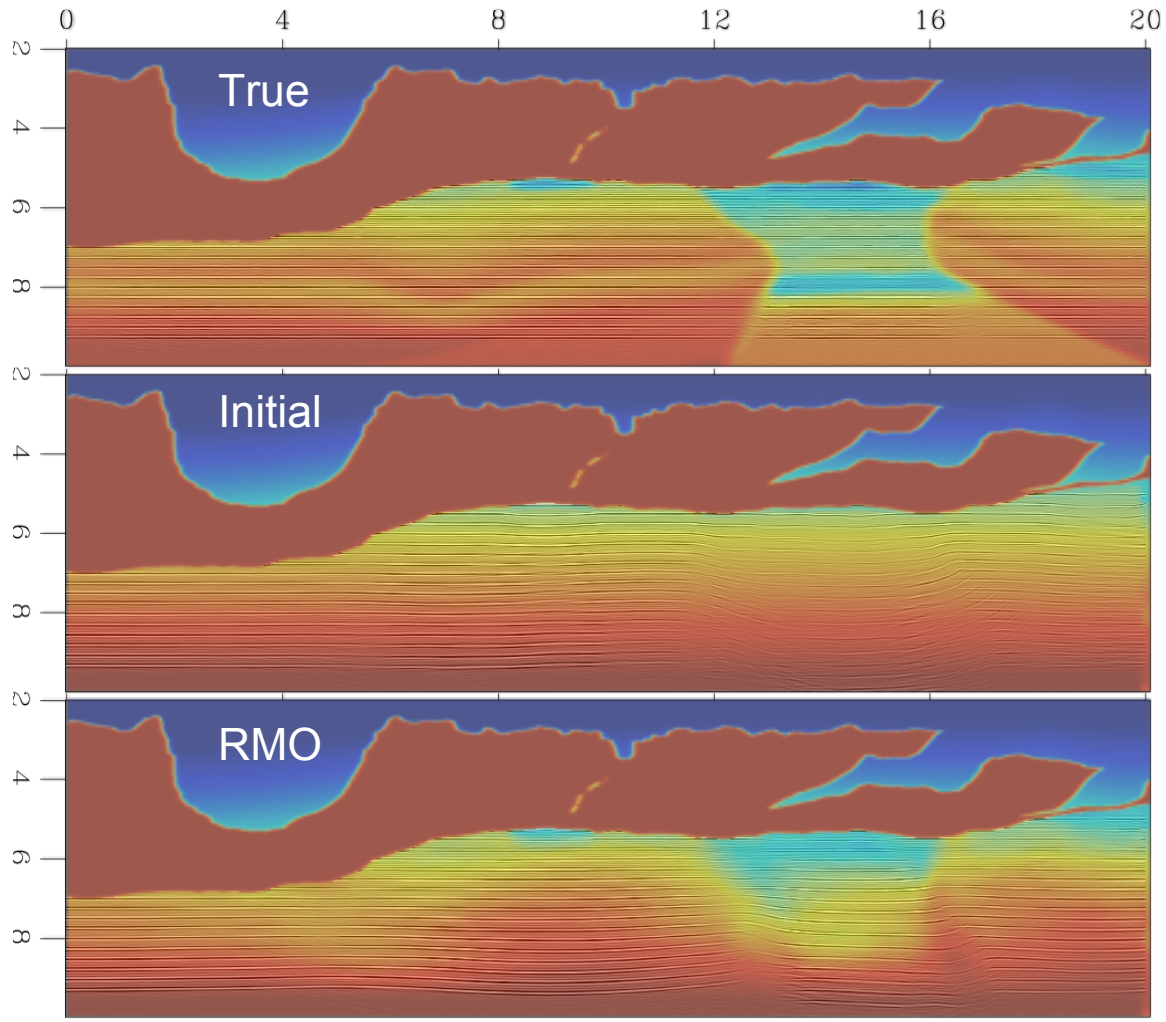
STANFORD EXPLORATION PROJECT

Ali Almomin, Ohad Barak, Biondo Biondi, Jason Chang, Jon Claerbout, Robert Clapp, Taylor Dahlke, Adam Halpert, Huy Le, Chris Leader, Stewart Levin, Yunyue Li, Musa Maharramov, Gary Mavko, Dave Nichols, Sjoerd de Ridder, Shuki Ronen, Guojian Shan, Xukai Shen, Yi Shen, Mandy Wong, and Yang Zhang

Report Number 149, June 2013



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/sep149
[2] http://sepwww.stanford.edu/public/docs/sepdatalib/toc_html
[3] http://sepwww.stanford.edu/research/redoc/

SEP149 — TABLE OF CONTENTS

Deconvolution and signal processing

Velocity estimation

Accelerated imaging

Imaging through inversion

Passive seismic

Modeling

Anisotropy and attenuation

Miscellaneous

Ricker-compliant and pseudo-unitary decon

Jon Claerbout

ABSTRACT

Ricker compliant deconvolution spikes the center lobe of the Ricker wavelet. It enables deconvolution to preserve and enhance seismogram polarities. It works by tapering at small lags the anti-symmetric part of the time-domain representation of the log spectrum. A byproduct of this decon is a pseudo-unitary (very clean) debubble filter where bubbles are lifted off the data while onset waveforms (usually Ricker) are untouched.

INTRODUCTION

Seismogram polarity becomes more apparent in seismic data when deconvolution removes the correct source wavelet. A problem is that predictive deconvolution does not do a good job defining that wavelet. Predictive decon assumes minimum phase while marine seismology typically exhibits the Ricker wavelet, a wavelet which is both theoretically and practically, marginally or doubtfully minimum phase. This problem is resolved in this paper by shifting the time origin to the middle main lobe of the Ricker wavelet and simultaneously estimating both this now noncausal shot waveform and its inverse. A byproduct of this approach is a debubble process giving results of outstanding clarity.

Textbook deconvolution produces a white output, but it cannot be true that we wish energy at the Nyquist frequency while the sampling rate is quite arbitrary. Recently I discovered that parameterizing the logarithm of the spectrum in the time domain lays the problem out in a manner where practical issues sort themselves out along the "quefrency" axis. I came to this approach because my inverse theory (with Guitton and with Fu) lacked a certain regularization this provides. Results here are excellent, and computed in $N \log_2 N$ time. I regard them as a final analytical stage before invoking iterative inverse theory.

BASICS OF LAG-LOG SPACE

Let $F(\omega)$ be a filter in frequency domain. Let $U(\omega) = \ln F(\omega)$ so the filter is $F(\omega) = e^{U(\omega)}$. With the definition $Z = e^{i\omega\Delta t}$ Fourier transforms become polynomials (Z-transforms). Thus $U(\omega)$ relates to the time function u_τ by a Fourier sum $U = \sum_{\tau=0}^{2048} u_\tau Z^\tau$. (The 2048 simply means "the whole trace".)/ The u_τ values will be our parameterization of the filter. Historically, the τ axis is called the "quefrency" axis though we sometimes refer to it as the "lag-log" axis.

The property of exponentials that $e^{A+B+C} = e^A\, e^B\, e^C$ has an interesting meaning when we exponentiate a Z-transform $\exp(A + B + C) = \exp(\sum_{\tau=1}^{2048} u_\tau Z^\tau)$. The Z-transform sum may be split up into small lags, medium lags, and large lags. This decomposes a filter (or waveform) into a sequence of three filters, each with its own meaning, for example, in

common marine seismology:

$$e^{A+B+C} = e^A\, e^B\, e^C \tag{1}$$

$$e^{\sum_{\tau=1}^{2048} u_\tau Z^\tau} = e^{\sum_1^2}\, e^{\sum_3^{15}}\, e^{\sum_{16}^{2048}} \tag{2}$$

$$\text{(wavelet)} = \text{(continuity)(Ricker)(bubble)} \tag{3}$$

Equation (2) defines the boundaries of the three regions abruptly although in practice we blend them smoothly. Changing the sign of $(A+B+C)$ changes a filter to its inverse. Both are parameterized by the same A, B, and C.

We may specify u_τ from prior knowledge, or from knowledge gained from various kinds of data averaging, or from some mixture of the two. Commonly, we begin using Kolmogoroff spectral factorization (next section) to give all the u_τ. We may design a filter e^{A+B+C} by overriding Kolmogoroff with $A = 0$ and $B = 0$. To see what happens, consider the filter $e^C = 1 + C + C^2/2! + \cdots \approx 1 + C$. Examine its leading coefficients. They are $(1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, u_{16}, u_{17}, u_{18}, \cdots)$. Figure 1 shows the application of such a filter. This operation on the data is called "debubbling". The 14 interval gap on 4ms data is 60ms, a number halfway between the "end of the ghosts" and the "onset of the bubbles". This result may be described as "textbook quality" (meaning it is the best I have ever produced).

Setting all of A, B, and C to zero makes a unitary filter that is simply an impulse. Setting to zero any of the three, or any combination, or suppressing any combination tends to make a filter more unitary, hence the appellation "pseudo-unitary". Debubbling in this manner seems to leave first arrivals untouched, further justifying the term. Before we go on to attack the middle-lag terms we review the starting point, Kolmogoroff spectral factorization.

Kolmogoroff spectral factorization

When a time function such as c_t vanishes at all negative time lags it is said to be causal. Its Z transform is $C(Z) = c_0 + c_1 Z + c_2 Z^2 + c_3 Z^3 + \cdots$. Observe that $C(Z)^2$ is also causal because it has no negative powers of Z, alternately, because the convolution of a causal with a causal is causal. Likewise e^C is causal because it is a sum of causals.

$$e^C = 1 + C + C^2/2! + C^3/3! + \cdots \tag{4}$$

Happily, this infinite series always converges because of the strong influence of the denominator factorials. The time-domain coefficients for e^C could be computed the hard way, putting polynomials into power series, or e^C may be computed by Fourier transforms. To do so, we would evaluate $C(Z = e^{i\omega})$ for many real ω, and then invoke an inverse Fourier transform program to uncover the time-domain coefficients.

Let $r = r(\omega)$, $\phi = \phi(\omega)$, and $Z^\tau = e^{i\omega\tau}$. Let us investigate the consequence of exponentiating a causal filter.

$$|r|e^{i\phi} = e^{\ln|r| + i\phi} = e^{\sum_\tau c_\tau Z^\tau} = \exp\left(\sum_\tau c_\tau Z^\tau\right) \tag{5}$$

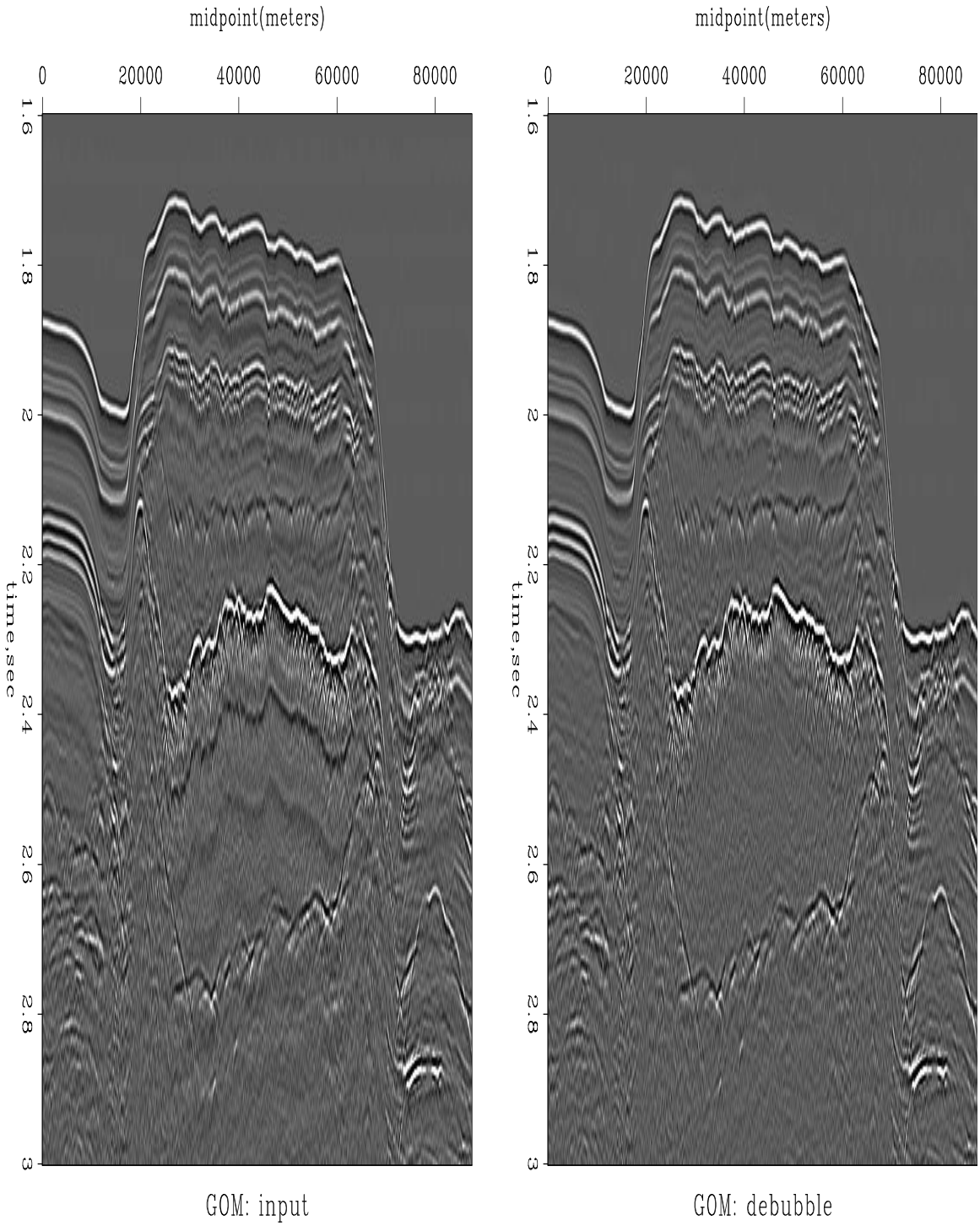


Figure 1: Gulf of Mexico data before and after pseudo-unitary debubble. This process preserves the wave onset while it lifts off the bubbles. Here the effect of the process is visible nearly everywhere after 2.4 sec, but also visible around 1.85 sec. On blinking displays it is easy to see bubble removed nearly everywhere. [**ER**] jon1/. GOM-unitary

Notice a pair of filters, both causal and inverse to each other.

$$|r|e^{i\phi} = e^{+\sum_\tau c_\tau Z^\tau} \tag{6}$$
$$|r|^{-1}e^{-i\phi} = e^{-\sum_\tau c_\tau Z^\tau} \tag{7}$$

A filter from any such pair is said to be "minimum phase". Many filters are not minimum phase because they have no causal inverse. For example the delay filter Z. It's inverse, Z^{-1} is not causal. Such filters do not relate to a causal complex logarithm. If they have a logarithm, it must be non causal.

Given a spectrum $r(\omega)$ we can construct a minimum-phase filter with that spectrum. Since $r(\omega)$ is a real even function of ω, the same may be said of its logarithm. Let the inverse Fourier transform of $\ln|r(\omega)|$ be u_τ, where u_τ is a real even function of time. Imagine a real odd function of time v_τ.

$$|r|e^{i\phi} = e^{\ln|r|+i\phi} = e^{\sum_\tau (u_\tau+v_\tau)Z^\tau} \tag{8}$$

The phase $\phi(\omega)$ transforms to v_τ. We can assert causality by choosing v_τ so that $u_\tau+v_\tau=0$ for all negative τ. This defines v_τ at negative τ. Since v_τ is odd, we also know its values at positive lags. This creates a causal exponent which creates a causal minimum-phase filter with the specified spectrum. The code does this by multiplying u_τ by a step function of height 2, the doubling accounting for the zeroing of half the axis. This computation is called Kolmogoroff spectral factoring. The word "factoring" enters because in applications one begins with an energy spectrum $|r|^2$ and factors it into an $re^{i\phi}$ times its conjugate (time reverse).

It is a exercise for the student to show that a complex-valued time function has a positive spectrum that is non-symmetrical in frequency and may also be factored.

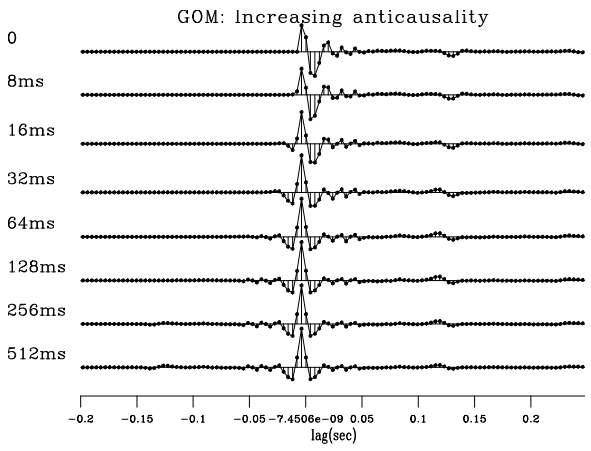


Figure 2: Gulf of Mexico Δt = 4ms: Increasing the anticausality in Ricker compliant decon. Too much causes half the bubble to appear before the shot. Numerical values of anticausality refer to the lag τ at which the sine-squared weight upon u_τ^{anti} reaches unity where weighting ceases. [**ER**] jon1/. GOM-laggard

Ricker compliant decon

Start with the u_τ resulting from a Kolmogoroff factorization. (Optionally you might weight down portions making it more unitary.) Split it into even and odd parts, $u_\tau^{\text{odd}} = (u_\tau - u_{-\tau})/2$ and $u_\tau^{\text{even}} = (u_\tau + u_{-\tau})/2$ whose sum is u_τ. The even part Fourier transforms to the logarithm of the amplitude spectrum. The odd part Fourier transforms to the phase spectrum. Here we monkey with the phase while not changing the amplitude. We simply taper u_τ^{odd} towards

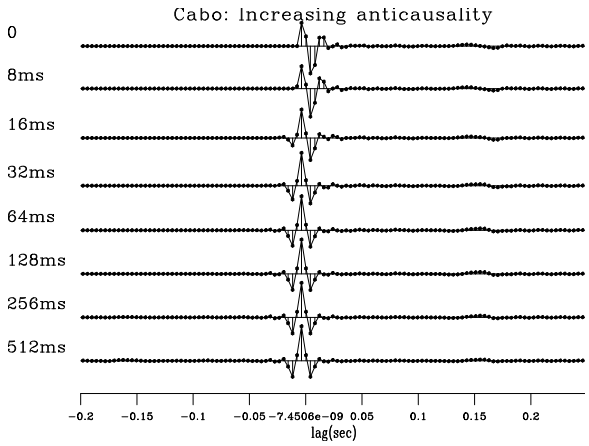


Figure 3: Gulf of California Δt = 4ms: Results similar to Figure 2. Notice in both cases the strictly causal wavelet looks as much like a doublet as like a Ricker wavelet. [**ER**] jon1/. Cabo-laggard

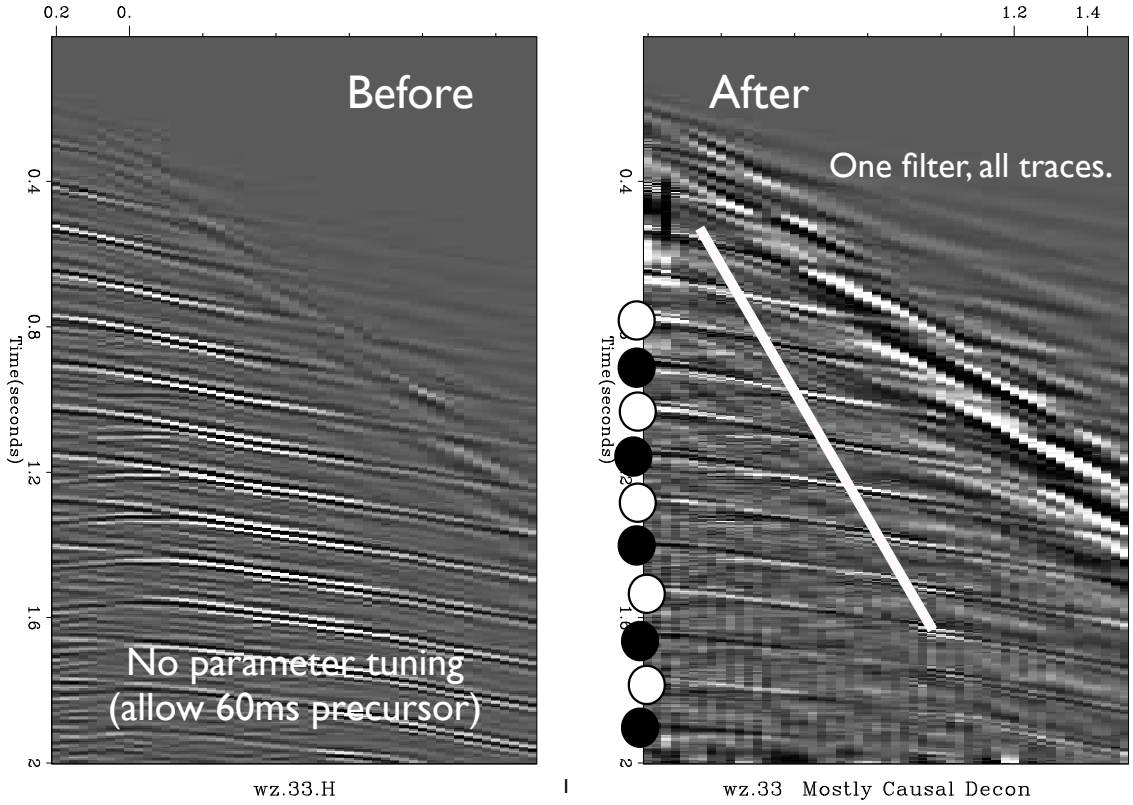


Figure 4: The spectrum of each trace in a shot gather was averaged and used to produce a single Ricker compliant deconvolution filter used on all traces. Because of changes with angle, consistently alternating polarity of multiples is best observed along a diagonal line. [**NR**] jon1/. wz33

zero for small lags. Figure 2 examines the consequences of various numerical choices of "small". As we increase the anticausality, the time function of the wavelet e^U increases in symmetry near $t = 0$. The example in Figure 2 is based on the average spectra of 1001 traces from a line in the Gulf of Mexico.

