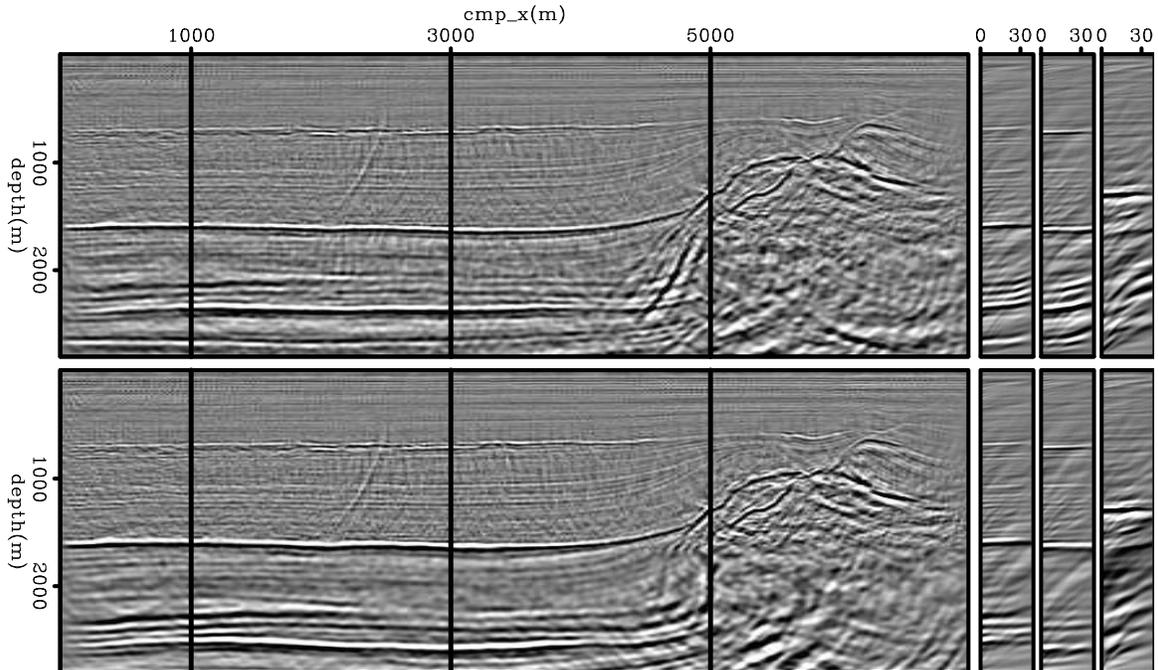


STANFORD EXPLORATION PROJECT

*Ali Almomin, Gboyega Ayeni, Ohad Barak, James Berryman, Biondo Biondi,
Jon Claerbout, Robert Clapp, Claudio Guerra, Adam Halpert, Benxi Ke, Chris Leader,
Yunyue Li, Nader Moussa, Sjoerd de Ridder, Shuki Ronen, Xukai Shen, Yaxun Tang,
Mandy Wong, and Yang Zhang*

Report Number 140, May 2010



Copyright © 2010

by the Board of Trustees of the Leland Stanford Junior University

Copying permitted for all internal purposes of the Sponsors of Stanford Exploration Project

Preface

The electronic version of this report¹ 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 library². We assume you have a UNIX workstation with Fortran, Fortran90, C, X-Windows system and the software downloadable from our website (SEP makerules, SEPlib, and the SEP latex package), or other free software such as SU. Before the publication of the electronic document, someone other than the author tests the author's claim by destroying and rebuilding all ER figures. Some ER figures may not be reproducible by outsiders because they depend on data sets that are too large to distribute, or data that we do not have permission to redistribute but are in the SEP data library.

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

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

Our testing is currently limited to LINUX 2.6 (using the Intel Fortran90 compiler), but the code should be portable to other architectures. Reader's suggestions are welcome. More information on reproducing SEP's electronic documents is available online³.

¹<http://sepwww.stanford.edu/private/docs/sep140>

²<http://sepwww.stanford.edu/public/docs/sepdata/lib/toc.html>

³<http://sepwww.stanford.edu/research/redoc/>

SEP140 — TABLE OF CONTENTS

Wave Equation Tomography

<i>Claudio Guerra and Biondo Biondi</i> , Fast 3D velocity updates using the pre-stack exploding reflector model.....	1
<i>Yaxun Tang and Biondo Biondi</i> , Target-oriented wavefield tomography using demigrated Born data.....	11
<i>Biondo Biondi</i> , Wave-equation tomography by beam focusing.....	23
<i>Xukai Shen</i> , Near-surface velocity estimation by weighted early-arrival waveform inversion.....	39
<i>Yunyue (Elita) Li and Biondo Biondi</i> , Wave-equation tomography for anisotropic parameters.....	49
<i>Ali Ameen Almomin</i> , Wave-equation travelttime tomography by global optimization.....	61

Imaging

<i>Yaxun Tang and Biondo Biondi</i> , Angle-domain illumination gathers by wave-equation-based methods.....	67
<i>Mandy Wong, Biondo Biondi, and Shuki Ronen</i> , Joint least-squares inversion of up- and down-going signal for ocean bottom data sets.....	83
<i>Robert G. Clapp</i> , More fun with random boundaries.....	95
<i>Ohad Barak</i> , Implicit finite difference in time-space domain with the helix transform.....	103
<i>Benxi Ke</i> , Predicting rugged water-bottom multiples through wavefield extrapolation with rejection and injection.....	119

Hybrid norm inversion

<i>Yang Zhang and Jon Claerbout</i> , Least-squares imaging and deconvolution using the hb norm conjugate-direction solver.....	129
<i>Ali Ameen Almomin</i> , Blocky velocity inversion by hybrid norm.....	143
<i>Yunyue Li, Yang Zhang, and Jon Claerbout</i> , Geophysical applications of a novel and robust L1 solver.....	155

Seismic Monitoring

<i>Gboyega Ayeni, Yaxun Tang, and Biondo Biondi</i> , Seismic reservoir monitoring with simultaneous sources.....	165
---	-----

<i>Sjoerd de Ridder</i> , Low frequency passive seismic interferometry for land data	181
<i>Gboyega Ayeni</i> , Cyclic 1D matching of time-lapse seismic data sets: A case study of the Norne Field	193

Interpretation and Visualization

<i>Adam Halpert</i> , A new method for more efficient seismic image segmentation	213
<i>Robert G. Clapp</i> , Hypercube viewer update	229
<i>Nader Moussa</i> , Short note: Enhanced visualization for seismic imaging .	233

Miscellaneous

<i>James G. Berryman</i> , Mechanics of stratified anisotropic poroelastic media	235
<i>Chris Leader and Sjoerd de Ridder</i> , Fourier methods of seismic data regularisation	251
SEP phone directory	263
Research personnel	265
(‘SEP article published or in press, 2010’.)	272

Fast 3D velocity updates using the pre-stack exploding reflector model

Claudio Guerra and Biondo Biondi

ABSTRACT

We show 3D real data results of migration velocity analysis by wavefield extrapolation using data synthesized by the pre-stack exploding-reflector model (PERM). PERM is a generalization of the exploding reflector model (ERM) in the sense that migration of PERM data can generate a pre-stack image, which is not achievable with ERM. PERM data allow fast migration velocity analysis, since its size can be orders of magnitude smaller than conventional data reduction approaches such as those used in plane wave migration. Also, because the modeling can be limited to the region where the velocity is to be updated, additional savings are possible by solving for the velocity in a target-oriented manner.

INTRODUCTION

In areas of complex geology, pre-stack depth migration is required not only for imaging purposes but also for velocity estimation. In such areas, migration by wavefield extrapolation has been widely used to produce the final image because it properly handles complex distortions of the wavefields. Migration velocity analysis by wavefield extrapolation (Sava and Biondi, 2004; Shen and Symes, 2008) promises to produce more reliable results than ray-based methods in those areas. However, due to the high computational cost, wavefield-extrapolation methods are rarely used to estimate the migration velocity model in 3D projects (Fei et al., 2009), and ray-based methods are the industry standards. In addition to the lower cost, ray-based methods are very flexible with respect to strategies for defining the velocity model (Stork, 1992; Kosloff et al., 1997; Billette et al., 1997). But despite their advantages, ray methods do not satisfactorily describe complex wave propagation in the presence of large lateral velocity contrasts. In this case, a more complete description of the wavefield complexities is needed, and therefore we face the challenge of decreasing the cost of migration velocity analysis by wavefield extrapolation while maintaining its robustness.

A typical way of decreasing the cost of wavefield extrapolation is to reduce the amount of input data by linearly combining the wavefields using plane-wave phase encoding (Whitmore, 1995; Liu et al., 2006) and random phase encoding (Romero et al., 2000; Sun et al., 2002). Combination of wavefields is exploited by the prestack-exploding reflector model (PERM) (Biondi, 2006) to significantly reduce the data size while keeping the migrated image crosstalk-free. This is achieved by selecting subsurface-offset common-image gathers (SODCIGs) separated by a decorrelation distance such that the wavefields from different SODCIGs in the same modeling experiment do not correlate during migration. Guerra et al. (2009) show that further reduction can be achieved by randomly phase encoding the modeling experiments, significantly decreasing the cost of migration velocity analysis iterations.

An interesting feature of PERM is that, because the wavefields are upward propagated, data can be collected at any depth level. In migration velocity analysis, PERM wavefields can be propagated only in the region with velocity inaccuracies. As a result the velocity update can be performed in a target-oriented way, which contributes to an additional cost reduction of migration velocity analysis.

In 3D, reduction of the data size can be considerable if the initial image used to model PERM wavefields has only in-line subsurface offsets, as in the common-azimuth approximation. We show that, in this case, 3D-PERM data size can be potentially up to two orders of magnitude smaller than 3D-plane wave data. The usefulness of PERM data to migration-velocity analysis with wavefield extrapolation is illustrated using a North Sea 3D dataset.

PRE-STACK-EXPLODING-REFLECTOR MODEL

The fundamental idea of PERM is to model data that describes the correct kinematics of an isolated SODCIG. When using Born modeling, since we do not know beforehand which shots contribute to forming the image at a point in the model space, we have to model several shots. Ideally, instead of performing many modeling experiments, we would like to synthesize a small amount of data with the condition that migration has the same kinematics as the initial SODCIG. This can be achieved by extrapolating source and receiver wavefields starting from a pre-stack image computed with wave-extrapolation methods.

The modeling of PERM source D_P and receiver U_P wavefields can be carried out by any wavefield-continuation scheme. Here, we use the following one-way wave equations:

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2 s_0^2(\mathbf{x}) - |\mathbf{k}|^2} \right) D_P(\mathbf{x}, \omega; \mathbf{x}_m) = I_D(\mathbf{x}_m, \mathbf{h}) \\ D_P(x, y, z = z_{\max}, \omega; \mathbf{x}_m) = 0 \end{cases}, \quad (1)$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s_0^2(\mathbf{x}) - |\mathbf{k}|^2} \right) U_P(\mathbf{x}, \omega; \mathbf{x}_m) = I_U(\mathbf{x}_m, \mathbf{h}) \\ U_P(x, y, z = z_{\max}, \omega; \mathbf{x}_m) = 0 \end{cases}, \quad (2)$$

where $I_D(\mathbf{x}_m, \mathbf{h})$ and $I_U(\mathbf{x}_m, \mathbf{h})$ is the isolated SODCIG at the horizontal location \mathbf{x}_m for a single reflector, suitable for the initial conditions for the source and receiver wavefields, respectively. The subsurface-offset \mathbf{h} is parameterized as $\mathbf{h} = (h_x, h_y)$. The initial conditions are obtained by rotating the original unfocused SODCIGs according to the apparent geological dip of the reflector (Biondi, 2007). By doing so, image-point dispersal is corrected such that the velocity information needed for migration velocity analysis is maintained.

To decrease the number of modeling experiments, linearity of wave propagation can be used to combine isolated SODCIGs and inject them simultaneously into one single model experiment, using the same modeling equations as above with the initial conditions replaced by the combined SODCIGs. The selection of SODCIGs can be thought of as the multiplication of the pre-stack image by spatial 2D *Comb* functions, which are shifted laterally to select new set of SODCIGs to initiate the modeling of another pair of combined wavefields. After shifting along one period of the sampling function in the x and y directions, all the points on the reflector are used in the modeling. Consequently, the number of modeling experiments equals the number of lateral shifts of the sampling function.

The choice of the sampling period determines the amount of crosstalk in the migrated image. To obtain a crosstalk-free image, the sampling period must be large enough that wavefields initiated at different SODCIGs do not correlate. PERM wavefields generated from SODCIGs at an interval equals to twice the subsurface-offset range still contribute to the image at the central SODCIG. For the same reason, no crosstalk is generated during migration if the period of the sampling function is larger than that interval. Since the focusing of energy in the SODCIG is velocity-error dependent, so is the the period of the sampling function. Therefore, for small velocity errors a small sampling period can be used and, consequently, a smaller number of combined modeling experiments is needed.

PERM using a common-azimuth migrated image

In the way PERM is formulated, there is no restriction on the dimensionality of the pre-stack image used as the initial condition for the modeling, which means that if the original data have sufficient cross-line offsets as in the acquisition geometries with wide range of azimuths (Regone, 2007; Kapoor et al., 2007), the initial conditions are a five-dimensional hypercube on \mathbf{x} , h_x and h_y . To synthesize PERM data starting with the five-dimensional initial conditions such that no crosstalk is generated during migration, the minimum number of modeling experiments is $2n_{h_x}n_{h_y}$, where n_{h_x} and n_{h_y} are the number of subsurface offsets in the x and y directions. Considering usual parameters, the number of modeling experiments may be as low as several hundreds. This data reduction is very substantial if we compare, for instance, with data reduction achieved by 3D-plane-wave decomposition. Using plane waves, to obtain artifact-free SODCIGs due to the lack of illumination from some propagation directions we need to migrate few thousands of plane waves. This means that 3D-PERM data size can be one order of magnitude smaller than 3D-plane wave data.

Despite the recent good migration results obtained in geologically complex areas using wide-azimuth data, narrow-azimuth acquisition is still the industry standard. Narrow-azimuth data can be efficiently imaged by common-azimuth wave-equation migration (CAM) (Biondi and Palacharla, 1996). By assuming zero cross-line offset in contrast with the full-azimuth migration, instead of a five-dimensional hypercube, CAM images are four-dimensional hypercubes in \mathbf{x} and h_x . Because of the lower dimensionality of CAM images, when using them as the initial conditions to synthesize PERM data, the SODCIGs in the cross-line direction can be sampled continuously, as depicted in Figure 1b. Contrast this case with the five-dimensional initial conditions for the full azimuth case of Figure 1a. The continuous sampling of SODCIGs in the cross-line direction yields one more order of magnitude of data reduction. Therefore, under the common-azimuth assumption, 3D-PERM data size can be two orders of magnitude smaller than 3D-plane wave data.

To illustrate the validity of the above assumptions, a split-spread data with maximum offset of 1587.5 m was computed using 3D-Born modeling on a 30° dipping reflector with 45° azimuth with respect to the acquisition direction. There are 96 in-lines and cross-lines spaced 25 m apart. The offset interval is 25 m. The velocity used in the modeling is the 1D function $v(z) = (1500 + 0.5z)$ m/s. The Born data are input to CAM with a 5% slower velocity. Migration results can be seen in Figures 2a and 2b for SODCIGs positioned at $(x = 750 \text{ m}, y = 600 \text{ m})$ and $(x = 750 \text{ m}, y = 1000 \text{ m})$, respectively. The panel on the left is the SODCIG, which contains 21 subsurface offsets ranging from -250 m to 250 m . The panel in the middle is the in-line at zero subsurface offset, with $y = 600 \text{ m}$ (Figure 2a) and

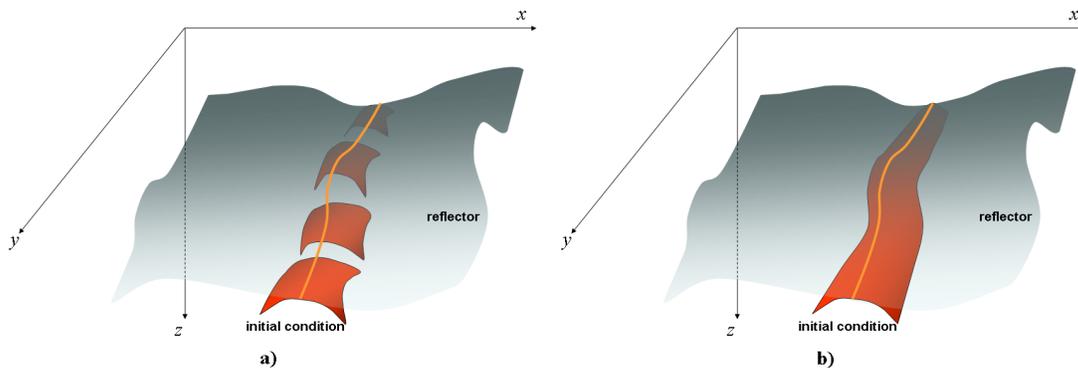


Figure 1: The initial conditions for synthesizing PERM data from CAM images can be specified as in b) because no pre-stack information exists in the cross-line direction, in contrast with the full azimuth situation in a). [NR] `claudio1/. cam01`

$y = 1000$ m (Figure 2b). The panel on the right is the cross-line at zero subsurface offset, with $x = 750$ m.

In the common-azimuth regime, the computation of the dip-independent initial conditions is performed by simply rotating the SODCIGs in the in-line direction, since no cross-line offset is computed in migration.

PERM source and receiver wavefields were modeled using as the initial conditions combined SODCIGs from the CAM image (Figure 2) with continuous sampling along the cross-line direction and sampling period of 48 in the in-line direction. This period is sufficient to avoid crosstalk during the areal-shot migration, given that the number of subsurface-offsets of the pre-stack image is 21. One synthesized 3D receiver wavefield is shown in Figure 3. The left panel is the in-line at $y = 1200$ m, the right panel is the cross-line at $x = 1400$ m, and the top panel is the time-slice at $t = 0.5$ s.

The 3D migration of the 48 areal shots with the velocity underestimated by 5% is shown in Figures 4a and 4b for SODCIGs positioned at $(x = 750$ m, $y = 600$ m) and $(x = 750$ m, $y = 1000$ m), respectively. The kinematics of the SODCIGs computed with PERM wavefields matches those of the SODCIGs computed with CAM. This enables the use of 3D PERM wavefields computed from CAM images in optimization of migration velocity, as will be shown next for a 3D survey from the North Sea.

3D VELOCITY UPDATE EXAMPLE

Migration velocity analysis by wavefield extrapolation using 3D PERM data was performed on a 3D real dataset from the North Sea with 3600 m maximum offset. This dataset was submitted to azimuth-moveout and imaged with CAM with an initial velocity sufficiently accurate up to the top of the chalk layer. From this reflector to deeper levels velocity is underestimated (Figure 5) as revealed by reflectors curving up in angle-domain common-image gathers at the bottom of Figure 6.

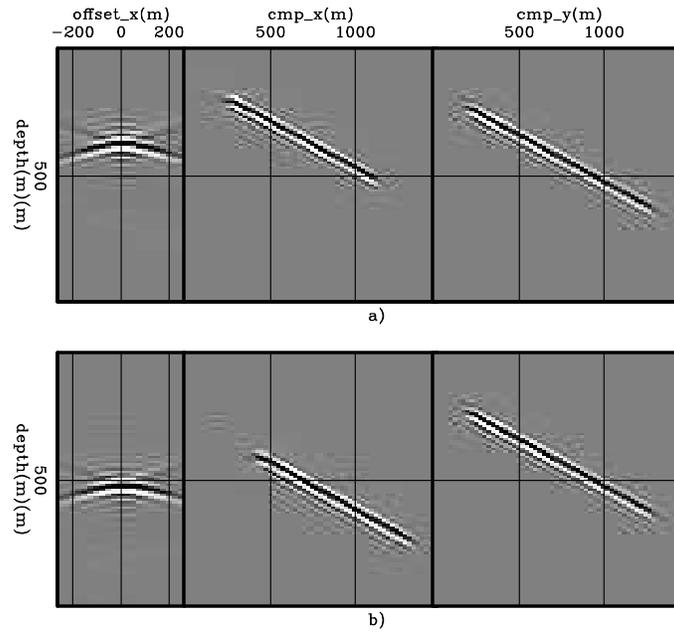


Figure 2: CAM of the 3D-Born data. Middle: in-line at zero-subsurface offset, and $y = 600$ m (Figure 2a) and $y = 1000$ m (Figure 2b). Right: cross-line at zero-subsurface offset, and $x = 750$ m. Left: SODCIGs. [CR] `claudio1/. cam02`

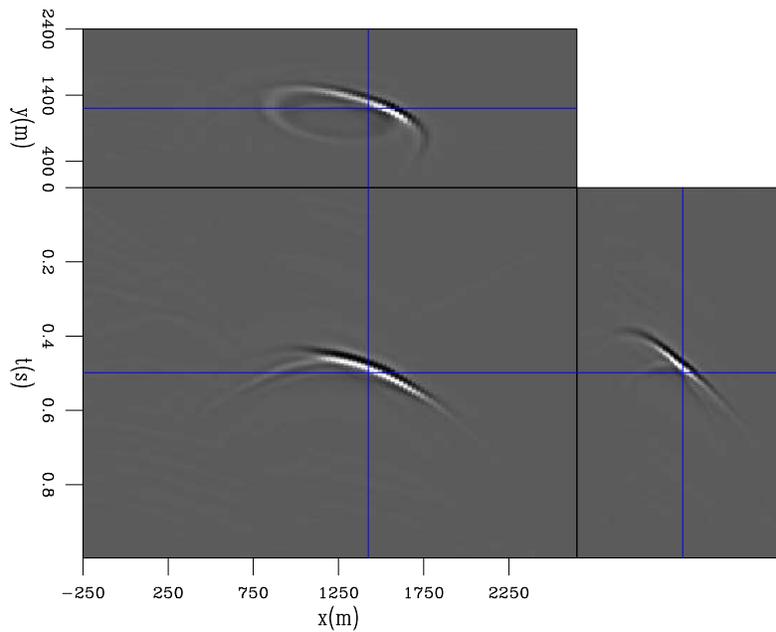


Figure 3: 3D-PERM receiver wavefield. The left panel is the in-line at $y = 1200$ m, the right panel is the cross-line at $x = 1400$ m, and the top panel is the time-slice at $t = 0.5$ s. [CR] `claudio1/. cam03`

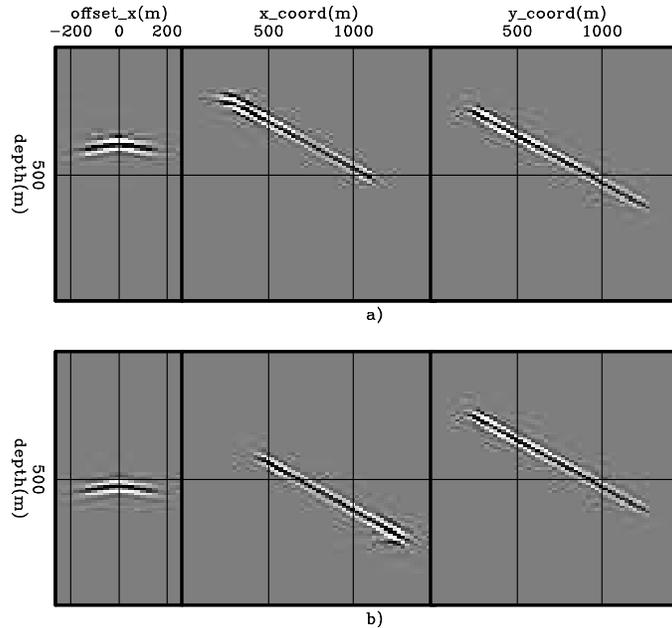


Figure 4: 3D-areal-shot migration of PERM data. Middle: in-line at zero-subsurface offset, and $y = 600$ m (Figure 4a) and $y = 1000$ m (Figure 4b). Right: cross-line at zero-subsurface offset, and $x = 750$ m. Left: SODCIGs. [CR] claudio1/. cam04

We compute velocity updates for the chalk layer in part of the 3D volume, which has 12 km in the in-line direction and 4 km in the cross-line direction. The velocity model definition for the entire 3D cube is ongoing work. We start by interpreting and windowing the base of chalk in the four-dimensional pre-stack image. The base of chalk is indicated by **BC** in Figure 6. The top of chalk, indicated by **TC** in Figure 6, is used as the upper limit for the velocity updates. The windowed reflector is then submitted to the rotation according to the apparent geological dip to correct for the image-point dispersal and it is used as the initial conditions to model 45 pairs of PERM source and receiver wavefields. Figure 7 shows the PERM receiver wavefield of one pair of wavefields. The wavefields are initiated by SODCIGs 1350 m apart. Since the offset range computed in the velocity optimization is 600 m, no crosstalk is generated. The wavefields are collected at 600 m depth, which is close to the top of the target-zone. Therefore, during the migration velocity analysis the wavefields are propagated between this depth level and the deeper level, which is 2800 m.

To optimize the velocity, we use a nonlinear conjugate gradient algorithm. The objective function we minimize is computed via differential-semblance optimization (DSO) (Symes and Carazzone, 1991; Shen and Symes, 2008), which corresponds to weighting the pre-stack image computed using the current velocity with the absolute value of the subsurface offset. The updated velocity is constrained to be within bounds 50%-lower and 50%-higher the initial velocity. Because of unbalanced amplitudes in the original data, it is necessary to smooth the gradient. We apply a B-spline smoothing, with node xy -spacing of 1000 m. The optimization stopped after 5 iterations.

The optimized velocity model is shown in Figure 8 for the in-line at 3760 m and cross-line at 3500 m. Notice that, as expected, the optimized velocity is higher than the initial

velocity. CAM using the optimized velocity confirms the correctness of the velocity update, by showing more focused reflectors and flatter ADCIGs than that obtained with the initial velocity (Figure 9). Compare with Figure 6. Some residual moveout is still present as can be seen in the rightmost ADCIGs of Figure 9. Since, the velocity of a salt body that occurs close to this region was not edited by a salt flooding procedure, we expect additional improvements in the image migrated with the final velocity model.

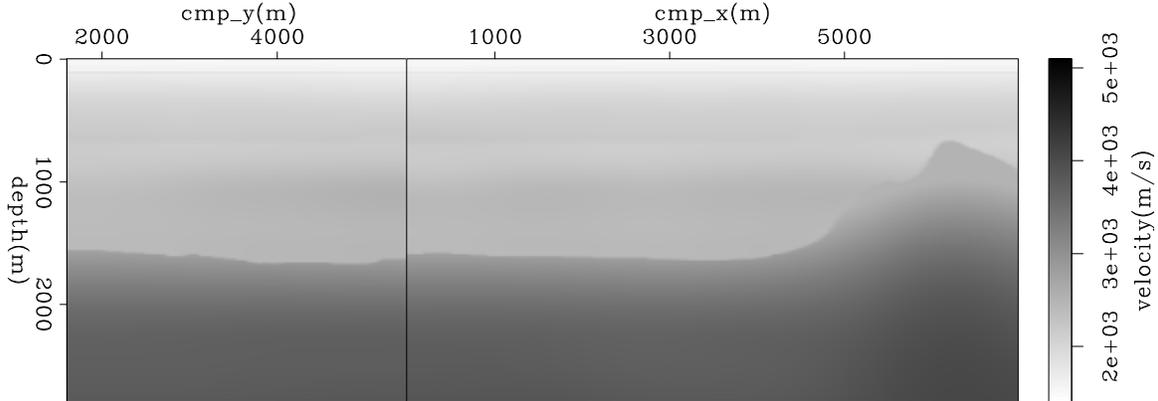


Figure 5: Initial velocity model. Left: cross-line at 3500 m. Right: in-line at 3760 m. [ER] claudio1/. nsea03

CONCLUSIONS

We show that pre-stack exploding-reflector model synthesizes wavefields that provides migrated images with correct kinematics while decreasing the data size. Data reduction is achieved by combining the modeling experiments and is controlled by the number of subsurface offsets that will be computed during areal-shot migration of PERM data. Implicit to PERM is that reflectors must be identified. 3D Pre-stack interpretation can be cumbersome, but it allows, for instance, avoiding the use of reflectors with low signal-to-noise ratios in migration velocity estimation. In 3D, the size of PERM data can be one order of magnitude smaller if cross-line subsurface-offsets are to be computed. Further data size reduction by one order of magnitude is achieved if the initial conditions are computed with common-azimuth migration.

In addition to the reduced data size, the definition of the velocity model is greatly accelerated due to target-oriented nature of PERM, which allows wavefields to be propagated only in the region where the velocity is to be updated. The 3D example with real data confirms the accuracy of the velocity solution achieved by migration velocity analysis by wavefield extrapolation using PERM wavefields.

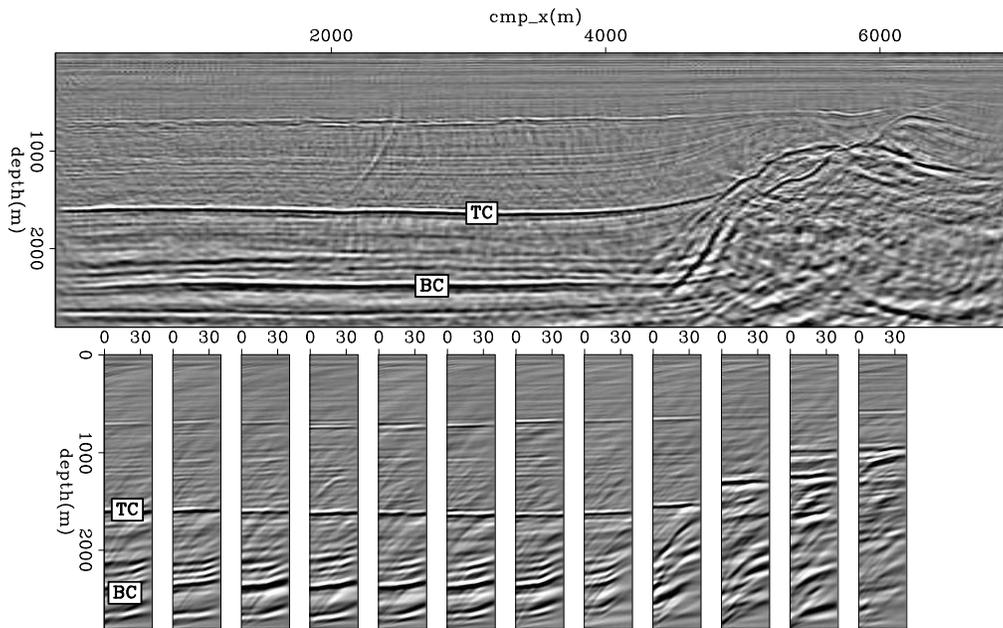


Figure 6: Angle-gathers after CAM with the initial velocity. Top: zero-angle section. Bottom: ACDCIGs from 0° to 40° selected roughly at the same x position in the section above. [CR] [claudio1/. nsea01](#)

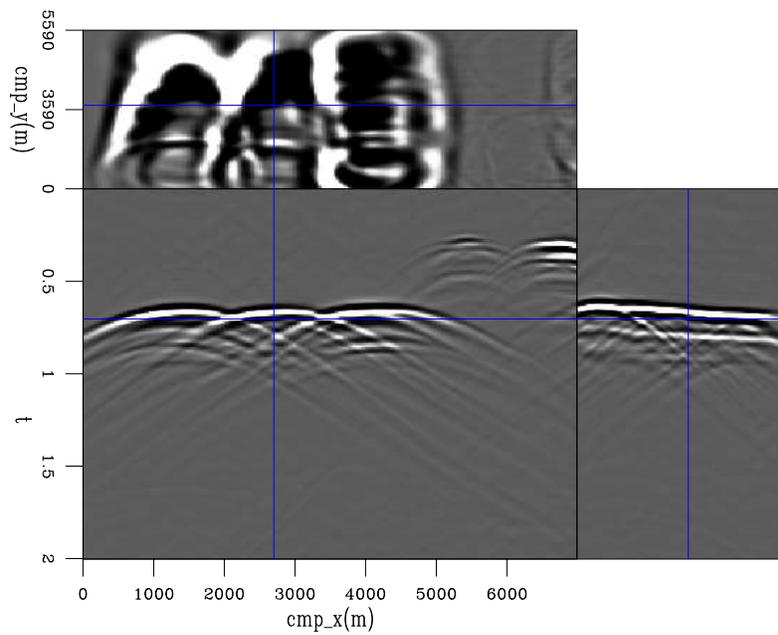


Figure 7: PERM receiver wavefield of one areal shot. [CR] [claudio1/. nsea00](#)

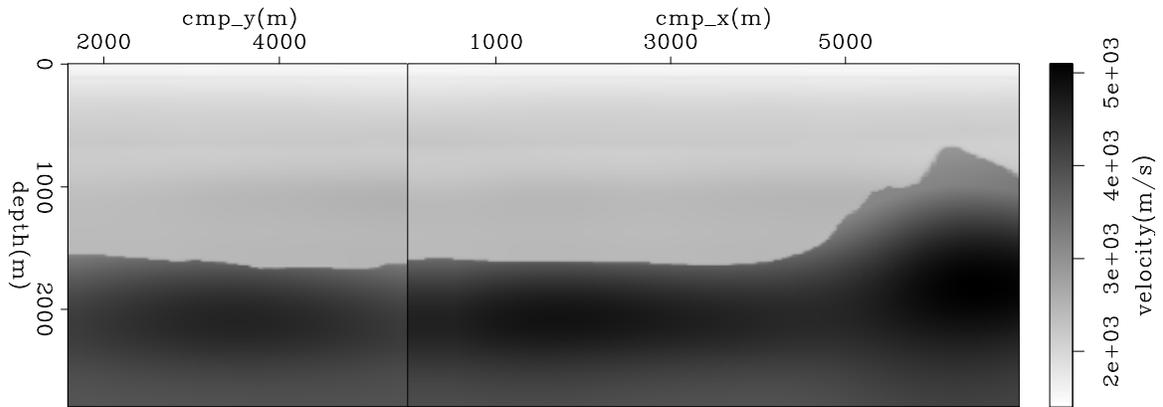


Figure 8: Optimized velocity model. Left: cross-line at 3500 m. Right: in-line at 3760 m.

[CR] claudio1/. nsea04

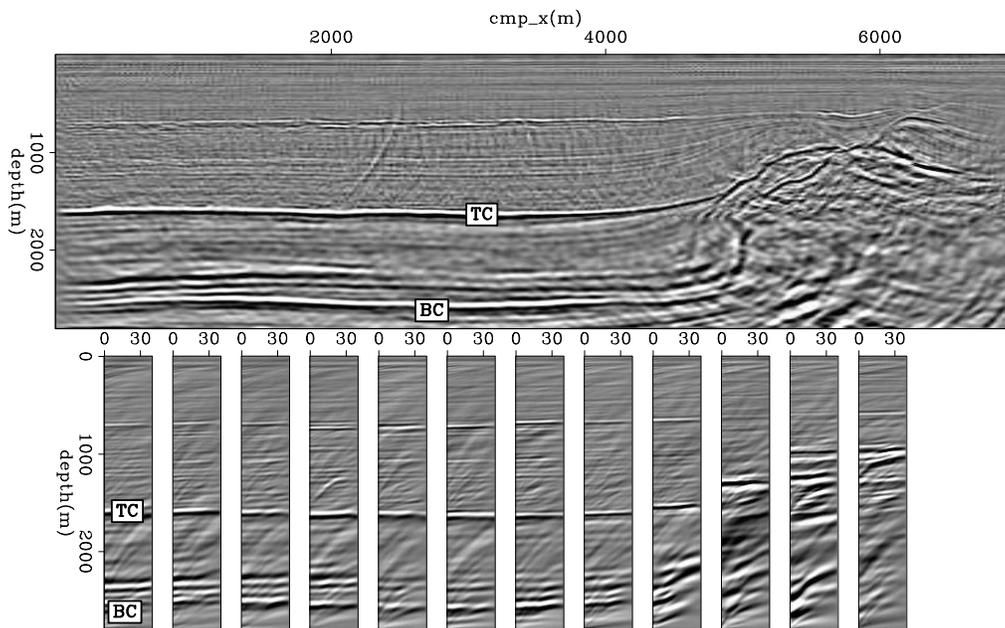


Figure 9: Angle-gathers after CAM with the optimized velocity. Top: zero-angle section. Bottom: ACDCIGs from 0° to 40° selected roughly at the same x position in the section above.

[CR] claudio1/. nsea02

ACKNOWLEDGMENTS

The authors thank TotalFinaElf for providing the data used in this report.

REFERENCES

- Billette, F., G. Lambare, and P. Podvin, 1997, Velocity macromodel estimation by stereotomography: SEG Technical Program Expanded Abstracts, **16**, 1889–1892.
- Biondi, B., 2006, Prestack exploding-reflectors modeling for migration velocity analysis: SEG Technical Program Expanded Abstracts, **25**, 3056–3060.
- , 2007, Prestack modeling of image events for migration velocity analysis: **SEP-131**, 71–89.
- Biondi, B. and G. Palacharla, 1996, 3-D prestack migration of common-azimuth data: Geophysics, **61**, 1822–1832.
- Fei, W., P. Williamson, and A. Khoury, 2009, 3-D common-azimuth wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **28**, 2283–2287.
- Guerra, C., Y. Tang, and B. Biondi, 2009, Wave-equation tomography using image-space phase encoded data: SEG Technical Program Expanded Abstracts, **28**, 3964–3968.
- Kapoor, S. J., M. O’Brian, D. Desta, I. Atakishiyev, and M. Tomida, 2007, Subsalt imaging — the raz/waz experience: SEG Technical Program Expanded Abstracts, **26**, 926–930.
- Kosloff, D., U. I. Zackhem, and Z. Koren, 1997, Subsurface velocity determination by grid tomography of depth migrated gathers: SEG Technical Program Expanded Abstracts, **16**, 1815–1818.
- Liu, F., D. W. Hanson, N. D. Whitmore, R. S. Day, and R. H. Stolt, 2006, Toward a unified analysis for source plane-wave migration: Geophysics, **71**, 129–139.
- Regone, C. J., 2007, Using 3D finite-difference modeling to design wide-azimuth surveys for improved subsalt imaging: Geophysics, **72**, SM231–SM239.
- Romero, L., D. Ghiglia, C. Ober, and S. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, 426–436.
- Sava, P. and B. Biondi, 2004, Wave-equation migration velocity analysis-I: Theory: Geophysical Prospecting, **52**, 593–606.
- Shen, P. and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.
- Stork, C., 1992, Reflection tomography in the postmigrated domain: Geophysics, **57**, 680–692.
- Sun, P., S. Zhang, and F. Liu, 2002, Prestack migration of areal shot records with phase encoding: 72nd Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1172–1175.
- Symes, W. W. and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.
- Whitmore, N. D., 1995, An imaging hierarchy for common angle plane wave seismograms, *in* Ph.D. thesis, University of Tulsa.

Target-oriented wavefield tomography using demigrated Born data

Yaxun Tang and Biondo Biondi

ABSTRACT

We present a method to reduce the computational cost of image-domain wavefield tomography. Instead of using the originally recorded data for velocity estimation, the proposed method simulates a new data set obtained using Born modeling or demigration based on the initial image and gathers. The modeling can be performed in a target-oriented fashion, and it can use arbitrary types of source functions and acquisition geometries. Hence the size of the new data set can be substantially smaller than the original one. We demonstrate with numerical examples that the new data set correctly preserves velocity information useful for velocity estimation, and that it generates wavefield-tomography gradients similar to that obtained using the original data set. We apply the proposed method to a modified version of the Sigsbee2A model, where two square anomalies below the salt have been successfully recovered in a target-oriented fashion at a much lower computational cost.

INTRODUCTION

Velocity estimation is always a challenging task in exploration seismology. In the past decade, ray-based tomography has been widely used in practice to derive velocity models. Although ray-based methods are efficient, the infinite-frequency approximation and the caustics inherent in ray theory prevent them from accurately modeling complicated wave phenomena (Hoffmann, 2001). As seismic exploration is moving towards structurally complex areas, ray-based methods become less reliable. On the other hand, wave-equation-based tomography (Tarantola, 1984; Mora, 1989; Woodward, 1992; Pratt, 1999; Sava, 2004; Shen, 2004) uses wavefields as carriers of information. It more accurately describes the bandlimited wave phenomena, and therefore more suitable for complex geologies.

Wavefield tomography can be implemented in either data domain or image domain. In this paper, however, we mainly focus on the image-domain wavefield tomography, which is also widely known as wave-equation migration velocity analysis (Sava, 2004; Shen, 2004). It derives an optimum velocity model by driving an objective function defined in the image domain to its minimum. Despite its advantages in modeling bandlimited wavefields, practical application of image-domain wavefield tomography is still rare and small in scale due to its huge computational cost (Biondi and Sava, 1999; Shen et al., 2005; Albertin et al., 2006). The high cost arises because of the use of more expensive wavefield modeling engines. The other reason is that it lacks flexibility and the recorded full data set is usually used for velocity estimation.

Several methods have been proposed to make wavefield tomography more cost effective. The main idea is to reduce the size of the data used for velocity estimation. One method is to

assemble the originally recorded point-source gathers into a smaller number of areal-source gathers. Among others, the plane-wave source gather (Whitmore, 1995; Zhang et al., 2005; Shen and Symes, 2008; Tang et al., 2008) is the most popular one because the plane-wave phase-encoding function is effective in attenuating the crosstalk artifacts (Liu et al., 2006; Tang, 2008). The other method is to model a new data set in a target-oriented fashion using the concept of prestack-exploding-reflector modeling with a bottom-up strategy (Biondi, 2006, 2007; Guerra et al., 2009). However, the modeling generates crosstalk when multiple image events are modeled simultaneously. Stochastic encoding methods, such as random-phase encoding, seem to be the only encoding methods available to attenuate the crosstalk produced when imaging the image-domain encoded gathers (Guerra and Biondi, 2008a,b).

In this paper, we present a new strategy to reduce the size of the data set used for image-domain wavefield tomography. The proposed strategy combines advantages of both prestack-exploding-reflector modeling and data-domain encoding: Not only can it model a new data set in a target-oriented fashion, but it also can use plane-wave sources to effectively attenuate the crosstalk. We start with an initial image and gathers obtained using a starting velocity model. The initial image and gathers are further normalized using the diagonal of the imaging Hessian, efficiently computed using the phase-encoding method (Tang, 2009), to optimally compensate for the uneven subsurface illumination and remove the effects of the original acquisition geometry. Instead of using prestack-exploding-reflector modeling, we then use Born modeling or demigration (Stolt and Benson, 1986) to simulate the new data set. The modeling procedure is based on the single-scattering approximation to the full wave equation. The resulting Born modeled data is obtained by convolving the source wavefield, computed using any type of source function (e.g. plane-wave sources), with the initial image and gathers and then propagating the convolved wavefields to receiver locations, which can be located anywhere in the model. The target-oriented data set is obtained by only modeling image points within a target zone or several key reflectors that carry important velocity information. This target-oriented velocity analysis strategy is useful, because it allows us to use the most powerful velocity estimation tool to focus on improving velocities in the most challenging areas, e.g., subsalt regions, provided that velocities at other locations are sufficiently accurate, e.g., regions above the salt, where the velocities are usually very accurately determined even by ray-based tomography thanks to the relatively simple geologies.

In the next section, we briefly review the theory of Born modeling. In the subsequent sections, we apply the proposed target-oriented velocity-estimation method to invert the local velocity anomalies in a modified Sigsbee2A model.

TARGET-ORIENTED BORN WAVEFIELD MODELING

Seismic images can be obtained by applying shot-profile migration to the recorded data as follows:

$$m(\mathbf{x}, \mathbf{h}) = \frac{1}{H(\mathbf{x}, \mathbf{h})} \sum_{\omega} \sum_{\mathbf{x}_s} G_0^*(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) \sum_{\mathbf{x}_r} W(\mathbf{x}_r, \mathbf{x}_s) \times G_0^*(\mathbf{x} + \mathbf{h}, \mathbf{x}_r, \omega) d(\mathbf{x}_r, \mathbf{x}_s, \omega) \quad (1)$$

where $*$ denotes adjoint, ω is the angular frequency, and $m(\mathbf{x}, \mathbf{h})$ is the migrated image as a function of both image point $\mathbf{x} = (x, y, z)$ and subsurface half offset $\mathbf{h} = (h_x, h_y, h_z)$.

The frequency-domain data $d(\mathbf{x}_r, \mathbf{x}_s, \omega)$ is recorded at receiver position $\mathbf{x}_r = (x_r, y_r, 0)$ due to a source located at $\mathbf{x}_s = (x_s, y_s, 0)$; $W(\mathbf{x}_r, \mathbf{x}_s)$ is the acquisition mask operator, which contains unity values where we record data and zero values where we do not; $G_0(\mathbf{x}, \mathbf{x}_s, \omega)$ and $G_0(\mathbf{x}, \mathbf{x}_r, \omega)$ are the Green's functions connecting the source and receiver, respectively, to image point \mathbf{x} . The Green's functions are obtained using a starting velocity model \mathbf{v}_0 . Operator $H(\mathbf{x}, \mathbf{h})$ is the diagonal of the subsurface-offset-domain imaging Hessian defined as follows (Plessix and Mulder, 2004; Valenciano, 2008; Tang, 2009):

$$H(\mathbf{x}, \mathbf{h}) = \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} W^2(\mathbf{x}_r, \mathbf{x}_s) |G(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega)|^2 |G(\mathbf{x} + \mathbf{h}, \mathbf{x}_r, \omega)|^2, \quad (2)$$

the inverse of which partially compensates for the image distortion caused by uneven subsurface illumination and removes the effects of the original acquisition geometry (Rickett, 2003; Plessix and Mulder, 2004; Symes, 2008; Tang, 2009). By doing so, the initial image $m(\mathbf{x}, \mathbf{h})$ contains only the effects of the initial velocity used for migration and is independent from the way the original data is recorded.

We then use Born modeling (Stolt and Benson, 1986) to simulate a new data set based on the initial image obtained using equation 1. The modeling can be performed using arbitrary source functions and arbitrary acquisition geometries. The modeled data set using encoded areal sources reads (Tang, 2008)

$$\tilde{d}(\mathbf{p}_s, \tilde{\mathbf{x}}_r, \omega) = \sum_{\tilde{\mathbf{x}}} \sum_{\mathbf{h}} G_0(\tilde{\mathbf{x}} - \mathbf{h}, \mathbf{p}_s, \omega) G_0(\tilde{\mathbf{x}} + \mathbf{h}, \tilde{\mathbf{x}}_r, \omega) m(\tilde{\mathbf{x}}, \mathbf{h}), \quad (3)$$

where $\tilde{\mathbf{x}} = (\tilde{x}, \tilde{y}, \tilde{z})$ is an image point within the selected target zone; $\tilde{\mathbf{x}}_r = (\tilde{x}_r, \tilde{y}_r, \tilde{z}_r)$ is the receiver location used for modeling. It can differ substantially from the original receiver location \mathbf{x}_r , which can be only on the surface, i.e., $\mathbf{x}_r = (x_r, y_r, 0)$; $G_0(\tilde{\mathbf{x}}, \mathbf{p}_s, \omega)$ is the Green's function obtained using the encoded areal source, it can be written as a weighted sum of point-source Green's functions as follows:

$$G_0(\tilde{\mathbf{x}}, \mathbf{p}_s, \omega) = \sum_{\tilde{\mathbf{x}}_s} G_0(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}_s, \omega) \alpha(\tilde{\mathbf{x}}_s, \mathbf{p}_s, \omega), \quad (4)$$

where $\alpha(\tilde{\mathbf{x}}_s, \mathbf{p}_s, \omega)$ and \mathbf{p}_s are the encoding function and encoding index, respectively, to be specified later. Similar to the new receiver location $\tilde{\mathbf{x}}_r$, the new source location $\tilde{\mathbf{x}}_s = (\tilde{x}_s, \tilde{y}_s, \tilde{z}_s)$ can also be substantially different from the original source location \mathbf{x}_s . Note in particular that the Green's function used for modeling ($G_0(\tilde{\mathbf{x}}, \mathbf{p}_s, \omega)$) is computed using the same velocity model (\mathbf{v}_0) as that used for migrating the original data. Therefore, the modeling process undoes the effect of the starting velocity model \mathbf{v}_0 (at least in kinematics), resulting in a starting-velocity-independent data set.

The synthesized areal data can be imaged by using conventional migration of areal-source data as follows:

$$\hat{m}(\tilde{\mathbf{x}}, \mathbf{h}) = \sum_{\omega} \sum_{\mathbf{p}_s} G^*(\tilde{\mathbf{x}} - \mathbf{h}, \mathbf{p}_s, \omega) \sum_{\tilde{\mathbf{x}}_r} G^*(\tilde{\mathbf{x}} + \mathbf{h}, \tilde{\mathbf{x}}_r, \omega) \tilde{d}(\mathbf{p}_s, \tilde{\mathbf{x}}_r, \omega), \quad (5)$$

where Green's functions G 's are obtained using velocity model \mathbf{v} , which can be the same or different from the starting velocity model \mathbf{v}_0 , to be discussed later. Imaging the areal-source

data using equation 5, however, generates crosstalk artifacts (Romero et al., 2000; Liu et al., 2006; Tang, 2008). To attenuate the crosstalk, the encoding function α is chosen such that

$$\sum_{\mathbf{p}_s} \alpha(\tilde{\mathbf{x}}_s, \mathbf{p}_s, \omega) \alpha(\tilde{\mathbf{x}}'_s, \mathbf{p}_s, \omega) \approx \delta(\tilde{\mathbf{x}}_s - \tilde{\mathbf{x}}'_s), \quad (6)$$

where $\delta(\cdot)$ is the Dirac delta function. For plane-wave-phase encoding, $\alpha(\tilde{\mathbf{x}}_s, \mathbf{p}_s, \omega) = A(\omega)e^{i\omega\mathbf{p}_s \cdot \tilde{\mathbf{x}}_s}$, with $A^2(\omega) = |\omega|$ in two dimensions and $A^2(\omega) = |\omega|^2$ in three dimensions, and $\mathbf{p}_s = (p_{sx}, p_{sy}, 0)$ is the ray parameter for the source plane waves at depth level \tilde{z}_s .

Substituting equations 4 and 3 into equation 5 yields

$$\hat{m}(\tilde{\mathbf{x}}, \mathbf{h}) = \sum_{\tilde{\mathbf{x}}'} \sum_{\mathbf{h}'} \Delta G(\tilde{\mathbf{x}}, \mathbf{h}, \tilde{\mathbf{x}}', \mathbf{h}') m(\tilde{\mathbf{x}}', \mathbf{h}'), \quad (7)$$

where

$$\begin{aligned} \Delta G(\tilde{\mathbf{x}}, \mathbf{h}, \tilde{\mathbf{x}}', \mathbf{h}') &= \sum_{\omega} \sum_{\tilde{\mathbf{x}}_s} \sum_{\tilde{\mathbf{x}}_r} G^*(\tilde{\mathbf{x}} - \mathbf{h}, \tilde{\mathbf{x}}_s, \omega) G_0(\tilde{\mathbf{x}}' - \mathbf{h}', \tilde{\mathbf{x}}_s, \omega) \\ &\times G^*(\tilde{\mathbf{x}} + \mathbf{h}, \tilde{\mathbf{x}}_r, \omega) G_0(\tilde{\mathbf{x}}' + \mathbf{h}', \tilde{\mathbf{x}}_r, \omega). \end{aligned} \quad (8)$$

When the same velocity model is used for Born modeling and migration ($\mathbf{v}_0 = \mathbf{v}$), $\Delta G(\tilde{\mathbf{x}}, \mathbf{h}, \tilde{\mathbf{x}}', \mathbf{h}')$ becomes the local Hessian operator or resolution function (Lecomte, 2008; Valenciano, 2008; Tang, 2009) under the new acquisition geometry. It has zero phase and is centered at $\tilde{\mathbf{x}}' = \tilde{\mathbf{x}}$ and $\mathbf{h}' = \mathbf{h}$. Therefore, $\hat{m}(\tilde{\mathbf{x}}, \mathbf{h})$ is a filtered version of the original image $m(\tilde{\mathbf{x}}, \mathbf{h})$. They have exactly the same kinematics. When the migration velocity is different from the modeling velocity ($\mathbf{v}_0 \neq \mathbf{v}$), the two images may substantially differ. Because we want to use migration results to estimate velocity, it is important to demonstrate that the velocity information contained in the prestack image obtained from the data modeled using the proposed procedure is consistent with the velocity information extracted from the prestack image obtained from migrating the originally recorded data set.

Throughout this paper, we perform numerical examples using Greens functions computed by means of one-way wavefield extrapolation (Claerbout, 1985; Stoffa et al., 1990; Ristow and Rühl, 1994; Biondi, 2002). Although not tested here, Greens functions obtained using other methods, such as by solving the two-way wave equation, can also be used under this framework. Since the one-way wavefield extrapolator has limited accuracy for large-angle propagation, we only compute horizontal half subsurface offset, i.e., $\mathbf{h} = (h_x, h_y, 0)$.

Figure 1(a) shows a modified Sigsbee2A velocity model, which contains two square anomalies below the salt: one 10% lower and the other 10% higher than the sediment velocity. We model 268 shots using the two-way wave-equation with a time-domain finite-difference scheme. The data is recorded with a marine acquisition geometry, the maximum offset for each shot is about 26000 ft. We refer this two-way data set as ‘‘original data’’ hereafter. The migrated prestack image and gathers using the original data and a starting velocity model (Figure 1(b)) are shown in Figure 2. The amplitude of the background image has been normalized by the diagonal of the phase-encoded Hessian (equation 3) to partially compensate for uneven illumination (Tang, 2009). Note the unfocused subsurface-offset-domain common-image gathers (SODCIGs) due to velocity errors. Then we use the target image (Figure 2) and the Born modeling described above to generate 41 plane-wave-source gathers from depth level $\tilde{z}_s = 15500$ ft, where the take-off angle is from -45° to 45° . The

same starting velocity model that was used for migration (Figure 1(b)) has been used for modeling, and the new data set is collected just above the target region, i.e., $\tilde{z}_r = \tilde{z}_s = 15500$ ft. We refer to the Born-modeled data as “new data” hereafter. The plane-wave migration result of the new data set using the starting velocity model is shown in Figure 4. Note the same kinematics shown in Figures 2 and 4. This suggests that the velocity information has been successfully preserved using the new data set, which is substantially smaller compared to the original one. Also note that Figure 4 is more blurry than Figure 2 due to the filtering effect of the resolution function ΔG (equation 8).

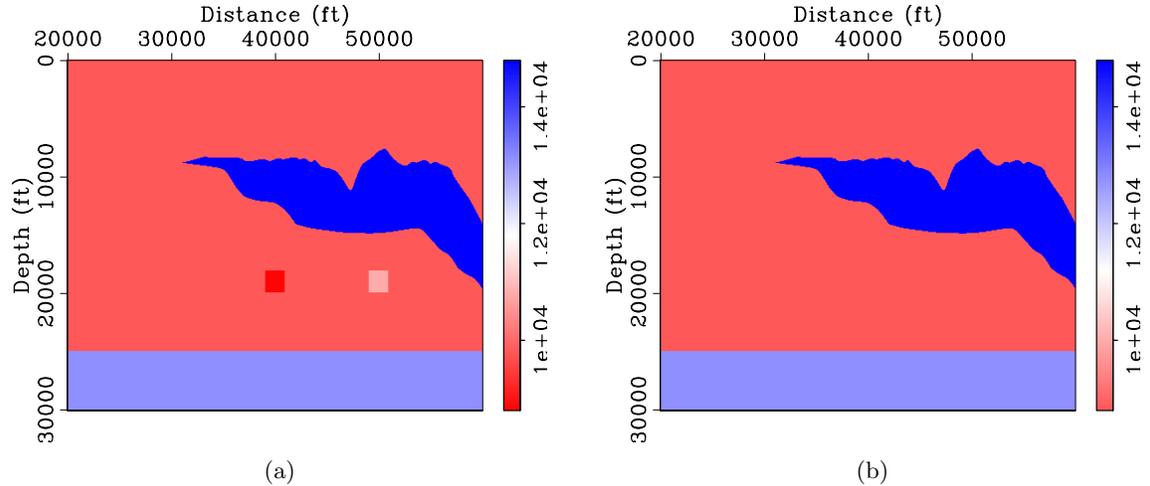


Figure 1: (a) The modified Sigsbee2A velocity model and (b) the starting velocity model. [\[ER\] yaxun1/. bwi-sigsb2c-vmod,bwi-sigsb2c-bvel](#)

TOMOGRAPHY USING DEMIGRATED DATA

The objective function of image-domain wavefield tomography can be defined as the ℓ_2 norm of either an image-domain residual field (Sava, 2004; Shen, 2004) or the negative image-stack power (or image coherence) (Toldi, 1985; Soubaras and Gratacos, 2007) or both (Shen and Symes, 2008). A widely used residual operator is the subsurface-offset-domain differential semblance optimization (DSO) operator, where the velocity model is optimized by penalizing energy at non-zero subsurface offset, utilizing the fact that the SODCIGs should be focused at zero subsurface offset if migrated using an accurate velocity model. The DSO objective function reads (Shen, 2004)

$$J_{\text{DSO}}(\mathbf{v}) = \frac{1}{2} \sum_{\tilde{\mathbf{x}}} \sum_{\mathbf{h}} |\mathbf{h}|^2 \hat{m}^2(\tilde{\mathbf{x}}, \mathbf{h}), \quad (9)$$

where $\hat{m}(\tilde{\mathbf{x}}, \mathbf{h})$ is the subsurface-offset-domain image migrated using velocity \mathbf{v} (equation 5) and the Born-modeled data described in the previous section. In contrast, minimizing the negative image-stack power (ISP) utilizes the fact that the stacked image should achieve maximum energy (or focus) when migrated with an accurate velocity (Toldi, 1985; Soubaras and Gratacos, 2007). Since the zero-subsurface-offset image is the stacked image, the ISP

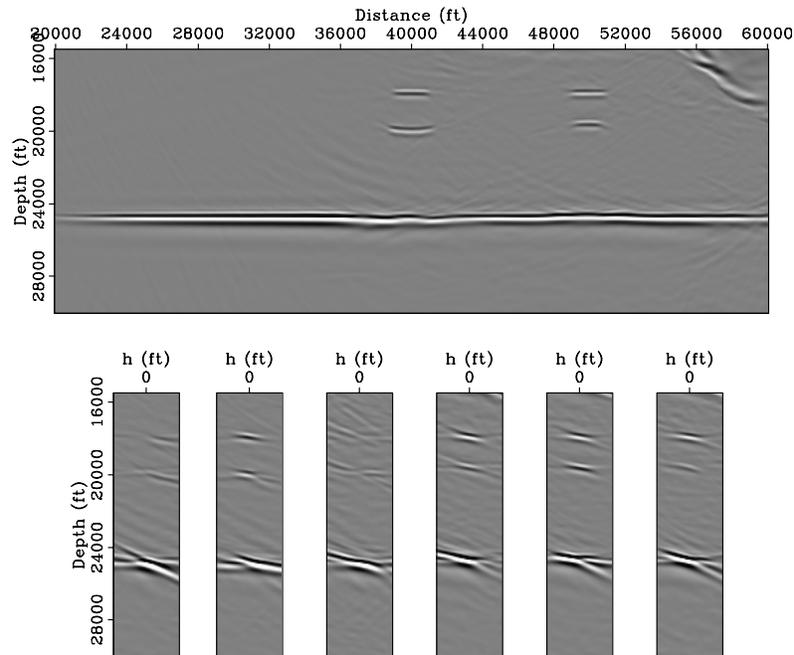


Figure 2: Initial image obtained using the original two-way finite-difference modeled data. The image has been normalized using the diagonal of the phase-encoded Hessian. Top: zero-subsurface offset image (stacked image), bottom: SODCIGs at surface location 38575,40000,41425,49300,50000 and 50700 ft, respectively. [CR] yaxun1/. bwi-sigsb2c-bimg-target-cpst

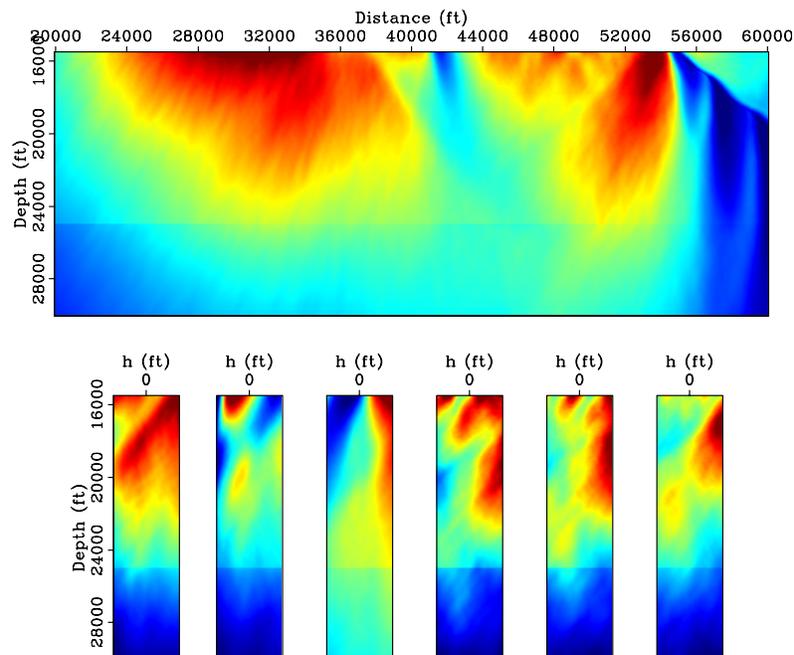


Figure 3: The diagonal of the phase-encoded Hessian obtained using the background velocity model. View descriptions are the same as in Figure 2. [CR] yaxun1/. bwi-sigsb2c-bhes-target

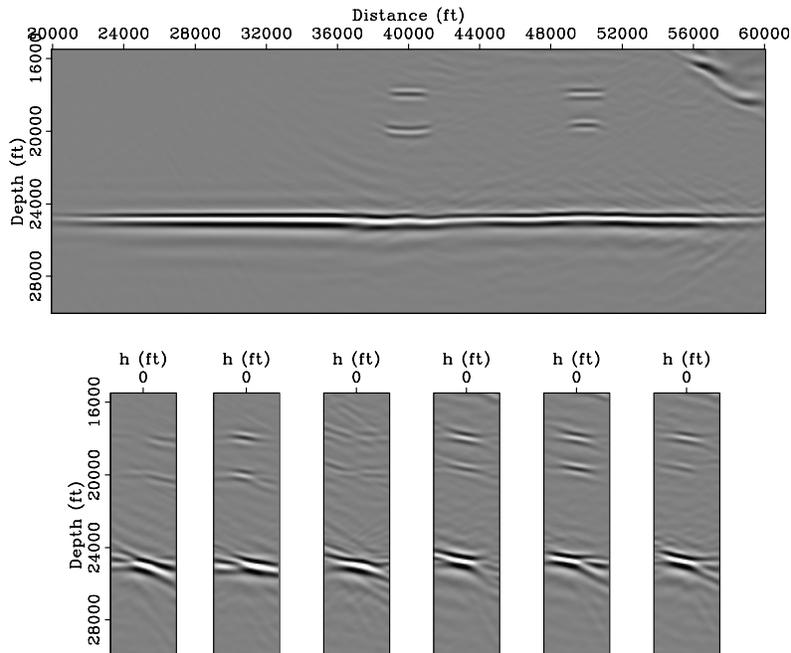


Figure 4: The migrated image and gathers using the new data set. View descriptions are the same as in Figure 2. [CR] [yaxun1/. bwi-sigsb2c-bimg-born-planes-target](#)

objective function reads

$$J_{\text{ISP}}(\mathbf{v}) = -\frac{1}{2} \sum_{\tilde{\mathbf{x}}} \hat{m}^2(\tilde{\mathbf{x}}, 0). \quad (10)$$

The above two objective functions can also be combined into one single objective function as follows (Shen and Symes, 2008):

$$J(\mathbf{v}) = J_{\text{DSO}} + \beta^2 J_{\text{ISP}}, \quad (11)$$

where β is a constant to trade off these two objective functions. The gradients can be calculated using the adjoint-state method without explicitly building the sensitivity matrix, see, e.g. Sava and Vlad (2008); Tang et al. (2008), for implementation details. The gradient is then used to update the velocity model with a suitable step length chosen by a line-search step. We iterate this process until an acceptable velocity model is obtained.

Figures 5 and 6 compare the normalized DSO and ISP gradients obtained using the original data set with those obtained using the new data set. These gradients are computed using the starting velocity model (Figure 1(b)). Note that both DSO and ISP gradients correctly identify those two anomalies and they also give correct directions for velocity updates. Also note that the gradients obtained using the new data set (Figures 5(b) and 6(b)) are very similar to those obtained using the original data set (Figures 5(a) and 6(a)), except that they are obtained with only 41 plane-wave-source gathers recorded at depth $\tilde{z}_r = 15500$ ft instead of 268 point-source gathers recorded at the surface ($z_r = 0$ ft). These examples demonstrate that the new data set can be used for wavefield tomography with much lower computational cost.

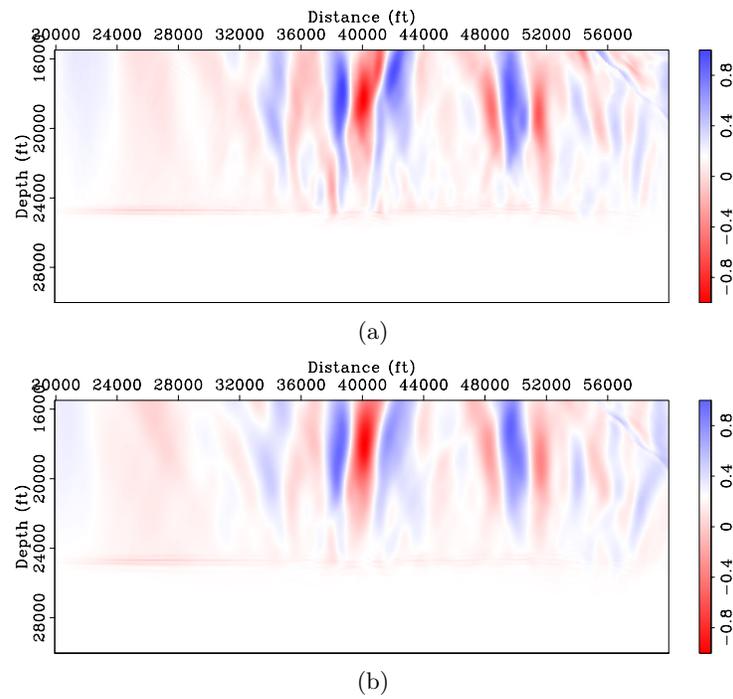


Figure 5: Normalized DSO gradient ∇J_{DSO} obtained using (a) the original data set and (b) the new data set. [CR]
 yaxun1/. bwi-sigsb2c-grad-offdso-target,bwi-sigsb2c-grad-offdso-born-planes-target

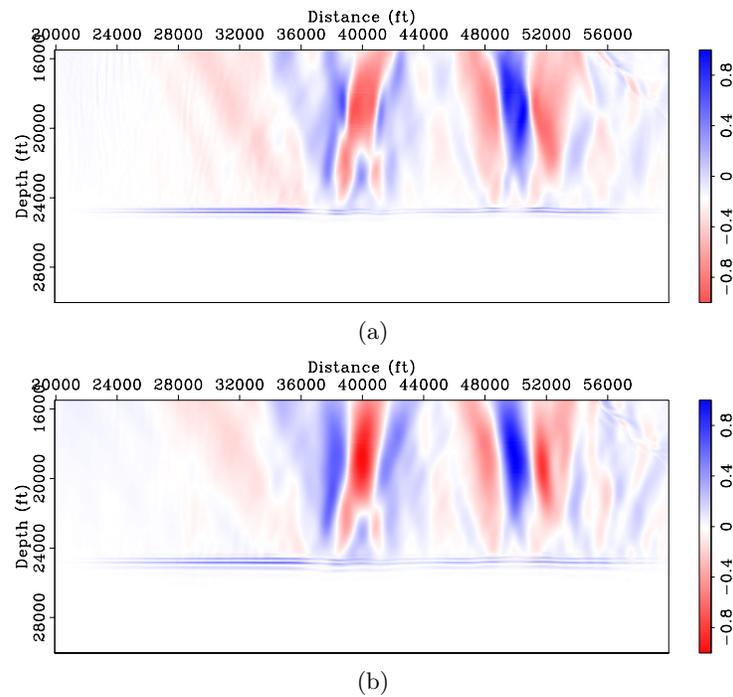


Figure 6: Normalized ISP gradient ∇J_{ISP} obtained using (a) the original data set and (b) the new data set. [CR]
 yaxun1/. bwi-sigsb2c-grad-imgpow-target,bwi-sigsb2c-grad-imgpow-born-planes-target

TOMOGRAPHIC INVERSION RESULTS

We have performed wavefield tomography using the new data set modeled from the selected target image as discussed in the previous section. We minimize the combined objective function J (equation 11) using a nonlinear conjugate-gradient solver. Figure 7(a) shows the inverted velocity model after 6 nonlinear iterations and Figure 7(b) shows the true velocity model. The nonlinear inversion has successfully recovered the two anomalies below the salt. The migrated image and gathers using the inverted velocity model and the original data set are shown in Figure 8 (compare it with the results obtained using the starting velocity model shown in Figure 2). The SODCIGs obtained using the inverted velocity model (Figure 8) are more focused at the zero subsurface offset and the horizontal reflector is flatter compared to those obtained using the starting velocity model (Figure 2). For comparison, Figure 10 shows the migration results obtained using the true velocity model.

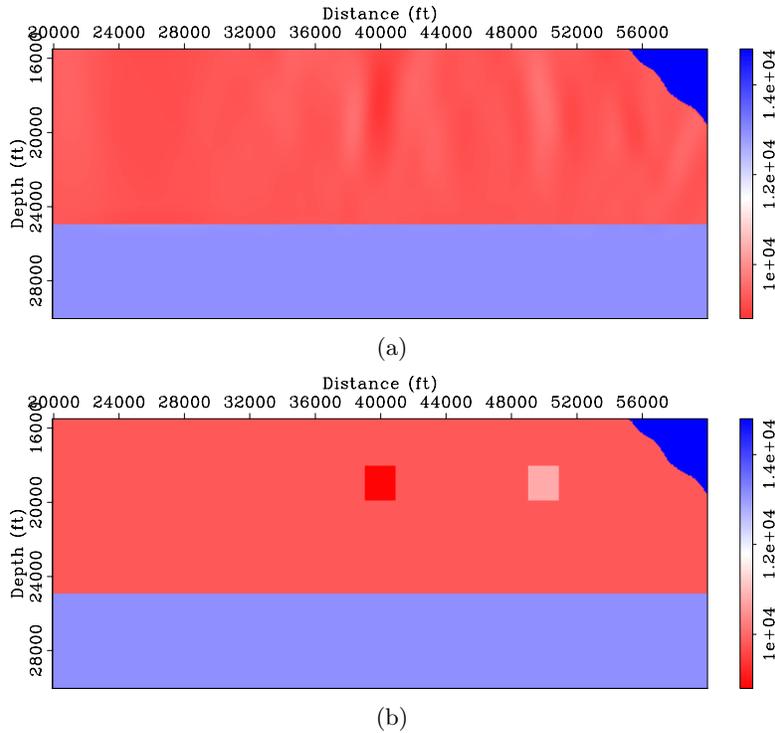


Figure 7: (a) The inverted velocity model and (b) the true velocity model. [CR] yaxun1/. bwi-sigsb2c-vmod-invt-target,bwi-sigsb2c-vmod-target

CONCLUSIONS

We have presented a cost-effective method for image-domain wavefield tomography. Instead of using the original data set for velocity estimation, our method uses demigrated Born data, which can be simulated in a target-oriented fashion and hence much smaller in size. Numerical examples demonstrate that the simulated new data set can successfully preserve velocity information that is useful for velocity analysis and can be used for velocity inversion with low computational cost.

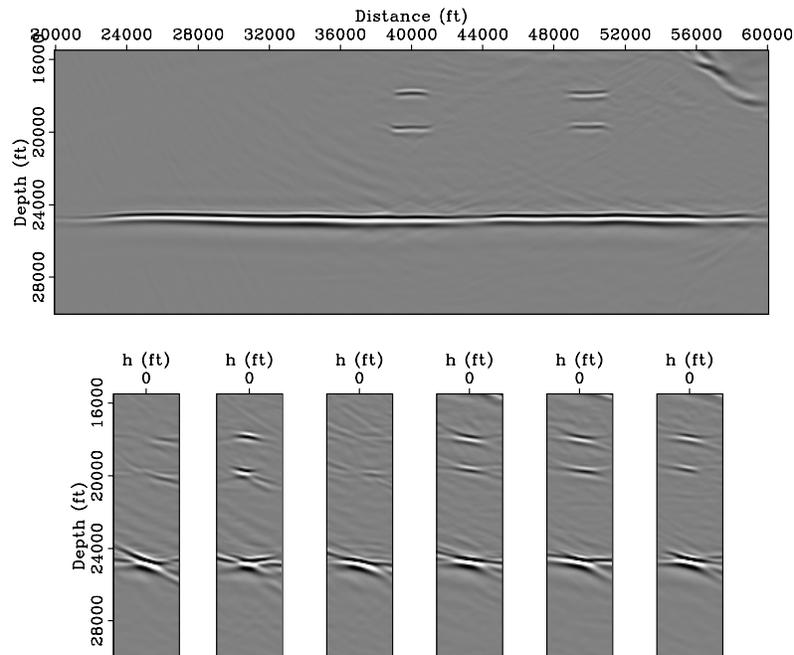


Figure 8: The migrated target image and gathers using the inverted velocity model (Figure 7(a)) and the original data. View descriptions are the same as in Figure 2. [CR] yaxun1/. bwi-sigsb2c-imag-Inv-target

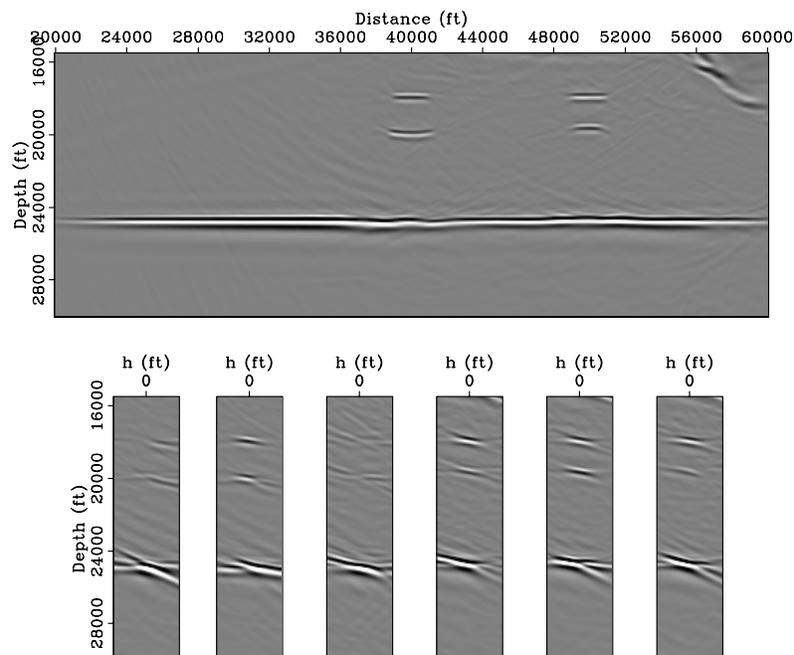


Figure 9: The migrated target image and gathers using the starting velocity model and the original data. View descriptions are the same as in Figure 2. [CR] yaxun1/. bwi-sigsb2c-bimg-target

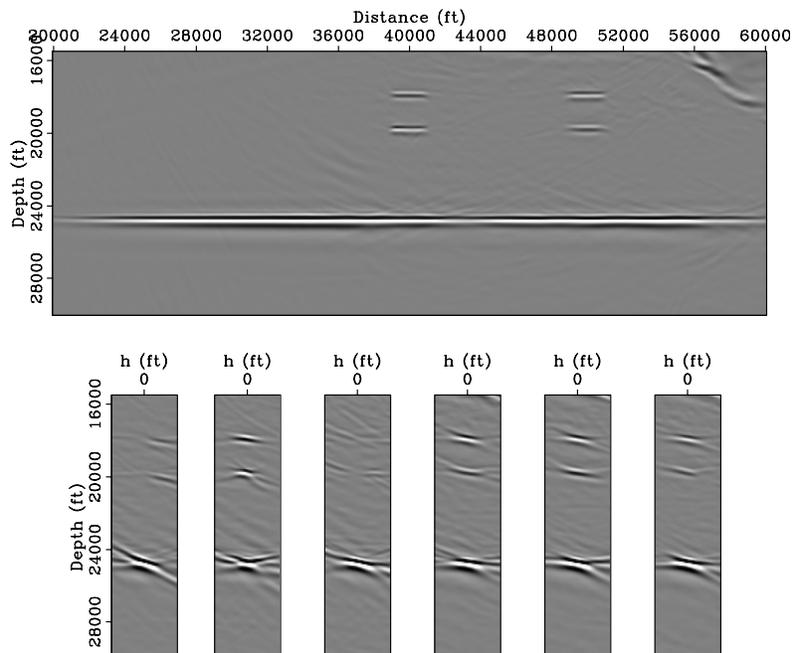


Figure 10: The migrated target image and gathers using the true velocity model and the original data. View descriptions are the same as in Figure 2. [CR] `yaxun1/. bwi-sigsb2c-imag-target`

REFERENCES

- Albertin, U., P. Sava, J. Etgen, and M. Maharramov, 2006, Adjoint wave-equation velocity analysis: SEG Technical Program Expanded Abstracts, **25**, 3345–3349.
- Biondi, B., 2002, Stable wide-angle Fourier finite-difference downward extrapolation of 3-D wavefields: *Geophysics*, **67**, 872–882.
- , 2006, Prestack exploding-reflectors modeling for migration velocity analysis: SEG Technical Program Expanded Abstracts, **25**, 3056–3060.
- , 2007, Prestack modeling of image events for migration velocity analysis: **SEP-131**, 101–118.
- Biondi, B., and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.
- Claerbout, J. F., 1985, *Imaging the earth’s interior*: Blackwell Scientific Publication.
- Guerra, C., and B. Biondi, 2008a, Phase-encoding with Gold codes for wave-equation migration: **SEP-136**, 23–42.
- , 2008b, Prestack exploding reflector modeling: The crosstalk problem: **SEP-134**, 79–92.
- Guerra, C., Y. Tang, and B. Biondi, 2009, Wave-equation tomography using image-space phase encoded data: SEG Technical Program Expanded Abstracts, **28**, 3964–3968.
- Hoffmann, J., 2001, Illumination, resolution, and image quality of PP- and PS-waves for survey planning: *The Leading Edge*, **20**, 1008–1014.
- Lecomte, I., 2008, Resolution and illumination analyses in PSDM: A ray-based approach: *The Leading Edge*, **27**, no. 5, 650–663.
- Liu, F., D. W. Hanson, N. D. Whitmore, R. S. Day, and R. H. Stolt, 2006, Toward a unified

- analysis for source plane-wave migration: *Geophysics*, **71**, S129–S139.
- Mora, P., 1989, Inversion = migration + tomography: *Geophysics*, **54**, 1575–1586.
- Plessix, R.-E., and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: *Geophys. J. Int.*, **157**, 975–987.
- 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.
- Rickett, J. E., 2003, Illumination-based normalization for wave-equation depth migration: *Geophysics*, **68**, 1371–1379.
- Ristow, D., and T. Rühl, 1994, Fourier finite-difference migration: *Geophysics*, **59**, 1882–1893.
- Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: *Geophysics*, **65**, 426–436.
- Sava, P., 2004, Migration and Velocity Analysis by Wavefield Extrapolation: PhD thesis, Stanford University.
- Sava, P., and I. Vlad, 2008, Numeric implementation of wave-equation migration velocity analysis operators: *Geophysics*, **73**, VE145–VE159.
- Shen, P., 2004, Wave-equation Migration Velocity Analysis by Differential Semblance Optimization: PhD thesis, Rice University.
- Shen, P., and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: *Geophysics*, **73**, VE49–VE59.
- Shen, P., W. W. Symes, S. Morton, A. Hess, and H. Calandra, 2005, Differential semblance velocity analysis via shot profile migration: *SEG Technical Program Expanded Abstracts*, **24**, 2249–2252.
- Soubaras, R., and B. Gratacos, 2007, Velocity model building by semblance maximization of modulated-shot gathers: *Geophysics*, **72**, U67–U73.
- Stoffa, P. L., J. T. Fokkema, R. M. de Luna Freire, and W. P. Kessinger, 1990, Split-step fourier migration: *Geophysics*, **55**, 410–421.
- Stolt, R. H., and A. Benson, 1986, *Seismic Migration: Theory and Practice*: Geophysical Press.
- Symes, W. W., 2008, Approximate linearized inversion by optimal scaling of prestack depth migration: *Geophysics*, **73**, R23–R35.
- Tang, Y., 2008, Modeling, migration and inversion in the generalized source and receiver domain: **SEP-136**, 97–112.
- , 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian: *Geophysics*, **74**, WCA95–WCA107.
- Tang, Y., C. Guerra, and B. Biondi, 2008, Image-space wave-equation tomography in the generalized source domain: **SEP-136**, 1–22.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266.
- Toldi, J., 1985, Velocity analysis without picking: PhD thesis, Stanford University.
- Valenciano, A., 2008, Imaging by Wave-equation Inversion: PhD thesis, Stanford University.
- Whitmore, N. D., 1995, An Imaging Hierarchy for Common Angle Plane Wave Seismogram: PhD thesis, University of Tulsa.
- Woodward, M. J., 1992, Wave-equation tomography: *Geophysics*, **57**, 15–26.
- Zhang, Y., J. Sun, C. Notfors, S. Grey, L. Chemis, and J. Young, 2005, Delayed-shot 3D depth migration: *Geophysics*, **70**, E21–E28.

Wave-equation tomography by beam focusing

Biondo Biondi

ABSTRACT

Velocity can be estimated using a wave-equation operator by maximizing an objective function that measures the flatness of the crosscorrelation computed between a source wavefield and a receiver wavefield. The proposed objective function depends on the parameters of a residual moveout applied to the computed correlation. It is composed of two terms: the first term maximizes the energy of the stack computed on local subarrays as a function of the local curvature. The second term maximizes the power of the stack computed globally as a function of time shifts applied to the stacks of the local subarrays. The first term is essential to assure global convergence in presence of large velocity errors. The second term plays a role in estimating localized velocity anomalies. Numerical examples of computation of the gradients of the proposed objective function confirm its potential for velocity estimation.

INTRODUCTION

Tomographic velocity estimation based on wave-equation operators can improve seismic imaging in areas where wavefield-continuation migration is needed. However, it is well-known that the straightforward application of waveform inversion to estimate migration velocity fails to converge to an accurate model when the starting model is too far from the correct one. This failure to converge is caused by the non-linear relationship between data amplitudes and velocity. To avoid this failure the velocity-estimation problem can be formulated in the image domain as the maximization (or minimization) of objective functions that are more sensitive to the data kinematics than to the data amplitudes. Two important examples of this approach are the Wave-Equation Migration Velocity Analysis (WEMVA) method (Biondi and Sava, 1999; Sava and Biondi, 2004a,b; Sava, 2004) and the Differential Semblance Optimization (DSO) (Symes and Carazzone, 1991; Shen, 2004; Shen et al., 2005). Luo and Schuster (1991) introduced a method based on a kinematic objective function to solve the problem of transmission tomography. Both their method and the WEMVA method suffer from the drawback that they require the picking of kinematic parameters: correlation lag in one case (Luo and Schuster, 1991), and a residual migration parameter for WEMVA.

In this paper I develop a framework to update migration velocity by maximizing an objective function defined in the image domain. The objective function is defined as a function of moveout parameters but velocity updating can be performed without explicit picking of the residual moveout parameters. Therefore, it overcomes one of the main difficulties of the WEMVA methods. The methodology is general and can be thus used to optimize the image as a function of arbitrary residual moveout parameters, and possibly of residual migration parameters.

I also introduce a new objective function that overcomes limitations of known methods. This novel objective function has two components: the first term measures the power of the stack over local subarrays (beams) as a function of the moveout curvature. The second term measures the power of the stack across the beams as a function of a bulk-shift of each beam. I then apply the general theory to the computation of the gradient of the proposed objective function with respect to velocity perturbations.

I develop the theory and show the results of numerical tests for a transmission tomography problem because transmission tomography is simpler than reflection tomography, and therefore better suited to the illustration of the basic concepts. I leave to future reports the application to reflection tomography of the method developed for transmission tomography in this report.

THEORY

In this section I develop the general theory for a transmission tomography problem because it is simpler than reflection tomography. In transmission tomography, the data are recorded after only one propagation path through the medium, as opposed to the downgoing and upgoing paths of a typical reflection tomography problems. Furthermore, in transmission tomography there is no need to image and locate reflectors in depth, which is a major hurdle in reflection tomography. However, the application to reflection tomography of the theory presented in this paper should be fairly straightforward. I propose to solve the transmission tomography problem by maximizing an objective function based on the correlation between recorded data and modeled data. This correlation is analogous to the correlation between source and receiver wavefields required by migration imaging condition.

To further simplify the theoretical development, I define an objective function that rewards consistency of the correlation computed independently for each source location. The objective function measures correlation consistency along the receiver axis. However, I use here the receiver axis as a proxy for the offset axis or the aperture-angle axis in reflection tomography. The application of the concepts developed in this paper to objective functions useful in reflection tomography should be straightforward, although it will require more complex notation and result in expressions for the gradients even more complex than the ones presented here.

I define the recorded data as $P_D(t, x_g, x_s)$, and the modeled data as $\tilde{P}(t, x_g, x_s; s)$, where t is the recording time, x_g is the receiver coordinate, x_s is the source coordinate, and $s(z, x)$ is the slowness model defined in depth z and along the horizontal coordinate x .

The cross-correlation $C(\tau)$ between the recorded data and the modeled data is defined as a function of the correlation time lag τ as

$$C(\tau) \left[\tilde{P}(t), P_D(t) \right] = \sum_t \tilde{P}(t - \tau) P_D(t). \quad (1)$$

I introduce an objective function that maximizes the flatness of the correlation function along the receiver axis for all values of the lag τ . In particular, I aim to maximize local correlation flatness after subdividing the receiver array into local subarrays. To extract the correlation for each subarray centered at \bar{x}_g , I apply a local beam-decomposition operators

$\mathbf{B}_{\bar{x}}$. Within in each subarray, traces are defined by the local offset Δx_g . The dimensions of each $\mathbf{B}_{\bar{x}}$ are thus $(N_{\Delta x_g} N_\tau \times N_{x_g} N_\tau)$.

Given a background slowness s_0 we can compute the correlation in equation 1. In each subarray, the correlation can be flattened by the application of $N_{\bar{x}_g}$ moveout operators $\mathcal{M}_{\bar{x}}$; that is

$$C(\tau + \theta(\boldsymbol{\mu}_{\bar{x}})) \left[\tilde{P}(t; s_0), P_D(t) \right] = \mathcal{M}_{\bar{x}}[\theta(\boldsymbol{\mu}_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s_0)], \quad (2)$$

where $\boldsymbol{\mu}_{\bar{x}}$ are the moveout parameters and θ are the corresponding time shifts. I further define the local stacking operator $\mathbf{S}_{\bar{x}}$ that sums the correlation traces along the local offset axis Δx_g .

I can now introduce the first, and local, term of the objective function that measures the flatness of the correlation within each subarray as:

$$J_{\text{Local}}(\boldsymbol{\mu}_{\bar{x}}(s)) = \frac{1}{2} \sum_{x_s} \sum_{\bar{x}_g} \|\mathbf{S}_{\bar{x}} \mathcal{M}_{\bar{x}}[\theta(\boldsymbol{\mu}_{\bar{x}}(s)), \mathbf{B}_{\bar{x}}C(\tau; s_0)]\|_2^2. \quad (3)$$

This objective function is not a direct function of the slowness s , but it depends indirectly from it through the moveout parameters $\boldsymbol{\mu}_{\bar{x}}$. These parameters are the solutions of $N_{x_s} \times N_{\bar{x}_g}$ independent fitting problems, one for each subarray and source location. These *auxiliary* objective functions measure the zero lag of the cross-correlation between the correlation computed for a realization of the slowness function s and and the moved-out correlation computed with the background slowness s_0 ,

$$\begin{aligned} J_{\text{FL}}(\boldsymbol{\mu}_{\bar{x}}) &= C(\tau = 0) [\mathcal{M}_{\bar{x}}[\theta(\boldsymbol{\mu}_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s_0)], \mathbf{B}_{\bar{x}}C(\tau; s)] \\ &= \langle \mathcal{M}_{\bar{x}}[\theta(\boldsymbol{\mu}_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s_0)], \mathbf{B}_{\bar{x}}C(\tau; s) \rangle, \end{aligned} \quad (4)$$

where with the notation $\langle \mathbf{x}, \mathbf{y} \rangle$ I indicate the inner product of the vectors \mathbf{x} and \mathbf{y} . This inner product spans the time-lag axis τ and the local offset axis Δx_g . The local moveout parameters are the solutions of the following maximization problem:

$$\max_{\boldsymbol{\mu}_{\bar{x}}} J_{\text{FL}}(\boldsymbol{\mu}_{\bar{x}}). \quad (5)$$

For velocity estimation, the most effective parametrization of the moveout within each beam is the curvature μ_C , that defines the following moveout equation

$$\theta(\boldsymbol{\mu}_{\bar{x}}) = \mu_C \Delta x_g^2. \quad (6)$$

Notice that when the slowness is equal to the background slowness s_0 , the corresponding best-fitting moveout parameters $\bar{\boldsymbol{\mu}}_{\bar{x}}$ are obviously the ones corresponding to no moveout; that is, $\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}) = 0$.

As the numerical examples I show in the next section demonstrate, the beam curvature is effective to capture the long-wavelength perturbations in the velocity model, but is less effective to capture the short-wavelength perturbations. Accordingly, a wave-equation tomography based solely on the objective function 3 may have difficulties to estimate short-wavelength velocity perturbations.

To address this shortcoming I introduce a second, and global, term to the objective function. This term measures flatness across the subarrays, after the local moveouts have

been applied, and is defined as,

$$J_{\text{Global}}(\boldsymbol{\mu}(s)) = \frac{1}{2} \sum_{x_s} \|\mathbf{S}\mathcal{M}\{\theta(\boldsymbol{\mu}(s)), \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathcal{M}_{\bar{x}}[\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s_0)]\}\|_2^2, \quad (7)$$

where $\boldsymbol{\Sigma}_{\bar{x}}$ assembles all the results of the stacking over the subarrays into a global array, \mathbf{S} is a global stacking operator, and \mathcal{M} is a global moveout operator function of the vector of parameter $\boldsymbol{\mu}$. They both operate on the result of the local stacking of the subarrays. As in the previous case, the moveout parameters are solutions of N_{x_s} independent fitting problems, one for each source location. Similarly, these auxiliary objective functions measure the zero lag of the cross-correlation between the local stack of the correlation computed using the current slowness function and the local stack of the moved-out correlation computed using the background slowness; that is,

$$J_{\text{FG}}(\boldsymbol{\mu}) = \langle \mathcal{M}\{\theta(\boldsymbol{\mu}), \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathcal{M}_{\bar{x}}[\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s_0)]\}, \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathcal{M}_{\bar{x}}[\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s)] \rangle. \quad (8)$$

In this case the inner product spans only the time-lag axis τ .

The global moveout parameters are the solutions of the following N_{x_s} maximization problems

$$\max_{\boldsymbol{\mu}} J_{\text{FG}}(\boldsymbol{\mu}). \quad (9)$$

I chose to parametrize the global moveout as simple time shifts for each beam center \bar{x} that is, the moveout equation is

$$\theta(\boldsymbol{\mu}) = \mu_{\theta}. \quad (10)$$

Notice that with this choice of moveout parameters each maximization problem in 9 is an ensemble of $N_{\bar{x}_g}$ independent problems. This consideration becomes important when computing the gradient of the objective function.

Combining the objective function in 3 and in 7 we define the maximization problem that we solve to estimate slowness:

$$\max_s [J_{\text{Local}}(\boldsymbol{\mu}_{\bar{x}}(s)) + \epsilon J_{\text{Global}}(\boldsymbol{\mu}(s))], \quad (11)$$

where the parameter ϵ can be tuned to find an optimal relative scaling between the local and global components, although in principle $\epsilon = 1$ should be effective.

Gradient of the objective function

I plan to solve the optimization problem defined in 11 by a gradient-based optimization algorithm. Therefore, the development of an algorithm to efficiently compute the gradient of the objective function with respect to slowness is an essential step to make the method practical. In this section I introduce the basic methodology to compute the gradients, and I leave some of the details to Appendix A.

The gradient of both the local objective function 3 and the global one 7 are computed using the chain rule. The first terms of the chains are the derivatives of the objective function with respect the moveout parameters. The second terms are the derivatives of the moveout parameters with respect to slowness; they are computed from the fitting objective functions 4 and 8.

Derivatives with respect to moveout parameters

The computation of the derivatives of 3 with respect to each vector of local-moveout parameters is easily evaluated using the following expression:

$$\frac{\partial J_{\text{Local}}}{\partial \boldsymbol{\mu}_{\bar{x}}} = \frac{\partial \mathcal{M}_{\bar{x}}'}{\partial \boldsymbol{\mu}_{\bar{x}}} \mathbf{S}'_{\bar{x}} \mathbf{S}_{\bar{x}} \mathcal{M}_{\bar{x}} [\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} C (\tau; s_0)]. \quad (12)$$

The linear operator $\frac{\partial \mathcal{M}_{\bar{x}}}{\partial \boldsymbol{\mu}_{\bar{x}}}$ has the dimensions $(N_{\Delta x_g} N_{\tau} \times N_{\boldsymbol{\mu}_{\bar{x}}})$ and is given by

$$\frac{\partial \mathcal{M}_{\bar{x}}}{\partial \boldsymbol{\mu}_{\bar{x}}} = \mathcal{M}_{\bar{x}} \left[\theta (\boldsymbol{\mu}_{\bar{x}}), \mathbf{B}_{\bar{x}} \dot{C} (\tau; s_0) \right] \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\bar{x}}}, \quad (13)$$

where $\dot{C} (\tau; s_0) = C (\tau) \left[\tilde{P} (t; s_0), \dot{P}_D (t) \right]$, with \dot{P}_D being the time derivative of the recorded-data traces. For the choice of moveout parameters expressed in equation 6 we have $\partial \theta / \partial \boldsymbol{\mu}_{\bar{x}} = \partial \theta / \partial \mu_C = \Delta x_g^2$.

Similarly, the evaluation of the derivatives of 7 with respect to each shift parameter $\boldsymbol{\mu}$ is easily carried out by the following:

$$\frac{\partial J_{\text{Global}}}{\partial \boldsymbol{\mu}} = \frac{\partial \mathcal{M}'}{\partial \boldsymbol{\mu}} \mathbf{S}' \mathbf{S} \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathcal{M}_{\bar{x}} [\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} C (\tau; s_0)], \quad (14)$$

where the linear operator $\frac{\partial \mathcal{M}}{\partial \boldsymbol{\mu}}$ is given by

$$\frac{\partial \mathcal{M}}{\partial \boldsymbol{\mu}} = \mathcal{M} \left\{ \theta (\boldsymbol{\mu}), \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathcal{M}_{\bar{x}} \left[\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} \dot{C} (\tau; s_0) \right] \right\} \frac{\partial \theta}{\partial \boldsymbol{\mu}}. \quad (15)$$

When the moveout parameters are simple trace-by-trace phase shifts, as defined in equation 10, it results that $\partial \theta / \partial \boldsymbol{\mu} = 1$.

On a practical note, the preceding expressions look more daunting than they are in practice. They greatly simplify in the important case when the gradient is evaluated for $\bar{\boldsymbol{\mu}}_{\bar{x}} = 0$ and $\bar{\boldsymbol{\mu}} = 0$. This simplifying condition is actually always fulfilled unless the optimization algorithm includes inner iterations for fitting the moveout parameters using a linearized approach. Under these conditions, equations 12 and 13 become, respectively,

$$\left. \frac{\partial J_{\text{Local}}}{\partial \boldsymbol{\mu}_{\bar{x}}} \right|_{\boldsymbol{\mu}_{\bar{x}}=0} = \frac{\partial \mathcal{M}_{\bar{x}}'}{\partial \boldsymbol{\mu}_{\bar{x}}} \mathbf{S}'_{\bar{x}} \mathbf{S}_{\bar{x}} \mathbf{B}_{\bar{x}} C (\tau; s_0), \quad (16)$$

and

$$\frac{\partial \mathcal{M}_{\bar{x}}}{\partial \boldsymbol{\mu}_{\bar{x}}} = \mathbf{B}_{\bar{x}} \dot{C} (\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\bar{x}}}. \quad (17)$$

Similarly, equations 14 and 15 become

$$\left. \frac{\partial J_{\text{Global}}}{\partial \boldsymbol{\mu}} \right|_{\boldsymbol{\mu}_{\bar{x}}=0, \boldsymbol{\mu}=0} = \frac{\partial \mathcal{M}'}{\partial \boldsymbol{\mu}} \mathbf{S}' \mathbf{S} \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathbf{B}_{\bar{x}} C (\tau; s_0), \quad (18)$$

and

$$\frac{\partial \mathcal{M}}{\partial \boldsymbol{\mu}} = \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathbf{B}_{\bar{x}} \dot{C} (\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}}. \quad (19)$$

Derivatives with respect to slowness

The evaluation of the derivatives of the moveout parameters with respect to slowness follows a slightly different procedure from the one above because the moveout parameters are solutions of the optimization problems 5 and 9. We take advantage of the fact that we need to evaluate the derivatives only at the solution points, where the objective functions are maximized and thus their derivatives with respect to the moveout parameters are zero. We can therefore write:

$$\left. \frac{\partial J_{\text{FL}}(\boldsymbol{\mu}_{\bar{x}})}{\partial \boldsymbol{\mu}_{\bar{x}}} \right|_{\boldsymbol{\mu}_{\bar{x}}=\bar{\boldsymbol{\mu}}_{\bar{x}}} = J_{\text{FL}}(\bar{\boldsymbol{\mu}}_{\bar{x}}) = 0 = \left\langle \frac{\partial \mathcal{M}_{\bar{x}}[\theta(\boldsymbol{\mu}_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s_0)]}{\partial \boldsymbol{\mu}_{\bar{x}}} \right|_{\boldsymbol{\mu}_{\bar{x}}=\bar{\boldsymbol{\mu}}_{\bar{x}}}, \mathbf{B}_{\bar{x}}C(\tau; s) \rangle,$$

and

$$\left. \frac{\partial J_{\text{FG}}(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}} \right|_{\boldsymbol{\mu}=\bar{\boldsymbol{\mu}}} = J_{\text{FG}}(\bar{\boldsymbol{\mu}}) = 0 = \left\langle \frac{\partial \mathcal{M}\{\theta(\boldsymbol{\mu}), \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathcal{M}_{\bar{x}}[\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s_0)]\}}{\partial \boldsymbol{\mu}} \right|_{\boldsymbol{\mu}=\bar{\boldsymbol{\mu}}}, \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathcal{M}_{\bar{x}}[\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s)] \rangle.$$

Using the rule for differentiating implicit functions, and taking advantage that the fitting problems are all independent from each other (i.e. the cross derivatives with respect to the moveout parameters are all zero), we can formally write:

$$\left. \frac{\partial \boldsymbol{\mu}_{\bar{x}}}{\partial s} \right|_{\boldsymbol{\mu}_{\bar{x}}=\bar{\boldsymbol{\mu}}_{\bar{x}}} = - \frac{\frac{\partial J_{\text{FL}}(\boldsymbol{\mu}_{\bar{x}})}{\partial s}}{\frac{\partial J_{\text{FL}}(\boldsymbol{\mu}_{\bar{x}})}{\partial \boldsymbol{\mu}_{\bar{x}}}}, \quad (20)$$

and

$$\left. \frac{\partial \boldsymbol{\mu}}{\partial s} \right|_{\boldsymbol{\mu}=\bar{\boldsymbol{\mu}}} = - \frac{\frac{\partial J_{\text{FG}}(\boldsymbol{\mu})}{\partial s}}{\frac{\partial J_{\text{FG}}(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}}. \quad (21)$$

Appendix A presents the analytical development of these expressions to compute the derivatives of the moveout parameters with respect to slowness. As for the derivatives of the main objective function with respect to moveout parameters, the final results for the special case of $\bar{\boldsymbol{\mu}}_{\bar{x}} = 0$ and $\bar{\boldsymbol{\mu}} = 0$ have a fairly simple analytical expression. The derivative of the local moveout parameters are (A-1):

$$\left. \frac{\partial \boldsymbol{\mu}_{\bar{x}}}{\partial s} \right|_{\boldsymbol{\mu}_{\bar{x}}=0} = - \frac{\left\langle \mathbf{B}_{\bar{x}} \dot{C}(\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\bar{x}}}, \mathbf{B}_{\bar{x}} \mathbf{P}_{\mathbf{D}} \frac{\partial \bar{P}}{\partial s} \right\rangle}{\left\langle \mathbf{B}_{\bar{x}} \ddot{C}(\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\bar{x}}}, \mathbf{B}_{\bar{x}} C(\tau; s) \right\rangle}, \quad (22)$$

and the derivative of the global moveout parameters are (A-2):

$$\left. \frac{\partial \boldsymbol{\mu}}{\partial s} \right|_{\boldsymbol{\mu}_{\bar{x}}=0, \boldsymbol{\mu}=0} = - \frac{\left\langle \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathbf{B}_{\bar{x}} \dot{C}(\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}}, \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathbf{B}_{\bar{x}} \mathbf{P}_{\mathbf{D}} \frac{\partial \bar{P}}{\partial s} \right\rangle}{\left\langle \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathbf{B}_{\bar{x}} \ddot{C}(\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}}, \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathbf{B}_{\bar{x}} C(\tau; s) \right\rangle}, \quad (23)$$

in which $\ddot{C}(\tau) = C(\tau) [\tilde{P}(t), \ddot{P}_D(t)]$. In both equations 22 and 23 the operator \mathbf{P}_D represents a convolution with the recorded data, whereas the operator $\partial\tilde{P}/\partial s$ is the basic wave-equation tomography operator that links perturbations in the slowness model to perturbations in the modeled data.

Combining the derivatives in equation 22 with the derivatives in equations 16-17 we can compute the gradient of the local objective function 3 with respect to slowness as:

$$\nabla J_{\text{Local}} = - \underbrace{\frac{\partial\tilde{P}'}{\partial s} \mathbf{P}_D' \mathbf{B}'_{\bar{x}}}_{\text{I}} \underbrace{\mathbf{B}_{\bar{x}} \dot{C}(\tau; s_0)}_{\text{II}} \frac{\partial\theta}{\partial\mu_{\bar{x}}} \underbrace{\left\langle \mathbf{B}_{\bar{x}} \ddot{C}(\tau; s_0) \frac{\partial\theta}{\partial\mu_{\bar{x}}}, \mathbf{B}_{\bar{x}} C(\tau; s) \right\rangle}_{\text{III}} \frac{\partial J_{\text{Local}}}{\mu_{\bar{x}}}. \quad (24)$$

I will now examine the effects of each of the terms in equation 24 starting from the right-most one. The third term (III) produces a scalar for each local curvature parameter μ_C . This scalar multiplies the traces in each beam, after they have been differentiated in time and scaled by $\partial\theta/\partial\mu_{\bar{x}}$, as described by the second term (II). Notice that the phase introduced by the time derivative of the correlation function in (II) is crucial for the successful backprojection into the slowness model that is accomplished by the first term (I). In this term, first $\mathbf{B}'_{\bar{x}}$ projects the traces of each individual beam into the space of the global array, then the convolution with the recorded data \mathbf{P}_D' time shifts the correlation function by the time delay of the events. Finally, the adjoint of the operator $\partial\tilde{P}/\partial s$ backprojects the perturbation in the wavefields at the receiver array into the slowness model.

The expression of the gradient of the global objective function 7 with respect to slowness is similarly derived by combining the derivatives in equation 23 with the derivatives in equations 18-19 and is the following three-terms expression:

$$\nabla J_{\text{Global}} = - \underbrace{\frac{\partial\tilde{P}'}{\partial s} \mathbf{P}_D' \mathbf{B}'_{\bar{x}} \mathbf{S}'_{\bar{x}} \Sigma'_{\bar{x}}}_{\text{I}} \underbrace{\Sigma_{\bar{x}} \mathbf{S}_{\bar{x}} \mathbf{B}_{\bar{x}}}_{\text{II}} \dot{C}(\tau; s_0) \frac{\partial\theta}{\partial\mu} \underbrace{\left\langle \Sigma_{\bar{x}} \mathbf{S}_{\bar{x}} \mathbf{B}_{\bar{x}} \ddot{C}(\tau; s_0) \frac{\partial\theta}{\partial\mu}, \Sigma_{\bar{x}} \mathbf{S}_{\bar{x}} \mathbf{B}_{\bar{x}} C(\tau; s) \right\rangle}_{\text{III}} \frac{\partial J_{\text{Global}}}{\mu}. \quad (25)$$

The structure of equation 25 is similar to the structure of equation 24 and the terms have similar explanations. The only important difference is that in term II the chain $\Sigma_{\bar{x}} \mathbf{S}_{\bar{x}}$ performs the stack over the local arrays and the assemblage of the stacked traces into the global array, whereas its adjoint in term I spreads the stacked traces back into the local arrays reforming the local beams.

