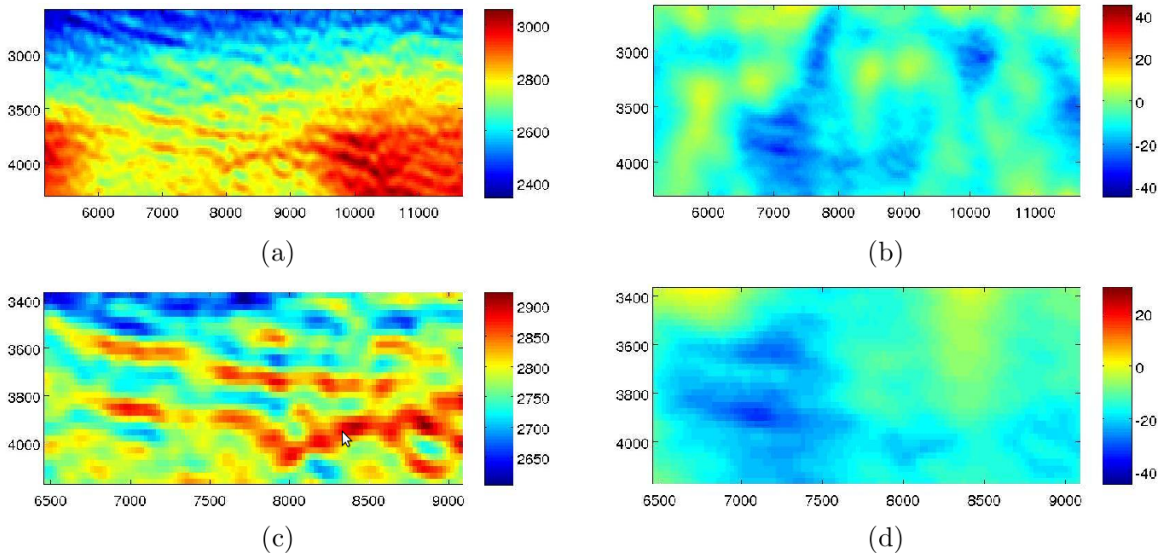
STANFORD EXPLORATION PROJECT

Jonathan Ajo-Franklin, Ali Almomin, Gustavo Alves, Ohad Barak, Guillaume Barnier, Biondo Biondi, Ettore Biondi, Alejandro Cabrales-Vargas, Jason Chang, Jon Claerbout, Robert Clapp, Steven Constable, Taylor Dahlke, Tom Daley, Shan Dou, Jack Dvorkin, Barry Freifeld, Antoine Guitton, Kerry Key, Huy Le, Chris Leader, Stewart Levin, Nate Lindsey, Yinbin Ma, Musa Maharramov, Eileen Martin, Mark Meadows, Paul Milligan, Michelle Robertson, Shuki Ronen, Yi Shen, Craig Ulrich, Anna Wagner, Yang Zhang, and Tieyuan Zhu

Report Number 158, May 2015



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, 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, MPI or CUDA based code, 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 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/sep158
[2] http://sepwww.stanford.edu/public/docs/sepdatalib/toc_html
[3] http://sepwww.stanford.edu/research/redoc/

SEP158 — TABLE OF CONTENTS

Velocity estimation

Modeling and imaging

Anisotropy and attenuation

Passive seismic

Sparse and L1 optimization

Acquisition processing and integration

Resolving the effects of production-induced overburden dilation using simultaneous TV-regularized time-lapse FWI

Musa Maharramov, Biondo Biondi and Mark Meadows

ABSTRACT

We present a field data application of the technique proposed by Maharramov and Biondi (2015) for reconstructing production-induced subsurface model changes from time-lapse seismic data using full-waveform inversion (FWI). The technique simultaneously inverts multiple survey vintages with total-variation (TV) regularization of the model differences. We apply it to the Gulf of Mexico, Genesis Field data, and successfully resolve negative velocity changes associated with overburden dilation and demonstrate that the results are stable with respect to the amount of regularization and consistent with earlier estimates of time strain in the overburden.

INTRODUCTION

Prevalent practice in time-lapse seismic processing relies on picking time displacements and changes in reflectivity amplitudes between migrated baseline and monitor images, and converting them into impedance changes and subsurface deformation (Johnston, 2013). This approach requires a significant amount of manual interpretation and quality control. One alternative approach uses the high-resolution power of full-waveform inversion (Sirgue et al., 2010a) to reconstruct production-induced changes from wide-offset seismic acquisitions (Routh et al., 2012; Zheng et al., 2011; Asnaashari et al., 2012; Raknes et al., 2013; Maharramov and Biondi, 2014a; Yang et al., 2014). However, while potentially reducing the amount of manual interpretation, time-lapse FWI is sensitive to repeatability issues (Asnaashari et al., 2012), with both coherent and incoherent noise potentially masking important production-induced changes. The joint time-lapse FWI proposed by Maharramov and Biondi (2013, 2014a) addressed repeatability issues by joint inversion of multiple vintages with model-difference regularization based on the L_2-norm and produced improved results when compared to the conventional time-lapse FWI techniques. Maharramov and Biondi (2015) extended this joint inversion approach to include edge-preserving total-variation (TV) model-difference regularization. The new method was shown to achieve a dramatic improvement over alternative techniques by significantly reducing oscillatory artifacts in the recovered model difference for synthetic data with repeatability issues. In this work, we apply this TV-regularized simultaneous inversion technique to the Gulf of Mexico, Genesis Field data and demonstrate a stable recovery of production-induced model changes.

FWI applications in time-lapse problems seek to recover induced changes in the subsurface model by using multiple seismic datasets from different acquisition vintages. For two surveys sufficiently separated in time, we call such datasets (and the associated models) "baseline" and "monitor". Time-lapse FWI can be conducted by separately inverting the baseline and monitor models ("parallel difference", Plessix et al. (2010)) or inverting them

sequentially with, e.g., the baseline supplied as a starting model for the monitor inversion ("sequential difference"). The latter may achieve a better recovery of model differences in the presence of incoherent noise (Asnaashari et al., 2012; Maharramov and Biondi, 2014a). Another alternative is to apply the "double-difference" method (Watanabe et al., 2004; Denli and Huang, 2009; Zheng et al., 2011; Asnaashari et al., 2012; Raknes et al., 2013). The latter approach may require significant data pre-processing and equalization (Asnaashari et al., 2012; Maharramov and Biondi, 2014a) across survey vintages. In all of these tech-

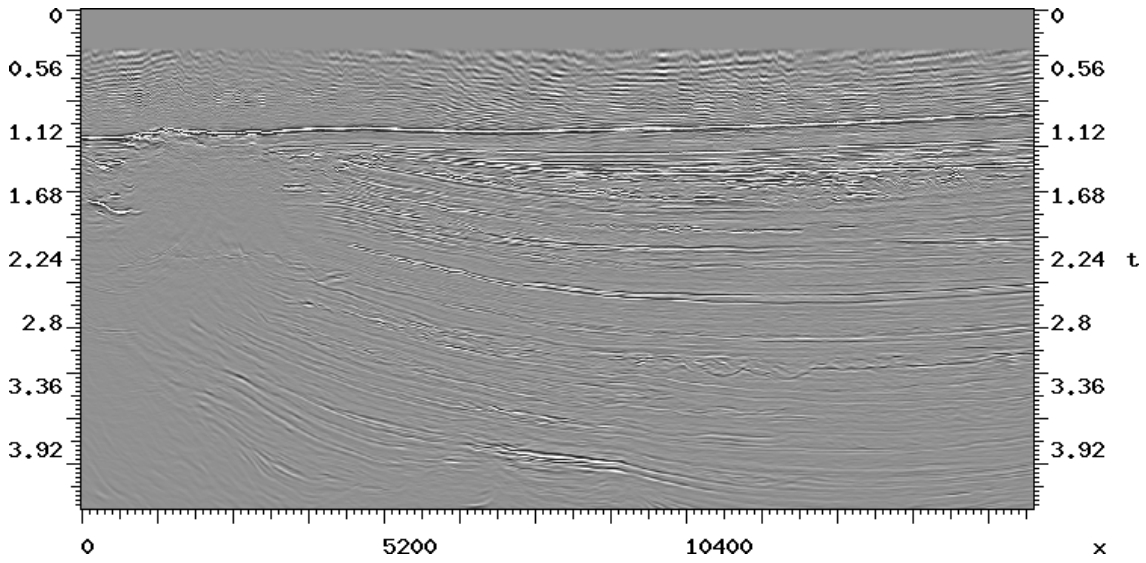


Figure 1: A north-south inline section of the baseline Genesis image produced by Chevron (vertical axis two-way travel time in seconds, horizontal axis inline meters). [**NR**] musa1/. cvxbase

niques, optimization is conducted with respect to one model at a time, albeit of different vintages at different stages of the inversion.

METHOD

We propose to invert the baseline and monitor models simultaneously by solving the following optimization problem (Maharramov and Biondi, 2015):

$$\alpha\|\mathbf{u}_b(\mathbf{m}_b)-\mathbf{d}_b\|_2^2+\beta\|\mathbf{u}_m(\mathbf{m}_m)-\mathbf{d}_m\|_2^2+ \tag{1}$$

$$\delta\|\mathbf{WR}(\mathbf{m}_m-\mathbf{m}_b)\|_1 \;\rightarrow\; \min \tag{2}$$

with respect to both the baseline and monitor models $\mathbf{m}_b$ and $\mathbf{m}_m$. Problem (1,2) describes time-lapse FWI with the L_1 regularization of the transformed model difference (2). The terms (1) correspond to separate baseline and monitor inversions with observed data $\mathbf{d}$ and modeled data $\mathbf{u}$. In (2), $\mathbf{R}$ and $\mathbf{W}$ denote regularization and weighting operators, respectively. If $\mathbf{R}$ is the gradient magnitude operator

$$\mathbf{R}f(x,y,z) \;=\; \sqrt{f_x^2+f_y^2+f_z^2}, \tag{3}$$

then (2) becomes the "Total Variation" (TV) seminorm. This case is of particular interest, because minimization of the gradient L_1 norm promotes "blockiness" of the model-difference, potentially reducing oscillatory artifacts (Rudin et al., 1992). Total-variation regularization, known in image processing as the "ROF Model", was applied earlier to full-waveform inversion as a way of resolving sharp geologic boundaries (Anagaw and Sacchi, 2012). The solution of a large-scale optimization problem based on the ROF model using conventional methods is computationally challenging, prone to the "staircasing effect" (Chambolle and Lions, 1997), and may require solution methods that involve splitting and gradient thresholding (Goldstein and Osher, 2009). However, time-lapse FWI appears to be a nearly ideal application for the ROF model, because significant production-induced subsurface model changes are spatially bounded and have magnitudes that can be roughly estimated *a priori* from geomechanical and production data (Maharramov and Biondi (2014a), supplementary material). More specifically, the weighting operator **W** may be obtained from prior geomechanical information. For example, a rough estimate of production-induced velocity changes can be obtained from time shifts (Hatchell and Bourne, 2005) and used to map subsurface regions of expected production-induced perturbation.

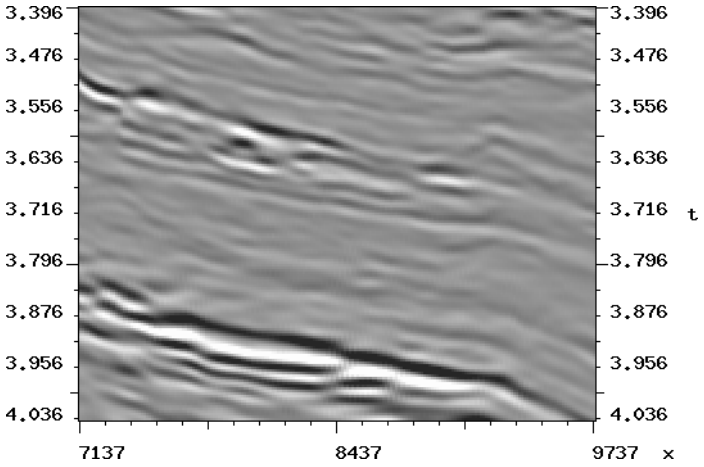


Figure 2: Monitor and baseline image-difference obtained from the 3D time-migration images provided by Chevron which corresponds to the inline section of Figure 1. Production-induced changes stand out at approximately 3.5 s (wet Illinoisan sands) and 4 s two-way travel times—stacked Neb 1, 2, and 3 reservoirs—compare with Hudson et al. (2005). [**NR**] musa1/. cvxdiff

APPLICATION TO FIELD DATA

The Genesis Field, operated by Chevron, is located 150 miles southwest of New Orleans in the Green Canyon area of the central Gulf of Mexico, in approximately 770-830m of water (Magesan et al., 2005). Oil was found in several late Pliocene through early Pleistocene deep-water reservoirs. Most of the field's oil and gas reserves are in the early Pleistocene Neb 1, Neb 2, and Neb 3 reservoirs that are the primary subject of this study. First oil production began in January 1999.

A 3D seismic survey was shot in 1990, and a time-lapse survey was shot in October 2002 with the aim of improving field management (Hudson et al., 2005; Magesan et al., 2005).

Cumulative production from the field at the time of the monitor survey was more than 57 MMBO, 89 MMCFG, and 19 MMBW (Hudson et al., 2005).

In addition to fluid substitution effects, producing reservoirs compact, increasing the depth to the top of the reservoirs and causing overburden dilation (Johnston, 2013). A time-lapse study performed by Chevron (Hudson et al., 2005) indicated significant apparent kinematic differences in the Pleistocene reservoir interval. Time shifts were observed both for the producing reservoirs and Illinoisan wet sands above Neb 1 (see Figure 3). Kinematic differences were attributed to a time shift caused by subsidence at the top of the uppermost reservoir, subsidence of the overburden, and overburden dilation (Hudson et al., 2005).

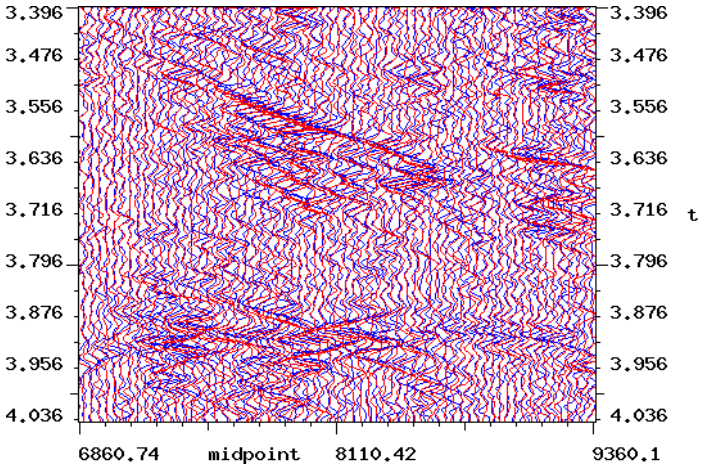


Figure 3: Production-induced changes resulted in measurable time-shifts between the surveys. Shown here are time-shifts between the baseline (blue) and monitor (red) common-offset gathers, 1074 m offset. [**CR**] musa1/. timeshifts

Processing parameters for the baseline and monitor surveys and subsequent time-lapse processing by Chevron were described by Magesan et al. (2005). The baseline survey had a maximum offset of 5 km, and the monitor survey had a maximum offset of 7.3 km. Both surveys used a bin size of 12.5 m by 37 m. For the purpose of time-lapse analysis, the acquired data had been subjected to pre-processing and imaging steps that included data equalization, spherical divergence correction, source and receiver statics, global phase rotation, time shift, amplitude scaling, global spectral matching, and cross-equalization (Magesan et al., 2005).

The data pre-processed for time-lapse analysis were used by Chevron in Kirchhoff time migration of the baseline and monitor surveys, producing 3D images. A single inline section of the baseline image is shown in Figure 1. The corresponding monitor and baseline image difference is shown in Figure 2. As noted by Hudson et al. (2005), the image difference is contributed to by time shifts at the Illinoisan sands (upper event) and Neb 1 (lower event) in Figure 2—compare with Figure 1 of Hudson et al. (2005).

The purpose of this application was to see whether or not joint regularized time-lapse FWI could resolve some of the production-induced model differences, thus providing additional insight into reservoir depletion patterns and optimal infill drilling strategies. As our first processing step, we performed separate baseline and monitor 2D full-waveform inversion of a single inline section. We extracted a single north-south inline corresponding

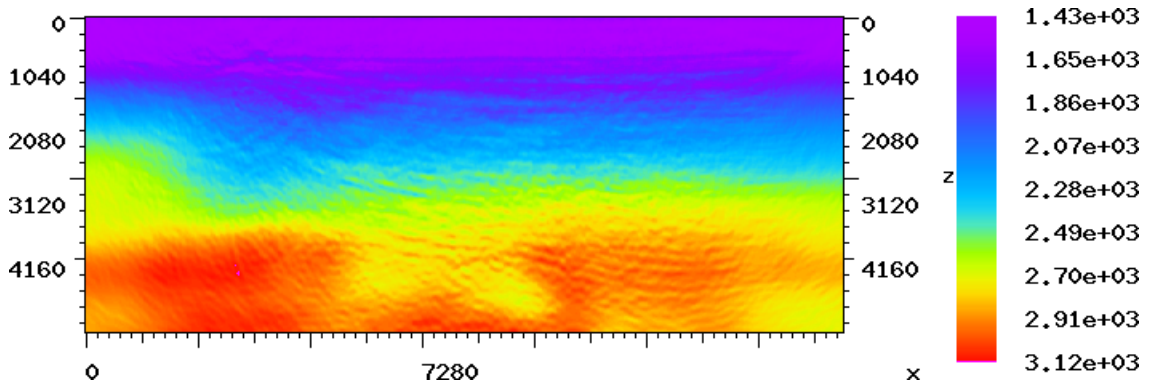


Figure 4: Inverted baseline velocity model. FWI resolved fine model features and oriented them along the dip structure of the image in Figure 1 (vertical axis depth meters). [**CR**] musa1/. base

to the image in Figure 1 from both surveys and sorted it into shot gathers with a minimum offset of 350 m and a maximum offset of 4,700 m. This provided 1,264 shots with up to 175 receivers per shot. A frequency-domain 2D FWI (Sirgue et al., 2008, 2010b) was conducted over the frequency range of 3-30.7 Hz. Frequency spacings were selected using the technique of Sirgue and Pratt (2004). The data provided to us had undergone amplitude pre-processing that included a spherical divergence correction. Furthermore, accurate handling of the amplitudes in 2D FWI of 3D field data requires a 3D-to-2D data transformation (Auer et al., 2013). Because the data exhibited significant time-shifts at the reservoir level (Hudson et al., 2005) that can be readily observed even at far offsets (see Figure 3), we decided to use a "phase-only" inversion and ignored amplitude information in the data (Fichtner, 2011).

The result of baseline inversion is shown in Figure 4. To build a starting model for the FWI, we converted Chevron's RMS time-migration velocity model to an interval velocity using the Dix equation, and smoothed the result using a triangular filter with a 41-sample window. Observe that FWI succeeded in resolving fine features, and oriented them consistently along the dip structure of the time-migrated image in Figure 1. Close-up views of the model area covering both the Illinoisan sands and the reservoirs are shown in Figures 6(a) and 6(c).

The result of parallel differencing is shown in Figure 5(a). Although significant model changes appear to be concentrated around the target area, this result is not interpretable, either qualitatively or quantitatively, because it is contaminated with oscillatory artifacts and overestimates the magnitudes of velocity perturbations. This result is consistent with our earlier assessment of conventional time-lapse FWI techniques tested on synthetic data (Maharramov and Biondi, 2014a, 2015).

Next, we solved the simultaneous, TV-regularized time-lapse full-waveform inversion problem (1,2). We set $\alpha = \beta = 1$ and carried out multiple experiments with the value of the regularization parameter δ ranging from $\delta = 100$ to $\delta = 1000$. The weighting operator $\mathbf{W}$ was set to 1 inside the larger target area shown in Figures 5(a) through 6(b), and tapered off to zero outside.

The results of inverting the model difference for $\delta = 100$, 500 and 1000 are shown in Figures 5(b), 5(c), and 5(d), respectively. Gradual increase of the regularization parameter results in the removal of most model differences with the exception of a negative velocity perturbation in the overburden, peaking at approximately 3.6 km and 3.9 km (see Figures 6(b) and 6(d)). Such perturbations are consistent with overburden dilation due to the compaction of stacked reservoirs, with more significant dilation in the wet Illinoisan sands than the surrounding shales (Rickett et al., 2007). The zone of negative velocity change appears to extend upward into the overburden in a direction roughly orthogonal to the reservoir dip—see Figure 5(d). Two negative velocity changes at approximately 10 and 11.5 km inline persist with increasing regularization, and may represent dilation effects associated with the production from deeper reservoirs—compare with Figure 3 of Rickett et al. (2007).

The estimated maximum negative velocity change of -45 m/s above the stacked reservoirs is consistent with the earlier estimates of time strain in the overburden (Rickett et al., 2007). Indeed, local time strain, physical strain and partial velocity change are related by the equation (Hatchell and Bourne, 2005)

$$\frac{d\tau}{dt} \approx \frac{\Delta t}{t} = \frac{\Delta z}{z} - \frac{\Delta v}{v}, \tag{4}$$

where τ, t, z, and v denote the observed time shift, travel time, depth, and velocity. Assuming, following Hatchell and Bourne (2005), that

$$\frac{\Delta v}{v} = -R\frac{\Delta z}{z}, \tag{5}$$

where the factor R is estimated to be 6 ± 2 for the Genesis overburden (Hodgson et al., 2007), we obtain

$$\frac{\Delta v}{v} = -\frac{R}{R+1}\frac{\Delta t}{t} \approx -\frac{\Delta t}{t} \approx -\frac{d\tau}{dt}. \tag{6}$$

Maximum time strains in the Genesis overburden are estimated to be around +2% (Rickett et al., 2007), yielding the maximum negative velocity change of

$$\Delta v \approx -.02 \times 2,800 \text{ m/s} = -56 \text{ m/s}, \tag{7}$$

where the estimated P-wave velocity of 2,800 m/s at a 3.6 km depth was taken from the output of our FWI.

CONCLUSIONS

Simultaneous time-lapse FWI with total-variation difference regularization can achieve robust recovery of production-induced changes, preserving the blocky nature of monitor-baseline differences. Velocity changes caused by overburden dilation are within the resolution of our method. However, velocity changes within thin reservoirs (e.g., caused by compaction) can no longer be characterized as "blocky" on a seismic scale, and recovering such changes may require a multi-norm model decomposition approach *a la* Maharramov and Biondi (2014b), and will be the subject of further research.

ACKNOWLEDGMENTS

The authors would like to thank the affiliate members of Stanford Exploration Project (SEP) for their support; Joseph Stefani, Stewart Levin, Shuki Ronen, and Dimitri Bevc for a number of useful discussions; Chevron, ExxonMobil, and BHP Billiton for providing the field data; and the Stanford Center for Computational Earth and Environmental Science for providing computing resources.

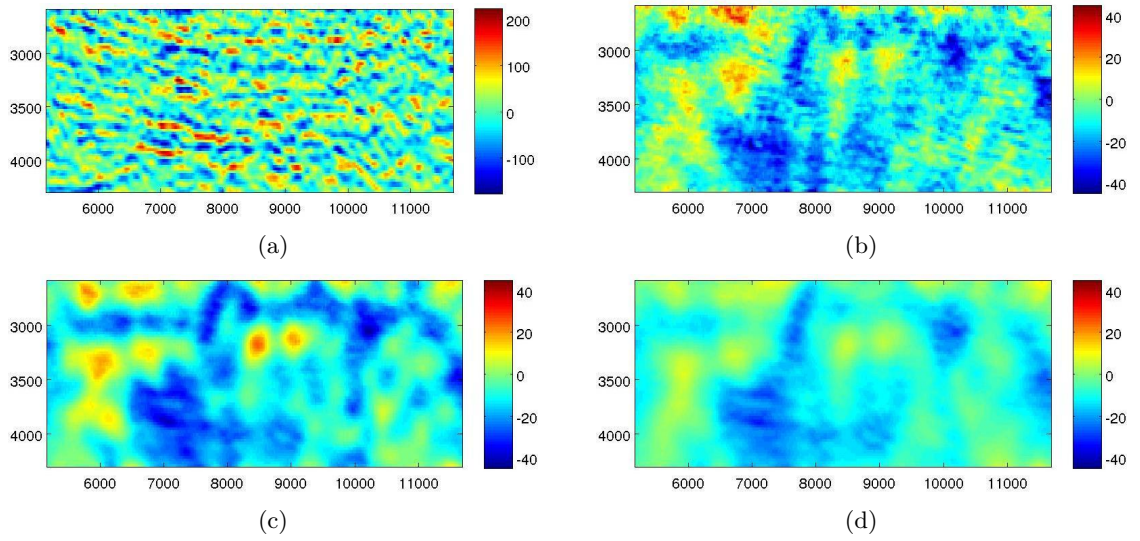


Figure 5: (a) Parallel difference and joint inversion results for (b) $\delta = 100$, (c) $\delta = 500$, and (d) $\delta = 1000$ in the target area. The parallel difference result is not interpretable because of the presence of artifacts. Increasing the regularization parameter δ results in gradual removal of most model differences except the negative velocity change in the overburden, peaking around the Illinoisan sands and near the top of the stacked reservoirs—see Figures 6(a) through 6(d). [**CR**] musa1/. pardiff,diff100,diff500,diff1000

REFERENCES

Anagaw, A. Y. and M. D. Sacchi, 2012, Edge-preserving seismic imaging using the total variation method: Journal of Geophysics and Engineering, **9**, 138.

Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, 2012, Time-lapse imaging using regularized FWI: A robustness study: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi:10.1190/segam2012-0699.1, 1–5.

Auer, L., A. M. Nuber, S. A. Greenhalgh, H. Maurer, and S. Marelli, 2013, A critical appraisal of asymptotic 3D-to-2D data transformation in full-waveform seismic crosshole tomography: Geophysics, **78**, no. 6, R235–R247.

Chambolle, A. and P. L. Lions, 1997, Image recovery via total variational minimization and related problems: Numerische Mathematik, **76**, 167–188.

Denli, H. and L. Huang, 2009, Double-difference elastic waveform tomography in the time domain: 79th Annual International Meeting, SEG, Expanded Abstracts, 2302–2306.

Fichtner, A., 2011, Full seismic modeling and inversion: Springer.

Goldstein, T. and S. Osher, 2009, The split Bregman method for L1-regularized problems: SIAM Journal on Imaging Sciences, **2**, 323–343.

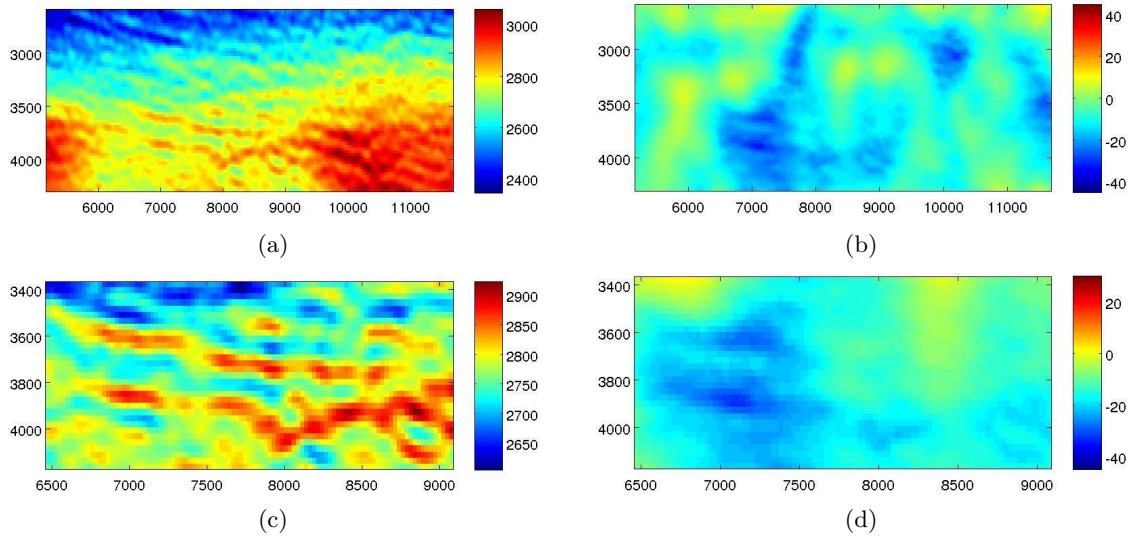


Figure 6: (a) Baseline target area and (b) estimated model difference for $\delta = 1000$. Close-up of (c) baseline target area and (b) estimated model difference for $\delta = 1000$. [**CR**] musa1/. baselargetarget,sdiff1000,basetarget,difftarget

Hatchell, P. and S. Bourne, 2005, Measuring reservoir compaction using time-lapse timeshifts: 75th Annual International Meeting, SEG, Expanded Abstracts, 2500–2503.

Hodgson, N., C. Macbeth, L. Duranti, J. Rickett, and K. Nihei, 2007, Inverting for reservoir pressure change using time-lapse time strain: Application to Genesis Field, Gulf of Mexico: The Leading Edge, **26**, 649–652.

Hudson, T., B. Regel, J. Bretches, J. Rickett, B. Cerney, and P. Inderwiesen, 2005, Genesis Field, Gulf of Mexico, 4D project status and preliminary lookback: 75th Annual International Meeting, SEG, Expanded Abstracts, 2436–2439.

Johnston, D., 2013, Practical applications of time-lapse seismic data: Society of Exploration Geophysicists.

Magesan, M., S. Depagne, K. Nixon, B. Regel, J. Opich, G. Rogers, and T. Hudson, 2005, Seismic processing for time-lapse study: Genesis Field, Gulf of Mexico: The Leading Edge, **24**, 364–373.

Maharramov, M. and B. Biondi, 2013, Simultaneous time-lapse full waveform inversion: SEP Report, **150**, 63–70.

———, 2014a, Joint full-waveform inversion of time-lapse seismic data sets: 84th Annual Meeting, SEG, Expanded Abstracts, 954–959.

———, 2014b, Multi-model full-waveform inversion: SEP Report, **155**, 187–192.

———, 2015, Robust simultaneous time-lapse full-waveform inversion with total-variation regularization of model difference: 77th EAGE Conference and Exhibition, Extended Abstracts (accepted).

Plessix, R.-E., S. Michelet, H. Rynja, H. Kuehl, C. Perkins, J. W. de Maag, and P. Hatchell, 2010, Some 3D applications of full waveform inversion: 72nd EAGE Conference and Exhibition, Workshop WS6 "3D Full Waveform Inversion—A Game Changing Technique?".

Raknes, E., W. Weibull, and B. Arntsen, 2013, Time-lapse full waveform inversion: Synthetic and real data examples: 83rd Annual International Meeting, SEG, Expanded Ab-

stracts, 944–948.

Rickett, J., L. Duranti, T. Hudson, B. Regel, and N. Hodgson, 2007, 4D time strain and the seismic signature of geomechanical compaction at Genesis: The Leading Edge, **26**, 644–647.

Routh, P., G. Palacharla, I. Chikichev, and S. Lazaratos, 2012, Full wavefield inversion of time-lapse data for improved imaging and reservoir characterization: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi:10.1190/segam2012-1043.1, 1–6.

Rudin, L. I., S. Osher, and E. Fatemi, 1992, Nonlinear total variation based noise removal algorithms: Physica D: Nonlinear Phenomena, **60**, 259–268.

Sirgue, L., O. Barkved, J. Dellinger, J. Etgen, U. Albertin, and J. Kommendal, 2010a, Full waveform inversion: The next leap forward in imaging at Valhall: First Break, **28**, no. 4, 65–70.

Sirgue, L., J. T. Etgen, and U. Albertin, 2008, 3D frequency domain waveform inversion using time domain finite difference methods: 70th EAGE Conference and Exhibition, Extended Abstract, F022.

Sirgue, L., J. T. Etgen, U. Albertin, and S. Brandsberg-Dahl, 2010b, System and method for 3D frequency domain waveform inversion based on 3D time-domain forward modeling. (US Patent 7,725,266).

Sirgue, L. and R. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.

Watanabe, T., S. Shimizu, E. Asakawa, and T. Matsuoka, 2004, Differential waveform tomography for time-lapse crosswell seismic data with application to gas hydrate production monitoring: 74th Annual International Meeting, SEG, Expanded Abstracts, 2323–2326.

Yang, D., A. E. Malcolm, and M. C. Fehler, 2014, Time-lapse full waveform inversion and uncertainty analysis with different survey geometries: 76th EAGE Conference and Exhibition, Extended Abstract, We ELI1 10.

Zheng, Y., P. Barton, and S. Singh, 2011, Strategies for elastic full waveform inversion of timelapse ocean bottom cable (OBC) seismic data: 81st Annual International Meeting, SEG, Expanded Abstracts, 4195–4200.

Residual-moveout-based wave-equation migration velocity analysis — field data example

Yang Zhang and Biondo Biondi

ABSTRACT

We apply our residual-moveout-based (RMO-based) wave-equation migration velocity analysis (WEMVA) method to an industry scale 3-D marine streamers wide-azimuth data set — E-Octopus III in the Gulf of Mexico. This 3-D field data set poses many challenges for our implementation, including irregular geometry, abnormal traces, complex 3-D salt geometry, and more importantly, huge data volume and large domain dimensions. To overcome these hurdles, we apply careful data regularization and preprocessing, and employ a target-oriented inversion scheme that focuses on the update of sediment velocities in specific regions of interest. Such target-oriented scheme significantly reduces the computational cost, allowing us to make our WEMVA method affordable on our academic cluster. Our experiment result on a subsalt sediments target region shows that the angles of illumination on the subsalt sediments are very limited (less than 25 degrees) because of the complex salt overburden and the depth of the target. Although the lack of angular illumination in this region severely reduces the capability of any reflection tomography method that tries to resolve a better velocity model, our RMO-based WEMVA method is still able to detect the curvatures of the angle gathers and produce good velocity model updates that further flatten the angle gathers and improve the quality of the structural image in the target region.

INTRODUCTION

Wave-equation migration velocity analysis (WEMVA) is a reflection tomography method which uses a wave-equation rather than a ray-based model to retrieve the velocity model from seismic data. The velocity information comes from the redundancy in the seismic data, since each reflector point in the subsurface is illuminated by wave energy from multiple directions. WEMVA exploits such redundancy by forming common-image gathers (CIG) and then enforcing coherence among the CIG to improve the velocity model.

We recently proposed a WEMVA method named RMO-based WEMVA (Zhang and Biondi, 2013) that describes the unflatness in the angle-domain common image gathers (ADCIG) using residual-moveout (RMO) parameters. This method tries to improve the velocity model using the moveout information. With the successful application of the RMO-based WEMVA on 2-D problems such as the 2006 BP synthetic model (Zhang et al., 2013), we have shown that this method does not suffer from cycle-skipping or require the picking of moveout parameters, and it can robustly improve the flatness of the angle gathers.

However, although the theory of this method can be easily extended to 3-D, a realistic 3-D data application remains very challenging. In this study, we examine a marine streamers

3-D wide-azimuth (WAZ) seismic recordings acquired from offshore Gulf of Mexico (GoM) by Schlumberger Multiclient. The remainder of this paper is organized as follows:

- In the first part, we give a brief recap of the 3-D theory of the RMO-based WEMVA. In addition to the basic formulation that assigns a RMO parameter to every image point, we present a refined formulation where the RMO parameters are assigned to individual identified events. In order to support the refined formulation in practical application, we propose the use of several image processing techniques to identify and extract the events in the image.
- In the second part, we focus on the application of our RMO-based WEMVA method on the aforementioned field data set. In the previous report (Zhang and Biondi, 2014), we presented the overview of this WAZ data set, our data pre-processing procedures and the initial 3-D migration images. Here we perform velocity estimation using the RMO-based WEMVA on the data set in a target-oriented manner. The results show improvements in the ADCIG flatness and in the migrated image, which has better defined continuity and coherency in terms of sedimentary structure.

THEORY OF 3-D RMO-BASED WEMVA

In this section, we summarize the theory of the proposed RMO-based WEMVA approach in 3-D. The detailed derivation (refer to Zhang (2015)) is beyond the scope of this paper. Here, we only show key definitions and equations in the method formulation. The model space is parameterized as slowness unless otherwise specified.

In the 3-D case, we denote the prestack image as $I(z, \gamma, \phi, x, y)$, in which (z, x, y) are the depth and horizontal axes, respectively, γ is the reflection opening angle, and ϕ is the subsurface azimuth. Assuming there are m samples along ϕ axis, then $\phi = \{\phi_i : i = 1, 2, ..., m\}$. Let us start from the classical stack-power maximization objective function:

$$J(s) = \frac{1}{2} \sum_{x,y} \sum_{z} \left[\sum_{i} \sum_{\gamma} I(z, \gamma, \phi_i, x, y; s) \right]^2, \tag{1}$$

where s is the model slowness, and $I(z, \gamma, x; s)$ is the prestack image in the reflection-angle domain obtained by migration using the slowness s.

RMO-based WEMVA objective function and its gradient

The objective function defined in equation (1) is known to be prone to the cycle-skipping problem. Recalling that the conventional ray-based reflection tomography does not show such disadvantages, we can approximate objective function (1) at some initial slowness s_0, with an alternative one that focuses on kinematic changes of the ADCIG. Given $I(s_0)$ and $I(s)$ as the prestack images with initial and updated slowness, respectively, we define a residual moveout parameter ρ and its moveout function $\rho \tan^2 \gamma$ that describes (as accurately as possible) the kinematic difference between $I(s)$ and $I(s_0)$. Hence ρ is a function of both s and s_0. Because s_0 will remain fixed during the calculation of the model gradient, we can denote the dependence of ρ simply with $\rho(s)$. We assign independent ρ to the ADCIG at

each image point (z, x, y) and each azimuth ϕ_i because there can be many events in the gathers, and each event can have a different moveout.

Next, we replace objective function (1) with one based on normalized moveout semblance:

$$J_{S_m}(\rho(s)) = \sum_{x,y}\sum_{z}\sum_{i} S_m(\rho_i(z,x,y),\phi_i,z,x,y;s_0), \tag{2}$$

in which

$$S_m(\rho_i,\phi_i,z,x,y;s_0) = \frac{\sum_{z_w}\left(\sum_{\gamma} I(z+z_w+\rho_i\tan^2\gamma,\gamma,\phi_i,x,y;s_0)\right)^2}{\sum_{z_w}\sum_{\gamma} I^2(z+z_w+\rho_i\tan^2\gamma,\gamma,\phi_i,x,y,;s_0)}. \tag{3}$$

Maximizing the semblance-based objective function achieves the same goal as maximizing the angle-stack because the optimal model s leads to optimal moveout parameters ρ such that the angle gathers in $I(s_0)$ become flat after being applied the moveout, and flat angle gathers will stack most coherently to yield maximum stacked amplitude. Note that in the above equation, I use a local window of length L for z_w (through this derivation, I assume the summation interval for variable z_w is always $[-L/2, L/2]$). When choosing a value for L, the rule of thumb is that L should not be smaller than the span of the event's wavelet, and it should not be too large such that it contains more than one event in each window.

We use gradient-based methods to solve this optimization problem. The model update given by the gradient of objective function (2) is

$$\frac{\partial J_{S_m}}{\partial s} = \sum_{x,y}\sum_{z}\sum_{i}\frac{\partial \rho_i}{\partial s}\frac{\partial J_{S_m}}{\partial \rho_i}. \tag{4}$$

In the above equation, $\partial\rho/\partial s$ describes the relation between slowness perturbation and moveout parameter perturbation, and $\partial S_m/\partial\rho$ indicates the search for better moveout parameters to flatten the gathers. Without going into the details described in Zhang (2015), we can summarize the calculation of these two terms as follows:

- $\partial J_{S_m}/\partial\rho_i$ is calculated by taking first order derivative along ρ_i on the semblance panel $S_m(\rho_i,\phi_i,z,x,y;s_0)$.
- As for $\partial\rho_i/\partial s$, we can derive a relationship from the moveout perturbation $\Delta\rho$ to the image perturbation ΔI, which is then back-projected into the model space as Δs using the image-space wave-equation tomographic operator.

The global objective function for RMO-based WEMVA

One problem in our RMO-based WEMVA formulation is that it is missing a global objective function. The reason is that the semblance-based objective function (2) is formulated around s_0, the slowness model at the current iteration. Consequently, s_0 will change as the WEMVA method iterates, which makes it difficult to use this objective function as the global objective function for model evaluation. In order to overcome this problem, we propose a global objective function directly based on the semblance (3):

$$J^G_{S_m}(s) = \sum_{x,y}\sum_{z}\sum_{i}\sum_{\rho_i}\rho_i^2 S_m(\rho_i,\phi_i,z,x,y;s). \tag{5}$$

Note that the moveout parameter ρ is simply sweeping through the possible moveout value range and is not associated with model s; in addition, the corresponding inversion aims to minimize instead of maximize equation (5).

The reason we can replace objective function (2) with (5) for easy model evaluation is that the two objective functions aim for the same goal. Equation (2) aims to find the best model s such that the corresponding $\rho(s)$ would flatten the angle gathers $I(s_0)$; this is equivalent to finding the best model s under which the gathers $I(s)$ are all flat. Meanwhile, equation (5) is enforcing the same goal because it penalizes non-zero moveout in $I(s)$, which means that the optimal model it looks for is such that $I(s)$ would have zero moveout everywhere, i.e., be flat.

In summary, the RMO-based WEMVA formulation can be interpreted in the following way: first it defines the global objective function (5) over the moveout semblance of the ADCIG, aiming to flatten all gathers (driving moveout to zero); and then, at every iteration, it approximates the global objective function with equation (2), which yields a superior gradient.

Refined formulation for RMO-based WEMVA using event detection

During the actual implementation of the RMO-base WEMVA, we have noticed a major aspect for improvement. Recalling the gradient calculation in equation (2) and (4), we made an assumption that every image point (in domain (z, x, y)) has an independent moveout parameter ρ. We also assumed the event windows (spanning over z axis) of all reflector events are roughly of equal length and therefore can be summarized with an average event window size L. The main rationale behind making such assumptions is for the simplicity of formulation and coding implementation. However, assigning ρ for every image point is not only unnecessary but also inconsistent with our method's physical interpretation. Indeed, we should assign moveout parameter ρ to individual events instead of individual image point locations; moreover, we should assign a different window size for each event because there can be large variations of the window sizes among different events.

Switching the assumption from moveout per image point to moveout per event is expected to improve the accuracy of the computed model gradient and also marginally reduce the proposed method's computation cost since it cuts the computation of image perturbation. However, it also brings additional complexity for implementation because now we need to identify individual events in the image gathers.

Automatic event detection

I propose a simple workflow for automatic event detection as follows:

- Our program first detects the center locations of individual events in the zero-offset image $I(z, x, y; s_0)$ as "*anchor points*", instead of blindly assuming that every image sample is its own event.
- For each anchor point detected, the program detects the event width by measuring the event signature (waveform) compactness around the anchor point location.

To avoid clutter, we leave the details of the event detection implementation in the appendix .

Formulation change after introducing event detection

Once we figure out how to do automatic event detection, the modification on our WEMVA formulation becomes simple. We only need to slightly modify the summation indices in the gradient calculation above. Instead of summing over every image point, our method now sums over each identified event anchor point. For example, the semblance-based objective function (2) now becomes

$$J_{S_m}\left(\rho(s)\right) = \sum_{k=1}^{|\mathbf{S_{anchor}}|} \sum_i S_m(\rho_i(z_k, x_k, y_k), \phi_i, z, x, y; s_0), \tag{6}$$

in which $\mathbf{S_{anchor}} = \{(z_k, x_k, y_k)|k = 1, 2, ..., \}$ is the set of event anchor points, and $|\mathbf{S_{anchor}}|$ denotes the size of this set. Similarly, the gradient formula (4) becomes

$$\frac{\partial J_{S_m}}{\partial s} = \sum_{k=1}^{|\mathbf{S_{anchor}}|} \sum_i \frac{\partial \rho_i(z_k, x_k, y_k)}{\partial s} \frac{\partial J_{S_m}}{\partial \rho_i(z_k, x_k, y_k)}. \tag{7}$$

In addition, I change the summation bounds of the local window variable z_w in the semblance formula (3). Originally the bounds are fixed at $[-L/2, +L/2]$, but now the event window sizes are automatically detected on individual bases. These formulation changes are very easy to implement in terms of programming.

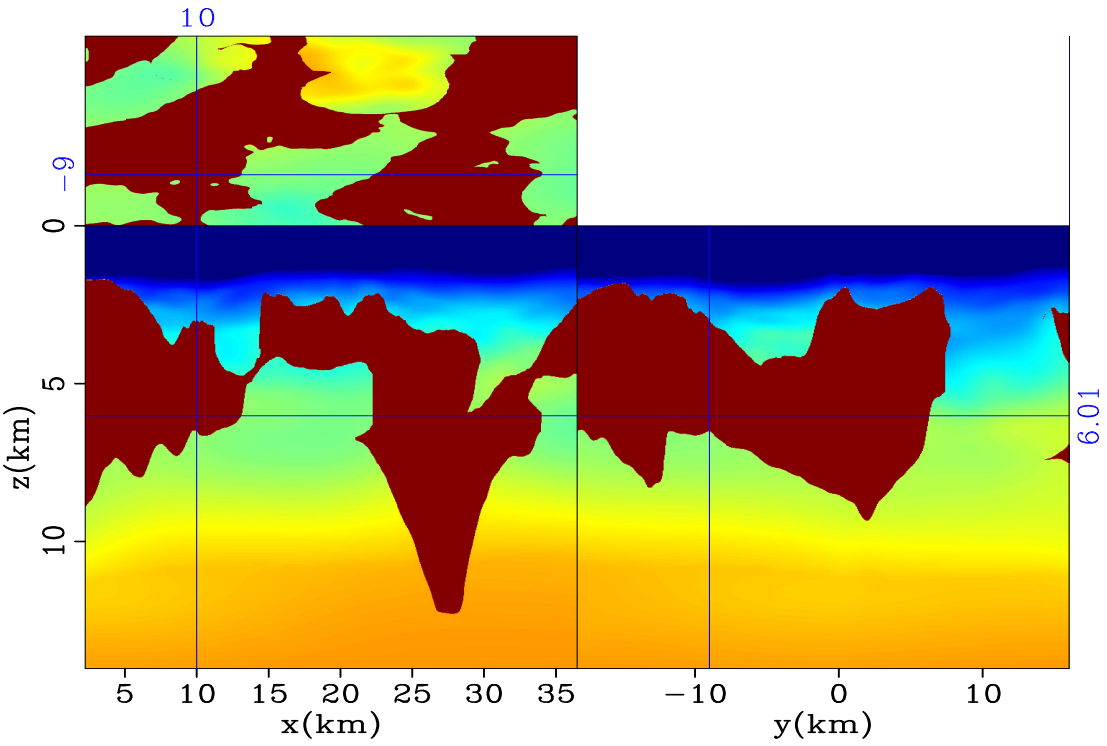
RMO-BASED WEMVA ON A 3-D WAZ FIELD DATA SET IN THE GULF OF MEXICO

In this section, we discuss how we apply the RMO-based WEMVA method for subsalt area velocity estimation on the E-Octopus WAZ field data set.

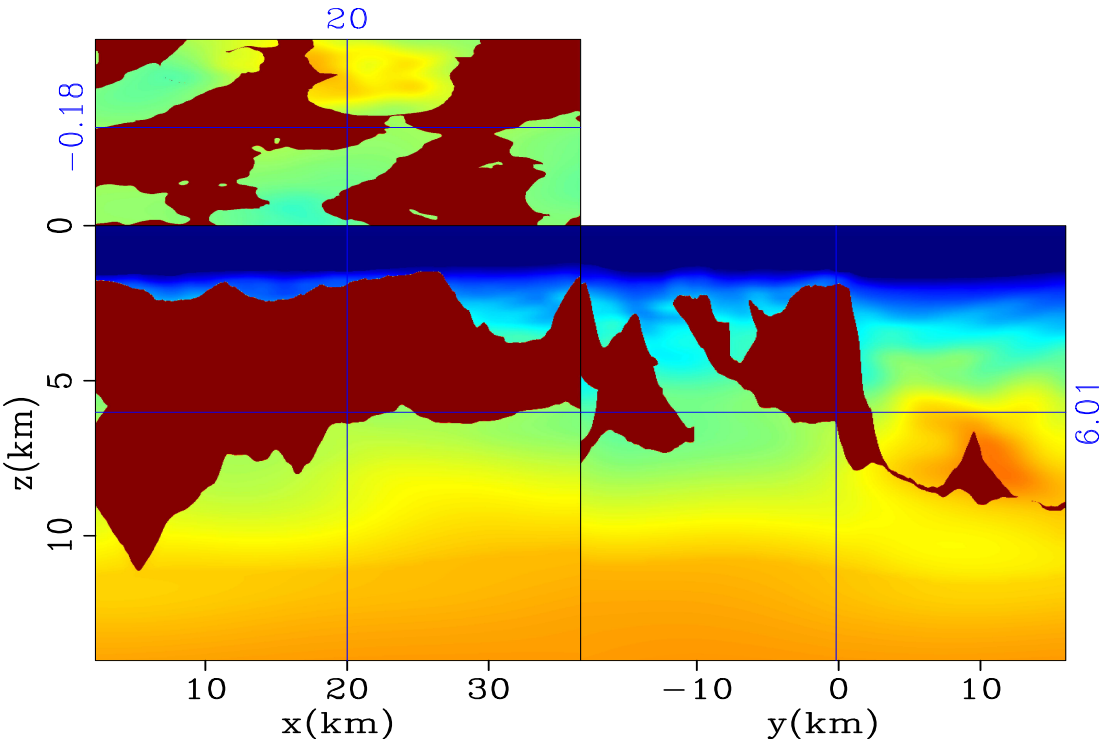
E-Octopus field data set overview

The data set we examine in this report is a WAZ streamer survey acquired offshore Gulf of Mexico (GOM) by Schlumberger. Specifically, it is part of the "E-Octopus phase III" survey in the Green Canyon area. The corresponding survey area is about 35 km by 30 km (inline by cross-line). The offset range of a shot gather is [-4.2 km, +4.2 km] cross-line and $[-7.0\,\text{km}, +7.0\,\text{km}]$ in-line. The recorded data contains $\sim 10,000$ shots, amounting to ~ 11 TB of raw disk space. The recording time per trace is 14 seconds. Please refer to our previous report (Zhang and Biondi, 2014) for more details on the WAZ survey's acquisition setup and our data preprocessing workflow.

Initial migration images on E-Octopus field data set



(a)



(b)

Figure 1: Three-panel display of the velocity model used for migration. Subplot (a) is sliced at x=10 km,y=−9 km, subplot (b) is sliced at x=20 km,y=−0.99 km. Notice the strong model variations along both X and Y directions. The color map ranges from 1450 m/s (deep blue) to 4480 m/s (deep red). [**ER**] yang1/. fig8,fig9

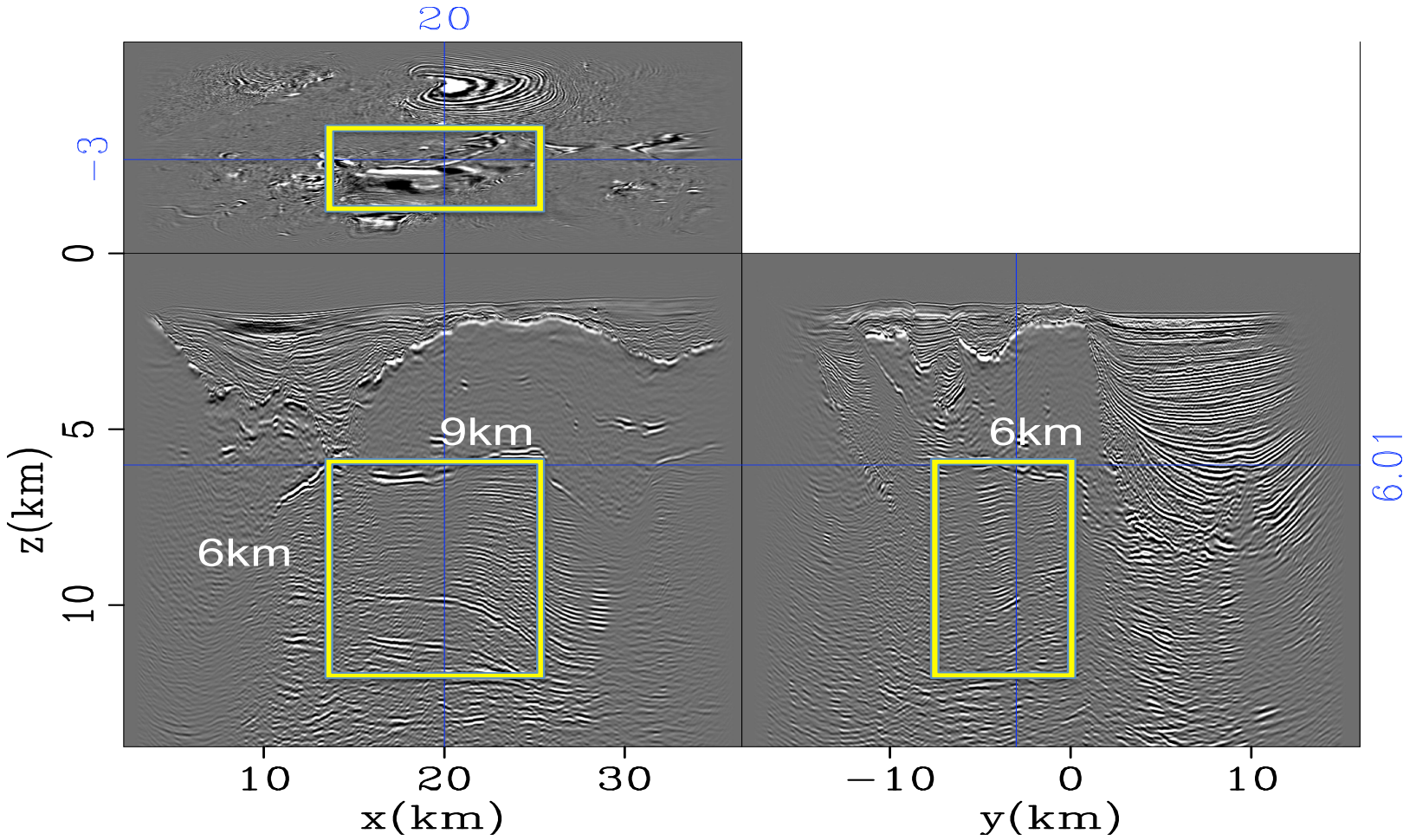


Figure 2: A three-panel view of the initial full 3-D image, in which the region of interest (ROI) for our target-oriented inversion is marked. The image has been applied a z-power (similar to t-power) gain in order to boost up the amplitudes of deeper reflectors. The dimension of ROI is 9 km by 6 km by 6 km. [**CR**] yang1/. fig12

Figure 1 shows the best velocity model we received from Schlumberger. The velocity plots show the sections of the same model at different slicing coordinates. As can be seen from figure 1, there are strong lateral variations along both X and Y directions. Figure 2 shows one section of the full 3-D migrated image. As we can see from the figure, the image quality of the sediments above the salt is very good, with continuous and coherent reflector geometries, indicating accurate velocity. However, the subsalt areas are not as well imaged. There are many discontinuities in the reflectors, as well as conflicting dips. We focus our efforts in these regions because of these indications of a less accurate velocity model.

Focusing on subsalt sediments with a target-oriented approach

Applying WEMVA method to a data set at such a large scale will occur prohibitively high computational cost. Even with all the precautions we make in the data preparation stage, a full migration on the entire domain (such as figure 2) costs ∼5000 node×hours on our academic cluster of 120 Intel Xeon nodes (two E5520 CPUs, each CPU is quad-core, 2.26 GHz). It amounts to $\sim$ 50 consecutive hours at 100% cluster usage in the optimal case. In practice, it takes us at least 5 days to complete a job at this scale. The turn-around time would be too long for practical applications, given that the wave-equation tomographic operator is even more expensive than the imaging operator, and that we have to perform tens of iterations in a typical WEMVA inversion. Therefore, to apply expensive wave-equation tomography methods on the target subsalt region, it is essential to adopt a target-oriented approach which greatly shrinks the problem domain, thus significantly reducing the computational cost.

Target-oriented tomography using generalized Born-modeling data

To make the computational cost affordable for our field data application, we use the target-oriented tomography scheme proposed by Tang (2011). This scheme synthesizes a new data set concentrating on the region of interest (ROI), which is able to preserve all kinematic information that is related to the ROI in the original data set. Therefore, in addition to shrinking the problem's physical domain, this scheme reduces the size of the data set used for WEMVA inversion, thus achieving high computational efficiency.

Tang (2011)'s method first computes the initial subsurface offset-domain common-image gathers (ODCIG) in the target region using a starting velocity model. The initial ODCIG are further normalized using the diagonal values of the imaging Hessian, which can be efficiently computed using the phase-encoding method (Tang, 2011), to optimally compensate for the uneven subsurface illumination. Tang (2011) demonstrates that the velocity information about the seismic data is now transformed into the ODCIG. Therefore this method uses the generalized Born modeling method, (which includes the ODCIG instead of merely the zero subsurface offset image in the modeling process), to simulate a new data set right on top of the target region. The user can design arbitrary acquisition geometry for the synthesized new data set. Once the new data set is generated, the wave-equation tomography is carried out exclusively on the new data set.

The target area we choose is a 9 km by 6 km by 6 km subsalt sediments region near the center of the full imaging domain, as shown in figure 2. Figure 3 shows a close-up view

of the velocity and stacked image of the same target region. We consider this area a good target for WEMVA-based velocity improvement because there are a lot of discontinuities among the imaged sediment layer interfaces, which indicates inaccurate medium velocity. Additionally, the salt overburden above this region is relatively well imaged, therefore we have more confidence in the correctness of the overhanging salt model. This leads us to infer that the velocity errors mainly exist within the subsalt sediments, and our WEMVA method is good at resolving these types of velocity errors.

The synthesized new data set

The first step in our target-oriented tomography workflow is to synthesize a data set that sits right on top of the target region while preserving all velocity information around the target region from the original data set. We choose plane-wave acquisition geometry for the new data set because it results in fewer total number of shots that need to be simulated, and the cross-talk effect usually caused by compound sources are kept minimum in plane-wave data. To determine the acquisition parameters of the plane-wave survey, we refer to the fact that the maximum subsurface illumination angle depicted in the ADCIG shown in Zhang and Biondi (2014) is no more than 25°, and the average velocity at the top of the target region is $\sim 3000\,\mathrm{m/s}$. The new acquisition samples 17 inlines and 17 crosslines (289 in total) plane-wave directions, with the ray parameters ranging at $\pm 160\,\mathrm{us}$, this corresponds to at least $\pm 30^\circ$ subsurface angle illumination with $\sim 3.75^\circ$ sampling along angle axis. Because of the reduction of the physical propagation domain, we reduce the recording time from 14.0 secs originally to 6.1 seconds. This translates to coarser sampling in the data frequency domain. Therefore we sample 99 frequencies between 4 Hz and 20 Hz for the new data set, compared to 208 frequencies for the original data set.

We take extra cautions dealing with the salt overburden when synthesizing the new data. First we build a binary mask from the starting velocity model to mark the area occupied by salt in the target region. Then we use this salt mask to erase out the subsurface ODCIG that will be used for the generalized Born modeling because the salt reflection does not provide any velocity information for the sediments below. In addition, as we can see from the velocity model in figure 3(a), over half of the source and receivers in the synthesized survey would be place in salt. The strong lateral variation in the velocity model at the recording depth would cause extra distortion to the plane-wave source wavefields, and it also forces us to use more reference velocities in our one-way propagator, which increases the computational cost. To address this potential issue, we create a *sediments-flooded* velocity model from the original model in which we replace the salt velocity with the velocity values of its surrounding sediments. We use the salt-excluded subsurface ODCIG combined with the sediments-flooded velocity model to generate our new synthetic data. Figure 4 shows the velocity model and zero subsurface offset section of the ODCIG that we use for the new data synthesis, respectively. We can manipulate the velocity model in this way because the actual salt area is considered not suitable for tomography updates and will remain fixed throughout our velocity estimation workflow. Therefore an ideal way to place the recording locations of the new survey is to put sources and receivers right at the salt/sediment boundary because preserving the sources/receivers wavefields at this boundary will preserve all velocity information about the underlying sediments, but this would cause extra complexity for our one-way propagator implementation. Using the

sediments-flooded velocity model avoids such extra complexity because now we can place sources/receivers at the same depth; more importantly, it will preserve the source/receiver wavefields at the salt/sediment boundary since we will not change the part of the velocity model that is originally occupied by salt throughout the tomography workflow.

The mapping from the original data set to this target-oriented plane-wave data set yields huge saving in computation. The size of the data set reduces by 90%, combined with the shrinkage of the modeling physical domain, the required computation time for each migration reduces from 5000 node hours to 150 node hours, a more than 30 times improvement, making overnight calculation per iteration feasible.

To verify that the newly synthesized data set contains the same velocity information as the original data, We recompute the subsurface ODCIG (both in h_x and h_y) using the new data set. In terms of the range of the offset dimensions, we use 21 points in h_x with 50 m spacing and 19 points in h_y with 60 m spacing, in order to capture most of the unfocused energy. The offset dimension is the same as the initial subsurface ODCIG I used to generate the new data set. Figure 5 and figure 6 show the comparison between the gathers migrated using the original data set and the synthesized new data set, respectively. As we can see, except a few disparities close to the boundary, their kinematic behaviors agree with each other very well. Particularly the defocuses and the shapes of the ODCIG are almost identical. Although the long-wavelength amplitude trends are slightly different due to different acquisition geometries, fortunately, our WEMVA inversion algorithm balances the amplitude of the reflectors so that the amplitude variance will not adversely affect the inversion.

Inversion result

We use the Polak-Riberie variant of nonlinear conjugate-gradient (CG) methods to performe the WEMVA iterations. Li (2014) presents a detailed recipe for this algorithm. Based on our empirical observation, we use only half of the data frequencies when applying the wave-equation tomographic operator to further save computation time. In addition, I re-parameterize the model space using coarsely sampled B-spline nodes (as introduced by Biondi (1990)) to control the resolution of the inversion among iterations. By gradually increasing the number of spline nodes as the number of iterations increases, I constrain the inversion to resolve the low-wavenumber part of the model first, then gradually move up to a higher wavenumber to retrieve more model details. The initial spline nodes spacing is set as 1.5 km in X, 1.5 km in Y, and 0.3 km in Z, and the spacing reduces by ~ 0.1 km in X,Y and 0.02 km in Z for each iteration. When computing the model gradient, we use the aforementioned salt area mask to zero out any updates inside the salt. For the ADCIG sampling, we use 21 points along the reflection angle axis ranging from -25.0° to $+25.0^\circ$, and we compute 7 azimuths from -67.5° to $+67.5^\circ$.

Figure 7 shows the preliminary velocity estimation result we have obtained so far. For this result, Our RMO-based WEMVA inversion stopped improving after 5 iterations, and the objective function reduces by $\sim 5\%$ in total. An interesting point about the inverted velocity model is that the tomography is suggesting a low velocity zone beneath the salt body, and the lowest value in that zone is ~ 2650 km/s, which is $\sim 10\%$ lower than the initial velocity model.

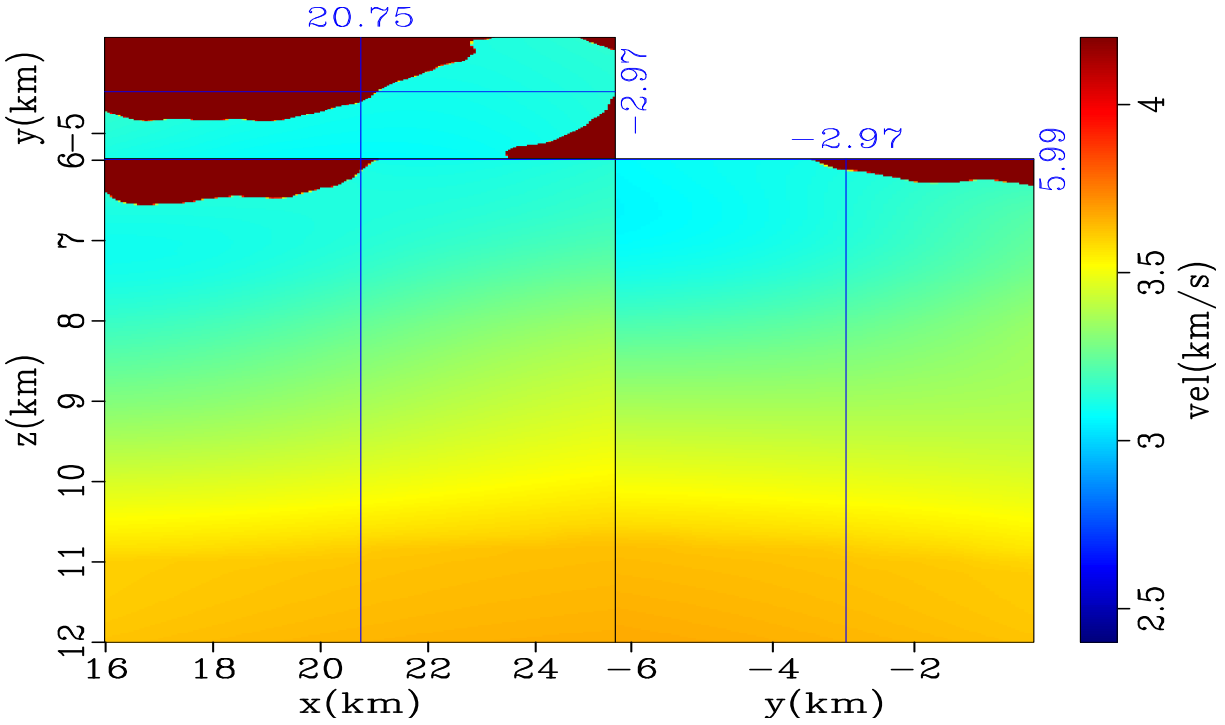
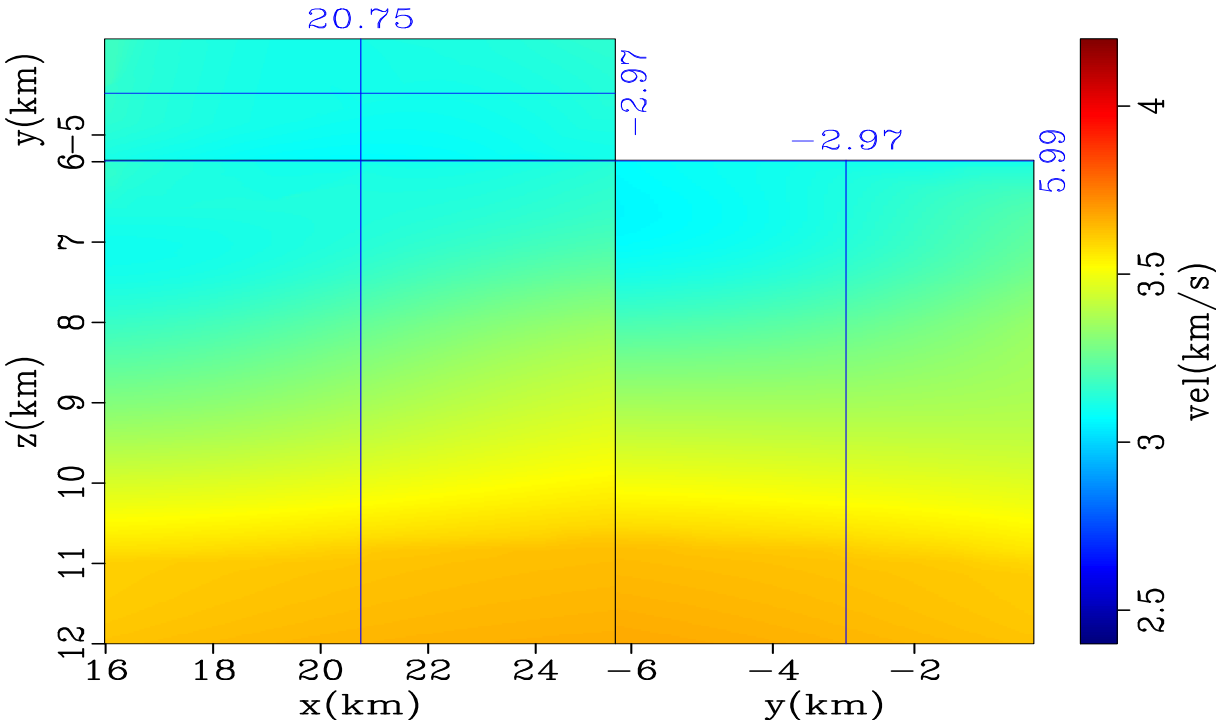
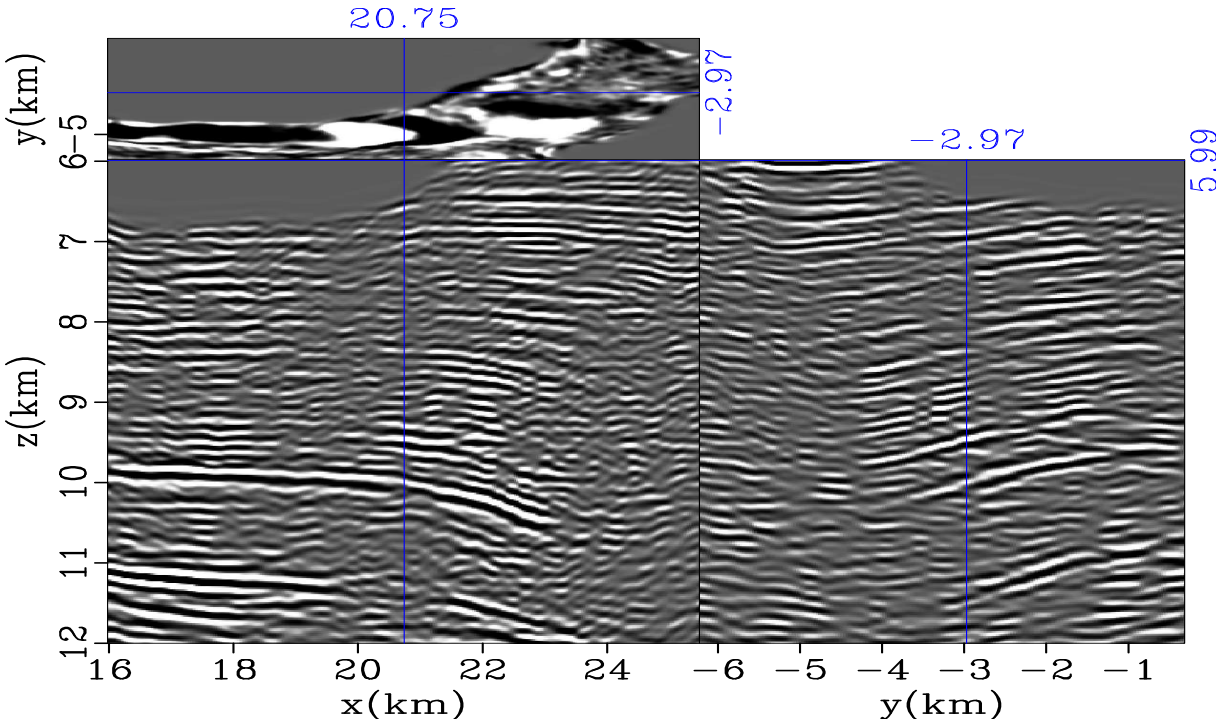


Figure 3: A close-up view of (a) the starting velocity model and (b) the zero subsurface-offset image at the target region (figure 2). [**CR**] yang1/. vel-roi,img-roi



(a)



(b)

Figure 4: A close-up view at the target region. (a): the sediments-flooded starting velocity model, which is used for new data synthesis as well as the starting velocity model in later WEMVA inversion; (b): the zero subsurface-offset image of the ODCIG that we use for new data synthesis. We remove the salt reflection from the original ODCIG and apply inverse of diagonal Hessian matrix to the masked ODCIG for amplitude balancing. Compare this figure with figure 3. [**CR**] yang1/. velfill-roi,img-roi-bal-saltmask

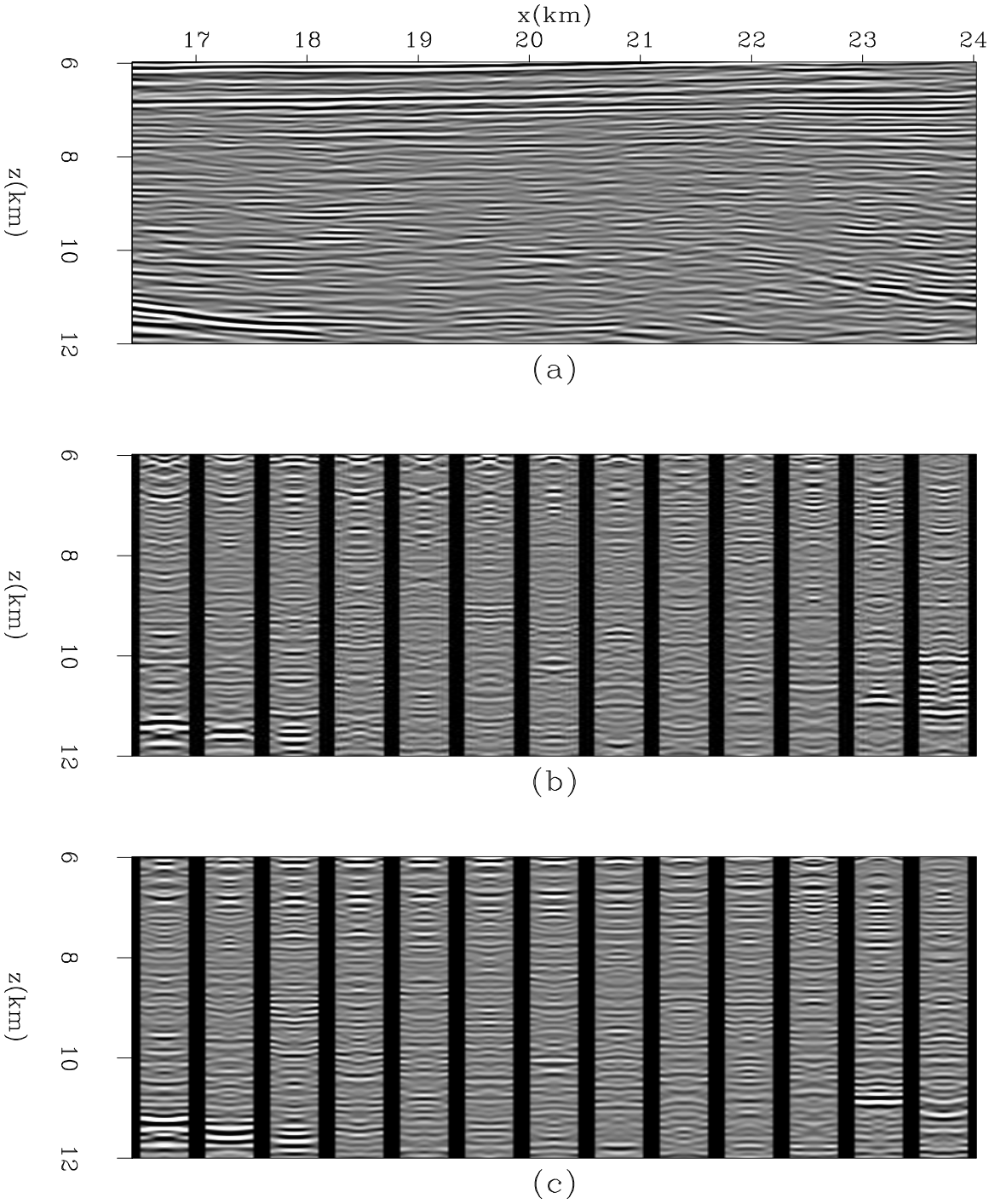


Figure 5: 3-D subsurface ODCIG in the target region, used as the input ODCIG for modeling the new data set. (a): An inline image section at $Y = -4.77\,\text{km}$; (b): the inline subsurface-offset (h_y is fixed at 0 km while h_x range spans $[-0.50\,\text{km}, +0.5\,\text{km}]$) CIG for different lateral locations in (a); (c): the crossline subsurface-offset (h_x is fixed at 0 km while h_y range spans $[-0.54\,\text{km}, +0.54\,\text{km}]$) CIG for different lateral locations in (a). [**CR**]
yang1/. odcigs-initmig1

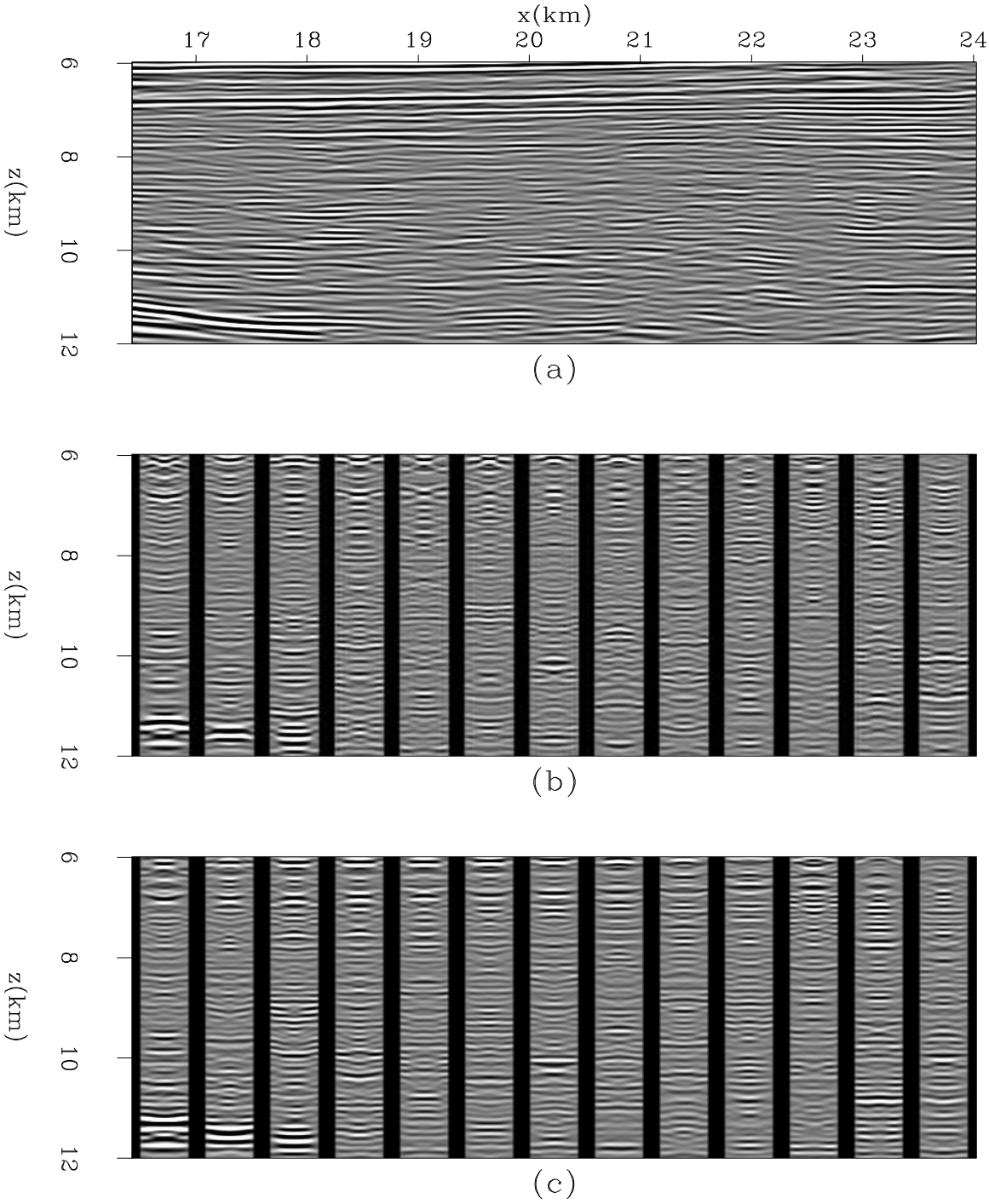


Figure 6: 3-D subsurface ODCIG in the target region, migrated using the generalized Born-modeling data set. Compare the figure with figure 5. (a): An inline image section at $Y = -4.77\,\text{km}$; (b): the inline subsurface-offset (h_y is fixed at 0 km while h_x range spans $[-0.50\,\text{km}, +0.5\,\text{km}]$) CIG for different lateral locations in (a); (c): the crossline subsurface-offset (h_x is fixed at 0 km while h_y range spans $[-0.54\,\text{km}, +0.54\,\text{km}]$) CIG for different lateral locations in (a). [**CR**] yang1/. odcigs-remig1

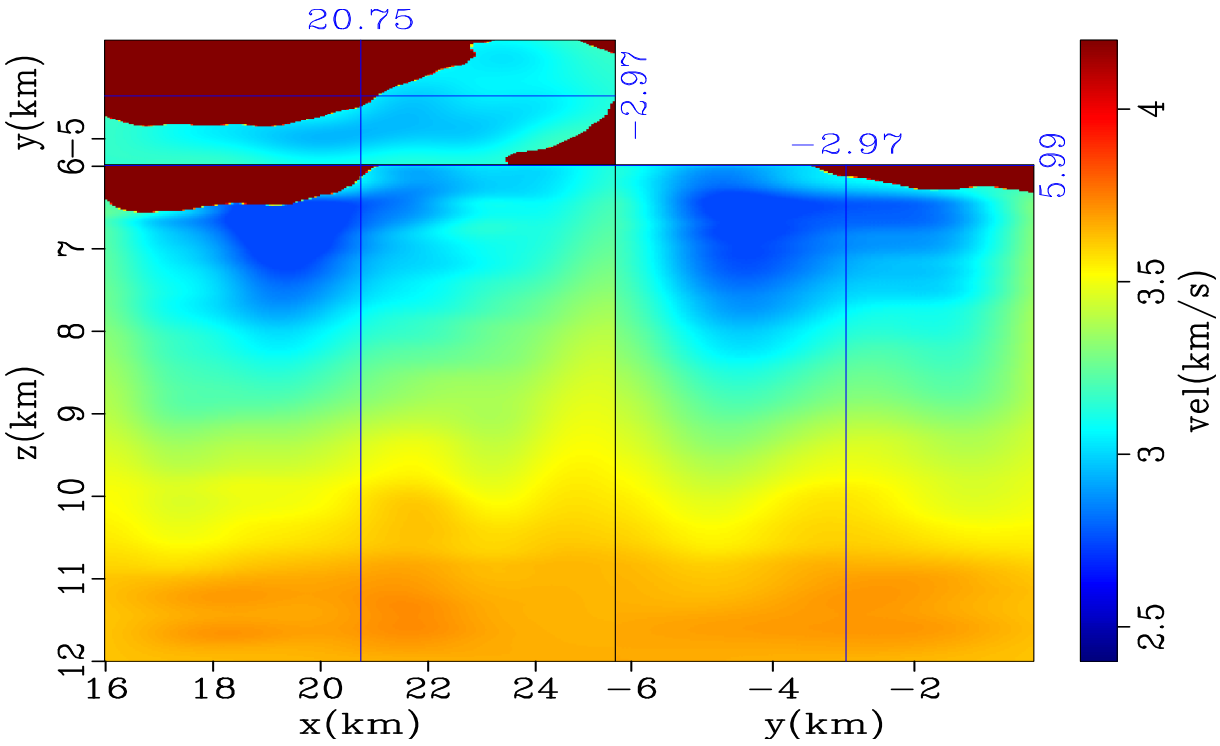
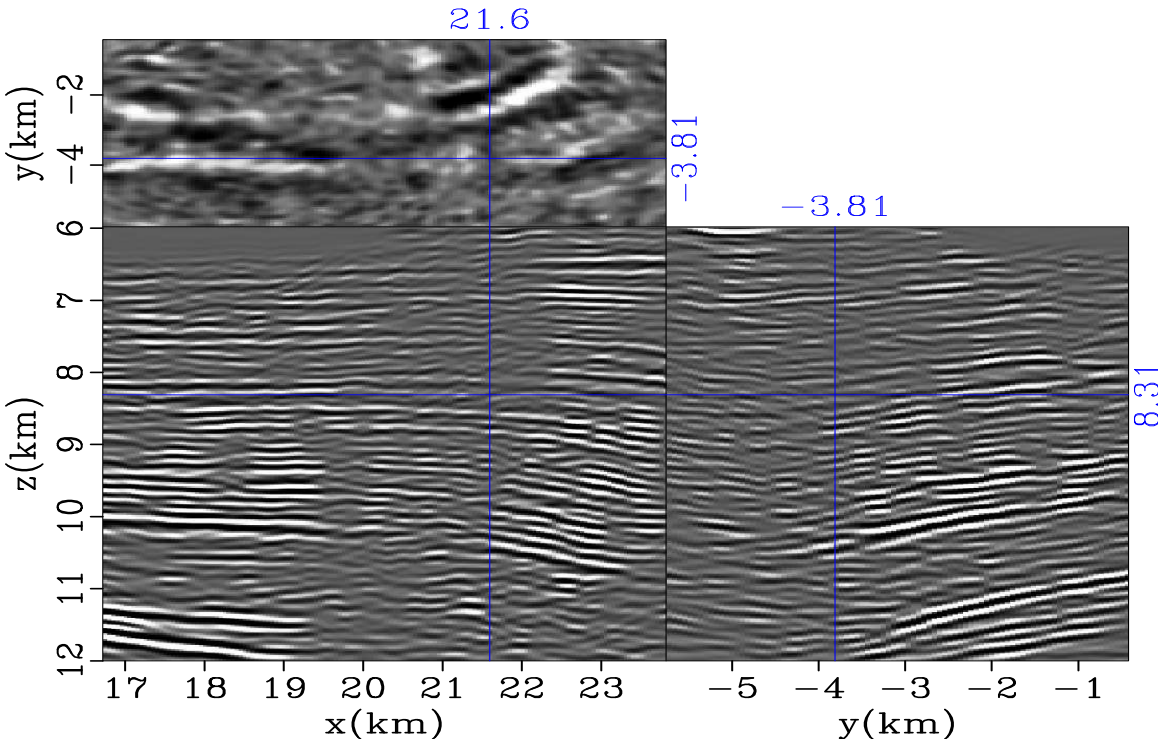
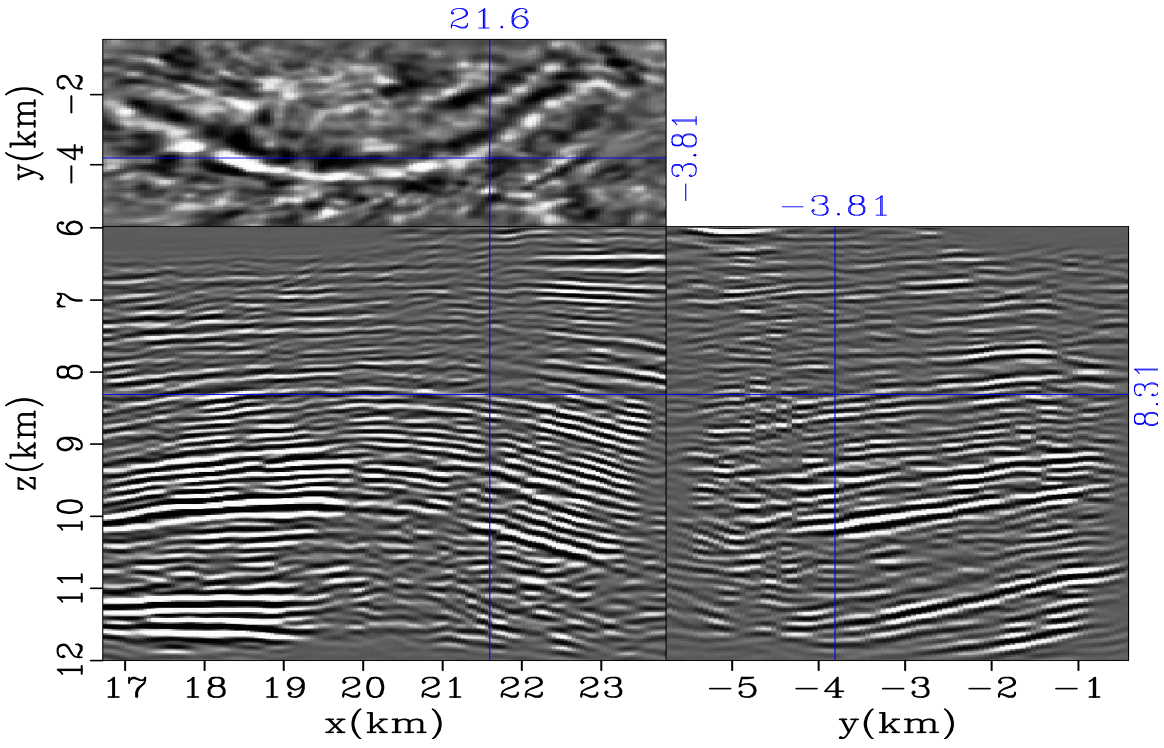


Figure 7: The inverted target region velocity model using the propose RMO-based WEMVA. The original salt area has been restored before plotting. Note the low velocity values immediately beneath the salt. [**CR**] yang1/. velinv-roi



(a)



(b)

Figure 8: Three-panel view showing the comparison between the zero subsurface-offset images migrated with (a) the starting velocity model (figure 4(a)) and (b) the inverted velocity model (figure 7). The coordinates of each section are annotated in the figure. Both figures are plotted using the same clip value. [**CR**] yang1/. img1-roi-beginv,img1-roi-endinv

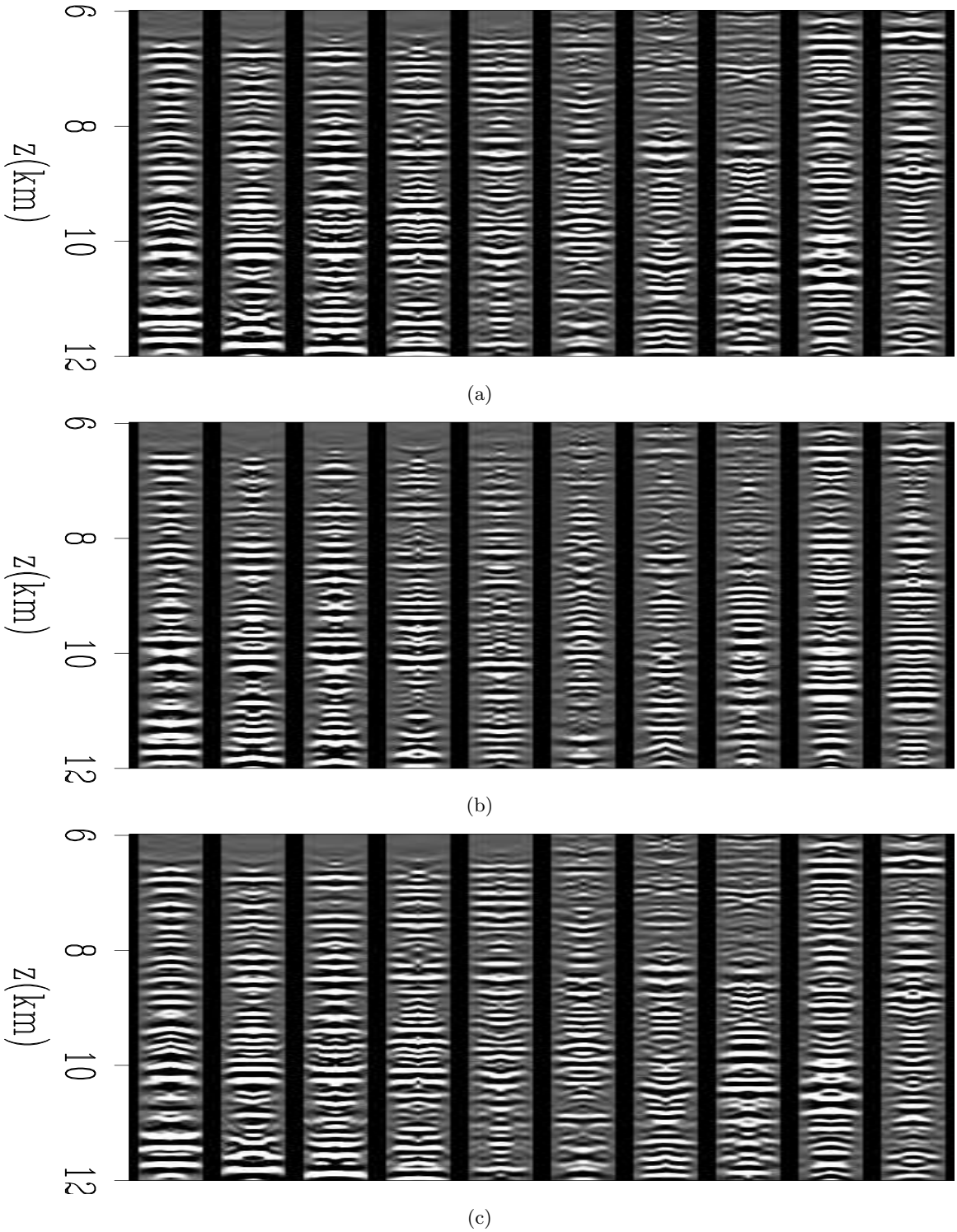


Figure 9: 3-D ADCIG migrated using the starting velocity model (figure 4(a)). The angle gathers are extracted from the same location where the X-Z image section in figure 8 is located. Each subplot shows the angle gathers at a certain azimuth for different lateral locations in X. The reflection angle range spans $[-25.0^\circ, +25.0^\circ]$. The three subplots ((a),(b) and (c)) show the angle gathers at 0°, -45.0° and 45.0° azimuth, respectively. yang1/. adcig-beg1-roi-azim-0,adcig-beg1-roi-azim-45,adcig-beg1-roi-azim–45

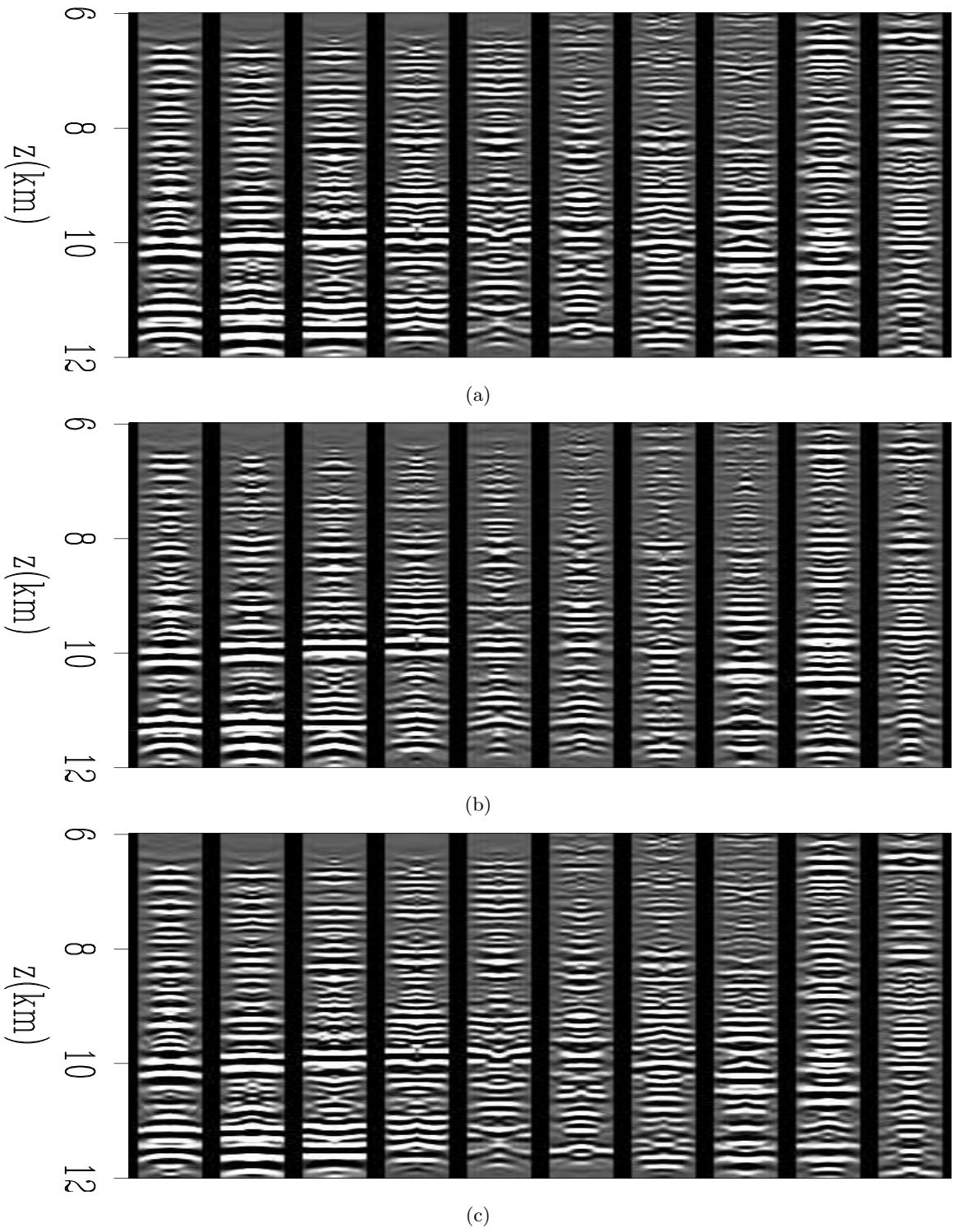
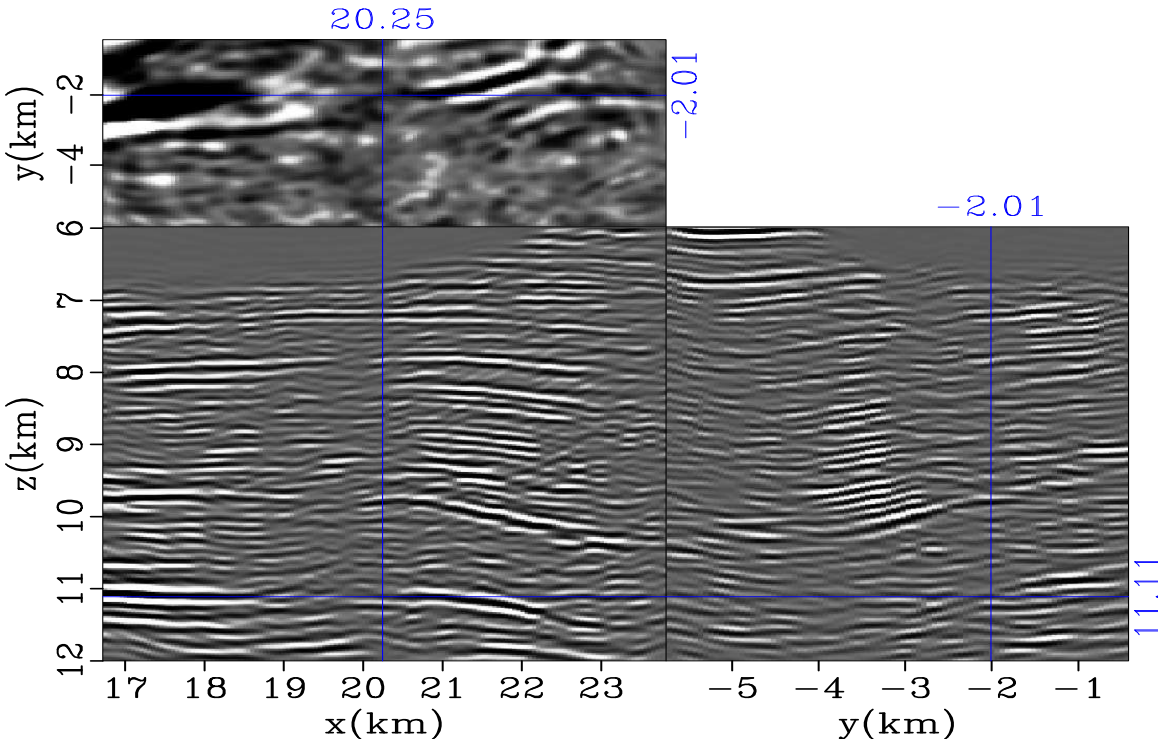
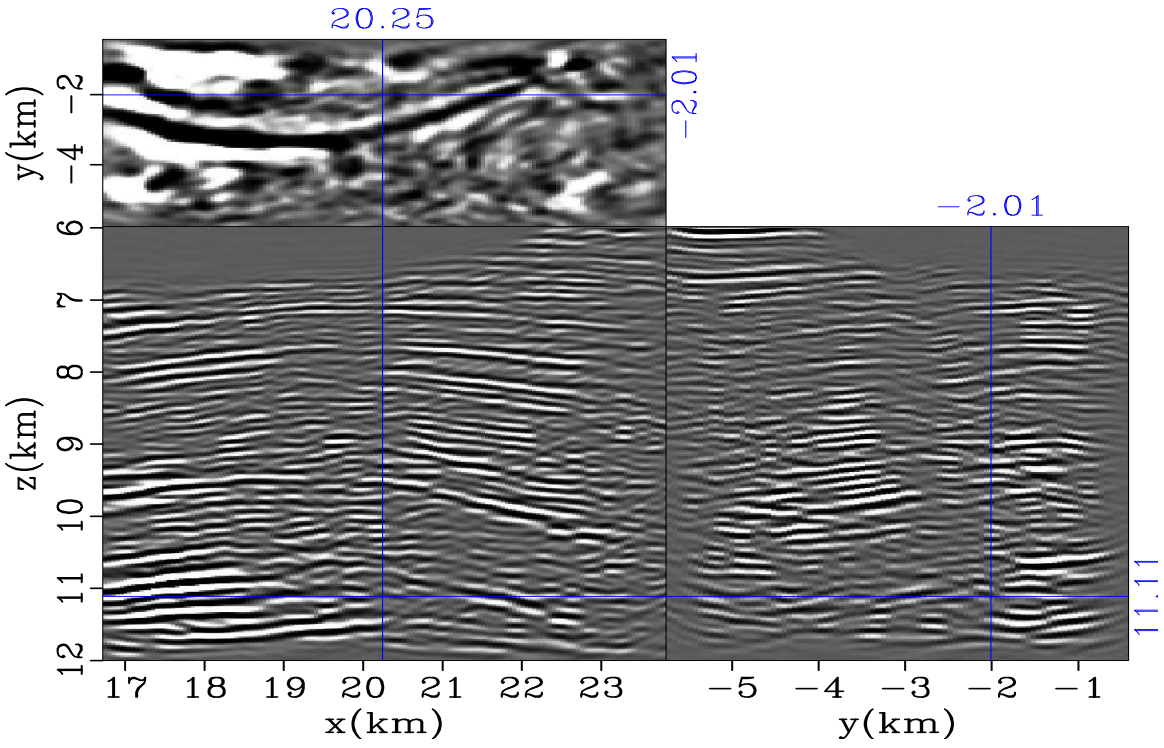


Figure 10: 3-D ADCIG migrated using the inverted velocity model (figure 7). The subplot description is the same as in figure 9. Compare this figure with figure 9. yang1/. adcig-end1-roi-azim-0,adcig-end1-roi-azim-45,adcig-end1-roi-azim–45



(a)



(b)

Figure 11: Another three-panel view showing the comparison between the zero subsurface-offset images migrated with (a) the starting velocity model (figure 4(a)) and (b) the inverted velocity model (figure 7). The coordinates of each section are annotated in the figure. Both figures are plotted using the same clip value. [**CR**]

yang1/. img2-roi-beginv,img2-roi-endinv

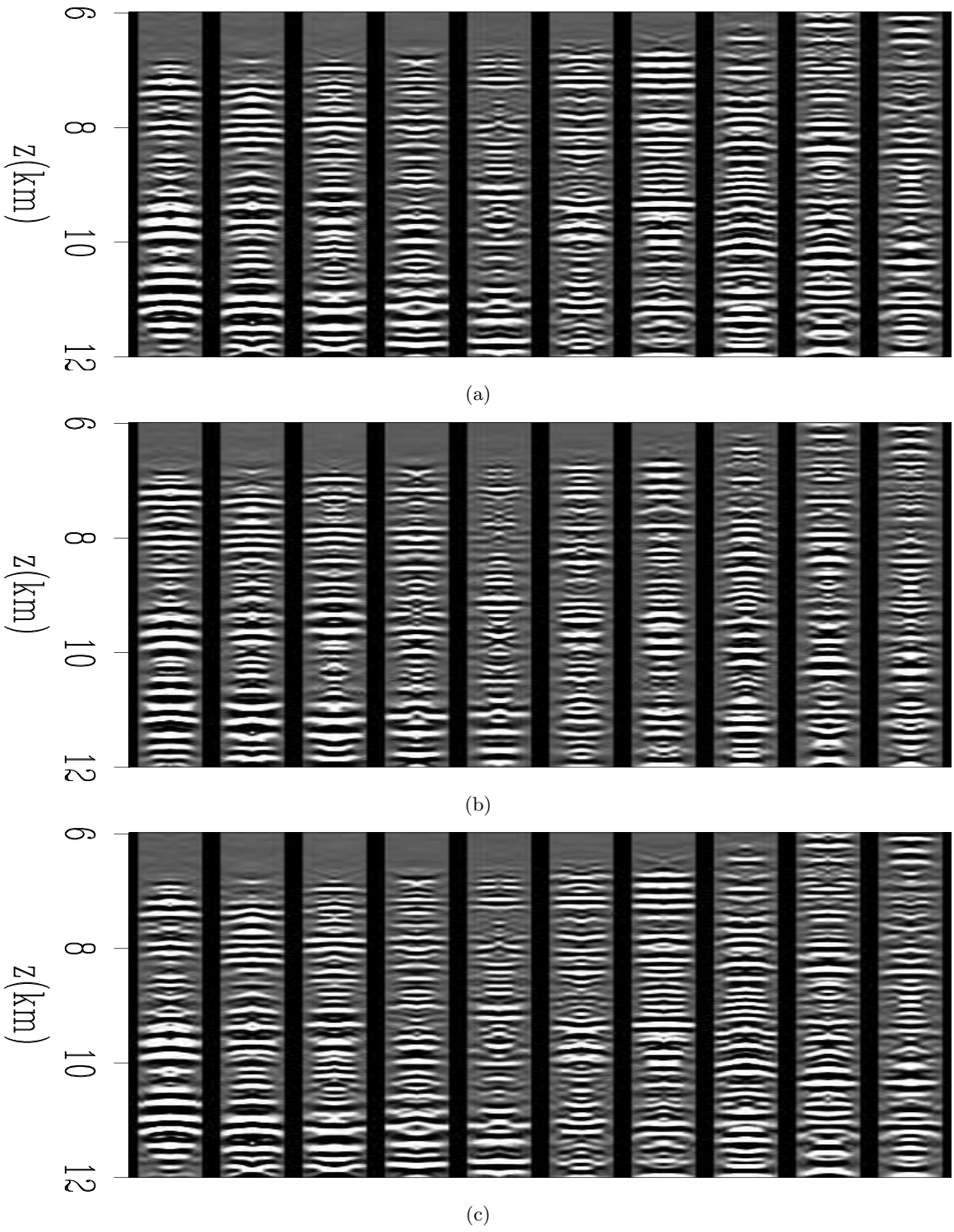


Figure 12: Same description as in figure 9, other than that the locations of these angle gathers are from the X-Z image section in figure 11. [**CR**] yang1/. adcig-beg2-roi-azim-0,adcig-beg2-roi-azim-45,adcig-beg2-roi-azim–45

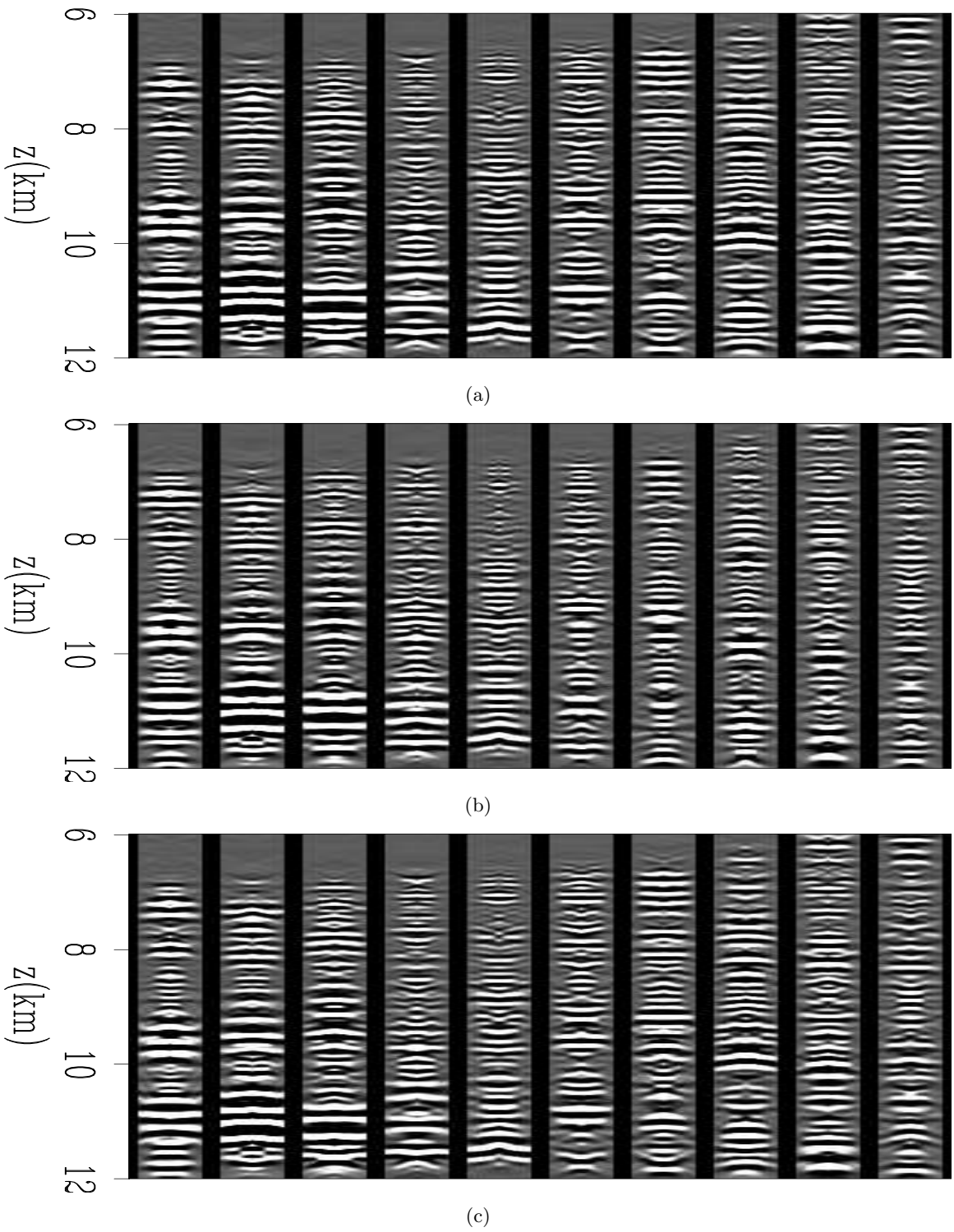


Figure 13: Same description as in figure 10, other than that the locations of these angle gathers are from the X-Z image section in figure 11. Compare this figure with figure 12. [**CR**] yang1/. adcig-end2-roi-azim-0,adcig-end2-roi-azim-45,adcig-end2-roi-azim–45

To further verify that our inverted model is indeed improving the subsalt images, we plot the comparisons between the initial image and the image migrated using the inverted model, as shown in figure 8 and figure 11. To compare the image focusness, we use the same clip value for the initial image and the updated image. As we can see from both figures, the events of the updated image become more coherent, thus having higher amplitudes. Especially in figure 11 an anticline structure is formed because of the low velocity zone, and the continuity of those reflectors are also significantly improved. Figure 9 and 10 show the comparison between the ADCIG before and after update. The locations of these ADCIG are the slicing locations shown in figure 8. (In this example, it is the Z-X slice at $Y = -3.81\,\text{km}$.) Similarly, figure 12 and 13 show the ADCIG comparison at a different location corresponding to figure 11. From these ADCIG comparisons, we can see that the majority of the ADCIG become more flat after velocity update, although there are certainly still some room for improvement even with the updated velocity.

CONCLUSION

We present a field data application of our target-oriented, RMO-based WEMVA tomography workflow using the E-Octopus III WAZ data set. The generalized Born-modeling based target-oriented approach we use enables huge computational savings, while preserving the velocity information in the original data losslessly. Although our current inversion result on a subsalt sediments region is preliminary, and the subsalt area is very challenging for velocity estimation; nonetheless, the inverted velocity model from our RMO-based WEMVA demonstrates convincing imaging improvement and uncovers an interesting low velocity zone beneath the salt, which might be worthy to investigate from geological aspects.

ACKNOWLEDGMENT

We thank Schlumberger for providing the E-octopus III data set used in this study.

APPENDIX A

AUTOMATIC EVENT DETECTION

In this appendix, we explain how we implement the automatic event detection module for the gradient calculation in the RMO-based WEMVA method.

Anchor points detection

The basic idea for our anchor point detection workflow is inspired by the methodology proposed in Cullison (2011). First our method uses several image processing techniques to extract all potential event anchor point candidates, and then filter out a portion of candidates based on several geophysical criteria before outputting the final set of anchor points. Let me denote the stacked image (zero subsurface offset image) $I(z, x)$. To find candidate anchor points, I perform the following steps:

1. Apply 1-D (along z axis) automatic gain control (AGC) to the stacked image $I_{bal}(z,x)$, in order to broadly balance the amplitude of reflectors in the image.

2. Take the energy envelope (magnitude of the input's Hilbert signal) along z axis on $I_{bal}(z,x)$, denoting the output as $I_E(z,x)$.

3. Apply *non-maximum suppression* (along z axis) to $I_E(z,x)$. As its name suggests, non-maximum suppression will suppress all samples that are not local maxima in the input image to zero. It will produce a binary image of the input size, and only the locations of local maxima in the input image are set to 1, while all the rest of the output are set to 0. From this binary image, I extract the locations of all non-zero values and place them in the candidates set of the anchor points, S_{cand}.

However, not all candidate points in S_{cand} are well-suited for WEMVA gradient back-projection. In other words, extra quality check (QC) is helpful to prevent noisy data going to the gradient calculation. Specifically, our implementation filters out any candidate points if they do not meet one of the criteria below:

1. Because the assumption of specular reflection is the theoretic foundation of all reflection tomography methods, I screen out all candidate points that do not satisfy specular reflection requirement. I achieve this screening by computing *linearity coefficients* over all candidate points and threshold the values of linearity coefficients.

   First introduced into the field of exploration Geophysics by Hale (2009a,b), the *linearity coefficient* is a good measure to quantify how specular a reflection point is. Briefly speaking, if the linear coefficient of the target reflection point is high (close to 1.0), it indicates that there is a strong local linear reflector structure around that point. Thus, we know that the specular reflector assumption holds well. Similarly, if the linear coefficient is low (close to 0.0), it indicates that there is no locally linear reflector structure around that point (for example, the target point is on fault boundary, or is just an isolated diffractor). Therefore, the specular reflector assumption does not hold.

   The local linearity coefficients of the stacked image can be conveniently estimated using structure tensor (Cullison, 2011).

2. I also filter out candidate points that have a large dipping angle, because neither the theory nor the way that I compute gathers handles steeply dipping reflectors well. The reason is explained as follows:

   - In the theoretic aspect, our derivation assumes that the ADCIG shift only vertically under velocity perturbation. In the case of reflector dips being large, this vertical shift assumption will be challenged because it is more accurate to assume each angle-domain gather shifts along the normal direction (perpendicular to the corresponding dipping interface).
   - On the implementation side, note that I compute the ADCIG indirectly by converting from subsurface offset common image gathers (Sava and Fomel, 2003). The horizontal subsurface offset (h_x) common image gathers are known to lose resolution and accuracy for angle-domain gathers on steeply dipping reflectors (Biondi, 2006).

Therefore our workflow also filters out steep reflectors. As for how to determine the steepness of reflectors, we can conveniently estimate local dipping angles of the stacked image using structure tensors as well (Cullison, 2011).

Variable event window size

After I identify the set of event anchor points, I can further detect the actual window size of each event instead of using a pre-set constant window size universally:

1. For each anchor point, extract the amplitude at that anchor location from the energy envelope image $I_E(z, x)$.
2. We can infer that each anchor point positions approximately at the center of the event it belongs to, because the anchor point samples the peak amplitude of the energy envelope. Therefore our code starts from the anchor point's location, then searches in both directions ($\pm z$) for the tails of the energy envelope.

Illustrative example on automatic event detection

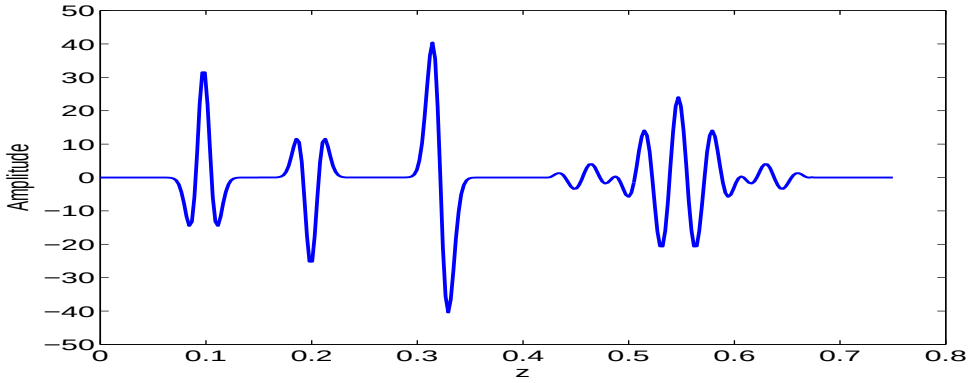
As an intuitive explanation for the algorithms used in our automatic event detection module, I present a simplified event detection example on a 1-D seismic trace input. In the RMO-based WEMVA case, we can think of the input as an image trace along the depth direction at a certain horizontal location, $I(z, x = x_0, y = y_0)$.

Figure 1(a) shows the input 1-D signal. The goal is to detect the individual events accurately and efficiently. As we can visually identify from the plot, I deliberately create 4 distinct events in the input, each with a different signature (waveform):

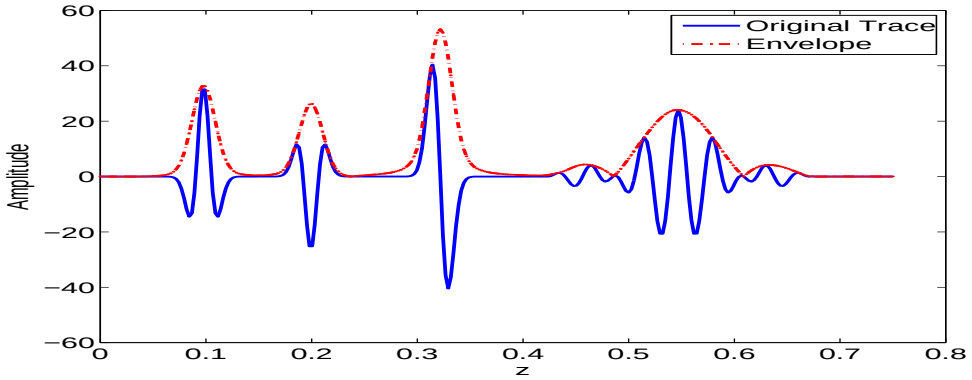
1. A typical Ricker2 wavelet (second derivative of Gaussian).
2. A Ricker2 wavelet with opposite polarity.
3. A Ricker1 wavelet (first derivative of Gaussian). This wavelet is asymmetric and has a 90° phase difference compared to the Ricker2 wavelet.
4. A typical Ormsby wavelet (Ryan, 1994). This wavelet usually has bigger side lobes (more rippling) compared to the Ricker wavelet due to the steeper tapering effect on the edges of the spectrum.

These events possess different signal characteristics (variant amplitude, polarity, phase and spectrum shape) that help to demonstrate the merits of the proposed approach.

Figure 1(b) shows the energy envelope overlaying the input. As we can see, the envelope strips out many of the complexities in the original signal (like asymmetric phase, negative polarity, vibrating waveform, etc.), and singles out the "wave packet" information for each event. The simplicity brought by the energy envelope makes it a much better choice for anchor point detection than the original signal, as demonstrated by the following example.

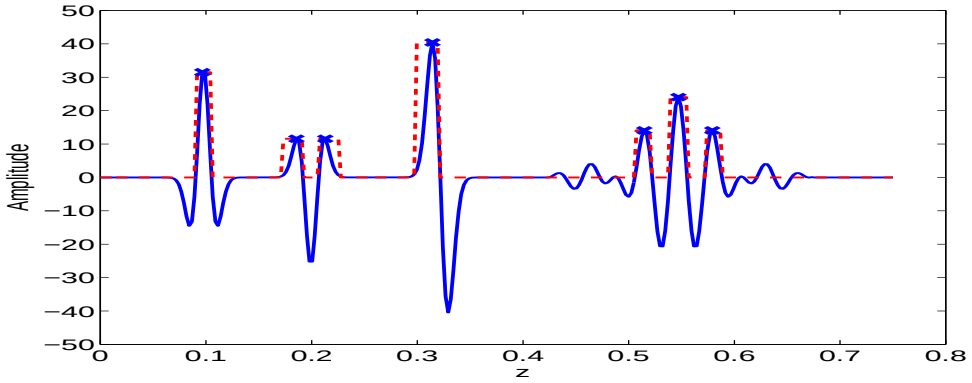


(a)

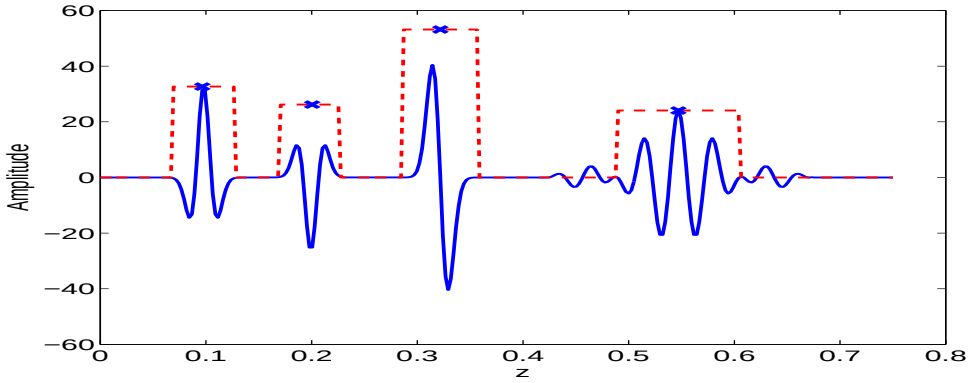


(b)

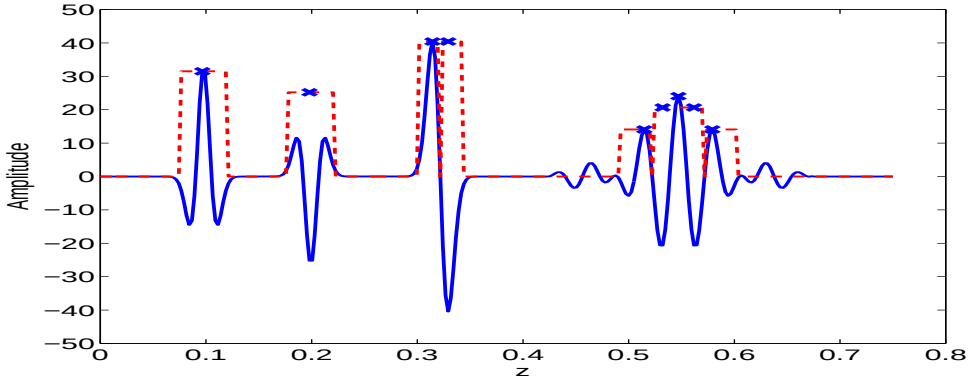
Figure A-1: An example to illustrate the event detection workflow. (a): the input 1-D signal; (b): the input with its energy envelop overlaid. [**CR**] yang1/. trace-input,trace-envelop-overlay



(a)



(b)



(c)

Figure A-2: The results of detected anchor points(shown as × marks) and event windows (shown as dash lines) using different inputs: (a) the original signal as input; (b) the energy envelope as input; (c) the absolute value of the original signal as input. [**CR**] yang1/. eventwid-ori-maxima,eventwid-env-maxima,eventwid-abs-maxima

Figure A-2 shows the anchor points found by applying non-maximum suppression to different input signals. More specifically, figure 2(a) uses the original signal as detection input, and figure 2(b) uses the energy envelope. For comparison, I add the result using the absolute value of the original signal as input, shown in figure 2(c). Apparently, the result using the energy envelope (figure 2(b)) is the most accurate. The result in figure 2(a) fails on the second event that has a negative polarity. Because of this negative polarity, the two local maxima it finds are indeed the peaks of two side lobes rather than the main lobe. This result also fails on the last event due to the severely vibrating waveform, that the computer program picks up both the peak of main lobe and the first two strong side lobes. The result in figure 2(c) is able to handle negative polarity in the second event because it uses the absolute value of the original signal as input. However, it fails on the third and fourth event. The result contains several false positives because the absolute value function converts local minima into local maxima.

After the event anchor points are located, figure A-2 also shows the detected event window sizes of every anchor point for each corresponding test case. The detected event windows are drawn as box functions with dash lines. As the figure indicates, the result using energy envelope is the most accurate.

REFERENCES

Biondi, B., 1990, Seismic velocity estimation by beam stack: PhD thesis, Stanford University.

Biondi, B. L., 2006, 3-D Seismic Imaging.

Cullison, T. A., 2011, An image-guided method for automatically picking common-image-point gather locations from seismic images: Master's thesis, Colorado School of Mines.

Hale, D., 2009a, Image-guided blended neighbor interpolation of scattered data: 1127–1131.

———, 2009b, Structure-oriented smoothing and semblance: CWP-Report, **635**, 261–270.

Li, Y., 2014, Wave-equation migration velocity analysis for VTI models with geological and rock physics constraints: PhD thesis, Stanford University.

Ryan, H., 1994, Ricker, Ormsby; Klauder, Butterworth - A choice of wavelets: CSEG Recorder, **19**, 8–9.

Sava, P. C. and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074.

Tang, Y., 2011, Imaging and velocity analysis by target-oriented wavefield inversion: PhD thesis, Stanford University.

Zhang, Y., 2015, Velocity model building using residual-moveout-based wave-equation migration velocity analysis: PhD thesis, Stanford University, (to be published).

Zhang, Y. and B. Biondi, 2013, Moveout-based wave-equation migration velocity analysis: GEOPHYSICS, **78**, U31–U39.

———, 2014, Residual-moveout-based WEMVA: A WAZ field data example - Part I : SEP-Report, **152**, 223–238.

Zhang, Y., B. Biondi, and R. Clapp, 2013, Accelerating residual-moveout-based wave-equation migration velocity analysis with compressed-sensing: SEP-Report, **149**, 95–106.

Amplitude normalization in TFWI

Ali Almomin

ABSTRACT

Tomographic Full Waveform Inversion (TFWI) provides a robust and accurate method to invert the seismic data by simultaneously inverting all scales of the model. However, maintaining simultaneous inversion of scales is hindered when the modeling operator cannot accurately match the amplitudes of the data. In this paper, I modify TFWI to increase robustness against amplitude inaccuracies in the modeling operator by normalizing the objective function using a running-window filter. Finally, I test the proposed algorithm on the Marmousi. The results of the modified TFWI show a major improvement in the accuracy and convergence rate of the inversion.

INTRODUCTION

TFWI, similar to other data-space inversion method, produces highly accurate results due to matching both the phase and amplitude of the data. This is achieved in two steps: first, extending the wave equation and adding an additional axis to the velocity model, and second, adding a regularization term that drives the solution towards a non-extended model. However, one limitation to TFWI occurs when the amplitude of the data cannot be accurately matched by the modeling operator. One solution is to only match the phase using a single frequency per iteration (Pratt, 1999; Shin and Ha, 2008). Using phase only will prevent simultaneous inversion of scales. Another approach is to normalize each trace by its norm, as presented in Shen (2014). The issue with trace normalization is that it does not take into account the large difference in amplitude behavior between the transmission and reflection data, which makes it only usable when inverting a few events to match.

In this report, I generalize the amplitude normalization inversion to use any nonlinear weighting function that is based on the data. Then, I propose a running window normalization that uses a Gaussian function to extract the local amplitude of the data. Finally, I compare different normalizations using a synthic Marmousi model.

THEORY

The conventional L2 objective function for data matching can be written as follows:

$$J_{\mathrm{FWI}}(\mathbf{m}) = \frac{1}{2}\|\mathbf{f}(\mathbf{m}) - \mathbf{d}_{\mathrm{obs}}\|_2^2, \tag{1}$$

where $\mathbf{m}$ is the model, $\mathbf{f}$ is the forward modeling operator and $\mathbf{d}_{\mathrm{obs}}$ is the observed surface data. I modify the conventional L2 objective function by adding a nonlinear weighting function as follows:

$$J_{\mathrm{W-TFWI}}(\widetilde{\mathbf{m}}) = \frac{1}{2}\|\mathbf{W}(\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}}))\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}}) - \mathbf{W}(\mathbf{d}_{\mathrm{obs}})\mathbf{d}_{\mathrm{obs}}\|_2^2, \tag{2}$$

where $\mathbf{W}$ is a diagonal weighting matrix. The elements on the weighting matrix can be described as follows:

$$\mathbf{W}(it, ir, is; \widetilde{\mathbf{f}}) = \frac{1}{\|\mathbf{A}(it, ir, is)\widetilde{\mathbf{f}}\|_2}, \tag{3}$$

where it, ir, is are the time, receiver and source coordinates and the matrix $\mathbf{A}$ represents a convolution by an abstract non-stationary filter. Notice that the norm used in the previous equation is only along the lag axis of the convolution. In other words, the weighting matrix divides each sample by the norm of a filtered version of the data. The purpose of the filter is to window and taper around each sample of the data.

The gradient $\mathbf{g}$ can be calculated as:

$$\widetilde{\mathbf{g}}(\widetilde{\mathbf{m}}) = \left(\frac{\partial \mathbf{r}(\widetilde{\mathbf{m}})}{\partial \widetilde{\mathbf{m}}}\right)^* \mathbf{r}(\widetilde{\mathbf{m}}). \tag{4}$$

Using the definition of the matrix $\mathbf{W}$ and the chain-rule, I can calculate the derivative of the residual with respect to the model as:

$$\frac{\partial \mathbf{r}(\widetilde{\mathbf{m}})}{\partial \widetilde{\mathbf{m}}} = \left(\mathbf{W}(\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}})) + \frac{\partial \mathbf{W}(\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}}))}{\partial \widetilde{\mathbf{f}}}\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}})\right)\frac{\partial \widetilde{\mathbf{f}}(\widetilde{\mathbf{m}})}{\partial \widetilde{\mathbf{m}}}. \tag{5}$$

The derivative of the weighting matrix with respect to the modeled data can be expressed as:

$$\frac{\partial \mathbf{W}(\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}}))}{\partial \widetilde{\mathbf{f}}} = -\mathbf{W}(\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}}))\frac{\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}})^*\mathbf{A}^*\mathbf{A}}{\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}})^*\mathbf{A}^*\mathbf{A}\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}})}. \tag{6}$$

I can now substitute the previous terms and the gradient as the backprojection of a virtual source $\mathbf{v}(\widetilde{\mathbf{m}})$ described as:

$$\mathbf{v}(\widetilde{\mathbf{m}}) = \mathbf{W}(\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}}))\left(\mathbf{r}(\widetilde{\mathbf{m}}) - \frac{\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}})^*\mathbf{A}^*\mathbf{A}\mathbf{r}(\widetilde{\mathbf{m}})}{\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}})^*\mathbf{A}^*\mathbf{A}\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}})}\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}})\right). \tag{7}$$

In order to interpret the gradient expression, I redefine the weighting matrix as

$$\mathbf{W}(it, ir, is; \widetilde{\mathbf{f_1}}, \widetilde{\mathbf{f_2}}) = \frac{1}{\sqrt{\widetilde{\mathbf{f_1}}^*\mathbf{A}^*(it, ir, is)\mathbf{A}(it, ir, is)\widetilde{\mathbf{f_2}}}}, \tag{8}$$

where the variable $\mathbf{f}_1$ and $\mathbf{f}_2$ are two data sets. This new definition gives the weighting matrix the freedom to divide by the product of two datasets instead of restricting the calculating to one. I can now write the virtual source as:

$$\mathbf{v}(\widetilde{\mathbf{m}}) = \mathbf{W}(\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}}), \widetilde{\mathbf{f}}(\widetilde{\mathbf{m}}))\left(\mathbf{r}(\widetilde{\mathbf{m}}) - \frac{\mathbf{W}^2(\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}}), \widetilde{\mathbf{f}}(\widetilde{\mathbf{m}}))}{\mathbf{W}^2(\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}}), \widetilde{\mathbf{r}}(\widetilde{\mathbf{m}}))}\widetilde{\mathbf{f}}(\widetilde{\mathbf{m}})\right). \tag{9}$$

We can see that the gradient is computed by the following steps. First, the modeled data is divided by the norm of the filtered modeled data. This can be seen as a way to "remove" the amplitudes from the modeled data. Then, the data is scaled by weights which are the result of the product of the modeled data and the residual, which scales the amplitudes of the normalized modeled data to those of the residual. At this point, the modeled data has a similar amplitude scale to the residual, which makes differencing them less prone to scale

differences. After calculating the difference, the resultant is divided by the amplitude of the modeled data again. This last division removes the amplitude imprint of the modeled data that will be imposed by the step to follow, which is the back projection of the data into the model space. It is worth noting that this normalization will automatically incorporate the preconditioning described in Almomin and Biondi (2014)

SYNTHETIC EXAMPLES

I now test the new algorithm on the Marmousi model. I create the observed data with variable density acoustic wave-equation. The first shot gather from the observed data is shown in figure 1. In order to create amplitude discrepancies between the observed and modeled data, I ignore the density variations in the modeled data.