Our favored choice of 60ms is larger than the Ricker width, about 20ms, but not as large as the bubble delay, about 150ms, an easy choice. The result of dividing such a wavelet from data is shown in Figures 4 and 5.

Band limiting

The very short lags in the lag-log space raise deep issues about the meaning of functions. Just the fact that a continuous function of time may be sampled densely on the time axis implies a large amplitude ratio between the central frequency band and the Nyquist frequency. Do we really want the deconvolution to produce a white output all the way to the Nyquist? Probably not. Consistent with our quefrency approach these issues show up at very small values of τ. They are there because the prediction error filter for a nearly continuous function should be something like $(1, -1)$ which vanishes at the Nyquist frequency, has a log spectrum will tend to blowing up (negatively).

An appealing aspect of deconvolution is that it defines in principle "the" wavelet which connects you to some canonical trace. So a second deconvolution of your data should take you nowhere because you are already "there". What other answer can there be but a white trace? This appears to invite some kind of an optimization formulation—beyond the scope of this study.

A pragmatic answer not meeting any philosophical goals is an adjustable parameter for the code. This parameter should repress the tendency of deconvolution to whiten near the Nyquist frequency. The parameter may be thought of as specifying our expected time resolution in milliseconds. The default `tresol=.01` says we are thinking of time resolution about 10ms. Specifying `tresol=0` would say we are envisioning great detail on the time axis, so we'd get a white output. This parameter is is implemented in the same manner as the pseudo-unitary decon. The only difference is the parameter τ_{tresol} defaults to 10ms instead of 60ms. It is installed in the code defaulting at $\tau_{\text{tresol}} = 10$ms. An example is shown in Figure 6.

WHAT IS LEFT TO DO?

This paper starts from a given spectrum, here an average of many traces. In reality, the spectrum might vary from trace to trace. The spectrum will vary from one offset to the next. These basic aspects are not addressed here.

Inverse theory incorporates other real-world complications beyond the scope of the present study. Our ongoing work with it shows it reveals polarity even more clearly. Perhaps so because it uses ℓ_1-like statistics. Also it correctly handles gain and filtering as non commuting operators. The method of this paper, however, wins when it comes to clear and simple parameter choice.

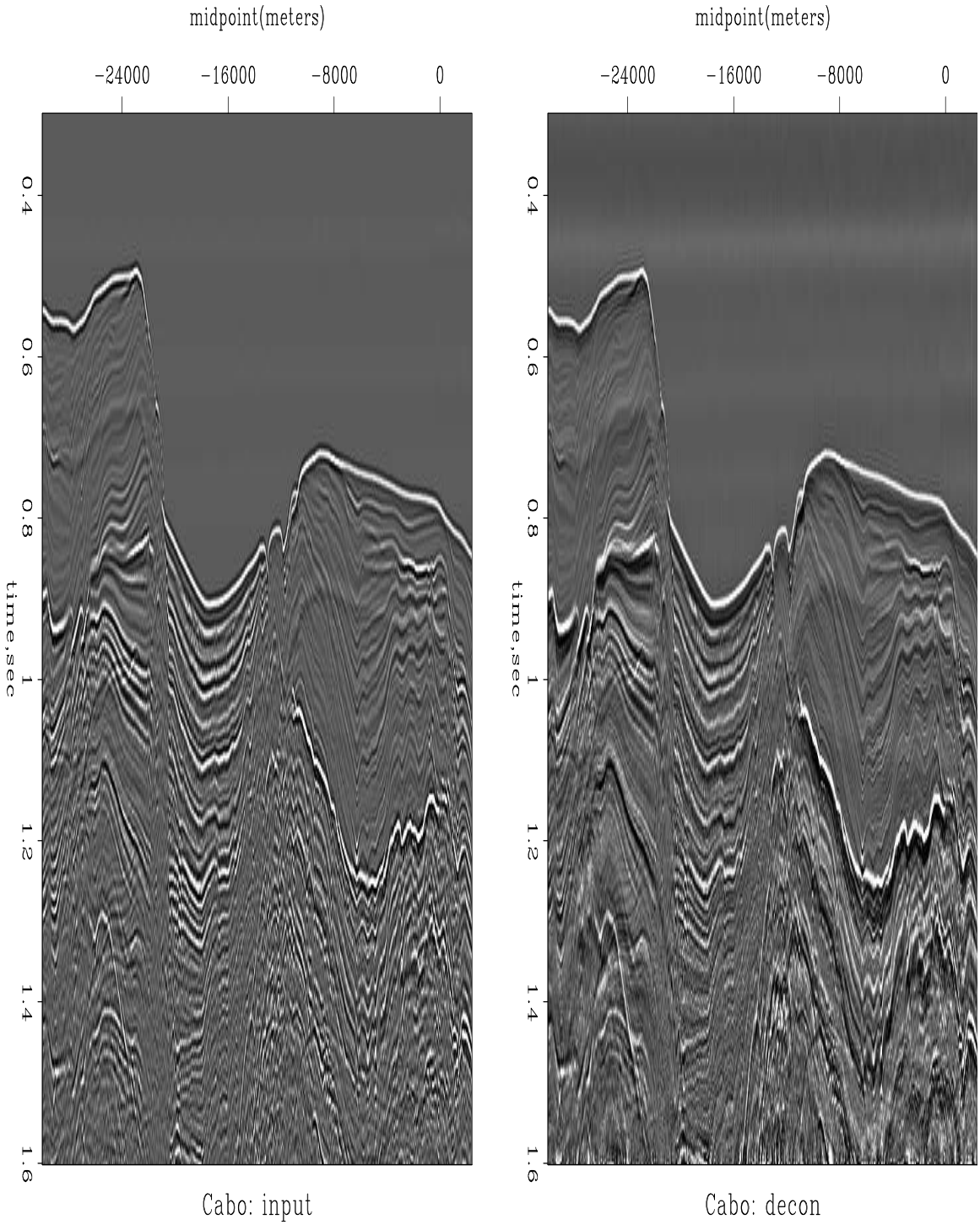


Figure 5: Gulf of California data and its decon. This deconvolution, called Ricker-compliant decon is converting the water bottom Ricker wavelet to a positive pulse. Imperfect debubbling may be explained by (1) guns changing along the line, and/or (2) the depth 550ms being comparable to the bubble delay 180ms (which calls for inverse theory.) [**ER**] jon1/. Cabo-decon

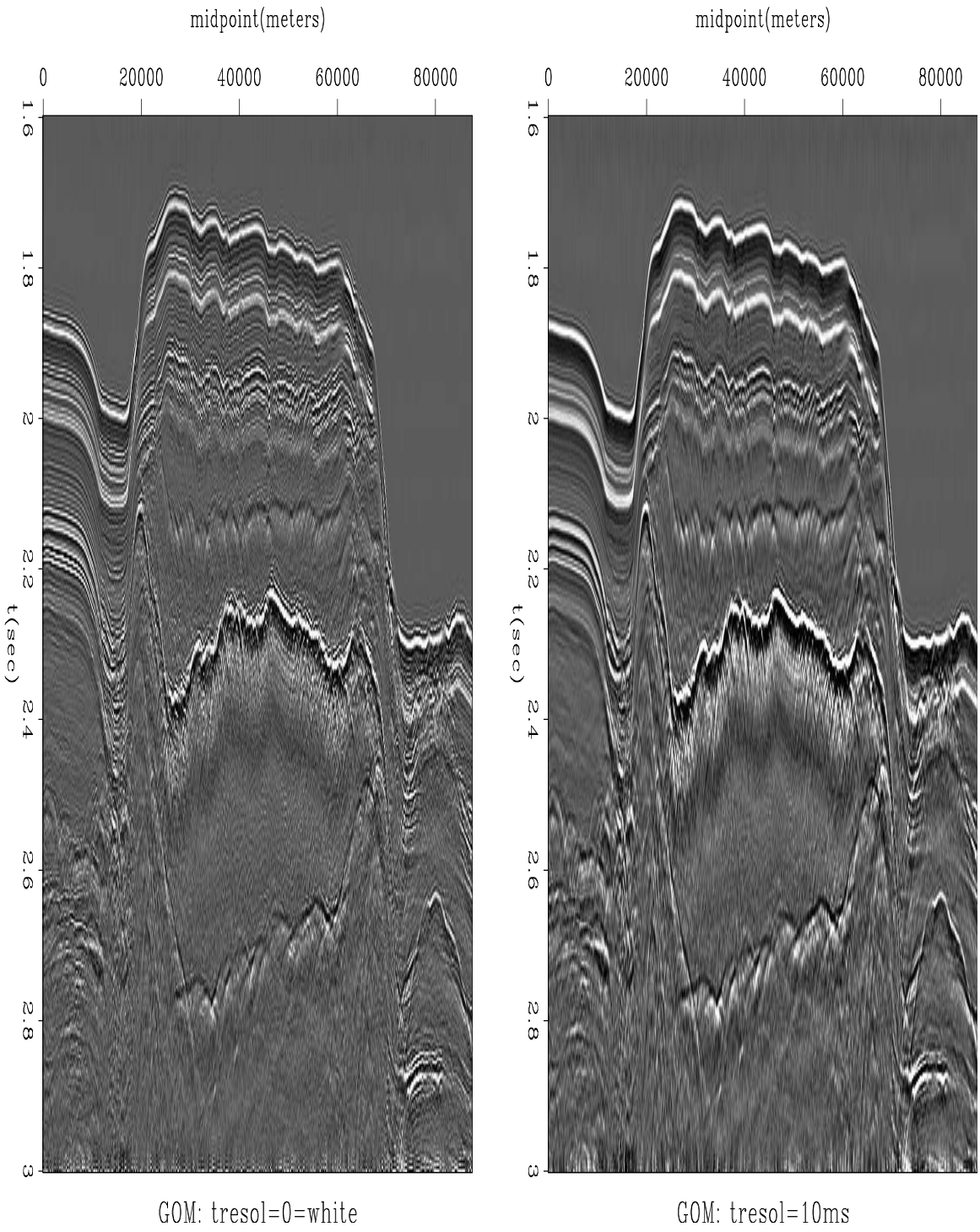


Figure 6: Getting a smoother output by suppressing lags less than about 10ms. [**ER**] jon1/. GOM-tresol2

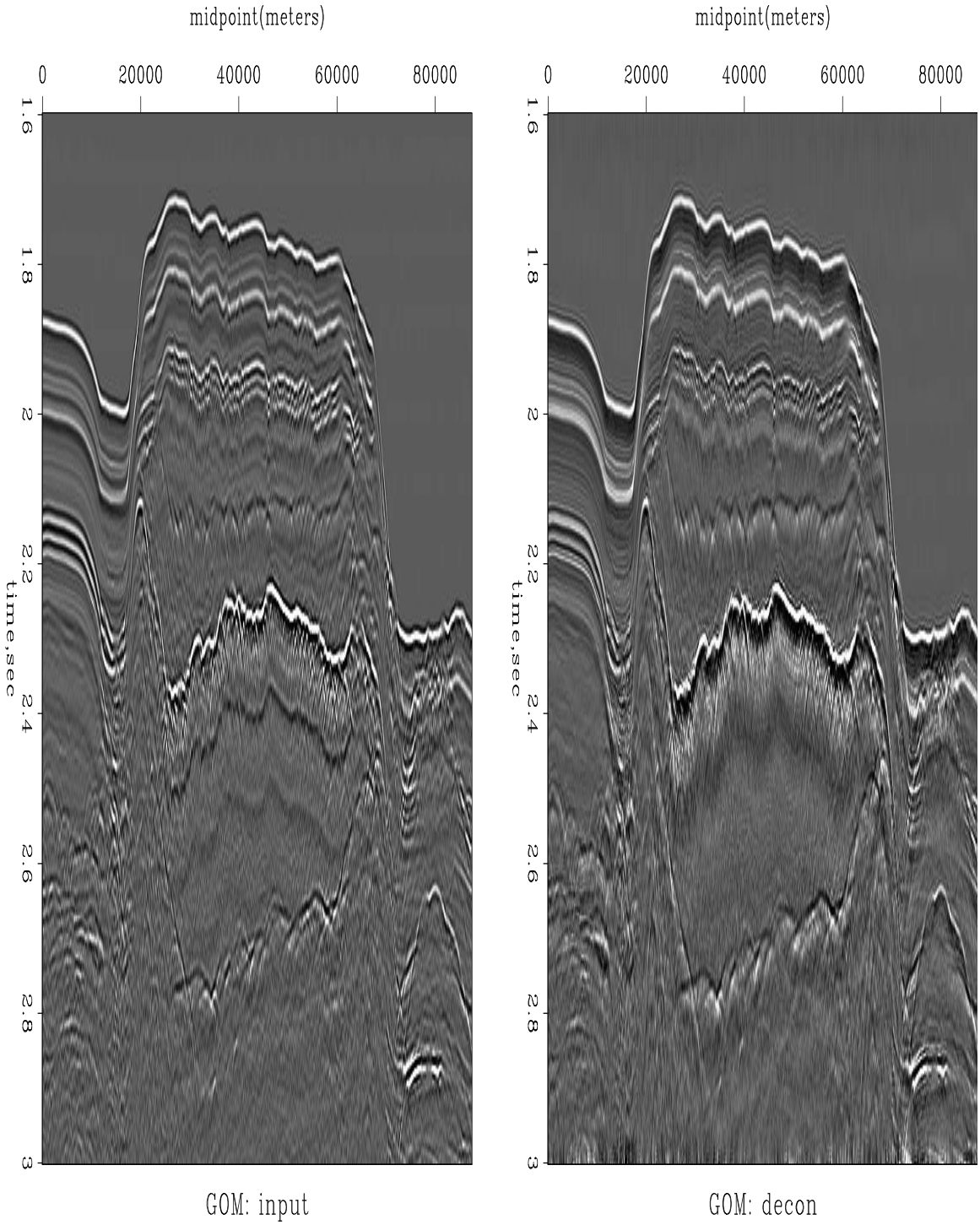


Figure 7: Gulf of Mexico data before and after decon. The Ricker wavelet should have been turned into a positive pulse. Unfortunately, very low frequencies have appeared making this less apparent. [**ER**] jon1/. GOM-decon

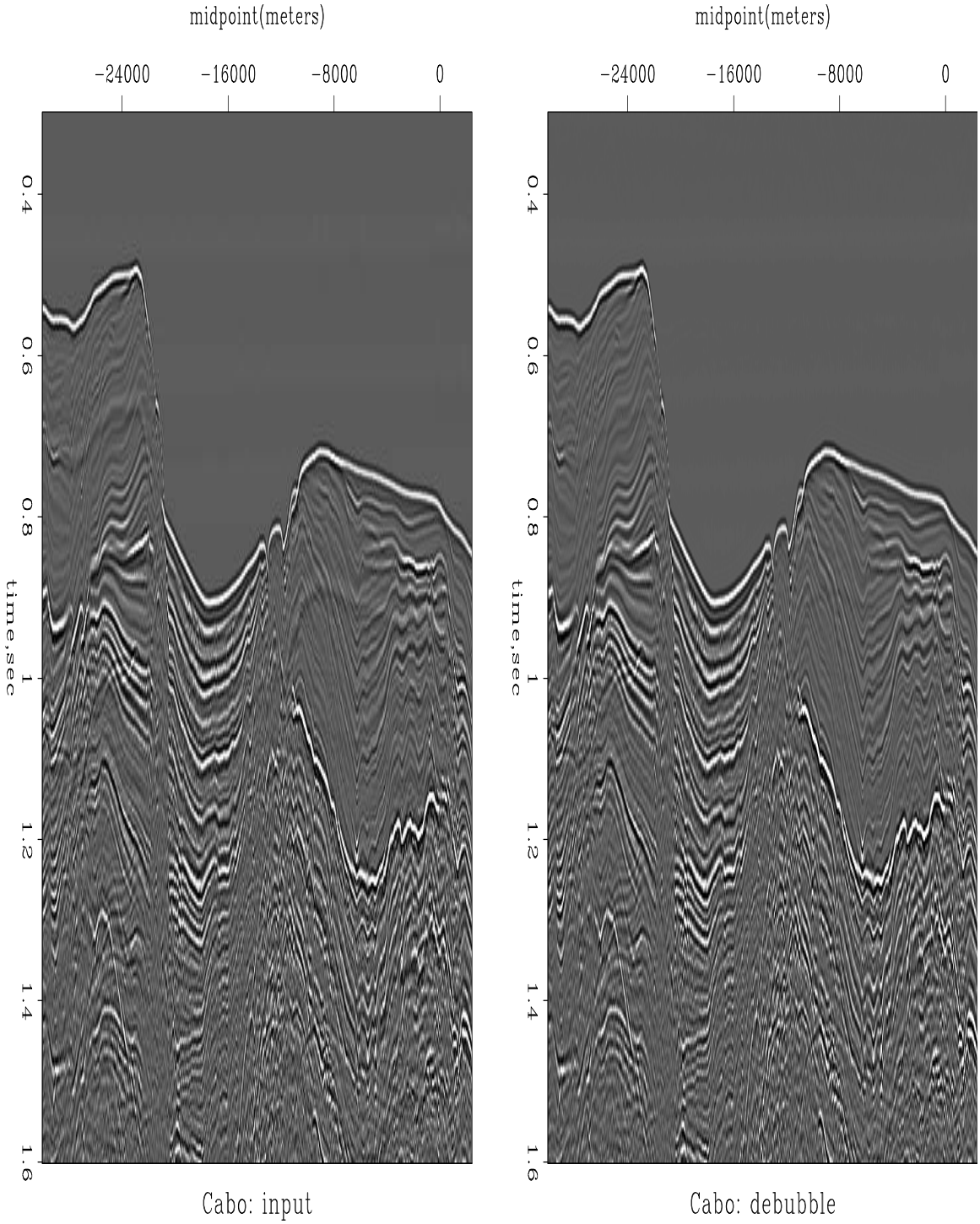


Figure 8: Bubble about 180ms after water bottom is obviously removed. Blink display shows bubble in many locations being lifted off the data. [**ER**] jon1/. Cabo-unitary

SELF DOCUMENTATION

```
# <in.H Rickdecon debubl=.06 ricker=.06 tresol=.01 shot=shot.H >out.H
# Inputs: d1=dt,debubl,ricker,tresol  must have identical units of time (seconds)
#         debubl=.06      =0 means no debubble
#         ricker=.06      =0 means no Ricker compliance
#         tresol=.01      time resolution, =0 for whitening decon
# Outputs: out.H is deconvolved data
# Outputs: shot(n)  centered at (n+1)/2.   Note n=2^^N >n1
#
# subroutine rickdecon( shot,n, data,n1,n2, d1,debubl,ricker,tresol)
```

APPENDIX

Subroutine `ftu` below is an ancient FT program from my book FGDP with conventional scaling consistent with Z-transforms. Data length must be a power of two. Subroutine `kolmogoroff` below was taken from my book PVI, converted from energy spectra to amplitude spectra.

While looking at the code you might notice that you could easily taper large lags to shorten your filter response. This might be useful should you want to crop off downgoing multiples from your source waveform. It could also be helpful when you have insufficient data to be estimating long source waveforms.

Beginning from the spectrum `cx(n)`, below is the code that makes the filter, also `cx(n)`.

```
subroutine kolmogoroff( n, cx, dt, debubl,ricker,tresol)  # Spectral factorization.
real                           dt, debubl,ricker,tresol, weight, tau
                                                  # Adapted from PVI, converted energy-->amplitude
integer i,              n                         # input:  cx = amplitude spectrum
complex cx(n)                                     # output: cx = FT of min phase wavelet

do i= 1, n
        cx(i) = clog( cx(i) )

call ftu( -1., n, cx)
do i= 2, n/2 {                      # Make it causal changing only the odd part.
        cx(i)     = cx(i) * 2.
        cx(n-i+2) = 0.
        }

## BEGIN weighting u
tau  = dt;  i=2;   while ( tau < debubl) {
        weight = sin( .5 * 3.14159265 * tau/(debubl+1.e-20))**2
        cx(i)     = cx(i)     * weight
        cx(n-i+2) = cx(n-i+2) * weight
        i = i+1;    tau = tau + dt
        }
tau  = dt;  i=2;   while ( tau < tresol) {
        weight = sin( .5 * 3.14159265 * tau/(tresol+1.e-20))**2
        cx(i)     = cx(i)     * weight
        cx(n-i+2) = cx(n-i+2) * weight
        i = i+1;    tau = tau + dt
        }
tau  = dt;  i=2;   while ( tau < ricker) {
        weight = sin( .5 * 3.14159265 * tau/(ricker+1.e-20))**2
        eve = (cx(i) + cx(n-i+2))/2.
```

```
        odd = (cx(i) - cx(n-i+2))/2.
        odd = odd  * weight
        cx(i)     = eve + odd
        cx(n-i+2) = eve - odd
        i = i+1;    tau = tau + dt
        }
# END weighting u

call ftu( +1., n, cx)
do i= 1, n
        cx(i) = cexp( cx(i))
return; end


subroutine ftu( signi, nx, cx )
#   complex Fourier transform with traditional scaling (FGDP)
#
#               1         nx          signi*2*pi*i*(j-1)*(k-1)/nx
#   cx(k)  =  -------- * sum cx(j) * e
#              scale     j=1             for k=1,2,...,nx=2**integer
#
#  scale=1 for forward transform signi=1, otherwise scale=1/nx
integer nx, i, j, k, m, istep, pad2
real    signi, arg
complex cx(nx), cmplx, cw, cdel, ct
if( nx != pad2(nx) )    call erexit('ftu: nx not a power of 2')
do i= 1, nx
        if( signi<0.)
                cx(i) = cx(i) / nx
j = 1;  k = 1
do i= 1, nx {
        if (i<=j) { ct = cx(j); cx(j) = cx(i); cx(i) = ct }
        m = nx/2
        while (j>m && m>1) { j = j-m; m = m/2 }         # "&&" means .AND.
        j = j+m
        }
repeat {
        istep = 2*k;   cw = 1.;   arg = signi*3.14159265/k
        cdel = cmplx( cos(arg), sin(arg))
        do m= 1, k {
                do i= m, nx, istep
                        { ct=cw*cx(i+k);  cx(i+k)=cx(i)-ct;  cx(i)=cx(i)+ct }
                cw = cw * cdel
                }
        k = istep
        if(k>=nx) break
        }
return; end
```

Ricker wavelet deconvolution of Western Australia data in the radial/time domain

Taylor Dahlke and Jon Claerbout

ABSTRACT

Performing deconvolution in the radial-time (r=x/t,t) domain is a method more consistent with the theory of the standard convolutional model than is deconvolution in the offset-time (x,t) domain. In this work we transform marine 2D seismic line data into the (r,t) domain, perform deconvolution along each common velocity panel, and transform back to the (x,t) domain. We also perform deconvolution on the same data (with no transform) along each common-offset panel in the (x,t) domain for comparison. Comparing in the (x,t) domain shows us that only limited differences in illumination are accomplished, even at far offset. Analysis of the spectra also shows us similar source and receiver frequency notches in the wavelet extracted from the data using both types of deconvolution.