NUMERICAL COMPUTATION OF SEARCH DIRECTIONS

To test the method presented in the previous section, I computed the gradient of the objective functions for two synthetic data sets. The first assumes a uniform slowness error,

Figure 1: Search direction computed using conventional full-waveform inversion with a positive small (0.1%) uniform slowness error: a) slowness perturbations and b) slowness perturbations averaged over the horizontal direction. [CR]

biondo1/. Avg-Dir-tiny-neg

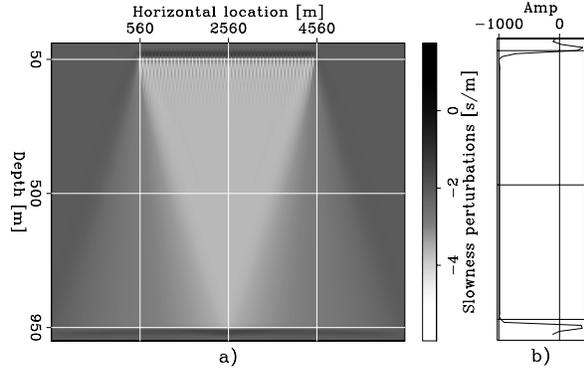
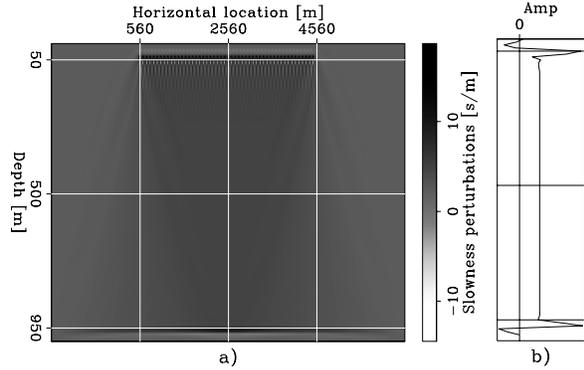


Figure 2: Search direction computed using conventional full-waveform inversion with a positive large (8.0%) uniform slowness error: a) slowness perturbations and b) slowness perturbations averaged over the horizontal direction. [CR]

biondo1/. Avg-Dir-large-neg



whereas the second assumes a localized velocity error. In both cases the correct velocity was constant and equal to 1 km/s. Both sets of experiments were computed assuming 50 shots at the surface (actually at 50 meters depth) spaced 80 meters apart, and the receiver array at a depth of 950 meters. The data modeled with a uniform slowness error were recorded by a split-spread array of 512 receivers spaced 10 meters apart, whereas the data modeled with a localized velocity error were recorded by a split-spread array of 720 receivers spaced 10 meters apart. In both cases, the receiver array was moving along with the shots, to maintain a uniform offset coverage for each shot. For the sake of clarity, for all the cases the figures display the search directions instead of the gradients. Also notice that I am showing separately the search directions computed for each term of the proposed objective function, without showing their sum. At the moment of writing this report, I am not confident of the relative scaling between the gradient computed for the local objective function and the gradient computed for the global one.

Uniform velocity error

I compare the results obtained for the uniform velocity error case with the corresponding gradient obtained from conventional full-waveform inversion. To compute the search directions Δs for full-waveform inversion I applied the following expression:

$$\Delta s = - \sum_{x_s} \sum_{x_g} \frac{\partial \tilde{P}'}{\partial s} \left[\tilde{P}(t, x_g, x_s; s_0) - P_D(t, x_g, x_s) \right]. \quad (26)$$

Figures 1–2 show the search direction obtained by applying expression 26. Panels a) show the velocity perturbations for each z and x of the model; panels b) show the velocity perturbation averaged over the horizontal direction. As expected, the search direction computed using expression 26 is pointing in the correct direction when the velocity error is small (0.1%), Figure 1. However, when the velocity error is large the search direction has the wrong polarity (8.0%), Figure 2.

Figure 2 illustrates the limitations of full-waveform inversion when applied to estimating the background velocity. I therefore applied the proposed method to the case with a large velocity error, to demonstrate that it overcomes these limitations of full-waveform inversion.

Figure 3 shows the derivatives of the objective function with respect to the moveout parameters computed using equations 16-17 and 18-19. The plot in panel a) displays the derivatives with respect to the local curvature, μ_C ; the plot in panel b) displays the derivatives with respect to the time shifts of the beam centers, μ_θ . In both cases the derivatives are plotted as a function of the beam center coordinate, \bar{x} , for the source location at $x_s = 2.56$ km; that is, in the middle of the model. As visible from the figure, I clipped to zero the derivatives for both parameters outside of the $-1.7 \text{ km} \leq \bar{x} \leq 1.7 \text{ km}$ range to avoid edge effects.

Figure 3a clearly shows that the local curvature is an appropriate parameters to measure the effects of large-scale velocity errors. In contrast, the alternating signs of the time shift derivatives shown in Figure 3b demonstrates that such perturbations are not well captured by the global moveout parameters.

These observations are confirmed by Figures 4 and 5. Figure 4 shows the search direction computed using equation 24. In contrast with the search direction computed with conventional waveform inversion (Figure 2), the search direction shown in Figure 4 has the correct polarity. It provides a good search direction, similar to the one provided by full-waveform inversion with the small velocity error that is shown in Figure 1.

The search direction obtained by computing the gradient of the global objective function using equation 25 is shown in Figure 5. It shows strong edge effects and its horizontal average (Figure 5b)) has the wrong polarity.

Localized velocity error

To analyze the interplay between the local and the global objective function, I computed the search directions in the case of a spatially localized slowness error. As for the previous examples, I computed the search directions provided by the gradient of both the local and the global objective functions. Figure 6 shows the slowness error that was assumed for the background slowness model s_0 . In addition to the horizontal average, this figure (and the ones that follow) shows the vertical average in panel c) at the bottom of the figure.

Similarly to Figure 3, Figure 7, shows the derivatives of the objective function with respect to the moveout parameters computed using equations 16-17 and 18-19. The plot in panel a) displays the derivatives with respect to the local curvature, μ_C ; the plot in panel b) displays the derivatives with respect to the time shifts of the beam centers, μ_θ . In both cases the derivatives are plotted as a function of the beam center coordinate, \bar{x} , for the source location at $x_s = 2.56$ km; that is, in the middle of the model. As visible from the figure, I

Figure 3: Derivatives of the objective functions with respect to the moveout parameters plotted as a function of beam center coordinate, \bar{x} , for the source location at $x_s = 2.56$ km: a) derivatives of J_{Local} with respect to the local beam curvatures, μ_C , b) derivatives of J_{Global} with respect to the time shifts of the beam centers, μ_θ . [CR]

biondo1/. Beam-DC-DT-large-neg

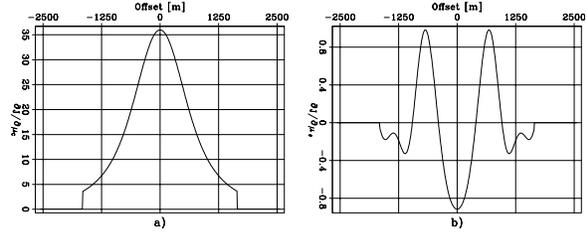


Figure 4: Search direction computed using the gradient of the local objective function J_{Local} with a positive large (8.0%) uniform slowness error: a) slowness perturbations and b) slowness perturbations averaged over the horizontal direction. [CR]

biondo1/. Avg-Dir-Beam-large-neg

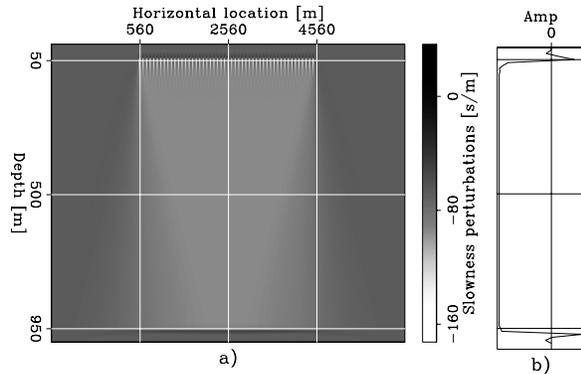


Figure 5: Search direction computed using the gradient of the global objective function J_{Global} with a positive large (8.0%) uniform slowness error: a) slowness perturbations and b) slowness perturbations averaged over the horizontal direction. [CR]

biondo1/. Avg-Dir-Glob-Beam-large-neg

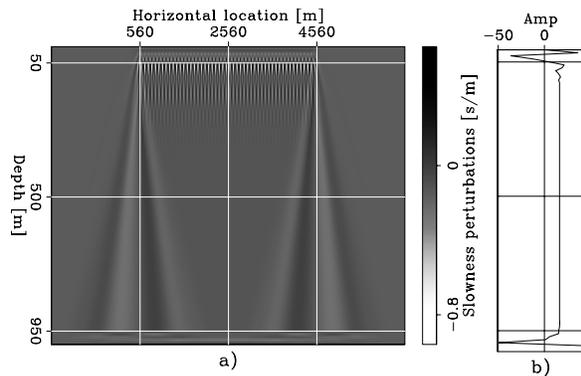
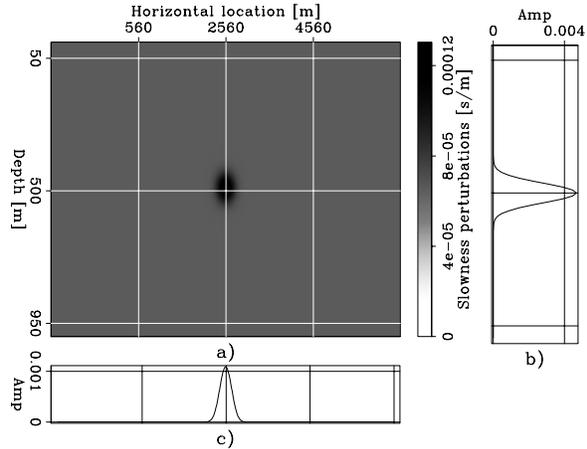


Figure 6: The slowness error that was assumed for the background slowness model s_0 : a) slowness perturbations, b) slowness perturbations averaged over the horizontal direction, and c) slowness perturbations averaged over the vertical direction. [ER]

biondo1/. VH-Avg-DSlow-panom



clipped to zero the derivatives for both parameters outside of the $-2.5 \text{ km} \leq \bar{x} \leq 2.5 \text{ km}$ range to avoid edge effects.

Figure 7a clearly shows that local curvature is not as an appropriate parameter for measuring the effects of small-scale velocity errors as it is for large-scale ones. The alternating signs of the objective-function derivatives causes the search direction to be highly oscillating as well. In contrast, the time-shifts derivatives shown in Figure 7b shows a large anomaly corresponding the localized velocity error and will provide useful slowness updates to localize the velocity anomaly.

These observations are confirmed by Figure 8 and 9. Figure 8 shows the search direction computed using equation 24. As expected it oscillating around the velocity anomaly. Figure 9 shows instead a nicely localized anomaly with the correct sign. It is useful to notice that the horizontal averages of the search directions shown in Figure 8 and 9 have opposite polarity. The average of the search direction provided by the global component has the wrong polarity, except at the depth of the anomaly. Whereas the average of the search direction provided by the local component has the correct polarity. This observation confirms the analysis that local curvature carries more reliable information for the long-wavelength component of the velocity updates than the global time shifts, as observed when analyzing the uniform velocity error example.

Another interesting observation can be made by computing the ratio between the amplitudes of the slowness updates in the two cases. In the case of uniform error, the update computed from the local curvature is larger than the other by approximately a factor of 200. In contrast, in the case of the localized anomaly, the ratio between amplitudes is only about 10. This difference in relative amplitudes confirms that the two components of the objective function switch in relative importance between the two cases.

CONCLUSIONS

To reliably estimate velocity using wavefield operators, I introduce a new objective function that rewards flatness of the correlation between source wavefield and receiver wavefield. The proposed objective function is maximized as a function of the slowness model through the application of residual moveout operators to the correlation. The first term of the objective

Figure 7: Derivatives of the objective functions with respect to the moveout parameters plotted as a function of beam center coordinate, \bar{x} , for the source location at $x_s = 2.56$ km: a) derivatives of J_{Local} with respect to the local beam curvatures, μ_C , b) derivatives of J_{Global} with respect to the time shifts of the beam centers, μ_θ . [CR]

biondo1/. Beam-DC-DT-panom

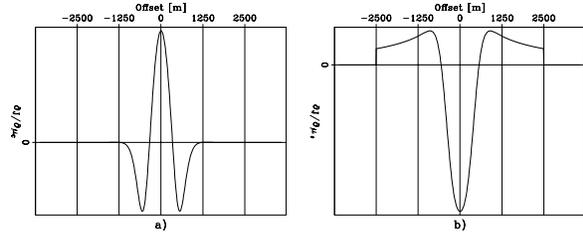


Figure 8: Search direction computed using the gradient of the local objective function J_{Local} with the localized velocity error shown in Figure 6: a) slowness perturbations, b) slowness perturbations averaged over the horizontal direction, and c) slowness perturbations averaged over the vertical direction. [CR]

biondo1/. VH-Avg-Dir-Beam-panom

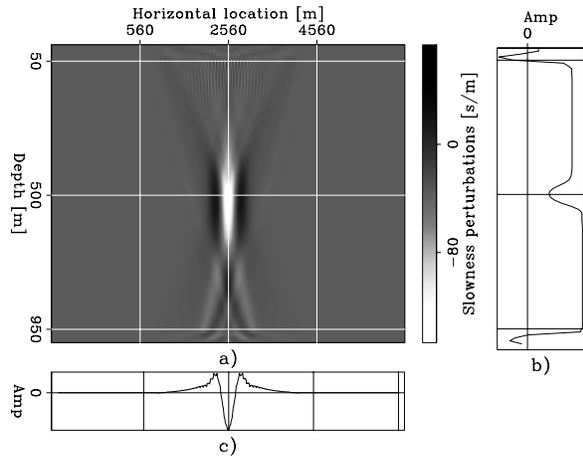
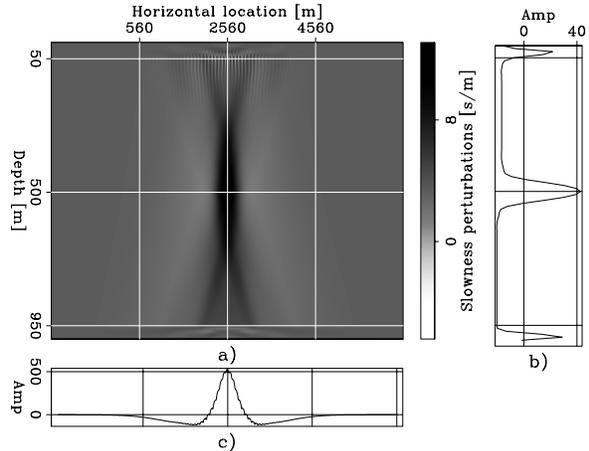


Figure 9: Search direction computed using the gradient of the global objective function J_{Global} with the localized velocity error shown in Figure 6: a) slowness perturbations, b) slowness perturbations averaged over the horizontal direction, and c) slowness perturbations averaged over the vertical direction. [CR]

biondo1/. VH-Avg-Dir-Glob-Beam-panom



function measures the power of the stack over local beams as a function of the local beam curvature. Maximization of this first term ensures global convergence in presence of large velocity errors. The second term maximizes the global power of the stack as a function of time shifts applied to the local stack over the beams. Maximization of this second term helps the estimation of localized velocity errors.

I tested the application of the proposed objective function by computing its gradients for two simple problems: the estimation of a large and spatially uniform velocity error and the estimation of a spatially localized velocity error. The computed search directions confirm the potential of the proposed method and illustrate the different roles played by the local and the global terms of the objective function.

REFERENCES

- Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.
- Luo, Y. and G. T. Schuster, 1991, Wave-equation traveltime inversion: *Geophysics*, **56**, 645–653.
- Sava, P., 2004, Migration and velocity analysis by wavefield extrapolation: PhD thesis, Stanford University.
- Sava, P. and B. Biondi, 2004a, Wave-equation migration velocity analysis—I: Theory: *Geophysical Prospecting*, **52**, 593–623.
- , 2004b, Wave-equation migration velocity analysis—II: Examples: *Geophysical Prospecting*, **52**, 607–623.
- Shen, P., 2004, Wave equation migration velocity analysis by differential semblance optimization: PhD thesis, Rice University.
- Shen, P., W. W. Symes, S. Morton, and H. Calandra, 2005, Differential semblance velocity analysis via shot profile migration: SEG, Expanded Abstracts, **24**, 2249–2252.
- Symes, W. W. and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: *Geophysics*, **56**, 654–663.

APPENDIX A

DETAILS OF GRADIENT COMPUTATION

In this appendix I present the analytical development needed to derive equations 22-23 from equations 20-21.

Equation 20 can be rewritten as

$$\begin{aligned} \frac{\partial \mu_{\bar{x}}}{\partial s} \Big|_{\mu_{\bar{x}}=\bar{\mu}_{\bar{x}}} &= - \frac{\partial J_{\text{FL}}(\mu_{\bar{x}})}{\partial s} \Big|_{\mu_{\bar{x}}=\bar{\mu}_{\bar{x}}} \\ &= - \frac{\left\langle \frac{\partial \mathcal{M}_{\bar{x}}[\theta(\mu_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s_0)]}{\partial \mu_{\bar{x}}} \Big|_{\mu_{\bar{x}}=\bar{\mu}_{\bar{x}}}, \mathbf{B}_{\bar{x}} \mathbf{P}_D \frac{\partial \tilde{P}}{\partial s} \right\rangle}{\left\langle \frac{\partial^2 \mathcal{M}_{\bar{x}}[\theta(\mu_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s_0)]}{\partial \mu_{\bar{x}}^2} \Big|_{\mu_{\bar{x}}=\bar{\mu}_{\bar{x}}}, \mathbf{B}_{\bar{x}}C(\tau; s) \right\rangle}, \end{aligned}$$

where,

$$\begin{aligned} \frac{\partial^2 \mathcal{M}_{\bar{x}} [\theta (\boldsymbol{\mu}_{\bar{x}}), \mathbf{B}_{\bar{x}} C (\tau; s_0)]}{\partial \boldsymbol{\mu}_{\bar{x}}^2} &= \mathcal{M}_{\bar{x}} \left[\theta (\boldsymbol{\mu}_{\bar{x}}), \mathbf{B}_{\bar{x}} \dot{C} (\tau; s_0) \right] \frac{\partial^2 \theta}{\partial \boldsymbol{\mu}_{\bar{x}}^2} \\ &+ \mathcal{M}_{\bar{x}} \left[\theta (\boldsymbol{\mu}_{\bar{x}}), \mathbf{B}_{\bar{x}} \ddot{C} (\tau; s_0) \right] \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\bar{x}}}, \end{aligned}$$

and in which $\ddot{C} (\tau) = C (\tau) \left[\tilde{P} (t), \ddot{P}_D (t) \right]$. Given the moveout parametrization expressed in 6, $\partial^2 \theta / \partial \boldsymbol{\mu}_{\bar{x}}^2 = 0$ and the previous expression simplifies into the following:

$$\frac{\partial^2 \mathcal{M}_{\bar{x}} [\theta (\boldsymbol{\mu}_{\bar{x}}), \mathbf{B}_{\bar{x}} C (\tau; s_0)]}{\partial \boldsymbol{\mu}_{\bar{x}}^2} = \mathcal{M}_{\bar{x}} \left[\theta (\boldsymbol{\mu}_{\bar{x}} (s)), \mathbf{B}_{\bar{x}} \ddot{C} (\tau; s_0) \right] \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\bar{x}}}.$$

Consequently, the general expression for the gradient of the local moveout parameters with respect to the slowness model is:

$$\left. \frac{\partial \boldsymbol{\mu}_{\bar{x}}}{\partial s} \right|_{\boldsymbol{\mu}_{\bar{x}} = \bar{\boldsymbol{\mu}}_{\bar{x}}} = - \frac{\left\langle \mathcal{M}_{\bar{x}} \left[\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} \dot{C} (\tau; s_0) \right] \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\bar{x}}}, \mathbf{B}_{\bar{x}} \mathbf{P}_D \frac{\partial \tilde{P}}{\partial s} \right\rangle}{\left\langle \mathcal{M}_{\bar{x}} \left[\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} \ddot{C} (\tau; s_0) \right] \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\bar{x}}}, \mathbf{B}_{\bar{x}} C (\tau; s) \right\rangle}.$$

When $\bar{\boldsymbol{\mu}}_{\bar{x}} = 0$, the general expression further simplifies into:

$$\left. \frac{\partial \boldsymbol{\mu}_{\bar{x}}}{\partial s} \right|_{\boldsymbol{\mu}_{\bar{x}} = 0} = - \frac{\left\langle \mathbf{B}_{\bar{x}} \dot{C} (\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\bar{x}}}, \mathbf{B}_{\bar{x}} \mathbf{P}_D \frac{\partial \tilde{P}}{\partial s} \right\rangle}{\left\langle \mathbf{B}_{\bar{x}} \ddot{C} (\tau; s_0) \frac{\partial \theta}{\partial \boldsymbol{\mu}_{\bar{x}}}, \mathbf{B}_{\bar{x}} C (\tau; s) \right\rangle}. \quad (\text{A-1})$$

Similar derivation can be developed for the derivative of the global moveout parameters with respect to slowness. Equation 21 can be rewritten as:

$$\begin{aligned} \left. \frac{\partial \boldsymbol{\mu}}{\partial s} \right|_{\boldsymbol{\mu} = \bar{\boldsymbol{\mu}}} &= - \frac{\frac{\partial J_{\text{FG}}(\boldsymbol{\mu})}{\partial s}}{\frac{\partial J_{\text{FG}}(\boldsymbol{\mu})}{\partial \boldsymbol{\mu}}} \\ &= - \frac{\left\langle \frac{\partial \mathcal{M} \{ \theta (\boldsymbol{\mu}), \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathcal{M}_{\bar{x}} [\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} C (\tau; s_0)] \}}{\partial \boldsymbol{\mu}} \Big|_{\boldsymbol{\mu} = \bar{\boldsymbol{\mu}}}, \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathcal{M}_{\bar{x}} \left[\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} \mathbf{P}_D \frac{\partial \tilde{P}}{\partial s} \right] \right\rangle}{\left\langle \frac{\partial^2 \mathcal{M} \{ \theta (\boldsymbol{\mu}), \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathcal{M}_{\bar{x}} [\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} C (\tau; s_0)] \}}{\partial \boldsymbol{\mu}^2} \Big|_{\boldsymbol{\mu} = \bar{\boldsymbol{\mu}}}, \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathcal{M}_{\bar{x}} \left[\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} C (\tau; s) \right] \right\rangle}, \end{aligned}$$

where,

$$\begin{aligned} \frac{\partial^2 \mathcal{M} \{ \theta (\boldsymbol{\mu}), \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathcal{M}_{\bar{x}} [\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} C (\tau; s_0)] \}}{\partial \boldsymbol{\mu}^2} &= \\ \mathcal{M} \left\{ \theta (\boldsymbol{\mu}), \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathcal{M}_{\bar{x}} \left[\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} \dot{C} (\tau; s_0) \right] \right\} \frac{\partial^2 \theta}{\partial \boldsymbol{\mu}^2} &+ \\ \mathcal{M} \left\{ \theta (\boldsymbol{\mu}), \boldsymbol{\Sigma}_{\bar{x}} \mathbf{S}_{\bar{x}} \mathcal{M}_{\bar{x}} \left[\theta (\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}} \ddot{C} (\tau; s_0) \right] \right\} \frac{\partial \theta}{\partial \boldsymbol{\mu}}. & \end{aligned}$$

Given the moveout parametrization in expressed in 10, $\partial^2\theta/\partial\boldsymbol{\mu}^2 = 0$ and the previous expression simplifies into:

$$\frac{\partial^2\mathcal{M}\{\theta(\boldsymbol{\mu}), \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathcal{M}_{\bar{x}}[\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s_0)]\}}{\partial\boldsymbol{\mu}^2} = \mathcal{M}\left\{\theta(\boldsymbol{\mu}), \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathcal{M}_{\bar{x}}\left[\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}}\ddot{C}(\tau; s_0)\right]\right\} \frac{\partial\theta}{\partial\boldsymbol{\mu}}.$$

The general expression for the gradient of the global moveout parameters with respect to the slowness model is:

$$\frac{\partial\boldsymbol{\mu}}{\partial s}\bigg|_{\boldsymbol{\mu}=\bar{\boldsymbol{\mu}}} = \frac{\left\langle \mathcal{M}\left\{\theta(\bar{\boldsymbol{\mu}}), \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathcal{M}_{\bar{x}}\left[\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}}\dot{C}(\tau; s_0)\right]\right\} \frac{\partial\theta}{\partial\boldsymbol{\mu}}, \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathcal{M}_{\bar{x}}\left[\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}}\mathbf{P}_D \frac{\partial\tilde{P}}{\partial s}\right]\right\rangle}{\left\langle \mathcal{M}\left\{\theta(\bar{\boldsymbol{\mu}}), \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathcal{M}_{\bar{x}}\left[\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}}\ddot{C}(\tau; s_0)\right]\right\} \frac{\partial\theta}{\partial\boldsymbol{\mu}}, \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathcal{M}_{\bar{x}}\left[\theta(\bar{\boldsymbol{\mu}}_{\bar{x}}), \mathbf{B}_{\bar{x}}C(\tau; s)\right]\right\rangle}.$$

When $\bar{\boldsymbol{\mu}}_{\bar{x}} = 0$ and $\bar{\boldsymbol{\mu}} = 0$ the general expression further simplifies into:

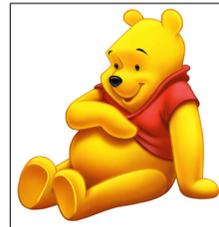
$$\frac{\partial\boldsymbol{\mu}}{\partial s}\bigg|_{\bar{\boldsymbol{\mu}}_{\bar{x}}=0, \boldsymbol{\mu}=0} = - \frac{\left\langle \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathbf{B}_{\bar{x}}\dot{C}(\tau; s_0) \frac{\partial\theta}{\partial\boldsymbol{\mu}}, \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathbf{B}_{\bar{x}}\mathbf{P}_D \frac{\partial\tilde{P}}{\partial s}\right\rangle}{\left\langle \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathbf{B}_{\bar{x}}\ddot{C}(\tau; s_0) \frac{\partial\theta}{\partial\boldsymbol{\mu}}, \boldsymbol{\Sigma}_{\bar{x}}\mathbf{S}_{\bar{x}}\mathbf{B}_{\bar{x}}C(\tau; s)\right\rangle}. \quad (\text{A-2})$$

Research Personnel

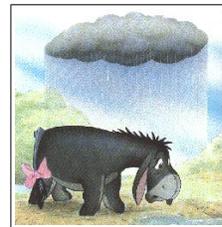
Ali Almomin graduated from Texas A&M University in 2007 with a BS in Geophysics. Then, he joined Saudi Aramco and worked in several exploration and research departments with a focus on 3D seismic processing and near surface imaging. He joined Stanford Exploration Project in 2009 to pursue a PhD in Geophysics and is currently working on seismic tomography. He is a member of SEG, EAGE, and SPE.



Gboyega Ayeni received his B.Sc. in Applied Geophysics from Obafemi Awolowo University, Nigeria in 2004. He was a Shell scholar at University of Leeds, where he obtained an M.Sc with Distinction in Exploration Geophysics. Gboyega joined SEP in September 2006 to work towards his Ph.D in Geophysics. He is a member of SEG, EAGE, AGU, SPE and AAPG.



Ohad Barak received a B.Sc. (2006) and an M.Sc. (2009) in Geophysics from Tel-Aviv University. In 2008 he joined Paradigm Geophysical and worked there as a developer. He joined SEP in 2009 and is currently pursuing a Ph.D. in geophysics at Stanford University, and a longer biography.



Near-surface velocity estimation by weighted early-arrival waveform inversion

Xukai Shen

ABSTRACT

In this paper, I present a modified version of the conventional waveform inversion objective function to bridge the gap between the acoustic waveform inversion engine and the more complicated physics in recorded data. The proposed method weighs the amplitude of observed and modeled data. In this way, I use more of the phase information rather than the amplitude information in recorded data. Since phase information is more robust in the presence of visco-acoustic or even elastic related physics. Synthetic examples with inversion of visco-acoustic data and elastic data show that using the proposed objective function can recover detailed velocity structures even with the presence of non-acoustic physics, while conventional waveform inversion tends to fail.

INTRODUCTION

Near-surface velocity is important for imaging deeper targets. Complex near-surface velocity can cause serious problems for imaging deeper targets if it is not accurately estimated. Conventionally, people use ray-based methods (Hampson and Russell, 1984; Olson, 1984; White, 1989) to derive the large-scale structure of near-surface velocity. Such smooth velocity structure may be adequate for areas with simple near-surface velocity. However, in geologically complex areas, smooth velocity is not accurate enough for imaging deeper reflectors (Marsden, 1993; Bevc, 1995; Hindriks and Verschuur, 2001). In such cases, waveform inversion (Tarantola, 1984; Pratt et al., 1998; Mora, 1987) tends to give more accurate results (Ravaut et al., 2004; Sheng et al., 2006; Sirgue et al., 2009) by using finite-frequency seismic wave propagation.

Yet there are several important factors for practical application of waveform inversion. Among these are the quality of the starting velocity model, the accuracy of the source wavelet estimation, and the complexity of physics in the recorded data and the inversion engine. In addition to the P-wave velocity, the density, S-wave velocity, anisotropy parameters, attenuation and other factors all affect the recorded data. When compared with data modeled from the constant-density acoustic wave equation, these extra parameters will not only change the amplitude and phase of existing events, but they also may add extra events, such as converted waves. However, there are several problems associated with inverting these parameters. First, they are much less well constrained than the P-wave velocity, so it is much more difficult to invert all these parameters than just P-wave velocity. Second, it is more computationally expensive to incorporate all these parameters into the wave-equation engine used in inversion; for example, modeling using the elastic wave equation is at least an order of magnitude more expensive than using the acoustic wave equation. Thus the acoustic wave equation is still the most practical waveform inversion

engine so far. With this choice, however, the objective function of conventional waveform inversion is inadequate to bridge the gap between the physics of the data and the physics of the inversion engine.

It is known that traveltimes/phase information is less sensitive to the presence of different physics in the recorded data and also carries the information about velocity field (Luo and Schuster, 1990; Shin and Min, 2006). In conventional waveform inversion, both phase and amplitude information of the modeled data are compared with recorded data. I modify the objective function of waveform inversion, giving added weight to the match of phase information in the modeled and the recorded data for each iteration of the inversion. I do not match the recorded and modeled data by subtracting their phases, thus avoiding the ambiguity caused by phase wrapping (Shin and Min, 2006).

THEORY

The generalized waveform inversion objective function can be written as

$$f(\mathbf{d}_{\text{obs}}, \mathbf{D}(\mathbf{m})) \approx \mathbf{0}, \quad (1)$$

where f is a function of \mathbf{d}_{obs} , observed data and $\mathbf{D}(\mathbf{m})$, forward modeled synthetic data from \mathbf{m} , the velocity model. Observed data can be in either the frequency domain or the time domain, depending on the actual form of \mathbf{f} . For example if we take \mathbf{f} as the L2 norm of $(\mathbf{d}_{\text{obs}} - \mathbf{D}(\mathbf{m}))$, we obtain the objective function of conventional waveform inversion (Tarantola, 1984; Pratt et al., 1998); if we take \mathbf{f} as the L2 norm of natural logarithm of $\mathbf{D}(\mathbf{m})/\mathbf{d}_{\text{obs}}$, we obtain the so called logarithmic objective function of waveform inversion (Shin and Min, 2006). However, direct comparison in the phase domain will encounter the problem of phase wrapping. I modify the objective function of conventional waveform inversion by using the following expression for \mathbf{f} :

$$\begin{aligned} f &= \sum_{\mathbf{s}, \mathbf{r}} \|\delta \mathbf{d}(\mathbf{s}, \mathbf{r})\|^2 \\ &= \left\| \frac{\mathbf{D}(\mathbf{s}, \mathbf{r}, \mathbf{m})}{\sqrt{\mathbf{D}^T(\mathbf{s}, \mathbf{r}, \mathbf{m})\mathbf{D}(\mathbf{s}, \mathbf{r}, \mathbf{m})}} - \frac{\mathbf{d}_{\text{obs}}(\mathbf{s}, \mathbf{r})}{\sqrt{\mathbf{d}_{\text{obs}}^T(\mathbf{s}, \mathbf{r})\mathbf{d}_{\text{obs}}(\mathbf{s}, \mathbf{r})}} \right\|^2 \end{aligned} \quad (2)$$

where \mathbf{m} is the model, which consists of near-surface velocity; \mathbf{d}_{obs} are the data, which consist of band-passed early arrivals of the wavefield. Data are filtered based on traveltimes difference before bandpassing to exclude events that come relatively late in the early arrivals. \mathbf{D} is the constant-density two-way acoustic wave-equation operator that generates synthetic early arrivals from source and near-surface velocity; \mathbf{s} and \mathbf{r} are source and receiver locations, respectively. In the new objective function, I weight both recorded data and forward-modeled data by their RMS energy trace by trace. This ensures that the recorded data and the forward-modeled data have relatively the same amplitudes, and waveform inversion in this case will focus more on phase comparison. To update the velocity, I use the nonlinear conjugate gradient method. The gradient of equation 2 with respect to \mathbf{m} is as follows:

$$\mathbf{J}_{\mathbf{F}}(\mathbf{m}) = 2\delta \mathbf{d}^T(\mathbf{s}, \mathbf{r}) \frac{\partial \left(\frac{\mathbf{D}(\mathbf{s}, \mathbf{r}, \mathbf{m})}{\sqrt{\mathbf{D}^T(\mathbf{s}, \mathbf{r}, \mathbf{m})\mathbf{D}(\mathbf{s}, \mathbf{r}, \mathbf{m})}} \right)}{\partial \mathbf{m}} \quad (3)$$

After some algebraic manipulations, equation 2 becomes:

$$\mathbf{J}_f(\mathbf{m}) = 2\mathbf{V}_{\text{src}} \frac{\partial \mathbf{D}(\mathbf{s}, \mathbf{r}, \mathbf{m})}{\partial \mathbf{m}} \quad (4)$$

where \mathbf{V}_{src} is defined as

$$\begin{aligned} \mathbf{V}_{\text{src}} &= \frac{\delta \mathbf{d}(\mathbf{s}, \mathbf{r})}{\sqrt{\mathbf{D}(\mathbf{s}, \mathbf{r}, \mathbf{m})^T \mathbf{D}(\mathbf{s}, \mathbf{r}, \mathbf{m})}} \\ &+ \frac{\delta \mathbf{d}^T(\mathbf{s}, \mathbf{r}) \mathbf{D}(\mathbf{s}, \mathbf{r}, \mathbf{m})}{\mathbf{D}(\mathbf{s}, \mathbf{r}, \mathbf{m})^T \mathbf{D}(\mathbf{s}, \mathbf{r}, \mathbf{m})} \mathbf{D}(\mathbf{s}, \mathbf{r}, \mathbf{m}) \end{aligned} \quad (5)$$

For gradient calculation, the only difference between conventional waveform inversion and weighted waveform inversion is how to calculate the virtual source. In conventional waveform inversion, the difference between observed data and modeled data is used as virtual source for reverse-time propagation. In weighted waveform inversion, a new virtual source defined in 5 is used for reverse-time propagation.

The step length calculation assumes local parabolic behavior of the objective function (Vigh and Starr, 2008). I first obtain two different perturbations of the velocity $\delta \mathbf{m}$ by scaling the gradient to about three percent of the minimum value in the current velocity; then I calculate the misfit function value, and determine the minimum of the parabola by using the two perturbation step length α_1 and α_2 and the existing velocity $\alpha_0 = \mathbf{0}$.

In the synthetic example shown next, I assume a known source wavelet in both examples. I carry out the waveform inversion using several frequency bands, starting from low frequency and gradually moving to higher frequency (Sirgue and Pratt, 2004). The inversion result of each frequency range is used as the starting model for inversion in the next frequency range. Using the low-frequency component of the data ensures good estimation of long-wavelength components of the velocity model, and subsequent higher-frequency inversion will retrieve finer velocity structures. As will be shown later, inversion results using this objective function are robust when we have attenuation or elastic waves in the recorded data.

EXAMPLES

To illustrate the idea, I show two examples of acoustic inversion of data generated with wave equations contain more complex physics than pure acoustics. The first example uses data generated by the visco-acoustic wave equation, which means there is attenuation in the subsurface that causes phase/amplitude change in the data. The second example uses data generated by the elastic wave equation. This not only changes the phase/amplitude of existing events relative to data generated by acoustic wave equation, but also add extra events which are converted modes. However in both cases, by carefully masking out non-acoustic events and focusing on the phase of "acoustic-equivalent" data, acoustic inversion can still produce a good estimate of near-surface velocity.

Data recorded with strong subsurface attenuation

Here I estimate near-surface velocity using recorded data with strong subsurface attenuation. The model is from Amoco-2.5D synthetic velocity model. The Q value is a scaled

version of velocity to maintain consistency between structures. The velocity model is 600 by 140 grid in x and z , with 10 m spacing in both directions. The acquisition geometry is relatively simple. Receivers are on all the grid points on the surface, there are 30 shots with uniform spacing of 100 m.

A shot gather at a single location from data is shown in figure 1; the left panel shows the data without attenuation, and the right panel shows the data with attenuation. It can be seen that the attenuated data has lower frequency content and much weaker amplitude. However, the kinematics of the data does not change much. Thus, in this case, phase information in the data is relatively insensitive to non-acoustic physics in the recorded data. Figure 2 shows the true near-surface velocity and starting velocity model. The results of estimated velocity using conventional waveform inversion fitting goals and weighted waveform inversion fitting goal are shown in figure 3. It can be seen that with strongly attenuated data, weighted waveform inversion is still able to recover the velocity model, while conventional fitting goal completely fails in this case. Since conventional waveform inversion use both phase and amplitude information in the recorded data, it tries to interpret weak amplitude as a result of a fast velocity layer at the surface. The new objective function, however, focuses more on phase information and partially ignores amplitude information to give a much better result.

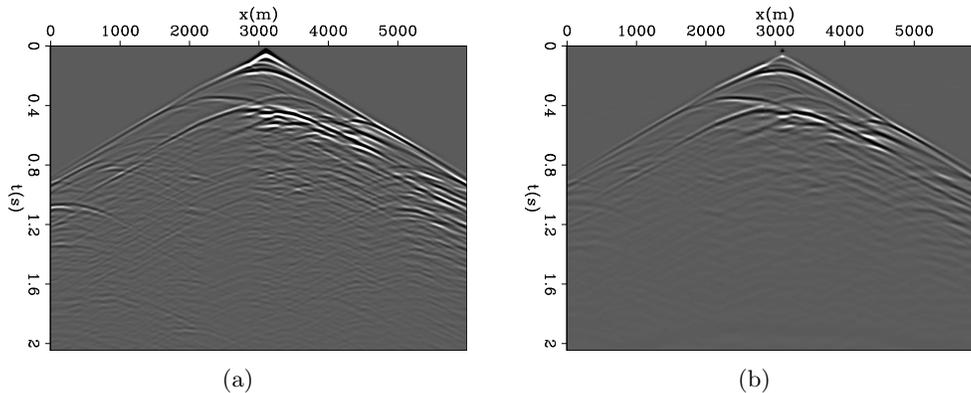


Figure 1: One synthetic shot with and without attenuation. Both panels have the same clip value. a) One shot generated by the acoustic wave equation. b) Same shot gather generated by the visco-acoustic wave equation with the amplitude of the entire shot gather scaled by a factor of ten. `xukai1/. nrmrcd,attencd`

Data recorded with elastic wave equation

Next I estimate the the near-surface velocity using recorded data that is generated by the elastic wave equation. The synthetic P-wave velocity model used in this case is provided by Saudi Aramco. It simulates typical near-surface velocity encountered in Saudi Arabia. For elastic modeling, the density is obtained by the Gardner equation (Gardner et al., 1974). S-wave velocity is scaled from P-wave velocity. The model is 1600 by 176 grid in x and z , with 5 m spacing in both directions. The acquisition geometry is also relatively simple. Receivers are on all the grid points on the surface, and there are 90 shots with equal spacing of 80 m. Inversion is done on a 10 m spacing grid, which is coarser than the grid used for

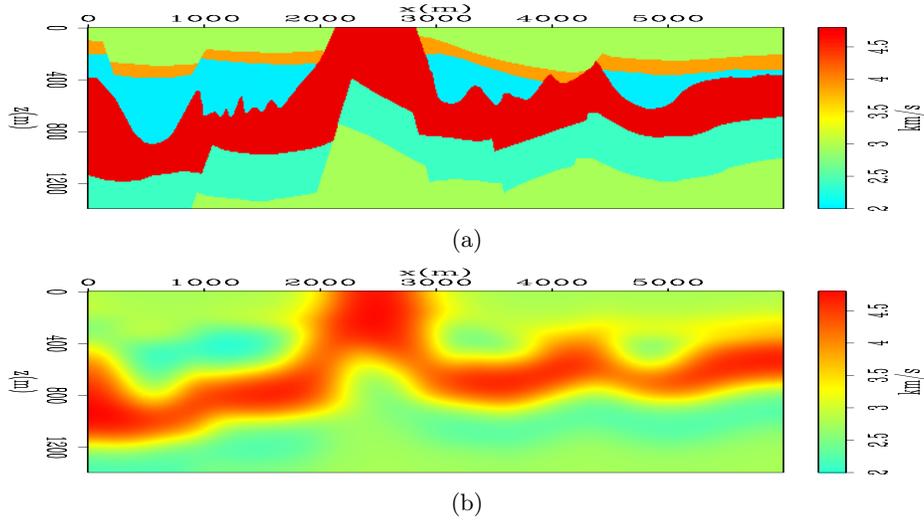


Figure 2: a) True near-surface velocity. b) Starting near-surface velocity model used for waveform inversion. `xukai1/. attenveltrue,attenvelstart`

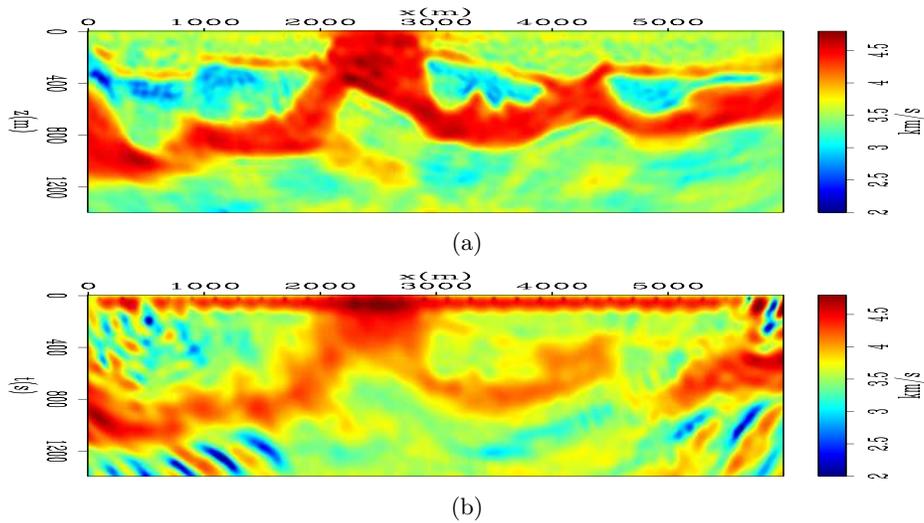


Figure 3: Estimated velocity using different objective functions. a) Estimated velocity using the new objective function. b) Estimated velocity using the conventional objective function. `xukai1/. velinpropattenred,velinconvattenrcd`

modeling. The major goal of acoustic inversion in this case is to see if waveform inversion can retrieve the velocity inversion layer.

One shot gather from data with acoustic and elastic modeling is shown in figure 4; the left panel is the data with acoustic modeling, and the right panel is the data with elastic modeling. It can be seen that data from elastic modeling has some amplitude/phase difference and also some extra events that are converted waves. In this case, I scaled the amplitude of both shots to approximately the same level for display purposes. Figure 5 shows refraction events for the same shot. When I carefully mask out most of the converted waves, the remaining refraction events looks very similar in terms of kinematics; the difference at this point is likely to be caused by different distribution of P-wave and S-wave energy at different reflector boundaries. Figure 6 shows the true near-surface velocity and the starting velocity model; the starting velocity model is a smoothed version of true velocity model. The result of estimated velocity using proposed waveform inversion fitting goal are shown in figure 7. The top panel is the result of inverting acoustic data, whereas the bottom is the result of inverting elastic data. It can be seen that even with strong elastic effects in data, after excluding most converted waves, the weighted fitting goal was still able to recover the velocity inversion structure. Compared with inversion of acoustic data, there are some small shifts of reflector locations. This kind of shift is likely due to the very sharp velocity contrast immediately below the slow velocity.

This example shows that without the presence of most converted wave, acoustic inversion of elastic data can still obtain a decent velocity estimation result.

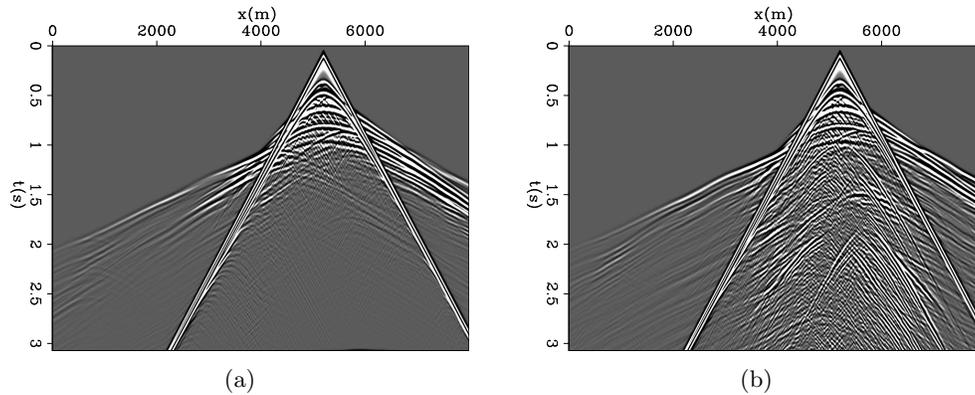


Figure 4: Shot gather generated by different modeling equations: a) shot generated by the acoustic wave equation b) the same shot generated by the elastic wave equation.

xukai1/. Arcd,Ercd

CONCLUSIONS

The conventional waveform inversion objective function is insufficient when the acoustic wave equation is used in inversion. The insufficiency mainly comes from the fact that non-acoustic physics tend to alter the amplitude and phase of existing events and add new events compared with acoustic data. A new weighted waveform inversion objective function is quite robust, despite the presence of non-acoustic physics in recorded data, and it can recover important velocity structure that can not be detected by conventional ray-based

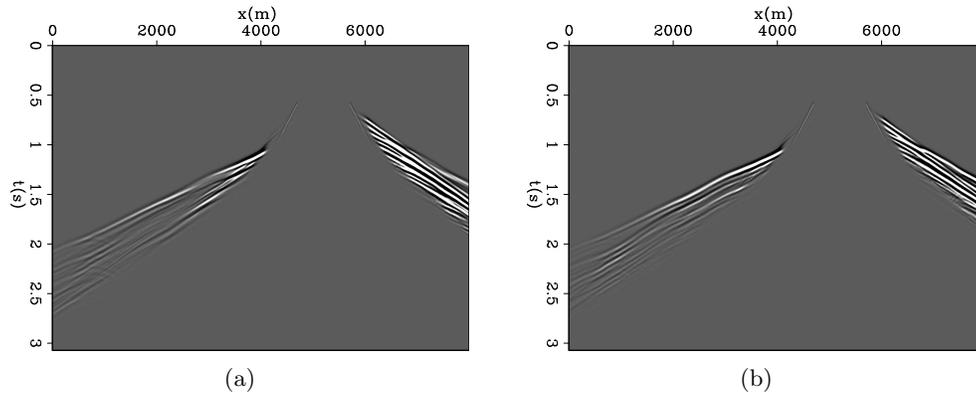


Figure 5: Refraction of the same shot generated by different modeling equations: a) refraction of shot generated by the acoustic wave equation b) refraction of same shot generated by the elastic wave equation. `xukai1/. Arefrac,Erefrac`

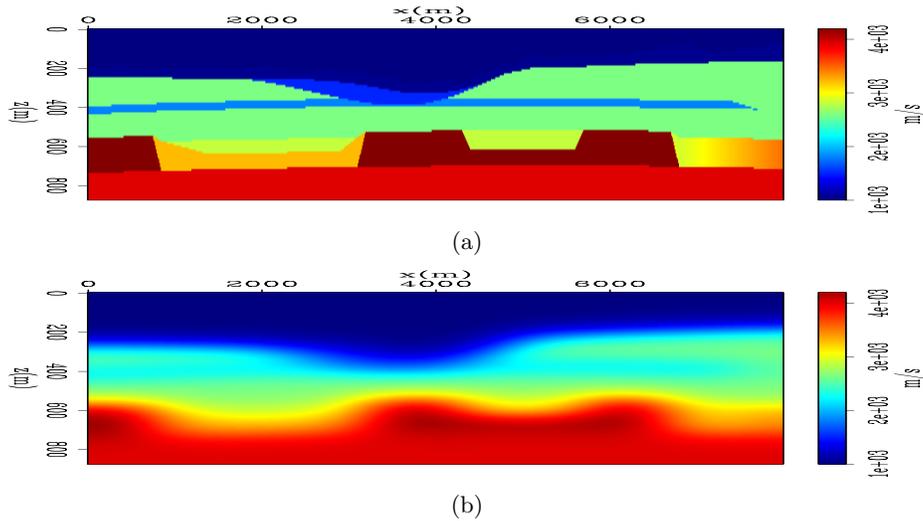


Figure 6: a) True near-surface velocity. b) Starting near-surface velocity model used for waveform inversion; the starting velocity model is a smoothed version of the true velocity model. `xukai1/. veltrue,velstart`

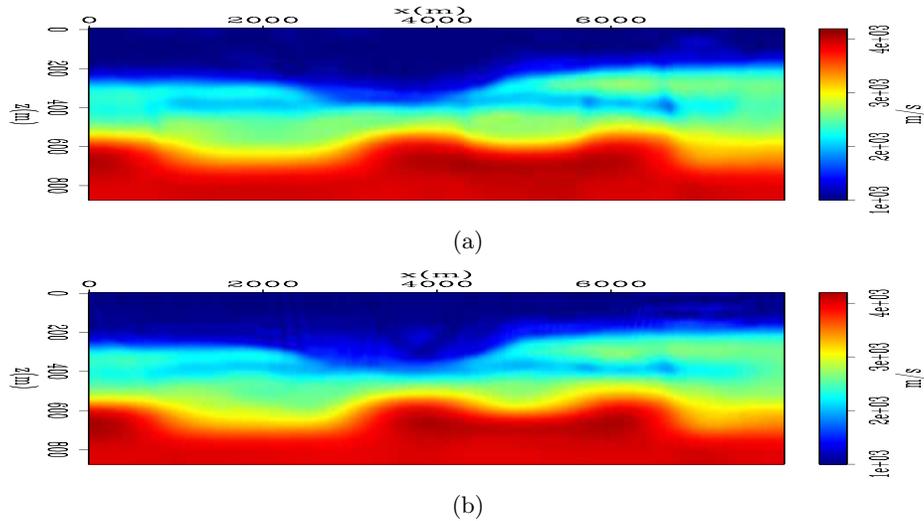


Figure 7: Estimated velocity using data from different modeling equations: a) estimated velocity using data generated by the acoustic wave equation. b) Estimated velocity using data generated by the elastic wave equation. `xukai1/. velinvArcd,velinvErcd`

tomography. For the purpose of near-surface velocity estimation, phase information is much less sensitive than amplitude information to the presence of non-acoustic parameters, and extra events caused by non-acoustic physics are usually late arriving and can be masked out during waveform matching. The velocity estimation results from this objective function are more stable than conventional waveform inversion.

ACKNOWLEDGMENTS

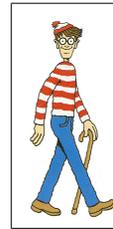
I would like to thank Stanford Exploration Project for the financial support of this research. Also I would like to thank Saudi Aramco for allowing me to use the synthetic velocity model.

REFERENCES

- Bevc, D., 1995, Imaging under rugged topography and complex velocity structure: SEP Ph.D Thesis.
- Gardner, G. H. F., L. W. Gardner, and A. R. Gregory, 1974, Formation velocity and density-the diagnostic basics for stratigraphic traps: *Geophysics*, **39**, 770–780.
- Hampson, D. and B. Russell, 1984, First-break interpretation using generalized linear inversion: *Journal of CSEG*, **20**, 40–54.
- Hindriks, C. and D. Verschuur, 2001, Common focus point approach to complex near surface effects: SEG Expanded Abstracts.
- Luo, Y. and G. T. Schuster, 1990, Wave-equation traveltime inversion: SEG Expanded Abstracts.
- Marsden, D., 1993, Statics corrections-a review: The Leading Edge.
- Mora, P., 1987, Elastic wavefield inversion: SEP Ph.D Thesis.

- Olson, K. B., 1984, A stable and flexible procedure for the inverse modelling of seismic first arrivals: *Geophysical Prospecting*, **37**, 455–465.
- Pratt, R. G., C. Shin, and G. Hicks, 1998, Gauss-Newton and full Newton methods in frequency domain seismic waveform inversion: *Geophysical Journal International*, **133**, 341–362.
- Ravaut, C., S. Operto, L. Imbrota, J. Virieux, A. Herrero, and P. Dell’Aversana, 2004, Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt: *Geophysical Journal International*, **159**, 1032–1056.
- Sheng, J., A. Leeds, M. Buddensiek, and G. T. Schuster, 2006, Early arrival waveform tomography on near-surface refraction data: *Geophysics*, **71**, U47–U57.
- Shin, C. and D. J. Min, 2006, Waveform inversion using a logarithmic wavefield: *Geophysics*, **71**, R31–R42.
- Sirgue, L., O.I.Barkved, J. V. Gestel, O. Askim, and R. Kommedal, 2009, 3d waveform inversion on valhall wide-azimuth obc: EAGE 71th Conference.
- Sirgue, L. and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: *Geophysics*, **69**, 231–248.
- 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**, 135–144.
- White, D. J., 1989, Two-dimensional seismic refraction tomography: *Geophysical Journal International*, **97**, 223–245.

Biondo L. Biondi graduated from Politecnico di Milano in 1984 and received an M.S. (1988) and a Ph.D. (1990) in geophysics from Stanford. SEG Outstanding Paper award 1994. During 1987, he worked as a Research Geophysicist for TOTAL, Compagnie Francaise des Petroles in Paris. After his Ph.D. at Stanford, Biondo worked for three years with Thinking Machines Co. on the applications of massively parallel computers to seismic processing. After leaving Thinking Machines, Biondo started 3DGeo Development, a software and service company devoted to high-end seismic imaging. Biondo is now Associate Professor (Research) of Geophysics and leads SEP efforts in 3-D imaging. He is a member of SEG and EAGE.



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.



Wave-equation tomography for anisotropic parameters

Yunyue (Elita) Li and Biondo Biondi

ABSTRACT

Anisotropic models are recognized as more realistic representations of the subsurface where complex geological environment exists. These models are widely needed by many anisotropic migration and interpretation schemes. However, anisotropic model building is still a challenging problem in the industry. In this paper, we propose an approach to build the anisotropic model using Wave-Equation Tomography (WETom) on surface seismic data in the image space. To reduce the null space of the inversion, we parametrize our model space using only vertical velocity (V_V) and the anellipticity parameter η . Numerical tests show that anisotropic WETom is effective in resolving model perturbations; however, ambiguity exists between the vertical velocity and the anellipticity parameter. Auxiliary information is needed to further constrain the inversion.

INTRODUCTION

Since first reported in exploration seismology in the 1930s (McCollum and Snell, 1932), the importance of anisotropy has been continuously increased in seismic imaging and exploration. This is partially due to acquisition with increasingly longer offsets and exploration in areas with strong geological deformation. Until now, the transverse isotropic (TI) model has been the most commonly used anisotropic model in seismic imaging. Postma (1955) and Helbig (1956) showed that a sequence of isotropic layers on a scale much smaller than the wavelength leads to an anisotropic medium. For the case of horizontal layers, the medium can be described by an equivalent vertical transverse isotropic (VTI) medium. When dip is present, the medium develops a tilted transverse anisotropy (TTI). Many authors (Shan, 2009; Fletcher et al., 2009; Zhang and Zhang, 2009; Fei and Liner, 2008) have developed various migration and processing schemes for VTI and TTI medium; however, the estimation of the anisotropy model is still challenging.

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

Wave-equation tomography (WETom) has been widely investigated in isotropic velocity building, and can be implemented either in the data space (Tarantola, 1984; Woodward, 1992) or in the image space (Sava and Biondi, 2004a,b; Shen, 2004; Shen and Symes, 2008; Guerra et al., 2009). Several advantages drive us to use the image-space WETom instead of the data-space version: First, the migration image is often much cleaner than the recorded wavefields. Second, we can use cheaper one-way extrapolators in the image space, compared

with expensive two-way extrapolators in the data space. Third, the objective function is directly related to the final image. Therefore, we choose to extend image-space WETom from isotropic velocity building to anisotropic model building.

In this paper, we first explain the parametrization of the inversion problem and then extend image-space WETom from the isotropic medium to the anisotropic medium. We show that theoretically the gradient of the tomographic objective functional for the anisotropic medium is similar to its isotropic version, with an extra term for the additional parameter. Then, we test the anisotropic WETom operator using a model with a localized anomaly. Finally, we invert for a 2-D VTI model using the proposed anisotropic WETom operator.

PARAMETERIZATION

In the VTI medium, Thomsen parameters ϵ and δ are commonly used to characterize the anisotropic seismic velocity. These two parameters define the relationships between the vertical velocity (V_V), the horizontal velocity (V_H), and the NMO velocity (V_N) as follows:

$$V_H^2 = V_V^2(1 + 2\epsilon), \quad (1)$$

$$V_N^2 = V_V^2(1 + 2\delta). \quad (2)$$

In the practice of surface seismic exploration, it is impossible to estimate the vertical velocity because depth of the reflectors is unknown, and there is no vertical offset information in the data. However, if we have long enough in-line and cross-line offsets, it may be possible to resolve the horizontal velocity and the NMO velocity. Therefore, the anellipticity parameter η is used to provide a direct link between V_N and V_H :

$$V_H^2 = V_N^2(1 + 2\eta), \quad (3)$$

where η is defined by the Thomsen parameters as follows:

$$\eta = \frac{\epsilon - \delta}{(1 + 2\delta)}. \quad (4)$$

To reduce the number of parameters, and thereby the null space of the resulting inversion procedure, we make an arbitrary assumption that $\delta = 0$. Hence, there are only two independent parameters:

$$\eta = \epsilon \quad (5)$$

and

$$V_V = V_N. \quad (6)$$

Therefore, we choose to use V_V and η as the model parameters that we will estimate during the inversion.

WETOM FOR ANISOTROPIC PARAMETERS

Anisotropic WETom is a non-linear inversion process that aims to find the anisotropic model that minimizes the residual field $\Delta\mathbf{I}$ in the image space. The residual image is derived from the background image \mathbf{I} , which is computed with current background model. In general, the residual image is defined as (Biondi, 2008)

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

where \mathbf{F} is a focusing operator acting on the background image.

In the least-squares sense, the tomographic objective function can be written as follows:

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

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

$$\begin{aligned} \mathbf{T} &= \left. \frac{\partial \mathbf{I}}{\partial \mathbf{m}} \right|_{\mathbf{m}=\hat{\mathbf{m}}} \\ &= \left. \frac{\partial \mathbf{I}}{\partial \mathbf{s}} \right|_{\mathbf{s}=\hat{\mathbf{s}}} + \left. \frac{\partial \mathbf{I}}{\partial \eta} \right|_{\eta=\hat{\eta}}, \end{aligned} \quad (9)$$

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

$$\Delta\mathbf{I} = \mathbf{T}\Delta\mathbf{m}. \quad (10)$$

In this paper, we evaluate the anisotropic tomographic operator in the shot-profile domain.

Both source and receiver wavefields are downward continued in the shot-profile domain using the one-way wave equations (Claerbout, 1971):

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

and

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

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

is the recorded shot gather at \mathbf{x}_s , which serves as the boundary condition of Equation 12. Operator Λ is the dispersion relationship for anisotropic wave propagation:

$$\Lambda = \omega s(\mathbf{x}) \sqrt{1 - \frac{|\mathbf{k}|^2}{\omega^2 s^2(\mathbf{x}) - 2\eta(\mathbf{x})|\mathbf{k}|^2}}, \quad (13)$$

where ω is the angular frequency, $s(\mathbf{x})$ is the slowness at \mathbf{x} , $\eta(\mathbf{x})$ is the anellipticity at \mathbf{x} ; $\mathbf{k} = (k_x, k_y)$ is the spatial wavenumber vector. Dispersion relationship 13 can be approximated with a rational function by Taylor series and Padé expansion analysis (Shan, 2009):

$$\Lambda = \omega s(\mathbf{x}) \left(1 - \frac{a|\mathbf{k}|^2}{\omega^2 s^2(\mathbf{x}) - b|\mathbf{k}|^2}\right) \quad (14)$$

where, to the second order, $a = 0.5$, $b = 2\eta + 0.25$. Using binomial expansion, Equation 14 can be further expanded to polynomials:

$$\Lambda = \omega s(\mathbf{x}) - \frac{a}{\omega s^2(\mathbf{x})} |\mathbf{k}|^2 - \frac{3ab}{\omega^3 s^4(\mathbf{x})} |\mathbf{k}|^4. \quad (15)$$

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

$$I(\mathbf{x}, \mathbf{h}) = \sum_{\mathbf{x}_s} \sum_{\omega} \overline{D(\mathbf{x} - \mathbf{h}, \mathbf{x}_s)} U(\mathbf{x} + \mathbf{h}, \mathbf{x}_s), \quad (16)$$

where the overline stands for the complex conjugate, and $\mathbf{h} = (h_x, h_y, h_z)$ is the subsurface half-offset.

Under the Born approximation, a perturbation in the model parameters causes a first-order perturbation in the wavefields. Consequently, the resulting image perturbation reads:

$$\Delta I(\mathbf{x}, \mathbf{h}) = \sum_{\mathbf{x}_s} \sum_{\omega} \left(\overline{\Delta D(\mathbf{x} - \mathbf{h}, \mathbf{x}_s)} \widehat{U}(\mathbf{x} + \mathbf{h}, \mathbf{x}_s) + \overline{\widehat{D}(\mathbf{x} - \mathbf{h}, \mathbf{x}_s)} \Delta U(\mathbf{x} + \mathbf{h}, \mathbf{x}_s) \right), \quad (17)$$

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

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

$$\begin{aligned} \Delta D(\mathbf{x}, \mathbf{x}_s) &= \sum_{\mathbf{h}} \Delta I(\mathbf{x}, \mathbf{h}) \widehat{U}(\mathbf{x} + \mathbf{h}, \mathbf{x}_s) \\ \Delta U(\mathbf{x}, \mathbf{x}_s) &= \sum_{\mathbf{h}} \Delta I(\mathbf{x}, \mathbf{h}) \widehat{D}(\mathbf{x} - \mathbf{h}, \mathbf{x}_s). \end{aligned} \quad (18)$$

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

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

and

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\Lambda \right) \Delta U(\mathbf{x}, \mathbf{x}_s) = \left(-i \frac{\partial \Lambda}{\partial \mathbf{m}} \widehat{U}(\mathbf{x}, \mathbf{x}_s) \right) \Delta \mathbf{m}^*(\mathbf{x}) \\ \Delta U(x, y, z = 0, \mathbf{x}_s) = 0 \end{cases}, \quad (20)$$

where \mathbf{m} is the row vector $[\mathbf{s} \ \eta]$, and \mathbf{m}^* is the transpose of \mathbf{m} .

When solving the optimization problem, we obtain the image perturbation by migrating the data with the current background model and performing a focusing operation (Equation 7). Then the perturbed image is convolved with the background wavefields to get the perturbed wavefields (Equation 18). The scattered wavefields are computed by applying the adjoint of the one-way wave-equations 19 and 20. Finally, the model space gradient is obtained by cross-correlating the upward propagated scattered wavefields with the modified background wavefields (terms in the parentheses on the right-hand sides of Equations 19 and 20).

NUMERICAL TEST OF THE ANISOTROPIC WETOM OPERATOR

To test the anisotropic WETom operator, we run the forward and adjoint WETom operator on a 2-D model. Figure 1 shows the background isotropic model, with one reflector in velocity and no anisotropy. The data are modeled with 4000m maximum offset, 8m receiver spacing, 80m source spacing and 41 split-spread shots. We use the two-way acoustic anisotropic modeling code in Madagascar to do the modeling, and the one-way SSF (Tang and Clapp, 2006) extrapolator to do the migration.

Figure 2 shows the model perturbations, with a rectangular slowness anomaly that is 10% lower than the background slowness on the left, and a rectangular anisotropic anomaly on the right. The perturbation in η within the rectangular block is constant ($\Delta\eta = 0.1$). Figure 3 shows the perturbed image at the zero lag of the subsurface offset due to the model perturbations after applying the forward WETom operator. Adjoint WETom operator back-projects the perturbed image into the model space, and outputs the gradient for the model perturbation, as shown in Figure 4. Comparing Figure 2 and Figure 4, we can see that the gradients provide the correct direction and shape of the perturbation to conduct a line search in a given inversion scheme.

INVERSION FOR 2-D VTI MODELS

From the gradient given by the anisotropic WETom operator in last section, we notice that the perturbations in slowness and η are co-located. This is an intrinsic characteristic of the operator, which may not be geologically realistic. Therefore, we design three tests to examine this effect in the inversion.

In these three tests, we use the same initial model (Figure 5), but data modeled using different true models. In Figure 6 we show the true model perturbations, which have one layer of perturbations in slowness only, in η only, and in both. Figure 7 shows the angle-domain common image gathers (ADCIGs) using the initial model, where ADCIGs are not flat due to the error in the model. Notice the image perturbation is small for the perturbation in η only.

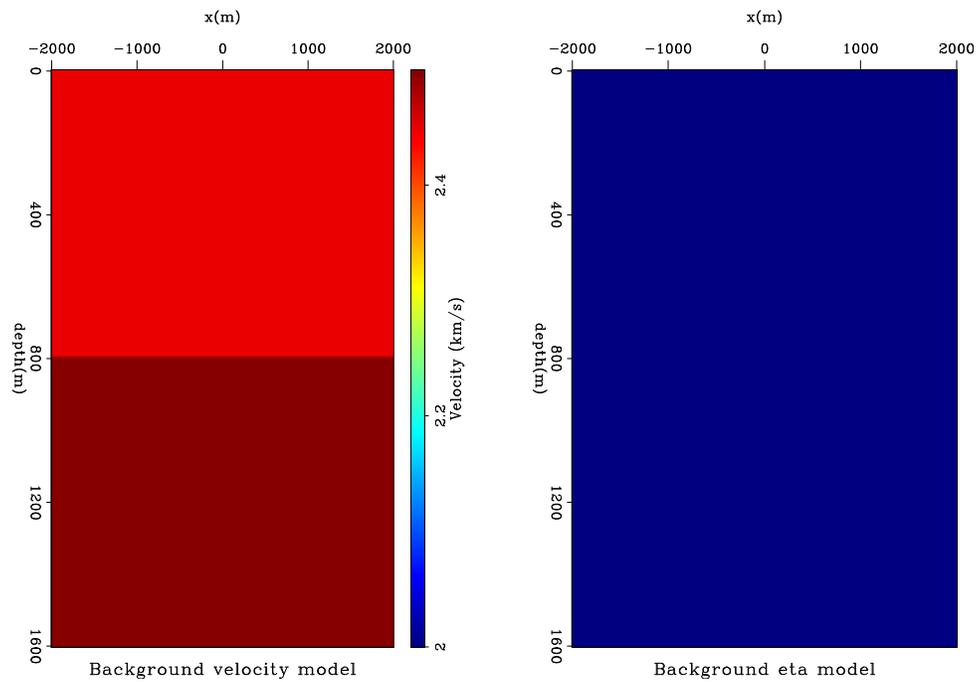


Figure 1: Background isotropic model. Left is the velocity model with one reflector, and right is the η model with constant zero. [ER] elita2/. background

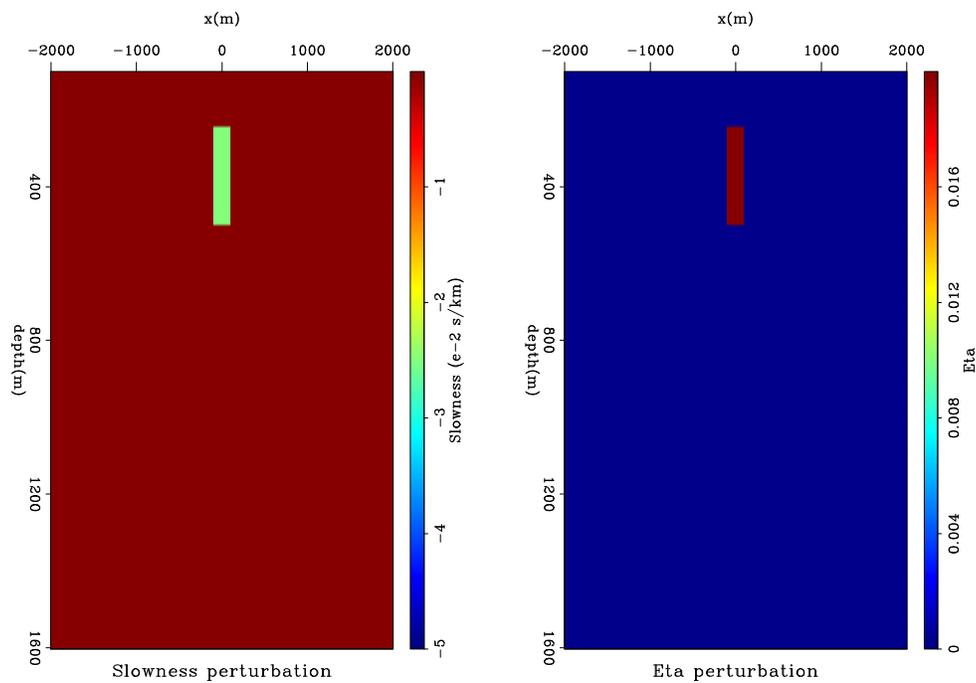


Figure 2: Model perturbations. Left is a rectangular slowness anomaly that is 10% lower than the background slowness, and right is a rectangular anisotropic anomaly with a constant value of $\Delta\eta = 0.1$. [ER] elita2/. perturb

Figure 3: Perturbed image from the forward anisotropic WETom operator. The image is extracted from the zero lag of the subsurface offset. [ER] `elita2/. Dimage`

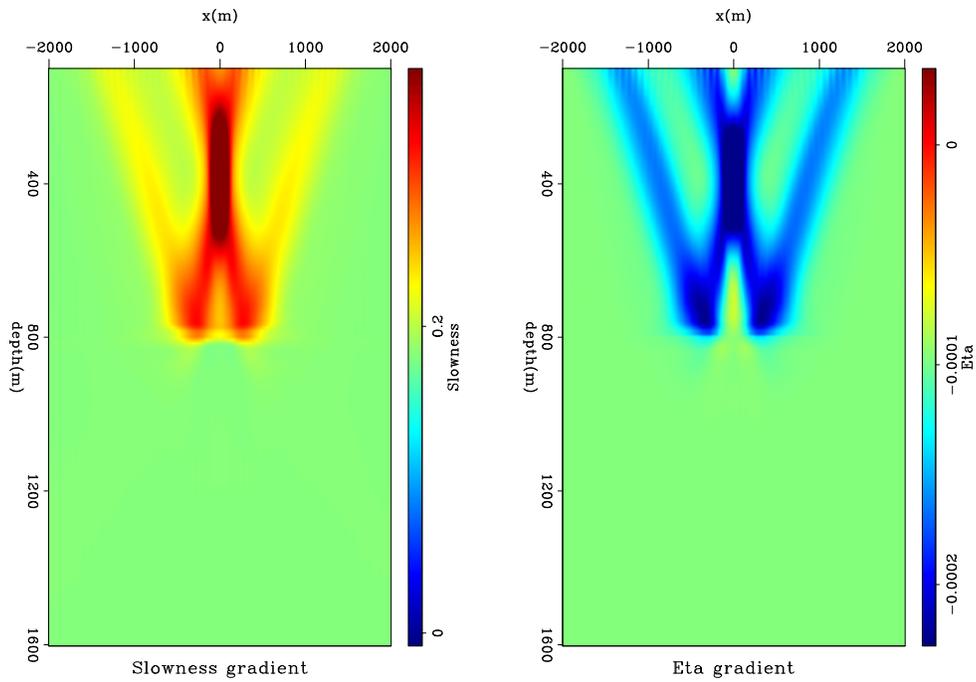
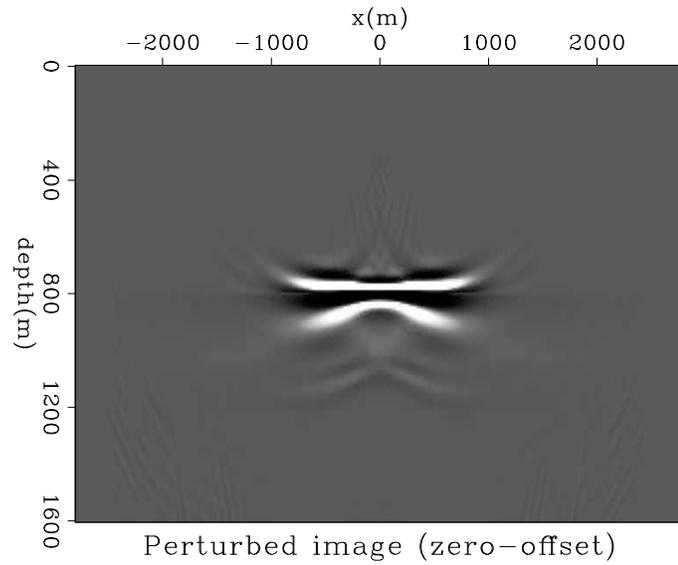


Figure 4: Back-resolved gradient for the model updates. Left is the gradient for slowness, and right is the gradient for η . [ER] `elita2/. Dmodel`

In the inversion, we define the focusing operator in the perturbed image (Equation 7) by the Differential Semblance Optimization (DSO) method (Shen, 2004):

$$\mathbf{F}(\mathbf{I}) = (\mathbf{1} - \mathbf{O})\mathbf{I}, \quad (21)$$

where $\mathbf{1}$ is the identity operator and \mathbf{O} is the DSO operator. Therefore, the objective function (Equation 8) becomes:

$$J = \frac{1}{2} \|\mathbf{O}\hat{\mathbf{I}}\|^2 = \frac{1}{2} \|h\hat{\mathbf{I}}\|^2, \quad (22)$$

where h is the subsurface half-offset. Since the DSO operator is independent of the model parameters, the gradient of J with respect to the model parameters is

$$\nabla J = \left(\frac{\partial \mathbf{I}}{\partial \mathbf{m}} \Big|_{\mathbf{m}=\hat{\mathbf{m}}} \right)^* \mathbf{O}^* \mathbf{O} \hat{\mathbf{I}} = \mathbf{T}^* \mathbf{O}^* \mathbf{O} \hat{\mathbf{I}}. \quad (23)$$

To minimize the objective function, we specifically use the steepest descent algorithm. To help convergence, we average the gradient at each depth to ensure a layered model and mute the shallow updates to avoid near-surface artifacts.

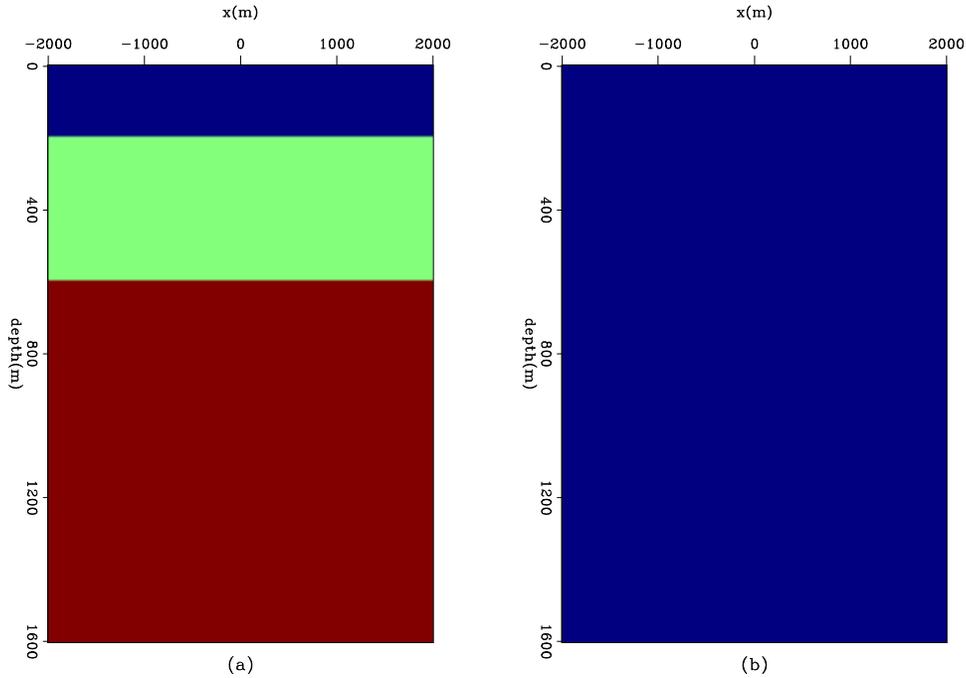


Figure 5: Initial model for inversion. Panel (a) is the initial velocity model with three layers; Panel (b) is the initial η model with constant value of zero. [ER] `elita2/. init`

Figure 8 shows the final model updates after 4 non-linear iterations. The results should be comparable to the model perturbations in Figure 6. In the final updates, we see a consistent over prediction of η . This is because error in η has a very small contribution in the image perturbation, as shown in the middle panel of Figure 7. Figure 9 shows the ADCIGs using the updated model. Comparing Figure 6 and Figure 8, Figure 7 with Figure 9, we can conclude that the inversion successfully identifies the layered perturbation and

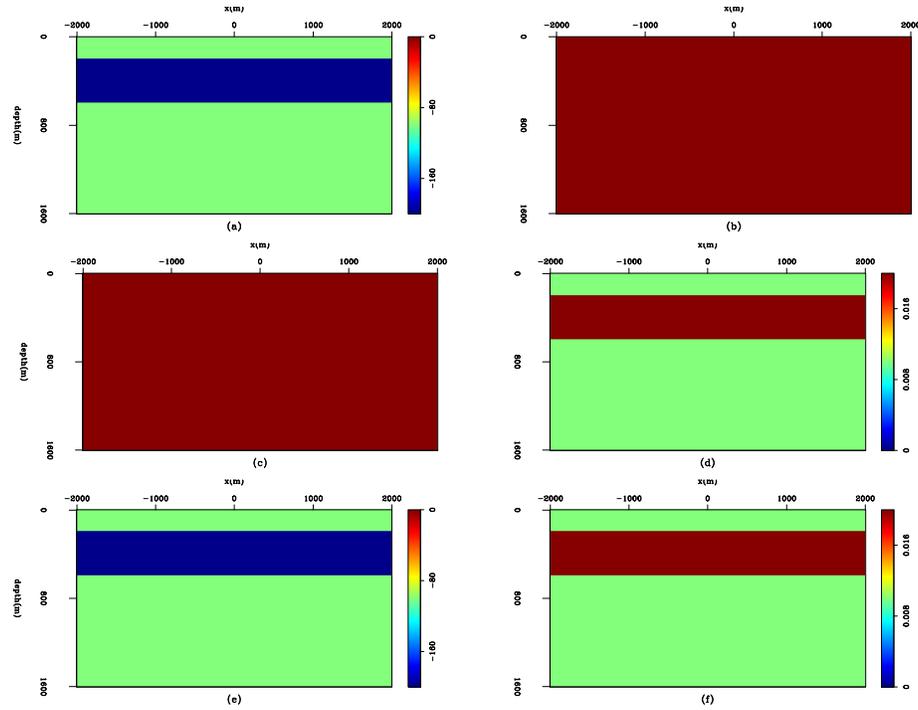


Figure 6: True model perturbation in three test cases. Panel (a) and (b): Perturbation in velocity only; Panel (c) and Panel (d): Perturbation in η only; Panel (e) and Panel (f): Perturbations in both. [ER] `elita2/. truep`

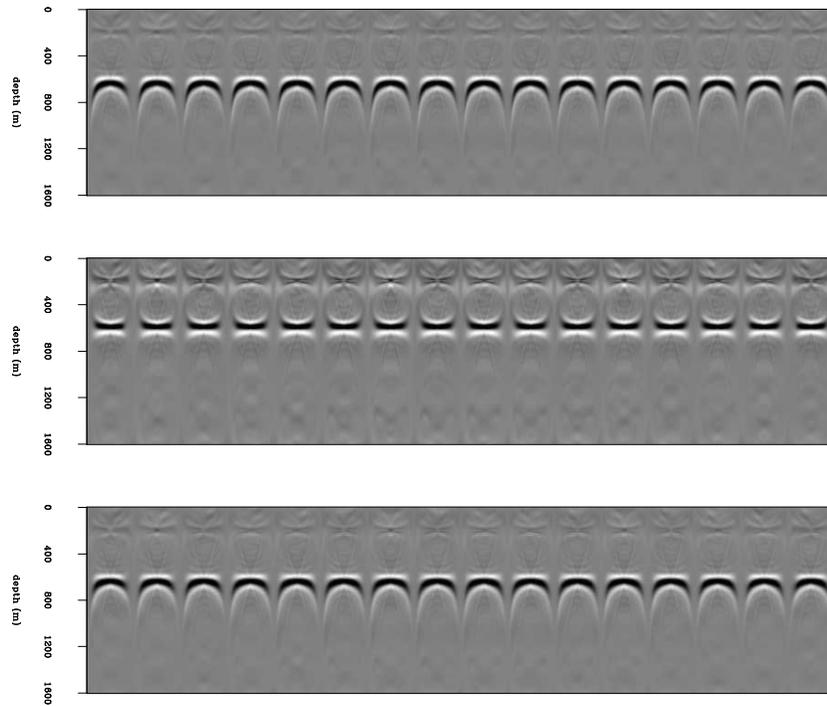


Figure 7: ADCIGs using the initial model in three cases. Top panel: Perturbation in velocity only; Middle panel: Perturbation in η only; Bottom panel: Perturbations in both. Curvature in the ADCIGs indicates errors in velocity or η . [ER] `elita2/. initimage`

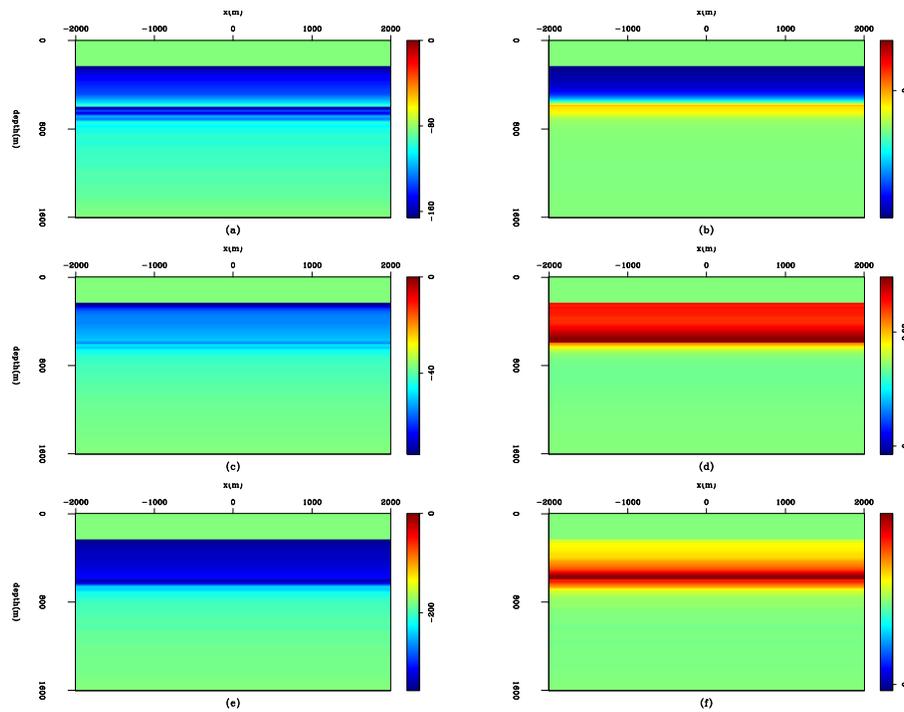


Figure 8: Inversion results of the three test cases. Panels are comparable to those in Figure 6. [CR] `elita2/.update`

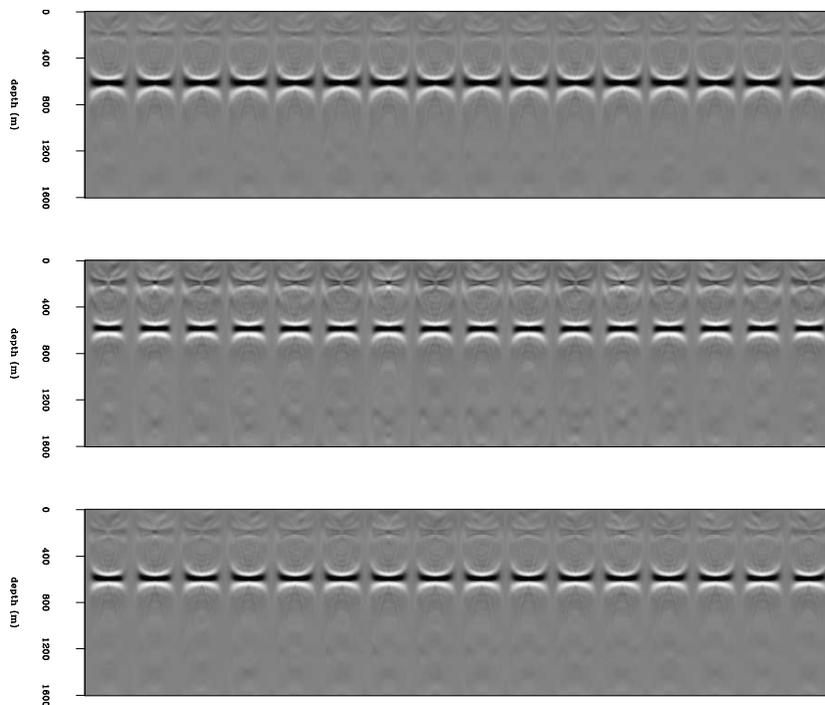


Figure 9: Common image gathers using the updated model in three cases. Top panel: Perturbation in velocity only; Middle panel: Perturbation in η only; Bottom panel: Perturbations in both. ADCIGs are flattened compared with Figure 7. [ER] `elita2/.finalimage`

flattens the ADCIGs. However, back-projection of the residual image shows up in both parameter spaces, although it is caused by perturbations in a single parameter. Also for the case of perturbing both spaces, the over-prediction of velocity perturbation and the under-prediction of η perturbation reconcile with each other and produce flat events in the ADCIGs. Therefore, we can conclude that the ambiguity between the velocity and the anellipticity cannot be resolved simply by the inversion, and auxiliary information is needed to further distinguish the difference.

CONCLUSIONS AND DISCUSSION

In this paper, we present a novel method to build the anisotropic velocity model using image-space wave-equation tomography. The mathematical formulation of the operator shows that by adding an additional term for η , the gradient of the anisotropic WETom shows similar structure as that for the isotropic case. Our numerical tests show that the anisotropic WETom operator is useful in identifying both localized and layered perturbation in simple cases. However, when only one parameter is perturbed, the WETom operator produces a model with perturbations in both parameters.

One possible way to resolve this ambiguity is to utilize the moveout information in the subsurface offset domain and/or angle domain. It has been shown that the RMO functions are different in ADCIGs for isotropic perturbations (Biondi and Symes, 2004) and anisotropic perturbations (Biondi, 2007). Therefore, adding the RMO information may help to distinguish the contributions from different parameters. Also, data other than surface seismic data, such as well logs and checkshots, can also be helpful in further constraining the problem. This will be investigated in the near future.

ACKNOWLEDGMENT

We thank James Berryman for fruitful discussion about the theory and the choice of the synthetic model.

REFERENCES

- Biondi, B., 2007, Residual moveout in anisotropic angle-domain common-image gathers: *Geophysics*, **72**, S93S103.
- , 2008, Automatic wave-equation migration velocity analysis: **SEP-134**, 65–78.
- Biondi, B. and W. Symes, 2004, Angle-domain common- image gathers for migration velocity analysis by wavefield- continuation imaging: *Geophysics*, **69**, 12831298.
- Cai, J., Y. He, Z. Li, B. Wang, and M. Guo, 2009, TTI/VTI anisotropy parameters estimation by focusing analysis, Part I: Theory: *SEG Expanded Abstracts*, **28**.
- Claerbout, J. F., 1971, Towards a unified theory of reflector mapping: *Geophysics*, **36**, 467–481.
- Fei, T. W. and C. L. Liner, 2008, Hybrid Fourier finite-difference 3D depth migration for anisotropic media: *Geophysics*, **73**, S27.
- Fletcher, R., X. Du, and P. J. Fowler, 2009, Stabilizing acoustic reverse-time migration in TTI media: *SEG Expanded Abstracts*, **28**.

- Guerra, C., Y. Tang, and B. Biondi, 2009, Wave-equation tomography using image-space phase-encoded data: SEP-report, **138**, 95.
- Helbig, K., 1956, Die ausbreitung elastischer Wellen in anisotropen Medien: Geophys. Prosp., **04**, 70–81.
- McCollum, B. and F. Snell, 1932, Asymmetry of sound velocity in stratified formations: Physics – A Journal of General and Applied Physics, **2**, 174–185.
- Postma, G. W., 1955, Wave propagation in a stratified medium: Geophysics, **20**, 780–806.
- Sava, P., 2004, Migration and velocity analysis by wavefield extrapolation: PhD thesis, Stanford University.
- Sava, P. and B. Biondi, 2004a, Wave-equation migration velocity analysis-I: Theory: Geophysical Prospecting, **52**, 593–606.
- , 2004b, Wave-equation migration velocity analysis-II: Examples: Geophysical Prospecting, **52**, 607–623.
- Shan, G., 2009, Optimized implicit finite-difference and Fourier finite-difference migration for VTI media: Geophysics, WCA189WCA197.
- Shen, P., 2004, Wave-equation Migration Velocity Analysis by Differential Semblance Optimization: PhD thesis, Rice University.
- Shen, P. and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.
- Tang, Y. and R. G. Clapp, 2006, Selection of reference anisotropic parameters for wavefield extrapolation: SEP-report, **124**.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.
- Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, 15–26.
- Yuan, J., X. Ma, S. Lin, and D. Lowrey, 2006, P-wave tomographic velocity updating in 3D inhomogeneous VTI media: SEG Expanded Abstracts, **25**.
- Zhang, Y. and H. Zhang, 2009, A stable TTI reverse time migration and its implementation: SEG Expanded Abstracts, **28**.
- Zhou, H., D. Pham, and S. Gray, 2004, Tomographic velocity analysis in strongly anisotropic TTI media: SEG Expanded Abstracts, **23**.
- Zhou, H., D. Pham, S. Gray, and B. Wang, 2003, 3-D tomographic velocity analysis in transversely isotropic media: SEG Expanded Abstracts, **22**.

Wave-equation traveltine tomography by global optimization

Ali Ameen Almomin

ABSTRACT

Wave-equation traveltine tomography is conventionally done by picking the maximum cross-correlation lags between the modeled and observed data. However, a trace-by-trace method of picking makes the velocity update more susceptible to local noise in the correlation as well as inconsistencies in the data. In this paper, I compare the local method of picking the maximum correlation to a global method based on maximizing the stacking power along an interpolated spline surface in the correlation window. The results show that the global scheme is more robust to local noise but sacrifices accuracy and convergence rate.

INTRODUCTION

Conventional full waveform inversion (FWI), which was first introduced by Tarantola (1984), has an objective function that is highly non-linear. The forward operator is linearized around the background velocity, which makes the initial model a determining factor for the convergence of the inversion. Therefore, a lot of previous work (Luo and Schuster, 1990; Symes and Carazzone, 1991; Biondi and Sava, 1999; Shen, 2004; Biondi, 2009) has focused on finding more tractable objective functions that have stronger dependence on the kinematics of the wavefield than on the amplitude of the waveform. One attractive method that uses such an objective function is wave-equation traveltine inversion (WT), which was first introduced by Luo and Schuster (1990). In this inversion, the objective function depends on the lag of maximum cross-correlation between the observed and modeled data. Conventionally, these lags are picked in a trace-by-trace scheme, which produce errors due to correlating noise, multiple events, and inconsistencies in the observed data. To overcome this problem, I cast the picking procedure as a global optimization problem in order to avoid local errors by making use of the redundancy of the data.

THEORY

The objective function of FWI can be written as follows:

$$\mathbf{J}_{FWI}(\mathbf{v}) = \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_g} \|\mathbf{d}_{cal}(t, \mathbf{x}_g, \mathbf{x}_s; \mathbf{v}) - \mathbf{d}_{obs}(t, \mathbf{x}_g, \mathbf{x}_s)\|_2^2, \quad (1)$$

where \mathbf{x}_s and \mathbf{x}_r are the source and receiver locations, d_{cal} is the modeled data with velocity \mathbf{v} , and d_{obs} is the observed data. By setting the first derivative of equation (1) around the velocity \mathbf{v}_0 to zero, the velocity update can be expressed as follows:

$$\Delta \mathbf{v} = \frac{-s}{\mathbf{v}_0^3} \sum_{\mathbf{x}_s} \sum_t \mathbf{L}(\mathbf{v}_0) \mathbf{S}(\mathbf{x}_s, t) \frac{\partial^2}{\partial t^2} \mathbf{L}^\dagger(\mathbf{v}_0) (\mathbf{d}_{cal}(t, \mathbf{x}_g, \mathbf{x}_s; \mathbf{v}) - \mathbf{d}_{obs}(t, \mathbf{x}_g, \mathbf{x}_s)), \quad (2)$$

where s is the step size, \mathbf{S} is the source signature, and \mathbf{L} and \mathbf{L}^\dagger are the forward wave propagation operator and its adjoint, respectively.

The objective function wave-equation traveltime tomography can be written as follows:

$$\mathbf{J}_{\Delta\tau}(\mathbf{v}) = \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_g} \|\Delta\tau(\mathbf{x}_g, \mathbf{x}_s; \mathbf{v})\|_2^2, \quad (3)$$

where $\Delta\tau$ is the lag of the maximum cross-correlation between the observed data and the data modeled by a velocity model \mathbf{v} . Again, the first derivative of equation (3) around the lags $\Delta\tau$ is set to zero to get the velocity update, which can be expressed as follows:

$$\Delta\mathbf{v} = \frac{s}{\mathbf{v}_0^3} \sum_{\mathbf{x}_s} \sum_t \mathbf{L}(\mathbf{v}_0)\mathbf{S}(\mathbf{x}_s, t) \frac{\partial^2}{\partial t^2} \mathbf{L}^\dagger(\mathbf{v}_0) \frac{\Delta\tau}{\xi} \frac{\partial}{\partial t} \mathbf{d}_{obs}(t + \Delta\tau, \mathbf{x}_g, \mathbf{x}_s), \quad (4)$$

where ξ is defined as follows:

$$\xi = \sum_t \frac{\partial}{\partial t} \mathbf{d}_{cal}(t, \mathbf{x}_g, \mathbf{x}_s; \mathbf{v}_0) \frac{\partial}{\partial t} \mathbf{d}_{obs}(t + \Delta\tau, \mathbf{x}_g, \mathbf{x}_s), \quad (5)$$

By examining equations (2) and (4), it can be shown that (WT) can handle much larger velocity errors than (FWI).

Now, I cast the picking procedure of the lags $\Delta\tau$ as a global optimization problem with an objective function as follows:

$$\mathbf{C}(\Delta\tau) = \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_g} f(\mathbf{A}\Delta\tau(\mathbf{y}_g, \mathbf{y}_s)), \quad (6)$$

where \mathbf{y}_g and \mathbf{y}_s are a sparse representation of the source and receiver locations, \mathbf{A} is a bicubic spline interpolation operator that maps the sparse coordinates \mathbf{y}_g and \mathbf{y}_s to the original coordinates \mathbf{x}_g and \mathbf{x}_s , and f evaluate the correlation value at $\Delta\tau(\mathbf{x}_g, \mathbf{x}_s)$.

The goal of the global optimization is to maximize the function described by equation (6), which is to maximize the stacking power along the interpolated spline surface. The searching procedure is a simulated annealing algorithm, which varies the spline points along the time axis in a stochastic sense until a satisfying solution is reached. In the following section, I show the results of using such global scheme to pick the correlation lags.

SYNTHETIC EXAMPLE

I use the velocity model shown in Figure 1 to create the synthetic data. The sampling interval of the model is 25 m for both the x-axis and z-axis. The data is modeled using a Ricker wavelet with a fundamental frequency of 15 Hz and a sampling interval of 3 ms. The acquisition geometry is a cross-well configuration, where sources are located in a vertical well on the left side of the model and the receivers are located in a vertical well on the right side of the model. Both the sources and the receivers start at the surface and cover the full depth of the model with a sampling interval of 25 m.

I start the velocity inversion with a constant background velocity of 2.9 km/s, which is very far from the true velocity model. After modeling the data with the background

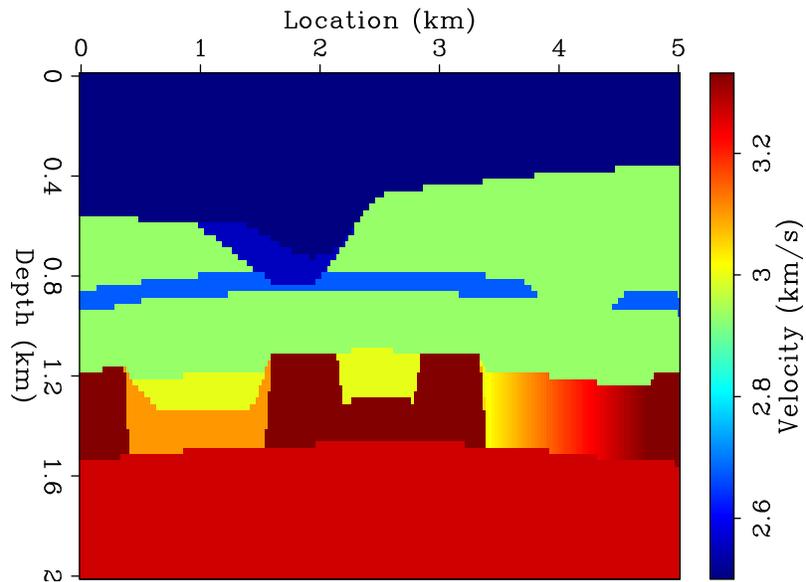


Figure 1: The true velocity model used to create the data. [ER] `ali2/. velali`

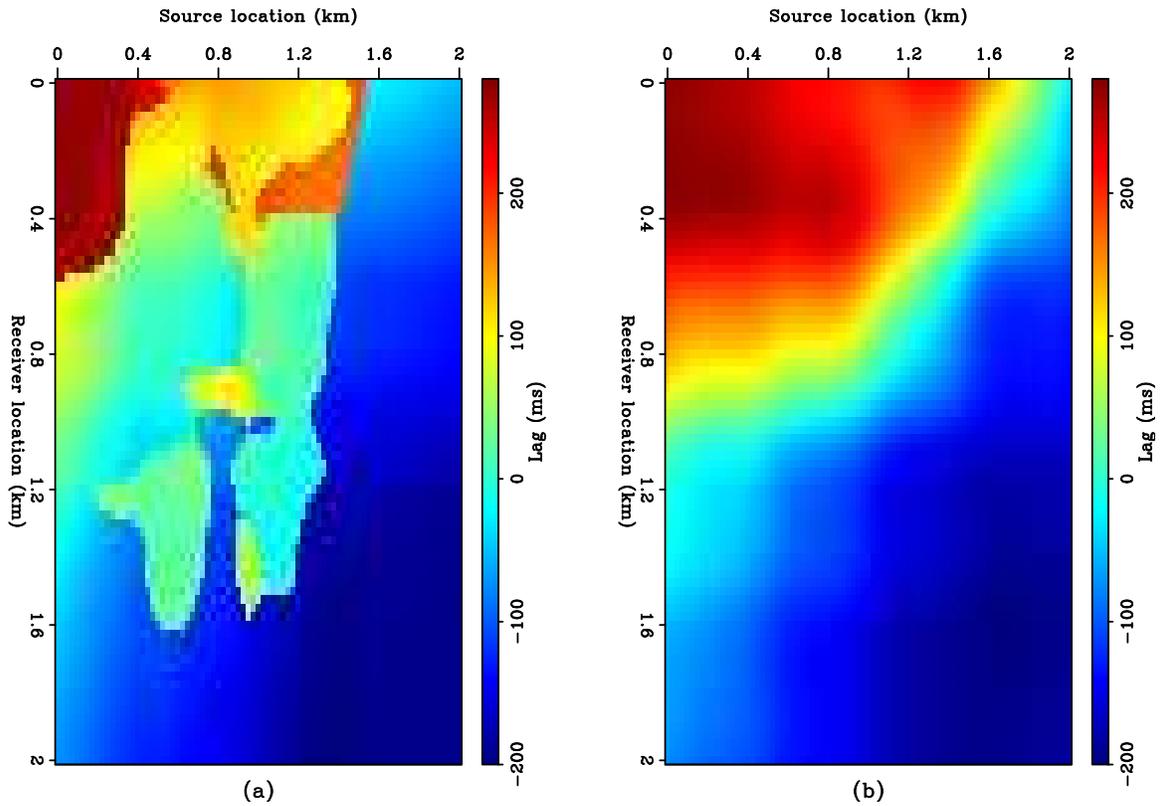


Figure 2: The correlation lags picked by: (a) the trace-by-trace maximum correlation method, and (b) the global spline-fitting method. [ER] `ali2/. lags`

velocity, I cross-correlate the modeled data with the observed data. Figure 2(a) shows the lags picked by maximum correlation at each trace. There is an overall trend from the top left corner to the bottom right corner, which is caused by the gradient in the original data. However, there is also some large anomalies in the picked lags with sharp discontinuities around them. These anomalies are caused by the events refracted by the large velocity contrasts in Figure 1. Figure 2(b) shows the lags picked by the global method, which are significantly different from those in Figure 2(a). The global algorithm ignored the local maximum caused by refracted energy and picked lags that are more consistent with their surroundings.

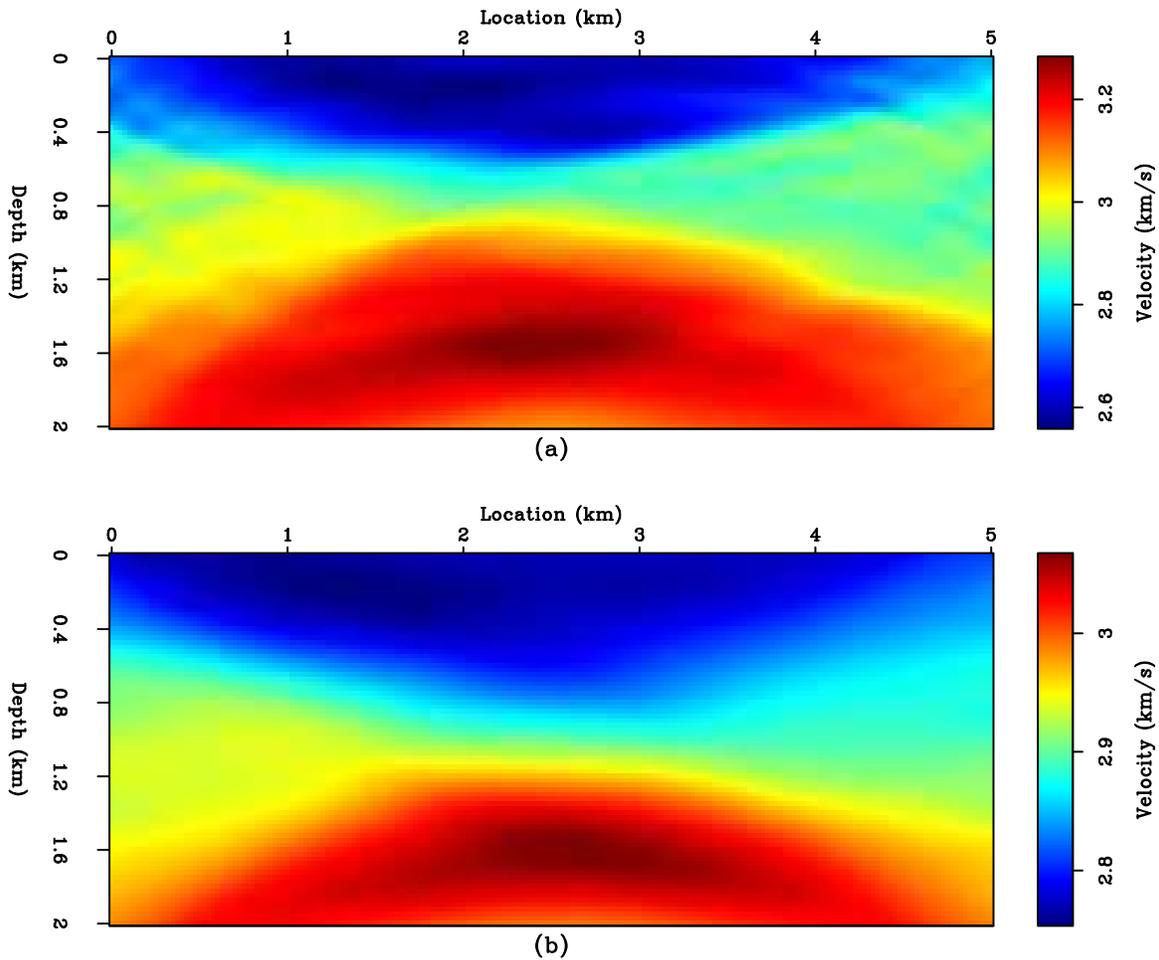


Figure 3: The velocity model after 1 iteration using: (a) the trace-by-trace maximum correlation method, and (b) the global spline-fitting method. [ER] ali2/. tomo-first

Now, I use the lags estimated by both methods to find a velocity update. The scale of the update is estimated by performing a line search. Figure 3 shows the updated velocity model using both methods after one iteration. The results of the local method shows some noise in the velocity model update, especially around the edges. However, since the velocity update is the sum of the updates of all the shots, the total results are satisfactory. On the other hand, the velocity model estimated by the global method looks cleaner and more consistent laterally. However, the velocity update by the global method shows less detail

than that of the local method. In addition, the global method converges at a slower rate than the local method.

Finally, I run both inversions for 10 iterations to further show the difference between the two methods. The results of both inversions are shown in Figure 4. The noise in the local method estimate grows even larger as I run more iterations, but the accuracy seems to be consistent with the true velocity model. On the other hand, the global method is still too smooth and did not pick the small details as well as the local method. Finally, there seems to be some bias in the global method toward lower velocities.