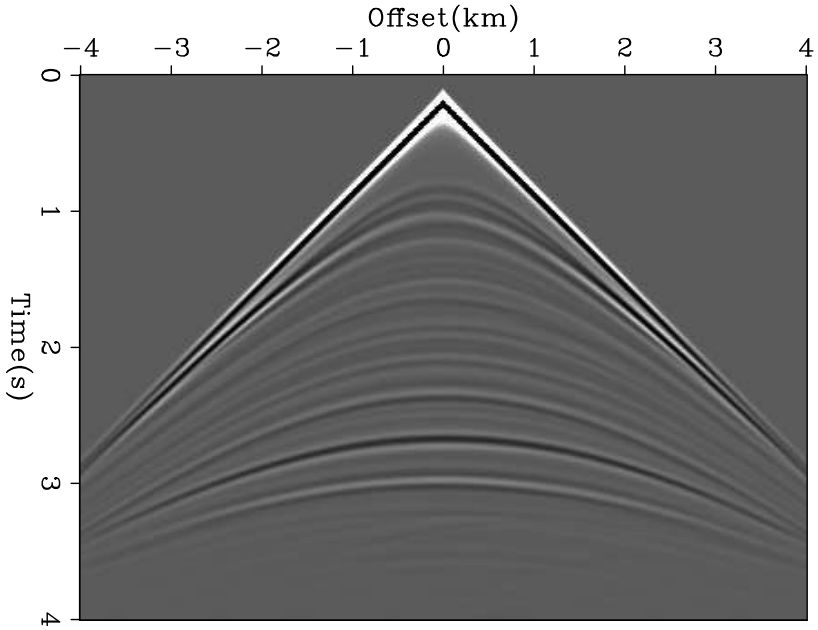


Figure 1: First shot gather from the observed data. [**ER**] ali1/. nl

Using the correct kinematics, the first shot of the modeled data is shown in figure 2. Notice that the amplitude behavior between the two figures 1 and 2 is variable both in offset and time. Next, I compare the first TFWI gradient using four objective functions: conventional L2 difference, gather-normalization, trace-normalization (Shen, 2014) and a running-window normalization using a Gaussian window as formulated in this report. Figure 3 shows the first gradient using the L2 difference. This gradient is dominated by large differences in amplitudes in the direct arrivals. Figures 4 and 5 shows the first gradient using the shot-gather-normalization and trace-normalization difference. This gradient has largely improved compared to the L2 gradient, but we can still notice the diminishing amplitudes of the deeper reflectors. Figure 6 shows the first gradient using the running window-normalization difference. This gradient has significantly improved compared to the other two gradients, especially when examining the deeper reflectors.

Next, I repeat the same experiment but use horizontal average of the velocity as the background, i.e., using the wrong kinematics. Figure 7 shows the first gradient using the

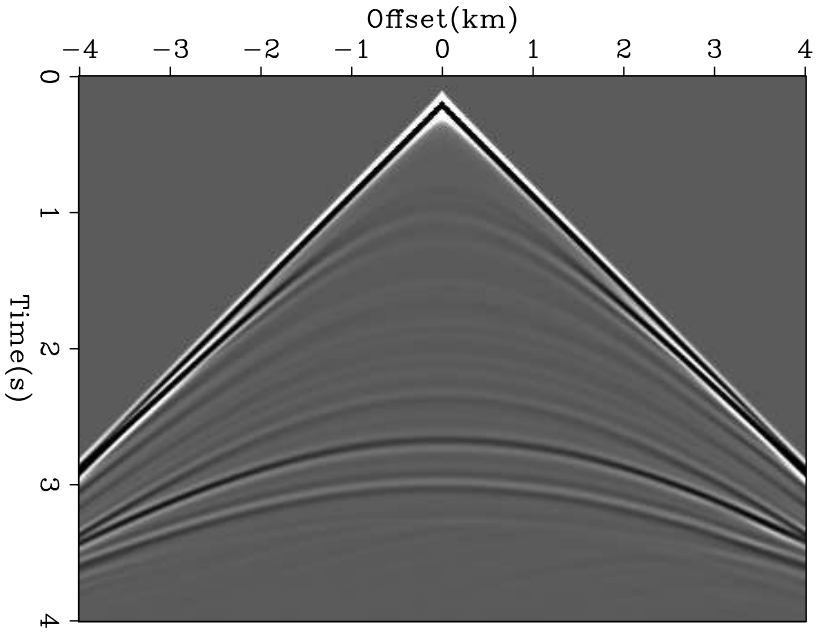


Figure 2: First shot gather from the modeled data using the correct velocity. [**ER**] ali1/. nl-0

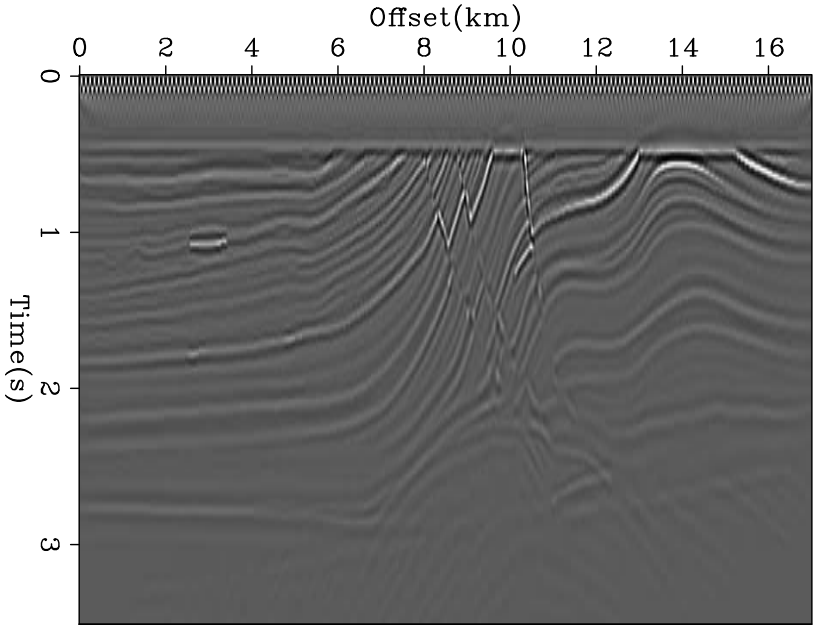


Figure 3: Model gradient using L2 residual and kinematically correct velocity. [**ER**] ali1/. r-0

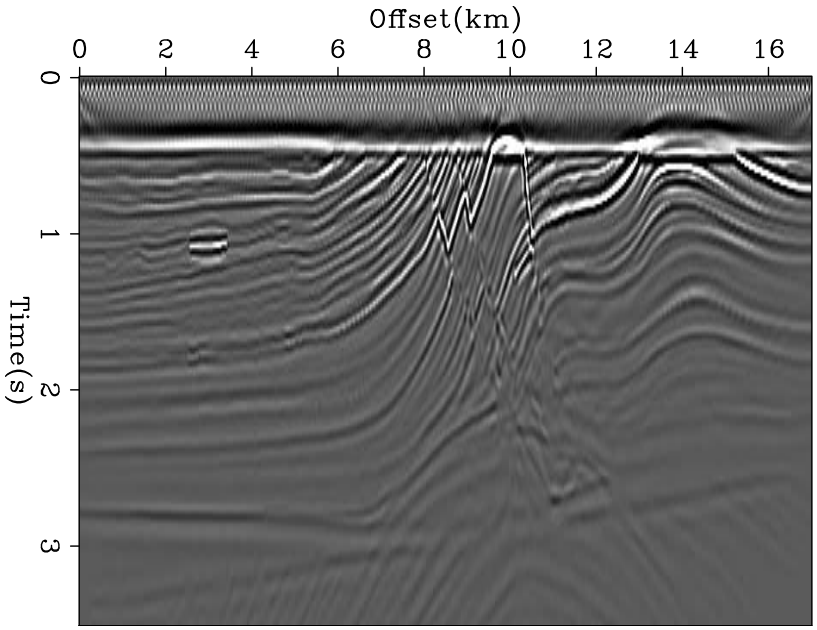


Figure 4: Model gradient using gather-normalized residual and kinematically correct velocity. [**ER**] ali1/. r2-0

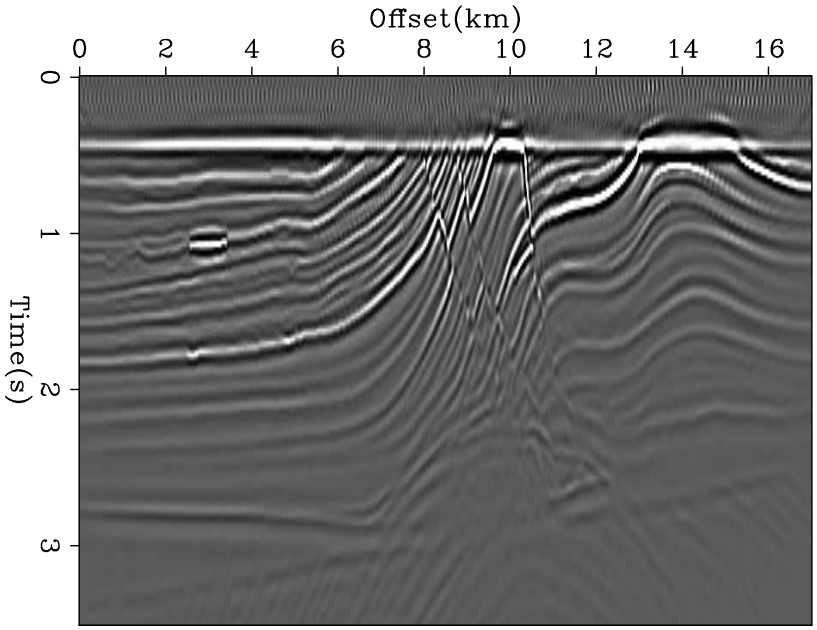


Figure 5: Model gradient using trace-normalized residual and kinematically correct velocity. [**ER**] ali1/. r3-0

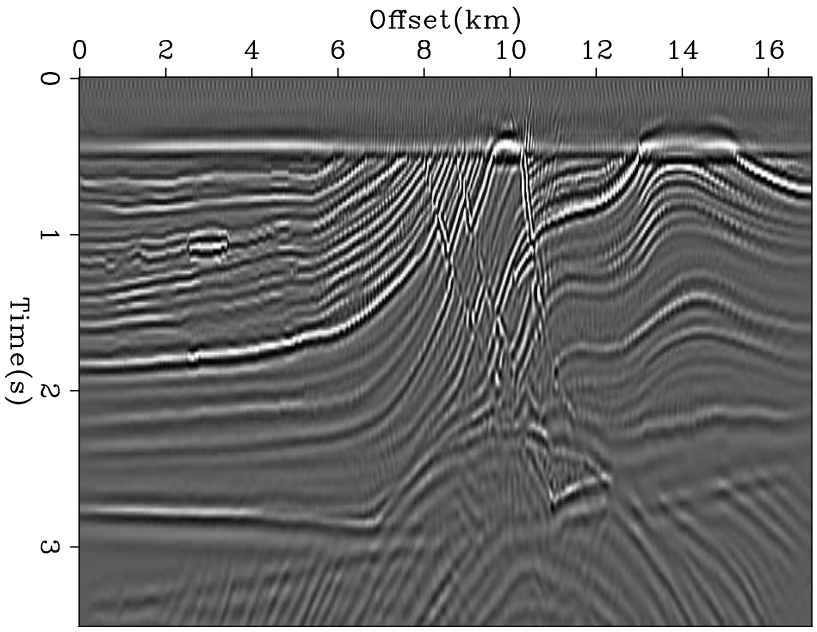


Figure 6: Model gradient using running window-normalized residual and kinematically correct velocity. [**ER**] ali1/. r4-0

L2 difference. Similar to the gradient calculated with the wrong kinematics, this gradient is also dominated by large differences in amplitudes in the direct arrivals. Figures 8 and 9 shows the first gradient using the shot-gather-normalization and trace-normalization difference. While some improvements can be seen compared to the L2 gradient, this gradient is dominated by the first reflector. Figure 10 shows the first gradient using the running-window normalization difference. Also, the low amplitude behavior for the deeper reflectors seems to be compensated for using the running-window normalization.

Figures 11 to 14 show the Angle-domain common-image gathers (ADCIGs) at four different locations of the gradients shown in figures 7 to 10. We can see that using gather-normalization or trace-normalization results in a loss of energy at deeper reflectors, especially near the zero angle. The running window-normalization improved the amplitude-versus-angle behavior of the gathers and produced more balanced gathers.

CONCLUSIONS

Matching the amplitudes of the data is important in the framework of TFWI in order to ensure a simultaneous inversion of scales. However, this task can prove challenging when trying to match the amplitudes of field data. An amplitude mismatch can have a large impact on the result, and can dominate the phase component of the data.

To mitigate this issue, I proposed a normalized objective function using an abstract non-linear weighting function. This formulation is very flexible and allows for several methods of normalization. This flexibility is capable of handing multiple events with varying amplitude-versus-offset behaviors. The preliminary results show a major improvement in the first gradient when compared to other methods of normalizing the objective function.

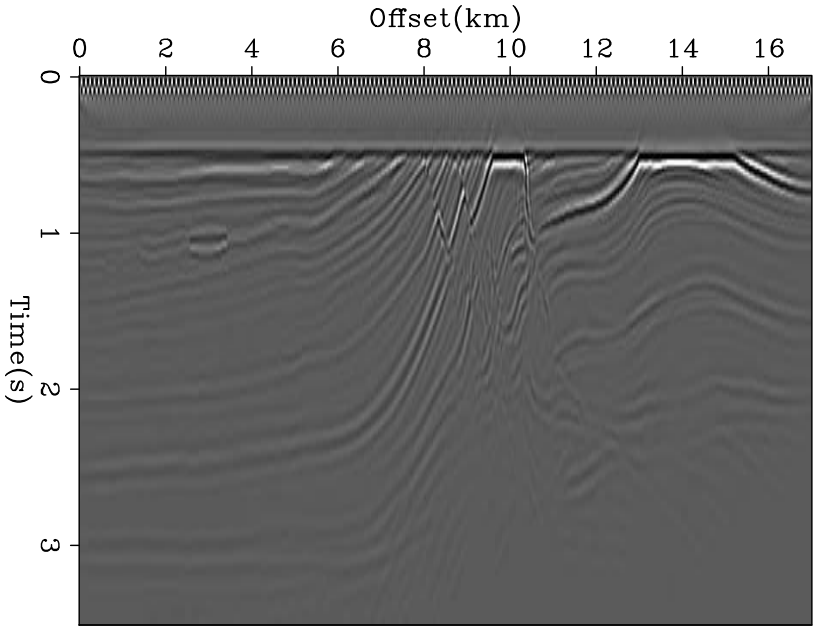


Figure 7: Model gradient using L2 residual and incorrect velocity. [**ER**] ali1/. r-1

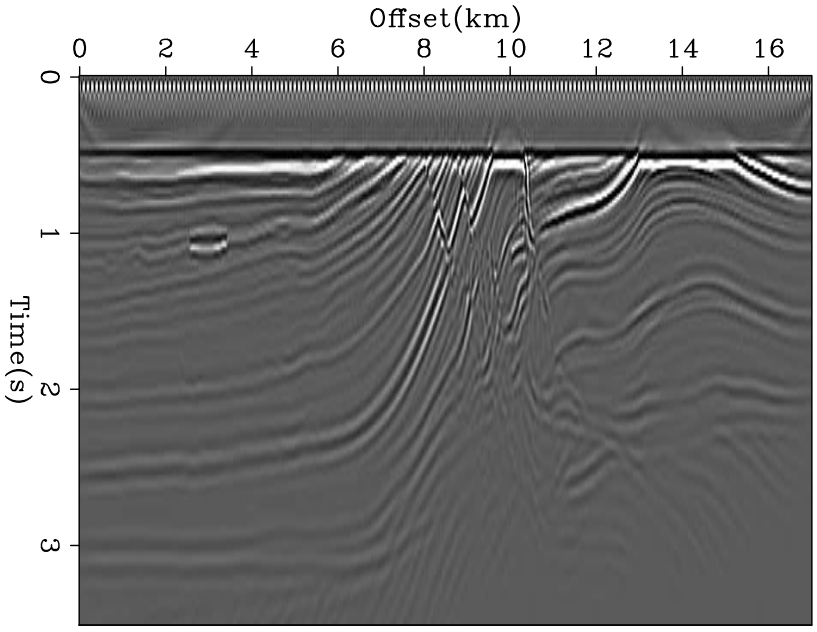


Figure 8: Model gradient using gather-normalized residual and incorrect velocity. [**ER**] ali1/. r2-1

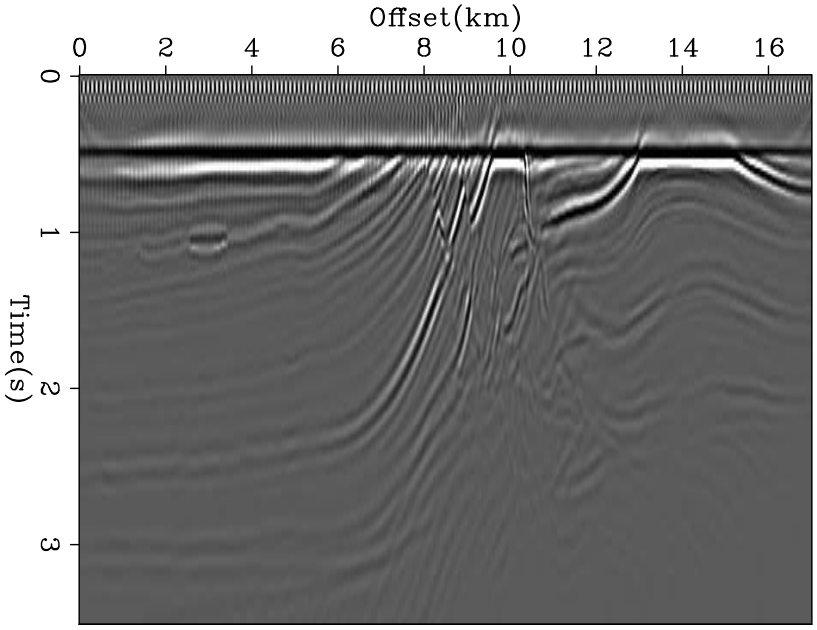


Figure 9: Model gradient using trace-normalized residual and incorrect velocity. [**ER**] ali1/. r3-1

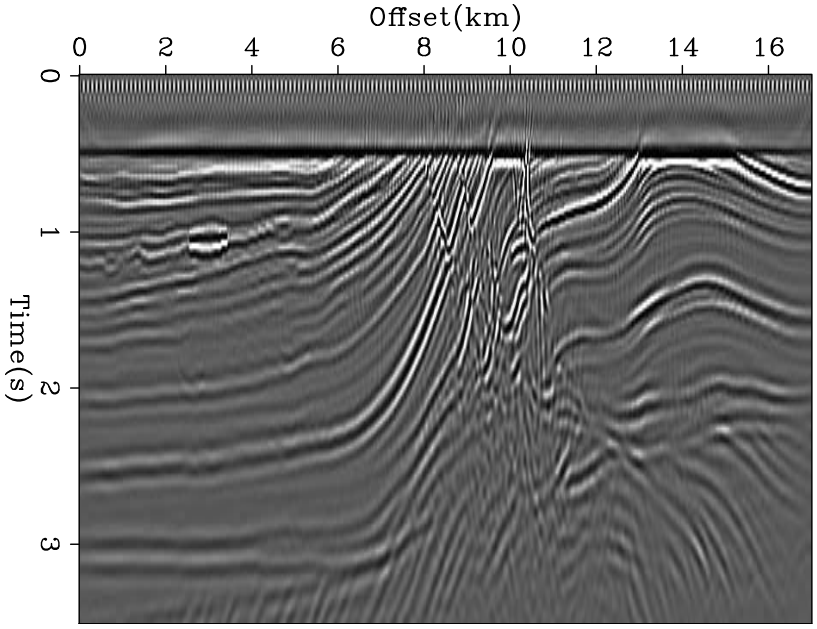


Figure 10: Model gradient using window-normalized residual and incorrect velocity. [**ER**] ali1/. r4-1

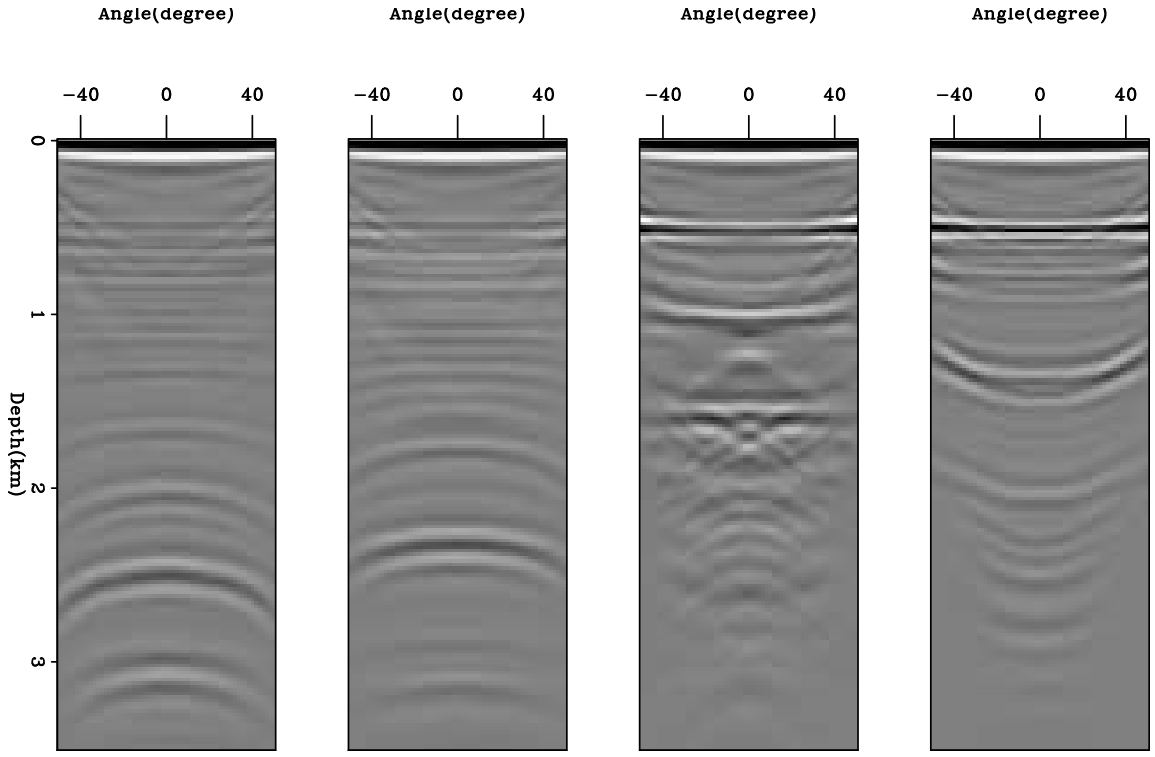


Figure 11: ADCIGs using L2 residual and incorrect velocity. [**ER**] ali1/. adcig-1

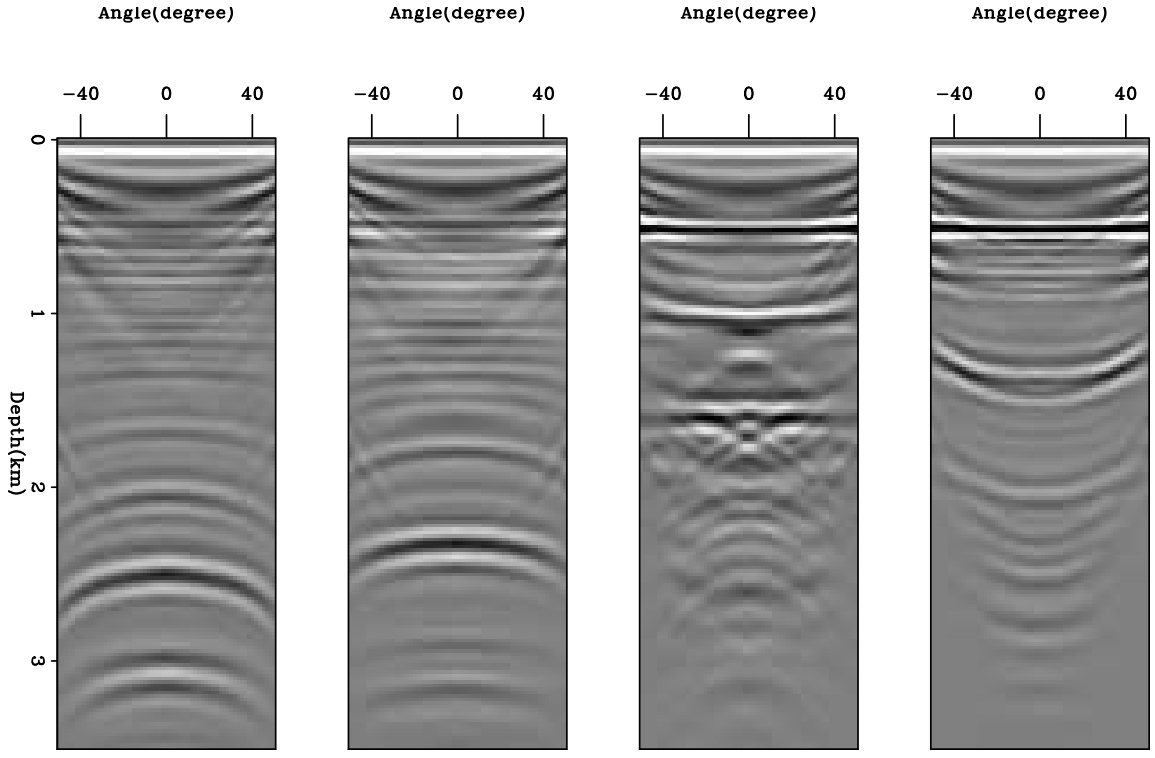


Figure 12: ADCIGs using gather-normalized residual and incorrect velocity. [**ER**] ali1/. adcig2-1

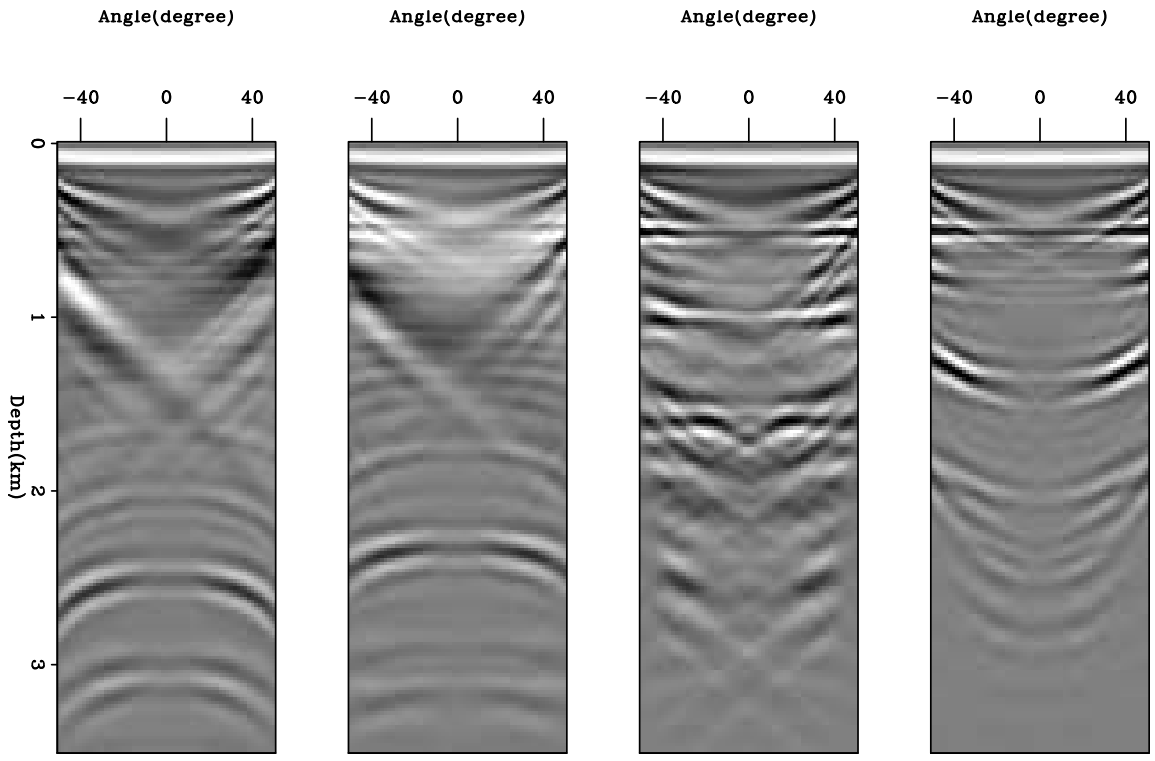


Figure 13: ADCIGs using trace-normalized residual and incorrect velocity. [**ER**] ali1/. adcig3-1

REFERENCES

Almomin, A. and B. Biondi, 2014, Preconditioned tomographic full waveform inversion by wavelength continuation: SEG Expanded Abstracts, 944–948.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain , Part 1 : Theory and verification in a physical scale model: Geophysics, **64**, 888–901.

Shen, X., 2014, Early-arrival waveform inversion for near-surface velocity estimation: PhD thesis, Stanford University.

Shin, C. and W. Ha, 2008, A comparison between the behavior of objective functions for waveform inversion in the frequency and Laplace domains: Geophysics, **73**, VE119–VE133.

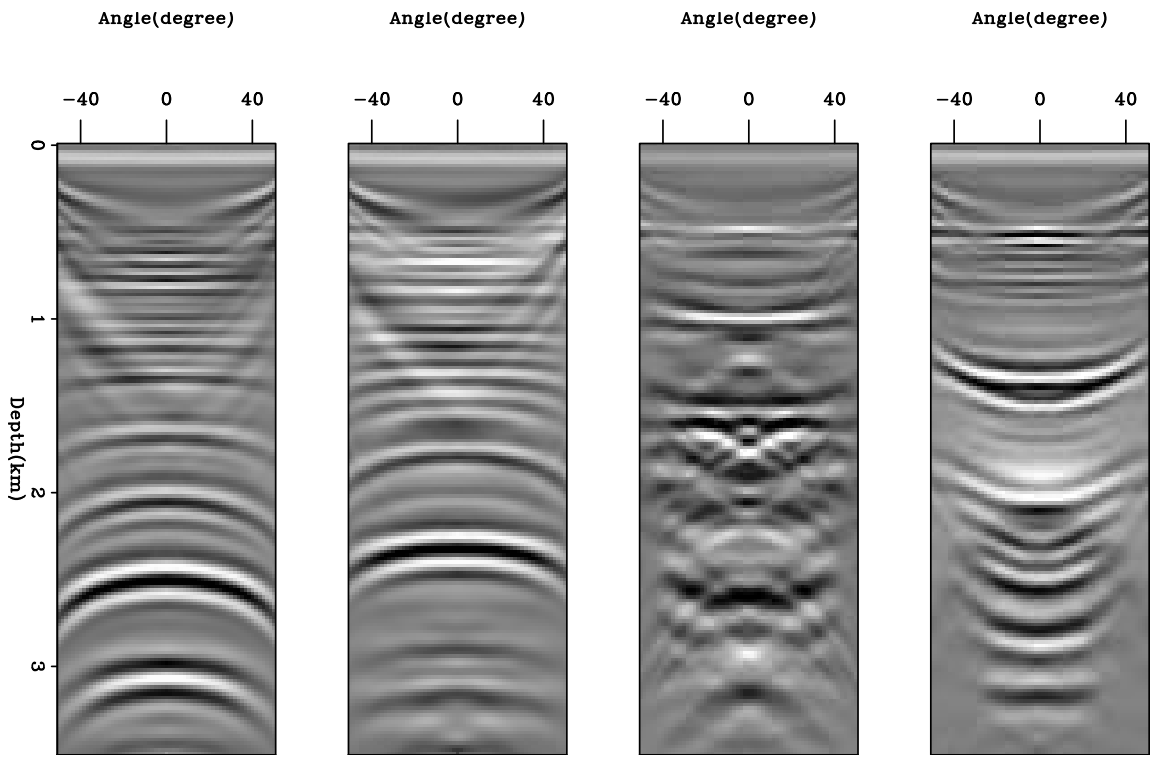


Figure 14: ADCIGs using window-normalized residual and incorrect velocity. [**ER**] ali1/. adcig4-1

Domain decomposition in shape optimization for segmenting salt bodies

Taylor Dahlke, Biondo Biondi and Robert Clapp

ABSTRACT

Level set methods can provide a sharp interpretation of the salt body by defining the boundary as an isocontour of a higher dimensional implicit representation, and then evolving that surface to minimize the Full Waveform Inversion objective function. I propose to take advantage of the benefits that the shape optimization approach has to offer. First, because the implicit surface update gradient is based on the tomographic update gradient, there is potential to utilize it to update the background velocity concurrently with the salt boundary. Second, we can decompose the update gradient into separate partitions with individual scaling parameters in order to better avoid local minima in our search and more effectively converge on the true model. Using a shape optimization approach on synthetic examples, we can achieve reasonable convergence both in terms of the residual L2 norm, as well as the evolution of the salt boundary and background velocity towards the true model, demonstrating the feasibility of this approach. Ultimately, this method can be integrated into the processing work flow as a tool that provides improved building and refining of the velocity models used for imaging.

INTRODUCTION

Oil producing regions like the Gulf of Mexico and offshore western Africa are known to have geologically complex salt body formations which can cause difficulties in producing seismic imagery. The velocity of these salt bodies often contrasts sharply with that of the background sediment layers. An inaccurate interpretation of the salt boundaries can cause significant errors in the velocity estimation process, because the formations themselves can act as lenses which focus or disperse seismic energy, influencing tomography. This can subsequently impact the imaging results that rely on accurate velocity models. Salt bodies can act as seals trapping hydrocarbons underneath, which are often the targets of imaging projects for reservior interpretation. For this reason the interpretation of salt body boundaries can also impact drilling and production activities.

Tomographic approaches to interpreting salt bodies can be less than effective, because the results tend to be too smooth to provide significantly accurate placement of the salt boundaries. Manual and semi-automatic picking of salt boundaries is a common approach to interpreting the desired sharp delineations, but these methods can be time-consuming and tedious since expert input is necessary for either the actual picking, or the oversight and correction. Furthermore, once a model has been produced, it must be used to generate an image, and then be refined as necessary. A robust method for further automating the salt interpretation procedure would greatly alleviate this bottleneck.

Some previous approaches to performing salt body segmentation use a shape optimization approach for identifying salt body boundaries (Guo and de Hoop, 2013; Lewis et al., 2012). The boundaries of a salt body can be represented as the zero-isocontour of a higher dimensional surface (for example, a 2D boundary as a contour of a 3D surface). A gradient can be derived to evolve this shape / isosurface according to the Full Waveform Inversion (FWI) objective function. Unlike the smooth boundaries produced by tomographic approaches, the isocontour resulting from the shape optimization provides a sharp boundary, which is a more appropriate way to classify most salt-sediment interfaces. Guo and de Hoop (2013) utilize this approach using a frequency domain forward wave operator to evolve a salt boundary and velocity model. Their approach creates and applies a global gradient update, which can create problems updating the base-of-salt (BOS) once the top-of-salt (TOS) has gotten close to convergence. This issue has been observed in recent work (Guo and de Hoop, 2013), and is an inherent problem with the global gradient update approach.

To address the problem with global updating, we introduce domain decomposition for the salt boundary update gradient by splitting it; demonstrated here using "upper" and "lower" partitions. This introduces an additional degree of freedom to our parameter searching so that local minima can be avoided. Using a shape optimization approach with time domain forward wave-propagation, we take advantage of using a continuous range of frequencies (rather than discrete frequencies) in each iteration, which allows for sharper delineation of the boundary. Further, we take advantage of the fact that our boundary update gradient is based on the tomographic update gradient, and make updates to both concurrently with the use of scaling parameters. In theory, this algorithm has the potential to be more efficient than an alternating update approach.

In this paper we will begin by discussing the fundamentals of the level set method and its key properties, followed by the derivation of the boundary update gradient. Next we will describe and demonstrate the general algorithm, and discuss the assumptions and fundamental limitations of this approach. Following, we will introduce the domain decomposition method, and demonstrate the improved results that this approach offers.

THEORY

While it may seem counter-intuitive to add an extra dimension to our problem, by doing so, we gain the advantage of easily merging/separating bodies as the evolution proceeds, as well as the ability to handle sharp corners and cusps in the lower-dimensional (2D) plane on which the boundary exists.

Osher and Sethian (1988) and Burger (2003) describe the level set of ϕ that represents the salt body boundary as

$$\phi(x_\Gamma, \tau) = 0,$$

where x_τ is the spatial domain, and τ is the iteration count. By taking the derivative of this equation with respect to τ (to find the $\delta\phi$ between iterations), applying the chain rule, and re-arranging terms we can get:

$$\frac{\partial \phi}{\partial \tau} = -V(x_\Gamma, \tau)\,|\nabla \phi|\,. \tag{1}$$

The scalar speed term $V(x_\Gamma, \tau)$ describes the magnitude of the variation of ϕ that is normal to the boundary Γ. It determines the evolution of the implicit surface, and ultimately the boundary implied by it. We derive this normal velocity such that the FWI objective function is minimized

$$\min \|F(m) - d\|_2^2\,, \tag{2}$$

where $F(\cdot)$ is the forward wavefield modeling operator, m is the velocity model, and d is the observed data.

Calculus of variations

The shape derivative we use is based on a formal calculus of variations outlined in Santosa (1996). The objective is to define the variation of the model m with respect to the boundary variation (represented implicitly by the surface, ϕ).

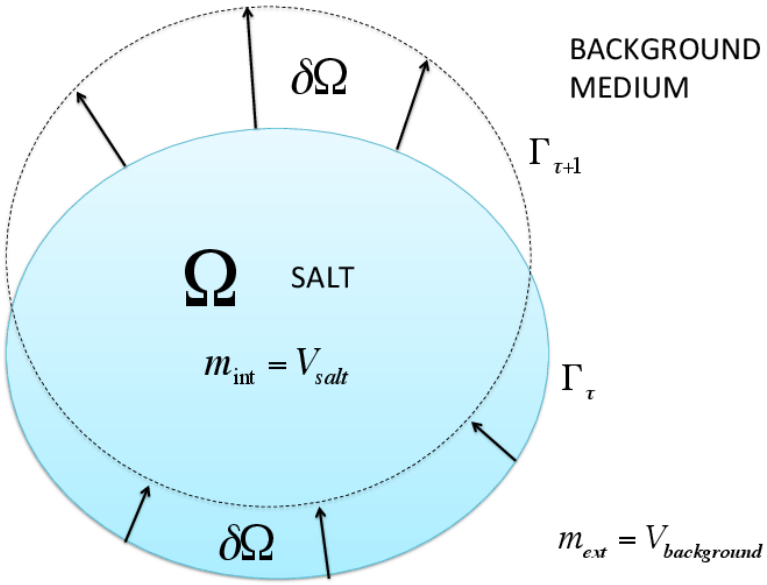


Figure 1: The geometry of the curve $\{x_\Gamma : \phi = 0\}$ for a variation $\delta\phi(x)$ for an evolution step τ. [**NR**] taylor1/. domain-fig

We begin by considering an inner product of velocity model perturbation δm with a test function $f(x)$. Formally, this can be written as,

$$\langle \delta m, f(x) \rangle = \int_{\mathbb{R}^2} \delta m(x) f(x) dx = \int_{\partial\Omega} \delta m(x) f(x) dx. \tag{3}$$

Because the $\delta m(x)$ term equals zero in $\mathbb{R}^2 \setminus \partial\Omega$, it does not contribute to the overall inner product when integrating over that domain; therefore, we only integrate over $\partial\Omega$ where $\delta m(x)$ is non-zero. We know that $\delta m(x)$ will be $\pm(m_{\text{int}} - m_{\text{ext}})$, depending on the relative values of m_{int} and m_{ext} or the direction of the normal vector $\vec{n}$. We only care about the component of δx_Γ that occurs in the normal direction, because a tangential variation of x_Γ does not affect m or ϕ. Furthermore, because δx_Γ is infinitesimal, we can replace dx with $\delta\vec{x}_\Gamma \cdot \vec{n}$ and simplify Equation 3 into

$$\langle \delta m, f(x) \rangle = \int_{\partial\Omega} (m_{\text{int}} - m_{\text{ext}}) \delta\vec{x}_\Gamma \cdot \vec{n} f(x) ds(x), \tag{4}$$

where $ds(x)$ is the incremental arc length along the boundary Γ. We can think of $\delta\vec{x}_\Gamma \cdot \vec{n} ds(x)$ as roughly the incremental area over which m varies at x.

We can identify δm from Equation 4. It can be considered a measure over $\partial\Omega$:

$$\delta m = (m_{\text{int}} - m_{\text{ext}}) \delta\vec{x}_\Gamma \cdot \vec{n} \,|_{x \in \partial\Omega}\,. \tag{5}$$

We remember that in the previous section we stated the goal of this derivation as being a solution of the scalar velocity function $V(x_\Gamma, \tau)$, such that the objective function is minimized. We recognize that the normal component of the variation δx_Γ satisfies:

$$\delta\vec{x}_\Gamma \cdot \vec{n} = V(x_\Gamma, \tau). \tag{6}$$

We can use the shape derivative formulation described in Santosa (1996) to find a $V(x_\Gamma, \tau)$ that minimizes the FWI objective function (Equation 2) that we insert into Equation 1 to get a final implicit surface update gradient of:

$$\frac{\partial \phi}{\partial \tau} = (m_{\text{int}} - m_{\text{ext}}) \frac{\partial F}{\partial m} \left| \vec{\nabla} \phi \right|. \tag{7}$$

The adjoint state method as described in Plessix (2006) is used to derive $\frac{\partial F}{\partial m}$, which can be shown to be equivalent to $F(m)^T(A(m) - d)$. Because our case uses the FWI objective function (Equation 2), this term can be interpreted as least squares migration, more specifically as reverse time migration (RTM). This result is formulated as:

$$\frac{\partial F(m)}{\partial m} = -\sum_s \int_0^T \int_{x \in \Gamma} h_s(x,t) \frac{\partial^2 u_s(x,t)}{\partial t^2} d\sigma dt, \tag{8}$$

where h_s is the backpropagated residual wavefield, and u_s is the source wavefield. This term by itself is the velocity model perturbation, which we can process and use to make tomographic updates as described in the following section.

General evolution algorithm

We begin with an initial background velocity, and a binary valued (-1, 1) function as the initial implicit surface ϕ. Since we assume a constant salt velocity, we use both of these inputs to create a full initial-guess velocity model m_o. Using this m_o, we forward model to get our d_{syn} and subsequently find our residual. The residual is used to calculate both a tomographic and a boundary update gradient, as described in the derivation section previously. We then do non-linear line searches for α and β in a manner that minimizes the FWI objective function. Afterwards we apply an explicit forward Euler scheme that updates the implicit surface (ϕ) and the background velocity V_{back}:

$$\phi^{j+1} = \phi^j + (\beta \frac{\partial \phi}{\partial j} + \mu G_{\text{reg}}) \tag{9}$$

$$V_{\text{back}}^{j+1} = V_{\text{back}}^j + \alpha \frac{\partial V_{\text{back}}}{\partial j}. \tag{10}$$

where β and α are the step sizes, j is the current iteration point, and μG_{reg} is the weighted regularization term (described subsequently). This work flow is graphically represented in Figure 2.

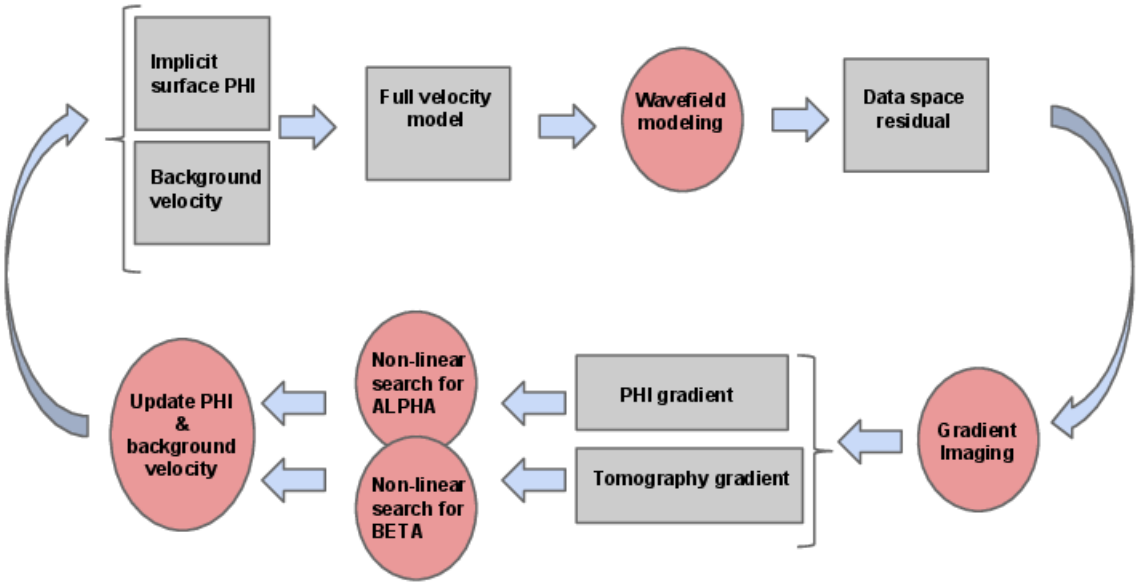


Figure 2: The general work flow used for shape optimization. [**NR**] taylor1/. workflowNL

Implicit surface stability: CFL condition

As the implicit surface is evolved, it is important to maintain stability of the evolution. One relevant aspect of maintaining stability is keeping the implicit surface update step size (β)

small enough to satisfy the Courant-Friedrich's-Levy (CFL) condition, which is stated by Chaudhury and Ramakrishnan (2007) (when applied to level set evolution) as being:

$$G_{\max} \cdot \beta \leq \min(h_x, h_y) \tag{11}$$

where h_x and h_y are the grid spacing in the x and y directions, and $G_{\max}$ is the maximum value of the update gradient.

Implicit surface stability: DRLSE regularization

It is well documented in level set literature that maintaining a smooth, differentiable implicit surface is important to maintaining stability as the surface evolves. Multiple approaches have been implemented towards this end, including freezing of the level set boundary and reinitializing the implicit surface to that of a signed distance function. This approach can be somewhat effective, but has the drawback of adding expense to the algorithm (since the boundary does not evolve during the iterations that reinitialization is occurring). Further, there is no clear approach for identifying how frequent the reinitialization should occur to maintain stability. For this reason, the more effective approach of distance-regularized level set evolution (DRLSE) introduced by Li et al. (2010), is more commonly implemented in recent work. Unlike the reinitialization scheme, Li et al. (2010) show how DRLSE applies regularization updates at every iteration to maintain the implicit surface gradient as equaling one ($|\nabla\phi| = 1$). We calculate this DRLSE regularization term, which is then scaled and added to the boundary gradient before each update is applied. By driving the gradient of the implicit surface to equal one, we minimize irregularities and are able to continue evolution without having to reinitialize a signed-distance function to the salt boundary contour.

Scaling parameter optimization

As shown previously, the salt boundary gradient is based on the adjoint of the linearized-Born operator, which is the tomographic update gradient. Since the gradient for both a tomographic and boundary update are calculated in each step regardless, we attempt to take advantage of this by finding scaling parameters to apply to these gradient updates such that we minimize the residual space objective function 12. We use non-linear line searches to find α, then update the tomography, and immediately after perform a non-linear line search for β using the updated background velocity.

$$\min \|F(m(\gamma) - d_{\text{obs}})\| . \tag{12}$$

Where γ is a placeholder for either α or β.

APPLICATION

My demonstration of the shape optimization algorithm in both cases was performed on a 2D model, with the implicit surface evolved being a 3D surface. For the forward wave propagation, a wavelet with a 15.0 Hz central frequency was propagated using a time domain

forward operator, so that continuous frequency information would be forward modeled in a single iteration.

Example: General algorithm

To implement the general algorithm, I used a model and acquisition that was symmetric on the horizontal (x) axis, with a salt model that is smaller than the true salt body. The background velocity had several layers, where the gradient in each layer was perturbed to be faster than actual, while preserving the reflector position (See Figure 3). The acquisition was composed of 11 shots and 75 receivers. The algorithm does a reasonably good job at converging on the top of the salt body, but the updating along the bottom and sides has still not been able to converge to the true model correctly.

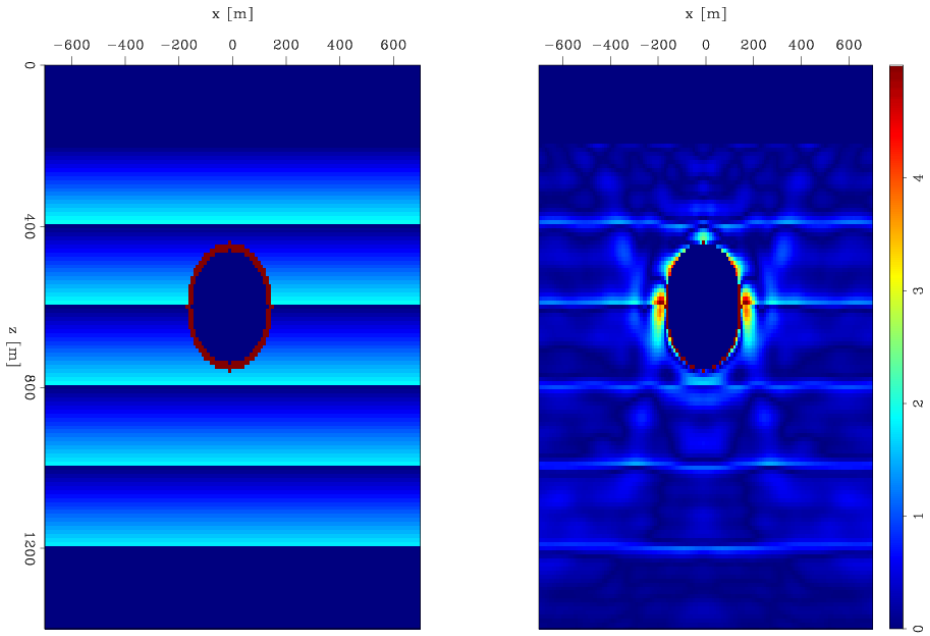


Figure 3: Percent velocity error for initial model (left), and for the model after 100 iterations (right) using general algorithm. [**CR**] taylor1/. dualNL4

domain decomposition of boundary update gradient

One notices in Figure 3 that the convergence of the general algorithm appears to stall before convergence of the base of salt occurs, even after a significant number of iterations. The reason for this is that the more strongly illuminated top of salt (TOS) is the primary influence in the line search for β, used to apply the implicit surface update gradient. For this reason, even when the gradient update would push the base of salt (BOS) towards convergence, the updating on the TOS (which is unnecessary since the TOS has converged) has a saturating effect on the step length search, forcing the update towards a very small value.

The implicit surface update would be successful if the gradient did not show an update

for the TOS after the TOS has converged. However, even when the TOS boundary and background velocity model are perfect, we will still realize an update for the TOS (see Figure 4).

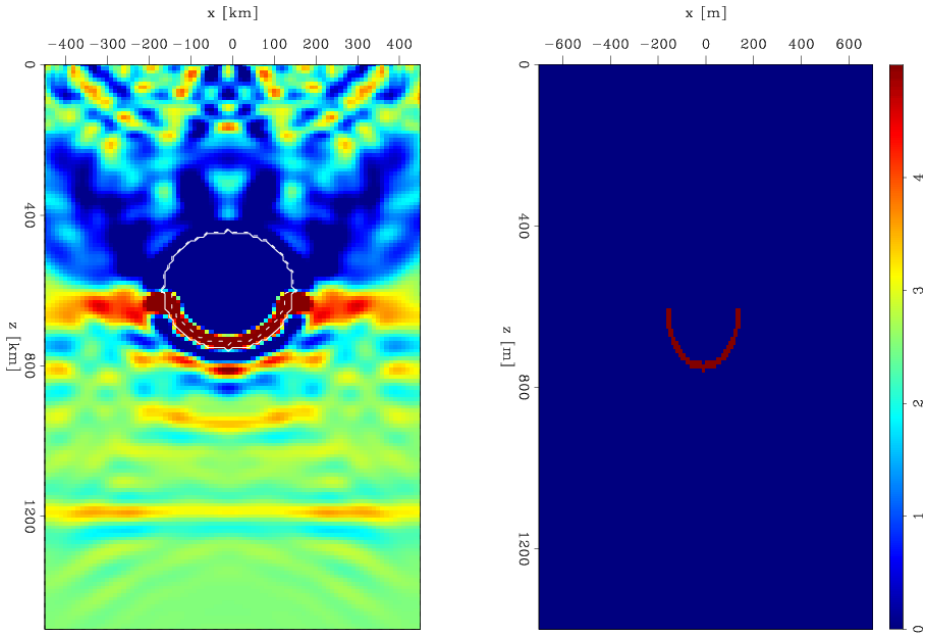


Figure 4: First gradient of FWI (left) for model with perfect TOS and background velocity. Velocity model perturbation (right). The FWI gradient wants to correct the residual by both increasing the velocity at the base of salt (red), while also reducing the velocity of the medium above it (blue). This RTM velocity updating maps to an implicit surface update that causes the TOS boundary to sink deeper in order to reach the same effect. [**CR**] taylor1/. perfectTOS-grad

The fundamental problem demonstrated in Figure 4 is that the reverse time migration (RTM) gradient contains information from both reflector position error as well as background velocity error. These can be difficult or impossible to separate. In this case, what could be correctly updated by changing the position of the bottom reflector can also be corrected by making a velocity update above it. The RTM calculation always produces a gradient that applies both these effects. For this reason, while an increase/decrease of velocity from the raw RTM gradient is often correctly mapped to a reflector position change, the algorithm ultimately stalls at a local minima when illumination for a reflector (like the BOS) is sourced from ray paths that must first travel through another reflector edge of the same body being updated (like the TOS). Fundamentally, the update gradient is flawed in terms of directing convergence to the global solution of our objective function. The 'false' TOS updates consistently lead us to a local minima instead. This is because the TOS updates are always much stronger in the residual space, and thus our non-linear line search for a global scaling parameter will always prefer to correct the TOS, typically with a very small update step (when the TOS is near the true model). This causes a lack of update for the base of salt reflector. Figure 5 shows that even after a significant number of iterations, this approach cannot converge to the true model, and in fact, even becomes worse for some sections of the boundary.

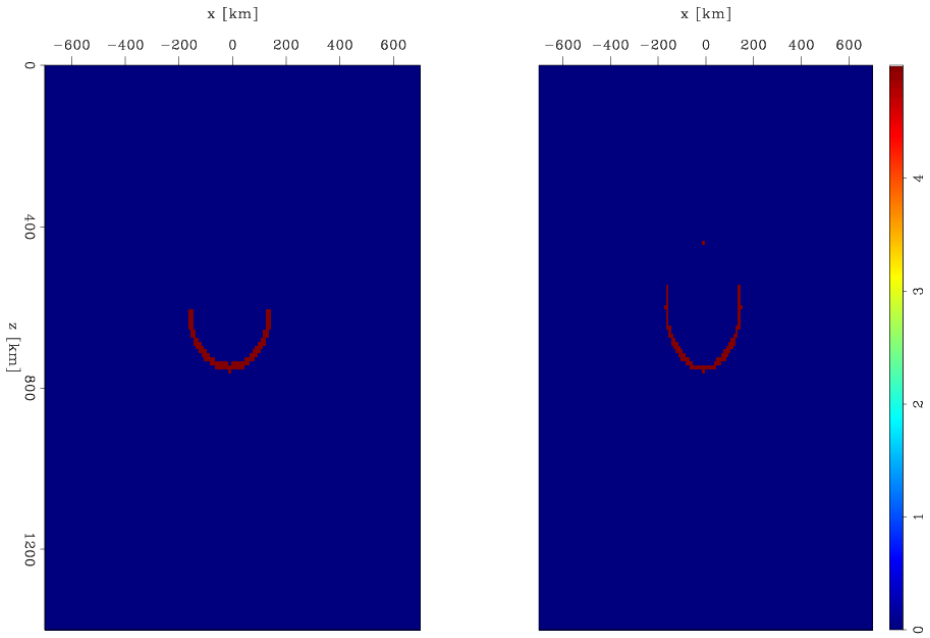


Figure 5: Velocity model error for starting model (left), and final model (right) after 200 iterations using general algorithm. [**CR**] taylor1/. perfectTOS-long200

Examples: Partitioning approach

Any gradient that updates both the entry and exit reflectors for a given ray path will eventually meet the limitation described above, because the gradient couples both updates together. A method to de-couple the entry and exit reflectors is necessary (typically top and bottom when the acquisition is above the salt, as in most industry seismic) so that reasonable convergence can still be reached when spurious updating occurs in the upper reflector due to the mixing of tomography and reflector position.

In order to decompose the ϕ update gradient into 'top' and 'bottom' components for the general case, we use an approach that makes use of straight rays by creating a mapping of the radial distance from each shot position. We take the gradient of this map in order to find the x and y gradient components of the radial vector field. We then find the x and y components of the normal vectors from the salt boundary. By taking the dot product of these straight ray and boundary normal vector fields, we generate a map that approximates the straight ray illumination of the body. However, more expensive techniques to produce ray path maps could be substituted to achieve better 'illumination' mapping. We perform these steps for each shot, and sum the dot-product fields to create a final cumulative map that is used for weighting on the ϕ gradient. Once this is done, thresholding is applied so that negative values in the weighting map are set to zero, while the positive values are set to 1.0. The inverse weighting map can be found by multiplying the original map by -1.0, and then thresholding as before. Figure 6 gives an example of what this boundary partitioning looks like.

This 'splitting' has the advantage of decomposing the boundary into sections that gen-

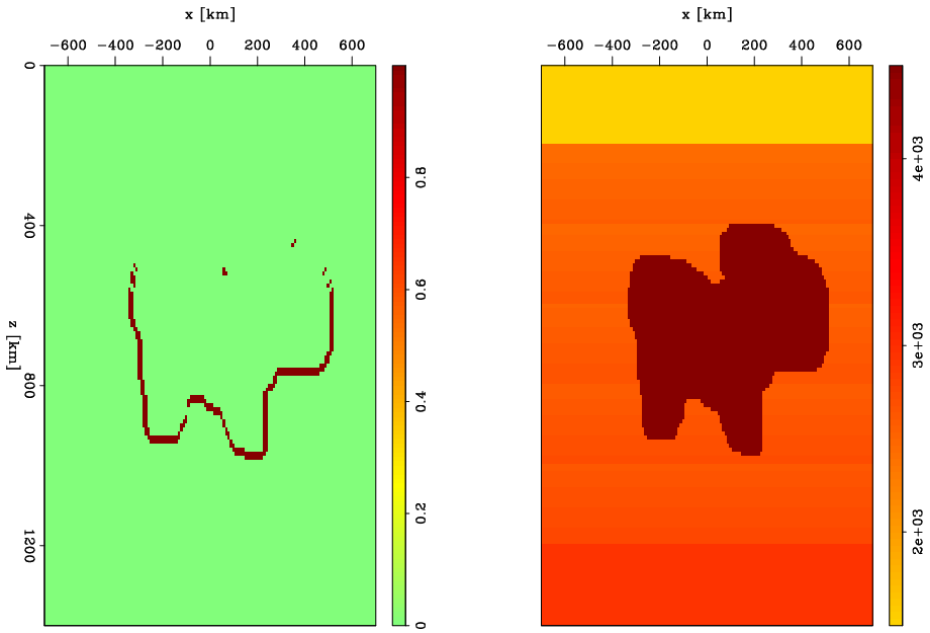


Figure 6: Example of weighting used for the 'bottom' ϕ gradient partition at first iteration (left), and velocity model that weighting is based on (right). [**CR**] taylor1/. smooth-split

erally face away from or towards the acquisition line. Further partitioning of the boundary update can be done by setting thresholding to selectively partition the facing, steep, and shadowed reflecting edges.

Examples

Isolating from the effects that tomography updating might have, we perform an experiment where the algorithm does no tomography updating, and begins with a perfect background velocity model. Figure 7 shows that the difference between the two approaches is clear. After 25 iterations, the partitioning algorithm has already converged on the true model, while the global gradient approach has trouble, particularly on the base of salt.

We test again with the same salt model, this time beginning with a perturbed background velocity model. We also have the algorithm perform tomography updates. Figure 8 shows that when we try this approach we get close convergence with the true model on both the top and base of salt.

We can compare the results of the algorithms shown in Figures 8 and 3 by plotting the differential error, as shown in Figure 9. From this we observe that while some of the tomographic updates are less accurate in the partitioning approach, the convergence on the salt model is significantly improved, especially for the base of salt and salt flanks. This further demonstrates that this domain decomposition approach can yield more accurate convergence, even when tomography is also concurrently being updated.

We next try the same comparison of the two algorithms on models with far more complex salt body geometry. Figure 10 also shows some slight improvements on the bottom of salt

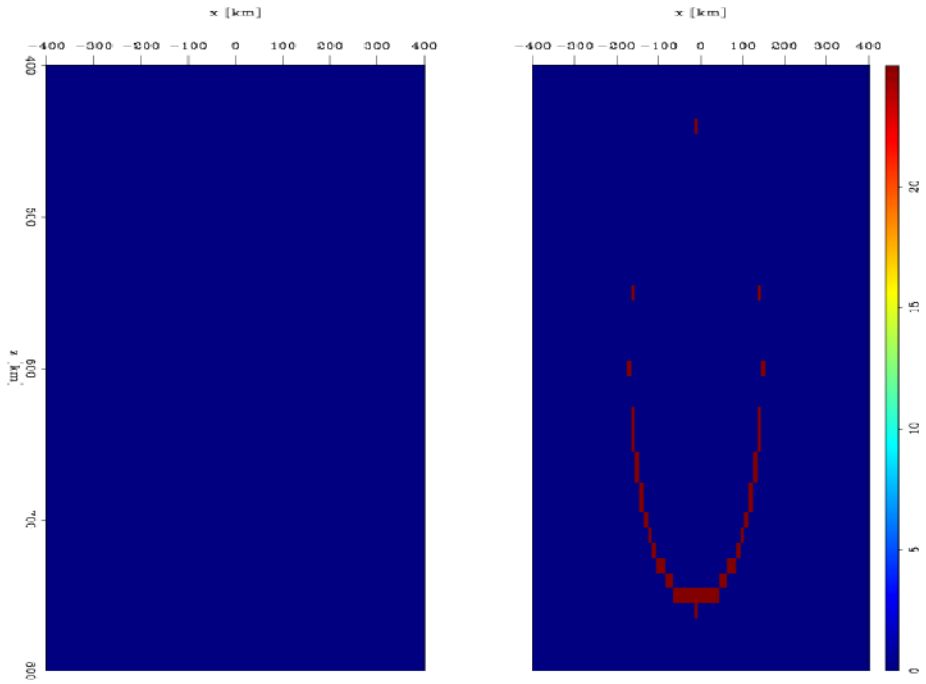


Figure 7: % velocity error at 25 iterations using domain decomposition algorithm (left) and general algorithm (right). Perfect background velocity model. [**CR**] taylor1/. compare-split-nosplit25

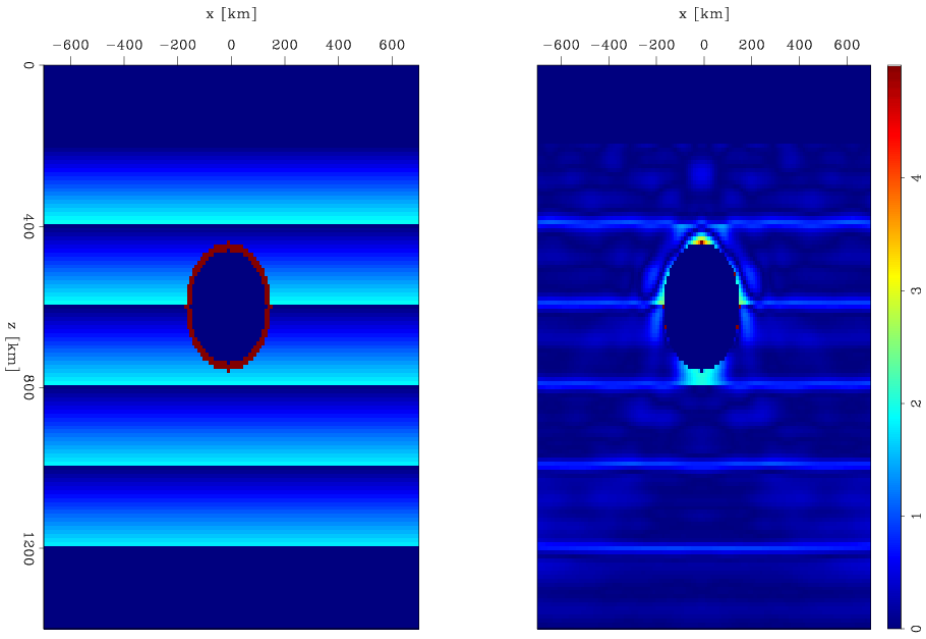


Figure 8: Initial model perturbation (left), and after 100 iterations (right) using split ϕ algorithm. [**CR**] taylor1/. dualNL3

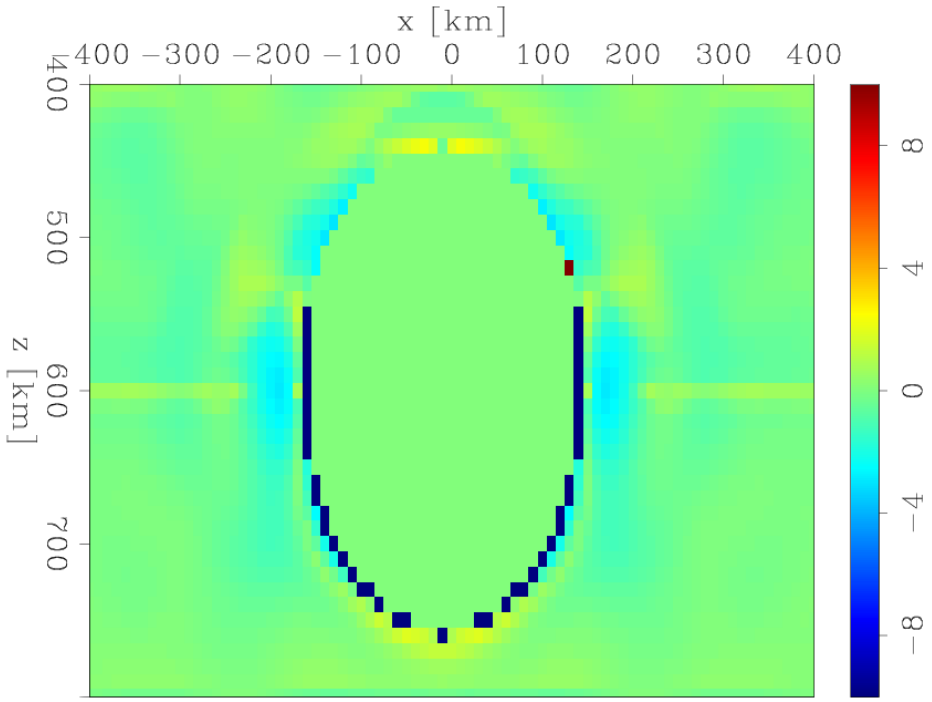


Figure 9: The differential error between the gradient partitioning approach and the general algorithms after 100 iterations, as expressed as a percentage error from the true model. Blue regions (negative differential) indicate areas where the domain partitioning approach performs better. Red (positive differential) indicates where it performs more poorly. [**CR**] taylor1/. method-compare100

and even some of the tomography updates along the top of salt.

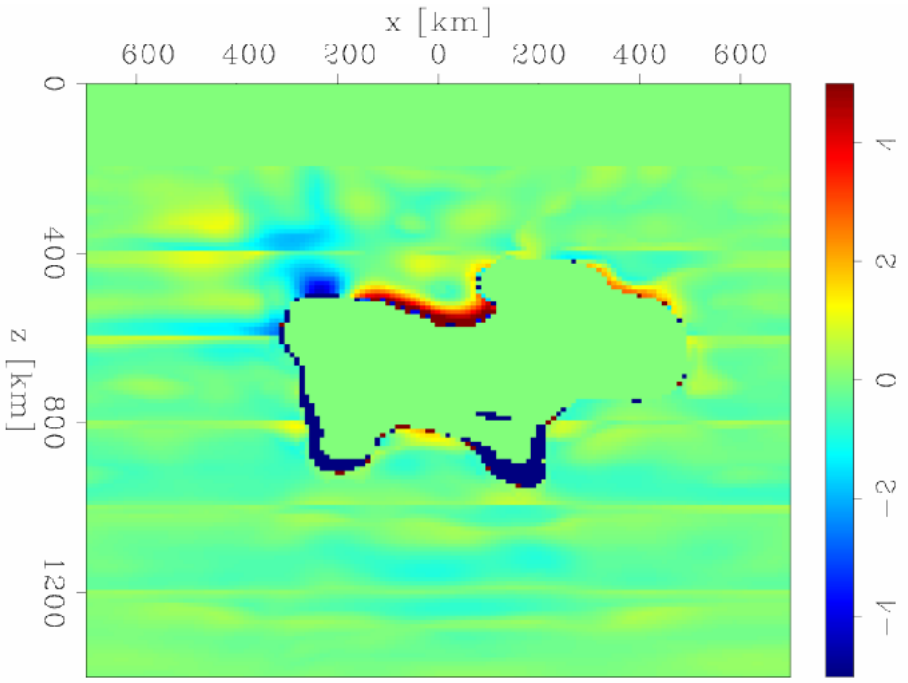


Figure 10: Velocity model difference between general and split ϕ methods, shown as a percentage error from the true model. Blue regions are where the domain partitioning approach performs better; red, more poorly. [**CR**] taylor1/. potato-vel-diff175

Figure 12 demonstrates improved convergence on the lower flanks of the salt body, as well as an improvement on the tomography above the TOS. Figure 11 shows that the gradient partitioning algorithm performs better in the underside areas of the left hand salt body, as well as on the flanks and underside of the right hand body. While the very bottom of the right hand side body has not converged as well as the other regions, further partitioning of the model (into top, sides, bottom) would allow for better updating on this region. A simple two part split may be insufficient since there are multiple entry/exit points on the ray paths that illuminate this lowest region of the body, which exasperates the problem of coupling this reflector with other reflector updates. This example demonstrates the potential of the domain decomposition algorithm to determine salt boundaries in more realistic models.

CONCLUSIONS

RTM imaging is fundamentally unable to differentiate between reflector position errors and tomography errors. By splitting the spatial domain of the implicit surface update, and performing separate non-linear parameter searches for each, we are able to mitigate occurence of local minima due to stronger TOS updates overwhelming the line search. The results from using this approach give us much better convergence on the base and flanks of the salt body. Further partitioning of the domain could improve convergence even more for complex salt models.

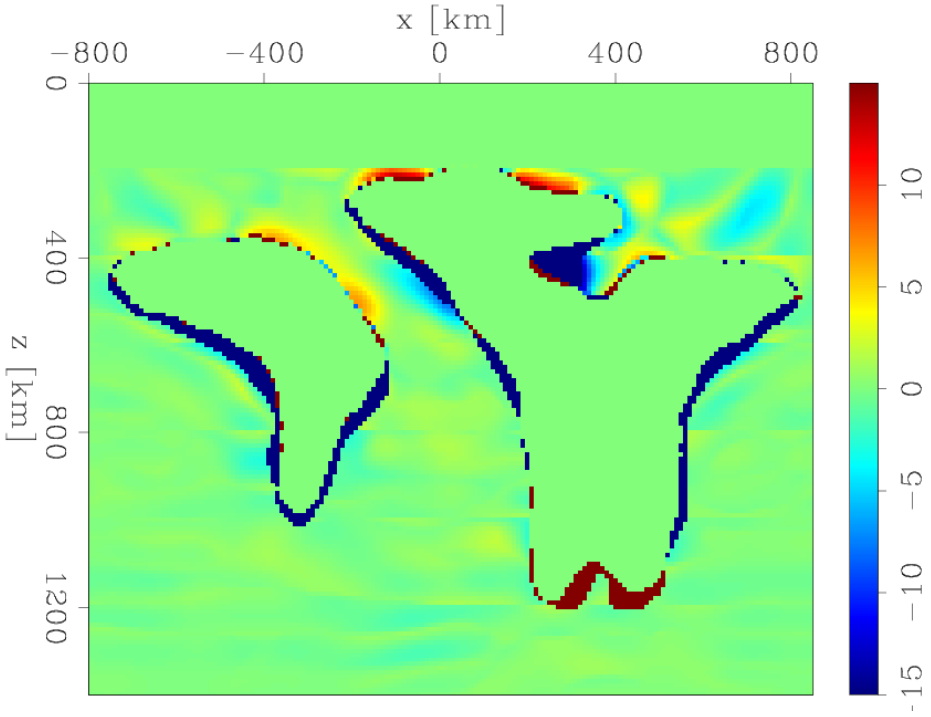


Figure 11: Velocity model difference between general and split ϕ methods, shown as a percentage error from the true model. Blue regions are where the domain partitioning approach performs better; red, more poorly. [**CR**] taylor1/. potato2-vel-diff100

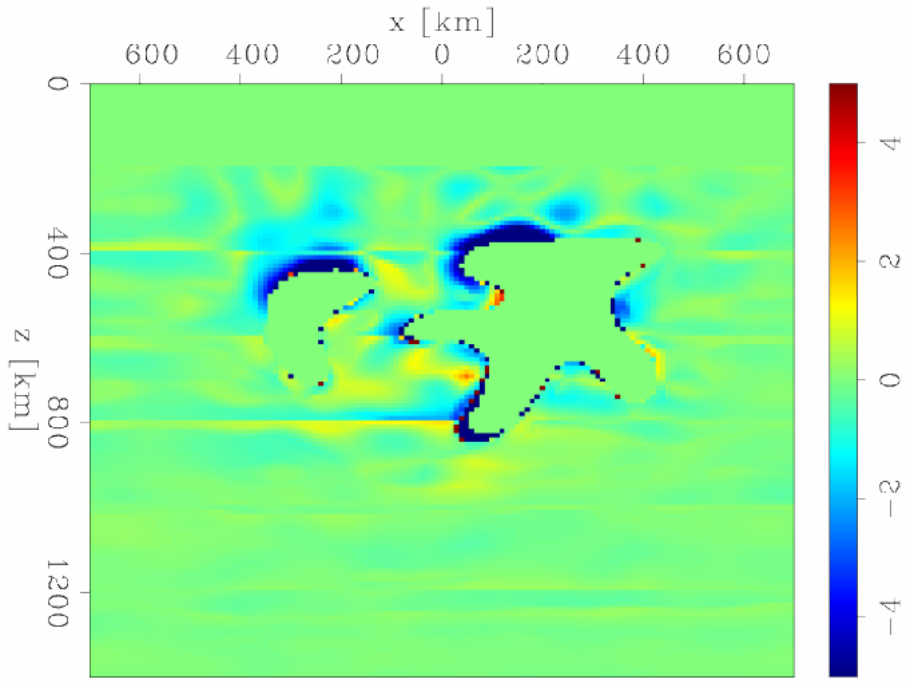


Figure 12: Velocity model difference between general and split ϕ methods, shown as a percentage error from the true model. Blue regions are where the domain partitioning approach performs better; red, more poorly. [**CR**] taylor1/. potato3-vel-diff175

REFERENCES

Burger, M., 2003, A framework for the construction of level set methods for shape optimization and reconstruction: Interfaces and Free boundaries, **5**, 301–330.

Chaudhury, K. N. and K. R. Ramakrishnan, 2007, Stability and convergence of the level set method in computer vision: Pattern Recognition Letters, **28**, 884–893.

Guo, Z. and M. de Hoop, 2013, Shape optimization and level set method in full waveform inversion with 3d body reconstruction: SEG Technical Program Expanded Abstracts 2013, 1079–1083.

Lewis, W., B. Starr, and D. Vigh, 2012, A level set approach to salt geometry inversion in full-waveform inversion: SEG Technical Program Expanded Abstracts 2012, 1–5.

Li, C., C. Xu, C. Gui, and M. Fox, 2010, Distance regularized level set evolution and its application to image segmentation: Image Processing, IEEE Transactions on Image Processing, **19**, 3243–3254.

Osher, S. and J. A. Sethian, 1988, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations: Journal of computational physics, **79**, 12–49.

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

Santosa, F., 1996, A level-set approach for inverse problems involving obstacles: ESAIM Control, Optimization, and Calculus of Variations, **1**, 17–33.

Addressing the effects of inaccurate top-salt delineation on subsalt seismic imaging

Guillaume Barnier and Biondo Biondi

ABSTRACT

Imaging in the presence of salt bodies is important for the oil and gas industry but, it is also very difficult. Current seismic imaging techniques may still fail to capture the fine-scale details of a rugose boundary between the top of a salt body and its overlying sediments. This lack of accuracy can have a destructive effect on images of potentially hydrocarbon-bearing layers located underneath the salt. In this paper, we give an overview of the main challenges encountered by the oil and gas industry in terms of subsalt imaging, and we explain the importance of addressing the problem of top-salt boundary delineation. We conduct synthetic tests, and show that a slight misinterpretation of a top-salt boundary significantly damages the image quality of underlying layers. Image gathers may become incoherent and uninterpretable, which prevents us from using standard migration velocity analysis techniques to improve imaging. There is a need to develop a technique that enables us to refocus the subsalt images. We propose an approach for future research.

INTRODUCTION

Oil and gas companies are expanding their search for hydrocarbons into zones with increasing geological complexity, such as sedimentary basins containing allochthonous salt bodies. Unfortunately, seismic imaging around salt is difficult. One of the main challenges is the lack of accuracy in the subsurface velocity model. More specifically, a problem that has recently drawn a great deal of attention is the limitation of currently existing techniques to accurately delineate the interface between the top of a salt body and its overlying sediments, and the impact is has on the image-quality of deeper layers (Halpert, 2014; Etgen et al., 2014a,b; Albertin et al., 2014). Because this interface is usually very rugose and may take complex geometrical shapes, it is difficult to resolve its very fine-scale features. The velocity contrast between the salt body and its surrounding sediments is so high that the quality of the seismic image beneath this interface is highly sensitive to the accuracy of the boundary location. Even a slight error in the interpreted location of this interface can have a disproportionate impact on the calculations of wave propagation beneath the interface, and lead to significant degradation of the resulting image (Halpert, 2014). In the first part of this paper, we explain why salt basins are relevant and attractive in hydrocarbon exploration, we describe where the main challenges lie, and we discuss the importance of getting more accurate and well-focused subsalt images. In the second part, we analyze into more detail the difficulty in precisely delineating the top-salt boundary, and we explain why present techniques may fail to do so. In the third and fourth part, we present synthetic examples that illustrate the effects of inaccurate top-salt delineation on the image-quality of subsalt

layers, and the need for a new algorithm to address this issue. In the last part, we discuss the possible approaches we intend to take to solve this problem.

SALT BASINS IN HYDROCARBON EXPLORATION

Salt basins are attractive zones for oil and gas exploration, and remain among the best places to find hydrocarbon reserves worldwide. The best-known salt basins are located in the Gulf of Mexico, the Persian Gulf, the North Sea, the West African margin, and the Brazilian margin. However salt basins, especially tertiary basins with allochthonous salts, are difficult places to explore because of poor subsurface images obtained around and below salt (Leveille et al., 2011). In this section, we propose a literature review of the main geological and geophysical properties of salt basins, and we summarize the challenges the industry currently faces when imaging in the presence of salt bodies.

Geological properties of salt bodies

Salt is an evaporite, and salt beds are formed by the natural evaporation of the sea water from an enclosed basin. The precipitated salt layer is then buried by successive layers of sediments over geologic time until segments of it begin to flow upward toward the surface of the Earth, thereby creating a salt dome (DNR, 2015). A salt dome is a mound or column of halite that has been pushed upward from below through the surrounding rocks and sediments into its present position. Rock salt (i.e., halite) has two properties that enable it to form salt domes.

- When it is buried to depths greater than a few 100 meters, salt will have a density of approximately 2.16 g/cm^3, which is much lower than most other sedimentary rocks (e.g., shales, limestones, etc.). These sedimentary rocks tend to have lower densities when they are deposited because they contain a lot of water. As the depth of burial increases, the density of salt remains about the same, but the mass density of shales and limestones increases as the water is squeezed from their pore space. Eventually, their mass density might reach values ranging from 2.4 g/cm^3 to 2.7 g/cm^3, which is much higher than the salt's mass density (King, 2015).
- Salt also has the ability to deform and flow like a high-viscosity fluid when it is under pressure.

The overlying sediments tend to exert a compressive, downward force; and if the rock sequence is also subjected to tectonic forces, salt may begin to flow upward like a plastic substance. A small fracture in the overlying, higher-density sediments can trigger the upward movement. While compression forces will produce folding, and salt domes might erupt through the crest of anticlines (Figure 1); extension forces will produce thinning and normal faulting, which might create weaknesses that the salt might exploit (Figure 2)(Hudec and Jackson, 2007).

Once this upward salt movement begins, salt from elsewhere in the salt bed moves into the region surrounding the salt plug to replace the salt that is flowing upward. The upward movement of the salt plug (or salt dome) continues as long as there is a sufficient source of

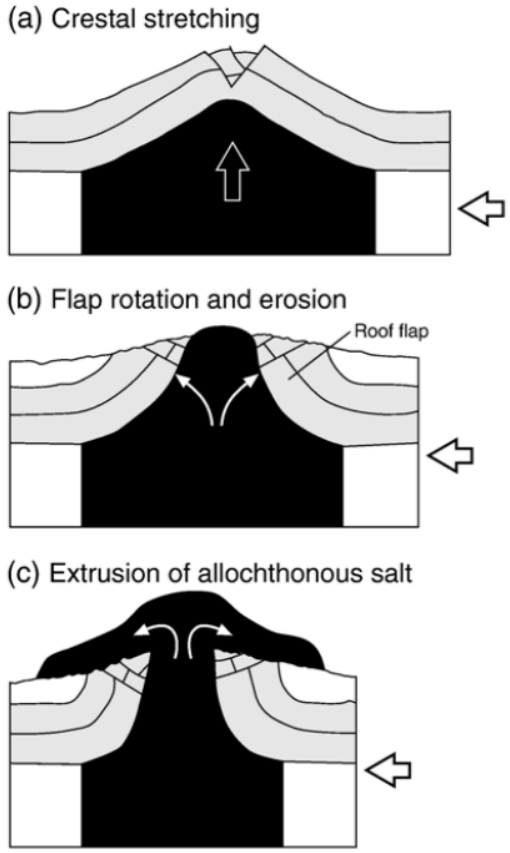


Figure 1: Example of a salt dome formation (diapir piercement) generated by compressional forces (regional shortening). This example assumes a pre-existing salt diapir, which is the most common scenario for piercement during shortening. Diapirs do not always progress through all these stages; whether or not they do depends on the magnitude of the shortening and the roof thickness above the diapir. (a) Arching of the diapir roof produces outer-arc extension, thinning, and weakening the roof. (b) The combination of a weakened roof and salt pressurized by lateral squeezing initiates the upward salt flow. Salt breaks through the thinned crest of the anticline, and the roof flaps on either side rotate away from the emerged diapir. (c) Rapid extrusion; salt displaced from the squeezed feeder flows out over the surface to form a salt glacier. Figures and captions from Hudec and Jackson (2007). [**NR**] guillaume/. saltdomecompressionforces1

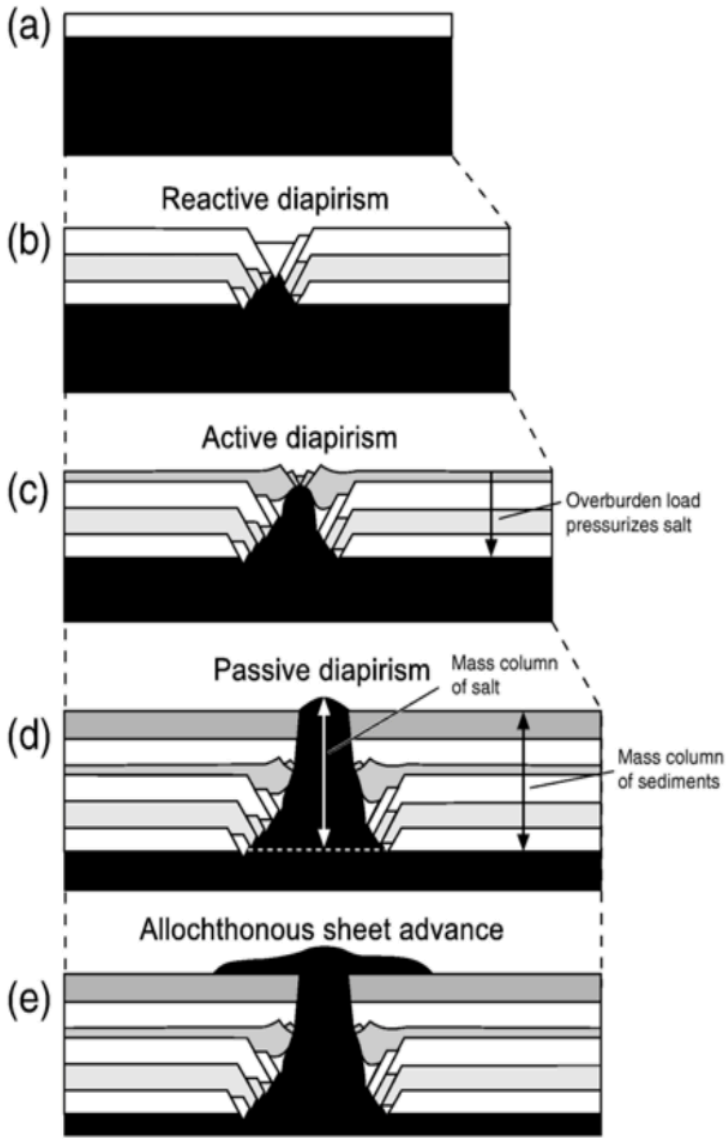


Figure 2: Example of a salt dome formation (diapir piercement) during regional extension. Diapirs do not necessarily progress through all these stages. The maturity of a given structure depends on the availability of salt, total amount of extension, and relative rates of extension and sedimentation. Figures and captions from Hudec and Jackson (2007). [**NR**] guillaume/. saltdomeextensionforces1

salt feeding the dome or until the upward movement is halted by a more rigid formation (Figure 3). Once equilibrium is reached, upward movement of the salt dome ceases, but may begin again if sufficient sediments are added to the weight of the overburden which again increases the load pressure on the parent salt mass (King, 2015).

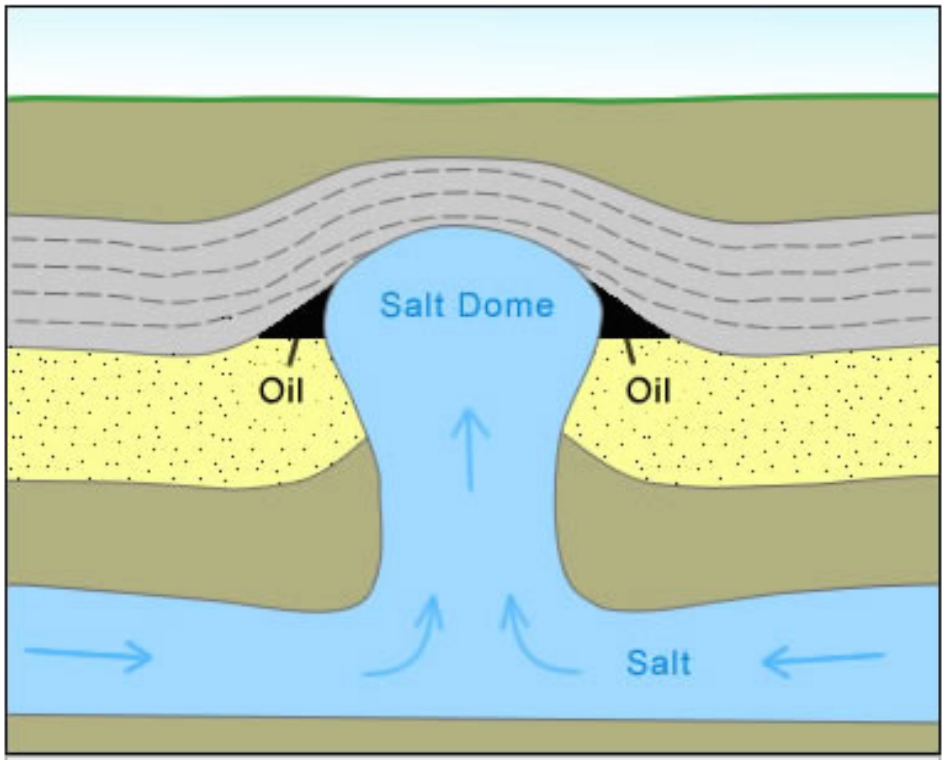


Figure 3: Schematic of the upward movement of the salt creating a salt dome. A salt dome is a type of structural dome formed when a thick bed of evaporite minerals (mainly salt or halite) intrudes vertically into surrounding rock strata, forming a diapir (defined as a mass of salt that has flowed ductilely and appears to have pierced or intruded the overburden). Figure and captions from King (2015). [**NR**] guillaume/. saltdomegeologydotcom

Salt bodies play an important role in the entrapment of oil and gas. First, their cap rock can serve as an oil or gas natural reservoir. Moreover, when salt flows toward the surface, the rocks that it penetrates are arched upwards along the sides of the dome, and faults might be created. This deformation may allow oil and gas to migrate towards the salt dome where it can accumulate in a structural trap, because salt also acts as a seal to fluid migration (Figures 3 and 4). Until relatively recently, energy companies did not search for prospects underneath salt bodies because of poor-quality subsalt seismic data, and because they believed that reservoir-quality rock or hydrocarbons did not exist below salt layers. Advances in seismic processing, imaging, and compelling drilling results from exploration wells encouraged companies to generate and drill prospects below salt layers, salt sheets, and other previously disregarded potential traps (Schlumberger, 2015). Therefore, a single salt body/dome can have many associated reservoirs at a variety of depths and locations in its vicinity.

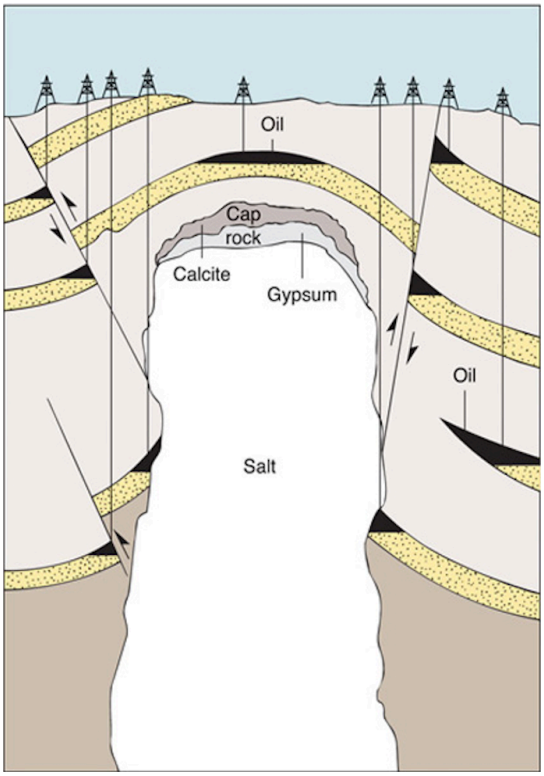


Figure 4: Salt dome, illustrating how the rising salt causes the surrounding sedimentary layers to be arched up, forming areas in which hydrocarbons might accumulate. Figure and captions from Levin (2009). [**NR**] guillaume/. saltdome2

Though potentially very rewarding, drilling through salt can be challenging, costly, and risky. Drillers have to address factors that cause openhole instability and accompanying problems, including borehole walls weakened by incompatible muds, restrictions and an undergauge hole caused by salt creep; or an enlargement caused by dissolution. Moreover, the actual salt rock is weak and undergoes continuous deformation. During the life of a well, salt movement can displace wellbore tubulars, possibly causing failure or restricted access (Farmer et al., 1996). Therefore, obtaining well-focused images of subsalt potential reservoirs is extremely important for risk assessment, as well as for the decision-making process.

Geophysical properties of salt bodies

Leveille et al. (2011) provide a description of a typical geophysical setup that one should expect when exploring a zone containing salt bodies. The salt dome is usually a thick salt sheet (1 to 5 km), with a very rugose top and a relatively smooth base with potential sharp discontinuities coming either from the suturing of two or more salt bodies or from normal faulting. The sediments at the top of the salt sheet are much younger than the salt, while the deeper (i.e., subsalt sediments) can be of any age. The flank of the salt dome is usually very steep (Figures 4 and 5). Pure salt P-wave velocity is typically 4,500 m/s, with a V_p/V_s ratio of 1.9, and a mass density of 2.16 g/cm^3. Salt is extremely light and very fast acoustically. For comparison, a shale with the same P-wave velocity would have a mass density of approximately 2.65 g/cm^3. Near the top of the salt, one should expect younger sediments; with P-wave velocities of approximately 2,000 m/s near the sea floor, and possibly up to 3,000 m/s near the top of the salt. Near the base of the salt, the range is from 3,000 m/s to as high as 4,000 to 4,500 m/s. Moreover, it is common to have highly overpressured sediments near the base of the salt, which can have velocities even lower than 3,000 m/s.

Challenges of seismic imaging in the vicinity of salt bodies

From available literature, we identified six major reasons why seismic imaging in the presence of salt bodies is a challenge (Leveille et al. (2011); Ritter et al. (2010); Jones and Davison (2014)).

1. The main physical problem with the presence of thick allochthonous salt sheets surrounded by acoustic softer sediments is an illumination issue related to the geometry of the sediment-salt interface. Because of the strong acoustic impedance and velocity contrast between the sediments and the salt, the critical angles between sediments and salt are greatly decreased (approximately 20 to 30 degrees when the top-salt is near the seafloor). Therefore, the energy from a shot is rapidly defocused as it passes through the salt and travels down the base. Only a fraction of the incident energy arrives at a subsalt reflector. Moreover, the subsalt energy that is reflected becomes subject to another critical effect barrier (i.e., the critical angle between subsalt sediments and salt) on its way upward, which can sometimes make certain reflector geometries impossible to image from various directions (such as a reflector dipping at the critical angle relative to the salt base). Variability in the salt-interface geometry, especially

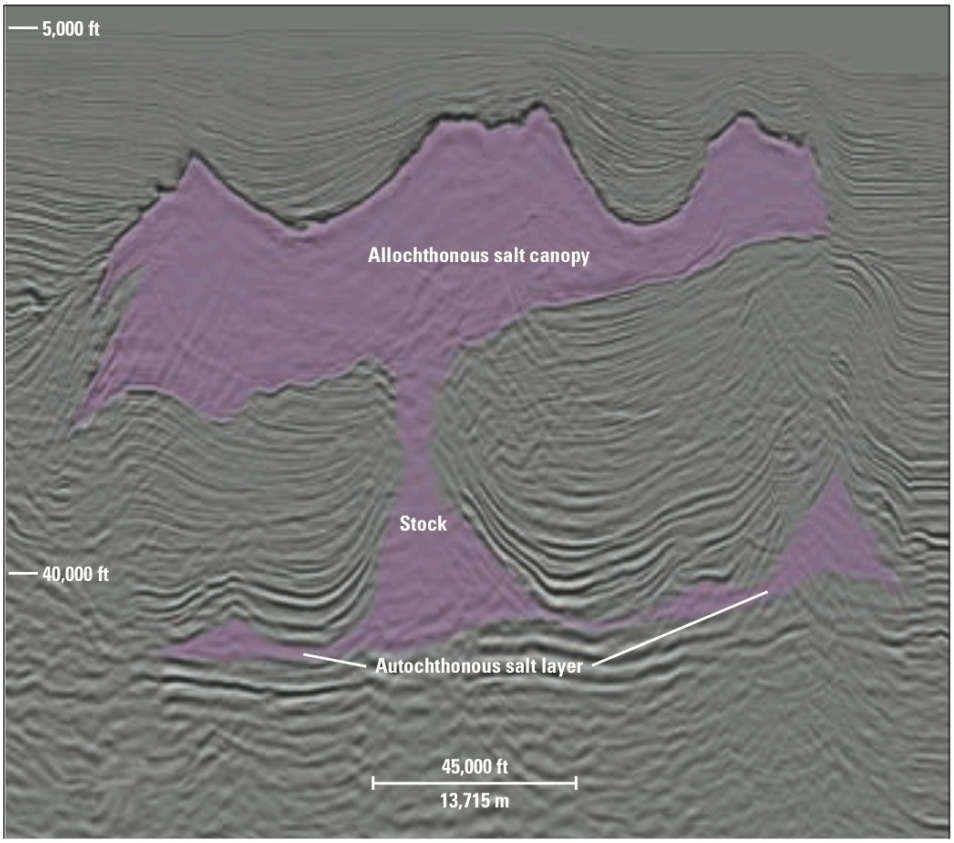


Figure 5: Interpreted seismic section of a salt body beneath the Continental Slope off Louisiana. This figure shows the evolving salt structure. Figure and captions from Dribus et al. (2008). [**NR**] guillaume/. saltseismic

rapid variability, causes defocusing of the raypaths, thereby producing holes in the illumination pattern subsalt. These holes are dependent on the local geometry of the salt surface, and therefore can vary rapidly with the azimuths of the raypaths (this effect has been used to attempt to fill in the illumination holes by finding potential azimuths where the holes do not exist). It is often stated that the salt acts as a partial "mirror," which is commonly misinterpreted to mean that the salt causes the energy losses simply by reflection/transmission losses, but it is not the case (Leveille et al., 2011).

2. The interpretation and the adjusting of fine-scale details around the salt boundaries can be very challenging and tedious. A small error in the positioning of this interface the can have severe consequences and distort the image-quality of underlying subsurface features. This issue is the main focus of our future research and we discuss it into more detail in the next section.

3. Subsalt reflectors are in general very weak reflectors, with reflection coefficients in the 5% range; at least for a large portion of the subsalt projects in the Gulf of Mexico (Leveille et al., 2011).

4. On top of having a strong acoustic impedance contrast with its surrounding sediments, salt bodies and diapirs may also have very complex geometrical shapes (e.g., steeply dipping and overturned flanks), which makes the travelpaths of seismic waves in the vicinity of salt bodies very complicated to model and to interpret. Therefore, it requires the use of costly imaging algorithms (e.g., two-way wave-equation engines that can model double bounces and turning wave reflections).

5. When a salt body moves, the presence of dirty salts can entrain sediments with it. These sediments can get trapped between two colliding salt sheets (referred to as "sutures"). There can be several sutures within what looks like a large continuous salt body. The result is that allochthonous salt sheets can vary in composition, from more or less pure salt (referred to as "dirty salt") containing trapped sediments, usually shale material. The composition and the seismic velocity of the material trapped in the sutures can vary widely, which makes it challenging to model accurately.

6. Lack of anisotropy representation and parametrization within a salt body: most of the time, there is not sufficient information to describe the anisotropic behavior inside salt bodies. But, all deformed salt bodies contain interbeds of different compositions and mineral grains, which are elongated in the salt flow direction (Hudec and Jackson, 2007). The direction is roughly vertical in salt diapirs, and parallel to strata in autochthonous layers (salt bed or layer located in the emplacement where it was originally deposited). An ultrasonic P-wave velocity anisotropy (wavelength less than 1 mm) has already been measured at up to 7% faster in the flow direction (Jones and Davison, 2014). However, to date, there has been little attempt to incorporate salt anisotropy into velocity model building.

CHALLENGES SPECIFIC TO TOP-SALT DELINEATION

Even with today's best seismic imaging algorithms, such as migration velocity analysis (MVA), full-waveform inversion (FWI), and tomographic full-waveform inversion (TFWI),

the very rugose geometry of the sediment-salt interface is difficult to model and delineate accurately to the scale needed to obtain high-quality images of subsalt reflectors. We analyze in this section why current algorithms may still fail to provide a high-resolution image of the boundary. From available literature, we identified four challenges that could explain the limitations of our current algorithms.

Complex transition zone

The transition between younger overlying sediments and the rock salt is rarely as clean as we may think. During the formation of salt domes or salt bodies, large blocks of sediments (referred to as "rafts" or "carapaces") can ride on top of the moving salt. These carapaces, usually shales or carbonates, can sit on the salt and be interposed between the younger overlying sediments and the top-salt. They are made of older sediments that have a very high density, velocity, and velocity range (more than 4,500 m/s). This wide velocity range is primarily influenced by other post-depositional, diagenetic processes, and not just by pure compaction at increasing burial (Anselmetti and Eberli, 1993; Leveille et al., 2011). Carbonate carapace layers are usually not very thick (up to several hundreds of meters (Leveille et al., 2011)), and are very common in the Gulf of Mexico. They generally have a velocity relatively close to salt, which means that they can negate the impedance contrast that usually marks the sediment-salt boundary (Ritter et al., 2010). In fact, the expected impedance contrast is often seen somewhere within the carapace and may lead to misinterpretation of the top-salt. An incorrect interpretation may create a distortion in the subsalt events caused by the misplacement of the top of salt and too slow overlying sediment velocities (Ritter et al., 2010). In some cases, it is also possible to have a reflection from the carapace-salt interface generated by a pure density contrast. Because most of the current seismic imaging algorithms assume a constant density Earth, they would not be able to accurately model these reflections.

Figure 6 shows a set of three well logs from two different wells drilled into the same salt body in the Gulf of Mexico (the red and green logs come from the same well). On the blue log (sonic P-wave velocity), a clear interface is visible at a relative depth of 300 m, but we can also identify an intermediate zone, which can be interpreted as a carapace, with a thickness of approximately 150 m interposed between the bottom of the sediments, and the top of the salt. This log illustrates the issue stemming from the existence of a transition zone between the young uncompacted sediments and the salt rock. In this case, an interpreter could identify the interface between the sediments and the carapace as the top-salt, and assign an inappropriate (too-low) velocity value to the few 100 meters of the carapace layer. The base salt interface, however, appears a lot cleaner. As expected, the subsalt sediments have a lower P-wave velocity (red curve) than the actual salt, but are denser (green curve). It is also important to notice that the salt sonic P-wave velocity appears to be higher in average on the red curve, which may indicate some compaction effect within the salt body.

Another potential difficulty can arise from the fact that many buried salt diapirs have a cap rock. It is a carapace of anhydrite, gypsum, and limestone (in ascending order) immediately above the salt rock of a salt dome (Figures 7(a) and (b)). The average cap rock is between 100 m to 150 m thick, but can be as thick as 300 m. It is a secondary product created as the salt dome ascends through the overlying materials. The top of the dome dissolves as it rises through the sediments and the material constituents of the salt

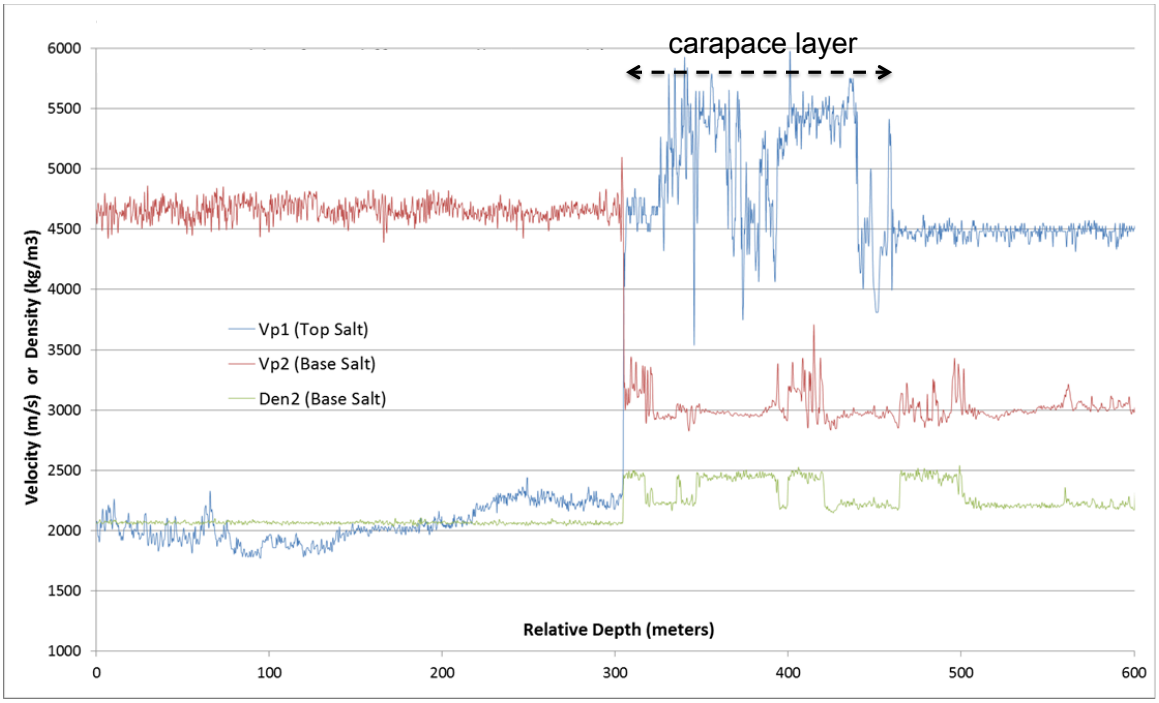


Figure 6: Well logs from one salt body in the Gulf of Mexico. The blue curve shows the sonic log for the P-wave velocity above and below the top-salt interface as a function of relative depth (the top-salt interface has been shifted to a relative depth of 300 m for proprietary reasons). The red curve shows the sonic log for P-wave velocity above and below the base-salt interface as a function of relative depth (the base-salt interface has also been shifted to a relative depth of 300 m). The green curve shows the density log around the base salt as a function of relative depth, and comes from the same well as the one for the red curve. The blue log comes from a different well. The wells were drilled from the same platform into the same salt body, but at different penetration points. Courtesy of Chevron Corporation. [**NR**] guillaume/. saltwelllogchevron

become concentrated at the top of the rising plug becoming three layers. The cap rock of some salt domes, and sometimes the actual rock salt, is found to overhang, or drape down the sides of the main salt mass, as shown in Figure 7(c). Salt is soluble and dissolves, but the cap rock is not very soluble and thereby left more intact. Thus the appearance of the cap rock overhanging the main salt body is the result of dissolution of the salt body within the subsurface, but not dissolving the cap rock. The presence of a cap rock may makes the sediment-salt interface difficult to accurately delineate.

Inaccurate velocity model of the overburden

To correctly image the top-salt, a very precise overburden velocity model is needed. Typically, a FWI inversion scheme could be used to build such a model. However, if the top-salt is buried too deeply into the subsurface, or if recorded data lack transmitted or refracted low-frequency energy, the inversion scheme might fail to deliver an accurate overburden velocity model down to the top of the salt.

Tortuous top-salt geometrical shapes

Current algorithms may fail to get an ideal image of what the top-salt is because of the complex geometry of certain salt bodies. Figures 8(a) and (b) show two cross-sections of a velocity model from the Gulf of Mexico wide-azimuth "E-Octopus" dataset, and illustrate the type of top-salt shapes encountered in salt basins. The circled zones are challenging to model accurately, even with a long-offset and wide-azimuth acquisition.

Lack of good-quality high-frequency data

Current algorithms can be limited by the low signal to noise ratio for higher frequencies contained in the data. This limitation may set an upper bound on the resolution of our inverted velocity model that may be too low to recover some of the fine-scale features of the top-salt interface.

EFFECTS OF INACCURATE TOP-SALT DELINEATION ON WAVEFIELDS

As shown in recent work (Etgen et al., 2014a,b; Albertin et al., 2014), misinterpreting the top-salt boundary can deteriorate the images of underlying layers. In this section, we illustrate the effect of a top-salt misinterpretation by looking at the actual wavefields. We propose two scenarios; one when the true top of the salt body is relatively smooth, and one when the true top-salt is rugose. We simulate a seismic reflection experiment in which we place a single source at a depth $z_s = 0$ m, a horizontal position $x_s = 500$ m, and 600 geophones equally spaced along the horizontal axis, at a depth $z_r = 0$ m. The reflector we try to image is located at a depth $z_{ref} = 4.7$ km. We perform a partial migration using a single shot, and a two-way wave-equation acoustic and isotropic engine. As we purposely decrease the accuracy of the top-salt picking, we analyze the change in behavior of the back-propagated receiver wavefield and the partially migrated images.

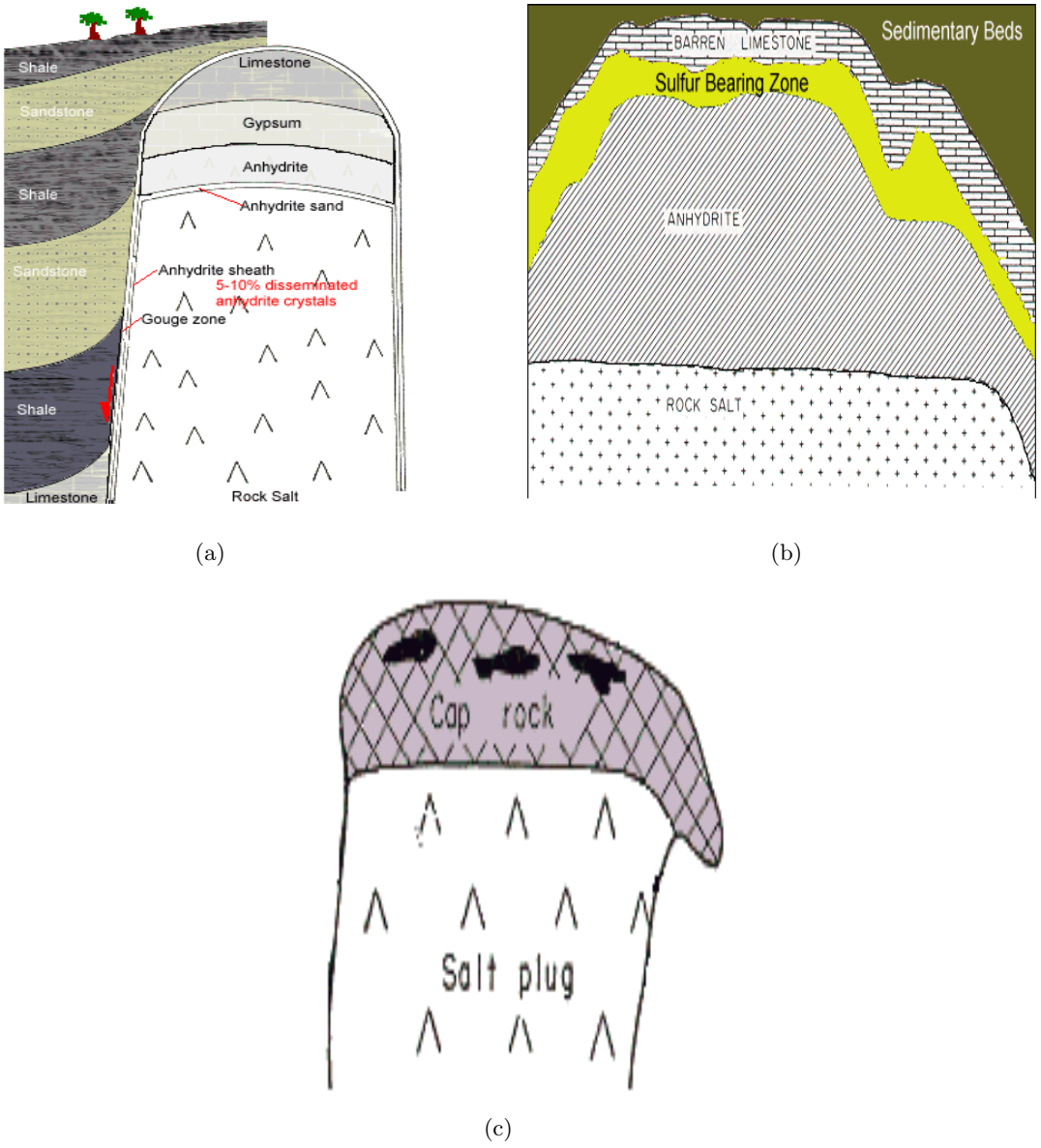
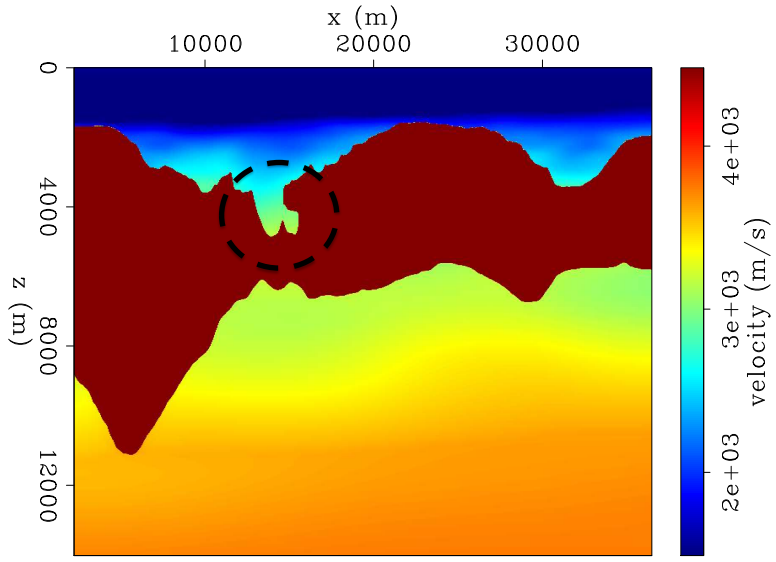
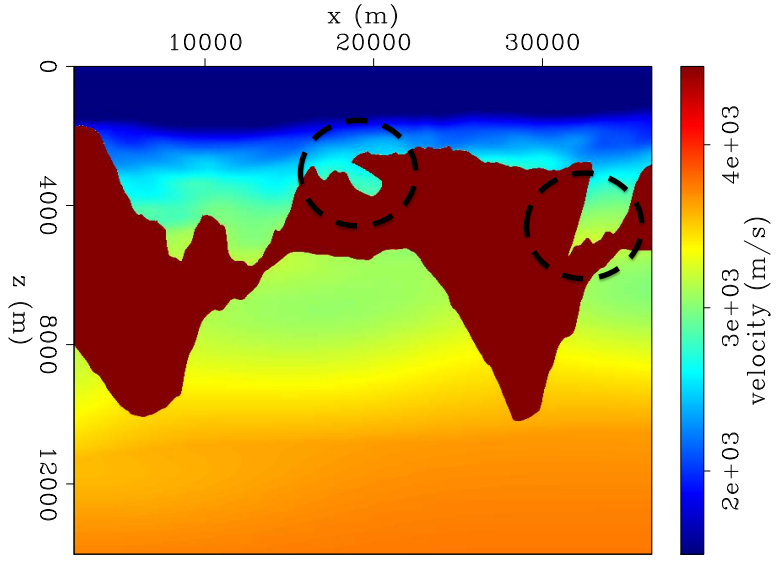


Figure 7: Schematic illustration of salt cap rocks. (a) The cap is composed of limestone located at the top of the dome followed by, in descending order, gypsum, anhydrite, and finally rock salt. (b) Cap rock has all the characteristics of a petroleum reservoir and often contains hydrocarbons. But cap rock also contains sulfur (i.e., gyspum) that is present in the cap rocks of almost all salt domes. The sulfur forms after the limestone is created and is derived from the destruction of the anhydrite. (c) The cap rock of some salt domes is found to overhang or drape down the side of the main salt mass. It is commonly found to overhang on one or two sides. Figures and captions from DNR (2015). [**NR**] guillaume/. saltdomecaprock,caprockcomposition,caprockoverhang



(a)



(b)

Figure 8: Cross-sections from E-Octopus data set (Gulf of Mexico) displaying the tortuosity of the top of salt bodies. Circled areas may suffer from a lack of illumination and are challenging to model accurately. Courtesy of WesternGeco. [**NR**] guillaume/. eoctopusframe1,eoctopusframe2

Smooth top-salt

Figures 9(a)-(d) show the true model, true background, inaccurate background (i.e., the background with a misinterpreted top-salt), and difference between the true and the incorrect backgrounds, respectively. Figure 10 (left and right columns) show three snapshots of the receiver wavefields back-propagated in time, and computed with the two different background models. The left column shows the wavefield computed with the correct background model. The receiver wavefield is then cross-correlated in time with the source wavefield to generate a partially migrated image (Figure 11(a)). The right column of Figure 10 shows the receiver wavefield computed with the inaccurate background model. The contribution of this shot to the migrated image is displayed in Figure 11(b). Even though the inaccuracy of the top-salt affects the behavior and the shape of the receiver wavefield, it does not damage it considerably. The partial image in Figure 11(b) is distorted, but it will likely be possible to improve the fully migrated image (i.e., the image generated with more shots) by applying a tomographic optimzation scheme (e.g., WEMVA).

Rugose top salt

In this test, the true top-salt boundary is rugose. Figures 12(a)-(d) show the true model, true background, inaccurate background (i.e., the model with the misinterpreted top-salt), and difference between the true and the incorrect backgrounds, respectively. Figure 13 (left column) shows three snapshots of the receiver wavefield computed with the correct background model, while Figure 13 (right column) displays the receiver wavefield computed with the inaccurate background. We see two major differences from the previous case.
Firstly, we notice that the receiver wavefield has a much more complex shape, especially at later physical times (Figures 13(a) and (b)), which is because the wavefield that illuminates the reflector propagates twice through the rugose top-salt before it is recorded at the surface.
Moreover, we can see that in Figures 13(c) and (e) the back-propagated receiver wavefield computed with the accurate top-salt has been well refocused: its wavefront is much smoother than it is at later physical times (Figure 13(a)). During the back-propagation of the receiver wavefield through the rugose top-salt, the fine-scale features of the salt boundary behave as point scatterers. Because we are using the correct top-salt, they collapse and remove most of the diffraction artifacts present in the wavefield at later physical times (Figure 13(a)). The partially migrated image from this shot (Figure14(a)) is accurate. However, in the case of a misinterpreted top-salt, the receiver wavefield does not refocus properly. The back-propagation through the inaccurate salt boundary creates even more asynchronized scattering (Figures 13(d) and (f)). Consequently, the partially migrated image is considerably damaged (Figure14(b)), and the image degradation is much more pronounced than in the previous example (Figure11(b)).

DEGRADATION OF SUBSALT REFLECTORS IMAGE QUALITY

We conduct three noise-free synthetic tests to illustrate the degradation of subsalt layer images when using an inaccurate top-salt model to migrate data. Computations are performed using a two-way wave-equation engine. We use a 25 Hz Ricker wavelet for the

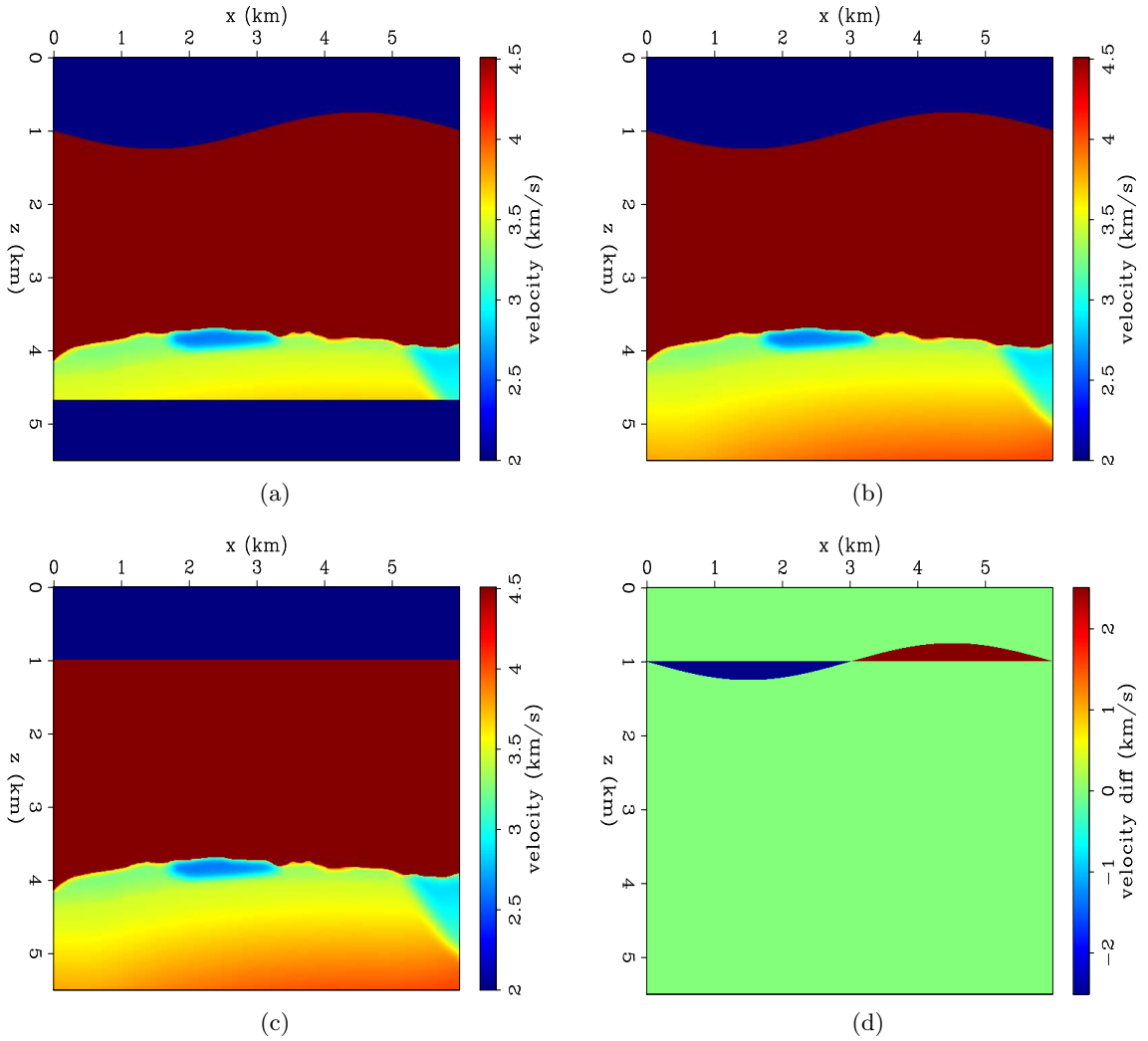


Figure 9: Synthetic velocity models, modified from the 2004 EAGE velocity benchmark study. (a) True velocity model. (b) True background velocity model. (c) Inaccurate background velocity model. (d) Difference between true background and inaccurate background velocity models. [**ER**] guillaume/. bp2Lref,bp2L,bp2Lsmooth,bp2Ldiff

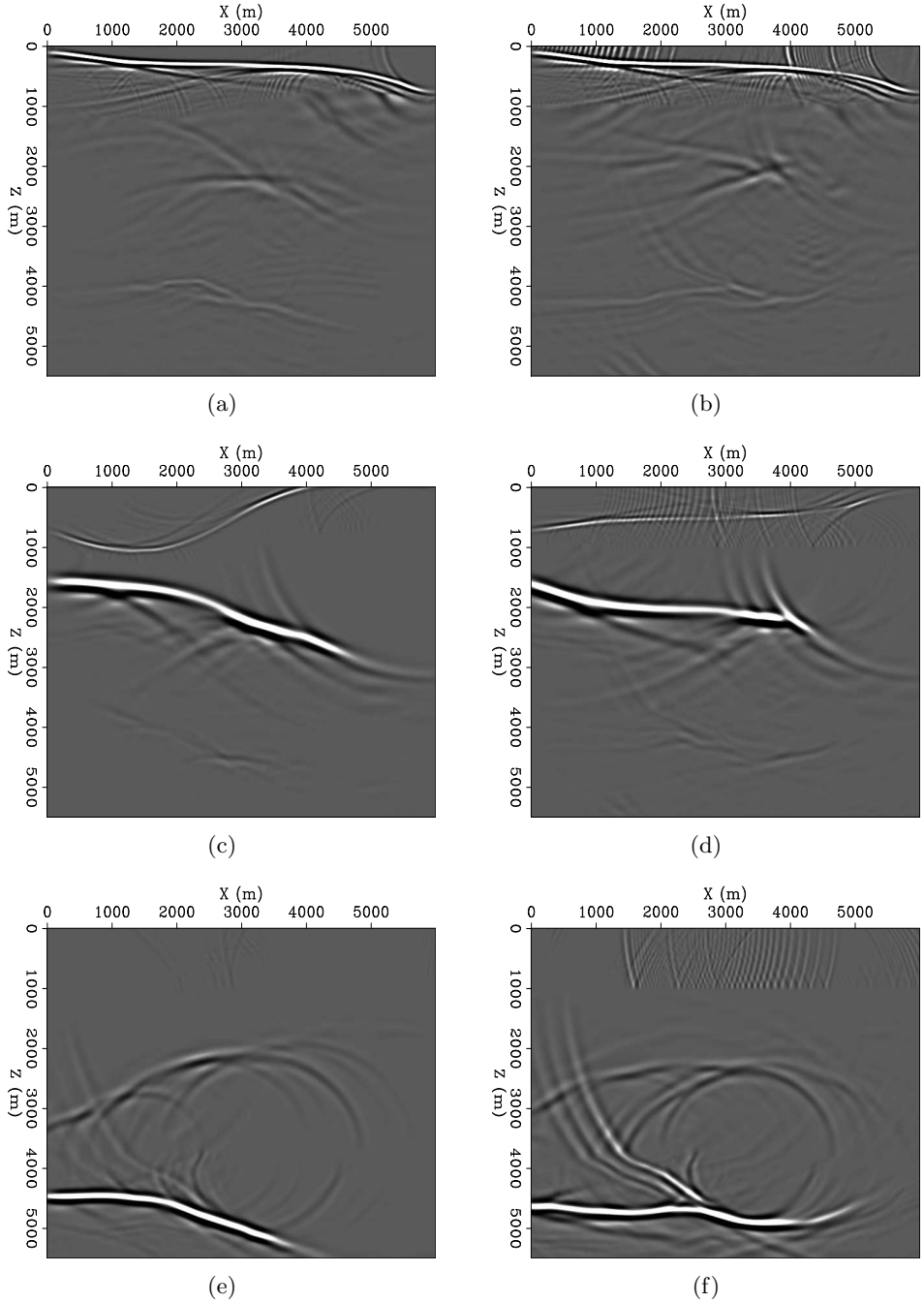


Figure 10: Snapshots of two receiver wavefields. Left column shows three snapshots of the receiver wavefield computed with the true background velocity model. Right column shows three snapshots of the receiver wavefield computed with the inaccurate background velocity model. Snaphots in first row are computed at $t_3 = 2.784$ s. Snapshots in second row are computed at $t_2 = 2.208$ s. Snapshots in last row are computed at $t_1 = 1.536$ s. [**ER**]

guillaume/. bpsimpleframec59,bpsimpleframeb59,bpsimpleframec47,bpsimpleframeb47,bpsimpleframec33,bps

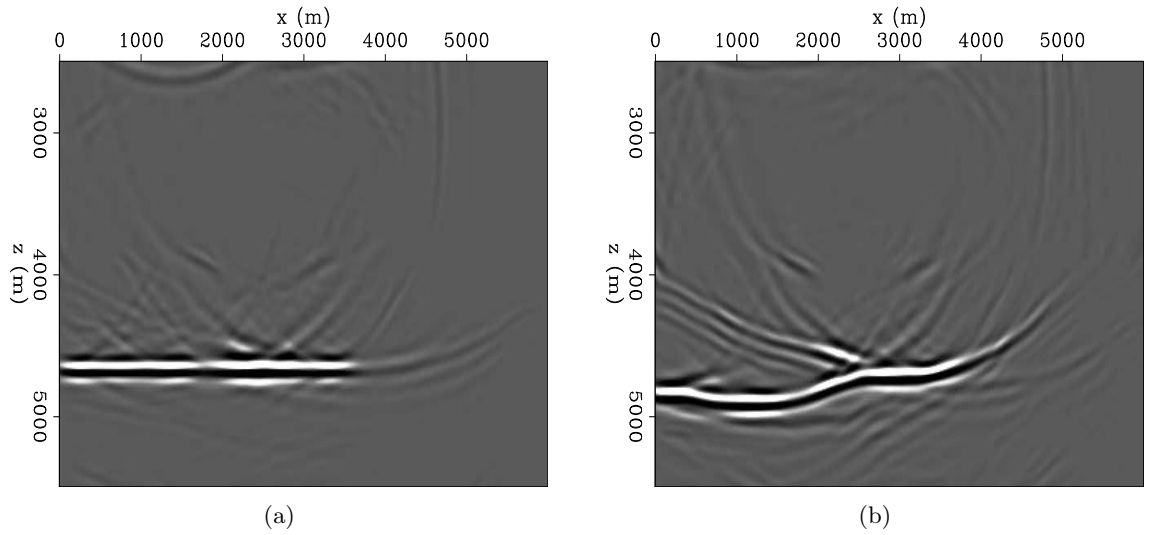


Figure 11: Partially migrated sections (one shot) using a two-way wave-equation engine. (a) Partial image computed using the true background velocity model. (b) Partial image computed using the inaccurate background velocity model. [**ER**] guillaume/. bpsimpleimxc,bpsimpleimxb

source function. Migrated images are generated with 100 sources located at the surface, and equally spaced along the horizontal axis. We set geophones at the surface at every grid point. Images are extended along the horizontal subsurface offset axis, and then converted to subsurface angle domain by a slant-stack operator.

Single flat layer

We perform two migrations with two different background velocity models. Our goal is to image the reflector at $z = 3,500$ m, and see how its image deteriorates with decreasing accuracy of the background model. Figure 15(a) shows the true model used to generate the synthetic data, while Figures 15(b) and (c) show the two different background velocity models used to image the subsalt reflector. Figure 15(d) shows the difference between the inaccurate background and the true background velocity model. The results of the migrated layer are shown in Figure 16(a) and (b). In Figure 16(b), we can see that the misinterpretation of the top-salt has damaged the image of the layer. The reflector shows a strong lack of lateral continuity, and its shape is highly distorted. Moreover, the angle gathers do not provide any coherent moveout information (Figures 17 and 18).