INTRODUCTION

An important objective in seismic imaging is to remove the acquisition response from our data so that the underlying earth response can be more clearly observed and interpreted. Deconvolution is the process of extracting the source wavelet from the data in order to reach this objective. While it has been common practice to perform deconvolution for a minimum phase wavelet, it has been shown to be both possible and beneficial to extract a Ricker source wavelet instead, so as to better match the ghosting phenomenon that is associated with marine seismic acquisition (Zhang, 2011). One problem that must be addressed is the non-stationarity of the source wavelet along the time axis when in the offset-time (x,t) domain. This non-stationarity occurs because of geometric spreading as well as attenuation along the raypath, which depends on parameters such as length of the raypath through different layers, and the spectrum change with angle caused by the ghost delay changing with angle.

Regardless, the non-stationarity of a source wavelet produces a time-variant spectrum (TVS) (Margrave, 1997). Methods to account for the time variance of the spectrum in deconvolution (such as windowing in time) can be computationally expensive, and can create "stitching" artifacts from recombining data windows into a final image. In this paper we demonstrate that the raypaths represented by each trace in the radial-time (r=x/t,t) domain have similar path lengths in each layer and similar angles of incidence, which means the attenuation along each trace is causal, rather than non-linear as in the (x,t) domain. The causal case is closer to the standard convolutional model, and performing deconvolution on the data in the (r,t) domain should result in the extraction of a better effective source wavelet. In this paper we will demonstrate how performing deconvolution in the (r,t) domain better fits the assumptions of the standard convolutional model. Next we will discuss the

pre-processing and deconvolution methods that we used. Last we will show the results of deconvolution in the (r,t) domain, and compare with the results of performing deconvolution on common-offset images in the (x,t) domain.

(r,t) Domain

Most forms of deconvolution arise from the assumptions in the standard convolutional model. This model is one-dimensional and assumes that seismic energy is transmitted along a single raypath and strikes reflectors at a normal incidence angle. These assumptions are regularly violated because of the constraints and realities of acquisition. Performing deconvolution in the (x,t) domain largely ignores these assumptions, with the raypath from each reflector varying widely in take-off angle and path length within each reflector bed. By converting to the (r,t) domain, the seismic energy in each trace shares a common down-going raypath, and the upward reflections travel parallel to each other within reflector beds (Figure 1). Furthermore, the take-off angles are the same for each reflector. The (r,t) domain geometry better satisfies the standard convolutional model assumptions, with the only true difference being that the upward raypath does not follow the original path of the downward raypath (Henley, 2004). However, assuming flat and horizontally uniform geology, we can think of the rays as following a common raypath, since the incidence angle and path length through each layer is the same. It should be noted that Figure 2 shows a case where the velocity varies with depth. For this paper our transform of $r = x/t$ assumes a constant velocity and ignores slant stacking, in order to take account of only the first order effects of the radial transform (Figure 1).

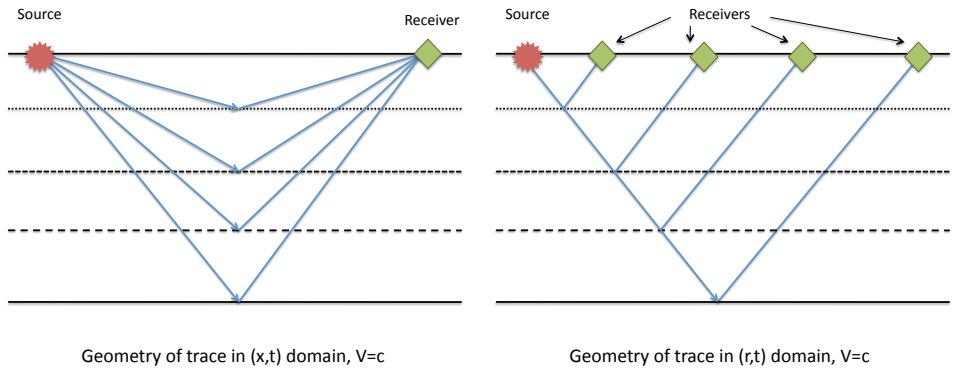


Figure 1: Diagram of raypath geometries in (x,t) and (r,t) domains. Constant velocity. [**NR**] taylor/. raygeom2

Ricker Wavelet

Marine data is typically collected using an air-gun source coupled with a receiver array that is towed at some depth below the water surface. While the source wavelet is often approximated as a Gaussian impulse, the wavelet recorded at the receiver is aggregated from this impulse reflection following four different ray paths. These ray paths include one direct reflection off the water bottom, one source and one receiver ghost reflection, and one

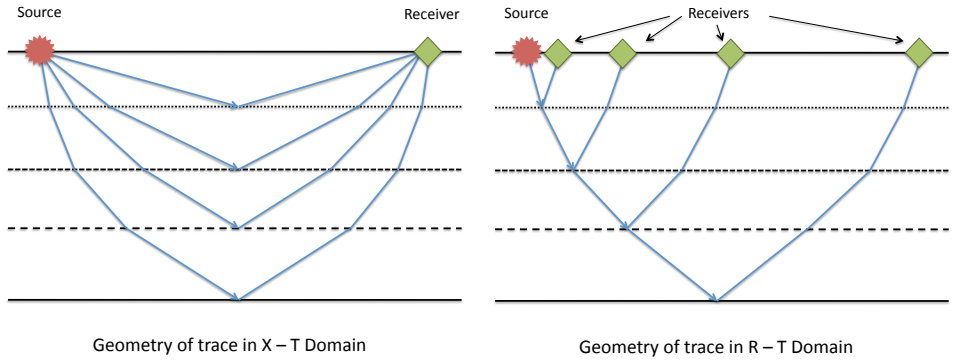


Figure 2: Diagram of raypath geometries in (x,t) and (r,t) domains. Velocity varies with depth only. [**NR**] taylor/. raygeom

double ghost reflection. These four reflections are superimposed, and the resulting pressure signal is captured at the receiver as a (opposite phase) Ricker wavelet (Liner, 2004).

While Figure 3 shows different ocean-bottom reflection points for the source ghost and receiver ghost, these reflection points are approximately the same for all four ray-paths when the ocean depth is much larger than the source/receiver depth. For this reason, the source and receiver ghost signals can be summed constructively, creating the main lobe of the superimposed wavelet observed at the receiver. Since the ghost reflections have opposite phase from the source impulse (due to the high impedance contrast at the water surface), these reflections destructively interfere with certain frequencies that depend on the depth of the streamer from the water surface. Notches are visible in the frequency spectra at the interference frequencies. These notch frequencies can be calculated with the following equation:

$$f_c = \frac{nc}{2z_c \cos\Theta}, \tag{1}$$

where z_c is the depth of the streamer, Θ is the direction of wave propagation from a reflection point (relative to the vertical), c is the velocity of the water, and n is an integer value. The information gathered from the observer's log of the survey data used in this paper indicates that the source depth was 5 m with +/- 0.5 m as an acceptable variance, while the receiver depth was recorded as varying from 6.2 - 7.4 m. Using equation (1), we found that the expected frequency notches at zero offset would range between 101 and 121 Hz for the receiver notch, and between 136 and 167 Hz for the source notch.

Deconvolution

The purpose of deconvolution is to separate the components of data due to acquisition effects from the components of the data that represent the response of the earth (Guitton, 2012). This requires that we create a filter that can perform this separation (Barak, 2012). When deconvolution is performed at increasing offset, the frequency notching from the

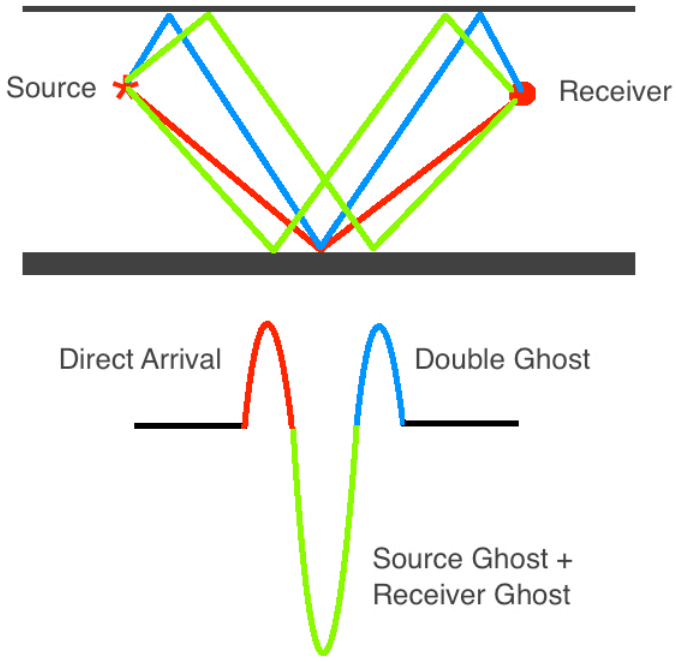


Figure 3: Diagram of Ricker wavelet construction. [**NR**] taylor/. ricker

ghost reflections influences the spectra of the source wavelet produced by our deconvolution (Figures 6(a) and 6(b)). Given our data $D(w)$, the deconvolved output can be defined as:

$$r_t = FT^{-1}\left[D_w e^{\sum_t u_t Z^t}\right], \tag{2}$$

where u_t = log spectrum variables, and $Z = e^{i\omega}$.

The method of deconvolution used on this data creates a filter based on parameters that act on the "lag-log" axis, as described by Claerbout (2013). If we begin with a wavelet $W(\omega)$ in the frequency domain, as defined using equations (3), (4) and (5) below, then we can derive equation (6) (where N is a positive integer), which describes the wavelet in the lag-log domain:

$$W(\omega) = e^{U(\omega)} \tag{3}$$

$$U(\omega) = \sum_{\tau=0}^{2048} u_\tau Z^\tau \tag{4}$$

$$Z = e^{i\omega\Delta t} \tag{5}$$

$$W(\tau) = \exp\left(\sum_{\tau=0}^{2^N} u_\tau Z^\tau\right). \tag{6}$$

This representation of the wavelet allows us to conceptually divide $W(\omega)$ into a composite of three wavelets, each defined by a range along the lag-log axis τ. The physical representation of these three wavelets can be thought of as:

$$e^{(A+B+C)} = e^A e^B e^C \tag{7}$$

$$e^{\sum_{\tau=1}^{2048} u_\tau Z^\tau} = e^{\sum_1^5} e^{\sum_6^{30}} e^{\sum_{31}^{2048}} \tag{8}$$

$$wavelet = (continuity)(Ricker)(bubble) \tag{9}$$

APPLICATION TO WESTERN AUSTRALIA DATA

The Chevron Western Australia data set that we deconvolved was collected from a marine 2D line survey, in about 950 m of water. Just over seven seconds of data with two-millisecond sampling was recorded for each of the 564 offset locations, each with a spacing of 12.5 m. In total, 1810 shots were recorded.

Preprocessing

Because of the high levels of sea-swell noise in the original data, one of the pre-processing steps was velocity filtering (Figure 4). A velocity filter is a far more appropriate method of eliminating this noise than a low-cut frequency filter, which would eliminate the sea-swell as well as compromise desirable components of the signal spectrum. The velocity filter we used was designed to mute all events that had a move-out velocity lower than the acoustic speed of salt water (1500 m/s). The filter also implemented a transition zone from the mute region of the velocity space, which was necessary to suppress the edge artifacts that occur when the filter has a hard "zero-to-one" filter boundary. This soft transition extended from the slope edge along the ω axis away from the origin, in order to preserve the zero weight at "zero" frequency.

DISCUSSION

Deconvolved Data

Once the velocity filter was applied to the data, both (x,t) and (r,t) domain deconvolution could be performed. For (r,t) deconvolution, the data was first transformed from the (x,t) domain to the (r,t) domain (Figure 5), and then deconvolution was performed along each common-angle panel. Likewise, in the (x,t) domain, deconvolution extracted the source wavelet from each common-offset panel. Since the geology is generally consistent horizontally, deconvolution along the common-offset and common-angle panels creates a filter that effectively averages the source wavelet across all shots. In order to compare the two methods, the (r,t) image cube was transformed back into the (x,t) space after deconvolution using a conjugate gradient solver.

Theory suggests that the low-velocity (low r value) traces will be similar to the near-offset traces, while the (r,t) transform differs most from the (x,t) at far offset. As a result, we expect to see the most significant difference between deconvolutional approaches in the

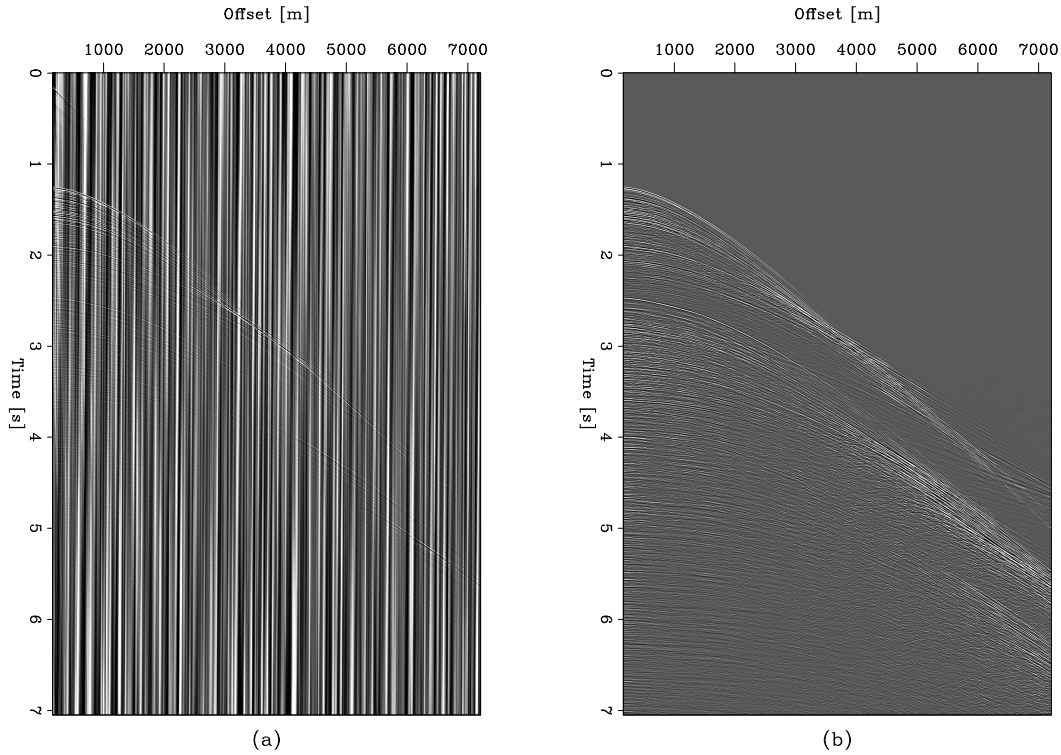


Figure 4: (a) Raw data input. (b) Velocity-filtered data. [**ER**] taylor/. compare1

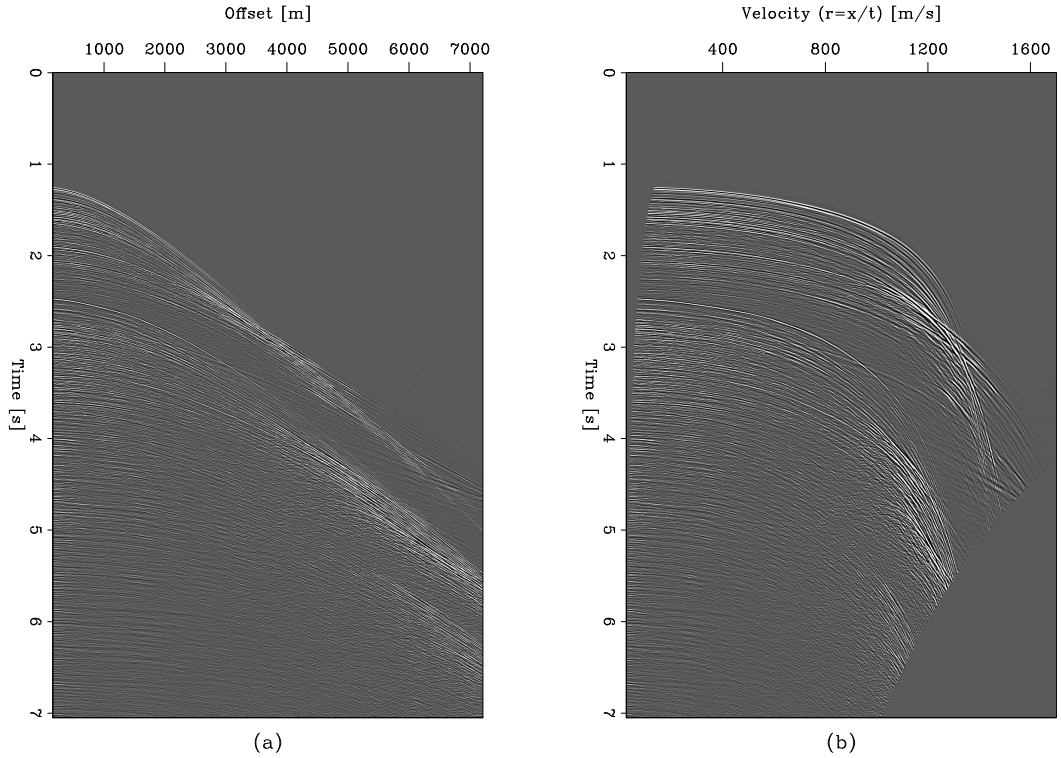
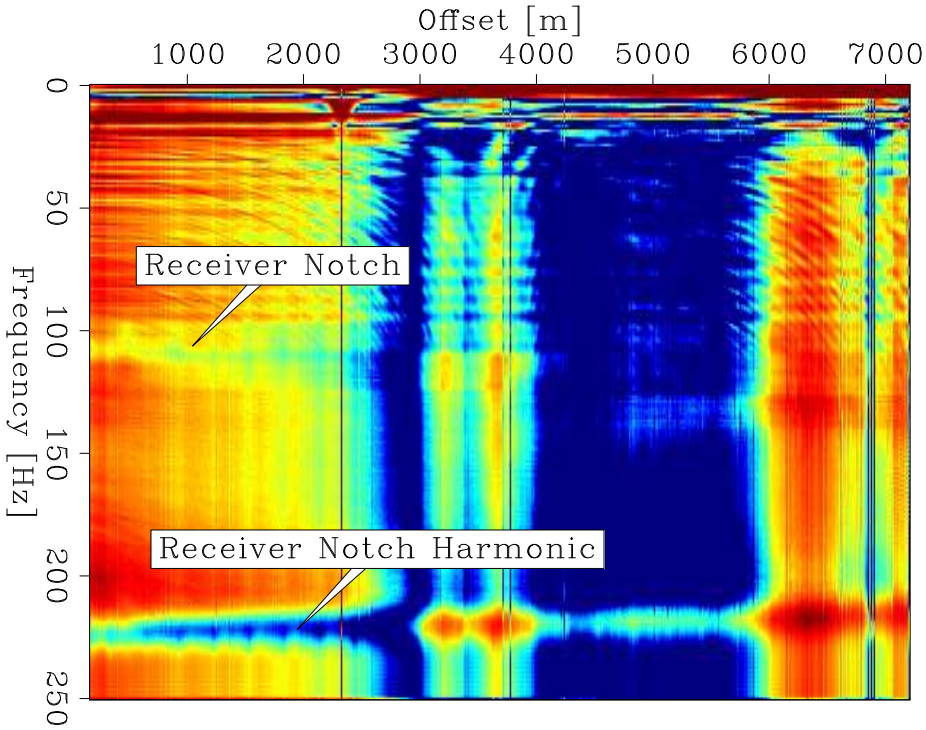
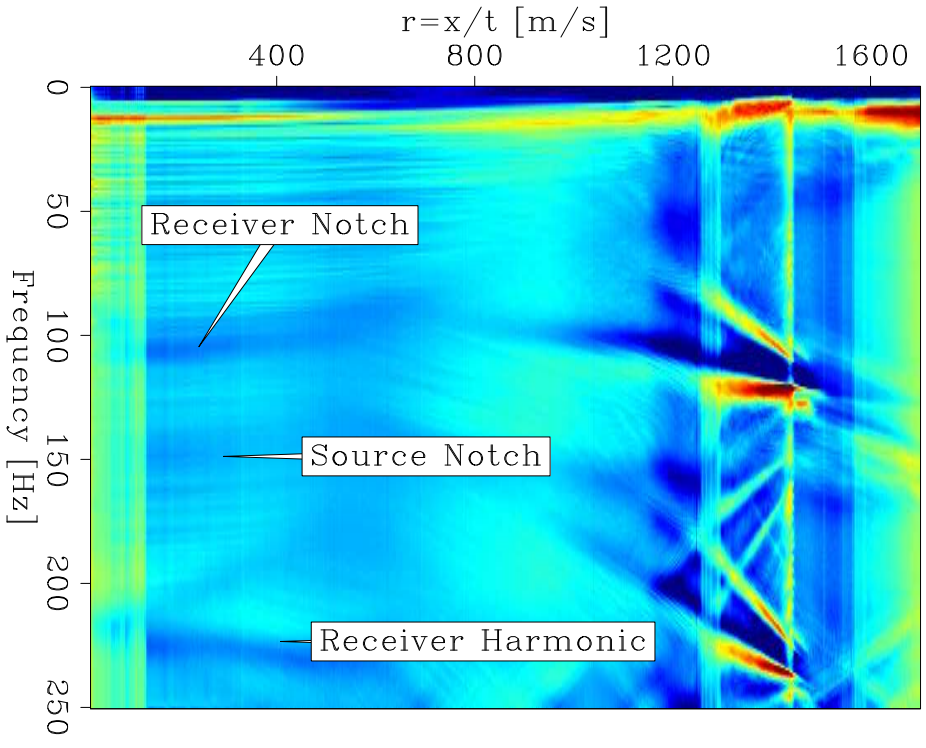


