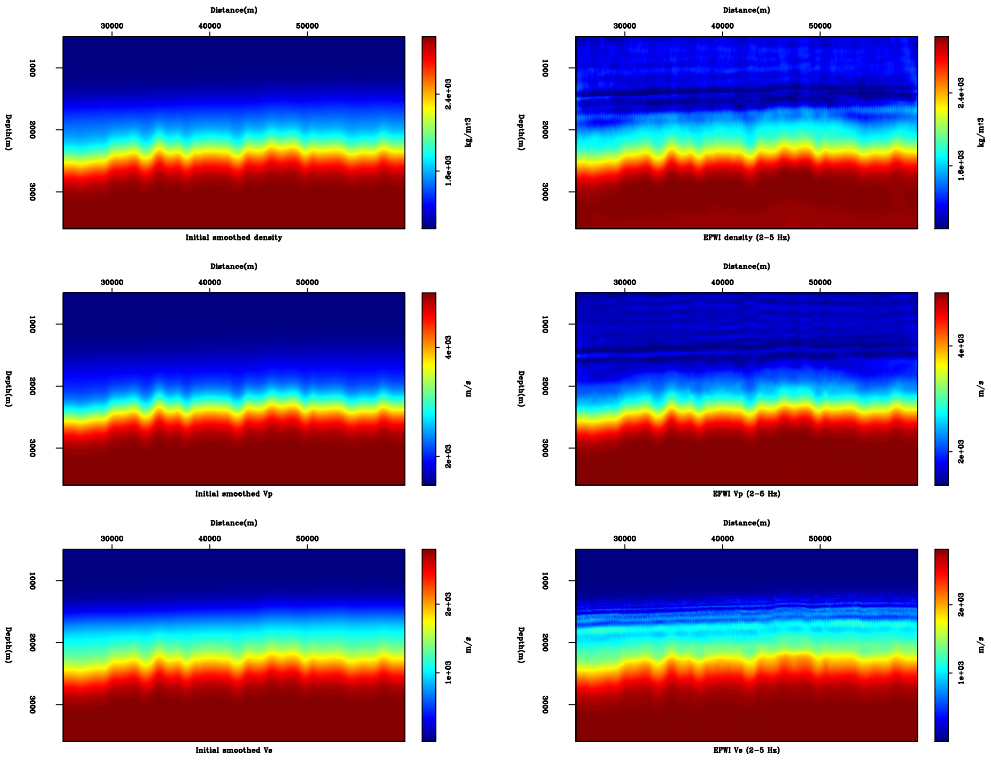
STANFORD EXPLORATION PROJECT

Gustavo Alves, Guillaume Barnier, Biondo Biondi, Ettore Biondi,
Alejandro Cabrales-Vargas, Jason Chang, Robert Cieplicki, Jon Claerbout, Robert Clapp,
Steve Cole, Taylor Dahlke, Eric Dunham, Stuart Farris, Jeremy Goldbogen, Fantine Huot,
Joseph Jennings, Martin Karrenbach, Huy Le, Stewart Levin, Yinbin Ma,
Musa Maharramov, Eileen Martin, Mark Meadows, Tapan Mukerji, Shuki Ronen,
Rahul Sarkar, and Leighton Watson

Report Number 168, April 2017



Preface

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

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

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

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

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

[1] http://sepwww.stanford.edu/private/docs/sep168
[2] http://sepwww.stanford.edu/public/docs/sepdatalib/toc_html
[3] http://sepwww.stanford.edu/research/redoc/

SEP168 — TABLE OF CONTENTS

Passive

Eileen Martin, Biondo Biondi, Steve Cole, and Martin Karrenbach, Overview of the Stanford DAS Array-1 (SDASA-1) 1

Biondo Biondi, Eileen Martin, Steve Cole, and Martin Karrenbach, Earthquakes analysis using data recorded by the Stanford DAS Array 11

Eileen Martin and Biondo Biondi, Ambient Noise Interferometry on Two-Dimensional DAS Arrays 27

Jason P. Chang, The search for P-waves at Forties 35

Leighton M. Watson, Shuki Ronen, Jeremy A. Goldbogen, and Eric M. Dunham, Comparing whales to marine seismic sources: low frequency sound generation by fin whales 51

Elastic imaging

Gustavo Alves, Elastic Full Waveform Inversion of the Moere Vest data . 63

Ettore Biondi, Guillaume Barnier, and Biondo Biondi, Preconditioned elastic full waveform inversion using approximated Hessian matrix 77

Gustavo Alves and Ettore Biondi, Reciprocity in elastic multi-component data 85

Stuart Farris, Ettore Biondi, and Guillaume Barnier, Phase Only Full Waveform Inversion of Elastic Data Using Acoustic Engine .. 95

Waveform inversion

Taylor Dahlke, Biondo Biondi, and Robert Clapp, Application of full Hessian on level set formulation and parameter inversion 101

Guillaume Barnier, Ettore Biondi, and Biondo Biondi, Tomographic Full Waveform Inversion for imaging under complex overburden .. 115

Yinbin Ma and Robert G. Clapp, An efficient time-lapse full waveform inversion by saving the wavefield at boundaries around the reservoir 141

Alejandro Cabrales-Vargas, Biondo Biondi, and Robert Clapp, Our progress towards linearized waveform inversion with velocity updating (LWIVU) 151

Ettore Biondi, Biondo Biondi, and Guillaume Barnier, Understanding the difference between full Newton and Gauss-Newton approximation of the full waveform inversion Hessian: a short note 169

Rahul Sarkar and Biondo Biondi, Incorporating optimal transport in Tomographic Full Waveform Inversion : Theory 181

Anisotropy

Huy Le, Biondo Biondi, Robert G. Clapp, Stewart A. Levin, and Tapan Mukerji, Source estimation and rock physics preconditioning for anisotropic full waveform inversion 201

Yinbin Ma, Stewart A. Levin, Biondo Biondi, Mark Meadows, and Robert G. Clapp, GENESIS dataset: 3D initial VTI model for time-lapse reverse time migration 213

Modeling

Jon Claerbout, Making Marine Data From Land Data: suppressing 2-D surface noise on a 1-D line 227

Leighton M. Watson, Joseph Jennings, and Shuki Ronen, Source signature deconvolution of marine seismic data using deterministic modeling of the bubble signature 235

Inversion and noise suppression

Joseph Jennings, Biondo Biondi, Robert G. Clapp, and Shuki Ronen, Linearized waveform inversion of multicomponent blended data with polarization filters 243

Jon Claerbout, Fitting while whitening nonstationary residuals 261

Musa Maharramov and Biondo L. Biondi, Subsurface model inversion: pushing the limits of resolution 269

Machine learning

Fantine Huot, Yinbin Ma, Robert Cieplicki, Eileen Martin, and Biondo Biondi, Automatic noise exploration in urban areas 283

Stuart Farris and Stewart Levin, Higher Order Principal Component Analysis for Identifying AVO Anomalies 295

Software

Ettore Biondi and Guillaume Barnier, A flexible out-of-core solver for linear/non-linear problems 301

Stewart A. Levin and Robert G. Clapp, make, schmake: CMake 315

Robert G. Clapp, Facilitating code distribution: Docker and Generic IO . 319

SEP phone directory 321
Research personnel 323
('SEP article published or in press, 2017',) 332

Overview of the Stanford DAS Array-1 (SDASA-1)

Eileen Martin, Biondo Biondi, Steve Cole, and Martin Karrenbach

ABSTRACT

Permanent dense seismic arrays are costly to install and maintain using traditional point sensors. Fiber optic Distributed Acoustic Sensing (DAS) arrays show promise as a low-cost alternative that can be left in place with thousands of sensors run by a single power source. However, the trenching process used in some surface DAS arrays can be costly and logistically prohibitive in some cases. To mitigate these issues, we designed an experiment to investigate the potential for fibers in slim holes underground, with an eye towards repurposing existing telecommunications infrastructure. Our experiment has three primary goals: ambient noise interferometry, earthquake detection, and recording active seismic shooting. In August 2016, 2.4 km of fiber optic cable was deployed in a two-dimensional array in existing telecommunications conduits underneath the Stanford University campus. This array has been continuously recording since Sep. 3, 2016, and is planned to continue for at least one year in the current configuration, known as the "Stanford DAS Array-1" or SDASA-1.

INTRODUCTION

Over the past decade, Distributed Acoustic Sensing (DAS) has become increasingly popular in the oil and gas industry for time-lapse and continuous monitoring, both for seismic event detection and imaging. DAS provides some advantages over other acquisition systems. Consider that as a traditional array of point sensors contains more sensors, the probability of having a number of sensors broken or out of power grows. But with a DAS array, there's a single power source to maintain. Each point sensor costs more than a channel of fiber optic cable. While the startup cost of a single DAS channel may be higher, the scalability to thousands of permanent sensors is much better. Additionally, fiber's flexibility means that a DAS interrogator unit may be plugged into a fiber run just under the surface as an array, then that same fiber may continue to run down the well(s) providing full field monitoring coverage throughout all stages of production. Furthermore, other types of laser interrogators may be plugged into either the same fiber or other fibers in the same bundle for information about other properties: temperature, static strain, potentially chemical composition.

The majority of work has been on using fibers in wells, but the same benefits of DAS have also led to multiple experiments on the use of DAS arrays buried in shallow trenches, both for active (Dou et al., 2016), (Kendall, 2014) and passive experiments (Ajo-Franklin et al., 2015), (Martin et al., 2016), (Zeng et al., 2017). Compared to trenching, telecommunications companies have found that the installation of slim boreholes within a few meters under the surface is often a more cost effective choice. In areas with soil contaminants or permafrost, it requires less environmental risk. In urban or suburban areas, trenching

may simply not be logistically possible, and there are often existing conduits underground previously installed for telecommunications fibers.

Using existing conduits can greatly reduce the cost of installing an array if one is willing to be constrained by existing conduit geometry. There are currently two ways to take advantage of existing conduits and/or fibers: (i) run a new cable in the empty space in those conduits, or (ii) plug into an unused fiber, called "dark fiber" in an existing bundle. For this experiment, we pursued option (i) so that we could more easily choose our conduit geometry without requiring fiber splices between distinct existing cables.

This report serves as an overview of the array and observations during the first 7 months of recording. First, we describe the array geometry and design process with respect to the angular sensitivity of DAS, then we describe the method for assigning spatial points to distances along the fiber given uncertainties in the instalation. We show some examples of a variety of noises and events recorded by the array and the spectral response of the array and heterogeneities in the background noise field throughout the site and over time. This is an incredibly versatile data set. We have only begun to scratch the surface with our analyses in other reports, so we provide a list of open questions that might be answered using these data.

ARRAY DESIGN AND GEOMETRY

After discussions with Stanford IT, we decided it would be more cost effective for us to run a new fiber in the existing telecomm conduits rather than splicing leftover fibers from multiple bundles together. If we had been willing to go with a more linear array design, or if we were in a different town with a different cost structure for installation work, this decision might have changed. This array was installed in the same way that all other fiber optic cables are installed in telecommunications conduits around campus. The fibers were spooled up and brought down into manholes, then pulled either by hand or with a machine along narrow (10-15 cm wide) conduits connected between the manholes. The fibers sit loosely in the conduits, and where they are inside manholes (small underground rooms roughly 8 feet high and 3-4 ft by 6-9 ft wide) the fibers are zip-tied to a bracket on the side of the wall. There were two locations with 150 feet of fiber spooled up and strapped to the wall (with a vertical and horizontal component): one at Campus Dr. and Via Ortega, and another just south of Allen on Via Pueblo.

Because a straight segment of fiber is only sensitive to extensional strain along its length, DAS has less sensitivity to plane waves at an angle than geophones (Mateeva et al., 2012) ($\cos^2(\theta)$ versus $\cos(\theta)$ for a planar P-wave), and no sensitivity to broadside waves. This is one of the biggest limitations of DAS, so we designed our array to include fibers in two orthogonal directions to have some sensitivity to waves in all angles as seen in Figure 2.

The array has two recording modes currently configured: active and passive.

- **Active** mode records 2500 samples per second at a gauge length of 7.14 m and channel spacing of 1.02 m.
- **Passive** mode records 50 samples per second at a gauge length of 7.14 m and channel spacing of 8.16 m.

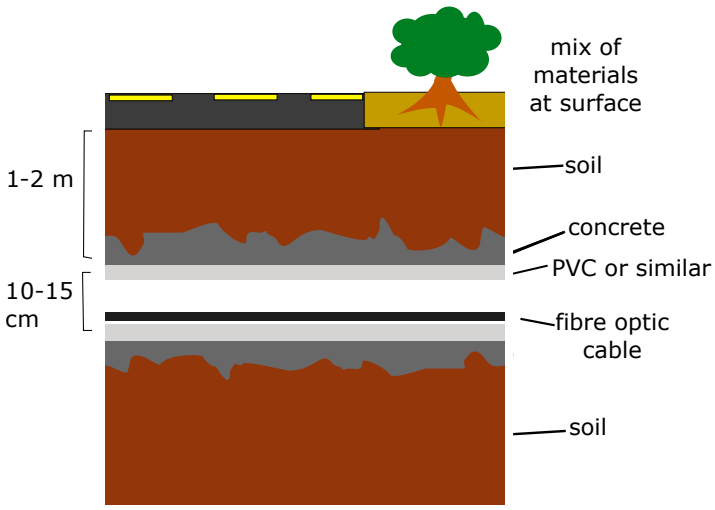


Figure 1: The campus has a mix of near surface materials, both natural and manmade. One to two meters below the surface sit conduits for telecommunications. These are generally 10-15 cm in diameter, and usually made of PVC or similar materials, and in some parts of campus are surrounded by concrete or cement slurry before being buried. Our fiber optic cable is roughly 1 cm in diameter and rests in the conduits loosely. [**NR**] eileen1/. installation

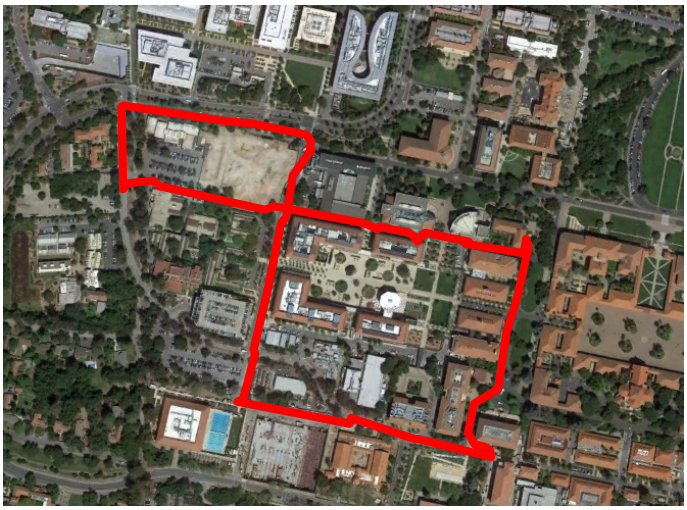


Figure 2: The layout of the fiber following telecommunications conduits overlaid on the map. The longest linear section is roughly 600 meters wide. Some deviations from straight lines had to occur due to existing conduit geometry constraints. [**NR**] eileen1/. fiberMap

Note that the gauge length is the length of the subset of fiber over which average strains are reported. The vast majority of the time, the array is recording in passive mode so as to keep the data size manageable. When we do active tests (including geometry calibration tap tests), we switch to active recording mode. The switch between these two configured modes can be handled remotely and no physical access to the box is required after installation. Although their gauge length is the same, the active mode data can have more options to add together neighboring channels to simulate a variety of gauge lengths (and thus, a variety of wavenumber sensitivity profiles, which can be beneficial). An example of a recording without any active shots but recorded in active mode can be seen in Figure 3. There is generally energy in anthropogenic frequency ranges rolling along from the northeast to the southwest, possibly coming from the town of Palo Alto, or from Highway 101, but this needs to be further investigated.

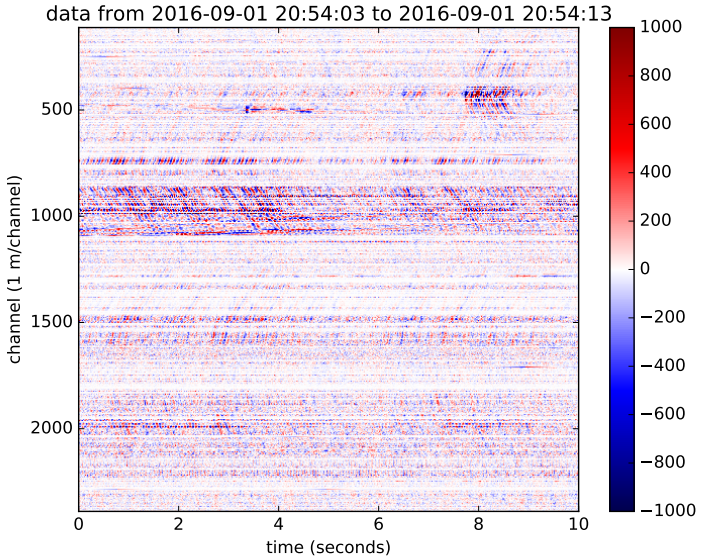
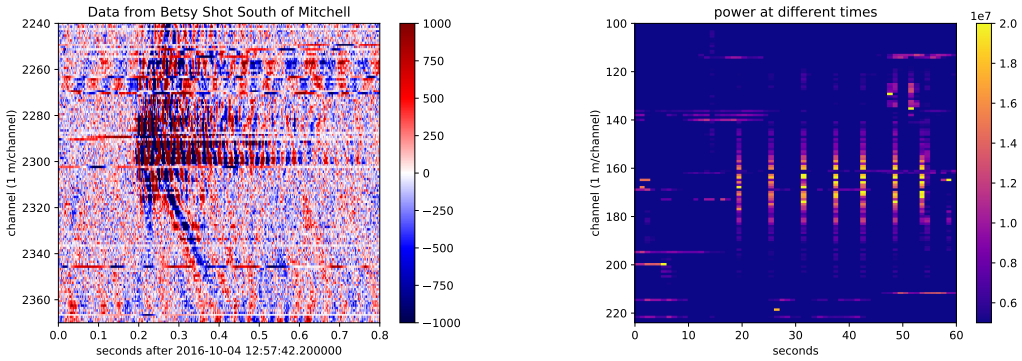


Figure 3: An example of 10 seconds of data recorded in active mode but without any controlled active sources. Active channels 400-500 are near a construction zone for Roble Parking Garage, which may explain the bump at 7-9 seconds. Active channels 800-1100 are around the area following Campus Dr. which gets quite a bit of vehicle traffic. [**CR**] eileen1/. activeEx

ARRAY LOCATION CALIBRATION

Unfortunately, there is currently no easy way to tie the data recorded on each channel to specific spatial locations without some manual labor. Stanford IT provided us with a scale map of manhole locations along our path, so we used many of these points for calibration. The channels in manholes tended to have poorer coupling since the fiber was strung partially along the side of a wall instead of sitting on the bottom of the conduit with gravity assisting it. We also did many sledgehammer tests, as well as a few betsy gun shots. Both types of tests can be seen in Figure 4(a) and 4(b). We recorded these in active mode, so we had to scale distances by roughly a factor of 8 for the passive channel points. Additionally, we

used the change in angle of waves coming in to the array as an indicator of array angle changes. An example of the kinds of passive noise revealing array geometry can be seen in Figure 3.



(a) (b)

Figure 4: (Left) A betsy gun shot south of Mitchell on 2016-10-04 as recorded on channels 2240 to 2370 in active recording mode starting 1016-10-04 12:57:42.2 UTC. (Right) The power of small windows of time on each channel during 8 lbs. sledgehammer tests west of Green Building from 2016-10-04 from 12:36:43 to 12:45:42 UTC. [**CR**] eileen1/. betsyShot,tapPower

SPECTRA

We often look at strain rate of the data, rather than the strain, by taking a forward difference in time for each channel. In large part, this is because the spectrum of each channel (a temporal derivative of a spatial derivative of displacement in the fiber direction) more closely matches spectra recorded by geophones, which measure velocities (a temporal derivative of displacement). Also, as seen in Figures 6(a), 6(b), 6(c), 6(d), the spectrum of the strain is so dominated by low frequencies that it is very difficult, even on a log scale, to visualize anthropogenic noise.

LIST OF ACTIVE RECORDING PERIODS

Some active recording periods of interest (in UTC time) include:

- Between 2016-08-30 23:00 and 2016-08-31 00:30, active recording of mid-afternoon mallet tests near a few manholes
- Between 2016-09-01 10:10 and 20:50, active recording of early afternoon Dropatron 5000 (a weight drop source designed by OptaSense) tests
- Between 2016-09-01 23:50 and 2016-09-02 00:30, passive recording of mid-afternoon Dropatron 5000 tests

Label	Channel Number	East UTM	North UTM
Start S. of Green	14	573088.0	4142497.0
Between Roble & HEPL	25	573000.58	4142531.35
Arrillaga Corner Start	48	572851.00	4142560.00
Arrillaga Corner End	49	572851.00	4142560.0
Via Ortega & Panama	58	572871.35	4142626.52
Via Ortega By Y2E2	70	572889.00	4142710.00
Via Ortega & Via Pueblo NS	83	572920.19	4142817.00
NW Corner of Allen	96	572961.29	4142909.74
Campus Dr. Coil Start	100	572942.64	4142936.69
Campus Dr. Coil End	107	572942.64	4142936.69
Panama & Campus	138	572694.11	4142985.43
Panama Near Pine	155	572693.26	4142866.36
Panama Curve Start	157	572695.88	4142872.24
Panama Curve End	165	572736.10	4142875.45
NW Corner of Pine	167	572740.38	4142868.55
Via Ortega & Via Pueblo EW	184	572922.4	4142826.52
Coil by Allen Start	203	573047.61	4142791.85
Coil by Allen End	209	573047.61	4142791.85
S of Hewlett Start	225	573172.14	4142767.36
S of Hewlett End	228	573172.14	4142767.36
Sequoia Jog Start	240	573249.00	4142746.00
Sequoia Jog Top	245	573258.00	4142770.00
Moore	261	573221.45	4142639.24
Bike Racks By Skilling	269	573205.43	4142563.1
NW Corner of Mitchell	274	573173.9	4142544.07
W Side of Mitchell	280	573161.0	4142501.0
S of Mitchell Start	283	573188.0	4142468.0
S of Mitchell End	288	573188.0	4142468.0
End S. of Green	302	573088.84	4142497.62

Table 1: List of physical points used to compare signals from particular channels to geometric locations

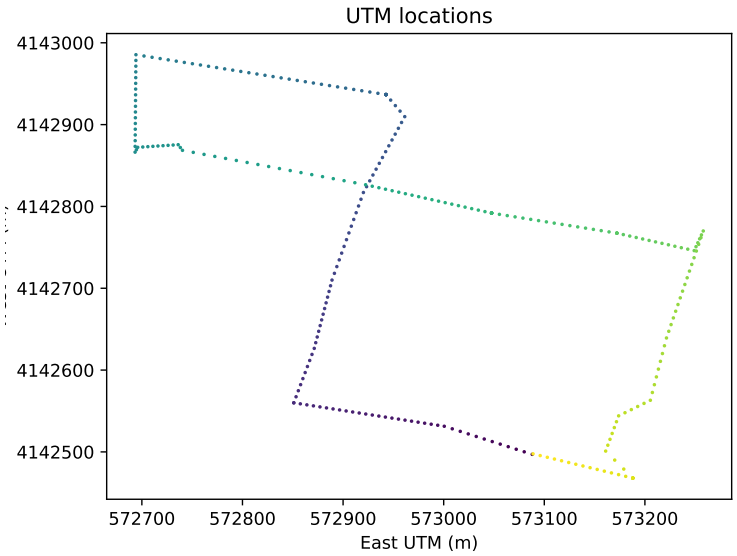


Figure 5: After including calibration points from Table 1, the center of each channel is marked in UTM coordinates without projection. [**ER**] eileen1/. channelMapping

- Between 2016-10-04 12:00 and 14:00, active recording of one betsy gun shot south of Mitchell as well as two mallet tests
- Between 2016-10-06 15:00 and 16:10, active recording during mallet tests on Lomita Mall and around Mitchell
- Between 2017-01-16 13:30 and 15:20, active recording during multiple betsy shots over a few blocks south of Mitchell
- Between 2017-03-18 18:00 and 2017-03-20 23:00, off-and-on active recording during multiple tests using 500+ sledgehammer hits and several dozen betsy gun shots, also concurrently recorded by a line of three-component nodes

DISCUSSION AND FUTURE WORK

Motivated by the successes of DAS in trenched arrays (Kendall, 2014), (Ajo-Franklin et al., 2015), (Martin et al., 2016), (Dou et al., 2016), (Lindsey et al., 2016), (Zeng et al., 2017), we have been collecting a multi-purpose data set testing the utility of DAS in telecommunications conduits. We have been testing it for ambient noise interferometry (Martin et al., 2017a), earthquake detection (Biondi et al., 2017), and active seismic survey recording (Martin et al., 2017b). We plan to publicly archive around ten earthquake recordings from the array in an IRIS assembled data set. We also plan to pull shots from the continuous recordings of the DAS array and 3C node array from the weekend of 2017-03-18, an experiment aimed at imaging the Stock Farm Monocline and comparing the two systems to improve our understanding of how to quantitatively use DAS. As described in (Huot et al., 2017), we are investigating methods to automatically aide geophysicists in exploring the

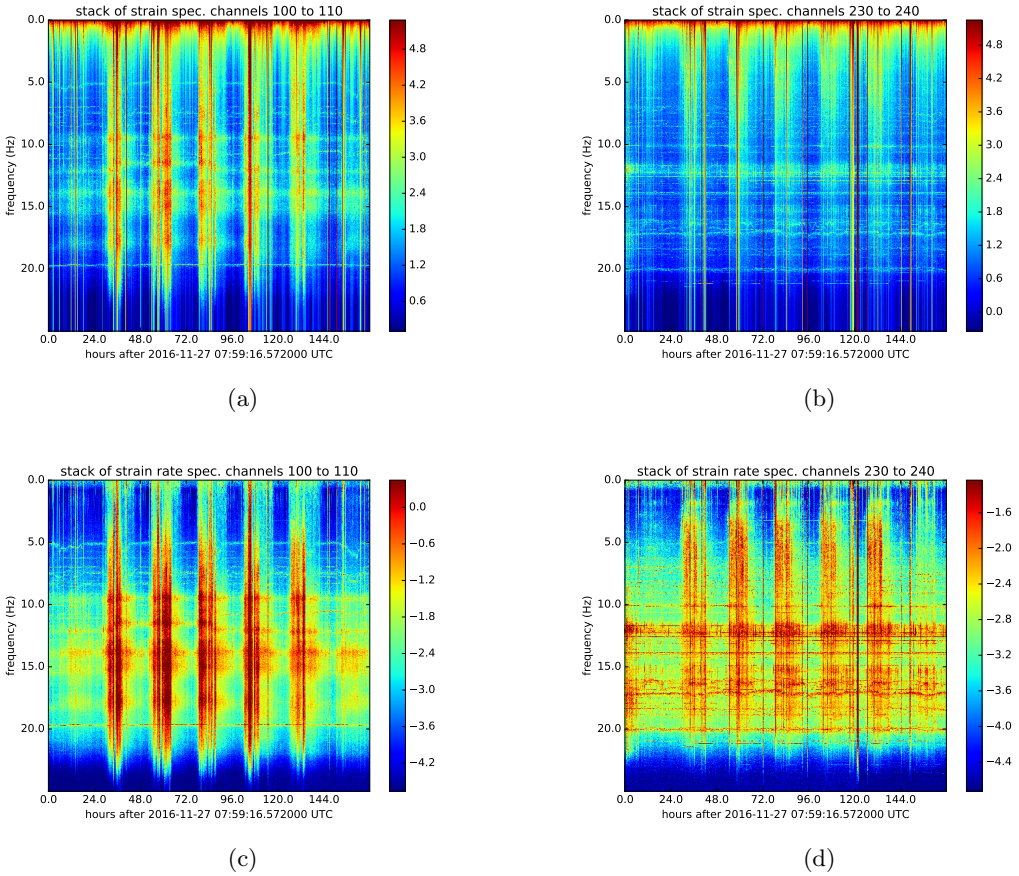


Figure 6: One week worth of the log of the average spectrum of the (top) strain and (bottom) strain rates of channels every 10 minutes (left) 100 to 110 and (right) 230 to 240. Channels 100 to 110 are along Campus Drive, and 230 to 240 are on Via Pueblo close to its intersection with Lomita Mall. [**CR**] eileen1/. strainSpec100To110,strainSpec230To240,strainRateSpec100To110,strainRateSpec230To240

noise field with the goal of quickly identifying potential calibration points and speeding up the process of developing ambient noise pre-processing.

ACKNOWLEDGEMENT

Acknowledge Carson Laing of OptaSense for installing the box and assisting in running it. We would like to thank Chris Castillo and Ethan Williams for help calibrating channel locations. We thank Subsea Systems for loaning us the trigger timing module used in calibrations. We thank Stanford IT for installing the fiber optic array, and Stanford SEEES IT for server room space and access. E. Martin conducted some of this work in part under DOE CSGF under grant number DE-FG02-97ER25308, and the Schlumberger Innovation Fellowship. We would also like to thank Jonathan Ajo-Franklin at Lawrence Berkeley National Lab and Nate Lindsey at University of California Berkeley for many useful discussions, particularly comparisons with other DAS experiments run by LBL.

REFERENCES

Ajo-Franklin, J., N. Lindsey, S. Dou, T. Daley, B. Freifeld, E. Martin, M. Robertson, C. Ulrich, and A. Wagner, 2015, A field test of distributed acoustic sensing for ambient noise recording: Expanded Abstracts of the 85th Ann. Internat. Mtg.

Biondi, B., E. Martin, S. Cole, and M. Karrenbach, 2017, Earthquakes analysis using data recorded by the Stanford DAS Array : SEP-Report, **168**, 11–26.

Dou, S., J. Ajo-Franklin, T. Daley, M. Robertson, T. Wood, and B. Freifeld, 2016, Surface orbital vibrator (sov) and fiber-optic das: Field demonstration of economical, continuous-land seismic time-lapse monitoring from the australian co2crc otway site: Expanded Abstracts of the 86th Ann. Internat. Mtg.

Huot, F., Y. Ma, R. Cieplicki, E. Martin, and B. Biondi, 2017, Automatic noise exploration in urban areas : SEP-Report, **168**, 277–288.

Kendall, R., 2014, A comparison of trenched distributed acoustic sensing (das) to trenched and surface 3c geophones - daly, manitoba, canada: Expanded Abstracts of the 84th Ann. Internat. Mtg.

Lindsey, N., D. Dreger, A. Wagner, S. James, and J. Ajo-Franklin, 2016, Distributed fiber optic sensing of earthquake wavefields: Presented at the , AGU Fall Meeting.

Martin, E., B. Biondi, M. Karrenbach, and S. Cole, 2017a, Ambient noise interferometry from das array in underground telecommunications conduits: Technical Programme of the 79th Conference & Exhibition.

———, 2017b, Continuous subsurface monitoring by passive seismic with distributed acoustic sensors- the "stanford array" experiment: Proceedings of the First EAGE Workshop on Practical Reservoir Monitoring.

Martin, E., N. Lindsey, S. Dou, J. Ajo-Franklin, A. Wagner, K. Bjella, T. Daley, B. Freifeld, M. Robertson, and C. Ulrich, 2016, Interferometry of a roadside das array in fairbanks, ak: Expanded Abstracts of the 86th Ann. Internat. Mtg.

Mateeva, A., J. Mestayer, B. Cox, D. Kiyashchenko, P. Wills, J. Lopez, S. Grandi, K. Hornman, P. Luments, A. Franzen, D. Hill, and J. Roy, 2012, Advances in distributed acoustic sensing (das) for vsp: Expanded Abstracts of the 82nd Annual International Meeting.

Zeng, X., C. Thurber, H. Wang, D. Fratta, E. Matzel, and P. Team, 2017, High-resolution shallow structure revealed with ambient noise tomography on a dense array: Proceedings, 42nd Workshop on Geothermal Reservoir Engineering.

Earthquakes analysis using data recorded by the Stanford DAS Array

Biondo Biondi, Eileen Martin, Steve Cole, and Martin Karrenbach

ABSTRACT

We analyze the data recorded by the "Stanford DAS Array" for 6 seismic events cataloged in the USGS online database. Two of these events are man made (blasts at a nearby quarry), and 4 are earthquakes spanning a range from a local event that occurred less than 4 km from the array to a large quake offshore Cape Mendocino in North California. The data from two seismometers installed on Stanford campus at Jasper Ridge help to validate and interpret the data recorded by our DAS array. The analysis of the two quarry-blasts demonstrates that both kinematics and waveforms are recorded by the DAS array with excellent repeatability. We show that the time of the first-break of P-wave arrivals can be reliably measured from the DAS array, notwithstanding the loose coupling of the fiber-optic cable with the ground, and the known limitations of DAS to record events with particle motion orthogonal to the fiber cable. P-wave waveforms are more challenging to analyze because of the complexity of the wavefield, probably caused by strong local scattering. All the events we analyzed show that kinematics of S and surface waves can be reliably measured from the data recorded by our DAS array. Because of noise and uncertain coupling, it is more challenging to assess the quality of the waveform shapes and amplitudes than the kinematics. We show that that DAS recording of propagating waves with particle displacement orthogonal to the propagation direction are subject to a phase rotation that is dependent on the fiber-cable direction. When we apply an approximate correction for this phase rotation the spatial coherency of S-wave and surface-wave arrivals substantially improves.

INTRODUCTION

One of the main goals of our "Stanford DAS Array-1" (SDASA-1) experiment (Martin et al., 2017b,a) is to evaluate the feasibility of using fiber-optics as Distributed Acoustic Sensors (DAS) to detect and analyze earthquake data. DAS arrays have the potential of enabling cost-effective continuous and spatially dense seismic monitoring over large areas. This application of DAS technology has implications for managing earthquake hazard in densely populated areas at high seismic risk (e.g. the San Francisco Bay Area) as well as for reservoir management when seismic events are caused by either injection or extraction of fluids. Our experiment is unique because we are recording data from a fiber-optic cable laying in a PVC conduit buried in the ground. Coupling between the fiber cable and the surrounding rocks relies exclusively on gravity and friction. This recording configuration could deliver huge savings by enabling the exploitation of existing infrastructure developed for telecommunication purposes (like in our experiment) or the use of ad-hoc inexpensive "slim-holes" with PVC casing buried sufficiently deep to bypass the near-surface and avoid its negative effects on the recorded data.

Because SDASA-1 has been continuously recording data starting in September 2016 and California is tectonically active, we have a large number of events to analyze. We focus on four naturally-occurring events that span the full range from weak and nearby earthquakes to strong and far away ones. In our analysis we also included two blasts from a nearby quarry that are ideal to test the repeatability of the recording. To validate the events observed in SDASA-1 data, we use the USGS earthquake online database and the data recorded at the "Jasper Ridge Biological Preserve" (JRSC station) by two seismometers that are managed by the Berkeley Digital Seismic Network. Data from JRSC are available online. The Jasper Ridge station is located about 6.4 km from the SDASA-1. Because near-surface conditions are different below SDASA-1 and JRSC the waveforms are not directly comparable. However, JRSC data provide a rough indication of the arrival time and relative strength of the signal corresponding to different arrivals (e.g. P-waves, S-waves, and surface waves).

In the next section we provide a brief description of the experimental setup and of the events that we used in our study. We then analyze the recordings by the SDASA-1 in details, and compare them with the corresponding data recorded at Jasper Ridge.

EXPERIMENT SETUP

Martin et al. (2017a) provide most of the relevant information on the SDASA-1 experiment. Here we provide additional information on the experiment setup that is important for interpreting the results discussed in the following section.

Figure 1 shows the layout of the array; its double-loop path is outlined by the red line. The annotations indicate the channel numbers at the array corner points. The channel numbering starts at #5, outside of the Green Earth Sciences Building where the laser interrogator is installed, and increases as we move westward along the path.

As mathematically demonstrated in Appendix A, elastic waves are recorded by the DAS array with a scaling and phase rotation dependent on the angle between the propagation direction and the fiber direction. In particular, waves with particle displacement normal to the propagation direction are recorded with opposite polarity by two segments of SDASA-1 oriented in orthogonal directions (N-S vs. E-W). To make the spatial continuity of some of the arrivals more apparent, when plotting some of the figures we have scaled the amplitudes of each channel by either -1 or +1, depending on whether the fiber cable was directed in the N-S direction (+1) or E-W direction (-1). Figure 2 displays this scaling factor as a function of the channel number.

Jasper Ridge seismic station (JRSC)

JRSC was installed in 1994 by a team from Stanford Geophysics (Robert Kovach and Greg Beroza), the USGS, and UC Berkeley. It is part of a network of seismic stations in northern and central California maintained and operated by the Berkeley Digital Seismic Network (BDSN). The station consists of two 3-components seismometers. A broad-band seismometer (JRSC-BB): Streckeisen STS-2 VBB Tri-Axial Seismometer, and a long-period accelerometer (JRSC-LP): TSA-100S (4 g max 5 v/g). Data are available for downloading sampled at 0.025 s for JRSC-BB and at 1 s for JRSC-LP. The seismometers sit on a concrete

pad inside a preexisting cave on the west shore of Searsville Lake. The cave is located 36 m above sea level and 6.4 km to the SWW of SDASA-1.

Seismic events

In the following section we analyze seismic data recorded by the SDASA-1 and compare them with data recorded by JRSC. Since we activated the array in September 2016 we recorded many seismic events. Our analysis is limited to 6 events that are representative of the whole range of events recorded by the array. The following list provide the main characteristics of these 6 events in order of increasing magnitude:

- *Blast #1*: September 14, 2016 – M 1.3 – z=-0.3 km – Δ=14.40 km to SSE,
- *Blast #2*: November 4, 2016 – M 1.6 – z=-0.3 km – Δ=14.05 km to SSE,
- *Ladera*: January 11, 2017 – M 2.0 – z=4.2 km – Δ=3.86 km to SW,
- *Bonny Doon*: October 27, 2016 – M 2.5 – z=13.1 km – Δ=37.80 km to SSW,
- *Piedmont*: September 12, 2016 – M 3.5 – z=4.0 km – Δ=42.00 km to N,
- *Ferndale*: September 2, 2016 – M 5.6 – z=28.6 km – Δ=443 km to NW.

The magnitude (M) and depth (z) of each event are based on the online USGS database. The distance from SDASA-1 (Δ) is estimated using Google map. The date is in local California time. In the following we will refer to each individual event by the corresponding label typeset in italic, as in the list above (e.g. *Blast #1*.)

The first two events are artificial blasts at a nearby quarry; their source locations are separated by only 350 meters, and thus their source waveforms and propagation paths are presumably similar. These two events are useful to evaluate the repeatability of the recording. The other 4 events are naturally occurring earthquakes, ranging from local and fairly shallow ones, to much stronger and deeper quakes. The USGS website provides further information on the source mechanism of events with magnitude equal or larger than 2. We are not reporting this information since our analysis is not using it, but more detailed analysis could correlate the observed waveforms to the estimated source mechanisms.

DATA ANALYSIS

We present the analysis of the data recorded by SDASA-1 and JRSC for the 6 events listed above by focusing on one important issue at the time: 1) signal repeatability, 2) P-wave arrivals, and 3) S-wave and surface-wave arrivals.

Signal repeatability

Because of the weak signals and strong environmental noise, in particular during the day, the first-order question that our analysis aimed to answer is whether the events that we observe in the recorded data correspond to actual seismic events. Fortunately, the vicinity

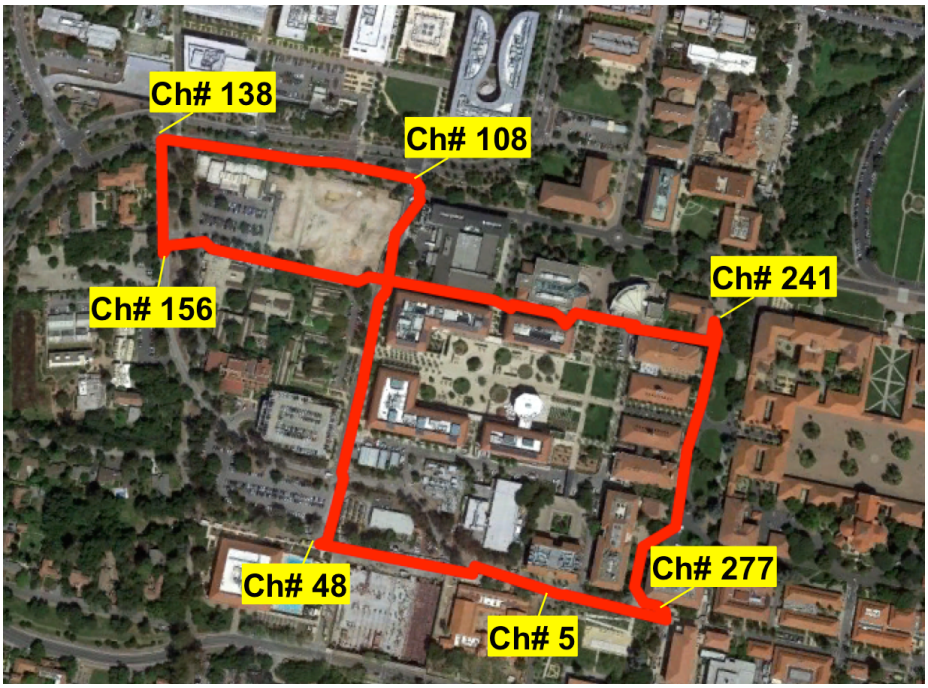


Figure 1: Map showing the path of SDASA-1 with the location of the corner points of the array marked and labeled with the corresponding channel numbers. [**NR**] biondo1/. MAP-channels-rot

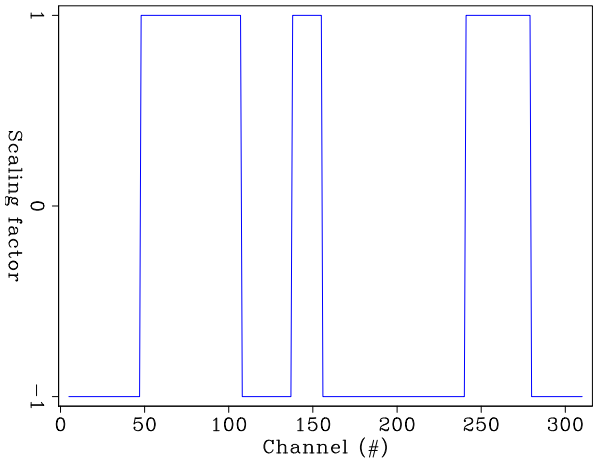


Figure 2: Scaling factor used for some of the data displays. Channels where the fiber cable is oriented in the N-S direction are weighted by +1, whereas the ones where the fiber cable is oriented in the E-W direction are weighted by -1. [**ER**] biondo1/. mask

with a quarry provided us with several recordings of artificial blasts that can be used to assess the repeatability of the observed arrivals. This analysis leads to the conclusions that indeed the recorded events are repeatable and directly correlated to quarry blasts.

The top panel in Figure 3 shows the data recorded by SDASA-1 after noise attenuation and bandpassing from 0.25 to 2.5 Hz. The DAS channels were scaled by the scaling function shown in Figure 2 before plotting. The origin of the time axis is the time of *Blast #1* as provided by the USGS online database. We use the same convention to set the origin of the time axis for all the data displays that follow. The trace at the bottom shows the data recorded by the North component of JRSC-BB. This trace was bandpassed with the same filter as the DAS data; that is, from 0.25 to 2.5 Hz. All the data displays that follow show JRSC data bandpassed using the same parameters used to bandpass the corresponding DAS data. The channel numbers in all the figures correspond to the channels numbers marked on the map in Figure 1. The strong slowly-propagating events visible around channel #150 were generated by vehicles passing close to the array. This vehicle-related noise is visible in most of the recording, with the exception of the data corresponding to the *Piedmont* event that occurred just before midnight.

Figure 4 shows the data recorded by SDASA-1 corresponding to *Blast #2* and the corresponding North component recorded by JRSC-BB. *Blast #2* occurred less than two month after *Blast #1* and had a higher magnitude (1.6 vs. 1.3); the source location was about 350 m to the North of *Blast #1* source location. The timing and waveform are consistent between these two events, with a slight time shift caused by the shift in source location. Because of the time delays from the blasts, we interpreted these events as surface waves generated by the blasts and being recorded by SDASA-1.

In both *Blast #1* and *Blast #2* recordings, the time shifts of the events across the array match the relative distances from the sources of the recording channels; the events arrive at the southern side of the array (edges) almost 0.5 s before they arrive at its Northern side (middle) (see Figure 1.)

At least for these quarry blasts, it is easier and more reliable to detect the events from the SDASA-1 data than from the data recorded by the seismometers at Jasper Ridge. It would be impossible to detect the blasts from any of the traces recorded by JRSC (all three components of JRSC-BB and JRSC-LP.) In these recordings the signal does not rise above the noise level, as exemplified by the North components displayed at the bottom of Figures 3 and 4.

P-wave kinematics and waveforms

Next issue that we address is whether we can extract reliable kinematic information about P-wave arrivals from data recorded by SDASA-1. Figures 5 and 6 demonstrate that the timing of the P-wave arrivals can be easily picked from the SDASA-1 data.

The top panel in Figure 5 shows the envelope of the data recorded by SDASA-1 corresponding to the *Piedmont* event. The envelope was computed after noise attenuation and high-passing above 0.25 Hz. The trace at the bottom shows the data recorded by the vertical component of JRSC-BB. The amplitude jump in the envelope function is easily identifiable and the arrival time is consistent with the arrival time at JRSC-BB. There is a time shift of approximately .25 s between SDASA-1 and JRSC data that is due to the

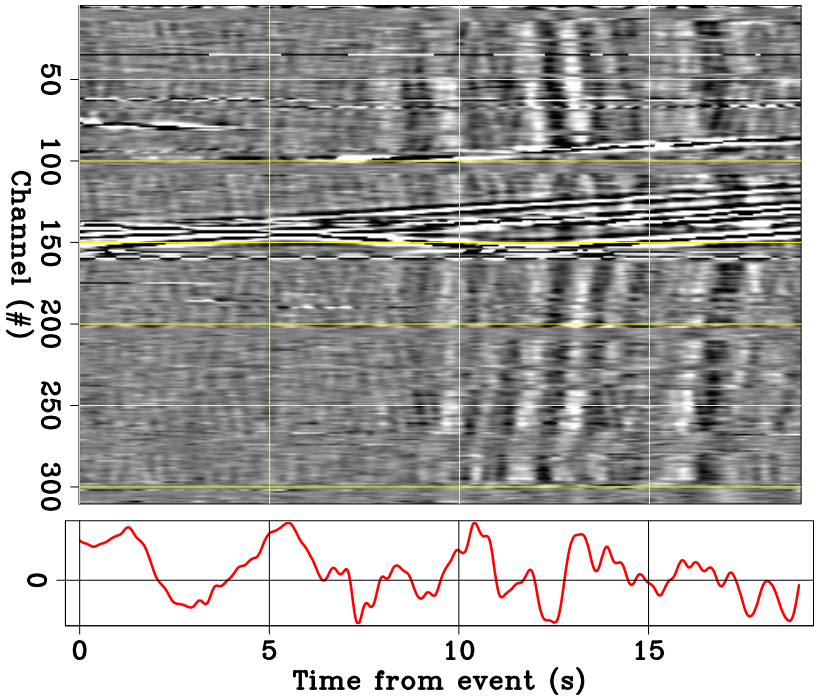


Figure 3: Data recorded for the *Blast #1* event after bandpassing (0.25–2.5 Hz) and channel-dependent scaling by the function shown in Figure2. Top: SDASA-1 array; bottom: JRSC-BB North component. [**ER**] biondo1/. S-N-blast14Sep-mask-hor

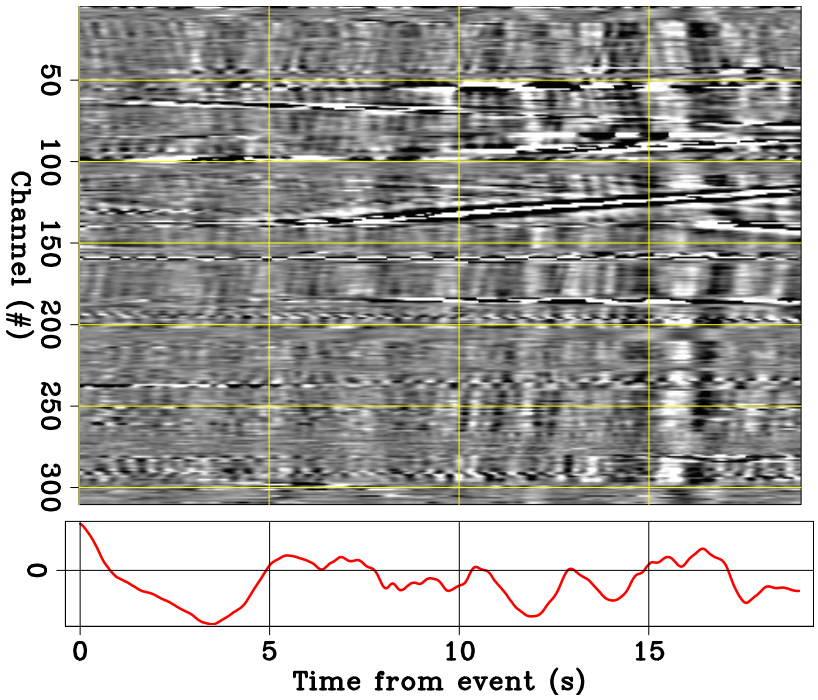


Figure 4: Data recorded for the *Blast #2* event after bandpassing (0.25–2.5 Hz) and channel-dependent scaling by the function shown in Figure2. Top: SDASA-1 array; bottom: JRSC-BB North component. [**ER**] biondo1/. S-N-blast4Nov-mask-hor

relative distance from the source location. Piedmont is almost perfectly North of Stanford and JRSC is approximately 2.5 km South of SDASA-1. The time-shifts within the array are also consistent with an event arriving from the North. They are much smaller than the time shifts of the surface waves observed in Figures 3 and 4 because P-waves generated at this distance (42 km) emerge at the surface at a fairly steep angle.

The top panel in Figure 6 shows the envelope of the data recorded by SDASA-1 corresponding to the *Ladera* event. In this case the event is much closer to JRSC than to SDASA-1, and thus the time of the first break at JRSC-BB (bottom trace) is about a third of the first-break time at SDASA-1. As in the *Piedmont* case, the relative arrival times across the array are consistent with a S-W direction of the source. However, the picking of first breaks would be more challenging in this case than in the previous one because of environmental noise. The first channels to record the event are around the SW corners of the array; that is, around Channel #48 and around Channel #156.

The vertical component recorded at JRSC-BB (bottom panel in Figure 6) shows clearly the arrival of the P-waves, the S-waves, and of the surface waves. The sudden increase in amplitudes of the envelope function at about 3.8 s is most probably caused by the arrival of the S-waves at SDASA-1. Unfortunately, because of the proximity of the *Ladera* event, the JRSC-BB data cannot be used to estimate the timing of the S-wave arrival at SDASA-1, and thus to validate this interpretation.

In contrast with the kinematics of the P-wave arrivals that are clearly measurable from the SDASA-1 data, the waveforms are not as well preserved. Figures 7 and 8 show the data from which the envelope functions shown in Figures 5 and 6 were computed. In Figure 8 we can identify several coherent arrivals between 2.5 s and 3.5 s that are related to the first P-wave arrival and to complex scattering along the wavepaths. The recorded wavefield is complex because of scattering, but the recording is also suffering from the effect of imperfect coupling and of the known strong directional sensitivity of the DAS sensor (Kuvshinov, 2016).

S-wave and surface-wave kinematics and waveforms

Figure 6 shows an example where the kinematic of S-waves can be estimated from the SDASA-1 data. However, it is more difficult to identify S-arrivals than surface-wave arrivals, not only from SDASA-1 data, but also from JRSC data. Figures 9 and 10 show an example of this challenge. The top panel in Figure 9 shows the data recorded for the *Bonny Doon* event after noise reduction and bandpassing from .25 to 1.0 Hz. The trace at the bottom shows the corresponding North component of JRSC-BB. Figure 10 shows the same data as in Figure 9, but after applying the scaling function shown in Figure 2 to the SDASA-1 data. The polarity flipping performed by the scaling helps the identification of S-waves and surface-waves as coherent arrivals across the array.

Surface waves arriving at about 16 s are clearly identifiable in both SDASA-1 and JRSC data. In the SDASA-1 data, the time differences between channels of these arrivals are similar to the ones observed from the quarry blasts (Figures 3 and 4.) There are a few notable differences related to the slight difference in arrival direction. The blasts arrived from the SEE direction, whereas the *Bonny Doon* event arrived from the SWW direction. The most obvious difference is that the *Bonny Doon* arrival has no moveout along the longest

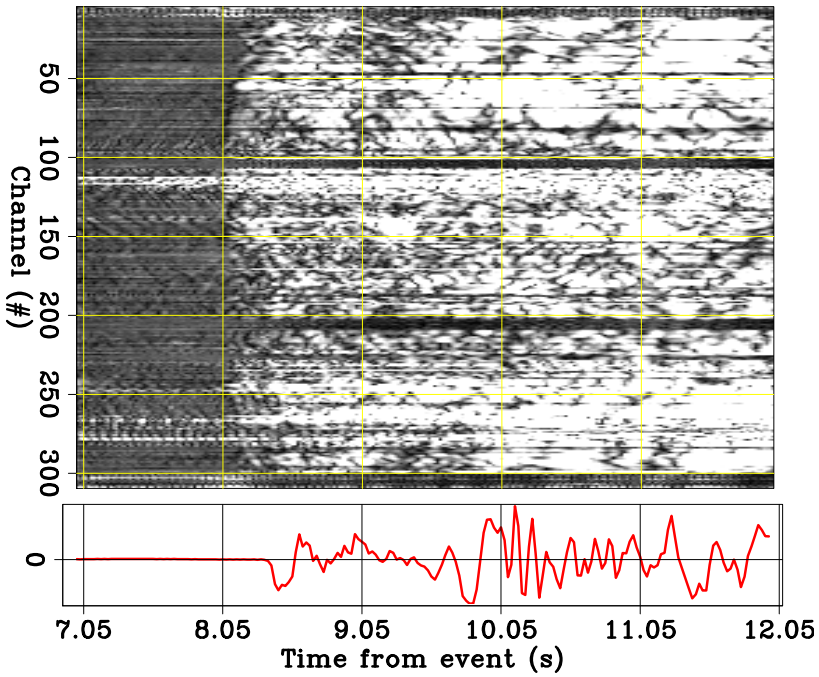


Figure 5: Data recorded for the *Piedmont* event after high-passing ($\geq$ 0.25 Hz). Top: Envelope of SDASA-1 array data; bottom: JRSC-BB vertical component. [**ER**] biondo1/. Env-hf-piedmont-hor

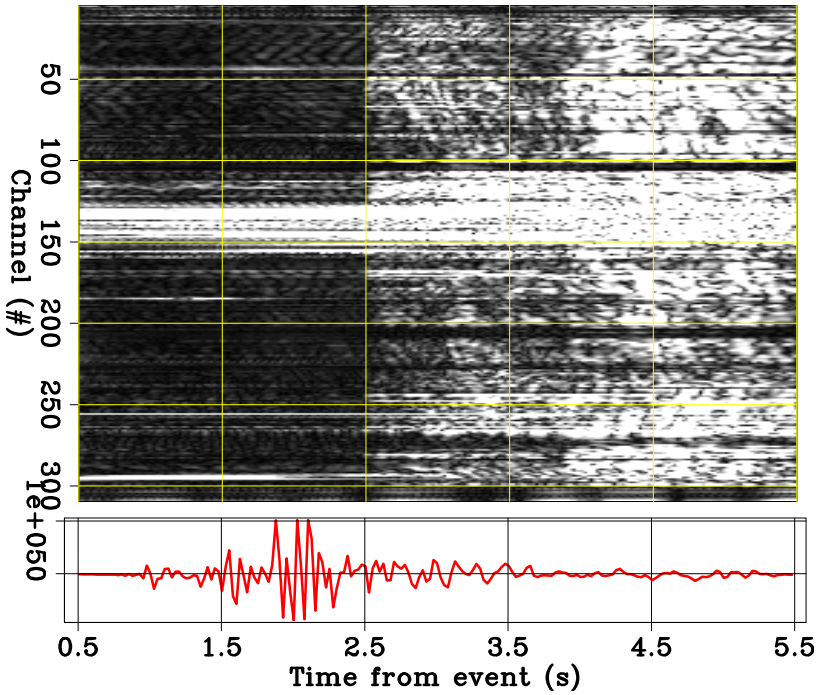


Figure 6: Data recorded for the *Ladera* event after high-passing ($\geq$ 0.25 Hz). Top: Envelope of SDASA-1 array data; bottom: JRSC-BB vertical component. [**ER**] biondo1/. Env-hf-ladera11Jan-hor

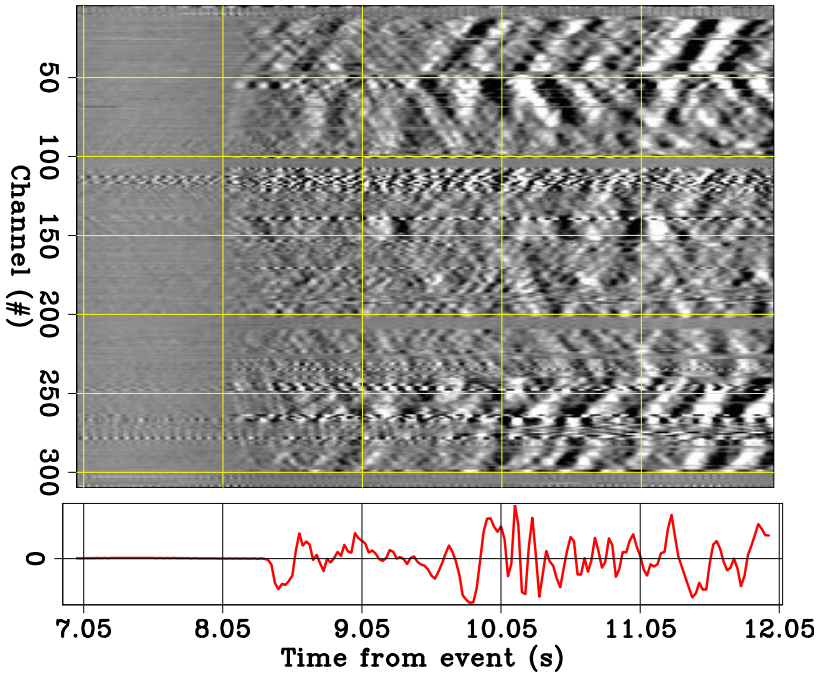


Figure 7: Data recorded for the *Piedmont* event after high-passing ($\geq$ 0.25 Hz). Top: SDASA-1 array; bottom: JRSC-BB vertical component. [**ER**] biondo1/. P-hf-piedmont-hor

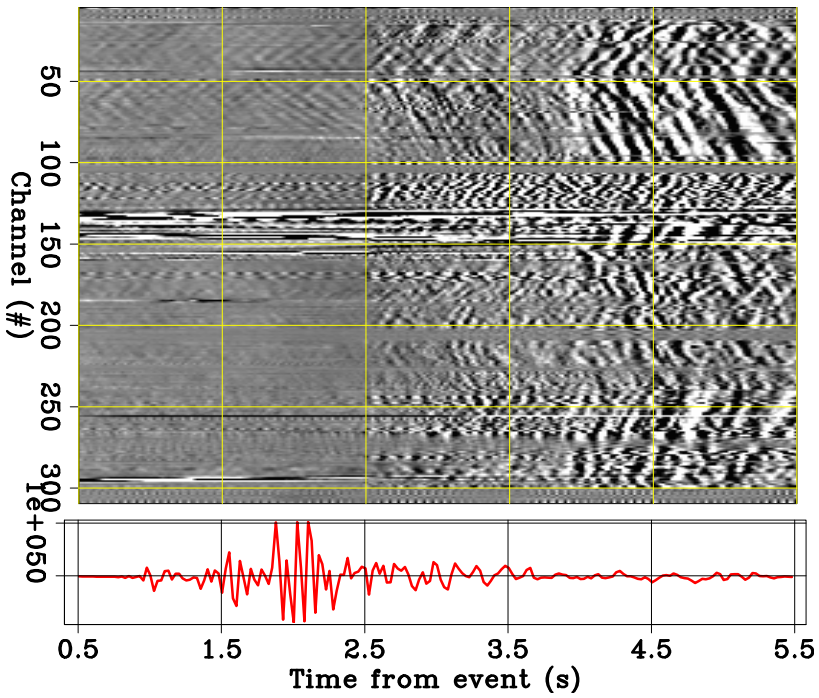


Figure 8: Data recorded for the *Ladera* event after high-passing ($\geq$ 0.25 Hz). Top: SDASA-1 array; bottom: JRSC-BB vertical component. [**ER**] biondo1/. P-hf-ladera11Jan-hor

leg of the array (Channel #156 to Channel #241) because it arrives along a direction almost perfectly orthogonal to the fiber-cable direction. On the contrary, the blasts events have a slight, but clearly observable, negative moveout along this same leg; the blasts arrive before at the eastern channels than at the western ones.

Both the SDASA-1 data and the JRSC data show some possible arrivals at around 5 s and 11 s. They could be interpreted the P-wave and S-wave arrivals, respectively. A large uncertainty would be associated with this interpretation. This uncertainty would be greatly reduced if we had data from a wider DAS array since we would be able to measure moveouts more reliably across a wider array than across SDASA-1. These moveouts are not measurable from our fairly narrow array because body-waves emerge at the surface almost vertically.

The last event we examine is *Ferndale*. This is a large (M=5.6) and deep event originated along the Mendocino Fault under the Pacific Ocean more than 400 km to the NW of Stanford. For this event we observe strong S-wave and surface-wave arrivals at both SDASA-1 and JRSC. Figure 11 shows the S-wave arrivals. The trace at the bottom shows the North component of JRSC-BB. The data were bandpassed from .25 to 1 Hz, and the data from the SDASA-1 were scaled before displaying. Fortunately, during the *Ferndale* event the vehicle noise was not as strong as during other events. Therefore, we could preserve the original data dynamic range when plotting the SDASA-1 data; that is, we clipped only the highest amplitudes (pclip=99.5). In both recordings, the S-arrival rises above the noise after a fairly quite period and it easily detectable.

The relative timing of the arrivals across the array is more challenging to interpret than for previous events. The arrival time has a local minimum close to the NW corner of the array (Channel #138). However away from that zone, the arrivals have more complex and not as easily interpretable moveouts. Given the long path from source to receivers, these behavior may be explained by ray bending, as well as by near-surface heterogeneities close to SDASA-1.

Finally, Figure 12 shows the data window around the surface-wave arrivals. As for the previous figure, we plotted the SDASA-1 data preserving the data dynamic range. Both SDASA-1 and JRSC data show clearly the first surface-waves arriving after a quite period. Similarly to the analysis of S-wave arrivals above, the kinematics across the array are more difficult to interpret but the arrival time has a local minimum close to the NW corner of the array (Channel #138).

CONCLUSIONS AND FUTURE WORK

The analysis of six seismic events recorded by SDASA-1 demonstrate the suitability of this kind of arrays for cost-effective recording of seismic events. We showed that a wealth of useful information can be extracted from the recorded data, notwithstanding the expected limitations in the fidelity of the recording caused by the strong directional sensitivity of DAS sensors, and imperfect coupling between the fiber cable and the ground in our open-conduit installation.

We have shown that reliable kinematic information is well preserved in the recorded data for all arrivals (P-waves, S-waves and surface waves.) A more detailed and quantitative analysis should be performed on the reliability of the information that can be extracted

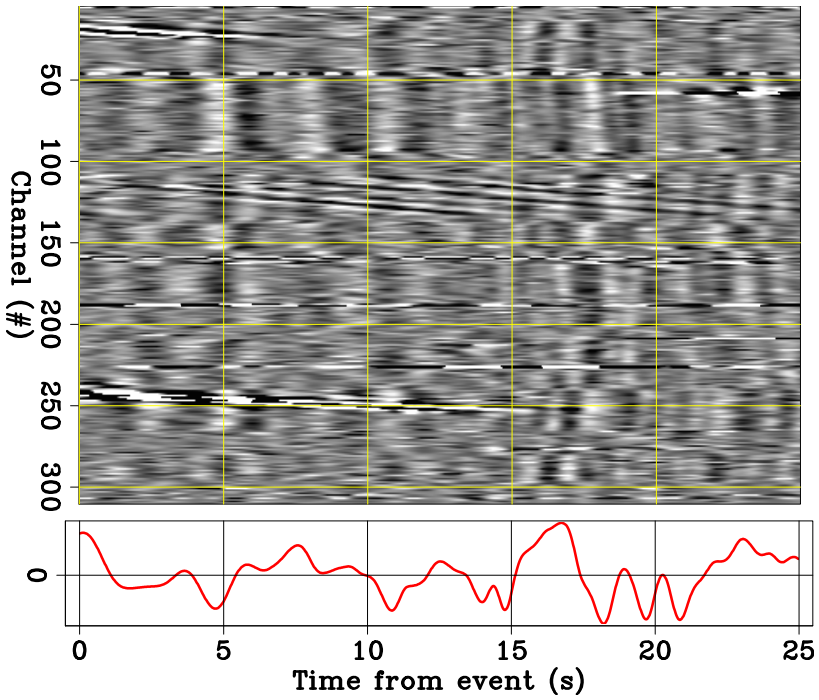


Figure 9: Data recorded for the *Bonny Doon* event after bandpassing (0.25–1.0 Hz). Top: SDASA-1 array; bottom: JRSC-BB North component. [**ER**] biondo1/. S-N-bd-hor

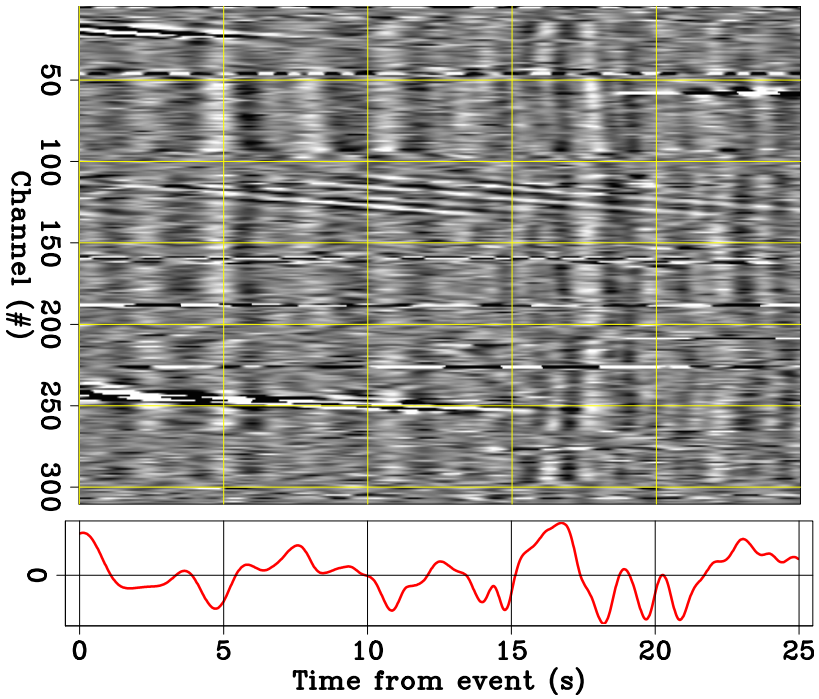


Figure 10: Data recorded for the *Bonny Doon* event after bandpassing (0.25–1.0 Hz) and channel-dependent scaling. Top: SDASA-1 array; bottom: JRSC-BB North component. [**ER**] biondo1/. S-N-bd-mask-hor

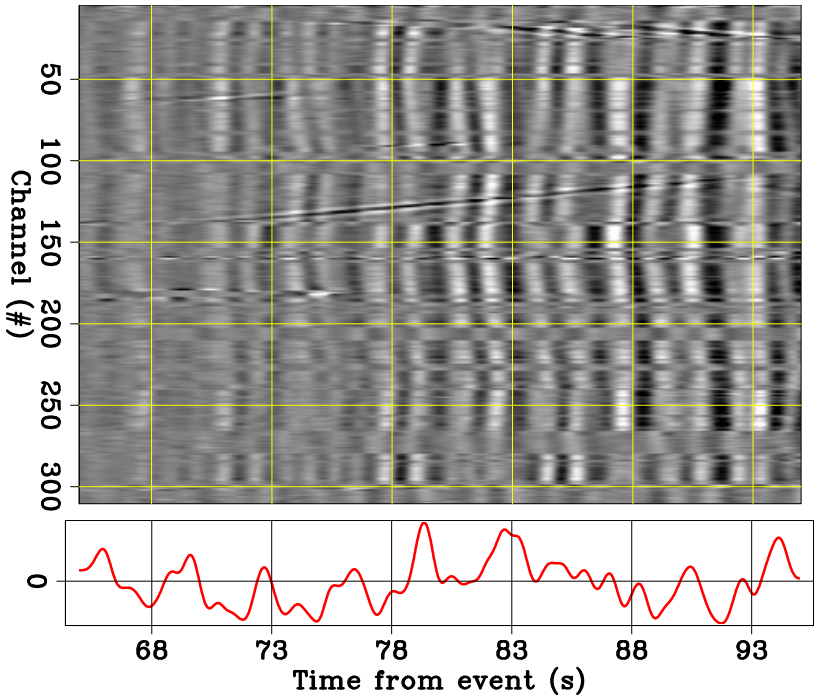


Figure 11: Data recorded for the *Ferndale* event after bandpassing (0.25–1.0 Hz) and channel-dependent scaling. Top: SDASA-1 array; bottom: JRSC-BB North component. [**ER**] biondo1/. S-N-ferndale-mask-hor

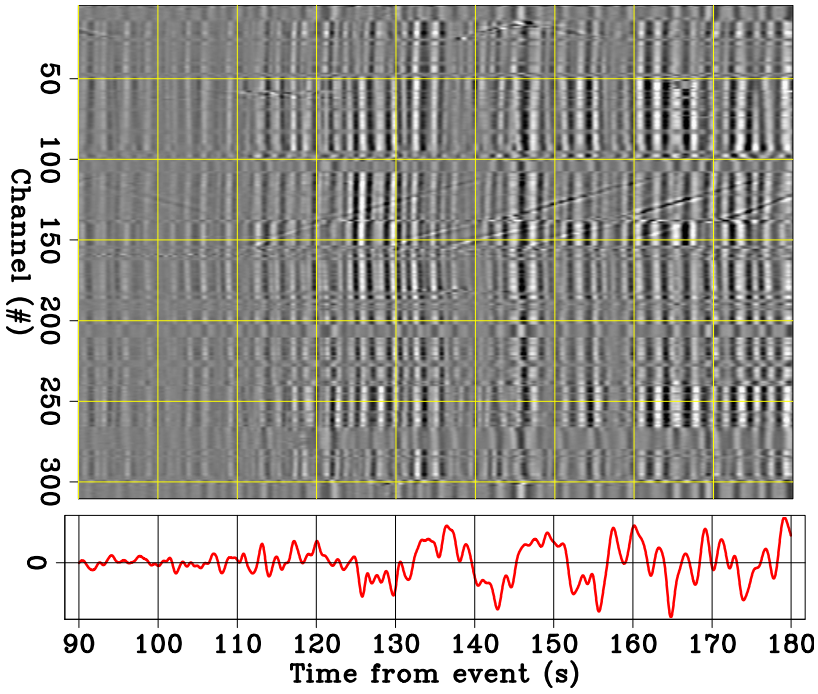


Figure 12: Data recorded for the *Ferndale* event after bandpassing (0.25–1.0 Hz) and channel-dependent scaling. Top: SDASA-1 array; bottom: JRSC-BB North component. [**ER**] biondo1/. Surf-N-ferndale-mask-hor

from the data on waveform shapes and amplitudes. We plan to install 2-3 broad-band seismometers in the manholes where the fiber cable is currently installed. These seismometers should provide invaluable reference data to perform a more quantitative analysis on the data collected by SDASA-1. Furthermore, as the array is continuously operating, we are accumulating a large database of many events of all magnitudes, coming from all directions, and with a broad range of source mechanisms. If appropriately mined, possibly by using novel data-analytic algorithms, this database will enable much more detailed analysis.

The comparison of the data recorded using our current laser interrogator with the data recorded using a latest-generation interrogator would also provide extremely useful information. In particular, we could leverage the repeatability of the quarry blasts to quantitatively analyze differences in data quality observed with different laser interrogators. This analysis could lead to quantitative evaluation of the limits on the data quality imposed by the coupling of the fiber cable (that we do not expect to improve in similar open-conduit installations) versus the limitations of the laser interrogators, that we expect to substantially improve as the technology progress.

Finally, data recorded from a broader array that could be easily deployed using Stanford's extensive fiber infrastructure would provide extremely useful information. As discussed in the paper, for body-wave arrivals the moveouts across the array are small, and thus do not enable the use of beamforming or other "array processing" methods. However, the reliability of the kinematic information that we measure from SDASA-1 suggests that a wider array (SDASA-2?) would enable much more detailed analysis of seismic events originating around, or below, the Bay Area.

ACKNOWLEDGMENTS

We would like thank OptaSense Ltd for making the DAS recording equipment available and supporting the work of two of the authors. We would like also to thank Gregory Kersey, Paul Narcisse, and Gary Gutfeld with the Fiber Team within Stanford IT Services for their support and patience.

REFERENCES

Kuvshinov, B., 2016, Interaction of helically wound fibre-optic cables with plane seismic waves: Geophysical Prospecting, **64**, 671–688.

Martin, E., B. Biondi, S. Cole, and M. Karrenbach, 2017a, Overview of the Stanford DAS Array-1 (SDASA-1) : SEP-Report, **168**, 1–10.

Martin, E., B. Biondi, M. Karrenbach, and S. Cole, 2017b, Continuous Subsurface Monitoring by Passive Seismic with Distributed Acoustic Sensors: the "Stanford Array" experiment: Presented at the "First EAGE Workshop on "Practical Reservoir Monitoring", Eur. Assoc. Expl. Geophys.

APPENDIX A

In this appendix we develop the theory showing that elastic waves impinging on a DAS array are recorded with different amplitude scaling and phase rotation depending on whether the

particle displacement is aligned with the propagation direction (e.g. P-waves) or orthogonal to it (e.g. S-waves and Love waves).

We present the case for 2D propagation. We have not fully developed the 3D case yet. However, it is probably required to understand the phase-rotation of Rayleigh waves recorded by SDASA-1.

2D propagation in the same direction as particle motion

The analysis that follows is relevant to P-waves and the horizontal component of Rayleigh waves. We'll use a coordinate system (r_1, r_2) where r_1 is the propagation direction as well as the particle displacement direction. The expression for displacement at position (r_1, r_2) and time t of a monochromatic plane wave with frequency ω and wavenumber k is

$$\mathbf{d}(r_1, r_2, t) = (\cos(kr_1 - \omega t), 0), \tag{A-1}$$

thus the strain is

$$\underline{\mathbf{s}}(r_1, r_2, t) = \begin{bmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{bmatrix} = \begin{bmatrix} -k\sin(kr_1 - \omega t) & 0 \\ 0 & 0 \end{bmatrix}. \tag{A-2}$$

Let's consider the case when we observe these strains from a rotated coordinate system (a fiber that sits at an angle θ from $(1, 0)$). By performing a tensor rotation we can write:

$$\begin{aligned} \underline{\mathbf{s}}(r_{1'}, r_{2'}, t) &= \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} -k\sin(kr_1 - \omega t) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \\ &= \begin{bmatrix} -k\sin(kr_1 - \omega t)\cos^2(\theta) & k\sin(kr_1 - \omega t)\sin(\theta)\cos(\theta) \\ k\sin(kr_1 - \omega t)\sin(\theta)\cos(\theta) & -k\sin(kr_1 - \omega t)\sin^2(\theta) \end{bmatrix} \\ &= -k\sin(kr_1 - \omega t) \begin{bmatrix} \cos^2(\theta) & -\sin(\theta)\cos(\theta) \\ -\sin(\theta)\cos(\theta) & \sin^2(\theta) \end{bmatrix}, \end{aligned}$$

so the fiber would measure $s_{1'1'}$ which is $-k\sin(kr_1 - \omega t)\cos^2(\theta)$.

Now let's consider the case when we observe these strains from a coordinate system rotated $\pi/2$ from that one, plugging in $\cos(\theta + \pi/2) = -\sin(\theta)$ and $\sin(\theta + \pi/2) = \cos(\theta)$:

$$\begin{aligned} \underline{\mathbf{s}}(r_{1''}, r_{2''}, t) &= -k\sin(kr_1 - \omega t) \begin{bmatrix} \cos^2(\theta + \pi/2) & -\sin(\theta + \pi/2)\cos(\theta + \pi/2) \\ -\sin(\theta + \pi/2)\cos(\theta + \pi/2) & \sin^2(\theta + \pi/2) \end{bmatrix} \\ &= -k\sin(kr_1 - \omega t) \begin{bmatrix} \sin^2(\theta) & \cos(\theta)\sin(\theta) \\ \sin(\theta)\cos(\theta) & \cos^2(\theta) \end{bmatrix}, \end{aligned}$$

so the fiber would measure $s_{1''1''}$ which is $-k\sin(kr_1 - \omega t)\sin^2(\theta)$. That means any measurements on this fiber are the same as measurements on the orthogonal fiber by a factor of $\tan^2(\theta)$. Therefore, while we expect the polarity of waves observed by these two fibers to be the same (since $\tan^2(\theta) \geq 0$), the amplitude ratio between their observations depends on the direction of the propagating wave.

2D propagation in orthogonal direction to particle motion

This is relevant to Love waves or S-waves. We'll use a coordinate system (r_1, r_2) where again r_1 is the direction of propagation, but now r_2 is the direction of particle motion. The expression for displacement at position (r_1, r_2) and time t of a monochromatic plane wave with frequency ω and wavenumber k is

$$\mathbf{d}(r_1, r_2, t) = (0, \cos(kr_1 - \omega t)), \tag{A-3}$$

thus the strain is

$$\underline{\mathbf{s}}(r_1, r_2, t) = \begin{bmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{bmatrix} = \begin{bmatrix} 0 & \frac{-k}{2}\sin(kr_1 - \omega t) \\ \frac{-k}{2}\sin(kr_1 - \omega t) & 0 \end{bmatrix}. \tag{A-4}$$

Let's consider the case when we observe these strains from a rotated coordinate system (a fiber that sits at an angle θ from $(1, 0)$). By performing a tensor rotation we can write:

$$\begin{aligned}
\underline{\mathbf{s}}(r_{1'}, r_{2'}, t) &= \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 0 & \frac{-k}{2}\sin(kr_1 - \omega t) \\ \frac{-k}{2}\sin(kr_1 - \omega t) & 0 \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \\
&= \begin{bmatrix} \sin(\theta)\frac{-k}{2}\sin(kr_1 - \omega t) & \cos(\theta)\frac{-k}{2}\sin(kr_1 - \omega t) \\ \cos(\theta)\frac{-k}{2}\sin(kr_1 - \omega t) & -\sin(\theta)\frac{-k}{2}\sin(kr_1 - \omega t) \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \\
&= -\frac{k}{2}\sin(kr_1 - \omega t) \begin{bmatrix} 2\sin(\theta)\cos(\theta) & \cos^2(\theta) - \sin^2(\theta) \\ \cos^2(\theta) - \sin^2(\theta) & 2\sin(\theta)\cos(\theta) \end{bmatrix}.
\end{aligned}$$

Hence, our fiber at angle θ would observe $s_{1'1'}$ of the rotated strain tensor which is $-k\sin(kr_1 - \omega t)\sin(\theta)\cos(\theta)$.

Now let's say we also wanted to observe the strains from a coordinate system that had been rotated $\pi/2$ from that one, we just have to plug in the identities $\cos(\theta + \frac{\pi}{2}) = -\sin(\theta)$ and $\sin(\theta + \frac{\pi}{2}) = \cos(\theta)$:

$$\begin{aligned}
\underline{\mathbf{s}}(r_{1"}, r_{2"}, t) &= \begin{bmatrix} \cos(\theta + \frac{\pi}{2}) & \sin(\theta + \frac{\pi}{2}) \\ -\sin(\theta + \frac{\pi}{2}) & \cos(\theta + \frac{\pi}{2}) \end{bmatrix} \begin{bmatrix} 0 & \frac{-k}{2}\sin(kr_1 - \omega t) \\ \frac{-k}{2}\sin(kr_1 - \omega t) & 0 \end{bmatrix} \cdots \\
&\quad \cdots \begin{bmatrix} \cos(\theta + \frac{\pi}{2}) & -\sin(\theta + \frac{\pi}{2}) \\ \sin(\theta + \frac{\pi}{2}) & \cos(\theta + \frac{\pi}{2}) \end{bmatrix} \\
&= -\frac{k}{2}\sin(kr_1 - \omega t) \begin{bmatrix} 2\sin(\theta + \frac{\pi}{2})\cos(\theta + \frac{\pi}{2}) & \cos^2(\theta + \frac{\pi}{2}) - \sin^2(\theta + \frac{\pi}{2}) \\ \cos^2(\theta + \frac{\pi}{2}) - \sin^2(\theta + \frac{\pi}{2}) & 2\sin(\theta + \frac{\pi}{2})\cos(\theta + \frac{\pi}{2}) \end{bmatrix} \\
&= -\frac{k}{2}\sin(kr_1 - \omega t) \begin{bmatrix} -2\cos(\theta)\sin(\theta) & \sin^2(\theta) - \cos^2(\theta) \\ \sin^2(\theta) - \cos^2(\theta) & -2\cos(\theta)\sin(\theta) \end{bmatrix}.
\end{aligned}$$

Therefore, our fiber at angle $\theta + \frac{\pi}{2}$ would observe $s_{1"1"}$ of the rotated strain tensor which is $k\sin(kr_1 - \omega t)\sin(\theta)\cos(\theta)$, which is exactly -1 times the observation of the fiber at angle θ.

In conclusions, no matter what direction our wave is coming from relative to the fibers, the two fibers in perpendicular directions will always observe this type of plane wave with opposite signs (one sees compression while the other sees extension, and vice-versa).

Ambient Noise Interferometry on Two-Dimensional DAS Arrays

Eileen Martin and Biondo Biondi

ABSTRACT

When using three-component geophones for ambient noise tomography, the process of extracting Rayleigh or Love waves requires first rotating data into radial and transverse components, respectively, before cross-correlation. When a two-dimensional distributed acoustic sensing (DAS) array is used, a single straight line can easily have Rayleigh waves extracted via cross-correlation within the line. However, only using straight lines significanly reduces our ray coverage, so we aim to garner meaninful information from all pairs of channels in the array. In practice, cross-correlation and cross-coherence of the data recorded at many pairs of channels in the Stanford DAS Array-1 (SDASA-1) yield coherence waveforms. But, we can not rotate every data point into radial and transverse components, so special treatment must be given to understand the coherent waveforms generated when non-colinear fibers' data are cross-correlated. This is the first extension of ambient noise theory given these geometric constraints.

INTRODUCTION

When ambient noise interferometry was first catching on, scientist cross-correlated single-component data (Shapiro et al., 2005), then researchers realized they could take any pair of three-component geophones, rotate their data to a radial component in-line with both receivers, then cross-correlate to get Rayleigh surface waves. Less than a decade after the use of amibent noise interferometry began, the theory was extended and shown to work in practice for extraction of Love waves at regional (Lin et al., 2008) and later reservoir (de Ridder, 2014), (Nakata et al., 2015) and engineering (Nakata et al., 2011) scales. The process is similar to Rayleigh wave extraction, although transverse components are cross-correlated.

Over the past few years, fiber optic Distributed Acoustic Sensing (DAS) arrays have begun to emerge as a potentially cost-effective option for permanent or long-term ambient noise collection (Ajo-Franklin et al., 2015), (Martin et al., 2015), (Martin et al., 2016), (Martin et al., 2017c), (Martin et al., 2017b), (Zeng et al., 2017), (Martin et al., 2017a). Although straight fiber optics keep the cost of DAS arrays lower, broadside insensitivity and lowered angular sensitivity relative to geophones is a major limiting factor in their utility (Kuvshinov, 2016). For extracting Rayleigh waves, the theory is not very different, so we can correlate channels along the same straight line (Martin et al., 2015), (Martin et al., 2016), (Martin et al., 2017c), (Martin et al., 2017b), (Zeng et al., 2017). All of the Rayleigh waves extracted from DAS arrays have been extracted from two-dimensional arrays, so we aim to use the vastly more numerous ray paths between different lines of these arrays so that velocities may be estimated throughout the extent of these arrays. This is the first extension of interferometry of a DAS array to non-colinear channels.

This report develops the theory of ambient noise cross-correlations between straight-fiber DAS channels at arbitrary distances and orientations from each other in section 2. The data extracted are predicted to be a mix of Love and Rayleigh waves in many cases. In section 3 we show results of cross-correlations and cross-coherences throughout the Stanford DAS Array-1 (SDASA-1) (Martin et al., 2017a). Then we conclude, and discuss future extensions of this work.

THEORY

We assume the signal recorded by a DAS channel is the average extensional strain rate along a gauge long subset of fiber, which is equivalent to the difference in the velocities in the direction of the fiber at the end points of that gauge length. We aim to describe the cross-correlation of two DAS signals in terms of cross-correlations of four three-component geophones, one sitting at each end of each of the two channels. Ideally we would follow the procedure used in Lin et al. (2008) to extract radial and transverse components that correspond to Rayleigh and Love waves, respectively, but with the geometric constraints of DAS, we must limit ourselves to only having measurements in one direction.

We start with $C_{geo:s,r}(\tau)$, the cross-correlation of the θ_s component of an $x-y$ oriented 3C geophone at (x_s, y_s) with the θ_r component of another $x-y$ oriented 3C geophone at (x_r, y_r), as pictured in Figure 1(a). Let $\mathbf{v}^{(\cdot)}(t)$ be the velocity vector at $(x_\cdot, y_\cdot)$ time t, and let $C_{xy:s,r}(\tau)$ be the cross-correlation of $v_x^s(t)$ with $v_y^r(t)$ (and likewise for other components). We drop the τ to shorten notation from here on. Let $c_\cdot = \cos(\theta_\cdot)$ and $s_\cdot = \sin(\theta_\cdot)$. Then our improperly rotated cross-correlation is:

$$\begin{aligned} C_{geo:s,r} &= \int (v_x^s(t)c_s - v_y^s(t)s_s)(v_x^r(t+\tau)c_r - v_y^r(t+\tau)s_r)dt \\ &= c_s c_r C_{xx:s,r} - c_s s_r C_{xy:sr} - s_s c_r C_{yx:s,r} + s_s s_r C_{yy:s,r} \end{aligned}$$

We can then use a 4x4 matrix as in Lin et al. (2008) to express these cross-correlations in terms of radial and transverse cross-correlations which can be more easily interpreted as Rayleigh and Love waves. This is achieved by rotating by the angle θ_T from the $x-y$ coordinate system into the $T-R$ transverse and radial coordinate system:

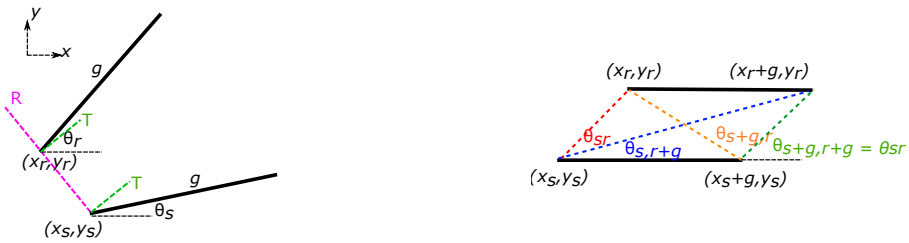
$$\begin{aligned} C_{geo:s,r} = C_{tt:s,r}(c_s c_T + s_s s_T)(c_r c_T + s_r s_T) + C_{tr:s,r}(c_s c_T + s_s s_T)(c_r s_T - s_r c_T) \\ + C_{rt:s,r}(c_s s_T - s_s c_T)(c_r c_T + s_r s_T) + C_{rr:s,r}(c_s s_T - s_s c_T)(c_r s_T - s_r c_T). \end{aligned}$$

Consider the case $\theta_r = \theta_s = 0$, which would be relevant to DAS channels on parallel lines, a common configuration (Martin et al., 2016), (Martin et al., 2017c), (Zeng et al., 2017). Note that $\theta_T = 0$ corresponds to two parallel channels directly across from each other ($\theta_{sr} = \pi/2$ in Figure 1(b)), and $\theta_T = \pi/2$ corresponds to two colinear channels ($\theta_{sr} = 0$ in Figure 1(b)). In this case, $C_{geos:s,r}$ simplifies to:

$$C_{geos:s,r} = C_{tt:s,r}c_T^2 + c_T s_T(C_{tr:s,r} + C_{rt:s,r}) + C_{rr:s,r}s_T^2$$

In fact, if $\theta_T = 0$, $C_{geos:s,r}$ simplifies to $C_{tt:s,r}$ which extracts a Love wave, and if $\theta_T = \pi/2$, $C_{geos:s,r} = C_{rr:s,r}$ which is a Rayleigh wave (Lin et al., 2008). For all θ_T in between, we

extract a mix of Love waves, Rayleigh waves and $(C_{tr:s,r} + C_{rt:s,r})$, quantities which have not been explained in the literature. The visuals in Lin et al. (2008) suggest these may simply increase the noise level, but further investigation is needed to understand the role of these terms.



(a) (b)

Figure 1: (Top) The geometry of two DAS channels (thick black lines) is drawn. Both channels use gauge length g, and if there were 3C geophones at their end points, radial R and transverse T components of the data could be used for easy Love and Rayleigh wave extraction. (Bottom) Multiple existing surface DAS arrays have some parallel lines of fibers in them. The angles between the end point pairs of two parallel DAS channels and the x-axis are labeled. [**NR**] eileen2/. geometryGeneral,parallelLines

Another common case relevant to perpendicular fibers is $\theta_s = 0$, $\theta_r = \pi/2$. In this case, $C_{geos:s,r}$ simplifies to:

$$C_{geos:s,r} = C_{tt:s,r} c_T s_T - C_{tr:s,r} c_T^2 + C_{rt:s,r} s_T^2 - C_{rr:s,r} s_T c_T$$

If $\theta_T = 0$ or $\pi/2$, i.e. we are at an orthogonal corner or crossing-point in the array, $C_{geos:s,r}$ simplifies to $-C_{tr:s,r}$ or $C_{rt:s,r}$ respetively. In future work, these corner and crossing points may be useful to recognizing these cross terms and understanding their influence throughout the rest of the array. At other angles, we would expect a linear combination of Rayleigh waves, Love waves, and the C_{rt} and C_{tr} cross-terms which are not well understood.

Now we can build from the expression for pairs of improperly rotated 3C geophones into the actual cross-correlation of two DAS channels. Consider two DAS channels of gauge length g, as in Figure 1(a), one extending from (x_s, y_s) to $(x_s + g\cos(\theta_s), y_s + g\sin(\theta_s))$, and the other from (x_r, y_r) to $(x_r + g\cos(\theta_r), y_r + g\sin(\theta_r))$. Using the notation that $s+g$ and $r+g$ are referring to positions $(x_s + g\cos(\theta_s), y_s + g\sin(\theta_s))$ and $(x_r + g\cos(\theta_r), y_r + g\sin(\theta_r))$, their cross-correlation, $C_{DAS:s,r}$ is:

$$C_{DAS:s,r} = C_{geo:s,r} - C_{geo:s+g,r} - C_{geo:s,r+g} + C_{geo:s+g,r+g}$$

Again consider the case of two parallel fibers, $\theta_r = \theta_s = 0$, as pictured in Figure 1(b), where channel $\cdot$ runs from $(x_\cdot, y_\cdot)$ to $(x_\cdot + g, y_\cdot)$. Let θ_{sr} be the angle between $(x_r, y_r) - (x_s, y_s)$ and $(x_s + g, y_s) - (x_s, y_s)$, and likewise for $s + g$ and $r + g$. There are two cases where $C_{DAS:s,r}$ simplifies. The first is for two colinear channels, $\theta_{sr} = 0$ (like the Rayleigh wave extraction that has been done in the past):

$$C_{DAS:s,r} = C_{rr:s,r} - C_{rr:s+g,r} - C_{rr:s,r+g} + C_{rr:s+g,r+g},$$

which leads to a Rayleigh wave which has a wavelet that is a bit more ringy than the typical $C_{rr:s,r}$. The other special case is for two parallel channels that line up, $\theta_{sr} = \theta_{s+g,r+g} = \pi/2$. As short notation, let X refer to $\theta_{s,r+g}$, so the DAS cross-correlation for parallel channels that line-up is:

$$\begin{aligned} C_{DAS:s,r} = C_{tt:s,r} - s_X^2(C_{tt:s+g,r} + C_{tt:s,r+g}) \; -s_X c_X(C_{tr:s+g,r} + C_{rt:s+g,r} - C_{tr:s,r+g} - C_{rt:s,r+g}) \\ -c_X^2(C_{rr:s+g,r} + C_{rr:s,r+g}) + C_{tt:s+g,r+g}. \end{aligned}$$

The cross-angle $\theta_{s,r+g} = \pi - \theta_{s+g,r}$ depends on the distance between the two fibers, so that $\theta_{s,r+g}$ goes to $\pi/2$ as the distance between the fibers grows or g shrinks. Thus, as $\theta_{s,r+g} \to \pi/2$, this simplifies to:

$$C_{DAS:s,r} \approx C_{tt:s,r} - C_{tt:s+g,r} - C_{tt:s,r+g} + C_{tt:s+g,r+g},$$

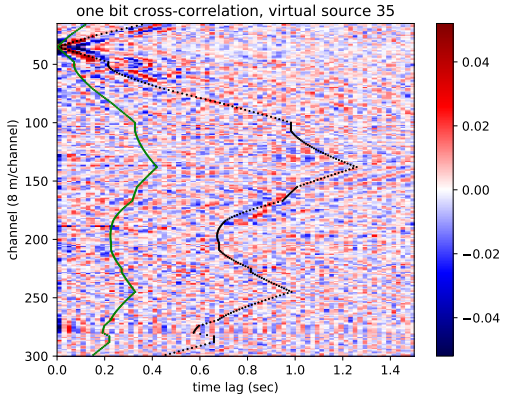
meaning that if the parallel fibers are at least a few gauge lengths apart, we expect to extract a Love wave with a wavelet that is a bit more ringy than the typical $C_{tt:s,r}$.

RESULTS

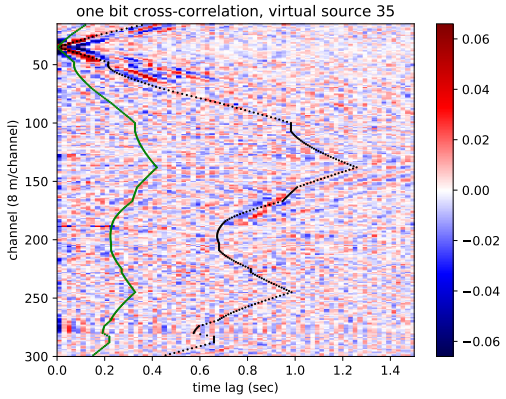
We tested cross-correlations between lines on SDASA-1. These are data recorded on an OptaSense ODH-3 interrogator unit at a 50 Hz sample rate, 8.16 m channel spacing, and 7.14 m gauge length. As seen in Martin et al. (2017c) and Martin et al. (2017b), we previously tested cross-coherence and cross-correlation as methods to extract Rayleigh waves from channels along the same straight line fiber. To start looking at response estimates between lines, we initally tried cross-correlation, but found it to not focus energy well enough. We next tested one-bit correlation. The data were first split into 300 second windows with 50% overlap, meaning a new windows starts every 150 seconds, then bandpassed from 0.2 to 24 Hz, then the median of each time sample was subtracted to remove laser noise before the data were thresholded to ± 1.

The correlations for a virtual source at channel 35 can be seen in Figures 2(a), 2(b), and 2(c). The greatest improvement can be seen in the response of channels 50 to 100, a segment perpendicular to 35 (so it should have a mix of Love and Rayleigh waves). In particular, there appear to be two distinct wave speeds along this segment. There's also a strong response on channels 160 to 180, a segment parallel to 35 which shows a much higher velocity event starting around 0.7 seconds lag. Further investigation is needed to determine whether this is truly an extracted Love wave, or something else, perhaps even an event with a higher apparent velocity due to 1D structure of that line.

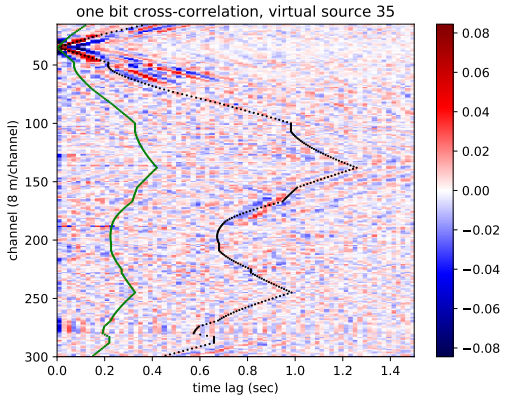
The correlations for a virtual source at channel 75 can be seen in Figures 3(a), 3(b), and 3(c). This channel is near the crossing point of Via Ortega and Via Pueblo, so in the 72 hour correlation, a slight V can be seen near channel 180 or 190, which is close to channel 75 but oriented in the orthogonal direction. More interestingly, after three days of correlations,



(a)



(b)



(c)

Figure 2: One-bit correlations of the array against a virtual source at channel 35 on the west-going line in the southwest corner of the array improve from (left) 24 hours of correlations to (right) 72 hours of correlations, and (bottom) to 120 hours. The arrival times of 1200 m/s and 400 m/s waves are overlaid. [**CR**]

eileen2/. oneBitXcorr24HrsSrcCh35,oneBitXcorr72HrsSrcCh35,oneBitXcorr120HrsSrcCh35

two events can be seen running along Via Ortega emanating from channel 75. A slow one that is stronger ata few hundred meters per second, and a fast one (particularly to the south of 75) closer to around one thousand meters per second. Further cross-correlations need to be done to get more certain picks on these velocities.

FUTURE WORK

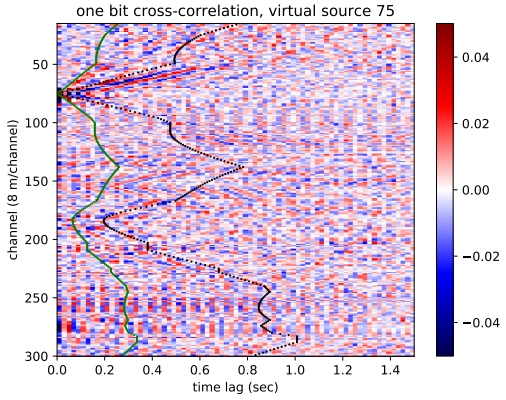
Further work needs to be done to quantify the convergence rates of the cross-correlations shown, and to analyze longer periods of time. It would be ideal to build further more scalable code into the framework already developed for pSIN (Chen et al., 2016), but so far, we have been unable to obtain the source code of pSIN. Although the cross-correlations worked well in many parts of the array, some areas, particularly near roads or along the big utility tunnel on Lomita Mall did not have very well-focused energy. To help with pre-processing decisions that may effect convergence rates and potential artifact removal, we are working on automatic-noise-detection/classification (Huot et al., 2017). It would also be interesting to see how site-specific these results are by testing these methods on the grid of fibers installed in Fairbanks, AK by LBL and the Corps of Engineers. After we understand the theory for straight fibers better (particularly the role of C_{RT} and C_{TR}, a separate theory could be built off of this for helically wound cables. This kind of theory may also be useful if one or two components of a geophone go out in a very sparse array where wavefield interpolation isn't possible.

CONCLUSIONS

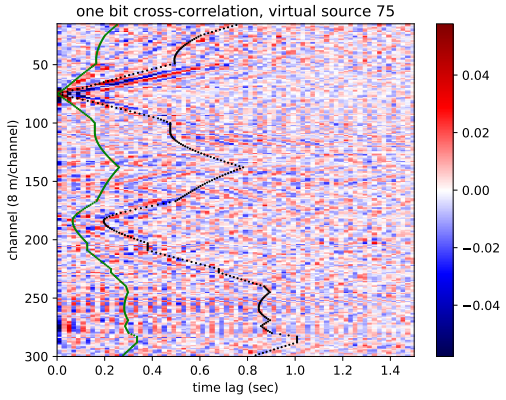
We have extended the theory of ambient noise interferometry to understand cross-correlations of all channels throughout two-dimensional DAS arrays. We have tested this theory on cross-correlations throughout SDASA-1, where we observed several features predicted by our theoretical model. Two particular predicted features observed were a mix of two wave components in cross-correlations between orthogonal lines, and a wave showing up at a reasonable time for a Love wave on a line parallel to the virtual source. Further investigation and verification of this theory should be done so that we can understand how to image with the many additional ray-paths provided by correlations throughout the whole array.

ACKNOWLEDGEMENTS

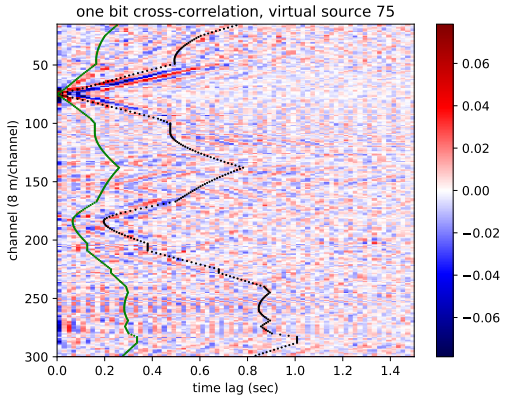
We thank OptaSense, Inc. for supplying the interrogator unit, and Martin Karrenbach and Steve Cole for their assistance in the installation and guidance in using the data. We thank affiliates of the Stanford Exploration Project for their support. Eileen Martin has also been supported in part by the U.S. DOE under grant No. DE-FG02-97ER25308 and a Schlumberger Innovation Fellowship. We thank our colleagues at Stanford for helpful discussions, particularly Jason Chang, Bob Clapp, Stew Levin and George Papanicolaou. We thank Stanford IT and SEEES IT for assistance running the array, and Stanford Center for Computational Earth and Environmental Sciences for computing resources. This work has been influenced by our collaborations with Lawrence Berkeley National Laboratory where E. Martin is an affiliate, and we have particularly benefited from discussions on DAS interferometry with Jonathan Ajo-Franklin, Nate Lindsey, and Shan Dou.



(a)



(b)



(c)

Figure 3: One-bit correlations of the array against a virtual source at channel 75 in the middle north-going line along Via Ortega. It improves from (left) 24 hours of correlations to (right) 72 hours of correlations, and (bottom) 120 hours. The arrival times of 1200 m/s and 400 m/s waves are overlaid. [**CR**] eileen2/. oneBitXcorr24HrsSrcCh75,oneBitXcorr72HrsSrcCh75,oneBitXcorr120HrsSrcCh75

REFERENCES

Ajo-Franklin, J., N. Lindsey, S. Dou, T. Daley, B. Freifeld, E. Martin, M. Robertson, C. Ulrich, and A. Wagner, 2015, A field test of distributed acoustic sensing for ambient noise recording: Expanded Abstracts of the 85th Ann. Internat. Mtg.

Chen, P., N. Taylor, K. Dueker, I. Keifer, A. Wilson, C. McGuffy, C. Novitsky, A. Spears, and W. Holbrook, 2016, psin: a scalable, parallel algorithm for seismic interferometry of large-n ambient noise data: Computers and Geosciences, **93**.

de Ridder, S., 2014, Passive seismic surface-wave interferometry for reservoir-scale imaging: PhD thesis, Stanford University.

Huot, F., Y. Ma, R. Cieplicki, E. Martin, and B. Biondi, 2017, Automatic noise exploration in urban areas : SEP-Report, **168**, 277–288.

Kuvshinov, B., 2016, Interaction of helically wound fibre-optic cables with plane seismic waves: Geophysical Prospecting, **64**, 671–688.

Lin, F., M. Moschetti, and M. Ritzwoller, 2008, Surface wave tomography of the western united states from ambient seismic noise: Rayleigh and love wave phase velocity maps: Geophysical Journal International.

Martin, E., J. Ajo-Franklin, S. Dou, N. Lindsey, T. Daley, B. Freifeld, M. Robertson, A. Wagner, and C. Ulrich, 2015, Interferometry of ambient noise from a trenched distributed acoustic sensing array: Expanded Abstracts of the 85th Ann. Internat. Mtg.

Martin, E., B. Biondi, S. Cole, and M. Karrenbach, 2017a, Overview of the Stanford DAS Array-1 (SDASA-1) : SEP-Report, **168**, 1–10.

Martin, E., B. Biondi, M. Karrenbach, and S. Cole, 2017b, Ambient noise interferometry from das array in underground telecommunications conduits: Technical Programme of the 79th Conference & Exhibition.

———, 2017c, Continuous subsurface monitoring by passive seismic with distributed acoustic sensors- the "stanford array" experiment: Proceedings of the First EAGE Workshop on Practical Reservoir Monitoring.

Martin, E., N. Lindsey, S. Dou, J. Ajo-Franklin, A. Wagner, K. Bjella, T. Daley, B. Freifeld, M. Robertson, and C. Ulrich, 2016, Interferometry of a roadside das array in fairbanks, ak: Expanded Abstracts of the 86th Ann. Internat. Mtg.

Nakata, N., J. Chang, J. Lawrence, and P. Boué, 2015, Body wave extraction and tomography at long beach, california, with ambient-noise interferometry: J. Geophys. Res. Solid Earth, **120**, 1159–1173.

Nakata, N., R. Snieder, T. Tsuji, K. Larner, and T. Matsuoka, 2011, Shear wave imaging from traffic noise using seismic interferometry by cross-coherence: Geophysics, **76**, SA97–SA106.

Shapiro, N., M. Campillo, L. Stehly, and M. Ritzwoller, 2005, High-resolution surface-wave tomography from ambient seismic noise: Science, **307**, 1615–1618.

Zeng, X., C. Thurber, H. Wang, D. Fratta, E. Matzel, and P. Team, 2017, High-resolution shallow structure revealed with ambient noise tomography on a dense array: Proceedings, 42nd Workshop on Geothermal Reservoir Engineering.

The search for P-waves at Forties

Jason P. Chang

ABSTRACT

Results from seismic interferometry at Apache Forties indicates the presence of P-wave events hidden in the ambient seismic noise field. I use phase-weighted stacking rather than linear stacking of correlations from quiet time periods to enhance this apparent P-wave energy. From source gathers with a virtual source located near the platform, there are three apparent events in the hydrophone-hydrophone and vertical-vertical geophone correlations between 40 and 80 Hz. From a tau-p transform of these gathers, there are events in the hydrophone component propagating at 1500 m/s and 3000 m/s with 0 s intercept time, while in the vertical-geophone component there are two events propagating at 3000 m/s with different intercept times. To determine the direction in which these events are traveling through the array, I look at source gathers along approximate lines of receivers in the north-south, east-west, and northeast-southwest directions. Using linear moveout with velocity estimated from the tau-p transform, I find that the slower event in the hydrophone correlations appears to be moving across the array from generally northeast to southwest, while the faster events in both component correlations appear to propagate from the platform. Based on the unlikelihood of interface waves traveling at such high velocities, I perform passive fathometry processing on ambient noise records in an attempt to recover body-wave energy. Preliminary results of passive fathometry appear to retrieve the water-column multiple along with some possible reflection events.

INTRODUCTION

Utilization of the interface-wave portion of Earth's ambient noise field to image the subsurface has been successful at the continental, regional, and local scales (e.g., Shapiro et al., 2005; Bensen et al., 2008; de Ridder and Dellinger, 2011). As a result, focus has shifted to extracting the body-wave portion of the Earth's ambient noise field for imaging. At the local scale and on land, Nakata et al. (2011) and Draganov et al. (2013) were able to recover reflection events from passive seismic data. Additionally, Nakata et al. (2015) was able to recover diving P-waves using a dense array at Long Beach, California, which were subsequently used to produce a 3-dimensional tomographic image of the subsurface. There has been similar success in finding non-interface waves in shallow marine ambient noise fields. Using continuous recordings from an ocean-bottom cable network in the Valhall oil field, Mordret et al. (2013) extracted an apparent acoustic wave generated by an operating platform. Brooks and Gerstoft (2009) extracted direct, sea-surface reflected, and sea-bottom reflected events in the water column from shipping and wave noise recorded by a small L-shaped (vertical and horizontal) hydrophone array in offshore New Jersey. Using a technique called passive fathometry, numerous studies (e.g., Gerstoft et al., 2008; Siderius et al., 2010) obtained shallow subsurface reflection images using energy generated by breaking waves at the sea surface recorded by a vertical array in the water column.

Here, I build off results from Chang (2016), who used seismic interferometry processing to extract P-waves that appeared to be propagating from the platform at Apache Forties. The goal of this report is to better classify the apparent P-waves in the hydrophone-hydrophone and vertical-vertical geophone correlations, and to determine whether these events can be used to image the subsurface. First, I introduce the continuous recordings from the Apache Forties dataset. Second, I apply seismic interferometry processing with phase-weighted stacking to generate high-frequency virtual source gathers throughout the array. To get a sense of velocities of the events, I perform a tau-p transform. To get a sense of the directionality of the events, I examine virtual source gathers along approximate lines of receivers in the north-south, east-west, and northeast-southwest directions. I then perform linear moveout on these source gathers using velocities estimated from the tau-p transform to better identify the different events. Finally, I show preliminary results from passive fathometry processing, which appears to retrieve the water-column multiple along with possible reflection events.

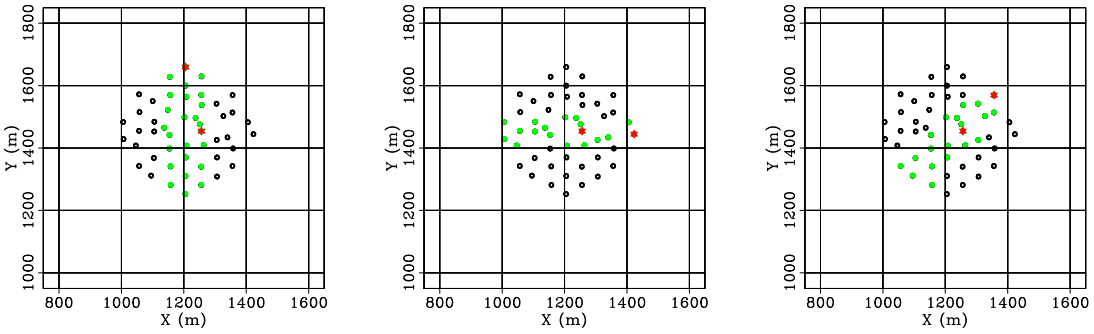
FORTIES CONTINUOUS RECORDINGS

The Forties data set, provided to SEP by the Apache Corporation, consists of three groups of ocean-bottom nodes (OBNs) centered at three different platforms in the North Sea. The four-component nodes were deployed as part of an active seismic survey aimed at imaging shallow gas pockets that could pose potential drilling hazards. The nodes were continuously recording, and because active seismic shooting had to be suspended for a couple of days due to severe weather conditions, there were enough quiet periods for ambient noise studies. Furthermore, Brooks and Gerstoft (2009) showed that stormy weather produces more breaking waves at the sea surface, which enable better extraction of vertically-propagating energy in the water column.

For this study, I examine the hydrophone and vertical-geophone components of the Bravo group of OBNs. This cluster of OBNs consists of 52 nodes arranged in a hexagonal shape and centered around an operating platform (Figure 1). The average node spacing is 50 m, and the maximum array offset is approximately 400 m. Each node continuously records for approximately 4 days at 2 ms sampling and is located roughly 120 m below the sea surface. To prepare the recordings for ambient noise processing, I truncate start times up to the nearest quarter of an hour and truncate the end times down to the nearest quarter of an hour. These times are not the same for all nodes, as they were not all deployed simultaneously. Additionally, I remove all times that contain clear active seismic shooting. For example, Figure 2(a) shows the full hydrophone power spectrogram for a node near the platform (centered red star in maps in Figure 1), while Figure 2(b) shows the hydrophone power spectrogram for the same node after the times of active seismic shooting have been removed. I use the data in the latter figure for ambient noise analysis.

PASSIVE SEISMIC INTERFEROMETRY

Passive seismic interferometry is performed by cross-correlating the recordings of ambient seismic noise at two receivers. Under certain conditions, the result is an estimate of the Green's function between the two receivers (Wapenaar et al., 2010). By correlating the recording at one receiver with recordings from all other receivers, we can create virtual



(a) (b) (c)

Figure 1: Maps of nodes at Forties. Red dots indicate virtual source locations. Green dots indicate receivers used for virtual source gathers. The operating platform is located in the center of the array. (a) North-south receiver line. (b) East-west receiver line. (c) Northeast-southwest receiver line. [**CR**] jasonpc1/. map-ns,map-ew,map-nesw

source gathers.

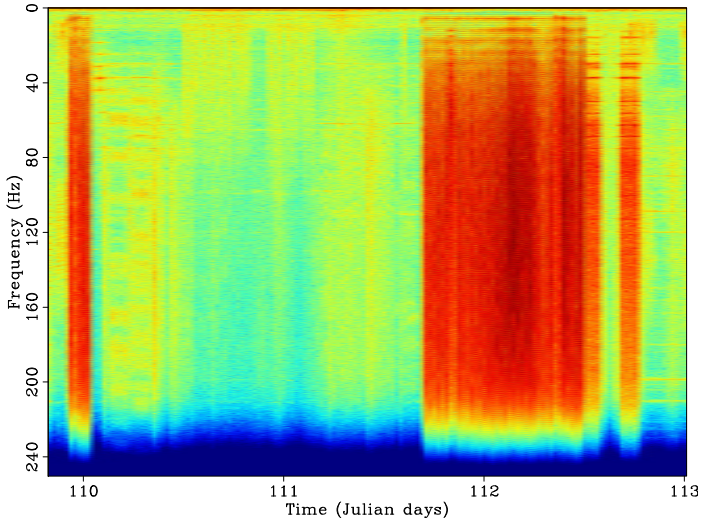
Processing

I perform the ambient noise cross-correlation technique on the hydrophone and vertical-geophone components of the data. Here, I modify the processing procedure outlined in Chang (2016). To recap, I first ensure that the recordings are synchronized in time and remove all spurious nodes (leaving 49 of 52 nodes). I then divide the recordings into 30-minute time windows with 50% overlap (for a total of 193 time windows, spanning over 2 days). Next, I perform ambient noise cross-correlation by calculating the averaged whitened coherency between each pair of nodes for each time window. This procedure is also referred to as calculating the cross-coherence. In the frequency domain, the procedure is generally expressed as:

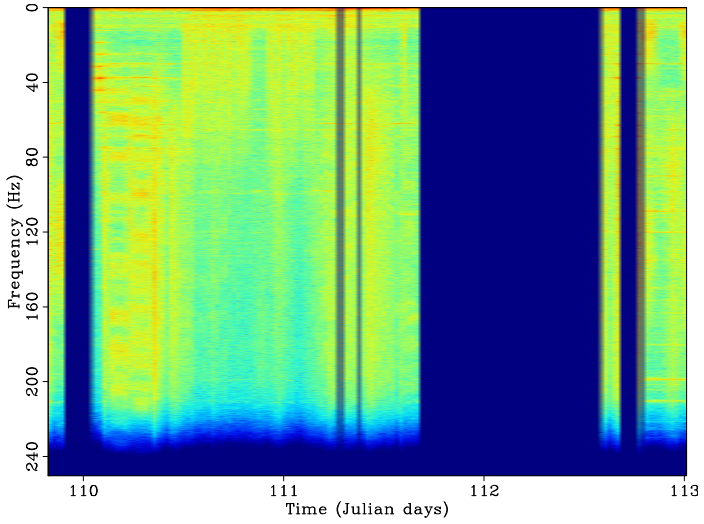
$$[G(x_B, x_A, \omega) + G^*(x_B, x_A, \omega)] = \left\langle \left(\frac{U(x_B, \omega)}{\{|U(x_B, \omega)|\}} \right) \left(\frac{U^*(x_A, \omega)}{\{|U(x_A, \omega)|\}} \right) \right\rangle, \tag{1}$$

where G is the Green's function between two receiver locations (x_A, x_B), $U(x, \omega)$ is the spectrum of the wavefield at a given receiver location x, $*$ is the complex conjugate, $\langle \cdot \rangle$ is an averaging operation, $|\cdot|$ is the magnitude of the spectrum, and $\{\cdot\}$ is a 0.003 Hz running window average used for normalizing the signal.

The processing procedure here differs from this point on. First, I scale the output of each correlation by the inverse of its maximum amplitude as an extra effort to suppress the effect of potential spurious instrument spikes and sudden large-amplitude events. I then perform phase-weighted stacking across all time windows rather than the traditional linear stack. Phase-weighted stacking is a method that suppresses incoherent noise better than a linear stack, and can thus be effective at detecting weak but coherent arrivals (Schimmel and Paulssen, 1997). The method essentially weighs each time sample from the linear stack by a stack based on the corresponding instantaneous phase (called a phase stack). Specifically,



(a)



(b)

Figure 2: Power spectrograms for a hydrophone near the platform (centered red star in maps in Figure 1). (a) Full spectrogram. (b) Spectrogram with active sources removed. Power is plotted in log scale. Colors in plots are scaled the same, so they can be compared to each other. [**CR**] jasonpc1/. sp-all,sp-keep

the weight is the stack of the corresponding sample-by-sample-normalized analytic traces. In equation form, the phase-weighted stacking procedure is expressed as:

$$p(t) = \frac{1}{N}\sum_{j=1}^{N} s_j(t) \left| \frac{1}{N}\sum_{k=1}^{N} \exp\left[i\Phi_k(t)\right] \right|^v , \tag{2}$$

where $p(t)$ is the phase-weighted stack result, N is the number of traces being stacked, $s_j(t)$ are the traces being stacked, $\Phi_k(t)$ is the instantaneous phase of the trace, and v is a phase sharpness factor (chosen empirically here to have a value of 2). Thus, time samples with incoherent instantaneous phase are weighted toward 0, while time samples with coherent instantaneous phase are weighted toward 1. Though this stack is not a linear process, the consistent signal in the stack should not be distorted much, while incoherent noise in the stack is damped to allow weaker coherent signals to become more apparent. I use this processing procedure to enhance the apparent P-waves in the virtual source gathers from Chang (2016).

Virtual source gathers

I first create source gathers with a virtual source located near the platform (centered red star in maps in Figure 1) for frequencies between 40 and 80 Hz. Traces are sorted by absolute offset. To compare the effects of different processing on virtual source gathers, I plot the results from processing with linear stacking from Chang (2016) and with phase-weighted stacking (described in the previous section). Compared to the former processing procedure (top row, Figure 3), phase-weighted stacking (bottom row, Figure 3) appears to sharpen the events in the virtual source gathers. The non-linear stacking procedure appears to particularly reduce noise in the vertical-vertical geophone correlations (right column, Figure 3), allowing the secondary event at later positive time lags to become more apparent. Though phase-weighted stacking appears to weaken the dominant arrivals in the hydrophone-hydrophone correlations (left column, Figure 3), it seems to also dampen the overall amount of noise as well as reveal a faster-propagating event at positive times lags not apparent in the original processing procedure.

Overall, we see clear arrivals in the correlations from both components. Looking at the hydrophone-hydrophone correlations (left column, Figure 3), we see significant energy at both positive and negative time lags. This suggests that there is seismic energy not only traveling away from the platform (as expected), but there is also a significant amount of energy traveling toward the platform. Additionally, while the phase-weighted stack may appear to degrade the phase at far offsets, it does reveal hints of a faster-propagating event at positive time lags not clearly observed when a linear stack is employed. By performing a tau-p transform on the hydrophone-hydrophone correlations (Figure 4(a)), I estimate that the primary event is propagating at approximately 1500 m/s (0.00067 s/m), while the faster event is propagating at approximately 3000 m/s (0.00033 s/m). Both events appear to have 0 s intercept time. Looking at the vertical-vertical geophone correlations (right column, Figure 3), there are two clear events at positive time lags. This suggests that both events are generated by the platform. By performing a tau-p transform on the vertical-vertical geophone correlations (Figure 4(b)), I estimate that both events are propagating at approximately 3000 m/s (0.00033 s/m), which is similar to the velocity of the faster event in the hydrophone-hydrophone correlations that was enhanced by phase-weighted stacking.