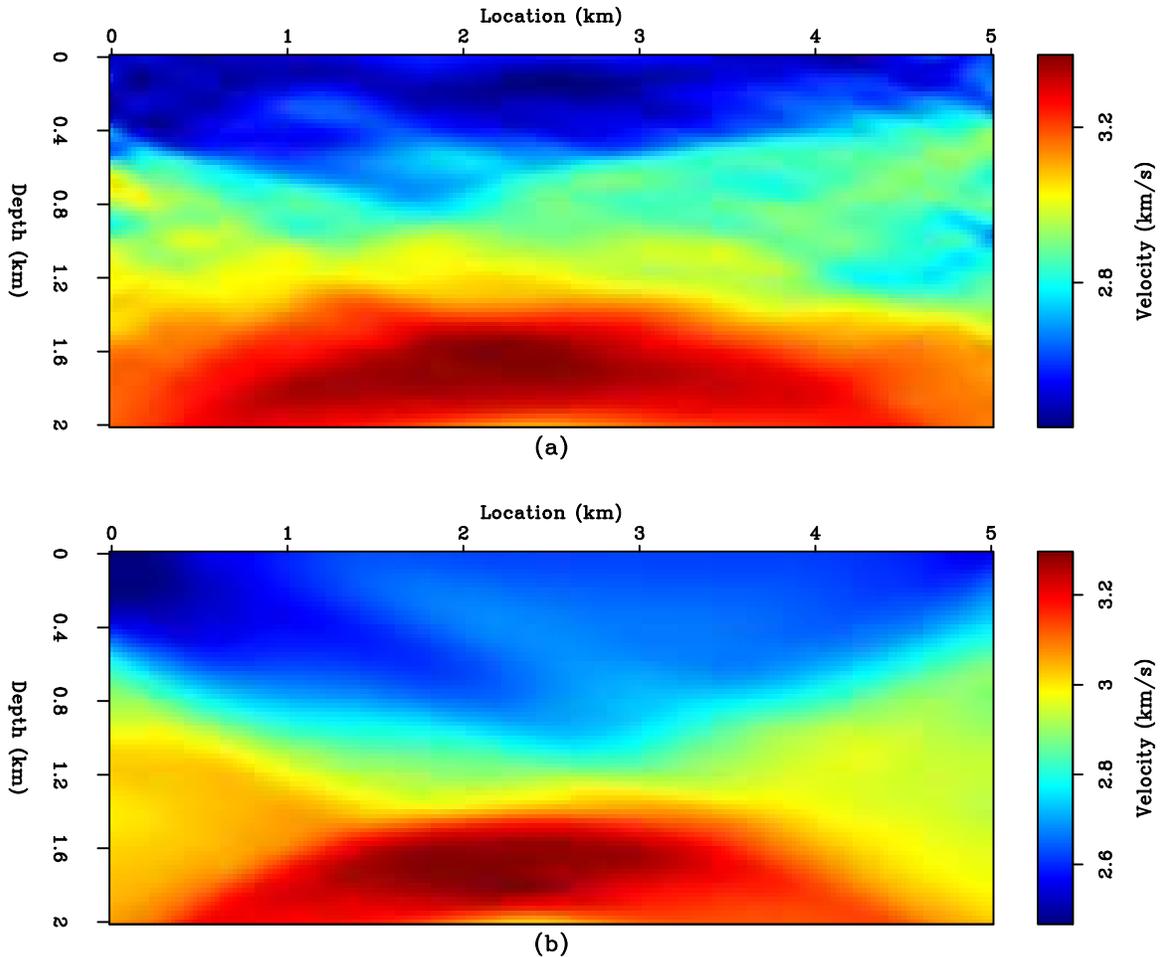


Figure 4: The inverted velocity model after 10 iterations using: (a) the trace-by-trace maximum correlation, and (b) the global spline-fitting method. [ER] ali2/. tomo-last

CONCLUSIONS AND DISCUSSIONS

We showed that using a global scheme for picking the correlation lags can better detect the correct events and is more robust to noise and multiple events than a trace-by-trace method. However, since the global picking is done on sparse spline points, the picking surface is too smooth, which causes the result to have less detail than that of the local method and converges at a lower rate. Nevertheless, the estimated velocity model did correctly approach

the true velocity model, even though we started with a constant velocity.

Moreover, the accuracy and smoothness of the global method is strongly influenced by the number of spline points specified by the user. Therefore, accurately estimating the smoothness of the picked cross-correlation surface can help in determining a good spline geometry.

FUTURE WORK

As we have seen, the global optimization is less accurate than the local fitting. One way to improve the results is to use a hybrid optimization method in which the algorithm either alternates between a global and a local iteration or runs a global search, followed by a local one. In addition, all testing has been done on transmission seismology. The next step is to extend the method to reflection seismology.

Finally, the tests case in this paper are synthetic data. In field data, the noise level is an issue that might degrade the local method even more, in which case the global method can show even more improvement.

ACKNOWLEDGMENTS

We thank Biondo Biondi, Robert Clapp and Xukai Shen for their valuable discussions and help throughout this project.

REFERENCES

- Biondi, B., 2009, Measuring image focusing for velocity analysis: SEP-Report, **138**, 59–79.
- Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **18**, 1723–1726.
- Luo, Y. and G. T. Schuster, 1990, Wave-equation traveltime inversion: SEG Expanded Abstracts.
- Shen, P., 2004, Wave equation migration velocity analysis by differential semblance optimization: PhD thesis, Rice University.
- Symes, W. W. and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: *Geophysics*, **56**, 654–663.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266.

Angle-domain illumination gathers by wave-equation-based methods

Yaxun Tang and Biondo Biondi

ABSTRACT

We present a method for computing the wave-equation-based angle-domain illumination for subsurface structures. It creates subsurface illumination for different scattering or/and dip angles for a given acquisition geometry, velocity model and frequency bandwidth. The proposed method differs from the conventional method in that it does not require local plane-wave decompositions for each source and receiver Green's functions. Instead, it transforms a precomputed subsurface-offset-indexed sensitivity kernel into angle domain using either a Fourier-domain mapping or a space-domain slant stack. We show that the computational cost can be significantly reduced by phase encoding the receiver-side Green's functions, or by simultaneously encoding both the source- and receiver-side Green's functions. Numerical examples demonstrate the accuracy and efficiency of our method. The main anticipated applications of our method are in areas of: (1) accurate amplitude-versus-angle (AVA) analysis by compensating depth-migrated images with angle-dependent illumination, (2) migration velocity analysis that incorporates angle-dependent illumination for robust residual parameter estimation, and (3) optimum seismic survey planning.

INTRODUCTION

Seismic image quality obtained using prestack depth migration is highly dependent on subsurface illumination, which can be affected by many factors, e.g., the limited acquisition geometry, the complex velocity model, and the frequency content of seismic waves. Uneven illumination causes distorted seismic images due to the fact that migration operator is nonunitary (Nemeth et al., 1999; Clapp, 2005; Valenciano, 2008; Tang, 2009). Seismic illumination analysis quantifies such image distortion and provides ways to optimally restore the reflectivity. More accurate amplitude-versus-angle (AVA) analysis can be obtained if the angle-domain common-image gather is corrected with angle-dependent illumination. Angle-domain velocity analysis may also benefit from angle-dependent illumination weighting. It can weight up well-illuminated angular images, while weight down poorly-illuminated ones for robust residual parameter estimation, reducing the uncertainty caused by poor illumination.

Angle-domain illumination analysis is often carried out using ray-based method due to its efficiency (Schneider and Winbow, 1999; Bear et al., 2000). However, the infinite-frequency approximation and the caustics inherent in ray theory prevent ray-based method from accurately modeling complex wave phenomena. Wave-equation-based method, on the other hand, more accurately describes bandlimited wave phenomena, and therefore more suitable for complex geologies, e.g., subsalt regions with complex overburdens. Wave-equation-based angle-domain illumination analysis is proposed by Xie et al. (2006), where

local plane-wave decomposition has been employed to extract the directional information for both source and receiver Green's functions before spatial and temporal correlations.

In this paper, we present an alternative way of computing the angle-dependent illumination. Instead of performing local plane-wave decomposition for each Green's function, our method transforms a subsurface-offset-indexed sensitivity kernel into the angle domain and then correlating the corresponding angle-domain sensitivity kernel to produce the angle-dependent illumination. The procedure closely resembles the method of Sava and Fomel (2003), which computes angle-dependent reflectivity image after imaging using a Fourier-domain mapping. We demonstrate in two dimensions that our method generates the scattering-angle illumination suitable for point scatterers if only horizontal subsurface offset has been computed. For planar reflectors, however, dip-dependent scattering-angle illumination is necessary, instead of scattering-angle illumination that averages over all dips. We show that the dip-dependent scattering-angle-domain illumination gather can be obtained by mapping either midpoint wavenumbers or subsurface-offset wavenumbers that contain both horizontal and vertical subsurface offsets.

In fact, the subsurface illumination that we often refer to is only a subset, or more precisely, the diagonal part of the imaging Hessian matrix. The off-diagonal components of the Hessian have been proven useful when applied to improve images of extremely poor illumination, e.g., subsalt regions with shadow zones (Valenciano, 2008; Tang, 2009). Our method discussed here also allows us to compute the off-diagonal components of the angle-domain imaging Hessian in a target-oriented fashion, hence it can be used to invert for the angle-domain reflectivity in a model-domain least-squares migration/inversion. A naive implementation of the proposed method, however, can be prohibitively expensive with the computational cost proportional to the number of sources, receivers and frequencies. We show how the phase-encoding technique (Tang, 2008b,a, 2009) is able to make the method more cost effective.

This paper is organized as follows: we first review the theory of computing the illumination, or more generally, the imaging Hessian, in the subsurface-offset domain. Then we demonstrate how to transform the sensitivity kernel into angle-domain and how to compute the scattering-angle-domain Hessian, as well as the dip-dependent scattering-angle-domain Hessian. Finally we apply our method to the Sigsbee2A model.

SUBSURFACE-OFFSET-DOMAIN IMAGE AND ILLUMINATION

Linearized modeling (Born modeling) from a prestack image parameterized as a function of subsurface offset can be described as follows:

$$d(\mathbf{x}_r, \mathbf{x}_s, \omega) = \sum_{\mathbf{x}} \sum_{\mathbf{h}} L_h(\mathbf{x}, \mathbf{h}, \mathbf{x}_r, \mathbf{x}_s, \omega) m_h(\mathbf{x}, \mathbf{h}), \quad (1)$$

where $d(\mathbf{x}_r, \mathbf{x}_s, \omega)$ is the seismic data with a source located at $\mathbf{x}_s = (x_s, y_s, z_s = 0)$ and a receiver located at $\mathbf{x}_r = (x_r, y_r, z_r = 0)$; ω is the angular frequency; $m_h(\mathbf{x}, \mathbf{h})$ is the prestack image located at $\mathbf{x} = (x, y, z)$ for a half subsurface offset $\mathbf{h} = (h_x, h_y, h_z)$; L_h is the sensitivity kernel defined as follows:

$$L_h(\mathbf{x}, \mathbf{h}, \mathbf{x}_r, \mathbf{x}_s, \omega) = \omega^2 f_s(\omega) G(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) G(\mathbf{x} + \mathbf{h}, \mathbf{x}_r, \omega), \quad (2)$$

where $f_s(\omega)$ is the source signature, and $G(\mathbf{x}, \mathbf{x}_s, \omega)$ and $G(\mathbf{x}, \mathbf{x}_r, \omega)$ are the Green's functions connecting the source and receiver, respectively, to the image point \mathbf{x} .

Reconstruction of the prestack image $m(\mathbf{x}, \mathbf{h})$ can be posed as an inverse problem by minimizing the following objective function defined in the data space:

$$F(\mathbf{m}) = \frac{1}{2} \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} |W(\mathbf{x}_r, \mathbf{x}_s) r(\mathbf{x}_r, \mathbf{x}_s, \omega)|^2, \quad (3)$$

where $r(\mathbf{x}_r, \mathbf{x}_s, \omega) = d(\mathbf{x}_r, \mathbf{x}_s, \omega) - d_{\text{obs}}(\mathbf{x}_r, \mathbf{x}_s, \omega)$ is the data residual and $W(\mathbf{x}_r, \mathbf{x}_s)$ is the acquisition mask operator, which contains unity values where we record data and zeros where we do not. The gradient of the objective function F reads

$$I_h(\mathbf{x}, \mathbf{h}) = \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} W^2(\mathbf{x}_r, \mathbf{x}_s) L_h^*(\mathbf{x}, \mathbf{h}, \mathbf{x}_r, \mathbf{x}_s, \omega) r(\mathbf{x}_r, \mathbf{x}_s, \omega), \quad (4)$$

where $*$ denotes complex conjugation. Equation 4 is similar to the prestack shot-profile migration formula and it produces migrated reflectivity images defined in the subsurface offset domain (Rickett and Sava, 2002).

The Hessian can be obtained by taking the second-order derivatives of F with respect to the model parameters as follows:

$$H_h(\mathbf{x}, \mathbf{x}', \mathbf{h}, \mathbf{h}') = \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} W^2(\mathbf{x}_r, \mathbf{x}_s) L_h^*(\mathbf{x}, \mathbf{h}, \mathbf{x}_r, \mathbf{x}_s, \omega) L_h(\mathbf{x}', \mathbf{h}', \mathbf{x}_r, \mathbf{x}_s, \omega). \quad (5)$$

When $\mathbf{x} = \mathbf{x}'$ and $\mathbf{h} = \mathbf{h}'$, we obtain the diagonal elements of the Hessian operator

$$H_h(\mathbf{x}, \mathbf{h}) = \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} W^2(\mathbf{x}_r, \mathbf{x}_s) |L_h(\mathbf{x}, \mathbf{h}, \mathbf{x}_r, \mathbf{x}_s, \omega)|^2. \quad (6)$$

The diagonal of the Hessian is often known as the illumination map of the subsurface, it contains illumination contribution from both sources and receivers for a given acquisition configuration.

ANGLE-DOMAIN IMAGE AND ILLUMINATION

Because the subsurface offset is linked to the scattering angle and dip angle through local ray parameters, the sensitivity kernel (and consequently the reflectivity image and the Hessian) can be transformed from the subsurface-offset domain into the angle domain through a simple Fourier domain mapping, or equivalently a space-domain slant stack (Sava and Fomel, 2003). In this section, we first demonstrate that a Fourier-domain mapping using the depth and horizontal-subsurface-offset wavenumbers produces scattering angles that are implicitly averaged over illuminated dips. It is useful for point scatterers. Dip-dependent scattering-angle illumination, however, is required for accurately predicting illumination strength for planar reflectors. We restrict our discussion only in 2-D for simplicity, where $\mathbf{x} = (x, z)$, $\mathbf{h} = (h_x, h_z)$, $\mathbf{x}_s = (x_s, z_s = 0)$ and $\mathbf{x}_r = (x_r, z_r = 0)$. But the extension to 3-D should be straight forward and would be discussed in further publications.

Scattering-angle-domain illumination

In a locally constant velocity medium (Figure 1), the midpoint ray parameter \mathbf{p}_m and the subsurface offset ray parameter \mathbf{p}_h can be expressed as follows:

$$\mathbf{p}_m = \mathbf{p}_s + \mathbf{p}_r \quad (7)$$

$$\mathbf{p}_h = \mathbf{p}_r - \mathbf{p}_s, \quad (8)$$

where \mathbf{p}_s and \mathbf{p}_r are the source and receiver ray parameters, respectively. Using trigonometric relations, we can further express the horizontal and vertical components of \mathbf{p}_m and \mathbf{p}_h as functions of the scattering angle γ that bisects the incident and the scattered rays (plane waves) and the corresponding dip angle α as follows:

$$\mathbf{p}_m = \begin{pmatrix} p_{m_x} \\ p_{m_z} \end{pmatrix} = \begin{pmatrix} 2s \cos \gamma \sin \alpha \\ -2s \cos \gamma \cos \alpha \end{pmatrix}, \quad (9)$$

and

$$\mathbf{p}_h = \begin{pmatrix} p_{h_x} \\ p_{h_z} \end{pmatrix} = \begin{pmatrix} 2s \sin \gamma \cos \alpha \\ 2s \sin \gamma \sin \alpha \end{pmatrix}, \quad (10)$$

where s is the slowness at the reflection point. Dividing p_{h_x} by p_{m_z} yields

$$\tan \gamma = -\frac{p_{h_x}}{p_{m_z}} = -\frac{k_{h_x}}{k_{m_z}}, \quad (11)$$

where k_{h_x} and k_{m_z} are vertical-subsurface-offset wavenumber and depth wavenumber, respectively. Equation 11 converts the sensitivity kernel from the subsurface-offset-domain into the scattering-angle domain. The action of the angle-domain sensitivity kernel to the data residual $r(\mathbf{x}_r, \mathbf{x}_s, \omega)$ gives the angle-domain reflectivity image. Since the Fourier-domain mapping is a linear operator and is independent from the data residual, the sources, receivers and frequencies, it can be postponed after the subsurface-offset-domain reflectivity image is obtained. Therefore, to obtain the angle-domain reflectivity image more efficiently, we perform the mapping on the image I_h after stacking over sources, receivers and frequencies, instead of the sensitivity kernel L_h for each source, receiver and frequency (Sava and Fomel, 2003).

We obtain the scattering-angle-domain Hessian by correlating the sensitivity kernels L_γ as follows:

$$H_\gamma(\mathbf{x}, \mathbf{x}', \gamma, \gamma') = \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} W^2(\mathbf{x}_r, \mathbf{x}_s) L_\gamma^*(\mathbf{x}, \gamma, \mathbf{x}_r, \mathbf{x}_s, \omega) L_\gamma(\mathbf{x}', \gamma', \mathbf{x}_r, \mathbf{x}_s, \omega). \quad (12)$$

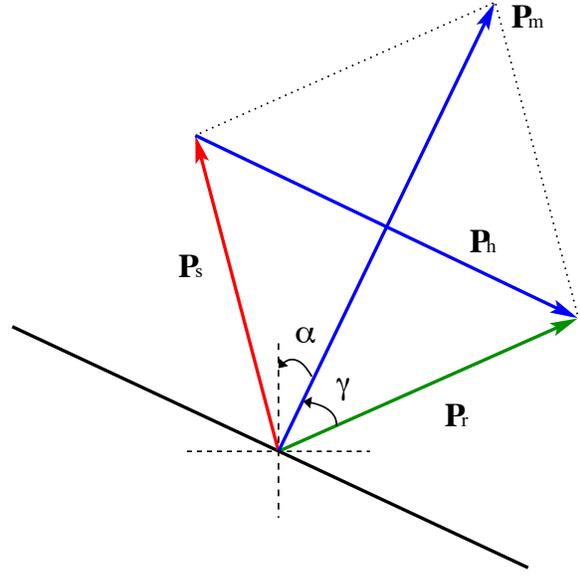
The diagonal of equation 12, or the scattering-angle-domain illumination is obtained when $\mathbf{x} = \mathbf{x}'$ and $\gamma = \gamma'$

$$H_\gamma(\mathbf{x}, \gamma) = \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} W^2(\mathbf{x}_r, \mathbf{x}_s) |L_\gamma(\mathbf{x}, \gamma, \mathbf{x}_r, \mathbf{x}_s, \omega)|^2. \quad (13)$$

Contrary to the case of computing angle-domain reflectivity image, postponing the Fourier-domain mapping after stack becomes less obvious for the angle-domain Hessian computation due to the correlation term inside the summation loop in equations 12 and 13. In the following numerical examples, we take a more straight forward way that directly transforms the sensitivity kernel from the subsurface-offset domain to the scattering-angle-domain to compute the angle-domain Hessian or illumination. The steps can be summarized as follows:

Figure 1: Geometric relations between ray vectors at a reflection point in a locally constant velocity medium. [NR]

yaxun2/. local-phase-relation



- for fixed x , \mathbf{x}_s , \mathbf{x}_r and ω , apply 2-D Fourier transform along axes z and h_x

$$L_h(x, z, h_x, \mathbf{x}_r, \mathbf{x}_s, \omega) \rightarrow L_h(x, k_z, k_{h_x}, \mathbf{x}_r, \mathbf{x}_s, \omega);$$

- perform the mapping

$$L_h(x, k_z, k_{h_x}, \mathbf{x}_r, \mathbf{x}_s, \omega) \rightarrow L_\gamma(x, k_z, \gamma, \mathbf{x}_r, \mathbf{x}_s, \omega)$$

according to relation 11;

- apply inverse 1-D Fourier transform along axis k_z

$$L_\gamma(x, k_z, \gamma, \mathbf{x}_r, \mathbf{x}_s, \omega) \rightarrow L_\gamma(x, z, \gamma, \mathbf{x}_r, \mathbf{x}_s, \omega)$$

to obtain the scattering-angle-domain sensitivity kernel.

- compute either the scattering-angle-domain Hessian using equation 12 or the illumination using equation 13.

As a simple example, Figure 3 shows the real part of the scattering-angle-domain sensitivity kernel, converted from its subsurface-offset-domain counterpart (Figure 2). The results are obtained by using a constant velocity model (2000 m/s), and only one source (−600 m), one receiver (600 m) and one frequency (19 Hz) are computed. Figure 4 shows the corresponding single-frequency scatter-angle-domain illumination for the given velocity model and acquisition configuration. Since there are only one source and one receiver, each subsurface point is illuminated by only one scattering angle.

The scattering-angle-domain illumination is useful for point scatterers, it, however, fails to accurately predict the illumination strength for planar reflectors, where the scattered waves have preferred orientations according to the local dips of the reflectors. The reason behind this is that the transformation (equation 11) is dip-independent, the resulting angle-domain illumination implicitly averages over all dip angles and measures the overall

scattering-angle illumination from all dips illuminated, instead of the illumination from one particular dip of the actual planar reflector. This point is further illustrated by Figures 5 and 6, where the computed scattering-angle-domain illumination (Figure 6(d)) accurately predicts the illumination for the point scatterer (Figures 5(c) and 6(c)), instead of the horizontal (Figures 5(a) and 6(a)) and the -30° dipping reflectors (Figures 5(b) and 6(b)).

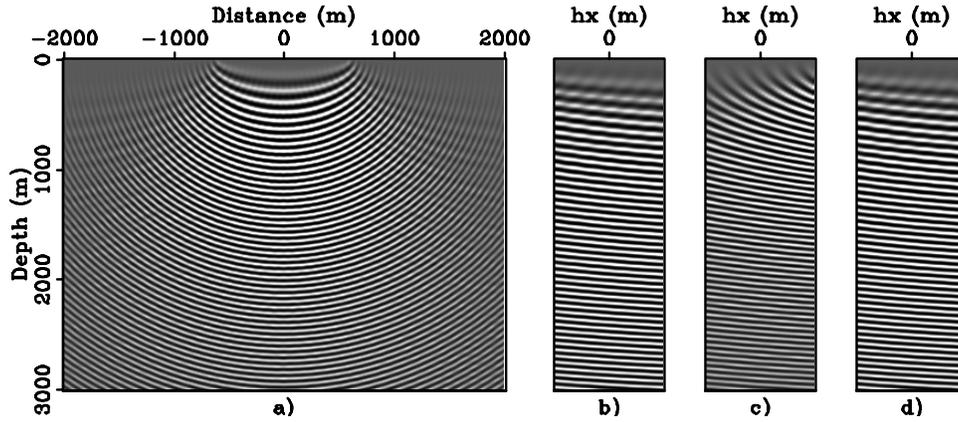


Figure 2: The real part of the subsurface-offset-domain sensitivity kernel for a constant velocity model (2000 m/s), one source at -600 m, one receiver at 600 m and one frequency (19 Hz). Panel (a) shows the kernel at zero-subsurface-offset; (b), (c) and (d) show the kernel for different spatial locations at -1000 m, 0 m and 1000 m, respectively. [CR]

[yaxun2/. const-kernel-odcig-1s1r-1w](#)

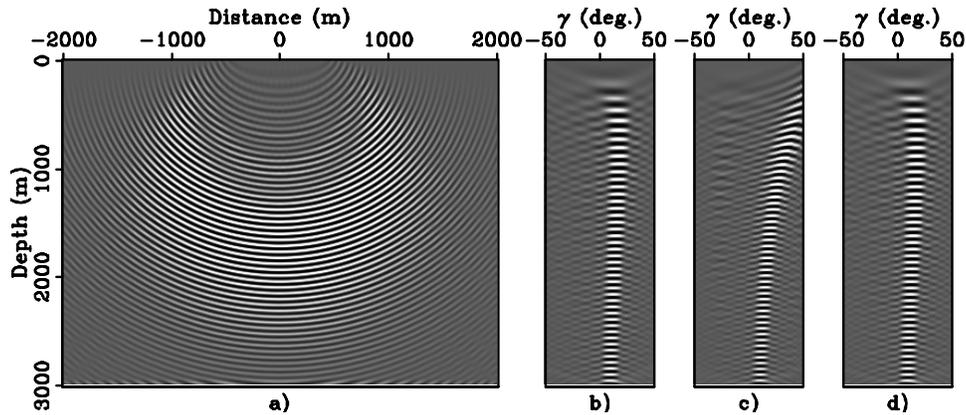


Figure 3: The real part of the scattering-angle-domain sensitivity kernel after conversion from the subsurface offset domain (Figure 2). Panel (a) shows the sensitivity kernel for a constant scattering angle 18.75° ; (b), (c) and (d) show the kernel for spatial locations at -1000 m, 0 m and 1000 m, respectively. [CR]

[yaxun2/. const-kernel-adcig-1s1r-1w](#)

Dip-dependent scattering-angle-domain illumination

To overcome the limitation of the scattering-angle-domain illumination for planar reflectors discussed in the preceding section, we further decompose the illumination into dip-angle

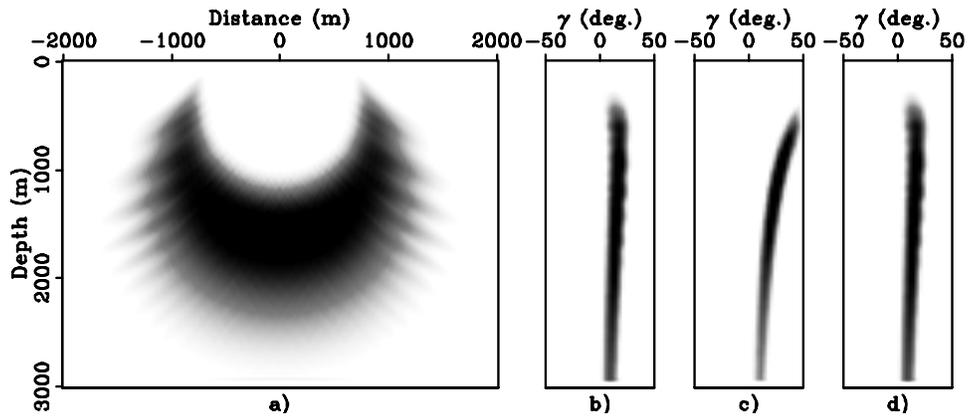


Figure 4: The single-frequency scattering-angle-domain illumination obtained using the sensitivity kernel shown in Figure 3. Panel (a) shows the illumination for scattering angle 18.75° ; (b), (c) and (d) show the illumination angle gathers for spatial locations at -1000 m, 0 m and 1000 m, respectively. [CR] yaxun2/. const-hess-adcig-1s1r-1w

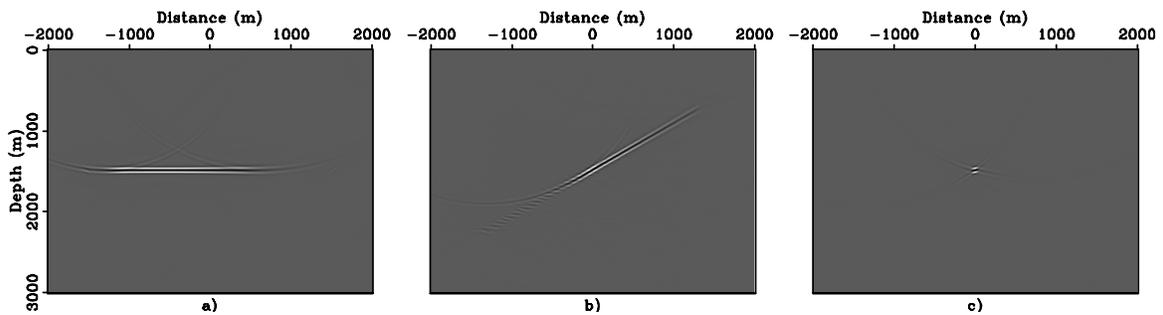
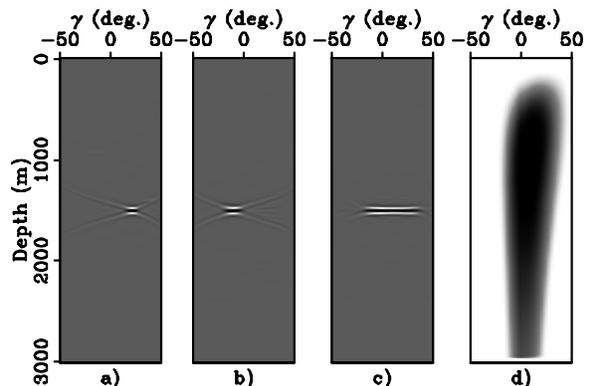


Figure 5: Migrated zero-subsurface-offset images (stacked images) for (a) a horizontal reflector, (b) a dipping reflector (-30°) and (c) a point scatterer. All images are obtained by migrating only one shot located at -600 m, where 401 receivers spread from -2000 m to 2000 m with a 10 m spacing. [CR] yaxun2/. const-imag-stack

Figure 6: The scattering-angle-domain image gathers extracted at spatial location 0 m for (a) the horizontal reflector, (b) the dipping reflector and (c) the point scatterer. Panel (d) shows the scattering-angle-domain illumination gather extracted at the same spatial location. [CR] yaxun2/. const-imag-adcig



domain, resulting in dip-dependent scattering-angle-domain illumination. From equations 9 and 10, it is easy to obtain the tangent of the dip angle using either

$$\tan \alpha = -\frac{p_{m_x}}{p_{m_z}} = -\frac{k_{m_x}}{k_{m_z}}, \quad (14)$$

or

$$\tan \alpha = \frac{p_{h_z}}{p_{h_x}} = \frac{k_{h_z}}{k_{h_x}}, \quad (15)$$

where k_{m_x} , k_{m_z} and k_{h_x} , k_{h_z} are the horizontal and vertical components of the midpoint wavenumber vector and the subsurface-offset wavenumber vector, respectively.

Dip decomposition using either equation 14 or 15 has its own pros and cons. Equation 15 is suitable for computing dip-angle gathers for sparsely isolated image points, because it does not require any CMP information, i.e., k_{m_x} and k_{m_z} , and outputting gathers for sparsely isolated image points may mitigate the extra computer time and storage spent in computing both the horizontal and vertical subsurface offsets, h_x and h_z . On the other hand, equation 14 is computationally less demanding, because it does not require computing vertical subsurface offsets. However, it estimates dips using the CMP information, hence a block of densely sampled image points in the CMP domain should be output to avoid dip aliasing. In the following numerical examples, we use equation 14 for dip decomposition due to the fact that it is relatively inconvenient to output vertical subsurface offsets by using one-way wave-equation-based extrapolators.

After transforming the subsurface-offset-domain sensitivity kernel into the dip-dependent scattering-angle domain, we can proceed to compute the corresponding Hessian using

$$\begin{aligned} H_{\gamma,\alpha}(\mathbf{x}, \mathbf{x}', \gamma, \gamma', \alpha, \alpha') = & \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} W^2(\mathbf{x}_r, \mathbf{x}_s) L_{\gamma,\alpha}(\mathbf{x}, \gamma, \alpha, \mathbf{x}_r, \mathbf{x}_s, \omega) \\ & \times L_{\gamma,\alpha}^*(\mathbf{x}', \gamma', \alpha', \mathbf{x}_r, \mathbf{x}_s, \omega), \end{aligned} \quad (16)$$

or the illumination using

$$H_{\gamma,\alpha}(\mathbf{x}, \gamma, \alpha) = \sum_{\omega} \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} W^2(\mathbf{x}_r, \mathbf{x}_s) |L_{\gamma,\alpha}(\mathbf{x}, \gamma, \alpha, \mathbf{x}_r, \mathbf{x}_s, \omega)|^2, \quad (17)$$

where $L_{\gamma,\alpha}$ is the dip-dependent scattering-angle-domain sensitivity kernel. The complete procedure can be summarized as follows:

- for fixed \mathbf{x}_s , \mathbf{x}_r and ω , apply 3-D Fourier transform along axes x , z and h_x (if only horizontal subsurface offsets are computed)

$$L_h(x, z, h_x, \mathbf{x}_r, \mathbf{x}_s, \omega) \rightarrow L_h(k_x, k_z, k_{h_x}, \mathbf{x}_r, \mathbf{x}_s, \omega),$$

or along z , h_x and h_z (if both horizontal and vertical subsurface offsets are computed)

$$L_h(x, z, h_x, h_z, \mathbf{x}_r, \mathbf{x}_s, \omega) \rightarrow L_h(x, k_z, k_{h_x}, k_{h_z}, \mathbf{x}_r, \mathbf{x}_s, \omega);$$

- perform the mapping

$$L_h(k_x, k_z, k_{h_x}, \mathbf{x}_r, \mathbf{x}_s, \omega) \rightarrow L_{\gamma,\alpha}(k_x, k_z, \gamma, \alpha, \mathbf{x}_r, \mathbf{x}_s, \omega)$$

according to relations 11 and 14, or

$$L_h(x, k_z, k_{h_x}, k_{h_z}, \mathbf{x}_r, \mathbf{x}_s, \omega) \rightarrow L_{\gamma, \alpha}(x, k_z, \gamma, \alpha, \mathbf{x}_r, \mathbf{x}_s, \omega)$$

according to relations 11 and 15;

- apply inverse 2-D Fourier transform along axes k_x and k_z

$$L_{\gamma, \alpha}(k_x, k_z, \gamma, \alpha, \mathbf{x}_r, \mathbf{x}_s, \omega) \rightarrow L_{\gamma, \alpha}(x, z, \gamma, \alpha, \mathbf{x}_r, \mathbf{x}_s, \omega),$$

or inverse 1-D Fourier transform along axis k_z

$$L_{\gamma, \alpha}(x, k_z, \gamma, \alpha, \mathbf{x}_r, \mathbf{x}_s, \omega) \rightarrow L_{\gamma, \alpha}(x, z, \gamma, \alpha, \mathbf{x}_r, \mathbf{x}_s, \omega)$$

to obtain the dip-dependent scattering-angle-domain sensitivity kernel.

- compute either the dip-dependent scattering-angle-domain Hessian using equation 16 or the illumination using equation 17.

For the same constant velocity example, Figures 7 and 8 show the dip-dependent scattering-angle-domain illumination for 0° and -30° dip angles, respectively. The acquisition geometry is the same as that in Figure 5, i.e., 1 shot and 401 receivers. The illumination gathers (Figures 7(b) and 8(b)) successfully predict the angle-dependent illumination for both the horizontal and dipping reflectors (Figure 7(c) and Figure 8(c)).

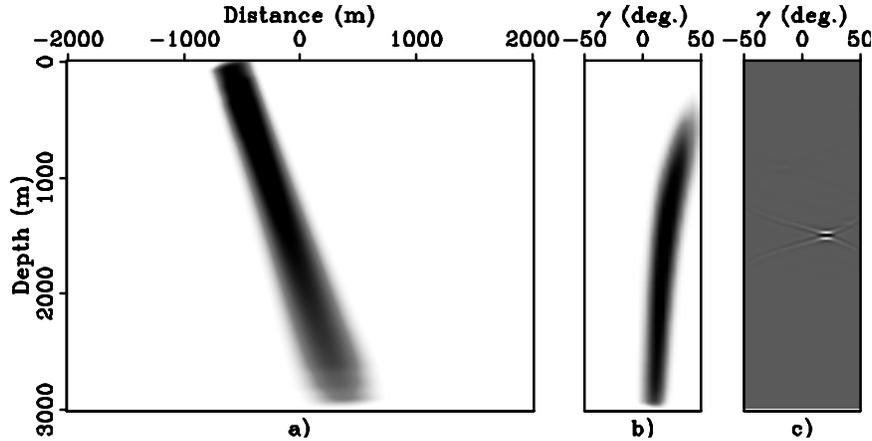


Figure 7: Dip-dependent scattering-angle-domain illumination. Panel (a) is the illumination for a constant dip angle $\alpha = 0^\circ$ and a constant scattering angle $\gamma = 18.75^\circ$; (b) is the illumination angle gather for a constant dip angle $\alpha = 0^\circ$ and at spatial location $x = 0$ m; (c) is the reflectivity angle gather for the horizontal reflector extracted at $x = 0$ m, it is the same as Figure 6(a). [CR] `yaxun2/. const-imag-hess-dip-adcig1`

REDUCING THE COST BY PHASE ENCODING

As mentioned in previous sections, equations 13 or 17 is intuitive to implement, however, it requires performing offset to angle transform for each component of the sensitivity kernel. This can be potentially expensive, since the cost is now proportional to the number of

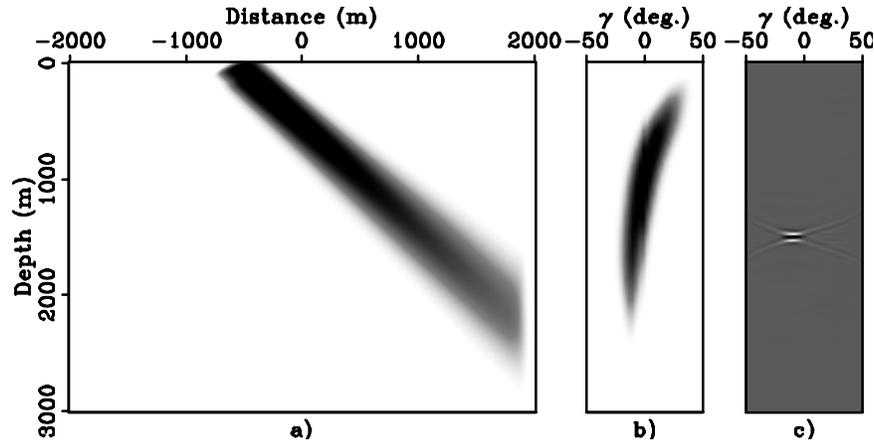


Figure 8: Dip-dependent scattering-angle-domain illumination. Panel (a) is the illumination for a constant dip angle $\alpha = -30^\circ$ and a constant scattering angle $\gamma = 18.75^\circ$; (b) is the illumination angle gather for a constant dip angle $\alpha = -30^\circ$ and at spatial location $x = 0$ m; (c) is the reflectivity angle gather for the dipping reflector extracted at $x = 0$ m, it is the same as Figure 6(b). [CR] `yaxun2/. const-imag-hess-dip-adcig2`

sources, receivers and frequencies. In this section, we show that the computational cost can be significantly reduced by using the phase-encoding technique, which was first introduced into wave-equation shot-record migration (Romero et al., 2000), and then adapted to Hessian computation by Tang (2009).

The basic idea behind phase encoding is simple, i.e., instead of computing the Green's functions sequentially with point sources as source functions, we now compute them simultaneously with encoded areal source as the source function. Thanks to the linearity of the wave equation with respect to the sources, the resultant wavefield computed using the encoded areal source can be expressed as the sum of the wavefields computed using the point sources. Therefore, instead of performing many wavefield propagations with the number of propagations being proportional to the number of point sources, by using the encoded areal sources, we reduce the number of wavefield propagations to just one, drastically reducing the computational cost. The encoding can be done in the receiver domain, which is suitable for any acquisition geometry, or in both receiver and source domain, which is suitable for land or ocean-bottom-node acquisition geometry (Tang, 2009). One drawback of the phase-encoding method, however, is that it generates crosstalk artifacts (Tang, 2009). The crosstalk can be attenuated by carefully choosing the encoding functions, such as plane-wave phase-encoding function, random phase-encoding function or a mix of the two (Tang, 2008a).

Figure 9 shows the exact scattering-angle-domain illumination for the same constant velocity model and the 1-shot-401-receiver configuration (Figure 6). It is obtained using the point source Green's function and is very expensive, and we show it here for benchmarking purposes. Figure 10 shows the result obtained by assembling the receiver-side Green's functions but without any phase-encoding function applied. The result is apparently dominated by crosstalk artifacts and has a completely wrong illumination pattern. The crosstalk artifacts can be effectively attenuated using the random phase-encoding function,

as shown in Figures 11, 12 and 13, which are computed using different number of random realizations. Although not shown here, similar results can be obtained for the dip-dependent scattering-angle-domain illumination.

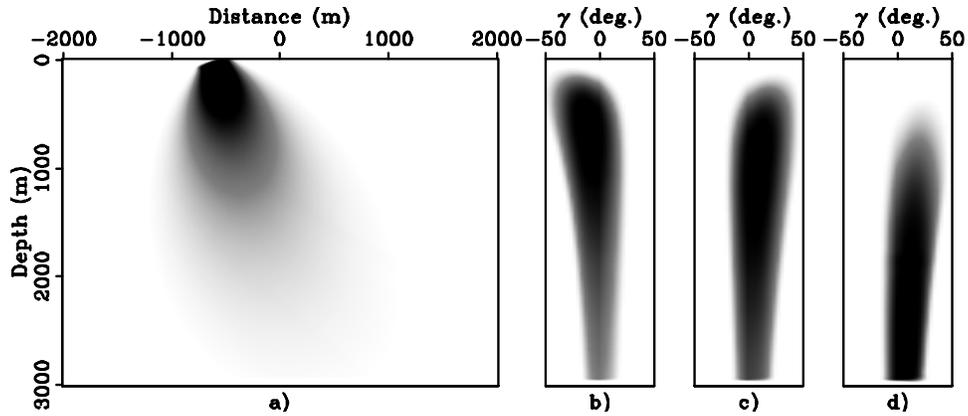


Figure 9: Exact scattering-angle-domain illumination. (a) shows the illumination for scattering angle 18.75° ; (b), (c) and (d) shows the illumination angle gathers for spatial locations -1000 m, 0 m and 1000 m, respectively. [CR] [yaxun2/. const-hess-adcig-1s401r-exact](#)

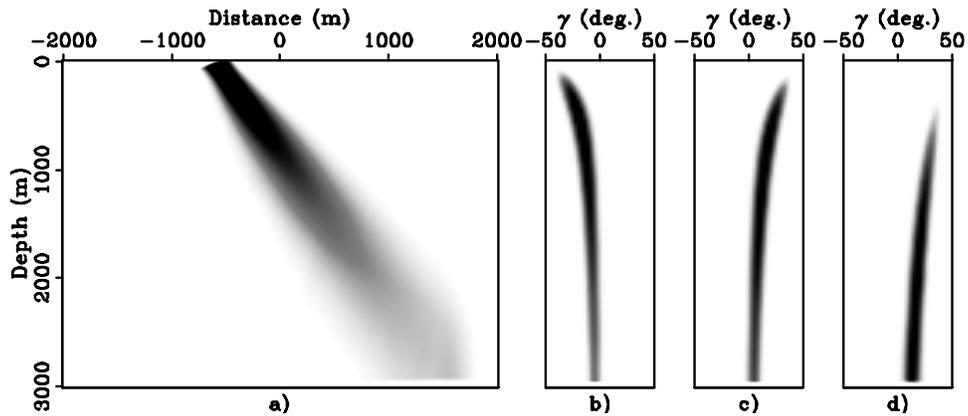


Figure 10: Scattering-angle-domain illumination with crosstalk artifacts. View descriptions are the same as Figure 9. [CR] [yaxun2/. const-hess-adcig-1s401r-xtalk](#)

NUMERICAL EXAMPLES

We test our methods on the Sigsbee2A model, where the complex salt body and limited acquisition geometry result in uneven subsurface illumination below the salt. The velocity field shown in Figure 14 is used for computing various angle-domain reflectivity images and illumination. Figure 15 shows the scattering-angle-domain reflectivity image. Note the holes in the scattering-angle gathers (Figures 15(b)-(g)) caused by uneven illumination. This phenomenon should conform with the angle-domain illumination. As expected, the scattering-angle-domain illumination (Figure 16) fails to accurately predict the illumination pattern for planar reflectors, e.g., the horizontal reflector at depth 29500 ft, although it is very accurate for the point scatterers located at depth 25000 ft. We further decompose

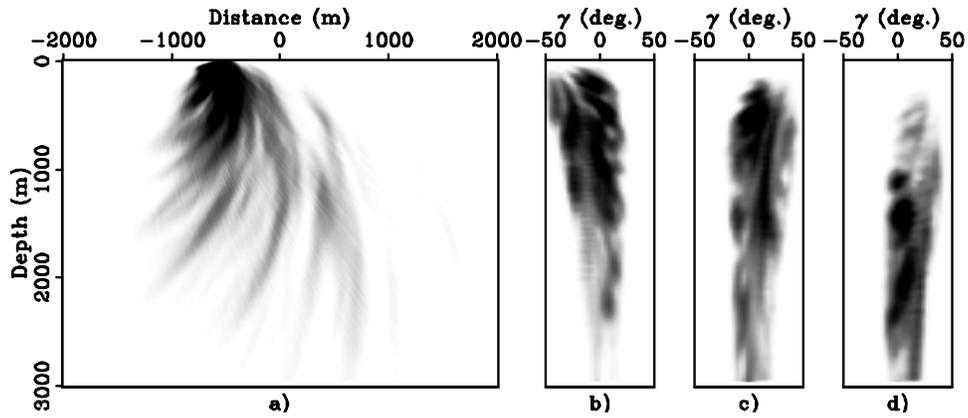


Figure 11: Receiver-side encoded scattering-angle-domain illumination with 1 random realization. View descriptions are the same as Figure 9. [CR]

`yaxun2/. const-hess-adcig-1s401r-stack1`

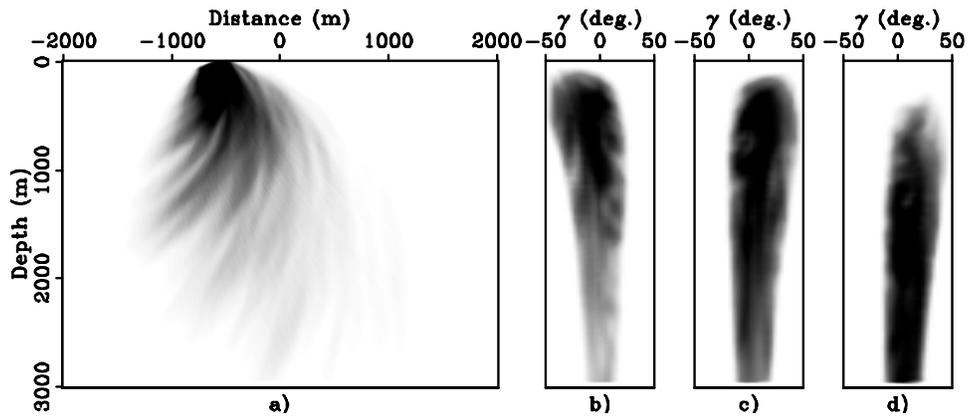


Figure 12: Receiver-side encoded scattering-angle-domain illumination with 5 random realizations. View descriptions are the same as Figure 9. [CR]

`yaxun2/. const-hess-adcig-1s401r-stack5`

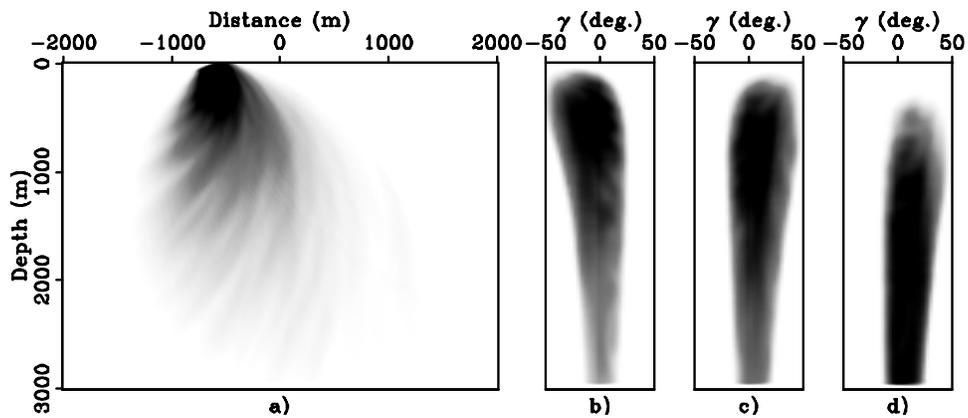


Figure 13: Receiver-side encoded scattering-angle-domain illumination with 20 random realizations. View descriptions are the same as Figure 9. [CR]

`yaxun2/. const-hess-adcig-1s401r-stack20`

the image and illumination into dip-dependent scattering-angle domain. The illumination pattern (Figure 18) now conforms very well with the image of the planar reflectors (Figure 17) as well as the reflectors that have a zero dip in Figure 15.

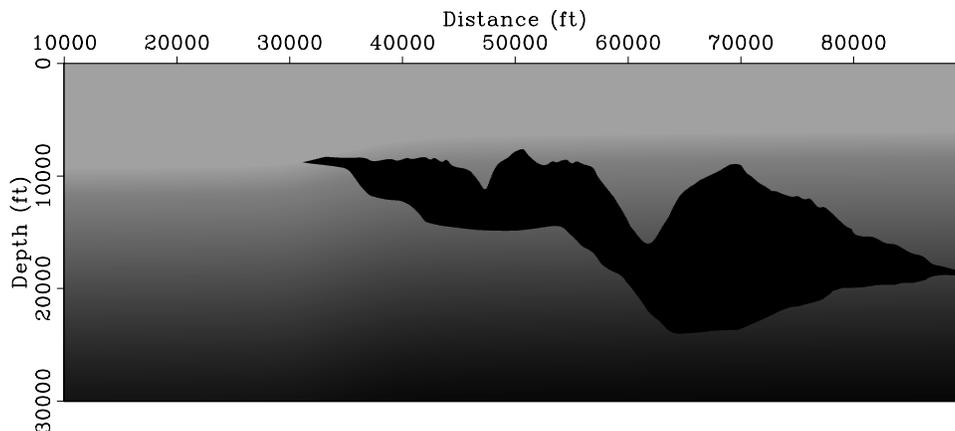


Figure 14: Sigsbee2A velocity model used for migration and illumination computation. [ER] `yaxun2/. sigsb2a-vmig`

CONCLUSIONS

We present a method for computing the angle-domain illumination, or more generally, the angle-domain Hessian. The method takes advantage of the relation among scattering angles, dip angles, CMP wavenumbers and subsurface-offset wavenumbers, and uses a simple Fourier-domain mapping to transform from subsurface offset domain into angle domain. Our method can decompose the illumination or Hessian into either scattering-angle domain suitable for point scatterers, or dip-dependent scattering-angle domain suitable for planar reflectors. A straight forward implementation of our method, however, can be very expensive. We show how the cost can be drastically reduced by using the phase-encoding technique. Numerical examples demonstrate that our method can produce accurate illumination estimation for both point scatterers and planar reflectors, and therefore, it is useful for application related to AVA analysis or robust migration velocity residual parameter estimation.

REFERENCES

- Bear, G., C.-P. Lu, R. Lu, D. Willen, and I. Watson, 2000, The construction of subsurface illumination and amplitude maps via ray tracing: *The Leading Edge*, **19**, 726–728.
- Clapp, M. L., 2005, *Imaging Under Salt: Illumination Compensation by Regularized Inversion*: PhD thesis, Stanford University.
- Nemeth, T., C. Wu, and G. Schuster, 1999, Least-squares migration of incomplete reflection data: *Geophysics*, **64**, 208–221.
- Rickett, J. E. and P. C. Sava, 2002, Offset and angle-domain common image-point gathers for shot-profile migration: *Geophysics*, **67**, 883–889.
- Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: *Geophysics*, **65**, 426–436.

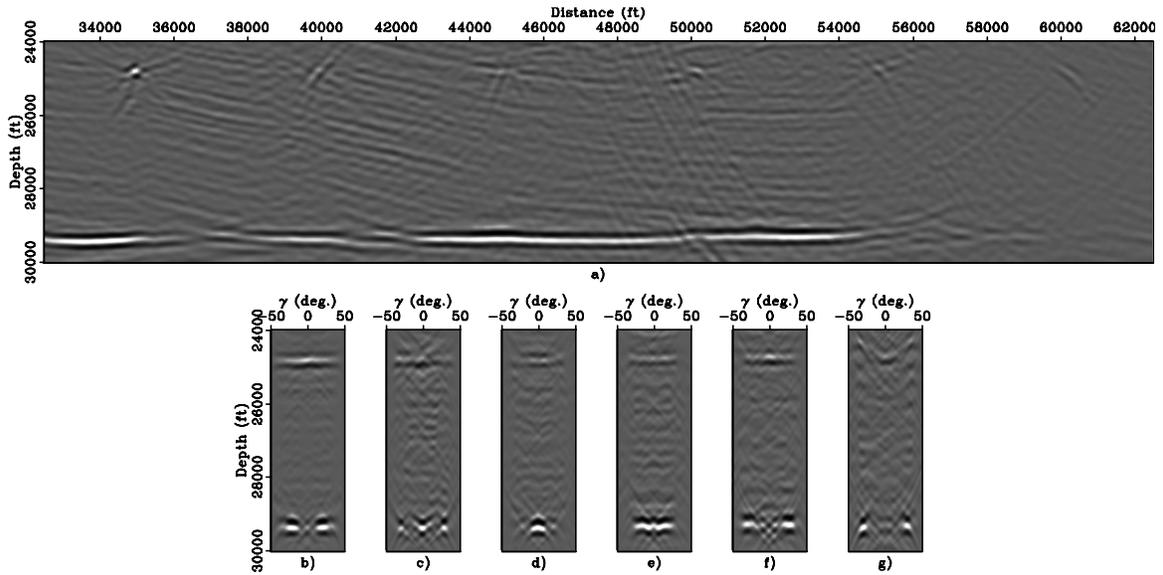


Figure 15: Scattering-angle-domain reflectivity image for the Sigsbee2A model. Panel (a) shows the image with scattering angle $\gamma = 0^\circ$; (b), (c), (d), (e), (f) and (g) show the scattering angle gathers extracted at spatial locations 35000 ft, 40000 ft, 45000 ft, 50000 ft, 55000 ft and 60000 ft, respectively. [CR] [yaxun2/. sigsb2a-imag-adcig](#)

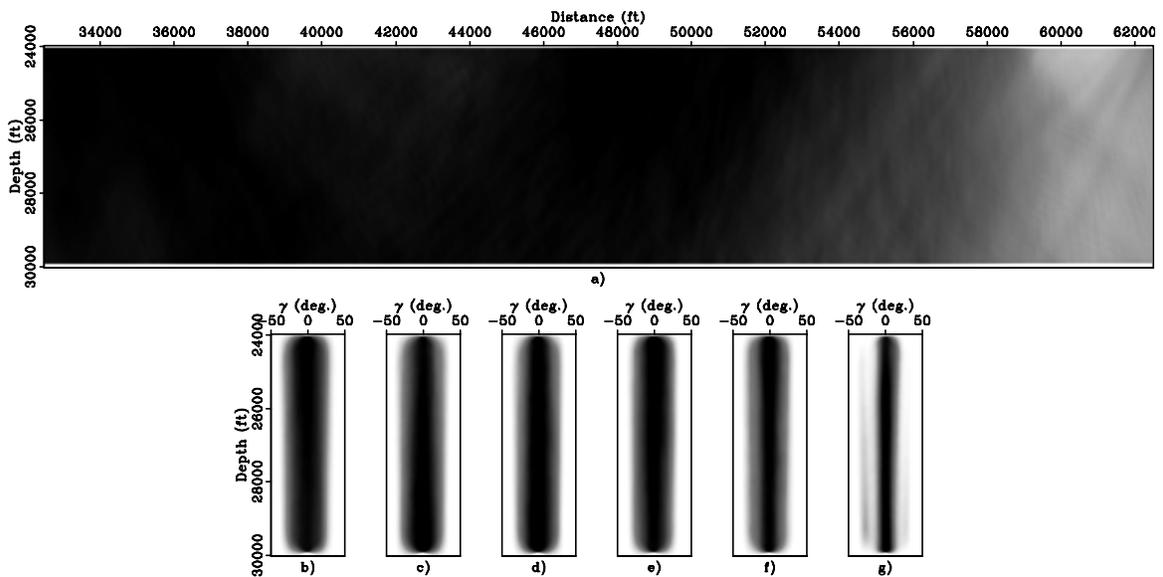


Figure 16: Scattering-angle-domain illumination for the Sigsbee2A model. View descriptions are the same as Figure 15. [CR] [yaxun2/. sigsb2a-hess-adcig](#)

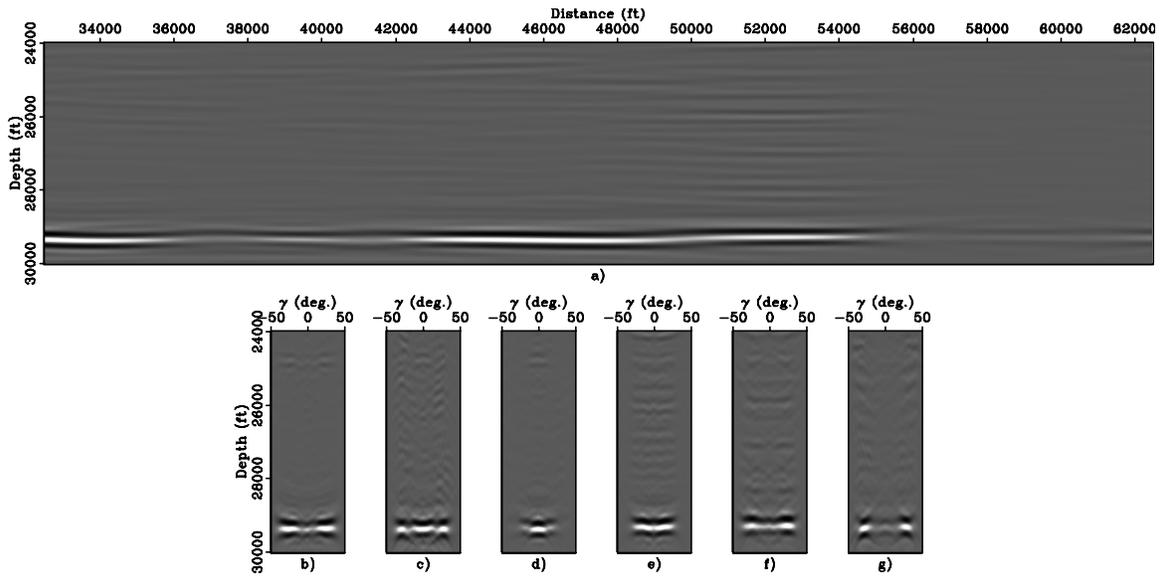


Figure 17: Dip-dependent scattering-angle-domain reflectivity image for the Sigsbee2A model. Panel (a) shows the image with scattering angle $\gamma = 0^\circ$ and dip angle $\alpha = 0^\circ$; (b), (c), (d), (e), (f) and (g) show the scattering angle gathers extracted at spatial locations 35000 ft, 40000 ft, 45000 ft, 50000 ft, 55000 ft and 60000 ft, respectively. [CR] yaxun2/. sigsb2a-imag-adcig-dip

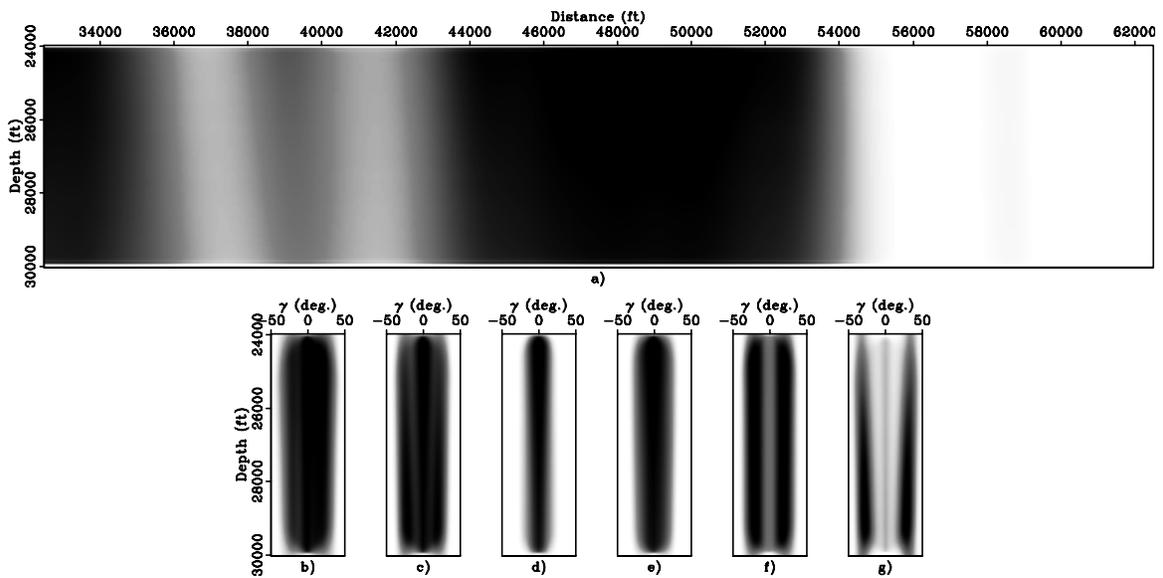


Figure 18: Scattering-angle-domain illumination for the Sigsbee2A model. View descriptions are the same as Figure 17. [CR] yaxun2/. sigsb2a-hess-adcig-dip

- Sava, P. C. and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: *Geophysics*, **68**, 1065–1074.
- Schneider, W. A. and J. G. A. Winbow, 1999, Efficient and accurate modelling of 3-d seismic illumination: *SEG Technical Program Expanded Abstracts*, **18**, 633–636.
- Tang, Y., 2008a, Modeling, migration and inversion in the generalized source and receiver domain: **SEP-136**, 97–112.
- , 2008b, Wave-equation Hessian by phase encoding: **SEP-134**, 1–24.
- , 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian: *Geophysics*, **74**, WCA95–WCA107.
- Valenciano, A., 2008, *Imaging by Wave-equation Inversion*: PhD thesis, Stanford University.
- Xie, X.-B., S. Jin, and R.-S. Wu, 2006, Wave-equation-based seismic illumination analysis: *Geophysics*, **71**, S169–S177.

Joint least-squares inversion of up- and down-going signal for ocean bottom data sets

Mandy Wong, Biondo Biondi, and Shuki Ronen

ABSTRACT

We present a joint least-squares inversion method for imaging the acoustic primary (up-going) and mirror (down-going) signals for ocean-bottom seismic processing. Joint inversion combines the benefits of wider illumination from the mirror signal and better signal-to-noise ratio from the primary signal into one image. Results from two modified 2D Marmousi models show a better illumination of the subsurface and improved resolution in geologically complex areas.

INTRODUCTION

Ocean-bottom seismic (OBS) acquisition is an established technology in which seismometers are placed at the sea bottom and shots are fired at the sea surface. In areas congested by platforms or other obstacles, ocean bottom seismic is advantageous because it is operated by small boats without cumbersome towed streamers. Such a geometry enables OBS acquisition to provide wide-azimuth illumination, shear-wave recording, a quieter recording environment, higher-resolution data and improved repeatability. Therefore, it is used for imaging in obstructed oilfields and for time-lapse monitoring of hydrocarbon reservoirs.

There are different processing schemes for ocean bottom data. The traditional way, inherited from surface seismic processing, is to remove all free-surface multiples and to migrate only with the primary signal (Wang et al., 2009). Therefore, initial work on OBS data processing has been dedicated to the removal of free-surface multiples. One way to attenuate strong free-surface multiples is to combine the geophone and hydrophone recordings to eliminate the receiver ghost and the water column reverberations, a technique known as PZ summation (Barr and Sander, 1989; Soubaras, 1996; Schalkwijk et al., 1999). Such a technique uses the polarity difference between a scalar measurement (pressure) and a vector measurement (velocity). As an alternative to PZ summation, Sonneland and Berg (1987) and Amundsen (2001) address free-surface multiples with the theory of up-down deconvolution in both layered and complex media. In this approach, not only are all free-surface multiples attenuated, but also de-ghosting and signature deconvolution are conducted in a single step.

While multiples are often treated as noise, they are formed by the same source signal as primaries but travel along different paths in the medium. The receiver ghost, also known as the mirror signal, is the next order of reflection beyond the primaries with an additional reflection off the sea surface. In a deep water OBS survey, the source grid has a much wider lateral extent than the receiver grid. Therefore, the subsurface reflection point of the receiver ghost is located at greater distances from the receiver station than the primaries Figure 1. Therefore, the mirror signal can provide a wider subsurface illumination than the

primaries if the energy is properly migrated. Several authors have used the mirror signal in the migration of OBS data (Godfrey *et al.*, 1998; Ronen *et al.*, 2005; Grion *et al.*, 2007; Dash *et al.*, 2009).

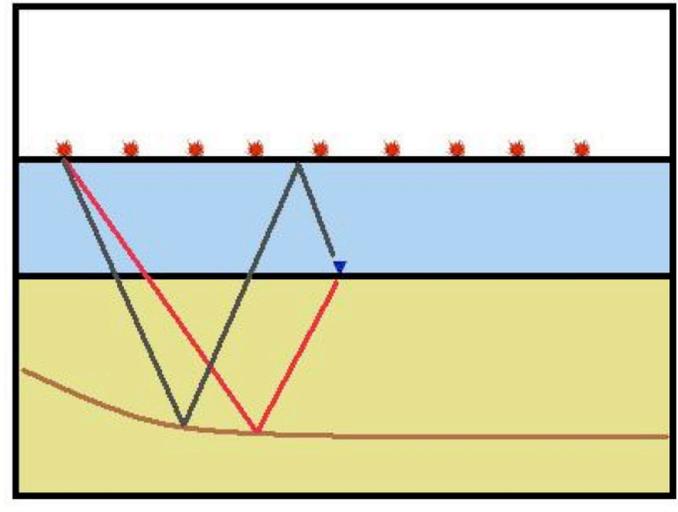


Figure 1: The subsurface reflection point of the receiver ghost, also known as the mirror signal (in black), is located at a greater distance from the receiver station than the primary signal (in red). For a deep water OBS survey, the source grid has a much wider lateral extent than the receiver grid. This translates to a wider subsurface illumination for the mirror signal than the primaries. [NR] mandy1/. illum

While most authors conclude that the mirror image gives a better result than the conventional primary image, the information in the primary image is also valuable. The primary ray path is shorter than the mirror ray path, which contributes to a higher signal-to-noise ratio in the primary signal. In a common receiver gather, the illumination points of the subsurface by each source-receiver pair are closer together in the primary reflection than in the mirror reflection. This translates to higher image quality in the region illuminated by the primary. Instead of treating the primary image and the mirror image separately, we propose an iterative linear least-squares inversion scheme that combines the primary and the mirror image. Such an inversion can improve the structure and aperture of the seismic images by using two sets of signals: the up-going primaries and the down-going mirror signals.

Muijs *et al.* (2007) made an early attempt to image primary and free-surface multiples together. It requires the data to be decomposed into up-going and down-going constituents, followed by downward extrapolation and a 2D deconvolution based imaging condition. While this technique is computationally efficient, its image contains crosstalk artifacts caused by interference of up-going and down-going waves not associated with the same subsurface reflector. In contrast to Muijs' method, joint inversion can optimally combine structural information provided by two types of reflection that are free from crosstalk.

In this paper, we focus on the joint inversion of the acoustic (P) wave signal. We first discuss the theory of the joint linear least-squares inversion. We then apply the inversion scheme to two modified versions of the 2D Marmousi model and show the overall improvement of the joint inversion result.

JOINT INVERSION OF UP/DOWN-GOING P WAVE

Joint inversion of up- and down-going signals for ocean-bottom data can potentially be a better imaging technique than migrating either signal alone, because it combines information from both sets of signals. Figure 4 summarizes the processing scheme for our algorithm. Ocean bottom data are first separated into acoustic up- and down-going components above the seafloor. The decomposed signals are then inverted to yield one optimally combined reflectivity image. The objective function for such an inversion is:

$$0 \approx \begin{bmatrix} \mathbf{L}_\uparrow \\ \mathbf{L}_\downarrow \end{bmatrix} \mathbf{m} - \begin{bmatrix} \mathbf{d}_\uparrow \\ \mathbf{d}_\downarrow \end{bmatrix}, \quad (1)$$

where \mathbf{L}_\uparrow and \mathbf{L}_\downarrow are modeling operators that produce up-going data (\mathbf{d}_\uparrow) and down-going data (\mathbf{d}_\downarrow) from the model space (\mathbf{m}). The up- and down-going operators can be defined in many ways with varying levels of difficulty and practicality. We use the adjoint of the acoustic reverse time migration (RTM) operator to formulate \mathbf{L}_\uparrow and \mathbf{L}_\downarrow . Two modified computational grids are used to forward model the lowest order of up- and down-going signals, namely the primary and the receiver ghost. The formulation of the modeling and its adjoint (RTM) operator is summarized in Figure 2 and Figure 3.

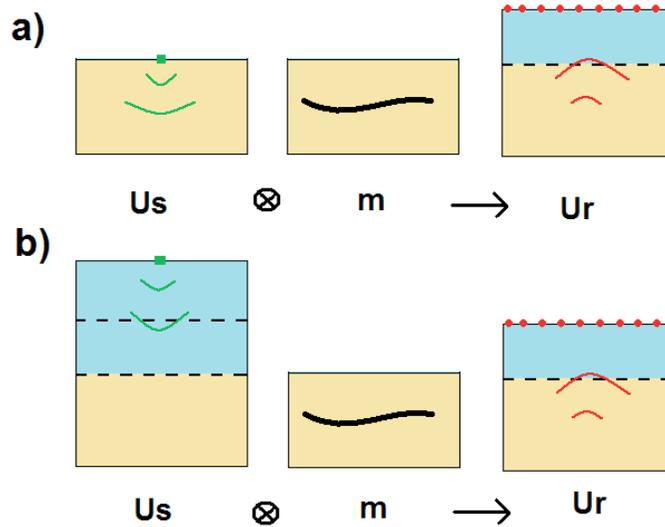


Figure 2: Forward modeling of (a) primary-only and (b) mirror-only data. The algorithm involves cross-correlating the source wavefield (U_s) with the reflectivity model (m) to generate the receiver wavefield (U_r). Reciprocity is used here where the data, in common-receiver domain, are injected at the source location while the source wavelet is injected at the receiver location. Cross-correlation is done only with grid points below the seabed. [NR]

mandy1/. forward

In the modified computational grid as shown in Figure 2, the primary signal is obtained by the cross-correlation of the source wavefields with the reflectivity. For the down-going receiver ghost, the receiver nodes are placed at twice the water depth, which effectively represents a reflection off the sea surface.

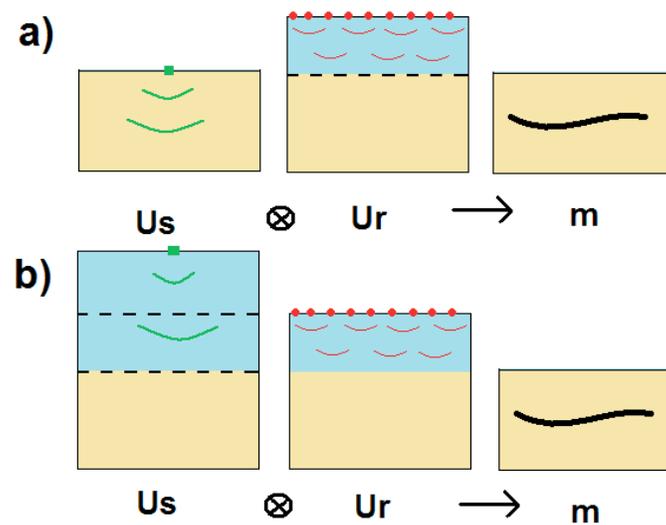


Figure 3: RTM of (a) primary-only and (b) mirror-only data. The algorithm involves cross-correlating the source wave field (U_s) with the receiver wave field (U_r) to generate the reflectivity model (m). Cross-correlation is done only with grid points below the seabed.

[NR] `mandy1/. reverse`

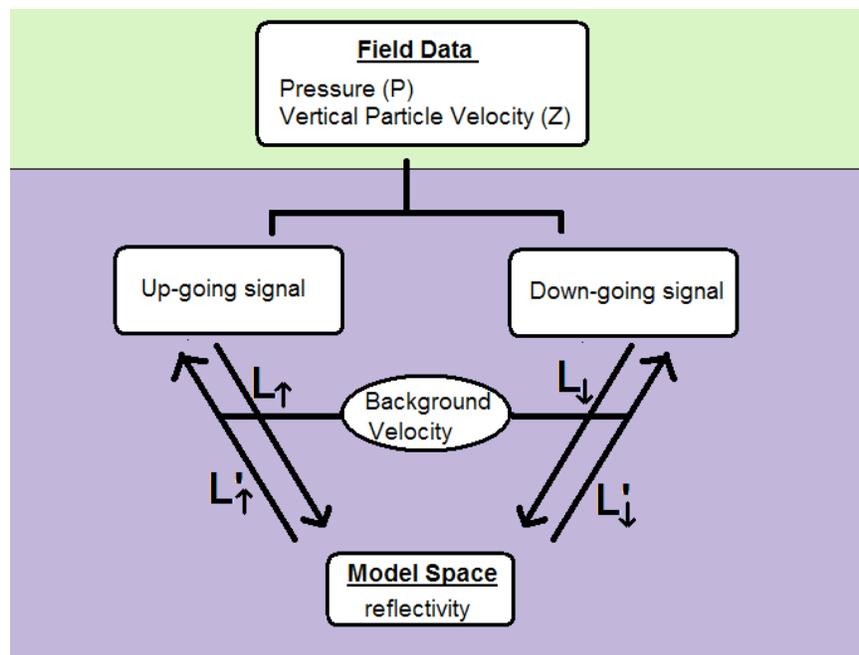


Figure 4: Pressure (P) and vertical particle velocity (Z) data are converted into up- and down-going data. The up- and down-going data are then migrated separately using a modified grid shown in Figure 3. Inversion is performed with residuals in the up/down data domain.

[NR] `mandy1/. Flowchart`

NUMERICAL EXAMPLE

We demonstrate the joint inversion of up- and down-going signals using two test cases, both modified from the 2D Marmousi model. The first example has a dense source and receiver spacing. It shows the improvement of joint inversion under optimal conditions. The second example explores the results of our algorithm when the source and receiver spacing are large.

Marmousi model with dense sampling

Figure 6 (d) shows the reflectivity model used for this example. The model is 2 km deep and 7 km wide with a spacing of 10 m. The ocean floor is not included in this model and is assumed to be flat. However, it is possible to adapt the method to handle an ocean floor with topography. We use a water layer of 500m. All the images are generated in a target-oriented way. That means when we apply the RTM operator, we only cross-correlate regions below the sea-bottom. For simplicity, a constant velocity of 2500 m/s is used for this model.

For the synthetic data, we use the \mathbf{L}_\uparrow and \mathbf{L}_\downarrow defined in the last section to forward model the lowest-order up-going (primary) and down-going (receiver ghost) data. In equation form, this is written as:

$$\begin{aligned} \mathbf{d}_\uparrow^{mod} &= \mathbf{L}_\uparrow \mathbf{m} \\ \mathbf{d}_\downarrow^{mod} &= \mathbf{L}_\downarrow \mathbf{m}. \end{aligned} \quad (2)$$

Shots run from 0 to 7000 m at the sea surface at an interval of 10 m. There are 700 shots in total. For the receiver geometry, we use a receiver spacing of 100 m with ocean bottom nodes located at every grid point from 2500 m to 4400 m. There are 20 nodes in total. Because reciprocity is used later, this geometry is equivalent to having 20 shots at the sea-bottom and 700 receivers at the sea-surface. Figure 5 shows the corresponding up-going and down-going common receiver gathers for a shot at $x=3200$ m.

To compare among conventional primary migration, mirror imaging, and joint inversion, we will first present the results of RTM on the conventional primary signal and on the mirror signal. The corresponding image will then be compared to the joint inversion result using both signals.

Reverse time migration on conventional primary and mirror signals

In this section, we define the term *up-image* (\mathbf{m}_\uparrow) to be applying the adjoint of \mathbf{L}_\uparrow to the up-going data. The term *down-image* (\mathbf{m}_\downarrow) is defined similarly. In equation form, this is written as:

$$\begin{aligned} \mathbf{m}_\uparrow &= \mathbf{L}'_\uparrow \mathbf{d}_\uparrow^{mod}, \\ \mathbf{m}_\downarrow &= \mathbf{L}'_\downarrow \mathbf{d}_\downarrow^{mod}. \end{aligned} \quad (3)$$

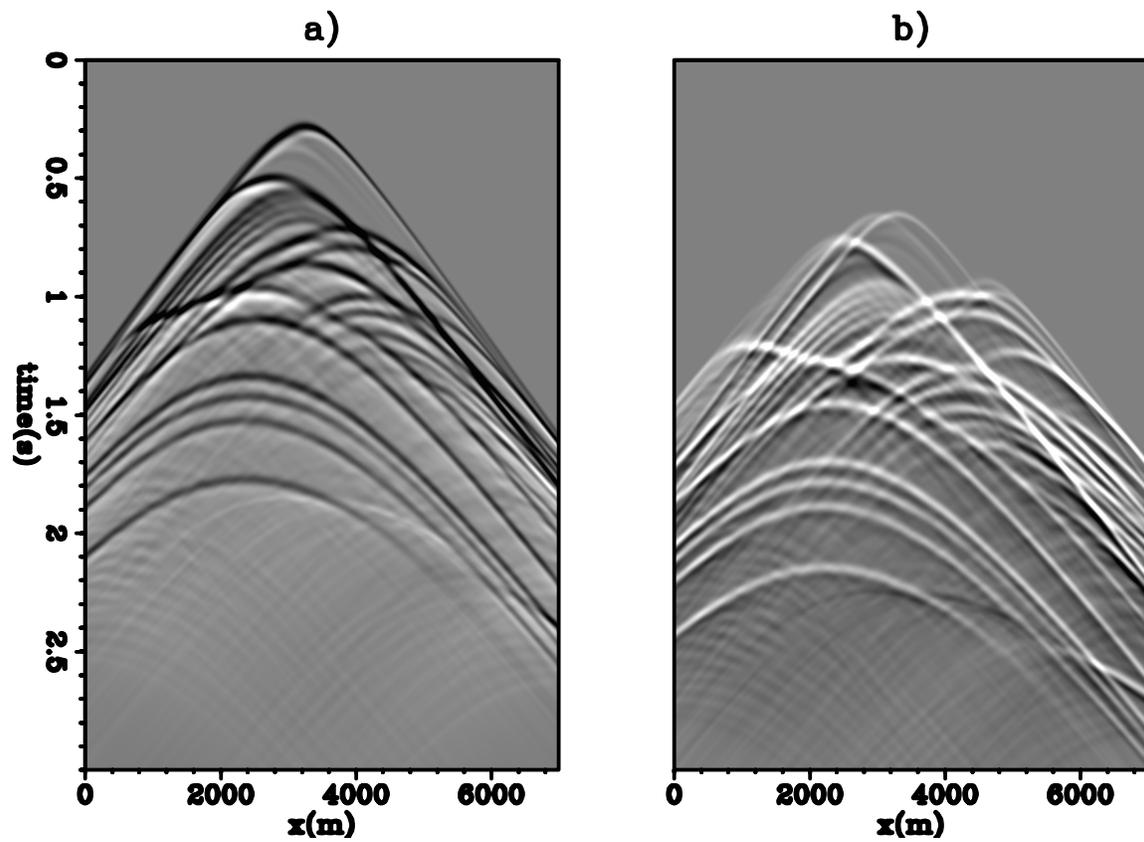


Figure 5: A common-receiver gather taken at $x=3200$ m: (a) synthetic up-going (primary) data obtained by applying \mathbf{L}_\uparrow to the model and (b) synthetic down-going (receiver ghost) data obtained by applying \mathbf{L}_\downarrow to the model. [CR] `mandy1/. modData`

Panels (a) and (b) of Figure 6 show the corresponding up-image and down-image. Comparing the two images, we can see that wider illumination is achieved by the mirror image. The benefit of the wider aperture is directly correlated with the depth of the sea-bottom. The deeper the sea-bottom, the wider the illumination. On the other hand, a close-up section of the images (Figure 7 (a) and (b)) shows that the up-image has a higher signal-to-noise ratio than the down-image. This is only a synthetic study and we expect the difference will be more apparent with a field dataset.

The goal of joint inversion is to get the best of both worlds. By manipulating two sets of data, we wish to produce a joint image that has both the wide illumination of the down-image and a high signal-to-noise ratio in the region covered by the up-image.

Joint Inversion Result

A joint inversion is performed in a least-squares sense with the objective goal described in equation 1. The initial guess is calculated by summing the up-image and the down-image. Panel (c) of Figure 6 shows the image after joint inversion (with 20 iterations). We can see an overall improvement from the migration images in panels (a) and (b) of Figure 6 to the inversion image. We have identified three areas of improvement with the close-up section shown in Figure 7:

1. In panel (c) of Figure 6, the near-surface reflector near $z=0-400$ m and $x=2400-3200$ m has a better relative amplitude and is more focused.
2. In panel (c) of Figure 6, the joint image has a wider illumination for the region from $x=5000$ m to the left and the region from $x=2000$ m to the right.
3. Figure 7 shows a higher signal-to-noise ratio for the joint image. In addition, the deeper reflector at $z=1200-1500$ m and $x=3200-4000$ m are better illuminated in the joint image than in the up-image or the down-image.

This example shows that joint inversion coherently combines information from primary and mirror signal to produce a better illuminated and resolved image. In the next section, we will explore the effect of a sparse geometry between shots and receivers.

Marmousi model with coarse sampling

In this second example, we increase the spacing of the source and receiver array. There are 22 shots with a spacing of 300m spanning from 500-6800m. We use only 10 receivers that are 400 m apart and span from 2000-5600m. The model grid has a different reflectivity as shown in panel (d) of Figure 8. The rest of the parameters are the same as in the previous example.

Panel (a) and (b) of Figure 8 show the image from migration with the primary signal and with the mirror signal. In the up-image, we can see the strong near-surface amplitude, which is characteristic of the RTM. The down-image is coarse-grained due to the large shot and receiver spacings. Panel (c) of Figure 8 shows the corresponding joint inversion image

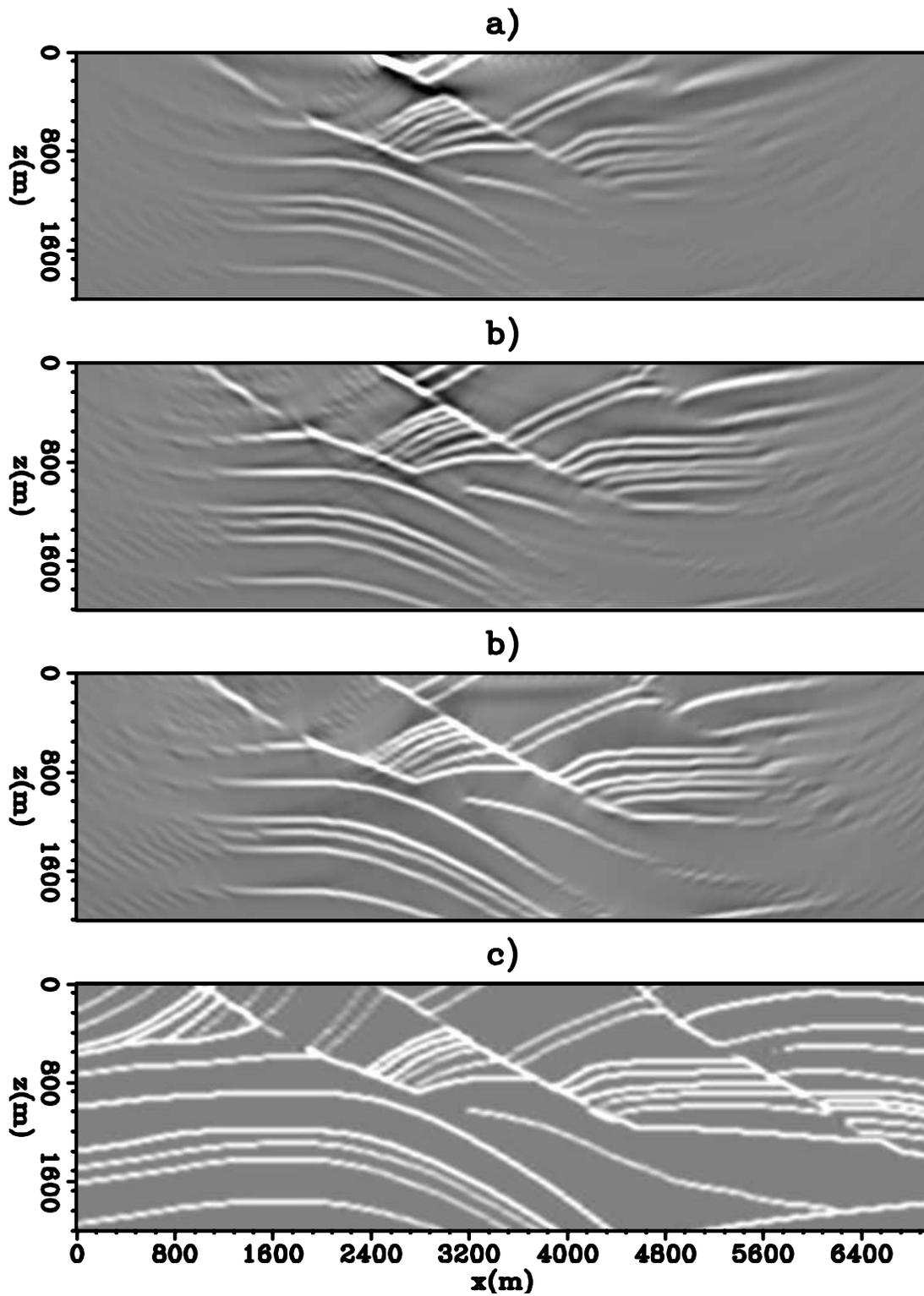


Figure 6: (a) Up-image obtained by calculating $\mathbf{L}'_{\uparrow} \mathbf{d}_{\uparrow}^{mod}$, (b) down-image obtained by calculating $\mathbf{L}'_{\downarrow} \mathbf{d}_{\downarrow}^{mod}$, (c) image obtained by joint inversion, and (d) reflectivity model. [CR]

mandy1/. RTM

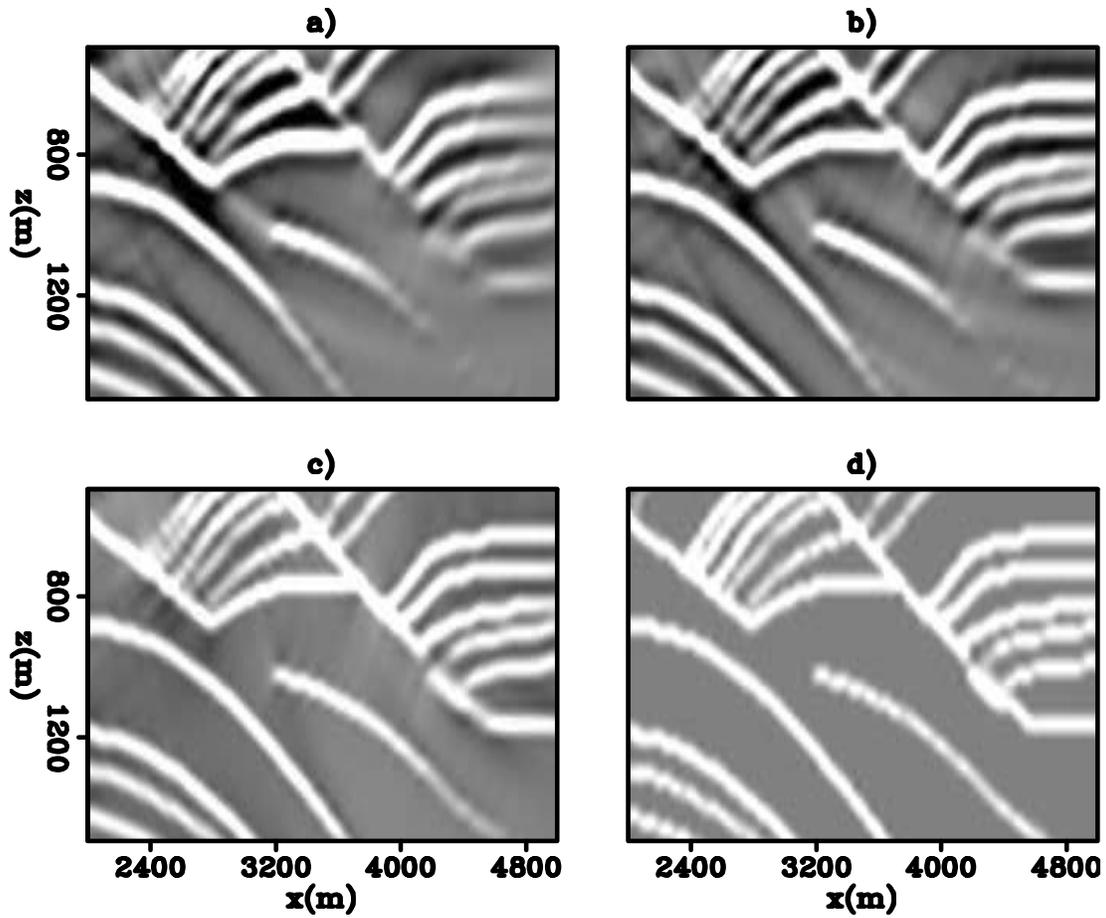


Figure 7: A section of the image cut from $x=2500-4500$ m and $y=500-1500$ m. Images are clipped at 90% to contrast the quality among (a) the up-image, (b) the down-image, (c) the joint-image, and (d) the model. [CR] `mandy1/. Dicom`

(with 20 iterations). We can see a dramatic improvement from the joint inversion. This example demonstrates that when the recorded information is limited with respect to the complexity of the subsurface, joint inversion of up- and down-going signal can provide a better illuminated and more refined image.

Figure 9 shows a close up section of Figure 8. It verifies the observation made in the previous example regarding the improvement of the joint-image.

DISCUSSION AND CONCLUSION

Free-surface multiples in OBS acquisition are often treated as noise. However, because they are formed by the same source signal as primaries but travel along different paths in the medium, they contain information not present in the primaries. To capitalize on the information provided by both the multiples and the primaries, we have developed a procedure that allows primary and mirror signals to be jointly imaged. While direct migration of the primary has limited illumination aperture, direct migration of the mirror signal is less resolved in complex areas. Joint inversion results in better illumination, improved resolution, and a more balanced amplitude image in geologically complex models.

REFERENCES

- Barr, F. J., and J. I. Sander, 1989, Attenuation of water-column reverberations using pressure and velocity detectors in a water-bottom cable: 59th SEG Annual Meeting Expanded Abstracts, 653–655.
- Dash, R., G. Spence, R. Hyndman, S. Grion, Y. Wang, and S. Ronen, 2009, Wide-area imaging from obs multiples: *Geophysics*, **74**, Q41–Q47.
- Godfrey, R., P. Kristiansen, B. Armstrong, M. Cooper, and E. Thorogood, 1998, Imaging the foynaven ghost: 68th SEG Annual Meeting Expanded Abstracts, 1333–1335.
- Grion, S., R. Exley, M. Manin, X. Miao, A. Pica, Y. Wang, P. Granger, and S. Ronen, 2007, Mirror imaging of obs data: *First Break*, **25**, 37–42.
- Ronen, S., L. Comeaux, and J. Miao, 2005, Imaging downgoing waves from ocean bottom stations: 75th SEG Annual Meeting Expanded Abstracts, 963–966.
- 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.
- Soubaras, R., 1996, Ocean bottom hydrophone and geophone processing: 66th SEG Annual Meeting Expanded Abstracts, 24–27.
- Wang, Y., S. Grion, and R. Bale, 2009, What comes up must have gone down: the principle and application of up-down deconvolution for multiple attenuation of ocean bottom data: *CSEG Recorder*, **34**, 16–20.

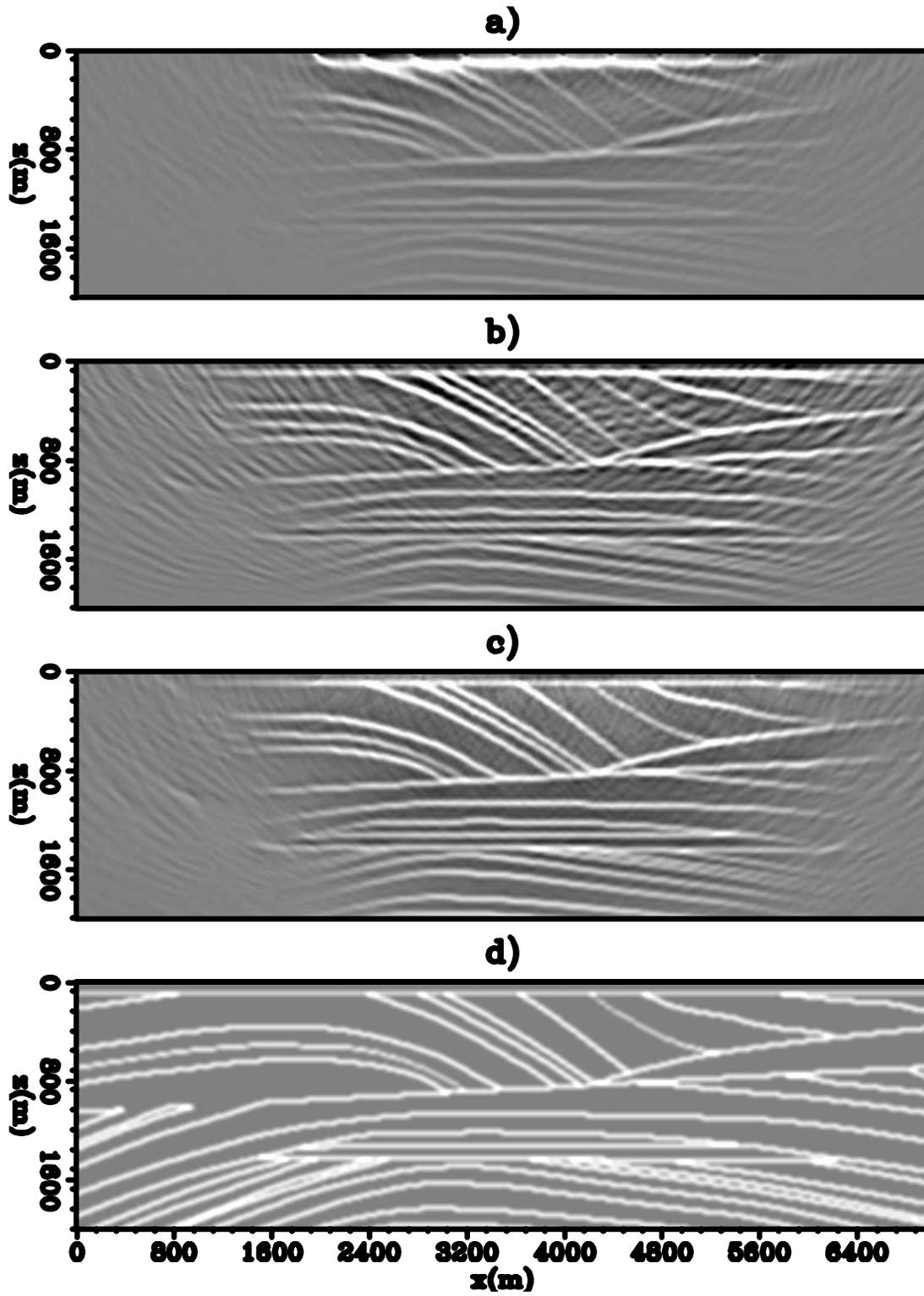


Figure 8: (a) The up-image, (b) the down-image, (c) the joint-image, and (d) the reflectivity model. [CR] [mandy1/. sparseRTM](#)

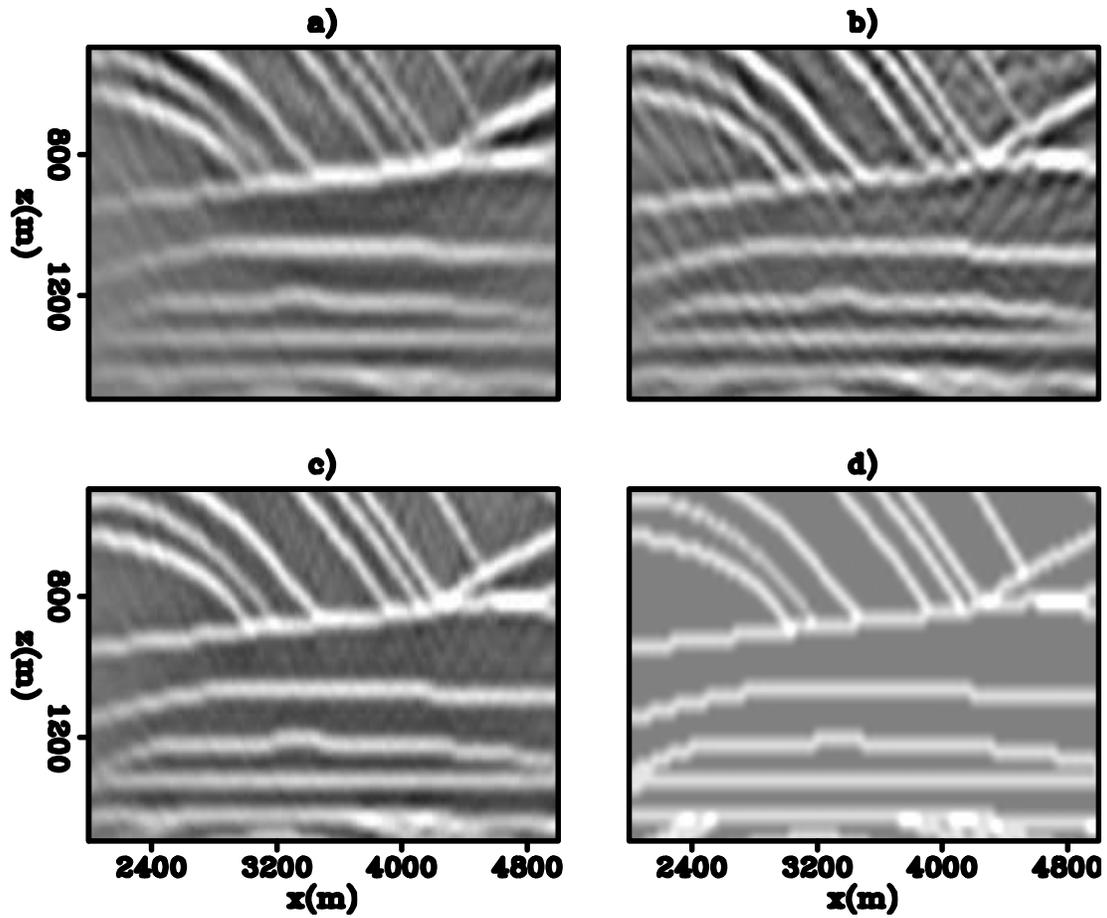


Figure 9: A close-up section of (a)The up-image (b) the down-image (c) the joint-image, and (d) the reflectivity model with coarse sampling. The sections are cut from $x=2000-5000$ m and $y=500-1500$ m. [CR] `mandy1/. sparseDeep`

More fun with random boundaries

Robert G. Clapp

ABSTRACT

Transferring wave-field checkpoints from disk to the compute engine is often a bottleneck in Reverse Time Migration (RTM). The need to create wave-field checkpoints can be eliminated by running the wave equation backwards and choosing a time reversible boundary condition. In acoustic propagation, velocity can be made more and more random within a boundary region. Reflections from the incoherent boundaries are random in nature and produce minimal coherent correlations when applying the RTM imaging condition. In Vertical Transverse Isotropic (VTI) media, the horizontal and Normal MoveOut (NMO) velocities can be modified to cause wavefronts to propagate parallel to the boundary region causing further degradation of coherent signals. Coherent reflections can be further reduced by decreasing the randomized velocity within the boundary region.

INTRODUCTION

Time domain finite difference is the most computationally efficient algorithm for both RTM (Baysal et al., 1983) and waveform inversion (Woodward, 1990). The kernel's simplicity, Single Instruction Multiple Data (SIMD), and high level of data reuse has led to high-performance implementations on Field Programmable Gate Arrays (FPGA) (Nemeth et al., 2008) and General Purpose Graphics Processing Units (GPGPU) (Micikevicius, 2008) along with conventional CPUs.

Both waveform inversion and RTM require source and receiver wave-fields at equivalent time to be correlated. Unfortunately, the source wave-field is propagated forward in time while the receiver wave-field is propagated backwards in time. As a result, the wave-fields are computed sequentially and one of the two wave-fields must be saved to disk. To avoid having to save the entire wave-field to disk, checkpointing Symes (2007) or boundary reinjection (Clapp, 2008; Dussaud et al., 2008) have been suggested. These schemes result in lower IO requirements at the cost of 50% more computation. Given that disk IO bandwidth is increasing at a much slower rate than computational power, all of these approaches seem likely to face an IO bottleneck. Clapp (2009) proposes to change the boundary condition on the computation kernel from one that damps to one that randomizes the wave-field hitting the boundary. Unlike the damping boundary, a randomizing boundary, which only modifies the velocity within the boundary region, is time reversible. As a result the source wave-field can be propagated from $t = 0$ to t_{\max} , and then the source and receiver wave-field can be propagated simultaneously backwards. This eliminates the need to write the entire wave-field to disk in RTM and requires only two time steps of the source wave-field to be saved to disk in the case of waveform inversion. Clapp (2009) showed that the RTM image produced using a random boundary was nearly identical to one using a damping boundary except at shallow depths.

In this paper, I propose several ways to improve the imaging result using a random boundary. I begin by proposing a modified random velocity boundary that not only becomes more random as energy propagates into it, but whose average velocity decreases, as a result causing larger time delays and less coherent reflections. In addition, I show for VTI propagation that if the horizontal and NMO velocities increase within the boundary region, energy can be turned parallel to the boundary, producing more randomness.

RANDOM BOUNDARY

Clapp (2009) used a simple scheme to introduce randomness into a wave-field hitting the boundary. Algorithm 1 summarizes the approach. The concept behind the increasing randomness was to draw in as much of the wavefront energy as possible into an ever-increasing chaotic system. Figure 1 shows several snapshots of the wave-field entering the randomized zone. Note how a very small portion of the wave-field is reflected from the beginning of the boundary and how chaotic the energy leaving the boundary ends up being. The end result is that energy reentering the non-boundary portion of the computational domain produces minimal coherent reflections with the receiver (or source) wave-field. Each point in the image is the result of stacking in multiple shot record migrations. We can think of the wave-field w at any time t as being the sum of the true wave-field signal s and the noise n introduced by energy leaving the random layer. By modifying the random layer at each shot, the noise changes from shot to shot (see Figure 2). By summing many shots, we significantly reduce the noise in the final image (see Figure 3).

Algorithm 1 Basic random bound construction

```

for all boundary points do
  found=false
  Closest velocity within computational domain (vel)
  Percentage into boundary region (pct)
  while found do
    Select uniform normal (rnd)
     $dev = rnd * 2.5 * dist * vel$ 
    if  $vel + dev < v_{max}$  then
      found=true
       $vel = vel + dev$ 
    end if
  end while
end for

```

Figures 1-3 are somewhat misleading because the source is in the middle of the computational domain. Figure 4 shows a more realistic scenario where the shot is at the edge of the computational domain. Note the similarity in the wave-fields at early times even with different randomized boundaries. As a result, reflections at these early times (at shallow depths) show many more artifacts than those at later times (Figure 3).

The shallow depth image can be improved by modifying the random velocity boundary. Algorithm 2 introduces a decreasing maximum velocity towards the outer edge of the computational domain. Energy now takes longer to leave the boundary creating a larger time gap and more random energy pattern between the true wave-field and the noise introduced

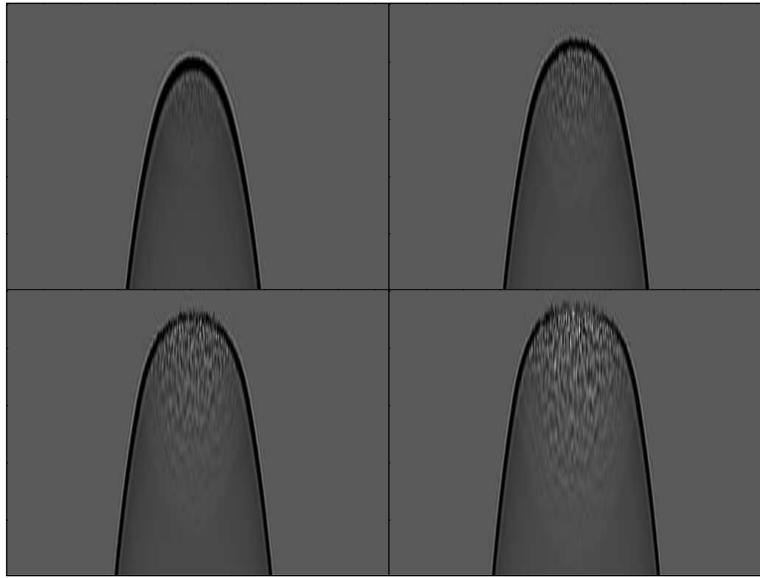


Figure 1: A zoom-in on the wave-field as it propagates into the random boundary. The order in time is top-left, top-right, bottom-left, bottom-right. Note how very little of the wavefront is perturbed at early times. [CR] bob2/. initial

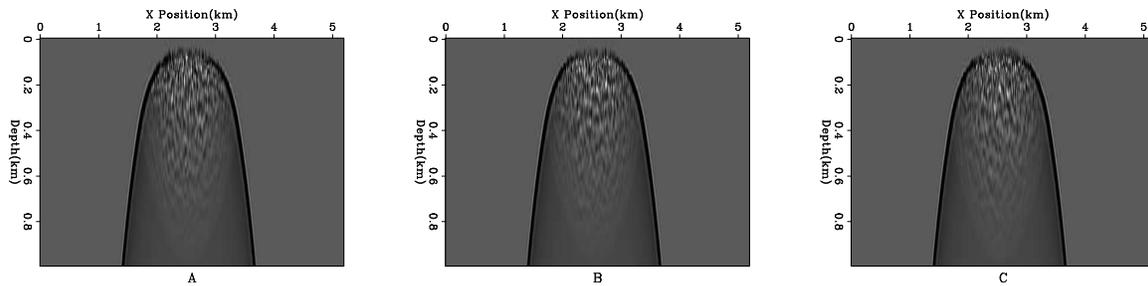
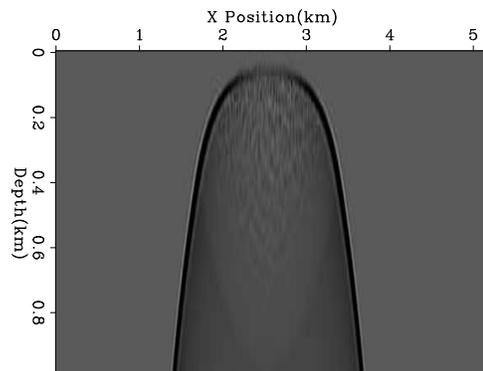


Figure 2: A zoom in on the wave-field as it passes into the random boundary. The three panels represent three different random boundaries. Note the differences in the wave-field leaving the boundary. [CR] bob2/. different

Figure 3: The result of stacking 16 different random realizations. Note how the noise introduced by the random boundary cancels out when summed. [CR] bob2/. stack1



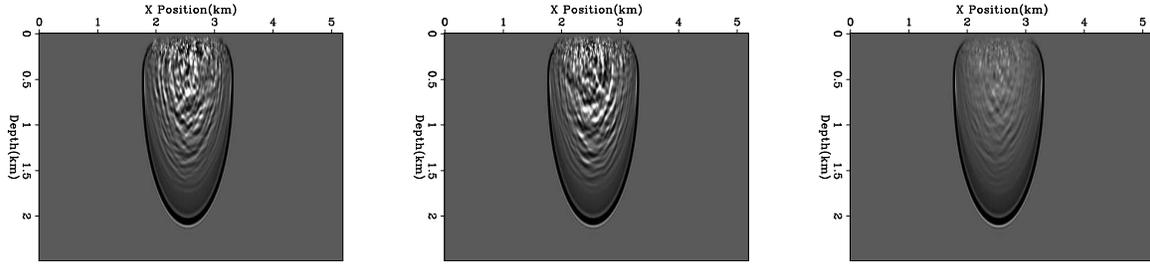


Figure 4: The left and center panels show the wave-field at the same time with two random boundaries generated using algorithm 1. The right panel shows the result of adding 16 realizations together. Note how some energy was stacked coherently. [CR] bob2/. edge

by the boundary. Figure 5 shows several snapshots with this new boundary condition. Note the increased separation in time and space (compared to the wave-fields seen in Figure 1). The increased time gap and randomness produces less noise at shallow depths (Figure 6).