Figure 5: (a) Velocity-filtered data input. (b) After radial-domain transform. [**ER**] taylor/. compare3



(a)



(b)

Figure 6: Extracted wavelet spectra for: (a) (x,t) deconvolution, and (b) (r,t) deconvolution. [**ER**] taylor/. off-spec-wavelet,rad-spec-wavelet

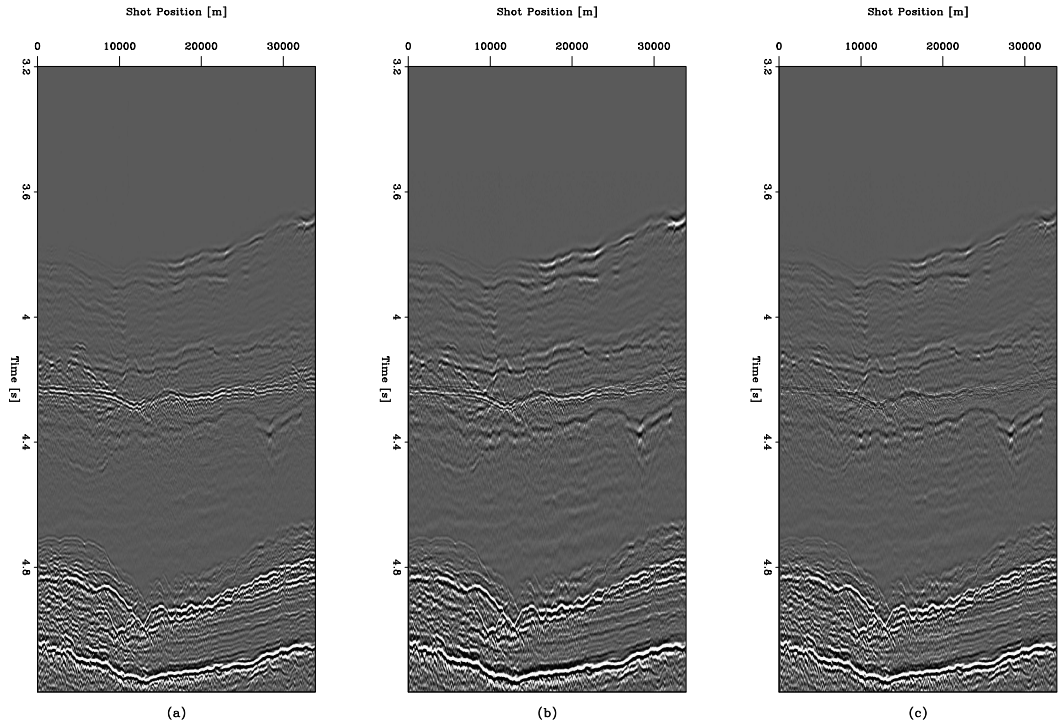


Figure 7: (a) (x,t) deconvolution: Dd. (b) Operator response: $R^{-T}R^{T}d$. (c) (x,t) deconvolution $(R^{-T}+\epsilon)DR^{T}d$. Offset = 6000 m. [**CR**] taylor/. compare-decon-og1

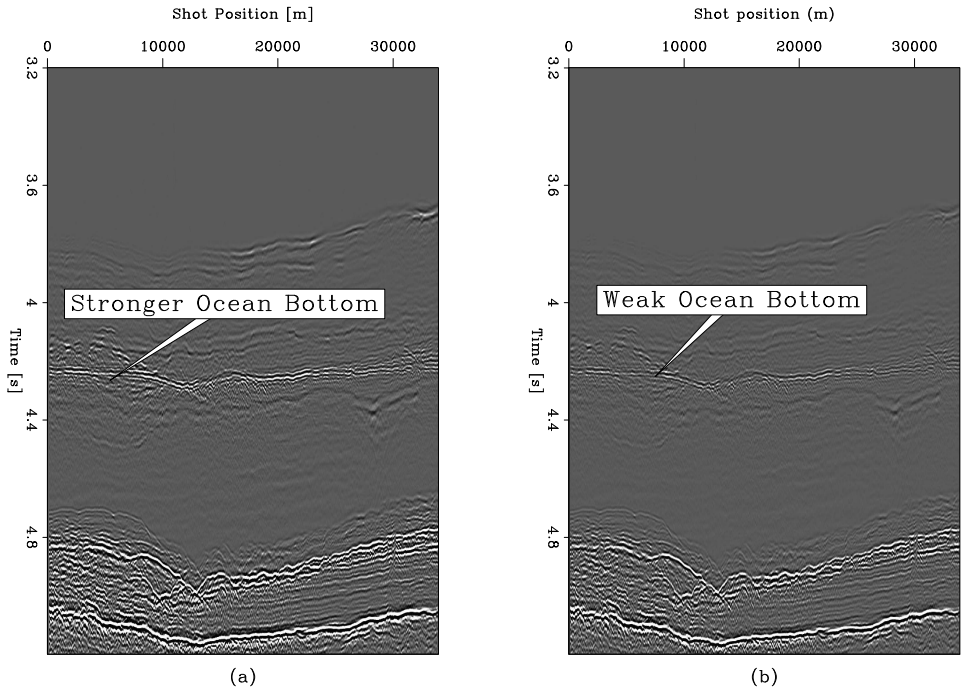


Figure 8: (a) (x,t) deconvolution: Dd. (b) Corrected (r,t) deconvolution: $R^{-T}DR^{T}d$. Offset = 6000 m. [**CR**] taylor/. diff1

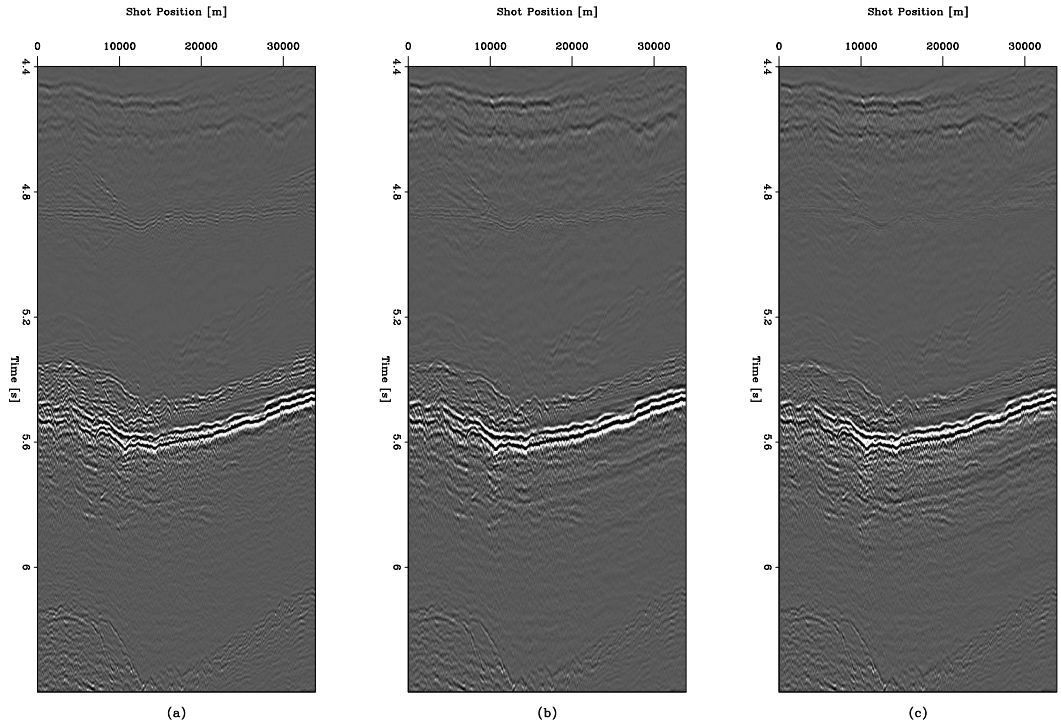


Figure 9: (a) (x,t) deconvolution: Dd. (b) Operator response: $R^{-T}R^{T}d$. (c) (x,t) deconvolution $(R^{-T}+\epsilon)DR^{T}d$. Offset = 7000 m. [**CR**] taylor/. compare-decon-og3

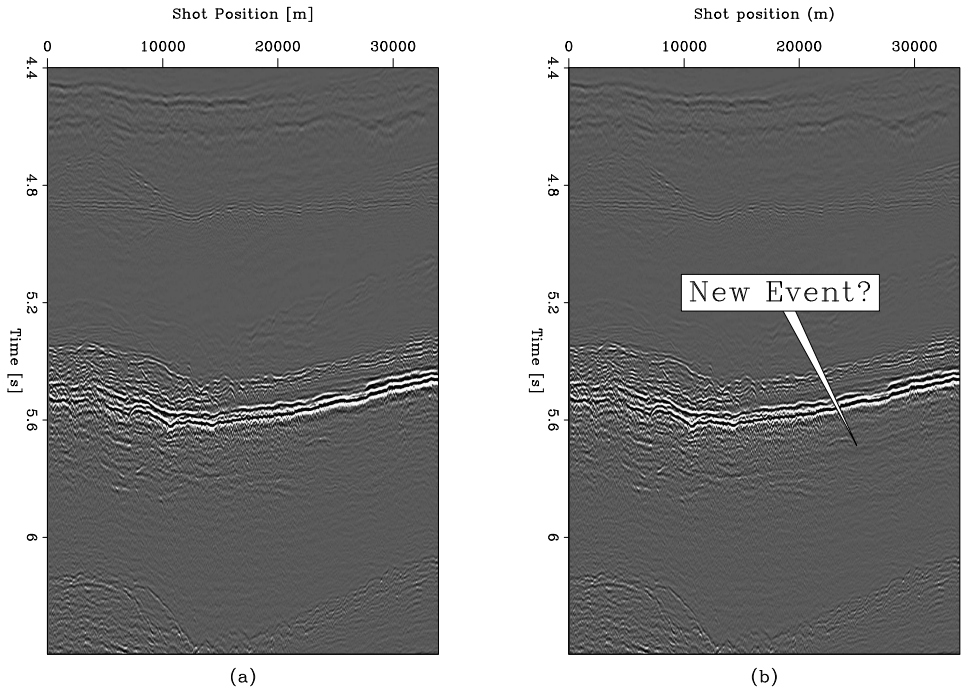


Figure 10: (a) (x,t) deconvolution: Dd. (b) Corrected (r,t) deconvolution: $R^{-T}DR^{T}d$. Offset = 7000 m. [**CR**] taylor/. diff3

far-offset areas of the shot gather, which is where the raypaths for the reflection events differ the most. For this reason, Figures 7 and 9 show offsets at 6000 m and 7000 m, respectively.

At far offset, deconvolution in the (x,t) domain is dominated by the stronger energy of the ocean-bottom reflection and head wave. This means the extracted wavelet will be best matched to this earlier, stronger energy. However, Figure 5 shows that the (r,t) transform shapes the reflection energy such that most of the traces in the new domain have similar time-lengths of reflection energy (for traces less than about r=1000 m/s). This means that for more traces, the first arrival/far-offset energy does not dominate in the deconvolution as much as it does in (x,t) space. The results of this are clear in Figures 10 and especially 8, where the earlier events in the (r,t) case are noticeably weaker than in the (x,t) case.

In side-by-side comparison, Figures 7 and 9 show that the image resulting from the response of the (r,t) transform is very similar to the final (r,t) image. When we correct for the response in the image due to the (r,t) transform and inversion, we get the results shown in Figures 8 and 10. These corrected (r,t) results confirm the general similarity between the (r,t) and (x,t) deconvolution methods that is implied in Figures 7 and 9.

These corrected (r,t) images with the operator response approximately removed are mathematically expressed by $I = F^{-1}DFd$, where $F = L'$ is the transform operator from (x,t) to (r,t) (thus the adjoint of L) and D is the deconvolution operator. However, since we use conjugate gradient methods to iteratively solve for the transform, what we really have is $I = (F^{-1} + \epsilon)DFd$, where ϵ represents the inherent error introduced from iteratively solving for the inverse of our forward operator. The following equations show how the effect of this error can be approximately removed. We make the assumption that $\epsilon DFd \approx D\epsilon Fd$, which is true when either ϵ or D is close to identity. Because the differences between the (r,t) deconvolution results and the operator response image is quite small, we allow the assumption of $D \approx I$.

$$
\begin{aligned}
I_{\text{response}} &= (F^{-1} + \epsilon)Fd \\
&= (F^{-1}Fd + \epsilon Fd)
\end{aligned}
\qquad\qquad
\begin{aligned}
I_{\text{original}} &= (F^{-1} + \epsilon)DFd \\
&= (F^{-1}DFd + \epsilon DFd)
\end{aligned}
$$

$$
\begin{aligned}
I_{\text{residual}} &= d - I_{\text{response}} \\
&= d - (F^{-1}Fd + \epsilon Fd) \\
&= -\epsilon Fd
\end{aligned}
\qquad\qquad
\begin{aligned}
I_{\text{correct}} &= I_{\text{original}} + DI_{\text{residual}} \\
&= (F^{-1}DFd + \epsilon DFd) - D\epsilon Fd \\
&\approx F^{-1}DFd
\end{aligned}
$$

When we examine the differences between the common offset images at far offset (Figures 8 and 10), the improvements in illumination are virtually unobservable. In Figure 10 some new information is visible, but it is not clear whether this is in fact a hidden reflector that is now exposed, or if it is an artifact of the deconvolution. On the other hand, Figure 8 seems to have less illumination in the indicated areas. This reflector is the ocean bottom, and lies almost completely along the radial trace at water velocity. In the radial domain, any reflection that travels at water velocity will have close to zero frequency, and the prediction error filter in our deconvolution dampens it accordingly.

Extracted Wavelet Spectra

The extracted wavelet from the (r,t) deconvolution (Figure 6(b)) displays the expected frequency notches similarly to the wavelet extracted using (x,t) deconvolution (Figure 6(a)). While the (x,t) results only seem to show a receiver notch and its harmonic, the (r,t) results show the source frequency notch slightly better. One particularly interesting observation is that the "harmonic" receiver notch seems to curve in the opposite direction of the primary receiver notch. We suspect that this non-linearity is due to the (r,t) transform. Another feature found in the (r,t) results is the presence of numerous crossing "slants" at larger radial values (beyond the approximate critical angle), which remain to be thoroughly understood. More consistent with our observations is the primary receiver notch, which curves towards lower frequencies at increasing radial values. Since increasing radial values mean higher take off and arrival angle, this means longer ray paths through each layer, which translates into the attenuation and loss of higher frequencies that we observe.

FUTURE WORK (JON)

Different angles of propagation should require different deconvolution filters. This is mainly because the effective depth of burial of the gun and the hydrophones varies with propagation angle. The air bubble resonance, however, is an effect that we expect to be independent of propagation angle.

We see two straight forward ways to deal with the angle effect on deconvolution filters. The first (begun in this paper) is to use radial traces. The second (which we intend to explore) is to multiply each shot gather by a family of wedge-shaped weighting functions. One such weighting function is shown in Figure 11.

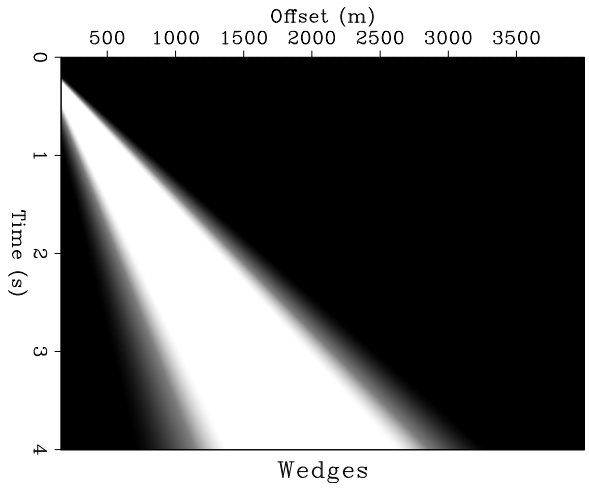


Figure 11: Wedge-shaped weighting function to select a zone of expected take-off and arrival angles. [**NR**] taylor/. wedge

More elaborate dip decompositions are possible (such as decomposition in the Fourier domain), but we shall begin with these two radial traces and a spectrum of weights such as in Figure 11.

Looking forward, the radial traces allow for rapid changes in angle but achieve statistical stabilization by averaging over shots (or midpoints). On the other hand, decomposition by zones has lower resolution by angle, but it achieves some statistical stabilization by averaging

over many traces in a gather.

CONCLUSION

By performing deconvolution in the (r,t) domain, the assumptions implicit in the standard convolutional model are satisfied far better than with deconvolution in the (x,t) domain. However, results from 2D field data indicate very limited differences in the amplitude, polarity contrast, and sharpness of events, and in some cases demonstrate a reduction of reflector illumination. As we predict with our theory regarding ghost reflections, we find in both cases that the frequency notches at about 110 Hz and 150 Hz are visible at the nearest offset and angle, with the spectra from the (r,t) domain extracted wavelet displaying these notches similarly to the (x,t) domain case. While in theory the (r,t) deconvolutional method would be expected to offer a more robust option for performing deconvolution, for this data set, the method does not seem to demonstrate any significant improvement in results.

ACKNOWLEDGMENTS

We want to thank Chevron for generously providing the data used in this paper. We also want to thank Antoine Guitton for the guidance and constructive criticism that he has provided throughout this work. He is also co-author of some of the programs used for the deconvolution steps.

APPENDIX (JON)

Air Gun Sources Lack Low Frequencies

Since raw marine data often shows sea swell with frequencies commonly as low as 0.1 Hz, it is tempting to conclude that low-frequency signal is present. It is not. The recorders record it, but the guns do not send it. The only signal we have at low frequencies is what we might be able to unscramble from the ambient noise. To see what is in the Chevron Australia data set, we applied a high-cut filter at 5 Hz. Figure 12 shows hardly anything left. At a little lower cutoff frequency it would not be possible to detect hyperbolic continuity.

For those of us hoping to estimate earth impedance, this is a sad state of affairs. One of us (Jon) advocates investigating the water cannon as a source. With such a source, one shoots water up into the sky. The water falls back to the surface only after data has been recorded.

If our goal is impedance estimation, we must have low frequencies. What are our alternatives?

1. Deeper guns
2. Water cannon source
3. Learn to use ambient noise as source. (Unfortunately, we don't have a surface receiver array suitable to destruction of water waves from all sides.)

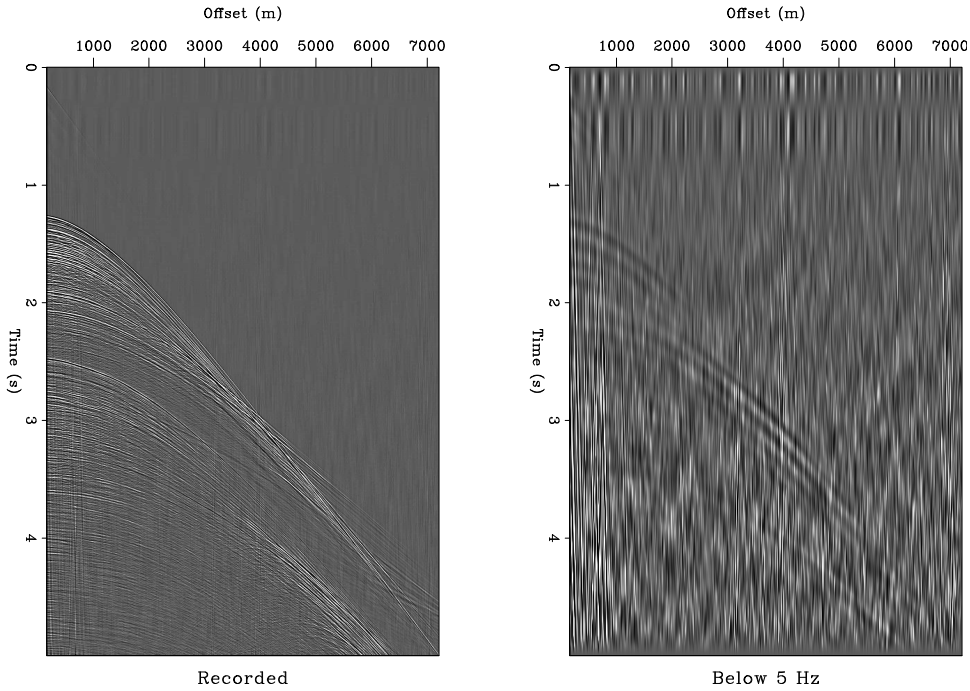


Figure 12: Data quality is poor at low frequencies. If the cut-off frequency is below about 5 Hz it is not possible to see the hyperbolic continuity. [**NR**] taylor/. lowest

REFERENCES

Barak, Claerbout, G. R., 2012, Decon in the log domain - practical considerations: SEP, **148**, 19–34.

Claerbout, 2013, Low-velocity deconvolution.

Guitton, C., 2012, Six tests of sparse log decon: SEP, **148**, 35–48.

Henley, 2004, More processing in the radial trace domain: 2004 CSEG National Convention: Great Explorations - Canada and Beyond, **148**, 1–4.

Liner, C. L., 2004, 3D seismology, 2 ed.: PennWell Books.