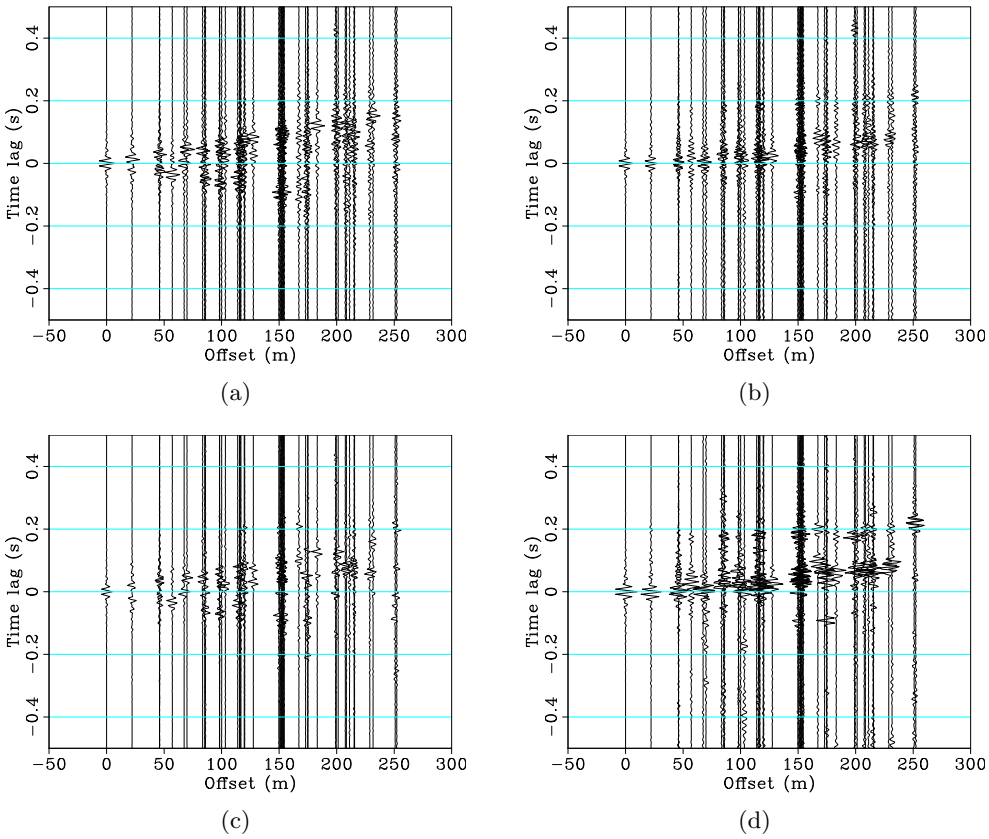
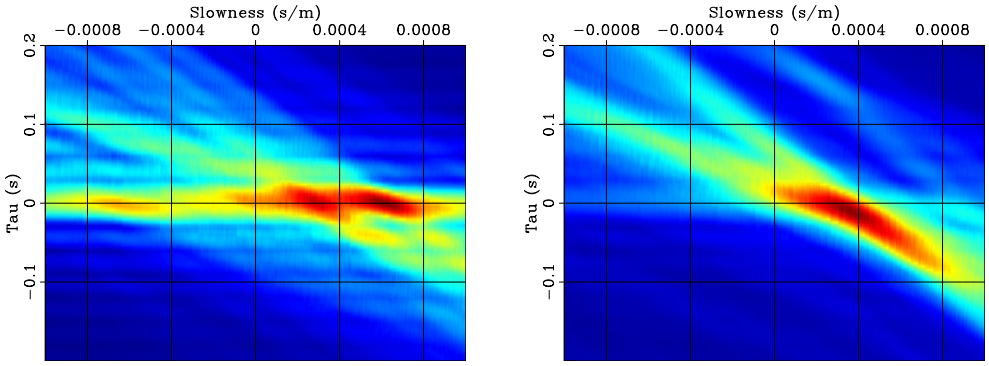


Figure 3: Source gathers for frequencies between 40 and 80 Hz with the virtual source located near the platform (centered red star in maps in Figure 1). Traces are sorted by absolute offset. Top row: cross-coherence processing with linear stacking. Bottom row: cross-coherence processing with temporal normalization and phase-weighted stacking. Left column: hydrophone-hydrophone correlations. Right column: vertical-vertical geophone correlations. [**CR**] jasonpc1/. s32-ww-abs-HH,s32-ww-abs-VV,s32-pws-abs-HH,s32-pws-abs-VV



(a) (b)

Figure 4: Tau-p transforms of the phase-weighted stacked virtual source gathers (bottom row, Figure 3). (a) Hydrophone-hydrophone correlations. (b) Vertical-vertical geophone correlations. Warmer colors indicate higher stacking power in the tau-p domain. The weaker event around intercept time 0.14 s and slowness 0.00033 s/m in (b) is likely related to the secondary event in Figure 3(d). [**CR**] jasonpc1/. s32-pws-taup-h,s32-pws-taup-v

Note that the secondary event around intercept time 0.14 s and slowness 0.00033 s/m in the tau-p transform is very weak compared to the primary event. I also estimate that the intercept time of the primary event is 0 s, while the intercept time of the weaker secondary event is approximately 0.15 s. Thus, from these virtual source gathers and the corresponding analysis, I determine that there are three different types of events:

1. A linear event propagating at 1500 m/s towards and away from the platform, observed in the hydrophone-hydrophone correlations (Figure 3(c)).

2. A linear event propagating at 3000 m/s with 0 s intercept time away from the platform, observed in both the hydrophone-hydrophone and vertical-vertical geophone correlations (Figures 3(c) and 3(d), respectively).

3. A weaker linear event propagating at 3000 m/s with 0.15 s intercept time away from the platform, observed in the vertical-vertical geophone correlations (Figure 3(d); most evident at far offsets).

To get a better sense of the directionality (not just towards and away from the platform), I create virtual source gathers with approximate lines of receivers in the north-south, east-west, and northeast-southwest directions (maps in Figure 1). All source gathers here use phase-weighted stacking rather than linear stacking. I first examine source gathers along the north-south direction with a virtual source located in the north of the array. From the hydrophone-hydrophone correlation (Figure 5(a)), we can see two types of events. At positive time lags, we see a clear event moving away from the virtual source. This suggests that this particular event is not moving north from the platform, but is moving south toward the platform. Because energy is moving away from the virtual source, I perform linear moveout (LMO) centered at the virtual source location using the previously estimated velocity of 1500 m/s from the tau-p transform and find that it flattens this particular event

(Figure 5(b)). At negative time lags we see a faster event with a V shape indicative of a strong source in the array (Chang et al., 2016). The peak of the V shape is located near the center of the array, which is where the platform is approximately located. Because energy appears to move away from the platform location, I perform LMO centered at the platform location (about -200 m offset) with a velocity of 3000 m/s to flatten this particular event (Figure 5(c)). Looking at the vertical-vertical geophone correlations (Figure 5(d)), we see that nearly all the energy is at negative time lags and that the moveout of both events follow a V shape centered at the platform location. Again, performing LMO with a velocity of 3000 m/s centered at the platform location flattens these particular events (Figure 5(e)).

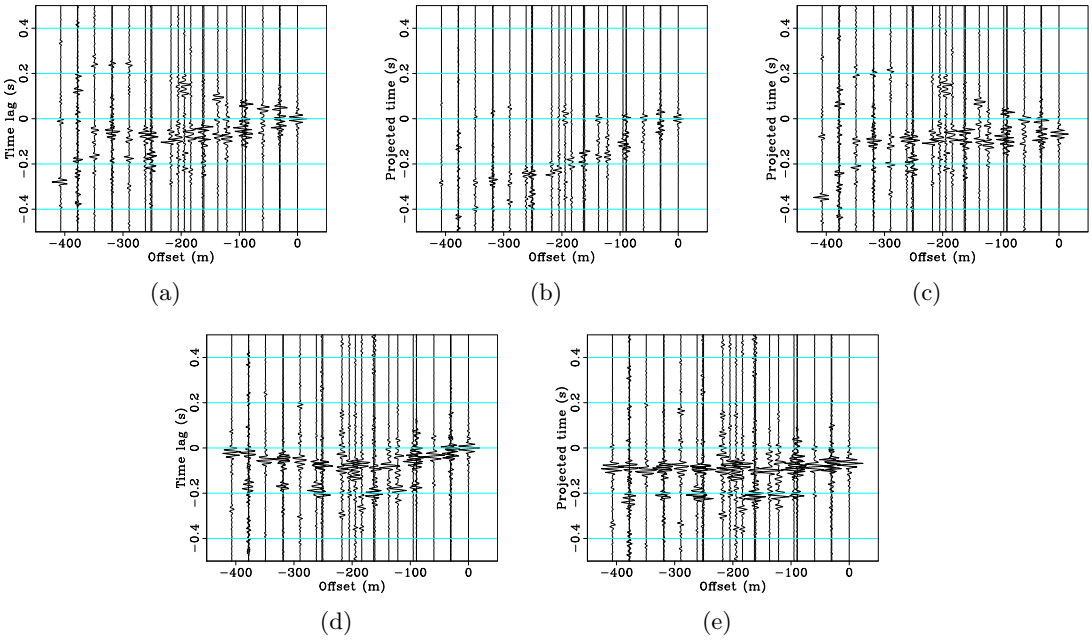


Figure 5: Virtual source gathers after phase-weighted stacking for frequencies between 40 and 80 Hz with receivers along approximate lines in the north-south direction. Virtual source is north in the array (see Figure 1(a)). Top row: hydrophone-hydrophone correlations. (a) Original source gather. (b) Source gather after LMO at 1500 m/s centered at the source location (0 m offset). (c) Source gather after LMO at 3000 m/s centered at the platform location (-200 m offset). Bottom row: vertical-vertical geophone correlations. (d) Original source gather. (e) Source gather after LMO at 3000 m/s centered at the platform location (-200 m offset). Gathers are sorted by offset in the north-south direction. [**CR**] jasonpc1/. s03-HH,s03-HH-1500,s03-HH-3000,s03-VV,s03-VV-3000

Similar observations can be made when looking at source gathers with an apparent line of receivers in the east-west direction and a virtual source in the east of the array. In the hydrophone-hydrophone correlations (Figure 6(a)), we can again see two events: the linear event moving away from the virtual source (toward the west), and the V-shaped event at negative time lags centered at the platform location. Performing LMO with a velocity of 1500 m/s centered at the source location flattens the event at positive time lags (Figure 6(b)). Performing LMO with a velocity of 3000 m/s centered at the platform location (about -200 m offset) flattens the event at positive time lags (Figure 6(c)). Much like the north-south correlations, we can see two events at negative time lags in the vertical-

vertical correlations along the east-west direction (Figure 6(d)). Again, performing LMO with a velocity of 3000 m/s centered at the platform location flattens these two events (Figure 6(e)).

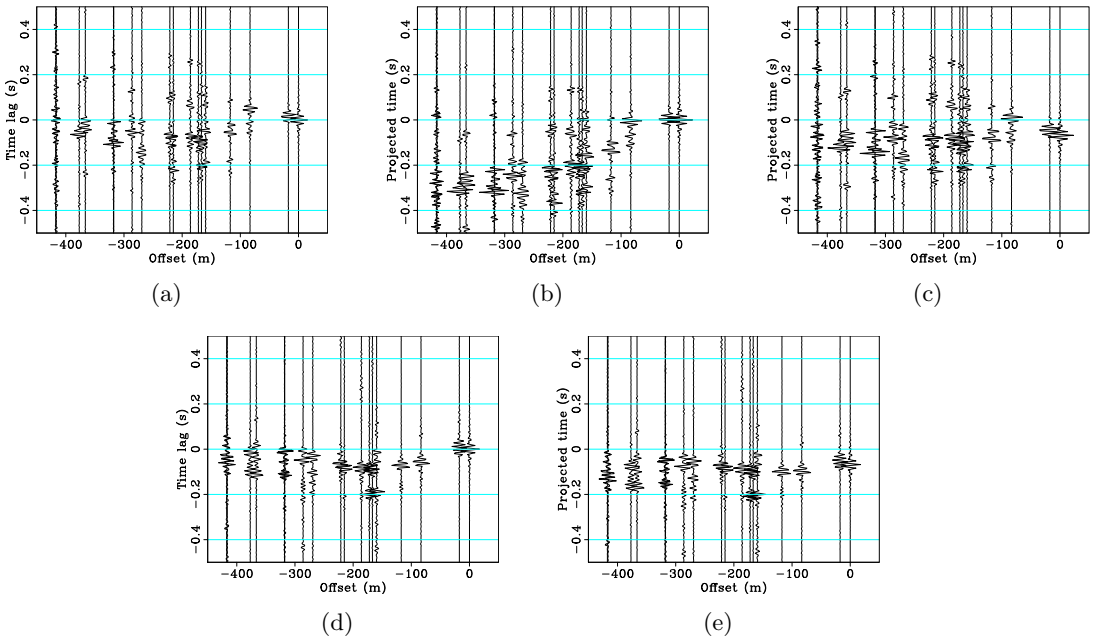


Figure 6: Virtual source gathers after phase-weighted stacking for frequencies between 40 and 80 Hz with receivers along approximate lines in the east-west direction. Virtual source is east in the array (see Figure 1(b)). Top row: hydrophone-hydrophone correlations. (a) Original source gather. (b) Source gather after LMO at 1500 m/s centered at the source location (0 m offset). (c) Source gather after LMO at 3000 m/s centered at the platform location (-200 m offset). Bottom row: vertical-vertical geophone correlations. (d) Original source gather. (e) Source gather after LMO at 3000 m/s centered at the platform location (-200 m offset). Gathers are sorted by offset in the east-west direction. [**CR**] jasonpc1/. s05-HH,s05-HH-1500,s05-HH-3000,s05-VV,s05-VV-3000

Given that the slower event in the hydrophone-hydrophone correlations appears to propagate at water velocity along the south and west directions, I examine source gathers with a virtual source in the northeast of the array and an apparent line of receivers along the northeast-southwest direction. The goal is to determine if the energy is propagating toward the southwest from a distinct source in the northeast or propagating along a number of azimuths between due south and due west. As in the hydrophone-hydrophone correlations along the other two directions, we see that there are two clear events: the linear event moving away from the virtual source (toward the southwest), and the faster event at negative time lags. Performing LMO with a velocity of 1500 m/s centered at the source location again flattens the event at positive time lags (Figure 7(b)). Therefore, it appears that this particular event propagates at water velocity (not slower than water velocity) along the northeast-southwest direction. This suggests that this event does not originate from a single source in the northeast, but rather originates from a number of sources along azimuths between due north and due east (leading to propagation along azimuths between due south and due west, respectively). For completeness, I also perform LMO centered at the virtual

source with a velocity of 3000 m/s to flatten the faster event at negative time lags (Figure 7(c)). As for the vertical-vertical geophone correlations, the results are similar to those from the other two apparent lines of receivers. There are multiple events that appear to originate from the platform location (Figure 7(d)), and performing LMO centered at the platform with a velocity of 3000 m/s appears to flatten those events (Figure 7(e)).

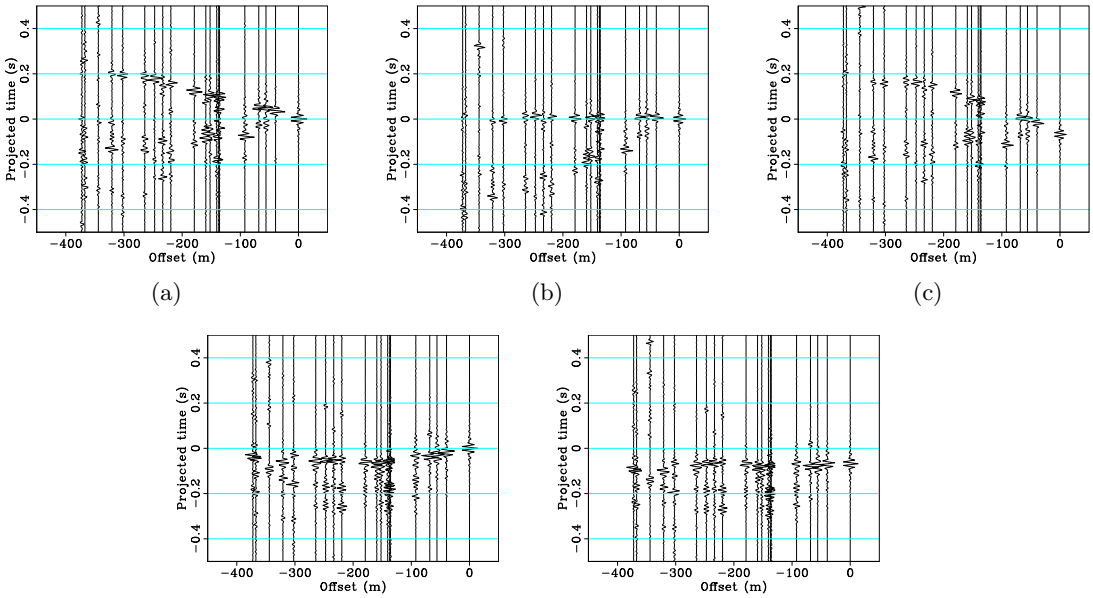


Figure 7: Virtual source gathers after phase-weighted stacking for frequencies between 40 and 80 Hz with receivers along approximate lines in the northeast-southwest direction. Virtual source is northeast in the array (see Figure 1(b)). Top row: hydrophone-hydrophone correlations. (a) Original source gather. (b) Source gather after LMO at 1500 m/s centered at the source location (0 m offset). (c) Source gather after LMO at 3000 m/s centered at the platform location (-200 m offset). Bottom row: vertical-vertical geophone correlations. (d) Original source gather. (e) Source gather after LMO at 3000 m/s centered at the platform location (-200 m offset). Gathers are sorted by offset in the northeast-southwest direction. [**CR**] jasonpc1/. s47-HH,s47-HH-1500,s47-HH-3000,s47-VV,s47-VV-3000

Therefore, from these virtual source gathers, we can elaborate on each of the three previously mentioned events:

1. The linear event propagating at 1500 m/s observed in the hydrophone-hydrophone correlation appears to be arriving from outside the array and propagating along multiple azimuths between due south and due west.

2. The linear event propagating at 3000 m/s with 0 s intercept time observed in both the hydrophone-hydrophone and vertical-vertical geophone correlations is generated by the platform.

3. The linear event propagating at 3000 m/s with 0.15 s intercept time observed in the vertical-vertical geophone correlations is also generated by the platform.

Given the apparent velocities and directions of these three events, we are likely looking at acoustic P-waves traveling in the water or body waves. The event traveling at 1500 m/s is potentially energy moving horizontally across the array at water velocity, which would explain why this event is so weak in the vertical-vertical geophone correlations. The origin of this signal could be some distant sources between due north and due east of the array. Since the sea-surface and sea-bottom act as a waveguide, signals from distant sources experience little attenuation and typically display mostly horizontally-propagating wavefronts (Brooks and Gerstoft, 2009). One potential source is distant shipping noise, which dominates the marine ambient seismic noise field at frequencies between 25 and 250 Hz (Wenz, 1972). Another potential source is active seismic shooting from another survey in the North Sea.

As for the events traveling at 3000 m/s, it is likely to be energy traveling with some vertical component of movement, as it is observed in both the hydrophone-hydrophone and vertical-vertical geophone correlations. Though it could be energy propagating through the subsurface, it seems more likely that it is energy traveling through the water. This is because the secondary event propagating at 3000 m/s in the vertical-vertical geophone correlations has an intercept time of approximately 0.15 s, which is approximately the two-way traveltime of a vertically-propagating event traveling through a water column 120 m deep at 1500 m/s. Additionally, the source of this high-velocity event appears to be the platform. Given that there was a severe storm during the continuous recordings, it is possible that ocean waves crashing against the platform could produce these types of events. Given an apparent velocity of 3000 m/s and assuming a plane wave traveling through the water at 1500 m/s, P-wave energy would be hitting the array at approximately 30 degrees with respect to normal. Assuming there is some vertical component to these high-velocity events, there is potential to obtain subsurface reflections. I turn to passive fathometry processing in an attempt to extract these types of reflections from the ambient seismic noise field.

PASSIVE FATHOMETRY

Passive fathometry is based on the idea that the cross-correlation of energy generated from breaking waves at the sea surface and its echo return from the seabed can be used to estimate travel times to subsurface reflectors (Gerstoft et al., 2008; Siderius et al., 2010). Therefore, the technique is similar to passive seismic interferometry in that they are both based on the cross-correlation of noise wavefields. Here, I apply passive fathometry processing on individual receivers in an attempt to extract reflection profiles beneath select OBNs.

Processing

Passive fathometry has been successfully applied to vertical arrays of hydrophones and geophones in shallow marine environments (Gerstoft et al., 2008; Siderius et al., 2010). The orientation of these arrays allowed for efficient separation of up- and down-going events using an adaptive beamforming approach. The cross-correlation of these two events can produce a reflection sequence beneath the array. Here, I adapt this processing procedure for use with horizontally-oriented ocean-bottom arrays.

Many ocean-bottom arrays contain both vertical-geophone and hydrophone components, making it possible to perform up- and down-going wavefield separation using PZ summation.

In the acoustic decomposition (Schalkwijk et al., 1999), the up- and down-going pressure wavefields are obtained by:

$$P_{\text{up}}(\omega,k) = \frac{1}{2}\left[P(\omega,k) + a(\omega)\frac{\rho}{q(\omega,k)}Z(\omega,k)\right] , \tag{3}$$

and

$$P_{\text{down}}(\omega,k) = \frac{1}{2}\left[P(\omega,k) - a(\omega)\frac{\rho}{q(\omega,k)}Z(\omega,k)\right] , \tag{4}$$

respectively, where P is the pressure component, Z is the vertical velocity component, ω is frequency, k is horizontal wavenumber, ρ is density of water, and q is vertical slowness in the water layer defined as $q(\omega,k) = \sqrt{c^{-2} - p^2(\omega,k)}$, where c is water velocity and p is the ray parameter.

I am interested in passive up-down separation at each individual node (not between nodes). For this report, I will assume that passive energy arrives at near-vertical incidence. This assumption allows some simplifications when estimating the calibration filter for PZ summation using active-source seismic data. First, I can focus on performing up-down separation on the zero-offset trace rather than the entire gather. Second, I can take the vertical slowness in the water layer to be the inverse of water velocity. Equations 3 and 4 thus simplify to:

$$P_{\text{up}}(\omega) = \frac{1}{2}\left[P(\omega) + a(\omega)\rho c Z(\omega)\right] , \tag{5}$$

and

$$P_{\text{down}}(\omega) = \frac{1}{2}\left[P(\omega) - a(\omega)\rho c Z(\omega)\right] . \tag{6}$$

To estimate the calibration filter, I follow the approach from Biondi and Levin (2014) to estimate the calibration filter from the zero-offset active-source data. This filter is typically estimated by minimizing the energy of the down-going pressure field in a window containing only up-going events. Here, I use a time window around a reflection event in the active-source data rather than a refraction event. First, I remove the instrument response from both the hydrophone and vertical-geophone recordings. Second, I manually select a time window around a reflection event near zero-offset for each node of interest. Finally, I compute a Wiener shaping filter that shapes the vertical-component reflection to the reflection recorded by the hydrophone within the picked window. This particular filter incorporates the density and slowness factors in Equations 5 and 6.

An example of the result of acoustic up-down separation on active source data is shown in the common receiver gathers for the receiver in the north of the array in Figure 8. Gathers have been bandpassed for frequencies between 20 and 220 Hz. Though the calibration filter is only applied to a time window in the zero-offset trace, the acoustic decomposition appears to be fairly effective throughout the gathers. The reflections just after the direct arrival are apparent in the upgoing pressure wavefield and nearly non-existent in the downgoing pressure wavefield. Additionally, the first water-column multiple (approximately 0.26 s at zero-offset) is weaker in the upgoing pressure wavefield than in the downgoing pressure wavefield.

With a reasonable calibration coefficient between the hydrophone and vertical-geophone recordings, I apply the same up-down separation processing to the same 30-minute time windows used for seismic interferometry in the previous section. As with the active source

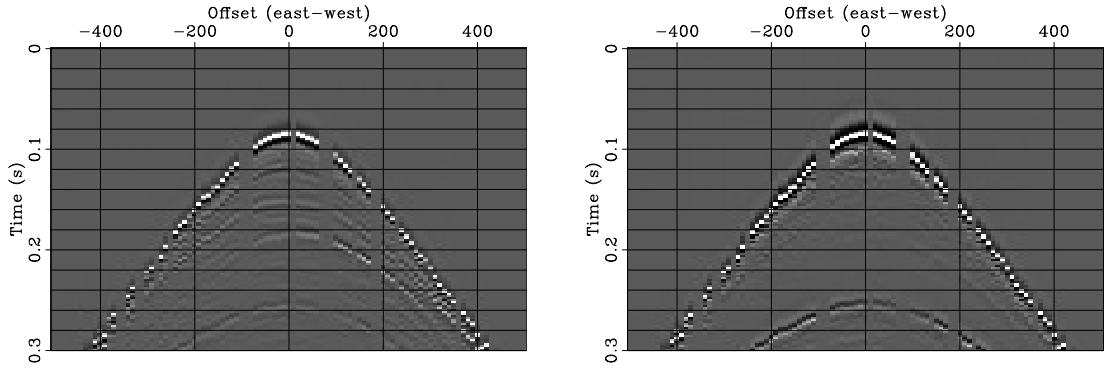


Figure 8: Common receiver gathers from active-source data for a receiver in the north of the array after acoustic up-down separation. (a) Upgoing acoustic wavefield. (b) Downgoing acoustic wavefield. Gathers are bandpassed for frequencies between 20 and 220 Hz. [**CR**] jasonpc1/. updesigbp-03,downdesigbp-03

data, I first remove the instrument responses from the passive recordings. I then bandpass the recordings for frequencies between 20 and 220 Hz. Typical passive fathometry uses acoustic energy up to kHz, so I examine up to fairly high frequency here. I then perform PZ summation on the passive recordings using the corresponding calibration coefficient for each node. Finally, I correlate the resulting up- and down-going wavefields from each time window and combine them with a phase-weighted stack.

1D profiles

Preliminary results from passive fathometry are shown for four node locations in Figure 9. A gain function t has been applied to each of the 1D profiles to limit the spike at zero-lag time and to enhance later arrivals. All of the 1D profiles share similar traits. One of the prominent events is the apparent arrival between 0.14 and 0.16 s. As discussed previously, the water depth in this area is approximately 120 m. Assuming a water velocity of 1500 m/s, the water-column multiple would take about 0.16 s to travel up and down. Additionally, there appears to consistently be an event between 0.3 and 0.32 s, which could be related to the second water-column multiple (two trips up and down). It should be noted that while these events arrive at the expected times of the water-column multiples, they do not display the expected opposite polarity. More investigation is needed to address this characteristic, as it might help determine whether these events are indeed water-column multiples. While there appear to be a few arrivals between the potential water-column multiples in each of the profiles, their arrival times do not obviously correspond to any events in the active-source data.

Furthermore, there appears to be an event before the first apparent water-column multiple, which is when we would expect to observe any near-surface reflections. Whether this event at 0.08 s is an actual reflection remains to be determined. The arrival time is consistent in all fathometry results, which could indicate a flat reflector or be an artifact of

processing. Additionally, this is the only clear event before the first water-column multiple, whereas the active-source data reveals multiple reflections before the first multiple. A careful comparison with the active-source data will be required to determine whether these preliminary fathometry results are indeed identifying subsurface reflection events. Overall, passive fathometry processing appears to extract the water-column multiples as well as potential reflection events.

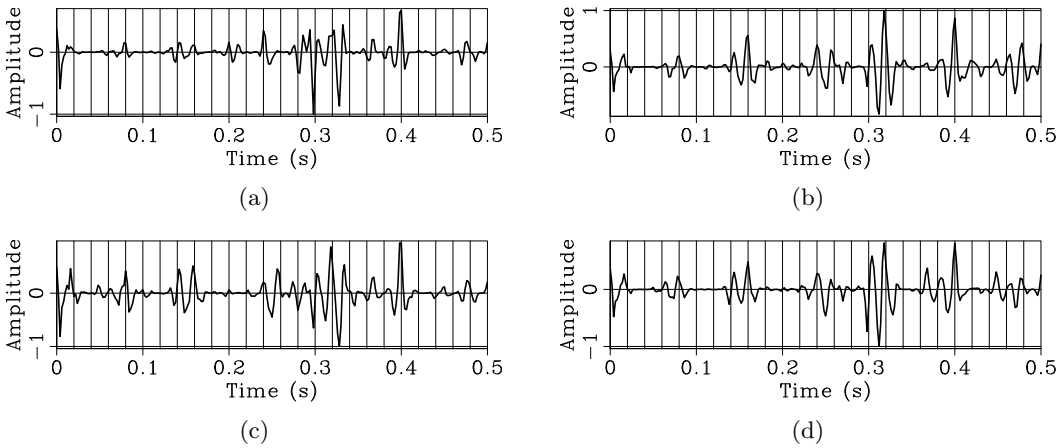


Figure 9: Results from passive fathometry at four different node locations. (a) Southwest edge in the array. (b) North edge in the array (corresponding to the receiver location in Figure 8). (c) East edge in the array. (d) Southwest near the platform. Results are for frequencies between 20 and 220 Hz. Gain function t is applied to limit the spike at zero-lag time and to enhance later arrivals. [**CR**] jasonpc1/. fathometer-00,fathometer-01,fathometer-02,fathometer-03

Clearly, there can be improvements to the method. One improvement would be the development of a spatial filter that could use the geometry of arrays to isolate vertically-propagating energy. The current method assumes that most of the energy is vertically-propagating. This would be true if most of the energy at these frequencies were actually generated by breaking waves at the sea surface. However, this is unlikely to be a valid assumption (particularly for stations further from the platform) given that the platform appears to be a dominant source of energy, and given that shipping noise is the dominant source of energy at these frequencies. Another improvement would be to use only times when there was a storm. As shown in Brooks and Gerstoft (2009), more vertically-propagating events are extracted during periods when waves are breaking violently. Though this would result in the use of less recording time, I would only be using times when the desired signal would theoretically be strong. Overall, passive fathometry appears to be picking up possible physical events, which is a promising development for any future investigation into P-waves at Forties.

CONCLUSIONS

I investigated the apparent P-waves in the ambient seismic noise recordings from Apache Forties using seismic interferometry and passive fathometry processing. By using a phase-weighted stack in the seismic interferometry workflow, I showed that there are three ap-

parent events in the hydrophone-hydrophone and vertical-vertical geophone correlations. I used tau-p transforms and virtual source gathers with approximate lines of receivers to determine the velocity and directionality of these events. One event travels at 1500 m/s and appeared to travel across the array along azimuths between due south and due west. Because it is most apparent in the hydrophone-hydrophone correlations, I hypothesized that this event was a horizontally-propagating event likely originating from distant sources between due north and due east. Another event traveled at approximately 3000 m/s with 0 s intercept time, and appeared in the correlations from both components. The last event also traveled at approximately 3000 m/s but with 0.15 s intercept time, and appeared in only the vertical-vertical geophone correlations. I hypothesized that the origin of these events was the platform, and that they were likely reverberations in the water column. Given this interpretation of the faster events, I performed passive fathometry processing on the ambient noise in an attempt to harness vertically-propagating energy for extracting subsurface reflections. Preliminary results suggest that I again recovered the water-column multiple, as well as possible subsurface reflections.

ACKNOWLEDGMENTS

The author would like to thank Apache North Sea Limited for access to the data set and permission to publish. The author also thanks Biondo Biondi, Shuki Ronen, Bob Clapp, and Stewart Levin for help processing the data and useful discussions. Finally, the author would like to thank the sponsors of the Stanford Exploration Project for their financial support.

REFERENCES

Bensen, G., M. Ritzwoller, and N. Shapiro, 2008, Broadband ambient noise surface wave tomography across the United States: Journal of Geophysical Research: Solid Earth, **113**.

Biondi, E. and S. A. Levin, 2014, Up-down separation of ocean bottom node data using calibration filter based on critically refracted waves and adaptive subtraction: SEP-Report, **155**.

Brooks, L. A. and P. Gerstoft, 2009, Green's function approximation from cross-correlations of 20–100 Hz noise during a tropical storm: The Journal of the Acoustical Society of America, **125**, 723–734.

Chang, J., 2016, Multicomponent ambient noise crosscorrelation at Forties: 86th Annual International Meeting, SEG, Expanded Abstracts, 2709–2714.

Chang, J. P., S. A. de Ridder, and B. L. Biondi, 2016, High-frequency Rayleigh-wave tomography using traffic noise from Long Beach, California: Geophysics, **81**, B1–B11.

de Ridder, S. and J. Dellinger, 2011, Ambient seismic noise eikonal tomography for near-surface imaging at Valhall: The Leading Edge, **30**, 506–512.

Draganov, D., X. Campman, J. Thorbecke, A. Verdel, and K. Wapenaar, 2013, Seismic exploration-scale velocities and structure from ambient seismic noise (¿ 1 Hz): Journal of Geophysical Research: Solid Earth, **118**, 4345–4360.

Gerstoft, P., W. S. Hodgkiss, M. Siderius, C.-F. Huang, and C. H. Harrison, 2008, Passive fathometer processing: The Journal of the Acoustical Society of America, **123**, 1297.

Mordret, A., M. Landès, N. Shapiro, S. Singh, P. Roux, and O. Barkved, 2013, Near-surface study at the valhall oil field from ambient noise surface wave tomography: Geophysical Journal International, ggt061.

Nakata, N., J. P. Chang, J. F. Lawrence, and P. Boué, 2015, Body wave extraction and tomography at long beach, california, with ambient-noise interferometry: Journal of Geophysical Research: Solid Earth, **120**, 1159–1173.

Nakata, N., R. Snieder, T. Tsuji, K. Larner, and T. Matsuoka, 2011, Shear wave imaging from traffic noise using seismic interferometry by cross-coherence: Geophysics, **76**, SA97–SA106.

Schalkwijk, K., C. Wapenaar, and D. Verschuur, 1999, Application of two-step decomposition to multicomponent ocean-bottom data: Theory and case study: Journal of Seismic Exploration, **8**, 261–278.

Schimmel, M. and H. Paulssen, 1997, Noise reduction and detection of weak, coherent signals through phase-weighted stacks: Geophysical Journal International, **130**, 497–505.

Shapiro, N. M., M. Campillo, L. Stehly, and M. H. Ritzwoller, 2005, High-resolution surface-wave tomography from ambient seismic noise: Science, **307**, 1615–1618.

Siderius, M., H. Song, P. Gerstoft, W. S. Hodgkiss, P. Hursky, and C. Harrison, 2010, Adaptive passive fathometer processing: The Journal of the Acoustical Society of America, **127**, 2193.

Wapenaar, K., D. Draganov, R. Snieder, X. Campman, and A. Verdel, 2010, Tutorial on seismic interferometry: Part 1–basic principles and applications: Geophysics, **75**, 75A195–75A209.

Wenz, G. M., 1972, Review of underwater acoustics research: Noise: The Journal of the Acoustical Society of America, **51**, 1010–1024.

Comparing whales to marine seismic sources: low frequency sound generation by fin whales

Leighton M. Watson, Shuki Ronen, Jeremy A. Goldbogen, and Eric M. Dunham

ABSTRACT

Seismic surveys have been shown to cause behavioral changes in whales and have the potential to cause temporary or permanent physical harm. Whale calls have a similar frequency content to seismic airguns, and hence communication between animals is likely to be disrupted by exploration activities. We demonstrate that whale calls recorded by tags attached to fin whales have comparable amplitude to a marine vibroseis observed at a distance of 100 m and a seismic airgun at 1000 m. In addition, we discuss how whales may generate sound through resonance of their lungs in an analogous fashion to how seismic airguns generate sound through the resonance of the bubble of air ejected from the gun.

INTRODUCTION

There is significant concern about the impact of anthropogenic noise, a substantial amount of which is due to seismic surveys (Hildebrand, 2005), on marine life. Several studies have documented behavioral changes and temporary or permanent threshold shifts in the hearing of marine mammals as a result of exposure to seismic surveys (e.g., Gordon et al., 2003; Weilgart, 2013).

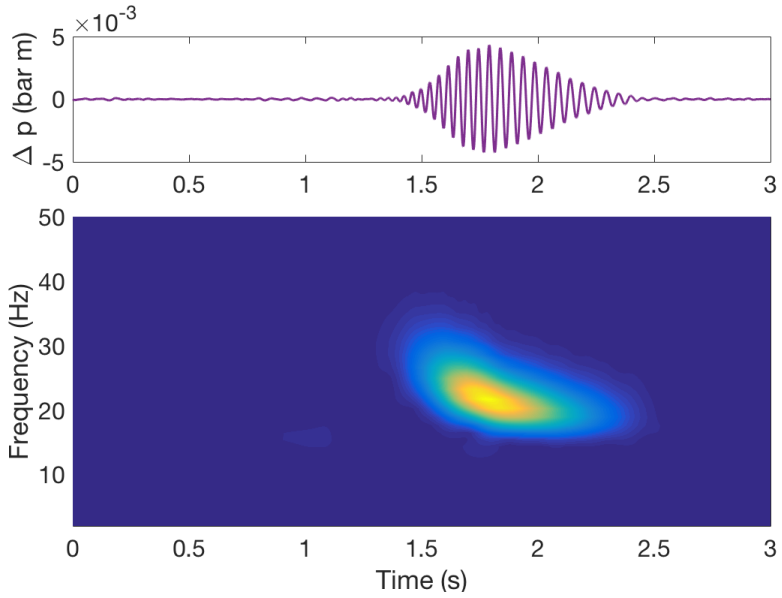
We compare the signals generated by seismic airguns and marine vibroseis with the observed whale calls. We focus on the impact on seismic surveys on fin whales. Fin whales generate a range of down-swept signals (Richardson et al., 1995; Aroyan et al., 2000) including characteristic low-frequency pulses around 20 Hz with a duration of approximately one second and higher frequency sounds sweeping from 75 to 40 Hz (Barham, 1973; Watkins, 1981; Stimpert et al., 2015). These signals overlap with the frequency range of seismic surveys, and hence environmental regulations have been imposed on marine seismic surveys out of concern that exploration activities may be harmful to whales or disruptive to their communication (Weir and Dolman, 2007; Compton et al., 2008).

We note that despite significant interest in the impact of seismic surveys on whale low frequency sound generation in whales remains an unsolved problem. Collecting data and performing experimental studies on marine mammals in general, and whales in particular, is fraught with difficulties. Novel studies have been performed on deceased animals in order to investigate physiological properties (e.g., Moore et al., 2011). However, these studies have been limited to smaller animals and the applicability of the results to live animals is debatable. Recent developments in multi-sensor acoustic recording tags (Johnson and Tyack, 2003; Johnson et al., 2009) and the ability to deploy these tags on live animals in the ocean (Goldbogen et al., 2014; Stimpert et al., 2015) have provided high-resolution acceleration and acoustic data that can be used to constrain models of sound generation.

Previous work has considered the possibility of cavities inside the whale acting as resonators (Barham, 1973; Aroyan et al., 2000). We further investigate this hypothesis and propose two possible mechanisms for sound generation: (1) impulsive excitation or (2) driven oscillation of the lungs.

WHALE CALLS

Multi-sensor acoustic recording tags, or DTAGs (Johnson and Tyack, 2003; Johnson et al., 2009) were attached to fin whales off the coast of southern California during the summer months of 2012 and 2013 (Goldbogen et al., 2014). The tags contain a hydrophone sampling at 240 kHz and a tri-axial accelerometer sampling at 200 Hz. They were attached to the outside of the whale in the vicinity of the lungs using suction cups.



(a)

Figure 1: Acoustic data from the DTAG on whale bp12-294a (Goldbogen et al., 2014). [**CR**] leighton1/. acousticData

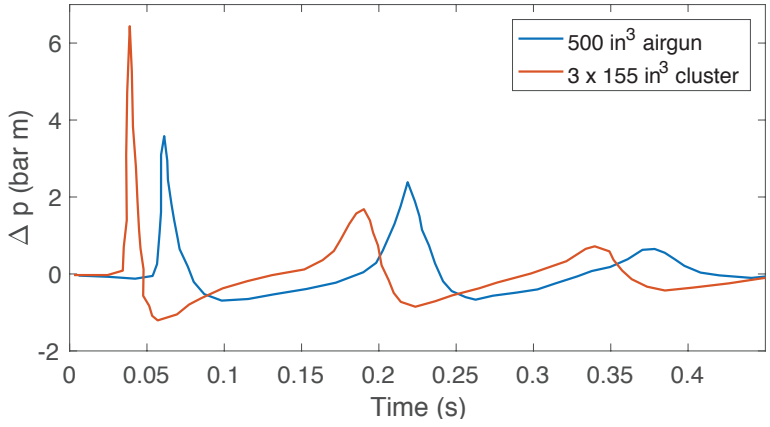
Tags were attached to the whales for several hours and recorded several hundred calls before the tags detached. The calls displayed remarkable consistency over time and between animals. The characteristic call, illustrated in Figure 1, was a down-swept signal from approximately 30 Hz to 15 Hz over a duration of slightly less than one second. The majority of calls were recorded while the whales were at depths of 15-20 m (Stimpert et al., 2015). For more details on the data collection the reader is referred to Goldbogen et al. (2014) and Stimpert et al. (2015).

SEISMIC SOURCES

Seismic airguns are the predominant source used in marine seismic surveys. An airgun functions by discharging highly pressurized air into the water forming a bubble that expands

and contracts. The signature is non-impulsive with a large peak from the initial expansion of the bubble and subsequent smaller peaks from the oscillations of the bubble. Tuned clusters of airguns (e.g., Dragoset, 2000) are used to increase the amplitude of the initial peak and generate destructive interference to reduce the bubble pulses.

Vibroseis works by moving an object up and down to transmit vibrations to the earth or water. Vibroseis is commonly used for land acquisition but the industry is still developing a fully operational marine vibroseis acquisition system. Vibroseis is thought to be less harmful to marine life (compared to seismic airguns) as it inputs a smaller amount of energy into the ocean over a longer period of time (Ogden, 2014).



(a)

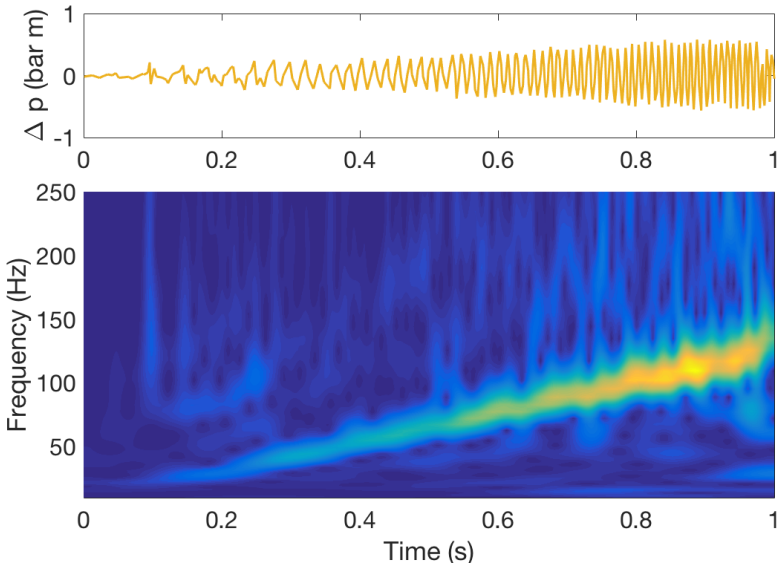
Figure 2: Pressure perturbation generated by a single airgun (blue) and by a cluster of three airguns (red). Data from IAGC (1995). [**CR**] leighton1/. seismicAirguns

COMPARING WHALE CALLS TO SEISMIC SOURCES

Seismic airguns (Figure 2) and marine vibroseis (Figure 3) are several orders of magnitude louder than the observed whale calls (Figure 1). We compare the whale calls recorded with the DTAG attached to the body of the whale with the pressure perturbation that would be recorded 100 m away from the marine vibroseis and 1000 m away from the two airgun signatures, assuming $1/r$ geometrical spreading (Figure 4).

We assume that a signal of comparable amplitude to the whale call will be unable to cause physical harm to the animal. This suggests that a whale 100 m away from a marine vibroseis or 1000 m away from a small array of airguns (a greater distance would be needed for a larger array of airguns) will not experience direct physical harm as a result of the seismic source. We note that this is only a rough estimation of how whales are harmed by seismic surveys. It is likely that the duration of exposure is important. Seismic surveys operate continuously for many hours whereas the whale signals are short lived and there are longer intervals between calls than between seismic shots. In addition, seismic airguns have a much faster onset than the emergent whale calls.

Seismic sources are likely to disrupt communication between whales. Figure 4 shows that seismic airguns generate substantial energy at the same frequency range as the fin



(a)

Figure 3: Signal generated by marine vibroseis. Data from IAGC (1995). [**CR**] leighton1/. marineVibe

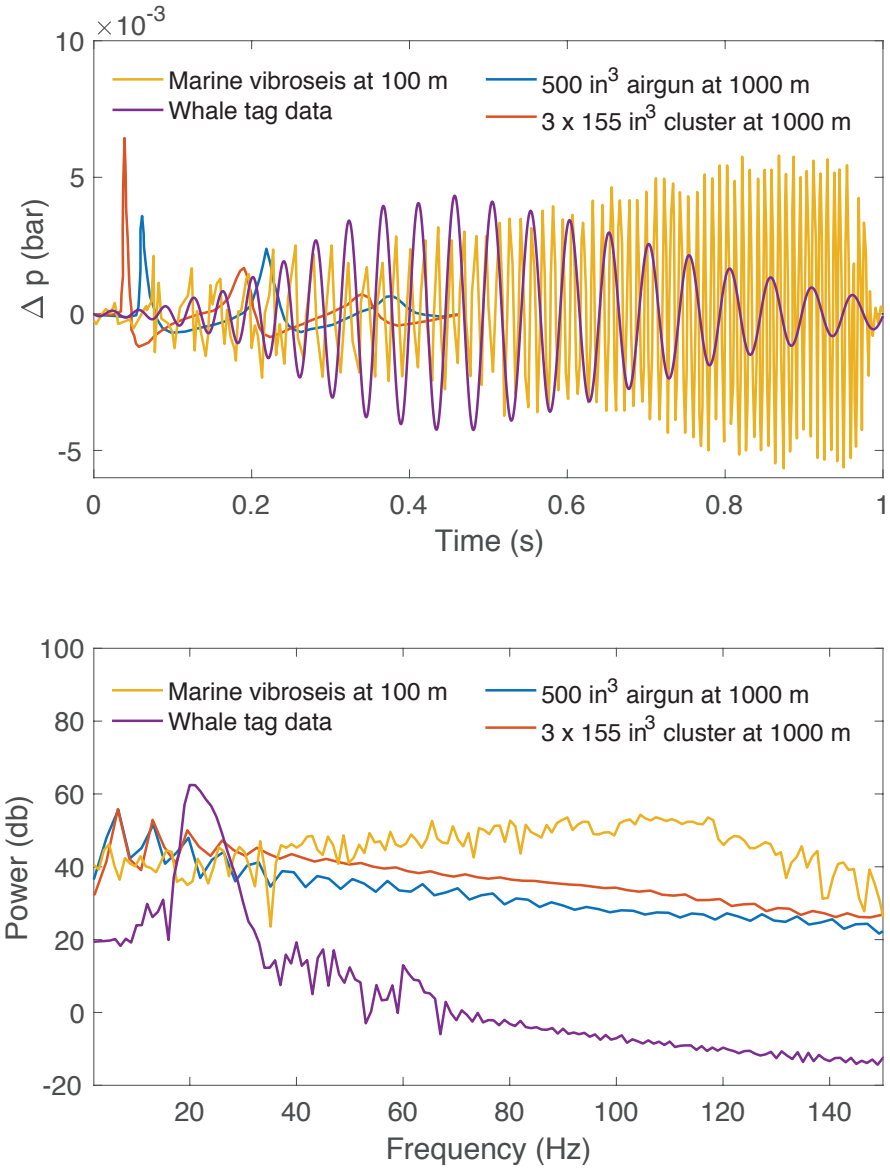
whale calls. The marine vibroseis sweeps from low frequencies to high frequencies with energy peaking around 100 Hz. Therefore, marine vibroseis may be less disruptive to whale communication than seismic airguns as the animals can likely distinguish between the signals based on the frequency content. However, in order for marine vibroseis to be commercially viable there is a need for improved low frequency content, which would be more disruptive to whale communication.

We note that the seismic sources are significantly more broadband than the whale call. Much of the high-frequency energy (above 150 Hz) generated by marine seismic sources is useless for seismic imaging as it is attenuated before it reaches the target or is scattered by the heterogeneous overburden (Watson et al., 2016). Hence, it is likely that marine seismic source designers will look at the lack of high frequencies in the whale calls with envy.

WHALE SOUND GENERATION

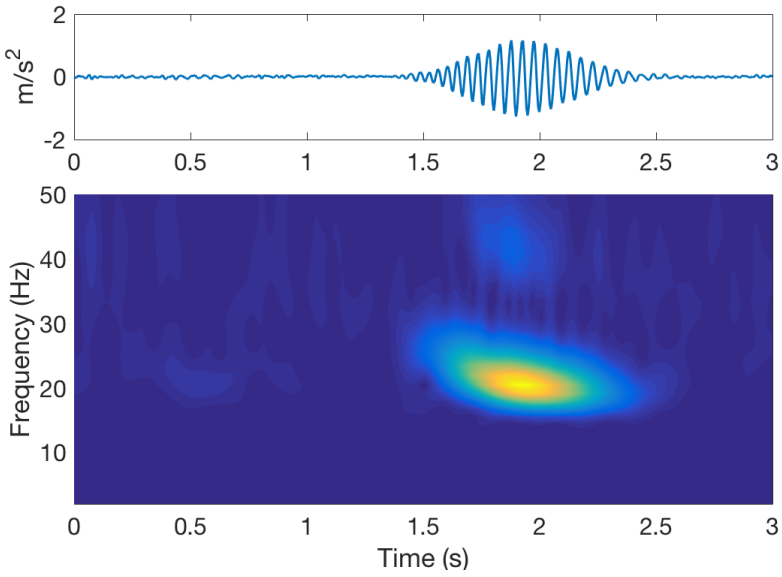
The mechanism of low-frequency sound generation in whales remains an open question. There has been some previous work investigating the possibility of cavities inside the whale acting as a resonator. Barham (1973) demonstrated that the free oscillation of the lungs could explain the 20 Hz peak frequency observed in fin whale calls, whereas Aroyan et al. (2000) argued that a Helmholtz resonator description with air moving between multiple cavities is a more appropriate description. We build upon the idea of Barham (1973) and develop a numerical model of the whale-lung system where the oscillation is forced by the contraction of muscles surrounding the lung.

The modeling approach is similar to that used to describe the oscillations of bubbles



(a)

Figure 4: Comparison of whale call and marine seismic sources in the time (top) and frequency (bottom) domain. [**CR**] leighton1/. compareWhalesSeismicSources



(a)

Figure 5: x-component of the accelerometer data from the DTAG on whale bp12-294a (Goldbogen et al., 2014) [**CR**] leighton1/. accelerometerData

produced by seismic airguns (e.g., Ziolkowski, 1970; Brennen, 1995; de Graaf et al., 2014; Watson et al., 2016) where conservation of momentum of the fluid is solved and evaluated on the bubble wall. The whale-lung system is modeled as a lung surrounded by a membrane representing the net effect of the whale's body on the dynamics of the lung. We approximate the lungs as a sphere with a volume equal to the total lung capacity. In reality a whale has two roughly elliptical lungs. However, the spherical lung approximation is reasonable because we are only interested in low frequencies ($\sim$ 20 Hz). The radius of the lungs ($\sim$ 1 m) is far smaller than the wavelength at these frequencies of interest ($\sim$ 75 m) and hence the precise shape of the lungs is not important (Aroyan et al., 2000).

There are two major differences between the model presented here and the models used in seismic airgun studies. Firstly, the airgun bubble models describe the oscillation of an air bubble surrounded by water. Here, we are interested in a similar problem but with the air bubble surrounded by the membrane of the lungs as well as the rest of the whale's body in addition to the water. This effect is parameterized into a surface tension term. Secondly, we consider the motion of the lungs to be a driven oscillation that is forced by the contraction of muscles around the lungs. This adds a forcing term to the governing equation. The resulting governing equation is a combination of the Rayleigh-Plesset equation (Rayleigh, 1917; Brennen, 1995) with the modified Herring correction for the compressibility of the water (Herring, 1941; Cole, 1948; Vokurka, 1986) and the addition of the driving force:

$$R\ddot{R} + \frac{3}{2}\dot{R}^2 = \frac{p - p_\infty}{\rho_\infty} + \frac{R\dot{p}}{\rho_\infty c_\infty} - \frac{2S}{\rho_\infty R} - \frac{\sigma}{\rho_\infty}, \tag{1}$$

where R, $\dot{R} = dR/dt$, and $\ddot{R} = d^2R/dt^2$ are the effective radius, velocity, and acceleration, respectively; p is the spatially uniform pressure inside the lung; and ρ_∞, p_∞ and c_∞ are

the constant ambient density, pressure, and speed of sound respectively. p_∞ is calculated assuming hydrostatic pressure and using the depth data provided by the tag (18.6 m for the data displayed here). S is the surface tension constant and σ is the potentially time-varying compressive stress from the muscles surrounding the lungs. The pressure is related to the effective radius by the ideal gas law,

$$p = p_0 \left(\frac{R_0}{R} \right)^{3k}, \tag{2}$$

where p_0 and R_0 are the equilibrium pressure and radius, respectively, and k is the polytropic index (equal to 1.4 for an adiabatic process).

The total lung capacity at the surface, V_a, is taken as 1540 L, which is the average of the measurements performed by Scholander (1940) and Leith and Lowe (1972) on excised lungs (Piscitelli et al., 2013). The equilibrium pressure inside the lung at depth, p_0, is equal to the hydrostatic pressure plus the contribution from surface tension, $p_\infty + 2S/R_0$. The equilibrium radius is determined by solving,

$$\frac{nQT_0}{\frac{4}{3}\pi R_0^3} = p_\infty + \frac{2S}{R_0}, \tag{3}$$

where T_0 is the body temperature of the whale, n is the number of moles of air inside the lungs and Q is the universal gas constant. The number of moles is determined from evaluating the ideal gas law at the surface, $p_a V_a = nQT_a$, where subscript a indicates atmospheric values.

Impulsive Excitation

A possible mechanism for sound generation in whales is an impulsive excitation of the lungs followed by decay with oscillations at the resonant frequency of the lungs, which is approximately 20 Hz. The measured displacements are extremely small (sub millimeter) compared to the radius of the lungs (several tens of centimeters). Therefore, the governing equation (equation 1) can be linearized around R_0 and p_0 to give,

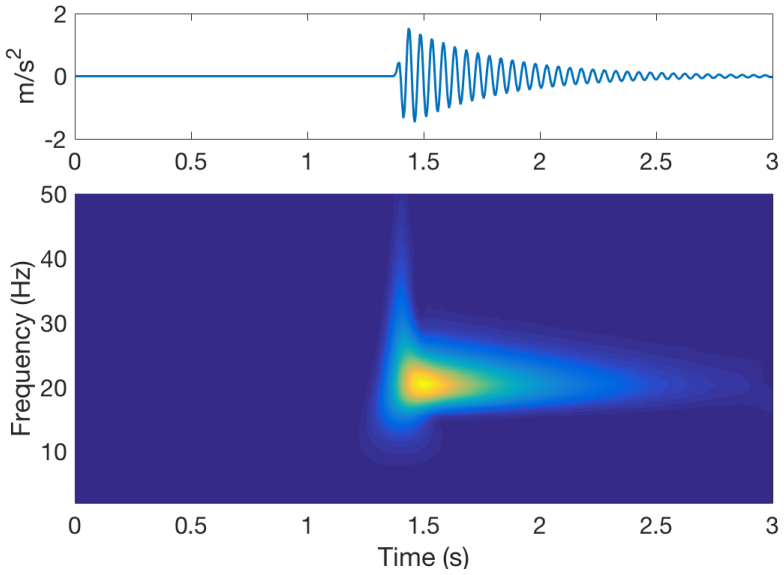
$$R_0 \Delta\ddot{R} + \frac{3kp_0}{\rho_\infty c_\infty}\Delta\dot{R} + \left(\frac{3kp_0}{\rho_\infty R_0} - \frac{2S}{\rho_\infty R_0^2} \right)\Delta R = -\frac{\sigma}{\rho_\infty}. \tag{4}$$

By comparison with the canonical equation for a damped harmonic oscillator, the characteristic decay time of the oscillations, τ, is,

$$\tau = \frac{2R_0}{3k}\frac{\rho_\infty c_\infty}{p_0}. \tag{5}$$

For the model parameters of interest, $\tau \approx 0.4$ s which is of the same order of magnitude as the duration of the observations, suggesting that an impulsive force and subsequent ringing could be the mechanism for fin whale sound generation.

The impulsive excitation model does not explain the emergent onset or the changing frequency throughout the call. Nonetheless, we note that the similarity between decay



(a)

Figure 6: Simulated acceleration of whale lung for impulsive excitation. The impulsive excitation is approximated as a gaussian with a mean of 1.4 s and a standard deviation of 0.009 s. [**CR**] leighton1/. impulsiveForcing

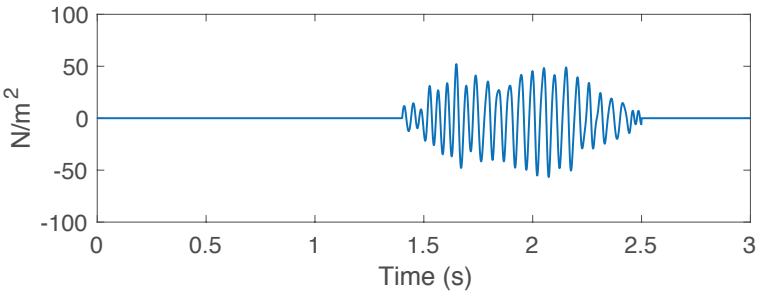
time and resonant frequency of a harmonic oscillator and the observations is an intriguing coincidence that warrants further investigation. We speculate that the emergent onset could be due to air flowing between cavities with a vibrating membrane in between while the resonant frequency of the lungs depends on the radius and hence would decrease expansion of the lungs, such as from mass flow into the lungs or muscular expansion. Further work is needed to investigate the validity of these mechanisms.

Driven Oscillation

An alternative sound generation mechanism is driven oscillation of the lungs through a time-varying compressive stress. The compressive stress is calculated from the data by taking the Fourier transform of equation 4 and solving for σ,

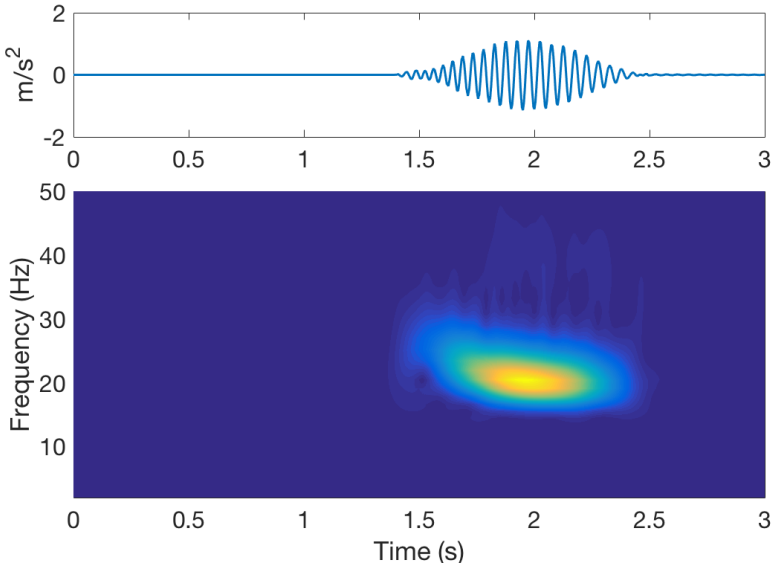
$$\left(-\omega^2 \rho_\infty R_0 - \frac{2S}{R_0^2} + \frac{3kp_0}{R_0} + i\omega \frac{3kp_0}{c_\infty} \right) \Delta \hat{R} = \frac{\hat{\sigma}}{\rho_\infty}, \tag{6}$$

where ω is the angular frequency. This model is able to match the observations exactly as there are an infinite number of free parameters (see Figure 8). Further physiological work is needed to determine if there are muscles in the whale capable of driving the oscillations of the lungs at ∼20 Hz.



(a)

Figure 7: ompressive stress as a function of time inverted from the data for $S = 8 \times 10^4$ N/m. It is assumed that whales have evolved to generate sound efficiently and the value of the surface tension parameter is chosen to minimize the amplitude of the compressive stress required to match the observations. Future work should compare this estimation to the mechanical properties of whale tissues. [**CR**] leighton1/. drivingForce



(a)

Figure 8: Simulated acceleration of whale lungs for the time-varying compressive stress shown in Figure 7 [**CR**] leighton1/. drivenOscillation

CONCLUSION

The pressure perturbation generated by seismic airguns and marine vibroseis are compared to fin whale calls. The seismic signals are several orders of magnitude louder than the whale calls. We present high resolution hydrophone and accelerometer data from tags attached to fin whales in the hope that those in the seismic industry may be able to utilize these data to better quantify the impact of seismic surveys on whales.

We speculate on possible mechanisms for generation of the whale calls. Further work needs to be done to link the mathematical model to whale biology. However, studies on smaller marine animals and on deceased whales have limited applicability to the physiology of live whales while performing measurements on live whales is extremely challenging. Therefore, mathematical models, such as the work presented here, can provide important insight and help to guide future physiological studies.

REFERENCES

Aroyan, J. L., M. A. McDonald, S. C. Webb, J. A. Hildebrand, D. Clark, J. Laitman, and J. S. Reidenberg, 2000, Acoustic models of sound production and propagation, *in* Au, W. W. L., A. N. Popper, and R. R. Fay, eds., Hearing by whales and dolphins, chapter 10, Springer.

Barham, E. G., 1973, Whales' respiratory volume as a possible resonant receiver for 20 Hz signals: Nature.

Brennen, C. E., 1995, Cavitation and Bubble Dynamics: Oxford University Press.

Cole, R. H., 1948, Underwater explosions: Princeton University Press.

Compton, R., L. Goodwin, R. Handy, and V. Abbott, 2008, A critical examination of worldwide guidelines for minimising the disturbance to marine mammals during seismic surveys: Marine Policy, **32**, 255–262.

de Graaf, K. L., I. Penesis, and P. A. Brandner, 2014, Modelling of seismic airgun bubble dynamics and pressure field using the Gilmore equation with additional damping factors: Ocean Engineering, **76**, 32–39.

Dragoset, B., 2000, Introduction to air guns and air-gun arrays: The Leading Edge, **19**, 892.

Goldbogen, J. A., A. K. Stimpert, S. L. Deruiter, J. Calambokidis, A. S. Friedlaender, G. S. Schorr, D. J. Moretti, P. L. Tyack, and B. L. Southall, 2014, Using accelerometers to determine the calling behavior of tagged baleen whales.: The Journal of Experimental Biology, **217**, 2449–2455.

Gordon, J., D. Gillespie, J. Potter, A. Frantzis, M. P. Simmonds, R. Swift, and D. Thompson, 2003, A Review of the Effects of Seismic Survey on Marine Mammals: Marine Technology Society Journal, **37**, 16–34.

Herring, C., 1941, Theory of the pulsations of the gas bubble produced by an underwater explosions: Technical report, Office of Scientific Research and Development.

Hildebrand, J. A., 2005, Impacts of anthropogenic sound, *in* Reynolds, J. E. e. a., ed., Marine mammal research: conservation beyond crisis, 101–124, The John Hopkins University Press.

IAGC, 1995, Marine seismic sources: Technical report, International Association of Geophysical Contractors.

Johnson, M., N. A. De Soto, and P. T. Madsen, 2009, Studying the behaviour and sensory

ecology of marine mammals using acoustic recording tags: A review: Marine Ecology Progress Series, **395**, 55–73.

Johnson, M. P. and P. L. Tyack, 2003, A digital acoustic recording tag for measuring the response of wild marine mammals to sound: IEEE Journal of Oceanic Engineering, **28**, 3–12.

Leith, D. E. and R. Lowe, 1972, Mechanics of baleen whale lungs: Fed Proc, 335.

Moore, M. J., T. Hammar, J. Arruda, S. Cramer, S. Dennison, E. Montie, and A. Fahlman, 2011, Hyperbaric computed tomographic measurement of lung compression in seals and dolphins.: The Journal of experimental biology, **214**, 2390–2397.

Ogden, L. E., 2014, Quieting marine seismic surveys: BioScience, **64**, 752.

Piscitelli, M. A., S. A. Raverty, M. A. Lillie, and R. E. Shadwick, 2013, A review of cetacean lung morphology and mechanics: Journal of Morphology, **274**, 1425–1440.

Rayleigh, L., 1917, On the pressure developed in a liquid during the collapse of a spherical cavity: Philosophical Magazine Series 6, **34**, 94–98.

Richardson, W. J., C. R. Greene, C. I. Malme, and D. H. Thomson, 1995, Marine mammals and noise: Academic Press.

Scholander, P. F., 1940, Experimental investigations on the respiratory function in diving mammals and birds: Hvalrådets Skrifter, **22**, 235–250.

Stimpert, A. K., S. L. Deruiter, E. A. Falcone, J. Joseph, A. B. Douglas, D. J. Moretti, A. S. Friedlaender, J. Calambokidis, G. Gailey, P. L. Tyack, and J. A. Goldbogen, 2015, Sound production and associated behavior of tagged fin whales (Balaenoptera physalus) in the Southern California Bight: Animal Biotelemetry, **3**, 1–12.

Vokurka, K., 1986, Comparison of Rayleigh's , Herring's , and Gilmore's Models of Gas Bubbles: Acta Acustica united with Acustica, **59**, 214–219.

Watkins, W. A., 1981, Activities and underwater sounds of fin whales: Scientific Report of Whale Research Institure, 83–117.

Watson, L. M., E. M. Dunham, and S. Ronen, 2016, Numerical modeling of seismic airguns and low-pressure sources: SEG Annual Meeting, Dallas, Texas, 219–224.

Weilgart, L., 2013, A Review of the Impacts of Seismic Airgun Surveys on Marine Life: Presented at the Convention on Biological Diversity Expert Workshop on Underwater Noise and its Impacts on Marine and Coastal Biodiversity.

Weir, C. R. and S. J. Dolman, 2007, Comparative Review of the Regional Marine Mammal Mitigation Guidelines Implemented During Industrial Seismic Surveys, and Guidance Towards a Worldwide Standard: Journal of International Wildlife Law & Policy, **10**, 1–27.

Ziolkowski, A., 1970, A Method for Calculating the Output Pressure Waveform from an Air Gun: Geophysical Journal International, **21**, 137–161.

Elastic Full Waveform Inversion of the Moere Vest data

Gustavo Alves

ABSTRACT

The extension of full waveform inversion (FWI) into elastic models can potentially address the limitations of acoustic FWI due to amplitude versus offset effects. I present results of applying an elastic FWI workflow to ocean-bottom nodes data. I show the steps used to pre-process the data and create an initial elastic model for inversion. Then, I apply these inputs into the inversion workflow and show how elastic inversion without pre-conditioning can be dominated by artifacts. Finally, I apply weighting operators in both residual space and image space and present results for the inverted model with data in the range of 2 to 5 Hz.

INTRODUCTION

Full Waveform Inversion (FWI) has become one of the most prolific research topics in seismic imaging in recent years. Although the fundamentals of FWI were proposed 30 years ago (Tarantola, 1987), computational demands limited its generalized adoption.

Elastic Full Waveform Inversion (EFWI) pushes even further the computational demands of the method. Even when considering 2D problems with the velocity-stress formulation, the number of wave fields to be computed increases fivefold compared to the acoustic case. Also, the number of model parameters in EFWI increases threefold when compared to acoustic FWI. Finally, numerical stability and dispersion in elastic modeling require finer spatial and time sampling, due to slower propagating shear waves.

In this work, I apply the methodology developed in Alves (2015) to 2D field data in an EFWI workflow. I describe the pre-processing steps applied to the data and detail the objective function, search direction and step-length criteria. The inversion workflow uses the generic non-linear solver library developed by Biondi and Barnier (2017).

The data provided to SEP by Seabed Geosolutions were recorded in the Moere Vest field, offshore Norway. It was acquired using four component ocean-bottom nodes (4C OBNs) in a 2D array. In this work, I use a subset of this data composed of 141 nodes spaced 250 meters apart, totaling a 35 km receiver line. There are 1145 shots in this data set, with 100 meters between each shot, totaling a 114.5 km shot line. The sail line overlaps the receiver line, which gives a maximum offset of 75 km. Figure 1 shows the bathymetry of the survey area and the locations of the receivers.

I perform EFWI with field data sorted in common-receiver gathers. To generate synthetic data in this domain, I apply reciprocity to find equivalent sources at node positions. Alves and Biondi (2017) describes how to obtain reciprocal data in the elastic case and its advantages.

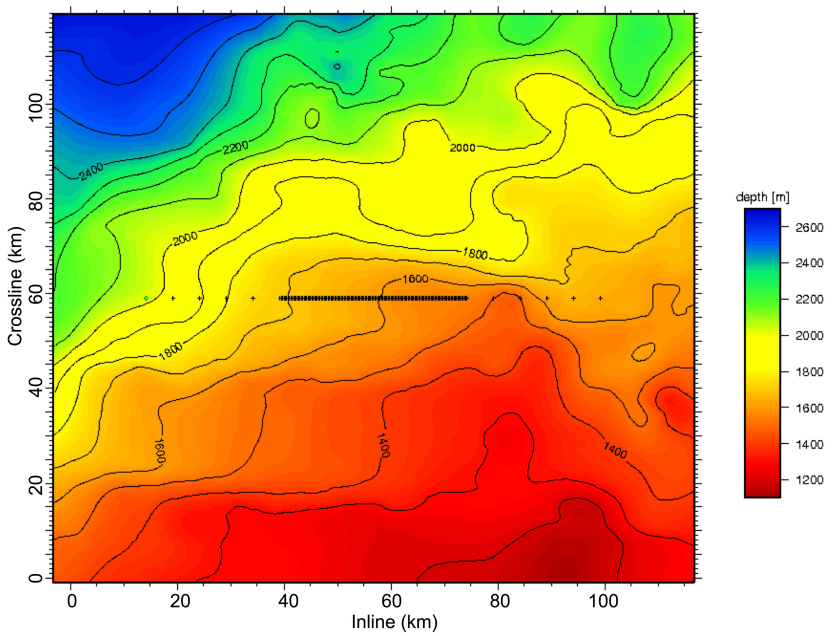


Figure 1: Bathimetry of the survey area and receiver positions. Receiver depths vary between 1942 m (West) to 1640 m (East). [**NR**] alves2/. bathimetry1

DATA PRE-PROCESSING

Before the data can be used for EFWI, it is important to pre-process it. By pre-processing, I mean the application of operators that remove features present in the field data but which are not modeled during the inversion. Since the EFWI process minimizes the L_2 norm of the data residual, it is paramount that field and synthetic data are as similar as possible. However, my goal is not to remove information related to the dynamics and kinematics of elastic events. Those should be unaffected by the pre-processing steps, so that they can be recovered by the EFWI.

Schematically, I can describe the pre-processing steps as the following chain of operators,

$$\mathbf{d_t} = \mathbf{F_{bub}F_{lc}S_{rec}Rd_{raw}}, \tag{1}$$

where $\mathbf{d_{raw}}$ is the original data, $\mathbf{R}$ is the rotation of horizontal geophones, $\mathbf{S_{rec}}$ is the receiver de-signature, $\mathbf{F_{lc}}$ is a low cut filter, $\mathbf{F_{bub}}$ is the bubble removal and $\mathbf{d_t}$ is the data prepared for EFWI.

Data rotation

I start the pre-processing by rotating the horizontal components of the geophones. My goal is to separate data into radial and transverse. Since the shot line and receiver lines overlap, this is the same as having one horizontal component aligned with the acquisition and the

other component orthogonal to it. To find the radial direction for each receiver, I apply an operator that rotates the horizontal components by 1 degree increments until I find the one that maximizes the energy of the first break in the radial component. Figure 2 shows data for horizontal geophones before and after rotation. I apply hyperbolic moveout (HMO) to the data and mark in red the areas above a fixed amplitude, to highlight the energy change. The rotation scheme applied here follows the methodology presented in Levin and Chang (2015).

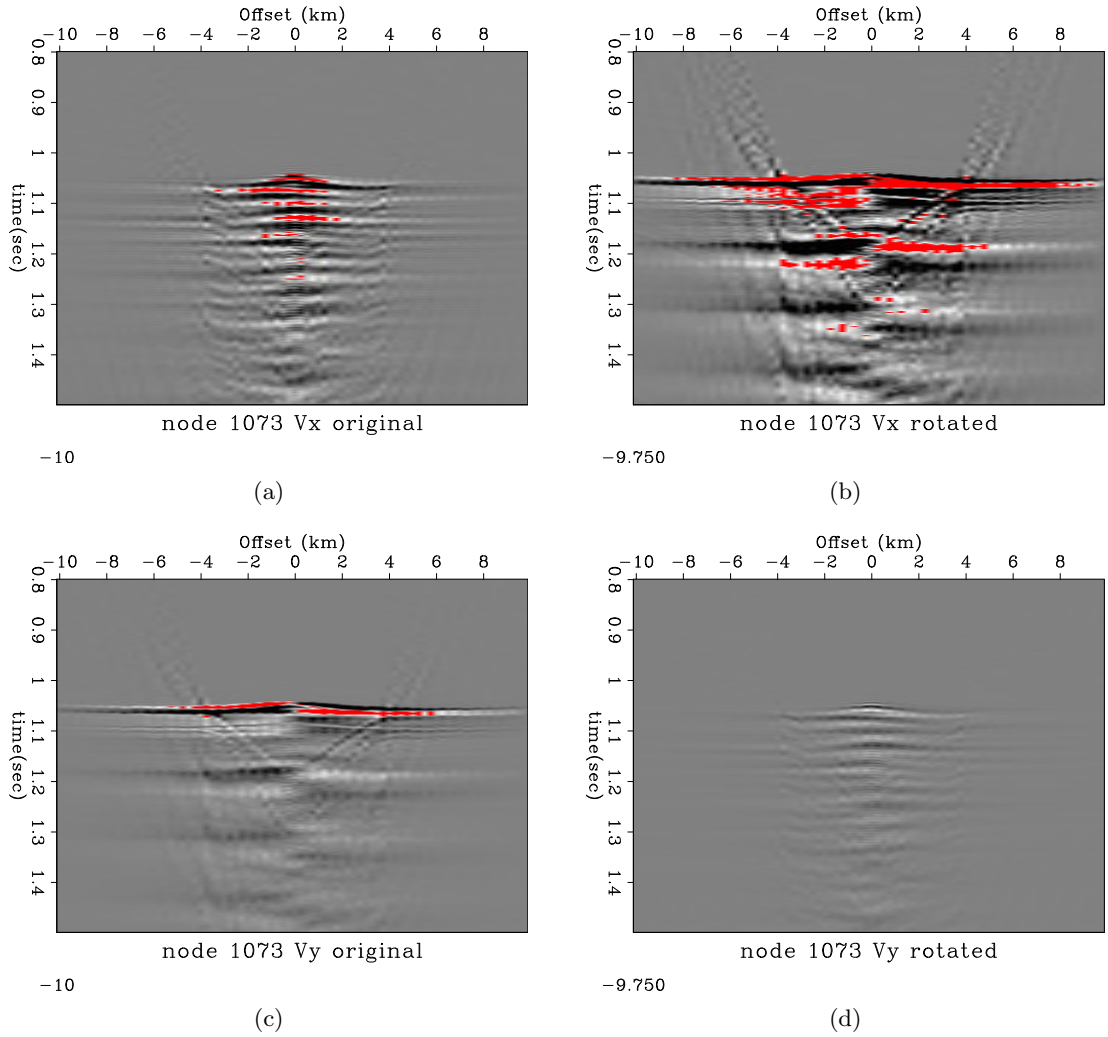


Figure 2: Common receiver gathers of the two horizontal geophones for a single node, before (a,c) and after (b,d) rotation. The red highlights indicate the amplitude above a clipping value. [**NR**] alves2/. rotation-original-vx,rotation-new-vx,rotation-original-vy,rotation-new-vy

Receiver de-signature

Data recorded in the field is subject to the response of receivers used in the acquisition. This response includes, but is not limited to, coupling between receiver and elastic medium, fre-

quency bandwidth and signature. I assume perfect coupling to the medium and combine all other effects as a convolution of wave field with receiver signature. I apply a de-convolution filter to the data by dividing it by the impulse response in the Fourier domain. A small stability term is added to avoid division by zero. Figure 3 shows the impulse responses of hydrophone and geophone receivers, as provided by Seabed.

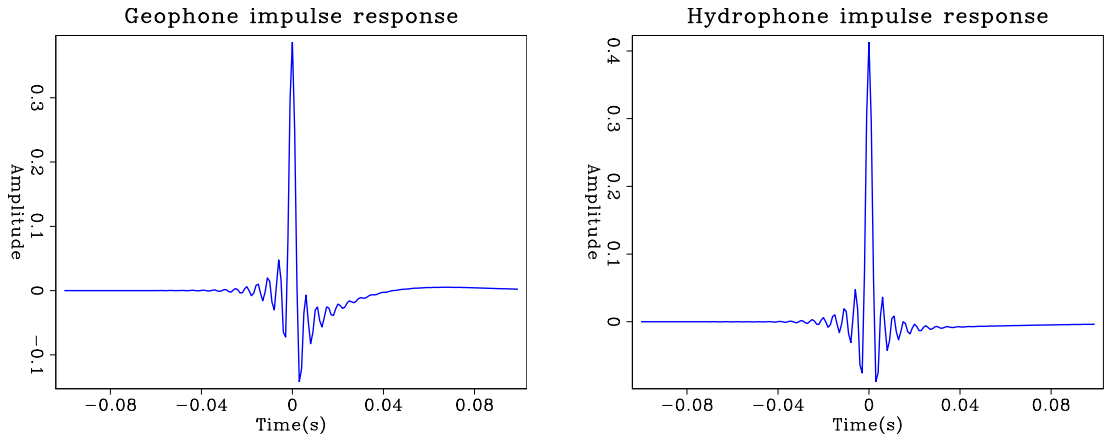


Figure 3: Receiver impulse responses for the (a) geophones and (b) hydrophones used in the survey. [**NR**] alves2/. rec-geo,rec-hyd

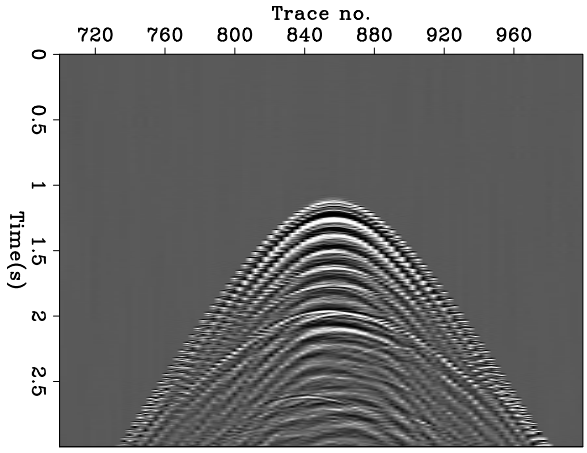
Bandpassing

I apply a low-cut filter at 2 Hz, aimed at removing swell noise. Figures 4(a) and 4(b) show the application of this frequency filter.

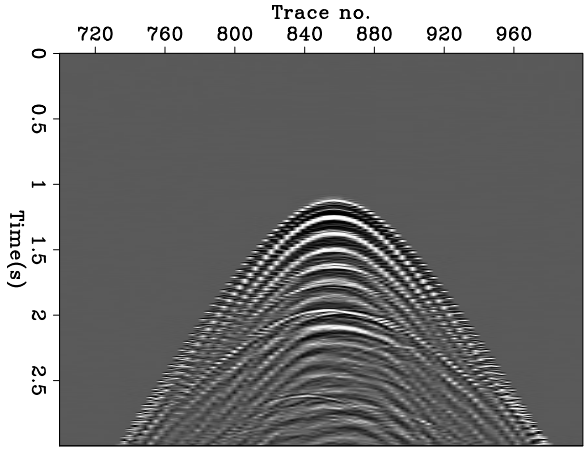
De-bubble

The last step in the field data pre-processing is the removal of the air gun bubbles. The air guns used in marine acquisition generate pressure waves by injecting a fast expanding air bubble in the water. This bubble oscillates as it surfaces and disperses, producing a periodic low frequency ringing in the data. While in theory this event could be introduced in the wavelet of the synthetic modeling, it would represent a novelty in traditional imaging with challenges of its own. Since the goal of this work is not to study this topic, I choose to remove the bubble effect.

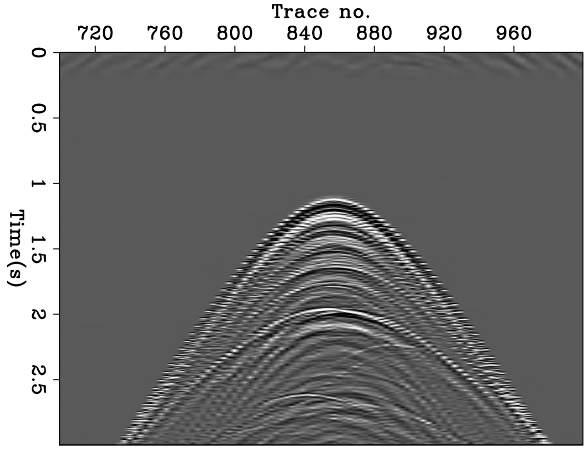
I remove the bubbles by applying an autoregression filter, also called a prediction error filter (PEF). This is a suitable method for removing events such as multiples or, in my case, bubble effects, where the time and length of the events is predictable. For a detailed description of PEFs, I refer to Claerbout and Fomel (2008). Figures 4(b) and 4(c) shows data before and after de-bubble, respectively.



(a)



(b)



(c)

Figure 4: Vertical geophone data (a) before and (b) after band-passing and (c) after bubble removal by prediction error filtering. [**ER**] alves2/. swell,noswell,dbub

DATA MODELING

While the pre-processing described in the previous section deals with the many undesired effects present in field data, the steps in this section are aimed at improving the generation of synthetic data.

Source wavelet estimation

The objective function is essentially a measure of the difference between the observed data in the field and the data synthetically generated by the modeling. Therefore, it is very important to find a good estimate of the source wavelet, to minimize the residual between equivalent events in each data.

A very robust method to approximate the source wavelet is by performing a linear inversion of the direct arrivals in the field data. In the case of marine data, the direct arrivals travel only through an approximately constant medium (water), so the background model is known and the only parameter to be inverted for is the source wavelet itself.

Figure 5 shows the data residual of the direct arrival for one gather before and after wavelet estimation. The initial estimated source signature is zero, so the initial residuals are just the observed data, while the final residuals are obtained after 100 iterations. Figure 6 displays the objective function for this inversion and 7 shows the estimated wavelet.

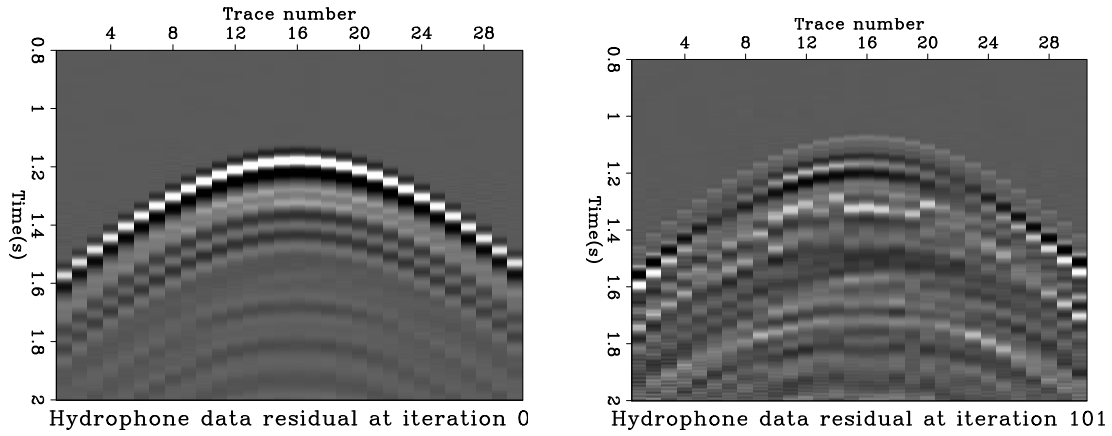


Figure 5: Data residual (a) before and (b) after wavelet estimation. [**ER**] alves2/. reswave0,reswave100

Data reciprocity

Both field and synthetic data used in the inversion are sorted in common-receiver gathers. To generate common-receiver gather synthetic data, I apply the elastic reciprocity theorem (Aki and Richards, 1980) to the hydrophone and geophone components. While the pressure (hydrophone) data is trivially obtained from acoustic reciprocity, vector data (geophones)

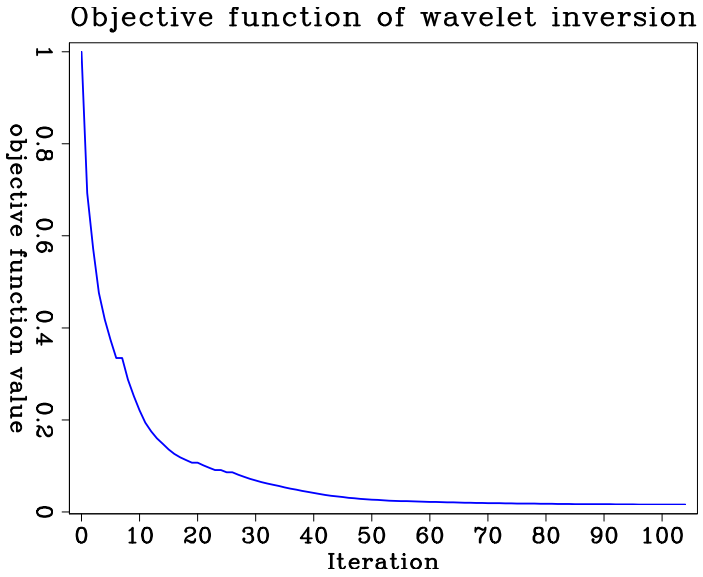


Figure 6: Objective function for the linear inversion of the source wavelet. [**ER**] alves2/. objwave

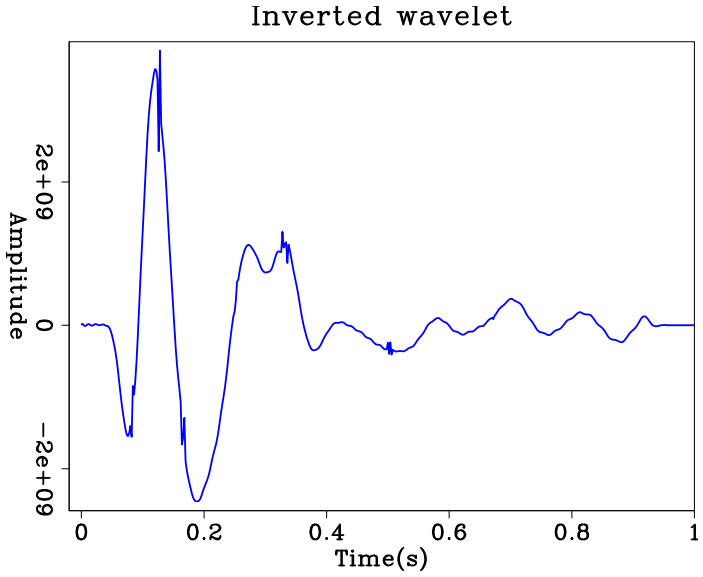


Figure 7: Final estimated source inverted from the direct arrivals in the field data. [**ER**] alves2/. srcwave

need to be understood more carefully. I refer to Alves and Biondi (2017) for an in-depth analysis of this problem.

INITIAL MODEL

The survey area is located in the Moere Basin, in the Norwegian-Greenland Sea. The Moere Basin is composed mainly of Lower Cretaceous to Late Oligocene shales, marine sandstone sequences during the Upper Cretaceous and Paleocene, and Early Eocene lavas, which are characteristic throughout most of the Voering and Moere Basins (Brekke et al., 1999).

V_p/V_s ratio

I obtain initial estimates for V_p and V_s from Linear Moveout (LMO) analyses of ocean bottom and top of basalt refractions. Estimated velocties are $V_p = 2300m/s$ and $V_s = 1340m/s$ for the overburden, which results in a V_p/V_s ratio of 1.72. A detailed explanation of how these velocities were estimated and how P-to-S conversions were used to botain V_p/V_s can be found in Moronfoyer et al. (2016).

Reverse time migration

I use the depths of the OBNs and velocities estimated previously to create a simple two-layer model (see figure 8). I apply Elastic Reverse Time Migration (ERTM) to the pre-processed data with a smoothed version of this model, as can be seen on figure 9. The ERTM code I use is essentially the solution of the adjoint Born operator for the elastic imaging condition, as I have described previously in Alves (2015).

I use the ERTM image to manually pick the top and bottom of the basalt layer. While I expect that the basalt bottom will be pulled up in this image due to the lower velocities, it is enough to create an improved properties model for the EFWI workflow. Figure 10 shows the improved initial model.

ELASTIC FULL WAVEFORM INVERSION

There are many different forms of implementing a FWI workflow. Every step of the iteration process includes many choices, such as, type of Hessian approximation, norm of the objective function, search criteria for both direction and step length, use of pre conditioners and regularizers, etc. An interesting review on some of these choices can be found in Hager and Zhang (2006).

The objective function that I minimize is given by,

$$\phi(\mathbf{m}) = \frac{1}{2} \left\| \mathbf{W}(\mathbf{m}) - \mathbf{d_t} \right\|_2^2 , \tag{2}$$

where $\mathbf{W}(\mathbf{m})$ is the solution of the non-linear operator for model $\mathbf{m}$ and $\mathbf{d_t}$ is the field data. The gradient at iteration n, $\mathbf{g_n}$, used for the model update is,

$$\mathbf{g_n} = \left(\frac{\partial W(m)}{\partial m} \bigg|_{m_n} \right)^T , \tag{3}$$

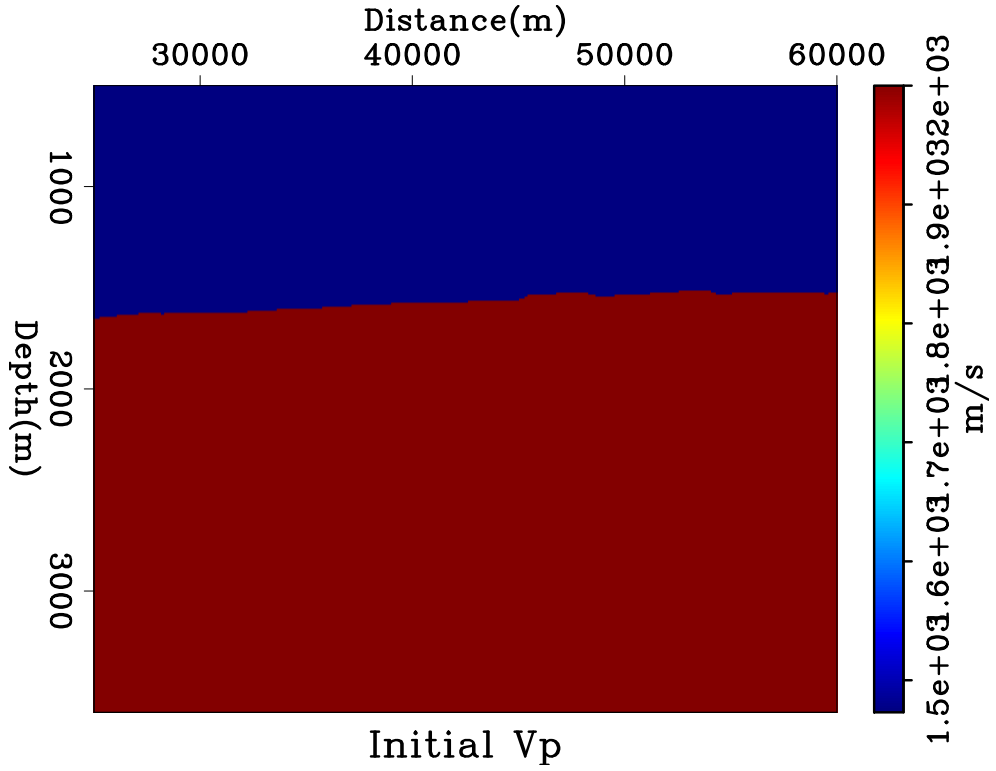


Figure 8: Initial V_p velocity calculated from linear moveout anlysis and used for reverse time migration. Components V_s and ρ are not displayed for brevity. These look similar to V_p, respecting the estimated V_p/V_s ratio and densities of $10^3 kg/m^3$ and $2.0 * 10^3 kg/m^3$ for water and overburden, respectively. [**ER**] alves2/. initvp

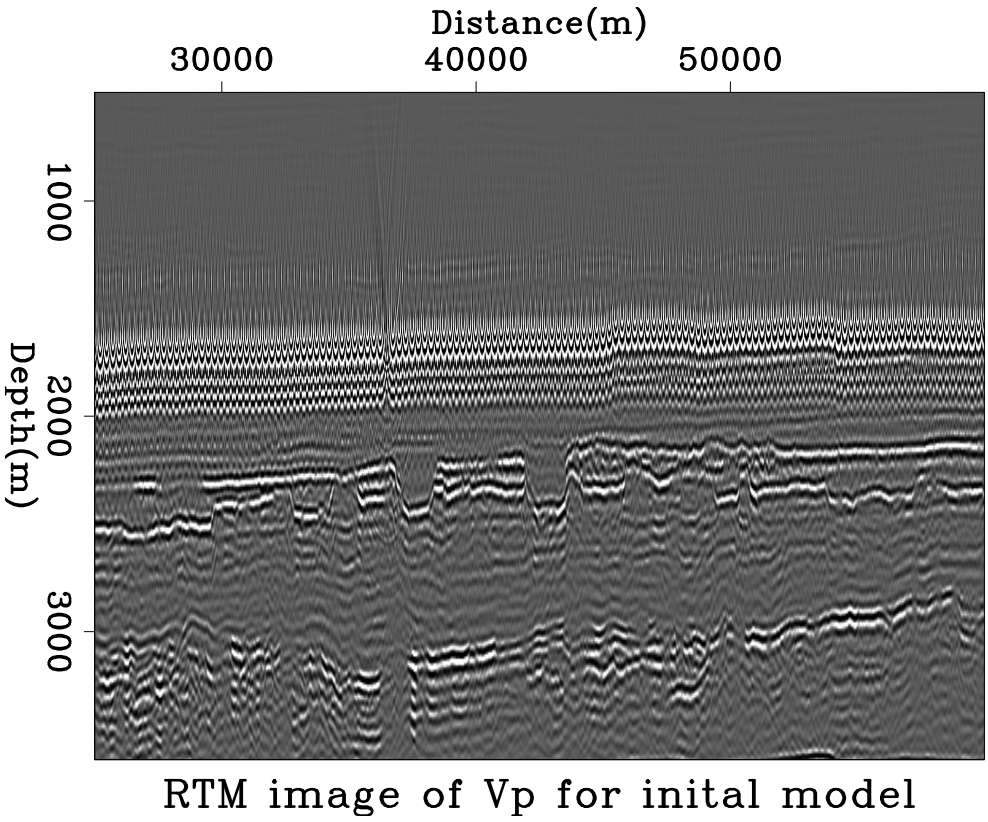


Figure 9: V_p component of ERTM image using the initial properties model. Notice the acquisition footprint due to receiver density. [**CR**] alves2/. rtm0

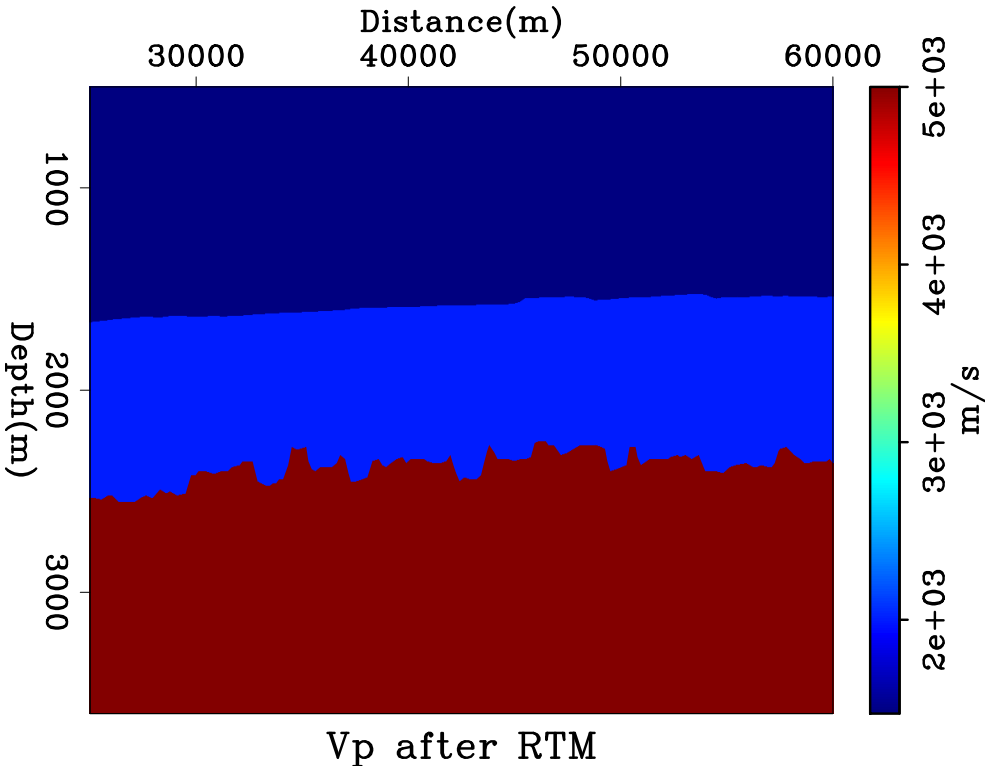


Figure 10: Improved initial velocity model with a basalt layer. [**NR**] alves2/. basaltvp

where the partial derivative with respect to model parameters is the Jacobian of the function $\mathbf{W}(\mathbf{m})$, evaluated at the current background model, $\mathbf{m_n}$. The search direction is obtained by steepest-descent on the first iteration, followed by conjugate-gradient at every other iteration (Aster et al., 2005). The step length is obtained by parabolic search.

I run 20 iterations of EFWI. Data and synthetics are band-passed between 2 and 5 Hz, in order to focus the initial inversion in the long wavenumber updates of the properties model and decrease the chances of cycle skipping. Figure 11 shows an example of a gradient obtained for this implementation.

It is clear from figure 11 that the gradient only contains information at the node positions. Even after several iterations, the gradient remained similar and did not guide the inversion towards updates in the subsurface. Therefore, I improve on the previous objective function by adding a weighting term to the data residual,

$$\phi(\mathbf{m}) = \frac{1}{2} \left\| \mathbf{M_{res}}(\mathbf{W}(\mathbf{m}) - \mathbf{d_t}) \right\|_2^2 , \tag{4}$$

and also a weighting term to the gradient,

$$\tilde{\mathbf{g}}_\mathbf{n} = \mathbf{M_{grad}} \mathbf{g_n}, \tag{5}$$

where $\mathbf{M_{res}}$ is a weighting function applied to the data residual that dampens the direct arrivals and applies a gain proportional to v_{rms} in time. $\mathbf{M_{grad}}$ is a weighting function to the gradient, that reduces the updates in the regions around the node positions.

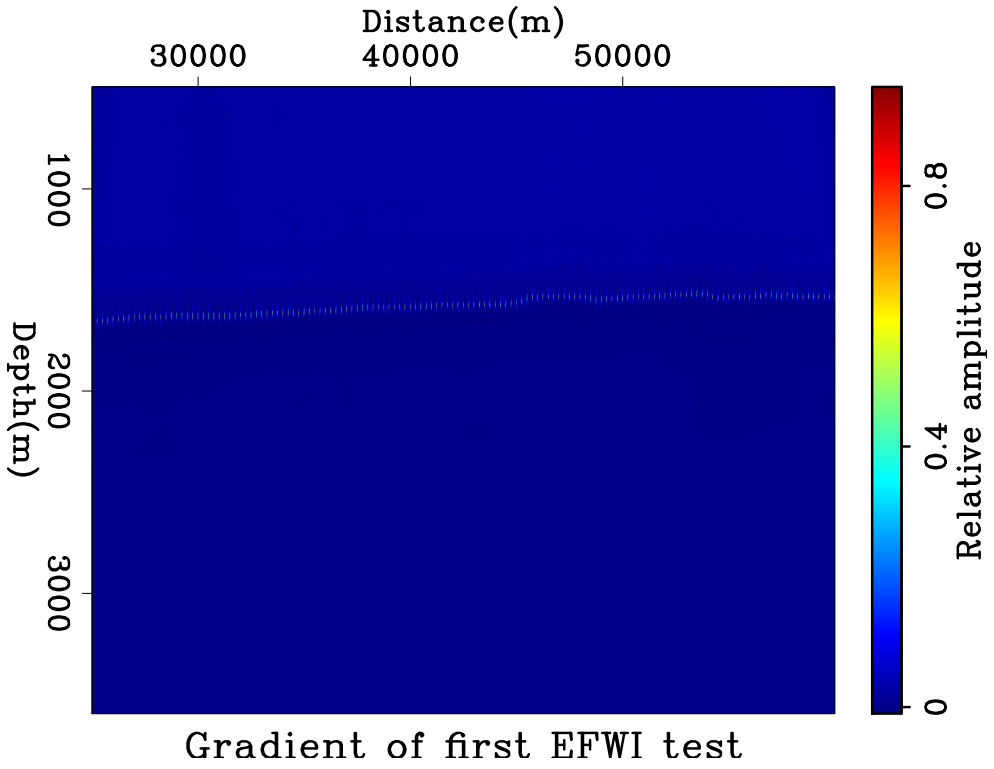


Figure 11: Example of gradient for first EFWI scheme without any data or gradient weighting. The gradient is dominated by updates at the nodes positions. [**CR**] alves2/. gefwi0

After 20 iterations using this modified method and the same frequency range of 2-5Hz, it is possible to notice improvements in the sub-surface part of the model. Figure 12 shows a comparison of the initial smoothed model and the improved inverted results. In this figure, a horizontal smoothing was applied to the final result to remove some of the footprint caused by the distance between nodes at the sea bottom.

CONCLUSION

Pre-processing was done successfully with a combination of different filters and tools. I applied a deconvolution filter for receiver de-signature, a PEF to remove the air gun bubble, and a linear inversion to estimate the source wavelet.

The use of linear move-out information gave an initial estimate for the V_p/V_s ratio. An initial RTM estimated the position for the top of the basalt layer with enough continuity to generate an improved model for inversion.

The use of a weighting operator was necessary to remove artifacts in the inversion. These artifacts were likely caused by discontinuities in the spatial first derivatives of the particle velocities at the nodes positions, although a more careful analysis of the problem would be required to confirm this hipotesis. It is important to recall that, because spatial reciprocity is being applied, the synthetic sources are positioned at the nodes coordinates.

The inverted properties show some improvements in the subsurface. However, the low

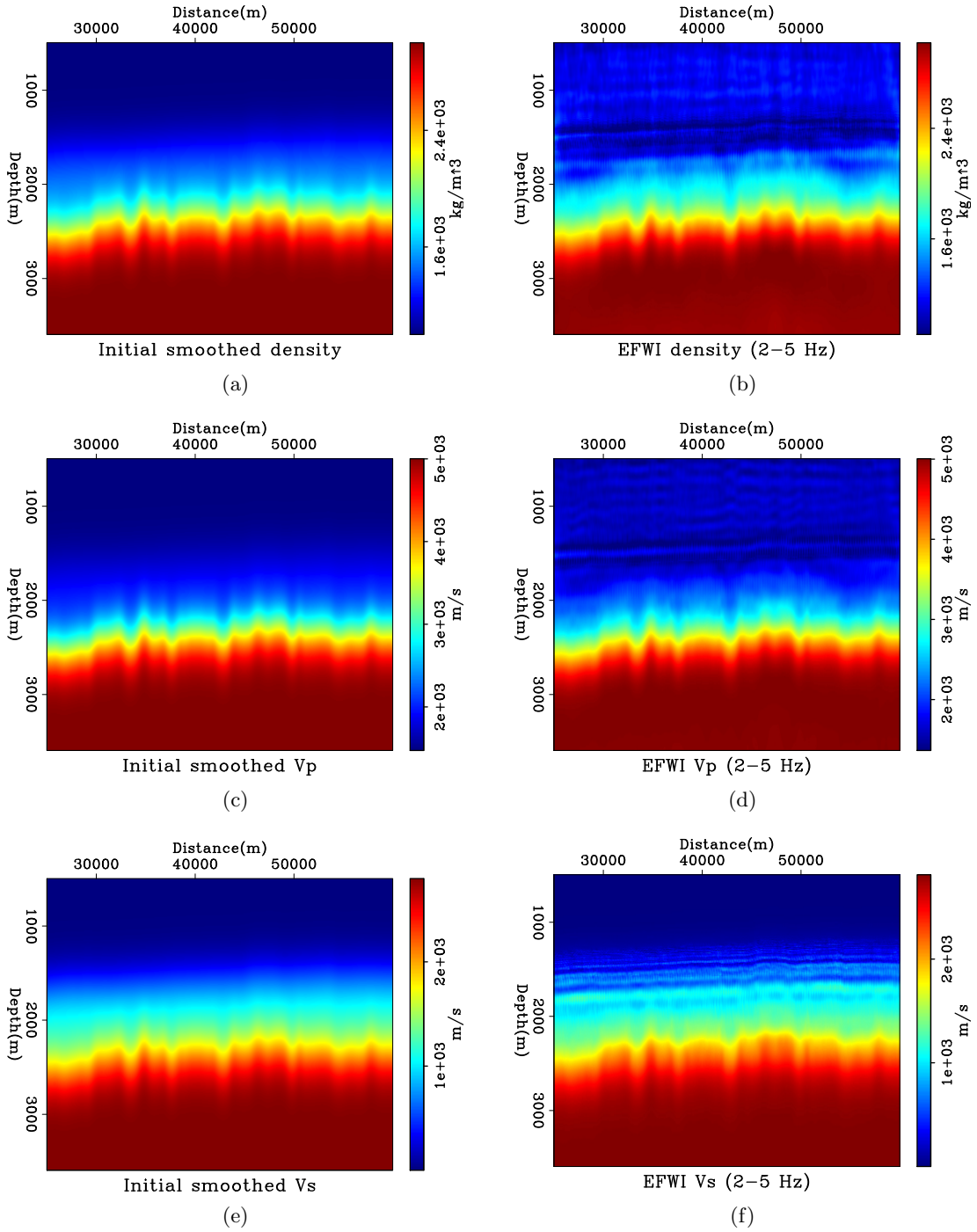


Figure 12: Comparison of (left) initial elastic properties and (right) after 2-5Hz EFWI. Inverted results have been smoothed horizontally to remove artifacts due to nodes geometry. [**CR**] alves2/. rho0,rho2-5,vp0,vp2-5,vs0,vs2-5

frequencies used still limit the updates to wavelengths in the order of hundreds of meters and therefore limit the resolution of the recovered layers. Further work in this problem will include higher frequencies in the inversion workflow, as well as better pre-conditioning of the gradients to address acquisition footprints and updates to the water column.

ACKNOWLEDGEMENTS

I would like to thank Seabed Geosolutions for the use of this data set. I also thank Petrobras for supporting my PhD and SEP sponsors for the ongoing support.

REFERENCES

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

Alves, G., 2015, Adjoint formulation for the elastic wave equation: Stanford Exploration Project Report, **158**, 133–150.

Alves, G. and E. Biondi, 2017, Reciprocity in elastic multi-component data: Stanford Exploration Project Report, **168**, –.

Aster, R. C., B. Borchers, and C. H. Thurber, 2005, Parameter estimation and inverse problems: Elsevier Academic Press.

Biondi, E. and G. Barnier, 2017, A flexible out-of-core solver for linear/non-linear problems: Stanford Exploration Project Report, **168**, –.

Brekke, H., S. Dahlgren, B. Nyland, and C. Magnus, 1999, The prospectivity of the Vøring and Møre basins on the norwegian sea continental margin: Geological Society, London, Petroleum Geology Conference series, 261–274.

Claerbout, J. F. and S. Fomel, 2008, Image estimation by example: Geophysical soundings image construction: multidimensional autoregression.

Hager, W. W. and H. Zhang, 2006, A survey of nonlinear conjugate gradient methods: Pacific journal of Optimization, **2**, 35–58.

Levin, S. A. and J. P. Chang, 2015, Stable reorientation for the forties dataset: Stanford Exploration Project Report, **160**, 189–194.

Moronfoyer, A. T., S. Ronen, S. L. Klemperer, and G. Alves, 2016, Upper-crustal structure of the Møre margin, offshore-norway: 20016 AGU Fall Meeting, San Francisco, –.

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

Preconditioned elastic full waveform inversion using approximated Hessian matrix

Ettore Biondi, Guillaume Barnier, and Biondo Biondi

ABSTRACT

We describe a simple way of estimating the elements of the Hessian matrix for elastic full waveform inversion (FWI) problems by applying it to spikes in the model space. We explain how to use this estimated matrix to precondition an elastic isotropic FWI problem and show results from a simple one layer model and a more complex subsurface elastic earth. We observe that already a main diagonal approximation is able to increase the inverse problem's convergence rate. This gradient preconditioning allows us to obtain a meaningful inversion result for the deeper part of the model in fewer iterations.

INTRODUCTION

Different preconditioning methods for FWI have been proposed in recent years. For example, Alkhalifah (2015b) describes a gradient preconditioning algorithm to avoid cycle skipping during FWI optimization. Based on the analysis performed by ten Kroode (2012) for Kirchhoff migration, Huang et al. (2016) shows how filtering FWI gradients in extended-image space can mimic the effect of a Newton optimization step. These two methods have been shown to be successful in the case of single-parameter inversions or when mostly the traveltimes are affected by the presence of multiple parameters (Alkhalifah, 2015a). However, it is missing a study of their advantage when also seismic wave amplitudes are influenced by different optimization parameters (e.g., elastic FWI).

Given the current interest in solving more complex and computationally intense multi-parameter FWI problems, such as fully orthorhombic elastic waveform inversion optimization (Albertin et al., 2016), it is critical to find methods to improve the convergence rate of any inversion algorithm employed. For this reason many authors have recognized the fundamental role of Hessian matrices as preconditioner for multi-parameter FWI problems (Forgues and Lambaré, 1997; Operto et al., 2013; Innanen, 2014). Tang and Lee (2015) show a simple way of estimating the Hessian matrix and demonstrate its value as preconditioner in a vertical transverse isotropic FWI example. In their implementation they take advantage of the sparse structure of the Gauss-Newton component of this matrix and compute its elements through matrix applications to spikes contained in the model space.
Korta et al. (2013) show how a block-diagonal Hessian approximation can improve convergence rate during FWI optimization when compressional and shear wave velocities are inverted simultaneously. Similarly, using a phase-encoding approach to estimate the matrix elements (Tang, 2008), Deuzeman and Plessix (2015) demonstrate how a similar Hessian approximation speed-ups the optimization algorithm on simple subsurface models.

In this study we simultaneously invert for elastic wave velocities and density following

the same approach proposed by Tang and Lee (2015) to estimate the Gauss-Newton Hessian elements. We show elastic FWI results on a complex subsurface elastic model and compare a main-diagonal approximation to un-preconditioned optimization.

THEORY

The 2D velocity-stress formulation of the elastic wave equation is given by the following set of relations (Virieux, 1986):

$$\rho(\vec{x})\frac{\partial v_x(\vec{x},t)}{\partial t} - \frac{\partial \sigma_{xx}(\vec{x},t)}{\partial x} - \frac{\partial \sigma_{xz}(\vec{x},t)}{\partial z}, = S_{vx}(\vec{x},t) \tag{1}$$

$$\rho(\vec{x})\frac{\partial v_z(\vec{x},t)}{\partial t} - \frac{\partial \sigma_{xz}(\vec{x},t)}{\partial x} - \frac{\partial \sigma_{zz}(\vec{x},t)}{\partial z} = S_{vz}(\vec{x},t), \tag{2}$$

$$\frac{\partial \sigma_{xx}(\vec{x},t)}{\partial t} - [\lambda+2\mu](\vec{x})\frac{\partial v_x(\vec{x},t)}{\partial x} - \lambda(\vec{x})\frac{\partial v_z(\vec{x},t)}{\partial z} = S_{xx}(\vec{x},t), \tag{3}$$

$$\frac{\partial \sigma_{zz}(\vec{x},t)}{\partial t} - \lambda(\vec{x})\frac{\partial v_x(\vec{x},t)}{\partial x} - [\lambda+2\mu](\vec{x})\frac{\partial v_z(\vec{x},t)}{\partial z} = S_{zz}(\vec{x},t), \tag{4}$$

$$\frac{\partial \sigma_{xz}(\vec{x},t)}{\partial t} - \mu(\vec{x})\left[\frac{\partial v_x(\vec{x},t)}{\partial z} - \frac{\partial v_x(\vec{x},t)}{\partial z}\right] = S_{xz}(\vec{x},t), \tag{5}$$

where λ, and μ are the Lamé parameters, ρ is density, v_x, and v_z are the particle velocities, and σ_{xx}, σ_{zz}, and σ_{xz} are the propagated stresses. The variables on the right-hand side of these equations represent the forcing terms. As shown by Alves and Biondi (2016), we can rewrite these equations as a non-linear operator:

$$\vec{d} = f(\vec{m}), \tag{6}$$

where $\vec{d} = [v_x\ v_z\ \sigma_{xx}\ \sigma_{zz}\ \sigma_{xz}]^T$, and $\vec{m} = [\lambda\ \mu\ \rho]^T$ define our data and model vectors, respectively. In real seismic acquisition only the hydrostatic pressure and particle velocities are recorded. Therefore, we apply a linear transformation to the data vector to simulate a real experiment:

$$\vec{d}_{obs} = \mathbf{R}\vec{d}, \tag{7}$$

where the matrix $\mathbf{R}$ is defined as follows:

$$\begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \tag{8}$$

and whose effect is to average the normal stresses and extract the particle velocities from the data vector. The choice of model parametrization influences the inversion results because of the presence of parameter crosstalk (Operto et al., 2013). We choose to parametrize our model space with the vector $\vec{m}' = [V_p\ V_s\ \rho]^T$ that contains the wave propagation velocities as opposed to elastic parameters. This change of variables introduces the following non-linear transformation:

$$\vec{m} = g(\vec{m}') = \begin{bmatrix} (V_p^2 - 2V_s^2)\rho \\ V_s^2\rho \\ \rho \end{bmatrix}, \tag{9}$$

where V_p and V_s are the compressional- and shear-wave propagation velocities, respectively. Given the previous equations we define our FWI objective function as:

$$\phi(\vec{m}') = \frac{1}{2}\left\|\mathbf{R}f\left(g(\vec{m}')\right) - \vec{d}_{obs}\right\|_2^2 = \frac{1}{2}\left\|\vec{r}\right\|_2^2, \tag{10}$$

where $\vec{d}_{obs}$ is the observed multi-component data, $\vec{r}$ is the residual vector, and we parametrized the modeling operator in terms of the $\vec{m}'$ vector. To minimize this function we compute the gradient of equation 10 that can be written as follows:

$$\nabla\phi = \mathbf{G}^T\mathbf{F}^T\mathbf{R}^T\vec{r}, \tag{11}$$

where $\mathbf{G}^T$ and $\mathbf{F}^T$ are the linearized adjoint operators of equations 9 and 6, respectively. The linearized adjoint operator $\mathbf{F}^T$ can be found using the adjoint state method (Fichtner, 2010), and the Jacobian matrix of the non-linear transformation of equation 9 is given by:

$$\mathbf{G} = \begin{bmatrix} 2V_p\rho & -4V_s\rho & V_p^2 - 2V_s^2 \\ 0 & 2V_s\rho & V_s^2 \\ 0 & 0 & 1 \end{bmatrix}. \tag{12}$$

Using equation 11 we apply any gradient-based optimization algorithm, such as non-linear conjugate gradient (CG) (Fletcher and Reeves, 1964). To improve the convergence rate of any optimization method we can apply an approximated inverse Hessian matrix to precondition the gradient. In fact, this operation would approximate a Newton's optimization step. The sparse structure of the Hessian enables us to estimate its elements by applying this matrix to impulses in the model space and interpolating for the unknown values (Tang and Lee, 2015). In this study we make use of the Gauss-Newton Hessian. For the objective function in equation 10 this matrix $\mathbf{H}_{GN}$ takes the following form:

$$\mathbf{H}_{GN} = \mathbf{G}^T\mathbf{F}^T\mathbf{R}^T\mathbf{R}\mathbf{F}\mathbf{G}, \tag{13}$$

where $\mathbf{F}$ is the Born operator that maps perturbations in the model space into data perturbations.

RESULTS AND DISCUSSION

We generate a 2D complex subsurface model using the model builder software proposed by Clapp (2014). Figure 1 shows this model in terms of compressional-wave velocity, shear-wave velocity, and density. In such model, we generate multi-component data by placing 50 explosive sources at the surface spaced by 100 m. As source signature we employ a Ricker wavelet with dominant frequency of 20 Hz. The hydrophones and geophones are also positioned at the surface but with an interval of 10 m. Figure 2 displays a representative shot multi-component record.

The initial model parameters used to start the FWI problem is constructed by smoothing the ones displayed in Figure 1 (Figure 3). To avoid local minima and increase the attraction basin of the global minimum, we follow a multi-scale approach (Bunks et al., 1995). We start the inversion with a bandwidth of maximum frequency of 5 Hz, and we increase the bandwidth by intervals of 5 Hz up to 20 Hz. The last inverted model from one band is going

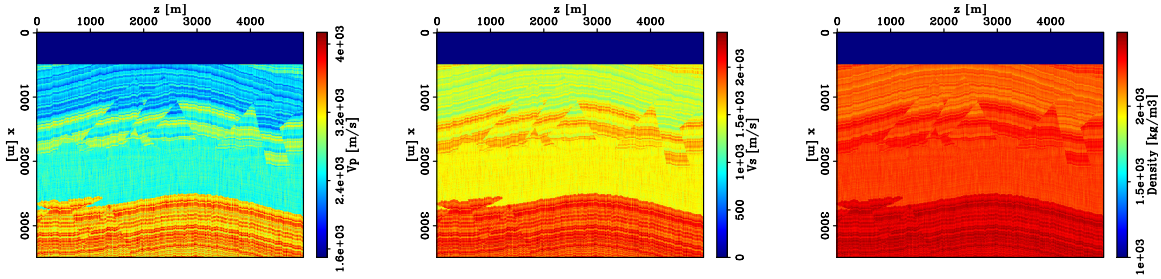


Figure 1: True subsurface model. Left panel: Compressional-wave velocity. Central panel: Shear-wave velocity. Right panel: Subsurface density. [**ER**] ettore2/. comb-true-model

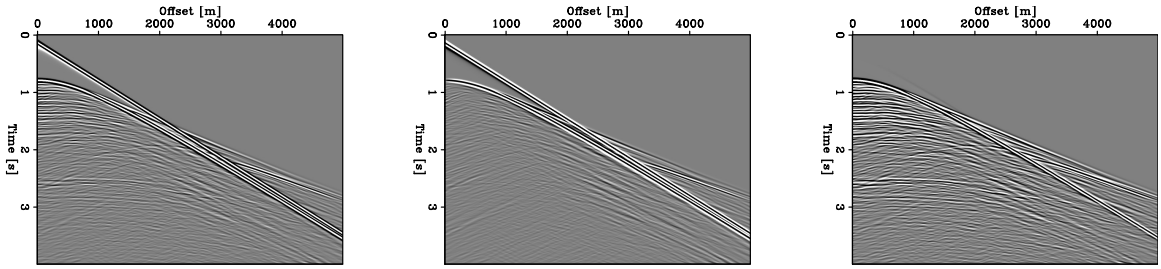


Figure 2: Multi-component data recorded for an explosive source placed on the surface at x = 0 m. Left panel: pressure component. Central panel: horizontal particle velocity component. Right panel: particle velocity component. [**CR**] ettore2/. comb-true-data

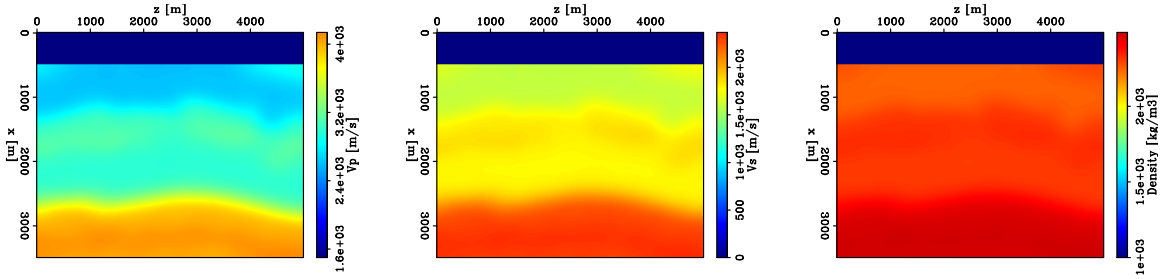


Figure 3: Starting FWI model created by smoothing the true model shown in Figure 1. Left panel: Compressional-wave velocity. Central panel: Shear-wave velocity. Right panel: Subsurface density. [**ER**] ettore2/. comb-start-model

to be our initial subsurface for the next one. For each frequency band we run 40 iterations of non-linear CG algorithm.