Sinusoidal-shape reflector

A similar experiment as described in the previous section is conducted, but the goal is to image a reflector with a sinusoidal shape. The velocity models are displayed in Figure 19. Figure 20 and shows the migrated images generated by an accurate top-salt model (Figure 20(a)), and an inaccurate top-salt (Figure 20(b)). Though the reflector image is

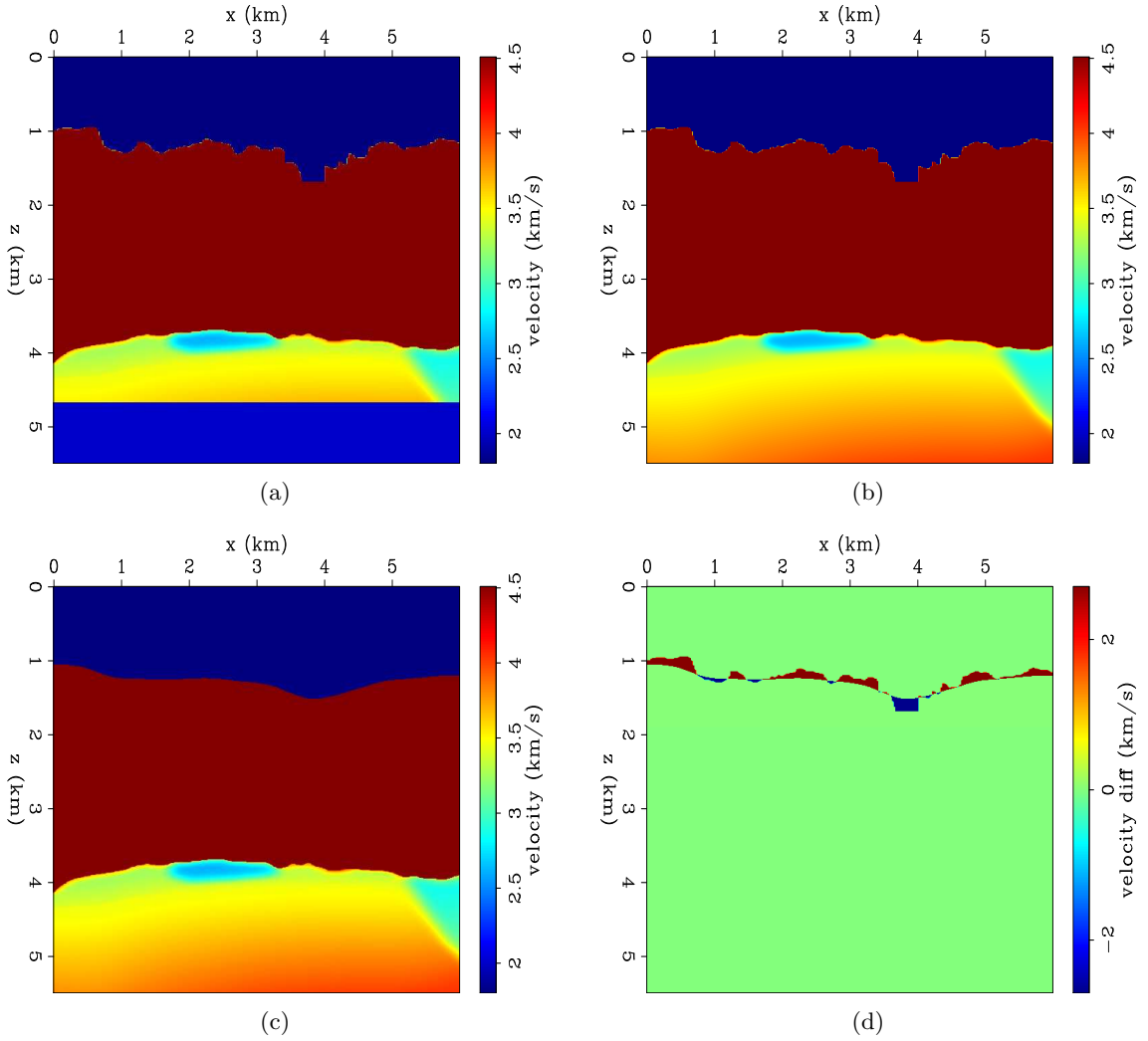


Figure 12: Synthetic velocity models, modified from the 2004 BP synthetic model. (a) True velocity model with a rugose top-salt. (b) True background velocity model. (c) Inaccurate background velocity model. (d) Difference between true background and inaccurate background velocity models. [**ER**] guillaume/. bpsimpleref,bpsimple,bpsimplesmooth,bpsimplediff

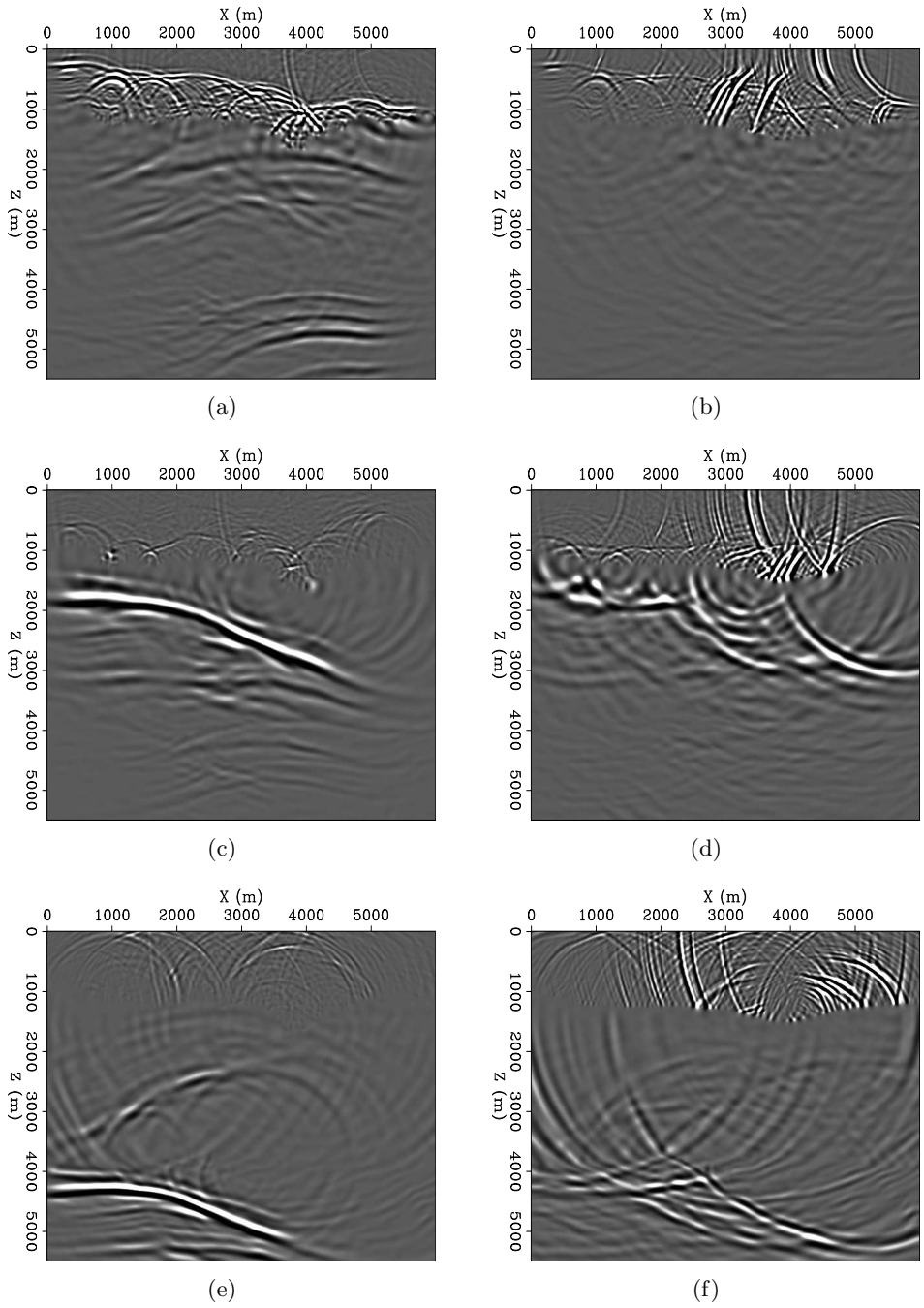


Figure 13: Snapshots of two receiver wavefields. Left column shows three snapshots of the receiver wavefield computed with the true background velocity model. Right column shows three snapshots of the receiver wavefield computed with the inaccurate background velocity model. Snapshots in first row are computed at $t_3 = 2.832$ s. Snapshots in second row are computed at $t_2 = 2.256$ s. Snpashots in last row is computed at $t_1 = 1.680$ s. [**ER**]

guillaume/. simpleframec60,simpleframeb60,simpleframec48,simpleframeb48,simpleframec36,simpleframeb36

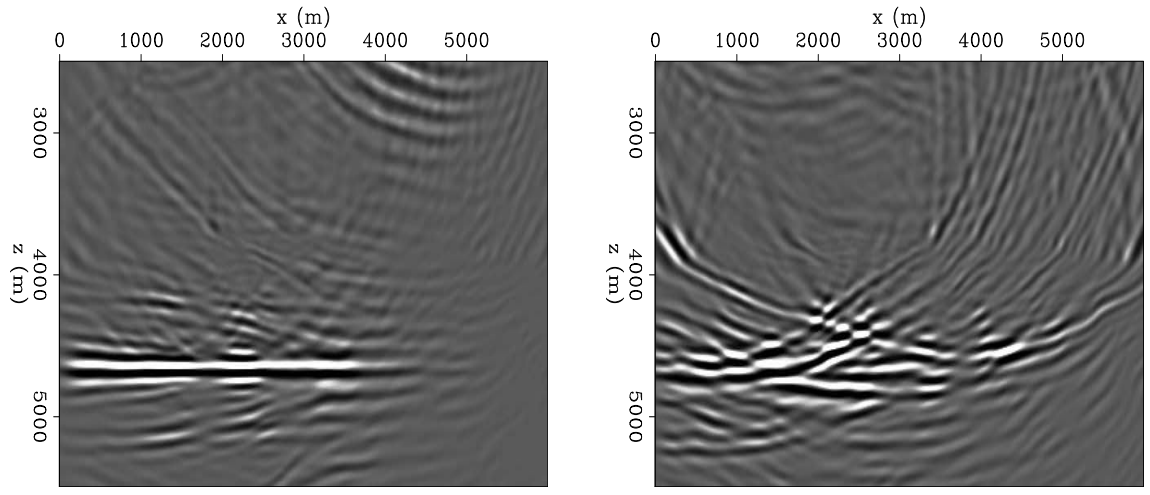


Figure 14: Partially migrated sections (one shot) using a two-way wave-equation engine. The true velocity model contains a salt body with a rugose top-salt. (a) Partial image computed using the true background velocity model. (b) Partial image computed using the inaccurate background velocity model. [**ER**] guillaume/. simpleimxc,simpleimxb

degraded, it is still interpretable and relatively continuous. However, the angle gathers still show no coherency (Figures 21 and 22).

Multiple layers

The two previous examples showed image degradations even though the models had quite simple geometries. In this test, we include two strong reflectors below the base salt (Figure23(a)) to simulate a more complex scenario. Images in Figures 24(a) and Figures 24(b) are computed with the background velocity models displayed in Figure 23(b) and Figure 23(c), respectively. As expected, we see a degradation of the two subsalt layers, as well as the angle gathers (Figure 25 and 26) are not coherent.

Conlusions on synthetic tests

We performed synthetic tests to assess how images of subsalt layers are damaged with inaccurate top-salt delineation. By looking at the actual wavefields, we showed that the lack of accuracy in the top-salt boundary does not enable the receiver wavefield to refocus correctly below the salt, which will in turn deteriorate the migrated image. Adding the contributions from all shots will not stack constructively and the resulting image will show very little coherency. This was confirmed by the second set of tests we performed. The images of subsalt layers are extremely sensitive to the accuracy of the salt boundary. Moreover, as mentioned by Etgen et al. (2014b), it is difficult to extract any useful information from extended images because the gathers show no coherency.

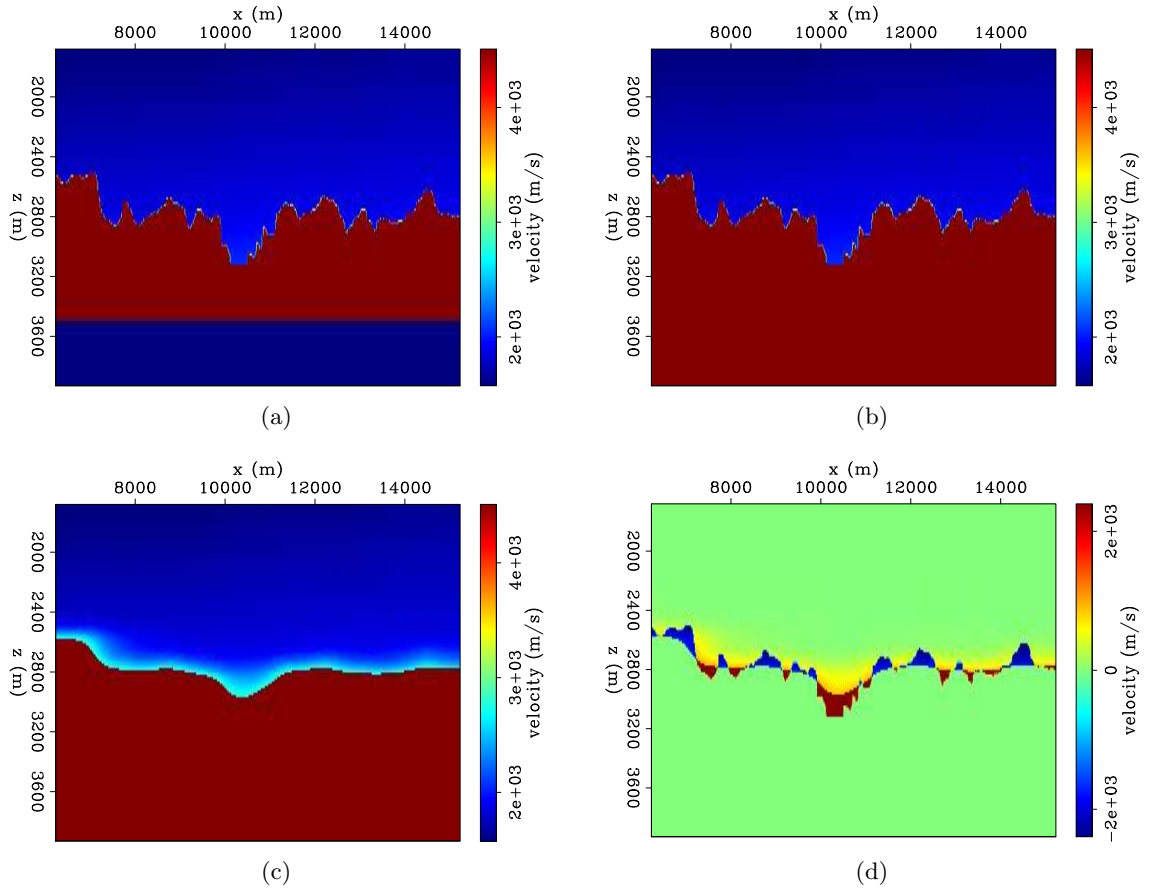


Figure 15: Velocity models modified from the the 2004 EAGE velocity benchmark study. (a) True velocity model. (b) True background velocity model. (c) Inaccurate background velocity model. (d) Difference between true and inaccurate background velocity models. [**ER**] guillaume/. bpr,bpt,bps,bpsdiff

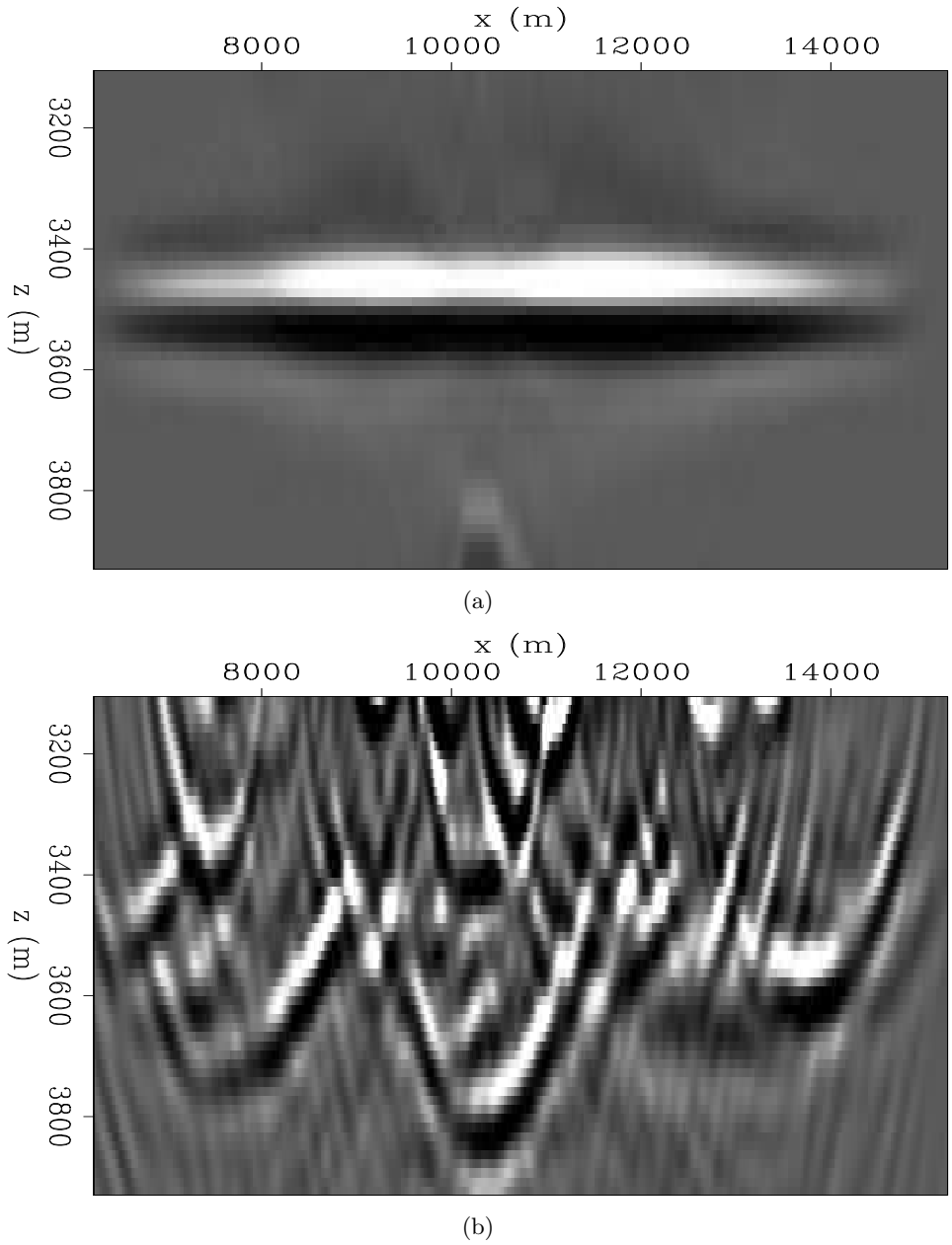


Figure 16: Migrated images of the subsalt reflector shown in Figure 15(a) using a two-way wave-equation engine. (a) Migrated layer using the true background velocity model. (b) Migrated layer using the inaccurate background velocity model. [**ER**] guillaume/. rimxcorrect,rimx

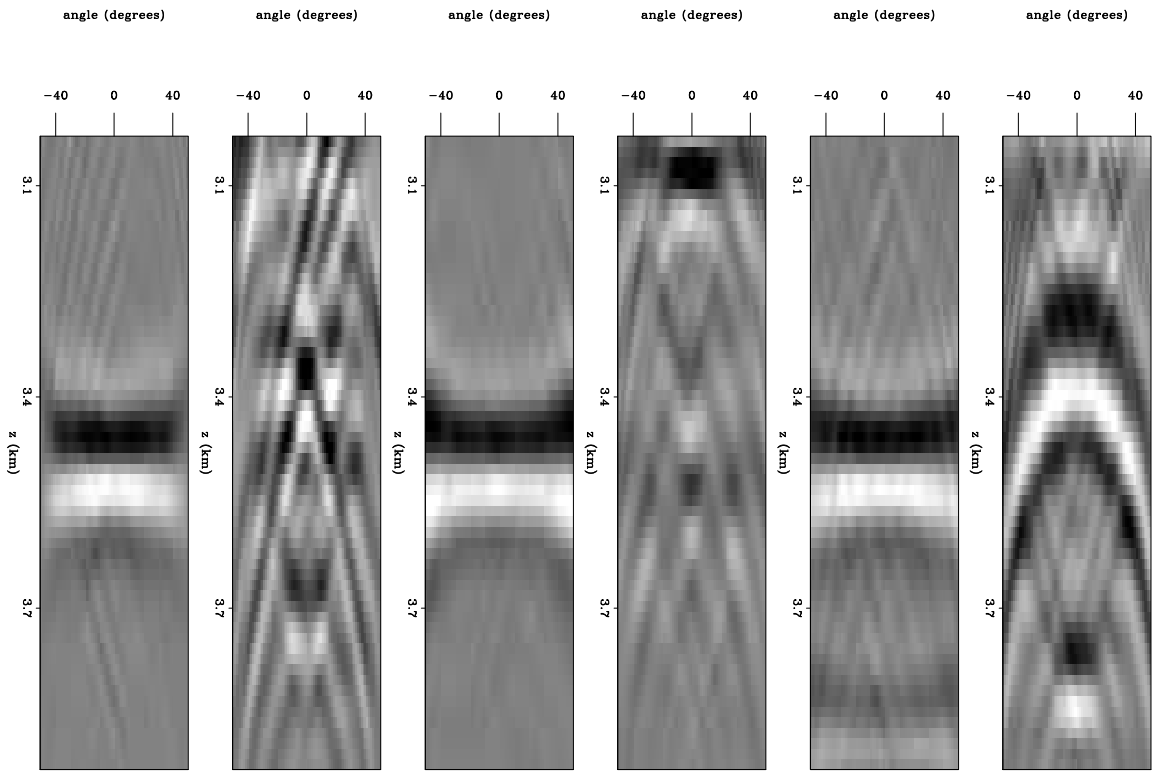


Figure 17: Three pairs of angle domain common image gathers (ADCIG). For each pair, the first image shows the gather computed from the accurate image (Figure 16(a)), and the second image shows the gather computed from the distorted image (Figure 16(b)). The gathers are computed at $x_1 = 7,617.5$ m, $x_2 = 9,177.5$ m, and $x_3 = 10,377.5$ m, respectively. [**ER**] guillaume/. rimagather1

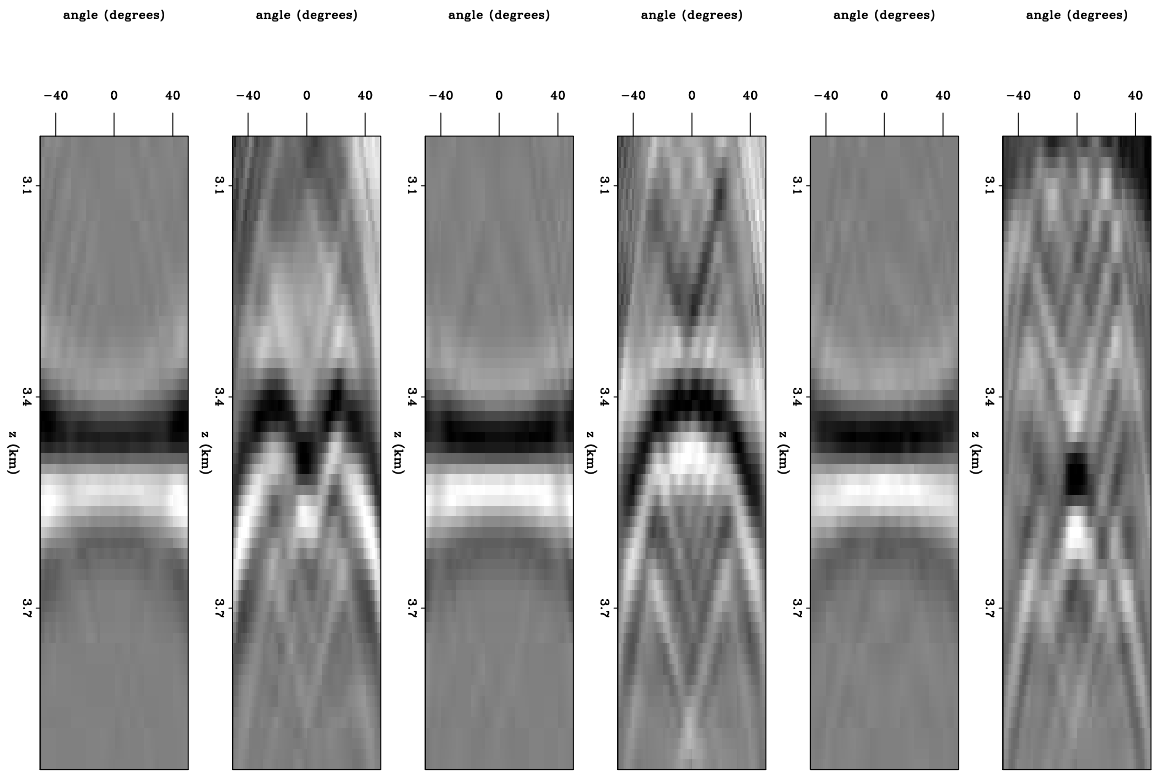


Figure 18: Three pairs of angle domain common image gathers. For each pair, the first image shows the gather computed from the accurate image (Figure 16(a)), and the second image shows the gather computed from the distorted image (Figure 16(b)). The gathers are computed at $x_4 = 11,352.5$ m, $x_5 = 12,267.5$ m, and $x_6 = 13,137.5$ m, respectively. [**ER**] guillaume/. rimagather2

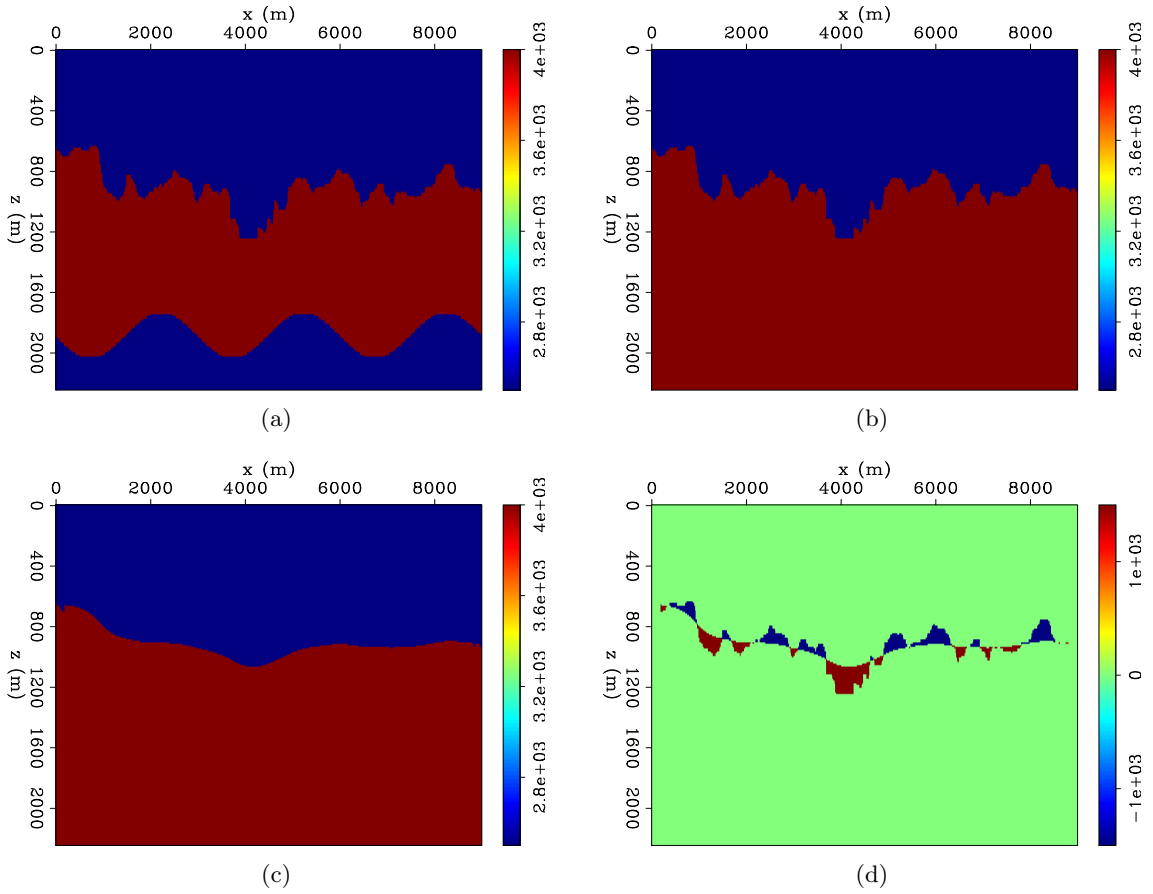


Figure 19: Velocity models modified from the 2004 EAGE velocity benchmark study. (a) True velocity model with a sinusoidal-shape reflector. (b) True background velocity model. (c) Inaccurate background velocity model. (d) Difference between true and inaccurate background velocity models. [**ER**] guillaume/. bprsinh,bptsinh,bptsinsmoothfullh2,bptsindiffh

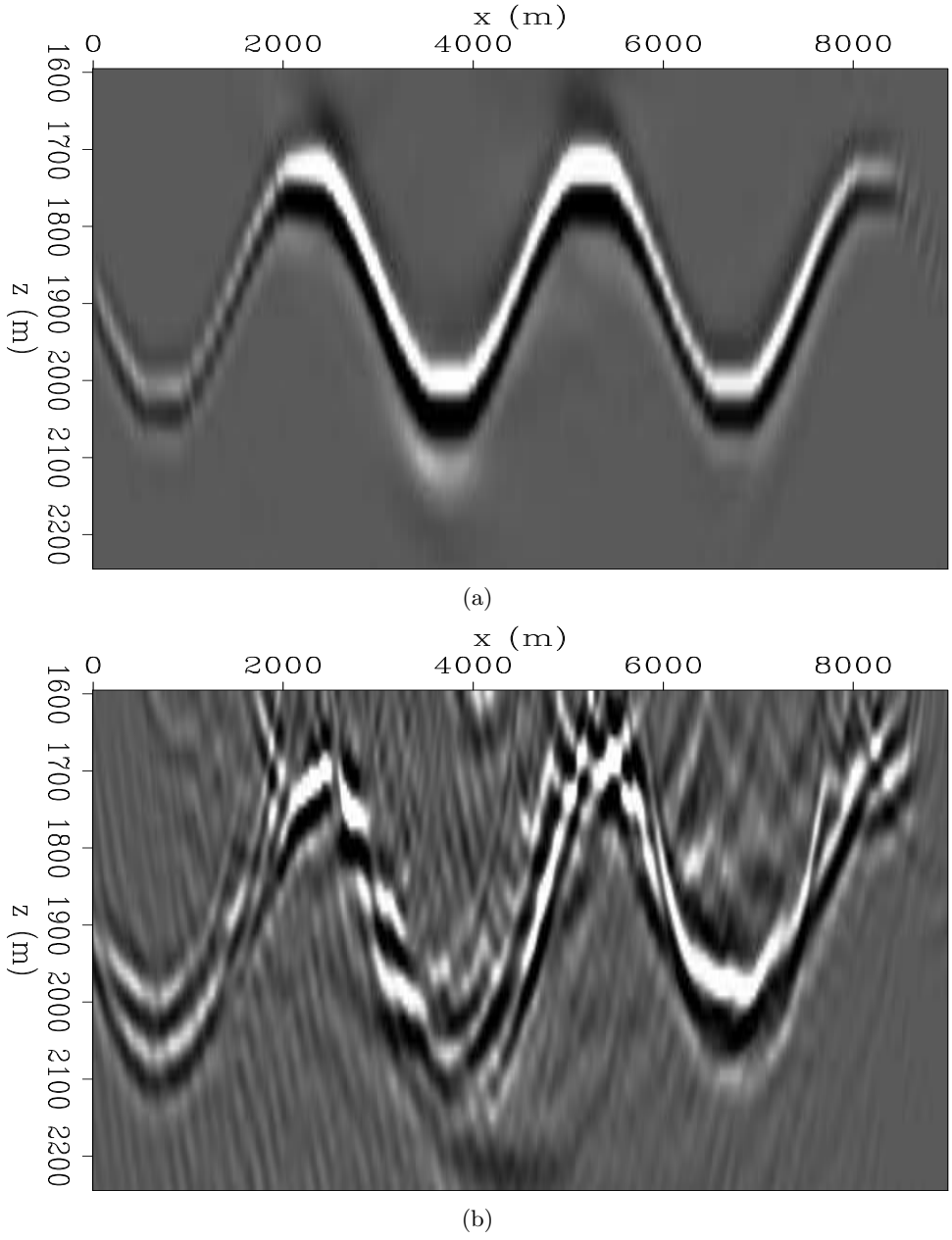


Figure 20: Migrated images. (a) Image computed with the correct background velocity model. (b) Image computed with the inaccurate background velocity model. [**ER**] guillaume/. bpsinimxc25,bpsinimxbf252

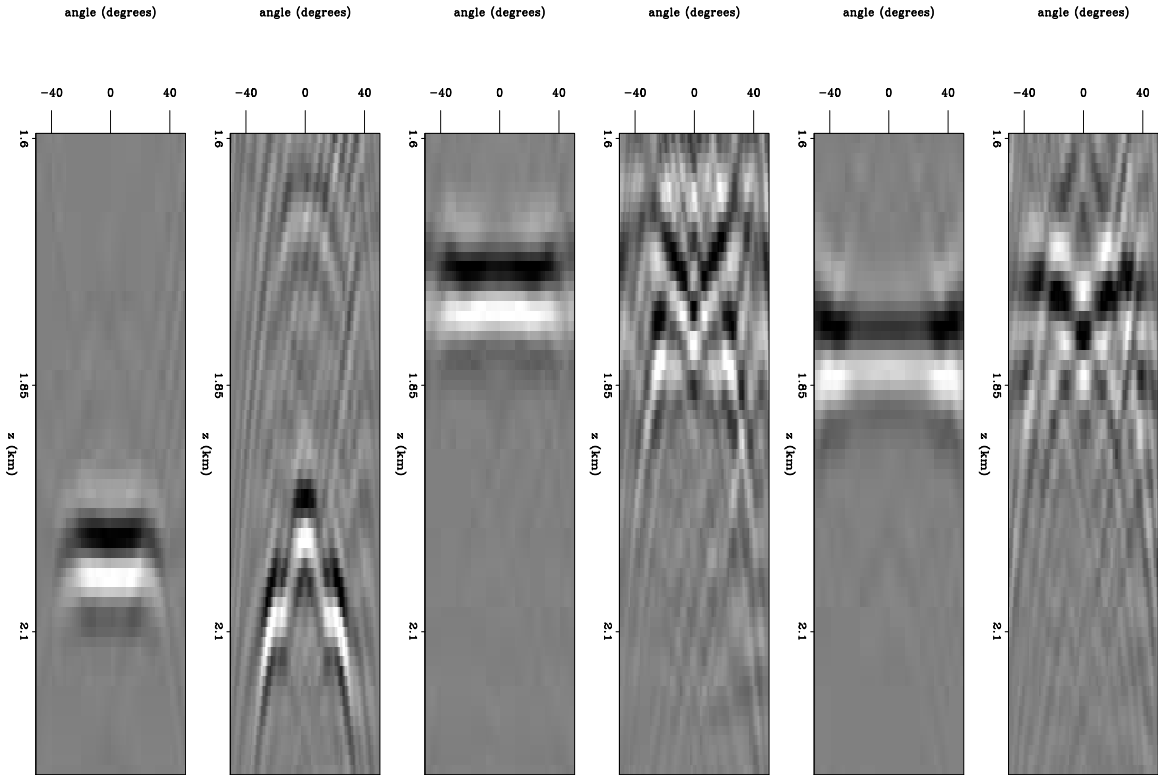


Figure 21: Three pairs of angle domain common image gathers. For each pair, the first image shows the gather computed from the image in Figure 20(a), and the second image shows the gather computed from the image in Figure 20(b). The gathers are computed at $x_1 = 920$ m, $x_2 = 1,196$ m, and $x_3 = 2,760$ m, respectively. [**ER**] guillaume/. bpsinagather1

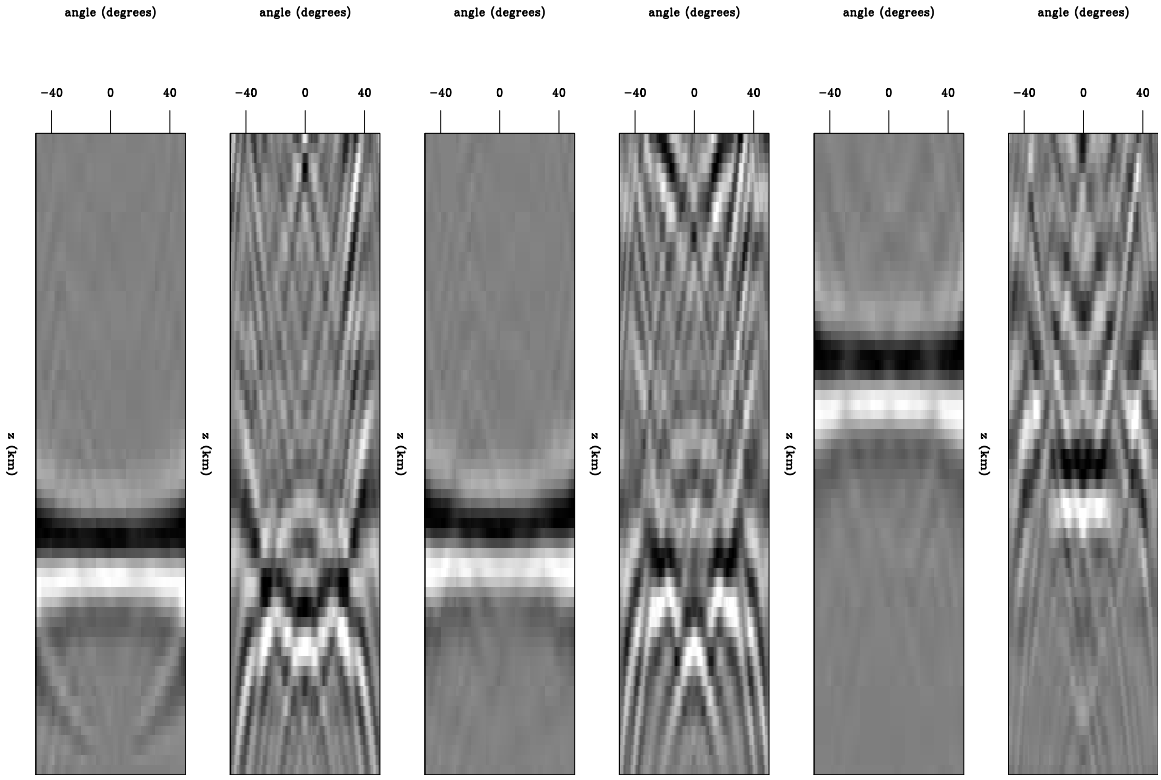


Figure 22: Three pairs of angle domain common image gathers. For each pair, the first image shows the gather computed from the image in Figure 20(a), and the second image shows the gather computed from the image in Figure 20(b). The gathers are computed at $x_4 = 3,730$ m, $x_5 = 4,020$ m, and $x_6 = 4,600$ m, respectively. [**ER**] guillaume/. bpsinagather2

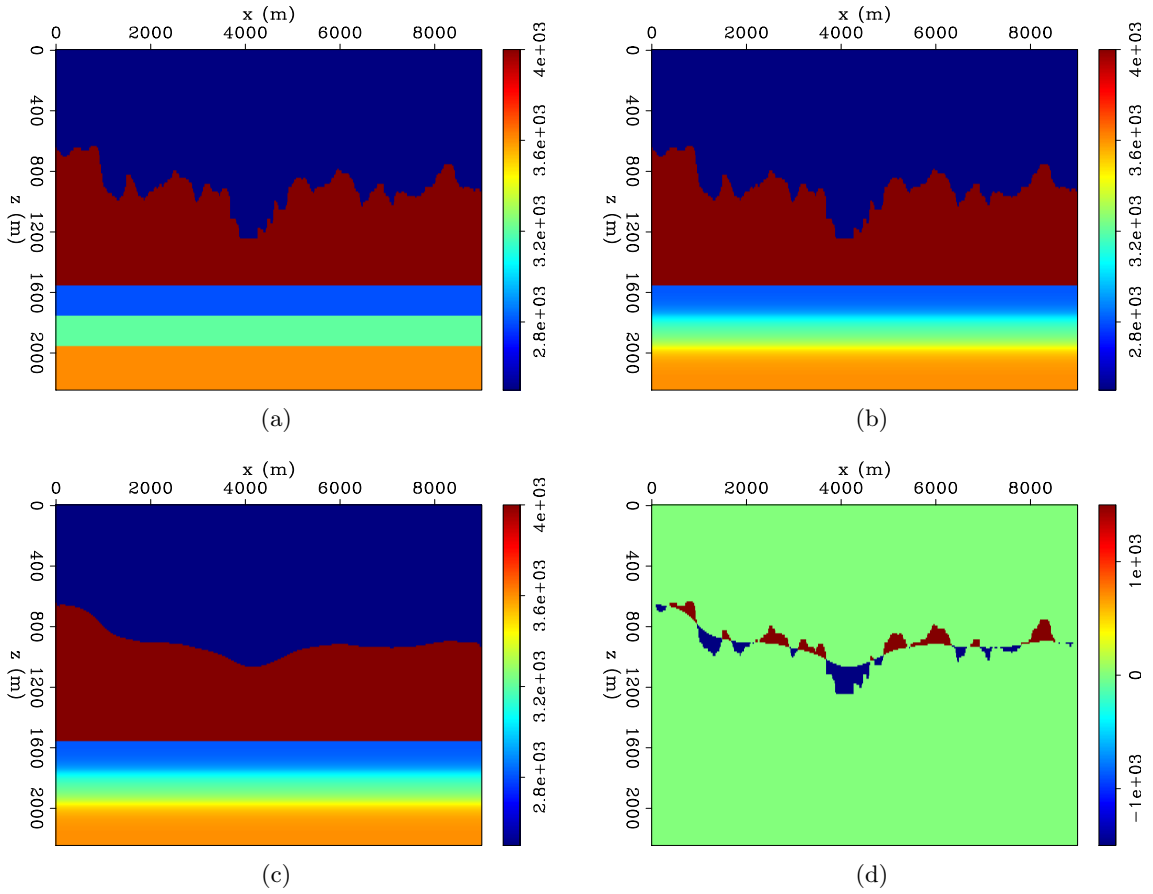


Figure 23: Velocity models modified from the 2004 EAGE velocity benchmark study. (a) True velocity model with two subsalt reflectors. (b) True background velocity model. (c) Inaccurate background velocity model. (d) Difference between true and inaccurate background velocity models. [**ER**] guillaume/. bprhflat3,bpthflat3,bpthsmoothflat3,bpthflat3diff

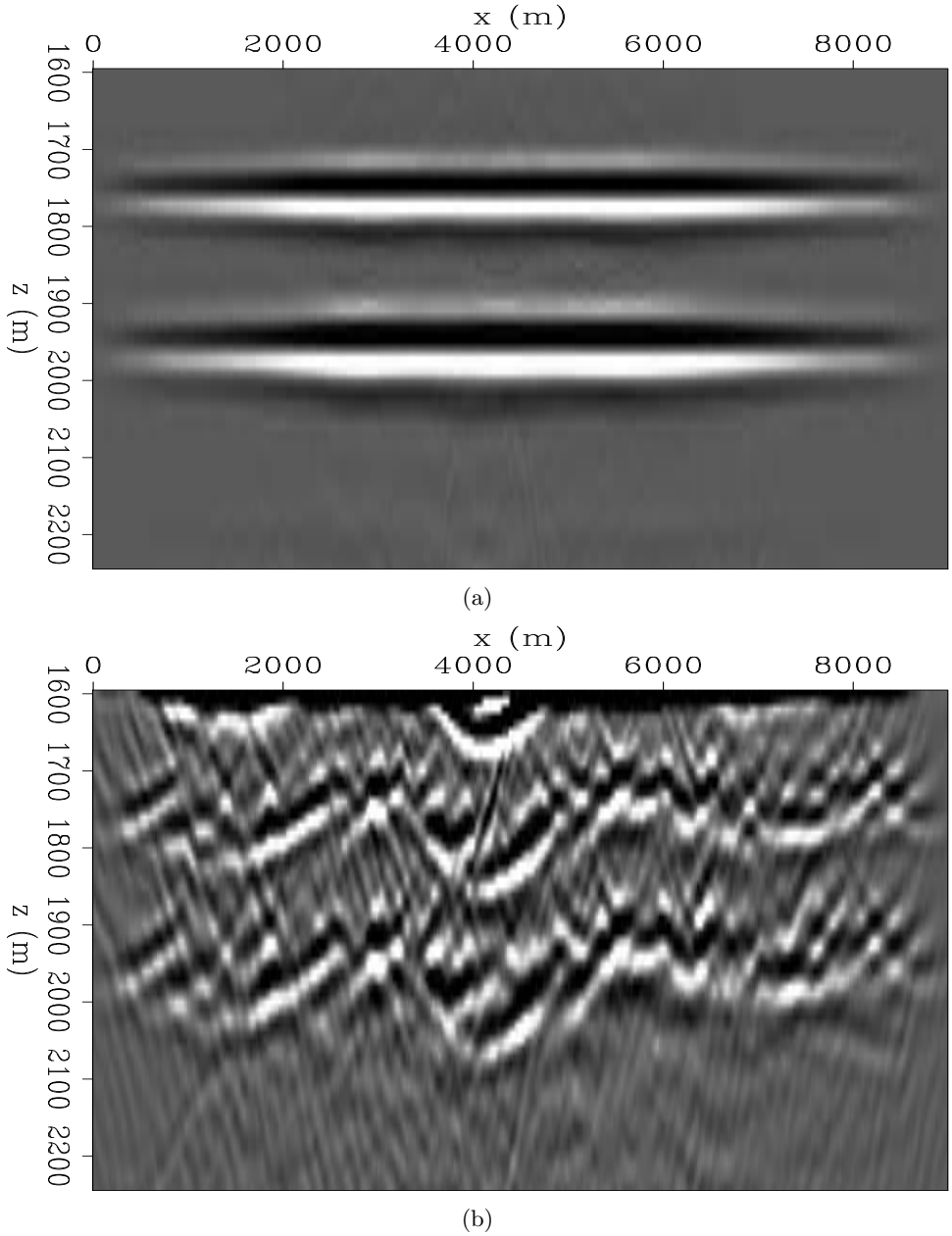


Figure 24: Migrated images of the velocity model in Figure23(a). (a) Image computed with the correct background velocity model. (b) Image computed with the inaccurate background velocity model. [**ER**] guillaume/. fbpsinimxc25,fbpsinimxb25

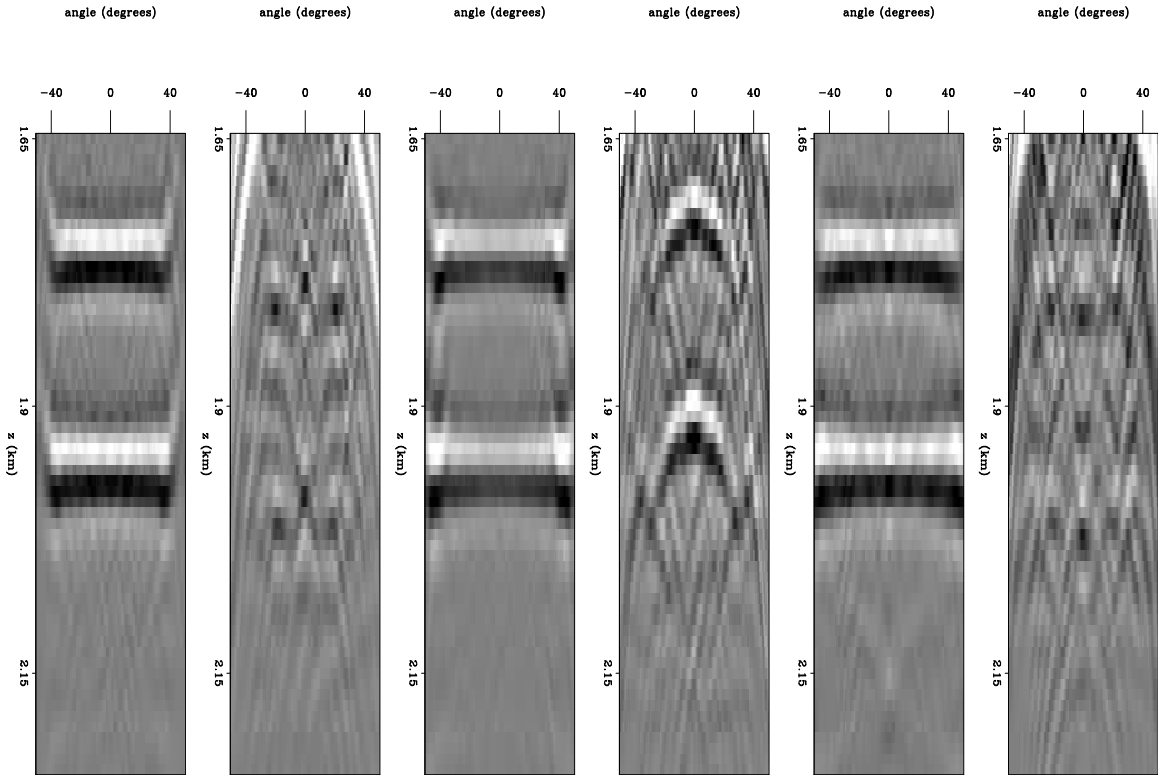


Figure 25: Three pairs of angle domain common image gathers. For each pair, the first image shows the gather computed from the image in Figure 24(a), and the second image shows the gather computed from the image in Figure 24(b). The gathers are computed at $x_1 = 1,730$ m, $x_2 = 2,580$ m, and $x_3 = 3,680$ m, respectively. [**ER**] guillaume/. fbpsinagather1

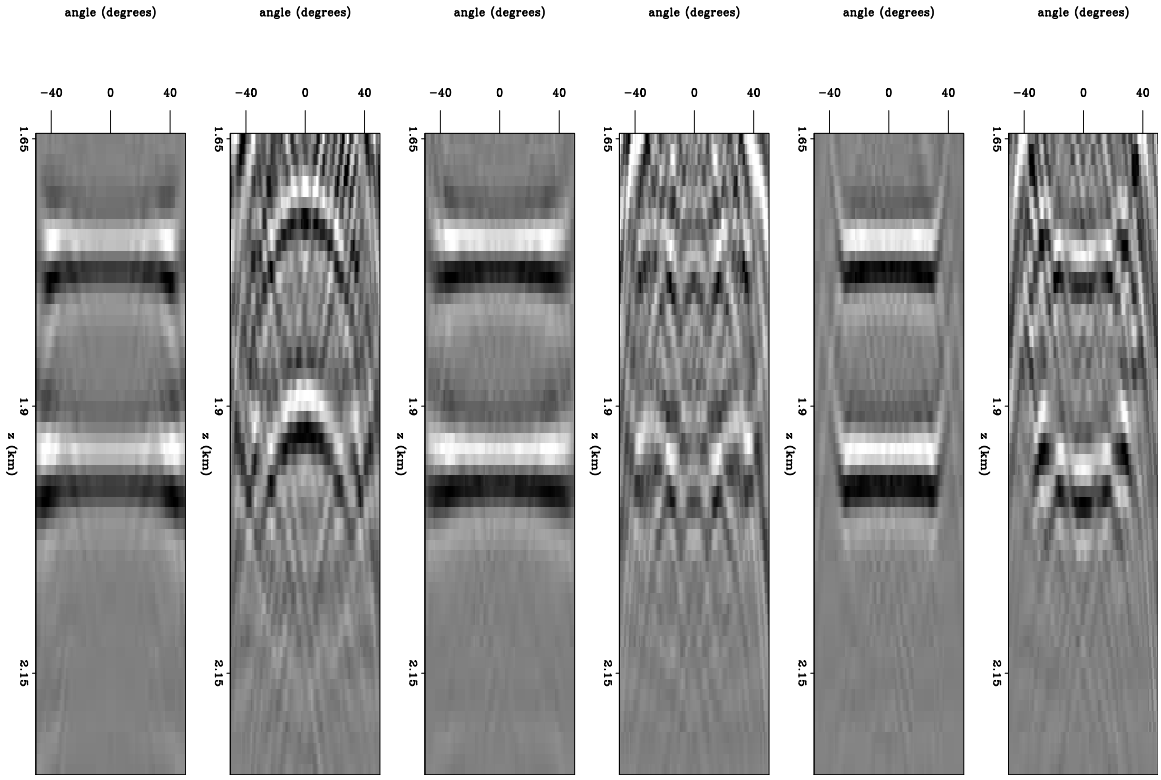


Figure 26: Three pairs of angle domain common image gathers. For each pair, the first image shows the gather computed from the image in Figure 24(a), and the second image shows the gather computed from the image in Figure 24(b). The gathers are computed at $x_4 = 5,980$ m, $x_4 = 6,6670$ m, and $x_5 = 7,550$ m, respectively. [**ER**] guillaume/. fbpsinagather2

CONCLUSIONS AND FUTURE WORK

We discussed why obtaining good-quality seismic images of zones that lie in the vicinity of salt bodies is becoming more and more crucial for the oil and gas industry. We summarized the main difficulties in doing so, and we chose to focus our research on the specific problem related to inaccurate top-salt boundary delineation. We argued that current techniques may still have trouble recovering a rugose sediment-salt interface with enough detail. The synthetic tests we performed confirmed that a slight misinterpretation of the top-salt deteriorates the images of subsalt reflectors to the point where almost no coherent energy is present, even by looking at angle domain common image gathers. This prevents us from successfully applying standard migration velocity analysis techniques. Therefore, there is a need for a new technique that can enable us to directly refocus subsalt images, or at least improve their coherency along an extended axis (e.g., time-lags, subsurface offsets, angle, etc.).

Moving forward, we would like to design an inversion scheme that allows us to bring back coherency along the extended axis of subsalt images, which would then allow us to successfully apply standard migration velocity analysis algorithms (e.g., TFWI). Considering our goal, we believe that an optimization scheme using an image-space objective function (e.g., WEMVA) could be appropriate. However, it can not rely on the presence of interpretable moveout in the extended image domain. A stack power maximization (SPM) scheme looks to offer the most potential. Though this implementation is prone to cycle-skipping, we could potentially circumvent this issue by taking a similar approach as the one done in Biondi and Almomin (2014), and introduce a linearization of the wave-equation based on the extension of the velocity model along the time-lag or subsurface offset axis.

ACKNOWLEDGMENTS

We would like to thank the Stanford Exploration Project affiliate companies for financial support. Guillaume Barnier would like to thank Ali Almomin, Gustavo Alves, Ettore Biondi, Musa Maharramov, Yi Shen, Joe Stefani, and Yang Zhang for their help and advice.

REFERENCES

Albertin, U., L. Zhang, et al., 2014, Migration optimization through local phase alignment of partial migration images: Presented at the 2014 SEG Annual Meeting.

Anselmetti, F. S. and G. P. Eberli, 1993, Controls on sonic velocity in carbonates: Pure and Applied Geophysics, **141**, 287–323.

Biondi, B. and A. Almomin, 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: Geophysics, **79**, WA129–WA140.

DNR, 2015, Reservoir traps: Louisiana Department of Natural Resources, http://dnr.louisiana.gov/assets/TAD/education/BGBB/4/traps.html (accessed April 2015).

Dribus, J., M. Jackson, J. Kapoor, and M. Smith, 2008, The prize beneath the salt: Oilfield Review, **15**, 4–17.

Etgen, J. T., I. Ahmed, M. Zhou, et al., 2014a, Seismic adaptive optics: Presented at the 2014 SEG Annual Meeting.

Etgen, J. T., C. Chu, T. Yang, M. Vyas, et al., 2014b, Adaptive image focusing: Presented at the 2014 SEG Annual Meeting.

Farmer, P., D. Miller, A. Pieprzak, J. Rutledge, and R. Woods, 1996, Exploring the subsalt: Oilfield Review, **8**, 50.

Halpert, A. D., 2014, Interpreter-driven automatic image segmentation and model evaluation: PhD thesis, Stanford University.

Hudec, M. R. and M. P. Jackson, 2007, Terra infirma: [u]nderstanding salt tectonics: Earth-Science Reviews, **82**, 1–28.

Jones, I. F. and I. Davison, 2014, Seismic imaging in and around salt bodies: Interpretation, **2**, SL1–SL20.

King, H., 2015, Salt domes: Geoscience News and Information, http://geology.com/stories/13/salt–domes/ (accessed April 2015).

Leveille, J. P., I. F. Jones, Z.-Z. Zhou, B. Wang, and F. Liu, 2011, Subsalt imaging for exploration, production, and development: A review: Geophysics, **76**, WB3–WB20.

Levin, H. L., 2009, The earth through time: John Wiley & Sons.

Ritter, G. et al., 2010, Interpretation driven velocity model building to improve subsalt imaging: Presented at the 2010 SEG Annual Meeting.

Schlumberger, 2015, Subsalt: Schlumberger Oilfield Glossary, http://www.glossary.oilfield.slb.com/en/Terms.aspx?LookIn=term=subsalt (accessed April 2015).

Inverse demigration for simultaneous source separation

Chris Leader and Biondo Biondi

ABSTRACT

Separating simultaneously acquired seismic data is the link between more efficient acquisition and conventional imaging techniques. Existing methods for separation rely on coherency measurements and work only for randomly delayed sources. By using the extended image space, data blended with a variety of time delay sequences can be well separated. Demigration will then output the separated, conventionally equivalent, dataset. A single pass of demigration can adequately recreate kinematic information. For amplitude balancing, however, an inverse process is required. By introducing a blending operator into a linearised inverse system posed in the extended image space, accurate data separation is observed for a variety of blending schemes, without the need for an accurate velocity model. Furthermore, such a system can be easily adapted to include velocity model updates, since extended imaging has already been applied.

INTRODUCTION

Contemporary seismic targets are increasingly associated with steeply dipping structures and strong velocity contrasts. In order to illuminate these difficult features, data are acquired with large offsets and multiple source boats (Verwest and Lin, 2007). Intuitively, this leads to both more expensive acquisition and an increase in field waiting time. This latter adverse consequence is due to the fact that it is necessary to allow the energy from the previous source to sufficiently dissipate before recording the next source point. If waiting time was not a restriction then denser sampling could be recorded per unit time and acquisition would be significantly more efficient (Beasley (2008); Hampson et al. (2008); Berkhout and Blacquiere (2008)). Practically, it is possible to disregard this waiting time and fire the next shot when in position; this is often called continuous recording. Recording overlapping data in this manner requires more processing time than conventionally acquired data, since separation is necessary to mitigate imaging artifacts. However, the economic gains from reduced acquisition time far outweigh this extra processing cost.

These simultaneously acquired data can be used to directly invert for model properties (Dai and Schuster (2009); Tang and Biondi (2009)). However, such methods require exact knowledge of the velocity model. Separation and subsequent imaging is a viable option, since this could be integrated into production data flows. Successful existing methods rely on random sampling in the source timings and locations (Abma and Yan (2009); Moore et al. (2008)). For example, constant receiver gathers can be transformed into the f-k or tau-p domain and iteratively thresholded (Doulgeris et al., 2011), removed in the parabolic random domain (Ayeni et al., 2011), removed by using a convex projection approach (Abma et al., 2010), or through compressive sensing methods (Herrmann et al., 2009).

Image domain processing has been used effectively for coherent energy removal / attenutation by posing the problem in the extended image space (Zhang and Duan (2013); Sava

and Guitton (2005)). It is possible to untangle certain events in this domain and recreate cleaner shot gathers by virtue of higher signal-to-noise ratio and reduced dimensionality. Additionally, when using the extended image space (Sava and Vasconcelos, 2011), event kinematics are preserved. Consequently, if the velocity model is inaccurate then demigration is still possible (Chauris and Benjemaa, 2010).

In the previous SEP report (Leader and Biondi, 2014), methods of distinguishing events in the angle domain were analysed, with the goal of using curvature-based penalties during demigration. Whilst this method worked very well for simple scattering models, for highly heterogeneous structures the ranges of curvature necessary for demigration were too high. Additionally, describing the curvature using a single parameter became less possible. In this study a mass inverse demigration scheme will be postulated. By introducing a blending operator and posing the problem as a linearised inversion, accurate separation (in terms of both amplitudes and kinematics) is observed after a small number of iterations.

A range of model complexities and velocity inaccuracies will be analysed to test the strength of this methodology. Additionally, three blending strategies will also be tested - constant time delays, pseudo-constant time delays and purely random time delays.

DEMIGRATION

To accurately recreate all wavefield complexities, Reverse Time Migration (RTM) (Baysal et al., 1983) is the choice of imaging operator. RTM uses solutions to the full two-way wave equation (within physical assumptions), making it a valuable option for highly heterogeneous Earth models. Furthermore, RTM is the adjoint of Born modeling, and it is possible to move between the data and image spaces by combining these operators. For zero-offset imaging, these are summarised in equation 1 and equation 2. Examples of a Born modelled shot gather and an RTM image are shown in Figure 1 and Figure 2 respectively.

$$m(\mathbf{x}) = \sum_{\mathbf{x}_s,\omega} f(\omega) G_0(\mathbf{x}, \mathbf{x}_s, \omega) \sum_{\mathbf{x}_r} G_0(\mathbf{x}, \mathbf{x}_r, \omega) d^*(\mathbf{x}_r, \mathbf{x}_s, \omega), \tag{1}$$

$$d(\mathbf{x}_r, \mathbf{x}_s, \omega) = \omega^2 \sum_{\mathbf{x}} f(\omega) G_0(\mathbf{x}, \mathbf{x}_s, \omega) m(\mathbf{x}) \sum_{\mathbf{x}} G_0(\mathbf{x}, \mathbf{x}_r, \omega). \tag{2}$$

here $\mathbf{x}$ represents the spatial coordinates, $m(\mathbf{x})$ the scattering field, $\mathbf{x}_s$ the current source coordinates, $\mathbf{x}_r$ the current receiver coordinates, ω the temporal frequency, $d^*(\mathbf{x}_r, \mathbf{x}_s, \omega)$ the complex conjugate of the data and G_0 the relevant Green's function. The aforementioned zero-offset image (Claerbout, 1971) is calculated here. For an accurate velocity model this will contain all necessary amplitude and kinematic information for demigration and hence data recovery.

However, for the problem of separating continuously acquired data a stringent requirement on the velocity model is undesirable. Direct application of equation 1 with an incorrect velocity model will result in the loss of certain events, and subsequent demigration will not represent the original dataset well. To preserve all event kinematics, extended imaging must be used. If zero-offset imaging can be described by equation 3, then extended imaging can be described by equation 4.

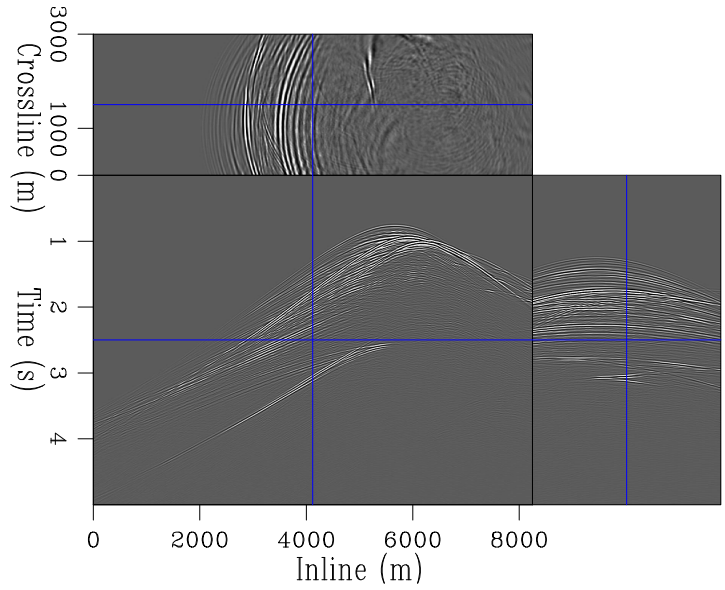


Figure 1: An example of a 3D shot gather produced from Born modeling. [**NR**] chris1/. seamdata3d

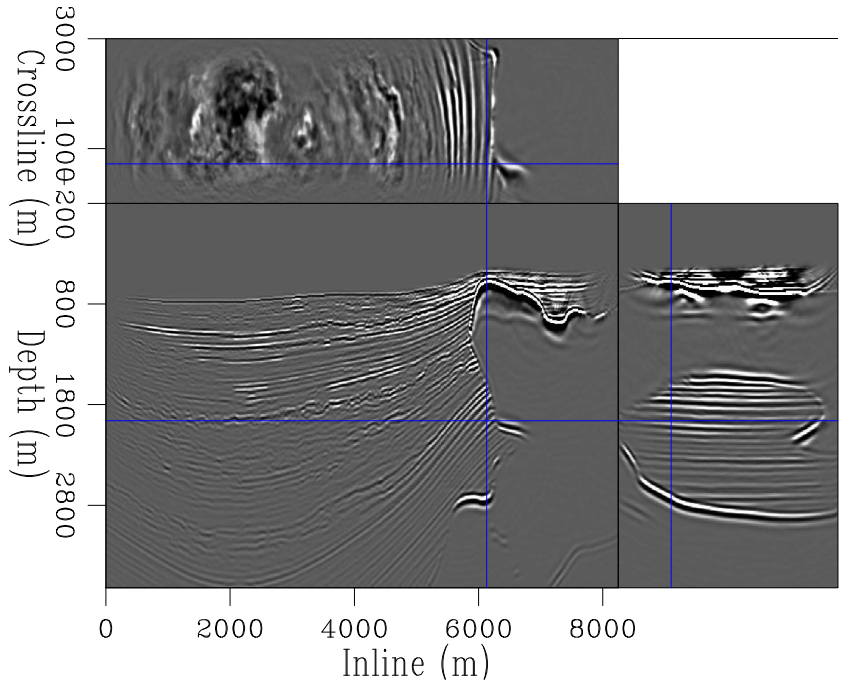


Figure 2: An example image produced from RTM over the SEAM velocity model, in 3D. [**NR**] chris1/. seamimg

$$I(x,y,z) = \sum_{i}^{nshots} \sum_{t} P_s(x,y,z,t;\mathbf{s}_i) P_r(x,y,z,t;\mathbf{s}_i), \tag{3}$$

$$I(x,y,z,x_h,y_h) = \sum_{i}^{nshots} \sum_{t} P_s(x+x_h,y+y_h,z,t;\mathbf{s}_i) * P_r(x-x_h,y-y_h,z,t;\mathbf{s}_i). \tag{4}$$

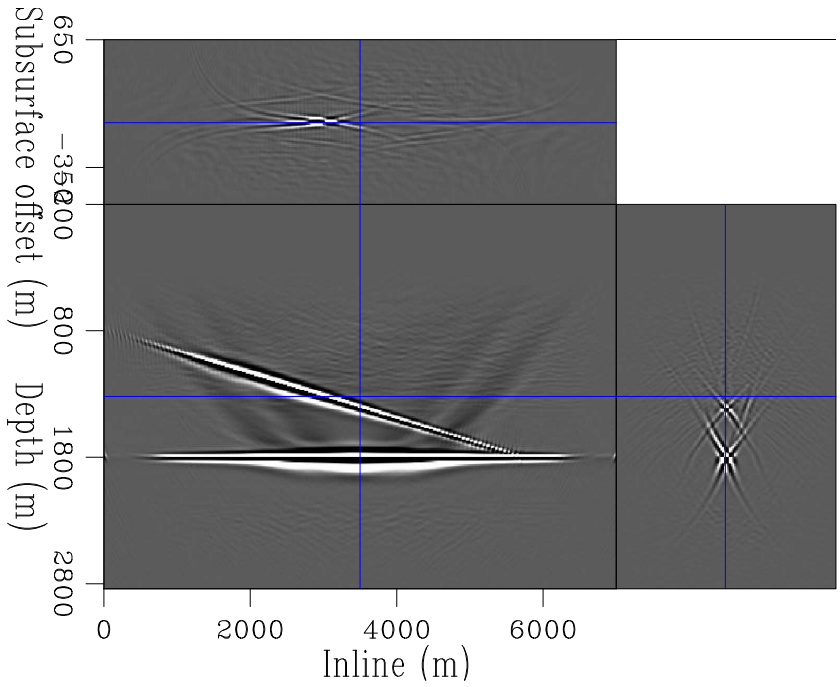


Figure 3: An extended image produced using an accurate velocity model. [**CR**] chris1/. extsimpimgcv

Here, $I(x,y,z)$ is the image in space, P_s is the source wavefield and P_r is the receiver wavefield; s_i represents the current shot of interest. If lag coordinates (known as subsurface offsets) in x and y are introduced (x_h and y_h), a 5D image can be created. It is possible to have lags in both t and z to create a 7D image, or any combination thereof. From hereon this discussion will be limited to subsurface offsets in the x direction only.

If the correct velocity model were used for imaging then the energy will be focused to a point in subsurface offset, as seen in Figure 3. If an incorrect model were used then the energy will be spread out over a range of offsets (Figure 4). Analysing this moveout as a function of the velocity model is the core concept of Wave Equation Migration Velocity Analysis (WEMVA) (Sava et al., 2003).

To create Figure 2, equation 3 was used, and for Figure 3, equation 4 was used. By taking the adjoint of these processes and creating an extended Born modeling operator, demigration can be performed. The result from demigrating Figure 3 from the extended

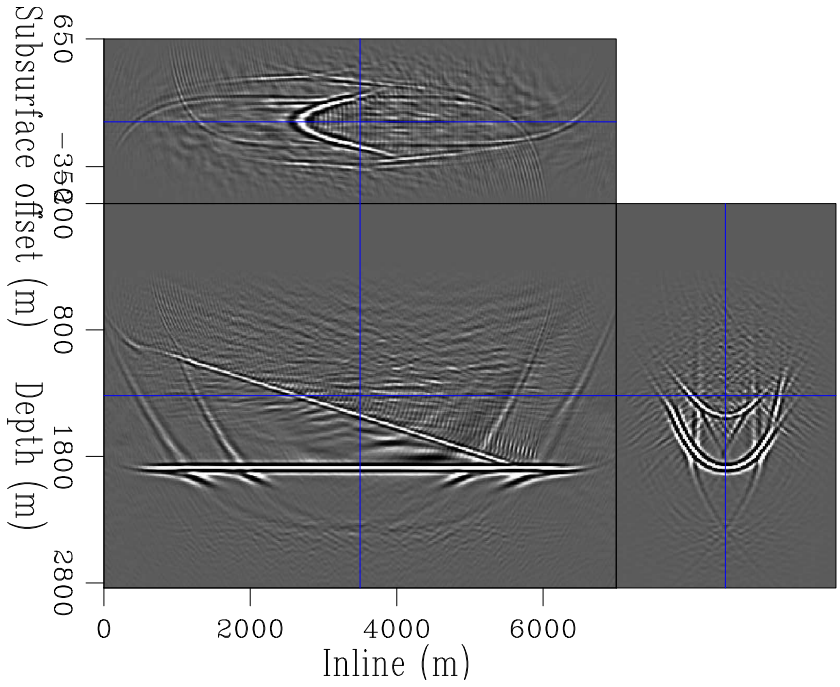


Figure 4: An extended image produced using an inaccurate velocity model. [**CR**] chris1/. extsimpimgincv

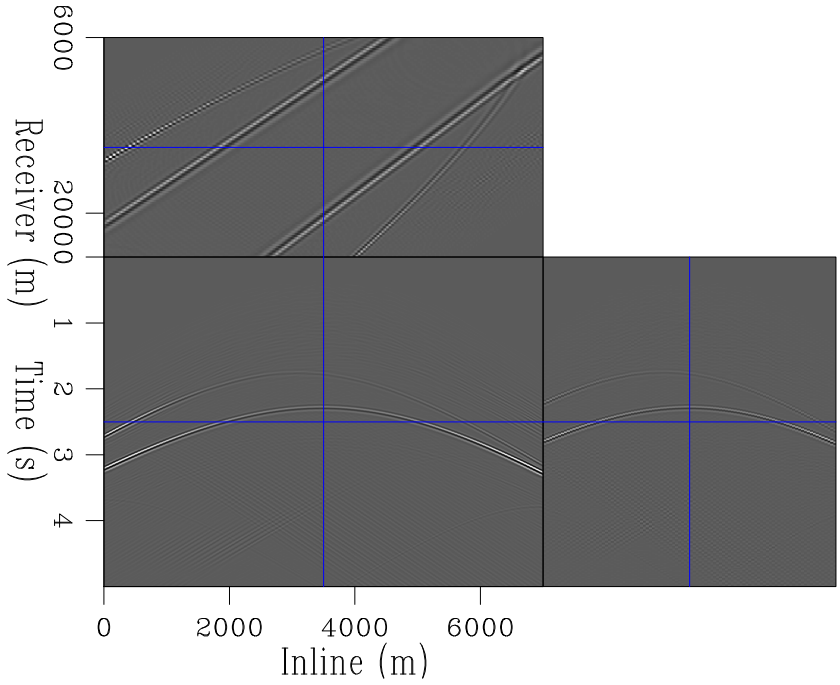


Figure 5: Adjoint demigration using the correct velocity model, in both directions. [**CR**] chris1/. simpdemcv

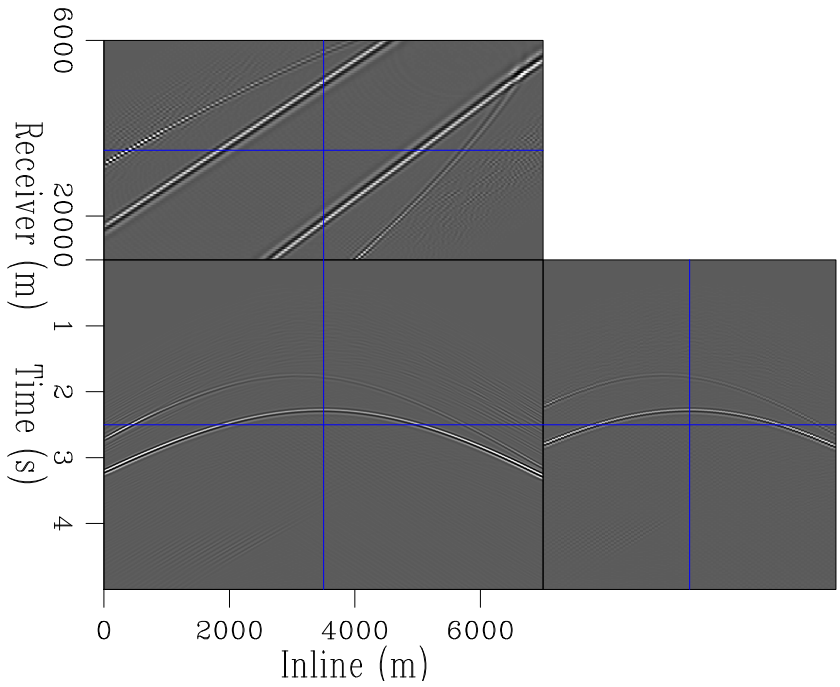


Figure 6: Adjoint demigration using an incorrect velocity model, in both directions. [**CR**] chris1/. simpdemincv

image domain to the data space is shown in Figure 5. Similarly, the result from demigrating Figure 4 is shown in Figure 6.

It is apparent that both passes of demigration have successfully recreated many of the data nuances - event kinematics are correctly positioned and all identifiable events in the input data are present in the demigrated data. However, frequency content has not been well preserved and event amplitudes at early times and short offsets are under represented. These imperfections are more pronounced when using the incorrect velocity model. Additionally, this result is more artifact laden. This is a consequence of performing extended Born modeling, where the image convolution is shifted over a range of offsets.

INVERSE DEMIGRATION

To recover the correct amplitudes, adapting demigration to become an inverse process is necessary. In conventional Least-Squares RTM (LSRTM), RTM is used as the adjoint procedure, and Born modeling as the forward operator. A solver, such as conjugate directions, can then be used for model and residual updates. For inverse demigration almost exactly the same procedure can be used, but with the forward and adjoint operators swapped. The input will now be the extended image and the procedure will aim to recover the dataset that best represents the given image.

Figure 7 and Figure 8 show the same recovered data as Figure 5 and Figure 6 respectively, after ten iterations of inverse demigration. Amplitudes are now consistently balanced and match the input data Furthermore, the vast majority of the artifacts from the incorrect

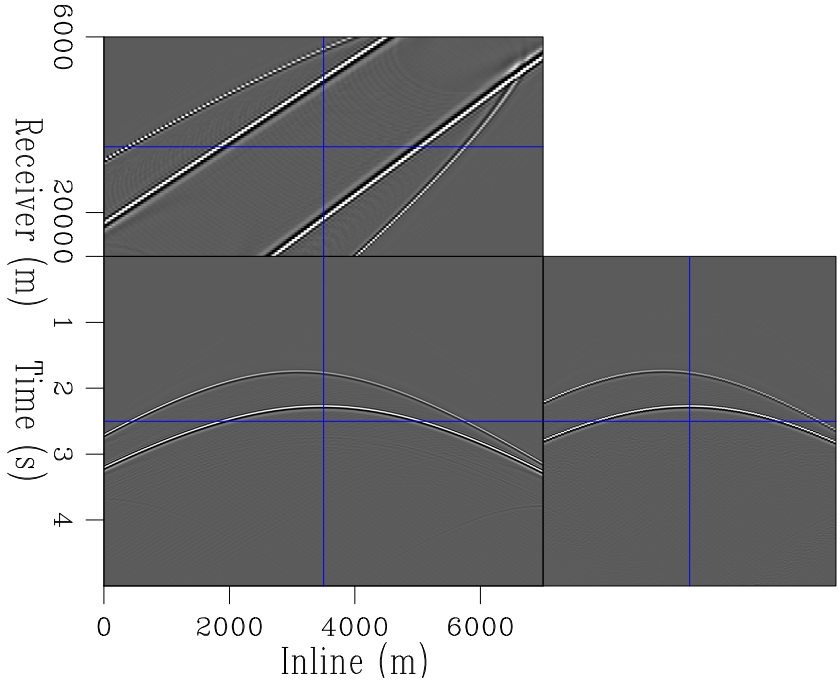


Figure 7: Inverse demigration, after ten iterations, using the correct velocity model. [**CR**]
chris1/. simpinvdemcv

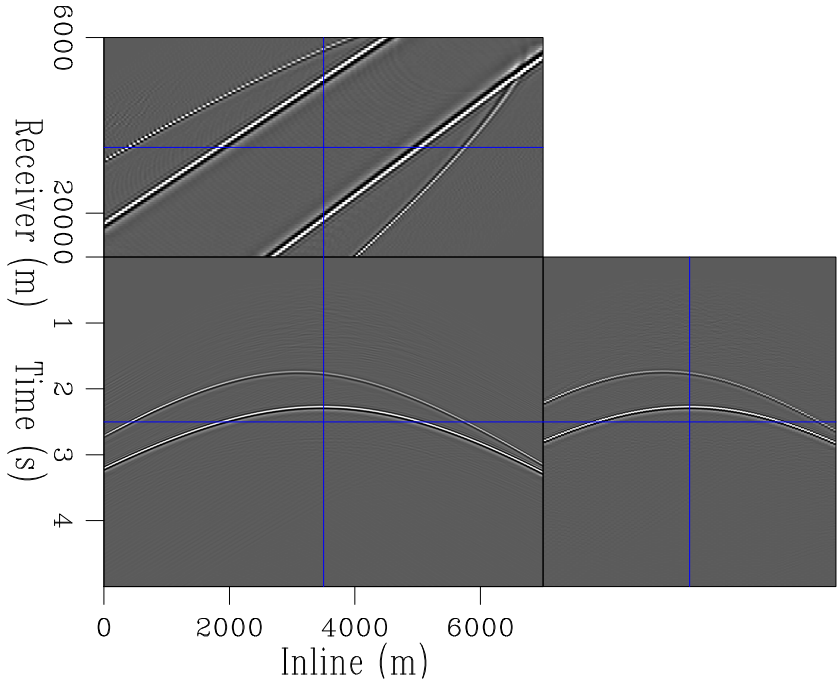


Figure 8: Inverse demigration, after ten iterations, using an incorrect velocity model. [**CR**]
chris1/. simpinvdemincv

velocity result are mitigated. The data-space residual, as a function of iteration number, can be seen in Figure 9.

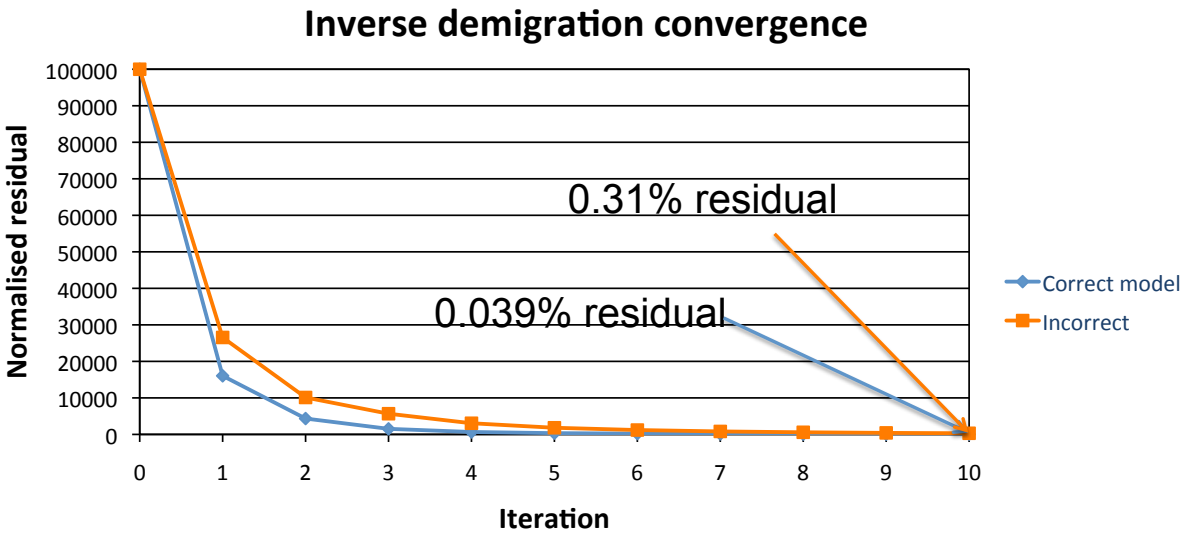


Figure 9: Data-space residual as a function of iteration number, for both correct and incorrect velocity model inverse demigration. [**NR**] chris1/. simpinvconv

SIMULTANEOUS SHOT SEPARATION

Now that inverse demigration has been demonstrated to accurately recover input reflection data, the system can be adapted to provide shot separation. The input data will be the overlapping, simultaneously recorded dataset, so an additional operator must be included to account for the blending. This operator, Γ, will be referred to as the 'blending' operator, for obvious reasons. The forward process of Γ will take a conventional dataset, then delay and sum the shots together to produce a continuous record. The adjoint process will take a continuous record and window the shots (according to an input recording length and sequence of shot delays) and output a discrete dataset. Thus the forward blending operator requires three inputs - the data, the desired recording length, and a record of shot delays. If the blended data is d_b, and the conventional / separated data is d_s, then the system can be described in equation 5.

$$d_b = \Gamma d_s \tag{5}$$

$$J(\mathbf{d}_s) = \|\mathbf{L}'\mathbf{d}_s - \mathbf{m}\| \tag{6}$$
$$J(\mathbf{d}_s) = \|\mathbf{L}'\mathbf{d}_s - \mathbf{L}'\mathbf{\Gamma}'\mathbf{d}_b\| \tag{7}$$

The objective functions for conventional and blended demigration are shown in equation 6. This operator Γ is necessary to calculate the data-space residual (actually the

model-space residual, in this case), since a recomputed blended dataset must be used for comparison. For shot separation, the output data will be the recovered data produced after demigration, before it is reblended for comparison.

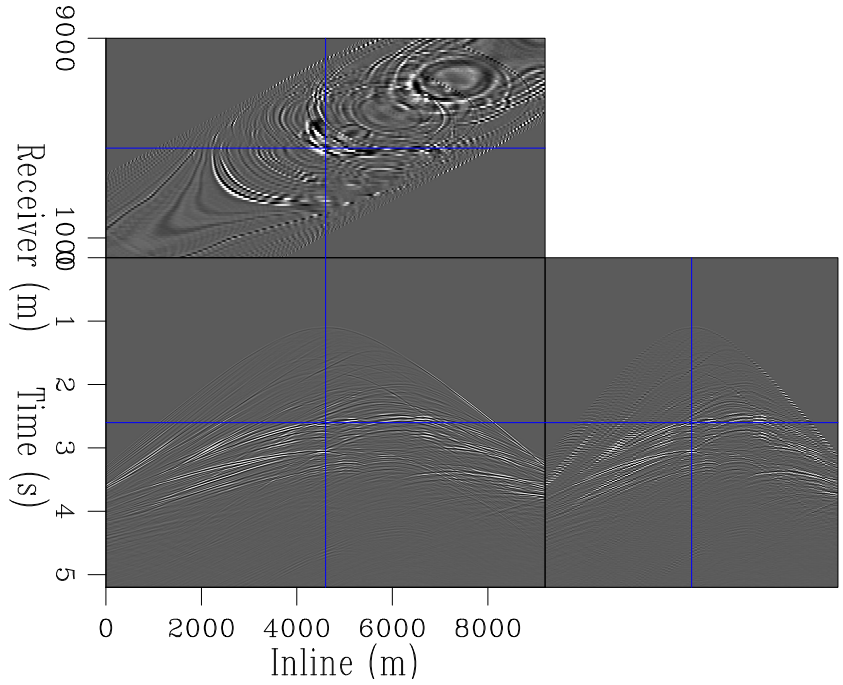


Figure 10: Data acquired over the Marmousi model. [**CR**] chris1/. marmdata

Three styles of data blending are studied, all with an Ocean Bottom Node (OBN) style geometry. These are: purely random time delays, constant (or linear) time delays, and pseudo-linear delays. These delays are linear in both time and space. For the last style, delays are roughly constant, with a 5% jitter added to the source timings. Example datasets for these three encoding functions are shown in Figure 11, Figure 12 and Figure 13, with the conventional data shown in Figure 10. For these upcoming examples 2D data were used with a fixed receiver geometry. This means the right-hand panel in these figures represents a constant receiver gather, whereas the left panel represents a constant source gather.

For the randomly delayed sequence, all secondary/overlapping energy is incoherent in the receiver domain. This is exploited by all of the aforementioned data-space inversion methods. For the pseudo-linear case, some coherency is induced in these receiver gathers, but the randomness is sufficient such that this may stack out. For the linear case there is no difference in coherency between these domains, so these data-space methods will all fail.

Correct velocity model solution

Figure 14, Figure 15 and Figure 16 show the extended images produced from these three datasets using the correct velocity. The differences in blending are manifested in the image space, although the coherency differences are not as pronounced as intuition may suggest. Even the linearly blended data becomes well dispersed in the image space. The artifacts are more coherent, but the focusing characteristics of the primary events, and the differences

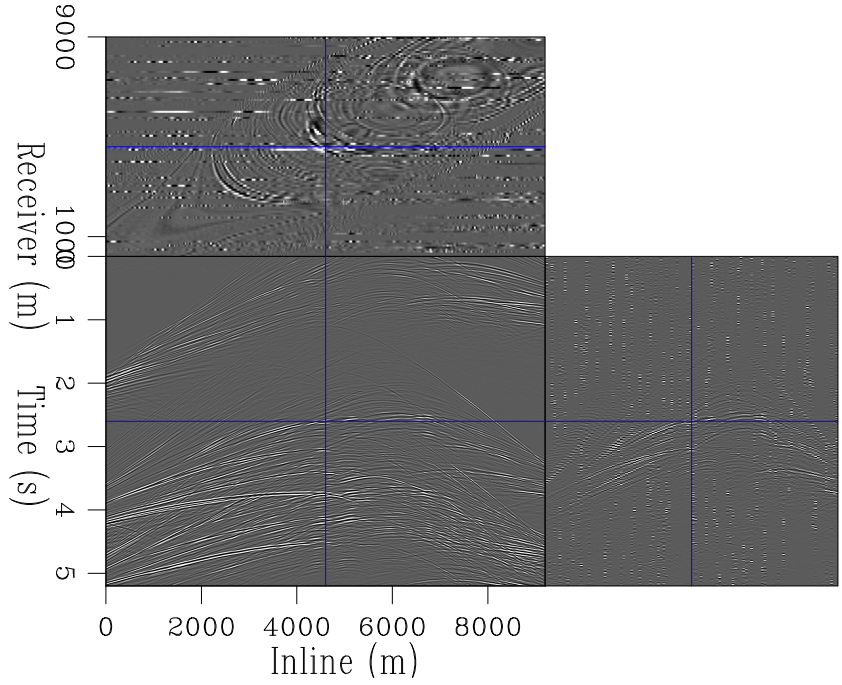


Figure 11: Data acquired over the Marmousi model using a random delay function. [**CR**]
chris1/. marmrndblnd

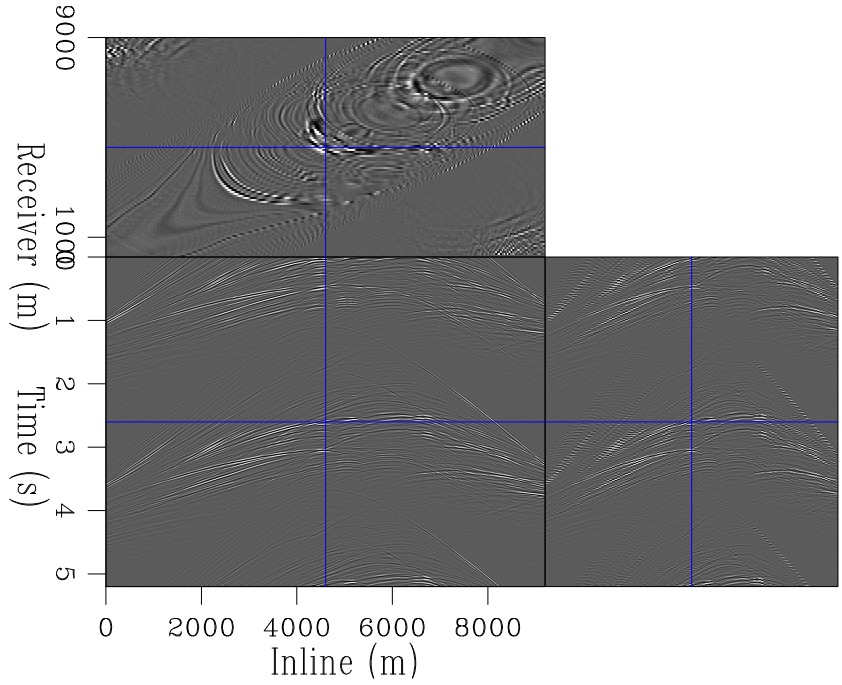


Figure 12: Data acquired over the Marmousi model using a linear delay function. [**CR**]
chris1/. marmlinblnd

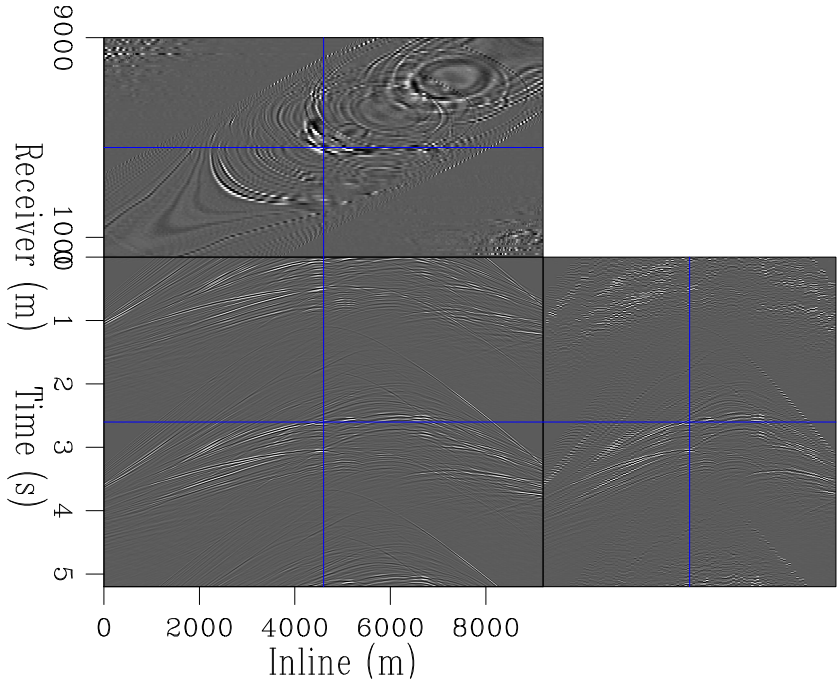


Figure 13: Data acquired over the Marmousi model using a pseudo-linear delay function. [**CR**] chris1/. marmplinblnd

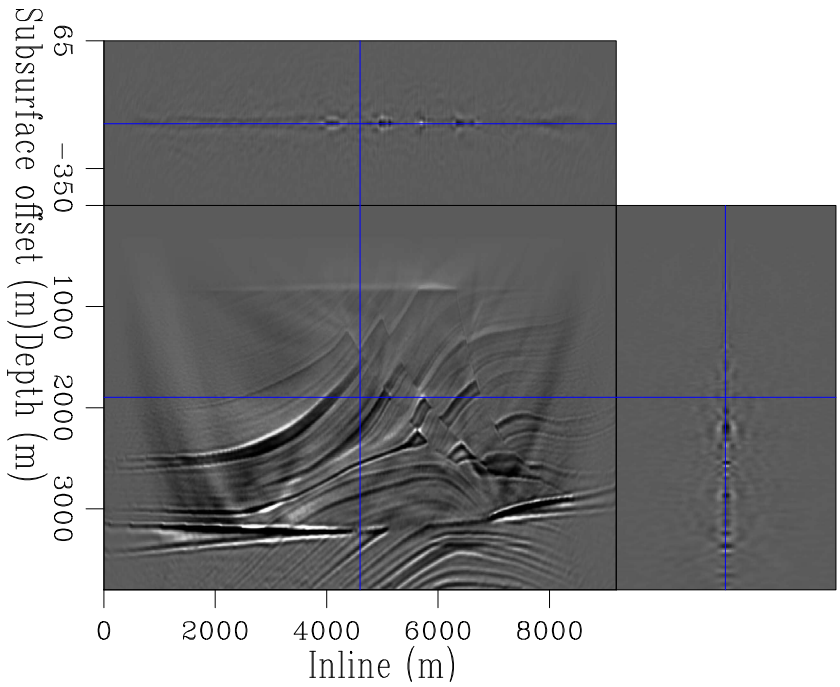


Figure 14: The extended image from migrating Figure 11 (data acquired using a random blending function) [**CR**] chris1/. marmrndblndim

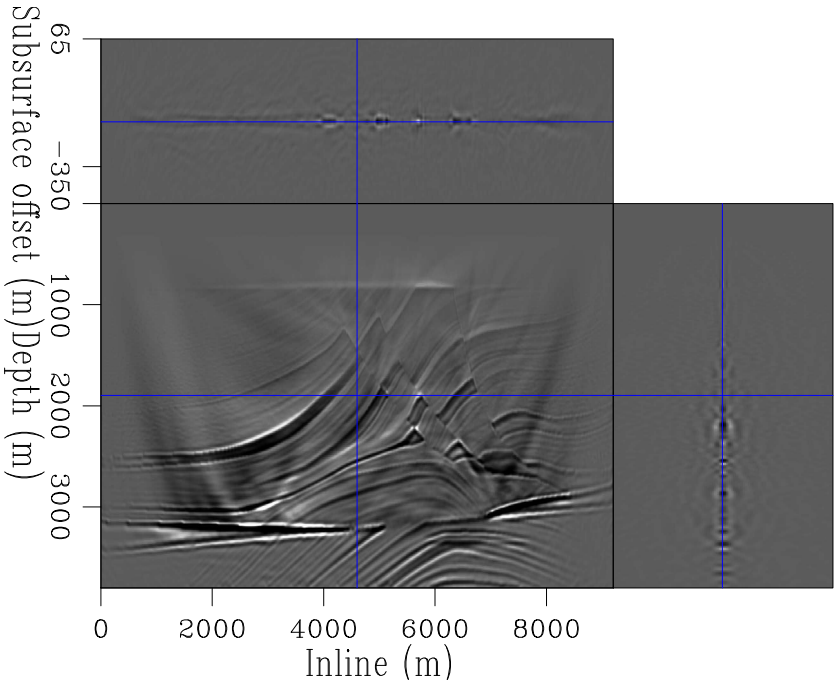


Figure 15: The extended image from migrating Figure 12 (data acquired using constant time delays). [**CR**] chris1/. marmlinblndim

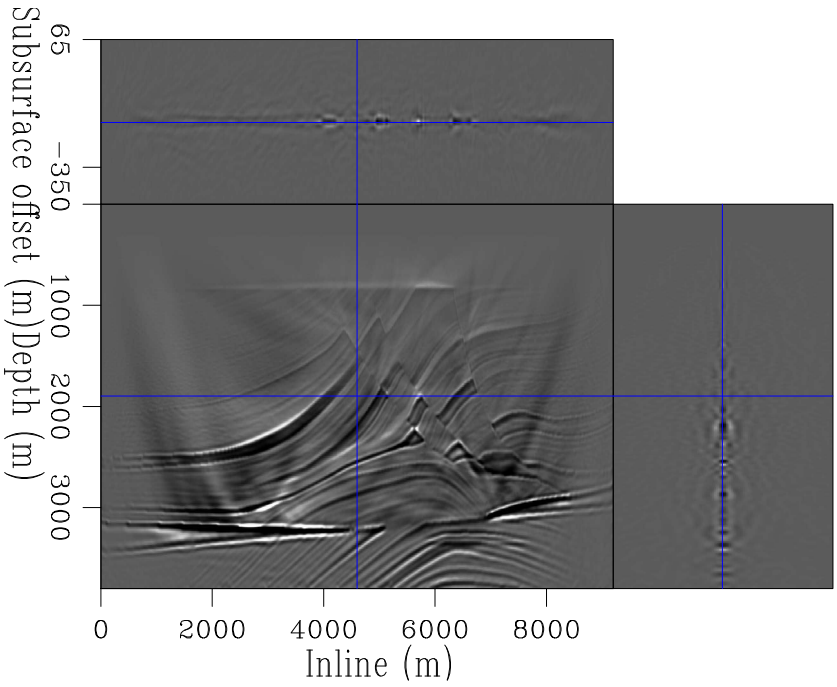


Figure 16: The extended image from migrating Figure 13 (data acquired using pseudo-linear delays). [**CR**] chris1/. marmplinblndim

in contrast, suggest that separation could be possible.

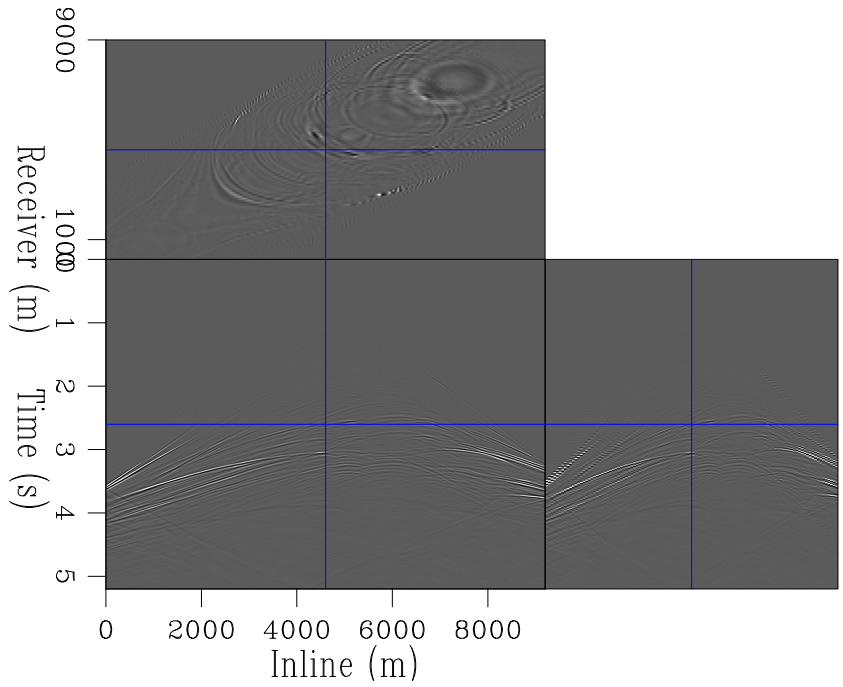


Figure 17: The output dataset after applying adjoint demigration to Figure 14, which was the image created from a randomly delayed dataset. [**CR**] chris1/. marmrndblndidem

The recovered datasets after adjoint separation can be seen in Figure 17, Figure 18 and Figure 19. Each of the blending schemes have been well separated and the resultant datasets could be used for conventional velocity estimation and imaging. Again, amplitudes at early times and shot offsets are weaker than in reality. This can be improved upon by using the inverse scheme.

It should be noted that there are some fractionally more coherent artifacts in the recovered data from the linear encoding. Nonetheless this methodology separated these linearly delayed data surprisingly easily.

Incorrect velocity model

Separation with an accurate velocity model is a relatively trivial problem. If simultaneous surveys are to be used for exploration then a processing scheme that does not assume strong velocity control is essential. This section will look at these same data, but imaged using a very inaccurate velocity model. The most informative and realistic scenario is pseudo-linear blending, so this will be the acquisition focus from hereon.

Figure 21 shows the result from migrating these data, but using a very rough velocity model. The primary events are now not well focused at zero-subsurface offset, and the focusing contrasts between primary and secondary events is far less. Energy of interest now spans many of the acquired subsurface offsets and distinguishing events from primary and secondary energy is less obvious.

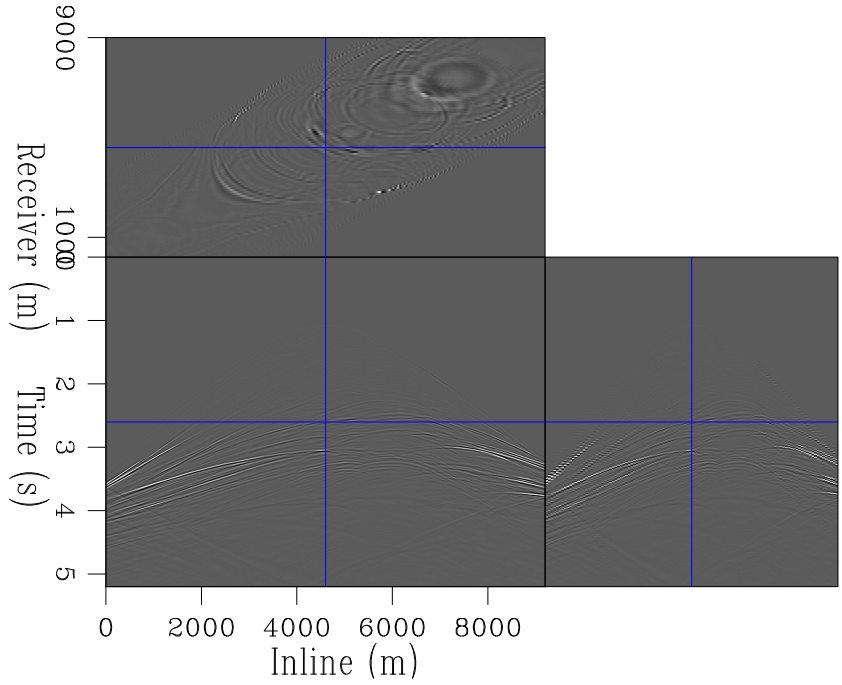


Figure 18: The output dataset after applying adjoint demigration to Figure 15, which was the image created from a linearly delayed dataset. [**CR**] chris1/. marmlinblndidem

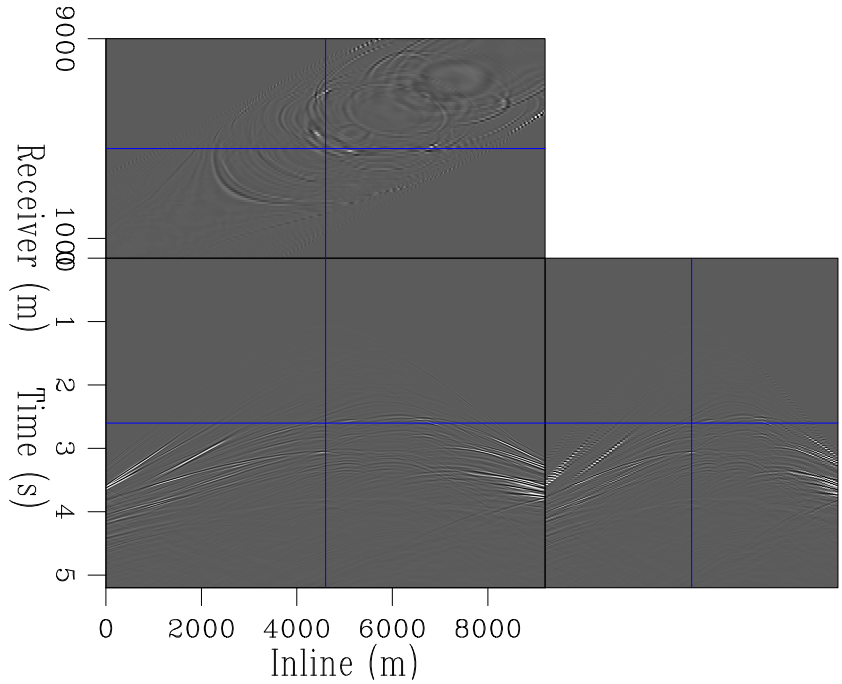


Figure 19: The output dataset after applying adjoint demigration to Figure 16, which was the image created from a pseudo-linearly delayed dataset. [**CR**] chris1/. marmplinblndidem

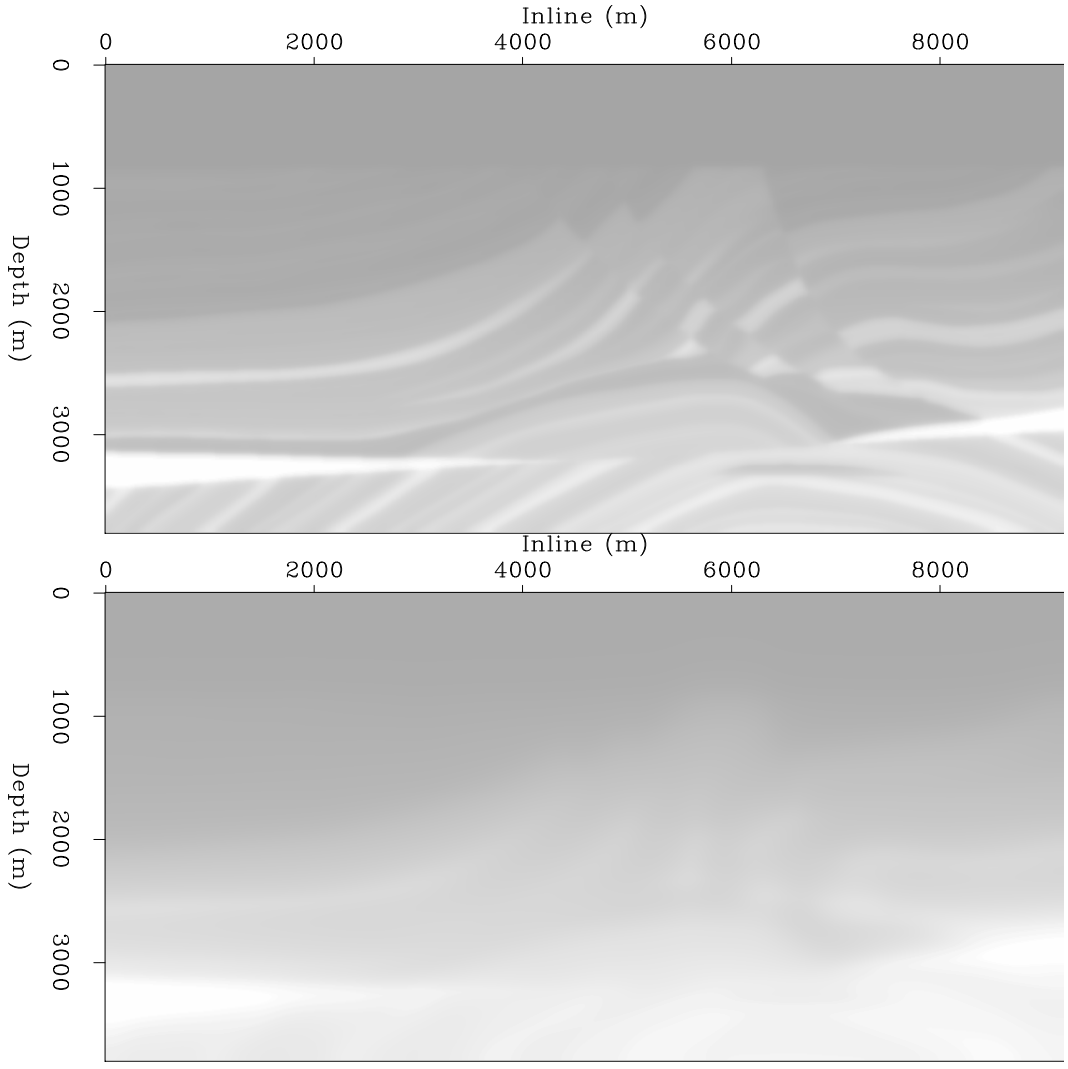


Figure 20: The velocity models used for inverse demigration, correct and incorrect respectively. [**ER**] chris1/. marmvels

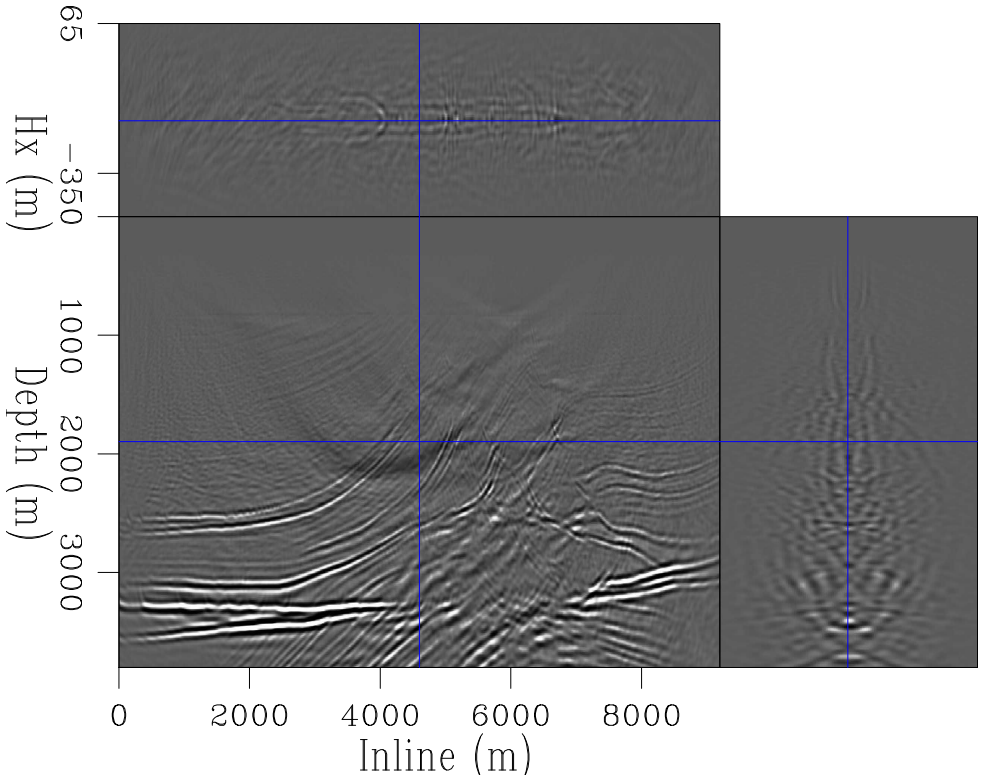


Figure 21: The extended image from migrating Figure 13 (data with constant delays) using a rough velocity model. [**CR**] chris1/. marmplinblndiminc

Using data acquired over the Marmousi model gives informative results about the suggested procedure, but to further confidence over this methodology a more difficult example must be used. A section of the SEAM model was windowed, featuring rugose reflectors, continuous reflectors and a steep salt body with carbonate top and sedimentary inclusions. These attributes comprise many of the difficulties of contemporary imaging targets. If this separation method can perform well over these data then much stronger conclusions can be made.

SEAM data

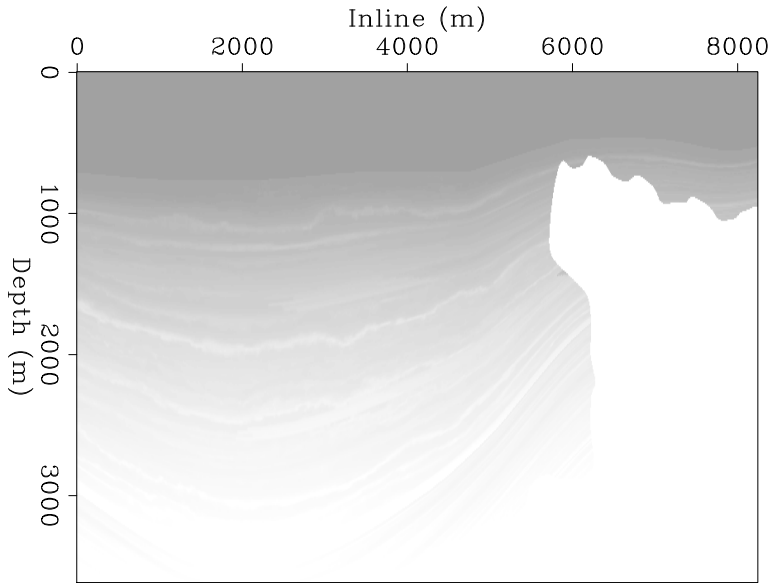


Figure 22: The velocity model windowed from the SEAM model, used for a more realistic separation test. [**ER**] chris1/. seamvel

Data were simulated over the velocity model shown in Figure 22. For this section only a pseudo-linear scheme will be studied, as this is the most realistic encoding that is recorded in the field. One hundred shots were simulated and then combined using a scheme with a blending power of three (similar to three source boats) with constant delays and a 5% timing randomness induced. These data are shown in Figure 24, while the output data after 10 iterations are shown in Figure 25.

These results demonstrate that for a complex model with a variety of impedance contrasts, the separation procedure performs very well. The convergence curve in Figure 26 demonstrates how accurately these data are simulated as a function of iteration number.

AUGMENTATION WITH WEMVA

Extended LSRTM is a very expensive process, and for many geometries and encoding functions this will be more expensive than many of the existing data-space methods. However, there are three circumstances where the suggested inverse demigration process can outperform these algorithms.

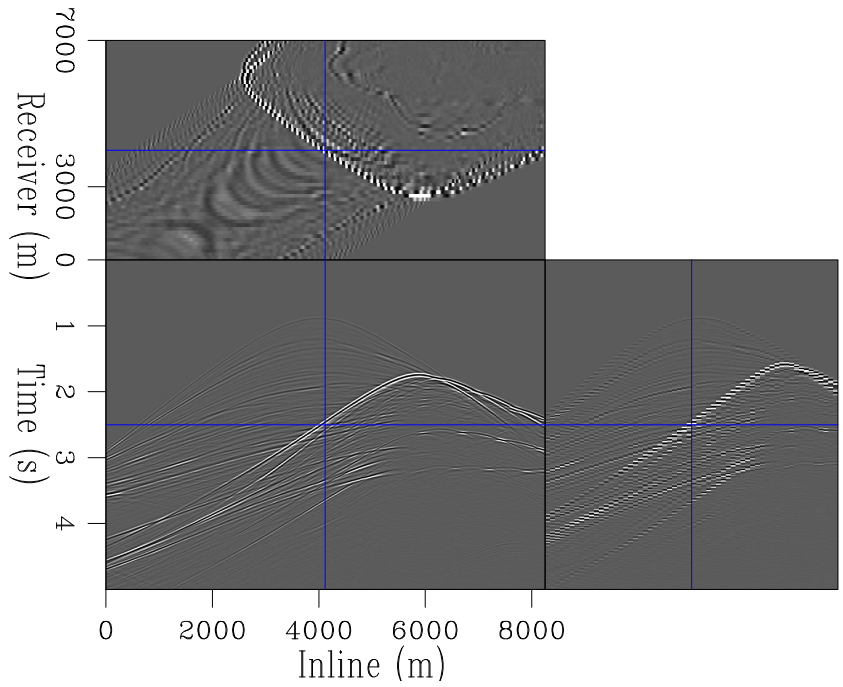


Figure 23: A conventional dataset acquired using the section of the SEAM model. [**CR**] chris1/. seamdata

The most successful of the data-space methods requires hundreds of 5D FFTs to be performed, so for geometries with sparse, large offsets, this will be a tough undertaking. Furthermore, this method entirely fails with constant time delays, and can induce large artifacts if there is any underlying predictability in the encoding sequence. So, under these two situations it is likely that extended inverse demigration will outperform a Projection Onto Convez Sets (POCS) type approach.

Thirdly, since extended imaging and modeling are applied, other approaches can take advantage of these spaces. Wave-Equation Migration Velocity Analysis (WEMVA) (Sava et al., 2003) relies on measuring these primary moveouts and using this back-projected information to update the velocity model. With this in mind, image-space shot separation could be easily augmented with WEMVA, where separation is the outer-loop and WEMVA the inner-loop process. This will also provide a positive feedback loop - the more accurate the velocity model has become, the better the separation results are, and fewer subsurface offsets must be collected.

The fact that this separation can be combined with WEMVA makes a powerful argument for the use of such an expensive scheme.

CONCLUSIONS

It is possible to accurately recover input seismic data from the image domain by adapting LSRTM. By using the extended image space, even if imaging was applied using an incorrect velocity model, these data are still recoverable to a high degree of accuracy.

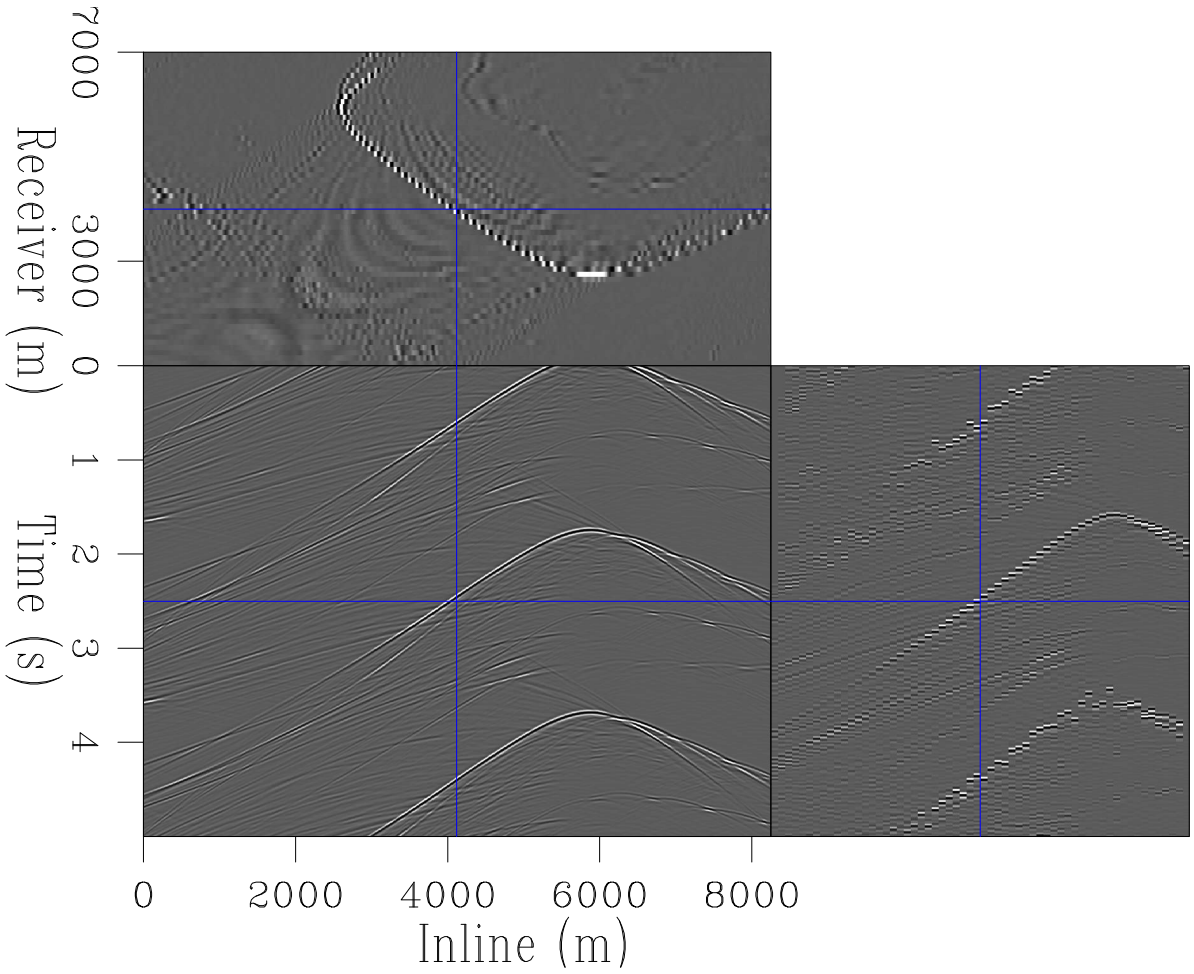


Figure 24: A pseudo-linearly blended dataset acquired using the section of the SEAM model. [**CR**] chris1/. seamdatain

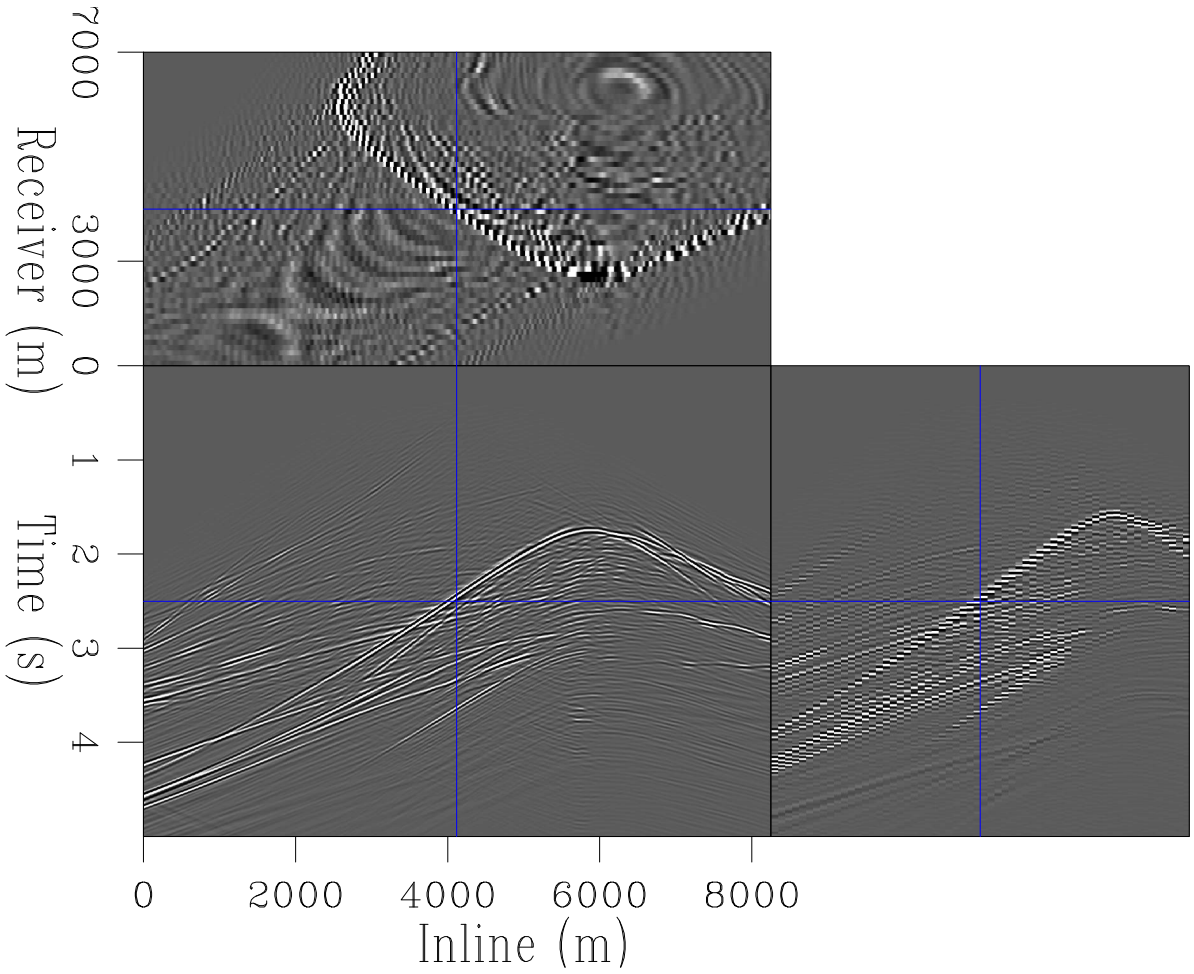


Figure 25: The output separated data after 10 iterations of inverse deimgration using a rough velocity model. [**CR**] chris1/. seamdataout

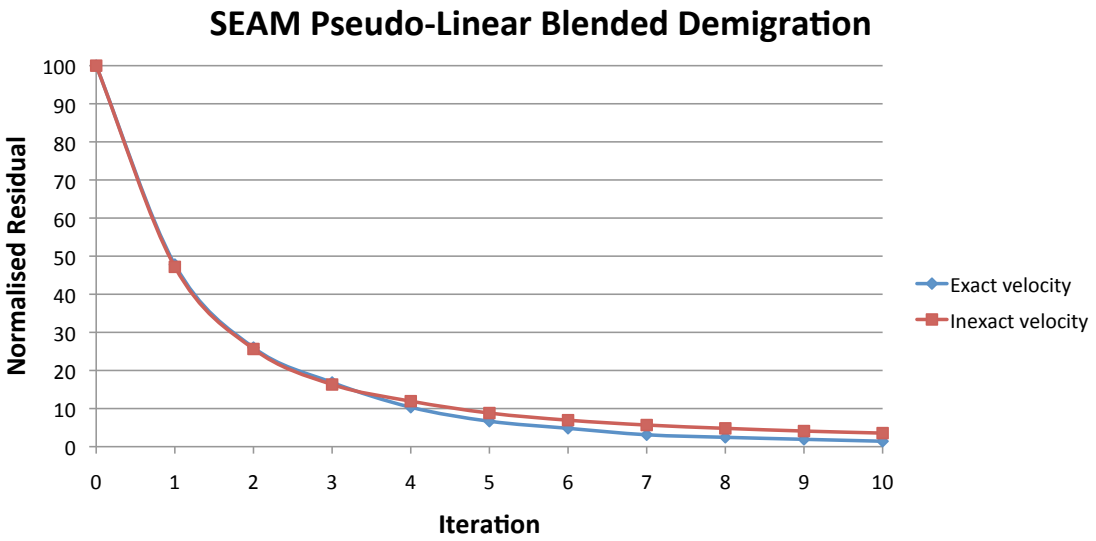


Figure 26: How the separation algorithm performs, for SEAM, as a function of iteration number. [**NR**] chris1/. seamconv

By including a blending operator into this suggested inverse demigration the system can be used to recover an equivalently separated dataset from a continuous, overlapping recording. These separated data are both kinematically accurate and amplitude real, and after re-blending match the input data to within a fraction of a percent. This is true of a synthetic example which features a variety of dips and impedance contrasts.

The fact that extended images have been constructed allows WEMVA to also be applied. Before separation a moveout estimation can be made, and the velocity model updated. This allows both separation and model updating to be applied for almost the same cost.

FUTURE WORK

Tests using field data are currently underway. These are in full 3D for an OBN dataset and the results will be available soon.

ACKNOWLEDGMENTS

The author would like to thank Ray Abma for many insightful discussions about simultaneously acquired data, and all sponsors of the Stanford Exploration Project.

REFERENCES

Abma, R., T. Manning, M. Tanis, J. Yu, and M. Foster, 2010, High quality separation of simultaneous sources by sparse inversion: Presented at the 72nd EAGE Conference & Exhibition.

Abma, R. and J. Yan, 2009, Separating simultaneous sources by inversion: Presented at the 71st EAGE Conference & Exhibition.

Ayeni, G., A. Almomin, D. Nichols, et al., 2011, On the separation of simultaneous-source data by inversion: Soc. Expl. Geophys., Expanded Abstracts, 20–25.

Baysal, E., D. Kosloff, and J. Sherwood, 1983, Reverse time migration: Geophysics, **45**, 1514–1524.

Beasley, C. J., 2008, A new look at marine simultaneous sources: The Leading Edge, **27**.

Berkhout, A., V. D. and G. Blacquiere, 2008, Processing of blended seismic data: SEG Technical Program Expanded Abstracts.

Chauris, H. and M. Benjemaa, 2010, Seismic wave-equation demigration/migration: Geophysics, **75**, S111–S119.

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

Dai, W. and J. Schuster, 2009, Least-squares migration of simultaneous sources data with a deblurring filter: SEG Expanded Abstracts, 2990–2993.

Doulgeris, P., A. Mahdad, and G. Blacquiere, 2011, Iterative separation of blended marine data: Discussion on the coherencepass filter, chapter 5, 26–31.

Hampson, G., J. Stefani, and F. Herkenhoff, 2008, Acquisition using simultaneous sources: The Leading Edge, **27**, 918–923.

Herrmann, F. J., Y. A. Erlangga, and T. T. Lin, 2009, Compressive simultaneous full-waveform simulation: Geophysics, **74**, A35–A40.

Leader, C. and B. Biondi, 2014, Image space separation of linearly blended data: SEP-155, **155**.

Moore, I., B. Dragoset, T. Ommundsen, D. Wilson, D. Eke, C. Ward, et al., 2008, Simultaneous source separation using dithered sources: Presented at the 2008 SEG Annual Meeting.

Sava, P., B. Biondi, et al., 2003, Wave-equation migration velocity analysis by inversion of differential image perturbations: 73rd Ann. Internat. Mtg Soc. of Expl. Geophys., submitted.

Sava, P. and A. Guitton, 2005, Multiple attenuation in the image space: Geophysics, **70**, V10–V20.

Sava, P. and I. Vasconcelos, 2011, Extended imaging conditions for wave-equation migration: Geophysical Prospecting, **59**, 35–55.

Tang, Y. and B. Biondi, 2009, Least square migration/inversion of blended data: SEG Expanded Abstracts, **28**, 2859–2863.

Verwest, B. and D. Lin, 2007, Modelling the impact of wide-azimuth acquisition on subsalt imaging: Geophysics, **72**, 241–250.

Zhang, Y. and L. Duan, 2013, Seismic data processing including predicting multiples using a reverse time demigration. (US Patent App. 13/860,567).

Inadequacy of inverse theory for images

Jon Claerbout and Antoine Guitton

ABSTRACT

Prior information generally enters inverse theory as regularization. In large scale problems such as image estimation where iteration does not continue to completion, an additional way to introduce a prior model is as the starting model. This lesson, hard won at Galilee, is widely applicable.

INTRODUCTION

Reflection seismology is a powerful tool in petroleum prospecting. It works so well it is often pushed beyond reasonable limits. That's when we mortals (mere data processors) get pushed into huge null spaces.

Inverse theory along with least-squares solving technology seduces us into thinking a good solution is at hand as soon we have a data fitting regression along with suitable model and data covariances. A drawn-out, humiliating experience taught us more is needed. The very size of our problems leads to a pitfall that is surely widespread.

Our most valuable datasets bring us into spaces of such high dimensionality we cannot know whether we have iterated long enough. A simple data-fitting problem where we know which answers are plausible and which impossible eventually taught us that what we generally ignore is what we very often need. The years we spent with this simple problem gives us fear that many colleagues, with their more difficult problems, produce solutions that are often wrong!

The convexity of least squares methods along with clever precondioning schemes trick us into feeling our final solution hardly depends on the starting guess. Actually, truth may often be the opposite.

THE CHALLENGING SETTING

An example of the simplest null space is one data value d to be divided into two models, m_1 and m_2. Should 10 be divided into $5+5$ or into $9+1$? This arbitrariness becomes obscure when $\mathbf{m}_1$ and $\mathbf{m}_2$ are families of complicated models competing to grab what's left (if anything) in multivariate data $\mathbf{d}$.

Less comprehensible scenarios in seismology arise when: (1) We gave them a great map of reflectivity, now they want density. (2) We gave them a fine map of velocity, now they want anisotropy. And a grand challenge, (3) There is a giant the null space between anisotropy and inhomogeneity. How should we characterize it?

The clear academic example examined here points to analogous, but much deeper, industrial examples. We take up a lake survey with a depth sounder. Data is the travel time from the water surface to the water bottom measured at many locations, possibly along survey lines that are somewhat organized. Starting from the presumption the lake surface is perfectly flat, unchanging during the survey, we find apparent survey tracks in our derived image (model) of the water bottom. This astonishes us and forces us to imagine a supplemental model looking like water surface level fluctuating during the surveying.

For 20 years we did not adequately solve the lake problem. The heart of it is that we cannot learn a better model without fully stating our prior model. Model covariances (even if we knew them!) are not enough.

The lake here is known as the "Sea of Galilee." A survey there gave data with complexities in many forms. Here we limit details to those central to the story. Looking back, why did it take us so long? When two different things look the same, they tend to have the same name, even though they differ. That and because theory and practice are worlds apart.

Prior model

There is a boat with a depth sounder, thirty year old navigation gear, and a recorder. Over the course of a season or several, the boat crosses the lake hundreds of times, along somewhat regular tracks, obtaining 131,514 triples (x_i, y_i, z_i), instances of observed depth $\mathbf{d} = (z_i)$ at surface (x_i, y_i) locations. The lake level fluctuates for many reasons (rain and drain?). We have been told (but cannot be certain) the given values of $\mathbf{d} = (z_i)$ have been properly corrected for lake level.

At boat locations (x_i, y_i) we have measured data depth z_i. A model is values of z on a uniform grid in (x, y) space. Linear interpolation (or nearest neighbor extraction) finds modeled depth z anywhere, hopefully a good approximation to z_i at each (x_i, y_i). We express data as $\mathbf{d}_{\text{modeled}} = \mathbf{Gm}$ where $\mathbf{m}$ is water depths on a regular 2-D mesh. The operator (matrix) $\mathbf{G}$ contains mainly information about navigation and interpolation. The navigation values (x_i, y_i) are not treated as data having noise (although we later suspect they should have been).

A number of data fitting issues need not concern us here, issues such as frequent spikes in the depth, zeros in the depth, gaps in the areal coverage, erratic spikes and a few surges in the navigation (x_i, y_i). What you do need to know is that brightness and darkness in Figure 1 do not directly represent depth itself; they represent a roughening of it, like its gradient (actually, after a 2-D operator like a gradient, the helix derivative). This roughening is needed because water bottom features are such small features on the overall trend of water depth.

Correction for apparent water level

Our first idea of water level fluctuation with time (including date) was simply "rain and drain." Since surveying is inactive at night and on holidays we first imagined surface elevation function an assemblage of step functions. We didn't have a ready way of modeling

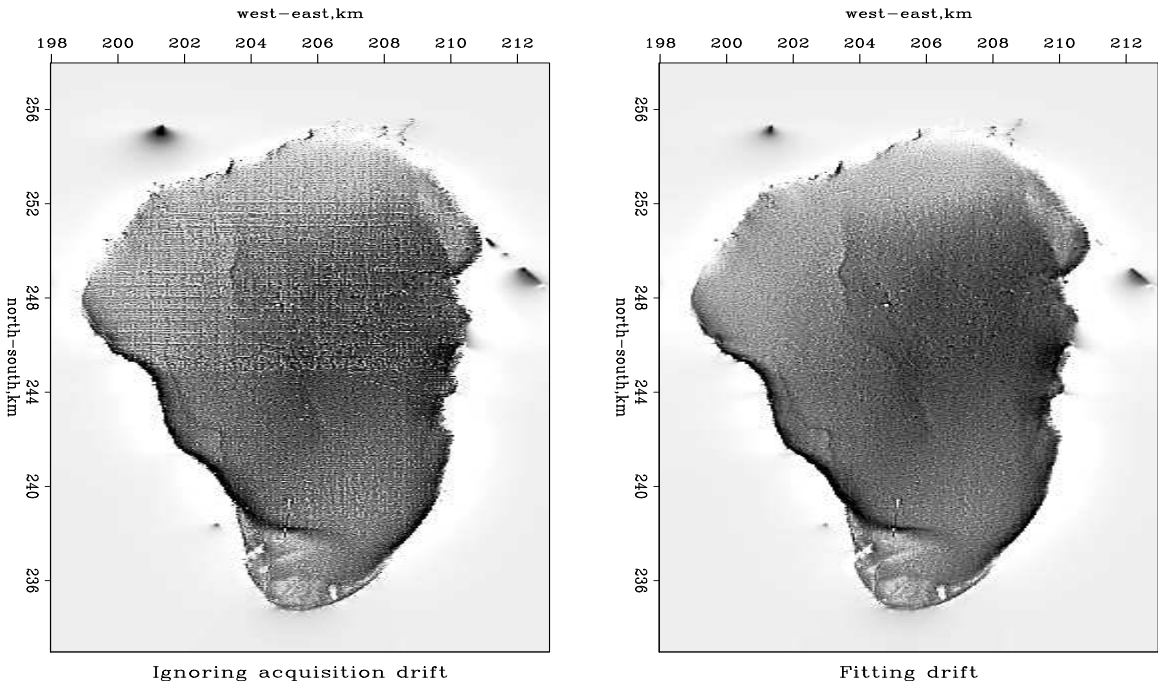


Figure 1: Left, roughened water depth **p** given a flat water surface. Right, roughened water depth **p** after compensating for apparent surface elevation variation. jon/. antoine4

those, so we tried a slowly-variable continuous function of measurement time.

