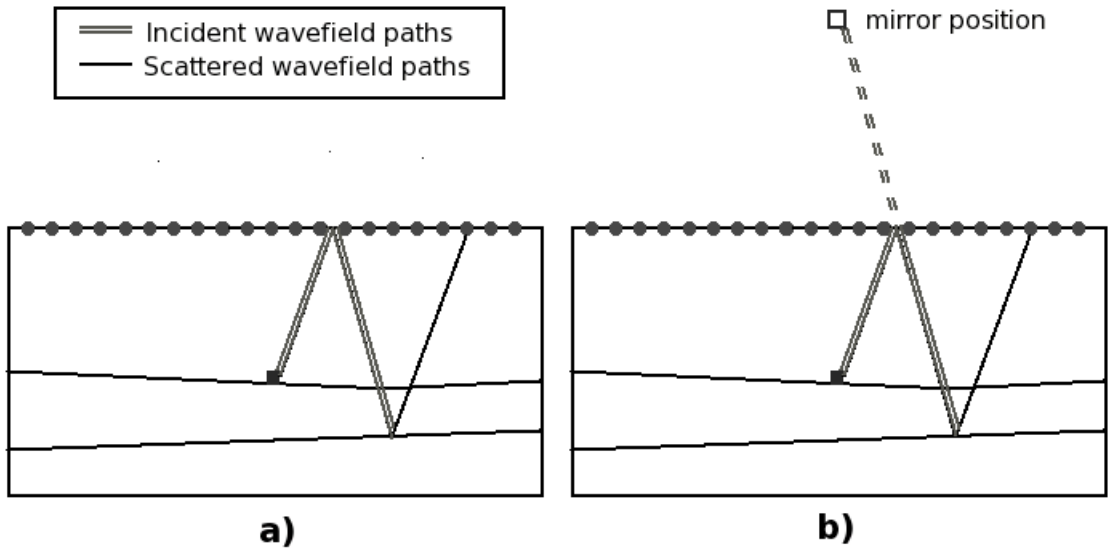


Figure 2: (a) The raypath of a mirror signal. (b) The raypath of the same signal in mirror imaging. The apparent position of the receiver is now at twice the ocean depth above the seabed. This assumes the sea surface is a perfect reflector. [**NR**] mandy1/. Figure2

Conventional mirror imaging

The essence of conventional mirror imaging is to account for the kinematics of the mirror signal (Figure 2 (a)) by placing the receiver position at twice the ocean depth above the seabed. If the sea surface is a perfect reflector with a reflectivity of -1, then the simulated mirror migration path (Figure 2 (b)) will be equivalent to the mirror signal.

One advantage with mirror imaging is that the mirror's illumination aperture is much wider than that of the primary signal. Another advantage is that mirror imaging can be easily adapted to existing migration algorithms by changing the receiver position.

Higher order mirror imaging

To image with higher-order multiples, we need to construct a migration operator that can account for the kinematics of the appropriate wave paths. To simulate the final down-going leg of the wave path, an areal shot is pre-calculated by first injecting the source wavelet at the receiver location, letting the wavefield propagate, and then capturing the signal at the sea surface (Figure 3). To generate the incident wavefield, the saved areal shot is re-injected at the sea-surface with a -1 factor. The re-injected signal is then allowed to travel back and forth in the water column using a reflecting top boundary and a well-defined velocity contrast at the sea-bottom. This algorithm can correctly simulate the wave paths traversed by the mirror signal, the double-mirror signal and even higher-order multiples.

Note that this technique does not migrate all orders of multiples. It only migrates surface-related multiples with a single reflection in the reflectivity model or with reflections

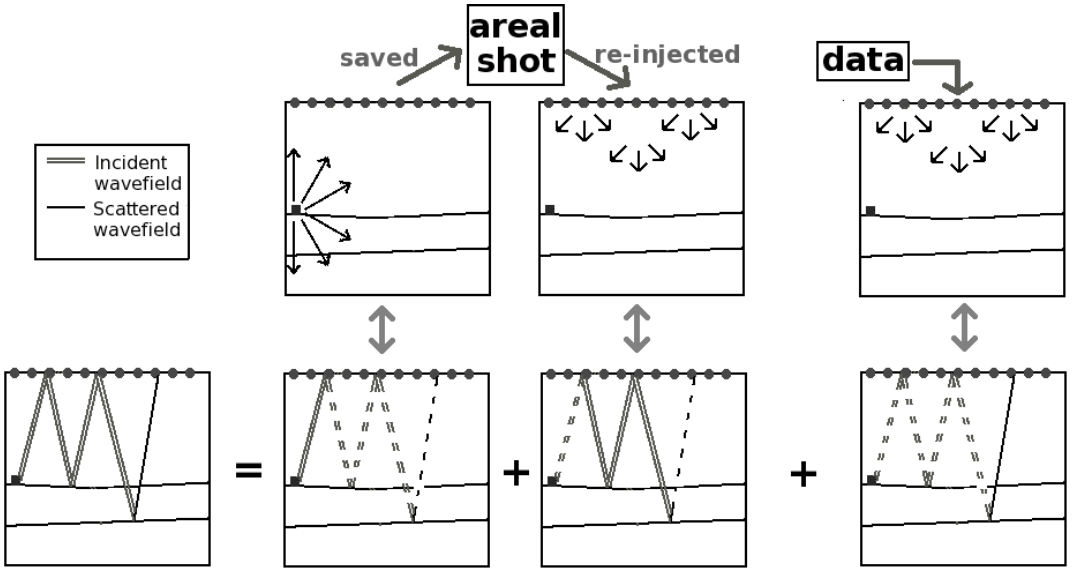


Figure 3: Illustration for the migration of the double-mirror signal. For the incident wavefield, an areal shot is pre-calculated to simulate the final down-going leg of the wave path. To generate the incident wavefield, the saved areal shot is re-injected at the sea-surface with a -1 factor. The re-injected signal is then allowed to travel back and forth in the water column using a reflecting top boundary and a well-defined velocity contrast at the sea-bottom. [**NR**] mandy1/. Figure3

in the acoustic modeling based on a sharp boundary in the velocity model.

Least-squares reverse time migration (LSRTM)

In OBS acquisition, pre-stack images are created from data in the common-receiver domain. This intrinsically requires that each trace in the CRG is de-signatured. In RTM, the migration image is a linear operator applied to the recorded data:

$$\mathbf{m}_{mig}(\mathbf{x}) = \sum_{\mathbf{x}_r,\mathbf{x}_s,\omega} \omega^2 f_s^*(\omega) G^*(\mathbf{x}_r,\mathbf{x},\omega) G^*(\mathbf{x},\mathbf{x}_s,\omega) d(\mathbf{x}_r,\mathbf{x}_s,\omega) \tag{1}$$

where ω is frequency, $m(\mathbf{x})$ represents the reflectivity at image point $\mathbf{x}$, $f_s(\omega)$ is the source waveform, and $G(\mathbf{x}_s,\mathbf{x},\omega)$ is the Green's function that is the solution to the two-way acoustic constant-density equation. In practice, the Green's function is calculated using a finite-difference time-domain technique, and the multiplication in the frequency domain is replaced by a zero-lag cross-correlation in the time domain.

The difference between mirror imaging and higher-order mirror imaging lies within the incident wavefield, $U_s(\mathbf{x}_r,\mathbf{x},\omega) = \omega^2 f_s(\omega) G(\mathbf{x}_r,\mathbf{x},\omega)$, which is calculated differently to the method described in the previous two sub-sections.

To obtain a better reflectivity image, we go beyond migration by formulating the imaging problem as a least-squares inversion problem. The solution $m_{inv}(\mathbf{x})$ is obtained by

minimizing the objective function $S(\mathbf{m})$, which is defined as the least-squares difference between the forward modeled data $\mathbf{d}^{mod}$ and the recorded data $\mathbf{d}^r$.

$$S(\mathbf{m}) = \| \mathbf{d}^{mod} - \mathbf{d} \|_2 = \| \mathbf{Lm} - \mathbf{d} \|_2 \tag{2}$$

In least-squares reverse time migration (LSRTM), the forward modeled data is defined to be the Born approximation of the linearized acoustic wave equation:

$$d^{mod}(\mathbf{x}_r, \mathbf{x}_s, \omega) = \sum_{\mathbf{x}} U_s(\mathbf{x}_r, \mathbf{x}, \omega) G(\mathbf{x}, \mathbf{x}_r, \omega) \mathbf{m}(\mathbf{x}) \tag{3}$$

It is important to point out that the forward modeling operator $\mathbf{L}$ is the adjoint of the reverse-time migration operator $\mathbf{L}^T$. Even for the case of modeling certain classes of surface-related multiples, the operator is still linear with respect to $m(\mathbf{x})$. This is because $\mathbf{L}$ simulates only events that would interact with the model space once.

FIELD EXAMPLE

Next we present results of LSRTM on an OBS survey located at the northern Cascadia continental margin offshore of western Canada. The area is known to contain gas hydrates, which have a characteristic structure known as the bottom-simulating reflector (BSR), that marks the base of the hydrate stability zone. OBS data were collected along five parallel lines normal to the margin (Figure 4). Line spacing was 500 m with ten ocean-bottom seismometers deployed with 100 m spacing at a water depth of about 1300 m.

Pre-processing

In our 2D study, we extract a 2D shot line from directly above the OBS line. Following the processing step in Dash et al. (2009), we apply a gapped deconvolution with a 12-ms gap length and 300-ms filter length to suppress the source bubble. Out of the ten OBS receivers, only six were used for final imaging. These receivers were selected according to their data quality. Three receivers were rejected due to large tilt angles, and one receiver was rejected because it was saturated with noisy traces.

Since we only use the down-going wavefield, only acoustic separation above the seafloor is needed. We used the adaptive decomposition method of Schalkwijk et al. (1999).

These data are bandpassed between 5 and 45 Hz to avoid dispersion in the time-domain finite-difference calculation. These data before and after pre-processing are shown in Figure 5.

Figure 6 (a) shows the migration velocity used for our RTM and LSRTM. The velocity ranges from 1480m/s in the water column to 1820m/s in the sediment layer. The BSR lies at an approximate depth of 1500-m with a velocity inversion.

Migration and Inversion result

Figure 6 (b) shows the mirror-imaging RTM result. Overall, we can identify some migration artifacts associated with RTM, such as the high-amplitude low-frequency artifacts. The

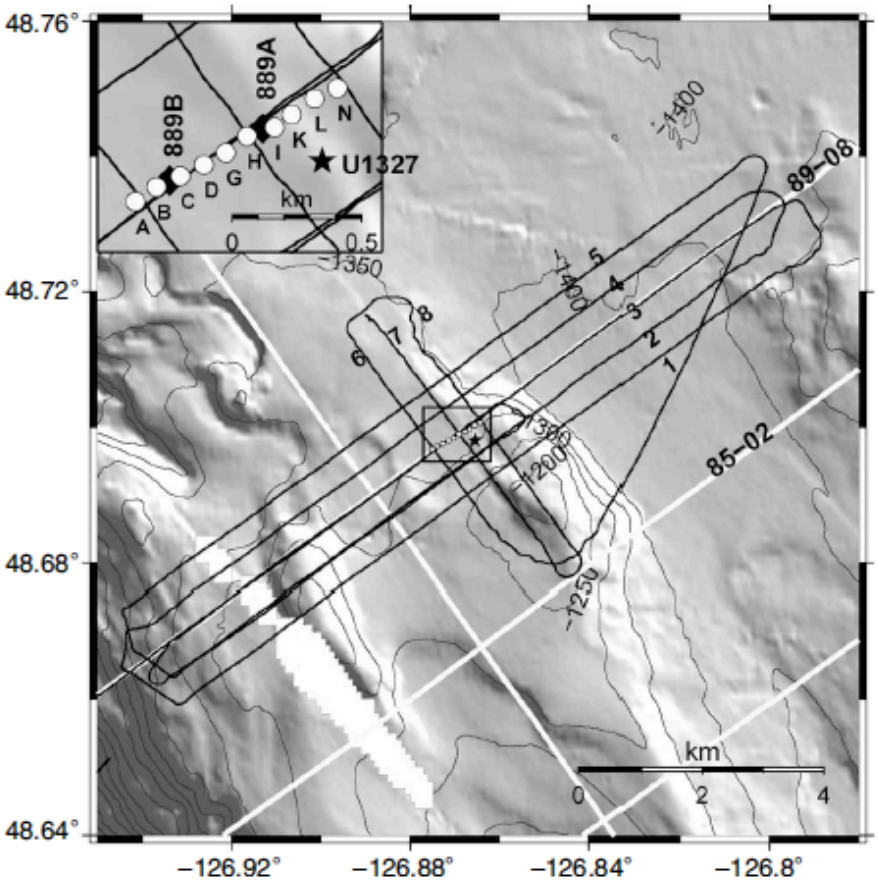


Figure 4: The geometry of the Northern Cascadia dataset. Ten ocean-bottom seismometers were deployed with 100 m spacing at a water depth of 1300 m. The shot line spacing was 500 m. [**NR**] mandy1/. geo

resolution is low, and the structural information below the BSR is generally difficult to identify.

Seismic inversion is an ill-posed problem. To prevent divergence to unrealistic solutions, we add an additional term to our objective function $S(\mathbf{m})$. In Nemeth et al. (1999); Ronen and Liner (2000), data weighting operators are used to remove the acquisition footprint by having zero weighting on regions corresponding to the acquisition gap. The overall fitting goal becomes

$$S(\mathbf{m}) = \| W(L\mathbf{m} - \mathbf{d}) \|^2 + \epsilon^2 \| \mathbf{m} \|^2, \tag{4}$$

where $\mathbf{W}$ is a diagonal data-weighting matrix, and ϵ tunes the level of damping in our objective function. The criterion for choosing ϵ is:

$$0.2 = \frac{\| \epsilon\mathbf{m} \|^2}{\| \mathbf{W}(L^T(\mathbf{d}^{mod} - \mathbf{d})) \|^2}. \tag{5}$$

In the first iteration, ϵ is defined by limiting the gradient contribution from the damping term to be 20% of the gradient contribution from the data fitting term. From experiments

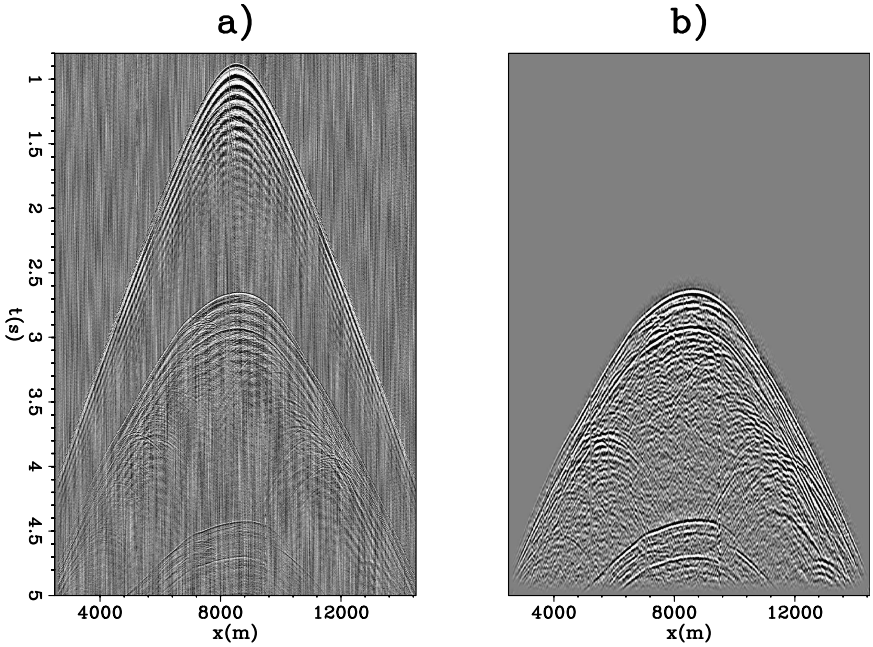


Figure 5: (a) A hydrophone common receiver gather (CRG) before pre-processing and (b) down-going data after pre-processing. [**CR**] mandy1/. dataCom

with both synthetic and field dataset, we found that this limit yields satifactory results. For this dataset, the damping term alone seems to adequately regularize the inversion and yield a satisfactory result. For more complex survey regions, other regularization terms are needed to further constrain the inversion.

We apply LSRTM in two cases: (1) using the mirror-imaging operator (LSRTM-M) and (2) using the higher-order mirror-imaging operator (LSRTM-HM). The results after 10 iterations are shown in Figure 6 (c) and (d) respectively. Overall, the inversion images have higher resolution and fewer artifacts than the RTM image. However, there is not much difference between the two inversion results. One reason is that this dataset records only a small portion of the double-mirror signal. We expect the LSRTM-HM to provide substantial improvement in the case of a shallow-water dataset in which higher-order multiples overlap with lower order multiples. Comparing between RTM and LSRTM, we have identified two areas of improvement with some close-up sections in Figure 7:

1. Panels (a) and (b) show the RTM and LSRTM-M images from a section at x=9400 to 11600 m and z=1350 to 1700 m. We can see that the reflectors under the seabed are better resolved with LSRTM.

2. Panel (c) and (d) shows the RTM and LSRTM-M images from a section at x=7500 to 9500 m and z=1400 to 1750 m. For regions beneath the BSR, the amplitude of the true reflector has been strengthened, and we can now identify some structures there.

Convergence

We use the iterative conjugate direction method to minimize our objective function. Figure 8 (a) shows the value of the objective function over 20 iterations. The inversion converges nicely with the residuals dropping to 50% of the original value. It is important to point out that the final residual drops is problem dependent. In general, if the raw dataset contains a low signal-to-noise ratio, then the residual drop will be less. Figure 9 (a) shows a common receiver gather for this data set. After 20 iterations, the corresponding forward modeled data 9 (b) is able to capture most of the prominent signal in the data. The difference of the two yields the data residual shown in Figure 9 (c).

Another way to evaluate convergence is to study the ratio of the model residual over data residual. The model residual is defined to be applying the migration operator to the data residual at each iteration. (It is also called the gradient.) In Figure 8 (b), we can see that the ratio almost drops down to zero over iterations, which also indicates convergence.

DISCUSSION AND FUTURE WORK

Although our modeling operator in LSRTM-HM can generate higher-order surface-related multiples, we do not see major improvement in the inverted images. One reason is related to the fact that there is limited signal coming from the double-mirror event in our recorded data. This dataset only has recording up to five seconds which results in some truncation in the double-mirror signal. It will be interesting to test on datasets with either a longer recording time or a shallower sea bottom. Another reason is related to the constant density assumption in our ascoutic modeling. In Figure 9 (b), we can see that the double-mirror signal is generated in the modeling. However, its amplitude is significantly weaker than that in the field dataset (Figure 9 (a)). With a constant density propagation kernel, the higher-order multiples are generated solely from reflection off sharp contrasts/interfaces in the migration velocity model. For the Cascadia dataset, it turns out that the seabed's reflectivity is attributed significantly to the density contrast, which is not accounted for in our modeling. As an result, the velocity constrast alone does not account for the full reflectivity of the sea bottom. We are currently trying to match the correct reflectivity by implementing our LSRTM using the full ascoutic wave equation. That is, without the constant density assumption. Nearby well log data will be used to estimate the density constrast at the sea bottom.

The computational cost of LSRTM is higher than that of RTM, with a factor proportional to the number of iterations. For an OBS survey, such additional computational costs can still be affordable. This is because the number of pre-stack migrations needed equals the number of OBS receivers in the survey. The number of migration needed is substantially smaller than that in a towed-streamer survey. Recently, Dai et al. (2010) suggested using phase encoding in LSRTM, which can reduce the computation substantially but introduces crosstalk into the image.

CONCLUSION

Although only 2D modeling and migration are used, we see improvements in the LSRTM image over conventional methods. Such improvements include suppression of migration

artifacts, enhancement of amplitudes along true reflectors, and better resolution. We expect the extension to 3D will bring further improvements.

ACKNOWLEDGMENTS

The authors are thankful to Professor George Spence from the University of Victoria and the Dalhousie University for collection and permission to publish the OBS data. MW thanks Ranjan Dash for information and previous work on the field dataset. We thank the sponsors of the Stanford Exploration Project for their financial support.

REFERENCES

Dai, W., C. Boonyasiriwat, and G. T. Schuster, 2010, 3d multi-source least-squares reverse time migration: SEG Expanded Abstracts, **29**, 3120–3124.

Dash, R., G. Spence, R. Hyndman, S. Grion, Y. Wang, and S. Ronen, 2009, Wide-area imaging from obs multiples: Geophysics, **74**, Q41–Q47.

Godfrey, R., P. Kristiansen, B. Armstrong, M. Cooper, and E. Thorogood, 1998, Imaging the foinaven ghost: 68th SEG Annual Meeting Expanded Abstracts, 1333–1335.

Grion, S., R. Exley, M. Manin, X. Miao, A. Pica, Y. Wang, P. Granger, and S. Ronen, 2007, Mirror imaging of obs data: First Break, **25**, 37–42.

Muijs, R., J. Robertsson, and K. Holliger, 2007, Prestack depth migration of primary and surface related multiple reflections: Part i - imaging: Geophysics, **72**, 59–69.

Nemeth, T., W. Chengjun, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221.

Ronen, S., L. Comeaux, and J. Miao, 2005, Imaging downgoing waves from ocean bottom stations: 75th SEG Annual Meeting Expanded Abstracts, 963–966.

Ronen, S. and C. Liner, 2000, Least-squares dmo and migration: Geophysics, **65**, 1364–1371.

Schalkwijk, K. M., C. P. A. Wapenaar, and D. J. Verschuur, 1999, Application of two-step decomposition to multicomponent ocean bottom data: Theory and case study: Journal of Seismic Exploration, **8**, 261–278.

Wang, Y., S. Grion, and R. Bale, 2009, What comes up must have gone down: the principle and application of up-down deconvoluiton for multiple attenuation of ocean bottom data: CSEG Recorder, **34**, 16–20.

Wong, M., B. Biondi, and S. Ronen, 2010, Joint least-squares inversion of up- and down-going signal for ocean bottom data sets: SEG Annual Meeting Expanded Abstracts, **29**, 2752–2756.

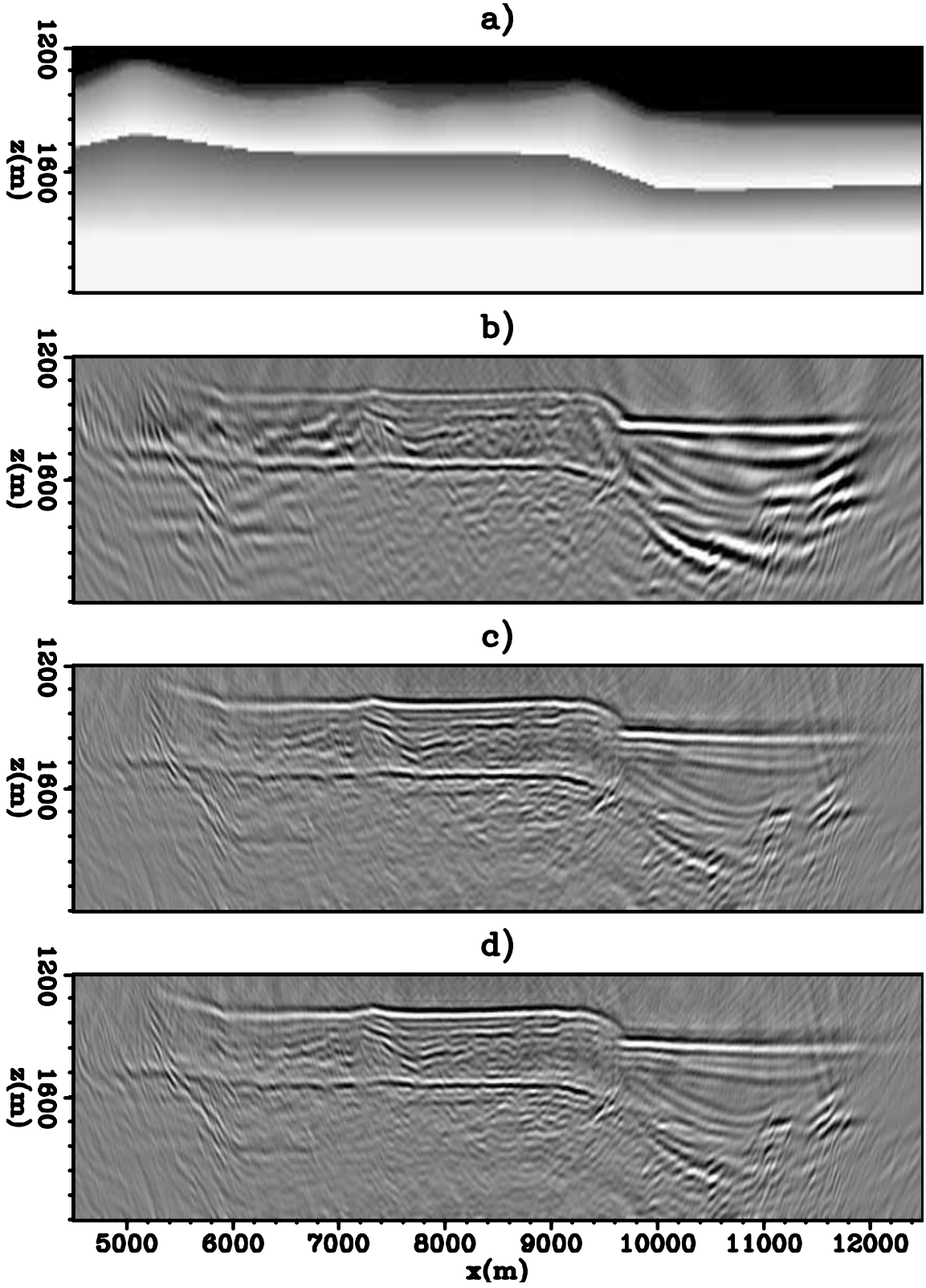


Figure 6: (a) Migration velocity model, (b) RTM image of the mirror signal, (c) LSRTM of the mirror signal (LSRTM-M), and (d) LSRTM of mirror and higher-order mirror signals (LSRTM-HM). [**CR**] mandy1/. RTMinv

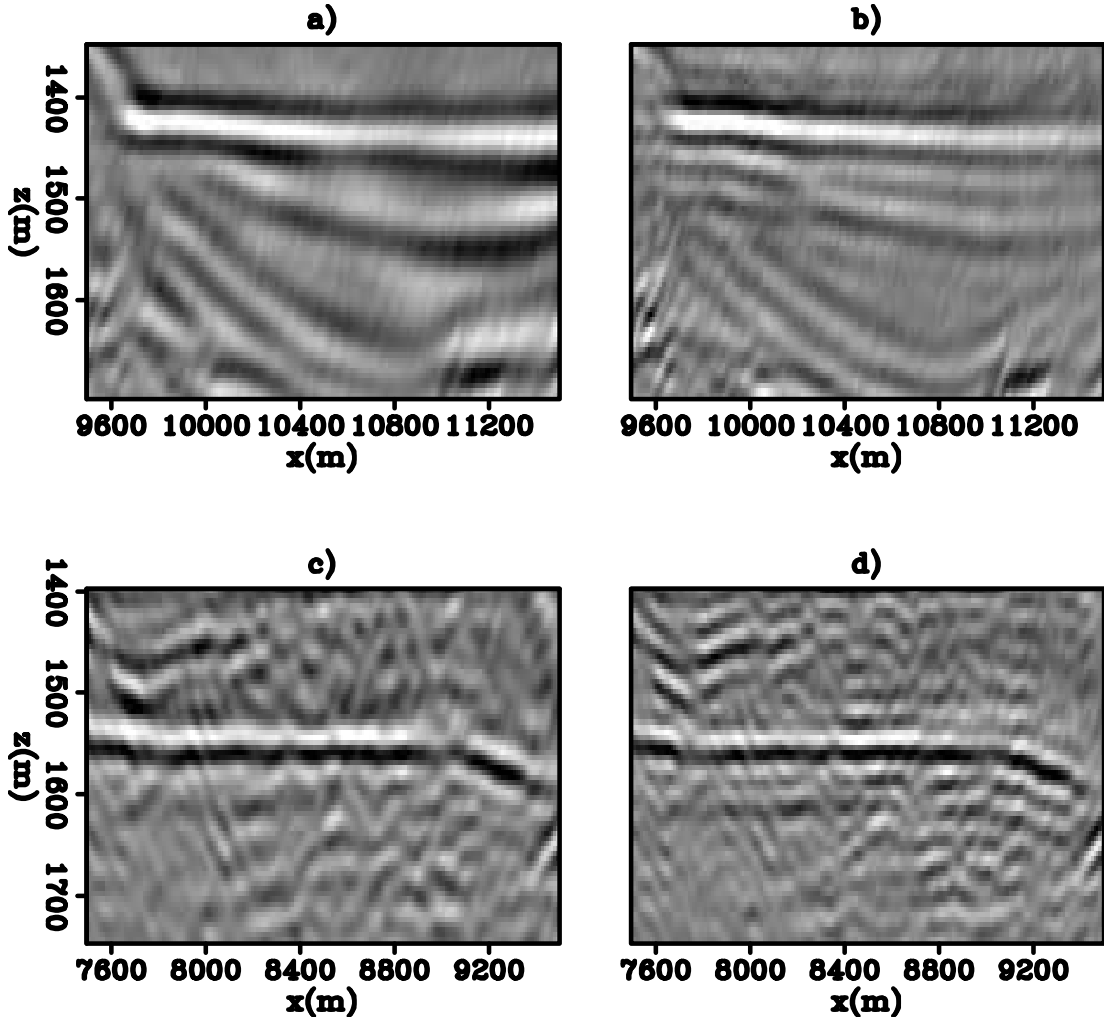


Figure 7: A section of the image cut from x=9400-11600 m and z=1350-1700 m after applying (a) RTM and (b) LSRTM-M. Another section of the image cut from x=7500-9500 m and z=1400-1750 m after applying (c) RTM and (d) LSRTM-M. [**CR**] mandy1/. Zoom

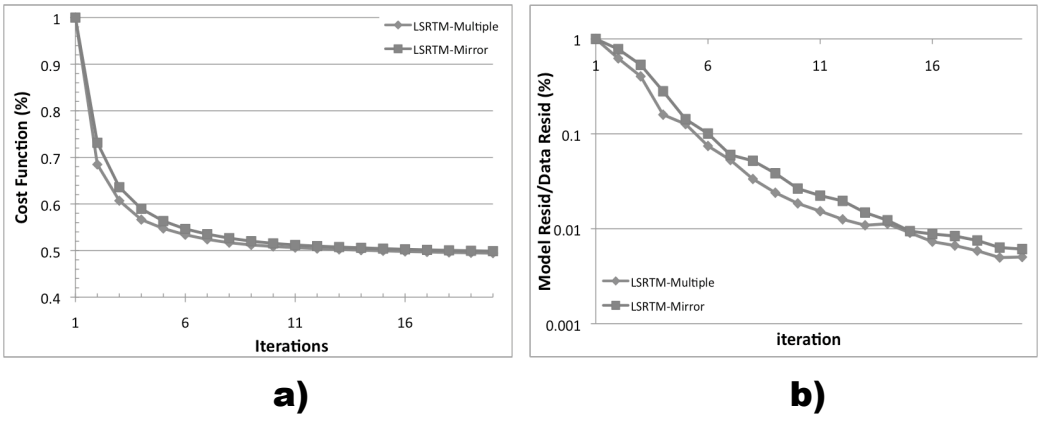


Figure 8: (a) Objective function over 20 iterations and (b) The ratio of model residual over data residual over 20 iterations (Note in log scale). [**NR**] mandy1/. ResidualPlot

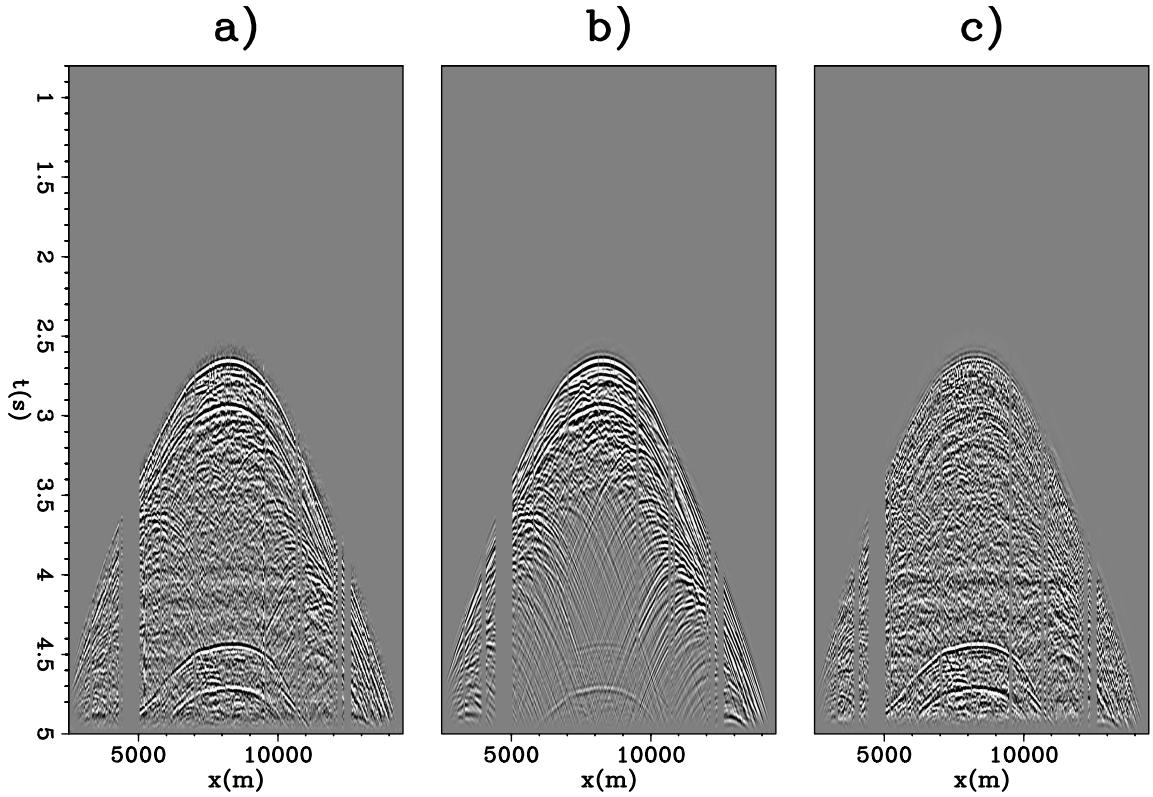


Figure 9: (a) A common receiver gather after pre-prosessing for one of the ocean-bottom seismometer, (b) the forward modeled data of the same gather using the LSRTM-M reflectivity image at the 20th iteration and(c) the data residual of the same gather at the 20th iteration, which is equivalent to the difference between (a) and (b). [**CR**] mandy1/. DatspaceResid

Imaging using compressive sensing

Robert G. Clapp

ABSTRACT

Constructing 3-D angle gathers is quickly becoming the computational bottleneck in Reverse Time Migration (RTM). Both the expansion in volume size (120-900x) and the cost of cross-correlating wavefields greatly increase the cost of RTM. Both the storage and computational cost can be greatly reduced by reformulating the imaging process as a compressive sensing problem. Preliminary results show that this approach holds promise but finding an acceptable L-1 inversion approach is still problematic.

INTRODUCTION

Reverse Time Migration is quickly becoming the standard high-end imaging technique. The last few years has seen numerous papers on how to speed up the finite-difference kernel on various platforms (Micikevicius, 2009; Nguyen et al., 2010; Nemeth et al., 2008; Clapp et al., 2010). These techniques was sufficient when RTM was simply used to produce a final image.

When RTM is used to construct angle gathers for velocity or rock property analysis, the finite difference kernel becomes a secondary concern. The construction of angle gathers, particularly 3-D angle gathers, through sub-surface offset correlation (Sava and Fomel, 2006) or time-shift gathers (Sava and Fomel, 2003) becomes the dominant cost. Some have proposed reducing the cost of 3-D angle gathers by constructing angle gathers along only a few azimuths. While these techniques are significantly less costly than full 3-D angle gathers, they are still expensive and not ideal. Compressive sensing (Donoho, 2006) offers a potential solution to this computation and storage problem. In compressive sensing, a random sub-set of the desired measurements are made. An inversion problem is then set up to estimate in an ℓ_1, or preferably ℓ_0, sense, a sparse basis function that fully characterizes the desired signal. For compressive sensing to work, a signal must be highly compressible. For compressive sensing to be worthwhile, the cost of inverting for the basis function must be significantly less than the cost of acquiring the full signal.

In this paper, I show how angle gather construction fits the criteria for compressive sensing. I demonstrate how angle gathers are highly compressible in the multi-dimensional wavelet domain. Further, I demonstrate how the cost of constructing a sub-set of the sub-surface offsets and then performing an ℓ_1 inversion is significantly less expensive.

COMPRESSIVE SENSING

Compressive sensing is a statistical technique whose start is usually traced back to Donoho (2006), but whose start could be placed as early as the basic pursuit work of Mallat and Zhang (1993). A compressive sensing problem at its heart is a special case of a missing data

problem. In geophysics, we often think of a missing data problem as solving for a model $\mathbf{m}$ given some data $\mathbf{d}$ which exist in the same vector space. We have a masking operator $\mathbf{M}$ (1 where the data is known, 0 elsewhere). We add in some knowledge of the covariance of the model through a regularization operator $\mathbf{A}$. We then estimate the best model from the following system of equations in a ℓ_2 sense,

$$\begin{aligned} \mathbf{0} \approx \mathbf{r_d} &= \ell_\mathbf{2}(\mathbf{d} - \mathbf{Mm}) \\ \mathbf{0} \approx \mathbf{r_m} &= \ell_\mathbf{2}(\mathbf{Am}), \end{aligned} \tag{1}$$

where $\mathbf{r_d}$ and $\mathbf{r_m}$ are the result of taking the ℓ_2 norm of the first and second equations. The success of this approach relies on the accuracy of $\mathbf{A}$ to describe the covariance of the model.

Compressive sensing approaches the problem from a different perspective. It starts from the notion that there exists a basis function that $\mathbf{d}$ can be transformed into through the linear operator $\mathbf{L}'$ in which very few non-zero elements are needed to represent the signal. The compressive sensing approach is then to set up the missing data problem in two phases. First, estimate the elements of the sparse basis function $\mathbf{m}$ through,

$$\mathbf{0} \approx \mathbf{r} = \ell_\mathbf{1}(\mathbf{d} - \mathbf{MLm}), \tag{2}$$

where we are now estimating $\mathbf{m}$ in the $\ell_\mathbf{1}$ sense. We can then apply $\mathbf{L}$ to recover the full model. Compressive sensing makes the claim that if the signal can be represented in this basis function by n points, if you collect bm *randomly sampled* points, where b is typically 4-5, significantly less than what the Nyquist-Shannon criteria would suggest.

Figure 1 shows an example of this technique applied to a 2-D missing data problem. In this case, we are trying to recover the image seen in Figure 1(a). We start from the data points seen in Figure 1(b), and recover the image seen in Figure 1(c). In this case, the image was subsampled by a factor of 8 and nearly perfect recovery was achieved.

(a) (b) (c)

Figure 1: The original image (a), the image sub-sampled by a factor of 8 (b), and the reconstructed image (c) using compressive sensing. Baraniuk et al. (2008). [**NR**] bob1/. before,random,after

Reconstruction using compressive sensing techniques is expensive. ℓ_1 solvers are significantly more expensive than their ℓ_2 counterparts, which in themselves can represent a

significant cost. As a result, only certain classes of problems benefit from compressive sensing techniques. For compressive sensing to be useful, the cost of acquiring the full dataset must be significant. In addition, the signal must be highly compressible. The following two sections address both criterion.

IMAGING CONDITIONS COSTS IN RTM

There are many different imaging condition choices in RTM, that can have dramatically different cost profiles depending on the compute engines' memory structure and size, along with implementation method. All RTM imaging conditions start from a source $s(x, y, z, t)$ and receiver $g(x, y, z, t)$ wavefields. Generally some kind of correlations of these two wavefields is done to produce the RTM image. The data handling problem becomes challenging because the source wavefield is propagated forward in time while the receiver wavefield is propagated backward in time. As a result, one of the wavefields is stored to disk, or a checkpointing or random boundary condition is applied.

Until recently, a zero-time, zero-subsurface offset imaging condition was the most commonly implemented method. This amounts producing an image $I(x, y, z)$ by

$$I(x, y, z) = \sum_{\text{shot}} \sum_{t} s(x, y, z, t) g(x, y, z, t). \tag{3}$$

With this imaging condition, we normally store $I(x, y, z)$ at a relatively low cost memory hierarchy (it exists in RAM on a CPU, in global memory in the GPGPU, or attached RAM on a FPGA). For simplicity, let's assume that $g(x, y, z, t)$ is written to disk in a non-compressed form and then read backwards during the source propagation. Given that we often need to propagate six to eight time steps between imaging steps due to stability constraints, to avoid an IO bottleneck we need the cost of reading the receiver wavefield to be 1/6 the cost of our finite difference stencil. As of this writing, we are nearing a cross-over point where this will no longer be valid.

The relative cost of implementing the actual imaging condition is a little trickier. For this exercise, let's assume that we are implementing acoustic RTM on a GPGPU (the relatively simple memory model simplifies the calculation). Let's assume a naive second-order in time, 14th order in space approach. In this case, at each sample we need to read in the previous wavefield, the velocity, and store the updated wavefield. In addition, we will access 43 values in the current wavefield, but because these are stored in shared memory they will have approximately $\frac{1}{10}$ latency. Therefore, the cost of propagating the wavefield eight times between imaging steps will be $nx * ny * nz * 8 * (3 + 43 * .1)$. The imaging step at each point will require the reading of the source and receiver wavefield from global memory, and the reading and writing of the image, a cost of $nx * ny * nz * 4$, or a magnitude less than the propagation cost.

Alternate imaging condition choices such as offset-gathers(Sava and Fomel, 2003), time-shift gathers(Sava and Fomel, 2006), and extended image gathers(Sava, 2007), dramatically alter this balance. For conciseness I will limit this discussion to sub-surface offset gathers, but similar limitations apply to all of the choices. To construct sub-surface offset gathers used to form angle gathers, the image, in its most general form, expands by one to three dimensions. In the extreme case, the image becomes $I(x, y, z, h_x, h_y, h_z)$, where h_x, h_y, and

h_z represent lags in x, y, and z. The imaging condition then becomes

$$I(x,y,z,h_x,h_y,h_z) = \sum_{\text{shot}} \sum_{t} s(x+h_x, y+h_y, z+h_z) g(x-h_x, y-h_y, z-h_z). \quad (4)$$

In this formulation, the cost of propagating the wavefield and reading the receiver wavefield from disk is the same as the zero-time, zero-offset imaging condition. The IO requirements and computational cost of the imaging steps changes dramatically. The image is now 20 (in the case where we only evaluate h_x) to several thousand times larger. As a result, it must now be read and written to disk at every imaging step. If we assume that we are constructing 400 sub-surface offsets at each point, the computational cost increases dramatically. We now benefit from transferring the source and receiver wavefields to shared memory, reducing their cost by possibly up to 80%, but we still end up with a cost of $nx*ny*nz*800*(2*.2+2)$, more than a magnitude more expensive than the propagation. However, that is not the real problem. We must read and write each image point at each image step, resulting in an increased IO requirement of 800x. As a result, only approaches that construct several different azimuths of 1-D angle gathers have been generally shown.

Sub-sampling sub-surface offsets offers benefits in both computation and storage. The reduction in offset calculation proportionally reduces the storage on large memory machines, potentially eliminating the need to go to disk. Total computation is reduced by sub-sampling, but given the required random nature of compressive sensing, cache misses will not be proportionally reduced.

COMPRESSIBILITY

There is significant literature on compressing seismic data. Relatively low compression ratios are achievable by compressing a trace. Significantly higher compression ratios are achieved by multi-dimensional approaches. Generally, the best results have used either multi-dimensional wavelets (Mallat, 1999) or its successor curvelets (Candes and Donoho, 1999). Villasenor et al. (1996) showed that compression ratios of 100:1 were achievable by compressing a 4-D volume (t, h_x, h_y, s). Further, Villasenor et al. (1996) states that the header information was the limiting factor in achieving even higher ratios.

Sub-surface offset gathers potentially represent an even higher, up to six, dimensional problem. To test compressibility, I used a 4-D volume (x, y, z, h_x) of dimensions (32,32,400,64). Figure 2 shows one of these sub-surface offset gathers and its neighbors, note the similarly. Following Villasenor et al. (1996), I chose the 9/7 bi-orthonormal transform(Antonini et al., 1992) used in JPEG compression. Figure 3 shows the resulting transform space and a histogram of the absolute values. I then used several different thresholds (throwing away 90%, 95%, 98%, and 99% of the data in the wavelet domain). Figure 4 shows the result of transforming this thresholding volume back into the space-domain. The resulting images are near-perfect at 95% and potentially acceptable at 98%. This translates into an acceptable compression ratio of approximately 30:1.

We get a more interesting result if we look at the compressibility of angle gathers. Figure 5 shows the same locations as Figure 2 but in the angle domain. Figure 6 shows the resulting wavelet domain space and a histogram of its absolute values. Note an even greater clustering around 0. Figure 7 shows the result of thresholding the wavelet domain

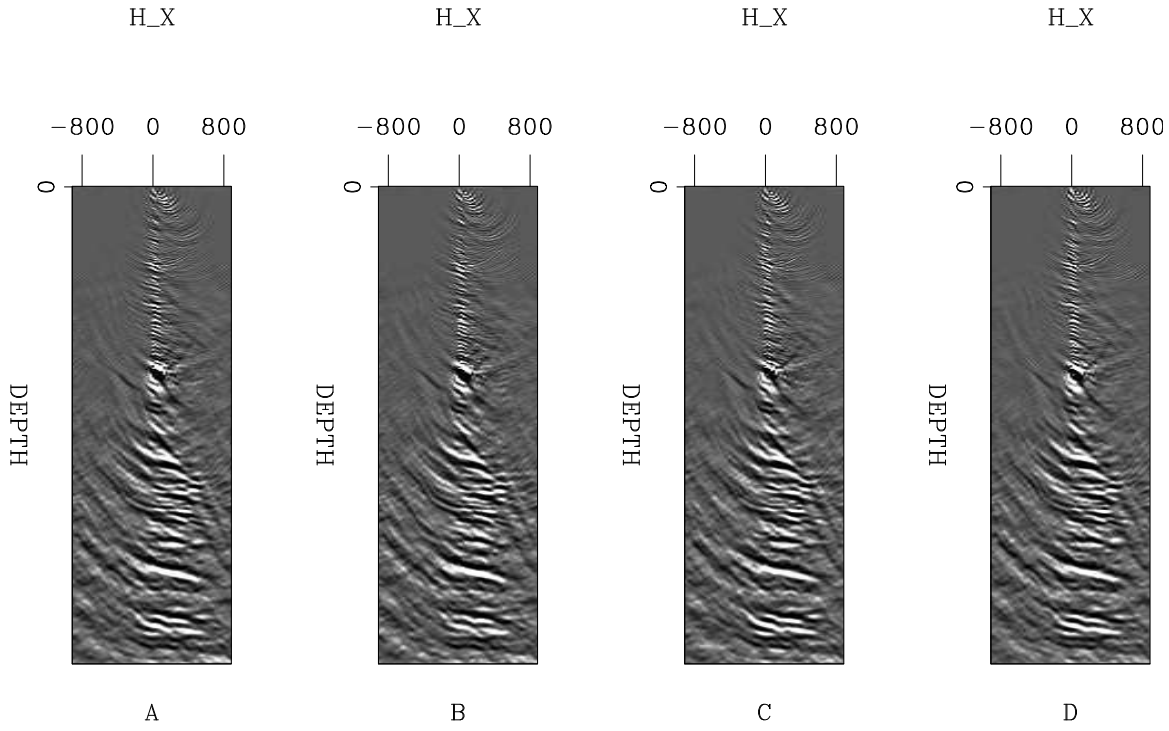


Figure 2: Five sub-surface offset gathers. B and C are one midpoint in X before and after A. E is one midpoint in Y before A. Note the spatial similarity, which lends itself to compression. [**ER**] bob1/. raw

at 95%, 98%, 99%, 99.5%. Note how we see near perfect reconstruction of the angle gather up to 99% and potential up to 99.5%. This is a 100:1 to 200:1 compression ratio (which translates into 20 to 50x reduction in sub-surface offsets).

Framing the compression in terms of angle gathers requires a change in formulation. The **L** in equation 2 now becomes chain of two operators. The first translates from sub-surface offset to angle. The second is multi-dimensional wavelet transform.

DISCUSSION AND CONCLUSIONS

Today, angle gather construction is acomputational bottleneck in terms of computation, cache misses, and IO. Compressive sensing, which reconstructs a sub-sampled signal by a ℓ_1 inversion of the data transformed into a sparse basis function, offers a potential solution. I showed that the angle gather constructions meets the two criterion for compressive sensing: the data is highly compressible using multi-dimensional wavelets and reducing the data size dramatically reduces computational cost. The next step in this work is to apply an ℓ_1 solver to reconstruct a sub-sampled offset gather.

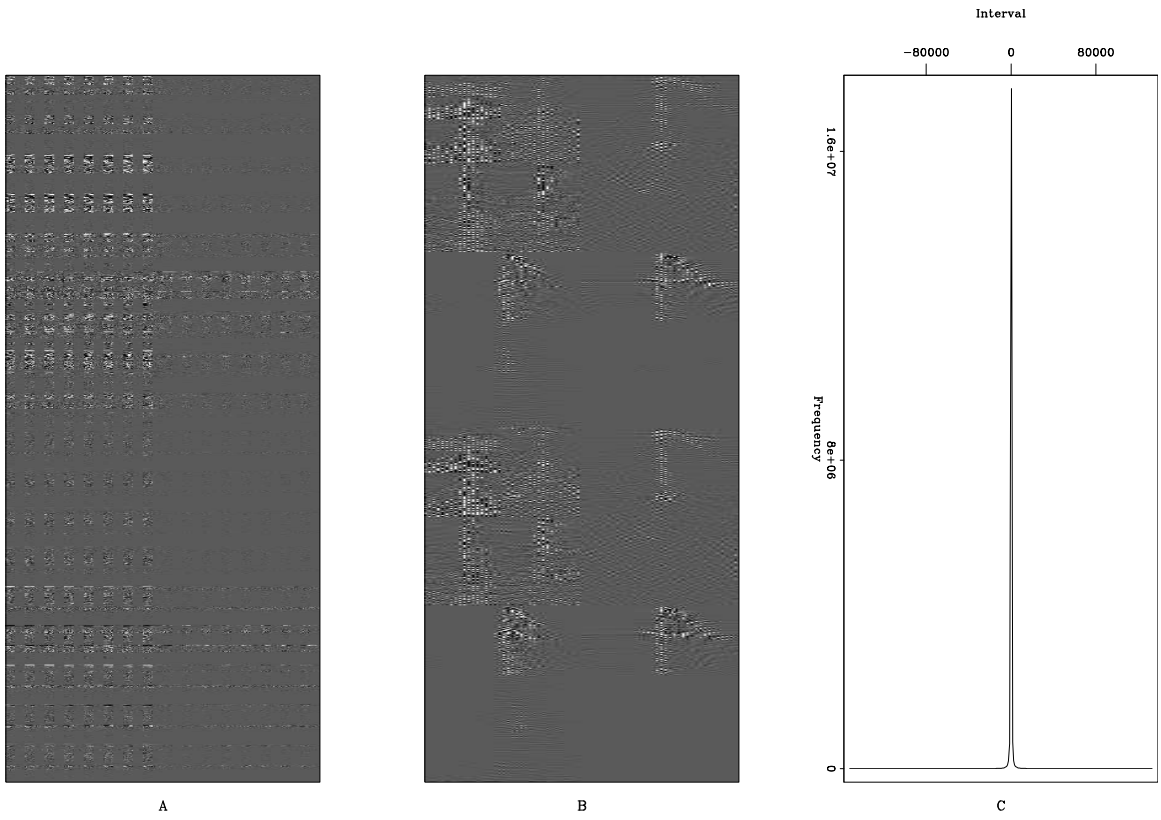


Figure 3: Panel A shows the wavelet domain representation of the 4-D volume used in this experiment. Panel B shows a zoom into portion of the wavelet domain. Panel C shows a histogram of the wavelet domain values. Note how the vast majority of the values are nearly zero. [**ER**] bob1/. wavelet1

REFERENCES

Antonini, M., M. . Barlaud, P. Mathieu, and I. Daubechies, 1992: IEEE Trans. on Image Proc., 205–220.

Baraniuk, R., J. Romberg, and M. Wakin, 2008, Tutorial on compressive sensing: `http://en.pudn.com/downloads154/sourcecode/zip/detail685097_en.html`.

Candes, E. J. . and D. L. Donoho, 1999, Curvelets - a surprisingly effective nonadaptive representation for objects with edges: Curves and Surfaces.

Clapp, R. G., H. Fu, and O. Lindtjorn, 2010, Sellecting the right hardware for reverse time migration: Leading Edge, **29**, 48–58.

Donoho, D. L., 2006, Compressive sensing: Information Theory, IEEE Transactions on, **52**, 1289–1306.

Mallat, S. G., 1999, A wavelet tour of signal processing: Academic Press.

Mallat, S. G. and Z. Zhang, 1993, Matching pursuits with time-frequency dictionaries: IEEE Transactions on Signal Processing, **41**, 3397–3415.

Micikevicius, P., 2009, 3d finite difference computation on gpus using cuda: GPGPU, **2**.

Nemeth, T., J. Stefani, W. Liu, R. Dimond, O. Pell, , and R. Ergas, 2008, An implementation of the acoustic wave equation on fpgas: 78th Ann. Internat. Meeting, Expanded Abstracts, 2874–2877, Soc. Expl. Geophys.

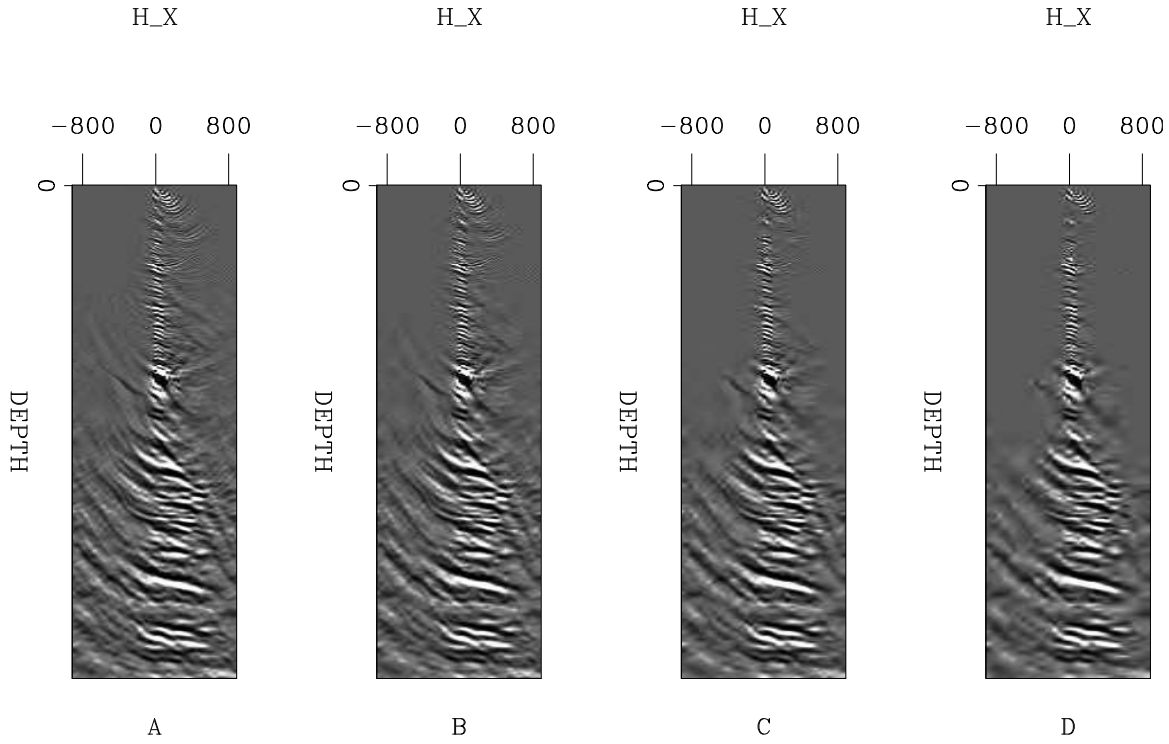


Figure 4: The result of zeroing the smallest values of the wavelet domain representation shown in Figure 3A. All four panels show the same sub-surface offset gather shown in Figure 2. A shows the result of clipping 90% of the values; B, 95%; C, 98%; and D,99%. Note how the reconstructed gather is nearly identical up to a 98% clip. [**ER**] bob1/. offset

Nguyen, A., N. Satish, J. Chhugani, C. Kim, and P. Dubey, 2010, 3.5-d blocking optimization for stencil computations on modern cpus and gpus: Super Computing 2010, Expanded Abstracts, 2874–2877, Super Computing.

Sava, P., 2007, Stereographic imaging condition for wave-equation migration: Geophysics, **72**, A87–A91.

Sava, P. and S. Fomel, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, 5209–5217.

Sava, P. C. and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074.

Villasenor, J. P., R. A. Ergas, and P. L. Donoho, 1996, Seismic data compression using high-dimensional wavelet transforms: Snowbird, UT, USA, Expanded Abstracts, 396–405, IEEE Computer Society Press.

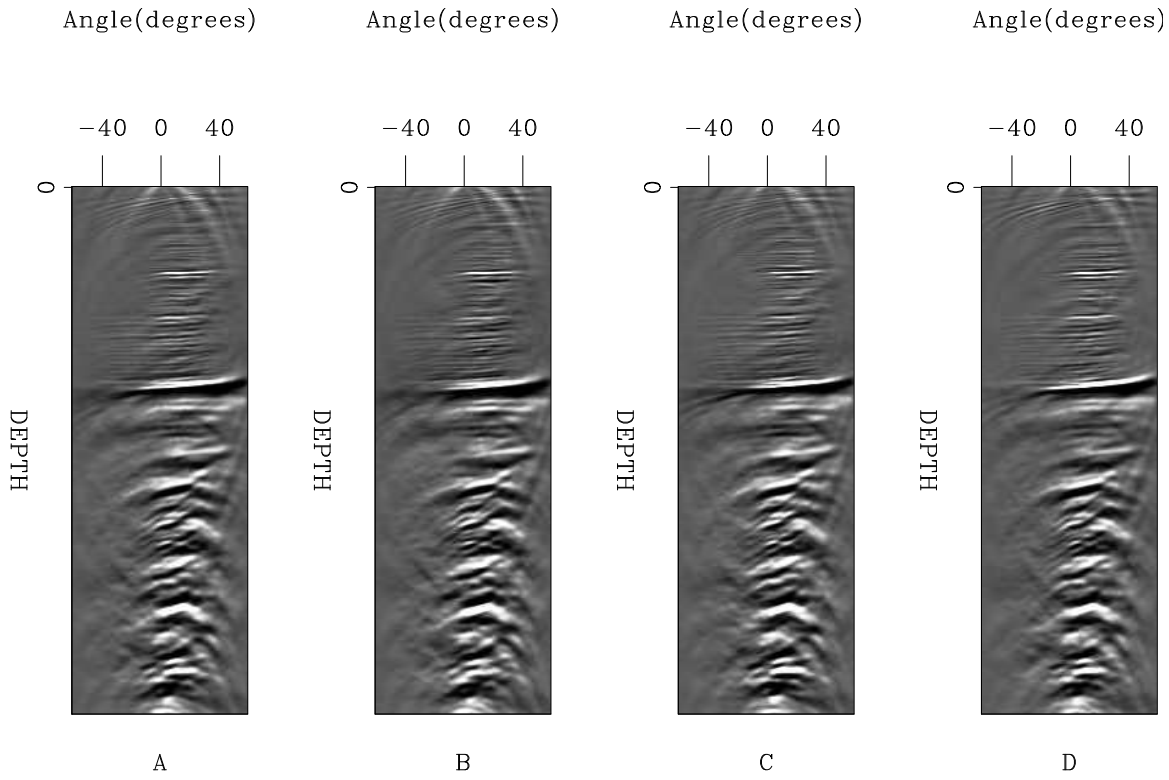


Figure 5: Four angle gathers at the same locations seen in Figure 2. [**ER**] bob1/. angle

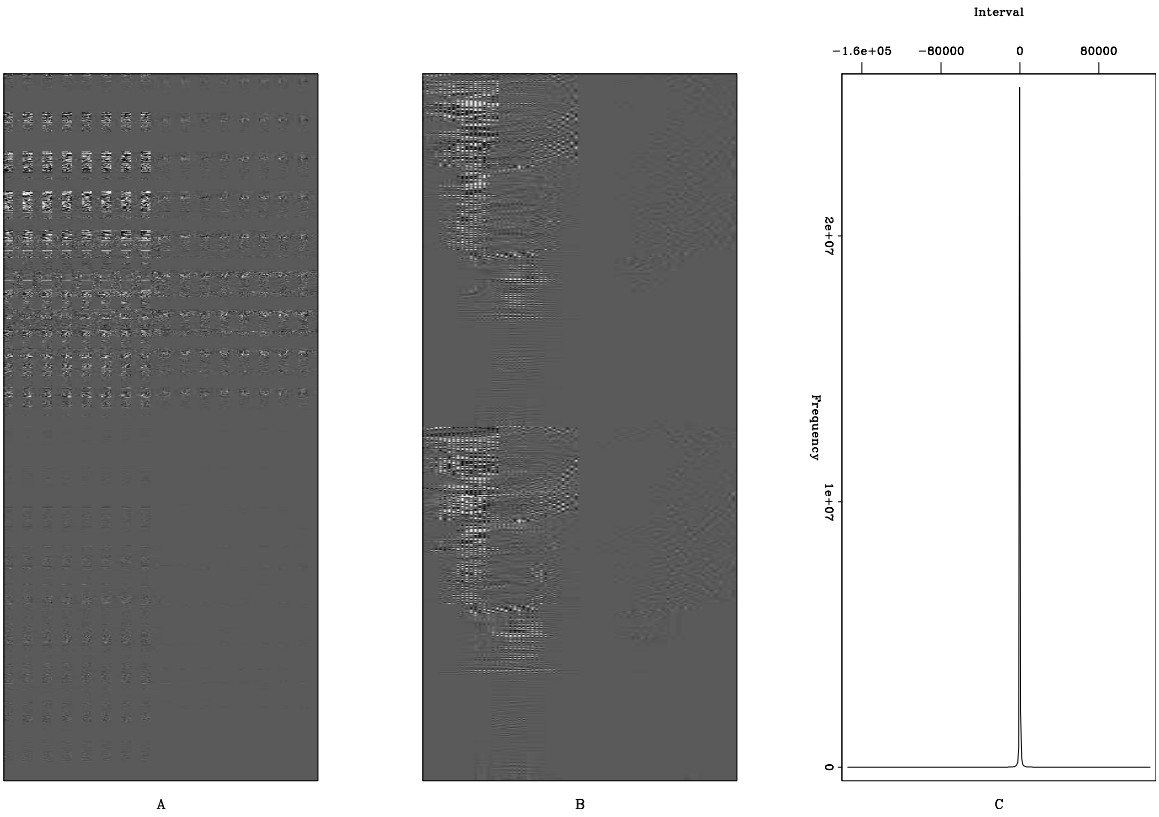


Figure 6: Panel A shows the wavelet domain representation of the 4-D volume used in this experiment transformed into the angle domain. Panel B shows a poriton of the wavelet domain. Panel C shows a histogram of the wavelet domain values. Note how the vast majority of the values are nearly zero. [**ER**] bob1/. wavelet2

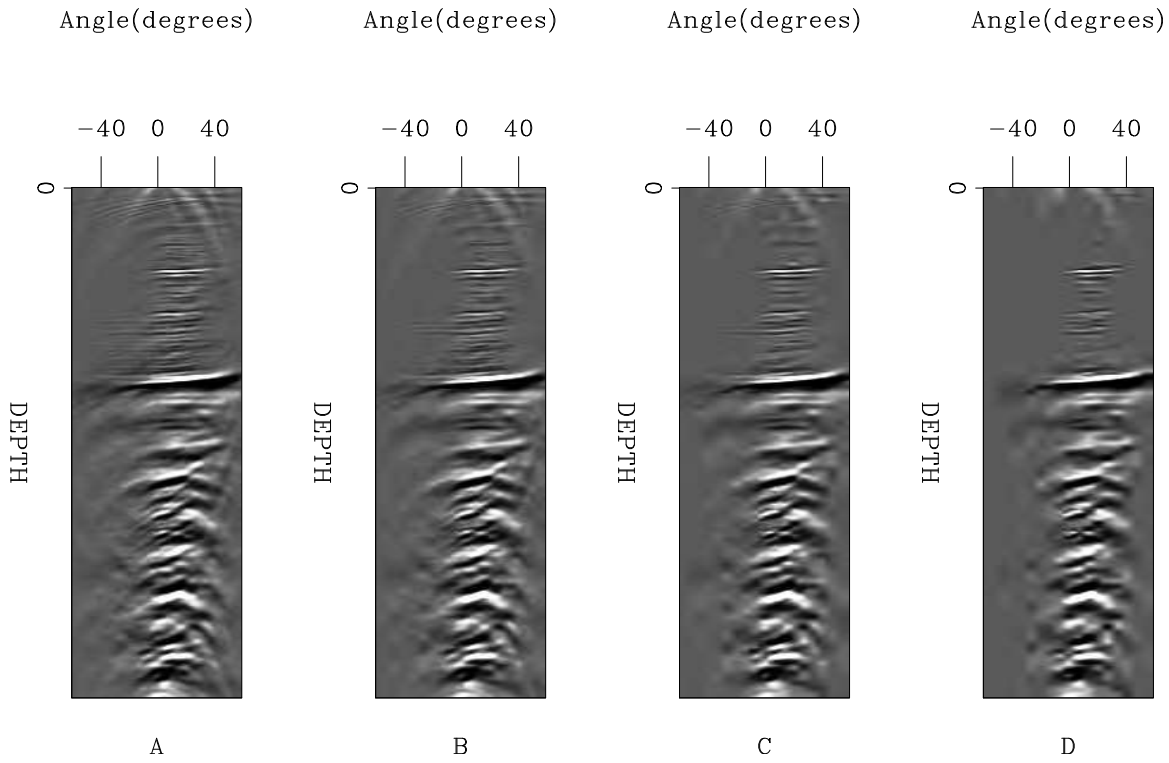


Figure 7: The result of zeroing the smallest values of the wavelet domain representation shown in Figure 3A. All four panels show the same sub-surface offset gather shown in Figure 5. A shows the result of clipping 95% of the values; B, 98%; C, 99%; and D,99.5%. Note how the reconstructed gather is nearly identical up to a 99% clip. [**ER**] bob1/. angle2

Memory efficient 3D reverse time migration

Chris Leader and Robert Clapp

ABSTRACT

Reverse time migration in three dimensions has two key bottlenecks - wavefield computation and IO limitations due to mass data transfer. Wave propagation and correlation in three dimensions imposes large computational constraints, both in terms of the number of floating point operations and the size of objects that need to be allocated. Furthermore, wavefields must be constantly written and read from disk since memory cannot possibly hold them all, this causes conventional RTM to become IO bound quickly. To address the computation requirements GPU propagation kernels are used to greatly reduce the computation time for source side modeling, achieving speeds in excess of 2.5 gigapoints calculated per second. Additionally, data handling requirements are vastly reduced by imposing pseudo-random boundaries on the velocity field, allowing time reversable source propagation. Achieving time reversable source propagation alleviates the requirement of checkpointing or boundary reinjection; this minimal data transfer results in GPU on-device operations becoming further accelerated.

INTRODUCTION

For advanced, accurate imaging Reverse Time Migration (RTM) is now the standard method. Despite the computational requirements relative to Kirchhoff or one-way wave-equation migration methods it has many characteristics that are desirable. In complex velocity fields the assumptions underlying one-way Wave Equation Migration (WEM) are insufficient, as only upward primary reflections are used, velocity approximations (FFD, PSPI etc) are made and only approximate phase and amplitude information is retained. Hence events with multiple reflections, overturning waves and steep dip information are not correctly positioned. Kirchhoff migration can image these steep dips, however the high frequency assumption results in sections of the data being incorrectly positioned and high angle artifacts in complex velocity structures. Conversely RTM images can resolve turning waves, horizontal waves and prismatic waves. In addition, no velocity, amplitude or phase approximations are necessary; though for practical purposes often such approximations are made.

For practical implementations, 3D RTM imposes large data handling problems. Generally, the source is fully propagated and saved in memory; the recorded data is then back propagated and correlated with the forward modeled shot at each imaging time step. Once geometries are 3D, it is not possible to hold this source wavefield fully in memory.