We compare an optimization result obtained without employing any preconditioning, and a different one where an approximated Gauss-Newton Hessian inverse is estimated for each frequency band and is used to construct a preconditioner. To estimate this matrix, we apply the matrix of equation 13 to twelve impulses placed in the subsurface and spaced by 500 m and 1000 m along the z-axis and x-axis, respectively (Figure 4). From these applications we extract the main diagonal elements and linearly interpolate the unknown Hessian values. We then create an approximated Gauss-Newton Hessian inverse that is used to precondition any computed FWI gradient. This preconditioner is estimated only once for each frequency band before starting the inversion. This approach assumes that this weighting is not changing as the FWI is modifying the model. This assumption is reasonable since the model perturbations introduced by the FWI scheme will not drastically change the Gauss-Newton Hessian matrix.

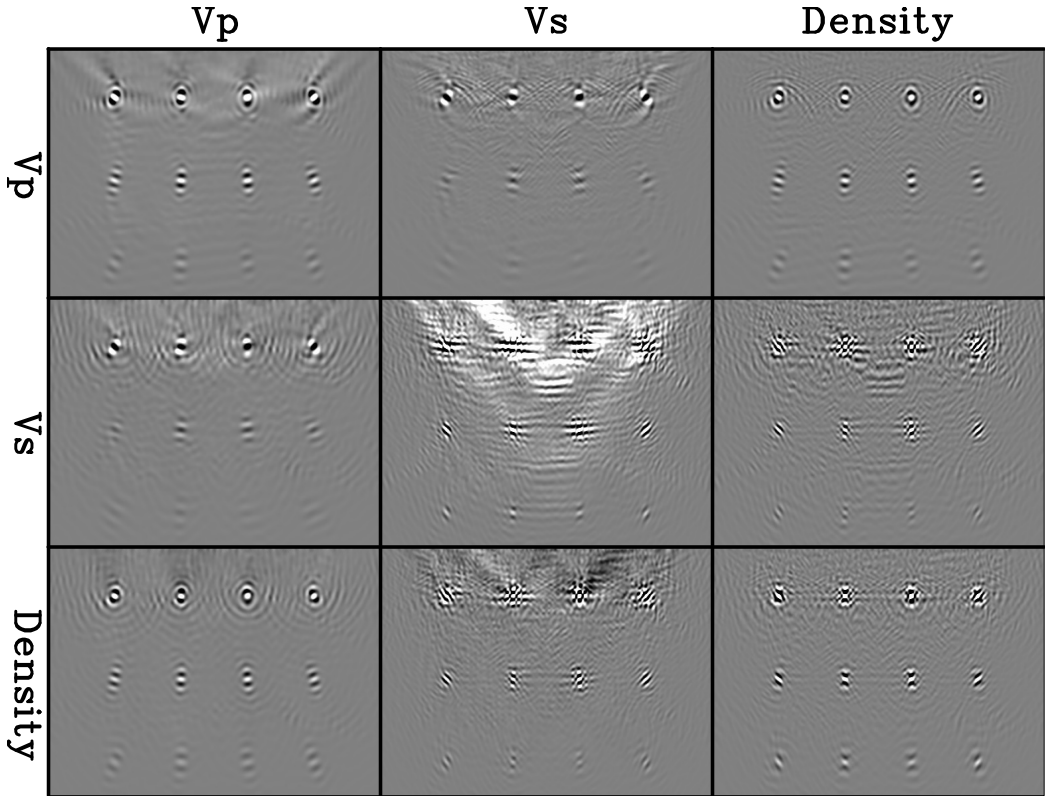
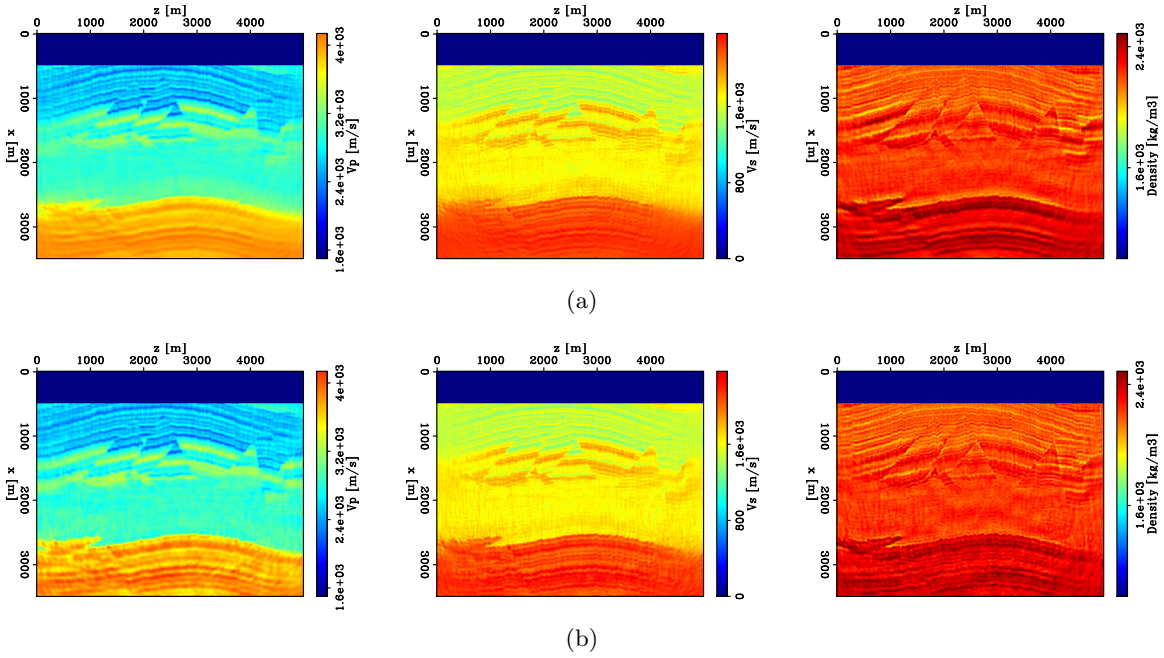


Figure 4: Application of the Gauss-Newton Hessian for the 15 Hz frequency band to spikes positioned in the subsurface. The top labels indicate the parameter image obtained by applying the Hessian matrix to impulses in the parameter class indicated by the left label. The diagonal panels enable us to estimate the main diagonal of the Gauss-Newton Hessian matrix. In these panels the water layer has been removed. [**CR**] ettore2/. hessian-sed15

Figure 5 shows the elastic FWI results. For all parameter classes, the inverted models are similar in the shallow portion of the subsurface, meaning both FWI optimizations have likely converged to the same minimum. However, in the deeper layers, they differ in terms of parameter resolution. In fact, when the optimization is preconditioned the inverted

parameters present more structural features. In addition, a low-density layer on top of the interface approximately at 2700 m does not affect the inverted density model. In terms of objective function values, both inversion results have similar behavior. From Figure 6 we observe a modest improvement in convergence rate when the problem is preconditioned with a Gauss-Newton main-diagonal approximation, especially as we increase the frequency content in the inversion. In this test, the effect the preconditioner is to mostly decrease the model crosstalk by properly scaling the gradients of the simultaneously inverted parameter classes.



(a)

(b)

Figure 5: Inverted FWI model comparison. (a) Inverted model without preconditioning. (b) Inverted model using approximated Gauss-Newton inverse. [**CR**] ettore2/. comb-CG-model,comb-GN-model

CONCLUSIONS

We discuss how to derive the Gauss-Newton Hessian approximation of elastic multi-component FWI as a series of linear operators when wave velocities and density parameterize the inverse problem. We describe how to estimate these matrix elements from applications to model vectors containing sparse impulses. Using linear interpolation we compute the missing matrix elements.

On a complex 2D synthetic model we show the use of this estimated matrix to precondition the FWI problem when multiple parameters are inverted simultaneously. We demonstrate that a simple main-diagonal approximation already provides a moderate convergence improvement and properly scales the FWI gradients. In fact, this scaling removes inversion artifacts in the final inverted model. Future works will explore the effect of introducing more elements when approximating the Hessian inverse as preconditioning matrix.

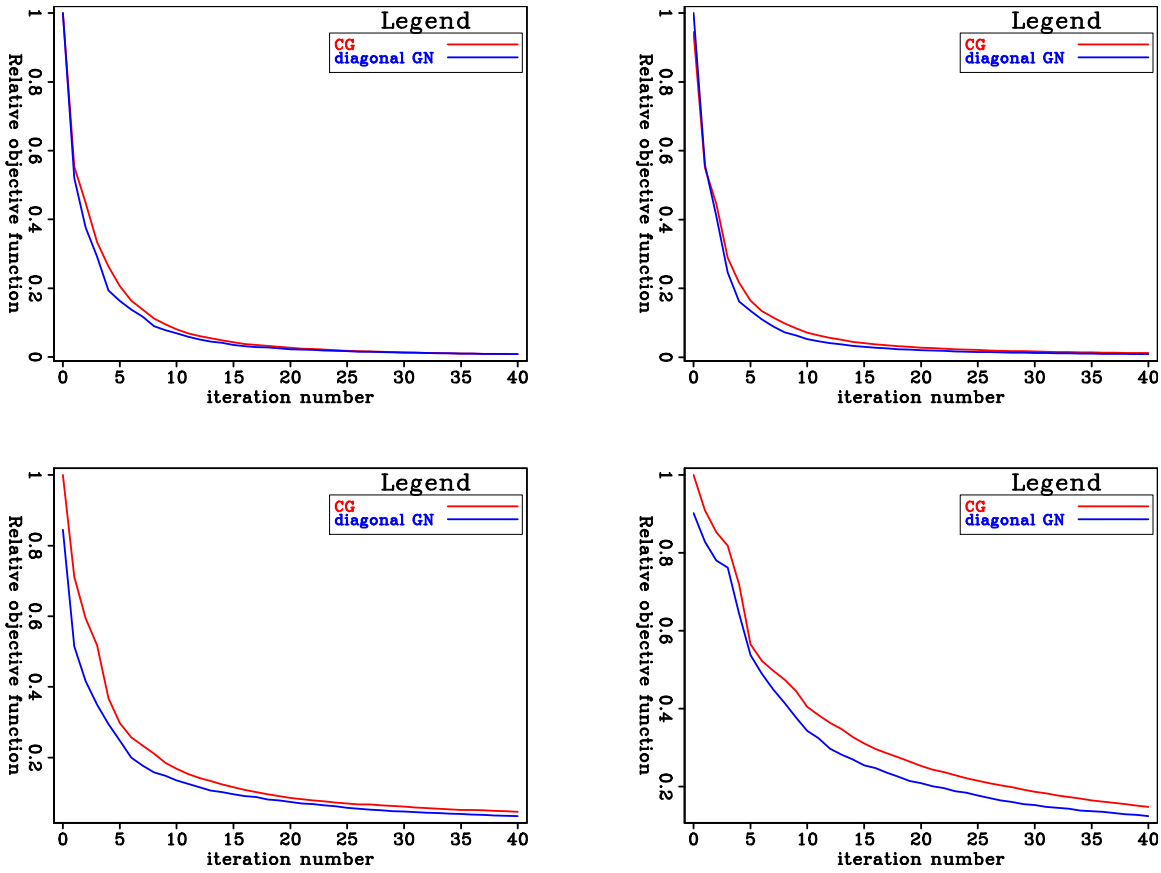


Figure 6: Relative objective function comparison between CG (red curve) and preconditioned CG (blue curve) for maximum frequency content of 5 (top-left), 10 (top-right), 15 (bottom-left), and 20 Hz (bottom-right) used during the inversion. [**CR**] ettore2/. comb-obj-rep

REFERENCES

Albertin, U., P. Shen, A. Sekar, T. Johnsen, C. Wu, K. Nihei, and K. Bube, 2016, 3D orthorhombic elastic full-waveform inversion in the reflection domain from hydrophone data: SEG Technical Program Expanded Abstracts 2016, 1094–1098, Society of Exploration Geophysicists.

Alkhalifah, T., 2015a, Conditioning the full-waveform inversion gradient to welcome anisotropy: Geophysics, **80**, R111–R122.

———, 2015b, Scattering-angle based filtering of the waveform inversion gradients: Geophysical Journal International, **200**, 363–373.

Alves, G. and B. Biondi, 2016, Imaging condition for elastic reverse time migration: SEG Technical Program Expanded Abstracts 2016, 4173–4177, Society of Exploration Geophysicists.

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

Clapp, R., 2014, Synthetic model building using a simplified basin modeling approach: SEP-Report, **155**, 143–150.

Deuzeman, A. and R.-E. Plessix, 2015, Block-diagonal approximation of the Hessian for multi-parameter FWI: 12th SIAM Conference on Mathematical and Computational Issues in Geosciences, Expanded Abstracts, MS2, Society of Industrial and Applied Mathematics.

Fichtner, A., 2010, Full seismic waveform modelling and inversion: Springer Science & Business Media.

Fletcher, R. and C. M. Reeves, 1964, Function minimization by conjugate gradients: The computer journal, **7**, 149–154.

Forgues, E. and G. Lambaré, 1997, Parameterization study for acoustic and elastic ray plus born inversion: Journal of Seismic Exploration, **6**, 253–277.

Huang, Y., R. Nammour, and W. Symes, 2016, Flexibly preconditioned extended least-squares migration in shot-record domain: Geophysics, **81**, S299–S315.

Innanen, K., 2014, Reconciling seismic AVO and precritical reflection FWI-analysis of the inverse Hessian: SEG Technical Program Expanded Abstracts 2014, 1022–1027, Society of Exploration Geophysicists.

Korta, N., A. Fichtner, and V. Sallarčs, 2013, Block-diagonal approximate hessian for preconditioning in full waveform inversion: Presented at the 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013.

Operto, S., Y. Gholami, V. Prieux, A. Ribodetti, R. Brossier, L. Metivier, and J. Virieux, 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice: The Leading Edge, **32**, 1040–1054.

Tang, Y., 2008, Wave-equation hessian by phase encoding: SEG Technical Program Expanded Abstracts 2008, 2201–2205, Society of Exploration Geophysicists.

Tang, Y. and S. Lee, 2015, Multi-parameter full wavefield inversion using non-stationary point-spread functions: SEG Technical Program Expanded Abstracts 2015, 1111–1115, Society of Exploration Geophysicists.

ten Kroode, F., 2012, A wave-equation-based Kirchhoff operator: Inverse Problems, **28**, 115013.

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

Reciprocity in elastic multi-component data

Gustavo Alves and Ettore Biondi

ABSTRACT

Reciprocity for source and receiver pairs has been routinely adopted in seismic processing to mitigate poor receiver density and improve computational costs. But while its implementation is straightforward for acoustic data, the correct combination of source and receiver components in elastic, multi-component data, can be challenging. We derive the Green's functions for the direct and reciprocal data for a source-receiver pair and show synthetic results demonstrating their equivalence. We also show how an explosive source can be represented as a combination of normal stresses or particle velocities. Finally, we calculate the respective wavelets for each formulation both analytically and by linear inversion.

INTRODUCTION

Spatial reciprocity in elastic systems, also known as Betti's theorem, relates applied forces and observed particle displacements between different points. It states that the Green's function associated with a particle displacement observed at point A due to a force at point B can be related to the Green's function of an equivalent problem where particle displacement and force positions have been interchanged (Aki and Richards, 1980).

Reciprocity can be advantageous in data processing techniques, where differences in spatial sampling of source and receiver domains can be mitigated by exchanging their positions, usually by re-sorting common-shot gathers into common-receiver gathers (Nowack et al., 2003). Such differences in spatial sampling are particularly common in Ocean Bottom Node (OBN) surveys, where the distance between receivers is in the order of hundreds of meters, while the distance between shots is in the order of tens of meters. In areas with low signal-to-noise (S/N) ratios, Regone et al. (2015) have shown that denser receiver grids are favored over denser source grids when producing Pre-Stack Depth Migration images.

Reciprocity in data processing can also lower the computational cost of imaging. When imaging involves propagation of individual shots, such as in Reverse Time Migration (RTM) or other wave field extrapolation techniques, exchanging shots and receivers can lower the cost of computation, since the number of shots in the direct problem is usually an order of magnitude greater than the number of receivers.

While reciprocity can be trivially implemented when dealing with acoustic data, vector data, as is the case of geophones, requires a more careful analysis of reciprocal components. In order to understand that, one should explicitly describe the sources and receivers in their respective forces and displacements, as well as the Green's functions that connect them.

Here, we start by showing the equivalence between normal stresses and particle velocities when describing explosive type sources. We use the velocity-stress formulation of the

elastic wave equation (Virieux, 1986) to analytically prove that equivalence. We then show numerically how the same analytical solution is found by a linear inversion of the source wavelet for a particle velocity quadrupole source.

Next, we describe the Green's functions that relate pairs of particle velocities at source and receiver positions and obtain their reciprocal functions. We show that a direct problem with a source in both vertical and horizontal velocities can be simulated by two reciprocal experiments where each particle velocity component is injected separately at the receiver position and recorded as a sum of components at the source. Finally, we show these results for a synthetic experiment in the Marmousi2 elastic model.

METHODOLOGY

In marine acquisition, the physical effect of airguns as seismic sources can be represented as a single point in which the moment tensor is diagonal (Aki and Richards, 1980). Using the velocity-stress formulation of the isotropic elastic wave equation, the injection of this kind of signal into the subsurface is given by the following system of partial differential equations:

$$\rho(x,z)\frac{\partial v_x(x,z,t)}{\partial t} = \frac{\partial \sigma_{xx}(x,z,t)}{\partial x} + \frac{\partial \sigma_{xz}(x,z,t)}{\partial z}, \tag{1}$$

$$\rho(x,z)\frac{\partial v_z(x,z,t)}{\partial t} = \frac{\partial \sigma_{xz}(x,z,t)}{\partial x} + \frac{\partial \sigma_{zz}(x,z,t)}{\partial z}, \tag{2}$$

$$\frac{\partial \sigma_{xx}(x,z,t)}{\partial t} = [\lambda+2\mu](x,z)\frac{\partial v_x(x,z,t)}{\partial x} + \lambda(x,z)\frac{\partial v_z(x,z,t)}{\partial z} + S_{ex}(x,z,t), \tag{3}$$

$$\frac{\partial \sigma_{zz}(x,z,t)}{\partial t} = \lambda(x,z)\frac{\partial v_x(x,z,t)}{\partial x} + [\lambda+2\mu](x,z)\frac{\partial v_z(x,z,t)}{\partial z} + S_{ex}(x,z,t), \tag{4}$$

$$\frac{\partial \sigma_{xz}(x,z,t)}{\partial t} = \mu(x,z)\left[\frac{\partial v_x(x,z,t)}{\partial z} + \frac{\partial v_x(x,z,t)}{\partial z}\right], \tag{5}$$

where λ, and μ represent the elastic parameters, ρ is the medium density, v_x, and v_z are the particle velocities, and σ_{xx}, σ_{zz}, and σ_{xz} are the propagated stresses. In equations 3 and 4 the forcing term is given by:

$$S_{ex}(x,z,t) = \delta(x-x_s, z-z_s)w(t), \tag{6}$$

where $w(t)$ is the source signature, and $\delta(x-x_s, z-z_s)$ is a spike positioned at the source location. In order to derive the reciprocal experiment based on the elastic reciprocity theorem shown in Aki and Richards (1980), it is easier to find the equivalent body force given by an explosive seismic source. To do so, we have to integrate in time and substitute equations 3 and 4 into 1 and 2, respectively:

$$\begin{aligned}\rho(x,z)\frac{\partial v_x(x,z,t)}{\partial t} = \quad & \frac{\partial}{\partial x}\int_{-\infty}^{t}\left\{[\lambda+2\mu](x,z)\frac{\partial v_x(x,z,\tau)}{\partial x} + \lambda(x,z)\frac{\partial v_z(x,z,\tau)}{\partial z}\right\}d\tau \\ & + \int_{-\infty}^{t}\frac{\partial S_{ex}(x,z,\tau)}{\partial x}d\tau + \frac{\partial \sigma_{xz}(x,z,t)}{\partial z},\end{aligned} \tag{7}$$

$$\begin{aligned}\rho(x,z)\frac{\partial v_z(x,z,t)}{\partial t} = \quad & \frac{\partial}{\partial x}\int_{-\infty}^{t}\left\{\lambda(x,z)\frac{\partial v_x(x,z,\tau)}{\partial x} + [\lambda+2\mu](x,z)\frac{\partial v_z(x,z,\tau)}{\partial z}\right\}d\tau \\ & + \int_{-\infty}^{t}\frac{\partial S_{ex}(x,z,\tau)}{\partial z}d\tau + \frac{\partial \sigma_{xz}(x,z,t)}{\partial x}.\end{aligned} \tag{8}$$

In these two equations we can clearly distinguish the forcing terms in the x and z components:

$$f_x(x,z,t) = \int_{-\infty}^{t} \frac{\partial S_{ex}(x,z,\tau)}{\partial x} d\tau = \frac{\partial \delta(x-x_s, z-z_s)}{\partial x} \int_{-\infty}^{t} w(\tau) d\tau, \tag{9}$$

$$f_z(x,z,t) = \int_{-\infty}^{t} \frac{\partial S_{ex}(x,z,\tau)}{\partial z} d\tau = \frac{\partial \delta(x-x_s, z-z_s)}{\partial z} \int_{-\infty}^{t} w(\tau) d\tau. \tag{10}$$

These two body forces have the same time signature that corresponds to the time integral of the original source wavelet of equation 6. On the other hand, the space signature is given by derivatives in the x and z directions of a delta function, respectively. This effect can be approximated by dipoles in the two directions, which is the same result found in the main diagonal of the moment tensor shown by Aki and Richards (1980).

When considering a multi-component marine ocean-bottom acquisition we have to be careful to separate the various recorded components to create the correct reciprocal experiments. In fact, it is important to understand that the hydrophone is coupled to the water; whereas, the geophones are coupled to the sea bed. Because of this differential coupling we have at least two reciprocal experiments to consider, one for an acoustic medium and a different one for an elastic subsurface. We are going to see that in reality we need three reciprocal propagations in this acquisition scenario. For the hydrophone component the true experiment can be represented by a time convolution of the explosive source response with the acoustic Green's function, as follows:

$$\begin{aligned} p(x_{r_h}, z_{r_h}, t) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_{r_h}, z_{r_h}, t; x, z) \star S_{ex}(x,z,t) dx dz \\ &= g(x_{r_h}, z_{r_h}, t; x_s, z_s) \star w(t), \end{aligned} \tag{11}$$

where (x_{r_h}, z_{r_h}) represents the hydrophone position, and $\star$ the time convolution operator. In this case the reciprocal experiment is given by interchanging the source position with the receiver location:

$$p(x_{r_h}, z_{r_h}, t) = g(x_s, z_s, t; x_{r_h}, z_{r_h}) \star w(t). \tag{12}$$

On the other hand, for the geophone component we have to employ the following property of elastic Green's functions (Aki and Richards, 1980):

$$g_{ij}(x_r, z_r, t; x_s, z_s) = g_{ji}(x_s, z_s, t; x_r, z_r), \tag{13}$$

where $g_{ij}(x_r, z_r, t; x_s, z_s)$ is the Green's function of a body force injected in the j-th direction from (x_s, z_s) and recorded in the i-th axis at (x_r, z_r). Using the representation theorem and the forcing terms of equations 9 and 10 we can describe the recorded particle velocities as:

$$\begin{aligned} v_x(x_{r_g}, z_{r_g}) = &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [g_{xx}(x_{r_g}, z_{r_g}, t; x, z) \star \dot{f}_x(x,z,t) \\ &+ g_{xz}(x_{r_g}, z_{r_g}, t; x, z) \star \dot{f}_z(x,z,t)] dx dz \\ = &[g_{xx}(x_{r_g}, z_{r_g}, t; x_s + dx, z_s) - g_{xx}(x_{r_g}, z_{r_g}, t; x_s - dx, z_s) \\ &+ g_{xz}(x_{r_g}, z_{r_g}, t; x_s, z_s + dz) - g_{xz}(x_{r_g}, z_{r_g}, t; x_s, z_s - dz)] \star w(t), \end{aligned} \tag{14}$$

$$\begin{aligned} v_z(x_{r_g}, z_{r_g}) = &[g_{zz}(x_{r_g}, z_{r_g}, t; x_s, z_s + dz) - g_{zz}(x_{r_g}, z_{r_g}, t; x_s, z_s - dz) \\ &+ g_{zx}(x_{r_g}, z_{r_g}, t; x_s + dx, z_s) - g_{zx}(x_{r_g}, z_{r_g}, t; x_s - dx, z_s)] \star w(t), \end{aligned} \tag{15}$$

in which (x_{r_g}, z_{r_g}) is the geophone position, and where we have approximated the derivatives of a delta function by:

$$\frac{\partial \delta(x,z)}{\partial x} \approx [\delta(x+dx,z) - \delta(x-dx,z)]/dx, \tag{16}$$

$$\frac{\partial \delta(x,z)}{\partial z} \approx [\delta(x,z+dz) - \delta(x,z-dz)]/dz. \tag{17}$$

By employing the result of equation 13 in relations 14 and 15 we have:

$$\begin{aligned} v_x(x_{r_g}, z_{r_g}) = [&g_{xx}(x_s+dx, z_s, t; x_{r_g}, z_{r_g}) - g_{xx}(x_s-dx, z_s, t; x_{r_g}, z_{r_g}) \\ &+ g_{zx}(x_s, z_s+dz, t; x_{r_g}, z_{r_g}) - g_{zx}(x_s, z_s-dz, t; x_{r_g}, z_{r_g})] \star w(t), \end{aligned} \tag{18}$$

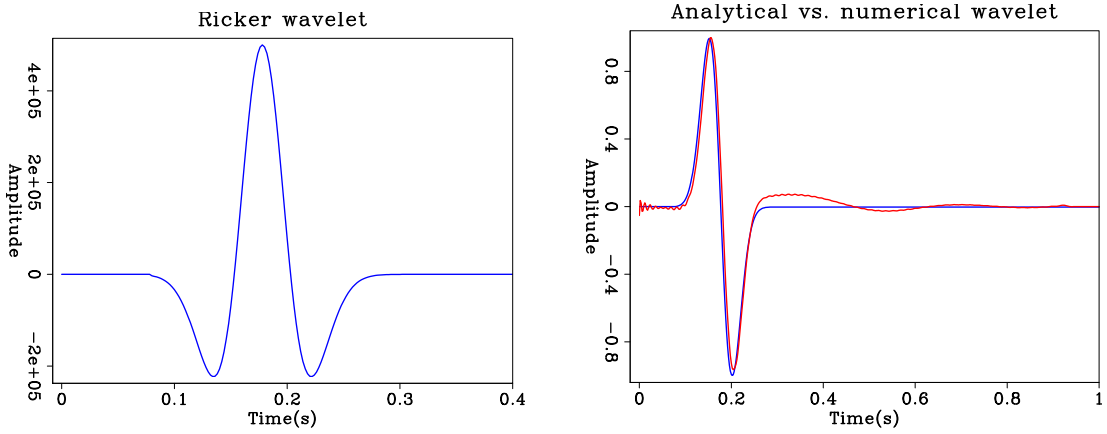
$$\begin{aligned} v_z(x_{r_g}, z_{r_g}) = [&g_{zz}(x_s, z_s+dz, t; x_{r_g}, z_{r_g}) - g_{zz}(x_s, z_s-dz, t; x_{r_g}, z_{r_g}) \\ &+ g_{xz}(x_s+dx, z_s, t; x_{r_g}, z_{r_g}) - g_{xz}(x_s-dx, z_s, t; x_{r_g}, z_{r_g})] \star w(t). \end{aligned} \tag{19}$$

These two equations describe the actual reciprocal experiment when an explosive source is employed and particle velocities are acquired. We notice that in both components we are injecting a force at the receiver location x_{r_g}, z_{r_g} only in one direction, that is in x and z, respectively. However, the recording of the particle velocities is performed using a quadrupole. For instance, the particle velocity $v_x(x_{r_g}, z_{r_g})$ component can be obtained by injecting a force in the horizontal axis and recording the propagated v_x and v_z with a dipole along the corresponding direction. The same is true for the vertical component but with an injection in the z axis. Because we are injecting a force along one axis per component in the reciprocal experiment, we have to perform two different propagations. This observation means that in order to perform the reciprocal experiment for an ocean-bottom multi-component survey we have to propagate three independent propagations, that is one for the pressure component and two for the geophone components. Therefore, as long as the number of receivers is less than a third the number of sources, the reciprocal experiment involves fewer propagations than the true acquisition. Despite this apparent issue we clearly see the advantage in this acquisition scenario. In fact, in this case we usually have thousands of sources and just few hundreds of receivers such that the aforementioned condition is potentially fulfilled.

RESULTS

Equivalence between stress and velocity explosive sources

We started by generating multi-component elastic data in a constant background model by applying an explosive source in the normal stress components of the wavefield, as described by equations 3 and 4. The source was a Ricker type wavelet with peak frequency 10Hz, as shown in Figure 1(a). Next, we run a linear inversion of the source wavelet, but now instead of finding a source for the stress components, we look for a wavelet that minimizes the objective function and is described by a quadrupole in the velocity components of the wavefield, as we have shown in equations 9 and 10. According to our analytical solution, the quadrupole wavelet should be the time integral of the original Ricker. Figure 1(b) shows a comparison of the analytical and inverted solutions and figure 2 shows the objective function for the linear inversion problem.



(a) (b)

Figure 1: (a) Ricker wavelet used to generate data and (b) comparison between the time integral of the Ricker wavelet (blue) and the numerical approximation obtained by linear inversion (red). [**CR**] alves1/. ricker,comp-wave

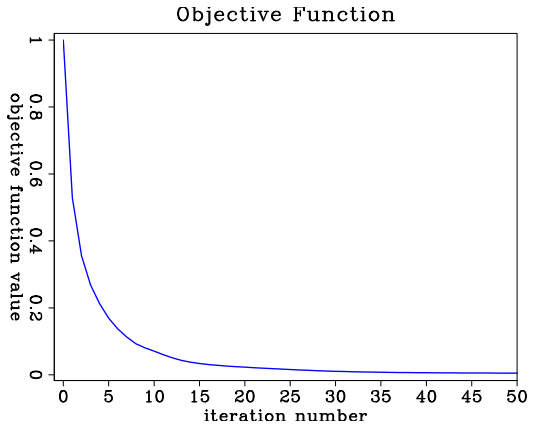


Figure 2: Objective function for the first 50 iterations of the linear inversion of the wavelet. The value at the 50th iteration is 0.5% of the initial value. [**CR**] alves1/. obj-function

Next, we run two synthetic experiments using the 2D elastic Marmousi model (Martin et al., 2002). In the first experiment, we inject an explosive source using the stress formulation in equations 3 and 4. In the second experiment, we repeat the positions for source and receiver, but now inject an explosive source according to equations 9 and 10. Our goal is to show that multi-component elastic data recorded at the ocean-bottom interface is equal for both source formulations, even with a fairly complex model. Figure 3 shows all data components overlaid for (red) stress sources and (blue) particle velocity sources.

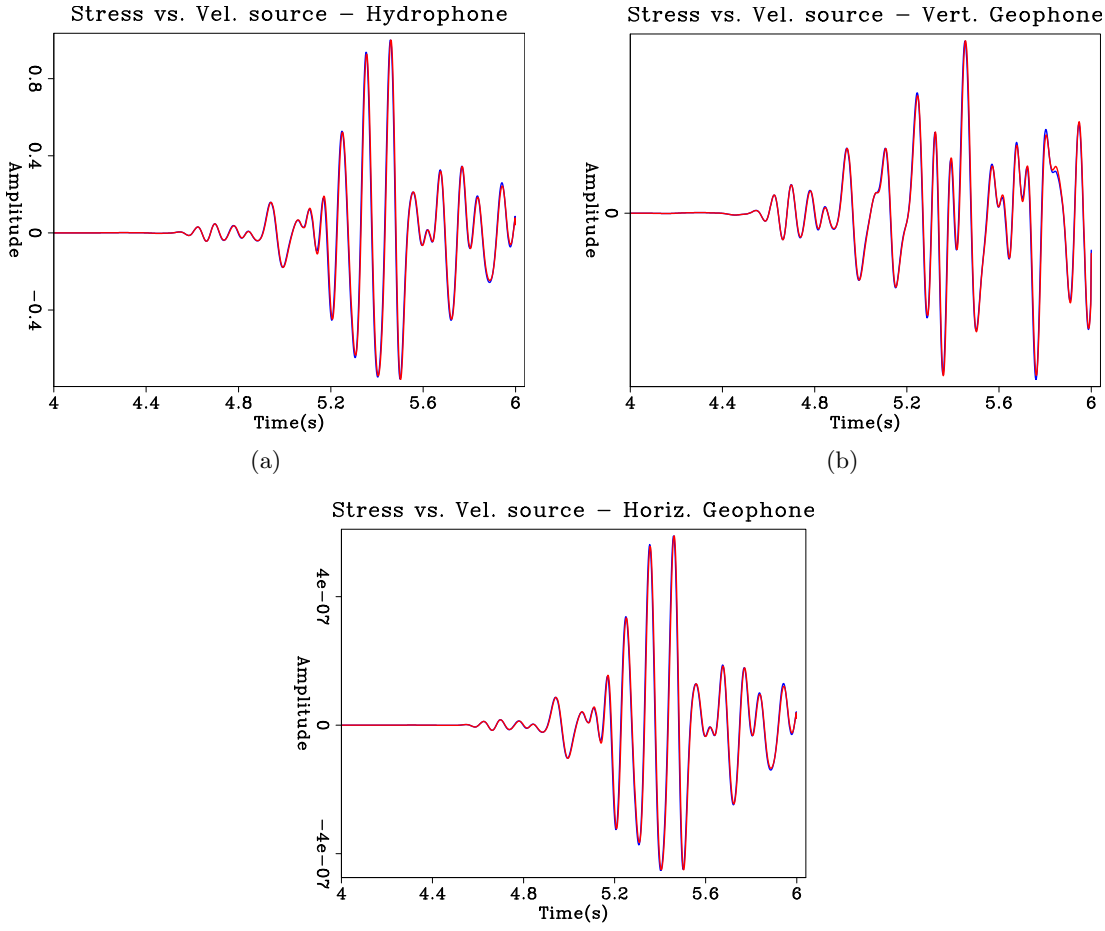
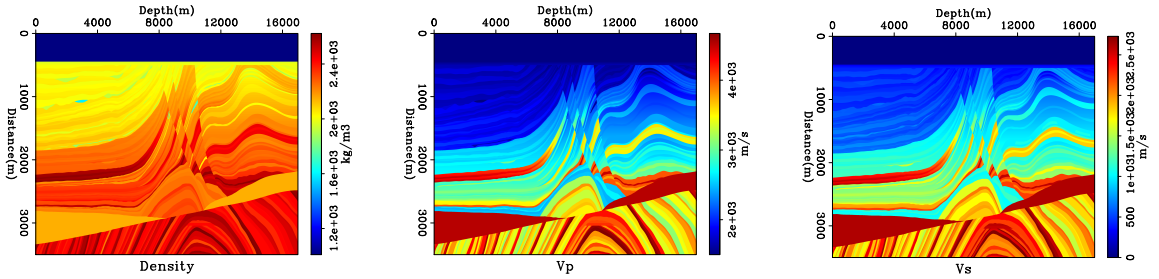


Figure 3: Data recorded for one receiver due to an explosive source injected in the stress components (red) or as a particle velocity quadrupole (blue). Graphs show (a) hydrophone, (b) vertical geophone and (c) horizontal geophone data. [**ER**] alves1/. comp-ex-f-hyd,comp-ex-f-ver,comp-ex-f-hor

Reciprocity in multi-component data

The 2D elastic Marmousi model is a fairly complex model even for shallow targets, which allows us to observe many reflections and mode conversions in the synthetic data and study the validity of the reciprocal data for different types of events. Figure 4 shows the elastic parameters for this model.



(a) (b) (c)

Figure 4: (a) Density, (b) V_p and (c) V_s for the 2D Marmousi 2 elastic model. [**ER**] alves1/. rho,vp,vs

In order to get the direct data, we place an initial explosive source close to the water surface, representing an air-gun. The synthetic source is modeled using the quadrupole in particle velocities. We position the receiver at the ocean-bottom interface, so it can mimic the complexity of the elastic coupling to the solid medium observed in ocean-bottom data. Source and receivers are not aligned in depth, so that amplitude versus angle relations are observed.

As we described in the previous section, we need three independent experiments to generate the reciprocal data for a pair of source and receiver. The first experiment injects the source at the reciprocal position as a quadrupole in velocities. This experiment generates the traditional acoustic reciprocity, giving us the hydrophone reciprocal data. Next, we run two modelings, a vertical and a horizontal velocity source and record the data as a quadrupole at the new receiver position. We summarize our results in figure 5, that shows the direct and reciprocal multi-component data. It can be seen that both data sets overlay almost perfectly. There are minor differences that we believe can be attributed to numerical accuracy.

CONCLUSION

We show that equivalent synthetic acoustic data can be generated in the velocity-stress formulation by either sources in the normal stress components or in the particle velocity components. When dealing with multi-component data, this is advantageous because a single source wavelet can be used for generating reciprocal data for both scalar (hydrophone) and vector (geophone) components.

We describe the Green's functions that relate pairs of particle velocities at the source to pairs of particle velocities at the receiver. Furthermore, we describe the reciprocal equations and show that, for a two-component source in the direct problem, the reciprocal problem is described by two independent solutions obtained by injecting single-component particle velocities at the receiver position and extracting a sum of both components at the source. While this increases the number of simulations from one in the direct problem to three in the reciprocal problem (acoustic data plus two vector components), we believe that such tradeoff is still computationally advantageous when the ratio between number of sources and number of receivers is greater than three. Also, as mentioned in the Introduction,

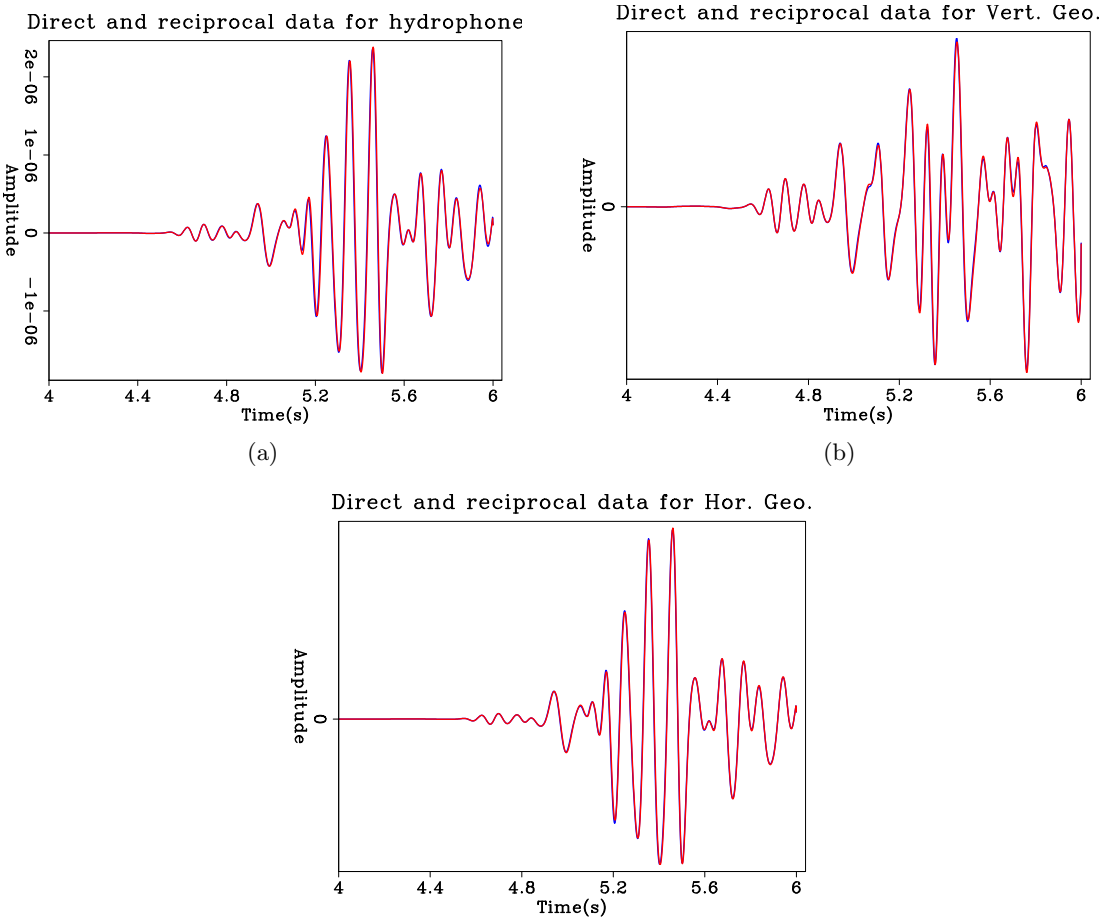


Figure 5: Direct (blue) and reciprocal (red) data for the (a) hydrophone, (b) vertical-geophone and (c) horizontal geophone. [**ER**] alves1/. comp-rec-hyd,comp-rec-ver,comp-rec-hor

spatial sampling in source and receiver domains can justify the use of reciprocal data.

ACKNOWLEDGEMENTS

We would like to thank SEP sponsors for the ongoing support.

REFERENCES

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

Martin, G. S., S. Larsen, K. Marfurt, et al., 2002, Marmousi-2: an updated model for the investigation of avo in structurally complex areas: Presented at the 2002 SEG Annual Meeting.

Nowack, R. L., M. K. Sen, and P. L. Stoffa, 2003, Gaussian beam migration for sparse common-shot and common-receiver data, *in* SEG Technical Program Expanded Abstracts 2003, 1114–1117, Society of Exploration Geophysicists.

Regone, C., M. Fry, and J. Etgen, 2015, Dense sources vs. dense receivers in the presence of coherent noise: a land modeling study, *in* SEG Technical Program Expanded Abstracts 2015, 12–16, Society of Exploration Geophysicists.

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

Phase Only Full Waveform Inversion of Elastic Data Using Acoustic Engine

Stuart Farris, Ettore Biondi, and Guillaume Barnier

ABSTRACT

Modeling the elastic earth with acoustic assumptions leads to large amplitude mismatches between observed and predicted data. Matching both amplitude and phase information between elastic field data and acoustic modeled data is more difficult than only matching phase information. Therefore, we attempt to perform acoustic full waveform inversion (FWI) with an objective function designed to favor minimizing phase variation between observed and predicted data.

INTRODUCTION

Seismic experiments observe wavefields that are propagated through an elastic earth. We try to create a model of the earth that predicts the elastic data we observe in the field. Full waveform inversion (FWI) provides an iterative technique to find such an earth model by matching observed field data to predicted model data (Tarantola, 1984). FWI has been shown to produce earth models with much higher resolution than competing methods (Sirgue et al., 2009). However, the high cost of FWI limits the complexity of its modeling capabilities. Therefore, FWI performed on large commercial datasets often models with acoustic wave propagation which is computationally cheaper than modeling with elastic propagation.

It should seem obvious that an inversion scheme that matches simplified, acoustic data with complex, elastic data will create incorrect models of the earth (Warner et al., 2012). Previous attempts have been made to rectify the differences between acoustic and elastic modeling. Namely, correcting acoustic propagation amplitudes by simulating elastic effects (Ben Veitch et al., 2012), using match filters to remove elastic effects from the initial observed data (Agudo et al., 2016), and conditioning the FWI objective function to minimize phase variations (Shen, 2014). The later approach will be applied here.

The method proposed by Shen (2014) uses an acoustic FWI engine, but performs an amplitude normalization within the objective function. This normalization steers the objective function to better match phase variations between the observed elastic and predicted acoustic data. The goal is to ignore the amplitude variations that cannot be matched because of the inherent differences between the two modeling techniques. In this report we evaluate the effectiveness of this technique by comparing its FWI results with those of standard acoustic FWI and elastic FWI.

THEORY

We redefine the objective function proposed by Shen (2014) in a more verbose manner:

$$\Phi(\mathbf{m}) = \frac{1}{2}\sum_{l=1}^{N_{traces}} \left|\left| \frac{\mathbf{d}_l^p}{||\mathbf{d}_l^p||} - \frac{\mathbf{d}_l^o}{||\mathbf{d}_l^o||} \right|\right|_2^2 = \frac{1}{2}\sum_{l=1}^{N_{traces}} \left|\left| \mathbf{r}_l^d \right|\right|_2^2 \tag{1}$$

where $\mathbf{m} \in \mathbb{R}^M$ defines our model space. $\mathbf{d}_l^p, \mathbf{d}_l^o \in \mathbb{R}^N$ are predicted traces and observed traces, respectively. $||\mathbf{d}_l||$ is the normalization term for a given trace, l, defined as:

$$||\mathbf{d}_l|| = \left(\sum_{t=1}^{N} d_{l,t}^2\right)^{\frac{1}{2}} \tag{2}$$

The residual $\mathbf{r}_l^d \in \mathbb{R}^N$ is the difference between one normalized predicted trace and one normalized observed trace. A physical interpretation of the objective function can be obtained by expanding Equation 1:

$$\Phi(\mathbf{m}) = \frac{1}{2}\left(\frac{\mathbf{d}_l^p}{||\mathbf{d}_l^p||} - \frac{\mathbf{d}_l^o}{||\mathbf{d}_l^o||}\right)^T \left(\frac{\mathbf{d}_l^p}{||\mathbf{d}_l^p||} - \frac{\mathbf{d}_l^o}{||\mathbf{d}_l^o||}\right) \tag{3}$$

$$= 1 - \frac{1}{2}\left(\frac{\mathbf{d}_l^p}{||\mathbf{d}_l^p||} \cdot \frac{\mathbf{d}_l^o}{||\mathbf{d}_l^o||}\right) \tag{4}$$

Minimizing the objective function can be interpreted as maximizing the zero lag cross correlation between the normalized predicted traces and normalized observed traces where the normalization term is the energy of each trace. The normalization weights the relative amplitudes of each trace so that maximizing the cross correlation corresponds to minimizing the phase difference between respective traces. This would be a true phase only inversion if we compared individual frequencies of individual arrival events. Here we find the limitation of this objective function since, of course, this is unfeasible in real applications. However, narrow band passes during this scheme's application make the comparison of multiple arrivals possible and move this objective function towards true phase only inversion (Shen, 2014).

To determine model update directions we need the derivative of the objective function. For simplicity we will consider a single trace:

$$\nabla_m \Phi(\mathbf{m}) = \frac{1}{2} \nabla_m ||\mathbf{r}^d||_2^2 \tag{5}$$

$$= \left(\frac{\partial \mathbf{r}^d}{\partial \mathbf{m}}\right)^T \mathbf{r}^d \tag{6}$$

$$= \left(\frac{\partial \mathbf{d}^p}{\partial \mathbf{m}}\right)^T \left(\frac{\mathbf{r}^d}{||\mathbf{d}^p||} - \frac{(\mathbf{r}^d)^T \mathbf{d}^p}{||\mathbf{d}^p||^3}\mathbf{d}^p\right) \tag{7}$$

$$= \mathrm{B}^T \tilde{\mathbf{r}}^d \tag{8}$$

We find that the derivative of the objective function is the adjoint of the Born operator applied to some normalized residual. The full derivation of the gradient can be found in the appendix. With the new objective function, Equation 1, and its derivative, Equation 7, we can perform a nonlinear conjugate gradient inversion scheme for our model $\mathbf{m}$.

OBSERVED DATA

To test the effectiveness of our inversion scheme, we conduct a seismic experiment in an elastic, synthetic model and invert for the pressure wave velocity. The elastic model is five kilometers wide, three kilometers deep, and includes a 500 meter water column. Figure 1 illustrates the pressure wave velocity, v_p, shear wave velocity, v_s, and density, ρ, of the model. The v_s/v_p ratio was chosen to be a constant, reasonable value of 1.75. The structure was created using a synthetic model generation code written by Clapp (2014). Using the true elastic model, seismic data is recorded at the surface with 50 shots spaced at 100 meters. Receivers are placed every 10 meters and record vertical and horizontal displacement.

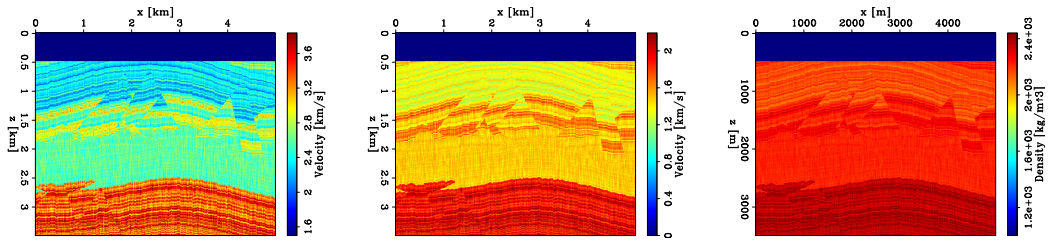


Figure 1: Elastic properties of synthetic model. Left: pressure wave velocity. Center: shear wave velocity. Right: density. [**CR**] sfarris2/. elasticModel

Figure 2 illustrates a sample shot gather with collected horizontal displacement, vertical displacement, and calculated pressure data. Pressure measurements can be calculated using,

$$P = \frac{1}{2}(\sigma_{xx} + \sigma_{zz}) \tag{9}$$

where σ_{xx} and σ_{zz} are the horizontal and vertical stress, respectively.

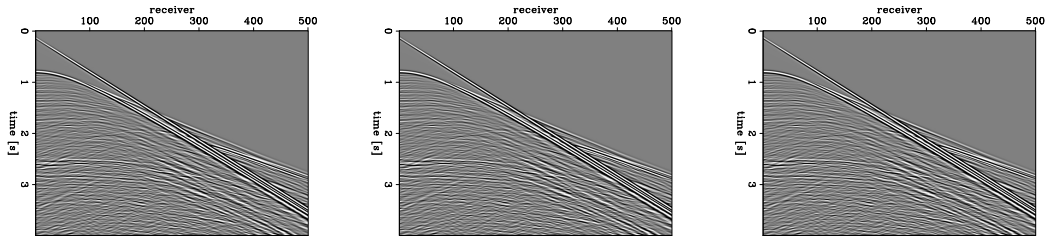


Figure 2: Example shot gather from elastic pressure data. Left: σ_{xx} Center: σ_{zz}. Right: Pressure. [**CR**] sfarris2/. elasticData

INVERSION PLAN

To evaluate the feasibility of the phase only acoustic FWI, we will compare our v_p model results with elastic FWI results and basic acoustic FWI results. Each FWI application was solved using the nonlinear solver created by Ettore Biondi and Guillaume Barnier (Biondi and Barnier, 2017).

When we perform elastic propagation to produce pressure data, we inject a source term with components in the x and z directions. But, when inverting with our acoustic phase

only engine we need to inject the an equivalent acoustic source wavelet that only has one component. Using a least squares regression we can invert for the acoustic wavelet that corresponds to the elastic wavelet used to create the observed data. The original elastic wavelet, a Ricker wavelet with a dominant frequency of 20 Hz, and the inverted acoustic wavelet are illustrated in Figure 3.

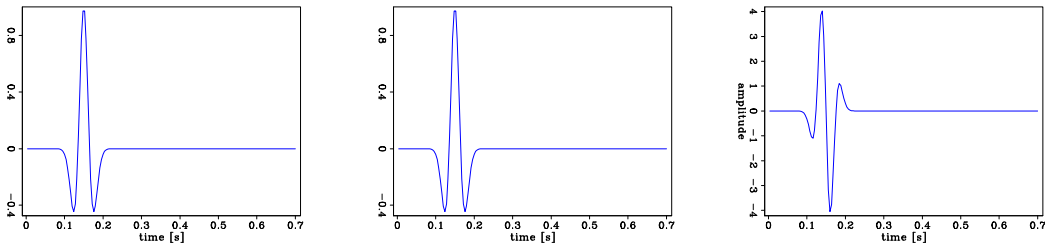


Figure 3: Left: x component of elastic source term. Center: z component of elastic source term. Right: inverted acoustic source. [**CR**] sfarris2/. waveletsCombined

The starting model for the FWI applications is a smoothed version of Figure 1. Bunks et al. (1995) showed an upscaling inversion is better at avoiding local minimum. Following this method, we will begin Each FWI scheme with a bandpass of 5 Hz and perform 40 conjugate gradient iterations. The result of this 5 Hz inversion will be the starting model for the next inversion which uses a 10 Hz bandpass. This scaling proceeds up to 25 Hz.

This is the framework we will use to compare our phase only acoustic FWI to traditional elastic and acoustic FWI.

ACKNOWLEDGMENTS

This investigation would be significantly delayed if not for all those who contributed to the nonlinear solver in Biondi and Barnier (2017). We would also like to thank Robert Clapp for creating the elastic model used in this report. We thank the affiliates of the Stanford Exploration Project for their continued support.

REFERENCES

Agudo, O. C., N. da Silva, M. Warner, and J. Morgan, 2016, Acoustic full-waveform inversion in an elastic world, *in* SEG Technical Program Expanded Abstracts 2016, 1058–1062, Society of Exploration Geophysicists.

Ben Veitch, James Rickett, and James Hobro, 2012, Imaging elastic media by corrections to acoustic propagation: Presented at the .

Biondi, E. and G. Barnier, 2017, A flexible out-of-core solver for linear/non-linear problems: SEP-Report, **168**.

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

Clapp, R., 2014, Synthetic model building using a simplied basin modeling approach: SEP-Report, **155**, 143–150.

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

Sirgue, L., O. I. Barkved, J. P. Van Gestel, O. J. Askim, and J. H. Kommedal, 2009, 3d waveform inversion on Valhall wide-azimuth OBC: Presented at the 71st EAGE Conference and Exhibition incorporating SPE EUROPEC 2009.

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

Warner, M., J. Morgan, A. Umpleby, I. Stekl, and L. Guasch, 2012, Which physics for full-wavefield seismic inversion?: Presented at the 74th EAGE Conference and Exhibition incorporating EUROPEC 2012.

Application of full Hessian on level set formulation and parameter inversion

Taylor Dahlke, Biondo Biondi, and Robert Clapp

ABSTRACT

Level sets are subsets of a domain that have the same value for a certain function. We can use them as a tool to update discrete boundaries of homogeneous bodies, which makes them particularly useful for updating salt models. Often, salt takes complicated geometries which causes a lack of direct illumination, as well as interactions between boundaries. Deriving a formulation of the Hessian which takes into account the level set parametrization should allow for better search directions than simpler methods. We find that by linearizing the velocity model perturbation with respect to the underlying background and level set parameters, we can derive a Hessian application operator suitable for a linear inversion scheme to get an improved search direction for updating the salt boundaries.

INTRODUCTION

Full Waveform Inversion (FWI) is generally used to update a continuous earth parameter like density or acoustic velocity. However, we often do not have enough frequency information to resolve features that have sharp boundaries, which is generally true for salt bodies in the Eastern Gulf of Mexico and other areas. It has been shown in previous work (see Santosa (1996), Lewis et al. (2012) and Burger (2003)) that level sets can provide a way to update these types of sharp boundaries.

However, this tool is meant to assist in our ability to invert for the entire earth model, which includes regions best approximated by continuous rather than discrete updates. This presents a problem of making updates to two domains; both an implicit surface that represents the sharp boundary, as well as the continuous background velocity field (as we do in traditional FWI). Previous work has demonstrated inversion schemes that update these domains sequentially Guo and de Hoop (2013), and also concurrently Dahlke et al. (2015) by finding global scaling variables for each gradient using line search methods. Both of these methods, while relatively inexpensive, are susceptible to local minima, and are ultimately first order updates that do not account for the interaction between model points, or the effects of acquisition.

By deriving a Hessian on a new model space that contains both level set (salt) and background (non-salt) parameters, we can better address the interaction between the salt and non-salt components of the model, and ultimately find a better update to these parameters.

We address this by first outlining the theory behind a model parametrization that contains both domains. Next we derive an application of the Hessian to this model. Afterwards, we describe the numerical application of this operator in an inversion scheme. Last we

show numerical results of using this Hessian operator in our inversion scheme, and compare against results from more simplistic approaches.

DERIVATION

The first step of this derivation is to describe the model space that we are working with. We will call our velocity model m which we define as:

$$m(\phi, b) = H(\phi)(c_{\text{salt}} - b) + b \tag{1}$$

where $H(\circ)$ is the Heaviside function, ϕ is the implicit surface, b is the background velocity model, and c_{salt} is the acoustic velocity of the salt. The dimensions of m, ϕ, and b are all NXxNZ since they exist over the same spatial domain (2D in this case). This means that the implicit surface is greater than zero in the salt body region, and less than zero outside of it. We generalize these parameters for the entire spatial domain (ignoring i, j), and expand this definition with a Taylor series as:

$$m_1 = m_0 + \frac{\partial m}{\partial \phi}\Big|_{m_0} \triangle\phi + \frac{\partial m}{\partial b}\Big|_{m_0} \triangle b + ...$$

By truncating this series and ignoring higher order terms, we can create a linear approximation for the perturbation of the velocity model m with respect to ϕ and b:

$$m_1 - m_0 = \triangle m \approx \frac{\partial m(\phi_o, b_o)}{\partial \phi}\triangle\phi + \frac{\partial m(\phi_o, b_o)}{\partial b}\triangle b. \tag{2}$$

This can be written as a matrix operation:

$$\triangle m \approx \begin{bmatrix} \frac{\partial m(\phi_o, b_o)}{\partial \phi} & \frac{\partial m(\phi_o, b_o)}{\partial b} \end{bmatrix} \begin{bmatrix} \triangle\phi \\ \triangle b \end{bmatrix}.$$

When we define $\triangle p$ as

$$\triangle p = \begin{bmatrix} \triangle\phi \\ \triangle b \end{bmatrix},$$

this gives us:

$$\triangle m \approx \begin{bmatrix} \frac{\partial m(\phi_o, b_o)}{\partial \phi} & \frac{\partial m(\phi_o, b_o)}{\partial b} \end{bmatrix} \triangle p.$$

Where we define operator D as:

$$\begin{aligned} D &= \begin{bmatrix} \frac{\partial m(\phi_o, b_o)}{\partial \phi} & \frac{\partial m(\phi_o, b_o)}{\partial b} \end{bmatrix} \\ &= \begin{bmatrix} \delta(\phi)(c_s - b) & 1 - H(\phi). \end{bmatrix} \end{aligned} \tag{3}$$

This operator D ultimately scales and masks the parameter fields $\triangle\phi$ and $\triangle b$. With this new approximation of the perturbation in our velocity model, the application of our Born operator (B) to our new model parameter space $\triangle p = [\triangle\phi \quad \triangle b]^T$ is:

$$\triangle d = B\triangle m$$
$$\triangle d \approx BD\triangle p.$$

Alternatively, we can find the gradient for our model parameters by applying the adjoint operations:

$$\triangle p \approx D^T B^T \triangle d.$$

Similarly we can find the application of the Hessian of the FWI objective function as:

$$D^T H D\triangle p \approx -D^T B^T \triangle d. \tag{4}$$

In equation 4, we can substitute H with either the full or Gauss Newton Hessian. Previous work by Fichtner (2010) shows that the full Hessian of the FWI objective function can be constructed by summing a WEMVA component with the Gauss-Newton component of the Hessian. It is this formulation of the full Hessian application that we use. The method I propose is to solve equation 4 for $\triangle p$ using a conjugate gradient inversion method.

Previous work (Dahlke et al. (2016)) showed a similar technique, with the transformation to $\triangle p$ being done after an inversion for the $\triangle m$ search direction is completed (without the "constraint" of projection operator D). This means instead of solving 4, we solve:

$$H\triangle m = -B^T \triangle d, \tag{5}$$

followed by:

$$\triangle p = D^T \triangle m. \tag{6}$$

The method using equations 5 and 6 will be referred to as unconstrained inversion, while the proposed method (the solution of equation 4) is constrained by D. We compare the two methods later in this paper.

NUMERICAL APPLICATION

Inside the D operator we have both a Heaviside function $H(\cdot)$ as well as a Dirac-delta function $\delta(\cdot)$. While computing the application of the Heaviside function on our implicit surface ϕ is relatively simple for a discrete case (masking with a threshold), the same cannot be said for the application of $\delta(\cdot)$ to our numerical problem. Most of our boundary is not explicitly represented as points where $\phi = 0$, but simply by the adjacency of positive and negative ϕ values. For this reason, we relax the definition of our operator to use an approximation of the Dirac-delta function. Previous derivations using calculus of variations (Santosa (1996), Dahlke et al. (2015)) have demonstrated that we can reduce our FWI objective function using an update for ϕ defined as:

$$\triangle\phi = |\nabla\phi|(c_{\text{salt}} - b)B^T\triangle d. \tag{7}$$

This is similar to the update we've derived using the linearization of the Taylor series expansion:

$$\triangle\phi = \delta(\phi)(c_{\text{salt}} - b)B^T\triangle d. \tag{8}$$

For this reason, we substitute $|\nabla\phi|$ for $\delta(\phi)$, as it suffices as an approximation of the Dirac-delta function when we keep the implicit surface slope regularized as we update it. In our case, this means using a double-well potential function fomulation of Distance Regularized Level Set Evolution (DRLSE) as described in Li et al. (2010). This step provides a forcing term to help keep the slopes of the implicit surface at either zero or one. When that is maintained, the only place where $|\nabla\phi| > 0$ is immediately surrounding the boundary. Using this substitution, the D operator we apply is:

$$D = \begin{bmatrix} |\nabla\phi|(c_{\text{salt}} - b) & .1 - H(\phi) \end{bmatrix} \tag{9}$$

RESULTS

To test the effectiveness of this linearization, I pose the following question: can I get a first-step update using this new method that is better than other methods? To answer this, I choose to find search directions using the proposed method as well as two other approaches, and then use a line search to find the step-length α for each search direction. These three search directions methods are steepest-descent, unconstrained inversion, and the proposed constrained inversion. I then do a model update and compare the model space residual to see which update performed better.

The first of the two comparison methods is simply steepest descent, where the search direction is my gradient: $\triangle p = D^T B^T \triangle d$. The second method is a little more subtle. With the second method, I use the application of the full Hessian as described in Fichtner (2010) and invert for $\triangle m$, followed by the application of D as described in equations 5 and 6.

To perform the inversion for the search direction in either case, we made use of the out-of-core solver developed by Biondi and Barnier (2017). We implemented a conjugate-gradient (CG) approach to take advantange of the symmetry of the Hessian operator (as opposed to conjugate gradient least-squares (CGLS), which doesn't require a symmetric operator). By using this method we were able to invert at the cost of one forward operation per iteration rather than both the forward and adjoint operations required by CGLS. One further requirement of using CG is that the operator must be positive definite, while for CGLS it need not be. For the Gauss-Newton Hessian we can use either safely, but with the full Hessian, we risk a failure in our inversion, depending on the velocity model that we are using. For the models we use in this paper, we found the CG inversion to be stable, so it was used instead of CGLS since it performs at half the cost. When CG is used, the objective function value will not necessarily converge to zero since it isn't a least squares formulation like CGLS (see Aster et al. (2011) for more). For all cases shown, we start at an objective function value of zero since we begin the inversion with a zero initial model.

We perform this comparison on two different velocity model examples. Both examples have the same true model, but have different initial model states (see Figure 1). In case one,

Table 1: Case one: Perturbed base of salt

Update method	Model residual norm
Steepest descent	8.617
Hessian inversion (unconstrained)	16.658
Hessian inversion (proposed)	1.863

the base of the circular salt is perturbed downwards. In case two, the top of the bottom reflector salt is perturbed upwards.

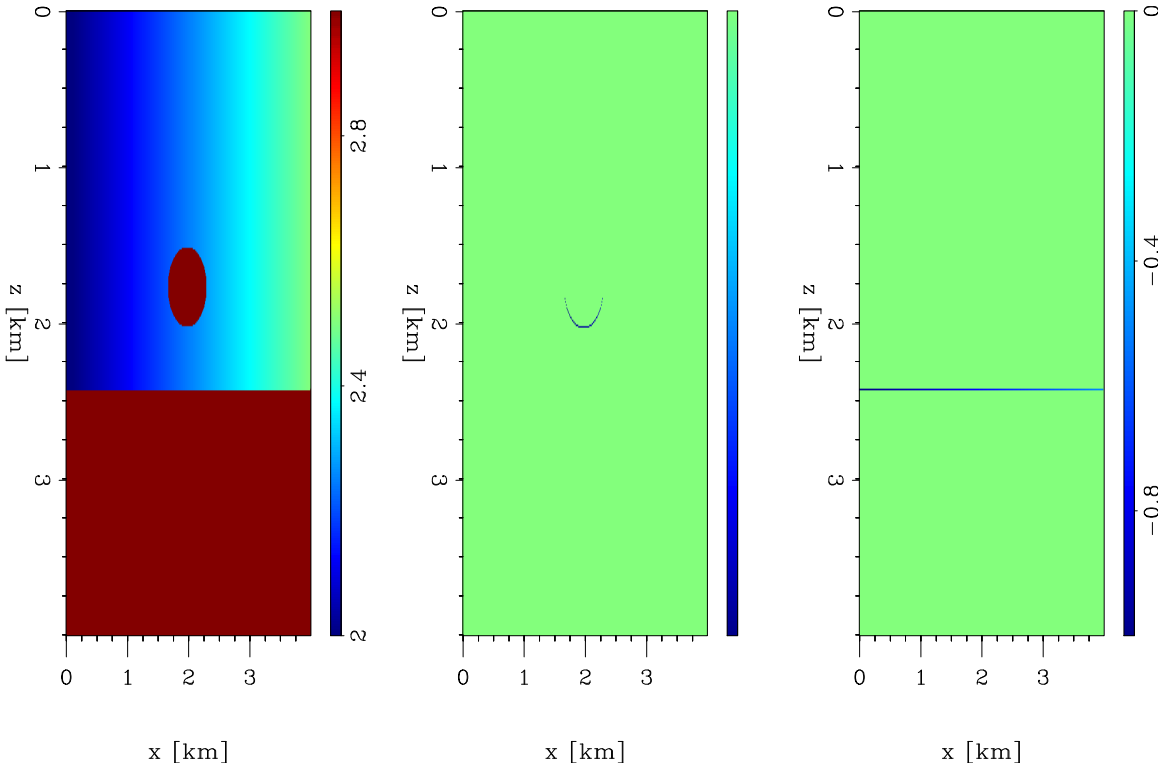


Figure 1: True model (left); the difference between the true and inital model for case one (middle); the difference between the true and inital model for case two (right). [**ER**] taylor/. velocity-model-ball

For the first case, we find a significant improvement by using the proposed method when compared to the other two methods. We can see from the search directions that the new Hessian inversion result (Figure 3, right) has a more correct focusing of energy on the base of salt rather than on the bottom reflector as with the other two methods (Figure 3 left, middle). This leads to a better update when we perform a line search on these search directions (Figure 4). When we difference the updated model with the true model and take the L2 norm of the residual, we see that the proposed method has an improved update when compared to either of the other approaches (Table 1).

One further question that we have is the effect of either the masking or scaling com-

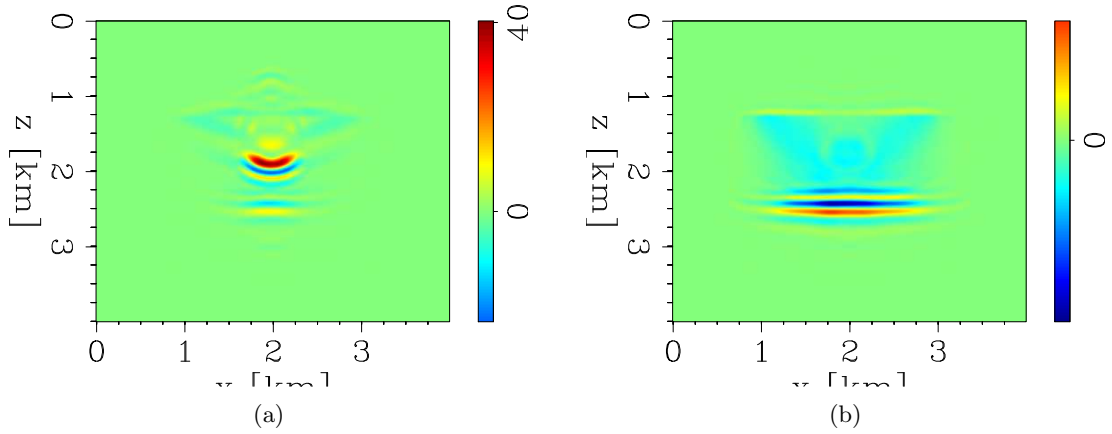


Figure 2: Adjoint Born images for (a) perturbed salt base case; (b) perturbed bottom reflector case; no clipping of amplitudes. [**CR**] taylor/. rtm-image-ball,rtm-image-ball2

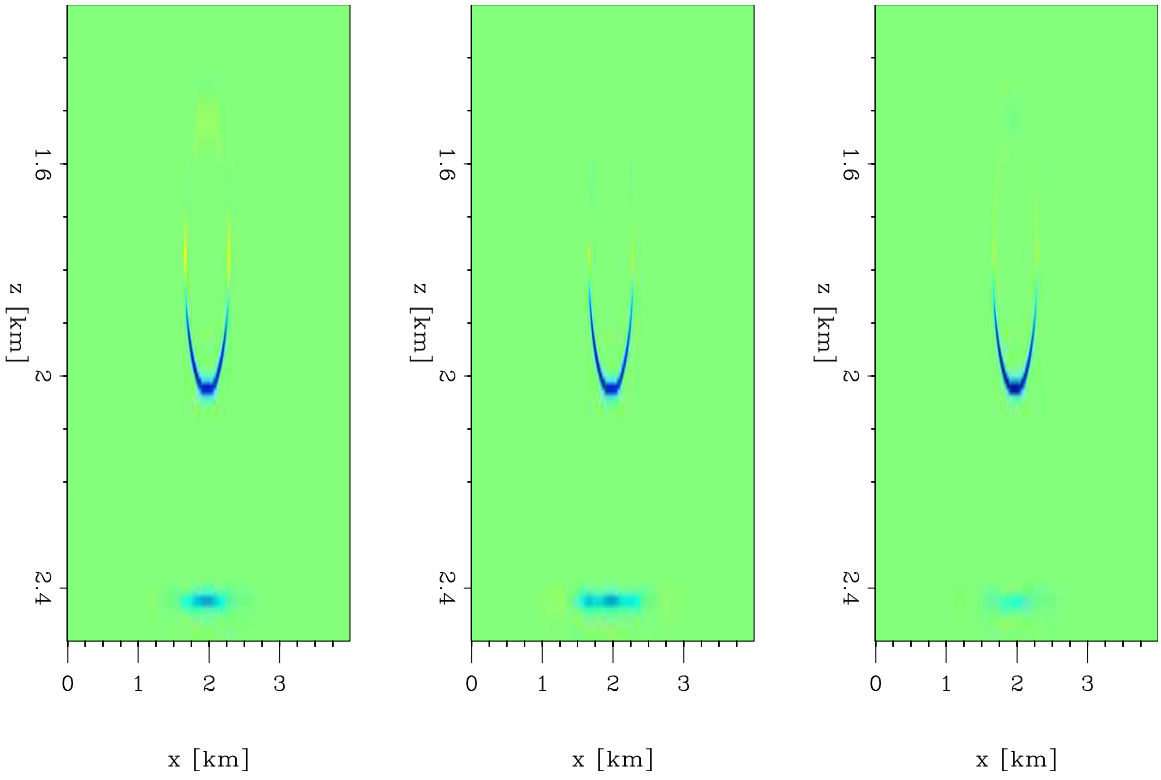


Figure 3: CASE 1: Search directions for the ϕ update resulting from Steepest descent approach (left), unconstrained Hessian inversion (middle), and proposed Hessian inversion scheme (right). [**CR**] taylor/. compare-searchdirections-ball-PHI

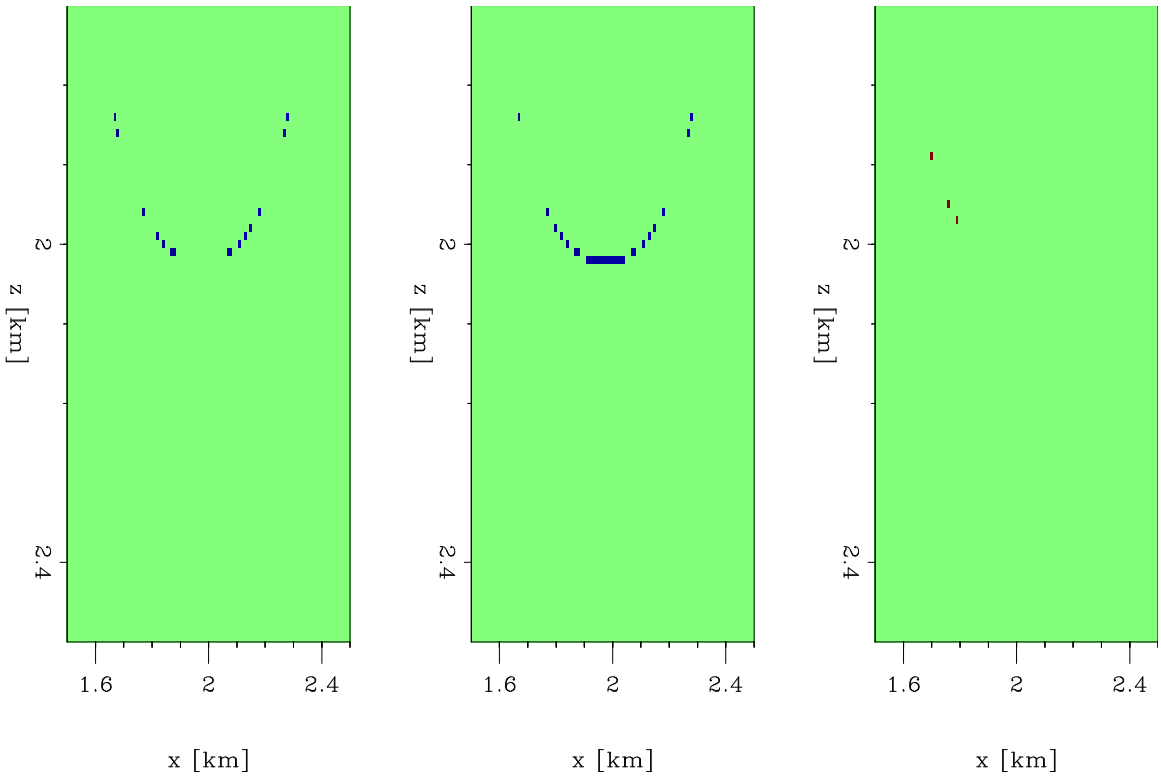


Figure 4: CASE 1: Difference between true model and velocity model applying first update to ϕ using a line search method to find step size α. (Left) based on steepest descent, (middle) based on unconstrained Hessian inversion, (right) based on proposed Hessian inversion incorporating projection operator D as a constraint. [**CR**] taylor/. compare-firststep-ball-PHI

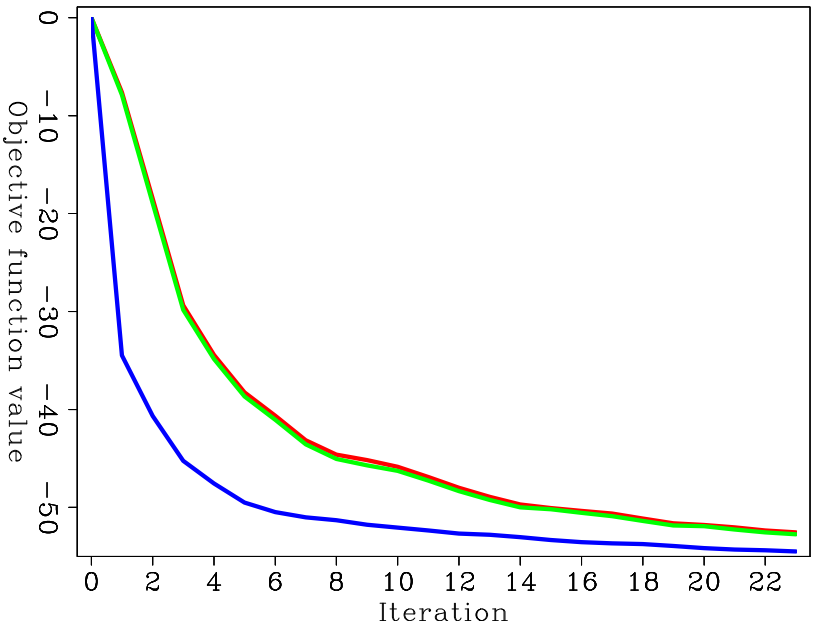


Figure 5: CASE 1: Objective function decrease for the three search direction inversions; unconstrained search direction inversion with no D operator (blue), Proposed method (red), and proposed method with scaling term equal to one (green). [**CR**] taylor/. compareobjfuncBALL

ponents inside the D operator. Which of these makes the most difference? In both cases, we have an unperturbed background velocity model where the scaling term inside D varies spatially (note the horizontal gradient in Figure 1, left), so the scaling term $(c_{\text{salt}} - b)$ inside D is non-trivial. To test this, we perform our proposed method as before, with the exception of the scaling term being equal to one across the whole spatial domain, eliminating its effect. We find that the search direction that results from the inversion is similar to the result from the originally proposed definition of D (Figure 6) in terms of amplitude distribution. This similarity is also reflected in the objective function descent (Figure 5), where the two curves are nearly identical. We do find that the lack of the scaling term increases the magnitude of $\triangle\phi$ in the correct direction by about 16% in the same areas that the original approach tries to update. This improved search direction found by ignoring the scaling requires further investigation.

For the second case with the perturbed bottom reflector, we find that we still have an improved result, but slightly less obvious. When we compare the objective function decrease for the two inversion approaches (Figure 9), we can see that these curves are far more similar than in the case one example (Figure 5). The steepest descent result (Figure 7, left) correctly has most of its energy on the bottom reflector rather than on the circular salt as the other two inversion methods do. However, the inverison methods both correctly update a larger portion of the perturbation, with the proposed method doing so with slightly less energy incorrectly applied to the circular salt above. This leads us to get an improved update after line search (Figure 8). Looking at the model residual norm again, we find that the proposed method again performs better, in this case with less of a margin than before (see Table 2).

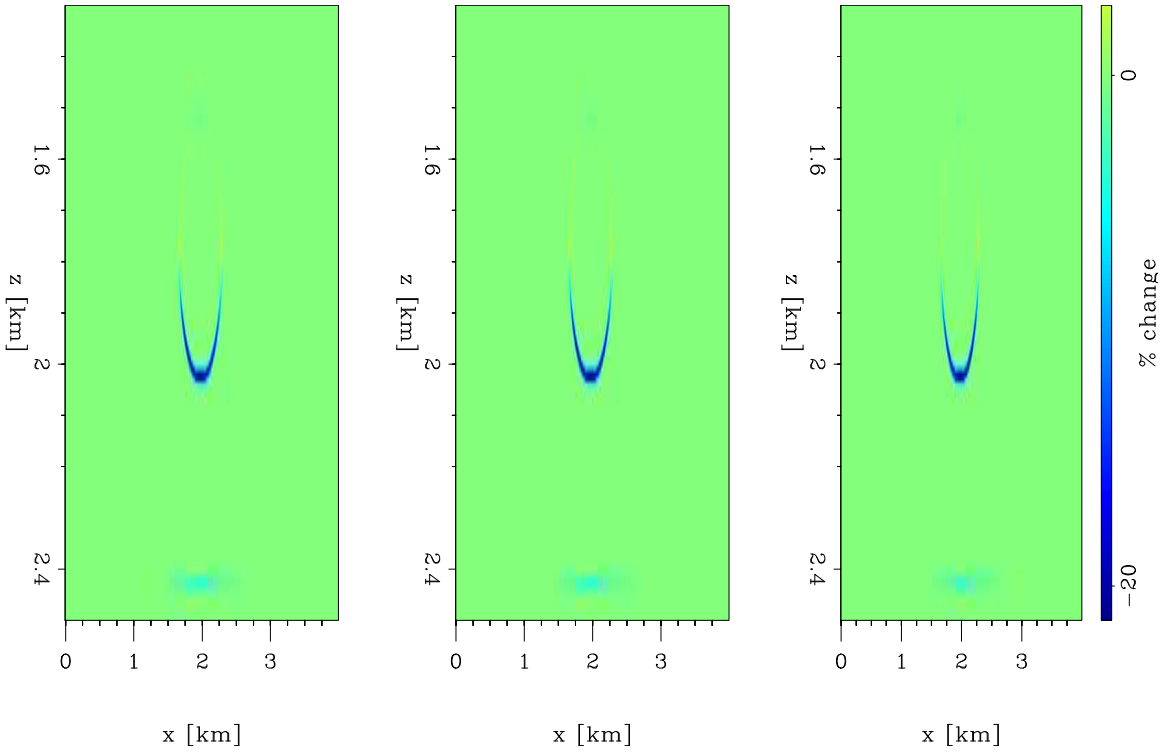


Figure 6: Search direction of proposed method (left), and proposed method without scaling component (middle). Percent difference between (left) and (middle) search directions (right). [**CR**] taylor/. testscaleeffect

Table 2: Case two: Perturbed bottom reflector

Update method	Model residual norm
Steepest descent	172.683
Hessian inversion (unconstrained)	144.088
Hessian inversion (proposed)	135.161

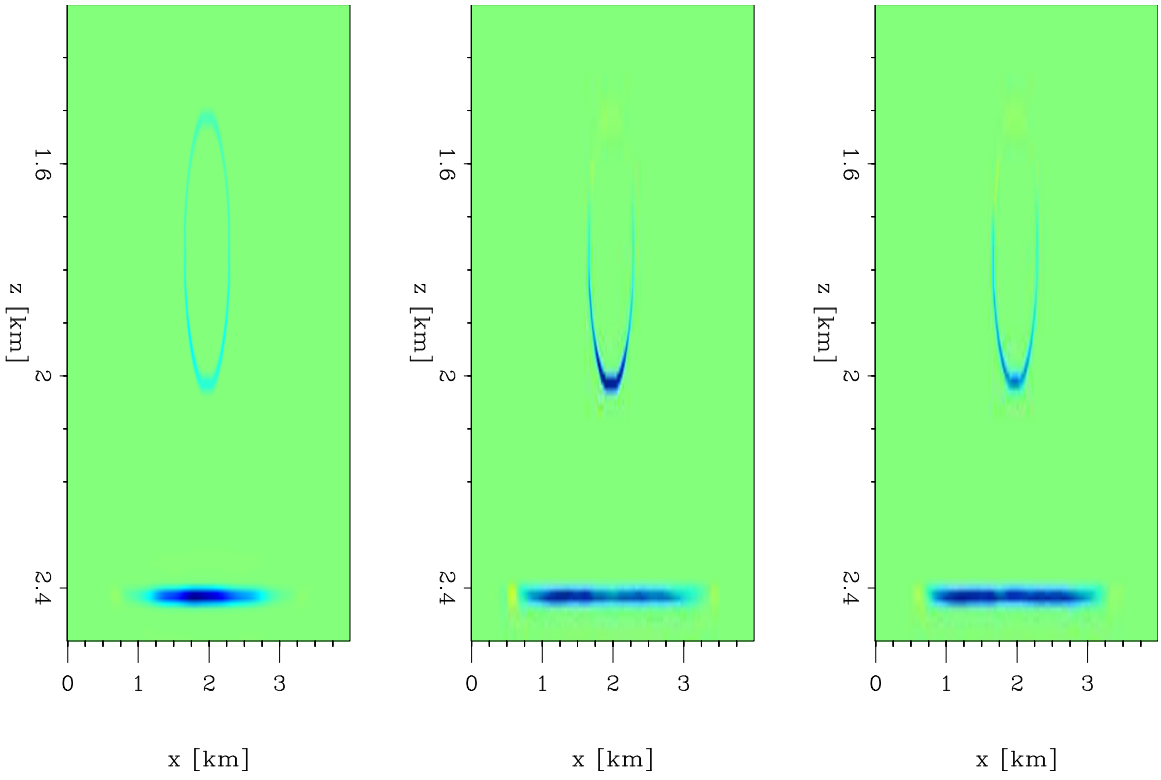


Figure 7: CASE 2: Search directions resulting from steepest descent approach based on gradient (left), from unconstrained Hessian inversion (middle) , and from proposed Hessian inversion scheme (right). [**CR**] taylor/. compare-searchdirections-ball2-PHI

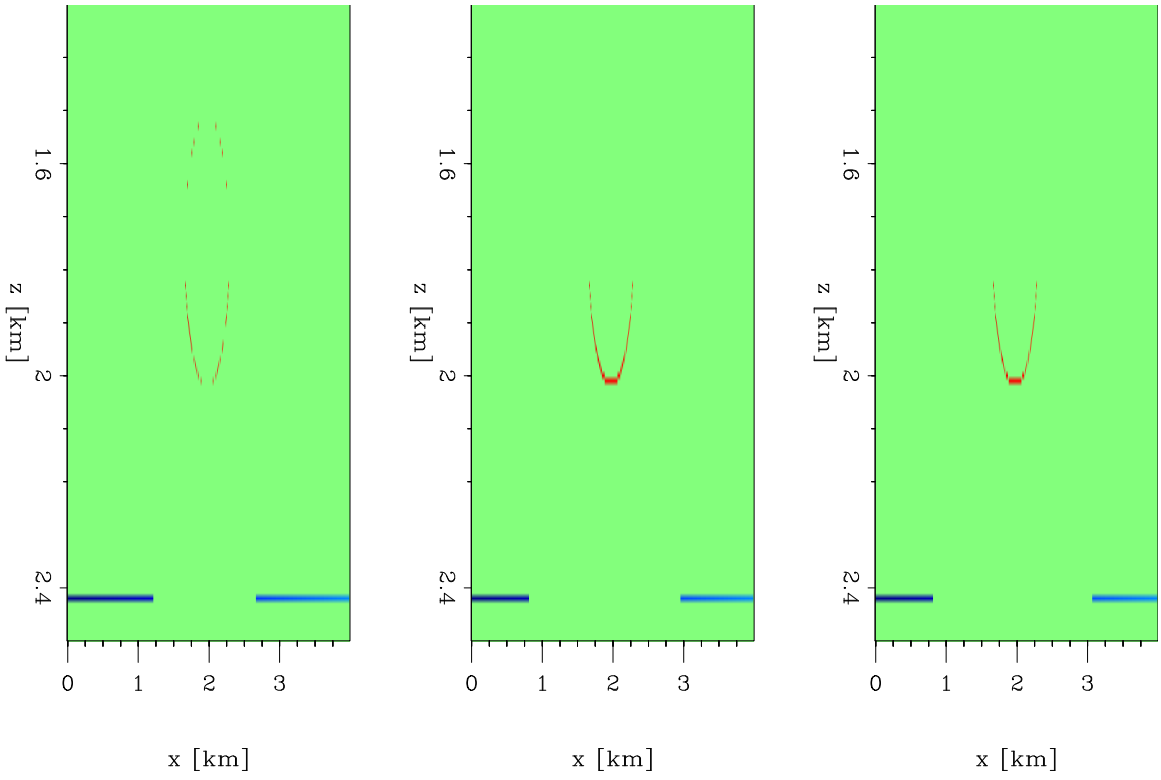


Figure 8: CASE 2: Difference between true model and velocity model for perturbed bottom reflector case after applying first update to ϕ using a line search method to find step size α. (Left) based on steepest descent, (middle) based on unconstrained Hessian inversion, (right) based on new Hessian inversion incorporating D operator directly. [**CR**] taylor/. compare-firststep-ball2-PHI

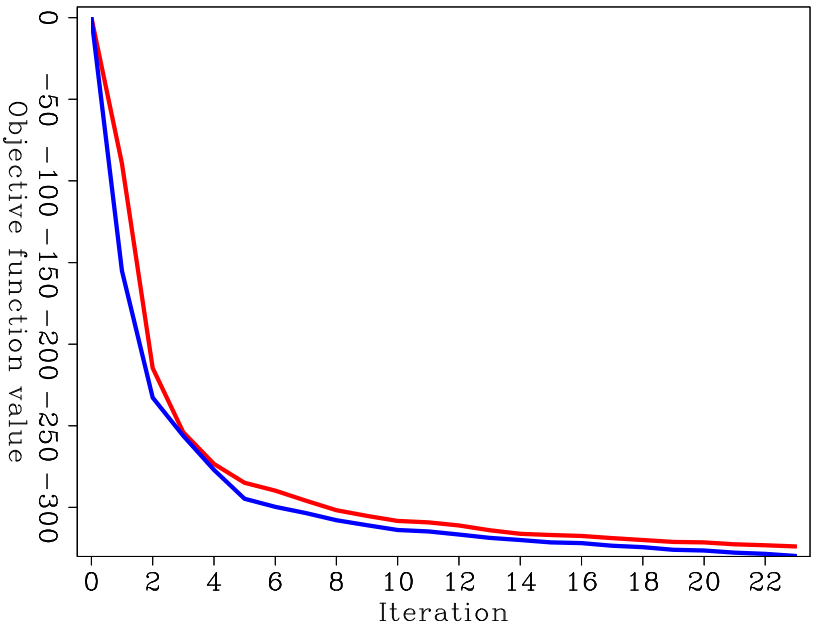


Figure 9: CASE 2: Objective function decrease for the two search direction inversions; Unconstrained search direction inversion (blue) and proposed method (red) using D operator. [**CR**] taylor/. compareobjfuncBALL2

CONCLUSIONS

The results of the two simple model cases show that we are able to get an improved search direction for the salt body boundary movement when we invert for our new model parameter updates $\triangle p$ using the new linearized derivative operator D defined earlier. We also show that the while the masking component seems to make the most significant difference when compared to the other two methods, there remains more to be learned about the significance of the scaling component of the D operator.

REFERENCES

Aster, R., B. Borchers, and C. Thurber, 2011, Parameter estimation and inverse problems. International Geophysics: Elsevier Science.

Biondi, E. and G. Barnier, 2017, A flexible out-of-core solver for linear/non-linear problems: SEP-Report, **168**.

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

Dahlke, T., B. Biondi, and R. Clapp, 2015, Domain decomposition of level set updates for salt segmentation, 1366–1371.

Dahlke, T., R. Clapp, and B. Biondi, 2016, Second-order updating in shape optimization for salt segmentation, 5369–5373.

Fichtner, A., 2010, Full seismic waveform modelling and inversion. Advances in Geophysical and Environmental Mechanics and Mathematics: Springer Berlin Heidelberg.

Guo, Z. and M. de Hoop, 2013, Shape optimization and level set method in full waveform

inversion with 3d body reconstruction: SEG Technical Program Expanded Abstracts, 1079–1083.

Lewis, W., B. Starr, D. Vigh, et al., 2012, A level set approach to salt geometry inversion in full-waveform inversion: Presented at the 2012 SEG Annual Meeting.

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

Santosa, F., 1996, A level-set approach for inverse problems involving obstacles: ESAIM Controle Optim. Calc. Var, **1**, 17–33.

Tomographic Full Waveform Inversion for imaging under complex overburden

Guillaume Barnier, Ettore Biondi, and Biondo Biondi

ABSTRACT

Tomographic Full Waveform Inversion (TFWI) is a waveform inversion technique that uses an extended modeling operator, allowing the predicted data to be in phase with the observed data, even when the starting velocity model is inaccurate. While being robust against cycle-skipping issues inherent to Full Waveform Inversion (FWI), it is still poorly understood whether this technique can be efficient in a complex overburden environment. And if so, which type of model extension is the most adequate to predict and capture the complexity of the waveforms such as the ones reflecting off a top-salt boundary. We first present three synthetic examples that illustrate the issues encountered in complex overburden environments. Then, we review the theoretical framework of TFWI and present our code that allows for space and time-lag extensions. Finally, we test for various model extensions to understand which combination manages to better capture the complexity of a wavefield that is reflected from a top-salt boundary.

INTRODUCTION

Producing accurate and high-resolution images of subsalt reflectors is a key challenge for the oil and gas industry. Companies are still looking for ways to reduce processing cycle-time and third-party costs associated with salt-model building. As shown in previous work, there are two major difficulties associated with subsalt imaging. On one hand, the velocity contrast between the salt rock and the surrounding sediments is so high that a slight misinterpretation can create huge time-shifts in the computation of wavefields, which in turn can translate into a total loss of coherency of migrated images (Etgen et al., 2014a,b; Barnier and Biondi, 2015). On the other hand, salt bodies tend to have rugose and complex boundaries (e.g., carapaces and cap rocks) that generate complicated waveforms typically coming from prismatic waves, multiple scattering, and point diffractors. Therefore, predicted data generated from an incorrect salt model will not match the complex recorded waveforms, and standard algorithms will fail to update the salt model efficiently.

To address these issues, we propose to apply the Tomographic Full Waveform Inversion (TFWI) workflow developed by Biondi and Almomin (2014) to improve images of subsalt reflectors. This technique has been shown to be efficient and immune to cycle-skipping in relatively simple overburden environments (Almomin, 2016). The modeling is based on an extended Born operator that allows the user to linearly predict data while ensuring that they stay in phase with the observed ones, even when the starting model is "far" from the truth. The non-physical energy in the extended axis is slowly reduced during the iterative process until an optimal and "physical" model is found. Since the algorithm requires the

energy in the extended axis to be somewhat coherent in order to converge towards the right solution, it is still poorly understood whether it may be successfully applied to subsalt imaging problems (as energy becomes incoherent even in the extended axis). Another problem may arise from the fact that salt boundaries generate such complicated events that even an extended linear modeling operator might have difficulties replicating them.

In this report, we first show three synthetic examples of FWI that illustrate the limitation of this algorithm when applied to more complex overburdens (e.g., models which contain sharp and high-velocity contrast anomalies). We provide a brief theoretical review of TFWI, we present the structure of our code that allows both subsurface offsets (vertical and horizontal) and time-lag extensions. We then test for different combinations of extensions to understand which one is more likely to be adequate for replicating waveforms generated at a salt-sediment interface. Finally, we apply TFWI on a synthetic example that contains an anomaly with a sharp velocity contrast. All optimization and inversion results in this reports were performed using the optimization library developed in Biondi and Barnier (2017).

LIMITATION OF FULL WAVEFORM INVERSION FOR IMAGING UNDER COMPLEX OVERBURDEN

FWI has been proved to be a very useful and accurate tool for velocity model building in relatively simple overburden environments (Shen, 2014). However, its robustness in the context of subsalt imaging is still an active research topic for the oil and gas industry (Williamson et al., 2016; Wang et al., 2015; Lewis et al., 2012). One of the main challenges is that in order to obtain accurate subsalt images, the sharpness of salt boundaries must be captured by our velocity model. Slight errors in their delineation may result in big distortions of the images (Etgen et al., 2014a,b). Moreover, obtaining higher resolution subsalt images requires the use of higher frequency energy, in which case the wavefields will be increasingly more sensitive to small inaccuracies at the salt-sediment interface. Beyond certain frequencies, the extreme velocity contrast at the interface and its rugosity (e.g., canyons and small-scale anomalies) may generate complicated waveforms (e.g., prismatic waves) which are difficult to predict and replicate with our standard linearized modeling operators (based on single scattering assumption). Failing to capture both the sharpness and the shape of the salt interface may cause convergence issues and FWI may fail to produce accurate velocity models.

The first tests we conduct are designed to assess how well we can recover the sharpness of an abrupt boundary in a simple synthetic setup based on Mora (1989), where we introduce a circular homogeneous anomaly in a background containing two horizontal layers. We show that if the velocity contrast of the anomaly is not too large, its sharpness and its shape can be well recovered. However, if the contrast is too large, FWI converges towards an incorrect solution, even for an elementary example (synthetic, noise-free and using very low frequency energy). Another issue for imaging in the vicinity of complex salt bodies arises from the fact that it is also still unclear how to choose and construct an efficient starting velocity model for FWI. On one hand, starting with a model containing a salt body with very sharp (but inaccurate) boundaries may be inefficient for FWI convergence because it would take many iterations to "move" and re-adjust these sharp boundaries (i.e., the high-wavenumber components of our model) to the correct location. On the other hand, a

starting velocity model with "not enough" salt would also be inefficient since the algorithm would spend a lot of iterations building the high velocity contrast between the salt and its surrounding sediments. Therefore, in our third example, we try to illustrate this issue with a synthetic salt basin environment that replicates a typical setup we may encounter in the Gulf of Mexico (GoM). We start with a smoothed version of the true velocity model, and we show that even with very long-offset data (containing diving waves and refracted energy off the top-salt) and very low frequency energy, the algorithm does not converge towards the correct solution.

Recovering a sharp boundary with a small velocity contrast

We apply FWI on a simple synthetic model which consists of one circular anomaly embedded into a two-layer velocity model shown in Figure 1(g), similar to the one proposed in Mora (1989). The velocity contrast between the anomaly and the top-layer is only 10 %, the shape of the anomaly is simple (circular), and data are modeled and inverted using the same acoustic engine. We perform a multi-scale FWI (Bunks et al., 1995), starting from a frequency band of 0.25-7 Hz, up to 0.25-45 Hz. For each frequency band, we perform 40 iterations of FWI. Though not realistic, this test provides an idea of how well we are able to recover a sharp boundary in a "best case scenario" (i.e., with energy at very low frequencies and noise-free data). In our first test, the initial model is the true model without the anomaly. We investigate how the sharp boundary evolves as we increase frequency and if the algorithm converges towards the true solution. We use a finite-difference grid size of $dz = dx = 10$ m. We place sources every 100 m, and receivers every 10 m at the surface. Results are shown in Figures 1(a)-(f), and we can see a good convergence as we increase the frequency content of our data. We extract a 1D velocity profile at $x = 1$ km, and display the results in Figures 2(a)-(f). We stop the inversion process at 45 Hz to avoid numerical dispersion but the inversion could be pursued with higher frequencies and a smaller finite-difference grid size. The interface between the anomaly and the surrounding background is well recovered, which shows that obtaining such sharpness with FWI is theoretically possible.

Figures 3(a)-(c) show the results of a similar multi-scale FWI inversion scheme, but using a constant velocity background (only the top-layer) as initial model. In this case, we observe (1) a decrease in the sharpness of the anomaly's boundaries (Figure 3(a)) and (2) an ambiguity in the velocity of the first and second layers, especially close to their interface (Figure 3(b)-(c)).

Recovering a sharp boundary with a large velocity contrast

We conduct a similar experiment but the velocity contrast between the anomaly and the top-layer is now 125 % (top-layer velocity is 2 km/s and anomaly velocity is 4.5 km/s), which is a typical contrast one can expect when exploring salt basins. Even though we start from a frequency band containing (unrealistic) low-frequency energy (0.25-5 Hz), the algorithm does not converge towards the correct solution. As in the previous example, we conduct a multi-scale FWI scheme up to 45 Hz. Finite-difference parameters are the same as in the previous example. Results are shown in Figures 4 and 5. For each frequency

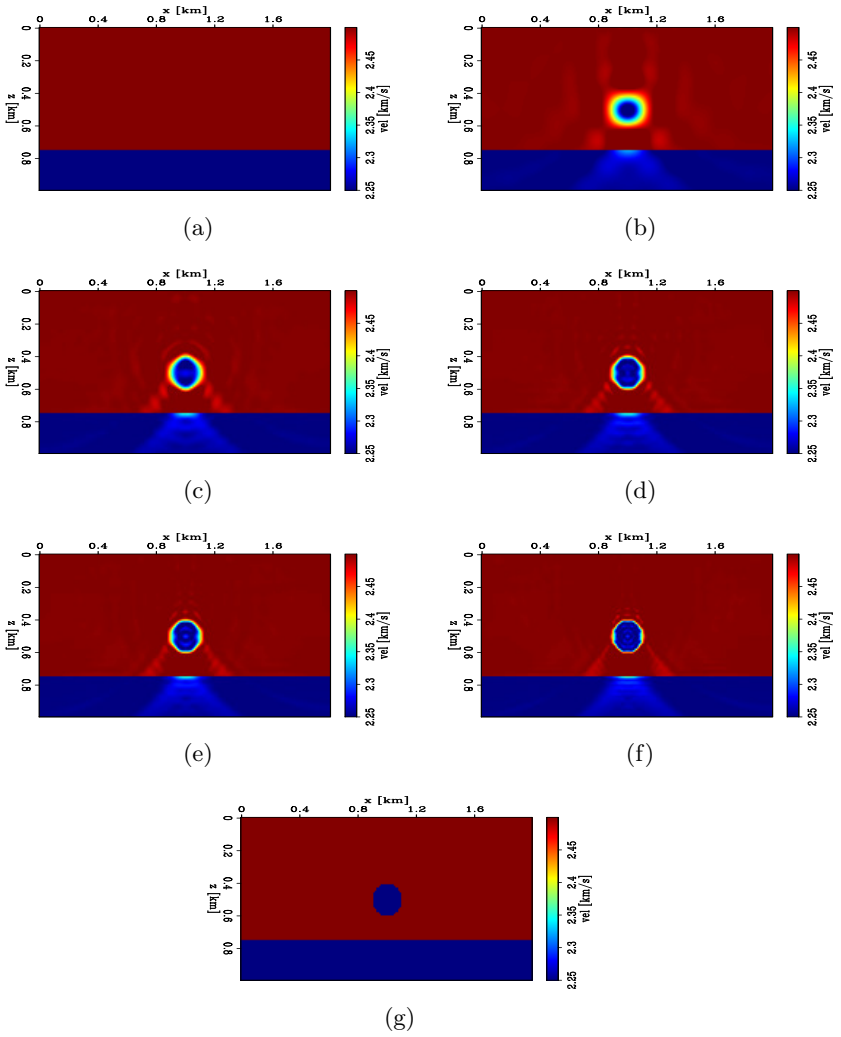


Figure 1: Inverted velocity models using a multi-scale FWI scheme. For each frequency band, we perform 40 iterations of FWI. (a) Initial model. (b) Inverted model using energy up to 7 Hz. (c) Inverted model using energy up to 15 Hz. (d) Inverted model using energy up to 25 Hz. (e) Inverted model using energy up to 30 Hz. (f) Inverted model using energy up to 45 Hz. (g) True model. [**CR**] guillaume1/. mora.vel.start,fwi.mora.vel.7,vel.15,vel.25,vel.30,vel.45,mora.vel

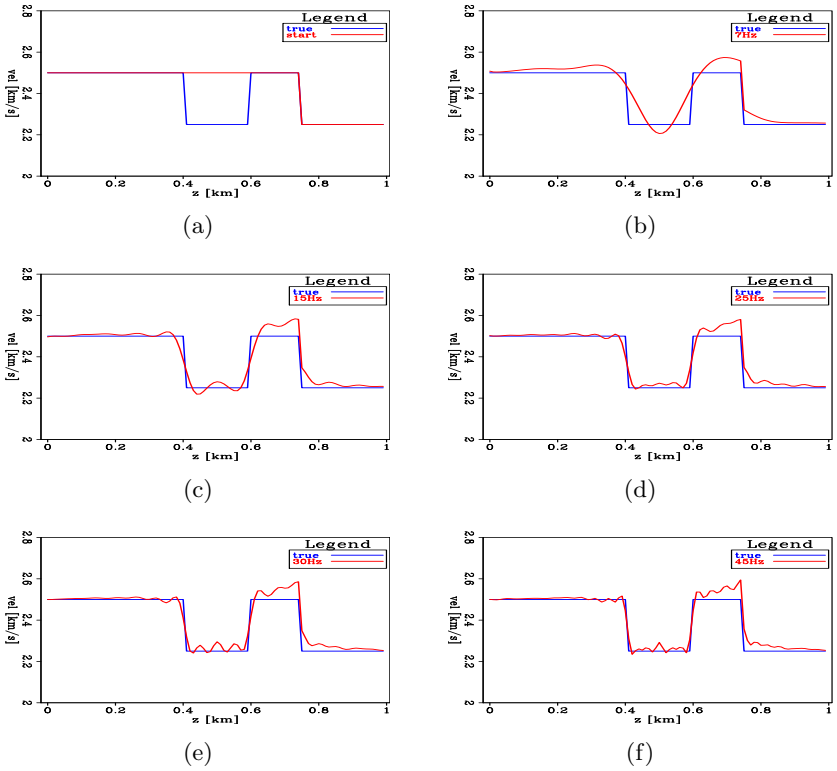


Figure 2: 1D velocity profile of inverted models using a multi-scale FWI scheme extracted at $x = 1$ km. For each frequency band, we perform 40 iterations of FWI. (a) Initial model. (b) Inverted model using energy up to 7 Hz. (c) Inverted model using energy up to 15 Hz. (d) Inverted model using energy up to 25 Hz. (e) Inverted model using energy up to 30 Hz. (f) Inverted model using energy up to 45 Hz. [**CR**] guillaume1/. fwi.mora.vel.start.1d,vel.7.1d,vel.15.1d,vel.25.1d,vel.30.1d,vel.45.1d

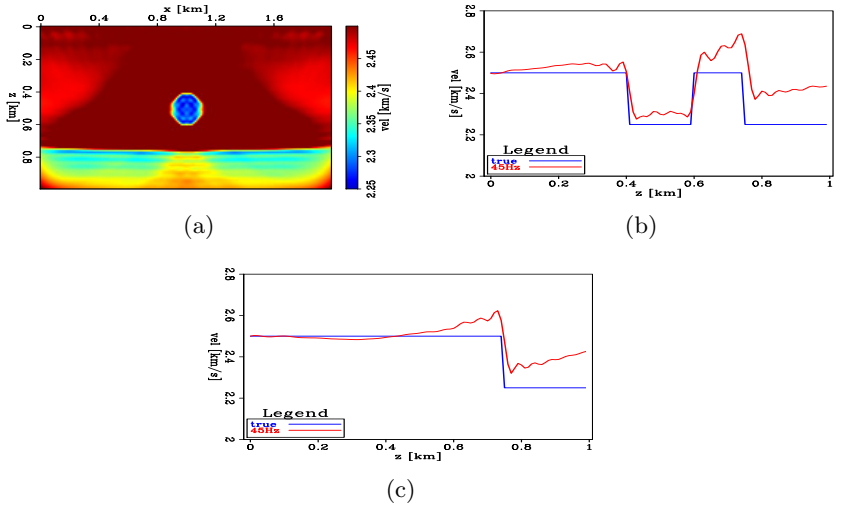


Figure 3: Inverted model with a multi-scale FWI scheme using energy up to 45 Hz and a constant initial velocity model. (a) Full 2D model. (b) 1D profile extracted at $x = 1$ km. (c) 1D profile extracted at $x = 0.4$ km [**CR**] guillaume1/. fwi.mora.ref.vel.40,fwi.mora.ref.vel.45.1d,fwi.mora.ref.vel.45.1d.2

band, we run 40 iterations of FWI. We can see that the algorithm fails to recover the strong anomaly and its sharp boundaries, even for such a simple setup.

FWI in the presence of a complex salt body

We apply FWI on a synthetic model that replicates a salt basin environment typically encountered in the GoM. We design our velocity model by combining a sediment model generated using the workflow from Clapp (2014) and by adding a complex salt body (Figure 6(a)). We simulate a long-offset acquisition (20 km), with shots (every 100 m) and receivers (every 10 m) at the surface. Our initial model which was obtained by smoothing the true velocity model, is shown in Figures 6(b)-(c). Figure 7 shows the shot record for a source positioned at $x = 0$ km, containing energy ranging from 0.25 Hz up to 3.5 Hz. In Figure 7, events arriving before the direct arrivals for longer offsets ($x \geq 10$ km) correspond to the refracted energy (traveling almost horizontally through the salt) caused by the high velocity contrast between the salt body and its surrounding sediments. Since the salt body has a rugose top, part of that refracted energy is also scattered and recorded at the surface. We purposely choose to include (unrealistic) low frequencies in our inversion scheme to assess whether the FWI algorithm will able to recover the boundaries of the salt body in the "ideal" case, or at least provide an improved velocity model. We run 30 iterations of FWI and display the results in Figures 8, 9, and 10. Data residuals at various iterations for the shot record generated by a source positioned at $x = 0$ km are shown in Figure 11. Results indicate that most of the updates occur in the shallowest part of the model, and are very small in amplitude (Figure 10(a)-(d)). Even though the top-salt boundary seems be somewhat correctly updated in some areas, we are already starting to see wrong updates just above the salt body (incorrect low velocity zones), which may indicate possible

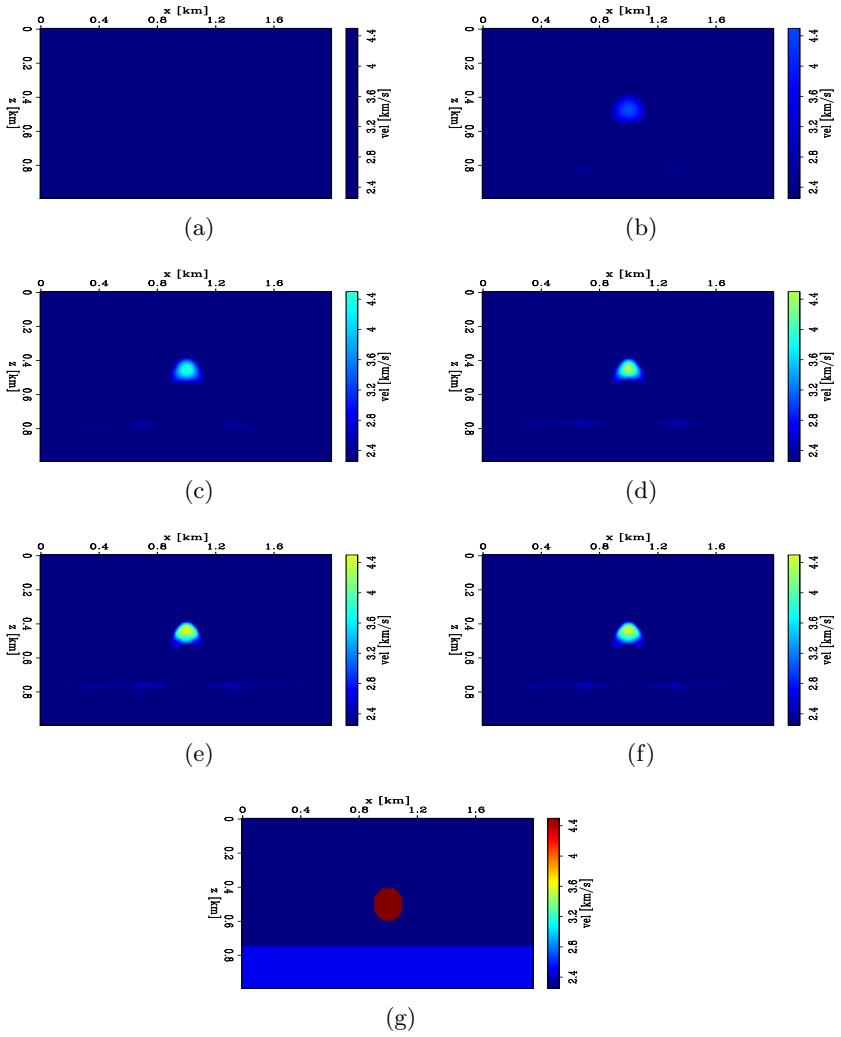


Figure 4: Inverted velocity models using a multi-scale FWI scheme. For each frequency band, we perform 40 iterations of FWI. (a) Initial model. (b) Inverted model using energy up to 5 Hz. (c) Inverted model using energy up to 12 Hz. (d) Inverted model using energy up to 20 Hz. (e) Inverted model using energy up to 30 Hz. (f) Inverted model using energy up to 45 Hz. (g) True model. [**CR**] guillaume1/. mora.contrast.start,cont.5,cont.12,cont.20,cont.30,cont.45,mora.contrast

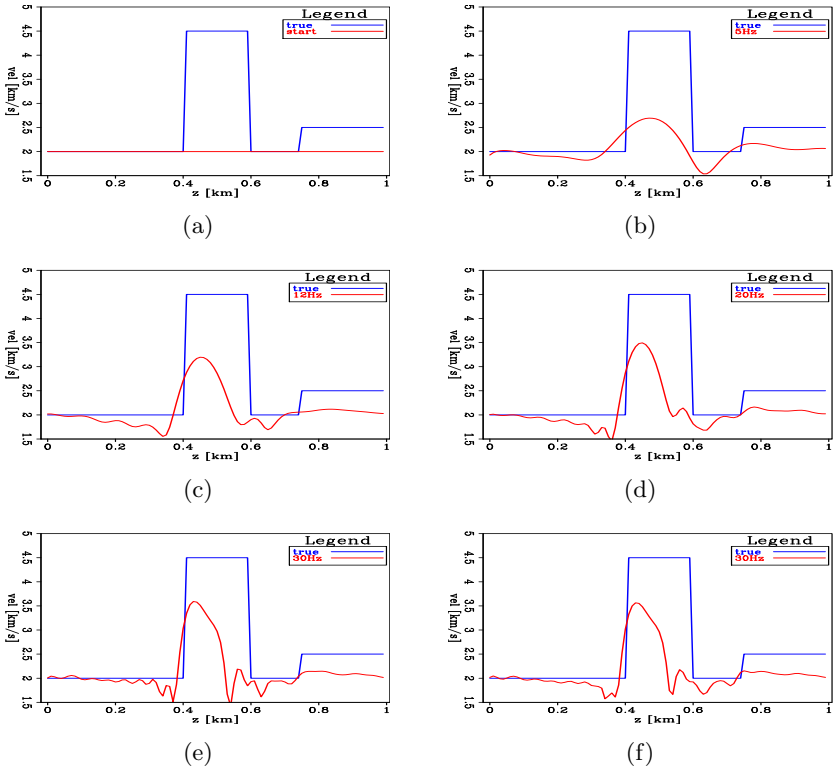


Figure 5: 1D velocity profile of inverted models using a multi-scale FWI scheme, extracted at $x = 1$ km. For each frequency band, we perform 40 iterations of FWI. a) Initial model. (b) Inverted model using energy up to 5 Hz. (c) Inverted model using energy up to 12 Hz. (d) Inverted model using energy up to 20 Hz. (e) Inverted model using energy up to 30 Hz. (f) Inverted model using energy up to 45 Hz. [**CR**] guillaume1/. fwi.mora.contrast.vel.start.1d,cont.5.1d,cont.12.1d,cont.20.1d,cont.30.1d,cont.45.1d

cycle-skipping. Therefore, the salt body boundaries are not being sharpened throughout the inversion process (or updated in the right direction), and the overlying sediments seemed to be wrongly updated in some locations, as seen in Figure 10(d). Future steps include performing a FWI scheme using a higher frequency content.

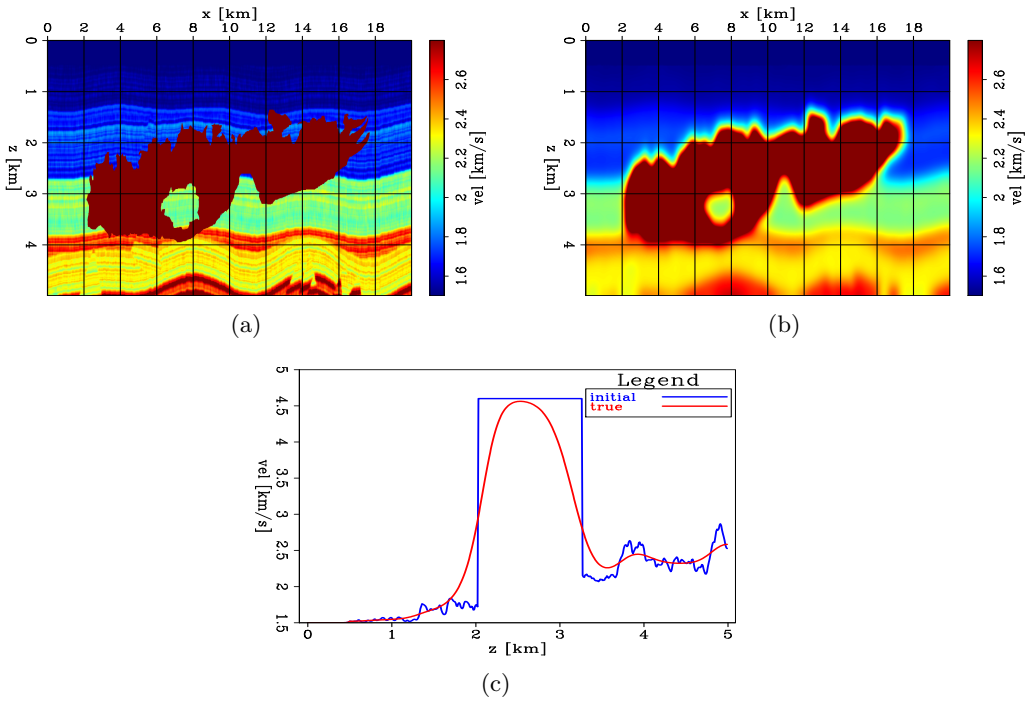
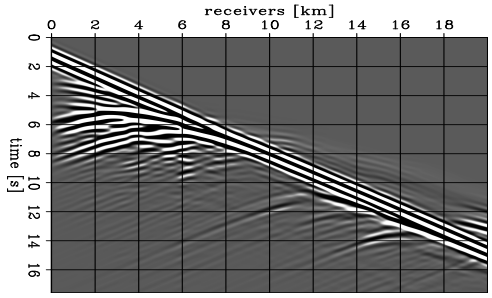


Figure 6: (a) True velocity model. (b) Initial velocity model. (c) 1D profile of the true and initial velocity models extracted at $x = 10$ km. [**NR**] guillaume1/. vel.salt.true,vel.salt.start,vel.salt.start.1d

TOMOGRAPHIC FULL WAVEFORM INVERSION

Almomin (2016) showed that TFWI can be robust and immune to cycle-skipping in relatively simple overburdens, but it is still unclear whether it can be applied to targets located under complex overburdens such as salt bodies. We identified two potential issues that may cause the algorithm to fail. (1) TFWI uses an extended Born modeling operator combined with an extended velocity model which allows the predicted data to stay in phase with the observed ones. The non-physical energy in the extended axis of the velocity model is slowly reduced during the inversion process. However, it requires this energy to be coherent in order to perform efficient tomographic updates to the model. Unfortunately, Etgen et al. (2014a,b); Barnier and Biondi (2015) showed that subsalt images are very sensitive to errors in the salt boundary delineation and rapidly lose coherency even in the extended space. (2) Salt bodies may have complex shapes, especially in the vicinity of the salt-sediments interface (e.g., impedance contrast solely due to a density change, carapaces and canyons). These structures can generate wavefields and wavepaths of such a complexity (e.g., prismatic waves, diffractions and wavefield triplications) that our extended modeling operator may not be able to capture.



(a)

Figure 7: Shot record generated by a source located at $x = 0$ km offset. [**ER**] guillaume1/. salt.shot

We first provide a brief theoretical review of TFWI, we apply the algorithm on the synthetic example shown in the previous section where we try to recover a sharp boundary with large velocity contrast, and we design a test to assess which combination of model extensions are more efficient at replicating wavefields generated by a complex top-salt interface.