Usually an erroneous low frequency in data is easy to handle by minimizing low-cut filtered residuals. Attempts to do this failed because of the erratic presence of spikes, zeros, and surges throughout the data. So, low-frequency measurement drift had to be modeled instead. Then spikes were easily managed by using an ℓ_1 styled data fitting procedure (hyperbolic penalty).

Modeling the water surface as a function of time led to beautiful track-free images resembling the right side of Figure 1. These images are delightful, but there was an underlying problem to be taken seriously. The water surface elevation **e** crossing the lake looked as sketched in Figure 2. At low frequencies the surface elevation **e** mimicked the water bottom **b**. The data (travel time) was distributing itself between the water bottom **b** and the surface top **e**. The water bulge in the middle of the lake came out unreasonably large. To get the tracks out of the water bottom image, a couple meters of water bulge was required!

Algebraic foundation

To "debug" the analysis we need to examine the regressions. Physical functions are smooth, both the water bottom map $b(x, y) = \mathbf{b}$ and the water surface elevation $e(t) = \mathbf{e}$. For regularization, **b** is roughened with the 2-D operator **A**, typically a helix derivative (square-root of the FT of a Laplacian) and **e** is roughened with a low-cut filter, typically $\mathbf{L}^{-1}$, where

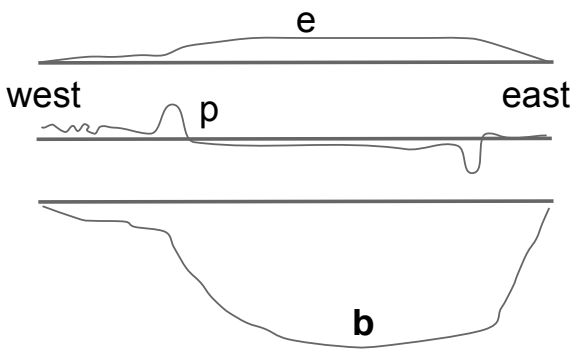


Figure 2: Cartoon of estimated lake surface elevation **e**, roughened bottom **p**, and estimated bottom **b**. jon/. mep4

L is leaky integration. The data fitting and two regularizing regressions are:

$$\mathbf{0} \approx_h \mathbf{Gb} + \mathbf{e} - \mathbf{d} \tag{1}$$
$$\mathbf{0} \approx_2 \mathbf{Ab} \tag{2}$$
$$\mathbf{0} \approx_2 \mathbf{L}^{-1}\mathbf{e} \tag{3}$$

The subscript 2 on $\approx_2$ means least squares, while the subscript h on $\approx_h$ means hyperbolic penalty (to soften noise bursts in the data).

A basic notion of statistics is that the regularizations should lead to residuals that tend to be roughly 2-D white (flat spectrum) in the (x, y) plane. Thus, **A** should be chosen so that **Ab** is roughly white. In practice a new variable $\mathbf{Ab} = \mathbf{p}$ called the preconditioner is introduced. Besides fulfilling theoretical desiderata, the variable **p** has interpretive use. It is the variable shown in Figure 1. The second regularization is that the elevation **e** be smooth along measurement time. Smoothed white noise **Ln** should give us a signal that looks like our preconceived water surface **e**, so the associated regularization is $\mathbf{0} \approx \mathbf{n} = \mathbf{L}^{-1}\mathbf{e}$.

Changing the formulation from the physical variables $(\mathbf{b}, \mathbf{e})$ to the computational/statistical variables $(\mathbf{p}, \mathbf{n})$, is called "preconditioning." Preconditioning speeds iterative solutions, and it handles matters as statistical theory instructs us. The preconditioned regressions are:

$$\mathbf{0} \approx_h \mathbf{GA}^{-1}\mathbf{p} + \lambda\mathbf{Ln} - \mathbf{d} \tag{4}$$
$$\mathbf{0} \approx_2 \epsilon_b \mathbf{p} \tag{5}$$
$$\mathbf{0} \approx_2 \epsilon_e \mathbf{n} \tag{6}$$

From a purely mathematical point of view, ϵ_b and ϵ_e are infinitesimals. Iteration would then resolve the data fitting before starting on the regularizations. Model space size is roughly $400^2 = 160,000$ while the iteration count is likely under 50. With so few iterations the regularizations seem hardly to come into play, except that they were earlier embedded by the preconditioning.

It might seem the ratio of the two unknown epsilons determines how much of the null space will end out on the water bottom and how much on the top. And, it might seem the parameter lambda λ is merely a scaling factor in the lowpass filter **L**. But, λ strongly affects

the balance of $\mathbf{p}$ and $\mathbf{n}$ in the gradient. The limited iteration count leads to the epsilon ratio being far less significant than the size of λ. Although we can easily include regularizations (5) and (6), at SEP they are generally ignored. What really matters is the data fitting (4). Here Antoine tried keeping the regularizations and found it made no difference.

Originally, we felt leaky integration $\mathbf{L}$ contained the only parameter needed to adjust the spectrum of the elevation, but we soon realized it always had too much short wavelength energy because of the sharp onset of the damped exponential in $\mathbf{L}$. So we switched to its autocorrelation $\mathbf{L}^{\mathrm{T}}\mathbf{L}$ in the low pass relation $\mathbf{e} = \mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{n}$. Thus, in practice, our code is iteratively working this lone regression:

$$\mathbf{0} \quad \approx_h \quad \mathbf{G}\mathbf{A}^{-1}\mathbf{p} \;+\; \lambda\mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{n} \;-\; \mathbf{d} \tag{7}$$

Finding the worst source of null space

We have a convex regression in data space. We have two regularizations in model space that we have taken into account. This should not fail, but it does. How does it fail? We can play with λ, and we did. We could find values of λ that were big enough to suppress the tracks in the image, but those values of λ still created giant bulges on the water surface. In other words, the water surface elevation $\mathbf{e}$ visually correlates with the data, both the observed data and the modeled data $\mathbf{G}\mathbf{b}$. In the middle of the lake the estimated elevation is a couple meters above that at the shoreline. Obviously wrong.

Why should we care that it is wrong, and why is it coming out wrong? We are seeing correlation between data made from $\mathbf{b}$ and data made from $\mathbf{e}$, If we cannot prevent apparent correlation within an estimated model $\mathbf{m} = (\mathbf{b}, \mathbf{e})$ where in real life there is none, how can we hope to know when such correlation is real? Imagine a map of seismic velocity correlating with a map of anisotropy. Are the two correlated in geology, or is the correlation a data processing artifact?

Regularization is not the only way to manage a null space. Choosing your starting solution carefully can make a difference—a huge difference. Textbook theory tells us with convex optimization (such as least squares) final solutions are independent of the starting model, but we learn otherwise from nonlinear problems, and we learn otherwise from linear problems that are too large for us to iterate to completion.

Any null space produces no perturbation in the modeled data so it cannot improve the fit to the observed data. Consequently, whatever null space may exist in the starting model will remain there throughout the iterative fitting process. When we included $\mathbf{e}$ in the data fitting, we introduced as many unknown model parameters as we have data values in $\mathbf{d}$, so we certainly know we now have a giant null space.

The starting model is not the Bayes' prior model. The starting model is simply one of many places to start the iterative solver. But, putting our prior model into our starting model assures us whatever null space it may contain will remain in our final solution. Bingo! The ultimately found elevation is shown in Figure 3, hardly an assemblage of step functions suggested by our original rain and drain ideas!

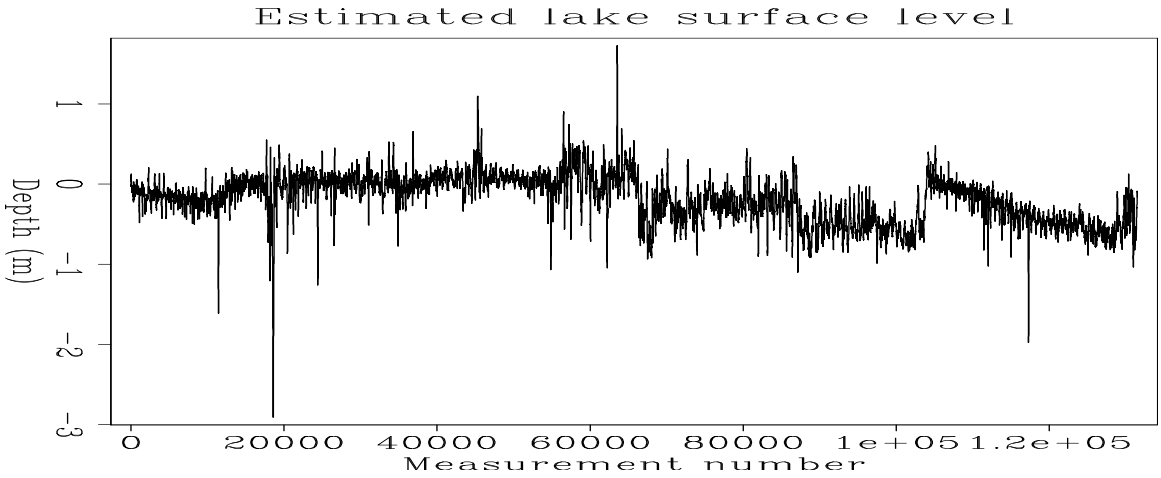


Figure 3: Apparent water surface elevation during the entire Galilee survey (an unknown number of months). Elevation ranges over almost a meter, but is no longer correlated with water depth. The 10cm fuzz corresponds to 10cm measurement z_i discretization. I'm left feeling (x_i, y_i) require a time-variable model, such as systematic navigation errors. We are astonished to see the apparent elevation containing big spikes. This might be explained by surges of navigation error. jon/. tide

Problem solved

So, after 20 years, how did we finally fix the Galilee fitting? We solved two problems, one after the other. Setting $\lambda = 0$ in regression (7)

$$\mathbf{0} \quad \approx_h \quad \mathbf{G}\mathbf{A}^{-1}\mathbf{p} \;+\; \lambda\mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{n} \;-\; \mathbf{d}$$

amounts to freezing a flat surface elevation $\mathbf{e} = \mathbf{L}\mathbf{n} = \mathbf{0}$. So, starting from $(\mathbf{p}, \mathbf{n}) = (\mathbf{0}, \mathbf{0})$ with $\lambda = 0$, we ran it getting our first solution for $\mathbf{p}$. This bottom image shows the ship tracks. We use that $\mathbf{p}$ to define $\mathbf{p}_0$ for the second pass. Iterating $(\mathbf{p}, \mathbf{n})$ starting from $(\mathbf{p}_0, (\mathbf{n}_0 = \mathbf{0}))$ gives the final bottom and top $(\mathbf{b}, \mathbf{e})$ estimates. We are free now to experiment with λ without developing the unholy correlation between $\mathbf{b}$ and $\mathbf{e}$.

The various choices of λ represent subjective guesses how to divide the data between bottom $\mathbf{b}$ and top $\mathbf{e}$. There's an opportunity here. Perhaps upon investigating the various λ choices, we might some day find a favorite, then find a reason for the favorite, then understand that reason is suggesting something lacking in the present analytic framework. Maybe we can sniff out an opportunity for building a model that better describes this data. That's real science. Inverse theory is merely a guide to parameterizing known models.

CONCLUSION

Books tell us we should specify prior information in the regularizations. Here we learned that we should specify it in the starting model too, particularly if we cannot iterate to completion. Finding a suitable prior model may, in many cases, be an easier and better way of regularizing. That's what Galilee teaches me. This experience also suggests that high

resolution models should generally be derived from lower resolution models (the assumption of scale invariance).

The process that led us to select regularizations ($\mathbf{A}$ and $\mathbf{L}^{-1}$) is highly subjective, sloppy even. We can do better simply by looking at the spectrum of $\mathbf{n}$ and $\mathbf{p}$. If they are not white, then adjusting the preconditioners to achieve it. This should have been done, but was not. If it had, would the overall story change? We don't know.

Our personal opinion is that our fellow image estimators get wrong answers much of the time, and they don't know it because they are not working on easy problems like Galilee, where "wrong answers" are easier to recognize.

What about velocity anisotropy?

Have we learned something from Galilee that we can carry over to estimating anisotropy? We think so, but have neither audacity nor time to explore details. You would be right to guess that we would start by solving for isotropic inhomogeneous material getting $\mathbf{m}_0$. We would use that $\mathbf{m}_0$ as the starting model in iterative fitting simultaneously for inhomogeneity and anisotropy. The regularization in the second problem would be minimizing filtered $\mathbf{m} - \mathbf{m}_0$.

Reference

This article is an adaptation of section 6.6 of the 2014 textbook "Geophysical Image Estimation by Example" by Jon Claerbout with Sergey Fomel. That book chapter is devoted to optimization with the hyperbolic penalty function using the Galilee dataset as an example. Many more details of the example are found there. The book is freely available at the website of Jon Claerbout along with his previous books.

ACKNOWLEDGEMENT

We thank Zvi Avraham for supplying the data. Jon wrote the words in this article, and prepared a video. Antoine cheerfully and enthusiastically persevered through the many versions of the computations.

Adjoint formulation for the elastic wave equation

Gustavo Alves

ABSTRACT

I present the formulation for the forward non-linear elastic wave equation and the linearized Born modeling approximation. The methodolgy used is similar to the adjoint formulation presented for the acoustic case by Almomin (2013). I also demonstrate how a perfectly matched absorbing layer (PML) can be applied to the formulation and derive the necessary adjoints for implementing inversion methods. Finally, I show results for a synthetic model using the linear approximation and compare it to the non-linear solution for a small localized perturbation. The linear solution converges to the non-linear one when the perturbation is small in relation to the background model.

INTRODUCTION

Seismic theory tells us that waves propagating in a linear elastic medium can be uniquely described by prescribing body forces and boundary conditions of a given problem (Aki and Richards, 1980). The solution is usually a combination of body waves, both pressure and shear, and surface waves. However, most current imaging techniques use only pressure waves. This is due both to the lack of available multicomponent data and the greater computational requirements of processing multicomponent data.

In recent years, however, this scenario has started to change. New technologies, such as ocean bottom cables (OBCs) and ocean bottom nodes (OBNs), have extended the ability to record multicomponent data to offshore acquisitions. New prospecting challenges also push the limits of current imaging methods, which in turn force the need for more detailed data. Consequently, a new collection of algorithms is required to process this new multicomponent data.

In this work, I present the velocity-stress formulation for the elastic wave-equation in the framework of adjoint methods. My goal is to cast the problem as a series of adjoint operators, so that both linearized forward methods and their adjoints can be clearly constructed.

I start by constructing two possible representations to the recursive operator for the 2D elastic wave equation, together with their respective adjoints. The recursive operator and its adjoint are the core elements for most wave propagation methods, such as Born modeling, Reverse Time Migration (RTM), Tomography and Wave Equation Migration Velocity Analysis (WEMVA). Therefore, proposing different solutions to the recursive operator helps in better understanding how to develop such methods.

I follow by extending both solutions to include an absorbing boundary condition. I propose the use of a Perfectly Matched Layer (PML), according to Collino and Tsogka (2001).

Finally, I describe the implementation of one of the proposed recursive operators for the Born operator. The Born operator is a linearized approximation of the non-linear wave equation and accurately estimates the non-linear solution for small perturbations of the background wavefield. I show synthetic examples for a point scatterer and compare the results to the non-linear case.

METHODOLOGY

The elastic wave equation for a two dimensional problem can be written in the velocity-stress formulation as a set of 5 equations,

$$\rho(\mathbf{x})\frac{\partial}{\partial t}V_x(\mathbf{x},t) - [\frac{\partial}{\partial x}\sigma_{xx}(\mathbf{x},t) + \frac{\partial}{\partial z}\sigma_{xz}(\mathbf{x},t)] = s_1(\mathbf{x},t) \tag{1}$$

$$\rho(\mathbf{x})\frac{\partial}{\partial t}V_z(\mathbf{x},t) - [\frac{\partial}{\partial x}\sigma_{xz}(\mathbf{x},t) + \frac{\partial}{\partial z}\sigma_{zz}(\mathbf{x},t)] = s_2(\mathbf{x},t) \tag{2}$$

$$\frac{\partial}{\partial t}\sigma_{xx}(\mathbf{x},t) = [\lambda(\mathbf{x}) + 2\mu(\mathbf{x})]\frac{\partial}{\partial x}V_x(\mathbf{x},t) + \lambda(\mathbf{x})\frac{\partial}{\partial z}V_z(\mathbf{x},t) + s_3(\mathbf{x},t) \tag{3}$$

$$\frac{\partial}{\partial t}\sigma_{zz}(\mathbf{x},t) = [\lambda(\mathbf{x}) + 2\mu(\mathbf{x})]\frac{\partial}{\partial z}V_z(\mathbf{x},t) + \lambda(\mathbf{x})\frac{\partial}{\partial x}V_x(\mathbf{x},t) + s_4(\mathbf{x},t) \tag{4}$$

$$\frac{\partial}{\partial t}\sigma_{xz}(\mathbf{x},t) = \mu(\mathbf{x})[\frac{\partial}{\partial z}V_x(\mathbf{x},t) + \frac{\partial}{\partial x}V_z(\mathbf{x},t)] + s_5(\mathbf{x},t), \tag{5}$$

where ρ, λ and μ are the model parameters, V_x and V_z are the particle velocities, σ_{xx}, σ_{zz} and σ_{xz} are the normal and shear stresses, respectively, s_1 and s_2 are the velocity components of the source and s_3, s_4 and s_5 are the stress components of the source.

My first goal is to write the forward non-linear modeling operator and its adjoint. To achieve that, I need to re-cast the previous set of equations as a recursive system. First, I apply a finite difference approximation to the time derivates, following the staggered-grid approach described by Virieux (1986) and Levander (1988). The equations become

$$V_x^n = V_x^{n-1} + \frac{\Delta t}{\rho}(\frac{\partial}{\partial x}\sigma_{xx}^{n-1/2} + \frac{\partial}{\partial z}\sigma_{xz}^{n-1/2} + s_1^{n-1}) \tag{6}$$

$$V_z^n = V_z^{n-1} + \frac{\Delta t}{\rho}(\frac{\partial}{\partial x}\sigma_{xz}^{n-1/2} + \frac{\partial}{\partial z}\sigma_{zz}^{n-1/2} + s_2^{n-1}) \tag{7}$$

$$\sigma_{xx}^{n+1/2} = \sigma_{xx}^{n-1/2} + \Delta t[(\lambda + 2\mu)\frac{\partial}{\partial x}V_x^n + \lambda\frac{\partial}{\partial z}V_z^n + s_3^{n-1/2}] \tag{8}$$

$$\sigma_{zz}^{n+1/2} = \sigma_{zz}^{n-1/2} + \Delta t[(\lambda + 2\mu)\frac{\partial}{\partial z}V_z^n + \lambda\frac{\partial}{\partial x}V_x^n + s_4^{n-1/2}] \tag{9}$$

$$\sigma_{xz}^{n+1/2} = \sigma_{xz}^{n-1/2} + \Delta t[\mu\frac{\partial}{\partial x}V_z^n + \mu\frac{\partial}{\partial z}V_x^n + s_5^{n-1/2}], \tag{10}$$

where $\mathbf{n}$ is the discretized time interval and $\mathbf{\Delta t}$ is the time step. For simplicity, I supress the spatial dependencies of the vectors.

The time staggering in this approach uses an alternating solution to the velocity and stress wave fields, which is concise and easy to implement. However, describing the set of equations as a single recursive relation becomes more difficult. Here, I choose two possible solutions to this problem. In the first method, which I call recursive operator by backward substitution, the stress equations are re-injected into the velocity equations, so that a single

recursive equation arises. The second method, which I refer as recursive operator by time step refinement, maintains the staggered-time structure, but redefines the time stepping of each wave field so that both velocities and stresses can be defined at every time step.

Recursive operator by backward substitution

My first step in this approach is to write the velocities and stresses as two separate data vectors,

$$\mathbf{d_1^n} = \begin{pmatrix} V_x^n \\ V_z^n \end{pmatrix} \qquad \mathbf{d_2^n} = \begin{pmatrix} \sigma_{xx}^{n-\frac{1}{2}} \\ \sigma_{zz}^{n-\frac{1}{2}} \\ \sigma_{xz}^{n-\frac{1}{2}} \end{pmatrix}. \tag{11}$$

I also define a few operators to make the representation more compact,

$$A = \tfrac{\Delta t}{\rho}\tfrac{\partial}{\partial x} \qquad B = \tfrac{\Delta t}{\rho}\tfrac{\partial}{\partial z} \qquad C = \tfrac{\Delta t}{\rho}$$

$$D = \Delta t(\lambda + 2\mu)\tfrac{\partial}{\partial x} \qquad E = \Delta t\lambda\tfrac{\partial}{\partial z} \qquad F = \Delta t(\lambda + 2\mu)\tfrac{\partial}{\partial z}$$

$$G = \Delta t\lambda\tfrac{\partial}{\partial x} \qquad H = \Delta t\mu\tfrac{\partial}{\partial x} \qquad J = \Delta t\mu\tfrac{\partial}{\partial z}\ .$$

I can now write my elastic equations as two matrix operations,

$$\mathbf{d_1^n} = \mathbf{d_1^{n-1}} + \begin{pmatrix} A & 0 & B \\ 0 & B & A \end{pmatrix} \mathbf{d_2^n} + C \begin{pmatrix} s_1^{n-1} \\ s_2^{n-1} \end{pmatrix} \tag{12}$$

$$\mathbf{d_2^n} = \mathbf{d_2^{n-1}} + \begin{pmatrix} D & E \\ F & G \\ H & J \end{pmatrix} \mathbf{d_1^{n-1}} + \Delta t \begin{pmatrix} s_3^{n-\frac{1}{2}} \\ s_4^{n-\frac{1}{2}} \\ s_5^{n-\frac{1}{2}} \end{pmatrix}. \tag{13}$$

Substituting equation 10 into 12 yields,

$$\mathbf{d_1^n} = \mathbf{d_1^{n-1}} + \begin{pmatrix} A & 0 & B \\ 0 & B & A \end{pmatrix} \mathbf{d_2^{n-1}} + \begin{pmatrix} A & 0 & B \\ 0 & B & A \end{pmatrix} \begin{pmatrix} D & E \\ F & G \\ H & J \end{pmatrix} \mathbf{d_1^{n-1}} + C \begin{pmatrix} s_1^{n-1} \\ s_2^{n-1} \end{pmatrix} + \Delta t \begin{pmatrix} A & 0 & B \\ 0 & B & A \end{pmatrix} \begin{pmatrix} s_3^{n-\frac{1}{2}} \\ s_4^{n-\frac{1}{2}} \\ s_5^{n-\frac{1}{2}} \end{pmatrix}. \tag{14}$$

I can now define a single vector $\mathbf{d^n}$

$$\mathbf{d^n} = \begin{pmatrix} V_x^n \\ V_z^n \\ \sigma_{xx}^{n-\frac{1}{2}} \\ \sigma_{zz}^{n-\frac{1}{2}} \\ \sigma_{xz}^{n-\frac{1}{2}} \end{pmatrix}, \tag{15}$$

which combines both subsets $\mathbf{d_1^n}$ and $\mathbf{d_2^n}$ into a single generalized recursive relation,

$$\mathbf{d^n} = \begin{pmatrix} I+AD+BH & AE+BJ & A & 0 & B \\ BF+AH & I+BG+AJ & 0 & B & A \\ D & E & I & 0 & 0 \\ F & G & 0 & I & 0 \\ H & J & 0 & 0 & I \end{pmatrix} \mathbf{d^{n-1}} + \begin{pmatrix} C & 0 & A & 0 & B \\ 0 & C & 0 & B & A \\ 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & \Delta t \end{pmatrix} \mathbf{S^{n-1}}, \tag{16}$$

where $\mathbf{I}$ is the identity operator and $\mathbf{S^n}$ is the source vector at time $\mathbf{n}$. For compactness, I define the matrices in 16 as $\mathbf{X}$ and $\mathbf{Y}$, so that the forward recursive relation can be written simply as

$$\mathbf{d^n} = \mathbf{X}\mathbf{d^{n-1}} + \mathbf{Y}\mathbf{S^{n-1}}. \tag{17}$$

While equation 17 correctly describes the forward elastic wave propagation, the adjoint of this equation cannot be taken directly since operators $\mathbf{X}$ and $\mathbf{Y}$ don't commute. The last step is to re-define the source term. The vector $\mathbf{q^n}$ is defined as the source term $\mathbf{S^n}$ after applying the operator $\mathbf{Y}$. Equation 18 is the final forward recursive operator

$$\mathbf{d^n} = \mathbf{X}\mathbf{d^{n-1}} + \mathbf{q^{n-1}}, \tag{18}$$

whose adjoint is described by

$$\mathbf{q^n} = \mathbf{X'}\mathbf{q^{n+1}} + \mathbf{d^{n+1}}, \tag{19}$$

where $\mathbf{X'}$ is the adjoint of $\mathbf{X}$.

Recursive operator by time step refinement

In this approach, instead of re-injecting the stress fields into the velocities to get them at integer time steps, I take the opposite approach. In other words, this method tries to obtain all wavefields at both integer and half-integer time steps, effectively refining the time stepping in the recursive relation.

Again, I start by definning a data vector $\mathbf{D^n}$, which represents the velocities and stresses at time $\mathbf{n}$. Notice that, unlike in the previous case, the stresses are not shifted in time in respect to the velocities.

$$\mathbf{d^n} = \begin{pmatrix} V_x^n \\ V_z^n \\ \sigma_{xx}^n \\ \sigma_{zz}^n \\ \sigma_{xz}^n \end{pmatrix}. \tag{20}$$

Next, I define a set of operator matrices that represent the original time staggered

equations,

$$\mathbf{d^n} = \begin{pmatrix} 0 & 0 & A & 0 & B \\ 0 & 0 & 0 & B & A \\ D & E & 0 & 0 & 0 \\ F & G & 0 & 0 & 0 \\ H & J & 0 & 0 & 0 \end{pmatrix} \mathbf{d^{n-\frac{1}{2}}} + \mathbf{I} \cdot \mathbf{d^{n-1}} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & \Delta t \end{pmatrix} \mathbf{S^{n-\frac{1}{2}}} +$$

$$\begin{pmatrix} C & 0 & 0 & 0 & 0 \\ 0 & C & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{S^{n-1}}, \qquad (21)$$

where $\mathbf{I}$ is the identity matrix and operators $\mathbf{A}$ through $\mathbf{J}$ are the same as those defined in the previous section. The forward recursive relation can be represented as

$$\mathbf{d^n} = \mathbf{X d^{n-\frac{1}{2}}} + \mathbf{I} \cdot \mathbf{d^{n-1}} + \mathbf{Y_1 S^{n-\frac{1}{2}}} + \mathbf{Y_2 S^{n-1}}, \qquad (22)$$

and its adjoint as

$$\mathbf{q^n} = \mathbf{X' q^{n+\frac{1}{2}}} + \mathbf{I} \cdot \mathbf{q^{n+1}} + \mathbf{Y'_1 d^{n+\frac{1}{2}}} + \mathbf{Y'_2 d^{n+1}}. \qquad (23)$$

While the recursive relations described here and in the previous section are at the core of most forward and inverse methodologies, it is important to take into account other important effects that appear when trying to numerically solve the elastic wave equation. In the next section, I re-derive the two previous solutions, taking into account border and interpolation effects.

PERFECTLY MATCHED LAYER IMPLEMENTATION

A perfectly matched layer (PML) is a type of absorbing boundary condition used in numerical modeling to avoid reflections of propagating waves off the corners of a finite numerical problem. It is more efficient than decaying exponential methods (Cerjan et al., 1985), because it splits the wavefront into its directional components, applying an absorbing factor in each direction instead of a single absorption normal to the model boundary.

In order to apply PML to my problem, I start by splitting the previous set of equations into its spatial derivatives. Since this is a 2D problem, the original equations are split into their $\mathbf{x}$ and $\mathbf{z}$ derivatives, namely: V_x^x, V_x^z, V_z^x and V_z^z for the particle velocities and σ_{xx}^x, σ_{xx}^z, σ_{zz}^x, σ_{zz}^z, σ_{xz}^x and σ_{xz}^z for the stresses. Also, I use the subscripts i, j to represent the grid position to which the variables belong. I need to include this because I implement the equations using the staggered-grid approach, similarly to what was done in the time domain

in the previous section. The decomposed set of 10 equations is

$$\frac{\partial}{\partial t}(V_x^x)_{i,j} = \frac{1}{\rho}[\frac{\partial}{\partial x}(\sigma_{xx})_{i+1/2,j}] + s_x^x \tag{24}$$

$$\frac{\partial}{\partial t}(V_x^z)_{i,j} = \frac{1}{\rho}[\frac{\partial}{\partial z}(\sigma_{xz})_{i,j+1/2}] + s_x^z \tag{25}$$

$$\frac{\partial}{\partial t}(V_z^x)_{i+1/2,j+1/2} = \frac{1}{\rho}[\frac{\partial}{\partial x}(\sigma_{xz})_{i,j+1/2}] + s_z^x \tag{26}$$

$$\frac{\partial}{\partial t}(V_z^z)_{i+1/2,j+1/2} = \frac{1}{\rho}[\frac{\partial}{\partial z}(\sigma_{zz})_{i+1/2,j}] + s_z^z \tag{27}$$

$$\frac{\partial}{\partial t}(\sigma_{xx}^x)_{i+1/2,j} = (\lambda + 2\mu)\frac{\partial}{\partial x}(V_x)_{i,j} + s_{xx}^x \tag{28}$$

$$\frac{\partial}{\partial t}(\sigma_{xx}^z)_{i+1/2,j} = \lambda\frac{\partial}{\partial z}(V_z)_{i+1/2,j+1/2} + s_{xx}^z \tag{29}$$

$$\frac{\partial}{\partial t}(\sigma_{zz}^x)_{i+1/2,j} = \lambda\frac{\partial}{\partial x}(V_x)_{i,j} + s_{zz}^x \tag{30}$$

$$\frac{\partial}{\partial t}(\sigma_{zz}^z)_{i+1/2,j} = (\lambda + 2\mu)\frac{\partial}{\partial z}(V_z)_{i+1/2,j+1/2} + s_{zz}^z \tag{31}$$

$$\frac{\partial}{\partial t}(\sigma_{xz}^x)_{i,j+1/2} = \mu\frac{\partial}{\partial x}(V_z)_{i+1/2,j+1/2} + s_{xz}^x \tag{32}$$

$$\frac{\partial}{\partial t}(\sigma_{xz}^z)_{i,j+1/2} = \mu\frac{\partial}{\partial z}(V_x)_{i,j} + s_{xz}^z, \tag{33}$$

where $V_x = V_x^x + V_x^z$, and so on. It is important to note that, following the method by Virieux (1986), the elastic properties of the model are defined at separate grid points. Therefore, they must be averaged to correspond to the equivalent values at the points being calculated. Here, I follow the method of harmonic averages described by Moczo et al. (2002).

The next step is to include the PML parameters. To do so, I follow the work of Collino and Tsogka (2001). Taking equation 24 as an example and representing the time derivative as a second order finite difference approximation, the elastic equation with PML becomes

$$\frac{(V_x^x)_{i,j}^{n+1} - (V_x^x)_{i,j}^n}{\Delta t} + d_i^x\frac{(V_x^x)_{i,j}^{n+1} + (V_x^x)_{i,j}^n}{2} = \frac{1}{\rho}[\frac{\partial}{\partial x}(\sigma_{xx})_{i+1/2,j}] + s_x^x, \tag{34}$$

where the superscript n refers to the discrete time in the finite differences method and d_i^x is given by

$$d_i^x = log(\frac{1}{R})(\frac{3V_P}{2\delta})(\frac{x}{\delta})^2, \tag{35}$$

where the index i refers to the grid position and the x represents the direction of absorption. R is an arbitrary parameter that is associated with the desired reflectivity at the outer boundary and δ is the boundary thickness. Typical values for these parameters are 0.001 and 10 grid points, respectively.

Rearranging the terms in equation 34, I get

$$(V_x^x)_{i,j}^{n+1} = (1 + \frac{\Delta t}{2}d_i^x)^{-1}[(1 - \frac{\Delta t}{2}d_i^x)(V_x^x)_{i,j}^n + \frac{\Delta t}{\rho}\frac{\partial}{\partial x}(\sigma_{xx})_{i+1/2,j}^{n+1/2}], \tag{36}$$

and similarly for the z component of the V_x velocity I get

$$(V_x^z)_{i,j}^{n+1} = (1 + \frac{\Delta t}{2}d_j^z)^{-1}[(1 - \frac{\Delta t}{2}d_j^z)(V_x^z)_{i,j}^n + \frac{\Delta t}{\rho}\frac{\partial}{\partial z}(\sigma_{xz})_{i,j+1/2}^{n+1/2}]. \tag{37}$$

Re-writing the set of equations as a chain of operators, I get

$$(V_x^x)_{i,j}^n = KL(V_x^x)_{i,j}^{n-1} + KA(\sigma_{xx})_{i+1/2,j}^{n-1/2} + KC(s_x^x)^{n-1} \tag{38}$$
$$(V_x^z)_{i,j}^n = MN(V_x^z)_{i,j}^{n-1} + MB(\sigma_{xz})_{i,j+1/2}^{n-1/2} + MC(s_x^z)^{n-1} \tag{39}$$
$$(V_z^x)_{i+1/2,j+1/2}^n = OP(V_x^x)_{i,j}^{n-1} + OA(\sigma_{xz})_{i,j+1/2}^{n-1/2} + OC(s_z^x)^{n-1} \tag{40}$$
$$(V_z^z)_{i+1/2,j+1/2}^n = QR(V_x^z)_{i,j}^{n-1} + QB(\sigma_{xx})_{i+1/2,j}^{n-1/2} + QC(s_z^z)^{n-1} \tag{41}$$
$$(\sigma_{xx}^x)_{i+1/2,j}^{n+1/2} = OP(\sigma_{xx}^x)_{i+1/2,j}^{n-1/2} + OD(V_x)_{i,j}^n + O\Delta t(s_{xx}^x)^{n-1/2} \tag{42}$$
$$(\sigma_{xx}^z)_{i+1/2,j}^{n+1/2} = QR(\sigma_{xx}^z)_{i+1/2,j}^{n-1/2} + QE(V_z)_{i+1/2,j+1/2}^n + Q\Delta t(s_{xx}^z)^{n-1/2} \tag{43}$$
$$(\sigma_{zz}^x)_{i+1/2,j}^{n+1/2} = OP(\sigma_{xx}^x)_{i+1/2,j}^{n-1/2} + OF(V_x)_{i,j}^n + O\Delta t(s_{zz}^x)^{n-1/2} \tag{44}$$
$$(\sigma_{zz}^z)_{i+1/2,j}^{n+1/2} = QR(\sigma_{xx}^z)_{i+1/2,j}^{n-1/2} + QG(V_z)_{i+1/2,j+1/2}^n + Q\Delta t(s_{zz}^z)^{n-1/2} \tag{45}$$
$$(\sigma_{xz}^x)_{i,j+1/2}^{n+1/2} = KL(\sigma_{xz}^x)_{i,j+1/2}^{n-1/2} + KH(V_x)_{i,j}^n + K\Delta t(s_{xz}^x)^{n-1/2} \tag{46}$$
$$(\sigma_{xz}^z)_{i,j+1/2}^{n+1/2} = MN(\sigma_{xz}^z)_{i,j+1/2}^{n-1/2} + MJ(V_z)_{i+1/2,j+1/2}^n + M\Delta t(s_{xz}^z)^{n-1/2}, \tag{47}$$

where the operators **A** through **J** have been defined previously and the PML operators are given by

$$\begin{array}{lll} K = (1 + \frac{\Delta t}{2} d_i^x)^{-1} & L = (1 - \frac{\Delta t}{2} d_i^x) & M = (1 + \frac{\Delta t}{2} d_j^z)^{-1} \\ N = (1 - \frac{\Delta t}{2} d_j^z) & O = (1 + \frac{\Delta t}{2} d_{i+1/2}^x)^{-1} & P = (1 - \frac{\Delta t}{2} d_{i+1/2}^x) \\ Q = (1 + \frac{\Delta t}{2} d_{j+1/2}^z)^{-1} & R = (1 - \frac{\Delta t}{2} d_{j+1/2}^z)\ . & \end{array}$$

Now that I have represented each equation as an idependent recursive relation, I need to describe the full data set as one recursive relation. Again, I will derive the necessary equations for both methods described earlier.

PML for the recursive operator by backward substitution

Similarly to the case without PML, I start by writing my wave fields as two data vectors,

$$\mathbf{d_1^n} = \begin{pmatrix} (V_x^x)^n \\ (V_x^z)^n \\ (V_z^x)^n \\ (V_z^z)^n \end{pmatrix} \qquad \mathbf{d_2^n} = \begin{pmatrix} (\sigma_{xx}^x)^{n-1/2} \\ (\sigma_{xx}^z)^{n-1/2} \\ (\sigma_{zz}^x)^{n-1/2} \\ (\sigma_{zz}^z)^{n-1/2} \\ (\sigma_{xz}^x)^{n-1/2} \\ (\sigma_{xz}^z)^{n-1/2}. \end{pmatrix} \tag{48}$$

The recursive relations then become

$$\mathbf{d_1^n} = \begin{pmatrix} KL & 0 & 0 & 0 \\ 0 & MN & 0 & 0 \\ 0 & 0 & OP & 0 \\ 0 & 0 & 0 & QR \end{pmatrix} \mathbf{d_1^{n-1}} + \begin{pmatrix} KA & KA & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & MB & MB \\ 0 & 0 & 0 & 0 & OA & OA \\ 0 & 0 & QB & QB & 0 & 0 \end{pmatrix} \mathbf{d_2^n} + \begin{pmatrix} KC & 0 & 0 & 0 \\ 0 & MC & 0 & 0 \\ 0 & 0 & OC & 0 \\ 0 & 0 & 0 & QC \end{pmatrix} \mathbf{S_1^{n-1}} \quad (49)$$

$$\mathbf{d_2^n} = \begin{pmatrix} OD & OD & 0 & 0 \\ 0 & 0 & QE & QE \\ OF & OF & 0 & 0 \\ 0 & 0 & QG & QG \\ KH & KH & 0 & 0 \\ 0 & 0 & MJ & MJ \end{pmatrix} \mathbf{d_1^{n-1}} + \begin{pmatrix} OP & 0 & 0 & 0 & 0 & 0 \\ 0 & QR & 0 & 0 & 0 & 0 \\ 0 & 0 & OP & 0 & 0 & 0 \\ 0 & 0 & 0 & QR & 0 & 0 \\ 0 & 0 & 0 & 0 & KL & 0 \\ 0 & 0 & 0 & 0 & 0 & MN \end{pmatrix} \mathbf{d_2^{n-1}} + \begin{pmatrix} O\Delta t & 0 & 0 & 0 & 0 & 0 \\ 0 & Q\Delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & O\Delta t & 0 & 0 & 0 \\ 0 & 0 & 0 & Q\Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & K\Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & M\Delta t \end{pmatrix} \mathbf{S_2^{n-1}}. \quad (50)$$

Substituting 10 into 12, I get

$$\mathbf{d_1^n} = \begin{pmatrix} KL+KAOD & KAOD & KAQE & KAQE \\ MBKH & MN+MBKH & MBMJ & MBMJ \\ OAKH & OAKH & OP+OAMJ & OAMJ \\ QBOF & QBOF & QBQG & QR+QBQG \end{pmatrix} \mathbf{d_1^{n-1}} +$$

$$\begin{pmatrix} KAOP & KAQR & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & MBKL & MBMN \\ 0 & 0 & 0 & 0 & OAKL & OAMN \\ 0 & 0 & QBOP & QBQR & 0 & 0 \end{pmatrix} \mathbf{d_2^{n-1}} +$$

$$\begin{pmatrix} KC & 0 & 0 & 0 \\ 0 & MC & 0 & 0 \\ 0 & 0 & OC & 0 \\ 0 & 0 & 0 & QC \end{pmatrix} \mathbf{S_1^{n-1}} +$$

$$\begin{pmatrix} KAO\Delta t & KAQ\Delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & MBK\Delta t & MBM\Delta t \\ 0 & 0 & 0 & 0 & OAK\Delta t & OAM\Delta t \\ 0 & 0 & QBO\Delta t & QBQ\Delta t & 0 & 0 \end{pmatrix} \mathbf{S_2^{n-1}}. \quad (51)$$

Finally, I can construct the recursive operator $\mathbf{W}$

$$\mathbf{d^n} = \begin{pmatrix} d_1^n \\ d_2^n \end{pmatrix} = \begin{pmatrix} KL+KAOD & KAOD & KAQE & KAQE & KAOP & KAQR & 0 & 0 & 0 & 0 \\ MBKH & MN+MBKH & MBMJ & MBMJ & 0 & 0 & 0 & 0 & MBKL & MBMN \\ OAKH & OAKH & OP+OAMJ & OAMJ & 0 & 0 & 0 & 0 & OAKL & OAMN \\ QBOF & QBOF & QBQG & QR+QBQG & 0 & 0 & QBOP & QBQR & 0 & 0 \\ OD & OD & 0 & 0 & OP & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & QE & QE & 0 & QR & 0 & 0 & 0 & 0 \\ OF & OF & 0 & 0 & 0 & 0 & OP & 0 & 0 & 0 \\ 0 & 0 & QG & QG & 0 & 0 & 0 & QR & 0 & 0 \\ KH & KH & 0 & 0 & 0 & 0 & 0 & 0 & KL & 0 \\ 0 & 0 & MJ & MJ & 0 & 0 & 0 & 0 & 0 & MN \end{pmatrix} \mathbf{d^{n-1}} +$$

$$\begin{pmatrix} KC & 0 & 0 & 0 & KAO\Delta t & KAQ\Delta t & 0 & 0 & 0 & 0 \\ 0 & MC & 0 & 0 & 0 & 0 & 0 & 0 & MBK\Delta t & MBM\Delta t \\ 0 & 0 & OC & 0 & 0 & 0 & 0 & 0 & OAK\Delta t & OAM\Delta t \\ 0 & 0 & 0 & QC & 0 & 0 & QBO\Delta t & QBQ\Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & O\Delta t & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & Q\Delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & O\Delta t & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & Q\Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & K\Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & M\Delta t \end{pmatrix} \begin{pmatrix} S_1^n \\ S_2^n \end{pmatrix} \tag{52}$$

PML for the recursive operator by time step refinement

Lastly, the PML implementation in the case of the refined time stepping method is

$$
\begin{pmatrix} (V_x^x)^n \\ (V_x^z)^n \\ (V_z^x)^n \\ (V_z^z)^n \\ (\sigma_{xx}^x)^n \\ (\sigma_{xx}^z)^n \\ (\sigma_{zz}^x)^n \\ (\sigma_{zz}^z)^n \\ (\sigma_{xz}^x)^n \\ (\sigma_{xz}^z)^n \end{pmatrix} =
\begin{pmatrix}
0 & 0 & 0 & 0 & KA & KA & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & MB & MB \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & OA & OA \\
0 & 0 & 0 & 0 & 0 & 0 & QB & QB & 0 & 0 \\
OD & OD & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & QE & QE & 0 & 0 & 0 & 0 & 0 & 0 \\
OF & OF & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & QG & QG & 0 & 0 & 0 & 0 & 0 & 0 \\
KH & KH & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & MJ & MJ & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix} \mathbf{d^{n-\frac{1}{2}}} +
$$

$$
\begin{pmatrix}
KL & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & MN & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & OP & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & QR & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & OP & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & QR & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & OP & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & QR & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & KL & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & MN
\end{pmatrix} \mathbf{d^{n-1}} +
$$

$$
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & O\Delta t & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & Q\Delta t & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & O\Delta t & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & Q\Delta t & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & K\Delta t & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & M\Delta t
\end{pmatrix} \mathbf{S^{n-\frac{1}{2}}} +
$$

$$
\begin{pmatrix}
KC & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & MC & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & OC & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & QC & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix} \mathbf{S^{n-1}}, \quad (53)
$$

which can be compactly described by the recursive relation

$$
\mathbf{d^n} = \mathbf{X d^{n-\frac{1}{2}}} + \mathbf{Z d^{n-1}} + \mathbf{Y_1 S^{n-\frac{1}{2}}} + \mathbf{Y_2 S^{n-1}}. \quad (54)
$$

NONLINEAR MODELING

Theoretically, both recursive operators defined in the previous sections could be used to construct a linear operator with respect to the source function. However, the first method yields a more cumbersome system, which would only grow in complexity as more elements are added. Therefore, in the next sections I focus only on the recursive operator by time stepping refinement. I start with a general linear operator summarized by

$$\mathbf{d} = \mathbf{F}\mathbf{s}, \tag{55}$$

where $\mathbf{d}$ is the output data, $\mathbf{F}$ is the forward modeling operator and $\mathbf{s}$ is the source.

Next, I include the necessary interpolation and padding operators, similarl to the work in Almomin (2013),

$$\mathbf{d} = \mathbf{K_r'}\mathbf{L_r'}\mathbf{W}\mathbf{Y}\mathbf{L_s}\mathbf{K_s}\mathbf{s}, \tag{56}$$

where $\mathbf{K_r}$ and $\mathbf{L_s}$ are the spatial padding and time interpolation operator, respectively. $\mathbf{W}$ is the recursive operator described by equation 54. Similarly, the adjoint modeling operator is given by

$$\mathbf{s} = \mathbf{K_s'}\mathbf{L_s'}\mathbf{Y}'\mathbf{W}'\mathbf{L_r}\mathbf{K_r}\mathbf{d}, \tag{57}$$

where $\mathbf{W}'$ is the recursive relation described by

$$\mathbf{q^n} = \mathbf{X}'\mathbf{q^{n+\frac{1}{2}}} + \mathbf{Z}'\mathbf{q^{n+1}} + \mathbf{d^{n+1}}, \tag{58}$$

where $\mathbf{q^n}$ is the scaled source

$$\mathbf{q^n} = \mathbf{Y_1}\mathbf{s^{n+\frac{1}{2}}} + \mathbf{Y_2}\mathbf{s^n}. \tag{59}$$

The adjoint of matrix $\mathbf{X}$ in the adjoint recursive relation $\mathbf{W}'$ is

$$\mathbf{X}' = \begin{pmatrix} 0 & 0 & 0 & 0 & D'O' & 0 & F'O' & 0 & H'K' & 0 \\ 0 & 0 & 0 & 0 & D'O' & 0 & F'O' & 0 & H'K' & 0 \\ 0 & 0 & 0 & 0 & 0 & E'Q' & 0 & G'Q' & 0 & J'M' \\ 0 & 0 & 0 & 0 & 0 & E'Q' & 0 & G'Q' & 0 & J'M' \\ A'K' & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ A'K' & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B'Q' & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B'Q' & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & B'M' & A'O' & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & B'M' & A'O' & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{60}$$

and the adjoints of matrices $\mathbf{Z}$, $\mathbf{Y_1}$ and $\mathbf{Y_2}$ are

$$\mathbf{Z}' = \begin{pmatrix} L'K' & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & N'M' & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & P'O' & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & R'Q' & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & P'O' & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & R'Q' & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & P'O' & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & R'Q' & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & L'K' & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & N'M' \end{pmatrix}, \tag{61}$$

$$\mathbf{Y_1'} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta t O' & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Delta t Q' & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \Delta t O' & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Delta t Q' & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Delta t K' & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Delta t M' \end{pmatrix}, \tag{62}$$

$$\mathbf{Y_2'} = \begin{pmatrix} C'K' & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & C'M' & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & C'O' & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & C'Q' & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{63}$$

The individual adjoint operators are

$$A' = \frac{\partial}{\partial x}\frac{\Delta t}{\rho} \qquad B' = \frac{\partial}{\partial z}\frac{\Delta t}{\rho} \qquad C' = \frac{\Delta t}{\rho}$$

$$D' = \frac{\partial}{\partial x}\Delta t(\lambda + 2\mu) \qquad E' = \frac{\partial}{\partial z}\Delta t\lambda \qquad F' = \frac{\partial}{\partial z}\Delta t(\lambda + 2\mu)$$

$$G' = \frac{\partial}{\partial x}\Delta t\lambda \qquad H' = \frac{\partial}{\partial x}\Delta t\mu \qquad J' = \frac{\partial}{\partial z}\Delta t\mu$$

$$K' = (1 + \tfrac{\Delta t}{2}d_i^x)^{-1} \qquad L' = (1 - \tfrac{\Delta t}{2}d_i^x) \qquad M' = (1 + \tfrac{\Delta t}{2}d_j^z)^{-1}$$

$$N' = (1 - \tfrac{\Delta t}{2}d_j^z) \qquad O' = (1 + \tfrac{\Delta t}{2}d_{i+1/2}^x)^{-1} \qquad P' = (1 - \tfrac{\Delta t}{2}d_{i+1/2}^x)$$

$$Q' = (1 + \tfrac{\Delta t}{2}d_{j+1/2}^z)^{-1} \qquad R' = (1 - \tfrac{\Delta t}{2}d_{j+1/2}^z)\ .$$

I apply the dot product test to validate that the adjoint of the recursive relation is correct. I test the case where the PML coefficient is equal to one (rigid boundary condition) and calculate the forward and adjoints for 5,000 time steps. For a 2D grid with 200 by 200 samples, the relative error in the dot product test is on the order of 10^{-12}. I run the tests in double precision, with random inputs for the forward and adjoint solutions. This method follows the one described in Claerbout (2010).

BORN OPERATOR

The Born method is a linear approximation to the non-linear wave propagation problem. Essentially, it aims to describe the non-linear problem in terms of a perturbation problem, where the solution is a combination of a non-scattering background wavefield and a

scattering term. Such an approximation is valid for small perturbations, where secondary scattering events are small compared to the background and first order scattering terms. To derive the Born approximation of the elastic wave equation, I start by writing the initial set of equations for a perturbed properties model,

$$(\rho+\Delta\rho)\frac{\partial}{\partial t}(V_x+\Delta V_x)-[\frac{\partial}{\partial x}(\sigma_{xx}+\Delta\sigma_{xx})+\frac{\partial}{\partial z}(\sigma_{xz}+\Delta\sigma_{xz})]=s_1 \tag{64}$$

$$(\rho+\Delta\rho)\frac{\partial}{\partial t}(V_z+\Delta V_z)-[\frac{\partial}{\partial x}(\sigma_{xz}+\Delta\sigma_{xz})+\frac{\partial}{\partial z}(\sigma_{zz}+\Delta\sigma_{zz})]=s_2 \tag{65}$$

$$\frac{\partial}{\partial t}(\sigma_{xx}+\Delta\sigma_{xx})=[(\lambda+\Delta\lambda)+2(\mu+\Delta\mu)]\frac{\partial}{\partial x}(V_x+\Delta V_x)+$$
$$(\lambda+\Delta\lambda)\frac{\partial}{\partial z}(V_z+\Delta V_z)+s_3 \tag{66}$$

$$\frac{\partial}{\partial t}(\sigma_{zz}+\Delta\sigma_{zz})=[(\lambda+\Delta\lambda)+2(\mu+\Delta\mu)]\frac{\partial}{\partial z}(V_z+\Delta V_z)+$$
$$(\lambda+\Delta\lambda)\frac{\partial}{\partial x}(V_x+\Delta V_x)+s_4 \tag{67}$$

$$\frac{\partial}{\partial t}(\sigma_{xz}+\Delta\sigma_{xz})=(\mu+\Delta\mu)(\frac{\partial}{\partial x}(V_z+\Delta V_z)+\frac{\partial}{\partial z}(V_x+\Delta V_x))+s_5. \tag{68}$$

Rearranging equation 64, I get

$$\rho\frac{\partial}{\partial t}V_x-[\frac{\partial}{\partial x}\sigma_{xx}+\frac{\partial}{\partial z}\sigma_{xz}]+\rho\frac{\partial}{\partial t}\Delta V_x+\Delta\rho\frac{\partial}{\partial t}V_x+\Delta\rho\frac{\partial}{\partial t}\Delta V_x-$$
$$[\frac{\partial}{\partial x}\Delta\sigma_{xx}+\frac{\partial}{\partial z}\Delta\sigma_{xz}]=s_1, \tag{69}$$

which is essentially equation 1 plus perturbation terms. If I ignore higher order perturbation terms and substract the original wave equation, equation 69 gives me the Born scattering term

$$\rho\frac{\partial}{\partial t}\Delta V_x-(\frac{\partial}{\partial x}\Delta\sigma_{xx}+\frac{\partial}{\partial z}\Delta\sigma_{xz}]=-\Delta\rho\frac{\partial}{\partial t}V_x. \tag{70}$$

Similarly, for the other equations I get

$$\rho\frac{\partial}{\partial t}\Delta V_z-(\frac{\partial}{\partial x}\Delta\sigma_{xz}+\frac{\partial}{\partial z}\Delta\sigma_{zz}]=-\Delta\rho\frac{\partial}{\partial t}V_z \tag{71}$$

$$\frac{\partial}{\partial t}\Delta\sigma_{xx}-[(\lambda+2\mu)\frac{\partial}{\partial x}\Delta V_x+\lambda\frac{\partial}{\partial z}\Delta V_z]=(\Delta\lambda+2\Delta\mu)\frac{\partial}{\partial x}V_x+\Delta\lambda\frac{\partial}{\partial z}V_z \tag{72}$$

$$\frac{\partial}{\partial t}\Delta\sigma_{zz}-[(\lambda+2\mu)\frac{\partial}{\partial z}\Delta V_z+\lambda\frac{\partial}{\partial x}\Delta V_x]=(\Delta\lambda+2\Delta\mu)\frac{\partial}{\partial z}V_z+\Delta\lambda\frac{\partial}{\partial x}V_x \tag{73}$$

$$\frac{\partial}{\partial t}\Delta\sigma_{xz}-[\mu(\frac{\partial}{\partial x}\Delta V_z+\frac{\partial}{\partial z}\Delta V_x)]=\Delta\mu(\frac{\partial}{\partial x}V_z+\frac{\partial}{\partial z}V_x). \tag{74}$$

The numerical implementation of the forward Born modeling can be broken into two steps. First, a virtual source is constructed by the forward propagation of the elastic wave fields in a background (unperturbed) properties model. I describe this background model by $\rho_{\mathbf{0}}$, $\lambda_{\mathbf{0}}$ and $\mu_{\mathbf{0}}$. This background wave field is constructed similarly to equation 56, but without the need to truncate the output to the data spatial sampling and with its time sampling equal to that of the source,

$$\mathbf{d_0}=\mathbf{L_s'WYL_sK_sf}. \tag{75}$$

After generating the background wave field, I use it as a virtual source to generate a new data set. This new data set, which I call the scattered data ($\mathbf{\Delta d}$), is given by

$$\mathbf{\Delta d} = \mathbf{K_r' L_r' W Y L_s Y \Theta d_0 r}, \tag{76}$$

where $\mathbf{\Theta}$ is a derivative operator that needs to be applied to the background wave field $\mathbf{d_0}$ and $\mathbf{r}$ is the perturbed properties model. The matrices that represent these operators are

$$\mathbf{\Theta} = \begin{pmatrix} \frac{\partial}{\partial t} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\partial}{\partial t} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial}{\partial t} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\partial}{\partial t} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial x} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial z} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial x} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial z} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\partial}{\partial z} & \frac{\partial}{\partial z} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial}{\partial x} & \frac{\partial}{\partial x} & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{77}$$

$$\mathbf{r} = \begin{pmatrix} -\Delta\rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\Delta\rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\Delta\rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\Delta\rho & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta\lambda + 2\Delta\mu & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Delta\lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \Delta\lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Delta\lambda + 2\Delta\mu & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Delta\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Delta\mu \end{pmatrix}. \tag{78}$$

A few observations should be made about equations 77 and 78. First, the operator $\mathbf{\Theta}$ has non-zero terms only for the particle velocity components. This means that the virtual sources for velocities and stresses, at least for this particular formulation, require only the velocity components from the background wavefield, which are then scaled correctly by the appropriate derivative. Second, $\mathbf{r}$ is a diagonal operator that has terms that depend not only on the perturbed density and each Lamé parameter individually, but also on a combination of the two Lamé parameters. If used in an iterative inversion method, this mixed term could lead to strong crosstalk between the gradient updates, in addition to the expected crosstalk from the equations coupling. However, this topic still needs further study.

RESULTS

I implement the Born forward modeling using a time domain, finite difference algorithm. The algorithm uses a second order approximation to the time derivatives and a 10^{th} order operator for the spatial derivatives, as described in Alves and Biondi (2014).

The model is a constant background with a gaussian perturbation in all model properties (λ, μ and ρ), as can be seen in figure 1(a).

Figure 2(a) shows a snapshot of the pressure wave field for the nonlinear modeling. In other words, this figure is generated by propagating the wave field in the correct gaussian model. For the Born modeling, I smooth the inital model (1(b)) and subtract it from the correct model (1(c)). This generates two wave fields: a background propagated one, shown in figure 2(b) and a scattered wavefield, shown in figure 2(c). All wave fields are scaled equally in order to compare the Born linear approximation and the non-linear modeling.

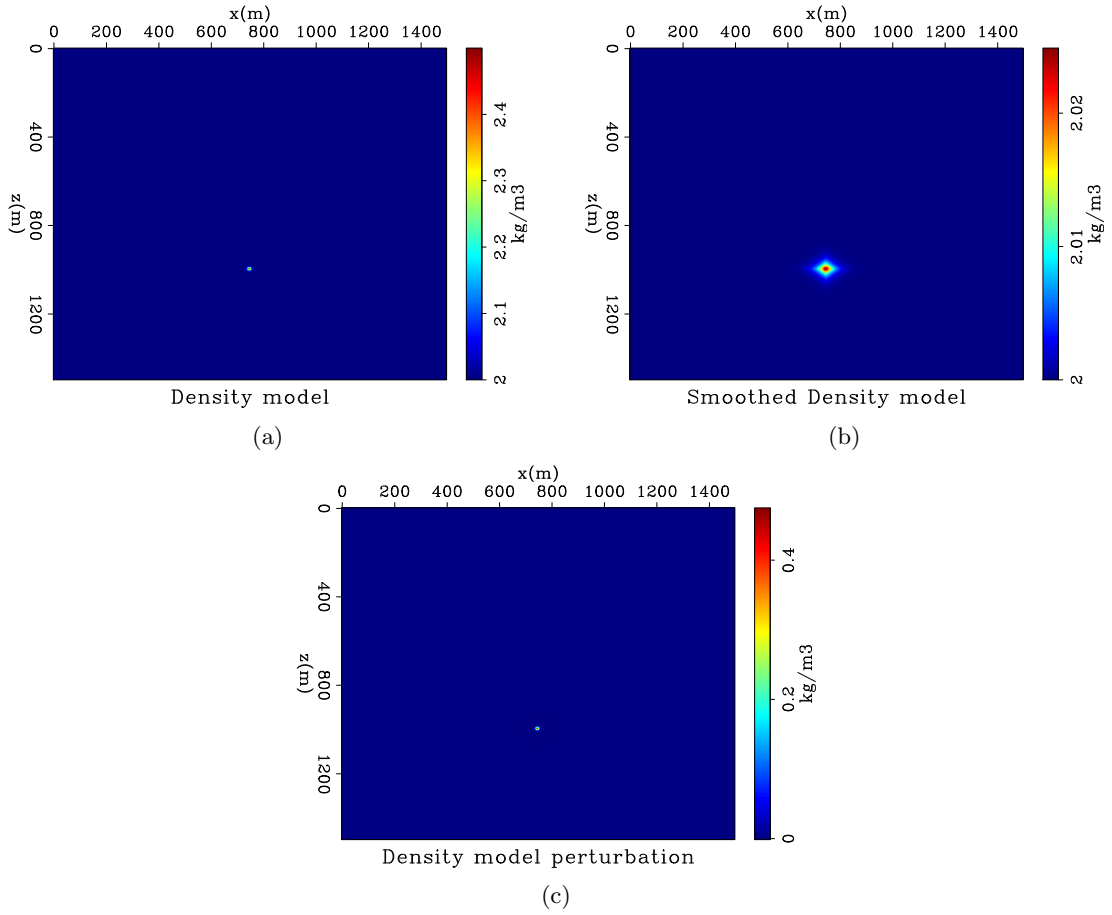


Figure 1: (a) True density model with a gaussian perturbation, (b) smoothed density model and (c) model difference between true and smoothed model. The other parameters display similar perturbations. [**ER**] gcalves/. rhomodel,rhomodelS,drhomodel

As can be seen from the examples, the Born modeling approximates the non-linear propagation not only qualitatively, but also quantitatively. It is important to point out that, while it is more customary to work with the pressure and shear velocities as model parameters for imaging, the formulation presented here is solved in terms of Lamè parameters.

CONCLUSIONS

The recursive operator by backward substitution might be computationally more efficient, due to the smaller number of time steps required, when compared to the recursive operator by time step refinement. However, the long chains of operators required in the former make

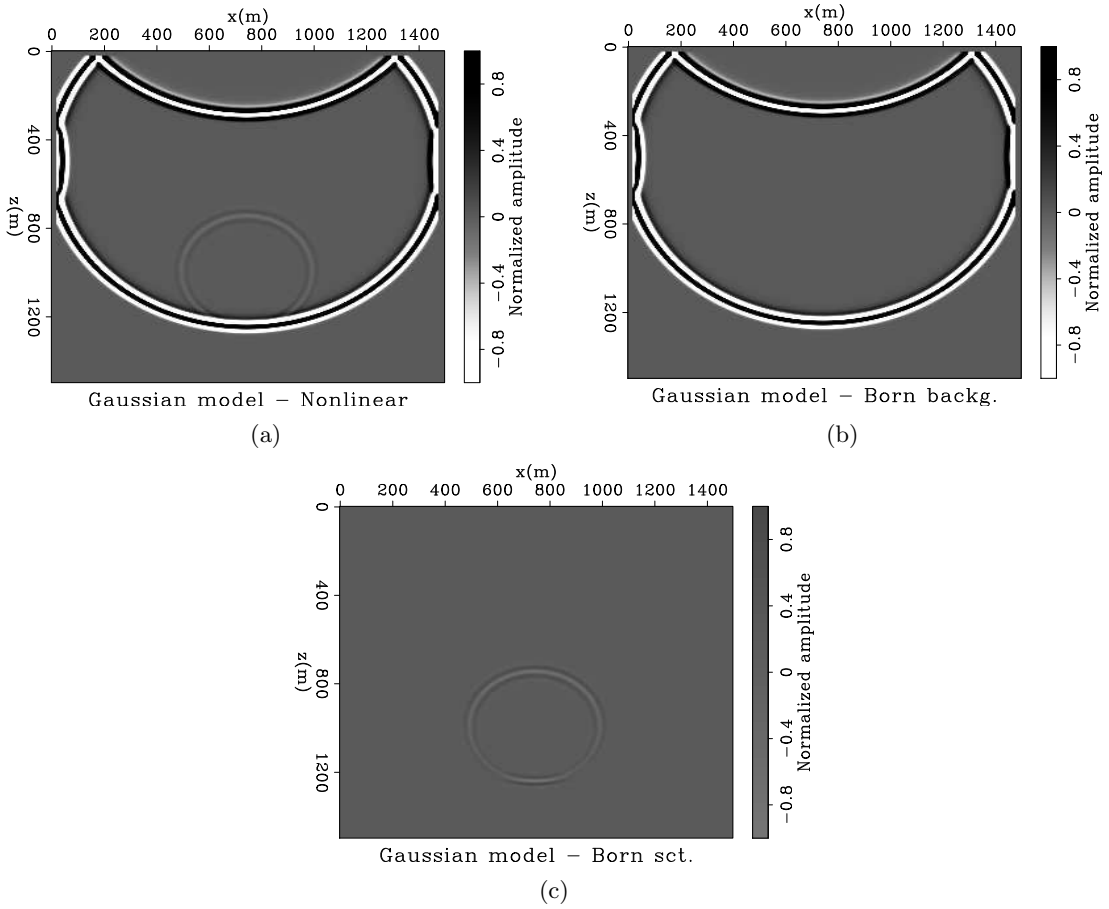


Figure 2: Snapshots of the pressure wavefield for (a) the non-linear modeling, (b) background modeling with the Born approximation and (c) scattering modeling with the Born approximation. [**ER**] gcalves/. gauss-nonlinear,gauss-born-bg,gauss-born-sct

its implementation more difficult.

The recursive operator by time refinement is simpler to implement and is shown to be stable both in its forward and adjoint formulations.

The PML implementation is efficient, but requires twice as many wavefield computations due to the required component splitting. When applying the proposed scheme to problems where computational efficiency is required, the PML method should be limited to the absorbing boundaries and not to the whole model domain.

The formulation presented here is the first step in describing elastic inversion schemes in the framework of the adjoint state method. While this is not the only possible approach to such schemes, this formulation is closer to those developed in the Stanford Exploration Project for acoustic methods. Such similarity allows for an easier comparison between acoustic and elastic methods.

REFERENCES

Aki, K. and P. G. Richards, 1980, Quantitative seismology: Theory and methods, 1: I: WH Freeman and Co.

Almomin, A., 2013, Accurate implmentation of two-way wave-equation operators: Stanford Exploration Project Report, **149**, 279–286.

Alves, G. and B. Biondi, 2014, High-order elastic finite-difference modeling: Stanford Exploration Project Report, **152**, 327–340.

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.

Claerbout, J. F., 2010, Basic earth imaging: Retrieved from http://sepwww.stanford.edu/sep/prof/bei11.2010.pdf.

Collino, F. and C. Tsogka, 2001, Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media: Geophysics, **66**, 294–307.

Levander, A. R., 1988, Fourth-order finite-difference P-SV seismograms: Geophysics, **53**, 1425–1436.

Moczo, P., J. Kristek, V. Vavryčuk, R. J. Archuleta, and L. Halada, 2002, 3D heterogeneous staggered-grid finite-difference modeling of seismic motion with volume harmonic and arithmetic averaging of elastic moduli and densities: Bulletin of the Seismological Society of America, **92**, 3042–3066.

Virieux, J., 1986, P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method: Geophysics, **51**, 889–901.

Reverse-time migration using the rapid expansion method (REM)

Alejandro Cabrales-Vargas

ABSTRACT

Reverse-time migration using the rapid expansion method allows the use of coarse time steps, and can be used as an alternative to high-order finite difference schemes. Its implementation using the pseudospectral method avoids the occurrence of frequency dispersion artifacts, and can be easily adapted to handling anisotropy. I use random boundary conditions to reduce memory usage.
Tests in two-dimensional synthetic models show the ability of this method to handle strong velocity and anisotropy contrasts in the presence of complex geology, regardless of the time resolution of the data.

INTRODUCTION

In recent years, reverse-time migration (RTM) has become a widely used imaging tool to produce high-quality images of the subsurface, primarily as a consequence of the petroleum exploration targeting ever more complex geological settings. Strong velocity contrasts, high-dip events, subsalt targets, and prismatic waves, among many other issues, constitute some of the challenges that seismic imagers have to deal with in their efforts to deliver the best possible subsurface images to seismic interpreters.

Although the first references about RTM can be traced back to the mid-70s, people usually refer to Baysal et al. (1983), Kosloff and Baysal (1983), Gazdag and Carrizo (1986), and McMechan (1983), among others, who set the basis for the post-stack RTM scenario. Meanwhile, Claerbout (1971) proposed the cross-correlation imaging condition to construct the subsurface image from downward- and upward-continued wavefields. This idea was later applied to forward and backward in time wavefield modeling, thereby constituting the basis for prestack RTM (Chang and McMechan, 1990; Loewenthal and Hu, 1991).

Solving the wave equation using implicit methods is prohibitive in practice for three-dimensional datasets (Claerbout, 1985). Therefore, geophysicists have opted for implementing explicit methods, accounting for keeping stability criteria, and avoiding frequency dispersion. One technique proposed to relax stability criteria and frequency dispersion restrictions is the rapid expansion method (REM) (Tal-Ezer et al., 1987; Etgen, 1989; Kosloff et al., 1989; Jastram and Behle, 1991; Stoffa and Pestana, 2009; Tessmer, 2011). This approximation is more accurate than the truncated Taylor series based methods. It has been generally proposed in conjunction with a pseudospectral scheme, thus avoiding the problem of frequency dispersion. Nevertheless, it is worth mentioning that Jastram and Behle (1991) implemented the REM using staggered grids to solve elastic wave propagation.

Forward wavefield propagation generally consists of solving the two-way wave equation forward in time, and injecting a source signature at appropriate locations. Conversely,

backward propagation consists of solving the same two-way wave equation, now backward in time, and injecting the receiver dataset of a single shot gather. The imaging condition is the zero-lag cross-correlation of such wavefields in the time domain for every shot profile. The final image is obtained by the superposition of the migrated shots. However, the cross-correlation imaging condition not only constructs reflection events related to the geology contrasts, but also low-frequency artifacts unique in RTM, caused by the cross-correlation of wavefronts propagating in the same direction in both wavefields. One solution proposed to tackle this problem is computing the Poynting vector (Yoon and Marfurt, 2006) to attenuate such artifacts by controling the angles of propagation of the wavefronts that are cross-correlated. Alternatively, Liu et al. (2011) proposed wavefield separation, recasting the imaging condition to perform the cross-correlation of wavefronts propagating in opposite directions exclusively. I implemented the latter solution to produce the RTM images for this report.

One problem associated with the cross-correlation of the source and the receiver wavefields is that every spatial frame (`x`,`z` plane in 2-D; `x`,`y`,`z` volume in 3-D) for all time steps has to be available, either loaded into memory (what easily overloads it), or stored on disk (which can make the I/O process rather slow). The reason is because the source and receiver wavefields are computed sequencially, forward and backward in time, respectively. Proposed solutions commonly involve checkpoint strategies (Symes, 2007). In this report I use random boundary conditions (RBC) (Clapp, 2009) to overcome this limitation, and compare the results to more conventional tapering boundary conditions (TBC) (Cerjan et al., 1985).

In the first part of this report, I present the formal derivation of the REM and the anisotropic solver (Zhan et al., 2012). Later, I present the application of the algorithm to 2D synthetic datasets: the Marmousi model with anisotropy (Alkhalifah, 1997), the Amoco 1997 2.5D model (Etgen and Regone, 1998; Dellinger et al., 2000), and the Hess VTI salt model. Then, I discuss the results and propose future steps. Finally, I present the conclusions of this report.

REVIEW OF THE RAPID EXPANSION METHOD THEORY

In this section I provide a review of the fundamentals of the REM.

Beginning with the 2D acoustic wave equation:

$$\frac{\partial^2 \mathbf{U}}{\partial t^2} = -\mathbf{L}^2\mathbf{U} + \mathbf{f}, \tag{1}$$

where $\mathbf{U}$ is the displacement field, $\mathbf{f}$ is the source term, the operator $-\mathbf{L}^2 = v(x,z)^2\nabla^2$ (the kernel), involving $v(x,z)$, which is the interval velocity, and the Laplacian operator

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}. \tag{2}$$

$\mathbf{L}^2$ constitutes a matrix in the physical domain, and a scalar in the Fourier domain. I refer to Equation 1 as the inhomogeneous wave equation because, in general, the source

term is not zero. When we explicitely set the source term equal to zero, Equation 1 becomes the homogeneous wave equation.

To perform RTM, we solve Equation 1 forward in time for the source wavefield, and backward in time for the receiver wavefield. Normally, source and receiver injection is employed to feed on the wavefields. However, using the REM we employ a different procedure to compute the source wavefield.

I find useful to set this problem from the point of view of linear algebra, because the set of solutions of the inhomogeneous wave equation constitutes a **linear variety**. It is represented as the superposition of the set of solutions for the homogeneous equation (the associated vector subspace), and one particular solution of the inhomogeneous equation (a known vector from the linear variety). The latter constitutes the period when the source is active, whereas the former constitutes the period after the source has shut down. This scheme is different to the source injection. On the other hand, the receiver wavefield backward propagation is conventionally solved with the homogeneous wave equation. There is no source in the sense of an independent term of the wave equation. Therefore, the receiver data are injected at every time step.

I elaborate more about these solutions in the following.

a) **Solution for the homogeneous equation**. This is the solution of Equation 1 without the source. We can propose a solution of the form (Stoffa and Pestana, 2009)

$$\mathbf{U}(t+\Delta t) = \mathbf{U}(t)\cos\mathbf{L}\Delta t + \frac{\partial \mathbf{U}(t)}{\partial t}\frac{\sin\mathbf{L}\Delta t}{\mathbf{L}}. \tag{3}$$

I dropped the spatial variables to simplify the notation. Now, we state the same solution backward in time:

$$\mathbf{U}(t-\Delta t) = \mathbf{U}(t)\cos\mathbf{L}\Delta t - \frac{\partial \mathbf{U}(t)}{\partial t}\frac{\sin\mathbf{L}\Delta t}{\mathbf{L}}. \tag{4}$$

and add it to Equation 3 to obtain:

$$\mathbf{U}(t+\Delta t) + \mathbf{U}(t-\Delta t) = 2\mathbf{U}(t)\cos\mathbf{L}\Delta t. \tag{5}$$

Equation 5 is structured as a three-point stencil to be solved either forward or backward in time. We need an appropriate approximation for the cosine term to prevent the occurrence of frequency dispersion. Kosloff et al. (1989) modified the original algorithm proposed by Tal-Ezer et al. (1987), thereby expanding the cosine term to obtain:

$$\mathbf{U}(t+\Delta t) + \mathbf{U}(t-\Delta t) = \sum_{k=0}^{M/2} C_{2k} J_{2k}(\Delta t R) Q_{2k}\left(\frac{i\mathbf{L}}{R}\right)\mathbf{U}(t), \tag{6}$$

where $C_0 = 1$ and $C_k = 2$ for $k > 0$, J_k are the Bessel functions of first kind of order k, Q_k are modified Chebyshev polynomials of order k given by the following recursive expressions:

$$Q_0\left(\frac{i\mathbf{L}}{R}\right) = 1, \tag{7}$$

$$Q_2\left(\frac{i\mathbf{L}}{R}\right) = 1 - \left(\frac{\mathbf{L}}{R}\right)^2, \tag{8}$$

$$Q_{k+2}\left(\frac{i\mathbf{L}}{R}\right) = \left[4\left(\frac{\mathbf{L}}{R}\right)^2 + 2\right]Q_k\left(\frac{i\mathbf{L}}{R}\right) - Q_{k-2}\left(\frac{i\mathbf{L}}{R}\right). \tag{9}$$

Finally, R is a constant that must be set larger than the largest eigenvalue of the operator $\mathbf{L}$ to ensure stability and convergence, and is given by:

$$R = v_{max}\pi\sqrt{\frac{1}{\Delta x^2} + \frac{1}{\Delta z^2}}. \tag{10}$$

Convergence is achieved by setting the number of terms of the series to $M > R\Delta t$.

As mentioned previously, Equation 6 can be solved both forward in time for the source wavefield and backward in time for the receiver wavefield, by isolating either $\mathbf{U}(t + \Delta t)$ or $\mathbf{U}(t - \Delta t)$, respectively. During the source wavefield modeling stage, we employ this solution for the period when the source is no longer active.

b) **Particular solution of the inhomogeneous equation**: This is the solution for the source wavefield propagation when the source term is active.

We first assume the source can be represented as:

$$\mathbf{f}(\mathbf{x}, t) = \mathbf{g}(\mathbf{x})h(t), \tag{11}$$

where $\mathbf{g}(\mathbf{x})$ is the spatial position of the source, usually represented as an impulse function in $\mathbf{x}$, and $h(t)$ is the source signature. Under this assumption, the formal solution of Equation 1 is (Jastram and Behle, 1991)

$$\mathbf{U}(\mathbf{x}, t) = \left[\int_0^t \frac{\sin(\tau\mathbf{L})}{\mathbf{L}}h(t - \tau)d\tau\right]\mathbf{g}(\mathbf{x}). \tag{12}$$

According to Stoffa and Pestana (2009), and Pestana and Stoffa (2010), we can approximate the solution as

$$\mathbf{U}(t, \mathbf{x}) = 2\sum_{k=1}^{M/2} b_{2k+1}(t)\frac{R}{i\mathbf{L}}Q_{2k+1}\left(\frac{i\mathbf{L}}{R}\right)\mathbf{g}(\mathbf{x}), \tag{13}$$

where

$$b_k(t) = \frac{1}{R}\int_0^t J_k(\tau R)h(t - \tau)d\tau. \tag{14}$$

The odd terms of the Chebyshev polynomials can be recursively obtained as (Chu and Stoffa, 2012):

$$\frac{R}{i\mathbf{L}}Q_1\left(\frac{i\mathbf{L}}{R}\right) = 1, \tag{15}$$

$$\frac{R}{i\mathbf{L}}Q_2\left(\frac{i\mathbf{L}}{R}\right) = 3 - 4\left(\frac{\mathbf{L}}{R}\right)^2, \tag{16}$$

$$\frac{R}{i\mathbf{L}}Q_{2k+1}\left(\frac{i\mathbf{L}}{R}\right) = 2\left[1 - 2\left(\frac{\mathbf{L}}{R}\right)^2\right]\frac{R}{i\mathbf{L}}Q_{2k-1}\left(\frac{i\mathbf{L}}{R}\right) - \frac{R}{i\mathbf{L}}Q_{2k-3}\left(\frac{i\mathbf{L}}{R}\right). \tag{17}$$

The factor $\frac{R}{i\mathbf{L}}$ is included to produce only real terms (even powers of $\frac{ii\mathbf{L}}{R}$).

When the source shuts down, the source wavefield propagation continues with Equation 6.

Remember that the number of terms in the summations (Equations 6 and 13) depends on the product $R\Delta t$. R is a function of grid spacing. The finer the grid, the larger R becomes, and so does the number of terms in the series. Therefore, a finer grid demands more terms for the summations, more computational time to perform the Fourier transformations, and more memory usage. On the other hand, reducing the time sampling makes the number of terms in the series smaller, but increases the time steps, thereby demanding more computational resources too. There appears to be no point about coarsening or refining the time axis, at least with regards to run time, although the proof of this statement is beyond the scope of this report. For practical purposes I keep the original time sampling of the data in all the examples.

ANISOTROPY AND RAPID EXPANSION METHOD: DECOUPLED EQUATIONS FOT TILTED TRANSVERSE ISOTROPY

I implement the anisotropy with the pseudodifferential tilted transverse isotropy (TTI) decoupled equations proposed by Zhan et al. (2012), which are fully compatible with the REM, and prevents the appearance of residual SV-wave artifacts in the image. I do not derive here the TTI decoupled equations (the reader is referred to Zhan et al. (2012) for an excellent and concise derivation), but I merely replicate the terms for two dimensions, to illustrate their application in this report:

$$A_0 = k_x^2 + k_z^2 \tag{18}$$

$$A_1 = (2\epsilon\cos^4\theta + 2\delta\sin^2\theta\cos^2\theta))\frac{k_x^4}{k_x^2 + k_z^2} \tag{19}$$

$$A_2 = (2\epsilon\sin^4\theta + 2\delta\sin^2\theta\cos^2\theta))\frac{k_z^4}{k_x^2 + k_z^2} \tag{20}$$

$$A_3 = (-4\epsilon\sin 2\theta\cos^2\theta + \delta\sin 4\theta)\frac{k_x^3 k_z}{k_x^2 + k_z^2} \tag{21}$$

$$A_4 = (-4\epsilon\sin 2\theta\sin^2\theta - \delta\sin 4\theta)\frac{k_x k_z^3}{k_x^2 + k_z^2} \tag{22}$$

$$A_5 = (3\epsilon\sin^2 2\theta - 2\delta\sin^2 2\theta + 2\delta\cos^2 2\theta)\frac{k_x^2 k_z^2}{k_x^2 + k_z^2}. \tag{23}$$

In the last expressions, ϵ and δ correspond to the Thomsen's parameters (Thomsen, 1986), and θ constitutes the dip field. With these functions we can define a new kernel as

$$-\mathbf{L}^2 = -v(x,z)^2\big[A_0 + A_1 + A_2 + A_3 + A_4 + A_5\big]. \tag{24}$$

I note that, other than A_0 (which constitutes the isotropic term) we do not always need apply all terms. In case of vertical transverse isotropy (VTI) media, we can drop all the sine terms as $\theta = 0$, retaining only shorter versions of A_1 and A_5. Likewise, in the presence of eliptical anisotropy ($\epsilon \neq 0, \delta = 0, \theta = 0$) only A_1 is required.

METHODOLOGY AND IMPLEMENTATION

I implement RTM using the REM using the pseudospectral method, with the TTI decoupled equations when anisotropy is present. I compare images obtained applying TBC (Cerjan et al., 1985) to images obtained applying RBC (Clapp, 2009). I implement them by extending the image space 30 points each side, where such boundary conditions are applied. I attenuate the low-frequency artifacts by separating the wavefields according to Liu et al. (2011), thereby correlating wavefronts that propagate in opposite directions, corresponding to seismic reflections.

The most time-consuming process in REM-based RTM is the series summation. Computing and applying every term of the series is impractical and inefficient, because Chebyshev polynomials of different order share the same powers of the kernel, $-\mathbf{L}^2$. Every second power of the kernel signifies one sequence of 1) 2D/3D Fourier transform in space; 2) multiplication by the kernel; 3) 2D/3D inverse Fourier transform in space; and 4) multiplication by the velocity model. A much better approach is to group the precomputed coefficients of both the Chebyshev polynomials and the Bessel functions, so the computation of every second power of the kernel is performed only once.

RESULTS

In this section I present the results of applying REM-based RTM to synthetic models.

The first model is the Amoco 1997 2.5D dataset (Etgen and Regone, 1998; Dellinger et al., 2000), which is particularly challenging because of its high geological complexity and strong velocity variability (top of Figure 1), and at the same time suitable to test robust imaging techniques. The RTM results using TBC and RBC are shown at the center and the bottom of Figure 1, respectively.

The second model is the anisotropic Marmousi dataset (Alkhalifah, 1997) (left side of Figure 2). The right side of Figure 2 shows the η field that I used instead of the ϵ field, while both δ and θ fields are set to zero (Alkhalifah, email communication). The results are shown in Figure 3.

The third model is the Hess VTI salt model (second version of the dataset obtained using the SEPlib) (Figure 4). I decimated the dataset to a grid of 160 ft by 160 ft to speed up the calculations. The results are shown in Figure 5. I compensated the amplitudes by applying a linear depth gain.

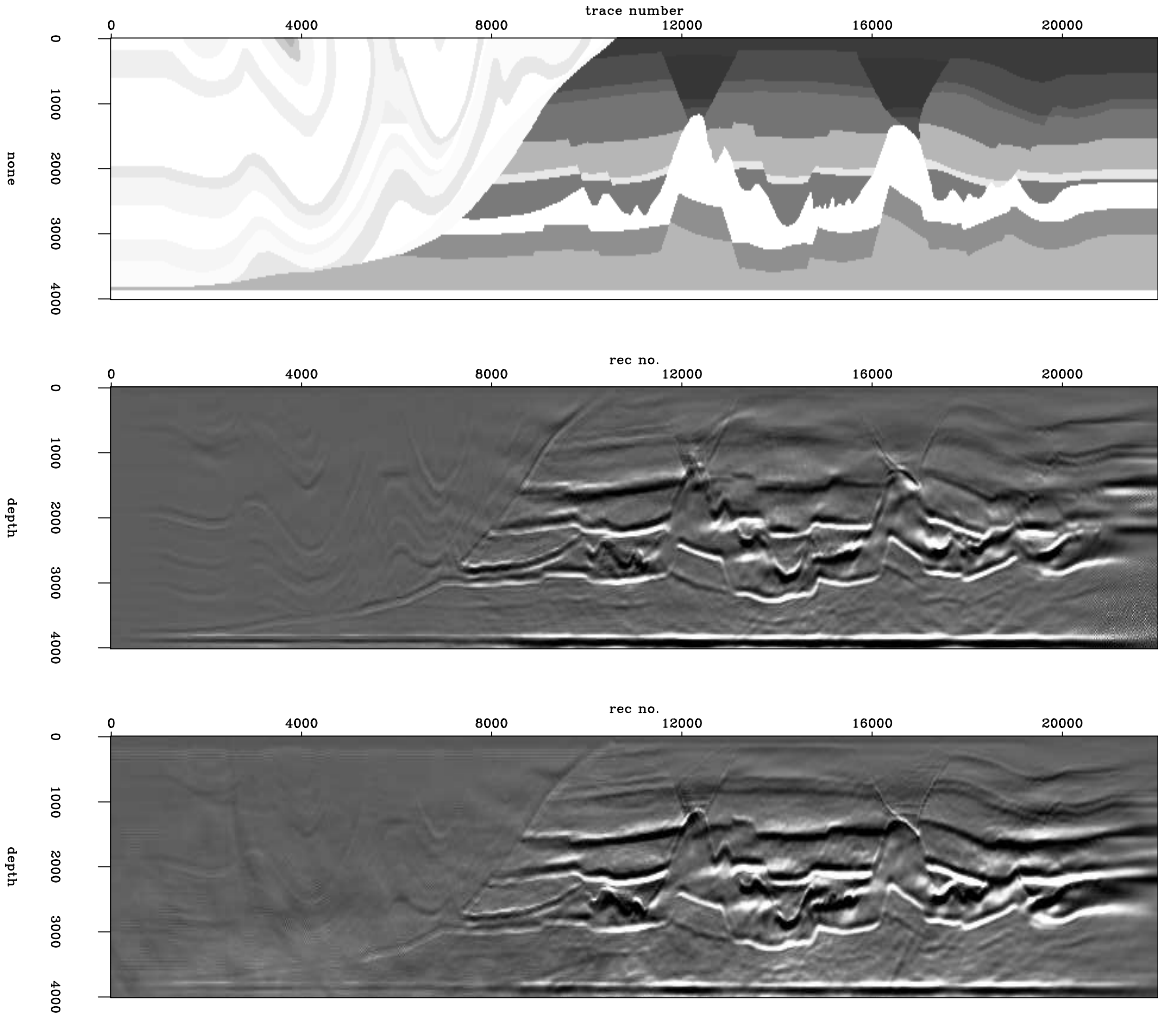


Figure 1: Amoco 1997 2.5D model. Top: Acoustic velocity field. Center: Reverse-time migration using tapering boundary conditions. Bottom: Reverse-time migration using random boundary conditions. alejandro/. 1997-stacks

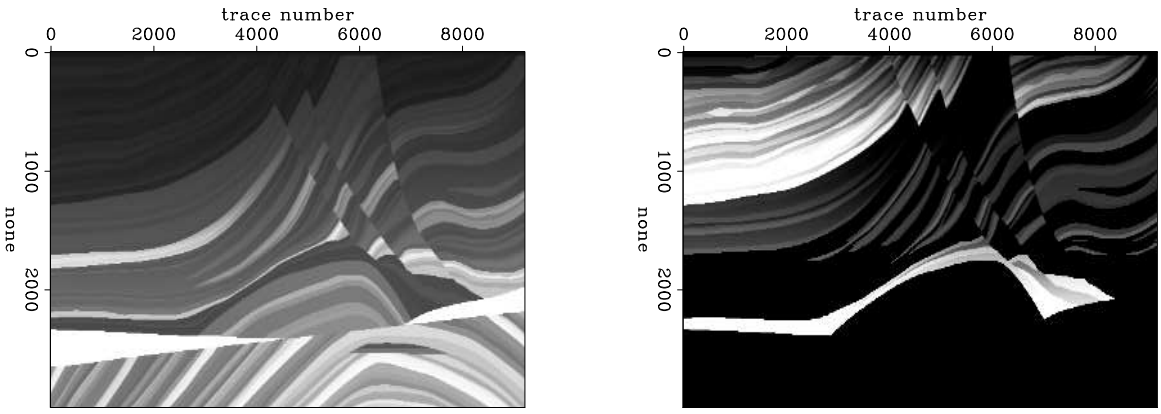


Figure 2: Anisotropic Marmousi model. Left: Acoustic velocity field. Right: η field. alejandro/. marmousi-ti-vels

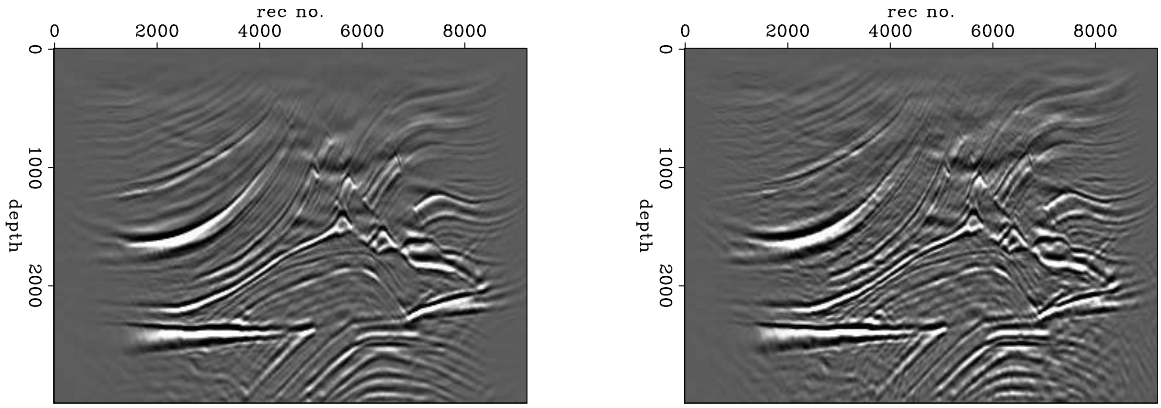


Figure 3: Reverse-time migration of the anisotropic Marmousi model. Left: Using tapering boundary conditions. Right: Using random boundary conditions. alejandro/. marmousi-ti-stacks

DISCUSSION

In general, the results in the synthetic models demonstrate that the implementation of REM-based RTM be useful as an alternative to high-order finite differences approximation, without requiring any refinement of the recording time sampling.

In the Amoco 1997 2.5D dataset, the REM was capable of handling the strong lateral velocity contrasts by adequately defining the complexity of the salt layer, and preserving the steep faults above it, and properly imaging the flat detachment at the bottom of the section. It is worth noting that the time sampling is coarser than usual: 9.9 ms. Nevertheless, the results obtained showed that this issue was automatically accounted for by incorporating enough terms to the expansion series. On the other hand, the anisotropic version of the algorithm correctly images the Marmousi model with good quality in the definition of the faults, the shallow reflectors, and the geologic features at the reservoir level. Likewise, the same algorithm adequately positions the salt body, the faults, and the reservoir reflections in the Hess VTI model. In all the migrated images there are still weak remnants of low-wavenumber artifacts.

Regarding the comparison between the boundary conditions, in the three datasets there is a tendency of the RBC to introduce both coherent and random noise, compared to the TBC. For instance, the Amoco 1997 2.5D RTM sections exhibit low-wavenumber noise and reverberating flat noise upon the the high-velocity hanging wall (on top of the thrust fault). The former appears to consist of remant low-wavenumber reflections that could not be "randomized" by the RBC. These different types of noise obscure this faint, low-amplitude part of the section (which has lower amplitude than the footwall in the original dataset). In contrast, the footwall exhibits some improvement details when using RBC: Layering around a 1 km depth is better defined, and there are sharper salt-sediment contrasts. Even part of the large-scale low-wavenumber shadow was attenuated in the footwall.

In the Marmousi RTM sections, we observe the inclusion of random noise when using RBC. There is neither apparent increase nor reduction of low-wavenumber artifacts.

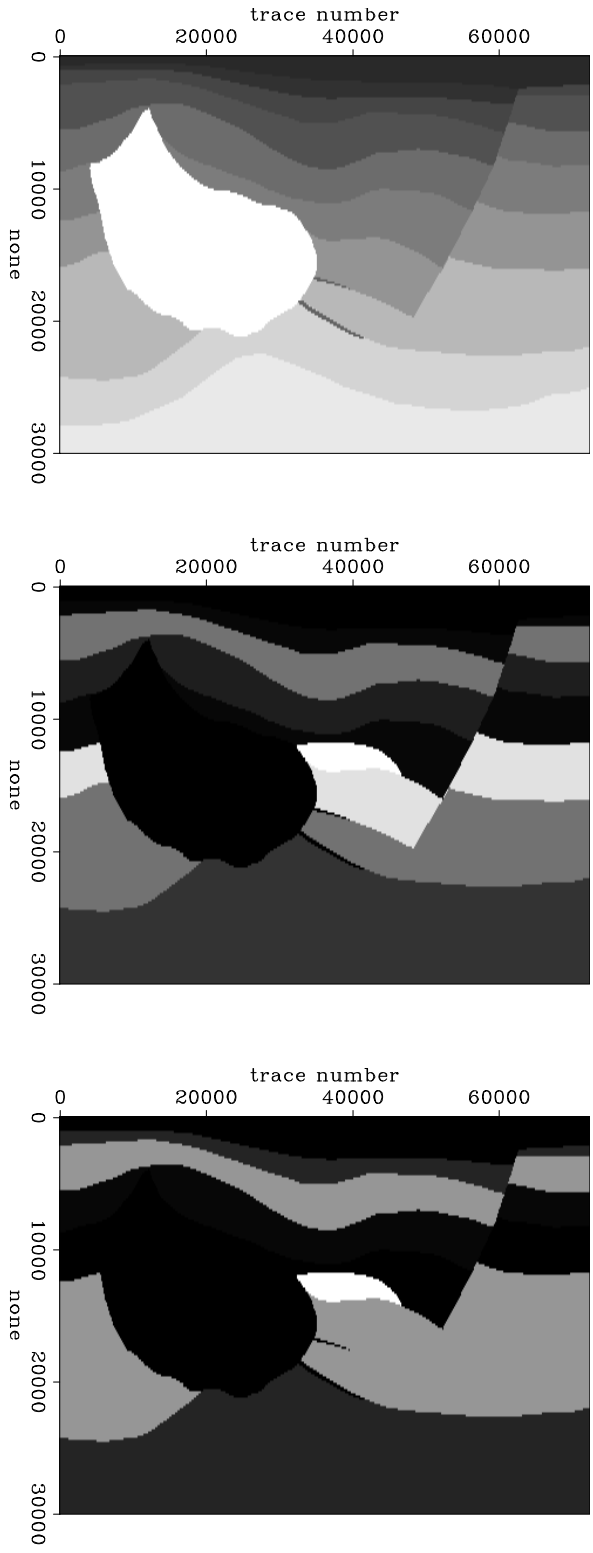


Figure 4: Hess VTI model. Top: Acoustic velocity field; center: ϵ field. Bottom: δ field. alejandro/. Hess-vels

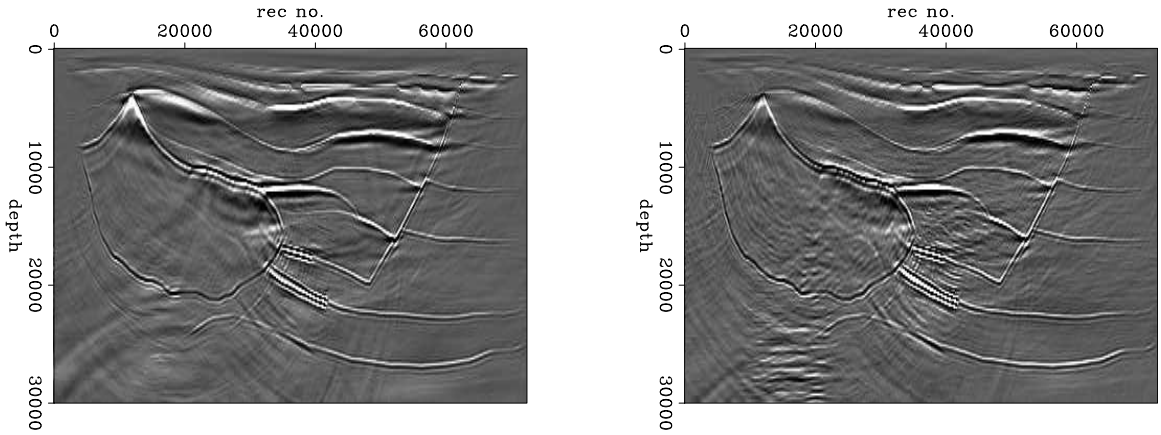


Figure 5: Reverse-time migration of the Hess VTI model. Left: Using tapering boundary conditions. Right: Using random boundary conditions. alejandro/. Hess-stacks

Finally, in the Hess VTI RTM sections we observe aliasing artifacts related to spatial sampling decimation. Additionally, using RBC introduces random noise, particularly severe below the salt body, and above and below the reservoir levels. Interestingly, low-wavenumber noise observed when using TBC is absent using RBC.

CONCLUSIONS AND FUTURE WORK

- RTM implemented via REM produces images of good quality without the necessity of refining the recording time sampling. The best illustration of this fact is the migration of the Amoco 1997 2.5D dataset, that is sampled every 9.9 ms.
- TTI anisotropy can be properly handled by REM using the pseudospectral approach, by means of the TTI decoupled equations.
- Using RBC highly reduces the memory usage, because only two spatial frames of every wavefield (corresponding to two time steps) are stored in memory.
- The comparisons between random boundaries reveals that in the current implementation there are still some artifacts related to the RBC, although some improvements in the sharpness of high-velocity contrasts (e.g., salt flanks) and better definition of reflections are observed in the Amoco 1997 2.5D RTM image, when using boundary conditions.
- Although RBC are very promising as an alternative to other optimization strategies to compute the imaging condition (e.g., checkpointing), in the current implementation care should be taken to minimize the artifacts.

Future actions include the following:

- Compare the REM performance to high-order finite differences
- Better suppression of the remaining low-frequency shadows in the final image

- Data-driven estimation of the source wavelet
- Computation of offset-domain and angle-domain common image gathers
- Exploration of more efficient anisotropy solvers
- Computation of the adjoint operator, for subsequent use in least-squares optimization

ACKNOWLEDGMENTS

I would like to thank Petroleos Mexicanos for the financial support, and Tariq Alkhalifah for the pertinent information about the usage of his Marmousi dataset. Very special thanks to Kurt Marfurt and the Attribute Assited Processing and Interpretation (AASPI) consortium, who allowed me to carry on the initial steps of this project using the University of Oklahoma servers. Last but not least, thanks to the SEP crew and sponsors.

REFERENCES

Alkhalifah, T., 1997, An acoustic wave equation for anisotropic media: SEP-Report, **95**, 283–307.

Baysal, E., D. Kosloff, and J. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524.

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.

Chang, W. and G. McMechan, 1990, 3D acoustic prestack reverse-time migration: Geophysical Prospecting, **38**, 737–755.

Chu, C. and P. Stoffa, 2012, Efficient 3D frequency response modeling with spectral accuracy by the rapid expansion method: Geophysics, **77**, no. 4, T117–T123.

Claerbout, J., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481.

Claerbout, J. F., 1985, Imaging the Earth's interior: Blackwell Scientific Publications, Inc.

Clapp, R. G., 2009, Reverse time migration with random boundaries: SEP-Report, **138**, 29–38.

Dellinger, J., S. Gray, G. Murphy, and J. Etgen, 2000, Efficient 2.5-D true-amplitude migration: Geophysics, **65**, 943–950.

Etgen, J., 1989, Accurate wave equation modeling: SEP-Report, **60**, 131–148.

Etgen, J. and C. Regone, 1998, Strike shooting, dip shooting, widepatch shooting - Does prestack depth migration care? A model study: SEG Technical Program Expanded Abstracts, 66–69.

Gazdag, J. and E. Carrizo, 1986, On reverse-time migration: Geophysical Prospecting, **34**, 822–832.

Jastram, C. and A. Behle, 1991, Elastic modeling by finite difference and the rapid expansion method (REM): SEG Technical Program Expanded Abstracts, 1573–1576.

Kosloff, D. and E. Baysal, 1983, Migration with the full acoustic wave equation: Geophysics, **48**, 677–687.

Kosloff, D., A. Filho, E. Tessmer, and A. Behle, 1989, Numerical solution of the acoustic and elastic wave equation by a new rapid expansion method: Geophysical Prospecting, **37**, 383–394.

Liu, F., Z. Guanquan, S. Morton, and J. Leveille, 2011, An effective imaging condition for reverse-time migration using wavefield decomposition: Geophysics, **76**, no. 1, S29–S39.

Loewenthal, D. and L. Hu, 1991, Two methods for computing the imaging condition for common-shot prestack migration: Geophysics, **56**, 378–381.

McMechan, G., 1983, Migration by extrapolation of time-dependent boundary values: Geophysical Prospecting, **31**, 413–420.

Pestana, R. and P. Stoffa, 2010, Time evolution of the wave equation using rapid expansion method: Geophysics, **75**, no. 4, T121–T131.

Stoffa, P. and R. Pestana, 2009, Numerical solution of the acoustic wave equation by the rapid expansion method (REM) - A one step time evolution algorithm: SEG Technical Program Expanded Abstracts, 2672–2676.

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

Tal-Ezer, H., D. Kosloff, and Z. Koren, 1987, An accurate scheme for seismic forward modelling: Geophysical Prospecting, **35**, 479–490.

Tessmer, E., 2011, Using the rapid expansion method for accurate time-stepping in modeling and reverse-time migration: Geophysics, **76**, no. 4, S177–S185.

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

Yoon, K. and K. Marfurt, 2006, Reverse-time migration using the Poynting vector: Exploration Geophysics, **37**, 102–107.

Zhan, G., R. Pestana, and P. Stoffa, 2012, Decoupled equations for reverse time migration in tilted transversely isotropic media: Geophysics, **77**, no. 2, T37–T45.

Image-based Q tomography using reverse time Q migration

Yi Shen and Tieyuan Zhu

ABSTRACT

We have developed a technique to tomographically estimate a Q model from migrated images using two-way wavefield continuation. When compared with our previously proposed technique that uses a one-way downward-continuation method, this technique better handles steep structures, e.g. salt flanks. Numerical results on a complex model with a salt body demonstrate the effectiveness of our two-way method for resolving attenuation in the presence of steep and overturned structures.

INTRODUCTION

Attenuation, parameterized by the seismic quality factor, Q, has considerable impact on surface seismic reflection data. It degrades the image quality by lowering the amplitude of events, distorting the phase of wavelet, removing the higher frequencies, and dispersing the velocity. These effects complicate reservoir interpretation. Therefore it is important to estimate the properties of this attenuation parameter and to use them for compensating the attenuated image.

Shen et al. (2013) and Shen et al. (2014) presented a new method, wave-equation migration Q analysis (WEMQA), to produce a reliable Q model. This method used one-way wavefield continuation to first migrate the viscoacoustic data with Q compensation for measuring the Q effects and then back-project these effects to update the current Q model. This method works well with gentle horizontal variations. However, the one-way method has limited abilities to represent the actual wave propagation around steep structures, because propagating the wavefield along only one direction of the depth axis cannot properly deal with overturned events. In some situations, overturned events provide extremely useful information, e.g. reservoir properties along a salt flank. In principle, two-way wavefield continuation is capable of modeling these overturned waves. Therefore, we apply two-way wavefield continuation migration with Q compensation (i.e. reverse time migration with Q compensation) to produce an image for Q-effect measurements (Zhu et al., 2014) and update the Q model using a two-way wave-equation tomographic operator to better handle the steep structures.

In this report, we first describe the theory of two-way migration with Q compensation and our two-way wave-equation tomographic operator. Then, we apply WEMQA with this method to a complex model with a salt body to demonstrate the effectiveness of the two-way method to resolve steep structures.

THEORY

We incorporate reverse time migration with Q compensation (Zhu et al., 2014) and two-way wave-equation tomographic operator into WEMQA (Shen et al., 2013, 2014) to estimate Q model and compensate Q effects.

Reverse time migration with Q compensation

Zhu and Harris (2014) first introduced the time-domain viscoacoustic wave equation based on the constant Q model (Kjartansson, 1979), which takes the form

$$\left(\eta\mathbf{L}+\tau\mathbf{H}\frac{d}{dt}-v^{-2}\frac{\partial^2}{\partial t^2}\right)P(t)=f(t), \tag{1}$$

where t is time, P is the propagated wavefield, and f is the source wavefield, v is the acoustic velocity at the reference frequency ω_0. Operators $\mathbf{L}=\left(-\nabla^2\right)^{\gamma+1}$ and $\mathbf{H}=\left(-\nabla^2\right)^{\gamma+1/2}$ are fractional Laplacians, where the variable γ is defined as $\gamma=1/\pi\tan^{-1}(1/Q)$. The first term on the left side of Equation 1 is related to dispersion effects, and the middle term on the left side of Equation 1 is related to the absorption effects (Zhu et al., 2014). The absorption and dispersion coefficients are given by $\eta=-v^{2\gamma}\omega_0^{-2\gamma}\cos\pi\gamma$ and $\tau=-v^{2\gamma-1}\omega_0^{-2\gamma}\sin\pi\gamma$. We use a pseudo-spectral method in our numerical implementation.

Attenuation damps the higher frequencies more than the lower frequencies of the propagating wave. Compensation through Q migration (Zhu et al., 2014), i.e. the inverse of forward wave propagation, therefore preferentially boosts higher frequencies. Such amplification may gain high frequency noise and make these frequencies dominate the image. Therefore, we add a low-pass filter to help mitigate high-frequency noise artifacts in migration:

$$\left(\nabla^2+\mathbf{N}(\eta\mathbf{L}-\nabla^2)+\mathbf{N}\tau\mathbf{H}\frac{d}{dt}-v^{-2}\frac{\partial^2}{\partial t^2}\right)P(t)=f(t), \tag{2}$$

where $\mathbf{N}$ is the low-pass filter in the spatial frequency domain.

Two-way wave-equation tomographic operator

Image-based wave-equation Q tomography is a nonlinear inversion process that aims to find the Q model that minimizes the residual field in the image space, $\Delta I(\mathbf{x},\mathbf{h})$. This residual image can be approximated by a linearized operator applied to the model perturbation ΔQ:

$$\Delta I(\mathbf{x},\mathbf{h})=\sum_{\mathbf{y}}\frac{\partial I(\mathbf{x},\mathbf{h})}{\partial Q(\mathbf{y})}\,|_{Q_0}\,\Delta Q(\mathbf{y}), \tag{3}$$

where $\mathbf{x}$ and $\mathbf{y}$ are the coordinates of the image and model, respectively. $\mathbf{h}$ is the subsurface offset and Q_0 is the background quality factor. The adjoint of this tomographic operator projects the image perturbation back into the Q model space. The back-projected changes in the model space are used as gradient directions to conduct a line search in an optimization scheme. This back-projection can be expressed as follows:

$$\Delta Q\left(\mathbf{y}\right)=\sum_{\mathbf{y}}\left(\frac{\partial I\left(\mathbf{x},\mathbf{h}\right)}{\partial Q\left(\mathbf{y}\right)}\mid_{Q_0}\right)^{*}\Delta I\left(\mathbf{x},\mathbf{h}\right) \tag{4}$$

The derivative of $I\left(\mathbf{x},\mathbf{h}\right)$ with respect to Q is the sum of the attenuation-induced perturbed source wavefield multiplied by the background receiver wavefield and the attenuation-induced perturbed receiver wavefield multiplied by the background source wavefield:

$$\begin{aligned}&\frac{\partial I\left(\mathbf{x},\mathbf{h}\right)}{\partial Q\left(\mathbf{y}\right)}\mid_{Q_0}\\&=\sum_{\mathbf{x_s},\mathbf{x_r}}\left(\frac{\partial G\left(\mathbf{x}-\mathbf{h},\mathbf{x_s};Q_0\right)}{\Delta Q\left(\mathbf{y}\right)}\right)^{*}G^{*}\left(\mathbf{x}+\mathbf{h},\mathbf{x_r};Q_0\right)d\left(\mathbf{x_r},\mathbf{x_s}\right)\\&+\sum_{\mathbf{x_s},\mathbf{x_r}}G^{*}\left(\mathbf{x}-\mathbf{h},\mathbf{x_s};Q_0\right)\left(\frac{\partial G\left(\mathbf{x}+\mathbf{h},\mathbf{x_r};Q_0\right)}{\Delta Q\left(\mathbf{y}\right)}\right)^{*}d\left(\mathbf{x_r},\mathbf{x_s}\right),\end{aligned} \tag{5}$$

where $\mathbf{x_s}$ and $\mathbf{x_r}$ are the source and receiver coordinates respectively, G is the Green's function, and d are the surface data. The attenuation-induced perturbed wavefield can be linearized using a Taylor expansion as follows:

$$\begin{aligned}&\frac{\partial G\left(\mathbf{x},\mathbf{x_s};Q_0\right)}{\partial Q\left(\mathbf{y}\right)}\\&=-\frac{\partial\eta\left(\mathbf{y};Q\right)\mathbf{L}\left(\mathbf{y};Q\right)}{\partial Q\left(\mathbf{y}\right)}\mid_{Q_0}G\left(\mathbf{y},\mathbf{x_s};Q_0\right)G\left(\mathbf{x},\mathbf{y};Q_0\right)\\&-\frac{\partial\tau\left(\mathbf{y};Q\right)\mathbf{H}\left(\mathbf{y};Q\right)}{\partial\gamma}\mid_{Q_0}\frac{d}{dt}G\left(\mathbf{y},\mathbf{x_s};Q_0\right)G\left(\mathbf{x},\mathbf{y};Q_0\right)\\&\frac{\partial G\left(\mathbf{x},\mathbf{x_r};Q_0\right)}{\partial Q\left(\mathbf{y}\right)}\\&=-\frac{\partial\eta\left(\mathbf{y};Q\right)\mathbf{L}\left(\mathbf{y};Q\right)}{\partial Q\left(\mathbf{y}\right)}\mid_{Q_0}G\left(\mathbf{y},\mathbf{x_r};Q_0\right)G\left(\mathbf{x},\mathbf{y};Q_0\right)\\&-\frac{\partial\tau\left(\mathbf{y};Q\right)\mathbf{H}\left(\mathbf{y};Q\right)}{\partial\gamma}\mid_{Q_0}\frac{d}{dt}G\left(\mathbf{y},\mathbf{x_r};Q_0\right)G\left(\mathbf{x},\mathbf{y};Q_0\right)\end{aligned} \tag{6}$$

where:

$$\begin{aligned}&\frac{\partial\eta\left(\mathbf{y};Q\right)\mathbf{L}\left(\mathbf{y};Q\right)}{\partial Q\left(\mathbf{y}\right)}\mid_{Q_0}\\&=-\frac{\left(2\eta\ln v-2\eta\ln\omega_0-\pi v\tau\right)\mathbf{L}+\eta\mathbf{L}\ln\left(-\nabla^2\right)}{\pi\left(Q_0^2+1\right)}\\&\frac{\partial\tau\left(\mathbf{y};Q\right)\mathbf{H}\left(\mathbf{y};Q\right)}{\partial\gamma}\mid_{Q_0}\\&=-\frac{\left(2\tau\ln v-2\tau\ln\omega_0+\pi v^{-1}\eta\right)\mathbf{H}+\tau\mathbf{H}\ln\left(-\nabla^2\right)}{\pi\left(Q_0^2+1\right)}.\end{aligned} \tag{7}$$

NUMERICAL RESULTS

Our numerical example employs a dataset generated by Schlumberger (Cavalca et al., 2013) using a 2D viscoacoustic version of the 2004 BP benchmark model (Billette and Brandsberg-Dahl, 2005). An attenuation model was added by Schlumberger to the original 2004 BP

models. This Q model is not released, but its location and value has been shown by Schlumberger (Cavalca et al., 2013). The attenuation model is a space- and depth-variant absorption model made of several Q heterogeneities and a nonattenuative background (1/Q = 0.0002). To test the effectiveness of our method on the steep structure, we only focus our work on the salt region, which has a large attenuative zone (1/Q = 0.02) near the left of the salt flank. A velocity model for this region is shown in Figure 1. The viscoacoustic surface seismic data generated by Schlumberger used a finite-difference modeling code based on standard linear solid theory. In my example, 248 shots are used with 100 m spacing, and the offsets range from -15,000 m to 15,000 m. Receivers are distributed on both sides of each shot at an increment of 25 m. The cut-off frequency that we used to filter out the noise in the reverse-time migration was 30 Hz, which is 60% of the maximum useful frequency. The workflow used in this numerical example has been presented by Shen et al. (2013, 2014).

Figures 2 and 3 are the attenuated images from the viscoacoustic data using one-way and two-way Q migration with a nonattenuating model, respectively. The results show that the events beside the salt flank are attenuated in terms of the amplitude dimming, incoherency of the events, and stretching of the wavelets. Both migrations image the structure with gentle horizontal variation well. However, Figure 2 shows a poorer image around the salt flank when compared with Figure 3. The salt flank in Figure 2 is not focused. The regions beside the salt have discontinuous events and are contaminated with high-frequency noises. Figure 3 shows a sharper and clearer salt flank. The image around the salt is cleaner with less high-frequency noise and the events are more coherent.

The next step is to invert for the Q model besides the salt flank from these attenuated images using the wave-equation migration Q analysis developed by Shen et al. (2013, 2014). The initial model for the inversion has no attenuation. We analyze the attenuation effects by calculating the slope of the logarithm of the spectral ratio between the windowed events of each trace and the events in the reference window at the same depth. The window size is 1500 m, and 100 sliding windows are used for each trace. The reference trace is the one at 24,000 m. Figure 5 and Figure 6 are the inverted Q model using one-way and two-way WEMQA, respectively. Figure 5 fails to resolve the Q model in area besides the salt flank, because one-way propagation is not able to accurately reflect the steep structure. Figure 6 shows that a two-way method retrieves the Q model in the reservoir region beside the salt well; especially at the upper part of the salt flank, which matches well with the true Q model. However, it still fails to update the Q model at the lower part of the salt flank because of the high-frequency image artifacts at that area. Figure 4 is the image using two-way migration with the inverted Q model shown in Figure 6 for compensation. The results show that the image at the lower region is not well compensated because of the weak updating of the Q model there. Therefore, a filter, e.g. a curvelet-transform-based filter, is needed to mitigate the image noise. But, the events in the rest of the part beside the salt flank are compensated. Their amplitudes are partially recovered, events are more coherent, and phases are corrected.

CONCLUSION

This report has presented a new wave-equation migration Q analysis based on a two-way wavefield continuation method. When compared with previous work that uses a one-way downward-continuation method, the two-way based method is better in handling the steep

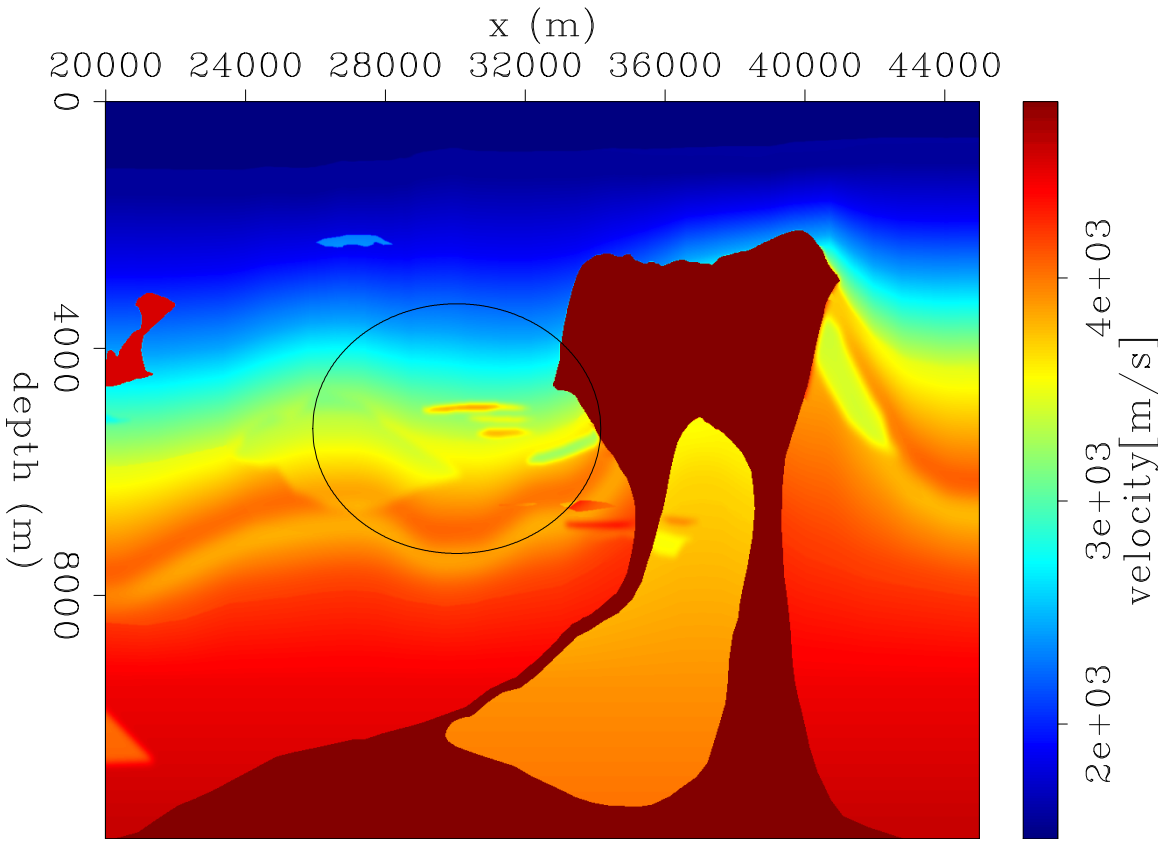


Figure 1: BP velocity model. The circles approximately show the location of the attenuation zone besides the salt flank, with the lowest Q value of $Q = 50$. [**ER**] yishen2/. bpVelAnn

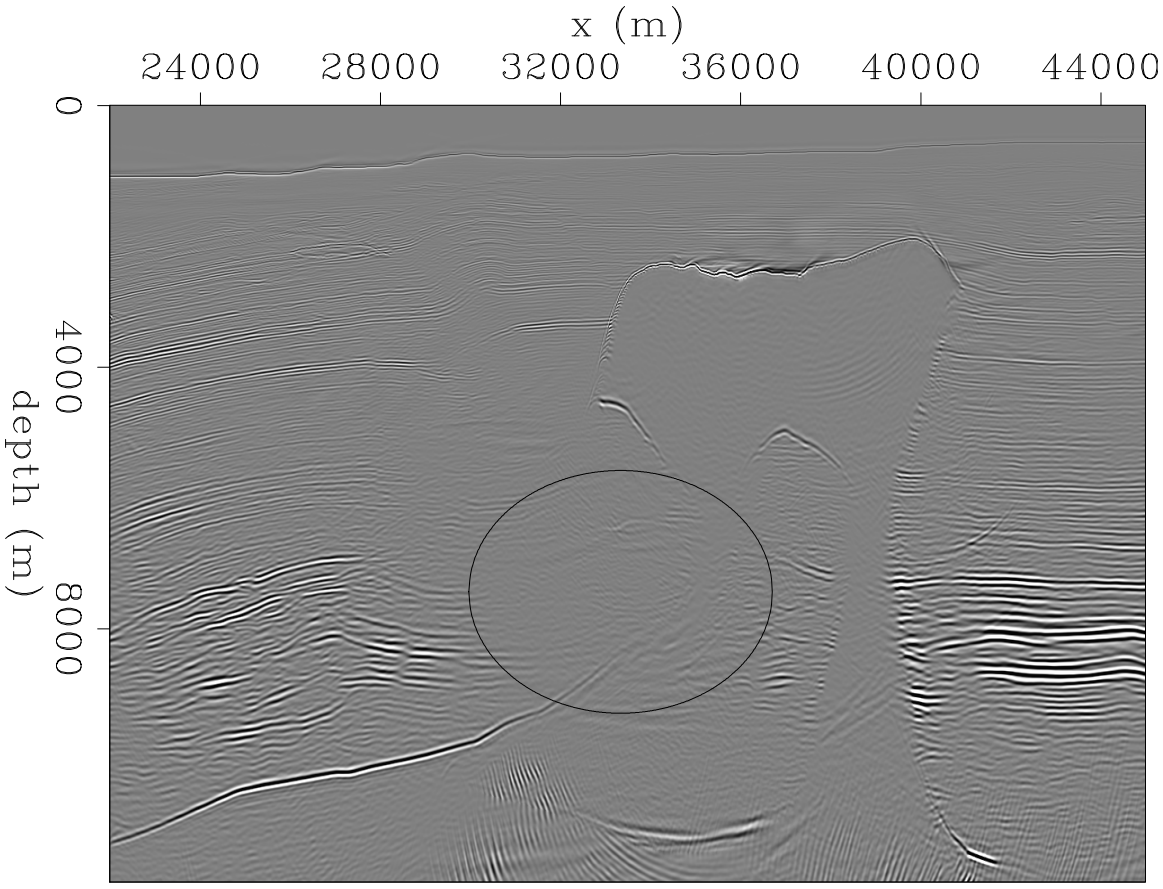


Figure 2: Attenuated images from the viscoacoustic data using one-way Q migration with a nonattenuating model. As shown in the circles, the salt flank is not focused, and the regions beside the salt have discontinued events and are contaminated with high-frequency noises. [**CR**] yishen2/. bpImgOneAnn

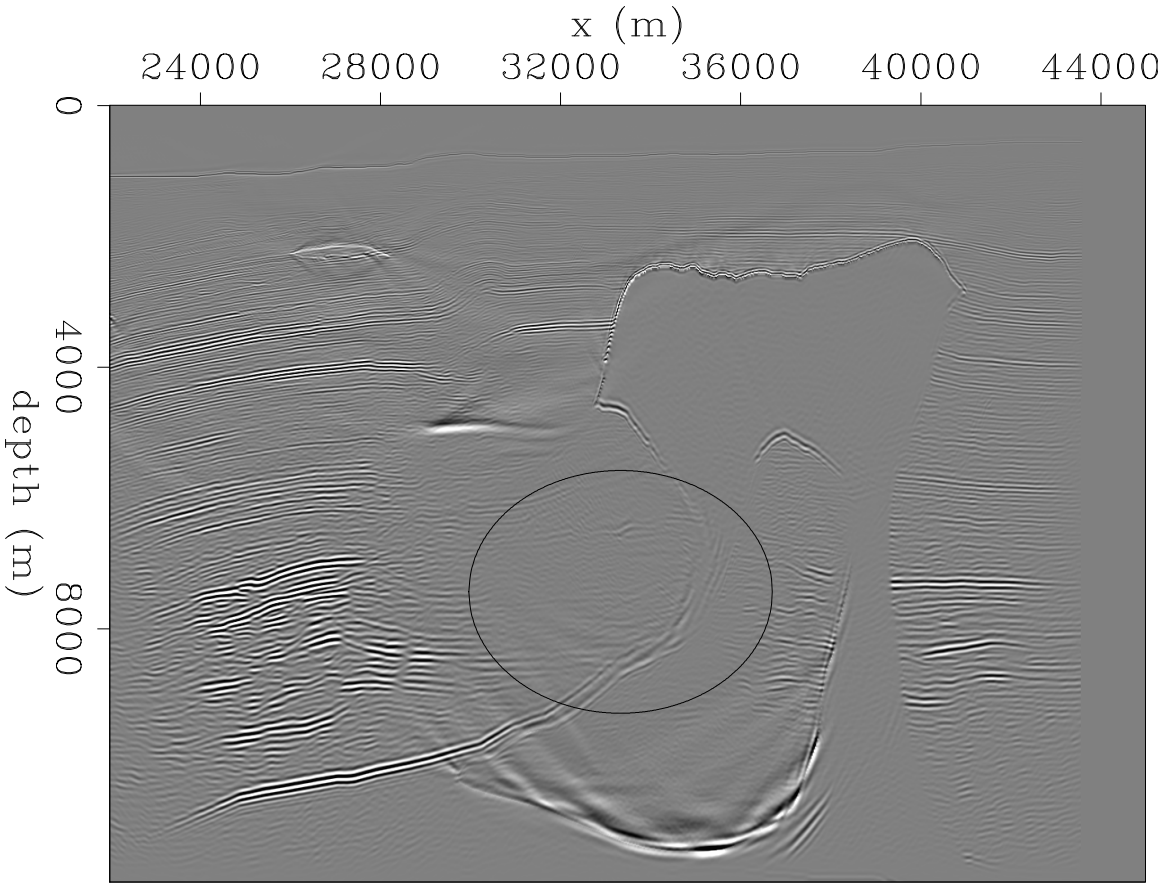


Figure 3: Attenuated images from the viscoacoustic data using two-way reverse-time Q migration with nonattenuating model. As shown in the circles, the salt flank is sharp and well-focused. The image around the salt is cleaner with less high-frequency noises, and the events becomes more coherent, when compared with Figure 2. [**CR**] yishen2/. bpImgRtmAnn

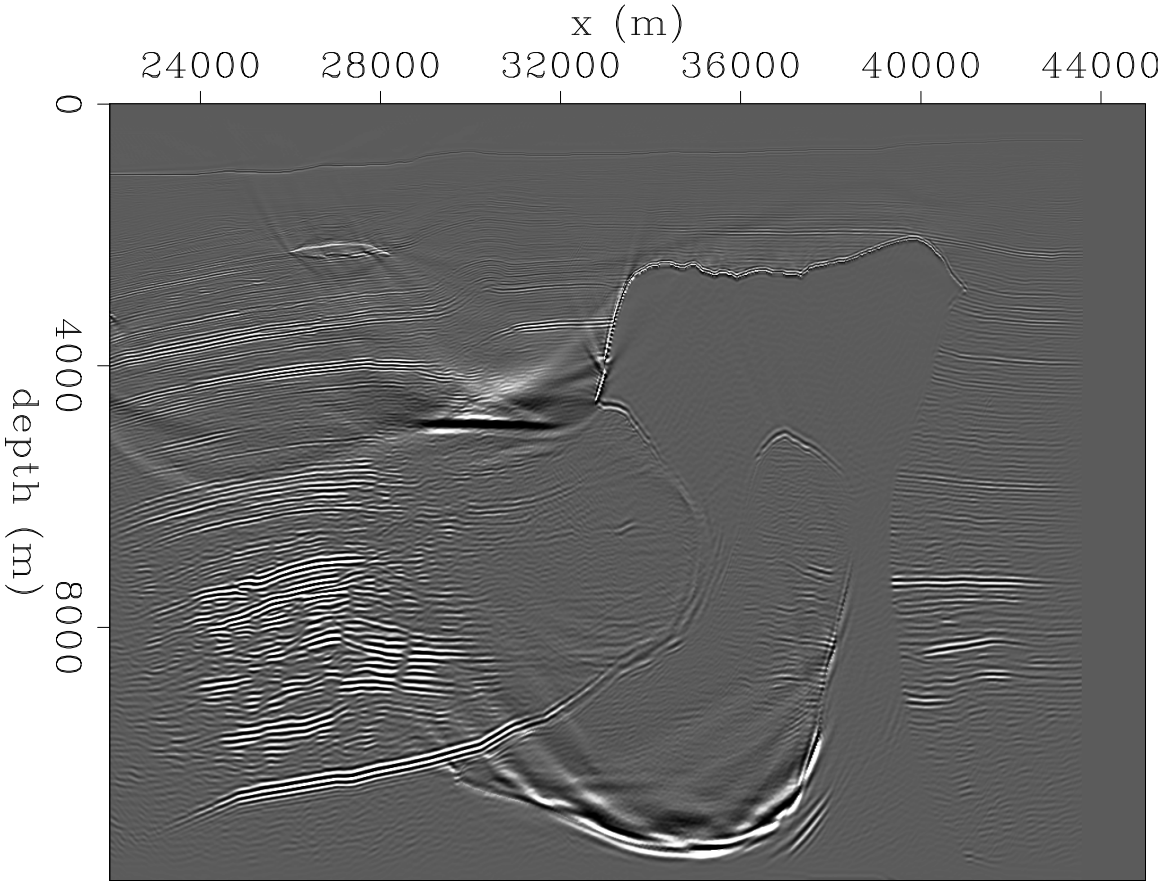


Figure 4: Compensated image using two-way reverse-time Q migration. The result shows the image at the lower region is not well compensated because of the weak updating of the Q model there. But the other events alongside the salt flank are compensated. [**CR**] yishen2/. bpImgQrtm

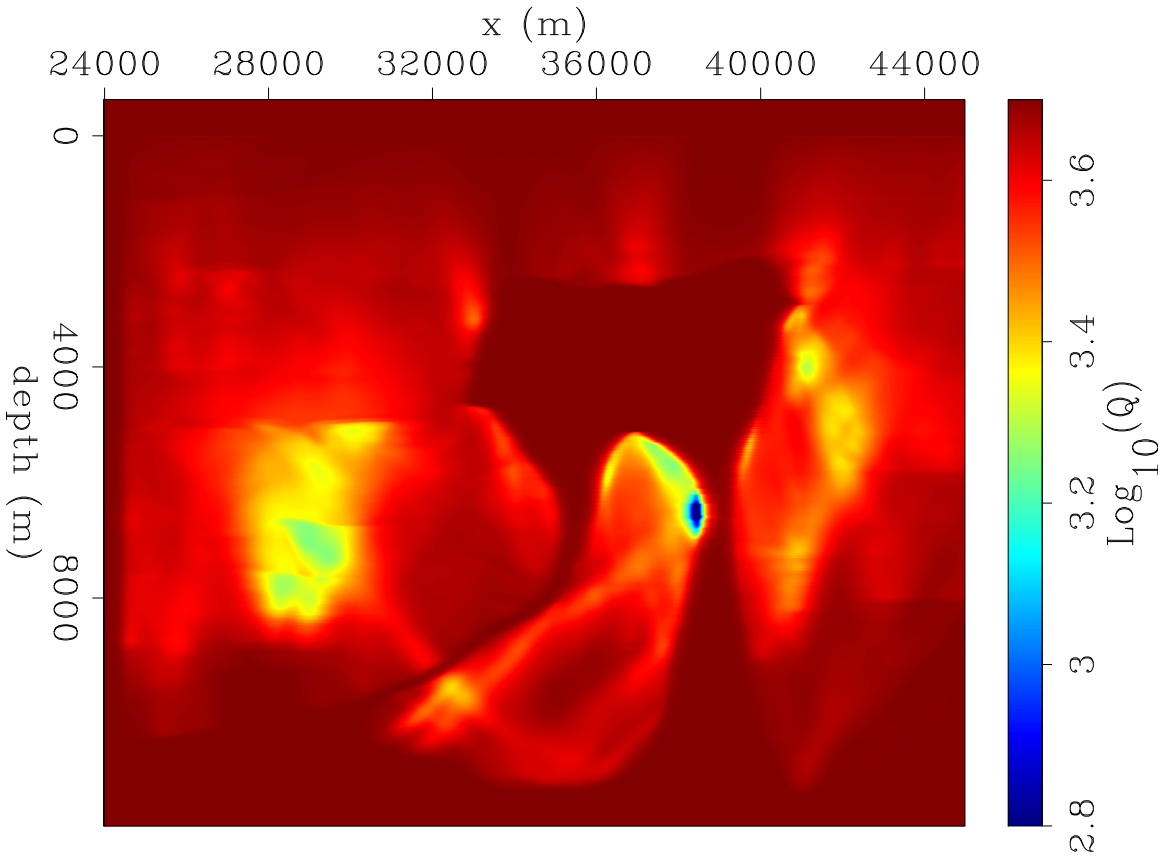


Figure 5: The inverted Q model using one-way wave-equation migration Q analysis. The result fails to resolve the Q model in the area besides the salt flank, because one-way propagation is not able to accurately reflect the steep structure. [**CR**] yishen2/. bpQOne

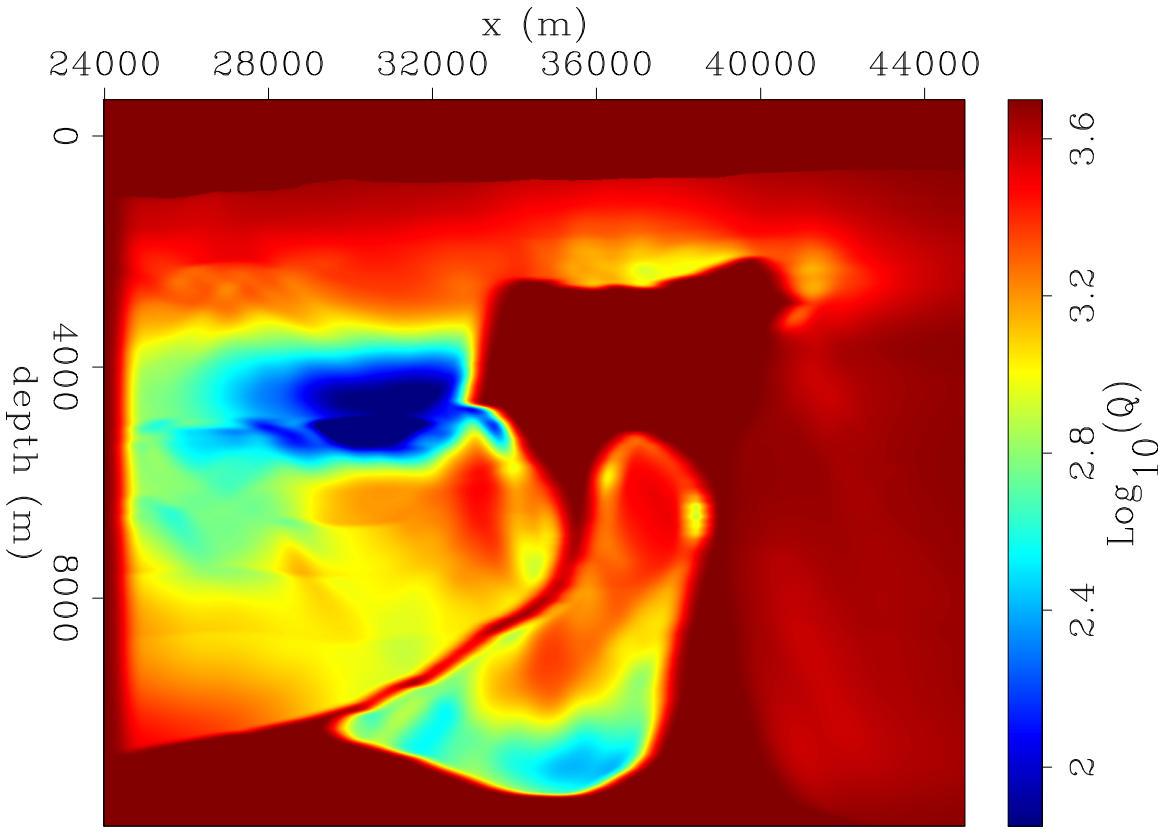


Figure 6: The inverted Q model using two-way wave-equation migration Q analysis. The result retrieves the Q model in the reservoir region beside the salt well; especially near the upper part of the salt flank. However, it still fails to update the Q model near the lower part of the salt flank because of the high-frequency image artifacts at that area. [**CR**] yishen2/. bpQRtm

structures, e.g. salt flanks. Numerical tests on a complex model with a salt body demonstrate the effectiveness of this two-way method on handling the overturned wave propagation caused by steep structures.