Algorithm 2 Basic random bound construction

```

for all boundary points do
  found=false
  Closest velocity within computational domain (vel)
  Percentage into boundary region (pct)
  while found do
    Select uniform normal (rnd)
     $dev = rnd * 1.3 * dist * vel$ 
    if  $vel + dev < v_{max} * (1. - pct) * .4$  then
      found=true
       $vel = vel + dev$ 
    end if
  end while
end for

```

VTI RANDOM BOUNDARIES

Both Clapp (2009) and the previous section describe how to create a boundary for acoustic propagation. VTI propagation is slightly more complicated. We start by defining the horizontal velocity v_h , the vertical velocity v_v , and the Normal MoveOut velocity v_n , defined in terms of Thompson parameters as

$$v_n = v_v(1 + 2\delta)^{\frac{1}{2}} \quad (1)$$

and

$$v_h = v_v(1 + 2\epsilon)^{\frac{1}{2}}. \quad (2)$$

In acoustic finite difference, a second derivative in x, y , and z ($p_{d2x}, p_{d2y}, p_{d2z}$) is calculated at all ix, iy , and iz at a given time step it of the wave-field p . The wave-field at the next

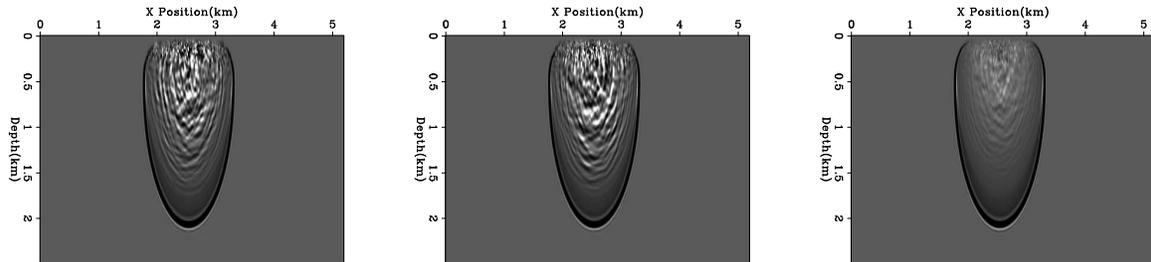


Figure 5: The left and center panels show the wave-field at the same time with two random boundaries generated using algorithm 2. The right panel shows the result of adding 16 realizations together. Note how some energy was stacked coherently. Note the larger gap in time between the main wavefront and the energy generated from the random boundary. Further note the decreased energy in the result of stacking multiple realizations. [CR]

bob2/. decrease

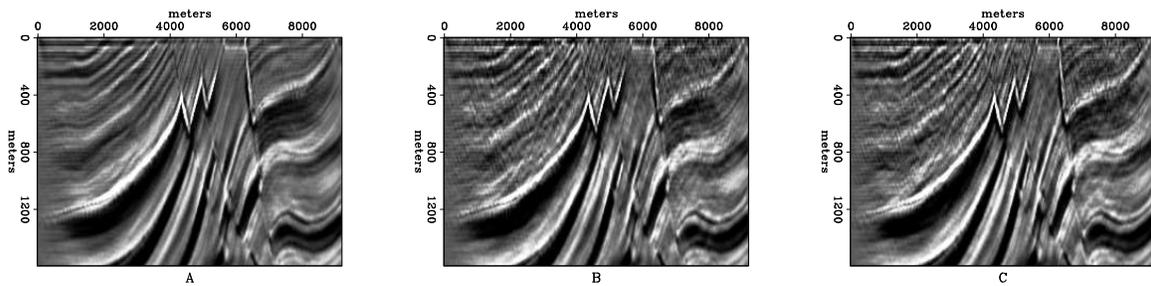


Figure 6: The result of migrating using a damped boundary condition(A), the boundary condition described in algorithm 1 (B), and the boundary condition described in algorithm 2 (C). Note how the shallow reflectors are preserved in (C). [CR]

bob2/. improved

time step is then calculated using these derivatives and the sampling in time dt , along with a source term s and the previous value of the wave-field. The computational kernel becomes

$$\begin{aligned} p(iz, ix, iy, it + 1) &= p(iz, ix, iy, it + 1) - p(iz, ix, iy, it - 1) + 2p(iz, ix, iy, it) \\ &+ v(iz, ix, iy)dt^2(p_{d2x} + p_{d2y} + p_{d2z}). \end{aligned} \quad (3)$$

Following the approach in Alkhalifah (2000) involves an auxiliary wave-field q ; Derivatives in x and y are calculated on p and derivatives in z are calculated on q (q_{d2z}). The computation kernel then becomes

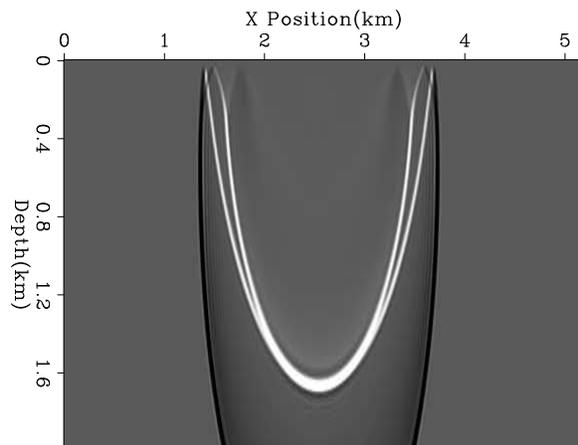
$$\begin{aligned} p(iz, ix, iy, it + 1) &= p(iz, ix, iy, it + 1) - p(iz, ix, iy, it - 1) + 2p(iz, ix, iy, it) \\ &+ dt^2(v_h(iz, ix, iy)(p_{d2x} + p_{d2y}) + v_v(iz, ix, iy)q_{d2z}) \end{aligned} \quad (4)$$

and

$$\begin{aligned} q(iz, ix, iy, it + 1) &= s(iz, ix, iy, it + 1) - q(iz, ix, iy, it - 1) + 2q(iz, ix, iy, it) \\ &+ dt^2(v_n(iz, ix, iy) * (p_{d2x} + p_{d2y}) + v_v(iz, ix, iy)q_{d2z}). \end{aligned} \quad (5)$$

Randomizing v_v , v_h , and v_n independently creates an unstable system. The most straightforward way to add stable random boundaries to the VTI problem is to create an acoustic random boundary layer by setting ϵ and δ to 0 (therefore $v_v = v_h = v_n$). A better strategy is to take advantage of the extra flexibility of having three parameters describing moveout. The longer a wavefront travels through the randomized layer, the more chaotic the resulting wave-field and the longer the delay between the true signal and the beginning of noise. By increasing ϵ and δ while decreasing v_v we can cause the wave-field to turn parallel to the random boundary. To see this effect, I have removed the random component of the boundary layer while still decreasing v_v and increasing ϵ and δ . Figure 7 shows the result of overlaying two wave-fields, one using an isotropic and one using anisotropic boundary. The anisotropic boundary results energy traveling longer in the boundary region. Figure 8 shows the randomized wave-field at several different time steps using both an isotropic and anisotropic boundary condition. Note how the noise pattern using the anisotropic boundary condition is much less regular.

Figure 7: The result of overlaying the wave-fields using an isotropic and anisotropic boundary. The anisotropic boundary results in energy traveling longer in the boundary region. bob2/. turn



CONCLUSIONS

By reducing IO requirements, random velocity boundaries have computational advantages for correlation-based imaging conditions, compared to damping boundary conditions. They

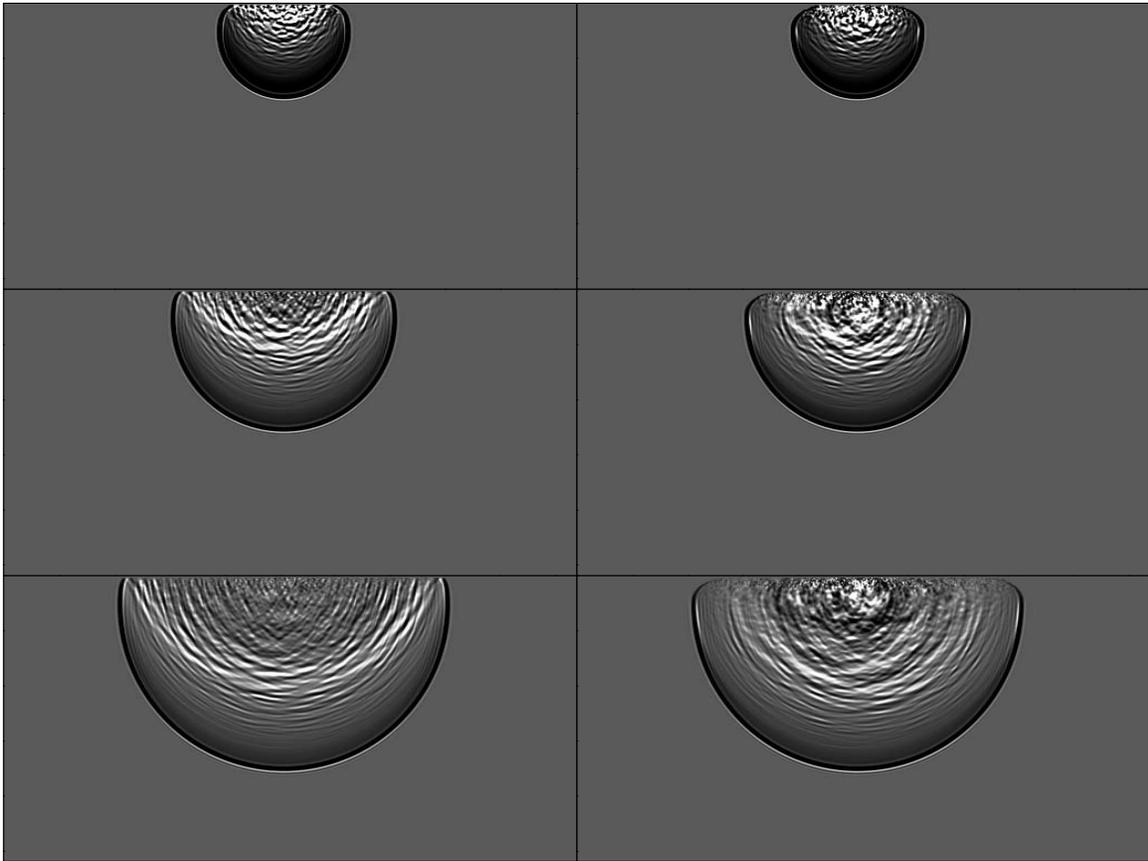


Figure 8: The left panels show a wave-field at several different time steps using an isotropic boundary condition. The right panel shows the wave-field using an anisotropic boundary condition. [CR] bob2/. vti

can be further improved by systematically decreasing the mean velocity in the boundary. In VTI modeling/migration, the boundaries' effectiveness can be improved by systematically modifying δ and ϵ .

REFERENCES

- Alkhalifah, T., 2000, An acoustic wave equation for anisotropic media: *Geophysics*, **65**, 1239–1250.
- Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: *Geophysics*, **48**, 1514–1524.
- Clapp, R. G., 2008, Reverse time migration: Saving the boundaries: 2008, **136**, 136–144.
- , 2009, Random boundaries: SEP-Report, **138**, 29–38.
- Dussaud, E., W. W. Symes, L. Lemaistre, P. Singer, B. Denel, and A. Cherrett, 2008, Computational strategies for reverse-time migration: 78th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, SPMI 3.3.
- Micikevicius, P., 2008, 3d finite difference computation on gpus using cuda: 2nd Workshop on General Purpose Processing on Graphics Processing Units, Expanded Abstracts, 79–84.
- Nemeth, T., J. Stefani, W. Liu, R. Dimond, O. Pell, , and R. Ergas, 2008, An implementation of the acoustic wave equation on fpgas: 78th Ann. Internat. Meeting, Expanded Abstracts, 2874–2877, Soc. Expl. Geophys.
- Symes, W. M., 2007, Reverse time migration with optimal checkpointing: *Geophysics*, **72**, SM213–SM221.
- Woodward, M., 1990, Wave equation tomography: **62**.

Implicit finite difference in time-space domain with the helix transform

Ohad Barak

ABSTRACT

Spectral factorization is a method of creating causal filters which have causal inverses. I use spectral factorization of an implicit finite-difference stencil of the two-way wave equation approximation in order to model wave propagation by a sequence of deconvolutions. I deconvolve this filter's coefficients with the wavefield propagating in a constant velocity medium using the helix approach. In comparison with explicit approximations, implicit approximations have unconditional stability, enabling the use of larger time steps during the modeling process. The advantages are both in reduced computation time, and in the extension and scalability to multiple dimensions enabled by the helix operator.

INTRODUCTION

Implicit finite difference is a widely used method in geophysical data processing, commonly utilized for the approximation of the differential wave equation when extrapolating wavefields. In comparison with explicit methods, implicit finite difference has unconditional numerical stability, thus enabling larger finite differencing steps during the computation. This is an attractive prospect, as the implication is a shorter processing time for wave extrapolation. However, implementing implicit finite difference in a multidimensional problem is not trivial. The method requires the solution of a sparse set of linear equations per each propagation step. The cost of solving these linear equations, and the computational complexity required, becomes unreasonable at anything greater than 2 dimensions. It is possible to split the solver so that only one dimension of the problem is computed at each propagation step, thus reducing the complexity and possibly the amount of resources required for the computation. However, this method may introduce azimuthal anisotropy to the solution if the actual differential equation being solved is non-separable.

Two concepts combine to greatly aid us in this matter. The first is the helix approach, envisaged in Claerbout (1997). The helix effectively enables us to treat multidimensional problems as one dimensional problems. Specifically, it enables execution of multidimensional convolutions as 1-D convolutions, and likewise for deconvolutions. Convolution equates to polynomial multiplication, while deconvolution equates to polynomial division. The application of convolution or deconvolution to a data set is likened by Claerbout to winding a coil (the filter coefficients) around the data, where the data is treated as a long set of traces combined end-to-end along their fast axis, as shown in Figure 1.

The second is the concept of spectral factorization. The purpose of spectral factoriza-

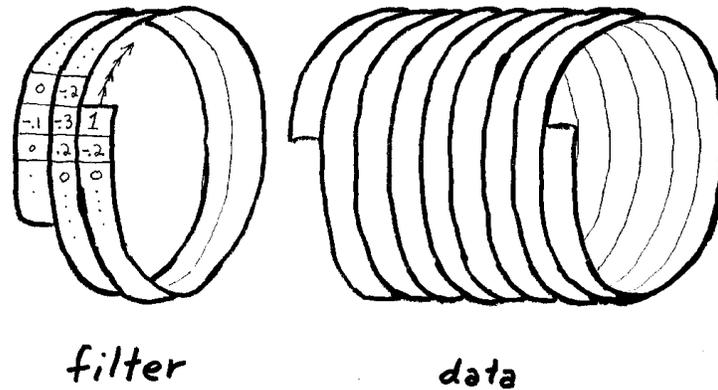


Figure 1: Sketch of the helix concept - convolution takes place by winding a "coil" of filter coefficients over a "coil" of data values (Claerbout (1997)) [NR] [ohad1/. coil](#)

tion is to input a series of coefficients, and create an alternate set of causal filter coefficients which have a causal inverse. The result will usually be a minimum-phase filter. The autocorrelation of this new set of filter coefficients recreates the original values of the input series. The upshot of this is that application of the original series' coefficients to a dataset is akin to convolving the data with the spectrally factorized filter coefficients in one direction, and then convolving again in the other direction ("coiling" and then "uncoiling" the filter coefficients over the data). This effectively applies the filter and its time reverse (adjoint) to the data, which amounts to multiplying the data by the original input series' coefficients. In the case of finite differencing, the "input" series might be the Laplacian, which when made to traverse over the data has the effect of a 2nd derivative approximation.

The spectral factorization concept enables us to represent a finite-difference operator as a forward and reverse convolution of filter coefficients. The helix concept disconnects us from the dimensionality of the problem, and enables simple application of 1D convolution and deconvolution to multidimensional problems. Together they enable an alternate method of propagating wavefields - by treating the finite-difference solution as a set of convolutions and deconvolutions.

My aim is to use the helix transform to propagate wavefields in the time-space domain, using an implicit finite-difference approximation of the 2-way acoustic wave equation. For this purpose, I formulate the proper implicit finite-difference weights with regard to the order of the difference approximation and the dimensionality of the problem, and use spectral factorization to create a causal filter with a causal inverse, whose convolution will equal those coefficients. These filter coefficients are then applied to the wavefield by deconvolution, using the helical coordinate system.

Previous implementations of the helix

Fomel and Claerbout (1997) show how spectral factorization of an implicit finite-difference scheme and application of the resulting filter coefficients with the helix operator can re-

cursively solve the heat conduction equation. They also present an implicit helix-based 3D velocity continuation method for post-stack data.

In Rickett et al. (1998) the one-way wave equation is iteratively solved in the frequency-space domain to downward-continue a wavefield. The extrapolation is robust and efficient. The shortcoming is that since factorization is done for coefficients in the frequency-space domain, the velocity is factorized into the resulting filter coefficients. This forces a workaround for dealing with lateral velocity variations. The suggested method is to create several filters, each with a reference velocity, to be applied to different parts of the wavefield.

EXPLICIT FINITE DIFFERENCE IN 2D USING SPECTRAL FACTORIZATION AND HELICAL COORDINATES

In order to test the general methodology and the programming modules, I first tested whether an explicit 2D finite-difference approximation in time-space domain done by spectral factorization and the helix transform properly emulates “standard” finite difference (i.e. using a differencing “star” which traverses over the data). The difference approximation was 2nd order in time and 2nd order in space:

$$\frac{P_{x,z}^{t+1} - 2P_{x,z}^t + P_{x,z}^{t-1}}{\Delta t^2} = C^2 \left(\frac{P_{x+1,z}^t - 2P_{x,z}^t + P_{x-1,z}^t}{\Delta x^2} + \frac{P_{x,z+1}^t - 2P_{x,z}^t + P_{x,z-1}^t}{\Delta z^2} \right) \quad (1)$$

where P is the pressure wavefield, t , x and z are the time and space coordinate indices, and Δt , Δx , and Δz are the temporal and spatial step sizes.

The 2D laplacian representing the spatial derivative’s coefficients is:

$$\nabla^2 \approx \begin{pmatrix} & & 1 & & \\ & 1 & -4 & 1 & \\ & & & & \\ & & & & 1 \\ & & & & \end{pmatrix} \quad (2)$$

In helical coordinates, the laplacian is represented as a sparse 1-D array:

$$\nabla^2 \approx (1 \quad \dots \quad 1 \quad -4 \quad 1 \quad \dots \quad 1) \quad (3)$$

The dots represent the number of elements required for the wrap-around of the helix along the “fast” axis of the 2D grid (see Figure 1). The greater the number of coefficients used by the spectral factorization, the closer the convolution of those coefficients will be to the desired operator, in this case the laplacian in Equation 3. In this experiment, I used the SEPlib `wilson` module to factorize the laplacian into 28 coefficients. The following is a partial selection of those coefficients:

$$h = \begin{pmatrix} & & & & 1.0 & -0.363 & -0.0024 & -0.0013 & \dots & \\ \dots & -0.0024 & -0.0048 & -0.113 & -0.3114 & & & & & \end{pmatrix}.$$

Convolving this series with it’s adjoint results in the following coefficients:

$$h' * h = \left(0.9999771 \quad \cdots \quad 0.9999771 \quad -4.0000000 \quad 0.9999771 \quad \cdots \quad 0.9999771 \right) \approx \nabla^2$$

The remaining coefficients at close proximity to the laplacian coefficients were of 4 or more magnitudes smaller than the ones displayed. The rest were all zeros. This led me to the conclusion that 28 coefficients are sufficient.

I used the SEPlib `helicon` module to convolve the coefficients in (4) with the data, coiling first in one direction, and then uncoiling in the other direction. The `helicon` module implements convolution by the linear operator:

$$y_k = x_k + \sum_{i=1}^{N_a} a_i x_{k-i}, \quad (4)$$

where x is the input data vector, a is a causal filter of length N_a , and y is the output vector. The module also implements the adjoint:

$$x_k = y_k + \sum_{i=1}^{N_a} a'_i y_{k+i}, \quad (5)$$

where a' is the time reverse of filter a .

The application of the forward and reverse convolution replicates the act of using the standard finite differencing “star” to approximate the spatial derivative. The spatial derivative is then used to estimate the temporal derivative, and forward propagate in time the values of the acoustic wavefield.

With reference to Equation 1, the propagation kernel is:

$$P_{x,z}^{t+1} = 2P_{x,z}^t - P_{x,z}^{t-1} + (h' * h)P_{x,z}^t. \quad (6)$$

A comparison of propagation of a source in the center of a 2D acoustic wavefield with regular finite differencing and with the helix derivative is shown in Figure 2.

Although qualitatively the two results are similar, the wavefield created by the helix spatial derivative displays some form of dispersion both leading and trailing the main wavelet. Furthermore, the wavefront itself and the dispersion appear to be propagating anisotropically, with a slightly faster component along the positive diagonal. A possible explanation of this outcome is that the filter operator which convolves the data does not have a symmetric angular coverage. Referring to Figure 3, the better coverage of updip angles is apparent in comparison with downdip angles, as the filter is coiled around the data (in the direction of the arrow). If this is indeed the cause of dispersion, one way it can be mitigated is by using a cube finite differencing stencil (Haohuan et al. (2009)).

Another problem which is not apparent in Figure 2 is the wrap-around effect of the helix operator on the data being propagated. Since the helix wraps the filter convolution

around one of the dimensions, it will invariably mix the information from one side of the wavefield with that from the other side. The longer and more accurate the filter will be, the more pronounced this effect will be as well. There is no immediate remedy for this inherent property of the helix, except for the standard model-padding solution. However, since the general intention is to use the helix for 2-way wavefield propagation in reverse time migration, one way the wrap around effect can be disregarded is by utilizing random boundaries in the wavefield, as shown in Clapp (2009).

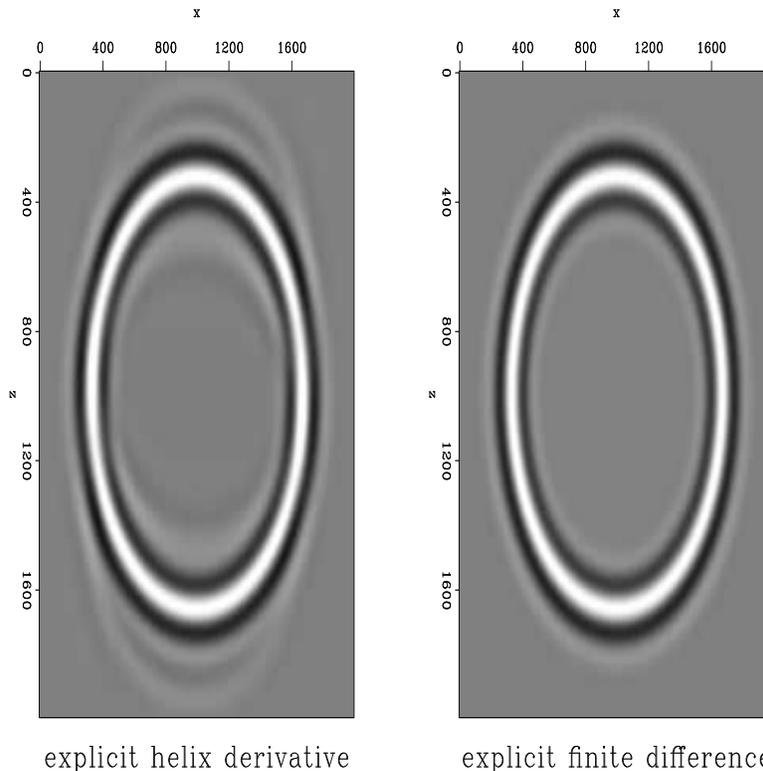


Figure 2: comparison of 2nd order in time and space finite difference Vs. helix derivative after spectrally factorizing the coefficients of the 2nd order spatial derivative and convolving them with the data using helical boundaries. The propagation here is with constant velocity $v = 3000m/s$. [ER] `ohad1/. exp-vs-hel-snap`

EXPLICIT VS. IMPLICIT FINITE DIFFERENCE APPROXIMATION OF THE 2-WAY ACOUSTIC WAVE EQUATION

Formulation for 1 dimension

The two-way acoustic wave equation in one dimension reads:

$$\frac{\partial^2 P}{\partial t^2} = C^2 \frac{\partial^2 P}{\partial x^2}. \quad (7)$$

In order to formulate it as an implicit finite-difference approximation, I first looked to the formulation of the implicit finite-difference operator of the 1-way wave equation in frequency-

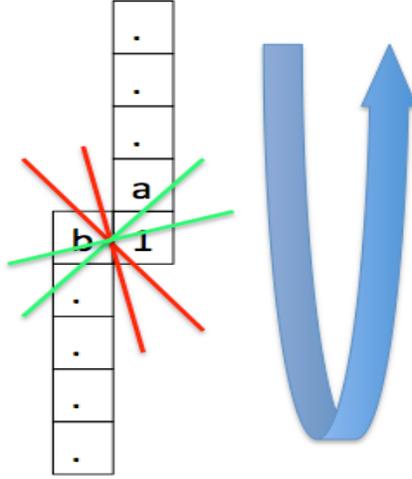


Figure 3: Angular anisotropy of helical filter operator [NR] `ohad1/. helical-aniso`

space domain in Claerbout (2009) Chapter 9. The Crank-Nicolson differencing method is used to create an implicit finite-difference approximation for downward-continuation of a wavefield:

$$\frac{P_{z+1}^x - P_z^x}{\Delta z} = \frac{v}{-i\omega 2} \left(\frac{P_z^{x+1} - 2P_z^x + P_z^{x-1}}{2\Delta x^2} + \frac{P_{z+1}^{x+1} - 2P_{z+1}^x + P_{z+1}^{x-1}}{2\Delta z^2} \right). \quad (8)$$

The Crank-Nicolson method achieves better accuracy by balancing the 2nd derivative between the current wavefield values (at depth z) and at the next as-of-yet unknown wavefield values (at depth $z+1$). Hence the division by 2 in the denominators of Equation 8. Borrowing from this methodology, I attempted to formulate a scheme which balances the values of the wavefield at known and unknown locations. Since Equation 7 has a second derivative on the left hand side (as opposed to first derivative in Equation 8), it implies that this balancing must be done over three time “locations”. In one dimension, these considerations led me to the following approximation:

$$\frac{P_x^{t+1} - 2P_x^t + P_x^{t-1}}{\Delta t^2} = \frac{C^2}{3\Delta x^2} [(P_{x+1}^{t+1} - 2P_x^{t+1} + P_{x-1}^{t+1}) + (P_{x+1}^t - 2P_x^t + P_{x-1}^t) + (P_{x+1}^{t-1} - 2P_x^{t-1} + P_{x-1}^{t-1})]. \quad (9)$$

Note that the division by 3 effectively averages the spatial derivative between the three time steps: $t-1$, t and $t+1$. In order to propagate the wavefield, the values of the the wavefield at time $t+1$ must be equated to the values at times t and $t-1$.

Application of spectral factorization and the helix operator

The linear system which must be solved according to Equation 9 has the form:

$$\begin{pmatrix} U_0 & U_1 & 0 & 0 \\ U_1 & U_0 & U_1 & 0 \\ 0 & U_1 & U_0 & U_1 \\ 0 & 0 & U_1 & U_0 \end{pmatrix} \begin{pmatrix} P_1^{t+1} \\ P_2^{t+1} \\ P_3^{t+1} \\ P_4^{t+1} \end{pmatrix} = \begin{pmatrix} V_0 & V_1 & 0 & 0 \\ V_1 & V_0 & V_1 & 0 \\ 0 & V_1 & V_0 & V_1 \\ 0 & 0 & V_1 & V_0 \end{pmatrix} \begin{pmatrix} P_1^t \\ P_2^t \\ P_3^t \\ P_4^t \end{pmatrix} + \begin{pmatrix} W_0 & W_1 & 0 & 0 \\ W_1 & W_0 & W_1 & 0 \\ 0 & W_1 & W_0 & W_1 \\ 0 & 0 & W_1 & W_0 \end{pmatrix} \begin{pmatrix} P_1^{t-1} \\ P_2^{t-1} \\ P_3^{t-1} \\ P_4^{t-1} \end{pmatrix}. \quad (10)$$

For simplicity, we can combine all the constants into one: $\alpha = \frac{C^2 \Delta t^2}{3\Delta x^2}$. The matrix coefficients in Equation 10 (the finite-difference weights) are then:

$$\begin{aligned} U_0 &= 1 + 2\alpha, & U_1 &= -\alpha; \\ V_0 &= 2 - 2\alpha, & V_1 &= \alpha; \\ W_0 &= -1 - 2\alpha, & W_1 &= \alpha. \end{aligned}$$

In shorter notation, Equation 10 reads:

$$UP^{t+1} = VP^t + WP^{t-1}. \quad (11)$$

The solution of this linear system is:

$$P^{t+1} = U^{-1}(VP^t + WP^{t-1}). \quad (12)$$

To solve this system, we must perform polynomial division. The system is tridiagonal (and easily solvable) only for 1 dimension. For multiple dimensions, matrix U is block diagonal. Additional non-zero elements appear at a certain offset from the diagonal, making the solution process more complicated. However, using spectral factorization, the finite-difference weights of matrix U can be factorized into a set of causal filter coefficients u and its time reverse u^T . Using the helical approach to deconvolution, the system can be recast as:

$$P^{t+1} = (u^T u)^{-1} (VP^t + WP^{t-1}); \quad (13)$$

$$P^{t+1} = u^{-1}(u^T)^{-1} (VP^t + WP^{t-1}). \quad (14)$$

As stated above, polynomial division is equal to deconvolution. This means that the polynomial division in Equation 12 can be achieved by a set of two deconvolutions of the spectrally factorized coefficients u of matrix U . One deconvolution is done along the data in the reverse direction (application of the adjoint of the filter):

$$y_k = x_k - \sum_{i=1}^{N_a} a'_i y_{k-i}, \quad (15)$$

where a' is the time reversed filter coefficients of u . The other deconvolution is done in the forward direction:

$$x_k = y_k - \sum_{i=1}^{N_a} a_i x_{k-i}. \quad (16)$$

I used the SEPlib module `polydiv`, which uses the helical coordinates to do the deconvolutions (the polynomial division) in equations 15 and 16.

The wavefield propagation is done by the following sequence:

1. Spectrally factorize the coefficients of matrix U .
2. Multiply the saved wavefield at time $t - 1$ by the coefficients of matrix W .
3. Multiply the saved wavefield at time t by the coefficients of matrix V .
4. Sum the results of the previous 2 steps into a result vector.
5. Uncoil the factorized coefficients u over the result vector (eq. 15).
6. Coil the factorized coefficients u over the result vector (eq. 16).

This sequence is repeated for each time step.

Formulation for 2 dimensions

The finite-difference weights for the 2-D case of matrix U (where $\Delta x = \Delta z$) are derived by using the same finite-difference approximation (2nd order in time and space) as in Equation 9. The derived linear equation system is similar to the one shown in (10), except that two off-diagonal bands appear at a certain offset from the main diagonal. The offset is equal to the number of elements of the “fast” axis of the 2D wavefield.

$$\begin{aligned} \begin{pmatrix} U_0 & U_1 & \dots & U_1 \\ U_1 & U_0 & U_1 & \dots \\ \dots & U_1 & U_0 & U_1 \\ U_1 & \dots & U_1 & U_0 \end{pmatrix} \begin{pmatrix} P_1^{t+1} \\ P_2^{t+1} \\ P_3^{t+1} \\ P_4^{t+1} \end{pmatrix} &= \begin{pmatrix} V_0 & V_1 & \dots & V_1 \\ V_1 & V_0 & V_1 & \dots \\ \dots & V_1 & V_0 & V_1 \\ V_1 & \dots & V_1 & V_0 \end{pmatrix} \begin{pmatrix} P_1^t \\ P_2^t \\ P_3^t \\ P_4^t \end{pmatrix} \\ &+ \begin{pmatrix} W_0 & W_1 & \dots & W_1 \\ W_1 & W_0 & W_1 & \dots \\ \dots & W_1 & W_0 & W_1 \\ W_1 & \dots & W_1 & W_0 \end{pmatrix} \begin{pmatrix} P_1^{t-1} \\ P_2^{t-1} \\ P_3^{t-1} \\ P_4^{t-1} \end{pmatrix}, \quad (17) \end{aligned}$$

where the finite-difference weights are:

$$\begin{aligned} U_0 &= 1 + 4\alpha, & U_1 &= -\alpha; \\ V_0 &= 2 - 4\alpha, & V_1 &= \alpha; \end{aligned}$$

$$W_0 = -1 - 4\alpha, \quad W_1 = \alpha,$$

and where $\alpha = \frac{C^2 \Delta t^2}{3\Delta x^2}$.

The propagation methodology is the same as shown in Equation 14 - creation of a solution vector, and then two deconvolutions using the SEPlib `polydiv` module.

The main advantage in comparison to a standard linear equation system solver is in the reduced number of operations required to propagate the wavefield by one time step. The computation time scales linearly with the size of the data and the number of factorized filter coefficients, instead of the square of the size of the data. Coupled with the larger time steps which are made possible by an implicit finite-difference scheme, this methodology has the potential to reduce total processing times for wavefield propagation algorithms. The further advantage is the separation of the solution from the dimensionality of the problem.

Validity tests of derived coefficients

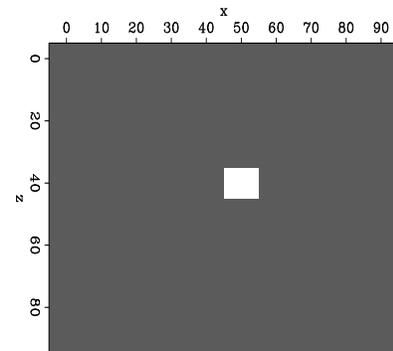
Before implementing the methodology discussed above, I tested whether the finite-difference weights derived in Equation 10 for the implicit approximation are indeed reversible by a set of convolution and deconvolution operators. The input impulse used for the tests is shown in Figure 4. The sequence of operations applied to this impulse were:

1. forward deconvolution with the spectrally factorized coefficients of the above finite-difference weights.
2. adjoint deconvolution.
3. adjoint convolution.
4. forward convolution.

The sequence and it's results are shown in Figure 5. The final result is not a perfect impulse, however the remaining values in the wavefield are 6 or more orders of magnitude smaller than the central impulse.

Figure 4: Input impulse used for helical convolution test [ER]

`ohad1/. imp-test3-in`



input - impulse at center

In order to test whether the derived implicit finite-difference coefficients are actually valid for wavefield propagation, the next step was to compare standard explicit finite difference to implicit propagation with these coefficients in one dimension. The implicit solution

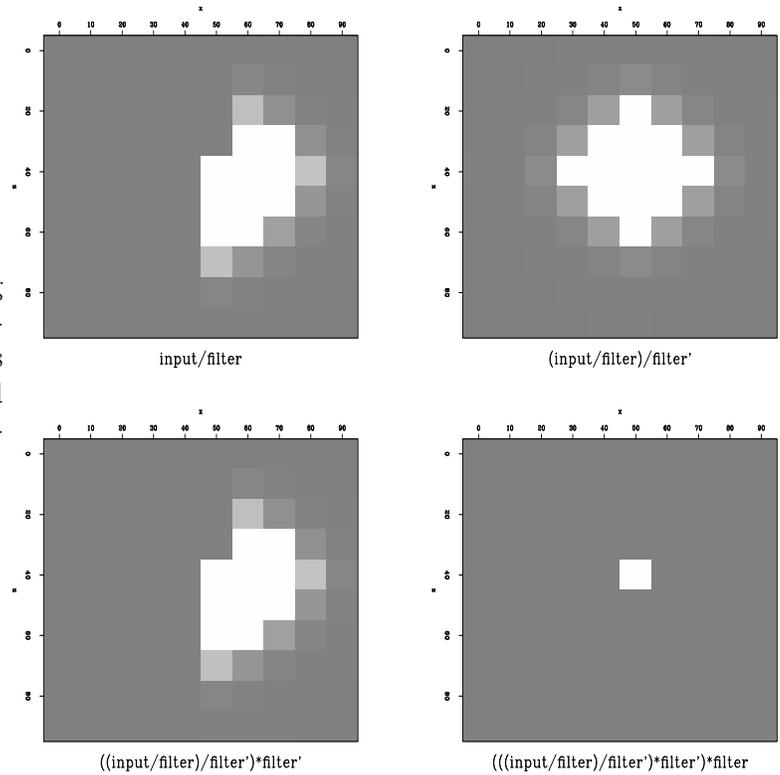


Figure 5: Results of convolving and deconvolving with filter coefficients whose convolution yields the finite-difference weights required by the implicit finite-difference approximation in Equation 17 [ER] `ohad1/. imp-test3-panels`

was done using the SEPlib module `rtris` which recursively solves a tridiagonal equation system. The results are in Figure 6. Severe dispersion is visible in the implicitly propagated wavefield. So far I have been unable to remove this dispersion without damaging the kinematics of the wavefield.

On the other hand, the wavefield derived by implicit finite difference does not diverge even when increasing the temporal time step beyond the stability limit of the explicit solver.

STANDARD IMPLICIT PROPAGATION VS. HELICAL IMPLICIT PROPAGATION

Implicit finite-difference propagation of a wavefield is the solution of the system shown in Equation 17. I tested wavefield propagation with constant velocity using a standard linear equation solver, and compared it to the wavefield propagated by repeated deconvolutions with spectrally factorized coefficients of the same finite-difference weights (Equation 14). Figure 7 shows this comparison - standard linear system solver is on the left, deconvolutions with spectrally factorized coefficients are on the right. The bottom figures were created with a larger time steps than the top ones. For the same time step size, the methods produce similar images. Both of the bottom images show greater dispersion than the top ones. This suggests that the dispersion (visible also in Figure 6) is an intrinsic property of the finite-difference weights derived by the formulation in Equation 9, and is not the result of the spectral factorization method.

Propagation with a linear equation solver is time consuming, which is why the wavefields

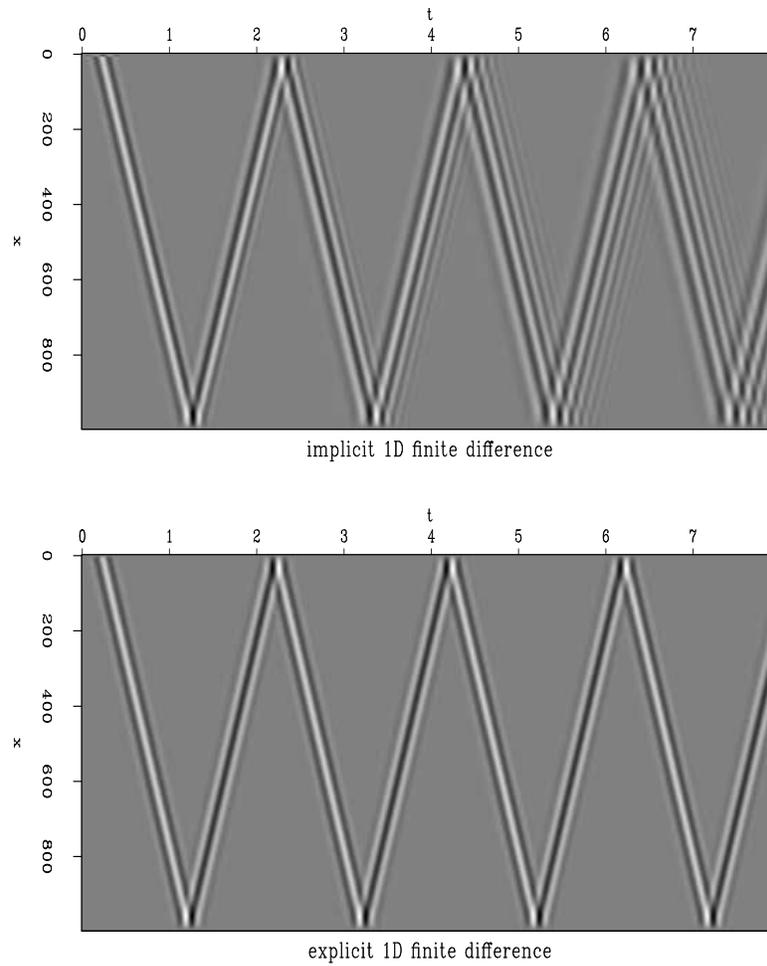


Figure 6: 1D Explicit Vs. Implicit finite difference with constant velocity = $1000m/s$. Source is a Ricker wavelet with central frequency = $12.5Hz$. The explicit time step was $\Delta t = 4msec$. The implicit time step was $\Delta t = 8msec$. $\Delta x = 10m$. [ER] [ohad1/. exp-vs-imp-1d](#)

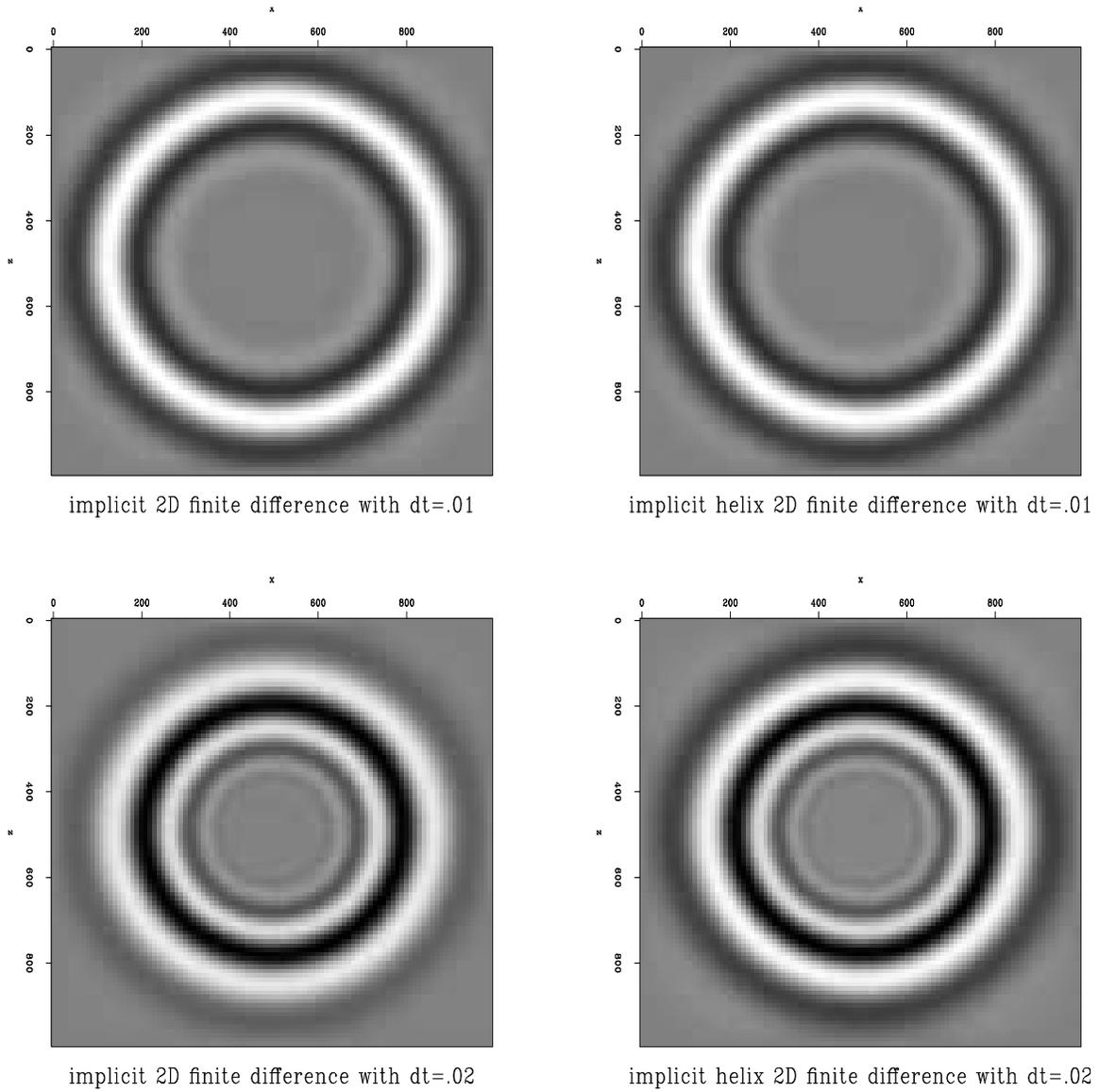


Figure 7: 2D Implicit Vs. Helical Implicit finite difference with constant velocity = $1000m/s$. Source is a Ricker wavelet with central frequency = $12.5Hz$. The time step for the top figures was $\Delta t = 10msec$, and for the bottom figures $\Delta t = 20msec$. $\Delta x = \Delta z = 10m$.

[CR] [ohad1/. imp-vs-helimp-snap](#)

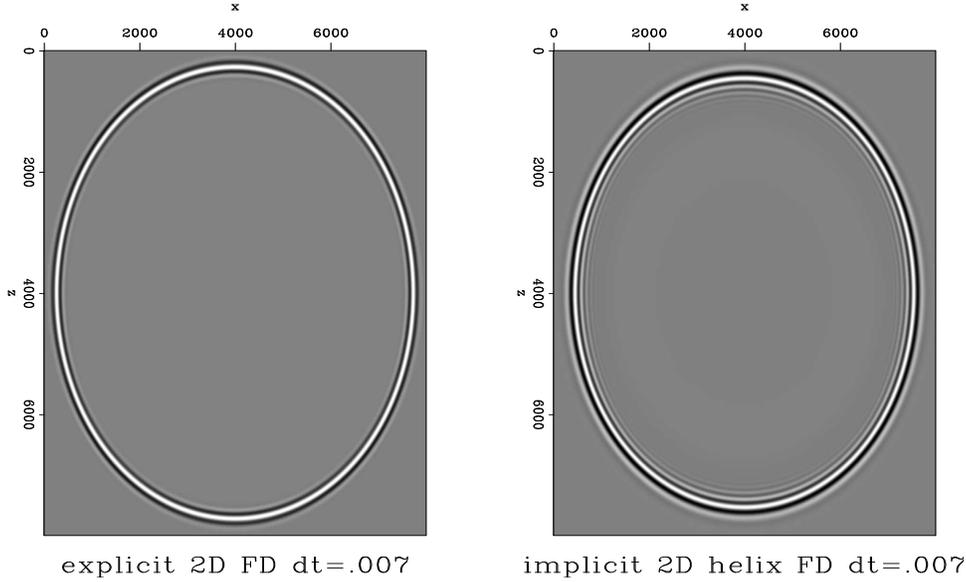


Figure 8: 2D Explicit Vs. Helical Implicit finite difference with constant velocity = $1000m/s$. Source is a Ricker wavelet with central frequency = $12.5Hz$. The time step was $\Delta t = 7msec$ - at the stability limit of the explicit scheme for the 2nd order in time and space approximation. $\Delta x = \Delta z = 10m$. [ER] [ohad1/. exp-vs-helimp-snap](#)

in Figure 7 are rather small (only 100×100 elements). To test whether this method is stable over longer periods, I compared explicit wavefield propagation to the helical implicit propagation. These are shown in Figure 8. The figure is after 4 seconds of propagation, with a time step almost equal to the stability limit of the 2nd order explicit finite-difference scheme. The explicit propagation exhibits less dispersion. However, it diverges for greater time step sizes, whereas the implicit solution does not. The dispersion problem however does get worse for the implicit solver as the time step increases.

Non-separability of velocity from factorized implicit coefficients

The initial impetus for using deconvolutions by spectrally factorized coefficients of an implicit finite-difference scheme to approximate the 2-way wave equation was that the velocity could be separated from the scheme's coefficients. The hope was that if this was possible, a filter with constant coefficients could be created to handle propagation through a variable velocity medium. This has turned out to not be the case, at least for the formulation of the implicit scheme in Equation 9. If we look at the left hand side of Equation 17:

$$\begin{pmatrix} 1 + 4\alpha & -\alpha & \dots & -\alpha & \dots \\ -\alpha & 1 + 4\alpha & -\alpha & \dots & -\alpha \\ \dots & -\alpha & 1 + 4\alpha & -\alpha & \dots \\ -\alpha & \dots & -\alpha & 1 + 4\alpha & -\alpha \\ \dots & -\alpha & \dots & -\alpha & 1 + 4\alpha \end{pmatrix} \begin{pmatrix} P_1^{t+1} \\ P_2^{t+1} \\ P_3^{t+1} \\ P_4^{t+1} \\ P_5^{t+1} \end{pmatrix}. \quad (18)$$

It is clearly impossible to divide the system by α or any of it's constituent parts (in par-

ticular - the velocity) without changing the value of the diagonal weight. If the velocity is variable, this will result in a different set of finite-difference weights at various lines of the linear system in (17), requiring a different set of spectrally factorized coefficients wherever the velocity changes.

The option of using a “filter bank” for different parts of the wavefield according to the local velocity has already been discussed in Rickett et al. (1998), for wave propagation in the frequency-wavenumber domain. This may also be applicable to propagation in the time-space domain, but I have not yet tested it. It was my hope that this would be unnecessary, and that a single set of filter coefficients could be utilized for the entire wavefield irrespective of velocity. This would make the propagation algorithm simpler, and more amenable to future parallelization schemes.

Stability analysis

The Von-Neumann stability analysis is useful in predicting the largest time steps possible for a particular order of a finite-difference scheme, for which the wavefield will not diverge. The application of this analysis to the 2nd order in time and space explicit finite-difference approximation of the 2-way wave equation follows. Assuming $h = \Delta x = \Delta z$, and using t as a time index, the approximation is:

$$U_{hj,hl}^{t+1} = 2U_{hj,hl}^t - U_{hj,hl}^{t-1} + \frac{C^2 \Delta t^2}{h^2} (U_{h(j+1),hl}^t + U_{h(j-1),hl}^t + U_{hj,h(l+1)}^t + U_{hj,h(l-1)}^t - 4U_{hj,hl}^t). \quad (19)$$

The field being propagated is some function of time combined with a harmonic function of space:

$$U^t = F^t e^{i(k_x h j + k_z h l)}. \quad (20)$$

inserting 20 into 19 and then dividing by $e^{i(k_x h j + k_z h l)}$ yields:

$$F^{t+1} = 2F^t - F^{t-1} + \alpha \left(F^t e^{ik_x h} + F^t e^{-ik_x h} + F^t e^{ik_z h} + F^t e^{-ik_z h} - 4 \right), \quad (21)$$

where $\alpha = \frac{c^2 \Delta t^2}{h^2}$.

In order to have stable propagation, the amplification factor - the amplitude ratio between the future wavefield and the current wavefield, must be smaller or equal to 1. This is also a requirement for the ratio between the past wavefield and the current wavefield, as the time reversed wavefield must also remain stable. From this consideration we have:

$$\frac{F^{t+1}}{F^t} = \frac{F^{t-1}}{F^t} = g(k_x, k_z). \quad (22)$$

Dividing (21) by F^t , and using the trigonometric identity for cosine we get:

$$g(k_x, k_z) = 1 + \alpha (\cos(k_x h) + \cos(k_z h) - 2) = 1 + \alpha R. \quad (23)$$

R is bounded by $-4 \leq R \leq 0$, and the requirement is that the amplification factor $|g| \leq 1$. It follows that

$$\begin{aligned} 1 - 4\alpha &\leq g \leq 1, \\ |1 - 4\alpha| &\leq 1, \\ \alpha &\leq \frac{1}{2}. \end{aligned} \tag{24}$$

Since $\alpha = \frac{C^2 \Delta t^2}{h^2}$, this analysis provides us with a way to determine the maximum time step for a given minimum velocity and spatial differencing step.

The same derivation for the 2D implicit finite differencing weights as derived in Equation 17 yields the amplification factor:

$$g(k_x, k_z) = \frac{1 + \alpha R}{1 - 2\alpha R}. \tag{25}$$

The boundaries for R remain $-4 \leq R \leq 0$. Because α is necessarily positive, it follows that $|g| \leq 1$ for any α in this 2D implicit finite-difference scheme. The time step can be arbitrarily large without causing the wavefield to diverge. This, of course, does not mean that we will get a *useful* wavefield with any arbitrary Δt .

CONCLUSION AND FUTURE WORK

It seems, at least in principle, that the general methodology described above for 2-way wave extrapolation with implicit finite difference by deconvolution using spectral factorization and the helix should be possible. The decrease in runtime in comparison to standard implicit finite difference application must be properly measured to assess the desirability of the method, although in my preliminary tests it is very evident that the method presented here is much faster than when using a standard solver.

The problem of dispersion of the wavefield when using the weights in Equation 10 must be overcome. It is not yet obvious to me what part of the observed dispersion can be attributed to the filter coefficients and what part to the finite-difference scheme. If the factorization is the main contributor, then one possible solution is to slightly increase the value of the diagonal element of matrix U in Equation 10, in order to improve the accuracy of the spectrally factorized filter coefficients. The connection between that and the number of factorized coefficients used in the filter is also as of yet unclear.

Another method of reducing the dispersion is to reformulate the implicit scheme with different weights. I have made one preliminary test in one dimension which suggests that doubling the value of the weights of matrix V in Equation 10 removes some of the dispersion. In any case, A more rigorous analysis of the phase velocity as a function of frequency of this or any other implicit scheme must follow.

Polynomial division is a necessarily sequential operation, as each output is dependent on the previous outputs (Equation 16). This seems to preclude any possibility of parallelization of the process within each time step. I do not see a way around this at the moment, but where there's a will there's a way.

ACKNOWLEDGMENTS

I wish to thank Ali Almomin for closing my linear algebra holes and explaining what polynomial division actually is, and also Biondo Biondi and Robert Clapp for suggesting this propagation method (and patiently explaining it...).

REFERENCES

- Claerbout, J., 2009, Basic earth imaging: Stanford University.
- Claerbout, J. F., 1997, Multidimensional recursive filters via a helix: Stanford Exploration Project, **Report 95**, 1–13.
- Clapp, R., 2009, Reverse time migration with random boundaries: SEG Expanded Abstracts, **28**, 2809.
- Fomel, S. and J. F. Claerbout, 1997, Exploring three-dimensional implicit wavefield extrapolation with the helix transform: Stanford Exploration Project, **Report 95**, 43–61.
- Haohuan, F., R. Clapp, O. Mencer, and O. Pell, 2009, Accelerating 3d convolutions using streaming architectures on fpgas: SEG Expanded Abstracts, **28**, 3035.
- Rickett, J., J. F. Claerbout, and S. Fomel, 1998, Implicit 3-d depth migration by wavefield extrapolation with the helical boundary conditions: SEG Expanded Abstracts, 1124–1127.

Predicting rugged water-bottom multiples through wavefield extrapolation with rejection and injection

Benxi Ke

ABSTRACT

Although convolution-based and WEM-modeling-based methods for predicting surface-related multiples are well-recognized in marine seismic data processing today, the effectiveness and efficiency of these methods are still a challenge in practice. In this paper, I present a WEM-modeling-based approach to multiple prediction. When wave-field rejection and injection are used during wave-field extrapolation, rugged water-bottom multiples can be accurately predicted when only the water-bottom elevation and water velocity are known.

INTRODUCTION

Seismic multiples can directly contaminate the stacking velocity and obscure the primary images in stack sections. Even worse, the energy of seismic multiples can smear into the target zones during seismic migration and result in a suboptimal result during seismic inversion. Therefore, seismic multiple attenuation has been a primary concern in marine seismic data processing.

In recent years, both the geophysical industry and academia have made significant progress in seismic multiple attenuation, including multiple prediction and adaptive multiple subtraction. As a result, a diverse toolkit for multiple attenuation is available in marine seismic data processing today. However, the effectiveness and efficiency of some multiple attenuation methods are still a challenge in practice.

In theory, 3D surface-related multiple elimination, or SRME, can predict surface-related multiples accurately from the seismic data itself without *a priori* knowledge of the subsurface (Verschuur and Berkhout, 1997; van Dedem and Verschuur, 2001). For a trace of a given shot-receiver pair, convolution-based SRME predicts the multiple model trace by convolving the common-shot gather related to the given shot point with the common-receiver gather related to the given receiver point. Ideally this method requires that the source and receiver are co-located, the source signatures are consistent, and the seismic data are completely acquired and well sampled. However, for most 2D and 3D marine streamer data, this is not the case, because of boat steering that deviates from the designed sail line, cable feathering and near offset gap, among other causes. Therefore, either data pre-processing for SRME or built-in processing during SRME is required, either of which reduces the computational efficiency.

As a complementary approach to 3D convolution-based multiple prediction, WEM-modeling-based multiple prediction employs downward and upward wavefield extrapolation between the sea surface and the water bottom to predict water-bottom multiples and

peg-legs based on a previously produced near-surface model (Wiggins, 1988; T. Weisser, 2006; Stork, 2006). This method theoretically is independent of seismic acquisition geometry and ideally requires a near-surface velocity model with water velocity, water-bottom topography and subsurface velocities. Unfortunately, in practice, it is not easy to estimate the subsurface velocities. Therefore, to allow wavefield extrapolation using only the water velocity and the water bottom topography, conventional wave equation modeling generally performs an approximation to the sea floor surface prior to wavefield continuation. In general, the Kirchhoff integral and finite-difference methods require dipping and horizontal flat sea floor surfaces respectively. Obviously, in the case of a structured or rugged sea floor, this approximation will result in suboptimal multiple modeling.

With the aim of improving the effectiveness and efficiency of wave-equation multiple modeling, I employ the so-called wavefield rejection and injection technique to perform wavefield extrapolation, so that I can use only the water velocity and the water-bottom topography to predict rugged water-bottom multiples and peg-legs. To best match this technique, I also present an algorithm for recursive Kirchhoff wavefield extrapolation in the space-frequency domain.

PREDICTING RUGGED WATER-BOTTOM MULTIPLES AND PEG-LEGS

Wave equation multiple modeling consists of downward wavefield extrapolation from the sea surface to the sea floor and upward wavefield extrapolation from the sea floor to the sea surface. In general, the water velocity and the water-bottom elevation can be measured with relative ease by field survey or estimated from seismic data. It is relatively difficult or expensive to get an accurate interval velocity of the subsurface by seismic velocity analysis when the lateral velocity varies strongly. Therefore, if I can use only the water velocity and water-bottom elevation to predict the water-bottom multiples and peg-legs, the efficiency will be improved.

One of the solutions to the above-mentioned issue is to make an approximation to the rugged water bottom. The approximation depends on which algorithm I choose. For example, if the wavefield extrapolation is based on the Kirchhoff integral algorithm, a dipping flat surface can be selected to approximate the rugged water bottom, so that the calculation of the Greens function does not involve the subsurface velocity. Similarly, if the wavefield extrapolation is based on the phase shift or finite difference algorithm, a horizontal flat surface can be chosen. Obviously, these approximations to the real water bottom will sacrifice effectiveness for efficiency.

To improve the efficiency without loss of effectiveness, I present a so-called wavefield extrapolation with rejection and injection.

Wavefield extrapolation with rejection and injection

The idea of this method comes from the fact that water-bottom multiples and peg-legs are trapped between the water bottom and sea surface, as shown in figure 1. Therefore, when we input the shot gather on the sea surface and perform downward wavefield extrapolation, the down-going wave below the water-bottom can be rejected. Similarly, when we inject

the recorded down-going wavefield at the water bottom and perform the upward wavefield extrapolation, the up-going wave below the water bottom also can be rejected.

Obviously, based on the above-mentioned idea, an auxiliary model presentation should be given. As shown in figure 2, in this auxiliary model presentation, the solid circle, rectangle and star represent the nodes of computation in water, in the subsurface, and on the water bottom, respectively.

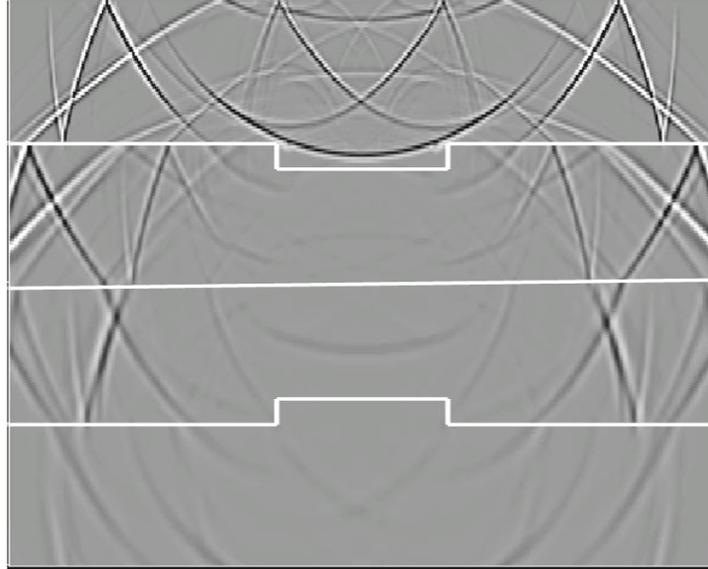


Figure 1: A snapshot of seismic modeling to show the water bottom multiples and illustrate wavefield rejection and injection. [NR] benxi1/. fig1

Recursive Kirchhoff wavefield extrapolation

Theoretically, wavefield rejection and injection can be carried out for any of the wavefield extrapolation algorithms. However, the computational efficiency can differ greatly. Recursive Kirchhoff wavefield extrapolation in both space-frequency and space-time domain is the most suitable algorithm.

Because the basic theory and extrapolation operator based on the Kirchhoff integral are widely known (Schneider, 1978; Berkhout, 1981; Margrave and Daley, 2001), I derive only the formulae corresponding to upward and downward wavefield extrapolation of both the up-going and down-going waves.

In general, the wavefield extrapolation from depth \mathbf{z} to $\mathbf{z} + \Delta\mathbf{z}$ can be written as follows:

$$P^+(k_x, k_y, z \pm \Delta z, \omega) = P^+(k_x, k_y, z, \omega) e^{-ik_z(\pm\Delta z)}, \quad (1)$$

$$P^-(k_x, k_y, z \pm \Delta z, \omega) = P^-(k_x, k_y, z, \omega) e^{+ik_z(\pm\Delta z)}, \quad (2)$$

where $P^+(k_x, k_y, z, \omega)$ and $P^-(k_x, k_y, z, \omega)$ are the Fourier transform over x , y and t of the down-gong and up-going waves at position (x, y, z) , respectively. The terms k_x , k_y and

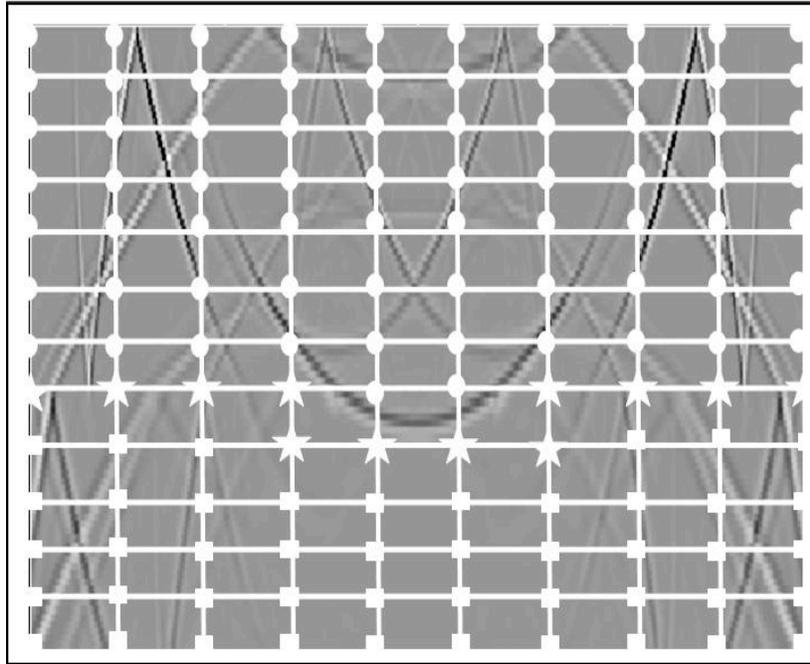


Figure 2: The auxiliary model presentation for wavefield extrapolation with wavefield rejection and injection. [NR] [benxi1/. fig2](#)

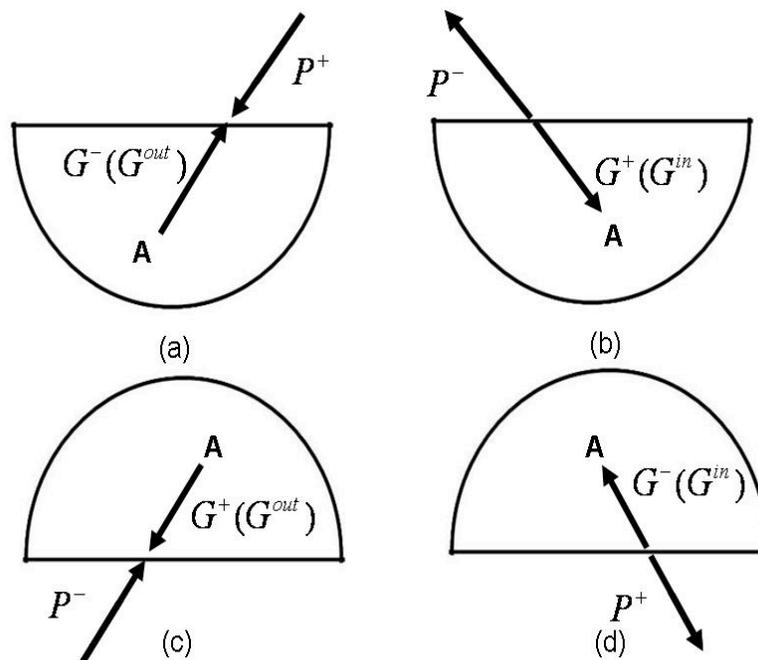


Figure 3: Schematic illustration of the four types of wavefield extrapolations in terms of Kirchhoff integral: (a) downward wavefield extrapolation of the down-going wave; (b) downward wavefield extrapolation of the up-going wave; (c) upward wavefield extrapolation of the up-going wave; and (d) upward wavefield extrapolation of the down-going wave. [NR] [benxi1/. fig3](#)

k_z are the three components of the wavenumber vector, and ω is the angular frequency. The sign \pm before the depth interval Δz relates to the upward and downward wavefield extrapolations. Thus, there are four types of wavefield extrapolations in total.

On the other hand, the Kirchhoff integral in the space-frequency domain is

$$\tilde{P}(r_A, \omega) = \oint_s \left[\tilde{G}(r, r_A, \omega) \frac{\partial \tilde{P}(r, \omega)}{\partial n} - \tilde{P}(r, \omega) \frac{\partial \tilde{G}(r, r_A, \omega)}{\partial n} \right] ds, \quad (3)$$

where r and r_A are the shorthand notations of (x, y, z) and (x_A, y_A, z_A) . The term n is the outward normal of the surface S . The wavefield $\tilde{P}(r, \omega)$ and Green function $\tilde{G}(r, r_A, \omega)$ satisfy the following Helmholtz equations:

$$\nabla^2 \tilde{P}(r, \omega) + \frac{\omega^2}{c^2} \tilde{P}(r, \omega) = 0, \quad (4)$$

$$\nabla^2 \tilde{G}(r, r_A, \omega) + \frac{\omega^2}{c^2} \tilde{G}(r, r_A, \omega) = \delta(r - r_A), \quad (5)$$

where c is wave propagation velocity. Obviously, it is not easy to relate equation 3 to the upward and downward wavefield extrapolation of both the up-going and down-going waves. Suppose S consists of a horizontal S_0 surface at $z = z_n$ and a hemisphere S_1 which contains point A and satisfies the Sommerfeld radiation condition. Transforming the Kirchhoff integral from the (x, y, z, ω) domain into (k_x, k_y, z, ω) domain, equation 3 becomes (Berkhout, 1989)

$$P(r_A, \omega) = \frac{1}{(2\pi)^2} \iint \left[G \frac{\partial P}{\partial z} - P \frac{\partial G}{\partial z} \right] dk_x dk_y, \quad (6)$$

where G and P represent $G(-k_x, -k_y, z, r_A, \omega)$ and $P(k_x, k_y, z, \omega)$. Furthermore, equation 6 can be mathematically expressed by

$$P(r_A, \omega) = \frac{1}{(2\pi)^2} \iint \left[(G^+ + G^-) \frac{\partial(P^+ + P^-)}{\partial z} - (P^+ + P^-) \frac{\partial(G^+ + G^-)}{\partial z} \right] dk_x dk_y, \quad (7)$$

where P^+ and G^+ are the down-going wave and Greens function respectively, and P^- and G^- are the up-going wave and Greens function respectively.

Using the definitions of the down-going and up-going waves,

$$\frac{\partial P^\pm}{\partial z} = \mp k_z P^\pm, \quad (8)$$

$$\frac{\partial G^\pm}{\partial z} = \mp k_z G^\pm, \quad (9)$$

Equation 7 can be simplified to

$$P(r_A, \omega) = \frac{1}{(2\pi)^2} \iint \left[\left(G^- \frac{\partial P^+}{\partial z} - P^+ \frac{\partial G^-}{\partial z} \right) + \left(G^+ \frac{\partial P^-}{\partial z} - P^- \frac{\partial G^+}{\partial z} \right) \right] dk_x dk_y, \quad (10)$$

and furthermore to

$$P(r_A, \omega) = \frac{1}{2\pi^2} \iint \left[P^+ \frac{\partial G^-}{\partial z} + P^- \frac{\partial G^+}{\partial z} \right] dk_x dk_y, \quad (11)$$

Thus the four types of wavefield extrapolations in terms of the Kirchhoff integral can be easily defined based on equation 11, as shown in figure 3.

In practice, generally we transform the equation 11 from the (k_x, k_y, z, ω) domain back into the (x, y, z, ω) or (x, y, z, t) domain. In this case, the down-going and up-going Greens functions will change accordingly, as shown in figure 3. The terms G^{out} and G^{in} are the causal and anticausal Greens functions, or the out-going and in-going waves respectively.

After the above derivation, we have all types of wavefield extrapolations in terms of the Kirchhoff integral. In water-bottom multiple prediction, we use only types (a) and (c), that is, the downward wavefield extrapolation of the down-going wave and the upward wavefield extrapolation of the up-going wave.

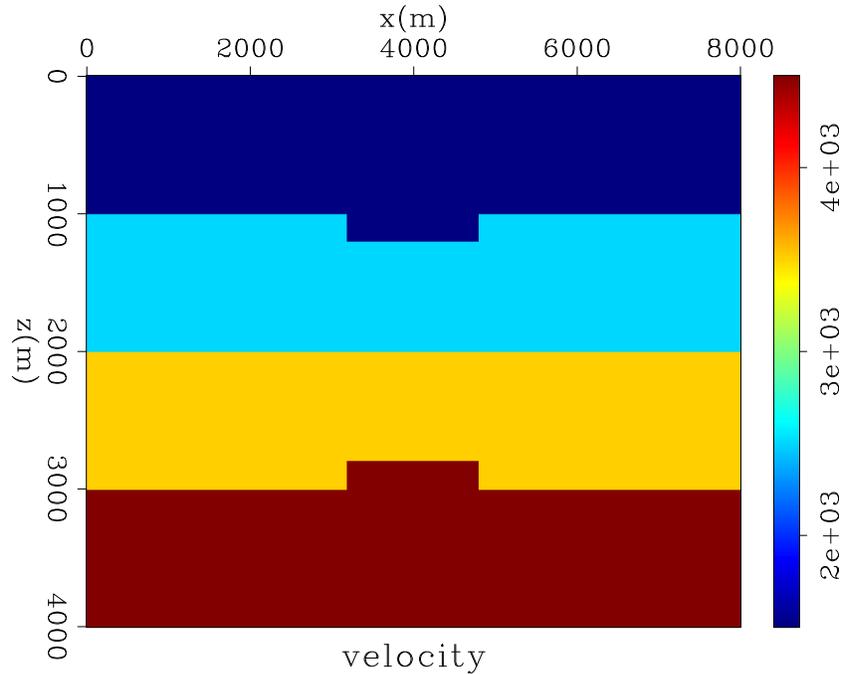


Figure 4: The structure and interval velocity of the geological model. [NR] benxi1/. vel

SYNTHETIC EXAMPLE

To evaluate the usefulness of multiple prediction and adaptive subtraction, we use the geologic model shown in figure 4 to produce the synthetic seismic data. In this model, there are four layers, including the water layer. The interval velocities from the top to bottom layer are 1500, 2500, 3500 and 4500 m/s respectively, and the depths of the first to third interface are 1000, 2000 and 3000 m respectively. As there is a notch (trench) on the first interface, the synthetic seismic data contain not only specularly reflected but also diffracted

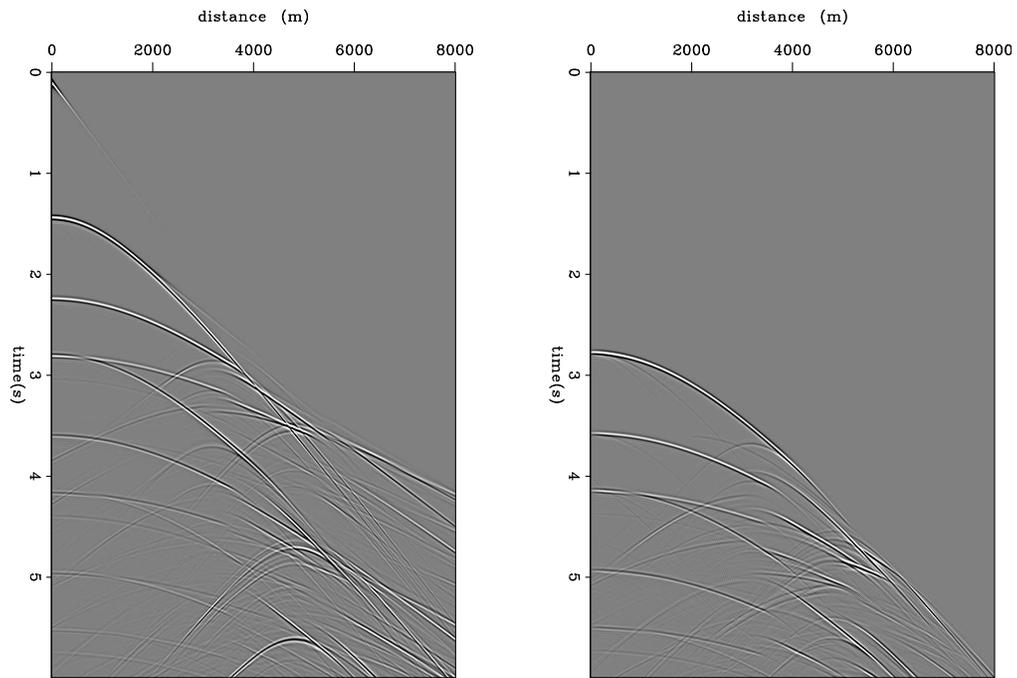


Figure 5: The shot gather at the leftmost location(left) and predicted multiples (right).
 [NR] `benxi1/. shotmult1`

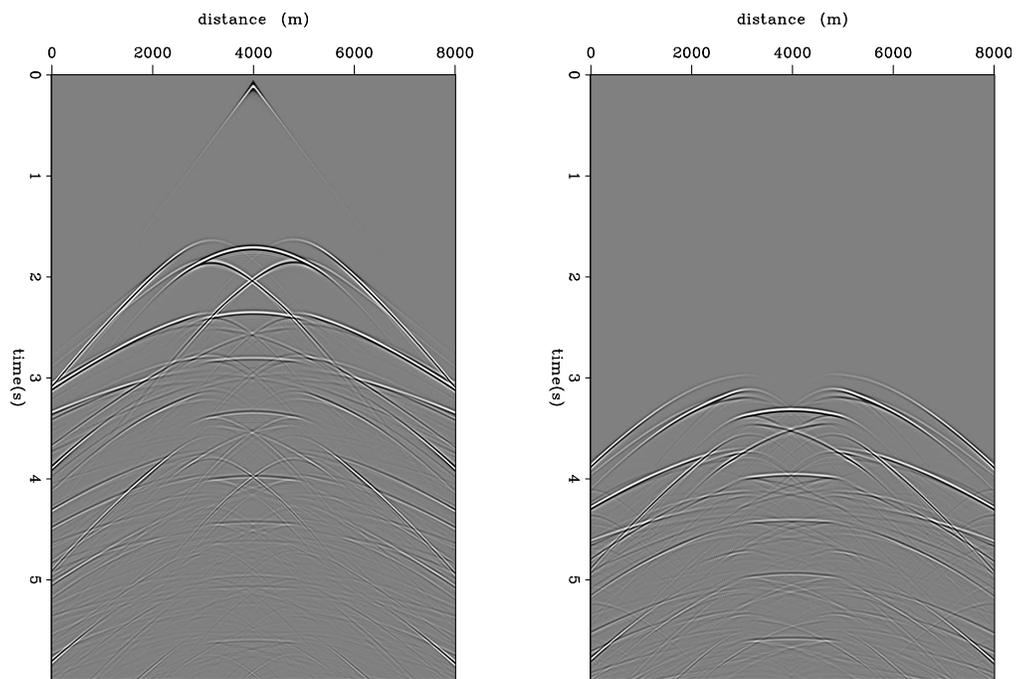


Figure 6: The shot gather at the central location (left) and predicted multiples (right).
 [NR] `benxi1/. shotmult2`

water-bottom multiples and peg-legs. Furthermore, the model has a convex feature (fault or salt) on the third interface to produce specularly reflected and diffracted interbed multiples. In addition, to test the ability to preserve the primary, we design the interface depths such that the zero-offset travel time of the first order water-bottom multiples are very close to the zero-offset reflection time of the third interface. Figure 5 and figure 6 show shot gathers and predicted multiples at the leftmost and central locations. Figure 7 shows the zero-offset sections of the original seismic data and predicted multiples. Comparing the original and predicted multiples both in the shot gathers and the zero-offset section, we can see that both the specularly reflected and diffracted water bottom multiples and peg-legs are well predicted. The results also demonstrate the limitation of both the convolution-based and WEM-modeling-based surface-related multiple prediction methods; neither can predict the interbed multiples.

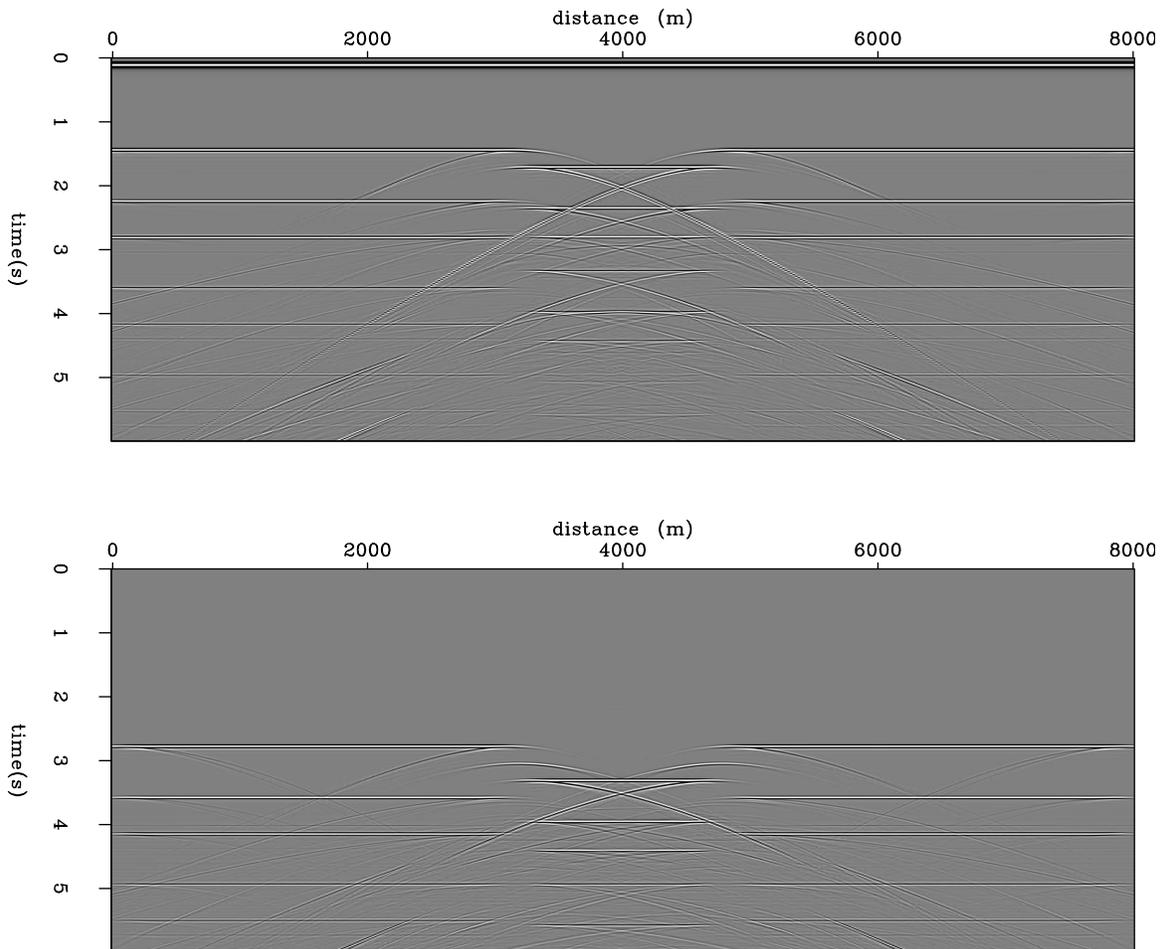


Figure 7: The zero-offset sections of the original seismic data (top) and the predicted multiples (bottom). [NR] `benxi1/. smcof1`

CONCLUSIONS

In the case of a structured or rugged sea floor, wavefield extrapolation with rejection and injection is an efficient and effective technique for the prediction of the rugged water-bottom multiple and peg-legs. The Kirchhoff integral is the algorithm best suited to both the downward wavefield extrapolation of the down-going wave and the upward wavefield extrapolation of the up-going wave. The synthetic data example shows that both the specularly reflected and diffracted water-bottom multiples and peg-legs are well predicted by our method.

ACKNOWLEDGMENTS

Thanks to Xukai Shen for performing seismic modeling and providing the synthetic seismic data.

REFERENCES

- Berkhout, A. J., 1981, Wave field extrapolation techniques in seismic migration, a tutorial: *Geophysics*, **46**, 1638–1656.
- , 1989, One-way versions of kirchhoff integral: *Geophysics*, **54**, 460–467.
- Margrave, G. F. and P. F. Daley, 2001, Recursive kirchhoff extrapolators: CREWES Research Report, **13**, 617–646.
- Schneider, W. A., 1978, Integral formulation for migration in two dimensions or three dimensions: *Geophysics*, **43**, 49–76.
- Stork, C., 2006, Predicting and removing complex 3d surface multiples with wem modeling—an alternative to 3d srme for wide azimuth surveys: SEG Expanded Abstract.
- T. Weisser, A. Pica, a. P. H., 2006, Wave equation based multiple modelling comparison of nominal and dense acquisition geometries: 68th EAGE Ann. Conference.
- van Dedem, D. J. and D. J. Verschuur, 2001, 3d surface multiple prediction using sparse inversion: 71st SEG International Exposition and Technical Meeting, Expanded Abstracts, 1285–1288.
- Verschuur, D. J. and A. J. Berkhout, 1997, Estimation of multiple scattering by iterative inversion, part ii: Practical aspects and examples: *Geophysics*, **62**, 1596–1611.
- Wiggins, J., 1988, Attenuation of complex water-bottom multiples by wave-equation based prediction and subtraction: *Geophysics*, **53**, 1527–1539.

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.



Claudio Guerra received his B.Sc. in Geology from Federal University of Rio de Janeiro, Brazil in 1988 and a M.Sc. from State University of Campinas, Brazil in 1999. Since 1989, he has been working for Petrobras, Brazil. He joined SEP in 2006 and is currently pursuing a Ph.D. in geophysics at Stanford University. He is member of SEG and SBGf.



Adam Halpert graduated from Rice University in May 2007 with a B.Sc. in Earth Science and a B.A. in Policy Studies. He joined SEP in the summer of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG and AGU.



Least-squares imaging and deconvolution using the hybrid norm conjugate-direction solver

Yang Zhang and Jon Claerbout

ABSTRACT

To retrieve a sparse model, we applied the hybrid norm conjugate-direction (HBCD) solver proposed by Claerbout to two interesting geophysical problems: least-squares imaging and blind deconvolution. The results showed that this solver is robust for generating sparse models.

INTRODUCTION

L1 norm optimization in many situations are more desirable than the conventional least squares (L2) optimization. However currently widely used methods like IRLS (Iterative Reweighted Least Squares) or weighted-median (Claerbout and Muir, 1973; Darche, 1989; Guitton, 2005) require the users to fine tune extra solver parameters in order to obtain a pleasing result. The sensitivity of such parameters make the solvers cumbersome to use since the users have to do trial-and-error. We developed a robust and efficient L1-type solver (Claerbout, 2009b) that uses a hybrid norm function to approximate the L1 norm, and implemented a generalized conjugate-direction (CD) method by using Taylor's expansion (Maysami and Mussa, 2009).

This solver is convenient to apply, because the function interface is almost the same as the traditional least-squares (L2) solver in the SEP library. The user must specify one additional parameter: the residual quantile. Fortunately this parameter has a clear physical meaning (Claerbout, 2009b). Users should assign this parameter according to prior observation or expectation of the model's spikiness/blockness.

In this paper we show the usefulness of the hybrid solver by applying it on the LSI imaging and deconvolution problems. The L1 inversion of LSI imaging (Least Squares Inverse) problem is preferable to L2 inversion, because it better preserves the spikiness/sparseness that are commonly encountered in reflectivity models. When the model regularization is posed with the L2 norm, it is hard to honor spikiness/sparseness, because the L2 norm cannot tolerate large values in the model. In contrast, the L1 type norm fits our regularization goal very well.

A similar motivation applies to the deconvolution problem; conventional deconvolution assumes the reflectivity series to be random (white spectrum), whereas we argue that a sparse reflectivity series is more appropriate (and often more desirable) in practice.

APPLICATION – LEAST-SQUARES INVERSE IMAGING

This application originated from the work of target-oriented wave-equation LSI imaging, as described in M. Clapp and Biondi (2005); Valenciano (2006); Tang (2008). The concept of LSI imaging starts with a simple inversion problem:

$$\mathbf{F}(\mathbf{m}) = \|\mathbf{L}\mathbf{m} - \mathbf{d}_{\text{obs}}\|_2, \quad (1)$$

where \mathbf{L} is a linearized wave-equation modeling operator, the adjoint of which is the imaging operator, \mathbf{m} is the subsurface *reflectivity* model, and \mathbf{d}_{obs} is the observed surface seismic data. In theory, the solution to this inversion problem can be written as follows:

$$\mathbf{m} = \mathbf{H}^+ \mathbf{L}' \mathbf{d}_{\text{obs}}, \quad (2)$$

Where $\mathbf{H} = \mathbf{L}'\mathbf{L}$ is called Hessian operator, and \mathbf{H}^+ is the pseudo inverse of \mathbf{H} . In practice, it is usually impossible to invert \mathbf{H} directly; thus a gradient-based optimization method is often used to find the solution.

One disadvantage of this data-space inversion scheme is that it can not be computed in a target-oriented way, since theoretically even a local perturbation in the model space will affect the entire data space and vice versa. To overcome this difficulty, Valenciano (2006) transformed (1) to a model space inversion based on (2):

$$\mathbf{J}(\mathbf{m}) = \|\mathbf{H}\mathbf{m} - \mathbf{L}'\mathbf{d}_{\text{obs}}\|_2.$$

Valenciano (2008) and Tang (2008) showed that unlike \mathbf{L} , matrix \mathbf{H} is usually very sparse (i.e., most of the non-zero elements are centered around the diagonal); thus despite the huge size of \mathbf{H} , it is feasible to store an approximation of \mathbf{H} matrix by keeping only a few off-diagonal elements without losing much accuracy.

If we write $\mathbf{m}_{\text{mig}} = \mathbf{L}'\mathbf{d}_{\text{obs}}$, and add a model regularization term (since most likely \mathbf{H} has a null space). Then the inversion formula is

$$\mathbf{J}(\mathbf{m}) = \|\mathbf{H}\mathbf{m} - \mathbf{m}_{\text{mig}}\|_2 + \epsilon \|\mathbf{m}\|_{\text{norm}},$$

in which we applied the hybrid norm to the regularization term.

Tang (2009) provided a way to efficiently compute the Hessian matrix using the phase-encoding technique, and this Hessian matrix is computed only once and stored for all iterations.

Numerical example

The reflectivity model we started with is a cropped subsalt region from the sigsbee2A reflectivity model, as shown in Figure 1(a). Notice that it is quite sparse.

Figure 1(b) shows the input migrated image \mathbf{m}_{mig} . While the data \mathbf{d} is modeled with a two-way wave equation, both the migrated image $\mathbf{m}_{\text{mig}} = \mathbf{L}'\mathbf{d}$ and the Hessian operator are computed using a one-way wave-equation propagator. Therefore the data contains non-linear information (e.g., multiples) that cannot be resolved by the linearized one-way

wave-equation operator. This explains some of the migration artifacts in the migrated image.

The explicit Hessian operator is computed using the receiver-side random-phase encoding method (Tang, 2008). The size of off-diagonal elements at each image point is 21×21 . After the Hessian is computed, we extract the portion corresponds to the above reflectivity model.

Figure 2(a) and 2(b) shows the inversion result of L2 and hybrid . The first thing to notice is that neither method can perfectly retrieve the original model; nonetheless, there is a significant improvement in the L1 inversion. The sharp boundary of the reflectors at the left of the image is better recovered in the hybrid result.

Figure 2(c) and 2(d) shows the data fitting errors of the two inversion are plotted. By evaluating the total energy of the fitting error (3.48% for both inversion results, $p_e = \|\mathbf{H}\mathbf{m} - \mathbf{m}_{\text{mig}}\|_2 / \|\mathbf{m}_{\text{mig}}\|_2$), we claim that hybrid inversion and L2 inversion fit the data almost equally well. This ensures that the major effect of the regularization is to fill the null space of \mathbf{H} , with little effect on the data-fitting. However it is true that the data fitting residual of hybrid inversion appears to be more correlated to the reflectivity model than that of L2 inversion is.

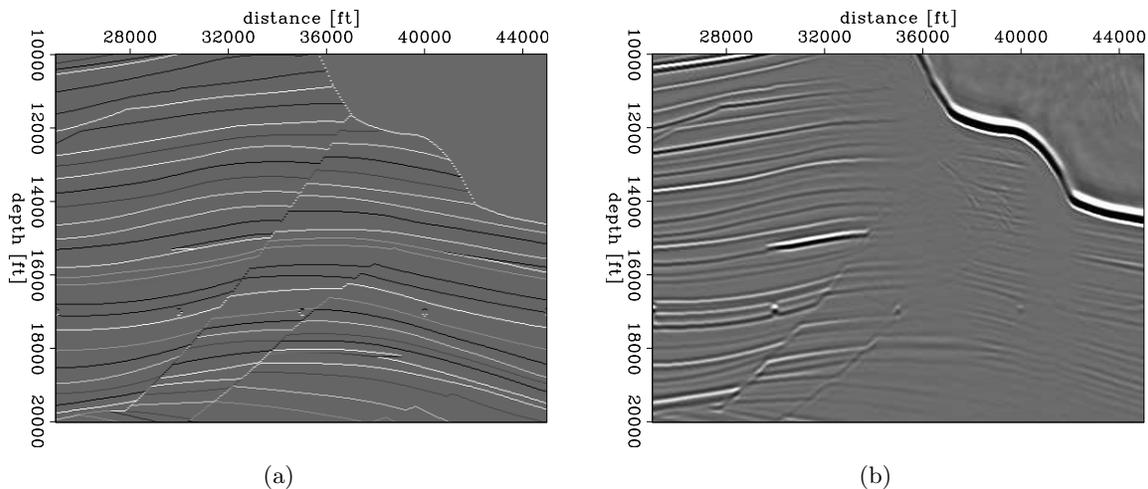


Figure 1: (a) Reflectivity model; (b) input migrated image. [ER]
 yang1/. refl-imaging,data-imaging-mig

In addition, notice that some parts of the sub-salt reflectors presented in the L2 result are missing in the hybrid result. The reason is clear: the hybrid norm is less tolerant of small values in the residual and always tries to put them down to zero (Claerbout, 2009b). This example shows that this feature of the hybrid norm is not always desirable.

APPLICATION – DECONVOLUTION

Deconvolution has been a well-known geophysical problem since the 1950s. We investigate the spiking deconvolution, which aims to compress the source wavelet, such that a reflectiv-

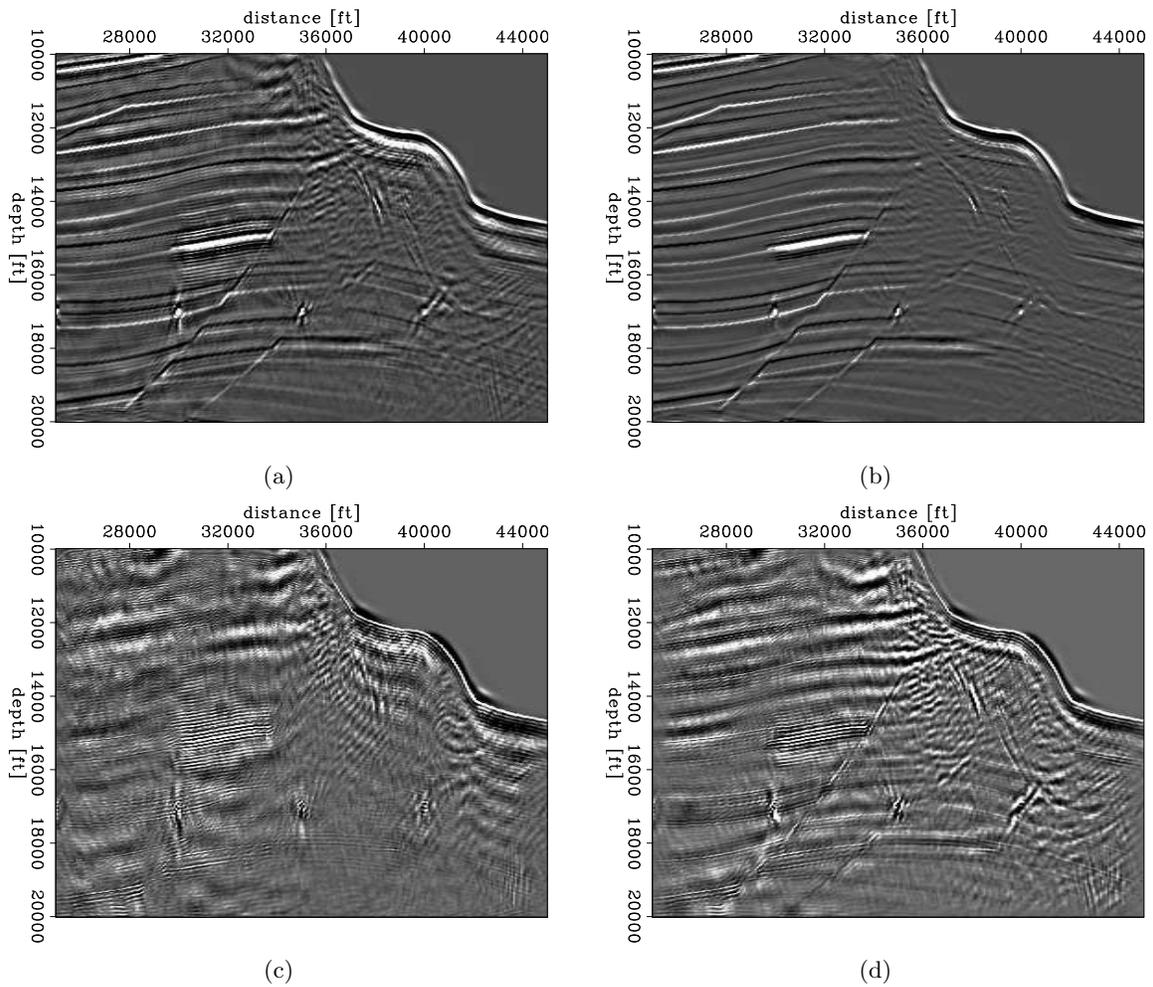


Figure 2: Comparison of the inversion results with L2 and hybrid regularization. (a): L2 inversion result; (b): hybrid inversion result; (c): L2 data fitting residual (d): hybrid data fitting residual. [ER]

yang1/. sigsb2a.mig.inv.l2,sigsb2a.mig.inv.hb,sigsb2a.mig.resd.l2,sigsb2a.mig.resd.hb

ity series with higher resolution can be obtained. The simple convolution model is expressed as follows:

$$m(t) * s(t) + n(t) = d(t) \quad (3)$$

where $r(t)$ is the reflectivity series, $s(t)$ is the source wavelet, $n(t)$ is random noise, and $d(t)$ is the seismic traces (we assume a certain kind of amplitude compensation has already been applied).

Intrinsically, this is an under-determined problem, because both $r(t)$ and $s(t)$ are unknown. Further assumptions about the reflectivity series are needed in order to get a deterministic answer. In the L2 scenario, the underlying assumption is that the reflectivity model is purely random (i.e., has a *white spectrum*). As mentioned before, the model may in fact be spiky, which is better matched by an L1 type inversion. Therefore the hybrid result should outperform the L2 result.

For simplicity, we also assume that source wavelet is minimum phased. The conventional spiking deconvolution can be defined as an inverse problem,

$$\mathbf{D}\mathbf{a} \approx \mathbf{0},$$

where \mathbf{D} is the data convolution operator, and \mathbf{a} is the filter. In this formulation, the filter is the only unknown, and in theory the data residual itself is the reflectivity model.

To incorporate the model regularization into the inversion framework, we generalize the formulation above by posing the deconvolution problem as such inversion problem:

$$\left\{ \begin{bmatrix} \mathbf{D} & -\mathbf{I} \\ \mathbf{0} & \epsilon\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{m} \end{bmatrix} \approx \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \right\}, \quad (4)$$

in which \mathbf{D} is the data convolution operator, \mathbf{a} is the filter, and \mathbf{m} is the reflectivity model. The parameter ϵ indicates the strength of the regularization. Since the source wavelet is assumed to be a minimum-phase wavelet, ideally the inversion gives the exact inverse of the source wavelet to \mathbf{a} .

The first equation in (4) (data fitting) implies that after convolving the data with the filter, we should get the reflectivity model; any values that cannot fit the reflectivity model are considered as noise in data. The second equation in (4) is the spiky regularization of the model; thus we apply the hybrid norm.

Deconvolution of synthetic data

We started from a simple synthetic reflectivity model from Basic Earth Imaging by Claerbout (2009a). Figure 3 shows the starting reflectivity model, which is quite sparse. Figure 4(a) shows the minimum-phase wavelet we designed. Figure 4(b) shows the data generated by convolving the reflectivity model with the wavelet. Convolution is done trace by trace.

Figure 5 compares the result of L2 inversion and L1 inversion (using the hybrid norm). As expected, the conventional L2 result gives a fuzzy model poorly correlated to the original one. In contrast, the hybrid result recovers the original model quite well. For verification, the data residuals are plotted at the bottom of Figure 5; the hybrid result has a fitting error

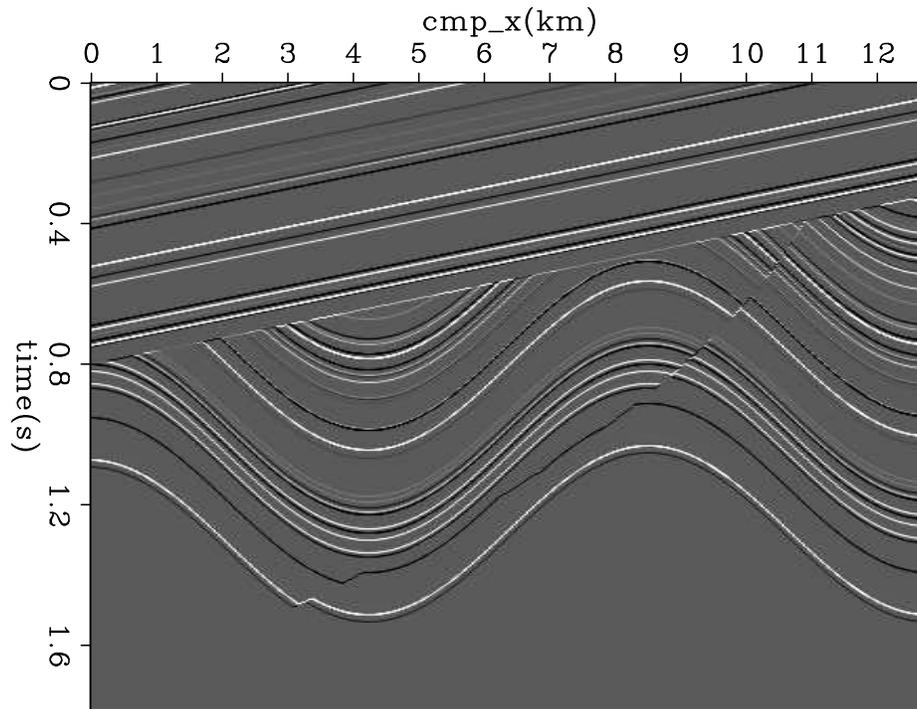


Figure 3: Reflectivity model. [ER] yang1/. refl-syn-decon

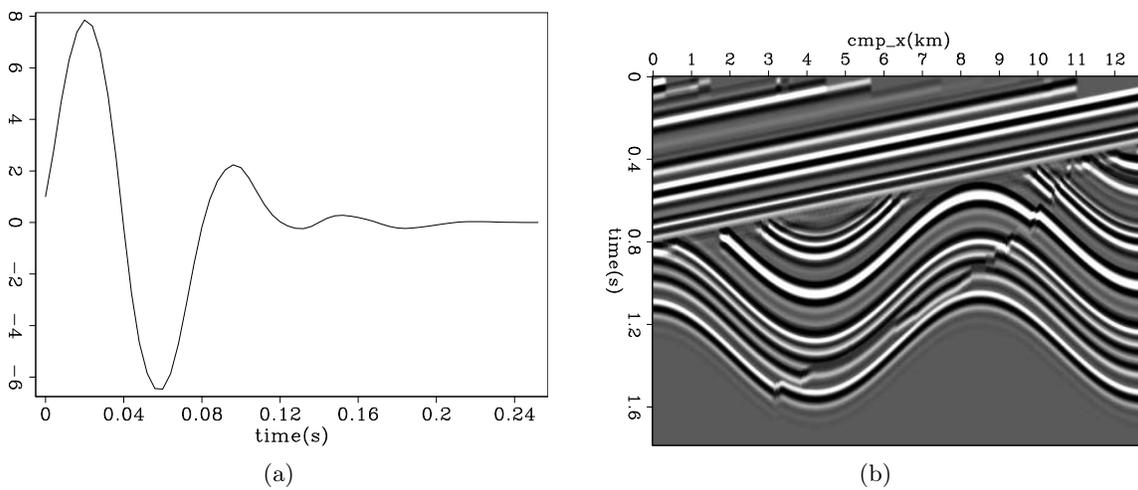


Figure 4: Synthetic Data for deconvolution. (a): The minimum phase wavelet; (b): generated input data. [ER] yang1/. wavelet-syn-decon,data-syn-decon

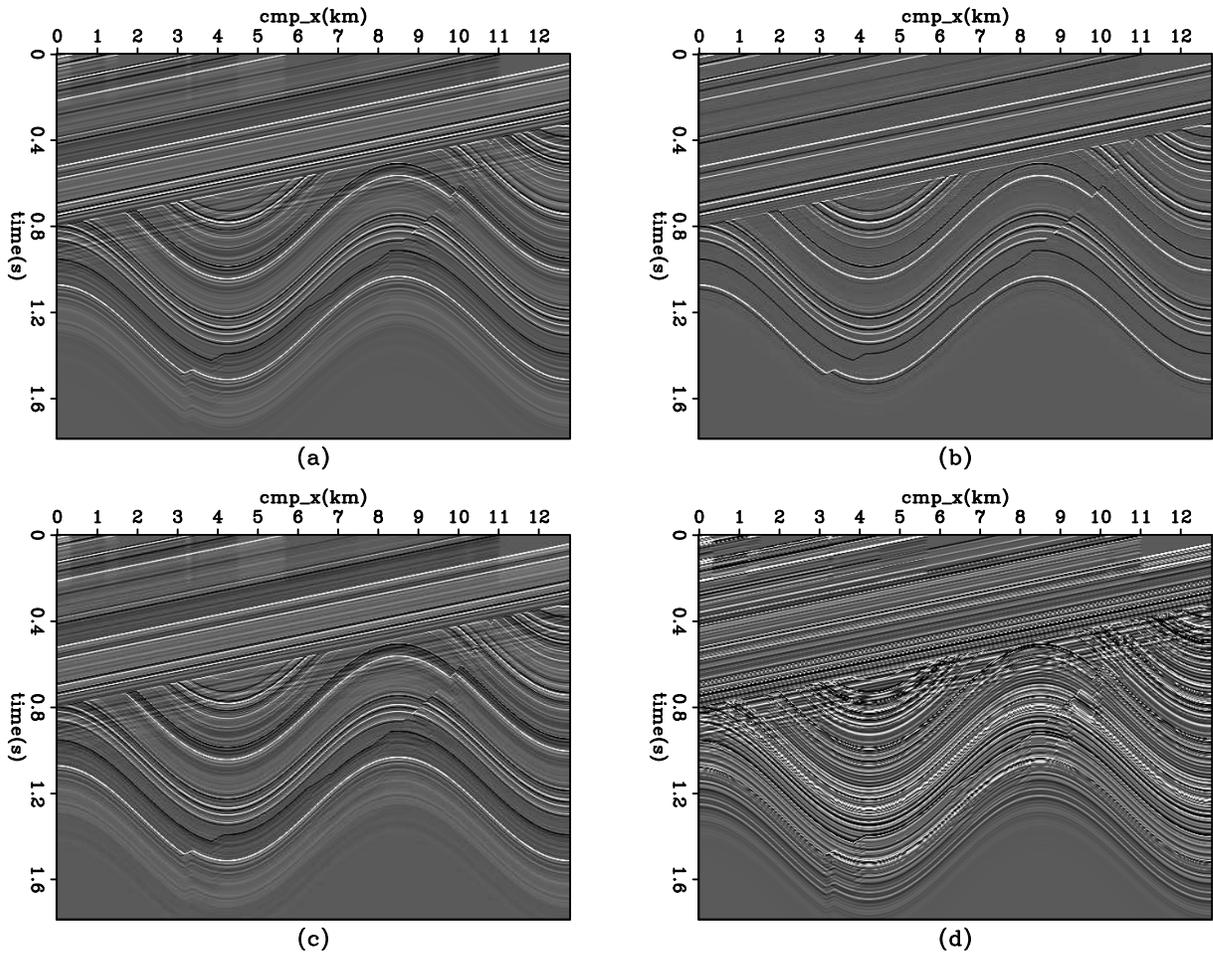


Figure 5: Comparison of deconvolution result for synthetic data. Top panels show the reflectivity model obtained by L2 inversion (a) and hybrid inversion (b); bottom panels (c)(d) show the respective data residual. [ER] yang1/. syn-decon

of 0.6%, indicating that the data fitting goal is well honored.