Margrave, S., 1997, An approach to nonstationary deconvolution.

Zhang, Claerbout, G., 2011, A new bi-directional sparse/spike deconvolution method that overcomes the minimum phase assumption: SEP, **143**, 271–282.

t-squared gain for deep marine seismograms

Jon Claerbout

ABSTRACT

Kjartansson's all-purpose gain function t^2 is updated for deep water and large offset. We examined several deep-water gain possibilities and finally nominate $(t - t_e + \Delta t)t$. The water path time (first-earth-arrival time) is t_e. The thickness Δt is defined as that of a $Q = 100$ layer required to convert a mathematically infinite-band source signal to a realistic band. A suitable value for Δt is 350ms. Using the nominated gain function reduces the need for plotting at various clip values, muting, or AGC. We wish we had code for routinely identifying t_e. Present code assumes the vertical travel time depth is known and assumes normal moveout predicts water-bottom travel time for non-zero offset. Program and code are ready for your use.

INTRODUCTION

The goal is to scale deep marine data so it is visible at all times and offsets. Ideally the scaling function, like Kjartansson's t^2 is independent of any parameters, but easily known parameters like offset and water-depth are acceptable. In zero-order seismology rays propagate vertically. In first-order seismology rays propagate on straight lines defined by shot and receiver locations. In second-order seismology one includes RMS velocity. Here we are looking for something like a first-order calculation of expected wave amplitude so we may divide data by it bringing a screen full of data to visibility everywhere at once. What we are looking for is a general purpose scaling function of time and space not tailored to a particular data set, but useful for most data sets. Our goal is to reduce need for AGC, for clip, and for mute. Besides general purpose data viewing, this scaling may be a useful weighting function for inverse problems.

Kjartansson (1979) points to the mathematical function t^2 as a general purpose gain function for seismic data. A simple mathematical model suggests his function. His model assumes constant Q absorption along the entire ray path. But there is no absorption on the water path to bottom which can be quite long in deep water, especially at wide offset. That fact motivates development of the improved gain function I find here. The water-path time t_e is a function of both shot and geophone locations. To find t_e, here I make simplifying assumptions that vertical water depth is known and normal moveout predicts t_e with offset. I'd rather have a code that for t_e estimates the time the first event reaching the bottom arrived back.

REVIEW OF KJARTANSSON'S T-SQUARED

One power of t arises from spherical spreading. Energy spreads out on a sphere whose area grows as t^2. But we are interested in amplitude, not energy, which implies the square root of t^2 namely t.

The other power of t comes from the constant Q model. Seismograms have their highest frequencies at early times. These damp out leaving only lower frequencies at later times. So data may be gained more at late times. Beginners often believe the way to compensate for such absorption is with exponential gain, but that is wrong because exponentials describe sinusoidal waves, not the broad spectral band that our data has.

The most basic absorption law is the *constant* Q model. According to it, energy diminishes in proportion to the number of wavelengths in space, or in proportion to the number of periods in time, the factor of proportionality being $1/Q$. For a downgoing wave the absorption is proportional to the frequency ω and proportional to time in the medium which is the distance z divided by the velocity v. Altogether the spectrum of a wave passing through a thickness z will be changed by the factor $e^{-|\omega|(z/v)/Q}$ where Q is called the Quality factor of the medium.

We may define spectral bandwidth by setting $e^{-|\omega|(z/v)/Q}$ to be some arbitrary cutoff constant, say, e^{-3}.

$$.05 \approx e^{-3} = e^{-|\omega|_{\text{cutoff}}(z/v)/Q} \tag{1}$$

$$\omega_{\text{cutoff}} = 3Qv/z = 3Q/t \tag{2}$$

This says the later you look on a seismogram, the narrower the spectral bandwidth. You compensate for this by gaining your data by another power of t, hence Kjartansson's t^2.

Fortuitously, this result is independent of velocity and the cutoff threshold e^{-3}; and it depends on Q only as a scaling factor (which merges itself with the usually-irrelevant plot scaling factor). As absorption $1/Q$ transits from zero to nonzero, the amplitude (plane wave) damping transits from 1 to Q/t. Not easy to connect those two functions!

NON-ZERO OFFSET AND DEEP WATER

For waves in the water path, regardless their direction of propagation we wish to delay the absorption effect until the time the waves first enter earth sediment t_e. My first guess at the gain rule $G(t)$ was this:

$$G(t) = \begin{cases} t & \text{for } t < t_e \\ t^2/t_e & \text{for } t > t_e \end{cases} \tag{3}$$

Notice $G(t)$ is continuous at $t = t_e$, has t behavior before, and t^2 after. I believed equation (3) for some weeks, but when time came to write about it I found I could not derive it or defend it.

We want a gain function that grows most rapidly where the energy dissipation is strongest — just below the water bottom. It feels like the gain function needs the time of the in-earth ray path $t - t_e$. So I considered this:

$$G(t) = \begin{cases} 0 \times t & \text{for } t < t_e \\ (t - t_e)t & \text{for } t > t_e \end{cases} \tag{4}$$

Equation (4) is also continuous at $t = t_e$ because it is zero there. The idea is that there is infinite bandwidth in the water path and at the water bottom reflection so waves there should be multiplied by zero. Observers would be upset if they could no longer see the water bottom! Head waves either? Well, they'd see both the head waves and the water bottom if they chose their t_e a little earlier—but how much?

Decon-friendly gain

Antoine Guitton suggested a "decon-friendly" gain function. By being non-zero before time t_e, it does not offer a hiding place for non-causal optimization decons to hide information before t_e.

$$G(t) \;=\; \begin{cases} 1 & \text{for } t < t_e \\ t^2/t_e^2 & \text{for } t > t_e \end{cases} \tag{5}$$

Unfortunately, on a seismic section where water depth is increasing along the traverse, water bottom arrivals are not getting their basic geometrical t gain.

Continuity gain function

Shuki Ronen came up with a gain function that starts from linear gain in the water, then converts to parabolic (polynomial containing t^2) in the sediment, with the condition that the two functions match in slope as well as value at t_e.

$$G(t) \;=\; \begin{cases} t & \text{for } t < t_e \\ t + (t-t_e)^2/t_e & \text{for } t > t_e \end{cases} \tag{6}$$

This gain function is linear before $t = t_e$ and quadratic after. It is continuous in value and derivative at $t = t_e$. There is no problem in having a continuous derivative, but there seems no reason for the derivative to be continuous.

An all-purpose source spectrum

Equation (4) seems best in theory, but worst in practice. Data is multiplied by zero at the water bottom. We won't see data there! Furthermore, we don't know exactly where the water bottom is. Getting started I simply picked t_e a little early. That expedient is barely suitable for personal software and wholly unsuitable for shared software.

After pondering this conundrum some weeks I came to realize it's connected with the assumption that the wave begins with infinite bandwidth. What bandwidth should I assume for the initial shot waveform? Rephrasing the question, suppose we begin with infinite bandwidth. How much earth of $Q = 100$ must that infinite bandwidth propagate through to reach the kind of bandwidth we normally see? Let us say we would like the spectrum to drop down to 5% at 150Hz.

$$.05 \;\approx\; e^{-\pi} = e^{-2\pi f \Delta t/Q} \tag{7}$$

$$\Delta t \;=\; Q/2f = 100/300 \;\approx 350\text{ms} \tag{8}$$

That's it in a nutshell. Put an infinite bandwidth into 1/3 second of earth and get out our defaulted standard seismic source spectrum. (Not a big reward for the builder of that infinite bandwidth source!) I nominate this default gain for our deep marine data:

$$G(t) \;=\; \begin{cases} 0 & \text{for } t < t_e - \Delta t \\ (t - t_e + \Delta t)t & \text{for } t > t_e - \Delta t \end{cases} \tag{9}$$

By this logic Kjartansson's all-purpose gain t^2 would be improved by $(t + \Delta t)t$.

RESULTS

Figure 1 shows the result on a shot gather from offshore western Australia. This result is delightful showing more improvement than expected. We might wish to incorporate this gain as an option in our routine plot facilities. Signal at inner offsets has grown to be comparable to that at wider offsets. There are dominating multiples in the middle, but primaries may be found from top to bottom. Critical-angle events formerly dominated and still do, but much less so. Back scatter events at 5s are now more evident. The need for subsequent mute, or AGC, or revised clip is now reduced or eliminated. Subsequent processes such as velocity analysis, migration, and stack should change in subtle, perhaps useful ways.

Figure 2 shows a near-trace section. Results are similar, but here the new deep-water gain is better in a simple way — stronger at early time. I had anticipated this good result, but had not anticipated the even better result in Figure 1.

At wide offset in Figure 1 we see events beyond the water asymptote, both head waves and deep reflections. Here is why: To avoid suppressing such good signal, I boosted the mute velocity parameter to 2000m/s. Formerly we might play with gain via the parameter `tpow`=γ in t^γ where we had no physical model for γ. Now we may play with gain where we have two physical parameters: (1) vertical travel time to water bottom `tau`=τ, and (2) `velocitymute` v, for a rough guess of first earth arrival time t_e as a function of offset.

Better yet should be an along-path first-arrival-time finding code. That would simplify our lives leaving only the minor sensitivity to the spectrally dependent Δt. Not only would we see our data better, but we might be more attuned to the notion that anomalous amplitudes have real meaning.

An expeditious approximation

Figure 3 shows the result of the expeditious approximation that the water path is only its horizontal distance ignoring the actual slanted path. The first earth reflection time t_e is not difficult to obtain, but it's not effortless either in view of the many complications arising in volume production. Since the offset is always known, this expeditious approximation $t_e = |x|/v$ is nearly effortless to install and use by default. Serendipitously, Figure 3 seems even better than the more accurate calculation in Figure 1! This result, however, is not likely to be typical. West Australia off-shore data is strongly dominated by multiples. The theory here is designed for primaries. Any gain function designed with multiples in mind must be customized according to their local strength. Examining the data here closely, we recognize weak primaries all throughout the range of time. Without the approximation the early primaries are gained to about the strength of the late ones. So the expedient result is not so desirable as it might first appear. Still, it is an expedient result, perhaps needed where depth varies within a survey and water-bottom tracking information is not available.

Test of the first-guess method

Figure 4 shows a test of the first guess method, equation (3). As expected, water bottom arrivals are very strong, dominating even the later multiples. The mathematical function

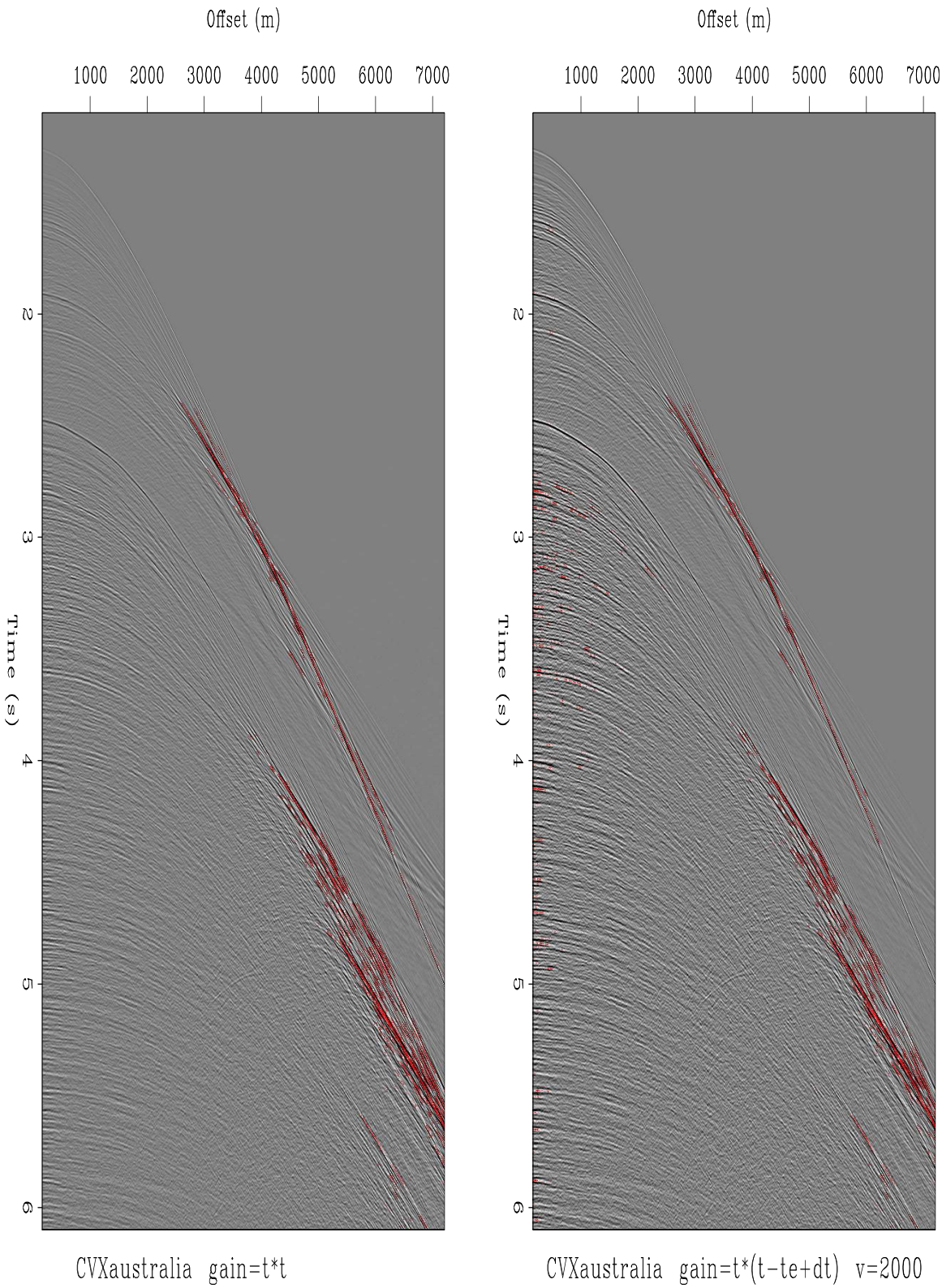


Figure 1: Chevron Australia shot gather gained by t^2 (left) and by $(t-\sqrt{1.2^2+x^2/v^2}+.35)t$ (right). These plots use our default clip percentile, 99%. If you see the color red, you see where it is clipped. [**ER**] jon2/. cvxAusG2

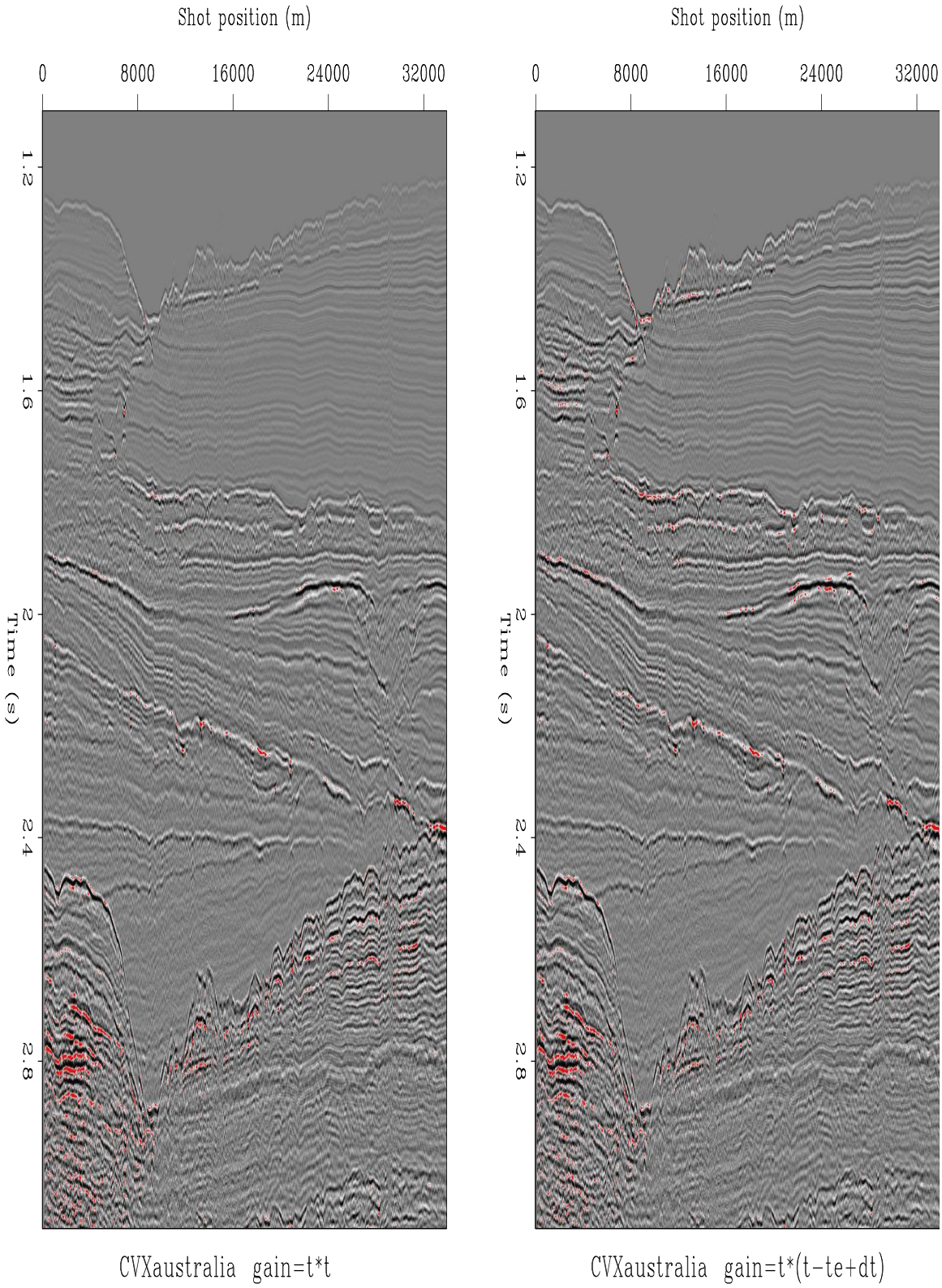


Figure 2: Chevron Australia near-trace section gained by t^2 (left) and by $(t - \sqrt{1.2^2 + x^2/v^2} + .35)t$ (right). The new deep-water gain is better at early time. [**ER**] jon2/. cvxAusS1

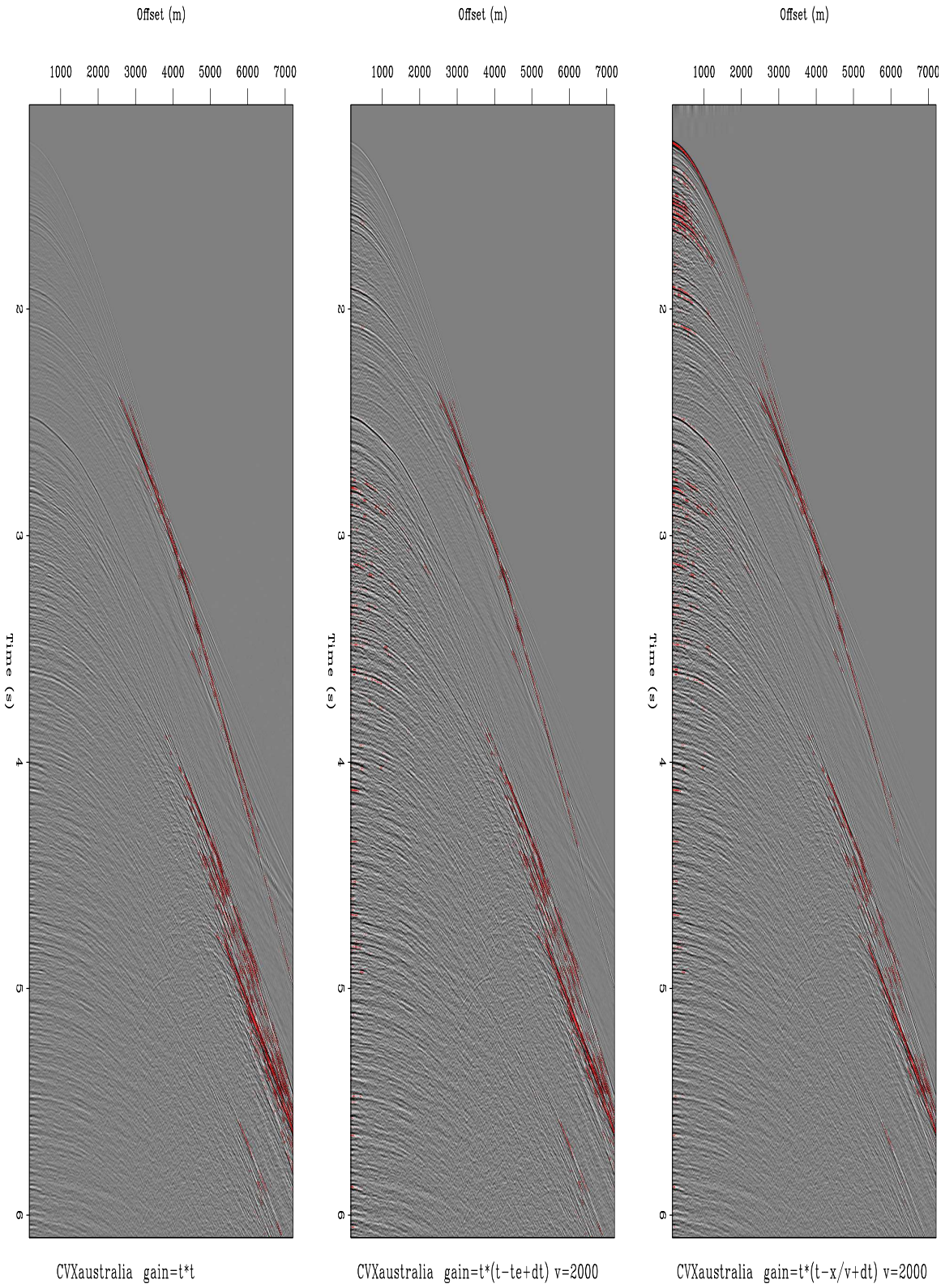


Figure 3: Test of expeditious gain: Kjartansson t^2 (left); Deep-water $(t - \sqrt{1.2^2 + x^2/v^2} + \Delta t)t$ (center); and Expeditious $(t - |x|/v + \Delta t)t$ (right). [**ER**] jon2/. cvxAusG3

t^2 has a powerful effect near $t = 0$ but a much weaker one near $t = t_e$. Data sets with weaker multiples should demonstrate more clearly the first guess weighting does not bring primaries into balance.