ACKNOWLEDGMENTS

The authors thank Biondo Biondi and Robert Clapp for their advice and suggestions, Schlumberger and BP for providing the 2D synthetic dataset, Yaxun Tang and Ali Almonin for help with programming, and Stanford Exploration Project sponsors for their financial support. The second author wish to thank the Jackson Distinguished Postdoctoral Fellowship at the University of Texas at Austin.

REFERENCES

Billette, F. and S. Brandsberg-Dahl, 2005, The 2004 BP velocity benchmark: 67th Conference and Exhibition, EAGE, Extended Abstracts, B035.

Cavalca, M., R. P. Fletcher, and M. Riedel, 2013, Q-compensation in complex media – ray-based and wavefield extrapolation approaches: 83rd Annual International Meeting, SEG, Expanded Abstracts, 3831–3835.

Kjartansson, E., 1979, Constant Q wave propagation and attenuation: Journal of Geophysical Research, **84**, 4737–4748.

Shen, Y., B. Biondi, R. Clapp, and D. Nichols, 2013, Wave-equation migration Q analysis (WEMQA): EAGE Workshop on Seismic Attenuation Extended Abstract.

———, 2014, Wave-equation migration Q analysis (WEMQA): SEG Technical Program Expanded Abstracts.

Zhu, T. and J. Harris, 2014, Modeling acoustic wave propagation in heterogeneous attenuating media using decoupled fractional Laplacians: Geophysics, **79**, no. 3, T105–T116.

Zhu, T., J. Harris, and B. Biondi, 2014, Q-compensated reverse-time migration: Geophysics, **79**, no. 3, S77–S87.

Approximating Q propagation to speed up finite differences

Robert G. Clapp

ABSTRACT

Propagating wave-fields using explicit finite difference is the kernel for Reverse Time Migration (RTM) and high end velocity analysis algorithms. To avoid grid dispersion artifacts, the cost of propagation is proportional to the frequency of the energy being propagated raised to the fourth power. Attenuation in the earth tends to decrease usable frequencies as a function of time. By using an approximation to the wave equation for attenuated media, to save computation, we can approximate the earth's behavior. As a result we can use coarser sampling at large time. Combined with limiting grid propagation to around the source at early times we can achieve large speedups in modeling, migration, and potentially velocity analysis.

INTRODUCTION

Imaging and velocity analysis are the most computationally intensive parts of seismic processing. As a results researchers are always trying to find ways to speedup these processes (Bednar and Neale, 2002; Stork, 2013). One approach used to speed up downward continuation based algorithms is to recognize that the earth attenuates seismic signals. As a result, as we push the wave-field down in depth, we can ignore higher and higher frequencies and still obtain an accurate image(Clapp, 2002). This approach lowers the cost as you increase in depth. This technique is well suited for downward continuation based approaches which are done frequency by frequency. Reducing the frequencies downward continued as a function of depth is particularly effective in combination with recognizing that there was no reason to propagate waves a large distance from the source at early times. While following the wavefield is used routinely in RTM, taking advantage of attenuation is not commonly used. Reasons include: the cost of propagation with an attenuated wave equation, attenuation is a function of medium parameter, and propagation is generally done in the the time, rather frequency domain.

In this paper I use a constant-Q approximation based on the work of Zhu and Harris (2014). As I propagate my source I resample my medium based on the maximum frequency that has not been significantly attenuated. Combining this approach with following the wave-field, I show that I can achieve significant computational speedups.

MODELING

Explicit finite difference modeling is constrained by figuring out a sampling in time and space that results in stable propagation and does not create dispersive events. For stability the Courant-Friedrichs-Lewy condition (Courant et al., 1967) must be met. Stability is a function of limiting what percentage of a grid cell energy can move in one time step. Stability

is therefore a function of the maximum velocity v_{max}, the minimum spatial sampling d_{min}, and the time step dt. For stability,

$$v_{max}\frac{dt}{d_{min}} < .5. \tag{1}$$

The stability condition pushes one to use larger spatial sampling (faster, but less resolution) and/or finer time sampling (more expensive). Dispersion, on the other hand, is a function of the minimum velocity v_{min}, the maximum frequency f_{max}, and the maximum spaital sampling d_{max}. To avoid grid dispersion we need to sample a given frequency with a minimum number of points. There isn't a consensus on the minimum number of points. For the purpose of this paper I will require 3.2 points therefore,

$$\frac{v_{min}}{f_{max}d_{max}} > 3.2. \tag{2}$$

The dispersion constraint pushes us towards smaller (more expensive) spatial sampling, because of the stability constraint, and results in smaller the steps. Minimizing dispersion is the real reason for the expense of finite differences. To avoid grid dispersion and achieve the same level of stability the number of operations increase by the fourth power (three due to space sampling and one for time).

From observation we know that the earth attenuates acoustic signals. Attenuation varies as a function of frequency and earth materials. The first approximation I am going to use is the concept of the constant Q model introduced by Kjartansson (1979). Q is defined as

$$Q = 2\pi\left(\frac{E}{\partial E}\right), \tag{3}$$

where $\frac{E}{\partial E}$ is the fraction of energy lost per cycle. The larger the Q value, the less energy loss per cycles. The constant Q assumption assumes that energy dies out is a function of the number of wavelengths traveled through a medium. The higher the frequency, the faster the energy is attenuated.

Constant-Q Formulation

Building on this foundation Zhu and Harris (2014) used a fractional Laplacian approach to perform attenuated propagation in the time domain. Given a forcing function $f(t)$ and the wave-field $P(t)$ they proposed the equation,

$$\left[\eta\mathbf{L} + \tau\mathbf{H}\frac{\mathbf{d}}{\mathbf{dt}} - \mathbf{v^{-2}}\frac{\partial^{\mathbf{2}}}{\partial\mathbf{t^2}}\right]P(t) = f(t), \tag{4}$$

where

$$\mathbf{L} = \left(-\nabla^{\mathbf{2}}\right)^{\gamma+\mathbf{1}} \tag{5}$$

$$\mathbf{H} = \left(-\nabla^{\mathbf{2}}\right)^{\gamma+\frac{\mathbf{1}}{\mathbf{2}}}. \tag{6}$$

The constants in equation 4 are defined as

$$\eta = -v^{2\gamma}w_0^{-2\gamma}\cos\pi\gamma, \tag{7}$$

$$\tau = -v^{2\gamma-1} w_0^{-2\gamma} \sin \pi\gamma, \tag{8}$$

and

$$\gamma = \frac{1}{\tan^{-1} \frac{1}{Q}}. \tag{9}$$

The first term in equation 4 deals the dispersive effect of attenuation. The middle term deals damping. For a constant Q approximation I can further simplify equation 4, approximating ∇^2 for $\mathbf{H}$. My resulting equation is then

$$\left[\nabla^2 - \tau\nabla^2 \frac{d}{dt} - v^{-2} \frac{\partial^2}{\partial t^2}\right] P(t) = f(t). \tag{10}$$

Figure 1 shows the wave-field using the standard acoustic wave equation (left) and an attenuation wave-field (right). Note the difference in the frequency content. This is can be more clearly seen in Figure 2 which shows the spectrum of wave-field at $.3, 1.3, 2.3$, and 3.3 seconds with $Q = 200$. Notice the energy decay of frequency over time. The key observation is that there is no reason to worry about dispersion at frequencies that have attenuated. Specifically, I redo my dispersion calculation (equation 1) several times while propagating a wave-field. At each time block I use a f_{max} based on frequencies whose energy has not been reduced more than some percentage (in this paper 96%). As a result, as I forward propagate in time my grid cells, and if I desire, time sampling, get larger. Figure 3 shows the speedup factor (defined as the number of grid cells times time steps) as a function of propagation time for different Q values. The longer the time record, the more using Q pays off in terms speed up. Using this approach the early time steps dominate the computation. For example assuming an initial maximum frequency of $90Hz$ and a constant velocity of $2000m/s$, a Q value of 200, the total speedup is 4 even though most of the propagation time shows significantly more speedup. Figure 4 shows the speedups for Q values ranging from 150 to 350.

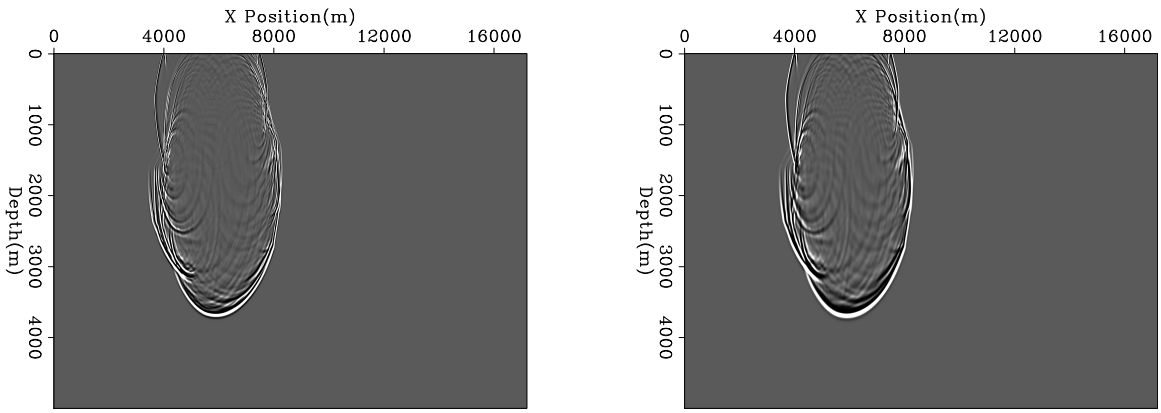


Figure 1: The wave-field with the standard acoustic wave equation (left) and the attenuated wave-field (right) at the same time. [**ER**] bob1/. qnoq

Another speedup trick used by many when doing modeling/migration is to recognize that there is no need to propagate the wave-fields significantly away from the source at early times. The number of cells we need to propagate increases as a power of three (expanding

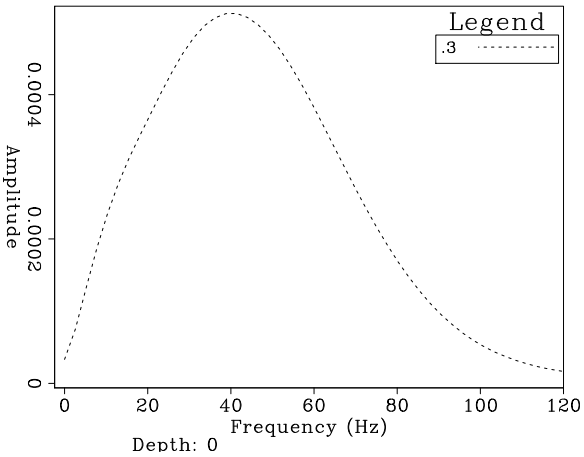


Figure 2: The frequency of the wavefield at various times. Note the reduction of the high frequencies over time. Note at .3 seconds, total amplitude is less because the source has not been fully injected into the medium. [**ER**] bob1/. spectra

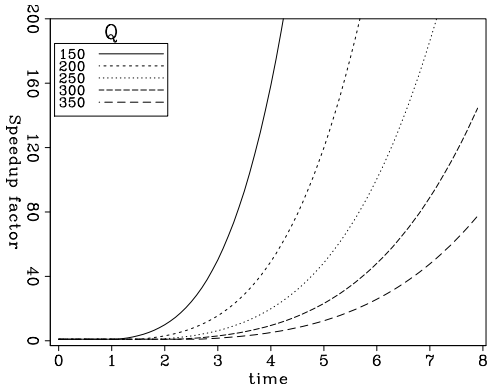


Figure 3: Speedup values as a function of propagation time for Q values ranging from 150 to 350. Note how the longer the time record the more the computation is sped up. [**ER**] bob1/. timeQ

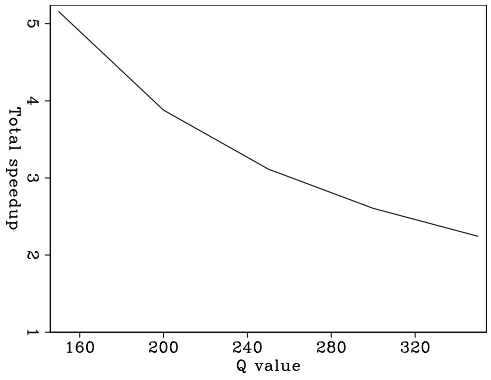


Figure 4: Total speedup for different values of Q. Note the relatively small speedups compared to what one might expect looking at Figure 3. Total time is the inverse of the sum of the inverses of speedups. [**ER**] bob1/. qtot

wave-field) in modeling, and somewhat less in migration due to the spatial extent of our receivers. This speedup trick is most effective at early times and useless at late times when energy has propagated throughout the model, the opposite behavior as the constant-Q trick. Figure 5 shows a typical wide azimuth geometry and the speedup as a function of time again assuming $2km/s$ medium. Total speedup for following the source is only 2-2.5. Combining the two tricks is where we begin to see big payoffs. The left plot of Figure 6 shows the result of combining the two approaches for different values of Q. The right plot shows the total speedup as a function of Q. Note how the maximum cost is in the 1 to 3 second range depending on the value of Q. Without following the wave-field the total speedup ranged from 1.8 to 4, we now get speedups between 15 and 250. Calculating speedup

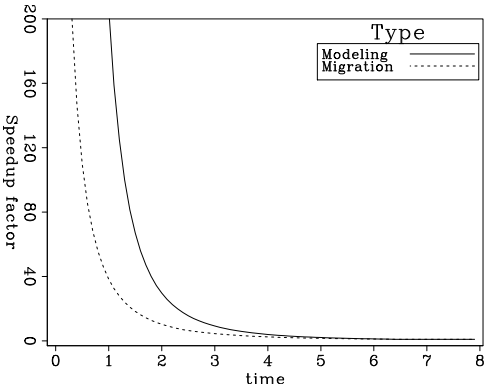


Figure 5: Speedup by following the wave-field in the case of modeling a single shot and in the case of an array of receivers. Note how these curves move in the opposite direction to those in Figure 3. [**ER**] bob1/. follow

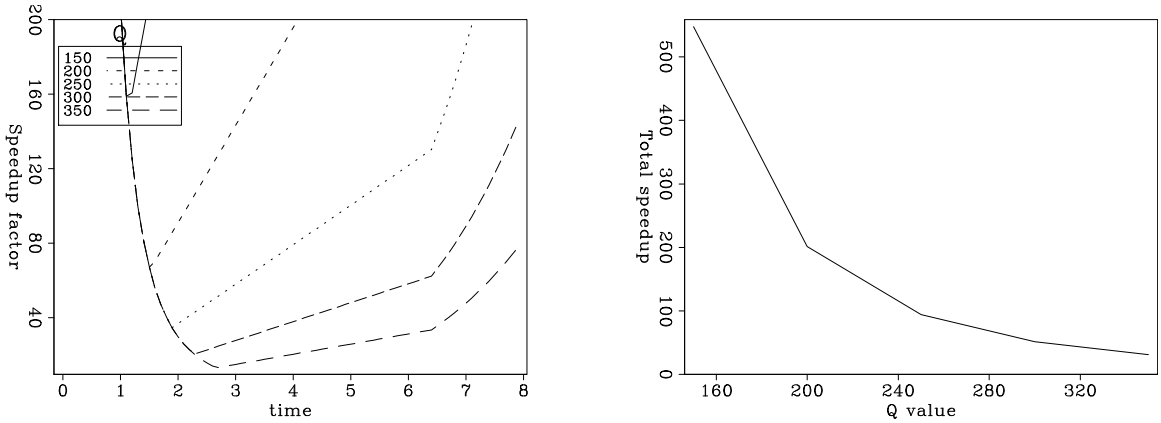


Figure 6: The left plot shows speed up as a function of propagation time for a series of Q values. The right plot shows the total speedup as a function of Q. Note on the left how total speedup is a combination of the trends shown in Figures 3 and 5. Note on the right the significant speedup compared to those shown in Figure 4. [**ER**] bob1/. timeTot

numbers is far from an exact science. My choice of a low constant velocity helps me by improving the speedups due to following the wave-field and hurts me by allowing larger time steps. Computational time and the number of grid cells do not completely linear relation. Smaller grids, means better cache behavior, and can lead to significant improvements in performance. In addition my propagator is more expensive than the standard acoustic propagator which isn't taken into account.

The basic algorithm for modeling is shown in Algorithm 1. Figure 7 shows the wave-field

Algorithm 1 Forward propagation

```
for timeblock 0...n do
  Calculate max frequency of interest relevant at the beginning of time window
  Calculate max extent of wave-field at end of time window
  Calculate sampling of medium and time based on stability/dispersion constraints
  Resample wave-field and velocity
  for t=0...n in timeblock do
    Propagate wave-field with Q approximation
    Inject source
    if Imaging step then
      Store wave-field
    end if
  end for
end for
```

propagating through a relatively complex synthetic after 2 seconds using resampling tricks (left) and a static grid (right).

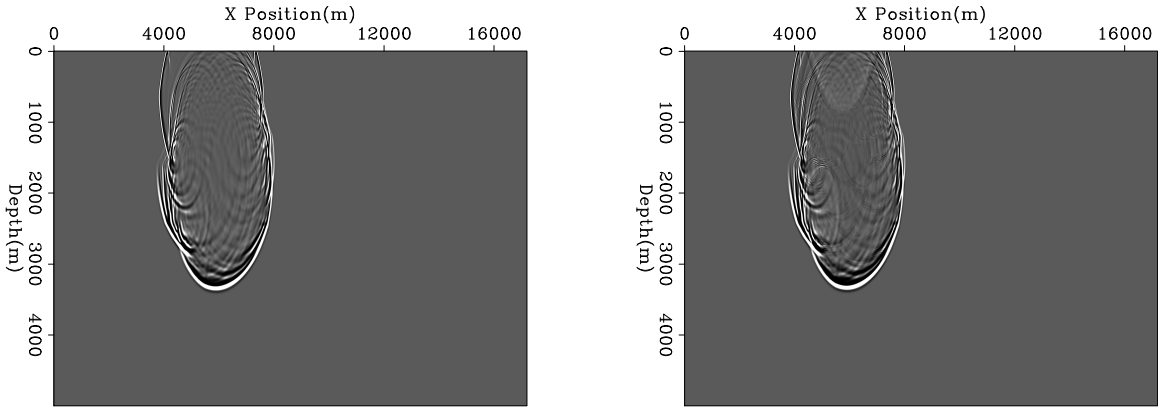


Figure 7: The left panel shows the wave-field using a static grid, the right panel using the approach outline in algorithm 1 both using equation 10. Note the low frequency artifact in the right panel, but otherwise the plots are nearly identical. [**ER**] bob1/. sampleCompare

MIGRATION

Applying the method described above to RTM is relatively straightforward. Modeling is done using the approach outlined in algorithm 1. Back propagation starts with low frequencies and then increases in frequency at smaller times. For real data there is no need to use the approximate Q propagator because the whole assumption of this approach is that there is no useful information at high frequencies at large times. The back projection step is described in algorithm 2.

Figure 8 shows the result of migrating a single shot using a standard static grid, left, and a changing grid, right. There is significantly more noise, particularly away from the

Algorithm 2 Backward propagation for real data

```
for timeblock n...0 do
    Calculate max frequency of interest relevant at the beginning of time window
    Calculate max extent of wave-field at end of time window
    Calculate sampling of medium and time based on stability/dispersion constraints
    Resample wave-field, velocity
    Resample temp image to full image and sum
    Create temp image
    for t=n..0 in timeblock do
        Propagate wave-field
        Inject source
        if Imaging step then
            Apply imaging condition store in temp image
        end if
    end for
end for
```

main energy train, using the variable grid, but the main energy train is nearly identical. Figure 9 shows the result of migrating an entire dataset using the changing grid method. After summing all the shots the artifacts seen in Figure 8 have disappeared.

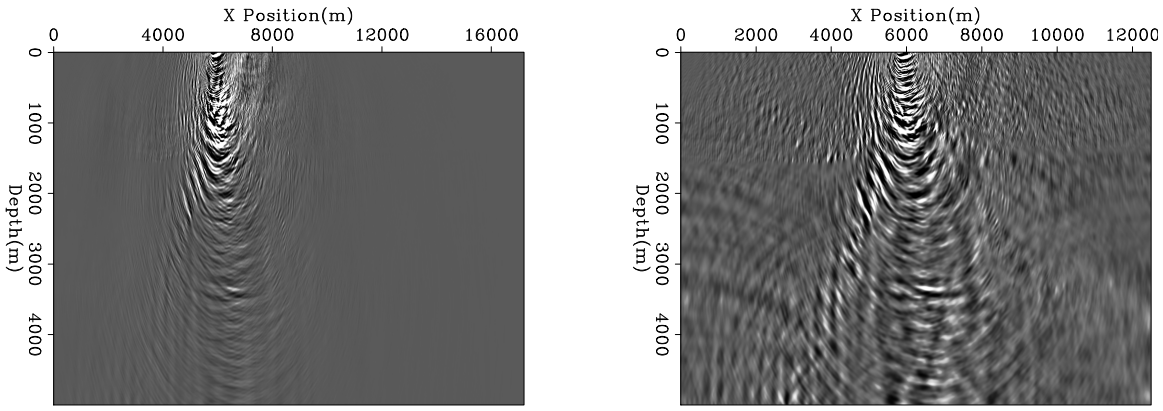


Figure 8: The left plot shows the result of migrating a single shot using a static grid, the right plot shows the result using the changing grid described by algorithms 1 and 2. Note the spurious energy away from the main energy train using the changing grid. [**ER**] bob1/. single

DISCUSSION

The tests shown in this paper are limited to modeling and migration. Another obvious area to apply these techniques is velocity analysis. For waveform inversion techniques that rely on low frequencies this approach will lead to minimal speedup advantages but the cheap approximation of attenuation might prove useful. The big advantage is for Wave Equation Migration Velocity Analysis(Zhang and Biondi, 2014) and Total Full Waveform

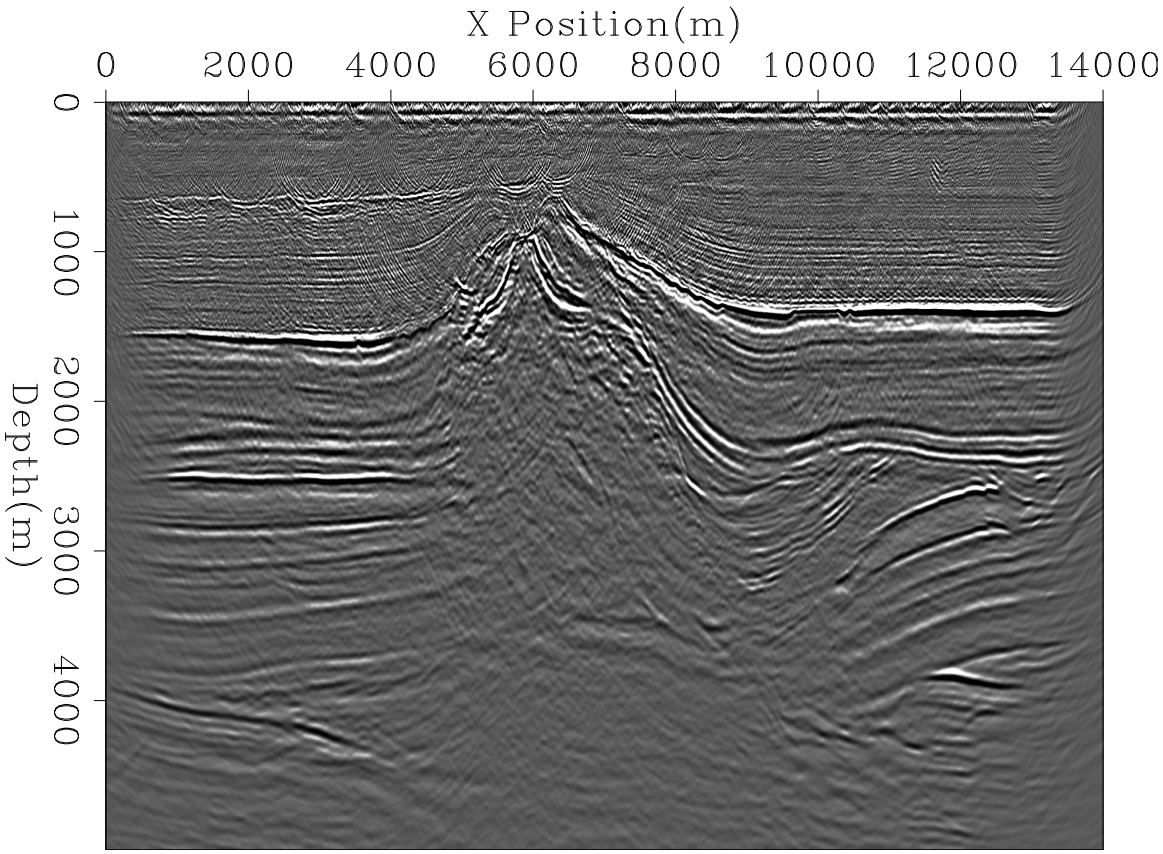


Figure 9: The result of migrating an entire dataset using the changing grid approach. Note how the artifacts seen in Figure 8 are not apparent in the final migration. [**CR**] bob1/. full

Inversion(Almomin, 2014) techniques. In addition to potentially a better approximation of matching the physics of the real data, these technique are interested in large time records and higher frequencies, which is where this approach leads to significant performance improvments. The reduction in data size, resulting from coarser sampling of the wave-field could also prove useful.

CONCLUSIONS

I use an approximation to Zhu and Harris (2014) to simulate propagation in an attenuated earth. I take advantage of the fact that frequencies decay with time by resampling my propagation grid at later times. Combining this approach with following the wave-field leads to significant speedups in modeling and RTM. The approach is also potentially useful for Wave Equation Migration Velocity Analysis and Total Full Waveform Inversion.

ACKNOWLEDGEMENTS

I would like to thank Yi Shen for the time-domain approximate-Q formulation and Gustavo Alves for the initial discussion that brought on this idea. I would like than Total for providing the data used in this paper.

REFERENCES

Almomin, A., 2014, Reformulating TFWI: SEP-Report, **155**, 193–198.

Bednar, J. B. and G. H. Neale, 2002, Limited vs. full frequency waveequation imaging, chapter 308, 1204–1207.

Clapp, R. G., 2002, Speeding up wave equation migration: SEP-Report, **111**, 233–241.

Courant, R., K. Friedrichs, and H. Lewy, 1967, On the partial difference equations of mathematical physics: IBM journal of Research and Development, **11**, 215–234.

Kjartansson, E., 1979, Attenuation of seismic waves in rocks and applications in energy exploration: **23**.

Stork, C., 2013, Eliminating dispersion from FD modeling is crucial for inversion: SEAM Workshop: Geoscience Advances from the use of SEAM Data.

Zhang, Y. and B. Biondi, 2014, Residual-moveout-based WEMVA: A WAZ field data example - Part I : SEP-Report, **152**, 223–238.

Zhu, T. and J. M. Harris, 2014, Modeling acoustic wave propagation in heterogeneous attenuating media using decoupled fractional laplacians: GEOPHYSICS, **79**, T105–T116.

Using a nonlinear wave equation for anisotropic inversion

Huy Le, Biondo Biondi, Robert G. Clapp and Stewart A. Levin

ABSTRACT

We present some preliminary results on a specific nonlinear pseudo-acoustic wave equation in anisotropic media, including forward modeling, linearization, and adjoint method. Our objective is to find a robust and efficient method for anisotropic full-waveform inversion (FWI). The wave equation is solved discretely by the rotated staggered finite-difference scheme (RSFD) in time and space domains. The solution is more accurate than the one obtained using the centered finite-difference (CFD) scheme. The linearized equation is derived by taking only the first-order dependence of the wavefield with respect to medium parameters. The nonlinearity of the pseudo-acoustic wave equation introduces an additional term in the linearized equation. The adjoint method provides a mean to compute the gradients of the least-squares misfit objective function with respect to medium parameters through the adjoint wavefield. As a result of solving the forward equation by RSFD, the medium parameters are located on two different grids. We show that the gradients computed by the derived adjoint equation are in fact collocated consistently with the medium parameters. Applications on simple models, in both vertically transverse isotropic (VTI) and orthorhombic media, show that they also lead to the correct update directions. These results show the potential of our method for anisotropic parameter estimation.

INTRODUCTION

Though anisotropy has been recognized by the industry to play an important role in seismic imaging and inversion, its multiparameter nature remains a great challenge. In an attempt to reduce the number of parameters, Alkhalifah (1998) introduced a pseudo-acoustic approximation, under which shear-wave velocities along the symmetry axes are set to zero. This approximation reduces the number of anisotropic parameters to three for VTI (v_{pz}, ϵ, and δ) and to six for orthorhombic media (v_{pz}, ϵ_i, and δ_i). The pseudo-acoustic approximation results in an equation for P-wave in frequency-wavenumber domain as follows:

$$\left(\omega^2 - v_{pz}^2 k^2 S\right) u = 0, \tag{1}$$

where u is the pressure wavefield, ω is the angular frequency, $\mathbf{k}$ is the wave vector, and $S = S\left(\mathbf{n}, \epsilon_i, \delta_i\right)$; with $\mathbf{n}$ being the normalized wave vector. Mathematically, S is a pseudo-differential operator, whose expressions for VTI and orthorhombic media can be found in the Appendix A (equations B-4 and B-5, respectively). Physically, it controls the degree of anisotropy along different propagation directions. For isotropic media, $S = 1$.

Equation 1 is a pseudo-differential equation, which can be computationally expensive to solve (Song and Alkhalifah, 2013; Le and Levin, 2014) because S incorporates all the

anisotropic parameters. To overcome the computational intensity of solving equation 1, Xu and Zhou (2014) introduced the following approximation:

$$\mathbf{n} \approx \frac{\nabla u}{|\nabla u|}. \tag{2}$$

By regarding the wavefront normal, $\mathbf{n}$, as the direction of greatest change in the pressure wavefield, ∇u, this approximation ignores any amplitude variation with angle, but is exactly correct for plane waves. For this reason, the approximation 3 might be called plane-wave approximation. Assuming local homogeneity, the linear pseudo-differential equation 1 now becomes a nonlinear differential equation:

$$\partial_t^2 u - v_{pz}^2 \nabla \cdot (S \nabla u) = 0. \tag{3}$$

The nonlinearity comes from the dependence of the scalar term, S, on the wavefield, u, as a result of the approximation 3.

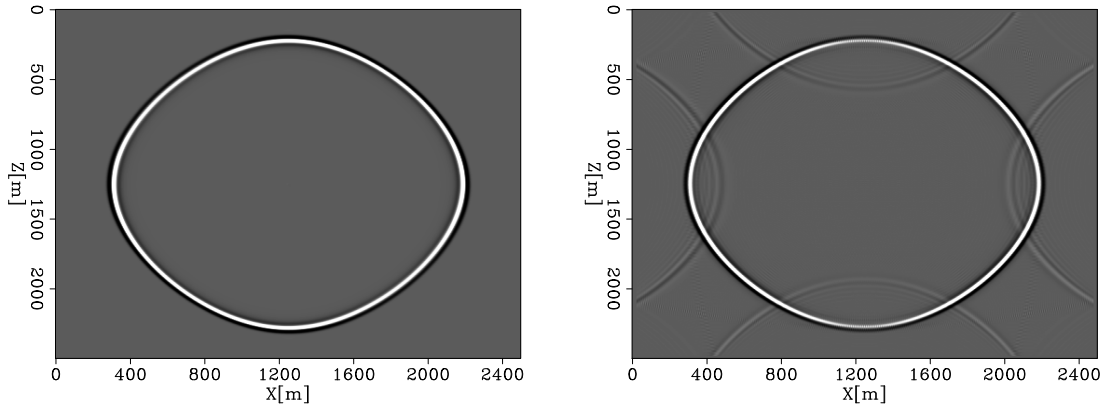
Forward modeling

Efficient solution to the forward modeling problem is crucial to any imaging and inversion process. Xu and Zhou (2014) solved equation 3 by a spectral method. However, using a spectral method to solve equation 3 is as expensive as solving equation 1. Taking advantage of the fact that equation 3 is no longer a pseudo-differential equation, we propose to use finite differences to solve this equation instead of spectral methods.

Equation 3 involves first derivatives, which can be inaccurately approximated by CFD. Figure 1 compares two snapshots of the wavefield solutions for pseudo-differential equation 1 by spectral method (panel a) and equation 3 by CFD (panel b). Similar artifacts as shown in Figure 1b were recognized by Ozdenvar and McMechan (1996) as a consequence of the nonlocality of the first-derivative CFD operator. To obtain an efficient and accurate solution, we use the RSFD scheme to solve equation 3 (Saenger et al., 2000).

Figure 2 shows a representative cell of the RSFD scheme in 2D, in which the pressure wavefield and its derivatives are located on two different grids that are staggered from each other: the main grid, denoted by circular nodes, and the staggered grid, denoted by square nodes. Consequently, the vertical P-wave velocity, which is associated with the wavefield, and the Thomsen parameters, which are associated with its first derivatives, are also on different grids. Figure 3a shows a snapshot of the wavefield using RSFD that is more accurate than the one using CFD (Figure 1b). However, when viewed at a harder clip, Figure 3b reveals high-frequency noise generated around the source injection area. This noise can be mitigated by a smaller time step or a larger source area.

Beside the high-frequency artifacts around the source area, the RSFD solution to equation 3 (Figure 3a) shows some amplitude differences in comparison to the solution to equation 1 by a spectral method (Figure 1a). These differences are expected because the plane-wave approximation 3 neglects amplitude variation along wavefronts. Despite these differences, the RSFD solution matches the traveltimes of the spectral solution.



(a) (b)

Figure 1: Wavefield snapshots of solutions to: (a) equation 1 by spectral method and (b) equation 3 by CFD in a homogeneous VTI medium. [**ER**] huyle/. acoustic.vti.k1,acoustic.vti.s11

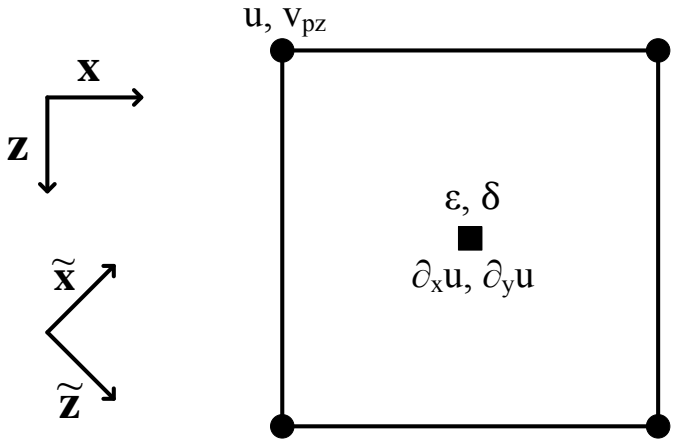
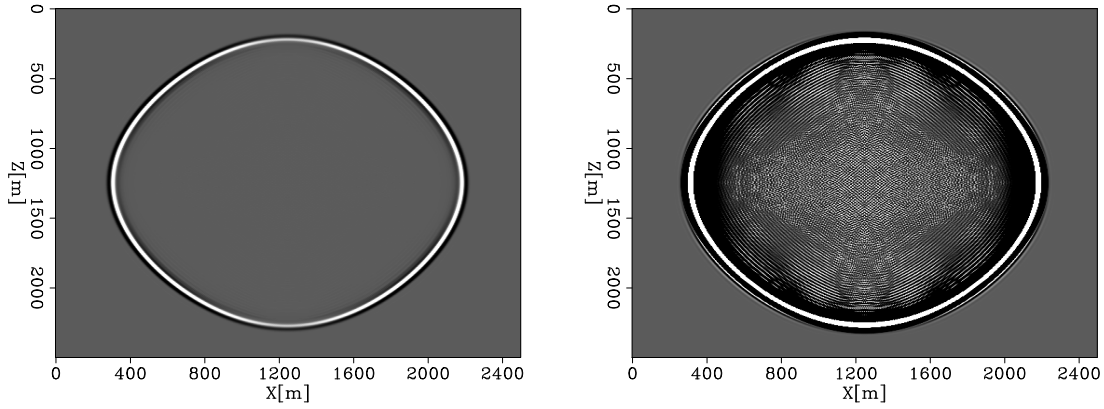


Figure 2: A representative cell of the RSFD scheme in 2D, in which the wavefield and velocity are located on the main grid while the anisotropic parameters and wavefield derivatives are located on the staggered grid. [**NR**] huyle/. rsgrid



(a) (b)

Figure 3: Modeling pseudo-acoustic wave equation in a homogeneous VTI medium using RSFD: (a) default clip and (b) hard clip revealing high-frequency noise around the source area. [**ER**] huyle/. acoustic.vti.s1,acoustic.vti.s1.clip

Linearization

Define $m_1 = \frac{1}{v_{pz}^2}$ and m_i with $i > 1$ being the anisotropic parameters, ϵ_i and δ_i, or some combinations of them. In VTI for example, one possibility is $m_2 = 1 + 2\epsilon$ and $m_3 = \epsilon - \delta$, whereas in orthorhombic media, m_i with $i = 2, ..., 6$ can just be the anisotropic parameters. With the source term, f, the wave equation 3 is re-written on the domain $\mathbf{\Omega}$, together with initial and boundary conditions, as:

$$m_1 \partial_t^2 u - \nabla \cdot (S \nabla u) = f, \tag{4a}$$
$$u(\mathbf{x}, 0) = 0, \quad \partial_t u(\mathbf{x}, 0) = 0, \tag{4b}$$
$$u|_{\partial \mathbf{\Omega}} = 0. \tag{4c}$$

These equations give a nonlinear relationship between the pressure wavefield and the model parameters. For the purpose of least squares reverse time migration and inversion, the first-order Born modeling operator is needed. The Born operator is the derivative of the modeling function with respect to model parameters. It is obtained by linearizing the forward modeling equations 4.

Perturb equations 4 to obtain (see Appendix A for a detailed derivation):

$$m_1 \partial_t^2 \delta u - \nabla \cdot (\mathbf{A} \nabla \delta u) = \delta f, \tag{5a}$$
$$\delta u(\mathbf{x}, 0) = 0, \quad \partial_t \delta u(\mathbf{x}, 0) = 0, \tag{5b}$$
$$\delta u|_{\partial \mathbf{\Omega}} = 0, \tag{5c}$$

where

$$\mathbf{A} = S\mathbf{I} + \nabla u \left(\frac{\partial S}{\partial \nabla u} \right)^T, \tag{6}$$

and

$$\delta f = -\partial_t^2 u \delta m_1 + \nabla \cdot \left[\left(\sum_{i>1} \frac{\partial S}{\partial m_i} \delta m_i \right) \nabla u \right]. \tag{7}$$

Equation 5 is a linear partial differential equation for the perturbed wavefield, δu, because the differential operator now does not depend on the perturbed wavefield but on the background wavefield, u. As a result, they give a linear relationship between the perturbed wavefield and the perturbed model pamareters. Compared to the nonlinear forward modeling equation 4a, the linearized equation 5a has an additional term, $\nabla u \left(\frac{\partial S}{\partial \nabla u} \right)^T$, incorporated in $\mathbf{A}$ (equation 6). This is a result of the nonlinearity of equation 4a.

Based on the fact that when the pertubation is small enough, the perturbed wavefield is a good approximation to the difference between the full and background wavefields to the first order, we design a test for the linearized equations 5 with a point perturbation. The amount of perturbation is 5% in velocity and 50% in Thomsen parameters. Figure 4 and Figure 5 are the test results for a VTI and an orthorhombic media respectively. Although these figures display high-frequency artifacts and amplitude differences as seen earlier (Figure 3), they show a similarity between the perturbed wavefield and the wavefield difference. This similarity validates the linearized equations 5.

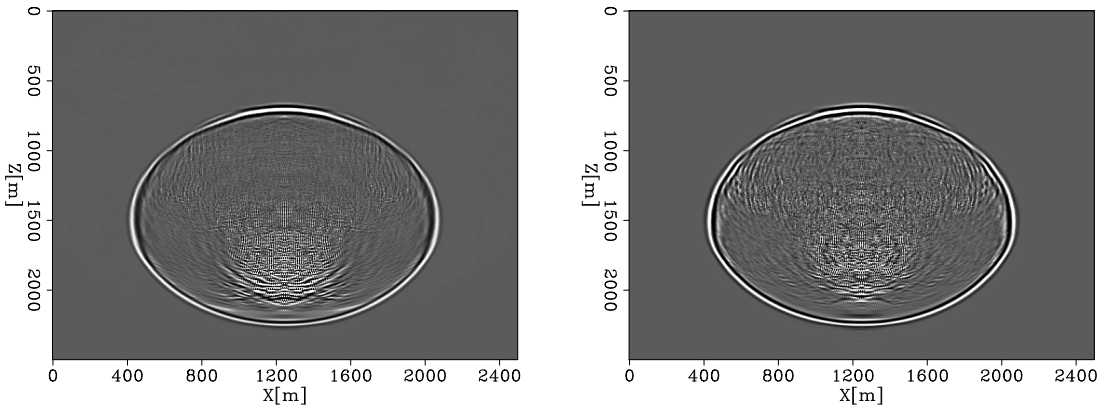


Figure 4: Linearization test in a VTI medium: (a) difference between the full and the background wavefields (b) the perturbed wavefield obtained by solving equation 5. [**ER**] huyle/. vti.d.snap.filtered,vti.l.snap.filtered

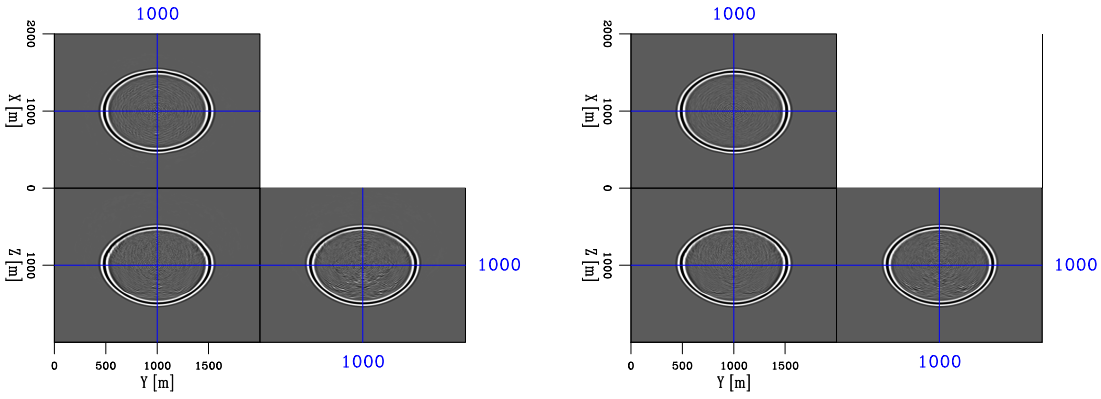


Figure 5: Linearization test in an orthorhombic medium: (a) difference between the full and the background wavefields (b) the perturbed wavefield obtained by solving equation 5. [**CR**] huyle/. filtered.ortho.diff.wfield1,filtered.ortho.lin.wfield1

Adjoint method

Gradients of the FWI objective function were originally derived by using Green's functions (Tarantola, 1984). Because the acoustic isotropic wave equation is a linear partial differential equation, it was possible to express its solution in terms of Green's functions. The equation at hand (equation 3) is, however, not linear; and therefore Green's theorem is not applicable in this case. As a result, we have to employ the adjoint method to derive the objective function's gradients.

The adjoint method provides an efficient way to obtain gradients of the FWI objective function through solution of an adjoint equation. Following Liu and Tromp (2006), the least-squares misfit function is defined as:

$$\chi = \frac{1}{2}\sum_r \int_0^T \|u(\mathbf{x}_r,t) - d(\mathbf{x}_r,t)\|^2 dt, \tag{8}$$

where $\mathbf{x}_r$ are the receivers' locations, T is the final time, and $d(\mathbf{x}_r,t)$ is the observed data.

The augmented Lagrangian is:

$$\chi = \frac{1}{2}\sum_r \int_0^T \|u(\mathbf{x}_r,t) - d(\mathbf{x}_r,t)\|^2 dt + \int_0^T \int_{\mathbf{\Omega}} \lambda \left[m_1 \partial_t^2 u - \nabla\cdot(S\nabla u) - f\right] dV dt, \tag{9}$$

where $\int_{\mathbf{\Omega}} dV$ is the volumetric integration.

The perturbed augmented Lagrangian is:

$$\begin{aligned}
\delta\chi = &\int_0^T \int_{\mathbf{\Omega}} \sum_r \left[u(\mathbf{x}_r,t) - d(\mathbf{x}_r,t)\right] \delta(\mathbf{x}-\mathbf{x}_r)\delta u dV dt \\
&+ \int_0^T \int_{\mathbf{\Omega}} \lambda \left[\delta m_1 \partial_t^2 u - \nabla\cdot(\delta S\nabla u) - \delta f\right] dV dt \\
&+ \int_0^T \int_{\mathbf{\Omega}} \lambda \left[m_1 \partial_t^2 \delta u - \nabla\cdot(S\nabla\delta u)\right] dV dt,
\end{aligned} \tag{10}$$

which, after some integration by parts and algebra manipulations (Appendix B), becomes:

$$\begin{aligned}
\delta\chi = &\int_0^T \int_{\mathbf{\Omega}} \sum_r \left[u(\mathbf{x}_r,t) - d(\mathbf{x}_r,t)\right] \delta(\mathbf{x}-\mathbf{x}_r)\delta u dV dt \\
&+ \int_0^T \int_{\mathbf{\Omega}} \left[\delta m_1 \lambda \partial_t^2 u + \nabla\lambda\cdot\left(\sum_{i>1}\frac{\partial S}{\partial m_i}\delta m_i\right)\nabla u - \lambda\delta f\right] dV dt \\
&+ \int_0^T \int_{\mathbf{\Omega}} \left[m_1\partial_t^2\lambda - \nabla\cdot\left(\mathbf{A}^T\nabla\lambda\right)\right] \delta u dV dt \\
&+ \int_{\mathbf{\Omega}} m_1 \left(\lambda\partial_t\delta u - \partial_t\lambda\delta u\right)|_T dV \\
&- \int_0^T \oint_{\partial\mathbf{\Omega}} \hat{\mathbf{n}}\cdot\lambda\left[\left(\sum_{i>1}\frac{\partial S}{\partial m_i}\delta m_i\right)\nabla u + \mathbf{A}\nabla\delta u\right] ds dt,
\end{aligned} \tag{11}$$

where $\delta(\mathbf{x}-\mathbf{x}_r)$ is the delta function centered at the receivers' locations and $\oint_{\partial\mathbf{\Omega}} ds$ is the surface integration.

Define the adjoint equations, with final and boundary conditions, as:

$$m_1 \partial_t^2 \lambda - \nabla \cdot \left(\mathbf{A}^T \nabla \lambda \right) = - \sum_r \left[u(\mathbf{x}_r, t) - d(\mathbf{x}_r, t) \right] \delta(\mathbf{x} - \mathbf{x}_r), \tag{12a}$$

$$\lambda(\mathbf{x}, T) = 0, \quad \partial_t \lambda(\mathbf{x}, T) = 0, \tag{12b}$$

$$\lambda|_{\partial \mathbf{\Omega}} = 0. \tag{12c}$$

Notice that, as with the linearized equation (equation 5), the above equation is also linear with respect to the adjoint wavefield because the differential operator of equation 12 does not depend on λ but on the background wavefield, u. Now equation 11 becomes:

$$\delta \chi = \int_0^T \int_{\mathbf{\Omega}} \left[\delta m_1 \lambda \partial_t^2 u + \nabla \lambda \cdot \left(\sum_{i>1} \frac{\partial S}{\partial m_i} \delta m_i \right) \nabla u - \lambda \delta f \right] dV dt. \tag{13}$$

Partial derivatives of the misfit function with respect to the model parameters can now be read from equation 13 as:

$$\frac{\partial \chi}{\partial m_1} = \int_0^T \lambda \partial_t^2 u dt, \tag{14}$$

and for $i > 1$:

$$\frac{\partial \chi}{\partial m_i} = \int_0^T \nabla \lambda \cdot \frac{\partial S}{\partial m_i} \nabla u dt. \tag{15}$$

These derivatives can be easily casted in terms of Thomsen paramaters, for example in VTI, as:

$$\frac{\partial \chi}{\partial \epsilon} = 2 \frac{\partial \chi}{\partial m_2} + \frac{\partial \chi}{\partial m_3}, \tag{16}$$

and

$$\frac{\partial \chi}{\partial \delta} = - \frac{\partial \chi}{\partial m_3}. \tag{17}$$

Equation 14 shows that the gradient of the objective function with respect to the squared vertical P-wave slowness is calculated from the adjoint wavefield and the second-order time derivative of the forward wavefield, both of which are located on the main grid. Equation 15 shows that the gradients with respect to m_i with $i > 1$ or the Thomsen parameters are computed from first derivatives of the forward and adjoint wavefields, which are located on the staggered grid. These computations are consistent with the locations where these parameters are defined (Figure 2).

We have solved the forward modeling and adjoint equations for simple cases. The first one is a model with a small perturbation point at the center of a homogeneous VTI medium. There are 800 receivers at 5 meters apart placed everywhere on the surface. Figures 6a through 6d show, respectively, the model, full data, and background data; as well as the residual for one shot. The full and background data are displayed at clip of 95% to reveal the reflection from the perturbation. These data and the residual are contaminated by the high-frequency artifacts mentioned earlier. Figures 7a through 7f show snapshots of the source wavefield and adjoint wavefield. Although the adjoint wavefield is degraded by noise, it is still able to focus at the perturbation point at 0.4 second. Cross-corelation and stacking over multiple shots help to reduce the effect of this noise. Figures 8a through 8c show, respectively, gradients of slowness squares, ϵ, and δ, using 40 shots at 100 meters spacing everywhere on the surface. These gradients focus clearly at the perturbation point.

In another model, the background consists of two homogeneous VTI layers and the perturbation is a rectangular box embedded in the top layer (Figure 9). We use positive perturbations in velocity and Thomsen parameters. There are 800 receivers and 40 shots placed everywhere on the surface with similar spacing as in the previous example. Figures 10a through 10c show to the full data, background data, and the residual for one shot. Figure 11 shows the gradients of the objective function with respect to: squared vertical P-wave slowness (panel a), ϵ (panel b), and δ (panel c). These figures show that the gradients accurately identify the perturbation, and indeed, lead to the correct update directions. Similarities between the slowness squares and ϵ gradients can be observed. These similarities might indicate cross-talks between these two parameters. The δ gradient is the most accurate in identifying the perturbation's shape but is the weakeast in terms of magnitude relative to the other two gradients. The small magnitude of the δ gradient is an indication of limited ability to invert for δ from surface reflection data. Additionally, notice the V-shaped sensitivity regions of the shots that are located on the edges of the model.

We carried out 100 inversion iterations in the time domain using a Ricker wavelet of 40-Hertz (Hz) fundamental frequency. We employ the steepest decent algorithm and calculate the step length from a parabolic interpolation in the search direction (Vigh and Starr, 2008). Figure 12 shows the model updates after the first iteration (left column) and after 100 iterations (right column) in velocity (top row), ϵ (middle row), and δ (bottom row). Improvements in resolving the magnitudes and shape of the perturbations are noticeable after 100 inversion iterations compared to the first-iteration results. Figure 13 shows the objective function decreasing with iterations.

For orthorhombic media, we only computed the sensitivity kernels for one shot and one receiver in different directions in a homogeneous subsurface. These kernels are shown in Figures 14, 15, and 16 for sources and receivers separated in the x-, y-, and z-directions, respectively. It can be noted from these figures that the sensitivity kernels of slowness squares are always nonzero, regardless of the source-receiver's direction. On the other hands, ϵ_i and δ_i are insensitive in certain planes when the source and receiver are apart in certain directions. This difference between the sensitivity kernels of slowness squares and Thomsen paramaters might be because slowness or velocity has first-order influence on seismic signature in anisotropic media in comparison to Thomsen parameters. The insensitivity of Thomsen parameters in certain planes, however, will not affect the ability to invert for these parameters when there are more sources and receivers covering wider apertures and azimuths. It can also be observed that the planes of insensitivity interchange between ϵ_i and between δ_i when the source and receiver are separated in different directions. For example, ϵ_1 is insensitive in the xz-plane when the source and receiver are apart in the x-direction; whereas, ϵ_2 is insensitive in the yz-plane when the source and receiver are apart in the y-direction. This interchangeability is consistent with how these parameters are defined: ϵ_1 is the Thomsen parameter in the yz-plane; whereas, ϵ_2 is the equivalent parameter in the xz-plane. By these definitions, the two parameters' roles are interchangeable.

CONCLUSIONS

With numerical examples for VTI and orthorhombic media, I have shown that the nonlinear pseudo-acoustic wave equation in anisotropic media can be modeled, linearized, and adjointed. The forward modeling equation can be solved accurately using RSFD. Its nonlin-

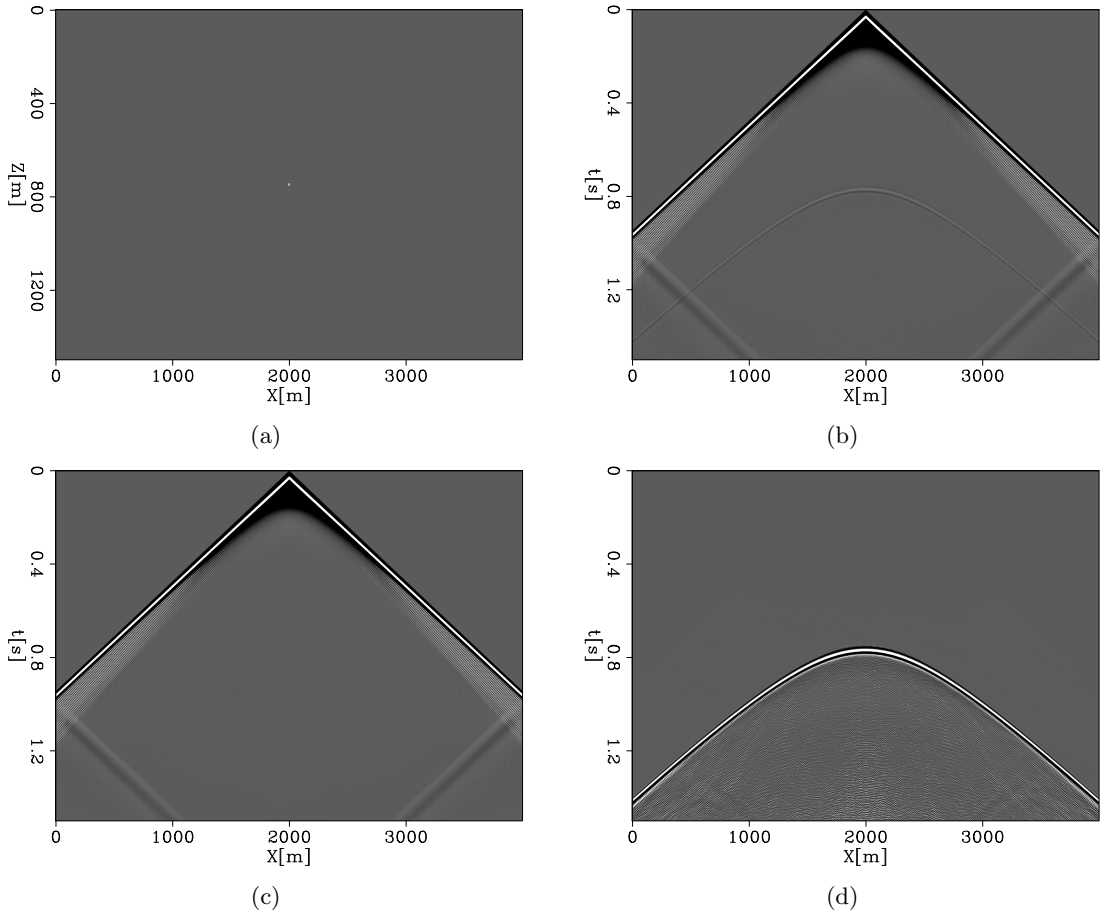


Figure 6: (a) Model with a small perturbation point in a homogeneous VTI medium, (b) full data, (c) background data, and (d) residual. The data, displayed at clip of 95%, and the residual are contaminated by high-frequency noise. [**ER**] huyle/. v.perturb1,data,bg.data,residual

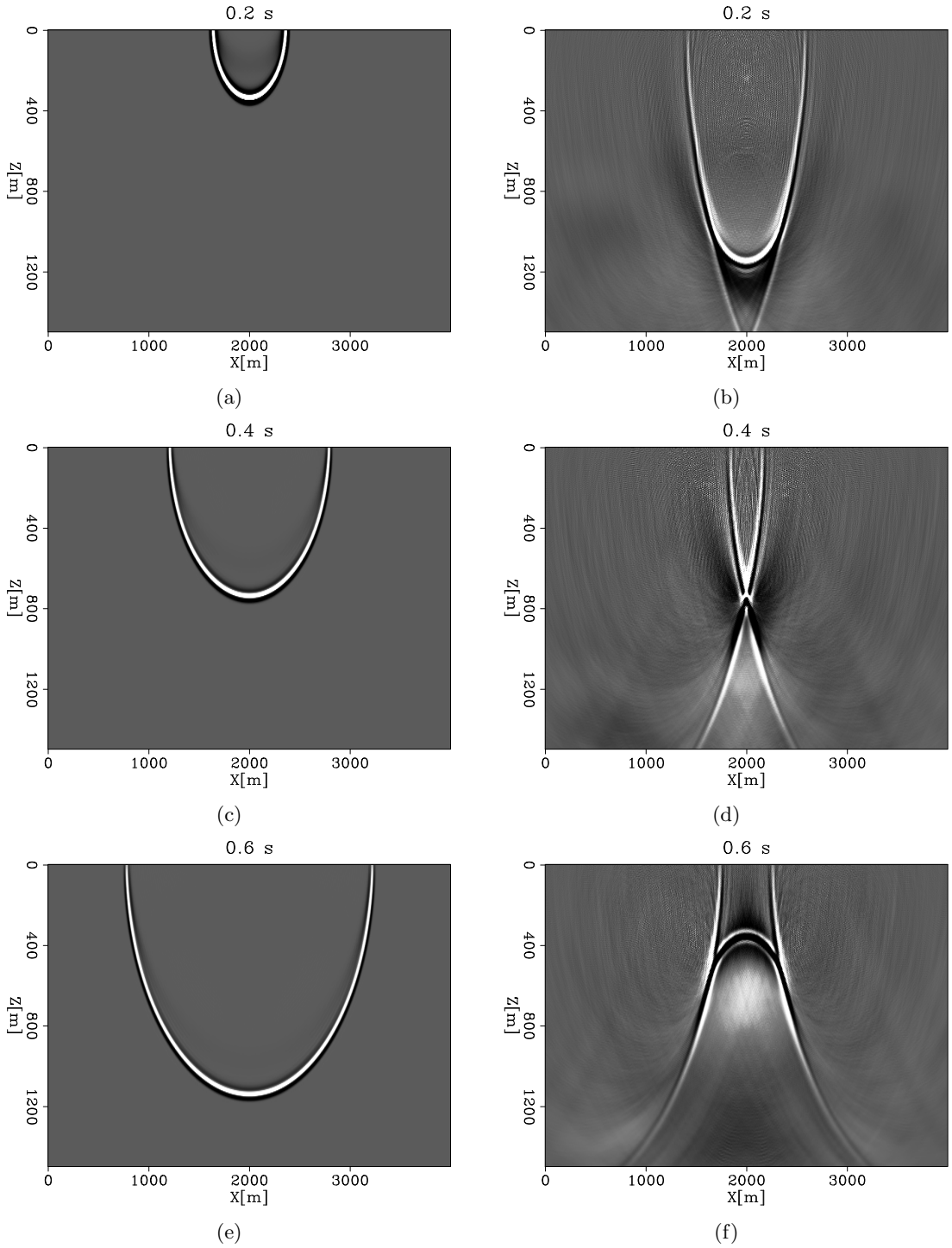


Figure 7: Snapshots of the source wavefield (left column) and adjoint wavefield (right column) at different times: (a) and (b) 0.2 second, (c) and (d) 0.4 second, and (e) and (f) 0.6 second. Although the adjoint wavefield is severely degraded by high-frequency noise, it is still able to focus at the perturbation point at 0.4 second (panel d). [**CR**] huyle/. sou1,rec1,sou2,rec2,sou3,rec3

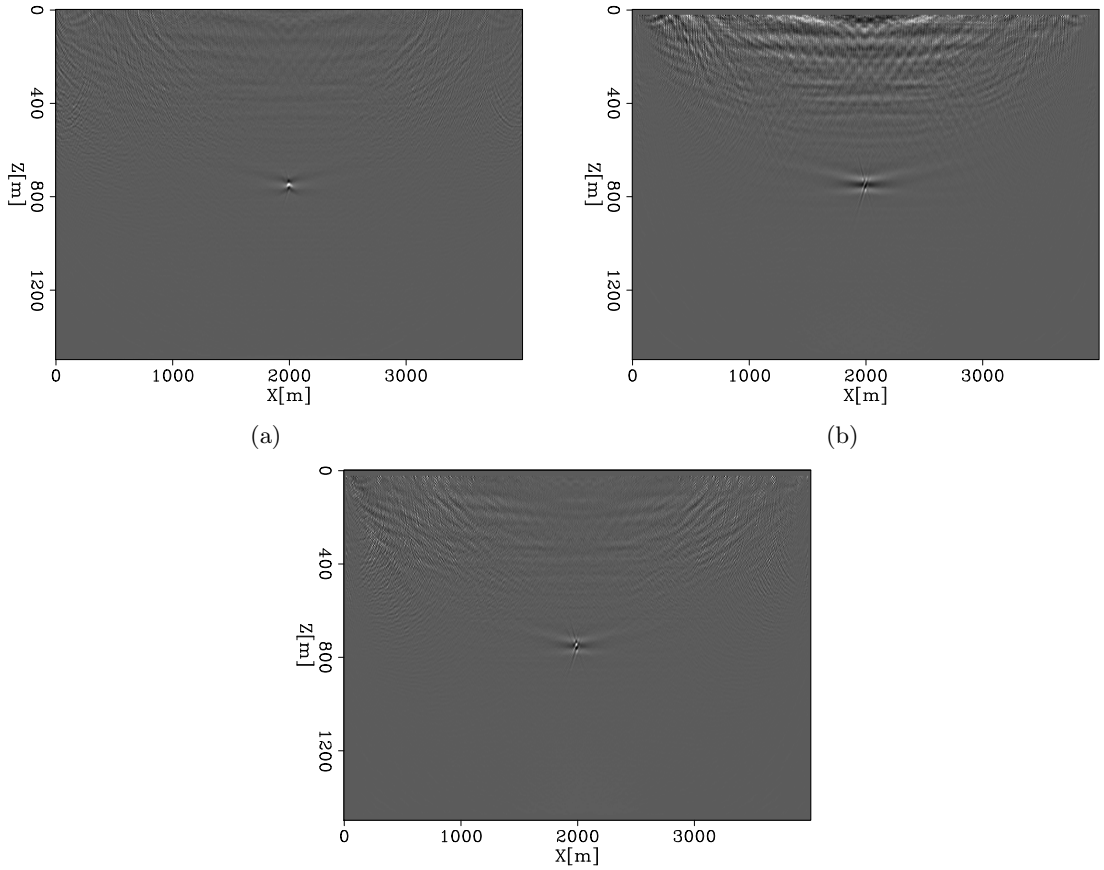


Figure 8: Gradients of the objective function with respect to: (a) squared vertical P-wave slowness, (b) ϵ, and (c) δ, for a simple model with a point perturbation in a homogeneous VTI medium. These gradients focus at the perturbation points. [**CR**] huyle/. v.grad1.all,eps.grad1.all,del.grad1.all

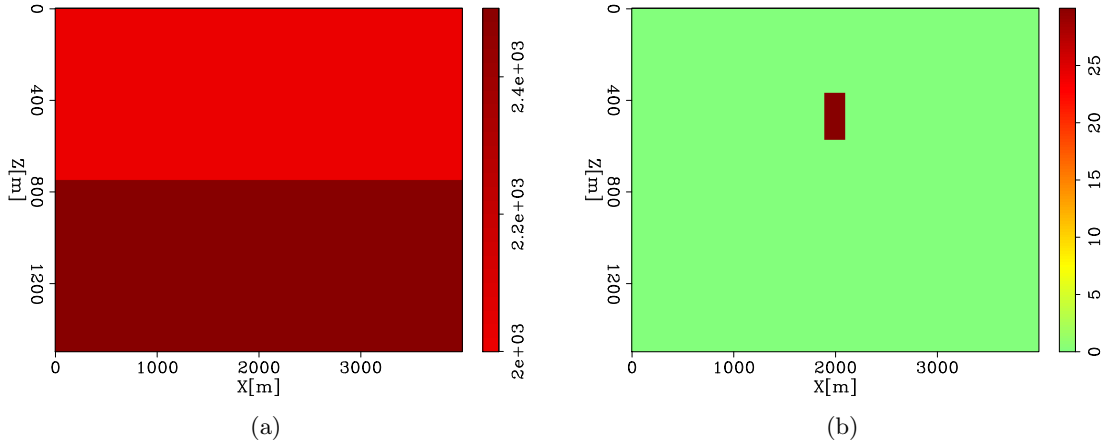


Figure 9: Model for testing gradient calculation: (a) background and (b) perturbed vertical P-wave velocity. Models for Thomsen parameters are similar. [**ER**] huyle/. v.bg,v.perturb

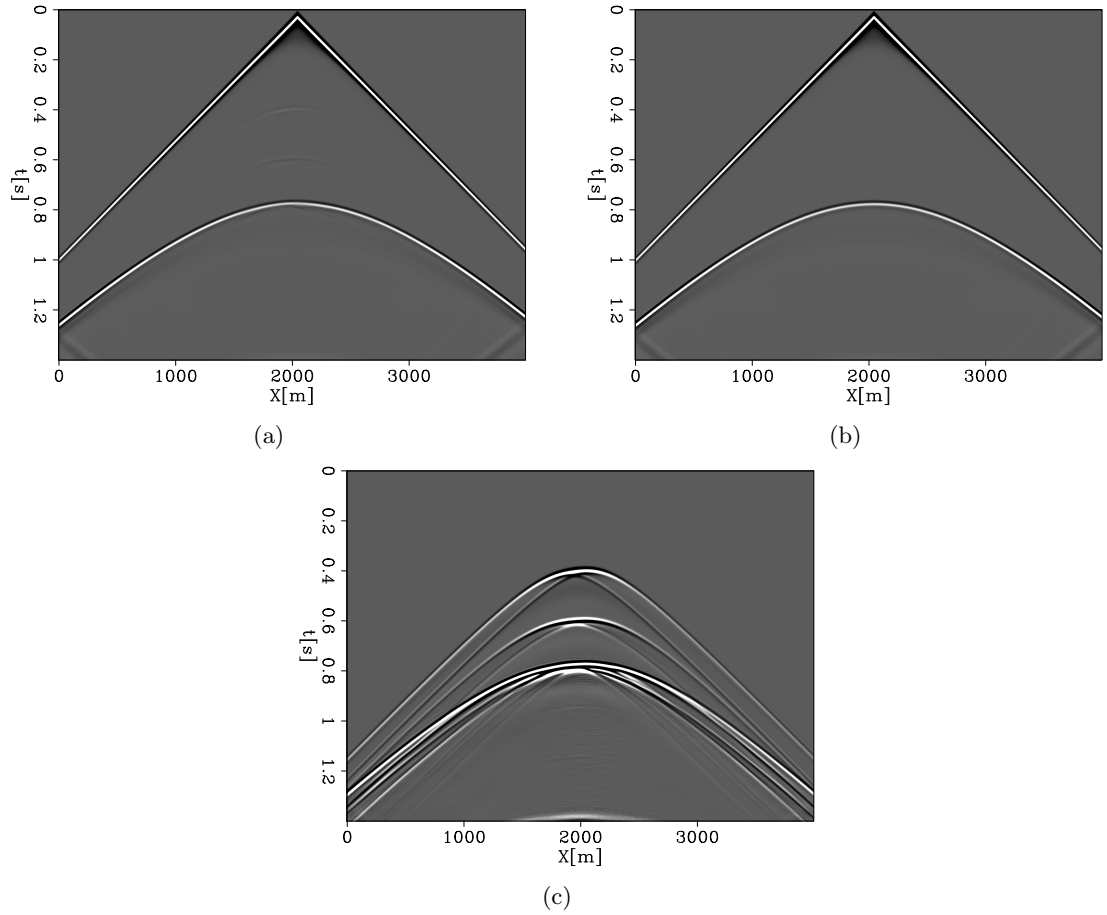


Figure 10: (a) Full data, (b) background data, and (c) residual for a model with two VTI layers and a rectangular perturbation. [**ER**] huyle/. data2,bg.data2,residual2

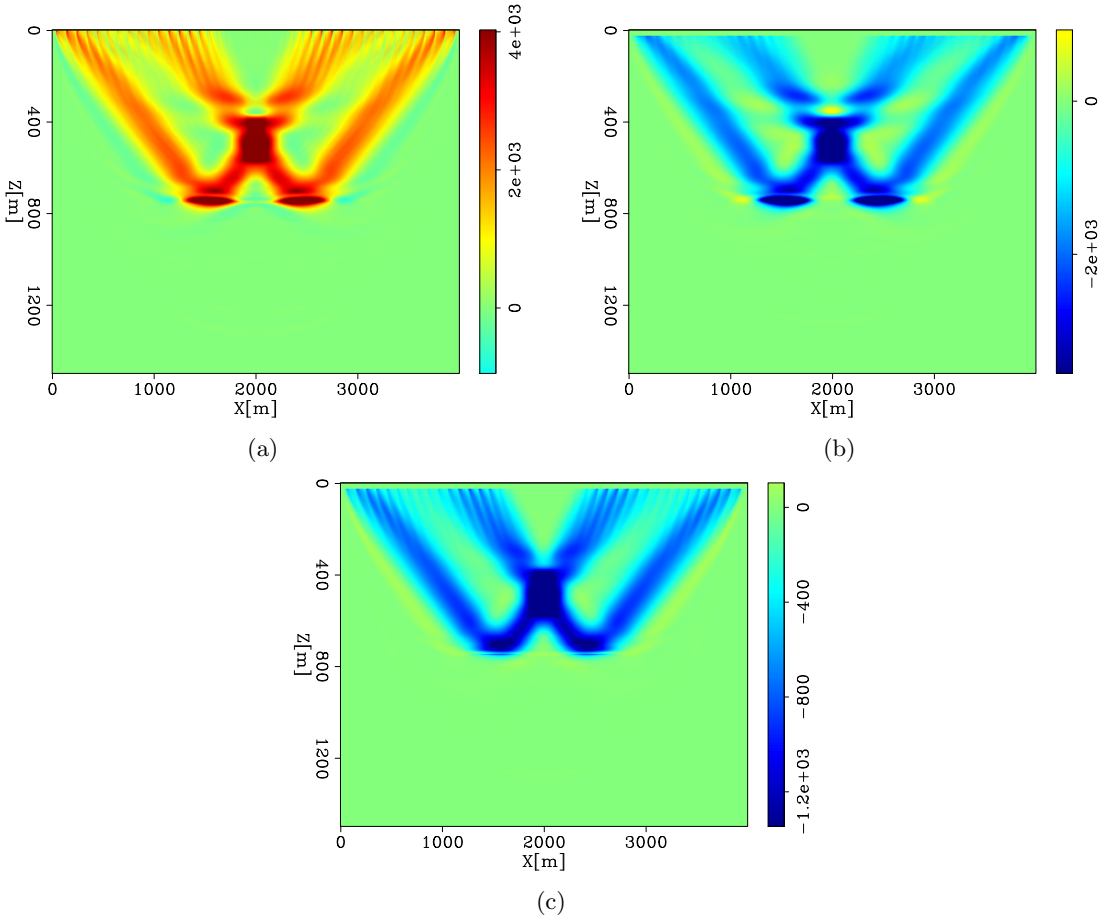


Figure 11: Gradients of the objective function with respect to: (a) squared vertical P-wave slowness, (b) ϵ, and (c) δ, for a model with two VTI layers and a rectangular perturbation. These gradients accurately locate the perturbation and have correct update directions. [**CR**] huyle/. s2.grad,eps.grad,del.grad

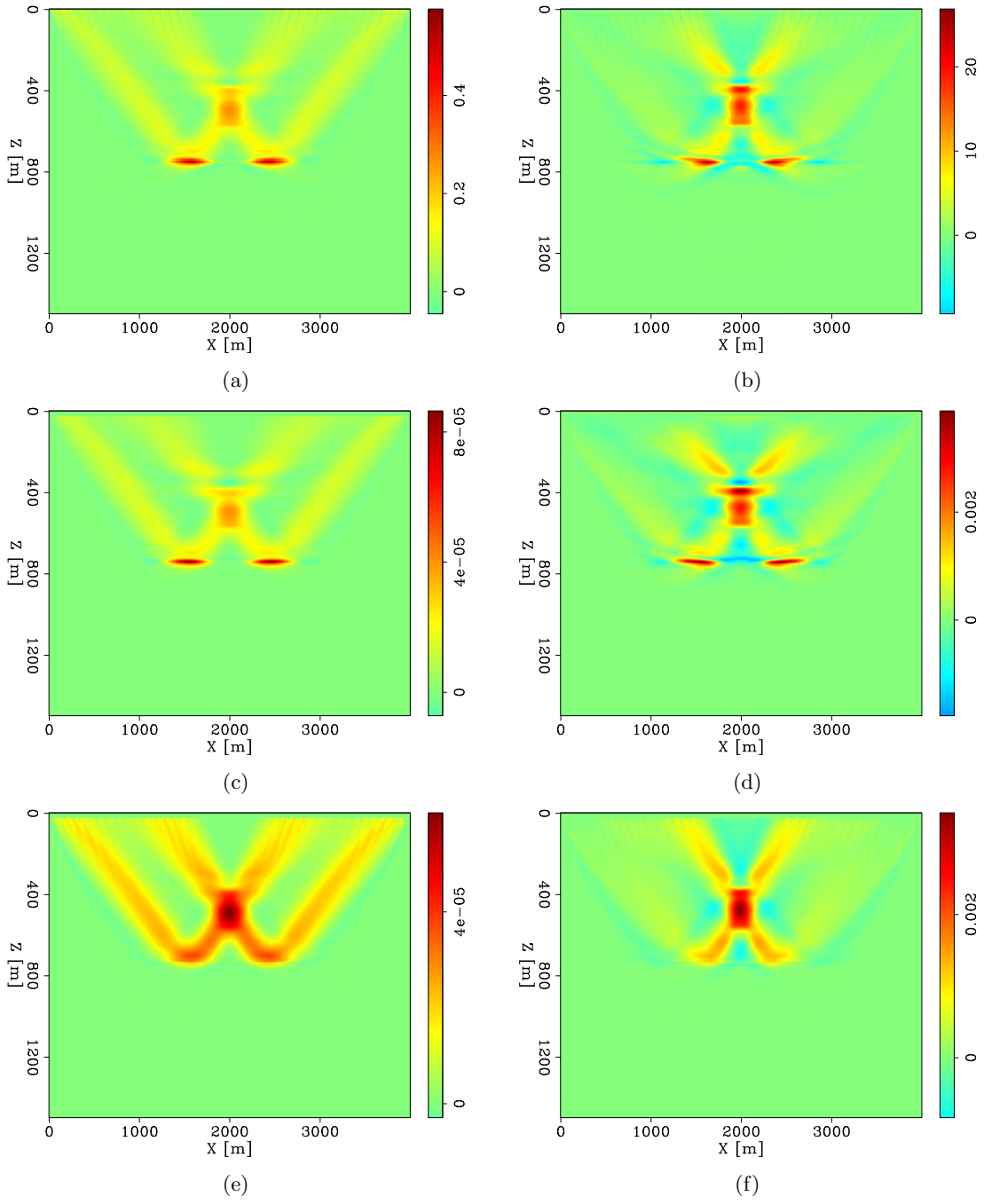


Figure 12: Model updates after the first iteration (left column) and after 100 iterations (right column) in: velocity (m/s) (top row), ϵ (middle row), and δ (bottom row). The updates show improvements after 100 iterations compared to the first-iteration updates. [**CR**] huyle/. v.1iter,v.100iter,eps.1iter,eps.100iter,del.1iter,del.100iter

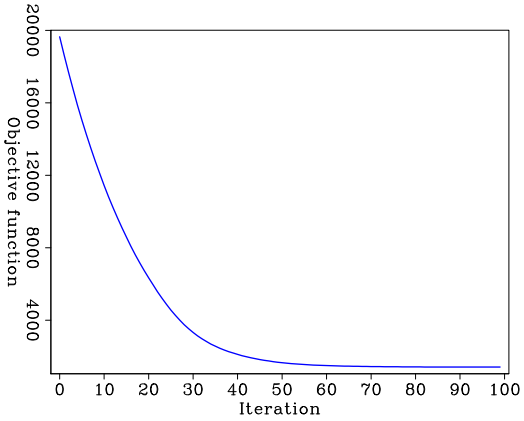


Figure 13: Plot of the objective function with iteration. [**CR**] huyle/. objfn1

earity introduces an additional term in the linearized and adjoint equations. Together, these three equations can be used in an inversion framework to estimate medium parameters.

APPENDIX A

In this appendix, I derive the linearized equation 5. The perturbed wave equation is:

$$(m_1 + \delta m_1)\partial_t^2(u + \delta u) - \nabla \cdot [(S + \delta S)\nabla(u + \delta u)] = f. \tag{B-1}$$

Neglect the second-order terms:

$$m_1\partial_t^2\delta u - \nabla \cdot (S\nabla\delta u) = -\partial_t^2 u\delta m_1 + \nabla \cdot (\delta S\nabla u). \tag{B-2}$$

After the plane-wave approximation (equation 3), $S = S(\nabla u, m_i)$ with $i > 1$. Consequently, to the first order:

$$\delta S = \left(\frac{\partial S}{\partial \nabla u}\right)^T \nabla\delta u + \sum_{i>1} \frac{\partial S}{\partial m_i}\delta m_i, \tag{B-3}$$

which is subtituted into equation B-2 to obtain:

$$m_1\partial_t^2\delta u - \nabla \cdot (S\nabla\delta u) = -\partial_t^2 u\delta m_1 + \nabla \cdot \left\{ \left[\left(\frac{\partial S}{\partial \nabla u}\right)^T \nabla\delta u + \sum_{i>1} \frac{\partial S}{\partial m_i}\delta m_i \right] \nabla u \right\}.$$

Collecting the terms that involve δu to obtain the linearized equation 5:

$$m_1\partial_t^2\delta u - \nabla \cdot \left\{ \left[S\mathbf{I} + \nabla u \left(\frac{\partial S}{\partial \nabla u}\right)^T \right] \nabla\delta u \right\} = -\partial_t^2 u\delta m_1 + \nabla \cdot \left[\left(\sum_{i>1} \frac{\partial S}{\partial m_i}\delta m_i \right) \nabla u \right].$$

In VTI (Xu and Zhou, 2014):

$$S = \frac{1}{2}\left[n_a^2 + \sqrt{n_a^4 - 8\left(\epsilon - \delta\right) n_x^2 n_z^2} \right], \tag{B-4}$$

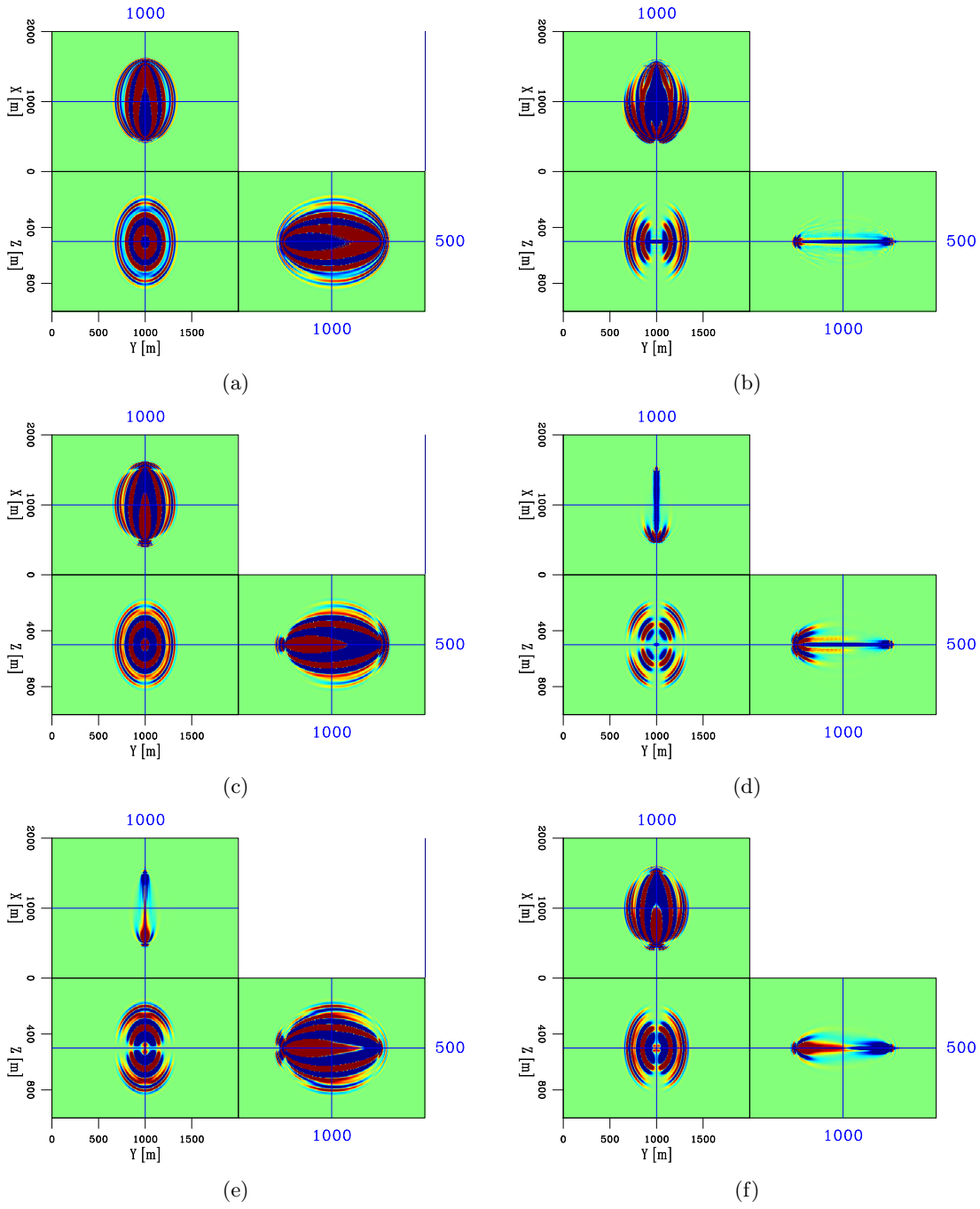


Figure 14: Sensitivity kernels for one shot and one receiver, separated in the x-direction, in a homogeneous orthorhombic subsurface of: (a) squared vertical P-wave slowness, (b) ϵ_1, (c) ϵ_2, (d) δ_1, (e) δ_2, and (f) δ_3. [**CR**] huyle/. g.s2.x,g.eps1.x,g.eps2.x,g.del1.x,g.del2.x,g.del3.x

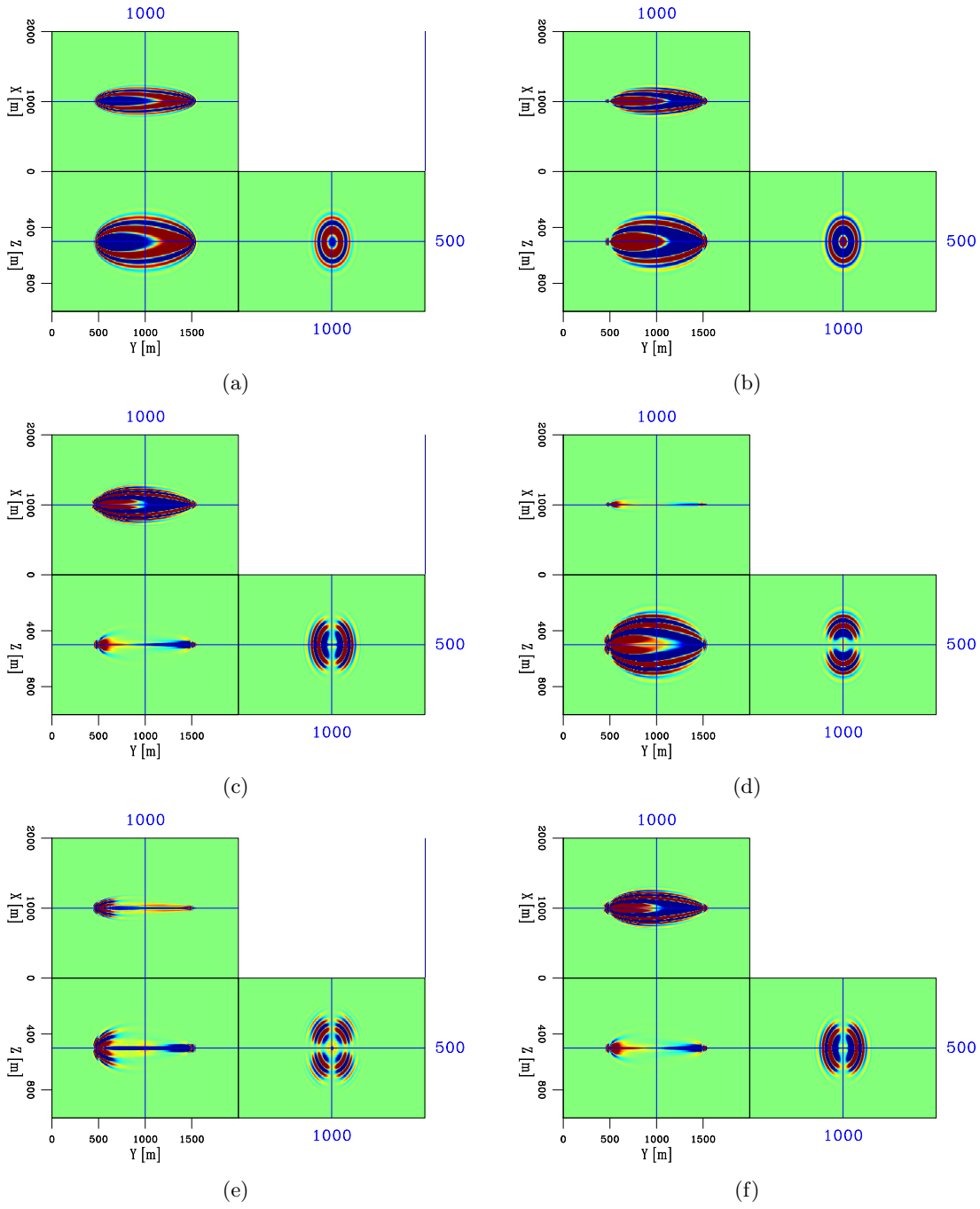


Figure 15: Sensitivity kernels for one shot and one receiver, separated in the y-direction, in a homogeneous orthorhombic subsurface of: (a) squared vertical P-wave slowness, (b) ϵ_1, (c) ϵ_2, (d) δ_1, (e) δ_2, and (f) δ_3. [**CR**] huyle/. g.s2.y,g.eps1.y,g.eps2.y,g.del1.y,g.del2.y,g.del3.y

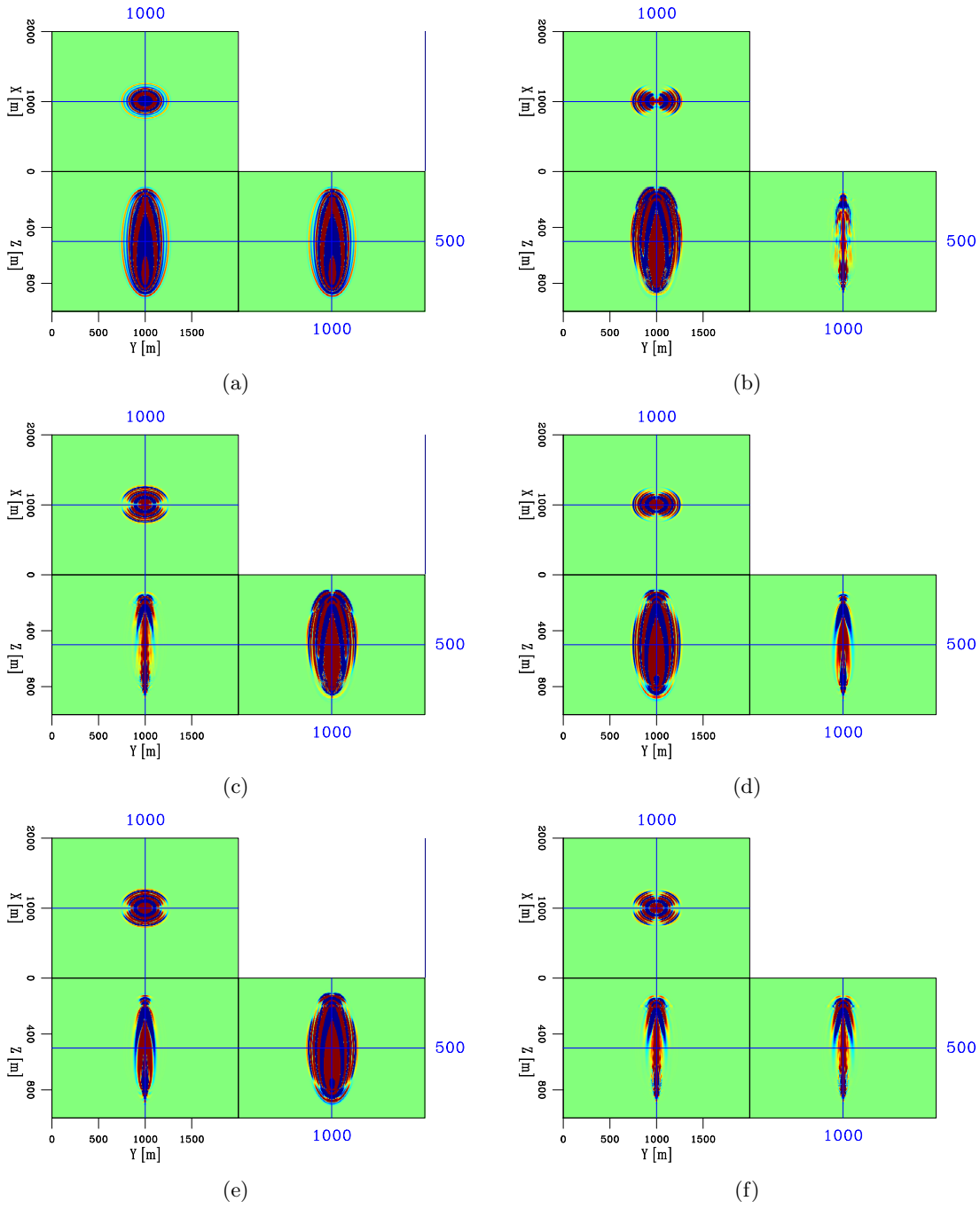


Figure 16: Sensitivity kernels for one shot and one receiver, separated in the z-direction, in a homogeneous orthorhombic subsurface of: (a) squared vertical P-wave slowness, (b) ϵ_1, (c) ϵ_2, (d) δ_1, (e) δ_2, and (f) δ_3. [**CR**] huyle/. g.s2.z,g.eps1.z,g.eps2.z,g.del1.z,g.del2.z,g.del3.z

where n_x and n_z are components of the normalized wave vector, $\mathbf{n}$, and $n_a^2 = (1+2\epsilon)\, n_x^2 + n_z^2$.

In orthorhombic media (Song and Alkhalifah, 2013):

$$S = \frac{1}{3}\left[a - \frac{d}{\sqrt[3]{2}} - \frac{\sqrt[3]{2}\left(a^2+3b\right)}{d}\right], \tag{B-5}$$

where:

$$a = (2\epsilon_2+1)\, n_x^2 + (2\epsilon_1+1)\, n_y^2 + n_z^2,$$
$$b = \left[(2\epsilon_2+1)^2 (2\delta_3+1) - (2\epsilon_2+1)(2\epsilon_1+1)\right] n_x^2 n_y^2 - 2(\epsilon_1-\delta_1)\, n_y^2 n_z^2 - 2(\epsilon_2-\delta_2)\, n_x^2 n_z^2,$$
$$d = \sqrt[3]{-2a^3 + 3(e-9c) - 9ab},$$
$$c = \left[2(2\epsilon_2+1)\sqrt{(2\delta_1+1)(2\delta_2+1)(2\delta_3+1)} - (2\epsilon_2+1)^2 (2\delta_3+1) + \epsilon_1\epsilon_2 - \epsilon_1\delta_2 - \epsilon_2\delta_1\right] n_x^2 n_y^2 n_z^2,$$
$$e = \sqrt{3b^2(a^2+4b) - 6ac(2a^2+9b) - 81c^2}.$$

APPENDIX B

In this appendix, I present a step-by-step derivation of equation 11. Substitute perturbation in S (equation B-3) into equation 10 to obtain:

$$\begin{aligned}\delta\chi = &\int_0^T \int_{\mathbf{\Omega}} \sum_r \left[u(\mathbf{x}_r,t) - d(\mathbf{x}_r,t)\right] \delta(\mathbf{x}-\mathbf{x}_r)\delta u dV dt \\ &+ \int_0^T \int_{\mathbf{\Omega}} \lambda \left\{ \delta m_1 \partial_t^2 u - \nabla\cdot\left[\left(\sum_{i>1} \frac{\partial S}{\partial m_i}\delta m_i\right)\nabla u\right] - \delta f \right\} dV dt \\ &+ \int_0^T \int_{\mathbf{\Omega}} \lambda \left[m_1 \partial_t^2 \delta u - \nabla\cdot(\mathbf{A}\nabla\delta u)\right] dV dt.\end{aligned} \tag{C-1}$$

Integration by parts gives, for $i > 1$:

$$\begin{aligned}\int_0^T \int_{\mathbf{\Omega}} \lambda \nabla\cdot\left(\frac{\partial S}{\partial m_i}\delta m_i \nabla u\right) dV dt = &\int_0^T \oint_{\partial\mathbf{\Omega}} \hat{\mathbf{n}}\cdot\left(\lambda \frac{\partial S}{\partial m_i}\delta m_i \nabla u\right) ds dt \\ &- \int_0^T \int_{\mathbf{\Omega}} \nabla\lambda\cdot\left(\frac{\partial S}{\partial m_i}\delta m_i \nabla u\right) dV dt,\end{aligned}$$

$$\begin{aligned}\int_0^T \int_{\mathbf{\Omega}} \lambda m_1 \partial_t^2 \delta u dV dt &= \int_{\mathbf{\Omega}} \lambda m_1 \partial_t \delta u|_0^T dV - \int_0^T \int_{\mathbf{\Omega}} m_1 \partial_t \lambda \partial_t \delta u dV dt \\ &= \int_{\mathbf{\Omega}} m_1 \left(\lambda \partial_t \delta u - \partial_t \lambda \delta u\right)|_0^T dV + \int_0^T \int_{\mathbf{\Omega}} m_1 \partial_t^2 \lambda \delta u dV dt \\ &= \int_{\mathbf{\Omega}} m_1 \left(\lambda \partial_t \delta u - \partial_t \lambda \delta u\right)|_T dV + \int_0^T \int_{\mathbf{\Omega}} m_1 \partial_t^2 \lambda \delta u dV dt,\end{aligned}$$

and:

$$\begin{aligned}
\int_0^T \int_{\mathbf{\Omega}} \lambda \nabla \cdot (\mathbf{A} \nabla \delta u) dV dt &= \int_0^T \oint_{\partial \mathbf{\Omega}} \hat{\mathbf{n}} \cdot \lambda \mathbf{A} \nabla \delta u ds dt - \int_0^T \int_{\mathbf{\Omega}} \nabla \lambda \cdot (\mathbf{A} \nabla \delta u) dV dt \\
&= \int_0^T \oint_{\partial \mathbf{\Omega}} \hat{\mathbf{n}} \cdot \lambda \mathbf{A} \nabla \delta u ds dt - \int_0^T \int_{\mathbf{\Omega}} (\mathbf{A}^T \nabla \lambda) \cdot \nabla \delta u dV dt \\
&= \int_0^T \oint_{\partial \mathbf{\Omega}} \hat{\mathbf{n}} \cdot \left(\lambda \mathbf{A} \nabla \delta u - \mathbf{A}^T \nabla \lambda \delta u \right) ds dt \\
&\quad + \int_0^T \int_{\mathbf{\Omega}} \nabla \cdot \left(\mathbf{A}^T \nabla \lambda \right) \delta u dV dt \\
&= \int_0^T \oint_{\partial \mathbf{\Omega}} \hat{\mathbf{n}} \cdot \lambda \mathbf{A} \nabla \delta u ds dt + \int_0^T \int_{\mathbf{\Omega}} \nabla \cdot \left(\mathbf{A}^T \nabla \lambda \right) \delta u dV dt.
\end{aligned}$$

where I have exploited the perturbed initial and boundary conditions (equations 5b and 5c). Substituting the above three integrations into equation C-1 results in equation 11.

REFERENCES

Alkhalifah, T., 1998, Acoustic approximations for processing in transversely isotropic media: Geophysics, **63**, 623–631.

Le, H. and S. A. Levin, 2014, Removing shear artifacts in acoustic wave propagation in orthorhombic media: SEG Annual International Meeting, Expanded Abstracts, 486–490.

Liu, Q. and J. Tromp, 2006, Finite-frequency kernels based on adjoint methods: Bulletin of the Seismological Society of America, **96**, 2383–2397.

Ozdenvar, T. and G. A. McMechan, 1996, Causes and reduction of numerical artefacts in pseudo-spectral wavefield extrapolation: Geophysical Journal International, **126**, 819–828.

Saenger, E. H., N. Gold, and S. A. Shapiro, 2000, Modeling the propagation of elastic waves using a modified finite-difference grid: Wave Motion, **31**, 77–92.

Song, X. and T. Alkhalifah, 2013, Modeling of pseudoacoustic p-waves in orthorhombic media with a low-rak approximation: Geophysics, **78**, no. 4, C33–C40.

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

Vigh, D. and E. W. Starr, 2008, 3D prestack plane-wave, full-waveform inversion: Geophysics, **73**, no. 5, VE135–VE144.

Xu, S. and H. Zhou, 2014, Accurate simulations of pure quasi-P-waves in complex anisotropic media: Geophysics, **79**, no. 6, T341–T348.

Pseudo-acoustic modeling for tilted anisotropy with pseudo-source injection

Musa Maharramov and Stewart A. Levin

ABSTRACT

We provide a general framework for deriving fast finite-difference algorithms for the numerical modeling of acoustic wave propagation in anisotropic media. We specialize this framework to the case of tilted transversely isotropic media to implement a kinematically accurate fast finite-difference modeling method. This results in a significant reduction of the shear artifacts compared to similar kinematically accurate finite-difference methods.

INTRODUCTION

Transverse isotropy and orthorhombic media are of significant interest for industrial applications such as seismic imaging and inversion in complex, fractured rocks (Grechka, 2009). While full elastic data and models are needed to fully understand and invert for parameters in such media, the lesser task of imaging in the presence of such anisotropy can often be accomplished under a pseudo-acoustic assumption. In particular, the pseudo-acoustic method of Alkhalifah (1998) is the anisotropic counterpart of isotropic acoustic modeling. However, this and similar anisotropic finite-difference methods suffer from shear artifacts or rely on approximations that break down for strong anisotropy (Fowler et al., 2010; Zhan et al., 2012). We note that both references discuss transverse isotropy but similar challenges exist for finite-difference modeling in orthorhombic media.

In this work we develop a computationally efficient finite-difference wave propagation modeling method for tilted transversely isotropic (TTI) media that is largely free of shear artifacts. The concept extends the approach that Maharramov (2014, 2015) formulated for vertically transversal isotropic (VTI) media, but is not limited to polar anisotropy.

Our derivation of pseudo-acoustic (systems of) equations for a specific medium symmetry can be described as a three-step process:

1) Derive a phase velocity surface (Musgrave, 1970) as a function of the angle of propagation.

2) Derive a dispersion relation from 1) (Alkhalifah, 1998).

3) Interpret the dispersion relation as an evolutionary pseudo-differential equation, and transform it into a form suitable for numerical solution.

The cause of numerical artifacts is that the pressure and shear wave velocity surfaces remain coupled after deriving computationally feasible equations in step 3 (more specifically, the pressure mode and one of the shear modes remain coupled).

Our method can be summarized as follows:

2′) After step 1) above, extract the branch of the phase velocity surface corresponding to the pressure wave velocity.

3′) Approximate the resulting $V^2 = F(\mathbf{m}, \theta)$, where V is the pressure wave velocity, $\mathbf{m}$ stands for medium parameters, and θ is the propagation direction, with a computationally efficient numerical Fourier operator. This can be a trigonometric polynomial in θ (Iserles, 2008) with coefficients depending on $\mathbf{m}$, as practiced in some of the existing spectral pseudo-acoustic modeling methods (Etgen and Brandsberg-Dahl, 2009). We opt to use a pseudo-differential operator spatially constrained to a narrow depth range of sources and receivers.

4) Derive a coupled pseudo-pressure, pseudo-shear differential equation system analogous to Alkhalifah (2000).

5) At each time step apply the spatial component of the pseudo-differential operator derived in step 3′) to the injected source[1] using a spectral method with spatial interpolation. This results in a "pseudo-source".

6) Inject appropriate linear combinations of the pseudo-source and the source into the primary and secondary component of the system derived in 4).

When we assume a VTI anisotropy, and that the system described in step 4) is that of (Alkhalifah, 2000), step 6) reduces to injecting the pseudo-source into the secondary component and the true source into the primary component of the system derived in step 4).

THE PSEUDO-DIFFERENTIAL MODELING OPERATOR

In step 1) we start with the equation for $V(\theta)$ in a VTI medium (Tsvankin, 1996)

$$\begin{aligned} \frac{V^2(\theta)}{V_P^2} &= 1 + \epsilon \sin^2\theta - \frac{f}{2} \pm \frac{f}{2}\sqrt{\left(1 + \frac{2\epsilon \sin^2\theta}{f}\right)^2 - \frac{2(\epsilon-\delta)\sin^2 2\theta}{f}}, \\ \text{with } f &= 1 - \frac{V_S^2}{V_P^2}, \text{ and} \\ \sin\theta &= \frac{V(\theta)\,[k_x]}{\left[\frac{\partial}{\partial t}\right]}, \quad \cos\theta = \frac{V(\theta)\,[k_z]}{\left[\frac{\partial}{\partial t}\right]} \end{aligned} \tag{1}$$

where V_P and V_S are vertical pressure and shear wave velocities, θ is the propagation angle measured from the transverse isotropy symmetry axis, and ϵ and δ are the Thomsen parameters (Thomsen, 1986). We assume that $V_S = 0$, as we are not interested in propagating shear modes, thus $f = 1$. Note that here and in the subsequent analysis we consider two-dimensional TTI, however, the results naturally extend to three dimensions by identifying k_x with the radial wavenumber—see, e.g., Maharramov and Nolte (2011). We use the equivalence $k_u = -i\frac{\partial}{\partial u}$ in (1), where u is an arbitrary variable, to stress that the phase velocity equation can be interpreted as both a dispersion relation and a pseudo-differential

[1]This includes back-propagating receiver data in applications such as reverse time migration.

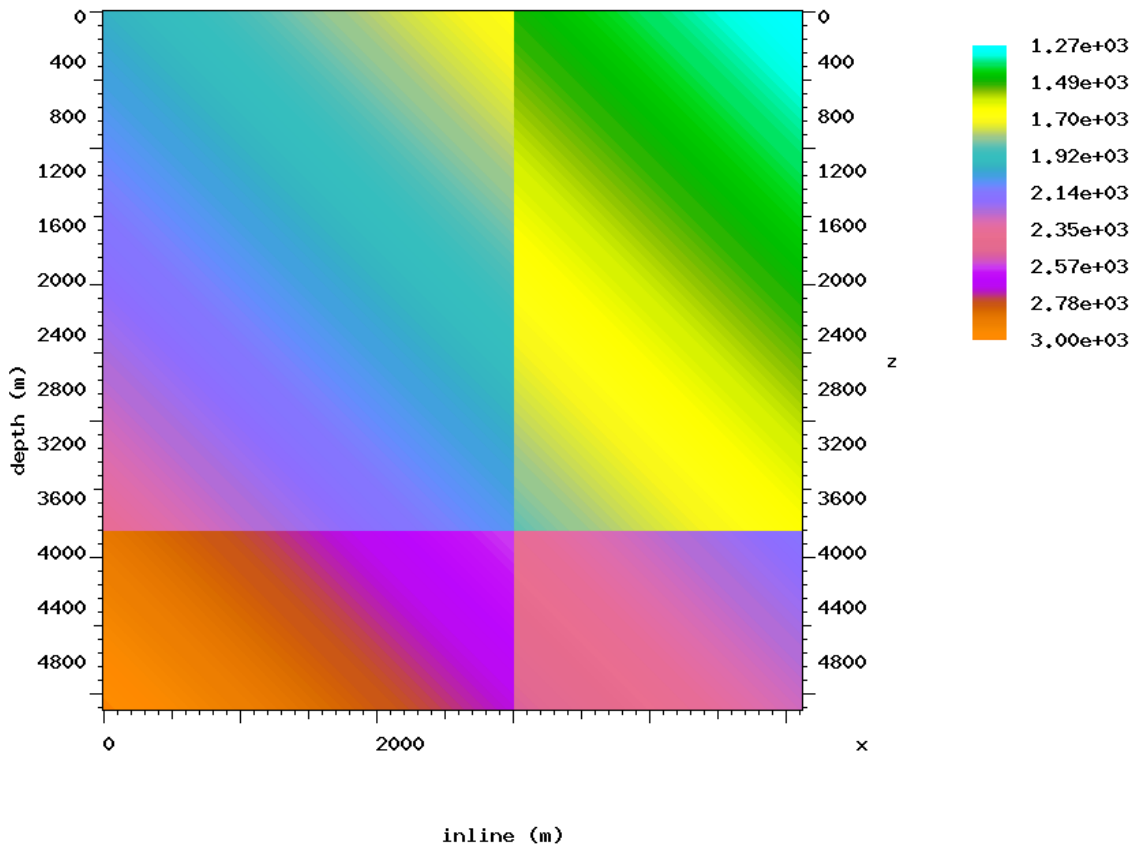


Figure 1: Test model with smooth and sharp V_P gradients and constant $\epsilon = 0.3$, $\delta = 0.1$, and tilt $\phi = 45°$. [**CR**] musa2/. model1

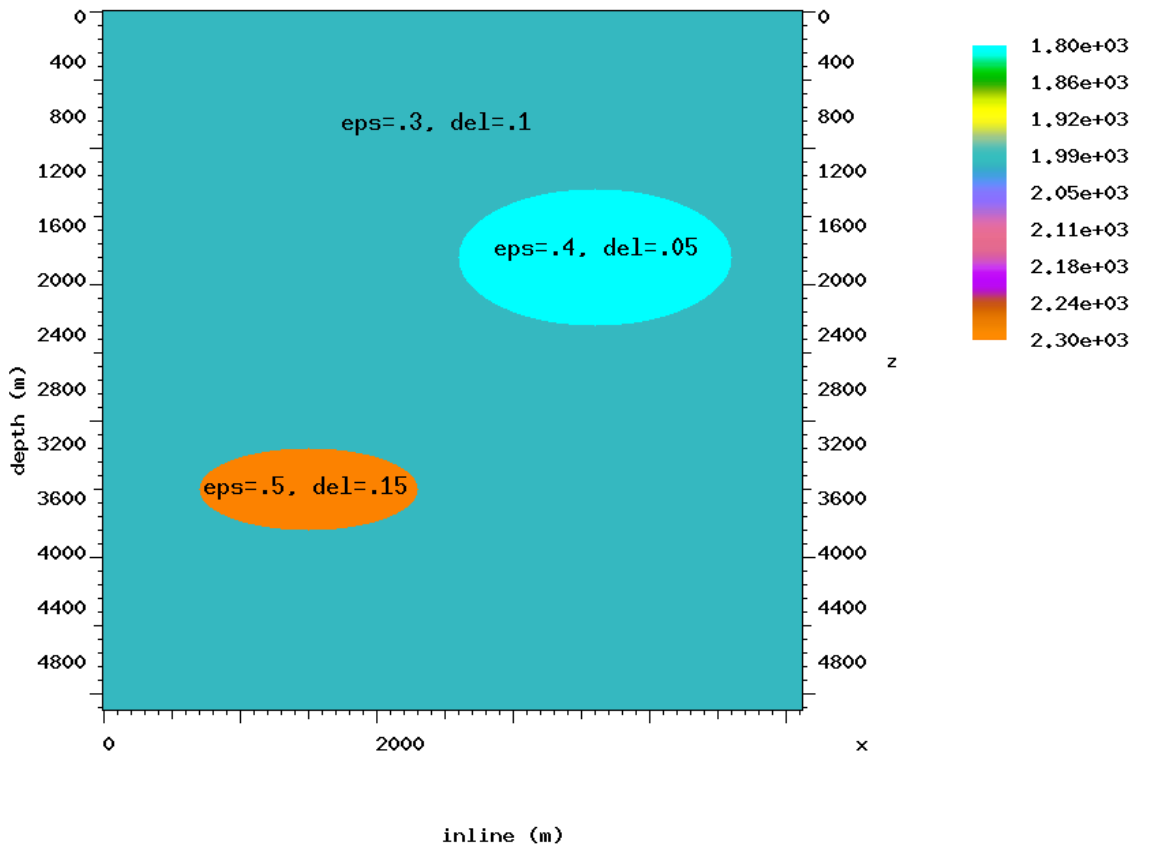


Figure 2: Test model with two anisotropic inclusions. The tilt angle is equal to 35° and 25° within the upper and lower inclusions, and is 30° in the background [**CR**] musa2/. model2

operator. In step 2′), we extract the branch of the square root with the positive sign in (1), corresponding to the (higher) compressional wave velocity. The resulting dispersion relation can be interpreted as an evolutionary pseudo-differential operator

$$\begin{aligned}&\frac{\partial^2}{\partial t^2} - V_P^2\left(\overset{2}{z},\overset{2}{x}\right)\frac{\Delta}{2} - \epsilon\left(\overset{2}{z},\overset{2}{x}\right)V_P^2\left(\overset{2}{z},\overset{2}{x}\right)\frac{\partial^2}{\partial x'^2} \;=\\ &V_P^2\left(\overset{2}{z},\overset{2}{x}\right)\frac{\Delta}{2}\sqrt{\left[1+2\epsilon\left(\overset{2}{z},\overset{2}{z}\right)\frac{\partial^2}{\partial x'^2}\frac{1}{\Delta}\right]^2 - 8\left(\epsilon\left(\overset{2}{z},\overset{2}{x}\right)-\delta\left(\overset{2}{z},\overset{2}{x}\right)\right)\frac{\partial^2}{\partial x'^2}\frac{\partial^2}{\partial z'^2}\frac{1}{\Delta^2}}\;,\end{aligned} \tag{2}$$

governing kinematically accurate propagation of the pressure wave, where x', z' are locally rotated coordinates with z' pointing along the tilted symmetry axis and x' pointing in the radial direction,

$$\Delta \;=\; \frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial z'^2} \;=\; \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}$$

is the Laplace operator (which is unaffected under rotation), and the "2" over x and z means that the multiplication by functions of spatial variables follows the application of differential operators in the pseudo-differential operator sense (Maslov, 1979). This is equivalent to "freezing" the operator coefficients, or assuming local homogeneity. Note that TTI in two dimensions requires one more parameter defined for each point of the subsurface: tilt angle ϕ from the vertical. This parameter implicitly enters (2) in the rotated coordinates x' and z'. Solving (2) for arbitrary heterogeneous media may be numerically challenging, because the Thomsen parameters $\epsilon(z,x)$ and $\delta(z,x)$ appear inside the square root of a pseudo-differential operator. However, operator (2) may simplify numerically if it is applied to a function with spatially bounded support – e.g., a source wavelet or receiver data. As noted earlier, an alternative to solving the full pseudo-differential operator equation (2) is to approximate, in step 3′), the extracted pressure velocity branch with a trigonometric polynomial:

$$V^2(\theta) \;\approx\; V_P^2\sum_{n=0}^{N} a_n \sin^{2n}(\theta), \tag{3}$$

where the coefficients $a_n,\ n=0,\dots,N$ depend on medium parameters. From the last line of (1) we can see that velocity surface (3) translates into the following pseudo-differential operator equation

$$\frac{\partial^2}{\partial t^2} \;=\; V_P^2\sum_{n=0}^{N} a_n \frac{\partial^{2n}}{\partial x'^{2n}}\Delta^{1-n}. \tag{4}$$

Equation (4) can be solved by applying the operators

$$\frac{\partial^{2n}}{\partial x'^{2n}}\Delta^{1-n}$$

to the wave field in the spatial Fourier domain, then summing up the results with spatially-dependent coefficients a_n in the spatial domain. Important particular cases of approximation (3) are the weak anisotropy approximation (Grechka, 2009)

$$V^2(\theta) \;\approx\; V_P^2\left(1+\delta\sin^2\theta + \frac{\epsilon-\delta}{1+2\delta}\sin^4\theta\right), \tag{5}$$

and the VTI approximation due to Harlan and Lazear (Harlan, 1998) used by Etgen and Brandsberg-Dahl (2009)

$$V^2(\theta) \;=\; V_P^2\cos^2\theta + \left(V_{P\mathrm{NMO}}^2 - V_{P\mathrm{Hor}}^2\right)\cos^2\theta\sin^2\theta + V_{P\mathrm{Hor}}^2\sin^2\theta, \tag{6}$$

where the subscripts PHor and PNMO denote the horizontal and NMO compressional wave velocities, respectively. Note that both (5) and (6) correspond to $N = 2$ in (3) and are suitable for weakly anisotropic VTI but break down in strong anisotropy. The case of $N = 3$ requires one additional inverse FFT for VTI but is accurate for a wide range of Thomsen parameters within (and beyond) practical requirements. Adapting (3) for TTI media would require the application at each time step of 5 additional inverse FFTs for $N = 2$ and extra 16 inverse FFTs for $N = 3$.

Solving (4) for $N = 2, 3$ using the described spectral method is an efficient modeling method in its own right, especially for VTI media where the number of FFTs at each time step is very low. However, in the next section we describe a finite-difference method that can outperform the spectral method for complex media and conceptually generalizes for other kinds of anisotropy.

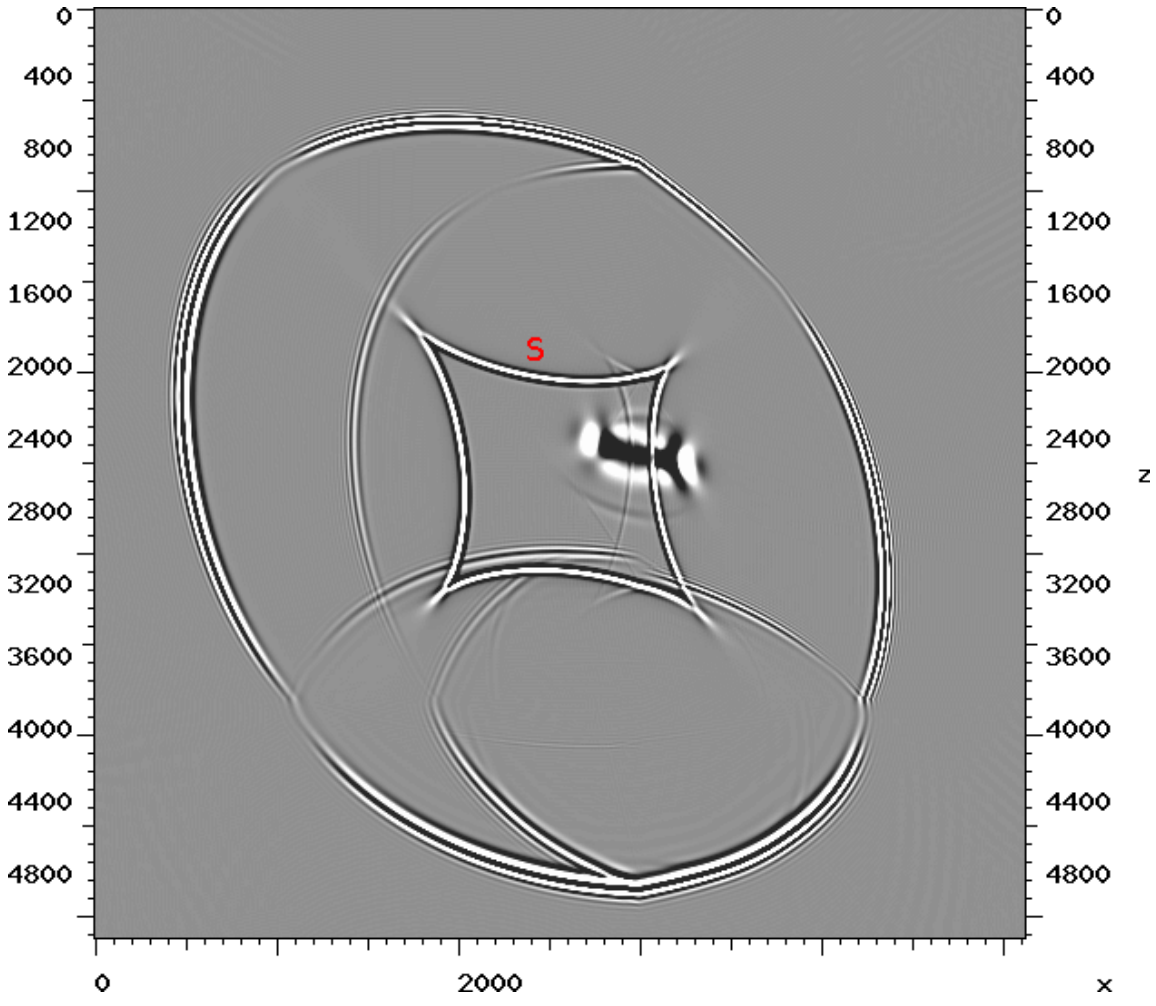


Figure 3: Shear artifact (marked with the "S") in the solution of (7) for the model of Figure 1 with sources injected in component r. Note that the shear artifact is causing numerical instability. [**CR**] musa2/. q1fdpdonored

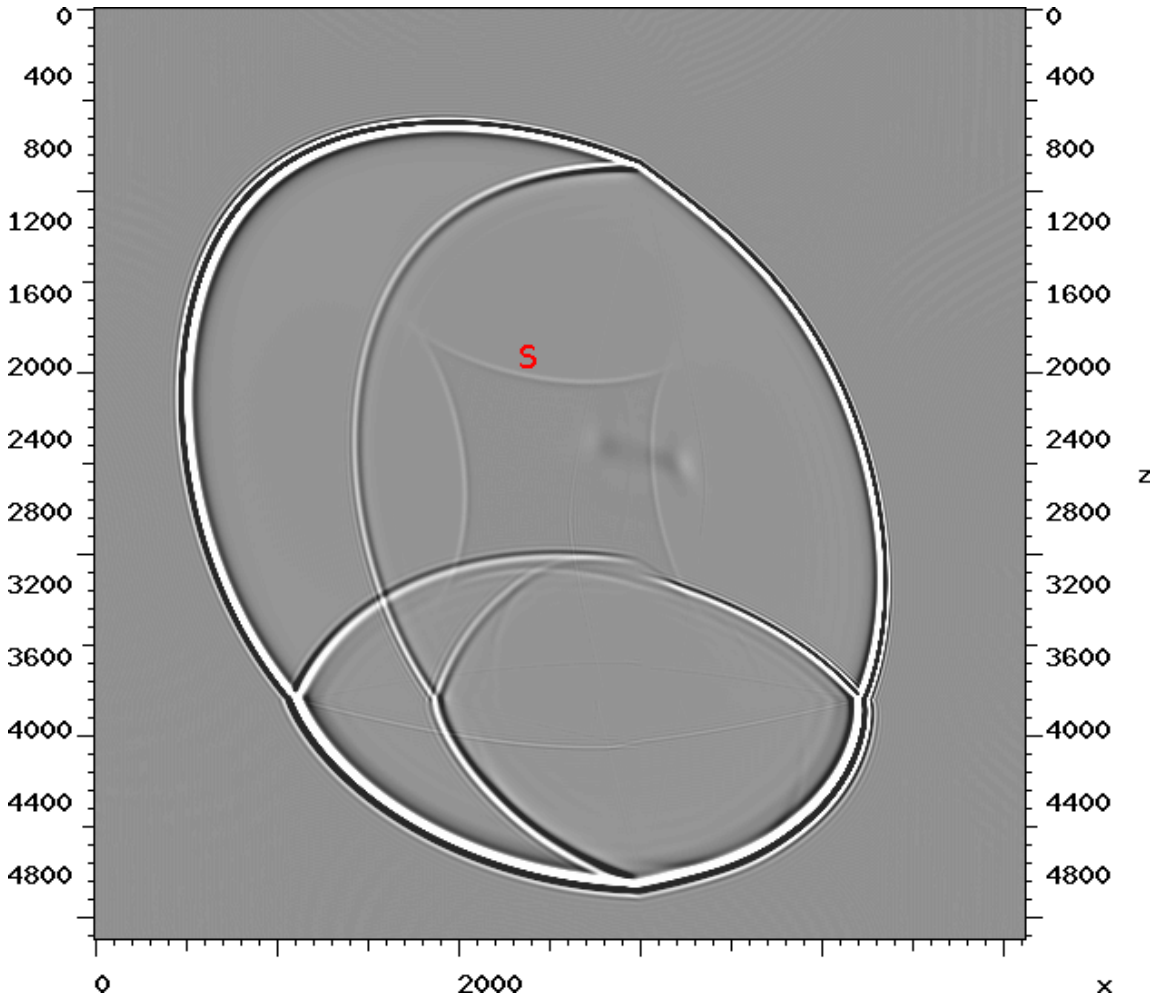


Figure 4: Shear artifact (marked with the "S") in the solution of (7) with sources injected in component q. Although significantly reduced, the shear artifact is still sufficiently strong to cause imaging cross-talk. [**CR**] musa2/. q1fdpdoqsource

THE FINITE-DIFFERENCE METHOD

In step 4) we square the pseudo-differential operator equation (2) so as to get rid of the square root, and obtain the following system of coupled second-order partial differential equations (Alkhalifah, 2000):

$$\begin{aligned} \frac{\partial^2 q}{\partial t^2} &= V_{P\mathrm{Hor}}^2 \frac{\partial^2 q}{\partial x'^2} + V_P^2 \frac{\partial^2 q}{\partial z'^2} - V_P^2 \left(V_{P\mathrm{Hor}}^2 - V_{P\mathrm{NMO}}^2\right) \frac{\partial^4 r}{\partial x'^2 \partial z'^2}, \\ \frac{\partial^2 r}{\partial t^2} &= q, \end{aligned} \tag{7}$$

where $r(z,x,t)$ and $q(z,x,t)$ are the pressure field and its second temporal derivative, and

$$V_{P\mathrm{Hor}}(z,x) = V_P(z,x)\sqrt{1+2\epsilon(z,x)},\ V_{P\mathrm{NMO}}(z,x) = V_P(z,x)\sqrt{1+2\delta(z,x)}.$$

Since the resulting system now includes the branch with the negative square root in (1), solution of this system may suffer from shear artifacts as shown in Figure 3. The artifacts can be reduced by injecting sources into the second component q (Fowler et al., 2010); however, they are still present—see Figure 4. However, the pseudo-differential operator equation (2) can be used to reduce the unwanted artifacts (appearing as the "diamond"-shaped inverted wavefront in the figure). Equation (1) and the corresponding pseudo-differential equation do not describe any pressure to shear conversion but rather govern the independent propagation of the pressure and shear waves. The same is true of the "coupled" system of differential equations. Consequently, any shear artifacts that appear in a solution to the coupled system of differential equations we attribute to the pseudo-shear modes present in the wave field. We can use the fact that the system of two coupled equations requires injecting two sources, to manufacture a pseudo-source to be injected into one of the components so as to suppress the shear modes. More specifically, if $\phi(z,x,t)$ is a time-dependent source function, then at each time step component r is injected with ϕ, and component q is injected with the result of applying the spatial part of the pseudo-differential operator (2) to $\phi(z,x,t)$:

$$\begin{aligned} r(z,x,t_n) &= r(z,x,t_n) + \phi(z,x,t_n), \\ q(z,x,t_n) &= q(z,x,t_n) + V_P^2 \left\{ \left(\overset{2}{z},\overset{2}{x}\right) \frac{\Delta}{2} + \epsilon \left(\overset{2}{z},\overset{2}{x}\right) \frac{\partial^2}{\partial x'^2} + \right. \\ &\left. + \frac{\Delta}{2} \sqrt{\left[1 + 2\epsilon \left(\overset{2}{z},\overset{2}{z}\right) \frac{\partial^2}{\partial x'^2} \frac{1}{\Delta}\right]^2 - 8 \left(\epsilon \left(\overset{2}{z},\overset{2}{x}\right) - \delta \left(\overset{2}{z},\overset{2}{x}\right)\right) \frac{\partial^2}{\partial x'^2} \frac{\partial^2}{\partial z'^2} \frac{1}{\Delta^2}} \right\} \phi, \end{aligned} \tag{8}$$

followed by a finite-difference time propagation step of system (7). This procedure ensures that the two-component source in the right-hand side of (8) satisfies equation (2). Since solutions of (2) are shear-free, the injected sources will not give rise to shear modes because the solution of (7) is effectively projected on to the space of solutions of (2).

NUMERICAL EXAMPLES

Figure 6 shows the result of applying the pseudo-source finite-difference method to the propagation in a heterogeneous VTI medium described by the model of Figure 1, with a Ricker source. The corresponding result obtained by solving the full pseudo-differential operator equation (2) is shown in Figure 5. Note the significant reduction of the shear artifacts and

that although we use the full pseudo-differential operator for generating the pseudo-source in (8), the fact that the source is localized makes this computationally efficient, obviating the need for approximations like (4).

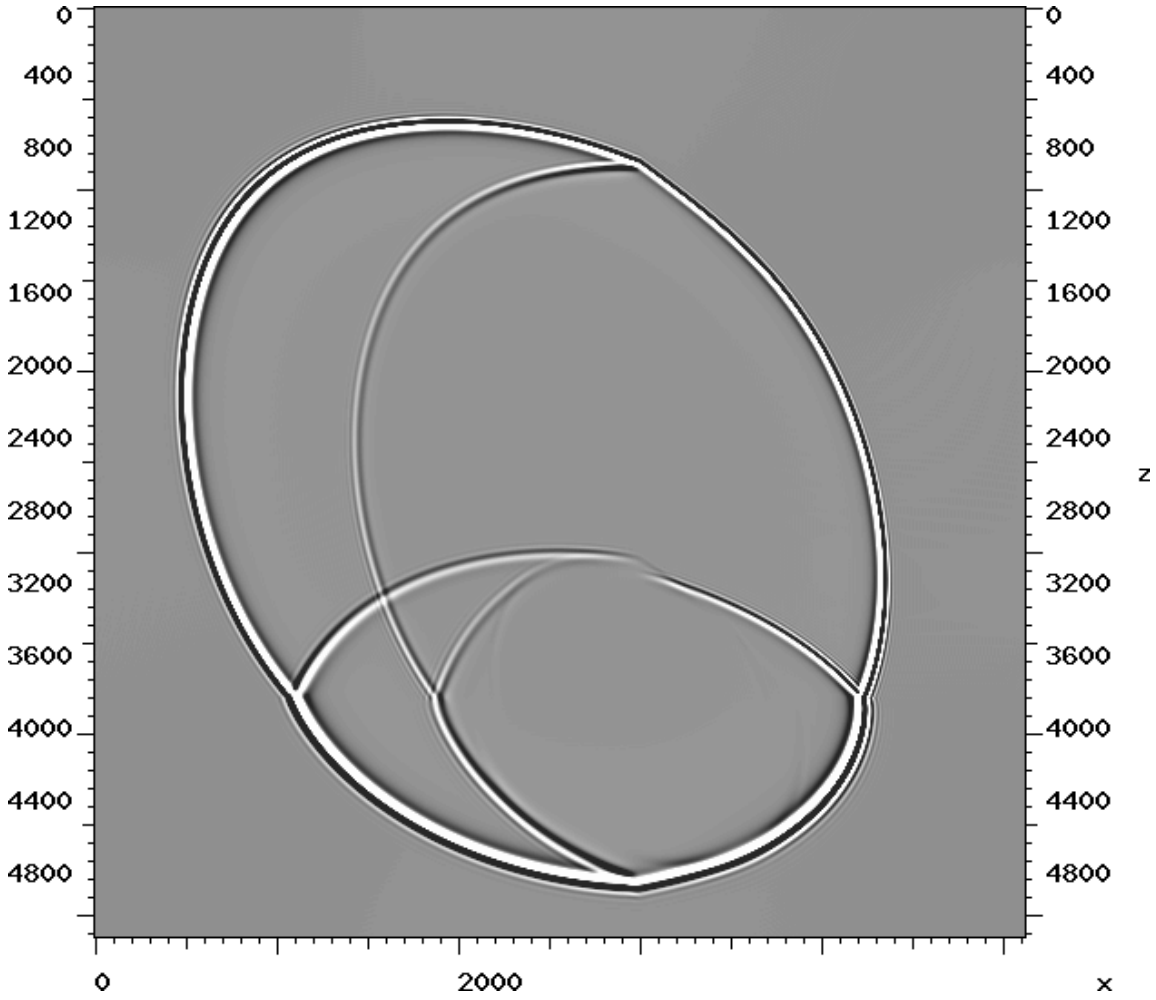


Figure 5: Solution of the full pseudo-differential operator equation (2) for the model of Figure 1. Note the good agreement with the result of finite-difference modelling using shear-reducing pseudo-sources shown in Figure 6. [**CR**] musa2/. p1pdo

The model of Figure 1, while featuring both sharp and smooth vertical velocity variation, assumes constant $\epsilon = 0.3$, $\delta = 0.1$, and tilt of 45°. While adding the pseudo-source (8) ensures that the solution of the coupled system (7) stays within the space of solutions of (2) in the continuous limit $\Delta t \to 0$, sharp contrasts in medium parameters may introduce numerical approximation errors that may contain a non-negligible shear component. Indeed, applying the method to the model of Figure 2, featuring two inclusions with significantly different Thomsen parameters, we can see weak artifacts (single lines) within the inclusion detail in Figure 10 for the finite-difference method that are absent from the result in Figure 9 obtained by solving the full pseudo-differential operator (2). Figure 11 shows the result of using the finite-difference method with pseudo-sources after smoothing the vertical velocity but keeping ϵ, δ and tilt contrasts unchanged. The result shows that the artifacts within the inclusions were almost completely removed.

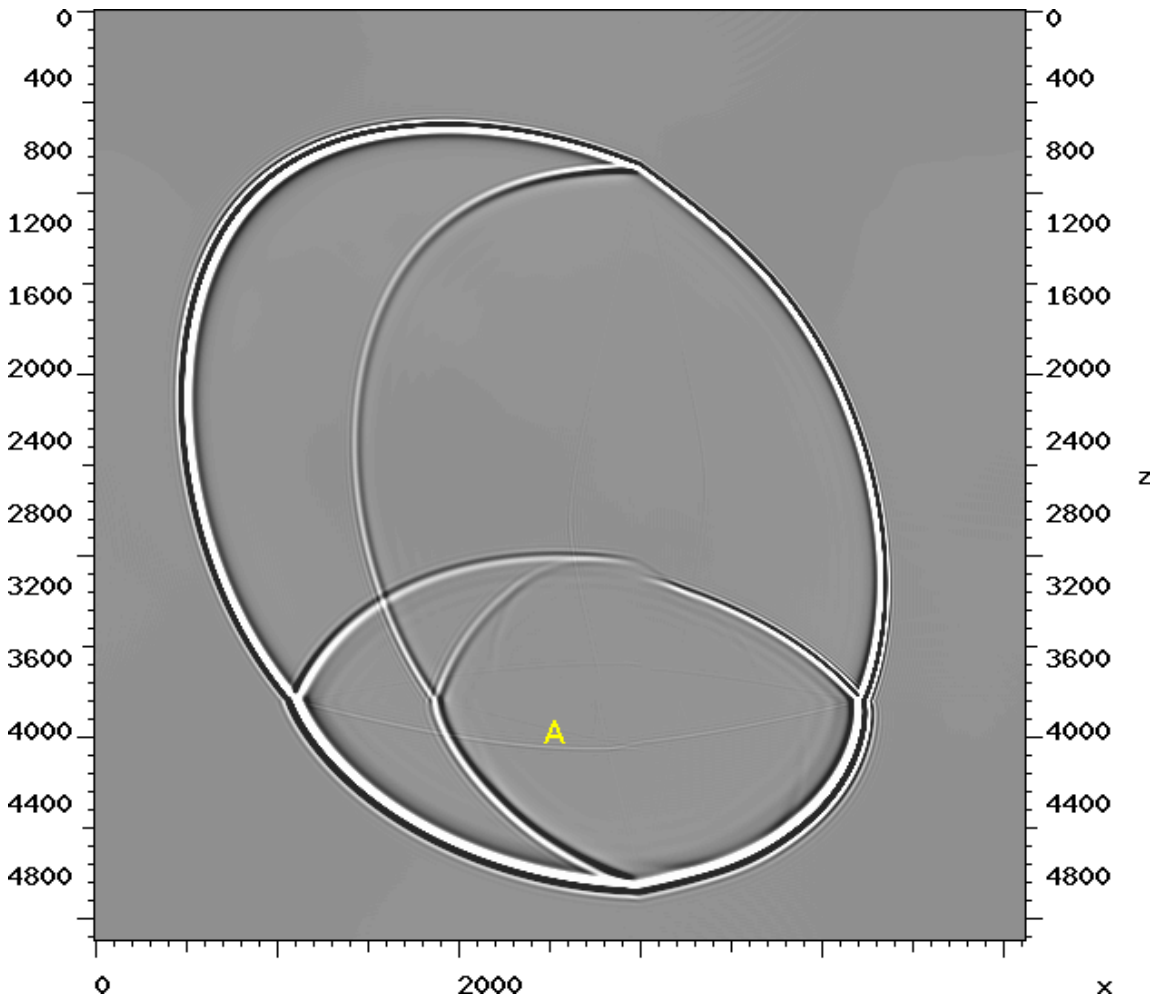


Figure 6: Solution of (7) for the model of Figure 1 with shear-reducing pseudo-sources is in kinematic agreement with Figure 5. A high-wavenumber computational artifact (marked with the "A") is caused by sharp model contrasts. [**CR**] musa2/. p1fdpdo

CONCLUSIONS AND PERSPECTIVES

The proposed pseudo-source finite-difference method allows us to take advantage of computationally efficient finite-difference solvers for the traditional pseudo-acoustic (fourth-order) systems while achieving a significant reduction in shear artifacts. The method is kinematically accurate for VTI media, and can be extended in principle to other kinds of anisotropy. While this implementation is based on using the coupled system (7) of Alkhalifah (2000), the method can be adapted to use equivalent systems such as that of Fowler et al. (2010). In that case the two-component source becomes a linear combination of the true source and the pseudo-source terms, with the coefficients of the linear combination determined by the relationship between the solution of the equivalent system and that of system (7).