One limitation of the formulation (4) is the assumption of a minimum phase wavelet. A non-minimum phase signal does not have a causal inverse in theory, thus breaking our formulation (4) since the filter \mathbf{a} that is supposed to be the inverse of source wavelet does not exist.

To see how much the non-minimum phase wavelet will affect the inversion result, we performed another experiment in which the wavelet is chosen to be a non-minimum phase one. Figure 6 shows the new wavelet and the synthetic data generated using this wavelet. Figure 7 shows the result of L2 inversion and L1 inversion (using the hybrid norm) of this data, using the same parameters as in the previous synthetic example. In contrast to the significant improvement obtained before, the hybrid result cannot yield a sparse model because it is impossible to find a filter \mathbf{a} that undoes the source wavelet.

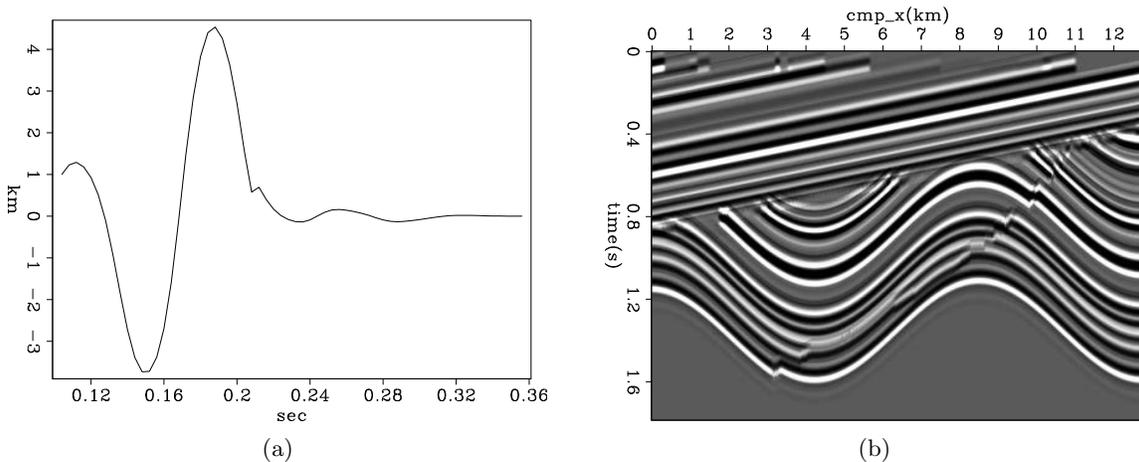


Figure 6: (a): the mixed-phase wavelet; (b): generated data. [ER] yang1/. wavelet-syn-mix-decon,data-syn-mix-decon

Deconvolution of a common-offset field data

The second example is a common-offset section of field marine data. Figure 8 shows the input data.

Figure 9 shows the deconvolution result using L2 and L1 inversion. Although in this case the L1 inversion gives a cleaner model, the model is less desirable. Some areas of interest (like the salt-bottom) are suppressed by regularization due to lower amplitude than do the salt-top and sea-bed. In other words, the regularization is too strong. The bottom panel of Figure 9 showing the data fitting residual further confirms this point. From the amplitude information of this plot, roughly 20% of the data was pushed into fitting error, the regularization distorts the data fitting too much rather than eliminating the model's null space.

Our predicament in this case is that if we set ϵ to a small value, then the hybrid result

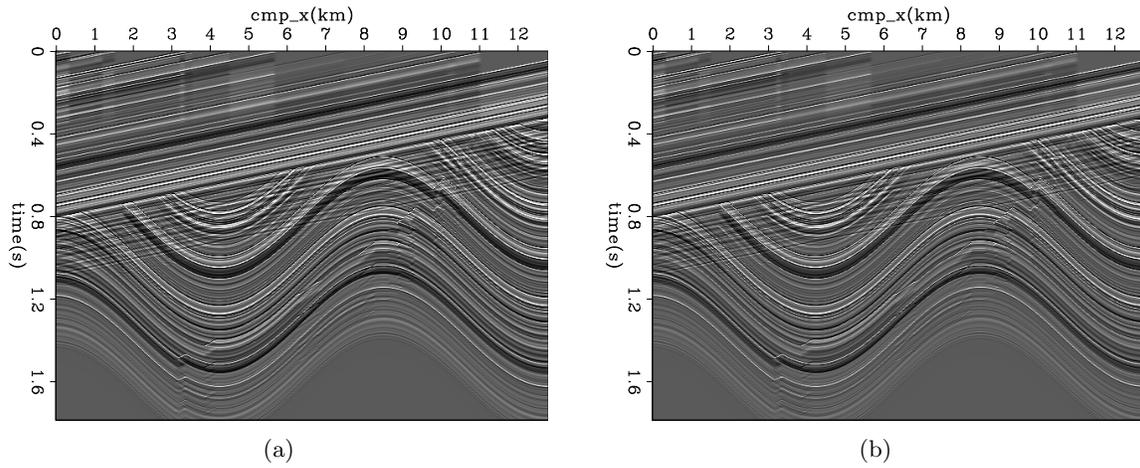


Figure 7: (a): L2 inversion result for the synthetic data generated with mixed-phase wavelet;(b): hybrid inversion result. [ER] yang1/. result-syn-mix-decon-l2,result-syn-mix-decon-hb

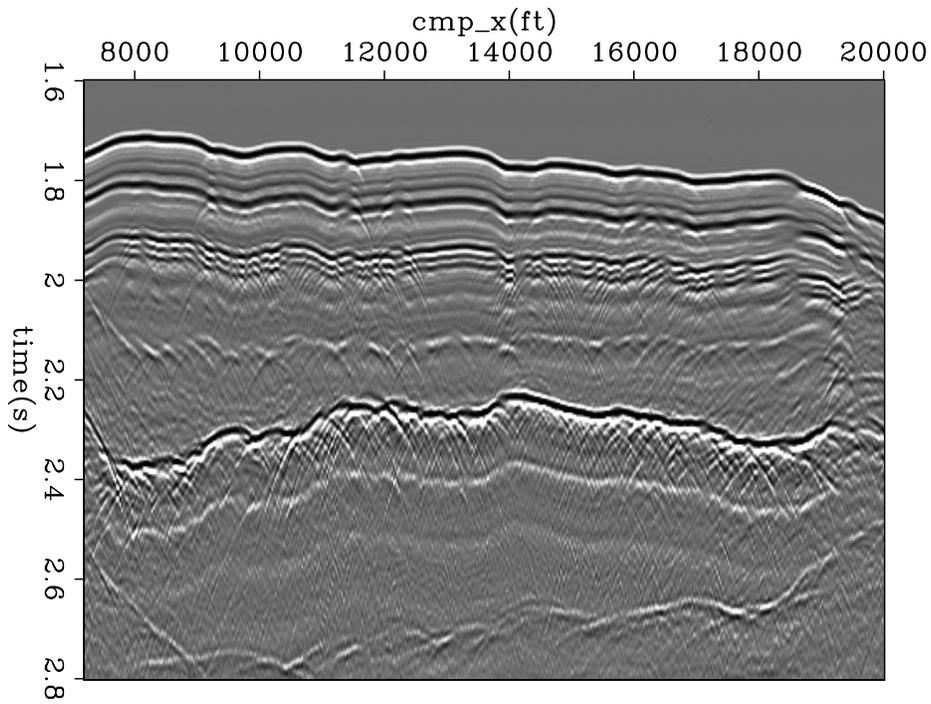


Figure 8: Input Common Offset data. [ER] yang1/. data-COF-decon

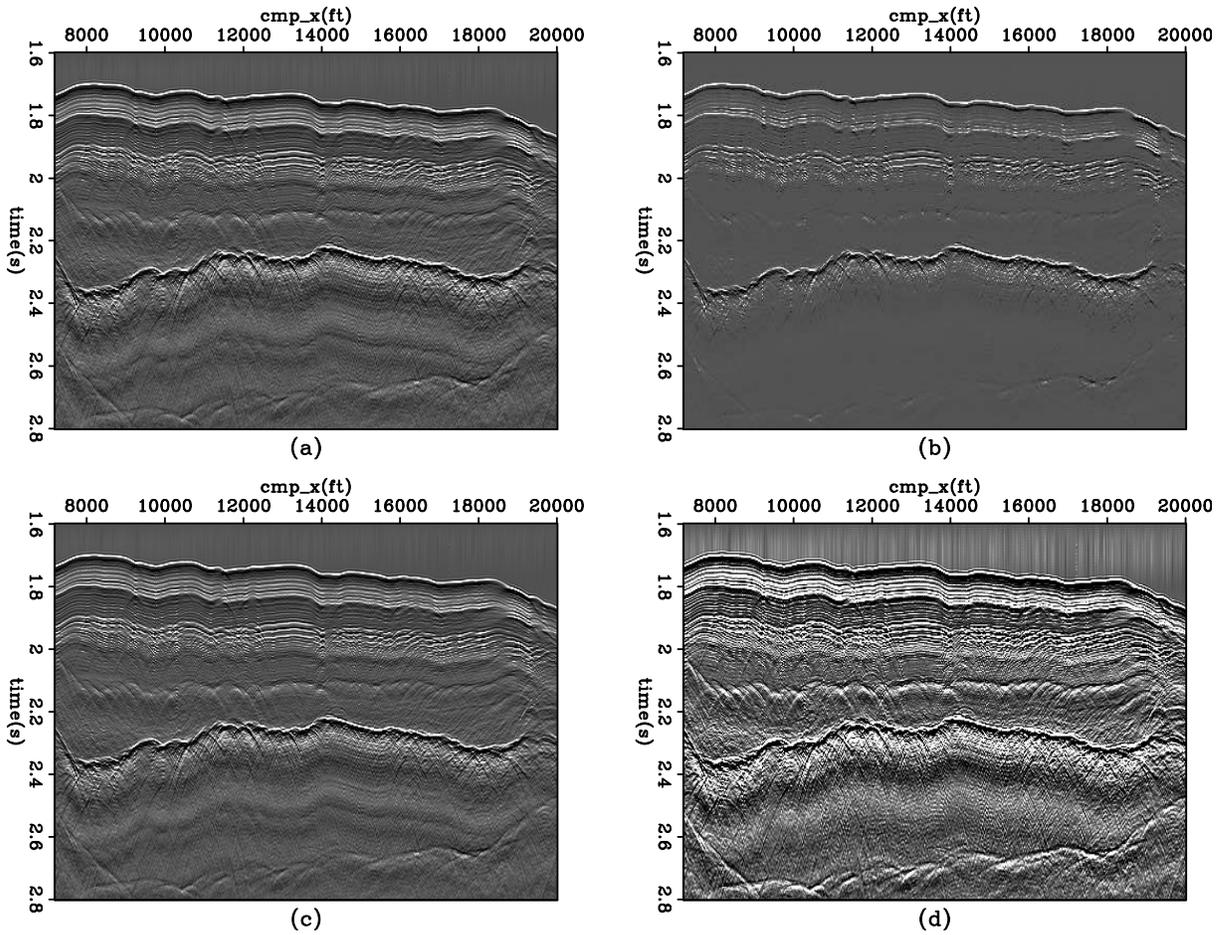


Figure 9: Top: L2 (a) and hybrid (b) deconvolution result for the Common Offset data; bottom: data fitting residuals of L2 (c) and hybrid (d) inversion. [ER] yang1/. result-COF-decon

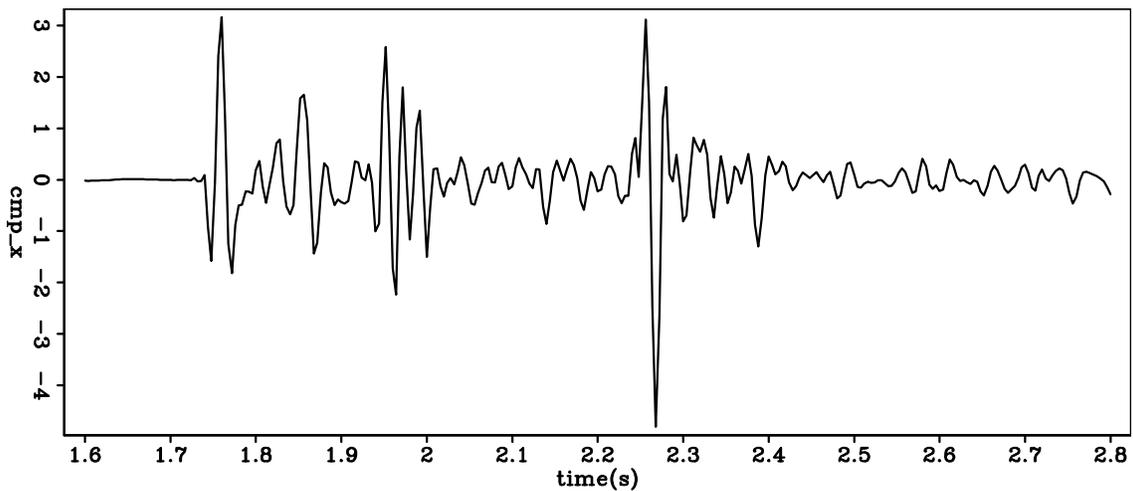
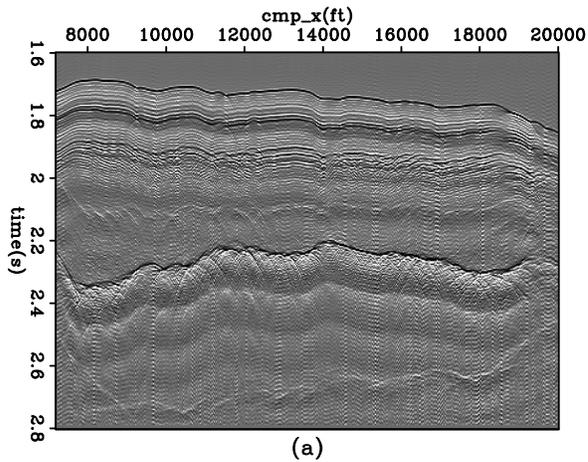
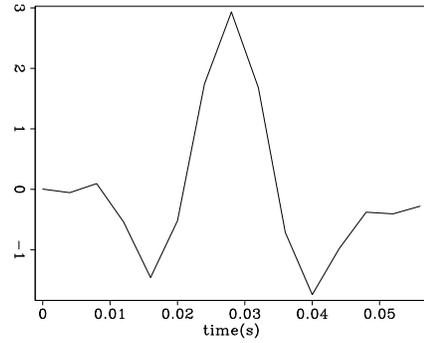


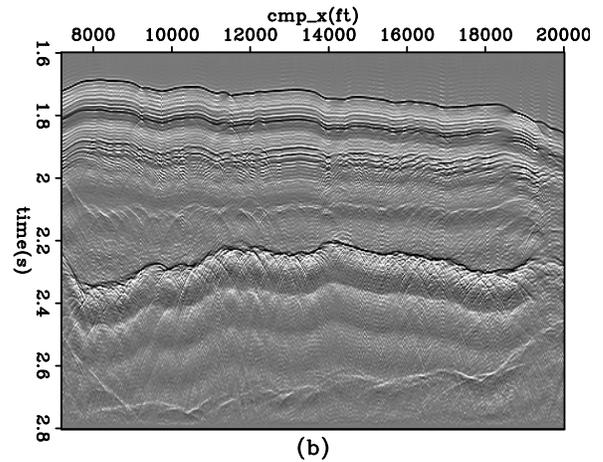
Figure 10: One single trace extracted at 12000m of the common offset section. [ER] yang1/. single-trc-COF

Figure 11: The extracted source wavelet from seismic data. [CR]

yang1/. wav-man-COF



(a)



(b)

Figure 12: (a): the L2 inversion result; (b) the hybrid inversion result. [ER]

yang1/. res-mandecon-COF

does not differ significantly from L2 result; on the other hand a large ϵ is not desirable either. Something in the data breaks our simple deconvolution model.

Figure 10 shows one trace extracted from the section. Thanks to the sparseness of the reflector, it is easy to identify the waveforms of several strong reflections. Take the very first sea bed reflection for instance, the wavelet is quite symmetric, more likely zero phase instead of minimum phase. From the lesson we have learned from the synthetic data, it is likely that the non minimum phase wavelet causes the failure of the method (4).

Fortunately, it is easy to identify the wavelets at several strong reflections, so we can roughly extract the wavelet from these locations and use this wavelet in the convolution model. The simplified inversion problem can be defined as follows:

$$\begin{cases} \mathbf{S}\mathbf{m} \approx \mathbf{d} \\ \epsilon\mathbf{m} \approx \mathbf{0} \end{cases} \quad (5)$$

in which \mathbf{S} is the convolution operator of the known wavelet \mathbf{s} , \mathbf{m} is the unknown model, and \mathbf{d} is the seismic data.

Figure 11 shows the extracted source wavelet by averaging the wavelet at the sea-floor reflection among all traces. Figure 12 shows the result of L2 inversion and hybrid inversion. Compared with the original data, both deconvolution results improve the spatial resolution; and the hybrid result is less noisy than the L2 result.

Discussion

It is not always true that wavelet can be extracted from the seismic data, in this case we have to perform *blind* deconvolution. To overcome the difficulty brought by non-minimum phase wavelet, we turn back to the original non-linear convolution model (3), and solve the non-linear inversion problem directly.

There are two ways to linearize this model. The first one is to use model perturbation and neglect the non-linear higher order terms in the following:

$$(\mathbf{s} + \Delta\mathbf{s}) * (\mathbf{m} + \Delta\mathbf{m}) \approx \mathbf{s} * \mathbf{m} + \mathbf{s} * \Delta\mathbf{m} + \mathbf{m} * \Delta\mathbf{s} = \mathbf{d},$$

in which \mathbf{m} , \mathbf{s} are the initial model and source wavelet respectively. $\Delta\mathbf{m}$, $\Delta\mathbf{s}$ are the perturbation of them, the linearized inversion will output $\Delta\mathbf{m}$, $\Delta\mathbf{s}$. The other way of linearization is a two-stage linear least squares formulation; i.e. alternately fixing one term (\mathbf{m} or \mathbf{s}) and inverting for the other one. First use an initial wavelet \mathbf{s} , keep \mathbf{s} unchanged and invert for model \mathbf{m}

$$\mathbf{S}\mathbf{m} = \mathbf{d}, \quad (6)$$

and then use the updated \mathbf{m} to invert for wavelet \mathbf{s}

$$\mathbf{M}\mathbf{s} = \mathbf{d}. \quad (7)$$

Repeat this process (6) and (7) for several iterations.

As is in all non-linear inversion problems, the difficulty in these methods is to find a good starting model. Another issue is to add proper constrain on the wavelet \mathbf{s} , for example, the wavelet should have constant energy during inversion, but this constrain does not fit the linear inversion framework.

CONCLUSION

We demonstrated that by using hybrid solver, it is robust and convenient to generate sparse models in Least-Squares Inverse imaging and deconvolution problems; although in the blind deconvolution formulation the L2 and hybrid inversion would yield similar results in the absence of minimum phase wavelet assumption.

ACKNOWLEDGMENTS

The authors would like to thank the sponsors of Stanford Exploration Project for the financial support, and also thank Bob Clapp, Luis Canales, Ali Almomin and Elita Li for the fruitful discussion.

REFERENCES

- Claerbout, J. F., 2009a, Basic earth imaging.
———, 2009b, Blocky models via the l_1/l_2 hybrid norm: SEP-Report, **139**, 1–10.
Claerbout, J. F. and F. Muir, 1973, Robust modeling with erratic data: *Geophysics*, **18**, 826–844.
Darche, G., 1989, Iterative l_1 deconvolution: SEP 61, 281–302.
Guitton, A., 2005, Multidimensional seismic noise attenuation: PhD thesis, Stanford University.
M. Clapp, R. C. and B. Biondi, 2005, Regularized least-squares inversion for 3-d subsalt imaging: *Soc. of Expl. Geophys.*, 1814–1817.
Maysami, M. and N. Mussa, 2009, Generalized-norm conjugate direction solver: SEP-Report, **139**, 11–22.
Tang, Y., 2008, Wave-equation hessian by phase encoding: SEP-Report, **134**, 1–25.
———, 2009, Inversion with phase-encoded hessian: *Geophysics*, **74**, WCA95–WCA107.
Valenciano, A., 2006, Target-oriented wave-equation inversion: *Geophysics*, **71**, A35–38.
———, 2008, Imaging by wave-equation inversion: PhD thesis, Stanford University.

SEP-140

269

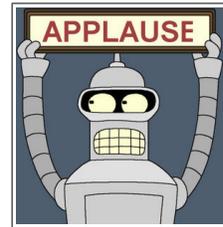
Chris Leader graduated from Oxford Univeristy in 2008 with a BA in Physics (with concentration on Astrophysics and Condensed Matter physics) and then from Imperial College London in 2009 with an MSc in Petroleum Geophysics (Distinction). He is currently a first year student in the Stanford Exploration Project on the PhD program working on Fourier methods of regularisation. Work experience involves 3D seismic processing for a Rio Tinto acquired dataset over summer 2009. He is a member of SEG, EAGE, PESGB and IOP.



Yunyue (Elita) Li graduated from China University of Petroleum, Beijing in July 2008 with a B.S. in Information and Computational Science. She joined SEP in the fall of 2008, and is currently working toward a Ph.D. in Geophysics. She is a student member of the SEG.



Nader Moussa graduated from North Carolina State University, where he studied Physics and Electrical and Computer Engineering. In 2008, he finished his M.Sc. at Stanford in Electrical Engineering, with focus on experimental radioscience and electromagnetic sensor instrumentation design. He is now working towards a Ph.D. in Geophysics with the Stanford Exploration Project.



Blocky velocity inversion by hybrid norm

Ali Ameen Almomin

ABSTRACT

Inverting a regularized Dix equation using the $L2$ norm produces smooth interval velocity models. To get a blocky interval velocity estimate that is more geologically reasonable, a regularized Dix inversion needs to be done using an $L1$ -like norm. In this paper, we compare different 2D regularizations using both the $L2$ norm and the hybrid $L1/L2$ norm to find the regularization that produces blocky velocity models while maintaining high accuracy and resolution. The results show that the hybrid $L1/L2$ norm successfully achieves blockiness when using the first derivative in multiple directions as a regularization.

INTRODUCTION

Most inversion problems try to correct for the fact that the adjoint is different than the inverse. However, in the case of the Dix formula (Dix, 1952), an exact inverse exists. The problem with the Dix formula is that it requires noise-free RMS velocity as an input. Real data always contains some level of noise. In addition, stacking velocity is measured from the data and used as the RMS velocity, even though these values can differ significantly. Therefore, the Dix formula is usually cast as an inverse problem. The conventional regularized Dix inversion (Koren and Ravve, 2006; Harlan, 1999; Clapp, 2001) produces smooth results because the inversion is optimized using the $L2$ norm. However, there are many situations, such as salt boundaries and faults, in which sharp boundaries (i.e. blocky models) are more geologically realistic.

Previous work has shown that the $L1$ norm can produce sparse and blocky results. However, $L1$ norm solvers can encounter stability and convergence issues, since the objective function defined by that norm is at the verge of convexity. Claerbout (2009) has proposed a hybrid $L1/L2$ norm in which a smooth transition between $L1$ and $L2$ can be set at any desired percentile; the hybrid norm has better convergence and stability properties than a pure $L1$ norm. Also, Maysami and Moussa (2009) describe a conjugate-direction solver in which the hybrid $L1/L2$ norm is optimized using Taylor's expansion.

Previously, Li and Maysami (2009) showed that blockiness can be achieved in 1D using a first derivative operator to regularize the problem. In this paper, we expand blockiness to 2D by testing different regularizations on two field datasets.

DIX INVERSION AS AN $L1$ -OPTIMIZATION PROBLEM

The Dix equation can be made linear by relating the square of interval velocity v to the square of RMS velocity V ,

$$v_{\tau}^2 = \tau V_{\tau}^2 - (\tau - 1)V_{\tau-1}^2, \quad (1)$$

where τ is the two-way traveltime. By defining $u_\tau = v_\tau^2$ and $d_\tau = \tau V_\tau^2$, we can set up the Dix inversion problem in an $L1$ sense as follows:

$$\|\mathbf{W}_d(\mathbf{C}\mathbf{u} - \mathbf{d})\|_{\text{hybrid}} \approx 0, \quad (2)$$

where \mathbf{W}_d is a weight function proportional to the pick strength in the velocity scan divided by τ , \mathbf{C} is the causal integration operator, and \mathbf{u} and \mathbf{d} are vectors containing all the values of u_τ and d_τ , respectively. The division by τ reduces the strength of the later events to balance the data fitting strength along the time axis.

The hybrid norm above defines the cost function as follows:

$$\mathbf{C}(\mathbf{r}) = \sqrt{\mathbf{r}^2 + \mathbf{R}^2} - \mathbf{R}, \quad (3)$$

where \mathbf{r} is the residual and \mathbf{R} is a threshold which defines a smooth transition between the $L1$ and $L2$ norms (Claerbout, 2009).

Fitting goal (2) is not enough to fully constrain the inversion, because it has a large null space (Li and Maysami, 2009). Moreover, picking errors can lead to incorrect RMS velocities and unreasonable interval velocities. Therefore, a second fitting goal (i.e. a regularization term) is required to constrain this inversion. The regularization term can be written as follows:

$$\|\epsilon\mathbf{A}\mathbf{u}\|_{\text{hybrid}} \approx 0, \quad (4)$$

where \mathbf{A} is typically a roughening operator, and ϵ is a scalar to balance the two fitting goals.

Notice that the norm in fitting goal (2) has a different effect than the norm in fitting goal (4). Using the hybrid norm in data fitting makes the inversion less sensitive to outliers. On the other hand, using the hybrid norm in model styling affects the general shape of the estimated model, which is the goal of this paper.

Li and Maysami (2009) successfully produced blockiness in 1D when using the first derivative as a regularization operator. In the following sections, we will try different regularization operators to achieve the same goals in 2D.

REGULARIZATION BY THE LAPLACIAN OPERATOR

First, we will use the Gulf of Mexico data provided by WesternGeco. This is a four-second dataset, at a sampling interval of 4 ms. The offset axis has 24 traces starting at 264 m with an increment of 134 m. There are 125 CMP gathers with a spacing of 67 m. Figure 1(b) shows the results of autopicking the dataset, and Figure 1(a) shows the results of direct Dix conversion. The direct conversion is done trace by trace after vertical smoothing which prevents the minimum interval velocity to go below water velocity. Figure 2 shows the strength of the picks in the velocity scans, which will be used as a weight as described above.

We start by choosing the symmetric Laplacian operator. Figure 3 shows the results of using the $L2$ norm and Figure 4 shows the results of using the hybrid norm. Although the two Figures look similar, the hybrid norm shows less smoothing and more detail than the $L2$ norm. The hybrid norm results are still not blocky, because a linear trend in velocity will also result in a zero second derivative. In the next section we attempt to more closely approach the first derivative by using the helix derivative.

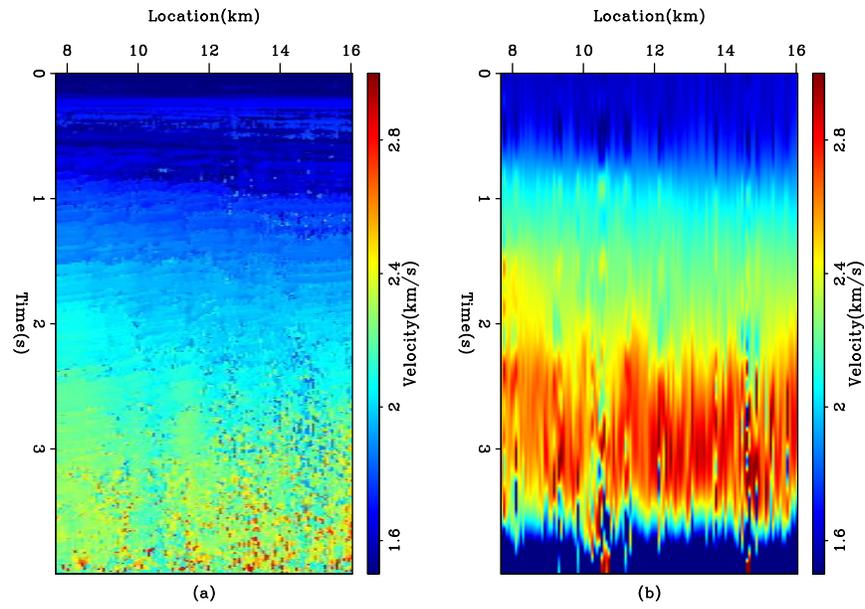
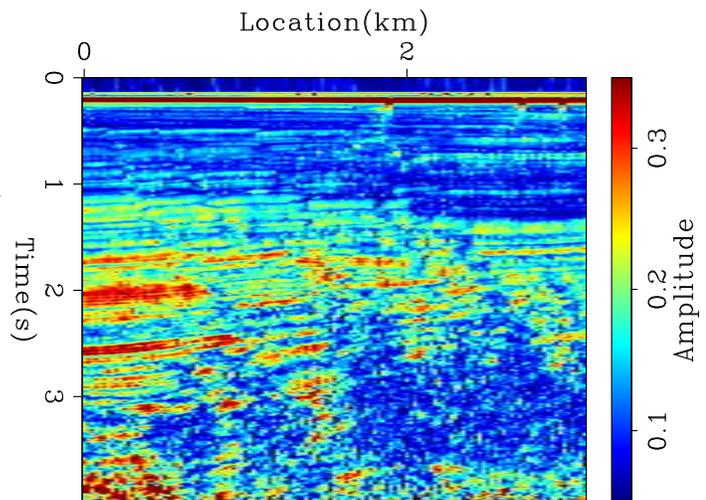


Figure 1: The WG dataset. (a) The input RMS velocity, which is automatically picked from the CMP gathers. (b) The interval velocity by direct dix conversion. [ER] ali1/. lab0

Figure 2: The strength of the picks in the velocity scan of the WG dataset, which is used as the weight before dividing by time. [ER] ali1/. lab-weight2d



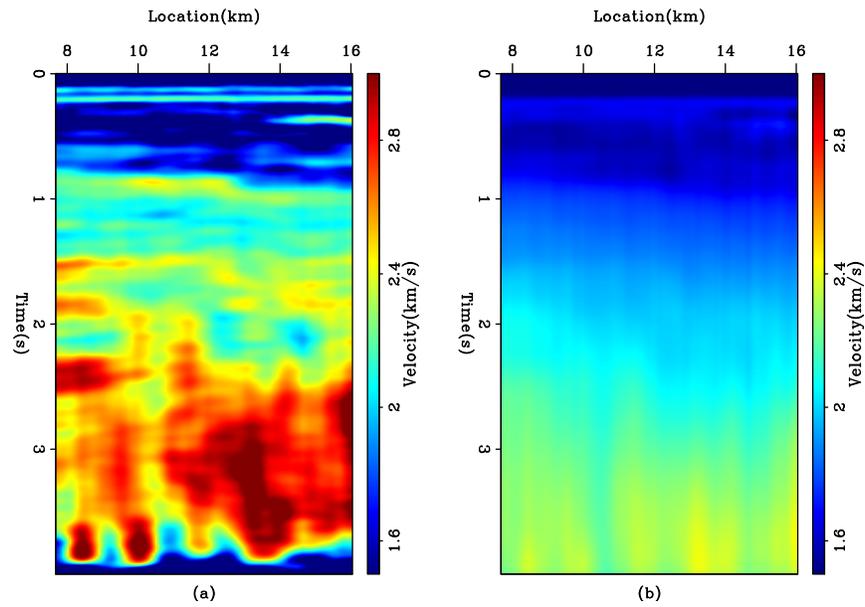


Figure 3: The WG dataset. (a) The interval velocity estimated by using the Laplacian operator as a regularization in the L_2 norm. (b) The reconstructed RMS velocity. [ER] ali1/. l2-lab18

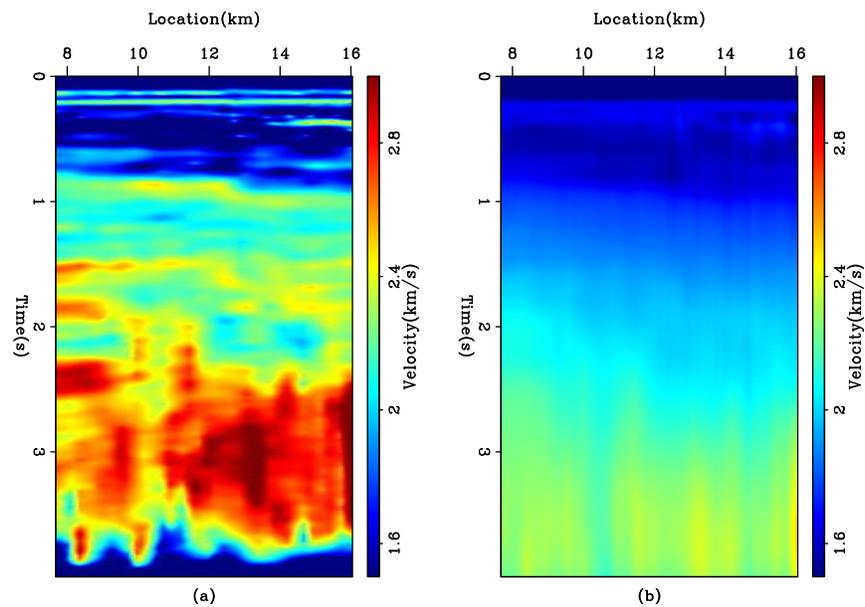


Figure 4: The WG dataset. (a) The interval velocity estimated by using the Laplacian operator as a regularization in the hybrid norm. (b) The reconstructed RMS velocity. [ER] ali1/. hbe-lab19

REGULARIZATION BY THE HELIX DERIVATIVE OPERATOR

Now we consider the helix derivative (Claerbout, 1997) as a regularization operator. Figure 5 shows the results of using the L_2 norm, and Figure 6 shows the results of using the hybrid norm. In this case, we see a dramatic difference between the two results. In the hybrid norm case, we can see the beginnings of blockiness, but only in one direction (toward the right). The reason for this asymmetry is that using an L1-like norm is similar to applying the regularization only once. On the other hand, we do not see this effect in the L_2 norm results, because the regularization in that norm is similar to applying the forward and the adjoint of an operator, which is a symmetric procedure.

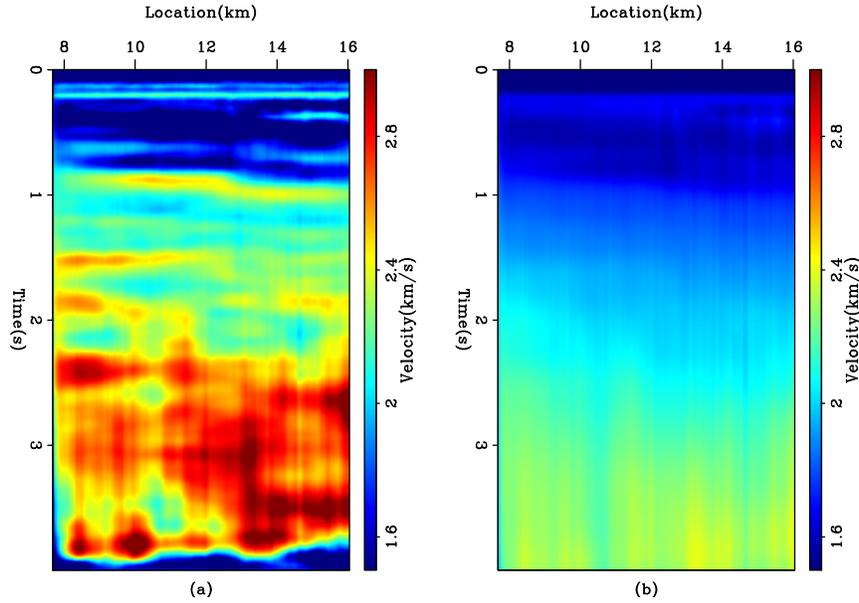


Figure 5: The WG dataset. (a) The interval velocity estimated by using the helix derivative operator as a regularization in the L_2 norm. (b) The reconstructed RMS velocity. [ER] ali1/. l2-lab27

REGULARIZATION BY THE FIRST DERIVATIVE OPERATOR IN TWO DIRECTIONS

The previous regularizations show that only a first derivative can create blockiness. However, using the first derivative means that we must pick a direction each time we apply the derivative. As a first test, we pick two directions: the vertical and horizontal as follows:

$$\|\epsilon \mathbf{D}_z \mathbf{u}\|_{\text{hybrid}} \approx 0, \quad (5)$$

$$\|\epsilon \mathbf{D}_x \mathbf{u}\|_{\text{hybrid}} \approx 0, \quad (6)$$

where \mathbf{D}_z and \mathbf{D}_x are the first derivative operators along the z- and x-axis, respectively. The derivative of each direction is applied in a separate regularization equation (i.e. we have two regularization equations in this case) in order to maintain symmetry. Combining these two filters in one regularization will cause an asymmetry in blockiness, similar to the previous result from the helix derivative regularization.

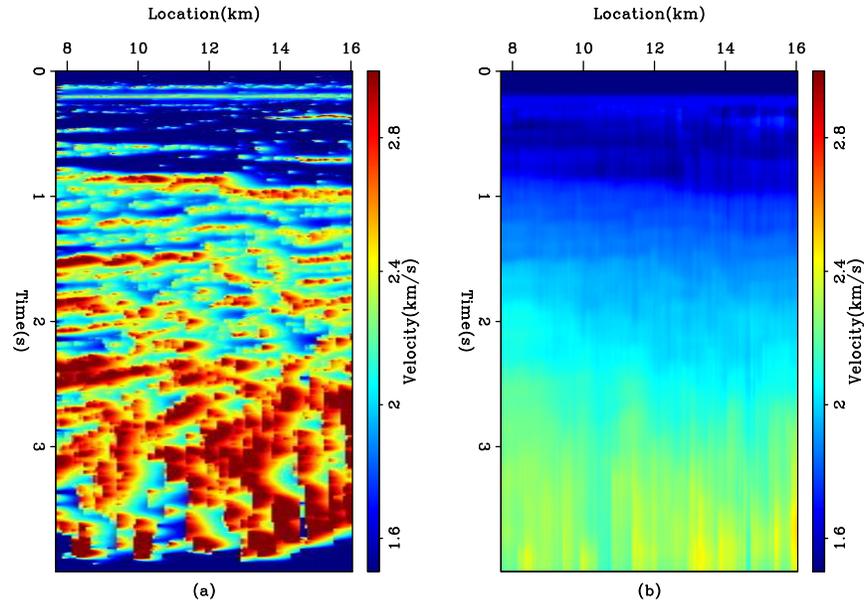


Figure 6: The WG dataset. (a) The interval velocity estimated by using the helix derivative operator as a regularization in the hybrid norm. (b) The reconstructed RMS velocity. [ER] ali1/. hbe-lab29

Figure 7 shows the results of using the $L2$ norm with two first derivative applications, and Figure 8 shows the results of using the hybrid norm. Blockiness is clearly present in the hybrid norm results. However, there seems to be a preference for the sharp boundaries to be either horizontal or vertical, which is due to the directions of the derivatives we chose.

REGULARIZATION BY THE FIRST DERIVATIVE OPERATOR IN FOUR DIRECTIONS

To reduce the bias in blockiness directions, we increased the directions of the first derivative to four: the previous two directions, plus two directions at 45 degrees to the vertical and horizontal axes. Figure 9 shows the results of using the $L2$ norm, and Figure 10 shows the results of using the hybrid norm. By comparing Figure 10 to Figure 8, we can clearly see a large improvement in the model. The details and blockiness are still preserved. However, the inversion now has more “freedom” to pick the direction of blockiness out of four directions instead of two directions.

ELF DATASET

In this section, we will estimate an interval velocity model of the North Sea data provided by ELF. There is a known salt body in the middle of this data, which makes it a proper test case for our blockiness goals. Also, this dataset has better spatial sampling than the previous dataset, and can thus better illustrate the differences between the different inversion results. This is also four-second dataset, at a sampling of 5.9 ms. The offset axis has 143 traces starting at 0 m with an increment of 25 m. There are 537 CMP gathers with a spacing of

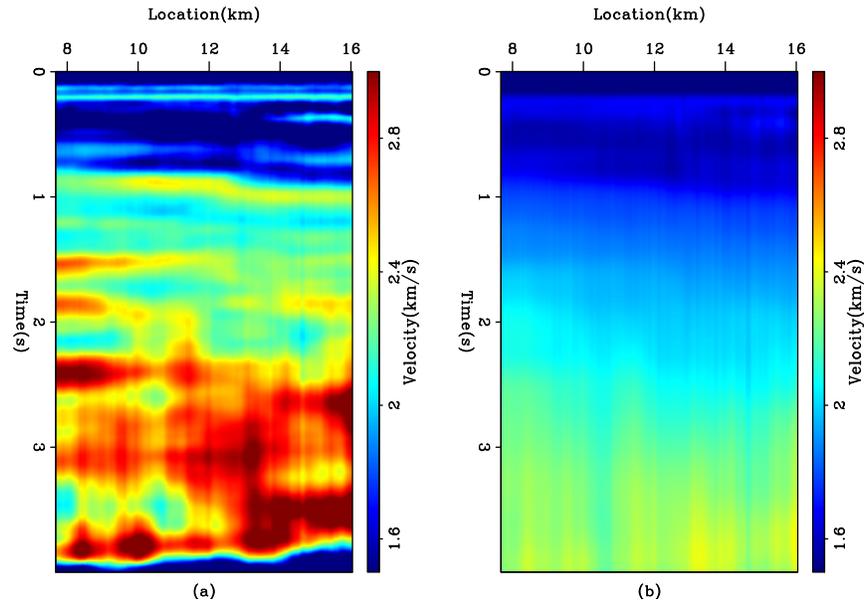


Figure 7: The WG dataset. (a) The interval velocity estimated by using the first derivative operator in two directions as a regularization in the L_2 norm. (b) The reconstructed RMS velocity. [ER] [ali1/. l2-lab39](#)

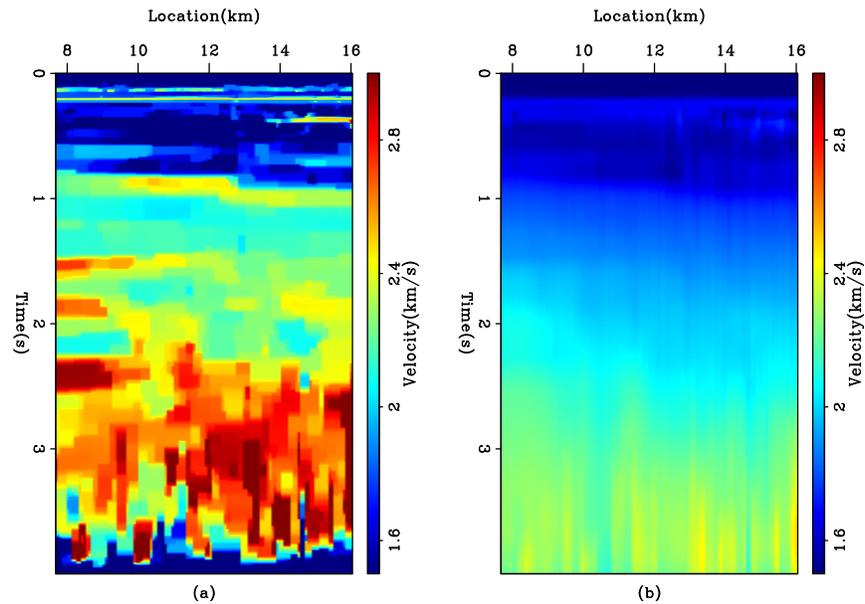


Figure 8: The WG dataset. (a) The interval velocity estimated by using the first derivative operator in two directions as a regularization in the hybrid norm. (b) The reconstructed RMS velocity. [ER] [ali1/. hbe-lab37](#)

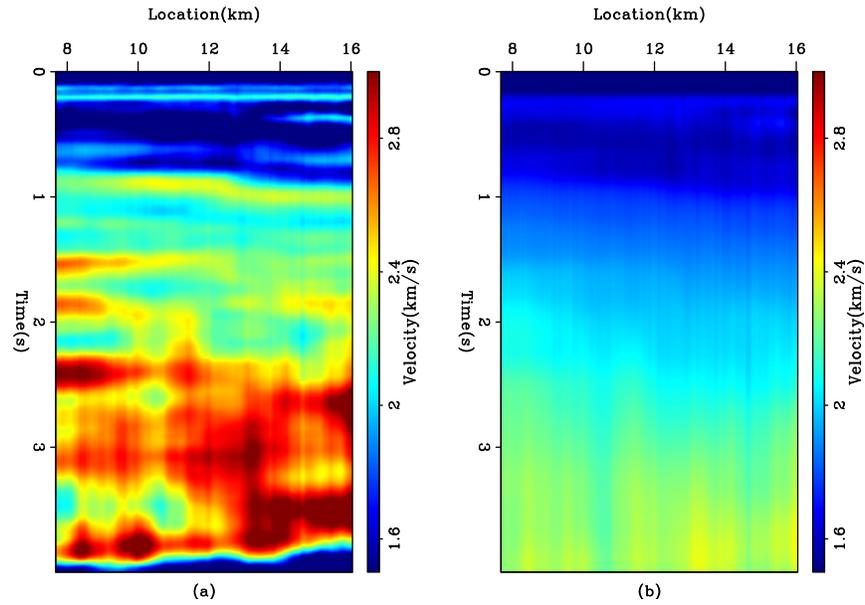


Figure 9: The WG dataset. (a) The interval velocity resulted by using the first derivative operator in four directions as a regularization in the L_2 norm. (b) The reconstructed RMS velocity. [ER] [ali1/. l2-lab46](#)

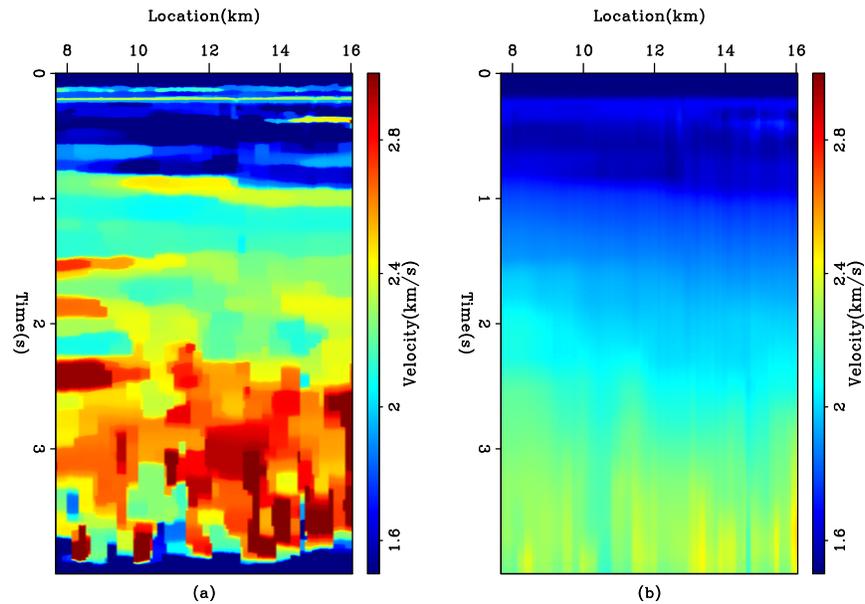


Figure 10: The WG dataset. (a) The interval velocity resulted by using the first derivative operator in four directions as a regularization in the hybrid norm. (b) The reconstructed RMS velocity. [ER] [ali1/. hbe-lab45](#)

25 m. Figure 11(b) shows the results of autopicking the dataset, which was constrained by the background RMS velocity, and Figure 11(a) shows the results of direct Dix conversion, as defined above. Figure 12 shows the strength of the picks in the velocity scans.

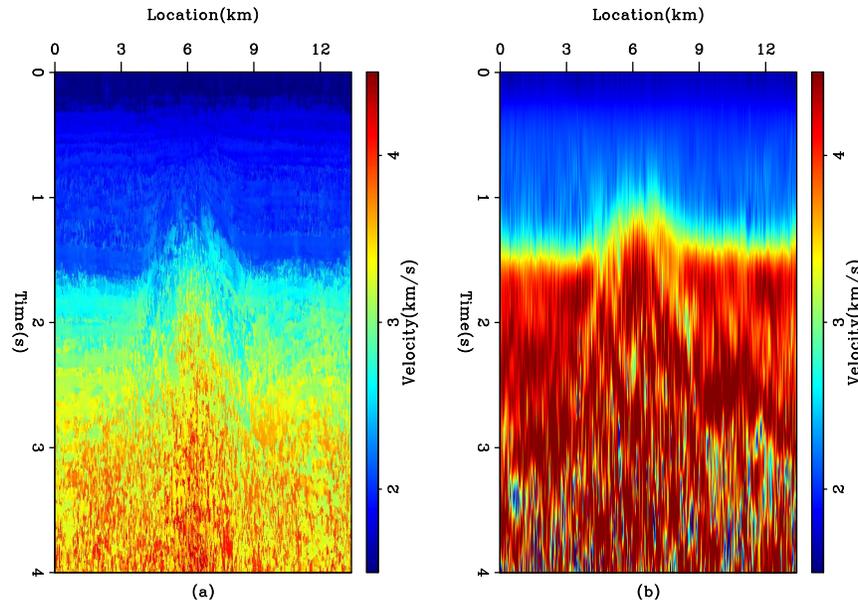


Figure 11: The ELF dataset. (a) The input RMS velocity which is automatically picked from the CMP gathers. (b) The interval velocity by direct Dix conversion. [ER] ali1/. elf0

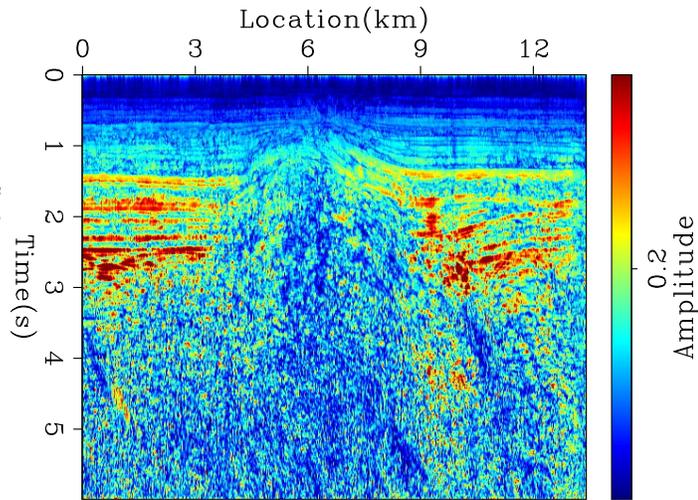


Figure 12: The strength of the picks in the velocity scan of ELF dataset, which is used as the weight before dividing by time. [ER] ali1/. elf-weight

For this dataset, we will only repeat the last two regularizations, which are the first derivative in two and four directions, since they showed the best results with the most symmetric blockiness. Figure 13 shows the results of regularizing in two directions in the L_2 norm and Figure 14 shows the results of the same regularization in the hybrid norm. Figures 15 and 16 show the results of using four-direction regularization.

Since the dataset is larger with smaller sampling, the improvement of using more directions is evident. Forcing the inversion to pick between two directions has an apparent effect of reducing the resolution. One obvious example is the chalk layer, which looks very

horizontal in Figure 14 but more detailed in Figure 16. In all cases, $L2$ always gives smooth results, which smear the model and lower the resolution of the inversion.

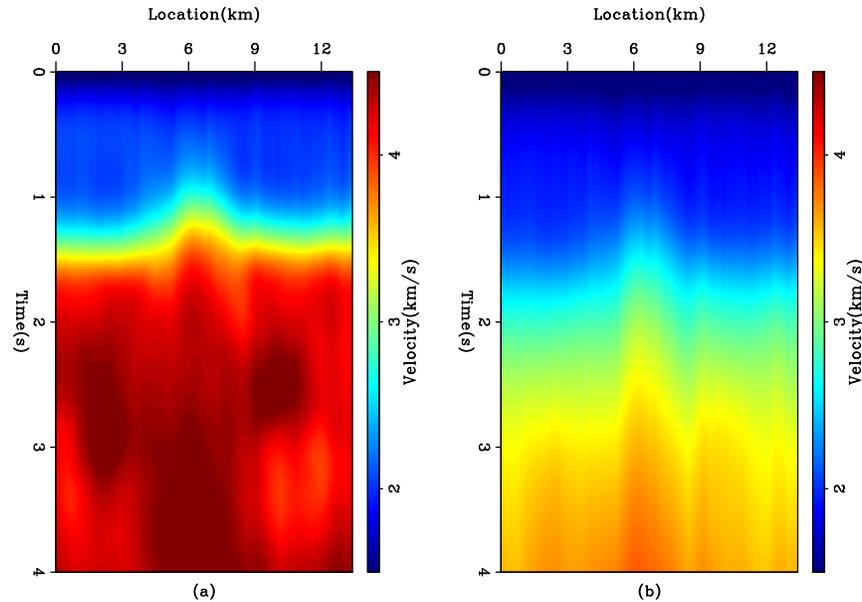


Figure 13: The ELF dataset. (a) The interval velocity estimated by using the first derivative operator in two directions as a regularization in the $L2$ norm. (b) The reconstructed RMS velocity. [ER] [ali1/. l2-elf14](#)

CONCLUSIONS AND DISCUSSIONS

We successfully achieved blocky velocity models by using the hybrid norm. We also showed that the choice of the regularization operator has a great impact on how blocky the results are. Nonetheless, the hybrid norm always showed more detail and resolution than the $L2$ norm, even when blockiness was not achieved. An example of this can be seen by comparing Figures 4 and 9. Although the first Figure uses a Laplacian operator for regularization and the later Figure uses the first derivative in four directions (which we showed has the best results), the hybrid norm was still superior to the $L2$ norm in preserving more details and showing higher resolution.

Another point to keep in mind is that the hybrid norm has a flexible threshold. In all previous cases, we set that threshold to 0.20, meaning that 80 percent of residuals (both data residuals and model residuals) are going to be in the $L1$ region and the rest in the $L2$ region. However, this is a parameter that can be adjusted based on the desired degree of blockiness.

FUTURE WORK

As we have seen, increasing the number of directional derivatives will increase the resolution and flexibility of the inversion results. However, increasing the number of directions will also slow the inversion, because each direction has a model residual the size of the model.

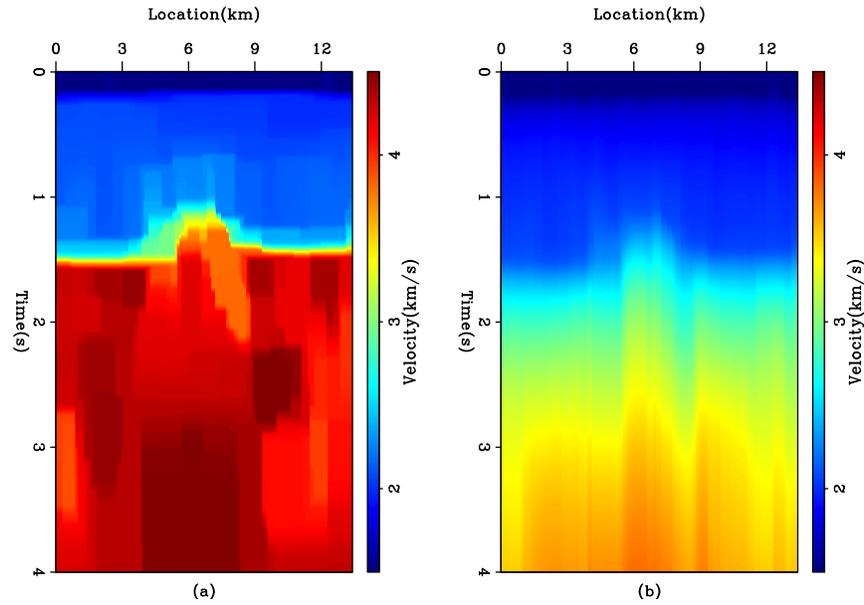


Figure 14: The ELF dataset. (a) The interval velocity estimated by using the first derivative operator in two directions as a regularization in the hybrid norm. (b) The reconstructed RMS velocity. [ER] ali1/. hbe-elf17

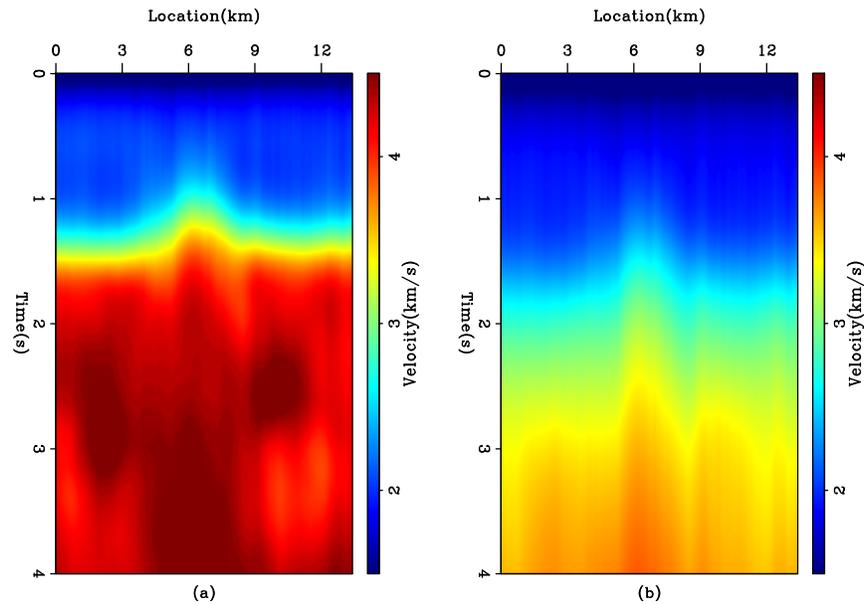


Figure 15: The ELF dataset. (a) The interval velocity estimated by using the first derivative operator in four directions as a regularization in the $L2$ norm. (b) The reconstructed RMS velocity. [ER] ali1/. 12-elf23

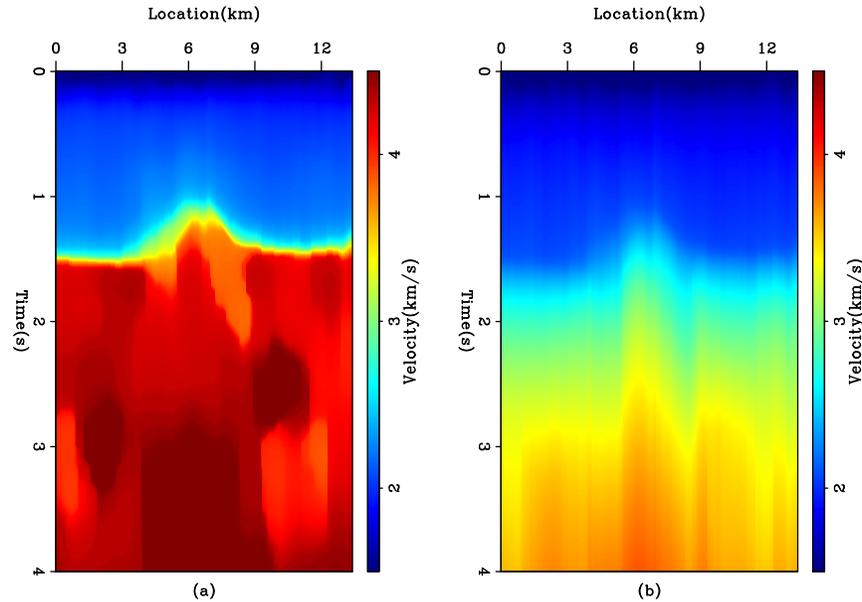


Figure 16: The ELF dataset. (a) The interval velocity estimated by using the first derivative operator in four directions as a regularization in the hybrid norm. (b) The reconstructed RMS velocity. [ER] `ali1/. hbe-elf29`

Instead of using many directions, it is possible to use steering filters to pre-define the local direction of maximum variance and then use that information to align the directions of the regularization to be parallel and perpendicular to it. This way, we might only need two directions in the regularization.

ACKNOWLEDGMENTS

We thank Jon Claerbout, Robert Clapp, Antoine Guitton, Louis Vaillant, Elita Li and Yang Zhang for their helpful suggestions and discussions throughout this project. We also thank WesternGeco and TotalFinaElf for providing the data.

REFERENCES

- Claerbout, J. F., 1997, Multidimensional recursive filters via a helix: SEP-Report, **95**, 1–13.
 ———, 2009, Blocky models via the l1/l2 hybrid norm: SEP-Report, **139**, 1–10.
 Clapp, R. G., 2001, Geologically constrained migration velocity analysis: PhD thesis, Stanford University.
 Dix, C. H., 1952, Seismic prospecting for oil.
 Harlan, W. S., 1999, Constrained dix inversion.
 Koren, Z. and I. Ravve, 2006, Constrained dix inversion: Geophysics, **71**.
 Li, Y. E. and M. Maysami, 2009, Dix inversion constrained by l1-norm optimization: SEP-Report, **139**, 23–36.
 Maysami, M. and N. Moussa, 2009, Generalized-norm conjugate direction solver: SEP-Report, **139**, 11–22.

Geophysical applications of a novel and robust L1 solver

Yunyue Li, Yang Zhang, and Jon Claerbout

ABSTRACT

L1-norm is better than L2-norm at dealing with noisy data and yielding blocky models, features crucial in many geophysical applications. In this report, we develop a hybrid-norm solver proposed by Claerbout (2009) to perform L1 regressions. The solver is tested on a 1-D field RMS velocity inversion, a 2-D regularized Kirchhoff migration inversion and a 2-D velocity analysis problem. The results of the inversions show that this solver can yield “blocky” models, and has the advantage of straightforward parametrization.

INTRODUCTION

L1 norm optimization is known to be a powerful estimator when the data are noisy or the model is sparse (Claerbout and Muir, 1973; Darche, 1989; Nichols, 0994; Guitton, 2005). However, the most widely used L1 solver—Iterated Reweighted Least-Squares (IRLS)—is cumbersome to use because users must specify numerical parameters with unclear physical meanings. To develop a robust, efficient L1 solver, Claerbout (2009) proposed a hybrid norm function to approximate the L1 norm (absolute value function), and he generalized the conjugate direction (CD) method using Taylor’s expansion to guide the plane search in the minimization.

The parametrization of this proposed hybrid norm is straightforward. Users specify thresholds for the data residual and model residual, respectively. These thresholds determine the transition point from L2 to L1. The threshold for the data residual (R_d) can be chosen according to the signal-to-noise ratio in data space; the threshold for the model residual (R_m) can be specified by the desired blockiness in the model space.

In theory, the convergence of this hybrid norm solver is guaranteed, because the objective function is strictly convex. Nevertheless, difficulties may occur as the hybrid-norm approaches the L1 limit.

To test the performance and analyze the stability of this hybrid solver, we apply the solver to a 1-D field RMS velocity inversion, a 2-D regularized Kirchhoff migration inversion, and a 2-D velocity analysis problem. The inversion results show that this hybrid solver is robust and simple to use.

GENERALIZED CONJUGATE DIRECTION METHOD FOR THE HYBRID NORM

The hybrid norm is defined as

$$h(r) = \sqrt{r^2 + R^2} - R. \quad (1)$$

where r is the residual, and R is the corresponding threshold. In the limit, the hybrid norm (1) becomes:

$$h(r) = \begin{cases} |r| - R, & \text{if } R \ll |r| \\ r^2/(2R), & \text{if } R \gg |r|. \end{cases} \quad (2)$$

It is obvious that when R is small, the hybrid norm (1) reduces to L1 norm; when R is big, it becomes the L2 norm. Therefore, threshold R behaves as the turning point where the objective function changes smoothly from L2 to L1.

The Conjugate Direction method is commonly used for solving immense linear regressions in exploration geophysics. The idea of the CD method is to search the plane determined by the gradient and the previous step for the best step direction and length, instead of moving along the gradient direction. The best direction in that plane is the linear combination of the gradient and previous step vector that decreases the measure of the residual the most. Traditionally, the measure is chosen to be L2, for its simplicity; however, we generalize the CD method for any arbitrary convex measure C . Readers can determine which measure to use to satisfy their own objectives.

Now let us examine the generalization of the CD method in detail. At each iteration, we have the residual vector \bar{r} , the gradient g and the previous step s . Therefore, the updated residual can be written as:

$$r_i = \bar{r}_i + \alpha g_i + \beta s_i. \quad (3)$$

where α and β are scalars controlling the relative weights of these two directions. To determine these two scaling parameters, we need to minimize the measure of the residual:

$$N(\alpha, \beta) = \sum_i C(\bar{r}_i + \alpha g_i + \beta s_i). \quad (4)$$

The system given by directly setting the partial derivatives of (4) to zero is transcendental, thus difficult to solve. Therefore, we use Taylor's expansion to approximate the original objective function. The polynomial estimation of (4) is given as follows:

$$N(\alpha, \beta) \approx \sum_i \left(C_i + (\alpha g_i + \beta s_i) C'_i + (\alpha g_i + \beta s_i)^2 C''_i / 2 \right). \quad (5)$$

Now, taking the derivatives of the parabolic function in (5) with respect to α and β and setting them to zero, we end up with a linear system of α and β :

$$\left\{ \sum_i C''_i \begin{bmatrix} g_i \\ s_i \end{bmatrix} (g_i \quad s_i) \right\} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = - \sum_i C'_i \begin{bmatrix} g_i \\ s_i \end{bmatrix}. \quad (6)$$

where (C_i, C'_i, C''_i) refer to a Taylor expansion of $C(r)$ about r_i .

Then we can obtain α, β by simply solving a set of 2×2 linear equations.

Notice that the α, β we get here is minimizing the approximated function (5), not the original objective function. Therefore, it is necessary to solve for α, β multiple times within each CD iteration. By doing this relatively cheap plane-search loop, we expect to save the number of iterations for the outer loop (Conjugate Direction), which is usually much more computational intensive (requiring application of both the forward and adjoint operator).

DIX INVERSION OF INTERVAL VELOCITY ESTIMATION

The Dix equation (Dix, 1952) inverts interval velocities from RMS velocity, which is picked during velocity scanning in prestack seismic data. The equation can be written as

$$v_{int(k)}^2 = kV_k^2 - (k-1)V_{k-1}^2, \quad (7)$$

or

$$\sum_{i=1}^k v_{int(i)}^2 = kV_k^2, \quad (8)$$

where v is interval velocity, V is RMS velocity, and k is the sample number, which can be regarded as travel-time depth. Directly calculating the interval velocity from this formula can easily yield wildly unreasonable results because of the error in the picked RMS velocity. Therefore, it is necessary to solve this problem as a regularized inversion. To linearize the problem, we choose the model space to be the squared interval velocity (v_{int}^2), instead of the interval velocity itself (v_{int}).

Thus we can formulate the Dix inversion problem as follows:

$$\mathbf{W}_d(\mathbf{C}\mathbf{u} - \mathbf{d}) \approx \mathbf{0}, \quad (9)$$

$$\epsilon \mathbf{D}_z \mathbf{u} \approx \mathbf{0}. \quad (10)$$

In the data-fitting goal (9), \mathbf{u} is the unknown model we are inverting for, \mathbf{d} is the known data computed from the RMS velocity, \mathbf{C} is the causal integration operator and \mathbf{W}_d is a data residual weighting function, which is a measure of our confidence in the RMS velocity. In the model-styling goal (10), \mathbf{D}_z is the vertical derivative of the velocity model and ϵ is the weight controlling the strength of the regularization.

The input RMS velocity with 1000 samples is shown in Figure 1. It is obvious that the violent variation at the end of the trace is not realistic. Thus, we use the hybrid-norm to ignore the large residuals in the data-fitting, which are considered to be noise. At the same time, to obtain a blocky interval velocity model, the large residual in the derivative of the interval velocity should be “invisible” to the measure. Therefore, the hybrid norm on the model styling appear to be the best choice.

To compare the inversion result, we also use the IRLS and L2 solver on the same data with comparable parameters. The inversion results are shown in Figure 2. The left column is the inverted interval velocity, while the right column is the corresponding reconstructed RMS velocity. The result shows that compared with the IRLS and L2 result, the hybrid solver successfully retrieved the most blocky velocity model, and the corresponding reconstructed RMS velocity contains less noise while keeping the trend of the original data.

SIMPLE KIRCHHOFF INVERSION

Kirchhoff migration was widely used before the era of wave-equation migration for marine data, and is still the principle migration method for land data. It always involves summing over or spreading along certain travel-time surfaces in 3-D, which reduce to curves in 2-D. For the purpose of testing our solver, we define the forward operator to be the Kirchhoff modeling operator, whose adjoint is the traditional Kirchhoff migration operator.

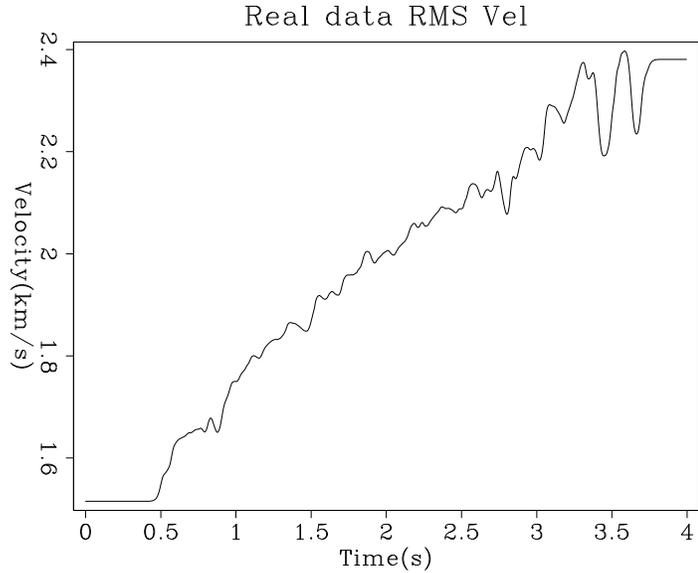


Figure 1: Input 1-D RMS velocity.

[ER] `elita1/. input-dix-real`

We formulate the inversion problem as follows:

$$\mathbf{H}\mathbf{m} \approx \mathbf{d} \quad (11)$$

$$\epsilon\mathbf{m} \approx \mathbf{0} \quad (12)$$

where \mathbf{H} is the forward Kirchhoff modeling operator, \mathbf{m} is the subsurface reflectivity model, and \mathbf{d} is the seismic response recorded at the surface. The second equation is a damping term, where the hybrid-norm is applied to retrieve the sparse model.

In field acquisition, data usually have denser sampling rates in the in-line direction than the cross-line direction. Therefore, the surface-recorded data are always aliased in the cross line direction. To illustrate the problem in the cross-line direction, figure 3 shows an example of highly aliased hyperbolas. The aliasing makes the inversion problem an underdetermined problem; therefore, the result of the inversion relies heavily on the regularization. With the model space sampling being 128×128 , the sampling of data space is only 128×16 . Also note that some of the hyperbolas are not symmetric; therefore the tops of the hyperbolas are shifted.

Same as the previous example, we experimented with different solvers: L2, IRLS and hybrid, to compare their results.

Figure 4 shows the inversion results with different schemes. The results show that the hybrid norm is superior for retrieving the spiky result that resembles the original model the best. Although severely aliased, the inversion result recovers the exact position, the correct size and most of the amplitude. Notice that the CD hybrid solver recovers the very low amplitude spike at the left edge. This promising result suggests that by choosing the regularization properly, we can overcome the aliasing problem in the presence of a sparse model.

Figure 5 shows the reconstructed data from the CD hybrid solver. The original data are accurately recovered. Notice the hyperbola with its top at the left edge is well resolved.

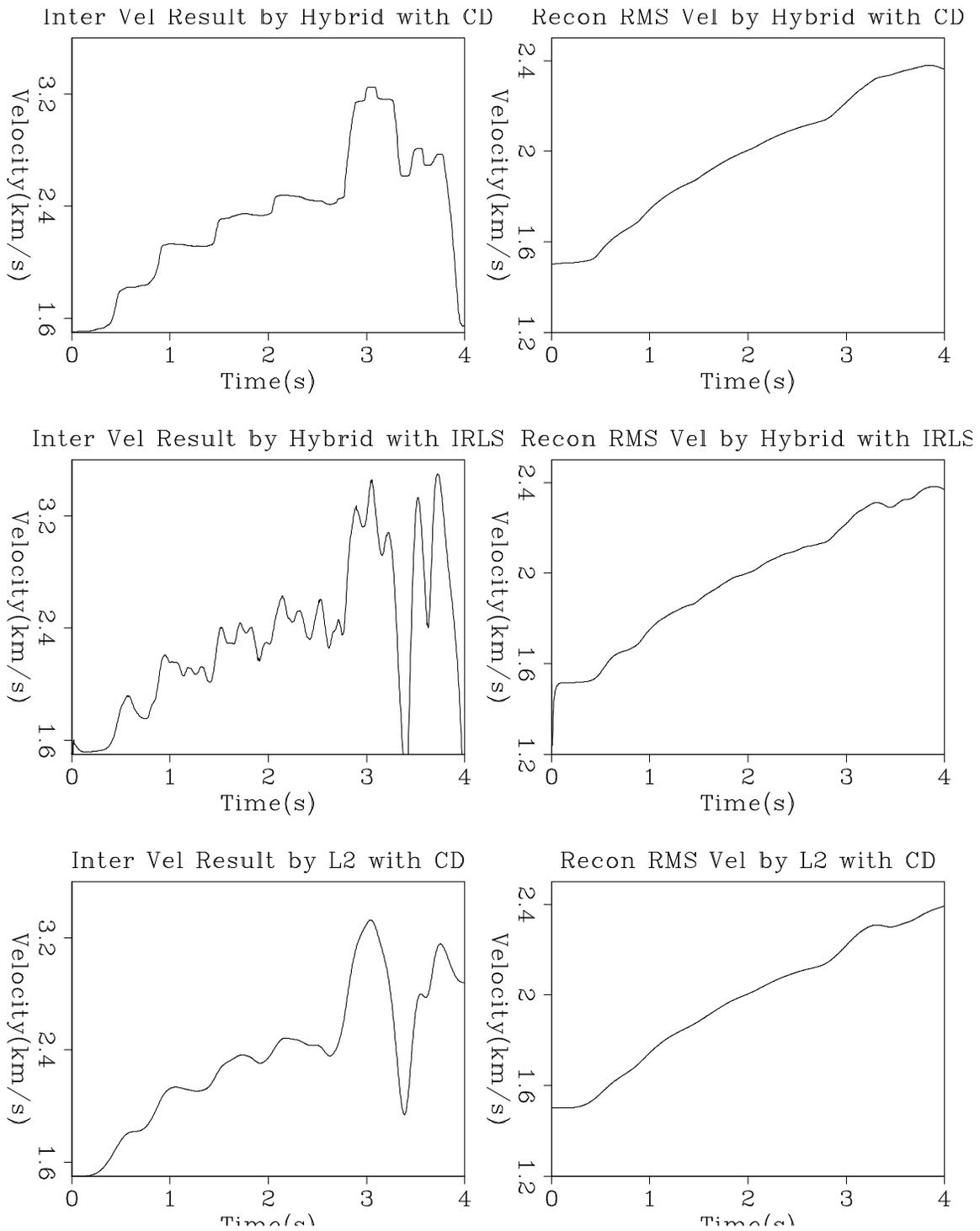


Figure 2: Comparison of the inversion results. Panels in the left column are the estimated interval velocity, while panels on the right are the corresponding reconstructed RMS velocity. Top panels: results of the hybrid with CD; Middle panels: results of the hybrid with IRLS; Bottom panels: results of the L2 norm with CD. Notice that although the reconstructed RMS velocity from the three methods are more or less the same, the interval velocity from hybrid CD is more blocky than the other two. [ER] `elita1/. dix-real`

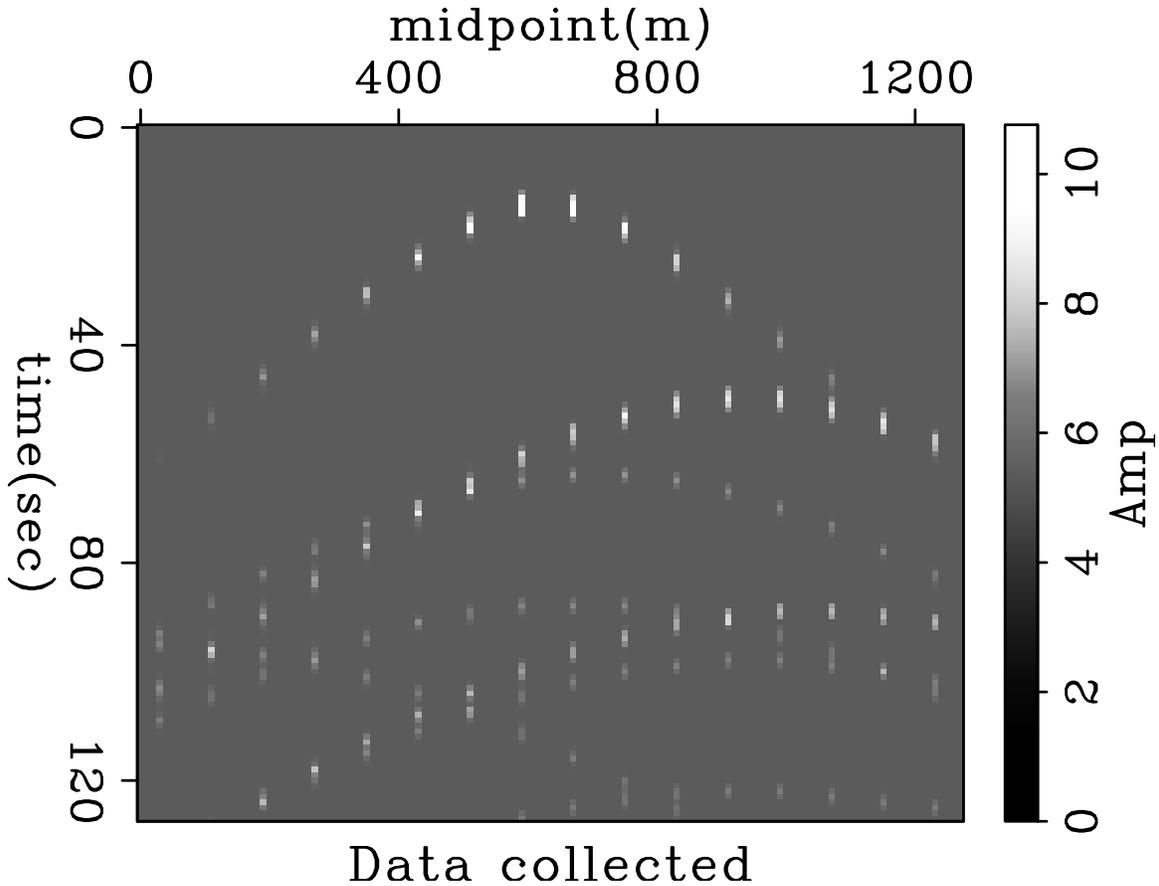


Figure 3: Highly aliased hyperbola. Input data for the Kirchhoff inversion. [ER] `elita1/. datasub`

VELOCITY ANALYSIS AS INVERSION

Velocity analysis is one the most critical and problematic procedures in seismic exploration. In data with noise bursts, velocity analysis is prone to error and even unrealistic results. Therefore, to handle this problem robustly, we formulate velocity analysis as an inversion problem as follows:

$$\mathbf{H}\mathbf{m} \approx \mathbf{d} \quad (13)$$

where \mathbf{H} is the modeling operator, whose adjoint operator is the slowness scan operator; \mathbf{m} is the slowness field, and \mathbf{d} is the data we collect after one shot.

Figure 6 shows a shot gather from Yilmaz's dataset. There are two distinct types of noise in these data: first is the linear noise caused by all kinds of surface waves, which can be attenuated by taking advantage of their physical properties; second is the abnormally high-amplitude burst noise at the near offsets, which is difficult to fit into a statistical model.

Figure 7 shows the inversion results for different methods. Because of the existence

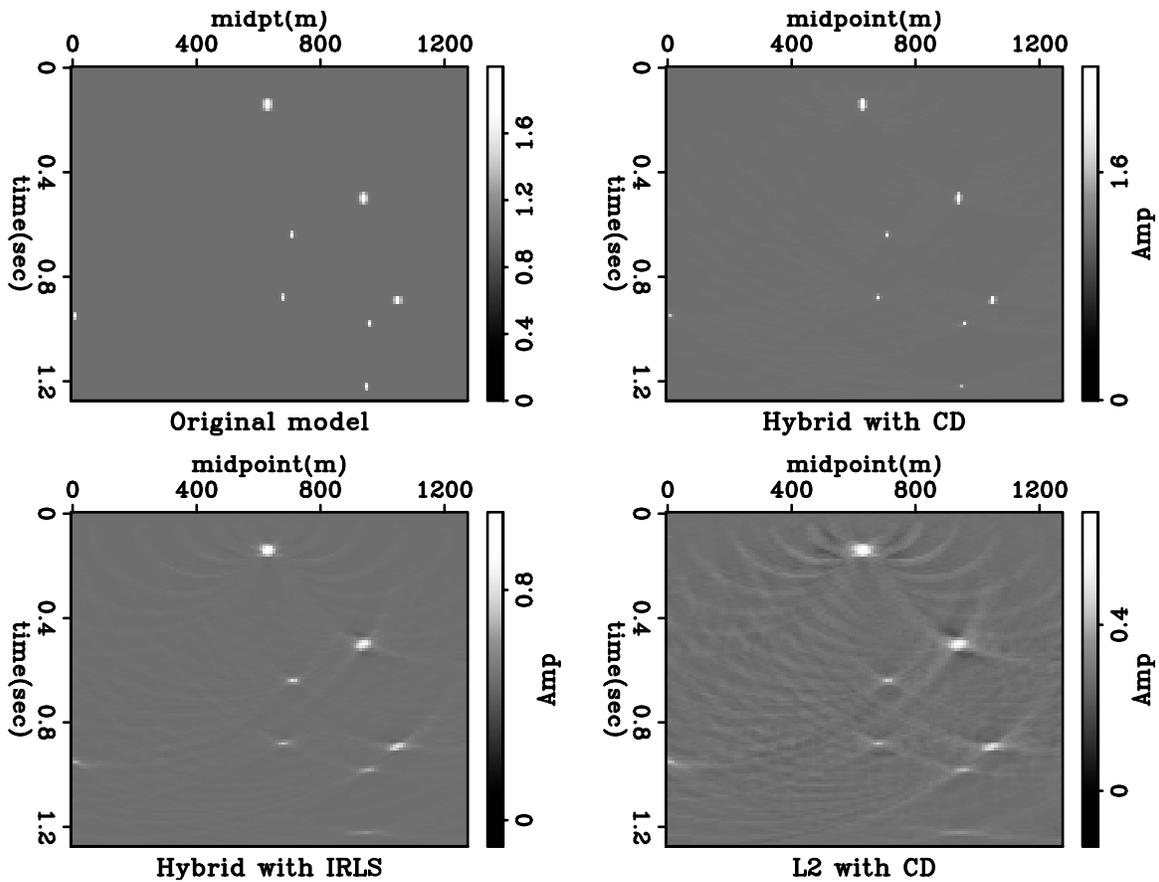


Figure 4: Original model and inversion results by different methods. Top left: original model; Top right: Hybrid result with CD; Bottom left: Hybrid result with IRLS; Bottom right: L2 result with CD. [ER] [elita1/. show-mod-rand-8](#)

of the high amplitude noise at the near-offset, a velocity scan without inversion yields no meaningful result. For the L2 inversion, the noise has contaminated the whole panel, making it impossible to see the velocity trend. The inversion results of both IRLS and HYCD show clear velocity trends, and the near-offset burst noise is successfully removed in the reconstructed data.

CONCLUSIONS

An L1-type optimizer is required to retrieve a blocky model or extract information from noisy data. To avoid the troublesome parameter tuning of IRLS, we develop a hybrid-norm Conjugate Direction solver with straightforward parametrization to achieve the goal in L1 sense. Tests of this solver on three different problems, with both synthetic and field data, show that this hybrid-norm solver is robust and simple to use. These promising results encourage us to apply this solver on large-scale, real-world problems.

ACKNOWLEDGMENTS

We thank Mohammad Maysami, Mandy Wong, and Nader Moussa for the earlier work. We like to thank Robert Clapp, Luis and Ali Almomin for fruitful discussions about the development of this solver and the choice of the test cases.

REFERENCES

- Claerbout, J. F., 2009, Blocky models via the l_1/l_2 hybrid norm: SEP-Report, **139**, 1–10.
Claerbout, J. F. and F. Muir, 1973, Robust modeling with erratic data: Geophysics, **18**, 826–844.
Darche, G., 1989, Iterative l_1 deconvolution: SEP 61, 281–302.
Dix, C. H., 1952, Seismic prospecting for oil.
Guitton, A., 2005, Multidimensional seismic noise attenuation: PhD thesis, Stanford University.
Nichols, D., 0994, Velocity-stack inversion using l^p norms: SEP 82, 1–16.

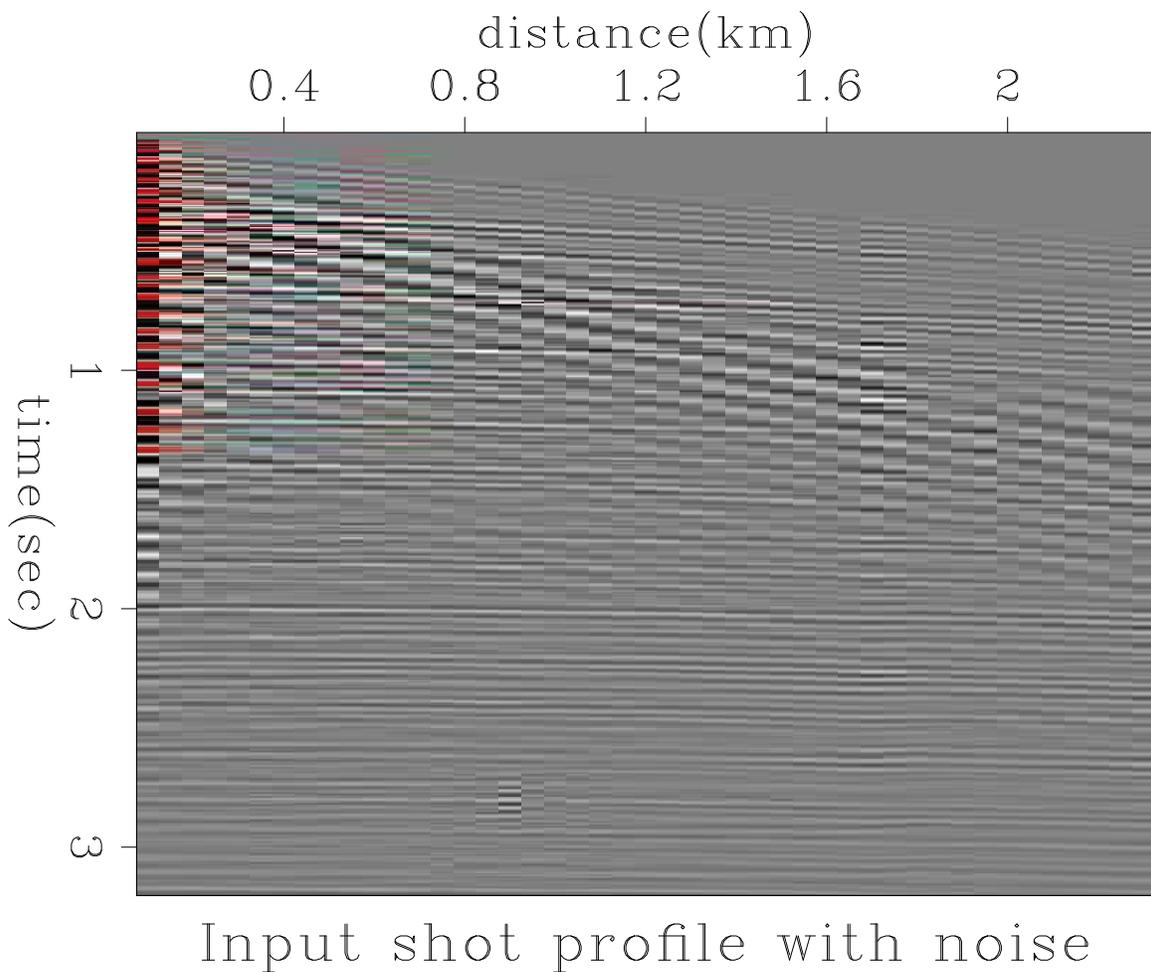
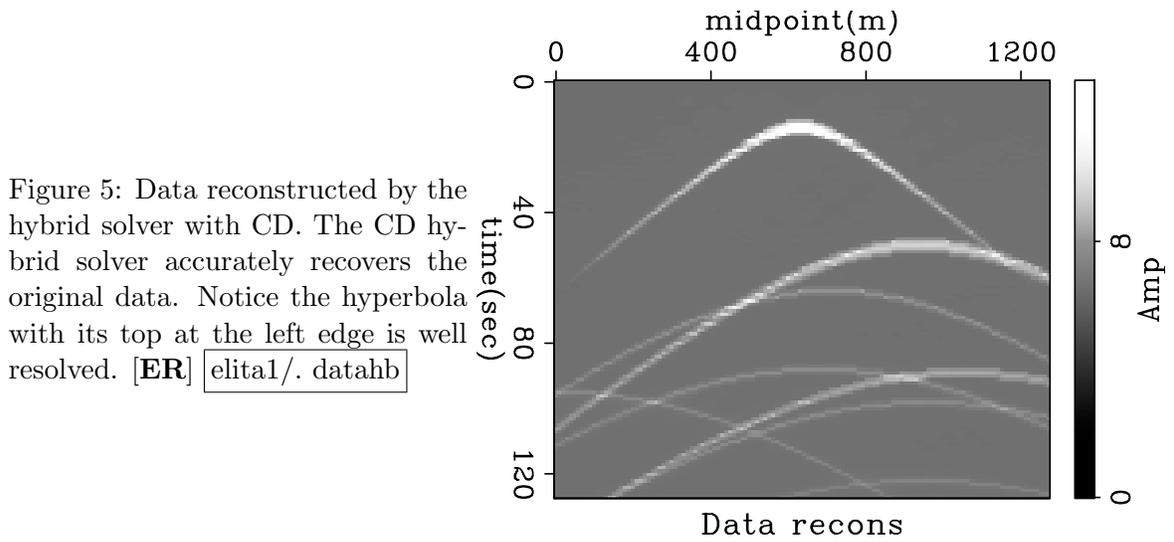


Figure 6: Input shot profile with noise. The *red* points in the figure indicates the data points with very high amplitude. [ER] `elita1/. noizin`

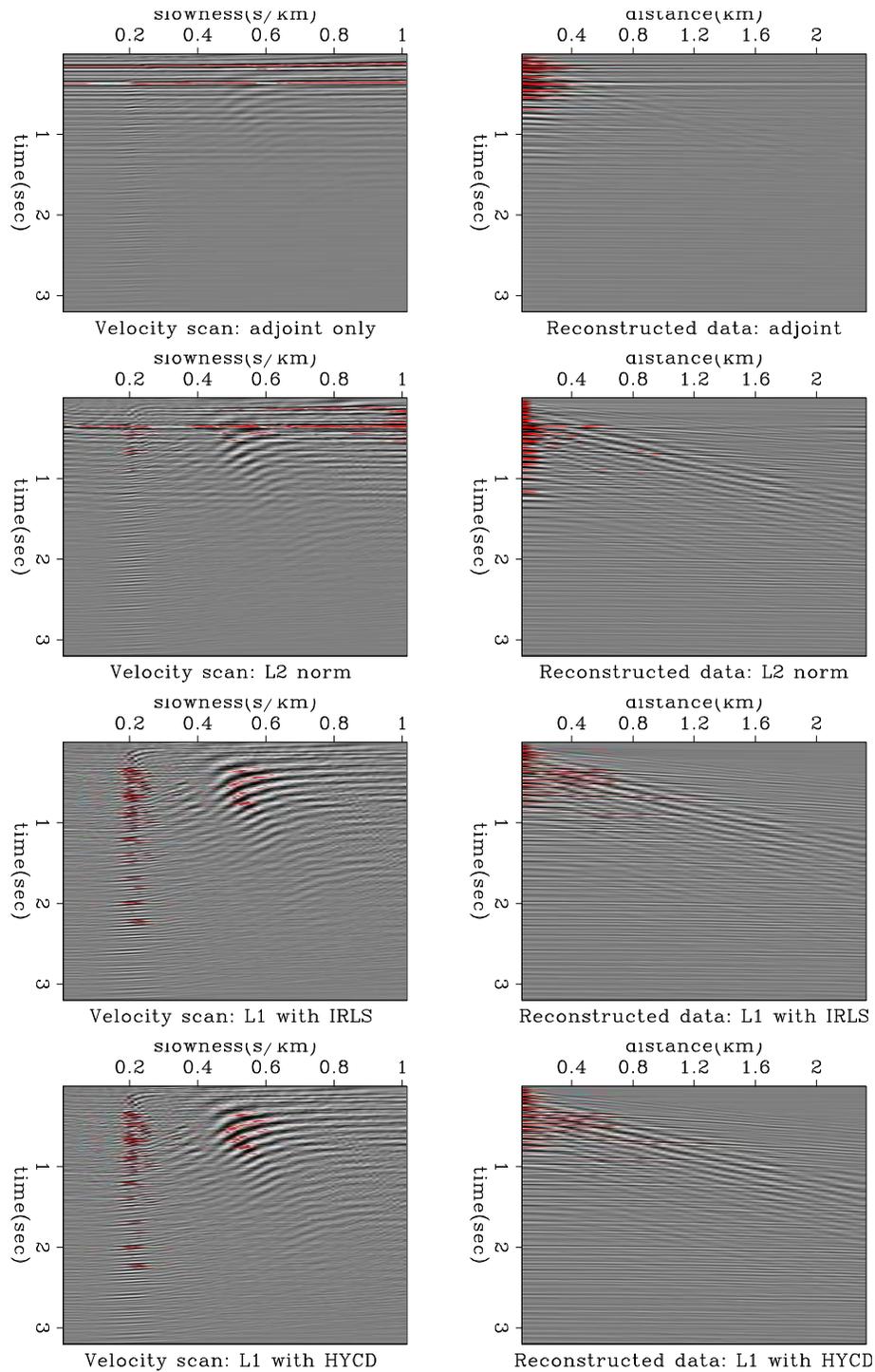


Figure 7: Inversion results of different methods. The *red* points in the figure indicates the data points with very high amplitude. Panels in the left column are the results of velocity scans, while panels on the right are the corresponding reconstructed data. First row: No inversion is applied (adjoint only); Second row: Inversion results by CD with L2; Third row: Inversion results by hybrid with IRLS; Bottom row: Inversion results by CD with L1.

[ER] elital/. real

Seismic reservoir monitoring with simultaneous sources

Gboyega Ayeni, Yaxun Tang, and Biondo Biondi

ABSTRACT

Hydrocarbon reservoirs can be efficiently monitored with simultaneous-source seismic data sets. Because simultaneous-source acquisition reduces time and cost requirements, seismic data sets can be recorded cheaply at short regular intervals, thereby allowing for near real-time monitoring. Although, in many cases, the recorded multiplexed data can be separated into independent records, we choose to leverage the efficiency of direct imaging of such data sets. However, direct imaging with a migration algorithm introduces cross-talk artifacts and does not account for differences in acquisition geometry and relative shot-timing between surveys. To attenuate cross-talk artifacts and acquisition discrepancies between data sets, we propose a joint least-squares migration/inversion method. By incorporating spatio-temporal and sparseness constraints in our inversion algorithm, we ensure that the resulting time-lapse images are geologically plausible. Using a 2D numerical model, we show that our method can give results of comparable quality to migrated single-source data sets.

INTRODUCTION

Conventional seismic data acquisition involves a single seismic source and a recording array of receivers. Although not a new idea (Womack et al., 1990), recent advances in acquisition technology enables seismic acquisition with multiple sources (Hampson et al., 2008; Beasley, 2008). This acquisition approach, also called simultaneous-shooting (or multi-shooting, or blended acquisition), can be used to achieve longer offsets, better shot-sampling, and improved time and cost efficiency (van Mastriigt et al., 2002; Berkhout et al., 2008; Howe et al., 2009). The recorded data can be separated into independent shot records and then imaged with conventional methods (Hampson et al., 2008; Spitz et al., 2008), or they can be imaged directly (Berkhout et al., 2008; Tang and Biondi, 2009).

Although time-lapse (4D) seismic is an established technology for monitoring hydrocarbon reservoirs (Rickett and Lumley, 2001; Whitcombe et al., 2004; Zou et al., 2006; Ebaid et al., 2009), it still has several limitations. First, because of the high cost of conventional (single-source) acquisition, it is impractical to acquire seismic data sets at short time intervals. Therefore, typical monitoring survey intervals may be too large to measure production-related, short-period variations in reservoir properties. Because of the large time intervals between seismic surveys, it may be difficult to match time-lapse seismic signatures to reservoir property changes derived from well-sampled sources (e.g. production history matching). Secondly, in many time-lapse seismic applications, inaccuracies in the replication of acquisition geometries for different surveys (*geometry non-repeatability*) are a recurring problem. Although modern acquisition techniques can improve repeatability of shot-receiver geometries, field conditions usually prevent perfect repetition. In order to isolate differences caused by changes in reservoir properties, non-repeatability effects must

be removed from time-lapse data sets. Furthermore, because of operational, climatic, and other limitations, the acquisition time-window may be too small for conventional seismic data acquisition. In such cases, it would be difficult to acquire conventional seismic data sets at desirable intervals.

Some limitations in current and conceptual time-lapse seismic applications can be overcome by simultaneous-shooting. First, by acquiring time-lapse data sets with multiple seismic sources, we can limit acquisition time and cost, and therefore acquire more data sets at shorter time intervals. Sufficiently small survey intervals will enable *quasi-continuous* monitoring of changes in reservoir properties. Other methods for quasi-continuous monitoring have been suggested (Arogunmati and Harris, 2009). Secondly, because we can account for differences in survey geometries during imaging, repetition of survey geometries is unnecessary. Furthermore, because of its high efficiency, simultaneous-shooting can be used for seasonal time-lapse seismic data acquisition in areas with short acquisition time-windows (Berkhout, 2008). Depending on operational limitations, an arbitrary number of seismic sources can be used for each survey. Figure 1 illustrates some scenarios where simultaneous-shooting concepts can be utilized.

There are two discrepancies in time-lapse seismic data sets recorded with multiple sources, namely, geometry and shot-timing non-repeatabilities (Ayeni et al., 2009). As mentioned above, geometry non-repeatability is a result of differences in acquisition geometries for different surveys. Shot-timing non-repeatability between different surveys is a result of mismatches in their relative shooting times. Neglecting survey geometry and shot-timing repeatability during acquisition ensures time and cost efficiency. However, if not accounted for, these two discrepancies will degrade the resulting time-lapse seismic images. Because conventional imaging and time-lapse processing methods are inadequate to account for such discrepancies, we propose a joint (global) least-squares imaging approach.

Least-squares migration/inversion can improve structural and amplitude information in seismic images (Nemeth et al., 1999; Kühl and Sacchi, 2003; Plessix and Mulder, 2004). Direct imaging of simultaneous-source data sets using least-squares migration/inversion methods has been discussed by previous authors (Ayeni et al., 2009; Dai and Schuster, 2009; Tang and Biondi, 2009). In this paper, we formulate time-lapse imaging of simultaneous-source data sets as a regularized joint least-squares problem. By avoiding separation of the recorded data into independent records, we reduce the data volume and processing cost. For each survey, we model the acquisition experiment with a phase encoding operator and the recorded shot-receiver geometries and relative shot-timings. We assume that the velocity and structural dips are known and that they change linearly between surveys. In addition, we assume that for each survey, the shot-receiver positions and relative shot timings are known. Finally, by including structural and temporal constraints in the inversion, we obtain geologically plausible time-lapse seismic images.

First, we consider the phase-encoding representation of simultaneous-shooting. We then introduce a joint inversion framework for simultaneous-source time-lapse data sets. Finally, using fifteen data sets from a 2D numerical model, we show that our method can give high quality images of reservoir property changes.

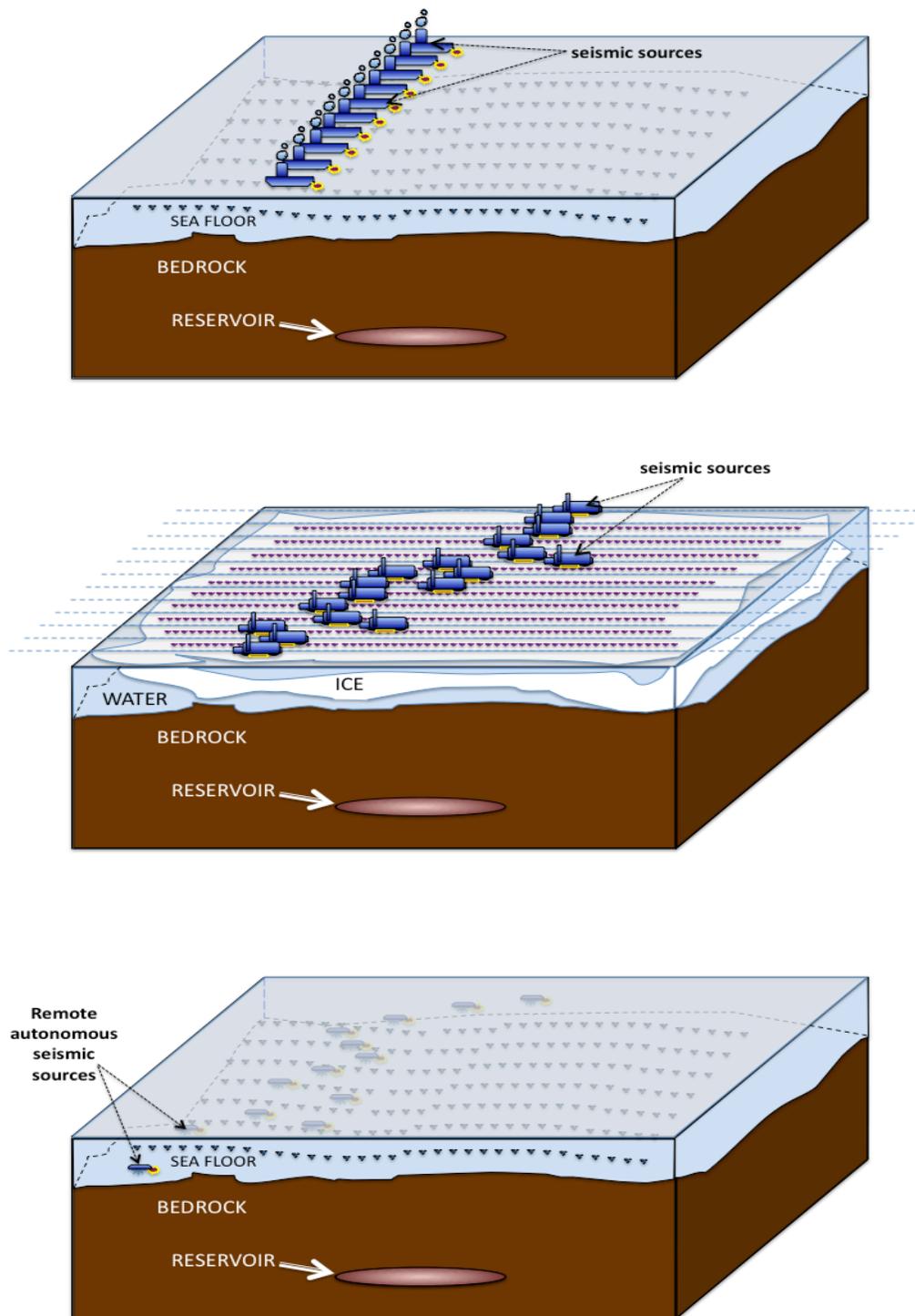


Figure 1: Some conceptual applications of simultaneous-shooting for ocean bottom cable/-seismometer acquisition (top), Frontier (e.g. Arctic) data acquisition (middle), and remote autonomous data acquisition (bottom). [NR]. [gayeni2/. examples](#)

LINEAR PHASE-ENCODED MODELING AND SIMULTANEOUS-SHOOTING

From the linearized Born approximation of the acoustic wave equation, the seismic data d recorded by a receiver at \mathbf{x}_r due to a shot at \mathbf{x}_s is given by

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

where ω is frequency, $m(\mathbf{x})$ is the *reflectivity* at image points \mathbf{x} , $f_s(\omega)$ is the source wavelet, and $G(\mathbf{x}_s, \mathbf{x}, \omega)$ and $G(\mathbf{x}, \mathbf{x}_r, \omega)$ are the Green's functions from \mathbf{x}_s to \mathbf{x} and from \mathbf{x} to \mathbf{x}_r , respectively. When there are multiple seismic sources, the recorded seismic data is due to a concatenation of phase-shifted sources. For example, the recorded data due to shots starting from $\mathbf{s} = \mathbf{q}$ to $\mathbf{s} = \mathbf{p}$, is given by

$$d(\mathbf{x}_{\mathbf{s}_{\mathbf{p}\mathbf{q}}}, \mathbf{x}_r, \omega) = \sum_{\mathbf{s}=\mathbf{p}}^{\mathbf{q}} a(\gamma_{\mathbf{s}}) \omega^2 \sum_{\mathbf{x}} f_{\mathbf{s}}(\omega) G(\mathbf{x}_{\mathbf{s}}, \mathbf{x}, \omega) G(\mathbf{x}, \mathbf{x}_r, \omega) m(\mathbf{x}), \quad (2)$$

where $a(\gamma_{\mathbf{s}})$ is given by

$$a(\gamma_{\mathbf{s}}) = e^{i\gamma_{\mathbf{s}}} = e^{i\omega t_{\mathbf{s}}}, \quad (3)$$

and $\gamma_{\mathbf{s}}$, the time-delay function, depends on the delay time $t_{\mathbf{s}}$ at shot \mathbf{s} .