In addition to these large memory requirements, computational problems must also be addressed. As 3D surveys continue to become larger, with longer crossline and inline offsets and denser source sampling, then to fully benefit from this additional data the source modeling and receiver back propagation have to be performed over large velocity models. This problem is further compounded by the fact that finite difference wave propagation

is less stable in as dimensionality increases, meaning smaller time steps have to be used (Dablain, 1986). Typical model dimensions can range from 500mb to tens of gigabytes, and the number of time steps needed to accurately model these data without dispersion is often around 10000. Using a spatial stencil order of 8 and temporal of 2, then over 1 giga point (Gpts) of calculations are needed per time step for both the source and the receiver.

In the following sections I will discuss how using GPUs can greatly accelerate finite difference time domain wave propagation, and how random velocity boundaries can improve the IO performance of RTM. Initially I will give a brief overview of GPUs. A standard measure for kernel performance is the number of model points calculated per second (Micikevicius, 2009), and this is the metric that will be used subsequently.

GENERAL PURPOSE GRAPHICS PROCESSING UNITS

Using a General Purpose Graphics Processing Unit (GPGPU) one can perform many independent parallel instructions simultaneously - far more than possible using a parallel CPU based system. These are known as SIMD (Single Instruction Multiple Data) devices, as they are capable or running one set of instructions many times over parallel threads. With the release of the programming language CUDA in 2006 Nvidia provided a way of harnessing the power of graphics processing units with limited knowledge of graphics processing.

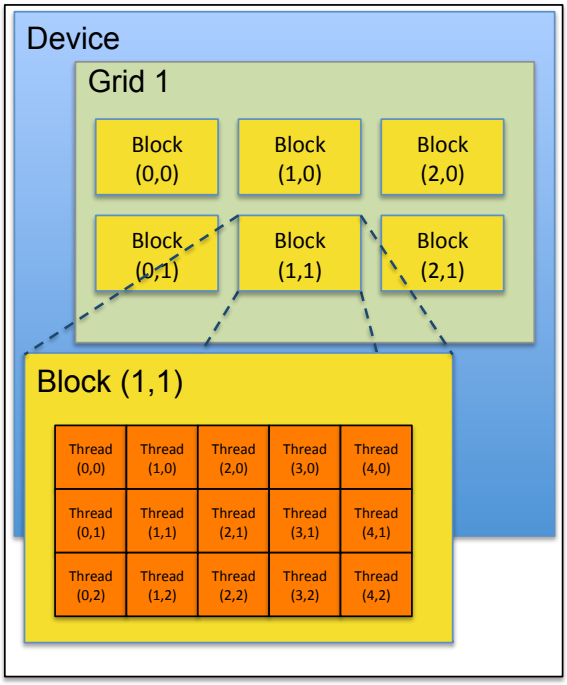


Figure 1: A schematic for the architechture of an Nvidia GPU [**NR**] chris1/. GPUarch

The GPU can be thought of as a 2D structure, or grid. This grid is broken into blocks, each block in term consists of a group of threads, Figure 1. Each of these thread-blocks has its own shared memory, which can be unique per thread-block, and it's own set of registers, which are the same between all thread-blocks. Each thread in this hierarchy can execute a set of instructions (a kernel) concurrently, allowing for fine grain parallelism. The true potential of the GPGPU architechture lies in how memory latency can be hidden. The GPU partitions resources using registers and shared memory, these both have a latency

of only a few cycles. Mass simultaneous execution effectively hides memory latency and context switching is (essentially) free. Parallel CPU execution, such as that possible when using OpenMP, MPI or POSIX, does not partition resources and features a memory latency at least an order of magnitude higher than the GPU equivalent, making this form of parallelisation vastly less efficient.

A schematic for how thread-block and global memories can interact is shown in Figure 2.

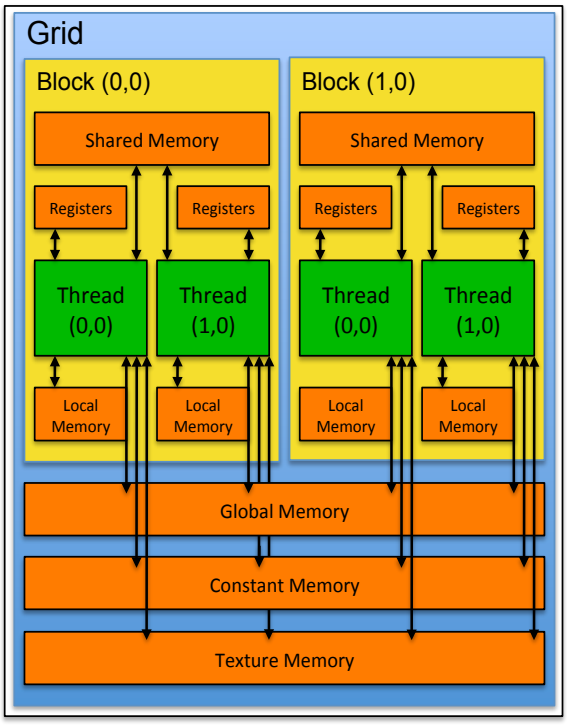


Figure 2: A schematic for the memory heirarchies within the GPU [**NR**] chris1/. GPUmem

3D WAVE PROPAGATION

The method used for wavefield forward and back propagation was time domain finite differencing of the 3D acoustic wave equation. This method approximates the temporal and spatial derivatives as a sum of Taylor series' about a symmetric sequence of points, such that this sum closely approximates the Taylor series of the desired derivative.

$$\frac{\partial^2 u(x,y,z)}{\partial t^2} = c(x,y,z)^2 \nabla^2 u(x,y,z) \tag{1}$$

The velocity model (Figure 3) used to test these propagation and imaging routines features velocities ranging between 1490 m/s to 4480 m/s; using a Ricker wavelet of central frequency 25 Hz then time steps of 0.5 ms and spatial sampling of 10m was used, adhering to the guidelines of Dablain (1986), extrapolated to 3D.

CPU Implementation

Solving Equation 1 numerically using a finite-difference time domain approach can be simply set up as a serial convolution repeated across the entire domain. Of course this is the most naive and least efficient way to do this. Over a very modest model size of 1 million points, for 8000 time steps this takes around 93 minutes to propagate. This serial methodology can be improved by blocking the domain into smaller 'pencil' shaped blocks about the fast axis; this will give a better cache hit rate, reduce calls to L3 memory and hide some latency. This can roughly half the computation time.

By parallelising over multiple cores a non-blocked speed up of around 3.5x is seen (over 8 cores, after this speed up saturates). Using the optimum blocking and parallelism discovered this computation time is reduced to around 16 minutes, giving a 5.5x - 6x speed up. Nonetheless, 3D model sizes are typically two or three orders of magnitude larger than this, meaning for a realistic model size this parallel, blocked CPU methodology will take several hours per shot. It should be stressed that this is the best speed up observed using a Fortran90 with OpenMP approach, by further vectorising loops and coding at assembly levels this can be further accelerated many times.

GPU Implementation

Using the identical serial kernel in CUDA this propagation takes 18 minutes. This is slower than the fastest CPU technique because GPU global memory latency is higher than on the CPU, to harness the power of the GPU shared memory and registers must be used.

Micikevicius (2009) describes an efficient way of setting up the base wave propagation kernel that utilises latency hiding by copying stencil location values into shared memory, dependent on thread location. Performing this propagation in 3D is about 27 times faster than the naive GPU approach, with speeds of over 3 Gpts/s achievable for typical model dimensions, and a run time of 41s (for the previously used model).

Operations purely on GPU global memory are slower than on a CPU, with latencies of 400-600 cycles, rather than 100-200 on a CPU. Due to this, a typical school of thought minimises data manipulation on the GPU. This approach constitutes sending to host (CPU) the full 3D wavefield at each desired modeling time step, causing the CPU wavefield object to be 4D. This can then be transposed and windowed accordingly. This method gives a speed of 24x over the best, blocked and parallelised CPU equivalent. Having to copy data this way is a shortcoming of interacting Fortran90, C and CUDA.

Storing the entire 4D wavefield as an allocated object on the CPU is unwise - memory is quickly saturated. For example, a model size of 500x500x500 dictates that only 64 time steps could be performed (on a 4 Gb card).

Several improvements can be made to make the GPU propagation as fast as possible. A new windowing kernel can be used to only send 2D wavefield slices back to the host (corresponding to desired acquisition geometries), this only marginally slows down the modeling and makes the full implementation vastly more flexible. In fact since now less data has to copied back to the CPU a speed up of about 1.3x is seen.

The final speed up thus far when compared to most efficient, blocked and parallelised

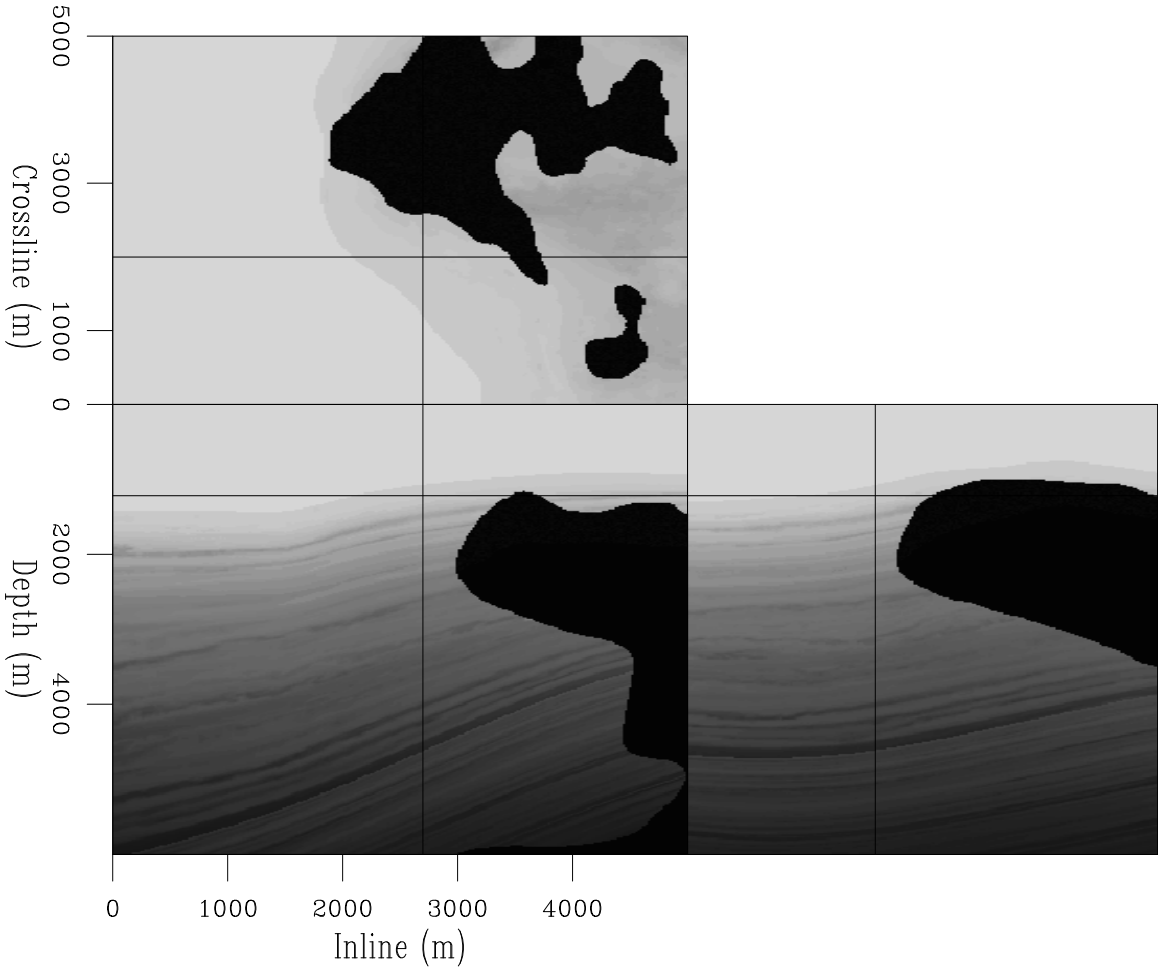


Figure 3: A 3D cubeplot of the velocity model used for 3D imaging [**ER**] chris1/. velslice3d

CPU routine is about 35x, and the current throughput is about 2.2 Gpoints/s. If the CPU modeling was fully optimised the observed speed up would be 15-20x. The evolution of computation speed (measured in million computed points per second) against propagation scheme can be observed in Figure 4.

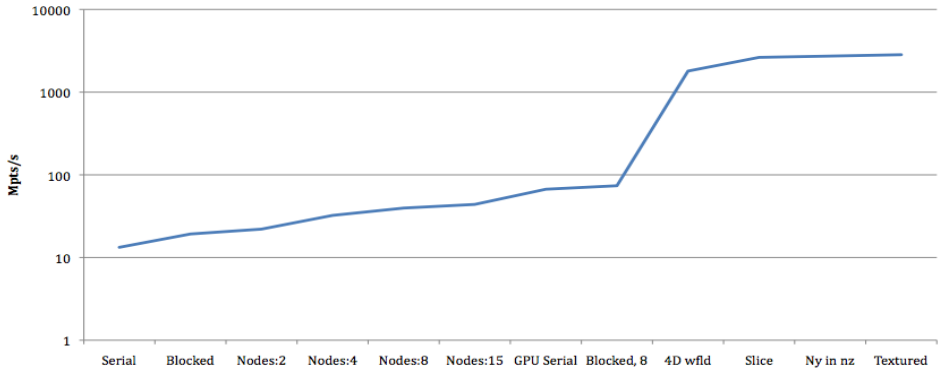


Figure 4: Graphing the speed up between various CPU and GPU propagation schemes [**NR**] chris1/. speed

The finite computational domain used when simulating wavefield propagation provides additional challenges. To accurately simulate real world physics, an infinite domain must be used, but of course this is not possible. If no boundary conditions are used then reflections from the computational domain are prevalent in the modeled data. There are many ways of reducing or hiding these domain reflections, such as zero value, zero order, absorbing, damping and Perfectly Matched Layer (PML) (Turkel and Yefet, 1998). For modeling an absorbtion scheme was used, which reduced throughput by about 10%.

This accelerated 3D wave propagation method will provide the engine for 3D RTM.

REVERSE TIME MIGRATION

Reverse time migration is a well known method, whereby reflection events are imaged in a kinematically accurate manner. Baysal and Sherwood (1983) and Claerbout (1985) pioneered the method, whereby an image is constructed by back propagating the receiver wavefield (recorded data) and forward propagating the source wavefield (modeled data, using assumed wavelet) and then cross correlating these at each imaging step. Mathematically this reduces to the imaging condition shown in Equation 2. It can be shown that this system is simply the adjoint of linearised modeling (Biondi, 2010).

$$I(z,x,y) = \sum_{i}^{nshots} \sum_{t}^{n} P_s(t,z,x,y;\mathbf{s}_i) P_g(t,z,x,y;\mathbf{s}_i) \tag{2}$$

$I(x,y,z)$ is the image at point (x,y,z), $P_s(t,z,x,y;\mathbf{s_i})$ is the source wavefield for shot $\mathbf{s_i}$ and $P_g(t,z,x,y;\mathbf{s_i})$ is the corresponding recorded wavefield. This can be extended to a prestack imaging condition where subsurface offsets are included. This is necessary velocity estimation and many inversion methods. Equation 2 simply extracts the zero subsurface offset image.

$$I(z,x,y,x_h,y_h) = \sum_{i}^{nshots} \sum_{t}^{n} P_s(t,z,x+x_h,y+y_h;\mathbf{s}_i) P_g(t,z,x-x_h,y-y_h;\mathbf{s}_i) \qquad (3)$$

In Equation 3, x_h and y_h represent the subsurface half offsets. For each set of subsurface offsets the two wavefields are laterally shifted and the imaging condition reapplied, hence a 5D prestack image can be constructed (Biondi, 2006).

In 3D the computational intensity of RTM can become a hinderance relative to simpler migration operators. This is because the modeled shot, $P_s(\mathbf{s_i})$, has to run forward in time, and the recorded data, $P_g(\mathbf{s_i})$, is reversed in time. Since domain boundary damping makes the shot modeling non-reversable then practically we model and save the entirity of $P_s(\mathbf{s_i})$, and hold this in memory as we reverse through $P_g(\mathbf{s_i})$, applying the imaging condition at each time step. Of course, for large wavefields this is unpractical. One partial solution is referred in literature as checkpointing (Symes (2007); Dussaud and Cherrett (2008)). Here we effectively block the propagation and periodically reinject the source model. For example, let some intermediate position in the propagation be t_{check}, initially $P_s(t_{max} - t_{check}; \mathbf{s_i})$ is read, the receiver incrementally back propagated to t_{check} and the correlations applied. This can then be done t_{max}/t_{check} times (Clapp, 2008a) and can be parallelised. The memory problem is thus solved, however this process is fundamentally IO bound.

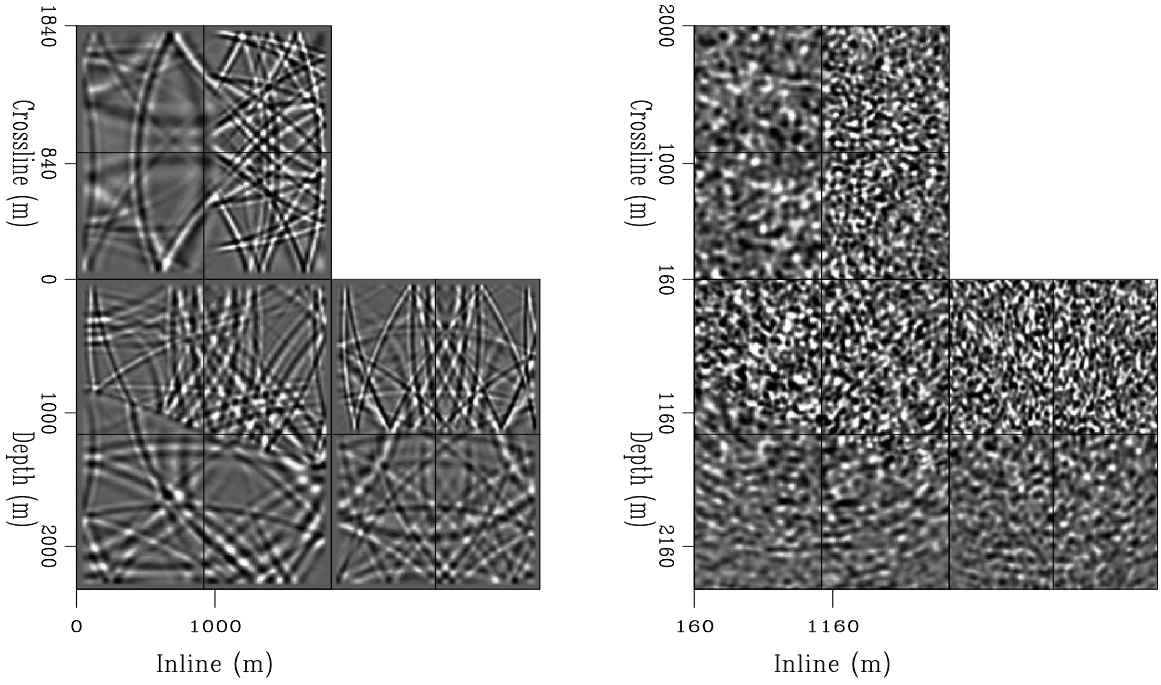


Figure 5: A slice of the 3D wavefield after long propagation times, without and with random boundaries [**CR**] chris1/. basic

Clapp (2008b) proposed a different scheme - to use pseudo-random boundaries optimised to scatter these wavefields incoherently (Figure 5). The RTM imaging principle operates on the fact that incoherent data are either correlated out in imaging, or stacked out subsequently. Pseudo-random boundaries create a lot of noise in the propagated shot, however

this will not correlate coherently with the receiver wavefield. Furthermore this propagation is now entirely time reversible, and by saving the velocity field the shot can be reverse propagated to its original state within machine precision (Figure 6). The benefits of this are twofold - computationally the source wavefield is time reversible, and so only two source wavefields are needed for any given time and no checkpointing is necessary; geophysically this algorithm provides accurate imaging, and any residual boundary scattered noise is well below the noise threshold. Thus a manner in which 3D RTM is both computationally and memory efficient can be constructed.

As seen in Figure 5, the low frequency component of the wavefield is not well scattered. For RTM this is not a significant problem, since low frequency artifacts are often prevalent and are removed by a low cut filter (or domain decomposition) before imaging (Taweesintananon (2011) provides an overview of these methods). However techniques such as waveform inversion methods rely on low frequency information. Shen and Clapp (2011) show how random boundaries can be further optimised, such that low and high frequencies become incoherent after scattering.

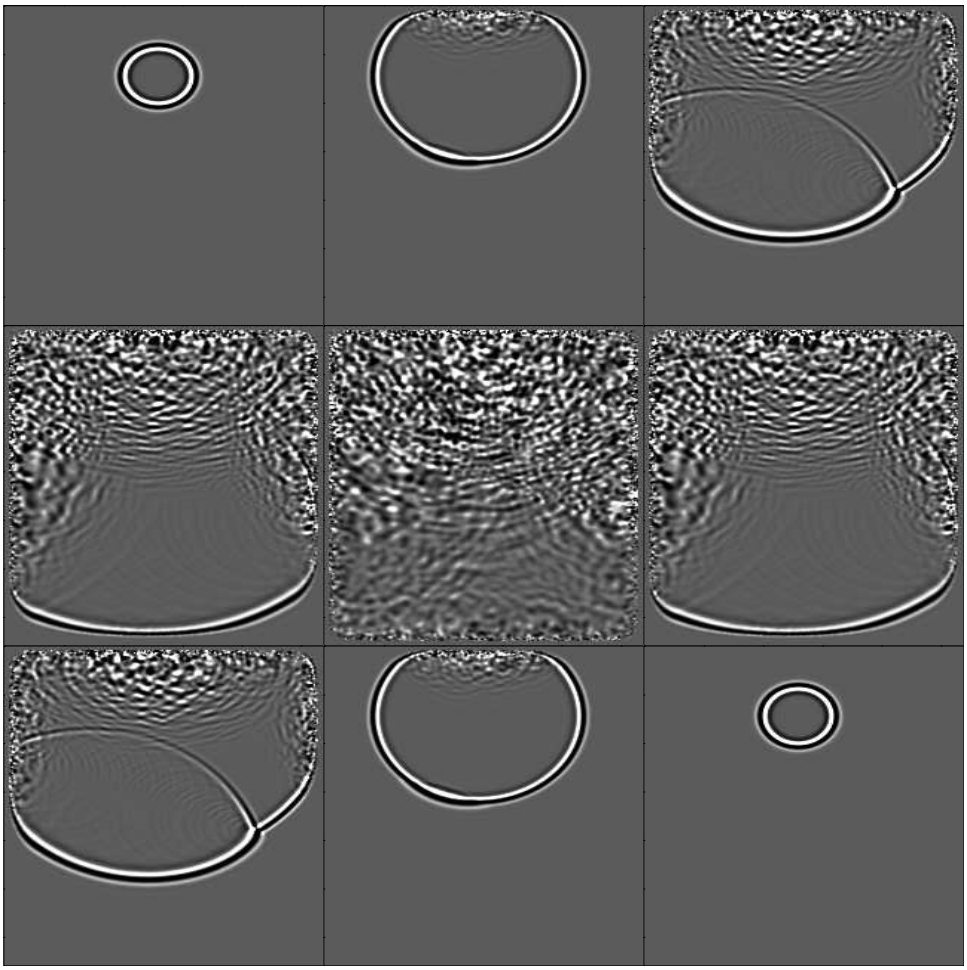


Figure 6: An example of the time reversable wave propagation [**CR**] chris1/. shotrand3

GPU implementation

Setting up this scheme optimally requires some thought. A random boundary of user input depth is imposed on this velocity section, which is then bound as a texture on the GPU. The source is then fully propagated and the final two 3D wavefields saved; now we have all the array inputs for our migration. This runs from t_{max} to t_0, and at each time step source and receiver wavefields are injected and propagated, at each imaging time step these fields are multiplied and the image stored and updated on the GPU. Damping boundaries are used for the receiver propagation, only one image object is used, and all correlations are stacked into this, then after propagation is completed the stacked image is sent back to the CPU. Under this scheme six 3D objects need to be stored - two wavefield slices for both the receiver and the source fields, the velocity model and the image. The fact that no device to host transfers occur during the time loop makes this kernel very efficient.

The extension to subsurface offsets adds data handling and numerical complexity. The aforementioned scheme produces a poststack image, since each image is stacked into one object on the GPU. When applying the imaging condition in Equation 3 the wavefields must be shifted and correlated mutliple times, for each time step, to produce the subsurface offset gathers. For practical execution we now need to transfer each imaging time step back to the host. Furthermore this multiplication is inefficient on the GPU, since the threads may need re-indexing after shifting, and a global memory operation is used. Future work is underway to see if these shifted correlations are better performed on the CPU. Other work at SEP on how compressive sensing can be used to reduce this new bottleneck is currently being performed (Clapp, 2011).

RESULTS

Using the latest SEAM velocity model both 2D and 3D random boundary migration routines were experimented with. A simple RTM result using a 2.5D dipping reflector is shown in Figure 7; this was created using a single shot at 20m depth and then applying a very low cut filter. Here very little background noise can be seen from the random boundaries, and with more shots this will reduce further. Few of the typical RTM artifacts are seen because of the simplistic nature the ray paths take in this model - all multiples and direct arrivals are muted and there is no multi-pathing or overturning of waves.

In the vastly more complex SEAM model multiple shots are required to suppress artifacts, especially since there are now many more conventional RTM artifacts due to the basic imaging condition used, such as those from correlating wavefields travelling in the same direction. Currently more work must be undertaken to ensure the velocity model is fully block-aligned in the GPU and the determine the best approach to begin to suppress these new imaging artifacts; a basic one shot RTM over the SEAM model can be seen in Figure 8. Some reflectors are noticeable and continuous, however there are several erroneous events and a lot of shallow noise.

CONCLUSIONS

The goal of this investigation was to try and reduce the RTM bottlenecks of 3D wavefield computation and memory transfers. The former of these was vastly reduced by using a

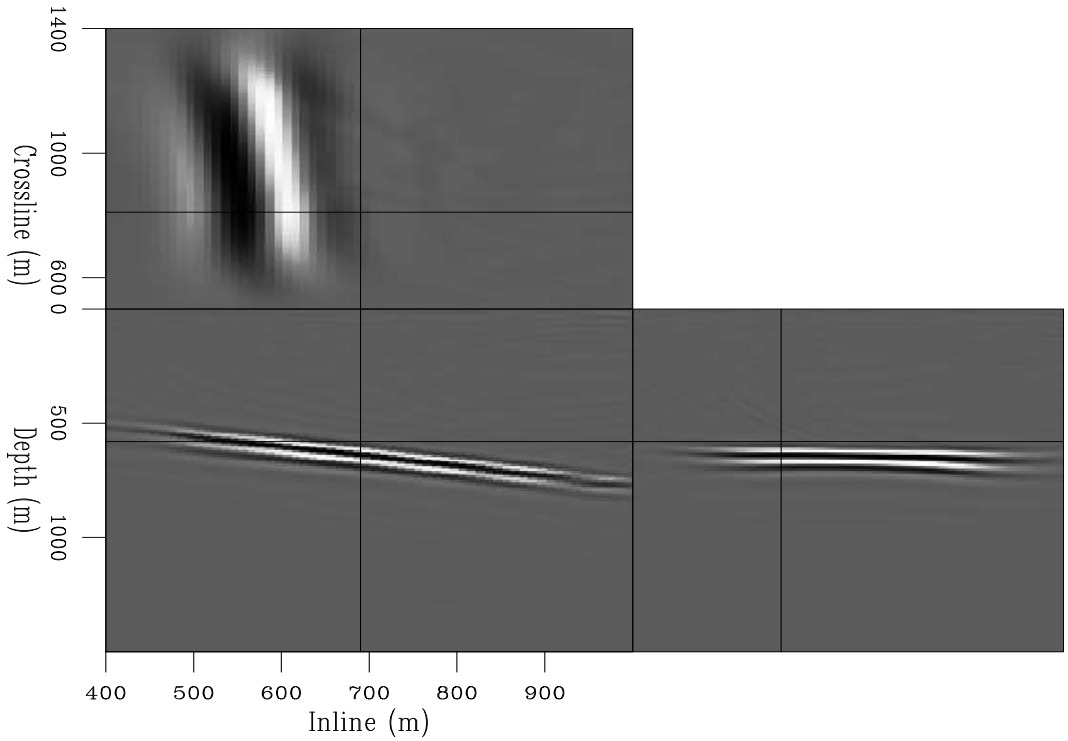


Figure 7: RTM result using one shot over a single dipping reflector [**CR**] chris1/. basicimg

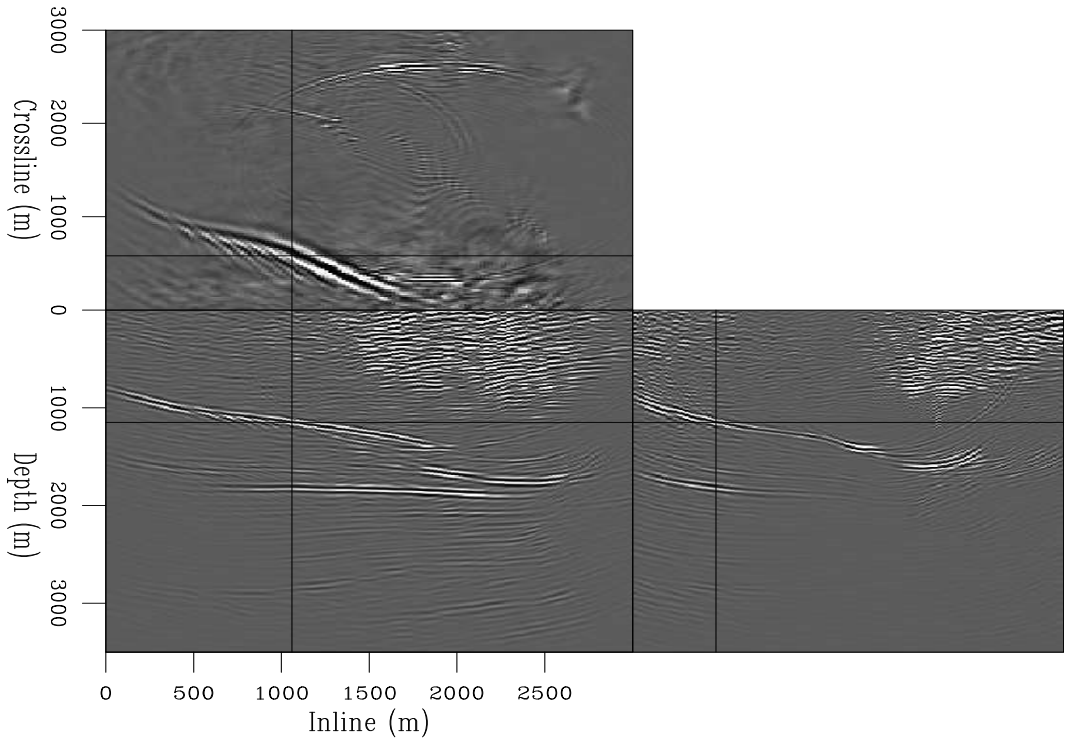


Figure 8: RTM result using one shot over a windowed section of the SEAM model [**CR**] chris1/. seamimg

series of GPU oriented kernels to perform wave propagation in a massively parallel sense, a total speed up of 30x was seen. The memory transfer data handling problem within RTM was diminished by using increasingly random boundaries on the velocity model to produce time-reversable wavefield modeling, alleviating the method from the intensely IO bound technique of checkpointing. A simple 2.5D single reflector model showed an encouraging RTM result, which demonstrates that random boundaries in 3D do indeed provide results comparable to checkpointing; work is underway to obtain detailed images of more complex velocity structures.

ACKOWLEDGMENTS

I would like to thank Bob Clapp for his patience and assistance on CUDA and C++ programming, and for providing an object oriented framework which inspired how this procedure was set up. I would also like to thank Xukai Shen and Ali Almomin for their CUDA debugging and receiver gather interpretation expertise.

REFERENCES

Baysal, E, K. D. and J. Sherwood, 1983, Reverse time migration: Geophysics, **45**, 1514–1524.

Biondi, B., 2006, 3d seismic imaging: Society of Exploration Geophysicists.

———, 2010, 3d seismic imaging, additional notes: Migration as adjoint of linearized modeling: Society of Exploration Geophysicists.

Claerbout, J., 1985, Basic earth imaging: Society of Exploration Geophysicists.

Clapp, R., 2008a, Reverse time migration: Saving the boundaries: SEP-136, **136**.

———, 2008b, Reverse time migration with random boundaries: SEP-138, **138**.

Clapp, R. G., 2011, Imaging using compressive sensing: SEP-Report, **143**, 149–158.

Dablain, M. A., 1986, The application of high-order differencing to the scalar wave equation: Geophysics, **51**, 54–66.

Dussaud, E, S. W. L. L. S. P. D. B. and A. Cherrett, 2008, Computational strategies for reverse time migration: SEG Expanded Abstracts, **78**, 2874–2877.

Micikevicius, P., 2009, 3d finite difference computation on gpus using cuda: GPGPU, **2**.

Shen, X. and R. G. Clapp, 2011, Random boundary condition for low-frequency wave propagation: SEP-Report, **143**, 85–92.

Symes, W., 2007, Reverse time migration with optimal check pointing: Geophysics, **72**, 213–221.

Taweesintananon, K., 2011, Reverse-time migration using wavefield decomposition: SEP-Report, **143**, 171–190.

Turkel, E. and A. Yefet, 1998, Absorbing pml boundary layers for wave-like equations: Applied numerical mathematics, **27**.


OPEN DAILY
HALLIBURTON

Reverse-time migration using wavefield decomposition

Kittinat Taweesintananon

ABSTRACT

Reverse-time migration (RTM) is capable of imaging very steeply dipping reflectors and overhangs. However, it usually produces strong artifacts that contaminate the shallow parts of the migrated images. These artifacts can be suppressed using an imaging condition with appropriate decomposed source and receiver wavefields. In this paper, such a technique is applied and examined. This imaging condition keeps only energy at the points where strong backscattering originates. The results show that RTM using wavefield decomposition is a promising remedy for attenuating artifacts compared to the implementation of a low-cut filter. However, some artifacts still remain in the decomposed RTM image. These residual artifacts are caused by the cross-correlation between the upgoing component of the direct source wavefield and the backscattered component of the receiver wavefield.

INTRODUCTION

RTM is a migration method based on wavefield extrapolation in time. As a two-way wave-equation migration, it can image very steeply dipping reflectors and even overturned events. This method is based on two consecutive steps: numerical propagation of wavefields from the recording surface into image space, followed by the image formation process using these wavefields. RTM was first used for poststack depth migration of 2D seismic data in the late 1970s (Baysal et al., 1983; Whitmore, 1983).

Poststack RTM involves the extrapolation of receiver wavefields from recorded seismic data by computing the full two-way solution for the acoustic wave equation with time running backwards. In contrast, prestack RTM involves the extrapolation of source and receiver wavefields.

Prestack RTM for a shot profile consists of three steps: 1) forward extrapolating the source wavefield $S(t, \vec{x})$ from time $t = 0$ to $t = t_{\max}$, 2) backward extrapolating the receiver wavefield $R(t, \vec{x})$ from time $t = t_{\max}$ to $t = 0$, and 3) applying a suitable imaging condition to construct the reflectivity image (Biondi, 2005). The conventional imaging condition for RTM is the zero-lag cross-correlation of source and receiver wavefields (Claerbout, 1985). Since RTM is a type of shot-gather migration, each shot profile is migrated independently. Thus, the final RTM image is the sum of the images from all individual profiles.

The extrapolated wavefields are full two-way solutions of the wave equation. These two-way solutions do not limit the propagation directions of wavefields. Therefore, RTM can correctly image very steeply dipping reflectors and even overturned events. In extremely complex areas, it can also handle strong lateral velocity variations and multi-path arrivals (Etgen et al., 2009). Thus, RTM can produce more accurate images than other popular

migration methods such as Kirchhoff, beam, and one-way wave-equation migration. Nevertheless, the two major drawbacks of RTM are strong low-frequency artifacts and high computational cost. Such artifacts are generated due to the nature of the conventional RTM imaging condition. Many remedies, such as velocity smoothing and post-migration filtering, have been proposed to suppress these artifacts, but they all have severe limitations.

This paper focuses on a promising technique for attenuating artifacts in correlation-based RTM images called wavefield decomposition. This technique is to decompose source and receiver wavefields into the appropriate components based on their propagation direction, so that the decomposed wavefields can produce the image without artifacts (Liu et al., 2007, 2011). In this paper, I first provide an overview of RTM, its artifacts, and some common solutions for attenuating them. Next, I examine RTM using wavefield decomposition. Finally, I discuss the benefits and drawbacks of this technique and its practical implications.

REVERSE-TIME MIGRATION

For a single shot profile, the conventional correlation-based imaging condition for 2-D prestack RTM can be written as

$$I(\vec{x}) = \sum_{t=0}^{t_{\max}} S(t,\vec{x})R(t,\vec{x}), \tag{1}$$

where $t_{\max}$ is the maximum recorded time and $\vec{x} = (x, z)$ is the vector position of a wavefield in 2-D space. Source and receiver wavefields are extrapolated in time by solving the acoustic wave equation

$$\left(\frac{1}{v^2(\vec{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right) P(t,\vec{x}) = 0, \tag{2}$$

where $P(t,\vec{x})$ is the pressure wavefield, and $v(\vec{x})$ is the acoustic velocity in the computational domain.

In this paper, I use an explicit finite-difference method using a 2nd-order forward difference in time and a 4th-order central difference in space for wavefield extrapolation. The stability and antidispersion conditions for this method were described by Dablain (1986).

In practice, we can restrict the computational domain for wavefield extrapolation to only parts of the physical domain. Such a limitation produces artificial reflectors at the boundaries of the computational domain. In this study, I applied a Gaussian taper (Cerjan et al., 1985) along the artificial boundaries in order to suppress these undesired reflections. At the end of each Gaussian tapering region, I also applied an absorbing boundary condition using paraxial approximations of the acoustic wave equation (Clayton and Engquist, 1980). In this study, I applied absorbing conditions to all computational boundaries for extrapolating source and receiver wavefields, so that additional surface-related artifacts were suppressed.

To illustrate the wavefield extrapolation used, Figures 1(a) and 1(b) show an extrapolated wavefield in a constant-velocity model with a reflecting top boundary, whereas Figures 1(c) and 1(d) show an extrapolated wavefield in the same model with an absorbing top boundary instead. The extrapolation using such an absorbing top boundary was used to generate seismic data without surface-related multiples (SRM).

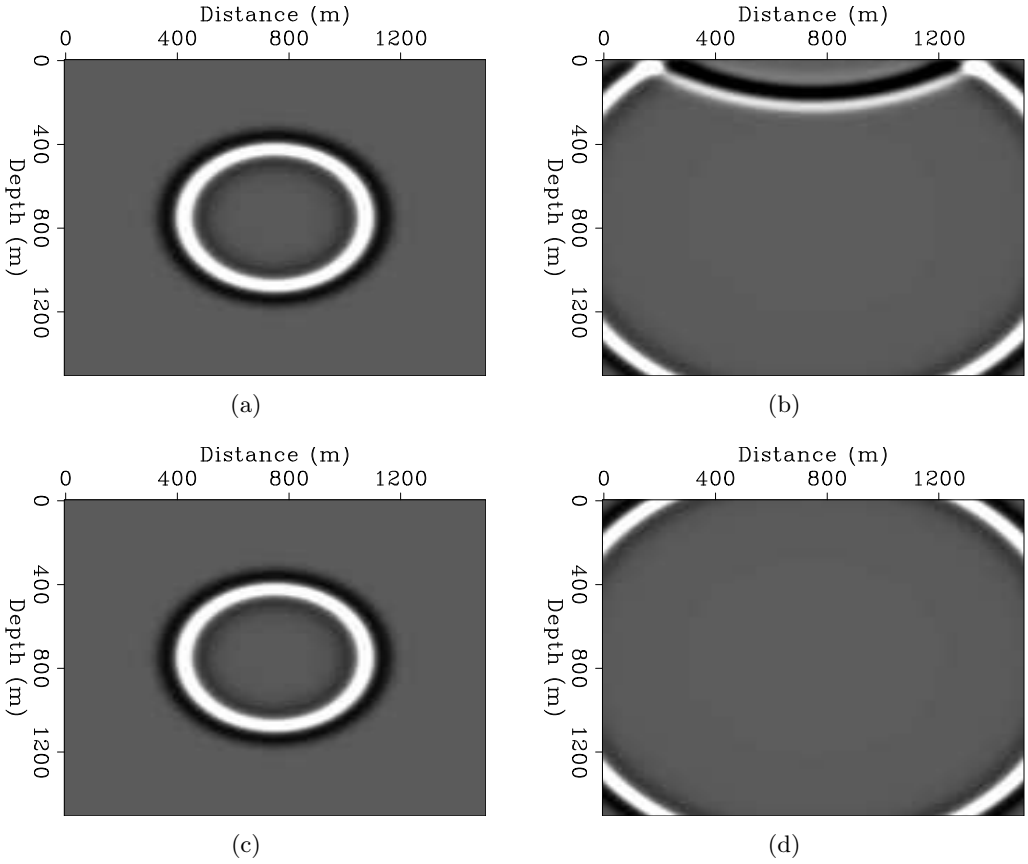


Figure 1: Wavefield snapshots illustrate the wavefield extrapolation used in this paper. (a) and (b) show snapshots from a single modeling experiment with a reflecting top boundary at 0.25 and 0.55 seconds, respectively. (c) and (d) show snapshots from a single modeling experiment with a nonreflecting top boundary at 0.25 and 0.55 seconds, respectively. [**ER**] kittinat/. wvfld0-mult1,wvfld0-mult2,wvfld0-nomult1,wvfld0-nomult2

RTM ARTIFACTS AND REMEDIES

Strong RTM artifacts result from the undesired cross-correlations of head waves, diving waves, and backscattered waves (Yoon et al., 2004; Yoon and Marfurt, 2006). These artifacts are strong when high velocity contrasts are present, especially in the shallow regions. In addition, undesired artifacts caused by SRMs can also confuse the interpretation.

Figure 2(a) shows a simple 1-D velocity model consisting of a strong velocity contrast, whereas Figure 2(b) is a shot profile corresponding to this velocity model. Figure 3(a) shows the conventional RTM image from this shot profile using the true interval velocity as the migration velocity. The figure contains many artifacts due to head waves, diving waves, and backscattered waves at shallow depths. In addition, there are also artifacts due to SRMs at the bottom of the figure. Muting the head waves, diving waves, and direct waves before RTM can suppress the corresponding artifacts as shown in Figure 3(b). In practice, the artifacts due to SRMs can be removed by using an appropriate SRM elimination program. Here, the synthetic data without SRM are generated by applying a nonflecting top boundary

to wavefield extrapolation. Figure 4(b) shows the RTM image from the muted shot profile without SRM shown in Figure 4(a). Although artifacts due to head waves, diving waves, and SRM have been removed, the adverse effect of backscattered waves still remains above the strong reflector (Figure 4(b)). Thus, only RTM artifacts due to backscattered waves cannot easily be suppressed by pre-migration processing.

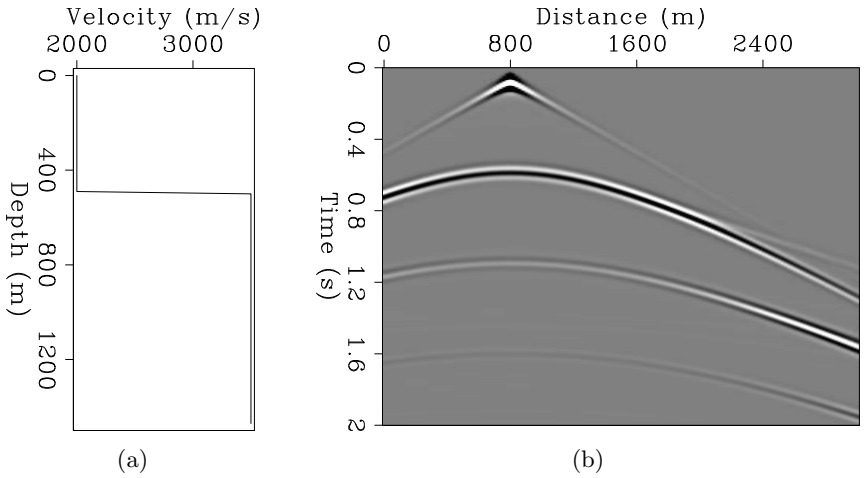


Figure 2: (a) 1-D true interval velocity model. (b) Synthetic shot profile from an explosive source on the surface in the velocity model (a). [**ER**] kittinat/. vel0,data0-srm

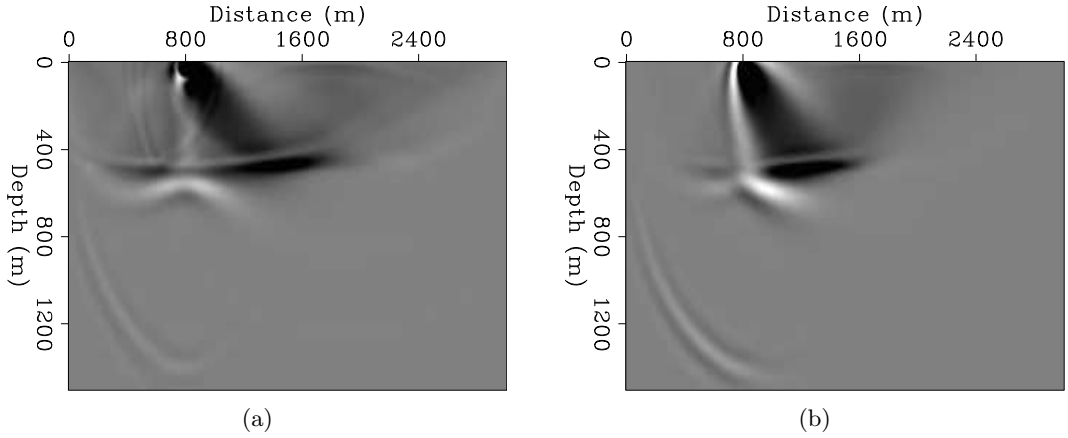


Figure 3: RTM images from the shot profile in Figure 2(b) before (a) and after (b) muting the head waves and diving waves. The true interval velocity shown in Figure 2(a) has been used as the migration velocity for both images. [**ER**] kittinat/. rtm-sharpv0-srm,rtm-sharpv0-srm-muted

Robein (2010) summarized four main approaches that have been used to suppress RTM artifacts due to backscattered waves: 1) smoothing the velocity model, 2) applying an appropriate filter to the post-migration image, 3) using nonreflecting wave propagation with directional damping, and 4) modifying the imaging condition.

Smoothing strong velocity contrasts can reduce the amplitude of undesired reflected waves; however, this can bias the accuracy of imaging due to incorrect velocity models.

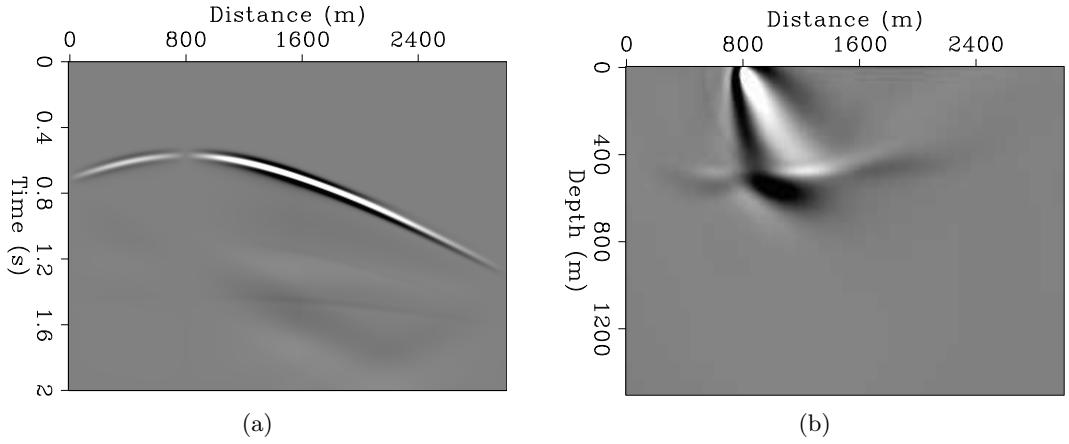


Figure 4: (a) A shot profile without SRM after muting the head waves and diving waves; weighting and tapering are also applied. (b) An RTM image from this shot profile; the true interval velocity shown in Figure 2(a) has been used as the migration velocity. [**ER**] kittinat/. data0-nosrm,rtm-sharpv0-muted-nosrm

Figure 5(b) shows the same RTM image as Figure 4(b), but using the smoothed migration velocity shown in Figure 5(a). As shown in Figure 5(b), velocity smoothing can partly reduce the artifacts caused by backscattered waves; however, the artifacts still have an adverse effect on the shallow parts of the image. Although velocity smoothing can also cause other kinds of artifacts (Fei et al., 2010), this approach is currently used due to its robustness and low computational cost.

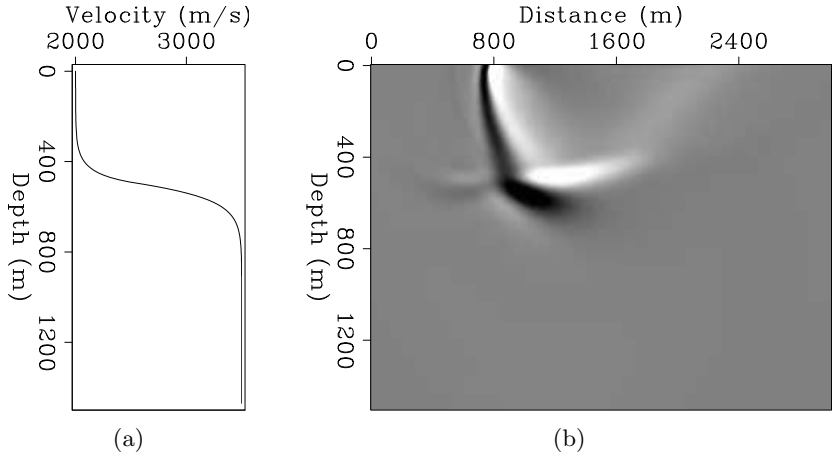


Figure 5: The effects of velocity smoothing on the RTM image. (a) Smoothed migration velocity corresponding to the true interval velocity in Figure 2(a). (b) RTM image from muted data without SRM using this smoothed migration velocity (a). [**ER**] kittinat/. vel0-mig,rtm-v0-nosrm-muted

There are many post-migration approaches whereby RTM artifacts are filtered after imaging. Low-cut filtering is an approach that can highly attenuate RTM artifacts as shown in Figure 6, where a causal low-cut filter was used. However, there still remains

some noise in the figure. Moreover, low-cut filters can reduce the amplitude of steeply dipping reflectors, which mainly contain low-frequency information. Guitton et al. (2006) presented another possible post-migration solution by applying a least-squares filter to RTM images. In addition, the application of a Laplacian filter with appropriate pre- and post-migration processing is an effective solution that can remove RTM artifacts without hurting steeply dipping reflectors (Zhang and Sun, 2009).

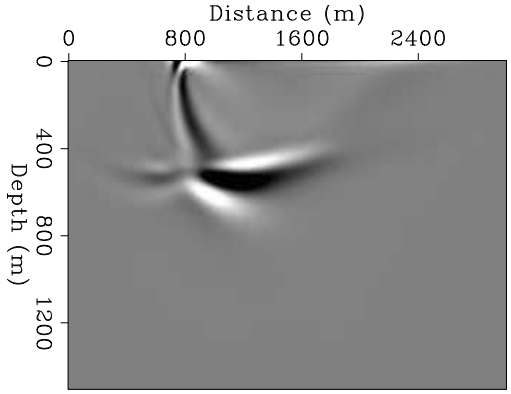


Figure 6: The RTM image in Figure 5(b) after applying a causal low-cut filter. [**ER**] kittinat/. rtm-v0-nosrm-muted-lowcut

Applying a directional damping term to the wave equation can suppress unwanted internal reflections, including backscattered waves (Fletcher et al., 2006). This approach is more effective than velocity smoothing, since it does not alter the migration velocity. However, picking the problematic interfaces, which cause artifacts due to backscattered energy, is required in order to obtain directional damping coefficients. Therefore, the approach using directional damping is time consuming.

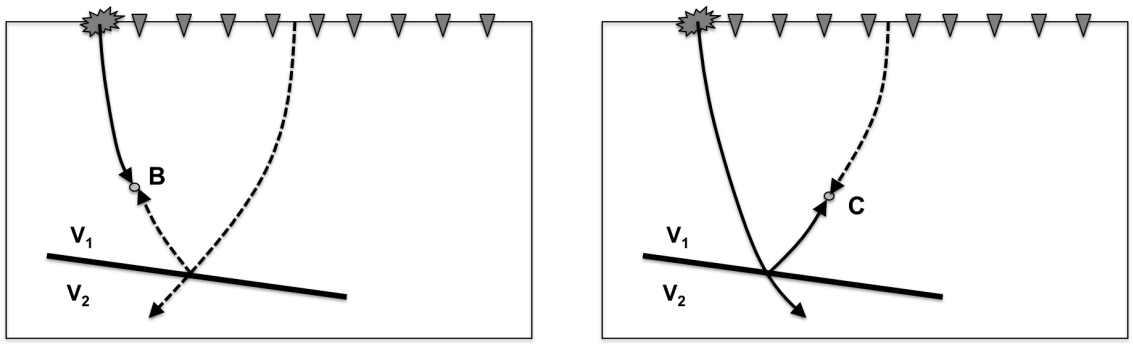
Modifying the imaging condition in RTM is another solution. This generally provides more effective noise attenuation than the other methods discussed above. The ultimate goal of this approach is to keep only energy along reflectors in the final image. There are various techniques used for modifying the imaging condition. The results of using different imaging conditions were discussed by Valenciano and Biondi (2003), Yoon et al. (2004), Liu et al. (2007), and Chattopadhyay and McMechan (2008).

RTM using wavefield decomposition is a method based on the modification of the imaging condition. This method modifies the imaging condition using wavefield decomposition in the Fourier domain, and was first introduced by Liu et al. (2007). This technique solves the problem of artifacts by suppressing their formation in the RTM algorithm. To demonstrate why this technique is worth investigation, we first need to understand the origin of RTM artifacts.

Origin of RTM artifacts due to backscattered waves

In order to investigate the origin of RTM artifacts due to backscattered waves, let us assume a flat interface with high velocity contrast as shown in Figure 7. The source wavefield (solid) and receiver wavefield (dashed) are extrapolated using a full propagator corresponding to Equation 2. When the propagated wavefields encounter this strong interface, part of the energy is transmitted and carries on downwards, while the remaining energy is reflected back to the surface. This phenomenon is observable for both source and receiver wavefields. Thus, at any point above the reflector, the corresponding source and receiver wavefields have

exactly the same propagation paths, along which non-zero cross-correlations are produced (Robein, 2010; Liu et al., 2011). Accordingly, the cross-correlation of these wavefields creates not only the reflectivity image but also noise at all nonreflecting points along the propagation paths, such as point B and point C in Figure 7. Thus, low-frequency noise is a common characteristic of the RTM image with a strong reflector.



(a) (b)

Figure 7: The solid lines represent the propagation paths of a source wavefield, where the arrows indicate the propagation direction with time running forward. The dashed lines represent the propagation paths of a receiver wavefield, where the arrows indicate the propagation direction with time running backwards. [**NR**] kittinat/. rtm-artifact1,rtm-artifact2

Note that, in this paper, I use the terms with the suffix *-going* such as *upgoing* and *downgoing* for describing the propagation direction of wavefields with time running **forward**. At point B in Figure 7(a), the downgoing source wavefield coincides in time with the propagated receiver wavefield that has been reflected. This receiver wavefield is indeed propagating downward with time running forward, i.e. downgoing receiver wavefield. On the other hand, at point C in Figure 7(b), the upgoing reflected source wavefield coincides in time with the upgoing receiver wavefield.

At either point B or point C in Figure 7, both source and receiver wavefields are propagating in the same direction and the same path with time running forward. Thus, the correlation-based images at point B and point C are caused by forward scattering events. These images are undesirable and can be considered to be noise, because they do not represent a reflected event. In contrast, at any point on the reflector, source and receiver wavefields are propagating in opposite directions with respect to the normal vector of the reflector. Thus, the cross-correlation of these source and receiver wavefields, which results from backward scattering events, produces the desired image.

To keep only energy corresponding to backward scattering events in the final image, we might apply the cross-correlation to only the source and receiver wavefields that are propagating in opposite directions with respect to the reflector normal. This approach requires us to compute the propagation direction of both the source and receiver wavefields at every point in the computational domain. Although this approach involves intensive computation, it can provide the reflectivity images without forward scattering noise.

The computation of the wavefield propagation direction at every point in the compuational domain can be done using a Poynting vector (Yoon et al., 2004). This works well for simple models, but it does not produce satisfactory results in complex subsurface structures

(Guitton et al., 2006). Xie and Wu (2006) applied the local one-way propagator, which is endowed with the Rayleigh integral, in order to decompose the wavefields. Another practical method is to decompose the wavefields in the Fourier domain in order to obtain the appropriate directional components of the wavefields based on Cartesian directions (Liu et al., 2007). This method is robust and worth investigation.

Wavefield decomposition based on Cartesian propagation directions can be used to modify the RTM imaging condition. This application was introduced by Liu et al. (2007, 2011), Suh and Cai (2009), and Fei et al. (2010). They all introduced methods to decompose source and receiver wavefields into one-way components along horizontal directions (leftgoing/rightgoing) and vertical directions (upgoing/downgoing). They then applied the zero-lag cross-correlation as the imaging condition to the appropriate combinations of these wavefield components. The results showed that such a method can effectively eliminate unwanted artifacts from conventional RTM images, and image quality was also improved.

Liu et al. (2007) applied wavefield decomposition in the F-K domain and obtained promising RTM results. The same authors later showed that a cheaper wavefield decomposition method in the wavenumber domain (T-K domain) can also provide the same results (Liu et al., 2011). Moreover, Suh and Cai (2009) applied a fan filter in the T-K domain before decomposing the filtered wavefields in the F-K domain. This approach produces better RTM images, but the computational cost significantly increases. In addition, Fei et al. (2010) illustrated how the decomposition method attenuates the artifacts in RTM images. However, Fei et al. (2010) did not provide the detail of the method they used to decompose the wavefields.

As mentioned above, Liu et al. (2007, 2011) advocated using wavefield decomposition for RTM in the Fourier domain. Next section, I describe this method with both horizontal and vertical propagation decompositions.

RTM USING WAVEFIELD DECOMPOSITION

A common method of wavefield decomposition is to obtain wavefields' one-way components with respect to two orthogonal directions, i.e. horizontal directions (leftgoing/rightgoing) and vertical directions (upgoing/downgoing). Hereafter I use the notations $z+$ to represent the downgoing wave component with respect to the vertical depth axis, and $x+$ to represent the rightgoing wave component with respect to the horizontal space axis. Note that a *downgoing* wavefield indicates a wavefield where the vertical direction of propagation is downward with time running forward, whereas a *rightgoing* wavefield indicates a wavefield where the horizontal direction of propagation is rightward with time running forward.

Source and receiver wavefields can be decomposed into their downgoing and upgoing components as follows (Liu et al., 2007):

$$S(t,\vec{x}) = S_{z+}(t,\vec{x}) + S_{z-}(t,\vec{x}), \tag{3}$$

$$R(t,\vec{x}) = R_{z+}(t,\vec{x}) + R_{z-}(t,\vec{x}), \tag{4}$$

where $S_{z+}(t,\vec{x})$, $S_{z-}(t,\vec{x})$, $R_{z+}(t,\vec{x})$ and $R_{z-}(t,\vec{x})$ are the downgoing and upgoing source and receiver wavefields, respectively. Similarly, they can also be decomposed into their horizontal components as follows:

$$S(t,\vec{x}) = S_{x+}(t,\vec{x}) + S_{x-}(t,\vec{x}), \tag{5}$$

$$R(t,\vec{x}) = R_{x+}(t,\vec{x}) + R_{x-}(t,\vec{x}), \tag{6}$$

where $S_{x+}(t,\vec{x})$, $S_{x-}(t,\vec{x})$, $R_{x+}(t,\vec{x})$ and $R_{x-}(t,\vec{x})$ are the rightgoing and leftgoing source and receiver wavefields respectively.

Thus, the conventional imaging condition can be separated into four imaging conditions based on combinations of vertical wavefield components:

$$\begin{aligned} I(\vec{x}) &= \sum_{t=0}^{t_{\max}} S_{z+}(t,\vec{x})R_{z-}(t,\vec{x}) + \sum_{t=0}^{t_{\max}} S_{z-}(t,\vec{x})R_{z+}(t,\vec{x}) \\ &\quad + \sum_{t=0}^{t_{\max}} S_{z+}(t,\vec{x})R_{z+}(t,\vec{x}) + \sum_{t=0}^{t_{\max}} S_{z-}(t,\vec{x})R_{z-}(t,\vec{x}) \\ &= I_{z,+-}(\vec{x}) + I_{z,-+}(\vec{x}) + I_{z,++}(\vec{x}) + I_{z,--}(\vec{x}). \end{aligned} \tag{7}$$

The first term is the cross-correlation of the downgoing source and upgoing receiver wavefields, which is equivalent to the result of downward continuing migration. The second term is the cross-correlation of the upgoing source and downgoing receiver wavefields. Thus, the first two terms can represent any reflector that causes vertical backscattering. The images corresponding to these two terms are caused by backward-scattered events with respect to the vertical axis. The subimages from the remaining two terms can be considered as noise, which results from vertical forward-scattered events where source and receiver wavefields are vertically propagating in the same direction. Thus, the last two terms contribute to artifacts in RTM. However, these are related to tomographic information in wave-equation migration velocity analysis (Almomin et al., 2011).

In order to suppress RTM artifacts, the imaging condition that contains only the first two terms in Equation 7 (Liu et al., 2007, 2011) as

$$I_{\text{vert}}(\vec{x}) = \sum_{t=0}^{t_{\max}} S_{z+}(t,\vec{x})R_{z-}(t,\vec{x}) + \sum_{t=0}^{t_{\max}} S_{z-}(t,\vec{x})R_{z+}(t,\vec{x}). \tag{8}$$

This condition might be called the *vertical backscatter-based imaging condition.*

The *horizontal backscatter-based imaging condition* can be derived in a similar way. This horizontal imaging condition contains only cross-correlation of the source and receiver wavefields that are propagating in horizontally opposite directions:

$$I_{\text{horiz}}(\vec{x}) = \sum_{t=0}^{t_{\max}} S_{x+}(t,\vec{x})R_{x-}(t,\vec{x}) + \sum_{t=0}^{t_{\max}} S_{x-}(t,\vec{x})R_{x+}(t,\vec{x}). \tag{9}$$

The first term is the cross-correlation of the rightgoing source and leftgoing receiver wavefields. The second term is the cross-correlation of the leftgoing source and rightgoing receiver wavefields. Thus, this imaging condition represents any reflector that causes horizontal backscattering.

To use the benefits of both vertical and horizontal backscatter-based imaging conditions, they are added togather (Liu et al., 2011):

$$I(\vec{x}) = I_{\text{vert}}(\vec{x}) + I_{\text{horiz}}(\vec{x}). \tag{10}$$

This imaging condition might be called the *Cartesian backscatter-based imaging condition.* Note that the subimages from both vertical and horizontal imaging condition partly overlap, because of the overlap between the vertical and horizontal wavefield components. For example, the wavefield that is moving in any downright direction can be considered to be either downgoing or rightgoing.

In addition, the Cartesian backscatter-based imaging condition described above only involves the decomposition of source and receiver wavefields into their vertical or horizontal components. Thus, some backward-scattered events are indistinguishable from forward-scattered events using only this decomposition scheme. Therefore, the image illumination is reduced. However, Equation 10 produces no forward-scattered artifacts based on Cartesian directions of propagation. Thus, this proposed imaging condition can suppress artifacts in RTM images.

WAVEFIELD DECOMPOSITION

The conventional method for wavefield decomposition operates in the Fourier domain. This method was first applied to vertical seismic profiles (Hu and McMechan, 1987). Here, wavefields are decomposed into their upgoing and downgoing components in the F-K domain by using a 2-D fast Fourier transform (FFT):

$$\tilde{P}_{z+}(f,k_z) = \begin{cases} \tilde{P}(f,k_z) & \text{for } fk_z \geq 0 \\ 0 & \text{for } fk_z < 0 \end{cases}, \tag{11}$$

$$\tilde{P}_{z-}(f,k_z) = \begin{cases} 0 & \text{for } fk_z \geq 0 \\ \tilde{P}(f,k_z) & \text{for } fk_z < 0 \end{cases}, \tag{12}$$

where $\tilde{P}(f,k_z)$ is the 2-D Fourier transform of the wavefield $P(t,z)$ at any horizontal position x, and f and k_z are the frequency and vertical wavenumber representations of the wavefield. Note that the wavefield $P(t,z)$ can be either the source wavefield $S(t,z)$ or the receiver wavefield $R(t,z)$.

A similar method can be used to obtain the leftgoing and rightgoing components of wavefields:

$$\tilde{P}_{x+}(f,k_x) = \begin{cases} \tilde{P}(f,k_x) & \text{for } fk_x \geq 0 \\ 0 & \text{for } fk_x < 0 \end{cases}, \tag{13}$$

$$\tilde{P}_{x-}(f,k_x) = \begin{cases} 0 & \text{for } fk_x \geq 0 \\ \tilde{P}(f,k_x) & \text{for } fk_x < 0 \end{cases}, \tag{14}$$

where $\tilde{P}(f,k_x)$ is the 2-D Fourier transform of the wavefield $P(t,x)$ at any depth z, and f and k_x are the frequency and horizontal wavenumber representations of the wavefield.

In this paper, I apply smooth-cut F-K filters to the wavefield decomposition instead of using the sharp-cut filters shown in Equations 3 to 6, so that noise due to the FFT of discontinuous functions is reduced. Using smooth-cut filters might slightly reduce the illumination of reflectivity images, but it is worth attentuating the noise from FFT.

The decomposed wavefields, as in Equations 3 to 6, are the the 2-D inverse Fourier transforms of the decomposed wavefields derived from Equations 11 to 14. Thus, the terms

on the right-hand sides of Equations 3 to 6 are all complex. However, in each equation, the summation of the imaginary parts becomes zero, and the wavefield is equal to the summation of the real parts of the decomposed wavefields; for example,

$$\begin{aligned} S(t,\vec{x}) &= S_{z+}(t,\vec{x}) + S_{z-}(t,\vec{x}), & (15)\\ &= \mathrm{Re}[S_{z+}(t,\vec{x})] + \mathrm{Re}[S_{z-}(t,\vec{x})]. & (16) \end{aligned}$$

In their proposed imaging condition, Liu et al. (2007) applied only the real parts of the decomposed source and receiver wavefields to the imaging conditions in Equations 8 and 9. Thus, a question is raised about the effect that the imaginary components might have on the decomposed RTM images. Considering decomposed wavefields as complex functions, I rewrite Equations 8 and 9 to be

$$I_{\mathrm{vert}}(\vec{x}) = \sum_{t=0}^{t_{\max}} S^*_{z+}(t,\vec{x}) R_{z-}(t,\vec{x}) + \sum_{t=0}^{t_{\max}} S^*_{z-}(t,\vec{x}) R_{z+}(t,\vec{x}), \quad (17)$$

and

$$I_{\mathrm{horiz}}(\vec{x}) = \sum_{t=0}^{t_{\max}} S^*_{x+}(t,\vec{x}) R_{x-}(t,\vec{x}) + \sum_{t=0}^{t_{\max}} S^*_{x-}(t,\vec{x}) R_{x+}(t,\vec{x}), \quad (18)$$

where an asterisk denotes the complex conjugate. Equations 17 and 18 can be computed using only the wavefield decomposition in the T-K domain (Liu et al., 2011). Thus, this method is practical, because there is no need to save wavefields at every time step.

RESULTS AND DISCUSSION

I applied RTM using the imaging conditions proposed in the previous section to a single shot profile corresponding to the 1-D velocity model in Figure 2(a). The same approach was also applied to a synthetic data set derived from a 2-D velocity model windowed from the SEAM model. Because of its complicated structures, this synthetic data can illustrate the effectiveness of the proposed imaging conditions.

I found that the imaging condition that uses only the real parts of the decomposed wavefields (Equations 8 and 9) can produce the same images as the imaging condition that utilizes the complex-valued decomposed wavefields (Equations 17 and 18). Indeed, the difference between the RTM images derived from these two imaging conditions can be seen; however, the difference is negligible with respect to the RTM image. Thus, the imaginary parts have a negligible effect on the decomposed RTM image. Therefore, the imaging condition using only real parts of decomposed wavefields as in Equations 8 and 9 can replace the imaging conditions in Equation 17 and 18.

Examples with 1-D model

Figures 8(a) to 8(d) illustrate the vertically decomposed RTM images from a single shot profile without SRM corresponding to the model in Figure 2(a). These decomposed images are related to the four terms on the right-hand side of Equation 7. Figure 8(b) results from the upgoing source and downgoing receiver wavefields. This figure was expected to

contain less artifacts; however, there are many artifacts above the reflector, as shown in Figure 8(b). I found that these artifacts in Figure 8(b) were caused by cross-correlation between the upgoing component of the direct source wavefield and the downgoing component of the backscattered receiver wavefield. The upgoing component of the source wavefield that produces severe artifacts indeed propagates in almost horizontal directions, where it is indistinguishable from the downgoing source wavefield.