Theoretical review

TFWI aims at minimizing the following objective function,

$$\Phi(\tilde{\mathbf{m}}) = \frac{1}{2}\|\tilde{\mathbf{f}}(\tilde{\mathbf{m}}) - \mathbf{d^{obs}}\|_2^2 + \frac{\epsilon^2}{2}\|\mathbf{A}\tilde{\mathbf{m}}\|_2^2, \tag{1}$$

where $\tilde{\mathbf{m}}$ is the extended velocity model, $\tilde{\mathbf{f}}$ is an extended wave-equation modeling operator, $\mathbf{d^{obs}}$ are the observed data, and $\mathbf{A}$ is an operator that enhances the non-physical energy of the extended model (e.g., differential semblance optimization (DSO)). The various types of model extensions are discussed in the next section. The first term of the objective function ensures that our predicted data match the observed ones, while the second term reduces the non-physical energy of our model. Due to computational limitation, modeling wavefields with the extended wave-equation operator $\tilde{\mathbf{f}}$ is impractical, and we linearize our forward modeling operator using the following Born approximation,

$$\tilde{\mathbf{f}}(\tilde{\mathbf{m}}) \approx \tilde{\mathbf{f}}(\tilde{\mathbf{m}}_\mathbf{0}) + \tilde{\mathbf{B}}(\tilde{\mathbf{m}}_\mathbf{0})\Delta\tilde{\mathbf{m}}, \tag{2}$$

where $\tilde{\mathbf{m}}_\mathbf{0}$ is an arbitrary model point, and $\tilde{\mathbf{B}}$ is an extended Born modeling operator, linear with respect to $\Delta\tilde{\mathbf{m}} = \tilde{\mathbf{m}} - \tilde{\mathbf{m}}_\mathbf{0}$. If we choose $\tilde{\mathbf{m}}_\mathbf{0}$ to be non-extended, we now have $\tilde{\mathbf{f}}(\tilde{\mathbf{m}}_\mathbf{0}) = \mathbf{f}(\mathbf{m_0})$, which is a standard (non-extended) wave-equation propagation. As done in Almomin (2016), we perform a scale separation, where the extended model is divided into two components,

1. a low wavenumber part $\tilde{\mathbf{b}}$ referred to as "background", which is responsible for the kinematic effects (transmissions) of the wavefields, and

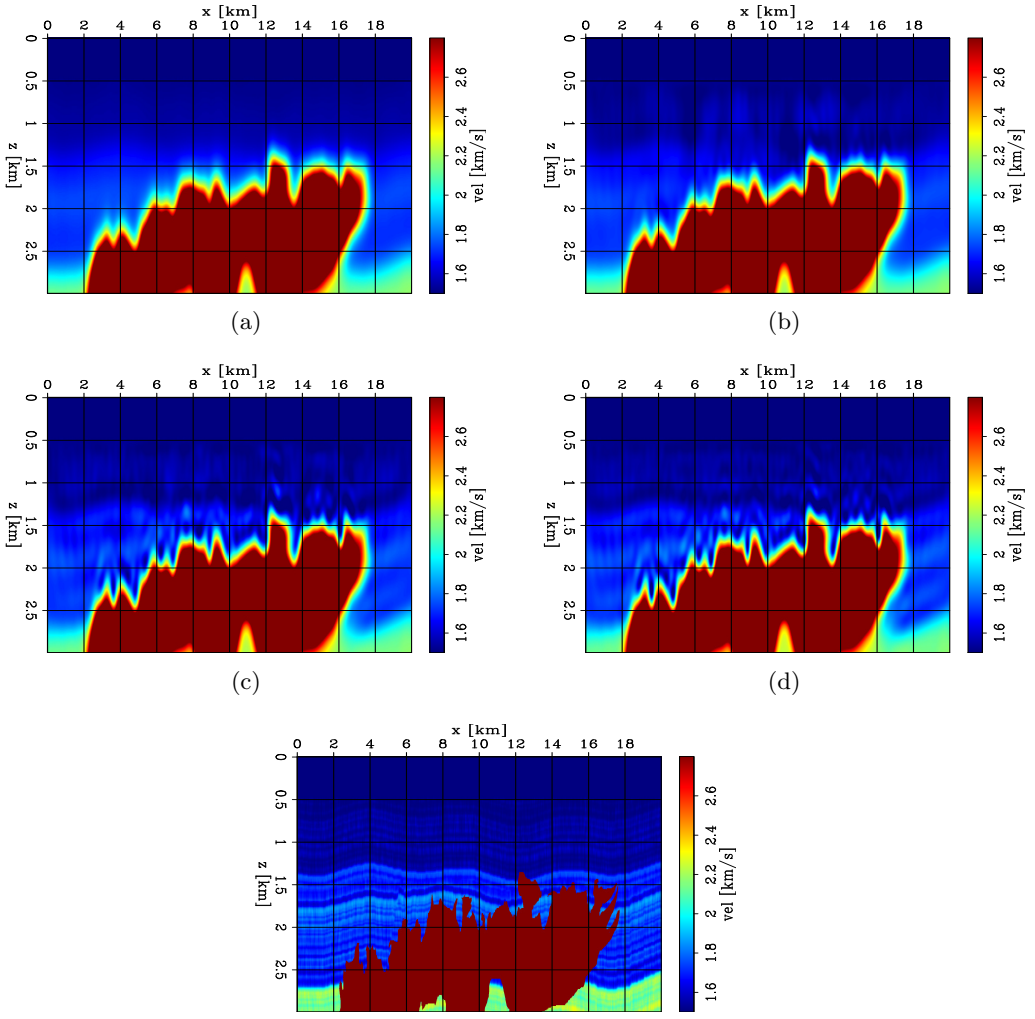


Figure 8: Inverted models at various iterations of the FWI workflow. (a) Starting model. (b) Model after 5 iterations. (c) Model after 15 iterations. (d) Model after 30 iterations. (e) True model. [**ER**] guillaume1/. vel.salt.start.cut,fwi.salt.vel.5,fwi.salt.vel.15,fwi.salt.vel.30,vel.salt.true.cut

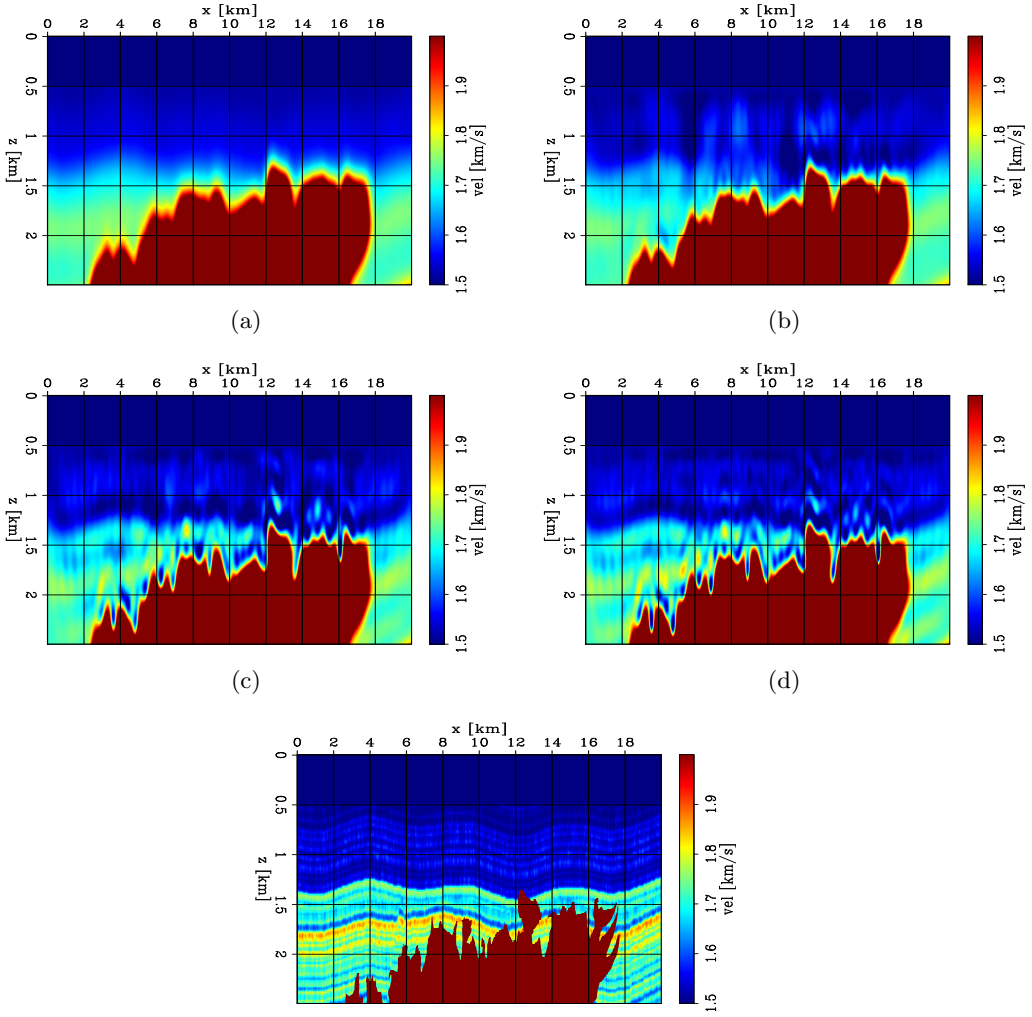


Figure 9: Inverted models at various iterations of the FWI workflow. Figures correspond to the same ones as in Figures 9 but have been clipped to a lower value in order to better show the updates made by the FWI scheme. (a) Starting model. (b) Model after 5 iterations. (c) Model after 15 iterations. (d) Model after 30 iterations. (e) True model. [**CR**]

guillaume1/. vel.salt.start.clip,fwi.salt.vel.5.clip,fwi.salt.vel.15.clip,fwi.salt.vel.30.clip,vel.salt.true.clip

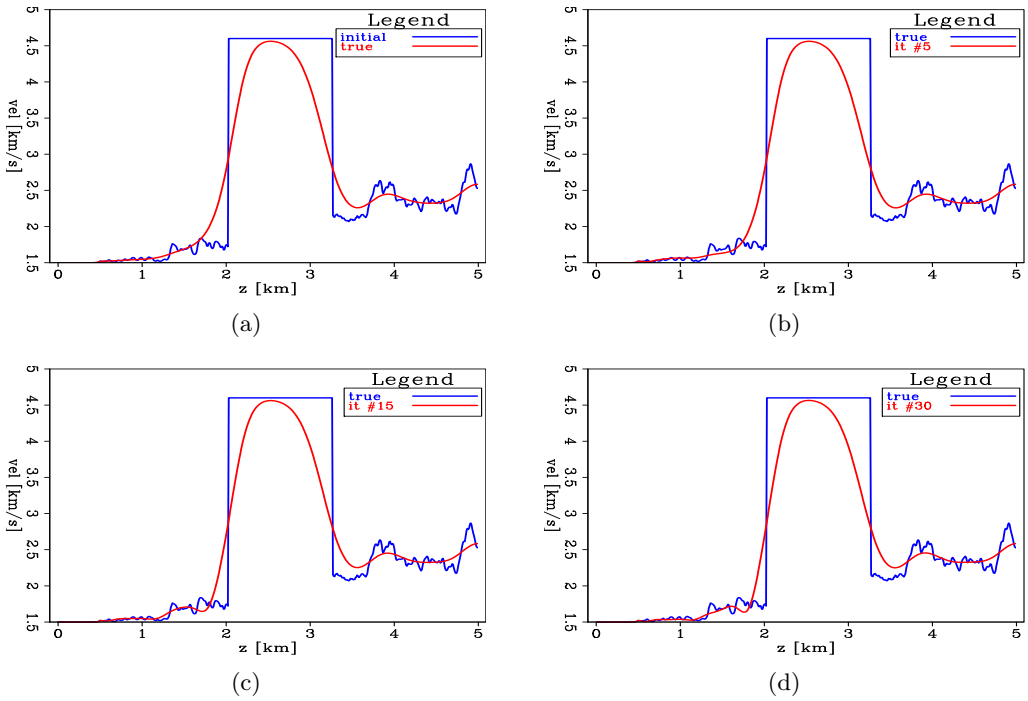


Figure 10: 1D velocity profiles of the inverted models at various iterations of the FWI workflow extracted at $x = 10$ km. (a) Starting model. (b) Model after 5 iterations. (c) Model after 15 iterations. (d) Model after 30 iterations. [**CR**] guillaume1/. vel.salt.start.1d,fwi.salt.vel.5.1d,fwi.salt.vel.15.1d,fwi.salt.vel.30.1d

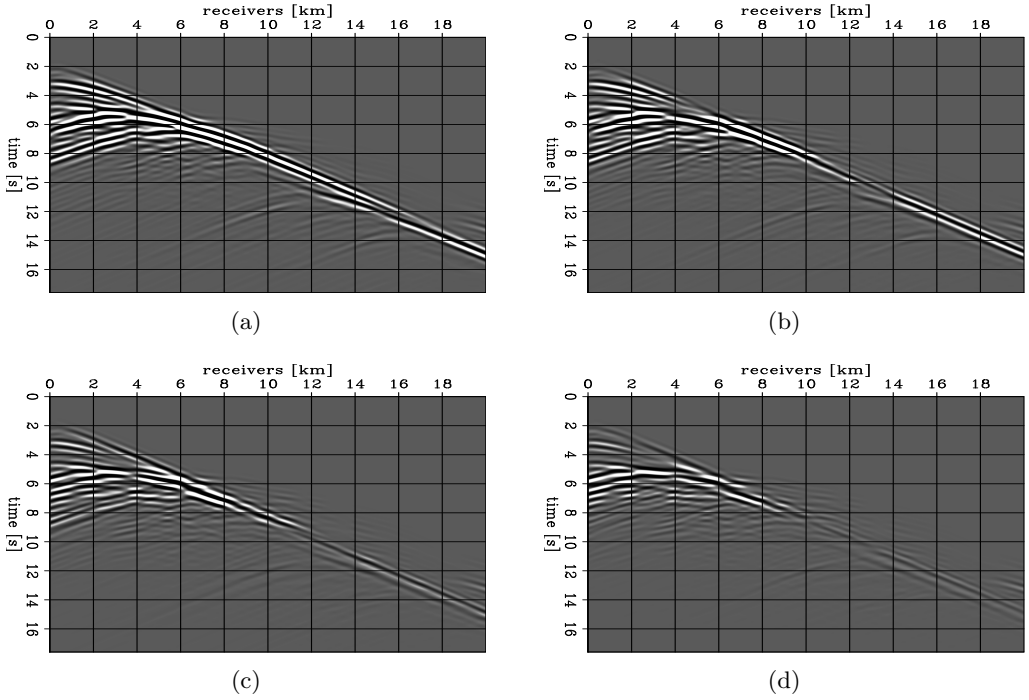


Figure 11: Data residuals for the shot record generated with a source positioned at $x = 0$ km at various iterations of the FWI workflow. (a) Initial residual. (b) Residual after 5 iterations. (c) Residual after 15 iterations. (d) Residual after 30 iterations. [**CR**] guillaume1/. salt.dat.res.1,salt.dat.res.5,salt.dat.res.15,salt.dat.res.30

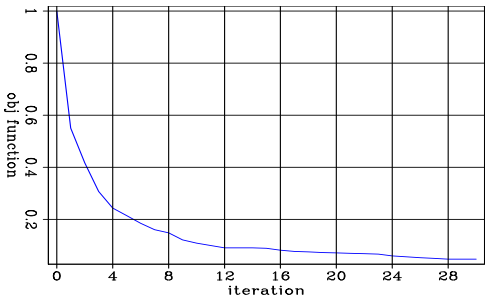


Figure 12: FWI objective function with iteration. [**ER**] guillaume1/. salt.obj

2. a high wavenumber part $\tilde{\mathbf{p}}$ referred to as "perturbation", which affects reflected and scattered energy.

Due to high computational costs, the background component is reduced to a non-extended model (while the reflectivity stays extended), and the total model can therefore be expressed by

$$\tilde{\mathbf{m}} \approx \mathbf{b} + \tilde{\mathbf{p}}, \tag{3}$$

where $\mathbf{b}$ is the non-extended background and $\tilde{\mathbf{p}}$ is the extended perturbation. We can re-formulate our forward modeling operator (equation 2) into the form

$$\tilde{\mathbf{f}}(\tilde{\mathbf{m}}) \approx \mathbf{f}(\mathbf{b}) + \tilde{\mathbf{B}}(\mathbf{b})\tilde{\mathbf{p}}. \tag{4}$$

equation 1 becomes

$$\Phi(\tilde{\mathbf{m}}) = \Phi(\mathbf{b}, \tilde{\mathbf{p}}) = \frac{1}{2}\|\mathbf{f}(\mathbf{b}) + \tilde{\mathbf{B}}(\mathbf{b})\tilde{\mathbf{p}} - \mathbf{d}^{\mathbf{obs}}\|_2^2 + \frac{\epsilon^2}{2}\|\mathbf{A}\tilde{\mathbf{p}}\|_2^2. \tag{5}$$

As explained in Almomin (2016), we now invert for two parameters ($\mathbf{b}$ and $\tilde{\mathbf{p}}$) that need to interact with each other in order to maintain a simultaneous inversion of scales, and both models need to be eventually combined into one physical model that can predict the recorded data. Almomin (2016) proposes a nested scheme composed of an outer and inner loop, described by the following workflow:

iterate outer loop
{
- compute data for inner loops: $\tilde{\mathbf{d}} = \mathbf{d}^{\mathbf{obs}} - \mathbf{f}(\mathbf{b})$
- initialize the extended perturbation to zero: $\tilde{\mathbf{p}} = \mathbf{0}$
- iterate inner loop
{
. set $\mathbf{r_d} = \tilde{\mathbf{B}}(\mathbf{b})\tilde{\mathbf{p}} - \tilde{\mathbf{d}}$
. set $\mathbf{r_m} = \mathbf{A}\tilde{\mathbf{p}}$
. set $\tilde{\Phi}(\mathbf{b}, \tilde{\mathbf{p}}) = 0.5\,\|\mathbf{r_d}\|_2^2 + 0.5\,\epsilon^2\|\mathbf{r_m}\|_2^2$
. compute $\nabla_{\mathbf{b}}\tilde{\Phi}$
. compute $\nabla_{\tilde{\mathbf{p}}}\tilde{\Phi}$
. apply scale mixing on both gradients to compute search direction
. apply nonlinear stepper
}
- extract the non-extended component of the perturbation: $\mathbf{p_{ne}}$
- update the background model: $\mathbf{b} \leftarrow \mathbf{b} + \mathbf{p_{ne}}$
- apply a low-wavenumber pass filter on $\mathbf{b}$
}

The data $\tilde{\mathbf{d}}$ are computed once for each outer loop iteration and serve as "observed data" for the inner loops. The gradient with respect to the background model $\nabla_{\mathbf{b}}\tilde{\Phi}$ is given by

$$\nabla_{\mathbf{b}}\tilde{\Phi} = \mathbf{T}^{*}\mathbf{r}_{\mathbf{d}}, \tag{6}$$

where $\mathbf{T}^{*}$ is the adjoint tomographic operator defined in Almomin (2016). The extended gradient with respect to the extended perturbation $\nabla_{\tilde{\mathbf{p}}}\tilde{\Phi}$ is given by,

$$\nabla_{\tilde{\mathbf{r}}}\tilde{\Phi} = \tilde{\mathbf{B}}^{*}\mathbf{r}_{\mathbf{d}} + \epsilon\mathbf{A}^{*}\mathbf{r}_{\mathbf{m}}, \tag{7}$$

where $\tilde{\mathbf{B}}^{*}$ denotes the adjoint extended Born modeling operator. It is assumed here that $\mathbf{A}$ is a linear operator with respect to $\tilde{\mathbf{p}}$. After computing both gradients (at each iteration of the inner loop), a scale mixing is applied in order to move the low-wavenumber model updates into a new background gradient $\nabla_{\mathbf{b}}\tilde{\Phi}_{\mathbf{mix}}$, and the high-wavenumbers updates into a new perturbation gradient $\nabla_{\tilde{\mathbf{p}}}\tilde{\Phi}_{\mathbf{mix}}$. $\nabla_{\mathbf{b}}\tilde{\Phi}$ is modified by extracting and adding the non-extended part of the perturbation gradient $\nabla_{\tilde{\mathbf{p}}}\tilde{\Phi}_{\mathbf{ne}}$ to the background gradient and by applying a low-wavenumber pass filter,

$$\nabla_{\mathbf{b}}\tilde{\Phi}_{\mathbf{mix}} = \mathbf{F}_{\mathbf{low}}\big(\nabla_{\mathbf{b}}\tilde{\Phi} + \nabla_{\tilde{\mathbf{p}}}\tilde{\Phi}_{\mathbf{ne}}\big). \tag{8}$$

where $\mathbf{F}_{\mathbf{low}}$ is a low-wavenumber pass filter. In an analogous fashion, we modify the extended perturbation gradient $\nabla_{\tilde{\mathbf{p}}}\tilde{\Phi}$,

$$\nabla_{\tilde{\mathbf{p}}}\tilde{\Phi}_{\mathbf{mix}} = \big(\mathbf{I}_{\mathbf{d}} - \mathbf{F}_{\mathbf{low}}\big)\big(\nabla_{\mathbf{b}}\tilde{\Phi} + \nabla_{\tilde{\mathbf{p}}}\tilde{\Phi}_{\mathbf{ne}}\big), \tag{9}$$

where $\mathbf{I}_{\mathbf{d}}$ is the identity operator. Note in equation 9, only the non-extended part of $\nabla_{\tilde{\mathbf{p}}}\tilde{\Phi}$ is modified as the scale mixing does not operate on the extended axes.

At the end of each outer loop, we have two inverted parameters, $\mathbf{b}$ (which contains the lowest wavenumber components) and $\tilde{\mathbf{p}}$ (which contains the highest wavenumber components) that we need to move into one background model $\mathbf{b}$. To do so, we extract the non-extended part of the inverted perturbation $\tilde{\mathbf{p}}$, add it to the inverted background, and apply a low-wavenumber pass filter. The reason for applying this filter is to control and slowly increase the wavenumber content of our inverted model $\mathbf{b}$ because incorrect high-wavenumber updates may affect the convergence efficiency of TFWI. Indeed, the filters' cut-off frequencies (both for the gradients and for $\mathbf{b}$) should be modified with outer loop iterations as we include higher and higher wavenumber components in the background during the inversion. At the end of the TFWI optimization scheme, (1) the inverted model $\mathbf{b}$ should contain all available wavenumbers recoverable from the data, and (2) the data predicted from model $\mathbf{b}$ should match the observed ones.

MODEL EXTENSION FOR COMPLEX WAVEFIELDS

As explained in Almomin (2016), cycle-skipping occurs for FWI (when using gradient-based optimization methods) because the model space is much smaller than the data space, resulting in the averaging and the destructive interference of many data components and the loss of kinematic information in the model space.

Avoiding cycle-skipping in FWI

In order to avoid cycle-skipping issues, the model and the modeling can be modified (e.g., extended) in order to preserve all kinematic information in the data. A nonlinear operator $\mathbf{f}$ with a linearized operator $\mathbf{F}$ (say about model point $\mathbf{m}_0$) satisfies the previous requirement if the following equation holds true for any model point $\mathbf{m}_0$ and any vector $\mathbf{d}$ (Biondi et al., 2016; Almomin, 2016),

$$\mathbf{F}(\mathbf{m}_0)\mathbf{F}^*(\mathbf{m}_0)\mathbf{d} \approx \mathbf{d} \tag{10}$$

In order to better understand the condition expressed in equation 10, we can start with the following FWI objective function

$$\Phi(\mathbf{m}) = \frac{1}{2}\|\mathbf{f}(\mathbf{m}) - \mathbf{d}^{\mathbf{obs}}\|_2^2. \tag{11}$$

Let $\mathbf{m}_0$ be our initial guess for the nonlinear problem. The first gradient is expressed by

$$\nabla\Phi(\mathbf{m_0}) = \mathbf{F}^*(\mathbf{m_0})(\mathbf{f}(\mathbf{m_0}) - \mathbf{d}^{\mathbf{obs}}), \tag{12}$$

and the first model update (using steepest descent) is given by

$$\Delta\mathbf{m_0} = -\alpha\nabla\Phi(\mathbf{m_0}). \tag{13}$$

The updated model is given by $\mathbf{m_1} = \mathbf{m_0} + \Delta\mathbf{m_0}$. To avoid cycle-skipping, the predicted data must be "close" to the observed data, so

$$\mathbf{f}(\mathbf{m_1}) = \mathbf{m_0} + \Delta\mathbf{m_0} \approx \mathbf{d}^{\mathbf{obs}}. \tag{14}$$

If the linearization holds true, we can write

$$\mathbf{f}(\mathbf{m_1}) \approx \mathbf{f}(\mathbf{m_0}) + \mathbf{F}(\mathbf{m_0})\Delta\mathbf{m_0} \approx \mathbf{d}^{\mathbf{obs}}. \tag{15}$$

Therefore, the condition for avoiding cycle-skipping can be re-written as

$$\alpha\mathbf{F}(\mathbf{m_0})\mathbf{F}^*(\mathbf{m_0})(\mathbf{f}(\mathbf{m_0}) - \mathbf{d}^{\mathbf{obs}}) \approx (\mathbf{f}(\mathbf{m_0}) - \mathbf{d}^{\mathbf{obs}}). \tag{16}$$

equation 16 should be true for any $\mathbf{m_0}$ and any data residual $\mathbf{d} = \mathbf{f}(\mathbf{m_0}) - \mathbf{d}^{\mathbf{obs}}$, and is therefore equivalent to equation 10 (up to a scaling factor α).

Optimal model extensions to predict complex waveforms

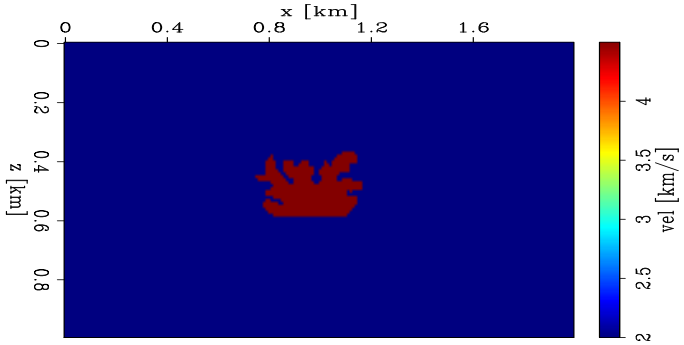
When our initial model is inaccurate, the condition expressed in equation 10 is not met and we are exposed to cycle-skipping. One way to solve that issue is to extend our linearized Born modeling operator and increase the dimension of the model space (Biondi and Almomin, 2014; Symes, 2008). Common extensions include vertical and horizontal subsurface offsets, and time-lags (Biondi and Almomin, 2014; Yang, 2013; Biondi et al., 1999). It has been shown that for reasonably complex overburdens, these extensions are able to predict data that stay in phase with the observed ones (Biondi and Almomin, 2014). However, wavefields and wavepaths become extremely complex when penetrating salt bodies and reflecting off their boundaries. Even an extended modeling operator might not be able to capture this complexity, which might lead to cycle-skipping.

In this section, we test the validity of equation 10 with various combination of extensions on a model containing a anomaly embedded in a constant velocity background (Figure 13(a)). The shape of anomaly tries to replicate the type of tortuosity we might expect at a top-salt interface. We use a wavelet containing energy up to 45 Hz and a finite-difference grid-size of 10 m x 10 m. We place sources and receivers at the surface spaced every 100 m and 10 m, respectively. Data residuals for a source located at $x = 1$ km are shown in Figures 14(a)-(b). The background data are computed using the true model without the anomaly, which only cancels the direct arrivals. Hence, the data residuals contains all the complex reflection events generated from the complex anomaly.

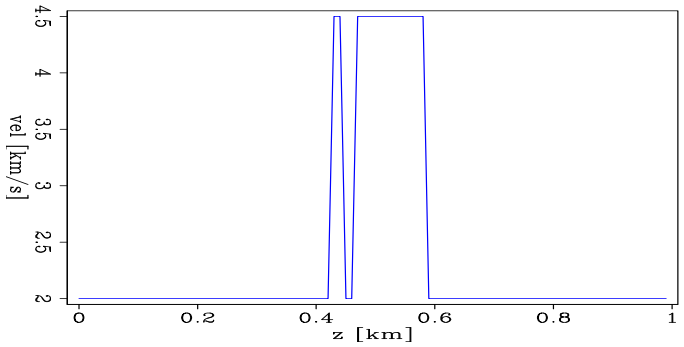
We compute the left side of equation 10 and we compare it to the data residuals (right side of the equation) for various combination of extensions. Our goal is to find a set of extensions for our Born modeling operator such that it satisfies equation 10. Results are shown in Figures 14(a)-(d). From Figures 15(a)-(c), we can see that the non-extended Born modeling operator fails to reconstruct the data residuals, while time-lags, z-offset extensions (and the combination of the two) seem to allow for a better reconstruction. We also observe that the reprojection produces artifacts (in the reconstructed data) that tend to be stronger for the x-offset extension (or combinations of x-offsets extensions with other extensions) (Figures 18, 19, and 21). We are not completely certain of the origin of such artifacts (and why it is stronger for the x-offset extension). However, we believe that it may be caused that the fact that we do not compensate for the illumination, the imprint of wavelet (e.g., the squaring of its amplitude spectrum due to the convolution introduced in the extended Born modeling) when applying a simple adjoint (equation 8). A possible way to test this hypothesis to be to apply an approximated Gauss-Newton step before reprojecting in the data space. Nevertheless, from these results, it seems that using combinations of extensions can potentially be efficient to satisfy the condition expressed in equation 10, even for complex salt bodies.

SUMMARY AND FUTURE WORK

We showed that FWI rapidly fails at recovering large and sharp velocity contrasts, such as the ones encountered in salt basins. In the synthetic examples conducted in this report (which contained noise-free, ultra-low frequency data and ideal acquisition geometries), FWI was not able to converge towards the correct solution. These results led us to pursue the work of Almomin (2016) and understand if and where it fails at inverting for very



(a)



(b)

Figure 13: (a) Velocity model containing. (b) velocity profile of the true model extracted at $x = 1$ km. [**CR**] guillaume1/. vel.mora.salt,vel.mora.salt.1d

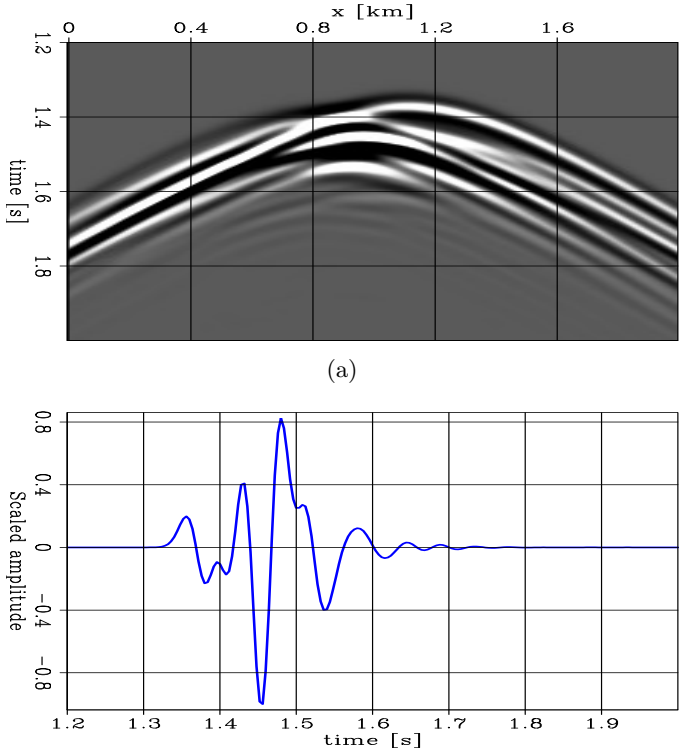
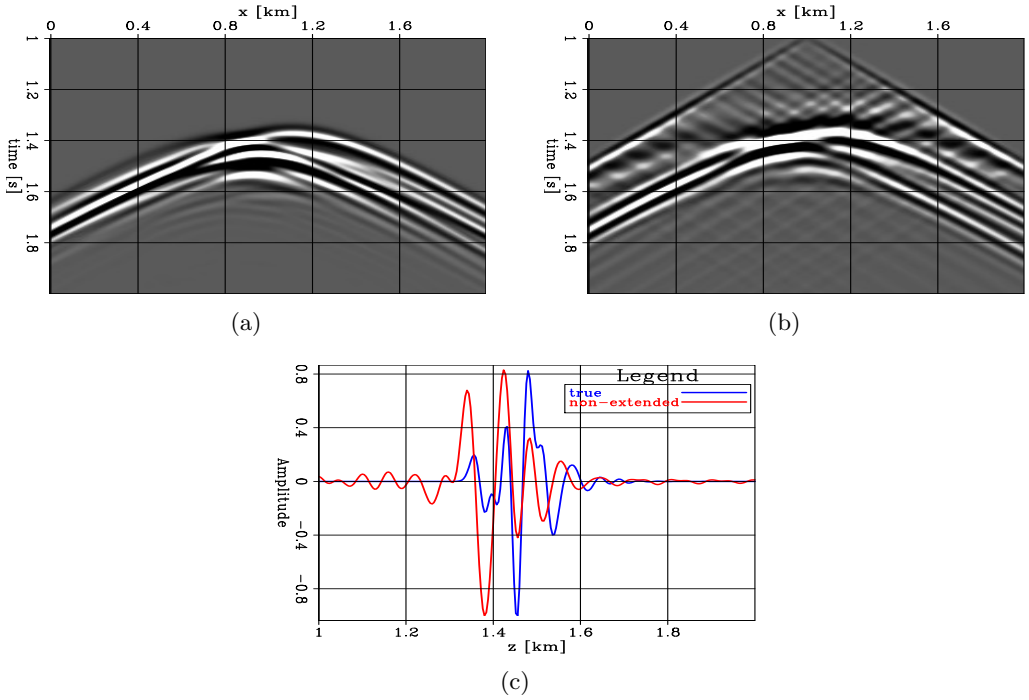
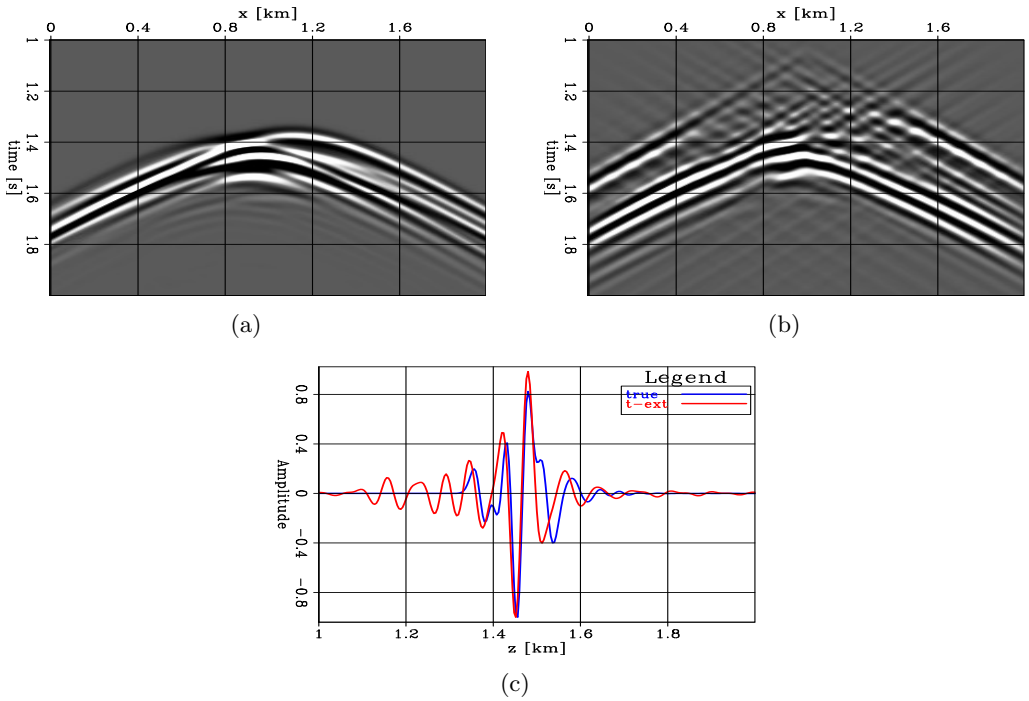


Figure 14: (a) Shot record generated by a source located at $x = 1$ km offset. (b) Trace extracted at $x = 1$ km. [**CR**] guillaume1/. dat.residuals,dat.residuals.1d



(a) (b) (c)

Figure 15: (a) Data residuals. (b) Data residuals after reprojection using a non-extended Born operator. [**CR**] guillaume1/. fwd.salt.true,fwd.salt,fwd.salt.1d



(a) (b) (c)

Figure 16: (a) Data residuals. (b) Data residuals after reprojection using a time-lag extended Born operator. [**CR**] guillaume1/. fwd.salt.true,fwd.salt.t,fwd.salt.t.1d

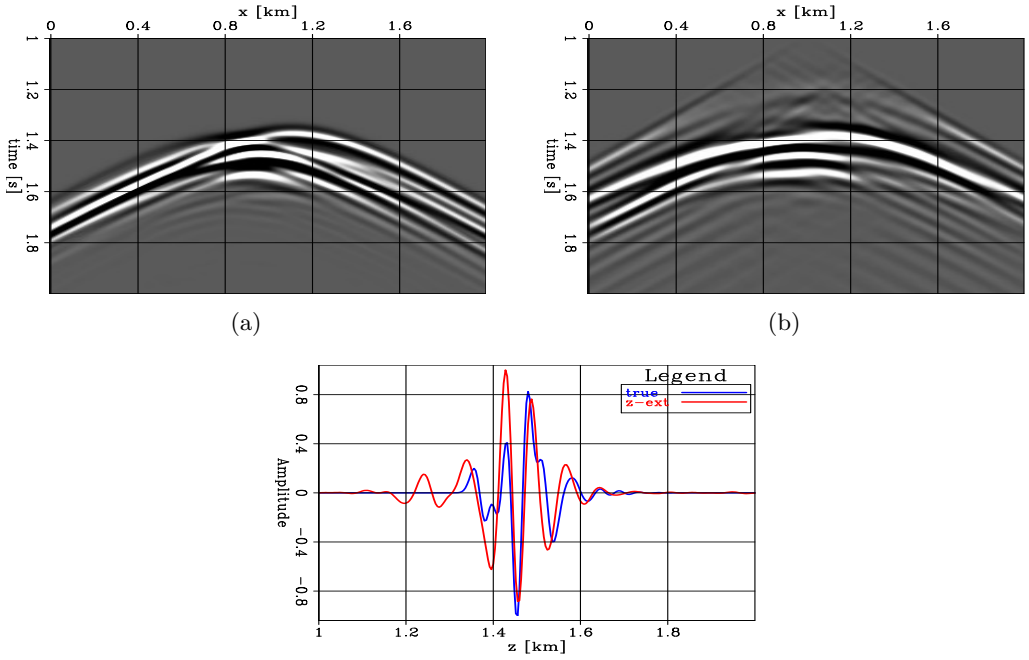


Figure 17: (a) Data residuals. (b) Data residuals after reprojection using a z-offset extended Born operator. [**CR**] guillaume1/. fwd.salt.true,fwd.salt.z,fwd.salt.z.1d

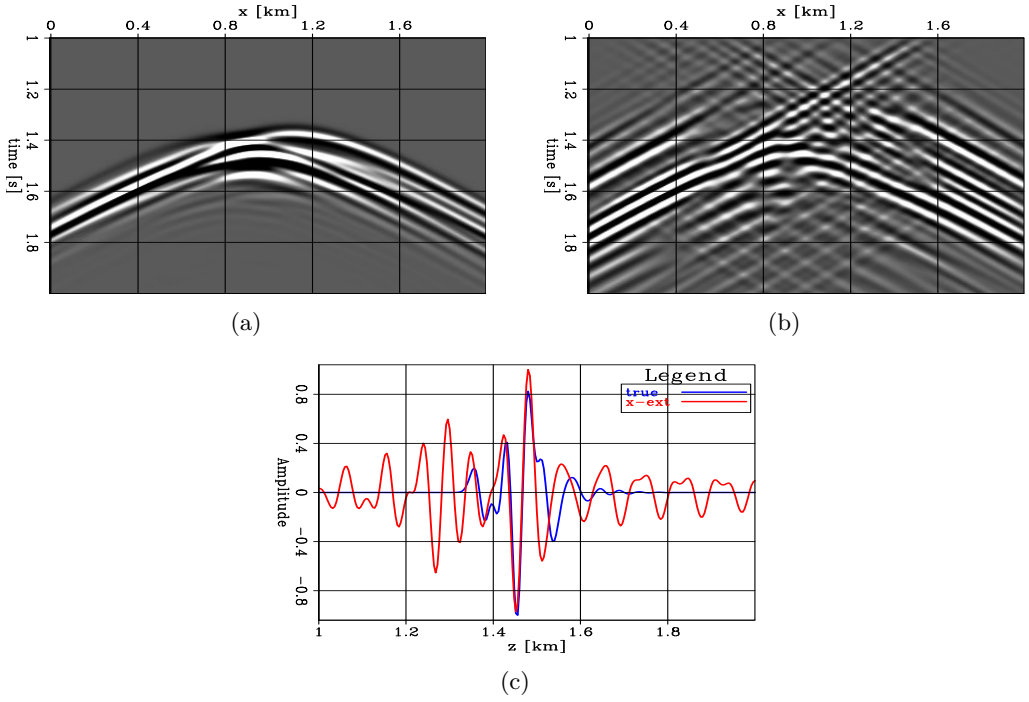
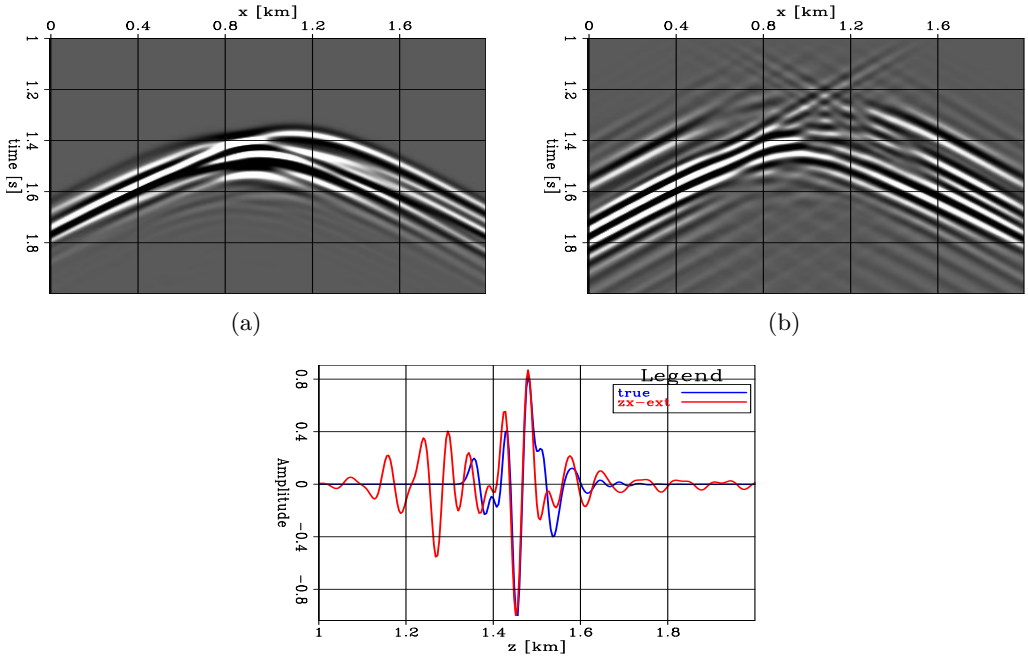
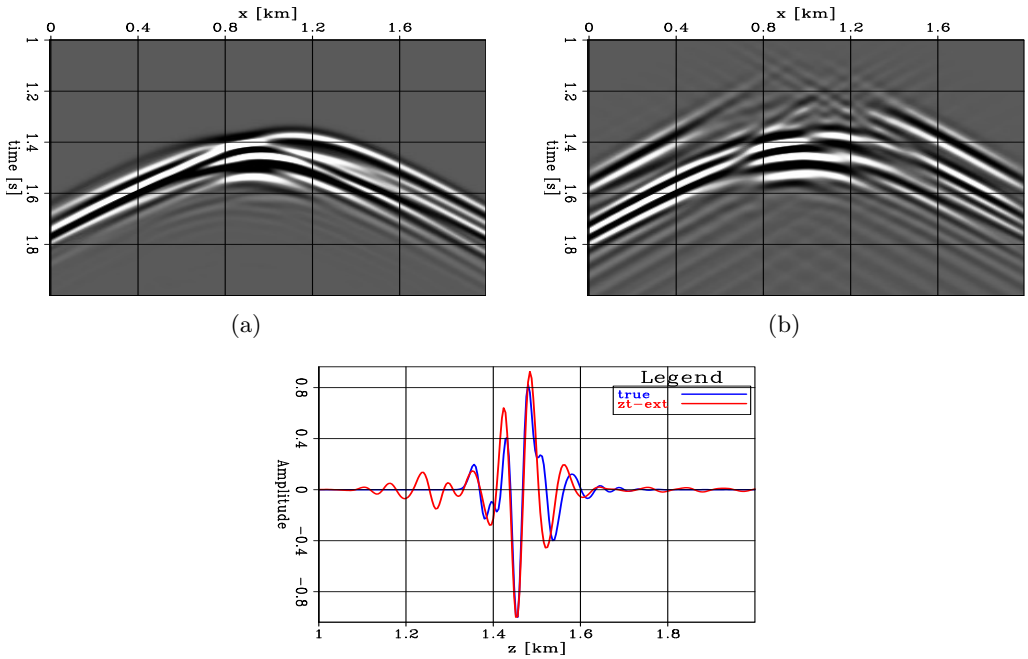


Figure 18: (a) Data residuals. (b) Data residuals after reprojection using a x-offset extended Born operator. [**CR**] guillaume1/. fwd.salt.true,fwd.salt.x,fwd.salt.x.1d



(a) (b) (c)

Figure 19: (a) Data residuals. (b) Data residuals after reprojection using a z-offset and x-offset extended Born operator. [**CR**] guillaume1/. fwd.salt.true,fwd.salt.zx,fwd.salt.zx.1d



(a) (b) (c)

Figure 20: (a) Data residuals. (b) Data residuals after reprojection using a z-offset and time-lag extended Born operator. [**CR**] guillaume1/. fwd.salt.true,fwd.salt.zt,fwd.salt.zt.1d

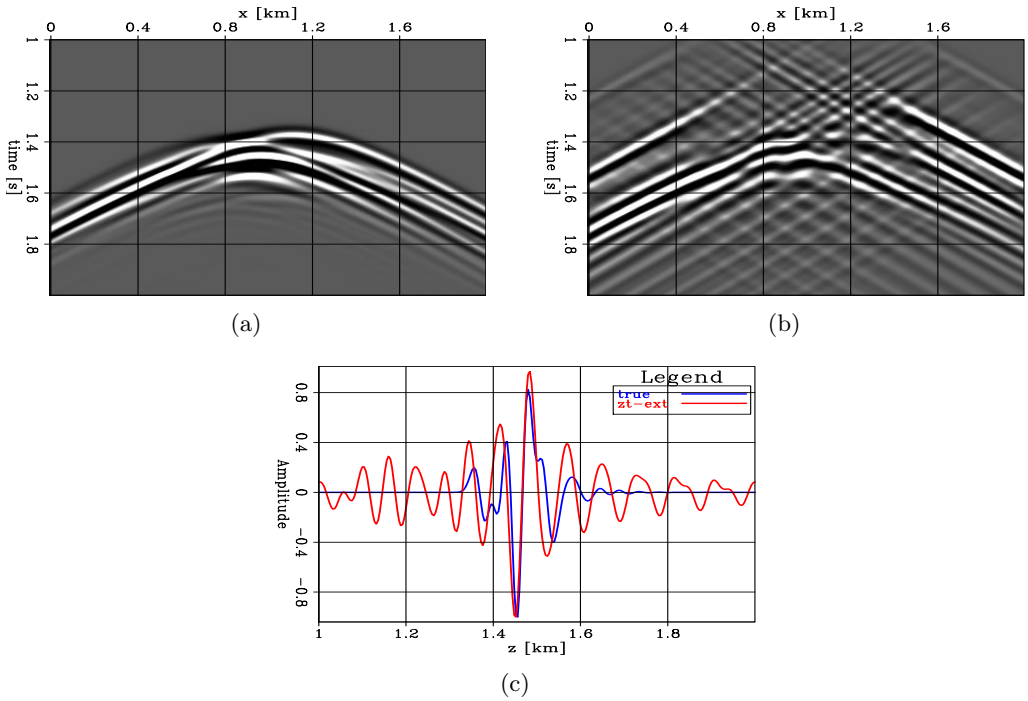


Figure 21: (a) Data residuals. (b) Data residuals after reprojection using all extensions. [**CR**] guillaume1/. fwd.salt.true,fwd.salt.zxt,fwd.salt.zxt.1d

complex salt boundaries. We are still benchmarking our TFWI code, investigating its optimal parameters (such as the ones used for filtering the gradients and the model updates), and currently testing the algorithm on model containing anomalies with large and sharp velocity contrasts. Future steps include testing TFWI on more realistic examples, such as the one conducted in this report which contains the complex salt structure.

REFERENCES

Almomin, A., 2016, Tomographic full waveform inversion: Stanford University.

Barnier, G. and B. Biondi, 2015, Addressing the effects of inaccurate top-salt delineation on subsalt seismic imaging: SEP-Report, **158**, 67–102.

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

Biondi, B., R. Sarkar, and J. Jennings, 2016, Solving nonlinear inverse problems by linearized model extension - a survey of possible methods: SEP-Report, **165**, 93–122.

Biondi, B., P. Sava, et al., 1999, Wave-equation migration velocity analysis: 69th Ann. Internat. Mtg Soc. of Expl. Geophys, 1723–1726.

Biondi, E. and G. Barnier, 2017, A flexible out-of-core solver for linear/non-linear problems: SEP-Report, **168**.

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

Clapp, R. G., 2014, Synthetic model building using a simplified basin modeling approach:

SEP-Report, **155**, 145–152.

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

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

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

Mora, P., 1989, Inversion= migration+ tomography: Geophysics, **54**, 1575–1586.

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

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

Wang, K., B. Deng, Z. Zhang, L. Hu, and Y. Huang, 2015, Top of salt impact on full waveform inversion sediment velocity update, *in* SEG Technical Program Expanded Abstracts 2015, 1064–1069, Society of Exploration Geophysicists.

Williamson, P., C. Rivera, and K. Mansoor, 2016, Acoustic fwi applied to subsalt imaging: an illuminating case study from offshore angola, *in* SEG Technical Program Expanded Abstracts 2016, 4896–4900, Society of Exploration Geophysicists.

Yang, T., 2013, Wavefield tomography using extended images: Colorado School of Mines.

An efficient time-lapse full waveform inversion by saving the wavefield at boundaries around the reservoir

Yinbin Ma and Robert G. Clapp

ABSTRACT

The need to propagate the source and receiver wavefield over the whole subsurface model is a challenge for time-lapse full-waveform inversion (FWI). We show that by saving the wavefield at the boundary enclosing the reservoir, we can estimate the wavefield around the reservoir after model perturbation. The perturbed wavefield can predict synthetic data at the surface through the pre-computed Green's function. We eliminate the need to propagate the wavefield through the overburden area and therefore accelerate the FWI procedure. Random phase-encoding can be incorporated into our proposed workflow and has the potential to further reduce the computational cost.

INTRODUCTION

Full-waveform inversion (Tarantola, 1984; Virieux and Operto, 2009) estimates the high-resolution subsurface models by minimizing the mismatch between the observed data and the synthetic data. FWI is a useful tool for time-lapse (4D) seismic imaging problems (Maharramov and Biondi, 2013). One of the computational challenges of time-lapse FWI is that we need to propagate the source and receiver wavefield through the whole subsurface model several times for each FWI iteration. This computational cost limits the number of iterations we can afford and therefore reduces the quality of 4D FWI results.

Previous studies (Robertsson and Chapman, 2000; Clapp, 2008; Borisov and Singh, 2013; Willemsen et al., 2016) show that by saving the wavefield at the boundaries of the computational domain, the wavefield inside the boundaries can be regenerated. Based on this idea, Clapp (2008) developed memory efficient reverse-time migration (RTM), and Robertsson and Chapman (2000) proposed an efficient algorithm for computing synthetic seismograms after model alternation. In this paper, we study the feasibility of accelerating time-lapse FWI by reconstructing the wavefield near the reservoir using similar techniques.

First we show that by saving the wavefield at the boundaries enclosing the reservoir, we can reconstruct the background wavefield and compute the perturbed wavefield. The wavefield within the boundaries is regenerated by injecting the saved boundaries, with the computational domain being the enclosed area which is much smaller than the whole subsurface model. The perturbed wavefield can predict the synthetic data at the surface with the pre-computed Green's function, without numerical wave propagation in the overburden area. Therefore FWI procedure is accelerated and more iterations are affordable.

Memory requirements and computational cost increase with the size of the seismic data, as the wavefield for each shot has to be saved. We aim to further reduce the computational cost by incorporating random phase-encoding (Tang, 2011) into our workflow.

METHODS

The FWI objection function measures the mismatch between the synthetic data and the recorded data,

$$J(\mathbf{m}) = \frac{1}{2}\|\mathbf{d}^{\text{syn}}(\mathbf{m}) - \mathbf{d}^{\text{obs}}\|_2^2, \tag{1}$$

where $\mathbf{m}$ is the subsurface model we want to estimate (velocity, anisotropic parameters, elastic parameters, etc), $\mathbf{d}^{\text{syn}}$ is the synthetic data obtained by forward propagating the source wavefield, and $\mathbf{d}^{\text{obs}}$ is the recorded data during the seismic survey. To minimize this objective function 1, the adjoint-state methods are commonly used (Tromp et al., 2005; Plessix, 2006), which require the backward propagation of the data residuals $\left(\mathbf{d}^{\text{syn}}(\mathbf{m}) - \mathbf{d}^{\text{obs}}\right)$.

For time-lapse FWI we are mostly interested in estimating production-induced model change in the vicinity of the reservoir. In comparison, there is little change in the overburden model during production. Direct application of the adjoint-state methods is computationally expensive as it requires wave propagations through the overburden area at every iteration. To reduce the cost, we modify the previous objective function, assuming the overburden has been estimated properly,

$$\begin{aligned} J(\mathbf{m}) &= \frac{1}{2}\| \left(\mathbf{d}^{\text{syn}}(\mathbf{m}) - \mathbf{d}^{\text{syn}}(\mathbf{m_0})\right) - \left(\mathbf{d}_{\text{obs}} - \mathbf{d}^{\text{syn}}(\mathbf{m_0})\right) \|_2^2 \\ &= \frac{1}{2}\|\tilde{\mathbf{d}}^{\text{syn}}(\mathbf{m}) - \tilde{\mathbf{d}}_{\text{obs}}\|_2^2, \end{aligned} \tag{2}$$

where $\mathbf{m_0}$ is the background subsurface model, and $\tilde{\mathbf{d}}^{\text{syn}}(\mathbf{m})$ is the perturbed data after model alteration $(\mathbf{m} - \mathbf{m_0})$. We assume $(\mathbf{m} - \mathbf{m_0})$ is non-zero only in the vicinity of the reservoir.

To accelerate the estimation of $\tilde{\mathbf{d}}^{\text{syn}}(\mathbf{m})$, we can set up virtual receivers at the top of the reservoir (as indicated in Figure 1), and record the perturbed wavefield $\tilde{\mathbf{p}}^{\text{syn}}(\mathbf{m})$. The perturbed data at the surface can then be predicted from the perturbed wavefield at the virtual receivers as,

$$\tilde{\mathbf{d}}^{\text{syn}}(\mathbf{m}) = \mathbf{G}\tilde{\mathbf{p}}^{\text{syn}}(\mathbf{m}), \tag{3}$$

where $\mathbf{G}$ is the pre-computed Green's function thats connect the virtual receivers to the surface receivers. $\mathbf{G}$ is a linear operator. Hence our new objection function can be rewritten as,

$$J(\mathbf{m}) = \frac{1}{2}\|\mathbf{G}\tilde{\mathbf{p}}^{\text{syn}}(\mathbf{m}) - \mathbf{d}^{\text{obs}}\|_2^2. \tag{4}$$

With the new objective function 4, we only need to propagate the wavefield to the virtual receivers instead of all the way to the surface. The wavefield $\tilde{\mathbf{p}}^{\text{syn}}(\mathbf{m})$ can be compute efficiently with the following wave equation (acoustic wave equation for example),

$$\begin{aligned} \left[\mathbf{m}\partial_t^2 - \nabla^2\right] \mathbf{p}^{\text{syn}}(\mathbf{m}) &= 0, \\ \mathbf{p}^{\text{syn}}(\mathbf{m}) &= \mathbf{p}^{\text{syn}}_{\text{saved}}, \text{ in boundary area B} \\ \tilde{\mathbf{p}}^{\text{syn}}(\mathbf{m}) &= \mathbf{p}^{\text{syn}}(\mathbf{m}) - \mathbf{p}^{\text{syn}}(\mathbf{m}_0), \end{aligned} \tag{5}$$

where $\mathbf{p}^{\text{syn}}_{\text{saved}}$ is the saved background wavefield in the boundary area B shown in Fig 1. Notice that equation 5 is different from the initial value problem and numerically we reinject the saved wavefield in the boundary area. We have eliminated the need to propagate

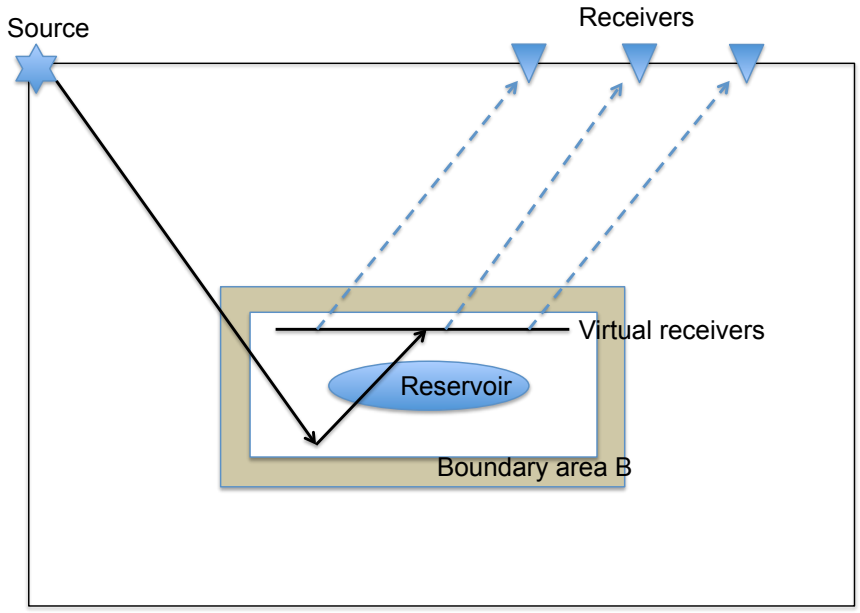


Figure 1: Cartoon illustration of the numerical experiment. The model perturbation is limited to the area within the reservoir. The wavefield in the boudary area B is saved in memory. The perturbed wavefield is numerically estimated near the reservoir and recorded at the virtual receivers. The synthetic data at the surface is estimated from the virtual receivers with the pre-computed Green's function in the overburden area. [**NR**] yinbin2/. cart

the wave in the overburden area and our algorithm runs faster than the conventional FWI approach.

The expense of this approach is that we need to save the wavefield at the boundaries for every single shot, therefore the memory cost is roughly:

$$\text{memory} \propto N_s \times N_t \times N_b, \tag{6}$$

where N_s is the number of shots in the seismic survey, N_t is the total number of time steps during wave propagation, and N_b is the size of boundaries we need to save. The memory cost is increasing as more shots are included in the FWI objective function.

The memory cost can be further reduce by incorporating the idea of random phase-encoding into our workflow. In equation 1, with $\mathbf{d}^{\text{obs}}$ being the recorded data for all shot, we can write the data in terms of shot gathers $\mathbf{d}^{\text{obs}} = \left[\mathbf{d}_1^{\text{obs}}, \cdots \mathbf{d}_i^{\text{obs}}, \cdots, \mathbf{d}_{N_s}^{\text{obs}}\right]$, with $\mathbf{d}_i^{\text{obs}}$ represent the data for shot i. We can blend the data from different shot,

$$\mathbf{d}^{\text{obs},\alpha} = \frac{1}{\sqrt{N_s}} \sum_{i=1}^{N_s} \mathbf{d}_i^{\text{obs}} \alpha_i, \tag{7}$$

where the weighting function α_i is a complex number and uniformly distributed on the unit circle. has absolute value $|\alpha_i|$.

The new objective function associated with the blended data can be written as,

$$J(\mathbf{m}, \alpha) = \frac{1}{2} \|\mathbf{G}\tilde{\mathbf{p}}^{\text{syn},\alpha}(\mathbf{m}) - \mathbf{d}^{\text{obs},\alpha}\|_2^2, \tag{8}$$

Multiple realizations of phase encoding can be used to attenuate the cross-talk, and our final solution is expressed as,

$$\mathbf{m}^* = \text{argmin}_{\mathbf{m}} \sum_{\alpha=1}^{N_\alpha} J(\mathbf{m}, \alpha), \tag{9}$$

where N_α is the number of realizations of phase encoding. Notice that for equation 9, the memory cost reduces to $\propto N_\alpha \times N_t \times N_b$, which does not grow as the number of shots increase.

NUMERICAL EXAMPLE

Regenerating the wavefield around the reservoir

We design a synthetic model as shown in Figure 2, with a local model perturbation. The wavefield at the boundary area enclosing the reservoir is saved. First we apply equation 5 to the unperturbed model to regenerate the background wavefield. The numerical results can be seen in Figure 3. The wavefield around the reservoir is perfectly regenerated in the absence of model perturbation.

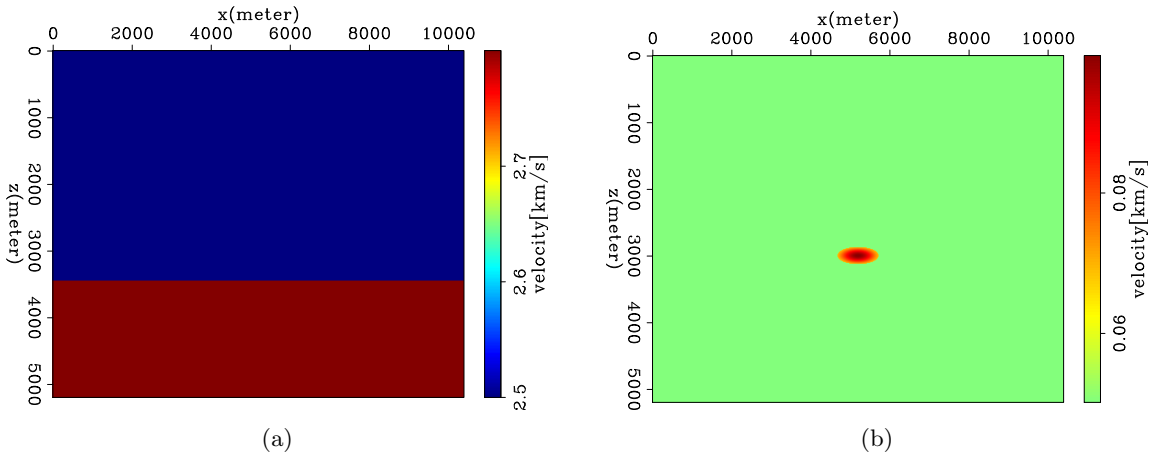


Figure 2: (a) Background velocity model. (b) Velocity model perturbation. [**ER**] yinbin2/. velM,vel4D

Next we apply equation 5 to perturbed model to predict the perturbed wavefield. The numerical results can be seen in Figure 4. We have captured most of the perturbed wavefield. The difference will be discussed in the following subsection.

Local time-lapse FWI solver

Next we construct a local solver for the 4D FWI problem. We demonstrate that we can predict the data residual at the surface and the gradient without requiring wave propagation in the overburden area, and use it to estimate the model perturbation.

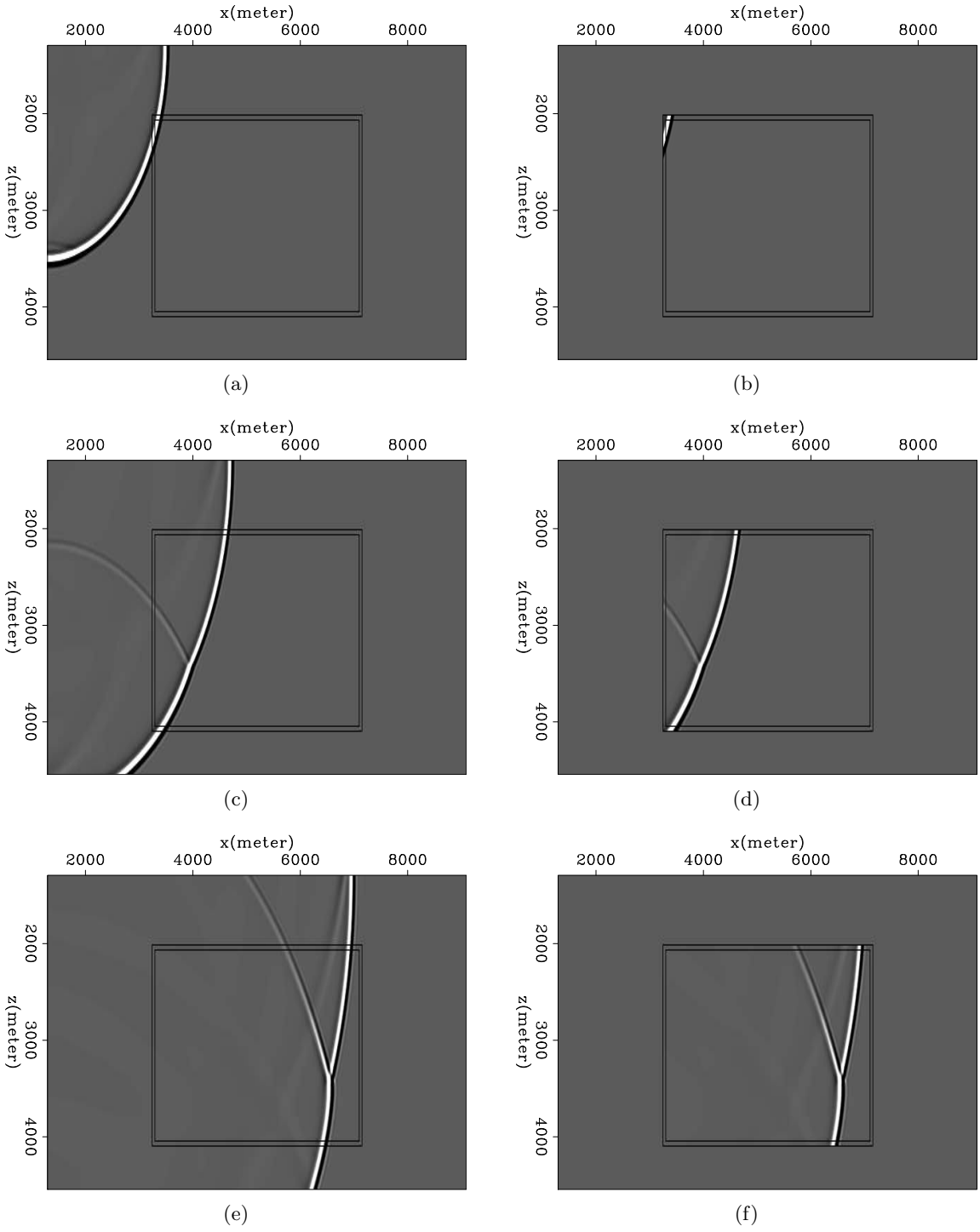


Figure 3: Left panel: (a), (c) and (e) are snapshots of wavefield by solving the wave equation in the whole subsurface. Right panel: (b), (d) and (f) are snapshots of wavefield by solving the wave equation around the reservoir with rejected boundaries. Black rectangles indicate the boundaries in which we save the wavefield. [**ER**]
yinbin2/. BGWavea,BGWaveb,BGWavec,BGWaved,BGWavee,BGWavef

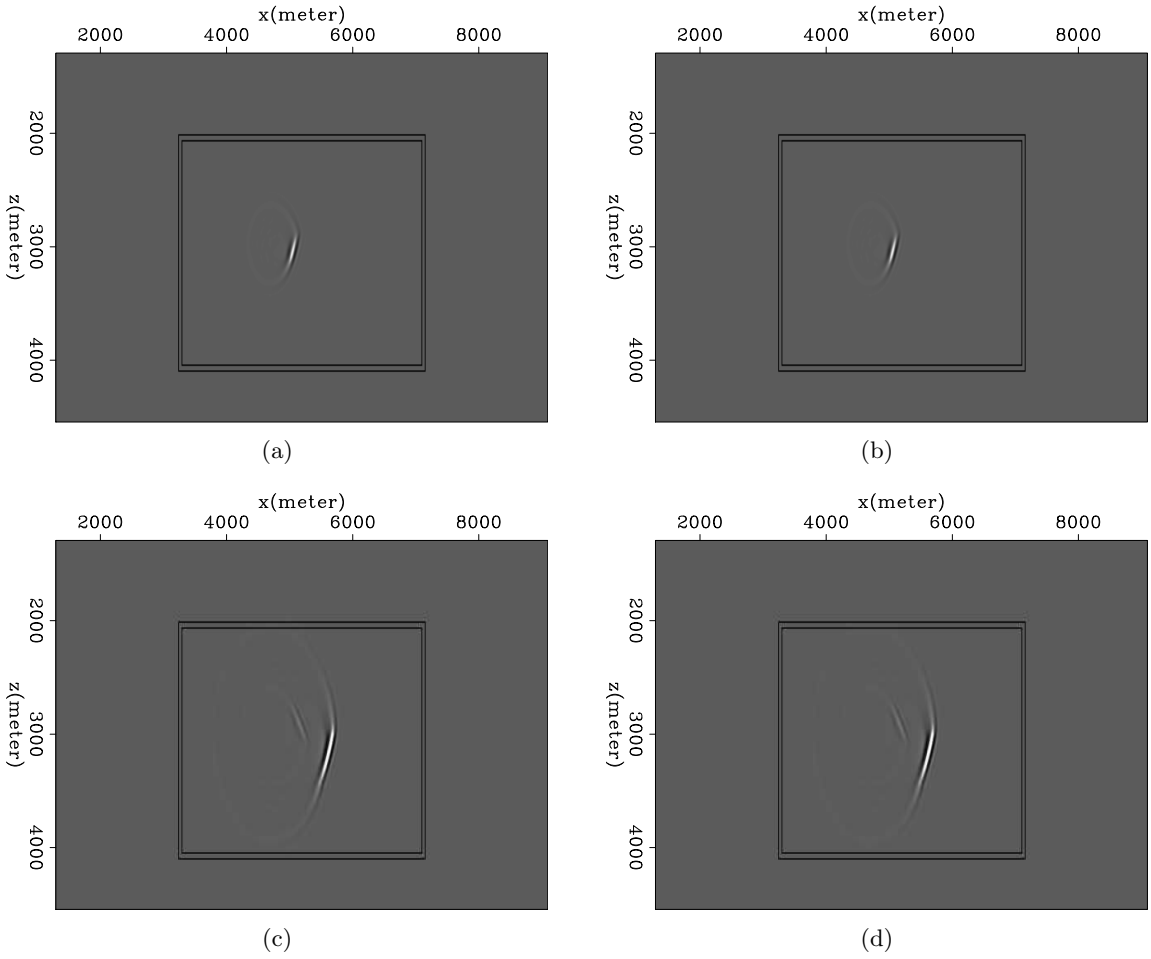
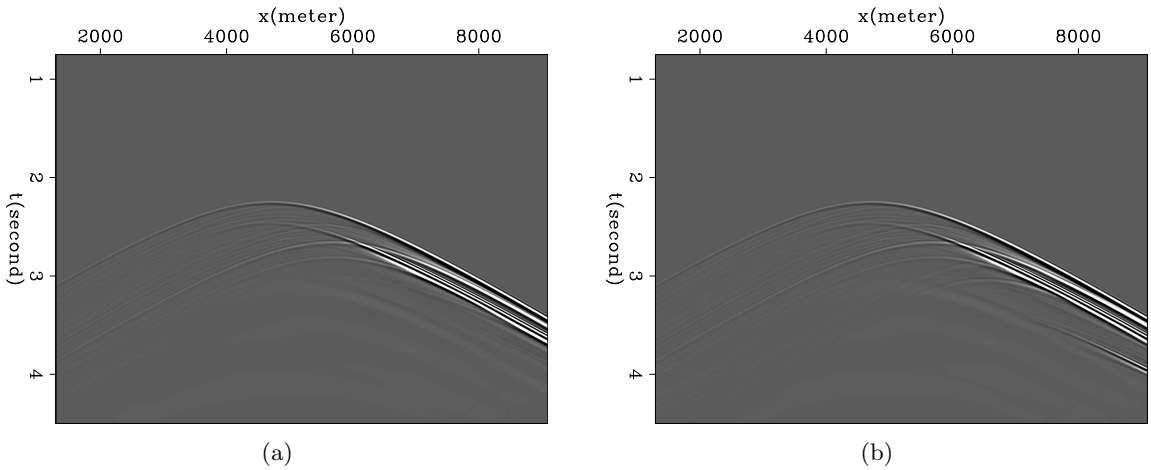


Figure 4: Left panel: (a) and (c) are snapshots of perturbed wavefield computed from the whole subsurface. Right panel: (b) and (d) are snaptshots of perturbed wavefield by solving the wave equation around the reservoir with rejected boundaries. [**CR**] yinbin2/. 4DWavea,4DWaveb,4DWavec,4DWaved

Once we compute and recorded the perturbed wavefield at the virtual receivers, the data residual at the surface can be approximated by $\tilde{\mathbf{d}} = \mathbf{G}\tilde{\mathbf{p}}^{\text{syn}}(\mathbf{m})$ where $\tilde{\mathbf{p}}^{\text{syn}}(\mathbf{m})$ is estimated by local wave propagation around the reservoir. In Figure 5, we compare the true data residual and our estimated data residual. Most the data residual has been captured, despite observable artifacts. The reason is that our local wave propagation is not designed to capture all the perturbed wavefield. For example, the perturbed wavefield from the reflector beneath the boundaries is not included in our workflow.

We compare the FWI results from conventional FWI with our proposed workflow, and the results are shown in Figure 6. Both methods estimate the model perturbation properly. The results from our proposed workflow requires much less computational time comparing with the one from FWI with wave propagation in the whole subsurface, although they do not match completely.



(a) (b)

Figure 5: (a) True data residual.(b) Estimated data residual from local wave propagation around the reservoir, and pre-computed Green's function in the overburden. [**CR**] yinbin2/. dataT,dataM

Reducing the cost further by random phase encoding

Our final numerical experiment tests the feasibility to reduce the memory cost by random phase-encoding. In Figure 7(a) and 7(b) we show the FWI results with one realization of the random phase-encoding ($N_\alpha = 1$), with observable artifacts from the cross-talk between different shots. The cross-talk will reduce when multiple realizations of random phase-encoding are included in the objective function. In Figure 7(c) we show the results with $N_\alpha = 5$, and in Figure 7(d) we show the results with $N_\alpha = 20$. The results approach the one without phase-encoding (Figure 6(b)), when N_α is sufficiently large, but much smaller than the number of shots N_s.

CONCLUSION

We show that by saving the wavefield at the boundary enclosing the reservoir, we can estimate the wavefield at the reservoir location after model perturbation. We eliminate the

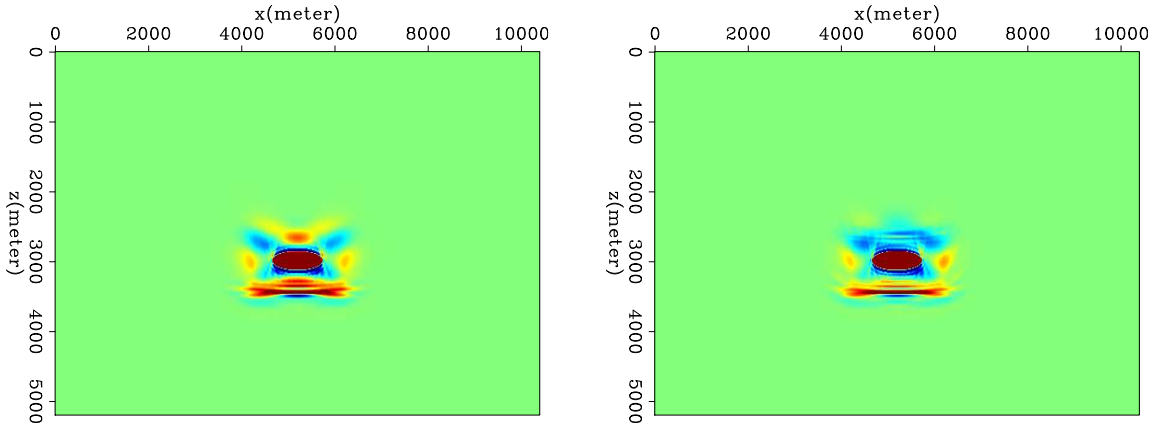


Figure 6: (a) Estimated model perturbation with conventional FWI method (wave propagation in the whole subsurface). (b) Estimated model perturbation using our proposed workflow (wave propagation around the reservoir). [**CR**] yinbin2/. fullFWIa,fullFWIb

need to propagate the wavefield through the overburden area and therefore accelerate the FWI procedure. Random phase-encoding can be incorporated into our proposed workflow and has the potential to further reduce the computational cost.

REFERENCES

Borisov, D. and S. C. Singh, 2013, An efficient 3d elastic full waveform inversion of time-lapse seismic data using grid injection method, 954–958.

Clapp, R., 2008, Reverse time migration: Saving the boundaries: SEP Report 136.

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

Robertsson, J. O. A. and C. H. Chapman, 2000, An efficient method for calculating finite-difference seismograms after model alterations: GEOPHYSICS, **65**, 907–918.

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

Tarantola, A., 1984, in the Acoustic Approximation: Geophysics, **49**, 1259–1266.

Tromp, J., C. Tape, and Q. Liu, 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: Geophysical Journal International, **160**, 195–216.

Virieux, J. and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1.

Willemsen, B., A. Malcolm, and W. Lewis, 2016, A numerically exact local solver applied to salt boundary inversion in seismic full-waveform inversion: Geophysical Journal International, **204**, 1703.

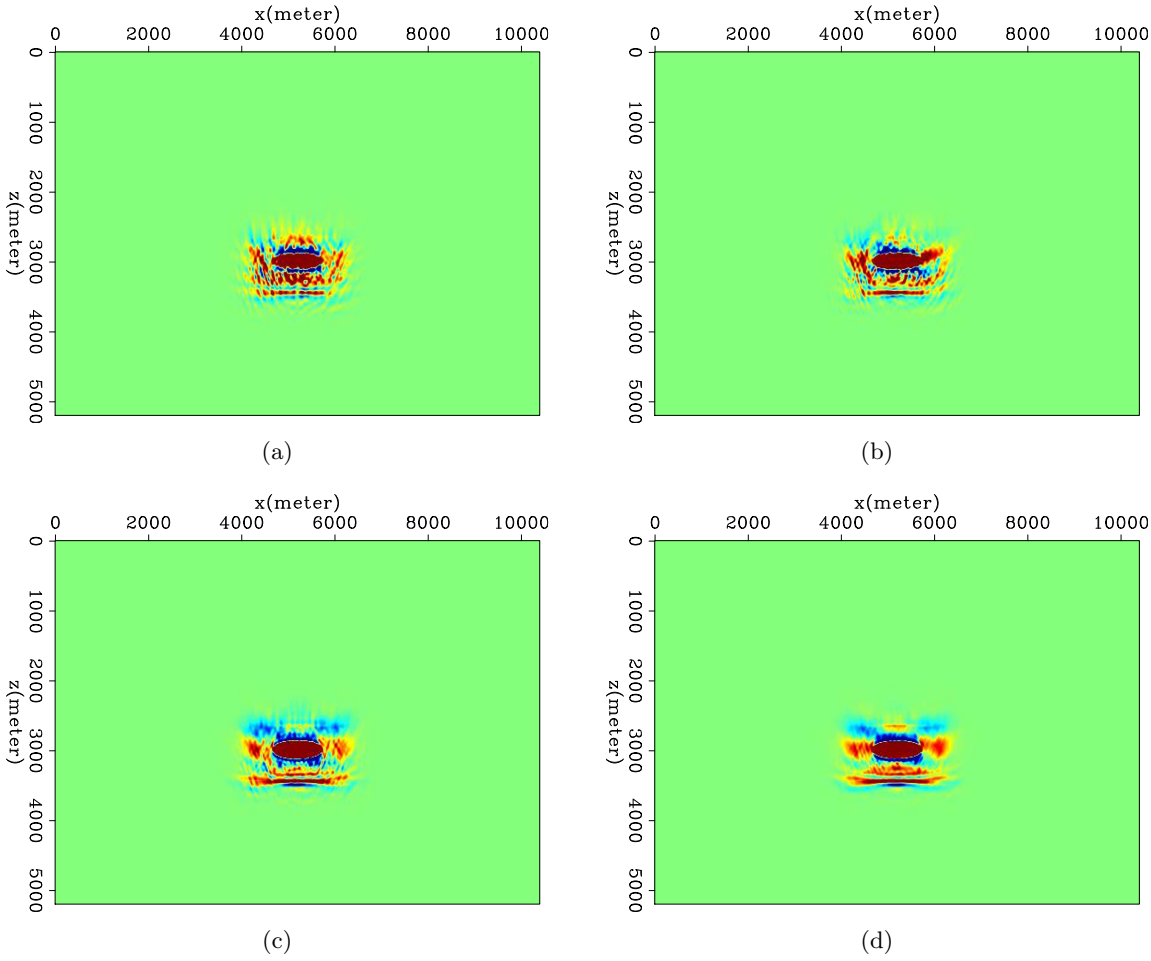


Figure 7: (a) estimated model perturbation with one realization of phase encoding $N_\alpha = 1$. (b) estimated model perturbation with another realization of phase encoding $N_\alpha = 1$. (c) estimated model perturbation with phase encoding $N_\alpha = 5$. (d) estimated model perturbation with phase encoding $N_\alpha = 20$. [**CR**] yinbin2/. phaseFWIa,phaseFWIb,phaseFWIc,phaseFWId

Our progress towards linearized waveform inversion with velocity updating (LWIVU)

Alejandro Cabrales-Vargas, Biondo Biondi, and Robert Clapp

ABSTRACT

Linearized waveform inversion with velocity updating demands the cooperation of Born modeling and wave-equation migration velocity analysis, both forward modeling operators and their adjoints. We successfully tested both of them for passing both the dot-product test and the linearization test. We previously re-organized the wave-propagation codes to operate with random boundary conditions, thus achieving more efficient RAM usage. Assuming the correctness of the operators after the tests, we use the Born modeling pair to illustrate how to pre-compute the Gauss-Newton Hessian by means of point-spread functions. We have now the elements to assemble a robust code for linearized waveform inversion with velocity updating.

INTRODUCTION

We introduced the concept of linearized waveform inversion with velocity updating (LWIVU) in the last SEP reports (Cabrales-Vargas et al., 2016a,b). This technique constitutes an attempt to improve the estimated reflectivity by accounting for remnant inaccuracies in the velocity model which, however too small to induce mispositioning of reflectors, can in turn produce inaccuracies in the estimated reflectivity. Whereas one may disregard such imprecisions as "trifles" for exploration purposes, well positioning in reservoir characterization can significantly rely upon careful and accurate handling of amplitude contrasts, which often reflect reservoir variations such as fluid content, sedimentary facies, dissolution and collapse features, etc. LWIVU can potentially minimize the occurrence of amplitude variations unrelated to reservoir rock properties. Hence, we envision its application mainly to small seismic volumes for detailed studies.

LWIVU is fundamentally based upon the fact that the full waveform inversion (FWI) Hessian (henceforth referred to as "full Hessian") can be split into the sum of the Gauss-Newton Hessian (henceforth referred to simply as "Hessian") and the wave-equation migration velocity analysis (WEMVA) operator (Biondi et al., 2015). These operators are applied to the search direction in the model parameters in the Newton's equation. Regarding the search direction as a *perturbation in the model parameters* we in turn split it into a perturbation in the background model (slowness squared) and a perturbation in the reflectivity. The other side of the Newton's equation, the negative of the gradient of the FWI misfit function, becomes the conventional reverse-time migration (RTM) of the field data. We refer the reader to the aforementioned reports for details. We have very succinctly summarized the derivation here to emphasize that LWIVU demands the interactivity of two operators (with their adjoints) that need be carefully tested, lest numerical errors and/or even unphysical results affect the inversion.

We report the progress in preparing the "pieces of the puzzle", *e.g.*, the operators that conform LWIVU. The aspects discussed in this report are a) Incorporation of random boundary conditions in Born modeling (forward and adjoint) and in WEMVA; b) Linearization test, complementary to the classic dot-product test; c) Use of point-spread-functions (PSF) to compute elements of the Hessian. A significant step, not addressed in this report, is that we re-wrote the codes in modern Fortran, including object-oriented programming. This paradigm improves organization and readability, and potentially facilitates the future transition to C++.

IMPLEMENTATION

Random boundary conditions

Random boundary conditions (RBC) in RTM (Clapp, 2009; Shen and Clapp, 2011; Clapp and Alves, 2016) constitute an alternative to checkpointing (*e.g.*, Symes, 2007). They consist of a halo or extended region around the velocity model. We progressively introduce random variations of velocity when entering such region, increasing randomness as going further inside. At the same time we make the velocity trend ramp down. The overall effect is "scrambling" and delaying the unphysical reflection from the ultimate computational frontier. In RTM we employ RBC in the source wavefield propagation, driving boundary reflections to becoming randomly incoherent from source to source. Stacking the migrated shots reinforces the signal and weakens the random and incoherent reflections from the boundary.

This procedure aims at making the source wavefield propagation reversible in time, unlike propagation using damping boundary conditions (DBC) (Cerjan et al., 1985). Figure 1 illustrates the comparison between RTM performed using DBC (top) and RBC (bottom). DBC require the storage of at least one wavefield (a), usually the source wavefield (orange boxes), as the receiver wavefield (blue boxes) is propagated backward (b-g). Cross-correlations are performed to construct the image (red rectangles). As wavefields may consist of hundreds or thousands of time frames, limitations in RAM or disk storage, particularly in 3-D, are obvious. On the contrary, with RBC we save two time frames of the source wavefield at the final time (a), then propagate it backward jointly with the receiver wavefield and performing the cross-correlations at each time step. Only four time frames are kept in memory during the process. The memory saving using RBC come at the price of an additional propagation of the source wavefield.

We too incorporate RBC in WEMVA. Using DBC, WEMVA requires four propagations: 1) source wavefield; 2) receiver wavefield; 3) source scattered wavefield; and 4) receiver scattered wavefield. With RBC we need an extra propagation of the source wavefield. Figure 2 shows a pseudocode of this procedure in three stages. Note that we use RBC when wavefields need be reversed. We *must not* use DBC in such cases. The scattered wavefields are not reversed, therefore they can be damped at the boundaries. The same situation occurs with the source wavefield during the third stage. Note the order of the injections of source/data and the scaling factors. We follow the scheme analyzed by Almomin (2013) in our RTM and WEMVA implementations, in order to ensure the correct linearization of the operators.

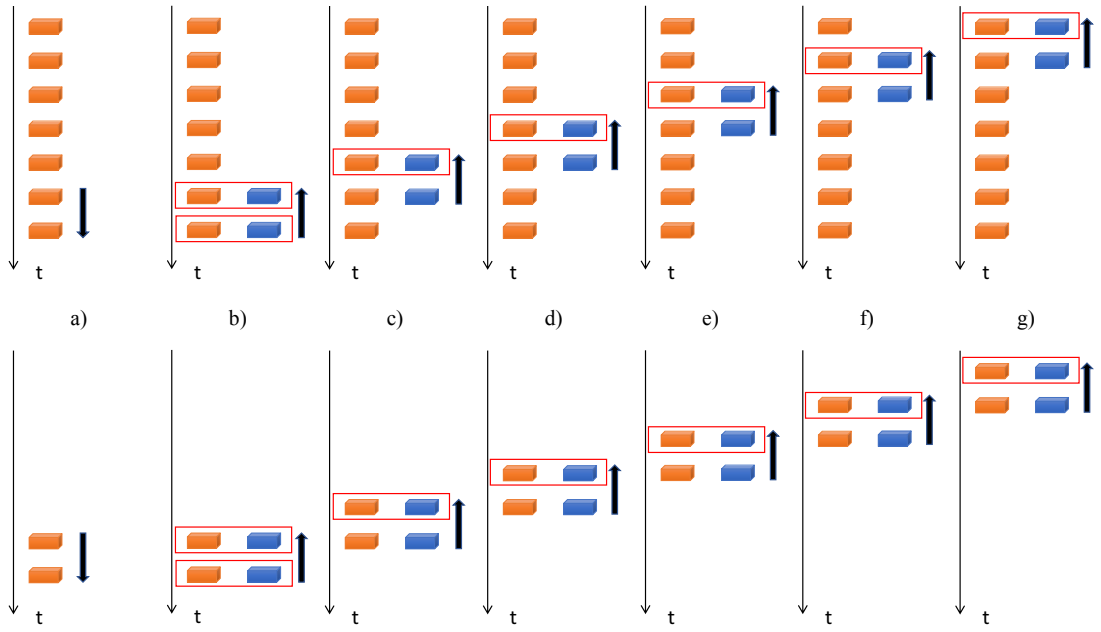


Figure 1: Schematic showing the comparison between RTM implementations using *top*: conventional damping boundary conditions; *bottom*: random boundary conditions. Orange boxes represent source wavefield time frames and blue boxes represent receiver wavefield time frames. See text for explanation. [**NR**]. alejandro1/. RBC

Linearization tests

For many years the SEP crew has been using (almost religiously) the dot-product test in code debugging (Claerbout, 2014) for inversion purposes. The formula is simple: If your code fails the test, this means that your linear forward/adjoint operators are not really adjoint of each other, and you ought to debug. However, if your code passes the test, you might have coded operators which, notwithstanding their mutual adjointness, may not be representative of the physical phenomenon that you intend to linearize. In order to prevent this scenario we perform the linearization test.

The linearization test consists on verifying that our linearized operator, $\mathbf{F}$, resembles the action of the full nonlinear operator, $\mathcal{F}$, in the presence of small perturbations in the model parameters, $\mathbf{\Delta m} = \mathbf{m} - \mathbf{m}_0$:

$$\mathcal{F}(\mathbf{m}) - \mathcal{F}(\mathbf{m}_0) \approx \mathbf{F}\mathbf{\Delta m}, \tag{1}$$

where $\mathbf{m}$ and $\mathbf{m}_0$ are the perturbed model and the background model, respectively.

In preparation for the LWIVU implementation we performed the linearization test on Born modeling and WEMVA after passing the dot-product test. In Born modeling, the perturbed model parameter is slowness squared, $\mathbf{s}^2 = \mathbf{s}_0^2 + \mathbf{\Delta s}^2$, which induces a perturbation in the data, $\mathbf{\Delta d} = \mathcal{L}(\mathbf{s}^2) - \mathcal{L}(\mathbf{s}_0^2) \approx \mathbf{L}\mathbf{\Delta s}^2$, where $\mathcal{L}$ represents the nonlinear acoustic wave-equation modeling operator, and $\mathbf{L}$ represents the Born modeling operator. In WEMVA

```
1st stage: Forward propagation of source wavefield
do in time steps (forward)
   Propagate the source wavefield one time step;
   Scale the source function;
   Inject the source function;
end do

2nd stage: Backward propagation of receiver wavefield+scattering
and cross-correlation with backward propagated source wavefield
do in time steps (backward)
   Inject the data;
   Propagate receiver wavefield and source wavefield one time step;
   Scale source function;
   Re-inject the source function;
   Scale the receiver wavefield;
   Scatter the receiver wavefield;
   Propagate scattered receiver wavefield;
   Damp scattered receiver wavefield at the boundaries;
   WEMVA_receiver_side += xcorr(source wavefield, scattered receiver wavefield)
end do

3rd stage: Forward propagation of source wavefield+scattering
and cross-correlation with forward propagated receiver wavefield
do in time steps (forward)
   Re-inject the data;
   Propagate source wavefield and receiver wavefield one time step;
   Scale the receiver wavefield;
   Scale source function;
   Inject the source function;
   Damp source wavefield at the boundaries;
   Scale the source wavefield;
   Scatter the source wavefield;
   Propagate the scattered source wavefield;
   Damp scattered source wavefield at the boundaries;
   WEMVA_source_side += xcorr(receiver wavefield, scattered source wavefield)
end do
```

Figure 2: Pseudocode for the implementation of WEMVA using RBC [**NR**]. alejandro1/. RBC-WEMVA

modeling, the perturbed model parameter is slowness squared too, this time inducing a perturbation in the migration image, $\mathbf{\Delta I} = \mathbf{I}(\mathbf{s}^2) - \mathbf{I}(\mathbf{s}_0^2) \approx \mathbf{W}\mathbf{\Delta}\mathbf{s}^2$, where $\mathbf{I}(\mathbf{s}^2)$ and $\mathbf{I}(\mathbf{s}_0^2)$ are the migration images corresponding to the perturbed and the background slownesses squared respectively, and $\mathbf{W}$ represents the WEMVA operator.

Figure 3 shows the results of the linearization test for our Born modeling operator. We pertubed a constant slowness squared model with a single impulse that represents a 3% increment of amplitude in the corresponding background velocity model. The corresponding perturbation in the data is a single diffraction. The comparison between the exact result (left panel) and the approximated (linearized) result (right panel) demonstrates that our Born modeling operator effectively represents the linearization of the full wave-equation modeling operator. The panels share the same scaling and are not clipped.

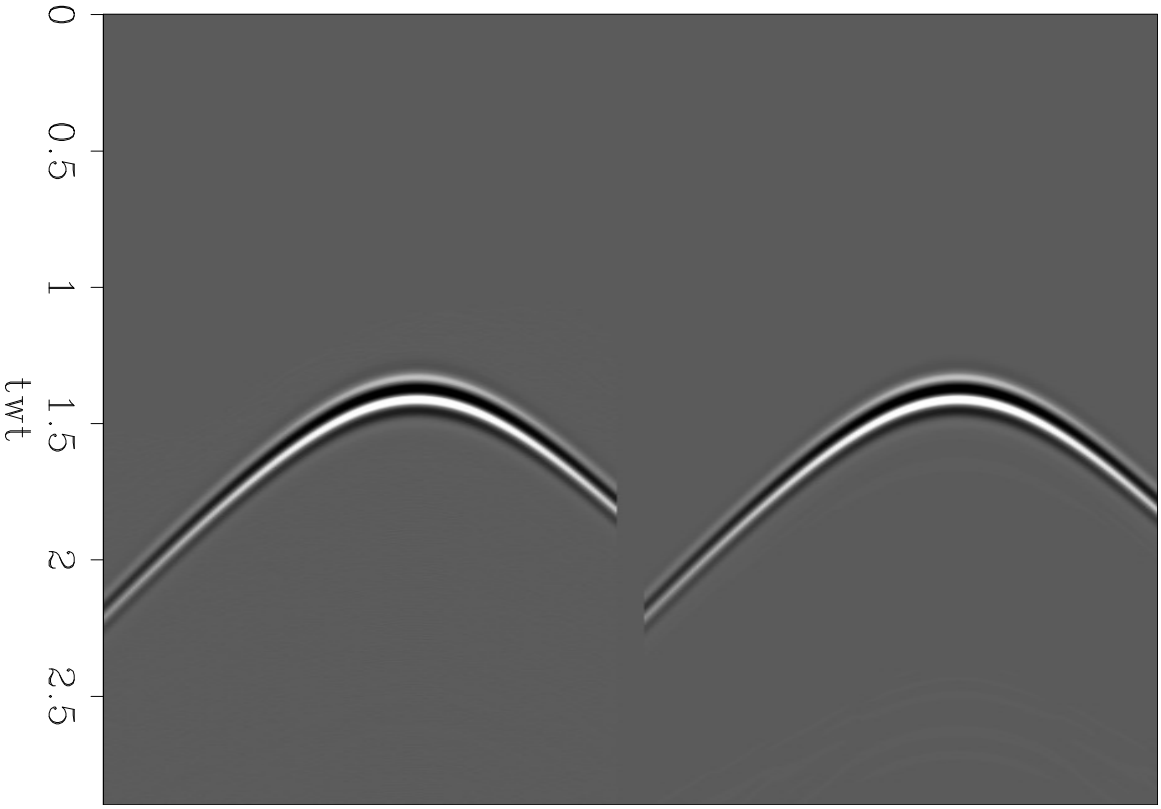


Figure 3: Linearization test of the Born operator. *Left panel*: Perturbation in the data obtained with the full nonlinear wave equation operator. *Right panel*: Perturbation in the data approximated with the Born modeling operator [**ER**]. alejandro1/. lnt-born

Similarly, Figure 4 shows the results of the linearization test for our WEMVA operator. The perturbation in the image is produced by the same perturbation in slowness squared as in the previous test. The dataset employed in this experiment is a single shot gather of the Marmousi model. Note the "banana-doughnut" kernel with foci at the source and the slowness squared perturbation positions. Except for some artifacts at the surface in the WEMVA panel (right), the exact result (left panel) and the approximated result (right panel) of the perturbation in the image are virtually equivalent.

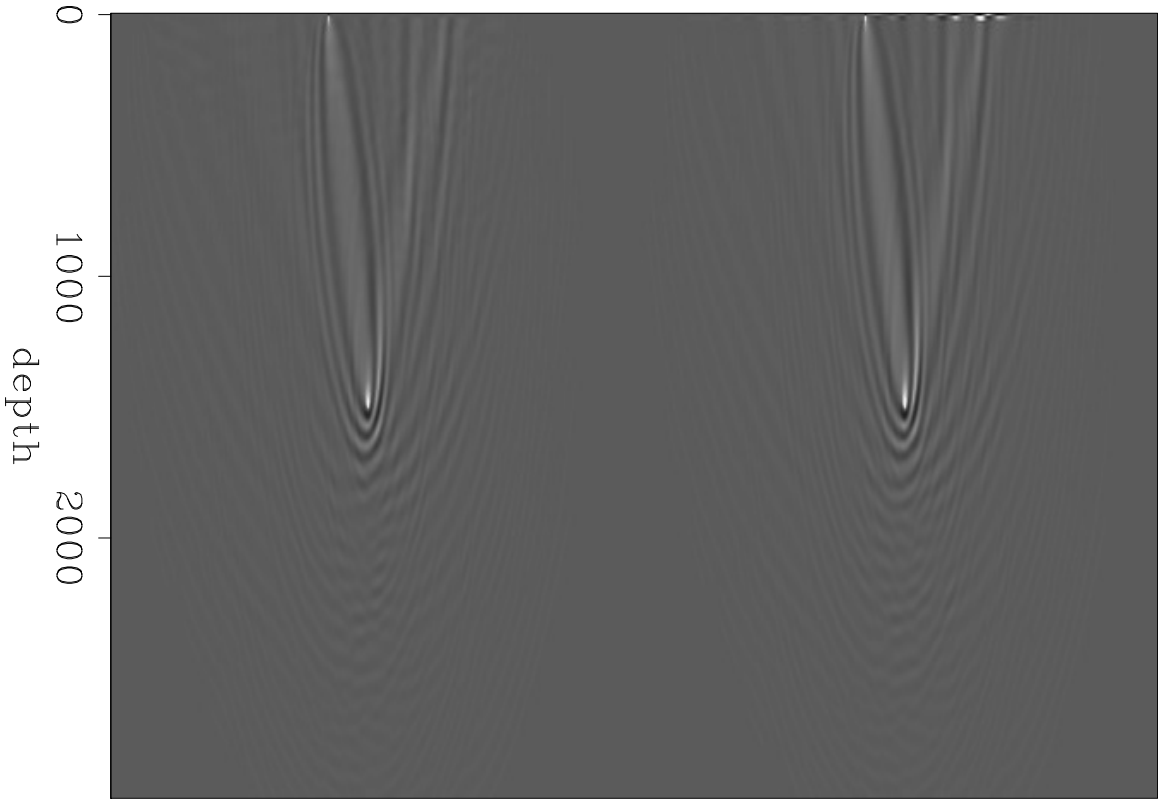


Figure 4: Linearization test of the WEMVA operator. *Left panel*: Perturbation in the image obtained with the RTM operator. *Right panel*: Perturbation in the image approximated with the WEMVA operator [**ER**]. alejandro1/. lnt-wemva

Precomputing the Gauss-Newton Hessian using point spread functions

Performing LWIVU requires precomputation and storage of the Hessian matrix in an affordable manner. For example, Tang (2011) computes Hessian terms using phase-encoding of Green's functions. We instead employ the concept of PSFs (Lecomte and Gelius, 1998; Fletcher et al., 2016) to pre-compute the Hessian. It consists of distributing unitary spikes over the model space, and cascade the Born forward and the Born adjoint modeling operators in order to determine impulse responses of the Hessian at the spikes' positions. Later on, we interpolate both the diagonal and the off-diagonal scattered elements of the Hessian over the model space.

Figure 5 shows a reflectivity model, $\mathbf{m}$, in a 9×9 model space, consisting of a single unitary impulse "seeded" at the center whereas all other elements are zero (hollow circles). The right panel schematically illustrates the PSF (orange circles) resulting from applying Born forward modeling followed by Born adjoint modeling: $\tilde{\mathbf{m}} = [\mathbf{L}^T\mathbf{L}]\mathbf{m}$. The red-bounded element corresponds to the position of the original spike. This PSF constitutes the impulse response of the Hessian ($\mathbf{H} = \mathbf{L}^T\mathbf{L}$) for the centered spike. The spike's amplitude is smeared around a region, oversimplified in Figure 5. The distribution of the smeared zone depends upon the seismic acquisition (the denser the acquisition, the less the smearing) and the frequency content of the source function. The column vectors at the right of the panels are employed for the matrix operation, $\tilde{\mathbf{m}} = \mathbf{H}\mathbf{m}$. Figure 6 illustrates such operation. The Hessian matrix consists of zero column vectors (large ovals) except the central column containing the elements of $\tilde{\mathbf{m}}$. Note that the red-bounded central element represents one element of the diagonal.

We now seed an additional spike as indicated in Figure 7. We use the same response as the central spike, thus the PSFs do not interfere with each other. The corresponding matrix operation is shown in Figure 8. Note that the PSF elements of the additional spike are transferred to the column that coincides with the spike position in the reflectivity vector. Note also that there are no elements of the two PSFs occupying a single row of the Hessian matrix.

Finally, Figure 9 illustrates the case of overlapping PSFs. Interference occurs at the position of the bicolored sample, which belongs to both PSFs. As consequence, in the Hessian matrix this element occupies two positions in a single row (Figure 10). Imperfect differentiation of independent impulse responses introduces innacuracies on the estimation of the Hessian elements. Increasing the sparsity of the spikes would reduce the interference of the PSFs, at the cost of losing geologic details during the interpolation. Splitting the spikes in panels increases the sparsity, avoiding the interference and preserving the details (*e.g.* Figure 11). The drawback is that we have to compute the PSFs in each panel.

We present a synthetic example of PSFs computation in the Marmousi model. Applying forward Born modeling followed by the adjoint (RTM) we obtain "blurred" or smeared spikes (Hu et al., 2001) that conform the PSFs. Figure 12 shows PSFs obtained after seeding spikes every 15 grid points both in depth and distance. Note the interference evidenced in zones between the PSFs, slightly resembling the migration image. As aforementioned, the diagonal of the Hessian matrix corresponds to amplitude values picked at the position of the original input spikes. Likewise, off-diagonal elements of the Hessian are picked at determined lags with respect to diagonal elements. We interpolate afterwards to get Hessian elements all over the model space for each lag, thus creating a four-dimensional array $\mathbf{H}(z, x, lag_z, lag_x)$,

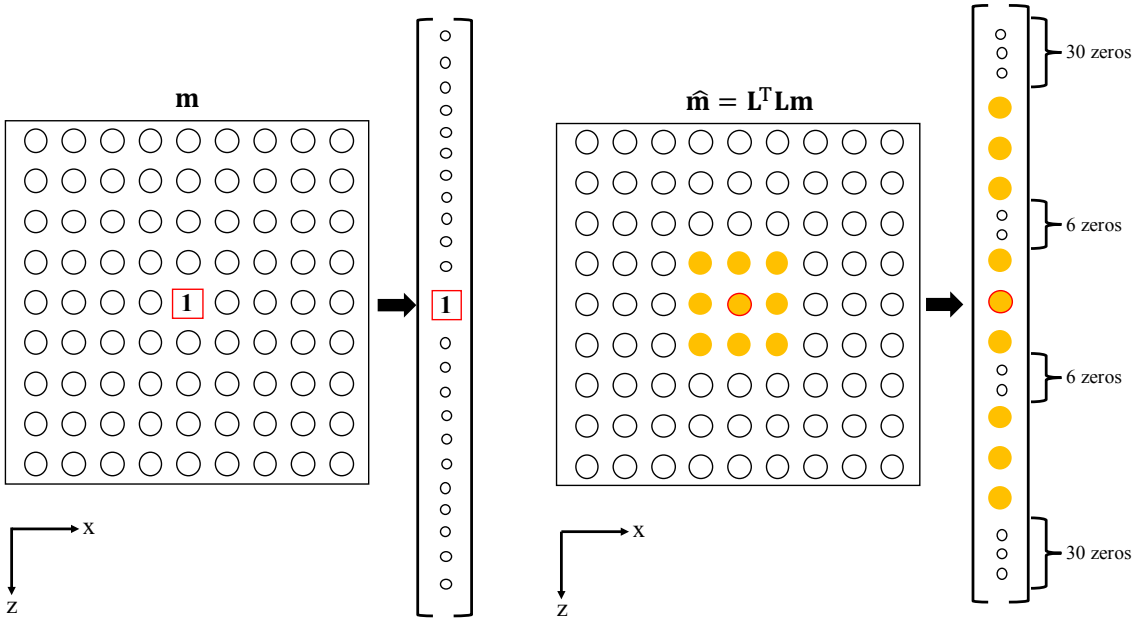


Figure 5: Point-spread function concept illustrated with a single spike in the reflectivity model [**NR**]. alejandro1/. psf-1

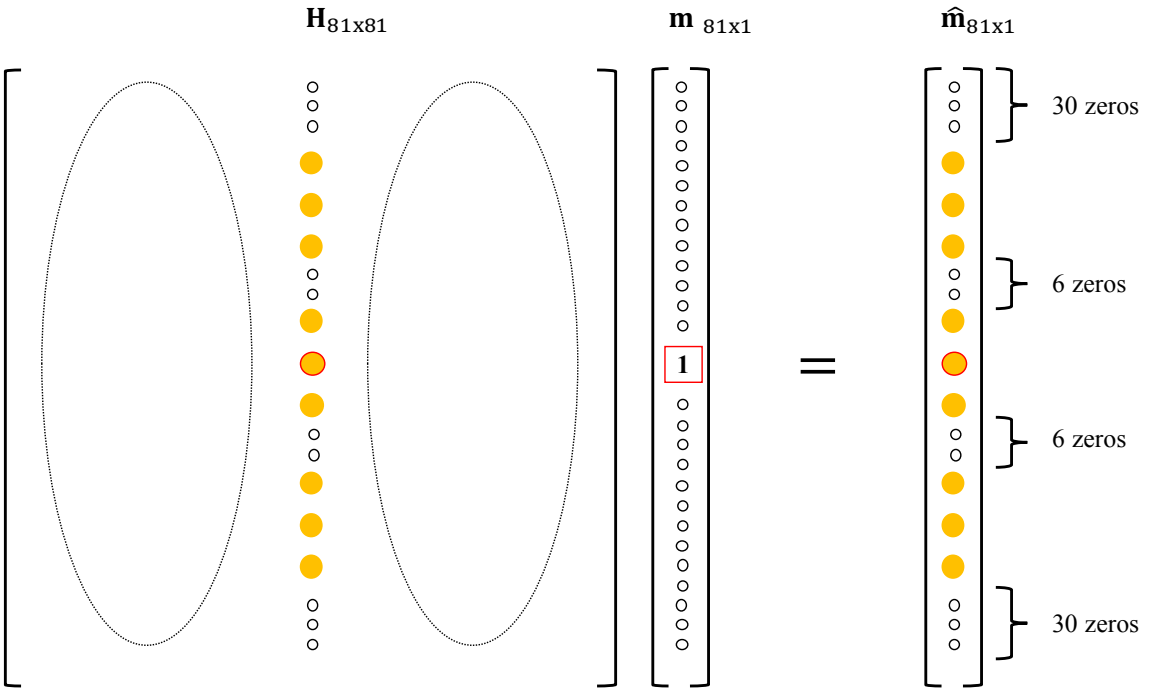


Figure 6: Hessian matrix correponding to the reflectivity model and the PSFs shown in Figure 5. The red-bounded elements lie at the diagonal of the Hessian matrix [**NR**]. alejandro1/. hess-op1

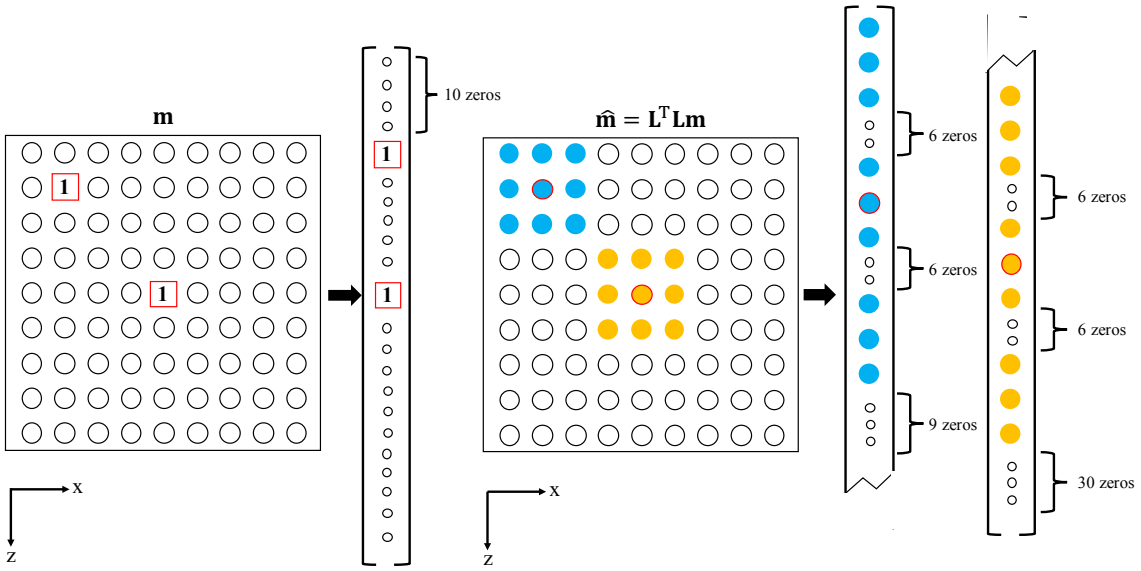


Figure 7: Point-spread function concept illustrated with two non-overlapping spikes in the reflectivity model [**NR**]. alejandro1/. psf-2

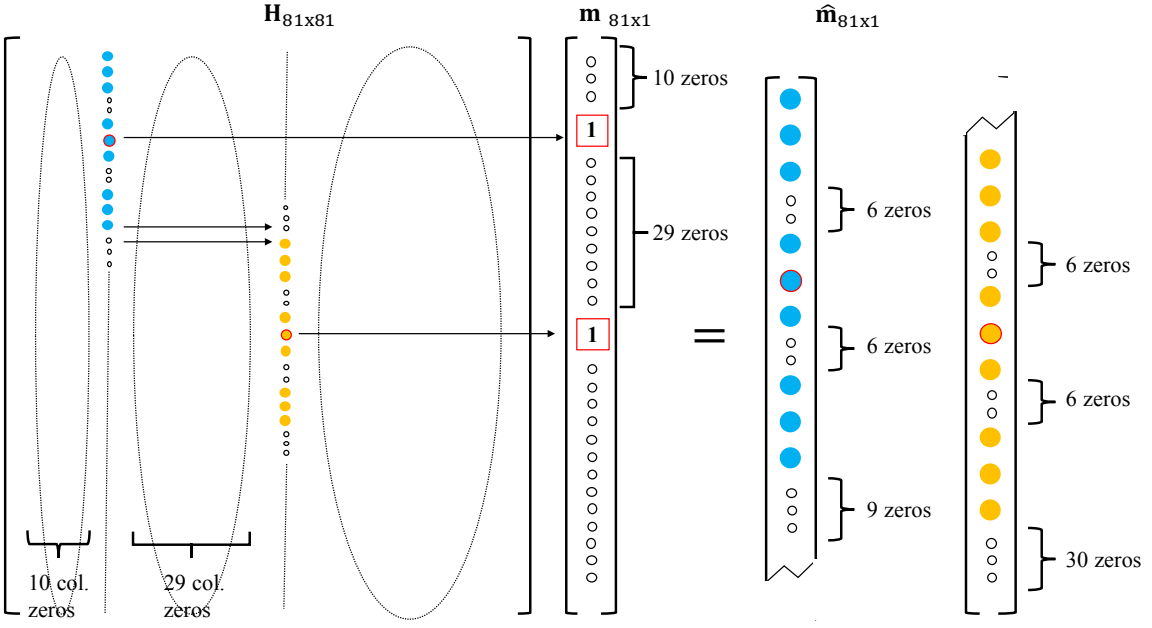


Figure 8: Hessian matrix correponding to the reflectivity model and the PSFs shown in Figure 7. The red-bounded elements lie in the diagonal of the Hessian matrix [**NR**]. alejandro1/. hess-op2

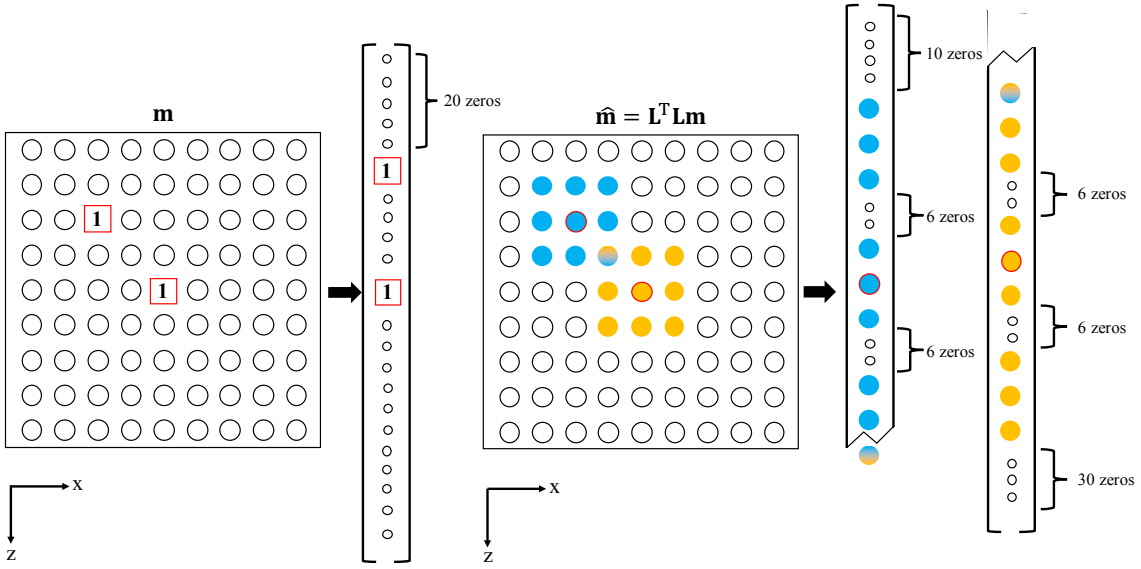


Figure 9: Point-spread function concept illustrated with two overlapping spikes in the reflectivity model [**NR**]. alejandro1/. psf-3

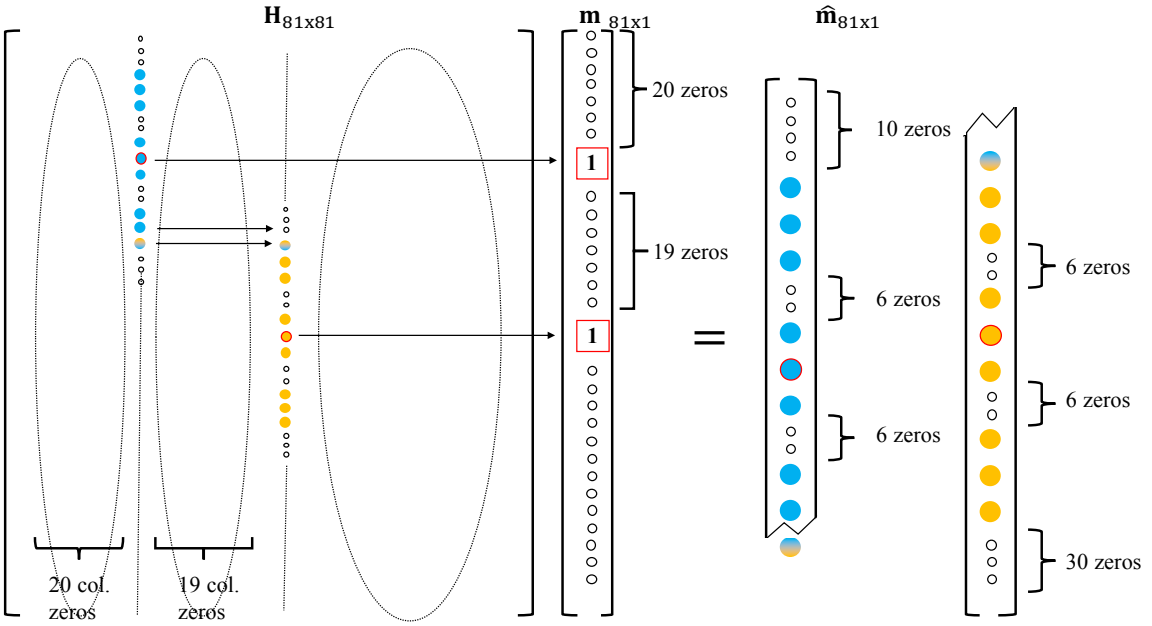


Figure 10: Hessian matrix correponding to the reflectivity model and the PSFs shown in Figure 9. The red-bounded elements lie in the diagonal of the Hessian matrix [**NR**]. alejandro1/. hess-op3

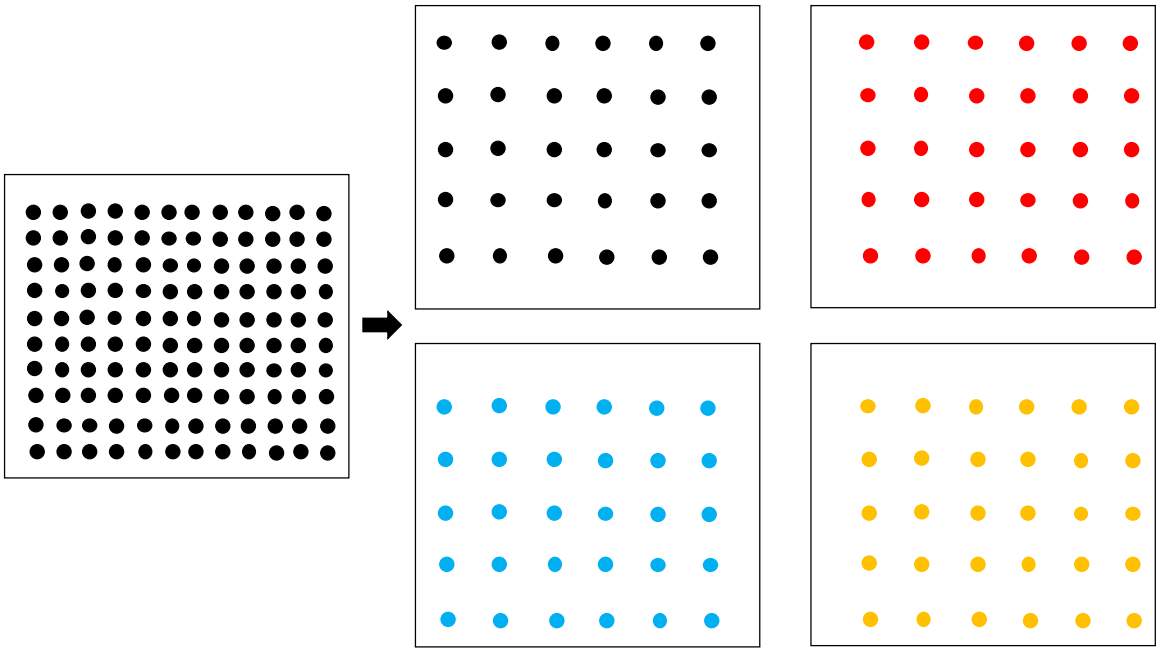


Figure 11: Left: Schematic representation of spikes densely "seeded" for PSF computation. Right: After splitting in four panels, incresing the sparsity to prevent interference [**NR**]. alejandro1/. panels

where lag_x and lag_z are lags with respect to the original spike position. Figure 13 shows the diagonal of the Hessian, *e.g.*, $\mathbf{H}(z, x, 0, 0)$ after 2-D linear interpolation.