For acquisition efficiency, it is unnecessary to repeat either the acquisition geometry or the relative shot timings for different surveys. By eliminating the cost associated with repeatability between surveys, we can significantly reduce the total acquisition cost. Because acquisition cost is usually several times higher than the processing cost, a reduction in acquisition cost will significantly reduce the total seismic monitoring cost. In addition, we achieve further cost reduction by imaging all the data sets directly. Figure 2 shows examples of wavefields from two configurations of simultaneous-shooting. In both figures, the third dimension represents the survey time, while the orthogonal lines indicate positions of the displayed slices within the cube.

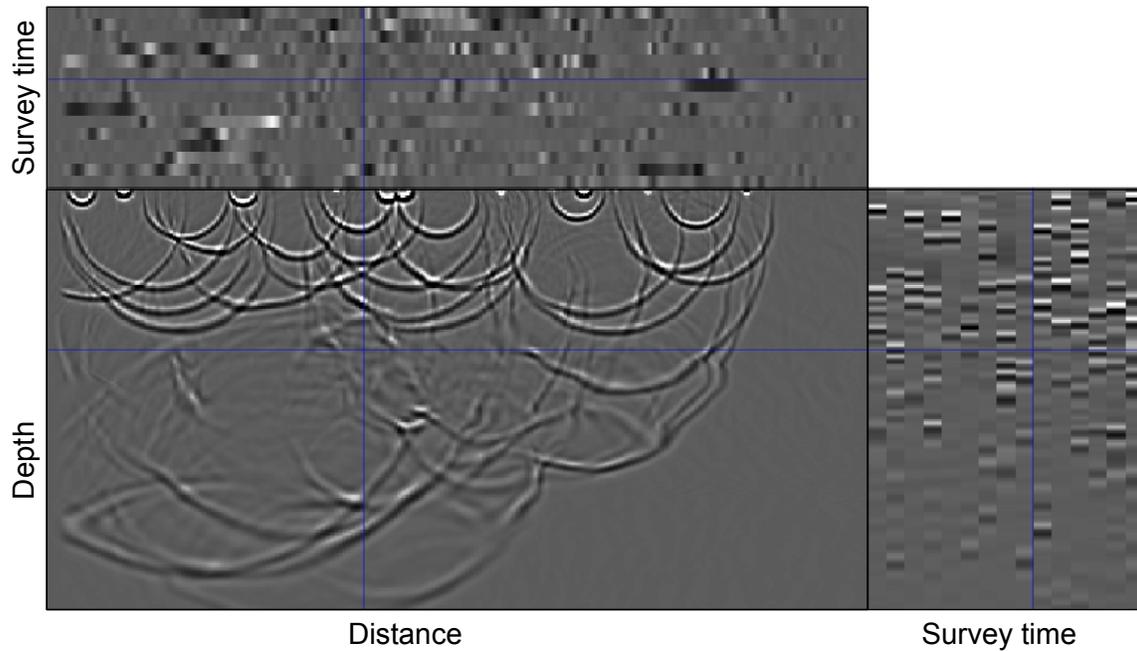
REGULARIZED JOINT INVERSION

For an arbitrary survey i , we can simplify the modeling equation into the form

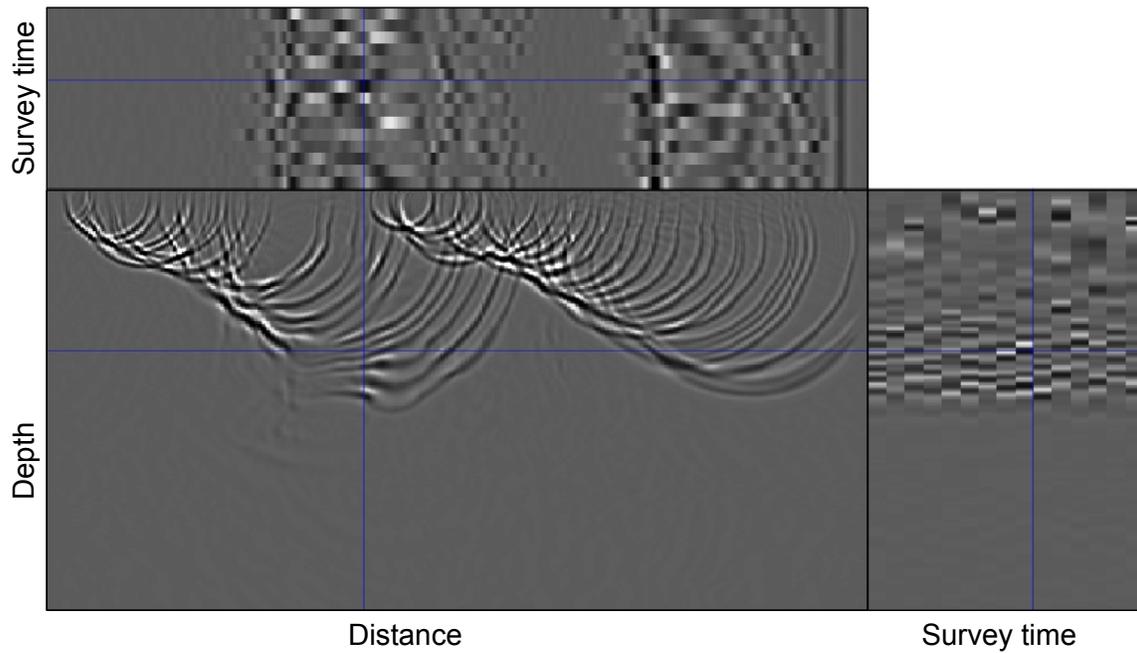
$$\tilde{\mathbf{d}}_i = \mathbf{B}_i \mathbf{L}_i \mathbf{m}_i = \tilde{\mathbf{L}}_i \mathbf{m}_i, \quad (4)$$

where $\tilde{\mathbf{d}}$ is the recorded data, \mathbf{B} is the encoding operator, \mathbf{L} is the modeling operator, \mathbf{m} is the earth reflectivity, and $\tilde{\mathbf{L}} = \mathbf{B}\mathbf{L}$. The migrated image, computed by applying the adjoint operator $\tilde{\mathbf{L}}^T$ to $\tilde{\mathbf{d}}$, will contain cross-term artifacts generated by cross-correlation between incongruous source and receiver wavefields (Romero et al., 2000; Tang and Biondi, 2009). In addition, because of the associated geometry and relative shot-time non-repeatability, different surveys have unique cross-term artifacts. To attenuate these artifacts, for N surveys, we minimize a joint (global) cost function S given by

$$S(\mathbf{m}_0, \dots, \mathbf{m}_N) = \sum_{i=0}^N \left\| \tilde{\mathbf{L}}_i \mathbf{m}_i - \tilde{\mathbf{d}}_i \right\|^2 + \sum_{i=0}^N \|\epsilon_i \mathbf{R}_i \mathbf{m}_i\|^2 + \sum_{i=1}^N \|\zeta_i \mathbf{\Lambda}_i(\mathbf{m}_{i-1}, \mathbf{m}_i)\|^2 + \sum_{i=1}^N \|\beta_i \mathbf{\Gamma}_i(\Delta m_i)\|_{hb}, \quad (5)$$



(a)



(b)

Figure 2: Wavefields from multiple randomized simultaneous sources (a), and from two continuously shooting seismic sources (b). In each figure, the blue line indicates intersecting positions of the the three slices that are displayed. In Panel (a), the geometry and relative shot-timing are different for all surveys, whereas in Panel (b), only the acquisition geometry differs between surveys. The third dimension denotes survey/recording time. [CR]. `gayeni2/. marm-wav1,marm-wav2`

where the parameters ϵ_i and ζ_i determine the strengths of the spatial and temporal regularization operators, \mathbf{R}_i and $\mathbf{\Lambda}_i$ respectively. Because only a small region in the model space contain desired in time-lapse signal, a sparseness requirement is desirable. Parameter β_i determines the strength of the sparseness operator $\mathbf{\Gamma}_i$. Related formulations have been applied to other time-lapse imaging problems (Ajo-Franklin et al., 2005). We compute the time-lapse image as the difference between the migrated or inverted image at time t and that at time 0. Because several shots are encoded and directly imaged, the computational cost of this approach is considerably reduced compared to non-encoded data sets.

In this paper, the spatial regularization operator is a system of non-stationary dip-filters, whereas the temporal regularization operator is a gradient between surveys. We compute dips using the plane-wave destruction method (Fomel, 2002), and we compute dip-filters using factorized directional Laplacians (Hale, 2007). To ensure stable transitions at sharp boundaries, the filter corresponding to any image point is scaled according to a dip-contrast-dependent variance. We estimate the spatial and temporal regularization parameters by scaling the maximum amplitude in each data set. Finally, we minimize the objective function using an iterative hybrid conjugate direction algorithm (Li et al., 2010) which enforces desired sparseness on the time-lapse images.

NUMERICAL EXAMPLE

The proposed method was applied to a modified 2D Marmousi model (Bourgeois et al., 1991). For simplicity, we neglect overburden geomechanical changes and assume no change in reflectivity, except within the reservoir (Figure 3(a)). Using a Born modeling algorithm, we simulated 15 data sets representing different production stages (Figure 4). Each data set comprises 56 randomly encoded shot records with unique shot positions and unique relative shot-timings (Figure 5). We estimated the dip-field and dip-contrast (Figures 3b and c) from the migrated baseline image. For data modeling and migration, we use a phase-encoding one-way wave-equation operator. For comparison, using the same number of shots and receivers and perfect repeatability, we modeled and migrated 15 conventional data sets. The migrated and inverted images, together with the corresponding time-lapse images, are shown in Figures 6 to 8.

DISCUSSION

If the temporal spacing between seismic surveys is small, we see that a near-continuous image of reservoir property change can be obtained (Figure 6). We can reduce the acquisition cost for these conventional seismic surveys by using multiple seismic sources. Instead of separating the recorded data from such an experiment, they can be imaged directly with a phase-encoding operator. However, direct imaging causes cross-talk artifacts that degrade the quality of migrated images (Figure 7(a)). In addition, if the the acquisition geometries and relative shot-timings are not repeated, the cross-term artifacts will degrade the quality of the time-lapse images (Figure 7(b)). Regularized joint inversion attenuates these artifacts (Figure 8(a)). Furthermore, inversion also produces high-quality time-lapse images (Figure 8(b)) that are of comparable quality but better resolution than perfectly repeated single-source data sets (Figure 6(b)). A careful choice of the regularization parameters ensures that the objective function is well behaved for all components of the global cost

function. This leads to a gradual reduction in the cross-term and non-repeatability artifacts with iteration (Figures 9 and 10).

CONCLUSIONS

We have proposed an efficient scheme for acquiring and processing time-lapse seismic data sets. This method can reduce the overall data acquisition and processing cost for seismic reservoir monitoring. We have shown that even if the survey geometries and relative shot timing are not repeated, our joint inversion method gives high-quality time-lapse images. These acquisition and processing approaches provide a realistic framework for efficient seismic reservoir monitoring in many scenarios. It can make several conceptual seismic monitoring technologies (e.g. autonomous seismic acquisition, Arctic seismic reservoir monitoring) possible. In the near future, we will incorporate a scheme to compensate for geomechanical reservoir changes.

REFERENCES

- Ajo-Franklin, J. B., J. Urban, and J. M. Harris, 2005, Temporal integration of seismic traveltime tomography: SEG Technical Program Expanded Abstracts, **24**, 2468–2471.
- Arogunmati, A. and J. M. Harris, 2009, An approach for quasi-continuous time-lapse seismic monitoring with sparse data: SEG Technical Program Expanded Abstracts, **28**, 3899–3903.
- Ayeni, G., Y. Tang, and B. Biondi, 2009, Joint preconditioned least-squares inversion of simultaneous source time-lapse seismic data sets: SEG Technical Program Expanded Abstracts, **28**, 3914–3918.
- Beasley, C. J., 2008, Simultaneous sources: A technology whose time has come: SEG Technical Program Expanded Abstracts, **27**, 2796–2800.
- Berkhout, A. J. G., 2008, Changing the mindset in seismic data acquisition: The Leading Edge, **27**, 924–938.
- Berkhout, A. J. G., G. Blacquièrre, and E. Verschuur, 2008, From simultaneous shooting to blended acquisition: SEG Technical Program Expanded Abstracts, **27**, 2831–2838.
- Bourgeois, A., M. Bourget, P. Lailly, M. Poulet, P. Ricarte, and R. Versteeg, 1991, Marmousi, model and data:: Proceedings of 1990 EAEG workshop on practical aspects of seismic data inversion.
- Dai, W. and J. Schuster, 2009, Least-squares migration of simultaneous sources data with a deblurring filter: SEG Technical Program Expanded Abstracts, **28**, 2990–2994.
- Ebaid, H., M. Nasser, P. Hatchell, and D. Stanley, 2009, Time-lapse seismic makes a significant business impact at Holstein: SEG, Expanded Abstracts, **28**, 3810–3814.
- Fomel, S., 2002, Applications of plane-wave destruction filters: Geophysics, **67**, 1946–1960.
- Hale, D., 2007, Local dip filtering with directional laplacians: CWP Project Reiew, **567**, 91–102.
- Hampson, G., J. Stefani, and F. Herkenhoff, 2008, Acquisition using simultaneous sources: SEG Technical Program Expanded Abstracts, **27**, 2816–2820.
- Howe, D., M. Foster, T. Allen, I. Jack, D. Buddery, A. Choi, R. Abma, T. Manning, and M. Pfister, 2009, Independent simultaneous sweeping in Libya-full scale implementation and new developments: SEG Technical Program Expanded Abstracts, **28**, 109–111.

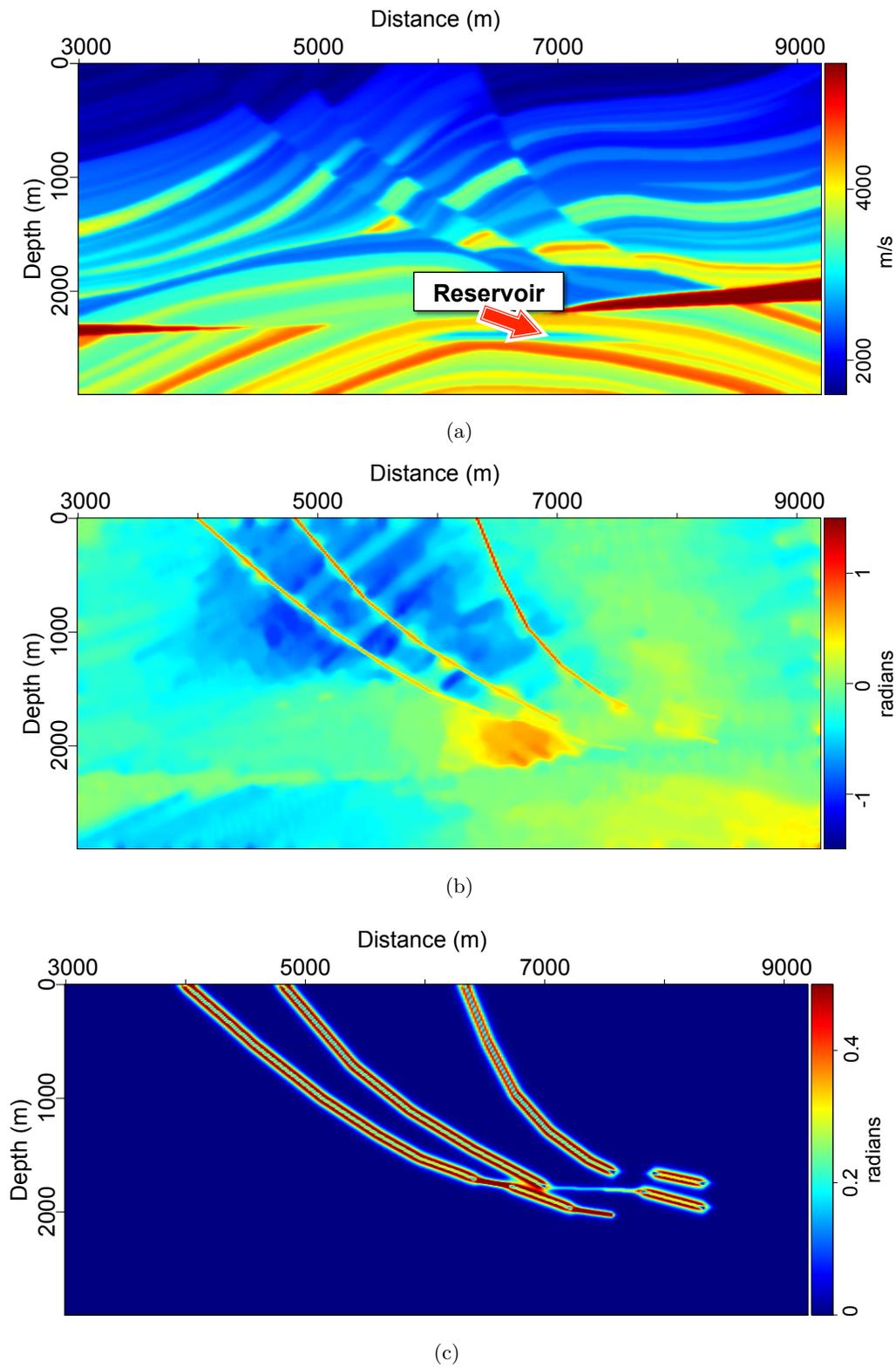


Figure 3: Baseline velocity model (a), dip-field computed from the migrated baseline image (b), and dip-variance estimated as a function of dip contrast (c). [CR].

gayeni2/. vel-0,dip-0,var-0

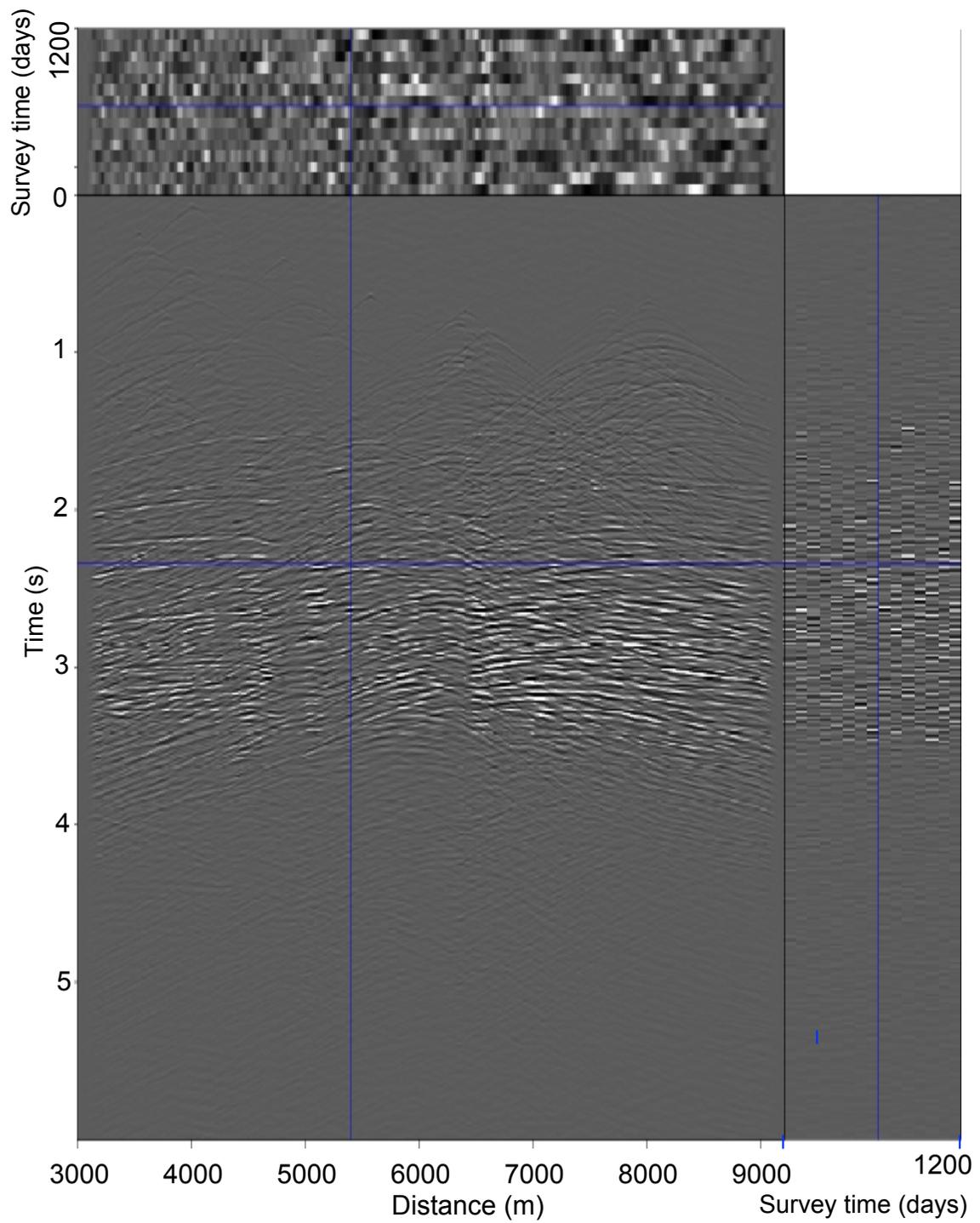


Figure 4: Synthetic data from multiple asynchronous sources. The third dimension denotes survey/recording time. [CR]. `gayeni2/. data-0`

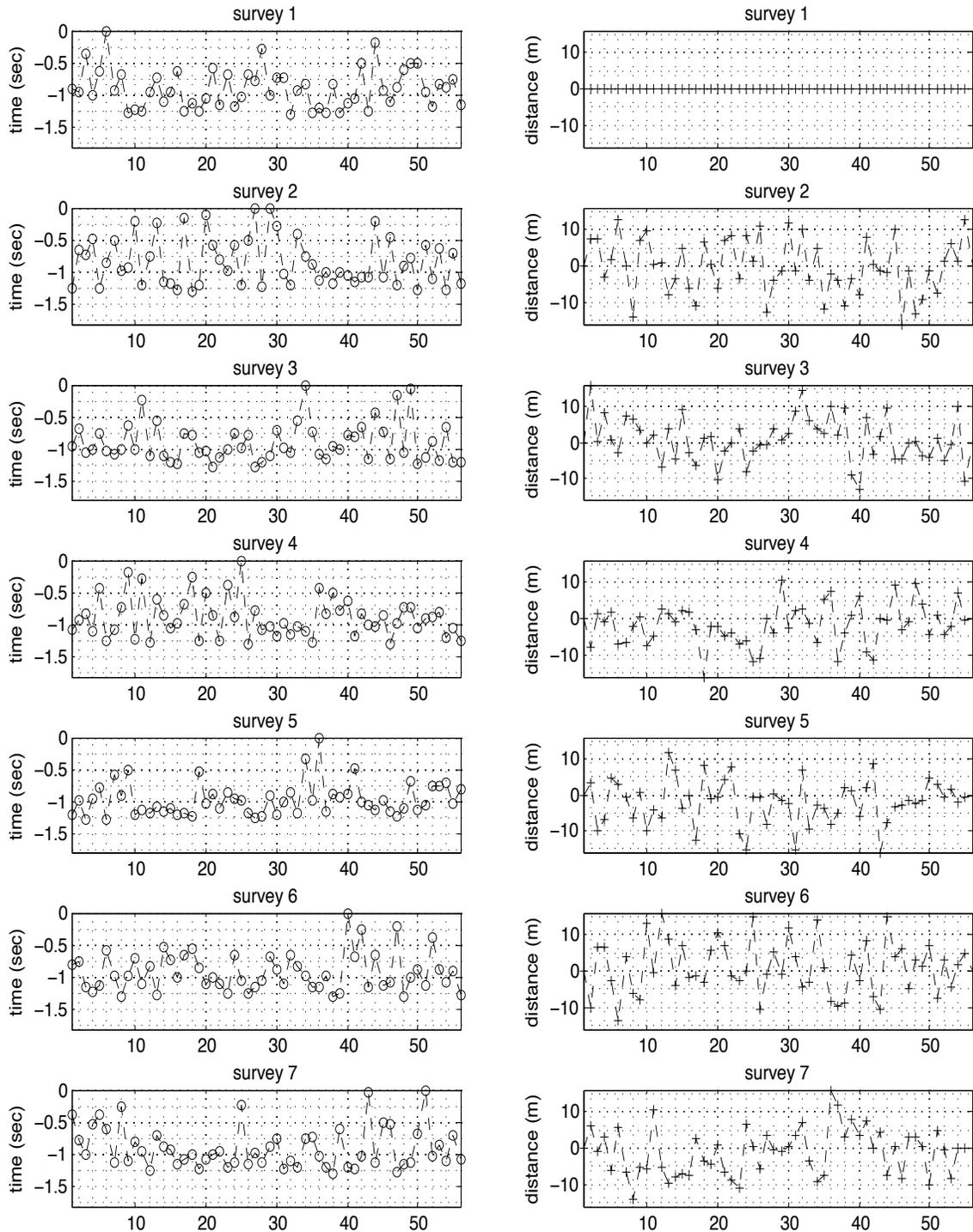
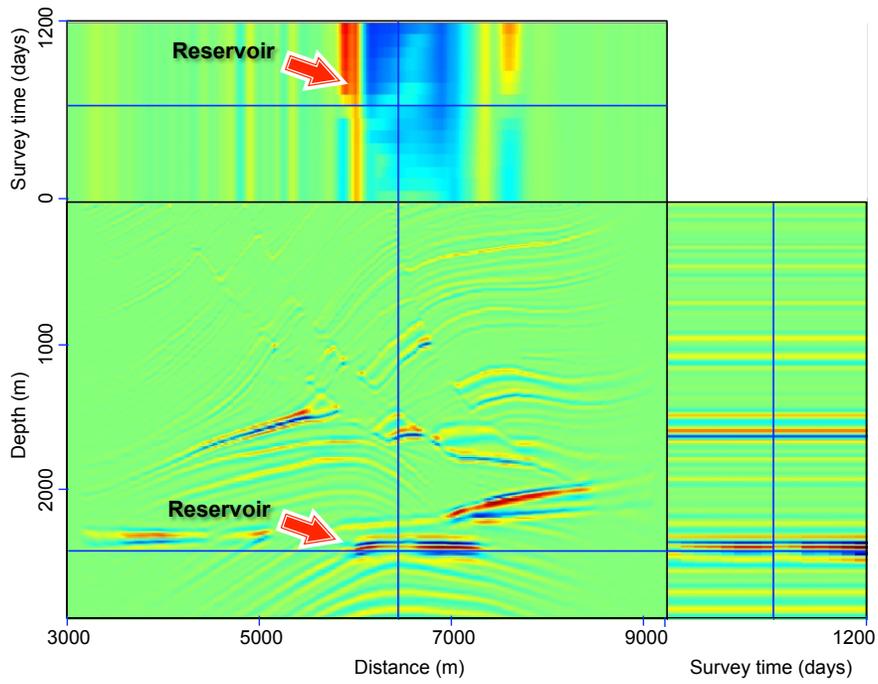
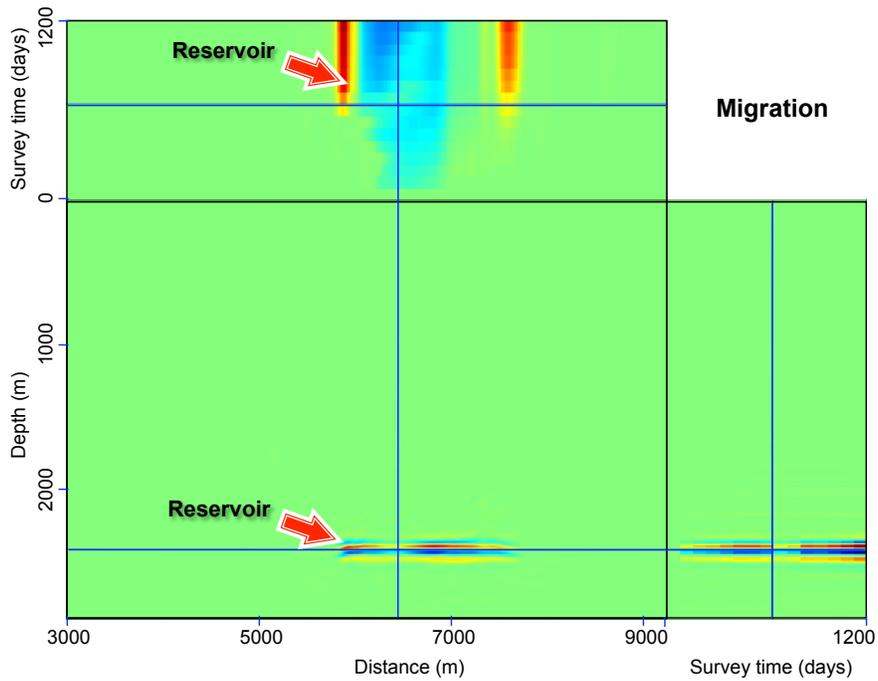


Figure 5: Plots of relative time-delays (left) and shot-displacements for seven out of the fifteen numerical models that were used to generate the data in Figure 4. In all plots, the horizontal axis indicates shot position. The relative shooting times are referenced to the earliest shot in each survey, whereas shot-displacements are referenced to the baseline shot positions. [NR]. [gayeni2/. source-time](#)

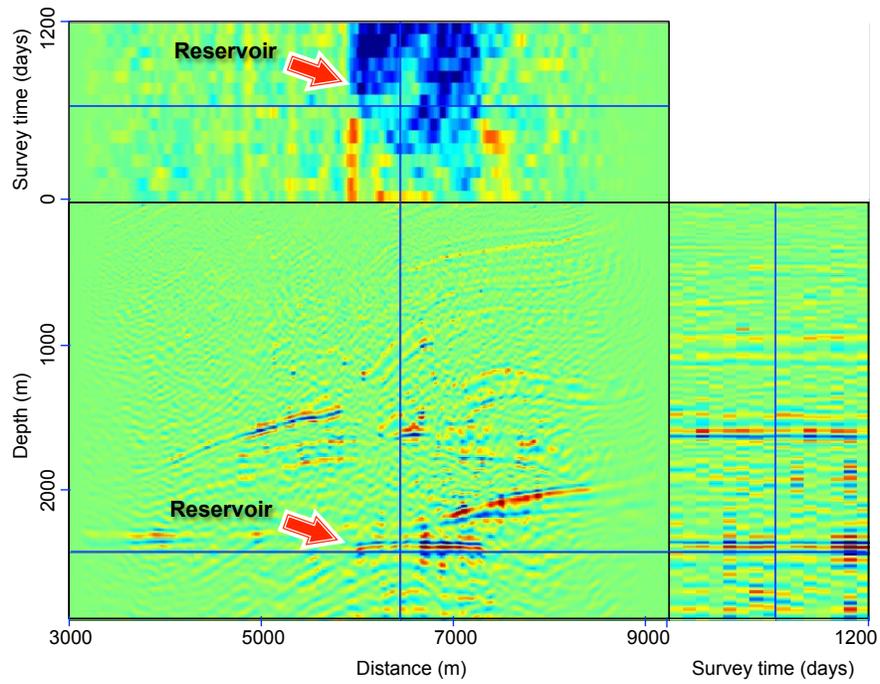


(a)

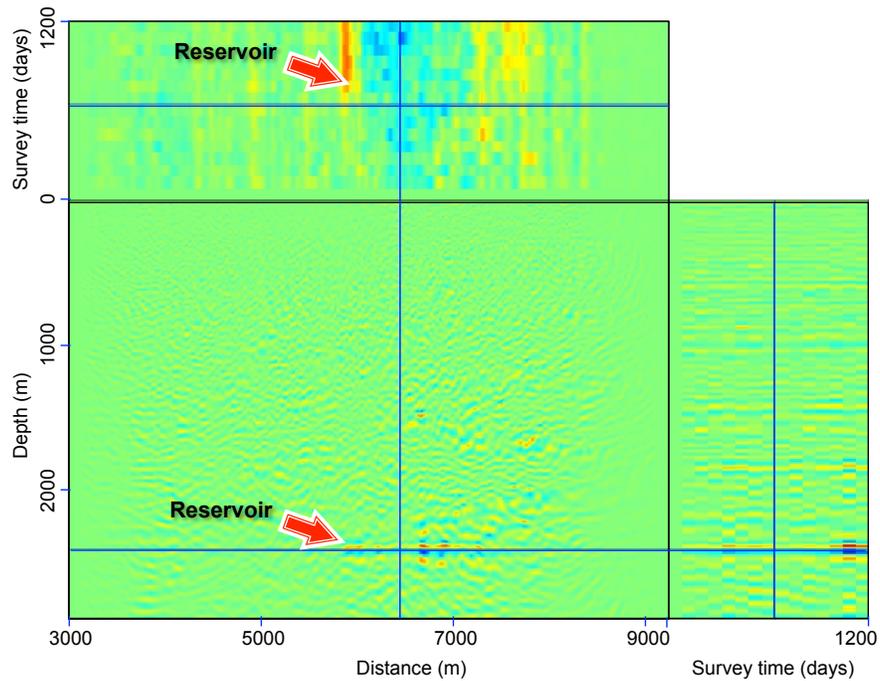


(b)

Figure 6: Images (a) and corresponding time-lapse estimates (b) obtained from migrating perfectly repeated conventional (single-source) data sets. In this (and in similar) Figures, the side panel (third axis) shows the seismic properties (a) and time-lapse changes (b) at a fixed spatial position, whereas the top panel shows the spatial-temporal distribution seismic properties. [CR]. `gayeni2/. nomig,4d-nomig`



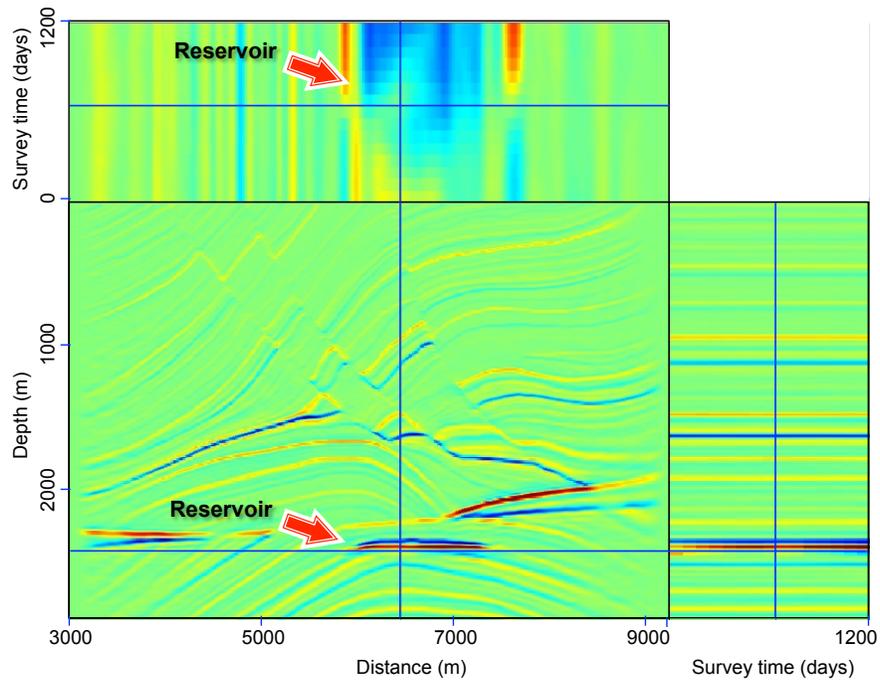
(a)



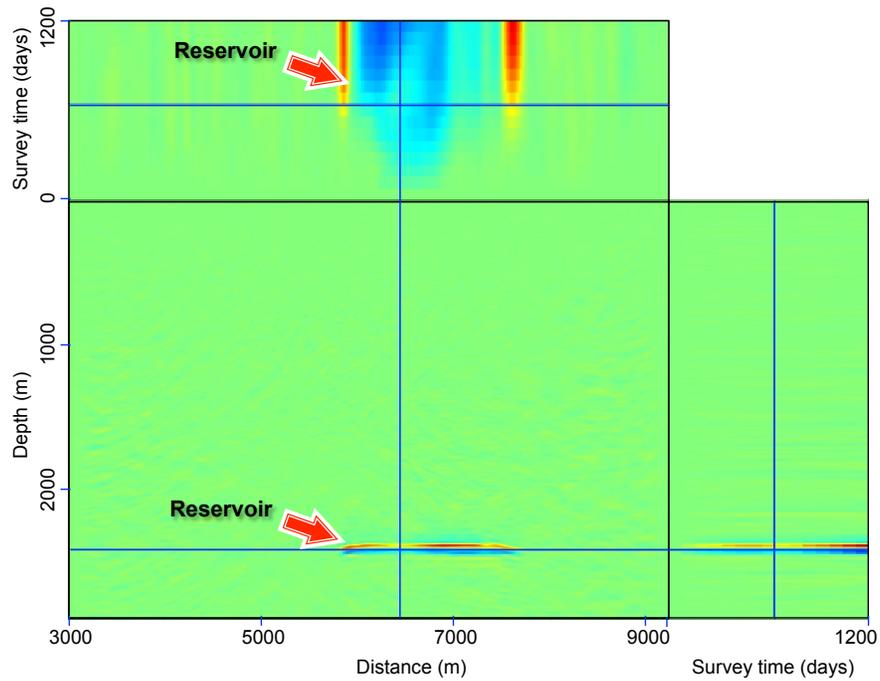
(b)

Figure 7: Images (a) and corresponding time-lapse estimates (b) obtained from migrating the data sets in Figure 4. In both Figures, note the numerous artifacts caused by geometry and shot-timing non-repeatability and cross-term artifacts. Without attenuating these artifacts, it would be difficult to accurately interpret the time-lapse information. [CR].

gayeni2/. mig,4d-mig

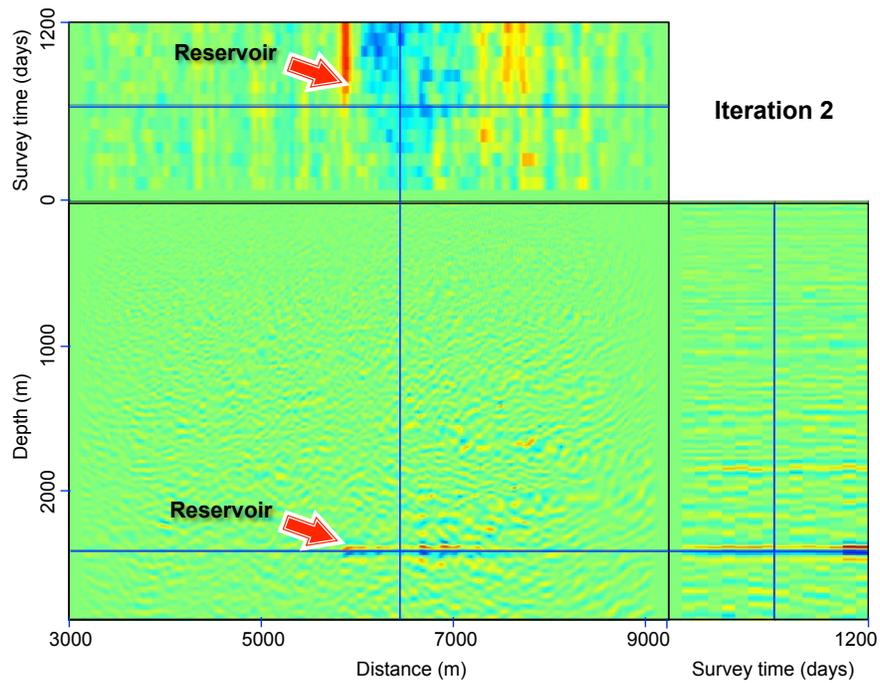


(a)

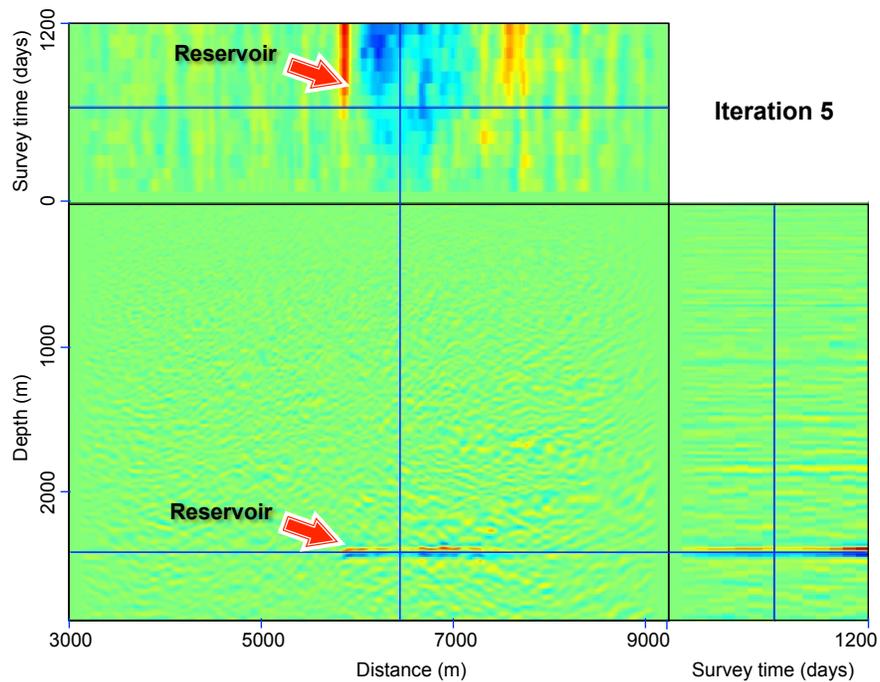


(b)

Figure 8: Images (a) and corresponding time-lapse estimates (b) obtained from inverting the simultaneous-source data sets in Figure 4. Note that the non-repeatability and cross-talk artifacts in the migrated images (Figure 7) have been attenuated by inversion. Also, note the better resolution of the inverted images compared to the migrated single-source data (Figure 6). `gayeni2/. inv,4d-inv`

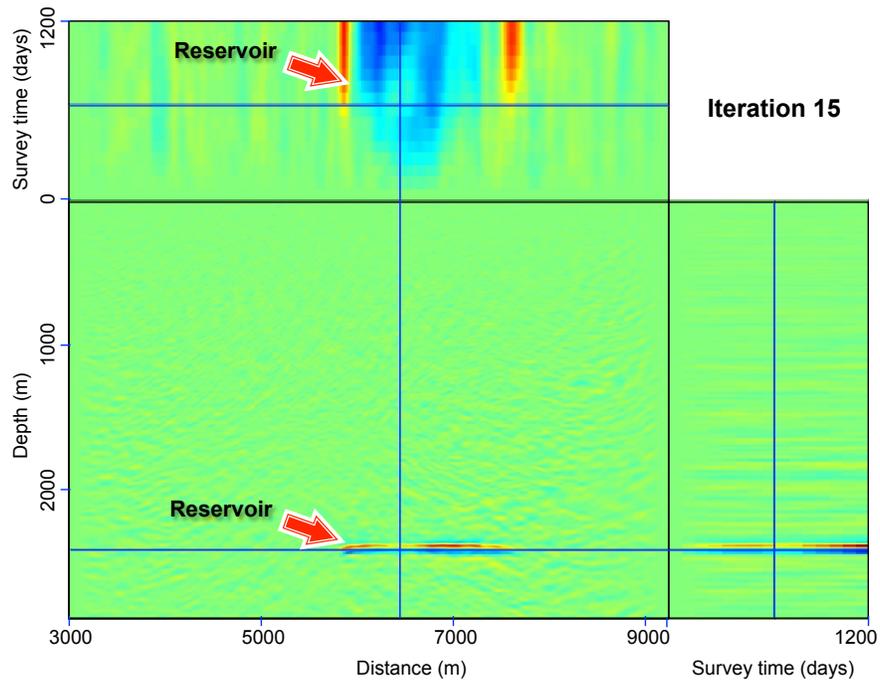


(a)

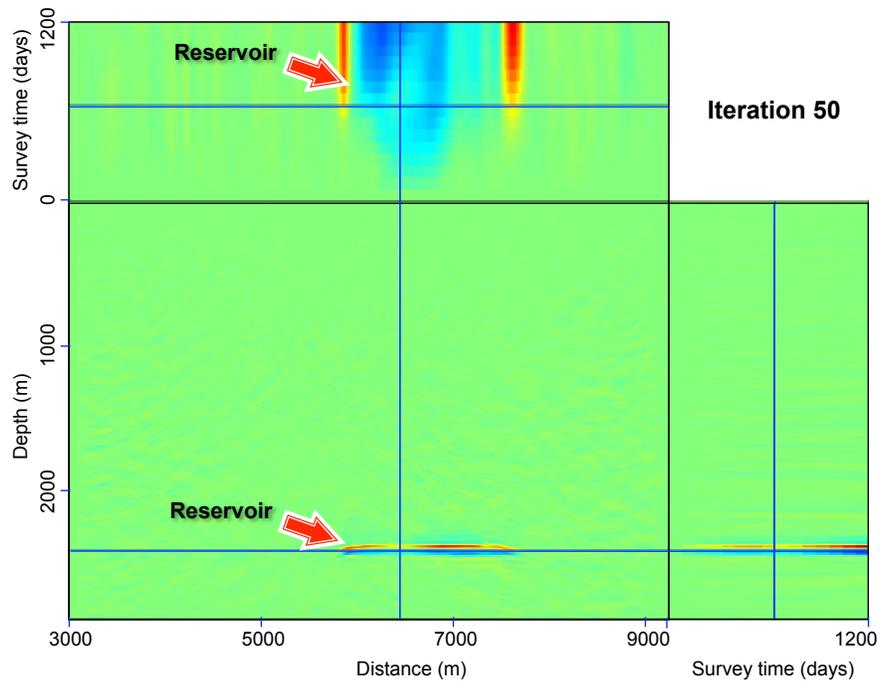


(b)

Figure 9: Time-lapse seismic images obtained after 2 and 5 conjugate gradient iterations (a) and (b) respectively. Note the gradual reduction in the artifacts compared to the time-lapse images from migration (Figure 7(b)). [CR]. `gayeni2/. 4d-inv-1,4d-inv-2`



(a)



(b)

Figure 10: Time-lapse seismic images obtained after 15 and 50 conjugate gradient iterations (a) and (b) respectively. Note the reduction in the artifacts compared to the time-lapse images from migration (Figure 7(b)). [CR]. [gayeni2/. 4d-inv-3,4d-inv-4](#)

- Kühl, H. and M. D. Sacchi, 2003, Least-squares wave-equation migration for avp/ava inversion: *Geophysics*, **68**, 262–273.
- Li, Y., Y. Zhang, and J. Claerbout, 2010, Geophysical applications of a novel and robust l1 solver: SEP Report, **140**.
- Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: *Geophysics*, **64**, 208–221.
- Plessix, R.-E. and W. Mulder, 2004, Frequency-domain finite-frequency amplitude-preserving migration: *Geophysical Journal International*, **157**, 975–985.
- Rickett, J. E. and D. E. Lumley, 2001, Cross-equalization data processing for time-lapse seismic reservoir monitoring: A case study from the Gulf of Mexico: *Geophysics*, **66**, 1015–1025.
- Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: *Geophysics*, **65**, 426–436.
- Spitz, S., G. Hampson, and A. Pica, 2008, Simultaneous source separation: A prediction-subtraction approach: SEG Technical Program Expanded Abstracts, **27**, 2811–2815.
- Tang, Y. and B. Biondi, 2009, Least-squares migration/inversion of blended data: SEG Technical Program Expanded Abstracts, **28**, 2859–2863.
- van Mastrigt, P., S. Vaage, M. Dunn, and B. Pramik, 2002, Improvements in 3-D marine acquisition using continuous long offset (CLO): *The Leading Edge*, **21**, 394–.
- Whitcombe, D. N., J. M. Marsh, P. J. Clifford, M. Dyce, C. J. S. McKenzie, S. Campbell, A. J. Hill, R. S. Parr, C. Pearse, T. A. Ricketts, C. P. Slater, and O. L. Barkved, 2004, The systematic application of 4D in BP's North-West Europe operations — 5 years on: SEG Technical Program Expanded Abstracts, **23**, 2251–2254.
- Womack, J. E., J. R. Cruz, H. K. Rigdon, and G. M. Hoover, 1990, Encoding techniques for multiple source point seismic data acquisition: *Geophysics*, **55**, 1389–1396.
- Zou, Y., L. R. Bentley, L. R. Lines, and D. Coombe, 2006, Integration of seismic methods with reservoir simulation, Pikes Peak heavy-oil field, Saskatchewan: *The Leading Edge*, **25**, 764–781.

Low frequency passive seismic interferometry for land data

Sjoerd de Ridder

ABSTRACT

Here we report results achieved by low-frequency seismic interferometry on a passive seismic land dataset recorded at a field in Saudi Arabia. Computed spectra for different portions of the data show a time varying ambient seismic wavefield displaying a diurnal pattern. At low frequencies (< 10 Hz) the ambient seismic wavefield mainly consists of surface waves in two modes, the fundamental mode propagates with a velocity of about 1250 ms. Results suggest that sufficient coherent energy is recorded between 1 Hz and 7 Hz for retrieval of a Rayleigh surface wave. The strength of the ambient seismic field affects the convergence rate of the correlations. The directionality in the ambient seismic field affects the radiation pattern of the virtual sources. Retrieved Rayleigh waves at low frequencies show spatial variation and dispersive behavior. Dispersion curve estimation opens opportunities for reservoir monitoring by background velocity estimation.

INTRODUCTION

Seismic interferometry aims to retrieve the Green's function between two receiver stations by correlating measurements of seismic responses at both stations, effectively turning one station into a virtual source (Claerbout, 1968; Wapenaar, 2004). Recently, a variety of applications have been developed for active seismic data, such as the virtual source method (Bakulin and Calvert, 2006), redatuming (Schuster et al., 2004; Schuster and Zhou, 2006), and imaging of multiples (Berkhout and Verschuur, 2006). Commonly, active seismic (controlled source) interferometry attempts to reconstruct the high-frequency (> 10 Hz) impulse response of the earth. Passive seismic interferometry has mainly focused on retrieving high-frequency virtual sources where acquiring data with real sources is undesirable. Thus far, results of passive seismic interferometry have been less than promising, partially due to directionality of the ambient seismic wavefield and poor sampling of the medium by passive seismic sources. Early attempts of Cole (1995), (Artman, 2006, 2007) yielded less than satisfactory results; more recently, Dragonov et al. (2007) retrieved high-frequency reflection events from pre-selected body-wave events. Meanwhile, global seismologists have been successful using seismic interferometry to retrieve Green's functions at much lower frequency. More recently, the ambient seismic field at lower frequencies (< 10) Hz has been shown to contain sufficient coherent and omnidirectional seismic energy to yield low-frequency Green's functions containing direct arrivals in a marine environment. Dellinger (2008) pointed out that conventional arrays record useable energy at low frequencies, while Dellinger and Yu (2009), Landes et al. (2009) and Bussat and Kugler (2009) have reported on the use of low-frequency seismic energy recorded in marine environments to image shallow seafloor structures.

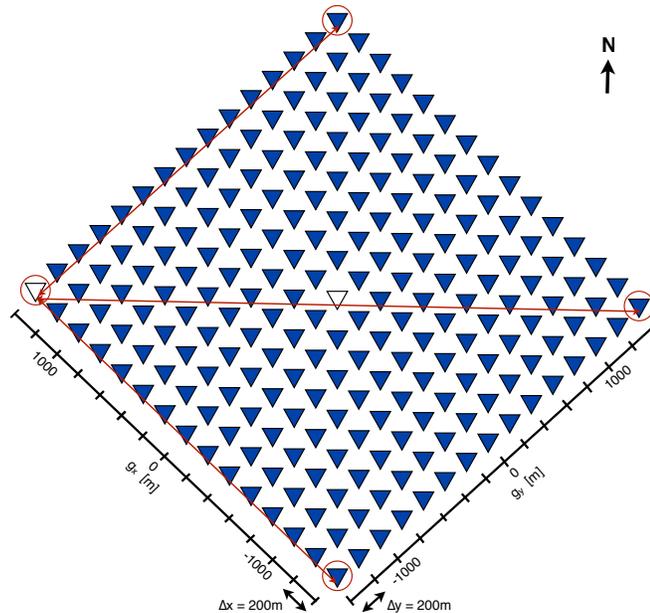
An ambient seismic field recorded above a field in Saudi Arabia has previously been

studied using novel imaging methods to locate microseismic events (Fu and Luo, 2009) and seismic interferometry to retrieve high frequency events (Xiao et al., 2009). Here we show interferometric retrieval of the Rayleigh surface waves at low frequencies. These surface waves are further studied for potential reservoir monitoring and imaging capabilities. Data availability limits the study to a conventional microtremor dispersion-curve analysis technique (Aki, 1957), which assumes laterally invariant media. However, more generally the retrieved Green's function can be studied and imaged without implicit lateral invariant assumptions.

AMBIENT SEISMIC WAVEFIELD

In 2007, Saudi Aramco initiated an experiment aiming to detect and characterize microseismic energy for reservoir monitoring (Jervis and Dasgupta, 2009). The surface array consisted of 225 buried 3-component stations placed in a 15 by 15 grid with station spacing of 200 m, spanning an area 3 km by 3 km.

Figure 1: Geometry of the Aramco passive experiment; 225 stations (denoted by triangles) placed in a 15 by 15 grid. The geographic North is indicated by N. Open triangles denote stations turned into virtual sources for Figures 7 and 8. Lines denote receiver lines for sections shown in Figure 7. [NR] [sjoerd1/. geometry](#)



Stanford University received a nearly continuous raw data record spanning 48 hours divided over 3 days, starting on day 1 at 18:00 ending on day 3 at 18:00. Only vertical components were used, after removing the arithmetic mean over 30 s time windows. The 48 hours of passive data were analyzed for their spectral characteristics. A frequency domain amplitude spectrogram averaged over the entire array is shown in Figure 2. Most energy was recorded between 2 Hz and 12 Hz and varies in a daily pattern with higher energy during the daylight hours and less energy at night. Figure 3 shows the frequency-domain amplitude spectrum, averaged for all recordings during the hour from 20:00 to 21:00 on day 1, drawn as curve (a). We identify a peak at very low frequencies, below 1 Hz.

To help determine what kind of seismic energy composes the low-frequency spectrum, the frequency-wavenumber ($f - k_x - k_y$) domain spectra were computed. A cross section through a cube of frequency-wavenumber spectra averaged over all recordings during the

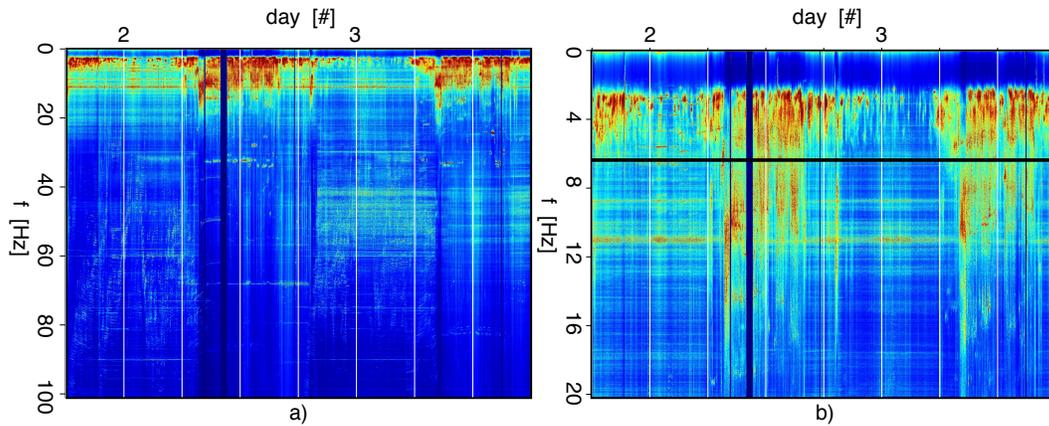


Figure 2: Frequency-domain amplitude spectrum averaged over the array as a function of time for 48 hours. White grid lines denote 6 hour blocks. a) Frequency-domain amplitude spectrum between 0 Hz and 100 Hz; b) frequency-domain amplitude spectrum between 0 Hz and 20 Hz; the black line denotes the frequency of the low-pass filter applied before interferometry. [CR] sjoerd1/. spectra1d-48-z

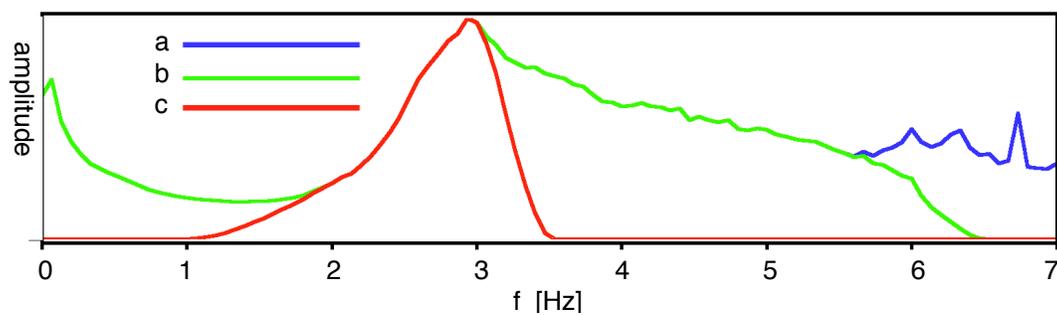


Figure 3: Normalized frequency-domain amplitude spectrum, averaged over the array and data recorded between 20:00 and 21:00 on day 1. Curve (a) is the spectrum of the original data record, curve (c) is the spectrum after low-pass filtering for interferometry and curve (b) is the spectrum of the data after band-pass filtering for beam-forming. [CR] sjoerd1/. spechbandpassed

hour from 20:00 to 21:00 on day 1 is shown in Figure 4. The panel in the upper right corner of Figure 4 denotes a beam-forming experiment (see below). A cross section through a cube of frequency-wavenumber spectra averaged over all recordings between 11:00 and 12:00 on day 2 is shown in Figure 5. Both cross-sections only show frequencies below 6 Hz; above 6 Hz no wave modes could clearly be identified. Up to 6 Hz most of the energy resides in the (Rayleigh) surface wave modes. The fundamental mode becomes aliased above 3 Hz. Between 20:00 and 21:00 on day 1 most energy comes from the west, while from 11:00 to 12:00 on day 2 most energy comes from the north. (The directionality of the tails in the frequency-wavenumber spectra is controlled by the sign of the Fourier transformations.) Another common technique to characterize directionality in a wavefield is beam-forming. This was performed, after bandpass filtering between 1 Hz and 3.5 Hz, by computing linear $\tau-p$ transformations over both directions. A beam is formed by averaging the amplitude in the $(\tau-p)$ domain over a certain τ -window. (Note τ denotes the interception times and p denotes the slownesses of the stacking lines of the $\tau-p$ transformation in (t, x) domain). The beams shown in the upper right corners of Figures 4 and 5 show that the fundamental mode travels with a slowness of slightly less than 1 ms/m (corresponding to a velocity of slightly greater than 1000 m/s). A higher mode visible in the frequency-wavenumber domain of Figure 4 can be observed (faintly) to travel with a slowness under 0.4 ms/m (corresponding with a velocity greater than 2500 m/s). Studying averaged frequency-wavenumber domains for other hours shows that the ambient seismic field at frequencies below 6 Hz is generally incident from the west and/or north.

SOURCES FROM INTERFEROMETRY

Since the coherent ambient seismic field (as recorded by the array) is dominated by surface waves, a low-pass filter of 6.5 Hz was applied to extract the low-frequency surface wave energy. The recordings at all stations are correlated with the recording at one master station, effectively creating a virtual source at that master station. This procedure is repeated using each station as a master station, creating a virtual seismic survey with sources at each station location. The theory of seismic interferometry indicates that the cross-correlation between two recordings should, under equipartitioned energy circumstances, retrieve a time-symmetric Green's function between the stations (Wapenaar, 2004). In Figure 6 the envelope of the Green's functions for a virtual source at $(s_x, s_y) = (0, 0)$ m is shown as a series of time slices, computed using all 48 hours of data. The Green's functions are not time-symmetric, which can be expected given the directionality of the ambient seismic field at frequencies below 6 Hz. However, the Green's functions are also not symmetric in space. The retrieved Rayleigh wave travels faster through the east than through the west part of the array. The capacity to retrieve longer offsets with less recording is studied by selecting a virtual source at the west end of the array and correlating all recordings between 20:00 and 21:00 on day 1. Figure 7 shows three slices through the correlation cube; a g_x -slice, a g_y -slice and a third diagonal slice (see Figure 1). Although the directionality in the ambient seismic field is favorable for retrieving along the diagonal of the correlation cube, correlating one hour of data was not sufficient to achieve convergence for a Green's function along the long offsets of the diagonal slice. For the smaller offsets along g_x and g_y -slices, one hour of data was sufficient to achieve reasonable convergence. The move-out of the arrivals indicate a velocity of approximately 1000 m/s.

The quality of the retrieved Green's functions depends on the portion of data used

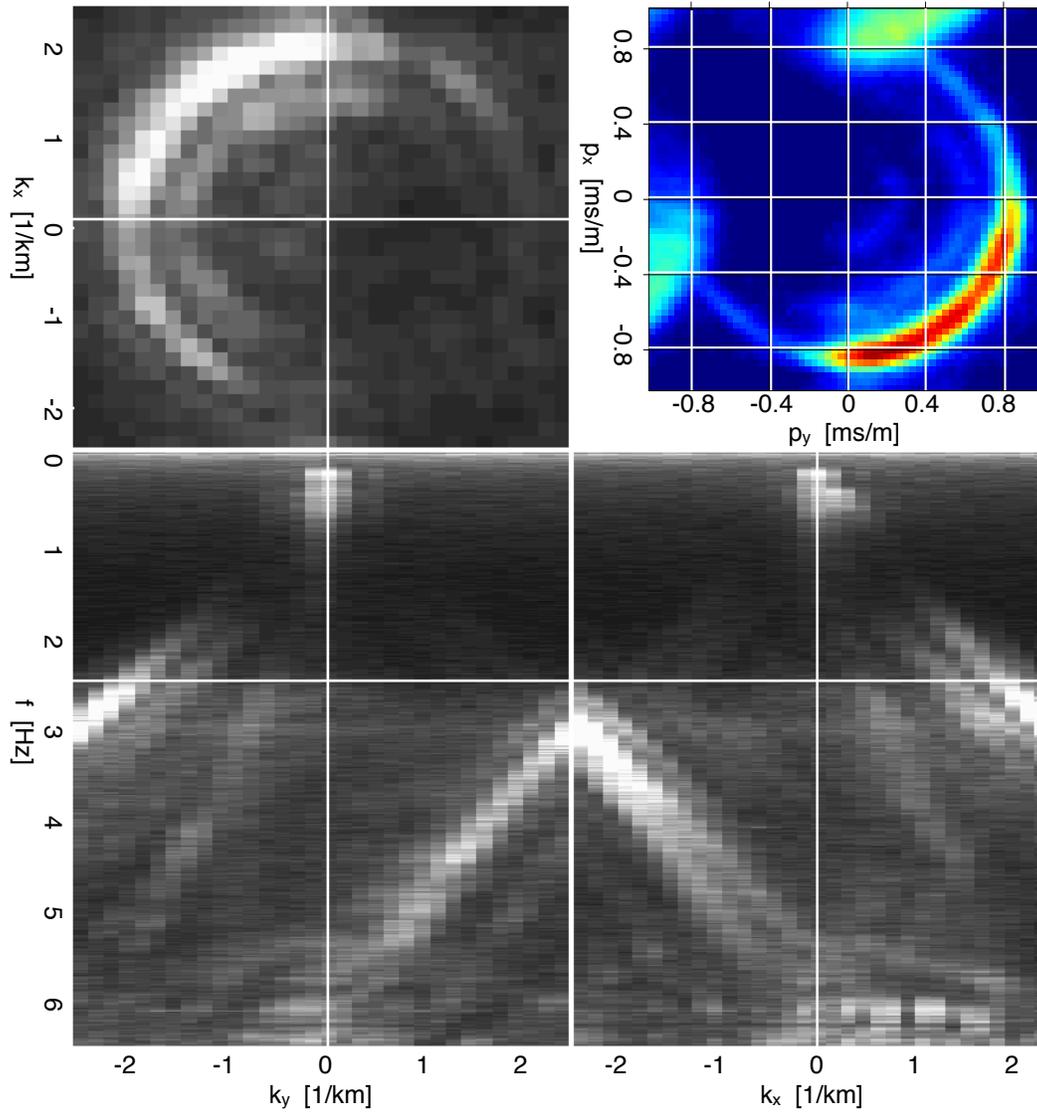


Figure 4: Cross-sections through frequency-wavenumber spectra cube and averaged over the data recorded between 20:00 and 21:00 on day 1. The top right panel contains a beam-forming experiment for the frequency band of 1 Hz to 3.5 Hz. [CR] [sjoerd1/. MDspectral1](#)

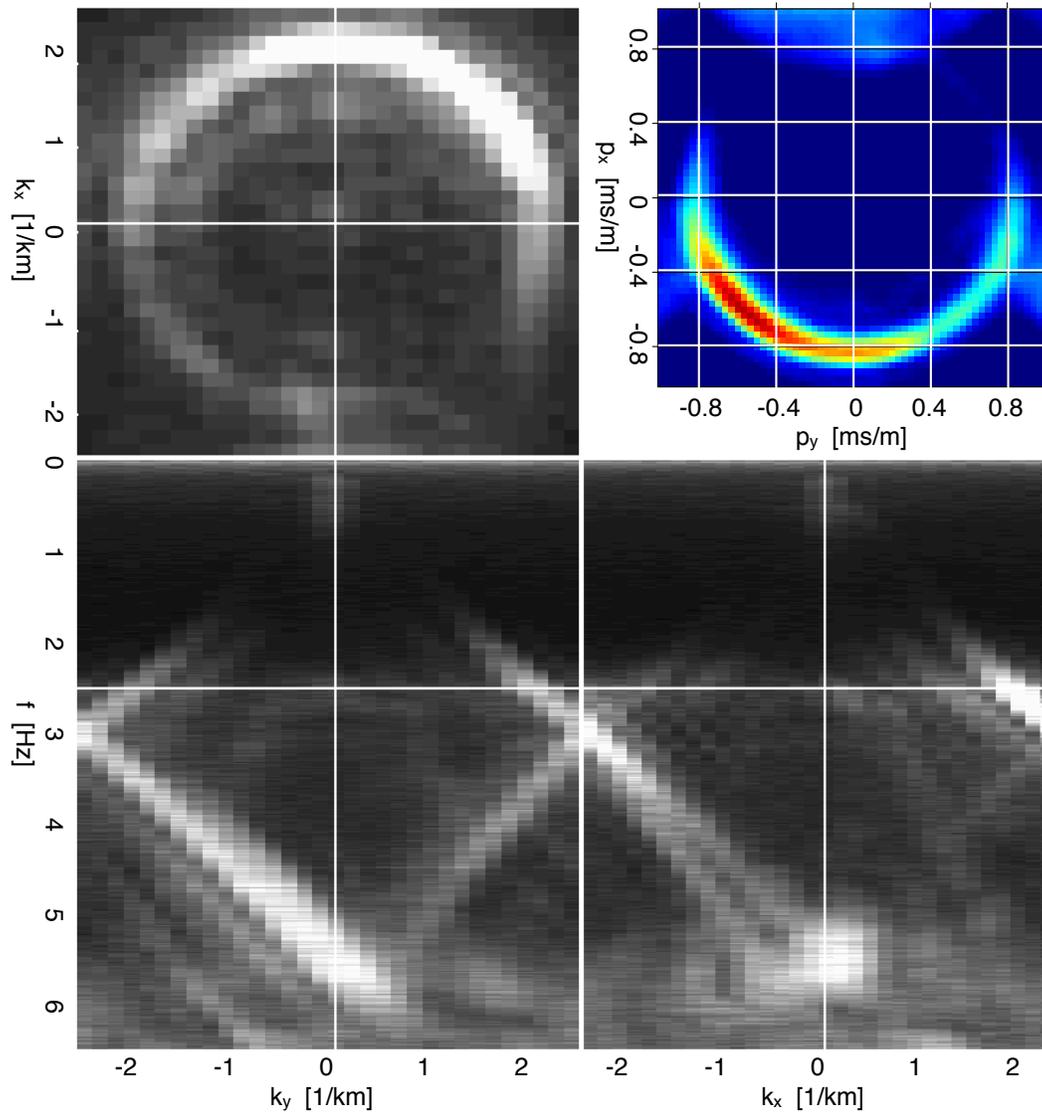


Figure 5: Cross-sections through frequency-wavenumber spectra cube and averaged over the data recorded between 11:00 and 12:00 on day 2. The top right panel contains a beam-forming experiment for the frequency band of 1 Hz to 3.5 Hz. [CR] [sjoerd1/. MDspectra2](#)

and the directionality in the ambient seismic field. To illustrate this, a correlation was computed using each hour of data, for a virtual source at $(s_x, s_y) = (0, 0)$ m and stations at the 4 corners of the array. The gathers are shown in Figure 8. The time asymmetry of the retrieved result can be linked to the directionality of the ambient seismic field. For example, for the retrieved signals on day 1 using the data recorded from 20:00 to 21:00, the energy in the correlations is dominantly causal for the station in the south and dominantly acausal for the station in the north, which is consistent with the observation that energy is traveling southwards for this time period. A high concentration of energy arriving at a station focused at approximately $t = \pm 2$ s corresponds to a good convergence rate. The observed convergence rates can be related to the strength of the ambient seismic field (see the frequency spectra of figure 2). Notice the crisp causal Green's functions for daylight hours arriving at the south station.

SPATIAL VARIABILITY AND DISPERSION

Stacking is required to further analyze the obtained virtual seismic survey. First, each spatial axis in the virtual seismic survey is transformed from source-geophone coordinates (s, g) , to midpoint (half)offset coordinates (m, h) . Then, under a local 1D approximation, the offset coordinates (h_x, h_y) are transformed into cylindrical coordinates and stacked over azimuth. The result is a virtual seismic survey at each midpoint, as a function of radial offset h_r . This result is analogous to the spatial auto-correlation of Aki (1957) (Asten, 2006; Yokoi and Margaryan, 2008). To investigate if there is any spatial variability despite the local 1D approximation, the retrieved Green's function for a full radial offset of $2h_r = 1640$ m are displayed in Figure 9 for two midpoint slices at $m_y = 100$ m and $m_x = 100$ m. Notice the Rayleigh waves arrive earlier for midpoints with negative m_x and positive m_y than for midpoints with positive m_x and negative m_y . This is consistent with the observation of higher velocities on the east side than on that the west end of the array, see previous section. Notice how the dominant phase of the arrival-train travels with a different velocity than the group velocity, indicating dispersion.

The fundamental mode of a Rayleigh surface wave is represented by a zero-order Bessel function of the first kind, stretched by phase velocity, station distance and frequency (Aki, 1957; Okada, 2003). In Figure 10, this is observed in a frequency range between 1.5 Hz and 6 Hz, for retrieved Rayleigh waves averaged over all midpoints in the array. The jump in the dispersion curve at 4.5 Hz is caused by a poor interpolation technique in the rotation from Cartesian to cylindrical offsets. The lowest retrieved frequencies travel with a wavelength of approximately 1250 m, suggesting sensitivity to a depth of approximately 1250 m. If more data was analyzed, convergence could have been achieved without averaging over azimuth or midpoint, preserving greater detail in subsurface properties for further study, and potentially yielding sufficient energy below 1 Hz and thus sensitivity to greater depth.

CONCLUSIONS

Rayleigh wave Green's functions were successfully retrieved from 48 hours of ambient seismic noise in the frequency range 1 Hz to 7 Hz. The radiation pattern and convergence rate of the retrieved Green's functions show dependence on the directivity and strength of energy in the ambient seismic field.

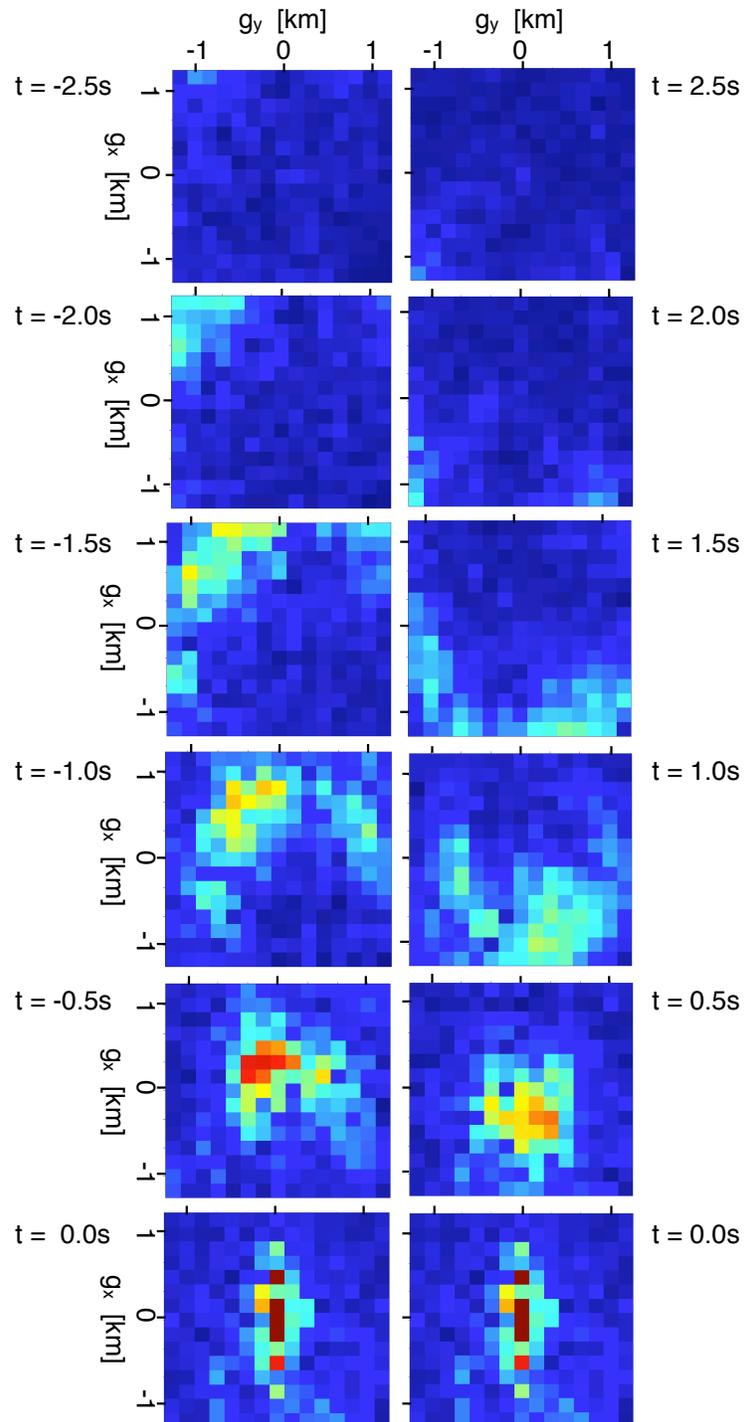


Figure 6: Snap shots of a movie of the acausal (left) and causal (right) parts of retrieved Green's functions for a virtual source at $(s_x, s_y) = (0, 0)$ m. [NR] [sjoerd1/. CenterG](#)

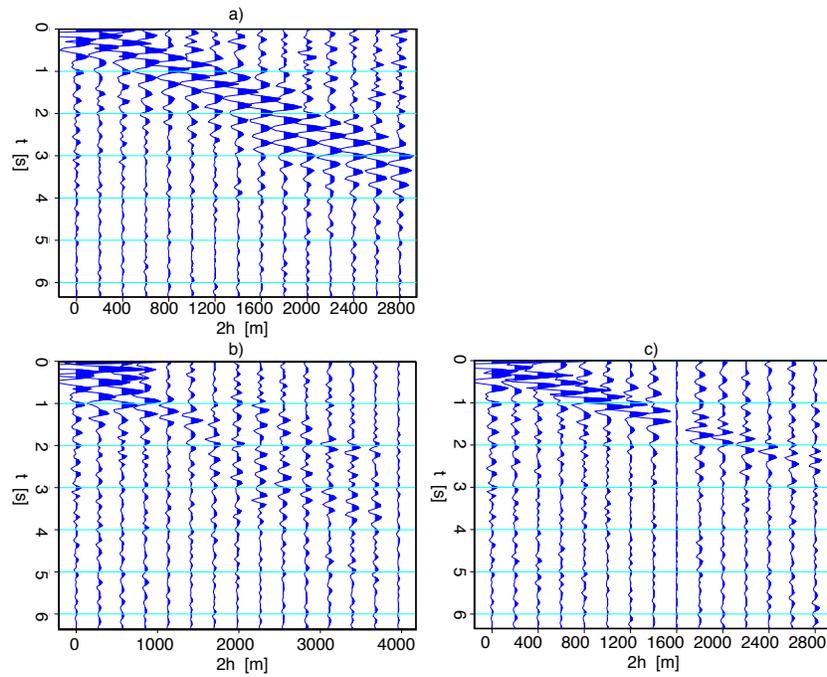


Figure 7: Causal retrieved Green's functions plotted as a function of full offsets ($2h$) for a source positioned at $(s_x, s_y) = (1400, -1400)$ m. a) A g_x -slice at $g_y = -1400$ m; b) a g_y -slice at $g_x = -1400$ m and; c) a diagonal slice. [CR] sjoerd1/. Glongoffset

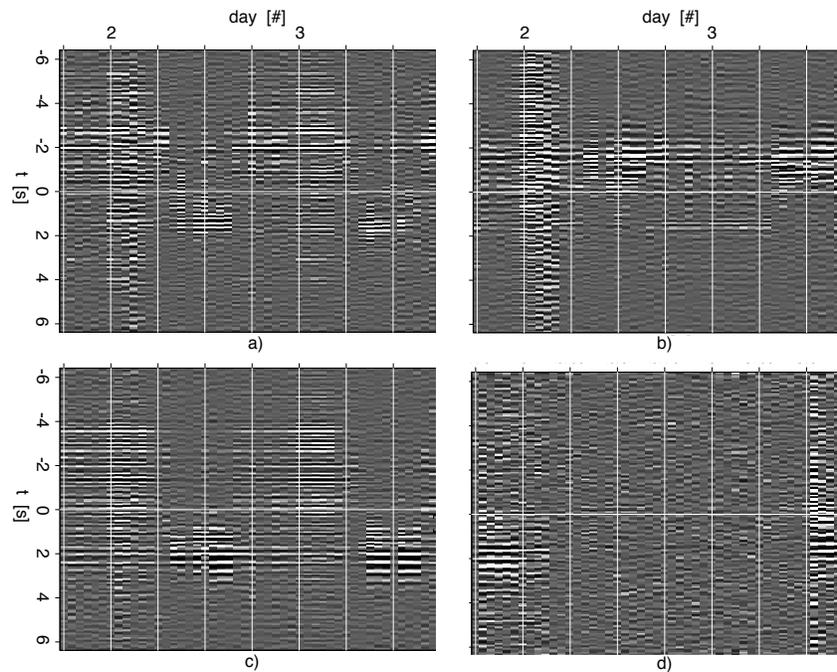


Figure 8: Green's functions for a virtual shot gather of a source positioned at $(s_x, s_y) = (0, 0)$ m and stations at; a) $(g_x, g_y) = (1400, -1400)$ m; b) $(g_x, g_y) = (1400, 1400)$ m; c) $(g_x, g_y) = (-1400, -1400)$ m; d) $(g_x, g_y) = (-1400, 1400)$ m. [CR] sjoerd1/. TGcomplete

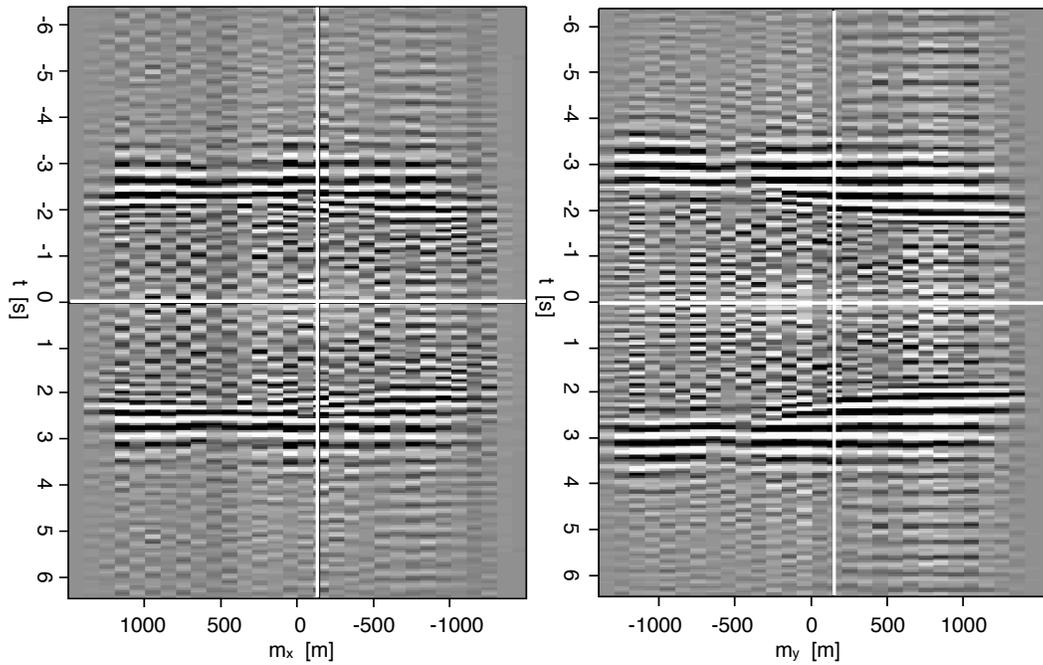


Figure 9: Two slices through a cube of retrieved Green's functions as a function of midpoints, for fixed full radial offset of $2h_r = 1640$ m. Left: $m_y = 100$ m, right: $m_x = 100$ m. [CR] sjoerd1/. Mvariability

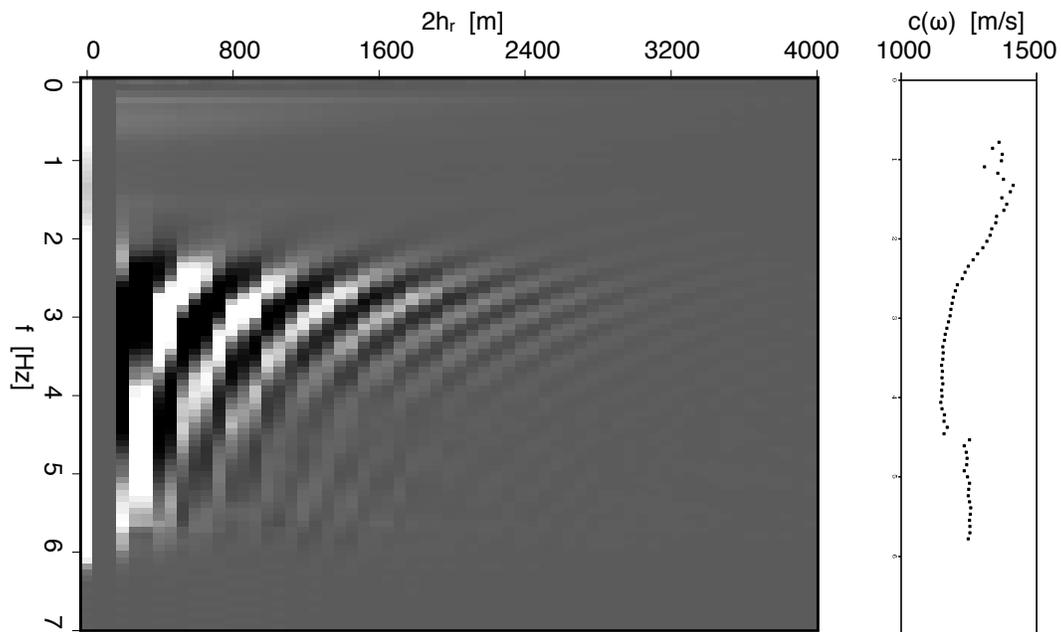


Figure 10: Left; real part of the retrieved Green's functions in the frequency-domain, shown as function of full radial offset, averaged over all midpoints. Right; estimated dispersion curve for the retrieved Green's functions shown in the left panel. [CR] sjoerd1/. CSdispersion

ACKNOWLEDGEMENTS

The authors would like to thank Jon Claerbout, Bob Clapp, Jesse Lawrence, Germán Prieto for useful discussions and suggestions. The authors would like to thank the sponsors of the Stanford Exploration Project for their financial support. Special thanks to Saudi Aramco's EXPEC Advanced Research Center for providing data and for permission to publish.

REFERENCES

- Aki, K., 1957, Space and time spectra of stationary stochastic waves, with special reference to microtremors: *Bulletin of the Earthquake Research Institute*, **35**, 415–456.
- Artman, B., 2006, Imaging passive seismic data: *Geophysics*, **71**, SI177–SI187.
- , 2007, *Passive seismic imaging*: PhD thesis, Stanford University.
- Asten, M. W., 2006, On bias and noise in passive seismic data from finite circular array data processed using spac methods: *Geophysics*, **71**, V153–V162.
- Bakulin, A. and R. Calvert, 2006, The virtual source method: Theory and case study: *Geophysics*, **71**, SI139–SI150.
- Berkhout, A. J. and D. J. Verschuur, 2006, Imaging of multiple reflections: *Geophysics*, **71**, SI209–SI220.
- Bussat, S. and S. Kugler, 2009, Recording noise - estimating shear-wave velocities: Feasibility of offshore ambient-noise surface-wave tomography (answt) on a reservoir scale: *SEG Technical Program Expanded Abstracts*, **28**, 1627–1631.
- Claerbout, J. F., 1968, Synthesis of a layered medium from its acoustic transmission response: *Geophysics*, **33**, 264–269.
- Cole, S. P., 1995, *Passive seismic and drill-bit experiments using 2-d arrays*: PhD thesis, Stanford University.
- Dellinger, J., 2008, Low frequencies using conventional sensors: “sign-bit” recording revisited: *SEG Technical Program Expanded Abstracts*, **27**, 149–153.
- Dellinger, J. A. and J. Yu, 2009, Low-frequency virtual point-source interferometry using conventional sensors: 71st Meeting, European Association of Geoscientists and Engineers, Expanded Abstracts, Expanded Abstracts, X047.
- Dragonov, D., K. Wapenaar, W. Mulder, J. Singer, and A. Verdel, 2007, Retrieval of reflections from seismic background-noise measurements: *Geophys. Res. Lett.*, **34**, L04305–1 – L04305–4.
- Fu, Q. and Y. Luo, 2009, Locating micro-seismic epicenters in common arrival time domain: *SEG Technical Program Expanded Abstracts*, **28**, 1647–1651.
- Jervis, M. and S. N. Dasgupta, 2009, Recent microseismic monitoring results from vsp and permanent sensor deployments in Saudi Arabia: *EAGE Workshop on Passive Seismic*, Expanded Abstracts, A10.
- Landes, M., N. M. Shapiro, S. Singh, and R. Johnston, 2009, Studying shallow seafloor structure based on correlations of continuous seismic records: *SEG Technical Program Expanded Abstracts*, **28**, 1693–1697.
- Okada, H., 2003, *The microtremor survey method*. Geophysical Monograph, No. 12: Society of Exploration Geophysicists.
- Schuster, G. T., J. Yu, J. Sheng, and J. Rickett, 2004, Interferometric/daylight seismic imaging: *Geophys. J. Int.*, **157**, 838–852.
- Schuster, G. T. and M. Zhou, 2006, A theoretical overview of model-based and correlation based redatuming methods: *Geophysics*, **71**, SI103–SI110.

- Wapenaar, K., 2004, Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation: *Phys. Rev. Lett.*, **93**, 254301-1 – 254301-4.
- Xiao, X., Y. Luo, Q. Fu, M. Jervis, S. Dasgupta, and P. Kelamis, 2009, Locate microseismic by seismic interferometry: EAGE Workshop on Passive Seismic, Expanded Abstracts, A22.
- Yokoi, T. and S. Margaryan, 2008, Consistency of the spatial autocorrelation method with seismic interferometry and its consequence: *Geophysical Prospecting*, **56**, 435-451.

Cyclic 1D matching of time-lapse seismic data sets: A case study of the Norne Field

Gboyega Ayeni

ABSTRACT

Seismic cross-equalization attenuates unwanted or non-production related artifacts from time-lapse seismic data sets. Two robust post-imaging cross-equalization methods are considered. First, an efficient cyclic 1D correlation method is used to estimate vertical and lateral displacement components between baseline and monitor images. To obtain accurate estimates of production-related changes, all displacement components must be taken into account. Next, a cyclic 1D match-filtering method that estimates the optimal filtering parameters attenuates residual artifacts. Application to the Norne time-lapse seismic data set shows that these methods provide a powerful time-lapse cross-equalization scheme.

INTRODUCTION

Time-lapse (4D) seismic provides important information about production-related changes in reservoir properties. Over the past decade, time-lapse seismic technology has developed into an integral part of most hydrocarbon reservoir management projects. Several authors have discussed important technical considerations and the business impact of time-lapse seismic (Johnston et al., 1998; Rickett and Lumley, 2001; Whitcombe et al., 2004; Zou et al., 2006; Ebaid et al., 2009).

Imperfect survey geometry repetition, non-repeatable ambient noise and other discrepancies generate artifacts that contaminate production-related differences between seismic data sets. Before reservoir property changes can be extracted from time-lapse seismic data, these artifacts must be attenuated. The process of attenuating artifacts from time-lapse data sets is generally called seismic cross-equalization (Rickett and Lumley, 2001). In this paper, two robust post-imaging cross-equalization steps are considered.

Because reservoir depletion changes the stress-states in and around a producing reservoir, seismic travel times and path lengths are different between surveys. However, in practice, the baseline and monitor data sets are migrated with a single (baseline) velocity, leading to *apparent* vertical and lateral displacements between seismic images. Because it is expensive to estimate all components of the displacement vector field, it is common practice to only consider vertical displacements. Instead of this *vertical-only* approximation, displacement components are efficiently computed with a cyclic 1D correlation method (Hale, 2009). It is shown that in order to obtain accurate estimates of reservoir property changes, all displacement components must be taken into account.

To ensure that only production-related changes within the reservoir are interpreted, differences in seismic signal in non-reservoir regions should be minimal. After aligning all

monitor images to the baseline (by removing the estimated apparent displacements), an optimal cyclic 1D match-filtering method is used to attenuate differences in the non-reservoir regions. This cyclic filtering method is an extension of the optimal matching method described by Ayeni and Nasser (2009). Together with the cyclic correlation method described above, this matching method forms a robust and effective seismic cross-equalization scheme.

In this paper, I first briefly discuss the cyclic shifting and match-filtering methods. Next, I apply these to four time-lapse seismic data sets from the Norne field. The Norne study shows that neglecting lateral displacements can cause spurious time-lapse signals. Finally, I show that these cross-equalization methods significantly improve the quality of the time-lapse images and their derivatives.

APPARENT DISPLACEMENT VECTORS BY CYCLIC 1D SEARCH

Components of the apparent displacement vectors between time-lapse seismic data sets can be efficiently computed using a cyclic 1D search approach (Hale, 2009). This method involves a sequential cyclic 1D search for cross-correlation peaks in different directions. This removes the need for a computationally expensive 3D search for cross-correlation peaks or the inaccurate assumption that displacements are only in the vertical direction. To improve the resolution of displacement estimates in all directions, the input data are first band-passed to remove very low frequencies. This preprocessing step serves a similar purpose as the prediction-error filtering step of Hale (2009).

This implementation utilizes correlation gates and lags of varying sizes and an acceptance criterion based on the relative values of the cross-correlation peaks. To first capture the low-frequency, high-amplitude displacements, this iterative procedure starts with large correlation gates and lags. The correlation gates and lags are then decreased as a function of iteration. Displacement estimates are filtered based on the displacement size and cross-correlation peak. For each iteration, the results (shifted image(s) and accumulated displacements) from one direction (e.g. vertical) are the inputs to the next direction (e.g. in-line). For each direction, shifts are accumulated after interpolating the image and displacement components from the previous step. Algorithm 3 summarizes this procedure.

The estimated displacements are used to align the baseline and monitor data sets prior to time-lapse image computation. They are also useful for estimating velocity and geomechanical changes.

Velocity, path-length and travel time changes

Depletion-related strain and velocity changes will perturb the travel path and length of a normally-incident ray. The distance dz travelled by the ray in time dt at velocity v is given by

$$dz = vdt. \quad (1)$$

For small perturbations $\delta z \ll dz$, $\delta v \ll dv$ and $\delta t \ll dt$, we can write

$$\frac{d(\delta z)}{dz} = \frac{\delta v}{v} + \frac{d(\delta t)}{dt}, \quad (2)$$

where, $d(\delta z)/dz = \epsilon_{zz}$ is the vertical “depth” strain and $d(\delta t)/dt = \epsilon_{tt}$ is the apparent “time” strain. The R -factor (Hatchell and Bourne, 2005a) relates the fractional change in velocity $\delta v/v$ to the depth strain ϵ_{zz} and time strain ϵ_{tt} :

$$\frac{\delta v}{v} = -R\epsilon_{zz} = -\frac{R}{1+R}\epsilon_{tt}. \quad (3)$$

The R -factor can be measured from core samples in the laboratory (Hatchell and Bourne, 2005a; Bathija et al., 2009), derived from the estimated time-shifts (Hawkins et al., 2007), or derived from other production-related data (Hodgson et al., 2007). For most reservoir rocks, R values of 4 to 6 have been shown to be reasonable (Hatchell and Bourne, 2005b; Hodgson et al., 2007). In this paper, a constant R value of 5 is assumed.

CYCLIC 1D MATCH-FILTERING BY EVOLUTIONARY PROGRAMMING

The match-filtering problem between two data sets, one baseline and one monitor, can be posed in the frequency domain as follows:

$$\mathbf{F}(\omega) = \frac{\overline{\mathbf{M}}(\omega)\mathbf{B}(\omega)}{\overline{\mathbf{M}}(\omega)\mathbf{M}(\omega) + \epsilon^2}. \quad (4)$$

where $\mathbf{F}(\omega)$ is the match filter, $\mathbf{B}(\omega)$ is the baseline data, and $\mathbf{M}(\omega)$ is the monitor data at frequency ω . The damping factor (ϵ) in the denominator prevents zero-division. The overline denotes complex conjugation. Match filters can be estimated globally (for all trace locations), locally (on a trace-by-trace basis) or both. In this paper, only local 1D match filters are considered.

At each trace location, local match filters are non-linear functions of the estimation parameters. Usually, these parameters are selected by a manual trial-and-error approach. This can be a tedious or impossible challenge due to the strong non-linearity of the problem and the large size of seismic data sets. Furthermore, in most cases a single set of parameters is used for all traces locations. While a particular parameter set may result in satisfactory filter estimates in some parts of the data volume, it may result in very poor filters in other parts.

Here, an evolutionary programming (EP) method is used estimate the optimal filter parameters at each trace location. EP belongs to a class of global optimization methods called evolutionary algorithms. These algorithms solve optimization problems using Darwinian evolutionary principles of natural selection. Although any global optimization method may be used, the EP method was chosen because of its easy implementation, easy parameterization and flexibility. The optimized estimation parameters include estimation window size and taper, filter-length and the damping factor. Ayeni and Nasser (2009) provide a more detailed description and an application of evolutionary programming to the match filtering problem. An extended version of this filter estimation approach that is used in this paper is presented in Algorithm 4.

Algorithm 3 Apparent displacements by cyclic 1D search

```

b  $\leftarrow$  preprocess(Base)
m  $\leftarrow$  preprocess(Monitor)
while  $s_{max}^{ia=1:3} \geq s_{stop}^{ia=1:3}; j = 1 : niter$  do
  for axis = 1:3 do
    if axis = 1 then
       $s^1 = s[xcorr_{peak}(m_j|b)]$  !compute displacements btw  $m_j$  &  $b$  along axis 1
       $z^1 \leftarrow s^1 \langle z^1 \rangle + s^1$  !interpolate and accum. displacements along axis 1
       $z^2 \leftarrow s^1 \langle z^2 \rangle$  !interpolate displacements along axis 2
       $z^3 \leftarrow s^1 \langle z^3 \rangle$  !interpolate displacements along axis 3
       $m_j^1 \leftarrow s^1 \langle m_j \rangle$  !interpolate monitor along axis 1
    else if axis = 2 then
       $s^2 = s[xcorr_{peak}(m_j^1|b)]$ 
       $z^1 \leftarrow s^2 \langle z^1 \rangle$ 
       $z^2 \leftarrow s^2 \langle z^2 \rangle + s^2$ 
       $z^3 \leftarrow s^2 \langle z^3 \rangle$ 
       $m_j^2 \leftarrow s^2 \langle m_j^1 \rangle$ 
    else if axis = 3 then
       $s^3 = s[xcorr_{peak}(m_j^2|b)]$ 
       $z^1 \leftarrow s^3 \langle z^1 \rangle$ 
       $z^2 \leftarrow s^3 \langle z^2 \rangle$ 
       $z^3 \leftarrow s^3 \langle z^3 \rangle + s^3$ 
       $m_j^3 \leftarrow s^3 \langle m_j^2 \rangle$ 
       $m_{j+1} \leftarrow m_j^3$ 
    end if
  end for
end while
  Monitor  $\leftarrow z^{1:3} \langle Monitor \rangle$ 

```

s: displacements

z: cumulative displacements

b: preprocessed base

m: preprocessed monitor

$xcorr_{peak}$: correlation peak

$a \langle b \rangle$: interpolation of b with displacements a

Algorithm 4 Optimal cyclic 1D matching by evolutionary programming

```

for  $irepeat = 1:nrepeat$  do
  for  $itrace = 1:ntrace$  do
    (Re)Initialize: estimation window, filter length, damping, etc.
    Evaluate: fitness (nrms)
    while  $nrms_{iter} \geq nrms_{stop}; iter = 1 : niter$  do
      Select: fittest individuals
      Mutate: new offsprings
      Evaluate: fitness (nrms)
    end while
  end for
  Apply constraints
end for

```

CASE STUDY

The Norne field is located in the Norwegian North sea. The main field is a 9 km x 3 km horst block comprised of high-porosity, high-permeability, high net-to-gross lower and middle Jurassic sandstones. The field was discovered in 1991 and production started in 1997. For this study, four high-resolution Q-marine time-lapse seismic data sets acquired in 2001, 2003, 2004 and 2006 are considered. The 2001 data is taken as the baseline. Osdal et al. (2006) and Aarre (2008) provide more detailed description of the geology, production history, time-lapse seismic acquisition and processing, and some interpretation of the time-lapse seismic response at the Norne.

The apparent vertical displacements (time-shifts) for three cycles between the 2006 and the 2001 data are shown in Figure 1. In this figure, and in similar ones throughout this paper, the orthogonal lines show slices through a 3D volume at the positions indicated by each line. Figure 2 shows the 2006 time-shifts together with the mapped top of the reservoir. The corresponding lateral displacements in the inline and cross-line directions (after three cycles) are shown in Figure 3. Figures 4 to 6 show the absolute vertical and lateral displacements overlain by arrows indicating the displacement direction. Estimated velocity changes at different production stages are shown in Figure 7. It is assumed that the time-shift estimates obtained by considering displacements in all directions are accurate. Differences between time-shifts, velocity changes and vertical strains obtained by considering all displacement components and those obtained by considering only vertical displacements are shown in Figure 8. The optimal match-filter parameters obtained by the evolutionary programming method are shown in Figure 9. Figures 10 and 11 show time-lapse images obtained after different processing steps. Amplitude errors resulting from considering only vertical displacements are shown in Figures 12.

DISCUSSION

The sequential cyclic 1D search method provides a robust alternative to a full 3D search for cross-correlation peaks. By iteratively estimating displacements in a cyclic manner

using variable correlation gates and lags, it is possible to obtain high-resolution estimates of displacement components (Figures 1 to 3). These displacement components contain different information about changes within and around the reservoir and may be interpreted qualitatively (Figures 4 to 6). Analysis of the estimated displacements show that in general, velocity within the reservoir decreases with time (Figure 7). A velocity decrease of up to 5 percent is observed between 2001 and 2006 (Figure 7(c)).

There are inherent errors caused by considering only vertical displacements in time-lapse seismic data sets (Figure 8(a)). In the current example, this approximation leads to time-shift overestimation below the reservoir, and time-shift underestimation above the reservoir. Errors in time-shift estimates, which are more pronounced around dipping reflectors, will map into errors in velocity change and vertical-strain estimates (Figures 8(b) and 8(c)).

Filter parameters derived from the proposed method (Figure 9) enable computation of the optimal filters at each trace location. Before computing time-lapse images, the monitor data sets are shifted into alignment with the baseline by interpolating them with the estimated displacements. After data alignment (Figures 10(b) and 11(b)), optimal match-filtering removes residual artifacts, thereby improving the time-lapse data-quality (Figures 10(c) and 11(c)). Finally, especially at steeply dipping flanks, considering only vertical displacements can lead to erroneous time-lapse amplitude changes (Figure 12).

CONCLUSIONS

Time-lapse seismic data sets can be efficiently cross-equalized by cyclic 1D matching methods. The cyclic 1D shifting and optimized match-filtering methods form an efficient and robust cross-equalization scheme. From the case study, it is apparent that in order to obtain accurate estimates of reservoir property changes, all displacement components must be taken into account. Where estimated parameters are important for decision making, considering only vertical displacements may result in serious undesirable consequences. Future work will consider ways to fully integrate all displacement components in time-lapse seismic imaging.

ACKNOWLEDGEMENTS

Thanks to Statoil for the release of the Norne data. Thanks to the NTNU/IO CENTRE for their cooperation and coordination of the Norne Cases. Special thanks to Trine Alsos (Staoil), Mohsen Dadashpour (NTNU) and David Echeverria (Stanford Smart Fields) for their continued assistance throughout this study.

REFERENCES

- Aarre, V., 2008, On the presence, and possible causes, of apparent lateral shifts below the Nome reservoir: SEG Technical Program Expanded Abstracts, **27**, 3174–3178.
- Ayeni, G., and M. Nasser, 2009, Optimized local matching of time-lapse seismic data: A case study from the gulf of mexico: SEG Technical Program Expanded Abstracts, **28**, 3939–3943.