Figure 8(e) illustrates the RTM image corresponding to Equation 8, i.e. the vertical backscatter-based imaging condition. As shown in the figure, the horizontal reflector has been imaged well. However, there are artifacts at shallow depths resulting from those in Figure 8(b).

On the other hand, Figures 9(a) to 9(d) show the four horizontally different decomposed RTM images from a single shot profile without SRM corresponding to the model in Figure 2(a). Figure 9(e) illustrates the RTM image using the horizontal backscatter-based imaging condition as in Equation 9. The figure shows strong artifacts aligned vertically above the reflector. These artifacts have a similar origin as those in Figure 8(b), i.e. cross-correlation between the downgoing direct source wavefield and the backscattered receiver wavefield. Specifically, the component of the source wavefield that cause severe artifacts indeed propagates in almost vertical directions, where the leftgoing and rightgoing components of the source wavefield overlap.

Figure 10 illustrates the sum of the vertical and horizontal backscatter-based imaging conditions (Equations 10), i.e. the Cartesian backscatter-based imaging condition. As shown in the figure, the RTM artifacts at shallow depths have been attenuated except for the vertically aligned artifacts. These severe artifacts result from the cross-correlation between direct source wavefields and backscattered receiver wavefields. It can be seen that the RTM image using wavefield decomposition as in Figure 10 contains less artifacts compared to the conventional RTM image as in Figure 5(b). However, there is no significant improvement when comparing Figure 10 to the image after applying a low-cut filter (Figure 6).

Examples with complex 2-D model

Figure 11(a) shows a complex 2-D velocity model windowed from the SEAM model. This model is used for modeling a 2-D synthetic data set, which contains 69 shot profiles. This data set involves the deployment of a 1-D array of regularly spaced geophones (every 20 meters) on the entire surface. For each shot record, a source is exploded on the surface with 80-meter shot spacing. In the pre-migration process, muting head waves and diving waves is appied to this data set. However, SRM elimination is not implemented. For RTM, the smoothed velocity in Figure 11(b) is used.

Figure 12(a) illustrates a conventional RTM image using the cross-correlation imaging condition in Equation 1, whereas Figure 12(b) shows the same image after applying a causal low-cut filter. We can see that RTM artifacts are partially attenuated using the low-cut filter.

Figures 13(a) to 13(d) illustrate the vertically decomposed RTM images corresponding to the four decomposed images on the right-hand side of Equation 7. Figure 13(b) results from the upgoing source and downgoing receiver wavefields. This figure contains artifacts at shallow depths which have the same origin as those in Figure 8(b).

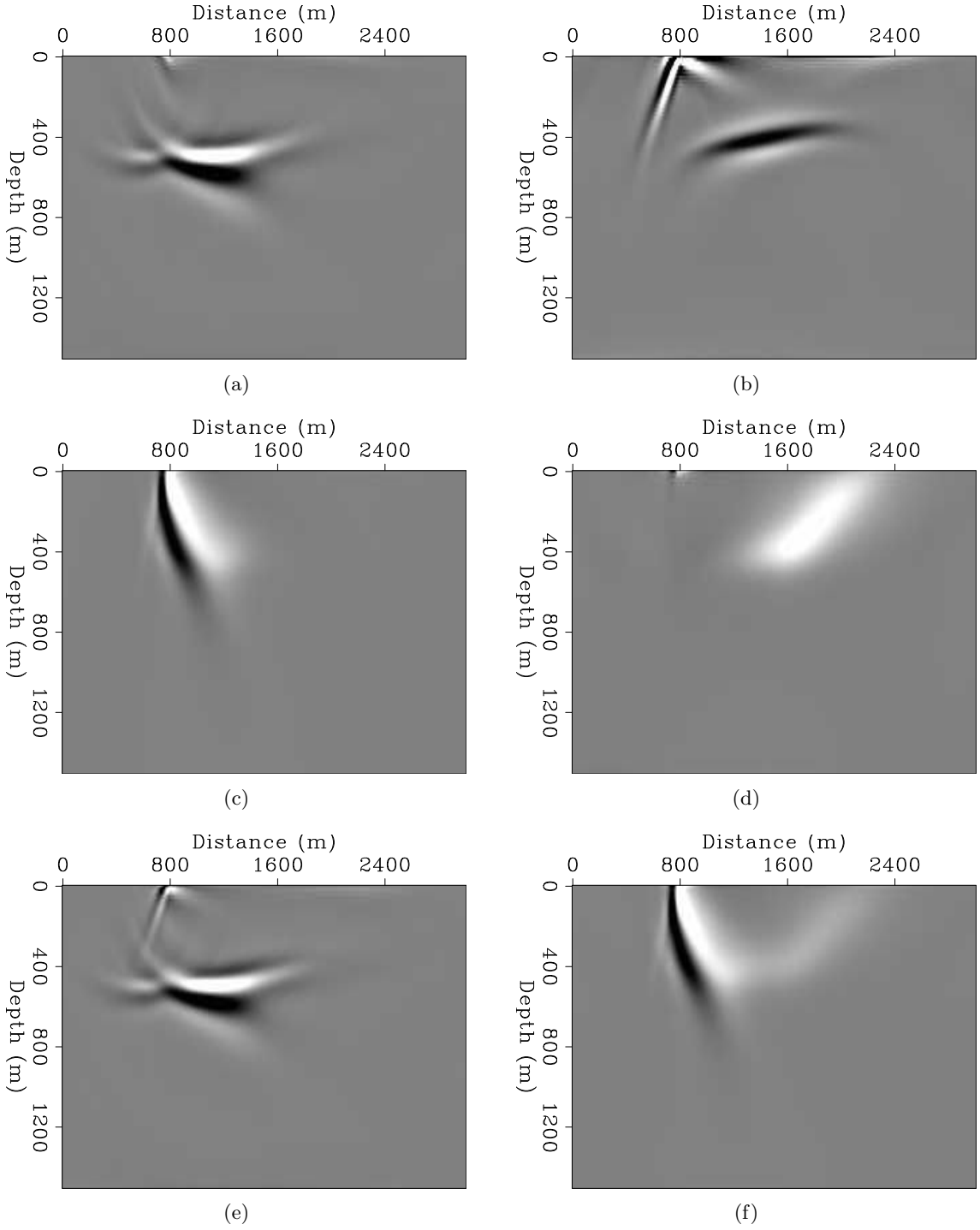


Figure 8: Different vertically decomposed RTM images from the single shot profile in Figure 4(a) along with smooth migration velocity in Figure 5(a): (a) using the downgoing source and upgoing receiver wavefields, (b) using the upgoing source and downgoing receiver wavefields, (c) using the downgoing source and downgoing receiver wavefields, and (d) using the upgoing source and upgoing receiver wavefields. (e) The sum of the images (a) and (b) equivalent to Equation 8. (f) The sum of the images (c) and (d). [**ER**] kittinat/. v0-SdRu,v0-SuRd,v0-SdRd,v0-SuRu,v0-SdRuPSuRd,v0-SdRdPSuRu

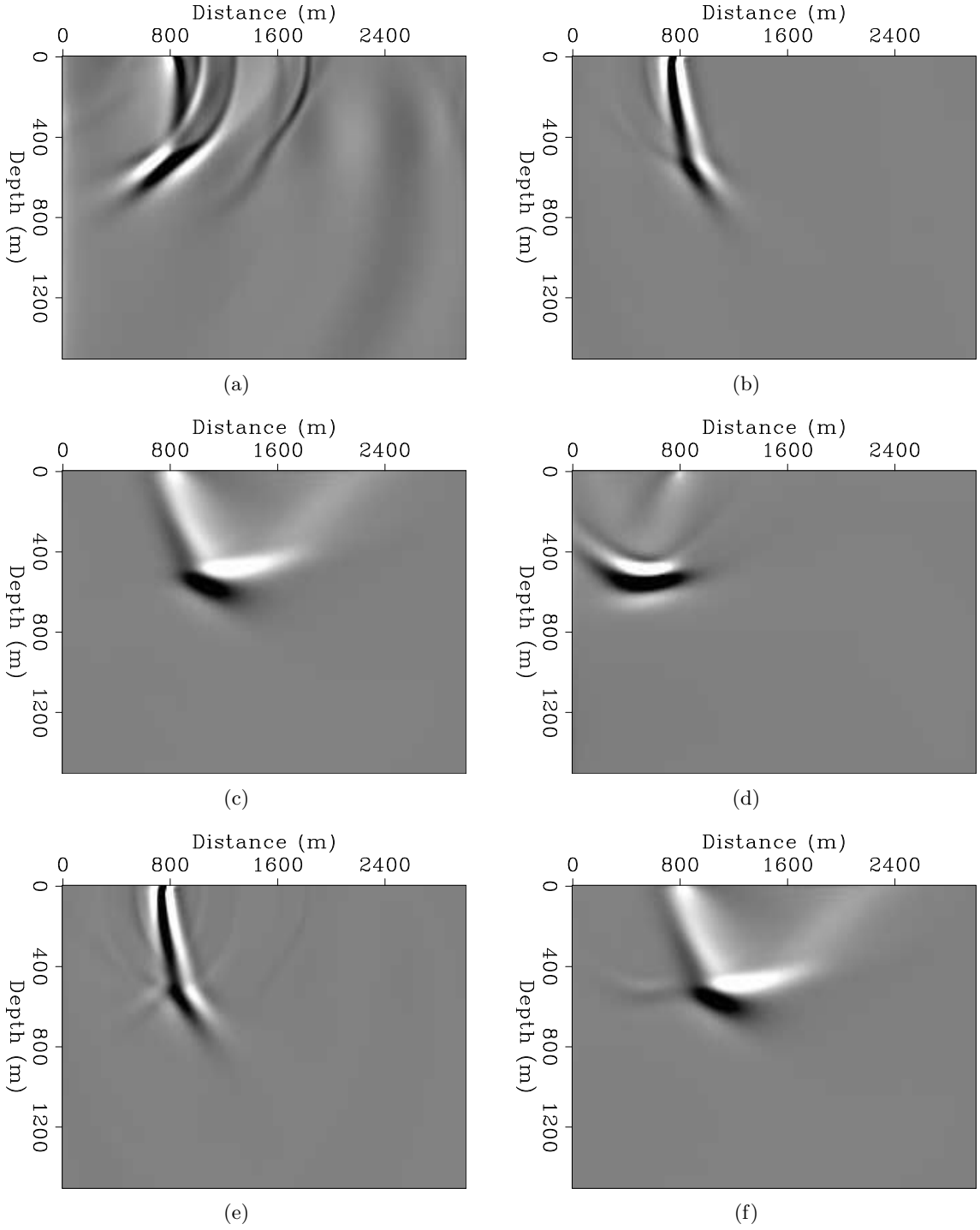


Figure 9: Different horizontally decomposed RTM images from the single shot profile in Figure 4(a) along with smooth migration velocity in Figure 5(a): (a) using the rightgoing source and leftgoing receiver wavefields, (b) using the leftgoing source and rightgoing receiver wavefields, (c) using the rightgoing source and rightgoing receiver wavefields, and (d) using the leftgoing source and leftgoing receiver wavefields. (e) The sum of the images (a) and (b) equivalent to Equation 9. (f) The sum of the images (c) and (d). [**ER**] kittinat/. v0-SrRl,v0-SlRr,v0-SrRr,v0-SlRl,v0-SrRlPSlRr,v0-SrRrPSlRl

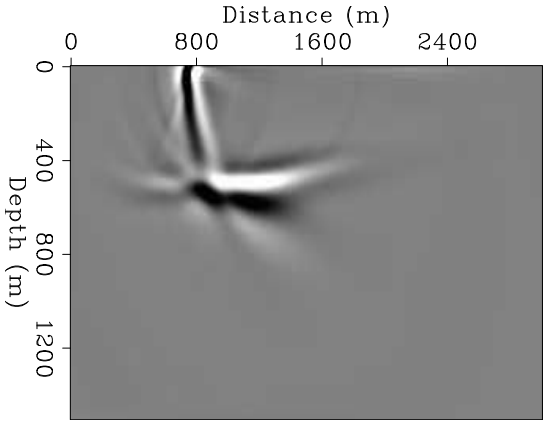


Figure 10: The RTM image from the Cartesian backscatter-based imaging condition as in Equation 10, i.e. the sum of Figure 8(e) and Figure 9(e). [**ER**] kittinat/. v0-vertPhoriz

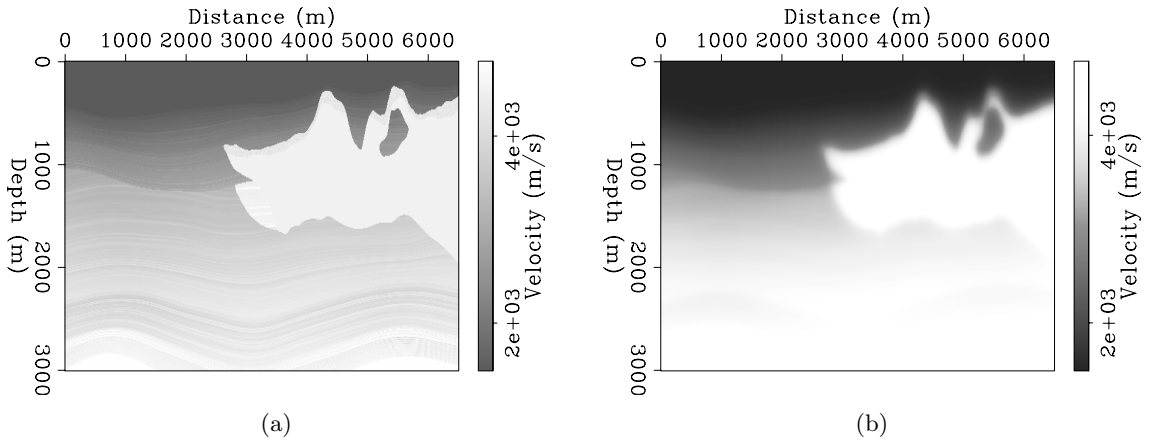


Figure 11: (a) A 2-D velocity model windowed from the SEAM model. (b) The smoothed velocity model used as the migration velocity. [**ER**] kittinat/. vel1,vel1-mig

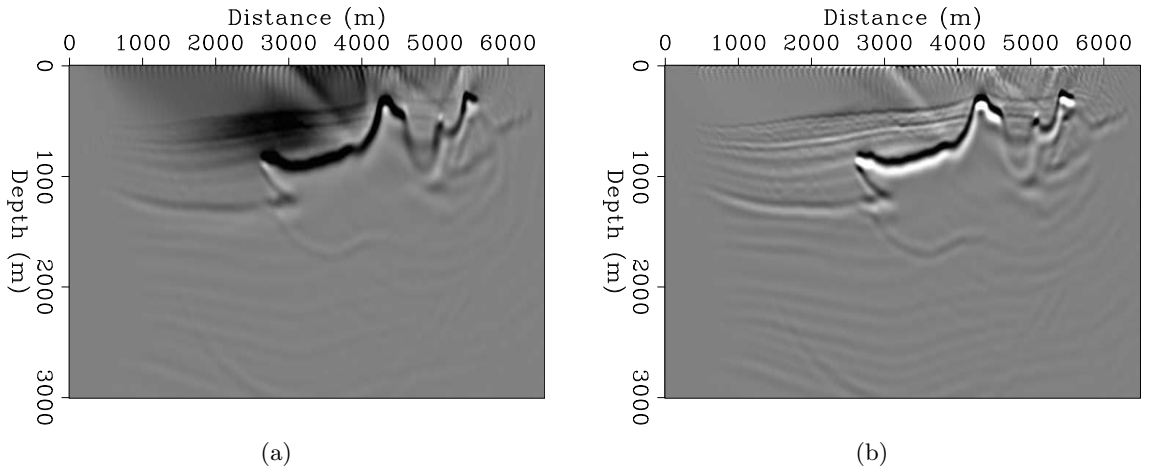


Figure 12: (a) A conventional RTM image of an entire 2-D synthetic data set using the cross-correlation imaging condition in Equation 1. (b) A causal low-cut filtered result of the image in (a). [**CR**] kittinat/. crtm0-v1,crtm0-v1-lowcut

Figure 13(e) illustrates the RTM images the vertical backscatter-based imaging condition. As shown in the figure, the illumination of steeply dipping reflectors is poor. In contrast, Figure 13(f) represents the decomposed images from the decomposed source and receiver wavefields that have been propagated in vertically opposite directions.

On the other hand, Figures 14(a) to 14(d) show the four different decomposed RTM images based on horizontal propagation directions of source and receiver wavefields. We can see that Figures 14(a) and 14(b) illustrate not only the images of steeply dipping reflectors but also artifacts in the shallow parts. These artifacts have the same origin as those in Figure 9(e).

Figure 14(e) is the result of using the imaging condition in Equation 9, i.e. the horizontal imaging condition. As shown in the figure, this imaging condition can image steeply dipping reflectors to an extent; however, most of low-angle reflectors cannot be imaged. In contrast, Figure 14(f), which involves the cross-correlation of the decomposed source and receiver wavefields caused by horizontally forward-scattered events, contains the major artifacts in the shallow parts of the model and the images of low-angle reflectors.

Figure 15 illustrates the result of using the Cartesian backscatter-based imaging condition, i.e. Equations 10. As shown in the figure, the result has the benefits of both vertical and horizontal imaging conditions. Both steeply dipping and low-angle reflectors are well imaged. In addition, Figure 15 contains artifacts in the shallow parts as a result of Figures 13(b), 14(a), and 14(b). However, strong RTM artifacts have significantly been attenuated.

It can be seen that the RTM image (Figure 15) using the Cartesian backscatter-based imaging condition, as in Equation 10, contains less strong artifacts compared to the conventional RTM image (Figure 12(a)) and the low-cut filtered RTM image (Figure 12(b)).

Thus, the combination of the vertical and horizontal imaging conditions is an effective technique for attenuating typical RTM artifacts, especially for complex velocity models. However, further investigation is necessary for removing the remaining artifacts in the shallow parts of the migrated images.

A possible solution is to modify the F-K filters in Equations 11 to 14, so there is no overlap between the decomposed wavefields with opposite propagation directions. In this study, I also tried using fan filters to decompose wavefields in the F-K domain. These filters aimed to limit the aperture of propagation directions of wavefields. However, as a result of the FFT of discontinuous functions, there still remained the overlap between the decomposed wavefields with opposite propagation directions.

CONCLUSIONS

RTM can image steeply dipping reflectors and overturned events. However, the conventional correlation-based imaging condition in RTM always produces strong low-frequency artifacts. To suppress these artifacts, an imaging condition using appropriate decomposed source and receiver wavefields based on their propagation direction has been applied and examined. This imaging condition contains only the backward scattered events based on Cartesian directions. The imaging condition discussed in this paper successfully attenuates such artifacts in complex velocity models compared to applying a low-cut filter. However,

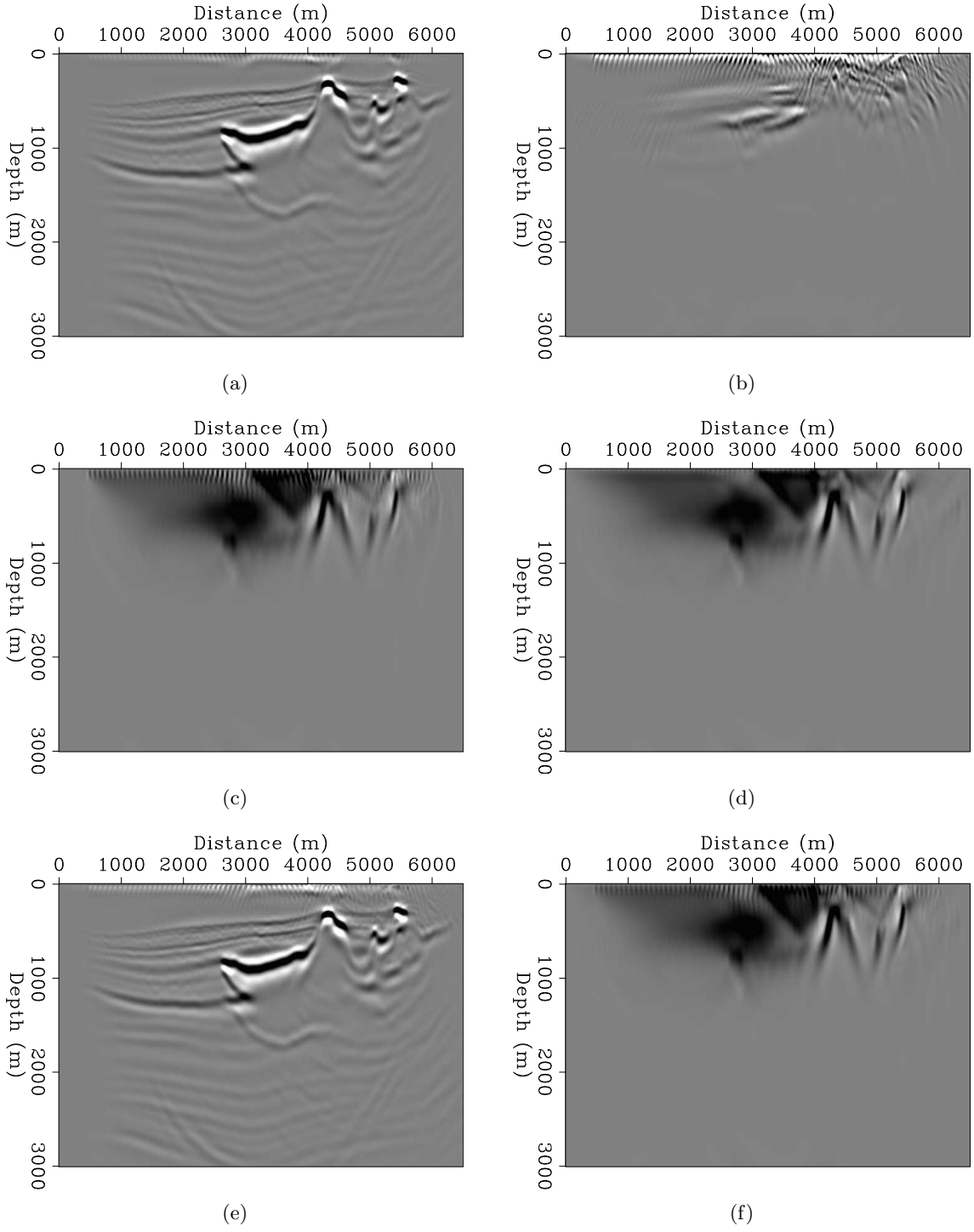


Figure 13: Different vertically decomposed RTM images from 2-D synthetic data along with the smoothed migration velocity in Figure 11(b): (a) using the downgoing source and upgoing receiver wavefields, (b) using the upgoing source and downgoing receiver wavefields, (c) using the downgoing source and downgoing receiver wavefields, and (d) using the upgoing source and upgoing receiver wavefields. (e) The sum of the images (a) and (b) equivalent to Equation 8. (f) The sum of the images (c) and (d). [**CR**] kittinat/. v1-SdRu,v1-SuRd,v1-SdRd,v1-SuRu,v1-SdRuPSuRd,v1-SdRdPSuRu

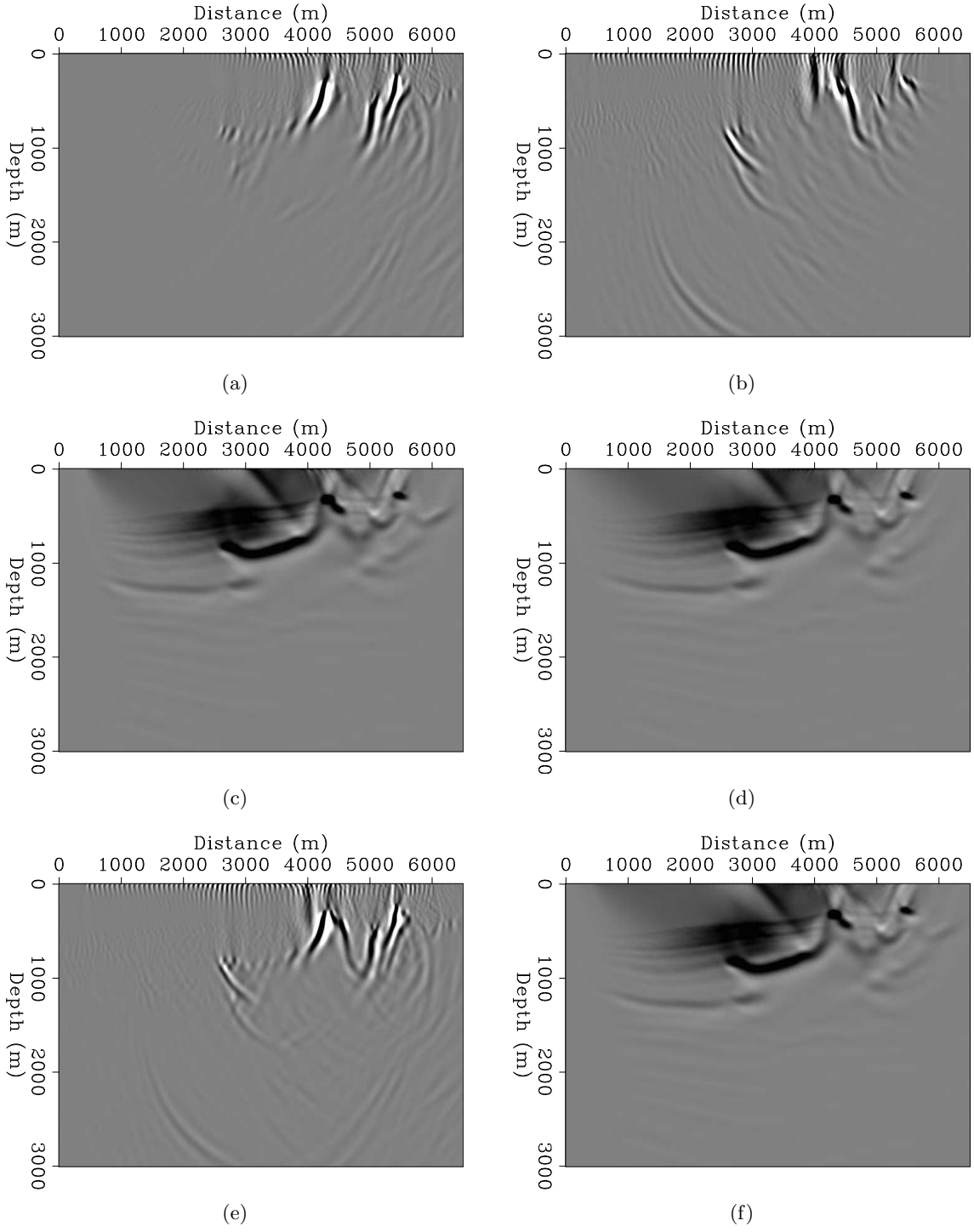


Figure 14: Different horizontally decomposed RTM images from 2-D synthetic data along with the smoothed migration velocity in Figure 11(b): (a) using the rightgoing source and leftgoing receiver wavefields, (b) using the leftgoing source and rightgoing receiver wavefields, (c) using the rightgoing source and rightgoing receiver wavefields, and (d) using the leftgoing source and leftgoing receiver wavefields. (e) The sum of the images (a) and (b) equivalent to Equation 9. (f) The sum of the images (c) and (d). [**CR**] kittinat/. v1-SrRl,v1-SlRr,v1-SrRr,v1-SlRl,v1-SrRlPSlRr,v1-SrRrPSlRl

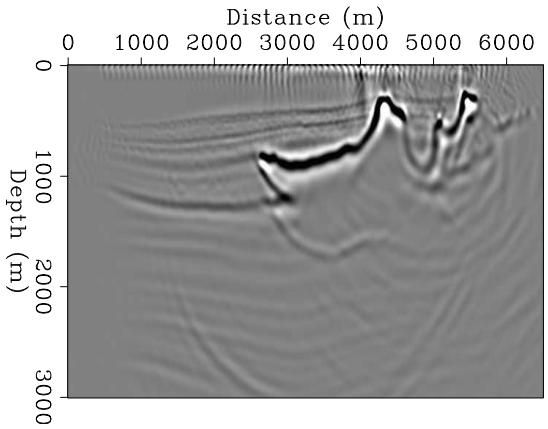


Figure 15: The RTM image from the Cartesian backscatter-based imaging condition as in Equation 10, i.e. the sum of Figure 13(e) and Figure 14(e). [**CR**] kittinat/. v1-vertPhoriz

there are still remaining artifacts in the decomposed RTM images. These are caused by the cross-correlation between the upgoing components of the direct source wavefields and the downgoing components of the backscattering of receiver wavefields. Thus, further investigation is still necessary to remove these residual artifacts.

ACKNOWLEDGMENTS

I would like to thank Biondo Biondi, Mandy Wong, Sjoerd de Ridder, Ali Almomin, and the other SEP colleagues for their valuable comments and suggestions. I also thank Exxon Mobil for providing the SEAM Phase 1 model used in this study.

REFERENCES

Almomin, A., B. Biondi, and R. Clapp, 2011, Combining forward-scattered and back-scattered wavefields in velocity analysis: SEP-Report, **143**, 77–84.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524.

Biondi, B. L., 2005, 3-d seismic imaging: Stanford Exploration Project.

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

Chattopadhyay, S. and G. A. McMechan, 2008, Imaging conditions for prestack reverse-time migration: Geophysics, **73**, S81–S89.

Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Science Inc.

Clayton, R. W. and B. Engquist, 1980, Absorbing boundary conditions for wave-equation migration: Geophysics, **45**, 895–904.

Dablain, M. A., 1986, The application of high-order differencing to the scalar wave equation: Geophysics, **51**, 54–66.

Etgen, J., S. H. Gray, and Y. Zhang, 2009, An overview of depth imaging in exploration geophysics: Geophysics, **74**, WCA5–WCA17.

Fei, T. W., Y. Luo, S. Aramco, and G. T. Schuster, 2010, De-blending reverse-time migration: SEG Technical Program Expanded Abstracts, **29**, 3130–3134.

Fletcher, R. F., P. Fowler, P. Kitchenside, and U. Albertin, 2006, Suppressing unwanted internal reflections in prestack reverse-time migration: Geophysics, **71**, E79–E82.

Guitton, A., B. Kaelin, and B. Biondi, 2006, Least-square attenuation of reverse-time migration artifacts: SEG Technical Program Expanded Abstracts, **25**, 2348–2352.

Hu, L.-Z. and G. A. McMechan, 1987, Wave-field transformations of vertical seismic profiles: Geophysics, **52**, 307–321.

Liu, F., G. Zhang, S. A. Morton, and J. P. Leveille, 2007, Reverse-time migration using one-way wavefield imaging condition: SEG Technical Program Expanded Abstracts, **26**, 2170–2174.

———, 2011, An effective imaging condition for reverse-time migration using wavefield decomposition: Geophysics, **76**, S29–S39.

Robein, E., 2010, Seismic imaging: A review of the techniques, their principles, merits and limitations: EAGE.

Suh, S. Y. and J. Cai, 2009, Reverse-time migration by fan filtering plus wavefield decomposition: SEG Technical Program Expanded Abstracts, **28**, 2804–2808.

Valenciano, A. A. and B. Biondi, 2003, 2-d deconvolution imaging condition for shot-profile migration: SEG Technical Program Expanded Abstracts, **22**, 1059–1062.

Whitmore, N. D., 1983, Iterative depth migration by backward time propagation: SEG Technical Program Expanded Abstracts, **2**, 382–385.

Xie, X.-B. and R.-S. Wu, 2006, A depth migration method based on the full-wave reverse-time calculation and local one-way propagation: SEG Technical Program Expanded Abstracts, **25**, 2333–2337.

Yoon, K. and K. J. Marfurt, 2006, Reverse-time migration using the poynting vector: Exploration Geophysics, **37**, 102–107.

Yoon, K., K. J. Marfurt, and W. Starr, 2004, Challenges in reverse-time migration: SEG Technical Program Expanded Abstracts, **23**, 1057–1060.

Zhang, Y. and J. Sun, 2009, Practical issues in reverse time migration: True amplitude gathers, noise removal and harmonic source encoding: First Break, **27**, 53–59.

Elastic wavefield directionality vectors

Ohad Barak

ABSTRACT

The direction of energy propagation at every point in a wavefield propagated using an elastic finite-difference method can be deduced from the displacement vectors. The pressure-wave amplitude and displacements can be separated from the shear-wave amplitude and displacements using a separation operator. Amplitude separation can be used to create a more informative image by correctly imaging converted waves. Displacement separation enables calculation of the propagation direction, which can be used to create angle gathers without utilizing extended imaging conditions.

INTRODUCTION

When modeling elastic wave propagation, both pressure and shear wave energy is propagated. This enables a more physical reverse-time propagation of 3-C receiver data. However, the cross-correlation imaging condition (Claerbout, 1971) is not sensitive to the type of wave being modeled, and this can result in a migrated image containing many artifacts without physical meaning. Therefore, when imaging with elastically propagated wavefields, a method of separating the two wave types is required. For isotropic media, the Helmholtz separation (Morse and Feshbach, 1953) separates the wavefield into a scalar P amplitude field and a vector S amplitude field. Yan and Sava (2008) show how this separation can be used to image primary modes (PP or SS) and converted modes (PS or SP) separately, resulting in an image with a more physical meaning. Yan and Sava (2009) extend this method to VTI and TTI media. Though these are effective methods, they do not separate the actual displacement fields into P displacements and S displacements. Zhang and McMechan (2010) deal with this problem directly, developing an elastic wavefield decomposition into P and S displacements, while preserving the original physical units, phase, amplitude and vector characteristics of the modeled field. They show an implementation for both isotropic and anisotropic media.

One method of constructing angle domain common-image gatherss by wavefield methods is the extended imaging condition by Sava and Fomel (2003), which requires performing cross-correlations between source and receiver wavefields at several horizontal spatial lags. A transform is then applied to these correlated wavefields to create the angle gathers. An alternate ray-based method is presented by Koren et al. (2008), which involves shooting ray-pairs from the subsurface points at which angle gathers are required, and saving the travel times to the surface and the surface location of the ray's emergence. The trace data at those locations is then interpolated from nearby recorded seismic data, and is copied into the subsurface image point's angle gather, at the opening angle of the ray pair. This methodology also incorporates dip gathers, by considering various tilting angles of the plane of the ray pairs, in relation to the vertical direction.

The angle gather construction methodology I wish to implement is a conceptual hybrid of these wavefield and ray-based methods. In the isotropic case, the direction of P wave propagation is tangent to the particle displacement direction, while the S wave's propagation direction is perpendicular to the displacements. If the vector displacement field (which is calculated as part of elastic wavefield finite-difference algorithm) can be decomposed into P and S wave displacements, it follows that the direction of polarization of each wave mode can, in principle, be determined. Under the isotropic assumption, the wave propagation direction can be readily deduced from the polarization direction. This is true only if, at a certain model point at a particular time window, only one wave exists. If this is indeed the case, then the displacement vector will be along a certain line, e.g. - the wave will be linearly polarized. Such information constitutes an "arrow" in space and time, indicating the wave's propagation direction within a time window. In this respect, this arrow can be likened to a ray, since it points toward the propagation direction of the wavefront. Having a directionality determination capability for finite-difference wavefield propagation methods is an exciting prospect, since the consequence is an ability to construct angle gathers *during wavefield propagation*, without applying any additional transforms to the wavefields, as in the extended imaging condition by Sava and Fomel (2003). It also makes the expensive data gathering step required by the method developed in Koren et al. (2008) unnecessary.

The scope of this report is much more limited than actual angle gather construction for elastic wavefields, and my purpose here is only to show whether it is possible to acquire P and S wave polarization directions during wavefield propagation. Furthermore, I limit myself to an isotropic medium assumption. I decompose each wavefield into P and S displacements using Zhang and McMechan (2010)'s method, and deduce from the displacements the angle of the polarization direction of each wave mode to the vertical axis, at each point in the modeling space, and at each time. From these angles, it is possible to ascertain the wave propagation direction. Angle gathers can eventually be constructed by summing the source and receiver wave's propagation angle to the vertical. This will constitute the "opening angle". The value of the angle gather at the particular opening angle will be the cross-correlation product of the source and receiver wavefields at the imaging point.

THEORETICAL BACKGROUND

Elastic wavefield modeling

The elastic isotropic wave equation in index notation reads:

$$\frac{\partial \sigma_{ii}}{\partial x_i} + \frac{\partial \sigma_{ij}}{\partial x_j} + f_i(\mathbf{X}, t) = \rho \ddot{U}_i, \tag{1}$$

where σ_{ii} are the normal stresses, σ_{ij} are the transverse stresses, f_i is the source function in direction i, X is the spatial source location operating at time t, ρ is density and U is the displacement. The stresses are propagated using the stress-displacement relation:

$$\sigma_{ii} = (\lambda + 2\mu)\frac{\partial U_i}{\partial x_i} + \lambda \frac{\partial U_j}{\partial x_j}, \tag{2}$$

$$\sigma_{ij} = \mu \left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} \right), \tag{3}$$

where λ and μ are the Lame elastic constants.

The finite-difference implementation follows the staggered grid methodology of Virieux (1986). The code I am using implements the variable grid size finite difference method, developed by Wu and Harris (2002), although I have not used the grid size variability so far. The code performs 3D elastic wavefield propagation, and was ported to Fortran90 by Robert Clapp of SEP.

P-wave and S-wave amplitude separation and displacement decomposition

The Helmholtz amplitude separation is based on the assumption that any vector field can be described as resulting from a combination of a scalar and vector potential fields:

$$\mathbf{U} = \nabla\Phi + \nabla \times \mathbf{\Psi}, \tag{4}$$

Where Φ is the scalar potential field and $\mathbf{\Psi}$ is the vector potential. $\mathbf{U}$ is the elastic displacement vector wavefield. The scalar potential generates pressure waves, and the vector potential generates shear waves. Therefore, the Helmholtz method of separating the P-wave amplitude from the S-wave amplitude is to apply a divergence operator and a curl operator to the displacement field:

$$P = \nabla \cdot \mathbf{U} = \nabla^2\Phi; \tag{5}$$

$$\mathbf{S} = \nabla \times \mathbf{U} = -\nabla^2\Psi. \tag{6}$$

The derivation of P-wave and S-wave displacement decomposition by Zhang and McMechan (2010) is done in the wavenumber domain. For an isotropic medium, the linear equation system they arrive at is:

$$\mathbf{K} \times \tilde{\mathbf{U}}^{\mathbf{P}} = 0, \tag{7}$$

$$\mathbf{K} \times \tilde{\mathbf{U}} = \mathbf{K} \times \tilde{\mathbf{U}}^{\mathbf{S}}, \tag{8}$$

$$\mathbf{K} \cdot \tilde{\mathbf{U}}^{\mathbf{S}} = 0, \tag{9}$$

and

$$\mathbf{K} \cdot \tilde{\mathbf{U}} = \mathbf{K} \cdot \tilde{\mathbf{U}}^{\mathbf{P}}, \tag{10}$$

where $\tilde{\mathbf{U}}$ is the 3D spatial fourier transform of the displacement field $\tilde{\mathbf{U}} = (\tilde{U}_x, \tilde{U}_y, \tilde{U}_z)$. $\tilde{\mathbf{U}}^{\mathbf{P}}$ and $\tilde{\mathbf{U}}^{\mathbf{S}}$ are the unknown P and S displacements. $\mathbf{K} = (K_x, K_y, K_z)$ is the wavenumber vector that describes the particle displacement direction. $K_i = \frac{\omega}{V_i}$, where V_i is the phase velocity in the i direction, and ω is angular frequency.

The solutions for P-wave displacements of these systems in 2D are:

$$\tilde{U}_x^P = K_x^2\tilde{U}_x + K_xK_z\tilde{U}_z, \tag{11}$$

$$\tilde{U}_z^P = K_z^2\tilde{U}_z + K_zK_x\tilde{U}_x. \tag{12}$$

The solutions for S-wave displacements are:

$$\tilde{U}_x^S = K_z^2\tilde{U}_x - K_xK_z\tilde{U}_z, \quad (13)$$
$$\tilde{U}_z^S = K_x^2\tilde{U}_z - K_zK_x\tilde{U}_x. \quad (14)$$

It is important to note that the K in these equations is normalized by the absolute value of the wavenumber $|\mathbf{K}|$. Therefore, if we use K' to designate the non-normalized wavenumbers, each wavenumber must be divided by $\frac{1}{\sqrt{K_x'^2+K_z'^2}}$. Equations 12 - 14 then take the form:

$$\tilde{U}_x^P = \frac{K_x'^2}{K_x'^2+K_z'^2}\tilde{U}_x + \frac{K_x'K_z'}{K_x'^2+K_z'^2}\tilde{U}_z, \quad (15)$$
$$\tilde{U}_z^P = \frac{K_z'^2}{K_x'^2+K_z'^2}\tilde{U}_z + \frac{K_z'K_x'}{K_x'^2+K_z'^2}\tilde{U}_x, \quad (16)$$

$$\tilde{U}_x^S = \frac{K_z'^2}{K_x'^2+K_z'^2}\tilde{U}_x - \frac{K_x'K_z'}{K_x'^2+K_z'^2}\tilde{U}_z, \quad (17)$$
$$\tilde{U}_z^S = \frac{K_x'^2}{K_x'^2+K_z'^2}\tilde{U}_z - \frac{K_z'K_x'}{K_x'^2+K_z'^2}\tilde{U}_x. \quad (18)$$

As mentioned previously, these are decomposition operators for isotropic media only. I decided to use them initially, in order to evaluate the possibility of using elastic wavefield directionality for the purposes of angle gather creation. The main difference between displacement decomposition and the Helmholtz amplitude separation is that the decomposition is reversible. The sum of the decomposed P and S displacement vectors will result in the original pre-decomposed displacements. However, there is no theoretical way to go from the separated P and S amplitude fields back to the original displacement fields from which they were calculated.

P-wave and S-wave decomposition in the space domain

The numerators in equations 16 - 18 can be calculated in the space domain by a finite-difference approximation to the 2^{nd} spatial derivative operator $\frac{\partial^2}{\partial x_i^2}$. This finite-difference operator can be convolved over the displacement field $\mathbf{U}$ in the space domain. However, the denominator in equations 16 - 18, a division in the wavenumber domain, can only be approximated in the space domain by a deconvolution. I chose to perform this deconvolution using spectral factorization and helical deconvolution, following the implementation in Claerbout (1997). This method has the advantage of treating multidimensional problems as one dimensional problems. Specifically, it enables execution of multidimensional deconvolutions as 1-D deconvolutions.

Using spectral factorization, a series of coefficients can be transformed to an alternate set of causal filter coefficients which have a causal inverse. The Wilson-Burg spectral factorization method (Fomel et al., 2003) ensures that the filter is minimum-phase. The autocorrelation

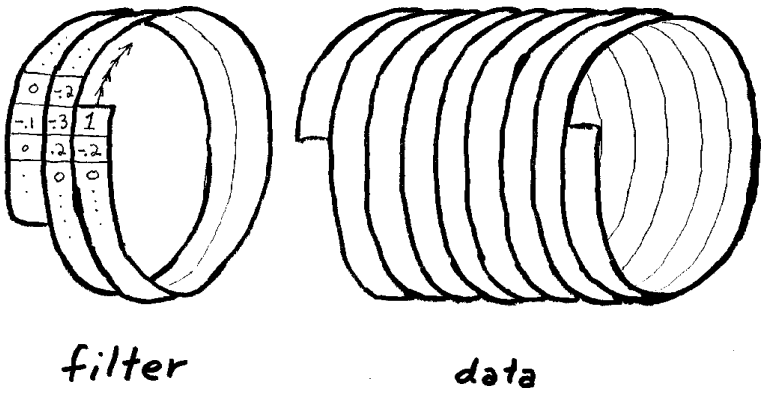


Figure 1: Sketch of the helix concept - convolution takes place by winding a "coil" of filter coefficients over a "coil" of data values (Claerbout, 1997). [**NR**] ohad1/. coil

of this new set of filter coefficients recreates the original values of the input series. The upshot of this is that application of the original series' coefficients to a dataset is akin to convolving the data with the spectrally factorized filter coefficients in one direction, and then convolving again in the other direction ("coiling" and then "uncoiling" the filter coefficients over the data). This effectively applies the filter and its time reverse (adjoint) to the data, which amounts to multiplying the data by the original input series' coefficients. In the case of finite differencing, the "input" series might be the Laplacian, which when made to traverse over the data has the effect of a 2^{nd} derivative approximation.

Application of a 2^{nd} derivative finite-difference operator to a dataset $\mathbf{U}$ is done by:

$$\frac{\partial^2}{\partial x_i^2} U_i = h^{'} * h * \mathbf{U}, \tag{19}$$

where h is the spectrally factorized filter coefficients and $h^{'}$ is the time-reversed filter. However, since equations 16 - 18 denote division of the displacement fields by a 2^{nd} derivative operator, a deconvolution with the filter coefficients is required:

$$\frac{1}{\frac{\partial^2}{\partial x_i^2}} U_i = \left(h^{'} * h \right)^{-1} \mathbf{U} = h^{-1} * (h^{'})^{-1} * \mathbf{U}. \tag{20}$$

The method of performing Zhang and McMechan (2010)'s displacement decomposition in the space domain is to first decide on the order of the 2^{nd} derivative finite difference operator, and then use spectral factorization to produce the filter coefficients h of this operator. Then, deconvolution with these filter coefficients must be applied to each displacement field, and the derivatives shown in equations 16 - 18 must then be performed on these deconvolved displacement fields:

$$U_x^P = \frac{\partial^2}{\partial x^2}\left(h^{-1} * (h^{'})^{-1} * U_x\right) + \frac{\partial^2}{\partial x \partial z}\left(h^{-1} * (h^{'})^{-1} * U_z\right), \quad (21)$$

$$U_z^P = \frac{\partial^2}{\partial z^2}\left(h^{-1} * (h^{'})^{-1} * U_z\right) + \frac{\partial^2}{\partial z \partial x}\left(h^{-1} * (h^{'})^{-1} * U_x\right), \quad (22)$$

$$U_x^S = \frac{\partial^2}{\partial z^2}\left(h^{-1} * (h^{'})^{-1} * U_x\right) - \frac{\partial^2}{\partial x \partial z}\left(h^{-1} * (h^{'})^{-1} * U_z\right), \quad (23)$$

$$U_z^S = \frac{\partial^2}{\partial x^2}\left(h^{-1} * (h^{'})^{-1} * U_z\right) - \frac{\partial^2}{\partial z \partial x}\left(h^{-1} * (h^{'})^{-1} * U_x\right). \quad (24)$$

I use the SEPlib `polydiv` module to perform the helical deconvolution of the spectrally factorized filter coefficients with the displacement wavefields, and then apply the spatial derivative operators, saving the decomposed results $\left(U_x^P, U_z^P, U_x^S, U_z^S\right)$ in separate arrays.

Determination of polarity and its angle to the vertical direction

My assumption is that if a single wave (P or S) exists within the modeled wavefield in a particular location, then the particle displacement there will have a distinct direction vector in space. In other words, the spatial components of the displacement vector will be linearly polarized when viewed along the time axis. As a measure of this polarization, which I call the *"Polarization Coefficient,"* I use the absolute value of statistical correlation between the displacement components:

My assumption is that if in any location within the modeled wavefield a single wave of a particular type (P or S) exists, then the particle displacement there will have a distinct direction vector in space. In other words, the spatial components of the displacement vector will be linearly polarized when viewed along the time axis. As a measure of this polarization, which I call the *"Polarization Coefficient"*, I use the absolute value of statistical correlation between the displacement components:

$$r_{ij} = \frac{|\mathrm{Cov}(U_i, U_j)|}{\mathrm{Var}(U_i)\mathrm{Var}(U_j)}. \quad (25)$$

The logic behind this is that if the energy going through a certain model point is indeed polarized, then there will be a linear dependence between the displacement components, and r will tend to 1. Otherwise, it will tend to 0. However, this measure is unstable where one of the displacement components is much smaller than the other (as shown in the following figures), where we would expect the polarization to actually be very specific (i.e. - along the larger component). The calculation of the polarization coefficient should be applied only within specific time windows during wavefield propagation, since conceptually, if a wavefield is complicated enough and is propagated for a certain length of time, then there will be many waves propagating in many directions, and no polarization can be expected at any one location for the entire propagation time.
Equation 25 can be used to produce a displacement correlation field. This field, in turn, will tell us to what degree the wavefield is polarized within a particular time window during propagation.

Determination of correlation time window size

The size of the time window within which the degree of polarization should be estimated is a major factor in the entire method, and can be considered as its Achilles' heel, it being a distinct point of failure (or success). On a conceptual basis, I think that the time window size should be related to the time period of the most energetic wave in the wavefield. Tying the window size to this parameter should enable us to decide on the resolution at which we wish to observe the wavefield's polarity, as a function of multiples of the most dominant frequency. However, this is useful only if there is a distinct dominant frequency. Though this may be the case for the source wavefield, the receiver wavefield may have a more variable spectrum. It is always possible to overestimate the correlation time window size, but that may result in a displacement correlation field with unuseful low values. This point remains unclear, and will likely require a lot of tuning to get right. At this stage, I define the time window size to be:

$$T = n_w/(f_c\sqrt{2}), \tag{26}$$

where T is the time window size in seconds, n_w is a parameter indicating the desired number of wave periods, and f_c is the central frequency of the source wavelet which, for the source wavefield only, is also the dominant frequency.

Estimation of angle between polarization direction and vertical direction

Only where there is a large polarization coefficient is it worthwhile to attempt to estimate the angle of the polarization direction to the vertical axis, within that time window. The slope of the line with the best fit to the crossploted displacements is estimated by linear regression. The angle to the vertical is the arctangent of this slope. Figure 2 explains this idea.

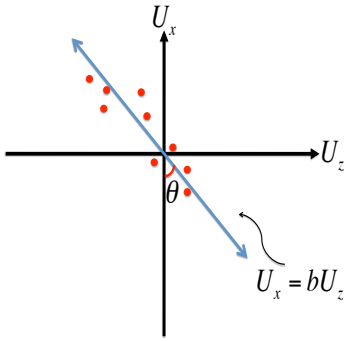


Figure 2: Crossplot of 2D displacement components of a nearly linearly polarized wave at some model location. The red dots represent samples of the displacements within a time window. The blue arrow indicates the polarization vector acquired by linear fitting of the displacement values. [**NR**] ohad1/. UxUz-example

Considering the displacement values U_x and U_z within a time window as two independent series, the standard linear regression equation can be applied to determine the slope of the line that best fits these series in the least-square sense:

$$U_x = bU_z, \tag{27}$$

where b is the slope, and is calculated by

$$b = \frac{\sum_{i=1}^{n} U_{x_i} U_{z_i} - n\overline{U_x}\,\overline{U_z}}{\sum_{i=1}^{n} U_{z_i}^2 - n\overline{U_z}}, \tag{28}$$

where $\overline{U_x}$ and $\overline{U_z}$ are the averages of the displacement values in the x and z directions respectively, within the the time window of length n time steps.

The angle between the linear fitting to the vertical axis is

$$\theta = \left| tan^{-1}(b) \right| . \tag{29}$$

Notice that the polarization line could "point" toward either direction, which is why I defined $\theta = \left| tan^{-1}(b) \right|$. This means that for $\theta = 0^o$, the polarization is in the $+U_z$ or in the $-U_z$ direction. In the case of a P wave, the interpretation is that within this time window, the wave at this particular location is propagating in the $+U_z$ or in the $-U_z$ direction. In the case of an S wave, the interpretation is that the wave is propagating in the $+U_x$ or in the $-U_x$ direction.

From propagation angles-to-vertical to actual dip-angle gathers

The construction of PP angle gathers from the saved angles to vertical during wavefield propagation may be technically difficult, but the physical principle is simple. For PP reflections, the incidence angle is equal to the reflection angle. Therefore, a simple summation of the source wavefield's angle to the vertical at an imaging point at a particular propagation time, with the receiver wavefield's angle to the vertical, is enough to determine opening angle. The result of the cross-correlation between the wavefields at that time can be placed into an angle gather at the resulting opening angle. Though we save only the angle to the vertical for each wavefield, this is simply a reference axis. Once the opening angle is determined, adding half of that opening angle to either the source or receiver wave's angle to the vertical will produce an angle which is perpendicular to dip angle at that (specular) reflection point.
For PS (or SP) angle gathers, the relation between the P incidence angle and the S reflection angle is more complicated, as shown in Rosales and Rickett (2001).

PROPAGATION AND DECOMPOSITION RESULTS

Basic elastic propagation test

As a test case for the amplitude separation and displacement decomposition operators, I used a homogeneous medium, with parameters as follows:

1. P-wave velocity: $V_p = 2500\frac{\mathrm{m}}{\mathrm{s}}$
2. S-wave velocity: $V_s = 1500\frac{\mathrm{m}}{\mathrm{s}}$
3. Density: $\rho = 2\frac{\mathrm{gr}}{\mathrm{cm}^3}$
4. Spatial sampling: $\Delta x = \Delta z = 5\mathrm{m}$
5. Temporal sampling: $\Delta t = 0.5\mathrm{ms}$
6. Model dimensions: $n_x = n_z = 200$

The source function was a minimum phase Ricker wavelet, with a central frequency of 10_{hz}, which was applied at the center of the model. I padded the model with absorbing boundaries - an additional 16 elements in each direction. The wavefield propagation snapshots shown below are all at propagation time $T = 250$ms, and are shown without the surrounding absorbing boundaries.

Figure 3 is a snapshot of the propagation of the 2D elastic wavefield, at time $T = 250$ms. In this case, the source is a pressure source, meaning that theoretically, only P-waves should be generated. The top figures show the displacements in the X direction (U_x), and in the Z direction (U_z). The bottom figures are the result of applying the Helmholtz separation to the wavefield (equation 6). The P-wave is shown on the bottom left, and has the typical appearance of an acoustic wave in a homogeneous acoustic medium. The S-wave is on the bottom right, and ideally we would see nothing there. However, the artifacts seen at the edges are the result of minute reflections off the absorbing boundaries. Since the propagation is elastic, a wave mode conversion occurs at these boundaries as a result of a very slight impedance contrast artificially produced by the absorbtion, and shear wave energy appears. This is actually a validation of the correct behaviour of the propagation code. However, as part of future work, it may be worth trying to do impedance matching along the boundaries to improve their effectiveness.

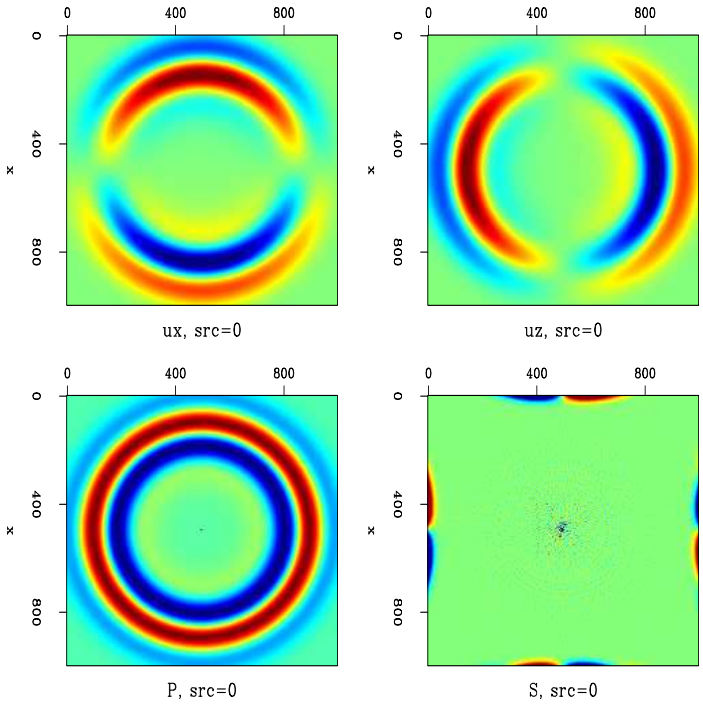


Figure 3: 2D elastic propagation snapshot at $T = 250$ms for a pressure source. X is the vertical direction, Z is horizontal. Top left: displacements in X direction. Top right: displacements in Z direction. Bottom left: P-wave amplitude, acquired by Helmholtz separation. Bottom right: S-wave amplitude. [**ER**] ohad1/. 0amp-and-disp

Figure 4 is similar to Figure 3, but in this case the source is a vertical dipole source, which excites both pressure and shear wave modes. These modes are apparent on the bottom row of the figure, which shows the result of applying Helmholtz separation to this wavefield. The pressure wave is radiating in the vertical direction, since the source was vertical. The shear wave is propagating in the horizontal direction, with the particle motion itself being vertical. The displacement snapshots on the top row of the figure show the U_x and U_z displacements, respectively. These displacements are a combination of P and S displacements in each direction. The vertical displacements (U_x) are more dominant than the horizontal ones (U_z), because the source excites motion in the vertical direction.

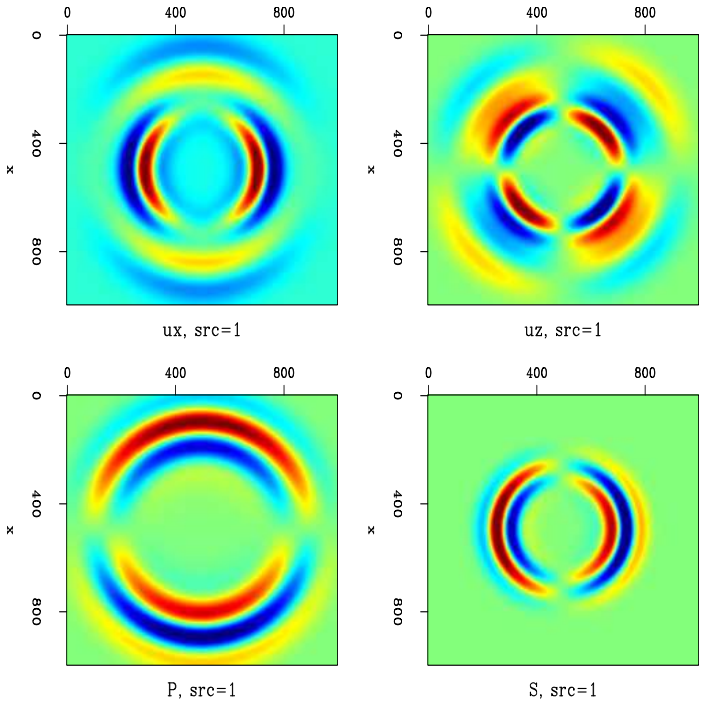


Figure 4: 2D elastic propagation snapshot at $T = 250$ms for a vertical dipole source. X is the vertical direction, Z is horizontal. Top left: displacements in X direction. Top right: displacements in Z direction. Bottom left: P-wave amplitude, acquired by Helmholtz separation. Bottom right: S-wave amplitude. [**ER**] ohad1/. 1amp-and-disp

Polarization coefficient and displacement decomposition in the wavenumber and space domain

The bottom right of Figure 5 is the polarization coefficient (equation 25) of the pressure-source wavefield (Figure 3), calculated at time $T = 200$ms. The time window size in which the polarization coefficient was calculated was 90ms, with its center at $T = 200$ms. The polarization coefficient is clipped below $r = 0.8$, which means that the field shows only model locations with very high polarization (clipped values are blue). As we may expect, this being a solitary pressure wave in a homogeneous medium, the field is indeed very linearly polarized.

The other panels in Figure 5 show crossplots (hodograms) of the vertical and horizontal displacement, in various locations in the wavefield within the time window centered at $T = 200$ms. For a linearly polarized wave, we expect this crossplot to appear as a thin line (the effect of linear dependence between displacement components). Otherwise, we should see an irregular shape in the crossplot. The panel on the bottom left of Figure 5 is the displacement crossplot at the center-left of the wavefield. The upper left panel is the displacement crossplot near the upper left corner of the wavefield, and the upper right panel is displacement at the center-top of the wavefield. Since there is only a single propagating P-wave present in the wavefield, all these locations exhibit linear dependence between the displacement components, indicating a high degree of polarization. The fundamental characteristic of P-waves, that the particle motion is tangent to the wave propagation direction, can be seen from this figure. All the crossplots seem to "radiate" away from the source, since there is only one source at the center. This is a demonstration of how polarization of displacements indicates the wavefield's directionality.

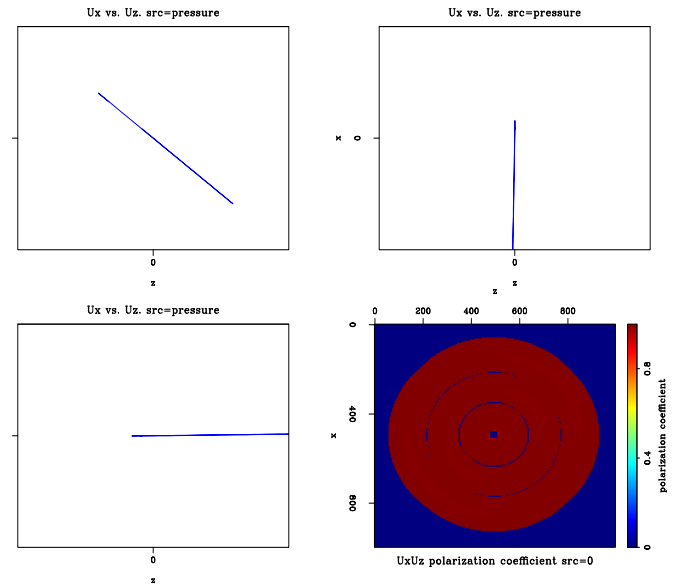


Figure 5: Bottom right: Polarization coefficient for a pressure source clipped below $r = 0.8$, and U_x vs. U_z displacement crossplots at certain locations in the wavefield at $T = 208$ms. Bottom left: crossplot at $x = 500, z = 330$. Top left: crossplot at $x = 330, z = 330$. Top right: crossplot at $x = 330, z = 500$. [**ER**] ohad1/. 0Uxzpol

Figure 6 displays the same panels as Figure 5, but in this case the source is a dipole source in the vertical direction, meaning that the displacement in the vertical axis (U_x) at the center of the wavefield is being forced. This results in the P and S waves shown in Figure 4. For this source, the polarization coefficient exhibits large regions of non-polarized waves, as seen by the polarization coefficient map on the bottom right of Figure 6. Examining the surrounding crossplot panels, a different picture of polarization appears: where in Figure 5 the polarization on the left of the model is horizontal, in Figure 6 it is vertical. The reason for this is that when a dipole source is excited, the wavefield contains a mixture of P and S waves, and thus a mixture of P and S displacements, as seen in Figure 4. Therefore, the linear polarization crossplot on the bottom left panel of Figure 6 is a result of shear wave displacements, while the linear polarization crossplot on the top right is a result of pressure wave displacements. The shear wave and its associated vertical particle motion propagates in the horizontal direction, while the pressure wave and its associated vertical particle motion propagates vertically. Consequently, the panel on the top left of Figure 6, which displays the crossplot at the upper left corner of the wavefield, is a mixture of shear and pressure displacements, and therefore exhibits particle motion which is not linearly polarized. This figure shows the need for a displacement decomposition method.

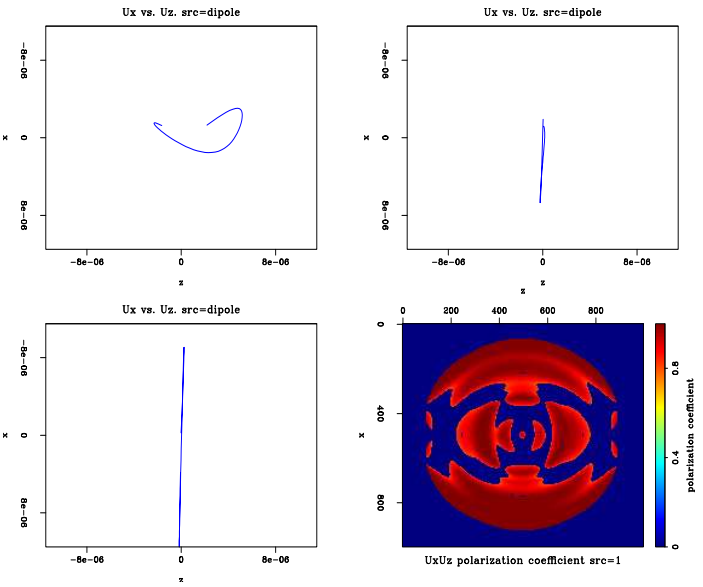


Figure 6: Bottom right: Polarization coefficient for a vertical dipole source clipped below $r = 0.8$, and U_x vs. U_z displacement crossplots at certain locations in the wavefield at $T = 208$ms. Bottom left: crossplot at $x = 500, z = 330$. Top left: crossplot at $x = 330, z = 330$. Top right: crossplot at $x = 330, z = 500$. [**ER**] ohad1/. 1Uxzpol

Figure 7 is an example of the implementation of P and S wave displacement decomposition in the wavenumber domain (equations 12 and 14), when a pressure source is applied. Shear waves are not expected to form in this case since the medium is homogeneous. The top row shows the vertical displacements, and the bottom row shows the horizontal displacements. From left to right, the panels are the pressure wave displacements after decomposition (U^P), the shear wave displacements after decomposition (U^S), the sum of the shear and pressure displacements ($U^P + U^S$), and the original displacements before decomposition. Figure 9 is the same, but in this case the source is a vertical dipole source. The expectation is that the sum of displacements after decomposition will be equal to the displacements before decomposition, and it is does appear to be so.

Figure 8 is the implementation of P and S wave displacement decomposition in the space domain (equations 24), when a pressure source is applied. The results are similar to those from the wavenumber domain decomposition, shown in Figure 7. Figure 10 shows the space domain decomposition for a vertical dipole source. In this case, the decomposed displacements are different from those seen by the wavenumber decomposition, shown in Figure 9. Notice the vertical streaks in the top left panels of Figure 9. I interpret these streaks as artifacts of the Fourier transform, since when I use more zero padding before applying the 2D Fourier transform, their magnitude decreases. However, it is interesting that the resulting summations in both cases appear similar, and likewise appear almost identical to the pre-decomposed displacements. This shows that the decomposition method in the wavenumber domain by Zhang and McMechan (2010), and the decomposition in the space domain both work in principle for an isotropic medium.