Figure 14 shows an approximation to LWI performed in model space, dividing the RTM image by the diagonal of the Hessian (Figure 13). For comparison, Figure 15 shows the result of LWI in data space after 10 iterations. As expected, the former simply constitutes a scaled version of the RTM image. Nonetheless, the amplitude tends to the correct values (compare the scale bars), encouraging us to definitely adopt this technique for the Hessian computation.

FUTURE WORK

In Cabrales-Vargas et al. (2016a,b) we expressed the full Hessian as the sum of the Gauss-Newton Hessian, $\mathbf{H}_{GN}$, and WEMVA, $\mathbf{H}_W$, (Biondi et al., 2015) and expanded the perturbation of the model parameters, $\mathbf{\Delta m}$, into a perturbation in the background model, $\mathbf{\Delta b}$, and a perturbation in the reflectivity, $\mathbf{\Delta r}$:

$$\mathbf{H\Delta m} = (\mathbf{H}_{GN} + \mathbf{H}_W)(\mathbf{\Delta r} + \mathbf{\Delta b}) = (\mathbf{H}_{GN}\mathbf{\Delta r} + \mathbf{H}_W\mathbf{\Delta b}) + (\mathbf{H}_W\mathbf{\Delta r} + \mathbf{H}_{GN}\mathbf{\Delta b}). \quad (2)$$

We kept the first part of the right-hand side of Equation 2, whereas neglected the second part as it contains terms that we do not wish to include in the inversion. We will design and performing numerical experiments aimed at analyzing each term's response, for the sake of understanding the significance of their absence from LWIVU.

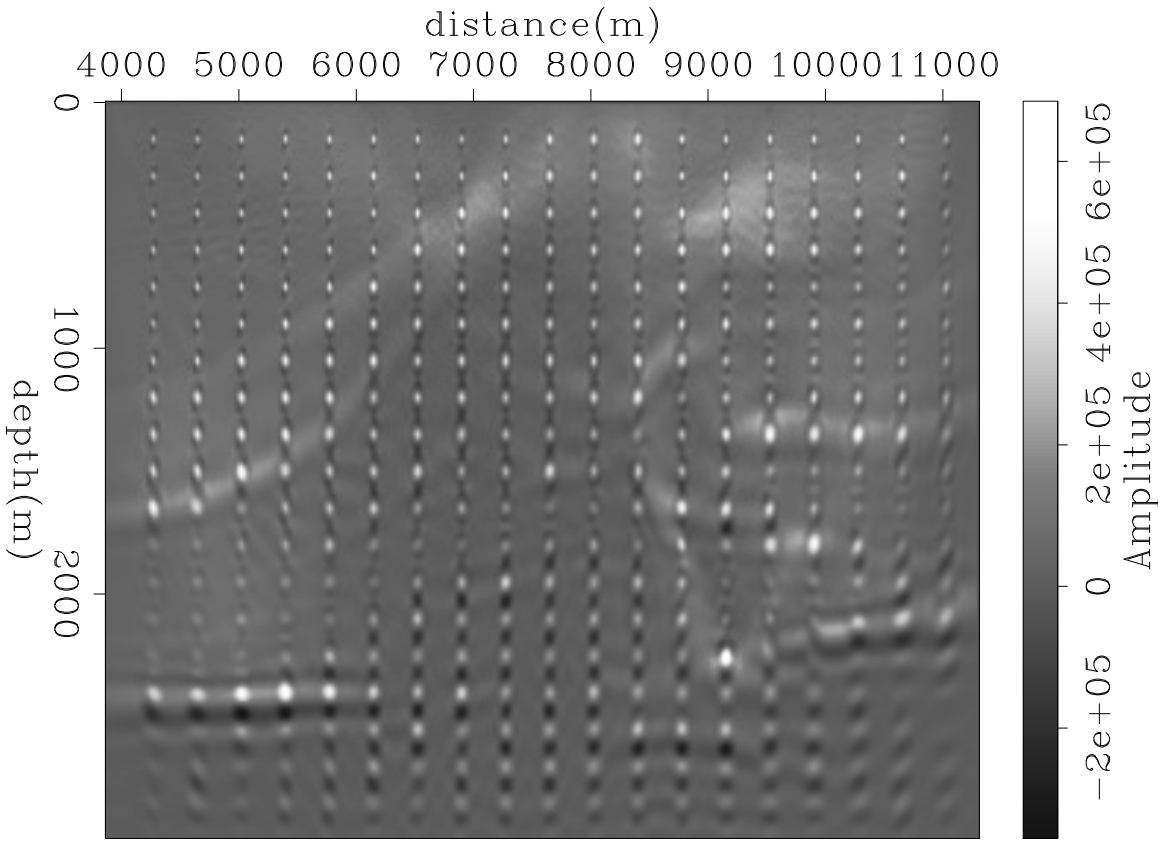


Figure 12: Point-spread functions computed in the Marmousi model. [**ER**]. alejandro1/. Marm-psf-stack

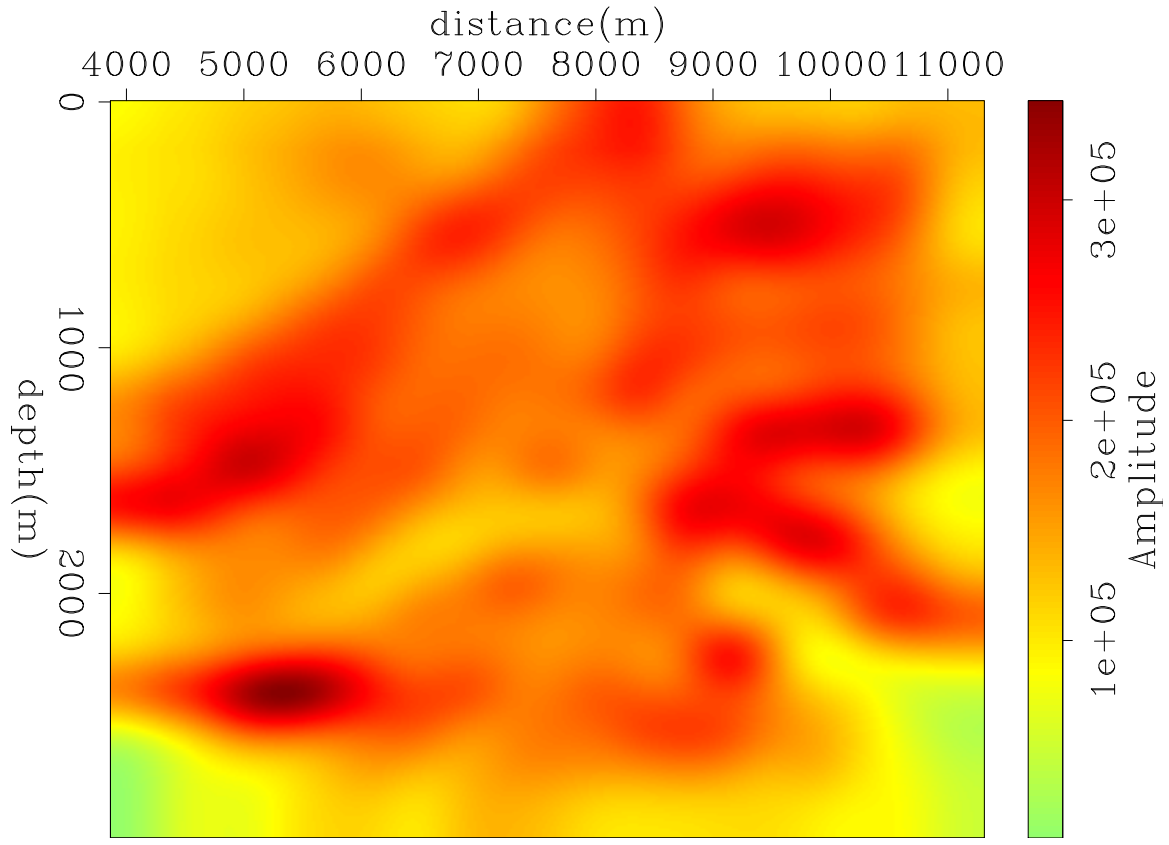


Figure 13: Diagonal of the Hessian approximated from the interpolation of scattered points picked from the PSF (Figure 12). [**ER**]. alejandro1/. Marm-hess-interp

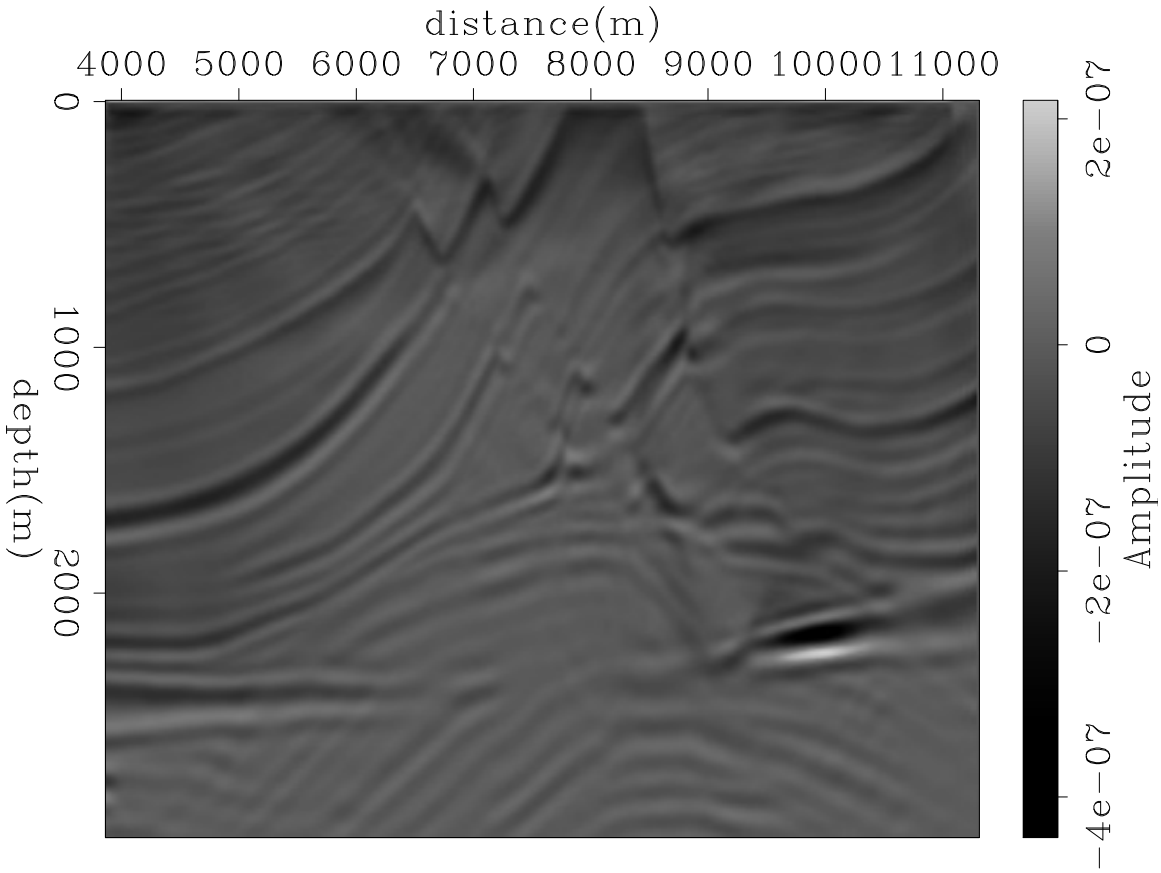


Figure 14: LWI of the Marmousi model in model space, using the diagonal of the Hessian. [**ER**]. alejandro1/. Marm-lsqrtmH-stack

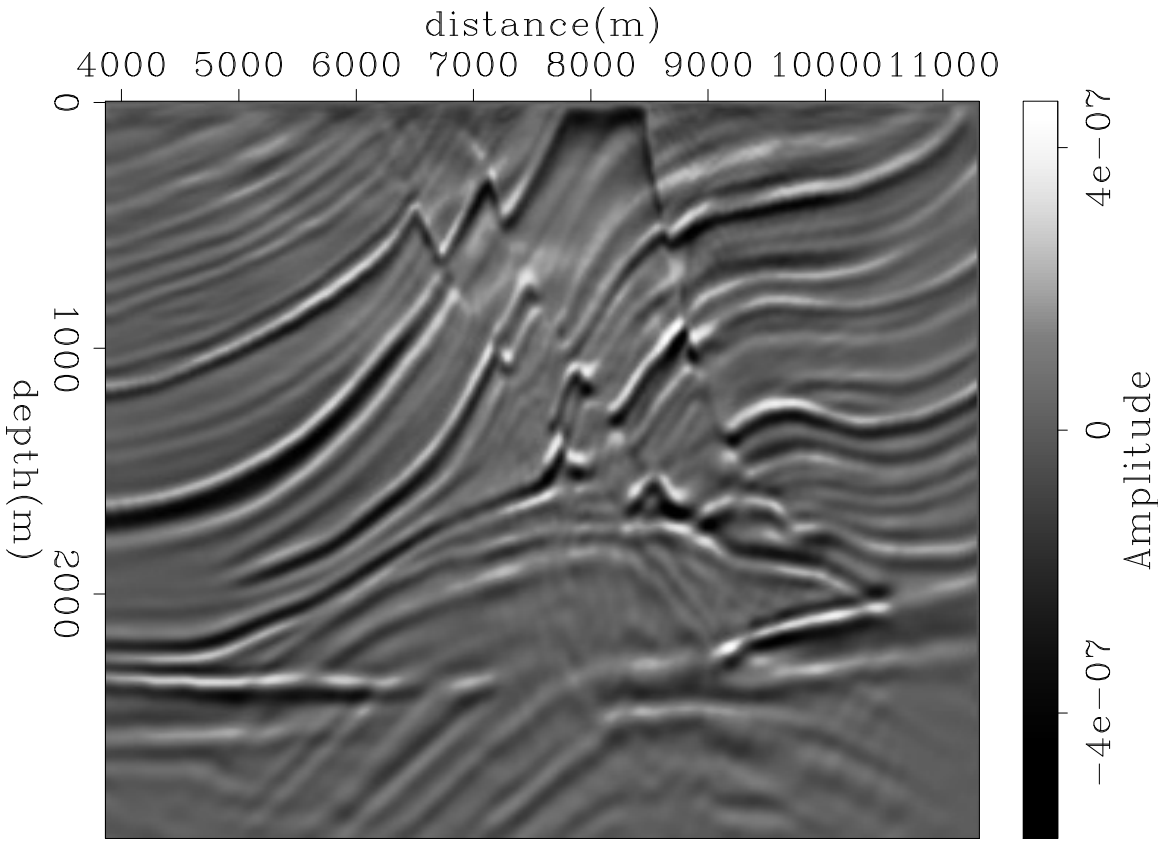


Figure 15: LWI of the Marmousi model in data space. [**CR**]. alejandro1/. Marm-lsqrtm-stack

The next step is the inclusion of off-diagonal elements of the Hessian in the inversion, followed by the construction of LWIVU with the operators shown in this report.

CONCLUSIONS

In this report we showed the progress in the construction of reliable Born modeling and wave-equation migration velocity analysis operators, intended to work together in the linearized waveform inversion with velocity updating process. The incorporation of random boundary conditions hugely reduced the computational storage burden, enabling us to perform future tests with large datasets. With the classical dot-product test and the recently adopted linearization test, we ensure both the "adjointness" of the operators and their correctness in portraying linearization at the presence of small perturbations in the model parameters. Finally, we showed the application of Born forward/adjoint pair in the computation of point-spread functions for fast and practical computation of the Gauss-Newton Hessian.

ACKNOWLEDGEMENTS

We would like to thank the SEP sponsors for their continuous support. Alejandro Cabrales would like to thank Petróleos Mexicanos for financing his studies in Stanford. Special thanks to Ettore Biondi for his assistance and advices with regards of the linearization test.

REFERENCES

Almomin, A., 2013, Accurate implementation of two-way wave-equation operators: SEP-Report, **149**, 281–288.

Biondi, B., E. Biondi, M. Maharramov, and Y. Ma, 2015, Dissection of the full-waveform inversion hessian: SEP-Report, **160**, 19–38.

Cabrales-Vargas, A., B. Biondi, and R. Clapp, 2016a, Linearized waveform inversion with (small) velocity updating : SEP-Report, **163**, 189–196.

———, 2016b, Linearized Waveform Inversion with Velocity Updating: Theory and first results : SEP-Report, **165**, 63–92.

Cerjan, C., D. Kosloff, R. Kosloff, and M. Reshef, 1985, A nonreflecting boundary condition for discrete acoustic and elastic wave equations: Geophysics, **50**, 705–708.

Claerbout, J. F., 2014, Geophysical image estimation by example: Jon Claerbout.

Clapp, R., 2009, Reverse time migration with random boundaries: SEG Technical Program Expanded Abstracts, 2809–2813.

Clapp, R. G. and G. Alves, 2016, Random boundaries for elastic medium : SEP-Report, **163**, 41–48.

Fletcher, R., D. Nichols, R. Bloor, and R. Coates, 2016, Least-squares migation - data domain versus image domain using point spread functions: The Leading Edge, **35**, no. 2, 157–162.

Hu, J., G. Schuster, and P. Valasek, 2001, Poststack migration deconvolution: Geophysics, **66**, 939–952.

Lecomte, I. and L. Gelius, 1998, Have a look at the resolution of prestack depth migration for any model, survey and wavefields: SEG Technical Program Expanded Abstracts, 1112–1115.

Shen, X. and R. Clapp, 2011, Random boundary condition for low-frequency wave propagation: SEG Technical Program Expanded Abstracts, 2962–2965.

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

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

Understanding the difference between full Newton and Gauss-Newton approximation of the full waveform inversion Hessian: a short note

Ettore Biondi, Biondo Biondi, and Guillaume Barnier

ABSTRACT

We continue the mixed theoretical-computational study of the difference between full Newton and Gauss-Newton approximation of the Hessian matrix in the context of full waveform inversion (FWI) started in the SEP report 160. We also continue to use an acoustic isotropic wave equation approximation during our discussion. We analyze the connection of the residual-dependent component of the full Hessian with the physical double scattering described by the wave equation. We explain how to avoid inversion instabilities of the full Newton Hessian when this matrix is not positive definitive. With the help of a simple two-perturbation model we study the advantages of full Newton compared to the Gauss-Newton approximation.

INTRODUCTION

Improving convergence rate of any FWI method is fundamental given the high computational cost of the required propagations during the problem optimization (Virieux and Operto, 2009). For this reason many authors have explored the use of the Hessian matrix as a potential preconditioner to reach the problems' optimal solution in the least number of iterations (Pratt et al., 1998; Epanomeritakis et al., 2008; Tang, 2008; Korta et al., 2013; Deuzeman and Plessix, 2015). Therefore, it is critical to understand the connection of each component of this matrix with the physical behavior of any wave equation to find new approximations to be used to precondition FWI optimization.

Newton-like algorithms for non-linear optimization are employed to account for the local curvature of the objective function such that the inverted models present a comparable resolution (i.e., the ability to retrieve a given parameter during the inversion) between model parameters. The differential parameter resolution in the context of multi-parameter FWI is know as inversion crosstalk or trade-off (Operto et al., 2013). Therefore, it is critical to develop an optimal inversion method that utilizes Hessian matrix information.
A detailed computational review of Hessian-based preconditioned FWI has been presented by Métivier et al. (2013). In their work the authors describe different algorithms to perform truncated Newton optimization steps. In their synthetic acoustic constant-density Marmousi test they find that truncated Gauss-Newton performs better than full Newton approximation. They attribute this difference to slower convergence rate of full Newton matrix during its linear inversion. On the other hand, when applying truncated full Newton to a simpler near-surface FWI problem, the inverted model presents higher spatial resolution when compare to truncated Gauss-Newton using the same number of iterations when inverting the Hessian matrix. From this study, it is therefore unclear whether full Newton has

any advantage compare to Gauss-Newton in the case of seismic exploration data inversion. In this report we continue the study shown in Biondi et al. (2015), and review the connection of the full Hessian with the double-scattering mechanism present in any wave equation. We also compare least-square full Newton and Gauss-Newton Hessian matrix inversions for a simple two-perturbation model.

BRIEF ANALYSIS OF THE FWI HESSIAN MATRIX

We start our analysis by defining the classical amplitude-matching FWI objective function as follows:

$$\phi(\mathbf{m}) = \frac{1}{2}\|\mathbf{f}(\mathbf{m}) - \mathbf{d}\|_2^2 = \frac{1}{2}\|\mathbf{r}(\mathbf{m})\|_2^2, \tag{1}$$

where $\mathbf{m}$ represents the model parameters such as wave-propagation velocities, attenuation, or density, $\mathbf{f}(\mathbf{m})$ is the wave-equation modeling operator, and $\mathbf{d}$ is the observed data. The difference between the latter two is represented by the data residual vector $\mathbf{r}(\mathbf{m})$. It is known that we can decompose our observed data into an infinite scattering series (Weglein et al., 2003). This representation of the data provides an intuitive way to look at the non-linearity present in a typical FWI problem. To do so we can expand the data vector as follows:

$$\mathbf{d} = \mathbf{f}(\mathbf{m}_0) + \frac{\partial \mathbf{f}(\mathbf{m}_0)}{\partial \mathbf{m}}\Delta\mathbf{m} + \frac{1}{2}\Delta\mathbf{m}^*\frac{\partial^2 \mathbf{f}(\mathbf{m}_0)}{\partial \mathbf{m}^2}\Delta\mathbf{m} + \dots, \tag{2}$$

where $\mathbf{m}_0$ is the background model, and $\Delta\mathbf{m}$ is the difference between the true model and the background. This infinite series is effectively a multivariate Taylor expansion of the data vector. The first two terms on the right-hand side of the equation represent the propagation in the background and the first-order scattering (i.e., Born approximation). The additional terms in the series represent the multiple scattering (e.g., multiples and non-linear amplitude behavior).

In the case $\mathbf{m}_0$ is our starting model for a FWI algorithm we see that our initial residuals are given by the terms on the right of the background data $\mathbf{f}(\mathbf{m}_0)$. If the residuals are approximately linear, meaning:

$$\mathbf{r}(\mathbf{m}_0) = \mathbf{d} - \mathbf{f}(\mathbf{m}_0) \approx \frac{\partial \mathbf{f}(\mathbf{m}_0)}{\partial \mathbf{m}}\Delta\mathbf{m}; \tag{3}$$

then, the inverse problem is almost linear, and thus easily solvable up to the null space of the linearized operator. This condition is usually satisfied when the initial model is reasonably close to the true one, meaning the background model explains most of the observed data in terms of both kinematic and amplitude responses (Biondi and Almomin, 2014).

When a Newton-like scheme is used to minimize equation 1 the following truncated expansion around the initial model is performed:

$$\phi(\mathbf{m}_0 + \Delta\mathbf{m}) \approx \phi(\mathbf{m}_0) + \frac{\partial \phi(\mathbf{m}_0)}{\partial \mathbf{m}}\Delta\mathbf{m} + \frac{1}{2}\Delta\mathbf{m}^*\frac{\partial^2 \phi(\mathbf{m}_0)}{\partial \mathbf{m}^2}\Delta\mathbf{m}. \tag{4}$$

The main approximation in the previous relation is to assume a convex quadratic behavior of the objective function close to the initial model. We already see that if this assumption is not fulfilled, any Newton method may fail completely to find the local minimum of the

objective function. The key for the success of this class of algorithms resides in the Hessian matrix of the objective function $\frac{\partial^2\phi(\mathbf{m}_0)}{\partial\mathbf{m}^2}$ being positive definitive. In fact, in any Newton-like optimization we try to solve the following linear system:

$$\frac{\partial^2\phi(\mathbf{m}_0)}{\partial\mathbf{m}^2}\Delta\mathbf{m} = -\frac{\partial\phi(\mathbf{m}_0)}{\partial\mathbf{m}}. \tag{5}$$

This system is obtained by setting the derivative of equation 4 with respect to the model perturbation $\Delta\mathbf{m}$ to zero. It is known that the full Hessian matrix can be decomposed into its Gauss-Newton approximation plus an additional residual-dependent component (Fichtner, 2010). In fact, we observed that:

$$\frac{\partial^2\phi(\mathbf{m})}{\partial\mathbf{m}^2} = \left(\frac{\partial\mathbf{f}(\mathbf{m})}{\partial\mathbf{m}}\right)^* \left(\frac{\partial\mathbf{f}(\mathbf{m})}{\partial\mathbf{m}}\right) + \left(\frac{\partial^2\mathbf{f}(\mathbf{m})}{\partial\mathbf{m}^2}\right)^* \mathbf{r}(\mathbf{m}), \tag{6}$$

where the first term is composed of the product between forward and adjoint Born operator, and the second one is connected to the second-order scattering in the series of equation 2. Because of this connection many authors see an additional value to the full Hessian when a Newton-like method is employed in solving FWI problems (Pratt et al., 1998; Epanomeritakis et al., 2008). It is still under debate whether this additional value is meaningful or not in seismic exploration (Métivier et al., 2013). In addition, as shown by Biondi et al. (2015) the residual-dependent component of the Hessian matrix is connected to the wave-equation migration velocity analysis (WEMVA) operator (Sava and Vlad, 2008).

We now discuss a few considerations when a single full-Newton step is performed during full waveform inversion. The Newton system of equations at the first iteration is given by:

$$\left[\left(\frac{\partial\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}}\right)^* \left(\frac{\partial\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}}\right) + \left(\frac{\partial^2\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}^2}\right)^* \mathbf{r}(\mathbf{m}_0)\right] \Delta\tilde{\mathbf{m}} = -\left(\frac{\partial\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}}\right)^* \mathbf{r}(\mathbf{m}_0). \tag{7}$$

The ideal solution of this linear inversion would be:

$$\Delta\tilde{\mathbf{m}} = \Delta\mathbf{m}, \tag{8}$$

meaning we converge to the true perturbation of the data expansion around the initial FWI model (equation 2) in one Newton step. Unless our problem is close to being linear this result is difficult to find. Assuming that most of the non-linearities of an FWI problem are given by the second-order scattering, we can truncate the Taylor expansion to the first three terms and make the inversion closer to linearity. The initial residuals are thus given by the following:

$$\mathbf{r}(\mathbf{m}_0) = -\frac{\partial\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}}\Delta\mathbf{m} - \frac{1}{2}\Delta\mathbf{m}^*\frac{\partial^2\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}^2}\Delta\mathbf{m}. \tag{9}$$

Substituting equation 9 into 7 we obtain:

$$\begin{aligned}&\left[\left(\frac{\partial^2\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}^2}\right)^* \left(-\frac{\partial\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}}\Delta\mathbf{m} - \frac{1}{2}\Delta\mathbf{m}^*\frac{\partial^2\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}^2}\Delta\mathbf{m}\right)\right] \Delta\tilde{\mathbf{m}} \\ &= \frac{1}{2}\left(\frac{\partial\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}}\right)^* \Delta\mathbf{m}^*\frac{\partial^2\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}^2}\Delta\mathbf{m},\end{aligned} \tag{10}$$

where we notice that the right-hand side term is basically the image produced by the first-order multiples in the data. The term on the left is more complicated, but can be

decomposed into two different operators; a WEMVA-like operator applied to the true Born data and the second-order scattering, respectively. We will converge in a single Newton step to the correct solution if equality 8 holds. In the next section we explore these considerations on a simple two-perturbation synthetic test.

SYNTHETIC TEST

In this section, we show an example of FWI in which a single truncated Newton step is performed on a simple synthetic model where two perturbations of 100 m/s are added to a constant velocity background of 2000 m/s (Figure 1). We inject a Ricker wavelet with central frequency of 20 Hz and use evenly spaced receivers and sources at the surface by 10 m and 100 m, respectively. A shot gather containing reflection only data is shown in Figure 2(a). The choice of this model introduces a mild non-linearity when the correct background velocity is considered as starting model for the inversion. In fact, comparing the observed data with the correct Born approximation (i.e., second term in the expansion 2) we clearly see that most of the recorded amplitudes are represented by this linear modeling operator (Figure 2(b)). The comparison of the difference between non-linear data and Born approximation with the second-order scattering term in the series provides us with an insight of the non-linearities in the treated FWI problem. From Figure 2(c) we observe that most of the higher-order scattering is contained in the second-order one. The amplitudes of these events are approximately one order of magnitude smaller than the first-order approximation (compare Figures 2(b) and 2(c)). Because of these observations we decide to consider only first- and second-order scattering in this test.

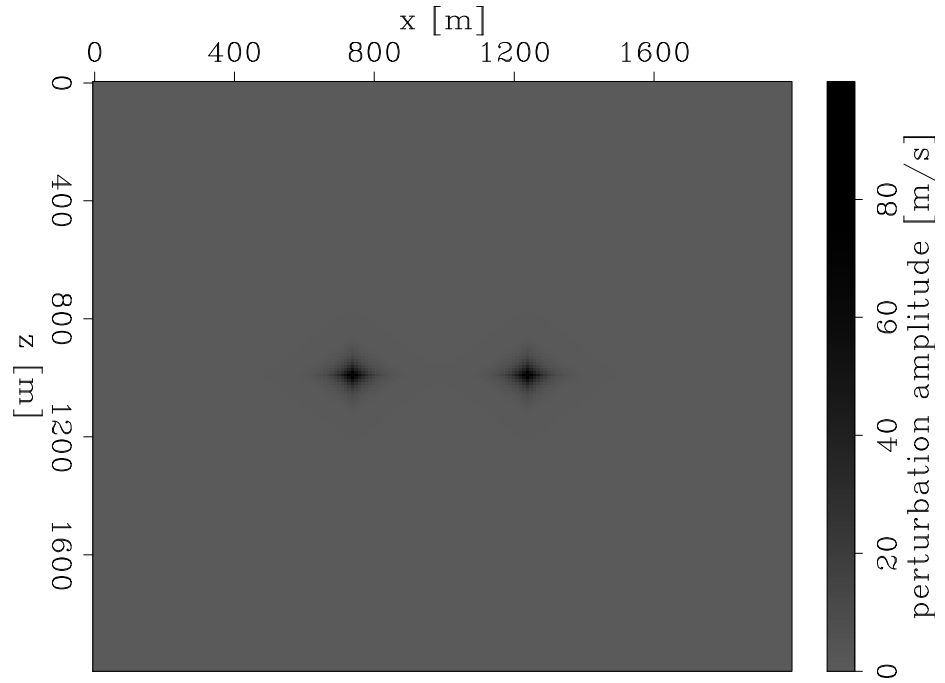
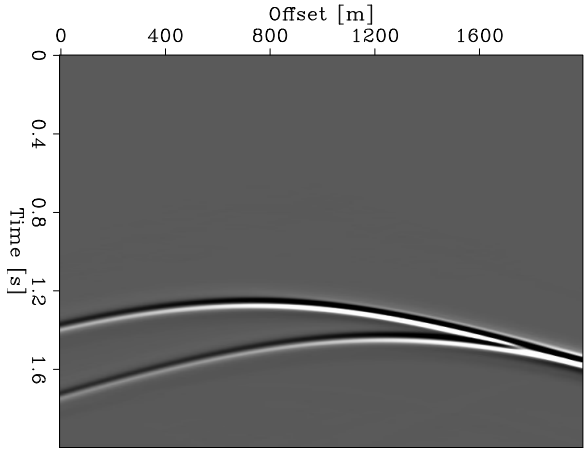
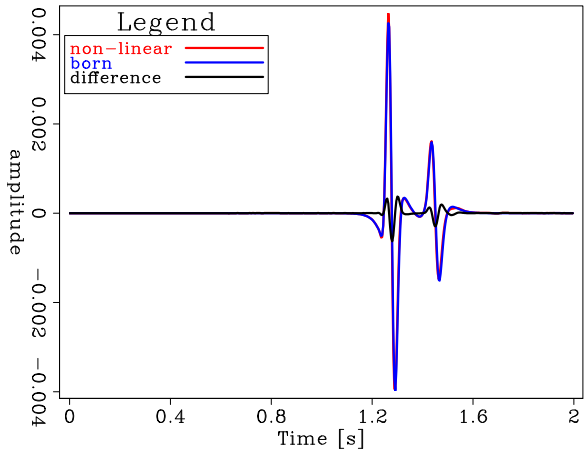


Figure 1: True velocity perturbation used during the reported synthetic test. The background model is a constant velocity of 2000 m/s. [**ER**] ettore3/. true-pert

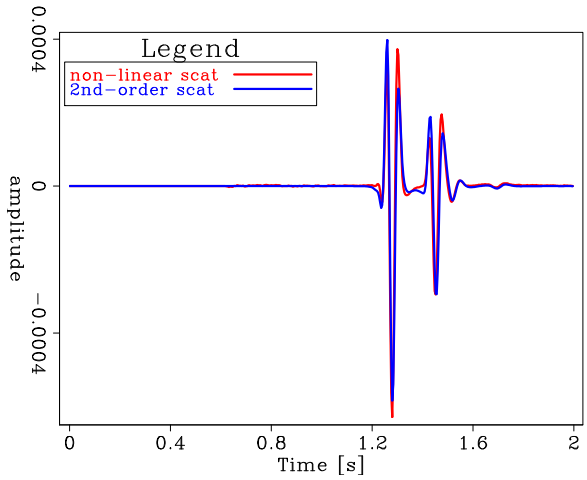
First of all we compute the initial search direction given the recorded data. Figure 3(a)



(a)



(b)



(c)

Figure 2: (a) Reflection data observed for one shot when the velocity perturbation in Figure 1 is used. (b) Comparison between non-linear data (red curve) and Born approximation (blue curve) using the true perturbation for a receiver at 1000 m. The black curve represents the higher-order scattering effects present in the observed data. (c) Comparison between true second-order scattering and black curve in Figure 2(b). The direct arrival has been removed because we assume to start the inversion with the correct background. [**ER**]

ettore3/. obs-data,data-born,ddata-tomo

shows the search direction obtained using both orders of scattering. The maximum of this search direction corresponds to the peaks of the used perturbations. To understand the effect of the presence of second-order scattering in the data we compute the image of this energy (Figure 3(b)). We first see that the relative amplitude is approximately 90 percent weaker compared to the total image. In addition, the phase of the imaged points is slightly different than the one in the total search direction. This plot describes the image mainly produced by the self-interaction of the scattered wavefield with the perturbation. Because of the chosen perturbation, the image of the interaction between the two anomalies is almost negligible.

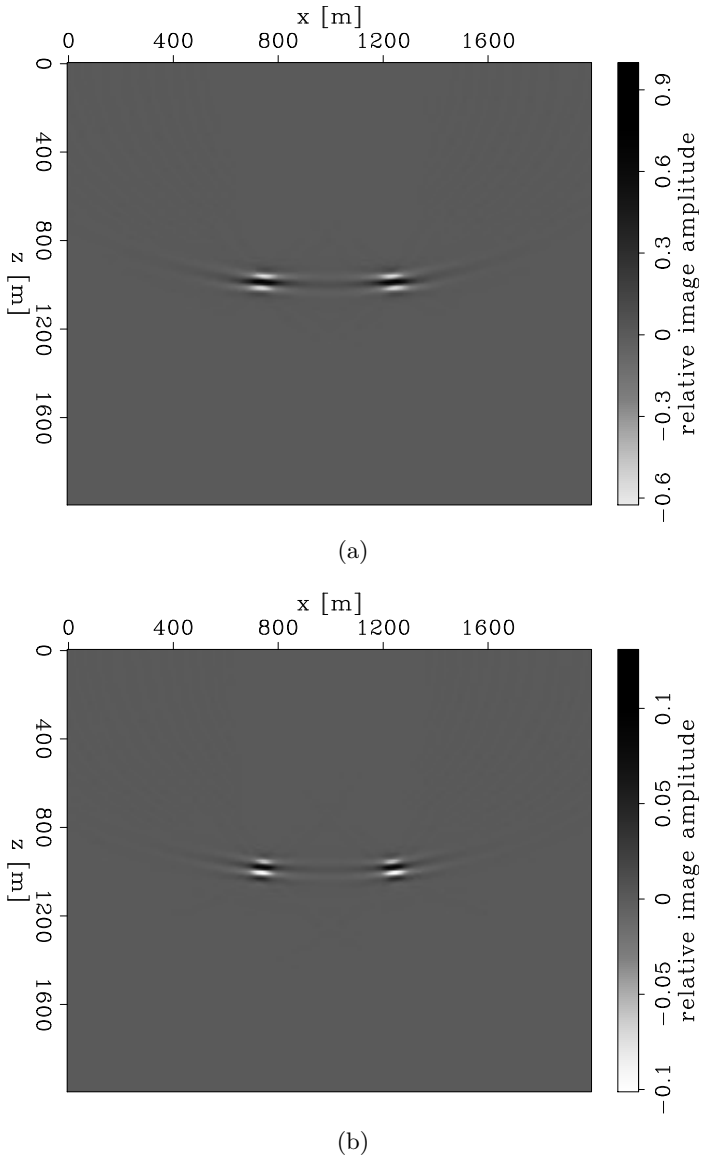
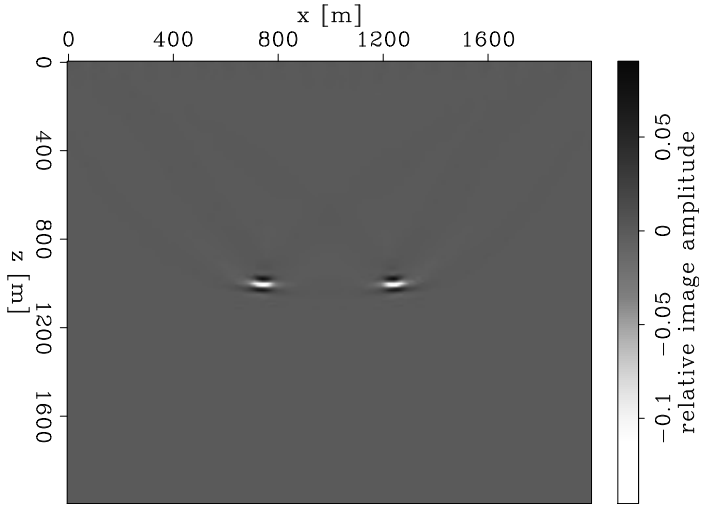
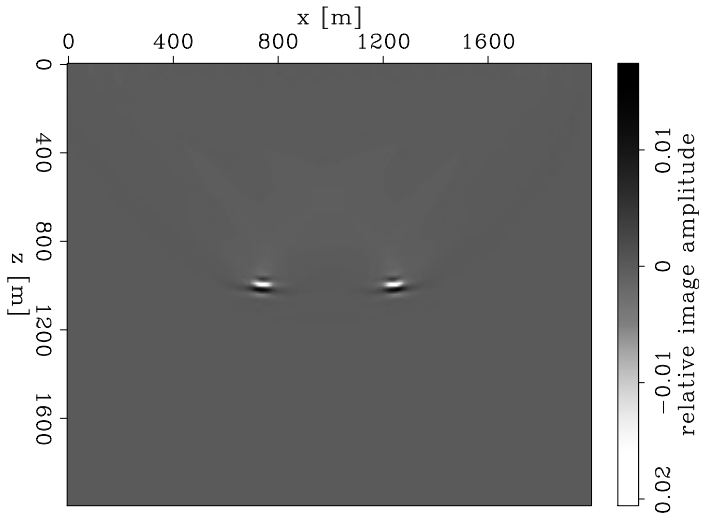


Figure 3: (a) First search direction when only first- and second-order scattering is considered to be present in the recorded data. (b) First search direction due to the second-order scattering in the data (i.e., right-hand side term in equation 10). [**ER**] ettore3/. total-search,tomo-search

As explained in the previous section, the hope of converging to the original model perturbation in a single Newton step depends on matching the image of the multiples by the application of the residual-dependent Hessian component to the true perturbation (equation 10). Figure 4(a) displays this application. Despite the similar amplitude behavior, the phase of the image is quite different; thus, it is unlikely that a single Newton step will provide the correct answer even though the mild non-linearity of the problem. Again, the contribution of the second-order scattering in the application of this Hessian component is weaker with respect to the first-order scattering residuals (compare Figures 4(a) and 4(b)).



(a)



(b)

Figure 4: Application of the residual-dependent hessian component to the true perturbation using the total residuals (a) and the second-order scattering component (b). The amplitudes are normalized with respect to the image shown in Figure 3(a) [**ER**] ettore3/. wemva-total,wemva-tomo

We test two different approximations of the FWI Hessian matrix against the full Newton

one. We compare Gauss-Newton, full Newton, and a Hessian approximation obtained by a matrix expansion. In fact, by decomposing the full Hessian $\mathbf{H}$ into its Gauss-Newton $\mathbf{H}_{GN}$ and residual-dependent $\mathbf{H}_r$ components, we can write the following inverse approximation:

$$(\mathbf{H})^{-1} = (\mathbf{H}_{GN} + \mathbf{H}_r)^{-1} = \left(\mathbf{I} + \mathbf{H}_{GN}^{-1}\mathbf{H}_r\right)^{-1} \mathbf{H}_{GN}^{-1} \approx \left(\mathbf{I} - \mathbf{H}_{GN}^{-1}\mathbf{H}_r\right) \mathbf{H}_{GN}^{-1}, \tag{11}$$

where we assume that the inverse for the Gauss Newton component exists, and that the matrix norm $\|\mathbf{H}_{GN}^{-1}\mathbf{H}_r\|$ is negligible compared to the norm of an identity matrix with the same size. From this equation we see that a correction factor, which is given by the product of a WEMVA-like operator and Gauss-Newton inverse matrix, is subtracted from the inverted perturbation.

In all of these linear inversions the different matrices are solved iteratively in the least-squares sense as follows:

$$\phi(\Delta\mathbf{m}) = \frac{1}{2} \left\|\mathbf{H}(\mathbf{m}_0)\Delta\mathbf{m} - \Delta\mathbf{m}_{mig}(\mathbf{m}_0)\right\|_2^2, \tag{12}$$

where $\mathbf{H}$ is the Hessian matrix to be inverted, and $\Delta\mathbf{m}_{mig}(\mathbf{m}_0)$ is the first FWI search direction (i.e., Figure 3(a) in this test). We choose this approach to mitigate the problem of the full Hessian matrix being symmetric indefinite (i.e., having positive and negative eigenvalues); an issue that has been also observed by Métivier et al. (2013). In fact, when we employ the symmetric solver proposed by Biondi and Barnier (2017) on the full Hessian matrix, after few iterations, the inversion starts diverging and becoming unstable. This least-squares approach stabilizes the inversion, although with this approach we are squaring the condition number of the actual inverted matrix.

In the first Hessian inversion we consider a Gauss-Newton approximation. Figure 5 shows the relative variation of the objective function 12 when 300 iterations of conjugate-gradient algorithm are performed. At the last iteration the objective function value is five orders of magnitude smaller than the initial one. Figure 6(a) shows the inverted perturbation at the last iteration. First, we observe that the amplitude of the anomalies is underestimated. Secondly, despite the considerably decrease in the objective function, the wavelet signature is still present in the inverted anomalies. When this inverted perturbation is reprojected non-linearly in the data space, the modeled amplitudes closely match the recorded ones (Figure 6(b)), meaning that most of the linear component of the data has been inverted and potentially we are encountering a null space of the linear operator. We then inverted iteratively the full Newton hessian and the matrix shown in equation 11. Figure 6(c) shows the comparison of these inversions on a depth profile passing through the left-hand anomaly. From this comparison we observe that practically there is no difference between full Newton and Gauss-Newton inversion results. The similar behavior of the solutions could be attributed to the values of the Gauss-Newton matrix elements that dominate the inversion. On the other hand, we the full Newton approximation shown in equation 11 provides a model slightly closer to the true perturbation. This fact potentially means that the correction term in this approximation could enhance the additional value of the full Newton matrix in the residual-dependent component and avoid the instabilities possibly present in the full Hessian. In fact, in this case we are inverting the more stable positive semi-definitive Gauss-Newton matrix (see equation 11).

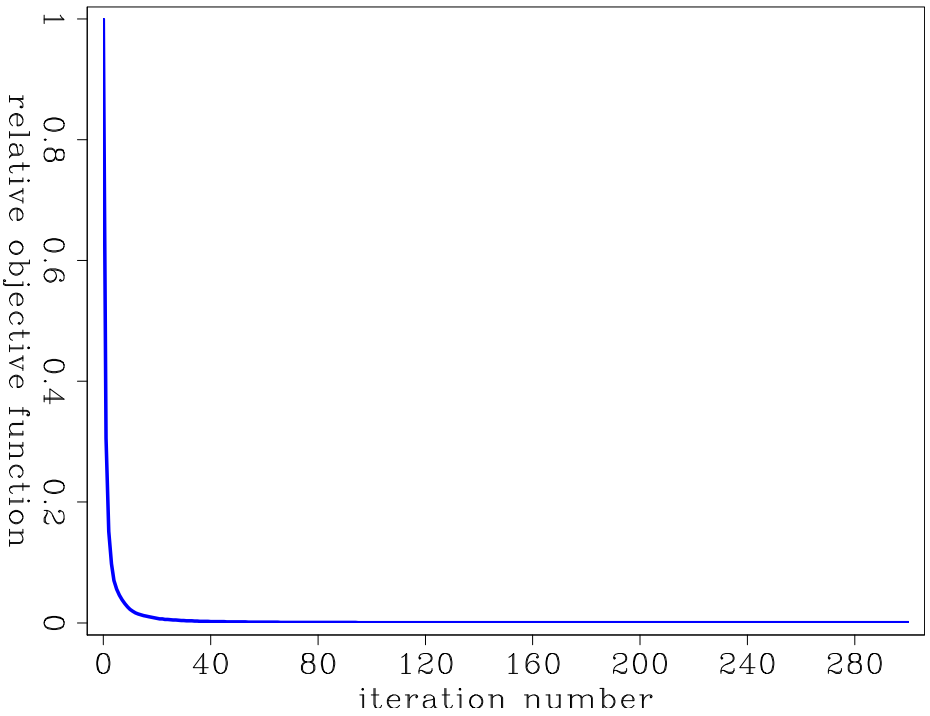


Figure 5: Relative objective function value for least-squares inversion of Gauss-Newton Hessian matrix. A similar behavior is seen for all the other described truncated Hessian inversions. [**CR**] ettore3/. obj-GN

CONCLUSIONS AND FUTURE DIRECTIONS

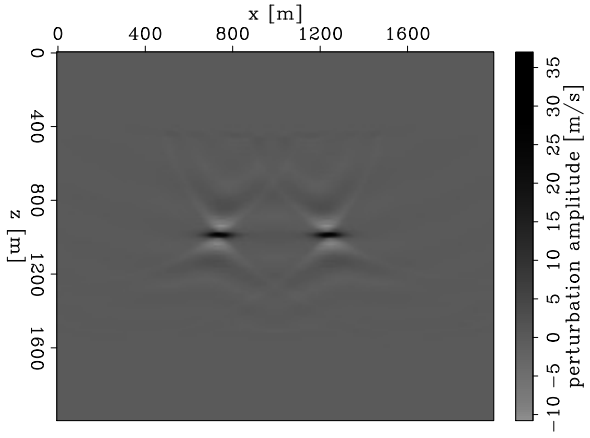
We discuss the difference between the full Hessian and its Gauss-Newton approximation in the context of FWI. We show the connection of the residual-component of this matrix to the second-order scattering present in the data. We also discuss the value given by truncated Newton steps and when those could potentially converge to the true perturbation in one single matrix inversion. With the help of a simple two-anomaly model we study the effect of a single Hessian matrix inversion step. Moreover, with the proposed approximation of the full Newton matrix we are able to enhance the residual-dependent component and avoid unstable inversion results. In the future, we are going to study the advantage of performing truncated Newton steps during a non-linear inversion framework. In addition, further studies of the proposed full Hessian approximation are going to be made to better evaluate its potential in more complex inversion scenarios.

REFERENCES

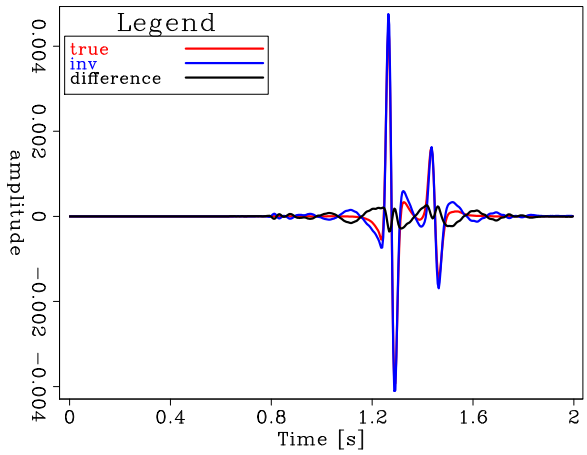
Biondi, B. and A. Almomin, 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: Geophysics, **79**, WA129–WA140.

Biondi, B., E. Biondi, M. Maharramov, and Y. Ma, 2015, Dissection of the full-waveform inversion Hessian: SEP-Report, **160**, 81–100.

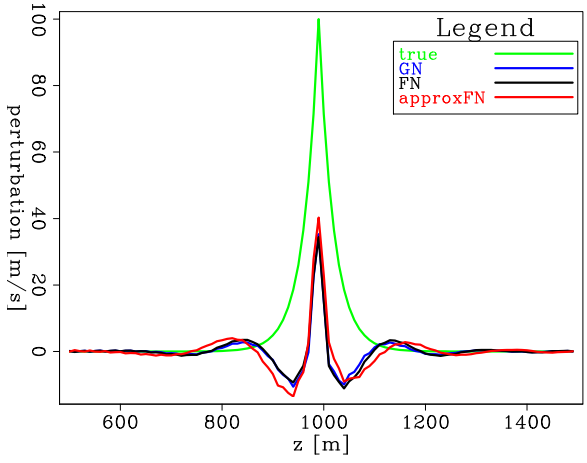
Biondi, E. and G. Barnier, 2017, A flexible out-of-core solver for linear/non-linear problems: SEP-Report, **168**.



(a)



(b)



(c)

Figure 6: (a) Inverted perturbations using a Gauss-Newton approximated Hessian. (b) Comparison between observed data (red curve), data generated using the inverted perturbations (blue curve) in Figure 6(a), and their difference (black curve). (c) Comparison between Gauss-Newton (blue curve), full Newton (black curve), and approximated full Newton (red curve) of the inverted perturbations against the true anomalies (green curve). [**CR**] ettore3/. gn-pert,compare-data-gn,compare-inversions

Deuzeman, A. and R.-E. Plessix, 2015, Block-diagonal approximation of the Hessian for multi-parameter FWI: 12th SIAM Conference on Mathematical and Computational Issues in Geosciences, Expanded Abstracts, MS2, Society of Industrial and Applied Mathematics.

Epanomeritakis, I., V. Akçelik, O. Ghattas, and J. Bielak, 2008, A Newton-CG method for large-scale three-dimensional elastic full-waveform seismic inversion: Inverse Problems, **24**, 034015.

Fichtner, A., 2010, Full seismic waveform modelling and inversion: Springer Science & Business Media.

Korta, N., A. Fichtner, and V. Sallarčs, 2013, Block-diagonal approximate hessian for preconditioning in full waveform inversion: Presented at the 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013.

Métivier, L., R. Brossier, J. Virieux, and S. Operto, 2013, Full waveform inversion and the truncated Newton method: SIAM Journal on Scientific Computing, **35**, B401–B437.

Operto, S., Y. Gholami, V. Prieux, A. Ribodetti, R. Brossier, L. Metivier, and J. Virieux, 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice: The Leading Edge, **32**, 1040–1054.

Pratt, R. G., C. Shin, and G. Hick, 1998, Gauss–newton and full newton methods in frequency–space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

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

Tang, Y., 2008, Wave-equation hessian by phase encoding: SEG Technical Program Expanded Abstracts 2008, 2201–2205, Society of Exploration Geophysicists.

Virieux, J. and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.

Weglein, A. B., F. V. Araújo, P. M. Carvalho, R. H. Stolt, K. H. Matson, R. T. Coates, D. Corrigan, D. J. Foster, S. A. Shaw, and H. Zhang, 2003, Inverse scattering series and seismic exploration: Inverse problems, **19**, R27.

Incorporating optimal transport in Tomographic Full Waveform Inversion : Theory

Rahul Sarkar and Biondo Biondi

ABSTRACT

In recent years, a great deal of work has gone into devising optimization strategies that are more robust to cycle skipping as compared to the conventional Full Waveform Inversion (FWI) objective function. Tomographic full waveform inversion (TFWI) is one such technique that involves the use of a non-physical extension of the whole velocity model and an extended modeling operator that is better capable of modeling the observed data. Inversion algorithms based on this concept have been shown to converge to reasonable models on field data, but with slow convergence rate. Another interesting direction of research that has emerged in this field is based on the use of optimal transport objective functions. These formulations have also been shown to be capable of converging to correct models from relatively inaccurate starting models, leveraging the fact that objective functions based on optimal transport distances are much less susceptible to local minima compared to their L^2 counterparts. In this report, we investigate an inversion strategy based on the idea of velocity model extension that attempts to leverage optimal transport distances to overcome the challenges of TFWI convergence. We derive the problem formulation and present some preliminary results in 2D for the acoustic case.

INTRODUCTION

The non-convexity of the acoustic FWI problem leads to gradient-descent based inversion schemes based getting stuck in local minima, unless one starts very close to the true velocity model. Even in the 2D case, the high dimensionality of the problem precludes brute force stochastic optimization techniques for typical model sizes of interest in the seismic industry. Several strategies, which can be broadly grouped into three categories, have been proposed and experimented with over the past few decades to overcome these problems.

The first category consists of frequency continuation (Pratt, 1999) and multiscale (Bunks et al., 1995) methods where the FWI inversion is performed in overlapping frequency bands, starting from low frequencies and progressively moving to higher frequencies. These schemes are motivated by the fact that for any velocity model, the modeled data and observed data will stop cycle skipping at sufficiently low frequencies. This type of approach works nicely if there is sufficient low frequencies in the data with high enough signal to noise ratio. However, due to limitations imposed by the physics of seismic sources, this requirement is quite often not fulfilled and valuable low frequency content is missing in the observed data.

The second category involves constructing objective functions which have a larger basin of attraction around the global minima, and are thus more "convex" with respect to the velocity model. Objective functions based on cross-correlation (Luo and Schuster, 1991),

dynamic-warping (Ma and Hale, 2013) and deconvolution (Luo and Sava, 2011) are examples that fall in this class of methods. However, the inverse problem of estimating model parameters from surface measurements is intrinsically non-convex because of the oscillatory nature of the source wavelet, and it may not be possible to design an objective function that is convex over the whole feasible set of velocities. However, these methods demonstrate that it may be possible to achieve some form of convexity over a subset of the feasible set that is still larger than that with conventional FWI. Another interesting development that belongs to this class of finding increasingly convex objective functions comes from the use of optimal transport based misfit functions (Engquist and Froese, 2014; Métivier et al., 2016; Yang et al., 2016). These misfit functions are more commonly known as Wasserstein distance functions, and they have been reported to be much more robust at avoiding local minima and also have better convergence properties.

The third category of methods involve extending the reflectivity model, such as extending subsurface offset and/or time lag, in the migrated image domain. The idea uses an extended imaging condition to form migrated images along the extended dimension and then attempts to focus the energy at zero along the extended axes with a misfit function. This is the concept behind approaches such as differential semblance optimization (Symes and Carazzone, 1991; Shen and Symes, 2008) which tries to focus the image at zero subsurface offset, or an equivalent formulation that attempts to flatten the image gathers in scattering angle (Sava and Biondi, 2004a,b; Zhang and Biondi, 2013). More recently, an interesting alternative has been proposed that involves extension of the whole velocity model (Symes, 2008; Biondi and Almomin, 2014) and is named tomographic full waveform inversion (TFWI). However, the speed of convergence of these methods can be quite slow for large scale problems (Almomin and Biondi, 2014).

In this paper, we develop the theoretical formulation of a new optimization scheme based on the model extension approach but one that attempts to overcome some of the convergence problems of TFWI, by leveraging the properties of the optimal transport objective function. The theory is developed for the 2D case, and its generalization to 3D is straightforward but computationally challenging. The new algorithm is motivated by the original TFWI algorithm proposed by Biondi and Almomin (2014) and follows it closely, but has some important differences in the way it attempts to focus the extended model at the origin of the extended model axis. Along the way, we illustrate the concept with some motivating examples and preliminary numerical results.

BACKGROUND

We first provide a concise introduction to the extended modeling operator in 2D for acoustic wave propagation that defines the forward problem. We will introduce some of the intuition behind such a formulation. Next, we briefly review Kolmogorov spectral factorization that is used to find the minimum phase equivalent of signal. We finish this section with a basic introduction to optimal transportation theory, with emphasis on the quadratic Wasserstein distance.

The acoustic wave equation

Let us represent the 2D spatial variables by $\mathbf{x} = (x, z)$ and time by t. We consider the acoustic wave equation in a constant density medium, written in the following form:

$$u_{tt}(\mathbf{x}, t) - v^2(\mathbf{x})\nabla^2 u(\mathbf{x}, t) = f(t)\delta(\mathbf{x} - \mathbf{x}_s) \ , t \geq 0 \ . \tag{1}$$

Here $u(\mathbf{x}, t)$ denotes the pressure at position $\mathbf{x}$ and time t, and ∇^2 denotes the 2D Laplacian operator $(\partial_x^2 + \partial_z^2)$. The source term $f(t)\delta(\mathbf{x}-\mathbf{x}_s)$ denotes a point source located at position $\mathbf{x}_s$ in space with a time dependent source function $f(t)$. The velocity of the medium is heterogenous and is denoted by $v(\mathbf{x})$. To compress notation going forward, we will drop the explicit dependence of the quantities appearing in equation (1) on space and time variables, and simply use u to denote $u(\mathbf{x}, t)$, m to denote the square of the velocity $v^2(\mathbf{x})$, and f to denote the source function $f(t)\delta(\mathbf{x} - \mathbf{x}_s)$. The dependence on space and time variables will be understood from context, and explicit dependence on these variables will only be given in cases deemed absolutely necessary. We will also refer to the quantity m as the "acoustic model parameter".

The solution u to equation (1) is given by the convolution of the source function with the Green's function (the solution to equation (1) when $f(t) = \delta(t)$), and is linear with respect to f. However the Green's function is not needed to understand what follows, and so we will simplify notation and introduce the solution operator $\mathcal{L}(m)$ that can be used to write u as indicated below:

$$u = \mathcal{L}(m)f \ . \tag{2}$$

In (2) we have abused notation slightly and written $\mathcal{L}(m)f$ to denote $\mathcal{L}(m)(f)$, where the first bracket denotes the fact that the solution operator depends on m, while the second bracket denotes the argument on which the solution operator $\mathcal{L}(m)$ acts on, namely the source function. We will be using both these notations interchangeably depending upon context. For a fixed value of m, $\mathcal{L}(m)$ is a linear operator on the space of source functions.

The Born linearization operator

A natural question that arises in the study of hyperbolic systems such as equation (1), is how the solution u changes upon perturbations in the acoustic model m for a fixed source function f. Under some assumptions about convergence, this question is conveniently answered using perturbation theory. We perturb the acoustic model by a small amount δm, parameterized by a parameter ϵ, so that the perturbed model is given by $m' = m + \epsilon\delta m$. We now seek a new solution u' of the equation $u'_{tt} - m'\nabla^2 u' = f$ in the power series expansion in ϵ

$$u' = u + \sum_{n=1}^{\infty} \epsilon^n \delta u_n \ , \tag{3}$$

assuming such a series exists and converges uniformly to u'.

Under the stated assumptions of existence and convergence, this series is known as the Born scattering series. Substituting the expressions for u' and m' into the wave equation:

$$(u + \sum_{n=1}^{\infty} \epsilon^n \delta u_n)_{tt} - (m + \epsilon\delta m)\nabla^2(u + \sum_{n=1}^{\infty} \epsilon^n \delta u_n) = f \ . \tag{4}$$

Next, following perturbation theory, we equate similar order terms in powers of ϵ appearing in equation (4) and finally obtain the system of equations

$$\begin{aligned} Zero^{th}\ order: &\quad u_{tt} - m\nabla^2 u = f\ , \\ 1^{st}\ order: &\quad (\delta u_1)_{tt} - m\nabla^2(\delta u_1) = \delta m \nabla^2 u\ ,\ \text{and} \\ n^{th}\ order: &\quad (\delta u_n)_{tt} - m\nabla^2(\delta u_n) = \delta m \nabla^2(\delta u_{n-1})\ ,\ n \geq 2\ . \end{aligned} \tag{5}$$

The $Zero^{th}$ order equation in (5) is the same as equation (1), and represents the acoustic wave equation satisfied by the unperturbed solution u, which is also called the zero order wavefield. The 1^{st} order and all subsequent higher order equations give us a recursive formula to compute the n^{th} order wavefield using the $(n-1)^{th}$ order wavefield. Using the solution operator $\mathcal{L}(m)$ introduced in equation (2), we can symbolically write the solution to the system of equations in (5) as

$$\begin{aligned} Zero^{th}\ order: &\quad u = \mathcal{L}(m) f\ , \\ 1^{st}\ order: &\quad \delta u_1 = \mathcal{L}(m)(\delta m \nabla^2 u)\ ,\ \text{and} \\ n^{th}\ order: &\quad \delta u_n = \mathcal{L}(m)(\delta m \nabla^2 \delta u_{n-1})\ ,\ n \geq 2\ . \end{aligned} \tag{6}$$

It is also clear that the solution operator allows us to write the perturbed solution directly as $u' = \mathcal{L}(m + \epsilon\delta m) f$. Hence, equating the Born series representation of the perturbed wavefield with $\mathcal{L}(m + \epsilon\delta m) f$ and using the relations derived in equation (6) we obtain:

$$\begin{aligned} &\mathcal{L}(m + \epsilon\delta m) f - \mathcal{L}(m) f = \epsilon \mathcal{L}(m)(\delta m \nabla^2 u) + \sum_{n=2}^{\infty} \epsilon^n \mathcal{L}(m)(\delta m \nabla^2 \delta u_{n-1}) \\ &\text{which taking limit } \epsilon \to 0 \text{ becomes, } \mathcal{L}(m + \epsilon\delta m) f - \mathcal{L}(m) f = \epsilon \mathcal{L}(m)(\delta m \nabla^2 u) + \mathcal{O}(\epsilon^2)\ . \end{aligned} \tag{7}$$

Equation (7) says that to first order the perturbation in the wavefield, which is the difference between the perturbed wavefield and zero order wavefield is linear in δm. Moreover, the Frechet derivative of the solution u with respect to m also exists as the following holds:

$$\lim_{\epsilon \to 0} \frac{||\mathcal{L}(m + \epsilon\delta m) f - \mathcal{L}(m) f - \epsilon \mathcal{L}(m)(\delta m \nabla^2 u)||}{||\epsilon \delta m||} = 0 \tag{8}$$

with any suitably defined norm on the respective spaces.

The existence of the Frechet derivative gives us the *Born linearization operator* $L(m)$ that, for fixed m and f linearly maps small perturbations in the model parameter to perturbations in the wavefield. For this operator we assume $||\epsilon\delta m||$ is small, drop quadratic and higher order terms in ϵ in equation (7), and finally absorb ϵ into the definition of δm to obtain:

$$\delta u = u' - u = \mathcal{L}(m + \delta m) f - \mathcal{L}(m) f = \mathcal{L}(m)(\nabla^2 u)(\delta m) = L(m)(\delta m)\ . \tag{9}$$

In the above equation, δu denotes the change in the wavefield, and $L(m)$ is the composition of two operators $\mathcal{L}(m) \circ \nabla^2 u$, where $\mathcal{L}(m)$ is the solution operator introduced earlier and $\nabla^2 u$ denotes another linear operator that produces a secondary source by multiplying the physical quantity $\nabla^2 u$ point-wise in space with the model perturbation δm for each time instance t. Also, the notation $L(m)$ signifies the dependence of the Born linearization operator on m which is the point in the model space around which linearization is being carried out.

The extended modeling operator

We are now in a position to introduce the extended model and the extended modeling operator. In seismic inversion problems, most of the time the starting models are simple models with only the low wavenumber components in them. In this case, the higher order terms in equation (7) are significant and cannot be neglected. This condition is also typically the case when we have cycle skipping between the observed data and the modeled data due to the presence of large time shifts, and in this regime the Born linearization operator is inadequate to model the perturbation in the wavefield δu. However, it is conceivable that in the latter scenario, it may possible to model the perturbation δu by creating a superposition of the secondary source wavefield $\nabla^2 u(\delta m)$ at different time delays or *"lags"*, and then using the solution operator $\mathcal{L}(m)$ to model the data using this superimposed secondary source. This idea is the intuition behind the extended Born modeling operator introduced in the context of TFWI by Biondi and Almomin (2014).

In what follows, we will use a tilde notation to differentiate the extended model and the extended operators from their non-extended counterparts. We start by defining the extended model perturbation which is a model perturbation that depends on a time lag parameter τ, and denote it $\delta\widetilde{m}(\tau)$. For conciseness, we will drop the explicit dependence on τ, and write it as simply $\delta\widetilde{m}$. It is convenient to think of $\delta\widetilde{m}$ as a model perturbation that depends on "time" where time is allowed to be positive or negative. The extended modeling operator $\widetilde{\mathcal{L}}(m)$ is then defined as follows:

$$\begin{aligned} \widetilde{\mathcal{L}}(m)(\delta\widetilde{m}, f) &= \mathcal{L}(m)f + \widetilde{L}(m)(\delta\widetilde{m}) = u + \widetilde{L}(m)(\delta\widetilde{m}) \ , \\ \text{where, } \widetilde{L}(m)(\delta\widetilde{m}) &= \mathcal{L}(m)(\nabla^2(\mathcal{L}(m)f) \star_\tau \delta\widetilde{m}) = \mathcal{L}(m)(\nabla^2 u \star_\tau \delta\widetilde{m}) \ , \\ \text{and, } (\nabla^2 u \star_\tau \delta\widetilde{m})(t) &= \int_\tau \nabla^2 u(\tau + t)\delta\widetilde{m}(\tau) d\tau \ . \end{aligned} \tag{10}$$

The operator $\widetilde{\mathcal{L}}(m)$ has two arguments $\delta\widetilde{m}$ and f, and is linear in the argument f. The output consists of the sum of the modeled wavefield u using the non-extended solution operator $\mathcal{L}(m)$, and the extended Born operator $\widetilde{L}(m)$ which is linear in the argument $\delta\widetilde{m}$. Notice that when $\delta\widetilde{m} = \delta(\tau)\delta m$, then $(\nabla^2 u \star_\tau \delta\widetilde{m})(t) = \nabla^2 u(\delta m)$, and then the extended Born operator reduces to the Born operator as $\widetilde{L}(m)(\delta\widetilde{m}) = L(m)(\delta m)$.

Kolmogorov spectral factorization

We now proceed to provide a very brief introduction to the task of estimating a minimum phase signal from a given spectrum. In effect, this process can also be used to construct the minimum phase equivalent of any given signal, by first computing its amplitude spectrum. We will limit ourselves to continuous real valued signals in time, and refer the reader to excellent references (Kolmogorov, 1939; Robinson and Treitel, 1980; Claerbout, 1985) for an in-depth treatment of this topic. We will also assume that the signal belongs to the Schwarz class of functions, which implies that it is *"nicely behaved"*, and we can Fourier transform it without needing to worry about existence, stability and convergence issues.

Let us first define a few quantities like the Fourier transform and its inverse, and state some important properties that hold in special cases, which are the ones we will need. Let $h(t)$

be a real valued signal in time. Then the Fourier transform $\hat{h}(\omega)$ is defined as:

$$\hat{h}(\omega) = \int_{\mathbb{R}} h(t) e^{-2\pi i \omega t} dt \ . \tag{11}$$

By the Fourier inversion formula, we can recover the function $h(t)$ from $\hat{h}(\omega)$ as follows:

$$h(t) = \int_{\mathbb{R}} \hat{h}(\omega) e^{2\pi i \omega t} d\omega \ . \tag{12}$$

A very special case occurs when $h(t)$ is an even function of t, i.e, $h(t) = h(-t)$. In this case, the Fourier transform $\hat{h}(\omega)$ is also an even function of ω, and so we have $\hat{h}(\omega) = \hat{h}(-\omega)$. The Fourier formulas in this case become:

$$\begin{aligned} \hat{h}(\omega) &= 2 \int_{\mathbb{R}^+} h(t) \cos(2\pi\omega t) dt \ , \text{and} \\ h(t) &= 2 \int_{\mathbb{R}^+} \hat{h}(\omega) \cos(2\pi\omega t) d\omega \ . \end{aligned} \tag{13}$$

The quantity $|\hat{h}(\omega)|$ is called the amplitude spectrum. Let us denote it by $A(\omega)$ and note that it satisfies $A(\omega) = \sqrt{\hat{h}(\omega)\hat{h}^*(\omega)}$, where $\hat{h}^*(\omega)$ is the complex conjugate of $\hat{h}(\omega)$. In the case that $h(t)$ is real, it follows from equation (11) that $A(\omega)$ is an even function of ω. This is a key property that is crucial in deriving the minimum phase signal that has the same amplitude spectrum $A(\omega)$. We also have by definition $A(\omega) \geq 0$. In addition, let us assume that the amplitude spectrum is strictly positive, i.e, $A(\omega) > 0$.

Under these assumptions, Kolmogorov spectral factorization performs the following steps in sequence:

(i) Take log of the amplitude spectrum $A(\omega)$ to get $B(\omega)$ which is even
$B(\omega) = \log(A(\omega))$.

(ii) Inverse Fourier transform $B(\omega)$ to time, and note that $B(t)$ is also even
$B(t) = \int_{\mathbb{R}} B(\omega) e^{2\pi i \omega t} d\omega = 2 \int_{\mathbb{R}^+} B(\omega) \cos(2\pi\omega t) d\omega$.

(iii) Create a causal signal from $B(t)$
$B_c(t) = 2B(t)H(t)$, where $H(t)$ is the Heaviside step function.

(iv) Fourier transform $B_c(t)$ to frequency domain
$B_c(\omega) = \int_{\mathbb{R}} B_c(t) e^{-2\pi i \omega t} dt = 2 \int_{\mathbb{R}} B(t) H(t) e^{-2\pi i \omega t} dt = 2 \int_{\mathbb{R}^+} B(t) e^{-2\pi i \omega t} dt$.

(v) Restore amplitude spectrum by exponentiation, and get Fourier transform of minimum phase signal
$\widetilde{h}(\omega) = e^{B_c(\omega)}$.

(vi) Recover the minimum phase signal by inverse Fourier transform
$\widetilde{h}(t) = \int_{\mathbb{R}} \widetilde{h}(\omega) e^{2\pi i \omega t} d\omega$.

Note that steps (i), (iv) and (v) use equation (13) to ensure that the minimum phase signal has the same amplitude spectrum as $A(\omega)$, as shown below:

$$\begin{aligned} B_c(\omega) &= 2 \int_{\mathbb{R}^+} B(t) e^{-2\pi i \omega t} dt = B(\omega) - 2i \int_{\mathbb{R}^+} B(t) \sin(2\pi\omega t) dt \ , \\ \implies |\widetilde{h}(\omega)| &= |e^{B_c(\omega)}| = e^{B(\omega)} = A(\omega) \ . \end{aligned} \tag{14}$$

It can be easily seen that $\widetilde{h}(\omega) \neq 0$, and hence $1/\widetilde{h}(\omega) = e^{-B_c(\omega)} \neq 0$. The signal obtained this way is front loaded, i.e., has most of the energy concentrated at the beginning of the signal. In certain cases, if the amplitude spectrum has zeros then the above algorithm can still be applied with a small thresholding parameter $\epsilon > 0$ which is added to $A(\omega)$ before performing step (i).

In the discrete case, the above algorithm goes through verbatim with the only exception that step (iii) is modified to assign the $t = 0$ sample a weight of 1, when forming the causal signal $B_c(t)$. Also the discrete Fourier transform and discrete inverse Fourier transform are used in the above calculations. The thresholding step is usually mandatory in the discrete case to achieve numerical stability. We will be using the discrete equivalent in the new inversion problem to be defined later.

Optimal transport

We finish this section by introducing the concept of optimal transport and the quadratic Wasserstein distance in 1D, which play a critical role in the new inverse problem formulation. We also modify the quadratic Wasserstein distance definition for non-positive signals that do not obey the *"mass conservation"* requirement. These concepts are introduced and explained next.

We follow a minimalist approach here and completely avoid the language of measure theory in this brief introduction to optimal transport. However, this approach keeps things simple and is more than sufficient for our case as we only deal with 1D optimal transport problem. The mathematically inclined reader is directed to excellent texts on this topic in the references (Villani, 2003, 2008; Engquist et al., 2016; Yang et al., 2016).

Consider two functions $v(x)$ and $w(x)$ defined on $\mathbb{R}$, and a positive integer N, which satisfy the properties:

$$\begin{aligned}
&(i)\ \ v(x) \geq 0 \text{ and } w(x) \geq 0,\ \forall\, x \in \mathbb{R}\ , && \text{(non-negativity)}\\
&(ii)\ \ \int_{\mathbb{R}} v(x)dx = \int_{\mathbb{R}} w(x)dx \neq 0\ , \text{ and} && \text{(mass conservation)}\\
&(iii)\ \ \int_{\mathbb{R}} |x|^p v(x)dx < \infty, \text{ and } \int_{\mathbb{R}} |x|^p w(x)dx < \infty,\ \forall\, 0 \leq p \leq N\ . && \text{(finite moments).}
\end{aligned} \tag{15}$$

In optimal transport literature, the mass conservation condition is stated to read $\int_{\mathbb{R}} v(x)dx = \int_{\mathbb{R}} w(x)dx = 1$, as it is always possible to do so by rescaling the functions $v(x)$ and $w(x)$, and this is the convention that we will also adopt. In fact, when this is true, these functions can be interpreted as probability measures on $\mathbb{R}$. As we will only be interested in the quadratic Wasserstein distance, it will also suffice for us to require $N = 2$.

The optimal transport problem is concerned with finding a transport plan to rearrange v to w given a cost function $c(x, y)$, such that the transportation cost is minimum. The cost function can be thought of as a penalty to move one unit of mass from position x to position y. For some rearrangement strategy T, called the transport plan, the transportation cost is given by:

$$\text{Cost}(T) = \int_{\mathbb{R}} c(x, T(x))v(x)dx\ . \tag{16}$$

We can now define the *quadratic Wasserstein distance* between the functions v and w (Villani, 2003), using the quadratic cost function $|x-y|^2$ (hence the name quadratic Wasserstein distance):

$$W_2(v,w) = \left(\inf_{T \in \mathcal{M}} \int_{\mathbb{R}} |x - T(x)|^2 v(x) dx \right)^{1/2} , \tag{17}$$

where $\mathcal{M}$ is the set of all possible maps that rearrange v to w. It can be shown that W_2 satisfies all the conditions of a metric, and hence is justifiably referred to as a "distance" function.

In 1D, the minimization problem in equation (17) can be solved exactly (see Villani (2003)). Moreover, this can also be done very quickly on a computer. This is a key feature that motivated us to consider W_2 in our work in this paper. The solution can be given in terms of the cumulative distribution functions of v and w, namely $V(x)$ and $W(x)$, which are defined below:

$$V(x) = \int_{-\infty}^{x} v(x) dx \ , \ \ W(x) = \int_{-\infty}^{x} w(x) dx \ . \tag{18}$$

With these quantities defined, the optimal transportation cost can be evaluated as follows (see Villani (2003) for details):

$$W_2^2(v,w) = \int_0^1 |V^{-1}(\gamma) - W^{-1}(\gamma)|^2 d\gamma = \int_{\mathbb{R}} |x - W^{-1}(V(x))|^2 v(x) dx \ . \tag{19}$$

Any of the formulas in equation (19) can be used for the computation of $W_2(v,w)$. It is also easy to see that the optimal transportation plan is given by the composite map $W^{-1} \circ V$. However, we will be trying to apply the quadratic Wasserstein distance to seismic inversion problems, where neither the positivity condition nor the mass conservation condition is satisfied. To get around this problem, we adopt the normalization technique introduced by Yang et al. (2016). The first step involves choosing a parameter $\delta \geq 0$, such that both the functions $v + \delta \geq 0$ and $w + \delta \geq 0$. An obvious such choice is $\delta = -\min(0, \min_{x \in \mathbb{R}}(v(x), w(x)))$. Next, once we have satisfied the non-negativity condition, we can scale both functions $v + \delta$ and $w + \delta$ to have area 1. We will refer to these normalized functions as $\widetilde{v}$ and $\widetilde{w}$. The W_2 metric can now be applied to these normalized functions. However, these operations in general no longer give rise to a metric with respect to the original functions v and w. For this reason, we will still refer to it as the quadratic Wasserstein distance (with abuse of the term "distance"), but to distinguish it from a true distance function we will denote it by $\widetilde{W}_2(v,w)$. Thus we have:

$$\widetilde{W}_2(v,w) = W_2(\widetilde{v}, \widetilde{w}) \ . \tag{20}$$

THE INVERSE PROBLEM

We now define the inverse problem mathematically and also explain the important differences between standard TFWI and our new formulation based on optimal transport. It is best to convert the problem to the discrete setting, as this operation is a necessary step in solving the problem on a computer. All parameters and operators will be represented in boldface to emphasize the fact they are defined on finite dimensional vector spaces after the discretization. Also, we will represent the field quantities like wavefield, source function

and all model parameters (extended and non-extended) as vectors using the canonical basis in their respective dimensions, and similarly the linear operators will become matrices represented with respect to the canonical basis of their arguments.

The discrete setting

We discretize space and time using $\Delta x, \Delta z,$ and Δt corresponding to the x, y and z dimensions respectively. We also assume that the number of grid points along these dimensions are respectively $N_x, N_z,$ and N_t. For the time lag axis τ, we discretize it using the same interval for time t, and thus $\Delta\tau = \Delta t$. To cover all possible time lags we will choose the number of grid points along τ as $N_\tau = 2N_t - 1$, with the N_t sample corresponding to $\tau = 0$. The model and its perturbation will be represented by $\mathbf{m}$ and $\boldsymbol{\delta}\mathbf{m}$, and belong to a vector space of dimension $N_x N_z$. The extended model perturbation will be denoted by $\boldsymbol{\delta}\widetilde{\mathbf{m}}$ that belongs to a vector space of dimension $N_x N_z N_\tau$. Source functions, wavefields and wavefield perturbations will be represented by $\mathbf{f}$, $\mathbf{u}$ and $\boldsymbol{\delta}\mathbf{u}$ respectively, and belong to a vector space of dimension $N_x N_z N_t$. The i^{th} order wavefield in the Born scattering series will be denoted by $\boldsymbol{\delta}\mathbf{u}_i$ and live in the same vector space as $\mathbf{u}$. All the operators are defined similarly in boldface, and we write their modeling equations as

$$\begin{aligned}
&\mathbf{u} = \boldsymbol{\mathcal{L}}(\mathbf{m})\mathbf{f}\ , \\
&\boldsymbol{\delta}\mathbf{u} = \boldsymbol{\mathcal{L}}(\mathbf{m})(\boldsymbol{\nabla}^2\mathbf{u})(\boldsymbol{\delta}\mathbf{m}) = \mathbf{L}(\mathbf{m})(\boldsymbol{\delta}\mathbf{m})\ , \text{ and} \\
&\widetilde{\boldsymbol{\mathcal{L}}}(\mathbf{m})(\boldsymbol{\delta}\widetilde{\mathbf{m}}, \mathbf{f}) = \boldsymbol{\mathcal{L}}(\mathbf{m})\mathbf{f} + \boldsymbol{\mathcal{L}}(\mathbf{m})(\boldsymbol{\nabla}^2\mathbf{u} \star_\tau \boldsymbol{\delta}\widetilde{\mathbf{m}}) = \boldsymbol{\mathcal{L}}(\mathbf{m})\mathbf{f} + \widetilde{\mathbf{L}}(\mathbf{m})(\boldsymbol{\delta}\widetilde{\mathbf{m}})\ .
\end{aligned} \tag{21}$$

The operators $\boldsymbol{\mathcal{L}}(\mathbf{m})$, $\mathbf{L}(\mathbf{m})$ and $\widetilde{\mathbf{L}}(\mathbf{m})$ are linear operators and thus can be represented as matrices with sizes easily understood from their range and domain. The elements of the matrices depend on the model $\mathbf{m}$. We recognize that they also depend on $\mathbf{f}$ but we will develop the inverse formulation for a fixed shot in the seismic survey and hence assume $\mathbf{f}$ to be fixed. For the sake of brevity, we will drop the explicit dependence of these matrices on $\mathbf{m}$ and simply write $\boldsymbol{\mathcal{L}}$, $\mathbf{L}$ and $\widetilde{\mathbf{L}}$ to refer to them respectively. We will denote their adjoints by $\boldsymbol{\mathcal{L}}^\dagger$, $\mathbf{L}^\dagger$ and $\widetilde{\mathbf{L}}^\dagger$ which we will need in the next section. The adjoints are nothing but the transposes of the matrices for each operator. We will also be using a few other adjoints which we will introduce later.

FWI and TFWI inverse problems

We can now formulate the inverse problem. For a seismic survey the objective function that we will develop is *additive* in each shot, hence all derived quantities like gradients and Hessians are also additive in the shots. It thus suffices for us to develop an inverse problem formulation for a single seismic shot.

Let the source be located at the point $\mathbf{x}_s = (x_s, z_s)$, and assume we have N_r receivers for the seismic shot records each with its own location $\mathbf{x}_r = (x_r, z_r)$. We denote the recorded data at all the receivers by $\mathbf{d}_r$, which is a vector with dimension $N_r N_t$. We also introduce the restriction operator $\mathbf{R}$ (also a matrix as it is a linear operator) which samples quantities such as the wavefield $\mathbf{u}$ at the receiver locations. We define the true model by $\mathbf{m}_{true}$, and thus we have $\mathbf{d}_r = \mathbf{R}\boldsymbol{\mathcal{L}}(\mathbf{m}_{true})\mathbf{f}$.

For reference, the standard FWI objective function $J_{FWI}(\mathbf{m})$ and the expression for its gradient $\frac{\partial J_{FWI}(\mathbf{m})}{\partial \mathbf{m}}$ are:

$$\begin{aligned} J_{FWI}(\mathbf{m}) &= \frac{1}{2}||\mathbf{R}\boldsymbol{\mathcal{L}}\mathbf{f} - \mathbf{d}_r||_2^2 \ , \text{ and} \\ \frac{\partial J_{FWI}(\mathbf{m})}{\partial \mathbf{m}} &= \mathbf{L}^\dagger \mathbf{R}^\dagger(\mathbf{R}\boldsymbol{\mathcal{L}}\mathbf{f} - \mathbf{d}_r) \ , \text{ noting that } \frac{\partial(\boldsymbol{\mathcal{L}}\mathbf{f})}{\partial \mathbf{m}} = \mathbf{L} \ . \end{aligned} \tag{22}$$

FWI attempts to recover $\mathbf{m}_{true}$ by minimizing J_{FWI} starting from some initial model $\mathbf{m}_0$. However, as is well known, if $||\mathbf{m}_{true} - \mathbf{m}_0||$ is too large, then FWI gets trapped in local minima.

Now, we examine the TFWI objective function introduced in Biondi and Almomin (2014). Their objective function is

$$J_{TFWI}(\mathbf{m}, \boldsymbol{\delta}\widetilde{\mathbf{m}}) = \frac{1}{2}||\mathbf{R}\widetilde{\boldsymbol{\mathcal{L}}}(\boldsymbol{\delta}\widetilde{\mathbf{m}}, \mathbf{f}) - \mathbf{d}_r||_2^2 + \epsilon||\,|\tau|\boldsymbol{\delta}\widetilde{\mathbf{m}}||_2^2 \ . \tag{23}$$

The data-fitting term in the TFWI objective function $||\mathbf{R}\widetilde{\boldsymbol{\mathcal{L}}}(\boldsymbol{\delta}\widetilde{\mathbf{m}}, \mathbf{f}) - \mathbf{d}_r||$ uses the extended modeling operator to match the recorded data. This is possible in part due to the introduction of additional variables into the problem in the form of the extended model $\boldsymbol{\delta}\widetilde{\mathbf{m}}$. The regularization term in the objective function $||\,|\tau|\boldsymbol{\delta}\widetilde{\mathbf{m}}||$ penalizes components of the extended model at $\tau \neq 0$. The goal of this term is to eventually drive the extended model to the approximate form $\boldsymbol{\delta}\widetilde{\mathbf{m}} \approx \delta(\tau)\boldsymbol{\delta}\mathbf{m}$. When this happens, i.e, both the data fitting and regularization terms are small simultaneously for some $\mathbf{m}$ and $\boldsymbol{\delta}\mathbf{m}$, it means we are in a regime where Born linearization is accurate and we thus have $J_{FWI}(\mathbf{m} + \boldsymbol{\delta}\mathbf{m}) \rightarrow 0$. The iterative inversion is carried out in both variables $\mathbf{m}$ and $\boldsymbol{\delta}\widetilde{\mathbf{m}}$ with scale mixing.

A new inverse problem formulation

Suppose we have a model $\mathbf{m}_0$ and we want to calculate $\boldsymbol{\delta}\widetilde{\mathbf{m}}$ that minimizes the data fitting term of TFWI $\frac{1}{2}||\mathbf{R}\widetilde{\boldsymbol{\mathcal{L}}}(\boldsymbol{\delta}\widetilde{\mathbf{m}}, \mathbf{f}) - \mathbf{d}_r||_2^2$. Let us start from $\boldsymbol{\delta}\widetilde{\mathbf{m}} = \mathbf{0}$, and calculate the gradient of the data fitting term with respect to $\boldsymbol{\delta}\widetilde{\mathbf{m}}$ by

$$\begin{aligned} &\frac{\partial}{\partial(\boldsymbol{\delta}\widetilde{\mathbf{m}})}(\frac{1}{2}||\mathbf{R}\widetilde{\boldsymbol{\mathcal{L}}}(\boldsymbol{\delta}\widetilde{\mathbf{m}}, \mathbf{f}) - \mathbf{d}_r||_2^2) = \widetilde{\mathbf{L}}^\dagger \mathbf{R}^\dagger(\mathbf{R}\widetilde{\boldsymbol{\mathcal{L}}}(\boldsymbol{\delta}\widetilde{\mathbf{m}}, \mathbf{f}) - \mathbf{d}_r) \ , \text{ as } \frac{\partial \widetilde{\boldsymbol{\mathcal{L}}}(\boldsymbol{\delta}\widetilde{\mathbf{m}}, \mathbf{f})}{\partial(\boldsymbol{\delta}\widetilde{\mathbf{m}})} = \widetilde{\mathbf{L}} \ , \\ \implies &\frac{\partial}{\partial(\boldsymbol{\delta}\widetilde{\mathbf{m}})}(\frac{1}{2}||\mathbf{R}\widetilde{\boldsymbol{\mathcal{L}}}(\boldsymbol{\delta}\widetilde{\mathbf{m}}, \mathbf{f}) - \mathbf{d}_r||_2^2)\Big|_{\boldsymbol{\delta}\widetilde{\mathbf{m}}=\mathbf{0}} = \widetilde{\mathbf{L}}^\dagger \mathbf{R}^\dagger(\mathbf{R}\boldsymbol{\mathcal{L}}\mathbf{f} - \mathbf{d}_r) \ . \end{aligned} \tag{24}$$

If we take a step along the negative gradient and use it as an estimate of $\boldsymbol{\delta}\widetilde{\mathbf{m}}$ we obtain

$$\boldsymbol{\delta}\widetilde{\mathbf{m}}(\mathbf{m}_0) = -\widetilde{\mathbf{L}}^\dagger(\mathbf{m}_0)\mathbf{R}^\dagger(\mathbf{R}\boldsymbol{\mathcal{L}}(\mathbf{m}_0)\mathbf{f} - \mathbf{d}_r) \ . \tag{25}$$

This $\boldsymbol{\delta}\widetilde{\mathbf{m}}$ obtained this way will in general have components at $\tau \neq 0$, and has an implicit dependence on $\mathbf{m}_0$. At this stage, the TFWI philosophy motivates us to find a way to change $\mathbf{m}_0$ such that $\boldsymbol{\delta}\widetilde{\mathbf{m}}$ becomes focused at $\tau = 0$. To that extent, we now define a focusing operator $\boldsymbol{\mathcal{T}}(\boldsymbol{\delta}\widetilde{\mathbf{m}})$ that achieves this using Kolmogorov spectral factorization. This method proceeds by isolating the strictly positive and strictly negative time lag parts of

$\boldsymbol{\delta\widetilde{\mathbf{m}}}$ at each grid point of space, and then applying minimum phase transformation to each part separately. This process is repeated at each spatial grid point to yield the final output $\boldsymbol{\mathcal{T}}(\boldsymbol{\delta\widetilde{\mathbf{m}}})$. This transformation is the first key step in how the new inversion framework differs from classical TFWI.

The next step in the proposed inversion framework involves devising an update strategy for $\mathbf{m}_0$. To achieve this, we minimize the objective function

$$J(\mathbf{m}) = \sum_{i=0}^{N_x-1}\sum_{j=0}^{N_z-1}\frac{1}{2}\widetilde{W}_2^2(\boldsymbol{\delta\widetilde{\mathbf{m}}}_{ij}(\mathbf{m}), \boldsymbol{\mathcal{T}}(\boldsymbol{\delta\widetilde{\mathbf{m}}}_{ij}(\mathbf{m}_0))) \tag{26}$$

based on the quadratic Wasserstein distance that was introduced in equation (20).

The quantity $\boldsymbol{\delta\widetilde{\mathbf{m}}}_{ij}$ appearing in equation (26) represents the part of the extended perturbation model corresponding to the spatial grid location indexed by (i, j), having physical coordinates $(i\Delta x, j\Delta z)$ with respect to the origin. An important idea that the notation in equation (26) tries to emphasize is that the second argument in the Wasserstein distance $\boldsymbol{\mathcal{T}}(\boldsymbol{\delta\widetilde{\mathbf{m}}}_{ij}(\mathbf{m}_0))$ only depends on the starting model $\mathbf{m}_0$, and not on $\mathbf{m}$. This step is expected to be important in achieving a faster convergence rate than conventional TFWI, and the intuition is based on prior work done in 1D using other objective functions (Biondi et al., 2016). Of course its efficacy can only be fully judged in the future once a full suite of numerical tests have been carried out. We also note that the objective function $J(\mathbf{m})$ is a sum over all the grid points in the physical space, i.e, the square of the quadratic Wasserstein distance is first evaluated at every spatial point, and then added up to yield the complete $J(\mathbf{m})$.

The minimization proceeds by starting from $\mathbf{m} = \mathbf{m}_0$, followed by using a suitable optimization algorithm. For our initial work, we propose to look at gradient descent as a viable descent type optimization algorithm. Using the chain rule, the gradient of $J(\mathbf{m})$ can be calculated as:

$$\frac{\partial J(\mathbf{m})}{\partial \mathbf{m}} = \sum_{i=0}^{N_x-1}\sum_{j=0}^{N_z-1}\widetilde{W}_2\frac{\partial \widetilde{W}_2}{\partial(\boldsymbol{\delta\widetilde{\mathbf{m}}}_{ij})}\frac{\partial(\boldsymbol{\delta\widetilde{\mathbf{m}}}_{ij})}{\partial \mathbf{m}}\ , \tag{27}$$

where we have dropped the explicit dependence of the quantities to make the notation concise. The quantity $\widetilde{W}_2\frac{\partial \widetilde{W}_2}{\partial(\boldsymbol{\delta\widetilde{\mathbf{m}}}_{ij})}$ can be calculated easily using finite differences. The other quantity $\frac{\partial(\boldsymbol{\delta\widetilde{\mathbf{m}}}_{ij})}{\partial \mathbf{m}}$ are columns of the matrix $\frac{\partial(\boldsymbol{\delta\widetilde{\mathbf{m}}})}{\partial \mathbf{m}}$. Using equation (25) this operator is given by

$$\begin{aligned}
\frac{\partial(\boldsymbol{\delta\widetilde{\mathbf{m}}})}{\partial \mathbf{m}} &= -\frac{\partial}{\partial \mathbf{m}}\left(\widetilde{\mathbf{L}}^\dagger\mathbf{R}^\dagger(\mathbf{R}\boldsymbol{\mathcal{L}}\mathbf{f} - \mathbf{d}_r)\right) \\
&= -\left(\frac{\partial\widetilde{\mathbf{L}}^\dagger}{\partial \mathbf{m}}\right)\mathbf{R}^\dagger(\mathbf{R}\boldsymbol{\mathcal{L}}\mathbf{f} - \mathbf{d}_r) - \widetilde{\mathbf{L}}^\dagger\mathbf{R}^\dagger\frac{\partial(\mathbf{R}\boldsymbol{\mathcal{L}}\mathbf{f})}{\partial \mathbf{m}} \\
&= -\left(\frac{\partial\widetilde{\mathbf{L}}^\dagger}{\partial \mathbf{m}}\right)\mathbf{R}^\dagger(\mathbf{R}\boldsymbol{\mathcal{L}}\mathbf{f} - \mathbf{d}_r) - \widetilde{\mathbf{L}}^\dagger\mathbf{R}^\dagger\mathbf{R}\mathbf{L}\ .
\end{aligned} \tag{28}$$

We now have all the machinery in place to solve the proposed inversion problem. We summarize the steps of the algorithm below:

- Assuming that at the k^{th} iteration the model is denoted by $\mathbf{m}^k$, we calculate the quantity $\boldsymbol{\delta}\widetilde{\mathbf{m}}^k$ as follows:

$$\boldsymbol{\delta}\widetilde{\mathbf{m}}^k(\mathbf{m}^k) = -\widetilde{\mathbf{L}}^\dagger \mathbf{R}^\dagger (\mathbf{R}\boldsymbol{\mathcal{L}}\mathbf{f} - \mathbf{d}_r) \; .$$

- Using Kolmogorov factorization, we then calculate the quantity $\boldsymbol{\mathcal{T}}(\boldsymbol{\delta}\widetilde{\mathbf{m}}^k_{ij}(\mathbf{m}^k))$, for all $0 \leq i \leq N_x - 1$ and $0 \leq j \leq N_z - 1$.

- We then perform gradient descent iterations to get a new model $\mathbf{m}^{k+1}$ starting from $\mathbf{m}^k$, on the objective function

$$J(\mathbf{m}) = \sum_{i=0}^{N_x-1} \sum_{j=0}^{N_z-1} \frac{1}{2} \widetilde{W}_2^2(\boldsymbol{\delta}\widetilde{\mathbf{m}}_{ij}(\mathbf{m}), \boldsymbol{\mathcal{T}}(\boldsymbol{\delta}\widetilde{\mathbf{m}}_{ij}(\mathbf{m}^k))) \; .$$

- Setting $k = k + 1$, we iterate till convergence / stopping criterion.

PRELIMINARY NUMERICAL RESULTS

In this section we present some preliminary numerical results with a software library under development, that can be used to solve a class of FWI problems in 2D, including the new proposed optimization algorithm. A key feature of this library is that it attempts to integrate numerical algorithms from the world of optimal transport into the TFWI scheme with modified objective functions, like the one proposed in this paper. We intend the library to also be capable of solving the standard FWI / TFWI problems, allowing us to numerically measure any gains in efficiency the optimal transport based algorithms provide.

We still have a lot of work to do with the library, and it is in its preliminary stages. As of now, we can only solve the standard FWI problem reliably. We are currently working on implementing efficient versions of standard TFWI solver along with the optimal transport variant. In the rest of this section, we present some results where we try to invert for a low-velocity Gaussian anomaly in a two layer background model. These results illustrate the failure of the FWI objective function.

Modeling synthetic data

We start by creating a 5 km x 3 km model along X and Z directions respectively with grid spacing 10 m x 10 m along each direction. This represents our simulation world. In Figure 1(a), we display the original model. The background is a two layer model with velocities 3 km/s and 3.5 km/s with the interface between the two layers located at $Z = 1.5$ km. There is a low velocity Gaussian anomaly present in the original model, centered at $X = 2.5$ km, $Z = 0.75$ km, with a peak velocity anomaly value of -0.5 km/s. We will refer to the original model as $\mathbf{m}_{true}$.

To model the simulated shots, we place receivers at every grid location on the surface of the model, i.e, at $Z = 0$ km. The source locations are placed at every 100 m intervals along X, also at $Z = 0$ km. An example shot is plotted in Figure 2 and corresponds to a shot located

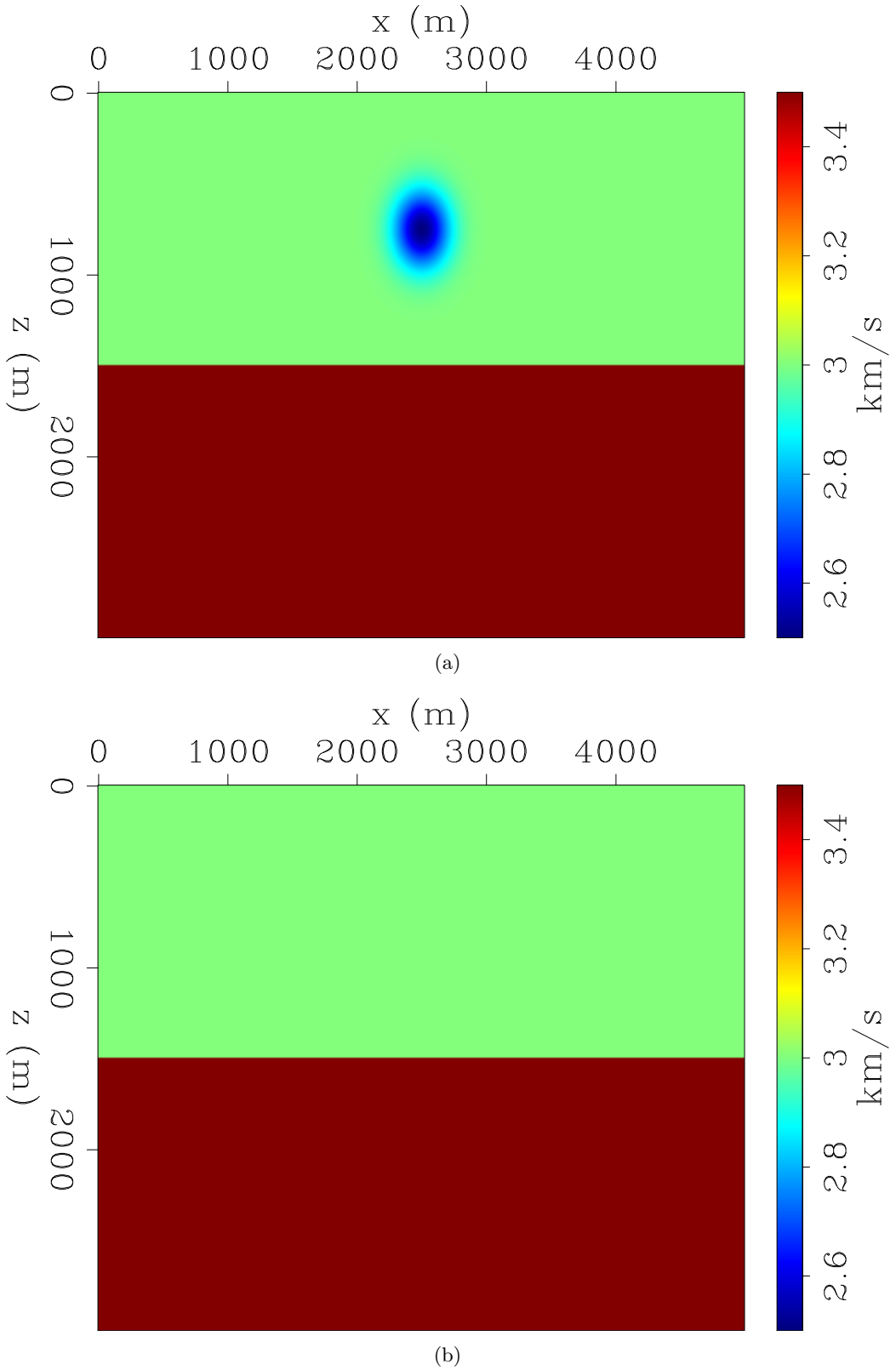


Figure 1: a) The original 2D model $\mathbf{m}_{true}$ used to generate synthetic shots for the study. b) The starting model $\mathbf{m}_0$ used for starting FWI. [**ER**] rahul1/. vel-orig,vel-starting-model

at $X = 2.5$ km, $Z = 0$ km. The data are modeled using a Ricker wavelet with dominant frequency 12.5 Hz, which is plotted in Figure 3. As can be clearly seen in the shot record, the low velocity Gaussian anomaly introduces triplication of raypaths that lead to the effect seen at around $X = 2.5$ km, $t = 1.2$ s.

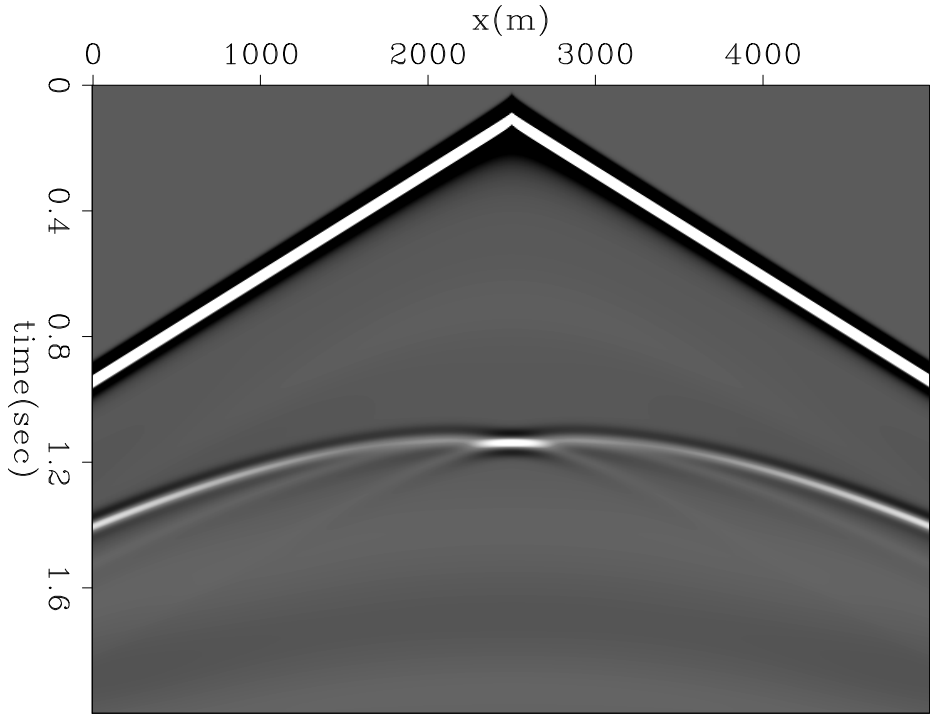


Figure 2: A shot record extracted from the center of the survey. The location of source is $X = 2.5$ km, $Z = 0$ km. [**ER**] rahul1/. one-shot-orig

Failure of FWI

Figure 1(b) represents the starting velocity model used in the FWI inversion, which is identical to the original velocity model (Figure 1(a)) without the Gaussian anomaly. We will refer to it as $\mathbf{m}_0$. In Figure 4(a), we plot the modeled data with $\mathbf{m}_0$ for the same central shot location given by $X = 2.5$ km, $Z = 0$ km. The residual for this shot is also shown in Figure 4(b). On the residual, we can clearly see cycle-skipping on the mid to far offsets. The FWI gradient with $\mathbf{m}_0$ is displayed in Figure 5.

We ran 10 iterations of FWI, starting from $\mathbf{m}_0$. The result of the update in the velocity after 10 iterations is shown in Figure 6(b). This can be contrasted with the true update required to recover the original model as shown in Figure 6(a). The maximum update is only about 20 m/s after 10 iterations. Running more FWI iterations does not improve the model significantly. Moreover we also observed that its convergence is really slow as we obtained very small step sizes during the line search process used in the inversion. Lastly, it is to be noted that this example was run to illustrate the limitation of FWI in the presence of cycle-skipping and thus we specifically did not use a multiscale inversion scheme. The inversion was carried out in the same frequency band as the original synthetic data.

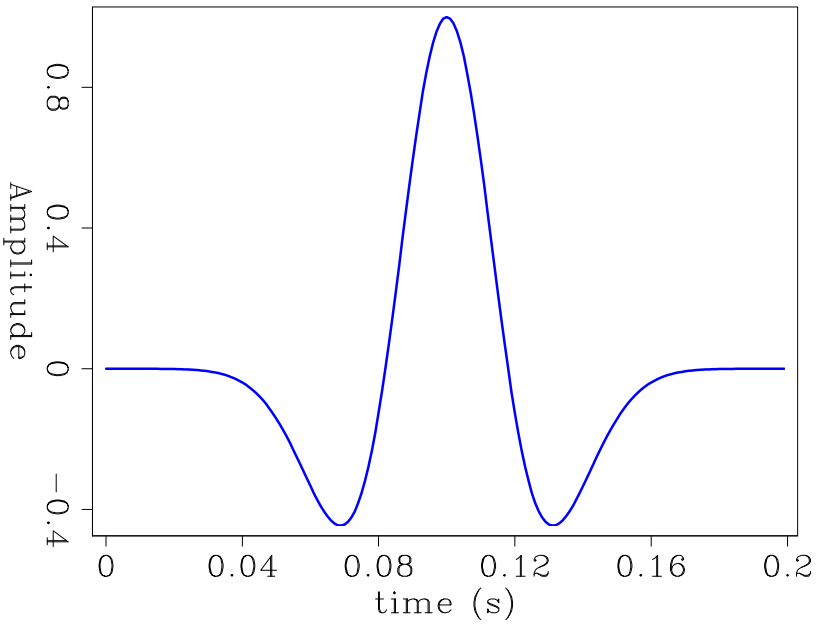


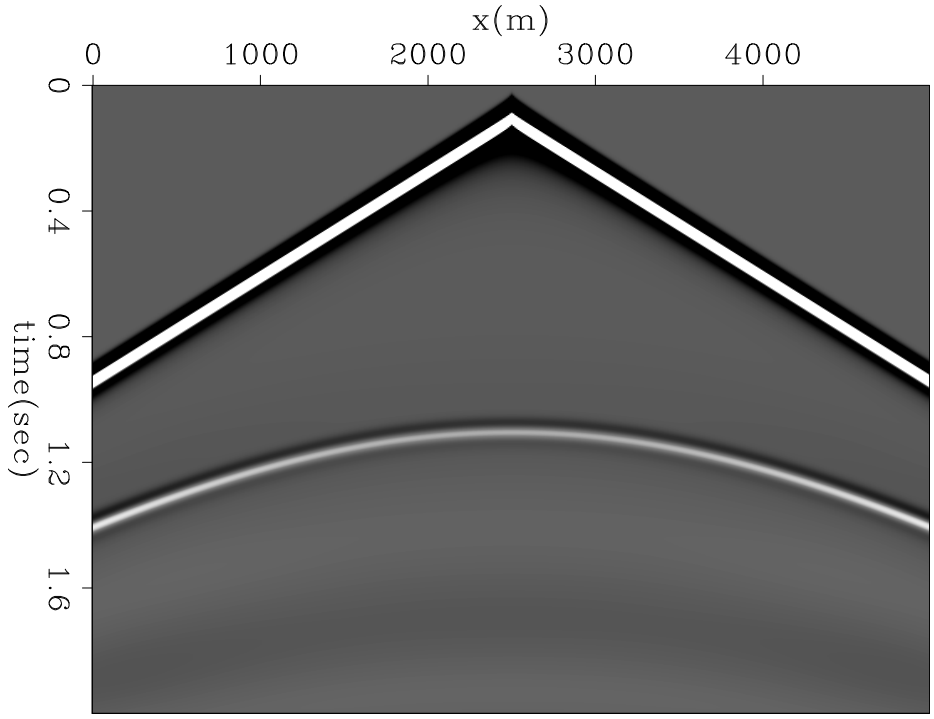
Figure 3: Ricker wavelet used in the modeling of synthetic data (12.5 Hz dominant frequency). [**ER**] rahul1/. ricker-peak12-5Hz

CONCLUSIONS

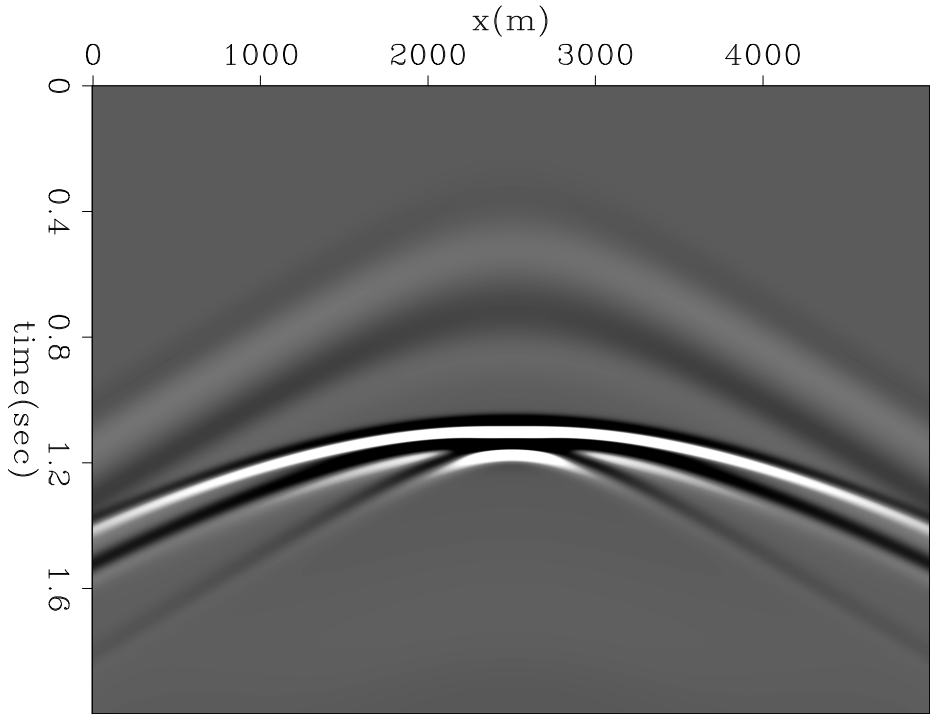
In this paper, we have introduced the concept of optimal transport as a tool to modify the standard TFWI inversion algorithm, and proposed a new inversion framework that can help address the issue of slow convergence in TFWI. We have also started to develop a 2D library that can be used to solve the new optimization problem. The library is intended to be capable of also solving the standard FWI and the standard TFWI problems, for efficient comparison. Some preliminary results with this library have been presented in this paper. In particular, we have demonstrated the need for devising better alternatives to the standard FWI algorithm, by showing its failure to converge, even in the case of a simple low-velocity Gaussian anomaly in a two layer model.

ACKNOWLEDGMENTS

Rahul Sarkar would like to thank Bob Clapp for his advise on several computational aspects of this work, that allowed him to write software to generate results in a reasonable amount of time. He would also like to thank Jason Chang, Ettore Biondi, Guillaume Barnier and Yinbin Ma for several discussions and their help with SEPlib. We thank Ettore Biondi and Stewart Levin for reading the manuscript and suggesting edits that greatly helped to improve the quality of the document. Finally, we thank all the sponsors of SEP for their generous support and funding.



(a)



(b)

Figure 4: a) Data modeled for the same central shot as in Figure 2 with the starting velocity model $\mathbf{m}_0$. b) The residual for the same shot with gain applied where we can see cycle-skipping on mid to far offsets. [**ER**] rahul1/. one-shot-starting-model,one-shot-starting-model-residual

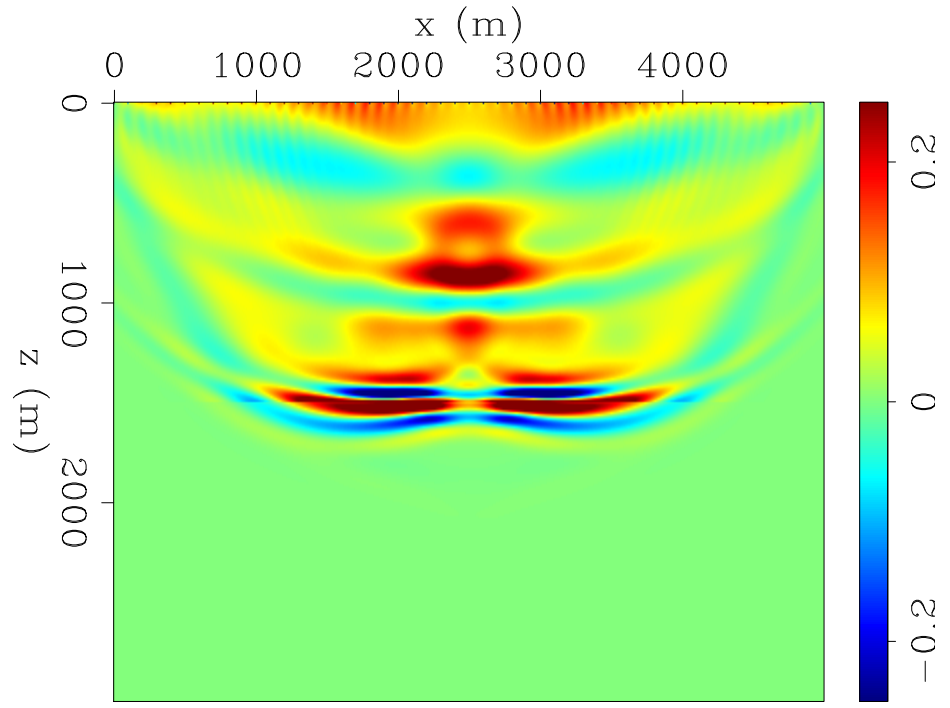


Figure 5: FWI gradient calculated using the starting model $\mathbf{m}_0$. [**CR**] rahul1/. fwi-vel-grad0

REFERENCES

Almomin, A. and B. Biondi, 2014, Preconditioned tomographic full waveform inversion by wavelength continuation: SEG Technical Program Expanded Abstracts, **33**, 944–948.

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

Biondi, B., R. Sarkar, and J. Jennings, 2016, Solving nonlinear inverse problems by linearized model extension — A survey of possible methods: SEP-Report, **165**, 93–122.

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

Claerbout, J. F., 1985, Fundamentals of geophysical data processing.

Engquist, B. and B. D. Froese, 2014, Application of the Wasserstein metric to seismic signals: Communications in Mathematical Sciences, **12**, 979–988.

Engquist, B., B. D. Froese, and Y. Yang, 2016, Optimal transport for seismic full waveform inversion: Communications in Mathematical Sciences, **14**, 2309–2330.

Kolmogorov, A. N., 1939, Sur linterpolation et extrapolation des suites stationnaires: C.R. Acad. Sci, **208**, 2043–2045.

Luo, S. and P. Sava, 2011, A deconvolution-based objective function for wave-equation inversion: SEG Technical Program Expanded Abstracts, 2788–2792.

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

Ma, Y. and D. Hale, 2013, Wave-equation reflection traveltime inversion with dynamic warping and full-waveform inversion: Geophysics, **78**, R223–R233.

Métivier, L., R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, 2016, Measuring the misfit

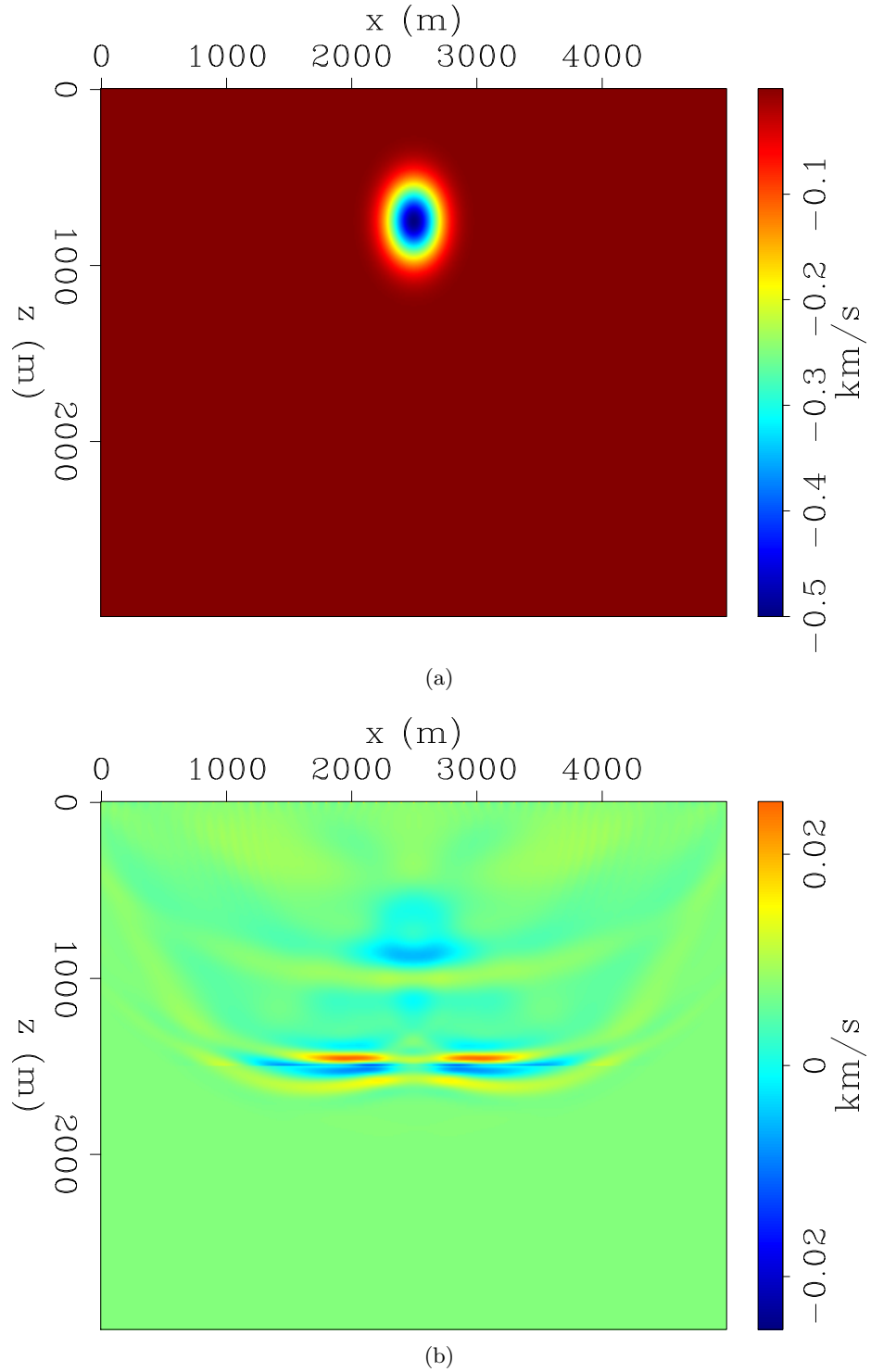


Figure 6: a) Difference between true model $\mathbf{m}_{true}$ and initial starting model $\mathbf{m}_0$. b) Difference between inverted model after 10 iterations and initial starting model $\mathbf{m}_0$. [**CR**] rahul1/. dvel-orig,dvel-iter10

between seismograms using an optimal transport distance: Application to full waveform inversion: Geophysical Journal International, **205**, 345–377.

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.

Robinson, E. and S. Treitel, 1980, Geophysical signal analysis: Prentice-hall: Englewood Cliffs.

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

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

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

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

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

Villani, C., 2003, Topics in optimal transportation: American Mathematical Soc.

———, 2008, Optimal transport: Old and New, volume **338**: Springer Science & Business Media.

Yang, Y., B. Engquist, J. Sun, and B. D. Froese, 2016, Application of optimal transport and the quadratic Wasserstein metric to full-waveform inversion: arXiv preprint arXiv:1612.05075.

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

Source estimation and rock physics preconditioning for anisotropic full waveform inversion

Huy Le, Biondo Biondi, Robert G. Clapp, Stewart A. Levin, and Tapan Mukerji

ABSTRACT

In the imaging and inversion of long-offset seismic data, it is neccessary to incorporate anisotropy to produce accurate images and subsurface models. Building an anisotropic model is a challenging inverse problem because of the large number of unknown parameters. To reduce the number of unknown parameters, it is common to assume a simple physics, for example, using the acoustic approximation or constant density. This leads to a mismatch in amplitude between the modeled data and the observed data. Estimation of the source wavelet by directly minimizing the inversion's objective function can help to mitigate this mismatch. Adding to the challenge of anisotropic model building is the fact that anisotropic parameters are often of different types, such as velocity, stiffness, or Thomsen's parameters. Consequently, one needs to account for their crosstalk and sensitivity. We follow an established workflow that stochastically simulates these parameters based on rock physics and compaction models. This simulation provides information about the model covariance, which can be used to reduce parameter crosstalk.

INTRODUCTION

The acquisition of long-offset three dimensional seismic data often requires imaging and inversion algorithms to go beyond isotropy. The simplest anisotropic model is transverse isotropy, which has one axis of rotational symmetry, i.e. seismic wave velocity depends on the angle between propagation direction and symmetry axis. This type of anisotropy is commonly found in layered media and shale reservoirs.