ACKNOWLEDGEMENTS

We would like to thank Biondo Biondi for a number of useful discussions.

REFERENCES

Alkhalifah, T., 1998, Acoustic approximations for processing in transversely isotropic media: Geophysics, **63**, 623–631.

———, 2000, An acoustic wave equation for anisotropic media: Geophysics, **65**, 1239–1250.

Etgen, J. and S. Brandsberg-Dahl, 2009, The pseudo-analytical method: Application of Pseudo-Laplacians to acoustic and acoustic anisotropic wave propagation: 79th SEG Annual International Meeting, Expanded Abstracts, 2552–2556.

Fowler, P. J., X. Du, and R. P. Fletcher, 2010, Coupled equations for reverse time migration in transversely isotropic media: Geophysics, **75**, no. 1, S11–S22.

Grechka, V., 2009, Applications of seismic anisotropy in oil and gas industry: EAGE.

Harlan, W., 1998, A convenient approximation of transverse isotropy for higher-order moveout, prestack time migration, and depth calibration. (http://www.billharlan.com/papers/aniso/aniso.html).

Iserles, A., 2008, A first course in the numerical analysis of differential equations: Cambridge University Press.

Maharramov, M., 2014, Artifact reduction in pseudo-acoustic modeling by pseudo-source injection: SEP Report, **152**, 95–104.

———, 2015, Efficient finite-difference modelling of acoustic wave propagation in anisotropic media with pseudo-sources: 77th EAGE Conference and Exhibition, Extended Abstracts (accepted).

Maharramov, M. and B. Nolte, 2011, Efficient one-way wave-equation migration in tilted transversally isotropic media: 73rd EAGE Conference and Exhibition, Extended Abstract F018.

Maslov, V., 1979, Operational methods: Nauka.

Musgrave, M., 1970, Introduction to the study of elastic waves and vibrations in crystals: Holden-Day.

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

Tsvankin, I., 1996, P-wave signatures and notation for transversely isotropic media: Geophysics, **61**, 467–483.

Zhan, G., R. C. Pestana, and P. L. Stoffa, 2012, Decoupled equations for reverse time migration in tilted transversely isotropic media: Geophysics, **77**, no. 2, T37–T45.

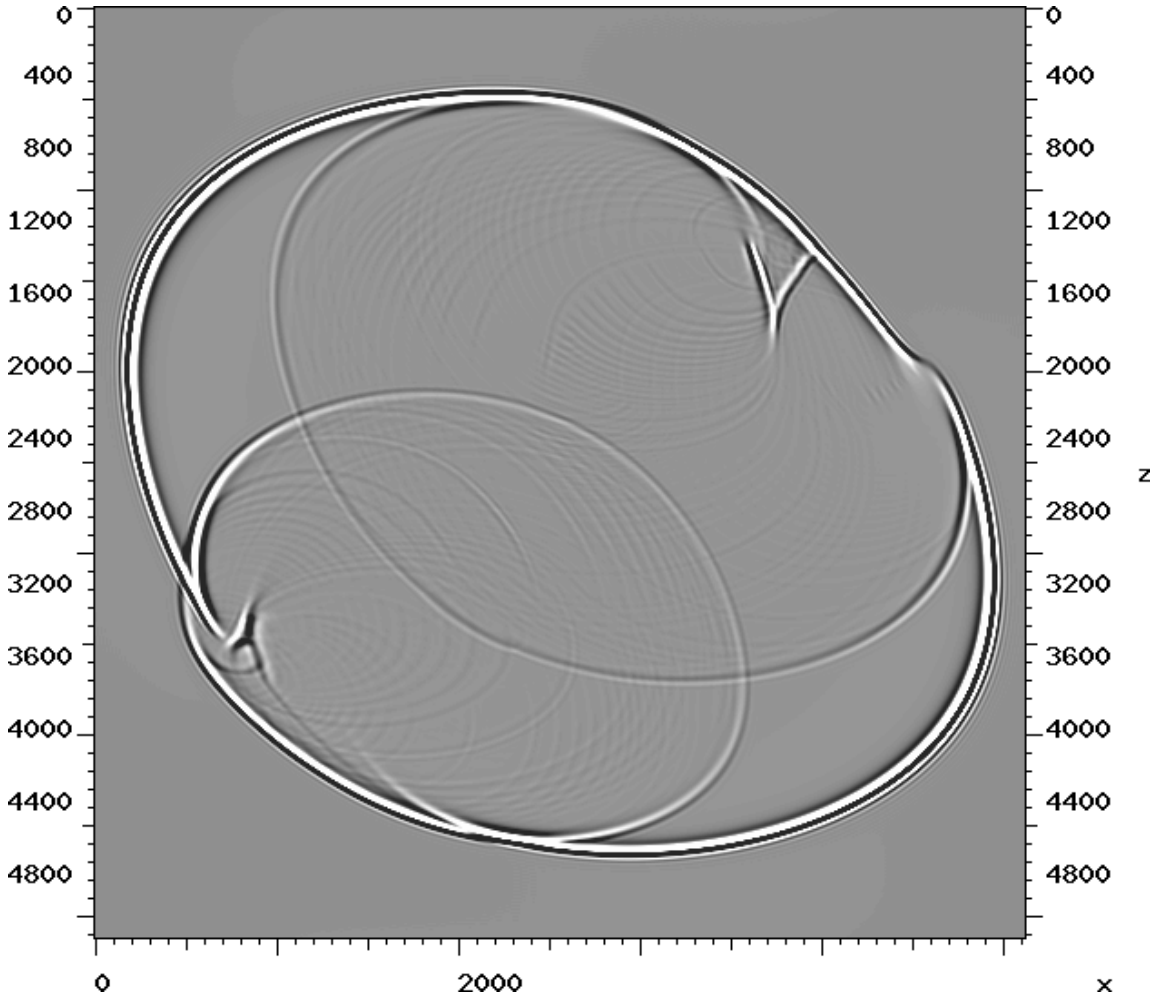


Figure 7: Solution of the full pseudo-differential operator equation (2) for the model of Figure 2. [**CR**] musa2/. p2pdo

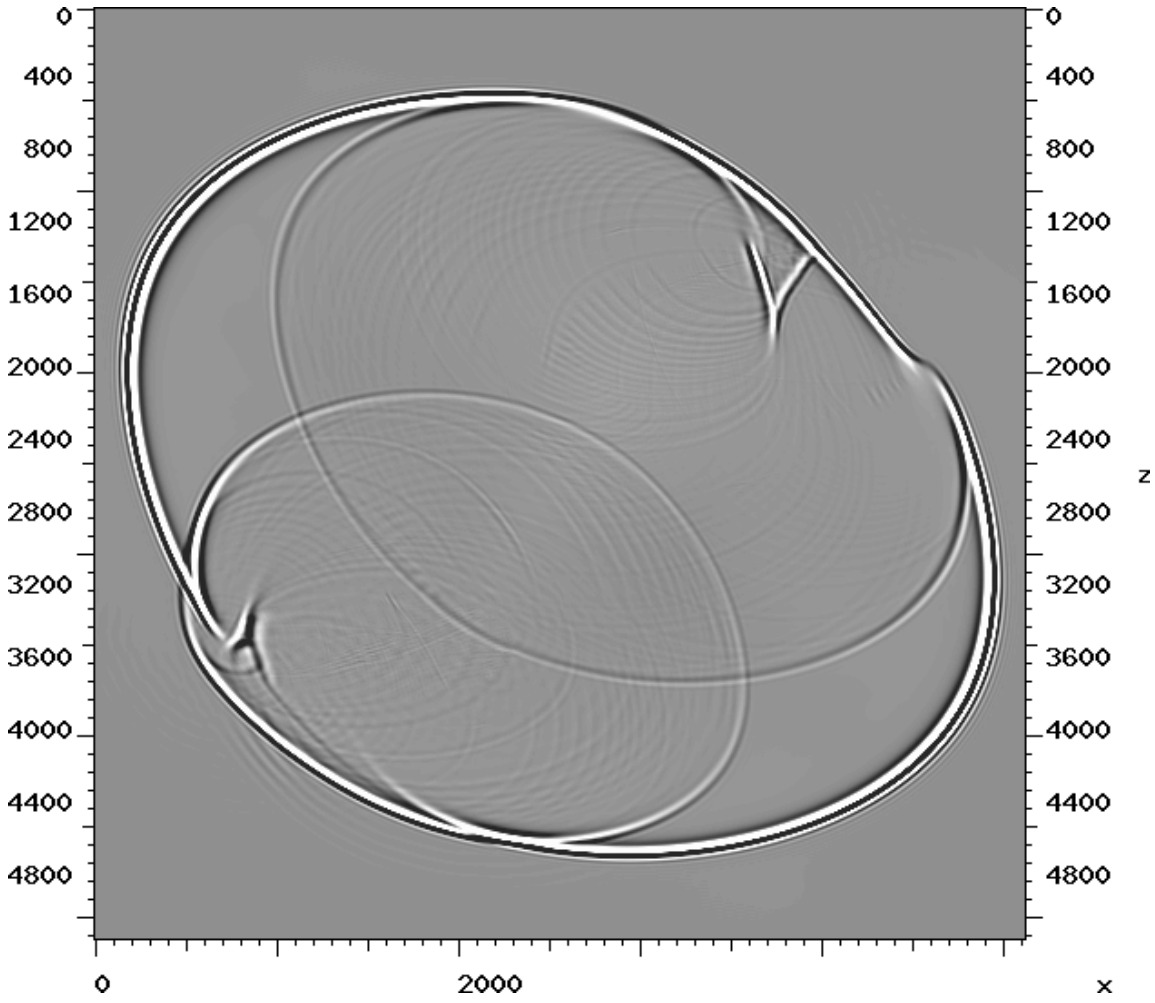


Figure 8: Solution of (7) for the model of Figure 2 with shear-reducing pseudo-sources. Note the good agreement with Figure 7. [**CR**] musa2/. p2fdpdo

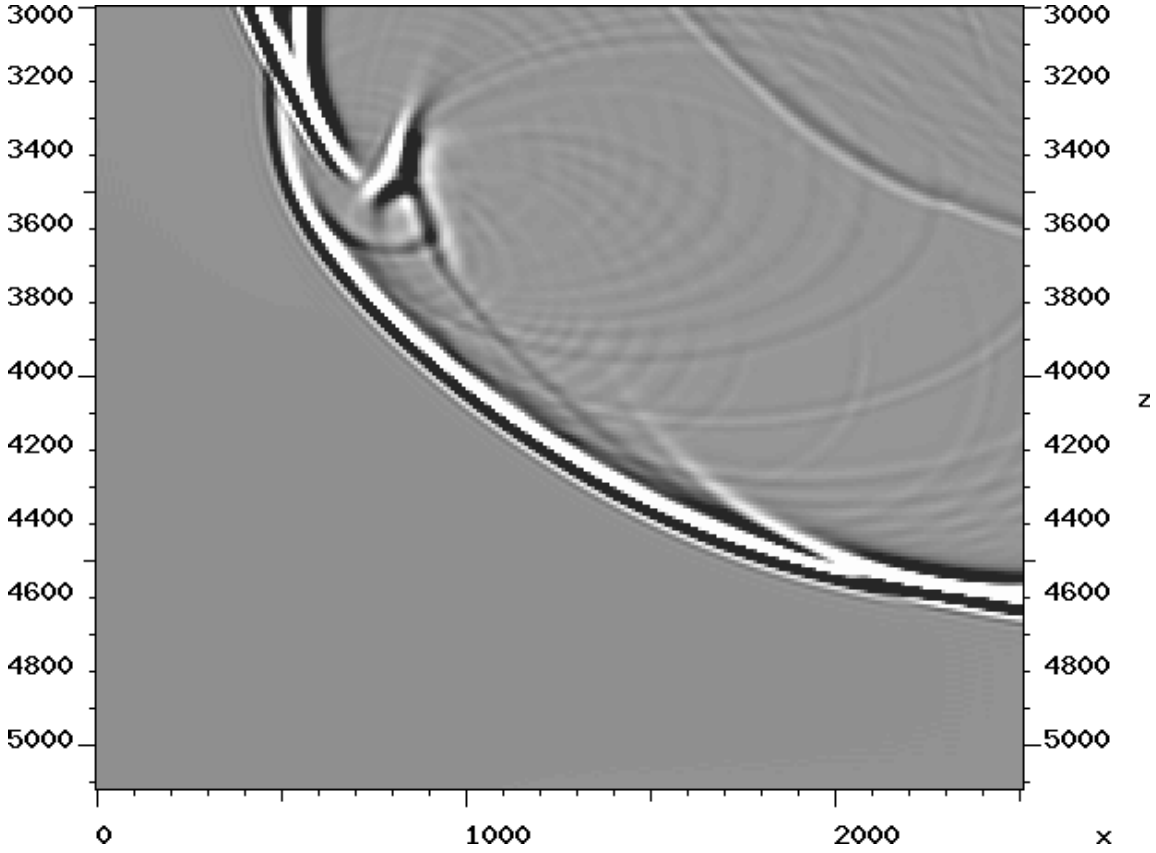


Figure 9: Artifact-free solution of the full pseudo-differential operator equation (2) showing multiple reflections within the lower inclusion of the model in Figure 2. [**CR**] musa2/. p2pdodetail

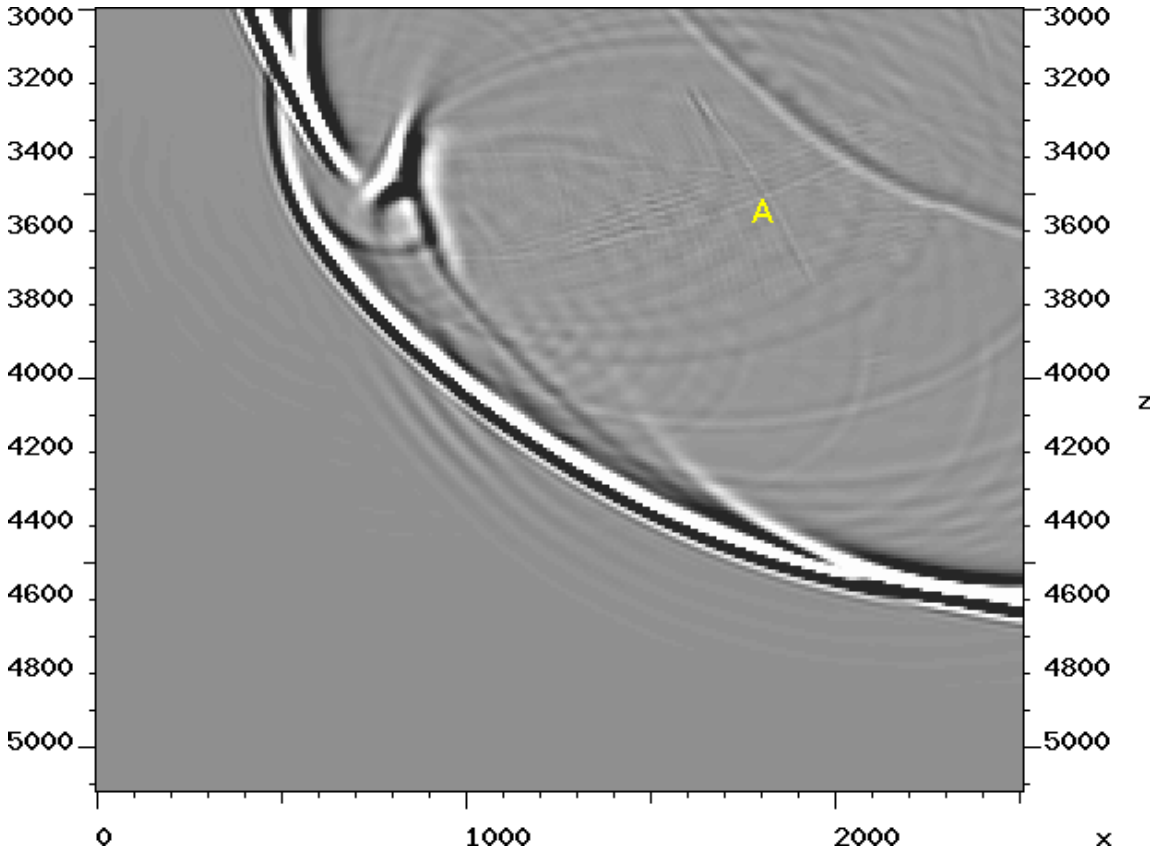


Figure 10: Solution of (7) for the model of Figure 2 with shear-reducing pseudo-sources. A sharp velocity contrast causes weak artifacts ("A"). [**CR**] musa2/. p2fdpdodetail

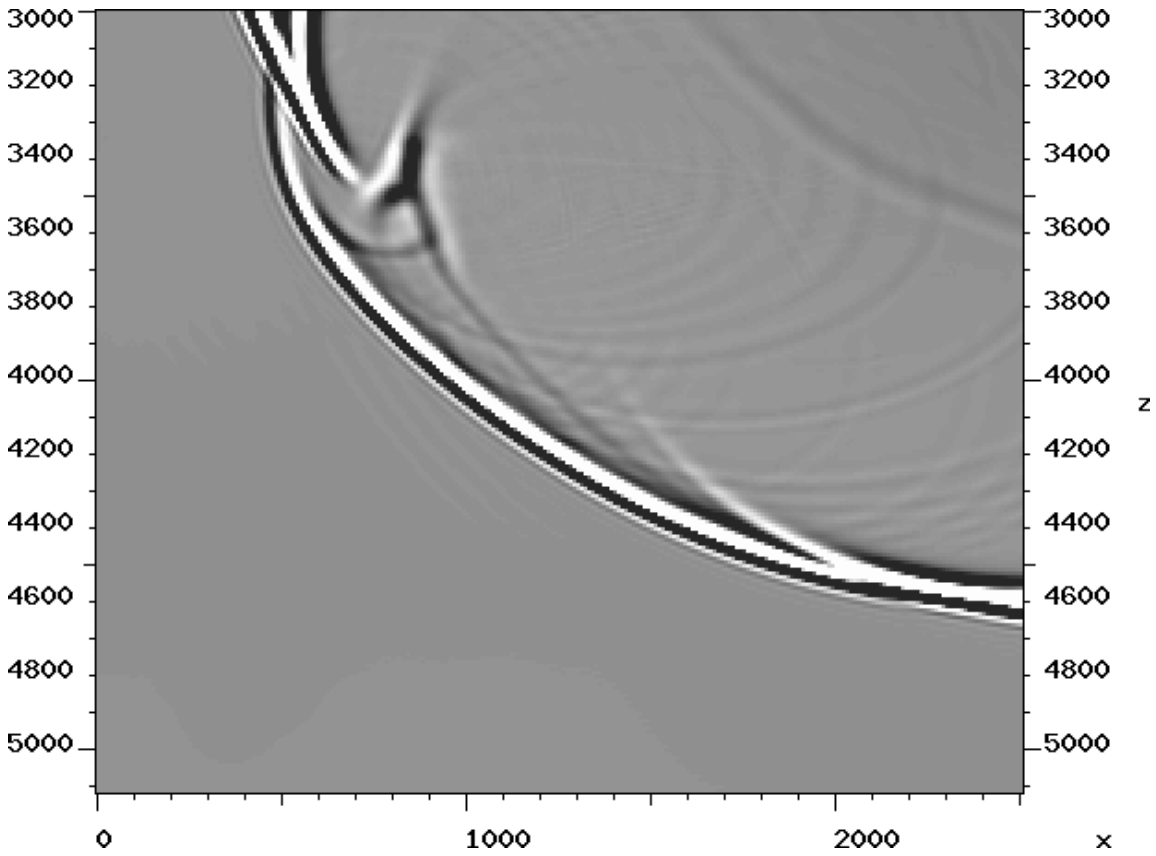


Figure 11: Moderate smoothing of the velocity contrasts remove the high-wavenumber artifacts. [**CR**] musa2/. p2fdpdosmdetail

Using rock physics to improve Q_p quantification in seismic data

Yi Shen and Jack Dvorkin

ABSTRACT

We derived an approximate closed-form solution relating V_p to Q_p using rock physics modeling. This solution is validated using well data in which the elastic properties were measured and Q was derived numerically using rock physics. We applied our new $Q_p - V_p$ equation to synthetic seismic data, which produced an improved Q model.

INTRODUCTION

Gas pockets or clouds are notoriously challenging problems for reservoir identification and interpretation (Billette and Brandsberg-Dahl, 2005), because strong attenuation and low-velocity anomalies are present in them. Attenuation degrades the seismic image quality by decaying the image amplitude, lowering the image resolution, distorting the phase of events, and dispersing the velocity. Low-velocity mispositions and distorts the events. These problems impede accurate image interpretation for hydrocarbon production and well positioning.

To mitigate the effects of gas accumulations on the image, we need a better understanding of the properties of the gas used for imaging the subsurface. Compressional velocity (V_p) and compressional Q (Q_p) play an important role in compensating for the gas-induced distortion in the image. However, an accurate estimation of these two properties is limited to the insufficient information from the acquired seismic data and to the ambiguity in the distorted effects of the images from these two parameters. A quantification of the relation between V_p and Q_p brings additional information to the seismic inversion for a better estimation of these two properties.

It has been observed from field data (He and Cai, 2012; Zhou et al., 2014) that high-compressional attenuation comes with low-compressional velocity in most situations. However, very few studies have analytically linked these two properties. Rock physics has built several models (Dvorkin et al., 2014; Muller et al., 2010) of V_p and Q_p based on rock properties (such as porosity, saturation and others), which may implicitly quantify the relation between these two parameters. But, these existing models for V_p and Q_p, based on rock physics properties, confine the measurements of these two parameters to a constrained area limited by the well locations. A model that provides an analytical relation between V_p and Q_p would allow us to approximately relate Q_p to V_p without going through direct rock physics modeling, which in turn would improve the accuracy of the seismic inversion of V_p and Q_p.

In this study, we first derived an approximate closed-form solution directly relating V_p to Q_p using rock physics modeling. Next, we validated this relation using field well data. Last, we applied our new $Q_p - V_p$ equation to synthetic seismic data, which produced an improved Q model.

THEORY

Modeling seismic attenuation

Seismic attenuation primarily occurs either at a gas reservoir or in the presence of shallow gas pockets. Wave-induced variations of pore pressure in the partially saturated rock results in oscillatory liquid flow. The viscous losses during this oscillatory liquid flow cause wave attenuation. The frequencies of the wave span broad frequencies and scales, which can be categorized as macroscopic, mesoscopic, and microscopic based on the spatial scale of the heterogenities. Mesoscopic flow, rather than macroscopic and microscopic flow, is engaged at the seismic exploration frequency range, and is considered to be the main mechanism for the fluid-related seismic attenuation.

Seismic attenuation, parameterized by quality factor Q, is a function of frequency. According to the standard linear solid model, we are able to obtain the relation between $1/Q$ and frequency, and therefore derive the maximum $1/Q$:

$$Q_{\max}^{-1} = \frac{M_\infty - M_0}{2\sqrt{M_\infty M_0}}, \tag{1}$$

where $Q_{\max}^{-1}$ is the maximum inverse quality factor, M_0 is the compressional modulus at very high frequency, and M_∞ is the compressional modulus at very low frequency. The compressional modulus is the product of the bulk density and P-wave velocity squared. This equation provides the upper bound for attenuation without addressing its frequency dependence. Therefore, we use $Q_{\max}^{-1}$ to describe the effects of seismic attenuation.

According to Dvorkin et al. (2014), in partially saturated rock, viscoelastic effects and attenuation may arise from the oscillatory liquid cross-flow between fully liquid-saturated patches and the surrounding rock with partial gas saturation. The reaction of rock with patchy saturation to loading by the elastic wave depends on the frequency. If it is low and the loading is slow, the oscillations of the pore pressure in a fully liquid-saturated patch and partially saturated domains next to it are equilibrating. The patch is "relaxed." Mavko et al. (1991) derived an approximation to Gassmann fluid substitution equation (Gassmann, 1951) for the compressional modulus of the partially saturated rock as follows:

$$M_0 = M_S \frac{\phi M_{Dry} - (1+\phi)K_F M_{Dry}/M_S + K_F}{(1-\phi)K_F + \phi M_S - K_F M_{Dry}/M_S}, \tag{2}$$

where M_S is the compressional modulus of the mineral phase, M_{Dry} is the compressional modulus of the dry frame of the rock, ϕ is the total porosity, and K_F is the effective bulk modulus of the fluid in the rock.

Conversely, if the frequency is high and the loading is fast, the resulting oscillatory variations of pore pressure cannot equilibrate between the fully saturated patch and the domain outside. The patch is "unrelaxed." For a high frequency, we can use the patchy saturation equation (Mavko et al., 1991), which expresses the unrelaxed compressional modulus as the harmonic average of the compressional moduli of the wet rock M_W and rock with only gas M_G:

$$\frac{1}{M_\infty} = \frac{S_W}{M_W} + \frac{1-S_W}{M_G}, \tag{3}$$

where

$$M_W = M_S \frac{\phi M_{Dry} - (1+\phi) K_W M_{Dry}/M_S + K_W}{(1-\phi) K_W + \phi M_S - K_W M_{Dry}/M_S}, \tag{4}$$

$$M_G = M_S \frac{\phi M_{Dry} - (1+\phi) K_G M_{Dry}/M_S + K_G}{(1-\phi) K_G + \phi M_S - K_G M_{Dry}/M_S}, \tag{5}$$

K_W and K_G are the bulk modulus of water and gas, respectively, and S_W is water saturation.

Modeling velocity

The compressional velocity of rock is related to the compressional modulus M_0 and its bulk density ρ as

$$V_p = \sqrt{\frac{M_0}{\rho}}. \tag{6}$$

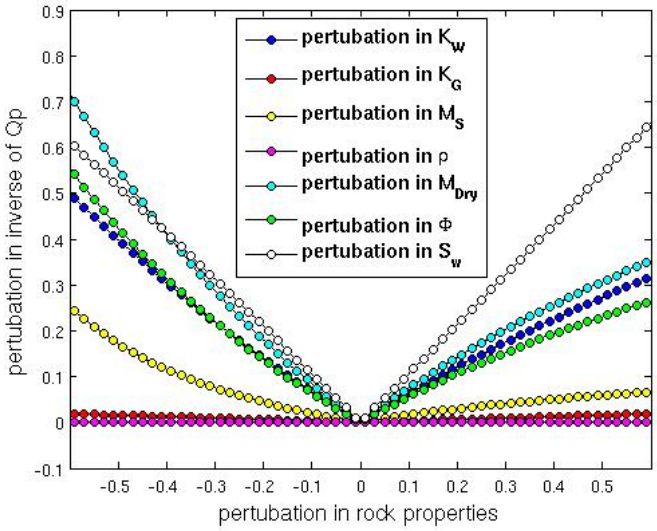
Linking seismic attenuation and velocity: closed-form solution

V_p and Q_p are functions of a number of rock properties: M_{Dry}, M_S, S_W, ϕ, K_W, K_G and ρ. Model building of V_p and Q_p based on these rock parameters is the intermediate step to link V_p with Q_p in our study. Because a change in these rock properties results in changes in V_p and Q_p to different degrees, we only linked V_p with Q_p using the rock properties which, when modified, generate a significant change in these two properties and therefore in their relation. We assigned a spatially constant value to the rest of the rock parameters. A prior knowledge of the lithology of the areas of our study and measurements from well data enabled us to obtain an average value to approximately quantify properties of little influence on V_p and Q_p. Figure 1 is the sensitivity curve showing the influence of a change in these rock properties on a change in V_p and Q_p. The variation of the rock properties are shown in Table 1. The results show that a perturbation in M_{Dry} leads to the most significant change in both V_p and Q_p among these rock parameters. Therefore, we linked V_p with Q_p using M_{Dry} in our study.

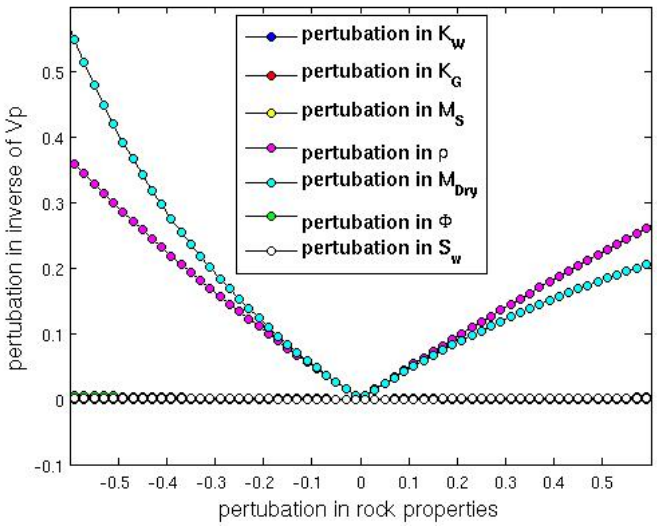
Table 1: The variation of the rock properties for Figure 1

Rock properties	Minimum value	Maximum value	background value
K_W	0.9521	3.9305	2.4413
K_G	0.0088	0.0364	0.0226
M_S	37.6740	155.5260	96.6000
ρ	0.8853	3.6547	2.2700
M_{Dry}	3.9000	16.1000	10.0000
ϕ	0.1170	0.4830	0.3000
S_W	0.1170	0.4830	0.3000

The rock property M_{Dry} is not a direct measurable parameter from the well log. The previous studies (Raymer et al., 1980; Dvorkin and Nur, 1996) built the rock physics model to compute M_{Dry} using other parameters. However, these models have not provided an



(a)



(b)

Figure 1: The sensitivity of (a) Q_p and (b) V_p to a change in the rock properties of M_{Dry}, M_S, S_W, ϕ, K_W, K_G, ρ. The background rock parameters are $K_W = 2.4413; K_G = 0.0226; M_S = 96.6; \rho = 2.27; M_{Dry} = 10; \phi = 0.3; S_W = 0.3$. [**NR**] yishen1/. NR/P.sen.Qp,NR/P.sen.Vp

analytical equation conveniently used to derive the relation between V_p and Q_p. Based on the idea that as the rock becomes softer, its moduli decreases as pore space increases, and the porosity-induced change in rock moduli should be proportional to ϕ, we propose a new model in this study for M_{Dry} with a simple equation:

$$\frac{M_{Dry}}{M_S - M_{Dry}} = \frac{\alpha_{Dry}}{\phi}, \tag{7}$$

where α_{Dry} is a positive number. The large α_{Dry} slowly increases the porosity-induced change of the rock moduli as the porosity increases. As a result, the dry rock moduli with a larger α_{Dry} is larger than the one with a smaller α_{Dry} for a fixed ϕ, as shown in Figure 2(a).

Equation 7 shows that M_{Dry} depends on M_S and ϕ. By substituting Equation 7 into Equations 1 and 6 to eliminate M_{Dry}, we obtained new sensitivity curves in Figure 3 showing the amount of changes of V_p and Q_p caused by a perturbation in the rock properties M_S, S_W, ϕ, K_W, K_G and ρ. The results show that a perturbation in M_S and ϕ results in the most significant change in both V_p and Q_p among these rock parameters. Because M_S can be estimated if the mineral content of the reservoir rock is known, we mainly studied the effects of the variation of ϕ on V_p and Q_p.

The porosity ϕ is a measure of the fraction of open space in the rock. Pore space softens the rock and makes its moduli decrease, which makes the compressional velocity decrease as shown in the lower panel of Figure 2(b). Also, the oscillatory liquid cross-flow between the gas and fluid in the pore space causes attenuation, as shown in the top panel of Figure 2(b). We linked V_p and Q_p by way of ϕ. Substituting ϕ with V_p, and assuming $M_\infty \approx M_0$ in the denominator of the right side of Equation 1, we were able to have the relation between V_p and Q_p

$$Q_p^{-1} = \frac{1}{2}\frac{c_1 V_p^{-2}}{c_2 V_p^{-2} + c_3} - \frac{1}{2}, \tag{8}$$

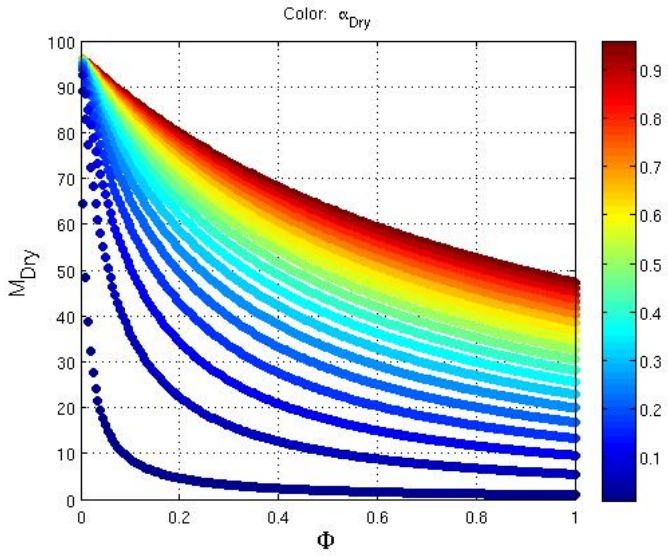
where

$$\begin{aligned} c_1 &= (\alpha_{Dry} + \alpha_W)(\alpha_{Dry} + \alpha_G), \\ c_2 &= S_W(\alpha_{Dry} + \alpha_F)(\alpha_{Dry} + \alpha_G) + (1 - S_W)(\alpha_{Dry} + \alpha_F)(\alpha_{Dry} + \alpha_W), \\ c_3 &= \frac{\rho}{M_S}\left(S_W(\alpha_{Dry} + \alpha_F)(\alpha_W - \alpha_G) + (\alpha_{Dry} + \alpha_W)(\alpha_G - \alpha_F)\right), \end{aligned} \tag{9}$$

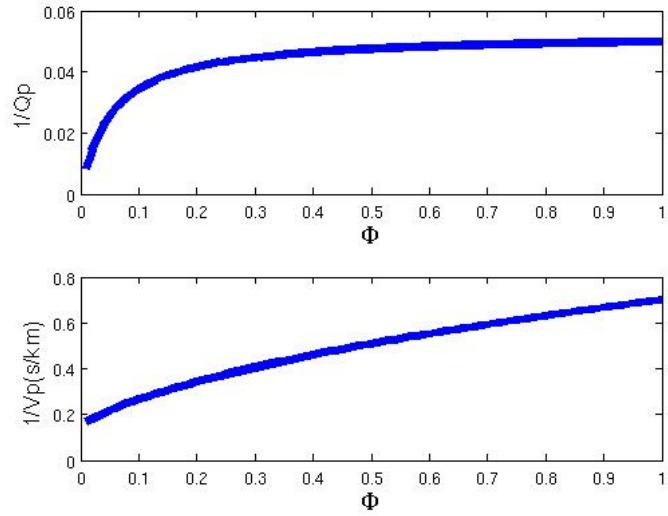
and

$$\begin{aligned} \alpha_F &= \frac{K_F}{(M_S - K_F)}, \\ \alpha_G &= \frac{K_G}{(M_S - K_G)}, \\ \alpha_W &= \frac{K_W}{(M_S - K_W)}. \end{aligned} \tag{10}$$

Figure 4 shows the relations between $1/V_p$ and $1/Q_p$. The solid circles are the exact relation using Equations 1 and 6, and the nonfilled circles are calculated by the approximated Equation 8. These two curves have no large bias overall, and match very well for

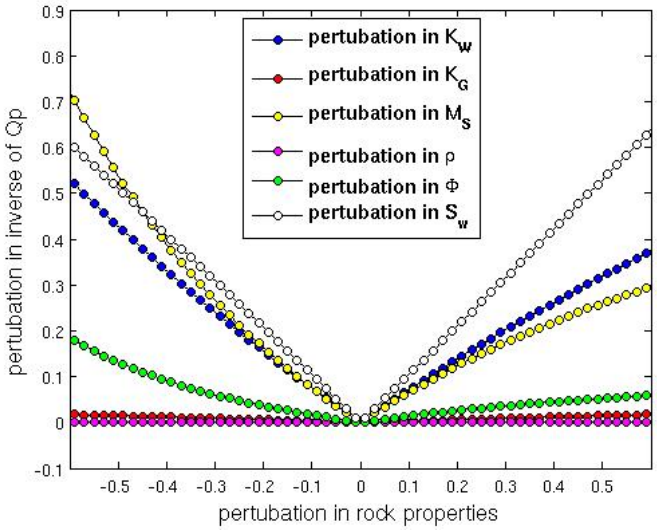


(a)

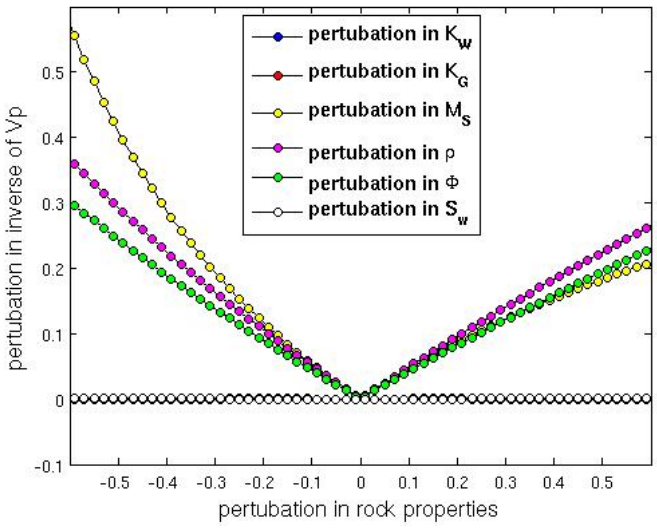


(b)

Figure 2: (a) Our new model for rock properties M_{Dry} as a function of ϕ and α_{Dry}, with $M_S = 96.6$ (b) The relations of ϕ with Q_p (top) and V_p (bottom), with $K_W = 2.4413; K_G = 0.0226; M_S = 96.6; \rho = 2.27; \alpha_{Dry} = 0.05; S_W = 0.3$. [**NR**] yishen1/. NR/P.Mdry,NR/P.Phi



(a)



(b)

Figure 3: The amount of changes of (a)Q_p and (b)V_p caused by a perturbation in the rock properties M_S, S_W, ϕ, K_W, K_G, ρ. The unchanged rock parameters are $K_W = 2.4413; K_G = 0.0226; M_S = 96.6; \rho = 2.27; \alpha_{Dry} = 0.05; \phi = 0.3; S_W = 0.3$. [**NR**] yishen1/. NR/P.sen.Qp.noMdry,NR/P.sen.Vp.noMdry

small porosity. From these results, we observe that the decrease of the compressional velocity corresponds to strong attenuation, in accordance with the observation from field data distorted by gas anomalies.

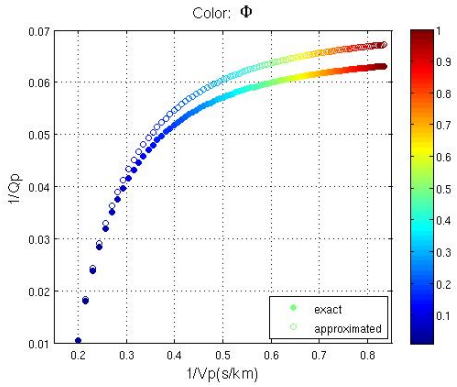


Figure 4: The relations between $1/V_p$ and $1/Q_p$, with $K_W = 2.4413; K_G = 0.0226, M_s = 96.6; \rho = 2.27$. [**NR**] yishen1/. NR/PQV

VALIDATING THE CLOSED-FORM SOLUTION USING WELL DATA

We use a field gas log to validate our theory. The black curve shown in Figures 5(a) through 5(g) are the rock properties measured from the log. We calculated the Q value using Equation 1 in Figure 5(h). The log has a low-water saturation at depths from 1.25 to 1.27 kilometers (km) and 1.31 to 1.32 km, indicating gas sand. We observe that both V_p and Q_p are small in the gas sand.

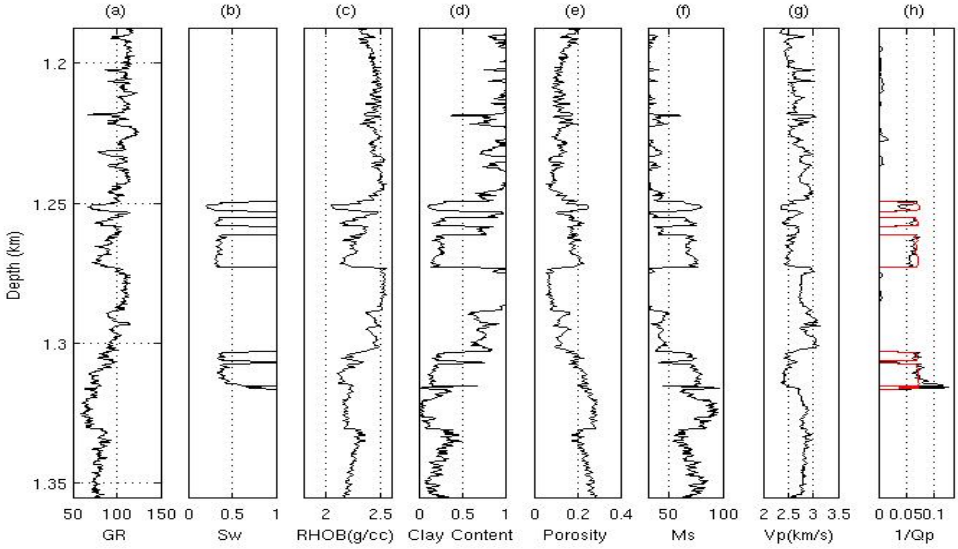


Figure 5: Well data used to verify our approximate relations. [**NR**] yishen1/. NR/Plog

We first computed the dry-rock compressional moduli M_{1Dry} from the provided rock

properties using the Gassmann-Mavko equation (Mavko et al., 1991). We then superimposed M_{Dry} using our new model on M_{1Dry} in Figure 6(a), finding a good match of our model with the gas sand shown in blue at $\alpha_{Dry} = 0.05$. Then, we computed the approximated Q_p for the gas sand using Equation 1, shown as the red curve in Figure 5(h), with the assumption that α_{Dry} is 0.05 for our M_{Dry} and that M_∞ equals to M_0 in the denominator of the right side of Equation 1.

Figure 6(b) shows the relation between V_p and Q_p. The blue dots are the log data in which the trend is approximated by red dots with a water saturation of $S_W = 0.37$. The approximated trend matches our predicted relation well by the black line using Equation 8. It slightly over-estimates direct computations because of our simplified new M_{Dry} model and our approximated equation (Equation 8).

SEISMIC APPLICATION

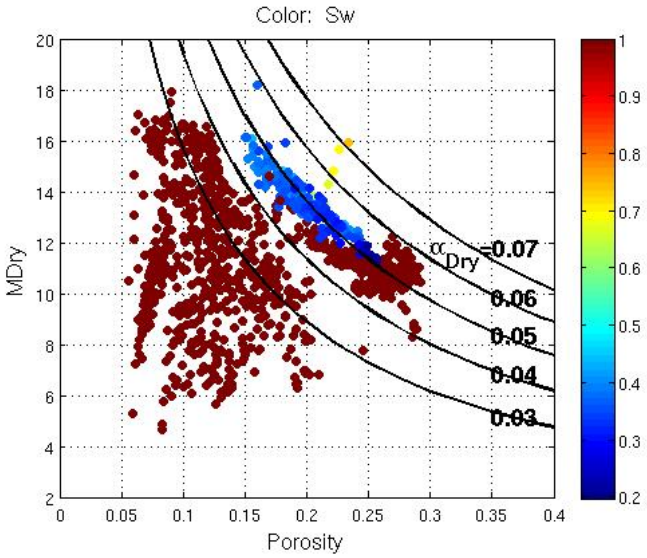
We forwarded model a synthetic seismic dataset using the V_p and Q_p models present in Li et al. (2015), as shown in Figure 7.Li et al. (2015) built the synthetic models based on the shallow unconsolidated sand reservoir model in Wang et al. (2013). A shallow gas pocket with 70% gas saturation is located in the upper part of the second layer, which has low velocity and high attenuation. Its rock properties porosity and gas saturation are shown in Figure 8. The shallow shale porosity in Figure 8 is arbitrarily assumed to be very small and unrelated to the velocity model. This assumption has no impact on our application in this study, because we do not use porosity in shale for our Q_p computations.

To synthesize the seismic data, we downward propagated and attenuated the wavefield (Shen et al., 2013, 2014) using the V_p and Q_p models in Figure 7 with 53 sources and 801 receivers uniformly distributed on the surface. A Ricker wavelet with 20 Hz central frequency was used as the source wavelet, and the density model was assumed to be spatial constant.

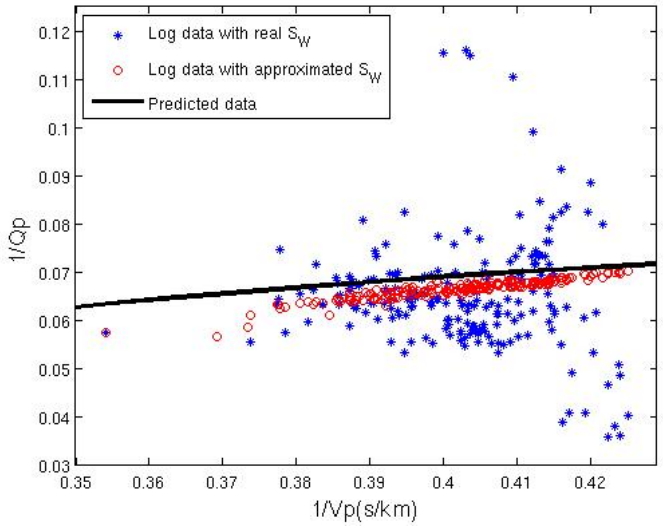
In practice, V_p and Q_p models are unknown in real fields. It is necessary to invert for these models that are important for generating a seismic migration image. The goal of our study is to invert for accurate Q_p models from the synthetic seismic data with the assumption that the correct V_p model is known. As a result, we were able to obtain a seismic migration image with a correction of its Q_p effects.

We inverted for the Q_p model shown in Figure 9(a) using wave-equation migration Q analysis (Shen et al., 2013, 2014), and used spatial constant $Q_p = 100,000$ as the initial model. The inverted Q_p model in Figure 9(a) highlights the area with high attenuation. However, the sparse reflectors and the limitation of this method result in a low resolution of the Q_p model, especially in the vertical direction. Therefore, we used the relation given by Equation 8 as a regularization term in the inversion workflow developed by Shen et al. (2013, 2014). In reference to the correct Q_p model in 7(b), the inverted Q_p model with regularization in Figure 9(b) has higher vertical resolution and a better shape than the one in Figure 9(a).

Figure 10(a) is the migration image without knowing the correct Q_p model for the gas sand. The events under the gas sand are attenuated, in terms of their dimming amplitude, stretching and distorted wavelets in Figure 10(a). Figure 10(b) is the migration image compensated using the inverted Q_p model in Figure 9(b). The results show that compensation



(a)



(b)

Figure 6: (a)The M_{Dry} model is superimposed on the log; (b)The relation between V_p and Q_p of the log. [**NR**] yishen1/. NR/P.Mdry.log,NR/P.QV.Log

adequately restores both the amplitude and frequency of the events below the gas sand.

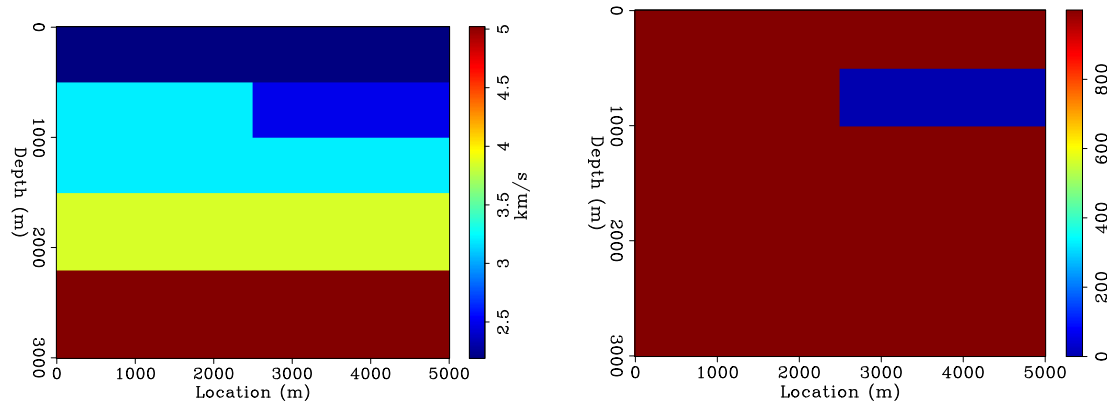


Figure 7: (a) Correct V_p model (b) Correct Q_p model. The Q values are clipped to 1,000 for a convenient display of the gas layer. The Q values are different in each layer and are larger than 1,000 in the nongas sand. [**ER**] yishen1/. layer4.vel,layer4.q

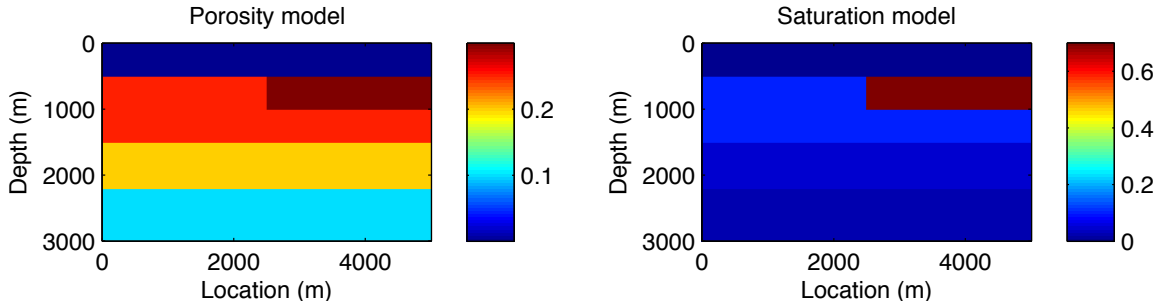


Figure 8: Rock properties: (a) porosity model; (b) gas saturation model. [**NR**] yishen1/. NR/layer4Rock

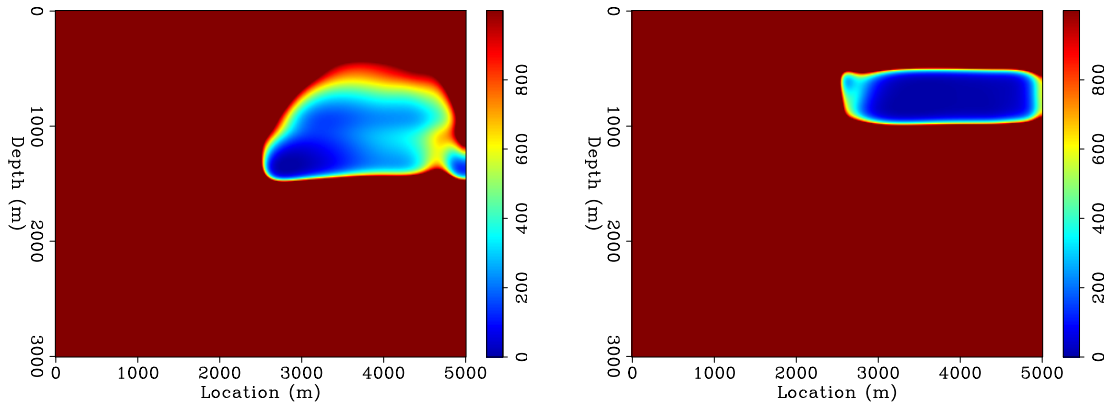
CONCLUSION

We derived an approximate closed-form solution relating V_p to Q_p using rock physics modeling. This solution is validated using well data in which the elastic properties were measured and Q was derived numerically using rock physics. Finally we applied our new $Q_p - V_p$ equation to synthetic seismic data, which produced an improved Q estimated model. We showed that this improved Q model leads to a better seismic migration image.

REFERENCES

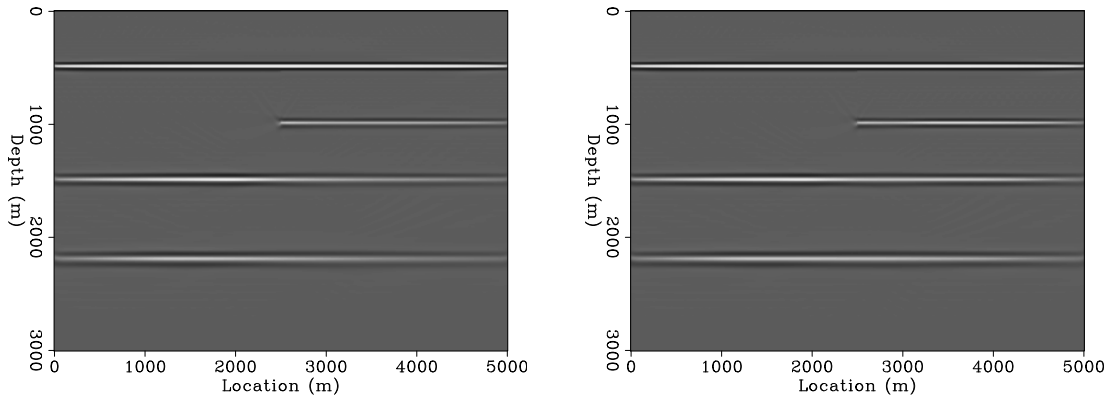
Billette, F. and S. Brandsberg-Dahl, 2005, The 2004 BP velocity benchmark: 67th Conference and Exhibition, EAGE, Extended Abstracts, B035.

Dvorkin, J., M. A. Gutierrez, and D. Grana, April 2014, Seismic reflections of rock properties: Cambridge University Press.



(a) (b)

Figure 9: (a) Inverted Q_p model without constraint. (b) Inverted Q_p model with constraint, with the parameters of gas sand:$c1 = 0.0038, c2 = 0.0029, c3 = 2.051 * 10^{-5}$ and non gas sand: $c1 = 0.0038, c2 = 0.0036, c3 = 5.97 * 10^{-6}$. [**ER**] yishen1/. layer4.bqtv.invq,layer4.vc.bqtv.invq



(a) (b)

Figure 10: Prestack migration image: (a) Attenuated image; (b) Compensated image. [**ER**] yishen1/. layer4.vc.bqtv.bimg,layer4.vc.bqtv.img

Dvorkin, J. and A. Nur, 1996, Elasticity of highporosity sandstones: Theory for two North Sea data sets: Geophysics, **61**, 1363–1370.

Gassmann, F., 1951, Uber die elastizitat poroser medien: Vierteljahrsschrift der Naturforschenden Gesellschaft in Zrich, **96**, 1–23.

He, Y. and J. Cai, 2012, Q tomography towards true amplitude image and improve sub-karst image: **521**, 1–5.

Li, Y., Y. Shen, and P. Kang, 2015, Integration of seismic and fluid-flow data: a two-way road linked by rock physics: EAGE – 77th EAGE Conference and Exhibition.

Mavko, G., T. Mukerji, and J. Dvorkin, 1991, The handbook of rock physics: Cambridge University Press.

Muller, T. M., B. Gurevich, and M. Lebedev, 2010, Seismic wave attenuation and dispersion resulting from wave-induced flow in porous rocks A review: Geophysics, **75**, A147–A164.

Raymer, L. L., E. R. Hunt, and J. S. Gardner, 1980, An improved sonic transit time-to-porosity transform: Society of Petrophysicists and Well-Log Analysts, **January**, no. 1.

Shen, Y., B. Biondi, R. Clapp, and D. Nichols, 2013, Wave-equation migration Q analysis (WEMQA): EAGE Workshop on Seismic Attenuation Extended Abstract.

———, 2014, Wave-equation migration Q analysis (WEMQA): SEG Technical Program Expanded Abstracts.

Wang, Y., S. Chen, L. Wang, and X. Li, 2013, Modeling and analysis of seismic wave dispersion based on the rock physics model: J. Geophys. Eng., **10**.

Zhou, J., X. Wu, K. H. Teng, Y. Xie, F. Lefeuvre, I. Anstey, and L. Sirgue, 2014, FWI-guided Q tomography and Q-PSDM for imaging in the presence of complex gas clouds, a case study from offshore Malaysia: **136**, 536–540.

Rayleigh-wave group velocity tomography using traffic noise at Long Beach, CA

Jason P. Chang and Biondo Biondi

ABSTRACT

Anthropogenic sources recorded by passive seismic arrays provide the opportunity for ambient noise cross-correlation techniques to effectively use frequencies well beyond the microseism band. Using data recorded by a dense array in Long Beach, California, we demonstrate that high-frequency (> 3 Hz) fundamental- and first-order-mode Rayleigh waves generated by traffic noise can be extracted from the ambient noise field and used for tomographic studies. Here, we use travel times of the fundamental-mode Rayleigh waves in a straight-ray tomography procedure to derive group velocity maps at 3.00 Hz and 3.50 Hz. The velocity trends in our results correspond to shallow depths and correlate well with lithologies outlined in a geologic map of the survey region. As expected, less-consolidated materials display relatively low velocities, while more-consolidated materials display relatively high velocities. Our results suggest important implications for research investigations concerned with the near-surface.

INTRODUCTION

Extracting surface waves from the ambient noise field for use in tomographic studies is well-established at both regional and continental scales (e.g., Shapiro et al., 2005; Yang et al., 2008; Bensen et al., 2008). The success of these studies has encouraged recent investigation of this technique at the exploration scale. In ocean-bottom environments, de Ridder et al. (2014) and Mordret et al. (2014) obtained reliable time-lapse group and phase velocity maps, respectively, of the Valhall overburden, while de Ridder et al. (2015) recovered phase velocity and anisotropy maps representative of subsidence patterns at Ekofisk. In those cases, they examined frequencies in the microseism band (< 2 Hz). In land environments, passive seismic arrays of sufficient density, size, and duration for these sorts of studies are rare.

One array that does meet these requirements was located in Long Beach, California (map in Figure 1). Deployed in January 2012 by NodalSeismic, the array spans an 8.5×4 km^2 region and consists of approximately 2400 vertical-component geophones. With an average geophone spacing of 100 m and a continuous 3-month recording period, the array is well-suited for exploration-scale tomography using ambient noise. Dahlke et al. (2014) were able to create phase velocity maps at low frequencies (approximately 1 Hz) that resolved the location of the Newport-Inglewood Fault. At the neighboring array, Lin et al. (2013) were able to create similar Rayleigh-wave phase velocity maps for frequencies up to 2 Hz that successfully imaged the same fault.

The focus of our present study is to create high-frequency (> 3 Hz) group velocity maps at the Long Beach array using passive seismic recordings. At these frequencies, traffic noise

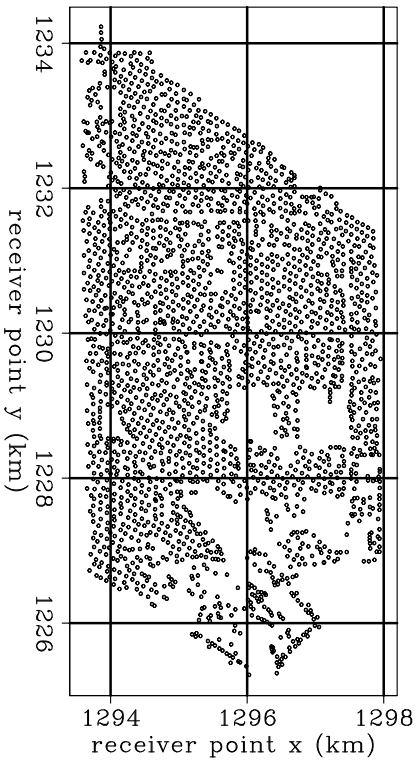


Figure 1: Map of stations at the Long Beach passive seismic array. Coordinates are NAD27, CA State Plane, Zone 7, kilometers. [**CR**] jason1/. map-black

dominates the ambient noise field. Power spectral density maps (Chang et al., 2013; Nakata et al., 2015) and beamforming of the ambient noise field (Chang et al., 2014) reveal Interstate 405 as the dominant source of seismic energy and local roads as the secondary source. We begin with an overview of our ambient noise processing procedure for harnessing this traffic noise. We then explain our fundamental-mode Rayleigh-wave travel time selection process and straight-ray tomography approach. To validate our results, we compare our group velocity maps to a geologic map of the region. Finally, we examine the resolution of our results with checkerboard tests.

ESTIMATING GREEN'S FUNCTIONS

To extract inter-station Green's functions from ambient noise recordings, we apply a processing procedure adapted from Bensen et al. (2007). Continuous recordings from all receiver pairs are divided into simultaneous, non-overlapping two-hour time windows. To compensate for expected variations in the source amplitude due to varying traffic conditions over time, we whiten the traces prior to cross-correlation. In the frequency domain, this process is expressed as:

$$[G(x_B, x_A, \omega) + G^*(x_B, x_A, \omega)] = \left\langle \left(\frac{U(x_B, \omega)}{\{|U(x_B, \omega)|\}} \right) \left(\frac{U^*(x_A, \omega)}{\{|U(x_A, \omega)|\}} \right) \right\rangle, \tag{1}$$

where G is the Green's function between two receiver locations (x_A, x_B), $U(x, \omega)$ is the spectrum of the wavefield at a given receiver location x, $*$ is the complex conjugate, $\langle \cdot \rangle$ is the time-averaged ensemble, $|\cdot|$ is the magnitude of the spectrum, and $\{\cdot\}$ is a 0.003 Hz running window average used for normalizing the signal. This procedure is equivalent to

calculating the cross-coherence between two traces. For this study, we sum 384 two-hour time windows (or 32-days worth) of cross-correlations.

Applying this processing procedure to a line of receivers running perpendicular to Interstate 405, Chang et al. (2014) verified that the extracted Green's functions above 3 Hz were dominated by fundamental- and first-order-mode Rayleigh waves generated by the highway and local roads. An example of an extracted Green's function with two Rayleigh-wave modes is shown in Figure 2. In this study, we use these correlations to perform group velocity tomography.

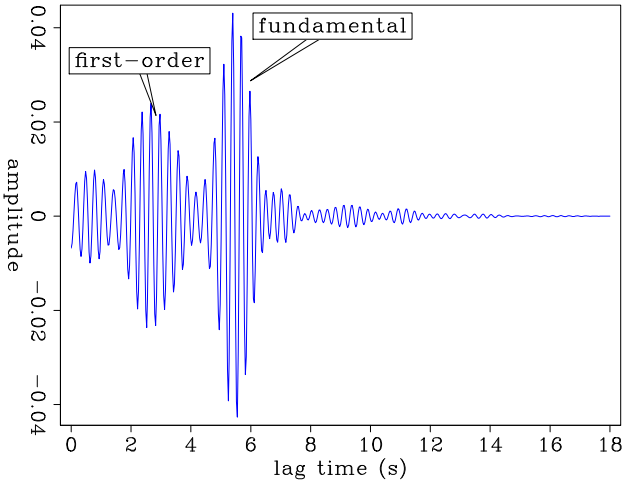


Figure 2: Example of an estimated Green's function between a virtual source and receiver at 3.50 Hz. The earlier peak corresponds to the first-order Rayleigh-wave mode, while the later peak corresponds to the fundamental mode. [**CR**] jason1/. high-trace

GROUP VELOCITY TOMOGRAPHY

Travel time selection

Group travel time between two stations is picked as the peak of the envelope of the bandpass-filtered trace that has been folded over zero time lag. In the frequency domain, the filter is applied as a Gaussian taper centered at the frequency of interest. For this study, we are interested in shallow velocity trends. Because the fundamental-mode Rayleigh wave is more sensitive to shallower depths than the first-order mode (for a given frequency), we are only interested in travel times of the fundamental mode. For this process, we calculate two linear moveout windows: one based on the fundamental mode, and the other based on the first-order mode. Both windows are between 2 and 3 seconds wide for the frequencies examined. Respective velocities are estimated from the frequency-wavenumber representation of a virtual source gather centered on Interstate 405 and sorted by radial offset (Figure 3). We first apply a mute to the moveout window associated with the first-order mode. This is done for two reasons. One is to avoid accidentally associating this earlier peak with the travel time of the fundamental mode. Another is to avoid bias in our signal-to-noise ratio (SNR) estimate due to the presence of multiple distinct peaks. Here, we calculate SNR as the ratio of the maximum amplitude of the envelope inside the fundamental-mode moveout window to the root-mean-square of the envelope outside both the fundamental and first-order moveout windows.

For the two central frequencies we examined (3.00 Hz and 3.50 Hz), we kept correlations with SNR greater than 5. Figure 4 shows maps of travel times obtained from traces

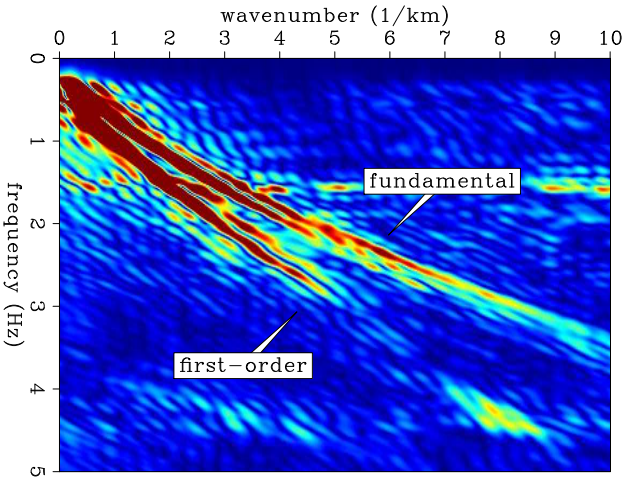


Figure 3: Frequency-wavenumber plot derived from a virtual source gather centered at Interstate 405. Note the two clear Rayleigh wave modes. The shallower sloping event corresponds to the fundamental mode, while the steeper sloping event corresponds to the first-order mode. [**CR**] jason1/. fk-freeway1-high

at 3.50 Hz that have passed the SNR criterion for four different virtual source locations. Because Interstate 405 effectively serves as an active source, it causes artificially early peaks in the correlations where the virtual source and receiver are on opposite sides of the highway. These earlier-than-expected peaks degrade the SNR and lead to the lack of accepted travel times on the opposite side of Interstate 405. We can see this effect in Figures 4(a) and 4(b). In the former figure, the virtual source is located north of the highway, and the truncated border of the travel time map toward the south correlates well with the path of the highway. A similar pattern is observed in the latter figure, where the virtual source is located south of the highway. Further away from the highway, the ambient noise field is dominated by Rayleigh waves generated by local traffic (Chang et al., 2014). Thus, it is no surprise that we still have suitable travel times when the virtual source is in the center of the array (Figure 4(c)). In the sparse southeast part of the array, there are relatively few sources of traffic noise due to the lack of busy roads and the presence of a marina. Thus, it is expected that there are very few suitable travel times in that part of the array (Figure 4(d)). Together, these travel time maps suggest that our fundamental-mode Rayleigh-wave travel time selection process is potentially reliable.

For our tomography study, we only use correlations with virtual source-receiver spacing between 0.5 and 1.7 km at 3.00 Hz and between 0.4 and 1.5 km at 3.50 Hz. The minimum spacing is approximately one wavelength longer than the suggested minimum of three wavelengths (Moschetti et al., 2007). The maximum spacing is chosen by looking at the travel time maps. Out of a possible $2,980,143$ traces, we kept $530,876$ traces (18%) at 3.00 Hz and $403,523$ traces (13%) at 3.50 Hz. Although we used significantly fewer traces than available, the inversion problem remains highly over-determined. The goal is to use the most reliable travel times, not the most number of travel times.

Inversion procedure

To obtain group velocity maps, we employ a straight-ray tomography procedure with regularization. The problem is posed as solving for the slowness perturbation, $\Delta\mathbf{m}$, with respect to an average slowness m_0. Therefore, the final slowness model, $\mathbf{m}$, is defined as $\mathbf{m} = m_0 + \Delta\mathbf{m}$. The average slowness is calculated from the travel time and offset infor-

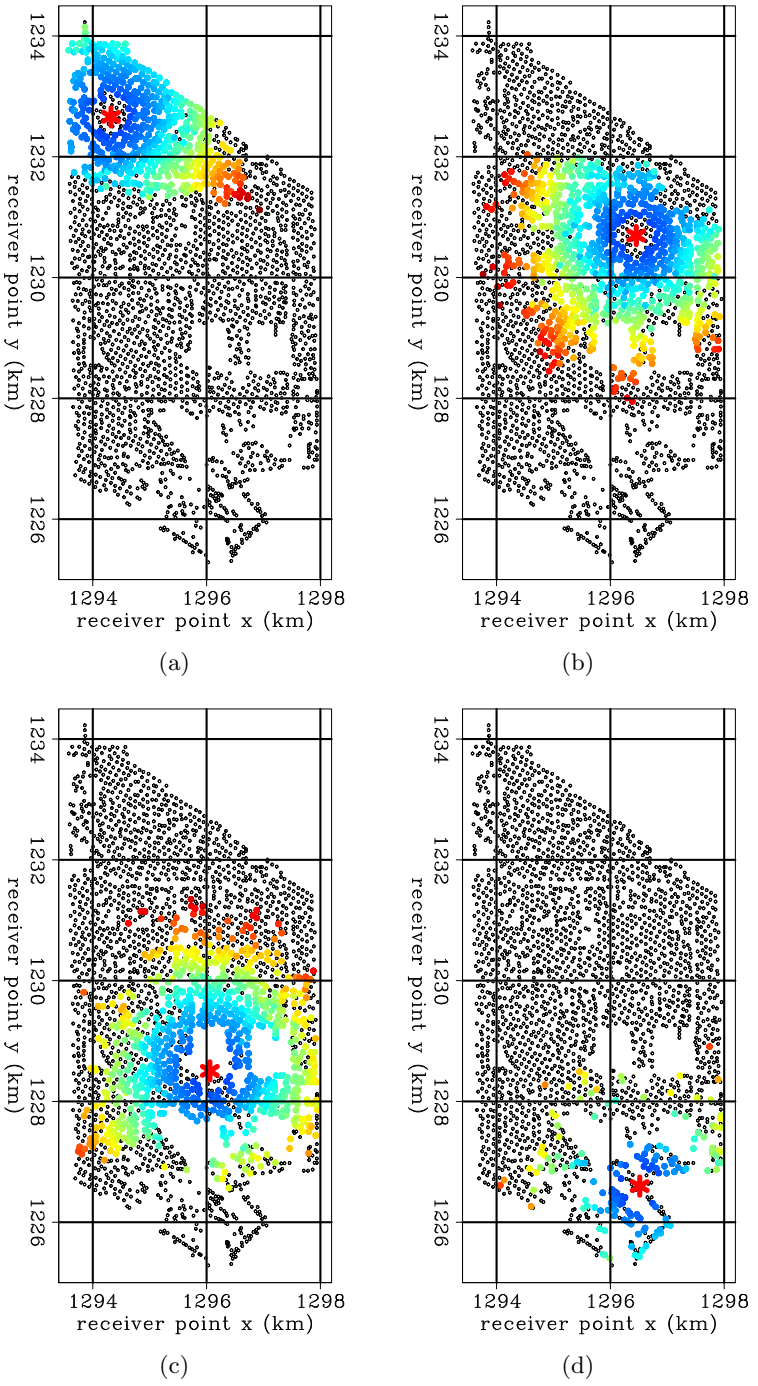


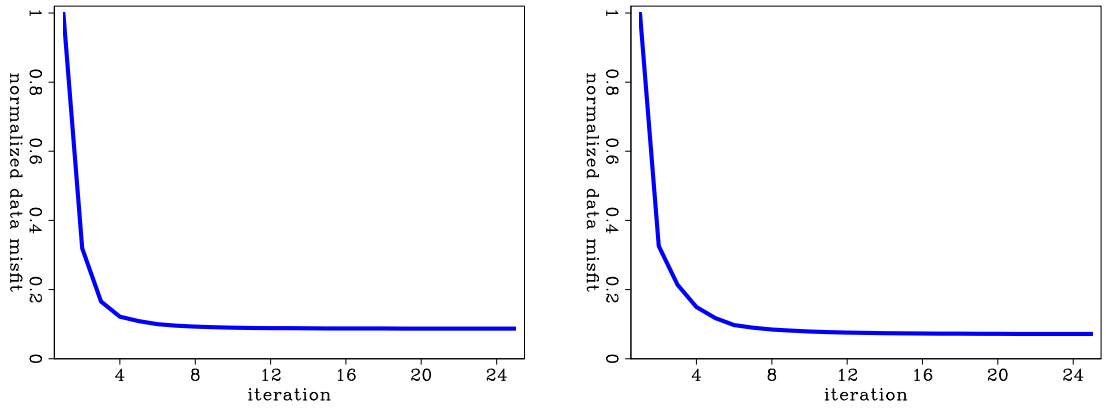
Figure 4: Fundamental-mode Rayleigh-wave travel time maps at 3.50 Hz for four different virtual source locations (indicated by red asterisks). Warmer colors correspond to later travel times. Only travel times obtained from correlations with SNR greater than 5 are displayed. (a) Virtual source north of Interstate 405. (b) Virtual source south of Interstate 405. In these cases, note the lack of suitable travel times on the other side of the highway. (c) Virtual source near the center of the array. (d) Virtual source near the coastline. The relative lack of suitable travel times in the southeast region of the array is likely due to the relative lack of traffic noise. [**CR**] jason1/. traveltime1,traveltime2,traveltime3,traveltime4

mation of all traces that pass quality control. We subtract the contribution of the average slowness from each travel time to obtain residual travel times, $\Delta\mathbf{t}$.

Modeled residual travel times are obtained by applying a straight-ray tomography operator, $\mathbf{F}$, to a slowness perturbation model. Rows of the operator contain the length of a straight ray through each model grid cell for a single virtual source-receiver path. For this data set, we construct a 110×220 model grid space, with each grid covering a 50×50 m^2 region. We use a conjugate gradient approach to minimize the following objective function:

$$J(\Delta\mathbf{m}) = ||\mathbf{F}\Delta\mathbf{m} - \Delta\mathbf{t}||_2^2 + \epsilon||\nabla^2\Delta\mathbf{m}||_2^2 \quad . \tag{2}$$

∇^2 represents the Laplace operator, which is the roughening operator used for regularization. ϵ represents regularization strength, which balances the data-fitting objective and the model-smoothing objective. We iterated 25 times, which is enough to converge to a solution (Figure 5).



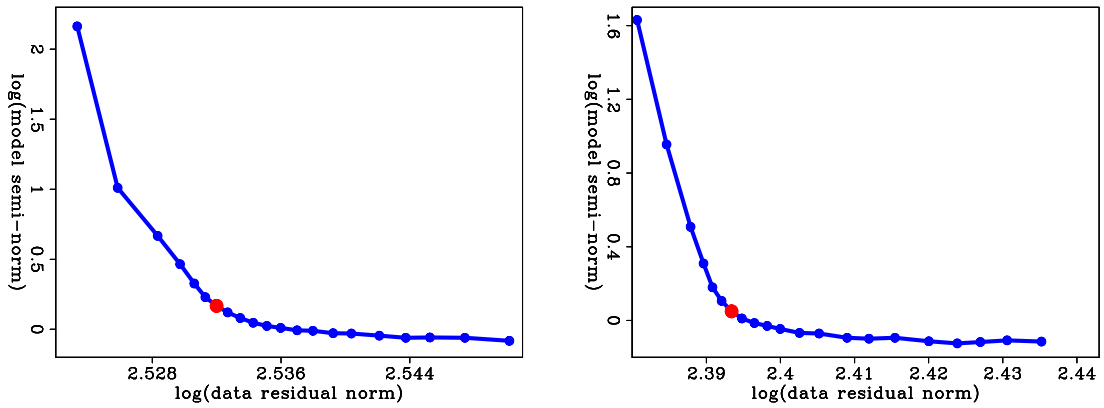
(a) (b)

Figure 5: Norm of the data residual as a function of iteration. Residuals are normalized by the initial residual. (a) 3.00 Hz. (b) 3.50 Hz. [**CR**] jason1/. error-3p00hz-eps4,error-3p50hz-eps4

Model selection

To choose the regularization strength (and hence a model), we examine the L-curve (Aster et al., 2013). An L-curve is a trade-off curve between the data residual misfit ($||\mathbf{Fm} - \mathbf{d}||_2$) and the model semi-norm ($||\nabla^2\mathbf{m}||_2$). Different regularization strengths will plot different points on this graph, with lower values sitting at the top left of the curve and higher values sitting at the bottom right of the curve. The resulting curve shape is typically an "L". Often, the ϵ that produces a solution closest to the corner of this curve is chosen, as it balances the model smoothing and data fitting parameters best.

For these tomography problems, we examine ϵ values from 0 to 5, in increments of 0.25. The resulting L-curves for 3.00 and 3.50 Hz are shown in Figures 6(a) and 6(b), respectively. To better locate the corner point, we plot the axes on a logarithmic scale. In both cases, we choose a regularization strength of 1.50 (indicated by the red dots on the curves).



(a) (b)

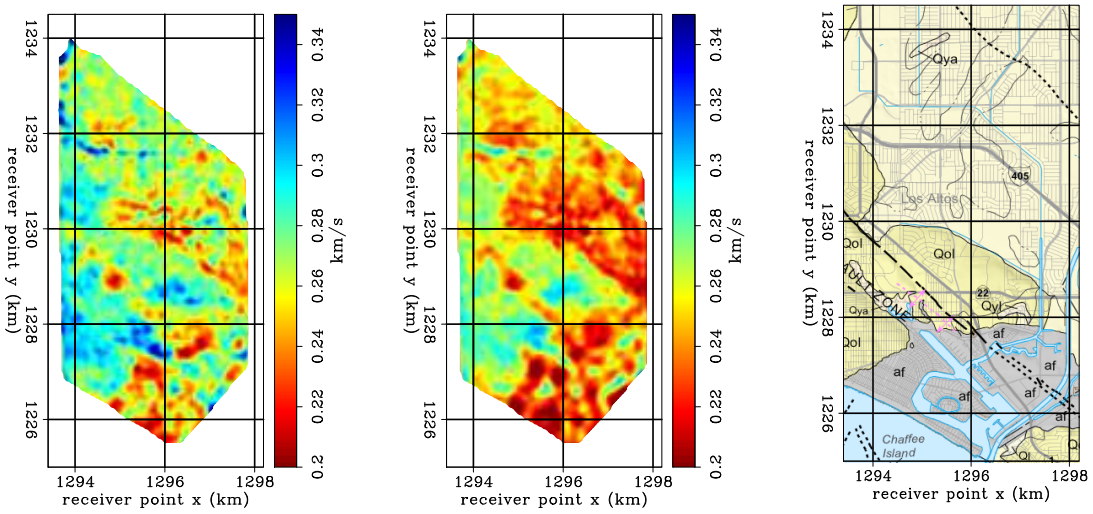
Figure 6: Trade-off curves between the norm of the data residual and the semi-norm of the resulting model for different regularization strengths ϵ (0 is at the top left of the curve and 5 is at the bottom right of the curve). (a) For 3.00 Hz. (b) For 3.50 Hz. The red dot corresponds to $\epsilon = 1.50$ and is used for the velocity maps shown in this paper. [**CR**] jason1/. lc-3p00hz-eps4,lc-3p50hz-eps4

GROUP VELOCITY MAPS

Due to the dispersive nature of Rayleigh waves, the group velocity map at 3.00 Hz (Figure 7(a)) shows generally higher velocities than those in the group velocity map at 3.50 Hz (Figure 7(b)). However, because the wavelengths at these two frequencies are similar, it is not a surprise that they show the same general velocity trends and features. We interpret our results with the guidance of a geologic map of the survey region in Figure 7(c) (California Department of Conservation, 2012).

One of the most prominent trends in our group velocity maps is a boundary cutting northwest-southeast across the array that separates a lower velocity region (north) from a higher-velocity region (south); it is most distinct south of the eastern portion of Interstate 405. The location of this boundary matches well with a boundary in the geologic map that separates Holocene to late Pleistocene alluvial valley deposits (north; light yellow) from late to middle Pleistocene lacustrine, playa, and estuarine deposits (south; deep yellow). The deposits in the north typically consist of unconsolidated to slightly consolidated material, while the deposits in the south typically consist of slightly to moderately consolidated material. Thus, it is no surprise that we find the northern region to be of a lower velocity than the southern region. Another prominent feature is a low velocity zone in the southeast region of the array. This zone corresponds to the location of Alamitos Bay, which is built on artificial fill (grey region in Figure 7(c)). Thus, our tomography results appear reliable, since artificial fill typically displays relatively low velocities (Wills and Silva, 1998).

Two other features are the high velocities overlaying Interstate 405 and the low velocities overlaying the Newport-Inglewood Fault. The former trend could be an artifact, as travel times from ray paths that cross the highway can be artificially early due to the highway acting as an active source. The latter trend is curious, as the fault displays relatively high velocities at low frequencies and deeper depths (Dahlke et al., 2014; Lin et al., 2013), whereas



(a) (b) (c)

Figure 7: Group velocity maps generated at (a) 3.00 Hz and (b) 3.50 Hz. (c) Geologic map of the survey region (California Department of Conservation, 2012). Note the similarities between the velocity trends in the tomography results and the lithologies outlined in the geologic map. [**CR**] [**CR**] [**NR**] jason1/. tomo-3p00hz,tomo-3p50hz,geo-coord

the fault seems to display relatively low velocities at higher frequencies and shallower depths. We require better understanding of fault systems to determine whether this low-velocity fault trend is reasonable.

MODEL RESOLUTION

We address the spatial resolution of the model in two ways. First, we examine the cumulative ray length in each model grid cell (Figure 8) when using acceptable travel times at 3.50 Hz. Longer ray lengths correspond to more rays passing through the cell. As expected, the ray coverage is highest in the center of the array. The relatively low-coverage areas correspond to gaps in the array. The lack of coverage is particularly evident in the southeast region of the array, where the effect of the lack of receivers is compounded by the relative lack of traffic noise.

Second, we run checkerboard tests. Although there are shortcomings when using this method to assess spatial resolution (Lévěque et al., 1993), we use it here for qualitative purposes. For this procedure, we first add a sinusoidal, or checkerboard, velocity perturbation to our final inverted model. For our study, we examine two checker sizes: 500 m $\times$500 m (Figure 9(a)) and 250 m $\times$250 m (Figure 10(a)). Peaks of the perturbation are set to $\pm$15% of the modeled velocity, which is smaller than the maximum percent difference between the major low- and high-velocity regions. As a result, the actual pattern is not uniform over the model. We then use our forward tomography operator to generate synthetic travel times between virtual source-receiver pairs for the perturbed velocity model. To simulate noise in our actual travel time picks, we add random travel time error to each synthetic travel time based on a Gaussian distribution with a mean of 0 and a standard deviation of 14%

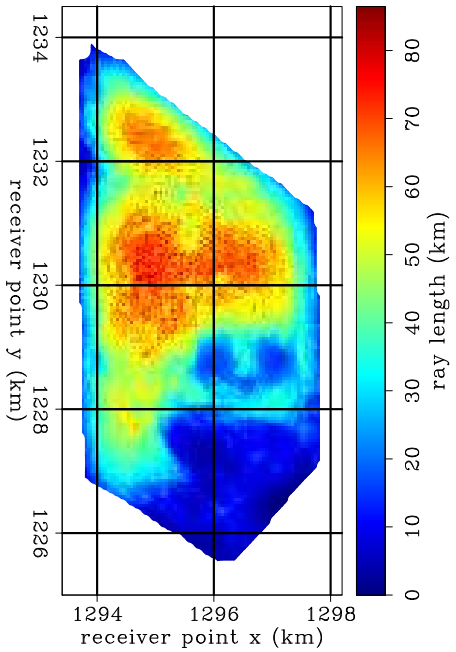


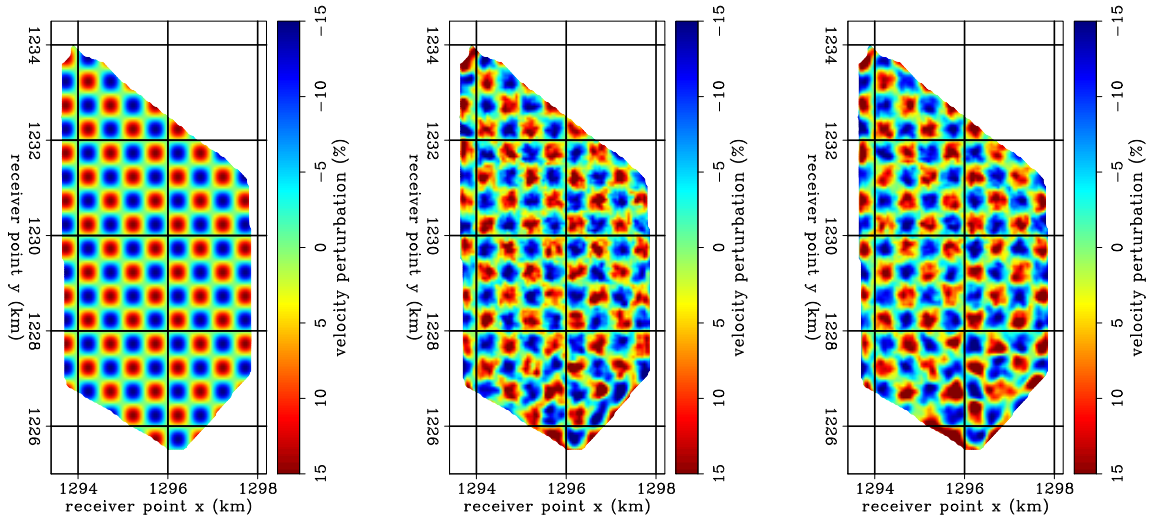
Figure 8: Cumulative ray length map for all station pairs used at 3.50 Hz. Although ray lengths in the southeast part of the array are relatively short, there is still enough coverage to trust the general low-velocity trend. [**CR**] jason1/. raylength-3p50hz

of the travel time. This standard deviation is calculated by taking the root-mean-square (RMS) of the absolute differences between each observed travel time and the corresponding synthesized travel time from the final velocity model, normalized by the observed travel time ($|t_{\text{obs}} - t_{\text{syn}}|/t_{\text{obs}}$). It is 14% for both velocity models. Finally, we invert these noisy synthetic travel times using the same parameters we used when inverting the actual data.

The inverted results for the larger checkers (500 m $\times$500 m) and smaller checkers (250 m $\times$250 m) are shown in Figures 9 and 10, respectively, for both 3.00 Hz and 3.50 Hz. For both checker sizes and both frequencies, it is clear that the high amount of noise in the travel times affects the smoothness of the recovered checkerboard pattern. This suggests we cannot completely trust the absolute velocity in every grid cell of our group velocity maps. Despite this, the overall checkerboard pattern is resolved to a reasonable degree in all cases. For larger checkers, the pattern is resolved throughout nearly the entire array. The exception is the southeast region of the array, where there is some smearing of the pattern. This is not a surprise since there was relatively low ray coverage in that region. For smaller checkers, there is very poor resolution in the same southeast region of the array. The resolution is worse when looking at 3.50 Hz (Figure 10(c)) than when looking at 3.00 Hz (Figure 10(b)). This is likely because there are fewer acceptable ray paths through that region at higher frequencies due to attenuation. These results show that we have up to 250 m $\times$250 m resolution in the center of the array, and up to nearly 500 m $\times$500 m resolution on the fringes of the array. This level of resolution is high enough to trust the broad velocity features we find in our group velocity tomography results.

CONCLUSIONS AND FUTURE WORK

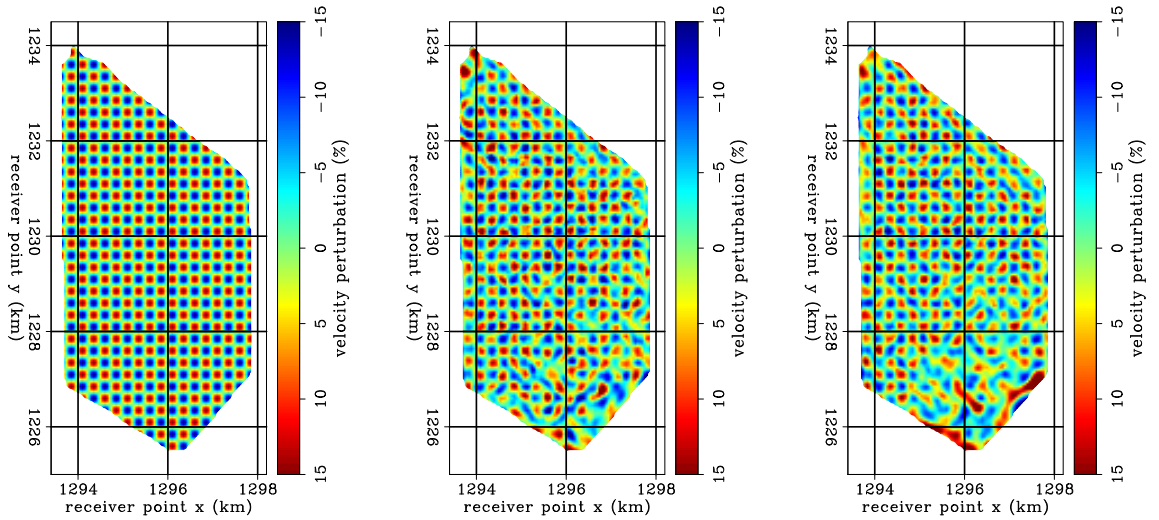
We showed that high-frequency Rayleigh waves primarily generated by traffic noise can be extracted from ambient noise recordings using cross-correlation techniques. The travel



(a) (b) (c)

Figure 9: Checkerboard test results using sinusoidal checkers 500 m ×500 m in size and a maximum velocity perturbation of 15%. (a) Modeled checkers. (b) Inverted results at 3.00 Hz. (c) Inverted results at 3.50 Hz. Note the poor resolution in the southeast region of the array. [**CR**]

jason1/. cb-3p00hz-500m-model,cb-3p00hz-500m-recon,cb-3p50hz-500m-recon



(a) (b) (c)

Figure 10: Checkerboard test results using sinusoidal checkers 250 m ×250 m in size and a maximum velocity perturbation of 15%. (a) Modeled checkers. (b) Inverted results at 3.00 Hz. (c) Inverted results at 3.50 Hz. Note the poor resolution in the southeast region of the array. [**CR**]

jason1/. cb-3p00hz-250m-model,cb-3p00hz-250m-recon,cb-3p50hz-250m-recon

times of the fundamental-mode Rayleigh waves appear to be reliable enough for performing group velocity tomography on the entire array. From our group velocity maps, we are able to distinguish velocity trends that agree with a geologic map of the region. Unconsolidated materials in the northern region appear to have relatively low velocities, while more consolidated materials in the southern region appear to have relatively high velocities. We are also able to resolve low-velocity material in Alamitos Bay, which is built on artificial fill, as well as a low velocity signature that aligns with the Newport-Inglewood Fault. Resolution tests indicate that we can trust these features. Because our investigation frequencies are high, and hence correspond to shallow depths, our results can potentially be useful for identifying regions that are susceptible to serious damage during earthquake-related shaking.

Our future goals are to improve the quality of our travel time picks, since they are currently noisy. A more careful selection of minimum and maximum virtual source-receiver distances could be beneficial, as we want to ensure that we avoid picking travel times when the fundamental and first-order modes interfere. We would also like to test our method of harnessing traffic noise by using a sparser subset of receivers because land arrays as dense as the Long Beach array are rare. Finally, it is highly unlikely we need 35-days worth of data to obtain stable correlations. We hope to create travel time maps from daily correlations for time-lapse monitoring purposes. Detectable changes in the subsurface could be due to earthquake events or precipitation.

ACKNOWLEDGMENTS

We thank Signal Hill Petroleum, Inc. for access to this data set and permission to publish. Thanks to Dan Hollis of NodalSeismic for his cooperation and help throughout this research. Thank you to Sjoerd de Ridder, Nori Nakata, Kevin Seats, Taylor Dahlke, Bob Clapp, and Stew Levin for helpful discussions. Many thanks to the sponsors of the Stanford Exploration Project for their financial support.

REFERENCES

Aster, R. C., B. Borchers, and C. H. Thurber, 2013, Parameter estimation and inverse problems: Academic Press.

Bensen, G., M. Ritzwoller, M. Barmin, A. Levshin, F. Lin, M. Moschetti, N. Shapiro, and Y. Yang, 2007, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements: Geophysical Journal International, **169**, 1239–1260.

Bensen, G., M. Ritzwoller, and N. Shapiro, 2008, Broadband ambient noise surface wave tomography across the United States: Journal of Geophysical Research: Solid Earth, **113**, 1–21.

California Department of Conservation, 2012, Geologic compilation of Quaternary surficial deposits in Southern California: Technical report.

Chang, J. P., S. de Ridder, B. Biondi, et al., 2013, Power spectral densities and ambient noise cross-correlations at Long Beach: 83rd Annual International Meeting, SEG, Expanded Abstracts, 2196–2200.

Chang, J. P., N. Nakata, R. G. Clapp, B. Biondi, S. de Ridder, et al., 2014, High-frequency surface and body waves from ambient noise cross-correlations at Long Beach, CA: 84th Annual International Meeting, SEG, Expanded Abstracts, 2235–2239.

Dahlke, T., G. Beroza, J. Chang, S. de Ridder, et al., 2014, Stochastic variability of velocity estimates using eikonal tomography on the Long Beach data set: 84th Annual International Meeting, SEG, Expanded Abstracts, 2352–2356.

de Ridder, S., B. Biondi, and R. Clapp, 2014, Time-lapse seismic noise correlation tomography at Valhall: Geophysical Research Letters, **41**, 6116–6122.

de Ridder, S., B. Biondi, and D. Nichols, 2015, Elliptical-anisotropic eikonal phase-velocity tomography: Geophysical Research Letters, **42**, 758–764.

Lévěque, J.-J., L. Rivera, and G. Wittlinger, 1993, On the use of the checker-board test to assess the resolution of tomographic inversions: Geophysical Journal International, **115**, 313–318.

Lin, F.-C., D. Li, R. W. Clayton, and D. Hollis, 2013, High-resolution 3D shallow crustal structure in Long Beach, California: Application of ambient noise tomography on a dense seismic array: Geophysics, **78**, no. 4, Q45–Q56.

Mordret, A., N. M. Shapiro, and S. Singh, 2014, Seismic noise-based time-lapse monitoring of the Valhall overburden: Geophysical Research Letters, **41**, 4945–4952.

Moschetti, M., M. Ritzwoller, and N. Shapiro, 2007, Surface wave tomography of the western United States from ambient seismic noise: Rayleigh wave group velocity maps: Geochemistry, Geophysics, Geosystems, **8**, Q08010.

Nakata, N., J. P. Chang, J. F. Lawrence, and P. Boué, 2015, Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry: Journal of Geophysical Research: Solid Earth, **120**, 1159–1173.

Shapiro, N. M., M. Campillo, L. Stehly, and M. H. Ritzwoller, 2005, High-resolution surface-wave tomography from ambient seismic noise: Science, **307**, 1615–1618.

Wills, C. J. and W. Silva, 1998, Shear-wave velocity characteristics of geologic units in California: Earthquake Spectra, **14**, 533–556.

Yang, Y., M. H. Ritzwoller, F.-C. Lin, M. Moschetti, and N. M. Shapiro, 2008, Structure of the crust and uppermost mantle beneath the western United States revealed by ambient noise and earthquake tomography: Journal of Geophysical Research: Solid Earth, **113**, 1–9.

Applying interferometry to ambient seismic noise recorded by a trenched distributed acoustic sensing array

Eileen Martin, Jonathan Ajo-Franklin, Nate Lindsey, Tom Daley, Barry Freifeld, Michelle Robertson, Craig Ulrich, Shan Dou and Anna Wagner

ABSTRACT

We deployed a shallow trenched distributed acoustic sensing (DAS) array consisting of multiple fiber optic cables and casings to assess the suitability of DAS to recording ambient noise for near-surface characterization. We briefly describe the acquisition of an ambient noise dataset, our processing workflow, and preliminary results of interferometry on a small data subset. The approximate virtual source responses show reasonable Rayleigh wave velocity estimates compared to geophones. The results of several types of cable casings are comparable, which is encouraging for surveys deployed in conditions requiring durable materials. We show coherent virtual source responses at a relatively high frequency range and short recording period relative to most passive seismic surveys.

INTRODUCTION

DAS uses a standard fiber optic cable as both a strain rate sensor and a means of transmitting data to a storage array. A laser probes the cable with a pulse the length of a channel (in our experiments 1 m), then optical Rayleigh backscattering is recorded (Daley et al., 2013). Modifications of the optical time-domain reflectometry (OTDR) technique are used to convert the optical backscattering profile into acoustic traces associated with each 1 m channel. Refer to Bakku (2015) for more details on DAS.

DAS is becoming more popular due to its low cost per sensor, its dense spatial sampling, the potential for flexible geometries, and the possibility of deployment in situations that would be prohibitive for traditional sensors (Mateeva et al., 2013). Its limitations include noise spikes, directional sensitivity, and power loss along long cables. DAS only detects strain changes in-line with the cable. Additionally, the method for generating acoustic traces acts as a spatial derivative, leading to a $\cos^2\theta$ sensitivity to plane waves at an angle θ with the cable. This increased directionality has been verified experimentally with active sources for both surface DAS arrays (Lancelle et al., 2014) and borehole deployments (Mateeva et al., 2012; Ajo-Franklin et al., 2014).

Most DAS surveys in the literature use cables deployed in wells and active sources. DAS has been used to passively record microseismic events in a reservoir (Webster et al., 2013). The detection and analysis of active sources and microseismic events is quite different from the use of interferometry on ambient microseismic noise. Ambient noise interferometry is a powerful and cost effective technique that shows promise for characterizing the near surface (Wapenaar et al., 2006; Bensen et al., 2007; Chang et al., 2014; de Ridder, 2014).

We present a novel survey designed as a pilot study on the suitability of DAS to ambient noise studies, particularly for applications requiring low-cost continuous monitoring of the near surface.

SURVEY DESIGN

We deployed a trenched DAS array made up of several types of fiber optic cables and recorded data using a Silixa iDAS interrogator unit at the Richmond Field Station (RFS) in Richmond, CA, as shown in Figure 1. We also deployed a helical cable, but only present the results of the straight cables here. The trenches were approximately 2 ft. deep and all cables were laid flat along the bottom before the trench was backfilled. Along the east-west trench, we installed 24 three-component geophones at approximately 3 m spacing for comparison purposes. Further details of the deployment are described in Ajo-Franklin et al. (2015).

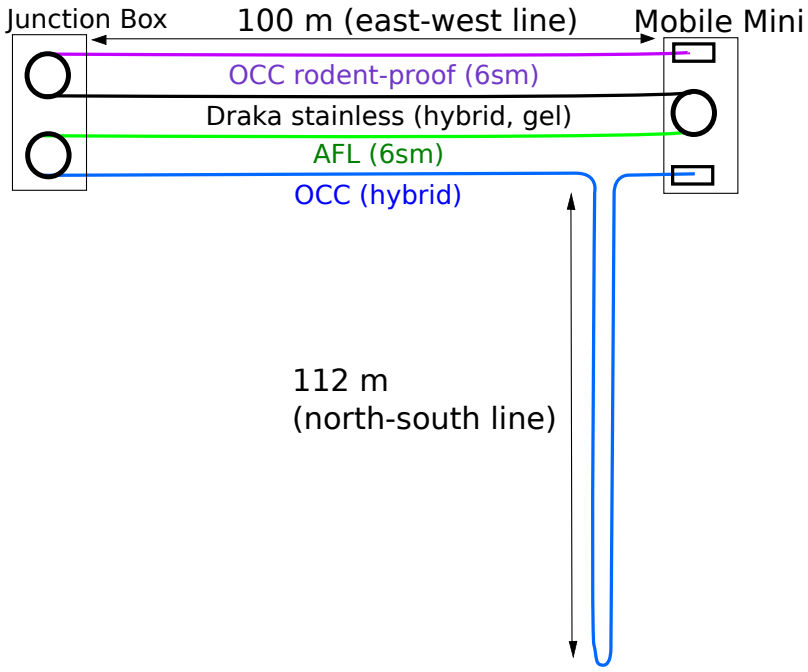


Figure 1: The trenched DAS survey at Richmond Field Station. We tested four types of straight fiber optic cable spliced end to end. The control unit and recording system are in the Mobile Mini. The cables pictured are: OCC hybrid 6 sm (blue), AFL 6 sm (green), Draka stainless-steel encased hybrid (black), and OCC rodent-proof 6 sm (purple). [**NR**] eileen1/. surveyDesign

Sensitivity analysis of DAS with respect to incoming plane wave angle and frequency shows survey design greatly influences the information detected. This survey was oriented to detect Rayleigh waves produced by noise sources including a nearby road and railway. The DAS data presented here were recorded at 2 kHz, and the geophone data were recorded at 1 kHz.

AMBIENT DAS DATA

As seen in Figures 2 and 3, the raw traces from a DAS channel and a geophone show some clear differences. There are differences in frequency sensitivity of DAS and geophones. An obvious difference is the presence of spikes in the DAS traces, a known issue in DAS data (Bakku et al., 2014). There are coherent events during some periods due to the presence of nearby large trucks and train tracks.

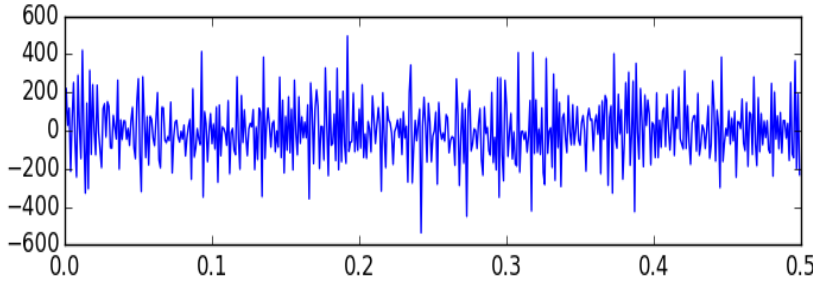


Figure 2: A raw ambient noise trace from the east end of the survey at channel 370 of the hybrid OCC cable. The DAS traces are recorded in DAS units so this is on a different scale from geophone traces. [**ER**] eileen1/. traceSampleDAS

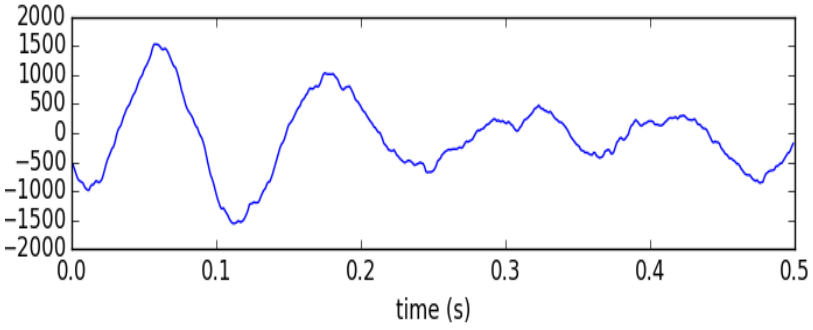


Figure 3: A raw ambient noise trace from the east end of the survey collected by the in-line component of a geophone. [**ER**] eileen1/. traceSampleGeo

Although we expected the response of all channels to be relatively uniform, we observed significant variation both in trace noise and time-varying gather noise. We are continuing to investigate the source of this variation. Channels were normalized against each other by scaling by the 1-norm of each trace.

PROCESSING WORKFLOW

For these initial results, we modified the standard ambient seismic noise interferometry workflow (Bensen et al., 2007) as seen in Figure 4. Despiking eliminates data points with absolute value more than twice the median absolute value in a small window (Bakku et al., 2014). Whitening boosts the magnitude of the spectrum up to a threshold chosen to be the median magnitude in the frequency band of interest.

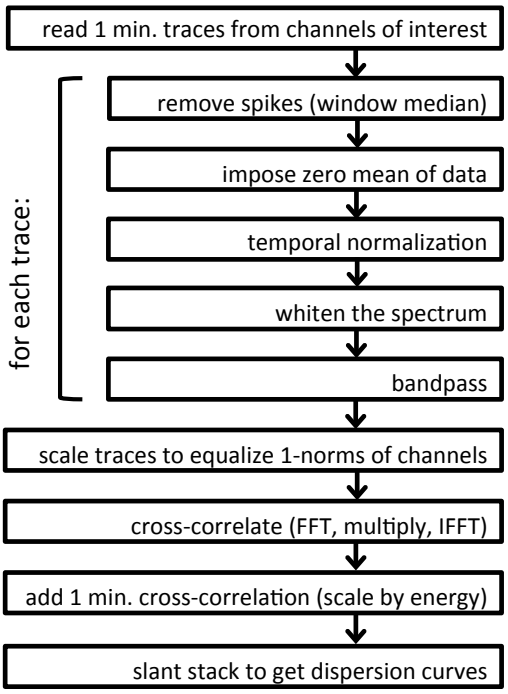
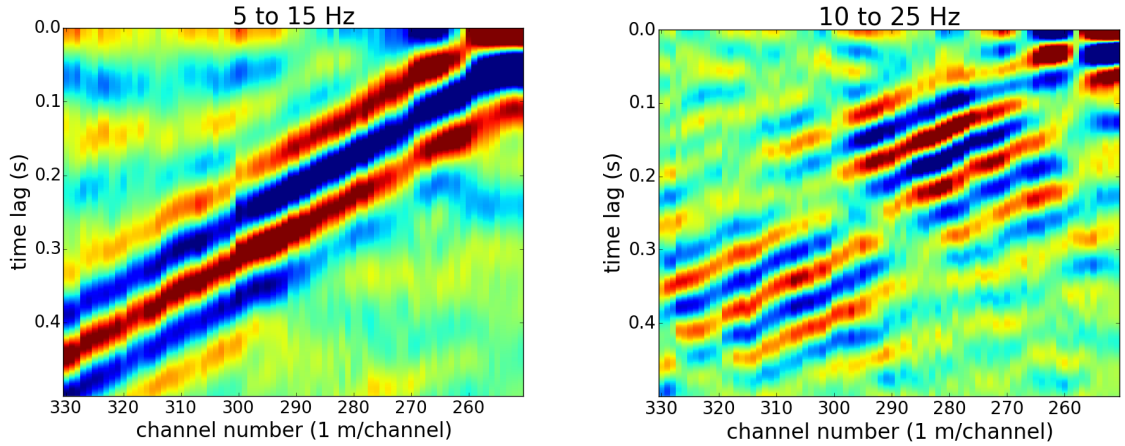


Figure 4: Our processing workflow is similar to the standard ambient seismic noise workflow. Despiking is necessary for DAS data. [**NR**] eileen1/. workflow

Geophones and DAS measure different physical quantities, so we did not expect the raw geophone and DAS traces to look the same, even after filtering. The estimators of virtual source responses (result of cross-correlations) are also not expected to look the same, but we hope to have some qualitative similarities between the DAS and geophone cross-correlations, particularly in the kinematics of observed events. In the final step of the processing workflow we hope to have dispersion curves revealing somewhat similar velocities. For each virtual source, the cross-correlation of the one minute windows are weighted by the inverse of the energy of the source (sum of the squared entries of the trace) then stacked.

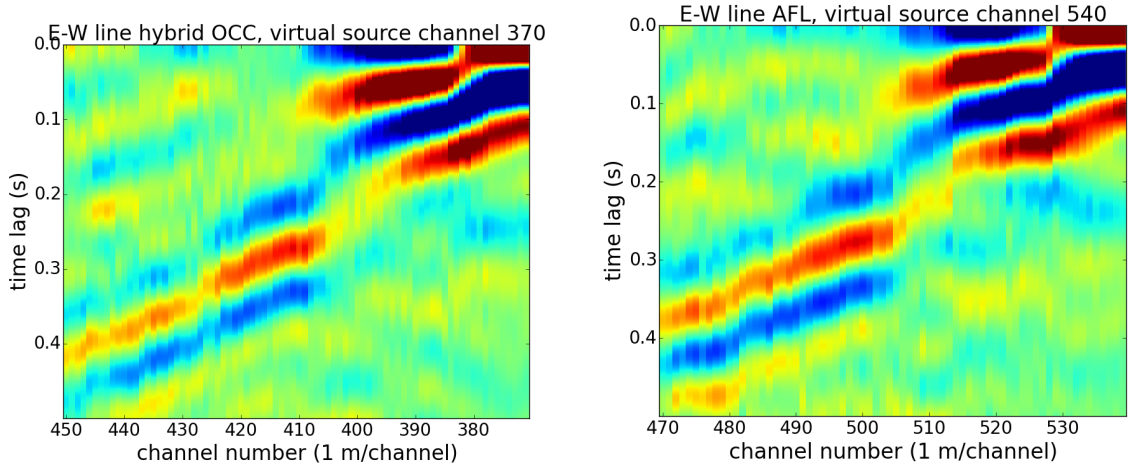
VIRTUAL SOURCE RESPONSES FROM DAS

After the preprocessing described in Figure 4 has been completed, we cross-correlate channels along the same side of the array. This is done for each channel, and the cross-correlations of each channel with every other channel serve as an estimator for the response to a virtual source located at that channel.



(a) (b)

Figure 5: Approximate virtual source responses from 10 minutes of data along the north-south hybrid OCC cable (cross-correlations with channel 250). The frequency ranges pictured are 5-15 Hz (left), 10-25 Hz (right). Positive and negative time lags are folded. [**ER**] eileen1/. xcorr250Lo5Hi15,xcorr250Lo10Hi25



(a) (b)

Figure 6: The estimated response to a virtual source at channel 370 on the east end of the east-west line of the hybrid OCC cable (left), and at channel 540 on the east end of the AFL cable (right). The frequency range pictured is 5-15 Hz, and this uses 10 minutes of ambient noise data. [**ER**] eileen1/. xcorr370,xcorr540

In Figures 5(a) and 5(b) we see the approximate response functions for 10 minute stacks of cross-correlations with a virtual source at the southern end of the hybrid OCC cable. We show this for two frequency ranges. In many ambient noise interferometry studies, a much longer recording time is needed to generate coherent virtual source response estimators, but these data were recorded in a particularly noisy area which contributed favorably to the convergence of the response estimators. There is little naturally occurring noise above the 20 Hz range, making it more difficult to pick a velocity from cross-correlations resulting from a short recording period. The response estimators show a dominant Rayleigh wave velocity between 200 and 400 m/s.

As seen in Figures 5(a), 5(b) and 6(a), the virtual source response approximations of both the north-south and east-west lines of the hybrid OCC cable show reasonable Rayleigh wave velocities. We expect different portions of the survey to display some variations in virtual sources response. It is promising that both trenches (north-south and east-west) can detect noise sources well enough to yield coherent virtual source responses, and that both the OCC hybrid and AFL cable show similar results.

DAS AND GEOPHONE COMPARISON

We deployed 24 three-component geophones at approximately 3 m spacing along the east-west trench, and analyzed the in-line components. Using the workflow from Figure 4 (except for the despiking step), we cross-correlated the in-line horizontal components of the geophones to estimate virtual source responses. The results in the 5-15 Hz frequency band for DAS channels and a geophone on the eastern end of the E-W line are shown in Figures 6(a) and 7, respectively. They appear qualitatively similar, although there are noticeable differences.

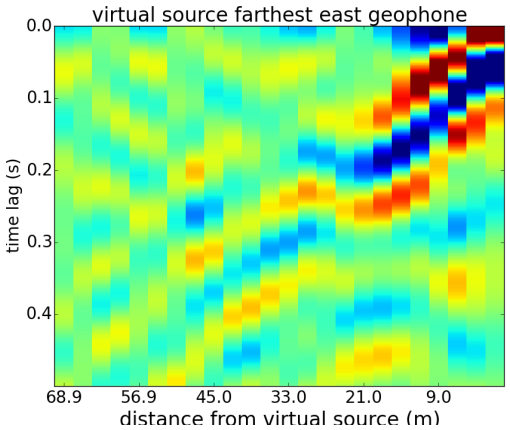


Figure 7: Cross-correlations of 40 minutes of ambient recording representing a virtual source from the eastern-most geophone's in-line component from 5-15 Hz. This is comparable to the cross-correlations from nearby DAS channels in Figure 6(a). [**ER**] eileen1/. xcorrGeos

The geophone cross-correlations look rough in comparison to the DAS cross-correlations. Even when these virtual source responses contain higher frequency information, the virtual source response estimates look very smooth compared to geophones. This is likely because the traces are calculated as a difference over a gauge length (about 10 m), which in theory

could cause a spatial spreading of any virtual source response estimates. Further analysis and data integration is needed to conclude whether these data reveal the same Rayleigh wave speed. The velocities appear to be in approximately the same range.

CONCLUSIONS

We deployed a novel survey design, a trenched DAS array to record ambient seismic noise. We developed a processing workflow to estimate virtual source responses using ambient noise interferometry. Using a small subset of the recorded ambient noise, this workflow resulted in estimated virtual source responses that showed Rayleigh wave velocities between 200 and 400 m/s for the frequency bands from 5-25 Hz. Compared to most passive seismic surveys, this is a relatively high frequency range.

The results presented here are the early analysis of a pilot test towards a system for low-cost continuous monitoring of the near surface. We aim to incorporate more data to estimate virtual source responses with more certainty, and to conduct statistical analyses to understand uncertainties and convergence. Although interferometry revealed reasonable results for Rayleigh wave velocities, the directional sensitivity of DAS effectively modifies a theoretical assumption justifying the use of interferometry. In particular, the theory assumes that random vibration sources are uniformly distributed around the receiver, but DAS is unable to detect vibrations perpendicular to the cable. Further theoretical work is needed to support the use of ambient noise interferometry on DAS data in a general context. In Summer 2015 we look forward to a larger scale deployment in Alaska, where we hope to image permafrost regions in the subsurface.

ACKNOWLEDGEMENTS

The field experiments are supported by the US Department of Defense under SERDP grant RC-2437 "Developing Smart Infrastructure for a Changing Arctic Environment Using Distributed Fiber-Optic Sensing Methods". Eileen Martin thanks the DOE CSGF fellowship grant number DE-FG02-97ER25308 for financial support, and Jason Chang, Shuki Ronen and Stew Levin for helpful suggestions.

REFERENCES

Ajo-Franklin, J., T. Daley, B. Freifeld, D. Tang, R. Zhang, A. Wagner, S. Dou, N. Lindsey, K. Bjella, and R. Pevzner, 2014, Development of a surface-wave imaging system for geotechnical applications based on distributed acoustic sensing (DAS) and ambient noise interferometry. Presented at AGU Fall Meeting.

Ajo-Franklin, J., N. Lindsey, S. Dou, T. Daley, B. Freifeld, E. Martin, M. Robertson, C. Ulrich, and A. Wagner, 2015, A field test of distributed acoustic sensing for ambient noise recording. Submitted.

Bakku, S., 2015, Fracture characterization from seismic measurements in a borehole: PhD thesis, Massachusetts Institute of Technology.

Bakku, S., P. Wills, M. Fehler, J. Mestayer, A. Mateeva, and J. Lopez, 2014, Vertical seismic profiling using distributed acoustic sensing in a hydrofrac treatment well: Expanded Abstracts of the 84th Ann. Internat. Mtg.

Bensen, G., M. Ritzwoller, M. Barmin, A. Levshin, F. Lin, M. Moschetti, N. Shapiro, and Y. Yang, 2007, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements: Geophysics Journal International, **169**, 1239–1269.

Chang, J., N. Nakata, R. Clapp, B. Biondi, and S. de Ridder, 2014, High frequency surface and body waves from ambient noise cross-correlations at Long Beach, CA: Expanded Abstracts of the 84th Ann. Internat. Mtg.

Daley, T., B. Freifeld, J. Ajo-Franklin, S. Dou, R. Pevzner, V. Shulakova, S. Kashikar, D. Miller, J. Goetz, J. Henninges, and S. Lueth, 2013, Field testing of fiber-optic distributed acoustic sensing (DAS) for subsurface seismic monitoring: The Leading Edge, **32**, 936–942.

de Ridder, S., 2014, Passive seismic surface-wave interferometry for reservoir-scale imaging: PhD thesis, Stanford University.

Lancelle, C., N. Lord, H. Wang, D. Fratta, R. Nigbor, A. Chalari, R. Karaulanov, J. Baldwin, and E. Castongia, 2014, Directivity and sensitivity of fiber-optic cable measuring ground motion using a distributed acoustic sensing array. Presented at AGU Fall Meeting.

Mateeva, A., J. Lopez, J. Mestayer, P. Wills, B. Cox, D. Kiyashchenko, Z. Yang, W. Berlang, R. Detomo, and S. Grandi, 2013, Distributed acoustic sensing for reservoir monitoring with VSP: The Leading Edge, **32**, 1278–1283.

Mateeva, A., J. Mestayer, B. Cox, D. Kiyashchenko, P. Wills, J. Lopez, S. Grandi, K. Hornman, P. Luments, A. Franzen, D. Hill, and J. Roy, 2012, Advances in distributed acoustic sensing (DAS) for VSP: Expanded Abstracts of the 82nd Ann. Internat. Mtg.

Wapenaar, K., E. Slob, and R. Snieder, 2006, Unified green's function retrieval by cross-correlation: Physical Review Letters, **97**, 234301.

Webster, P., J. Wall, C. Perkins, and M. Molenaar, 2013, Micro-seismic detection using distributed acoustic sensing: Expanded Abstracts of the 83rd Ann. Internat. Mtg.

Fast dispersion curves from ambient noise

Eileen Martin

ABSTRACT

To calculate dispersion curves, phase velocity vs. frequency, from ambient seismic noise, many researchers calculate virtual source responses in the time domain, then stack. This algorithm scales as $O(n^2)$, where n is the number of sensors. If the data are regularly sampled in space and time, an existing $O(n \log n)$ algorithm involves an F-K transform of the data cube followed by a transform into velocity vs frequency. I propose a new $O(n)$ algorithm which only takes the Fourier transform of the data in time. The new algorithm is conceptually simple, parallelizes easily, and does not require regular spacing between sensors. I show similar results for dispersion curves resulting from both the $O(n)$ algorithm and the $O(n^2)$ algorithm applied to data collected in a field trial of a trenched distributed acoustic sensing array at the Richmond Field Station. There are fewer opportunities for accumulating numerical error in the new $O(n)$ algorithm, so it yields a sharper image than the $O(n^2)$ algorithm.

INTRODUCTION

Dispersion curves are a simple way to synthesize ambient seismic noise data. Dispersion images roughly show how much surface wave energy is traveling at any given frequency and velocity. Dispersion curves are the result of picking curves along the peaks in frequency and velocity in these images. They sometimes allow us to see not only fundamental surface wave modes, but also higher order modes (de Ridder, 2014).

Let n be the number of sensors in a linear array that passively records seismic data. We are interested in continuous monitoring applications over large regions with many sensors (ideally, n should scale well into the 10,000's). In particular, scaling with the number of sensors of interest when processing ambient seismic noise data collected by a distributed acoustic sensing (DAS) array due to the dense sensor spacing. We assume ambient seismic noise is processed in small chunks (on the order of 1 minute with sampling on the order of 1 kHz).

We can reduce the communication and computation cost of calculating dispersion images (sometimes the peaks of these images are referred to as dispersion curves) from many sensors acting as virtual sources from $O(n^2)$ to $O(n)$ by solving this problem in the Fourier domain instead of the time domain. Following the derivation of this new algorithm, we show dispersion images on a small field dataset collected by a distributed acoustic sensing array.

First, we review some fundamental properties of the Fourier transform that are used in the derivation of the new $O(n)$ algorithm:

- The Fourier transform is linear, so

$$\mathcal{F}_t \sum_{r=0}^{n} u(x_r, t) = \sum_{r=0}^{n} \mathcal{F}_t u(x_r, t) = \sum_{r=0}^{n} \hat{u}(x_r, \omega).$$

I use the notation that $\hat{u}(\omega)$ is the Fourier transform of a function $u(t)$ throughout this report.

- A time shift is a frequency shift of the Fourier transform, i.e.

$$e^{2\pi i \omega p x} \hat{u}(x, \omega) = \mathcal{F}_t u(x, t + px)$$

- A cross-correlation in the time domain is a multiplication in the frequency domain. If we define cross-correlation as

$$d(x_1, t) \times d(x_2, t) = \int_{-\infty}^{\infty} d^*(x_1, \tau) d(x_2, \tau + t) d\tau,$$

then the Fourier transform of the cross-correlation is

$$\mathcal{F}_t \left(d(x_1, t) \times d(x_2, t) \right) = \hat{d}^*(x_1, \omega) \hat{d}(x_2, \omega)$$

where d^* denotes complex conjugation.

SUMMARY OF EXISTING ALGORITHMS

A common and conceptually simple algorithm to calculate dispersion curves requires calculating source responses in the time domain then slant-stacking, but this method scales as $O(n^2)$. A slightly less intuitive but more efficient $O(n \log n)$ algorithm involves a Fourier transform of the data in both space and time, followed by a transform into slowness vs. frequency. The new $O(n)$ algorithm is inspired by an $O(n^2)$ algorithm, so we describe this existing $O(n^2)$ in detail before moving to the derivation of the $O(n)$ algorithm.

Let $d(x_r, t)$ be an ambient seismic noise trace recorded for some time period at a sensor in position x_r, perhaps with some filtering and preprocessing applied (see Bensen et al. (2007) for general outline of preprocessing). Then let

$$u_s(x_r, t) = d(x_s, t) \times d(x_r, t)$$

be a cross-correlation that is some realization of a random variable with a mean that is the response to a virtual source placed at x_s (this is sometimes loosely referred to as a Green's function). Let p represent slowness (if v is the velocity then $p = 1/v$). One conceptually simple method for calculating the dispersion curve, c_s, from the response to virtual source, s, would be to calculate slant stacks along the responses of all sensors to the virtual source at x_s as is done in Chang (2013):

$$\begin{aligned} c_s(p, t) &= \sum_{r=1}^{n} u_s(x_r, t + p(x_r - x_s)) \\ \hat{c}_s(p, \omega) &= \mathcal{F}_t(c_s(p, t)) \end{aligned}$$

Clearly, this approach requires calculating $O(n)$ virtual source responses per virtual source, and must be carried out for $O(n)$ virtual sources to get an understanding of spatial variability across the array.

PROPOSED $O(N)$ FREQUENCY DOMAIN PROCESSING

We will use the basic facts about cross-correlations and the Fourier transform to rewrite the dispersion curve c_s for the response to a virtual source at x_s:

$$\begin{aligned}
\hat{c}_s(p,\omega) &= \mathcal{F}_t\left(\sum_{r=1}^{n} u_s(x_r, t + p(x_r - x_s))\right) \\
&= \sum_{r=1}^{n} \mathcal{F}_t(u_s(x_r, t + p(x_r - x_s))) \\
&= \sum_{r=1}^{n} \hat{u_s}(x_r,\omega) e^{2\pi i p(x_r - x_s)\omega} \\
&= \sum_{r=1}^{n} \hat{d}^*(x_s,\omega)\hat{d}(x_r,\omega) e^{2\pi i p(x_r - x_s)\omega} \\
&= \hat{d}^*(x_s,\omega) \sum_{r=1}^{n} \hat{d}(x_r,\omega) e^{2\pi i p(x_r - x_s)\omega} \\
&= \hat{d}^*(x_s,\omega) e^{-2\pi i p x_s \omega} \sum_{r=1}^{n} \hat{d}(x_r,\omega) e^{2\pi i p x_r \omega}
\end{aligned}$$

Let $\sigma := \sum_{r=1}^{n} \hat{d}(x_r,\omega) e^{2\pi i p x_r \omega}$. Clearly, σ can be reused in calculating the dispersion curves for all sources $\hat{c}_s(p,\omega)$, and it only takes $O(n)$ calculations. Also, these calculations are highly parallelizable over the number of sensors. Only a single round of communication is required for the reduction to calculate σ. A point-wise multiplication in the frequency domain must be calculated for each source, but that is a constant. Thus, we can calculate dispersion curves for n sensors in $O(n)$ time.

Algorithm 1 $O(n)$ algorithm:

Given a short chunk of traces $d(x_i, t)$ from receivers at $x_1, x_2, \ldots, x_n$
Initialize $\sigma(p,\omega) = 0$
Calculate σ, the sum of phase shifted data spectra
for $i = 1, \ldots, n$ **do**
  Despike and temporal normalization of $d(x_r, t)$
  Bandpass and whitening of $\hat{d}(x_r,\omega) = FFT(d(x_r, t))$
  $\sigma(p,\omega) += \hat{d}(x_r,\omega) e^{2\pi i p x_r \omega}$
end for
for $i = 1, \ldots, n$ **do**
  If stored, make filtered version of $\hat{d}(x_s,\omega)$ available. If not stored, do any filtering needed to calculate $\hat{d}(x_s,\omega)$
  Dispersion curve for virtual source at x_s is $\hat{c}_s(p,\omega) = \hat{d}^*(x_s,\omega) e^{-2\pi i p x_r \omega} \sigma(p,\omega)$
end for

FIELD DATA EXAMPLE

Using just ten minutes of ambient noise data, we were able to extract coherent virtual source responses from a trenched distributed acoustic sensing (DAS) array deployed at Richmond

Field Station. The channel length was 1 meter and the guage length was 10 meters. The preprocessing of traces is detailed in the companion paper Martin et al. (2015). Although we demonstrate this algorithm on data collected by a distributed acoustic sensing array, the algorithm can also be applied to traditional point sensors.

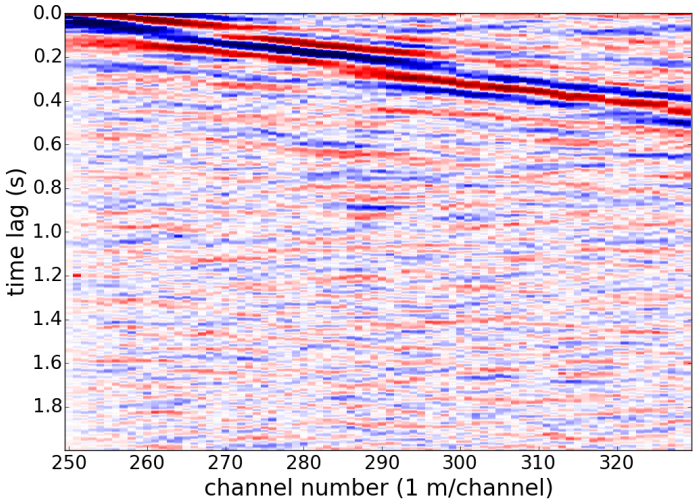


Figure 1: Estimated response to virtual source at channel 250 from 10 minutes of data filtered in the 5 to 50 Hz range. Virtual source response estimates such as this one must be calculated for each virtual source in the $O(n^2)$ algorithm, but can be avoided in the $O(n)$ algorithm. [**ER**] eileen2/. response

In the $O(n^2)$ algorithm, the first step is to calculate virtual source responses. We show these response estimates (folded over zero time lag) in Figure 1 for one virtual source at the end of the array based on only ten minutes of ambient data. The next step in the $O(n^2)$ algorithm is to stack the response estimates over a variety of time lags and velocities, as seen in Figure 2.

In the end, we are really interested in what the dispersion image looks like. For this data set we can see the dispersion image that only includes a virtual source at channel 250 in Figures 3(a) and 3(b). In the $O(n^2)$ algorithm, we can get this dispersion image by taking the FFT of the $\tau - p$ transform along the τ direction. However, in the $O(n)$ algorithm we can skip calculating the virtual source response estimates and the $\tau - p$ transforms, and directly calculate the dispersion image. Both algorithms yield two strong modes at approximately the same velocity and frequency.

There are some significant differences in these results due to numerical error despite the results being equivalent in theory for exact arithmetic. Switching between the time and frequency domains multiple times in the $O(n^2)$ algorithm leads to small nonzero values below 5 Hz, despite the bandpass filter, which should cut out energy below 5 Hz. However, the new $O(n)$ algorithm respects the bandpass filter. The new $O(n)$ algorithm results in a much sharper dispersion image, most likely because it has fewer Fourier transforms and fewer opportunities for numerical error to accumulate.

There are some features visible in the $O(n)$ algorithm's dispersion image that are not ap-

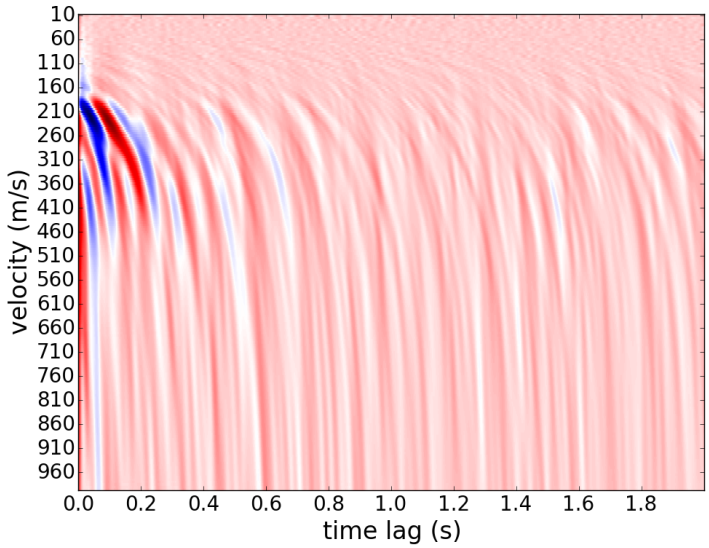


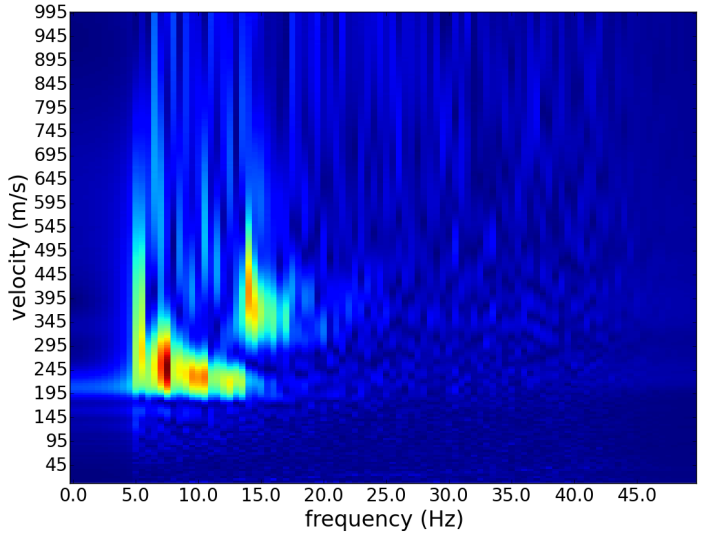
Figure 2: Velocity versus time lag resulting from $\tau - p$ transform of response estimate to virtual source at channel 250. This is a necessary step in the $O(n^2)$ algorithm, but can be avoided in the $O(n)$ algorithm. [**ER**] eileen2/. velLag

parent in the $O(n^2)$ algorithm's dispersion image. Between the two most readily apparent modes in the $O(n)$ algorithm's dispersion image, there appears to be another smaller peak for each frequency. It is possible this is another mode, but more data must be included before we can say anything conclusive.

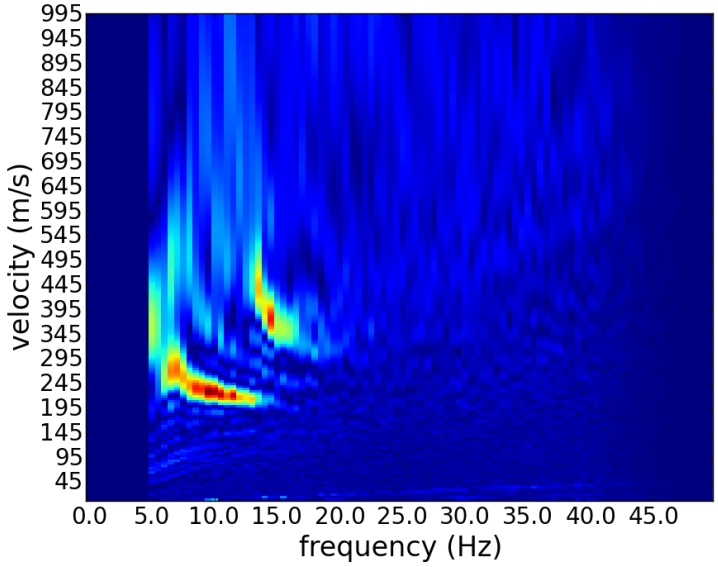
In the $O(n)$ algorithm's dispersion curve, there are some faint lines running from the low velocity & low frequency regime up to the high velocity & high frequency regime. The slope of these lines is approximately 10 meters. More data need to be incorporated before we can draw any more conclusions about whether this is a coherent feature, and what the interpretation of this feature could be.

CONCLUSIONS

We propose a fast algorithm to calculate dispersion images from the data spectra of receivers in a passive seismic survey. This new algorithm scales linearly with the number of sensors. As demonstrated on an ambient noise data set collected by a distributed acoustic sensing array, the new algorithm also yields sharper images. The dispersion images resulting from the new $O(n)$ algorithm are most likely sharper because the simplified algorithm requires fewer operations that may accumulate numerical errors and spread them over the dispersion image space.



(a)



(b)

Figure 3: (Top) Velocity versus frequency dispersion image calculated by the $O(n^2)$ algorithm which takes an FFT of the $\tau - p$ transform in the τ direction. (Bottom) Dispersion image calculated by the $O(n)$ algorithm directly from the data spectra, then binned into 1 Hz intervals. Both plots show the two strongest modes at approximately the same velocity and frequency, but the $O(n)$ algorithm results in a much sharper image. [**ER**] eileen2/. dispCurve,dispCurveFrq

ACKNOWLEDGEMENTS

I would like to thank Jason Chang, Bob Clapp, and Sjoerd de Ridder for helpful conversations about existing methods to calculate dispserion curves and this algorithm. I would also like to thank my collaborators Jonathan Ajo-Franklin, Nate Lindsey, Shan Dou, Tom Daley, Barry Freifeld, Michelle Robertson, Craig Ulrich, and Anna Wagner, who collected the data with the support of the US Department of Defense under SERDP grant RC-2437 "Developing Smart Infrastructure for a Changing Arctic Environment Using Distributed Fiber-Optic Sensing Methods". I am supported by the Department of Energy Computational Science Graduate Fellowship, provided under grant number DE-FG02-97ER25308.