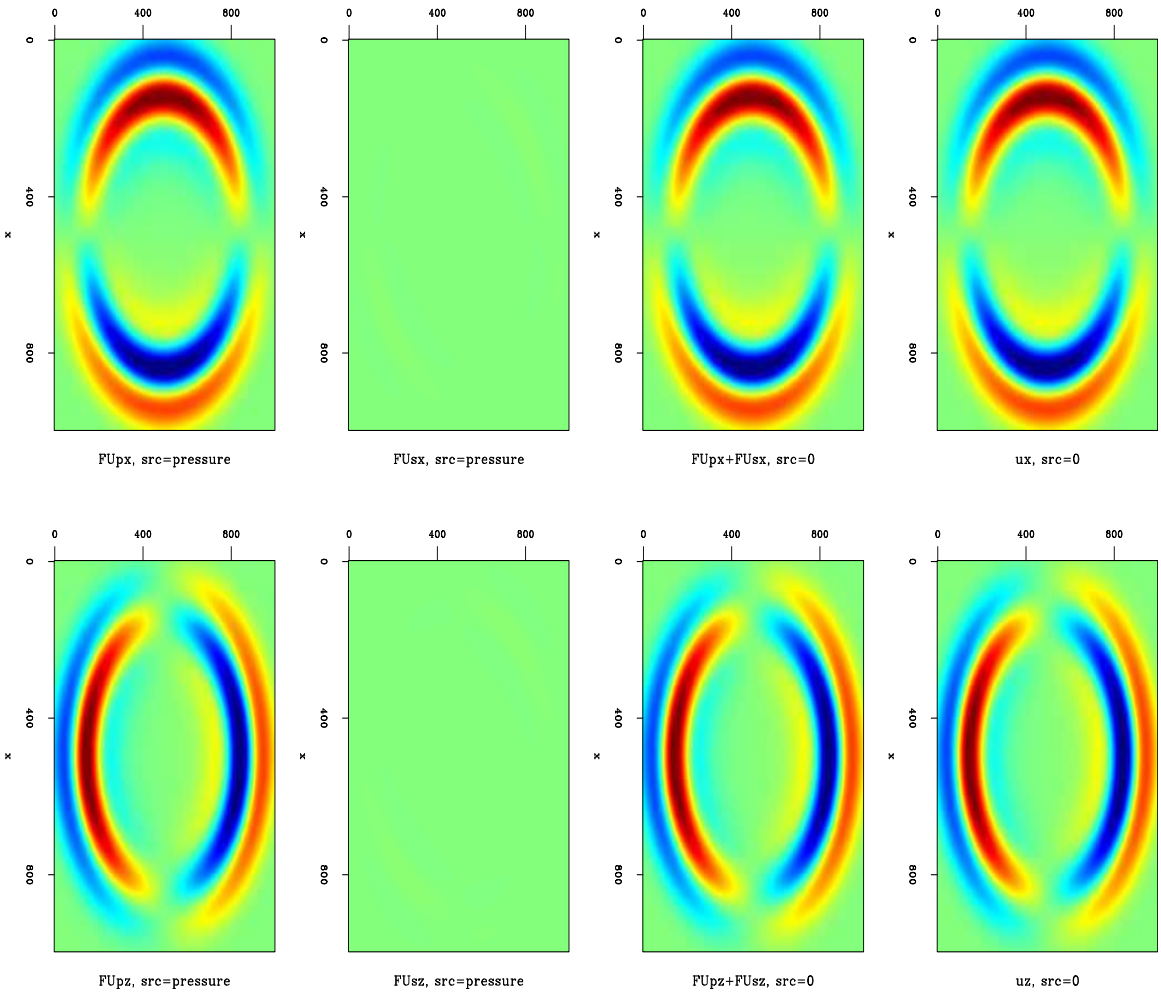


Figure 7: Decomposed displacements snapshot at $T = 250ms$ for a pressure source. Displacement decomposition is in wavenumber domain. X is the vertical direction, Z is horizontal. Top row: displacements in X direction. Bottom row: displacements in Z direction. From left to right: P displacements, S displacements, P+S displacements and original displacements. [**ER**] ohad1/. 0Fseparated-combined

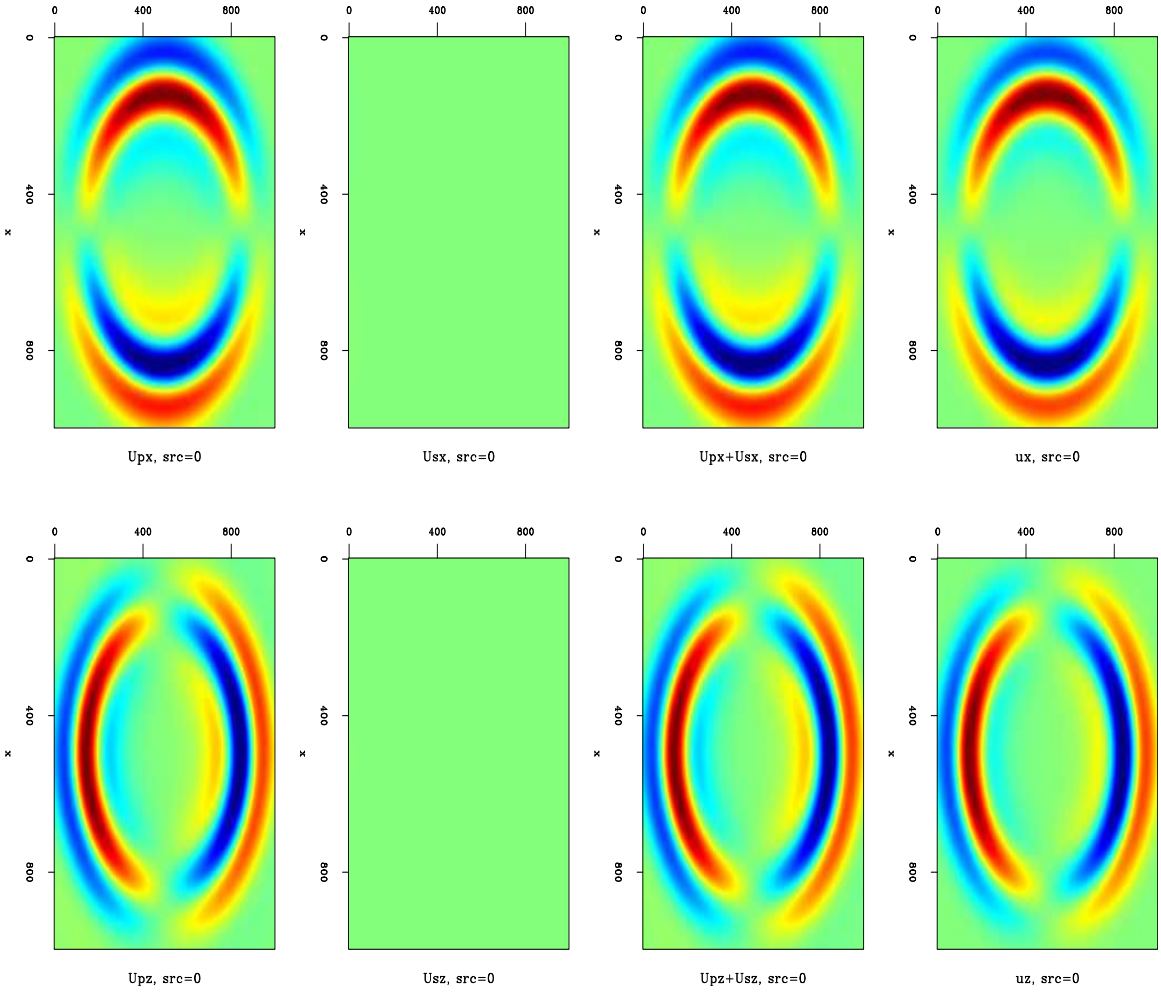


Figure 8: Decomposed displacements snapshot at $T = 250ms$ for a pressure source. Displacement decomposition is in space domain. X is the vertical direction, Z is horizontal. Top row: displacements in X direction. Bottom row: displacements in Z direction. From left to right: P displacements, S displacements, P+S displacements and original displacements. [**ER**] ohad1/. 0separated-combined

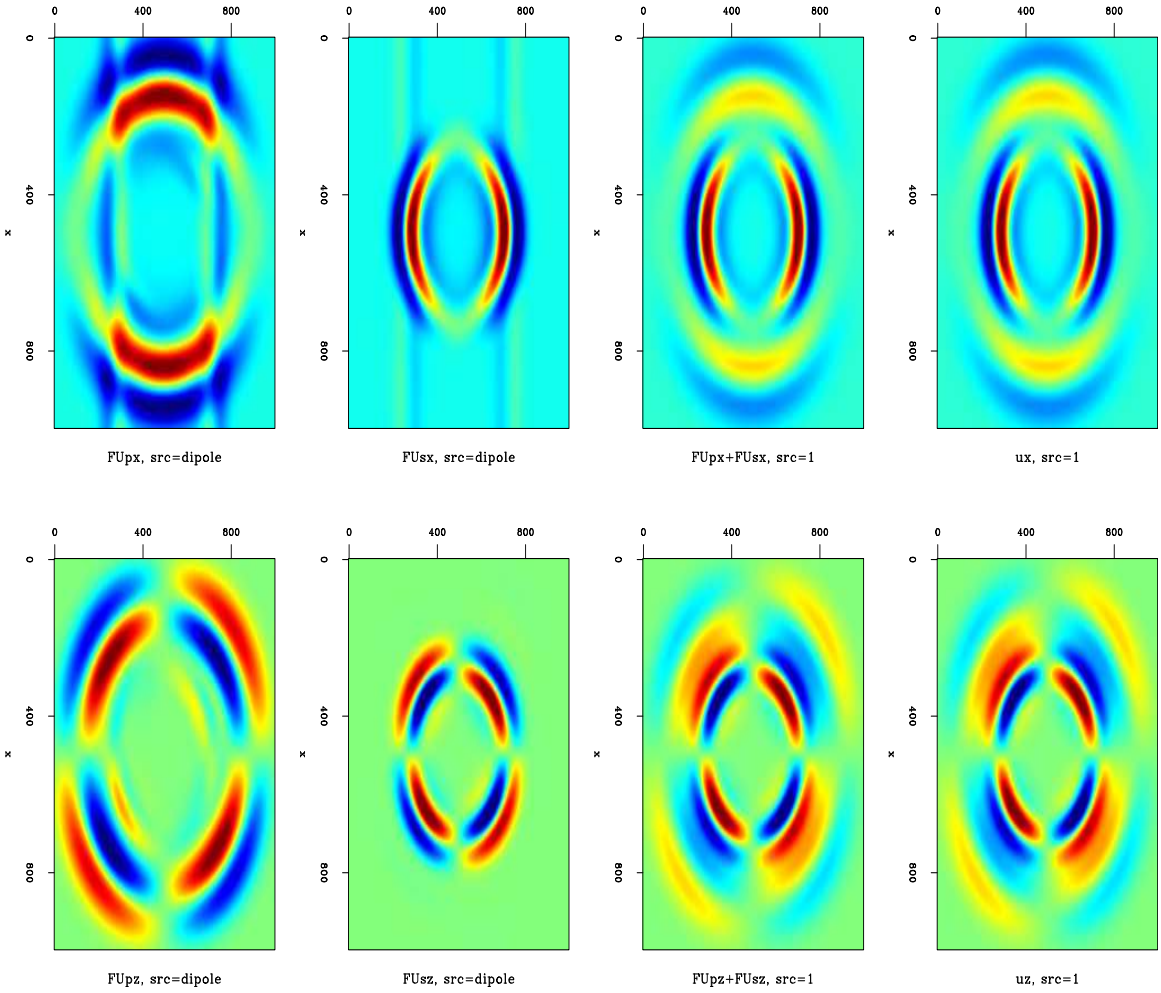


Figure 9: Decomposed displacements snapshot at $T = 250ms$ for a vertical dipole source. Displacement decomposition is in wavenumber domain. X is the vertical direction, Z is horizontal. Top row: displacements in X direction. Bottom row: displacements in Z direction. From left to right: P displacements, S displacements, P+S displacements and original displacements. [**ER**] ohad1/. 1Fseparated-combined

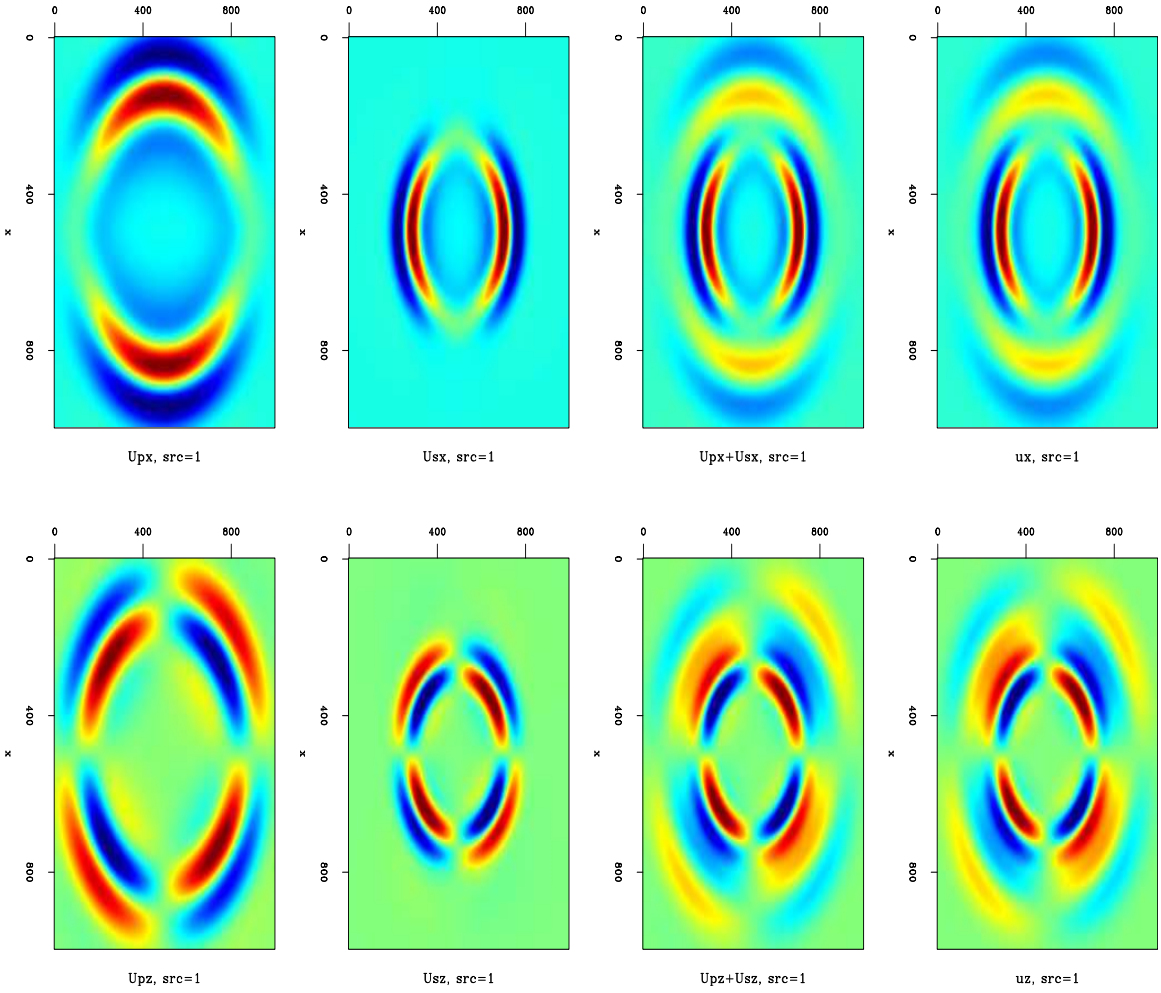


Figure 10: Decomposed displacements snapshot at $T = 250ms$ for a vertical dipole source. Displacement decomposition is in space domain. X is the vertical direction, Z is horizontal. Top row: displacements in X direction. Bottom row: displacements in Z direction. From left to right: P displacements, S displacements, P+S displacements and original displacements. [**ER**] ohad1/. 1separated-combined

However, a large difference between the wavenumber and space domain methods appears when comparing the errors of the sum of the decomposed displacement fields vs. the pre-decomposed displacement fields. The comparison is done by calculating the absolute value of the difference between the summed-decomposed displacements and the undecomposed displacements for each model point at each propagation time, and then summing all that into a single number:

$$E_i = \sum_{t=1}^{T} \sum_{n=1}^{N} \left| U_i^{n,t} - (U_{p_i}^{n,t} + U_{s_i}^{n,t}) \right|, \tag{30}$$

where $U_i^{n,t}$ is the displacement in direction i before decomposition, at model point n at propagation time t, $U_{p_i}^{n,t}$ is the decomposed P wave displacement, and $U_{s_i}^{n,t}$ is the decomposed S wave displacement. Figures 11 and 12 show a comparison of the absolute difference field at $T = 200ms$ for the wavenumber and space domain decompositions. The figures are misleading, since the maximum value of the wavenumber domain difference is 6 orders of magnitude less than the space domain decomposition difference. The expectation is that the sum of displacements after decomposition will be equal to the displacements before decomposition. It appears that this is the case for decomposition in the wavenumber domain, at least from the relative error point of view.

The following table shows the summations of the absolute displacement fields and the absolute decomposition errors, over the entire propagation time, for a pressure source and a dipole source:

Source	$\lvert U_x \rvert$	E_x space	Rel. err	E_x wavenumber	Rel. err
Pressure	3.54e-10	5.51e-11	0.15	6.06e-17	0.17e-6
Dipole	5.19	0.88	0.17	1.68e-6	0.32e-6
Source	$\lvert U_z \rvert$	E_z space	Rel. err	E_z wavenumer	Rel. err
Pressure	3.54e-10	7.56e-11	0.21	5.95e-17	0.16e-6
Dipole	3.32	0.36	0.10	1.03e-6	0.31e-6

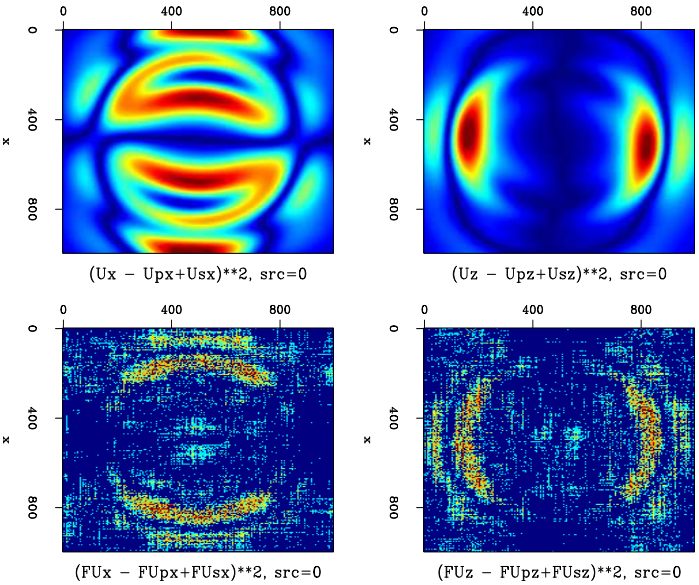


Figure 11: Absolute errors of decomposed displacements in comparison to original displacements for a pressure source. $T = 250$ms. Top: displacement decomposition in space domain. Bottom: displacement decomposition in the wavenumber domain. Wavenumber domain errors are 6 orders of magnitude smaller than space domain. [**ER**] ohad1/. 0space-vs-wavenumber-diff

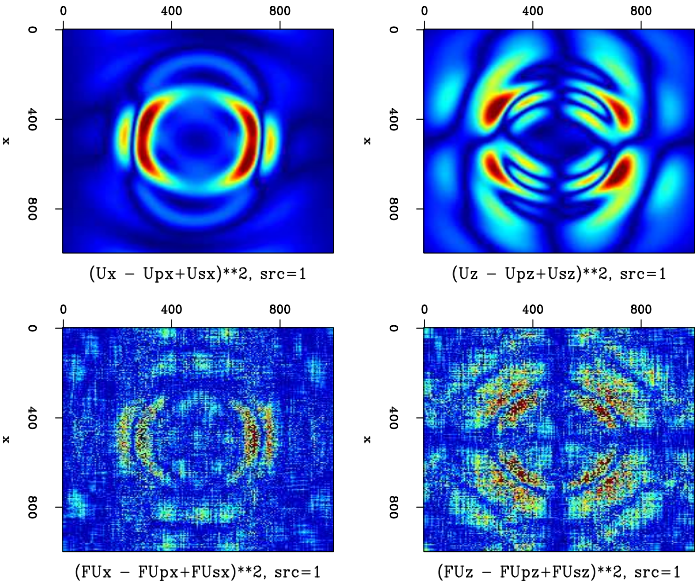


Figure 12: Absolute errors of decomposed displacements in comparison to original displacements for a vertical dipole source. $T = 250$ms. Top: displacement decomposition in space domain. Bottom: displacement decomposition in the wavenumber domain. Wavenumber domain errors are 6 orders of magnitude smaller than space domain. [**ER**] ohad1/. 1space-vs-wavenumber-diff

The correctness of the displacement decomposition operators can be verified when comparing Figure 5 to Figures 13 and 14. Since there is only a pressure wave in the wavefield, we expect the decomposition operator to not have an effect on the resulting pressure wave polarization. Looking at these two figures, we can see they are similar, indicating that the decomposition operators do not introduce a significant amount of numerical inaccuracies. However, it is clear that decomposition in the wavenumber domain has introduced less errors into the P wave displacements.

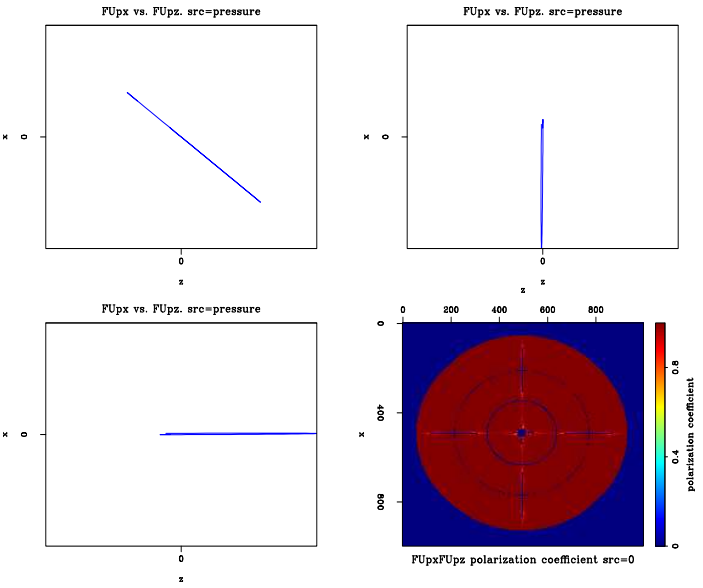


Figure 13: Bottom right: Polarization coefficient for a pressure source clipped below $r = 0.8$, and U_x^P vs. U_z^P displacement crossplots at certain locations in the wavefield at $T = 208$ms. Decomposition is in the wavenumber domain. Bottom left: crossplot at $x = 500, z = 330$. Top left: crossplot at $x = 330, z = 330$. Top right: crossplot at $x = 330, z = 500$. [**ER**] ohad1/. 0FUpxzpol

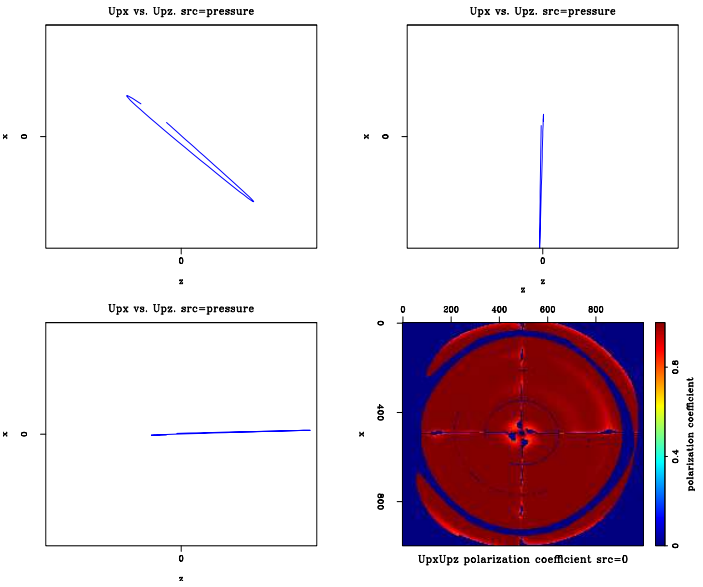


Figure 14: Bottom right: Polarization coefficient for a pressure source clipped below $r = 0.8$, and U_x^P vs. U_z^P displacement crossplots at certain locations in the wavefield at $T = 208$ms. Decomposition is in the space domain. Bottom left: crossplot at $x = 500, z = 330$. Top left: crossplot at $x = 330, z = 330$. Top right: crossplot at $x = 330, z = 500$. [**ER**] ohad1/. 0Upxzpol

The effect of decomposition on polarization can be seen when comparing Figure 6 to Figures 15 - 18. Notice in particular the top left panel in these figures, indicating the crossplot of vertical and horizontal displacements on the upper left corner of the wavefield. Where in Figure 6 this panel displayed non-linear polarization as a result of a mixture of P and S displacements, in Figures 15 and 16 the polarization appears more linear, similar to the same panel in Figure 5 where only a P wave existed. In Figures 17 and 18, the same panel likewise shows a more linear polarization, but in this case the polarization is flipped by 90 degrees compared to that in Figure 15. This is exactly what we would expect to see - for a shear wave, the particle motion is perpendicular to the wave's propagation direction. Looking at the bottom left panel of Figures 15 and 16, we can see that the U_x^p displacement amplitude is about 20% of that in the top right panel, since at that location the P displacement should be quite small (refer to Figure 4). Similarly, looking at the top right panel of Figures 17 and 18 we can see that the U_x^s displacement amplitude is significantly lower than in the bottom left panels of these figures. Again - refer to Figure 4 to understand why that should be so.

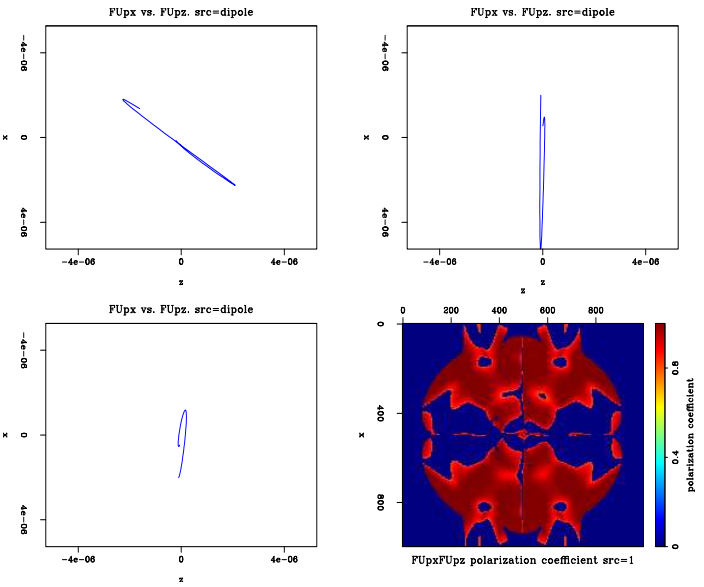


Figure 15: Bottom right: Polarization coefficient for a vertical dipole source clipped below $r = 0.8$, and U_x^P vs. U_z^P displacement crossplots at certain locations in the wavefield at $T = 208$ms. Decomposition is in the wavenumber domain. Bottom left: crossplot at $x = 500, z = 330$. Top left: crossplot at $x = 330, z = 330$. Top right: crossplot at $x = 330, z = 500$. [**ER**] ohad1/. 1FUpxzpol

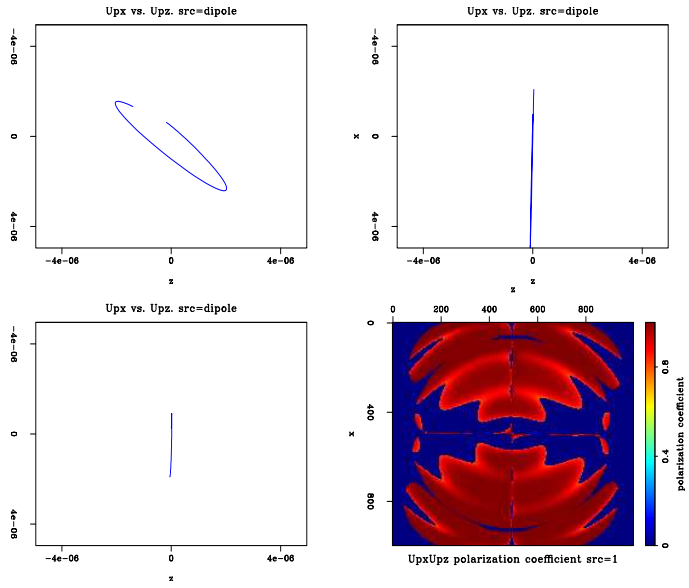


Figure 16: Bottom right: Polarization coefficient for a vertical dipole source clipped below $r = 0.8$, and U_x^P vs. U_z^P displacement crossplots at certain locations in the wavefield at $T = 208$ms. Decomposition is in the space domain. Bottom left: crossplot at $x = 500, z = 330$. Top left: crossplot at $x = 330, z = 330$. Top right: crossplot at $x = 330, z = 500$. [**ER**] ohad1/. 1Upxzpol

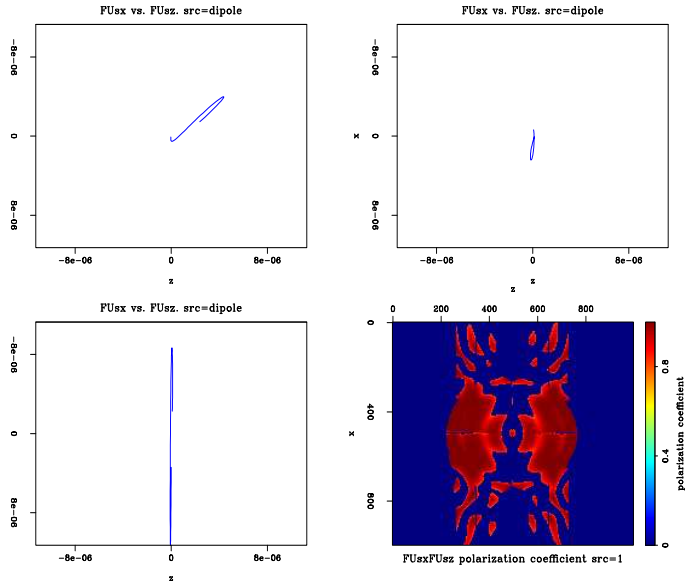


Figure 17: Bottom right: Polarization coefficient for a vertical dipole source clipped below $r = 0.8$, and U_x^S vs. U_z^S displacement crossplots at certain locations in the wavefield at $T = 208$ms. Decomposition is in the wavenumber domain. Bottom left: crossplot at $x = 500, z = 330$. Top left: crossplot at $x = 330, z = 330$. Top right: crossplot at $x = 330, z = 500$. [**ER**] ohad1/. 1FUsxzpol

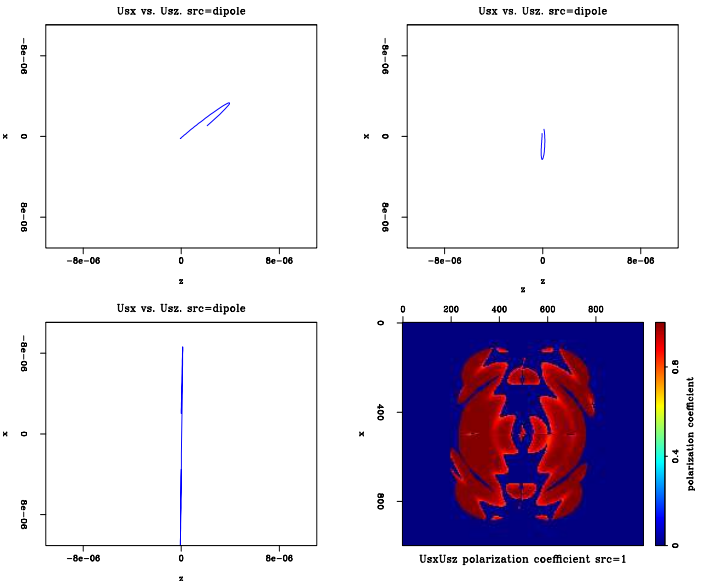


Figure 18: Bottom right: Polarization coefficient for a vertical dipole source clipped below $r = 0.8$, and U_x^S vs. U_z^S displacement crossplots at certain locations in the wavefield at $T = 208$ms. Decomposition is in the space domain. Bottom left: crossplot at $x = 500, z = 330$. Top left: crossplot at $x = 330, z = 330$. Top right: crossplot at $x = 330, z = 500$. [**ER**] ohad1/. 1Usxzpol

Observing Figures 15 - 18, we can see that when there is more than one wave mode in the field, decomposition in the space domain produces a smoother polarization correlation field than the wavenumber method. This is in accordance with the errors visible in Figures 11 and 12 for the wavenumber domain decomposition errors. It seems that though the error is smaller when decomposing in the wavenumber domain, its spread around the true result is very "ringy." The error introduced by the space domain deconvolution method may be larger, but it is much smoother. Because what I need is a measure of polarization correlation and subsequently its direction, I think that it will be more advantageous to have a smooth error.

Determination of polarization angle to the vertical

The following figures are the result of applying linear regression to the displacement fields and then calculating the polarization's angle to the vertical (equations 28 and 29), within the same time windows at which the polarization coefficient is calculated. If the polarization coefficient is below $r = 0.8$, then the angle is set to a clipped value (greater than 90^o; clipped angles are red), so that the polarization coefficient figures and the angle figures display similar clipped regions. This is a visualization of the eventual methodolgy I intend to implement: when polarization cannot be estimated, I do not wish to bring the wavefield data from that location into an angle gather.

Figure 19 is a snapshot at propagation time $T = 208$ms of the undecomposed displacements (top), the polarization coefficient field (bottom-left) and the polarization angle-to-vertical field (bottom-right). There is a single pressure source at the center. The color in the angle field ranges from white to black, where black designates vertical polarization ($\theta = 0^o$), and white designates horizontal polarization $\theta = 90^o$). In this figure, because there is only a single P-wave, the field is highly linearly polarized. Also, the polarization direction is the propagation direction. The angle field therefore is a good indication of the propagation direction, and this can be verified when observing the displacement fields. However, notice the "waviness" along the diagonals radiating from the center of the field, where the source is located. If the polarization estimation were perfect, we would expect all lines radiating from the source to have the same color, meaning the same polarization angle.

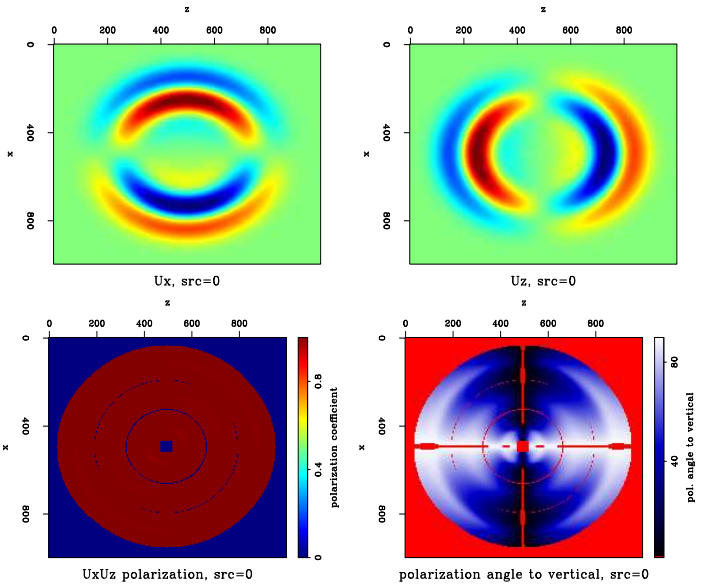


Figure 19: Displacement snapshot at $T = 208$ms for a pressure source. Top: displacements in X and Z directions. Bottom left: polarization coefficient clipped below $r = 0.8$. Bottom right: angle of polarization direction to the vertical (Dark blue= 0^o, white= 90^o). [**ER**] ohad1/. 0U-disp-and-ang

Figures 20 and 21 are the result of wavenumber domain decomposition and space domain decomposition, respectively. As the decomposition should, theoretically, have no effect on this single mode wavefield, we expect these figures to be identical to Figure 19. The polarization coefficient and the polarization angle panels in Figure 20 are very similar to their undecomposed counterparts, but in Figure 21 a circular artifact is evident. I attribute this artifact to the helical deconvolution by which the decomposition was done (equation 24). As a result of the length of the filter, data is "spread" from one part of the model to another, in this case from the surrounding absorbing boundary region to the internal part of the wavefield. As the filter length changes, the amplitude and location of this artifact

changes. If the filter is too long, data from the bottom edge of the wavefield could actually be incorporated into the top edge, due to the wrap-around of the helical coordinate system.

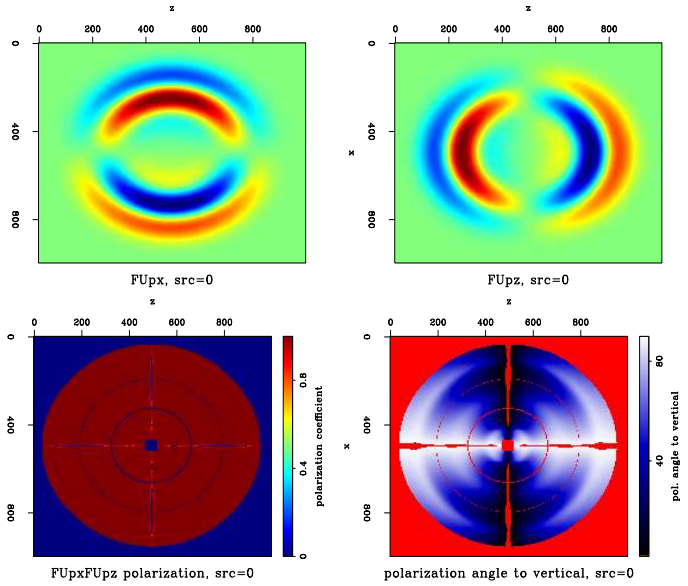


Figure 20: P wave displacement snapshot at $T = 208$ms for a pressure source, decomposition in wavenumber domain. Top: displacements in X and Z directions. Bottom left: polarization coefficient clipped below $r = 0.8$. Bottom right: angle of polarization direction to the vertical (Dark blue= 0^o, white= 90^o). [**ER**] ohad1/. 0FUp-disp-and-ang

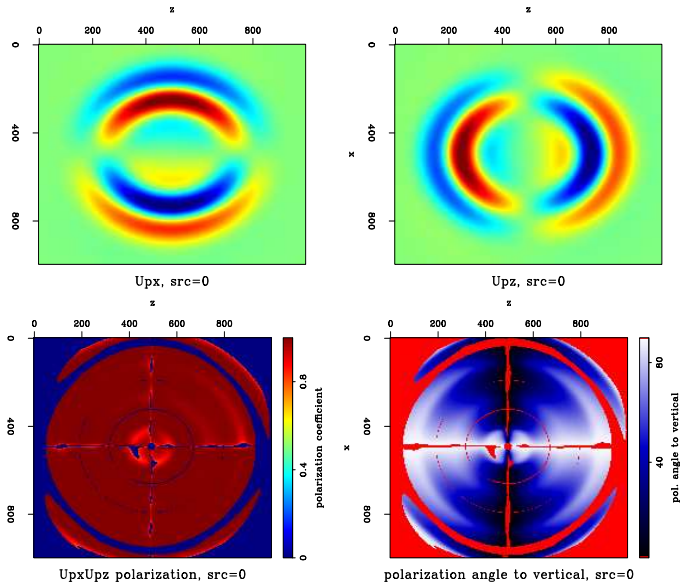


Figure 21: P wave displacement snapshot at $T = 208$ms for a pressure source, decomposition in space domain. Top: displacements in X and Z directions. Bottom left: polarization coefficient clipped below $r = 0.8$. Bottom right: angle of polarization direction to the vertical (Dark blue= 0^o, white= 90^o). [**ER**] ohad1/. 0Up-disp-and-ang

Figure 22 is the undecomposed displacements snapshot at propagation time $T = 208$ms, for a vertical dipole source at the center of the field. This excites both P and S wave modes, and therefore the displacements are a mixture of P and S. This is the reason for the smaller region of linear polarization seen in the bottom-left of the figure, in comparison to Figure 19. The bottom-right of the figure shows the polarization's angle to vertical. A dark region, indicating vertical polarization, is visible along both directions away from the source. The vertical polarization to the right and left of the source are a result of vertical S displacements, and those that are over and under the source are a result of vertical P displacements.

Figures 23 and 24 are the decomposed P displacements of the dipole-source generated field, by wavenumber and by space domain decomposition, respectively. They both exhibit

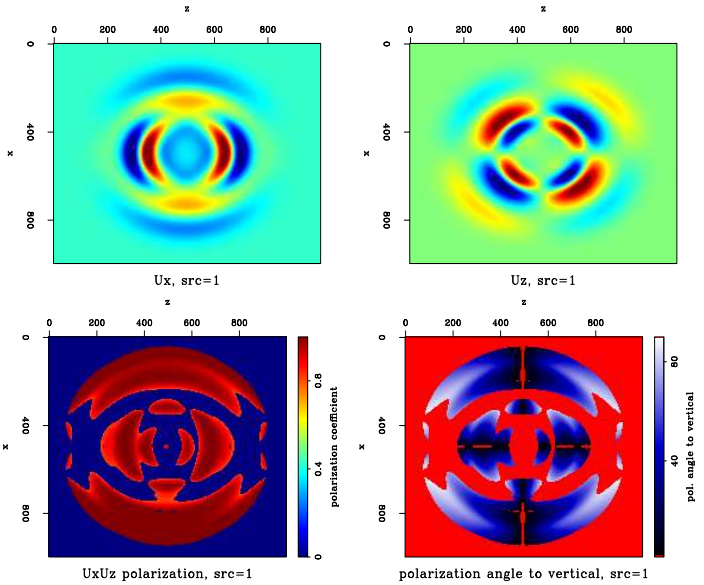


Figure 22: Displacement snapshot at $T = 208$ms for a vertical dipole source. Top: displacements in X and Z directions. Bottom left: polarization coefficient clipped below $r = 0.8$. Bottom right: angle of polarization direction to the vertical (Dark blue= 0^o, white= 90^o). [**ER**] ohad1/. 1U-disp-and-ang

the generally correct behaviour for the P-wave - it propagates mostly vertically, and the polarization direction should be likewise mostly vertical, radiating vertically from the source at the center (refer to Figure 4). However, the smoothness of the polarization coefficient field derived by space domain decomposition is much greater than the one derived by wavenumber domain decomposition. This in turn affects the smoothness of the polarization angle field.

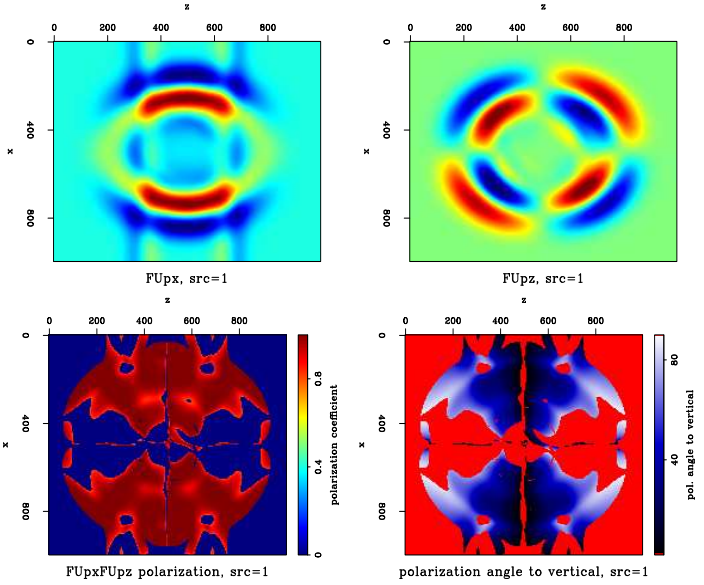


Figure 23: P wave displacement snapshot at $T = 208$ms for a vertical dipole source, decomposition in wavenumber domain. Top: displacements in X and Z directions. Bottom left: polarization coefficient clipped below $r = 0.8$. Bottom right: angle of polarization direction to the vertical (Dark blue= 0^o, white= 90^o). [**ER**] ohad1/. 1FUp-disp-and-ang

Figures 25 and 26 are the decomposed S displacements of the dipole-source generated field, by wavenumber and by space domain decomposition, respectively. In this case, we expect there to be a mostly vertically polarized wave, propagating to the sides away from the source at the center. This effect can be seen in both figures, on the bottom-right panel. Just as was shown above for the P-wave, the relative smoothness of the space domain decomposition compared to the wavenumber domain decomposition is evident.

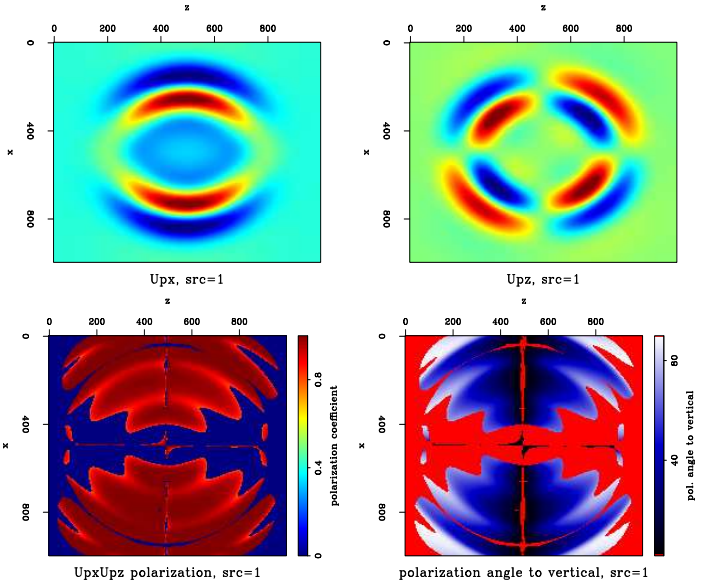


Figure 24: P wave displacement snapshot at $T = 208$ms for a vertical dipole source, decomposition in space domain. Top: displacements in X and Z directions. Bottom left: polarization coefficient. Bottom right: angle of polarization direction to the vertical (Dark blue= 0^o, white= 90^o). [**ER**] ohad1/. 1Up-disp-and-ang

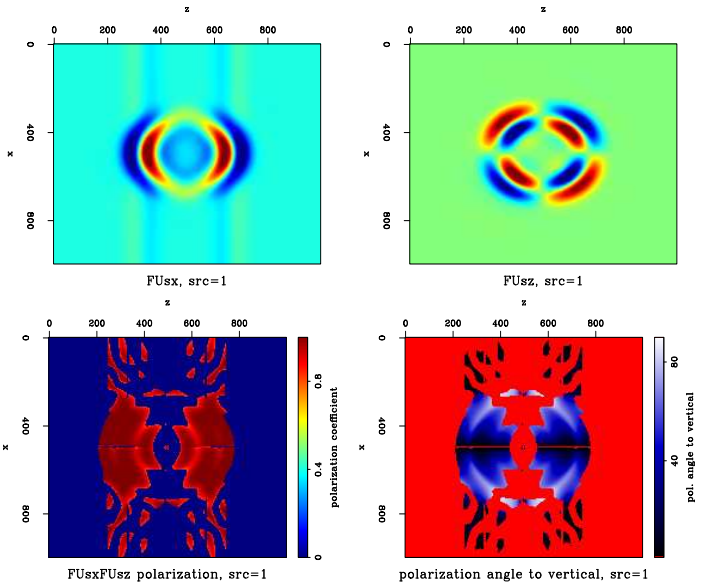


Figure 25: S wave displacement snapshot at $T = 208$ms for a vertical dipole source, decomposition in wavenumber domain. Top: displacements in X and Z directions. Bottom left: polarization coefficient clipped below $r = 0.8$. Bottom right: angle of polarization direction to the vertical (Dark blue= 0^o, white= 90^o). [**ER**] ohad1/. 1FUs-disp-and-ang

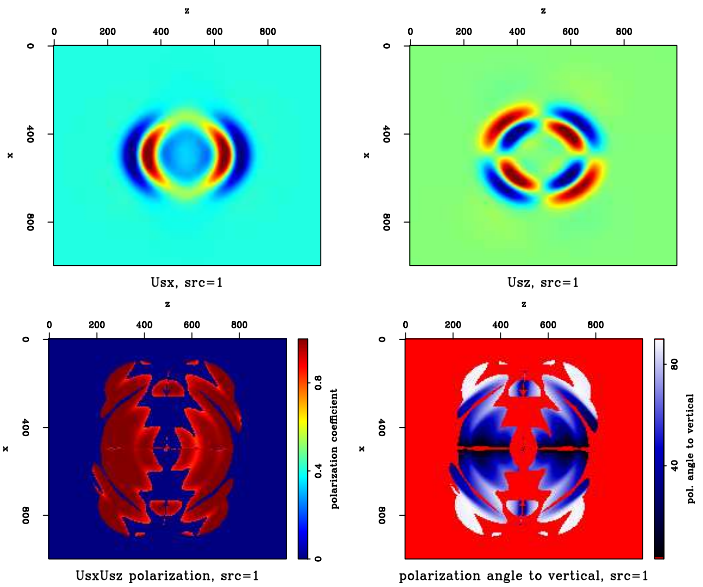


Figure 26: S wave displacement snapshot at $T = 208$ms for a vertical dipole source, decomposition in space domain. Top: displacements in X and Z directions. Bottom left: polarization coefficient clipped below $r = 0.8$. Bottom right: angle of polarization direction to the vertical (Dark blue= 0^o, white= 90^o). [**ER**] ohad1/. 1Us-disp-and-ang

DISCUSSION AND CONCLUSION

I have verified that the displacement decomposition operator (Zhang and McMechan, 2010) works for an isotropic elastic wavefield in a homogeneous medium. Currently, I have implemented it in the wavenumber domain, as stated in equations 12 and 14, and in the space domain, as stated in equation 24. Since the wave propagation itself is in the time-space domain, executing transforms to wavenumber domain at each time step may be more costly in the eventual 3D implementation than using the space domain decomposition by filter deconvolution. For the moment, it appears that both methods produce similar results as far as understanding of the polarization direction is concerned. This is despite the fact that the error introduced by the wavenumber domain displacement decomposition is six orders of magnitude smaller than the one introduced by space domain displacement decomposition. It appears that the space domain error is large but smooth, while the wavenumber domain error is small, but highly variable.

The size of the sliding time window within which the polarization coefficient is calculated is a critical decision to the derivation of polarization direction, and therefore to the calculation of the wave propagation direction. The ability to calculate the propagation direction during propagation is the cornerstone of the method I wish to eventually implement for constructing angle gathers. Therefore, a robust formulation of the time window size as a function of the wavefield parameters is required.

The cut-off value for the correlation between the displacements on the two axes, which I use as a determination of the level of local polarization, is another critical factor. Obviously, where the displacement values are very small or where they are not correlated there is no point in trying to estimate polarization. At those locations in the wavefield, either there is no significant energy anyway, or there is more than one wave at the same time. It should be noted that this same limitation exists for current angle gather construction methodology, but it is not addressed until the result is viewed. The use of statistical correlation (equation 25) does not seem to be the most robust way of measuring the linearity of polarization in the propagating wavefield. There are other methods, such as executing SVD on a matrix

consisting of the displacement components of an image point over within a time window (Meersman et al., 2006), which may provide a more stable polarization estimation.

I calculate the slope of the line that best fits, in the least-square sense, the crossplotted displacements U_x and U_z. This slope is then used to determine the angle between the polarization direction and the vertical. According to the results (Figures 19 - 25), this method is imperfect. I would expect that along the line radiating from the source in Figure 19, the polarization angle would be constant. I have a nagging suspicion that the cause of this is the staggered grid methodology, which in effect means that the U_x and U_z displacements are not calculated at the same model location. I will need to find a way to test this possibility. If this is a critical issue, some correction to the displacements will be necessary so they'll be calculated at the same grid points. Zhang and McMechan (2010) mention that they interpolate the values from one staggered grid to another, to ensure that displacement information for different axes is located at the same grid points.

As for the use of this directionality information - I believe knowing the wavefield's direction at each time during propagation incorporates the advantages of wavefield propagation methods and ray-based methods. It may thus enable a method of angle gather creation in the style shown in (Koren et al., 2008), but without the restrictions imposed by ray theory. Furthermore, knowing the wavefield's direction will make it unnecessary to "shoot" rays in many directions from various CDP's in the model space, and then create the angle gather using only the data from those rays that reached the surface near shot-receiver points, where data exists. This can result in a significant saving of computational and I/O resources.

Using the displacement information inherent in elastic wavefield propagation can also make redundant the angle gather creation method by spatial and temporal lag cross-correlation (Sava and Fomel, 2003). The extended imaging condition there requires cross-correlation at various spatial lags between the receiver and source wavefields, but the reason for doing this is exactly because there is no information as to the wavefield's propagation direction. Therefore, a "scan" of a CIG's local neighbourhood is executed, to cover all directions where the source and receiver wave energy could have arrived from. Just as for the ray-based method, a possible reduction in computational resources can be achieved by using the pre-existing wavefield directionality information in the displacement fields.

Another use for wavefield directionality information is for the cross-correlation imaging condition itself. If we have the source and receiver propagation direction, we can ensure that we cross-correlate only those events traveling in opposite directions, such as the up-going receiver wavefield with the downgoing source wavefield. By doing that, we remain within the assumptions of the cross-correlation imaging condition, and the resulting image should contain fewer artifacts. Many authors acknowledge the need for this separation, and useful methods exist to carry it out. What I am suggesting is that the same directionality information acquired by displacement decomposition, which I wish to use for angle gather construction, can also be used in conjunction with the imaging condition, in order to image the events which the imaging condition was designed for.

ACKNOWLEDGEMENTS

My most heartfelt gratitude to my two lovely officemates: Manchu Wong and Yunyue Li, who have had to listen to me ramble on and on about this subject for the last 4 months,

and never once complained. Their ideas and good nature have enabled me to do this work.

REFERENCES

Claerbout, J., 1997, Multidimensional recursive filters via a helix: SEP-Report, **95**, 1–13.

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

Fomel, S., P. Sava, J. Rickett, and J. F. Claerbout, 2003, The wilson-burg method of spectral factorization with application to helical filtering: Geophysical Prospecting, **51**, 409–420.

Koren, Z., I. Ravve, E. Ragoza, A. Bartana, and D. Kosloff, 2008, Full-azimuth angle domain imaging: SEG Las-Vegas, 78th Ann. Internat. Mtg.

Meersman, K. D., M. van der Baan, and J. Kendall, 2006, Signal extraction and automated polarization analysis of multicomponent array data: Bulletin of the Seismological Society of America, **96**, 2415–2430.

Morse, P. and H. Feshbach, 1953, Methods of theoretical physics: Mcgraw-Hill Book Company.

Rosales, D. and J. Rickett, 2001, Ps-wave polarity reversal in angle domain common image gathers: 71st Ann. Internat. Mtg., expanded abstracts, 1843–1846.

Sava, P. and S. Fomel, 2003, Angle domain common image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074.

Virieux, J., 1986, P-sv wave propagation in heterogeneous media: Velocity-stress finite difference method: Geophysics, **51**, 889–901.

Wu, C. and J. Harris, 2002, An optimized variable-grid finite-difference method for seismic forward modeling: SEG Workshop, 72nd Ann. Internat. Mtg.

Yan, J. and P. Sava, 2008, Isotropic angle-domain elastic reverse-time migration: Geophysics, **73**, S229–S239.

———, 2009, 3d elastic wave mode separation for tti media: SEG Houston, 79th Ann. Internat. Mtg.

Zhang, Q. and G. McMechan, 2010, 2d and 3d elastic wavefield vector decomposition in the wavenumber domain for vti media: Geophysics, **75**, D13–D26.

Ambient seismic noise eikonal tomography for near-surface imaging at Valhall

Sjoerd de Ridder and Joe Dellinger

ABSTRACT

We demonstrate that in three passive seismic datasets recorded by an ocean-bottom-cable array at the Valhall field in the Norwegian North Sea, virtual-source interferometry over the frequency range 0.35-1.75 Hz produces strong omnidirectional Scholte-wave sources. We then use these virtual-sources gathers to image the shallow structure in high resolution using eikonal traveltime tomography. Unlike conventional tomography, which determines velocities between source and receiver using an inversion scheme, eikonal tomography computes the local traveltime gradient at each receiver, and thus directly estimates the local velocity in the neighborhood of each receiver. The Scholte-wave images produced from imaging the passive noise reveal many of the same features visible in the active P-wave data. These results suggest that a permanent recording system using passive seismic noise might be useful for real-time surveillance of shallow shear-wave velocity anomalies.

INTRODUCTION

The permanent ocean-bottom array at the Valhall field in Norway provides an excellent source of passive seismic data to test what might be accomplished with seismic interferometry. The array, installed in 2003 (Kommedal et al., 2004), can record data for long periods in all weather conditions. The subsurface structure is well known, both from numerous wells and from seismic imaging. During periods without active seismic acquisition at Valhall, there is abundant passive energy in the data over a wide range of frequencies. Previous attempts to synthesize omnidirectional point sources at conventional seismic exploration frequencies (3-60 Hz) failed at Valhall (Artman, 2007; Landes et al., 2009), so passive data will not replace active data in the higher frequencies ranges there. However, at lower frequencies (0.18-1.75 Hz), omnidirectional virtual Scholte-wave sources have been successfully generated (Dellinger and Yu, 2009; Bussat and Kugler, 2009; Dellinger et al., 2010). We process three passive datasets from the Valhall permanent array in Norway, generate virtual Scholte-wave source gathers using seismic interferometry, and then use these sources to image the shallow structure using eikonal traveltime tomography (Lin et al., 2009). Finally, we compare our results to the P-wave velocity structure at Valhall as determined from active data (Sirgue et al., 2010).

THE AMBIENT SEISMIC NOISE FIELD AT VALHALL

We use three OBC recordings of ambient seismic noise recorded at Valhall: 29 hours in 2004, 6.5 hours in 2005 in conditions that transitioned from calm to stormy, and 2 hours in

2008 recorded under stormy conditions as the remnants of tropical storm Laura passed over. Unlike the previous passive datasets, the 2008 dataset was recorded without the customary low-cut recording filter. Figure 1 shows spectra of the vertical component of the particle velocity for all three datasets. The ambient seismic field is excited by various noise sources that dominate in different frequency regimes (Olofsson, 2010). In the Valhall recordings, we distinguish five frequency regimes based on their distinctive spectral amplitudes and excitation mechanisms (Dellinger and Yu, 2009). Band A (< 0.18 Hz) contains energy from ocean swells. This energy has tremendous amplitude in the pressure component (which explains why low-cut recording filters are customary) but is virtually absent in the vertical component of the particle velocity sensor. Bands B (0.18-0.35 Hz) and C (0.35-1.75 Hz) contain microseism energy generated by the interaction of weather-generated ocean waves and the sea floor (Longuet-Higgins, 1950; Rhie and Romanowicz, 2004, 2006). In ocean acoustics, this band is referred to as the double-frequency microseism peak. This energy has proven to be suitably random in propagation direction, as required for interferometry.

Bands D (1.75-18 Hz) and E (> 18 Hz) contain human-generated noise associated with field operations, boats and distant seismic shooting. Note that repetitive artificial sources can be identified by the spiky character of their spectra. Natural sources are more random and result in relatively smooth spectra. When weather is calm and the microseism peak is weak, human-generated noise sources overwhelm the microseism energy at the upper end of the microseism band. This pushes the upper limit of useful microseism energy down during calm weather conditions. Thus ambient seismic-noise data is best recorded during stormy conditions and/or during periods with few noise-making field operations.

Weather records (NMI, 2011) indicate that the 2005 recording began in calm weather that then transitioned to stormy, a change that is readily apparent in the change in the spectrum over time windows. Weather records for the 2004 dataset are not available, but the data show that the 2004 dataset started under weather conditions even calmer than the start of the 2005 dataset, and ended under weather conditions somewhat less stormy than at the end of the 2005 dataset. Comparing the spectrum of the (unfiltered) 2008 recording with the (filtered) 2004 and 2005 recordings shows that the low-cut filter suppresses much of the microseism band of interest. In this paper we wish to compare results for all three datasets, and so we will perform our analysis for energy in frequency band C, between 0.35-1.75 Hz. Band C is highlighted in Figure 1. Seismic interferometry for microseism energy also succeeds in band B for the 2008 recording (Dellinger and Yu, 2009; Dellinger et al., 2010).

EIKONAL TOMOGRAPHY

Tomography produces an estimated velocity field from a large number of measurements of travel time (or phase). Traditionally, a forward-modeling operator is constructed by solving the eikonal equation in a background velocity model for each observable data point. These data points then act as a global constraint on the inversion. The resulting linear system is typically solved by matrix inversion techniques, regularized by spatial smoothing and Tichonov regularization. In contrast, eikonal tomography (Lin et al., 2009) is a procedure that uses the spatial gradient of a travel-time map as a local constraint on the wave-speed map. This relation is described by the scalar eikonal equation (Aki and Richards, 2002):

$$|\nabla t_i(x)| = c_i^{-1}(x). \tag{1}$$

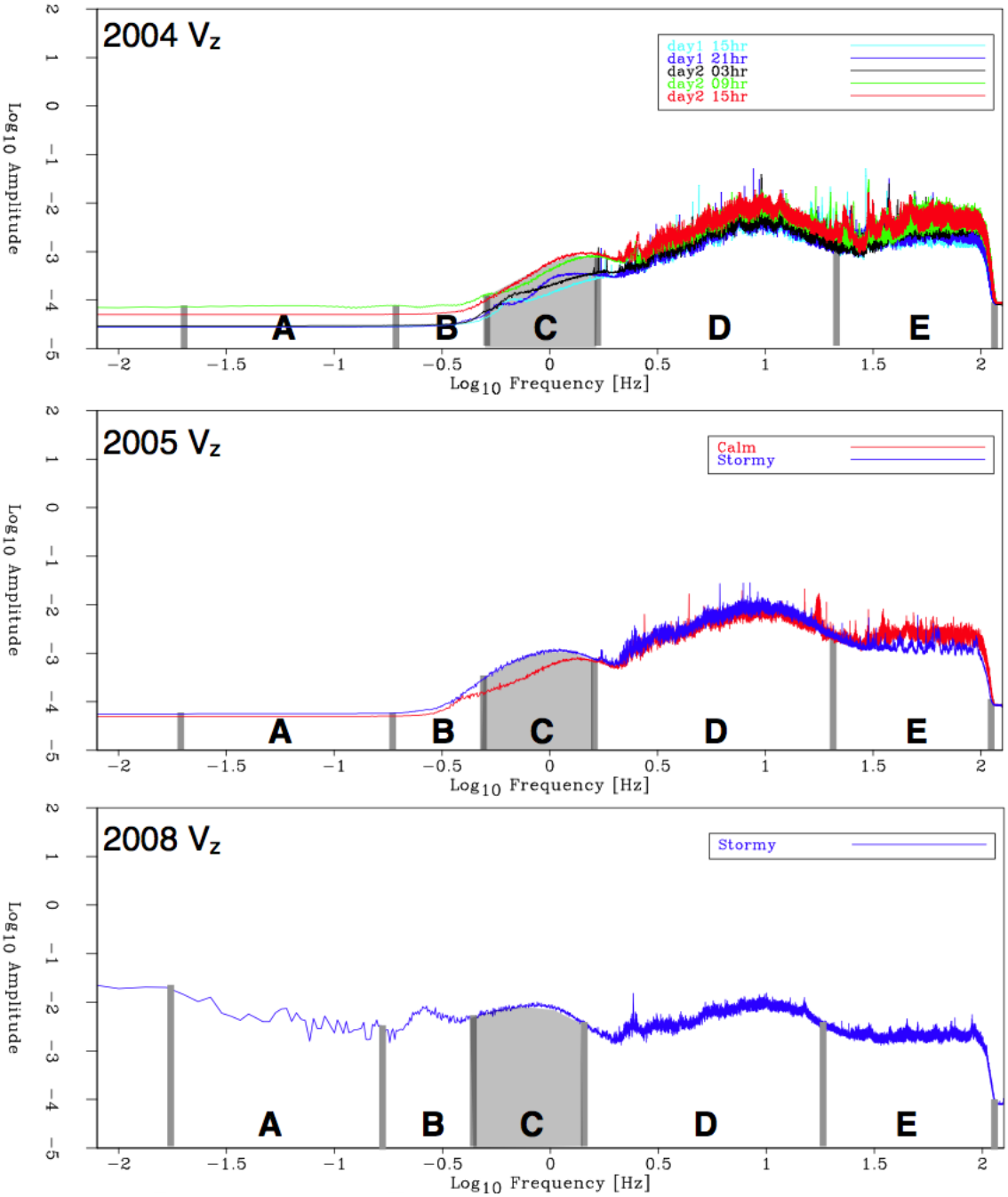


Figure 1: Frequency spectra of the vertical component of the particle velocity for the 2004 (top panel), 2005 (middle panel) and 2008 (bottom panel) datasets. We divide the data into five frequency bands labeled A to E. Band A (< 0.18 Hz) is the frequency regime for ocean swells, bands B (0.18-0.35 Hz) and C (0.35-1.75 Hz) contain microseism energy, and in bands D (1.75- 18 Hz) and E (> 18 Hz) human-generated noise dominates. The 2005 data have been subdivided into calm and stormy spectra. Five spectra for the 2004 dataset are shown, computed over consecutive 6-hour intervals. [**NR**] sjoerd1/. Valhall-AMS-Spectra

Evaluating this equation for one virtual-source gather leads to an estimate of the local wave velocity.

We perform ambient-seismic-noise eikonal tomography as follows: First, the data are bandpassed (for band C) and virtual-seismic source gathers are constructed by correlation. Figure 2a shows 30 seconds of bandpassed seismic noise recorded in the 2008 dataset by one line of the Valhall array. Figure 2b shows a virtual-source gather created by correlating 2 hours of noise along that line with the data for the station at the offset marked 0, located towards the right end of the line. For ideal ambient seismic noise, positive and negative time lags of the correlation should contain a symmetrized impulse response (Wapenaar and Fokkema, 2006). We stacked positive and negative time lags to enhance the quality of the created virtual-source gathers.

Next, traveltimes are estimated by picking the maximum value of the envelope function of the virtual-seismic-source data within a defined moveout window. Figure 2c shows picks for the virtual-source gather in figure 2b. We also calculate a signal-to-noise ratio (SNR) quality factor, defined as the amplitude of the envelope peak within the moveout window divided by the average amplitude outside the window. We keep only the traveltime picks exceeding a specified SNR and regularize them using biharmonic spline interpolation (Sandwell, 1987) (see Appendix). The traveltime map should only be interpreted for portions surrounded by picks that we kept. Next, we use equation 1 to estimate the local velocity at each point for waves from each virtual-source gather. The final step is to produce tomographic maps by statistically characterizing the estimates at each position in the model.

$$s_0(x) = \frac{1}{N}\sum_{i=1}^{N} s_i(x) \tag{2}$$

and

$$\sigma_s^2(x) = \frac{1}{N(N-1)}\sum_{i}^{N} \left(s_i(x) - s_0(x)\right)^2 . \tag{3}$$

Figure 3 shows how this works. Figure 3a shows an amplitude snapshot for a virtual-source gather gather in the southern portion of the Valhall data for the 2008 dataset. Notice that the impulse response is nearly perfectly omnidirectional. Figure 3b shows regularized travel times across the array for this virtual-source. Note that more distant picks did not have a sufficient SNR and were discarded. Figure 3c shows a map of velocity estimates calculated using those regularized picked travel times. This process is repeated for each virtual-source gather, potentially producing a large number of velocity estimates.

EXPLICIT INVERSION BY STATISTICS

Whereas for conventional seismic imaging, the inversion contains a wavefield extrapolation step (Claerbout, 1985), eikonal tomography finds an estimate of the model directly after merely regularizing the data. This possible because measurements are available throughout the domain of the model space. Different illuminations of the model by diferent sources are still added together through a weighted sum to find a composite model. Evaluating equation 1 for N virtual-source gathers, one at each receiver station, and applying the

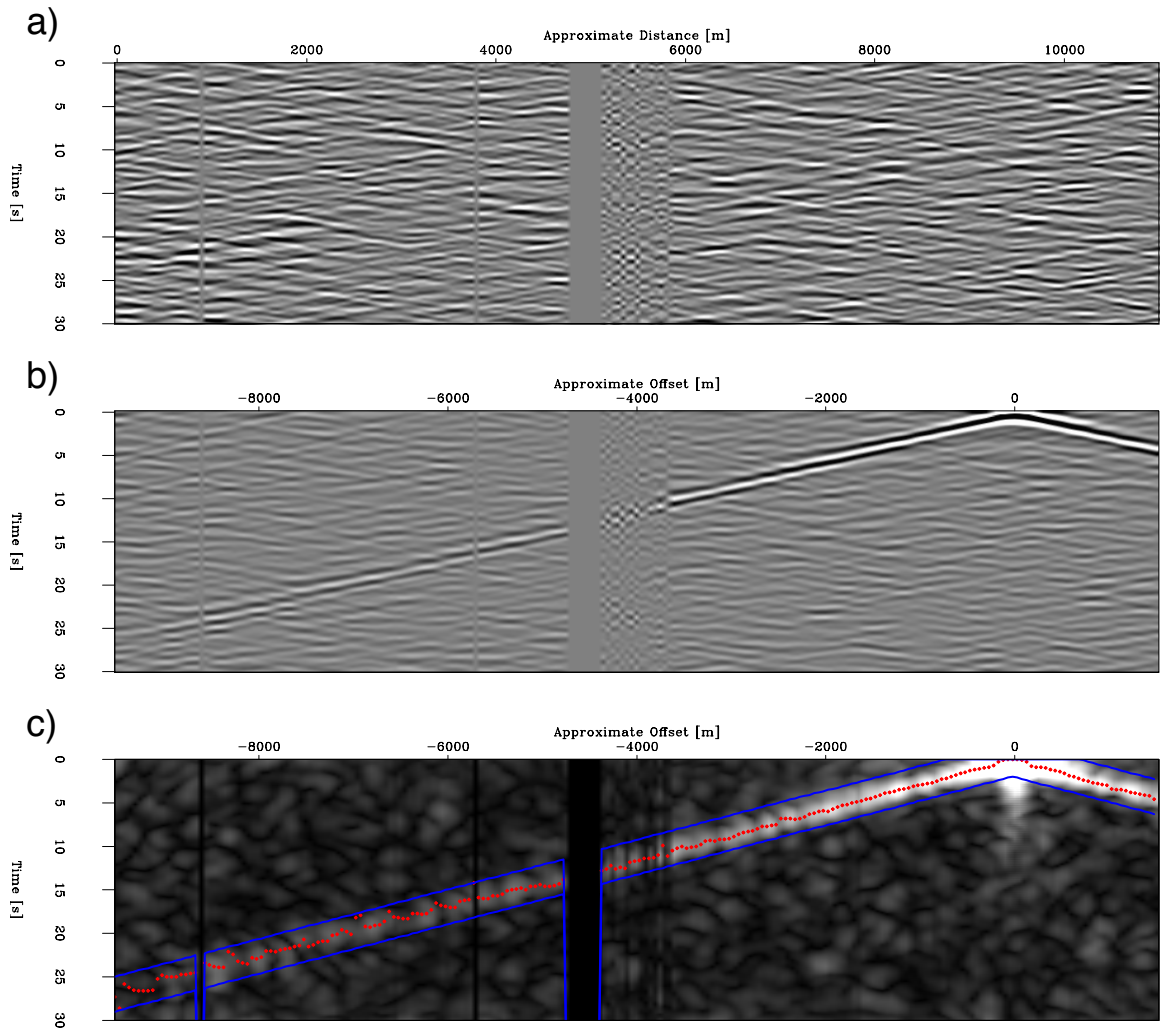


Figure 2: Three panels demonstrating the first half of the processing flow of ambient seismic noise eikonal tomography for a receiver line in the 2008 dataset: (a) 30 seconds of ambient seismic noise after bandpassing between 0.35 Hz and 1.75 Hz; (b) data from a virtual-source created at a station near the right end of the line; (c) travel-time picks superimposed on top of the envelope function of the virtual-seismic-source data in the middle panel. [**CR**] sjoerd1/. Laura-V-picked

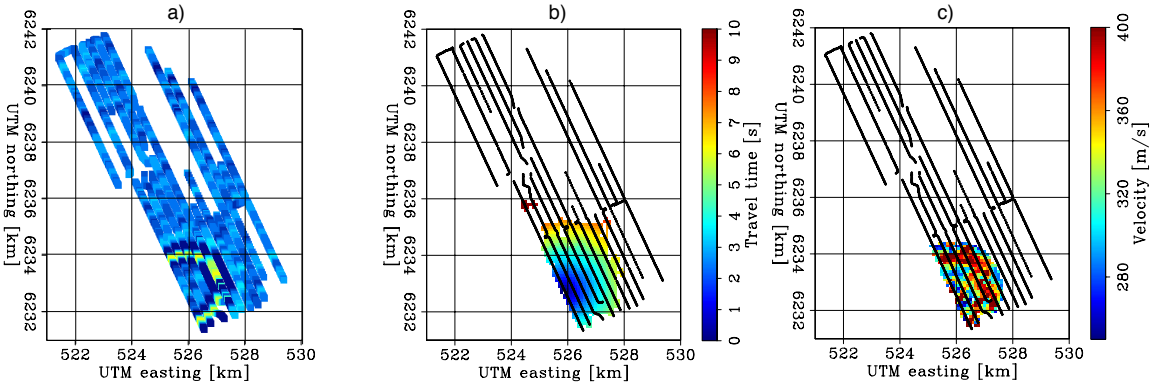
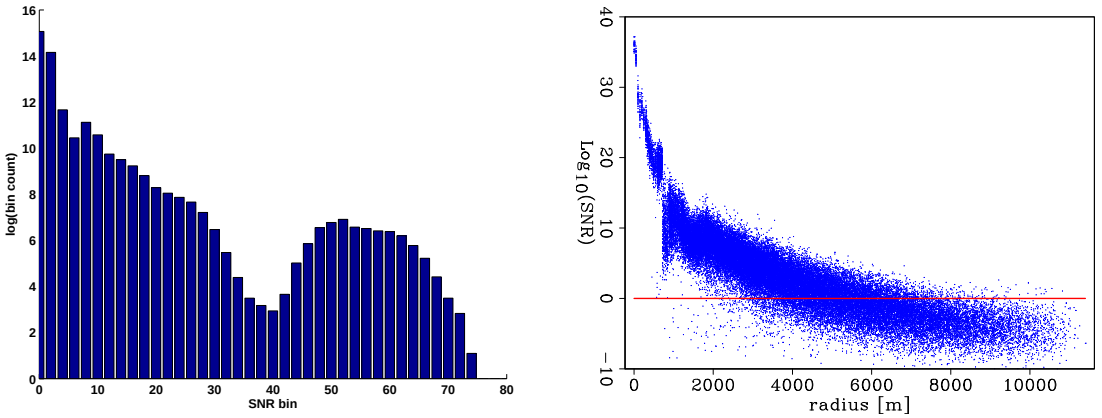


Figure 3: Three panels showing the continuation of the processing flow begun in Figure 2, traveltime map construction and velocity estimation for one virtual-source: (a) A snapshot showing Scholte-waves propagating away from a virtual-source in the southern portion of the Valhall array. The virtual-source was created using data from the 2008 dataset, bandpassed between 0.35 Hz and 1.75 Hz. (b) Regularized travel time surfaces picked from this virtual-source gather. (c) The resulting velocity estimates. [**NR**] sjoerd1/. Eikonal-tomo

acceptence criteria mentioned above leads to a set of estimates of the local wave velocity at each receiver location **x**.