To fully describe a transverse isotropic medium, we need six parameters, including five stiffnesses and one density. In the context of waveform inversion and model building, it is hardly possible to estimate that many number of parameters. As a result, the industry commonly uses approximations, for example, constant density and acoustic media. For anisotropy, the resulting wave equations from such approximations were derived by Alkhalifah (2000) and are usually refered to as pseudo-acoustic.

The pseudo-acoustic anisotropic wave equations have been widely used and demonstrated to accurately capture the kinematics of seismic waves (Fletcher et al., 2009; Zhang et al., 2011; Bube et al., 2012; Warner et al., 2013). The assumptions of constant density and acoustic media can lead to an amplitude mismatch between the simulated data and the observed data, which can be harmful in waveform inversion. One approach to reduce this mismatch is to directly estimate the source wavelet by minimizing the objective function. In this paper, we derive a method to estimate the source wavelet using the pseudo-acoustic wave equations in anisotropic media.

The second half of this paper deals with the problem of crosstalk and sensitivity among different anisotropic parameters. We adopt an workflow introduced by Bachrach (2010) that stochastically models anisotropic parameters using rock physics and compaction trends of sand and shale. In this workflow, sand's mineral matrix is modeled with quartz while shale's matrix is modeled as a composition of smectite and illite. Both sand's pore and shale's pore are brine-filled. Temperature is used as an indication of volumes fractions of smectite and illite, i.e. the higher the temperature, the greater the illite fraction (Freed and Peacor, 1989). Anisotropy of shale is determined by compaction, as clay minerals align, via porosity. A similar workflow has been incorporated successfully to wave equation migration velocity analysis (Li et al., 2016). We here extend it to full waveform inversion (FWI).

SOURCE INVERSION

Derivation

Using the second-order pseudo acoustic anisotropic wave equations, Le (2016) derived adjoint equations and recursive relations to compute the derivatives of the FWI objective function with respect to model parameters. To estimate the source, $s(t)$, we minimize the same objective function as when inverting for velocity models, $\mathbf{m}$, i.e. the l_2-norm of the different between the modeled, d and the observed data, d_0:

$$f(\mathbf{m}, s) = \frac{1}{2}\|d(\mathbf{x}_r, t; \mathbf{m}, s) - d_0(\mathbf{x}_r, t)\|_2^2. \tag{1}$$

Since the wave equations are linear differential equations, so is the modeled data with respect to the source function:

$$d = \mathbf{RPQ}s, \tag{2}$$

where:

- $\mathbf{Q}$ is the injection of the source function at shot locations,
- $\mathbf{P}$ is the forward propagation operator,
- and $\mathbf{R}$ is the recording of wavefields at receivers.

The linear relationship between the modeled data and the source function makes our objective function quadratic, whose minima can be reached using methods like conjugate gradient if we iterate long enough, i.e. equal to number of source function samples.

The gradient of the objective function 1 can be easily computed:

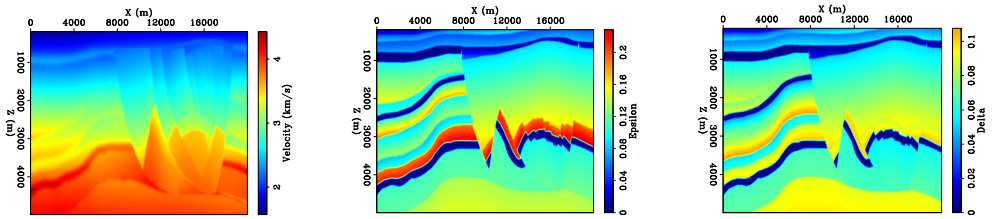
$$\frac{\partial f}{\partial s} = \mathbf{Q}^T\mathbf{P}^T\mathbf{R}^T(d - d_0), \tag{3}$$

where:

- $\mathbf{R}^T$ is the injection of receiver locations,
- $\mathbf{P}^T$ is the backward propagation operator,
- and $\mathbf{Q}^T$ is the recording at source locations.

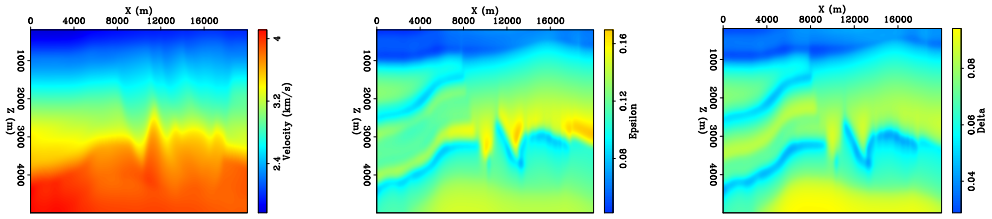
Synthetic example

We perform a synthetic test of source inversion on the BP 2007 models. Figure 1 shows the true models that are used to create the synthetic observed data and Figure 2 shows the smooth background models used in the inversion. Sources and receivers are placed on the surface. We used a band-limited wavelet from a real airgun signature as an input. Figure 3 shows the inversion results. The inverted wavelet matches the input wavelet really well, both in terms of waveform (Figure 3(a)) and frequency amplitude (Figure 3(b)). The inverted wavelet, however, shows some high-frequency ringing noise that can be observed in its waveform and spectrum at about 50 and 70 Hz. This is due to the band-limited nature of the input wavelet that high-frequency noise leaks in from the null space. The objective function reduced 95% after only 14 iterations (Figure 3(c)).



(a) (b) (c)

Figure 1: True models used to generate synthetic observed data: (a) vertical velocity in km/s, (b) ϵ, and (c) δ. [**ER**] huyle/. v.bp2007,eps.bp2007,del.bp2007

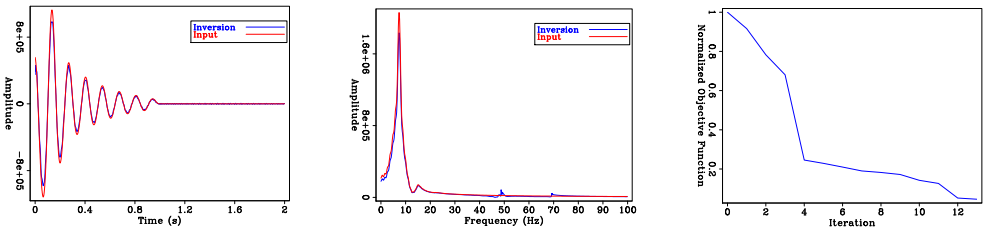


(a) (b) (c)

Figure 2: Smooth background models used in the inversion: (a) vertical velocity in km/s, (b) ϵ, and (c) δ. [**ER**] huyle/. bgv.bp2007,bgeps.bp2007,bgdel.bp2007

ROCK PHYSICS TO PRECONDITION FWI

To deal with the problem of parameter crosstalk in anisotropic FWI, one approach is to incorporate additional information, such as well logs or seismic attributes, to constrain the inversion. Particularly, combining with rock physics and compaction models, one can sample the model space of anisotropic parameters. This sampling provides us with covariance information of model parameters, which can be used as a preconditioner to separate the crosstalk. Bachrach (2010) presented a workflow that stochastically simulates anisotropic models using rock physics principles of shale compaction and anisotropy described in Bandyopadhyay (2009). We extend this workflow to model a sand-shale layered medium.



(a) (b) (c)

Figure 3: Inversion results: (a) input (red) and inverted (blue) wavelets, (b) spectra of input (red) and inverted (blue) wavelet, and (c) objective function reduction with iterations. [**CR**] huyle/. souInvWavelets,souInvSpectra,souInvObjfunc

Inputs

Our rock physics workflow requires subsurface temperature and shale volume as direct inputs. Temperature controls the transition from smectite to illite, which are the two clay minerals that compose shale (Bachrach, 2010):

$$P_{\text{illite}}(T) = \frac{1}{2} + \frac{1}{2}\tanh\left(\frac{T - T_T}{2s_T}\right). \tag{4}$$

Here P_{illite} is the percentage of illite, T_T is the transition temperature, where 50% of the smectite-illite reaction has occured, and s_T is the spread of the transition (Freed and Peacor, 1989).

While temperature can be obtained from basin modeling using borehole data or by assuming a constant depth gradient, shale volume is an attribute that can be obtained from seismic inversion and rock physics transforms. Besides these two direct inputs, it is critical to have well logs, such as gamma ray, sonic logs, and density, for the calibration of rock physics model parameters.

Compaction model

Compaction influences physical properties of rocks, such as porosity and anisotropy. Here we use a normal compaction trend that is commonly adopted for porosity of water-saturated sand and shale in the Gulf of Mexico (Dutta et al., 2009):

$$\phi = ae^{bz} + ce^{dz}. \tag{5}$$

In sand, porosity directly affects density and stiffness of rocks. In shale, compaction determines the alignment of clay minerals, which controls the degree of anisotropy. The more a shale rock is compacted, the more aligned its clay minerals are, and therefore the more anisotropic it is. This relation makes stiffnesses of clay in the subsurface different from those of a single clay mineral. Sayers (1994) derived equations for anisotropy of shale

as functions of its orientation distribution:

$$
\begin{aligned}
c_{11} &= \lambda + 2\mu + \frac{4\sqrt{2}}{105}\pi^2(2\sqrt{5}a_3W_{200} + 3a_1W_{400}),\\
c_{33} &= \lambda + 2\mu - \frac{16\sqrt{2}}{105}\pi^2(\sqrt{5}a_3W_{200} - 2a_1W_{400}),\\
c_{44} &= \mu - \frac{2\sqrt{2}}{315}\pi^2\left[\sqrt{5}(7a_2 + a_3)W_{200} + 24a_1W_{400}\right],\\
c_{12} &= \lambda - \frac{4\sqrt{2}}{315}\pi^2\left[2\sqrt{5}(7a_2 - a_3)W_{200} - 3a_1W_{400}\right],\\
c_{13} &= \lambda + \frac{4\sqrt{2}}{315}\pi^2\left[\sqrt{5}(7a_2 - a_3)W_{200} - 12a_1W_{400}\right],
\end{aligned}
\tag{6}
$$

where:

$$
\begin{aligned}
a_1 &= c^a_{11} + c^a_{33} - 2c^a_{13} - 4c^a_{44},\\
a_2 &= c^a_{11} - 3c^a_{12} + 2c^a_{13} - 2c^a_{44},\\
a_3 &= 4c^a_{11} - 3c^a_{33} - c^a_{13} - 2c^a_{44},\\
\lambda &= \frac{1}{15}(c^a_{11} + c^a_{33} + 5c^a_{12} + 8c^a_{13} - 4c^a_{44}),\\
\mu &= \frac{1}{30}(7c^a_{11} + 2c^a_{33} - 5c^a_{12} - 4c^a_{13} + 12c^a_{44}).
\end{aligned}
\tag{7}
$$

In the above equations, c^a_{ij} are the stiffnesses of a domain, i.e. a single clay mineral, either smectite or illite, and c_{ij} are the averaged stiffnesses of aligned clay particles. W_{200} and W_{400} are the orientation distribution parameters and related to compaction via porosity (Bachrach, 2011):

$$
\begin{aligned}
W_{200}(\phi) &= W^{\max}_{200}\left(1 - \frac{\phi}{\phi_0}\right)^m,\\
W_{400}(\phi) &= W^{\max}_{400}\left(1 - \frac{\phi}{\phi_0}\right)^n,
\end{aligned}
\tag{8}
$$

where $W^{\max}_{200} = \frac{\sqrt{2.5}}{4\pi^2}$ and $W^{\max}_{400} = \frac{\sqrt{4.5}}{4\pi^2}$ are the maxima of W_{200} and W_{400} respectively, when clay particles are completely aligned, m and n are empirical parameters, and ϕ_0 is the critical porosity, i.e. the porosity above which mineral grains are no longer in contact but suspended in the fluid.

Rock physics model

We model a rock as composed of layers of sand and shale. When porosity is less than critical porosity, to model sand, we use the Differential Effective Medium (DEM) equations with water-filled inclusion in quartz background (Hornby et al., 1994). Modeling shale follows a similar procedure as does sand, except that the background mineral is now clay instead of quartz. Elastic properites of clay are computed as Backus's average of smectite and illite

(Backus, 1962):

$$
\begin{aligned}
C_{11}^{\text{clay}} &= \langle\frac{c_{13}}{c_{33}}\rangle^2\langle\frac{1}{c_{33}}\rangle^{-1} - \langle\frac{c_{13}^2}{c_{33}}\rangle + \langle c_{11}\rangle,\\
C_{12}^{\text{clay}} &= C_{11}^{\text{clay}} - \langle c_{11}\rangle + \langle c_{12}\rangle,\\
C_{13}^{\text{clay}} &= \langle\frac{c_{13}}{c_{33}}\rangle^2\langle\frac{1}{c_{33}}\rangle^{-1},\\
C_{33}^{\text{clay}} &= \langle\frac{1}{c_{33}}\rangle^{-1},\\
C_{44}^{\text{clay}} &= \langle\frac{1}{c_{44}}\rangle^{-1},
\end{aligned}
\tag{9}
$$

where c_{ij} are stiffnesses of smectite and illite computed from Equations 6 and $\langle . \rangle$ stands for the average of the enclosed quantity weighted by their volume fractions. After modeling sand and shale separately, we use Backus's average again to obtain the elastic stiffnesses of our sand-shale rock layers.

When porosity is greater than critical porosity, we use Reuss average of water and the mineral matrix:

$$\frac{1}{C_{ij}} = \frac{1-\phi}{C_{ij}^{\text{matrix}}} + \frac{\phi}{C_{ij}^{\text{water}}}. \tag{10}$$

One important parameter, and input to DEM, is the pore aspect ratio. In our workflow, we choose to model pores with elliptical shapes and define aspect ratio as the ratio between the vertical axis and the horizontal axis. An aspect ratio of one, i.e. spherical pores, would result in a hard and isotropic rock, if the background matrix is also isotropic. Under compaction and deposition, however, pores tend to be deformed and elongated horizontally, resulting in an anisotropic rock that is soft in the z-direction but harder in the x-direction. Bandyopadhyay (2009) defined aspect ratio as a function of porosity:

$$a = \frac{(1-\phi_0)\phi}{(1-\phi)\phi_0}. \tag{11}$$

In reality, not all pores are deformed and oriented in the same direction, so we use Equation 11 to model aspect ratio for half of the pores and choose random aspect ratios for the other half.

Stochastic rock physics modeling workflow

Many parameters in the above workflow are either not certainly determined, well studied, empirical, or theoretical, such as smectite-illite reaction temperature parameters, T_T and s_T (Equation 4), critical porosity, ϕ_0, exponent parameters, m and n (Equation 8), or pore aspect ratio (used in DEM). We can account for this uncertainty and explore the model space by varying these parameters within some ranges and simulating a number of realizations of subsurface anisotropic models. The means and covariances of these models can be used to constrain anisotropic FWI. We summarize the rock physics workflow in the following:

Mineral	ρ (kg/m^3)	v_p (m/s)	v_s (m/s)	ϵ	δ	γ
Quartz	2650	6000	4000	0	0	0
Smectite	2400	3250	1500	0.1	0.05	0.18
Illite	2600	4250	2519	0.2	0.1	0.28

Table 1: Mineral properties used in our study based on Bachrach (2011).

Parameter	Value/range
Critical porosity ϕ_0	0.45
Thermal gradient	2.13^o C/100 m
Transition temperature T_T	60-90^o C
Transition spread s_T	40-80^o C
Exponents n and m	0.5-2
Aspect ratio a	0.5-1.5

Table 2: Values and ranges of parameters used in our study. Value for thermal gradient is taken from Forrest et al. (2005). Ranges for smectite-illite transition temperature and spread are based on Freed and Peacor (1989). Ranges for exponents n and m are from Bachrach (2010).

1. From uniform distributions, randomly select values for parameters like reaction temperatures, T_T and s_T (Equation 4), exponent parameters, m and n (Equations 8), and aspect ratio. Tables 1 and 2 show values and ranges of parameters used in our workflow.

2. Compute porosity for sand and shale (Equation 5) and the total porosity.

3. Model sand and shale using either DEM or Reuss average depending on porosity.

4. Compute the average stiffnesses of sand-shale layers using Backus's average (Equations 9).

5. Compute the mean models and covariances.

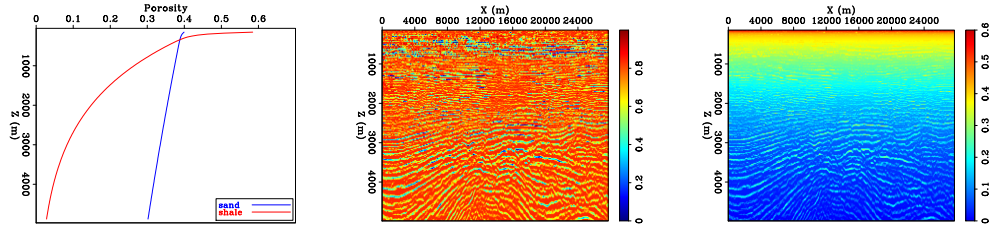
Choosing the ranges of the input parameters into our workflow is also an important task. Here is where the well data come in handy. The above simulation is usually performed several times and the input parameters are tuned so that the outputs match well data.

ROCK PHYSICS MODELING ON A FIELD DATA

We apply the above described workflow to a field data collected in the Gulf of Mexico (courtesy of WesterGeco). Figure 4(a) shows the porosity-depth trends for sand and shale (Equations 5). Note that shale's porosity decreases mush faster with depth than sand's porosity. Figures 4(b) and 4(c) show the shale volume and the total porosity. Figure 5 shows the mean velocity, ϵ, and δ. These mean models share similar structures with the shale's volume (4(b)). Anisotropy generally increases with depth, as we expect.

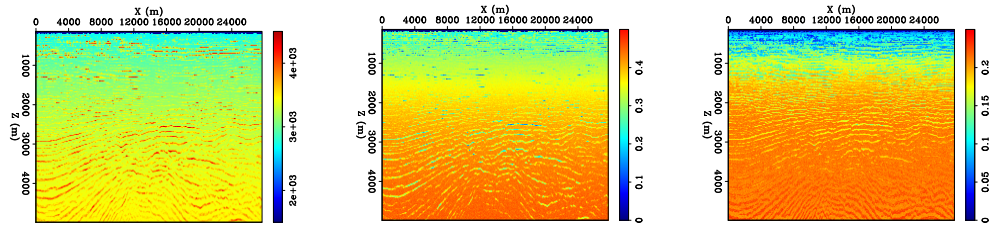
Figure 6(a) compares the probability density functions (PDF) of velocity from our rock physics simulation and well data. The simulated PDF does not match the data well. Particularly our simulation does not capture velocity range from 1600-2500 m/s. We anticipate this is because of the abrupt transition from unconsolidated rocks near the water bottom (porosity greater than critical porosity) to consolidated rocks (porosity lower than critical porosity). Figures 6(b) and 6(c) show the simulated PDFs of ϵ and δ. One noticiable feature among these simulated PDFs is that they are all trimodal.

Figure 7 shows the model variances. Notice that velocity variance is much greater than those of ϵ and δ. For all three parameters, variances increase with depth. It is more interesting to observe that velocity has negative correlations with Thomsen's parameters, indicated by negative covariances between velocity and ϵ (Figure 8(a)) and beween velocity and δ (Figure 8(c)). The two Thomsen's parameters, on the other hand, are positively correlated (Figure 8(b)). Note that we only compute these model variances and covariances at individual locations, and therefore, do not capture any spatial correlation information. Nevertheless, these covariances can still be used as preconditioners for FWI to separate crosstalk among parameters.



(a) (b) (c)

Figure 4: (a) Porosity-depth trends for sand and shale; (b) Provided shale volume (courtesy of WesternGeco); (c) Total porosity. [**CR**] huyle/. sandshaleporo,vshale,totalporo

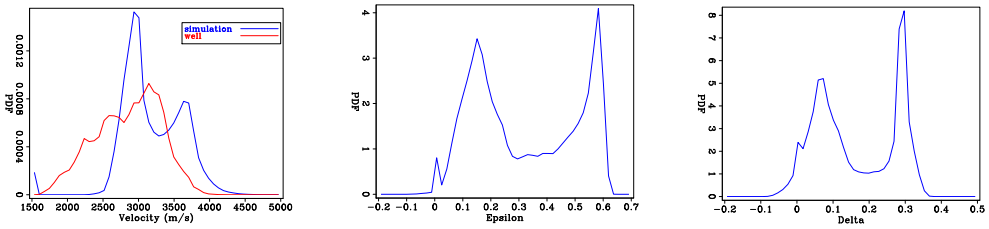


(a) (b) (c)

Figure 5: Mean models of: (a) velocity, (b) ϵ, and (c) δ [**CR**] huyle/. vmean2,epsmean2,delmean2

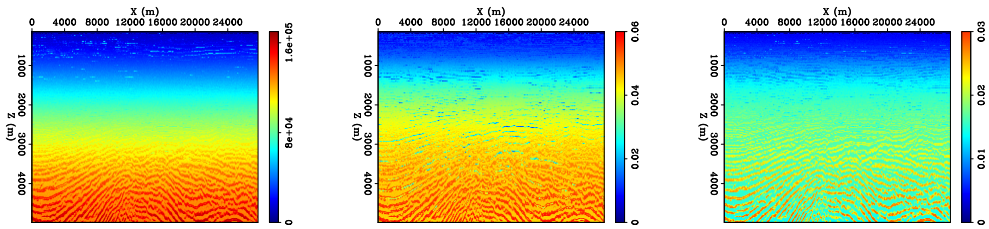
CONCLUSIONS

In this report, we present two different approaches to improve anisotropic FWI. The first approach is to estimate the source wavelet directly from the observed data by minimizing the FWI objective function. This is particularly helpful to mitigate any amplitude difference



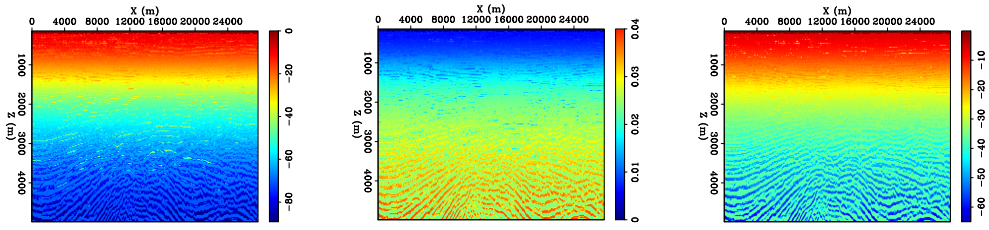
(a) (b) (c)

Figure 6: Probability density functions (PDF) of: (a) velocity (including data supplied by IHS Energy Log Services, Copyright 2017, IHS Energy Log Services Inc.), (b) ϵ, and (c) δ [**CR**] huyle/. vhis,epshis,delhis



(a) (b) (c)

Figure 7: Variances of : (a) velocity, (b) ϵ, and (c) δ [**CR**] huyle/. Cvv2,Cee2,Cdd2



(a) (b) (c)

Figure 8: Covariances of : (a) velocity and ϵ, (b) ϵ and δ, and (c) velocity and δ [**CR**] huyle/. Cve2,Ced2,Cvd2

between the modeled data and the observed data, caused by assumptions and approximations in the modeling engine. A simple synthetic example validates our implementation of source wavelet inversion. More complicated tests are to be carried out to evaluate how source estimation could improve FWI.

The second approach is to use rock physics modeling to estimate model covariances and preconditioners. We present a workflow that models sand-shale layers using Reuss average for unconsolidated rocks and Differential Effective Medium theory for consolidated rocks. The model covariances are computed by varying parameters such as smectite-illite transition temperatures, pore aspect ratio, and compaction factors. stochastically simulating many anisotropic earth models. Our application on a field data, however, does not produce satisfactory results. In particular, our simulated probability density function of velocity does not match well data. We suspect that the density function mismatch might be due to the sharp transition from unconsolidated rocks to consolidated rocks that is not captured well by Reuss average and Differential Effective Medium theory. This also signifies the importance of data calibration in order to correctly describe model parameters' statistics.

ACKNOWLEDGEMENTS

We would like to thank WesternGeco for their donation of the shale volume and IHS Energy Log Services Inc. for the well logs. We also would like to thank Iris Yang from the Stanford Rock Physics group for providing the DEM codes.

REFERENCES

Alkhalifah, T., 2000, An acoustic wave equation for anisotropic media: Geophysics, **65**, 1239–1250.

Bachrach, R., 2010, Applications of deterministic and stochasitc rock physics modeling to anisotropic velocity model building: SEG Annual International Meeting, Expanded Abstracts, 2436–2440, Society of Exploration Geophysicists.

———, 2011, Elastic and resistivity anisotropy of shale during compaction and diagenesis: Joint effective medium and field observations: Geophysics, **76**, No. 6, E175–E186.

Backus, G. E., 1962, Long-wave elastic anisotropy of hirizontal layering: Journal of Geophysical Research, **67**, 4427–4440.

Bandyopadhyay, K., 2009, Seismic anisotropy: Geological causes and its implications to reservoir gephysics: PhD thesis, Stanford University.

Bube, K. P., T. Nemeth, J. P. Stefani, R. Ergas, W. Liu, K. T. Nihei, and L. Zhang, 2012, On the instability in second-order systems for acoustic VTI and TTI media: Geophysics, **77**, No. 5, T171–T186.

Dutta, T., G. Mavko, T. Mukerji, and T. Lane, 2009, Compaction trends for shale and clean sandstone in shallow sediments, Gulf of Mexico: The Leading Edge, **28**, 590–596.

Fletcher, R. P., X. Du, and P. J. Fowler, 2009, Reverse time migration in tilted transversely isotropic (TTI) media: Geophysics, **74**, No. 6, WCA179–WCA187.

Forrest, J., E. Marcucci, and P. Scott, 2005, Geothermal gradients and subsurface temperatures in northern Gulf of Mexico: GCAGS Transactions, **55**, 233–248.

Freed, R. L. and D. R. Peacor, 1989, Variability in temperature of the smectute/illite reaction in guld coast sediments: Clay Minerals, **24**, 171–180.

Hornby, B. E., L. M. Schwartz, , and J. A. Hudson, 1994, Anisotropic effectivemedium modeling of the elastic properties of shales: Geophysics, **59**, 1570–1583.

Le, H., 2016, Anisotropic full waveform inversion : SEP-Report, **163**, 155–162.

Li, Y., B. Biondi, R. Clapp, and D. Nichols, 2016, Integrated VTI model building with seismic data, geological information, and rock-physics modeling-Part 1: Theory and synthetic test: Geophysics, **81**, C177–C191.

Sayers, C. M., 1994, The elastic anisotropy of shale: Journal of Geophysical Research, **99**, 767–774.

Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. Stekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand, 2013, Anisotropic 3D full-waveform inversion: Geophysics, **78**, No. 2, R59–R80.

Zhang, Y., H. Zhang, and G. Zhang, 2011, A stable TTI reverse time migration and its implementation: Geophysics, **76**, No. 3, WA3–WA11.

GENESIS dataset: 3D initial VTI model for time-lapse reverse time migration

Yinbin Ma, Stewart A. Levin, Biondo Biondi, Mark Meadows, and Robert G. Clapp

ABSTRACT

We estimate a 3D vertical transverse isotropy (VTI) model for the GENESIS data set based on stable inverse Dix formula. We reprocess the time-lapse seismic data set to attenuate the spatial aliasing problem. The common shot gathers and the common receiver gathers are created to enhance the subsurface illumination because the surveys were conducted with the towed streamers. Time-lapse reverse-time migration (RTM) to the GENESIS data set, with the isotropic velocity model, suggests observable velocity change near the reservoir and the overburden area as a result of production. Time-lapse RTM with the VTI model has also been studied.

INTRODUCTION

The GENESIS field has experienced reservoir compaction during production (Magesan et al. (2005)). In a recent study, negative velocity change has been observed via full-waveform inversion (FWI) and has been associated with overburden dilation (Maharramov (2016)). Therefore it is reasonable to assume the anisotropic parameters near the reservoir also change because of geomechanical effects associated with production.

We estimate an initial 3D VTI model for the purpose of wave propagation with VTI wave equation. The interval normal moveout (NMO) velocity and η model as a function of travel time are obtained by a stable Dix inversion with a smoothness constraint applied on the model. Anisotropic parameter δ is constructed as a function of depth/travel time only because we do not have prior knowledge on δ. The δ has value that is typical in the Golf of Mexico environment. Finally, the interval vertical velocity, η, δ and ϵ model are constructed as functions of depth using the estimated time-to-depth table.

The time-lapse surveys with data collected by the towed streamers over the GENESIS field have been cross equalized to improve the similarity between baseline survey and monitor survey (Magesan et al. (2005)). Each mid-point gather has 30 different offsets, ranging from 1146 ft to 15414 ft with spacing 492 ft. From the prospective of the wave propagation, the data is aliased in space and therefore reprocessing is required. We design a workflow to interpolate the data along the offset axis in order to attenuate the spatial aliasing problem. Subsurface angular illumination is limited because of the acquisition acquisition geometry (towed streamer). We extract both source gathers and receiver gathers to enhance the illumination using the principle of reciprocity.

We analysis the quality of VTI models and data processing procedure by examining the subsurface angle gathers with the isotropic and VTI wave equation (Zhang et al. (2011)). RTM with the isotropic velocity model is first computed on the baseline and

monitor datasets. RTM images near the reservoir are shifted vertically, indicating velocity change near the reservoir and possibly in the overburden area. Time-lapse RTM with our constructed VTI models is also studied.

CREATING 3D INITIAL VTI MODEL FOR THE GENESIS DATASET

We receive stacking NMO velocity and η from the previous time-lapse study on the GENESIS dataset (Magesan et al. (2005)), as functions of mid-point and travel time. For the purpose of wave propagation, we need smooth interval vertical velocity and anisotropic parameters in the depth domain. In this section, we describe our approach to extract a 3D initial VTI model for the GENESIS dataset.

The stacking NMO velocity and parameter η are obtained as (Wang and Tsvankin, 2009),

$$\left(V_{NMO,stack}(N)\right)^2 = \frac{1}{t_0(N)} \sum_{i=1}^{N} \left(V_{NMO,int}(i)\right)^2 \left(t_0(i) - t_0(i-1)\right), \quad (1)$$

$$\eta_{stack}(N) = \frac{\sum_{i=1}^{N} \left(V_{NMO,int}(i)\right)^4 \left(1 + 8\eta_{int}(i)\right) \left(t_0(i) - t_0(i-1)\right)}{8 \left(V_{NMO,stack}(N)\right)^4 t_0(N)} - \frac{1}{8}, \quad (2)$$

where $V_{NMO,stack}(N)$ is the stacking NMO velocity at layer N, $t_0(N)$ is the zero-offset travel time at layer N, $V_{NMO,int}(i)$ is the interval NMO velocity at layer i, $\eta_{stack}(N)$ and $\eta_{int}(i)$ are the stacking and interval parameter η separately. Direct estimation of interval parameters from equations 1 and 2 is unstable and leads to spatially non-smooth model because each mid-point is processed independently.

We design a stable algorithm to construct the smooth VTI model based on equations 1 and 2. Interval NMO velocity is estimated first because it does not depend on parameter η. We define $\mathbf{m} = V^2_{NMO,int}$, $\mathbf{d} = V^2_{NMO,stack}$, where both $\mathbf{m}$ and $\mathbf{d}$ are functions of mid-point and travel time. Equation 1 can be represented as $\mathbf{d} = \mathbf{L}\mathbf{m}$ with $\mathbf{L}$ representing the Dix equation. We construct a smooth interval NMO velocity by solving the optimization problem,

$$\mathbf{m} = \text{argmin}_{\mathbf{m}} \left\{ \frac{1}{2} \|\mathbf{L}\mathbf{m} - \mathbf{d}\|_2^2 + \frac{c_x}{2} \|\nabla_x^2 \mathbf{m}\|_2^2 + \frac{c_y}{2} \|\nabla_y^2 \mathbf{m}\|_2^2 + \frac{c_t}{2} \|\nabla_t^2 \mathbf{m}\|_2^2 \right\}, \quad (3)$$

where ∇_x^2, ∇_y^2 and ∇_t^2 are the Laplacian operators along spatial direction x, y and travel time separately, which promote the smoothness of the model. The coefficients c_x, c_y and c_t control the strength of constraint on smoothness.

Ideally, smooth constraint on the slowness leads to better results as the travel time is preserved. The downside is that in equation 3, the data fitting term and the regularization terms cannot be linear at the same time. A nonlinear optimization problem must be solved which increases the challenge to get a reasonable initial model.

Once the interval NMO velocity $V_{NMO,int}$ is computed, the interval η model can be con-

structed in a similar fashion. The other VTI parameters can be estimated with,

$$V_x^2 = V_z^2(1+2\varepsilon), \tag{4}$$
$$V_{\text{NMO}}^2 = V_z^2(1+2\delta), \tag{5}$$
$$\eta = \frac{\varepsilon-\delta}{1+2\delta}, \tag{6}$$

where η, δ, ε are the anisotropic parameters, V_x is the horizontal velocity, V_z is the vertical velocity. In Figure 1 and 2, we show our inverted VTI model.

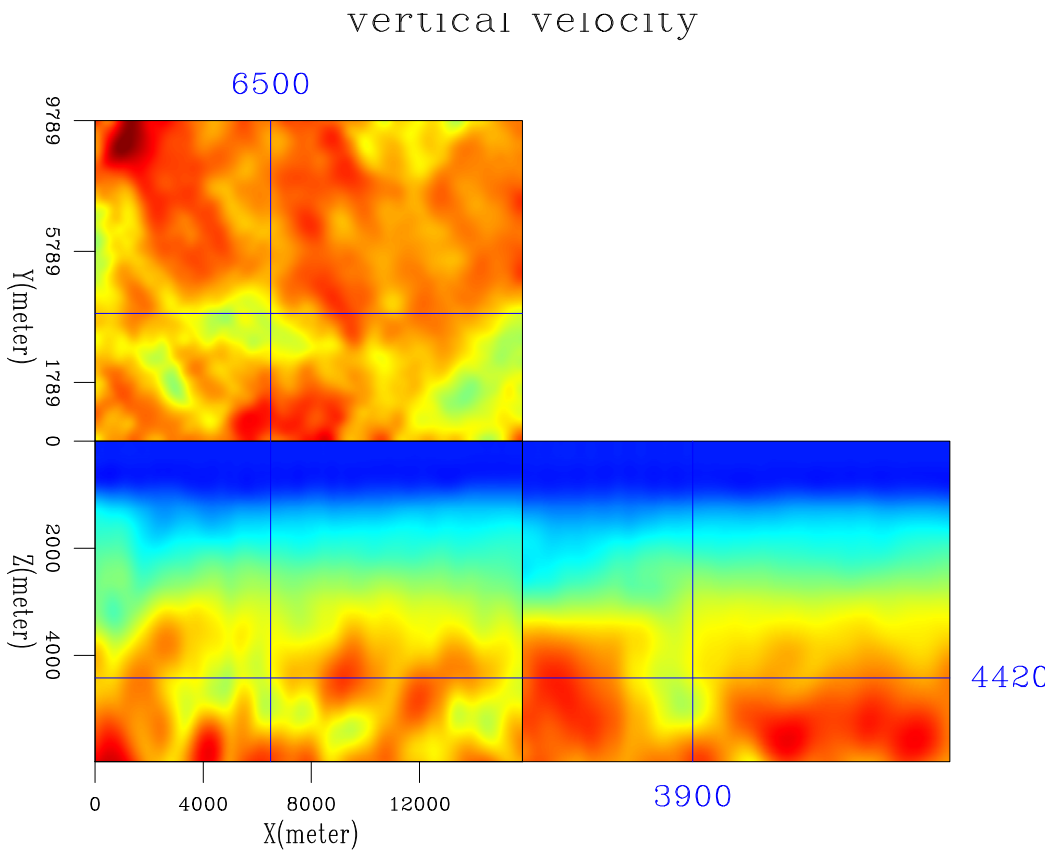


Figure 1: Initial vertical velocity model. [**ER**] yinbin1/. VzInitial

DATA PROCESSING

To improve the quality of wave propagation and RTM images, we interpolate the data along the offset axis to attenuate the spatial aliasing problem. We design the following workflow:

1. estimate RMS velocity from CMP gathers;
2. apply NMO correction to flatten the gathers;
3. interpolate to dense grid in the CMP domain using sinc interpolation;
4. after interpolation, apply inverse NMO to recover correct kinematics;

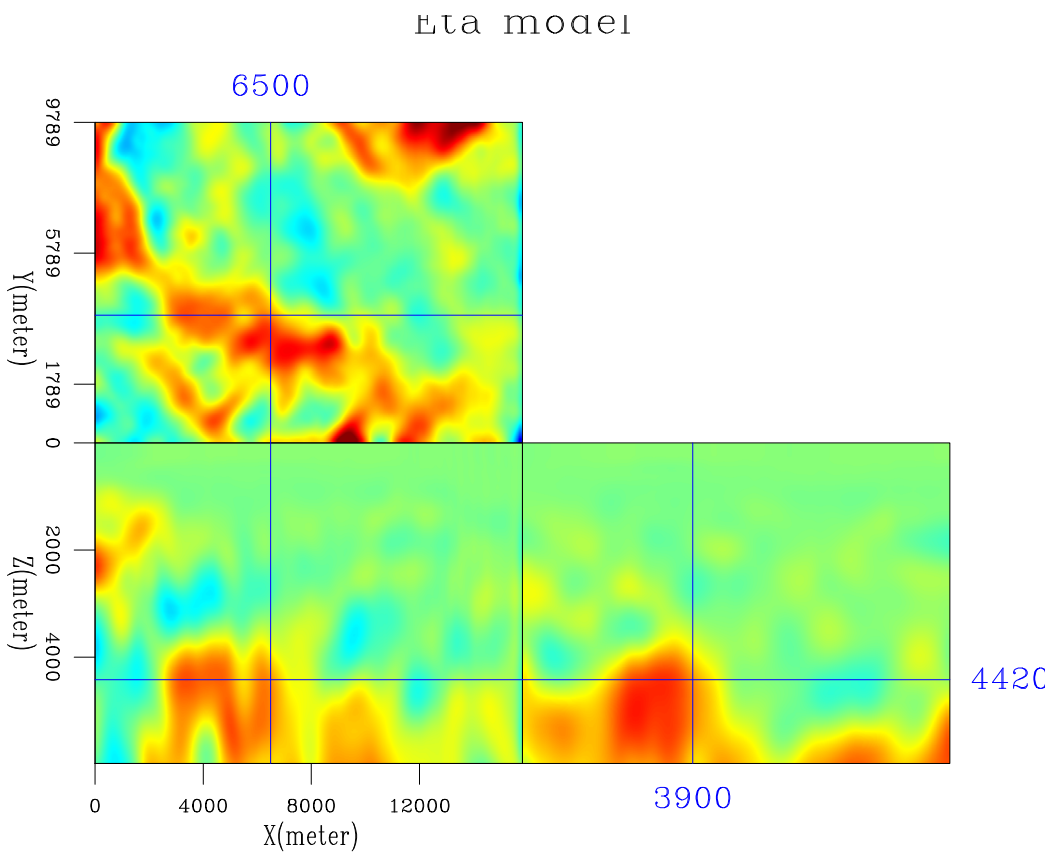


Figure 2: Initial η model. [**ER**] yinbin1/. EtaInitial

5. sort CMP gathers to shot gathers and receiver gathers.

After the proposed workflow, we lose far offset data from shallow target, because a stretch has been applied during the NMO correction, as can be seen in Figure 3. Reflection from the reservoir is not affected though.

The towed streamer acquisition has illumination primarily from positive or negative subsurface angle. To enhance the quality of RTM image, especially common image gathers (CIG), both source gathers and receiver gathers are used in the RTM process. Unfortunately, the results from the source gathers are not ready by the time the paper is written, therefore not included in this report.

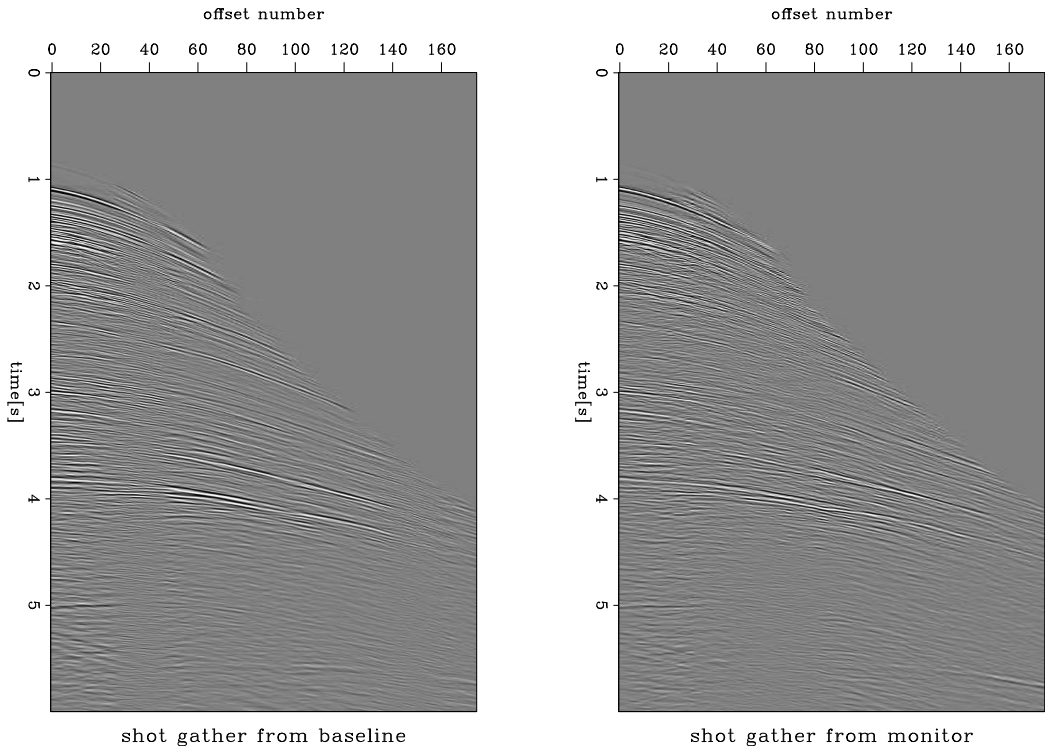


Figure 3: Left: one shot gather from the baseline survey after our processing workflow. Right: one shot gather from the monitor survey. [**NR**] yinbin1/. dataSample

TIME-LAPSE RTM BASED ON VTI WAVE EQUATION

In this section we describe the implementation of RTM based on VTI wave equation. We use the the following equation,

$$\frac{1}{V_z^2}\partial_t^2 \begin{pmatrix} p \\ r \end{pmatrix} = \begin{pmatrix} 1+2\varepsilon & \sqrt{1+2\delta} \\ \sqrt{1+2\delta} & 1 \end{pmatrix} \times \begin{pmatrix} \partial_x^2+\partial_y^2 & 0 \\ 0 & \partial_z^2 \end{pmatrix} \begin{pmatrix} p \\ r \end{pmatrix}, \tag{7}$$

where p and r are defined as the horizontal and vertical stress components. ε and δ are the anisotropic parameters. Equation (7) reduces to isotropic case when $\varepsilon = 0, \delta = 0$.

The imaging condition for the velocity perturbation, can be described with the following equation,

$$I(z, x, h_x) = \sum_t U_s(t, z, x - h_x)\partial_t^2 U_r(t, z, x + h_x), \tag{8}$$

where h_x is the subsurface offset, $I(z, x, h_x)$ is the subsurface image in the offset domain, U_s is the source pressure wave field (average of horizontal and vertical stress components), and U_r is the receiver pressure wavefield.

We implement the wave propagation in equation 7 and imaging condition in equation 8 on a GPU node with 8 Tesla K80 available. To reduce the data transfer between the GPU memory and the main memory, we use random boundary condition for the RTM, with the following workflow,

1. generate a velocity model from the initial velocity with random boundary;
2. forward propagate the source wavefield to the end, and save the last 2 snapshots of wavefield;
3. back propagate the source wave field with random boundary, and in the mean time back propagate the receiver wavefield with absorbing boundary;
4. apply imaging condition in equation 8 every 4 ms.

4D RTM with isotropic velocity model and VTI model

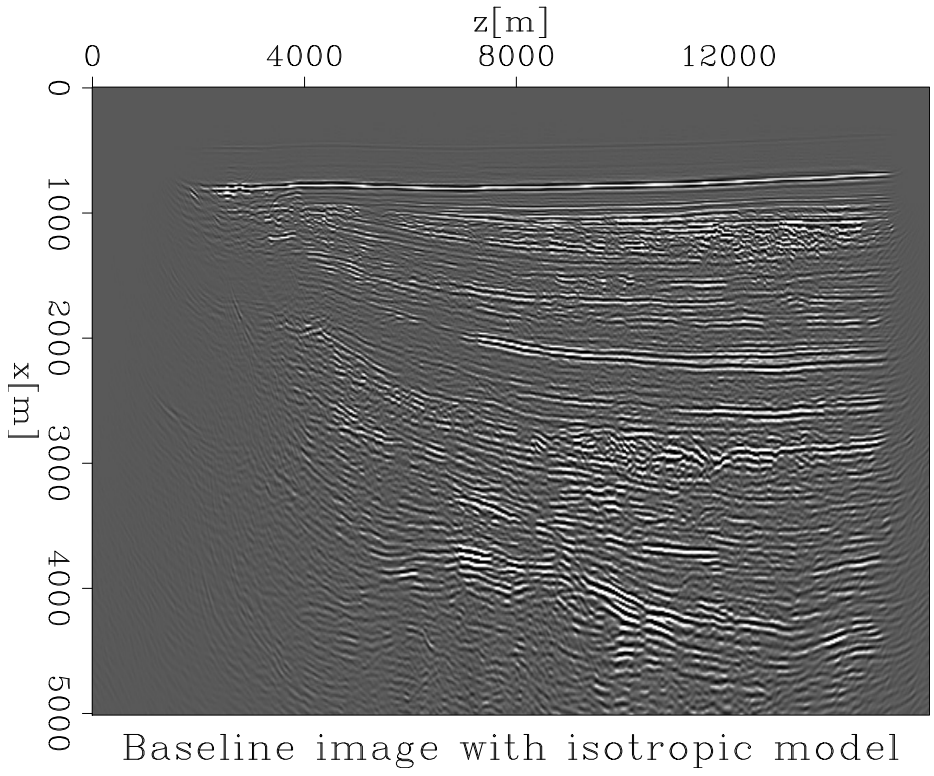
In this subsection, we study the time-lapse RTM with the isotropic model and VTI model. For isotropic model, the vertical velocity is replaced by NMO velocity, and anisotropic parameters are set to zero in the VTI wave equation.

The baseline images at zero subsurface offset are shown in Figure 4. The RTM image with isotropic velocity model is stretched comparing with the VTI model case because isotropic model has higher vertical velocity. The monitor images are shown in Figure 5. The time-lapse differents are shown in Figure 8. To get a clear view of the image near the reservoir, we show a closer view of baseline and monitor image in Figure 6 and Figure 7.

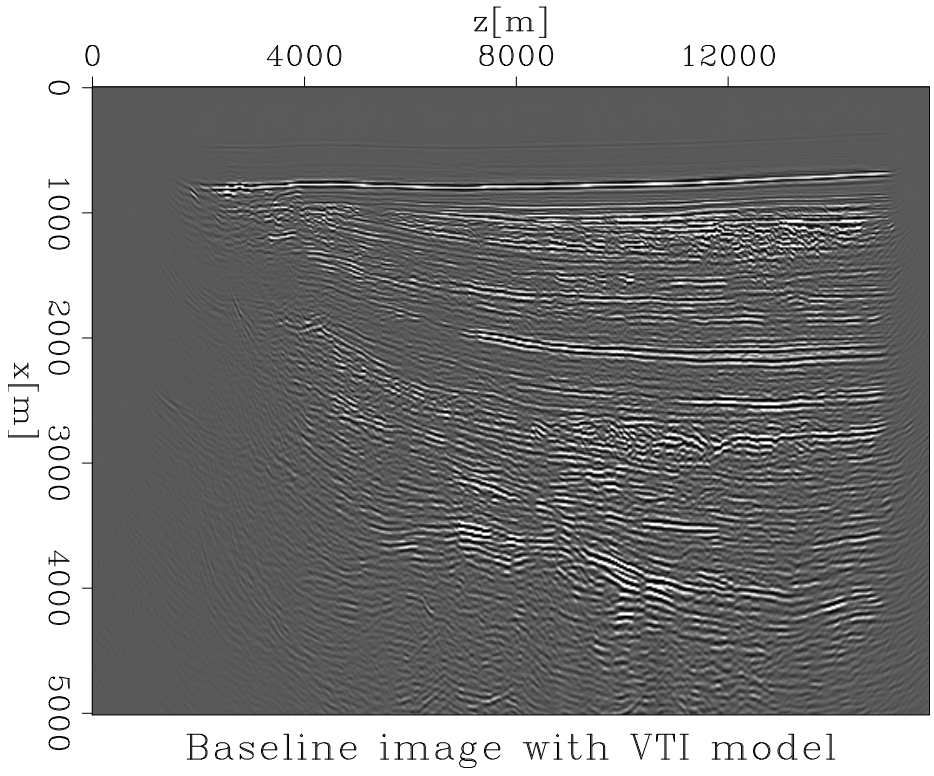
The effect of anisotropic is more observable in the common image gathers. We obtain the angle domain common image gathers (ADCIGs) for baseline and monitor dataset, with isotropic and VTI wave equation separately. The ADCIGs for the baseline dataset are shown in Figure 9, and ADCIGs for the monitor dataset are shown in Figure 10. The results are still preliminary and no conclusion has been made at this stage.

CONCLUSION

In conclusion, we estimate a 3D VTI model for the GENESIS dataset. We reprocess the time-lapse seismic data set to attenuate the spatial aliasing problem. Time-lapse RTM with the isotropic model and VTI model are both examined. ADCIGs suggest more work need to be done before we can make conclusion on the dataset.

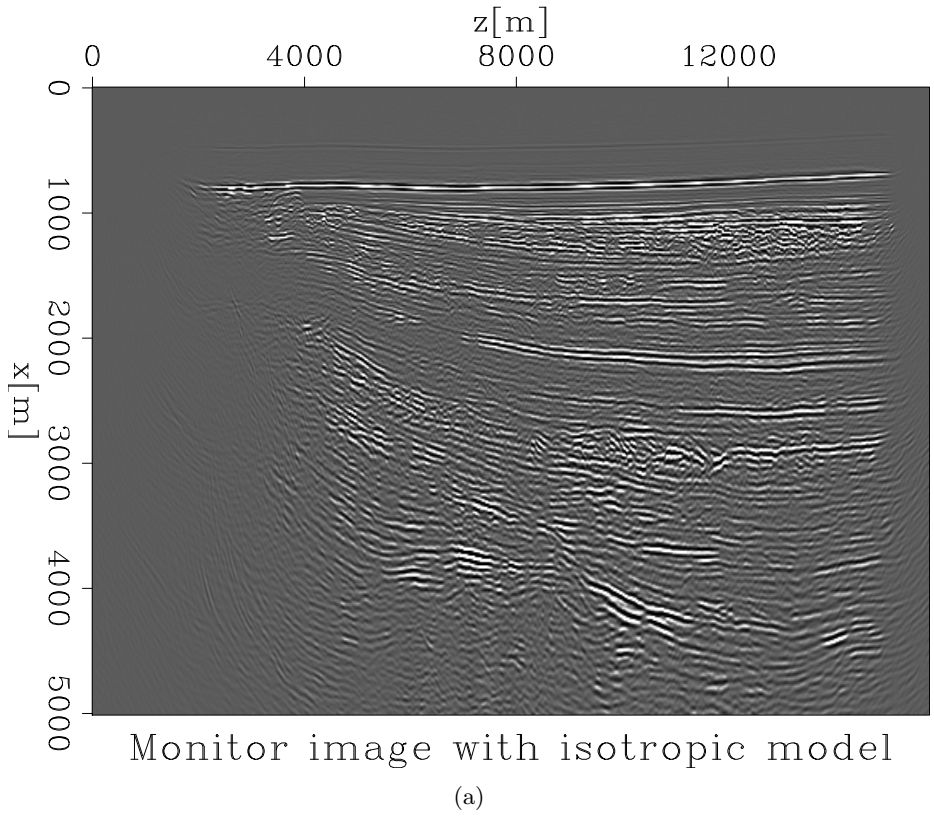


(a)

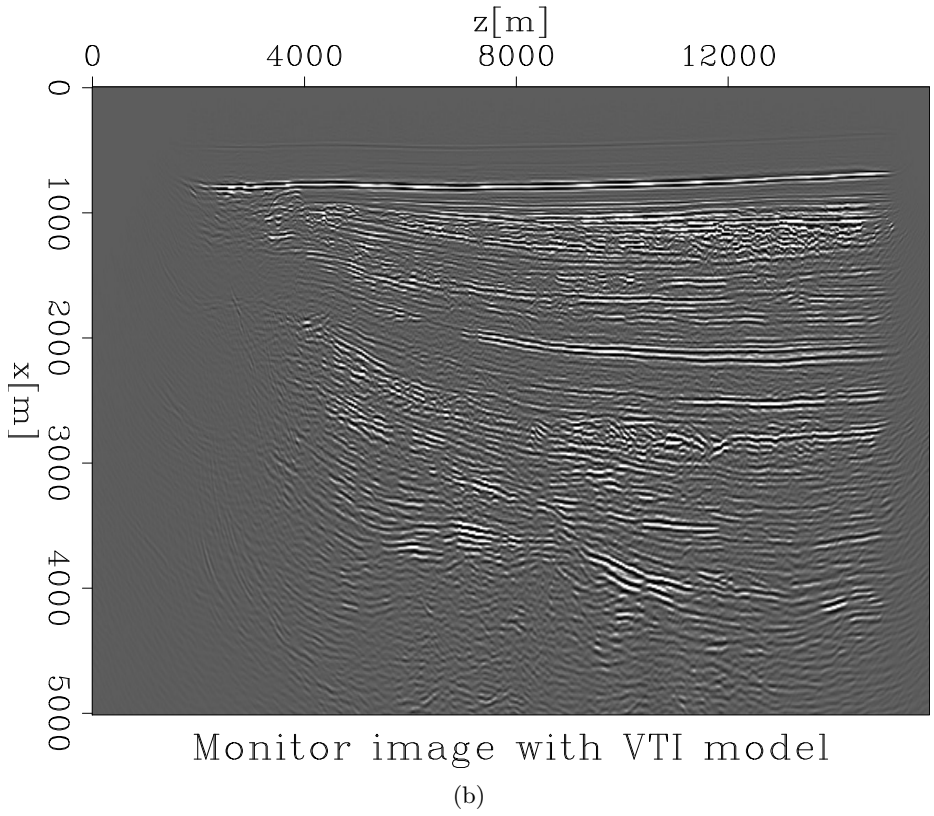


(b)

Figure 4: Baseline RTM image. Top: RTM with isotropic wave equation. Bottom: RTM with VTI wave equation. [**CR**] yinbin1/. Base.ISO.RTM,Base.VTI.RTM

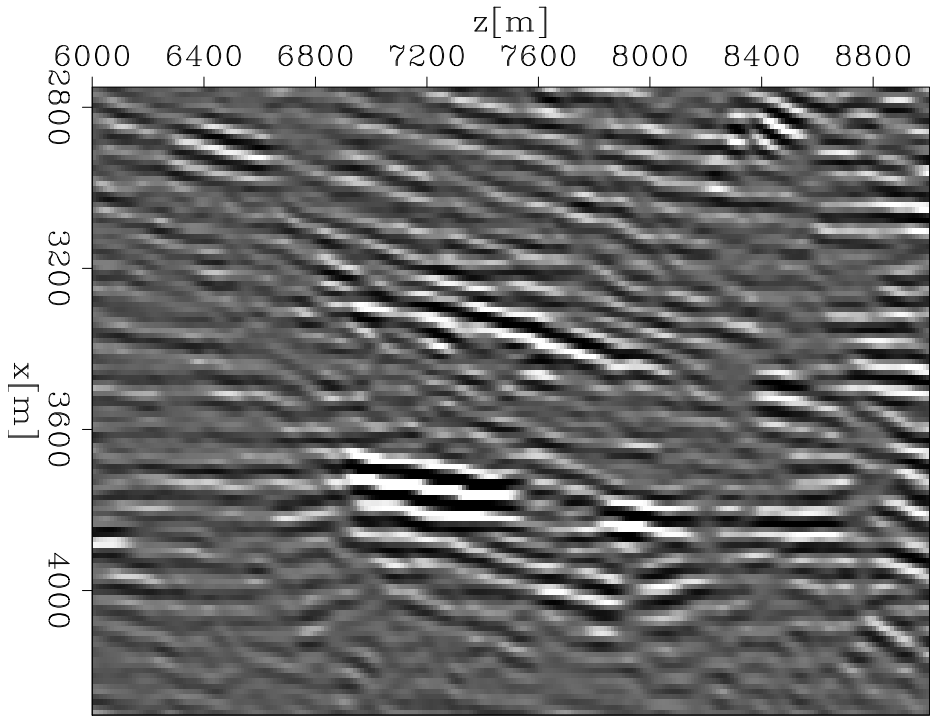


(a)

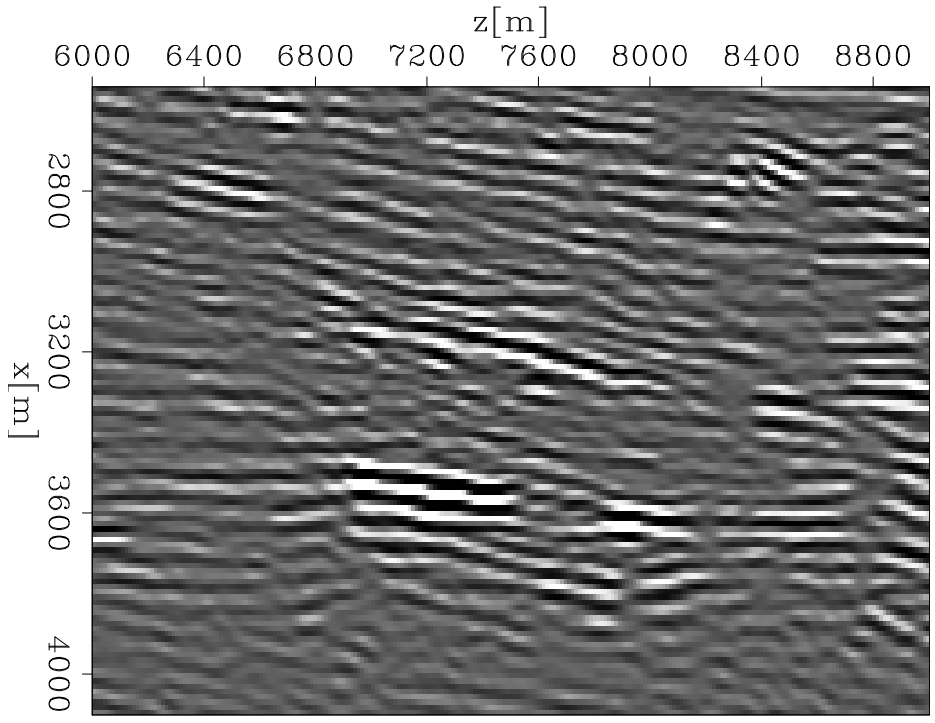


(b)

Figure 5: Monitor RTM image. Top: RTM with isotropic wave equation. Bottom: RTM with VTI wave equation. [**CR**] yinbin1/. Mon.ISO.RTM,Mon.VTI.RTM

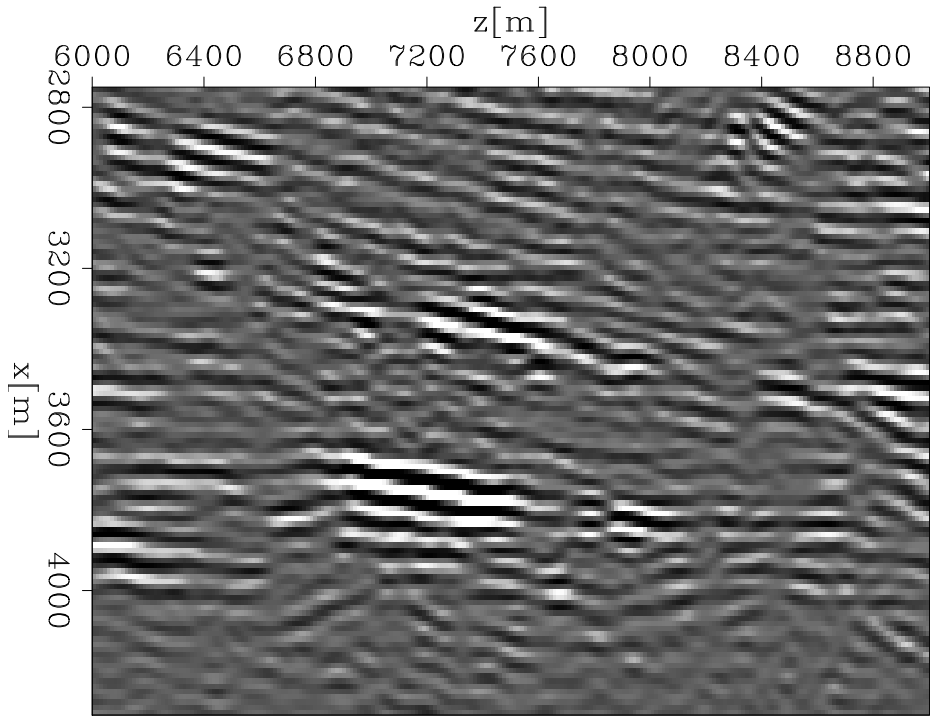


(a)

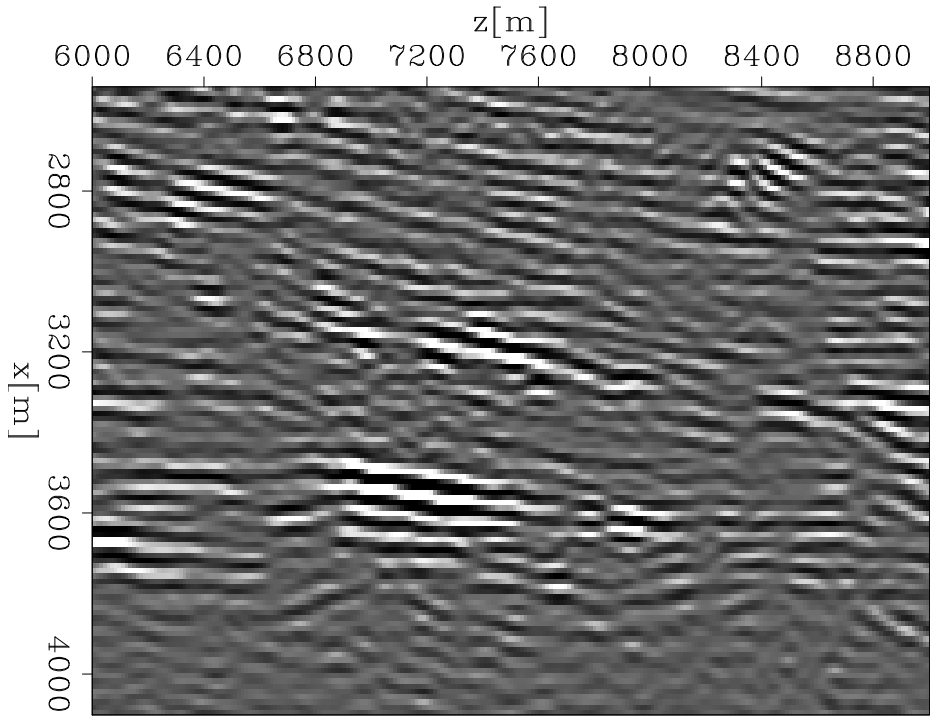


(b)

Figure 6: Baseline RTM image near the reservoir. Top: RTM with isotropic wave equation. Bottom: RTM with VTI wave equation. [**CR**] yinbin1/. base.ISO.Zoom,base.VTI.Zoom

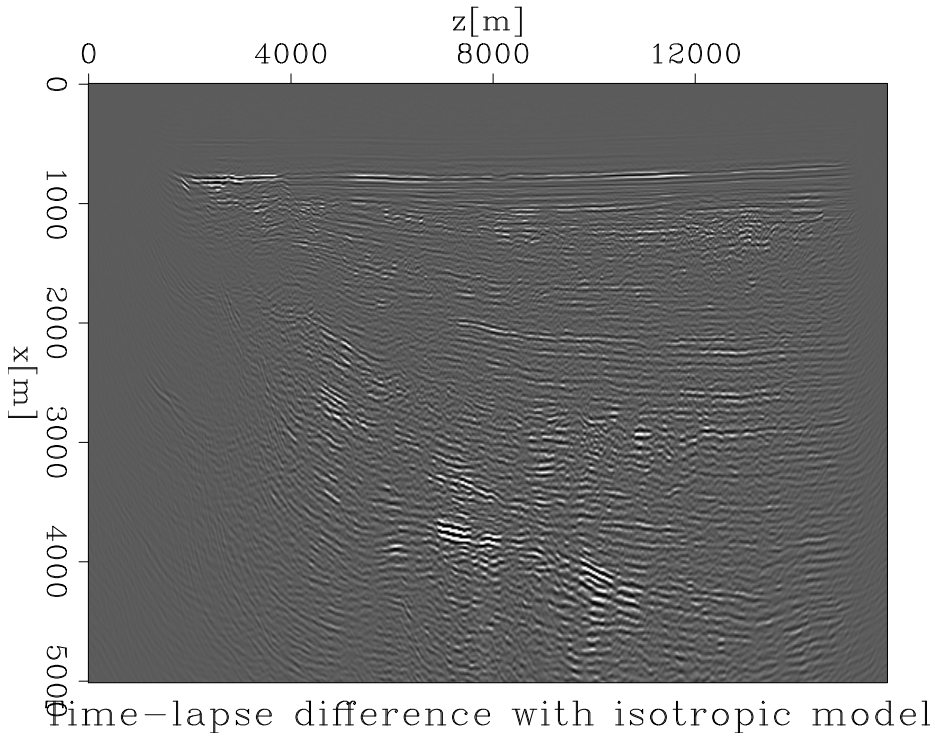


(a)

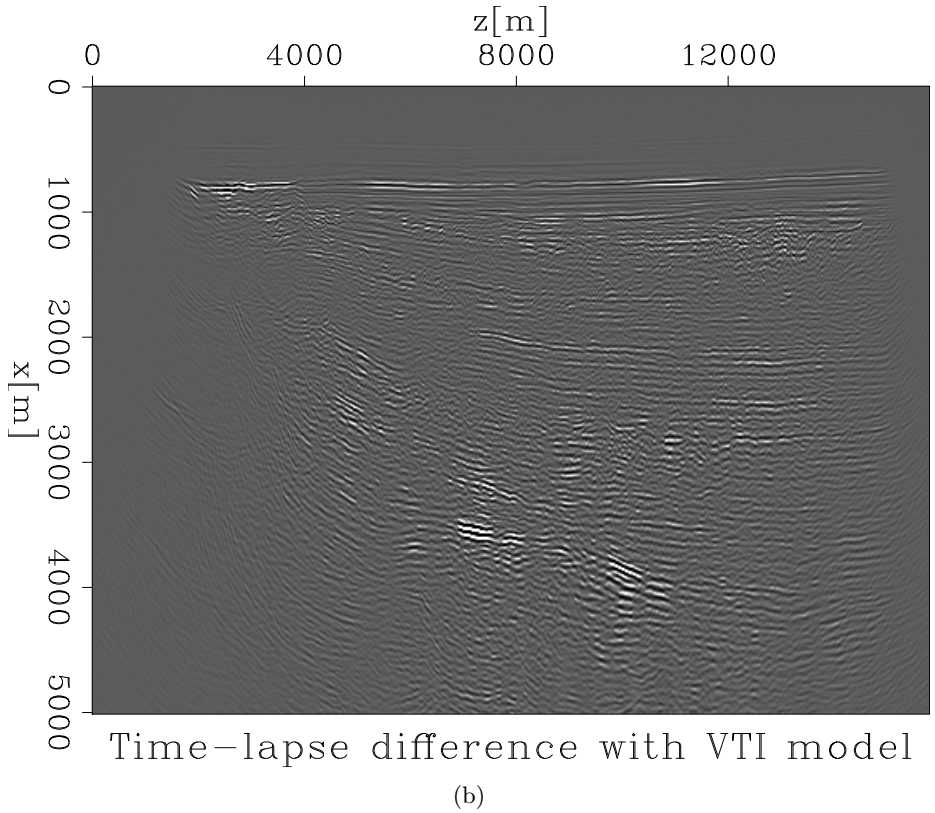


(b)

Figure 7: Monitor RTM image near the reservoir. Top: RTM with isotropic wave equation. Bottom: RTM with VTI wave equation. [**CR**] yinbin1/. mon.ISO.Zoom,mon.VTI.Zoom

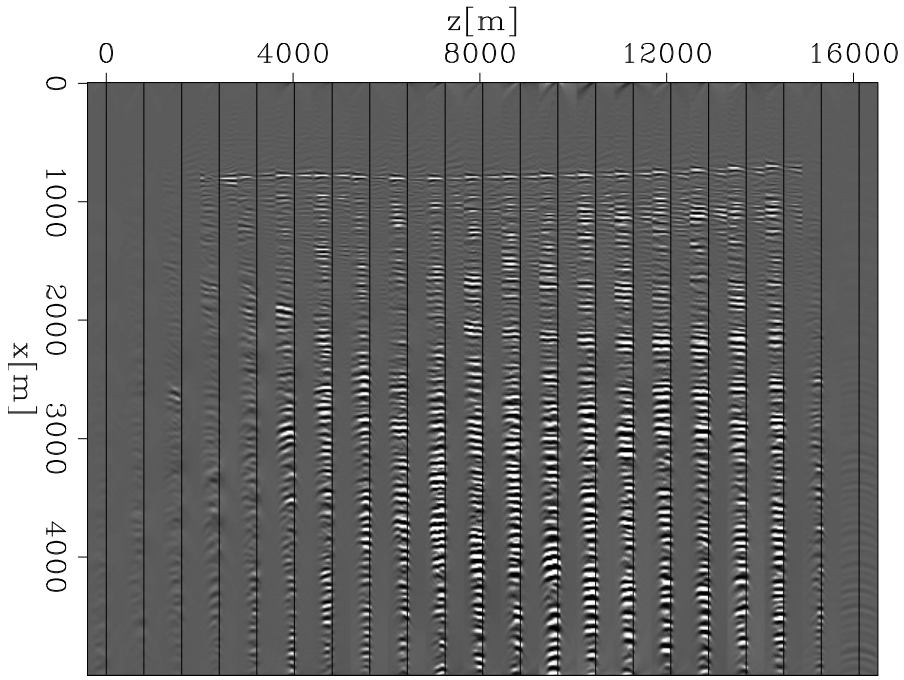


(a)

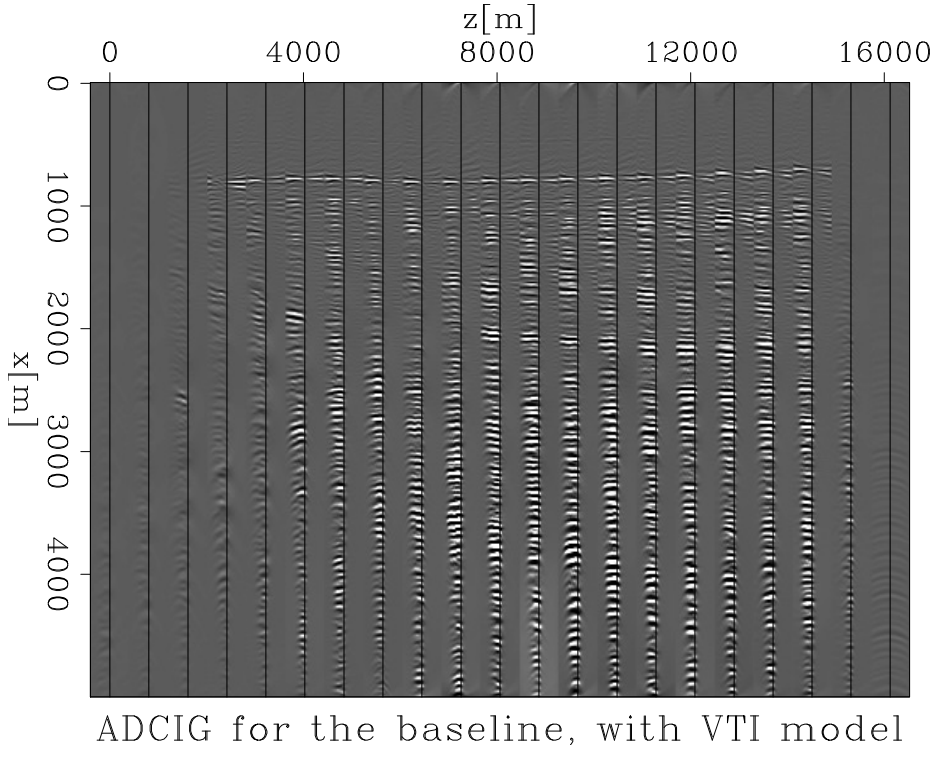


(b)

Figure 8: Time-lapse RTM image, pclip=99.9 is applied to identify the location of change. Top: RTM with isotropic wave equation. Bottom: RTM with VTI wave equation. [**CR**] yinbin1/. 4D.ISO.RTM,4D.VTI.RTM

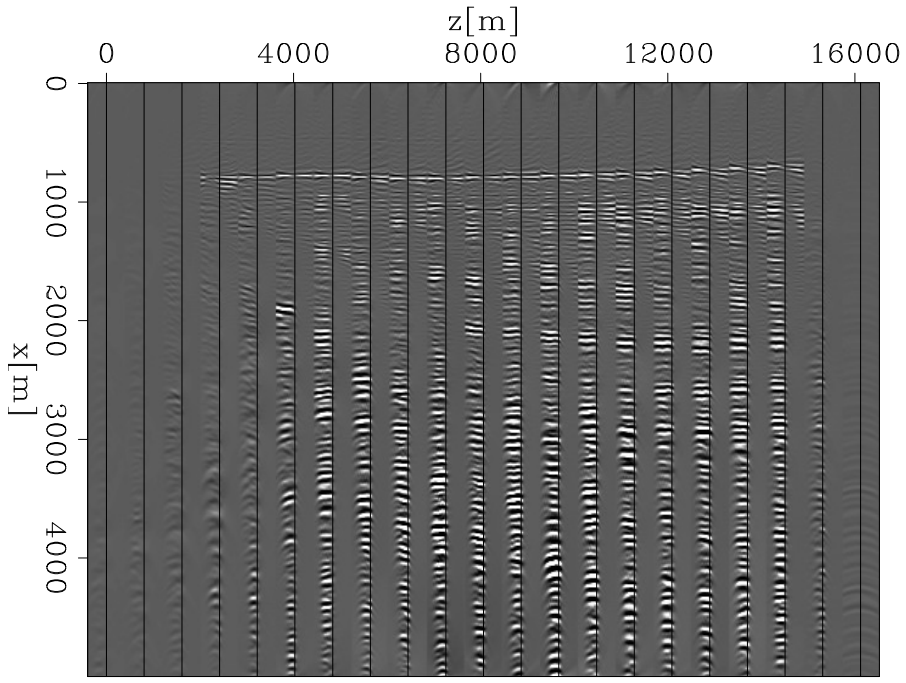


(a)

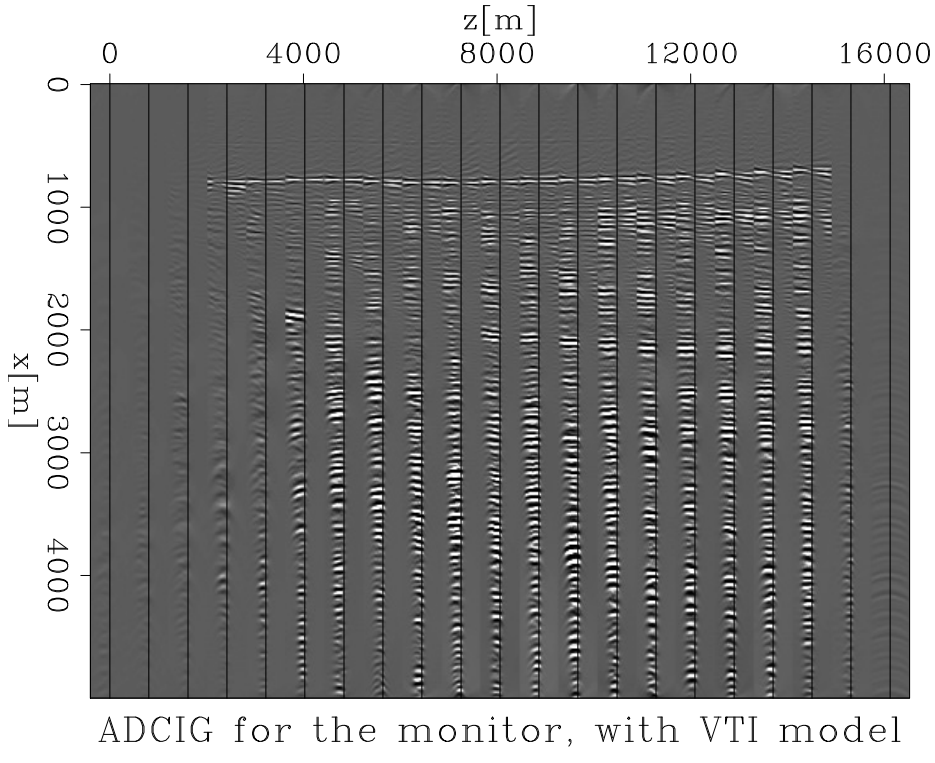


(b)

Figure 9: ADCIG for the baseline survey. White line indicate the mid-point of each angle gather, and the angle range from $-30°$ to $+30°$. Top: ADCIG with isotropic wave equation. Bottom: ADCIG with VTI wave equation. [**CR**] yinbin1/. baseISO.ADCIG,baseVTI.ADCIG



(a)



(b)

Figure 10: ADCIG for the monitor survey. White line indicate the mid-point of each angle gather, and the angle range from $-30°$ to $+30°$. Top: ADCIG with isotropic wave equation. Bottom: ADCIG with VTI wave equation. [**CR**] yinbin1/. monISO.ADCIG,monVTI.ADCIG

ACKNOWLEDGEMENT

We would like to thank Chevron to share the GENESIS data set with us.

REFERENCES

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

Maharramov, M., 2016, Time-Lapse Inverse Theory: SEP Report-165, Ph.D. Thesis.

Wang, X. and I. Tsvankin, 2009, Estimation of interval anisotropy parameters using velocity-independent layer stripping: GEOPHYSICS, **74**, WB117–WB127.

Zhang, Y., H. Zhang, and G. Zhang, 2011, A stable tti reverse time migration and its implementation: GEOPHYSICS, **76**, WA3–WA11.

Making Marine Data From Land Data: suppressing 2-D surface noise on a 1-D line

Jon Claerbout

ABSTRACT

We model a land survey data line (t, x) to be a marine line with the addition of out-of-line $(x, y, z = 0)$ surface scatter. To begin, take a single surface line scatterer crossing the survey line at some angle. Its effect in the (t, x)-space is to add an upside-down Vee shaped event for every reflector. The many earth layers produce a stack of nested Vees, each the same but for a shift corresponding to the reflector depth and amplitude that of the reflector. Mathematically, the 2-D Vee patterns convolve on the 1-D layers. With many surface line scatterers we obtains many stacks. The making of marine data from land data is essentially a 2-D deconvolution problem. Luckily, in principle, 2-D PEFs can handle spatially aliased data (so long as there are not "too many" events). Luckily we also have recently developed techniques for nonstationary PEFs. Such PEFs facilitate curved surface scattering lines.

INTRODUCTION

In the 1980s the oil industry moved offshore. Drilling got a whole lot more expensive. Seismology got a whole lot cheaper. Hooray!

Now the industry is going back on shore. Uh, oh! Maybe we had better start thinking about land data again!

Land data is commonly understood to be really noisy. But this data is also known to be repeatable to good precision. The real trouble is that our ability to model such data is so poor. Although we say the data is noisy, we actually mean we don't have feasible modeling methods. A feasible method is one we can compute so readily that iterative fitting becomes practical.

I present here a feasible way to model land data. I model land data as marine data supplemented by surface scatterers. The surface scatterers could be *point* scatterers, but surface *line* scatterers are also a possibility. Indeed, surface line scatterers are more geologically plausible and lead to a more parsimonious parameter space.

MOTIVATION FROM AN OLD SURVEY LINE

The difference between land seismic data and marine seismic data is the complex near-surface land material. It affects waves both in transmission (statics) and by scattering. Upcoming waves mainly reflect back down, but they also scatter sideways into surface ground roll.

Figure 1: A bit of a shot gather with 3.4m receiver separation. Focus on the scatterer near the middle of the line near the top. See an upside-down Vee pattern. Each layer gives rise to a Vee. The slope of each Vee is about 130 m/s. That happens to match the slope of the second arriving ground roll train, theoretically a very shallow penetrating surface wave. [**NR**] jon2/. shotgath

Figure 1 is windowed from an illustration on page 127 in my textbook BEI[1]. It shows a survey line (shot gather) containing a single strong surface scatterer, an object halfway down the line. Here the upcoming wave from each deep layer scatters waves outgoing and incoming (toward the shot). Thus the data shows a slow inverted Vee topped at each of the many subsurface reflectors.

That each reflector shows the same shaped Vee suggests statistical averaging to produce the best estimate of the 2-D Vee waveform. Vee patterns are highly predictable and destructible beyond aliasing by (t, x) filtering with a 2-D PEF (prediction-error filter). The scattering point source is defined by applying the inverse Vee filter.

This data shows only a single surface location scatterer making all the Vees, but we may presume that every other surface location along the survey line likewise scatters (or could scatter) more Vees. This gives a clear model for why land data looks so much noisier than marine data; furthermore it suggests a potential inversion: land data to marine data.

Out-of-line surface scatterers

We do not know whether this dataset results from a point scatterer on the survey line, or if there is a line scatterer intersecting the survey line. We need a model for 3-D data (an $(x, y, z = 0)$ surface containing scattering elements). We get it here next!

Consider a surface point scatterer off to the side of the survey line. Its consequences are not simple. There is no Vee response, instead there is the flank of a (t, x, y)-space

[1] `http://sep.stanford.edu/sep/prof/bei11.2010.pdf`

hyperbola. The 3-D world suddenly looks far more complicated than the 2-D world. But, there is an escape! We have the option of parameterizing the surface scatterers by lines instead of points. This simplifies matters immensely. Which is better, lines or points?

It is easy to visualize line scatterers geologically. A river or fossil river in near surface sediments is defined by a curved line on the surface. Luckily, curving wavefronts have recently become an easily manageable complication in data processing. Using nonstationary PEFs we can destroy curved arrivals with a space-variable line-destruction PEF. Two numbers locate a point. Two numbers locate a line. Would lines normally allow a more parsimonious fitting than points? I think so.

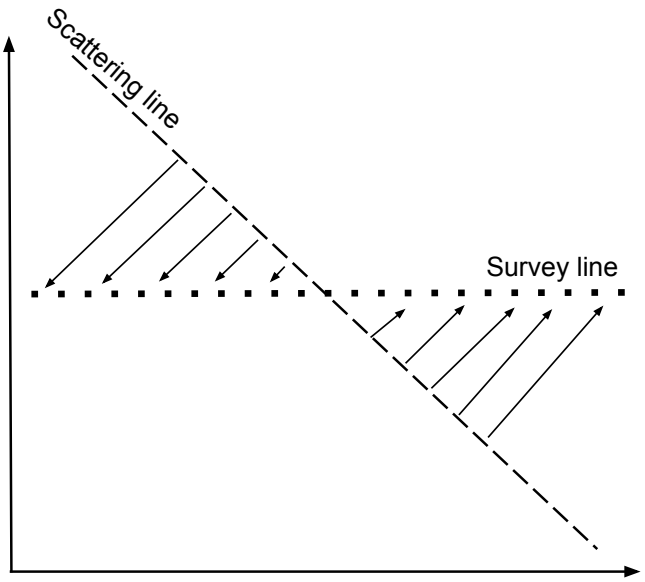


Figure 2: View of a surface plane containing a line scatterer. A vertically incident impulsive plane wave hits the scattering line. Travel time to the survey line is seen in the length of the rays (the arrows). The data (t, x) plane contains an impulse along the trajectory $v(t - t_0) = |x - x_0| \cos\theta$ where v is the surface wave velocity and x_0 is the location of the line crossing. [**NR**] jon2/. vertical

The simplest case to consider is deeply reflected waves emerging straight up. Figure 2 shows the rays propagating from the scattering line to the survey line. Travel time is the ray length, so the observations are again Vee patterns centered where the scattering line crosses the survey line. What differs from the elementary 2-D case is the angle of the Vees. The Vees now become faster, i.e. closer to the horizontal in (t, x). (A special case is a scattering line parallel to the survey line.)

Figure 3 introduces the complication of a wavefront not emerging straight up, say an impulse emerging at an angle. Let it propagate along the scattering line from northwest to southeast. Wavefronts are no longer parallel to the scattering line. Rays are perpendicular to the wavefronts. Travel time is ray length from the scattering line to the survey line. The Vees are no longer symmetric in x. Minimum surface wave travel time remains at the Vee top where survey and scattering lines cross. The backscatter wavefront on the left races faster over the survey line than the forward scatter wavefront on the right. Thus, the Vee tilts pointing more to the right. Perhaps (I don't know.) the Vee is simply picking up the stepout of the upcoming wave.

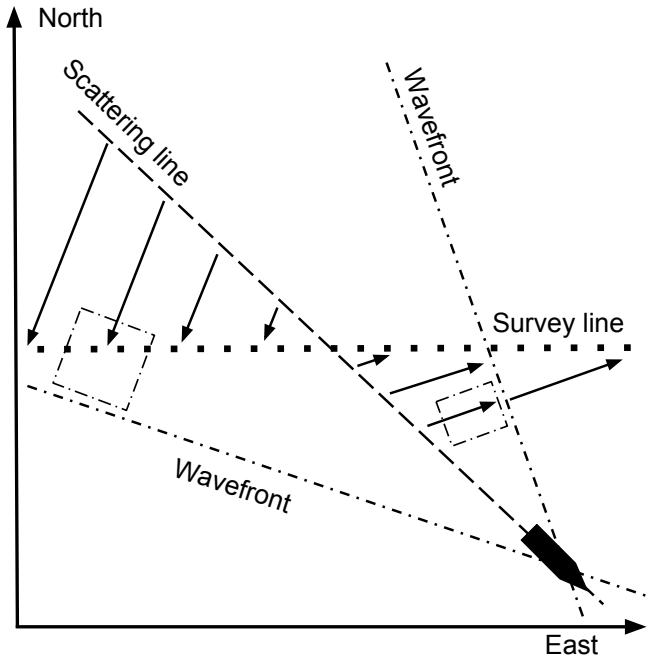


Figure 3: View of a surface plane containing a line scatterer. A non-vertically incident impulsive plane wave hits the scattering line. The impulse flies like a rocket along the scattering line. Wavefronts are no longer parallel to the scattering line. The arrows are perpendicular to the wavefront. Travel time to the survey line is again the length of the rays. The Vee pattern in (t, x) space is now a tilted Vee. [**NR**] jon2/. tilted

2-D SPECTRUM

Various surface scattering lines introduce variously sloped Vee's. It might be that Normal Moveout Correction does a good job of symmetrizing the Vee patterns. To begin with we may address data with limited numbers of curved surface scattering lines. This situation is ripe for estimating 2-D PEFs from the 2-D spectrum of moved out data. Further, the PEFs may be spatially variable to account for curving surface scatterers.

Complications and generalizations

What if the earth really is more like point scatterers than like line scatterers? Then we will not see Vees. These Vees become nested identical hyperbolas. These are unlike the nested hyperbolas we see on the usual reflection gather. Each of these hyperbolas has the same asymptotic slope but not the same asymptote. These repeated hyperbolas all have the same $\tau = z/v$-value. They differ only in their time location (that being determined by the time to deep layer structures). So, the nature of the problem remains time convolutional although the 2-D filters are no longer simple Vee shapes. We can still readily estimate such filter inverses by familiar techniques.

What about statics? Does our approach require data with no statics on it? Good question. It happens that the data shown here has no apparent statics, but what if it did? We might find statics simply imposed upon the Vees and hyperbolas. The problem then amounts to deconvolution with an unstructured 2-D filter. The simple Vee structure is lost, but the method appears much the same. Perhaps this approach may be seen as a reformulation of the traditional statics problem.

CONCLUSION

On what important truth do very few people agree with you? —Peter Thiel

I feel this is a winner of a project for graduate students. It should produce interesting results on field data almost immediately. Intriguingly, there are many prospects for extensions: (1) what about midpoint-offset space? (2) what about 3-D coverage?

The traditional product that comes closest to these ideas is (f, k) steep dip rejection. But that process does not cope with spatial aliasing. Nor would it handle the hyperbola tops from point scatterers. Nor does it survive statics.

ANOTHER EXAMPLE

Years ago Sam Allen had a seismic recording company that could manage 600 channels of sign bit data. He had it strung out on a road about 200 miles from the Nevada nuclear test site. The military shot off some kilotons of explosive. Data shows the first 10 sec of the arrival. Essentially, we see a plane wave arriving. Fortuitously the recording line was perpendicular to the direction to the source (except for a bend in the road about 800m from the west end). The thing to notice is patterns of slow events, roughly 1.0 km/sec. And, the patterns have some resonance. (Sorry, I have very little densely sampled land data.)

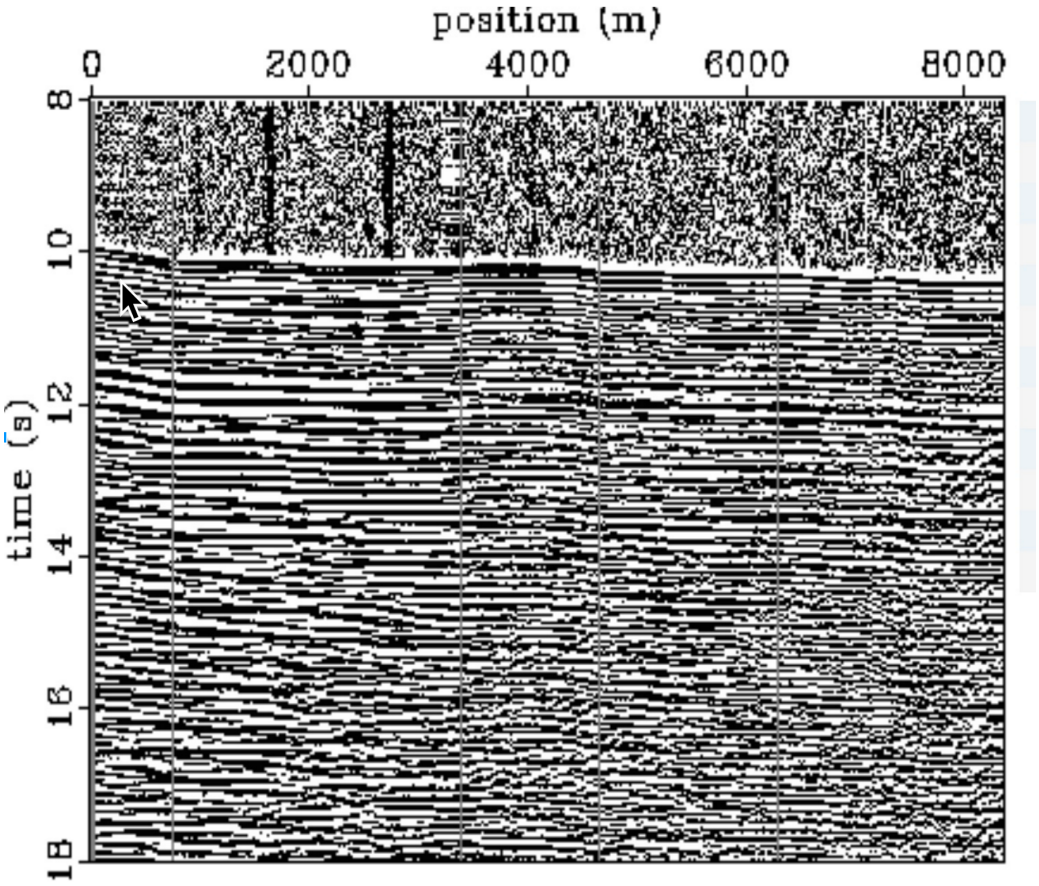


Figure 4: Multiple tons of old military explosives were detonated at the Nevada nuclear test site. Here see first arrivals at a couple hundred kilometers distance. Observe the "plane wave" is soon followed by cross-hatching slow waves, roughly 1 km/sec. [**NR**] jon2/. sjoerd

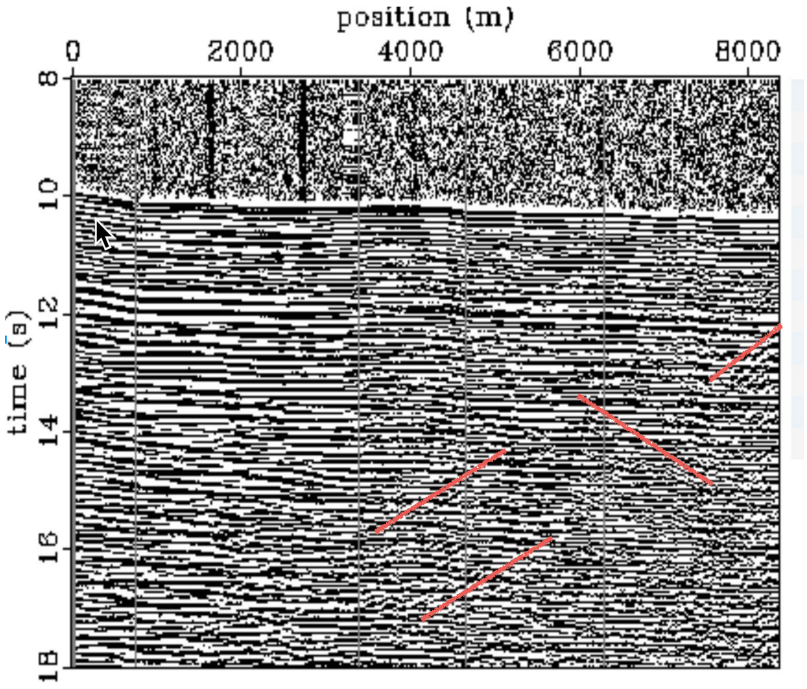


Figure 5: Multiple tons of old military explosives were detonated at the Nevada nuclear test site. Here see first arrivals at a couple hundred kilometers distance. Observe the "plane wave" is soon followed by cross-hatching slow waves, roughly 1 km/sec. [**NR**] jon2/. sjoerdAnnot

Source signature deconvolution of marine seismic data using deterministic modeling of the bubble signature

Leighton M. Watson, Joseph Jennings, and Shuki Ronen

ABSTRACT

Seismic airguns are not impulsive sources and hence marine seismic data must be designatured before interpretation. We demonstrate how deterministic modeling of the airgun signature can be used to designature field data. A heuristic approach is used to generalize the single airgun model to describe the signature of a small array of closely spaced guns where the bubbles from the different guns coalesce. The simulated signatures are used to designature data from a near-surface survey around a production rig. The results compare favorably to prediction-error filtering and the importance of using the correct source signature is illustrated.

INTRODUCTION

The data recorded in a seismic survey are the convolution of the source signature, wave propagation and the reflectivity. Estimating the reflectivity, or equivalently the velocity and density structure of the subsurface, is important for geological interpretation. In this paper, we discuss how marine seismic data can be designatured using deterministic modeling of the airgun signature.

Seismic airguns are the main source used in marine seismic surveys. An airgun functions by discharging highly pressurized air into the water, forming a bubble that expands and contracts. When the airgun discharges it produces a source signature with an initial peak due to the rapid expansion of the bubble and subsequent oscillations associated with the dynamics of the bubble expansion and collapse (Figure 1). Tuned airgun arrays have been used to minimize the bubble oscillations and make signature more impulsive (e.g., Dragoset, 2000), thereby simplifying the problem of source deconvolution. However, the signature of the airgun array can often still be seen in the data, especially for smaller arrays which cannot be tuned as well.

There are several methods for source designaturing. Statistical methods such as prediction-error filtering (PEF) are commonly used in the industry (Yilmaz, 2011). PEF attempts to solve the blind-deconvolution in which the spectrum of the signature is estimated and the deconvolution of the source signature are performed at the same time. Alternatively, the signature can be measured directly from near-field hydrophones and deconvolved from the data (Ziolkowski et al., 1981, 1984; Ziolkowski, 1991; Landrø et al., 1991; Landrø and Sollie, 1992; Laws et al., 1998; Kryvohuz and Campman, 2016) although there are some issues with the reliability of the near-field hydrophones and propagation of high frequency noise from the near-field to the far-field. Finally, physics-based modeling of the airgun-bubble dynamics can be used to estimate the source signature (Ziolkowski, 1970; Li et al., 2010; de Graaf et al., 2014; Watson et al., 2016). Commercial software packages like *Nucleus*®and

Gundalf®use a combination of numerical modeling and measured signatures of individual airguns to determine the signature for an array of airguns.

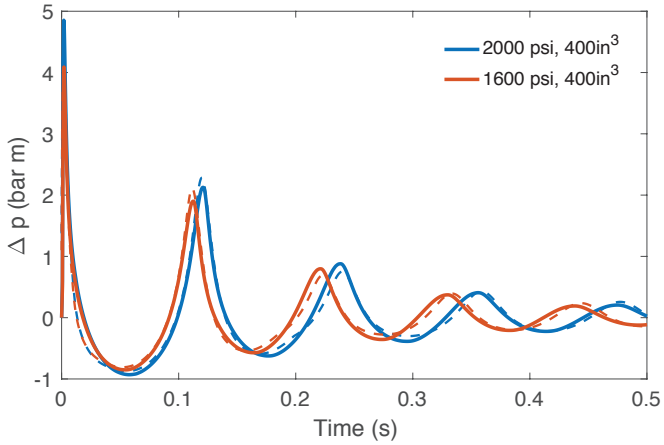
Here, we examine how physics-based modeling of the airgun dynamics can be used to estimate the source signature. We briefly discuss the numerical model and show that the simulated signature is comparable to that predicted by *Nucleus* before demonstrating how the simulated source signature can be deconvolved from the observed data in order to improve the quality of the observations.

MODEL

We model the airgun dynamics using a similar treatment to that outlined in the seminal paper by Ziolkowski (1970). We solve the compressible Euler equations in the water and evaluate the solution on the bubble wall to obtain the modified Herring equation (Herring, 1941; Cole, 1948; Vokurka, 1986),

$$R\ddot{R} + \frac{3}{2}\dot{R}^2 = \frac{p - p_\infty}{\rho_\infty} + \frac{R}{\rho_\infty c_\infty}\dot{p} - \alpha\dot{R}, \tag{1}$$

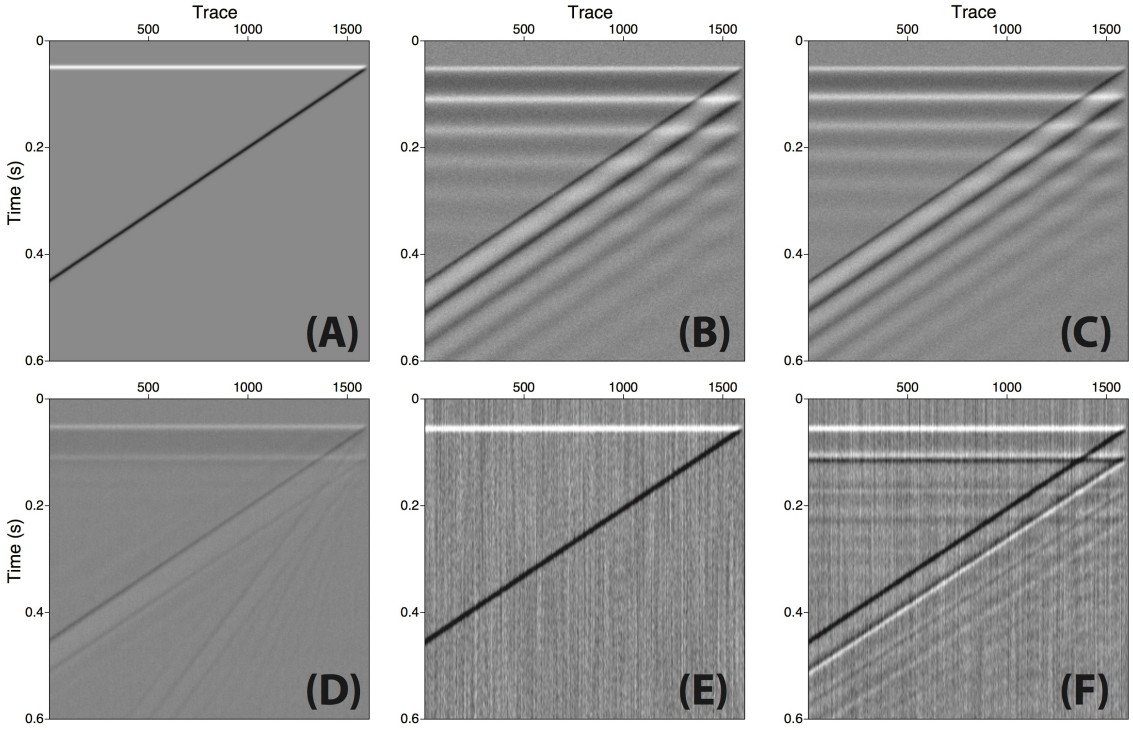
where R and $\dot{R} = dR/dt$ are the radius and velocity of the bubble wall, respectively, p is the pressure inside the bubble, and p_∞, ρ_∞ and c_∞ are the pressure, density, and speed of sound, respectively, in the water infinitely far from the bubble. Numerical models of airgun signatures typically under-predict the damping observed in the data. Hence, α is included as a damping constant (Langhammer and Landrø, 1996) to obtain a better fit between simulation and observations.



(a)

Figure 1: Our simulated signatures (solid) are in good agreement with the *Nucleus* signatures (dashed). [**CR**] leighton2/. compareNucleus

The pressure perturbation in the water is related to the bubble dynamics by (Keller and Kolodner, 1956)



(a)

Figure 2: (A) Reflectivity model. (B) Synthetic data generated by convolving the reflectivity model with the simulated signature of the 2000 psi airgun shown in Figure 1.(C) Synthetic data generated by convolving the reflectivity model with the simulated signature of the 1600 psi airgun shown in Figure 1. (D) Prediction-error filter applied to the synthetic data shown in (B). (E) Synthetic data shown in (B) deconvolved with the correct signature (2000 psi airgun). (F) Synthetic data shown in (C) deconvolved with the wrong signature (2000 psi airgun). [**ER**] leighton2/. synthetics

$$\Delta p(r,t) = \rho_\infty \left[\frac{\ddot{V}(t - r/c_\infty)}{4\pi r} - \frac{\dot{V}(t - r/c_\infty)^2}{32\pi^2 r^4} \right], \tag{2}$$

where Δp is the pressure perturbation in the water, r is the distance from the center of the bubble and $V = 4/3\pi R^3$ is the volume of the bubble. For further details on the numerical model the reader is referred to Watson et al. (2016) and the references therein.

SYNTHETICS

We illustrate the importance of using the correct signature in the designature process with a synthetic model of the subsurface consisting of a wedge of high acoustic impedance within a background medium with a lower acoustic impedance. There is a flat negative reflector at the top of the wedge and a dipping positive reflector at the base (Figure 2A).

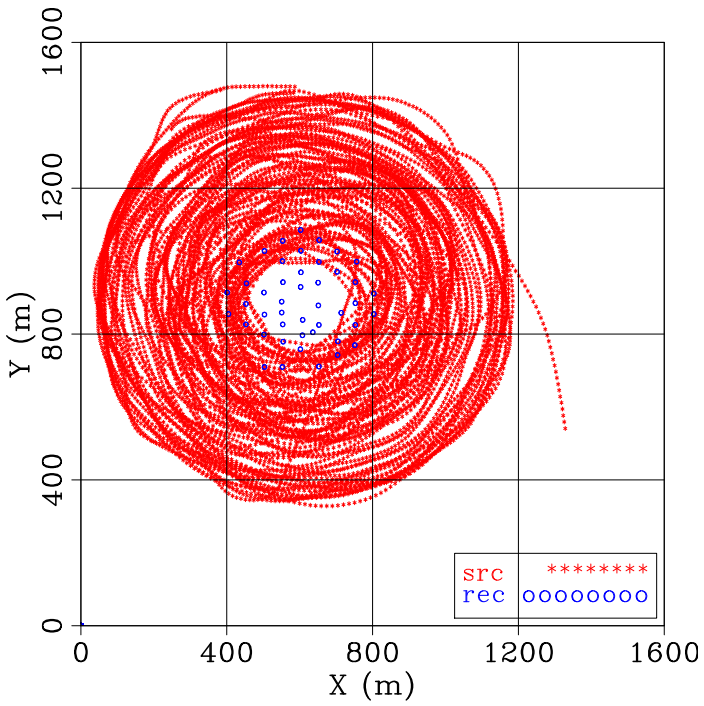
Synthetic data are generated by convolving the reflectivity model with the two simulated signatures shown in Figure 1: a 400 in^3 airgun pressurized to 2000 psi (Figure 2B) and a 400 in^3 airgun that has developed an air leak and is pressurized to 1600 psi (Figure 2C). Random noise is added to give a signal-to-noise ratio of 2.

Figure 2E shows the reflectivity model produced when the synthetic data are deconvolved with the correct source signature. The result, as expected, is, apart from the addition of random noise, the same as the initial reflectivity model. Figure 2F shows the reflectivity model produced when the synthetic data are deconvolved with the wrong source signature. Here synthetic data were generated with the 1600 psi airgun signature but the deconvolution was performed with the 2000 psi airgun signature. The reflectivity model still contains significant ringing. Figure 2F illustrates the challenges faced in designaturing field data where the source signature may be unknown or may differ from what are expected. This underscores the importance of deconvolving with the correct source signature as well as the significance of having a source signature estimate when working with real data. This can be obtained by modeling or from direct measurements, such as with a near-field hydrophone.

Prediction-error filters (PFEs) are used to remove predictable information from a dataset. Figure 2D shows the application of a PEF to the synthetic data shown in Figure 2B. The PEF removes some of the ringing but introduces additional artificial noise. This is because the synthetic data involves two reflectors with different dips which introduces non-stationarity. As we estimate a single prediction-error filter for each trace, we cannot properly account for this non-stationarity which prevents the prediction-error filter from fully predicting all of the ringing and leads to additional artifacts.

DATA

The Apache Forties oil field data were acquired as part of an ocean-bottom node survey around three production rigs in the North Sea. The goal of the survey was to image gas clouds and other shallow drilling hazards near the rigs (Alves, 2015; Jennings and Ronen, 2016). We focus on the data collected around the Delta platform. The survey consisted of 14485 shots and 48 nodes. The source vessel circled the platform and spiraled inwards. Receivers were arranged in a roughly hexagonal pattern around the platform with a 50 m node spacing (Alves, 2015). The parabolic reflectors seen in the data in the time domain



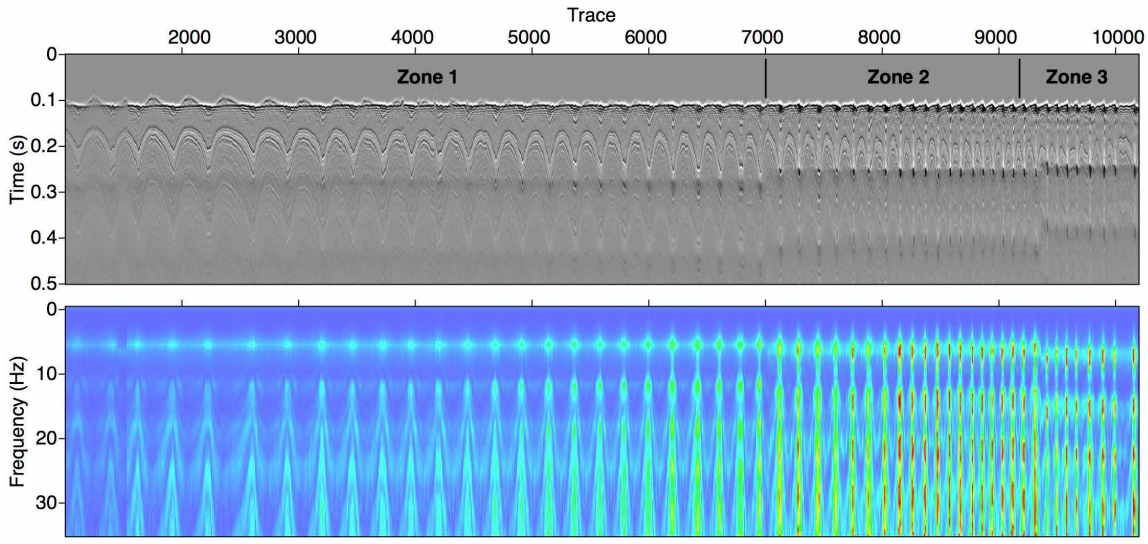
(a)

Figure 3: Shot-receiver geometry for the Apache Forties survey around the Delta platform. [**ER**] leighton2/. acquisitionGeometry

and the clear vertical bands of increased energy in the frequency domain are a result of the spiral acquisition geometry. The horizontal banding is a result of the non-impulsive source signature.

The seismic source was a cluster of three 250 in^3 airguns. There were issues with the airgun array during acquisition. A leak developed around trace 7000 and the array operated with reduced pressure to at least one of the guns. Eventually, one gun failed, approximately at trace 9300 reducing the array to two 250 in^3 airguns. The changing source signature is clearly visible in the data shown in Figure 4. Once the air leak develops the dominant frequency of the source shifts to higher frequencies because less mass is ejected from the array into the water. This means that the bubble generated is smaller and hence oscillates faster. For analysis, we divide the data into three zones as shown in Figure 4.

The source signature is simulated in each of the three zones. A heuristic approach is taken to estimate the source signature for the array of airguns from the single airgun model described above. Initially the airguns do not interact. Therefore, the first instances of the signal are the summation of the pressure perturbations from the multiple airguns. Based on the detailed calculations of Barker and Landrø (2014), we introduce a scaling factor were the amplitude of the summation is reduced by 10%. By the time of the first bubble peak the bubbles from the multiple airguns have coalesced (the distance between the airguns is 1 m) and subsequently behave as a single large bubble. This can be simulated as a single airgun with a volume equal to the total volume of the array. The signal is tapered between these two limits to give a smooth source signature. Note that this approximation can be justified



(a)

Figure 4: Data in the time (top) and frequency (bottom) domain. Note the discontinuities around trace 7000 and trace 9300 due to the airgun failure. [**ER**] leighton2/. displayData

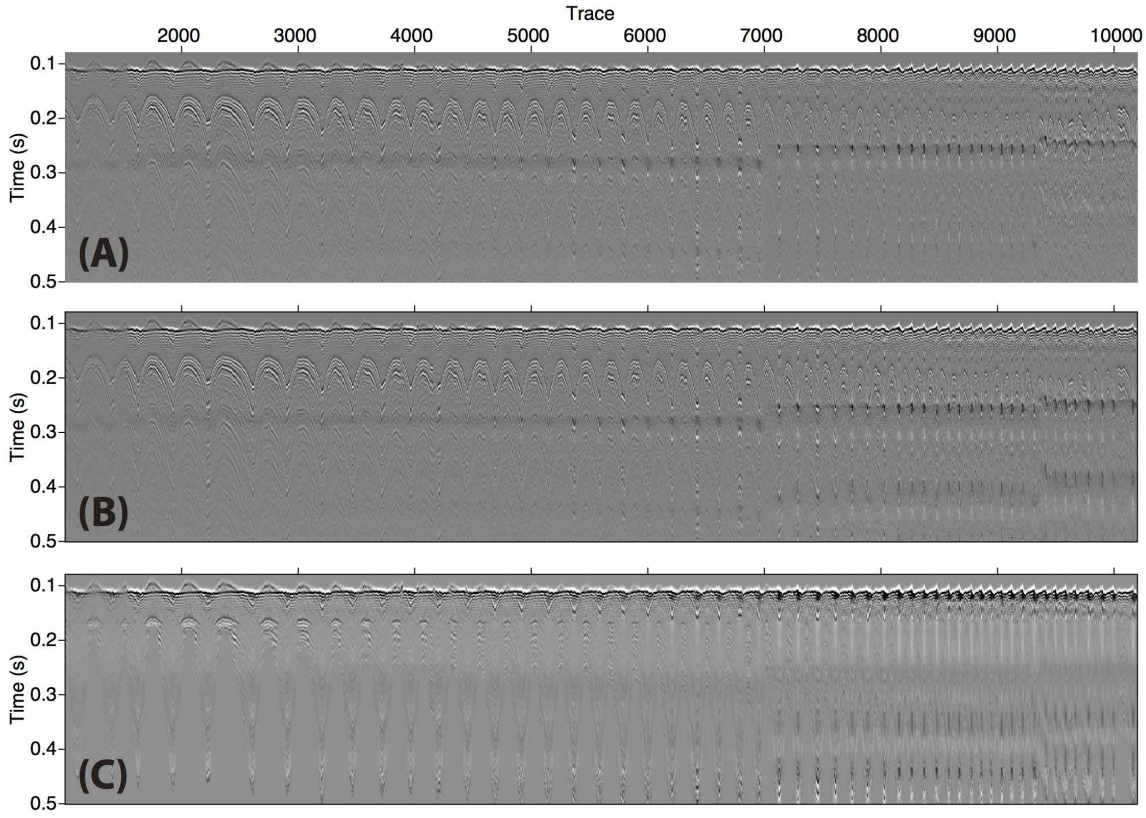
for a small array of closely spaced airguns. This approach is not valid for larger arrays where directivity effects become important and bubble coalescence is more complicated.

Figure 5 shows the deconvolution of the data using the simulated signatures (5A and 5B) and using a prediction-error filter (5C). Figure 5A is designatured using a different simulated signature for each of the three zones. Figure 5B is designatured using the simulated signature from zone 1. The results are the same for zone 1. For zones 2 and 3 bubble energy remains in Figure 5B that is not present in Figure 5A. As with the synthetics, this demonstrates the importance of deconvolving with the correct source signature.

Note that for both Figure 5A and Figure 5B a dark band of energy around 0.25 s remains after designaturing. The energy increasing to the right in each zone. This may be due to the source signature changing within each zone or may be an artifact of the spiral acquisition geometry.

CONCLUSION

We show how deterministic modeling of the airgun signature can be used to designature ocean bottom node data. We show that using an incorrect signature, such as not accounting for a pressure drop due to an air leak, may do more harm than good. We also show that using nondeterministic prediction error filter may create spurious events.



(a)

Figure 5: (A) Data deconvolved using the simulated source signature. The source signature is simulated for each of the three zones. (B) Data deconvolved using the simulated source signature from zone 1 to deconvolve from all three zones. (C) Data deconvolved using a prediction-error filter. The prediction error filter is applied zone by zone. [**ER**] leighton2/. deconvolution

ACKNOWLEDGEMENTS

We thank Apache North Sea Limited for access to the data set and permission to publish. We acknowledge the Stanford Exploration Project and their sponsors for financial support.

REFERENCES

Alves, G., 2015, Overview of the Apache Forties data set: SEP-Report, **160**.

Barker, D. and M. Landrø, 2014, An alternative method for modeling close-range interactions between air guns: Geophysics, **79**, P1–P7.

Cole, R. H., 1948, Underwater explosions: Princeton University Press.

de Graaf, K. L., I. Penesis, and P. A. Brandner, 2014, Modelling of seismic airgun bubble dynamics and pressure field using the Gilmore equation with additional damping factors: Ocean Engineering, **76**, 32–39.

Dragoset, B., 2000, Introduction to air guns and air-gun arrays: The Leading Edge, **19**, 892.

Herring, C., 1941, Theory of the pulsations of the gas bubble produced by an underwater explosions: Technical report, Office of Scientific Research and Development.

Jennings, J. and S. Ronen, 2016, Separation of simultaneous source blended data using radiality and source similarity attributes: SEP-Report, **163**.

Keller, J. B. and I. I. Kolodner, 1956, Damping of underwater explosion bubble oscillations: Journal of Applied Physics, **27**, 1152–1161.

Kryvohuz, M. and X. Campman, 2016, Optimization of sea surface reflection coefficient and source geometry in conventional dual source flip/flop marine seismic acquisition: SEG Annual Meeting, Dallas, Texas, 188–192.

Landrø, M. and R. Sollie, 1992, Source signature determination by inversion: Geophysics, **57**, 1633–1640.

Landrø, M., S. Strandenes, and S. Vaage, 1991, Use of near-field measurements to compute far-field marine source signatures: evaluation of the method: First Break, **9**, 375–385.

Langhammer, J. and M. Landrø, 1996, High-speed photography of the bubble generated by an airgun: Geophysical Prospecting, 153–173.

Laws, R., M. Landrø, and L. Amundsen, 1998, An experimental comparison of three direct methods of marine source signature estimation: Geophysical Prospecting, **46**, 353–389.

Li, G. F., M. Q. Cao, H. L. Chen, and C. Z. Ni, 2010, Modeling air gun signatures in marine seismic exploration considering multiple physical factors: Applied Geophysics, **7**, 158–165.

Vokurka, K., 1986, Comparison of Rayleigh's , Herring's , and Gilmore's Models of Gas Bubbles: Acta Acustica united with Acustica, **59**, 214–219.

Watson, L. M., E. M. Dunham, and S. Ronen, 2016, Numerical modeling of seismic airguns and low-pressure sources: SEG Annual Meeting, Dallas, Texas, 219–224.

Yilmaz, Ö., 2011, Seismic data analysis: Processing, inversion, and interpretation of seismic data: Society of Exploration Geophysicists.

Ziolkowski, A., 1970, A Method for Calculating the Output Pressure Waveform from an Air Gun: Geophysical Journal International, **21**, 137–161.

———, 1991, Why dont we measure seismic signatures?: Geophysics, **56**, 190–201.

Ziolkowski, A. M., G. E. Parkes, L. Hatton, and T. Haugland, 1981, British patent appl. no. 8 116 527.

———, 1984, U.S. patent no. 4,476,553.

Linearized waveform inversion of multicomponent blended data with polarization filters

Joseph Jennings, Biondo Biondi, Robert G. Clapp, and Shuki Ronen

ABSTRACT

We develop a new inversion scheme for linearized waveform-inversion (LWI) of blended data. We introduce polarization filters, computed from the independently modeled multicomponent data, which filter the blended data at each iteration of the inversion. We compare an approach of polarization filtering based on the Singular Value Decomposition (SVD) to our own new approach in which we extend the application of prediction-error filtering (PEF) to multicomponent data. We show that estimating PEFs using all data components provides the best deblended data which should provide much better imaging during the LWI process.

INTRODUCTION

The most difficult challenge in the direct imaging of simultaneous-source blended data is the mitigation of the artifacts that appear in the image due to cross-terms (Jiang et al., 2010). These artifacts are well understood especially in the context of LWI (least-squares migration) of phase-encoded data. While imaging of blended data is in fact quite similar to imaging of phase-encoded data, the two problems should be treated differently. This is due to the fact that with blended field data, we have a fixed encoding which cannot be changed from iteration to iteration. Changing the encoding function at each iteration, also known as dynamic phase-encoding, proves to be very useful in LWI of phase-encoded data (Boonyasiriwat et al., 2010).

While more work has been done on mitigating artifacts in phase-encoded imaging, there has been progress towards removing artifacts in imaging of blended data. In some of the earlier work, Tang et al. (2009) mitigated the artifacts with a horizontal Laplacian by adding a regularization term to the LWI objective function. Dai et al. (2011) proposed to use a deblurring filter in order to precondition the inversion and therefore achieving faster convergence. Lastly and also in the direction of regularizing the inversion, Xue et al. (2014) approached the problem by adding a shaping regularization term to the objective function.

Like the aforementioned work, we also propose to use a filter in order to mitigate the cross-term artifacts. But unlike the other approaches and following the work of Jennings and Ronen (2016), we compute our filter from the horizontal components of multicomponent data (e.g. ocean-bottom node). Because the horizontal components have a directional dependence (as opposed to hydrophones which only measure pressure, a scalar quantity), we can know *a priori* the directionality of the source. We can estimate filters to separate the blended data based on the direction. These directional filters are known as polarization filters that are commonly used in multicomponent signal and noise separation. At each iteration of the LWI process, we estimate these polarization filters on the independently

modeled shots, which serve as proxies to the deblended data. We then propose to apply these polarization filters to the blended data each each iteration of the LWI.