CODE DOCUMENTATION

We would like to acquire code that deduces the first arrival time, but so far we have not done it. Consequently we require knowledge of `offsetAxis`, a way to know in a hypercube which axis is the offset. Additionally, for the same reason, we require water depth `tau` and a mute velocity `mutevelocity`.

```
#  Deep Marine Gain
#  Dmgain <in.H > out.H tau=0 velocitymute=2000 offsetaxis=2 tspec=.350
#  Gain 3-D data by t*(t-te+tspec) where  te = sqrt(tau**2+x*x/velocitymute**2)
#       tau is travel time to water bottom
#       velocitymute is used to estimate travel time to water bottom
#       tspec=.350 = time before water bottom where mute ends and gain begins.
#       2-D data with offsetaxis=3 (section) requires n3=1 and o3 to be present.
```

CONCLUSIONS

1. The code is 35 lines long plus the above documentation. Located in /homes/sep/jon/res/-futterman/Dmgain.rst After I get some user feedback, we can think about installing it.

2. Testing should be done on data less dominated by multiples. The same makefile can be used.

3. If someone codes identification of t_e from the data itself, results should be better from several points of view. Then the code would work equally well on shot gathers as midpoint gathers.

4. The gain function should have an optional radial x/t dependence to account for critical angle strength at water velocity. Need a wave propagation theorist to tell me the equation to use!

5. Adam's good idea is that the required parameters might be estimatable from the scaled data by some optimization code. This could be our default, as `Tpow` is supposed to compute a `tpow` if you don't specify one.

REFERENCES

Kjartansson, E., 1979, Constant Q-wave propagation and attenuation: J. Geophys. Res., **84**, 4737–4748.

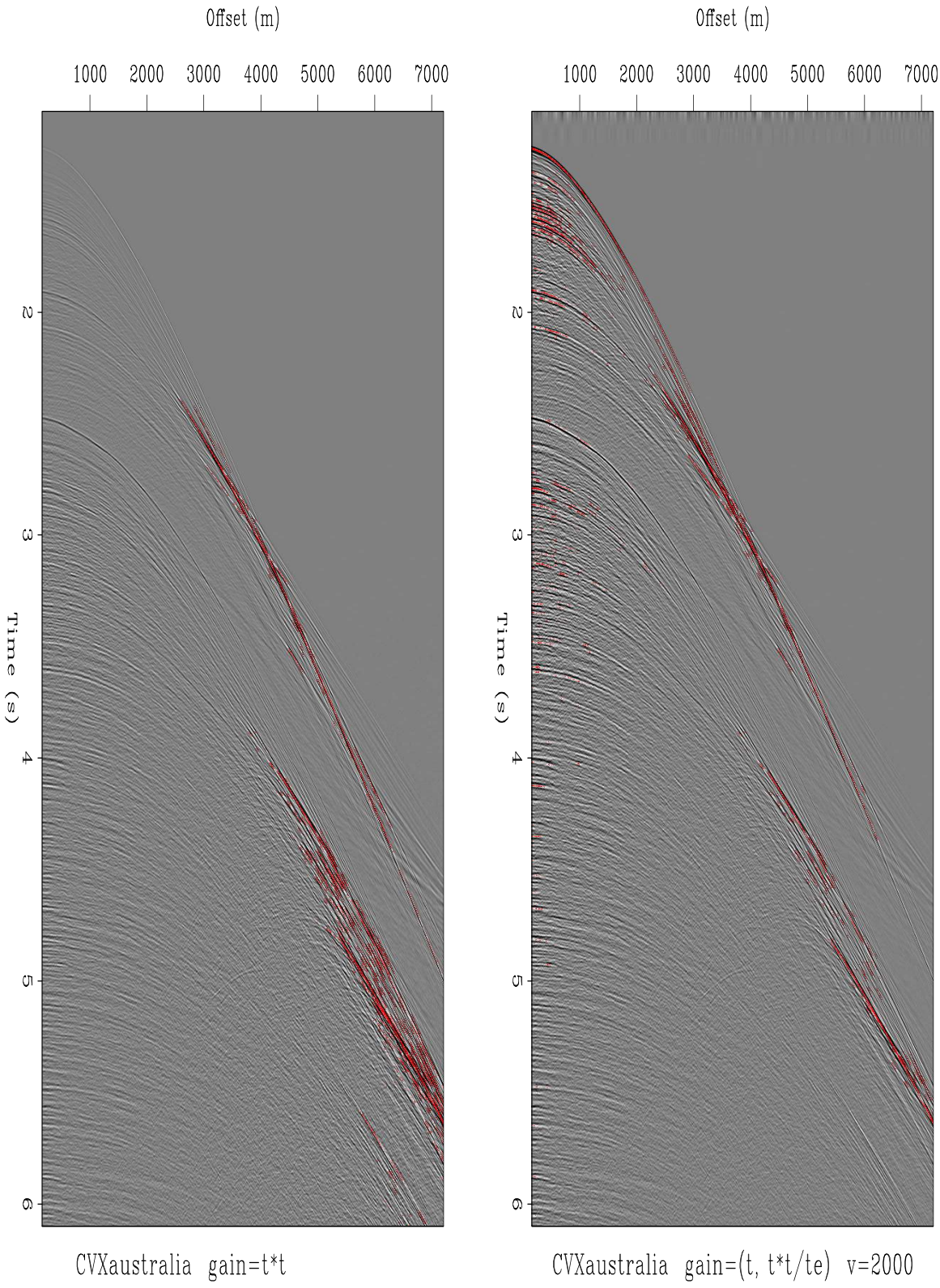


Figure 4: Chevron Australia shot gather gained by first guess weighting, t in the water path, t^2/t_e later. Early arrivals are too strong. [**ER**] jon2/. cvxAusG4

Tomographic full waveform inversion (TFWI) by extending the velocity model along the time-lag axis

Biondo Biondi and Ali Almomin

ABSTRACT

Convergence of full waveform inversion can be improved by extending the velocity model along either the subsurface-offset axis or the the time-lag axis. The extension of the velocity model along the time-lag axis enables us to linearly model large time shifts caused by velocity perturbations. The extension is based on a new linearization of the scalar wave equation where the extended-velocity perturbation is convolved in time with the Laplacian of the background wavefield. This linearization is accurate for both reflected events and transmitted events, and in particular for diving waves recorded at large offsets. The modeling capabilities of the proposed linearization enable the simultaneous inversion of reflections and diving waves even when the starting velocity model is far from being accurate. Numerical tests performed on synthetic data modeled on the "Caspian Sea" portion of the well-known BP model shows the global-convergence properties as well the high-resolution potential of the method.

INTRODUCTION

Conventional seismic imaging relies on a separation of scales between migration velocity model (long-wavelength components) and reflectivity (short-wavelength components). Figure 1 shows a simplified 1D graphical representation of the separation of scales concept. The black line represents the two disjointed wavelength ranges (mapped into corresponding temporal-frequency bands) and the consequent gap in information between long wavelengths and short wavelengths. This conceptual understanding leads naturally to a sequential approach for seismic imaging; the velocity model is estimated first, and then it is used as input to migration for imaging reflectivity. In current velocity-estimation practice, reflectivity is used only indirectly to measure the focusing power of the velocity model. The only important exceptions occur when migrated volumes are used to interpret boundaries of geobodies (e.g. salt bodies) and to estimate predominant dips in the geologic layering that are then used to constraint a tomographic velocity update.

The sequential imaging process is slowly being undermined by three long-standing trends in the industry: 1) acquisition of lower-frequency data, 2) imaging under complex overburden which requires higher-resolution velocity models to focus and correctly position reflectors, and 3) acquisition of longer-offset data. As the industry strives to widen the data frequency band at both the low and high end, the reflectivity band is extended at the low end, as graphically represented by the green line in Figure 1. The high end of the velocity band is also pushed upward (blue line in Figure 1) by the application of sophisticated tomographic methods that enable the estimation of the high-resolution velocity models needed to focus reflectors located under complex overburden. However, tomography (either ray

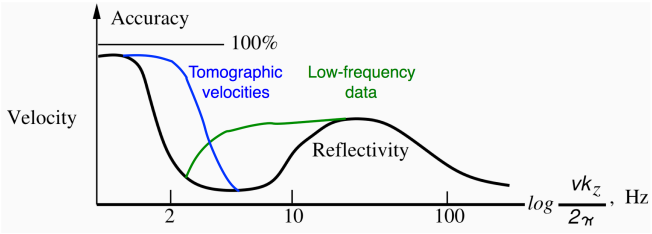


Figure 1: Simplified 1D graphical representation of the separation of scales in seismic imaging (black line) and how current industry trends are narrowing the gap between the estimation of long wavelengths and short wavelengths (blue and green lines). (Adapted from Jon Claerbout's *Imaging the Earth Interior*. [**NR**] biondo1/. iei-rely-new-ppt

or wave-equation based) is a more challenging task than migration, and thus often it falls short of providing the accuracy and resolution necessary to satisfactorily image the high frequencies in the data. The acquisition of longer-offset data enables the recording of diving waves and refracted arrivals that provide a complementary illumination of the velocity components in the crucial scale gap and blur the distinction between migration and tomography since they contain forward-scattering perturbations to transmission events.

As the information gap narrows, imaging methods that simultaneously estimate the velocity and reflectivity model by taking advantage of all the information in the data are becoming more attractive. The renewed interest in full waveform inversion (FWI) can be explained as an attempt to overcome the limitations imposed by the sequential imaging approach, as well as the availability of computational power sufficient for practical applications of FWI. FWI has been the most successful when applied to the low frequencies in the data (green line in Figure 1) to improve the velocity-model estimation needed to image the data high frequencies under complex overburden. FWI has been less successful in using the high-frequencies in the data to tomographically estimate the long-wavelengths in the model.

Since the eighties (see for example (Mora, 1989)) it has been recognized that FWI has both a migration component and a tomographic component. However, to ensure convergence of the tomographic component the recorded and modeled data must be almost in phase with each other; the rule of thumb being that the residual time-shifts must be shorter than the half period of the dominant frequency in the data. Bootstrapping the inversion by starting from the low frequencies may ameliorate the convergence problems, but it still depends on conventional velocity estimation methods to deliver starting models sufficiently accurate to satisfy the convergence criterion for the FWI tomographic component. It also undermines the goal of simultaneous estimation because the high frequencies in the data contain the high-resolution tomographic information that facilitates the estimation of the velocity components represented by the blue line in Figure 1.

To perform true simultaneous and synergistic inversion for all the model scales we must address the convergence problems of the tomographic term in FWI. These convergence problems are related to the non-linearity of the wave-equation with respect to perturbations in the long wavelengths of the velocity model. Long-wavelengths perturbations cause substantial time shifts of the propagating wavefields that are poorly approximated by the linearization of the wave equation based on first order Born approximation. In this paper we

introduce a linearization of the wave equation based on an extension of the velocity model along the time lag axis (τ). This extension enables the linear modeling of large time shift in the propagating wavefields, and consequently in the data. Based on this extension we define an objective function that has a model-focusing term in addition to the conventional FWI data-fitting term. Numerical examples with realistically complex velocity models demonstrate that this objective function has excellent convergence behavior, although currently convergence is unsatisfactorily slow.

The usefulness of extending the reflectivity model (as prestack images in the angle or offset domain) to manage the non-linearities in wave-equation velocity analysis was demonstrated in the context of differential semblance optimization (DSO) (Symes and Carazzone, 1991; Shen and Symes, 2008) and wave-equation migration velocity analysis (WEMVA) (Biondi and Sava, 1999; Sava and Biondi, 2004a,b). The generalization of extending the model to the propagation component of the model (long wavelength) was first introduced by Symes (2008) and more recently successfully applied by Sun and Symes (2012); Almomin and Biondi (2012); Biondi and Almomin (2012). These methods are based on a velocity extension along the subsurface-offset or plane-wave axes. We propose an extension along the time lag axis (τ) because it is better suited to describe large time shifts in wave propagation, that are at the root of FWI convergence problems. Furthermore, extending the velocity along the time lag axis can easily handle forward-scattered events recorded at long offsets as well as the reflections recorded at near and intermediate offsets. We thus believe that the time-lag formulation has the potential to deliver high-quality results also for modern long-offset data sets. Furthermore, a one-dimensional extension along time is computationally more efficient than a two-dimensional extension along subsurface offsets. Yang and Sava (2009, 2010) have discussed the use and the computational advantages of time-lag gathers to perform WEMVA for reflected events.

TOMOGRAPHIC FULL WAVEFORM INVERSION (TFWI)

Conventional full waveform inversion in the acoustic constant-density approximation is performed by solving the following optimization problem

$$\min_{\mathbf{v}^2} J_{\mathrm{FWI}}\left(\mathbf{v}^2\right) \tag{1}$$

where:

$$J_{\mathrm{FWI}}\left(\mathbf{v}^2\right) = \frac{1}{2}\left\|\mathcal{L}\left(\mathbf{v}^2\right) - \mathbf{d}\right\|_2^2, \tag{2}$$

$\mathbf{v} = v\left(\vec{x}\right)$ is the velocity vector, $\mathcal{L}$ is a wave-equation operator non linear with respect to velocity perturbations and the data vector $\mathbf{d}$ is the pressure field $\mathbf{P} = P\left(t, \vec{x}\right)$ measured at the surface.

The wave-equation operator is evaluated by recursively solving the following finite difference equation

$$\left[\mathbf{D_2} - \mathbf{v}^2\nabla^2\right]\mathbf{P} = \mathbf{f}, \tag{3}$$

where $\mathbf{D_2}$ is a finite-difference representation of the second derivative in time, ∇^2 is a finite-difference representation of the Laplacian, and $\mathbf{f}$ is the source function.

Efficient solution of the optimization problem expressed in equation 1 is performed by gradient-based methods, and thus requires the evaluation of the linear operator $\mathbf{L}$, which is

the linearization of $\mathcal{L}$ with respect to velocity perturbations $\delta\mathbf{v}^2$. This linear operator can be derived by perturbing equation 3 as follows

$$\left[\mathbf{D_2} - \left(\mathbf{v_o}^2 + \delta\mathbf{v}^2\right)\nabla^2\right](\mathbf{P_o} + \delta\mathbf{P}) = \mathbf{f}, \tag{4}$$

where $\mathbf{P_o}$ and $\mathbf{v_o}$ are the background wavefield and velocity, respectively, and $\delta\mathbf{P}$ is the scattered wavefield.

Equation 4 can be rewritten as the following two equations:

$$\begin{aligned}\left[\mathbf{D_2} - \mathbf{v_o}^2\nabla^2\right]\mathbf{P_o} &= \mathbf{f}, &(5)\\ \left[\mathbf{D_2} - \mathbf{v_o}^2\nabla^2\right]\delta\mathbf{P} &= \delta\mathbf{v}^2\nabla^2(\mathbf{P_o} + \delta\mathbf{P}), &(6)\end{aligned}$$

which represent a nonlinear relationship between velocity perturbations and scattered wavefield. In conventional FWI, to linearize this relationship we drop the term multiplying the perturbations with each other; that is, we drop the scattered wavefield from the right hand side of equation 6 and obtain the following coupled equations:

$$\begin{aligned}\left[\mathbf{D_2} - \mathbf{v_o}^2\nabla^2\right]\mathbf{P_o} &= \mathbf{f}, &(7)\\ \left[\mathbf{D_2} - \mathbf{v_o}^2\nabla^2\right]\delta\mathbf{P} &= \delta\mathbf{v}^2\nabla^2\mathbf{P_o}. &(8)\end{aligned}$$

The linear operator $\mathbf{L}$ used to compute the gradient of the FWI objective function 2 is evaluated by recursively propagating the background wavefield $\mathbf{P_o}$ and the scattered wavefield $\delta\mathbf{P}$ by solving equations 7–8. The scattered wavefield $\delta\mathbf{P}$ is now a linear function of the velocity perturbations $\delta\mathbf{v}^2$, equation 8 has the limitation that it takes into account only first-order scattering, and thus it is unsuited to model large time shifts between the background wavefield and the scattered wavefield.

To improve the capability of the linearization to model large time shifts, we extend the velocity model along the time lag axis τ and convolve its perturbations $\delta\tilde{\mathbf{v}}^2(\tau)$ with the Laplacian of the background wavefield; we rewrite equation 6 as:

$$\left[\mathbf{D_2} - \tilde{\mathbf{v}}^2(\tau = 0)\nabla^2\right]\delta\mathbf{P} = \delta\tilde{\mathbf{v}}^2(\tau) \overset{\tau}{*} \nabla^2\mathbf{P_o}, \tag{9}$$

which defines the linear operator $\widehat{\mathbf{L}}$ and where $\overset{\tau}{*}$ denotes convolution in τ. The combination of the wave equation non-linear operator $\mathcal{L}(\mathbf{v})$ and of the linear operator $\widehat{\mathbf{L}}$ defines the extended non-linear operator

$$\tilde{\mathcal{L}}(\tilde{\mathbf{v}}) = \mathcal{L}(\tilde{\mathbf{v}}(\tau = 0)) + \widehat{\mathbf{L}}(\tilde{\mathbf{v}}(\tau = 0))\,\delta\tilde{\mathbf{v}}^2, \tag{10}$$

that can be used to define the TFWI objective function as

$$J_{\mathrm{TFWI}}(\tilde{\mathbf{v}}) = \frac{1}{2}\left\|\tilde{\mathcal{L}}(\tilde{\mathbf{v}}) - \mathbf{d}\right\|_2^2 + \epsilon\,\||\tau|\,\tilde{\mathbf{v}}\|_2^2. \tag{11}$$

The second term in the equation 11 rewards focusing of the data around zero time lag. It introduces a strong tomographic component, which is necessary to constrain the optimization problem because the velocity extension relaxes the constraints on the modeled data kinematics imposed by the data fitting term in equation 11. This objective function can be minimized using the nested optimization algorithm with scale mixing that we discussed in Almomin and Biondi (2013).

We will use a simple 1D numerical example to analyze some of the characteristics of the TFWI method we introduced above. Figure 2 shows the difference between the background wavefield propagated with $\mathbf{v}$=1.2 km/s and the wavefield propagated with the true velocity of $\mathbf{v}$=1.13 km/s. The source function is a zero-phase wavelet bandlimited between 5 and 20 Hz. The difference wavefield is displayed as a function of propagation distance and traveltime. The velocity error is sufficiently high that the wavefields are completely out of phase after propagating for a couple of kilometers. This is therefore a situation like the ones described where the first order Born linearization (equation 8) would fail to to model the data residuals and conventional FWI would have troubles to converge, even if the problem is extremely simple.

Figure 3 shows the conventional FWI objective functions when the data are recorded with a single receiver located at 7 km for a total of 4 km offset from the source. The plot shows the initial value of the objective function for several 1D transmission problems sharing the same starting velocity (1.2 km/s) and with different true velocities. If the true velocity is lower than $\approx$ 1.18 km/s or larger than $\approx$ 1.22 km/s a gradient-based method starting from a velocity of 1.2 km/s will not converge to the right solution. An explanation for the lack of convergence of a gradient-based method is that conventional first-order Born approximation cannot model the data residuals when the true velocity is outside this range. On the contrary, the linearized modeling equation defined in equation 9 would have no troubles to model the data residual.

For example, we can easily reproduce the wavefield difference shown in Figure 2 by setting the extended-velocity perturbation to be a delta function along the τ axis, where the shift of the delta function linearly increases with the distance from the origin. This linear shift is computed by integrating the difference in slowness between the background model and the true model. The extended-velocity perturbation is shown in Figure 4. Figure 5 shows the result of solving equation 9 with the model shown in Figure 4. The approximation of the scattered wavefield $\Delta\mathbf{P}$ is almost identical to the wavefield difference shown in Figure 2.

DIVING-WAVE MODELING EXAMPLE

One of the advantages of extending the velocity model along the time-lag axis τ is the capability to model with a linear operator large time shifts in the diving waves recorded by modern long-offset acquisition geometries. The capability of modeling time shifts in these events enables robust convergence of the inversion even when the starting velocity model is far from the correct one. To show these modeling capabilities we use one long-offset shot profile recorded over a half space with a vertical velocity gradient. The starting velocity model is assumed to be uniform and equal to the velocity at the surface. Figure 6a shows the data residual; both the recorded diving wave as well as the data modeled with the starting velocity are clearly visible.

The backprojection of data residuals shown Figure 6a by the application of $\widehat{\mathbf{L}}'$ generates the velocity perturbation cube shown in Figure 7. The front panel of the cube shown in Figure 7 displays the zero time lag of the velocity perturbations. A substantial amount of the energy in the residual has been backprojected away from the zero time-lag panel.