A logarthmic histogram of the SNR of all picks in the Valhall array (Figure 4(a)) shows the large majority of picks have very low SNR. Generally the SNR ratio decreases quite rapidly with increasing source-receiver distance due to geometrical spreading (Figure 4(b)). The deficiency between SNR's of 30 to 40 and the jump in the SNR scatter plots at a radius of 750 m are related to the dispersion of the Scholte wave in the correlation gathers. They coincide with phase cycles, which are strongly reflected in the amplitude of the envelope function.



(a) (b)

Figure 4: a) Histogram of SNR. b) Scatter plot of 1% of all SNR's versus radius, line is the lower acceptence level. [**NR**] [**CR**] sjoerd1/. snrhist,snrscat

A complication with Gaussian statistics on measurements of velocity or slowness is that

they are both always positive. Tarantola (2006) suggests these so called *Jeffreys quantities* are more safely analyzed in their logarithmic magnitudes. Here we briely disuss the model estimate distributions. The expected value and variance of the estimate sets are calculated on the estimates in three magnitudes: linear slowness, linear velocity and logarithm of velocity. For linear slowness magnitudes we have (again)

$$s_0(x) = \frac{1}{N}\sum_{i=1}^{N} s_i(x) \tag{4}$$

and

$$\sigma_s^2(x) = \frac{1}{N(N-1)}\sum_{i}^{N} \left(s_i(x) - s_0(x)\right)^2 . \tag{5}$$

For linear velocity magnitudes we calculate

$$c_0(x) = \frac{1}{N}\sum_{i=1}^{N} c_i(x) \tag{6}$$

and

$$\sigma_c^2(x) = \frac{1}{N(N-1)}\sum_{i}^{N} \left(c_i(x) - c_0(x)\right)^2 . \tag{7}$$

When characterizing logarithmic slowness or velocity magnitudes an expected value and variance can be obtained from either slowness or velocity as

$$\ell_0(x) = -\frac{1}{N}\sum_{i=1}^{N} \log_{10}\left(s_i\right) = \frac{1}{N}\sum_{i=1}^{N} \log_{10}\left(c_i\right) \tag{8}$$

and

$$(\ell_\sigma)^2(x) = \frac{1}{N(N-1)}\sum_{i}^{N} \left(\log_{10}\left(s_i\right) + \ell_0\right)^2 = \frac{1}{N(N-1)}\sum_{i}^{N} \left(\log_{10}\left(c_i\right) - \ell_0\right)^2 . \tag{9}$$

The results of either approach are most intuitively compared in units of velocity (m/s). So the expected values are compared through:

$$c_0 \text{ with } s_0^{-1} \text{ and } 10^{\ell_0}. \tag{10}$$

The variances are compared through:

$$\sigma_c \text{ with } \frac{1}{s_0^2}\sigma_s \text{ and } \frac{1}{2}\ 10^{\ell_0}\left[10_\sigma^{\ell} - 10^{-\ell_\sigma}\right]. \tag{11}$$

The last term is the average between two velocity values corresponding to the expected value plus and minus the square-root of variance in logarithmic space.

Normal distributions are fitted to the estimates in Figures 5 and 6 on the respective magnitudes. For the estimates near a high-velocity anomaly we find $c_0 = 348$ m/s and $\sigma_c =$

38.8 m/s, $s_0 = 0.00290 \Rightarrow 345$ m/s and $s_\sigma = 0.000328 \Rightarrow 38.7$ m/s, and $\ell_0 = 2.54 \Rightarrow 346$ m/s and $\ell_\sigma = 0.0487 \Rightarrow 38.9$ m/s. For the estimates near a low-velocity anomaly we find $c_0 = 299$ m/s and $\sigma_c = 36.36$ m/s, $s_0 = 0.00339 \Rightarrow 295$ m/s and $s_\sigma = 0.000410 \Rightarrow 35.6$ m/s, and $\ell_0 = 2.47 \Rightarrow 297$ m/s and $\ell_\sigma = 0.0527 \Rightarrow 36.1$ m/s. For both estimate sets we find the expected value on the logarithmic velocity magnitude to be within the values of the expected value on the linear slowness and linear velocity magnitudes. However the deviations on these expected values are about 10 times smaller then the standard deviations on the estimates, regardless of which magnitudes are analyzed.

In Figures 5(d) and 6(d) Q-Q plots indicate how well the estimates fit normal distributions (plus symbols for slowness magnitudes, circles for velocity magnitudes and quandrangles for logarithmic velocity magnitudes). Q-Q plots for the different magnitudes carry a different unit on their vertical axis. The diagonal corresponds to a perfect fit. The intermediate quantiles follow Gaussian statistics quite well. In the lower and higher quantiles the data differs from a normal distribution. In the high-velocity anomaly (Figure 5(d)) the slowness magnitudes fit a normal distribution better in the lower quantiles and the velocity magnitudes better in the higher quantiles (not surprising because velocity is inverse slowness). In the low-velocity anomaly (Figure 6(d)) the logarithmic velocity magnitudes fit a normal distribution best in the lower quantiles while the velocity magnitudes better in the higher quantiles.

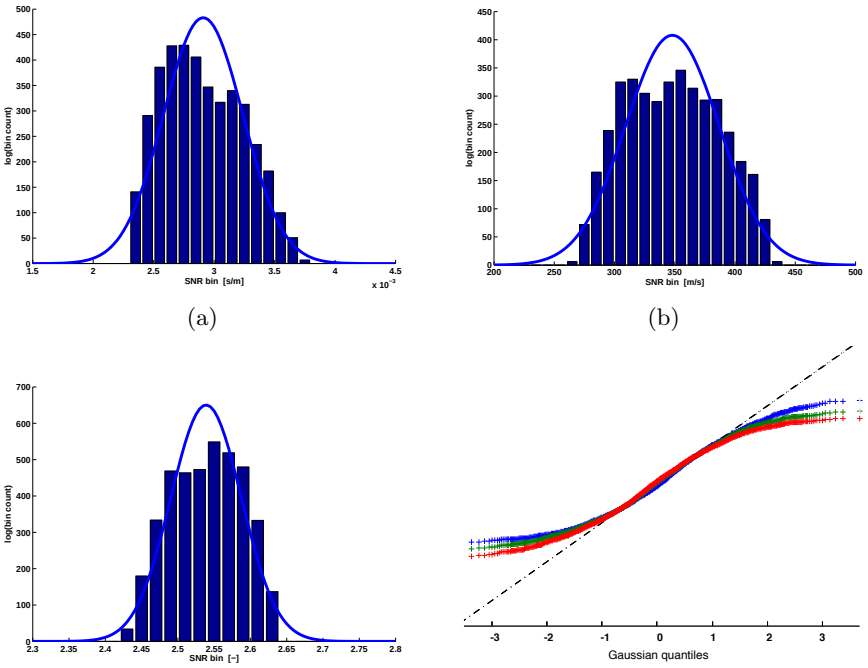


Figure 5: Histograms and their normal distribution fits of the gradient estimates in a high velocity region: a) linear slowness magnitude, b) linear velocity magnitude, c) logarithmic velocity magnitude, d) normal Q-Q plots of the gradient magnitudes (plus symbols: linear slowness, circles: linear velocity, quandrangles: logarithmic velocity). [**CR**] [**CR**] [**CR**] [**NR**] sjoerd1/. estimates1a,estimates1b,estimates1d,fits1

These observations might not hold in general for low- and high-velocity anomalies. For-

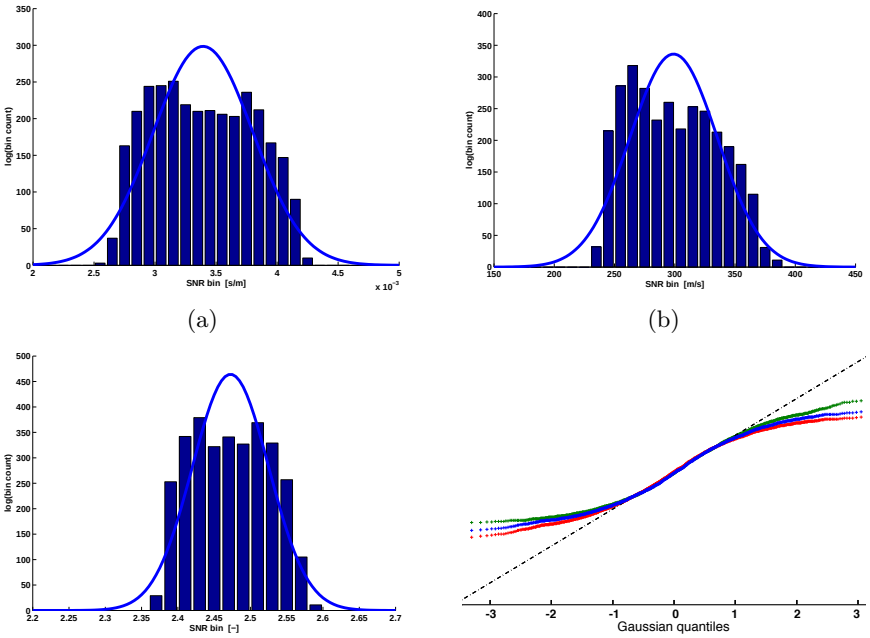


Figure 6: Histograms and their normal distribution fits of the gradient estimates in a low-velocity region: a) linear slowness magnitude, b) linear velocity magnitude, c) logarithmic velocity magnitude, d) normal Q-Q plots of the gradient magnitudes (plus symbols: linear slowness, circles: linear velocity, quandrangles: logarithmic velocity). [**CR**] [**CR**] [**CR**] [**NR**] sjoerd1/. estimates2a,estimates2b,estimates2d,fits2

tunately, the choice of which magnitude to use to analyze the gradient estimates does not significantly change the resulting tomography maps. Gaussian statistics do not adequately fit the tails of the estimate sets. For simplicity we chose to base the tomographic images on the expected value of a linear slowness basis. For this data, the variance is a very limited measure and does not infer an estimate certainty in a Gaussian statistics sense. Thus two additional quality measures of the expected value are the density of accepted picks and the velocity-estimate density. The first is a (regularized) map of the number of accepted travel time picks per station. The second is simply a map of N, the number of raw travel time measurements available at that location.

SCHOLTE-WAVE VELOCITY MAPS

Correlations were performed for time blocks of at most 3 hours. Signal to noise ratios were calculated using a moveout window 4 seconds wide centered on a moveout-velocity trajectory of 350 m/s. For the inversion we only used picks with SNRs of over 2 that were located within a radius of 1750 m of the virtual-source gather. We found the eikonal tomography results to be more improved by excluding picks at larger offsets than by elevating the SNR threshold. Further offsets might get sensitive to deeper structures or contain other wavemodes. Thus the gradients would fundamentally start to conflict. Figures 7-9 contain images of Scholte-wave velocities and quality factors obtained by performing eikonal tomography on the 2004, 2005 and 2008 datasets, respectively. The top left panels show the expected-value velocity maps, and the top right panels show the associated standard deviations. The lower left and right panels show maps of the regularized pick density and the number of estimates (N), respectively. For comparison, Figure 10 shows maps of estimated P-wave (not Scholte-wave) velocities between 60 and 105 m (in Figure 10a) and between 150 and 195 m (in Figure 10b) beneath the ocean floor. This result was obtained by Sirgue et al. (2010) using full-waveform inversion (FWI) between 3.5 Hz and 7 Hz on a conventional active seismic dataset.

Several high- and low-velocity anomalies can be distinguished in the Scholte-wave maps from the 2004 and 2005 data. These appear to correspond to structures also observable in the P-wave velocity map. The high-velocity anomalies are interpreted to be channel features (indicated by arrows). Dotted circles indicate two distinct low-velocity zones. The FWI inversion was performed for frequencies higher than those in this eikonal tomography study, so it is not surprising that the tomographic images of Scholte-wave velocities obtained from the eikonal tomography lack the short-wavelength resolution of the P-wave velocity map obtained from FWI.

The traveltime map estimated from the 2008 dataset contains generally higher velocitiesthan the traveltime maps from the 2004 and 2005 datasets. This might be because the (unfiltered) 2008 dataset contains stronger low frequencies, and thus is more sensitive to higher velocities deeper in the Earth. However, parts of the 2008 map (especially in the South) bear little resemblance to the 2004 or 2005 maps, or to the structures in the P-wave velocity map. Only 2 hours of data were available from 2008, and as a result the SNRs of the virtual-source gathers generally were lower. The 2008 Scholte-wave velocity map was thus constructed using far fewer accepted picks than the maps of 2004 and 2005, so the map from the 2008 dataset may simply be unreliable. Availability of longer data sets recorded without a low-cut filter would enable further investigation of the application of the lower

frequencies in the microseism energy to subsurface imaging.

Note the isolated low near the center of the array in the pick-density map for the 2004 eikonal inversion. This receiver station was probably malfunctioning (or possibly was simply mislocated) and slipped through initial quality control. It did not damage the results, however, because it was discarded by the automated picking algorithm. The P-wave velocity map in Figure 10b prominently displays a large meandering channel that passes under the extreme southeastern part of the array. This channel is deeper than the narrower channels marked by the meandering dotted arrows in Figure 10b. There is no corresponding anomaly in the eikonal result for the 2004 data. There are anomalies in the 2005 and 2008 data at approximately the right location, but they are certainly not conclusive matches. It is possible that the lower-frequency, deeper-sensing 2008 dataset might be detecting that channel. Unfortunately the edges of the array have relatively few travel times to work with, so this remains inconclusive. Note that FWI was able to use active sources, which covered a wider area, to image well outside the receiver array even at these shallow depths. Interferometry uses the receivers as sources and so has a more restricted image space.

CONCLUSIONS

The ambient seismic field at Valhall contains omnidirectional noise suitable for seismic noise tomography at low frequencies between 0.35 Hz and 1.75 Hz. Scholte-waves traveling along the ocean floor can provide a tomographic image of features in the near surface (between 0 and 105 meters beneath the ocean floor). We have demonstrated that eikonal tomography is an effective tool for studying surface waves at Valhall, even though the gradient estimate sets do not closely resemble Gaussian distributions. This research shows that there is potential for developing a low-cost near-surface monitoring system using passive noise recorded by permanent ocean-bottom arrays.

ACKNOWLEDGEMENTS

This research was initiated as part of a summer internship at BP America by S. de Ridder under the supervision of Joe Dellinger. The authors would like to thank Ray Abma, Olav Barkved, John Etgen, Phuong Vu, and Jianhua Yu for helpful discussions and suggestions. We also thank BP and the partners of the Valhall Field (BP Norge and Hess Norge) for permission to publish this paper. S. de Ridder would like to acknowledge Biondo Biondi, Jon Claerbout, Bob Clapp and Andre Journel for fruitful discussions and the sponsors of the Stanford Exploration Project for their financial support.

REFERENCES

Aki, K. and P. G. Richards, 2002, Quantitative Seismology - second edition: University Science Books.

Artman, B., 2007, Passive seismic imaging: PhD thesis, Stanford University.

Bussat, S. and S. Kugler, 2009, Recording noise - estimating shear-wave velocities: Feasibility of offshore ambient-noise surface-wave tomography (answt) on a reservoir scale: SEG Technical Program Expanded Abstracts, **28**, 1627–1631.

Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Scientific Publications.

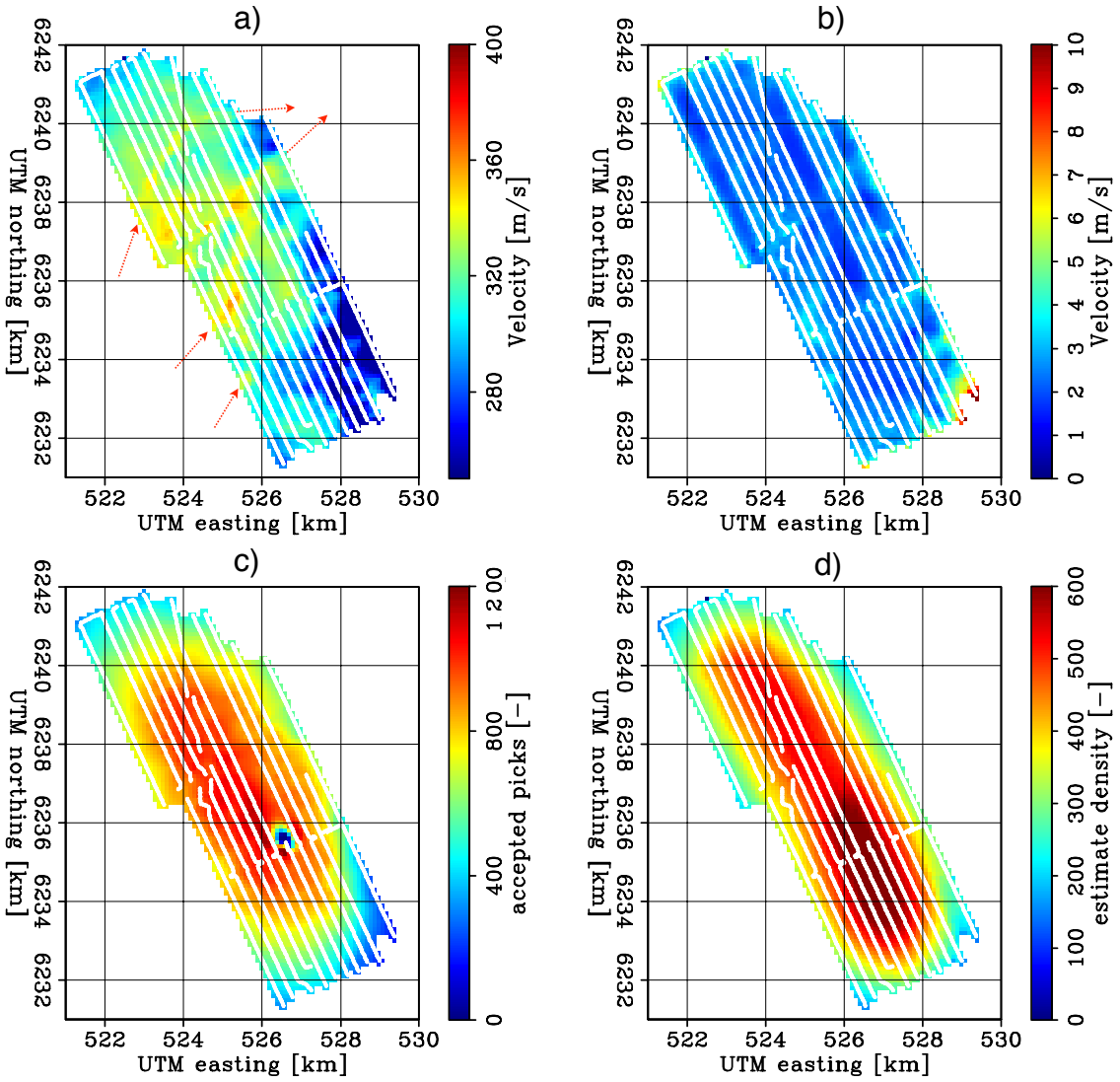


Figure 7: Images of Scholte-wave velocities and uncertainties obtained from eikonal tomography on the 2004 dataset; the expected value (a), the standard deviation of velocity (b), pick density at each station (c), and the number of velocity estimates at each grid point (d). Dotted arrows in (a) indicate channel features. [**CR**] sjoerd1/. Artman-V-Etom

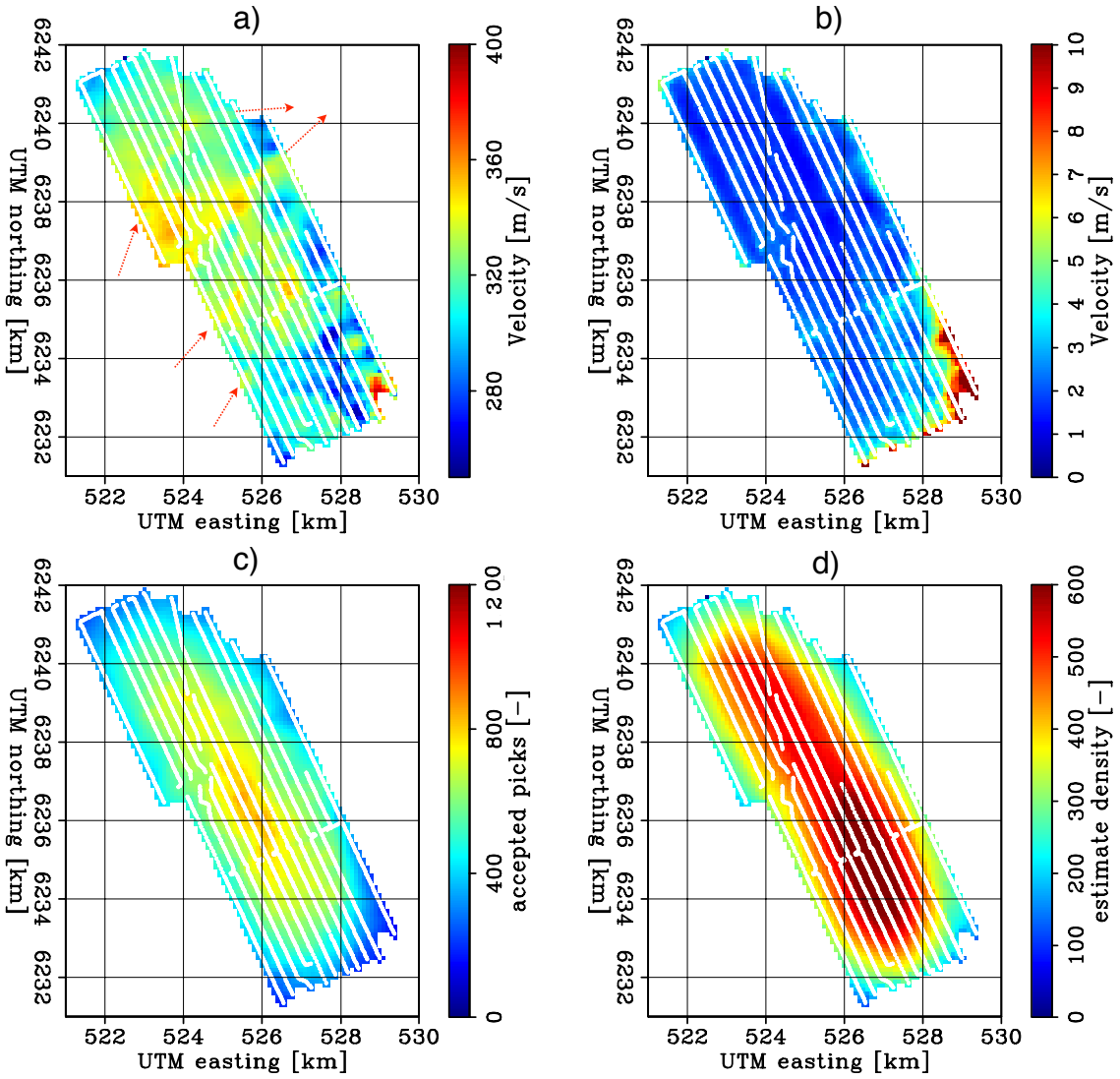


Figure 8: Images of Scholte-wave velocities and uncertainties obtained from eikonal tomography on the 2005 dataset; the expected value (a), the standard deviation of velocity (b), pick density at each station (c), and the number of velocity estimates at each grid point (d). Dotted arrows in (a) indicate channel features. [**CR**] sjoerd1/. Jianhua-V-Etom

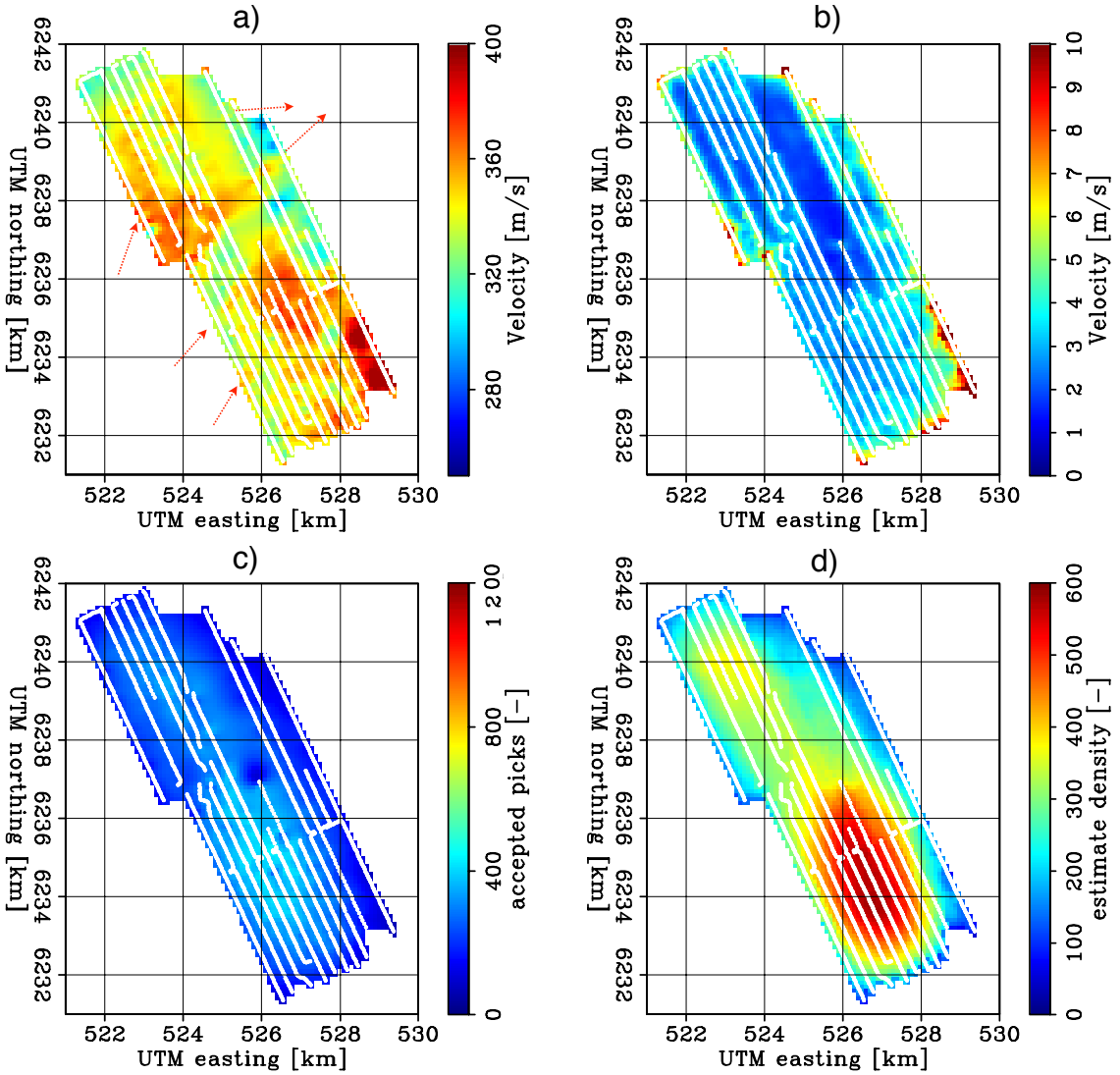


Figure 9: Images of Scholte-wave velocities and uncertainties obtained from eikonal tomography on the 2008 dataset; the expected value (a), the standard deviation of velocity (b), pick density at each station (c), and the number of velocity estimates at each grid point (d). Dotted arrows in (a) indicate channel features. This dataset produces a distinctly different result from the other two, possibly because it was recorded without a low-cut recording filter. [**CR**] sjoerd1/. Laura-V-Etom

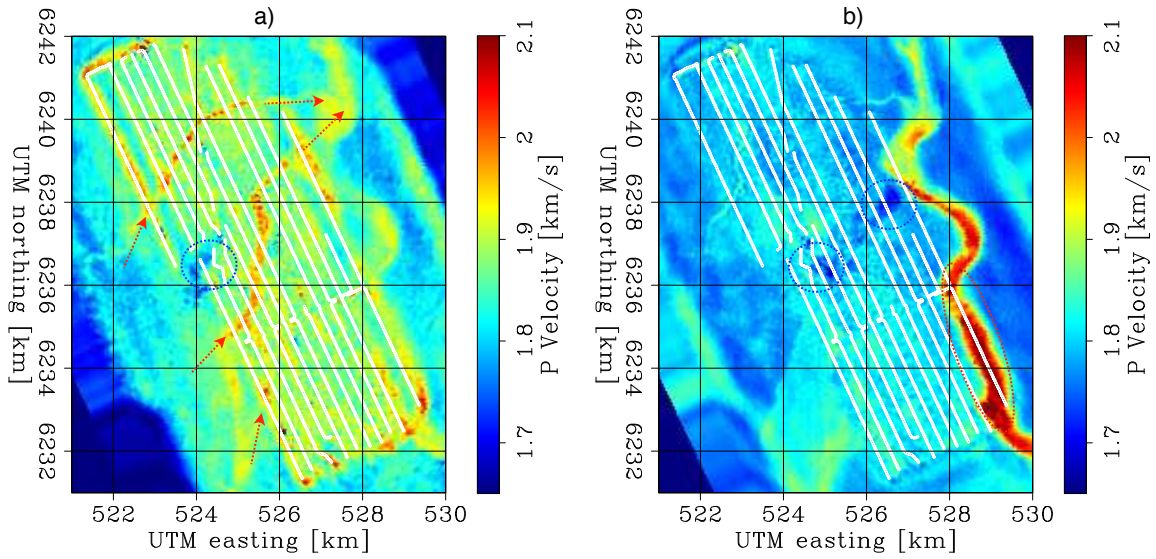


Figure 10: Images of average P-wave velocities obtained using waveform inversion (Sirgue et al., 2010) of active P-wave data; (a) between 60 and 105 m beneath the ocean floor, and (b) between 150 and 195 m beneath the ocean floor. Dotted arrows indicate channel features; dotted circles indicate two distinct low-velocity zones. [**NR**] sjoerd1/. Valhall-Laurent-FWI

Dellinger, J., J. Yu, and S. de Ridder, 2010, Virtual-source interferometry of 4C OBC data at Valhall without a low-cut recording filter: SEG low-frequency workshop, Expanded Abstracts, Session 6.

Dellinger, J. A. and J. Yu, 2009, Low-frequency virtual point-source interferometry using conventional sensors: 71st Meeting, European Associated of Geoscientists and Engineers, Expanded Abstracts, Expanded Abstracts, X047.

Kommedal, J. H., O. I. Barkved, and D. J. Howe, 2004, Initial experience operating a permanent 4C seabed array for reservoir monitoring at Valhall: SEG Technical Program Expanded Abstracts, **23**, 2239–2242.

Landes, M., N. M. Shapiro, S. Singh, and R. Johnston, 2009, Studying shallow seafloor structure based on correlations of continuous seismic records: SEG Technical Program Expanded Abstracts, **28**, 1693–1697.

Lin, F. C., M. H. Ritzwoller, and R. Snieder, 2009, Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array: Geophys. J. Int., **177**, 1091–1110.

Longuet-Higgins, M. S., 1950, A theory of the origin of microseisms: Phil. Trans. R. Soc. Lond. A, **243**, 135.

Moreau, J.-P., 2011, Programs concerning matrixes in fortran 90: http://jean-pierre.moreau.pagesperso-orange.fr/. (last visited: January 12, 2011).

NMI, 2011, Norwegian Meteorological Institute climate data database: http://sharki.oslo.dnmi.no/. (last visited: January 12, 2011).

Olofsson, B., 2010, Marine ambient seismic noise in the frequency range 1–10 hz: The Leading Edge, **29**, 418–435.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1986, Numerical recipes: Cambridge University Press.

Rhie, J. and B. Romanowicz, 2004, Excitation of earth's continuous free oscillations by atmosphere-ocean-seafloor coupling: Nature, **431**, 552–556.

———, 2006, A study of the relation between ocean storms and the earths hum: Geochem. Geophys. Geosyst., **7**, Q10004.

Sandwell, D. T., 1987, Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data: Geophys. Res. Let., **14**, 139–142.

Sirgue, L., O. I. Barkved, J. Dellinger, J. E. U. Albertin, and J. H. Kommedal, 2010, Full waveform inversion: the next leap forward in imaging at Valhall: First Break, **28**, 65–70.

Tarantola, A., 2006, Elements for Physics - Quantities, Qualities, and Instrinsic Theories: Springer.

Wapenaar, K. and J. Fokkema, 2006, Green's function representations for seismic interferometry: Geophysics, **71**, SI33–SI46.

Wessel, P., 2009, A general-purpose Green's function-based interpolator: Computers & Geosiences, **35**, 1247–1254.

APPENDIX: BICUBIC SPLINE REGUARIZATION

Bicubic spline interpolation is method of interpolation and regularization that relies on fitting the data by a set of weighted Green's functions for cubic cplines (Sandwell, 1987). It is intuitively comparable to bending a metal plate to fit through desired points, by applying and positioning different weights at positions along the plate. The Green's function for a cubic spline with forcing at $\delta(\mathbf{x})$ satisfies

$$\nabla^4 G(\mathbf{x}) = \delta(\mathbf{x}). \tag{12}$$

To fit N datapoints using N forcing functions weighted by w_j, we have the system

$$\nabla^4 d_i(\mathbf{x}) = \sum_{j=1}^{N} w_j \delta(\mathbf{x}_i - \mathbf{x}_j). \tag{13}$$

Using the defined Green's function, we have the system

$$d_i(\mathbf{x}) = \sum_{j=1}^{N} w_j G(\mathbf{x}_i - \mathbf{x}_j), \tag{14}$$

or in matrix notation

$$\mathbf{d} = \mathbf{G}\mathbf{w}, \tag{15}$$

where $G_{ij} = G(\mathbf{x}_i - \mathbf{x}_j)$ is a kernel with Green's functions. The system is solved using an f90 library that performs LU decomposition (Moreau, 2011; Press et al., 1986). Green's function solutions for cubic splines in various dimensions have been derived and are summarized by Wessel (2009). This paper uses the two dimensional solution

$$G(\mathbf{x}_i - \mathbf{x}_j) = r^2 \left(\ln r - 1 \right), \tag{16}$$

where $r = |\mathbf{x}_i - \mathbf{x}_j|$.

Interpreter input for seismic image segmentation

Adam Halpert

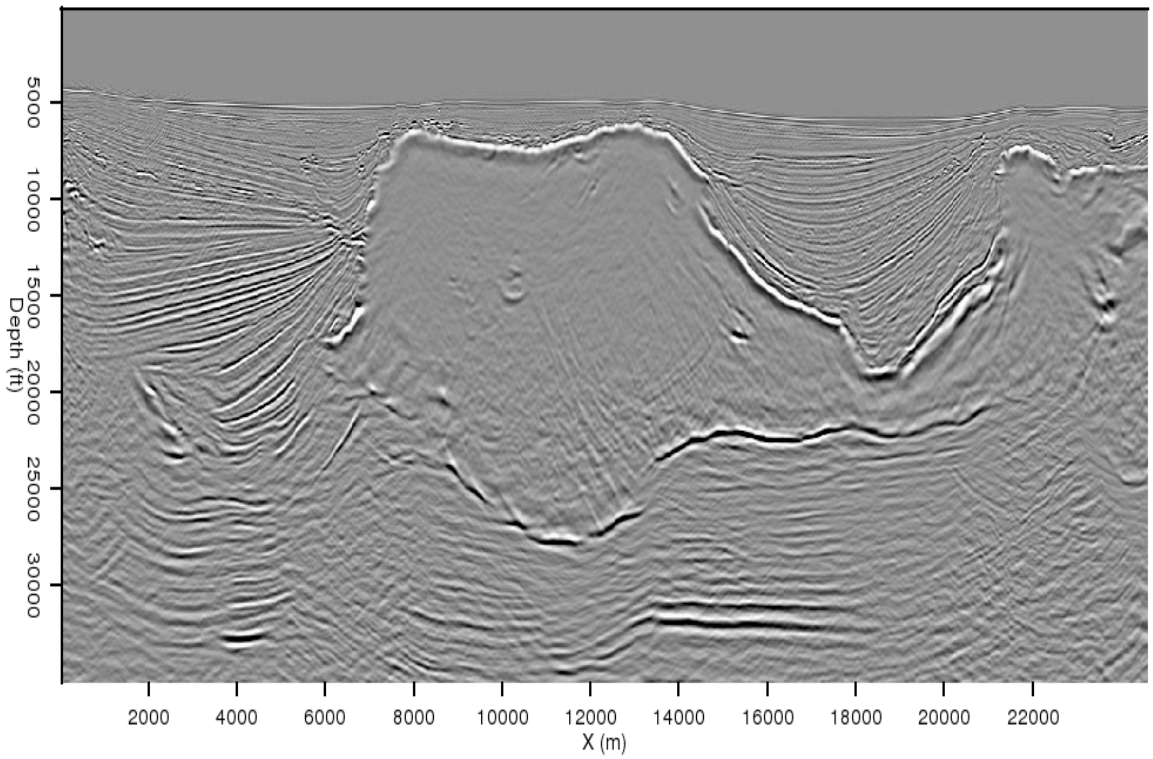
ABSTRACT

While automatic segmentation of seismic images can dramatically speed up interpretation of salt bodies and alleviate a major model-building bottleneck, human expertise at this task is valuable and should be included. Here, I demonstrate a strategy to incorporate such expertise into the highly-efficient Pairwise Region Comparison (PRC) segmentation algorithm. By supplying a limited manual interpretation of the salt boundary, interpreters can correct local inaccuracies and guide the automatic result in both two and three dimensions. In the 3D case, accuracy of the segmentation improves even in areas far from the manual picks. Examples from a wide-azimuth Gulf of Mexico survey demonstrate the effectiveness of this procedure.

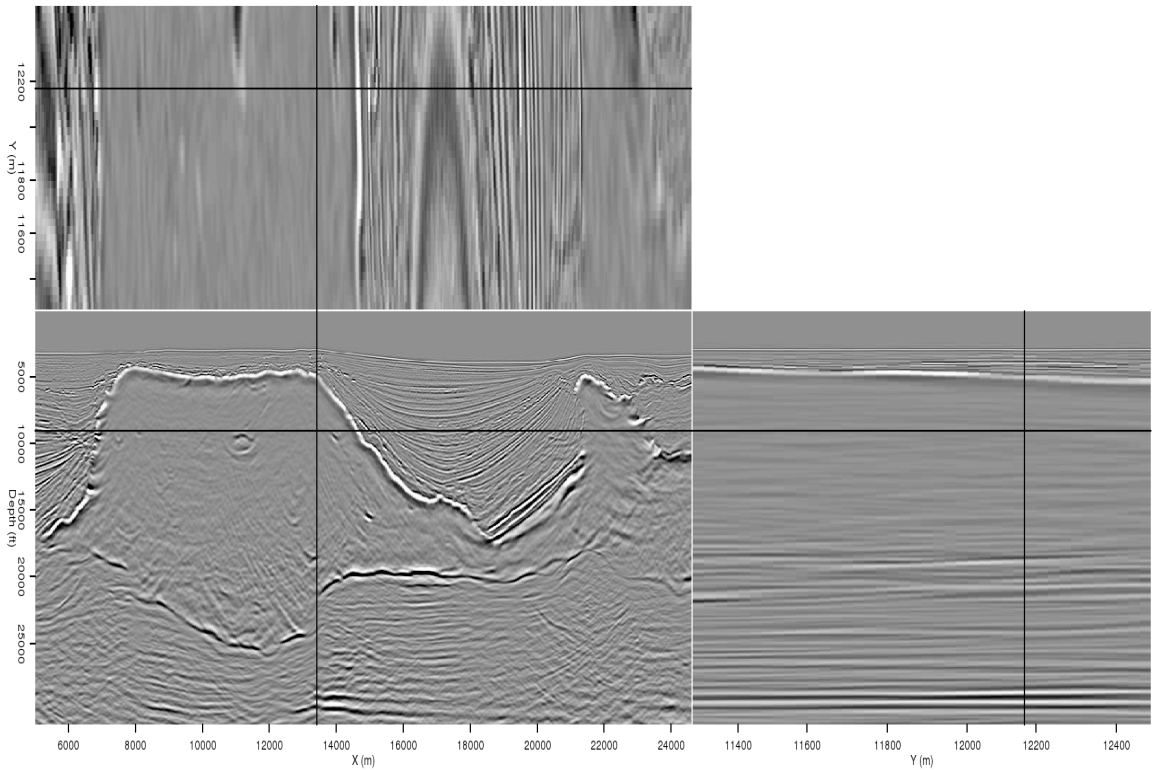
INTRODUCTION

Automated seismic image segmentation allows for fast interpretation of regions within seismic images, and is especially useful for identifying large subsurface salt bodies that are tedious to interpret manually. This automation helps to alleviate a substantial bottleneck within the iterative imaging, interpretation and model-building workflow. At this point, however, complete automation is not a feasible or even desirable goal. Experienced human interpreters offer a great deal of expertise, especially in complex geological settings where computers are unable to provide an accurate automatic interpretation. Here, I discuss a strategy for incorporating such expertise into the framework of automated image segmentation.

Beyond relatively simple horizon auto-picking, which tends to get lost along chaotic or discontinuous boundaries, a variety of options exist for *global* image segmentation. One useful approach, developed by Lomask et al. (2007), uses the eigenvector-based Normalized Cuts Image Segmentation (NCIS) method (Shi and Malik, 2000). However, this method is relatively inefficient; large seismic images require substantial preprocessing, and the computational domain must be windowed around a prior best-guess of the boundary to make the method computationally feasible. More recent work (Halpert et al., 2010) adopts a "Pairwise Region Comparison" (PRC) approach based on the method of Felzenszwalb and Huttenlocher (2004). This method holds several advantages over the NCIS approach, including computational efficiency and the ability to operate on full seismic images. In this paper, the PRC method is used as a basis to explore how an interpreter's own top- or base-salt picks can influence automated segmentation results in two or three dimensions. Throughout the paper, I show 2D (Figure 1(a)) and 3D (Figure 1(b)) field data examples from a wide-azimuth Gulf of Mexico dataset provided by WesternGeco.



(a)



(b)

Figure 1: (a) 2D and (b) 3D images taken from a field dataset that will be used for examples throughout this paper. [**ER**] adam1/. oct-2d,oct-3d

SEGMENTATION METHOD

The PRC segmentation algorithm is a graph-based method that relies heavily on the concept of "minimum spanning trees" (Zahn, 1971). In short, the algorithm makes comparisons between pairs of nearby pixels in a migrated image, and determines whether the pixels belong in the same image segment, or in separate regions. See Halpert (2010) for full details on how this method is adapted for seismic images. This algorithm is very efficient compared to other segmentation techniques such as NCIS; for example, the 1000 x 1000 pixel 2D result here was computed in less than a minute, while the 3D result (1000 x 800 x 40) required 10 minutes on a single processor. Parallelization of this scheme is made possible by operating simultaneously on different chunks of the input cube, although additional pre- and post-processing is necessary. Segmentation results (for example, Figures 2(a) or 5(a)) are shown as random-color segments overlying the image. Very few parameters are required; the user controls the minimum segment size, and can quickly choose segments to merge together in two or three dimensions. All results shown here required merging 5-10 segments within the salt body.

INTERPRETER GUIDANCE

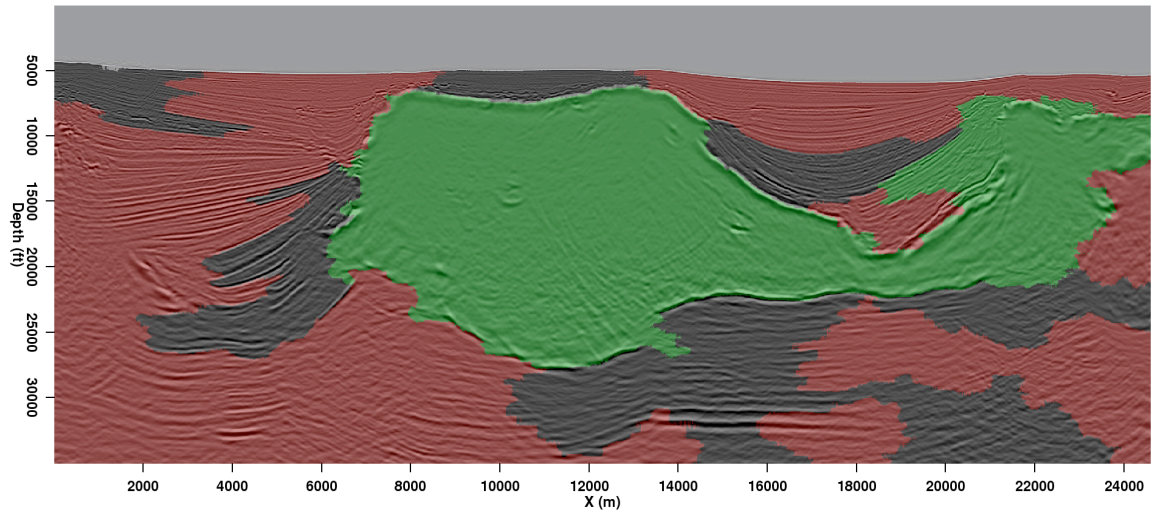
It is clear from the initial 2D segmentation result in Figure 2(a) that the algorithm did not perform ideally in several locations. Specifically, the salt segment "leaks" through the boundary at two locations, one along the top of the salt body and one along the base. In such situations, we want to use interpreter input to rectify the inaccuracies. Figure 2(b) shows manual interpretations of the salt boundary in the region where the leakages occurred. These boundary segments were quickly picked using a few seed points and an auto-picker.

At this point, these manual interpretations must somehow be incorporated into the automated framework of the PRC method. Broadly speaking, this can be done in two ways. First, we can modify the algorithm itself to check for the manual interpretations when assigning weights to pixel pairs. However, this option adds unnecessary complexity to the process and would likely severely curtail the algorithm's efficiency (one of its major advantages). Alternatively, we could modify the input data according to the picks, and allow the segmentation algorithm to proceed normally. Since this option preserves the algorithm's efficient structure, I focus here on modifying the input amplitude of the envelope volumes to accommodate the manual interpretations.

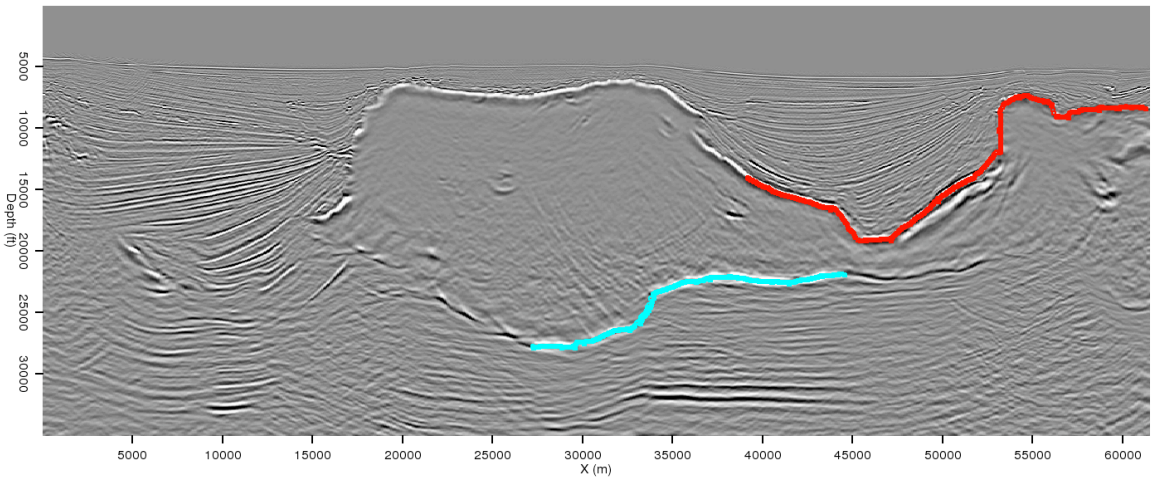
A simple solution to the problem would involve assigning a high amplitude value to the manually picked points. While this is the basic idea behind the solution described here, it is unfortunately not quite so straightforward. Assigning high amplitudes to a string of points without regard to neighboring amplitude values causes the algorithm to treat the manual interpretation as a segment distinct from those surrounding it, even those also inside a salt body. Instead, I define a new amplitude value (A) for a "picked" pixel at position (x,y,z) in terms of the highest-amplitude pixel in a neighborhood surrounding it and a scaling factor α:

$$A_{xyz} = \alpha \max_{ijk} A_{ijk}, \tag{1}$$

where i, j, and k are constrained by $|x-i| \leq 5, |y-j| \leq 5, |z-k| \leq 5$. This approach ensures that a boundary will be emphasized on the input data where it has been manually

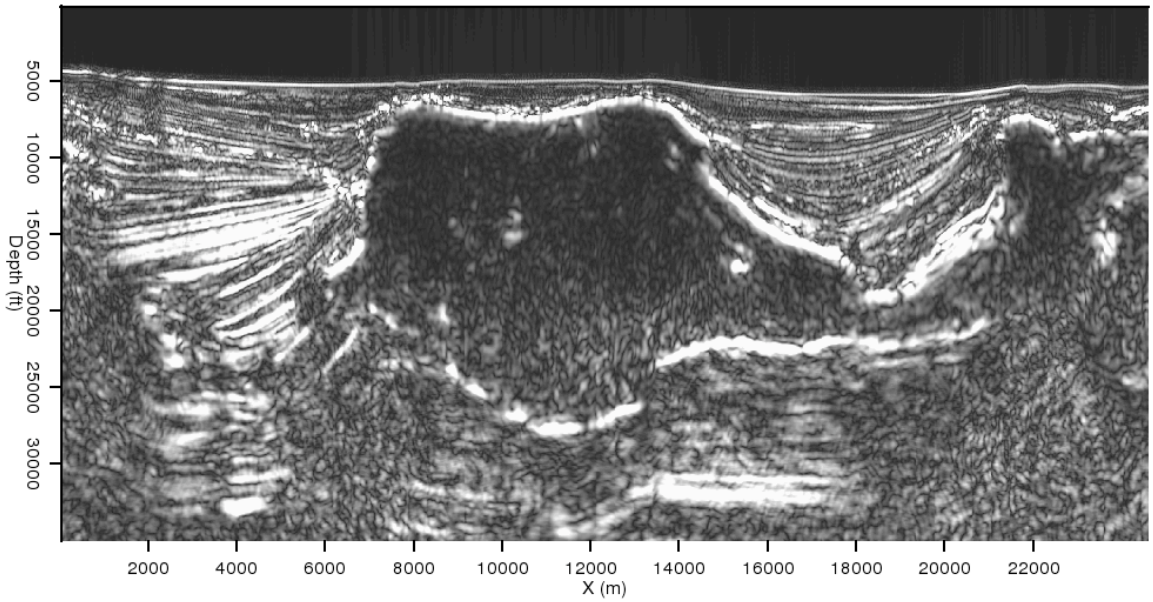


(a)

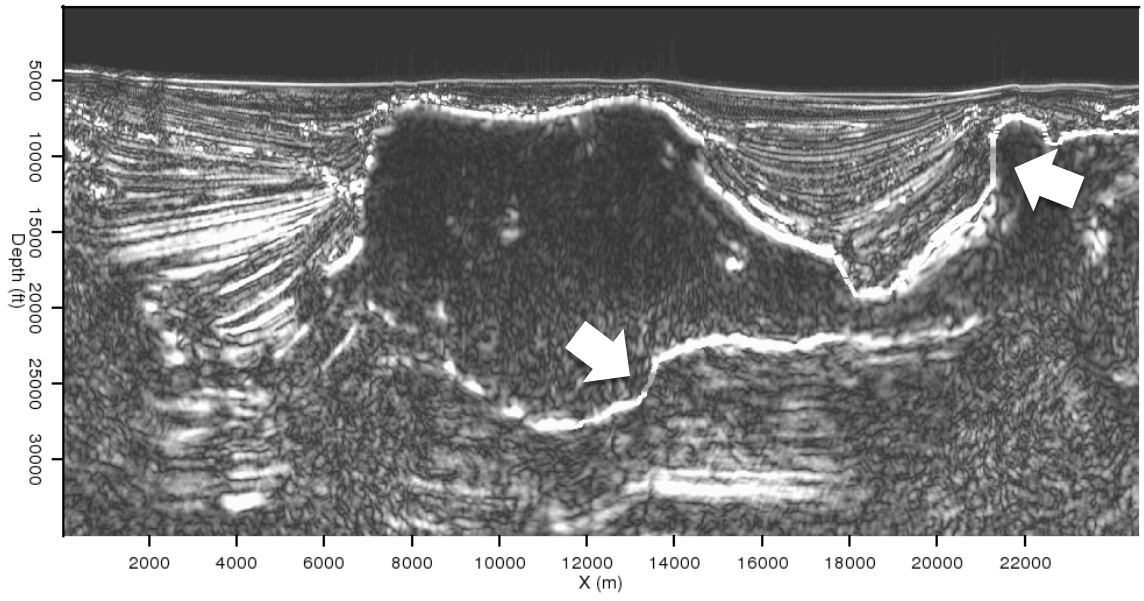


(b)

Figure 2: (a) 2D segmentation result without interpreter guidance. The salt body segment shows significant "leakage" at the locations indicated. (b) Manual interpretations of the salt boundary near where the leakages occurred. [**ER**] adam1/. 2d-merge,2d-topbase



(a)



(b)

Figure 3: Amplitude of the envelope calculations from the original image (a), and modified according to the manual salt picks in Figure 2(b). The salt boundary is clearer and more continuous in the modified image, especially at the indicated locations. [**ER**] adam1/. 2d-env-orig,2d-env-new

picked, but will not appear radically different from its surroundings. Figures 3(a) and 3(b) illustrate this idea; the manually picked points are now clearly visible in panel (b), but the character of the boundary does not change noticeably. Now, segmenting the new input image with parameters identical to the original segmentation yields the result seen in Figure 4. The segments conform to the manual picks seen in Figure 2(b), while the rest of the image is segmented as accurately as the original result in Figure 2(a).

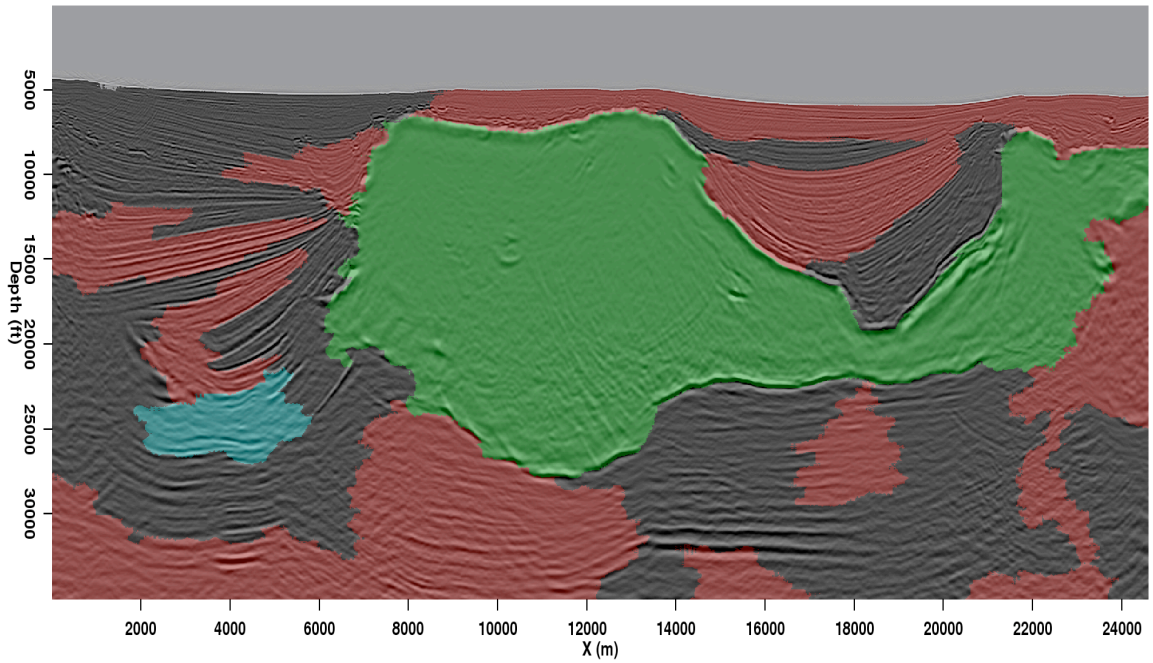


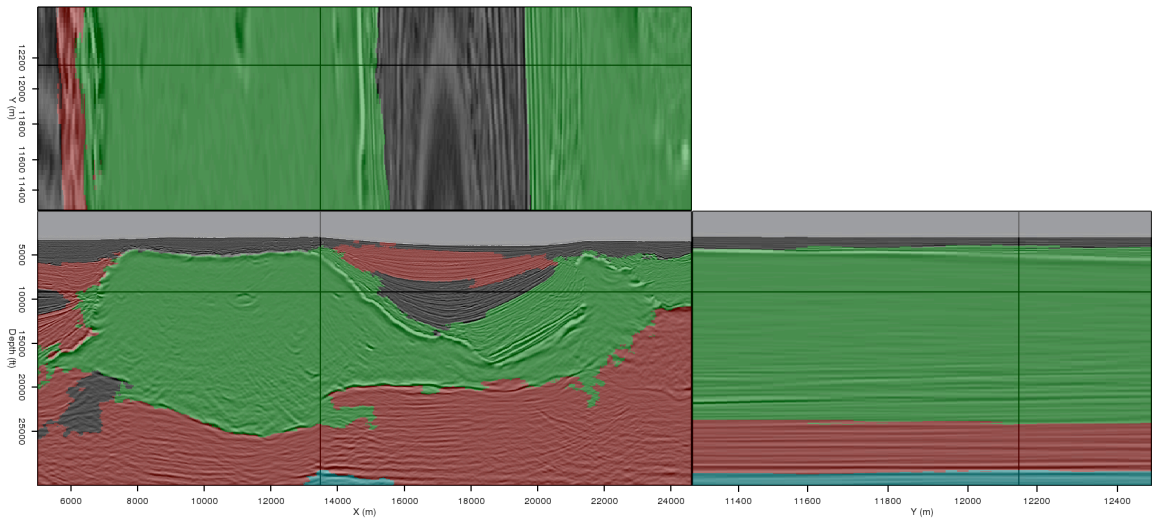
Figure 4: 2D interpreter-guided segmentation results. This result incorporates the manual picks seen in Figure 2(b). [**ER**] adam1/. 2d-merge-new

THREE DIMENSIONS

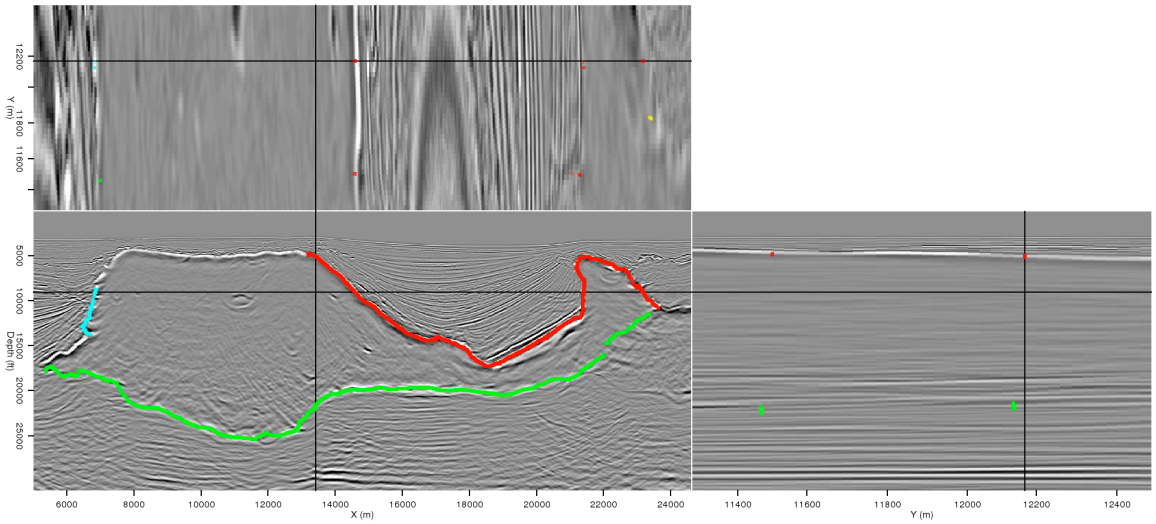
While the PRC algorithm suffers from the same leakage problem in 3D (Figure 5(a)) seen in the 2D example, the same solution cannot be applied. Because segments are much larger in 3D, amplitude changes on a single 2D section are not significant enough to alter 3D segmentation results. Instead, we must "project" an interpreter's manual picks on an inline section, like those seen in Figure 5(b), into the third (crossline) dimension. If we make the assumption that the dip of the salt flank does not fluctuate by more than two pixels per slice in the crossline direction, we can construct a square pyramid (Figure 6) in the crossline direction with sides of length $2h$, where h is the number of crossline samples between the base of the pyramid and its apex, which is the manually interpreted point P. Now, the new amplitude value for any pixel Q that falls within the pyramid is

$$A^Q_{\text{new}} = A^Q_{\text{orig}} + \frac{A_0}{||PQ||}, \tag{2}$$

where A_0 is the amplitude value at point P as determined by equation (1), and $||PQ||$ is the distance between the two points. As seen in Figure 7(b), the interpreter's picks now influence



(a)



(b)

Figure 5: (a) 3D segmentation result prior to any interpreter guidance. (b) Manually interpreted salt boundary picks. Manual picks from one or more sections are used to guide the automatic 3D segmentation. [**ER**] adam1/. 3d-origseg,oct-3d-picks

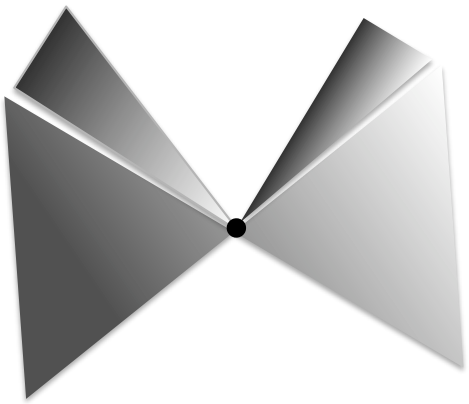
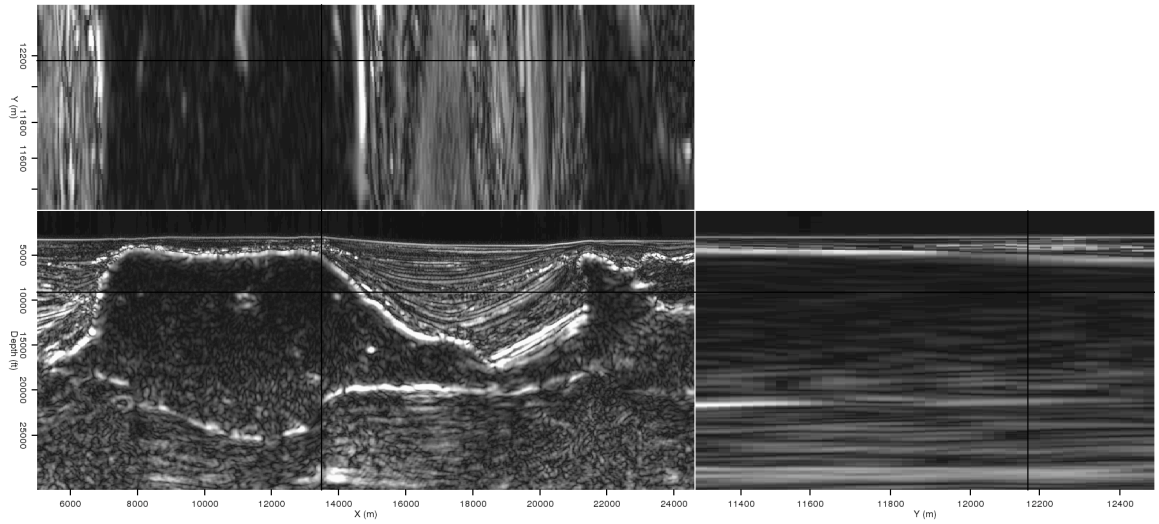
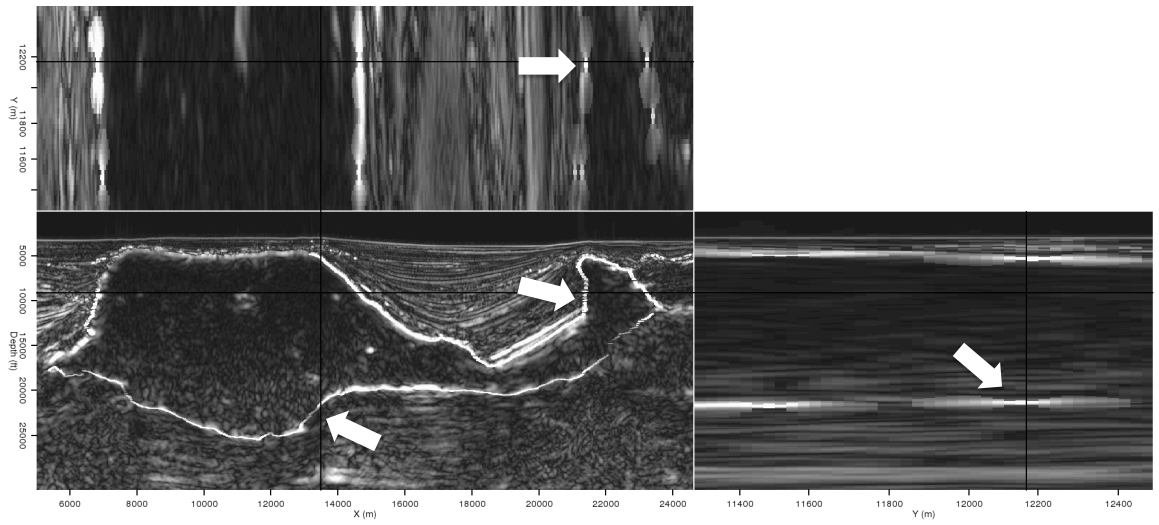


Figure 6: Schematic illustrating the concept of "spraying" manual interpretations of a 2D section into the third dimension. The black dot represents a pixel along a manually-picked horizon. [**NR**] adam1/. pyramid



(a)



(b)

Figure 7: 3D amplitude of the envelope calculations from (a)the original image, and (b) modified according to the manual picks seen in Figure 5(b). The 3D effects of the procedure to incorporate the picks are evident on the crossline and depth slices. [**ER**]

adam1/. 3d-env-orig,3d-env-new

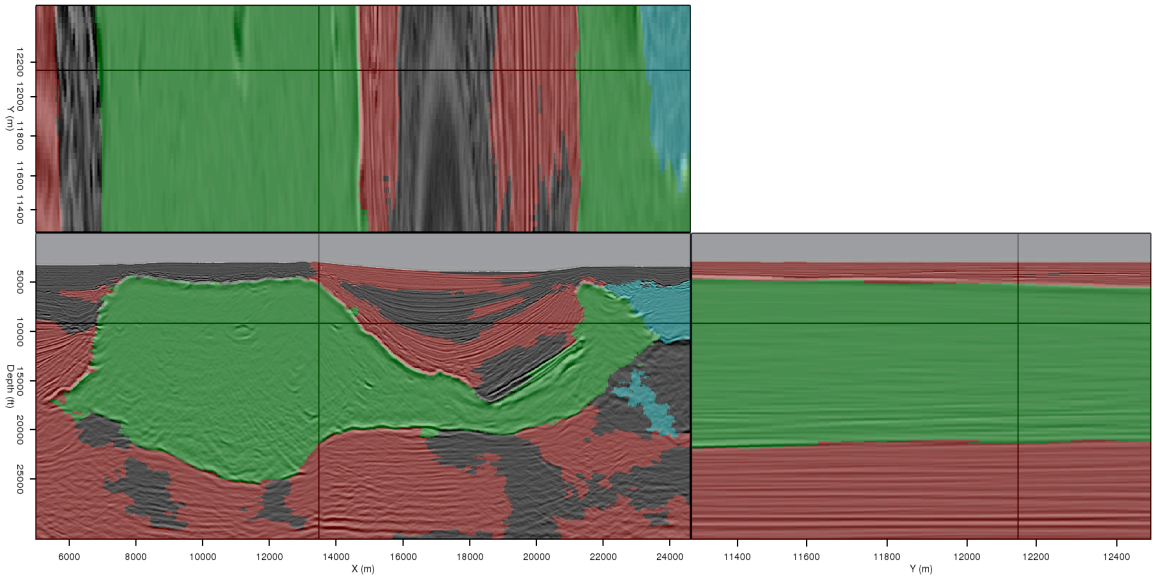


Figure 8: 3D interpreter-guided segmentation results, incorporating the picks seen in Figure 5(b). Note that in this result, the segmentation is more accurate in all three directions. [**ER**] adam1/. 3d-newseg

amplitude values in all three directions, but the magnitude of that influence decays with distance from the manual picks. Accordingly, the updated 3D segmentation result in Figure 8 is improved throughout the image cube, and not just on the two inline sections where manual picks were provided. For example, Figure 9 shows image slices far away from the two crosslines for which manual interpretations were supplied. In Figures 10(a) and 10(b), the effects of the 3D interpreter input procedure on the envelope volumes are apparent. Figures 11(a) and 11(b) compare the original and interpreter-guided segmentation results for this location, demonstrating that the 3D segmentation results can improve dramatically even far away from any manual pick locations.