REFERENCES

Bensen, G., M. Ritzwoller, M. Barmin, A. Levshin, F. Lin, M. Moschetti, N. Shapiro, and Y. Yang, 2007, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements: Geophysics Journal International, **169**, 1239–1269.

Chang, J. P., 2013, Velocity dispersion at Long Beach: SEP-Report, **150**, 101–108.

de Ridder, S., 2014, Passive seismic surface-wave interferometry for reservoir-scale imaging: PhD thesis, Stanford University.

Martin, E., J. Ajo-Franklin, N. Lindsey, T. Daley, B. Freifeld, M. Robertson, C. Ulrich, S. Dou, and A. Wagner, 2015, Applying interferometry to ambient seismic noise recorded by a trenched distributed acoustic sensing array: SEP-Report, **158**, 247–254.

Double-difference time-lapse FWI with a total-variation regularization

Musa Maharramov, Yinbin Ma and Biondo Biondi

ABSTRACT

In this paper we study double-difference FWI with a total-variation (TV) model-difference regularization (Maharramov and Biondi, 2014c). This data-space technique reduces the sensitivity of time-lapse FWI to inaccurate velocity model reconstruction. We describe a computational framework for conducting a TV-regularized double-difference FWI both as a simultaneous inversion and as an extension of single-model inversion. The method is demonstrated on linearized time-lapse waveform inversion of production effects for a synthetic example of compacting sub-salt reservoirs. We demonstrate the resolution of production effects and discuss stability of the results with respect to inaccuracies in the background velocity model.

INTRODUCTION

Simultaneous time-lapse (4D) full-waveform inversion (FWI) with a total-variation (TV) regularization achieved a considerable success in resolving production-induced changes both in synthetic and field-data tests (Maharramov and Biondi (2015b, 2014c, 2015a), and Maharramov and Biondi (2014a), supplementary material). As a model-space technique, the simultaneous inversion is stable with respect to repeatability issues such as different acquisition parameters, while the conventional double-difference method (Watanabe et al., 2004; Denli and Huang, 2009; Zheng et al., 2011; Asnaashari et al., 2012; Raknes et al., 2013) is quite sensitive to survey repeatability and may require a significant preprocessing effort to achieve data equalization (Maharramov and Biondi, 2014b,a). However, one important potential advantage of the double-difference FWI is that once the baseline and monitor data sets are equalized and measurable production effects can be observed and isolated in the data difference, the double-difference FWI seeks to resolve the model difference by matching the data difference only, rather than matching the separate acquisition data sets. This suggests the double-difference FWI as a potentially useful technique in situations where resolution of the baseline model is still subject to considerable uncertainty but the production-induced effects in the data are significant enough to estimate model perturbations that are causing them. Maharramov and Biondi (2014c) proposed a formulation of the double-difference method that allows matching the observed data difference by simultaneously inverting subsurface models of different vintage while imposing a blockiness-promoting total-variation

regularization on the model difference:

$$\alpha\|\mathbf{F}(\mathbf{m}_b) - \mathbf{d}_b\|_2^2 \; + \beta\|\mathbf{F}(\mathbf{m}_m) - \mathbf{d}_m\|_2^2 \; + \tag{1}$$

$$\gamma\| \left(\mathbf{F}(\mathbf{m}_m) - \mathbf{F}(\mathbf{m}_b)\right) - \left(\mathbf{d}_m - \mathbf{d}_b\right) \|^2 \; + \tag{2}$$

$$\alpha_1\|\mathbf{W}_b\mathbf{R}_b(\mathbf{m}_b - \mathbf{m}_b^{\mathrm{PRIOR}})\|_2^2 \; + \tag{3}$$

$$\beta_1\|\mathbf{W}_m\mathbf{R}_m(\mathbf{m}_m - \mathbf{m}_m^{\mathrm{PRIOR}})\|_2^2 \; + \tag{4}$$

$$\delta\|\mathbf{WR}(\mathbf{m}_m - \mathbf{m}_b - \Delta\mathbf{m}^{\mathrm{PRIOR}})\|_1 \; \rightarrow \; \min, \tag{5}$$

In the above equations, subscripts b and m denote baseline and monitor acquisitions, $\mathbf{d}_{b,m}$ is a vector of observations (survey data), $\mathbf{m}_{b,m}$ are the baseline an monitor models, $\mathbf{F}$ is the forward-modeling operator, $\mathbf{W}_{b,m}, \mathbf{W}$ and $\mathbf{R}_{b,m}, \mathbf{R}$ are weighting and regularization operators for the baseline, monitor, and the model difference. Model and model-difference priors can be explicitly specified in the objective function as shown in (3,4,5), however, in this work we assume no prior information. For a TV-regularized model-difference inversion, we use the model-difference regularization operator $\mathbf{R}$ such that

$$\mathbf{R}f(x,y,z) \; = \; |\nabla f|, \tag{6}$$

i.e., $\mathbf{R}$ computes the spatial gradient of its argument function at each point of the subsurface. This means that we seek a monitor model $\mathbf{m}_m$ that differs from the inverted baseline by a spatially bounded or *blocky* component (Rudin et al., 1992). The assumption of blockiness and, more generally, spatial boundedness of production-induced model perturbations is consistent with the physical effects of fluid substitution, reservoir compaction and overburden dilation (Johnston, 2013). An example of a TV-regularized simultaneous FWI applied to estimating spatially localized production-induced overburden dilation from Gulf of Mexico time-lapse data is provided by Maharramov and Biondi (2015a) in this report. In this work, we assess the feasibility of TV-regularized double differencing with the terms (1,3,4) omitted, i.e., TV-regularized double-differencing without simultaneously fitting data of different vintage. In our initial tests we consider a linearized formulation with adaptive sparsity-promoting steering-filter regularization (Ma et al., 2015a), and compare the results with the simultaneous linearized inversion of Ma et al. (2015b). Of particular interest to us is the effect on the two methods of inaccuracies in the baseline model, and whether matching the data difference only can achieve a greater robustness with respect to the uncertainty in the background model.

METHOD

We consider the following optimization problem

$$\| \left(\mathbf{F}(\mathbf{m}_m) - \mathbf{F}(\mathbf{m}_b)\right) - \left(\mathbf{d}_m - \mathbf{d}_b\right) \|^2 \; + \tag{7}$$

$$\delta\|\mathbf{WR}(\mathbf{m}_m - \mathbf{m}_b)\|_1 \; \rightarrow \; \min, \tag{8}$$

with operator $\mathbf{R}$ given by (6). The regularized optimization problem (7,8) can be solved with respect to the combined baseline/monitor model vector $(\mathbf{m}_b, \mathbf{m}_m)$ (i.e., by inverting the two models simultaneously) or using the traditional double-difference approach by e.g. fixing the baseline model $\mathbf{m}_b$ and minimizing with respect to the monitor model $\mathbf{m}_m$ (Watanabe et al., 2004; Denli and Huang, 2009; Asnaashari et al., 2012; Zheng et al., 2011; Maharramov and Biondi, 2014b). The latter approach reduces the size of optimization problem and

alleviates the null-space issues associated with resolving two subsurface models from the data difference alone. Note that the latter issue can be remedied by adding back the terms (1)—i.e., effectively combining the double-differencing with a simultaneous baseline and monitor model inversion. However this combined approach is outside the scope of this work.

Our method can be summarized as follows:

1) Invert the baseline model $\mathbf{m}_b$:

$$\|\mathbf{F}(\mathbf{m}_b) - \mathbf{d}_b\|_2^2 \;\rightarrow\; \min . \tag{9}$$

2) Generate new synthetic monitor survey data $\mathbf{d}_2$ by adding the observed data difference $\mathbf{d}_m - \mathbf{d}_b$ to the forward-modeled baseline data $\mathbf{F}(\mathbf{m}_b)$:

$$\mathbf{d}_2 \;=\; \mathbf{F}(\mathbf{m}_b) \;+\; (\mathbf{d}_m \;- \mathbf{d}_b) , \tag{10}$$

3) Invert the monitor model $\mathbf{m}_m$ from the new synthetic data $\mathbf{d}_2$:

$$\| \left(\mathbf{F}(\mathbf{m}_m) - \mathbf{d}_2\right) \|_2^2 + \tag{11}$$

$$\delta \|\mathbf{WR}(\mathbf{m}_m - \mathbf{m}_b)\|_2^2 \;\rightarrow\; \min . \tag{12}$$

Method (9,10,11,12) can be used with both 4D FWI and the linearized waveform inversion. Note that in the latter case, because the forward-modeling operator $\mathbf{F}$ is linear, the above procedure is the only correct approach to solving (7,8) in the absence of extra constraints, as the linearized inversion problem has a null space dimension of (at least) the subsurface model: adding the same perturbation to both the baseline and monitor models does not affect the data difference.

Because for linearized inversion the inverted models are qualitatively interpreted as reflectivity, the total-variation regularization operator (6) should be replaced in this case with an operator that promotes blockiness of the reflectivity only along reflector dips while enforcing sparsity in the orthogonal direction (Ma et al., 2015a). This is achieved by replacing the full gradient in (6) with a directional gradient

$$\mathbf{R}f(x,y,z) \;=\; |\nabla_{\boldsymbol{\xi}} f|, \tag{13}$$

where $\boldsymbol{\xi} = \boldsymbol{\xi}(x,y,z)$ is the dip direction at (x,y,z) that is updated at each iteration of (11,12). This approach is effectively equivalent to promoting blockiness only along reflector surfaces, and therefore we will call this approach *Steering TV* (STV) regularization.

RESULTS

We applied the proposed method to a sub-salt time-lapse reflectivity inversion problem studied by Ma et al. (2015b) in this report. Synthetic baseline and monitor acoustic velocity models are shown in Figures 1(a) and 1(b), respectively. The monitor model has been chosen to simulate the effects of gas substitution with water (a higher velocity below the reservoir top) and overburden dilation due to reservoir compaction (lower velocity above the reservoir top) at two locations below and on the left side of the salt body. The time-lapse

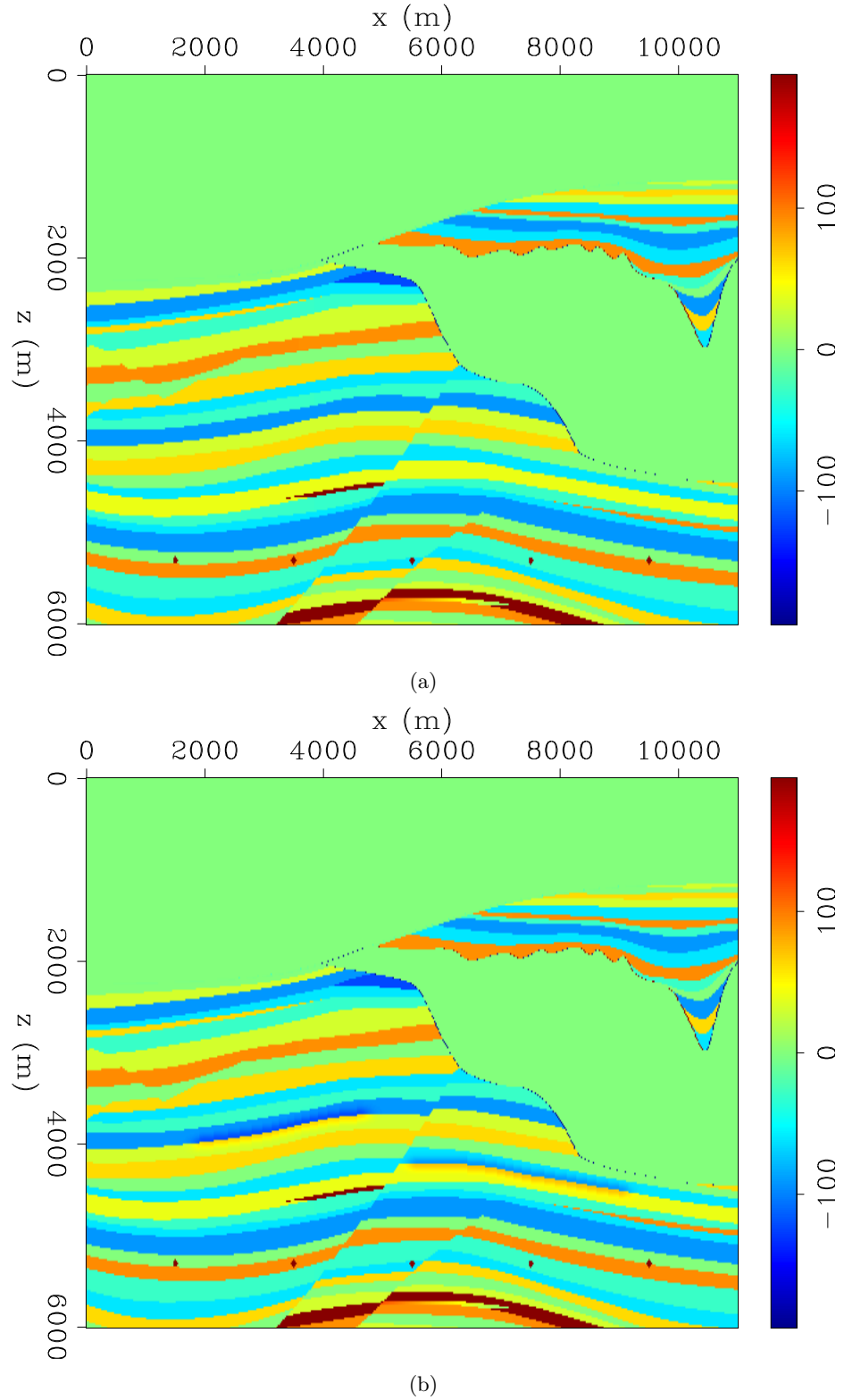


Figure 1: (a) Baseline velocity model. (b) Monitor velocity model simulating production-induced fluid substitution effects (positive change) and overburden dilation effects (negative change) for two reservoirs. [**ER**] musa4/. simple.diffT0,simple.diffT2

linearized waveform inversion seeks to recover image difference between monitor and baseline migrations, typically using the same (baseline) background velocity model for both baseline and monitor images. The result of applying the simultaneous STV-regularized linearized inversion (Ma et al., 2015b) is shown in Figure 4(b). The corresponding STV-regularized double-difference inversion is shown in Figure 2(b). In both cases a target-oriented inversion was conducted within a target window shown in the figures, using the exact background velocity model matching the baseline velocity model. Both methods are expected to achieve similar results because under the assumption that problem (9) is solved exactly, the two methods are mathematically equivalent. Production-induced reflectivity changes for both reservoirs stand out prominently in both images. The double-difference result was obtained using identical acquisition geometries to match the effect of data equalization that is part of standard time-lapse processing (Maharramov and Biondi, 2014b). However, when the background model is inaccurate, the simultaneous inversion and double-difference solve two different problems: the first method seeks to match two different sets of reflection data using the wrong background velocity, while the regularized double difference seeks to match the observed *relative* data difference by perturbing the predicted (and inaccurate) baseline reflectivity model. Inversion results for the two methods using a 10% overestimated velocity model in the target zone are shown in Figures 3(a) and 3(b). While the wrong background velocity still results in a slight mispositioning of the target reflectors, the double-difference reflectivity inversion appears to be more robust with respect to inaccurate velocity. The double-difference inversion is still able to resolve the reflectivity change along two isolated reflectors corresponding to the two reservoirs, but the simultaneous inversion result using the wrong background velocity is contaminated with artifacts that can be misinterpreted as production effects (e.g. the artifacts marked with red circles in Figure 3(a))

CONCLUSIONS AND PERSPECTIVES

Double differencing with steering TV (STV) regularization may yield more robust inversion of reflectivity changes in the presence of velocity uncertainty. Better imaging of reflectivity changes due to fluid-substitution effects leads to improved infill strategies and reservoir monitoring, and therefore is of paramount importance for reservoir production management. While linearized waveform inversion presents a useful initial application of the regularized double-differencing technique, application to full-waveform inversion, especially in a hybrid approach involving simultaneous inversion of the baseline and monitor data-fitting terms (1,2), requires further analysis and will be subject of our future work.

ACKNOWLEDGMENTS

The authors would like to thank Stewart Levin for a number of useful discussions.

REFERENCES

Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, 2012, Time-lapse imaging using regularized FWI: A robustness study: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi:10.1190/segam2012-0699.1, 1–5.

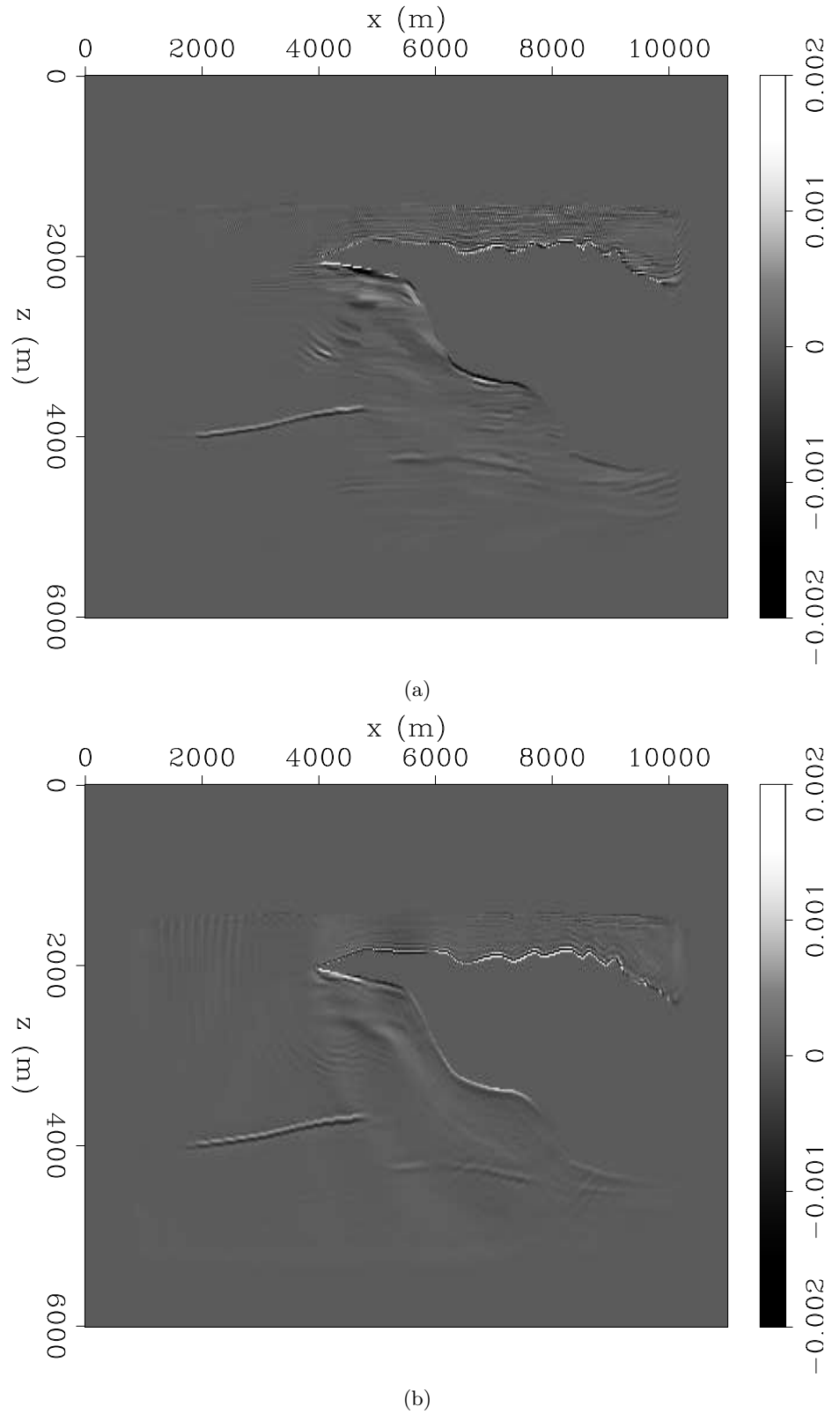


Figure 2: (a) Steering TV-regularized simultaneous linearized inversion (Ma et al., 2015b). (b) Steering TV-regularized double-difference method (9-12). In both cases the exact baseline velocity and reflectivity models were used. Both methods resolved reflectivity changes for the two reservoirs. [**CR**] musa4/. simple.IRLSRTMDIFF,DDexactv

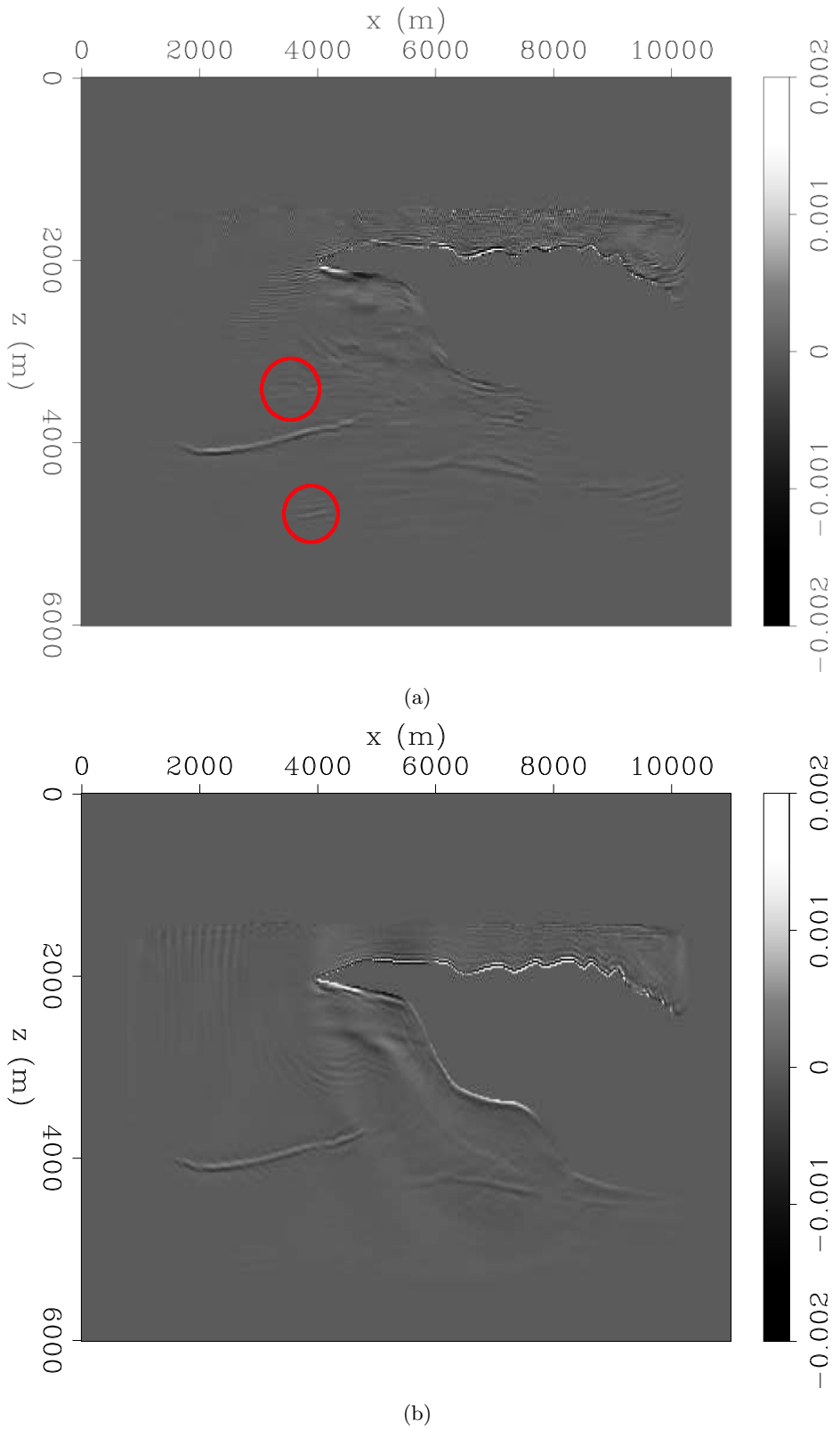


Figure 3: (a) Steering TV-regularized simultaneous linearized inversion using a 10% overestimated velocity model. (b) Steering TV-regularized double-difference method (9-12). The simultaneous inversion result is now contaminated with artifacts (e.g. marked with red circles) that are absent from the double-difference result. [**CR**] musa4/. marked,DDwrongv

Denli, H. and L. Huang, 2009, Double-difference elastic waveform tomography in the time domain: 79th Annual International Meeting, SEG, Expanded Abstracts, 2302–2306.

Johnston, D., 2013, Practical applications of time-lapse seismic data: Society of Exploration Geophysicists.

Ma, Y., M. Maharramov, R. Clapp, and B. Biondi, 2015a, Illumination compensation by L1 regularization and steering filters: SEP Report, **158**.

———, 2015b, Time lapse seismic imaging with L1 regularization and steering filters: SEP Report, **158**.

Maharramov, M. and B. Biondi, 2014a, Joint full-waveform inversion of time-lapse seismic data sets: 84th Annual Meeting, SEG, Expanded Abstracts, 954–959.

———, 2014b, Joint full-waveform inversion of time-lapse seismic data sets: SEP Report, **152**, 19–28.

———, 2014c, Robust joint full-waveform inversion of time-lapse seismic datasets with total-variation regularization: SEP Report, **155**, 199–208.

———, 2015a, Resolving the effects of production-induced overburden dilation using simultaneous TV-regularized time-lapse FWI: SEP Report, **158**.

———, 2015b, Robust simultaneous time-lapse full-waveform inversion with total-variation regularization of model difference: 77th EAGE Conference and Exhibition, Extended Abstract (accepted).

Raknes, E., W. Weibull, and B. Arntsen, 2013, Time-lapse full waveform inversion: Synthetic and real data examples: 83rd Annual International Meeting, SEG, Expanded Abstracts, 944–948.

Rudin, L. I., S. Osher, and E. Fatemi, 1992, Nonlinear total variation based noise removal algorithms: Physica D: Nonlinear Phenomena, **60**, 259–268.

Watanabe, T., S. Shimizu, E. Asakawa, and T. Matsuoka, 2004, Differential waveform tomography for time-lapse crosswell seismic data with application to gas hydrate production monitoring: 74th Annual International Meeting, SEG, Expanded Abstracts, 2323–2326.

Zheng, Y., P. Barton, and S. Singh, 2011, Strategies for elastic full waveform inversion of timelapse ocean bottom cable (OBC) seismic data: 81st Annual International Meeting, SEG, Expanded Abstracts, 4195–4200.

Total-variation minimization with bound constraints

Musa Maharramov and Stewart A. Levin

ABSTRACT

We present a powerful and easy-to-implement algorithm for solving constrained optimization problems that involve L_1/total-variation regularization terms, and both equality and inequality constraints. We discuss the relationship of our method to earlier works of Goldstein and Osher (2009) and Chartrand and Wohlberg (2010), and demonstrate that our approach is a combination of the augmented Lagrangian method with splitting and model projection. We test the method on a geomechanical problem and invert highly compartmentalized pressure change from noisy surface uplift observations. We conclude the paper with a discussion of possible extension to a wide class of regularized optimization problems with bound and equality constraints.

INTRODUCTION

The primary focus of this work is a class of least-squares fitting problems with a total-variation (TV) regularization and bound model constraints:

$$\begin{aligned} &\||\nabla \mathbf{m}|\|_1 \;+\; \frac{\alpha}{2}\|\mathbf{F}(\mathbf{m})-\mathbf{d}\|_2^2 \;\rightarrow\; \min, \\ &\mathbf{m}_1 \;\leq\; \mathbf{m} \;\leq\; \mathbf{m}_2. \end{aligned} \tag{1}$$

In (1) we seek a model vector $\mathbf{m}$ such that forward-modeled data $\mathbf{F}(\mathbf{m})$ match observed data $\mathbf{d}$ in the least squares sense, while imposing blockiness-promoting total-variation (TV) regularization (Rudin et al., 1992) and lower ($\mathbf{m}_1$) and upper ($\mathbf{m}_2$) model bounds. Rather than using a regularization parameter as a coefficient of the regularization term, we use a data-fitting weight α. TV regularization (also know as the Rudin-Osher-Fatemi, or ROF, model) acts as a form of model styling that helps to preserve sharp contrasts and boundaries in the model, even when spectral content of input data has limited resolution. Examples of successful geophysical application of unconstrained TV-regularized optimization are included in this report (Maharramov and Biondi, 2015; Maharramov et al., 2015; Ma et al., 2015a,b). The regularization provided by bounded total-variation sometimes produces sufficient smoothing side-effect on the inverted model that obviates explicit bound constraints. However, many applications still require the imposition of additional constraints regardless of the regularization. For example, reservoir pore-pressure inversion problems often come with *a priori* bounds on the estimated pore pressure change, such as the pore pressure change being non-negative as a result of fluid injection (lower bound) or never exceeding a hydraulic fracturing pressure (upper bound). TV regularization is a key tool in imaging and de-noising applications (Rudin et al., 1992; Chambolle and Lions, 1997; Goldstein and Osher, 2009; Chartrand and Wohlberg, 2010) and require an efficient mechanism for including *a priori* model constraints that can significantly reduce model space (Chartrand and Wohlberg, 2010). While barrier or penalty function methods, such as nonlinear

interior-point methods (Nocedal and Wright, 2006), can be used to tackle the general constrained formulation (1), the presence of a non-differentiable L_1-norm total-variation term and non-quadratic penalty terms pose considerable challenges to practical implementation. A log-barrier function such as

$$-\text{const} \times \sum_{i=1}^{n} \log \frac{m_2^i - m^i}{\delta} + \log \frac{m^i - m_1^i}{\delta}, \tag{2}$$

where n is the model space dimension, can be added to the right-hand side of the objective function to keep solution iterates away from the rectangular bounds. However, this adds a non-quadratic term to the objective function. For large-scale inversion problems with $n > 10^5$ (such as typical in geophysical applications) often only iterative gradient-based solution techniques like the nonlinear conjugate gradients (Nocedal and Wright, 2006) are available, and adding non-quadratic terms may significantly affect convergence properties. Note that this is in addition to the challenges associated with handling the non-differentiable TV-regularization term.

Chartrand and Wohlberg (2010) used a splitting approach to decouple the TV-regularized problem from enforcing the constraints. In our approach, we perform three-way splitting of problem (1) into a smooth optimization, gradient thresholding and projection steps using the Alternating Direction Method of Multipliers (ADMM) (Boyd et al., 2010). For unconstrained TV-regularized problems this approach is equivalent to the split-Bregman method of Goldstein and Osher (2009). However, we integrate the projection step associated with enforcing the bound constraints into the TV-minimization loop and avoid unnecessary iterations in the minimization of a proximal term (Parikh and Boyd, 2013) associated with the projection.

METHOD

First, we recast the TV-regularization part of (1) as a constrained optimization problem following the approach of Goldstein and Osher (2009). We introduce an auxiliary variable $\mathbf{x}$ and operator $\mathbf{\Phi} : \mathbf{m} \rightarrow \mathbf{x}$ such that for isotropic TV regularization we have a vector of the model-space dimension

$$\mathbf{\Phi}(\mathbf{m}) \;=\; \sqrt{(\nabla_x \mathbf{m})^2 + (\nabla_y \mathbf{m})^2}, \tag{3}$$

and for anisotropic regularization a vector twice the model-space dimension

$$\mathbf{\Phi}(\mathbf{m}) \;=\; \begin{bmatrix} \nabla_x \mathbf{m} \\ \nabla_y \mathbf{m} \end{bmatrix}. \tag{4}$$

Problem (1) can now be reformulated with an additional equality constraint:

$$\begin{aligned} &\|\mathbf{x}\|_1 \;+\; \frac{\alpha}{2}\|\mathbf{F}(\mathbf{m}) - \mathbf{d}\|_2^2 \;\rightarrow\; \min, \\ &\mathbf{x} \;=\; \mathbf{\Phi}(\mathbf{m}), \\ &\mathbf{m}_1 \;\leq\; \mathbf{m} \;\leq\; \mathbf{m}_2. \end{aligned} \tag{5}$$

Problem (5) is still a bound-constrained problem. Introducing the projection operator

$$\mathbf{\Pi}(\mathbf{m}) \;=\; \max\{\min\{\mathbf{m}, \mathbf{m}_2\}, \mathbf{m}_1\}, \tag{6}$$

where min and max are applied component-wise, we reduce (5) to a fully equality-constrained formulation:

$$\begin{aligned} &\|\mathbf{x}\|_1 \;+\; \frac{\alpha}{2}\|\mathbf{F}(\mathbf{m})-\mathbf{d}\|_2^2 \;\rightarrow\; \min, \\ &\mathbf{x} \;=\; \mathbf{\Phi}(\mathbf{m}), \\ &\mathbf{m} \;=\; \mathbf{y}, \\ &\mathbf{y} = \mathbf{\Pi}(\mathbf{m}). \end{aligned} \tag{7}$$

Following the augmented Lagrangian recipe for (7) while assuming the last equality constraint still enforced explicitly, we obtain a sequence of problems (Nocedal and Wright, 2006)

$$\begin{aligned} &(\mathbf{x}^{k+1},\mathbf{m}^{k+1}) \;=\; \operatorname{argmin}\; \|\mathbf{x}\|_1 \;+\; \frac{\alpha}{2}\|\mathbf{F}(\mathbf{m})-\mathbf{d}\|_2^2 \;+ \\ &\frac{\lambda}{2}\|\mathbf{x} \;-\; \mathbf{\Phi}(\mathbf{m})\|_2^2 - \boldsymbol{\mu}_k{}^T\left(\mathbf{x} \;-\; \mathbf{\Phi}(\mathbf{m})\right) + \frac{\delta}{2}\|\mathbf{m}-\mathbf{y}\|_2^2 - \boldsymbol{\nu}_k{}^T\left(\mathbf{m}-\mathbf{y}\right) \;\rightarrow\; \min, \\ &\boldsymbol{\mu}_{k+1} \;=\; \boldsymbol{\mu}_k - \lambda\left[\mathbf{x}^{k+1} - \mathbf{\Phi}(\mathbf{m}^{k+1})\right], \\ &\boldsymbol{\nu}_{k+1} \;=\; \boldsymbol{\nu}_k - \delta\left[\mathbf{m}^{k+1} - \mathbf{y}\right],\; k = 0,1,2,\ldots \end{aligned} \tag{8}$$

Coefficients λ and δ are any positive constants above certain problem-specific "threshold" values (Nocedal and Wright, 2006), and can be selected experimentally. Vectors $\boldsymbol{\mu}_k$ and $\boldsymbol{\nu}_k$ are vectors of multipliers that converge to the set of Lagrange multipliers for the first two equality constraints of problem (7). At each step, (8) solves an L_1-regularized problem with respect to the combined model vector $(\mathbf{x},\mathbf{m})$. Introducing new scaled multiplier vectors

$$\mathbf{b}^k \;=\; \frac{\boldsymbol{\mu}_k}{\lambda},\; \mathbf{c}^k \;=\; \frac{\boldsymbol{\nu}_k}{\delta},\; k = 0,1,2,\ldots \tag{9}$$

a little algebra shows that (8) is equivalent to

$$\begin{aligned} &(\mathbf{x}^{k+1},\mathbf{m}^{k+1}) \;=\; \operatorname{argmin}\; \|\mathbf{x}\|_1 \;+\; \frac{\alpha}{2}\|\mathbf{F}(\mathbf{m})-\mathbf{d}\|_2^2 \;+ \\ &\frac{\lambda}{2}\|\mathbf{x} \;-\; \mathbf{\Phi}(\mathbf{m}) - \mathbf{b}^k\|_2^2 + \frac{\delta}{2}\|\mathbf{m} \;-\; \mathbf{y} - \mathbf{c}^k\|_2^2 \;\rightarrow\; \min, \\ &\mathbf{b}^{k+1} \;=\; \mathbf{b}^k + \Phi(\mathbf{m}^{k+1}) - \mathbf{x}^{k+1}, \\ &\mathbf{c}^{k+1} \;=\; \mathbf{c}^k + \mathbf{y} - \mathbf{m}^{k+1},\; k = 0,1,2,\ldots \end{aligned} \tag{10}$$

Here we used the fact that adding a constant term $\lambda/2\|\mathbf{b}^k\|_2^2 + \delta/2\|\mathbf{c}^k\|_2^2$ to the objective function obviously does not change the minimizing solution.

Problem (7) can be solved by iteratively projecting the current model vector $\mathbf{m}$ onto $\mathbf{y}$, then conducting the iterations (10) to convergence, then repeating the process. However, presence of the proximal term $\delta/2\|\mathbf{m}-\mathbf{y}\|_2^2$ in (8) due to the constraint $\mathbf{m}=\mathbf{y}$ means that a very accurate solution of (10) at early iterations is wasteful and unnecessary. We instead carry out a single iteration of (10) followed by the model projection:

$$\begin{aligned} &(\mathbf{x}^{k+1},\mathbf{m}^{k+1}) \;=\; \operatorname{argmin}\; \|\mathbf{x}\|_1 \;+\; \frac{\alpha}{2}\|\mathbf{F}(\mathbf{m})-\mathbf{d}\|_2^2 \;+ \\ &\frac{\lambda}{2}\|\mathbf{x} \;-\; \mathbf{\Phi}(\mathbf{m}) - \mathbf{b}^k\|_2^2 + \frac{\delta}{2}\|\mathbf{m} \;-\; \mathbf{y}^k - \mathbf{c}^k\|_2^2 \;\rightarrow\; \min, \\ &\mathbf{b}^{k+1} \;=\; \mathbf{b}^k + \Phi(\mathbf{m}^{k+1}) - \mathbf{x}^{k+1}, \\ &\mathbf{c}^{k+1} \;=\; \mathbf{c}^k + \mathbf{y}^k - \mathbf{m}^{k+1}, \\ &\mathbf{y}^{k+1} \;=\; \mathbf{\Pi}(\mathbf{m}^{k+1}) \;=\; \max\{\min\{\mathbf{m}^{k+1},\mathbf{m}_2\},\mathbf{m}_1\},\; k = 0,1,2,\ldots \end{aligned} \tag{11}$$

The iterative process (11) still requires soling an L_1-regularized problem. However, the L_1-norm term now involves only the vector $\mathbf{x}$. Therefore, we can apply Douglas-Rachford splitting, minimizing

$$\|\mathbf{x}\|_1 \ + \ \frac{\alpha}{2}\|\mathbf{F}(\mathbf{m})-\mathbf{d}\|_2^2 \ + \frac{\lambda}{2}\|\mathbf{x} \ - \ \mathbf{\Phi}(\mathbf{m})-\mathbf{b}^k\|_2^2 + \frac{\delta}{2}\|\mathbf{m} \ - \ \mathbf{y}^k-\mathbf{c}^k\|_2^2 \tag{12}$$

alternately with respect to $\mathbf{m}$ and $\mathbf{x}$ in an inner loop of $N_1 \geq 1$ cycles. Because the proximal constraint $\mathbf{m}=\mathbf{y}$ renders good fitting accuracy at early stages unnecessary, N_1 can be small. Further we note that the minimization of (12) with respect to $\mathbf{x}$ (in a splitting step with $\mathbf{m}$ fixed) is given trivially by the "shrinkage" operator (Goldstein and Osher, 2009):

$$\mathbf{x}^{k+1} \ = \ \text{shrink}\left\{\mathbf{\Phi}(\mathbf{m})+\mathbf{b}^k, \frac{1}{\lambda}\right\}, \tag{13}$$

where

$$\text{shrink}\left\{\mathbf{x},\gamma\right\} \ = \ \frac{\mathbf{x}}{|\mathbf{x}|}\max\left(|\mathbf{x}|-\gamma,0\right), \tag{14}$$

and is effectively thresholding the model gradient. Our algorithm can be described by the following 5 steps:

1 Initialization

$$\begin{aligned}
\mathbf{m}^0 \ &= \ \text{starting guess},\\
\mathbf{x}^0 \ &= \ \mathbf{0},\\
\mathbf{y}^0 \ &= \ \max\{\min\{\mathbf{m}^0,\mathbf{m}_2\},\mathbf{m}_1\},\\
\mathbf{b}^0 \ &= \ \mathbf{0},\\
\boldsymbol{c}^0 \ &= \ \mathbf{0},
\end{aligned} \tag{15}$$

2 **Outer loop**. Repeat steps 3-5 for $k=0,1,2,\ldots$

3 **Inner loop**. Iterate (16) $N_1 \geq 1$ times.

$$\begin{aligned}
\mathbf{m}^{k+1} \ &= \text{argmin} \ \frac{\lambda}{2}\|\mathbf{x}^k-\mathbf{\Phi}(\mathbf{m})-\mathbf{b}^k\|_2^2+\frac{\alpha}{2}\|\mathbf{F}(\mathbf{m})-\mathbf{d}\|_2^2+\\
&\quad \frac{\delta}{2}\|\mathbf{m}-\mathbf{y}^k-\boldsymbol{c}^k\|_2^2,\\
\mathbf{x}^{k+1} \ &= \text{shrink}\left\{\mathbf{\Phi}(\mathbf{m}^{k+1})+\mathbf{b}^k,\frac{1}{\lambda}\right\}, \ \mathbf{x}^k \ = \ \mathbf{x}^{k+1},
\end{aligned} \tag{16}$$

4 Update the multipliers and project the model onto the bounding rectangle:

$$\begin{aligned}
\mathbf{b}^{k+1} \ &= \ \mathbf{b}^k+\mathbf{\Phi}(\mathbf{m}^{k+1})-\mathbf{x}^{k+1},\\
\boldsymbol{c}^{k+1} \ &= \ \boldsymbol{c}^k+\mathbf{y}^k-\mathbf{m}^{k+1},\\
\mathbf{y}^{k+1} \ &= \ \max\{\min\{\mathbf{m}^{k+1},\mathbf{m}_2\},\mathbf{m}_1\}.
\end{aligned} \tag{17}$$

5 Terminate if the target accuracy is reached

$$\frac{\|\mathbf{m}^{k+1}-\mathbf{m}^k\|_2}{\|\mathbf{m}^k\|} \ \leq \ \text{target accuracy}. \tag{18}$$

or go back to step 2 otherwise.

Optimizing (16) with respect to $\mathbf{m}$ is in itself a large-scale optimization problem, nonlinear for a nonlinear modeling operator $\mathbf{F}$. Solving the optimization problem (16) exactly is unnecessary because for small k (i.e., at early stages of the inversion) vector $\mathbf{y}^k$ is not the true model, vector $\mathbf{x}^k$ is far from the true model gradient, and the multipliers $\mathbf{b}^k, \mathbf{x}^k$ could be far from scaled Lagrange multipliers, s. Therefore, for large-scale problems only a few steps of an iterative method like conjugate gradients need be carried out. As the solution converges to the true solution and critical sharp contrasts in the model are identified, an iterative solver can begin to take advantage of the objective function curvature information collected at previous iterations of the *outer* loop, potentially leading to a significantly faster convergence. Optimal strategies for spanning iterations of nonlinear conjugate gradients across iterations of the *outer* loop of our algorithm are the subject of an upcoming report.

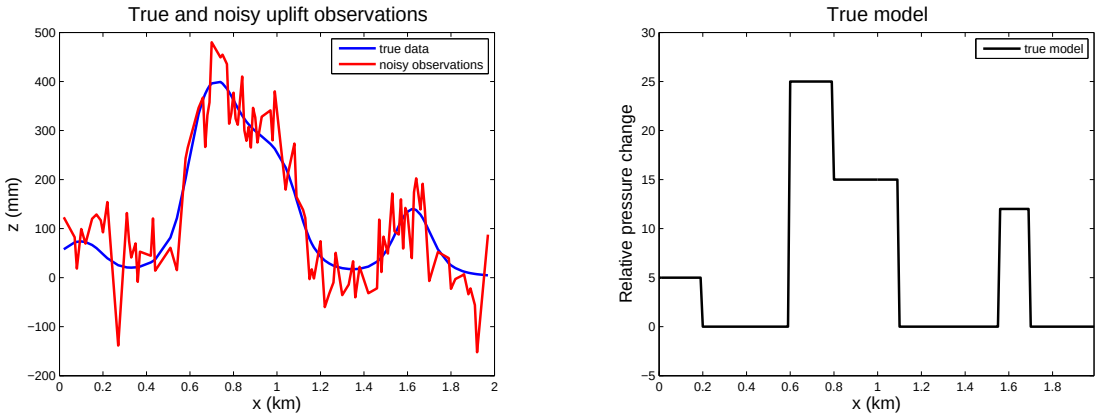
RESULTS

We demonstrate our method with a test problem that simulates vertical surface uplift in response to distributed dilatational sources, mathematically equivalent to surface deformation due to pore pressure change (Segall, 2010). Our modeling operator is defined as

$$\mathbf{F}(\mathbf{m}) \;=\; u(x),\; u(x) \;=\; \int_0^A \frac{D^3 m(\xi) d\xi}{\left(D^2 + (x-\xi)^2\right)^{3/2}}, \tag{19}$$

where we assume that $\mathbf{m} = m(\xi), \xi \in [0, A]$ is a relative pore pressure change along a linear segment $[0, A]$ of a reservoir at a constant depth D, and $\mathbf{u} = u(x), x \in [0, A]$, within a proportionality factor determined by poroelastic medium properties (Maharramov and Zoback, 2015), is the induced vertical uplift on the surface. For demonstration purposes we consider a one-dimensional model but the results trivially extend to realistic reservoir and surface geometries. Operator (19) is a smoothing operator, and recovering sharp pressure contrasts e.g. due to permeability barriers requires model "styling" or regularization such as blockiness-promoting ROF model. As a true model we used a highly compartmentalized pressure model of Figure 1(b). In our experiments, we set $D = 100$m $A = 2$km, and discretized both the model and data space using a 200-point uniform grid. Random Gaussian noise with $\sigma = 15\%$ of the maximum clean data amplitude was added to the clean forward-modeled data to produce the noisy observations shown in Figure 1(a).

The result of a TV-regularized *unconstrained* inversion is shown in Figure 2(a) against the true model and a Tikhonov-regularized inversion. This result was obtained using the above algorithm by setting $\delta = 0$ (no bound constraints) and using the values of $\alpha = 1$ and $\lambda = 2$. The TV-regularized result captures the compartmentalized picture of pressure distribution better than the highly smoothed Tikhonov regularization result. However, due to absence of bound constraints, lower pressure bounds are not honored, resulting in negative pressure areas that are not present in the true model. The result of running our bound-constrained TV-regularization algorithm is shown in Figure 2(b). The imposition of bound constraints not only removed the negative relative pressure areas, but also removed the pressure spike at about $x \approx 1$km in the unconstrained inversion of Figure 2(a) that apparently had resulted from compensating negative pressures. In both the constrained and the unconstrained runs we conducted 1000 outer loop iterations with 2 inner loops cycles. However, the algorithm converged quickly, with only a few initial iterates outside a tight neighborhood of the final curve, as shown in Figure 3(b). Finally, we note that many



(a) (b)

Figure 1: (a) True and noisy uplift observations. Random Gaussian noise with $\sigma = 15\%$ of maximum clean data amplitude was added to the clean data. (b) True model exhibibits a highly compartmentalized "blocky" behavior. [**CR**] musa5/. tvdata,true

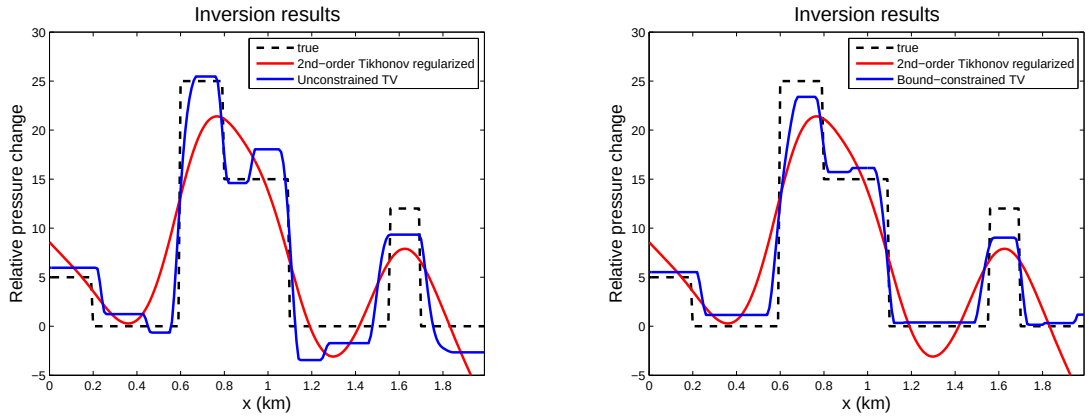
practical implementations of bound constraints often resort to a simplistic way of enforcing the constraints: the inverted model is projected onto the bounding rectangle either once after applying a direct unconstrained solver, or at each iteration of an unconstrained solver. In this case variable $\mathbf{y}$ and the associated quadratic regularization term are not introduced into the objective function. This may result in a violation of the KKT optimality conditions where the bound constraints are active (Nocedal and Wright, 2006), and is demonstrated by the blue plot in Figure 3(a). While the bound constraints are honored, the solution is both qualitatively and quantitatively far from optimal.

CONCLUSIONS AND PERSPECTIVES

Our algorithm combines the advantages of the blockiness-promoting and edge-preserving ROF model with the ability to impose bound constraints. The splitting mechanism used for enforcing the bound constraints is naturally integrated into the split-Bregman iterations and results in no extra computational cost. The method was able to resolve compartmentalized subsurface pressure changes from noisy surface uplift observations despite the highly diffusive nature of the underlying deformation process. The method can be implemented around any large-scale nonlinear solver such as conjugate gradients or quasi-Newton methods. Additional equality and inequality constraints can be incorporated into the algorithm using the general ADMM framework.

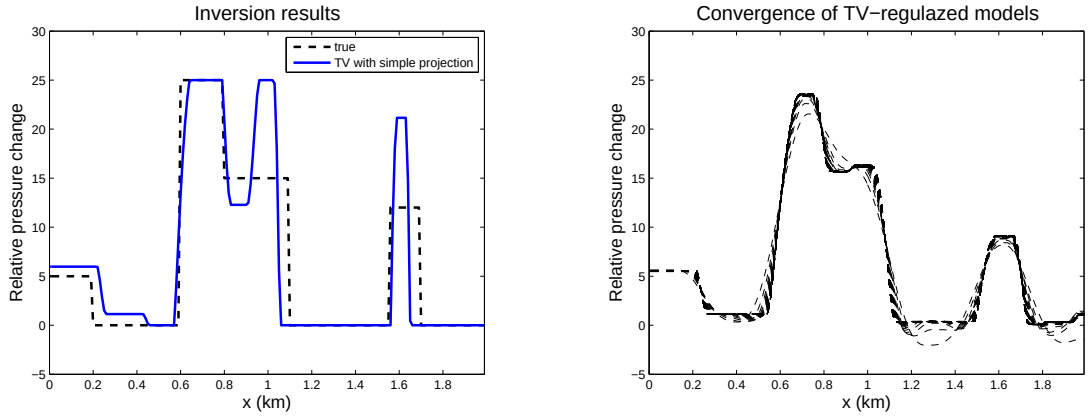
REFERENCES

Boyd, S., N. Parikh, E. Chu, B. Peleato, and J. Eckstein, 2010, Distributed optimization and statistical learning via the alternating direction method of multipliers: Foundations and Trends in Machine Learning, **3**, 1–122.



(a) (b)

Figure 2: (a) Unconstrained TV-regularized inversion. The algorithm tries to fit the data by allowing negative relative pressure changes. (b) Bound constrained TV-regularized result. Note that enforcing lower bounds resulted in a more accurate shape matching of the true model. [**CR**] musa5/. tvinv,boundtvinv



(a) (b)

Figure 3: (a) Direct imposition of the bound constraints at each iteration of the unconstrained solver resulted in a qualitatively and quantitatively wrong inversion. (b) Convergence of TV-regularized inverted models with bound constraints. The method quickly resolves both sharp contrasts and active bounds as only a few initial curves out of 1000 iterates lie outside a small neighborhood of the final curve. [**CR**] musa5/. simplefailboundtvinv,boundtvconv

Chambolle, A. and P. L. Lions, 1997, Image recovery via total variational minimization and related problems: Numerische Mathematik, **76**, 167–188.

Chartrand, R. and B. Wohlberg, 2010, Total-variation regularization with bound constraints: Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference, 766–769.

Goldstein, T. and S. Osher, 2009, The split bregman method for L1-regularized problems: SIAM Journal on Imaging Sciences, **2**, 323–343.

Ma, Y., M. Maharramov, R. Clapp, and B. Biondi, 2015a, Illumination compensation by L1 regularization and steering filters: SEP Report, **158**.

———, 2015b, Time lapse seismic imaging with L1 regularization and steering filters: SEP Report, **158**.

Maharramov, M. and B. Biondi, 2015, Resolving the effects of production-induced overburden dilation using simultaneous TV-regularized time-lapse FWI: SEP Report, **158**.

Maharramov, M., Y. Ma, and B. Biondi, 2015, Double-difference time-lapse FWI with a total-variation regularization: SEP Report, **158**.

Maharramov, M. and M. Zoback, 2015, Reservoir monitoring by inverting pore pressure change from surface deformation: SEP Report, **158**.

Nocedal, J. and S. J. Wright, 2006, Numerical optimization: Springer.

Parikh, N. and S. Boyd, 2013, Proximal algorithms: Foundations and Trends in Optimization, **1**, 123–231.

Rudin, L. I., S. Osher, and E. Fatemi, 1992, Nonlinear total variation based noise removal algorithms: Physica D: Nonlinear Phenomena, **60**, 259–268.

Segall, P., 2010, Earth and volcano deformation: Princeton University Press.

Illumination compensation by L1 regularization and steering filters

Yinbin Ma, Musa Maharramov, Robert Clapp and Biondo Biondi

ABSTRACT

L1/L2-regularization techniques often generate better results than the conventional least-squares solutions for inverse problem in geophysics. We implement a method to combine L1 regularization with steering filters. We obtain steering filters iteratively from input data without using any prior information. The numerical examples show significant improvement in comparison with the standard least squares. We demonstrate our method is robust with respect to inaccurate steering filters.

INTRODUCTION

L1-regularized optimization often yields more robost results in comparison with the standard least-squares optimization and is useful in geophysics when data is contaminated by high-amplitude noise. It is also necessary, if we will apply a filter to the model after inversion to seperate data from a different physical process, like the primary-multiple separation. L1 regularization, as a sparsity-promoting technique, is especially useful for time-lapse inversion, because the area of production-induced change is bounded (Maharramov and Biondi, 2014a,b, 2015).

In our previous work (Ma et al., 2014), we solved L1-regularized linearized waveform inversion (Tang, 2008, 2011) using a number of solution techniques: least-squares with conjugate gradient (CGLS), iterative reweighted least squares (IRLS), alternating direction method of multipliers (ADMM) and Split-Bregman method (Goldstein and Osher, 2009; Boyd, 2010), hyperbolic penalty function (HPF) with conjugate directions (Claerbout, 2009; Zhang and Claerbout, 2010). While all the methods performed well, we demonstrated that L1 regularization delivered a significant improvement over standard least squares. However, the results still suffer from lingering effects of illumination gaps in the data. In this work, we address this issue by combining L1-regularized inversion with the concept of steering filters (Clapp et al., 2005; Prucha and Biondi, 2002).

Steering filters can improve the quality of inversion, especially in complex overburden. We style the L1 regularization term in a way that favors specified dip directions and sparsity in the direction orthogonal to the dip. We study the same test problem as Zhang and Claerbout (2010), and demonstrate that with suitable prior knowledge of the dip structure, we can effectively compensate insufficient illumination and greatly improve the inversion results.

In this work we demonstrate that even without prior geological information, and using only seismic data, we can iteratively construct L1-regularization steering filters, thereby achieving a quality of inversion comparable to that when good prior information is available.

We begin by solving a least squares problem with the zero-order Tikhonov regularization (minimal norm solution) that yields an estimate of the gradient field. We then use this estimate to construct an initial steering filter. This steering filter is incorporated into the L1 regularization term as described in Section 2.2, and we use the results of subsequent regularized inversions to update the steering filter iteratively. We discuss techniques of regularizing estimated gradient field for the purpose of constructing robust steering filters in the **appendix**.

OPTIMIZATION METHOD AND REGULARIZATION

We have implemented several L1/L2 solvers and an HPF solver in our Stanford Exploration Project (SEP)-Vector library. Previously we have tested solvers on a few geophysical examples (Ma et al., 2014). We have shown that the quality of inversion results and computational costs are comparable for all the solvers in the examples we considered. The ADMM/Split-Bregman methods are deemed to be better than IRLS for compressed sensing and denoising problems. However, because it is hard to select the proper parameters of the ADMM/Split-Bregman methods, in this work we conduct L1-regularized linearized waveform inversion using our IRLS solver.

Regularization plays an important role in geophysical inverse problems. Regularization prevents overfitting noisy measurements, and ameliorates ill-posedness due to lack of data. Our mathematical model is only an approximation of a real physical process, and a suitable regularization may mitigate the mismatch between model and observation.

Regularization can be considered as "model styling", imparting on our model certain features considered desirable in the context of our problem (Claerbout, 2014). In this sense, L1-regularized steering filters can be considered a constraint on model inversion favoring the continuity of dip structures.

Adaptive regularization has been proposed before, e.g. based on Baysian analysis (Zamanian et al., 2014). This typically requires solving auxiliary inverse problems. In our deterministic approach, the L1-regularization term is updated between solution iterations.

LINEARLIZED WAVEFORM INVERSION

Description of the problem and challenges

The target-oriented linearlized waveform inversion has been explored in several previous works (Clapp et al., 2005; Valenciano, 2006; Tang, 2008, 2011; Zhang and Claerbout, 2010; Ma et al., 2014). We are trying to solve:

$$\mathbf{Hm} \approx \mathbf{m_{mig}}, \tag{1}$$

where $\mathbf{H}$ is the Hessian operator,$\mathbf{m_{mig}}$ is called migrated data which is known, and $\mathbf{m}$ is the model we want to compute.

A standard method to solve the problem is least squares inversion. Regularization terms are usually included in the objective function for least squares inversion when the modeling

operators are singular and data contains noise.

In this specific example, we have several challenges, as can be seen from the true reflectivity model and data used for inversion in Figure 1. The forward modeling is a linear approximation of the true operator. We have poor illumination under the salt. The existence of faults prevent us from getting a high quality model considering that Hessian operator acts like "convolution" operator.

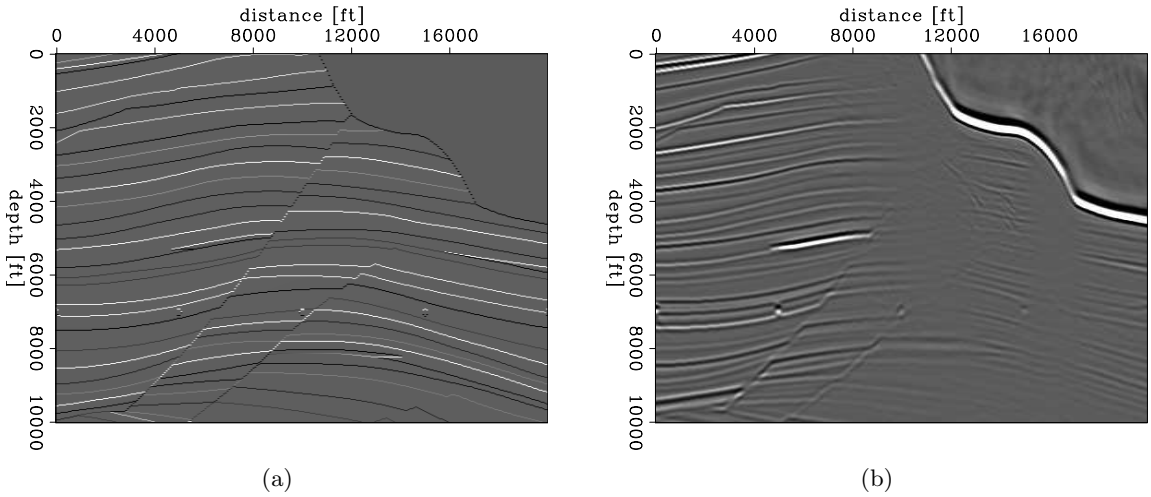


Figure 1: (a) Reflectivity model. (b) input migrated image. [**NR**] yinbin1/. refl-imaging,data-imaging-mig

Power of L1 regularization and steering filters

Previous results (Zhang and Claerbout, 2010; Ma et al., 2014) using objective function:

$$\mathbf{J(m)} = \frac{1}{2}||\mathbf{Hm} - \mathbf{m_{mig}}||_2^2 + \varepsilon||\mathbf{m}||_1, \tag{2}$$

suggest that we can get more sparse reflectors in the well-illuminated zone (the left part of the model) using L1/L2 solvers. However, we do not get significant improvement under the salt. In addition, the data residues from L1/L2 solvers are more correlated comparing with CG method, suggesting our results are not convincing.

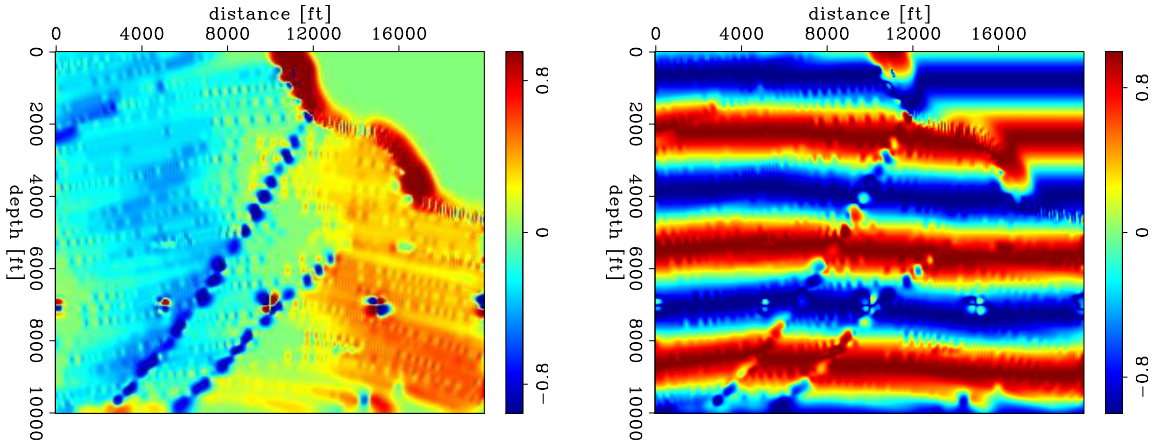
To further improve the results, the regularization term is replaced by more geologically plausible constraints, namely steering filters,

$$\mathbf{J(m)} = \frac{1}{2}||\mathbf{Hm} - \mathbf{m_{mig}}||_2^2 + \varepsilon_\mathbf{a}||\mathbf{W_a}\nabla_\mathbf{r}\mathbf{m}||_1 + \varepsilon_\mathbf{b}||\mathbf{W_b}\mathbf{m}||_1, \tag{3}$$

where $\mathbf{W_a}$ and $\mathbf{W_b}$ control the strength of regularization at each point, because we do not have equal illumination. The derivative $\nabla_\mathbf{r}$ is taken along the gradient prior. The geological information is contained in $\mathbf{W_a}$, $\mathbf{W_b}$ and $\mathbf{r}$.

To test the power of L1 regularization and steering filters, we assume we already have a good prior for $\mathbf{W_a}$, $\mathbf{W_b}$ and $\mathbf{r}$ from external knowledge. $\mathbf{r}$ in Figure 2(a) is used as a prior

gradient field. The choice is natural, because previous constraints $\nabla_{\mathbf{x}}\mathbf{m}$ would penalize dipping reflectors. Next, we set $\mathbf{W_b} = 1.$ for $\mathbf{x} < 10000ft$ and $\mathbf{W_b} = 0.$ for $\mathbf{x} > 10000ft$, and $\mathbf{W_a} = 1.$. We made this choice based on knowing the area under the salt is poorly illuminated, and we need more compensation from steering filters.



(a) (b)

Figure 2: (a) x component of correct gradient prior. (b) x cpomponent of wrong gradient prior. [**ER**] yinbin1/. CorrectPriorDx,WrongPriorDx

The final results in Figure 3(e) are better comparing with CG method (Figure 3(a)) and previous L1/L2 results (Figure 4(a)). We have obtained sparse reflectors as expected from hybrid L1/L2 solvers. The faults are correctly recovered and the reflectors under the salts are properly inverted. The data residue in Figure 3(f) looks more random comparing with previous results. We have run 30 by 20 CG steps to obtain the results in Figure 3(e), which is ~ 10 times more expensive than the CG method. The cost is acceptable considering the improvement of inversion results.

Obtaining steering filters from the data iteratively

In the previous subsection, we have demonstrated L1 regularization and steering filters can help in the inverse problems. However, it is not trivial to obtain proper geological constraints. We can add external knowledge that is independent of seismic data: geophysical, geological, and geomechanical information. Before we have enough knowledge, we can squeeze the data and construct reasonable steering filters from the seismic data.

Numerically, assume a prior gradient field is not available, though a little "external information" is known: the strongly dipping events at $z \sim 4000ft$ and $x \sim 12000ft$ are artifacts. We use the following workflow. First, we do inversion with uniform constraints as in equation 8. Next, we extract the gradient field $\mathbf{r}$ from the inversion results and mute the dipping reflectors near $z \sim 4000ft$ and $x \sim 12000ft$. Then we use the gradient field as prior to construct steering filters and do the inversion again with equation 3 and repeat the process until the results converge. $\mathbf{W_a}$, $\mathbf{W_b}$ should also be changed iteratively based on previous inversion results; however, we adjust them manually for simplicity.

We run the IRLS algorithm 6 times, and we can see the results in Figure 4. Significant

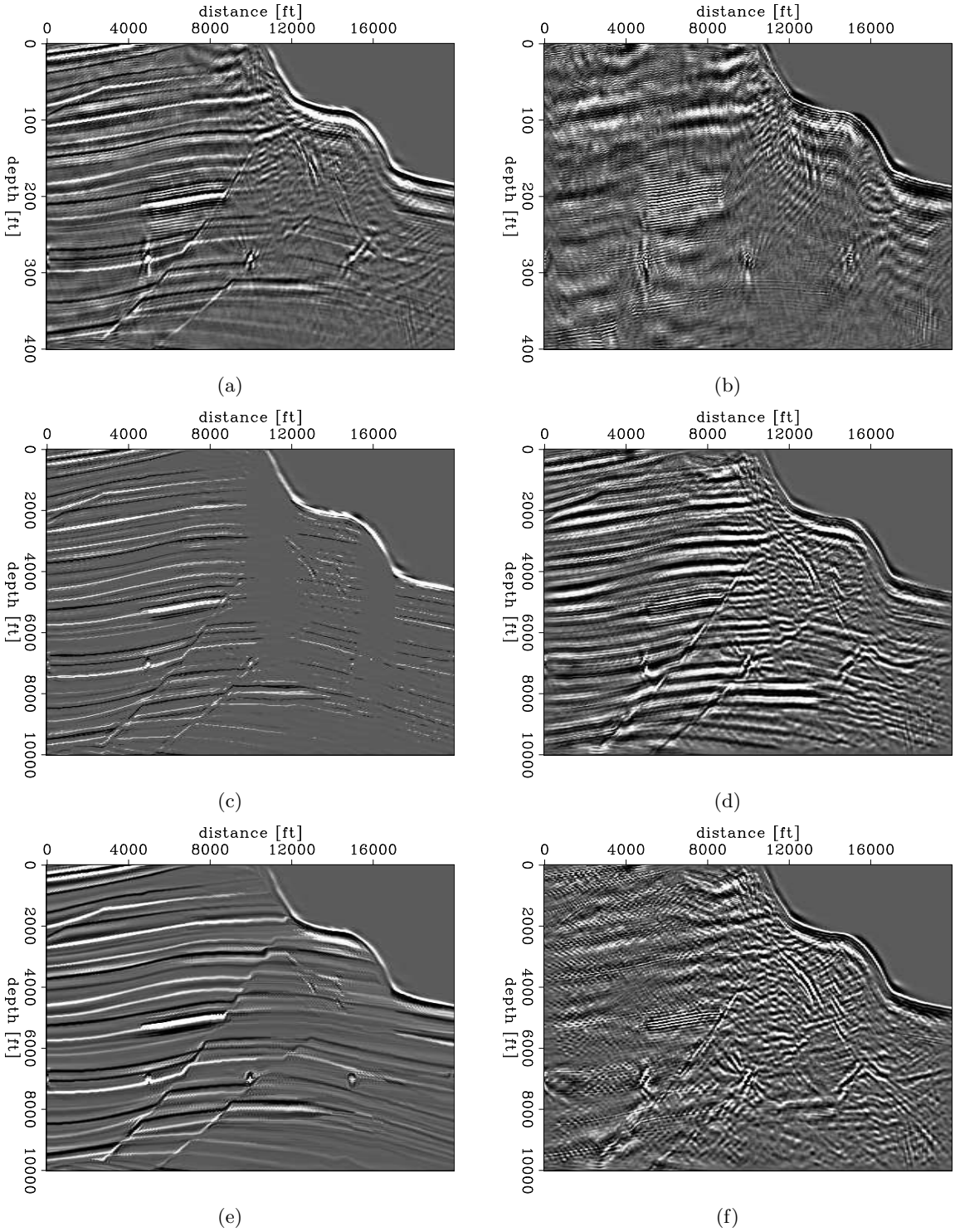


Figure 3: Comparison of linearized waveform inversion results and residues: (a) CGLS inversion result. (b)CGLS inversion residue. (c) IRLS inversion result . (d) IRLS inversion residue. (e) IRLS inversion result with gradient prior. (f) IRLS inversion residue with gradient prior. All figures are clipped at the same level. [**CR**]

yinbin1/. LSICGmod,LSICGresd,L1Sparsemod,L1Sparseresd,L1SteeringFiltermod,L1SteeringFilterresd

improvement can be observed which means we can really get more information from the data if we go beyond simple least squares inversion. However, we could not match the results in the previous section (where we assume we know good prior).

Do the wrong steering filters ruin the inversion results?

We put our knowledge into the objective function as prior and we want to know how our choice affects the results. Firstly, we are geophysicists, not data scientists. When we solve an inverse problem, we rarely obtain our results exclusively from the data. Even in the simplest case of geophysical inverse problem with regularization, the extra terms $\nabla_{\mathbf{r}}\mathbf{m}$ or $\mathbf{m}$ come from our geophysical knowledge of the previous subsurface studies. Secondly, we iteratively construct steering filters and there is inevitable inaccuracy.

Ideally, we want proper prior to promote correct reflectors in the poorly illuminated area, and wrong prior has negligible effects. We can justify our assumption qualitatively. Because the true model lies close to the null spaces of forward modeling operator $\mathbf{H}$ and correct regularization operator $\nabla_{\mathbf{r}}$, the correct objective yields minimum norm solution within the intersection of the two null space. If the null space is not large, we would expect the inversion results to be close to the true model. However, if we provide wrong constraints $\nabla_{\tilde{\mathbf{r}}}$, then $Null(\mathbf{H})$ would be far away from $Null(\nabla_{\tilde{\mathbf{r}}})$, and there is no way to fit them at the same time. Considering the weight on the regularization term is small comparing with the weight of the data fitting term, the wrong constraints will be negligible.

We use the same objective function as described in the previous subsection with the wrong gradient prior as in Figure 2(b). We set $\mathbf{W_b} = 1$ everywhere, because in this case as we assume we do not know whether the gradient prior is good or bad. We also twist the gradient prior as in Figure 2(b), and we can see the inversion result in Figure 5(c). The correct gradient prior yields better results at the gaps under the salt, while the wrong gradient prior does not introduce too many artifacts at the gaps, which proves our method is robust against inaccurate steering filters.

CONCLUSIONS

In conclusion, we present a workflow to obtain better results from geophysical inverse problems by combining L1 regularization with steering filters. We prove improvement is possible when no prior knowledge is available. When applying the methods to a real problem, we can use external knowledge like geomechanics to construct more reasonable steering filters.

ACKNOWLEDGMENT

We would like to thank Dr. Stewart Levin for help on constructing local filters.

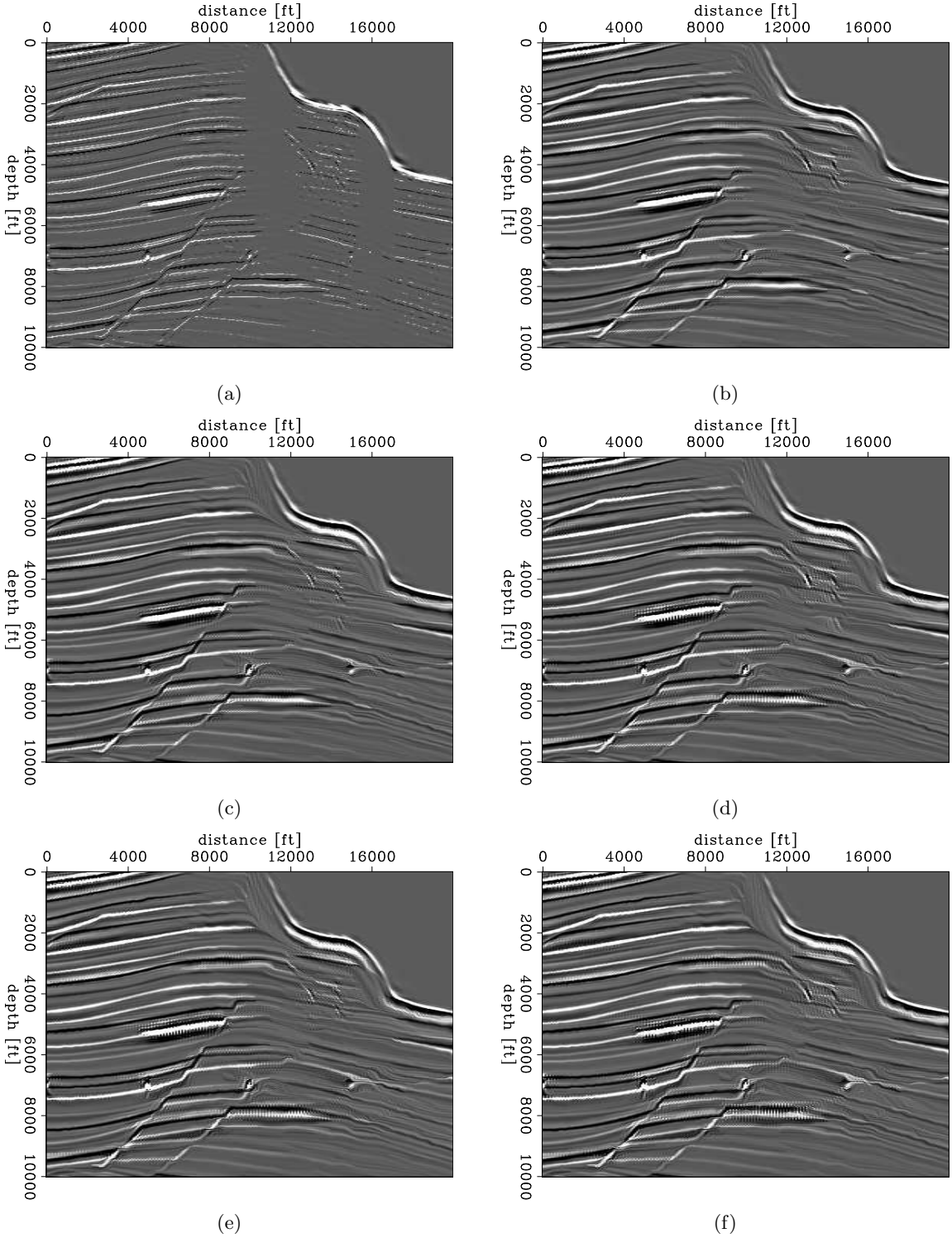


Figure 4: Comparison of linearized waveform inversion results and residues: (a) 1st iteration of inversion with no prior. (b) 2nd iteration of inversion with prior. (c) 3rd iteration of inversion with prior. (d) 4th iteration of inversion with prior. (e) 5th iteration of inversion with prior. (f) 6th iteration of inversion with prior. [**NR**]

yinbin1/. L1Sparsemod,L1SteeringFilter2mod,L1SteeringFilter3mod,L1SteeringFilter4mod,L1SteeringFilter5

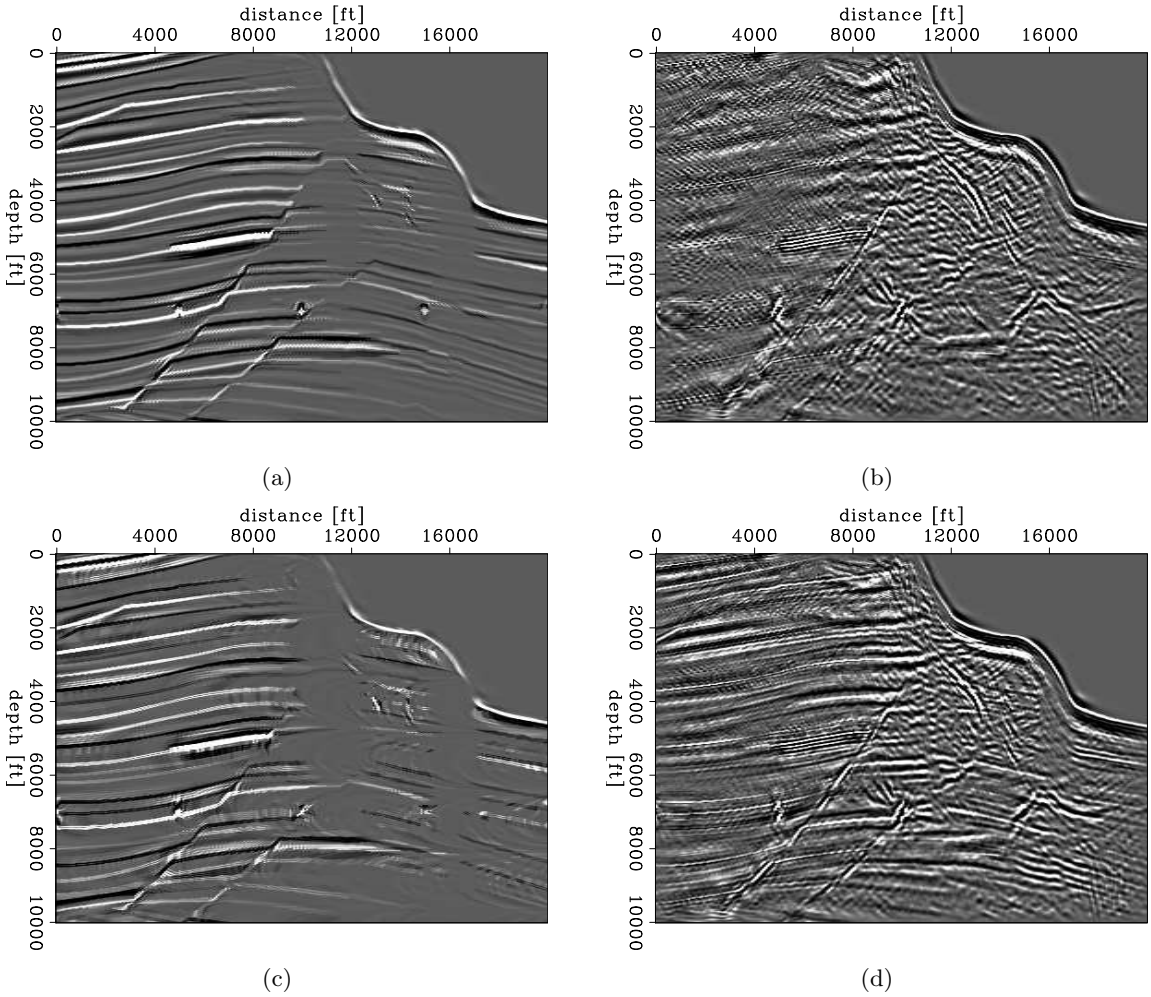


Figure 5: Comparison of linearized waveform inversion with different gradient prior: (a), (b) Inversion result and residue with gradient prior in Figure 2(a). (c), (d) Inversion result and residue with gradient prior in Figure 2(b). All figures are clipped at the same level. [**CR**] yinbin1/. CorrectPriormod,CorrectPriorresd,L1WrongPriormod,WrongPriorresd

APPENDIX A: CONSTRUCTION OF STEERING FILTERS BASED ON LEAST SQUARES

In this appendix, a method to construct steering filters from images is presented. The filters are used as prior information for the next iteration, before we understand how to incorporate geomechanics into regularization.

We are interested in obtaining a smoothly varying gradient field. Correct gradient direction boosts the desired events in the area of poor illumination. Incorrect prior of gradient direction is ignored by the optimization process thus does not create many artifacts, as proved previously by (Prucha and Biondi, 2002). In order to test the strength and weakness of our method, we apply it to the example in Figure 6, following the idea from Hale (2007).

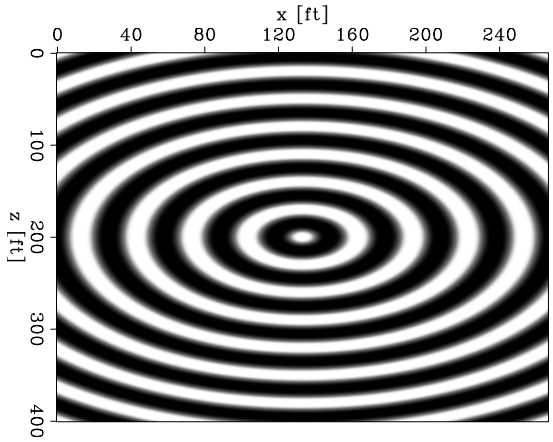


Figure 6: Test model with all dipping directions. [**ER**] yinbin1/. RingRefl

Suppose we have a 2D image $\mathbf{m}_0(x,z)$ from which we want to extract the gradient field. We first use the Sobel filter to estimate $\partial_x \mathbf{m}_0$ and $\partial_z \mathbf{m}_0$, from which we can compute the norm of gradient $\|\nabla \mathbf{m}_0(x,z)\|_2$ and dipping angle $\theta_0(x,z)$. It is obvious that $\theta_0(x,z)$ cannot be directly used as prior gradient field for the next iteration. θ_0 is not smooth because of the noise in the image and crossing events, etc. θ_0 is not reliable when $\|\nabla \mathbf{m}_0(x,z)\|_2 \approx 0$.

To fix those problems, we need to construct a weighting function $\mathbf{W}(x,z)$ to suppress unreliable θ_0. We choose

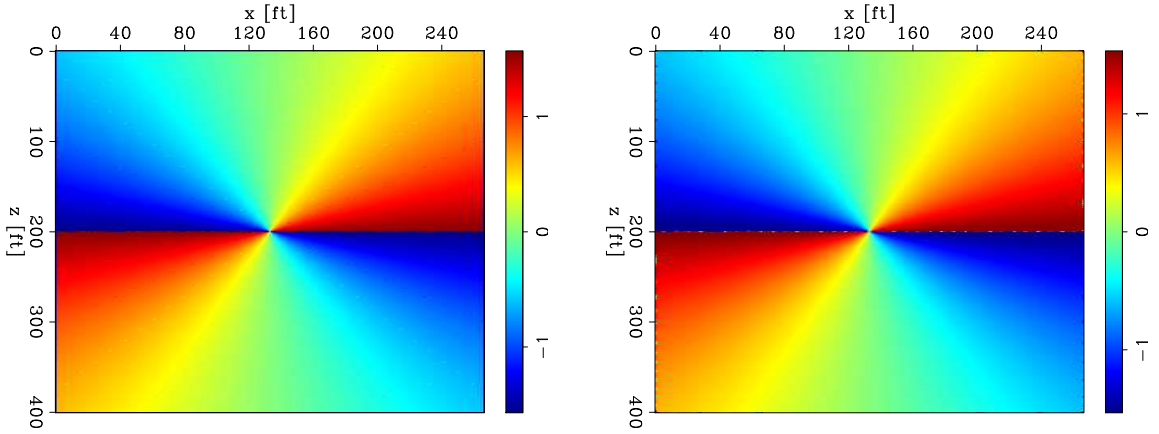
$$\mathbf{W}(x,z) = \|\nabla \mathbf{m}_0(x,z)\|_2 \times (\lambda + |\mathbf{m}_0(x,z)|), \tag{4}$$

where we have a smaller weight when $\|\nabla \mathbf{m}_0(x,z)\|_2 \approx 0$ or $|\mathbf{m}_0| \approx 0$. With this weight function we can set up a linear inverse problem with an objective function,

$$\mathbf{J}(\theta) = \frac{1}{2}\|\mathbf{W}(\theta - \theta_0)\|_2 + \frac{\varepsilon}{2}\|\nabla^2 \theta\|_2. \tag{5}$$

The first term is data fitting with larger weight on the reliable estimation of gradient direction. The second term is regularization which promotes a smooth solution.

We can solve the optimization problem to obtain θ^*, we can see θ_0 and θ^* in Figure 7. There are some artifacts in θ_0 caused by the numerical issue with 3 by 3 Sobel filter, and the artifacts are removed in θ^*.



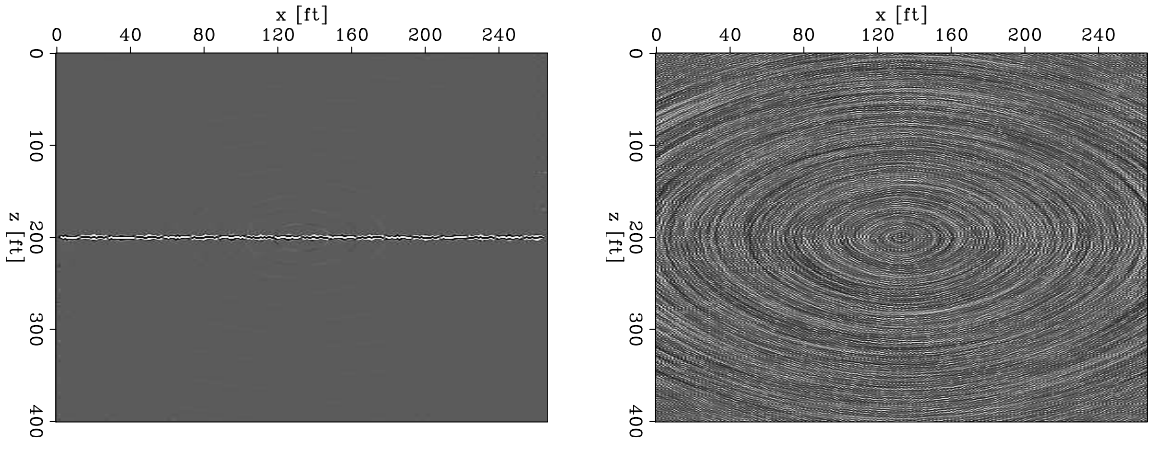
(a) (b)

Figure 7: Left: $\theta_0 - \pi/2$. Right: $\theta^* - \pi/2$. [**ER**] yinbin1/. RingDip0,RingDip

Once we obtain a smooth gradient field characterized by dipping angle, we can construct local steering filter,

$$\mathbf{L}_{\theta^*} = \nabla_{r_\perp^*}, \tag{6}$$

where $r_\perp^*$ is perpendicular to the gradient direction associated with θ^*. We apply the local filter to the original image, and obtain the result in Figure 8(a). We can see that our method works very well except for the vertical-dipping direction. The reason is that our objective function is built on angle and suffers from branch cut problem. Removing the branch cut will lead to a nonlinear optimization problem that is beyond the scope of this report.



(a) (b)

Figure 8: (a) Apply local filter to original image.(b) Reconstructed image from noise. [**ER**] yinbin1/. RingDipDestruct,RingDipConstruct

We also test the local filter by solving an inverse problem with

$$\mathbf{F}(\mathbf{m}) = \frac{1}{2}\|\mathbf{L}_{\theta^*}\mathbf{m} - \mathbf{d}\|_2^2 + \frac{\alpha}{2}\|\mathbf{m}\|_2^2, \tag{7}$$

where $\mathbf{d}$ is uniform random noise. We would expect $\mathbf{m}^*$ obtained by minimizing $\mathbf{F}(\mathbf{m})$ should have the same curvature as the original image. We can see $\mathbf{m}^*$ in Figure 8(b).

Finally, in this paper, we apply our method to the example from the Sigbee2A model and the results can be seen in Figure 9(c). It is interesting to see that after applying the local filter, we are able to see the fault and multiple more clearly.

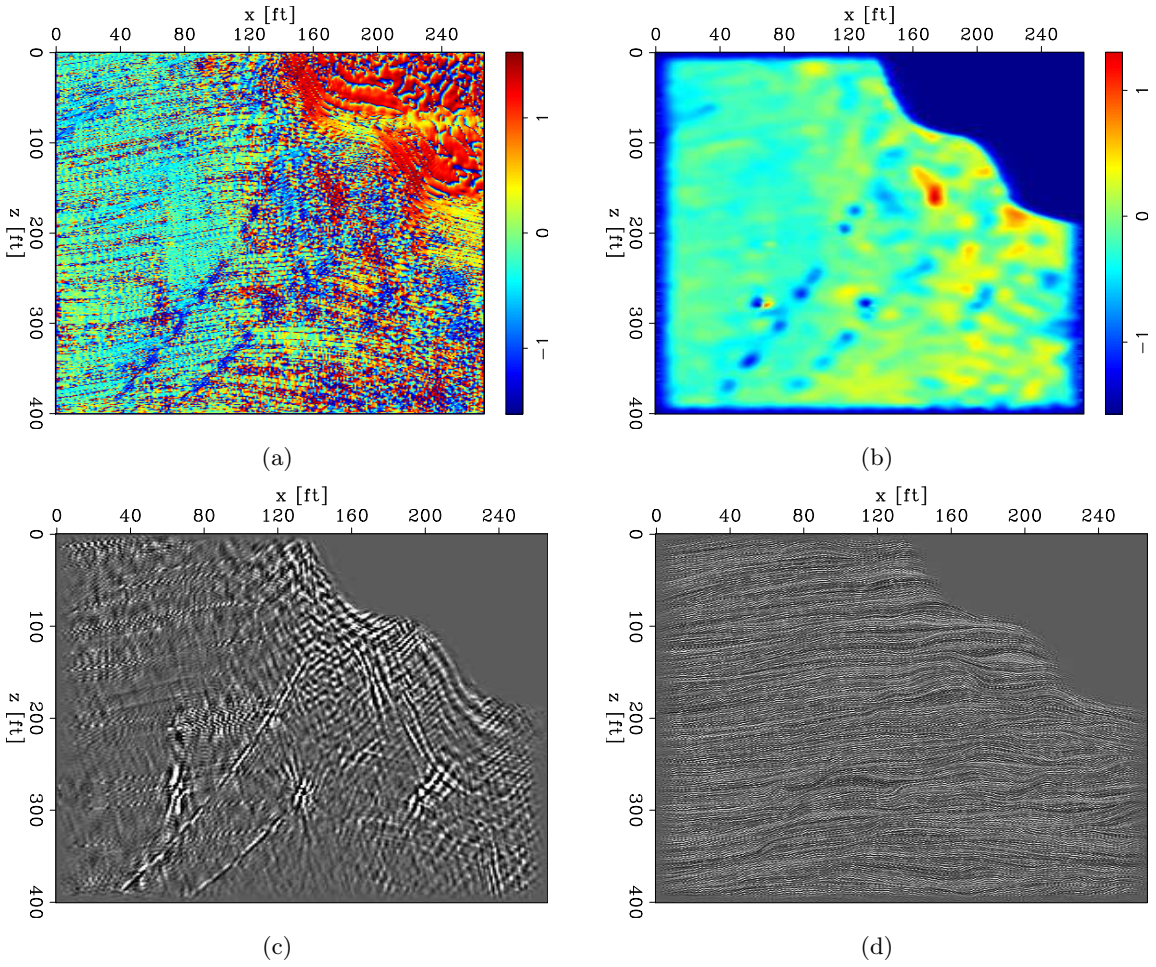


Figure 9: (a) $\theta_0 - \pi/2$. (b) $\theta^* - \pi/2$. (c) Apply steering filter to original image. (d) reconstructed image from noise. [**ER**] yinbin1/. SigbeeDip0,SigbeeDip,SigbeeDipDestruct,SigbeeDipConstruct

REFERENCES

Boyd, S., 2010, Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers: Foundations and Trends in Machine Learning, **3**, 1–122.

Claerbout, J. F., 2009, Blocky models via the L1/L2 hybrid norm: SEP Report, **139**, 1–10.

———, 2014, GEOPHYSICAL IMAGE ESTIMATION BY EXAMPLE.

Clapp, M. L., R. G. Clapp, and B. L. Biondi, 2005, Regularized least-squares inversion for 3-D subsalt imaging: SEG Technical Program Expanded Abstracts 2005, 1814–1817.

Goldstein, T. and S. Osher, 2009, The Split Bregman Method for L1-Regularized Problems: SIAM Journal on Imaging Sciences, **2**, 323–343.

Hale, D., 2007, Local dip filtering with directional Laplacians: CWP Report-567, 91–102.

Ma, Y., M. Maharramov, and B. Biondi, 2014, Geophysical applications of hyperbolic and hybrid L1/L2 optimization: SEP Report, **155**, 133–144.

Maharramov, M. and B. Biondi, 2014a, Joint full-waveform inversion of time-lapse seismic data sets: SEG Technical Program Expanded Abstracts 2014, 954–959.

———, 2014b, Robust joint full-waveform inversion of time-lapse seismic datasets with total-variation regularization: SEP Report, **155**, 199–208.

———, 2015, Robust Simultaneous Time-lapse Full-waveform Inversion with Total-variation Regularization of Model Difference: 77th EAGE Conference and Exhibition, Extended Abstracts (accepted).

Prucha, M. L. and B. L. Biondi, 2002, Subsalt event regularization with steering filters: SEG Technical Program Expanded Abstracts 2002, 171–174.

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

———, 2011, Imaging and velocity analysis by target-oriented wavefield inversion: PhD thesis, Stanford University.

Valenciano, A., 2006, Target-oriented wave-equation inversion: Geophysics, **71**, A35–38.

Zamanian, S. A., W. L. Rodi, M. I. T. Earth, J. A. Kane, and S. I. E, 2014, Hierarchical and Variational Bayesian Approaches to Least-Squares Migration: SEG Technical Program Expanded Abstracts 2008, 3731–3736.

Zhang, Y. and J. F. Claerbout, 2010, Least-squares imaging and deconvolution using the hybrid norm conjugate-direction solver: SEP Report, **140**, 129–142.

Time Lapse seismic imaging with L1 regularization and steering filters

Yinbin Ma, Musa Maharramov, Robert Clapp and Biondo Biondi

ABSTRACT

We propose a new L1-regularized simultaneous time-lapse linearized waveform inversion method. We test the proposed method on models exhibiting production induced reservoir compaction and overburden dialation. We demonstrate that L1 regularization and steering filters significantly improve image quality in the presence of acquisition nonrepeatability issues, such as those caused by different acquisition parameters.

INTRODUCTION

Time-lapse seismic imaging and velocity analysis are challenging in complex subsurface structures because of various repeatability issues, such as different acquisition geometries, noise in the data, and insufficient illumination under the salt (Ayeni and Biondi, 2010, 2011). In a companion paper, we showed (Ma et al., 2015) that L1 regularization with steering filters can greatly improve the quality of inversion in areas of poor illumination. In this paper, we apply this technique to time-lapse seismic imaging.

First, we implement the acoustic wave equation and Born modeling for constant density, and set up four-dimensional (4D) seismic imaging as an optimization problem. It is known that the adjoint state method is an efficient way to compute the gradient for partial-differential-equation (PDE) constrained optimization. This paper is limited to the study of linearized waveform inversion, therefore the adjoint operators can be written explicitly. Both forward and adjoint operators are expressed in matrix form, confirming the conclusion from the adjoint state method.

A simple synthetic model is used to test our codes. We take a shallow part from Sigbee2A model away from the salt. The velocity is perturbed near the sea bottom and used as monitor velocity model. Numerical results show that L1 regularization is able to recover clean 4D signals. The algorithm is robust against different acquisition geometries and noise.

Our method is also tested on a synthetic model with complex subsurface structure. We construct our model from Sigbee2A model with salt body, and then, perturb the velocity in the area with poor illumination under the salt. Least squares reverse time migration (LSRTM) results suggest it is already difficult to obtain the reflectors under the salt as can be seen in Figure 7, not to mention 4D signals. L1 regularization with steering filters is used to improve the quality of time lapse signals. We could not completely solve the time-lapse imaging problem with our method, thereby indicating the necessity to introduce geomechanical knowledge in the time-lapse study.

METHOD

In the first part of this section, the wave-equation operators used for seismic imaging are derived. In the second part, we show the objective function and optimization strategies incorporating L1 regularization with steering filters.

Acoustic wave equation and Born modeling for constant density media

We start with the wave equation from Tarantola (1984) as follows:

$$[\frac{1}{K(\mathbf{r})}\frac{\partial^2}{\partial t^2} - \nabla\left(\frac{1}{\rho}\nabla\right)]\mathbf{u}(\mathbf{r},t) = f(\mathbf{r},t), \tag{1}$$

where $K(\mathbf{r})$ is the bulk modulus, ρ is the density and $u(\mathbf{r},t)$ is the pressure field and $f(\mathbf{r},t)$ is the source function. In this paper, we assume constant density in our numerical examples, and we have $v^2 = K/\rho$. The wave equation can be simplified as:

$$[\partial_t^2 - v^2(r)\Delta]\mathbf{u}(\mathbf{r},t) = s(\mathbf{r},t), \tag{2}$$

where we define $s(\mathbf{r},t) \equiv v(\mathbf{r})^2 f(\mathbf{r},t)$.

In a seismic imaging problem, we want to find the velocity perturbation, namely the difference between the true velocity (stratigraphic velocity) and migration velocity (smooth velocity). For constant density media, there is direct relationship between velocity perturbation and acoustic reflectivity. Assuming the true velocity to be $v(\mathbf{r})$ and the smooth velocity used for migration/inversion to be $v_0(\mathbf{r})$, two different solutions can be obtained for each velocity using equation (1):

$$[\partial_t^2 - v^2(\mathbf{r})\Delta]\mathbf{u}_{\text{true}}(\mathbf{r},t) = s(\mathbf{r},t) \tag{3}$$
$$[\partial_t^2 - v_0^2(\mathbf{r})\Delta]\mathbf{u_0}(\mathbf{r},t) = s(\mathbf{r},t). \tag{4}$$

Substract equation (4) from equation (4), we can get:

$$[\partial_t^2 - v_0^2(\mathbf{r})\Delta]\mathbf{u}(\mathbf{r},t) = \mathbf{m}(\mathbf{r})\partial_t^2\mathbf{u_0}(\mathbf{r},t), \tag{5}$$

where we define $\mathbf{u} \equiv \mathbf{u}_{\text{true}} - \mathbf{u_0}$, $\mathbf{m}(\mathbf{r}) = -2\delta v(\mathbf{r})/v(\mathbf{r})$ and $\delta v(\mathbf{r}) = v(\mathbf{r}) - v_0(\mathbf{r})$.

To set up an inverse problem, both the forward and adjoint of modeling operators are needed. Based on the adjoint state method, the forward operator can be written as:

$$\begin{cases} \mathbf{u}(\mathbf{r},0) & = 0 \\ \partial_t\mathbf{u}(\mathbf{r},0) & = 0 \\ (\partial_t^2 - v_0(\mathbf{r})^2\Delta)\mathbf{u}(\mathbf{r},t) & = s(\mathbf{r},t), \end{cases} \tag{6}$$

and the corresponding adjoint operator is:

$$\begin{cases} \mathbf{u}(\mathbf{r},T) & = 0 \\ \partial_t \mathbf{u}(\mathbf{r},T) & = 0 \\ (\partial_t^2 - v(\mathbf{r})^2\Delta)^\dagger \mathbf{u}(\mathbf{r},t) & = s(\mathbf{r},t). \end{cases} \tag{7}$$

The forward and adjoint operators for nonlinear modeling with varying density are computed in **Appendix A**, which allows us to extend our work to velocity estimation in the near future. The Born modeling operator in matrix form is derived in **Appendix B**, which serves as pseudocode for the numerical examples in this paper.

Joint inversion with steering filters and L1 regularization

Fixing the migration velocity $v_0(\mathbf{r})$, the data recorded at surface $\mathbf{d}$ is a linear function of $\mathbf{m}(\mathbf{r})$, as can be seen from equation (5). We use $\mathbf{L}$ to represent the linear mapping,

$$\mathbf{L}\mathbf{m} = \mathbf{d}. \tag{8}$$

In the following of the paper, we use subscripts b and m to denote baseline and monitor, respectively. Data for baseline $\mathbf{d}_b$ and monitor $\mathbf{d}_m$ are collected using true velocity model. Reverse time migration (RTM) is used to obtain images $\mathbf{m}_b^{\text{RTM}}$ and $\mathbf{m}_b^{\text{RTM}}$ as initial guess for inversion. We do an inversion for baseline and monitor separately as a first step using the following objective function,

$$\begin{aligned} \mathbf{J}_1(\mathbf{m}_b,\mathbf{m}_m) &= \frac{1}{2}||\mathbf{L}_b\mathbf{m}_b - \mathbf{d}_b||_2^2 + \frac{1}{2}||\mathbf{L}_m\mathbf{m}_m - \mathbf{d}_m||_2^2 && (9)\\ &+ \frac{\varepsilon}{2}||\mathbf{m}_b||_2^2 + \frac{\varepsilon}{2}||\mathbf{m}_m||_2^2, && (10) \end{aligned}$$

where $\varepsilon \to 0$ in the numerical sense to fit the data. After a fixed number of iterations, we obtain images $\mathbf{m}_b^{\text{LSRTM}}$ and $\mathbf{m}_b^{\text{LSRTM}}$. This step is crucial, because it balances the amplitude and attenuates the acquisition related artifacts in the RTM images. LSRTM images are also used to extract a prior gradient field, which is part of the joint-inversion objective function.

The final step is to resolve the time-lapse signal using the joint inversion, with following objective function:

$$\begin{aligned} \mathbf{J}_2(\mathbf{m}_b,\mathbf{m}_m) &= \frac{1}{2}||\mathbf{L}_b\mathbf{m}_b - \mathbf{d}_b||_2^2 + \frac{1}{2}||\mathbf{L}_m\mathbf{m}_m - \mathbf{d}_m||_2^2 && (11)\\ &+ \frac{\alpha}{2}||\mathbf{W}(\mathbf{m}_m - \mathbf{m}_b)||_1 && (12)\\ &+ \frac{\beta}{2}||\mathbf{W}(\nabla_{\mathbf{r}}(\mathbf{m}_m - \mathbf{m}_b))||_1 && (13)\\ &+ \frac{\varepsilon}{2}||\mathbf{m}_b||_2^2 + \frac{\varepsilon}{2}||\mathbf{m}_m||_2^2, && (14) \end{aligned}$$

where $\mathbf{W}$ is weight function and $\nabla_{\mathbf{r}}$ is the derivative perpendicular to $\mathbf{r}$. In the objective function, The terms (11) are data fitting and the terms (14) are used to fill the null space. The terms (12) promotes sparse 4D signals and further reduces acquisition related artifacts. Line (13) is the steering filter that promotes 4D signals along particular directions $\mathbf{r}$. In (Ma et al., 2015), we showed that steering filters are able to compensate illumination changes under the salt, and we proposed a method to construct a steering filter that can be used as prior information for inversion.

SYNTHETIC EXAMPLE 1

We select part of Sigbee 2A model as our true baseline velocity model. We perturb the velocity at the boundary of subsurface layers and use it as true monitor velocity model. Migration velocity is the same for both baseline and monitor. The velocity differences between true velocity and migration velocity are shown in Figure 1.

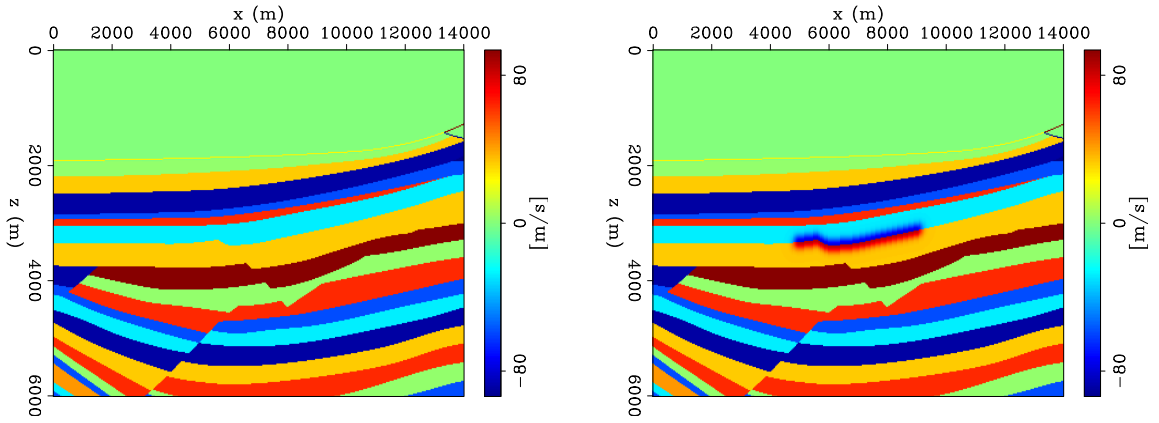


Figure 1: (a) baseline velocity perturbation. (b) monitor velocity perturbation. [**ER**] yinbin2/. simple.diffT0,simple.diffT2

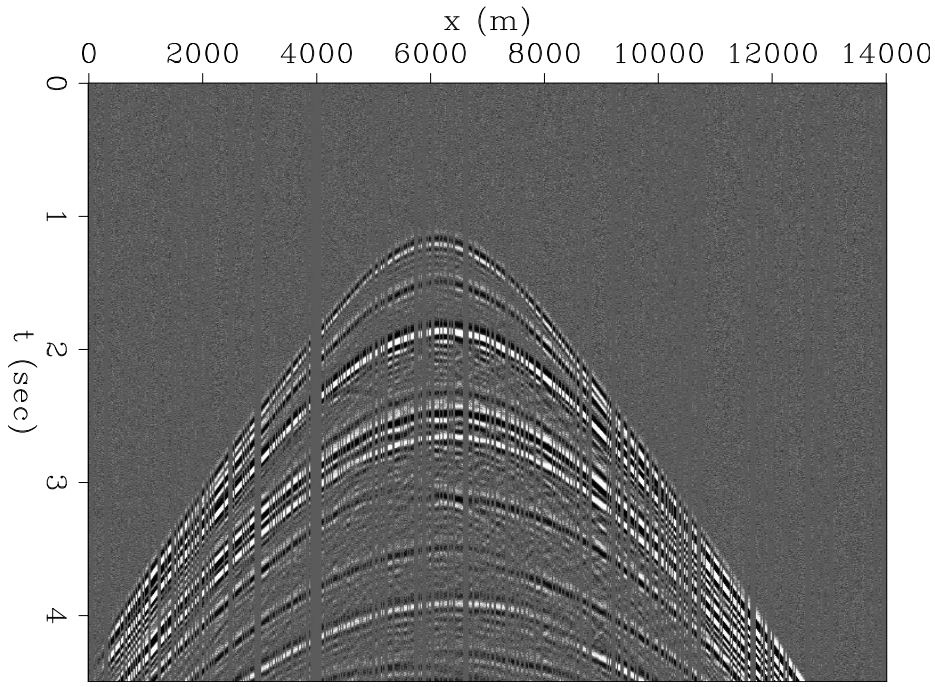
Model size is 301×701 with 20 meters (m) between neighboring pixels. 48 sources spaced every 200 meters with up to 701 receivers spaced every 20 meters, are used to create the synthetic data. To simulate the acquisition nonreproducibility issues, we turn off receivers randomly for baseline and monitor surveys. We also assume there are 100 receivers missing in the monitor recording because of the presence of a platform right above the reservoir. Uniform random noise is added to the data. The synthetic data can be seen in Figure 2. For one shot gather, we have 397 traces for baseline, and 224 traces for monitor.

RTM image from baseline model is used as initial guess in LSRTM step for both baseline and monitor. Figure 3 shows the LSRTM results after 500 conjugate gradient (CG) steps. The amplitude difference at the reservoir is observable at this stage. Substracting the two images, as seen in Figure 4(a), we get very noisy results contaminated by footprints from the acquisition gap in the monitor survey. Nevertheless, the time-lapse signal can be interpreted.

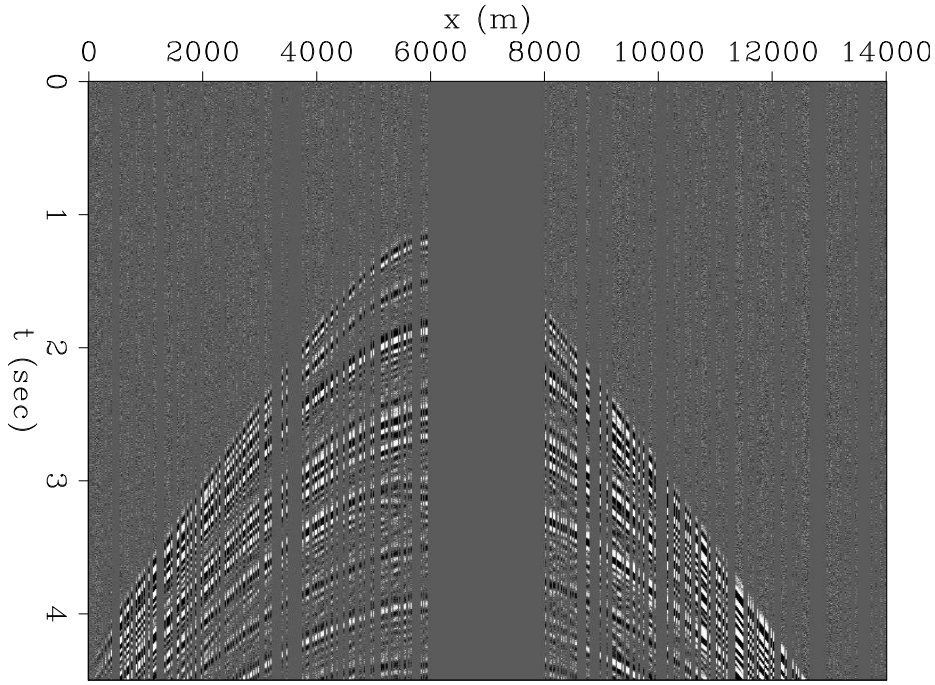
Simultaneous joint inversion is then solved using iteratively reweighted least squares (IRLS) for 20 by 10 CG steps, and the results are shown in Figure 4(b). Weight function $\mathbf{W}$ has non-zero values in a rectangle around the area with production induced velocity change. Because of the simplicity of the model, $\nabla_{\mathbf{r}} = \nabla_{\mathbf{x}}$ would be sufficient. In this example, 4D signals with L1 regularization can be recovered.

SYNTHETIC EXAMPLE 2

In this section, the target area is selected under the salt. Because of the complexity, we add acquisition nonrepeatability issues gradually to identify the boundary of failure and

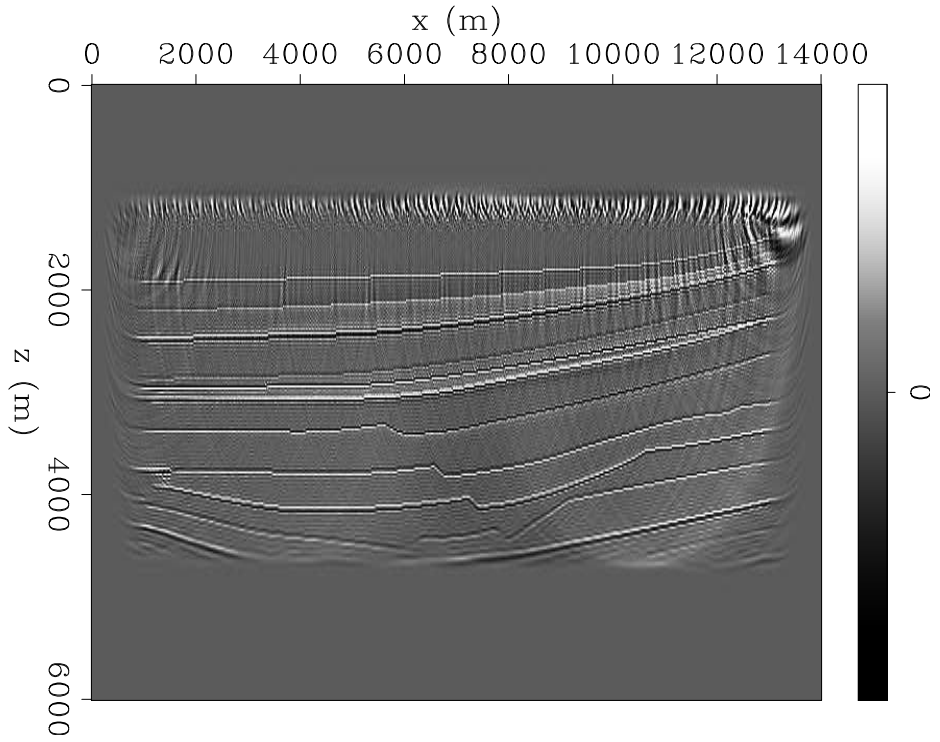


(a)

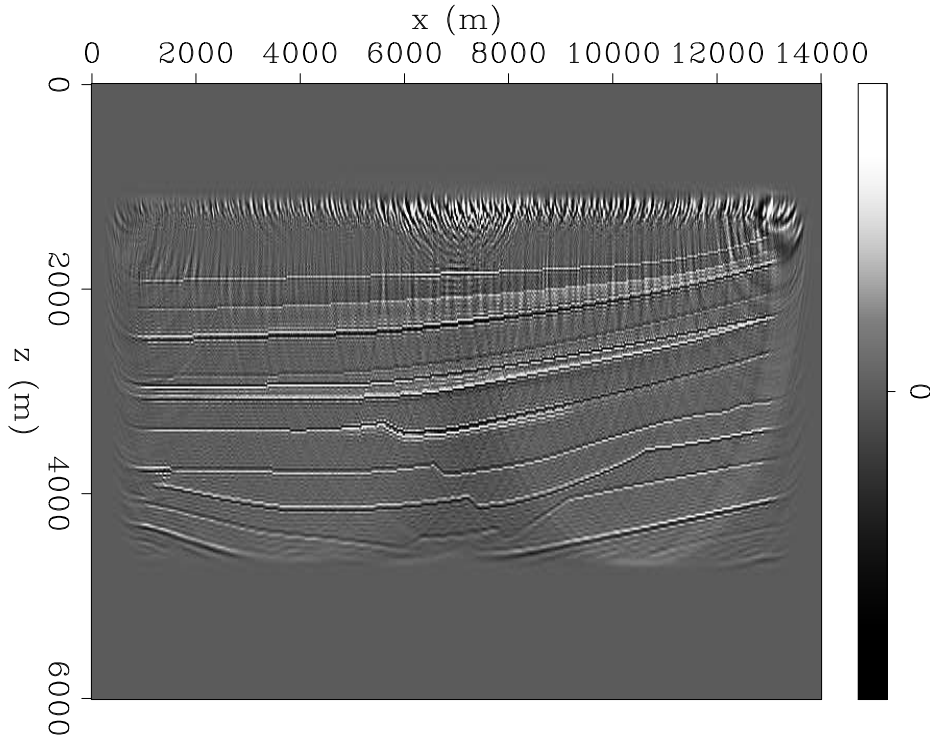


(b)

Figure 2: (a) baseline shot gather. (b) monitor shot gather. (same percentile clip) [**CR**] yinbin2/. simple.DataShot1,simple.DataShot2

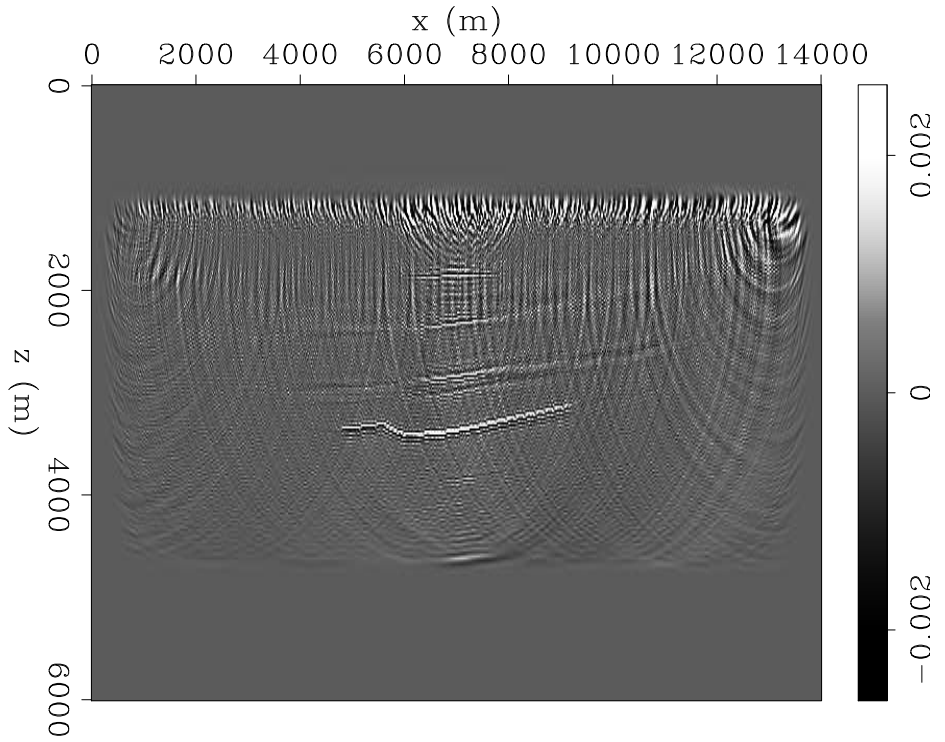


(a)

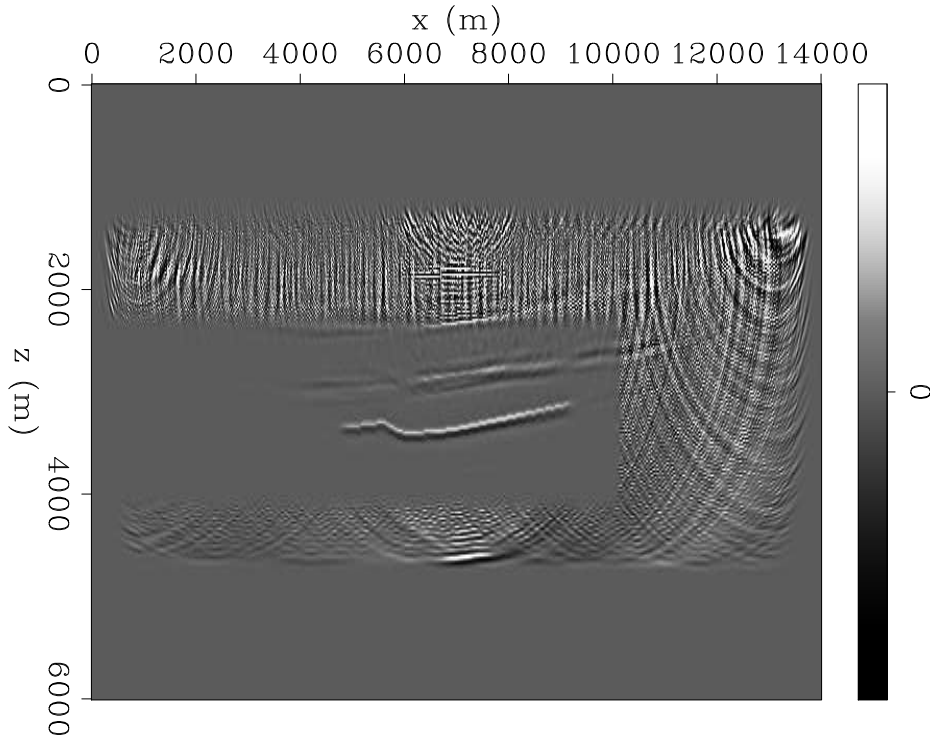


(b)

Figure 3: (a) baseline image from LSRTM. (b) monitor image from LSRTM. (same percentile clip) [**CR**] yinbin2/. simple.LSRTM1,simple.LSRTM2



(a)



(b)

Figure 4: (a) difference bewteen baseline and monitor images from LSRTM. (b) difference between baseline and monitor from Joint IRLS. The clean rectangle in the middle is where $\mathbf{W}(\mathbf{r}) \neq 0$. [**CR**] yinbin2/. simple.LSRTMDIFF,simple.IRLSRTMDIFF

success. The first subsection shows results with different acquisition geometries. The second subsection shows results with an additional acquisition gap in the monitor.

The true velocity models for baseline and monitor are constructed from part of the Sigbee 2A model with salt. Model size is 301×501 with 20 meters (m) between neighboring pixels. 48 sources spaced every 180 meters with up to 501 receivers spaced every 20 meters, are used to create the synthetic data. Migration velocity is the same for both baseline and monitor. The velocity differences between true velocity and migration velocity are shown in Figure 5. Two reservoirs are included in our model, one in the relatively easy part and another one under the salt.

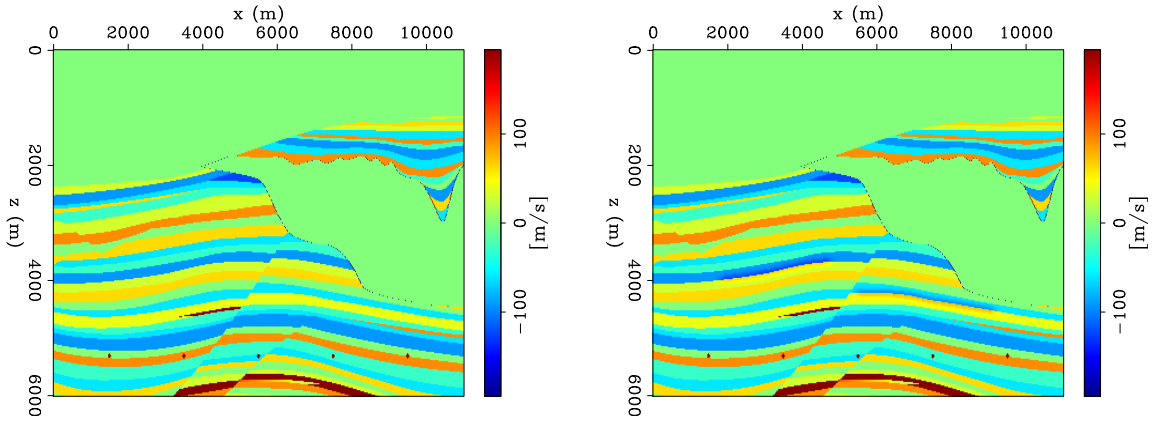


Figure 5: Left: baseline velocity perturbation; right: monitor velocity perturbation. [**ER**] yinbin2/. subsalt.diffT0,subsalt.diffT2

Subsalt 4D imaging with different acquisition geometry

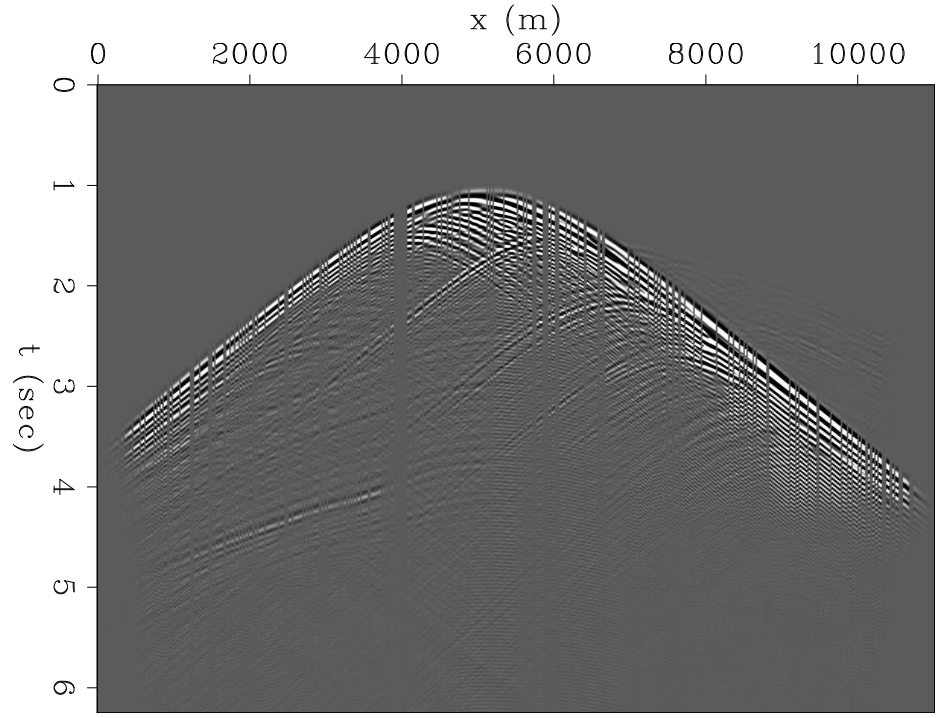
Each shot gather contains 383 traces for the baseline and 330 traces for the monitor. Two shot gathers are shown in Figure 6, from baseline and monitor survey seperately. RTM images from baseline and monitor are created seperately and used as initial guesses for LSRTM. Images from LSRTM in Figure 7 are computed after 200 CG steps. The predicted 4D signals are shown in Figure 8(a).

From the LSRTM results, a smooth gradient field is constructed using the method in Ma et al. (2015), and used as prior information for the joint inversion.

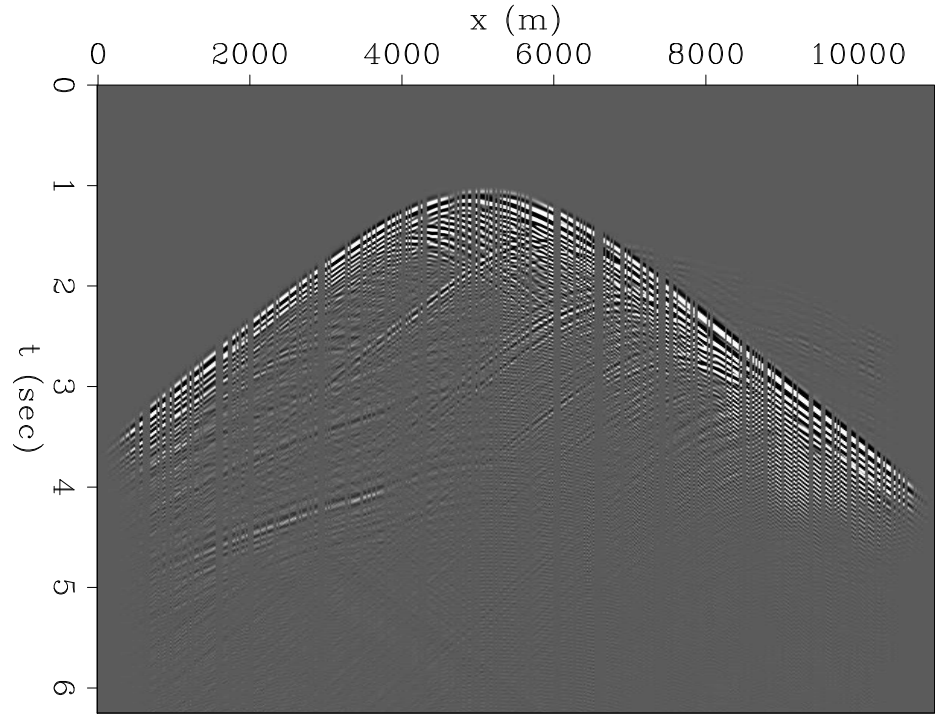
Finally, we solve the joint inversion problem using IRLS for 20 by 10 iterations, and the result can be seen in Figure 8(b).

Subsalt 4D imaging with different acquisition geometry and acquisition gap

50 receivers are removed from the monitor survey to create an acquisition gap. Two shot gathers are shown in Figure 9. RTM images from baseline and monitor are created seperately and used as initial guesses for LSRTM. Images from LSRTM in Figure 10 are computed

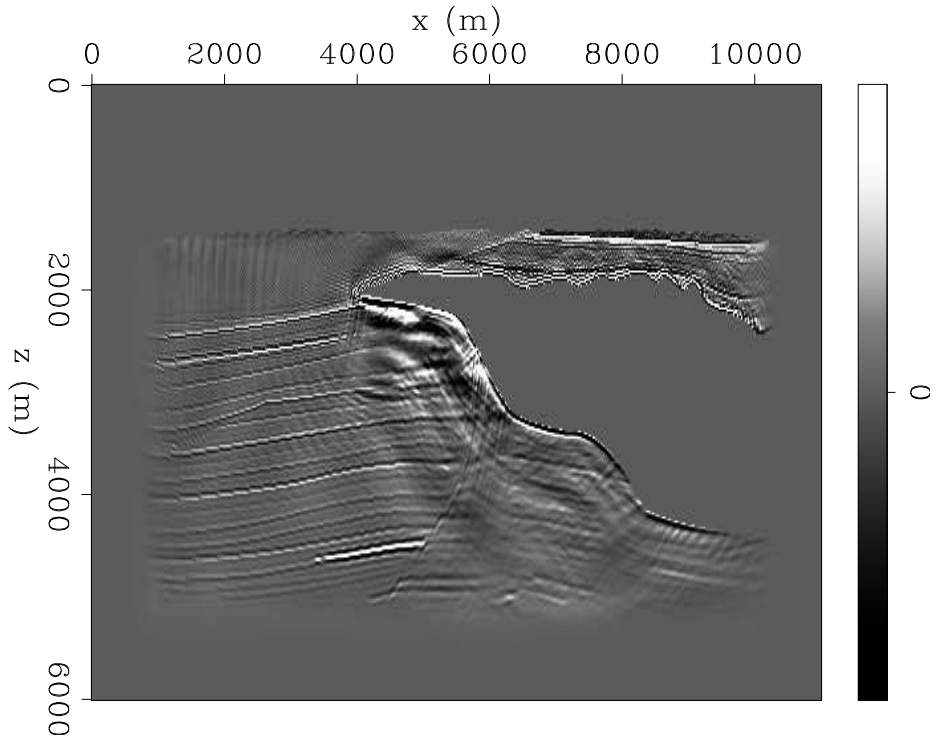


(a)

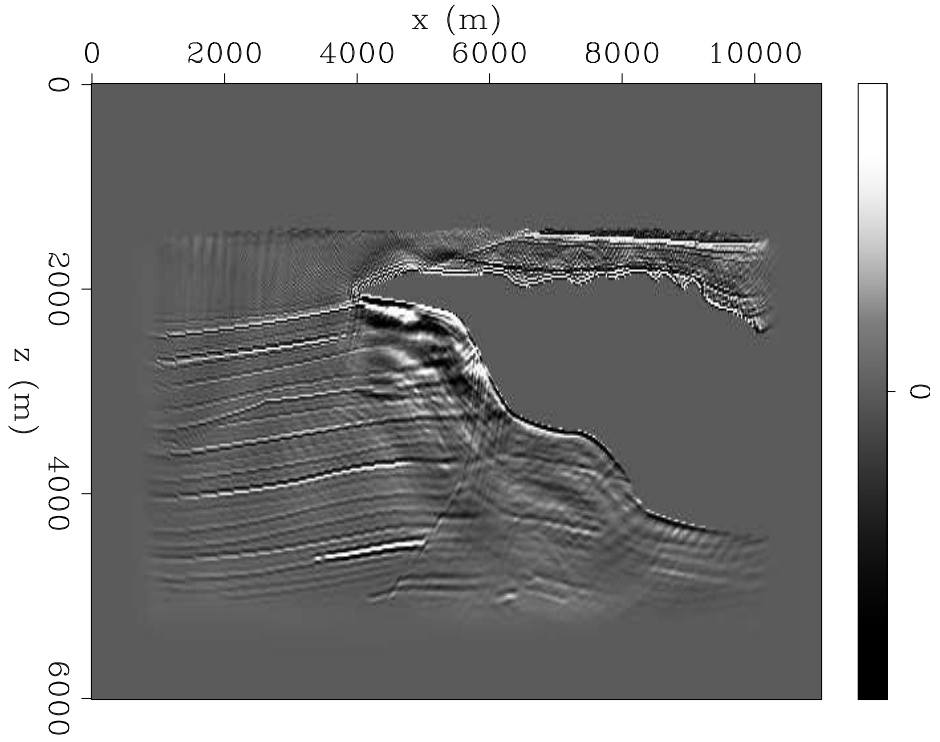


(b)

Figure 6: (a) baseline shot gather. (b) monitor shot gather. (same percentile clip) [**CR**] yinbin2/. subsalt1.DataShot1,subsalt1.DataShot2



(a)



(b)

Figure 7: (a) baseline image from LSRTM. (b) monitor image from LSRTM. (same percentile clip) [**CR**] yinbin2/. subsalt1.LSRTM1,subsalt1.LSRTM2

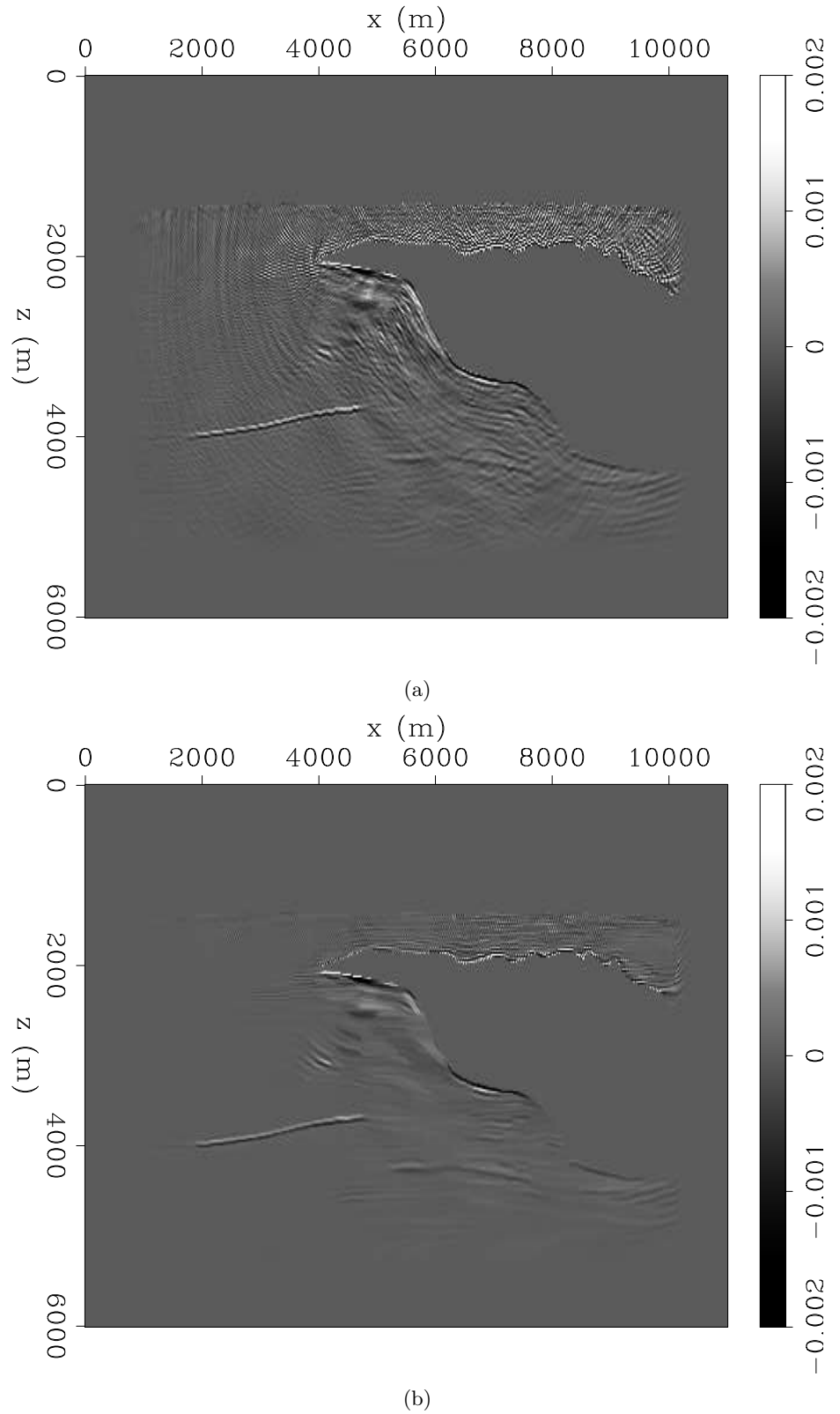


Figure 8: (a) difference bewteen baseline and monitor images from LSRTM. (b) difference between baseline and monitor from Joint IRLS. [**CR**] yinbin2/. subsalt1.LSRTMDIFF,subsalt1.IRLSRTMDIFF

after 200 CG steps. The predicted 4D signals are shown in Figure 11(a). The acquisition gap leads to a incorrect prediction of 4D signals under the salt. Around $x = 5000m$, LSRTM results suggest there are reflectivity changes both above and below the true velocity perturbation, near $z = 3000m$ and $z = 5000m$.