Figure 6c displays the result of forward modeling the data residuals by the application

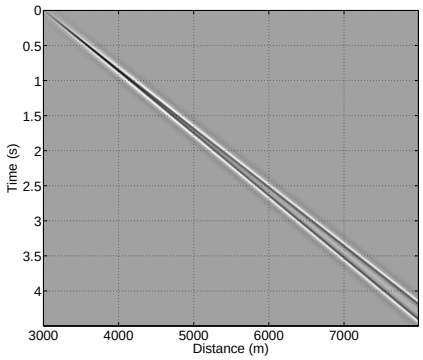


Figure 2: Difference between background wavefield computed with the starting velocity (1.2 km/s) and the wavefield propagated with the true velocity (1.13 km/s). [**CR**] biondo1/. Init-Residual

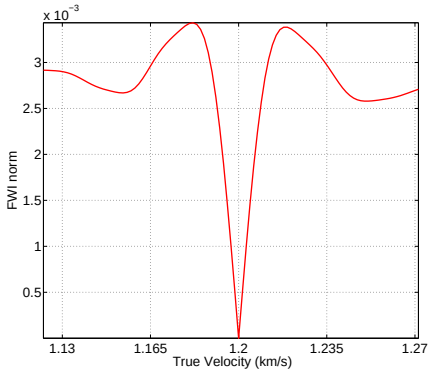


Figure 3: FWI norm as a function of the true velocity, when the starting velocity is equal to 1.2 km/s. [**CR**] biondo1/. FWI-Norm-new

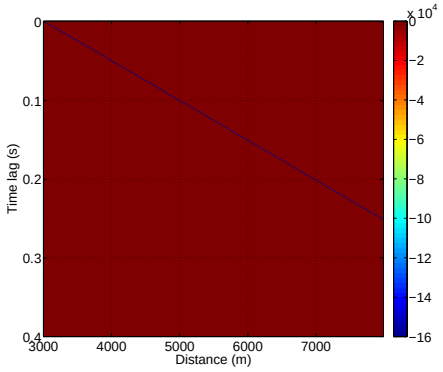


Figure 4: Extended velocity perturbation chosen to approximately model the wavefield difference shown in Figure 2. [**CR**] biondo1/. Delta-Vel-Extend

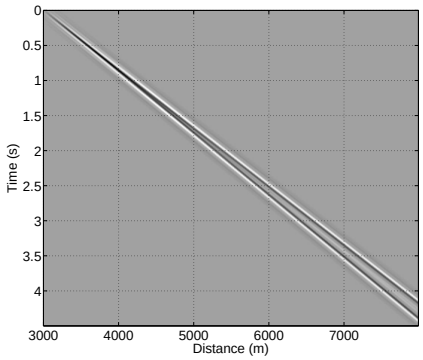


Figure 5: Perturbed wavefield computed by solving equation 9 with the model shown in Figure 4. [**CR**] biondo1/. Lin-Forw-Ext-new

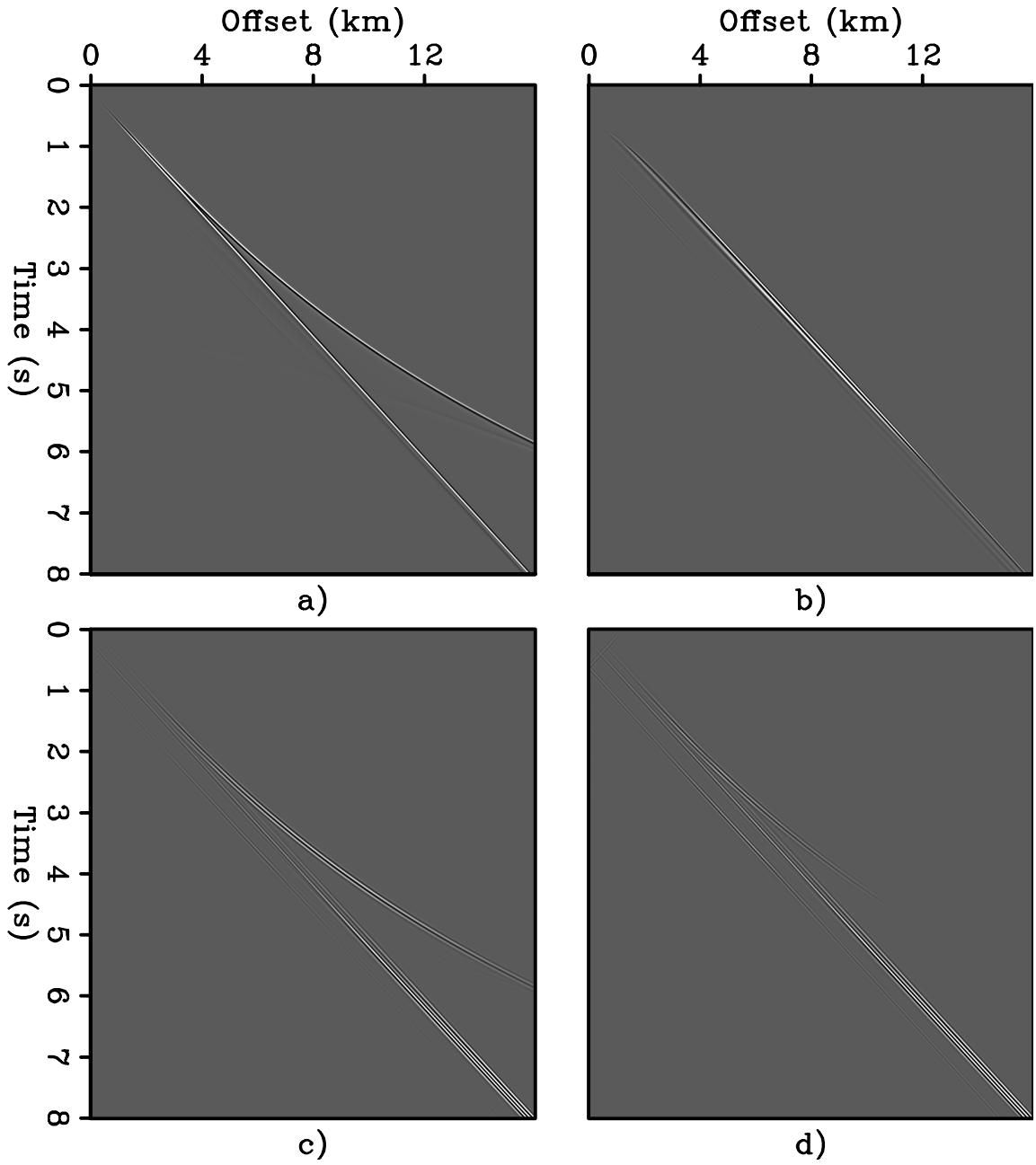


Figure 6: a) Difference between background wavefield computed with the starting velocity and the wavefield propagated with the true velocity, b) data residual modeled from zero lag of the velocity perturbation cubes (front panel in cubes shown in both Figure 7 and Figure 8), c) data residual modeled from the velocity perturbations extended along the time-lag axis, (Figure 7), and d) data residual modeled from the velocity perturbations extended along the horizontal subsurface offset axis, (Figure 8). [**CR**] biondo1/. diving-data-all

of $\widehat{\mathbf{L}}$ to the extended velocity perturbation shown in Figure 7. The kinematics of these modeled residuals are very close to the kinematics of the true residuals shown Figure 6a. In contrast, when only the zero time lag of the velocity perturbation (front panel of the cube shown in Figure 7) is used to model the data residuals, we obtain the seismograms displayed in Figure 6b. The diving wave is totally missing from these modeled residuals because the background wavefield propagates with constant velocity along the horizontal direction.

Figure 8 shows the velocity perturbation cube when the velocity is extended along the horizontal subsurface offset axis. The front panel of the cube displays the zero subsurface offset, and thus it is identical to the front panel of the cube shown in Figure 7. Figure 6d displays the result of forward modeling the data residuals starting from the extended velocity perturbation shown in Figure 8. The diving wave event is present in these modeled residuals. However, it dies out at larger offsets, starting at about 8 kilometers offset. A subsurface offset extension of the velocity has difficulties in modeling large time shifts in transmitted events propagating in directions orthogonal to the subsurface offset axes. Consequently, as the propagation paths of the diving waves have longer vertical components, the less accurate the modeled residuals are. To address this limitation we could use vertical subsurface offset in addition to the horizontal ones (Biondi and Symes, 2004), but the computational cost would increase accordingly.

TFWI OF LONG-OFFSET DATA

To verify the capabilities of the TFWI method based on time-lag extension of the velocity model we tested the method on a synthetic data set recorded with long offsets. The data were generated over the "Caspian Sea" portion of the well-known BP velocity model, as shown in Figure 9. The receiver array was assumed to be fixed at the surface, and thus data with more than 20 kilometers long offsets were recorded. The source was bandpassed wavelet between 5 and 10 Hz. No energy was present below 3 Hz.

Figure 10 shows the data recorded for the leftmost shot location. Strong and complex diving waves and refracted arrivals are visible in the data starting from approximately 8 kilometers offsets. These events carry useful information on the velocity, in particular in the shallow part of the section. In this data set, they are extremely useful to define the low-velocity anomalies present around the depth of two kilometers.

The starting model for the TFWI inversion was obtained by a strong horizontal smoothing of the true model, after the low and high velocity anomalies were removed. Figure 11 shows the shot gather located at the same location as the one shown in Figure 10, but modeled with the starting model, which is shown in Figure 12. As a direct comparison of Figure 10 with Figure 11 demonstrates, the differences between the true and starting models cause large time shifts in the diving-waves arrivals. The inaccuracies of the starting model, together with the lack of low frequencies in the data, prevent conventional FWI from converging to any useful model.

The straightforward application of an optimization algorithm designed for non-quadratic problems to the TFWI minimization problem defined by equation 11 would be very expensive. We therefore solved the problem by a nested optimization algorithm as described in Almomin and Biondi (2013). The nested inversion converged towards the accurate model shown in Figure 13. The main features of the true model are accurately reconstructed.

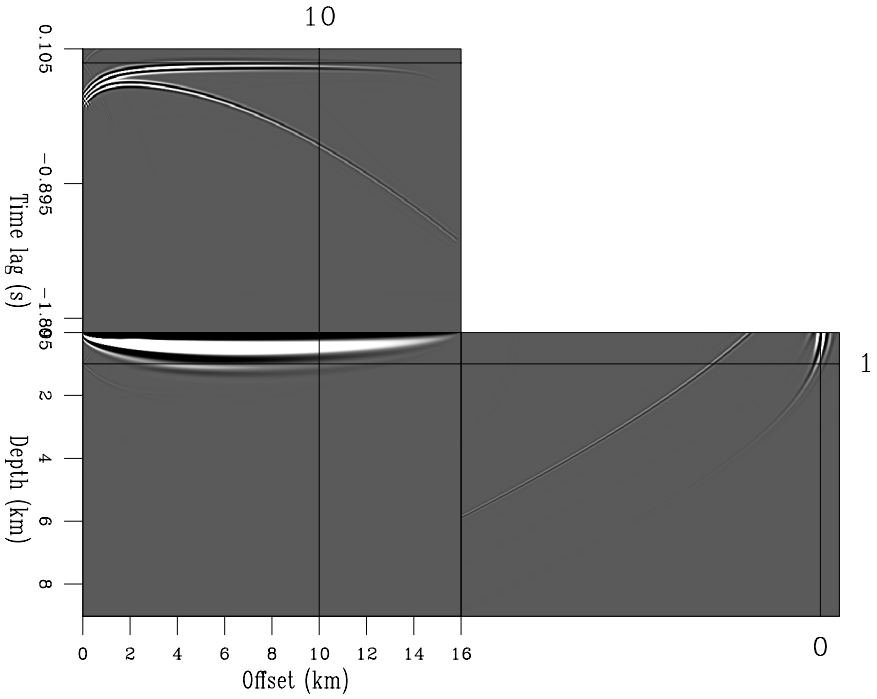


Figure 7: Velocity-perturbation cube extended along the time-lag axis and computed by backprojecting the data residuals shown Figure 6a. [**CR**] biondo1/. diving-image-t-cube

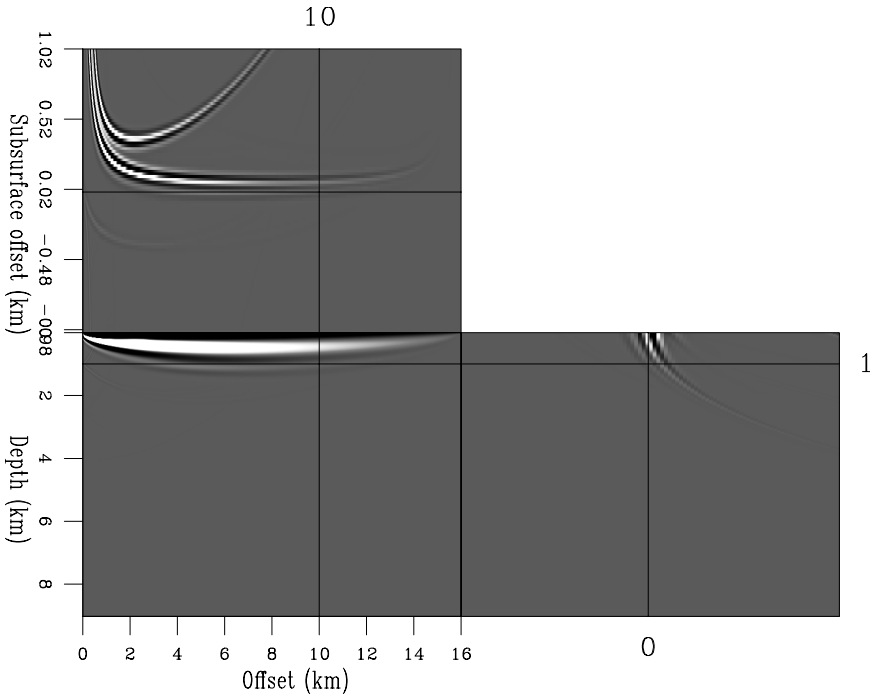


Figure 8: Velocity-perturbation cube extended along the horizontal subsurface-offset axis and computed by backprojecting the data residuals shown Figure 6a. [**CR**] biondo1/. diving-image-x-cube

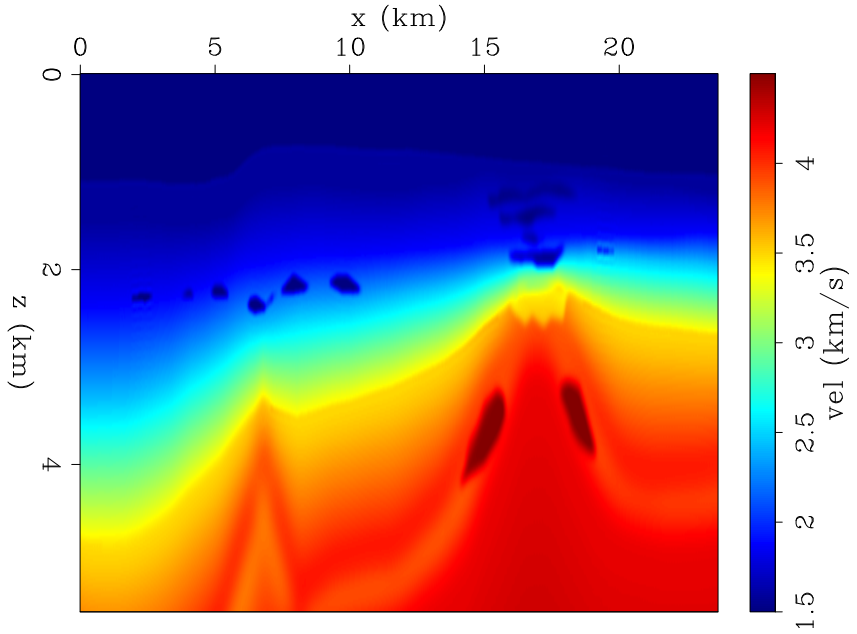


Figure 9: Portion of the BP velocity model used for the numerical test of the proposed TFWI method. The model contains both low-velocity anomalies (shallow gas) as well as high-velocity anomaly on the flanks of the mud volcano. [**ER**] biondo1/. bp-true

Some edge artifacts are present; they are caused by the finite span of the receiver array and the finite range of source locations. No sources were activated outside the displayed model, and the receiver array was fixed and covering the whole model.

Although the resolution of the model and the robustness of convergence are extremely attractive, the rate of convergence is slower than ideal; 500 iterations of the outer loop of the nested TFWI inversion were required to estimate the model shown in Figure 13. We believe that several numerical techniques could be applied to speed up the convergence, but further experiments are needed.

CONCLUSIONS

The integration of FWI and WEMVA into TFWI promises to enable robust convergence to high-resolution models. We introduced a TFWI method based on the extension of the velocity along the τ axis. This extension is based on a linear operator capable of correctly modeling transmitted events with large time shifts, as we demonstrated by a simple numerical example. The results of the inversion of a long-offset data set recorded over the BP "Caspian Sea" demonstrate the strong convergence properties of new method when both reflections and diving waves are present.

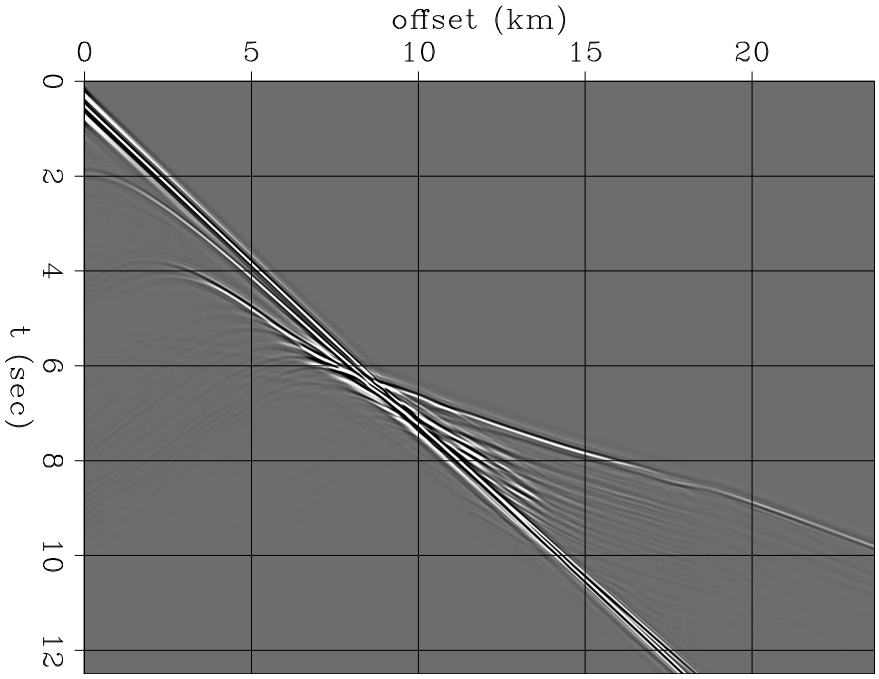


Figure 10: Leftmost shot profile recorded on the model shown in Figure 9. Notice several diving waves and refractions present in the data at offset larger than 8 kilometers. These events carry useful information for the estimation of the velocity anomalies present in the model. [**CR**] biondo1/. bp-data-obs

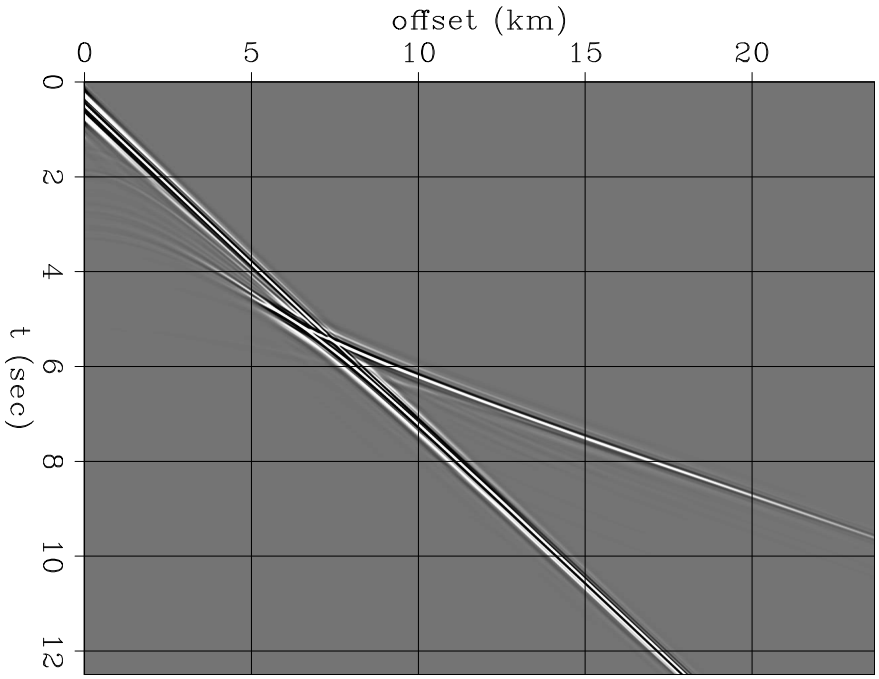


Figure 11: Shot gather modeled assuming the starting model shown in Figure 12 at the same shot location as the data shown in Figure 10. Notice the large time shifts between the diving-wave arrivals in this gather with the one shown in Figure 10. [**CR**] biondo1/. bp-data-init

ACKNOWLEDGMENTS

We would like to thank BP for making publicly available the velocity model of which we used a portion. Almomin would like the thank Saudi Aramco for supporting his Ph.D. studies at Stanford.

REFERENCES

Almomin, A. and B. Biondi, 2012, Tomographic full waveform inversion: Practical and computationally feasible approach: SEG Technical Program Expanded Abstracts, **31**, 500–505.

———, 2013, Tomographic full waveform inversion (TFWI) by successive linearizations and scale separations: SEP-Report, **149**, 51–58.

Biondi, B. and A. Almomin, 2012, Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **31**, 547–552.

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

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

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

Sava, P. and B. Biondi, 2004a, Wave-equation migration velocity analysis—I: Theory: Geophysical Prospecting, **52**, 593–623.

———, 2004b, Wave-equation migration velocity analysis—II: Examples: Geophysical Prospecting, **52**, 607–623.

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

Sun, D. and W. Symes, 2012, Waveform inversion via nonlinear differential semblance optimization: SEG Technical Program Expanded Abstracts, **31**, 497–502.

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.

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

———, 2010, Moveout analysis of wave-equation extended images: Geophysics, **75**, S151–S161.