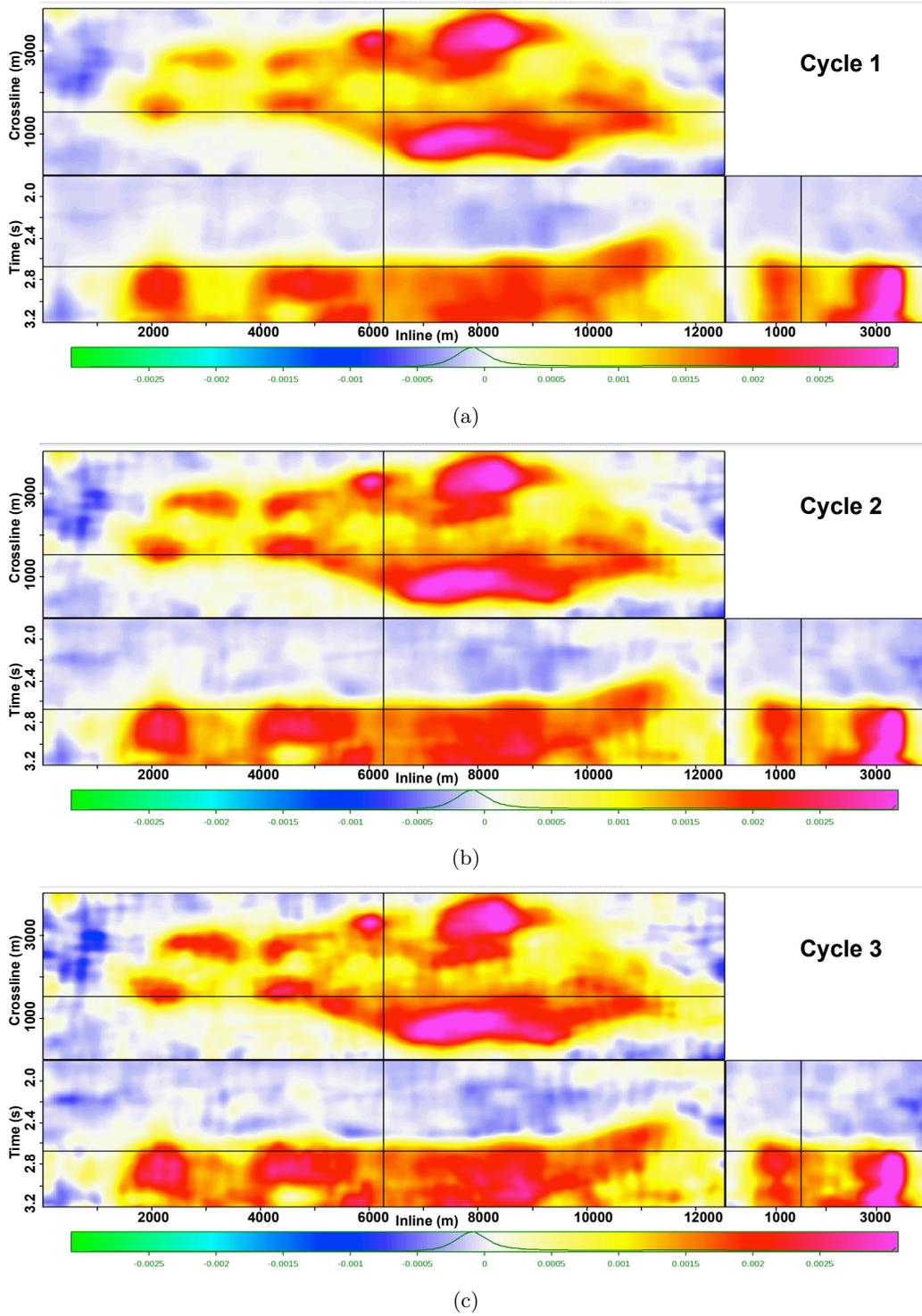


Figure 1: Apparent vertical (time-) shifts estimates between the baseline and the 2006 data sets after one (top), two (middle) and three (bottom) cycles. Note the improvement in resolution with the number of cycles. [CR]. `gayeni1/. z-shift-1,z-shift-2,z-shift-3`

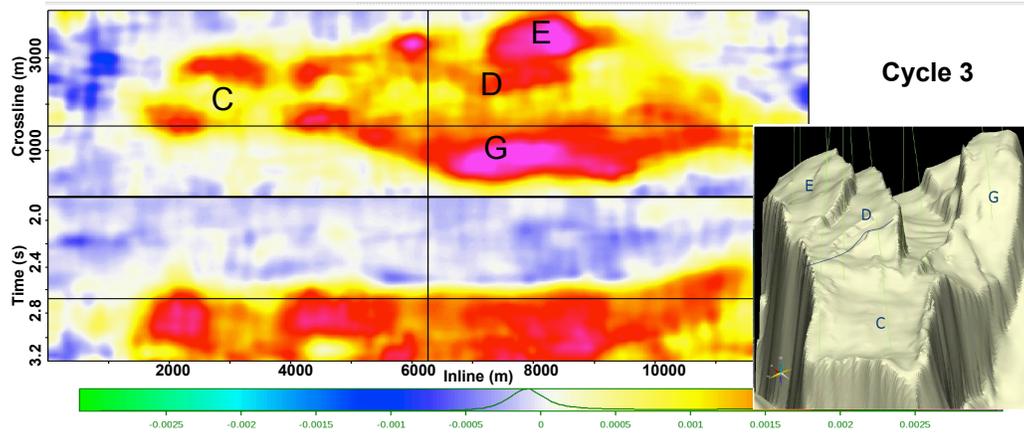
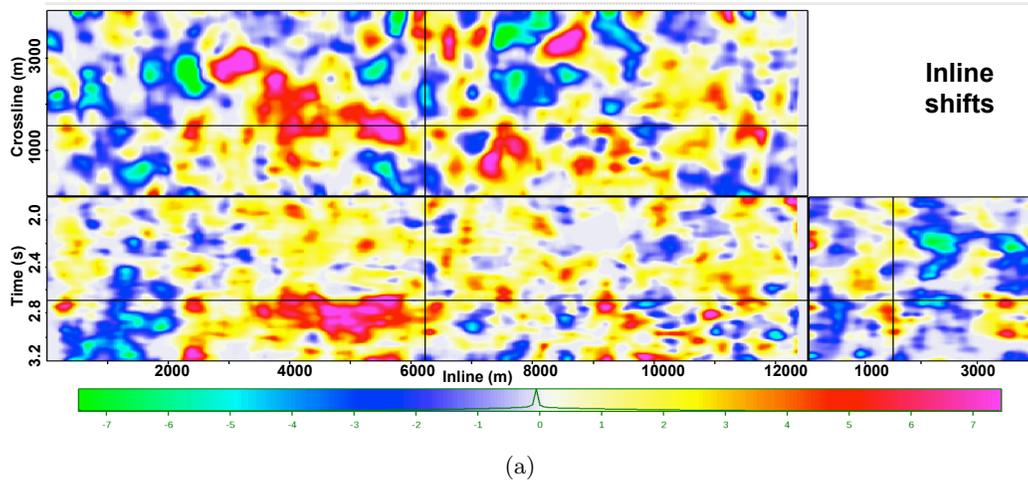
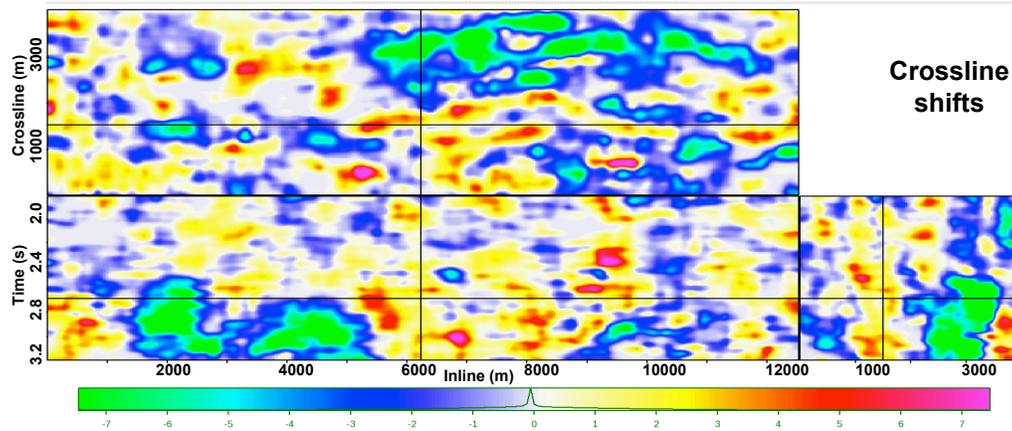


Figure 2: Time-shifts between the baseline and the 2006 data and the top of the reservoir. Note the similarity between the displacements correspond to known producing segments in the filed. [CR]. gayeni1/. z-shift-3-TopNot



(a)



(b)

Figure 3: Apparent lateral displacements between the baseline and the 2006 data. [CR]. gayeni1/. x-shift-3,y-shift-3

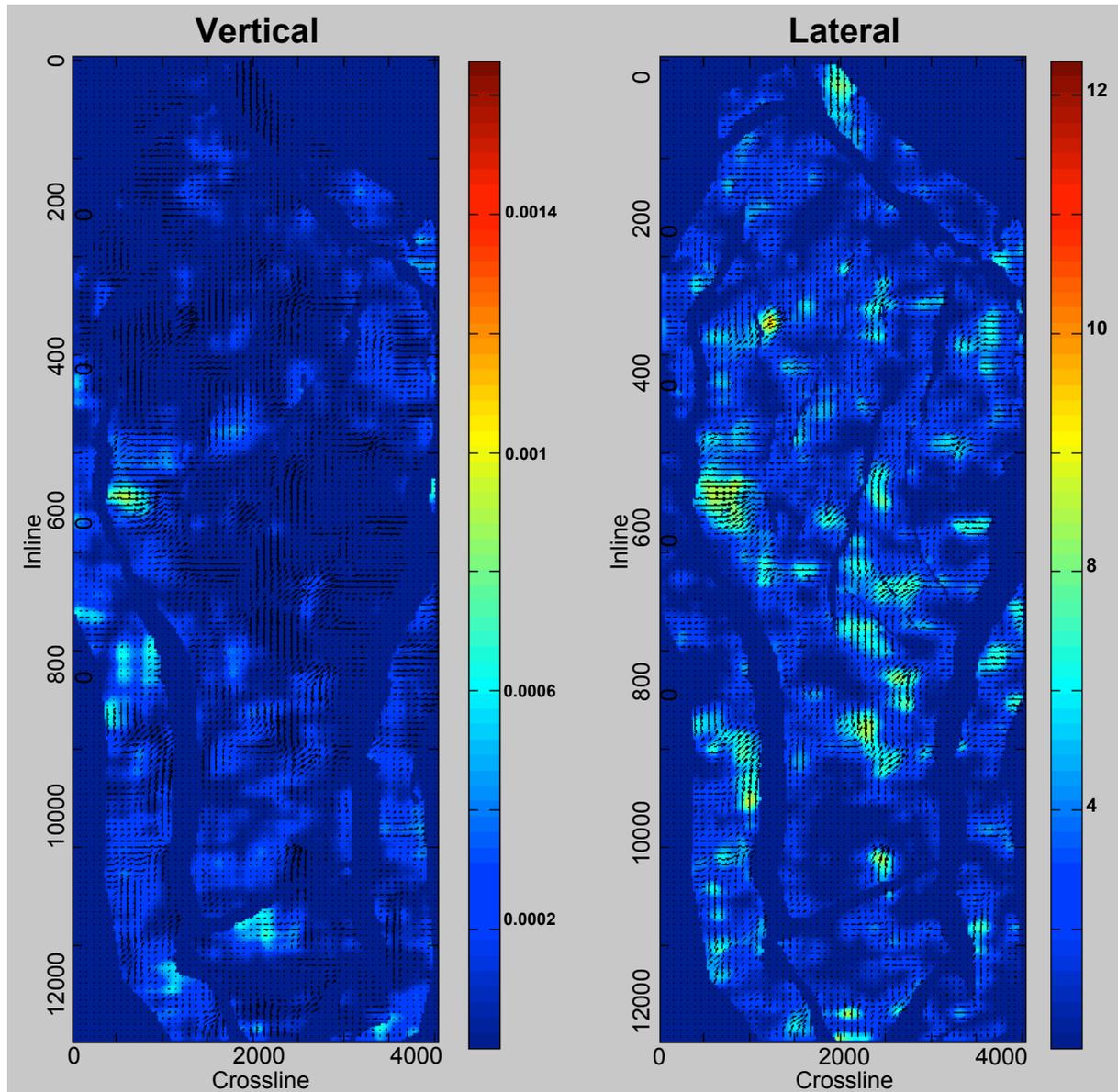


Figure 4: Absolute vertical and lateral displacements at the top of the reservoir between the baseline and the 2003 monitor. The arrows indicate the displacement direction. [CR].
`gayeni/. shift-compare1`

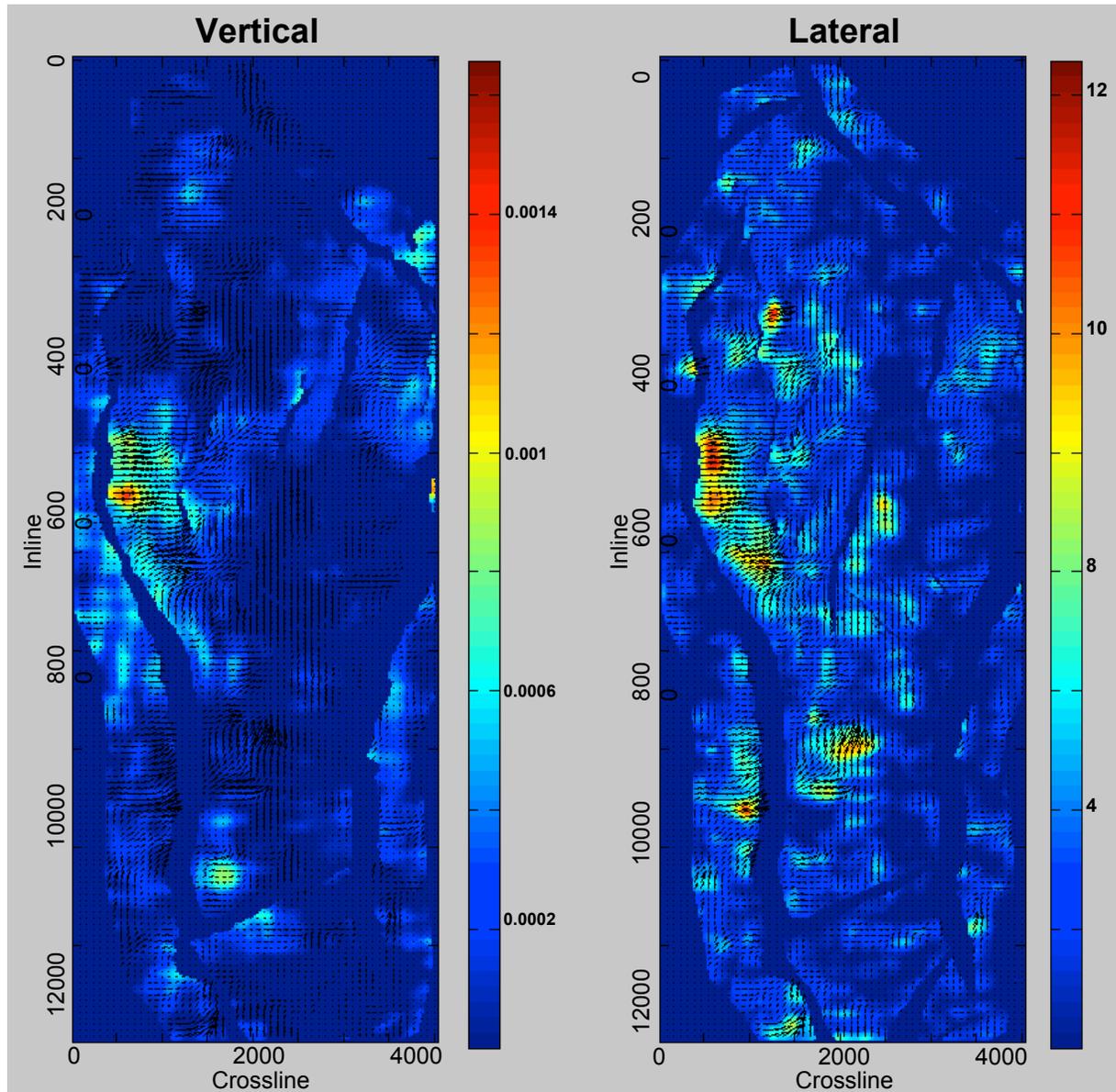


Figure 5: Absolute vertical and lateral displacements at the top of the reservoir between the baseline and the 2004 monitor. The arrows indicate the displacement direction. [CR].
`gayeni1/. shift-compare2`

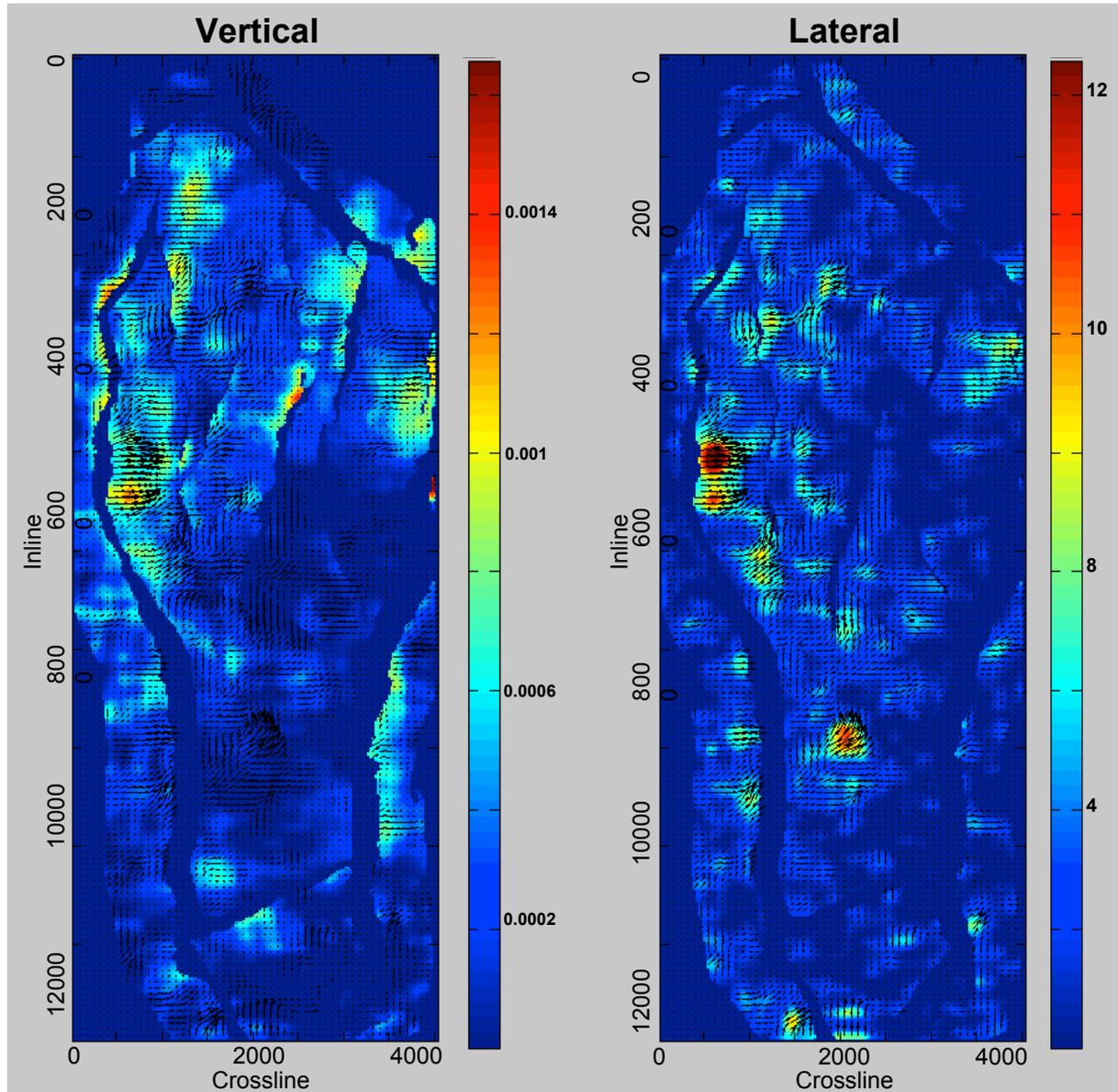


Figure 6: Absolute vertical and lateral displacements at the top of the reservoir between the baseline and the 2006 monitor. The arrows indicate the displacement direction. [CR].
`gayeni/. shift-compare3`

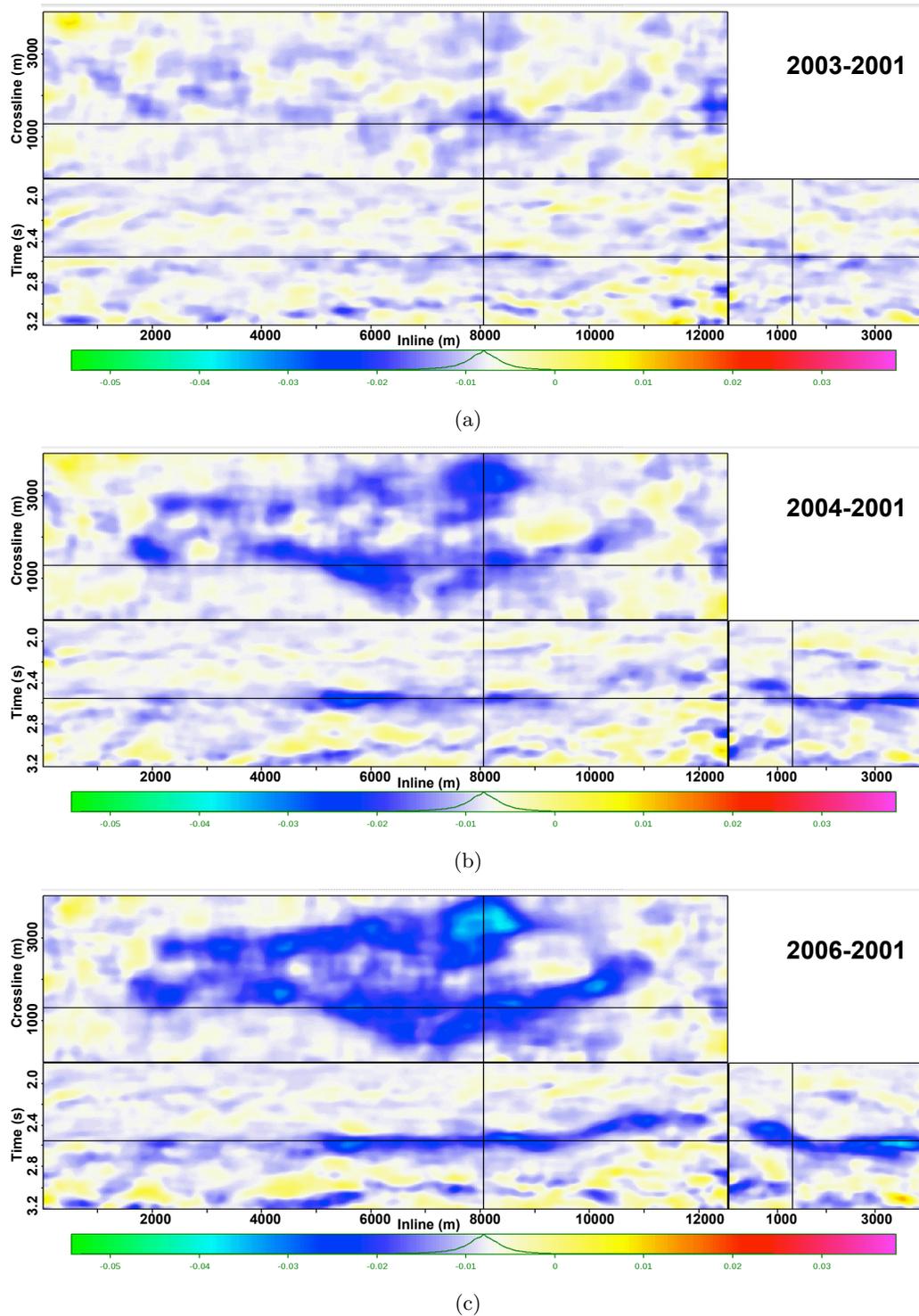


Figure 7: Fractional velocity change between the baseline and the 2003 (a), 2004 (b) and 2006 (c) monitor data. Note that the velocity within the reservoir decreases with time.

[CR]. `gayeni1/.dvv-shift-2003,dvv-shift-2004,dvv-shift-2006`

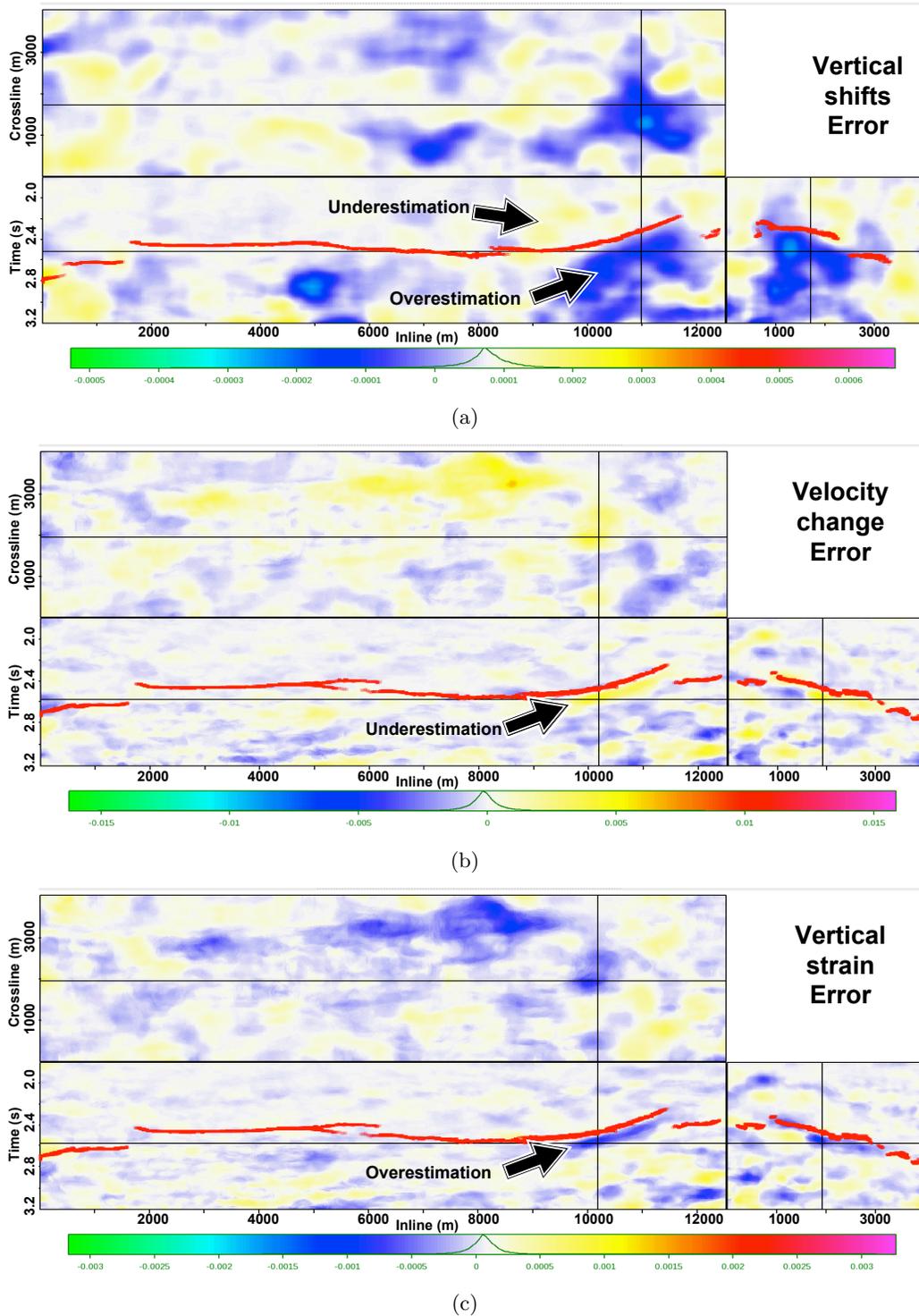


Figure 8: Time-shift (a), velocity change (b) and vertical strain (c) errors between the baseline and the 2006 data as a result of assuming on vertical displacements. In each case, the result obtained by considering only vertical displacements were subtracted from those obtained by considering all directions with the cyclic search method. The red line indicates the reservoir top. [CR]. `gayeni/. z-shift-3-error,dvv-shift-2006-error,ez-shift-2006-error`

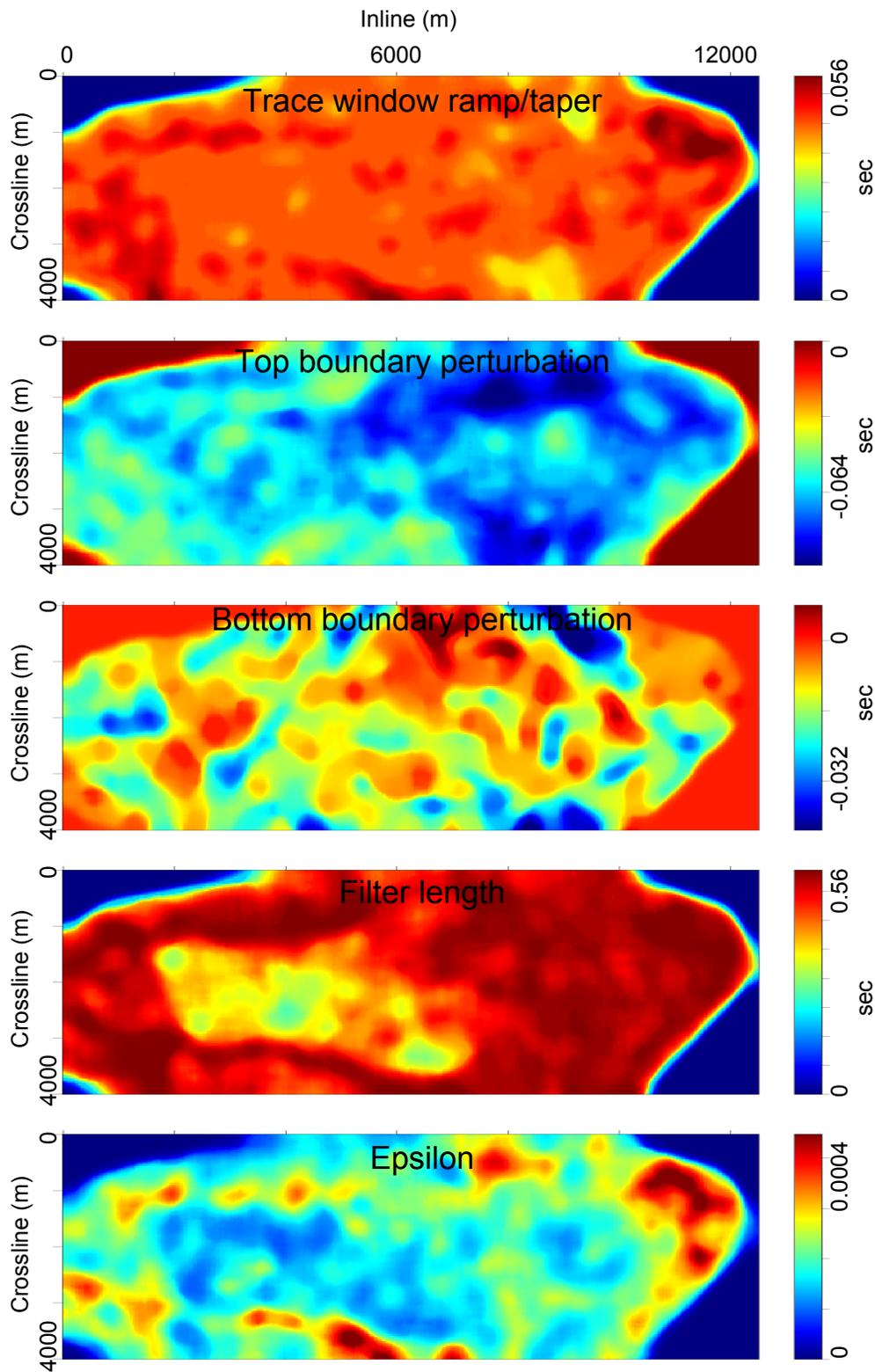


Figure 9: Optimal filter estimation parameters obtained by the evolutionary programming method. [CR]. [gayeni1/. optimal-1](#)

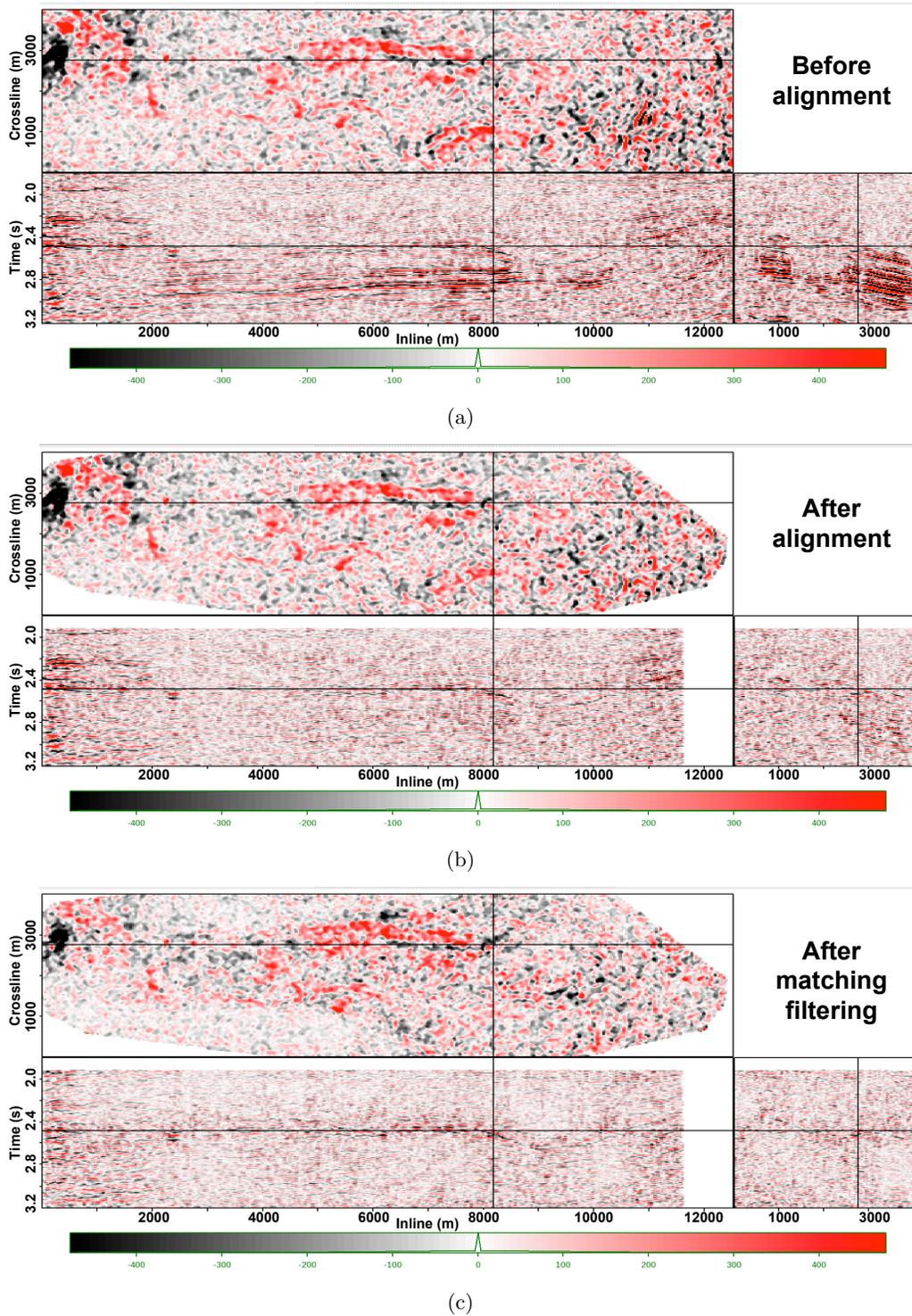


Figure 10: Time-lapse images after different processing steps. The top panel is a time-slice through the reservoir. Note the improvements in the quality of time-lapse amplitudes within the reservoir from top to bottom. [CR].
 gayenil/. amp-shift-2006,amp-shift-2006-shifted,amp-shift-2006-matched

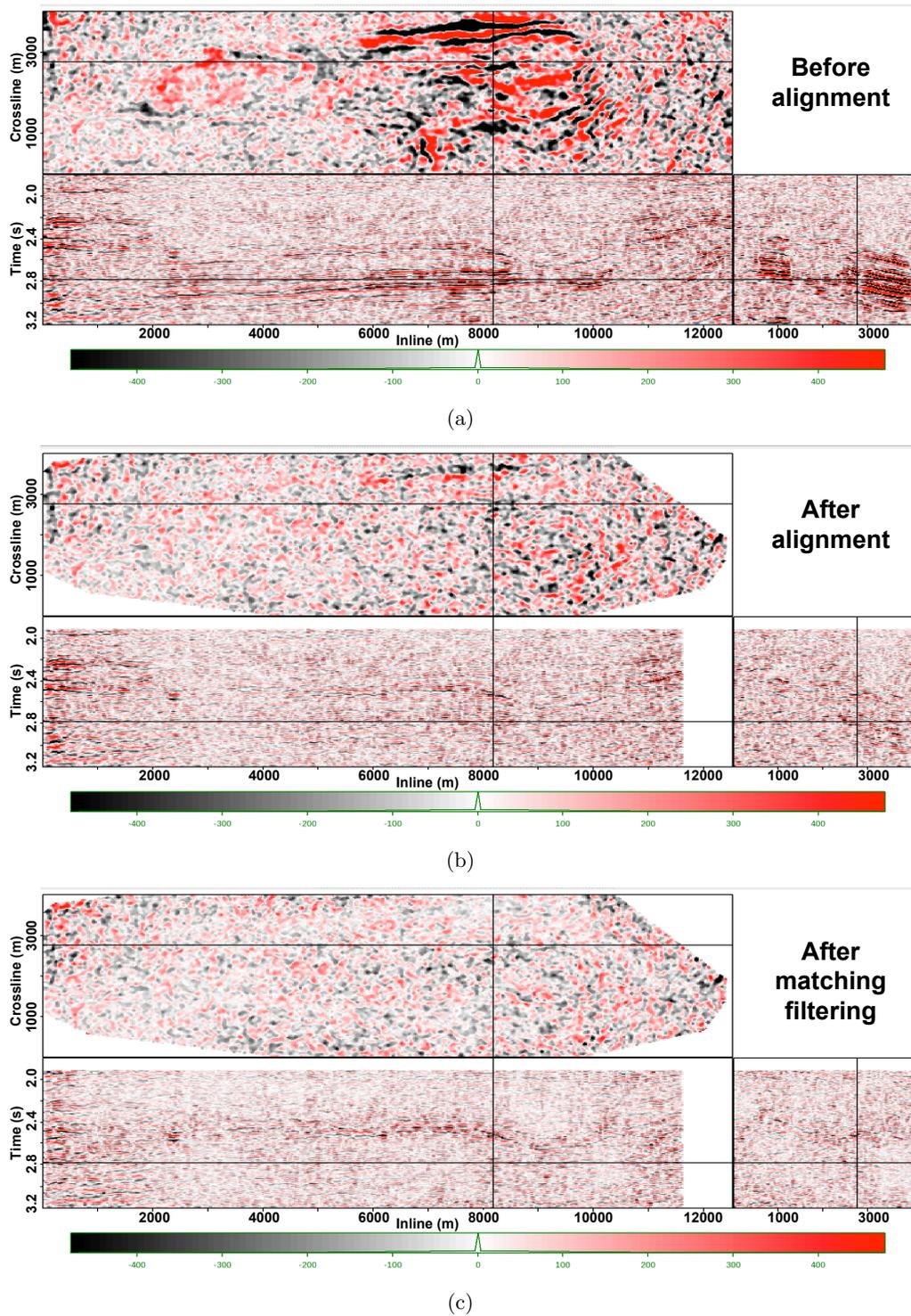


Figure 11: Time-lapse images after different processing steps. The top panel is a time-slice below the reservoir region where no time-lapse amplitude changes are expected. Note the improvement in the data quality from top to bottom. [CR].
 gayeni1/. amp-shift-2006-b,amp-shift-2006-b-shifted,amp-shift-2006-b-matched

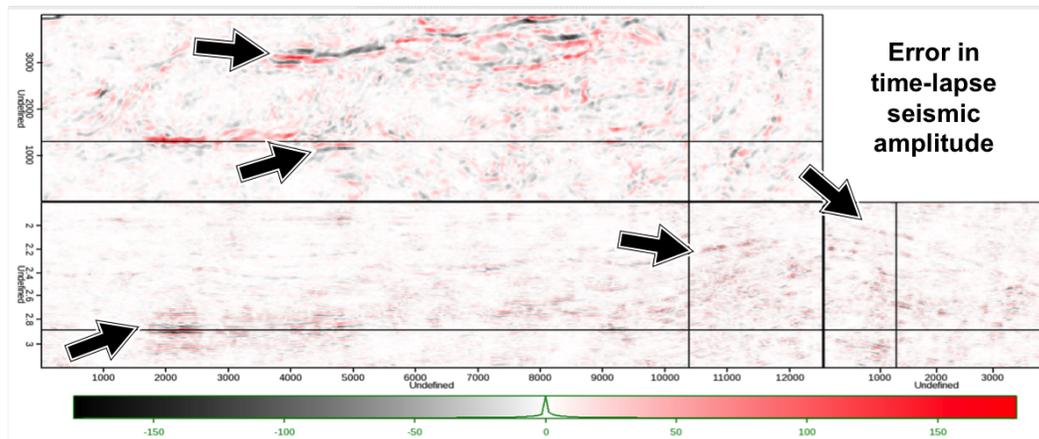


Figure 12: Errors in time-lapse amplitude computed as a difference between the time-lapse image obtained by considering only vertical displacements and from considering all displacement components. [CR]. [gayeni1/. amp-shift-2006-b-error](#)

- Bathija, A. P., M. L. Batzle, and M. Prasad, 2009, An experimental study of the dilation factor: *Geophysics*, **74**, E181–E191.
- Ebaid, H., M. Nasser, P. Hatchell, and D. Stanley, 2009, Time-lapse seismic makes a significant business impact at Holstein: SEG, *Expanded Abstracts*, **28**, 3810–3814.
- Hale, D., 2009, A method for estimating apparent displacement vectors from time-lapse seismic images: *Geophysics*, **74**, V99–V107.
- Hatchell, P., and S. Bourne, 2005a, Rocks under strain: Strain-induced time-lapse time shifts are observed for depleting reservoirs: *The Leading Edge*, **24**, 1222–1225.
- Hatchell, P. J., and S. J. Bourne, 2005b, Measuring reservoir compaction using time-lapse timeshifts: SEG Technical Program *Expanded Abstracts*, **24**, 2500–2503.
- Hawkins, K., S. Howe, S. Hollingworth, G. Conroy, L. Ben-Brahim, C. Tindle, N. Taylor, G. Joffroy, and A. Onaisi, 2007, Production-induced stresses from time-lapse time shifts: A geomechanics case study from franklin and elgin fields: *The Leading Edge*, **26**, 655–662.
- 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.
- Johnston, D. H., R. S. McKenny, J. Verbeek, and J. Almond, 1998, Time-lapse seismic analysis of fulmar field: *The Leading Edge*, **17**, 1420–1428.
- Osdal, B., O. Husby, H. A. Aronsen, N. Chen, and T. Alsos, 2006, Mapping the fluid front and pressure buildup using 4D data on norne field: *The Leading Edge*, **25**, 1134–1141.
- Rickett, J. E., and D. E. Lumley, 2001, Cross-equalization data processing for time-lapse seismic reservoir monitoring: A case study from the Gulf of Mexico: *Geophysics*, **66**, 1015–1025.
- Whitcombe, D. N., J. M. Marsh, P. J. Clifford, M. Dyce, C. J. S. McKenzie, S. Campbell, A. J. Hill, R. S. Parr, C. Pearse, T. A. Ricketts, C. P. Slater, and O. L. Barkved, 2004, The systematic application of 4D in BP's North-West Europe operations — 5 years on: SEG Technical Program *Expanded Abstracts*, **23**, 2251–2254.
- Zou, Y., L. R. Bentley, L. R. Lines, and D. Coombe, 2006, Integration of seismic methods with reservoir simulation, Pikes Peak heavy-oil field, Saskatchewan: *The Leading Edge*,

25, 764–781.

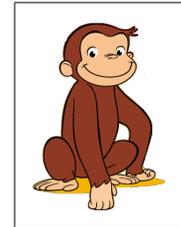
270

SEP-140

Sjoerd de Ridder received his B.Sc. (Dec. 2004) in Earth Sciences at Utrecht University, the Netherlands. He was a graduate exchange student at the Department of Geophysics, Colorado School of Mines, USA, in the fall of 2004 and the spring of 2005. He obtained an M.Sc. (Aug. 2007) at Delft University of Technology, the Netherlands, for his thesis 'Simulation of interferometric seismoelectric Green's function recovery' in Applied Earth Sciences with specialisation in applied geophysics. He joined SEP in the fall of 2007 and is currently pursuing a Ph.D. in Geophysics at Stanford University. Recipient of the "Jon F. Claerbout - Chevron Fellowship". Currently working on exploration seismology with low-frequency ambient seismic noise.



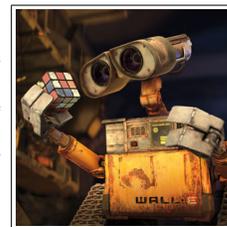
Xukai Shen graduated from Tsinghua University in July 2007 with a B.E. in Electrical Engineering. He joined SEP in the fall of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG.



Yaxun Tang received his B.Sc. (Jul. 2003) and M.Sc. (Aug. 2005) in Geophysics from School of Ocean and Earth Science, Tongji University, Shanghai. He joined SEP in 2005 and is currently working towards a Ph.D. in Geophysics at Stanford University. He is a student member of SEG and EAGE.



Kittinat Taweessintananon received his B.Sc. (March 2009) in Physics from Chulalongkorn University, Thailand. He joined SEP in 2009 and is working toward an M.S. in Geophysics. During the time at Stanford, he is granted by PTT Exploration and Production plc., which is the national petroleum exploration company in Thailand. He is a student member of SEG.



A new method for more efficient seismic image segmentation

Adam Halpert

ABSTRACT

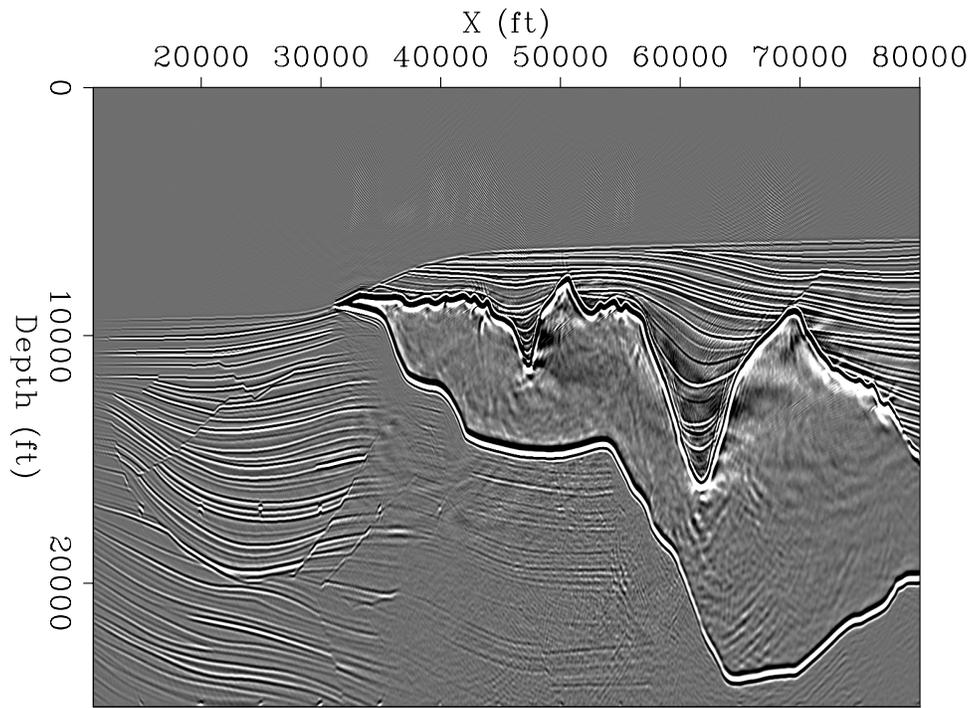
Seismic images of the subsurface are often very large and tedious to interpret manually; as such, automatic segmentation algorithms can be highly useful for tasks such as locating large, irregularly-shaped salt bodies within the images. However, seismic images present unique challenges for image segmentation algorithms. Here, a new segmentation algorithm using a “pairwise region comparison” strategy is implemented and tested on seismic images. Numerous modifications to the original algorithm are necessary to make it appropriate for use with seismic data, including: (1) changes to the nature of the input data, (2) the way in which the graph is constructed, and (3) the formula for calculating edge weights. Initial results, including a preliminary 3D implementation, indicate that the new method compares very favorably with an existing implementation of the eigenvector-based normalized cuts approach, both in terms of accuracy and efficiency.

INTRODUCTION

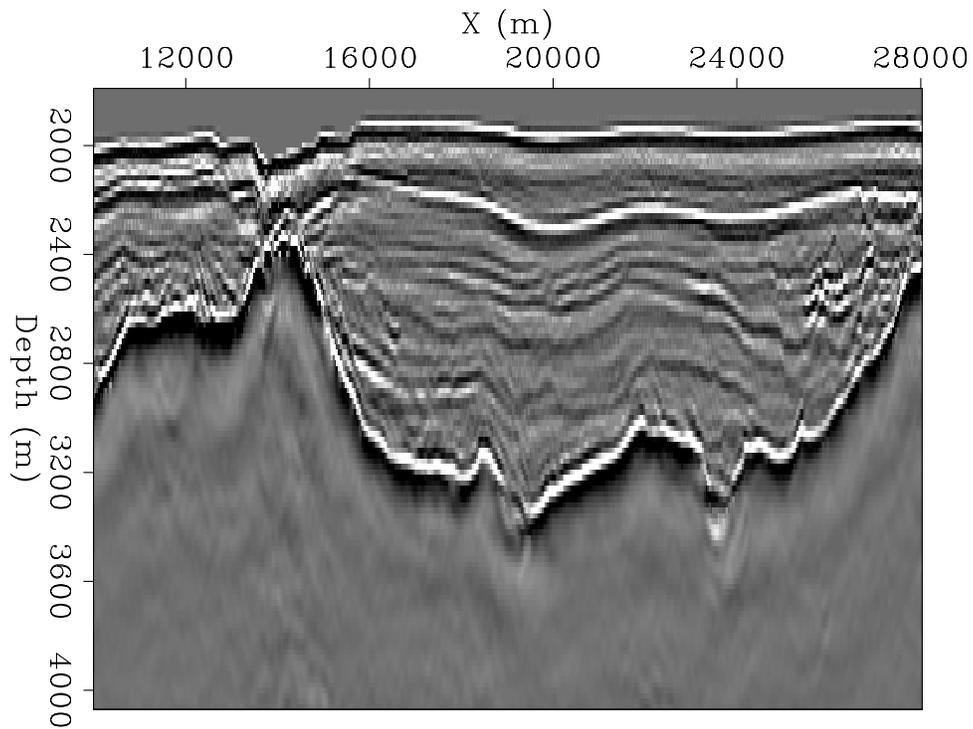
A proper salt interpretation is one necessary component of any imaging project where salt bodies play a prominent role in the subsurface geology. Because of the sharp velocity contrast between salt and nearly any other material, inaccurate placement of salt boundaries has a disproportionate effect on the accuracy of the resulting velocity model. Such errors can have damaging imaging and engineering consequences. Unfortunately, interpreting salt boundaries is not only crucial, but also extremely tedious and time-consuming when undertaken manually. Thus, while some degree of automation would be ideal for salt picking, any such method must be highly accurate as well as efficient.

One approach to implementing automatic salt-picking is to use graph-based image segmentation. In this method, each pixel in a seismic image is treated as a node or vertex in a graph; then edges are constructed between specific pixels and weighted according to some property. Image segments are created by partitioning the graph – here, a partition is a salt boundary.

In this paper, I will review a previous effort to apply a graph-based image segmentation algorithm to seismic data, and then introduce a new, more efficient method using an algorithm from Felzenszwalb and Huttenlocher (2004). Since seismic images are very different from more “conventional” images (such as photographs) for which this newer algorithm was designed, I will detail several modifications necessary for the algorithm to be useful for seismic images. Finally, I demonstrate the results of applying the new method to the examples seen in Figures 1(a) and 1(b), and (for the field data example) compare these results with those obtained using the previous algorithm.



(a)



(b)

Figure 1: A perfect-velocity migration of the Sigsbee synthetic model (a), and a field seismic image (b) that will be used as examples throughout this paper. [ER]

adam1/. zig-img,uno-img

BACKGROUND

As mentioned previously, graph-based segmentation methods are popular for automating salt interpretation. More specifically, the eigenvector-based Normalized Cuts Image Segmentation (NCIS) algorithm (Shi and Malik, 2000) has attracted a great deal of interest because of its capability to capture global aspects of the image, rather than just track local features. Local feature trackers are ubiquitous in seismic interpretation software (for example, tracking reflections between manually-picked “seed” points), but often struggle when they encounter situations such as a discontinuous boundary, or one which varies in intensity. Such situations are extremely common in seismic data, especially along the boundary between salt bodies and other types of rock. Recent research has involved implementing the NCIS algorithm for the purpose of tracking these salt boundaries [e.g., Lomask et al. (2007); Lomask (2007); Halpert et al. (2009)]. Results of this line of research are encouraging; however, there are significant limitations, most notably computational. The NCIS algorithm calls for an edge weight matrix of size n^2 , where n is the number of pixels in the image; this matrix quickly grows very large, especially for 3D surveys. Computationally, calculation of eigenvectors for such a large matrix is an extremely demanding task. As such, this method is limited to relatively small images; alternatively, we can restrict the computational domain to a specific region around a previously interpreted boundary. However, this means the method is of limited utility if there is no “best guess” model available, or if the accuracy of that model is in question.

Thus, a more efficient global segmentation scheme that can include the entire image in the computational domain would be a very useful tool for interpretation of seismic images. One candidate for such a scheme is the algorithm from Felzenszwalb and Huttenlocher (2004), who write:

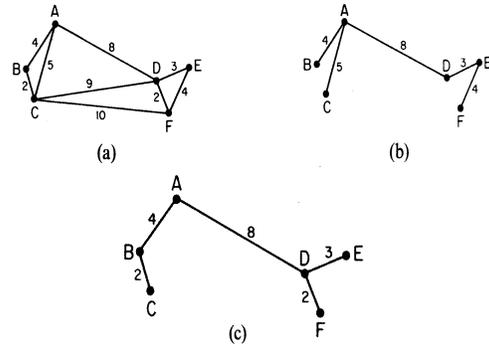
“Our algorithm is unique, in that it is both highly efficient and yet captures non-local properties of images.”

These two features are crucial for the task of seismic image segmentation. The algorithm is designed to run in $O(n \log n)$ time, where n is the number of pixels in the graph; in contrast, other methods such as NCIS require closer to $O(n^2)$ time to run. This represents a significant cost savings, especially for very large 3D seismic datasets that are becoming increasingly common.

The algorithm proposed by Felzenszwalb and Huttenlocher (2004) relies heavily on the concept of the “Minimum Spanning Tree” [see Zahn (1971)]. A graph’s edges may be weighted using a measure of dissimilarity between vertex pairs; a connected graph is defined as one in which all such edges are assigned a weight value. If a spanning tree is a connected graph which connects all vertices of the graph without forming a circuit, the minimum spanning tree (MST) of a graph is the spanning tree with the minimum sum of edge weights (see Figure 2). In Zahn (1971), partitioning of a graph was achieved simply by cutting through edges with large weights. However, this approach is inadequate for images with coherent regions that are nonetheless highly heterogeneous (for example, consider the heterogeneous nature of the intensity values within the salt bodies in the examples above). However, the MST concept allows Felzenszwalb and Huttenlocher (2004) to develop what they term a “pairwise region comparison” predicate. They define the *internal difference* of

Figure 2: Modified from Zahn (1971). A graph with weighted edges (a); a spanning tree of that graph (b); and the minimum spanning tree of the graph (c). [NR]

adam1/. MST



a region (C) in the graph to be the largest edge weight of the MST of that region:

$$Int(C) = \max_{e \in MST} w(e), \quad (1)$$

where e is a graph edge and $w(e)$ is the edge's weight, defined according to some simple algorithm. When comparing two regions (such as C_1 and C_2), they define the *minimum* internal difference for the two regions to be

$$MInt(C_1, C_2) = \min(Int(C_1) + \tau(C_1), Int(C_2) + \tau(C_2)), \quad (2)$$

where τ is a thresholding function that in a sense determines the scale at which the segmentation problem is approached, and thus indirectly the size of the regions in the final segmentation. Finally, they define the *difference* between the two regions to be the smallest edge weight that connects them:

$$Dif(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2} w((v_i, v_j)), \quad (3)$$

where v_i and v_j are vertices (or pixels) in the two different regions. When determining whether these two regions should be considered separate segments of the graph, or merged into a single region, they simply compare the values of $Dif(C_1, C_2)$ and $MInt(C_1, C_2)$. If $Dif(C_1, C_2)$ is greater, the “pairwise comparison predicate” is determined to be true, and the two regions are separated. While this is a relatively simple procedure, it is designed to allow highly heterogeneous regions to be segmented as a single component of an image. Additionally, Felzenszwalb and Huttenlocher (2004) note that their algorithm produces segmentations that are “neither too coarse nor too fine,” referring to the global capabilities of the segmentation process.

In the next section, I will provide more detail about the algorithm itself, and explain the modifications necessary to make it suitable for use with seismic images.

APPROACH

The goal of this paper is to apply the algorithm of Felzenszwalb and Huttenlocher (2004), introduced above, to seismic data. Publicly-available code from Felzenszwalb (2010) allows for relatively easy implementation for standard images; however, seismic images are very different from photographs or other types of images. The following sections describe the rationale and procedure for modifying and adding additional features to the algorithm in order to apply it to seismic data.

Data input and manipulation

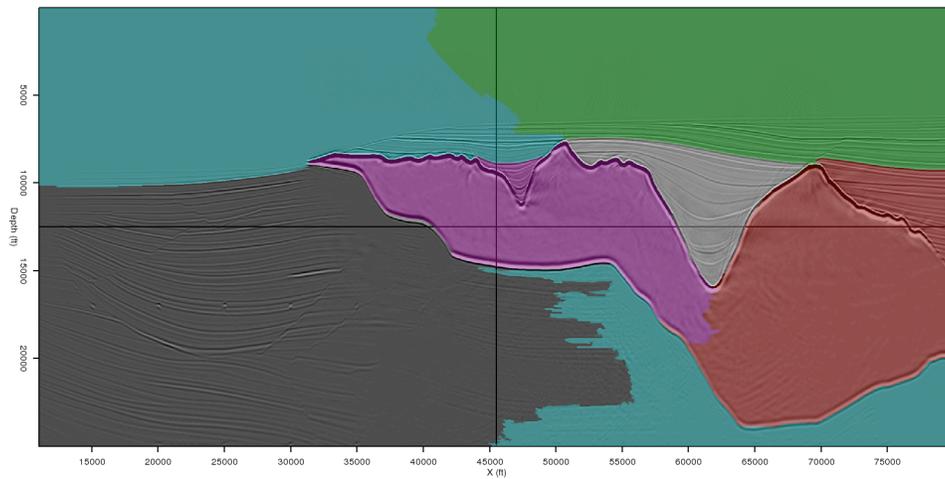
After modifying the original algorithm to work with seismic data rather than integer RGB values, seismic images may be segmented according to the same rules used by Felzenszwalb and Huttenlocher (2004) to segment RGB images. The results can be seen in Figure 3(a) for the synthetic seismic image, and Figure 3(b) for the field data image. In these figures, each segment or region is assigned a random color, and the segments are overlain on the seismic image itself for reference. The results from the (relatively) unaltered algorithm are promising, but require improvement. In Figure 3(a), the salt body is fairly well-resolved. However, it appears that the smaller canyon along the left of the salt body has been improperly included, along with another portion adjacent to the salt body at the far right of the image. The segmentation of the field data example in Figure 3(b) is especially poor; however, it is interesting to note that portions of the salt boundary itself have been recognized as segments. This may be related to the nature of the input data, an issue treated in the next section.

Transformation of input data

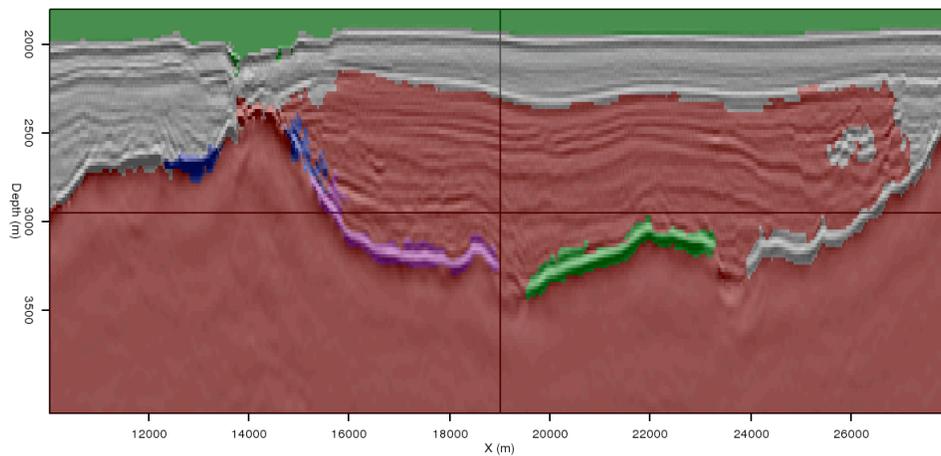
Seismic data may be thought of as signals with amplitude and phase varying as a function of time (or depth). This could present problems for any segmentation algorithm, and we may see an indication of this in Figure 3(b). At the boundary between the salt body and the surrounding rocks, the seismic waves change phase rapidly; this is common behavior when the waves encounter an interface and reflect back to the surface. As originally written, the algorithm may interpret the area around the boundary as several regions, instead of an interface between just two regions. In this case, the boundary itself becomes its own “region” in several locations. To avoid this situation, we would like the seismic image to be represented as amplitude information only, since this would indicate a single boundary between two regions. As Taner et al. (1979) point out, seismic data may be represented as a complex valued function:

$$A(z)e^{i\phi(z)}, \quad (4)$$

where z can be time or depth. The exponential term in this expression represents the phase information for the seismic data, while the leading term represents the amplitude information. By transforming the data such that the amplitude information is the only information present, the problem described above may be avoided. Figures 4(a) and 4(b) show the result of this process, also known as taking the “amplitude of the envelope” of the data, for the synthetic and field seismic images, respectively. We see that in both instances the phase information is no longer present, and the boundaries delineating the salt bodies



(a)



(b)

Figure 3: Segmentation of the example seismic images from Figure 1, using the original algorithm from Felzenszwalb and Huttenlocher (2004). [NR] `adam1/. zig-origseg,uno-origseg3`

are more clearly visible. By using these transformed images as inputs to the segmentation algorithm, it is likely that much of the unwanted behavior seen in the original examples can be avoided.

Creating the graph

The original implementation of the pairwise region comparison algorithm from Felzenszwalb (2010) creates a graph with eight edges per node (pixel). This graph is constructed by looping over every pixel, and performing four calculations at each vertex. The left side of Figure 5 illustrates this process – if the “active” pixel is the one in red, edges are drawn or built to each of the blue pixels. Since every pixel in the image undergoes this process, a form of reciprocity allows for each pixel to be connected to its eight immediate neighbors via edges. While this process allows for the extreme efficiency of the algorithm, the unique

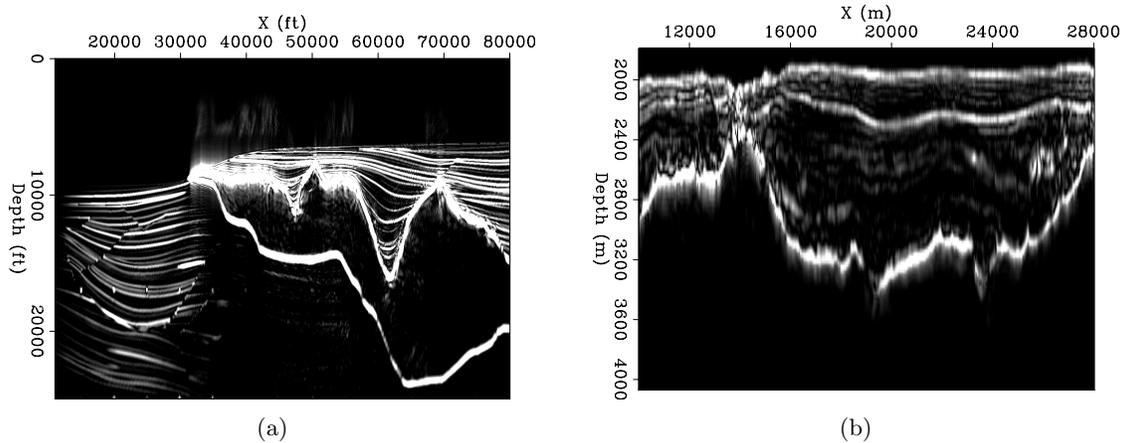
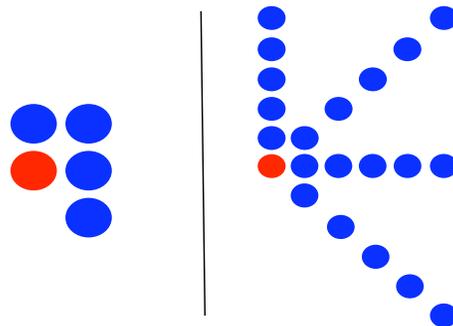


Figure 4: Result of calculating the amplitude envelope of the example images seen in Figure 1. These become the input to the new segmentation algorithm. [ER] `adam1/. zig-env,uno-env`

and often irregular nature of seismic data does not lend itself well to segmentations using so few edges per vertex or pixel. Instead, a much larger “stencil,” shown on the right of Figure 5, has been implemented. Rather than building edges that extend only one pixel in each direction, this stencil creates five edges extending in each horizontal, vertical and diagonal direction from the center pixel. This scheme allows for many more comparisons (40) per pixel, and a far greater amount of information goes into the segmentation algorithm. Near the boundaries of the image, the stencil shrinks to the largest size allowable by the image dimensions. While this approach obviously decreases the efficiency of the algorithm, the increased accuracy seen in the final results appears to make it a worthwhile trade-off. Even with the sharply increased number of edges per node, this algorithm is still far less computationally intensive than the NCIS algorithm from Shi and Malik (2000).

Figure 5: Stencils used for comparing pixel values and assigning edge weights for the graph. At left, the five-point stencil (8 edges per pixel) used in the original implementation from Felzenszwalb and Huttenlocher (2004); at right, a modified 21-point stencil (40 edges per pixel) used for the seismic images. [NR] `adam1/. stencils`



Calculating edge weights

The previous section details the construction of a graph used for segmentation of seismic images containing 40 edges per pixel. However, perhaps the most important part of any graph-based segmentation scheme is the calculation of weights for each of these edges. It is in

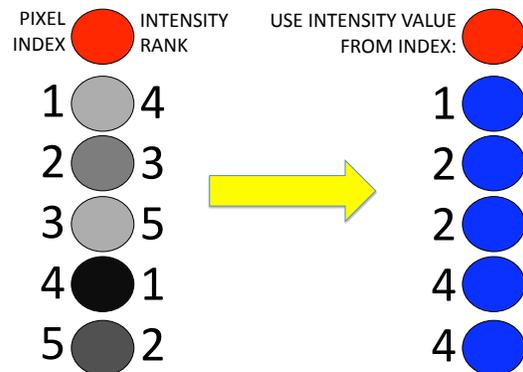
this aspect that the algorithm described here differs most from the original implementation provided by Felzenszwalb (2010).

When using the original stencil seen on the left in Figure 5, determining edge weights is relatively straightforward. Since each pair of vertices are at most one pixel apart in the image, we can simply compare the adjacent pixel values and use some expression to determine the likelihood that the two pixels reside in different regions or segments. This process is further simplified by the fact that the original implementation was designed to segment regions with coherent interiors; that is, even if the interior of a region is relatively chaotic, it is still distinct from the interiors of other regions of the image. However, this is not necessarily the case for seismic images – the interior of a salt body might not be distinct from other areas in the image that lack reflections. In this case a region is defined by its *boundary* rather than the character of its interior. Therefore, we must create a process for calculating edge weights that treats a boundary *between* two vertices as more convincing evidence for the existence of two regions than simply the difference in intensity at the two pixels themselves.

As described in the previous section, the stencil used for determining graph edges forms what are essentially four line “segments” at each vertex – one horizontal, one vertical, and two diagonal. Recall that once the entire graph is constructed, each vertex will actually be connected with eight such segments due to reciprocity of the calculations. Since two vertices that are the endpoints of a graph edge are no longer necessarily adjacent in the image, determining the existence of a boundary between them is no longer straightforward. However, the construction of distinct line segments extending from each pixel suggests one method of searching for a boundary: existence of a large amplitude value at any point along the line segment between the two pixels comprising the edge is evidence of a boundary. In other words, the intensity values of the two vertex pixels themselves will not be used to determine the edge weight; rather, the largest intensity value of any pixel along the line segment connecting the two endpoint vertices will be used. This strategy follows the approach of Lomask (2007) in his NCIS implementation. Figure 6 illustrates the logic behind this process. If all pixels in a segment are ranked according to intensity, the highest ranked pixel between the two edge vertices will be used for the weight calculation.

Figure 6: Diagram illustrating the logic behind deciding which pixel intensity value to use when calculating edge weights. Pixel intensities are shown and ranked on the left; the numbers in the right column indicate which intensity value will be used when calculating the edge weight between the pixel in red and the adjacent blue pixel. [NR]

adam1/. pixels



This process obviously involves some degree of algorithmic complexity, as it requires sorting and searching the pixel intensity values along each segment. Algorithm 5 illustrates the steps for carrying out the process shown graphically in Figure 6. After creating the edges

linking each pixel in a line segment to the “active” pixel, sort the line segment’s pixels in decreasing order of pixel intensity. Once this is done, compare the index value of the edge vertex pixel with the intensity-ranked list of pixel indices. To find the highest-intensity pixel value between the two vertices, simply take the value of the first pixel index on the sorted list that is less than the index of the vertex pixel.

Algorithm 5 Calculating graph edge weights

```

for each pixel ipix in image do
  create four line segments with five pixels per segment;
  record relative position (path.ind) and intensity (path.val) of each pixel;
  for each line segment do
    sort the segment in decreasing order of pixel intensity;
    for each pixel ix in the segment (nearest to furthest from pixel ipix) do
      for ij = 1..5 do
        if path[ij].ind <= ix then
          calculate edge weight using path[ij].val;
        end if
      end for
    end for
  end for

```

Once we have selected the intensity value to use for determining the edge weight, we must still calculate the weight value itself. The original implementation from Felzenszwalb (2010) used a simple Euclidean “distance” between adjacent pixel values. For RGB images, this expression is the equivalent of taking the square root of the sum of the squared differences for each of the three color components at each pixel. However, for the purposes of seismic image segmentation I have found it more appropriate to use an exponential function:

$$w_{ij} = \exp((\max I(\mathbf{p}_{ij}))^2), \quad (5)$$

where \mathbf{p}_{ij} is the vector of all pixels between i and j . Additionally, since the edges in the graph can now be much longer than with the adjacent-pixels-only approach taken in the original implementation, it makes sense to include a distance-weighting term to the edge weight calculation:

$$w_{ij} = \exp((\max I(\mathbf{p}_{ij}))^2) \exp(d_{ij}), \quad (6)$$

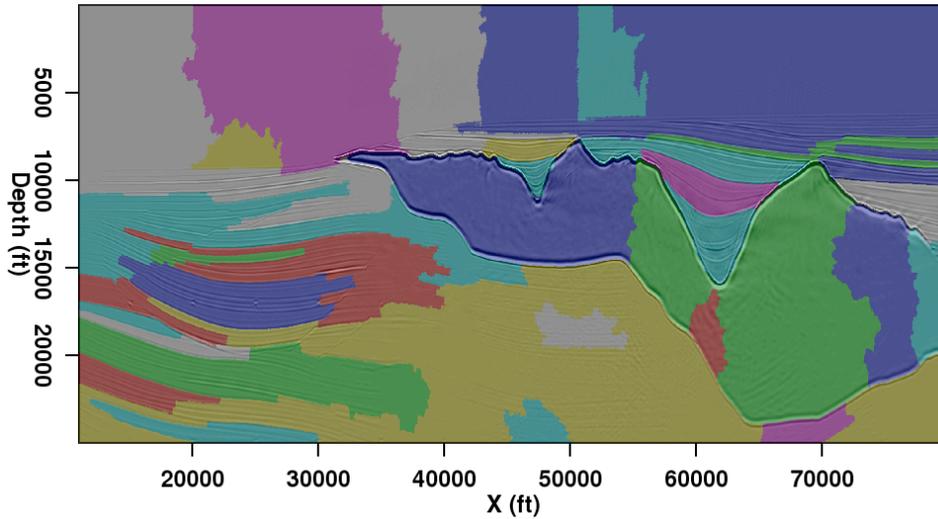
where d_{ij} is simply the Euclidean distance (in samples) between the two pixels.

Once each of the edges is assigned a weight, the segmentation of the image can proceed as described in Felzenszwalb and Huttenlocher (2004). In summary, the process begins with each pixel as its own image segment; then individual pixels, and eventually, groups of pixels, are merged according to the criteria set forth in section 2. Segments can also be merged in post-processing if they are smaller than a “minimum segment size” parameter specified by the user.

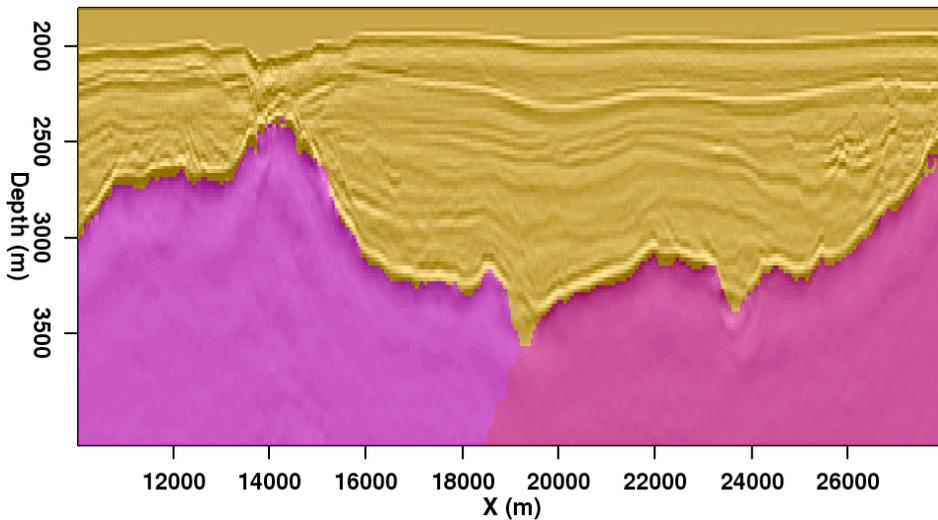
RESULTS

The results of the procedure detailed in the sections above can be seen in Figures 7(a) and 7(b). In Figure 7(a), although the salt body has been segmented into several regions, all of

these regions are contained within the salt body and fully conform to its boundaries. This represents an improvement over the original segmentation in Figure 3(a). The improvement is even more dramatic for the field data example (original image in Figure 1(b)). In Figure 7(b), we see that the salt body has been segmented virtually perfectly; no longer do the salt body segments spill over the boundary, nor are parts of the boundary itself treated as individual segments as we saw in the original example.



(a)



(b)

Figure 7: Results of applying the modified segmentation algorithm to the example images from Figure 1. Both (a) and (b) are a significant improvement over the original results in Figures 3(a) and 3(b). [ER] `adam1/. zig-segfinal2,uno-segfinal2`

A major motivation behind this research was to improve on segmentation results using the NCIS method from Shi and Malik (2000), and adapted for use with seismic images by Lomask et al. (2007). The first indication that the newer method does indeed represent a substantial improvement is the fact that the synthetic image from Figure 1(a) is simply *too*

large to be segmented on a single processor using the existing NCIS implementation. The necessity of holding a giant, sparse weight matrix in memory and calculating an eigenvector precludes problems of this size from being feasible. The smaller field data example, however, is well-suited for the NCIS algorithm, and will allow us to make a relatively fair comparison. Figure 8(a) shows the eigenvector produced by the NCIS method to segment the image. In this case, the graph partition will occur between the negative (red) and positive (blue) values of the eigenvector. The boundary resulting from this partition is drawn on the image in Figure 8(b). Recall that in order to make the problem more computationally feasible, the computational domain is already limited around a “prior guess” of the boundary. The pairwise region comparison method developed here requires no such limitation, which is another indication of its superiority.

Efficiency comparison

One of the primary means of comparison for the relative effectiveness of these two approaches to image segmentation is the computational efficiency of the method. The following table summarizes the computational expense required to create the examples seen in the paper.

Table 1: Comparison of CPU times for two segmentation methods

Image type	Pixels	CPU time (s)	
		NCIS	PRC
Synthetic data	2761000	n/a	31
Field data	55000	156	1

Again, due to memory constraints the existing NCIS implementation is unable to segment an image the size of Figure 1(a). The implementation described here, however, produces an accurate segmentation in 31 seconds; during this time, approximately 55 million edges are created, weighted, and used to segment the graph. The efficiency advantage for the new implementation is quantified using the field data example; in this case, the image is segmented over 150 times faster using the new implementation. These differences are extremely significant and represent a huge savings of time and computational expense, especially for larger problems.

Accuracy comparison

Of course, computational efficiency means little if the resulting segmentation is not accurate. For the synthetic image result in Figure 7(a), we can get a qualitative sense for the accuracy of the new segmentation implementation; while the salt body is divided into multiple segments, these segments do indeed fit almost exactly inside the boundaries of the salt body. A more direct comparison of the relative accuracy of the NCIS and PRC methods can be obtained via analysis of the field data results. Figure 9 shows both calculated salt boundaries overlain on the image: the NCIS boundary in green, and the PRC boundary in pink. Visually, we can see very little difference between these two results; in many locations, they are almost exactly on top of one another. The most noticeable difference between the

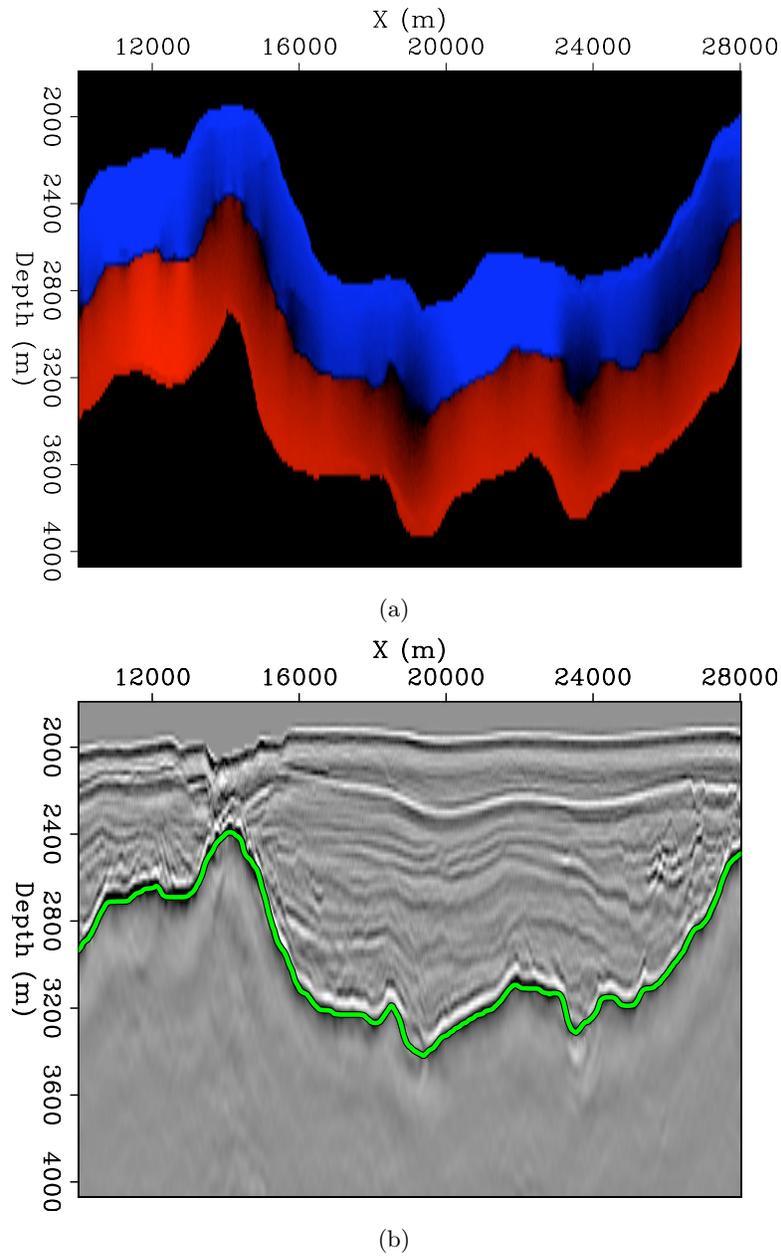


Figure 8: Eigenvector (a) used to segment the image in Figure 1(b) according to the NCIS algorithm of Shi and Malik (2000) and adapted for seismic data by Lomask et al. (2007), and the resulting salt boundary (b). [CR] `adam1/. uno-eig,uno-segeig`

two results is near $X = 20000m$, where the PRC boundary dips deeper than the NCIS boundary. Examination of the input image in Figure 1(b) suggests that in this location, an error in the migration velocity model has led to a discontinuity in the boundary image. The new method appears to do a better job of “correcting” the error. This result serves to increase confidence in the viability of this new segmentation scheme as an alternative to the existing NCIS implementation.

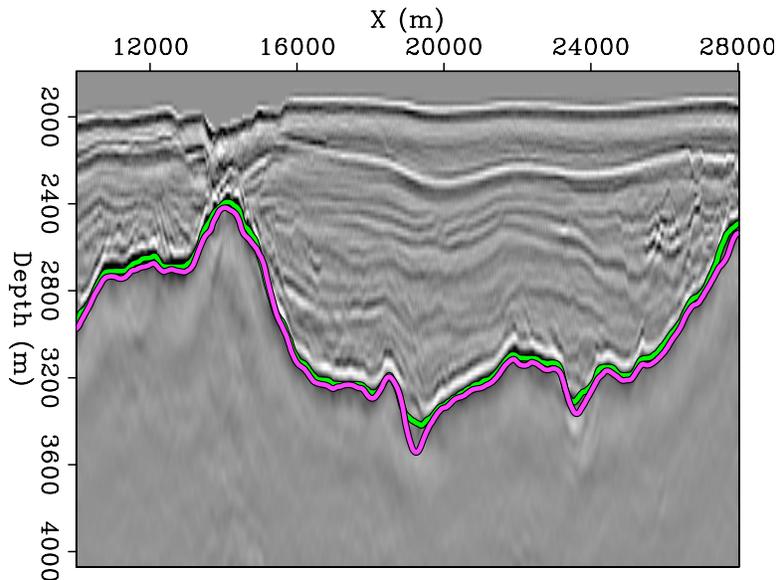
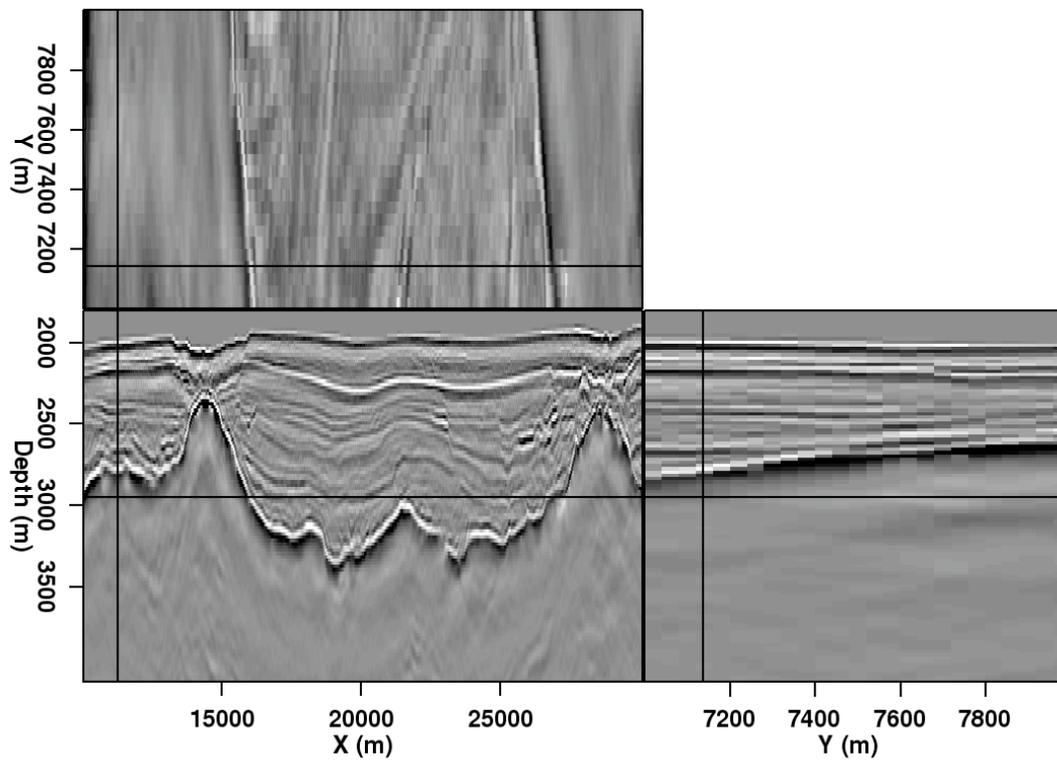


Figure 9: Comparison of the boundaries obtained using the NCIS eigenvector method (green) and the pairwise region comparison method (pink). [CR] [adam1/. uno-segcomp](#)

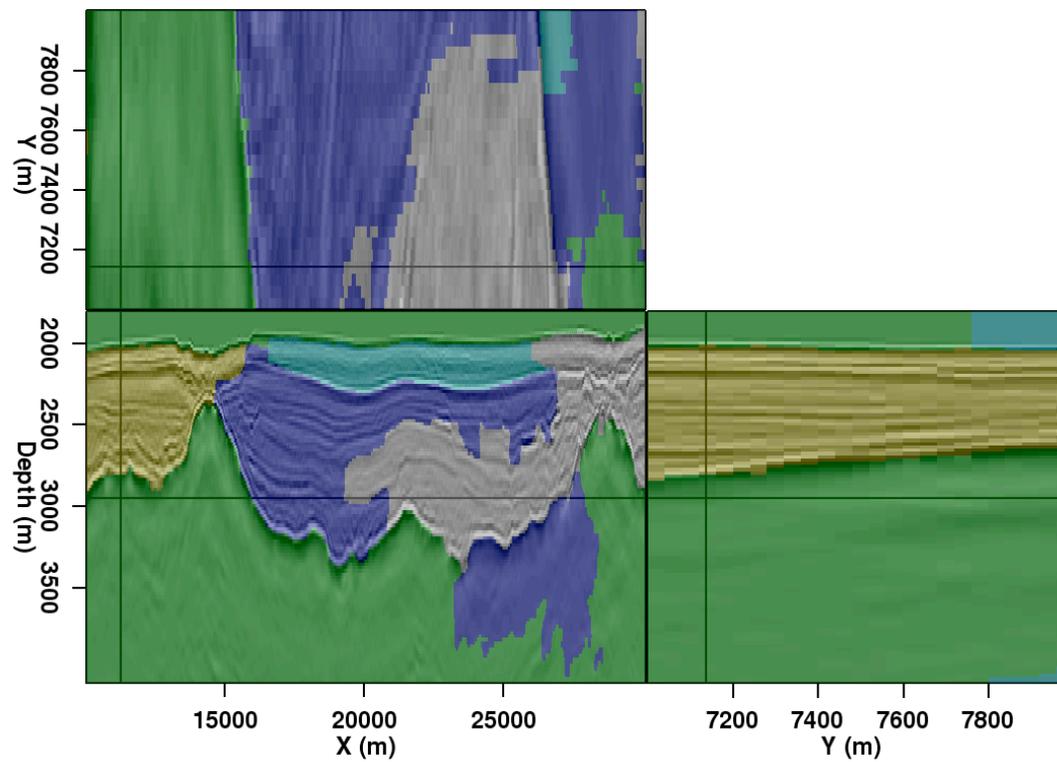
3D implementation

The pairwise region comparison strategy described here is easily extendable to three dimensions. While the theory and general approach remain the same as with the two-dimensional case, in three dimensions the stencil used to construct the graph must change. Along with the four co-planar segments extending from each pixel used for the second stencil in Figure 5, a 3D stencil must incorporate four additional segments that extend the graph edges into the extra dimension. Fortunately, the additional information this brings to the segmentation algorithm allows for the stencil’s segments to shrink in length; initial tests indicate that the most accurate results are obtained with segments that are three pixels in length. Thus, the computational impact of extending any algorithm from two to three dimensions is somewhat mitigated in this case.

Figures 10(a) and 10(b) show a segmentation example resulting from a preliminary 3D implementation of the new algorithm. While further improvements seem necessary, note that both the salt body (in multiple segments) and the water column have been accurately identified.



(a)



(b)

Figure 10: Slices through a 3D image (a) and the resulting 3D segmentation (b) using the new PRC algorithm. [ER] `adam1/. 3dimg,3dseg`

CONCLUSIONS

This paper presented an implementation of the pairwise region comparison (PRC) scheme of Felzenszwalb and Huttenlocher (2004) for segmenting seismic images. Numerous modifications were made to the original algorithm, including structural changes to allow for seismic images as inputs, a change in the way edges are constructed for the graph, and a change in the weighting calculation for each edge. Each of these modifications increased the accuracy of the method when applied to seismic data.

Initial results from applying the modified algorithm to both synthetic and field seismic images are extremely encouraging. Segmentation of a synthetic image accurately located the boundaries of a complex salt body, although several different segments were required. Segmentation of both 2D and 3D field seismic data was even more successful. Compared to an existing implementation of the Normalized Cuts algorithm from Shi and Malik (2000), the new method performs extremely well – it required only half a minute to segment the synthetic data image (which is too large for the NCIS implementation to handle on a single processor), and only one second to segment the smaller field data example, over 150 times faster than NCIS. An additional advantage is that the newer algorithm is able to operate on the entire image, rather than only within a certain windowed radius of a previously interpreted boundary. This approach has many advantages, not least of which is the opportunity to identify segments other than only salt bodies. Instead of a binary salt/no-salt determination, the ability to identify coherent sedimentary “segments” as well would be tremendously useful for constructing seismic velocity models.

While these results are promising, there are many potential improvements that remain to be explored. First, the fact that the salt body is in some cases divided into multiple segments is a situation that should be examined. One possibility is to modify the edge weight calculation to include the relative importance of the amplitude/intensity and distance factors; right now, both terms are weighted equally (see equation 6). Another option is to change either the shape or size of the stencil (Figure 5) used to build the graph. Yet another potential improvement is the incorporation of multiple seismic attributes – for example, dip and frequency attributes. This enhancement can be accomplished relatively simply by taking multiple volumes as inputs, and calculating edge weights using some weighted combination of the different attributes at each pixel.

Interpreting subsurface salt bodies in large, 3D seismic images is an incredibly complex and tedious task. With further improvement, however, an accurate, *efficient* automatic segmentation scheme such as this one has the potential to be an extremely useful and powerful tool for processing and interpreting seismic images.

ACKNOWLEDGMENTS

I thank SMAART JV and Unocal (now Chevron) for providing the data used for examples in this paper, and Robert Clapp for many helpful suggestions.

REFERENCES

- Felzenszwalb, P. F., 2010, Image segmentation. (<http://people.cs.uchicago.edu/~pff/segment/>).
- Felzenszwalb, P. F. and D. P. Huttenlocher, 2004, Efficient graph-based image segmentation: *International Journal of Computer Vision*, **59**, 167–181.
- Halpert, A., R. G. Clapp, and B. L. Biondi, 2009, Seismic image segmentation with multiple attributes: *SEG Technical Program Expanded Abstracts*, **28**, 3700–3704.
- Lomask, J., 2007, Seismic volumetric flattening and segmentation: PhD thesis, Stanford University.
- Lomask, J., R. G. Clapp, and B. Biondi, 2007, Application of image segmentation to tracking 3d salt boundaries: *Geophysics*, **72**, P47–P56.
- Shi, J. and J. Malik, 2000, Normalized cuts and image segmentation: *Institute of Electrical and Electronics Engineers Transactions on Pattern Analysis and Machine Intelligence*, **22**, 838–905.
- Taner, M. T., F. Koehler, and R. E. Sheriff, 1979, Complex seismic trace analysis: *Geophysics*, **44**, 1041–1063.
- Zahn, C. T., 1971, Graph-theoretical methods for detecting and describing gestalt clusters: *IEEE Transactions on Computers*, **C-20**, 68–86.

Hypercube viewer update

Robert G. Clapp

ABSTRACT

Efficient viewing and interacting with multi-dimensional data volumes is an essential part of many scientific fields. This interaction ranges from simple visualization to steering computationally demanding tasks. SEP uses a multi-dimensional slice viewer called **Hypercube**. I describe improvements to the **Hypercube** viewer including rotation, 3-D picking, and the ability to create surfaces.

INTRODUCTION

Viewing and interacting with multi-dimensional volumes is necessary when working with 3-D data. SEP wrote its first movie program 30 years ago and has continually expanded on this initial idea (Claerbout, 1981; Sword, 1981; Ottolini, 1982, 1983, 1988, 1990). These movie programs have progressed from simply showing a series of frames to allowing greater and greater levels of interactivity. In Clapp et al. (2008), a new viewer, **Hypercube**, based on the QT * graphics library was introduced. In this paper I describe improvements to the **Hypercube** viewer. I begin by giving an overview of the viewer and then I describe improvements in its ability to do data auto-picking, rotate data, and create surfaces from picks.

UPDATES

A great number of additional features and bug fixes have been introduced since Clapp et al. (2008). In this paper I am highlighting a few of the more useful features.

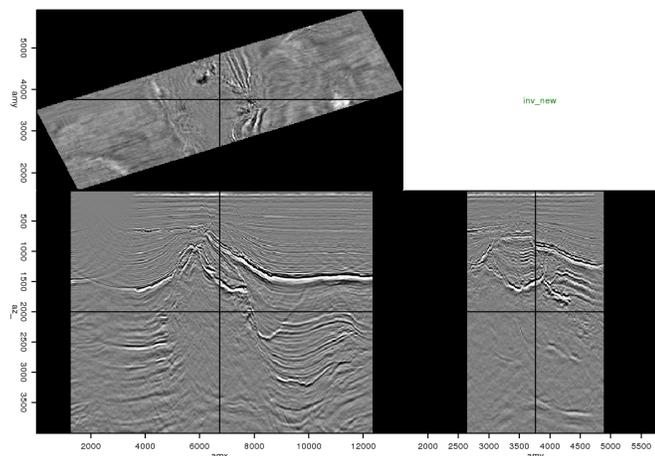
Rotation

Hypercube is based on a regular grid concept. This has the advantage of allowing for simple book keeping, lower graphic card requirements, and makes it more amenable to interaction with applications/approaches that rely on regular sampled functions. It has the disadvantage of not allowing data that does not fit either its grid or line (through pick-sets) data concept. **Hypercube** now allows the grid to be rotated along any plane. The rotation amounts to rotating the gridding of the data. The e basic slice viewer concept is maintained. As a result, if you rotated your view by 45 degrees you would see a single trace at the edge of the cube. Figure 1 demonstrates this concept.

*<http://www.trolltech.com/products/qt>

Figure 1: The result of rotating the cube viewing angle. [NR]

`bob1/. rotate`



Auto-picking

Picking every X-Y location is too human-intensive to be practical. As a result, auto-pickers are used to speed up what is still a human-intensive task. **Hypercube** has two 2-D auto picking options. Both of these options use a dynamic programming Liner and Clapp (2004) approach to find the most likely path between a set of user picked points. The difference between the two approach is how the dynamic programming *score matrix* is constructed.

The user begins by selected several points along a given interface (Figure 2(a)). **Hypercube** has the concept of a *single* axis for each pick set. The best way to understand the **single** axis is through a couple examples. When picking NMO velocities you do not want more than one velocity at a given time and midpoint. As a result, the velocity axis would be the *single* axis. When picking surfaces, **Hypercube** forces you to pick an axis (such as depth) where you will not have multiple depths at any given location. The definition of a *single*

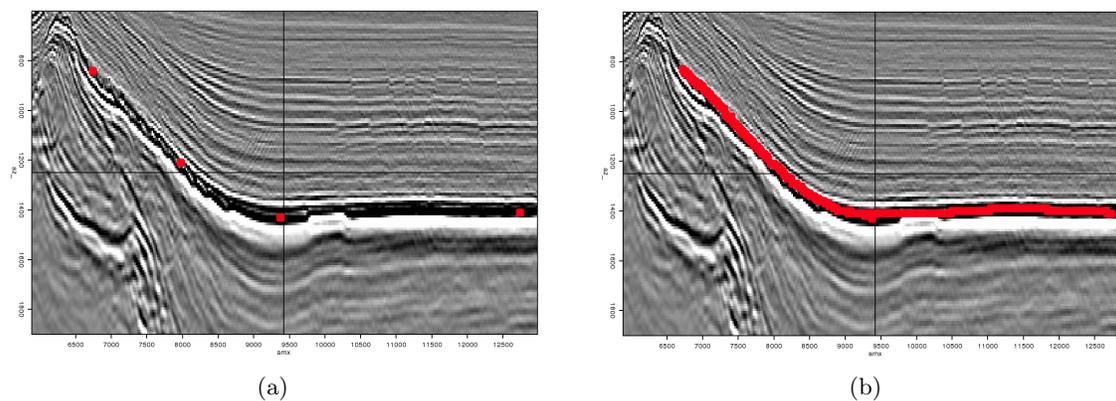


Figure 2: a) Several picks on an inline section. b) The result of using the modified Brown auto-picker on the picks shown in a). [NR] `bob1/. pick,auto`

axis allows the viewer viewer to construct a linear path between pairs of user specified points (the linear path is along the non-single axis). Points are then extracted by taking points that are to either side (along the *single* axis) of the line segments making a 2-D matrix.

The 2-D matrix is as wide as the distance between the first and last picked points along the non-*single* axis and is as high as $2nc + 1$, where nc is the user specified number of points to either of the line segment that are extracted. At this stage the two auto-picking approaches (correlation Viterbi and Brown) diverge. In the correlation approach, the score matrix is constructed by cross correlating a vector at a picked location with the vector at the test location. The Brown method (Brown et al., 2006) substitutes amplitude for velocity and uses an Eikonal solver to find the best path between picked points. Figure 2(b) demonstrates the result of running the Brown auto-picker on the picks shown in Figure 2(a).

The auto-picker provides some level of 3-D picking through extending a series of 2-D picked lines. Figure 3 show the result of auto-picking three in-lines. The auto picker will loop through the planes perpendicular to picked lines using the specified auto-picking methods to create a dense pick-set. The left panel of Figure 4 shows the result of extending a series of picked lines.

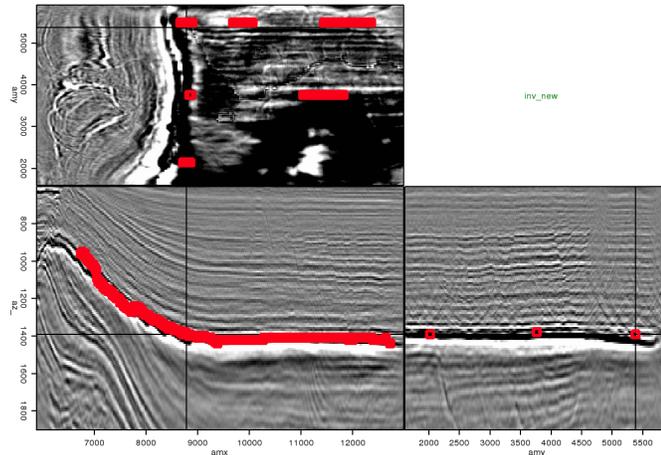


Figure 3: The result of picking two additional inline using the auto-pickers. [NR] `bob1/. multi`

Surfaces

Surface viewing is a new option available in `Hypercube` by specifying `do_surface=1` on the command line. A `Hypercube` surface is a new dataset created from a pick-set. It is one dimension less in size than the original data (based on the *single* axis). The values of the dataset are either the location of a pick along the *single* axis or the amplitude at that location. Figure 4 shows a depth map from a set of auto-picked points. Note how we have depth information only at X-Y locations where we have a pick.

WEBSITE

A technical report is not the best medium to describe how to use an interactive tool. As a result I have set up a website[†] that describes how to install and use the `hypercube` viewer. It contains a complete description of the menu, keyboard, and mouse options available in `hypercube`. It also contains several movies demonstrating how to use a wide variety of `hypercube`'s features.

[†]<http://sep.stanford.edu/data/media/public/sep/bob/doku.php?id=hypercube:hypercube>

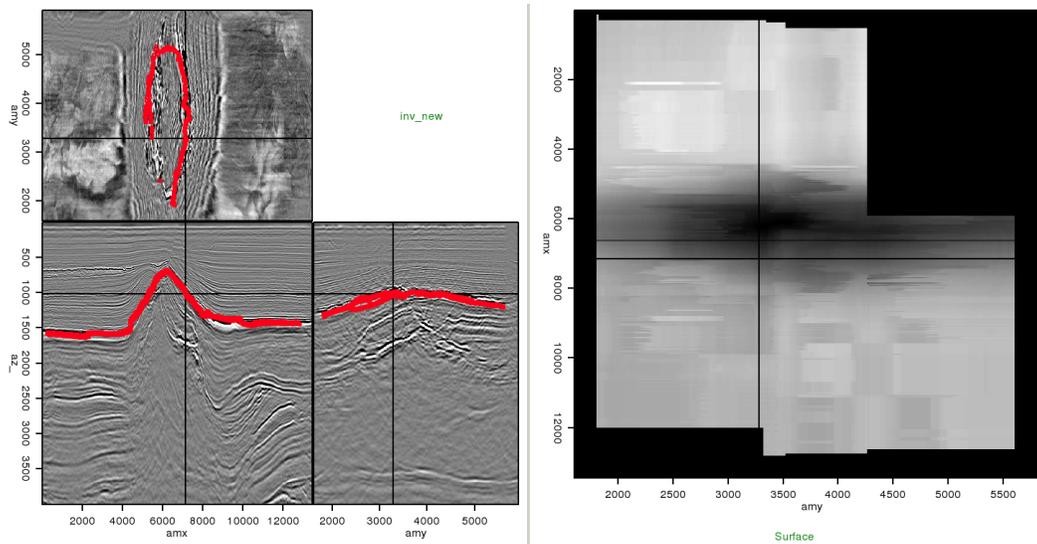


Figure 4: The left panel shows a dense set of picks based on extending three user picked in-lines. The right panel shows the depth of each picked point. [NR] bob1/. surface

CONCLUSIONS

SEP's hypercube viewer is a powerful tool to visualize regularly sampled fields. The viewer now allows rotation and expanded picking features. A website has been created that describes the viewer in more detail.

ACKNOWLEDGEMENTS

I would like to than TotalFinaElf for the data displayed in this paper.

REFERENCES

- Brown, M. P., S. A. Morton, and G. Whittle, 2006, Seismic event tracking by global path optimization: SEG, Expanded Abstracts, **25**, 1063–1067.
- Claerbout, J. F., 1981, On-line movies: SEP-Report, **28**, 1–4.
- Clapp, R. G., D. M. Chen, and S. Luo, 2008, Hypercube viewer: SEP-Report, **134**, 179–192.
- Liner, C. L. and R. G. Clapp, 2004, Nonlinear pairwise alignment of seismic traces: Geophysics, **69**, 1552–1559.
- Ottolini, R., 1982, Interactive movie machine user's documentation: SEP-Report, **32**, 183–196.
- , 1983, Movie cubes: SEP-Report, **35**, 235–240.
- , 1988, Movies on the Macintosh II: SEP-Report, **59**, 255–268.
- , 1990, Seismic movies on the XView graphics system: SEP-Report, **65**, 315.
- Sword, C. H., 1981, SEP goes to the movies: SEP-Report, **28**, 11–20.

Short note: Enhanced visualization for seismic imaging

Nader Moussa

ABSTRACT

I have developed a new visualization package for seismic images. SEP's existing visualization toolkit, including Grey, Graph, Wiggle, Contour, and Cube, will be integrated into a novel, unified, open-source tool. Enhanced features include interactive manipulation of graphics, efficient file-access for large data sets, network-streaming for large files, 3D acceleration options, and platform portability.

This new suite of technologies will be capable of tight integration with processing workflows, enabling interactive processing. At the same time, it preserves the best features of our legacy tools, and provides a compatibility layer for seamless integration into existing environments. JTube charges forward with new features, and capitalizes on a quarter-century of technological advances since the last major design of SEP's visualization tools.

INTRODUCTION

Stanford Exploration Project is an academic consortium whose primary purpose is to improve the theory and practice of constructing 3-D and 4-D images of the Earth from seismic echo soundings. In Clapp et al. (2008), a new viewer, **Hypercube**, based on the QT graphics library was introduced. Existing SEP plotting and graphing utilities, while sufficient for basic visualization, have not kept pace with the rapidly changing technological landscape of modern computer hardware, operating systems, and commercial visualization tools.

In response to this issue, I have completely redesigned the existing SEP visualization utilities: Graph, Grey, Wiggle, Contour, Cubeplot, and the Tube utility. The new software is designed to be modular, object-oriented, and extensible. The resulting utility is a tightly-integrated, well-managed, and efficient suite of seismic visualization tools that enable more advanced image creation.

Practical details of the new implementation address common usability issues that have plagued SEPlib users, both within the research group and among our industrial sponsor user-base. Behind the scenes, the code is designed to be object-oriented and easily extended if new plotting features are required. The implementation takes advantage of the Java platform, allowing portable and efficient graphical interfaces.

The new project has tentatively been named JTube, wrapping all visualization utilities into a single front-end interface. Individual plot tools are accessible from the command line or via an interactive interface.

REQUIREMENTS

The JTube software requires a Java 6 virtual machine to run. To build the project, a Java 6 software development kit (SDK) is required, and the Apache Ant build tool is recommended (but not required). The system will compile with any Java 6 SDK, and has been tested with Sun, OpenJDK, and GCJ.

BUILDING AND RUNNING

The command to build the distribution package is `ant dist`. This will compile all source code, package it as an executable archive, and copy the entire output and utility scripts into the distribution output directory.

To run the program, change directory to the distribution directory, or add this directory to the path. The command `JTube` will start the program. A complete list of command-line arguments can be found in the extended documentation. This documentation can be viewed with any text editor or web browser, and will be automatically generated during the build. It can also be manually built using the command `ant doc`.

Future revisions of JTube will be available through SEP or at

<http://www.nimur.com/projects/JTube>.

REFERENCES

Clapp, R. G., D. M. Chen, and S. Luo, 2008, Hypercube viewer: SEP-Report, **134**, 179–192.