In this report, we first describe a new objective function that incorporates polarization filters into the LWI algorithm. We then compare two multicomponent filters, one from SVD and the other based on the theory of multidimensional non-stationary linear prediction (prediction-error filters or PEFs). We show that estimating the PEF on all components at the same time leads to better deblending results.

THEORY

Linearized blended modeling and data

For the purposes of simplified analysis, we represent a single shot gather observed in recorded seismic data by the following linear operation:

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

where $\mathbf{L}_i = \mathbf{L}(\mathbf{b}, \mathbf{x}_i)$ is the linearized wave-equation (Born) modeling operator for a specific background velocity model $\mathbf{b}$ and a source location $\mathbf{x}_i$. As $\mathbf{L}$ is the linearized modeling operator, then the model vector $\mathbf{m}$ represents the reflectivity and $\mathbf{d}_i$ represents the ith shot gather of a seismic survey. With this representation of a single shot gather, we can then form a column vector of linear operators $\mathbf{L}_i$ to represent n shots and recorded data in a single seismic survey as:

$$\begin{bmatrix} \mathbf{L}_1 \\ \mathbf{L}_2 \\ \vdots \\ \mathbf{L}_n \end{bmatrix} \mathbf{m} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_n \end{bmatrix},$$

$$\mathbf{L}\mathbf{m} = \mathbf{d}. \tag{2}$$

In order to express linearized blended modeling, we introduce the blending operator $\mathbf{\Gamma}$. This operator is linear and its primary function is to describe the blending of the data performed during the acquisition. In matrix form, it can be written as a rectangular matrix with many more columns than rows ($n > m$) where the number of rows (m) represents the extent of blending and the number of columns (n) is determined by the number of shots. A generic form of $\mathbf{\Gamma}$ can be written as:

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{G} & \mathbf{G} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{G} & \mathbf{G} & \mathbf{G} & \dots & \mathbf{0} & \mathbf{0} & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \dots & \dots & \dots & \dots & \mathbf{G} & \mathbf{G} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \mathbf{G} & \mathbf{G} \end{bmatrix}_{m \times n},$$

where $\mathbf{G}$ is a phase-shift operator. For ease of notation, we use the same phase-shift operator ($\mathbf{G}$) for all shots but in general, the phase-shifts will vary from shot to shot. Applying $\mathbf{\Gamma}$ to equation 2 we arrive at an expression for modeling blended data:

$$\begin{aligned} \mathbf{\Gamma}\mathbf{L}\mathbf{m} &= \mathbf{\Gamma}\mathbf{d}, \\ \widetilde{\mathbf{L}}\mathbf{m} &= \widetilde{\mathbf{d}}, \end{aligned} \tag{3}$$

where $\widetilde{\mathbf{L}}$ and $\widetilde{\mathbf{d}}$ are the linearized blended modeling operator and data respectively.

LWI of blended data

With our previously-defined blended modeling operator and data, we can write the objective function for LWI of blended data as:

$$J(\mathbf{m}) = \frac{1}{2}||\widetilde{\mathbf{L}}\mathbf{m} - \widetilde{\mathbf{d}}||_2^2. \tag{4}$$

Differentiating with respect to the model parameters we can express the gradient of equation 4 as:

$$\nabla J(\mathbf{m}) = \underbrace{\sum_{i=1}^{n} \mathbf{L}_i^* \mathbf{r}_i}_{\text{conventional}} + \underbrace{\sum_{j \neq i}^{n} \sum_{i=1}^{n} \mathbf{L}_j^* \mathbf{r}_i}_{\text{cross-terms}}, \tag{5}$$

where $\mathbf{r}_i$ is the data residual for the ith shot gather. As expected, at each iteration, our gradients will be contaminated with artifacts, seriously affecting the convergence of our LWI. These artifacts, also known as cross-talk, are due to the cross correlation of source-side Green's functions between different sources and the relative time delay between sources (Jiang et al., 2010). To suppress these artifacts, we propose to apply polarization filters estimated from the independently modeled data and applied to the to multicomponent blended data $\widetilde{\mathbf{d}}$ at each iteration resulting in an approximate deblended data. This process will be performed at each iteration resulting in iterative imaging with minimal cross-talk during the LWI process.

LWI of blended multicomponent data with polarization filters

As sources will have different azimuths in simultaneous source surveys, the recorded multicomponent data will be polarized differently providing information on the directionality of the source. Of course we can only be certain of the polarization of the direct arrival and not of the reflections and other recorded events. To ensure that we can accurately estimate the polarization for each source, we estimate these polarization filters on independently modeled data. The independently modeled shots can be thought of as proxy data on which we estimate these polarization filters. To incorporate the use of multicomponent data and polarization filters into LWI, we propose the following objective function:

$$J(\mathbf{m}) = \frac{1}{2}||\mathbf{L}\mathbf{m} - \mathbf{d_p}||_2^2, \tag{6}$$

where $\mathbf{L}$ models multicomponent data (e.g. linearized elastic wave-equation operator), and $\mathbf{d_p}$ is the multicomponent data, approximately deblended via polarization filtering. Note the absence of the tilde on the linearized modeling operator. This indicates that during the LWI process, each shot will be modeled independently allowing for the computation of polarization filters on proxy shots at each iteration. In terms of the blended data, we can write the computation of $\mathbf{d_p}$ as:

$$\mathbf{d_p} = \mathbf{P}(\tilde{\mathbf{d}}) \approx \mathbf{d} \tag{7}$$

where $\mathbf{P}$, is a filtering operation that is either linear or non-linear in the blended data and utilizes the polarization in order separate one shot from another. In the remainder of this report, we describe two possible choices for $\mathbf{P}$, one based on a more traditional polarization filtering approach that utilizes the Singular Value Decomposition (SVD) of the data matrix, and the other, a new method, in which we extend the estimation of non-stationary multidimensional prediction error-filters (PEFs) to multicomponent data.

COMPARISON OF FILTERING APPROACHES

An SVD-based approach

Theory

Filtering of multicomponent data via eigendecompositions was first developed in the earthquake seismology community in order to remove elliptically polarized events from the data (e.g. surface waves) as well as to obtain an estimate of the direction-of-arrival of the earthquake source (Flinn, 1965). A key step in each of these methods is the computation of what are known as the polarization vectors (Vidale, 1986) of the data which can then be used to provide an estimate of the polarization at each point in the dataset. While the computation of these polarization vectors has been extended into other domains that are more suited for capturing non-stationary information (Barak, 2017), we only describe a time-domain implementation.

To compute the polarization vectors of our data we first construct a $nt \times nc$ size matrix where nt is length of a time window to be placed on the data and nc is the number components we wish to analyze at a time. If we consider the case in which we only have two components (e.g. the horizontal geophone components), then we may write our data matrix $\mathbf{D}$ as:

$$\mathbf{D} = \begin{bmatrix} \mathbf{y} & \mathbf{x} \end{bmatrix},$$

where $\mathbf{y}$ and $\mathbf{x}$ are two column vectors of length nt that contain the data measured on the Y and X components respectively. If we then perform a reduced-SVD on this data matrix, we can write out the decomposed matrix as:

$$\mathbf{D} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \mathbf{U} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{bmatrix},$$

where $\mathbf{U}$ is a $nt \times 2$ matrix that contains the left singular vectors, σ_1 and σ_2 are the singular values and $\mathbf{v_1}$ and $\mathbf{v_2}$ are the right singular vectors, or our orthornormal polarization vectors. As in an SVD, the first singular value (σ_1) is defined to be the largest singular value, then $\mathbf{v_1}$, its associated polarization vector, will describe the dominant polarization on that data trace for the chosen time window. To develop a better understanding of why the right singular vectors of the SVD of the data matrix give an estimate of the polarization in the data, it is helpful to recall that the right, scaled singular vectors that result from an SVD of a 2×2 matrix will give the principal axes of an ellipse (Trefethen and Bau III, 1997). Notice also that in its simplest form, the hodogram of a multicomponent seismogram can be described by an ellipse and the largest principal axis of the hodogram ellipse describes the dominant polarization direction of the data. Therefore, the right singular vectors of a data matrix $\mathbf{D}$ will provide an estimate of the maximum direction of polarization in

the data. A visual example of this for two phase-shifted, monochromatic signals is shown in Figure 1. For a more rigorous explanation on the connection between polarization in multicomponent seismic data and the right singular vectors obtained via SVD we refer the reader to Montalbetti and Kanasewich (1970) and Jackson et al. (1991).

Application to filtering multicomponent blended data

As these polarization vectors give an estimate of the direction of the data, we can use them to design filters that will remove parts of the data that have an undesired direction. In the context of simultaneous-source acquisition, when sources have different azimuths their direct arrivals will be polarized differently and therefore can be suppressed with a polarization filter. In addition, depending on the geometry of the subsurface, we can also remove other events that originate from that source if they have different polarization than events coming from the other shot.

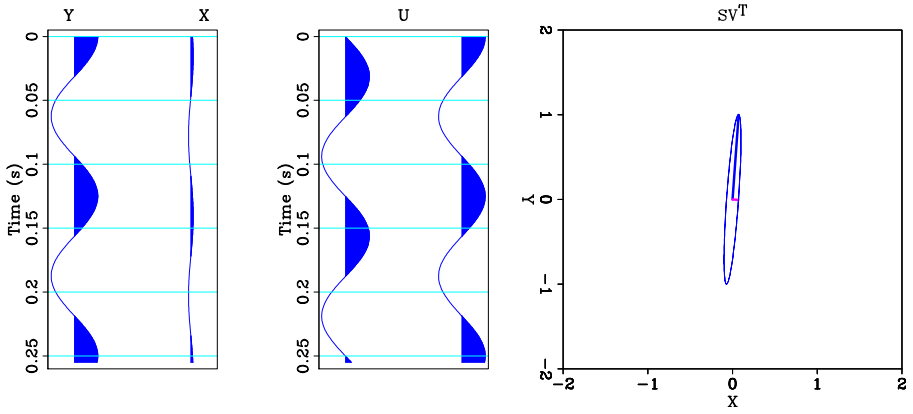
Figure 2 shows the acoustic wavefields and the recorded particle velocities (v_y, v_x) as a result of two shots. The blue line indicates the receiver locations. One shot was positioned cross-line to receiver positions (Figure 2(a)) and another directly inline with the receiver positions (Figure 2(b)). Because of these shot locations, the data are highly polarized in the Y and X directions as is evident by the shot gathers of the horizontal components (Figures 2(c)-2(f)). In another experiment, these two shots occurred with a much shorter shot interval resulting in blended data. The recorded blended particle velocities are shown in Figures 3(a)-3(b). Because these blended data are highly polarized, it provides a simple example in which we can use polarization filters to separate the two arrivals.

In order to design polarization filters for these blended data and blended data in general, let us first examine the SVD of the data matrices of the independently modeled shots $\mathbf{D}_1$ and $\mathbf{D}_2$ (Figure 2):

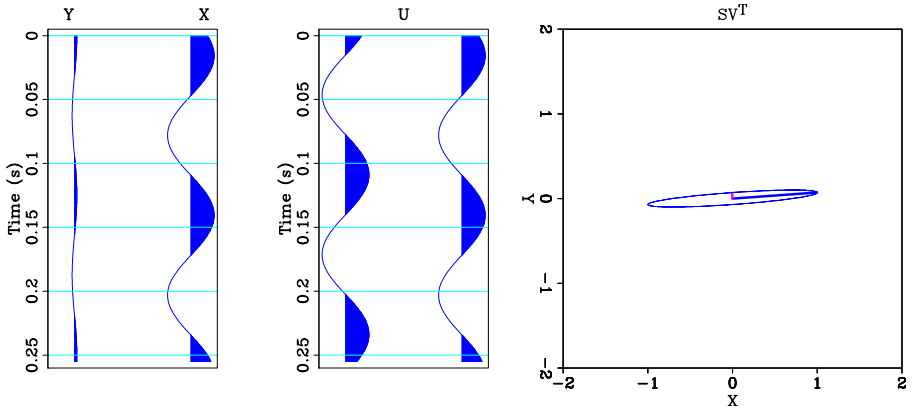
$$\mathbf{D_1} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{x}_1 \end{bmatrix} = \mathbf{U_1 \Sigma_1} \begin{bmatrix} \mathbf{v}_{11}^T \\ \mathbf{v}_{12}^T \end{bmatrix}, \mathbf{D_2} = \begin{bmatrix} \mathbf{y}_2 & \mathbf{x}_2 \end{bmatrix} = \mathbf{U_2 \Sigma_2} \begin{bmatrix} \mathbf{v}_{21}^T \\ \mathbf{v}_{22}^T \end{bmatrix},$$

where $\mathbf{v}_{ij}$ indicates the ith shot for the jth polarization vector. When the shots have different azimuths, then the polarization vectors of each shot $(\mathbf{v}_{11}, \mathbf{v}_{12})$ and $(\mathbf{v}_{21}, \mathbf{v}_{22})$ will point in different directions. In fact, when the two shots have orthogonal azimuths, then their polarization vectors will also be orthogonal. For the data shown in Figure 2, we can compute the polarization vectors for each and every time sample. If we then take both components of the largest polarization vector and plot them for each and very time sample of the gathers shown in Figure 2, we can visualize the strength of the Y and X polarization at each point in the data. Figure 4 shows result of carrying-out this process. It is clear that for the majority of the points in the gather, the polarization vectors are indeed orthogonal.

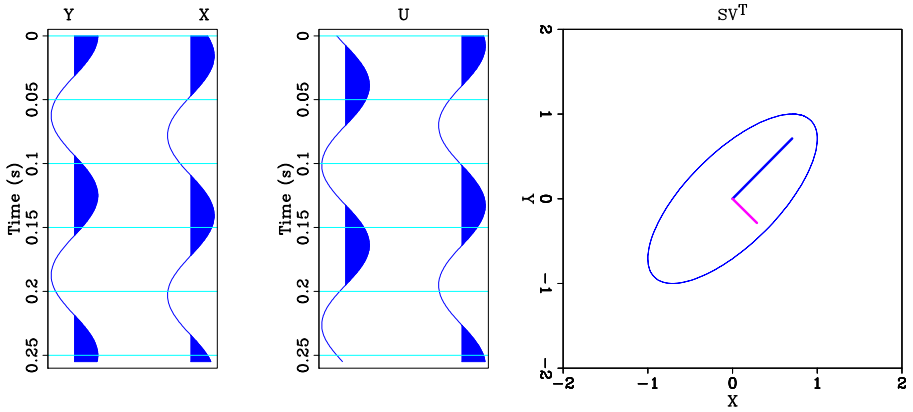
Using the fact that sources with different azimuths will lead to different polarization vectors, we can construct filters based on the inner-product of the independently modeled polarization vectors with the polarization vectors of the blended data. In places where the blended data are polarized in the same direction as one of the shots, the inner-product will be large and in places where they are nearly orthogonal, the inner-product will be small. From this, we can arrive at an expression for the polarization filter coefficients. First, we compute the



(a)



(b)



(c)

Figure 1: Three simple examples of polarization analysis via SVD. The left wiggle plot of each subpanel shows $\mathbf{D} = [\mathbf{y} \quad \mathbf{x}]$. The right wiggle plot in each panel shows the computed $\mathbf{U}$ of the SVD. The rightmost plot shows the scaled singular vectors $\mathbf{\Sigma V}^T$ overlain on the hodogram. Panel (a) shows an example in which the majority of the energy is on the Y component. This is evident by the hodogram and the right singular vector that point in the Y direction. Panel (b) shows now a similar situation but now the majority of the energy is located on the X component. Panel (c) shows the polarization analysis now of a multicomponent signal of mixed polarization. Note that the singular vectors are in fact the principal axes of the hodogram ellipse (as is defined by the SVD) thus providing the polarization information. [**ER**] joseph1/. svdelly,svdellx,svdellmix

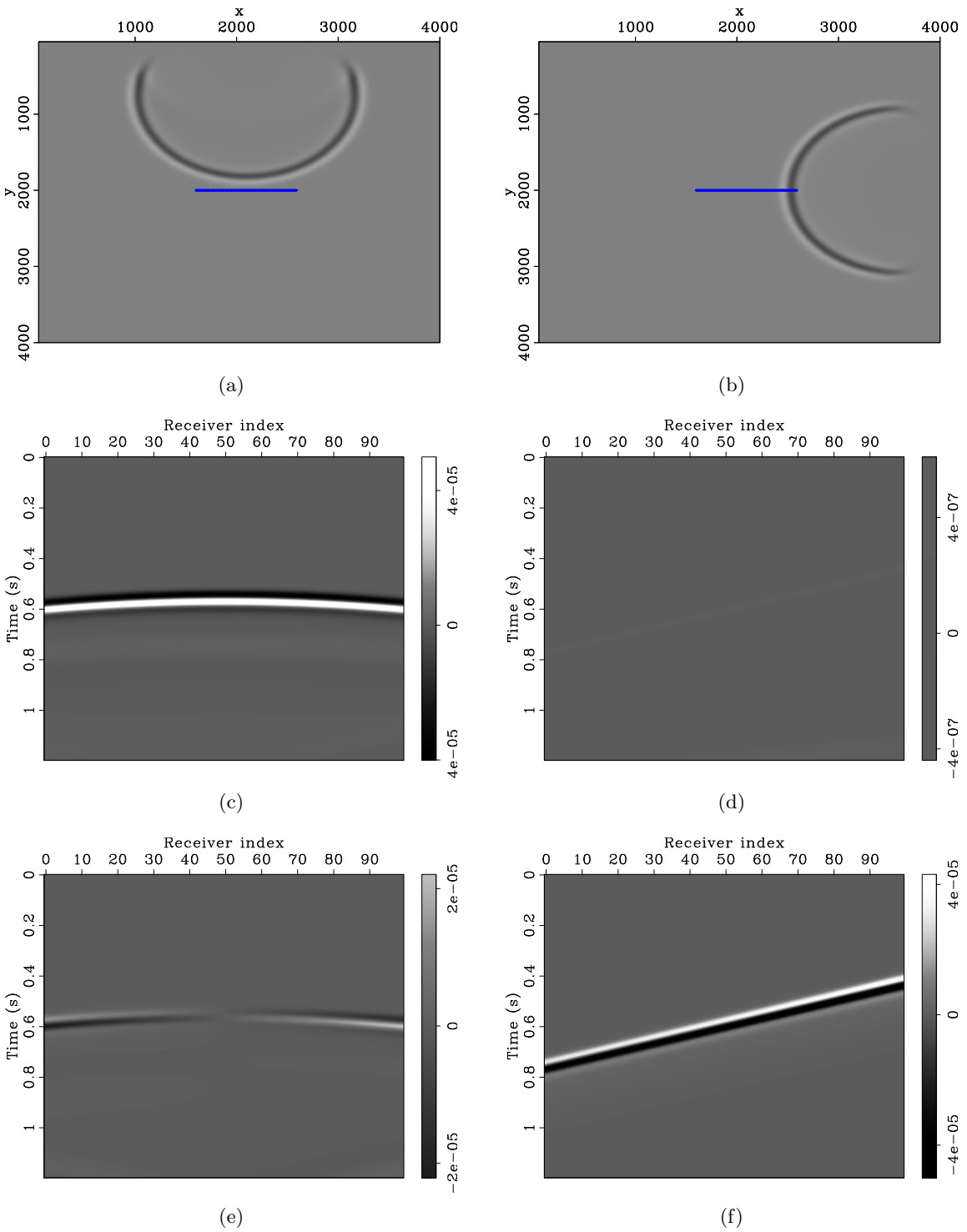


Figure 2: Two independently modeled shots positioned cross-line (panel (a)-shot one) and inline (panel(b)-shot two) to the line of the receivers (blue) resulting in two wavefields and highly polarized multicomponent acoustic data. (a) Pressure wavefield propagating in the X-Y (surface) plane as a result of shot one. (b) Pressure wavefield propagating along the surface as a result of shot two. (c) Recorded Y-component of the particle velocity field (v_y) from shot one. (d) Recorded v_y from shot two. (e) Recorded X-component of the particle velocity field (v_x) from shot one. (e) Recorded v_x from shot two. [**ER**] joseph1/. wfld0frame115,wfld1frame115,vyda0,vyda1,vxda0,vxda1

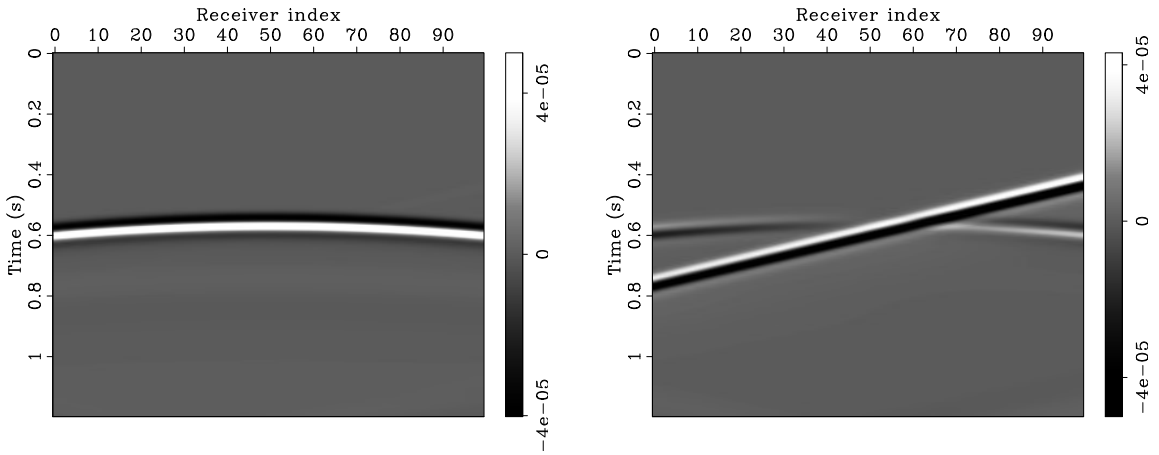
polarization vectors of the blended data:

$$\widetilde{\mathbf{D}} = \widetilde{\mathbf{U}}\widetilde{\mathbf{\Sigma}} \begin{bmatrix} \widetilde{\mathbf{v}_1}^T \\ \widetilde{\mathbf{v}_2}^T \end{bmatrix}.$$

Then computing the inner-products of the largest polarization vectors of the independently modeled data with the largest polarization vector of the blended data, we can construct a 2×1 vector of weights that can be applied to each component of the blended data for each time sample:

$$f_1 = \mathbf{v}_{11}^T \widetilde{\mathbf{v}_1}, \quad f_2 = \mathbf{v}_{21}^T \widetilde{\mathbf{v}_1},$$
$$\begin{bmatrix} v_{y_1} \\ v_{y_2} \end{bmatrix}_{it} = \begin{bmatrix} |f_1|^n \\ |f_2|^n \end{bmatrix}_{it} \widetilde{v_{y_{it}}},$$
$$\begin{bmatrix} v_{x_1} \\ v_{x_2} \end{bmatrix}_{it} = \begin{bmatrix} |f_1|^n \\ |f_2|^n \end{bmatrix}_{it} \widetilde{v_{x_{it}}},$$

where v_{y_1}, v_{y_2}, v_{x_1} and v_{x_2} are the deblended Y and X components of the particle velocity of shot one and shot two, f_1 and f_2 are the two weights and n is an integer to increase the strength of the filtering. We apply this filtering scheme to blended data shown in Figure



(a) (b)

Figure 3: Multicomponent blended data. (a) Y component of particle velocity. (b) X component of particle velocity. Note that the energy from shot two was not recorded on the Y component because it was completely polarized in the X direction (Figure 2(e)). [**ER**] joseph1/. bvyda,bvxda

3 and obtain the approximately deblended data shown in Figure 5.

Multicomponent PEFs

In addition to the SVD filter, we also investigated the use of non-stationary PEFs in order to approximately deblend the data. Because PEFs act as annihilation filters when the prediction-error is small, they are well-suited for signal and noise separation which include deblending applications (Spitz et al., 2008). Because we desire that the PEF use the

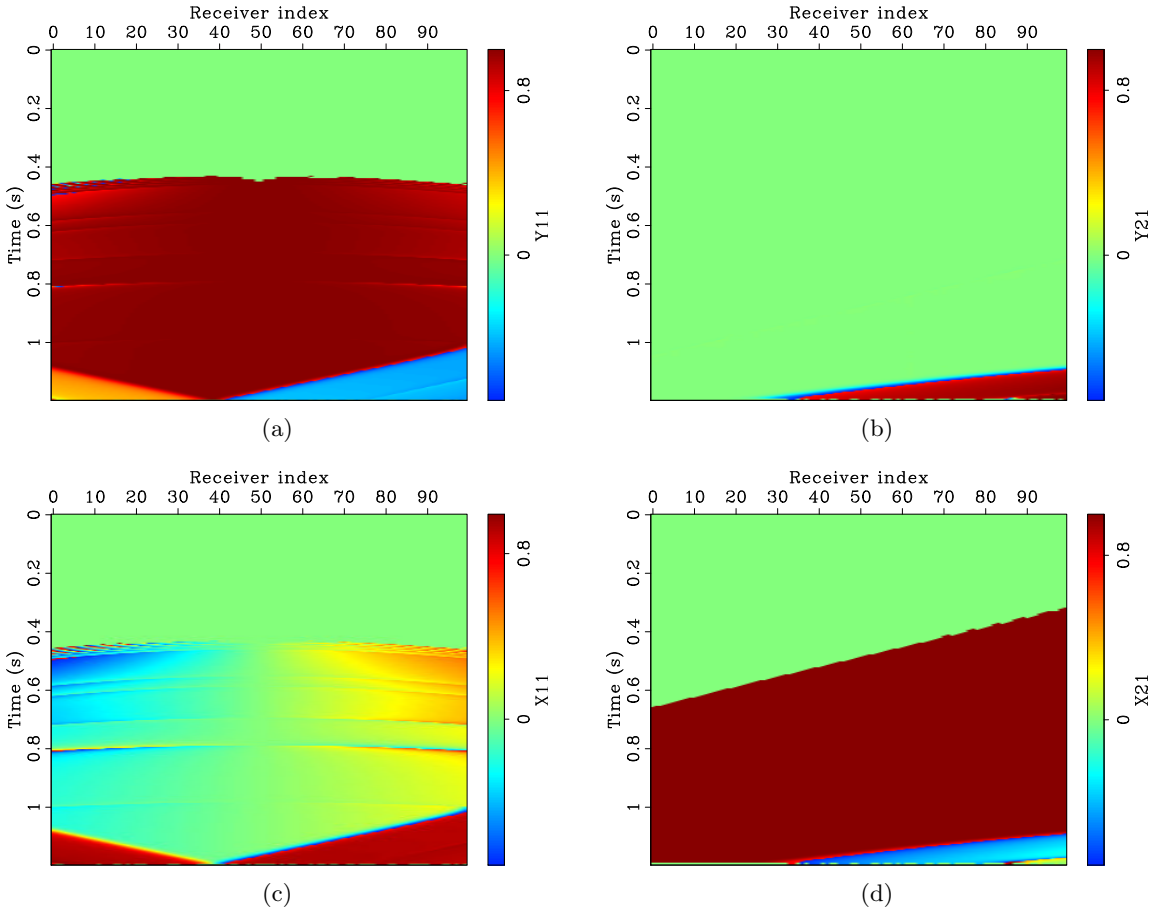


Figure 4: Y and X components of the largest polarization vector of shots one and two. (a) Y component of the strongest polarization vector of shot one. (b) Y component of the largest polarization vector of shot two. (c) X component of the largest polarization vector of shot one. (d) X component of the largest polarization vector of shot two. Comparing the magnitudes of each of the components of the vectors, it is clear that at almost every point in the data the vectors $\mathbf{v}_{11}$ and $\mathbf{v}_{21}$ are nearly orthogonal. [**ER**] joseph1/. y1,y2,x1,x2

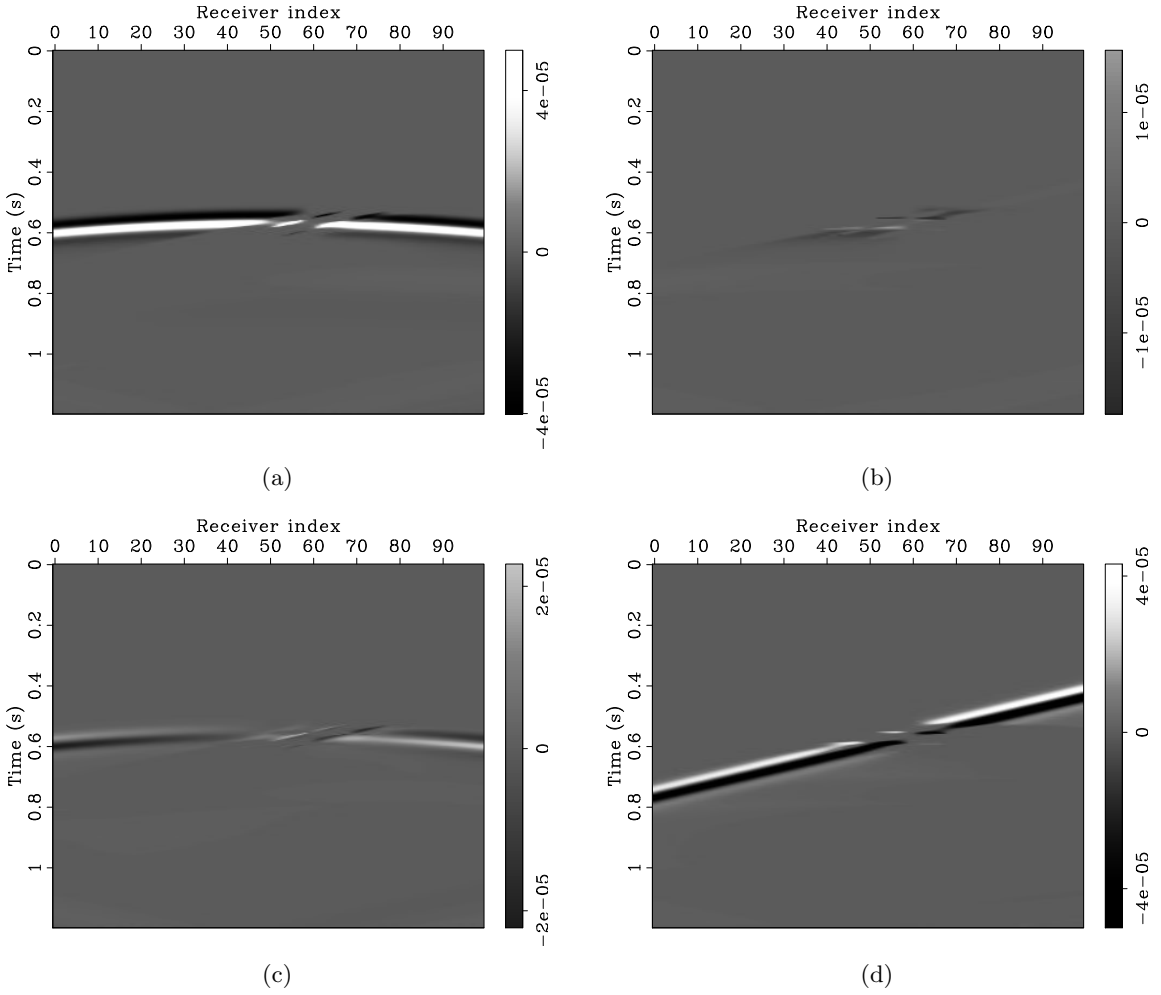


Figure 5: Approximately deblended data via the SVD-based polarization filtering. (a) Deblended Y component of shot one. (b) Deblended Y component of shot two. (c) Deblended X component of shot one. (d) Deblended X component of shot two. It appears that the filtering algorithm did a reasonable job at what it was designed to do: remove events that are not polarized in the direction of the independently modeled shots. Note that the method performs poorly when the two events overlap. [**ER**] joseph1/. vyfilt1,vyfilt2,vxfilt1,vxfilt2

polarization information to aid in the prediction, we estimate PEFs on all of the components at the same time. To verify our implementation of multicomponent PEF estimation, we compare the PEF estimated on all components simultaneously with PEFs estimated on the components separately. For a description of the algorithm we use for estimating non-stationary PEFs, we refer to Ruan et al. (2015).

Theory

Again, starting from the independently modeled data we can estimate non-stationary PEFs on the independently modeled shots, which serve as annihilators for each shot. To do this, we estimate the PEFs $\mathbf{f}_1$ and $\mathbf{f}_2$ from the following system of regressions:

$$\overline{\mathbf{D}}_1\mathbf{f}_1 \approx \mathbf{0}, \tag{8}$$

$$\overline{\mathbf{D}}_2\mathbf{f}_2 \approx \mathbf{0}, \tag{9}$$

where $\overline{\mathbf{D}}_1$ and $\overline{\mathbf{D}}_2$ are the data from shot gather one and two respectively formed into convolution matrices. Upon solving for these filters, we can then use them to separate shots one and two by minimizing the following objective function:

$$J(\mathbf{d}_1, \mathbf{d}_2) = \frac{1}{2}||\mathbf{d}_1 + \mathbf{d}_2 - \widetilde{\mathbf{d}}||_2^2 + \frac{1}{2}||\overline{\mathbf{F}}_1(\mathbf{d}_2 - \widetilde{\mathbf{d}})||_2^2 + \frac{1}{2}||\overline{\mathbf{F}}_2(\mathbf{d}_1 - \widetilde{\mathbf{d}})||_2^2, \tag{10}$$

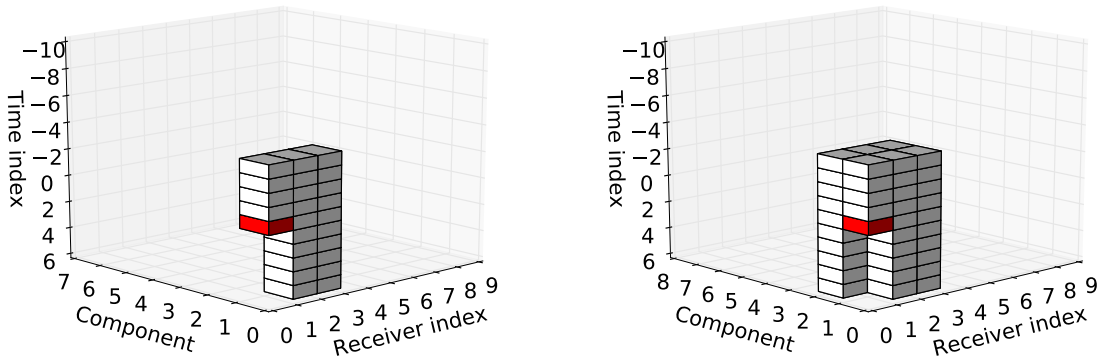
where $\mathbf{d}_1$ and $\mathbf{d}_2$ are the data from shots one and two respectively now written as column vectors and $\overline{\mathbf{F}}_1$ and $\overline{\mathbf{F}}_2$ are the PEFs estimated from expressions 8 and 9 now expressed as convolutional operators. The first term of this objective function is the familiar deblending objective that requires that the sum of the estimated shots be equal to the blended data. The other two terms act to enforce the separation of the two shots (Abma, 1995).

Single vs. multicomponent PEFs

Although the above algorithm may provide a manner in which to deblend the data at each iteration of the LWI process, we need to extend the current use of single component PEFs, to multicomponent data, in which the PEF is estimated simultaneously using all data components. In doing so, we expect that the PEF has the possibility of incorporating polarization information into its coefficients that could annihilate unwanted noise not only by a non-stationary temporal or spatial signature, but also based on the direction from which it originated. In doing so, these multicomponent PEFs become powerful tools not only for separating multicomponent blended data, but also for other multicomponent signal processing applications typical in land and ocean-bottom node settings. While all of the theory has yet to be developed, we currently extend the 2D single component PEF along the component axis. This then makes the 2D single component PEF now a 3D PEF estimated on 2D multicomponent data. A typical single component PEF is shown in Figure 6(a) and the new multicomponent PEF is shown in Figure 6(b).

Application to filtering multicomponent blended data

With the data shown in Figure 3 as our $\widetilde{\mathbf{d}}$, we minimized the objective function in equation 10 to compute the approximately deblended data. The results of minimizing expression 10



(a) (b)

Figure 6: (a) 2D Single component and (b) 2D multicomponent PEF design. The red sample indicates the location of the zero lag (where the prediction will be made). In order to use the additional direction information provided by the multicomponent data, we extend the current 2D single component PEF along the component axis, to create the multicomponent PEF. In addition to using previous temporal samples and the two spatial lags, the multicomponent PEF will use the spatial lags as well as previous and future samples from the other component to make predictions on the current component. [**ER**] joseph1/. pef2d,pef3d

using single component PEFs are shown in Figure 7 and the results using multicomponent PEFs are shown in Figure 8. Both inversion results completely recover the two shots on each component and therefore would provide a good $\mathbf{d}_p$ during the LWI process.

DISCUSSION

Single and multicomponent PEF comparison

While the single and multicomponent PEFs provide similar filtering results, we find differences when we examine the model updates at each iteration. Figure 9 shows the comparison of the model updates of the X component of shot one after 60 iterations of both the single and multicomponent PEF approaches. Observing the model residual plot shown in Figure 9(c), it is readily apparent that the deblending using the multicomponent PEF greatly outperforms the deblending with the single component PEF. Moreover, examining the deblending misfit at each iteration shown in Figure 9(d), we see that using the multicomponent PEF also provides an overall better fit to the data. While we do not yet understand all of the reasons for why this occurs, a simple explanation is that extending the PEF along the component axis allows the filter to better distinguish the signal and noise (Abma, 1995). This is because in addition to the temporal and spatial axes, we have the component axis. It is also possible that the additional information provided on the component axis that the PEF uses for prediction is in fact polarization information thus making these multicomponent PEFs polarization filters. We are currently in the process of running more numerical

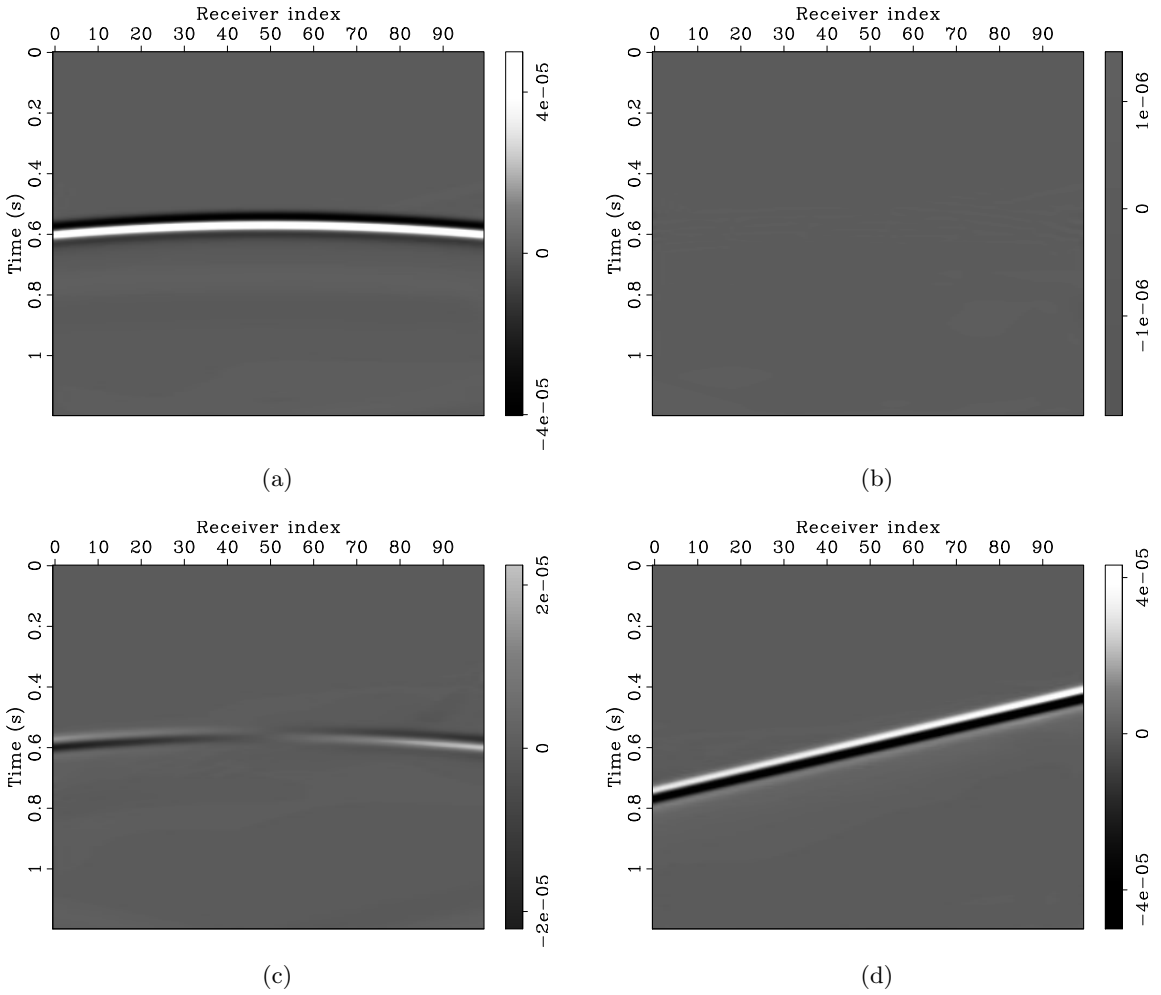


Figure 7: Approximately deblended data via minimizing equation 10 using single component PEFs (Figure 6(a)). (a) Deblended Y component of shot one. (b) Deblended Y component of shot two. (c) Deblended X component of shot one. (d) Deblended X component of shot two. It appears that the inversion completely recovered both shots on each of the components. [**ER**] joseph1/. sepsigvyda,sepnoizvyda,sepsigvxda,sepnoizvxda

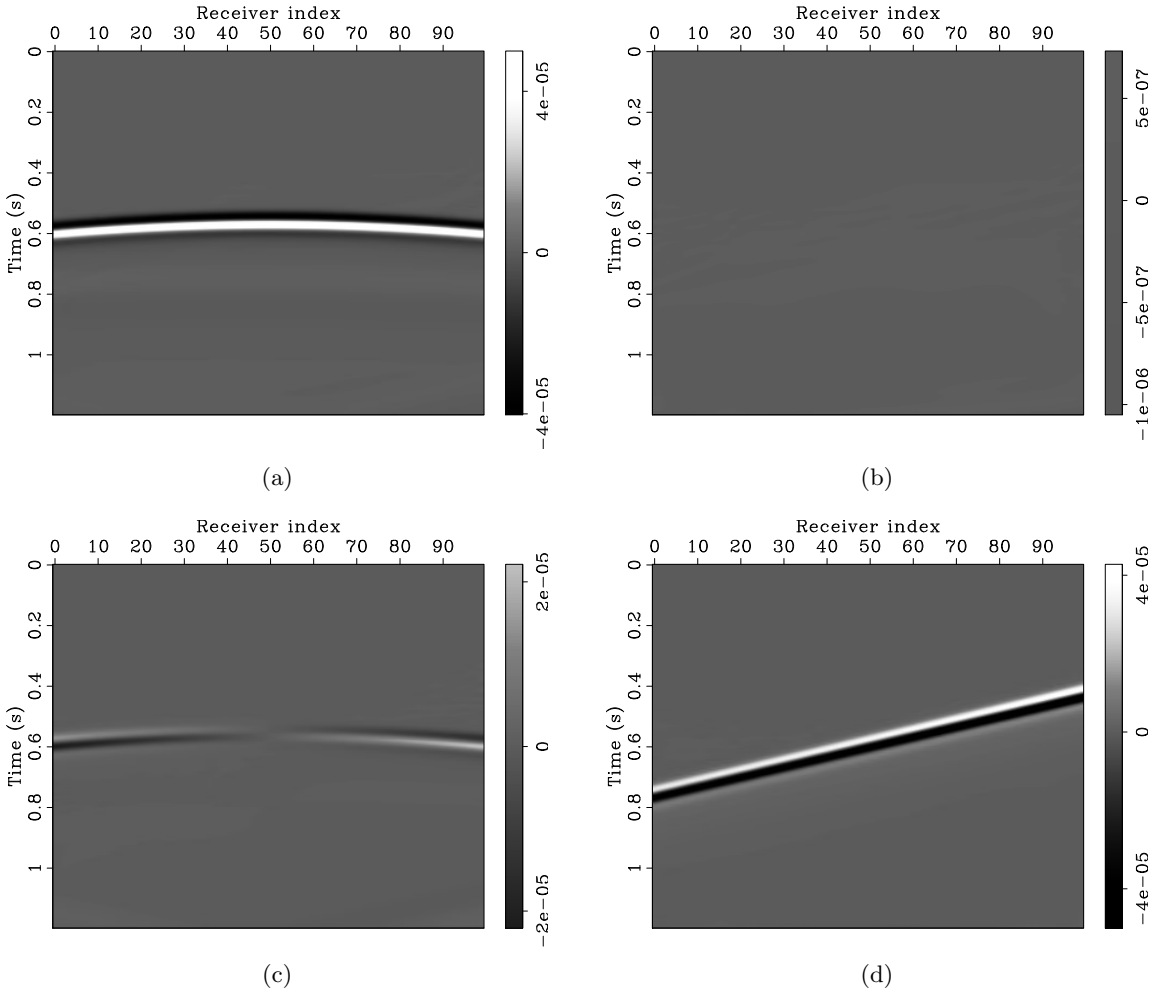


Figure 8: Approximately deblended data via minimizing equation 10 using multicomponent PEFs (Figure 6(b)). (a) Deblended Y component of shot one. (b) Deblended Y component of shot two. (c) Deblended X component of shot one. (d) Deblended X component of shot two. The results are quite similar to the single component results and completely recovered both shots on each of the components. [**ER**] joseph1/. sepsigxyy,sepnoizxyy,sepsigxyx,sepnoizxyx

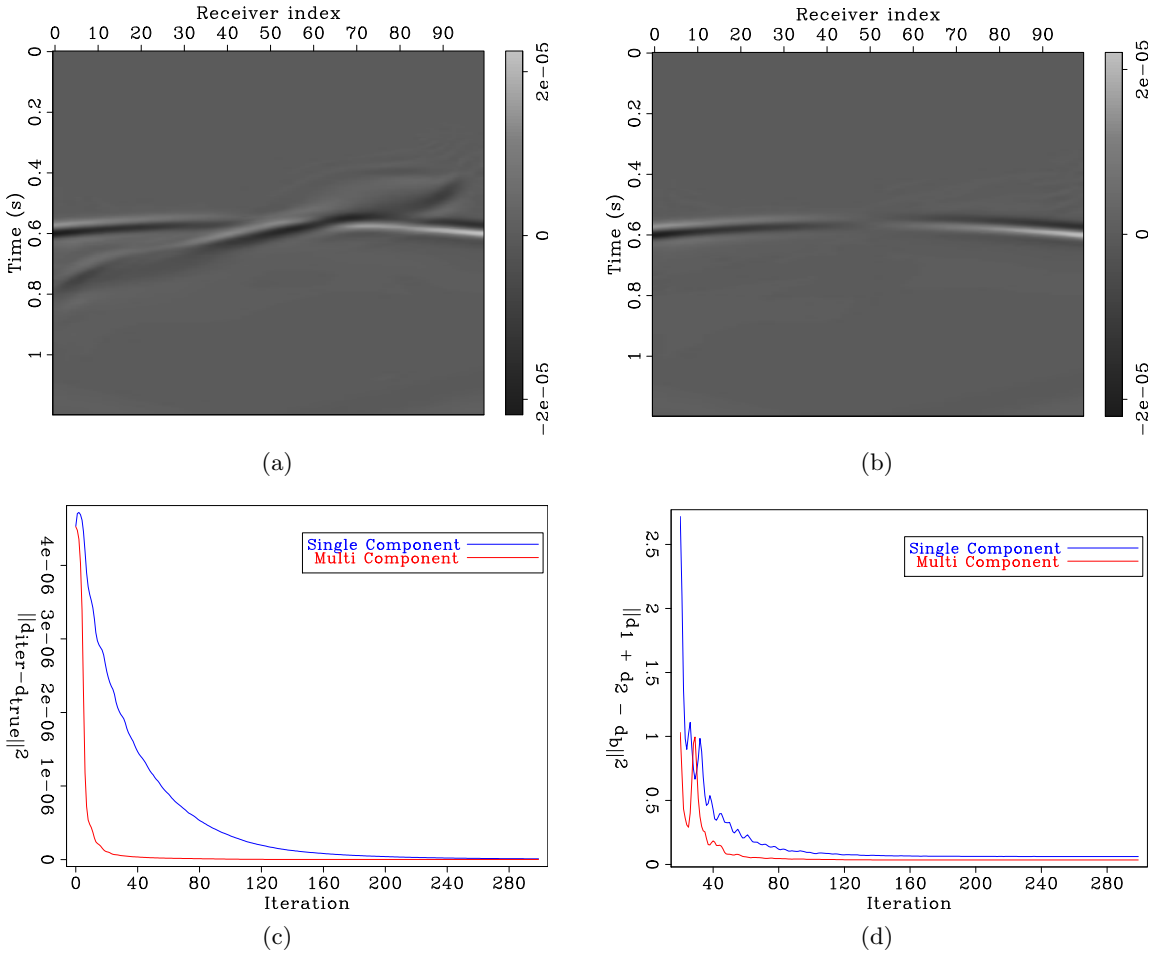


Figure 9: (a) The updated X component of shot one after 60 iterations of minimizing the objective function in equation 10 using a single component PEF. (b) The updated X component of shot one after 60 iterations of the same inversion, but using the multicomponent PEF. (c) The model residual of the single component (blue) and the multicomponent (red) PEF inversions. (d) The deblending data residual (first term in equation 10) of the single component (blue) and multicomponent (red) PEF inversions. After 60 iterations, while the single component PEF result still has a significant amount of noise from the other shot, the multicomponent appears to have nearly recovered the signal. [**ER**] joseph1/. vxda60iter,xy60iter,modnrmcmp,datnrmcmp

experiments to gain a better intuition for these multicomponent PEFs.

SVD and multicomponent PEF comparison

Comparing the results of the two multicomponent filters (SVD vs multicomponent PEF) shown in Figures 5 and 8, it is apparent that the PEF-based approach provided much better results. While it is possible to extend the SVD-based approach beyond a 1D diagonal weighting operator (Schimmel and Gallart, 2003; Paulus and Mars, 2006) that would lead to better filtering results, we prefer the PEF approach as it more naturally extends to a convolutional filter (linear operator) formulation that can used for multidimensional signal and noise separation. Additionally, the PEF-based approach is entirely linear in the data, whereas the SVD introduces additional non-linearity into our new LWI process. In spite of the results obtained using the multicomponent PEF, we are still uncertain of how much of the directional information it uses in the signal and noise separation, while the SVD-based filtering is designed to remove noise that has a specific polarization. While there do exist parallels between the SVD-based filter and the multicomponent PEF, much more research has yet to be done on this topic.

CONCLUSION

We have presented a new algorithm that has the potential to provide better images resulting from LWI of simultaneous-source blended data. In addition to the traditional LWI of blended data, this algorithm requires that at each iteration the shots be modeled independently and that polarization filters be estimated on these proxy shots. These polarization filters will be later applied on the blended data providing a preconditioned data. As this occurs at each iteration, this in fact turns the linear inversion, into a non-linear inversion where the objective function changes at each iteration. We presented a potentially new polarization filter, the multicomponent PEF, that provided the best filtering results when compared to traditional SVD-based polarization filters and single component PEFs. The estimation and application of these multicomponent PEFs at each iteration of LWI can potentially reduce the cross-talk artifacts that are introduced into the gradients at each iteration of LWI therefore providing overall better convergence of LWI of blended data.

ACKNOWLEDGEMENTS

We would like to thank the sponsors of the SEP for their financial and intellectual support. Additionally, the first author would like to thank Ohad Barak for helping him through the basics of SVD-based polarization filtering.

REFERENCES

Abma, R. L., 1995, Least-squares separation of signal and noise using multidimensional filters: PhD thesis, Stanford University.

Barak, O., 2017, Seismic rotational data: Acquisiton, processing and applications: PhD thesis, Stanford University.

Boonyasiriwat, C., G. T. Schuster, et al., 2010, 3d multisource full-waveform inversion using dynamic random phase encoding: Presented at the 2010 SEG Annual Meeting.

Dai, W., X. Wang, and G. T. Schuster, 2011, Least-squares migration of multisource data with a deblurring filter: Geophysics, **76**, R135–R146.

Flinn, E., 1965, Signal analysis using rectilinearity and direction of particle motion: Proceedings of the IEEE, **53**, 1874–1876.

Jackson, G., I. Mason, and S. Greenhalgh, 1991, Principal component transforms of triaxial recordings by singular value decomposition: Geophysics, **56**, 528–533.

Jennings, J. and S. Ronen, 2016, Deblending using the radiality attribute: SEP-Report, **165**, 123–132.

Jiang, Z., R. Abma, et al., 2010, An analysis on the simultaneous imaging of simultaneous source data: Presented at the 2010 SEG Annual Meeting.

Montalbetti, J. F. and E. R. Kanasewich, 1970, Enhancement of teleseismic body phases with a polarization filter: Geophysical Journal International, **21**, 119–129.

Paulus, C. and J. I. Mars, 2006, New multicomponent filters for geophysical data processing: IEEE transactions on geoscience and remote sensing, **44**, 2260–2270.

Ruan, K., J. Jennings, E. Biondi, R. G. Clapp, et al., 2015, Industrial scale high-performance adpative filtering with PEF applications: SEP-Report, **160**, 181–192.

Schimmel, M. and J. Gallart, 2003, The use of instantaneous polarization attributes for seismic signal detection and image enhancement: Geophysical Journal International, **155**, 653–668.

Spitz, S., G. Hampson, and A. Pica, 2008, Simultaneous source separation: A prediction-subtraction approach, *in* SEG Technical Program Expanded Abstracts 2008, 2811–2815, Society of Exploration Geophysicists.

Tang, Y., B. Biondi, et al., 2009, Least-squares migration/inversion of blended data: Presented at the 2009 SEG Annual Meeting.

Trefethen, L. N. and D. Bau III, 1997, Numerical linear algebra, volume **50**: Siam.

Vidale, J. E., 1986, Complex polarization analysis of particle motion: Bulletin of the Seismological society of America, **76**, 1393–1405.

Xue, Z., Y. Chen, S. Fomel, J. Sun, et al., 2014, Imaging incomplete data and simultaneous-source data using least-squares reverse-time migration with shaping regularization: Presented at the 2014 SEG Annual Meeting.

Fitting while whitening nonstationary residuals

Jon Claerbout

ABSTRACT

While fitting a model to given data, I simultaneously whiten nonstationary residuals in both data space and model space. Last year we learned, in the presence of nonstationary noise in multidimensional space, how to solve three classes of problems: (1) deconvolution, (2) missing data, and (3) simulating data with the nonstationary spectrum of given data. This year I extend that class of problems to a general model $\mathbf{m}$ building for data fitting, $\Delta\mathbf{m} = \mathbf{F}^T\mathbf{A}^T\mathbf{A}(\mathbf{Fm} - \mathbf{d})$. In a space of any dimension we need only store $\mathbf{A}$, a Prediction Error Filter (PEF) in the zone of the transition from the filtered to the unfiltered. I propose nine test cases and projects.

INTRODUCTION

Multidimensional PEFs are powerful data analysis tools that can deal effectively with spatially aliased data. Most of the time-series literature (perhaps all of its textbooks) presume data spectra are time invariant (*aka* stationary) while most applications involve spectra that change in time and space. Hence there is a strong need for nonstationary tools. Textbook theory such as my book GIEE[1] tells us that PEF output tends to spectrally white.

We design here a time variable PEF (TV-PEF) that pushes output towards a steady whiteness even while the input spectrum varies. Statistical theory tells us we need Independent Identically Distributed (IID) residuals. In signal and image estimation practice the statistical term "IID" means that signals and images are spectrally white (the leading "I") and of uniform variance (the "ID"). Here I deal only with the first "I". The approach here is "streaming," meaning that the entire data volume need not be kept in memory—it all flows through the box that I define here. Whether your entire process allows streaming naturally depends on whether your $\mathbf{F}$ operator in $\mathbf{0} \approx \mathbf{Fm} - \mathbf{d}$ allows it.

We also overcome a fundamental conundrum. Inverse theory, invented by mathematicians, says it will guide us to the "best" solution, but first we must supply a good deal of statistical information about the "best" solution. Whoa! Additionally, we must supply a good deal of statistical information about the misfit of theory and experiment. Hold your horses! If we haven't seen a solution yet, how can we know the statistics of the misfit of models to data?

We're pragmatic people. Let us not be so fussy. We can always begin by guessing that the data misfit statistics will resemble the data statistics. Since we study seismic image estimation it seems reasonable to start with the assumption that the data variance, decreasing as it does with t^{-2}, matches the data residual variance. That's for the time

[1] `http://sep.stanford.edu/sep/prof/gee/book-sep.pdf` page 182

domain. Likewise for space. In this paper we work only with multidimensional filtering, which amounts to weighting in the multidimensional frequency domain.

In principle, PEFs whiten signals (GIEE). We are just starting out. We'd like to know the PEF that whitens the ultimate fitting residuals that we don't have yet. So, let us change the PEF at every iteration. By whitening our current residuals we whiten our final residuals. This is more proper than guessing that the final residual spectrum is the same as the data spectrum! (Theory states that iteration-variable weighting harms some solvers, while others are immune.) What happens to people who use suboptimal weights and PEFs? They do not use their data "efficiently" (a statistical term). We are ready to go, wondering if we might consistently outperform our forebears in iteration count as well as solution quality. Even if not, we remain enthusiastic. The tool here handles both **nonstationarity** and **spatial aliasing**, massive issues in seismic imaging.

BEGIN IN ONE DIMENSION

Begin with 1.5 pages of review of work I did last year with Sergey. Suppose we have a PEF that represents all previous moments in time. Call it $\bar{\mathbf{a}} = (1, \bar{a}_1, \bar{a}_2, \bar{a}_3, \cdots)$. Say that $\bar{\mathbf{a}}$ represents data values $(d_1, d_2, d_3, \cdots, d_{98})$. We seek to define the $\mathbf{a}$ that represents that data with an appended data value d_{99}.

Consider the regression:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \approx \left[\begin{array}{cccc} d_{99} & d_{98} & d_{97} & d_{96} \\ \hline \gamma & \cdot & \cdot & \cdot \\ \cdot & \gamma & \cdot & \cdot \\ \cdot & \cdot & \gamma & \cdot \\ \cdot & \cdot & \cdot & \gamma \end{array}\right] \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} - \gamma \left[\begin{array}{c} 0 \\ \hline 1 \\ \bar{a}_1 \\ \bar{a}_2 \\ \bar{a}_3 \end{array}\right] \tag{1}$$

The top row says we are trying to fit a new data point d_{99}. The bottom block says the new PEF $\mathbf{a}$ should be pretty similar to the PEF that fit earlier data, $\bar{\mathbf{a}}$. The parameter γ should be big enough that the new data point d_{99} does not change $\mathbf{a}$ very much. Rewrite equation (1) as

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \approx \begin{bmatrix} d_n & d_{n-1} & d_{n-2} \\ \gamma & 0 & 0 \\ 0 & \gamma & 0 \\ 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} - \begin{bmatrix} -d_{n+1} \\ \gamma\,\bar{a}_1 \\ \gamma\,\bar{a}_2 \\ \gamma\,\bar{a}_3 \end{bmatrix} \tag{2}$$

or, in a shortened block-matrix notation, we have the residual to minimize

$$\mathbf{0} \approx \mathbf{r} = \begin{bmatrix} \mathbf{d}^T \\ \gamma\,\mathbf{I} \end{bmatrix} \mathbf{a} - \begin{bmatrix} -d_{n+1} \\ \gamma\,\bar{\mathbf{a}} \end{bmatrix}, \tag{3}$$

where $\mathbf{I}$ is the identity matrix and

$$\mathbf{d} = \begin{bmatrix} d_n \\ d_{n-1} \\ d_{n-2} \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix},$$

For decades Bernard "Bernie" Widrow (Wikipedia) attacked problems of this nature by defining a quadratic form and finding its gradient. Then he repeatedly made small steps down the gradient (not up). How big are the small steps? Experience teaches. The quadratic form is $\mathbf{r}^T\mathbf{r}$. We take its derivative to find the search direction.

$$\Delta\mathbf{a} \quad = \quad -\,(\text{some constant})\; \frac{\partial}{\partial \mathbf{a}^T}\bigg|_{\mathbf{a}=\bar{\mathbf{a}}} \mathbf{r}^T\mathbf{r} \tag{4}$$

Form the transpose of the residual (3) and then differentiate by $\mathbf{a}^T$. (By $\mathbf{a}^T$ we mean the complex conjugate transpose of $\mathbf{a}$.)

$$\frac{\partial \mathbf{r}^T}{\partial \mathbf{a}^T} \quad = \quad \frac{\partial}{\partial \mathbf{a}^T}\; \Big\{\mathbf{a}^T[\mathbf{d} \quad \gamma\mathbf{I}] \;-\; [-d_{n+1} \quad \gamma\bar{\mathbf{a}}]\Big\} \quad = \quad [\mathbf{d} \quad \gamma\mathbf{I}] \tag{5}$$

and multiply that onto $\mathbf{r}$ from (3) keeping in mind that $\mathbf{d}^T\bar{\mathbf{a}}$ is a scalar.

$$\Delta\mathbf{a} \quad \propto \quad \frac{\partial \mathbf{r}^T}{\partial \mathbf{a}^T}\quad \mathbf{r} \quad = \quad [\mathbf{d} \quad \gamma\mathbf{I}]\; \left\{ \begin{bmatrix} \mathbf{d}^T \\ \gamma\,\mathbf{I} \end{bmatrix} \mathbf{a} \;-\; \begin{bmatrix} -d_{n+1} \\ \gamma\,\bar{\mathbf{a}} \end{bmatrix} \right\} \tag{6}$$

$$= \quad \mathbf{d}(\mathbf{d}^T\mathbf{a}) + \gamma^2\mathbf{a} + \mathbf{d}d_{n+1} - \gamma^2\bar{\mathbf{a}} \tag{7}$$

$$\Delta\mathbf{a} \quad \propto \quad \frac{\partial \mathbf{r}^T}{\partial \mathbf{a}^T}\bigg|_{\mathbf{a}=\bar{\mathbf{a}}}\quad \mathbf{r} \quad = \quad (\mathbf{d}^T\bar{\mathbf{a}} + d_{n+1})\; \mathbf{d} \tag{8}$$

It is certainly surprising that the analytic solution to the regression (1) computationally amounts to a single step of the optimization strategy (8), a strategy so crude as to be absent from textbooks. Yet that is so (REFER TO SEP REPORT)[2] Experimentalists will first notice that (1) demands we supply a not-given constant γ while (8) and its coming equivalent (9) demands a not-given constant λ.

Now we recast (8) so as to best understand how to choose the not-given step size. The expression $(\mathbf{d}^T\bar{\mathbf{a}} + d_{n+1})$ in (8) is the prediction error. The prediction filter $\mathbf{a}$ takes data to predicted data, so it has no physical units. The parameter we define for the aggressiveness of PEF adaptation will be unit free. The gradient (8) has units of data squared so it needs a normalizing factor of the same units such as a local data variance $\hat{\sigma}_d^2$. Now we choose a dimensionless parameter λ that we will learn from experience. I put λ in the denominator so the larger the value of λ the slower the adaptation rate. Thus λ represents something like the size of the statistical averaging region measured in pixels. I would begin experimentation with $\lambda = 100$.

$$\Delta\mathbf{a} \quad = \quad \mathbf{a} - \bar{\mathbf{a}} \quad = \quad -\,\frac{1}{\lambda}\left(\frac{\text{prediction error}}{\hat{\sigma}_d^2}\right)\; \mathbf{d} \tag{9}$$

We likely wish to define local variance $\hat{\sigma}_d^2$ with the same averaging region that we are defining the PEFs, so

$$\hat{\sigma}_{d_t^2} \quad = \quad \frac{\lambda - 1}{\lambda}\; \hat{\sigma}_{d_{t-1}^2} \;+\; \frac{1}{\lambda}\; d_t^2 \tag{10}$$

Perhaps the divisor in equation (9) should scale by $\sigma_r\sigma_d$ instead of σ_d^2. Experience will show.

It is almost obvious how and why Equation (9) works. It adds a little piece of data $\mathbf{d}$ to the filter $\Delta\mathbf{a}$. That same little piece of data is going into the filter and $\mathbf{d}\cdot\mathbf{d} > 0$.

[2]http://sep.stanford.edu/sep/jon/streamingSergey.pdf

Whether that piece should be scaled positively or negatively depends on the polarity of the prediction error (which depends on d_{n+1}). Stunningly simple, is it not? And the cost is near nothing.

APPLYING THE ADJOINT OF A STREAMING FILTER

Those of us with a long history of filtering think of a filter adjoint as running the filter backwards. That view arises with recursive filters whose adjoint must indeed run backwards. With nonrecursive filters, such as prediction error, there is a more basic view. In a (nonrecursive) linear operator program, the inputs and outputs can be exchanged to produce the adjoint. For example the pseudocode below applies a filter "a" to data "d" to get a residual "r".

```
do it = na, nt {
        do ia = 1, na {
                if operator itself
                        y(it) += x(it-ia+1) × a(ia)        # one output y(t) pulls
                if adjoint
                        x(it-ia+1) += y(it) × a(ia)        # one input y(t) pushes
                }
        }
```

Observe the time axis runs forward for both the operator and its adjoint $\mathbf{A}^T$. To update a model requires both a PEF $\mathbf{A}$ and its adjoint. Define

$$\mathbf{r} = (\mathbf{Fm} - \mathbf{d}) \tag{11}$$
$$\mathbf{q} = \mathbf{A}(\mathbf{Fm} - \mathbf{d}) \tag{12}$$
$$\mathbf{s} = \mathbf{A}^T\mathbf{A}(\mathbf{Fm} - \mathbf{d}) \tag{13}$$
$$\Delta\mathbf{m} = \mathbf{F}^T\mathbf{A}^T\mathbf{A}(\mathbf{Fm} - \mathbf{d}) \tag{14}$$

You take care of $\mathbf{F}$ and $\mathbf{r}$. I next give you pseudocode for $\mathbf{s} = \mathbf{A}^T\mathbf{Ar}$.

```
do it = na, nt {
        Using equation (9), update filter a(ia=1,na)
        do ia = 1, na {
                q(it) += r(it-ia+1) × a(ia)        # pull
                }
        do ia = 1, na {
                s(it-ia+1) += q(it) × a(ia)        # push
                }
        Move a to abar.
        }
```

Notice that the program also works when the time axis is run backwards. In two dimensions,

either or both the axes may be run backwards. Flipping axes flips the region in which statistics are gathered.

AVERAGING IN TIME AND SPACE

A streaming 1-D prediction filter is an average of earlier prediction filters, however these earlier filters need not be all saved in memory. Since they vary smoothly we may simply use the most recent one. Take the $\bar{\mathbf{a}}$ of one dimension. In two dimensions, for example time and space, it becomes some average of its previous value on each of those two axes.

$$\bar{\mathbf{a}}(t,x) \quad = \quad \cos^2\theta\ \mathbf{a}(t-\Delta t,x) \ + \ \sin^2\theta\ \mathbf{a}(t,x-\Delta x) \tag{15}$$

where θ is an adjustable parameter for the user to specify the shape of the region for gathering statistics. With equation (15), equation (9) becomes

$$\mathbf{a}(t,x) \quad = \quad \bar{\mathbf{a}}(t,x) \quad - \quad \frac{1}{\lambda}\left(\frac{\text{prediction error}}{\hat{\sigma}_d^2}\right)\mathbf{d} \tag{16}$$

Sweeping across the x axis for all t requires in memory all PEFs at $x-\Delta x$, but only the one PEF at $t-\Delta t$. In 3-D it looks like we will need a plane of PEFs. In higher dimensional spaces we need store PEFs only in the zone of the transition from the filtered to the unfiltered. Thus in 5-D we need to store a 4-D volume of PEFs. Don't let that trouble you though. Since the PEFs are all smoothly variable they could be linearly interpolated from a sparse mesh. PEFs on the previous trace $\mathbf{a}(t,x-\Delta x)$ might be smoothed symmetically on the time axis so the region of averaging expands from a quadrant to a halfspace.

PEFS FOR BOTH FITTING AND STYLING

Having PEF $\mathbf{A}$ on the regularization and PEF $\mathbf{B}$ on the fitting, the gradient is

$$\Delta\mathbf{m} = \mathbf{F}^T\mathbf{B}^T\mathbf{B}\mathbf{r} + \epsilon^2\mathbf{A}^T\mathbf{A}\mathbf{m} \tag{17}$$

which may be coded as we did equations (11)-(14). Naturally, the PEF $\mathbf{B}$ is applicable only on those data-space axes that are regularly sampled such as along a marine cable.

We have ignored preconditioning. It's important for covering large gaps such as at cable ends. But in most applications we have more modest goals such as data sampling irregularities and gaps the size of streamer separations. Furthermore, the speed of this method might render preconditioning irrelevant even on larger gaps.

Big stoppers for conventional analysis are two: (1) nonstationarity (2) spatial aliasing. Our PEFs can handle both.

CONCLUSIONS AND FUTURE WORK

The theory and programming sketch above seems fairly complete and ready to run. Some of us may use `cgstep` while new students use `sepsolve`. Here are some preliminary tests, some of them pathfinders of textbook value.

1. **Extend a shot gather:** We should see conflicting dips extended appropriately off the cable ends. We could always try preconditioning (polynomial division), but because the new methods are ultra fast they alone might compensate for lack of preconditioning?
2. **Velocity transform:** It's always fun to transform (t, x) to (τ, s). Do we automatically get sharp resolution? Can we obtain more sparsity in model space simply by introducing softclip? This is really promising and should be easy to test. Is it relevant or irrelevant to estimate off-end traces in data space?
3. **Streaming interpolation while stacking** Could we regularize and interpolate data, then add it into the final image without ever storing the interpolated data? Start by assuming one of every three traces is missing.
4. **Flip flop shooting** In multistreamer seismic data acquisition left- and rightside guns fire alternating pops. How is this paper relevant?
5. **Madagascar:** Should be a pleasing demonstration of nonstationarity. There are a few spikes in the interior, and many around the periphery but code exists to mask out those data values.
6. **Vesuvius:** Are we correct to expect different PEFs on different sides of the mountain? Textbook method lacks weights, but they are easy to include. Textbook also coarse binned the given high density data partly to reduce noise but mainly to speed computation. Is coarse binning still necessary?
7. **Galilee:** To avoid all the noise we could use the solution from the book to make noise-free data. On the real data, large noise is the main issue. It would be delightful should we discover replacing the data residual by its softclip solves the problem of nonstationary model with bursty noise.
8. **Flipping axes:** Running a time or space axis backwards flips the quadrant in which statistics are gathered. What are the disadvantages and possible advantages? When should we be running both ways?
9. **Deep learning:** Start from the assumption that one of the above projects works well. By testing a wide range of shapes and sizes (λ and θ) we should gain insights about the appropriate choice. Nonstationary filters adapt to changing changing spectra, but they don't know how slow (big size) to adapt. Can we quantify it, teaching them to learn? See Bernard Widrow at Wikipedia.

SIGNAL GRIDDING

Old thoughts

A universal problem in applied geophysics is signals (time functions) on irregular (x, y)-space axes. Denote the signals $\mathbf{d}_t(x, y)$. With $\mathbf{L}$ being linear interpolation and $\mathbf{A}$ being a regularizer, the gridding problem might be posed as a collection of many 2-D regressions, each with the same operator.

$$\begin{bmatrix} \mathbf{L} \\ \mathbf{A} \end{bmatrix} [\cdots \mathbf{m}_t \mathbf{m}_{t+1} \mathbf{m}_{t+2} \cdots] \quad \approx \quad \begin{bmatrix} \cdots \mathbf{d}_t \mathbf{d}_{t+1} \mathbf{d}_{t+2} \cdots \\ \cdots \mathbf{0} \cdots \end{bmatrix} \tag{18}$$

Write the collection of desired solutions as $\mathbf{M} = [\cdots \mathbf{m}_t \mathbf{m}_{t+1} \mathbf{m}_{t+2} \cdots]$ and likewise for $\mathbf{D}$. The analytic solution to the regression (18) is $\mathbf{M} = (\mathbf{L}^T\mathbf{L} + \epsilon^2 \mathbf{A}^T\mathbf{A})^{-1}\mathbf{L}^T\mathbf{D}$. Iterative solvers easily give a solution for any one of the $\mathbf{m}_t$, but we had no sensible way to come up with the inverse of the 2-D operator $(\mathbf{L}^T\mathbf{L} + \epsilon^2\mathbf{A}^T\mathbf{A})^{-1}$. Given it, we would easily have $\mathbf{m}_t$ for all t, though in principle it's a full, dense matrix, so most likely would be approximated by a low order product of $\mathbf{L}$'s and $\mathbf{A}$'s which starts to look like solving the original regression at each t.

New thoughts

Let us try another approach. Define $\mathbf{m} = [\cdots \mathbf{m}_t \mathbf{m}_{t+1} \mathbf{m}_{t+2} \cdots]$ as a 3-D (t, x, y) space. Like equations (17) and (11)-(14) we may deduce PEFs $\mathbf{B}$ and $\mathbf{A}$ and the gradient $\Delta\mathbf{m}$

$$\Delta\mathbf{m} = \mathbf{L}^T\mathbf{B}^T\mathbf{B}(\mathbf{L}\mathbf{m} - \mathbf{d}) + \epsilon^2\mathbf{A}^T\mathbf{A}\mathbf{m} \tag{19}$$

Naturally, the PEF $\mathbf{B}$ is nontrivial only on data-space axes that are regularly sampled. Equation (19) is fundamentally better than (18) because we do not impose the PEFs but find them. Additionally, in (19) all the operators are local, suitable for streaming. Admittedly, we are iterating 2-D operators over the entire t-space. We recall the cost of finding PEFs is merely double the cost of applying them. But surely, this cannot fill big data gaps in (x, y) without iterating (19). I would first try $\epsilon = 1$.

As for the time axis, in principle we handle it with an infinitesimal adaptation rate, though in practice we'd find some other way to the limit. I won't understand this until I code it.

Subsurface model inversion: pushing the limits of resolution

Musa Maharramov and Biondo L. Biondi

ABSTRACT

Detecting and inverting the imprint of changing subsurface elastic parameters on seismic data lies at the heart of time-lapse seismic imaging for reservoir monitoring. In this work we demonstrate that the recently proposed technique of simultaneous time-lapse full-waveform inversion with a model-difference regularization can be used to extract high-resolution information on magnitude and location of subsurface velocity and stress anomalies, potentially providing valuable input for reservoir monitoring and assessment of geohazards.

INTRODUCTION

Fluid movement in the subsurface and the associated changes in saturation translate into changes of the subsurface elastic parameters. Stress changes, whether due to fluid extraction/injection or deformation (such as slips on preexisting faults) affect the elastic parameters as well. Detecting and inverting the imprint of changing subsurface elastic parameters on seismic data is key to time-lapse seismic imaging for reservoir monitoring (Johnston, 2013). A recently developed methodology for time-lapse full-waveform inversion (FWI) based on a simultaneous inversion with a total-variation (TV) model-difference regularization (Maharramov et al., 2015, 2016) has been demonstrated to achieve multi-scale inversion of subsurface changes in the presence of strong repeatability issues. The main objective of this work is to demonstrate that even the weakest subsurface changes can still be recovered by the joint full-waveform inversion with regularization. More specifically, using a synthetic example with added noise, we demonstrate that reliable indicators of the subsurface stress change and location of the stress anomaly can be extracted from seismic data. Moreover, relative magnitudes of the inverted model changes at various vintages represent the actual relative changes. We discuss the implications of this observation for reservoir monitoring and seismic hazard assessment.

A change of the subsurface stress field (more specifically, change of the effective stress) results in changes of the acoustic velocity (Johnston, 2013). The following empirical relation is often used to describe this dependence in reservoir rocks,

$$V \;=\; V_\infty \left(1 - A \exp -\frac{P}{P_0}\right), \tag{1}$$

where V is acoustic velocity, P is *isotropic* effective stress, and V_∞, A, P_0 are positive fitting constants for various types of rocks (Domenico, 1977; Zimmer, 2003; Lee, 2003; Johnston, 2013)—see Figure 1. Equation (1) means that velocities increase in "compacting" rocks with increasing effective stress, flattening out at high effective stress—see, for example, Figure

The reproducibility of the results presented in this article will be tested by the end of May 2017.

20 from Chapter 3 of Johnston (2013). Typically, a few megapascal change in the effective stress results in a few tens of meters per second change in the acoustic velocity within the affected rocks. The evolution of subsurface elastic moduli in response to changes in the effective isotropic stress as a result of compaction of pore pressure change is a central topic of seismic time-lapse analysis, and the subject of continuing active research. Going beyond the effects of isotropic stress change, the effects of shear strain on the elastic properties of granular and clay-rich layered materials (for example, expected compositions of fault gouges) may be quite significant (Knuth et al., 2013), potentially allowing us to study the evolution of tectonic fault zones using time-lapse seismic.

Successfully applying time-lapse seismic to the analysis of subtle subsurface deformation effects requires pushing the limits of resolution. Can relatively small changes of the acoustic and elastic velocities be detected from seismic data? We show that the simultaneous FWI with a sparsity-promoting model-difference regularization can indeed reliably resolve such small changes in the presence of noise.

THEORY

In the first subsection we describe some of the key geomechanical models that govern the effect of subsurface deformation on the elastic moduli, and propose a computational framework for integrating such models with seismic inversion. In the second subsection we conduct a resolution analysis of our proposed multi-scale FWI technique.

Sensitivity of elastic moduli to stress

Experimentally obtained empirical trends of equation (1) and Figure 1 provide the most direct way of relating the effective stress and seismic velocities. However, theoretical rock physics models may still provide a useful and powerful tool in combination with seismic inversion. For the P-wave and S-wave velocities we have

$$V_P \;=\sqrt{\frac{K+4/3\mu}{\rho}},\quad V_S \;=\; \sqrt{\frac{\mu}{\rho}}, \tag{2}$$

where K, μ, and ρ are the bulk and shear rock moduli, and ρ is the rock density. A simple rock model based on an elastic sphere pack subject to confining pressure is given by the Hertz-Mindlin theory (Mindlin, 1949; Mavko et al., 2009), which yields the following expressions for the dry rock elastic moduli:

$$K_{HM} \;=\; \sqrt[3]{\frac{n^2(1-\phi)^2\mu}{18\pi^2(1-\nu)^2}P}, \tag{3}$$

$$\mu_{HM} \;=\; \frac{5-4\nu}{5(2-\nu)}\sqrt[3]{\frac{3n^2(1-\phi)^2\mu}{2\pi^2(1-\nu)^2}P}, \tag{4}$$

where μ and ν are the solid phase shear modulus and Poisson ratio; P is the effective stress; ϕ is the depositional porosity, and n is the coordination number (the average number of contacts per grain) empirically estimated from porosity by the Murphy equation (Avseth et al., 2005),

$$n \;=\; 20-34\phi+14\phi^2. \tag{5}$$

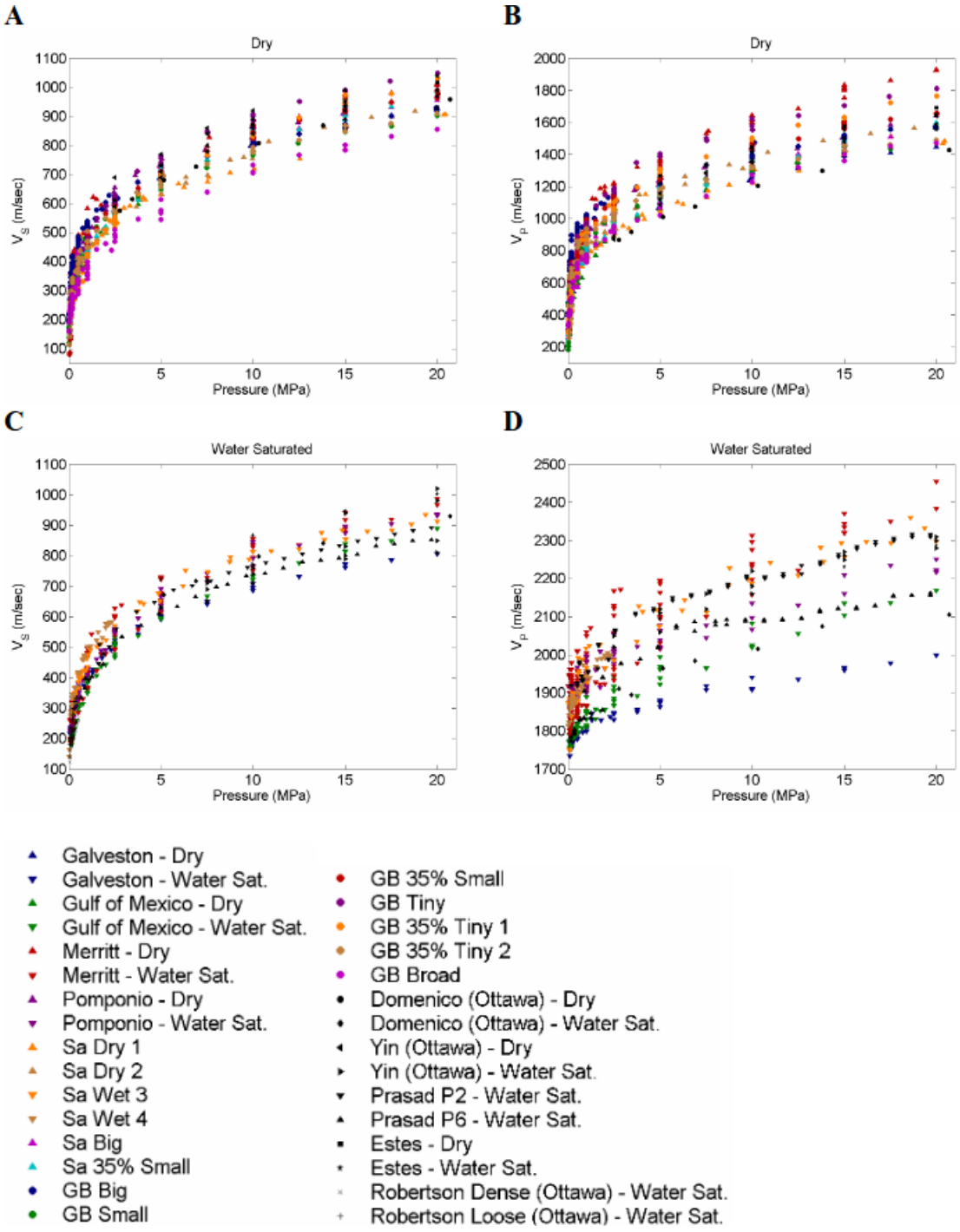


Figure 1: Experimental velocity measurement results and comparison presented by Zimmer (2003) (reproduced from Figure 2.9 of Zimmer (2003)). Laboratory measurements by Zimmer (2003) compared to measurements made on clean sands by Domenico (1977), Prasad and Domenico (1992), Yin (1992), Estes et al. (1994), and Robertson et al. (1995): A) dry shear-wave velocities, B) dry compressional-wave velocities, C) water-saturated shear-wave velocities, and D) water-saturated compressional wave velocities. [**NR**] musa1/. zimmerthesis03

The Gassmann fluid substitution equations can then be used to obtain the bulk modulus of a fluid-saturated rock (Mavko et al., 2009). Rock density in (2) is obviously always an arithmetic mean of the constituent densities.

Equations (2-5) provide simple analytical relations between the seismic velocities on the one hand, and the effective stress, porosity and solid phase moduli, on the other hand. Such relations can be used, in principle, to invert the effective stress from seismic data. However, this simplicity is misleading as the Hertz-Mindlin theory describes an extreme example of a single-mineral sphere packing. However, it may still provide an adequate description of a single constituent part of a composite rock. Estimating the elastic moduli of a sediment from its constituent parts depends on the sequence of depositions, diagenesis and compaction (Mavko et al., 2009) that formed the sediment. For example, a simple harmonic (Reuss) average (Avseth et al., 2005) can be used for estimating the composite moduli of unconsolidated rocks and suspensions. A more realistic compositional model is provided by the Hashin-Shtrikman bounds (Hashin and Shtrikman, 1963; Mavko et al., 2009),

$$K_{HS} = K_1 \; + \; \frac{f_2}{(K_2 - K_1)^{-1} + f_1(K_1 + 4\mu_1/3)^{-1}}, \tag{6}$$

$$\mu_{HS} = \mu_1 \; + \; \frac{f_2}{(\mu_2 - \mu_1)^{-1} + 2f_1(K_1 + 2\mu_1)\left[5\mu_1(K_1 + 4\mu_1/3)\right]^{-1}}, \tag{7}$$

where $K_{1,2}$ and $\mu_{1,2}$ are the bulk and shear moduli of the two phases, and $f_{1,2}$ are the volume fractions of the individual phases. Equations (6,7) are obtained from a theoretical model of a composite rock schematically depicted in Figure 2 from Mavko et al. (2009). The second material forms the darker core spheres, and the first material forms the lighter outer shells. Volume fractions of the shells and cores are f_1 and f_2, respectively. The model assumes that spheres of arbitrarily small sizes completely fill up the entire space.

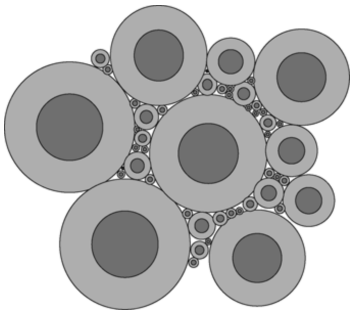


Figure 2: Physical interpretation of the Hashin-Shtrikman bounds for the bulk modulus of a two-phase material (reproduced from Figure 4.1.2 of Mavko et al. (2009)). [**NR**] musa1/. HS

Avseth et al. (2016) used the Hertz-Mindlin theory, Gassmann fluid substitution and nested Hashin-Shtrikman bounds to express the acoustic velocity as a function of the effective stress in patchy, poorly to moderately consolidated sands. The resulting acoustic velocity models were used to estimate seismic time-shifts due to reservoir compaction and production at an offshore field using the following vertical traveltime approximation,

$$\Delta\tau \; = \; 2\int_{z_0}^{z_1} \left(s_b^P - s_m^P\right) dz, \tag{8}$$

in combination with the Gassmann fluid substitution equations that assumes constant stress in all constituent parts

where z_0 and z_1 are the reservoir top and bottom depths, and s_b^P and s_m^P are the baseline and monitor acoustic slowness models. Avseth et al. (2016) then used (8) to invert effective stress changes within the reservoir from observed time shifts.

We demonstrate in the subsequent sections that time-lapse FWI with a model-difference regularization can achieve a stable resolution of small velocity anomalies. Instead of using the equation (8), we propose to use FWI *simultaneously* with equations (2-7) to produce spatially variable estimates of stress changes. While quantitative accuracy of time-lapse FWI may be limited due to, for example, aggressive model-difference regularization, relative magnitudes of velocity changes and their *spatial heterogeneity* can still be well recovered (Maharramov and Biondi, 2017), leading to a new FWI technique of heterogeneous relative stress-change inversion.

So far we have been focusing on the effect of isotropic stress on seismic velocities. However, important geomechanical models of earth deformation, such as the antiplane stress model of an infinite strike-slip fault (Segall, 2010), involve only shear strains. Our ability to observe and measure a build-up of shear stress energy near seismogenic fault zones as a result of interseismic deformation is key to the timely assessment of geohazards. There is a growing body of evidence that suggests that both the velocities and amplitudes of transmitted acoustic and shear waves are affected by shear strains. While we are unaware of any theoretical rock physics models for predicting the evolution of elastic moduli under shear strain, the latest experimental results involving granular/clay mixtures typical for fault gouges (Knuth et al., 2013) point to significant measurable effects of shear strain on the acoustic velocity (see Figure 3).

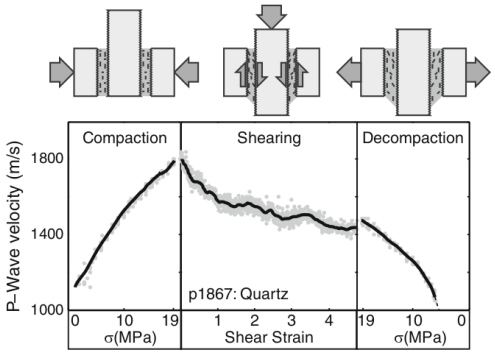


Figure 3: A V_P versus shear strain (middle panel) dependence obtained by Knuth et al. (2013) for granular .1-1 cm thick quartz layers with $\approx 100\mu m$ grain sizes (reproduced from Figure 5 of Knuth et al. (2013)). [**NR**] musa1/. shear

These results imply that high-resolution time-lapse seismic inversion may be able to detect shear deformation—for example, near fault gouges. This information may potentially provide insights into the micromechanics at the fault interface that control earthquake nucleation and rupture propagation.

Multiscale subsurface model inversion

In our method, we invert the baseline and monitor models *simultaneously* by solving the following optimization problem:

$$\text{minimize } \alpha \| \exp i\arg \mathbf{u}_b - \exp i\arg \mathbf{d}_b \|_2^2 + \beta \| \exp i\arg \mathbf{u}_m - \exp i\arg \mathbf{d}_m \|_2^2 + \tag{9}$$

$$\delta \|\mathbf{WR}(\mathbf{m}_m - \mathbf{m}_b)\|_1 \tag{10}$$

with respect to both the baseline and monitor models $\mathbf{m}_b$ and $\mathbf{m}_m$. Problem (9,10) describes a time-lapse FWI with a model difference regularization (10) (Maharramov et al., 2015); $\mathbf{d}_b$ and $\mathbf{d}_m$ are the observed baseline and monitor data; $\mathbf{u}_b$ and $\mathbf{u}_m$ are the predicted baseline and monitor data; $\mathbf{W}$ denotes an optional weighting operator. When $\mathbf{R} = \nabla$, the regularization or "model-styling" term (10) represents a *total-variation (TV) regularization* that promotes "blockiness" of the model-difference, potentially reducing oscillatory artifacts (Rudin et al., 1992). If $\mathbf{R} = \mathbf{I}$ is the identity map, we obtain a sparsity-promoting L_1 model-difference regularization. Problems with $\mathbf{R} = \nabla$ and $\mathbf{R} = \mathbf{I}$ can be solved in a cascaded fashion, resolving "blocky" changes first, followed by spiky velocity-difference anomalies (Maharramov et al., 2016). The data misfit terms (9) are formulated in the frequency domain and represent misfits between the amplitude-normalized observed and predicted wave fields at different vintages, in other words representing phase misfits. Maharramov et al. (2016) showed that minimization of the phase misfits in (9) is to a first order and under the travel-time approximation equivalent to a tomographic inversion of travel-time delays due to changes in the subsurface model. Once the tomographic (background-velocity) component has been resolved, spiky or oscillatory components that may represent amplitude effects can be resolved by replacing the phase-only misfits (9) with the conventional phase-and-amplitude L_2 misfits (Maharramov et al., 2016). However, in this work we focus on a tomographic inversion of time-lapse effects that can be observed in minuscule time-shifts between the monitor and baseline data.

Regularization term (10) plays a dual role in our method: it penalizes oscillatory artifacts in the model difference that may be due to acquisition and computational repeatability issues, and it constrains the inverted model by fitting the sparsest model difference that explains the data. Since problem (9,10) is to a first order equivalent to tomographic model-difference inversion, conceptually we can study the limits of its resolution by considering the constrained tomographic inversion problem

$$\begin{aligned} \|\mathbf{A}\delta\mathbf{s} \; - \; \delta\boldsymbol{\tau}\|_2 \; &< \; \sigma, \\ \|\delta\mathbf{s}\|_0 \; &= \; k, \end{aligned} \tag{11}$$

where $\delta\boldsymbol{\tau}$ is a vector of observed time shifts, $\delta\mathbf{s}$ is the unknown slowness change, $\mathbf{A}$ is the travel-time modeling operator, σ is the 2-norm of the estimated measurement error, and $\|\cdot\|_0$ is the L_0 norm (the number of non-zero components) of a vector. Parameter k in (11) is effectively a sparsity measure of the slowness perturbation. In problem (11) we fit the observed time shifts with a slowness difference of a given sparsity (see Elad (2010) for a discussion of the relation between L_0 and L_1-regularized optimization). If $\delta\mathbf{s}_0$ is a true solution of (11) for some observed time shifts $\delta\boldsymbol{\tau}$, any minimizer $\delta\mathbf{s}$ of (11) satisfies the estimate:

$$\|\delta\mathbf{s} - \delta\mathbf{s}_0\|_2 \; \leq \; \frac{2}{c_{2k}}\sigma, \tag{12}$$

where c_{2K} is the lower restricted isometry constant of operator $\mathbf{A}$ (Candes et al., 2006; Demanet and Nguyen, 2015) defined as

$$c_{2k} \; = \; \min_{J:|J|=2k} \lambda_{\min}\left(\mathbf{A}_J\right), \tag{13}$$

where J is a subset of $2k$ columns of $\mathbf{A}$, $\mathbf{A}_J$ is the operator made up of those columns, $\lambda_{\min}$ is its minimal singular value, and σ is measurement noise. Equation (12) relates

valid for isolated transmission or reflection events

recoverability of a $k-$sparse slowness model to acquisition via the isometry constant (13) and noise in the data. Note that simply increasing dimension of the data (the number of receivers) does not improve the estimate (12): the singular value in (13) and the L_2 norm of the noise grow at the same asymptotic rate with the number of receivers so long as noise distribution of an individual time-shift measurement is the same. However, according to the Central Limit Theorem, temporally redundant measurements reduce the noise level: for the temporal average of N repeated measurements at the same receiver locations

$$\delta\boldsymbol{\tau} \;=\; \frac{1}{N}\sum_{i=1}^{N}\delta\boldsymbol{\tau}^i, \tag{14}$$

the corresponding measurement error in the first line of (11) decreases asymptotically as $\sigma/\sqrt[4]{N} \to 0$. In other words, given sufficiently redundant observations, we should be able to recover the geometry and (qualitatively) the magnitude of a sparse model difference. Averaging in (14) does not necessarily mean multiple surveys; we envisage the use of emergent continuous-source technologies (Kurosawa and Kato, 2015) as a cost-effective alternative to repeated surveys. It should be noted that the magnitude of the subsurface slowness change is recovered qualitatively because the regularization results in a penalization and underestimation of the model difference, however, enhancements exist that can address this phenomenon (Maharramov et al., 2016).

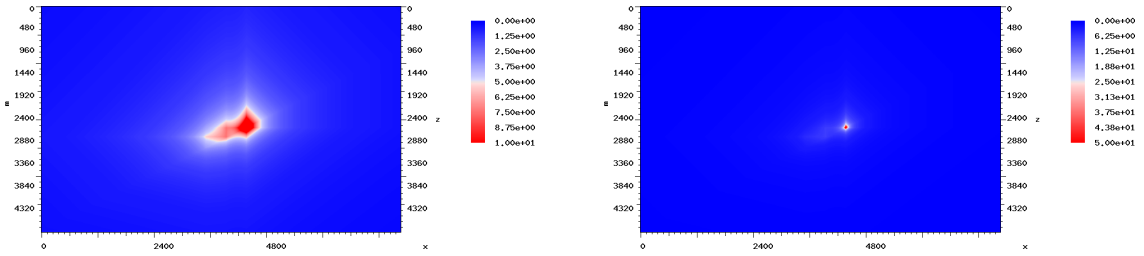
One potentially important implication of estimate (12) for velocity-stress relations of the form (1) is that continuous observations in combination with a robust simultaneous FWI (9,10) can detect relative magnitudes of the subsurface stress changes, such as a "flattening" of (1) for large changes in the effective stress. Our method is merely an inversion tool for detecting small changes in the subsurface, but enhanced monitoring capabilities that can be delivered by this technology open up new, if somewhat speculative at this point, possibilities. For example, can a "flattening" of the velocity-stress curve near a locked fault undergoing stress change due to natural or man-made phenomena indicate an impending slip? While study of viable precursors is well beyond the scope of this work, the inversion technique that we propose may prove instrumental for both conventional and novel techniques of subsurface monitoring.

EXAMPLES

We demonstrate the kind of high-resolution inversion product that can be achieved by our method (9,10) on the synthetic baseline model used by Maharramov et al. (2016). We assume very small velocity changes between first monitor and baseline (Figures 4(a),4(b)) and second and first monitor (Figures 5(a),5(b)).

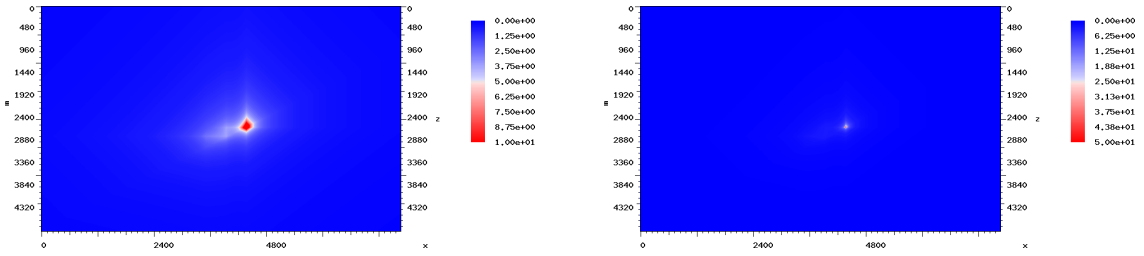
The figures show the true changes at 10 m/s and 50 m/s clip. The true change is modeled to imitate the stress change near a locked segment of a fault undergoing an interseismic slip, with a singularity near the tip of the locked segment. Note that both compressive and tensile stresses are present around a locked fault segment, and according to (1) that means both positive and negative velocity changes (Segall, 2010). However, in this experiment we extracted only a single positive lobe of the velocity change. Naive application of FWI often results in non-physical oscillatory artifacts, or "side-lobes", that in our case can be easily

slowing down of the velocity change rate with increasing stress



(a) (b)

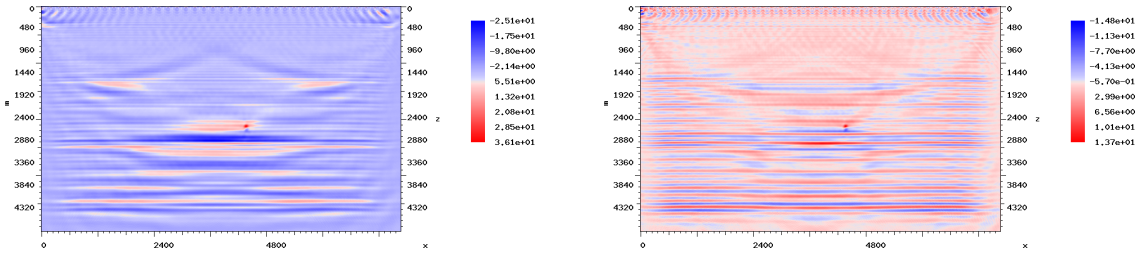
Figure 4: True model difference between the first monitor and base at 10 (a) and 50 (b) m/s clip. [**CR**] musa1/. range10trued1diff,range50trued1diff



(a) (b)

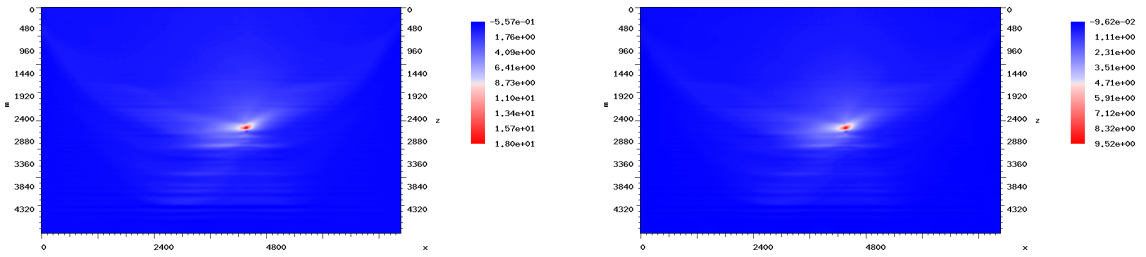
Figure 5: True model difference between the second and first monitors at 10 (a) and 50 (b) m/s clip. [**CR**] musa1/. range10trued2diff,range50trued2diff

mistaken for the effects of stress regime changing from compressive to tensile. Therefore, we demonstrate the robustness of our method (9,10) by recovering a strictly positive velocity-difference anomaly without producing oscillatory artifacts.



(a) (b)

Figure 6: Parallel-difference inversion of the first (a) and second (b) difference from clean synthetics. [**CR**] musa1/. cpard1diff,cpard2diff

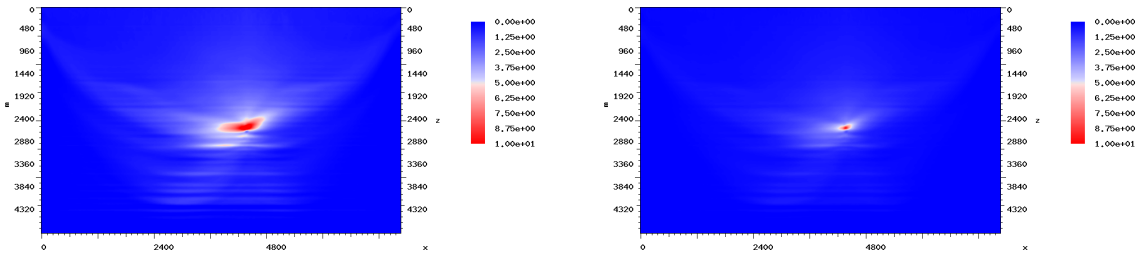


(a) (b)

Figure 7: First (a) and second (b) difference inversion from clean data using simultaneous FWI with a TV model-difference regularization. [**CR**] musa1/. ctvd1diff,ctvd2diff

The results of applying a naive parallel-difference FWI to invert the first and second model difference from clean synthetic data are shown in Figures 6(a),6(b). The clean synthetic was generated using the synthetic model of Maharramov et al. (2016) with 39 shots at 192 m spacing and 320 receivers per shot with a 24 m receiver spacing. Absorbing boundary conditions were applied at the top of the model to avoid surface-related multiples, and a Ricker wavelet centered at 12 Hz was used as the source. FWI was conducted with a frequency continuation from 4 to 20 Hz, starting from a smoothed model of Maharramov et al. (2016). Note that while the parallel-difference FWI resolved location of the stress singularity for both first and second monitor acquisitions, the results are contaminated with oscillatory artifacts of significant magnitude that render them uninterpretable. Solutions of the simultaneous FWI with a TV model-difference regularization for the first and second model differences are shown in Figures 7(a),7(b). Note that these results are now free of oscillatory artifacts, correctly locate the stress singularity (peak velocity change), and largely recover the velocity difference while indicating their relative magnitudes (see the inverted

model differences plotted at 10 m/s clip in Figures 8(a),8(b)).



(a) (b)

Figure 8: First (a) and second (b) difference inversion from clean data using simultaneous FWI with a TV model-difference regularization. 10m/s clip [**CR**] musa1/. range10ctvd1diff,range10ctvd2diff

We repeated the same experiment after adding random noise to the data. Signal-to-noise ratio of the noisy data peaked at about 8 dB (see Figure 9) however the signal-to-noise ratio was below 1 outside of the 4-20 Hz range.

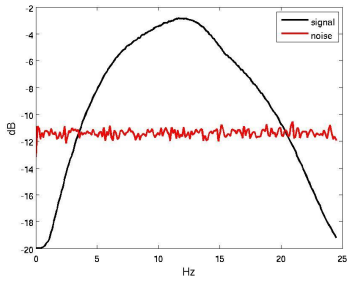
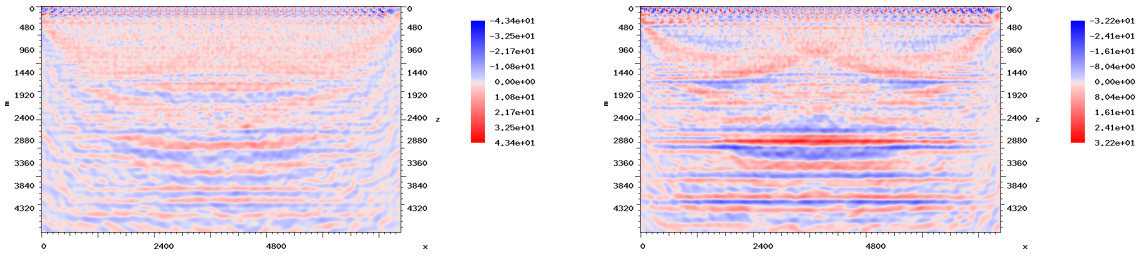


Figure 9: Spectra of the clean data and added noise. [**CR**] musa1/. noise

The results of a parallel-difference FWI with the same inversion parameters as before are shown in Figures 10(a),10(b). Note that the second model difference is completely masked by oscillatory artifacts while the difference on the left panel only hints at the location of the stress singularity.

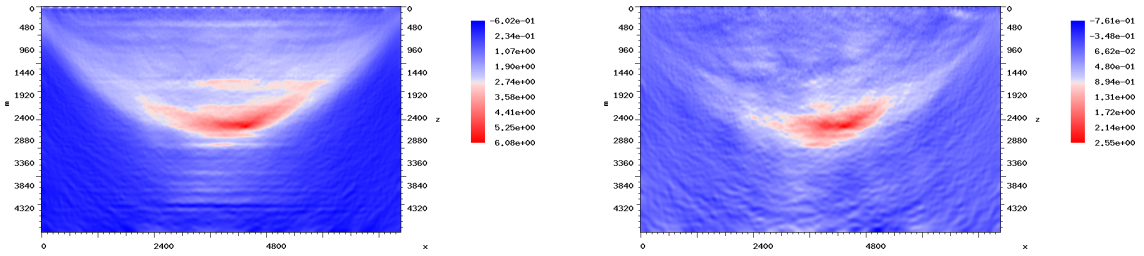
Solving (9,10) with $\mathbf{R} = \nabla$, on the other hand, largely recovered the velocity changes and their relative magnitudes (see Figures 11(a),11(b)), although the absolute magnitudes are underestimated. The TV regularization flattened the velocity peak, resulting in a more ambiguous location of the stress singularity.

However, supplying the result of the TV-regularized simultaneous inversion as the starting model for problem (9,10) with a sparsity-promoting L_1 regularization ($\mathbf{R} = \mathbf{I}$ in (10)) results in the recovery of the sparse velocity peak corresponding to the stress singularity—see Figures 12(a),12(b).



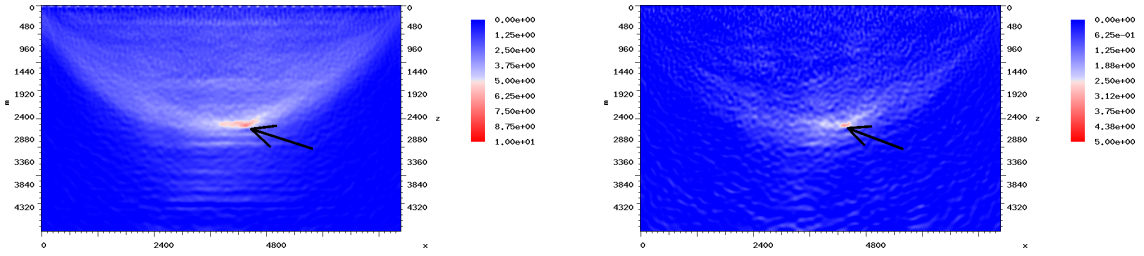
(a) (b)

Figure 10: Parallel-difference inversion of the first (a) and second (b) difference from noisy (right) synthetics. [**CR**] musa1/. npard1diff,npard2diff



(a) (b)

Figure 11: First (a) and second (b) difference inversion from noisy data using simultaneous FWI with a TV model-difference regularization. [**CR**] musa1/. minmaxntvd1diff,minmaxntvd2diff



(a) (b)

Figure 12: First (a) and second (b) difference inversion from noisy data using cascaded inversion with a TV and L_1 model-difference regularization. [**CR**] musa1/. markedrange10nl1d1diff,markedrange5nl1d2diff

CONCLUSIONS

Given continuous or redundant seismic observations, simultaneous time-lapse FWI with a sparsity-promoting model-difference regularization can be used to detect small changes in the subsurface model induced by changes in the stress field or fluid content. Even in the presence of strong noise, model-difference regularization removes oscillatory artifacts from the inverted model difference while retaining useful information. Cascaded inversion with TV and L_1 regularization helps to achieve multi-scale inversion of subsurface changes, potentially pinpointing locations of significant stress change. While absolute values of the inverted model difference are underestimated due to regularization, relative magnitudes are indicative of changing rates of stress, and may be potentially used in seismic hazard studies.

ACKNOWLEDGEMENTS

The authors would like to thank Paul Segall and Jack Dvorkin for very useful discussions, and Stanford CEEES for computing support.

REFERENCES

Avseth, P., T. Mukerji, and G. Mavko, 2005, Quantitative seismic interpretation: Applying rock physics tools to reduce interpretation risk: Cambridge University Press.

Avseth, P., N. Skjei, and G. Mavko, 2016, Rock-physics modeling of stress sensitivity and 4d time shifts in patchy cemented sandstones application to the visund field, north sea: The Leading Edge, **35**, 868–878.

Candes, E. J., J. Romberg, and T. Tao, 2006, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information: IEEE Transactions on Information Theory, **52**, 489–509.

Demanet, L. and N. Nguyen, 2015, The recoverability limit for superresolution via sparsity: arXiv:1502.01385.

Domenico, S. N., 1977, Elastic properties of unconsolidated porous sand reservoirs: Geophysics, **42**, 1339–1368.

Elad, M., 2010, Sparse and redundant representations: Springer.

Estes, C. A., G. Mavko, H. Yin, and T. Cadoret, 1994, Time-lapse inverse scattering theory: SRB Annual Report, **55(B)**, G1–1–G1–9.

Hashin, Z. and S. Shtrikman, 1963, A variational approach to the elastic behavior of multiphase materials: Journal of the Mechanics and Physics of Solids, **11**, 127–140.

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

Knuth, M. W., H. J. Tobin, and C. Marone, 2013, Evolution of ultrasonic velocity and dynamic elastic moduli with shear strain in granular layers: Granular Matters, **15**, 499–515.

Kurosawa, I. and A. Kato, 2015, Permanent seismic source for continuous reservoir monitoring: Reservoir Monitoring Consortium Semi-Annual Review meeting, University of Southern California.

Lee, M. W., 2003, Elastic properties of overpressured and unconsolidated sediments: U.S. Geological Survey Bulletin, **2214**, 1–14.

Maharramov, M., B. Biondi, and S. Ronen, 2015, Robust simultaneous time-lapse full-waveform inversion with total-variation regularization of model difference: 77th EAGE Conference and Exhibition, Extended Abstract, We P3 09.

Maharramov, M. and B. L. Biondi, 2017, Full waveform inversion for reservoir monitoring—pushing the limits of subsurface resolution: First EAGE Workshop on Practical Reservoir Monitoring, Session: Increasing Value through Processing and Imaging, 6 March 2017, Amsterdam, the Netherlands.

Maharramov, M., B. L. Biondi, and M. A. Meadows, 2016, Time-lapse inverse theory with applications: Geophysics, **81**, R485–R501.

Mavko, G., T. Mukerji, and J. Dvorkin, 2009, The rock physics handbook: Cambridge University Press.

Mindlin, R. D., 1949, Compliance of elastic bodies in contact: Journal of Applied Mechanics, **16**, 259–268.

Prasad, M. and R. Domenico, 1992, Attenuation mechanisms in sands: Laboratory versus theoretical (Biot) data: Geophysics, **57**, 710–719.

Robertson, P. K., S. Sasitharan, J. C. Cunning, and D. C. Sego, 1995, Shear-wave velocity to evaluate in-situ state of Ottawa sand: J. Geotech. Eng., **121**, 262–273.

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

Segall, P., 2010, Earthquake and volcano deformation: Princeton University Press.

Yin, H., 1992, Acoustic velocity and attenuation of rocks: Isotropy, intrinsic anisotropy, and stress induced anisotropy: PhD thesis, Stanford University.

Zimmer, M. A., 2003, Seismic velocities in unconsolidated sands: PhD thesis, Stanford University.

Automatic noise exploration in urban areas

Fantine Huot, Yinbin Ma, Robert Cieplicki, Eileen Martin, and Biondo Biondi

ABSTRACT

Near-surface imaging with ambient noise has grown into an increasingly common tool over the past decade thanks to the virtual source method. However, if non-ideal noise sources are present, experts must manually analyze the noise to look for any issues they suspect, then design filters to remove these non-ideal noises. Up until now, the deployment and maintenance of the receiver array was the primary cost, but this is changing with advancements in Distributed Acoustic Sensing (DAS), an emerging technology that repurposes a fiber optic cable as a series of strain sensors. On the Stanford campus we have shown that we can record seismic waves with fiber optic cables sitting loosely in existing telecommunications conduits. As we look forward at the possibility of easily plugging into unused fibers in telecom bundles on-demand, it is clear that manual selection of non-ideal noise sources is the next bottleneck. Herein we show a variety of methods, mixing traditional signal processing and machine learning, to automatically assist geophysicists in analyzing the ambient noise recorded and selecting non-ideal noises. We demonstrate that we can identify different types of noise using clustering algorithms and that template matching can be used for detecting specific events.

INTRODUCTION

By measuring the speed of seismic waves propagating in the Earth's near-surface, we can image the top tens to hundreds of meters of the subsurface, with deeper features being resolved by lower frequencies. These seismic velocity images can be interpreted to evaluate earthquake or landslide risk, to detect permafrost, to find sinkholes or tunnels, or to track near-surface changes related to drilling activities. Additionally, in cases of complex near-surface conditions, resolving this complexity is a prerequisite to obtaining a high-quality image of the deeper subsurface.

By cross-correlating noise recorded at a selected receiver with noise recorded by all other receivers in an array, we can extract signals mimicking an active seismic survey with a source at the selected receiver, called its virtual source response function (Lobkis and Weaver, 2001; Lin et al., 2008; Wapenaar et al., 2010a,b). Thus, when active sources are too costly or logistically prohibitive, passive seismic can be a good option for near-surface imaging. However, the theory is limited by the assumption of homogeneous uncorrelated sources (Wapenaar et al., 2010a). Non-ideal sources can cause artifacts in extracted velocities, but with careful processing experts can overcome these limitations (Bensen et al., 2007; Daskalakis et al., 2016; Zhan et al., 2013), even for anthropogenic sources occurring at higher frequencies (Girard and Shragge, 2016; Martin* et al., 2015; Martin et al., 2016; Nakata et al., 2011, 2015; Zeng et al., 2016). Acquisition cost is a further issue: nearly all of the past ambient noise studies on dense arrays have been temporary arrays which were

labor-intensive to install and maintain (Lin et al., 2008; Nakata et al., 2011, 2015; Zeng et al., 2016).

Distributed acoustic sensing (DAS) is a new acquisition technology being increasingly adopted in the energy industry for microseismic monitoring (Webster et al., 2013) and time-lapse seismic surveys (Daley et al., 2013; Daley* et al., 2014; Mateeva et al., 2014; Bakku, 2015). DAS probes a fiber-optic cable with a laser interrogator unit (IU) to repurpose that fiber as a series of strain sensors. Motivated by avoiding the maintenance cost of node arrays, there have been several recent ambient noise experiments using fiber optics buried in shallow trenches directly coupled to the ground (Ajo-Franklin* et al., 2015; Martin* et al., 2015; Martin et al., 2016; Zeng et al., 2016). In fact, we can sacrifice some ground-to-sensor coupling in favor of easier installation by running fibers in existing telecommunications conduits (Martin et al., 2017). By running fibers in existing conduits, or even plugging into unused fibers in existing telecommunications bundles, easy, on-demand, repeatable seismic studies (even in urban areas) will soon be a reality.

With ambient noise data becoming increasingly easy to record, data volumes are increasing, and we can only extract their full value if we further automate the processing workflow. Herein we introduce some tools and metrics, mixing traditional signal processing and machine learning, to automatically assist geophysicists in analyzing the ambient noise recorded and selecting non-ideal noises.

We present this methodology in the context of a case study: a figure-eight-shaped array of 2.4 km of fiber optics lying loosely in existing telecommunications conduits underneath the Stanford University campus. This particular experiment is ideal because it demonstrates the wide variety of issues coming up on the horizon as we push for broader use of ambient noise. The array detects a wide variety of seismic noise sources that do not conform to the ideals of existing ambient noise theory: it sits in a seismically active region, 20 km from the Pacific ocean, 7 km from the San Francisco bay, with major highways on either side, a variety of roads with differing levels of traffic near the fiber, regular quarry blasts within 15 km, plumbing and HVAC systems throughout the site, multiple construction sites near the array, and foot and bicycle traffic throughout. With over 600 sensors continuously recording 50 samples per second since September 2016, manual inspection of most data is infeasible, making automation tools critical to extracting subsurface information from the data.

After providing a brief overview of the recorded data and the pre-processing steps, we present how unsupervised learning algorithms can help us identify the main types of seismic noise present in the data. We then focus on the high amplitude noise and use template matching to detect specific events.

THE STANFORD DAS ARRAY

The fiber optic cable is deployed in Stanford's telecommunication tunnels in a double loop pattern (Figure 1). Every 8 meters of cable acts as a receiver and records vibration at a sampling rate of 50 Hz, creating a data matrix of 600 channels distributed in space. As a result, we obtain neighboring time series which conveniently lend themselves to image processing (Figure 2). As the fiber optic cable is sitting loosely in existing telecommunications conduits, it is very poorly coupled with the ground. We balance the amplitudes by

bandpass filtering (0.5 and 20 Hz) followed by automatic gain control (AGC). As a result, more weight is given to less noisy channels and we obtain balanced amplitudes across the whole array (Figure 2(c)).

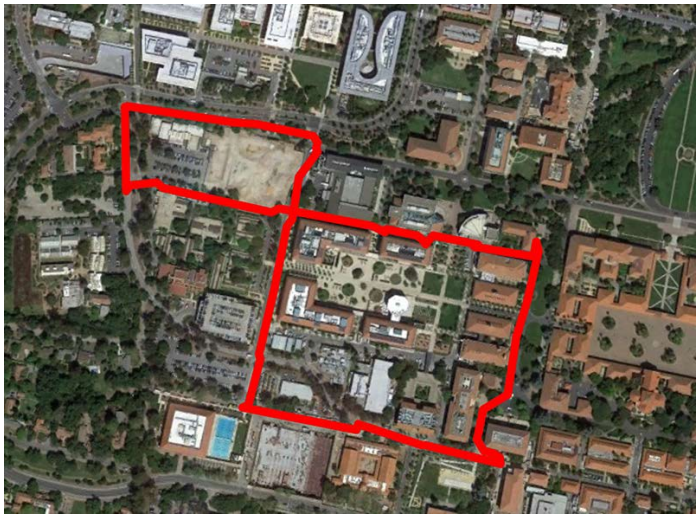


Figure 1: The fiber optic cable (red line) is deployed under Stanford campus in the telecommunication infrastructure tunnels. It forms a North-East and a South-West loop. [**NR**] fantine1/. map

AUTOMATICALLY IDENTIFYING DIFFERENT TYPES OF SEISMIC NOISE

Statistical learning has become a highly explored field in many scientific areas as well as marketing, finance, and other business disciplines. In recent years, new and improved software packages have significantly eased the implementation burden for many statistical learning methods, providing scientists and practitioners with complete toolkits for training, testing, and deploying models with well-documented examples for all these tasks (Abadi et al., 2016; Pedregosa et al., 2011; Collobert et al., 2002; Jia et al., 2014; James et al., 2013). As a consequence, machine learning techniques are increasingly adopted to improve accuracy and speed up processing for seismic applications (Fisher et al., 2016; Yoon et al., 2015)

In order to identify the main types of noise in the DAS data, we performed unsupervised learning on a week-long subset of the data, encompassing the daily variations in the noise field. We applied continuous wavelet transforms (CWT) both along the time axis and the spatial axis to extract features. The CWT is commonly used in pattern recognition, as it has the ability to decompose complex patterns into elementary forms (Mallat, 1999; Sinha et al., 2005). CWT measures the similarity between a signal and an analyzing wavelet by comparing the input signal to shifted and compressed or stretched versions of the wavelet. For an input signal f, and a mother wavelet ψ, the CWT can be expressed as follows:

$$C(a,\, b\,;f(t),\, \psi(t)) = \int_{-\infty}^{\infty} f(t)\, \frac{1}{\sqrt{a}}\, \psi^{*}\left(\frac{t-b}{a}\right)\, dt\ ,$$

where a is the scaling factor, b the time shift, and $*$ denotes the complex conjugate.