CONCLUSIONS

The Pairwise Region Comparison algorithm offers a highly efficient means of automatically segmenting salt bodies in seismic images, but its accuracy can suffer in areas with discontinuous boundaries or poor illumination. In such cases, altering the input amplitude data according an interpreter's manual picks can successfully guide the segmentation to a more accurate result, even in areas far from the manual picking locations. Combining interpreter expertise with automation in this manner could help alleviate a major bottleneck in the model-building and imaging workflow.

ACKNOWLEDGMENTS

I thank WesternGeco for providing the field dataset used for examples, and all sponsors of the Stanford Exploration Project for their support.

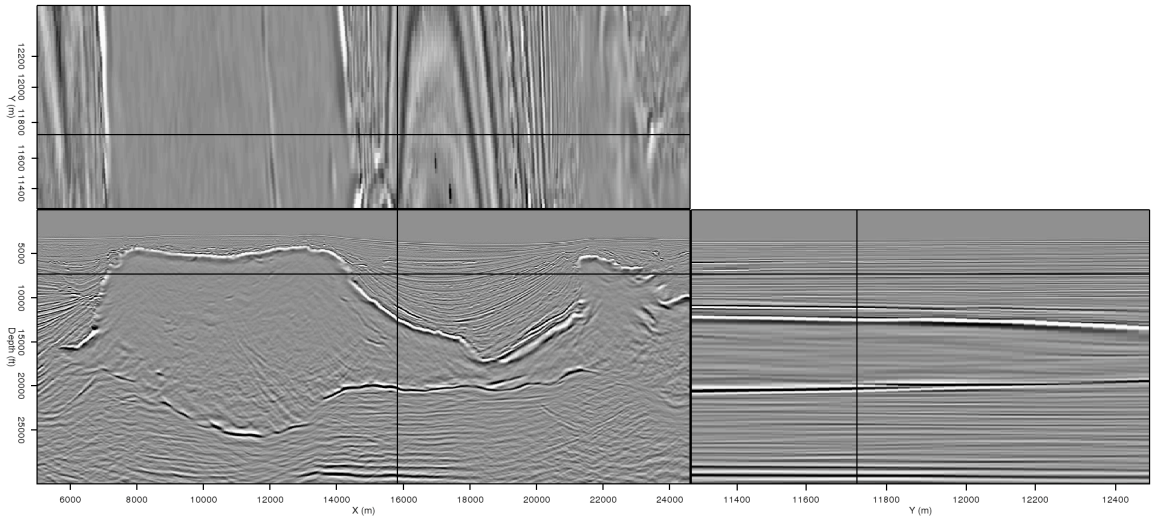


Figure 9: Slices from the 3D image cube far away from any interpreter-supplied salt picks. [**ER**] adam1/. o3d-far

REFERENCES

Felzenszwalb, P. F. and D. P. Huttenlocher, 2004, Efficient graph-based image segmentation: International Journal of Computer Vision, **59**, 167–181.

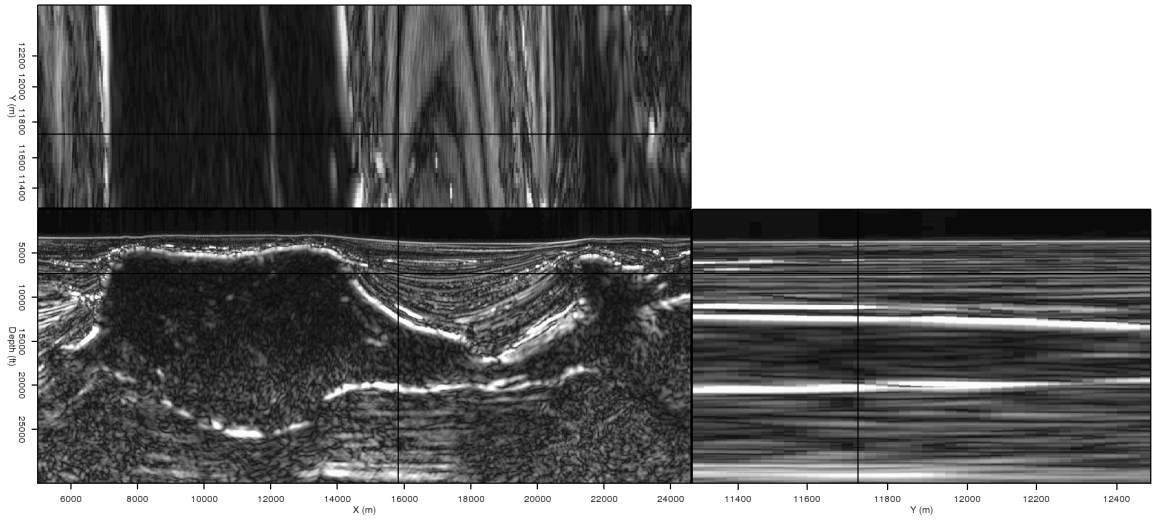
Halpert, A., 2010, A new method for more efficient seismic image segmentation: SEP-Report, **140**, 213–228.

Halpert, A., R. G. Clapp, and B. L. Biondi, 2010, Speeding up seismic image segmentation: SEG Technical Program Expanded Abstracts, **29**, 1276–1280.

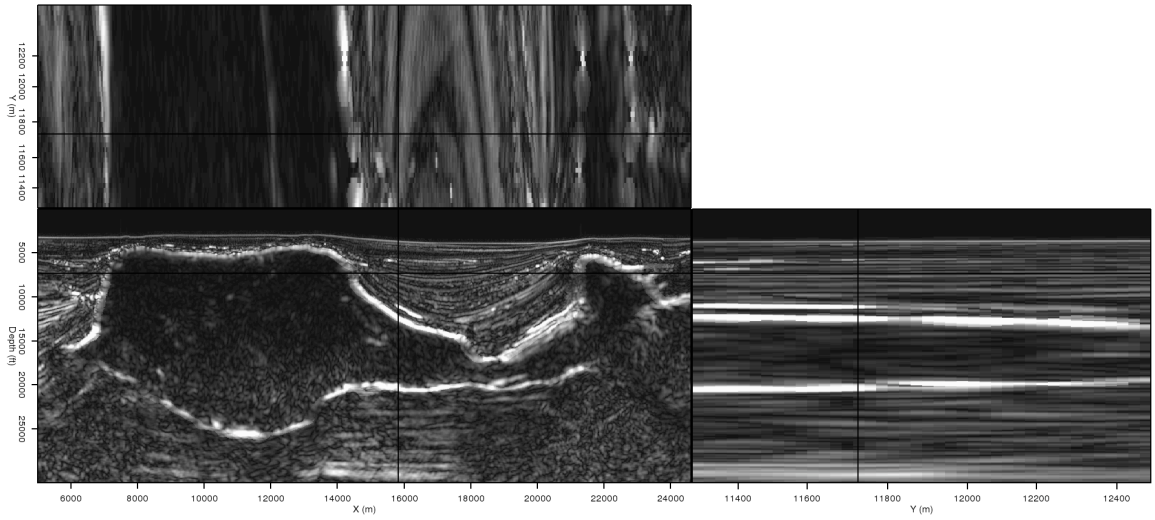
Lomask, J., R. G. Clapp, and B. Biondi, 2007, Application of image segmentation to tracking 3d salt boundaries: Geophysics, **72**, P47–P56.

Shi, J. and J. Malik, 2000, Normalized cuts and image segmentation: Institute of Electrical and Electronics Engineers Transactions on Pattern Analysis and Machine Intelligence, **22**, 838–905.

Zahn, C. T., 1971, Graph-theoretical methods for detecting and describing gestalt clusters: IEEE Transactions on Computers, **C-20**, 68–86.

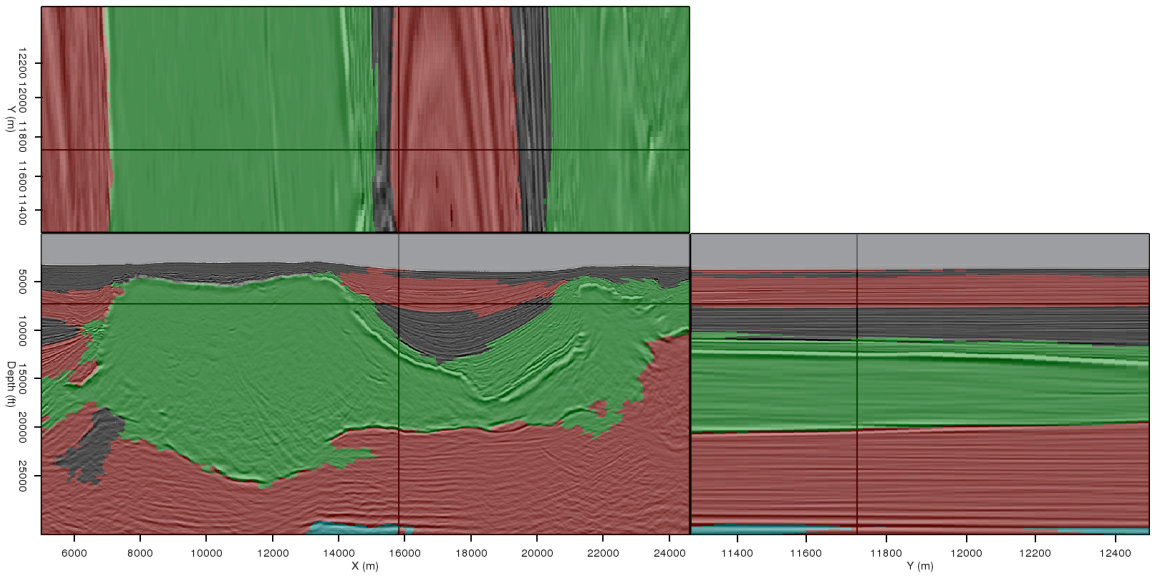


(a)

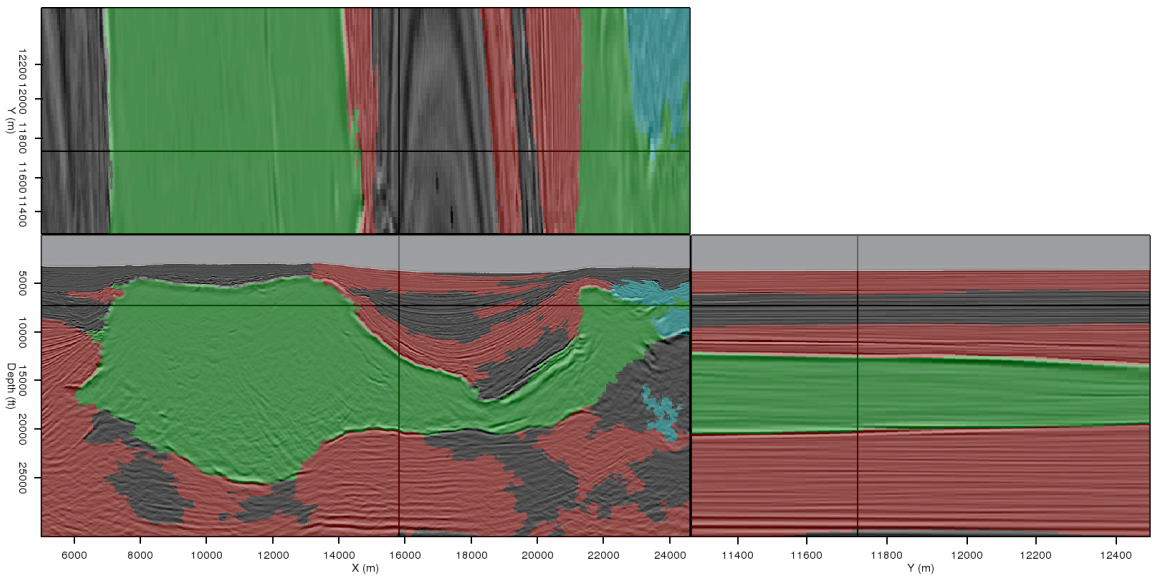


(b)

Figure 10: Envelope volumes at the position of the image slices seen in Figure 9, (a) before and (b) after incorporating interpreter input from a distant location. In (b), continuity of the salt boundary is improved, allowing a more accurate segmentation result. [**ER**] adam1/. o3d-env-orig-far,o3d-env-new-far



(a)



(b)

Figure 11: A comparison of segmentation results (a) without using interpreter input, and (b) after incorporating information from the picks in Figure 5(b). Even though the data slices shown here are far from the location of the manual picks, the strategy of spreading information from 2D picks into the third dimension allows for a much more accurate result in (b). [**ER**] adam1/. 3d-origseg-far,3d-newseg-far

Anisotropic tomography with rock physics constraints

Yunyue Li, Dave Nichols, Konstantin Osypov, and Ran Bachrach

ABSTRACT

Anisotropic model building is a challenging problem, well-known for its non-linear and underdetermined nature. To reduce the null-space and stabilize the inversion, we propose a new preconditioning scheme in linearized tomography to include rock physics prior information. We introduce the rock physics information in the form of covariance among P-wave vertical velocity (v_0), ϵ and δ, and is generated by stochastic realizations of a compacting shale model. We design a VSP synthetic survey with the common industry geometry on two different examples, of which one fulfills the assumption of our rock physics model and the other does not. The results show that by utilizing the proper rock physics prior information, tomography can better resolve the anisotropy parameters, especially in the area where inversion is poorly constrained by the data. However, precautions should be taken when the lithology of the subsurface is largely unknown. Finally, we perform a posterior uncertainty analysis to evaluate the contribution of the rock physics prior information. The results show that the null-space is greatly reduced by introducing the prior information.

INTRODUCTION

Anisotropic model building tries to resolve more than one parameter at each grid point of the subsurface. This number could be 3 for a vertical transverse isotropic (VTI) media, and increase to 5 for a tilted transverse isotropic (TTI) media. Traditional surface seismic tomography may be able to produce accurate isotropic earth models efficiently for a large area when the acquisition is dense and the earth is well-illuminated by rays at a wide range of angles. However surface seismic data inversion becomes ill-posed and highly underdetermined due to the rapidly increasing dimensionalities of the model space with the increasing complexity of the subsurface.

One big disadvantage of the surface seismic tomography is the lack of the depth information. During tomography, not only is the low wavenumber earth model estimated, but the depth of the reflectors is unknown as well. To add the depth dimension into the inversion, several localized tomography experiments around the wells are analyzed (Bakulin et al., 2010d,c). In these studies, joint inversion of surface seismic data and borehole data (check-shots, walkaway VSPs) shows great potential to yield better defined earth models. However, due to the ambiguity among the parameters, even the borehole aided localized tomography has difficulty in resolving a reliable, unique anisotropic model in 3D (Bakulin et al., 2009).

To constrain the inversion further, we need to consider some prior knowledge of the subsurface. This prior knowledge can be characterized by the covariance of the model space and is independent of the data. There are many ways to obtain the covariance information

based on different assumptions. For example, we often smooth our earth model horizontally and vertically, which implies a certain user-defined spatial correlation lag. More realistically, we can use the geological information as a prior and shape our estimate accordingly. This model shaping can be posed as a decomposition of the earth model into different layers and horizons before tomography (Bakulin et al., 2010a), or as a regularization/preconditioning operator during tomography (Bakulin et al., 2010b). We can obtain the geological information either by interpreting and picking the horizons or by building a set of steering filters (Clapp, 2000) according to the current subsurface image.

In addition to the spatial covariance, for a multi-parameter estimation, a point-by-point cross-parameter covariance is also needed to fully describe the subsurface. One source of the cross-parameter covariance comes from rock physics study (Hornby et al., 1995; Sayers, 2004, 2010; Bachrach, 2010b). In particular, Bachrach (2010a) develops both deterministic and stochastic modeling schemes based on the rock physics effective media models for compacting shale and sandy shale. Along with appropriate laboratory core measurements, the parameters needed by the rock physics model are limited in a certain range, which greatly reduces the correlation lag in the earth model parameters. These rock physics modeling results can be used to construct the initial earth model and the covariance relationships among the earth model parameters. When all of these four ingredients - surface seismic, borehole traveltime measurements, geological information and rock physics priors - are available, we will have a better chance to resolve anisotropic models that both flatten the gathers and follow the geological and rock physics principles at the same time.

In this paper, we assume the spatial covariance and the local cross-parameter covariance can be fully separated and focus on utilizing the rock physics modeling results to constrain the anisotropic tomography. A VSP survey with a common industry geometry is simulated on two different models, one with only shale (sandy shale), and the other with one layer of pure sand (isotropic). We compare the inversion results using the unconstrained tomography, current constrained tomography and the rock physics constrained tomography. Finally, we perform a-posteriori uncertainty analysis (Osypov et al., 2008) for the shale example to evaluate the contribution of rock physics prior knowledge to the reduction of the null-space.

THEORY

Here, we define a VTI tomography in using event traveltime as data. To the first order, when we assume the ray path does not change with perturbation in the background velocity model, the traveltime misfit is related to the velocity perturbation as follows:

$$\Delta t = \int_{\text{raypath}} \frac{1}{v^2} \Delta v \, \mathrm{d}\, l. \tag{1}$$

For weak anisotropy, the angle dependent velocity in VTI media can be described by vertical velocity v_0 and the Thomson anisotropic parameters ϵ and δ as follows (Tsvankin and Thomsen, 1994):

$$v(\theta) = v_0(1 + 2\delta \sin^2\theta \cos^2\theta + 2\epsilon \sin^4\theta), \tag{2}$$

where θ is the angle between propagation direction and the vertical direction. The angles are defined by the raypath in equation 1. Consequently, the velocity perturbation can be

expressed as follows:

$$\Delta v(\theta) = \frac{\partial v(\theta)}{\partial v_0}\Delta v_0 + \frac{\partial v(\theta)}{\partial \epsilon}\Delta\epsilon + \frac{\partial v(\theta)}{\partial \delta}\Delta\delta. \tag{3}$$

Now we can see explicitly that the traveltime misfit has three contributions from v_0, ϵ and δ, respectively:

$$\Delta t = \int \frac{1}{v^2(\theta)}\frac{\partial v(\theta)}{\partial v_0}\Delta v_0 \,\mathrm{d}\, l + \int \frac{1}{v^2(\theta)}\frac{\partial v(\theta)}{\partial \epsilon}\Delta\epsilon \,\mathrm{d}\, l + \int \frac{1}{v^2(\theta)}\frac{\partial v(\theta)}{\partial \delta}\Delta\delta \,\mathrm{d}\, l. \tag{4}$$

In practice, we usually formulate the seismic tomography problem as an inversion problem with another model space regularization term. Then the objective function reads:

$$S(\Delta m) = ||\mathbf{T}\Delta m - \Delta t||_D^2 + ||m_0 + \Delta m - m_{\text{prior}}||_M^2, \tag{5}$$

where D denotes the data space whose covariance is $\mathbf{C_D}$ and M denotes the model space whose covariance is $\mathbf{C_M}$. In the data space fitting goal, operator $\mathbf{T}$ is a row operator vector $[\mathbf{T}_{v_0}\ \mathbf{T}_\epsilon\ \mathbf{T}_\delta]$, each element corresponding to one of the three terms in equation 4 respectively; Δm is a column model vector $[\Delta v_0\ \Delta\epsilon\ \Delta\delta]'$; Δt is the traveltime misfit. In the model space fitting goal, m_0 is the background (initial) model $[v_{0_0}, \epsilon_0, \delta_0]'$; m_{prior} is the mean of the prior distribution in the model space $[v_{0p}, \epsilon_p, \delta_p]'$. We evaluate both fitting goals in L2 using an LSQR solver.

We can also obtain the normal equation representation of the objective function by taking the derivative of equation 5 with respect to Δm:

$$\begin{vmatrix} \mathbf{C_D^{-1/2}T} \\ \lambda\ \mathbf{C_M^{-1/2}} \end{vmatrix} \Delta m = \begin{vmatrix} \mathbf{C_D^{-1/2}}\Delta t \\ \lambda\ \mathbf{C_M^{-1/2}}(m_{\text{prior}} - m_0), \end{vmatrix} \tag{6}$$

where λ is the balancing factor between the data fitting equation and the model fitting equation. For the synthetic study here, we may assume $\mathbf{C_D} = \mathbf{I}$, since there is no noise in these data.

To speed up the convergence, we turn this regularized problem into a preconditioned problem by doing a variable substitution:

$$\begin{vmatrix} \mathbf{TC_M^{1/2}} \\ \lambda\ \mathbf{I} \end{vmatrix} \Delta m' = \begin{vmatrix} \Delta t \\ \lambda\ \mathbf{C_M^{-1/2}}(m_{\text{prior}} - m_0) \end{vmatrix}. \tag{7}$$

Now LSQR inverts the preconditioning variable $\Delta m'$, and the real model updates are obtained by $\Delta m = \mathbf{C_M^{1/2}}\Delta m'$ after inversion.

For a multi-parameter model inversion, we assume the model covariance can be fully decoupled into two different parts: spatial covariance for each parameter, and the local cross-parameter covariance. Mathematically, that translates into:

$$\mathbf{C_M^{1/2}} = \mathbf{S}\sigma = \begin{vmatrix} \mathbf{S_v} & 0 & 0 \\ 0 & \mathbf{S}_\epsilon & 0 \\ 0 & 0 & \mathbf{S}_\delta \end{vmatrix} \begin{vmatrix} \sigma_{vv} & \sigma_{v\epsilon} & \sigma_{v\delta} \\ \sigma_{\epsilon v} & \sigma_{\epsilon\epsilon} & \sigma_{\epsilon\delta} \\ \sigma_{\delta v} & \sigma_{\delta\epsilon} & \sigma_{\delta\delta} \end{vmatrix}, \tag{8}$$

where $\mathbf{S}$ is a spatial smoothing matrix, and σ is the square-root of the point-by-point cross-parameter covariance matrix. Each element of $\mathbf{S}$ can be estimated according to the

user's assumption for the smoothness of different parameters. For example, the steering filters (Fomel, 1994; Clapp et al., 2004) provide a good choice to incorporate the structural information of the subsurface. The local cross-parameter covariance can be estimated from the rock physics modeling (Bachrach, 2010a).

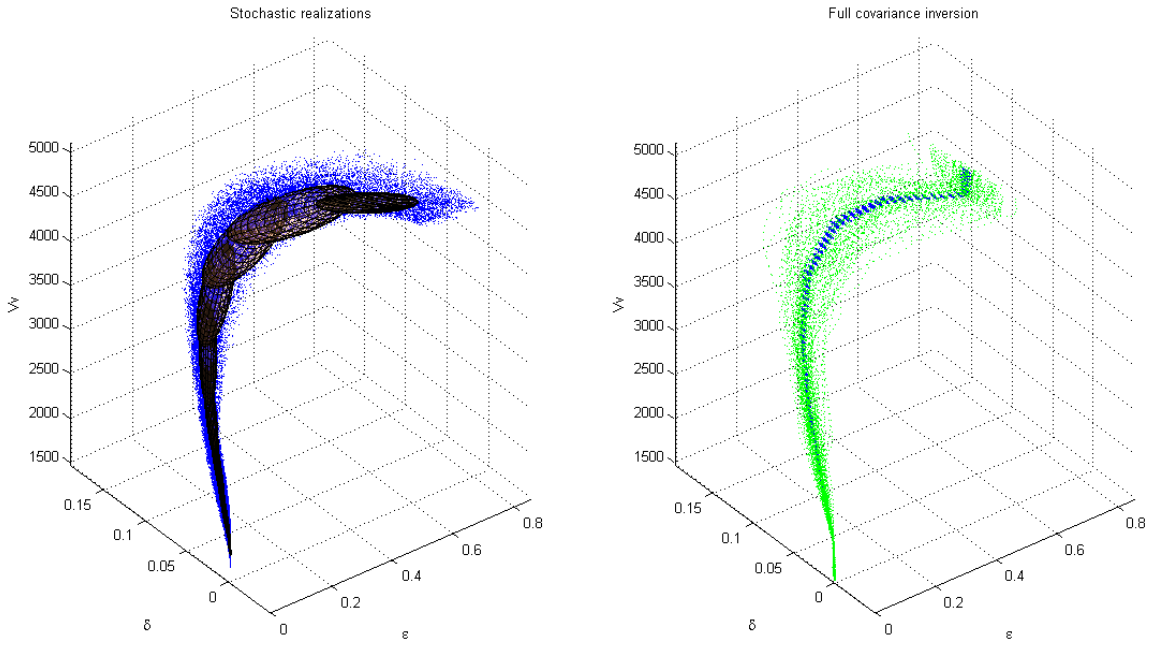


Figure 1: Left: Stochastic realizations of a rock physics modeling for shale and sandy shale. The ellipsoids are fitted locally to describe the multi-Gaussian relationship among the parameters. Right: Small dots in the background are the estimated prior distribution by the operator using the ellipsoids on the left. Larger dots are the null-space projection of the operator. [**NR**] elita2/. covariance

The left panel on Figure 1 shows an example of the stochastic realizations of a rock physics model for shale and sandy shale. The dots are scattered v_0, ϵ and δ results of hundreds of realizations, while the nine ellipsoids are fitted locally with v_0 the controlling variable to describe the multi-Gaussian relationship among the parameters. With a statistical study of the laboratory measurements on the cores, the range of the required input parameters for the rock physics model has been limited to a relatively narrow range, which leads to tight ellipsoids in Figure 1. We refer interested readers to details in Bachrach (2010a).

In this paper, we build the cross-parameter covariance matrix for each grid point in the subsurface according to the velocity, and linearly interpolate between the ellipsoids to which the velocity value belongs. Notice that additional non-linearity has been introduced during this process. Better methods to utilize the stochastic realizations could fit ellipsoids centered at every velocity value, which can be more precise but outside the scope of this study.

NUMERICAL TESTS

We design a walkaway VSP survey on a 2D study area shown in Figure 2. There are 51 sources distributed evenly along a 5 km line on the surface with 2.5 km maximum offset, and 10 receivers fixed every 1 km down the borehole. This acquisition geometry is designed to be similar to the industry standard VSP surveys, which have good constraints for vertical velocity and δ, but fewer constraints for ϵ in the deeper part due to the limited propagation angles.

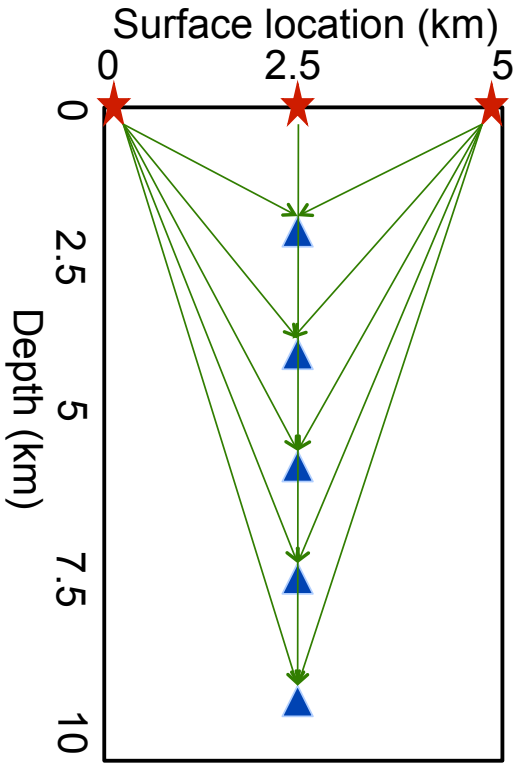


Figure 2: Walkaway VSP acquisition geometry. [**NR**] elita2/. AcquiGeo

Two 1.5D models are evaluated using this method. One is a shale (sandy shale) model which completely follows the covariance matrix we generated from the rock physics modeling; the other is the same except for a layer of isotropic sand where the prior information is "wrong".

Now we are ready to test our method using different prior information. For different tests, we apply the same smoothing operator $\mathbf{S}$, but different estimates of point-by-point cross-parameter covariance σ. In the notations below, "no prior" means $\sigma = \mathbf{I}$; "column weighting" means σ has only diagonal elements which are constant for every grid point in the subsurface; "diagonal covariance" means σ has only diagonal elements which vary according to the rock physics modeling results; "full covariance" means σ has all nine elements which vary according to the rock physics prior for each grid point in the subsurface.

Figure 3 shows the inversion results of the shale (sandy shale) model. It is obvious that vertical velocity is the best constrained variable, therefore, all inversion schemes yield good estimations for vertical velocity. However, instability is seen in the results of ϵ and δ

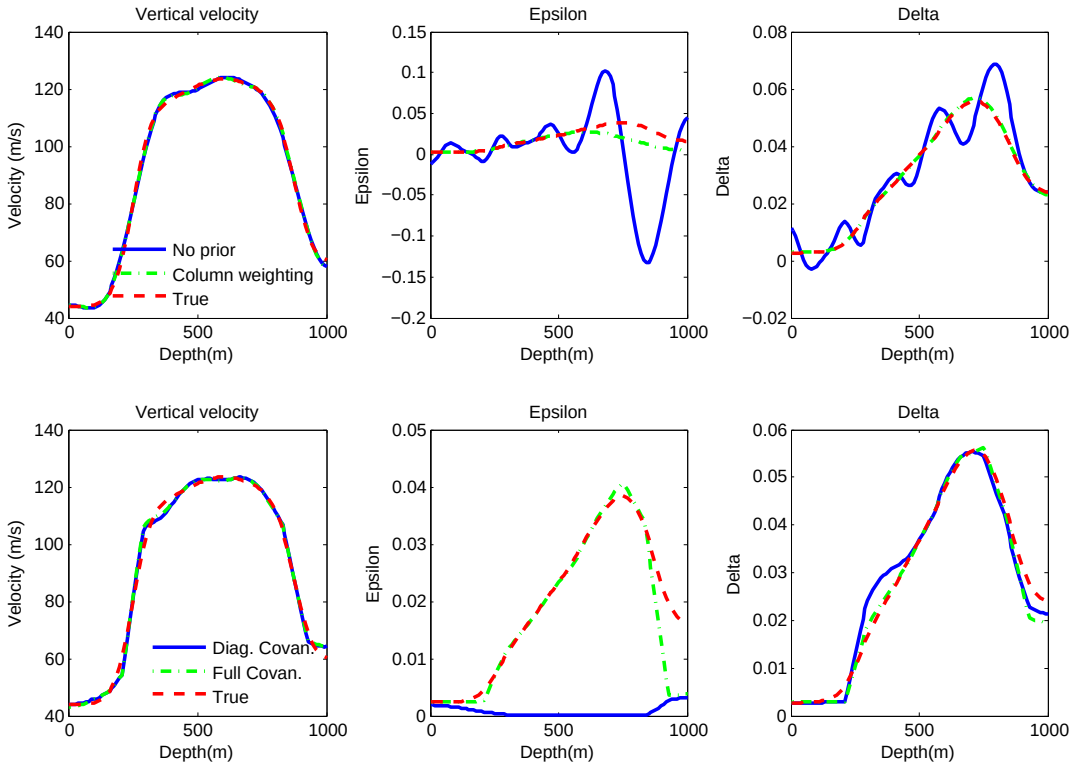


Figure 3: Inversion results of the shale (sandy shale) model. Panels on the left show the velocity perturbation, in the middle ϵ perturbation, and on the right δ perturbation. The top row shows the inversion results without rock physics prior knowledge: Solid line: No prior; Dashed dot line: Column weighting; Dashed line: True model. The bottom row shows the inversion results with some rock physics knowledge: Solid line: Diagonal covariance; Dashed dot line: Full covariance; Dashed line: True model. [**NR**] elita2/. model1

when no prior information is included. The oscillations in ϵ and δ are out-of-phase, which is the numerical proof for the theoretical predicted trade-off between these two parameters. Inversion with column weighting yields more stable results for δ and the shallow part of ϵ. For the deeper part and also less constrained part of ϵ, column weighting gives a less satisfactory result. When rock physics prior knowledge is introduced, the inversion is further stabilized. Both the diagonal and the full covariance give good estimations for velocity and δ, while superior result for ϵ is obtained by full covariance since correct prior knowledge adds information to the inversion. The fact that a much closer estimation for ϵ was produced using full covariance rather than the diagonal one suggests that large cross-terms exist in the covariance matrix.

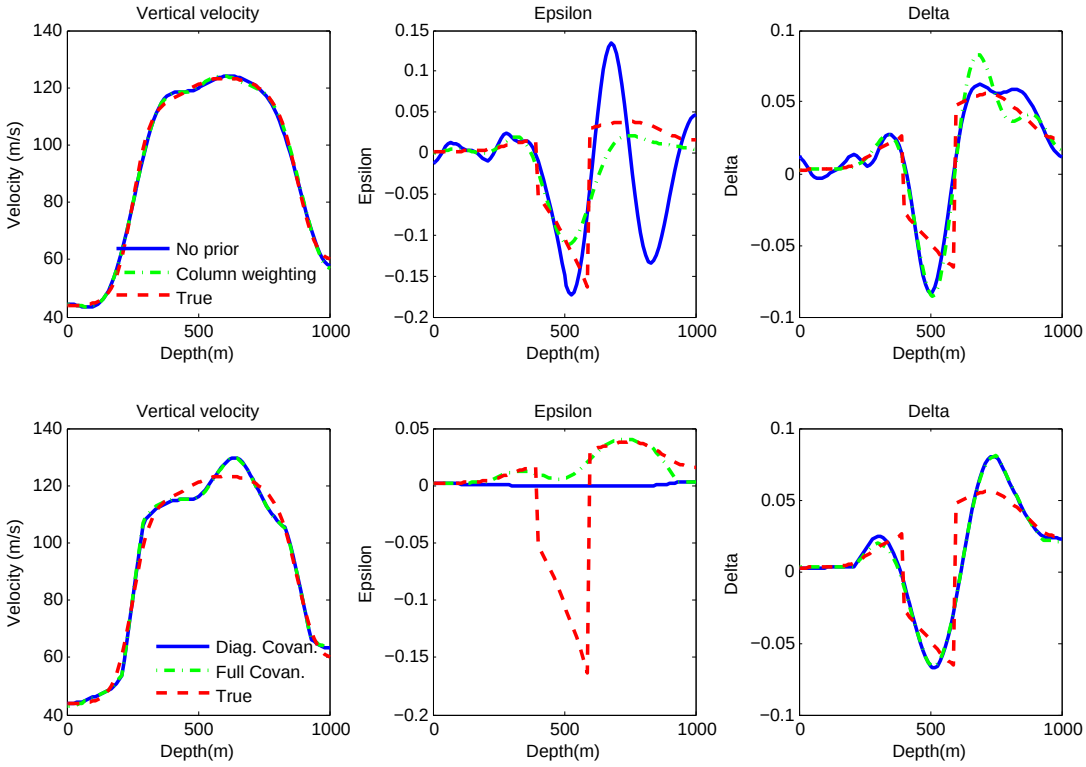


Figure 4: Inversion results of the shale model with an isotropic sand layer. Panels are arranged in the same order as Figure 3. [**NR**] elita2/. model2

Figure 4 shows the inversion results of the shale (sandy shale) model with an isotropic layer in the middle. Similar stability conclusions can be drawn as for the shale (sandy shale) case. Notice that for the well-constrained variable δ, inversion is able to resolve the isotropic layer even though "wrong" prior information is provided. However, the inversion result is highly biased towards the prior information for ϵ where it is not so well constrained.

POSTERIOR UNCERTAINTY ANALYSIS

For a regularized problem, an L-curve analysis is often useful to determine the damping parameter λ in equation 7 and investigate the posterior distribution of the inversion (Hansen and O'Leary, 1993). Typical L-curve has two distinct parts: one vertical part where the

solution is dominated by the data fitting and one horizontal part where the solution is dominated by the model styling. The corner of the L-curve corresponds to a good balance between minimization of both fitting goals. In Figure 5, we obtain the L-curve for each inversion scheme for the shale (sandy shale) model in log-log scale by varying λ from $1e-11$ to 0.1. The model residual is defined in the preconditioning space. Therefore, the shape of the L-curve also depends on the covariance matrix. A good estimation of the covariance and a proper λ will place the solution right at the corner of the L-curve, as in the case of using the full covariance. The "7-curve" shape in the log-log scale (which is still an "L-curve" in absolute scale) shows a relatively poor estimation of the covariance matrix, hence indicating difficulties in finding a proper damping parameter λ.

Finally, to move beyond the deterministic inversion which produces only one solution, we perform the null-space analysis following the work flow proposed by Osypov et al. (2008). The effective null-space of an operator can be sampled by an iterative Lanczos eigen-decomposition method. The right panel on Figure 1 shows the null-space projection (darker dots) overlaid on the approximate prior distribution (smaller dots) when the full covariance scheme is used. It is obvious that the full covariance matrix representation produces a good estimation to the true prior (by the similarity of the cloud shape on the left panel and the right panel). Also, the null-space projection suggests that higher uncertainty in anisotropic parameters for higher velocities, which often means greater depth, is embedded in rock physics knowledge. The reduced volume of the cloud shows the value of information that the data bring into the inversion.

CONCLUSIONS

In this paper, we have proposed a new formulation to incorporate rock physics prior information with the anisotropic tomography. Two models were analyzed using this method, and the inversion results demonstrate the trade-off among the parameters and the instability due to the huge null-space when no prior information is included. Any estimation of the local cross-parameter distribution (column weighting, diagonal covariance and full covariance) is helpful to stabilize the inversion and leads to a better representation of the subsurface. However, we should be careful in using too tight of a prior distribution when the lithology is uncertain, especially for areas where parameters are not well-constrained by the data. The posterior distribution analysis shows that by adding the rock physics prior information, we will obtain a better estimation of the true prior statistics in the inversion and a smaller uncertainty in the posterior statistics.

The experiment of rock physics constrained tomography suggests to us a new workflow in anisotropic model building.

- First, Build an initial model using the deterministic rock physics modeling and obtain the initial image.
- Second, Build the point-by-point cross-parameter covariance according to the stochastic rock physics modeling. Build the spatial covariance using geology information and/or the initial image.
- Third, run the rock physics constrained joint tomography with surface seismic data and borehole data.

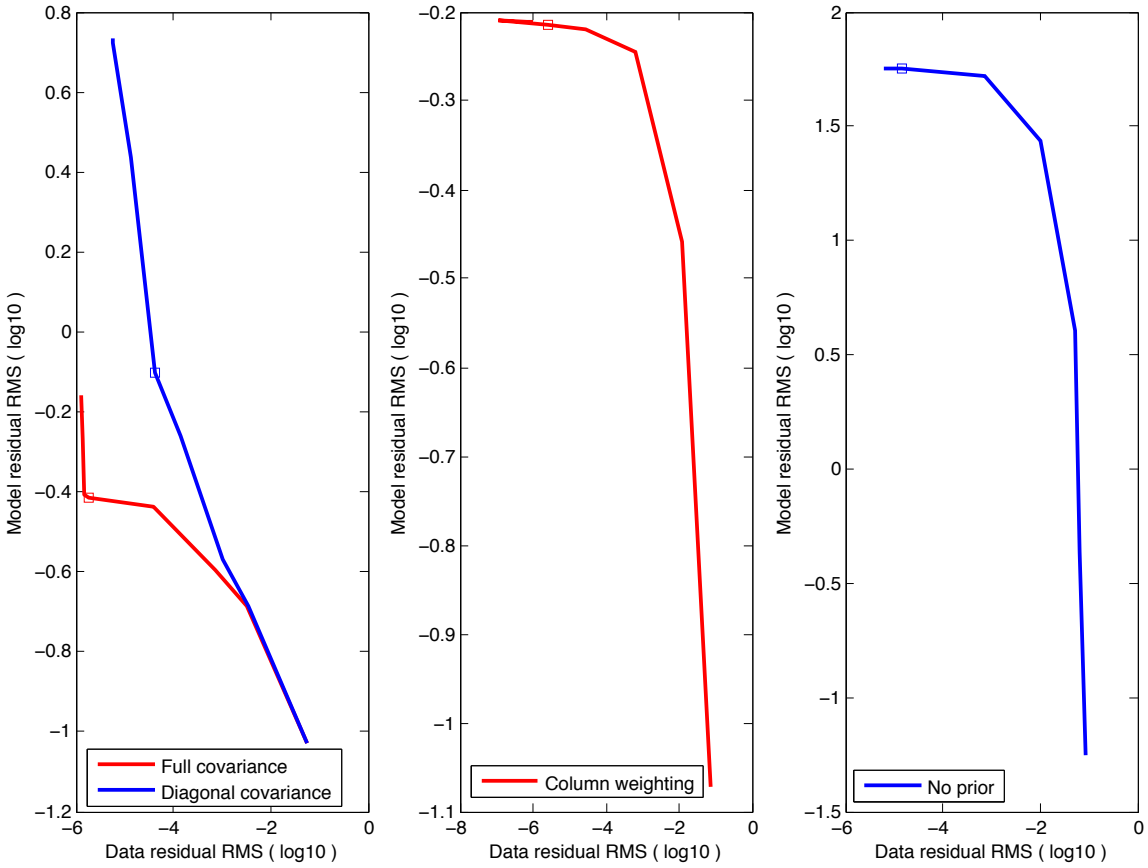


Figure 5: L-curve for the shale (sandy shale) model inversion. [**NR**] elita2/. L-curve

Repeat the workflow if necessary. Up to now, it is possible to use all the information: surface reflection seismic, borehole data, geological estimation and the rock physics covariance in the tomography to produce a unique earth model that explains the seismic data and satisfies the geological and rock physics theory at the same time.

ACKNOWLEDGEMENT

This work was completed during my summer internship with WesternGeco in 2010. The results were produced based on the C++ library in WesterGeco, and hence NR to non-WesternGeco sponsors. The authors thank WesternGeco for the permission to publish this work. Thanks to Marta Woodward, Can Evren Yarman, Yangjun Liu, Xiang Xiao, Feng Qiao for helpful discussions.

REFERENCES

Bachrach, R., 2010a, Applications of deterministic and stochastic rock physics modeling to anisotropic velocity model building: SEG Expanded Abstracts, **29**, 2436–2440.

———, 2010b, Elastic and resistivity anisotropy of compacting shale: Joint effective medium modeling and field observations: SEG Expanded Abstracts, **29**, 2580–2584.

Bakulin, A., Y. K. Liu, O. Zdraveva, and K. Lyons, 2010a, Anisotropic model building with wells and horizons: Gulf of mexico case study comparing different approaches: The Leading Edge, **29**, 1450–1460.

Bakulin, A., M. Woodward, Y. Liu, O. Zdraveva, D. Nichols, and K. Osypov, 2010b, Application of steering filters to localized anisotropic tomography with well data: SEG Expanded Abstracts, **29**.

Bakulin, A., M. Woodward, D. Nichols, K. Osypov, and O. Zdraveva, 2009, Can we distinguish TTI and VTI media?: SEG Expanded Abstracts, **28**, 226–230.

———, 2010c, Building tilted transversely isotropic depth models using localized anisotropic tomography with well information: Geophysics, **75**, 27–36.

———, 2010d, Localized anisotropic tomography with well information in VTI media: Geophysics, **75**, 37–45.

Clapp, R., 2000, Geologically constrained migration velocity analysis: PhD thesis, Stanford University.

Clapp, R., B. Biondi, and J. Claerbout, 2004, Incorporating geologic information into reflection tomography: Geophysics, **69**, 533–546.

Fomel, S., 1994, On model-space and data-space regularization: A tutorial: Stanford Exploration Project Report, **94**, 141–160.

Hansen, P. C. and D. P. O'Leary, 1993, The use of the the L-curve in the regularization of discrete ill-posed problems: SIAM J. Sci. Comput., **14**, 1487–1503.

Hornby, B., D. Miller, C. Esmersoy, and P. Christie, 1995, Ultrasonic-to-seismic measurements of shale anisotropy in a North Sea well: SEG Expanded Abstracts, **14**, 17–21.

Osypov, K., D. Nichols, M. Woodward, O. Zdraveva, and C. E. Yarman, 2008, Uncertainty and resolution analysis for anisotropy tomography using iterative eigendecomposition: SEG Expanded Abstracts, **27**, 3244–3249.

Sayers, C., 2004, Seismic anisotropy of shales: What determines the sign of Thomsen's delta parameter?: SEG Expanded Abstracts, **23**, 103–106.

———, 2010, The effect of anisotropy on the Young's moduli and Poisson's ratios of shales: SEG Expanded Abstracts, **29**, 2606–2611.

Tsvankin, I. and L. Thomsen, 1994, Nonhyperbolic reflection moveout in anisotropic media: Geophysics, **59**, 1290–1304.

Model-building with image segmentation and fast image updates

Adam Halpert and Gboyega Ayeni

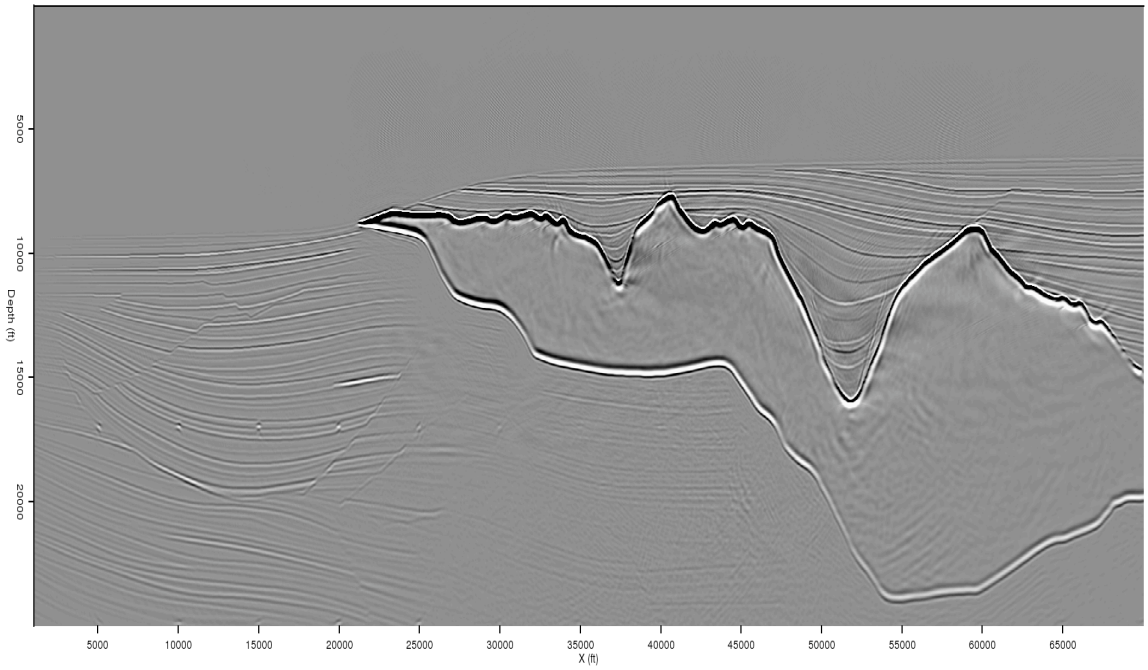
ABSTRACT

An accurate salt interpretation is an essential component of velocity model-building in areas dominated by complicated salt geology. The Pairwise Region Comparison (PRC) image segmentation algorithm can automatically pick salt bodies on seismic images, and be used as part of an iterative sediment-flood and salt-flood model-building workflow. In areas where the salt interpretation is highly uncertain, however, human expertise is needed to judge the relative accuracy of two or more possible models. We demonstrate a fast image updating scheme based on shot-profile migration that can be used to investigate how different salt interpretations provided by the PRC algorithm affect the final image. With further efficiency improvements, this process could allow for an interactive interpretation, model-building, and imaging workflow that substantially reduces cycle time for large-scale iterative imaging projects.

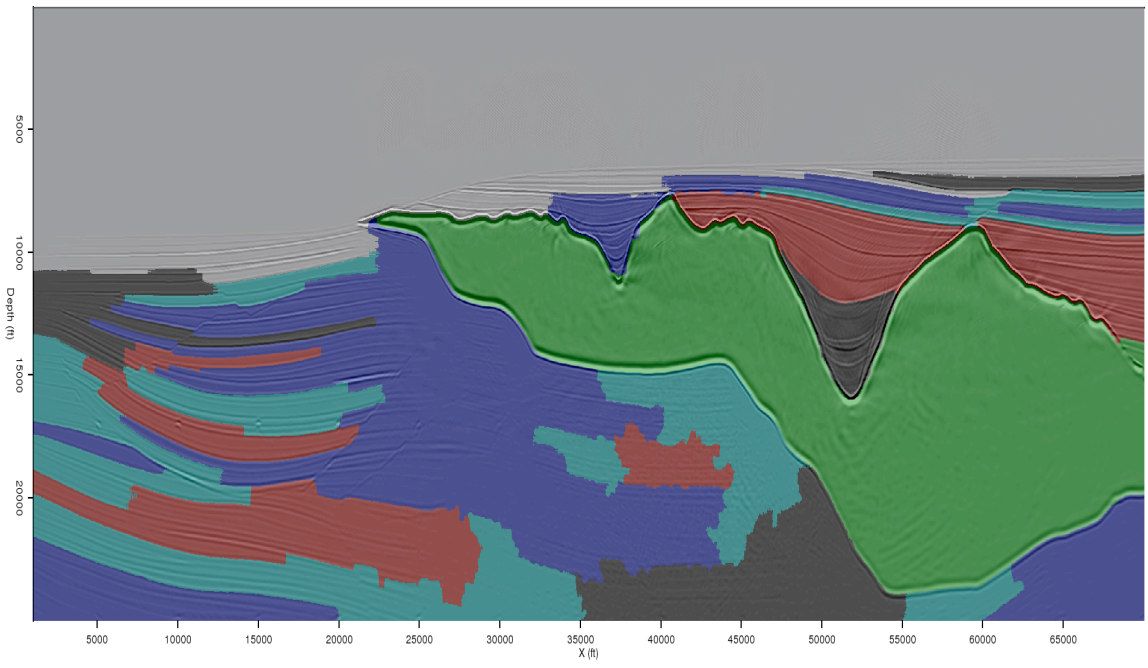
BACKGROUND

The goal of any automated seismic image segmentation scheme is to automatically pick out or delineate regions within a seismic image. Salt interpretation is an especially useful application of this technology; large, subsurface salt bodies can be tedious and very time-consuming to interpret manually, but the sharp salt/sediment velocity contrast makes an accurate interpretation essential for building a satisfactory velocity model. This means that we should not rely solely on an automatic salt interpretation, but should instead draw on the expertise of experienced human interpreters to supplement the process. This can be accomplished in a variety of ways; for example, a manual 2D interpretation can be used to guide an automatic 3D segmentation (Halpert, 2011). A complementary option is to use human expertise to choose the best of several possible interpretations provided by an automatic algorithm. Here, we demonstrate one such strategy using fast image updates to examine differences between possible models.

The "Pairwise Region Comparison" (PRC) seismic image segmentation approach (Halpert et al., 2010; Halpert, 2010) relies on the graph-based segmentation method of Felzenszwalb and Huttenlocher (2004). This method is more efficient than eigenvector-based algorithms such as Normalized Cuts Image Segmentation (Shi and Malik, 2000; Lomask, 2007), and is capable of operating on full seismic images. Here, the PRC method is used on images from the Sigsbee 2A synthetic model to provide two possible top-salt interpretations. Then, a fast image-update scheme based on shot-profile migration is used to examine how an image changes as a result of the two interpretations.



(a)



(b)

Figure 1: Perfect-velocity image (a) and automatic segmentation result (b) from the Sigsbee synthetic model. In (b), the salt body has been interpreted nearly perfectly. [**ER**] adam2/. zig-perf,seg-perf

SEGMENTATION PROCESS

Under ideal cirmucmstances, automated segmentation using the PRC method can provide a quick and accurate interpretation of regions within a seismic image. For example, the image in Figure 1(a) is a perfect-velocity migration of data from the Sigsbee synthetic model. In the corresponding segmentation result (Figure 1(b)), the salt body is delineated clearly and accurately. Unfortunately, perfect velocity models are non-existent in the real world. Instead, a common approach to building a velocity model in areas dominated by complex salt geology is to pick the top salt after migration with a sediment-flood velocity model, and then pick the base of salt after migration with the resulting salt-flood velocity model (Mosher et al., 2007). It is clear from Figure 2(a), a sediment-flood migration zoomed in to the left half of the salt body, that the salt boundary of a sediment-flood image is not likely to be imaged as well as the perfect-velocity example in Figure 1(a). The resulting segmentation of the image into above- and below-salt regions (Figure 2(b)) is therefore not as accurate.

Since the major source of inaccuracy in the Figure 2(b) result is within the salt canyon indicated on the figure, we can restrict our analysis to that region. By setting a smaller minimum segment size, the user can be much more selective about what regions to include as salt. Figures 3(a) and 3(b) display two possible salt interpretations generated by the automatic PRC algorithm for the region in question. In Figure 3(b), an extra segment (indicated by the arrow) has been included in the salt. By replacing segments interpreted as salt with salt velocities on the original sediment flood velocity model, two different velocity models can also be created (Figures 4(a) and 4(b)).

IMAGE UPDATES

Image updates and comparisons will be performed using *shot-profile migration* (see Biondi (2005)). In general, this type of migration uses downward-continued source (P^s) and receiver (P^g) wavefields:

$$P^s_z = P^s_{z=0} e^{-ik_z z} \tag{1}$$

$$P^g_z = P^g_{z=0} e^{ik_z z}, \tag{2}$$

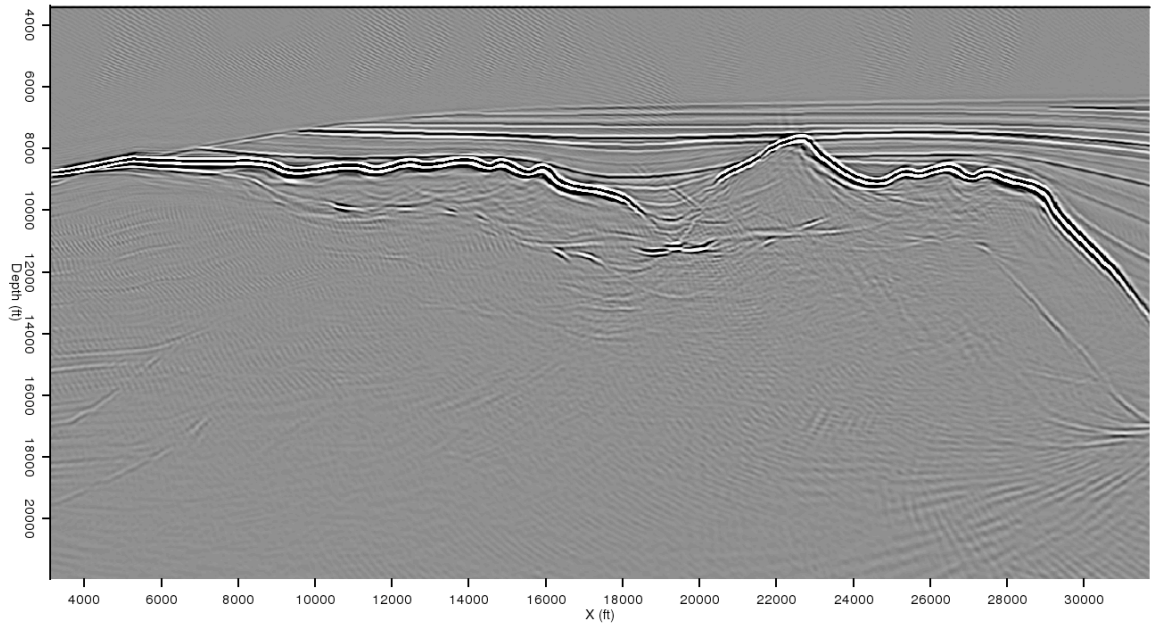
where z is the depth and k_z is the vertical wavenumber calculated in a split-step fashion (Stoffa et al., 1990) for a given frequency ω according to

$$k_z = \sqrt{\frac{\omega^2}{v^2_{\text{ref}}} - |\mathbf{k}|^2} + \left(\frac{\omega}{v(z,x)} - \frac{\omega}{v_{\text{ref}}} \right), \tag{3}$$

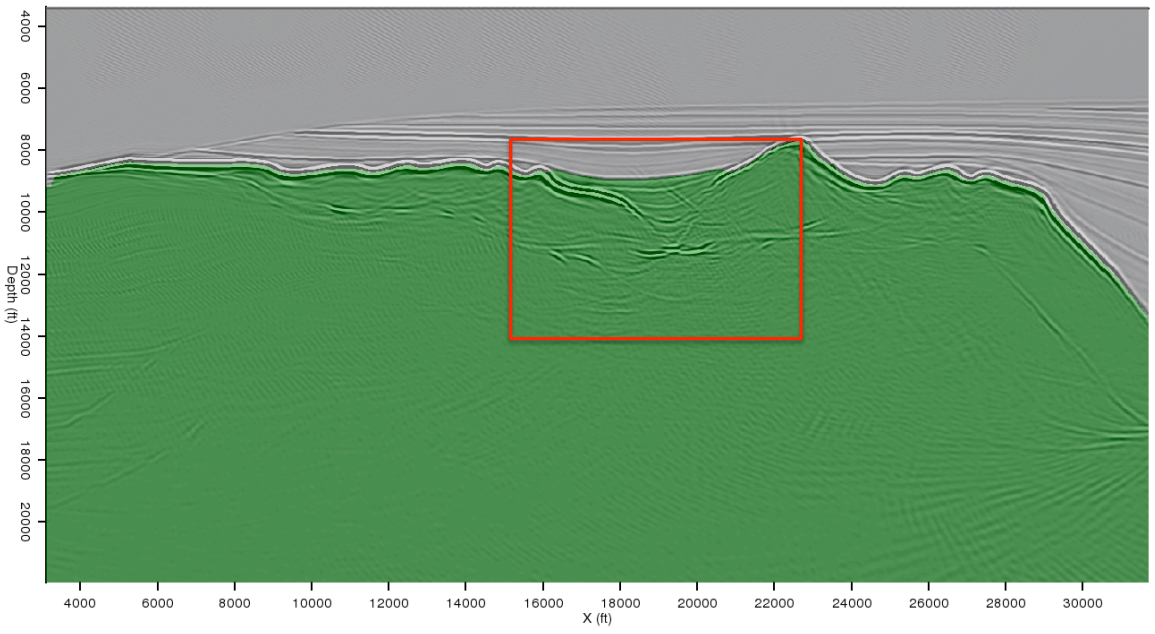
where $\mathbf{k}$ is a vector containing the horizontal wavenumbers. Here, v_{ref} is a reference velocity that is constant at each depth step, while $v(z,x)$ is the actual estimated velocity. An image is formed by correlating the two wavefields:

$$I(z,x) = \sum_i \sum_\omega P^g_z(\omega, x; \mathbf{s}_i) \overline{P^s_z(\omega, x; \mathbf{s}_i)}. \tag{4}$$

Performing full migrations is impractical for our stated purpose of allowing interpreters to quickly judge the relative accuracy of two or more possible velocity models. One way

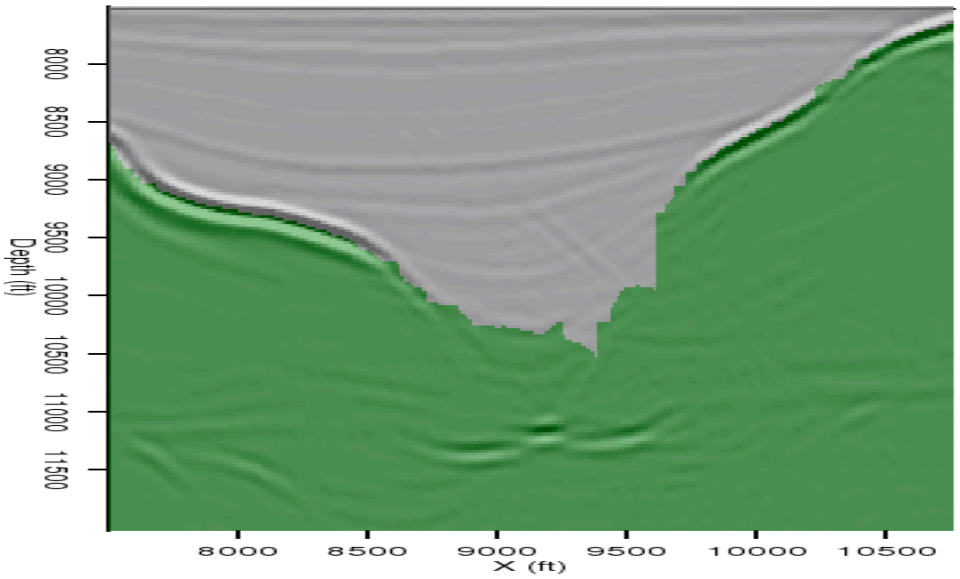


(a)

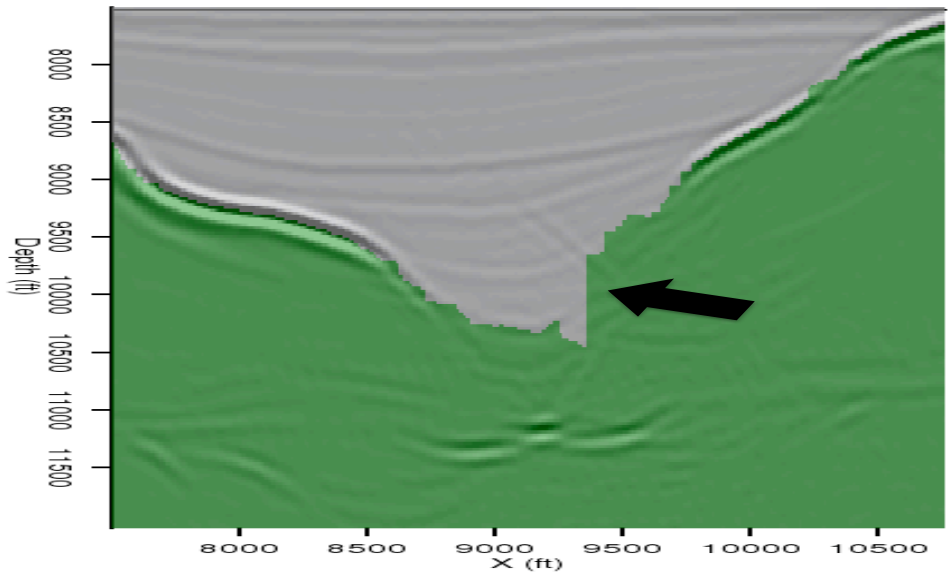


(b)

Figure 2: Sediment-flood migration (a), and corresponding segmentation result (b). The segmentation algorithm performs poorly within the area indicated on (b). [**CR**] adam2/. zig-sfimg,zig-largesfseg2

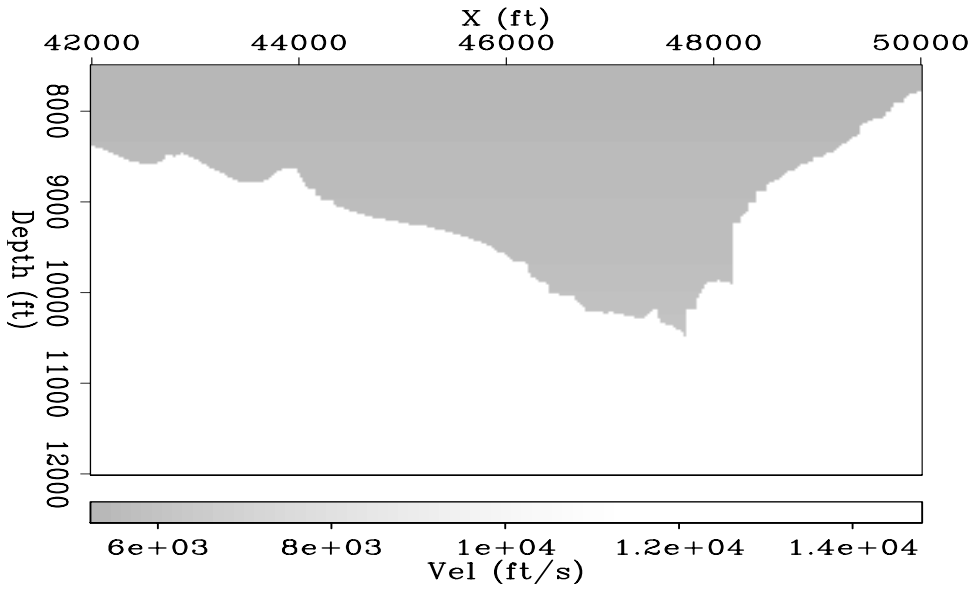


(a)

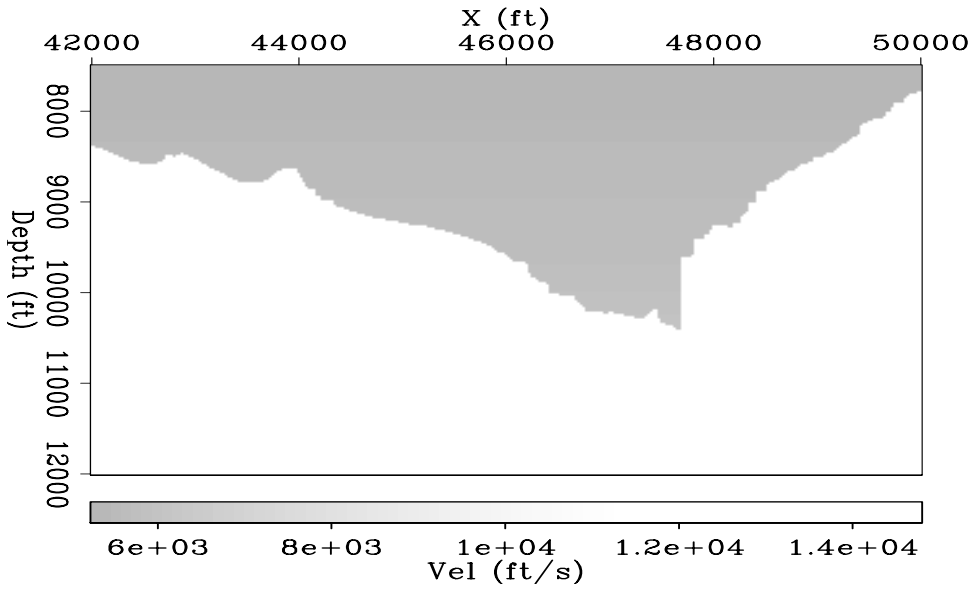


(b)

Figure 3: Two possible salt interpretations provided by the segmentation algorithm. In (b), an extra segment (indicated by the arrow) has been included in the salt. [**CR**] adam2/. zig-canyonseg1,zig-canyonseg2a



(a)



(b)

Figure 4: Salt-flood velocity models corresponding to the interpretations in Figures 3(a) and 3(b). The models were created from a sediment-flood velocity model by assigning salt velocities below the interpreted salt boundary. [**CR**] adam2/. velstitch1,velstitch2

to speed up the process is to re-datum the wavefields to a depth just above the region of interest – in this case, just above the salt canyon near $z = 7500$. If we only wish to investigate changes in a specific area of the image, we can downward continue both wavefields to this level, and inject areal source and receiver gathers. This allows us to obtain comparison images like those in Figure 5(a) and 5(b), at a computational cost an order of magnitude less than performing full migrations like the one in Figure 1(a). In this case, it is clear that the salt canyon flanks in Figure 5(a) are more sharply focused, so the salt interpretation and velocity model in Figures 3(a) and 4(a), respectively, are more accurate.