Mechanics of stratified anisotropic poroelastic media

James G. Berryman

ABSTRACT

The mechanics of vertically stratified porous media has some similarities to and some differences from the more typical layered analysis for purely elastic media. Assuming welded solid contact at the solid-solid interfaces implies the usual continuity conditions, which are continuity of the horizontal strain components and the vertical stress components. These conditions are valid for both elastic and poroelastic media. Differences arise through the conditions for the pore pressure and the increment of fluid content in the context of fluid-saturated porous media. The two distinct conditions most typically considered between any pair of contiguous layers are: (1) an undrained fluid condition at the interface, meaning that the increment of fluid content is zero (*i.e.*, $\delta\zeta = 0$), or (2) fluid pressure continuity at the interface, implying that the change in fluid pressure is zero across the interface (*i.e.*, $\delta p_f = 0$). Depending on the types of measurements being made on the system and the pertinent boundary conditions for these measurements, either (or neither) of these two conditions might be directly pertinent. But these conditions are sufficient nevertheless to be used as thought experiments to determine the expected values of all the poroelastic coefficients. For quasi-static mechanical changes over long time periods, drained conditions hold, so the pressure must then be continuous. For high frequency wave propagation, the fluid typically acts essentially as if it were undrained – or nearly so, with vanishing of the fluid increment at the boundaries being appropriate. The poroelastic analysis of both these end-member cases is treated in detail.

INTRODUCTION

Traditional poroelastic analysis (Gassmann, 1951; Biot and Willis, 1957; Biot, 1962; Brown and Korrington, 1975; Rice and Cleary, 1976; Thigpen and Berryman, 1985; Zimmerman, 1991; Cheng, 1997; Wang, 2000; Pride et al., 2002) usually progresses from assumed knowledge of dry or drained porous media to the predicted behavior of fluid-saturated and undrained porous media. This class of problems is characterized by a single upscaling step, taking the homogeneous fluid and solid constituent properties and deducing the macroscopic behavior of such systems. In recent work (Berryman, 2010), the author has shown in detail how the poroelastic coefficients are related to the microstructural constants of the solid constituents when the overall behavior varies from isotropic to orthotropic. The focus of the present work is on stratified (*i.e.*, layered) poroelastic materials, which are therefore heterogeneous at the mesoscale but fairly homogeneous within each layer. In particular, individual layers are assumed to satisfy the same assumptions as the class of problems considered by Berryman (2010), which is basically limited to orthotropic poroelastic media with a known set of symmetry axes.

The main issue addressed here concerns how the interface boundary conditions between

anisotropic porous layers should be treated. For very low frequency (say quasi-static) analysis, this issue is clear since then the boundary conditions must be drained conditions and therefore the fluid pressure is continuous across the boundary. However, for high frequency wave propagation, it is expected to be more appropriate to treat the system as locally undrained, since the pressure of the pore-fluid does not have time to equilibrate via the drainage mechanism, which can take much longer than is appropriate to these quasi-static analyses. The most accurate way to treat these situations is to consider the variables to be frequency dependent and complex. This approach has been taken for example by Pride and Berryman (2003a,b); Pride et al. (2004) for mixtures of isotropic poroelastic materials. But the problem becomes harder for the anisotropic case, as there were simple exact results for the two-isotropic-component case, but simple results are not available for the anisotropic problems. And more importantly, the interest in layered media is not just for two-component examples, but ultimately for multi-component layered media. So it is important to consider these cases separately, as is being done here.

The analysis is restricted to anisotropic systems. The nature of the grains themselves composing the solid frame material will not be a focus of the present paper. This issue does matter, but it is most important for determining the relationship between the grain constants and the off-diagonal coefficients that are called the β 's in this formulation. These issues have been fully addressed in the earlier contribution of the author (Berryman, 2010), and will therefore not be treated again in this paper. Our focus here is on heterogeneous poroelastic media when the heterogeneity is well-represented via layered porous-medium modeling.

BASICS OF ANISOTROPIC POROELASTICITY

Orthotropic poroelasticity

If the overall porous medium is anisotropic either due to some preferential alignment of the constituent particles or due to externally imposed stress (such as a gravity field and weight of overburden, for example), then consider the orthorhombic anisotropic version of the poroelastic equations:

$$\begin{pmatrix} e_{11} \\ e_{22} \\ e_{33} \\ -\zeta \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{13} & -\beta_1 \\ s_{12} & s_{22} & s_{23} & -\beta_2 \\ s_{13} & s_{23} & s_{33} & -\beta_3 \\ -\beta_1 & -\beta_2 & -\beta_3 & \gamma \end{pmatrix} \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ -p_f \end{pmatrix}. \quad (1)$$

Throughout most of the paper, I will not introduce δ 's preceding the stresses and strains, as is sometimes done to emphasize their smallness, since this extra notation is truly redundant when they are all being treated as quantities pertinent to seismic wave propagation (and therefore resulting in linear effects) as I do here, for very small deviations from an initial rest state.

The e_{ii} (no summation over repeated indices) are strains in the $i = 1, 2, 3$ directions. The σ_{ii} are the corresponding stresses, assumed to be positive in tension. The fluid pressure is p_f , which is positive in compression. The increment of fluid content is ζ , and is often defined via:

$$\zeta \equiv \frac{\delta(\phi V) - \delta V_f}{V} \simeq \phi \left(\frac{\delta V_\phi}{V_\phi} - \frac{\delta V_f}{V_f} \right), \quad (2)$$

where $V = V_\phi/\phi \simeq V_f/\phi$ is the pertinent local volume (within a layer in present circumstances) of the initially fully fluid-saturated porous layer at the first instant of consideration, $V_\phi = \phi V$ is the corresponding pore volume, with ϕ being the fluid-saturated porosity of the same volume. V_f is the volume occupied by the pore-fluid, so that $V_f = \phi V$ before any new deformations begin. The δ 's here do indicate small changes in the quantities immediately following them. For “drained” systems, there would ideally be a reservoir of the same fluid just outside the volume V that can either supply more fluid or absorb any excreted fluid as needed during the nonstationary phase of the poroelastic process. The amount of pore fluid (*i.e.*, the number of fluid molecules) can therefore either increase or decrease from that of the initial amount of pore fluid; at the same time, the pore volume can also be changing, but — in general — not necessarily at exactly the same rate as the pore fluid itself. The one exception to these statements is when the surface pores of the layer volume V are sealed, in which case the layer is “undrained” and $\zeta \equiv 0$, identically. In such circumstances, it is still possible that both V_f and $V_\phi = \phi V$ are changing; but, because of the imposed undrained boundary conditions, they are necessarily changing at the same rate. The drained compliances are $s_{ij} = s_{ij}^d$, with or without the d superscript. Undrained compliances (not yet shown) are symbolized by s_{ij}^u . Coefficients

$$\beta_i = s_{i1} + s_{i2} + s_{i3} - 1/3K_R^g, \quad (3)$$

where K_R^g is again the Reuss average modulus of the grains. The drained Reuss average bulk modulus is defined by

$$\frac{1}{K_R^d} = \sum_{ij=1,2,3} s_{ij}^d. \quad (4)$$

For the Reuss (1929) average undrained bulk modulus K_R^u , undrained compliances have replaced drained compliances in a formula analogous to (4). A similar definition of the effective grain modulus K_R^g is:

$$\frac{1}{K_R^g} = \sum_{i,j=1,2,3} s_{ij}^g. \quad (5)$$

with grain compliances replacing drained compliances as discussed earlier by Berryman (2010). The alternative Voigt (1928) average [also see Hill (1952)] of the stiffnesses will play no role in the present work. And, finally, $\gamma = \sum_{i=1,2,3} \beta_i / BK_R^d$, where B is the second Skempton (1954) coefficient, which will be defined carefully later in my discussion.

The shear terms due to twisting motions (*i.e.*, strains e_{23} , e_{31} , e_{12} and stresses σ_{23} , σ_{31} , σ_{12}) are excluded from this poroelastic discussion since they typically do not couple to the modes of interest for anisotropic systems having orthotropic symmetry, or any more symmetric system such as those being either transversely isotropic or isotropic. I have also assumed that the true axes of symmetry are known, and make use of them in my formulation of the problem. Note that the s_{ij} 's are the elements of the compliance matrix \mathbf{S} and are all independent of the fluid, and therefore would be the same if the medium were treated as elastic (*i.e.*, by ignoring the fluid pressure, or assuming that the fluid saturant is air – or vacuum). In keeping with the earlier discussions, I typically call these compliances the drained compliances and the corresponding matrix the drained compliance matrix \mathbf{S}^d , since the fluids do not contribute to the stored mechanical energy if they are free to drain into a surrounding reservoir containing the same type of fluid. In contrast, the undrained compliance matrix \mathbf{S}^u presupposes that the fluid is trapped (unable to drain from the system into an adjacent reservoir) and therefore contributes in a significant and measurable way

to the compliance and stiffness ($\mathbf{C}^u = [\mathbf{S}^u]^{-1}$), and also therefore to the stored mechanical energy of the undrained system.

Although the significance of the formula is somewhat different now, I find again that

$$\beta_1 + \beta_2 + \beta_3 = \frac{1}{K_R^d} - \frac{1}{K_R^g} = \frac{\alpha_R}{K_R^d}, \quad (6)$$

if we also define (as we did for the isotropic case) a Reuss-averaged effective stress coefficient:

$$\alpha_R \equiv 1 - K_R^d/K_R^g. \quad (7)$$

Furthermore, I have

$$\gamma = \frac{\beta_1 + \beta_2 + \beta_3}{B} = \frac{\alpha_R}{K_R^d} + \phi \left(\frac{1}{K_f} - \frac{1}{K_R^\phi} \right), \quad (8)$$

since I have the rigorous result in this notation that Skempton's B coefficient is given by

$$B \equiv \frac{1 - K_R^d/K_R^u}{1 - K_R^d/K_R^g} = \frac{\alpha_R/K_R^d}{\alpha_R/K_R^d + \phi(1/K_f - 1/K_R^\phi)}. \quad (9)$$

Note that both (8) and (9) contain dependence on the distinct *pore* bulk modulus K_R^ϕ that comes into play when the pores are heterogeneous (Brown and Korrington, 1975), regardless of whether the system is isotropic or anisotropic. I emphasize that all these formulas are rigorous statements based on the earlier anisotropic analyses. The appearance of both the Reuss average quantities K_R^d and α_R is not an approximation, but merely a useful choice of notation.

Determining off-diagonal β_i coefficients

I will now provide several results for the β_i coefficients, and then follow the results with a general proof of their correctness.

In many useful and important cases, the coefficients β_i are determined by

$$\beta_i = s_{i1}^d + s_{i2}^d + s_{i3}^d - \frac{1}{3K_R^g}. \quad (10)$$

Again, K_R^g is the Reuss average of the grain modulus, since the local grain modulus is not necessarily assumed uniform here as discussed previously. Equation (10) holds as written for homogeneous grains, such that $K_R^g \equiv K^g$. It also holds true for the case when K_R^g is determined instead (Reuss, 1929; Wood, 1955; Hashin, 1962) by

$$\frac{1}{K_R^g} \equiv \sum_{m=1, \dots, n} \frac{v_m}{K_m}, \quad (11)$$

where v_m is the volume fraction (out of all the solid material present, so that $\sum_m v_m = 1$). However, when the grains themselves are uniform but anisotropic, I need to allow again

for this possibility, and this can be accomplished by defining three directional grain bulk moduli determined by:

$$\frac{1}{3\overline{K}_i^g} \equiv s_{i1}^g + s_{i2}^g + s_{i3}^g = s_{1i}^g + s_{2i}^g + s_{3i}^g, \quad (12)$$

for $i = 1, 2, 3$. The second equality follows because the compliance matrix is always symmetric. I call these quantities in (12) the “partial grain-compliance sums,” and the \overline{K}_i^g are the directional grain bulk moduli. Note that the factors of three have again been correctly accounted for because in agreement with (5):

$$\sum_{i=1,2,3} \frac{1}{3\overline{K}_i^g} = \frac{1}{K_R^g}, \quad (13)$$

I can also simplify and symmetrize our notation somewhat by introducing a similar concept for the drained constants, so that

$$\frac{1}{3\overline{K}_i^d} \equiv s_{i1}^d + s_{i2}^d + s_{i3}^d = s_{1i}^d + s_{2i}^d + s_{3i}^d, \quad (14)$$

for $i = 1, 2, 3$. Then, the formula for (10) is replaced by

$$\beta_i = \frac{1}{3\overline{K}_i^d} - \frac{1}{3\overline{K}_i^g}. \quad (15)$$

If the three contributions represented by (12) for $i = 1, 2, 3$ happen to be equal, then clearly each equals one-third of the sum (13).

The preceding results are for perfectly aligned grains. If the grains are instead perfectly randomly oriented, then it is clear that the formulas in (10) hold as before, but now K_R^g is determined instead by (5).

All of these statements about the β_i 's are easily proven by considering the situation when $\sigma_{11} = \sigma_{22} = \sigma_{33} = -p_c = -p_f$, where p_c is uniform external confining pressure, and p_f is the internal fluid pressure. Because then, from (1), I have (since by my assumption $p_c = p_f$):

$$-e_{ii} = \frac{1}{3\overline{K}_i^d} p_c + \beta_i (-p_f) = (s_{i1}^g + s_{i2}^g + s_{i3}^g) p_f \equiv \frac{p_f}{3\overline{K}_i^g}, \quad (16)$$

in the most general of the three cases discussed, and holding true for each value of $i = 1, 2, 3$. This is a statement about the strain e_{ii} that would be observed in this situation, as it must be the same if these anisotropic (or inhomogeneous) grains were immersed in the fluid, while measurements were taken of the strains observed in each of the three directions $i = 1, 2, 3$, during variations of the fluid pressure p_f . I consider this proof to be a thought experiment for determining these coefficients, in the same spirit as those proposed originally by Biot and Willis (1957) and Biot (1962) for the isotropic and homogeneous case.

The β_i coefficients and effective stress

Making use of the previous definitions, it is easy to see that the coefficients β_i are closely related to different sort of effective stress coefficient, for the individual principal strain

coefficients:

$$e_{ii} = -\frac{1}{3\bar{K}_i^d}(p_c - D_i p_f), \quad \text{for } i = 1, 2, 3, \quad (17)$$

where

$$D_i = 3\bar{K}_i^d \beta_i = 1 - \frac{\bar{K}_i^d}{\bar{K}_i^g}, \quad \text{for } i = 1, 2, 3, \quad (18)$$

and $-p_c = \sigma_{11} = \sigma_{22} = \sigma_{33}$ in the case of uniform applied confining pressure p_c . Then clearly, the D_i 's are completely analogous to the usual Biot [or Biot and Willis (1957)] coefficient $\alpha_R = 1 - K_R^d/K_R^g$, as commonly defined for isotropic poroelasticity.

Coefficient γ

The relationship of coefficient γ to the other coefficients is easily established because I have already discussed the main issue, which involves determining the role of the various other constants contained in Skempton's second coefficient B (Skempton, 1954). This result is

$$B = \left(\frac{1}{K_R^d} - \frac{1}{K_R^g} \right) \left[\left(\frac{1}{K_R^d} - \frac{1}{K_R^g} \right) + \phi \left(\frac{1}{K_f} - \frac{1}{K_R^\phi} \right) \right]^{-1} \quad (19)$$

From (1), I find that

$$-\zeta = 0 = -(\beta_1 + \beta_2 + \beta_3) \sigma_c - \gamma p_f, \quad (20)$$

for undrained boundary conditions. Thus, I find again that

$$\frac{p_f}{p_c} \equiv B = \frac{\beta_1 + \beta_2 + \beta_3}{\gamma}, \quad (21)$$

where $p_c = -\sigma_c$ is the confining pressure. Thus, the scalar coefficient γ is determined immediately and given by

$$\gamma = \frac{\beta_1 + \beta_2 + \beta_3}{B} = \frac{\alpha_R/K_R^d}{B} = \alpha_R/K_R^d + \phi \left(\frac{1}{K_f} - \frac{1}{K_R^\phi} \right). \quad (22)$$

Alternatively, I could say that

$$B \equiv \frac{\alpha_R}{\gamma K_R^d}. \quad (23)$$

I have now determined the physical/mechanical significance of all the coefficients in the poroelastic matrix (1). These results are as general as possible without considering poroelastic symmetries that have less than orthotropic symmetry, while also taking advantage of my assumption that I do typically know the three directions of the principal axes of symmetry.

Inverting poroelastic compliance

Being in compliance form, the matrix in (1) has extremely simple poroelastic behavior in the sense that all the fluid mechanical effects appear only in the single coefficient γ . I can simplify the notation a little more by lumping some coefficients together, combining the 3×3 submatrix in the upper left corner of the matrix in (1) as \mathbf{S} , and defining the column vector \mathbf{b} by

$$\mathbf{b}^T \equiv (\beta_1, \beta_2, \beta_3). \quad (24)$$

The resulting 4×4 matrix and its inverse are now related by:

$$\begin{pmatrix} \mathbf{S} & -\mathbf{b} \\ -\mathbf{b}^T & \gamma \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{q} \\ \mathbf{q}^T & z \end{pmatrix}^{-1}, \quad (25)$$

where the elements of the inverse matrix can be shown to be written in terms of drained stiffness matrix $\mathbf{C}^d = \mathbf{C} = \mathbf{S}^{-1}$ by introducing three components: (a) scalar $z = [\gamma - \mathbf{b}^T \mathbf{C} \mathbf{b}]^{-1}$, (b) column vector $\mathbf{q} = z \mathbf{C} \mathbf{b}$, and (c) undrained 3×3 stiffness matrix (*i.e.*, the pertinent one connecting the principal strains to principal stresses) is given by $\mathbf{A} = \mathbf{C} + z \mathbf{C} \mathbf{b} \mathbf{b}^T \mathbf{C} = \mathbf{C}^d + z^{-1} \mathbf{q} \mathbf{q}^T \equiv \mathbf{C}^u$, since \mathbf{C}^d is drained stiffness and $\mathbf{A} = \mathbf{C}^u$ is clearly undrained stiffness by construction. This result is the same as that of Gassmann (1951) for anisotropic porous media, although his results were presented in a form somewhat harder to scan than the form shown explicitly here.

Also, note the important fact that the observed decoupling of the fluid effects occurs only in the compliance form (1) of the equations, and never in the stiffness (inverse) form for the poroelasticity equations.

From these results, it is not hard to show that

$$\mathbf{S}^d = \mathbf{S}^u + \gamma^{-1} \mathbf{b} \mathbf{b}^T. \quad (26)$$

This result emphasizes the remarkably simple fact that the drained compliance matrix can be found directly from knowledge of the inverse of undrained stiffness, and the still unknown, but sometimes relatively easy to estimate, values of γ , together with the three distinct orthotropic β_i coefficients, for $i = 1, 2, 3$.

There are clearly many measurements required to determine all these various poroelastic coefficients. Furthermore, the strategy for finding the coefficients depends on available data sets, and whether the porous media of interest are constructed from a homogeneous or heterogeneous set of solid materials, and whether the individual grains are isotropic or anisotropic. It also makes some difference if the pores are approximately rounded (for granular media) or flat (for fractured media). All these issues have been discussed previously at length, and this discussion will not be repeated here.

The remainder of the paper will concentrate on making use of the general poroelastic equations in situations where at least two and possibly many distinct layers of porous materials obeying these equations are under stress (either quasi-static or dynamic as would occur in a wave propagation scenario). As will be shown, the layered poroelastic equations behave somewhat differently from layered elastic equations because there are two distinct additional boundary conditions (drained and undrained) that can occur depending on the details of the excitation itself.

AVERAGING RESULTS FOR ALL DRAINED OR ALL UNDRAINED BOUNDARY CONDITIONS

The two most common boundary conditions to consider in poroelastic media are the drained and undrained conditions. Drained conditions imply that the fluid pressure change is zero while the increment of fluid content in the individual layers may be considered arbitrary. Of course, the total amount of fluid present needs to be properly conserved in the analysis to be presented, but the usual idea for drained conditions is that the poroelastic systems is immersed in an infinite reservoir of fluid so that pore fluid is freely available to move in and out of the region of interest. For present studies, this situation implies that the layer increments ζ can take arbitrary (small) values, but the fluid pressure is constrained to be a constant value p_f everywhere. So changes in p_f always vanish for drained conditions.

Undrained boundary conditions place the hard constraint on the fluid increment ζ , requiring no flow at the boundaries, so $\zeta = 0$ at all boundaries. These conditions ensure that the fluid pressure p_f does change, since as the boundaries move in or out the pressure on the confined fluid is increasing or decreasing.

Both of these conditions must be approximations to conditions in a generally realistic earth model. It is easy to imagine situations where some boundaries between layers (the vertical direction) are plugged, so undrained boundary conditions $\zeta_z \equiv 0$ might be correct while neighboring layers (horizontal direction) might be open to fluid flow (so ζ_x and/or $\zeta_y \neq 0$). I will consider these more general situations in later work, but for now limit the analysis to that for either *all drained* conditions or *all undrained* conditions. All undrained conditions are also appropriate, as mentioned previously, regardless of the physical boundary conditions if the probe changing the physical variables is a passing high frequency acoustic or seismic wave train or pulse.

General analysis for layered poroelastic systems

I will next formulate the layered poroelastic earth problem in a way so that both of these standard boundary conditions can be imposed, as needed in any particular modeling problem.

Now I assume throughout the rest of the paper that the porous layers are stacked vertically (3- or z -axis) and for this geometry it is easy to see that the three horizontal strains e_{11} , e_{22} , and e_{12} must be continuous if the layers are in solid-welded contact. Furthermore, the vertical stress σ_{33} , and rotational stresses involving the vertical direction σ_{13} and σ_{23} must also be continuous. These conditions are equivalent to an assumption of welded contact between layers. If contact is not welded, then the system can have more complicated behaviors than those being considering here.

Appendix A summarizes the Backus (1962) and/or Schoenberg and Muir (1989) approach to elastic layer averaging. The method I present here provides a small generalization of this approach, taking the presence of the pore-fluid into account. For the drained situation, the influence of the fluid on the system mechanics is minimal (as will be shown). But I should nevertheless have this result available to compare it with the more interesting case of the undrained layers.

Although the shear moduli normally associated with the twisting shear components e_{23} ,

e_{31} , and e_{12} usually do not interact with the pore-fluid itself in systems as symmetric or more symmetric than orthotropic, I nevertheless need to carry these terms along in the poroelastic formulation for layered systems because of possible boundary effects due to welded contact at interfaces. To accomplish this goal, I will generalize the form of equation (42) from Appendix A. In compliance form, the equations will relate the strains

$$E_T \equiv \begin{pmatrix} e_{11} \\ e_{22} \\ e_{12} \end{pmatrix}, \quad \text{and} \quad E_N \equiv \begin{pmatrix} e_{33} \\ e_{32} \\ e_{31} \end{pmatrix}, \quad (27)$$

and fluid increment ζ to the stresses

$$\Pi_T \equiv \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \end{pmatrix}, \quad \text{and} \quad \Pi_N \equiv \begin{pmatrix} \sigma_{33} \\ \sigma_{32} \\ \sigma_{31} \end{pmatrix}, \quad (28)$$

and the fluid pressure change p_f .

The required general relationship is:

$$\begin{pmatrix} E_T \\ -\zeta \\ E_N \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{TT} & -\mathbf{g}_{12} & \mathbf{S}_{TN} \\ -\mathbf{g}_{12}^T & \gamma & -\mathbf{g}_3^T \\ \mathbf{S}_{NT} & -\mathbf{g}_3 & \mathbf{S}_{NN} \end{pmatrix} \begin{pmatrix} \Pi_T \\ -p_f \\ \Pi_N \end{pmatrix}, \quad (29)$$

where, for example, in the orthotropic media considered here I have:

$$\mathbf{S}_{TT} \equiv \begin{pmatrix} s_{11} & s_{12} & s_{16} \\ s_{21} & s_{22} & s_{26} \\ s_{61} & s_{62} & s_{66} \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & \\ s_{21} & s_{22} & \\ & & s_{66} \end{pmatrix}, \quad (30)$$

$$\mathbf{S}_{NN} \equiv \begin{pmatrix} s_{33} & s_{34} & s_{35} \\ s_{43} & s_{44} & s_{45} \\ s_{53} & s_{54} & s_{55} \end{pmatrix} = \begin{pmatrix} s_{33} & & \\ & s_{44} & \\ & & s_{55} \end{pmatrix}, \quad (31)$$

and

$$\mathbf{S}_{NT} \equiv \begin{pmatrix} s_{31} & s_{32} & s_{36} \\ s_{41} & s_{42} & s_{46} \\ s_{51} & s_{52} & s_{56} \end{pmatrix} = \begin{pmatrix} s_{31} & s_{32} & \\ & 0 & \\ & & 0 \end{pmatrix}, \quad (32)$$

with $\mathbf{S}_{TN} = \mathbf{S}_{NT}^T$ (the T superscript indicates the matrix transpose). Here all these expressions for elastic compliance refer specifically to drained compliances $s_{ij} = s_{ij}^d$, for all $i, j = 1, \dots, 6$ within each poroelastic anisotropic layer.

All the poroelastic contributions to (29) are determined by γ , \mathbf{g}_{12} , and \mathbf{g}_3 . The scalar γ within the 7×7 matrix in (29) was defined earlier in (8), and is the only term in the 7×7 matrix that includes fluid effects directly through K_f . The remaining pair of vectors contained within the 7×7 matrix in (29) is defined by:

$$\mathbf{g}_{12}^T = (\beta_1, \beta_2, 0) \quad (33)$$

and

$$\mathbf{g}_3^T = (\beta_3, 0, 0), \quad (34)$$

where the β 's were defined previously following (1).

I now consider two examples of special uses of the general equation (29) for different choices of boundary conditions. These two physical circumstances covered in the cases considered are distinct end-members. For relatively high-frequency wave propagation, it is appropriate to consider that the fluids do not have time to equilibrate, and therefore fluid pressures can be different in distinct layers, while the fluid particles do not have time to move very far during wave passage time, so the fluid increment $\zeta = 0$ essentially everywhere. This situation is called the “undrained” condition. An alternative condition is the fully drained condition, in which the fluid particles have as much time as they need to achieve fluid-pressure equilibration, so that $p_f = \text{constant}$. These two limiting situations are clearly connected physically via Darcy’s law, which provides the mechanism to move fluid particles, and ultimately to guarantee that the fluid pressure reaches an equilibrium state. Bringing Darcy’s law actively into play in the equations would result in Biot-style equations which are beyond my current scope. So I consider only the end-member conditions for the present contribution.

Drained scenario ($p_f \equiv 0$)

Now, recall that, in the drained scenario, changes in pore-fluid pressure are assumed to be zero (or at least negligibly small), so $p_f \equiv 0$ in these equations. Accounting for this condition, the results should (and do) recover the Backus (1962) and Schoenberg and Muir (1989) results for the elastic parts of the system (found in Appendix A) exactly. Also, I find the additional (expected) result for the poroelastic case that the average fluid increment is:

$$\langle \zeta \rangle = \langle \beta_1 \sigma_{11} \rangle + \langle \beta_2 \sigma_{22} \rangle + \langle \beta_3 \rangle \sigma_{33}, \quad (35)$$

if σ_{33} is nearly constant. Or, if σ_{33} is not uniform from one layer to the next (as might happen due to weight of solid overburden pressure), then the third expression in (35) should be modified, by moving σ_{33} inside the averaging operator. So then I have

$$\langle \zeta \rangle = \langle \beta_1 \sigma_{11} \rangle + \langle \beta_2 \sigma_{22} \rangle + \langle \beta_3 \sigma_{33} \rangle, \quad (36)$$

whenever σ_{33} taken constant is a poor approximation. The results shown in (35) and (36) are easy to reconcile with the definitions of the β 's, and the meaning of averaging operator $\langle \cdot \rangle$ across all layers. When p_f vanishes everywhere, the final results for the averaging and the various stresses and strains are identical to the results in Appendix A. For the drained scenario, the only difference is the addition of equations (35) or (36).

Undrained scenario ($\zeta \equiv 0$)

Now consider that the fluid pressure might vary across the stack of layers (as should be expected to happen either because of hydrostatic overburden, or due to fluid injection or extraction at certain chosen depths). Then I can treat this case as well, assuming undrained circumstances, by averaging the fluid pressure itself via $\langle p_f \rangle$. In this case, some knowledge of the fluid-pressure distribution along the stack of layers would be required, as well as some information about whether the undrained condition applies at every interface, or just at some interfaces. Variations might occur if a sealing layer were present to close off flow at

the top, or bottom. Both ends might be sealed for some range of porous layers forming a heterogeneous, layered anisotropic reservoir including cap rocks. For this undrained case, the fluid pressure in each undrained layer is free to vary compared to all the others; so there is no constancy of p_f for this scenario. The averaging condition resulting from the formulation for such a reservoir according to (29) is:

$$\langle p_f \rangle = - \left\langle \frac{1}{\gamma} (\beta_1 \sigma_{11} + \beta_2 \sigma_{22}) \right\rangle + \left\langle \frac{\beta_3 \sigma_{33}}{\gamma} \right\rangle. \quad (37)$$

Proper choice of the range of depth for averaging will clearly depend on the details of each reservoir, and the type of physical probe being used. For example, either quarter- or half-wavelength for seismic waves, when used as the probe, would be typical choices of the averaging depth in this case.

While the preceding part of the averaging for undrained boundary conditions was straightforward, I still need to check what happens when averaging the remainder of the equations. I show the work in Appendix B leading to the general undrained result (61), but just quote the final answer here – being valid for each undrained layer in the overall system:

$$\begin{pmatrix} E_T \\ E_N \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{TT}^u & \mathbf{S}_{TN}^u \\ \mathbf{S}_{NT}^u & \mathbf{S}_{NN}^u \end{pmatrix} \begin{pmatrix} \Pi_T \\ \Pi_N \end{pmatrix}, \quad (38)$$

where

$$\mathbf{S}_{TT}^u \equiv \begin{pmatrix} s_{11}^u & s_{12}^u \\ s_{21}^u & s_{22}^u \\ & & s_{66} \end{pmatrix}, \quad \mathbf{S}_{NN}^u \equiv \begin{pmatrix} s_{33}^u & & \\ & s_{44} & \\ & & s_{55} \end{pmatrix}, \quad \text{and} \quad \mathbf{S}_{NT}^u \equiv \begin{pmatrix} s_{31}^u & s_{32}^u \\ & 0 \\ & & 0 \end{pmatrix}, \quad (39)$$

while $\mathbf{S}_{TN}^u = (\mathbf{S}_{NT}^u)^T$. Once these definitions are used for the undrained matrices, the layer analysis for the system follows exactly the same steps as in Appendix A. Note that I arrived at these results in another (step-by-step) way in Appendix B in order to prove that this is the right answer for the undrained problem. Fortunately, the intuitive answer is also the same as the right answer.

SUMMARY AND CONCLUSIONS

This paper has presented a method of generalizing the elasticity-based methods of Backus (1962) and of Schoenberg and Muir (1989) to poroelastic systems, having additional dependencies on pore-pressures and pore-fluid physical properties. The approach taken is very close in spirit to that of Schoenberg and Muir (1989). The results are necessarily a little more complicated due to the presence of the pore-fluid, which permits different types of elastic-poroelastic behavior to occur than can appear in simply elastic stratified systems. The drained results agree completely with the elastic results, as they should. The undrained results permit studies of complicated layered poroelastic systems to proceed in an analytical framework that should be helpful for better understanding of a wide range of problems in oil and gas reservoir exploration.

APPENDIX A: THE SCHOENBERG-MUIR METHOD

The quasi-static elasticity equations are often written in compliance form using the Voigt 6×6 matrix notation as:

$$\begin{pmatrix} e_{11} \\ e_{22} \\ e_{33} \\ e_{23} \\ e_{31} \\ e_{12} \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16} \\ s_{21} & s_{22} & s_{23} & s_{24} & s_{25} & s_{26} \\ s_{31} & s_{32} & s_{33} & s_{34} & s_{35} & s_{36} \\ s_{41} & s_{42} & s_{43} & s_{44} & s_{45} & s_{46} \\ s_{51} & s_{52} & s_{53} & s_{54} & s_{55} & s_{56} \\ s_{61} & s_{62} & s_{63} & s_{64} & s_{65} & s_{66} \end{pmatrix} \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{31} \\ \sigma_{12} \end{pmatrix} \equiv \mathbf{S} \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{31} \\ \sigma_{12} \end{pmatrix}, \quad (40)$$

where \mathbf{S} is the symmetric 6×6 compliance matrix. The numbers 1,2,3 always indicate Cartesian axes (say, x,y,z respectively). The z -direction is usually chosen as the layering direction, which could be oriented any direction in the earth. But, in many geological and geophysical applications, the 3-axis (or z -axis) is also taken to be the vertical direction, and I conform to this convention here. The principal stresses are σ_{11} , σ_{22} , σ_{33} , in the directions 1,2,3, respectively. Similarly, the principal strains are e_{11} , e_{22} , e_{33} . The stresses σ_{23} , σ_{31} , σ_{12} are the torsional shear stresses, associated with rotation-based strains around the 1, 2, or 3 axes, respectively. The corresponding torsional strains are e_{23} , e_{31} , and e_{12} , where the torsional motion is again a rotational straining motion around the 1, 2, or 3 axes. The compliance matrix is symmetric, so $s_{ij} = s_{ji}$, and this fact could have been used when displaying the matrix. The axis pairs in the subscripts 11, 22, 33, 23, 31, and 12, are often labelled (again following the conventions of Voigt) as 1,2,3,4,5,6, respectively.

The important contribution made by Backus (1962) [also see Postma (1955)] is the observation that, in a horizontally layered system, there are certain strains e_{ij} and stresses σ_{ij} that are necessarily continuous across boundaries between layers, while the others are not necessarily continuous. I have been implicitly (and now explicitly by calling this fact out) assuming that the interfaces between layers are in welded contact, which means practically that the in-plane strains are always continuous: so if axis 3 (or z) is the symmetry axis (as is most often chosen for our layering problem), I have e_{11} , $e_{12} = e_{21}$, and e_{22} are all continuous. Similarly, in welded contact, I must have continuity of the all the stresses involving the 3 (or z) direction: so σ_{33} , $\sigma_{13} = \sigma_{31}$, and $\sigma_{23} = \sigma_{32}$ must all be continuous.

Then, following Backus (1962) and/or Schoenberg and Muir (1989) but — for present purposes considering instead the compliance (inverse of stiffness) matrix — I have rearranged the statement of the problem so that:

$$\begin{pmatrix} e_{11} \\ e_{22} \\ e_{12} \\ e_{33} \\ e_{32} \\ e_{31} \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{16} & s_{13} & s_{14} & s_{15} \\ s_{21} & s_{22} & s_{26} & s_{23} & s_{24} & s_{25} \\ s_{61} & s_{62} & s_{66} & s_{63} & s_{64} & s_{65} \\ s_{31} & s_{32} & s_{36} & s_{33} & s_{34} & s_{35} \\ s_{41} & s_{42} & s_{46} & s_{43} & s_{44} & s_{45} \\ s_{51} & s_{52} & s_{56} & s_{53} & s_{54} & s_{55} \end{pmatrix} \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \\ \sigma_{33} \\ \sigma_{32} \\ \sigma_{31} \end{pmatrix}. \quad (41)$$

Note that this equation, although similar to (40) is nevertheless quite different because of the rearrangement of the matrix elements and the reordering of the strains and stresses. The expression in (41) is general for all elastic media. In the main text I restrict the discussion to orthotropic media. Assuming then that I am using the correct set of axes as

the symmetry axes in the presentation, all *off-diagonal* compliances having subscripts 4, 5, or 6 in (40) vanish identically. The diagonal shear compliances s_{44} , etc., generally do not vanish however.

Expression of (41) can be made more compact by writing it as:

$$\begin{pmatrix} E_T \\ E_N \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{TT} & \mathbf{S}_{TN} \\ \mathbf{S}_{NT} & \mathbf{S}_{NN} \end{pmatrix} \begin{pmatrix} \Pi_T \\ \Pi_N \end{pmatrix}, \quad (42)$$

where

$$\mathbf{S}_{TT} \equiv \begin{pmatrix} s_{11} & s_{12} & s_{16} \\ s_{21} & s_{22} & s_{26} \\ s_{61} & s_{62} & s_{66} \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & \\ s_{21} & s_{22} & \\ & & s_{66} \end{pmatrix}, \quad (43)$$

$$\mathbf{S}_{NN} \equiv \begin{pmatrix} s_{33} & s_{34} & s_{35} \\ s_{43} & s_{44} & s_{45} \\ s_{53} & s_{54} & s_{55} \end{pmatrix} = \begin{pmatrix} s_{33} & & \\ & s_{44} & \\ & & s_{55} \end{pmatrix}, \quad (44)$$

and

$$\mathbf{S}_{NT} \equiv \begin{pmatrix} s_{31} & s_{32} & s_{36} \\ s_{41} & s_{42} & s_{46} \\ s_{51} & s_{52} & s_{56} \end{pmatrix} = \begin{pmatrix} s_{31} & s_{32} & \\ & 0 & \\ & & 0 \end{pmatrix}, \quad (45)$$

with $\mathbf{S}_{TN} = \mathbf{S}_{NT}^T$ (with T superscript indicating the matrix transpose). Also I have

$$E_T \equiv \begin{pmatrix} e_{11} \\ e_{22} \\ e_{12} \end{pmatrix}, \quad \text{and} \quad E_N \equiv \begin{pmatrix} e_{33} \\ e_{32} \\ e_{31} \end{pmatrix}, \quad (46)$$

and

$$\Pi_T \equiv \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \end{pmatrix}, \quad \text{and} \quad \Pi_N \equiv \begin{pmatrix} \sigma_{33} \\ \sigma_{32} \\ \sigma_{31} \end{pmatrix}. \quad (47)$$

It is important to distinguish between “slow” and “fast” variables in this analysis, since this distinction makes it clear when and how averaging should be performed. The “slow” variables, *i.e.*, those that are continuous across the (here assumed horizontal) boundaries and also essentially constant for the present quasi-static application, are those contained in E_T and Π_N . So, after averaging $\langle \cdot \rangle$ along the layering direction, I have:

$$\begin{pmatrix} E_T \\ \langle E_N \rangle \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{TT}^* & \mathbf{S}_{TN}^* \\ \mathbf{S}_{NT}^* & \mathbf{S}_{NN}^* \end{pmatrix} \begin{pmatrix} \langle \Pi_T \rangle \\ \Pi_N \end{pmatrix}, \quad (48)$$

where $\mathbf{S}_{TN}^* = (\mathbf{S}_{NT}^*)^T$, and all the starred quantities are the *nontrivial* average compliances I seek. They are defined in terms of layer-average quantities where the symbol $\langle \cdot \rangle$ indicates a simple volume average of all the layers. By this notation I mean that a quantity Q that takes on different values in different layers has the layer average $\langle Q \rangle \equiv x_a Q_a + x_b Q_b + \dots$. The definition is general and applies to an arbitrary number of different layers where the fraction of the total volume occupied by layer a is x_a , etc. Total fractional volume is $x_a + x_b + \dots \equiv 1$.

Of the three final results, the two easiest ones to compute are:

$$\mathbf{S}_{TT}^* = \langle \mathbf{S}_{TT}^{-1} \rangle^{-1}, \quad (49)$$

$$\mathbf{S}_{TN}^* = (\mathbf{S}_{NT}^*)^T = \langle \mathbf{S}_{TT}^{-1} \rangle^{-1} \langle \mathbf{S}_{TT}^{-1} \mathbf{S}_{TN} \rangle = \mathbf{S}_{TT}^* \langle \mathbf{S}_{TT}^{-1} \mathbf{S}_{TN} \rangle, \quad (50)$$

where $\langle \cdot \rangle$ is the layer average of some quantity. These results follow from this equation:

$$\langle \mathbf{S}_{TT}^{-1} \rangle E_T = \langle \Pi_T \rangle + \langle \mathbf{S}_{TT}^{-1} \mathbf{S}_{TN} \rangle \Pi_N, \quad (51)$$

which also followed immediately from the formula

$$E_T = \mathbf{S}_{TT} \Pi_T + \mathbf{S}_{TN} \Pi_N \quad (52)$$

multiplying through first by the inverse of \mathbf{S}_{TT} , and then performing the layer average. [Note that \mathbf{S}_{TT} and \mathbf{S}_{NN} are both normally square and invertible matrices, whereas for most systems the off-diagonal matrix \mathbf{S}_{NT} is not invertible. But, this fact does not cause problems in the analysis, because I do not need to invert \mathbf{S}_{NT} in order to solve the averaging problem at hand.] These averages are meaningful because, when the matrix equations presented are multiplied out, there never appear any cross products of two quantities that are both unknown. [From this view point, Eq. (51) is an equation for $\langle \Pi_T \rangle$, just as the unaveraged version of (51) is an equation for Π_T in each layer.] So simple layer-averaging suffices (thereby providing the main motivation and value of this method). Multiplying (51) through by $\langle \mathbf{S}_{TT}^{-1} \rangle^{-1}$ then gives the results (49) and (50).

The remaining result is more tedious to compute, since it requires several intermediate steps in its derivation. But the final result is given by the formula:

$$\mathbf{S}_{NN}^* = \langle \mathbf{S}_{NN} \rangle - \langle \mathbf{S}_{NT} \mathbf{S}_{TT}^{-1} \mathbf{S}_{TN} \rangle + \mathbf{S}_{NT}^* (\mathbf{S}_{TT}^*)^{-1} \mathbf{S}_{TN}^*. \quad (53)$$

To provide some clues to the derivation, again consider:

$$\Pi_T = \mathbf{S}_{TT}^{-1} E_T - \mathbf{S}_{TT}^{-1} \mathbf{S}_{TN} \Pi_N, \quad (54)$$

which is just a rearrangement of (52). The point is that $\langle \Pi_T \rangle$ is then given immediately in terms of the quantities E_T and Π_N , which are both “slow” variables and therefore essentially constant. An intermediate result that helps to explain the form of this relation (53) is:

$$\mathbf{S}_{NT}^* (\mathbf{S}_{TT}^*)^{-1} \mathbf{S}_{TN}^* = \langle \mathbf{S}_{NT} \mathbf{S}_{TT}^{-1} \rangle \langle \mathbf{S}_{TT}^{-1} \rangle^{-1} \langle \mathbf{S}_{TT}^{-1} \mathbf{S}_{TN} \rangle = \langle \mathbf{S}_{NT} \mathbf{S}_{TT}^{-1} \rangle \mathbf{S}_{TN}^*. \quad (55)$$

Substituting for Π_T from (54) into

$$E_N = \mathbf{S}_{NT} \Pi_T + \mathbf{S}_{NN} \Pi_N \quad (56)$$

and then averaging, I find that

$$\langle E_N \rangle = \langle \mathbf{S}_{NT} \mathbf{S}_{TT}^{-1} \rangle E_T + \langle \mathbf{S}_{NN} - \mathbf{S}_{NT} \mathbf{S}_{TT}^{-1} \mathbf{S}_{TN} \rangle \Pi_N, \quad (57)$$

an expression that completely determines the remaining coefficients. After some more algebra, the formula giving the final result is:

$$\begin{aligned} \langle E_N \rangle &= \langle \mathbf{S}_{NT} \mathbf{S}_{TT}^{-1} \rangle \langle \mathbf{S}_{TT}^{-1} \rangle^{-1} [\langle \Pi_T \rangle + \langle \mathbf{S}_{TT}^{-1} \mathbf{S}_{TN} \rangle \Pi_N] \\ &+ [\langle \mathbf{S}_{NN} \rangle - \langle \mathbf{S}_{NT} \mathbf{S}_{TT}^{-1} \mathbf{S}_{TN} \rangle] \Pi_N \\ &= \mathbf{S}_{NT}^* \langle \Pi_T \rangle + \mathbf{S}_{NN}^* \Pi_N. \end{aligned} \quad (58)$$

Equation (58) contains all the information needed to produce the third and final result found in (53).

Another check on these formulas is to compare them directly to those found by Schoenberg and Muir (1989). However, direct comparison is not so easy, since their analysis focuses on the stiffness version of these equations. My treatment makes use of the compliance version instead. Since the symmetries of the two forms of the equations nevertheless are nearly identical, cross-checks and comparisons will be left to the interested reader.

APPENDIX B: POROELASTIC FORMULAS FOR UNDRAINED BOUNDARY CONDITIONS IN LAYERED SYSTEMS

Using equation (29) as our starting point, I next treat the boundary condition $\zeta = 0$ for undrained layers (meaning that the fluid is actually physically trapped in the layer, or the physical process is so fast – such as high frequency wave propagation – that the fluid inertia prevents rapid movement of fluid particles over non-infinitesimal distances). Depending on the application scenario, this boundary condition might be applied to all layers, or only to just one or a few layers.

I consider first a single layer having the undrained boundary condition. For this case, the condition from Eq. (29) becomes

$$0 = \mathbf{g}_{12}^T \Pi_T + \gamma p_f + \mathbf{g}_3^T \Pi_N, \quad (59)$$

within the layer. Next, the equation can be solved to express the fluid pressure p_f strain dependence in each undrained layer (the layer labels are suppressed here for simplicity) as

$$p_f = -\frac{1}{\gamma} (\mathbf{g}_{12}^T \Pi_T + \mathbf{g}_3^T \Pi_N). \quad (60)$$

Then, substituting this condition back into the expressions for E_T and E_N from (29), I find that

$$\begin{pmatrix} E_T \\ E_N \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{TT} - \gamma^{-1} \mathbf{g}_{12} \mathbf{g}_{12}^T & \mathbf{S}_{TN} - \gamma^{-1} \mathbf{g}_{12} \mathbf{g}_3^T \\ \mathbf{S}_{NT} - \gamma^{-1} \mathbf{g}_3 \mathbf{g}_{12}^T & \mathbf{S}_{NN} - \gamma^{-1} \mathbf{g}_3 \mathbf{g}_3^T \end{pmatrix} \begin{pmatrix} \Pi_T \\ \Pi_N \end{pmatrix}. \quad (61)$$

To understand the significance of (61), I next consider that it is straightforward to show that each of these composite matrix elements corresponds exactly to the undrained version of the Schoenberg-Muir matrices. So that,

$$\mathbf{S}_{TT}^u \equiv \mathbf{S}_{TT} - \gamma^{-1} \mathbf{g}_{12} \mathbf{g}_{12}^T, \quad (62)$$

$$\mathbf{S}_{NN}^u \equiv \mathbf{S}_{NN} - \gamma^{-1} \mathbf{g}_3 \mathbf{g}_3^T, \quad (63)$$

and

$$\mathbf{S}_{TN}^u \equiv \mathbf{S}_{TN} - \gamma^{-1} \mathbf{g}_{12} \mathbf{g}_3^T = (\mathbf{S}_{NT}^u)^T. \quad (64)$$

All these expressions follow directly from the form of (61).

Thus, I arrive at a result that might have been anticipated, which is that the undrained layers respond according to the usual undrained conditions in each individual layer. The part of the result that is new concerns the forms of the undrained matrices \mathbf{S}_{TT}^u , $\mathbf{S}_{NT}^u = (\mathbf{S}_{TN}^u)^T$, and \mathbf{S}_{NN}^u , in the now modified Schoenberg-Muir formalism.

This analogy can be pushed somewhat further to include the effective values for the undrained moduli \mathbf{S}_{TT}^{u*} , $\mathbf{S}_{NT}^{u*} = (\mathbf{S}_{TN}^{u*})^T$, and \mathbf{S}_{NN}^{u*} , with formulas entirely analogous to (49), (50), and (53), and undrained constants replacing drained constants everywhere. Since there is nothing subtle about this step, I will therefore leave these details to be filled in by the interested reader.

REFERENCES

- Backus, G. E., 1962, Long-wave elastic anisotropy produced by horizontal layering: *Journal of Geophysical Research*, **67**, 4427–4440.
- Berryman, J. G., 2010, Poroelastic measurement schemes resulting in complete data sets for granular and other anisotropic porous media: *International Journal of Engineering Science*, **48**, 446–459.
- Biot, M. A., 1962, Mechanics of deformation and acoustic propagation in porous media: *Journal of Applied Physics*, **33**, 1482–1498.
- Biot, M. A. and D. G. Willis, 1957, The elastic coefficients of the theory of consolidation: *Journal of Applied Mechanics*, **24**, 594–601.
- Brown, R. J. S. and J. Korrington, 1975, On the dependence of the elastic properties of a porous rock on the compressibility of the pore fluid: *Geophysics*, **40**, 608–616.
- Cheng, A. H.-D., 1997, Material coefficients of anisotropic poroelasticity: *International Journal of Rock Mechanics*, **34**, 199–205.
- Gassmann, F., 1951, Über die Elastizität poröser Medien: *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, **96**, 1–23.
- Hashin, Z., 1962, Elastic moduli of heterogeneous materials: *Journal of Applied Mechanics*, **29**, 143–150.
- Hill, R., 1952, The elastic behaviour of crystalline aggregate: *Proceedings of the Physical Society of London*, **A65**, 349–354.
- Postma, G. W., 1955, Wave propagation in a stratified medium: *Geophysics*, **20**, 780–806.
- Pride, S. R. and J. G. Berryman, 2003a, Linear dynamics of double-porosity dual-permeability materials: I. governing equations and acoustic attenuation: *Physical Review E*, **68**, 036693.
- , 2003b, Linear dynamics of double-porosity dual-permeability materials: II. fluid transport equations: *Physical Review E*, **68**, 036694.
- Pride, S. R., J. G. Berryman, and J. M. Harris, 2004, Seismic attenuation due to wave-induced flow: *Journal of Geophysical Research*, **109**, B01201.
- Pride, S. R., E. Tromeur, and J. G. Berryman, 2002, Biot slow-wave effects in stratified rock: *Geophysics*, **67**, 271–281.
- Reuss, A., 1929, Berechnung der Fließgrenze von Mischkristallen: *Z. Angew. Math. Mech.*, **9**, 55.
- Rice, J. R. and M. P. Cleary, 1976, Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents: *Reviews of Geophysics and Space Physics*, **14**, 227–241.
- Schoenberg, M. and F. Muir, 1989, A calculus for finely layered anisotropic media: *Geophysics*, **54**, no. 4, 581–589.
- Skempton, A. W., 1954, The pore-pressure coefficients A and B: *Geotechnique*, **4**, 143–147.
- Thigpen, L. and J. G. Berryman, 1985, Mechanics of porous elastic materials containing multiphase fluid: *International Journal of Engineering Science*, **23**, 1203–1210.
- Voigt, W., 1928, *Lehrbuch der Kristallphysik*: Teubner, Leipzig.
- Wang, H. F., 2000, *Theory of Linear Poroelasticity with Applications to Geomechanics and Hydrology*: Princeton University Press, Princeton, NJ.
- Wood, A. W., 1955, *Lehrbuch der Kristallphysik*: Bell, London.
- Zimmerman, R. W., 1991, *Compressibility of Sandstones*: Elsevier, Amsterdam.

Fourier methods of seismic data regularisation

Chris Leader and Sjoerd de Ridder

ABSTRACT

Acquisition geometries of seismic data invariably lead to an irregular sampling in space of midpoint, offset and azimuth, however many seismic processing theories and algorithms work best if not exclusively on regularly sampled data. Here, three methods of Fourier regularisation are implemented and contrasted each with their own benefits and drawbacks. Standard theory is extended to the use of discrete spheroidal prolate sequence multitapers to help to diminish spectral contamination caused by an irregular input grid. Early results show that the use of multitapers alone have not yet proved to adequately reconstruct the input data, however the use of these to augment other types of Fourier regularisation have shown promise in increasing accuracy and reducing the number of required iterations.

INTRODUCTION

Regularisation of input data is required, or at least assumed, for a variety of advanced seismic data processing methods. Consequently if the input data is irregularly sampled then the effectiveness of many methods are compromised. Particularly algorithms such as 3D surface related multiple elimination (SRME) and pre-stack wave-equation depth migration have been shown to suffer considerably when applied to irregular data, and often induce new processing artifacts (Canning and Gardner, 1996). Furthermore irregular sampling can also lead to poor repeatability between 4D seismic surveys (at least for conventional 4D processing). To help alleviate these problems a common seismic data processing step involves some form of regularisation and/or interpolation.

Whilst precise definitions can vary, the concept of regularisation involves transferring samples from their irregular input grid to new locations on a regular output grid. Often interpolation will follow regularisation, whereby missing samples in the output grid are filled in; this is also often used to interpolate to a finer grid. A regularisation operator that can handle aliased and non-stationary events is desired, as this will assist in the preservation of all signal.

Several advanced regularisation methods have been suggested to overcome the problem of irregularly sampled data. One such technique is to use convolution operators known as prediction error filters (PEFs) that can predict absent traces honouring the spectra of known traces, however this can be limited due to an assumption of local linearity (Abma, 1995). As discussed in Biondi (2005) the problem can also be solved by AMO regularisation, whereby data is reconstructed at an arbitrary (regular) azimuth.

Fourier based methods can also be used, where the regularisation is performed by preserving the spectrum of the data and inverse transforming to a desired grid; however this suffers from the phenomena known as spectral leakage. The problem arises due to the non-orthogonality of the global basis functions (the *sinc* function) that are exhibited by an irregular grid (Xu and Pham, 2004a) and is discussed in detail in the next section. This phenomenon has been addressed by many disciplines that involve time series analysis, and several solutions to remove or reduce spectral leakage have been created.

A further possibility is to design an appropriate taper to minimise spectral leakage about a certain frequency or band of frequencies, and these are known as Slepian tapers, as discussed in Slepian (1978) and Percival and Walden (1993). The concept is that the taper parameters are designed to maximise spectral concentration within a predefined resolution bandwidth window, and so the use of these should help to greatly reduce the number of iterations required in other Fourier regularisation methods, or possibly replace them altogether. Such methods have been shown to work on random signals, synthetic earthquake signals and for bathymetry profiles (Prieto, 2007), and so the extension of this concept to active seismic data could help to solve the problem of Fourier regularisation.

SPECTRAL LEAKAGE

The occurrence of spectral leakage is a fundamental product of data reconstruction theory, which can be summarised to say that if a continuous function is sampled by at least the Nyquist interval, such that $\Delta t \leq \pi/\omega_{Ny}$, then total reconstruction of the input series is completely determined. However a fundamental assumption of this theory is that the basis functions used for reconstruction are orthogonal, and to summarise this assumption can be described as

$$\langle \varphi_k, \varphi_l \rangle = \int_{-\infty}^{\infty} \text{sinc}\left(\frac{x}{T} - k\right) \text{sinc}\left(\frac{x}{T} - l\right) dx, \quad (1)$$

$$= \text{sinc}(k - l) = \delta_{kl} \quad (2)$$

where δ_{kl} is the Kronecker delta function and k, l are integers (Xu and Pham, 2004b). On a regular grid this ensures that the sinc function has value 1 on the original data position, and 0 elsewhere. However, on an irregular grid this condition is violated, and indeed the Nyquist concept is less meaningful. The non-orthogonality of these basis functions leads to a non-orthogonality of Fourier coefficients, and this causes energy from large Fourier coefficients to ‘leak’ to others in the frequency domain, resulting in a denser, noisier spectrum. An example of this can be seen in Figure 2.

In light of fact that for real data there is always irregularity, whether it is due to cable feathering, obstacles or statics, Fourier regularisation methods must try and reduce or remove spectral leakage. To do this an irregular Fourier transform, its inverse and an appropriate weighting scheme must be devised.

IRREGULAR FOURIER TRANSFORM

The regular unitary discrete Fourier transform can be expressed as follows

$$F(k) = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} f(x_l) e^{-2\pi i k x_l}, \quad k = 0, \pm 1, \pm 2, \dots, \pm N \quad (3)$$

where $F(k)$ denotes the size of the k^{th} Fourier coefficient, x_l is the position of the input data and $f(x_l)$ denotes the value of the input data series corresponding to this position. To extend this expression to handle an irregular x_l series we must define a new transform, and care must be taken with the transform weights to ensure the forward and inverse transforms are unitary; this is simply done by dividing by \sqrt{N} in the regular case.

The weights must be related to the relative positioning of x_l on the input axis, with some form of normalisation. This can be done by letting

$$F(k) = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} f(x_l) \frac{\Delta x_l}{\Delta x_{av}} e^{-2\pi i k x_l}, \quad (4)$$

$$= \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} \frac{1}{\Delta x_{av}} \sum_{l=0}^{N-1} f(x_l) \Delta x_l e^{-2\pi i k x_l}, \quad (5)$$

$$= \frac{\sqrt{N}}{\Delta X} \sum_{l=0}^{N-1} f(x_l) \Delta x_l e^{-2\pi i k x_l}, \quad (6)$$

ΔX is the range of the input axis, $x_N - x_1$. The weighting here has taken into account the data range, size, and relative irregularity. Upon multiple tests it is easy to show that this transform is indeed unitary, and correctly preserves amplitudes after hundreds of domain transformations. The Δx_l values are calculated using a symmetric differencing star for stability, such that

$$\Delta x_l = \frac{1}{2}(x_{l+1} - x_{l-1}) \quad (7)$$

a weighting scheme for applying the transform in multiple dimensions is more difficult to estimate, but the problem can be broken down and each dimension considered separately for this application.

However designing an appropriate weight does not alleviate spectral leakage. One first method used was setting up the problem as a least squares inverse, using the Cauchy norm of $F(k)$ as a regularisation term within the cost function (Saachi and Ulrych, 1995), however a key problem with this approach was that the data is not sufficiently honoured.

ANTI-LEAKAGE FOURIER TRANSFORM

The first documented solution to the leakage problem was performed by Hogbom (1974) as a solution within Astrophysics. The situation was similar to that of seismic data regularisation

- radio interferometric baselines after aperture synthesis did not fall on a regular grid as desired, and the adaptive process known as ‘The CLEAN Algorithm’ was devised.

An adaptation of this can give a robust solution to Fourier regularisation, as discussed by Xu and Pham (2004a), Xu and Pham (2004b) and Schonewille (2009), and has been named the ‘anti-leakage Fourier transform’ due to its ability to suppress spectral leakage. This is an iterative algorithm, which exploits the fact that leakage is caused by high energy components contaminating other Fourier components. The maximum component of a given spectrum is calculated and added to an ‘estimated’ spectrum, then the contribution of this component is removed from the input data in data space, by inverse transforming the single component (taking care to reconstruct the same irregular axis) and subtracting. The new spectrum will see this peak removed, and then the next highest energetic component is found, added to the estimated spectrum and removed from the input data, and then iterations continue. The number of iterations is problem dependent, but the maximum required to rebuild a leakage free spectrum will be N , the number of Fourier components, and often far fewer will be required. The final spectrum is free of leakage, however some attention must be paid to Fourier artifacts; in this case cosine tapering is applied to the edges of the data

This algorithm can be summarised as follows

1. Compute the irregular, weighted DFT of the input data
2. Select the strongest Fourier component
3. Add this value to the estimated spectrum
4. Inverse irregular DFT this component (onto the irregular input axis)
5. Subtract this component from the input data
6. Iterate until the desired number of Fourier components have been estimated

It is clear from this set up that this algorithm is very limited by problem size. If the input axis is of length N and N_k components are to be estimated then $O(NxNxN_k)$ operations are required. Weights are only axis dependent so these are calculated and stored outside of the main loop.

One Dimensional Example

The power of the ALFT is best observed using a 1-D example and examining how the input series and corresponding spectra evolve with iterations. A 1-D irregular axis is constructed, where the irregularity is limited to half a bin size about the corresponding ‘regular’ position such that $x(ix) = ox + (ix - 1) * dx + (rand - 0.5) * dx$, where ix in the position index, ox an arbitrary origin, dx the receiver/trace spacing and $rand$ a random number between 0 and 1. This is representative of the problem for land data.

Using this axis a superposition of sine waves is created of varying frequencies, amplitude and phase. For the presented example two different waves are used for viewing simplicity.

Figure 1: Part of the irregularly sampled sine wave superposition used as input data/ [ER] chris1/. input

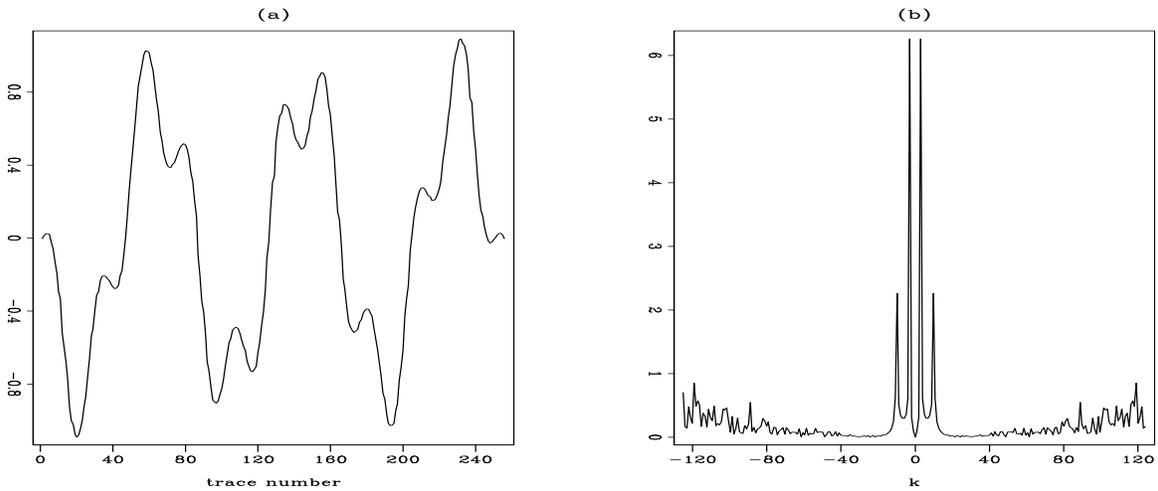
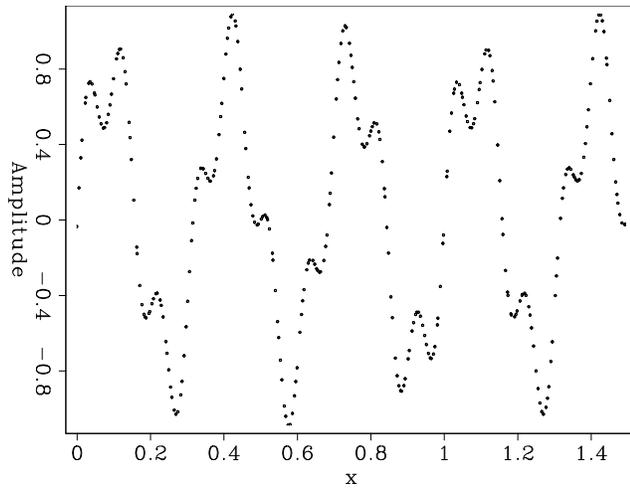


Figure 2: A section of the DFT reconstructed sine wave and its corresponding spectrum. [ER] chris1/. 1ddata

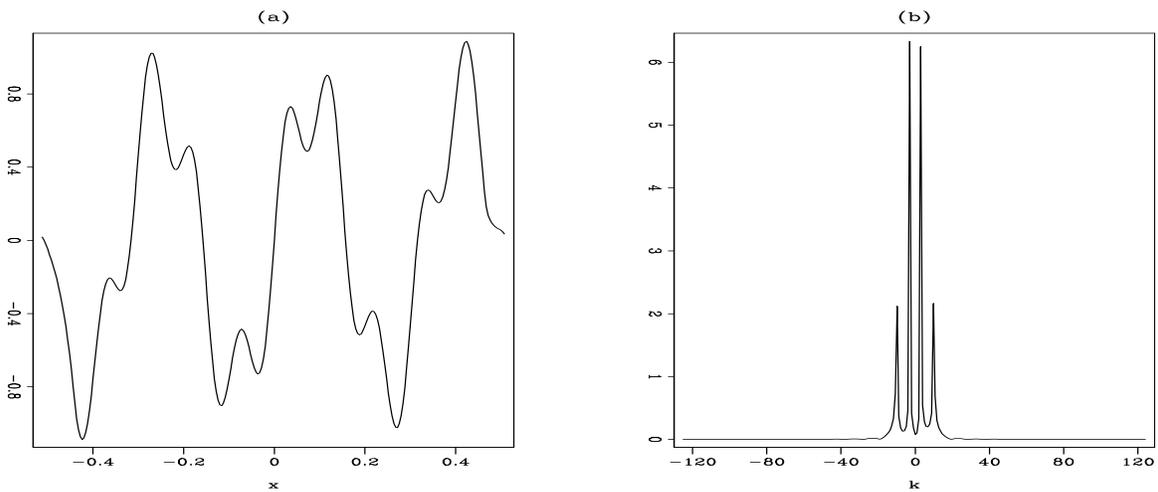


Figure 3: A section of the ALFT reconstructed sine wave and its corresponding spectrum. [ER] chris1/. 1dalft

As can be seen in Figure 3 the ALFT has perfectly reconstructed a regular signal from the irregular input, and the spectrum is free from leakage except for adjacent to the main peaks. In this oversimplified example only four iterations are really required. Also note the amplitude preservation between Figure 2 and Figure 3. For further insight a more complicated, higher dimensional example is required.

Two Dimensional Plane Wave Example

Next a 2-D synthetic data example was created. A series of plane waves of different dips and frequency contents was used and can be seen in Figure 4, along with the corresponding F-K spectrum in Figure 6. One event is strongly aliased. The input data was regularised using both a standard nonuniform DFT approach and the ALFT algorithm, and the results and F-K spectra can be seen in Figure 5

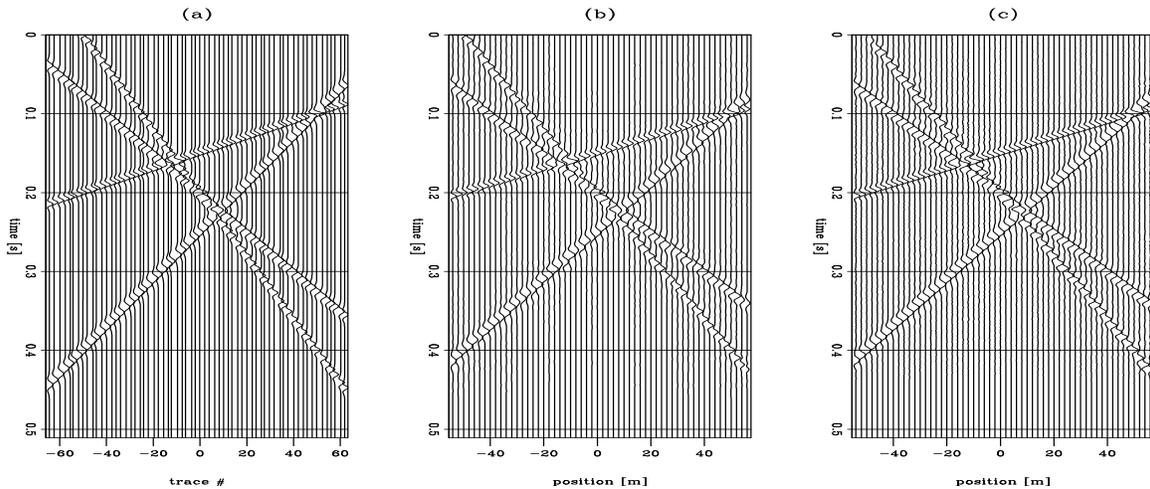


Figure 4: Input irregular data (a), DFT regularisation result (b) and ALFT regularisation result (c). [ER] `chris1/. 2dcomp`

It is apparent that the anti-leakage Fourier transform gives more uniform amplitudes and improves event alignment when compared to using a nonuniform DFT. Some small artifacts are still induced from multiple Fourier transformations, however these artifacts are not more pronounced than in the DFT case. The spectra show that most of the incoherent noise in the ALFT regularised image is induced from the aliased event, giving a mostly clean spectrum elsewhere. Wrongly dipping events and noise in the corresponding DFT spectrum is significantly reduced in the ALFT spectrum.

QDome Data Example

To use a more realistic seismic example the QDome synthetic data from Claerbout's Image Estimation By Example (Claerbout, 1995) was adapted to contain an irregular x-axis, details of the data computation can be found in Claerbout (1995). This dataset features flat beds at the surface, a Gaussian centre with a fault and dipping beds at the bottom - this model is useful since it incorporates events that are flat, dipping, curving, aliased and

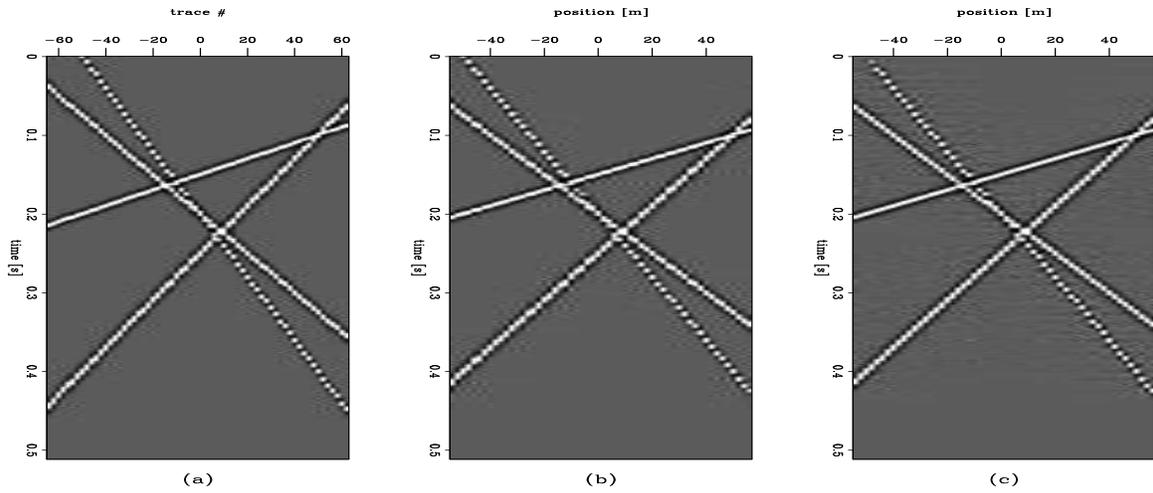


Figure 5: Input irregular data (a), DFT regularisation result (b) and ALFT regularisation result (c). [ER] `chris1/. 2dcompg`

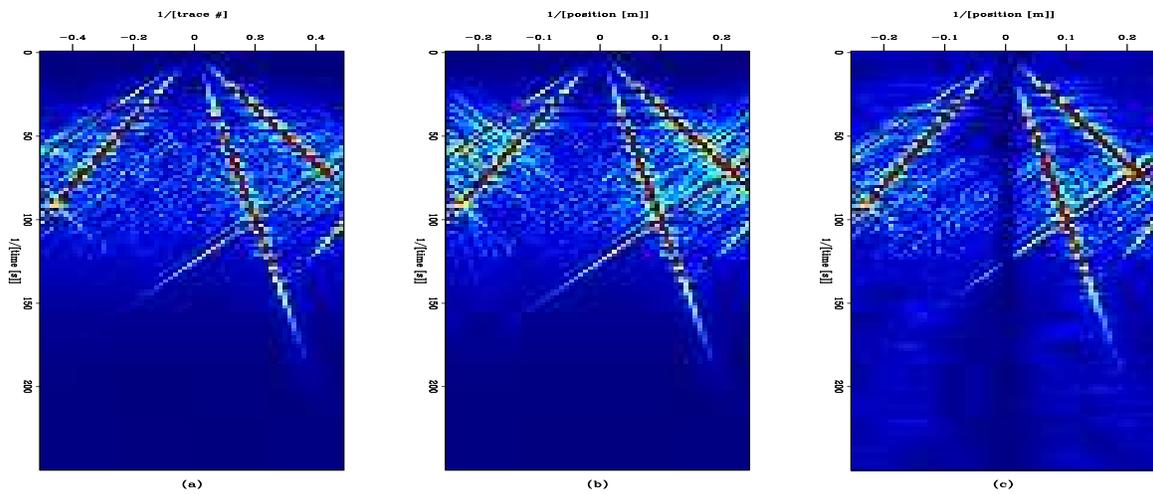


Figure 6: FK Spectra of the input data (a), data with DFT regularisation (b) and data with ALFT regularisation (c). [ER] `chris1/. allspec`

discontinuous, and so is an appropriate model to test the robustness of the algorithm.

An irregular axis was created as in the previous examples and the QDome data was calculated for this irregular axis. A 3D cube of data was created and presented herein are two slices from the cube orthogonal to the (regular) y axis - one slice containing strongly aliased events and one containing a discontinuous fault

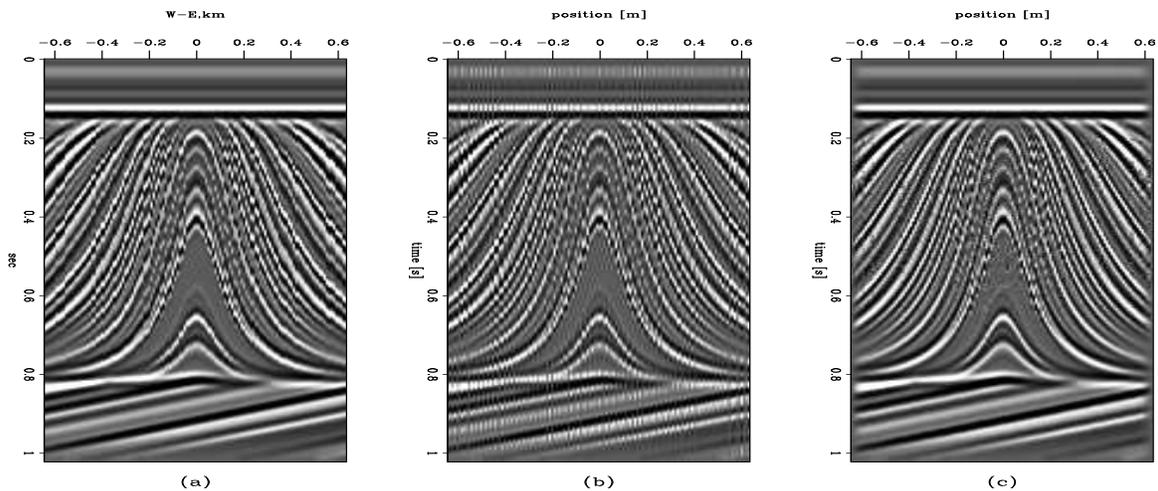


Figure 7: QDome slice containing aliased data (a), DFT regularisation result (b) and ALFT regularisation result (c). [ER] `chris1/. aliasg`

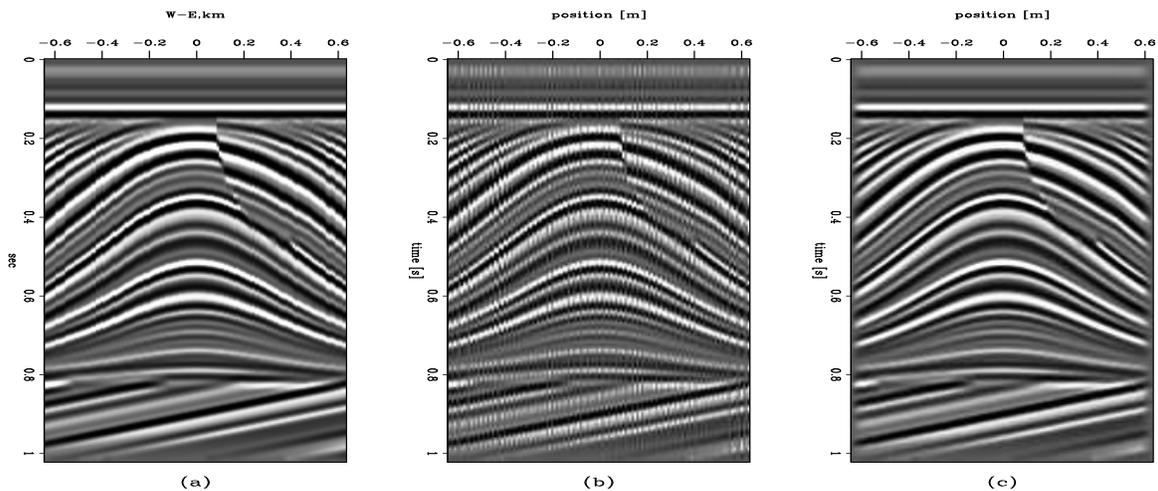


Figure 8: QDome slice containing a fault (a), DFT regularisation result (b) and ALFT regularisation result (c). [ER] `chris1/. faultg`

The ALFT shows a much improved result in both of these cases. In this example the DFT particularly struggles with the amplitudes of the traces and some appear much diminished - these are the traces that were the most irregular. The fault appears more continuous and defined in the ALFT result than in the irregular input image, as do the step events

in Figure 8. However the algorithm still struggles somewhat in the aliased sections of the data, and the spectrum (as seen in Figure 7), is noisier in aliased regions. Schonewille (2009) discusses how this problem can be addressed, and this can be done by extrapolating the known spectrum to higher ω and k values, estimating weights for these higher frequencies and computing the ALFT for these.

Whilst the results of using the ALFT show a marked improvement over DFT regularisation and a robustness in the presence of aliased and discontinuous events it is fundamentally limited by its reliance on multiple domain transformations. For large datasets the number of operations required by this technique is unreasonable and so a different method, or maybe one that augments the ALFT to require fewer transforms, is desired. It is possible that the answer to this problem lies in tapering the input data in such a way that Fourier components can be made more orthogonal. This is discussed in the subsequent section.

SLEPIAN TAPERS

Equation 6 can be slightly modified to include a new series of weights, or tapers, $a(x_l)$,

$$F_w(k) = \frac{\sqrt{N}}{\Delta X} \sum_{l=0}^{N-1} f(x_l) a_w(x_l) \Delta x_l e^{-2\pi i k x_l}, \quad (8)$$

if $a(x_l)$ is a boxcar then equation 8 will reduce to 6. It is possible to design a series of tapers $a(x_l)$ in order to reduce spectral leakage, and this is a classical time series analysis solution. Each w subscript indicates a possible different taper.

Thomsen (1982) proposed the multitaper method, which is now widely used in geophysical applications. In this method $a(x_l)$ represents a set of orthogonal tapers that can be averaged over to give an estimation of the power spectral density function (the absolute value squared of the Fourier sequence.) This can outperform using a single taper method for several reasons; one being that using a single taper can discard a significant amount of the signal. When using the multitaper method this is less of a problem since information discarded by one taper can be (at least partially) recovered by the others (Prieto, 2007).

Slepian (1978) showed that by stipulating a requirement of spectral leakage minimisation the design of the tapers becomes an eigenfunction problem. Ultimately the end result resembles

$$\mathbf{D}\mathbf{a} - \lambda\mathbf{a} = \mathbf{0}, \quad (9)$$

where \mathbf{D} is the matrix described by:

$$D_{n,m} = \frac{\sin 2\pi W(n-m)}{\pi(n-m)} \quad n, m = 0, 1, 2, \dots, N-1, \quad (10)$$

where W is a predefined resolution bandwidth such that $1 < NW < N/2$, and NW is known as the time bandwidth product (or the number of Rayleigh bins) and controls the amount

of smoothing in the frequency domain. λ here represents a measure of spectral concentration and \mathbf{a} the set of eigenfunctions associated with (4), which will be the components of the multitaper estimate (Percival and Walden (1993)). The closer that λ is to 1 then the greater the expected suppression of spectral leakage is, and indeed $1 - \lambda$ is a good measure of taper 'strength' and will be a core component of the weighting scheme used. The resulting sequence of tapers are known as Slepian tapers, or Discrete Prolate Spheroidal Sequences (DPSSs), each DPSS is dependent on the choice of NW . The first four tapers in the sequence $NW=4$ are shown in Figure 9.

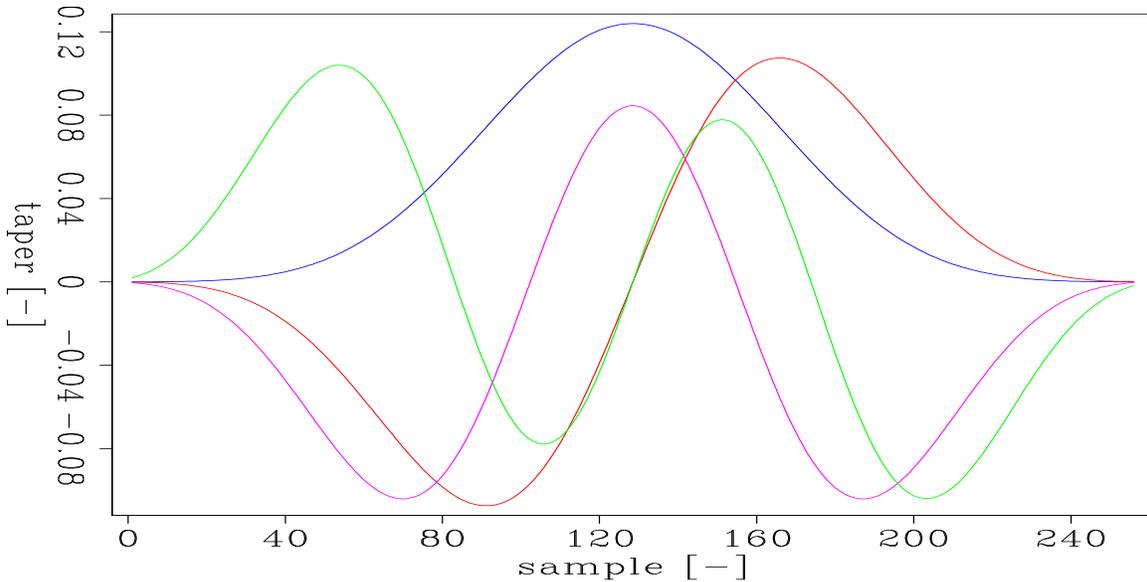


Figure 9: The first 4 Slepian tapers calculated with $NW=4$, also known as the 4π Slepian sequence. [ER] `chris1/. tapers`

Using a very similar method Prieto (2007) showed that this can be effective in determining and quantifying the spectra of earthquake and bathymetry data. Whilst it is true that estimating the tapers from the data is computationally intensive, once estimated the tapers are stored and reused (if iterations are required,) and so a substantially improved run time relative to the ALFT is shown.

As can be seen from Eq. (5) the matrix \mathbf{D} is dense and symmetric. This symmetry exists because the matrix can be rewritten as $D_{n,m} = 2W \text{sinc } 2\pi W(n - m)$ and the sinc function is symmetric. The eigenfunctions need to be found and so \mathbf{D} is reduced to tridiagonal form using Householder's reduction. Then using QL decomposition the eigenvalues and eigenvectors of this real, symmetric, tridiagonal matrix can be computed and hence the tapers calculated (Sparks and Todd, 1973).

However the manner in which these tapers are computed does not easily lead to calculating the tapers over an irregular grid, and since the tapers are applied within the forward Fourier transform kernel they must lie on the same grid as the input data. This problem can be solved by creating a very dense taper and then subsampling to this to the input

transformation axis. Moreover this process is made more accurate by linearly interpolating the irregular taper value between the two known values on the dense grid. Since the tapers only require calculating once for a given time series length and time bandwidth product, creating a dense series of tapers does not cause any significant computation time.

One issue that requires much thought is the design of weighting functions for the procedure that will not cause local bias within the data. A well designed weighting function will generate a smooth estimate from the multitaper scheme, with less variance than the corresponding single taper method and also will try and reduce any local bias. Several methods have been suggested such as that by Thomsen (1982), Riedel and Sidorenko (1995) and Prieto (2007). The method used herein follows the derivation by Thomsen, where the taper weights are frequency dependent and are calculated using λ .

$$|F(k)|^2 = \frac{\sum_{w=0}^{N-1} d_k^2 |F_w(k)|^2}{\sum_{w=0}^{N-1} d_k^2}, \quad (11)$$

$$d_w(k) = \frac{\sqrt{\lambda_k} |F_w(k)|^2}{\lambda_k |F_w(k)|^2 + (1 - \lambda_k) \sigma^2} \quad (12)$$

$$(13)$$

where as before $F_w(k)$ the result of Fourier transforming the product of the input series with the taper a_w , λ is the eigenvalue of the given taper and σ is the variance of the input series. This can be proved by seeking to estimate the discrepancy of the different eigencomponents, $F_w(k)$, and an involved proof can be seen in Percival and Walden (1993).

These classical weights are designed for weighting the power spectral density function (PSD), which is the square of the amplitudes of the spectrum. However in taking the power of the spectrum the phase information is lost, and so the phase spectrum is computed and stored before the PSD is computed and weighted, and now the weighted spectrum can be (roughly) estimated by combining the root of the power with the unweighted phases.

Results using Slepian tapers

The use of Slepian tapers for seismic data regularisation requires some initial ingenuity, since previously their purpose has been in estimating the spectra of complex processes and not in reconstructing data. Figure 10 shows a windowed irregular spectrum containing two main frequencies and its equivalent spectrum after applying four π Slepian tapers and weighting appropriately. The irregular spectrum indicates leakage has occurred, both around the peak and at higher frequencies, whilst the spectrum after tapering shows less contamination around the frequencies of interest and an overly cleaner spectrum. In this case the tapering has reduced the spectral leakage with little computation time. It should be noted that the ALFT can often not remove the leaked frequencies surrounding the main peaks, whilst the DPSS tapers have reduced this.

The choice of number of tapers and the time bandwidth product is very situation dependent. In a lot of cases a poor choice of the time bandwidth product merely serves the

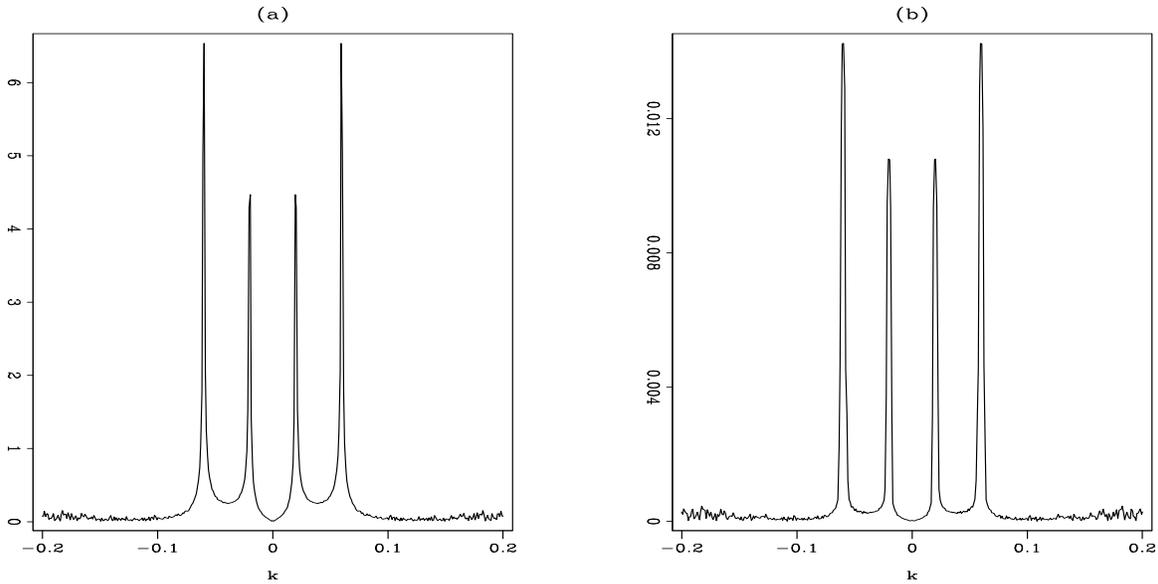


Figure 10: The result of applying the four tapers in the π DPSS sequence to a spectrum exhibiting leakage. [ER] `chris1/. tape`

smear spectral peaks and enhance leakage, even giving zero frequency energies.

The taper weights are designed to reduce the bias of the power spectrum, and not the amplitude. Using these weights on the amplitude gives peak splitting, which theory tells us should happen, and if the root of the power is used then phase information of the data is lost. Ongoing efforts involve calculating amplitude and/or phase weights to reduce bias, rather than within the power. Currently in data reconstruction taper imprints are still visible, which again was expected by theory. Increasing the number of tapers alters this problem, and a method of amplitude preservation is yet to be finalised. Currently the method outlined by Park (1992) where the reconstruction is posed as an inverse problem is being attempted.

CONCLUSIONS

This paper has shown that Fourier methods of seismic data regularisation can be a powerful and robust method in regularising data, which is useful for future processing steps and for image viewing. A standard irregular DFT has limited application since the contaminated spectra lead to a large number of artifacts and amplitudes are not reconstructed correctly. Using a recursive ALFT approach gives a much improved result - amplitudes are now correct along events, events are exactly aligned in the correct places and there is a slight reduction in artifacts. Discontinuous events are imaged and whilst aliased events are not perfectly reconstructed they do not contaminate the image as strongly as in the standard DFT case, and this method can be extended to handle these sorts of frequencies. However the ALFT approach is slow, and could be improved by implementing a non-uniform FFT and better tapering between transforms, to help reduce artifacts.

Examination of multitaper methods with Slepian tapers has yielded some intriguing

results - the spectra after tapering and stacking are shown to have reduced spectral leakage without iterations around frequencies of interest, however leakage noise at higher frequencies is preserved.

Key issues that still require resolving before this is proved to be a viable regularisation method are those of adapting taper weights to optimise the amplitude spectrum and not the power, a better method of reconstructing and weighting the phase spectrum and a more robust method of removing the taper imprints in the reconstructed data. Early tests are showing that these problems are gradually being overcome.

ACKNOWLEDGMENTS

I would like to thank Jim Berryman for his insightful discussions about spectral tapers and his useful literary recommendations.

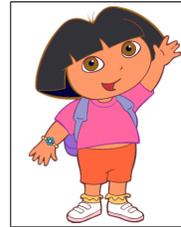
REFERENCES

- Abma, R. C. J., 1995, Lateral prediction for noise attenuation by t-x and f-x techniques: *Geophysics*, **60**, 1887–1896.
- Biondi, B., 2005, 3d seismic imaging: Society of Exploration Geophysicists.
- Canning, A. and G. Gardner, 1996, Another look at the question of azimuth: *The Leading Edge*, **15**, 821–823.
- Claerbout, J., 1995, Image estimation by example: Society of Exploration Geophysicists.
- Hogbom, J., 1974, Aperture synthesis with a non-regular distribution of interferometer baselines: *Astronomy and Astrophysics*, **15**, 417–426.
- Park, J., 1992, Envelope estimation for quasi-periodic geophysical signals in noise: a multitaper approach: Edward Arnold, London.
- Percival, D. and A. Walden, 1993, Spectral analysis for physical applications: Cambridge University Press.
- Prieto, G. e. a., 2007, Reducing the bias of multitaper spectrum estimates: *International Journal of Geophysics*, **171**, 1269–1281.
- Riedel, K. and A. Sidorenko, 1995, Minimum bias multiple taper spectral estimation: *IEEE Transactions on Signal Processing*, **43**, 188–195.
- Saachi, M. and T. Ulrych, 1995, High-resolution velocity gathers and offset space reconstruction: *Geophysics*, **60**, 1169–1177.
- Schonewille, M. e. a., 2009, Seismic data regularisation with the anti-alias anti-leakage Fourier transform: *The First Break*, **27**, 85–92.
- Slepian, D., 1978, Prolate, spheroidal wavefunctions, Fourier analysis and uncertainty v: the discrete case: *Bell System Technical Journal*, **57**, 1371–1430.
- Sparks, D. and A. Todd, 1973, Latent roots and vectors of a symmetric matrix: *Applied Statistics*, **22**, 260–265.
- Thomsen, D., 1982, Spectrum estimation and harmonic analysis: *Proceedings of the IEEE*, **70**, 1055–1096.
- Xu, S. and D. Pham, 2004a, Seismic data regularisation with anti-leakage Fourier transform: EAGE 66th Conference and Exhibition.
- Xu, S. Z. Y. and D. Pham, 2004b, On the orthogonality of anti-leakage Fourier transform based seismic trace interpolation: SEG Int'l Exposition and 74th Annual Meeting.

SEP-140

271

Mandy Wong graduated in 2004 with a B.Sc. in Physics and Mathematics from the University of British Columbia (UBC) in Vancouver, Canada. In 2006, she obtained a M.Sc. degree in Condensed Matter Theory at UBC. Afterward, Mandy worked for a geophysical consulting company, SJ Geophysics, based in Vancouver, Canada. Mandy joined SEP in 2008, and is working towards a Ph.D. in Geophysics. Her main research interest is imaging with multiples.



Yang Zhang graduated from Tsinghua University in July 2007 with a B.E. in Electrical Engineering. He took an internship in Microsoft Research Asia during 2007-2008. He joined SEP in 2009, and is currently pursuing a Ph.D. in Geophysics.



SEP PHONE DIRECTORY

Name	Phone	Login Name
Almomin, Ali	723-3187	ali
Ayeni, Gboyega	723-6006	gayeni
Barak, Ohad	723-3187	ohad
Berryman, James	–	berryman
Biondi, Biondo	723-1319	biondo
Cardoso, Claudio	723-1250	claudio
Claerbout, Jon	723-3717	jon
Clapp, Bob	725-1334	bob
de Ridder, Sjoerd	723-1250	sjoerd
Guitton, Antoine	–	antoine
Halpert, Adam	723-6006	adam
Lau, Diane	723-1703	diane
Leader, Chris	723-0463	chris1
Li, Elita	723-9282	myfusan
Moussa, Nader	723-0463	nwmoussa
Shen, Xukai	723-0463	xukai
Tang, Yaxun	723-1250	tang
Taweesintananon, Kittinat	723-3187	kittinat
Wong, Mandy	723-9282	mandyman
Zhang, Yang	723-3187	yang

SEP fax number: (650) 723-0683

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, 2009-2010

Name	Company	Tel #	E-Mail
Raymond Abma	BP	(281)366-4604	abmar1@bp.com
Biondo Biondi	SEP	(650)723-1319	biondo@sep.stanford.edu
Robert Bloor	ION/GX Technology	(281)781-1141	robert.bloor@iongeo.com
Jon Claerbout	SEP	(650)723-3717	jon@sep.stanford.edu
Richard Cook	Shell	(713)245-7195	richard.cook@shell.com
Stewart A. Levin	Halliburton	(303)488-3062	stewart.levin@halliburton.com
Yi Luo (Co-chair, 2nd year)	Saudi Aramco	(966)38738520	yi.luo@aramco.com
Marta Woodward (Co-chair, 1st year)	WesternGeco	(832)798-6204	martajo@slb.com

Research Personnel

Ali Almomin graduated from Texas A&M University in 2007 with a BS in Geophysics. Then, he joined Saudi Aramco and worked in several exploration and research departments with a focus on 3D seismic processing and near surface imaging. He joined Stanford Exploration Project in 2009 to pursue a PhD in Geophysics and is currently working on seismic tomography. He is a member of SEG, EAGE, and SPE.



Gboyega Ayeni received his B.Sc. in Applied Geophysics from Obafemi Awolowo University, Nigeria in 2004. He was a Shell scholar at University of Leeds, where he obtained an M.Sc with Distinction in Exploration Geophysics. Gboyega joined SEP in September 2006 to work towards his Ph.D in Geophysics. He is a member of SEG, EAGE, AGU, SPE and AAPG.



Ohad Barak received a B.Sc. (2006) and an M.Sc. (2009) in Geophysics from Tel-Aviv University. In 2008 he joined Paradigm Geophysical and worked there as a developer. He joined SEP in 2009 and is currently pursuing a Ph.D. in geophysics at Stanford University, and a longer biography.



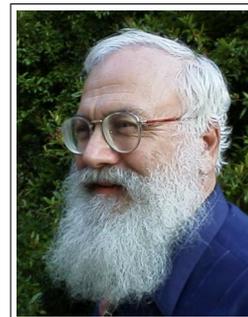
James G. Berryman received a B.S. degree in physics from Kansas University (Lawrence) in 1969 and a Ph.D. degree in physics from the University of Wisconsin (Madison) in 1975. He subsequently worked on seismic prospecting at Conoco. His later research concentrated on seismic waves in rocks and sediments – at AT&T Bell Laboratories (1978-81) and at Lawrence Livermore National Laboratory (1981-), where he is currently a physicist in the Energy and Environment Directorate. He received the Maurice Anthony Biot Medal of the ASCE in May, 2005, for his work in the mechanics and acoustics of porous media containing fluids. Continuing research interests include acoustic, seismic, and electrical methods of geophysical imaging and studies of waves in porous media. He is a member of APS, AGU, ASA, and SEG.



Biondo L. Biondi is Associate 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. In 2004 the Society of Exploration Geophysicists (SEG) has honored Biondo with the Reginald Fessenden Award. Biondo recently published a book, 3-D Seismic Imaging, that is the first text book to introduce the theory of seismic imaging from the 3-D perspective. The book is published by SEG in the Investigations in Geophysics series. During 2007 gave a one-day short course in 28 cities around the world as the SEG/EAGE Distinguished Short Course Instructor (DISC) . He is a member of AGU, EAGE, SEG and SIAM.



Jon F. Claerbout (M.I.T., B.S. physics, 1960; M.S. 1963; Ph.D. geophysics, 1967), professor at Stanford University, 1967. Emeritus 2008. Best Presentation Award from the Society of Exploration Geophysicists (SEG) for his paper, *Extrapolation of Wave Fields*. Honorary member and SEG Fessenden Award “in recognition of his outstanding and original pioneering work in seismic wave analysis.” Founded the Stanford Exploration Project (SEP) in 1973. Elected Fellow of the American Geophysical Union. Authored three published books and five internet books. Elected to the National Academy of Engineering. Maurice Ewing Medal, SEG’s highest award. Honorary Member of the European Assn. of Geoscientists & Engineers (EAGE). EAGE’s highest recognition, the Erasmus Award.



Robert Clapp received his B.Sc. (Hons.) in Geophysical Engineering from Colorado School of Mines in May 1993. He joined SEP in September 1993, received his Masters in June 1995, and his Ph.D. in December 2000. He is a member of the SEG and AGU.



Claudio Guerra received his B.Sc. in Geology from Federal University of Rio de Janeiro, Brazil in 1988 and a M.Sc. from State University of Campinas, Brazil in 1999. Since 1989, he has been working for Petrobras, Brazil. He joined SEP in 2006 and is currently pursuing a Ph.D. in geophysics at Stanford University. He is member of SEG and SBGf.



Adam Halpert graduated from Rice University in May 2007 with a B.Sc. in Earth Science and a B.A. in Policy Studies. He joined SEP in the summer of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG and AGU.



Benxi Ke is currently a visiting scholar at SEP. He graduated from Dalian University of Technology in 1983 with a B.Sc. in applied Mathematics. He received His M.Sc. in applied geophysics from China University of petroleum (Huadong) and is currently working towards a Ph.D. in solid geophysics at Institute of Geology and Geophysics, Chinese Academy of Science. He is a Professor at BGP, CNPC since 2003. He is an active member of SEG, a number of CPS and a member of editorial board of the journal Oil Geophysical prospecting (China).



Chris Leader graduated from Oxford Univeristy in 2008 with a BA in Physics (with concentration on Astrophysics and Condensed Matter physics) and then from Imperial College London in 2009 with an MSc in Petroleum Geophysics (Distinction). He is currently a first year student in the Stanford Exploration Project on the PhD program working on Fourier methods of regularisation. Work experience involves 3D seismic processing for a Rio Tinto acquired dataset over summer 2009. He is a member of SEG, EAGE, PESGB and IOP.



Yunyue (Elita) Li graduated from China University of Petroleum, Beijing in July 2008 with a B.S. in Information and Computational Science. She joined SEP in the fall of 2008, and is currently working toward a Ph.D. in Geophysics. She is a student member of the SEG.



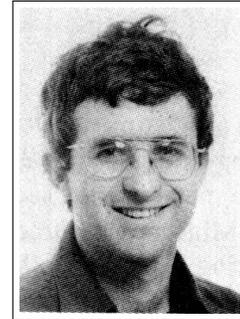
Nader Moussa graduated from North Carolina State University, where he studied Physics and Electrical and Computer Engineering. In 2008, he finished his M.Sc. at Stanford in Electrical Engineering, with focus on experimental radioscience and electromagnetic sensor instrumentation design. He will be leaving Stanford Exploration Project in 2010 to pursue a position as the chief executive for Nimur Enterprises, a privately funded technology research corporation.



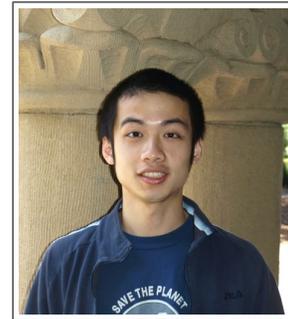
Sjoerd de Ridder received his B.Sc. (Dec. 2004) in Earth Sciences from Utrecht University in the Netherlands. During in the fall of 2004 and the spring of 2005 he was a graduate exchange student in geophysics at Colorado School of Mines. He obtained an M.Sc. (Aug. 2007) in Applied Earth Sciences with specialization in applied geophysics from Delft University of Technology in the Netherlands for his thesis: 'Simulation of interferometric seismoelectric Greens function recovery'. He joined SEP in the fall of 2007 and is currently pursuing a Ph.D. in Geophysics. He received the 'Jon F. Claerbout - Chevron Fellowship' in 2009-2010. During the summer of 2010 he did an internship at China University of Petroleum, Beijing. Currently he is working on exploration seismology with low-frequency ambient seismic noise. In his free time he likes to ride his bike, discuss politics and history or travel.



Shuki Ronen is a geophysicist with Geco-Prakla working on various aspects of seismic data processing and acquisition. Previously he worked with GeoQuest on seismic data interpretation; with Schlumberger on reservoir characterization; with the Institute of Petroleum Research and Geophysics on seismic data processing and acquisition; with the Colorado School of Mines as a visiting professor; and with Saxpy Computer company as an engineer. He has a Ph.D. from Stanford in Geophysics, and a B.Sc. in Physics and Geology from Hebrew University. (Photo: December 1985, Geophysics, p. 2919)



Xukai Shen graduated from Tsinghua University in July 2007 with a B.E. in Electrical Engineering. He joined SEP in the fall of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG.



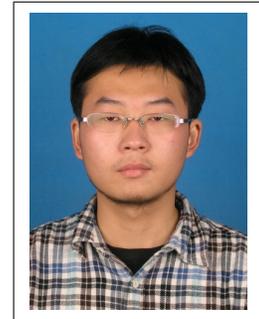
Yaxun Tang received his B.Sc. (Jul. 2003) and M.Sc. (Aug. 2005) in Geophysics from School of Ocean and Earth Science, Tongji University, Shanghai. He joined SEP in 2005 and is currently working towards a Ph.D. in Geophysics at Stanford University. He is a student member of SEG and EAGE.



Mandy Wong graduated in 2004 with a B.Sc. in Physics and Mathematics from the University of British Columbia (UBC) in Vancouver, Canada. In 2006, she obtained a M.Sc. degree in Condensed Matter Theory at UBC. Afterward, Mandy worked for a geophysical consulting company, SJ Geophysics, based in Vancouver, Canada. Mandy joined SEP in 2008, and is working towards a Ph.D. in Geophysics. Her main research interest is imaging with multiples.



Yang Zhang graduated from Tsinghua University in July 2007 with a B.E. in Electrical Engineering. He took an internship in Microsoft Research Asia during 2007-2008. He joined SEP in 2009, and is currently pursuing a Ph.D. in Geophysics.



SEP ARTICLES PUBLISHED OR IN PRESS

- Ayeni, G., A. Huck, and P. de Groot, 2008, Extending reservoir property prediction with pseudo-wells: First Break. **26**, 11–15.
- Ayeni, G., Y. Tang, and B. Biondi, 2009, Joint preconditioned least-squares inversion of simultaneous source time-lapse seismic data sets: SEG Technical Program Expanded Abstracts, **28**, 3914–3918.
- Ayeni, G. and Nasser, M., 2009, Optimized local matching of time-lapse seismic data: A case study from the Gulf of Mexico, SEG Technical Program Expanded Abstracts, **28**, 3939–3943.
- Ayeni, G., and Biondi, B., 2009, Joint target-oriented wave-equation inversion of multiple time-lapse seismic data sets, SEG Technical Program Expanded Abstracts, **28**, 3825–3829.
- Ayeni, G., Tang, Y., and Biondi, B., 2010, Efficient seismic monitoring of hydrocarbon reservoirs using multiple shooting vessels, Proceedings of the Offshore Technology Conference (Accepted).
- Ayeni, G., and Biondi, B., 2010, Continuous reservoir monitoring with asynchronous simultaneous-source seismic data, 72nd Conference & Technical Exhibition, EAGE, Extended Abstracts (Accepted).
- Ayeni G., and Biondi, B., 2010, Target-oriented joint least-squares migration/inversion of time-lapse seismic data sets, Geophysics (Accepted).
- de Ridder, S., Prieto, G.A., 2008, Seismic Interferometry and the Spatial Auto-Correlation Method on the Regional Coda of the Non-Proliferation Experiment: AGU Fall Meeting, S31A-1885.
- de Ridder, S., Green's function retrieval by iterated coda correlations, Eos, Trans. AGU, 90(52), Fall Meet. Suppl., Abstract S21A-1699.
- de Ridder, S., B. Biondi, G. Papanicolaou, 2009, Kinematics of iterative interferometry in a passive seismic experiment: SEG Technical Program Expanded Abstracts **28**, 1622-1626.
- de Ridder, S., N. Crook, S. S. Haines and S. T. Ide, 2010, Seismic Investigation of Underground Coal Fires; A Feasibility Study at the Southern Ute Nation Coal Fire Site, Symposium on the Application of Geophysics to Engineering and Environmental Problems **23**, 630–638.
- Halpert, A., Clapp, R.G., Lomask, J., and B. Biondi, 2008, Image segmentation for velocity model construction and updating: SEG Technical Program Expanded Abstracts **27**, 3088–3092.
- Halpert, A., Clapp, R.G., and B. Biondi, 2009, Seismic image segmentation with multiple attributes: SEG Technical Program Expanded Abstracts **28**, 3700-3704.
- Maysami, M., and F. J. Herrmann, 2008, Lithology constraints from seismic waveforms: application to opal-A to opal-CT transition: SEG Technical Program Expanded Abstracts **27**, 2011-2015.
- Shragge, J., 2008, Riemannian wavefield extrapolation: Nonorthogonal coordinate systems: Geophysics, **73**, T11–T21.
- Tang, Y., 2008, Wave-equation Hessian by phase encoding: SEG Technical Program Expanded Abstracts **27**, 2201-2205.
- Tang, Y. and B. Biondi, 2009, Least-squares migration/inversion of blended data: SEG Technical Program Expanded Abstracts **28**, 2859-2863.
- Tang, Y., 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian: Geophysics, **74**, WCA95-WCA107.

- Tang, Y. and S. Lee, 2010, Efficient Gauss-Newton Hessian for full waveform inversion: 72nd Conference & Technical Exhibition, EAGE, Extended Abstracts (Accepted).
- Vyas, M. and Y. Tang, 2010, On the gradient of wave-equation migration velocity analysis: 72nd Conference & Technical Exhibition, EAGE, Extended Abstracts (Accepted).