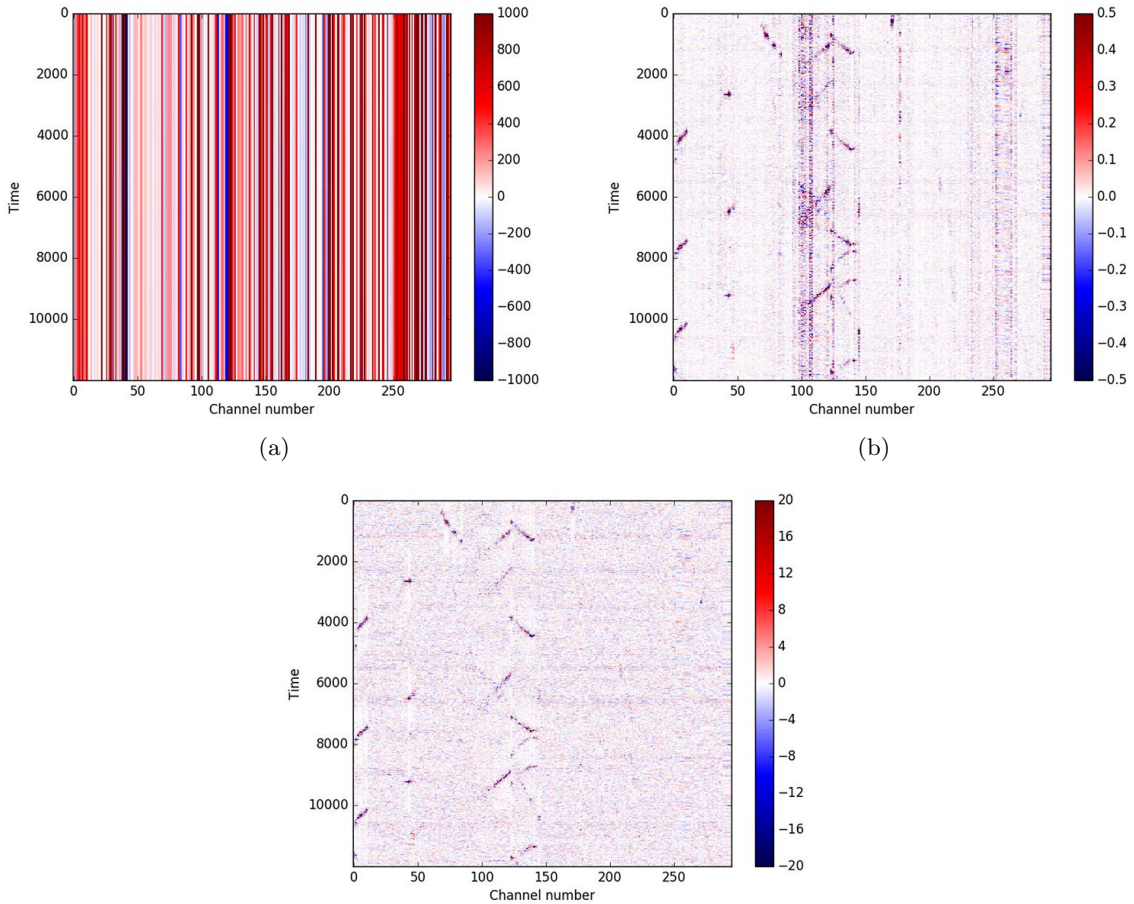


Figure 2: (a) 4 min of raw data (50 samples per second); (b) data after bandpass filter between 0.5 and 20 Hz; (c) data after AGC with a sliding window of 40 seconds. [**ER**] fantine1/. dataraw,databp,dataagc

We used the Morlet wavelet as the analyzing function and subsequently took the amplitude of the resulting complex numbers. For each sample we obtained 100 CWT scale factors, half of them resulting from decomposing the data along the time axis, the other half from decomposing along the spatial dimension. At this stage, we subsampled the features by averaging the CWT scales over windows of 0.5 s, as any seismic events that last less than 0.5 s would be hard to interpret physically. This resulted in a data matrix of 100 features for over a million samples.

Common clustering methods for wavelet domain time series include K-means, agglomerative clustering and self-organizing maps (Liao, 2005; Köhler et al., 2009). Herein we use K-means for faster computation. Figure 4 presents the detected clusters over time, projected over the array's geometry at two different time stamps. We ran the algorithm for different numbers of clusters and qualitatively settled for 4 clusters, each capturing a different portion of the CWT scales, and therefore, easier to interpret. A higher number of clusters merely yielded subdivisions of these 4 main clusters. After examining each cluster's CWT and frequency components, the 4 main types of identified noise can be described as follows (Figure 4): Dark blue represents the incoherent ambient noise field; Light green corresponds to laser noise, responsible for disruptions with large spatial extent over the array; Cyan corresponds to traffic noise. Note how the clustering algorithm captured moving sources such as cars without being provided any information related to the geometry of the array; Orange indicates other sources of coherent noise of medium amplitude, mostly localized on the North-West part of the fiber optic array. These coherent noise sources could be linked to remote traffic noise (from cars that are not running exactly above the fiber), or construction noise (the area circled by the North-West loop is under construction).

While the exact sources of these main types of noise might be difficult to interpret, these results have an important consequence for seismic processing. By automatically detecting different types of coherent noise, we can selectively filter them out in CWT space, and apply inverse CWT to map the data back to the time domain to obtain ambient field data without non-ideal noise sources.

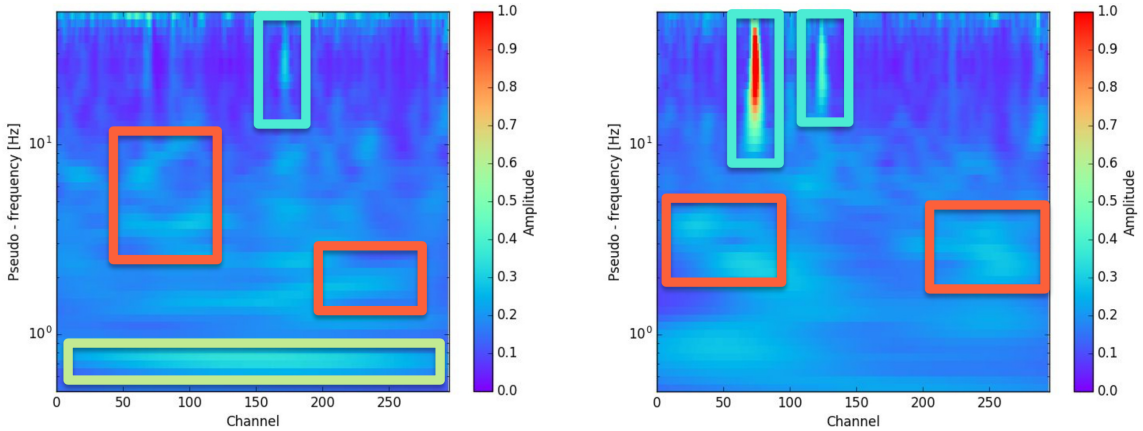
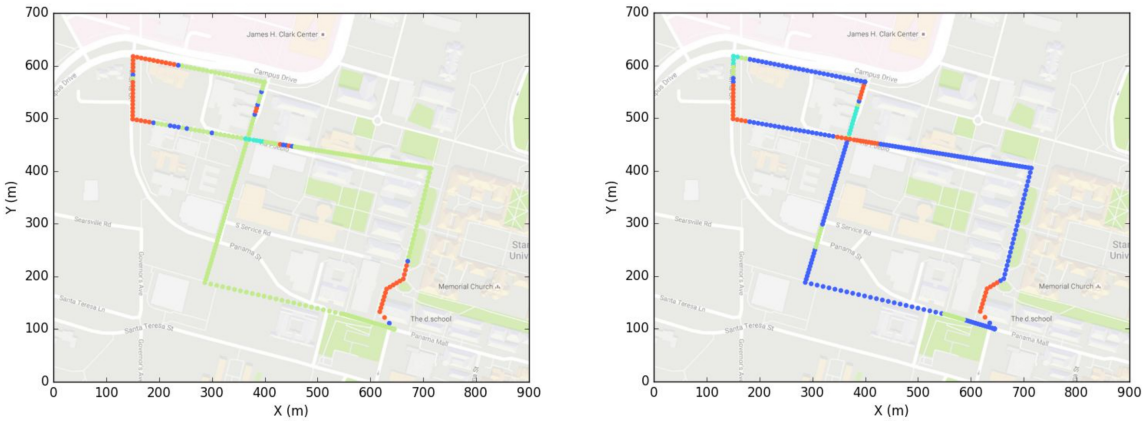


Figure 3: Continuous wavelet transform (CWT) scales obtained over the spatial dimension at two different time stamps. The colored rectangles highlight the different types of seismic signature that were detected by the clusters. [**CR**] fantine1/. cwt1,cwt2



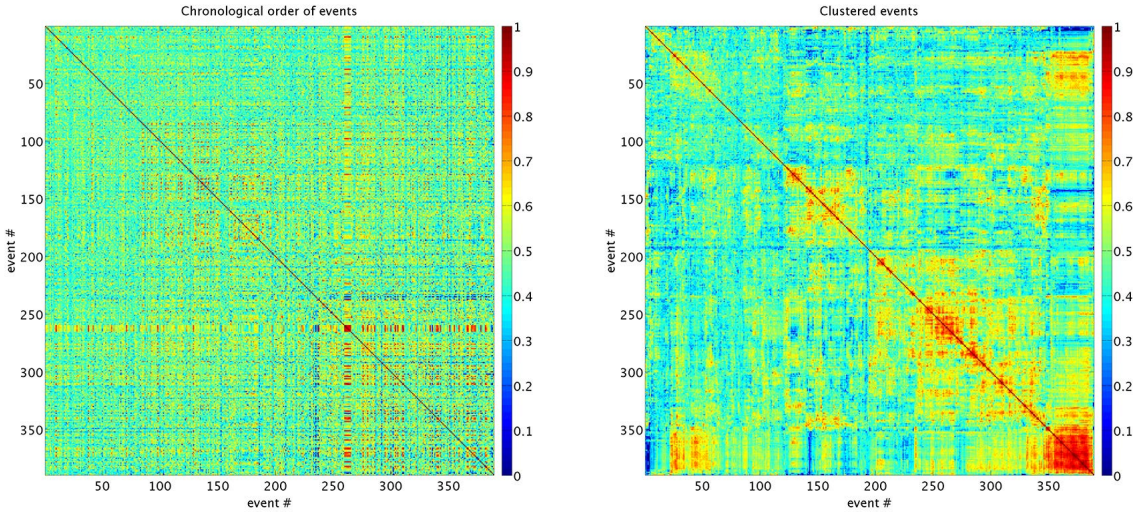
(a) (b)

Figure 4: Clustering results projected over the array's geometry at two different time stamps. The time stamps correspond to those from Figure 3. Dark blue represents the incoherent ambient noise field; Light green corresponds to laser noise, responsible for disruptions with large spatial extent over the array; Cyan corresponds to traffic noise. Note how the clustering algorithm captured moving sources such as cars without being provided any information related to the geometry of the array; Orange indicates other sources of coherent noise of medium amplitude, mostly localized on the North-West part of the fiber optic array. These coherent noise sources could be linked to remote traffic noise (from cars that are not running exactly above the fiber), or construction noise (the area circled by the North-West loop is under construction). [**CR**] fantine1/. cluster1,cluster2

CLUSTERING OF HIGH-AMPLITUDE EVENTS

In this section, we narrow down our study to the high-amplitude events. We performed event detection by computing an edge density map using the gradient (Prewitt's operator) followed by convolution with a stencil of ones (200 samples long in time and 20 channels wide in space). We arbitrarily chose a conservative threshold, at the discretion of the interpreter, to select events that stand above the background noise.

Having selected potential events, we computed dot products between all possible pairs which gave us a measure of similarity used for clustering (Figure 5). We chose hierarchical trees to obtain clusters with varying number of events (James et al., 2013) (Figure 5). Although we assumed an input number of clusters between 10-20, there would generally be only a couple of clusters which captured most of the data variability. Figure 6 represents the four largest. Note visual consistency for inter-cluster events manifesting in the slope and the distribution of energy.



(a) (b)

Figure 5: (a) Similarity matrix between pre-selected high-energy events; (b) events from (a) sorted by a hierarchical clustering algorithm. Although we assumed an input number of clusters between 10-20, a small number of clusters already captures most of the data variability. [**CR**] fantine1/. matrix1,matrix2

EVENT DETECTION USING TEMPLATE MATCHING

We then attempted to identify active sources automatically using template matching. As it is not feasible to manually label all the data, we selectively hand-picked several classes of templates: one class contained events along a road which is only open to school buses, as shown in Figure 7 (a), (b) and (c); another class contained events near a road crossing where the fiber optic cable changes direction, as shown in Figure 7 (d), (e) and (f). The size of the templates is 100 seconds by 30 channels (240 m). We performed template matching using normalized cross-correlation (Lewis, 1995). The data is assigned to a particular class

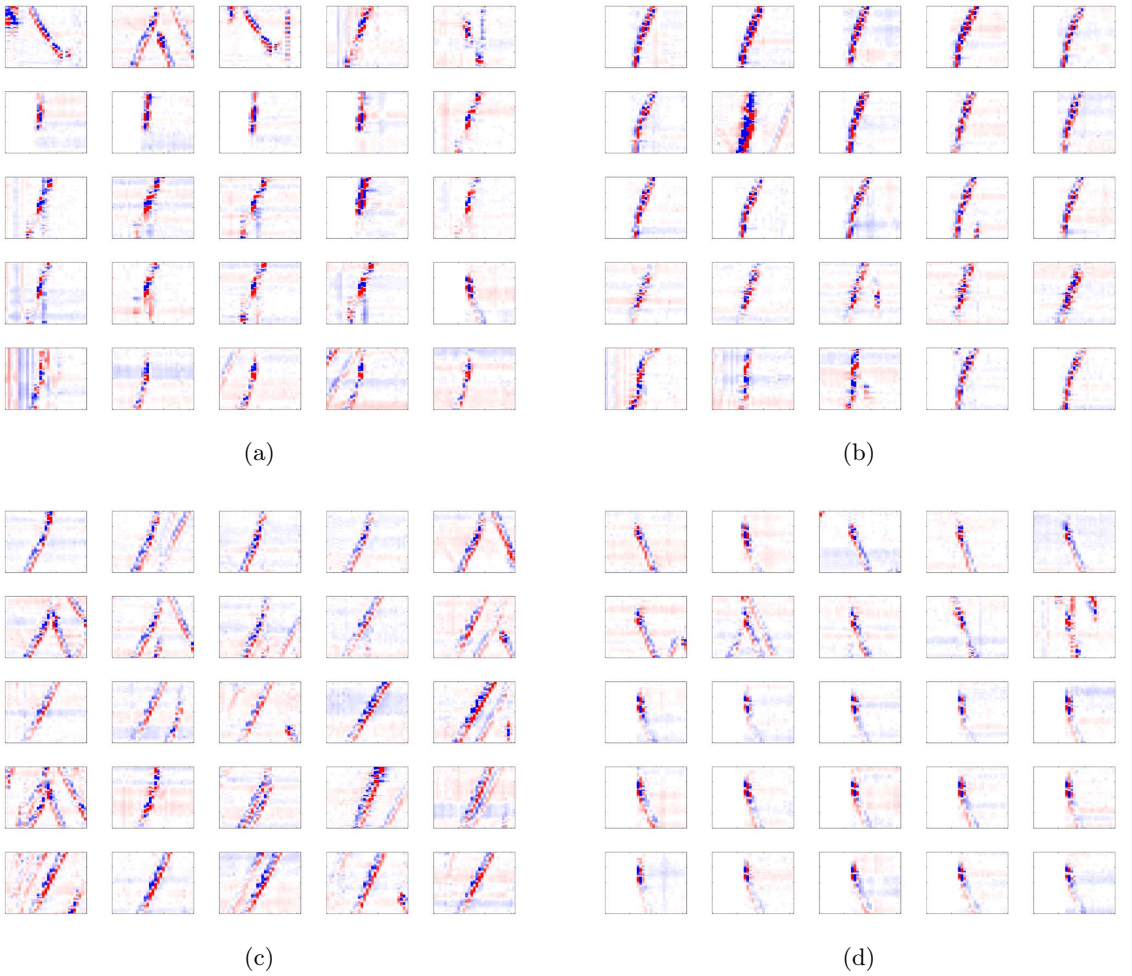


Figure 6: Example 25 events corresponding to each of the four largest identified clusters. Every subimage in this figure is 300 samples long in time (vertical axis) and 30 samples wide in space (horizontal axis). Note visual consistency for inter-cluster events manifesting in the slope and the distribution of energy. [**CR**] fantine1/. events1,events2,events3,events4

when the normalized cross-correlation with the templates from that class exceeds a certain threshold (0.5 in this study).

The detected sources are shown in the blue rectangles in Figure 8. The official schedule of the corresponding bus line is shown in the red rectangles. We see that using only a limited number of templates, we can detect events that approximately match the official bus schedule. The time discrepancy can be due to the bus not running exactly on time. The detected events that do not correspond to the schedule could be empty buses or delivery trucks.

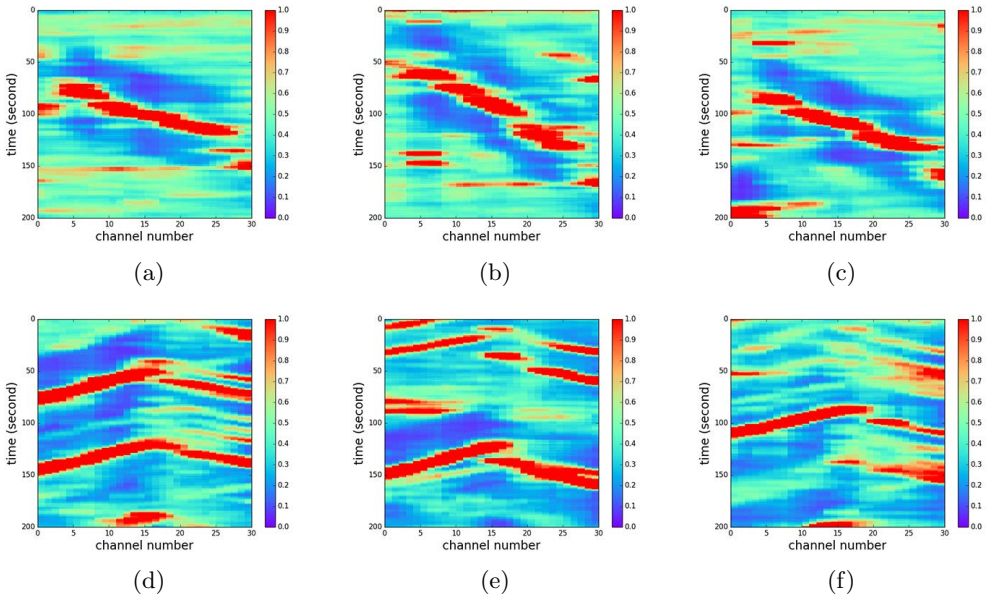
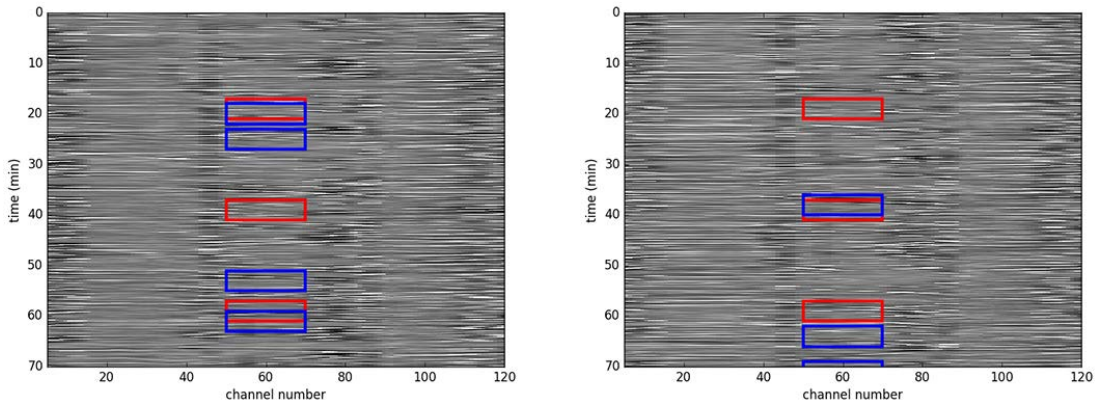


Figure 7: Manually picked events corresponding to vehicles at various locations on the Stanford University campus where only buses are allowed to circulate. [**CR**] fantine1/. template1,template2,template3,template4,template5,template6

CONCLUSIONS

In this study we employed unsupervised and supervised learning techniques to characterize seismic data recorded by a fiber optics array deployed underneath the Stanford University campus. First, we borrowed conventional processing techniques from the field of seismology and applied clustering algorithms to distinguish between different types of noise, automatically separating noise generated by cars from incoherent background noise without requiring any information related to the geometry of the array. Clustering performed on the high amplitude events showed us we could automatically sort traffic noise depending on the car's travel direction. Template matching allowed us to detect active sources and compare them to the local bus schedule.



(a) (b)

Figure 8: Blue rectangles indicate the times of the identified active sources, while red rectangles correspond to the official timetable of the local bus line. We see that using only a limited number of templates, we can detect events that approximately match the official bus schedule. The time discrepancy can be due to the bus not running exactly on time. The detected events that do not correspond to the schedule could be empty buses or delivery trucks. [**CR**] fantine1/. matching1,matching2

REFERENCES

Abadi, M., A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, et al., 2016, Tensorflow: Large-scale machine learning on heterogeneous distributed systems: arXiv preprint arXiv:1603.04467.

Ajo-Franklin*, J., N. Lindsey, T. Daley, B. Freifeld, M. Robertson, C. Ulrich, S. Dou, E. Martin, and A. Wagner, 2015, A field test of distributed acoustic sensing for ambient noise recording, *in* SEG Technical Program Expanded Abstracts 2015, 2620–2624, Society of Exploration Geophysicists.

Bakku, S. K., 2015, Fracture characterization from seismic measurements in a borehole: PhD thesis, Massachusetts Institute of Technology.

Bensen, G., M. Ritzwoller, M. Barmin, A. Levshin, F. Lin, M. Moschetti, N. Shapiro, and Y. Yang, 2007, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements: Geophysical Journal International, **169**, 1239–1260.

Collobert, R., S. Bengio, and J. Mariéthoz, 2002, Torch: a modular machine learning software library: Technical report, IDIAP.

Daley*, T., M. Robertson, B. Freifeld, D. White, D. Miller, F. Herkenhoff, and J. Cocker, 2014, Simultaneous acquisition of distributed acoustic sensing vsp with multi-mode and single-mode fiber optic cables and 3-component geophones at the aquistore co2 storage site, *in* SEG Technical Program Expanded Abstracts 2014, 5014–5018, Society of Exploration Geophysicists.

Daley, T. M., B. M. Freifeld, J. Ajo-Franklin, S. Dou, R. Pevzner, V. Shulakova, S. Kashikar, D. E. Miller, J. Goetz, J. Henninges, et al., 2013, Field testing of fiber-optic distributed acoustic sensing (das) for subsurface seismic monitoring: The Leading Edge, **32**, 699–706.

Daskalakis, E., C. Evangelidis, J. Garnier, N. Melis, G. Papanicolaou, and C. Tsogka, 2016, Robust seismic velocity change estimation using ambient noise recordings: Geophysical

Journal International, **205**, 1926–1936.
Fisher, W. D., T. K. Camp, and V. V. Krzhizhanovskaya, 2016, Anomaly detection in earth dam and levee passive seismic data using support vector machines and automatic feature selection: Journal of Computational Science.
Girard, A. and J. Shragge, 2016, Extracting body waves from ambient seismic recordings, *in* SEG Technical Program Expanded Abstracts 2016, 2715–2719, Society of Exploration Geophysicists.
James, G., D. Witten, T. Hastie, and R. Tibshirani, 2013, An introduction to statistical learning, volume **6**: Springer.
Jia, Y., E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, 2014, Caffe: Convolutional architecture for fast feature embedding: Proceedings of the ACM International Conference on Multimedia, 675–678.
Köhler, A., M. Ohrnberger, and F. Scherbaum, 2009, Unsupervised feature selection and general pattern discovery using self-organizing maps for gaining insights into the nature of seismic wavefields: Computers & Geosciences, **35**, 1757–1767.
Lewis, J., 1995, Fast normalized cross-correlation: Vision Interface.
Liao, T. W., 2005, Clustering of time series data: Pattern recognition, **38**, 1857–1874.
Lin, F.-C., M. P. Moschetti, and M. H. Ritzwoller, 2008, Surface wave tomography of the western united states from ambient seismic noise: Rayleigh and love wave phase velocity maps: Geophysical Journal International, **173**, 281–298.
Lobkis, O. I. and R. L. Weaver, 2001, On the emergence of the greens function in the correlations of a diffuse field: The Journal of the Acoustical Society of America, **110**, 3011–3017.
Mallat, S., 1999, A wavelet tour of signal processing: Academic press.
Martin*, E., J. Ajo-Franklin, S. Dou, N. Lindsey, T. Daley, B. Freifeld, M. Robertson, A. Wagner, and C. Ulrich, 2015, Interferometry of ambient noise from a trenched distributed acoustic sensing array, *in* SEG Technical Program Expanded Abstracts 2015, 2445–2450, Society of Exploration Geophysicists.
Martin, E., B. Biondi, M. Karrenbach, and S. Cole, 2017, Continuous subsurface monitoring by passive seismic with distributed acoustic sensors-the stanford array experiment: Presented at the First EAGE Workshop on Practical Reservoir Monitoring.
Martin, E., N. Lindsey, S. Dou, J. Ajo-Franklin, T. Daley, B. Freifeld, M. Robertson, C. Ulrich, A. Wagner, and K. Bjella, 2016, Interferometry of a roadside das array in fairbanks, ak, *in* SEG Technical Program Expanded Abstracts 2016, 2725–2729, Society of Exploration Geophysicists.
Mateeva, A., J. Lopez, H. Potters, J. Mestayer, B. Cox, D. Kiyashchenko, P. Wills, S. Grandi, K. Hornman, B. Kuvshinov, et al., 2014, Distributed acoustic sensing for reservoir monitoring with vertical seismic profiling: Geophysical Prospecting, **62**, 679–692.
Nakata, N., J. P. Chang, J. F. Lawrence, and P. Boué, 2015, Body wave extraction and tomography at long beach, california, with ambient-noise interferometry: Journal of Geophysical Research: Solid Earth, **120**, 1159–1173.
Nakata, N., R. Snieder, T. Tsuji, K. Larner, and T. Matsuoka, 2011, Shear wave imaging from traffic noise using seismic interferometry by cross-coherence: Geophysics, **76**, SA97–SA106.
Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al., 2011, Scikit-learn: Machine learning in python: The Journal of Machine Learning Research, **12**, 2825–2830.
Sinha, S., P. S. Routh, P. D. Anno, and J. P. Castagna, 2005, Spectral decomposition of

seismic data with continuous-wavelet transform: Geophysics, **70**, P19–P25.

Wapenaar, K., D. Draganov, R. Snieder, X. Campman, and A. Verdel, 2010a, Tutorial on seismic interferometry: Part 1basic principles and applications: Geophysics, **75**, 75A195–75A209.

Wapenaar, K., E. Slob, R. Snieder, and A. Curtis, 2010b, Tutorial on seismic interferometry: Part 2underlying theory and new advances: Geophysics, **75**, 75A211–75A227.

Webster, P., J. Wall, C. Perkins, and M. Molenaar, 2013, Micro-seismic detection using distributed acoustic sensing, *in* SEG Technical Program Expanded Abstracts 2013, 2459–2463, Society of Exploration Geophysicists.

Yoon, C. E., O. OReilly, K. J. Bergen, and G. C. Beroza, 2015, Earthquake detection through computationally efficient similarity search: Science advances, **1**, e1501057.

Zeng, X., C. Thurber, H. Wang, D. Fratta, E. Matzel, and P. Team, 2016, High-resolution shallow structure revealed with ambient noise tomography on a dense array.

Zhan, Z., V. C. Tsai, and R. W. Clayton, 2013, Spurious velocity changes caused by temporal variations in ambient noise frequency content: Geophysical Journal International, ggt170.

Higher Order Principal Component Analysis for Identifying AVO Anomalies

Stuart Farris and Stewart Levin

ABSTRACT

We use principal component analysis (PCA) to identify amplitude versus offset (AVO) anomalies in a legacy 2D seismic data set. The traditional methods of AVO analysis rely on inversion techniques that require simplified models of the subsurface. Simple models of the earth make assumptions that may lead to incorrect hydrocarbon identification. By making no assumptions about the earth model or how waves propagate in the model, PCA could provide a more robust technique for finding AVO anomalies.

INTRODUCTION

The underlying principles of AVO analysis are presented in Ostrander (1984). In essence, rock formation boundaries with highly contrasting Poisson's ratios can create "bright spots" in seismic offset gathers. These anomalies often indicate trapped hydrocarbon gas which may overlie oil reserves. The promise of having oil reserves light up on seismic gathers has lead the industry to invest significant effort into developing robust detection techniques (Keys and Foster, 1996). But, the propagation of seismic waves through the earth is too complex to model perfectly. Consequently, models make assumptions regarding elasticity, anisotropy, attenuation, and other physical properties (Hill, 2015). These physical assumptions in the modeling process can lead to false-positive or missed reservoir identifications.

Bougher and Herrmann (2016) presented a statistical method for AVO anomaly identification that requires no rock property assumptions. The method transforms migrated common midpoint gathers (CMPs) into their principal components and correctly identifies gas AVO anomalies. Their abstract shows, in synthetic 2 dimensional examples, this PCA method can correctly identify gas AVO anomalies with results comparable to traditional AVO inversion techniques.

We attempt to expand upon their method in multiple ways. Bougher and Herrmann (2016) only examined the first and second principal components for anomalies. We examine higher order principal components with hopes to identify more complex AVO anomalies. Furthermore, we conduct our study on the Mobil Viking Graben data set. This allows us to test PCA on real data which has already been highly scrutinized for AVO anomalies. The legacy AVO results will provide a benchmark for our new method.

METHOD

We use higher order principal components of CMP gathers to identify AVO anomalies.

The goal of PCA is to linearly transform data into a coordinate system where the projection of the data into that system has the highest variance on the first axis, the second highest variance on the second axis, and so on. These axis are known as the principal components. By keeping those components that are most significant, we reduce the number of features in a data set while maximizing the variance of those features. For this application we take single, prestack-migrated CMP gathers, and transform them into their principal components. Therefore, we transform some CMP gather $D \in \mathbb{R}^{i \times j}$, with i time samples and j offsets, into $\tilde{D} \in \mathbb{R}^{i \times j}$, with i time samples and j principal components. Each time sample in a CMP now has values for each principal component. By cross-plotting principal components, we can find time samples that do not match the time vs. offset variance trend of all the time samples. It was shown by Bougher and Herrmann (2016) that outliers in crossplots of the first and second principal components can correspond to AVO anomalies. The crossplot outliers can be visualized by mapping the time samples into image space for each CMP gather.

We extend this idea by cross-plotting higher principal components and analyzing their outliers. We hypothesize that they correspond to more subtle AVO anomalies resulting from anisotropy or other local rock property changes. The more difficult task will be linking newly identified anomalies to the physical phenomenon that caused them. Modeling rock property changes in synthetic data and observing the effects on the principal components will be the focus of future work with this topic.

DATA

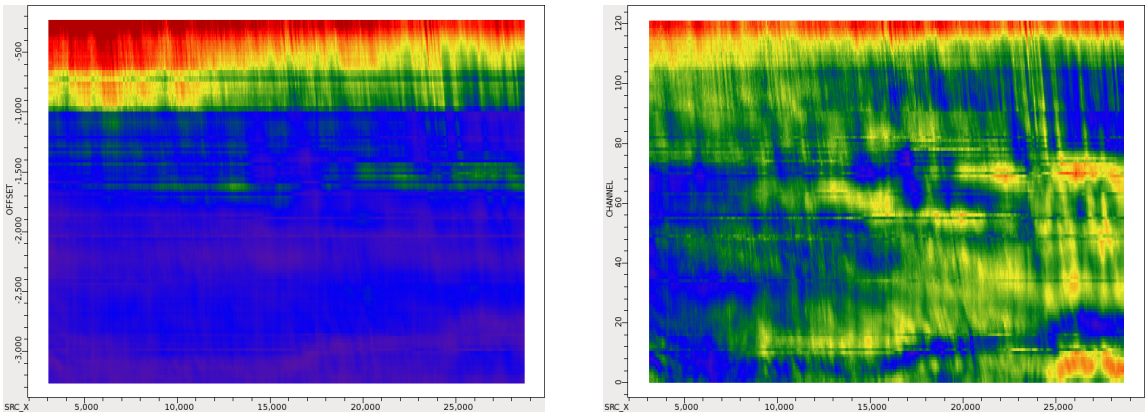
Here we evaluate our method using the public Viking Graben data set collected by Mobil and provided by the Society of Exploration Geophysicists (SEG). This data set was released for industry and academic institutions to test "new ideas about the use of seismic inversion or AVO methods for detecting hydrocarbons" (Keys and Foster, 1996). We continue the investigation of AVO methods on this data set with our new PCA technique. The Stanford Exploration Project (SEP) is very familiar with this data set, namely in reports 80 through 84. With the legacy results from SEG and SEP reports, we will be able to assess the feasibility of our PCA anomaly detection.

DATA PROCESSING

To prepare the Viking Graben data for PCA, preprocessing steps must be considered. Notice that these steps do not preserve amplitudes for quantitative AVO analysis.

Amplitude variations caused by inconsistent source strengths and hydrophone sensitivities can distort AVO results Berlioux and Lumley (1994). To mitigate this issue we balance the amplitudes of traces for both shot and receiver variations using OpenCPS (Open Geophysical, Inc., 2016). To balance shot strengths, we compute the root-mean-square (RMS) for all shot gathers. Then we subtract the amplitudes of a particular shot from the average RMS of all shots. Likewise, we compute the RMS of all receiver gathers then subtract amplitudes of each receiver from the average RMS of all receiver gathers. This ensures no individual trace can bias the average shot or receiver RMS value. Figure 1 illustrates the RMS amplitudes for each trace before and after amplitude balancing. Most of the streaks in

the vertical direction, caused by source strength variation, and horizontal direction, caused by hydrophone sensitivity variation, have been removed.



(a) (b)

Figure 1: RMS amplitude values plotted in shot vs offset domain using a) original data b) data after shot and receiver strength correction [**CR**] sfarris1/. originalAmp,correctedAmp

Our PCA method assumes flattened, migrated CMP gathers. To obtain these gathers, we use the prestack Kirchhoff migration in OpenCPS with the velocity model found in (Ecker and Lumley, 2001) illustrated in Figure 2. This migration method is advantageous since it preserves amplitudes reasonably well and keeps our data in the CMP domain.

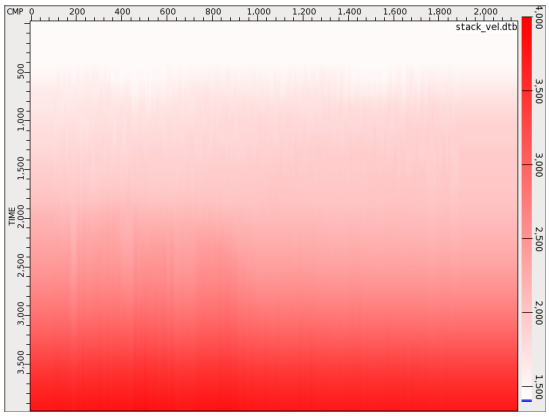


Figure 2: Stacking velocities used for SMRE and prestack Kirchhoff migration [**CR**] sfarris1/. eckerVel

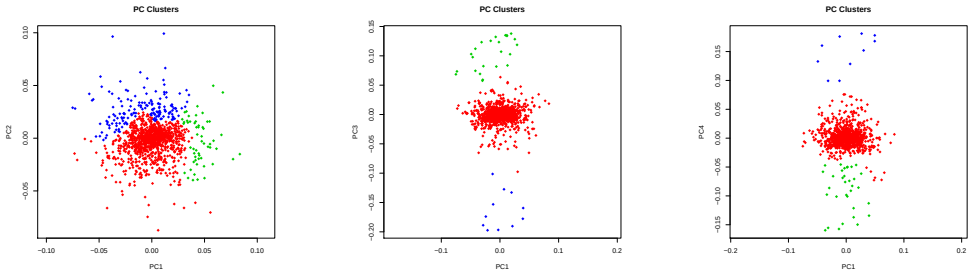
We also attempt surface-related-multiple-elimination (SRME) using the workflow outlined in OpenCPS. It is unclear if multiples affect our PCA method. Intuitively, it seems multiples contaminate migrated gathers and distort the amplitude variations we wish to identify. Regardless, we will attempt our method with and without SRME.

Finally, we are ready for PCA. We use the PCA code in the University of Tokyo's C Clustering Library (Michiel de Hoon et al., 2013).

PRELIMINARY RESULTS

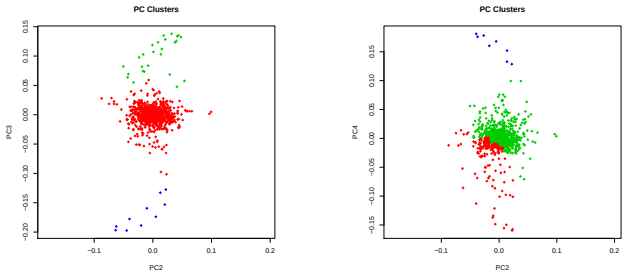
Figures 3, 4, and 5 show crossplots of various principal components for a single CMP gather. The CMP gather was selected from the center of the line to test the viability of this method before applying the technique to the full collection of gathers.

In order to automatically identify outliers in the crossplots, we used a hierarchical clustering technique to group samples based on euclidean distance in the principal component space. Here we have chosen to cluster the samples into three groups which are identified by their color in the principal component crossplots. The optimal number of clusters needs to be determined if we want to quickly identify anomalies in many CMP gathers.



(a) (b) (c)

Figure 3: Crossplots of principal components a) 1 vs. 2, b) 1 vs. 3, and c) 1 vs. 4 [**CR**] sfarris1/. PC1PC2cluster,PC1PC3cluster,PC1PC4cluster

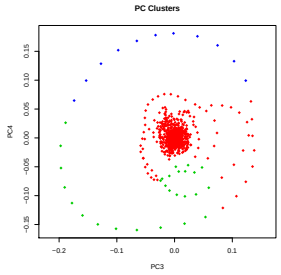


(a) (b)

Figure 4: Crossplots of principal components a) 2 vs. 3 and b) 2 vs. 4 [**CR**] sfarris1/. PC2PC3cluster,PC2PC4cluster

DISCUSSION

In Figure 3a we do not see outlier clusters that are as distinct as we expect. If there was a significant AVO anomaly, we would expect it to be a clear outlier in the crossplot of the principal components 1 vs 2, as shown in Bougher and Herrmann (2016). It is possible that the CMP gather selected for testing did not pass through a significant anomaly. It is also possible a step in processing prior to PCA dampened any existing anomalies.



(a)

Figure 5: Crossplot of principal components 3 vs. 4 [**CR**] sfarris1/. PC3PC4cluster

Higher order principal component crossplots show more promise. The clusters clearly separate in many of the crossplots shown in Figures 3b-4b. But, it is difficult to rely on these results until we can reproduce the results we expect in principal components 1 vs. 2.

WHAT'S NEXT

The next steps for this project are clear. First, we need to refine our preprocessing until we are confident that amplitudes are not compromised beyond the point of AVO analysis. Next, we need to replicate the PCA cross-plotting with a larger number of gathers that we know interact with AVO anomalies. We know these crossplots should show distinct outliers. Assuming successful testing on a subset of CMP gathers, we will apply PCA to all gathers. The outliers identified from all the gathers will then be mapped and visualized in the migrated image space. The results in the image space can then be compared to the legacy AVO analysis results found in the Viking Graben data.

The more difficult task will be matching anomalies found with PCA to the physical phenomenon that caused them. A synthetic model will have to be used to induce various physical anomalies, such as density, velocity, or anisotropy. With this model we can forward model a 2D seismic experiment and evaluate principal component anomalies.

ACKNOWLEDGMENTS

We must first acknowledge Mobil and SEG for providing the legacy Viking Graben data. Further, borrowing the PCA code from the University of Tokyo allowed us to avoid reinventing the wheel with regards to a statistical method that is well understood. Finally, we must thank all of the sponsors of the SEP for their continued funding and support.

REFERENCES

Berlioux, A. and D. Lumley, 1994, Amplitude balancing for AVO analysis: SEP-80: Stanford Exploration Project, 349–360.

Bougher, B. and F. Herrmann, 2016, AVA classification as an unsupervised machine-learning

problem, *in* SEG Technical Program Expanded Abstracts 2016, 553–556, Society of Exploration Geophysicists.

Ecker, C. and D. E. Lumley, 2001, Velocity analysis of the Mobil AVO dataset: SEP-82: Stanford Exploration Project.

Hill, S., 2015, Topics in Seismic Analysis.

Keys, R. G. and D. J. Foster, 1996, Comparison of Seismic Inversion Methods on a Single Real Data Set. Open file publications, No. no. 3: Society of Exploration Geophysicists.

Michiel de Hoon, Seiya Imoto, and Satoru Miyano, 2013, The C Clustering Library: Technical report, The University of Tokyo, Institute of Medical Science, Human Genome Center.

Open Geophysical, Inc., 2016, OpenCPS 3.3 Manual.

Ostrander, W., 1984, Plane-wave reflection coefficients for gas sands at nonnormal angles of incidence: Geophysics, **49**, 1637–1648.

A flexible out-of-core solver for linear/non-linear problems

Ettore Biondi and Guillaume Barnier

ABSTRACT

Based on the SEPlib non-linear solver library described previously, we implement an out-of-core Python-based generic solver for linear/non-linear inverse problems. We closely follow the object-oriented structure of the previous library and add new features in order to handle any programming language describing the application of user-defined operators. Our out-of-core feature addresses many treating memory-intensive inverse problems frequently encountered in geophysical applications. We demonstrate the new solver capability with simple linear problems and a non-linear acoustic isotropic full-waveform inversion (FWI) example.

INTRODUCTION

Geophysicists deal with inverse problems routinely. Many operations such as deconvolution or velocity estimation are inverse problems in disguise (Claerbout, 1992). Clearly it is advantageous to have at our disposal a generic solver capable of handling most of the inverse problems encountered. Métivier and Brossier (2016) present a generic non-linear solver written in *FORTRAN 90* based on reverse communication. Although their library offers many methods to solve a given inverse problem, it lacks the flexibility of being able to work with programming languages different than *FORTRAN*. This constraint is greatly restrictive. In fact, nowadays modeling/processing routines are written in many different programming languages and combined together to fully exploit the available computation architecture (e.g., core libraries written in *C++* or *CUDA* handled by a *Python* or *Shell* interface). For this reason we chose to implement a generic solver written in the free-ware scripting language *Python* that can be used for any inverse problem in which its core operators are coded with any programming language. We achieved this goal by using a *system-called operator* concept. In addition, the presented solver is implemented using an out-of-core philosophy, making it capable of handling memory-intensive inverse problems.

We first start by summarizing the object-orient structure of the solver, which is closely related to the one present in Almomin et al. (2015). Secondly, we explain the concept of a system-called operator that allows us to combine the solver with any other programming language. Finally, we describe the usage of the generic linear and non-linear solvers along with a few examples demonstrating their capabilities.

SOLVER STRUCTURE

The great flexibility of the object-oriented structure of the library described in Almomin et al. (2015) is also found in this new solver implementation. In fact, we decided to retain the same class structure:

- problem class: object describing the inverse problem to solver,
- solver class: object that takes problem class and solves the given problem,
- stepper class: object containing the method to use in case a line search is necessary,
- stopper class: object that quits the inversion in case any stopping criterion is met.

This implementation allows us to define any inverse problem and separate it from the method that we want to utilize for finding the problem's solution. The only requirement is that the problem has to be defined as an indirect problem (i.e., the solution has to be found using an objective function (Weglein et al., 2003)). Therefore, the problem class has to contain an objective function method. If a gradient-based method is employed, then linear forward and adjoint operator methods have to be present. We will explain this necessity when describing the generic non-linear solver. The stepper class is used whenever a line search algorithm is necessary to perform an optimization step during the inversion (e.g., whenever we are treating a non-linear problem). The definition of this class is avoidable when a linear problem has to be solved because step lengths are usually already embedded into the solver algorithm (e.g., the conjugate-gradient (CG) method (Aster et al., 2005)). Lastly, the stopper class just defines different stopping criteria that the user can use to stop the inversion. For a more detailed description of the *FORTRAN 2003* version see Almomin et al. (2015).

The typical inverse problem encountered in geophysics is the *L2*-norm inverse problem:

$$\phi(\mathbf{m}) = \frac{1}{2}\|\mathbf{f}(\mathbf{m}) - \mathbf{d}\|_2^2 = \frac{1}{2}\|\mathbf{r}(\mathbf{m})\|_2^2, \tag{1}$$

where $\mathbf{m}$ is the model vector, $\mathbf{f}(\mathbf{m})$ is the modeling operator, $\mathbf{d}$ is the observed data vector, $\mathbf{r}(\mathbf{m})$ is the residual vector, and $\phi(\mathbf{m})$ is the objective function to be minimized. The modeling operator can be either linear or non-linear. Obviously, the two problems are treated with different solvers. The two proposed solvers minimize the objective function shown in equation 1. In addition, in the current solver version, we also included the minimization of a *L2*-norm regularized problem defined as follows:

$$\phi_{reg}(\mathbf{m}) = \frac{1}{2}\|\mathbf{f}(\mathbf{m}) - \mathbf{d}\|_2^2 + \frac{\epsilon^2}{2}\|\mathbf{A}(\mathbf{m} - \mathbf{m}_{ref})\|_2^2 = \frac{1}{2}\|\mathbf{r}_d(\mathbf{m})\|_2^2 + \frac{\epsilon^2}{2}\|\mathbf{r}_m(\mathbf{m})\|_2^2, \tag{2}$$

where $\mathbf{A}$ is a linear regularization operator, and $\mathbf{m}_{ref}$ is a reference model.
In both cases, when a gradient-based method is used to optimize these objective function we have to compute their gradients as follows:

$$\nabla\phi(\mathbf{m}) = \left(\frac{\partial \mathbf{f}(\mathbf{m})}{\partial \mathbf{m}}\right)^* \mathbf{r}(\mathbf{m}), \tag{3}$$

$$\nabla\phi_{reg}(\mathbf{m}) = \left(\frac{\partial \mathbf{f}(\mathbf{m})}{\partial \mathbf{m}}\right)^* \mathbf{r}_d(\mathbf{m}) + \epsilon^2 \mathbf{A}^*\mathbf{A}\mathbf{r}_m(\mathbf{m}), \tag{4}$$

where * denotes the adjoint operation. We now turn our attention to the concept with which we are able to perform an out-of-core language-independent inversion.

THE SYSTEM-CALLED OPERATOR CONCEPT

From the previous equations we notice that the most complex operation to be performed is applying different operators (i.e., the modeling operator $\mathbf{f}(\mathbf{m})$, its linearization $\left(\frac{\partial \mathbf{f}(\mathbf{m})}{\partial \mathbf{m}}\right)$, and the linear regularization operator $\mathbf{A}$ forward and adjoint). In fact, in order to solve our inverse problem, we just need to tell the computer how to apply different operators to given vectors. To perform simple vector operations we use python-system calls to the standard SEPlib program *Solver_ops*; on the other hand, to apply the different required operators we define a new object called *operator*.

Whenever we create a new operator object we have to specify a series of terminal/shell commands that describes the application of the given operator to a provided input file and places the result of this operation into an output file. With this template the computer knows how to apply any supplied operator to a given vector and where the result will be placed. Therefore, to perform any inversion we have to just specify the solver the set of commands that we would run from terminal to apply the operators for our inverse problem into parameter files. As an example the following lines describe the application of a generic forward modeling operator that employs SEP-formatted files with the solver-operator definition:

```
#This is a comment
#Applying forward modeling operator
RUN: ./bin/forward.x < input.H par=Par/parfile.p > output.H
```

Note that in front of the command that has to be run it is placed the keyword `RUN:`. This syntax tells the operator object how to parse the parameter file that describes the given operator. In this case, we see that whenever the solver needs to apply the forward modeling operator $\mathbf{f}(\mathbf{m})$, it has to substitute `input.H` and `output.H` with the actual input-file and output-file names, respectively.

The operator object handles multiple lines when the operator is composed as a set of commands. In fact, a more complex operator could be the following:

```
#Applying foward modeling operator as a series of commands
RUN: ./bin/forward1.x < input.H par=Par/parfile.p > tmp1.H
RUN: ./bin/forward2.x < input.H par=Par/parfile.p > tmp2.H
#Combining the results of the previous commands
RUN: ./bin/combine.x file1=tmp1.H file2=tmp2.H output=output.H
#Cleaning temporary files
RUN: rm tmp1.H tmp2.H
```

In this case, the operator is composed of two different programs where their outputs are combined after their application. Note that the commands are run sequentially. Hence, whenever we start a command the subsequent one is going to be run after exit code of this process is returned. As we can see from this description, there is no assumption on what programming language the user employed to code their executables. In fact, with this definition of system-called operator, the python solver can handle operators written in any

programming language. In addition, because the solver uses filenames, it is not necessary to store any array inside the solver itself. The actual python class implementation is in file `operator_obj.py` contained in the folder `python_solver/` of the related *zip* file of this report.

GENERIC LINEAR SOLVER AND EXAMPLES

In this section we provide a brief explanation of how to use the generic linear solver. The name of the python script is `generic_linear_prob.py` and can be found in the folder `python_solver/`. If the script is run without any input it provides a documentation of all the possible parameters that can be used during a linear inversion. Before running any inversion it is necessary to run `make install` from the main folder, which compiles two necessary programs related to the solver, and add the screen-printed path to the user's environment variable *PYTHONPATH*. This operation has to be performed only once. A simple *L2*-norm linear inversion can be performed by calling the script in the following manner:

```
./python_solver/generic_linear_prob.py fwd_cmd_file=./Par/tmp_fwd.txt \
adj_cmd_file=./Par/tmp_adj.txt data=data.H init_model=init_model.H \
inv_model=inv_model.H suffix=_problem iteration_movies=obj,model,residual \
niter=10 log_file=Generic_lin_test.txt wrk_dir=inversion_test
```

Because we want to solve a non-regularized linear problem we just need to specify the linear forward and adjoint operators (i.e., $\mathbf{f}(\mathbf{m}) = \mathbf{F}\mathbf{m}$, and $\mathbf{F}^*$ necessary for the gradient computation). These two operators are provided in the command line with `fwd_cmd_file=./Par/tmp_fwd.txt` and `adj_cmd_file=./Par/tmp_adj.txt` parameters, respectively. All inversion files are going to be placed in the work directory `wrk_dir=inversion_test`. Along with the inverted model `inv_model=inv_model.H`, the solver can return separate different files containing the objective function value, inverted model, residual vector, and gradient as function of iteration number (see parameter `iteration_movies`). In the logfile `log_file=Generic_lin_test.txt` the user can find information about the inversion run, such as step lengths, elapsed time, and others.

Simple matrix tests

In this report's *Makefile* there are several simple matrix tests that can be run to understand the solver usage. Here, we report three simple matrix tests.

In the first test we want to minimize the objective function of equation 1 when the forward operator is a lower-triangular matrix of size 100 by 100 elements with linear CG method. The makerule to run this example is *Generic_lin_test*. Figure 1(a) shows the operator matrix used during the inversion. To ensure an invertible linear system, the main diagonal of this matrix is different than zero. The data vector is simply constant and equal to one. After running the solver for 50 iterations the initial objective function value decreases by more 99.4 percent (Figure 1(b)). Figure 1(c) displays the residual vector at the last iteration. We can see that the residual elements oscillate around zero and that the

first element is furthest from the true data. This behaviour is probably due to the matrix structure chosen in this test.

In our second linear test we minimize the objective function of equation 2 for the same matrix of our first example but adding a model-norm regularization term. As discussed by Claerbout (2014), there are different ways for estimating a balancing weight for the regularization term. We implement the solver so that it can provide an estimated ϵ that balances the first gradient $\mathbf{g}(\mathbf{m}_0)$ in the data space:

$$\epsilon = \|\mathbf{F}\mathbf{g}(\mathbf{m}_0)\| \,/\, \|\mathbf{A}\mathbf{g}(\mathbf{m}_0)\| \,. \tag{5}$$

With this ϵ definition we correctly balance the two objective function components by using a model (the gradient) that has proper model-scale size (i.e., numerical values). In fact, the step-length scaling is removed by the division of the two factors. Moreover, this approach prevents the denominator of this equation to be zero when the user-provided initial model is contained in the null space of the operator $\mathbf{A}$. In the described example we use a regularization weight of 1.0. Because of the regularization term, the solver cannot find a step length after 45 iteration and stops when the total objective function decreases by 98.79848 percent (Figure 2(a)). The behavior of the residual vector at the last iteration is similar to that of the previous test (Figure 2(b)). The make rule to run this example is *Generic_lin_reg_test*.

In the last linear test using matrices, we want to solve a symmetric negative definite system of equations. In this case the objective function to be solved is the following (Aster et al., 2005):

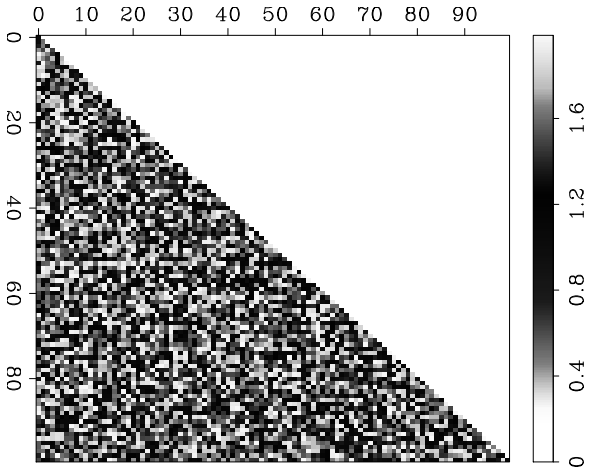
$$\phi_{sym}(\mathbf{m}) = \frac{1}{2}\mathbf{m}^*\mathbf{H}\mathbf{m} - \mathbf{m}^*\mathbf{d}, \tag{6}$$

where $\mathbf{H}$ is a symmetric positive or negative definite matrix. If the matrix is positive definite, this objective function is monotonically decreasing during the inversion. For this test we choose a second-order derivative operator (Figure 3(a)). Because of the boundary conditions employed the system is invertible. For data vector $\mathbf{d}$ we again use a constant value of one for all the elements. This choice results in finding a function whose second-order derivative is always constant and positive (i.e., a convex parabola). Indeed, this is the result that we obtain after running for 44 iterations the linear symmetric system CG Figure 3(b). To run this example the makerule is *Generic_lin_test_symmetric*.

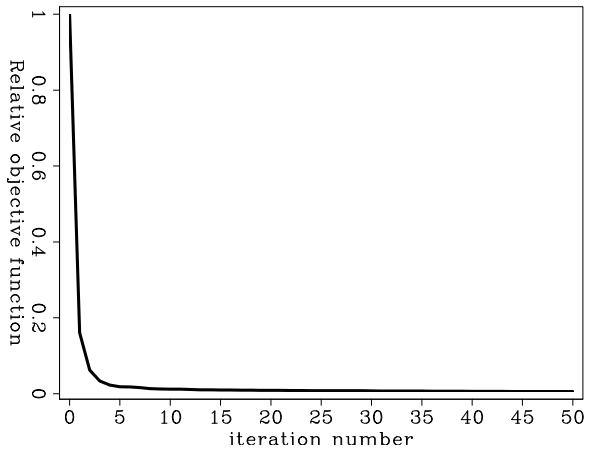
GENERIC NON-LINEAR SOLVER AND EXAMPLES

The same installation procedure applies for the generic non-linear solver. To run a non-linear inversion we ensure that whenever we are evaluating the gradient of the objective function we have to apply the adjoint linearized operator to the data residuals (see equations 3 and 4). In addition, we need to specify the forward linearized operator that is used to estimate the initial step length for any line search algorithm. Indeed, as input parameters the user has to provide three different operator-template parameter files. A simple call of the non-linear solver takes the following form:

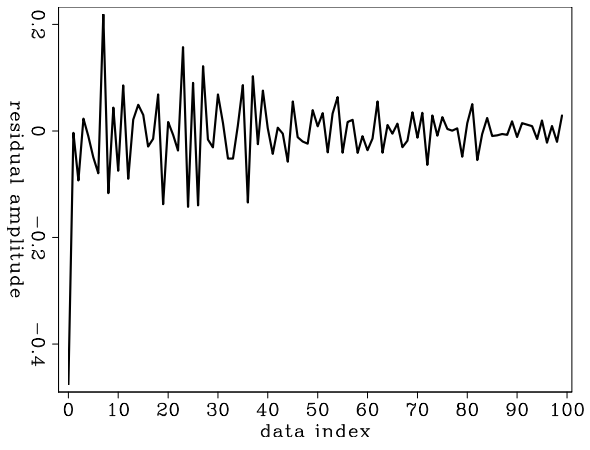
```
./python_solver/generic_non_linear_prob.py  \
```



(a)

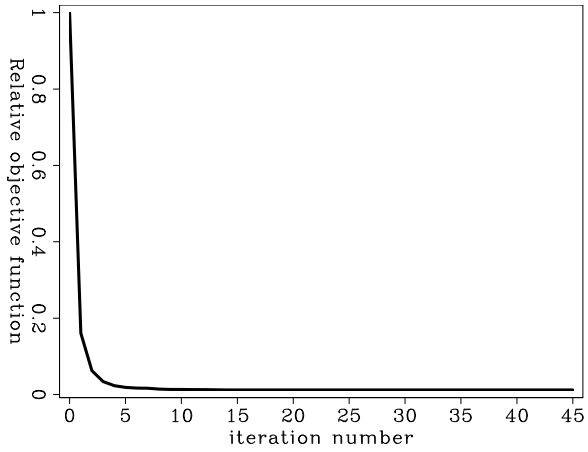


(b)

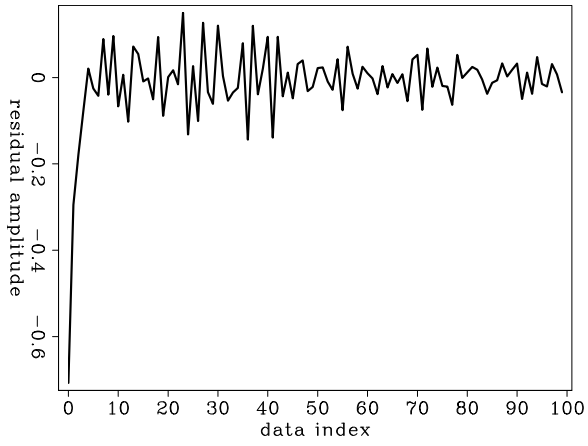


(c)

Figure 1: (a) Lower-triangular square matrix used in the first linear inversion test. (b) Relative objective function value as function of iteration number. (c) Residual vector at the last iteration. [**ER**] ettore1/. lower-matrix,test1-obj,test1-res



(a)



(b)

Figure 2: (a) Relative objective function value as function of iteration number. (b) Residual vector at the last iteration. Both plots belong to the second linear test described in the text. [**ER**] ettore1/. test2-obj-reg,test2-res-reg

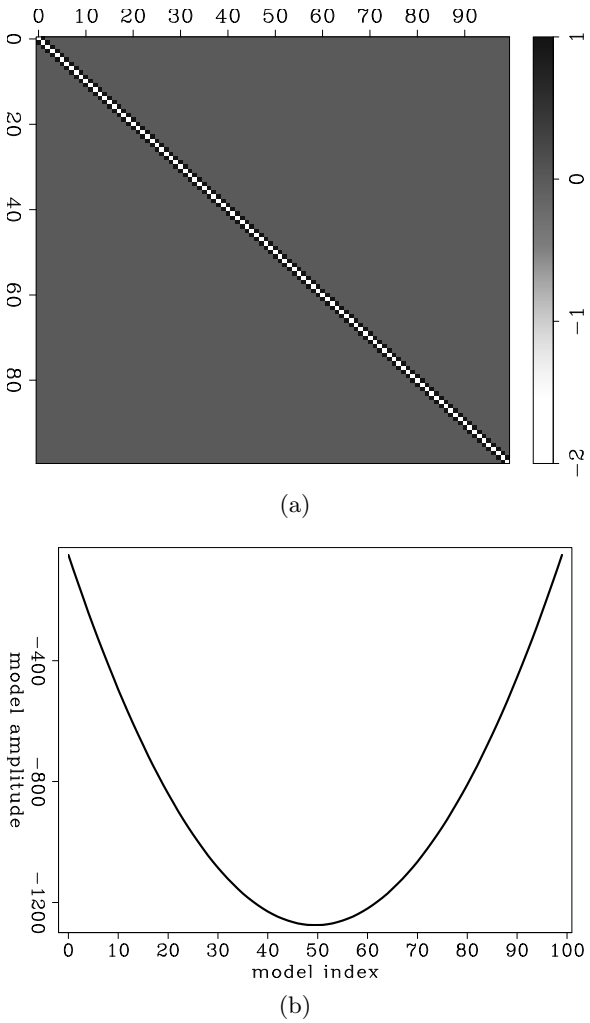


Figure 3: (a) Symmetric negative definite matrix used in the third linear inversion test. (b) Inverted model at the last iteration. [**ER**] ettore1/. sym-matrix,sym-model

```
fwd_nl_cmd_file=${P}/nl_fwd.txt fwd_cmd_file=${P}/lin_fwd.txt \
adj_cmd_file=${P}/lin_adj.txt data=data.H init_model=init_model.H \
inv_model=inv_model.H suffix=_nlproblem \
iteration_movies=obj,model,residual niter=10 \
log_file=inversion_log.txt wrk_dir=nl_inversion
```

Aside from this complication, the solver call is similar to the linear one. The required forward linearized operator `fwd_cmd_file` is necessary when an initial step-length guess has to be made during the line-search algorithm. In fact, a simple way to estimate the value of this guess is to linearize the modeling operator $\mathbf{f}(\mathbf{m})$ around the current model $\mathbf{m}_0$ and find the step length α that minimizes the given objective function along the current search direction $\Delta\mathbf{m}$. For instance, this procedure for the objective function in equation 1 would take the following form:

$$\phi(\mathbf{m}_0 + \alpha\Delta\mathbf{m}) = \frac{1}{2}\left\|\mathbf{f}(\mathbf{m}_0) + \alpha\frac{\partial\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}}\Delta\mathbf{m} - \mathbf{d}\right\|_2^2. \tag{7}$$

The linearization of the non-linear operator enables us to analytically compute the optimal step length by minimizing the derivative of the previous objective function with respect to α:

$$\frac{d\phi(\mathbf{m}_0 + \alpha\Delta\mathbf{m})}{d\alpha} = 0 => \alpha = -\frac{\mathbf{r}(\mathbf{m}_0)^*\frac{\partial\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}}\Delta\mathbf{m}}{\left\|\frac{\partial\mathbf{f}(\mathbf{m}_0)}{\partial\mathbf{m}}\Delta\mathbf{m}\right\|_2^2}, \tag{8}$$

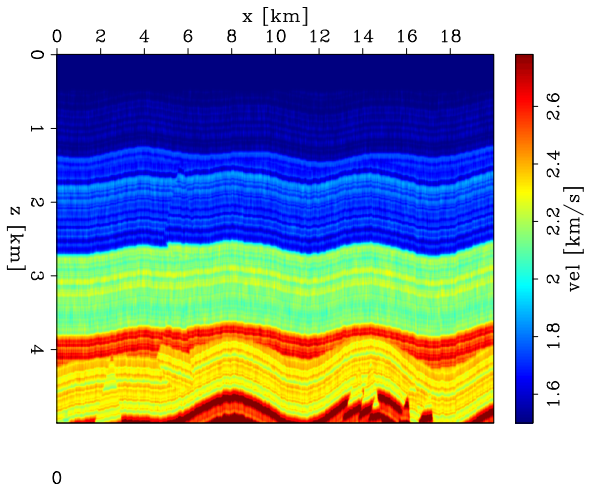
where the search direction is projected into the data space through the forward linearized operator.

In the next section we describe an application of the solver in a full waveform inversion problem. Although we do not show here the usage of any regularization term during the inversion, you can see its use in Barnier et al. (2017). In their examples the solver provides them with a balancing weight for their regularization term. In this case, the solver finds an ϵ that balances the two terms in the objective function of equation 2. If the initial model is in the null space of the regularization operator, the solver will perform a steepest-descent search and recompute the estimated weight.

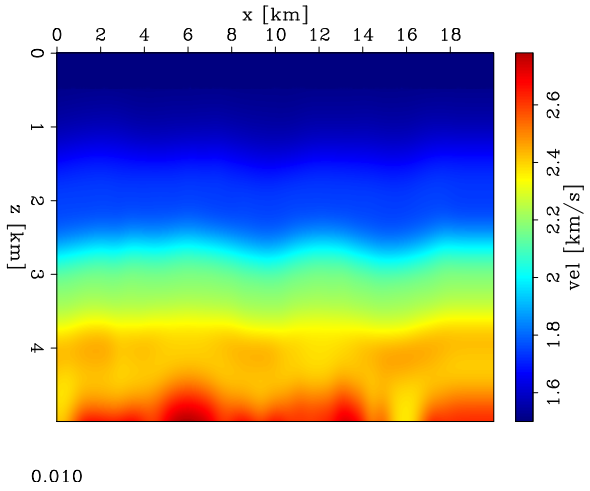
Full waveform inversion

We test the non-linear solver on a synthetic FWI example where we use an isotropic acoustic wave-equation modeling operator. Here, we minimize a simple $L2$-norm data difference (equation 1). Figure 4(a) shows the true velocity model used in this test. We place 200 shots and 2000 receivers spaced 100 m and 10 m at the surface, respectively. To generate the observed data we inject a Ricker wavelet of dominant frequency of 15 Hz. We then apply a bandpass filter to th observed data to a maximum frequency of 5 Hz. We invert this bandwidth only in this example. The starting model is generated by smoothing the true model (Figure 4(b)).

We apply the same preconditioning technique proposed by Biondi et al. (2017) in which an approximated Gauss-Newton Hessian inverse is estimated from model-space Hessian matrix applications. We run 14 iterations of a non-linear conjugate-gradient method (Fletcher



(a)



(b)

Figure 4: True (a) and initial (b) velocity models used in the described FWI example. [**ER**] ettore1/. fwi-true-model,fwi-start-model

and Reeves, 1964). From Figure 5 we clearly distinguish a difference when we precondition the gradient. In fact, this simple approach improves the convergence rate of the algorithm. In terms of inverted models, we also see a noticeable difference (Figure 6). We observe that the preconditioned inverted model presents more details, especially, in the deeper portion. In this example, the approximated Hessian inverse matrix provides an amplitude-balancing effect when applied to the computed gradient during the inversion.

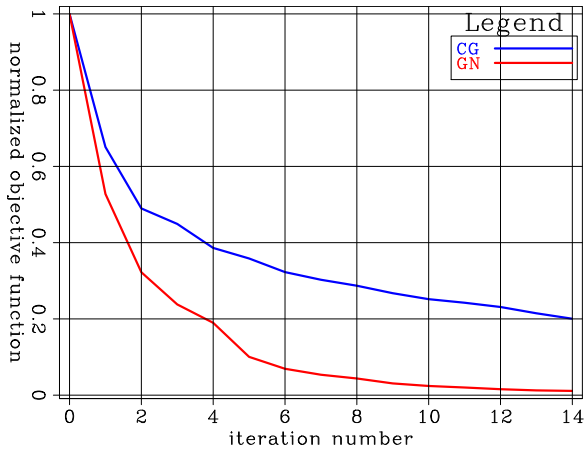


Figure 5: Relative objective function comparison between conjugate-gradient method using an approximated Gauss-Newton Hessian inverse as preconditioner (red curve) and without it (blue curve). [**CR**] ettore1/. fwi-obj

CONCLUSIONS AND FUTURE DIRECTIONS

We present a flexible out-of-core Python-based linear/non-linear solver library that uses of system-called operators to compute applications of these to given vectors. This feature provides the user a solver in which inverse-problem operators can be written with any programming language. Because the solver runs storing only filenames it has the potential of handling memory-demanding problems. We demonstrate its capabilities on different linear and non-linear inverse problems. In all these tests we show the convergence of the solver to a reasonable optimal model.

Currently, the solver employs only linear and non-linear CG methods, and different line-search algorithms are implemented within the non-linear optimization (i.e., sampler, parabolic, and linear steppers). In the future, memory-limited BFGS method is going to be added into the non-linear solver frame. Also, truncated Gauss-Newton and Newton steppers are going to be fully deployed into the solver. Finally, to overcome the limiting factor of applicability to SEP-formatted file we are going to use the *Generic IO* library proposed by Clapp (2017).

REFERENCES

Almomin, A., E. Biondi, Y. Ma, K. Ruan, J. Jennings, R. Clapp, M. Maharramov, and A. Cabrales-Vargas, 2015, SEPLib nonlinear solver library – Manual: SEP-Report, **160**,

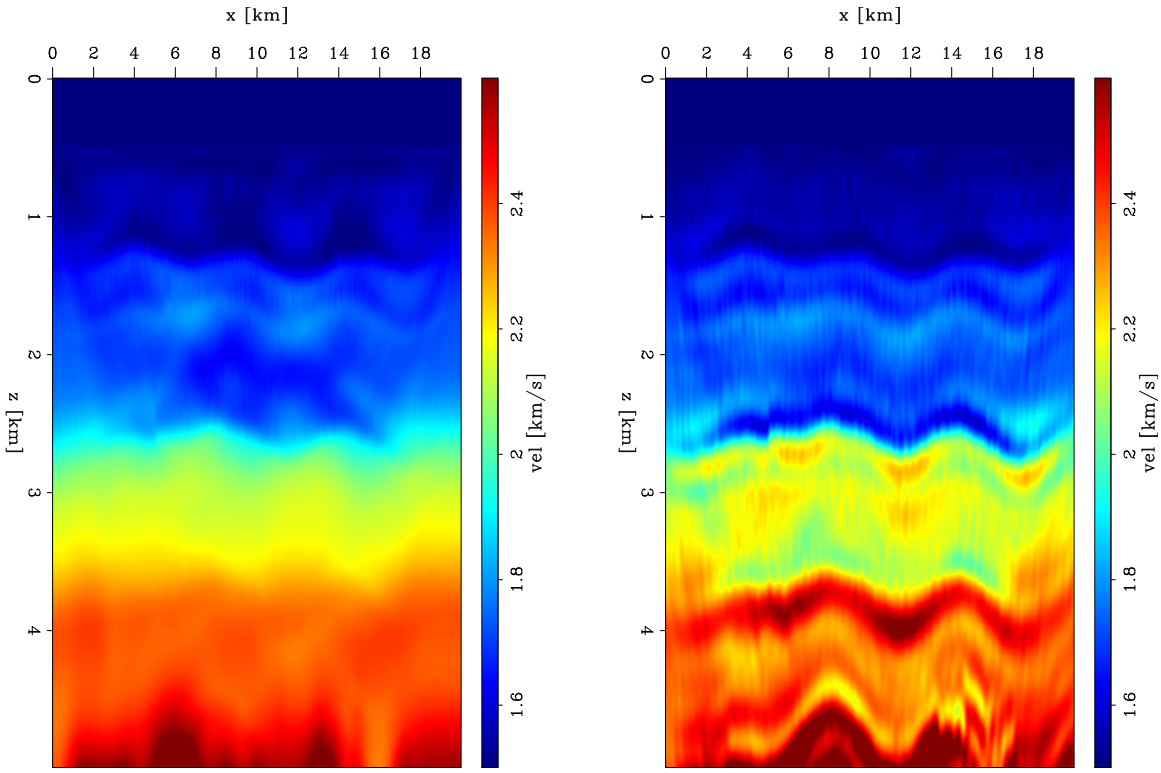


Figure 6: Inversion result comparison. Left panel: inverted model without preconditioner. Right panel: inverted model using an approximated Hessian inverse as preconditioner. [**CR**] ettore1/. fwi-inv-models

39–71.

Aster, R., B. Borchers, and C. Thurber, 2005, Parameter Estimation and Inverse Problems: Elsevier.

Barnier, G., E. Biondi, and B. Biondi, 2017, Applying TFWI for imaging under complex overburden: SEP-Report, **168**.

Biondi, E., G. Barnier, and B. Biondi, 2017, Preconditioned elastic FWI: SEP-Report, **168**.

Claerbout, J., 2014, Geophysical image estimation by example: Lulu. com.

Claerbout, J. F., 1992, Earth soundings analysis: Processing versus inversion, volume **6**: Blackwell Scientific Publications, Cambridge, Massachusetts, USA.

Clapp, R., 2017, Generic IO: SEP-Report, **168**.

Fletcher, R. and C. M. Reeves, 1964, Function minimization by conjugate gradients: The computer journal, **7**, 149–154.

Métivier, L. and R. Brossier, 2016, The SEISCOPE optimization toolbox: A large-scale nonlinear optimization library based on reverse communication: Geophysics, **81**, F1–F15.

Weglein, A. B., F. V. Araújo, P. M. Carvalho, R. H. Stolt, K. H. Matson, R. T. Coates, D. Corrigan, D. J. Foster, S. A. Shaw, and H. Zhang, 2003, Inverse scattering series and seismic exploration: Inverse problems, **19**, R27.

make, schmake: CMake

Stewart A. Levin and Robert G. Clapp

ABSTRACT

This year SEP has switched from GNU *automake/autoconf/libtool* builds of our public source tree to *CMake*. Here we highlights the advantages, and annoyances, of using CMake and an assessment of whether it brings significant advantages to our current makefile-based seismic processing and inversion environment.

INTRODUCTION

Since as least as far back as 2010, our public SEPlib source distribution has been configured and built with the sophisticated GNU portability tools *automake*, *autoconf* and, more recently, *libtool*. These let the end user build and install our distributions by running `./configure` in the top level of the source tree followed by `make && make install`. Those tools, however, are unwieldy to configure and require fairly intimate knowledge of close to a thousand pages of documentation at `gnu.org/software` to deploy consistently and correctly. They also have a lot of overhead—it takes over an hour to build SEPlib from scratch.

About five years ago, co-author Clapp, began using *CMake* to build his own software, most notably his QuickTime-based 3D seismic viewer. *CMake* is designed around modern programming paradigms such as plugins (modules) and object-oriented (C++) features. Over that period he ported the SEPlib source build system to *CMake* as well. This proved remarkably easy and co-author Levin was able to make additions, deletions, and edits, despite substandard documentation, much more easily than with the GNU-based system. Additionally, a bottom up SEPlib build takes only ten minutes.

At this juncture, CMake-based SEPlib distribution is basically ready for prime time and will be the default build system for the forseeable future. This raises some interesting issues and questions.

ISSUES AND QUESTIONS

Warts *aka* novice usage

CMake is the joint development of a large community of developers, most volunteer, and, as such, contains a vast number of features that the ordinary user, and even more so the novice, would not typically be using. This panoply can be quite daunting and is only partially alleviated by the basic seven step tutorial provided at `http://cmake.org`. Such issues include:

- How to specify output and installation directories.

- How to specify compilers and linkers and their options.
- How to define and use CMake variables.
- Where to find such information about CMake.

For the more sophisticated, the book **Mastering CMake** (Martin and Hoffman, 2015) is intended to provide authoritative information on *CMake*. While it does so to a great extent, more than half of it consists of reprints of a tutorial and manual pages from `cmake.org` and its index is almost useless for searching. Frankly not worth the 50 buck price. We have mostly resorted to the alternative of finding examples elsewhere that are similar to what we want to accomplish and tailoring them to the task at hand.

Software distribution

One of the requirements for using CMake in our software distribution is the need to have the `cmake` utility on the receiving end. The GNU setup only required `/bin/sh` or an equivalent such as `bash` in order to run the `configure` script plus a reasonably modern version of `make` such as `gmake` to build and install the libraries and executables. In the near future we will use the *CPack* utility to build one or more binary installers for SEPlib. While this will not obviate the need for *CMake* for system development, it will allow precompiled executables to be conveniently distributed.

Fortunately, SEP research will soon be distributed via `docker` (Clapp, 2016) files which will contain `cmake` and can be run safely and securely on most any Intel-compatible computer architecture.

SEP data processing

SEP goes to great lengths to meet the gold standard of research reproducibility: "Burn, Build, View." That is, whenever possible, all results and intermediate files can be deleted, a processing sequence rerun automatically, and the resulting output identical to the results in the research report. This is done at the Stanford Exploration Project by using makefiles to specify inputs, dependencies, outputs, and command lines that transform inputs into outputs. For example, the makefile excerpt

```
output.H : input.H process.par
           $(SEP)/bin/Filter < input.H > output.H par=process.par
```

generates `output.H` from `input.H` using the SEP `Filter` program with keyword parameters given in the `process.par` parameter file. An equivalent in cmake syntax is

```
add_custom_command(
   OUTPUT output.H
   DEPENDS input.H process.par
   COMMAND ${SEP}/bin/Filter < input.H par=process.par > output.H
   )
```

which is somewhat more verbose. It is also subtly different. First, this is not executed by `cmake`. `cmake` simply converts it into a makefile and ancillary files. The user still has to type `make output.H` in order to run the commands. Second, the `output.H` file is marked as GENERATED which means that a subsequent `make clean` will remove it. This relieves the user from explicitly listing intermediate and final outputs individually in a makefile "burn" rule. Third, there is an automatic dependency of `output.H` on the cmake input (`CMakeLists.txt`) so that if the command or various inputs are modified, `cmake` is rerun automatically to create a new makefile.

DISCUSSION

Both theoretically and practically, albeit with rather inadequate reference documentation, *CMake* offers significant advantages in ease of use and efficiency in source code builds and installs. For reproducible research, there are advantages and disadvantages. On the plus side, cmake syntax is more verbose, hence more readable, than makefile syntax. On the minus side, the default GENERATED property presents the hazard of deleting intermediate or final outputs that are only conditionally reproducible whether for lack of access to its inputs or the cost of recomputing key results.

REFERENCES

Clapp, R. G., 2016, Reproducibility through containers : SEP-Report, **165**, 193–198.
Martin, K. and B. Hoffman, 2015, Mastering CMake Version 3.1: Kitware Inc.

Yes it can be unset, but the documentation is silent on this.

Facilitating code distribution: Docker and Generic IO

Robert G. Clapp

ABSTRACT

Disseminating research through computational software is of growing importance in academia. The growth of cloud computing and object-oriented programming has led to the development of new paradigms. We take advantage of two of these paradigms, containers and generalized IO, as a new way to distribute. software to SEP sponsors.

INTRODUCTION

For more than 25 years SEP has included the concept of reproducible research in its mission. SEP's initial approach was to use `Cake` (Claerbout, 1990; Claerbout and Nichols, 1990). Later Schwab and Schroeder (1995) outlined our replacement of `Cake` with the standard Unix tool `make`. Fomel and Hennenfent (2007) translated and extended these reproducible concepts using `Scons`.

Reproducible research serves multiple purposes such as integrity of research results and a way to distribute research ideas. All of the above approaches have significant challenges. First, they require the installation of a significant software stack. Clapp (2016) described initial efforts of overcoming this obstacle using containers. Containers can capture the entire software stack and guarantee reproducibility. The second mission, making code simpler to use by others, is complicated by the fact that the seismic community does not agree on a single software solution. Instead companies and universities use one of a multitude of software packages.

In this paper we attempt to deal with the issue of producing reproducible and interruptible software. To solve the first problem, we build on Clapp (2016), describing how dockers can be used to encapsulate a research paper. We also introduce a library, `genericIO`, a C++ and Fortran 2003 library which abstracts parameter handling and IO. We include three concrete versions of the library using SEPlib, Madgascar, and Java Script Object Notation (JSON). We also show several examples of applications using the library.

GENERIC-IO

Abstraction is a powerful concept in object-oriented programming. Software written to function with a generic abstract object will work with all objects that inherit from the abstract base class. An example of using abstraction is writing solvers. which separate the mathematics of inversion from the physics of operators. Parameter handling and input/output (IO) can also be thought of as being separate from the mathematics/science that application is performing. If all of the IO/parameter handling in a program is written in terms of an abstract base class, all that is required to change to a new processing system

is to write a concrete child class that inherits from the abstract base class. The library `genericIO` attempts to create an abstract interface for IO/parameter handling. The entire library is documented using `Doxygen`. In the next portion we describe the basic building blocks of the library in C++ then Fortran2008.

The `paramObj` abstract class reads parameters for the program. It has functions to return `bool, float, int, string`, and vectors of each given the parameter name. It allows the user specify a default if the parameter isn't found.

The `genericRegFile` base class can read the same type of parameters from a file. In addition it has the ability to write parameters to a file, and to read/write float, complex, and byte data to/from a file. The reading and writing can be in terms of stream (read the next X bytes), or a subcube of the data object.

The `genericIO` base class is able to return a `genericRegFile` object given a filename or tag, and how the file is going to be used (in,out, in/out). It also has the ability to return a `genericIrregFile` object which still in the design stage. In addition, it keeps a list of all the files it has opened and when the object goes out of scope closes all of its files.

The `IOModes` class knows about all of the base classes of `genericIO` along which the default IO mode (can be selected at compile time). A programmer asks for either the default or a specific IO mode.

Currently three concrete classes inherit from abstract base classes: `rsfIO, sepIO,` and `JSONIO`. YOU ONLY LIST 3 CLASSES. The first two are simply wrap the corresponding RSF and SEPlib function. The `JSONIO` class uses JSON file that describe command line and file parameters.

Rather than maintaining a parallel FOrtran2003 library for all potential IO types, we used a singleton design pattern to allow us to create a single instance of the `ioModes` object to be created. By writing C functions that all use the singleton ioModes object we were able to support the same IO mechanisms in Fortran while creating a single set of Fortran2003 classes.

An example of using the `genericIO` library to window a dataset can be found in the first appendix.

DOCKERS

Clapp (2016) discussed how containers, specifically dockers, were designed to streamline the development to application process. A docker can be thought of as a virtual machine with minimal overhead. Once a developer has built a docker that runs on their local machine, that same docker will run in a production environment. The latest versions of the `gcc` compilers docker is conventionally built through a `Dockerfile`. A `Dockerfile` starts with a `From` statement which is another docker to build from. For example, we might start from a compact version of an operating system such as a specific version of CentOS.

```
From centos:7
MAINTAINER Bob Clapp <bob@sep.stanford.edu>
```

We then line by line describe how to build our environmnet. Each line represents a different

layer in the docker filesytstem A delta of the filesytesm is calculated and stored. The final docker is the summation of all of these deltas. The storing of deltas means that we don't want to create files in one layer that we are going to erase in another layer. As a result we are careful to remove all temporary and unneeded files. For example, every time we run `yum` it creates a temporary cache. So after we install a list of packages with `yum` we clean up these cache files. As an example, imagine we want to build the latest version of the `gcc` compilers. There are several prequisites that we need to install.

```
RUN yum -y install gmp-devel mpfr-devel libmpc-devel;\\
   yum -y clean  all
RUN yum -y install make cpp wget gcc-c++ zlib-devel \\
curl wget; yum -y clean all
```

Once we've installed these packages we can download the latest versions of the compilers. Then configure, compile, install, and finally clean up. We need to do all of this on one line to minimize the delta size.

```
RUN mkdir   /tmp/gcc-build
RUN cd /tmp; \\
wget ftp://208.118.235.20/gnu/gcc/gcc-6.2.0/gcc-6.2.0.tar.gz;  \\
tar xf gcc-6.2.0.tar.gz; cd /tmp/gcc-build;  cd /tmp/gcc-build;\\
 ../gcc-6.2.0/configure --prefix=/usr/local --disable-multilib \\
  -with-system-zlib --enable-languages=c,c++,fortran;  \\
  cd /tmp/gcc-build; \\
  make  -j 12; cd /tmp/gcc-build make install;  rm -rf /tmp/gcc*
```

Dockers for reproducibility

The docker concept of layers is analogous to the way we currently do reproducible research. We start with our current environment, the base image in dockerspeak. We then, through make dependencies, build the software we need, and run a series of commands to build a result. There is a significant level of redundancy in a paper, let alone a report. On the report level, each paper depends on the same basic environment. At the paper level, results often depend on the same executables so those layers will get reused. To convert to using dockers for reproducibility, and to enable others to use our code more easily, we need to significantly improve the way we build our environment.

For years SEP has used autoconf/automake for building SEPlib. The autoconf/automake combination also attempted to configure makefiles to ease the building process for SEP reports. This process worked somewhat successfully inside SEP but proved challenging for outside users due to the incompleteness of the makefile configuration process, the use of additional libraries (such as FFTW), and the vagarities of autoconf. Levin and Clapp (2017) describes our reasoning for converting/methodology of converting from autoconf to `cmake` for building SEPlib. Simplifying the SEPlib building process was the necessary first step in our complete overhaul of how we define reproducible software. The same argument that supports the shift to `cmake` for building SEPlib is true for the software in individual report articles. By switching to `cmake` we make installation at outside locations much simpler.

REPRODUCIBILITY TEMPLATE

The key for internal and external reproducibility will be the combination of a gitlab server and a docker repository. Each reproducible figure will be its own docker. All report articles, stored as git repositories, will contain the Dockerfiles to build each reproducible figure and make commands to build the dockers from these images.

While the bar described above seems high, it should, once internalized, make doing reproducible research easier. We will begin by creating a docker containing the entire SEP environment at report time. The student will then start from this base image, clone from their git repository the source code for the paper and build it using cmake, creating a report article image. The Dockerfile's for each figure starts from this report article image and runs the commands to the create an individual figure.

Their are several advantages of this approach. First, the student can switch between systems with minimal effort as long as docker is installed. Second, guaranteed reproducibility over time. As long as docker exists the results are completely reproducible. If docker does not exist the Dockerfile contains a complete description of how it was created. Third, reprodcubility testing is simplified for both the author and the checker. Building a docker is not environment dependent. As long as it builds for the author, it will build for the checker.

FUTURE WORK

One of the longer term goals of this project is for all SEP research papers to be written in this form. To accomplish this goal the generic IO library will be extended to handle irregular dataasts beyond its current capabilities. In addition, we would like to have all of the labs for SEP classes to be done in this genericIO/docker format. An even longer term goal is to replace all of the IO in SEPlib using this genericIO framework. This would allow us to update the core of SEPLib to a more more modern design without affecting any applications.

CONCLUSION

In this paper we described a new way to perform research that is guaranteed to be easily reproducible while it the same time allowing someone to quickly incorporate research ideas into their software environment. We accomplish this goal by writing programs in terms of a generalized IO library, using git and cmake for code development and installation, and dockers to guarantee reproduciblity.

REFERENCES

Claerbout, J. F., 1990, Active documents and reproducible results: SEP-Report, **67**, 139–144.

Claerbout, J. F. and D. Nichols, 1990, Why active documents need cake: SEP-Report, **67**, 145–148.

Clapp, R. G., 2016, Reproducibility through containers : SEP-Report, **165**, 193–198.

Different locations for data/scratch space, etc. will change the docker invocation.

Fomel, S. and G. Hennenfent, 2007, Reproducible computational experiments using scons: Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on, IV–1257.

Levin, S. A. and R. G. Clapp, 2017, make, schmake: CMake : SEP-Report, **168**, 309–312.

Schwab, M. and J. Schroeder, 1995, Reproducible research documents using GNUmake: SEP-Report, **89**, 217–226.

A

```
#include<ioModes.h>
using namespace SEP;
std::shared_ptr<hypercube> calc_params(std::shared_ptr<paramObj> par,
  std::shared_ptr<hypercube> hyperIn, const std::vector<int> ng,
  std::vector<int>&nw,std::vector<int>&fw,std::vector<int>&jw){

  int ndim=(int)ng.size();
  std::vector<axis> axes;
  for(int idim=0,i=1; idim<ndim; idim++,i++){
    axis ain=hyperIn->getAxis(i);
    nw[idim]=par->getInt(std::string("n")+std::to_string(i),nw[idim]);
    fw[idim]=par->getInt(std::string("f")+std::to_string(i),0);
    jw[idim]=par->getInt(std::string("j")+std::to_string(i),1);
    fprintf(stderr,"i=%d n=%d f=%d j=%d \n",idim,nw[idim],fw[idim],jw[idim]);
    if(fw[idim] >= ng[idim]) par->error(std::string("Invalid f")+
     std::to_string(i));
    if(nw[idim] > ng[idim]) par->error(std::string("Invalid n")+
      std::to_string(i));
    if(jw[idim] ==-1) jw[idim]=1;
    if(nw[idim] ==-1) {
      if(fw[idim]==-1){
        fw[idim]==0;
      }
      nw[idim]=(ng[idim]-1-fw[idim])/jw[idim]+1;
    }
    else if(fw[idim]==-1) fw[idim]=0;
    if(ng[idim] < 1+ fw[idim] +jw[idim]*(nw[idim]-1)){
       fprintf(stderr,"f=%d j=%d n=%d ng=%d \n",fw[idim],jw[idim],nw[idim],ng[idim]);
       par->error(std::string("Invalid window parameters axis ")+std::to_string(i));
     }
    float o=fw[idim]*ain.d+ain.o;
    float d=jw[idim]*ain.d;
    axes.push_back(axis(nw[idim],o,d,ain.label));
  }
  for(int idim=ndim; idim<8; idim++){
    nw[idim]=1; fw[idim]=0; jw[idim]=1;
  }
  std::shared_ptr<hypercube> h(new hypercube(axes));
```

```
  return h;
}
int main(int argc, char **argv){

  ioModes modes(argc,argv);



  std::shared_ptr<genericIO>  io=modes.getDefaultIO();
  std::shared_ptr<paramObj> par=io->getParamObj();
  std::string in=par->getString(std::string("in"));

 std::string out=par->getString(std::string("out"));

 std::shared_ptr<genericRegFile> inp=io->getRegFile(in,usageIn);


 std::shared_ptr<hypercube> hyperIn=inp->getHyper();
 std::vector<int> ng=hyperIn->getNs();

 int ndim=(int)ng.size();

 if(ndim>8) par->error("Currently handle maximum of 8 dimensons");
 std::vector<int> nw(8,-1),fw(8,-1),jw(8,-1);

 std::shared_ptr<hypercube> hyperOut=calc_params(par,hyperIn,ng,nw,fw,jw);

 std::shared_ptr<genericRegFile> outp=io->getRegFile(out,usageOut);
 outp->setHyper(hyperOut);
 outp->writeDescription();

 int maxSize=par->getInt(std::string("maxsize"),8);
 long long maxS=(long long) maxSize*1024*1024*1024;

 std::vector<int> nloop(8,1);
 std::vector<int> fsend=fw,jsend=jw,nsend=nw;


 long long n123= hyperOut->getN123();
  int iaxis=ndim-1;
  while(n123 > maxS && iaxis>0){
    n123=n123/(long long) ng[iaxis];
    nloop[iaxis]=nw[iaxis];
    iaxis--;
  }
  for(int i=iaxis+1; i<8; i++){
     fsend[i]=0; jsend[i]=1; nsend[i]=1;
   }
```

```
  float *buf=new float[n123];
  for(int i8=0; i8 < nloop[7]; i8++){
    if(iaxis<7){fsend[7]=i8*jw[7]+fw[7]; nsend[7]=1;}
  for(int i7=0; i7 < nloop[6]; i7++){
    if(iaxis<6){fsend[6]=i7*jw[6]+fw[6]; nsend[6]=1;}
  for(int i6=0; i6 < nloop[5]; i6++){
    if(iaxis<5){fsend[5]=i6*jw[5]+fw[5]; nsend[5]=1;}
  for(int i5=0; i5 < nloop[4]; i5++){
    if(iaxis<4){fsend[4]=i5*jw[4]+fw[4]; nsend[4]=1;}
  for(int i4=0; i4 < nloop[3]; i4++){
    if(iaxis<3){fsend[3]=i4*jw[3]+fw[3];  nsend[3]=1;}
  for(int i3=0; i3 < nloop[2]; i3++){
  if(iaxis<2){fsend[2]=i3*jw[2]+fw[2];  nsend[2]=1;}
for(int i2=0; i2 < nloop[1]; i2++){
if(iaxis<1){ fsend[1]=i2*jw[1]+fw[1]; nsend[1]=1;}

inp->readFloatWindow(nsend,fsend,jsend,buf);
outp->writeFloatStream(buf,n123);
}}}}}}}
delete [] buf;

}
```

SEP PHONE DIRECTORY

Name	Phone	Login Name
Alves, Gustavo	723-9282	gcalves
Barnier, Guillaume	723-6006	gbarnier
Baroni, Claudia	723-5002	
Biondi, Biondo	723-1319	biondo
Biondi, Ettore	723-6006	ettore
Cabrales, Alejandro	723-9282	cabrales
Chang, Jason	724-4322	jasonpc
Claerbout, Jon	723-3717	jon
Clapp, Bob	725-1334	bob
Dahlke, Taylor	724-4322	taylor
Farris, Stuart	723-6006	stuart
Huot, Fantine	723-0463	fantine
Jennings, Joe	723-1250	joseph29
Le, Huy	723-1250	huyle
Levin, Stewart	726-1959	stew
Ma, Yinbin	723-0463	yinbin
Martin, Eileen	723-0463	ermartin
Sarkar, Rahul	725-1625	rahul

Dept fax number: (650) 725-7344

E-MAIL

Our Internet address is "*sep.stanford.edu*"; i.e., send Jon electronic mail with the address "*jon@sep.stanford.edu*".

WORLD-WIDE WEB SERVER INFORMATION

Sponsors who have provided us with their domain names are not prompted for a password when they access from work. If you are a sponsor, and would like to access our restricted area away from work, visit our website and attempt to download the material. You will then fill out a form, and we will send the username/password to your e-mail address at a sponsor company.

STEERING COMMITTEE MEMBERS, 2016-2017

Name	**Company**	**Tel #**	**E-Mail**
Raymond Abma	BP	(281)366-4604	abmar1@bp.com
Francois Audebert	TOTAL	(33)-6-1278-4195	francois.audebert@total.com
Biondo Biondi	SEP	(650)723-1319	biondo@sep.stanford.edu
Jon Claerbout	SEP	(650)723-3717	jon@sep.stanford.edu
Thomas Dickens	ExxonMobil	(713)431-6011	tom.a.dickens@exxonmobil.com
Carlo Fortini (Co-chair, 1st year)	ENI	+39 02-520-61599	carlo.fortini@eni.com
Faqi Liu	Hess	(823)231-9218	fliu@hess.com
Yi Luo	Saudi Aramco	–	yi.luo@aramco.com
Guojian Shan (Co-chair, 1st year)	Chevron	–	guos@chevron.com
Alejandro Valenciano	PGS	–	alejandro.valenciano@pgs.com
Kathy Zou	PGS	–	kathy.zou@pgs.com

Research Personnel

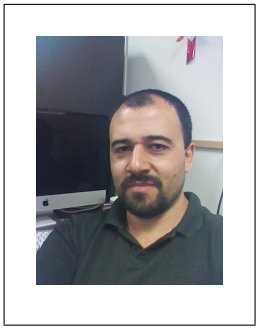
Gustavo Catao Alves graduated from the Federal University of Minas Gerais, Brazil, in 2004 with a BS in Physics. In 2006, he got a MS in Solid State Physics from the same Institution and was awarded Best Student Presenter on the 3rd Conference of the Brazilian Carbon Association. That same year he joined Petrobras, where he worked at the Company's research center for the next 7 years, on projects focused on seismic illumination studies, Reverse Time Migration and Full Waveform Inversion. In 2008, he received an Honorable Mention for Best Poster Paper at SEG's Annual Meeting, with the work entitled "Target oriented illumination analysis using wave equation". He joined Stanford Exploration Project in 2013 to pursue a PhD in Geophysics and is currently working on amplitude studies in long offset data. He is a member of SBGf, SEG, EAGE and AAPG.

Guillaume Barnier is from Nice, France. He graduated from Telecom Paristech in 2007 with a MSc in telecommunications and signal processing. After working as a fixed income trader for JP Morgan (London) from 2007 to 2010, he decided to change career path by joining the geophysics MSc program in Colorado School of Mines (2011), where he focused his research on seismoelectric coupling in poro elastic media. In September 2013, he joined the Stanford Exploration Project to pursue his Ph.D. Guillaume focuses on improving the image quality of subsalt layers when they have been damaged by an inaccurate top salt interface delineation.

Biondo L. Biondi is professor of Geophysics at Stanford University. Biondo graduated from Politecnico di Milano in 1984 and received an M.S. (1988) and a Ph.D. (1990) in geophysics from Stanford. He is co-director of the Stanford Exploration Project and of the Stanford Center for Computational Earth and Environmental Science. In 2004 the Society of Exploration Geophysicists (SEG) has honored Biondo with the Reginald Fessenden Award. Biondo recently published a book, 3-D Seismic Imaging, that is the first text book to introduce the theory of seismic imaging from the 3-D perspective. The book is published by SEG in the Investigations in Geophysics series. During 2007 gave a one-day short course in 28 cities around the world as the SEG/EAGE Distinguished Short Course Instructor (DISC) . He is a member of AGU, EAGE, SEG and SIAM.

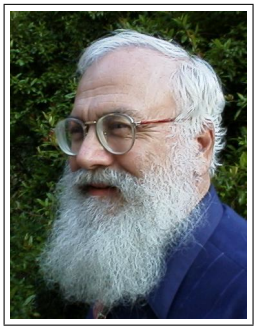
Ettore Biondi received a B.S. (2010) in geology from the University of Genoa, an M.S. (2012) in geophysics from the University of Pisa, and a diploma (2013) in computational chemistry from the Scuola Normale Superiore of Pisa. He spent almost two years as a research fellow at the University of Milan within the geophysics section of the Earth Science Department. In the summer of 2014, he joined the SEP and started his Ph.D. at the Stanford University. He is interested in multi-parameter inversion and wave-equation velocity estimation methods.

Alejandro Cabrales-Vargas obtained his Bachelors degree in Geophysics in the University of Mexico in 2002. He has being working for Petroleos Mexicanos since 2002, initially in seismic interpretation for oil and gas exploration, and more recently in the supervision of depth imaging processes. He obtained his Masters Degree in Geophysics in the University of Oklahoma in 2011. He joined SEP in the fall of 2014, and is currently working towards his PhD.



Jason Chang received his B.A. in geophysics from the University of California, Berkeley, in 2010. He joined SEP in autumn of 2011 and is currently working toward a Ph.D. in geophysics. He is a student member of SEG and AGU.

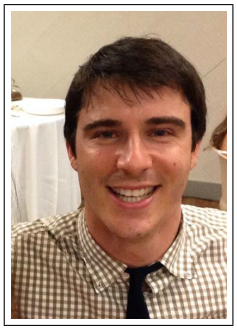
Robert Cieplicki is a Ph.D. student in the Zoback Stress and Crustal Mechanics Research Group (Geophysics Department, Stanford University). He applies machine learning to horizontal wells drilled in shale hydrocarbon reservoirs to predict the efficiency of hydraulic fracturing.



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.

Steve Cole is Manager of Integrated Analysis at OptaSense. He has worked in the oil and gas industry for over 30 years at companies including Fugro, Chevron, and 4th Wave Imaging, which he co-founded. He has a Ph.D. in geophysics from Stanford.



Taylor Dahlke is a fifth year student with SEP. He received his B.S. in civil engineering from the University of California, Berkeley in 2012, and joined SEP in July 2012. Currently, he is working towards a Ph.D. in geophysics with his research focused on applying level set methodologies to perfom salt body segmentation. Taylor is a student member of SPE and SEG.

Eric M. Dunham is an Associate Professor in the Department of Geophysics at Stanford University and an affiliated faculty member with Stanford's Institute for Computational and Mathematical Engineering. He received his PhD in Physics from the University of California, Santa Barbara, in 2005, before moving to Harvard University as a Reginald A. Daly postdoctoral fellow and later as a Lecturer on Applied Math. He has been a professor at Stanford in 2009. Prof. Dunham uses modeling and computation to study natural hazards like earthquakes, volcanoes, and tsunamis, and is now exploring related problems in the oil and gas industry. He is an Alfred P. Sloan Fellow in Physics and a recipient of the National Science Foundation CAREER award and the Stanford School of Earth Sciences Excellence in Teaching Award.

Stuart Farris received his B.S. in Geophysical Engineering from the Colorado School of Mines with an added minor in Computer Science. After graduating in May of 2016, he joined the Stanford Exploration Project that summer and began his Ph.D. at Stanford University. He is interested in the unique problems associated with processing land seismic data and tomographic full waveform inversion.

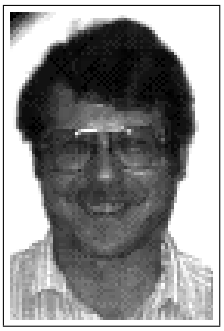
Jeremy A. Goldbogen is an Assistant Professor in the Department of Biology at Stanford University based at the Hopkins Marine Station. He received his PhD in Zoology from the University of British Columba in 2010. He works on the physiology and biomechanics of marine organisms and is interested in the effects of anthropogenic activities of marine life.

Fantine Huot received her M.S. in Science and Engineering from the Ecole Normale Supérieure des Mines de Paris in 2013. Following graduation, she worked as a software developer in France. In fall 2015, started her Ph.D. in Geophysics at Stanford University and joined the Stanford Exploration Project. She leverages machine learning algorithms to address earth science problems.

Joseph Jennings Joseph received his BS in Geophysical Engineering from Colorado School of Mines in 2014. Following graduation, he worked in land seismic processing R&D at ION Geophysical. In the summer of 2015, he joined the SEP as a PhD student where he hopes to pursue research in simultaneous source imaging and imaging of unconventional reservoirs.

Martin Karrenbach is the senior Manager of Innovation at OptaSense. He received his Ph.D. in geophysics from Stanford University in 1995.

Huy Le earned his B.S. degree in Geophysics from the University of Oklahoma in 2012 and his M.S. degree in Computational Geoscience from Stanford University in 2014. He is now a PhD candidate in Geophysics. His research interests include anisotropy, full waveform inversion, finite difference method, and high-performance computing.



Stewart A. Levin was acting director of the Stanford Exploration Project during Jon Claerbout's 1993-4 sabbatical year. After a distinguished career in industry at Mobil and Halliburton, he has returned to Stanford as a senior research scientist in the Department of Geophysics.

Yinbin Ma graduated from University of Science and Technology of China in 2009 with a B.S. in physics. In 2013, he obtained a M.S. in Condensed Matter Theory at Rice University. Yinbin moved to Stanford in 2013 and is currently working towards a Ph.D. in the Institute for Computational and Mathematical Engineering. He started doing research at SEP from June 2014. He is a student member of the SEG.

Musa Maharramov received his diploma c.l. in Mathematics from Baku State University in 1993. In 1995, he received a Ph.D. in Mathematics (Differential Equations) from the Azerbaijan Institute of Mathematics for his thesis "Asymptotic Solutions of Quasilinear Parabolic Equations". In 1995-2001, he worked as an IT& Comms consultant, later IT manager, at Fluor Corporation in Baku. In 2001, he joined BP Azerbaijan, and in 2006 he started his work at BP High Performance Computing as a Computational Scientist. In that role he worked with the BP Imaging Team on mathematical, algorithmic and computational aspects of seismic migration and inversion. Musa joined SEP in the fall of 2011 and is pursuing his second Ph.D., in Geophysics. He is currently working on regularization of seismic inversion problems through the application of Geological/Geomechanical constraints and advanced optimisation techniques. Musa is a member of the SEG, EAGE and SIAM.

Eileen Martin graduated from the University of Texas at Austin in 2012 with a B.S. in Mathematics and Computational Physics. She is a Department of Energy Computational Science Graduate Fellow, and is working towards her Ph.D. in the Institute for Computational and Mathematical Engineering. She Joined SEP in the summer of 2013.

Tapan Mukerji is an Associate Professor (Research) in the Department of Energy Resources Engineering and the Department of Geophysics at Stanford University. He co-directs the Stanford Center for Reservoir Forecasting (SCRF), and is associated with the Stanford Rock Physics and Borehole Geophysics (SRB) and Stanford Basin and Petroleum System Modeling (BPSM) research groups. Tapan obtained his Ph.D. in Geophysics from Stanford University, in 1995, and M.S. (Geophysics) and B.S. (Physics) degrees from Banaras Hindu University, India. His research interests include rock physics, geostatistics, wave propagation, stochastic methods for quantitative reservoir characterization, time-lapse reservoir monitoring, and geomodeling applications. He was awarded the Karcher Award in 2000 by the Society of Exploration Geophysicists. He is a co-author of "The Rock Physics Handbook", "Quantitative Seismic Interpretation", and "The Value of Information in the Earth Sciences" all published by Cambridge University Press. In 2014 Tapan was awarded the ENI award - the so called "Energy Nobel Prize" - for pioneering innovations in theoretical and practical rock physics for seismic reservoir characterization.

Shuki Ronen is currently (2015) external and collaborative research manager of Dolphin Geophysical and a consulting faculty at SEP.

Rahul Sarkar graduated with a B.S and M.S from the Indian Institute of Technology, Kharagpur in 2011. After that he worked for Schlumberger for four years in the United States and Mexico. Since 2015, he is pursuing a M.S in the Institute for Computational and Mathematical Engineering at Stanford University.

Leighton Watson received his B.Sc. (Hons) in geophysics and mathematics from the University of Auckland in 2014. He is a third year student in the geophysics department working on volcano infrasound with Eric Dunham. He is working on numerical modeling of seismic airguns and whale sound generation for his second project with the Stanford Exploration Project. He is a member of SEG and AGU.

SEP ARTICLES PUBLISHED OR IN PRESS

Almomin, A., and Biondi, B., 2012, Tomographic full waveform inversion: Practical and computationally feasible approach. SEG Technical Program Expanded Abstracts, 1–5.

Almomin, A., and Biondi, B., 2013, Tomographic full waveform inversion (TFWI) by successive linearizations and scale separations. SEG Technical Program Expanded Abstracts, 1048–1052.

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

Almomin, A., and Biondi, B., 2014, Generalization and amplitude normalization of tomographic full waveform inversion. SEG Technical Program Expanded Abstracts.

Barak, O., Jaiswal, P., de Ridder, S., Giles, J., Brune, R. and Ronen, S., 2014, Six-component seismic land data acquired with geophones and rotation sensors: wave-mode separation using 6C SVD: SEG Technical Program Expanded Abstracts (Submitted).

Chang, J., de Ridder, S., and Biondi, B., 2013, Power spectral densities and ambient noise cross-correlations at Long Beach: SEG Technical Program Expanded Abstracts **32**, 2196–2200.

Chang, J., Nakata, N., Clapp, R.G., de Ridder, S., and Biondi, B., 2014, High-frequency surface and body waves from ambient noise cross-correlations at Long Beach, CA: SEG Technical Program Expanded Abstracts **33**.

Chang, J., and B. Biondi, 2015, Rayleigh-wave tomography using traffic noise at Long Beach, CA: SEG Technical Program Expanded Abstracts **34**.

Chang, J.P., de Ridder, S.A.L., and Biondi, B.L., 2016, High-frequency Rayleigh-wave tomography using traffic noise from Long Beach, California: Geophysics, 81, No. 2, B43–B53.

Chang, J.P., 2016, Multicomponent ambient noise crosscorrelation at Forties: SEG Technical Program Expanded Abstracts **35**.

Dahlke, T., Beroza, G., Chang, J., and de Ridder, S., 2014, Stochastic variability of velocity estimates using eikonal tomography on the Long Beach data set: SEG Technical Program Expanded Abstracts **33** .

Dahlke, T., Biondi, B., Clapp, R., 2015, Domain decomposition of level set updates for salt segmentation: SEG Technical Program Expanded Abstracts **33** (Submitted).

Grobbe, N., F.C. Schoemaker, M.D. Schakel, S.A.L. de Ridder, E.C. Slob, and D.M.J. Smeulders, 2012, Electrokinetic fields and waves: Theory, experiments, and numerical modeling. Geophysical Research Abstracts, Vol. **14**, EGU2012–10636.

de Ridder, S., and B. Biondi, 2012, Continuous passive seismic monitoring of CCS projects by correlating seismic noise: A feasibility study. 74th Conference & Technical Exbition, EAGE, Extended Abstracts (Accepted).

de Ridder, S., and B. Biondi, 2012, Reservoir monitoring by passive seismic interferometry. SEG/SPE/AAPG Joint Workshop, La Jolla, CA, 24-29 June 2012 (Accepted).

de Ridder, S., 2012, Ambient seismic noise correlations for reservoir monitoring. SEG Technical Program Expanded Abstracts **31** (Submitted).

de Ridder, S., and B. Biondi, 2012, Continuous reservoir monitoring by ambient seismic noise tomography. SEG Technical Program Expanded Abstracts **31** (Submitted).

Schoemaker, F.C., N. Grobbe, M.D. Schakel, S.A.L. de Ridder, E.C. Slob, and D.M.J. Smeulders, 2012, Experimental validation of the electrokinetic theory and development of seismoelectric interferometry by cross-correlation: International Journal of Geophysics (Accepted).

Farghal, M.S., and Levin, S.A., 2012, Hunting for microseismic reflections using multiplets: SEG Technical Program Expanded Abstracts **31**.

Guitton, A., Ayeni, G., and Esteban, D.A., 2012, Constrained full-waveform inversion by model reparameterization: Geophysics, 77, No. 2, R117–R127.

Guitton, A., and Esteban, D.A., 2012, Attenuating crosstalk noise with simultaneous source full waveform inversion: Geophysical Prospecting (In Press).

Guitton, A., 2012, Blocky regularization schemes for full waveform inversion: Geophysical Prospecting (In Press).

Halpert, A., 2012, Edge-preserving smoothing for segmentation of seismic images: SEG Technical Program Expanded Abstracts **31**, 1–5.

Halpert, A., Clapp, R.G., and B. Biondo, 2014, Salt delineation via interpreter-guided seismic image segmentation, Interpretation: **2**, T79–T88.

Arevalo, Humberto and Stewart A. Levin, 2014, Well and Seismic matching with ArcGIS and ProMAX via KML: Esri International User Conference Paper 901, San Diego, 15 July.

Leader, C., and R. Clapp, 2012, Least squares reverse time migration on GPUs - balacing IO and computation: 74th Conference & Technical Exbition, EAGE, Extended Abstracts (Accepted).

Leader, C., and A. Almomin, 2012, How incoherent can we be? Phase encoded linearised inversion with random boundaries: SEG Technical Program Expanded Abstract **31** (Submitted).

Levin, Stewart A. and Fritz Foss, 2014, Downward continuation of Mars SHARAD data: SEG Technical Program Expanded Abstract **33** (Submitted).

Li, Y., Y. Zhang, and J. Claerbout, 2012, Hyperbolic estimation of sparse models from erratic data: Geophysics **77**, 1–9.

Li, Y., P. Shen, and C. Perkins, 2012, VTI migration velocity analysis using RTM: SEG Technical Program Expanded Abstract **31**.

Li, Y., Image-guided WEMVA for azimuthal anisotropy: SEG Technical Program Expanded Abstract **32**.

R.E. Plessix and Y. Li, 2013, Waveform acoustic impedance inversion with spectral shaping: Geophysical Journal International **195**(1), 301–314, 2013.

Li, Y., B. Biondi, D. Nichols, and R. Clapp, 2014, Wave equation migration velocity analysis for VTI models: Accepted for publication in Geophysics, 2014.

Li, Y., M. Wong, and R. Clapp, 2014, Equivalent accuracy by fractional cost: overcoming temporal dispersion: Submitted to SEG 2014.

Li, Y., D. Nichols, and G. Mavko, 2014, Stochastic rock physics modeling for seismic anisotropy: Submitted to SEG 2014.

Li, Y., B. Biondi, R. Clapp, and D. Nichols, 2014, Rock physics constrained anisotropic wave-equation migration velocity analysis: Submitted to SEG 2014.

Maharramov M., and Biondi, B., 2014, Joint full-waveform inversion of time-lapse seismic data sets. SEG Technical Program Expanded Abstracts, 954–959.

Maharramov M., and Biondi, B., 2014, Joint 4DFWI with model-difference regularization. SEG-AGU Summer Research Workshop. Advances in Active+Passive "Full Wavefield" Seismic Imaging: From Reservoirs to Plate Tectonics.

Maharramov M., and Biondi, B., 2015, Robust Simultaneous Time-lapse Full-waveform Inversion with Total-variation Regularization of Model Difference. EAGE Technical Program Expanded Abstracts (accepted).

Maharramov M., 2015, Efficient Finite-difference Modelling of Acoustic Wave Propaga-

tion in Anisotropic Media with Pseudo-sources. EAGE Technical Program Expanded Abstracts (accepted).

Maharramov M., and Biondi, B., 2015, Simultaneous TV-regularized time-lapse FWI with application to field data. SEG Technical Program Expanded Abstracts (submitted).

Martin E.R., B. Biondi, M. Karrenbach, and S. Cole, 2017, Ambient noise interferometry from DAS array in underground telecommunications conduits. EAGE Technical Programme Expanded Abstracts (accepted).

Martin E.R., B. Biondi, M. Karrenbach, and S. Cole, 2017, Continuous subsurface monitoring by passive seismic with distributed acoustic sensors- the "Stanford Array" experiment. Proceedings of the First EAGE Workshop on Practical Reservoir Monitoring.

Martin E.R., P. Wills, D. Hohl, and J.L. Lopez, 2017, Using Machine Learning to Predice Production at a Peace River Thermal EOR Site. Proceedings of SPE Reservoir Simulation Conference.

Martin E.R., N.J. Lindsey, S. Dou, J.B. Ajo-Franklin, A. Wagner, K. Bjella, T.M. Daley, B. Freifeld, M. Robertson, C. Ulrich, 2016, Interferometry of a roadside DAS array in Fairbanks, AK. SEG Technical Program Expanded Abstracts.

Martin E.R., J. Ajo-Franklin, S. Dou, N. Lindsey, T.M. Daley, B. Freifeld, M. Robertson, A. Wagner, and C. Ulrich, 2015, Interferometry of ambient noise from a trenched distributed acoustic sensing array. SEG Technical Program Expanded Abstracts.

Ajo-Franklin, J., N. Lindsey, S. Dou, T.M. Daley, B. Freifeld, E.R. Martin, M. Robertson, C. Ulrich, and A. Wagner, 2015, A field test of distributed acoustic sensing for ambient noise recording. SEG Technical Program Expanded Abstracts.

Li, Y., H. Yang, E.R. Martin, K. Ho, and L. Ying, 2015, Butterfly factorization. Multiscale Modeling and Simulation, 13, 714–732.

Shen, Y., B. Biondi, R. Clapp, and D. Nichols, 2013, Wave-equation migration Q analysis (WEMQA): EAGE Workshop on Seismic Attenuation Extended Abstract 2013.

Shen, Y., B. Biondi, R. Clapp, and D. Nichols, 2014, Wave-equation migration Q analysis (WEMQA): SEG Technical Program Expanded Abstracts.

Li, Y., Y. Shen and P. Kang, 2015, Integration of seismic and fluid-flow data: a two-way road linked by rock physics :77th EAGE Conference and Exhibition.

Wong, M., and Ronen, S., and Biondi, B.L., 2012, Joint imaging with streamer and ocean bottom data. SEG Technical Program Expanded Abstracts, 1–5.

Wong, M., and Ronen, S., and Biondi, B.L., 2012, Imaging with multiples using linearized full-wave inversion. SEG Technical Program Expanded Abstracts, 1–5.

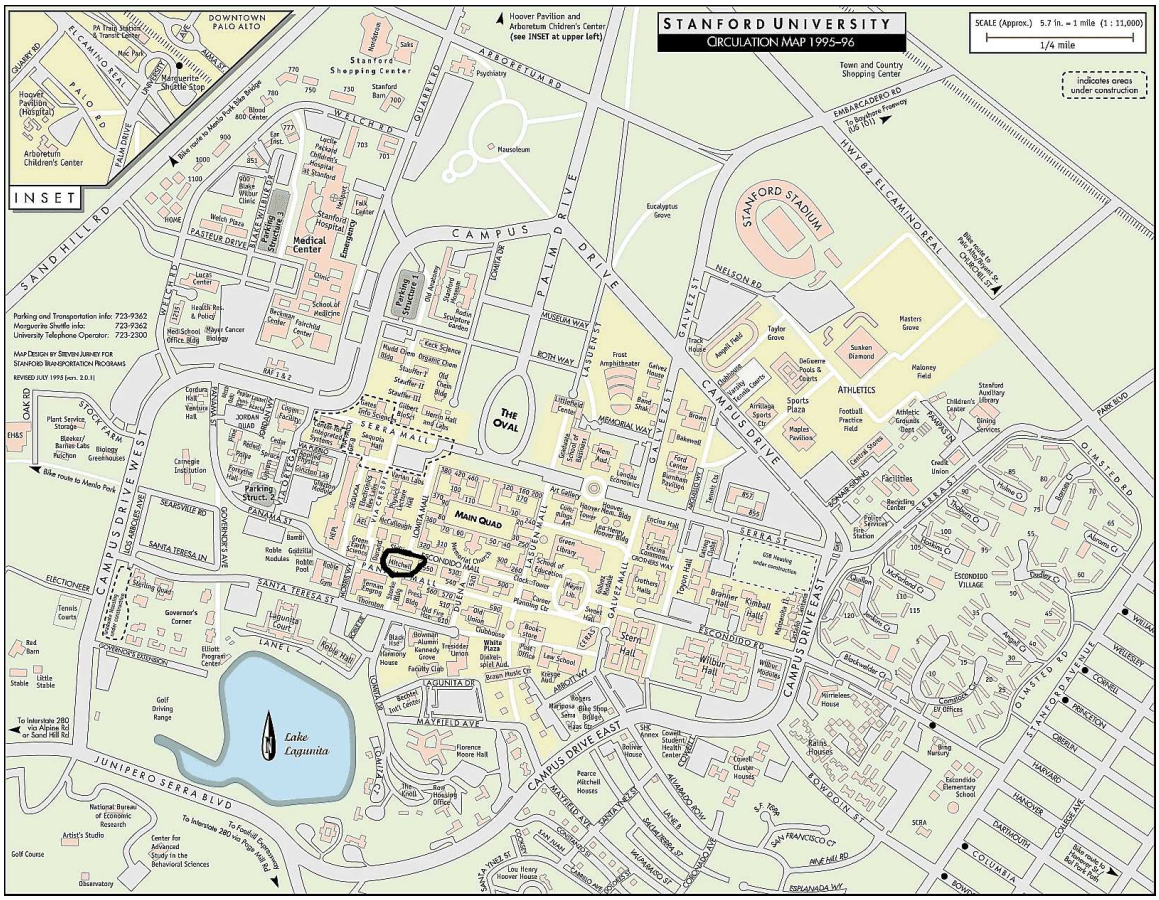
Wong, M., 2013, Handling salt reflection in Least-squares RTM SEG Technical Program Expanded Abstracts, 3921–3925.

Wong, M., B. Biondi, S. Ronen, C. Perkins, M. Merritt, V. Goh, and R. Cook, 2014, Robbust least squares RTM on the 3D Deimos ocean bottom node dataset SEG Technical Program Expanded Abstracts, (Submitted)

Wong, M., B. Biondi, and S. Ronen, 2014, Imaging with multiples using least-squares reverse-time migration. The Leading Edge, (submitted)

Huy Le and Stewart A. Levin, 2014, Removing shear artifacts in acoustic wave propagation in orthorhombic media: SEG Technical Program Expanded Abstracts, 486–490.

Huy Le, Biondo Biondi, Robert G. Clapp, and Stewart A. Levin, 2015, Using a nonlinear acoustic wave equation for anisotropic inversion: SEG Technical Program Expanded Abstracts (Submitted).


STANFORD UNIVERSITY
CIRCULATION MAP 1995–96
SCALE (Approx.) 5.7 in. = 1 mile (1 : 11,000)
1/4 mile
indicates areas under construction
INSET
Hoover Pavilion and Arboretum Children's Center (see INSET at upper left)
DOWNTOWN PALO ALTO
Stanford Shopping Center
Town and Country Shopping Center
Medical Center
Stanford Hospital
Hoover Pavilion (Hospital)
Arboretum Children's Center
Mausoleum
Eucalyptus Grove
STANFORD STADIUM
THE OVAL
MAIN QUAD
Lake Lagunita
Golf Driving Range
Golf Course
Observatory
Stern Hall
Wilbur Hall
Lou Henry Hoover House
Parking and Transportation info: 723-9362
Marguerite Shuttle info: 723-9362
University Telephone Operator: 723-2300
MAP DESIGN BY STEVEN JURNEY FOR STANFORD TRANSPORTATION PROGRAMS
REVISED JULY 1995 (rev. 2.0.1)