GENERALIZED WAVEFIELDS AND PHASE ENCODING

While restricting the domain and datuming the wavefields as described above significantly lessened the computational expense of updating the image, the process still occurs over a matter of minutes, rather than the seconds required to approach the level of interactivity we seek. One possibility to improve the performance of shot-profile migration is to use *phase encoding* (Romero et al., 2000), in which data from N shots are combined into a generalized source gather:

$$\hat{P}(\omega, x; j) = \sum_{i=1}^{N} \epsilon_{i,j}(\omega) P(\omega, x; \mathbf{s}_i). \tag{5}$$

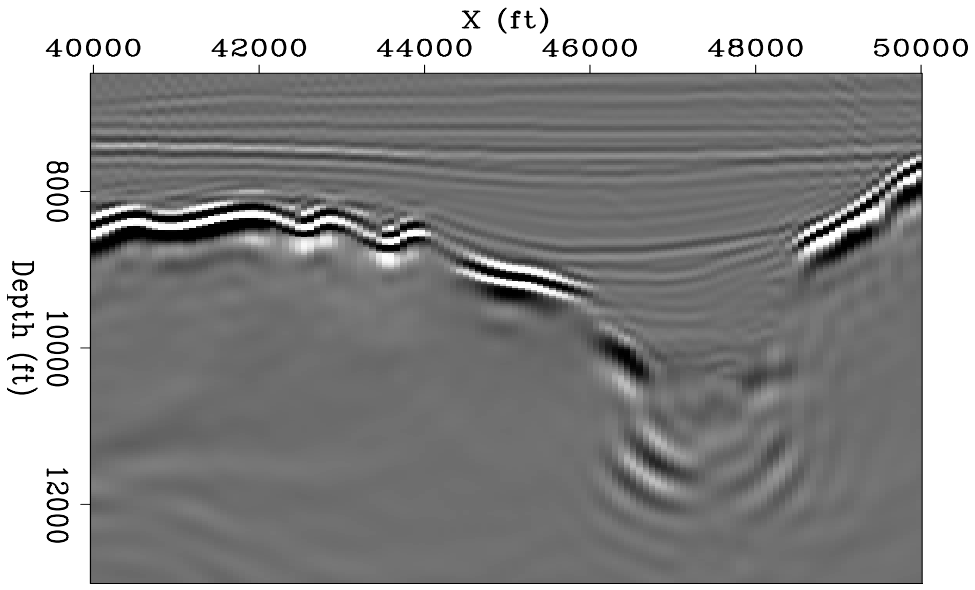
Here, the ϵ term is a complex weight value assigned to each shot.

A simple experiment combines all shots into a single generalized source gather, and uses a single generalized plane-wave source function for migration. Figure 6 is the perfect-velocity image resulting from this procedure. While information from unwanted crosstalk terms have significantly degraded the image, the salt body and its boundaries are still visible. When the two possible salt-flood velocity models in Figures 4(a) and 4(b) are used, we obtain the zoomed-in images in Figures 7(a) and 7(b), respectively. While the differences between these two images is not as apparent as for Figures 5(a) and 5(b), the salt canyon walls appear more continuous for the first model, especially near the location indicated by the arrow. These migrations were completed in less than five seconds; although this is only a 2D example, this is approaching a level at which interactive imaging becomes feasible.

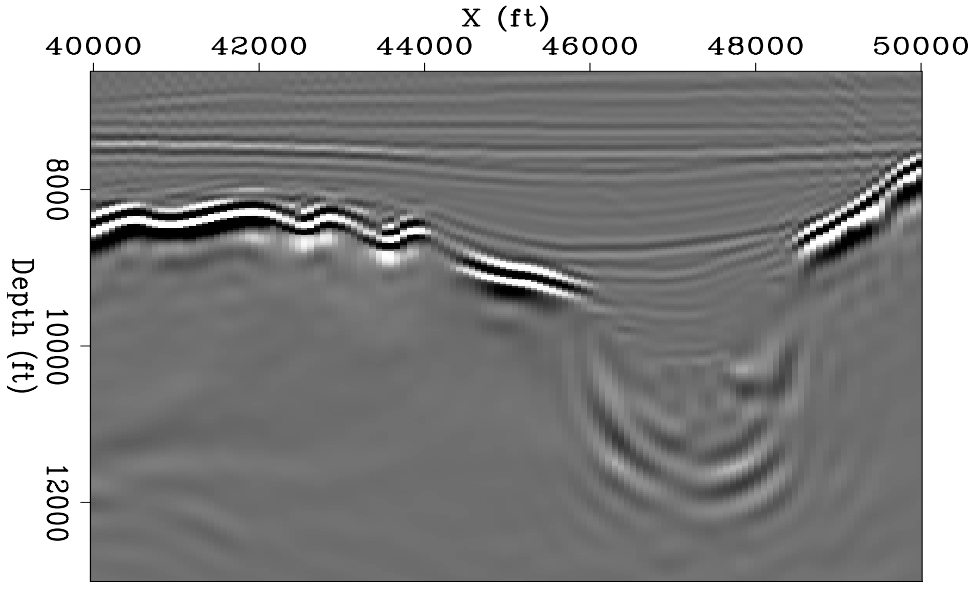
Further improvents are necessary to obtain cleaner images than in Figures 6, 7(a) and 7(b). One option is to define the weighting coefficients from equation 5 as having only imaginary (phase) components:

$$\epsilon_{i,j}(\omega) = \frac{e^{i\phi_{i,j}}(\omega)}{\sqrt{M}}, \tag{6}$$

where M is the number of generalized sources. By making ϕ a random phase function, it is possible to attenuate the crosstalk terms that arise from combining information from different shots (Morton and Ober, 1998). The implementation of a scheme combining image segmentation, re-datumed wavefields and phase-encoding could allow interpreters to interactively view high-quality images of several salt-interpretation scenarios.



(a)



(b)

Figure 5: Images resulting from the velocity models in Figures 4(a) and 4(b). The salt canyon walls are more focused in (a), indicating the the salt interpretation in Figure 3(a) is more accurate. [**CR**] adam2/. shortimg1,shortimg2

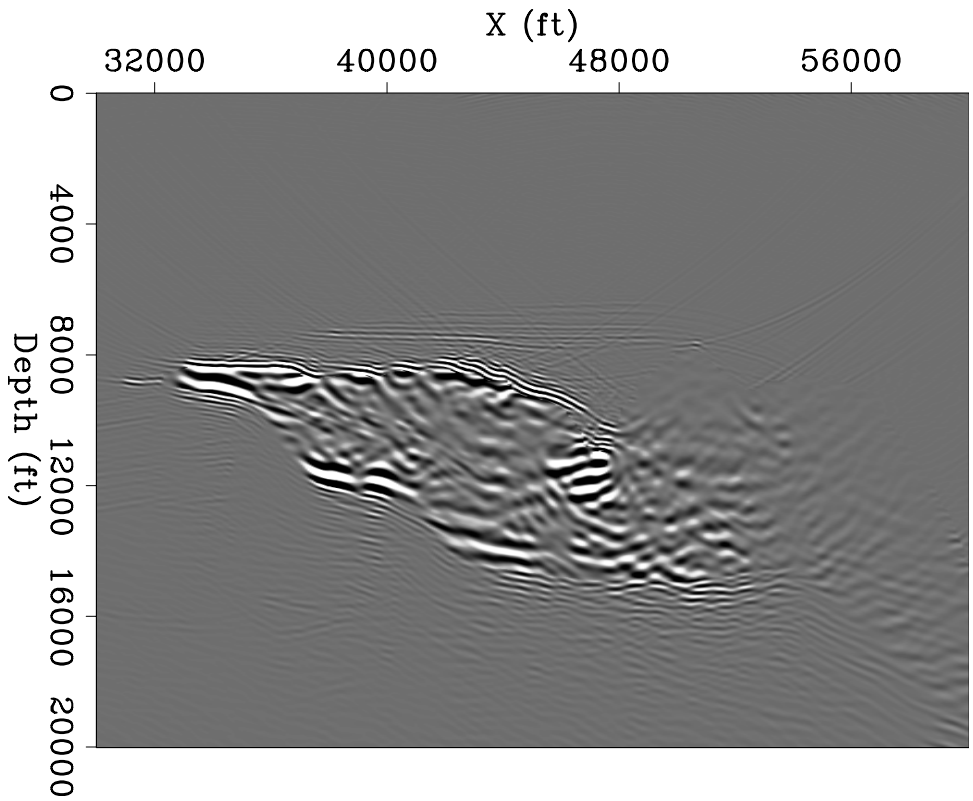
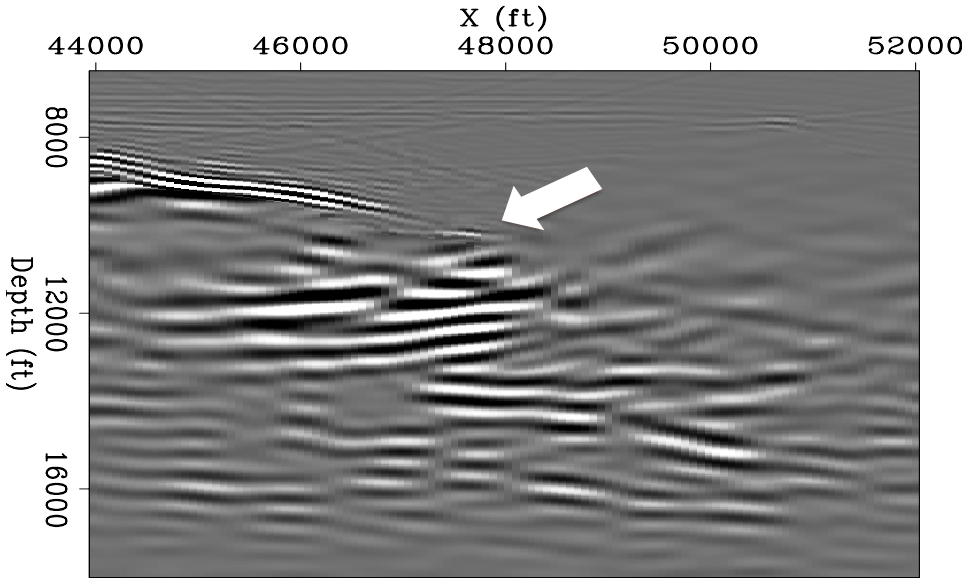
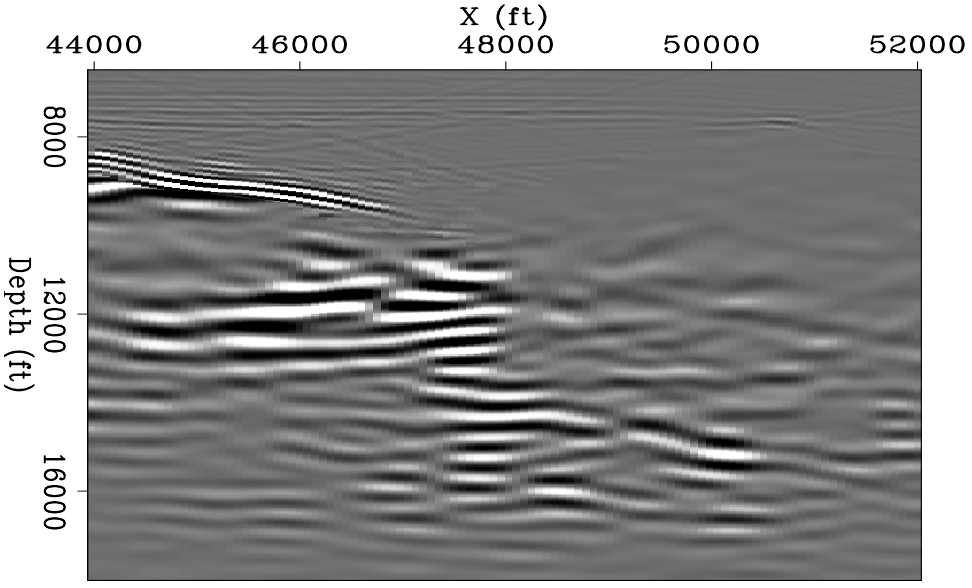


Figure 6: A perfect-velocity migration in which all shots have been combined into a single generalized source gather, and a single plane-wave is used as the generalized source function. Crosstalk artifacts have significantly degraded the image, but the salt body is still clearly visible. [**CR**] adam2/. zmig-1p



(a)



(b)

Figure 7: Generalized wavefield migrations corresponding to the velocity models in Figures 5(a) and 5(b). While crosstalk artifacts obscure the differences between the two images to a much greater extent than in Figures 5(a) and 5(b), the salt canyon wall is still noticeably more continuous near the indicated location in (a). [**CR**] adam2/. zmig-mod1a,zmig-mod2

CONCLUSIONS

The Pairwise Region Comparison (PRC) image segmentation algorithm can function effectively as part of an iterative salt interpretation and model-building workflow. However, segmentation results have a higher degree of uncertainty when a boundary is faint or discontinuous (for example, on a sediment-flood image). In such cases, a targeted shot-profile migration scheme using datumed wavefields can test two or more possible models relatively quickly, allowing an interpreter to judge each model's accuracy either qualitatively or quantitatively (e.g., flatness of angle gathers). The use of generalized wavefields further reduces the computational expense of testing multiple salt-interpretation scenarios. With additional improvements, such as the inclusion of phase-encoding migration to improve the quality of images, this scheme could form the basis of an interactive interpretation and model-building workflow.

ACKNOWLEDGMENTS

We thank SMAART JV for providing the Sigsbee 2A model used for examples, and the sponsors of the Stanford Exploration Project for their support.

REFERENCES

Biondi, B., 2005, 3-D seismic imaging: Stanford University.

Felzenszwalb, P. F. and D. P. Huttenlocher, 2004, Efficient graph-based image segmentation: International Journal of Computer Vision, **59**, 167–181.

Halpert, A., 2010, A new method for more efficient seismic image segmentation: SEP-Report, **140**, 213–228.

———, 2011, Interpreter input for seismic image segmentation: SEP-Report, **143**, 237–248.

Halpert, A., R. G. Clapp, and B. L. Biondi, 2010, Speeding up seismic image segmentation: SEG Technical Program Expanded Abstracts, **29**, 1276–1280.

Lomask, J., 2007, Seismic volumetric flattening and segmentation: PhD thesis, Stanford University.

Morton, S. A. and C. C. Ober, 1998, Fastershot-record depth migrations using phase encoding: SEG Technical Program Expanded Abstracts, **17**, 1131–1134.

Mosher, C., E. Keskula, J. Malloy, R. Keys, H. Zhang, and S. Jin, 2007, Iterative imaging for subsalt interpretation and model building: The Leading Edge, **26**, 1424–1428.

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.

Shi, J. and J. Malik, 2000, Normalized cuts and image segmentation: Institute of Electrical and Electronics Engineers Transactions on Pattern Analysis and Machine Intelligence, **22**, 838–905.

Stoffa, P. L., J. Fokkema, R. M. de Luna Freire, and W. P. Kessinger, 1990, Split-step Fourier migration: Geophysics, **55**, 410–421.

A new algorithm for bidirectional deconvolution

Yi Shen, Qiang Fu and Jon Claerbout

ABSTRACT

We introduce a new algorithm for bidirectional deconvolution. In our method, we estimate the causal filters and anti-causal filters simultaneously instead of alternately. We test three data examples (1D synthetic, 2D synthetic and 2D field data). The results show that the wavelet can be compressed almost into a spike using our method. The two filters can be estimated equally when we are dealing with a zero-phase wavelet. In addition, our method has a lower computational cost and faster convergence rate than the method discussed by Zhang and Claerbout (2010).

INTRODUCTION

In a previous report, Zhang and Claerbout (2010) introduced a bidirectional deconvolution method that overcomes the minimum-phase assumption of the conventional deconvolution. They factored the mixed-phase wavelet into two parts, the minimum-phase part and the maximum-phase part, which can be estimated by a causal filter and an anti-causal filter, respectively. Since such deconvolution is a non-linear problem, a pair of conventional linear deconvolutions were utilized to invert these two filters alternately and iteratively. In their paper, both theory and data examples showed that the mixed-phase wavelet can be accurately inverted using this bidirectional deconvolution.

However, there are some obstacles to inverting these two filters sequentially. There is a battle between these two filters competing for the spectrum. This competition makes the solution jump back and forth between the causal filter and the anti-causal filter, which may lead to a low convergence rate and an unstable deconvolution result. In addition, when Zhang and Claerbout (2010) inverted a zero-phase wavelet by this method, they produced two different filters; that is, the causal part and the anti-causal part are different, which is contrary to the nature of the zero-phase wavelet.

To avoid these problems, we invert these two filters at the same time instead of sequentially, hoping this simultaneous inversion will lead to a faster convergence rate and more stable solutions.

THEORY

In this paper, we still rely on the idea of bidirectional deconvolution to deal with the mixed-phase wavelet. The wavelet can be factored into a minimum-phase part and a non-minimum-phase part. The deconvolution problem can be defined as follows:

$$d * a * b^r = r, \tag{1}$$

where d is the data, a and b are the unknown causal filters, and superscript r denotes the time reverse of filter b. Again, the hybrid norm is applied to r, and the reflectivity model is simply r plus a time shift. Now consider perturbations Δa and Δb:

$$d * (a + \Delta a) * (b^r + \Delta b^r) = r. \tag{2}$$

If we assume the the nonlinear part $\Delta a \Delta b$ is relatively small, we can neglect this term:

$$d * a * b^r + d * a * \Delta b^r + d * b^r * \Delta a \approx r. \tag{3}$$

We use matrix algebraic notation to rewrite the fitting goal. We also want to guarantee filter a to be causal and filter b^r to be anti-causal during the iterations. For this we need mask matrices (diagonal matrices with ones on the diagonal where variables are free and zeros where they are constrained). The free-mask matrix for Δa is denoted $\mathbf{K}$, whose first diagonal element is zero, and that for Δb^r is denoted $\mathbf{Y}$, whose last diagonal element is zero:

$$\begin{bmatrix} \mathbf{d} * \mathbf{a} & \mathbf{d} * \mathbf{b}^r \end{bmatrix} \begin{bmatrix} \mathbf{Y} & \mathbf{0} \\ \mathbf{0} & \mathbf{K} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{b}^r \\ \Delta \mathbf{a} \end{bmatrix} + \mathbf{d} * \mathbf{a} * \mathbf{b}^r \approx \mathbf{0}. \tag{4}$$

From equation (4), we have our new model $\mathbf{m} = \begin{bmatrix} \Delta \mathbf{b}^r & \Delta \mathbf{a} \end{bmatrix}^T$ and new operator $\mathbf{F} = \begin{bmatrix} \mathbf{d} * \mathbf{a} & \mathbf{d} * \mathbf{b}^r \end{bmatrix}$. Now we can acquire these two filters only by applying the conventional inversion method and hybrid norm solver. The pseudocode for minimizing this new objective function by the hyperbolic conjugate-direction method developed by Claerbout (2010) is:

```
non - linear iteration
{
  r = −d ∗ a ∗ b^r
  F = [ d ∗ a   d ∗ b^r ]
  linear iteration
  {
    g = (FJ)^T H'(r)
    Δr = FJg
    m ← Hyperbolic_cgstep(g, m, Δr, r)
  }
  a ← a + Δa
  b^r ← b^r + Δb^r
}
```

where $H'(\mathbf{r})$ is defined as the first derivative of the hybrid norm $\sqrt{R^2 + \mathbf{r}^2} - R$, where R is the l_1/l_2 threshold parameter, $\mathbf{J}$ is the mask matrix $\begin{bmatrix} \mathbf{Y} & \mathbf{0} \\ \mathbf{0} & \mathbf{K} \end{bmatrix}$, and $\mathbf{g}$ is the gradient.

From the template we notice that both linear and non-linear iterations are needed. Perturbations $\Delta \mathbf{a}$ and $\Delta \mathbf{b}^r$ are inverted by the hyperbolic conjugate-direction method in each

linear iteration. Filters $\mathbf{a}$ and $\mathbf{b}^r$ are updated in the non-linear iteration, which generates a new operator $\mathbf{F}$ to update the model. However, this method requires only 2 linear iterations to reach convergence, instead of the 100 linear iterations required by the previous method, greatly speeding convergence. In addition, there is no need to reverse the filters in the non-linear iteration, which makes our implementation more convenient.

Although the fitting goal is linearized, we still need the initial model to be close enough to get a good result. Here we expect an impulse function for both filters a and b. The following sections will show the application of this new method and demonstrate its effectiveness and limitations, when compared with the previous method discussed by Zhang and Claerbout (2010).

APPLICATION

Single wavelet

The first data example is the simplest mixed-phase wavelet, which only has three points [3,7,2]. We use it to verify the ability of our method to deal with the mixed-phase wavelet. The input data and its bidirectional result are shown in Figures 1 and 2. In this case, our new method is able to compress the simple mixed phase wavelet into a spike.

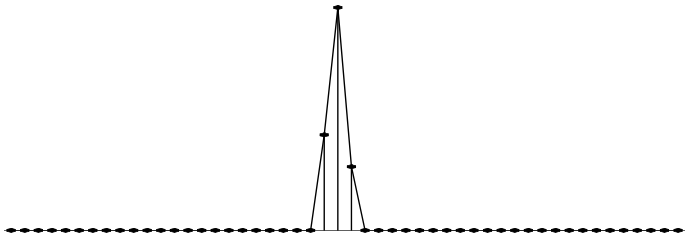


Figure 1: The three data points [3,7,2]. [**ER**] yishen/. data1

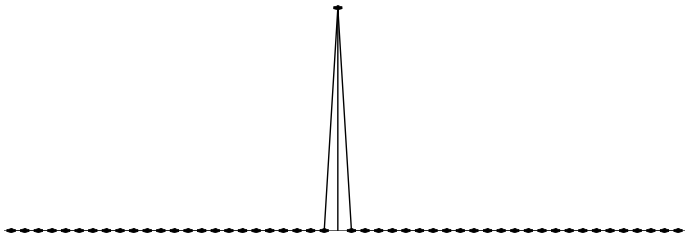


Figure 2: The result of our deconvolution method. [**ER**] yishen/. mod1

To illustrate the capabilities and limitations of our new method, we analyze the results obtained by inverting the zero-phase wavelet. This wavelet is created by convolving the minimum-phase with its own time-reversed wavelet.

Figures 3(a) and 3(b) show the filters estimated by our method. Figures 3(c) and 3(d) show the filters inverted by the previous method of bidirectional deconvolution. Here we time-reverse the anti-causal filter b^r into a causal filter b for easier comparison. Ideally, filter a and filter b should be identical, because the zero-phase wavelet is symmetric, with its minimum-phase part the same as its maximum-phase part, but time-reversed. The results of our method perfectly satisfy the theory, which shows the extreme similarity between filter a and filter b. When we invert the filters, the update direction is the same for both filters, because the searching gradients are equal. However, using the method from Zhang

and Claerbout (2010) yields a filter a that is quite different from filter b, because they are inverted separately.

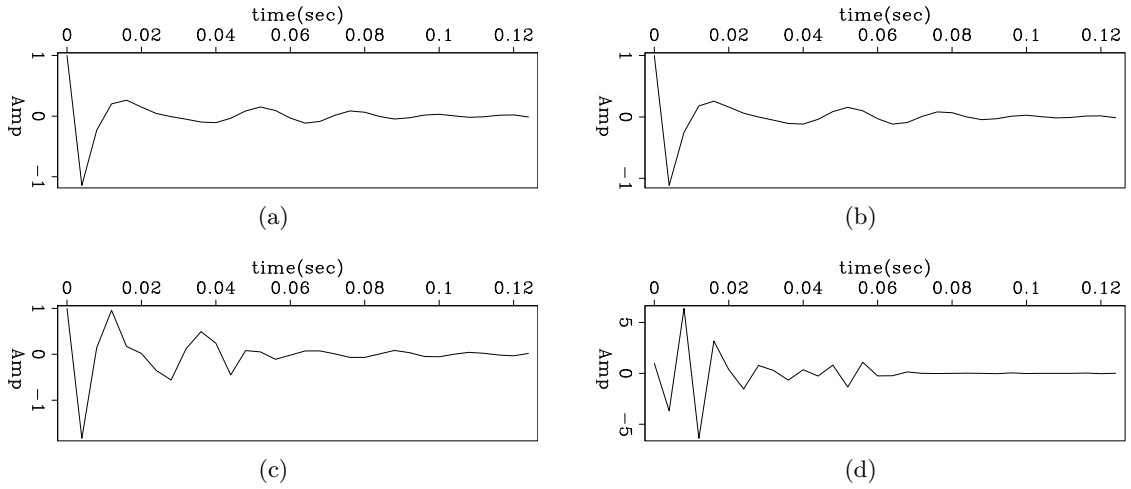


Figure 3: For zero-phase wavelet inversion, (a) filter a estimated by our method; (b) filter b estimated by our method; (c) filter a estimated by the previous method; (d) filter b estimated by the previous method. [**ER**] yishen/. filt2anew,filt2bnew,filt2aold,filt2bold

Figures 4, 5 and 6 show the zero-phase wavelet and its bidirectional deconvolution, using our new algorithm and the method of Zhang and Claerbout (2010). The results show that the wavelet is almost compressed into a spike by our method, but it is not as spiky as the result of the previous method. One possible reason may be the nature of the non-linear problem. There may be multiple minima in this problem, and due to our additional condition that filter a and filter b should be the same in this case, we find a different minimum, which leads to a different result.

Thus a good starting guess may help us to get a better result. Or perhaps the preconditioning can also make the solution fast converge to the global minima by utilizing prior information.

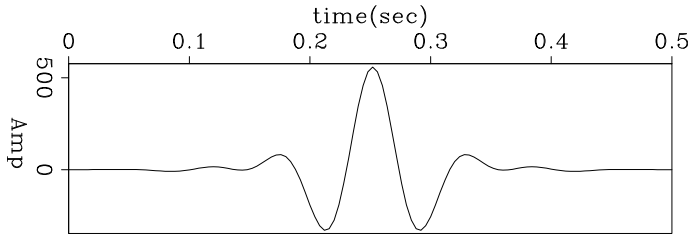


Figure 4: Zero-phase wavelet. [**ER**] yishen/. data2

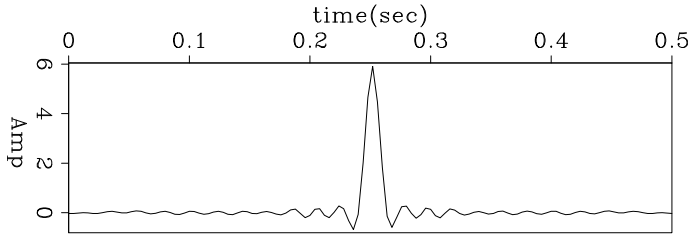


Figure 5: Deconvolution result by our method. [**ER**] yishen/. mod2new

Figures 7 and 8 show the wavelet estimated by our method and the previous method. We

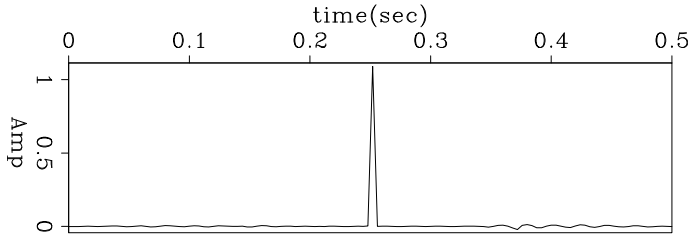


Figure 6: Deconvolution result by the previous method. [**ER**] yishen/. mod2old

notice that both of the wavelets approximate the input zero-phase wavelet, which is shown in Figure 4. However our estimated wavelet looks more symmetric and cleaner with less side lopes than the one estimated by the previous method. The reason is that our estimated filters are identical, which make the causal part and anti-causal part of the inverted wavelet the same.

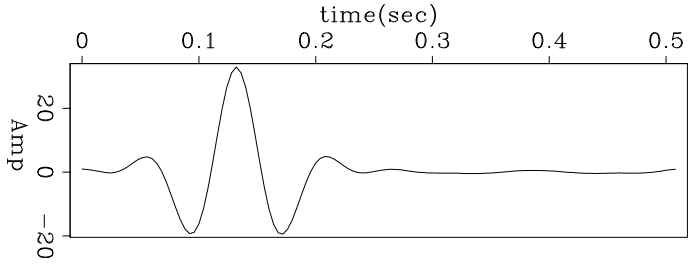


Figure 7: Shot wavelet estimated by our method yishen/. wavelet2new

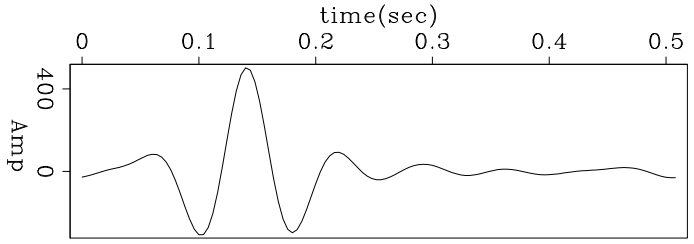


Figure 8: Shot wavelet estimated by the previous method yishen/. wavelet2old

2D synthetic data

After applying deconvolution on the simple 1D case, we test the Zhang and Claerbout (2010) method and our new method on more complicated 2D synthetic data. Figure 9(a) shows the starting reflectivity model. Figure 9(b) shows the data generated by convolving the reflectivity model with the zero-phase wavelet in the previous section. All traces use the same wavelet when generating the data, and all traces share the same wavelet when we are doing the deconvolution.

Figures 10(a) and 10(b) show the bidirectional deconvolution using our method and the older method. Both methods retrieve the sparse reflectivity model and compress the wavelet into a spike, but the deconvolution model produced by the previous method is more spiky than ours, just as in the previous section, because of the wavelet we use.

We also compare the computational costs of these two methods. To deal with this synthetic data, we use 2 linear iterations and 280 non-linear iterations to reach convergence, whereas Zhang and Claerbout (2010) used 100 linear iterations and 20 non-linear iterations, a total of 2000 iterations. In fact, our code is almost 6 times faster than the previous method.

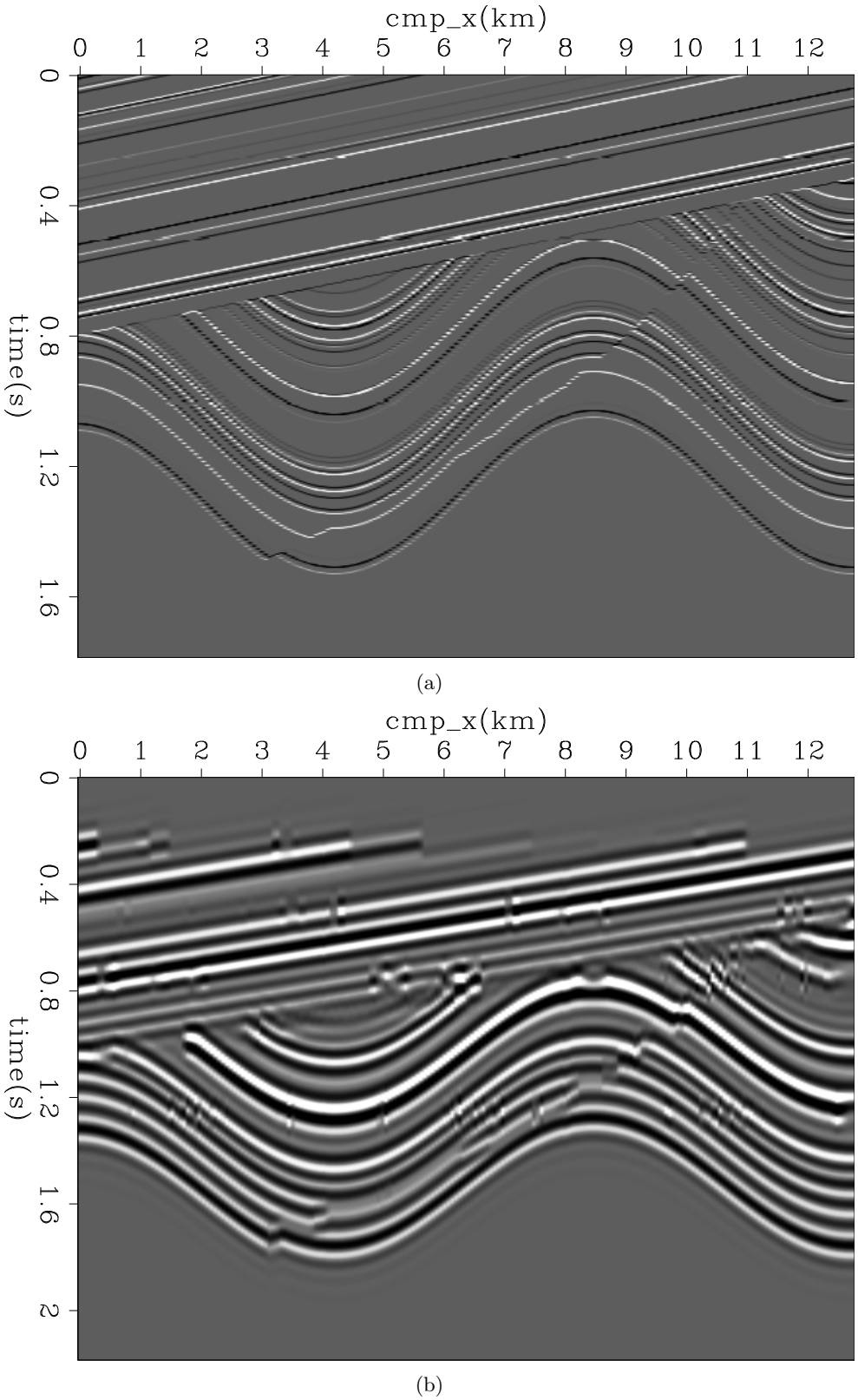


Figure 9: (a) The 2D synthetic reflectivity model; (b) the synthetic data generated using the zero-phase wavelet. [**ER**] yishen/. syn3,data3

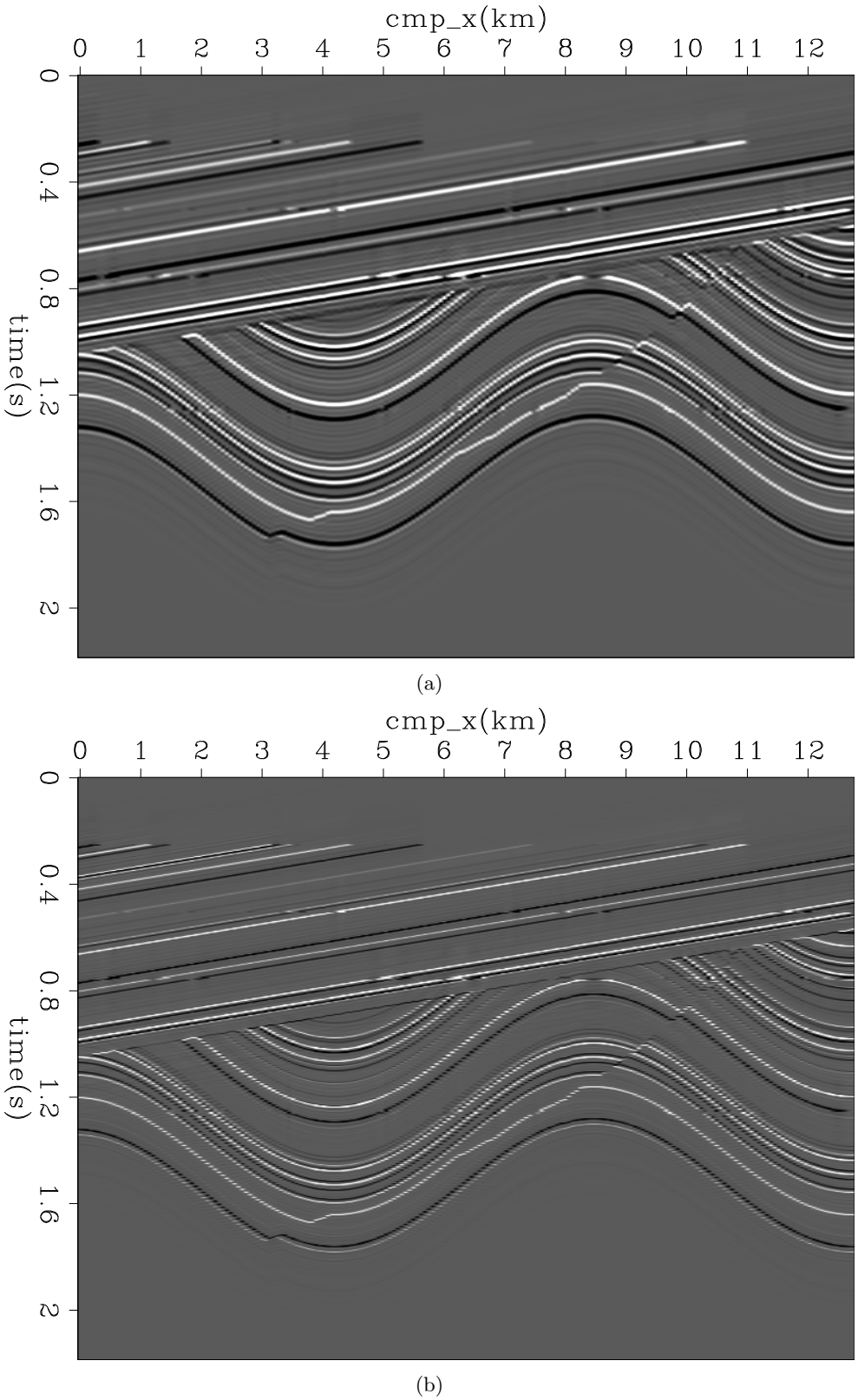


Figure 10: Given the 2D synthetic data in Figure 9(b), (a) reflectivity model retrieved using our method; (b) reflectivity model retrieved using the previous method. [**ER**] yishen/. mod3new,mod3old

2D field data

The field data we use in this example is a common-offest section of marine field data. Figure 11 shows the input data. Figures 12(a) and 12(b) show the bidirectional deconvolution using our method and the previous method. Both perform well to retrieve the sparse reflectivity in this field data.

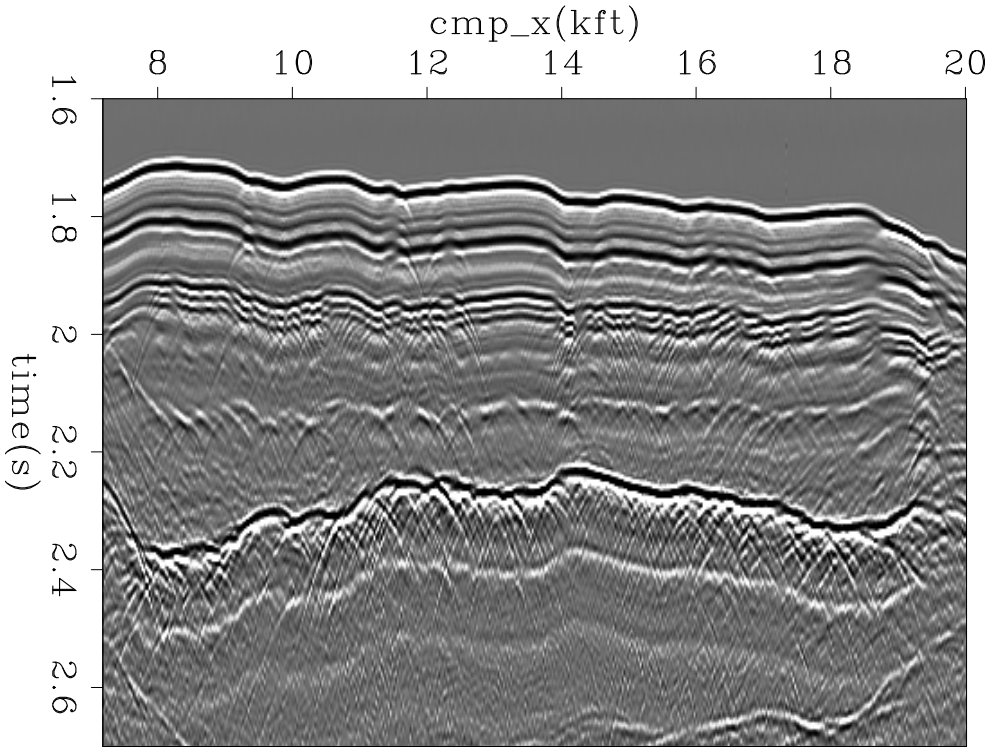
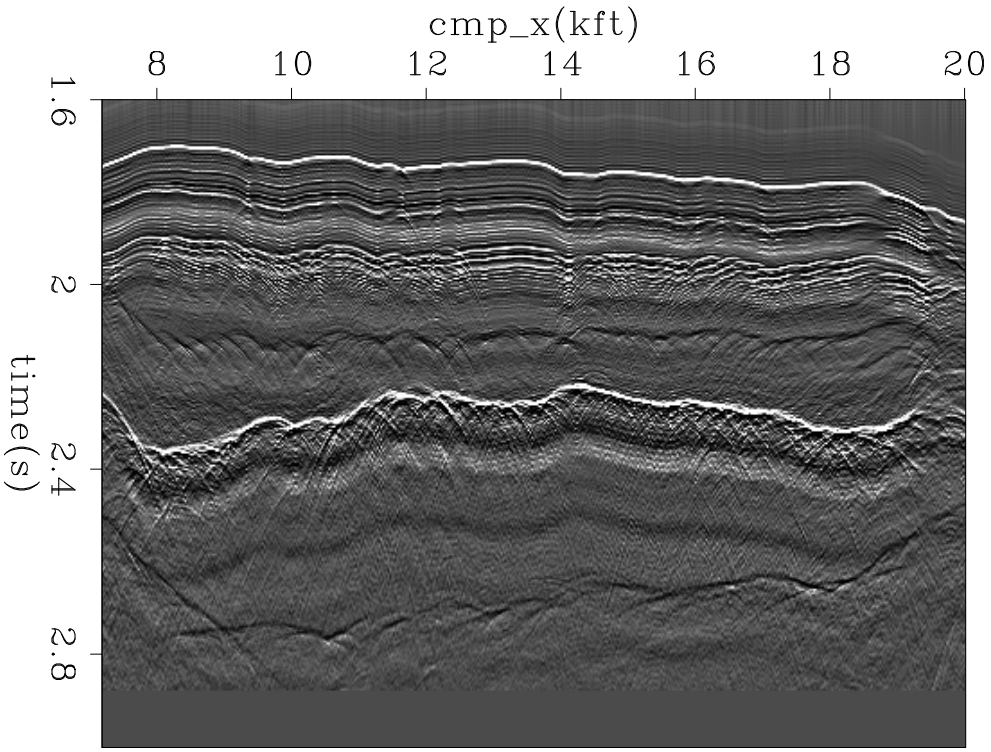


Figure 11: Input Common Offset data. [**ER**] yishen/. data4

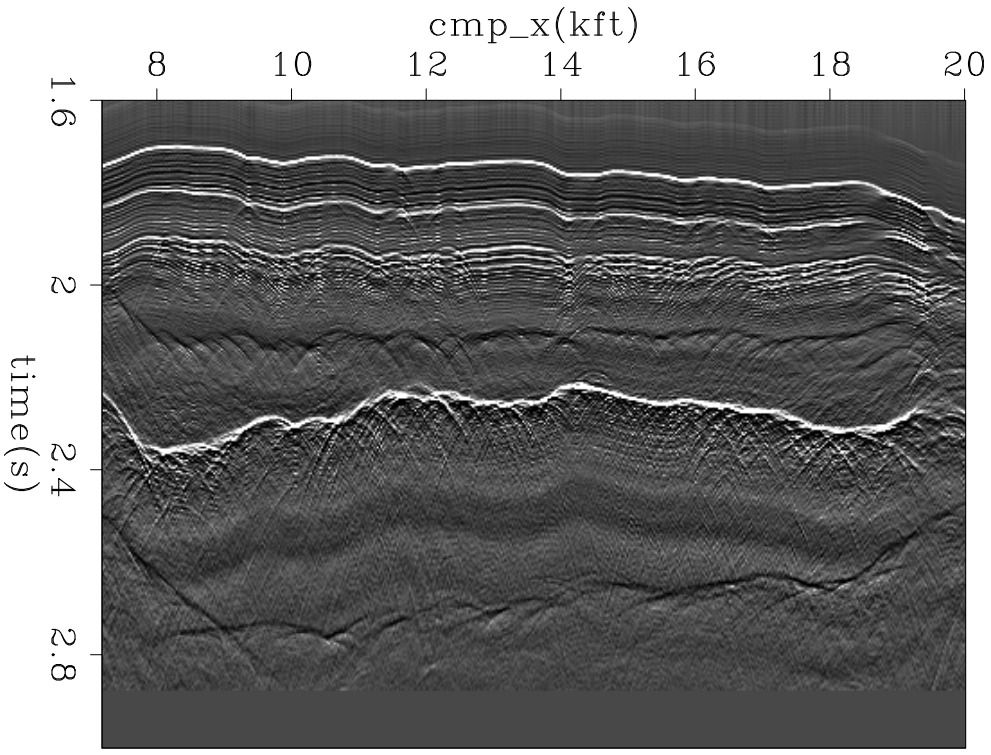
The filters estimated by our method are shown in Figures 14(a), and 14(b), and the filters estimated by the previous method are plotted in Figures 14(c) and 14(d). Because the wavelet we aim to invert is not symmetric, filter a and filter b are not equal. However, the strong events look like a double ghost (white, black,white), which approximates a symmetric wavelet. Thus we would like our filters to resemble each other. From the result, we notice that our filters satisfy these expectation.

Figures 13(a) and 13(b) show the estimated shot waveform. The data sampling is 4ms. We notice that both of our method estimate the bubbles and the double ghost. However, our inverted waveform is more like a double ghost, which can be noticed in the data. The reason is that our filters more resemble each other than the ones estimated by the previous method.

Our method requires 2 linear iterations and 100 non-linear iterations, or only 200 iterations in total. Zhang and Claerbout (2010) used 100 iterations and 8 non-linear iterations, or 800 iterations in total. Therefore, our method is four times faster, which is a large reduction in computational cost.



(a)



(b)

Figure 12: Given the common offset data in Figure 11, (a) reflectivity model retrieved using our method; (b) reflectivity model retrieved using the previous method. [**ER**] yishen/. mod4new,mod4old

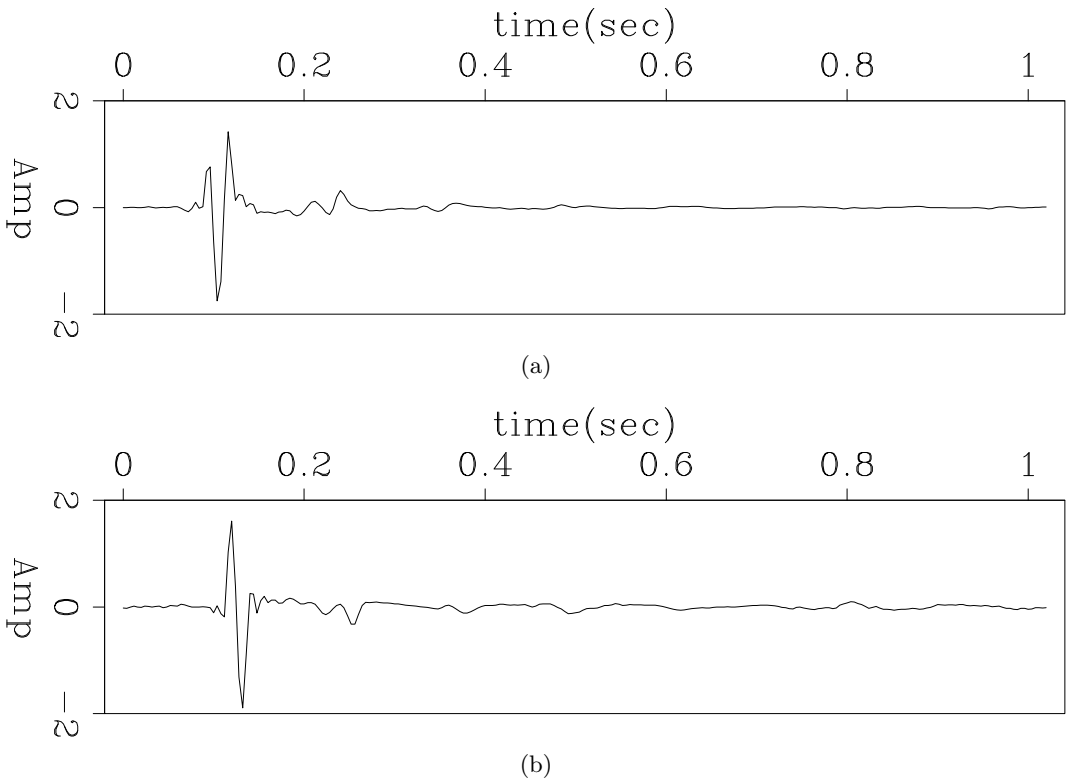


Figure 13: For 2D synthetic data, (a) shot wavelet estimated by our method; (b) shot wavelet estimated by the previous method. [**ER**] yishen/. wavelet4new,wavelet4old

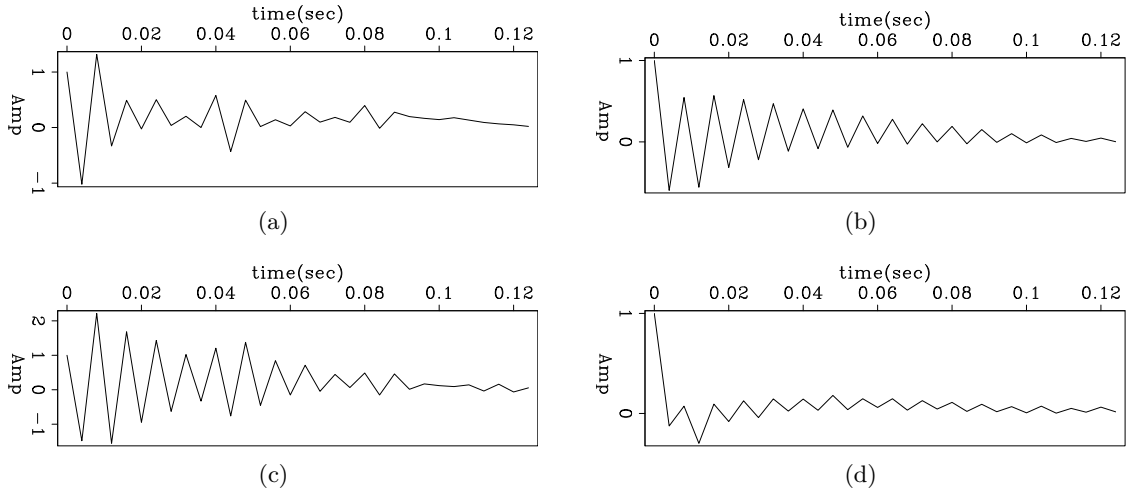


Figure 14: Given the common offset data in Figure 11, (a) filter a estimated by our method; (b) filter b estimated by our method; (c) filter a estimated by the previous method; (d) filter b estimated by the previous method. [**ER**] yishen/. filt4anew,filt4bnew,filt4aold,filt4bold

CONCLUSION

In this paper, we introduce an algorithm for bidirectional deconvolution that estimates the two filters simultaneously. We test the new method on three data examples. The results show that the wavelet can be compressed almost into a spike. When we are dealing with the zero-phase wavelet, we obtain two identical filters, a major improvement compared with the previous bidirectional method. Another important advantage is the low computational cost and fast convergence rate due to the reduced number of linear iterations. However, we are surprise to see that our results are not as spiky as the ones produced by Zhang and Claerbout (2010). One possible reason may be the nature of the non-linear problem. Perhaps we need a good initial guess or preconditioning to achieve acceptable results.

FUTURE WORK

As mentioned previously, the nature of the nonlinear problem strongly affects our results. Thus, a good initial guess is needed to obtain a better sparse reflectivity. In most cases, data will resemble the Ricker wavelet, as is true for the band-limited marine seismic data with ghosts and the for the land response of an accelerometer. For this situation, we can use the Ricker wavelet to approximate the data and derive the initial filter from this wavelet. Since the Ricker wavelet vanishes at zero frequency and at the Nyquist frequency, it has no stable inverse. Therefore, we use the approximated Ricker wavelet instead of the true one.

Another potential solution is to do the preconditioning, which utilizes prior information. In this non-linear problem, we hope it can guide the gradient along sensible pathways thus avoiding potential local minima.

ACKNOWLEDGMENTS

The authors thank the sponsors of Stanford Exploration Project for their financial support, and also thank Yang Zhang, Antoine Guitton, Shuki Ronen, Mandy Wong and Elita Li for fruitful discussion.

REFERENCES

Claerbout, J. F., 2010, Image estimation by example.

Zhang, Y. and J. Claerbout, 2010, A new bidirectional deconvolution method that overcomes the minimum phase assumption: SEP-Report, **142**, 93–103.

An approximation of the inverse Ricker wavelet as an initial guess for bidirectional deconvolution

Qiang Fu, Yi Shen and Jon Claerbout

ABSTRACT

Bidirectional deconvolution is a powerful tool for performing blind deconvolution on a signal that contains a mixed-phase wavelet, such as seismic data. Previously, we used a prediction error filter (PEF) as the minimum-phase filter and an impulse function as the maximum-phase filter as initial guesses. This surprised us by apparent instability as the solution would jump from spiking the first pulse of a Ricker wavelet to get larger middle pulse. In this paper, we propose new initial guesses for the causal and anti-causal deconvolution filters that are more effective on data with a Ricker-like wavelet. We test these on both synthetic data and field data. The results demonstrate that the new starting filters do a better job than the previous initial guesses.

INTRODUCTION

Zhang and Claerbout (2010) proposed a new method for blind deconvolution to overcome the minimum-phase assumption, called "bidirectional deconvolution". A seismic data trace can be represented by a convolution of a wavelet with a reflectivity series,

$$d = r * w, \tag{1}$$

where d denotes the seismic data trace, r denotes the reflectivity series, and w denotes a wavelet.

In conventional blind deconvolution, we assume that w is a minimum-phase wavelet. However this assumption is not applicable for field seismic data. If w denotes a mixed-phase wavelet, it can be represented by a convolution of two parts: $w = w_a * w_b^r$, where w_a is a minimum-phase wavelet and w_b is a reversed minimum-phase wavelet; hence w_b^r itself is a maximum-phase wavelet. (The superscript r denotes reversed in time.) Thus equation 1 can be rewritten as

$$d = r * (w_a * w_b^r). \tag{2}$$

If we know the inverse filters f_a and f_b^r for w_a and w_b^r, respectively, to satisfy

$$\begin{cases} w_a * f_a = \delta(n) \\ w_b * f_b = \delta(n) \end{cases}, \tag{3}$$

we can recover the reflectivity series:

$$r = d * f_a * f_b^r, \tag{4}$$

where filter f_a is the inverse signal of w_a and filter f_b is the inverse signal of w_b

Now we can use nonlinear inversion to solve this blind deconvolution problem for a mixed-phase wavelet by solving the two equations below alternately:

$$\begin{cases} (d * f_b^r) * f_a = r_a \\ (d * f_a)^r * f_b = r_b^r \end{cases} \tag{5}$$

where both f_a and f_b are minimum-phase signals.

Shen et al. (2011) proposed another method to solve equation 4. Instead of solving for f_a and f_b alternately, they solve f_a and f_b simultaneously. Using this new approach allows us to estimate results with similar waveforms for f_a and f_b, which is a natural characteristic for data with a Ricker-like wavelet. In addition, this new method is faster than previous one. Hence we will use this to perform bidirectional deconvolution.

Since bidirectional deconvolution is a nonlinear problem, it requires that the starting model be close to the true one, and it is highly sensitive to the initial guess for both f_a and f_b. Shen et al. (2011) uses a simple one-spike impulse function for both filters. However, sometimes the true model does not resemble an impulse function. Therefore, we attempt to find a better initial guess.

APPROXIMATION OF THE INVERSE RICKER WAVELET

Generally, it is a complicated issue to find a good initial guess for bidirectional deconvolution. However, in most field and synthetic geophysical data the wavelet is similar to a Ricker wavelet. For example, band-limited marine seismic data with ghosts and the land response of an accelerometer are both Ricker-like. Hence Ricker-like wavelets have broad applicability. For this reason, we choose a Ricker wavelet to approximate the wavelet of the data and derive the initial filters from the inverse Ricker wavelet. If we could derive the inverse of the Ricker wavelet, it would provide a suitable initial guess.

In theory, however, Ricker wavelets do not have a stable inverse. Therefore we must find an approximate inverse to use as the initial guesses for filters f_a and f_b. Since we need two initial guesses, one for each filter, our approximate inverse should consist of two symmetric parts.

We have three tasks: first we must find a finite approximation for the continuous Ricker wavelet; second we must separate the approximate form into two symmetric parts; and third we must find a way to avoid the singularity problems we encounter when inverting these two parts directly in the frequency domain. Let's address these tasks one by one.

Finite approximation

It is known that a Ricker wavelet is the second-order derivative of a Gaussian function. For computation, we use a finite and discrete approximation to a Ricker wavelet as a replacement to the infinite and continuous real second-order derivative of a Gaussian function. We use a second-order finite-difference operator to approximate a second-order derivative and binomial coefficients to approximate a Gaussian function.

In the Z domain,

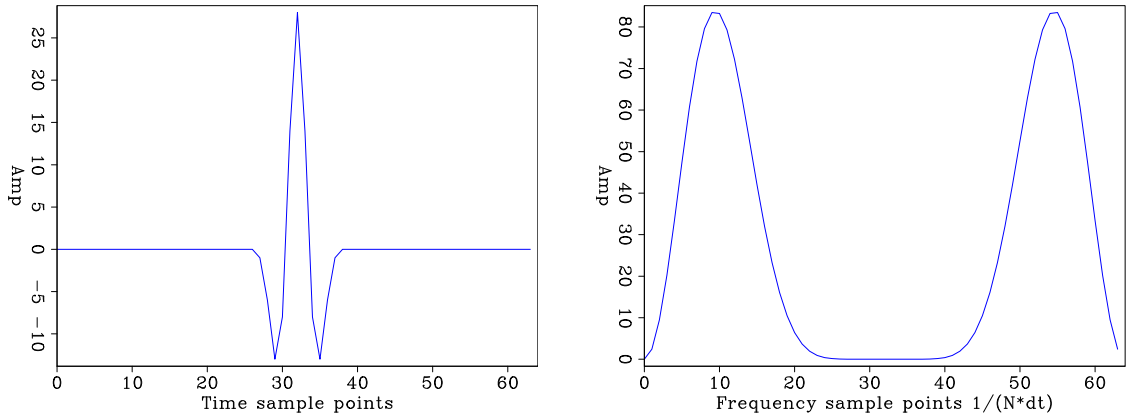
$$\text{Ricker} = -(1-Z)^2(1+Z)^{2N}. \tag{6}$$

The parameter N is half the order of the binomial we used. Here we use $2N$ in equation 6 instead of N simply to keep the order of the binomial even to facilitate the later separation. In practice, we would choose the value of N parameter according to the wavelength (or principle frequency component) of the wavelet in our data. The larger the value, the wider the wavelet.

Figures 1(a) and 1(b) show this fourth-order (N=4) finite approximation of the Ricker wavelet in the time and frequency domains. Here we use the fourth-order as an example, but we can use a different-order implementation as long as the approximate Ricker wavelet has the same wavelength (or principle frequency component) as the wavelet of our data.

Separation

We can manipulate equation 6 and decompose it into two symmetric parts. First, we shift the Ricker wavelet to the center of the axes:



(a) (b)

Figure 1: The fourth-order finite approximation of the Ricker wavelet: (a) in the time domain; (b) in the frequency domain. [**ER**] qiang/. ricker,ricker-freq

$$\text{Ricker} = -\frac{(1-Z)^2}{Z}\frac{(1+Z)^{2N}}{Z^N} \tag{7}$$

which rearranges to

$$\text{Ricker} = [(1-\frac{1}{Z})(1+\frac{1}{Z})^N][(1-Z)(1+Z)^N]. \tag{8}$$

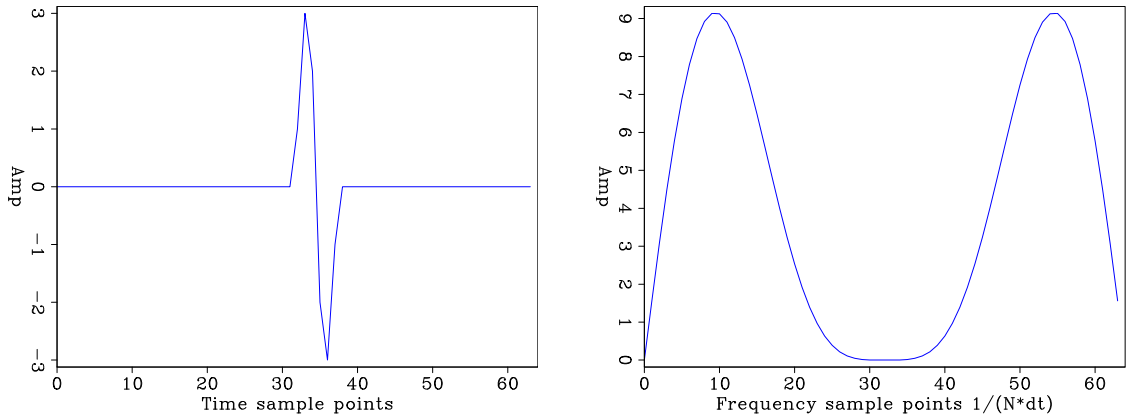
We call one of these symmetric parts a "half-Ricker wavelet":

$$H(Z) = (1-Z)(1+Z)^N \tag{9}$$

Figures 2(a) and 2(b) show this finite approximation of the Ricker wavelet in the time and frequency domains.

We denote the inverse of the half-Ricker wavelet, which is our candidate as the initial guess for both filter f_a and filter f_b^r, as $\hat{f}_a$ and $\hat{f}_b^r$ in the time domain and $\hat{F}_a$ and $\hat{F}_b^r$ in the Z domain:

$$\begin{cases} \hat{F}_a(Z) = \hat{F}_b(Z) = \frac{1}{H(Z)} = \frac{1}{(1-Z)(1+Z)^N} \\ \overline{\hat{F}_b^r(Z)} = \hat{F}_b^r(\frac{1}{Z}) = \frac{1}{H(\frac{1}{Z})} = \frac{1}{(1-\frac{1}{Z})(1+\frac{1}{Z})^N} \end{cases} \tag{10}$$



(a) (b)

Figure 2: The fourth-order half-Ricker wavelet: (a) in the time domain; (b) in the frequency domain. [**ER**] qiang/. half-ricker,half-ricker-freq

Avoiding the singularity

In the Z domain, the half-Ricker wavelet has two zeros, which makes it impossible to directly invert. Therefore, we modify the formula with two new factors to avoid errors caused by dividing by zero:

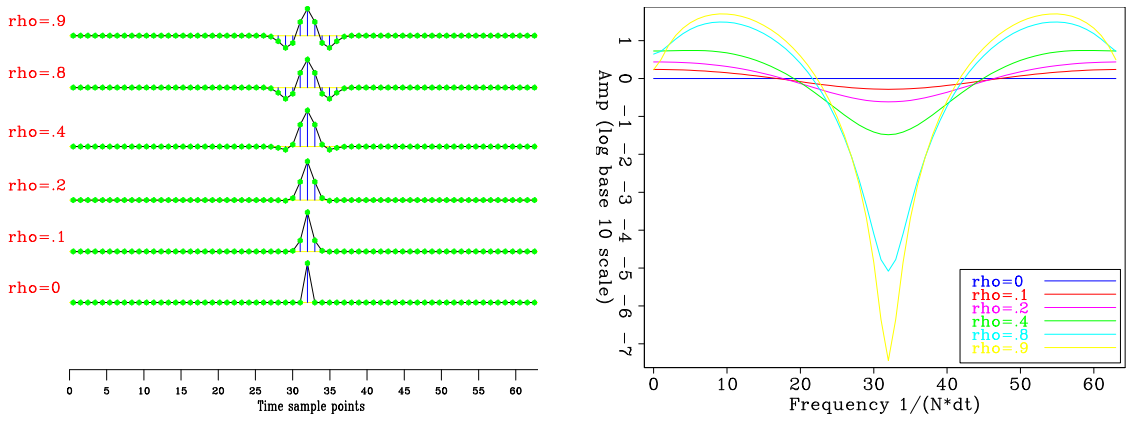
$$\hat{F}_a(Z) = \hat{F}_b^r(Z) = \frac{1}{H(Z)} = \frac{1}{(1-\rho_1 Z)(1+\rho_2 Z)^N} \tag{11}$$

For simplicity, we use the same value for ρ_1 and ρ_2, $\rho = \rho_1 = \rho_2$.

Figures 3(a) and 3(b) show the fourth-order Ricker wavelet in the time and frequency domains with different ρ values.

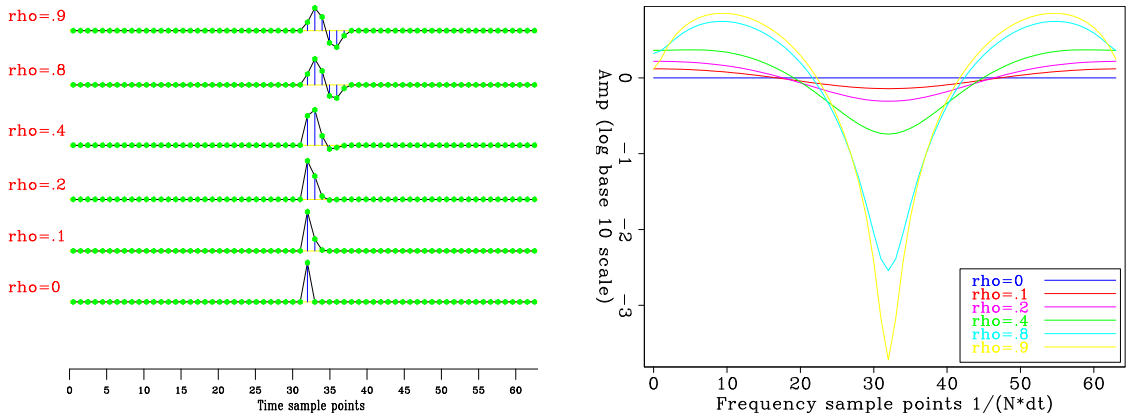
Figures 4(a) and 4(b) show the fourth-order half-Ricker wavelet in the time and frequency domains with different ρ values.