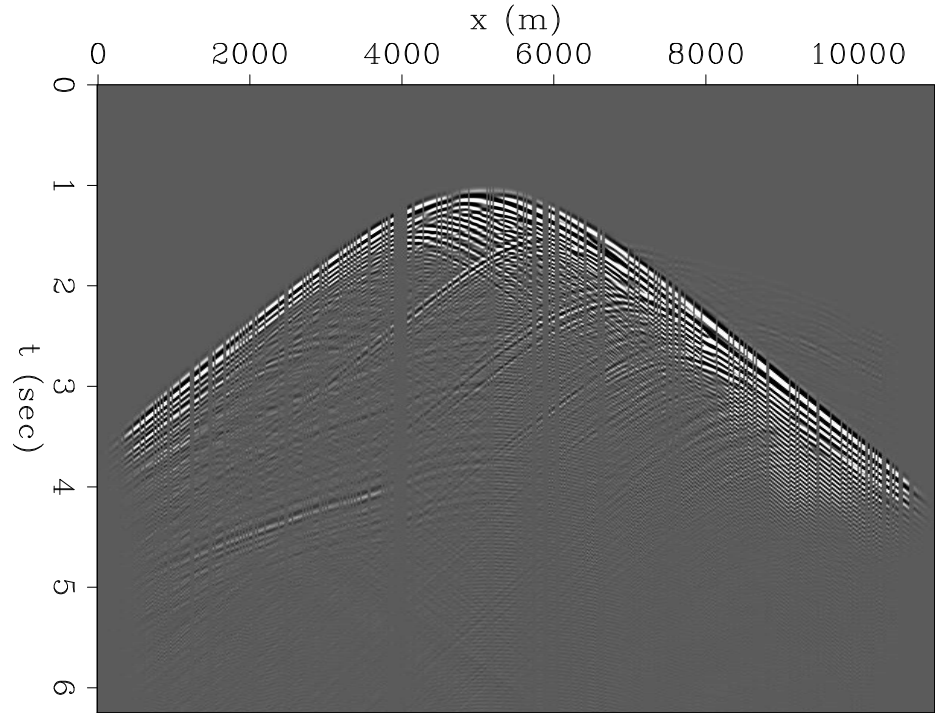
We then solve the joint inversion problem using IRLS for 20 by 10 iterations, and the result can be seen in Figure 11(b). The incorrect prediction of 4D signals is not completely removed from the joint inversion with steering filters, maybe because they are in the null space of steering filter and forward modeling operator $\mathbf{L}$. The limitation of steering filters indicates that more sophisticated regularization terms are necessary in time-lapse study.

CONCLUSIONS AND FUTURE WORK

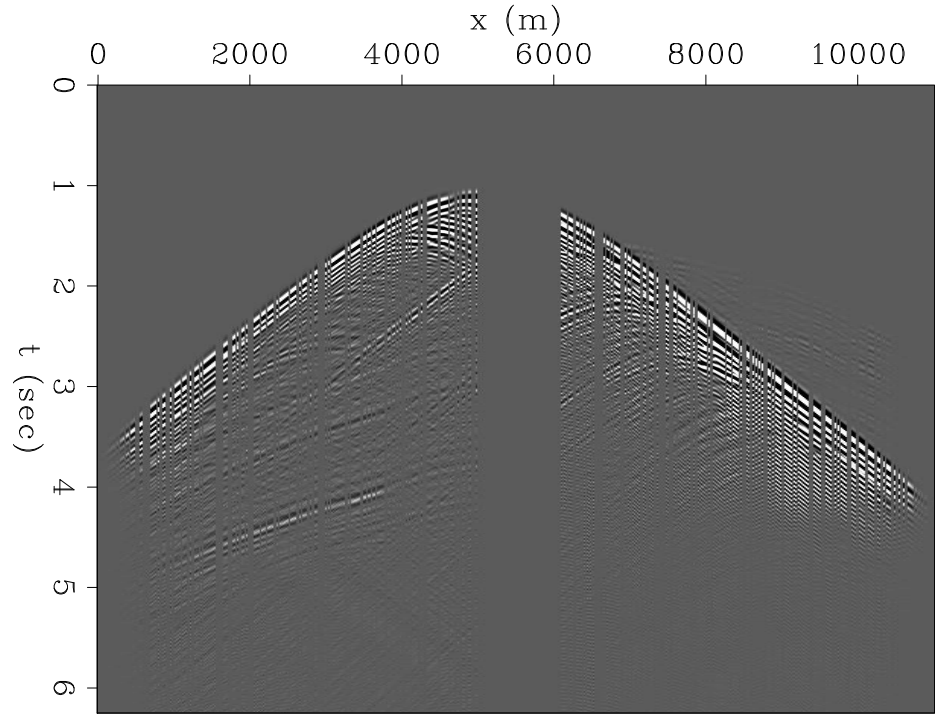
We applied L1 regularization and steering filters to the time-lapse imaging problem. Our numerical results suggest great improvements can be achieved when the subsurface structure is relatively simple. We also demonstrated enhancement of 4D signals under the salt can be achieved. The methods in this paper extract information purely from the seismic data. Subsalt imaging is still challenging when reliable data is not present. The bright side is that in 4D we have additional data from reservoir simulation, well logs, rock physics and geomechanics etc. In our next work, we need to bring in geomechanical information to build more robust algorithms.

ACKNOWLEDGMENTS

We would like to thank Huy Le, Gustavo Alves, Chris Leader and Stewart Levin for their valuable comments and numerious discussions.

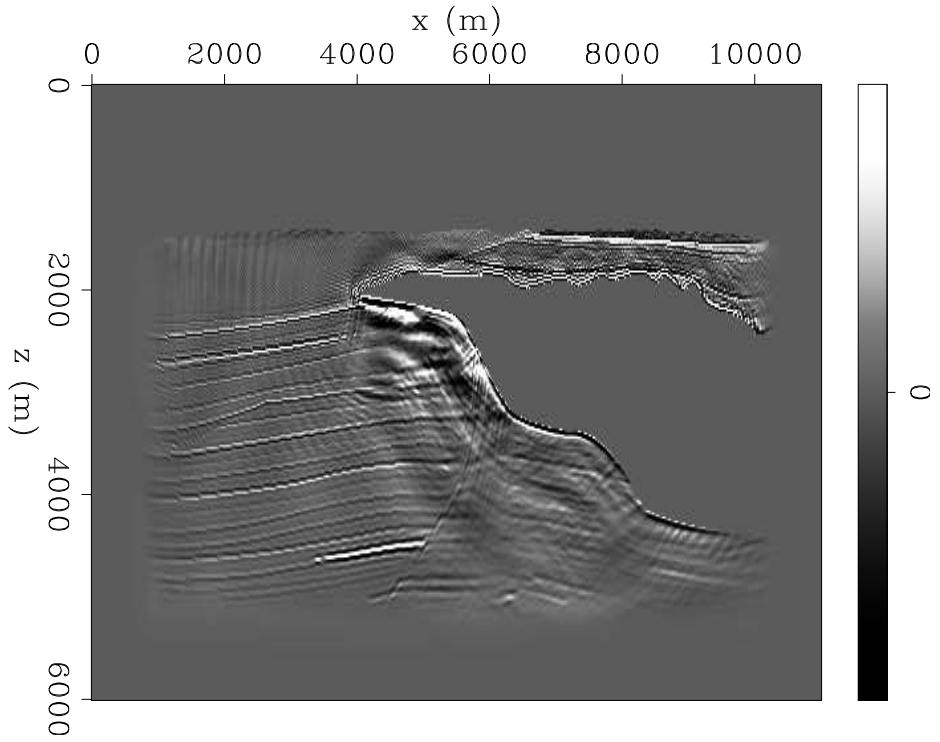


(a)

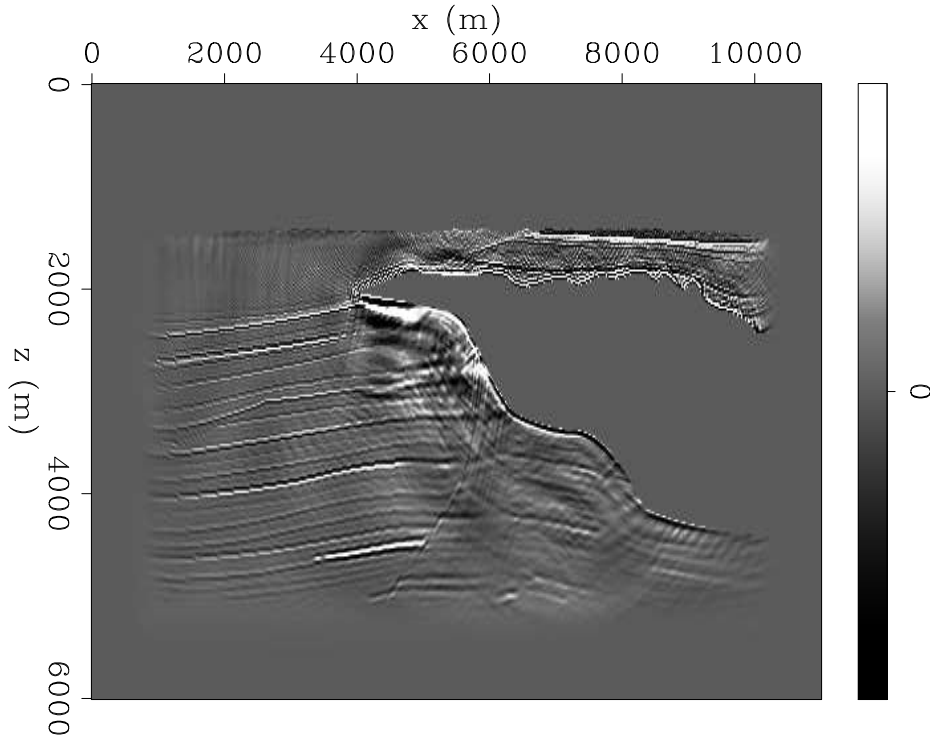


(b)

Figure 9: (a) baseline shot gather. (b) monitor shot gather. (same percentile clip) [**CR**] yinbin2/. subsalt2.DataShot1,subsalt2.DataShot2



(a)



(b)

Figure 10: (a) baseline image from LSRTM. (b) monitor image from LSRTM. (same percentile clip) [**CR**] yinbin2/. subsalt2.LSRTM1,subsalt2.LSRTM2

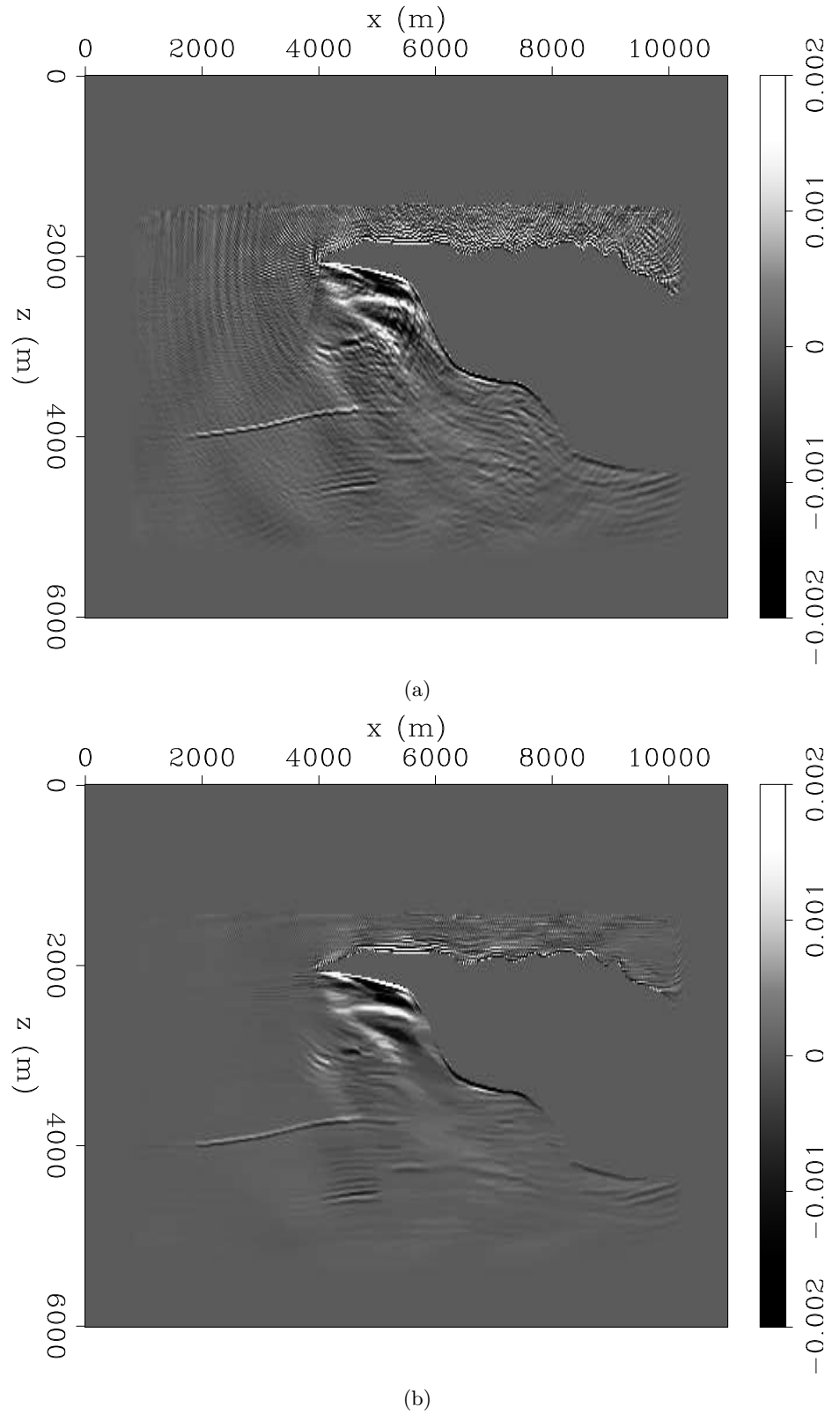


Figure 11: (a) difference bewteen baseline and monitor images from LSRTM. (b) difference between baseline and monitor from Joint IRLS. [**CR**] yinbin2/. subsalt2.LSRTMDIFF,subsalt2.IRLSRTMDIFF

APPENDIX A: FULL WAVEFORM INVERSION GRADIENT AND LINEARIZED VERSION

We aim at inverting velocity and density simultaneously. In this appendix, we derive the Frechet derivative of objective function for a full-waveform inversion. We follow the notations of (Liu and Tromp, 2006; Plessix, 2006; Bai and Yingst, 2014). We compute the gradient for a single-shot gather. Our objective function is:

$$J(v,\rho) = \frac{1}{2}\sum_r \int_0^T \|u(x_r,t) - d(x_r,t)\|_2^2 dt, \tag{15}$$

where $u(x_r,t)$ is a computed wavefield and $d(x_r,t)$ is recorded data. Our objective function is constrained by the following PDE:

$$\begin{cases} [\frac{1}{v^2}\partial_t^2 - \rho\nabla(\frac{1}{\rho}\nabla)]u & = f \\ u(x,0) & = 0 \\ \partial_t u(x,0) & = 0. \end{cases} \tag{16}$$

The Lagrangian:

$$\begin{aligned} L(v,\rho,\lambda,\mu_1,\mu_2) &= \frac{1}{2}\sum_r \int_0^T \|u(x_r,t) - d(x_r,t)\|_2^2 dt & (17)\\ &- \int_0^T \int \langle \lambda, \frac{1}{v^2}\partial_t^2 u - \rho\nabla(\frac{1}{\rho}\nabla u) - f\rangle d_\Omega^3 x dt & (18)\\ &- \int_\Omega \langle \mu_1, u(x,0)\rangle d^3x & (19)\\ &- \int_\Omega \langle \mu_2, \partial_t u(x,0)\rangle d^3x. & (20) \end{aligned}$$

Perturb Lagrangian and we can get:

$$\begin{aligned} \delta L &= \sum_r \int_0^T \int \langle u(x_r,t) - d(x_r,t), \delta u(x,t)\rangle \delta_3(x-x_r) d^3x dt & (21)\\ &- \int_0^T \int_\Omega \langle \lambda, -\frac{2}{v^3}\delta v \partial_t^2 u - \delta\rho\nabla(\frac{1}{\rho}\nabla u) + \rho\nabla(\frac{\delta\rho}{\rho^2}\nabla u) - \delta f\rangle d^3x dt & (22)\\ &- \int_0^T \int_\Omega \langle \lambda, \frac{1}{v^2}\partial_t^2 \delta u - \rho\nabla(\frac{1}{\rho}\nabla\delta u)\rangle d^3x dt & (23)\\ &- \int_\Omega \langle \mu_1, \delta u(x,0)\rangle d^3x & (24)\\ &- \int_\Omega \langle \mu_2, \partial_t \delta u(x,0)\rangle d^3x. & (25) \end{aligned}$$

Before we go further, we want to briefly mention the effects of each term in the previous equation. Line (22) is used to derive the Frechet derivative. Line (21) and (23) form another

PDE by properly choosing the Lagrangian multipliers, which are used in the computation of gradient. Line (24) and (25) vanish with proper μ_1 and μ_2.

We will first derive the PDE constraints for λ. Define A to be the sum of line (23), (24), and (25) of δL,

$$\begin{aligned} A &= -\int_0^T \int_\Omega \langle \lambda, \frac{1}{v^2}\partial_t^2 \delta u - \rho \nabla (\frac{1}{\rho}\nabla \delta u)\rangle d^3x dt \\ &= -\int_\Omega \langle \mu_1, \delta u(x,0)\rangle d^3x - \int_\Omega \langle \mu_2, \partial_t \delta u(x,0)\rangle d^3x \\ &= -\int_0^T \int_\Omega \langle \frac{1}{v^2}\partial_t \lambda - \nabla^\dagger(\frac{1}{\rho}\nabla^\dagger(\rho\lambda)), \delta u\rangle d^3x dt \\ &\quad -\int_\Omega \langle \mu_2 - \frac{1}{v^2}\lambda, \partial_t \delta u(x,0)\rangle d^3x - \int_\Omega \langle \mu_1 + \frac{1}{v^2}\partial_t\lambda, \delta u(x,0)\rangle d^3x \\ &\quad -\int_\Omega \langle \lambda, \frac{1}{v^2}\partial_t \delta u(x,T)\rangle d^3x \\ &\quad -\int_\Omega \langle -\partial_t \lambda, \frac{1}{v^2}\delta u(x,T)\rangle d^3x. \end{aligned} \tag{26}$$

We can choose $\mu_2 = \frac{1}{v^2}\lambda$, $\mu_1 = -\frac{1}{v^2}\partial_t\lambda$ to simplify the previous equation. Define B to be the sum of line 22 of δL, assuming $\delta f = 0$ and proper boundary condition, we get:

$$\begin{aligned} B &= -\int_0^T \int_\Omega \langle \lambda, -\frac{2}{v^3}\delta v \partial_t^2 u - \delta\rho\nabla(\frac{1}{\rho}\nabla u) + \rho\nabla(\frac{\delta\rho}{\rho^2}\nabla u) - \delta f\rangle d^3x dt \\ &= \int_0^T \int_\Omega \langle \lambda, \frac{2}{v^3}\partial_t^2 u\rangle \delta v d^3x dt \\ &\quad \int_0^T \int_\Omega \langle \lambda, \frac{1}{\rho}(\nabla\rho)(\nabla u)\rangle \delta\rho d^3x dt. \end{aligned} \tag{27}$$

Substitute A and B back into δL,

$$\begin{aligned} \delta L &= \int_0^T \int_\Omega \langle \lambda, \frac{2}{v^3}\partial_t^2 u\rangle \delta v d^3x dt \\ &\quad \int_0^T \int_\Omega \langle \lambda, \frac{1}{\rho}(\nabla\rho)(\nabla u)\rangle \delta\rho d^3x dt \\ &\quad -\int_0^T \int_\Omega \langle \frac{1}{v^2}\partial_t\lambda - \nabla^\dagger(\frac{1}{\rho}\nabla^\dagger(\rho\lambda)) - \sum_r (u(x_r,t) - d(x_r,t))\delta_3(x-x_r), \delta u(x,t)\rangle d^3x dt \\ &\quad -\int_\Omega \langle \lambda, \frac{1}{v^2}\partial_t \delta u(x,T)\rangle d^3x \\ &\quad -\int_\Omega \langle -\partial_t\lambda, \frac{1}{v^2}\delta u(x,T)\rangle d^3x. \end{aligned} \tag{28}$$

Choose the λ such that the last three lines vanish, and we get the gradients,

$$\begin{cases} \frac{\partial J}{\partial v} & = \int_0^T \int_\Omega \langle \lambda, \frac{2}{v^3} \partial_t^2 u \rangle d^3x dt \\ \frac{\partial J}{\partial \rho} & = \int_0^T \int_\Omega \langle \lambda, \frac{1}{\rho} (\nabla \rho)(\nabla u) \rangle d^3x dt, \end{cases} \tag{29}$$

subject to:

$$\begin{cases} \frac{1}{v^2} \partial_t \lambda - \nabla^\dagger (\frac{1}{\rho} \nabla^\dagger (\rho \lambda)) & = \sum_r (u(x_r, t) - d(x_r, t)) \delta_3(x - x_r) \\ \lambda(x, T) & = 0 \\ \partial_t \lambda(x, T) & = 0. \end{cases} \tag{30}$$

λ satisfies the adjoint of wave equation, and we emphasize that $\nabla^\dagger \neq \nabla$.

In this paper, we only need linearized waveform inversion with constant density, and we can simplify the gradient expression:

$$\frac{\partial J}{\partial v} = \int_0^T \int_\Omega \langle \lambda, \frac{2}{v^3} \partial_t^2 u \rangle d^3x dt \tag{31}$$

$$\frac{1}{v^2} \partial_t \lambda - (\nabla^\dagger)^2 \lambda = \sum_r (u(x_r, t) - d(x_r, t)) \delta_3(x - x_r) \tag{32}$$

$$\lambda(x, T) = 0 \tag{33}$$

$$\partial_t \lambda(x, T) = 0 \tag{34}$$

$$\rho(x) = \rho_0. \tag{35}$$

APPENDIX B: UNDERSTANDING FORWARD/ADJOINT OF WAVE PROPAGATION IN MATRIX FORM

In this section, Born modeling for constant density in matrix form is derived, following the idea of (Ji, 2009; Almomin, 2013) . All the necessary modules are included and they should be ready to translate into real code.

We discrete time into $[0, \Delta t, 2\Delta t, \cdots, N\Delta t]$, with $T = N\Delta t$. Assume we have obtained the wavefield $\tilde{u}(x, t)$ for $t = 0, \Delta t, \cdots, N\Delta t$, we can then approximate $\partial_t^2 u_0(x, t)$ for $t = 0, \Delta t, \cdots, (N-1)\Delta t$, using a second order finite difference.

Use u^i to represent the wavefield at $t = i \times \Delta t$, and use $U = [u^0, u^1, \cdots, u^N]^T \in \mathbb{R}^{N \times dom}$ to represent wavefield at the domain in which we are interested. First, we need to create the source term $S = [s^0, s^1, \cdots, s^{N-1}, s^N = 0]^T$, it can be described as:

$$S = BVm, \tag{36}$$

where $B = [D_t^2 u^0, D_t^2 u^1, \cdots, D_t^2 u^{N-1}, 0]^T$, and $V = v(r)^2$. D_t^2 is 2nd order finite difference operator.

An implicit trick we want to impose is a shift operator T_1 with the effect:

$$T_1 S = T_1 \begin{bmatrix} s^0 \\ s^1 \\ \vdots \\ s^{N-1} \\ s^N = 0 \end{bmatrix} = \begin{bmatrix} 0 \\ s^0 \\ s^1 \\ \vdots \\ s^{N-1} \end{bmatrix}. \tag{37}$$

We need this operator to compute $u_1, \cdots, u_N$; and we need $s_0, \cdots, s_{N-1}$, and the shift operator to simplify the following expression. It is also important to obtain the correct final state for the adjoint operator.

The explicit form for time marching is:

$$u^{t+1} = (2 - \frac{\Delta t^2}{\Delta x^2} v^2 \Delta) u^t - u^{t-1} + s^t. \tag{38}$$

After the source wavefield is obtained, we claim the following operator is time marching:

$$M_t \equiv \begin{bmatrix} I_{(t-1)\times(t-1)} & 0 & 0 \\ 0 & A_{3\times 3} & 0 \\ 0 & 0 & I_{(N-t-2)\times(N-t-2)} \end{bmatrix}. \tag{39}$$

which maps $[u^0, \cdots, u^t, s^t, \cdots, s^{N-1}]^T$ to $[u^0, \cdots, u^t, u^{t+1}, s^{t+1}, \cdots, s^{N-1}]$. The time marching kernel:

$$A = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -P & PT & P \end{bmatrix} \tag{40}$$

$$T = 2 - \frac{\Delta t^2}{\Delta x^2} v^2 \Delta, \tag{41}$$

with P as absorbing boundary operator. The effect of M_t can be checked by induction, and the boundary term need some attention.

Now, we have:

$$M_{N-1} M_{N-2} \cdots M_0 T_1 B V m = U = \begin{bmatrix} u^0 = 0 \\ u^1 \\ \vdots \\ u^{N-1} \\ u^N \end{bmatrix}. \tag{42}$$

Thus, it is obvious that $M_{N-1} M_{N-2} \cdots M_0 T_1 B V$ is the forward operator for Born modeling.

It is interesting to see the exact adjoint of this method:

$$T_1^T M_0^T \cdots M_{N-1}^T \begin{bmatrix} u^0 \\ u^1 \\ \vdots \\ u^{N-1} \\ u^N \end{bmatrix} = T_1^T \begin{bmatrix} s^0 = \text{weird state} \\ s^1 \\ \vdots \\ s^{N-1} \\ s^N \end{bmatrix} = \begin{bmatrix} \tilde{s}^0 \\ \tilde{s}^1 \\ \vdots \\ \tilde{s}^{N-1} \\ \tilde{s}^N = 0 \end{bmatrix}, \tag{43}$$

which is exactly what the adjoint formula predicts, with the correct final condition.

REFERENCES

Almomin, A., 2013, Accurate implementation of two-way wave-equation operators: SEP Report, **149**, 279–286.

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

———, 2011, Wave-equation inversion of time-lapse seismic data sets: SEP Report, **143**, 117–134.

Bai, J. and D. Yingst, 2014, Simultaneous inversion of velocity and density in time-domain full waveform: SEG Technical Program Expanded Abstracts 2014, 922–927.

Ji, J., 2009, An exact adjoint operation pair in time extrapolation and its application in least-squares reverse-time migration: Geophysics, **74**, H27–H33.

Liu, Q. and J. Tromp, 2006, Finite-frequency kernels based on adjoint methods: Bulletin of the Seismological Society of America, **96**, 2383–2397.

Ma, Y., M. Maharramov, R. Clapp, and B. Biondi, 2015, Illumination compensation with L1 regularization and steering filters: SEP Report, **158**.

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

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

Acquiring rotation data on the ocean bottom without rotation sensors

Ohad Barak, Kerry Key, Steven Constable, Paul Milligan and Shuki Ronen

ABSTRACT

There are currently no widely available rotation sensors that can operate on the ocean-bottom. We derive rotation data on the ocean bottom from two surveys that were not originally designed to record them: 1) from geophone recordings in the Moere Vest ocean-bottom survey by differencing adjacent geophones; and 2) from magnetometer recordings in the SERPENT CSEM ocean-bottom survey by extrapolating from the deviations in magnetic field projections on the magnetometer components.

INTRODUCTION

Rigid bodies in a three dimensional world have six degrees of freedom: three components of translation and three components of rotation. The time derivatives of translations are the particle velocities and the rotations are the pitch, roll and yaw, as shown in the following table:

Axis	Translation		Rotation	
Z	Vertical	v_z	Yaw	r_z
X	Radial	v_x	Roll	r_x
Y	Transverse	v_y	Pitch	r_y

where v_i are particle velocities along the i axis, and r_i are rotation rates around the i axis.

In ocean-bottom node acquisition, multicomponent geophones that are coupled to the seafloor record the vertical and the two horizontal components of particle velocity $\vec{v}$. Additionally, a hydrophone records the divergence of the translation wavefield $P = \kappa\,(\nabla \cdot \vec{u})$, where $\vec{u}$ are particle displacements and κ is the bulk modulus of the water to which the hydrophones are coupled. Rotation rates are a measurement of the curl of the particle velocity wavefield $\vec{r} = \frac{1}{2}(\nabla \times \vec{v})$, and are a recording of the anti-symmetric strains of the medium (Cochard et al., 2006).

Vassallo et al. (2012) use hydrophones together with pressure gradient sensors in marine streamer acquisition to interpolate the pressure wavefield in the crossline direction, between streamer cables. Similarly, the rotational components can be used to interpolate vertical geophone data (Edme et al., 2014), and spatial aliasing of high-wavenumber arrivals can thus be mitigated. Barak et al. (2014b) show that rotation data are extra information, are independent of geophone data, and can be used in conjunction with geophone data to identify and separate wave-modes on land using singular-value decomposition polarization analysis.

As of yet there are no industry-grade solutions for recording rotational motion on the ocean bottom, though a few such recording stations have been deployed previously by Pillet et al. (2009). The objective of this paper is to show how rotation data can be extracted from existing ocean-bottom recordings.

THE MOERE VEST OCEAN-BOTTOM SURVEY

The Moere Vest data include a group of 26 four-component ocean-bottom nodes, a "microspread," which have a unique geometry in that they are spaced at 2 m intervals. We estimated the three-component rotational motion by differencing adjacent geophones of these microspread nodes. Geophone differencing as a method of estimating rotational motion has been shown previously in Muyzert et al. (2012) and Barak et al. (2014a). In the case of the microspread, the short 2 m interval between receivers ensures that most of the data are sampled well enough to prevent spatial aliasing. Therefore, we assume that a differencing of the data recorded by adjacent nodes pertains to differences of displacements within half a wavelength.

Though the acquired data may be of high quality, any differencing of data coming from physically separate sensors must be done with the caveat that we are in effect decreasing the signal to noise ratio in the resulting differenced data. Each sensor may have a slightly different coupling to the medium, reducing the reliability of the difference signal. Also, the data and the data-difference are not collocated in space. The proper way of obtaining a recording of any physical quantity is to design a sensor that can measure that particular quantity directly at one point in space. The resulting rotation data we get from differencing are an estimate of the data which would have been recorded with rotation sensors. We are able to obtain this estimate due to the special geometry of the microspread.

Estimating rotational motion from geophone data

The stress-displacement relation for tangential stresses reads:

$$\sigma_{ij} = \mu \left(\partial_j u_i + \partial_i u_j \right), \tag{1}$$

where σ_{ij} are the tangential stresses, u_i are particle displacements and μ is the shear modulus.

At a free surface, or at an interface between a medium with shear strength and one without shear strength (such as the ocean-bottom interface), the tangential stresses σ_{ij} are zero. Therefore, assuming we have receivers laid out on a flat, horizontal sea bottom, we have

$$\begin{aligned} \partial_z u_y &= -\partial_y u_z, \\ \partial_z u_x &= -\partial_x u_z, \end{aligned} \tag{2}$$

meaning that the vertical derivative of the horizontal displacement component is equal to the horizontal derivative of the vertical displacement component.

Rotation is defined as the curl of the wavefield. Since our geophones record the time derivative of displacement (particle velocity), we use the time derivative of rotation, or rotation rate:

$$\begin{aligned}\vec{r} = \frac{1}{2}(\nabla \times \vec{v}) &= \hat{X}\left(\partial_y v_z - \partial_z v_y\right)/2 \\ &+ \hat{Y}\left(\partial_z v_x - \partial_x v_z\right)/2 \\ &+ \hat{Z}\left(\partial_x v_y - \partial_y v_x\right)/2, \end{aligned} \tag{3}$$

where $\hat{X}$, $\hat{Y}$ and $\hat{Z}$ are the rotation axes. Substituting equation 2 into 3, we see that on the sea bottom

$$r_x = \partial_y v_z, \quad r_y = -\partial_x v_z, \quad r_z = \frac{1}{2}\left(\partial_x v_y - \partial_y v_x\right), \tag{4}$$

i.e., the horizontal rotation-rate components can be derived from the vertical geophones, and the vertical rotation-rate component can be derived from the horizontal geophones.

Microspread geometry

Since we intend to perform geophone differencing, the receiver positions are of high importance. An error in positioning could lead to an error in the derived rotation data. There were two sets of fields in the SEGY files that indicated receiver positions. One of them was the "as-laid" positions, which are the coordinates of the underwater Remotely Operated Vehicle (ROV) that deployed the nodes on the sea bottom. The other set of receiver positions were calculated using the first-break arrival time at each node from all shots in the survey. Figure 1(a) shows these two sets of receiver positions.

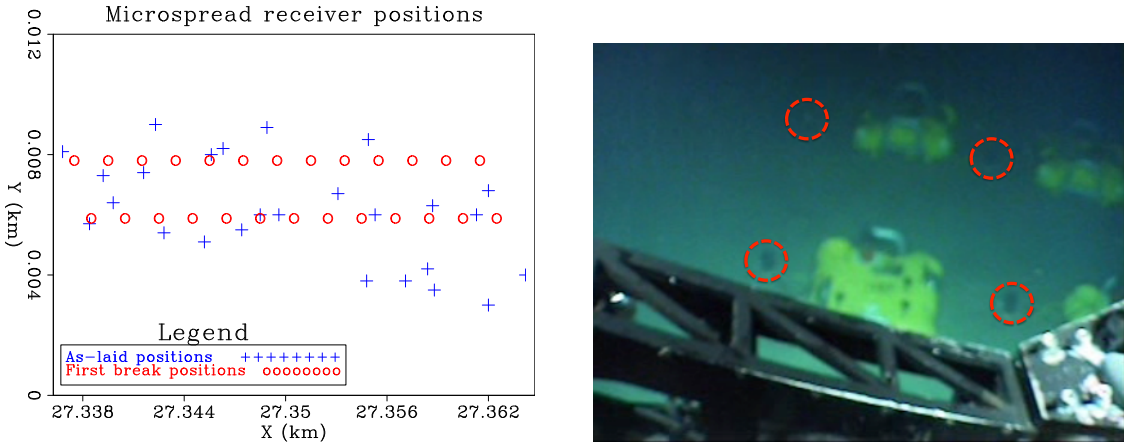
Despite the fact that the first-break positions appear too "regular", we opted to use these node positions for the microspread, since we also have video footage of the node deployment showing a very regular geometry. An example is shown in Figure 1(b). The ability of the ROV operator to see all nodes during deployment contributed to the positioning accuracy.

The length of the shot line we used was 55 km, with a shot interval of 50m, traversing almost directly above the nodes of the microspread. The microspread nodes were positioned near the center of the shot line. More than 90% of the shots have less than 5 m of crossline offset. Since the receivers were at a depth of 1.6 km, the data are effectively 2D. We removed the receiver instrument signature and aligned the horizontal geophone components to the 2D survey coordinates.

6-component data from the Moere Vest survey

To generate the three-component rotation-rate data we differenced adjacent receiver stations, effectively executing a finite-difference approximation to equations 4. Figures 2(a), 2(c) and 2(e) are the vertical (v_z), inline (v_x) and crossline (v_y) geophone components of the receiver gather of one node of the microspread. Figures 2(b), 2(d) and 2(f) are the yaw (r_z), roll (r_x) and pitch (r_y) rotational components. Notice that adjacent to each particle velocity component is the rotational component around that geophone's axis.

We are not displaying the direct arrival and some of the associated bubbles that are between $t = 1.08$s and $t = 1.75$s. The water-bottom multiple appears at $t = 3.25$s, and can be seen on the v_z and v_x sections.



(a) (b)

Figure 1: (a) Microspread receiver "as-laid" positions (blue) vs. first-break positions (red circles). The length of the receiver line is 26 m, and the nominal receiver spacing is 2 m. (b) Photo taken by ROV operator showing four of the ocean-bottom nodes of the microspread. The sensors are indicated by the red circles. Note the regularity of the deployment. [**NR**] ohad1/. gxgy-all,node1587-photo

The v_z component seems to contain mostly high-frequency reflections with a moveout consistent with P-wave velocity, but there are some lower frequency events that have a much slower moveout. These events are commonly called "Vz noise", and may be caused by shear-wave scattering in an inhomogeneous seabed, which in turn can generate Scholte waves on the seabed. The v_x component contains mostly those shear-induced events, but some of the P-wave events apparent on v_z are also visible. The v_y component is much weaker than the other two geophone components, but a shear-induced event similar to the one on the v_x component at $t = 3.3$s is prominent. It is possible that the shear energy is indeed coming from shear-wave reflections, however on a receiver gather it is difficult to tell the difference between reflected shear waves and scattered Scholte waves based solely on moveout information.

Observing the rotational components, we see that the one with the greatest energy is r_y. This fits with our expectation. Since the survey geometry is practically 2D, most of the translation should occur in the vertical and inline directions, which means that most of the rotational motion should occur around the crossline direction. Note also the generally increased noise level on the rotational components, which we attribute to the geophone differencing operation.

Compare the v_z and r_y components, and observe how the P-waves are almost not visible on r_y, even though this section was obtained using two adjacent vertical geophones. This indicates that the P-waves generate a similar response on adjacent vertical geophones, and are removed by the differencing. Another way of saying the same thing is that the P-waves do not generate a strong rotational deformation of the surface. Instead, we see a section that is more similar to v_x, with events that have shear-wave moveouts (though slightly delayed compared to v_x). Shear and Scholte waves generate a shear deformation of the surface, which manifests itself as rotational motion. Therefore, rotation data should preferentially

record shear waves, and indeed the P events on the r_y component are much weaker than those visible on the v_z and v_x components.

The r_x rotational component seems to also contain some shear-induced events. They are slightly weaker than the events on r_y, indicating that if these are indeed the result of scattered shear waves, then these waves are causing mainly rotation around the crossline axis. The r_z section is the weakest of the rotations (by a factor of 2 compared with r_y). This component should record events that cause a horizontal deformation around the vertical axis. The energy on the r_z component seems to also be related to the shear-induced waves. One explanation could be that multiple scatterings in the near surface are generating horizontal shear deformations of the medium.

THE SERPENT ELECTROMAGNETIC SURVEY

The SERPENT controlled source electromagnetic survey (Key et al., 2012; Naif et al., 2013) took place offshore Nicaragua on May 2010. 55 ocean-bottom EM nodes were deployed along a line perpendicular to a sudbuction zone. Each node had two horizontal induction-coil magnetic field sensors, and two horizontal electric field sensors. A composite node also had a three-component geophone in addition to the EM sensors. Information from the USGS website regarding this earthquake is shown in Figure 3, displaying the earthquake, node line and composite node positions.

As expected, the earthquake was recorded on the geophone components of the composite node. The data are shown in Figure 5(a). The P-wave arrives at about $t = 3186.5$s, and the S-wave comes in 2 seconds later. Unexpectedly however, the data of the two magnetic components in Figure 5(b) show a remarkably similar behaviour to the geophone data.

We interpret the magnetic data as resulting from a rotation of the ground caused by the seismic waves generated by the earthquake. The node body is coupled to the seabed via a 150 kg slab of concrete. Therefore, the magnetic sensors in the node body rotate together with the ground. The Earth's magnetic field, however, does not rotate and is effectively constant in direction and in amplitude for the duration of the earthquake. The ground rotation therefore manifests itself as a change in the projection of the Earth's magnetic field on the node's magnetic sensor components. An illustration of this is shown in Figure 4. This concept was explored previously in Kappler et al. (2006) using land data recorded by USGS permanent stations.

From magnetic field projections to ground rotations

After designature, the magnetometer data are in terms of deviation of the magnetic field strength on the two horizontal components over time:

$$\Delta\vec{H}(t) = [\Delta H_x(t), \Delta H_y(t)]. \tag{5}$$

We deduced the vertical magnetic component by rotating the horizontal node components to geographic North. We then added the ambient magnetic field values for the North and East magnetic components, as given by the British Geological Survey's World Magnetic Model, to the rotated magnetic horizontal components. We then use the total magnetic

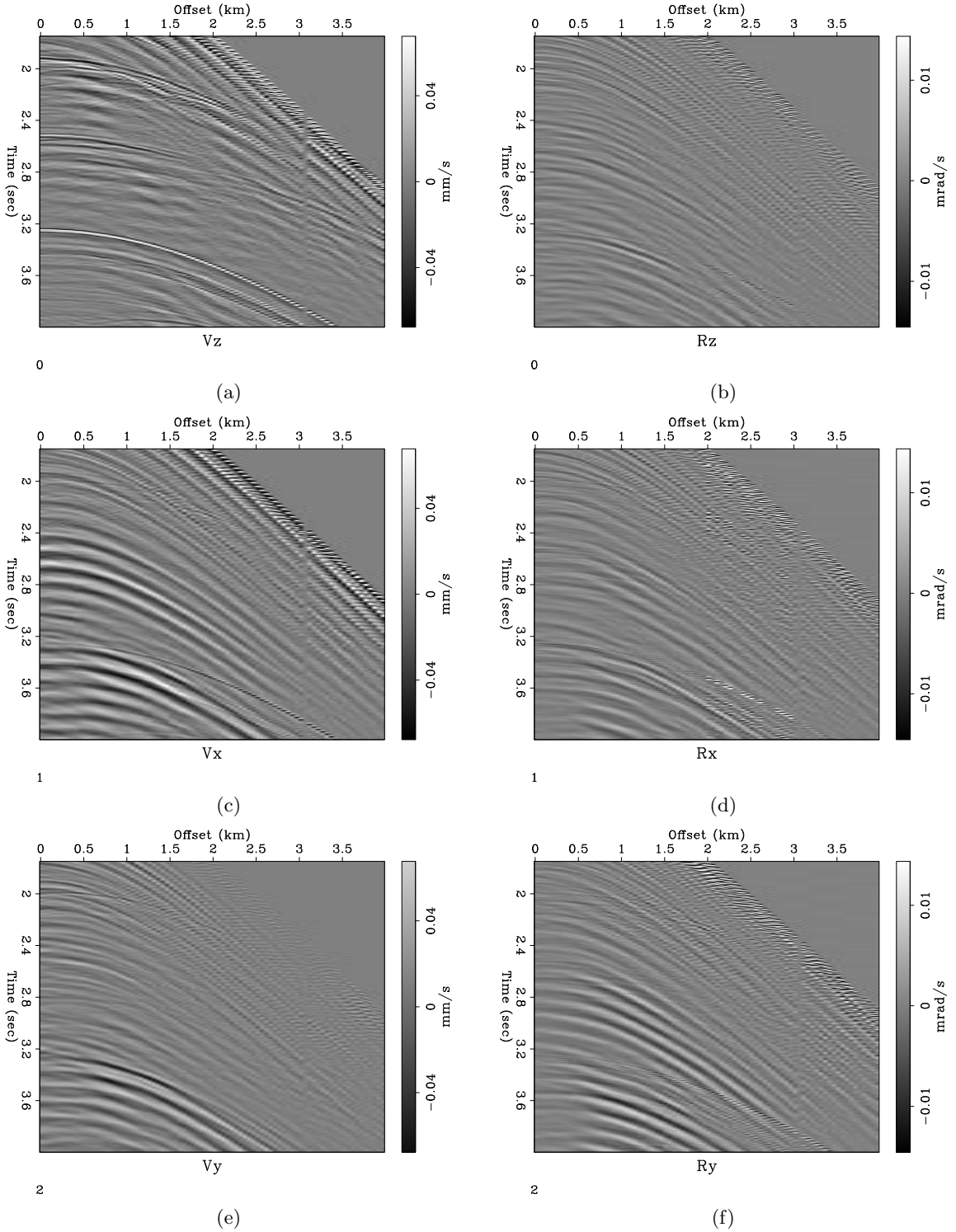


Figure 2: Six-component receiver gather of one node of the microspread of the Moere Vest data. (a) Vertical particle velocity v_z. (b) Yaw rotation r_z. (c) Inline particle velocity v_x. (d) Roll rotation r_x. (e) Crossline particle velocity v_y. (f) Pitch rotation r_y. The rotational components contain mostly the shear-induced energy, and have a lower signal to noise ratio. [**CR**]

ohad1/. 6c-node1730-vz,6c-node1730-rz,6c-node1730-vx,6c-node1730-rx,6c-node1730-vy,6c-node1730-ry

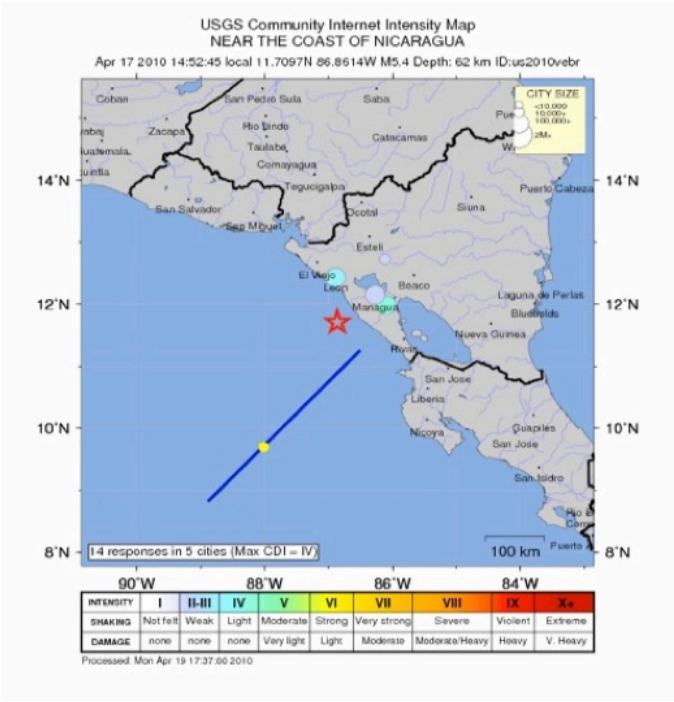


Figure 3: USGS report pertaining to the earthquake (indicated by the red star) that occurred during the SERPENT EM survey. The blue line indicates the EM ocean-bottom node line, where the yellow dot is the position of the composite sensor which had a three-component geophone in addition to the EM sensors. [**NR**] ohad1/. M54EQ

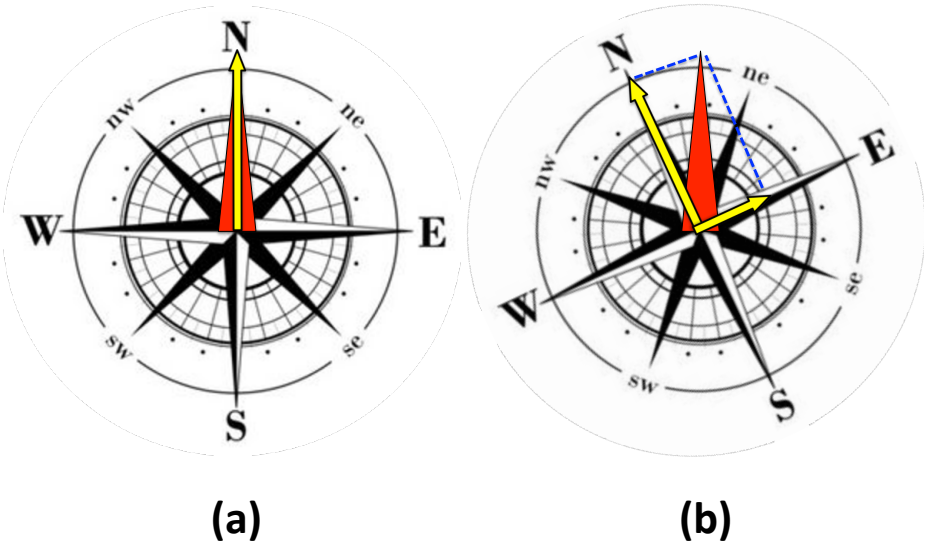


Figure 4: Illustration of how ground rotation is recorded on the magnetic field sensor as represented by the compass, which is coupled to the ground. (a) Before ground rotation, the magnetic field (red) is recorded only by the North component (yellow). (b) During ground rotation, the magnetic field (red) does not change, but its projection on the North and East components (yellow) changes. We can calculate the amount of rotation from the change in projection. Note that translations of the ground will not result in a change of the projection of the magnetic field on the magnetic components. [**NR**] ohad1/. compass-rotation

field strength at the location of the node and at the time of the survey to obtain the vertical magnetic component:

$$H_z(t) = \sqrt{H^2_{\text{total}} - H^2_x(t) - H^2_y(t)}. \tag{6}$$

The angle of rotation between consecutive time steps can now be calculated by

$$\theta(t) = \cos^{-1}\left(\frac{\vec{H}(t+\Delta t)\cdot\vec{H}(t)}{\left|\vec{H}(t+\Delta t)\right|\left|\vec{H}(t)\right|}\right), \tag{7}$$

while the unit vector describing the axis of rotation is

$$\hat{\nu}(t) = \frac{\vec{H}(t)\times\vec{H}(t+\Delta t)}{\left|\vec{H}(t)\times\vec{H}(t+\Delta t)\right|}. \tag{8}$$

In order to have the rotation-rate data in terms of Euler angles in the reference frame of the geophone component axes, we must use a quaternion representation. Our quaternion four-vector system state begins with no rotation, i.e.

$$\mathbf{q}_{t=0} = \begin{pmatrix} q_w \\ q_x \\ q_y \\ q_z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{9}$$

We use equations 7 and 8 to get the rotation angle θ and the rotation axis $\vec{u}$, and then we convert to a quaternion representation by

$$\mathbf{p}(t) = \begin{pmatrix} \cos\frac{\theta(t)}{2} \\ u_x(t)\cdot\sin\frac{\theta(t)}{2} \\ u_y(t)\cdot\sin\frac{\theta(t)}{2} \\ u_z(t)\cdot\sin\frac{\theta(t)}{2} \end{pmatrix}. \tag{10}$$

In order to rotate our system from its state at time t to its new state at time $t+\Delta t$, we need to apply quaternion multiplication ($\star$) of the quaternion $\mathbf{q}$ by $\mathbf{p}$:

$$\mathbf{q}(t+\Delta t) = \mathbf{p}(t+\Delta t)\star\mathbf{q}(t) = \left(p_w q_w - \vec{p}\cdot\vec{q},\ p_w\vec{q} + q_w\vec{p} + \vec{p}\times\vec{q}\right). \tag{11}$$

We can now retrieve the change in rotation for every time step in terms of Euler angles around each axis using:

$$\Delta\vec{r}(t) = \begin{pmatrix} \arctan\left(\frac{q_y q_z + q_w q_x}{\frac{1}{2} - \left(q_x^2 + q_y^2\right)}\right) \\ \arcsin\left(-2\left(q_x q_z - q_w q_y\right)\right) \\ \arctan\left(\frac{q_x q_y + q_w q_z}{\frac{1}{2} - \left(q_y^2 + q_z^2\right)}\right) \end{pmatrix}. \tag{12}$$

To get the rotation rate, we must divide $\Delta\vec{r}$ by the time step: $\dot{\vec{r}}(t) = \frac{\Delta\vec{r}}{\Delta t}$

Six-component earthquake data

Figure 5(c) shows the three rotational components as derived from the changes in the projection of the magnetic field on the magnetometer components. The radial geophone component seems to correspond with the rotational pitch components, while the roll and yaw components seem relatively weaker. The ratio between mm/s of particle velocity and mrad/s of rotation rate in these data is approximately 10:1. This is similar to what we see for offsets of a few hundred meters in previous rotational studies done with active land seismic surveys. Additionally, assuming a nominal total magnetic field of 50,000 nT, we expect rotation rates to be on the order of 1 mrad/s for magnetic deviations on the order of 1 nT, as occurs here. However, we do not know if the ground rotation is the only source of the changes in the magnetic field projections in the seismic frequency band. Other effects such as the seismoelectric effect may be occurring in conjunction with the ground rotation and contributing to the recorded changes.

DISCUSSION

In current seismic acquisition, geophones record only the displacements but not rotations. However, with the advent of a new generation of seismic sensors, these physical values will be measurable at each receiver position, providing us with 7-component seismic data: 1 pressure, 3 displacements and 3 rotations. Rotation sensors exist and have seen some very limited use in seismic test surveys on land. On the ocean-bottom though, there are as of yet no industry-grade rotation sensors. We therefore derived the rotational components by alternate means for the datasets mentioned in this paper.

We used the fact that the receivers were deployed with small spacings in the Moere Vest survey to difference their data and estimate the rotational motion that would have been recorded had we instruments that were able to measure these physical variables directly on the ocean-bottom, thereby generating seven-component data. We also derive six-component ocean-bottom earthquake data from a three-component seismometer and a two-component horizontal magnetometer deployed for the SERPENT CSEM survey.

Both methods we show have possible sources of error in the derived rotation data, and without proper experimentation there is, in principle, no way to validate them. We intend to conduct a land experiment to confirm the derivation of rotations from magnetometer recordings. The experiment will take place in the Mojave desert in California, and will include 3-component rotation sensors, 3-component magnetometers and vertical geophones. We will compare the rotation data as derived from the magnetometers to the rotation data recorded by the rotation sensors. If successful, we may tentatively envision future 7-component seismic acquisition comprising hydrophones, geophones and magnetometers.

ACKNOWLEDGEMENTS

We thank the sponsors of the Stanford Exploration Project, the Seafloor Electromagnetic Methods Consortium, and Seabed Geosolutions for releasing the Moere Vest data. We also thank Robert Brune for many fruitful discussions. The SERPENT project was funded by National Science Foundation grant OCE-0841114.

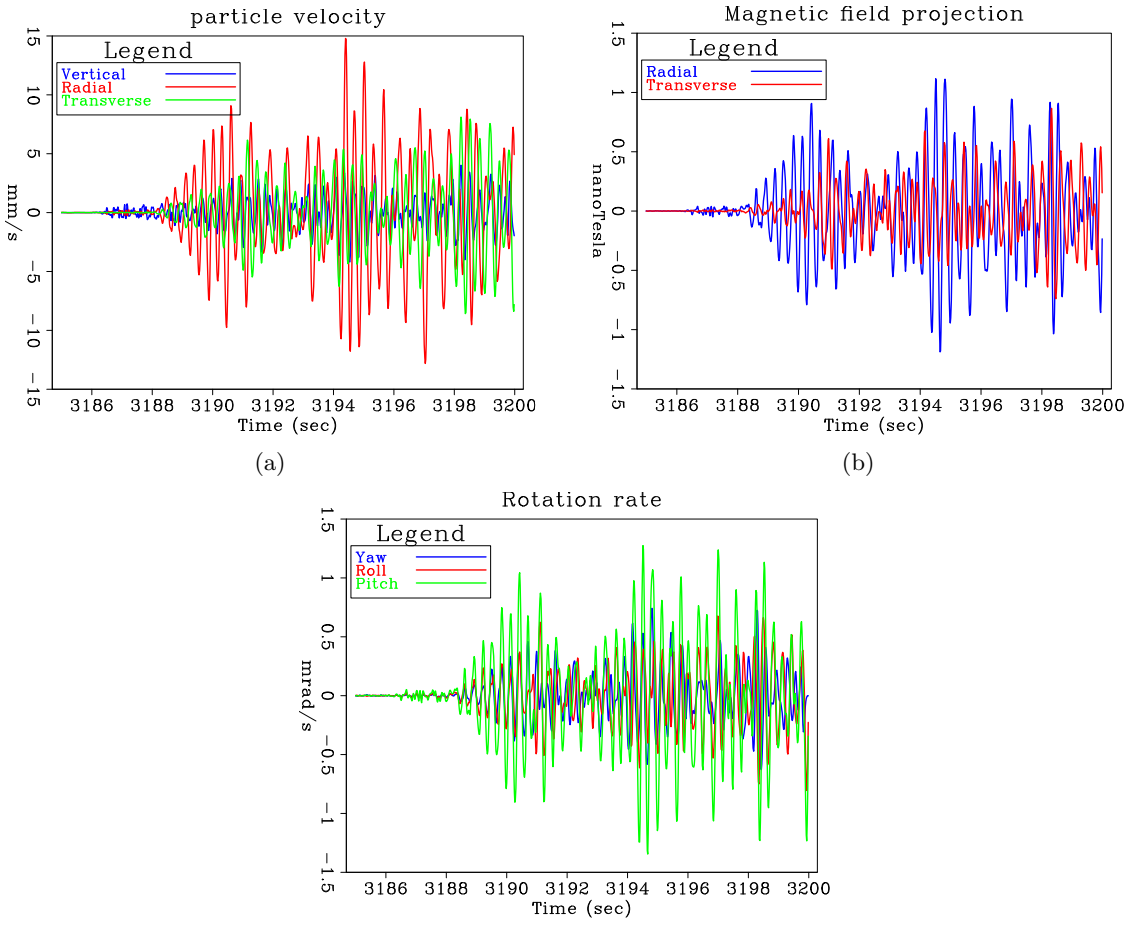


Figure 5: Data of the first 14 second of the earthquake that occurred during the SERPENT survey. Horizontal component data are rotated to minimize the energy on the geophone transverse component. (a) Particle velocity recorded by 3C geophone. (b) Deviations in projection of magnetic field on 2C horizontal magnetometer components. (c) 3 components of rotation rate derived from the 2C magnetometer data. Note the strong amplitudes of the radial geophone component and how there is some correspondence between it and the pitch rotational component. [**ER**] ohad1/. geo-win2,mag-win2,rot-win2

REFERENCES

Barak, O., S. de Ridder, J. Giles, P. Jaiswal, R. Brune, and S. Ronen, 2014a, Six-component seismic land data acquired with geophones and rotation sensors: Wave-mode separation using 6C SVD: SEG Technical Program Expanded Abstracts, 1863–1867.

Barak, O., F. Herkenhoff, R. Dash, P. Jaiswal, J. Giles, S. de Ridder, R. Brune, and S. Ronen, 2014b, Six-component seismic land data acquired with geophones and rotation sensors: Wave-mode selectivity by application of multicomponent polarization filtering: The Leading Edge, **33**, 1224–1232.

Cochard, A., H. Igel, B. Schuberth, W. Suryanto, A. Velikoseltsev, U. Schreiber, J. Wassermann, F. Scherbaum, and D. Vollmer, 2006, Rotational motions in seismology: Theory, observations, simulation: Earthquake Source Asymmetry, Structural Media and Rotation Effects, 391–411.

Edme, P., E. Muyzert, and E. Kragh, 2014, Efficient land seismic acquisition sampling using rotational data: 76th EAGE Conference and Exhibition, Seismic Noise Attenuation Session, **ELI1 08**.

Kappler, K., N. Cuevas, and J. W. Rector, 2006, Response of induction coil magnetometers to perturbations in orientation: SEG Technical Program Expanded Abstracts, 899–903.

Key, K., S. Constable, T. Matsuno, R. L. Evans, and D. Myer, 2012, Electromagnetic detection of plate hydration due to bending faults at the middle america trench: Earth and Planetary Science Letters, **351**, 45–53.

Muyzert, E., A. Kashubin, E. Kragh, and P. Edme, 2012, Land seismic data acquisition using rotation sensors: 74th Conference and Exhibition, EAGE, Extended Abstracts.

Naif, S., K. Key, S. Constable, and R. L. Evans, 2013, Melt-rich channel observed at the lithosphere-asthenosphere boundary: Nature, **495**, 356–359.

Pillet, R., A. Deschamps, D. Legrand, J. Virieux, N. Bethoux, and B. Yates, 2009, Interpretation of broadband ocean-bottom seismometer horizontal data seismic background noise: Bulletin of the Seismological Society of America, **99**, 1333–1342.

Vassallo, M., K. Eggenberger, D. J. van Manen, K. Ozdemir, J. Robertsson, and A. Ozbek, 2012, Contributions of the horizontal and vertical components of particle velocity in 3D pressure wavefield reconstruction on dense receiver grids using generalized matching pursuit: SEG Technical Program Expanded Abstracts, 1–5.

Toward PZ summation without Z

Ettore Biondi and Stewart A. Levin

ABSTRACT

We examine the possibility of separating up-going and down-going wavefields of ocean-bottom data using only one component. This possibility relies on differential static shifts of the up-going events caused by near sea-bottom inhomogeneities. We downward continue a survey to the sea bottom to recover these shifts while leaving non-primary arrivals smoothly curved. We then explore whether or not these static shifts are detectable in the curvelet domain. In addition, we show how suppressing fine-scale curvelet coefficients affects an event distributed along a disrupted hyperbolic curve. From our synthetic examples, we demonstrate that the curvelet domain has the potential to separate up-going from down-going events.

INTRODUCTION

The separation of up-going and down-going waves of a multicomponent dataset is one of the fundamental preprocessing steps in ocean-bottom acquisition (OBN) (Grion, 2010). This wavefield decomposition is usually performed by summing the pressure (P) and vertical velocity (V_z) components (Schalkwijk et al., 1999), a process commonly known as 'PZ summation'. Before summing the two components, the difference in instrument coupling response between the geophone and hyrophone must be properly accounted for by applying a calibration filter to the recorded vertical velocity (Melbø et al., 2002; Biondi and Levin, 2014).

Many applications have already demonstrated the advantages and the importance of such wavefield separation. For example, water reverberations can be attenuated by using the extracted up-going events (Rosales and Guitton, 2004). Furthermore, a source wavelet can be easily estimated from the down-going separated direct arrival (Wong and Ronen, 2009). In addition, the combined information of the two wavefields can be used in the context of improving the images obtained by least-squares reverse-time migration (LSRTM)(Wong et al., 2010).

Current up-down wavefield separation techniques rely on the assumption that the data contain noise-free up- and down-going events. However, it is well known that the vertical velocity component can be contaminated by shear wave energy (Paffenholz et al., 2006), a phenomenon commonly called Vz noise. The leakage of shear energy into the vertical velocity, which is not recorded by hydrophones, can degrade PZ summation results and subsequent imaging (Campman et al., 2005). Different techniques have been proposed to suppress or dampen the effect of Vz noise present in the vertical component (Shatilo et al., 2004). We discuss how the up-down wavefield separation may be performed using only the pressure field of the recorded data. We start by explaining the main assumption, seafloor statics, on which this single component separation relies. We show that by downward

continuing a survey line to the sea bottom we can retrieve the static shifts of the up-going events caused by the near sea-bottom inhomogeneities. We then briefly review the curvelet transform and apply this transformation to two hyperbolic events, one of which is affected by static shifts. We show the differences between these events in the curvelet domain. The possibility of separating up- and down-going energy without the combination of the pressure and vertical components can enable application of the wavefield separation to the horizontal components as well.

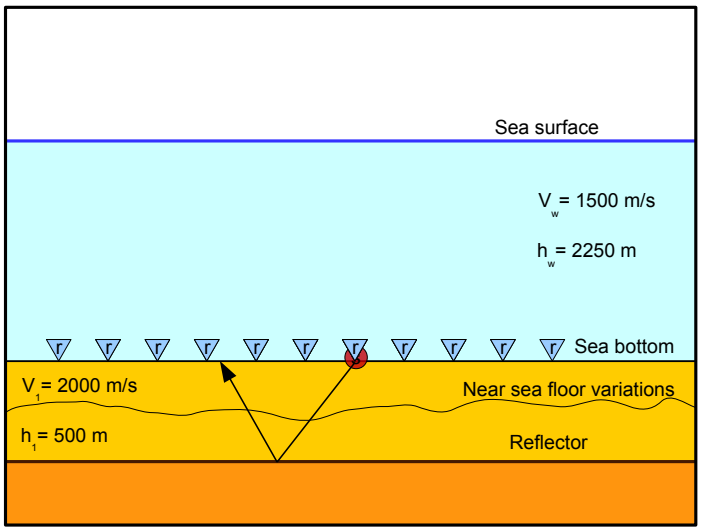
UP-GOING AND DOWN-GOING STATIC DIFFERENCES

We recall the concept of wavefield focusing shown by Claerbout (1976) to exlain why up-going events should exhibit static shifts not present on the down-going arrivals after downward continuation of the survey to the sea bottom. Consider a wavefront that has been disrupted by small-scale inhomogeneities present in the subsurface. When this wavefront propagates through a homogeneous medium (e.g., a water layer), the wavefront heals and the energy spreads out smoothly. The longer the wave propagates through the homogeneous medium, the more it will heal.

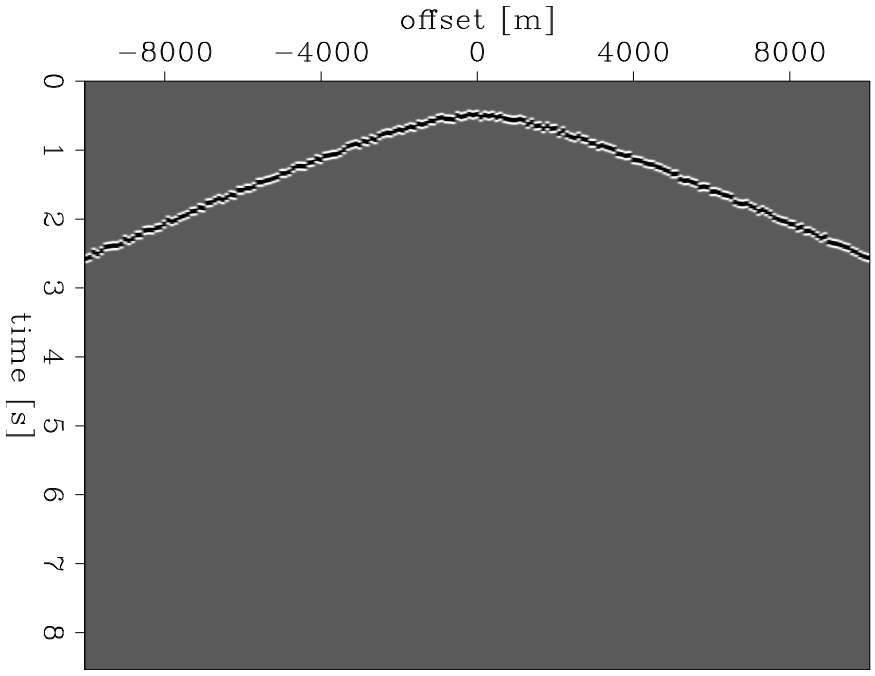
We start with the layered earth model shown in Figure 1a. This figure displays a reciprocal ocean-bottom acquisition scenario where the shots and receiver are on the sea bottom. The black arrow indicates a primary event that we would record with this survey. The recorded reflection is disrupted by the near sea-floor variations as we see in the common-shot gather of Figure 1b. We then upward continue the receivers to the sea surface (Figure 2a) to simulate the physical reciprocal experiment of an OBN acquisition. The red and green arrows in this figure indicate the direct arrival and its first-order surface related multiple respectively, which are not affected by the near sea-bottom inhomogeneities. Figure 2b displays the common-shot gather of these events and the upward continued primary reflection. Ignoring the minor artifact introduced by the upward continuation, we note that the static shifts present on the primary event are eliminated, and the energy is distributed along a smooth hyperbolic curve. This effect comes as a result of wavefront healing as the wave propagates through the homogeneous water layer.

Now, if we downward continue the gather shown in Figure 2b, we expect to focus the energy of the direct event at the source location and at zero time, and also start collapsing the energy of the other events (Figure 3a). This is indeed what we see in the downward continued gather of Figure 3b. In fact, the energy of the direct event is tightly focused, and the two other events have started focusing. The main point to note is that the static shifts, initially present in the primary reflection, have been retrieved by sinking the receivers to the sea floor. In contrast, the energy of the other down-going event is still distributed along smooth curves or lines.

This simple illustrative example shows that the difference in static shifts between the up-going and down-going events can be recovered by downward continuing the survey line to the sea floor. This difference is the key element on which our separation criterion relies.

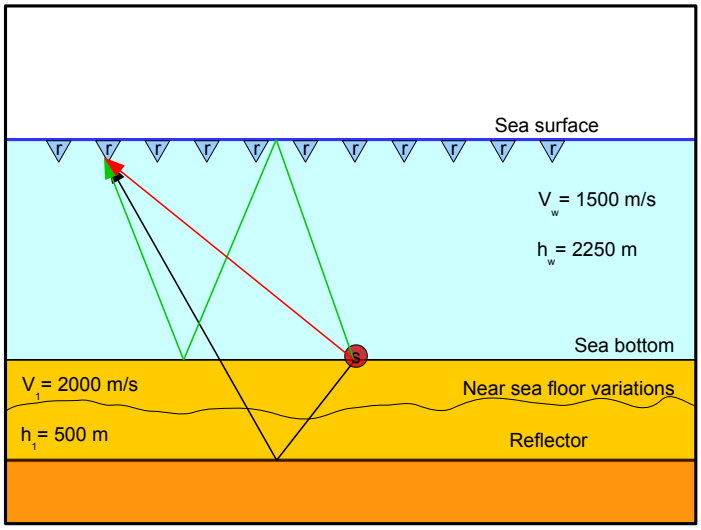


(a)

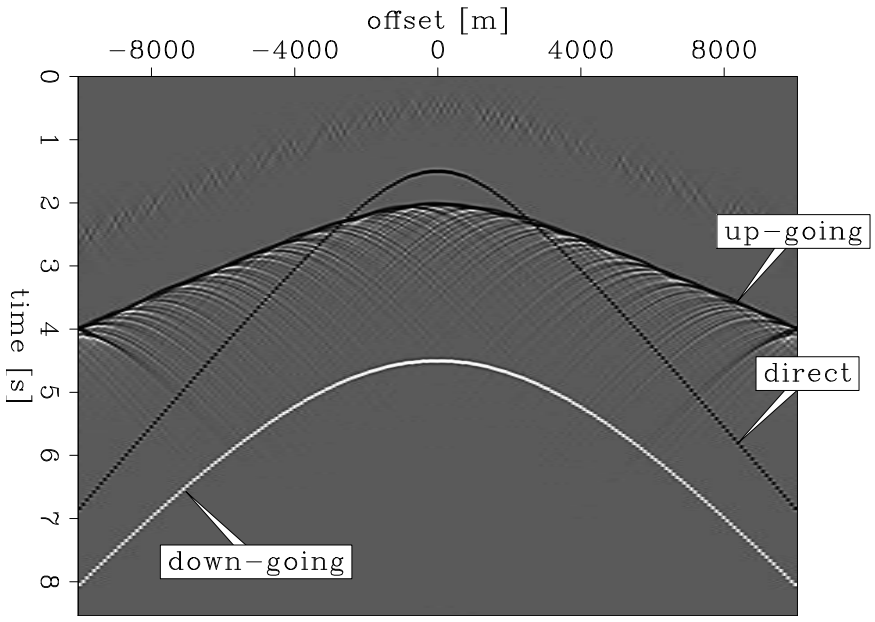


(b)

Figure 1: Initial acquisition setting with source and receivers at the sea bottom. (a) Earth model used for generating one up-going event displayed by the black arrow. [**NR**](b) Single up-going event recorded at the sea floor. The effect of the near sea-floor variations are visible in the statics present in the hyperbolic event. [**ER**] ettore1/. startNR,data

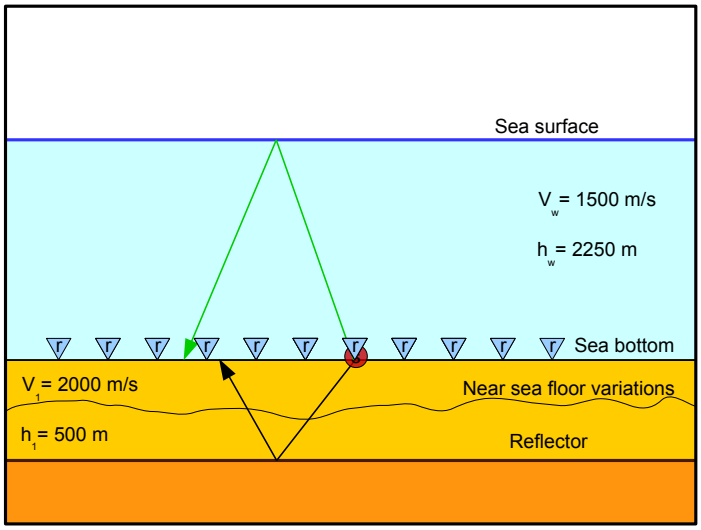


(a)

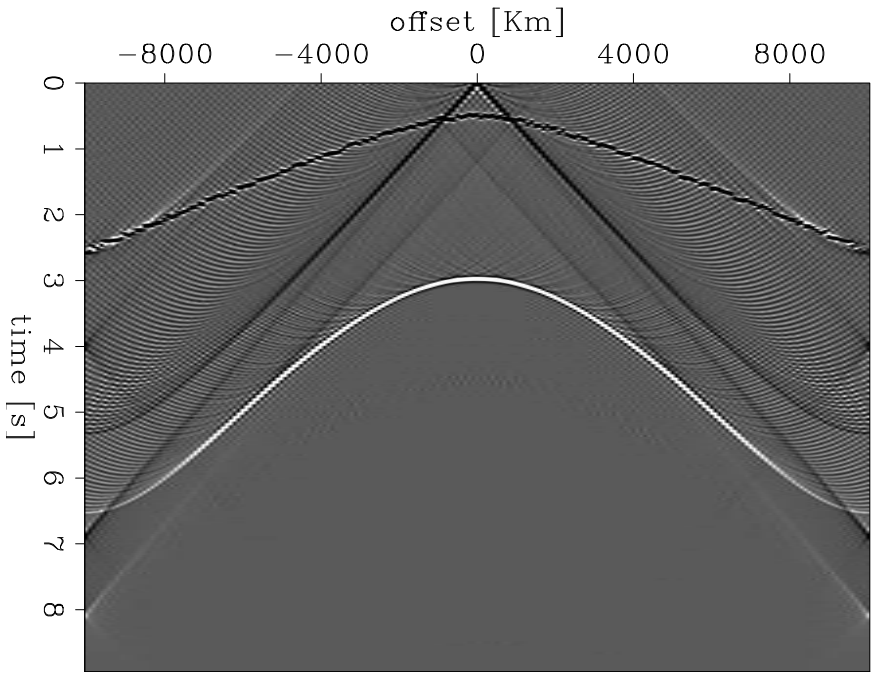


(b)

Figure 2: Events recorded when the receivers are at the sea surface. (a) Illustration of the recorded events. The red, green, and black arrows represent the direct arrival, its multiple, and the upward continue primary event, respectively. [**NR**](b) Common-shot gather showing the upward continue primary event, the direct arrival, and its first-order multiple. The statics present on the primary reflection of Figure 1b are eliminated by the propagation of the event toward the sea surface. Some artifacts of the upward continuation are also present. [**ER**] ettore1/. upNR,total-ann



(a)



(b)

Figure 3: Downward continuation of the events shown in Figure 2. (a) Illustration of how the recorded events change when the receivers are downward continued to the sea-floor surface. [**NR**](b) Downward continued common-shot gather of Figure 2b. The direct event tends to focus at zero time; while, its multiple starts collapsing. The statics of up-going primary event are recovered when the receivers are downward continued to the same depth of the source. [**ER**] ettore1/. downNR,downtotal

THE CURVELET TRANSFORM

Before explaining how the curvelet transform can capture the differences between events disrupted by static shifts and arrivals distributed along smooth curves, we briefly show the basic concept behind the curvelet transform. The curvelet domain was originally developed in the context of image processing, and it has been widely used for image denoising (Starck et al., 2002; Candés et al., 2006). This transform has also been applied to seismic imaging to attenuate the effect of Vz noise on the PZ summation results (Peng et al., 2013).

The main idea of the curvelet transform is to take an original signal $f(x)$, usually an image, and decompose into it a sum of wavelets $\varphi_{j,l,k}(x)$ with a different j scale, l orientation, and k position,

$$f(x) = \sum_{j,l,k} c(j,l,k)\varphi_{j,l,k}(x), \tag{1}$$

where $c(j,l,k)$ coefficients are given by the scalar product of the original signal with the basis functions

$$c(j,l,k) = \langle f, \varphi_{j,l,k} \rangle = \int f(x)\varphi_{j,l,k}(x)dx. \tag{2}$$

Figure 6 shows the curvelet transform flowchart by Starck et al. (2002). The image is divided into a different number of blocks for each scale. As the level of the detail that we want to represent increases, the number of blocks increases as well. At first, we take the whole image and apply a wavelet transform. Then, we divide the image into a fixed number of blocks, and apply the wavelet transform on each section. This process continues until we reach the desired number of scales. The wavelet transform is performed as a combination of Fourier and Radon transforms followed by a ridgelet transform along lines in the Radon domain. All the mathematical details can be found in Starck et al. (2002). The ability of the curvelet transform to analyze different level of details on an image, thanks to the inclusion of the j scale factor in the wavelet expansion, enables us to capture the difference between events disrupted by static shifts and events unaffected by them as we see in the next section.

THE EFFECT OF STATICS IN THE CURVELET DOMAIN

In this section, we show the differences of two hyperbolic events in the curvelet domain, one disrupted by static shifts and one unaffected by static shifts. Figure 5 displays these two hyperbolic events. The left panel shows the down-going sea-bottom multiple event, which is unaffected by statics. The right panel depicts the up-going reflection event with static shifts resulting from the near sea-floor variations.

By applying the curvelet transform described in the previous section, we obtain the panels of Figure 6. These images display the coefficients of the curvelets necessary to represent the original functions. The central square panel represents the coefficients of the

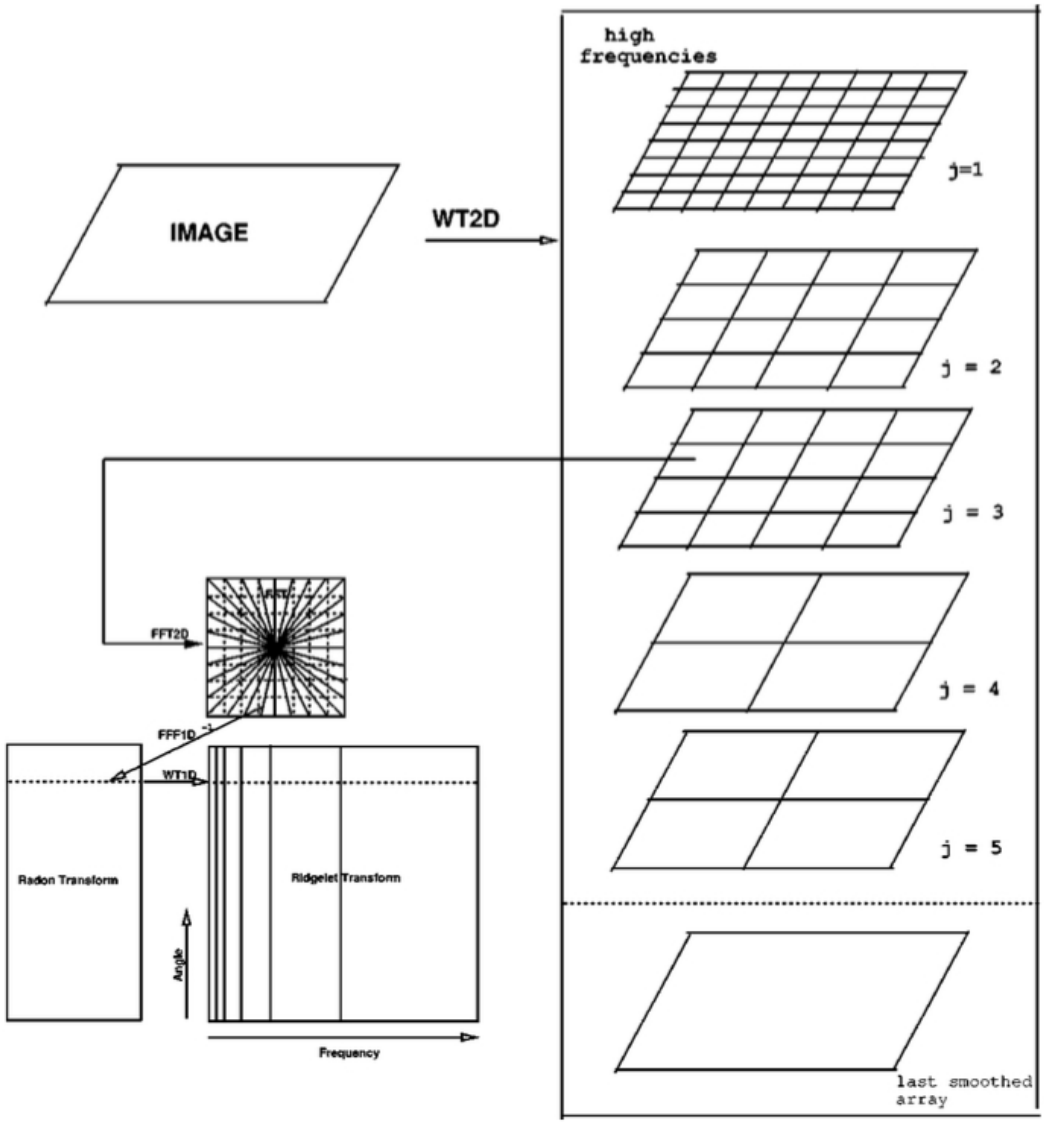


Figure 4: Figure from Starck et al. (2002). Graphical explanation of the curvelet transform. An input image is decomposed into a different number of blocks as a function of the scale to be analyzed. The wavelet/ridgelet transform is then applied to each block. [**NR**] ettore1/. curvelet

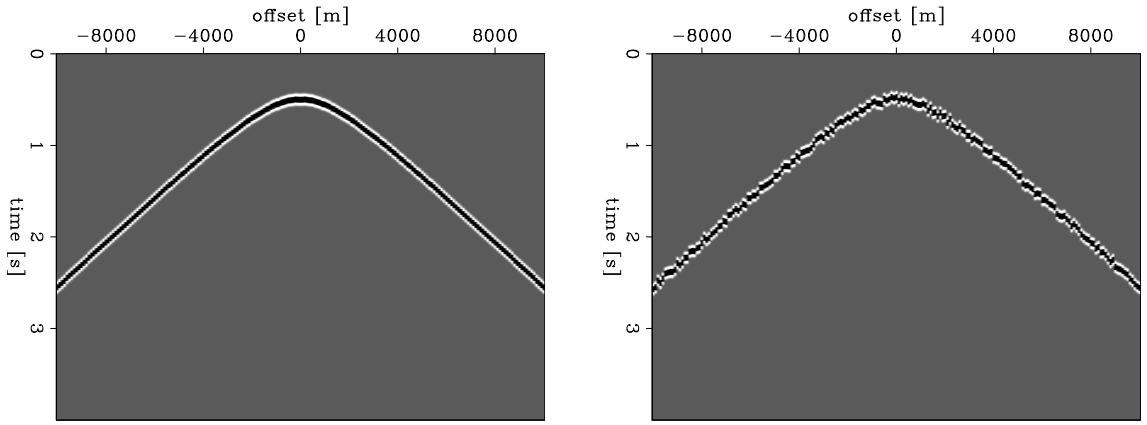


Figure 5: Hyperbolic events used for analyzing apparent differences in the curvelet domain. (a) Event with no static shifts. (b) Event affected by statics. [**ER**] ettore1/. clean,statics

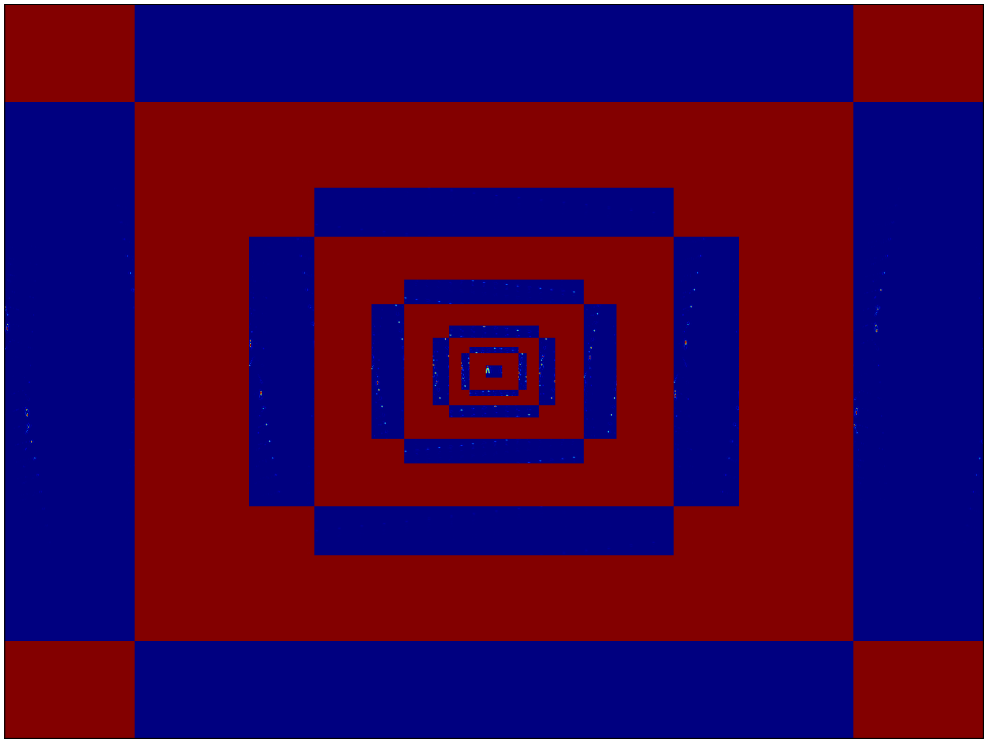
curvelets of the whole image scale. The surrounding red zone is present just to separate the coefficients from one scale to another. The next four surrounding blue rectangles are the coefficients of the curvelets of a finer scale. As we move to the external rectangles, we find the multiplicative factors of finer scale curvelets. Comparing the curvelet transform factors of the two events, we observe that the finer scale coefficients are more scattered for the hyperbola with static shifts (Figure 6b) than for the hyperbola with no static shifts (Figure 6a).

The same conclusion can be drawn from the closeups of the central curvelet domain coefficients, shown in Figure 7. However, from these closeups it is apparent that at a coarse scale, the two events are similar. The scattering behavior of the fine scale coefficients of the disrupted hyperbolic event can be a distinctive factor that enables us to separate up-going energy from down-going arrivals, after the downward continuation of the survey line.

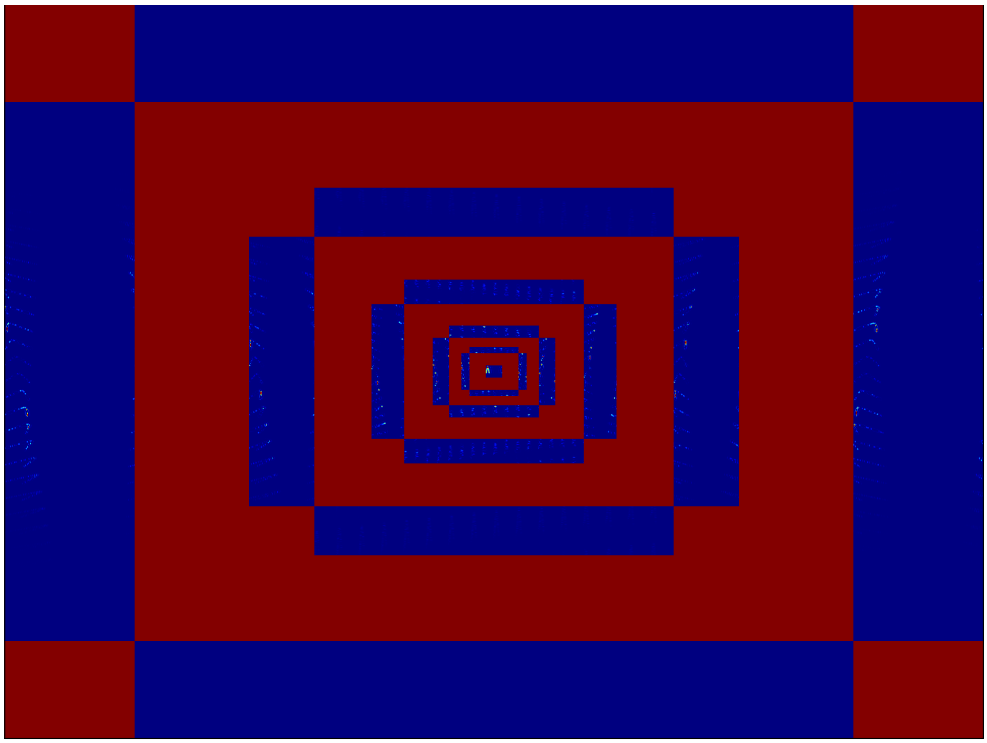
In the following synthetic example, we show how to employ the curvelet domain to separate the two hyperbolic events, one with statics and one without them. Figure 8(a) shows the gather containing these two events, and Figure 8(b) is the curvelet domain representation of the gather. From the previous observations, we can say that the curvelet domain coefficients of the event with statics are more dispersed than the ones for the smooth hyperbolic event (Figure 6b). Figure 9 is the result of filtering the scattered coefficients of the last four fine curvelet scales and applying the inverse curvelet transform. Observe that we are able to correctly reconstruct the static-free event, and make the disrupted event more similar to random noise distributed along a hyperbolic curve. This random energy can be then attenuated by applying a denoising technique (Kahoo and Siahkoohi, 2009; Han et al., 2014), and thus separating the two events.

FUTURE WORKS AND CONCLUSIONS

Performing a wavefield up-down separation by using just one component of a multicomponent ocean bottom dataset is an attractive goal, especially when PZ summation algorithms

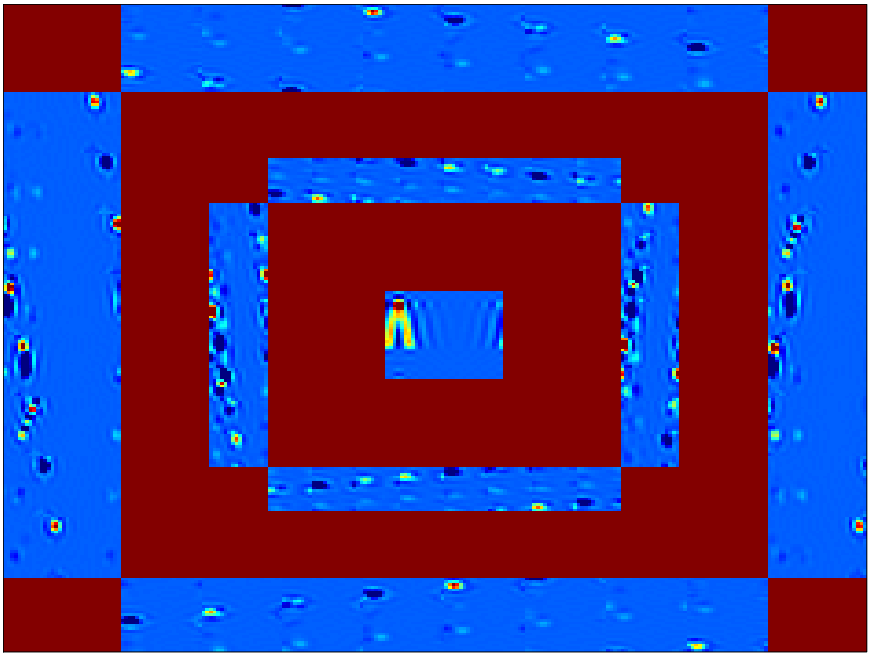


(a)

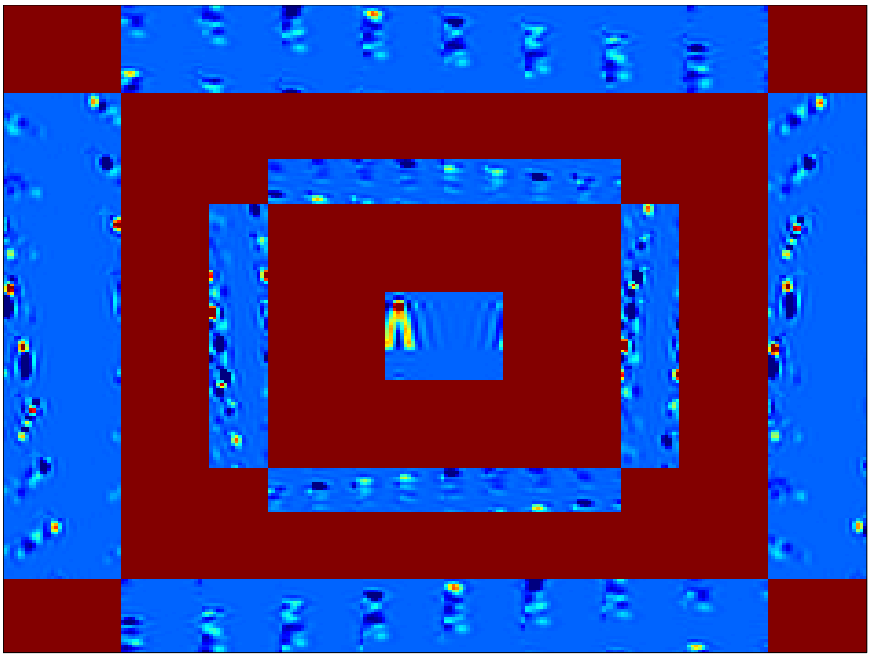


(b)

Figure 6: Curvelet domain coefficients of the hyperbolic events of Figure 5. (a) Coefficients of static-free event. (b) Coefficients of event affected by statics. As we look toward finer scale factors (external blue rectangles), the coefficients of the disrupted event become more scattered. [**ER**] ettore1/. cur-clean,cur-static

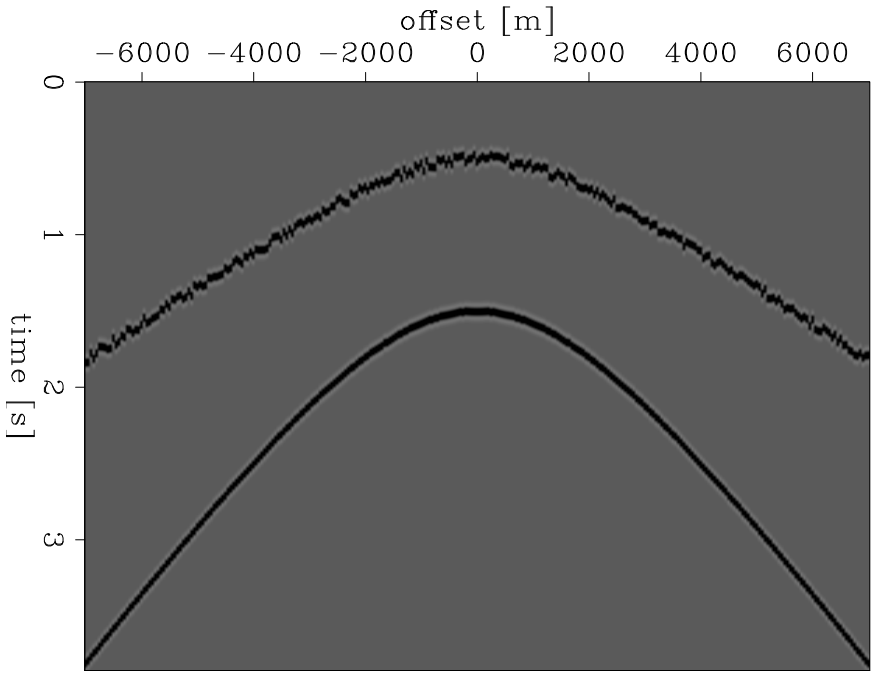


(a)

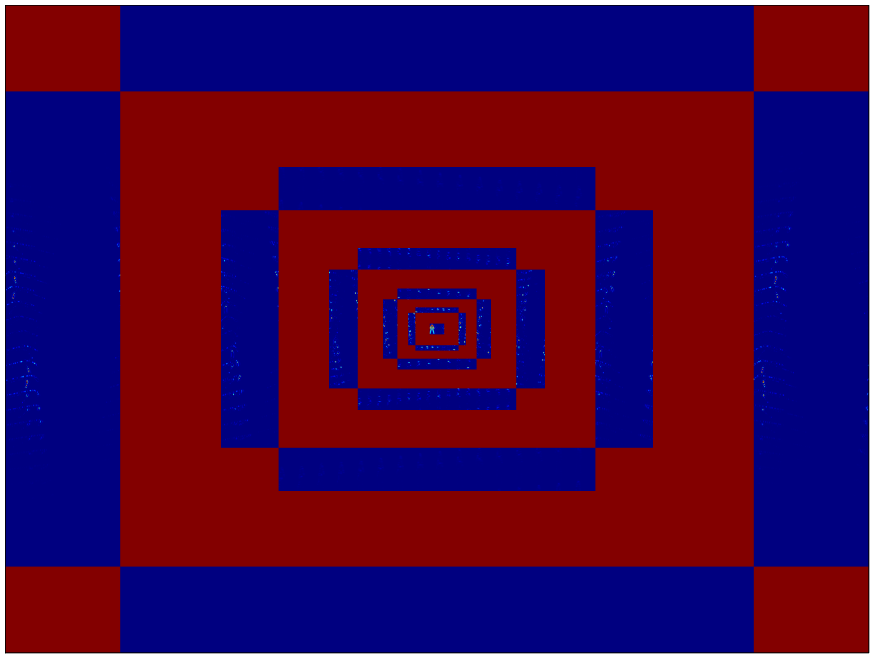


(b)

Figure 7: Closeups of the first three scales of the curvelet transform factors of Figure 6. (a) Closeup of curvelet factors of static-free event. (b) Closeup of curvelet factors of event affected by statics. Comparing the whole image scale curvelet factors, we note that the two image are similar. However, as we move to finer scale factors, the coefficients for the disrupted event tend to become less concentrated on a single area. [**ER**] ettore1/. clean-zoom,static-zoom



(a)



(b)

Figure 8: Synthetic example for explaining how to separate an event affected by statics from an event unaffected by them. (a) Gather with the two hyperbolic events. (b) Curvelet domain representation of the two events. The scattered coefficients at fine scales represent the amplitudes of the curvelets of the event disrupted by the static shifts. [**ER**] ettore1/. events,cur-events

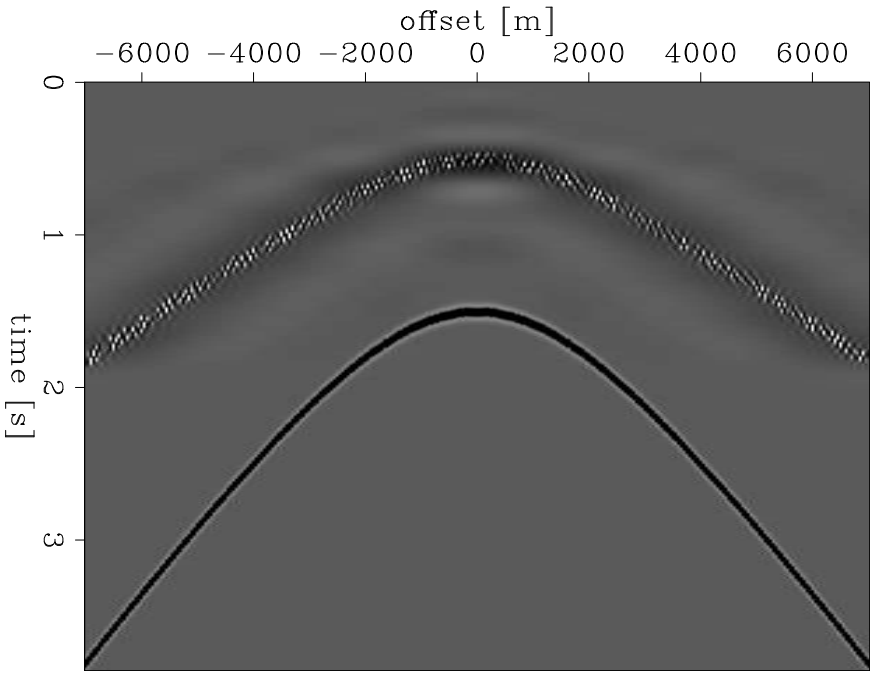


Figure 9: Resulting gather after filtering out the fine-scale curvelet coefficients of the event affected by statics. The event distributed along the smooth hyperbola is completely reconstructed. The hyperbolic trend of the event with statics can still be seen, but its energy now appears as random noise. [**ER**] ettore1/. events-clean

fail because of the presence of Vz noise in the recorded vertical velocity. We discussed the possibility of doing up-down wavefield separation by downward continuing the survey line to the ocean bottom, and then separating up- and down-going energy in the curvelet domain. We explained how to retrieve the static shifts caused by near sea-floor inhomogeneities that affect up-going events by downward continuation. We observe that the curvelet domain is able to capture the differences between events affected by statics and events without them. We have also shown that by suppressing the fine scale curvelets coefficients of the disrupted event, we were able to make this event appear as a random noise feature, making it easier to filter it out with denoising techniques. This example demonstrates that the curvelet domain has the potential to separate up-going events from down-going ones. In the future, we will use these observations to separate up-going energy from down-going events, both on a more complex synthetic gather and a real common-receiver gather of an OBN dataset. Furthermore, because this wavefield decomposition is based on a single component, we will be able to apply this curvelet domain separation on the pressure and vertical components geophone components separately, and possibly also on the horizontal geophone components.

REFERENCES

Biondi, E. and S. A. Levin, 2014, Application of the up-down separation using PZ calibration filter based on critically refracted waves.

Campman, X. H., K. van Wijk, J. A. Scales, , and G. C. Herman, 2005, Imaging and suppressing near-receiver scattered surface waves: Geophysics, **70**, no. 2, V21–V29.

Candés, E., L. Demanet, D. Donoho, and L. Ying, 2006, Fast discrete curvelet transforms: Multiscale Modeling & Simulation, **5**, 861–899.

Claerbout, J. F., 1976, Fundamentals of Geophysical Data Processing: Blackwell Scientific Publications.

Grion, S., 2010, From PZ Sum to wavefield separation, mirror imaging and up down deconvolution: The Evolution of OBS Data Processing: Annual International Meeting, KazGeo, 83–87.

Han, L., M. D. Sacchi, and L. Han, 2014, Spectral decomposition and de-noising via time-frequency and space-wavenumber reassignment: Geophysical Prospecting, **62**, 244–257.

Kahoo, A. R. and H. R. Siahkoohi, 2009, Random noise suppression from seismic data using time-frequency peak filtering: Annual International Meeting, 71st EAGE Conference and Exhibition, P317.

Melbø A. H. S., J. O. A. Robertsson, and D. Manen, 2002, PZ calibration by applying the equation of motion to critically refracted waves, 2002 SEG Annual Meeting, 1030–1033. Society of Exploration Geophysicists.

Paffenholz, J., R. Shurtleff, D. Hays, and P. Docherty, 2006, Shear wave noise on OBS Vz Data - Part I Evidence from field data: Annual International Meeting, 68th EAGE Conference and Exhibition, B046.

Peng, C., R. Huang, and B. Asmerom, 2013, Shear noise attenuation and PZ matching for OBN data with a new scheme of complex wavelet transform: Annual International Meeting, 75th EAGE Conference and Exhibition, Tu 02 07, 250–254.

Rosales, D. A. and A. Guitton, 2004, Ocean-bottom hydrophone and geophone coupling.

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.

Shatilo, A., R. Duren, and T. Rape, 2004, Effect of noise suppression on quality of 2C OBC Image: Annual International Meeting, 2004 SEG Annual Meeting,, 917–920.

Starck, J.-L., E. J. Candés, and D. L. Donoho, 2002, The curvelet transform for image denoising: IEEE Transactions on Image Processing, **11**, 670–684.

Wong, M., B. Biondi, and S. Ronen, 2010, Joint least-squares inversion of up- and down-going signal for ocean bottom data sets.

Wong, M. and S. Ronen, 2009, Source signature and static shifts estimations for multi-component ocean bottom data.

SEP PHONE DIRECTORY

Name	Phone	Login Name
Almomin, Ali	723-0463	ali
Alves, Gustavo	723-9282	gcalves
Barak, Ohad	723-9282	ohad
Barnier, Guillaume	723-6006	gbarnier
Biondi, Biondo	723-1319	biondo
Biondi, Ettore	725-1625	ettore
Cabrales, Alejandro	723-3187	cabrales
Chang, Jason	724-2974	jason
Claerbout, Jon	723-3717	jon
Clapp, Bob	725-1334	bob
Dahlke, Taylor	723-3187	taylor
Lau, Diane	723-1703	diane
Le, Huy	723-3187	huyle
Leader, Chris	723-0463	chrisl
Levin, Stewart	726-1959	stew
Ma, Yinbin	723-3187	yinbin
Maharramov, Musa	723-1250	musa
Martin, Eileen	723-0463	ermartin
Ruan, Kaixi	723-3187	kaixi
Shen, Yi	723-6006	yishen
Zhang, Yang	723-6006	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, 2014-2015

Name	**Company**	**Tel #**	**E-Mail**
Raymond Abma	BP	(281)366-4604	abmar1@bp.com
Francois Audebert (Co-chair, 2nd year)	TOTAL	(33)-6-1278-4195	francois.audebert@total.com
Biondo Biondi	SEP	(650)723-1319	biondo@sep.stanford.edu
Jon Claerbout	SEP	(650)723-3717	jon@sep.stanford.edu
Thomas Dickens	ExxonMobil	(713)431-6011	tom.a.dickens@exxonmobil.com
Faqi Liu (Co-chair, 1st year)	Hess	(823)231-9218	fliu@hess.com
Yi Luo	Saudi Aramco	–	yi.luo@aramco.com
Alejandro Valenciano	PGS	–	alejandro.valenciano@pgs.com

Research Personnel

Jonathan Ajo-Franklin is a research scientist in the geophysics department of Lawrence Berkeley National Laboratory.

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.

Gustavo Catao Alves graduated from the Federal University of Minas Gerais, Brazil, in 2004 with a BS in Physics. In 2006, he got a MS in Solid State Physics from the same Institution and was awarded Best Student Presenter on the 3rd Conference of the Brazilian Carbon Association. That same year he joined Petrobras, where he worked at the Company's research center for the next 7 years, on projects focused on seismic illumination studies, Reverse Time Migration and Full Waveform Inversion. In 2008, he received an Honorable Mention for Best Poster Paper at SEG's Annual Meeting, with the work entitled "Target oriented illumination analysis using wave equation". He joined Stanford Exploration Project in 2013 to pursue a PhD in Geophysics and is currently working on amplitude studies in long offset data. He is a member of SBGf, SEG, EAGE and AAPG.

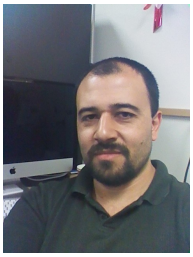
Ohad Barak received a B.Sc. (2006) and an M.Sc. (2009) in Geophysics from Tel-Aviv University. In 2008 he joined the R&D team of Paradigm Geophysical, maintaining and developing the production codes. He joined SEP in 2009 and is currently pursuing a Ph.D. in geophysics at Stanford University, and a longer biography. His current research involves seismic rotation data and the various methods of acquiring them, on land and on the ocean bottom. He is working on showing how rotation data can be utilized in combination with standard geophone and hydrophone data for applications such as wave-mode separation and data interpolation.

Guillaume Barnier is from Nice, France. He graduated from Telecom Paristech in 2007 with a MSc in telecommunications and signal processing. After working as a fixed income trader for JP Morgan (London) from 2007 to 2010, he decided to change career path by joining the geophysics MSc program in Colorado School of Mines (2011), where he focused his research on seismoelectric coupling in poro elastic media. In September 2013, he joined the Stanford Exploration Project to pursue his Ph.D. Guillaume focuses on improving the image quality of subsalt layers when they have been damaged by an inaccurate top salt interface delineation.

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.

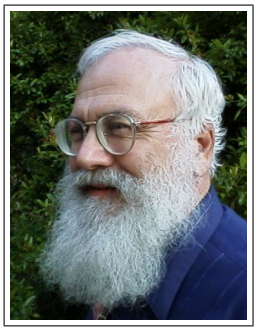
Ettore Biondi received a B.S. (2010) in geology from the University of Genoa, an M.S. (2012) in geophysics from the University of Pisa, and a diploma (2013) in computational chemistry from the Scuola Normale Superiore of Pisa. He spent almost two years as a research fellow at the University of Milan within the geophysics section of the Earth Science Department. In the summer of 2014, he joined the SEP and started his Ph.D. at the Stanford University.

Alejandro Cabrales-Vargas obtained his Bachelors degree in Geophysics in the University of Mexico in 2002. He has being working for Petroleos Mexicanos since 2002, initially in seismic interpretation for oil and gas exploration, and more recently in the supervision of depth imaging processes. He obtained his Masters Degree in Geophysics in the University of Oklahoma in 2011. He joined SEP in the fall of 2014, and is currently working towards his PhD.



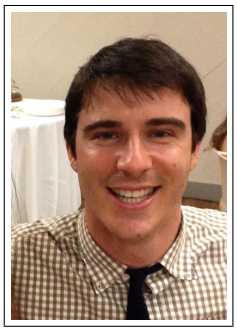
Jason Chang received his B.A. in geophysics from the University of California, Berkeley, in 2010. He joined SEP in autumn of 2011 and is currently working toward a Ph.D. in geophysics. He is a student member of SEG and AGU.

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.



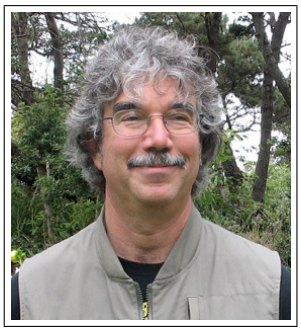
Robert Clapp received his B.Sc. (Hons.) in Geophysical Engineering from Colorado School of Mines in May 1993. He joined SEP in September 1993, received his Masters in June 1995, and his Ph.D. in December 2000. He is a member of the SEG and AGU.

Steven Constable received a Ph.D. in geophysics from ANU in 1983 and is currently a professor at Scripps Institution of Oceanography. Constable is interested in all aspects of electrical conductivity, and has made contributions to inverse theory, electrical properties of rocks, mantle conductivity, magnetic satellite induction studies, global lightning, and instrumentation. However, his main focus is marine electromagnetic methods; recent efforts have involved the commercialization of marine EM for hydrocarbon exploration, in which he played an important role in the early adoption of marine EM methods by industry.



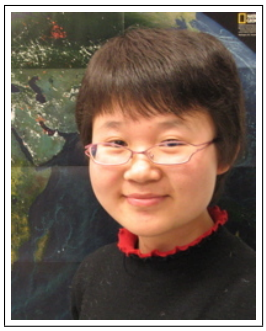
Taylor Dahlke is a third year student with SEP. He received his B.S. in civil engineering from the University of California, Berkeley in 2012, and joined SEP in July 2012. Currently, he is working towards a Ph.D. in geophysics with his research focused on applying levelset methodologies to perfom salt body image segmentation. Taylor is a student member of SPE and SEG.

Tom Daley is currently Geophysics Department Head and a staff scientist in the Earth Sciences Division of Lawrence Berkeley National Laboratory. He has been with Berkeley Lab since 1987. He received a Bachelors degree in Geophysics from the University of California, Berkeley in 1980 and a Masters degree in Engineering Geoscience from the University of California, Berkeley in 1987. He worked from 1980 to 1985 with Seismograph Service Corporation performing borehole seismic surveys and managing a district office in Ventura, CA.



Tom's research work is focused on the acquisition and analysis of borehole seismic data from field scale experiments. Current work includes testing of distributed acoustic sensing (DAS) using fiber optic cables. Problems addressed include continuous travel time monitoring to detect stress changes, monitoring geologic storage of CO2, characterization of fracture content and dominant fracture orientation in geothermal and oil fields, high resolution imaging of shallow surface materials, imaging fracture flow zones in contaminated aquifers, and geophysical characterization of volcanic tuff flows for nuclear waste isolation at Yucca Mountain.

Tom has been a member of the American Geophysical Union (AGU) since 1987, a member of the Society of Exploration Geophysicists (SEG) since 1980, the European Association of Geoscientists and Engineers (EAGE), and is currently on the SEG Council, the SEG Research Committee, and chair of the CO2 research subcommittee.



Shan Dou is a Ph.D. student in the department of earth and planetary science at University of California, Berkeley. Her main areas of research center upon mapping and characterizing permafrost with seismic methods- including field-scale surface wave inversion and lab-scale rock physics modeling.

Jack Dvorkin received his MS (1974) and Ph.D. (1980), both in continuum mechanics, from Moscow University, Department of Mechanics and Mathematics. His disseration topic was "Impact of Elastic Shells upon the Surface of Compressible Fluid." From 1974 to 1989 Jack worked at the Moscow Research Oil and Gas Institute, and the Moscow Offshore Oil Institute. Jack joined the Stanford Rock Physics Laboratory in July 1989. He is an AGU, SEG, SPE, and AAPG member.

Barry Freifeld is a hydrogeologist in the Earth Sciences Division of Lawrence Berkeley National Laboratory. He is a specialist in subsurface monitoring and fiber-optic technology.

Antoine Guitton received a M.Sc. in geophysics from Université de Strasbourg in 1996. He also received a M.Sc. and Ph.D in geophysics from Stanford University in 1998 and 2005. He is a recipient of the EAGE Arie van Weelden Award in 2004 and SEG J. Clarence Karcher Award in 2007. He received the Best Paper Presented by a Student at the Annual Meeting Award in 1999 for his work on robust norms. Antoine is currently working for Geoimaging Solutions Inc. as a senior research geophysicist working on imaging and velocity estimation problems. He was a consulting professor of geophysics at Stanford University from August 2005 to August 2014. He is a member of the EAGE, SEG and AGU.

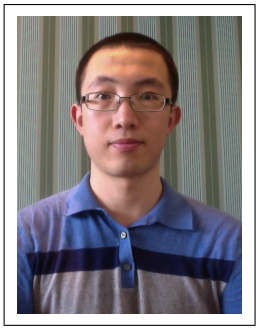
Kerry Key received a B.S. in Earth science from the University of California, San Diego and a Ph.D. in geophysics from Scripps Institution of Oceanography at UC San Diego. He is an associate professor in the Institute of Geophysics and Planetary Physics (IGPP) at Scripps Institution of Oceanography. Keys research focuses on using electromagnetic geophysical methods for the remote detection of geologic fluids contained in the crust and mantle underneath the oceans at mid-ocean ridges, subduction zones and on the continental shelves. Key pioneers new electromagnetic technologies for oil and gas exploration such as the marine controlled-source electromagnetic and the marine magnetotelluric methods. He has published numerous papers in the field of geophysics and has participated in more than 30 seagoing scientific expeditions.

Huy Le earned his B.S. degree in Geophysics from the University of Oklahoma in 2012 and his M.S. degree in Computational Geoscience from Stanford University in 2014. He is now a PhD candidate in Geophysics. His research interests include anisotropy, full waveform inversion, finite difference method, and high-performance computing.

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.

Stewart A. Levin was acting director of the Stanford Exploration Project during Jon Claerbout's 1993-4 sabbatical year. After a distinguished career in industry at Mobil and Halliburton, he has returned to Stanford as a consulting professor in the Department of Geophysics.

Nate Lindsey is a senior research associate in the earth sciences division of Lawrence Berkeley National Laboratory.



Yinbin Ma graduated from University of Science and Technology of China in 2009 with a B.S. in physics. In 2013, he obtained a M.S. in Condensed Matter Theory at Rice University. Yinbin moved to Stanford in 2013 and is currently working towards a Ph.D. in the Institute for Computational and Mathematical Engineering. He started doing research at SEP from June 2014. He is a student member of the SEG.

Musa Maharramov received his diploma c.l. in Mathematics from Baku State University in 1993. In 1995, he received a Ph.D. in Mathematics (Differential Equations) from the Azerbaijan Institute of Mathematics for his thesis "Asymptotic Solutions of Quasilinear Parabolic Equations". In 1995-2001, he worked as an IT& Comms consultant, later IT manager, at Fluor Corporation in Baku. In 2001, he joined BP Azerbaijan, and in 2006 he started his work at BP High Performance Computing as a Computational Scientist. In that role he worked with the BP Imaging Team on mathematical, algorithmic and computational aspects of seismic migration and inversion. Musa joined SEP in the fall of 2011 and is pursuing his second Ph.D., in Geophysics. He is currently working on regularization of seismic inversion problems through the application of Geological/Geomechanical constraints and advanced optimisation techniques. Musa is a member of the SEG, EAGE and SIAM.

Eileen Martin graduated from the University of Texas at Austin in 2012 with a B.S. in Mathematics and Computational Physics. She is a Department of Energy Computational Science Graduate Fellow, and is working towards her Ph.D. in the Institute for Computational and Mathematical Engineering. She Joined SEP in the summer of 2013.

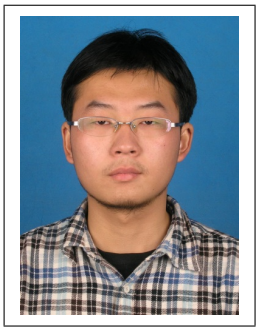
Paul Milligan graduated from the University of California at Berkeley with both M.Sc. and Ph.D. degrees in Engineering Geoscience in 1999. He has worked extensively in the petroleum industry specializing in seismic data acquisition and processing, including marine streamer seismic, ocean bottom seismic, VSP, and crosswell seismic, while working for the following companies: GSI, Sonics Exploration, Geco-Prakla, Baker Hughes, Paulsson Geophysical services, Seabird Exploration, and currently Seabed Geosolutions. He also has experience working further downstream on velocity inversion and wave equation migration imaging while working for 4th Wave Imaging and 3DGeo. He is currently working offshore for Seabed Geosolutions, processing ocean bottom seismic data, and developing quality control methods and software for the multi-component data.

Michelle Robertson is a project scientist in the Earth Sciences Division of Lawrence Berkeley National Laboratory.

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.

Craig Ulrich is a scientific engineering associate in the Earth Sciences Division of Lawrence Berkeley National Laboratory.

Anna Wagner is a research environmental engineer at the U.S. Army Cold Regions Research and Engineering Laboratory in Fort Wainwright, Alaska.



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.

Tieyuan Zhu is a Postdoctoral Fellow in the Jackson School of Geosciences at the University of Texas at Austin. He is an exploration geophysicist who is interested in using seismic waves to produce detailed 3D images of subsurface structure and its physical properties. He got his PhD degree from Dept. of Geophysics,Stanford in 2014.

SEP ARTICLES PUBLISHED OR IN PRESS

Almomin, A., and Biondi, B., 2012, Tomographic full waveform inversion: Practical and computationally feasible approach. SEG Technical Program Expanded Abstracts, 1–5.

Almomin, A., and Biondi, B., 2013, Tomographic full waveform inversion (TFWI) by successive linearizations and scale separations. SEG Technical Program Expanded Abstracts, 1048–1052.

Almomin, A., and Biondi, B., 2014, Preconditioned tomographic full waveform inversion by wavelength continuation. SEG Technical Program Expanded Abstracts, 944–948.

Almomin, A., and Biondi, B., 2014, Generalization and amplitude normalization of tomographic full waveform inversion. SEG Technical Program Expanded Abstracts (Submitted).

Barak, O., Jaiswal, P., de Ridder, S., Giles, J., Brune, R. and Ronen, S., 2014, Six-component seismic land data acquired with geophones and rotation sensors: wave-mode separation using 6C SVD: SEG Technical Program Expanded Abstracts (Submitted).

Chang, J., de Ridder, S., and Biondi, B., 2013, Power spectral densities and ambient noise cross-correlations at Long Beach: SEG Technical Program Expanded Abstracts **32**, 2196–2200.

Chang, J., Nakata, N., Clapp, R.G., de Ridder, S., and Biondi, B., 2014, High-frequency surface and body waves from ambient noise cross-correlations at Long Beach, CA: SEG Technical Program Expanded Abstracts **33**.

Chang, J., and B. Biondi, 2015, Rayleigh-wave tomography using traffic noise at Long Beach, CA: SEG Technical Program Expanded Abstracts **34** (Submitted).

Dahlke, T., Beroza, G., Chang, J., and de Ridder, S., 2014, Stochastic variability of velocity estimates using eikonal tomography on the Long Beach data set: SEG Technical Program Expanded Abstracts **33** .

Dahlke, T., Biondi, B., Clapp, R., 2015, Domain decomposition of level set updates for salt segmentation: SEG Technical Program Expanded Abstracts **33** (Submitted).

Grobbe, N., F.C. Schoemaker, M.D. Schakel, S.A.L. de Ridder, E.C. Slob, and D.M.J. Smeulders, 2012, Electrokinetic fields and waves: Theory, experiments, and numerical modeling. Geophysical Research Abstracts, Vol. **14**, EGU2012–10636.

de Ridder, S., and B. Biondi, 2012, Continuous passive seismic monitoring of CCS projects by correlating seismic noise: A feasibility study. 74th Conference & Technical Exbition, EAGE, Extended Abstracts (Accepted).

de Ridder, S., and B. Biondi, 2012, Reservoir monitoring by passive seismic interferometry. SEG/SPE/AAPG Joint Workshop, La Jolla, CA, 24-29 June 2012 (Accepted).

de Ridder, S., 2012, Ambient seismic noise correlations for reservoir monitoring. SEG Technical Program Expanded Abstracts **31** (Submitted).

de Ridder, S., and B. Biondi, 2012, Continuous reservoir monitoring by ambient seismic noise tomography. SEG Technical Program Expanded Abstracts **31** (Submitted).

Schoemaker, F.C., N. Grobbe, M.D. Schakel, S.A.L. de Ridder, E.C. Slob, and D.M.J. Smeulders, 2012, Experimental validation of the electrokinetic theory and development of seismoelectric interferometry by cross-correlation: International Journal of Geophysics (Accepted).

Farghal, M.S., and Levin, S.A., 2012, Hunting for microseismic reflections using multiplets: SEG Technical Program Expanded Abstracts **31** (Submitted).

Guitton, A., Ayeni, G., and Esteban, D.A., 2012, Constrained full-waveform inversion by model reparameterization: Geophysics, 77, No. 2, R117–R127.

Guitton, A., and Esteban, D.A., 2012, Attenuating crosstalk noise with simultaneous source full waveform inversion: Geophysical Prospecting (In Press).

Guitton, A., 2012, Blocky regularization schemes for full waveform inversion: Geophysical Prospecting (In Press).

Halpert, A., 2012, Edge-preserving smoothing for segmentation of seismic images: SEG Technical Program Expanded Abstracts **31**, 1–5.

Halpert, A., Clapp, R.G., and B. Biondo, 2014, Salt delineation via interpreter-guided seismic image segmentation, Interpretation: **2**, T79–T88.

Arevalo, Humberto and Stewart A. Levin, 2014, Well and Seismic matching with ArcGIS and ProMAX via KML: Esri International User Conference Paper 901, San Diego, 15 July.

Leader, C., and R. Clapp, 2012, Least squares reverse time migration on GPUs - balacing IO and computation: 74th Conference & Technical Exbition, EAGE, Extended Abstracts (Accepted).

Leader, C., and A. Almomin, 2012, How incoherent can we be? Phase encoded linearised inversion with random boundaries: SEG Technical Program Expanded Abstract **31** (Submitted).

Levin, Stewart A. and Fritz Foss, 2014, Downward continuation of Mars SHARAD data: SEG Technical Program Expanded Abstract **33** (Submitted).

Li, Y., Y. Zhang, and J. Claerbout, 2012, Hyperbolic estimation of sparse models from erratic data: Geophysics **77**, 1–9.

Li, Y., P. Shen, and C. Perkins, 2012, VTI migration velocity analysis using RTM: SEG Technical Program Expanded Abstract **31**.

Li, Y., Image-guided WEMVA for azimuthal anisotropy: SEG Technical Program Expanded Abstract **32**.

R.E. Plessix and Y. Li, 2013, Waveform acoustic impedance inversion with spectral shaping: Geophysical Journal International **195**(1), 301–314, 2013.

Li, Y., B. Biondi, D. Nichols, and R. Clapp, 2014, Wave equation migration velocity analysis for VTI models: Accepted for publication in Geophysics, 2014.

Li, Y., B. Biondi, D. Nichols, and R. Clapp, 2014, Wave equation migration velocity analysis for VTI models: Accepted for publication in Geophysics, 2014.

Li, Y., M. Wong, and R. Clapp, 2014, Equivalent accuracy by fractional cost: overcoming temporal dispersion: Submitted to SEG 2014.

Li, Y., D. Nichols, and G. Mavko, 2014, Stochastic rock physics modeling for seismic anisotropy: Submitted to SEG 2014.

Li, Y., B. Biondi, R. Clapp, and D. Nichols, 2014, Rock physics constrained anisotropic wave-equation migration velocity analysis: Submitted to SEG 2014.

Maharramov M., and Biondi, B., 2014, Joint full-waveform inversion of time-lapse seismic data sets. SEG Technical Program Expanded Abstracts, 954–959.

Maharramov M., and Biondi, B., 2014, Joint 4DFWI with model-difference regularization. SEG-AGU Summer Research Workshop. Advances in Active+Passive "Full Wavefield" Seismic Imaging: From Reservoirs to Plate Tectonics.

Maharramov M., and Biondi, B., 2015, Robust Simultaneous Time-lapse Full-waveform Inversion with Total-variation Regularization of Model Difference. EAGE Technical Program Expanded Abstracts (accepted).

Maharramov M., 2015, Efficient Finite-difference Modelling of Acoustic Wave Propagation in Anisotropic Media with Pseudo-sources. EAGE Technical Program Expanded Abstracts (accepted).

Maharramov M., and Biondi, B., 2015, Simultaneous TV-regularized time-lapse FWI with application to field data. SEG Technical Program Expanded Abstracts (submitted).

Martin E.R., J. Ajo-Franklin, S. Dou, N. Lindsey, T.M. Daley, B. Freifeld, M. Robertson, A. Wagner, and C. Ulrich, 2015, Interferometry of ambient noise from a trenched distributed acoustic sensing array. SEG Technical Program Expanded Abstracts (submitted).

Ajo-Franklin, J., N. Lindsey, S. Dou, T.M. Daley, B. Freifeld, E.R. Martin, M. Robertson, C. Ulrich, and A. Wagner, 2015, A field test of distributed acoustic sensing for ambient noise recording. SEG Technical Program Expanded Abstracts (submitted).

Li, Y., H. Yang, E.R. Martin, K. Ho, and L. Ying, 2015, Butterfly factorization. Multiscale Modeling and Simulation (submitted). http://arxiv.org/abs/1502.01379

Shen, Y., B. Biondi, R. Clapp, and D. Nichols, 2013, Wave-equation migration Q analysis (WEMQA): EAGE Workshop on Seismic Attenuation Extended Abstract 2013.

Shen, Y., B. Biondi, R. Clapp, and D. Nichols, 2014, Wave-equation migration Q analysis (WEMQA): SEG Technical Program Expanded Abstracts.

Li, Y., Y. Shen and P. Kang, 2015, Integration of seismic and fluid-flow data: a two-way road linked by rock physics :77th EAGE Conference and Exhibition.

Wong, M., and Ronen, S., and Biondi, B.L., 2012, Joint imaging with streamer and ocean bottom data. SEG Technical Program Expanded Abstracts, 1–5.

Wong, M., and Ronen, S., and Biondi, B.L., 2012, Imaging with multiples using linearized full-wave inversion. SEG Technical Program Expanded Abstracts, 1–5.

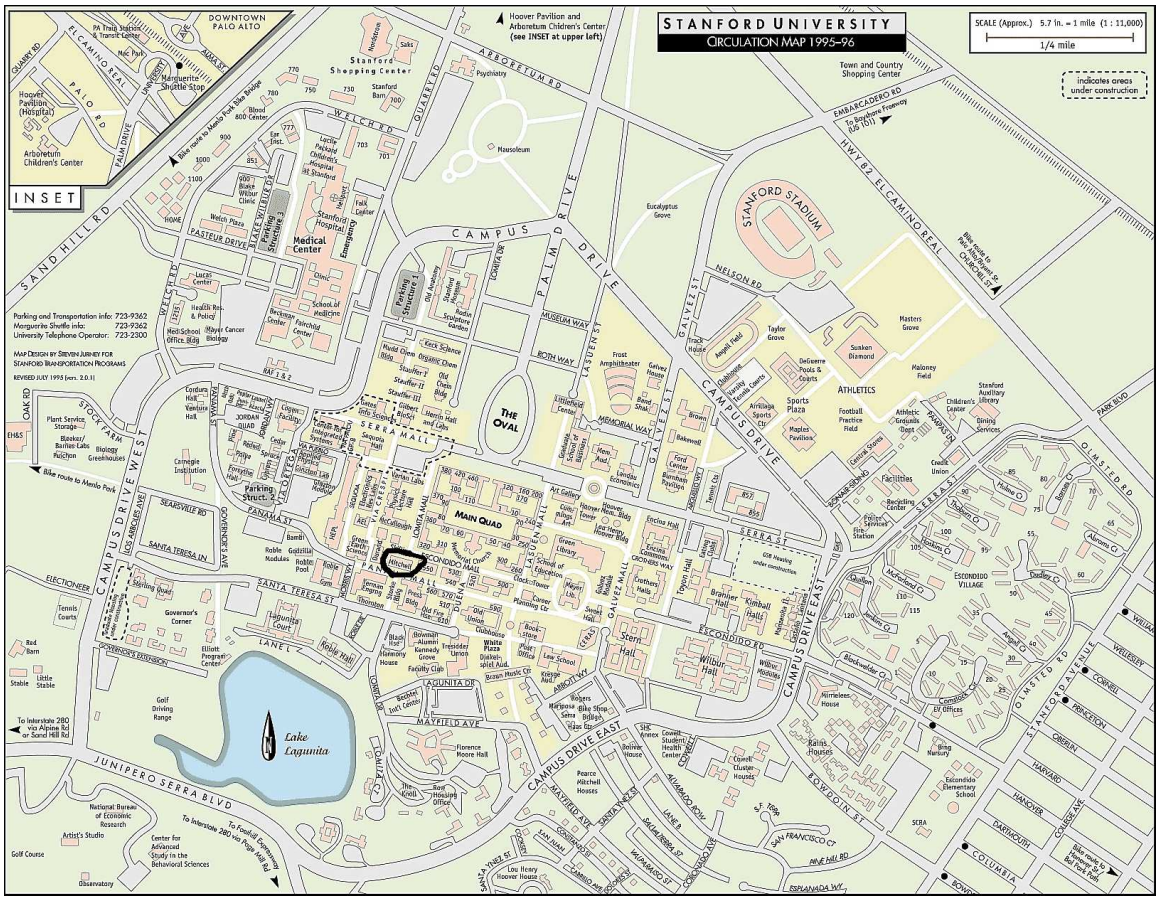
Wong, M., 2013, Handling salt reflection in Least-squares RTM SEG Technical Program Expanded Abstracts, 3921–3925.

Wong, M., B. Biondi, S. Ronen, C. Perkins, M. Merritt, V. Goh, and R. Cook, 2014, Robbust least squares RTM on the 3D Deimos ocean bottom node dataset SEG Technical Program Expanded Abstracts, (Submitted)

Wong, M., B. Biondi, and S. Ronen, 2014, Imaging with multiples using least-squares reverse-time migration. The Leading Edge, (submitted)

Huy Le and Stewart A. Levin, 2014, Removing shear artifacts in acoustic wave propagation in orthorhombic media: SEG Technical Program Expanded Abstracts, 486–490.

Huy Le, Biondo Biondi, Robert G. Clapp, and Stewart A. Levin, 2015, Using a nonlinear acoustic wave equation for anisotropic inversion: 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
Hoover Pavilion (Hospital)
Arboretum Children's Center
Mausoleum
Eucalyptus Grove
STANFORD STADIUM
THE OVAL
MAIN QUAD
Lake Lagunita
Golf Driving Range
Golf Course
Observatory
Stern Hall
Wilbur Hall
Lou Henry Hoover House
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 (ver. 2.0.1)