Since we have $\hat{f}_a$ and $\hat{f}_b^r$ in the Z domain ($\hat{F}_a$ and $\hat{F}_b$), we have two methods to recover time-domain filter coefficients. The first one is to use $Z = e^{iw}$ to get frequency domain filters and then transfer them into the time domain by an inverse Fourier transform. Another method is to use a Taylor expansion to get time-domain filter coefficients. In theory, the



(a) (b)

Figure 3: The fourth-order Ricker wavelet with different ρ values: (a) in the time domain; (b) in the frequency domain. [**ER**] qiang/. ricker-with-rho,ricker-with-rho-freq

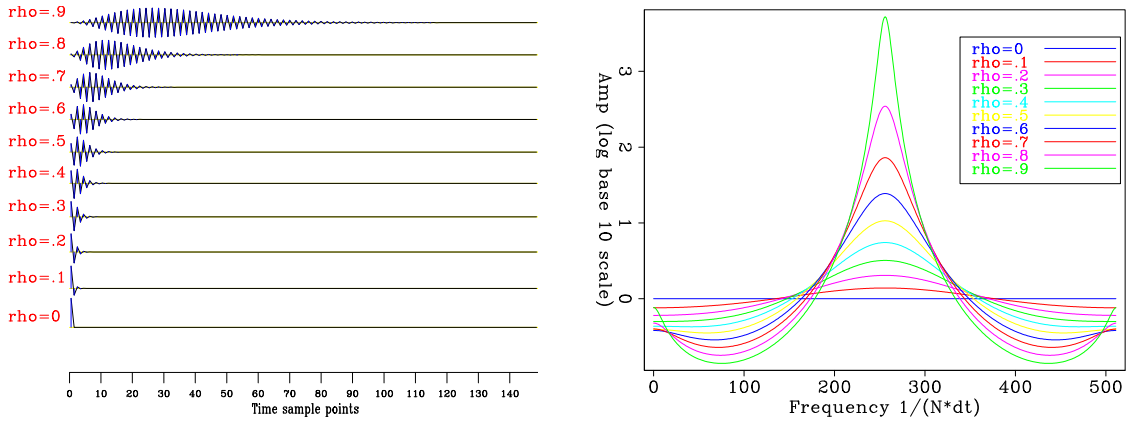


(a) (b)

Figure 4: The fourth-order half-Ricker wavelet with different ρ values: (a) in the time domain; (b) in the frequency domain. [**ER**] qiang/. half-ricker-with-rho,half-ricker-with-rho-freq

two methods should be the same. For less accuracy problem, we always use the frequency method in this paper.

Finally we obtain the approximate inverse Ricker wavelet. Figures 5(a) and 5(b) show the inverse fourth-order causal half-Ricker wavelet ($\hat{f}_a$) in the time and frequency domains with different ρ values. The inverse anti-causal half-Ricker wavelet ($\hat{f}_a^r$) can be generated by reversing $\hat{f}_a$ along the time axis. $\hat{f}_a$ and $\hat{f}_a^r$ will be our initial guesses for bidirectional deconvolution.



(a) (b)

Figure 5: The inverse fourth-order causal half-Ricker wavelet ($\hat{f}_a$) with different ρ values: (a) in the time domains ; (b) in the frequency domain. [**ER**] qiang/. freq-domain-a,freq-domain-a-freq

To test how good our approximation of the inverse Ricker wavelet is, we convolve our fourth-order approximate Ricker wavelet (with no ρ factor) with our inverse Ricker wavelet (equation 12).

$$\text{test result} = \text{Ricker} * \hat{f}_a * \hat{f}_b^r \tag{12}$$

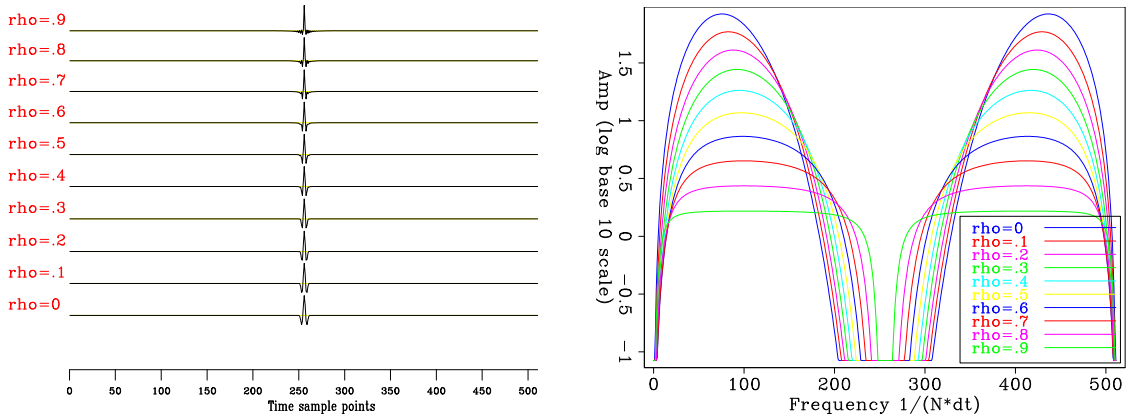
If our inversion is good, the output should be a spike in the time domain and a flat spectrum in the frequency domain.

Figures 6(a) and 6(b) show the test result in the time domain and the frequency domain with different ρ values. We see for $\rho = .9$ we get a good result. In particular, the frequency spectra are flatten as ρ increases.

Large ρ value can yield even better results. For large ρ, the frequency spectra are almost flat, except for two zero-value points (at 0 frequency and the Nyquist frequency). Figures 7(a) and 7(b) show these results.

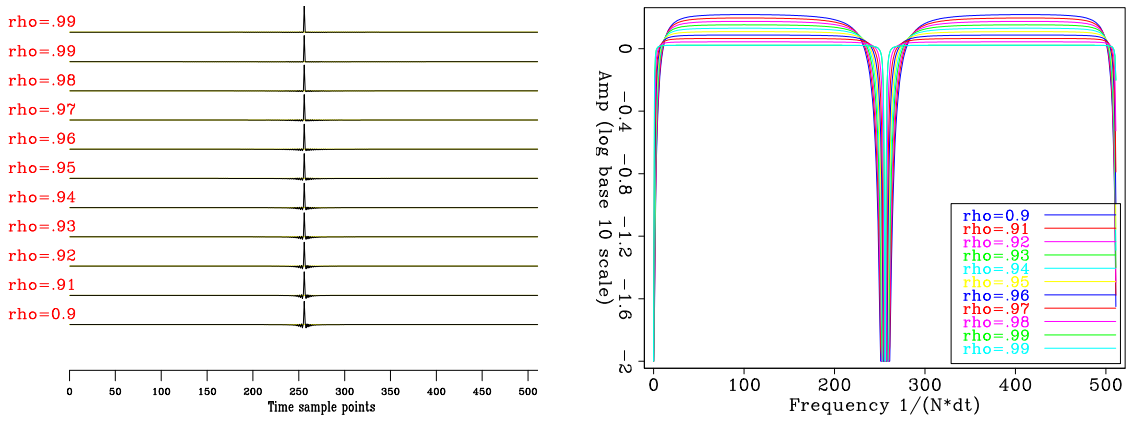
However, for our goal (to use the approximate inverse Ricker wavelet as an initial guess for bidirectional deconvolution), we do not need such large ρ values. With large ρ values, the filter will be long and can easily to lead to an unstable result in our deconvolution scheme.

We now have an approximate inverse of the Ricker wavelet separated into two symmetric parts. This matches our requirements for initial guesses for filters f_a and f_b^r. In the next section we will test the approximation using synthetic and field data.



(a) (b)

Figure 6: The test results with different ρ values: (a) in the time domain; (b) in the frequency domain. [**ER**] qiang/. freq-domain-result,freq-domain-result-freq



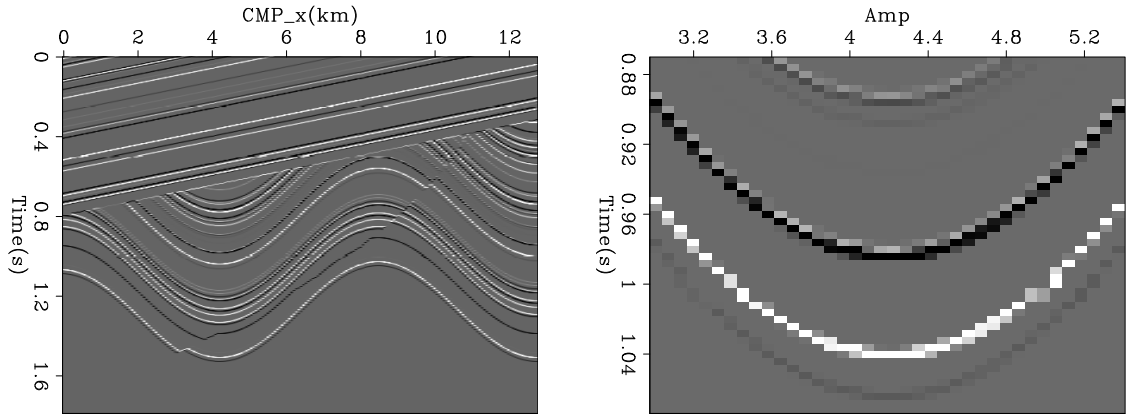
(a) (b)

Figure 7: The test results with large ρ values: (a) in the time domain; (b) in the frequency domain. [**ER**] qiang/. freq-domain-9-99-result,freq-domain-9-99-result-freq

EXAMPLES

Synthetic data example

First we test our initial guess on synthetic data. Because we hope to improve the deconvolution of data with a Ricker-like wavelet, we will test the synthetic data with a Ricker wavelet. We use a 2D reflectivity model to generate our synthetic data. Figure 8(a) shows the model we used, and Figure 1(a) shows our approximate Ricker wavelet (fourth-order) that we used to generate our data. The final synthetic data is shown in Figure 9(a). The wavelet we used is causal, hence we have some time lag in our synthetic data then the spiky model. We use the same deconvolution filter for all traces in our synthetic data example. Figures 10(a) and 10(b) are the result of bidirectional deconvolution using a spike as initial filters for f_a and f_b^r. Figures 11(a) and 11(b) are the result of deconvolution using our approximate inverse Ricker wavelet with ρ=0.7. We see both of the initial guesses (the simple spike and our inverse Ricker wavelet) did a reasonably good job if we see the global view. If we examine the result carefully, we find that the result with the inverse Ricker wavelets as initial guesses has less ghosting, especially in the vicinity of 1.0 s to 1.2 s and 3 km to 5.4 km. This may be more obvious if we see the magnified views of the results. We can also compare them to the magnified views of the synthetic data (Figure 9(b)) and the model (Figure 8(b)), and we find both of bidirectional deconvolution results improve the resolution of events. Figures 12(a) and 12(b) are magnified wiggle plots in the same time window but for just a single trace in the vicinity of 4000 m of CMP_x, from which we can see the the result with inverse Ricker initial guess has less ghosting in wiggle. Although there is improvement, it is not very significant for our synthetic case.

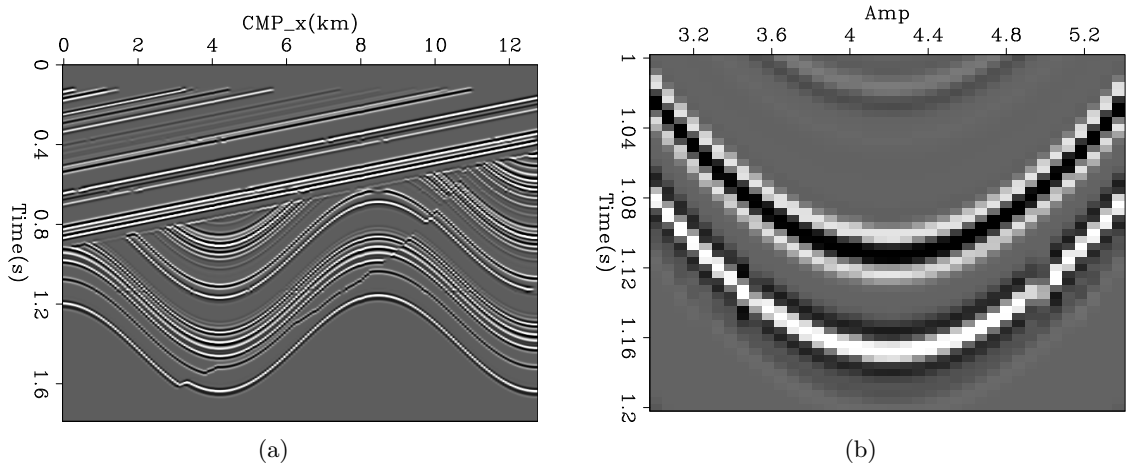


(a) (b)

Figure 8: 2D reflection model we used to to generate our synthetic data: (a) the global view; (b) the magnified view in the time window from .872 s to 1.072 s and in CMP_x range from 3 km to 5.4 km. [**ER**] qiang/. 2d-synthetic-model,2d-synthetic-model-local

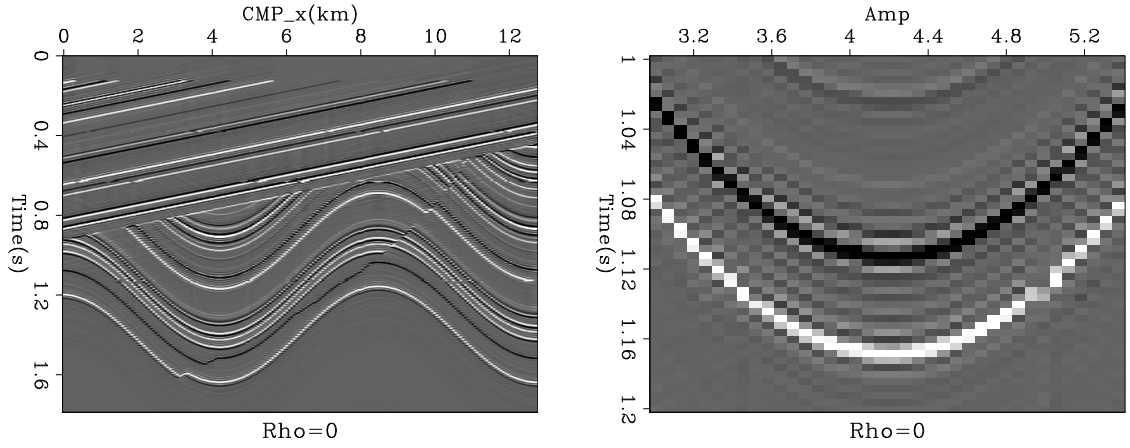
Field data example

We also test our initial guess on a marine common-offset 2D data set. We use the same deconvolution filter for all traces in our field data example as well. Figure 13 shows the 2D



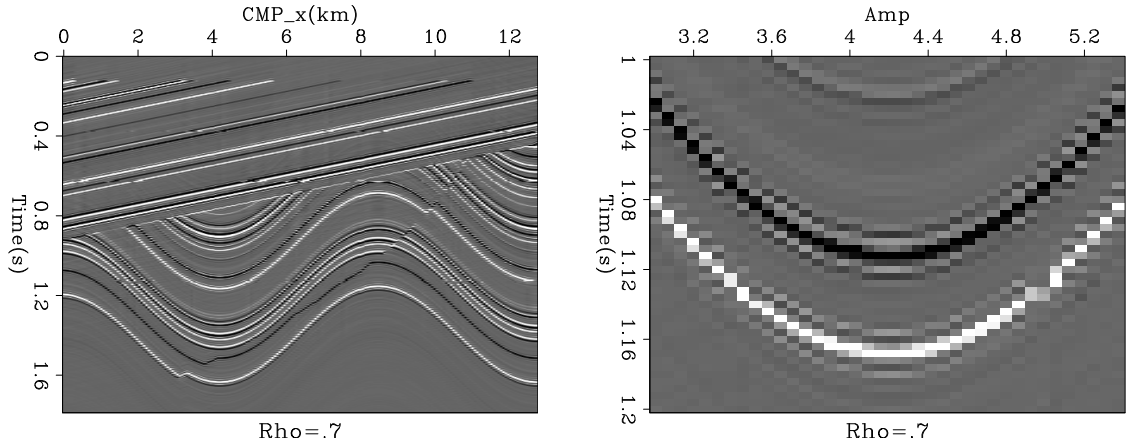
(a) (b)

Figure 9: The synthetic data we used for our test: (a) the global view; (b) the magnified view in the time window from 1 s to 1.2 s and in CMP_x range 3 km to 5.4 km. [**ER**] qiang/. 2d-synthetic-data,2d-synthetic-data-local



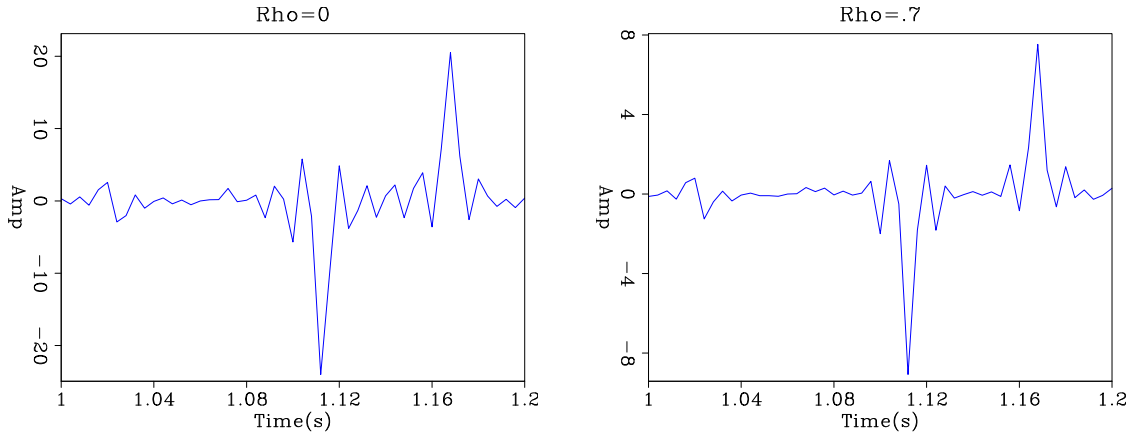
(a) (b)

Figure 10: The result of bidirectional deconvolution using a spike as initial filters f_a and f_b^r on synthetic data: (a) the global view; (b) the magnified view in the time window from 1 s to 1.2 s and in CMP_x range 3 km to 5.4 km. [**ER**] qiang/. result-rho-0,result-local-rho-0



(a) (b)

Figure 11: The result of bidirectional deconvolution using our approximate inverse Ricker wavelet with ρ=0.7 on synthetic data: (a) the global view; (b) the magnified view in the time window from 1 s to 1.2 s and in CMP_x range 3 km to 5.4 km. [**ER**] qiang/. result-rho-7,result-local-rho-7



(a) (b)

Figure 12: The magnified wiggle plot of the bidirectional deconvolution results at cmp=4.3 km in the time window from 1 s to 1.2 s: (a) with a spike as the initial guess; (b) with an inverse Ricker wavelet as the initial guess. [**ER**] qiang/. wiggle-local-rho-0,wiggle-local-rho-7

input data set. Figure 14 shows the bidirectional deconvolution result with a simple spike as the initial guesses and Figure 15 shows the bidirectional convolution with our inverse Ricker wavelet as the initial guesses, using ρ=0.8.

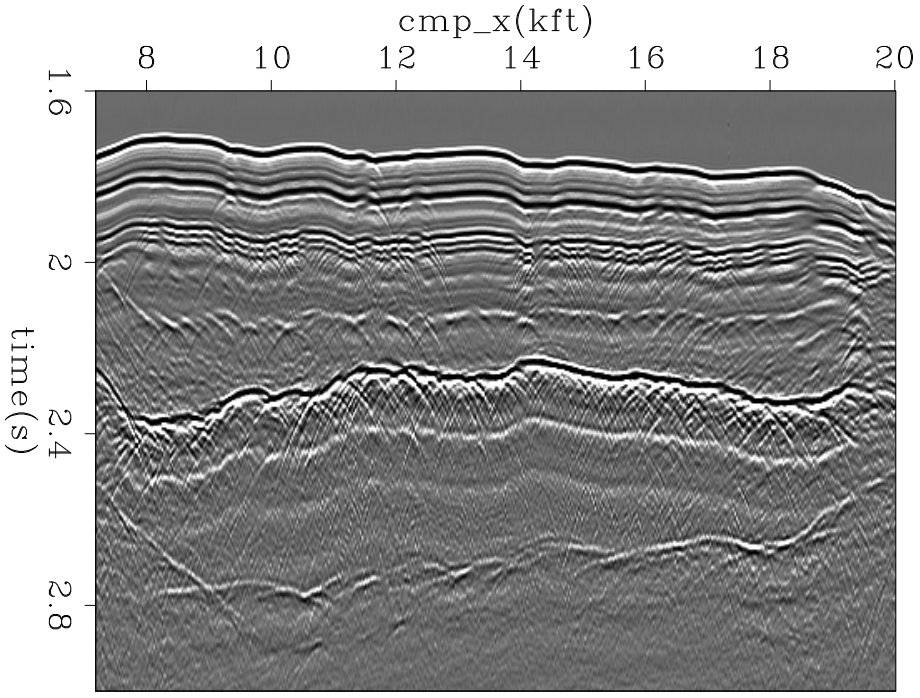


Figure 13: A marine common-offset 2D data set we used for our test. [**ER**] qiang/. 2d-field-data

It is obvious that the result of using the inverse Ricker wavelet as an initial guess is much better. For example, the ghost events before the first reflection are much suppressed, and the upper boundary of the salt body, which is the event from about 2.2 s to 2.4 s, is condensed. There appears to be a polarity change plus a time shift in our two bidirectional deconvolution output results with different initial guesses. This is not an error. Given the wavelets estimated from deconvolution process, it is a reasonable result.

Figures 16(a) and 16(b) are the wavelets estimated by a simple spike initial guess and our inverse Ricker wavelet initial guess with ρ=0.8. On these wavelet, we can clearly see the airgun bubbles. And the time interval between bubbles are corelate the data shown in 13 very well. The wavelets are causal-like, or in other words, we do not have too much energy in the negative time before the first major impulse. These two evidence indicate that we estimate the wavelet of our field data correctly.However, we still have some problem here. We must mention that we tried to enlarge the filter length to get a longer estimated wavelet by the same scheme, but we got a noisy result. It seems we have some instability here, but we do not know the reason of this now. We are researching this and will explain the reason in furtrue report after we figure it out.

The two wavelets in Figures 16(a) and 16(b) look very similar, but have opposite polarity and a different origin in the time axis. The flipped polarity is caused by the origin shift. Since we have the same input data and wavelet for deconvolution, even with different initial guesses, we should get the same (or at least similar) estimated wavelet and deconvolution

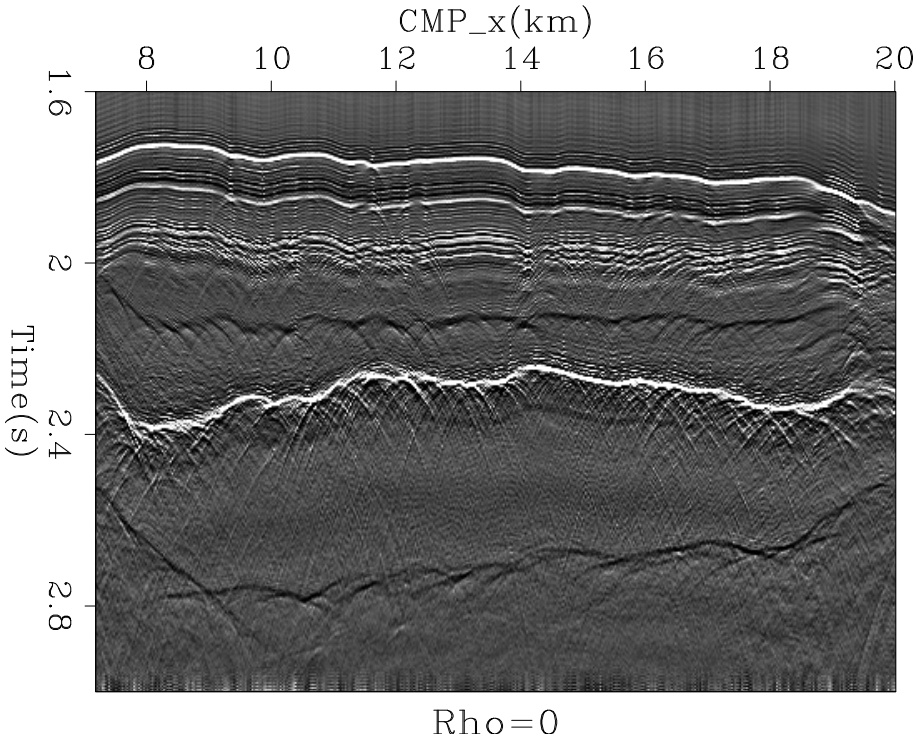


Figure 14: The result of deconvolution using a spike as initial filters f_a and f_b^r on field data. [**ER**] qiang/. result-0

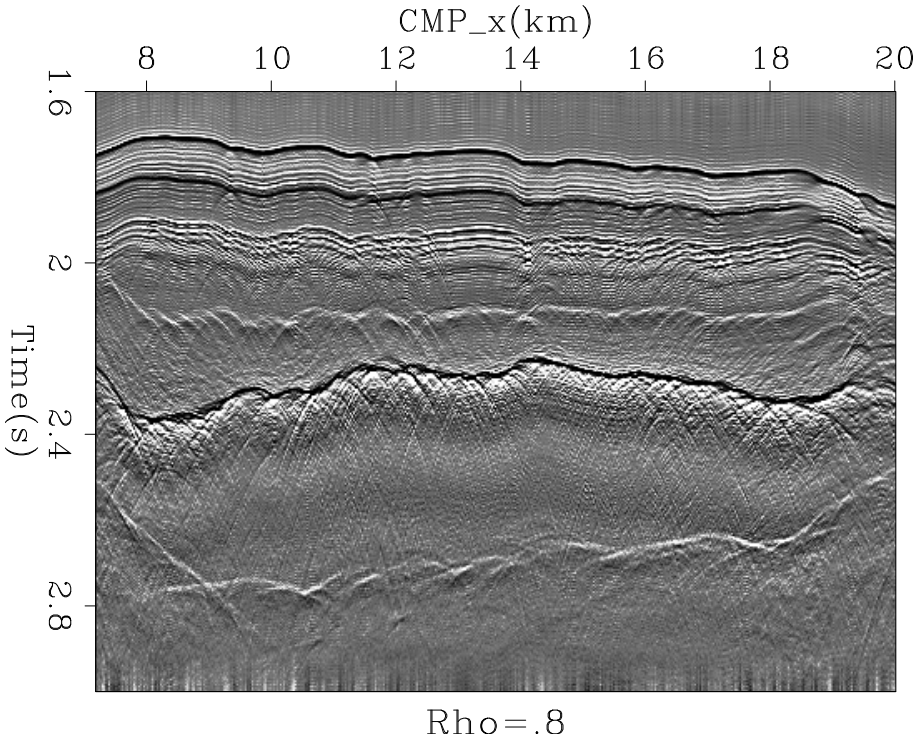


Figure 15: The result of deconvolution using our approximate inverse Ricker wavelet with ρ=0.8 on field data. [**ER**] qiang/. result-8

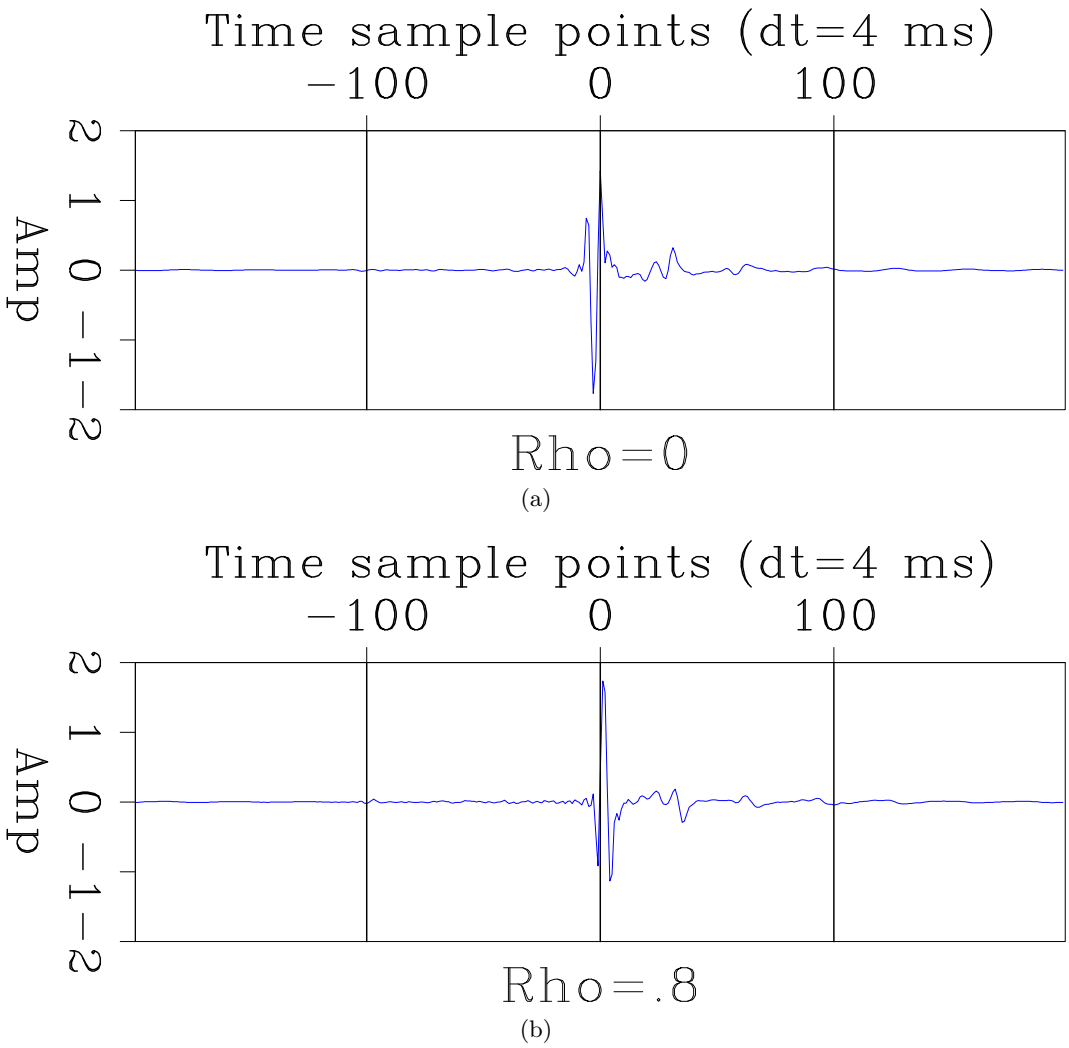


Figure 16: Estimated wavelet from bidirectional deconvolution on field data: (a) using a spike as initial filters f_a and f_b^r; (b) using our approximate inverse Ricker wavelet with ρ=0.8. [**ER**] qiang/. wavelet-0,wavelet-8

filter after the process. Bidirectional deconvolution forces the polarity of both the estimated wavelet and the deconvolution filter to be positive at the origin of the time axis. So if the origin of the estimated wavelet has been shifted from positive to negative, the polarity will be flipped. This estimated wavelet origin shift will cause the output of the deconvolution result to shift on the time axis as well.

CONCLUSIONS

We tested a numerical inversion of the Ricker wavelet as an initial guess for a bidirectional deconvolution scheme. This Ricker wavelet inversion has two symmetric parts, which makes it suitable to be the initial guess. We tested our inversion of the Ricker wavelet on a field data example, and the the inversion of the Ricker wavelet initial guess did improve the

bidirectional deconvolution result (with a suitable ρ parameter). We found there is some instability in our scheme if we use different filter length for our deconvolution filter and do not have a explanation of it. We will continue researching on this.

FUTURE WORK

Preconditioning is a well-established technique in linear problems to hasten convergence by utilizing prior information. In non-linear problems such as this, preconditioning may be essential to guide the descent along sensible pathways thus avoiding potential local minima. By adding prior information in the problem, the preconditioning can not only accelerate the convergence but also has the ability to stabilize the result. Because we are facing a unstable problem in our deconvolution implementation, it is our plan to add a preconditioning in the bidirectional deconvolution method for next step. We hope it will help us to increase both the stability and convergence speed of bidirectional deconvolution method.

ACKNOWLEDGMENTS

We would like to thank Yang Zhang, Antoine Guitton, Shuki Ronen, Mandy Wong and Elita Li for helpful discussion of our research.

REFERENCES

Shen, Y., Q. Fu, and J. Claerbout, 2011, A new algorithm for bidirectional deconvolution: SEP-Report, **143**, 271–282.

Zhang, Y. and J. Claerbout, 2010, A new bidirectional deconvolution method that overcomes the minimum phase assumption: SEP-Report, **142**, 93–104.

A log spectral approach to bidirectional deconvolution

Jon Claerbout, Qiang Fu, and Yi Shen

ABSTRACT

The blind-deconvolution problem for non-minimum-phase-source was established in the time domain. This is a Fourier domain formulation. Changing variables from $A(Z)B(1/Z)$ to $U(Z) = \ln(A(Z)B(1/Z))$ leads to a different kind of whiteness — the output being orthogonal to its shifted soft clip.

INTRODUCTION

We three have been impressed by the excellent field data results of Zhang and Claerbout (2010), making us feel that proper blind deconvolution will give much more reliable presentation of seismogram polarity, and hopefully impedance. Additionally Fu et al. (2011) have estimated a field-data shot waveform of remarkable plausibility. Exploring with programs similar to Zhang and Claerbout (2010), however, we discovered a variety of unexpected irregularities that have sent us back to many investigations with synthetic data. Indeed, we now have three quite independent approaches. We believe the heart of the problem lies in the fundamental non-linearity of the basis of the formulation. We are not aware of any realistic linear formulation of it. To allieviate the difficulties we have (1) come up with a much improved starting solution in the form of Ricker wavelet inverses, (2) come up with a smooth "simultaneous" descent method, and (3) come up with a log spectral formulation. Unfortunately, we are not out of the woods yet. Hopefully, adding preconditioning to our codes will guide us away from spurious solutions enabling us to make more firm conclusions about the theory which would then allow us to plunge confidently into many field data investigations.

THE LOG SPECTRAL APPROACH

A minimum phase wavelet can be made from any causal wavelet by taking it to Fourier space, and exponentiating. The proof is straightforward: Let $U(Z) = 1 + u_1 Z + u_2 Z^2 + \cdots$ be the Z transform ($Z = e^{i\omega}$) of any causal function. Then $e^{U(Z)}$ will be minimum phase. Although we would always do this calculation in the Fourier domain, the easy proof is in the time domain. The power series for an exponential $e^U = 1 + U + U^2/2! + U^3/3! + \cdots$ has no powers of $1/Z$, and it always converges because of the powerful influence of the denominator factorials. Likewise e^{-U}, the inverse of e^U, always converges and is causal. Thus both the filter and its inverse are causal. Q.E.D.

We seek to find two functions, one strictly causal the other strictly anticausal (nothing at $t = 0$).

$$U^+ = u_1 Z + u_2 Z^2 + \cdots \tag{1}$$

$$U^- = u_{-1}/Z + u_{-2}/Z^2 + \cdots \tag{2}$$

Notice U, U^2, etc do not contain Z^0. Thus the coefficient of Z^0 in $e^U = 1 + U + U^2/2! + \cdots$ is unity. Thus $a_0 = b_0 = 1$.

$$e^{U^+} = A = 1 + a_1 Z + a_2 Z^2 + \cdots \tag{3}$$

$$e^{U^-} = B = 1 + b_1/Z + b_2/Z^2 + \cdots \tag{4}$$

Define $U = U^- + U^+$. The decon filter is $AB = e^U$ and the source waveform is its inverse e^{-U}. With the Fourier transform of the data $D(\omega)$, the decon output is:

$$\mathbf{r} = (r_t) = \mathrm{FT}^{-1}\ D(\omega)\ e^{U(Z(\omega))} \tag{5}$$

where U is found with a penalty function, our choice being the hyperbolic penalty function.

$$\text{argmin(U)} = \text{hyp}(\mathbf{r}) = \sum_t H(r_t) \tag{6}$$

where $H(r) = \sqrt{r^2 + R^2} - R$, and R is the ℓ_1/ℓ_2 threshold parameter.

Take the gradient of the penalty function assuming there is only one variable, u_3 giving a single regression equation:

$$0 \approx \sum_t \frac{\partial H}{\partial r}\frac{\partial r}{\partial u_3} = \frac{\partial r}{\partial u_3}\frac{\partial H}{\partial r} \tag{7}$$

$$0 \approx \sum_t (\mathrm{FT}^{-1}\ D(\omega)\ \frac{\partial}{\partial u_3}\ e^{U(Z)})_t\ H'(r_t) \tag{8}$$

$$0 \approx \sum_t (\mathrm{FT}^{-1}\ D(\omega)\ Z^3\ e^{U(Z)})_t\ H'(r_t) \tag{9}$$

so the deconvolution output selected at time $t + 3$ multiplies $H'(r_t)$ (also known as the "soft-clip" function).

Equation (9) requires us to do an inverse Fourier transform to find the gradient for only u_3. For u_4 there is an analogous expression, but it is time shifted by Z^4 instead of Z^3. Clearly we need only do one Fourier transform and then shift it to get the time function required for other filter lags. Thus the gradient for all nonzero lags is:

$$0 \approx \Delta\mathbf{u} = (\Delta u_\tau) = (\sum_t\ r_{t+\tau}\ H'(r_t)) \tag{10}$$

$$\Delta U = \overline{\mathrm{FT}(\mathbf{r})}\ \mathrm{FT}(\text{softclip}(\mathbf{r})) \tag{11}$$

where τ measures some filter lag. Actually, equation (11) is wrong as it stands. Conceptually it should be brought into the time domain and have Δu_0 set to zero. More simply, the mean can be removed in the Fourier domain.

Equation (10) says if we were doing least squares, the gradient would be simply the autocorrelation of the residual. When the gradient at nonzero lags drops to zero, the residual is white. Hooray! We long understood that limit. What is currently new is that we now have a two-sided filter. Likewise the ℓ_1 limit must be where the output is uncorrelated with the clipped output at all lags but zero lag.

(I'm finding it fascinating to look back on what we did all these years with the causal filter $A(Z)$ and comparing it to the non-causal exponential filter $e^{U(Z)}$. In an ℓ_2 norm

world for filter $A(Z)$ we easily saw the shifted output was orthogonal to the fitting function *input*. For filter $e^{U(Z)}$ we easily see now the shifted output is orthogonal to the *output*. The whiteness of the output comes easily with $e^{U(Z)}$ but with $A(Z)$ the Claerbout (2011) contains a lengthy and tricky proof of whiteness.)

Let us figure out how a scaled gradient $\alpha\Delta\mathbf{u}$ leads to a residual modification $\alpha\Delta\mathbf{r}$. The expression e^U is in the Fourier domain. We first view a simple two term example.

$$e^{\alpha\Delta U} = e^{\alpha(\Delta u_1 Z + \Delta u_2 Z^2)} \tag{12}$$

$$e^{\alpha\Delta U} = 1 + \alpha(\Delta u_1 Z + \Delta u_2 Z^2) + \alpha^2(\cdots) \tag{13}$$

$$\mathrm{FT}^{-1}\ e^{\alpha\Delta U} = (1, \alpha\Delta u_1, \alpha\Delta u_2) + \alpha^2(\cdots) \tag{14}$$

With that background, ignoring α^2, and knowing the gradient $\Delta\mathbf{u}$, let us work out the forward operator to find $\Delta\mathbf{r}$. Let "$*$" denote convolution.

$$\mathbf{r} + \alpha\Delta\mathbf{r} = \mathrm{FT}^{-1}(De^{U+\alpha\Delta U}) \tag{15}$$

$$= \mathrm{FT}^{-1}(De^{U}e^{\alpha\Delta U}) \tag{16}$$

$$= \mathrm{FT}^{-1}(De^{U}) * \mathrm{FT}^{-1}(e^{\alpha\Delta U}) \tag{17}$$

$$= \mathbf{r} * (1, \alpha\Delta\mathbf{u}) \tag{18}$$

$$= \mathbf{r} + \alpha\mathbf{r} * \Delta\mathbf{u} \tag{19}$$

In familiar ℓ_2 problems we would find α as $\alpha = -(\mathbf{r}\cdot d\mathbf{r})/(d\mathbf{r}\cdot d\mathbf{r})$, now we must find it by

$$\mathrm{argmin}(\alpha) = H(\mathbf{r} + \alpha\mathbf{r} * \Delta\mathbf{u}) \tag{20}$$

This we do by the Newton method which iteratively fits the hyperbola to a parabola.

```
iterate {
  alpha = - (Sum_i H'(r_i) dr_i ) / ( Sum H''(r_i) dr_i^2 )
  r = r + alpha dr
  }
```

In the pseudocode above, in the ℓ_2 limit $H'(r) = r$ and $H''(r) = 1$, so so the first iteration gets the correct α, and changes the residual accordingly so all subsequent α values are zero.

We are not finished because we need to assure the constraint $u_0 = 0$. In a linear problem it would be sufficient to set $\Delta u_0 = 0$, but here we soon do a linearization which breaks the constraint. In the frequency domain the constraint is $\sum_\omega U = 0$. We meet this constraint by inserting a constant β in equation (15) and chosing β to get a zero sum over frequency of $U + \alpha\Delta U + \beta$. Let $\sum$ denote a normalized summation over frequency. By normalized, I mean $\sum \beta = \beta$. We must choose β so that $0 = \sum U + \alpha \sum \Delta U + \beta$. Clearly, $\beta = -\sum U - \alpha \sum \Delta U$. Pick up again from equation (15) including β.

$$\mathrm{argmin}(\alpha) = H(\ \mathrm{FT}^{-1}(De^{U}\ e^{\alpha\Delta U}\ e^{\beta})\ ) \tag{21}$$

$$= H(\ \mathrm{FT}^{-1}(De^{U}\ e^{\alpha\Delta U}\ e^{-\sum U}\ e^{-\alpha\sum \Delta U})\ ) \tag{22}$$

$$= H(\ \mathrm{FT}^{-1}(De^{U-\sum U}\ \ e^{\alpha(\Delta U - \sum \Delta U)})\ ) \tag{23}$$

$$= H(\ \mathrm{FT}^{-1}(De^{U-\sum U})\ *\ \mathrm{FT}^{-1}(e^{\alpha(\Delta U - \sum \Delta U)})\ ) \tag{24}$$

Proceed now along the lines of equation (17) through (20), but with means removed in Fourier space leading to slightly different vectors $\tilde{\mathbf{r}}$ and $\Delta\tilde{\mathbf{u}}$

$$\mathrm{argmin}(\alpha) = H(\ \tilde{\mathbf{r}} + \alpha\tilde{\mathbf{r}} * \Delta\tilde{\mathbf{u}}\ ) \tag{25}$$

which is solveable by the method of the same pseudocode above.

ALGORITHM

Here we fill in more details of the algorithm. After we are certain of its behavior we would naturally switch over to conjugate directions.

```
D(omega,x) = FT d(t,x)
u=0;
iteration {
        U = FT(u)
        remove mean from U(omega)
        exp(U(Z))
        dU = 0
        for all x
                r(t,x) = IFT( D exp(U) )
                softclip( r )
                dU += conjg(FT(r)) * FT(softclip)        #  "*" means multiply
        remove mean from dU(omega)
        for all x
                dR = FT(r) * dU                          #  "*" means multiply
                dr = IFT(dR)
        argmin(alpha) = H(r+alpha*dr)
        u = u + alpha du
        }
```

CONCLUSIONS

It is too early to draw reliable conclusions about this theory. It arose a few weeks before the progress report deadline, and got coded a few days before. The coding (steepest descent) illuminated a bug in the initial theory regarding the constraint. A few test cases were successfully run, but we do not know how this method compares to Shen et al. (2011) and Fu et al. (2011).

The problem formulation itself is nonlinear and thereby susceptible to many minima, mostly bad ones. We don't know whether preconditioning applies in the usual way to this method. Perhaps some kind of tapering of Δu_t is a means of directing the method towards a more appropriate optimum, but for now this is no more than speculation.

REFERENCES

Claerbout, J. F., 2011, Image Estimation by Example: `http://sepwww.stanford.edu/sep/prof/`.

Fu, Q., Y. Shen, and J. Claerbout, 2011, An approximation of the inverse ricker wavelet as an initial guess for bidirectional deconvolution: SEP-Report, **143**, 283–296.

Shen, Y., Q. Fu, and J. Claerbout, 2011, A new algorithm for bidirectional deconvolution: SEP-Report, **143**, 271–282.

Zhang, Y. and J. Claerbout, 2010, A new bidirectional deconvolution method that overcomes the minimum phase assumption: SEP-Report, **142**, 93–104.

SEP PHONE DIRECTORY

Name	Phone	Login Name
Almomin, Ali	723-0463	ali
Ayeni, Gboyega	723-6006	gayeni
Barak, Ohad	723-9282	ohad
Berryman, James	–	berryman
Biondi, Biondo	723-1319	biondo
Claerbout, Jon	723-3717	jon
Clapp, Bob	725-1334	bob
de Ridder, Sjoerd	723-1250	sjoerd
Fu, Qiang	723-3187	qiang
Guitton, Antoine	–	antoine
Halpert, Adam	723-6006	adam
Lau, Diane	723-1703	diane
Leader, Christopher	723-0463	chrisl
Li, Elita	723-9282	myfusan
Shen, Xukai	723-0463	xukai
Shen, Yi	723-3187	yishen
Tang, Yaxun	723-1250	tang
Taweesintananon, Kittinat	723-3187	kittinat
Wong, Mandy	723-9282	mandyman
Zhang, Yang	723-1250	yang

Dept fax number: (650) 725-7344

E-MAIL

Our Internet address is "*sep.stanford.edu*"; i.e., send Jon electronic mail with the address "*jon@sep.stanford.edu*".

WORLD-WIDE WEB SERVER INFORMATION

Sponsors who have provided us with their domain names are not prompted for a password when they access from work. If you are a sponsor, and would like to access our restricted area away from work, visit our website and attempt to download the material. You will then fill out a form, and we will send the username/password to your e-mail address at a sponsor company.

STEERING COMMITTEE MEMBERS, 2010-2011

Name	Company	Tel #	E-Mail
Raymond Abma	BP	(281)366-4604	abmar1@bp.com
Biondo Biondi	SEP	(650)723-1319	biondo@sep.stanford.edu
Robert Bloor	ION/GX Technology	(281)781-1141	robert.bloor@iongeo.com
Luca Cazzola (Co-chair, 1st year)	Eni E&P	–	luca.cazzola@eni.com
Jon Claerbout	SEP	(650)723-3717	jon@sep.stanford.edu
Richard Cook	Shell	(713)245-7195	richard.cook@shell.com
Stewart A. Levin	Halliburton	(303)488-3062	stewart.levin@halliburton.com
Yi Luo	Saudi Aramco	–	yi.luo@aramco.com
Marta Woodward (Co-chair, 2nd year)	WesternGeco	(832)798-6204	martajo@slb.com

Research Personnel

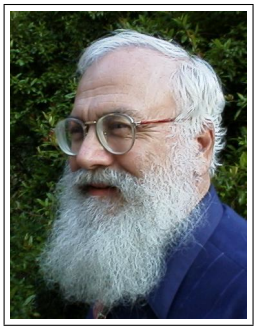
Ali Almomin graduated from Texas A&M University in 2007 with a BS in Geophysics. Then, he joined Saudi Aramco and worked in several exploration and research departments with a focus on 3D seismic processing and near surface imaging. He joined Stanford Exploration Project in 2009 to pursue a PhD in Geophysics and is currently working on seismic tomography. He is a member of SEG, EAGE, and SPE.

Gboyega Ayeni received his B.Sc. in Applied Geophysics from Obafemi Awolowo University, Nigeria in 2004. He was a Shell sholar at University of Leeds, where he obtained an M.Sc with Distinction in Exploration Geophysics. Gboyega joined SEP in September 2006 to work towards his Ph.D in Geophysics. He is a member of SEG, EAGE, AGU, SPE and AAPG.

Ohad Barak received a B.Sc. (2006) and an M.Sc. (2009) in Geophysics from Tel-Aviv University. In 2008 he joined Paradigm Geophysical and worked there as a developer. He joined SEP in 2009 and is currently pursuing a Ph.D. in geophysics at Stanford University, and a longer biography.

Biondo L. Biondi is professor of Geophysics at Stanford University. Biondo graduated from Politecnico di Milano in 1984 and received an M.S. (1988) and a Ph.D. (1990) in geophysics from Stanford. He is co-director of the Stanford Exploration Project and of the Stanford Center for Computational Earth and Environmental Science. In 2004 the Society of Exploration Geophysicists (SEG) has honored Biondo with the Reginald Fessenden Award. Biondo recently published a book, 3-D Seismic Imaging, that is the first text book to introduce the theory of seismic imaging from the 3-D perspective. The book is published by SEG in the Investigations in Geophysics series. During 2007 gave a one-day short course in 28 cities around the world as the SEG/EAGE Distinguished Short Course Instructor (DISC) . He is a member of AGU, EAGE, SEG and SIAM.

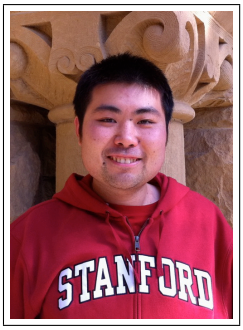
Jon F. Claerbout (M.I.T., B.S. physics, 1960; M.S. 1963; Ph.D. geophysics, 1967), professor at Stanford University, 1967. Emeritus 2008. Best Presentation Award from the Society of Exploration Geophysicists (SEG) for his paper, *Extrapolation of Wave Fields.* Honorary member and SEG Fessenden Award "in recognition of his outstanding and original pioneering work in seismic wave analysis." Founded the Stanford Exploration Project (SEP) in 1973. Elected Fellow of the American Geophysical Union. Authored three published books and five internet books. Elected to the National Academy of Engineering. Maurice Ewing Medal, SEG's highest award. Honorary Member of the European Assn. of Geoscientists & Engineers (EAGE). EAGE's highest recognition, the Erasmus Award.



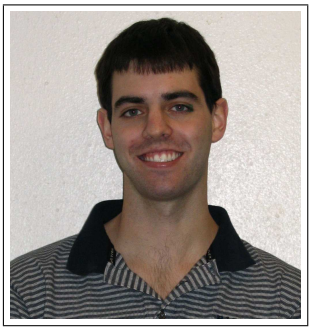
Robert Clapp received his B.Sc. (Hons.) in Geophysical Engineering from Colorado School of Mines in May 1993. He joined SEP in September 1993, received his Masters in June 1995, and his Ph.D. in December 2000. He is a member of the SEG and AGU.

Joe Dellinger was born in Tulsa, spent his teenage years in Dallas and graduated from Texas A&M University in 1983 with B.S. degrees in Mathematics and Geophysics. He then joined SEP and graduated with a Ph.D. in Geophysics from Stanford in 1991. While at Stanford Joe worked on a variety of subjects, but the ones that left a lasting legacy are his work on anisotropy and building SEP's graphical software infrastructure. Joe did a post-doc at the University of Hawaii 1991-1994. The algorithm he invented while there to attempt to salvage a multicomponent VSP with bad coupling is now widely used for processing multicomponent data in the industry. In 1994-1999 Joe worked at Amoco's research laboratory in Tulsa, Oklahoma (all of 100 yards from where he first lived). The tides of the great oil-industry consolidations swept him to BP in Houston in 1999 where he remains to the present day.

Qiang Fu received his Bachelors degree in Computer Science from Beijing Institute of technology, China in 2003. From 2007 to 2010 he worked for Saudi Aramco in Saudi Arabia. He joined SEP in the fall of 2010 and is currently working toward a M.S. in Geophysics at Stanford University. He is a student member of the SEG.

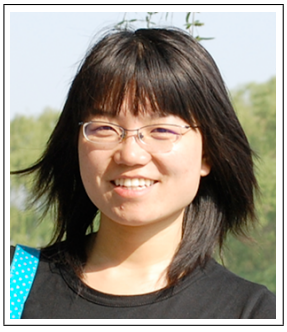


Adam Halpert graduated from Rice University in May 2007 with a B.Sc. in Earth Science and a B.A. in Policy Studies. He joined SEP in the summer of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG and AGU.



Chris Leader graduated from Oxford Univeristy in 2008 with a BA in Physics (with concentration on Astrophysics and Condensed Matter physics) and then from Imperial College London in 2009 with an MSc in Petroleum Geophysics (Distinction). He is currently a first year student in the Stanford Exploration Project on the PhD program working on Fourier methods of regularisation. Work experience involves 3D seismic processing for a Rio Tinto acquired dataset over summer 2009. He is a member of SEG, EAGE, PESGB and IOP.

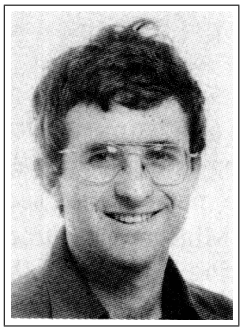
Yunyue (Elita) Li graduated from China University of Petroleum, Beijing in July 2008 with a B.S. in Information and Computational Science. She joined SEP in the fall of 2008, and is currently working toward a Ph.D. in Geophysics. She is a student member of the SEG.



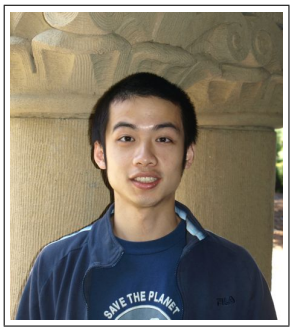
Dave Nichols received a B.A. in physics from Cambridge University in 1982; a M.Sc. in geophysics from Imperial College, London in 1983; and a Ph.D. in geophysics from Stanford University in 1994. From 1983 to 1987, he was employed by Western Geophysical in London and joined SEP in 1987. During the summer of 1991, he worked for Chevron Oil Field Research Company in La Habra, California. He currently works for Schlumberger in the Schlumberger Reservoir Completions Center in Rosharon, Texas.

Sjoerd de Ridder received his B.Sc. (Dec. 2004) in Earth Sciences from Utrecht University in the Netherlands. After an exchange year at Colorado School of Mines he attended Delft University of Technology. In August of 2007, he obtained an M.Sc. in Applied Earth Sciences with specialization in applied geophysics from Delft University of Technology for his thesis: 'Simulation of interferometric seismoelectric Greens function recovery'. He joined SEP in the fall of 2007 and is pursuing a Ph.D. in Geophysics. He received the 'Jon F. Claerbout - Chevron Fellowship' in 2009-2010. During the summer of 2010 he did an internship at China University of Petroleum in Beijing. Currently he is working on exploration seismology with low-frequency ambient seismic noise. In his free time he likes to run, bike and hike while traveling and discussing phylosophy, politics and history.

Shuki Ronen is a geophysicist with Geco-Prakla working on various aspects of seismic data processing and acquisition. Previously he worked with GeoQuest on seismic data interpretation; with Schlumberger on reservoir characterization; with the Institute of Petroleum Research and Geophysics on seismic data processing and acquisition; with the Colorado School of Mines as a visiting professor; and with Saxpy Computer company as an engineer. He has a Ph.D. from Stanford in Geophysics, and a B.Sc. in Physics and Geology from Hebrew University. (Photo: December 1985, Geophysics, p. 2919)



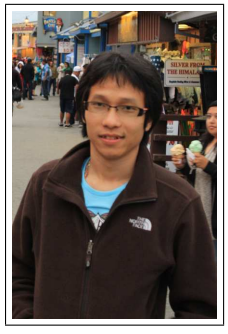
Xukai Shen graduated from Tsinghua University in July 2007 with a B.E. in Electrical Engineering. He joined SEP in the fall of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG.



Yi Shen got her Bachelors degree in Acoustics, from Dept. of Electronic Science and Engineering, Nanjing University, China in 2010. She joined SEP in the fall of 2010, and is currently working toward a Ph.D. in Geophysics at Stanford University. She is a student member of the SEG.

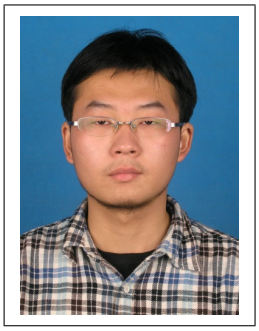
Yaxun Tang received his B.Sc. (Jul. 2003) and M.Sc. (Aug. 2005) in Geophysics from School of Ocean and Earth Science, Tongji University, Shanghai. He joined SEP in 2005 and is currently working towards a Ph.D. in Geophysics at Stanford University. He is a student member of SEG and EAGE.

Kittinat Taweesintananon received his B.Sc. (March 2009) in Physics from Chulalongkorn University, Thailand. He joined SEP in 2009 and is working toward an M.S. in Geophysics. During the time at Stanford, he is granted by PTT Exploration and Production plc., which is the national petroleum exploration company in Thailand. He is a student member of SEG.



Mandy Wong graduated in 2004 with a B.Sc. in Physics and Mathematics from the University of British Columbia (UBC) in Vancouver, Canada. In 2006, she obtained a M.Sc. degree in Condensed Matter Thoery at UBC. Afterward, Mandy worked for a geophysical consulting company, SJ Geophysics, based in Vancouver, Canada. Mandy joined SEP in 2008, and is working towards a Ph.D. in Geophysics. Her main research interest is imaging with multiples.

Yang Zhang graduated from Tsinghua University in July 2007 with a B.E. in Electrical Engineering. He took an internship in Microsoft Research Asia during 2007-2008. He joined SEP in 2009, and is currently pursuing a Ph.D. in Geophysics.



SEP ARTICLES PUBLISHED OR IN PRESS

Ayeni, G., Y. Tang, and B. Biondi, 2009, Joint preconditioned least-squares inversion of simultaneous source time-lapse seismic data sets: SEG Technical Program Expanded Abstracts, **28**, 3914–3918.

Ayeni, G. and Nasser, M., 2009, Optimized local matching of time-lapse seismic data: A case study from the Gulf of Mexico, SEG Technical Program Expanded Abstracts, **28**, 3939–3943.

Ayeni, G., and Biondi, B., 2009, Joint target-oriented wave-equation inversion of multiple time-lapse seismic data sets, SEG Technical Program Expanded Abstracts, **28**, 3825–3829.

Ayeni, G., Tang, Y., and Biondi, B., 2010, Efficient seismic monitoring of hydrocarbon reservoirs using multiple shooting vessels. In Proceedings of the 2010 Offshore Technology Conference.

Ayeni, G., and Biondi, B., 2010, Continuous reservoir monitoring with asynchronous simultaneous-source seismic data, 72nd Conference & Technical Exbition, EAGE, Extended Abstracts.

Ayeni, G., 2010, Seismic reservoir monitoring with permanent encoded seismic arrays: SEG Technical Program Expanded Abstracts **29**, 4221–4226.

Ayeni G., and Biondi, B., 2010, Target-oriented joint least-squares migration/inversion of time-lapse seismic data sets: Geophysics, **75**, no. 3, R61–R75.

Ayeni, G., 2011, Cyclic 1D matching of time-lapse seismic data sets: A case study of the Norne Field, SEG Technical Program Expanded Abstracts (Submitted).

Ayeni, G., and Biondi, B., 2011, Wave-equation inversion of time-lapse seismic data sets, SEG Technical Program Expanded Abstracts (Submitted).

Ayeni, G., Almomin, A. and Nichols, D., 2011, On the separation of simultaneous-source data by inversion, SEG Technical Program Expanded Abstracts (Submitted).

Guerra, C. and Biondi, B., 2011, Fast 3D velocity updates using the pre-stack exploding reflector model. SEG Technical Program Expanded Abstracts **29**, 4375–4379.

Guerra, C. and Biondi, B., 2010, Fast 3D velocity updates using the pre-stack exploding reflector model. Geophysics (Accepted).

Biondi, B., 2010, Velocity estimation by image focusing analysis: Geophysics, **75**, pp. 49–60.

Biondi, B., 2010, Velocity estimation by image focusing analysis: SEG Technical Program Expanded Abstracts **29**, 4066–4070.

de Ridder, S., Green's function retrieval by iterated coda correlations, Eos, Trans. AGU, 90(52), Fall Meet. Suppl., Abstract S21A–1699.

de Ridder, S., B. Biondi, G. Papanicolaou, 2009, Kinematics of iterative interferometry in a passive seismic experiment: SEG Technical Program Expanded Abstracts **28**, 1622–1626.

de Ridder, S., N. Crook, S. S. Haines and S. T. Ide, 2010, Seismic Investigation of Underground Coal Fires; A Feasibility Study at the Southern Ute Nation Coal Fire Site, Symposium on the Application of Geophysics to Engineering and Environmental Problems **23**, 630–638.

de Ridder and B. Biondi, 2010, Low-frequency passive seismic interferometry for land data: SEG Technical Program Expanded Abstracts **29**, 4041–4046.

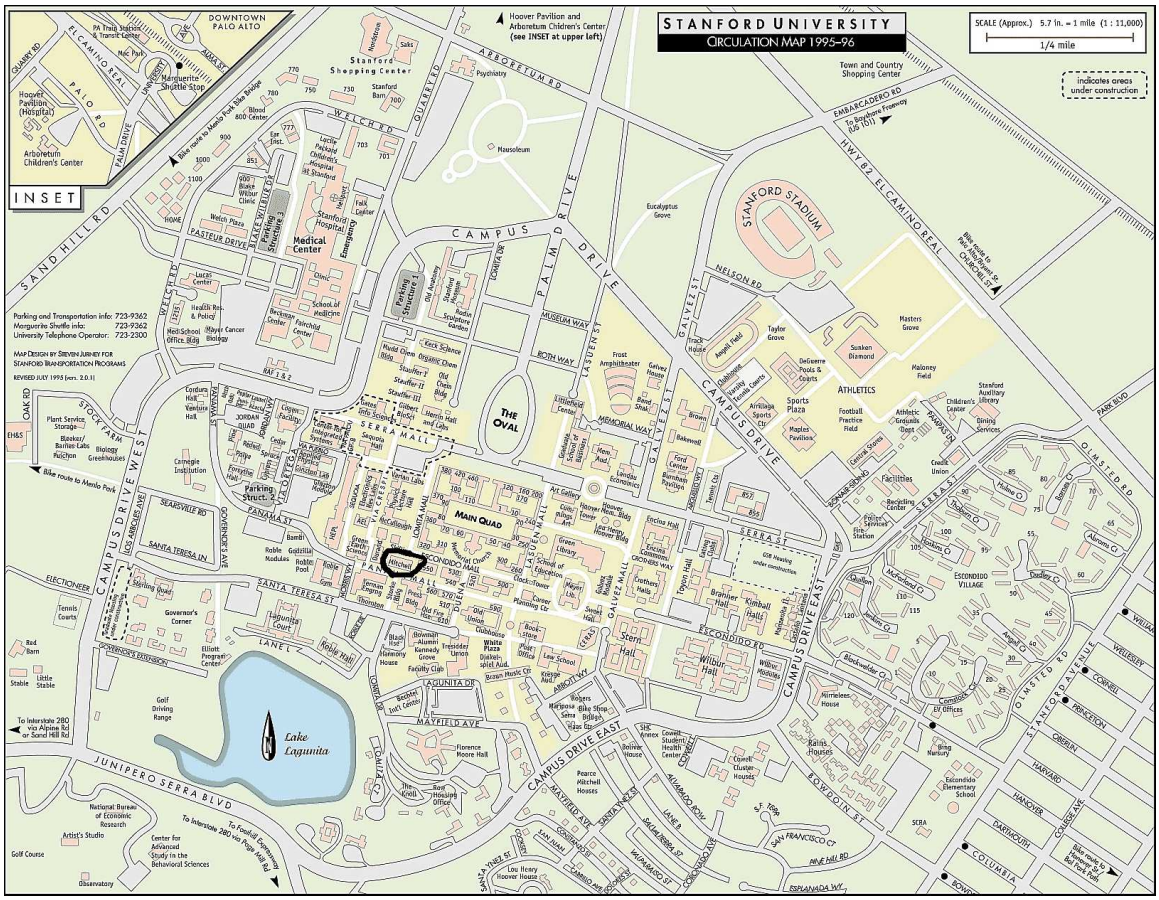
Dellinger, J., J. Yu and S. de Ridder, 2010, Virtual-source interferometry of 4C OBC data at Valhall without a low-cut recording filter: SEG low-frequency workshop.

de Ridder, S. and J. Dellinger, 2011, Imaging of Near-Surface Structures at Valhall using Ambient-Seismic-Noise Eikonal Tomography: EAGE Third Passive Seismic Workshop,

Athens, Greece, PSP13.
de Ridder, S. and J. Dellinger, 2011, Ambient seismic noise eikonal tomography for near-surface imaging at Valhall. The Leading Edge, **30**, no. 5, pp. 936–942.
Halpert, A., Clapp, R.G., and B. Biondi, 2009, Seismic image segmentation with multiple attributes: SEG Technical Program Expanded Abstracts **28**, 3700–3704.
Halpert, A., Clapp, R.G., and B. Biondo, 2010, Speeding up seismic image segmentation: SEG Technical Program Expanded Abstracts **29**, 1276–1280.
Halpert, A., Clapp, R.G., and B. Biondo, 2011, Interpreter guidance for automated seismic image segmentation: 73rd Conference & Exhibition, EAGE, Extended Abstracts (Accepted).
Leader, C., Shen, X., and Clapp, R., 2011, Memory efficient reverse time migration in 3D: Proceedings of Nvidia GPU Technology Conference 2011/2012 (Submitted).
Li, Y., Y. Zhang and J. Claerbout, 2010, Geophysical applications of a novel and robust L1 solver: SEG Technical Program Expanded Abstracts **29**, 3519–3523.
Li, Y., and Biondi, B., 2011, Migration velocity analysis for anisotropic models: SEG Technical Program Expanded Abstracts **30** (Submitted).
Li, Y., Y. Zhang and J. Claerbout, 2011, Hyperbolic estimation of sparse and blocky models: Geophysics (Submitted).
Li, Y., Nicoles, D., Osypov, K., and Bachrach, R., 2011,Anisotropic tomography with rock physics constraints: 73rd Conference & Technical Exbition, EAGE, Extended Abstracts.
Shen, X., 2010, Near-surface velocity estimation by weighted early-arrival waveform inversion: SEG Technical Program Expanded Abstracts **29** (Accepted).
Zhang, Y., Claerbout, J., Guitton, A., 2011, A new bidirectional deconvolution that overcomes the minimum phase assumption: 73rd Conference & Technical Exbition, EAGE, Extended Abstracts.
Tang, Y. and B. Biondi, 2009, Least-squares migration/inversion of blended data: SEG Technical Program Expanded Abstracts **28**, 2859–2863.
Tang, Y., 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian: Geophysics, **74**, WCA95–WCA107.
Tang, Y. and S. Lee, 2010, Efficient Gauss-Newton Hessian for full waveform inversion: 72nd Conference & Technical Exbition, EAGE, Extended Abstracts.
Tang, Y. and B. Biondi, 2010, Target-oriented wavefield tomography using demigrated Born data: SEG Technical Program Expanded Abstracts **29**, 4280–4285.
Tang, Y. and S. Lee, 2010, Preconditioning full waveform inversion with phase-encoded Hessian: SEG Technical Program Expanded Abstracts **29**, 1034–1038.
Tang, Y. and B. Biondi, 2011, Target-oriented wavefield tomography using synthesized Born data: Geophysics (Accepted).
Tang, Y. and B. Biondi, 2011, Subsalt velocity analysis by target-oriented wavefield tomography: A 3-D field-data example: SEG Technical Program Expanded Abstracts (Submitted).
Vyas, M. and Y. Tang, 2010, On the gradient of wave-equation migration velocity analysis: 72nd Conference & Technical Exbition, EAGE, Extended Abstracts.
Vyas, M. and Y. Tang, 2010, Gradients for wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts **29**, 4077–4081.
Wong, M., Biondi, B. L., and Ronen, S., 2010, Joint inversion of up- and down-going signal for ocean bottom data: SEG Technical Program Expanded Abstracts **29**, 2752–2756.
Wong, M. and S. Ronen and B. Biondi , 2010, Least-squares reverse time migration/inversion for ocean-bottom data: A case study (Submitted).

Shen, X., 2010, Near-surface velocity estimation by weighted early-arrival waveform inversion: SEG Technical Program Expanded Abstracts **29**, 1975–1979.

Shen, X. and R. Clapp, 2011, Random boundary condition for low-frequency wave propagation: SEG Technical Program Expanded Abstracts (Submitted).


STANFORD UNIVERSITY
CIRCULATION MAP 1995–96
SCALE (Approx.) 5.7 in. = 1 mile (1 : 11,000)
1/4 mile
indicates areas under construction
INSET
Hoover Pavilion and Arboretum Children's Center (see INSET at upper left)
DOWNTOWN PALO ALTO
Stanford Shopping Center
Town and Country Shopping Center
Medical Center
Stanford Hospital
Emergency
Lake Lagunita
Golf Driving Range
Golf Course
Observatory
Artist's Studio
Stable
Tennis Courts
Governor's Corner
Governor's Extension
Golf Course
Parking and Transportation info: 723-9362
Marguerite Shuttle info: 723-9362
University Telephone Operator: 723-2300
MAP DESIGN BY STEVEN JURNEY FOR STANFORD TRANSPORTATION PROGRAMS
REVISED JULY 1995 (rev. 2.0.1)
Bike route to Menlo Park
Bike route to Menlo Park via Alma St
Bike route to Palo Alto
Bike route to Bay Area Bicycle Path
To Interstate 280 via Alpine Rd or Sand Hill Rd
To Interstate 280 via Page Mill Rd
To Foothill Expressway
To Highway 101
Parking Structure 1
Parking Structure 2
Parking Struct. 3
Parking Struct. 4
Parking Structure 5
Parking Structure 6
Parking Structure 7
Parking Structure 8
Parking Structure 9
Parking Structure 10
Parking Structure 11
Parking Structure 12
Parking Structure 13
Parking Structure 14
Parking Structure 15
Parking Structure 16
Parking Structure