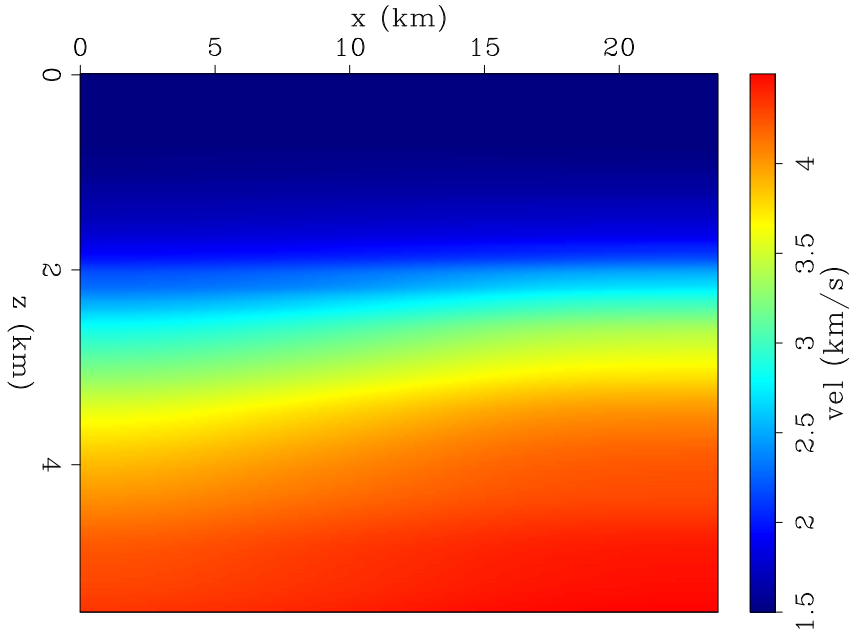


Figure 12: Starting model for the TFWI inversion. This model was obtained by strong horizontal smoothing of the model shown in Figure 9, after the low and high velocity anomalies were removed. The lack of low frequencies in the data makes this model inappropriate for starting a conventional FWI inversion. [**ER**] biondo1/. bp-init

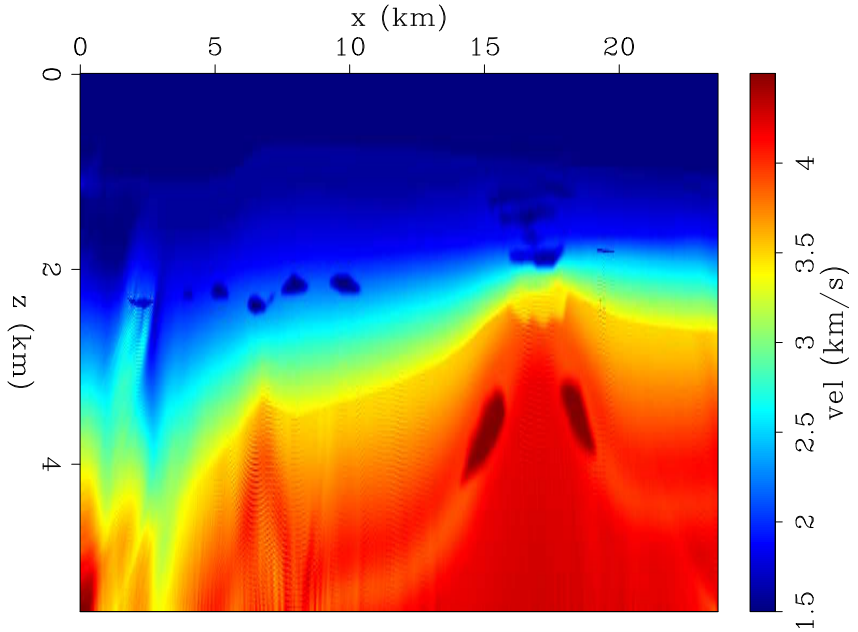


Figure 13: Estimated model after 500 iterations of the outer loop of the TFWI inversion based on time-lag extension of the velocity model. The main features of the true model are accurately reconstructed. Some edge artifacts are present; they are caused by the finite span of the receiver array and the finite range of source locations. No sources were activated outside the displayed model, and the receiver array was fixed and covering the whole model. [**CR**] biondo1/. bp-inv

Tomographic full waveform inversion (TFWI) by successive linearizations and scale separations

Ali Almomin and Biondo Biondi

ABSTRACT

Tomographic full waveform inversion (TFWI) provides a framework to invert the seismic data that is immune to cycle-skipping problems. This is achieved by extending the wave equation and adding an offset axis to the velocity model. However, this extension makes the propagation considerably more expensive because each multiplication by velocity becomes a convolution. We provide an alternative formulation which computes the backscattering and the forward scattering components of the gradient separately. To maintain high resolution results from TFWI, the two components of the gradient are first mixed and then separated based on a Fourier domain scale separation. This formulation is based on the Born approximation where the physical medium parameters are broken into a long-wavelength and short-wavelength components. The inversion setup includes two steps that maintain the high resolution results of TFWI. First, the linearized residual are updated in a nested inversion scheme. This step corrects for the underlying assumption that the data contain only primaries and no multiples. Second, the two components of the gradient are first mixed and then separated based on a Fourier domain scale separation to allow for a fully simultaneous inversion of model scales. After deriving the equations, we test the theory with two synthetic examples. The results of both the Marmousi and BP models show that convergence is possible even with large errors in the initial model that would have prevented convergence in conventional FWI.

INTRODUCTION

Seismic velocity-analysis methods can be divided into two major groups. First, there are techniques that aim to minimize misfit in the data domain, such as full-waveform inversion (Tarantola, 1984; Pratt, 1999; Luo and Schuster, 1990). Second, there are other techniques that aim at improving the quality in the image domain, such as migration velocity analysis (MVA) (Symes and Carazzone, 1991; Biondi and Sava, 1999; Shen, 2004; Zhang et al., 2012). All these techniques try to measure the quality of the image in several ways and then invert the estimated image perturbation using a linearized wave-equation operator.

There are several advantages to minimizing the residual in the image-space, such as global convergence, increasing the signal-to-noise ratio, and decreasing the complexity of the data (Tang et al., 2008). However, a common drawback in doing velocity analysis in the image domain is that only the transmission effects of the velocity are used. This results in a loss of vertical resolution in the estimated model updates. On the other hand, full-waveform inversion (FWI) does not have that problem, since it utilizes the information from both the forward-scattered and back-scattered wavefields, i.e. both their kinematics

and their dynamics. This results in higher resolution in the model estimates. Moreover, the data misfit is computed in the data spaces directly without the need to go to another domain or to separate the data into several components. This direct computation of the errors results in a relatively simple relationship between the data residuals and the model updates. However, FWI has the disadvantage that its objective function is far from being smooth and convex which, in the absence of very low frequency, requires the starting model to be very close to the true model to avoid converging to a local minimum.

One conventional solution is to first invert for the velocity model using MVA techniques and then use the output as the initial model for FWI. However, this practice might not work if the results of MVA are not an accurate enough initial model for FWI. This could be a result of the larger null space that forward-scattered wavefields do not constrain. Moreover, the convergence rate of the MVA techniques is very likely sub-optimal, since they do not use all of the information contained in the data.

Tomographic Full Waveform Inversion (TFWI) (Symes, 2008; Biondi and Almomin, 2012) provides a way to overcome this problem by combining both FWI and wave-equation migration velocity analysis (WEMVA) techniques in a generalized framework. This generalized approach utilizes the components of all seismic data to invert for the medium parameters without cycle-skipping. This is achieved in two steps: first, extending the wave equation and adding a subsurface offset axis to the velocity model, and second, adding a regularization term that drives the solution towards the zero subsurface offset. Biondi and Almomin (2013) presented an alternative extension using time shift instead of subsurface offset. In either setting, this velocity extension makes the propagation considerably more expensive because each multiplication by velocity becomes a convolution over the extended axis.

In a previous abstract (Almomin and Biondi, 2012), we presented an approximation that significantly reduces the computational cost of TFWI by breaking the velocity model into a background component and a perturbation component. Unfortunately, there were several issues with that approach. First, it was based on the Born approximation where the data are assumed to contain only primaries and only utilizes the linearized operator. Second, the inversion output a background model and a perturbation model but not the velocity model. Finally, dividing the two models depended on a heuristic scale separation that might not be accurate across different wavelengths.

In this report, we present a modified approach that mitigates the previously mentioned limitations while maintaining the advantages and high resolution results of TFWI. We achieve this in two steps. First, we set up a nested inversion scheme that utilizes the nonlinear modeling operator to update the residuals. Second, the two components of the gradient are first mixed and then separated based on a Fourier domain scale separation.

NESTED INVERSION

Although the derivation that follows is carried out in frequency domain, our actual implementation is in the time domain. We start with the full waveform objective function, J_{FWI}, which we write as

$$J_{\text{FWI}}(\mathbf{v}) = \|\mathbf{d}(\mathbf{v}) - \mathbf{d}_{\text{obs}}\|_2^2, \tag{1}$$

where $\mathbf{v}$ is the velocity model, $\mathbf{d}(\mathbf{v})$ is the modeled data, and $\mathbf{d}_{\text{obs}}$ is the observed surface data. The modeled data $\mathbf{d}(\mathbf{v})$ is computed as

$$d(\mathbf{x}_s, \mathbf{x}_r, \omega; \mathbf{v}) = f(\mathbf{x}_s, \omega) G(\mathbf{x}_s, \mathbf{x}, \omega; \mathbf{v}) \delta(\mathbf{x}_r - \mathbf{x}), \tag{2}$$

where $f(\omega)$ is the source function, ω is frequency, $\mathbf{x}_s$ and $\mathbf{x}_r$ are the source and receiver coordinates, and $\mathbf{x}$ is the model coordinate. In the acoustic, constant-density case the Green's function $G(\mathbf{x}_s, \mathbf{x}, \omega; \mathbf{v})$ satisfies:

$$\left(v^{-2}(\mathbf{x})\omega^2 + \nabla^2\right) G(\mathbf{x}_s, \mathbf{x}, \omega) = \delta(\mathbf{x}_s - \mathbf{x}). \tag{3}$$

This first objective function represents the outer loop of the inversion. We now compute the nonlinear residuals $\mathbf{d}'$ as:

$$\mathbf{d}'(\mathbf{v}) = \mathbf{d}(\mathbf{v}) - \mathbf{d}_{\text{obs}}. \tag{4}$$

The nonlinear residual will be used as the "observed" data for the inner loop. In the inner loop of the inversion, we need to separate the velocity model into a background and a perturbation as follows:

$$v^{-2}(\mathbf{x}) = b(\mathbf{x}) + p(\mathbf{x}, \mathbf{q}), \tag{5}$$

where $\mathbf{q}$ is the extension axis with lags in either space or time, $b(\mathbf{x})$ is the background component, which is a smooth version of the slowness squared and $p(\mathbf{x}, \mathbf{q})$ is the perturbation component. The perturbation component can extend across several subsurface offsets or time shifts so it is important to keep its extended axis. On the other hand, the background component is not expected to generate reflections that would be grossly time shifted with respect to the recorded data, and it thus safe to reduce its extension. This greatly reduces our cost since the convolution with velocity in propagation becomes a multiplication. The model separation allows us to use a linearized (Born) operator $\mathbf{L}$ to model "linearized" data. The linearized objective function J_{Born} can be written as:

$$J_{\text{Born}}(\mathbf{b}, \mathbf{p}(\mathbf{q})) = \|\mathbf{L}(\mathbf{b})\mathbf{p}(\mathbf{q}) - \mathbf{d}'\|_2^2 + \epsilon\|\mathbf{A}\mathbf{p}(\mathbf{q})\|_2^2, \tag{6}$$

where $\mathbf{L}$ is the Born modeling operator and $\mathbf{A}$ is the regularization operator. The linearized wave equation defines the data as follows:

$$d'(\mathbf{x}_s, \mathbf{x}_r, \omega; \mathbf{b}, \mathbf{p}(\mathbf{q})) = -\omega^2 f(\omega) \sum_{\mathbf{x},\mathbf{q}} G(\mathbf{x}_s, \mathbf{x}, \omega, -\mathbf{q}; \mathbf{b}) p(\mathbf{x}, \mathbf{q}) G(\mathbf{x}, \mathbf{x}_r, \omega, +\mathbf{q}; \mathbf{b}), \tag{7}$$

where the Green's functions are lagged by $\mathbf{q}$ and satisfy the conventional acoustic wave equation as follows:

$$\left(\omega^2 b(\mathbf{x}) + \nabla^2\right) G(\mathbf{x}_s, \mathbf{x}, \omega) = \delta(\mathbf{x}_s - \mathbf{x}), \tag{8}$$

$$\left(\omega^2 b(\mathbf{x}) + \nabla^2\right) G(\mathbf{x}, \mathbf{x}_r, \omega) = \delta(\mathbf{x} - \mathbf{x}_r). \tag{9}$$

The forward modeling can be written in a compact notation as follows:

$$\mathbf{d}' = \mathbf{L}(\mathbf{b})\mathbf{p}(\mathbf{q}). \tag{10}$$

The Born modeling operator is linear with respect to perturbation but nonlinear with respect to the background component. Therefore, another linearization around the "background" background is required to compute the gradient. First, we rewrite the background as the sum of two components as follows:

$$b(\mathbf{x}) = b_0(\mathbf{x}) + \Delta b(\mathbf{x}), \tag{11}$$

where $b_0(\mathbf{x})$ is the current background model and $\Delta b(\mathbf{x})$ is the perturbation of the background. The Born approximation is used again to linearize the $\mathbf{L}$ operator with respect to the background resulting in a data-space tomographic operator. The data perturbation with respect to the background perturbation is now defined as:

$$\begin{aligned}
&\Delta d'(\mathbf{x}_s, \mathbf{x}_r, \omega; \mathbf{b}_0, \mathbf{p}(\mathbf{q})) = \\
&\sum_{\mathbf{x},\mathbf{y},\mathbf{q}} \omega^4 f(\omega) G(\mathbf{x}_s, \mathbf{y}, \omega, -\mathbf{q}; \mathbf{b}_0) p(\mathbf{y}, \mathbf{q}) G(\mathbf{y}, \mathbf{x}, \omega, +\mathbf{q}; \mathbf{b}_0) \Delta b(\mathbf{x}) G(\mathbf{x}, \mathbf{x}_r, \omega, 0; \mathbf{b}_0) \\
&+ \omega^4 f(\omega) G(\mathbf{x}_s, \mathbf{x}, \omega, 0; \mathbf{b}_0) \Delta b(\mathbf{x}) G(\mathbf{x}, \mathbf{y}, \omega, -\mathbf{q}; \mathbf{b}_0) p(\mathbf{y}, \mathbf{q}) G(\mathbf{y}, \mathbf{x}_r, \omega, +\mathbf{q}; \mathbf{b}_0), \qquad (12)
\end{aligned}$$

where $\mathbf{y}$ is the perturbation coordinate. Observe that in the previous equation, the tomographic operator correlates a background and a scattered wavefield from both the source and receiver sides. The scattered wavefields are computed by correlating a background wavefield with the perturbation model $\mathbf{p}$ and then propagating again to all model locations. This operator is similar to the WEMVA operator except that in the WEMVA operator, we keep the data fixed and vary the image while here it is the other around for this tomographic operator. The forward tomographic operator can be written in a compact notation as follows:

$$\Delta \mathbf{d}' = \frac{\partial \mathbf{L}}{\partial \mathbf{b}} \mathbf{p}(\mathbf{q}) \Delta \mathbf{b} = \mathbf{T} \Delta \mathbf{b}, \qquad (13)$$

Where $\mathbf{T}$ is the tomographic operator that relates changes in the background model to changes in the data. We can now compute the perturbation gradient as follows:

$$g_{\mathbf{p}}(\mathbf{x}, \mathbf{q}) = \frac{\partial J}{\partial \mathbf{p}(\mathbf{q})} = \mathbf{L}^* \Delta \mathbf{d}', \qquad (14)$$

which is simply migration of the residuals. Then, we compute the background gradient as follows:

$$g_{\mathbf{b}}(\mathbf{x}) = \frac{\partial J}{\partial \mathbf{b}} = \mathbf{T}^* \Delta \mathbf{d}'. \qquad (15)$$

When we start the inner loop, we set $\mathbf{b} = \mathbf{v}^{-2}$ and $\mathbf{p} = 0$. Once we iterate enough times within the inner loop, we update the velocity model with both the background and the perturbation.

As previously mentioned, what is considered observed data in the inner loop is simply the residual of the outer loop. This nested scheme has many benefits. The first benefit is that the limitation of the linearized, first-order Born operator will not prevent the optimization from inverting higher order scattering. The reason is that in this setup, the Born operator will attempt to match the first-order scattering from the background, which is not necessarily primary data since the background itself can generate many orders of scattering if it contains sharp boundaries (and it will in later iterations). In other words, the linearized operator will account for one order of scattering at a time, and keep going to higher and higher scattering orders as we continue iterating in the outer loop until we eventually invert all scattering orders that are present in the data. This also means that higher scattering orders will initially be wrongly positioned but later iterations will correct for that. The second benefit of this scheme is that we result in one model because we keep pushing both background and perturbation into the velocity model, in contrast to the method presented in Almomin and Biondi (2012). Therefore, our goal is to drive the perturbation to a minimum and have the background explain the data.

SCALE MIXING

In the inner loop, a straight forward use of the gradients is to update their corresponding models directly. However, this hinders the simultaneous inversion of different wavelengths of the model. This becomes apparent when we examine the result of the two operators in the inner loop. Just migrating the data could give a tomographic update which shows as a low wavenumber update. In a conventional migration, this low wavenumber part is considered noise and filtered out. However, it is actually a tomographic component that should feed into the background model. The opposite argument is also true for a tomographic operator containing a migration component. Therefore, to improve our inversion results, we first mix the two gradients and then separate them in Fourier domain to get the update of each model as follows:

$$s_{\mathbf{b}}(\mathbf{x}) = \mathbf{C_b}(g_{\mathbf{b}}(\mathbf{x}) + g_{\mathbf{p}}(\mathbf{x}, \mathbf{q} = 0)), \tag{16}$$

where $s_{\mathbf{b}}(\mathbf{x})$ is the search direction of the background model and $\mathbf{C_b}$ is a low-pass filter. Similarly, we compute the update of the perturbation model as

$$s_{\mathbf{p}}(\mathbf{x}, \mathbf{q}) = \mathbf{C_p}(g_{\mathbf{b}}(\mathbf{x}) + g_{\mathbf{p}}(\mathbf{x}, \mathbf{q})), \tag{17}$$

where $s_{\mathbf{p}}(\mathbf{x}, \mathbf{q})$ is the search direction of the perturbation model and $\mathbf{C_p}$ is a high-pass filter. In order to sum the two gradients properly, both of them need to have the same units as well as the same scale. This requires careful implementation of each operator at each linearization.

In the example in this paper, we used a radial cut-off in the Fourier domain with a cosine squared taper. The wavelength cut-off is based on the dominant frequency in the data as well as the average velocity of the initial model. The two filters were designed such that they always sum to unity at all wavelengths so as to maintain the energy of the gradients. It is possible to design a more accurate filter that varies spatially, but it is not necessary because both models will eventually be added to the velocity. This is another benefit of having this nested scheme as oppose to inverting two separate models where the scale separation become very important to the final results.

SYNTHETIC EXAMPLE

We modified the Marmousi model for our synthetic example by adding 500m of water layer to the top. Figure 1 shows the true velocity model. We used a bandpass wavelet with a frequency range between 5 Hz to 10 Hz and a small taper on both ends. The purpose of using this wavelet was to completely eliminate unrealistically low frequencies in the data. There were 461 fixed receivers with a spacing of 20 m and 93 sources with a spacing of 100 m. The initial model is shown in Figure 1, which we obtained by strongly smoothing the true model laterally after clipping out the salt anomalies. For this example, the perturbation was extended over a time lag axis. The inversion results after 900 outer loop iterations are shown in Figure 1. The inversion shows remarkable results in reconstructing most features in the velocity model. The results are most accurate in the top and middle of the bottom because the have the best illumination and coverage.

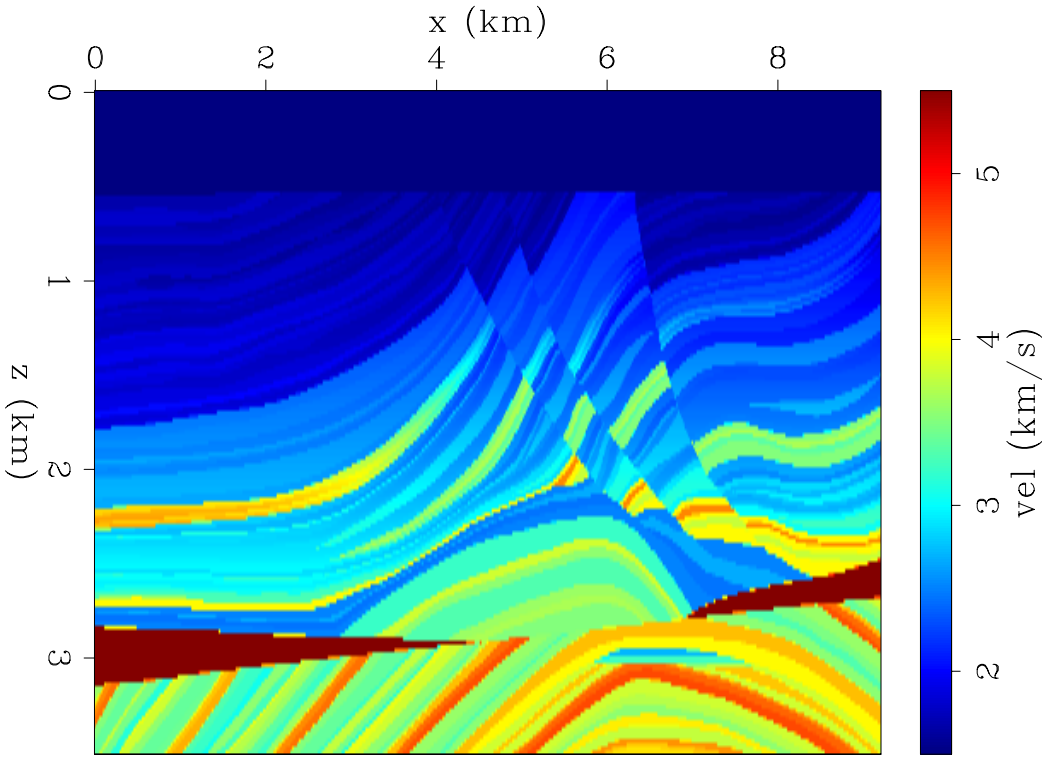


Figure 1: The true velocity of the Marmousi example. [**ER**] ali1/. marmousi-true

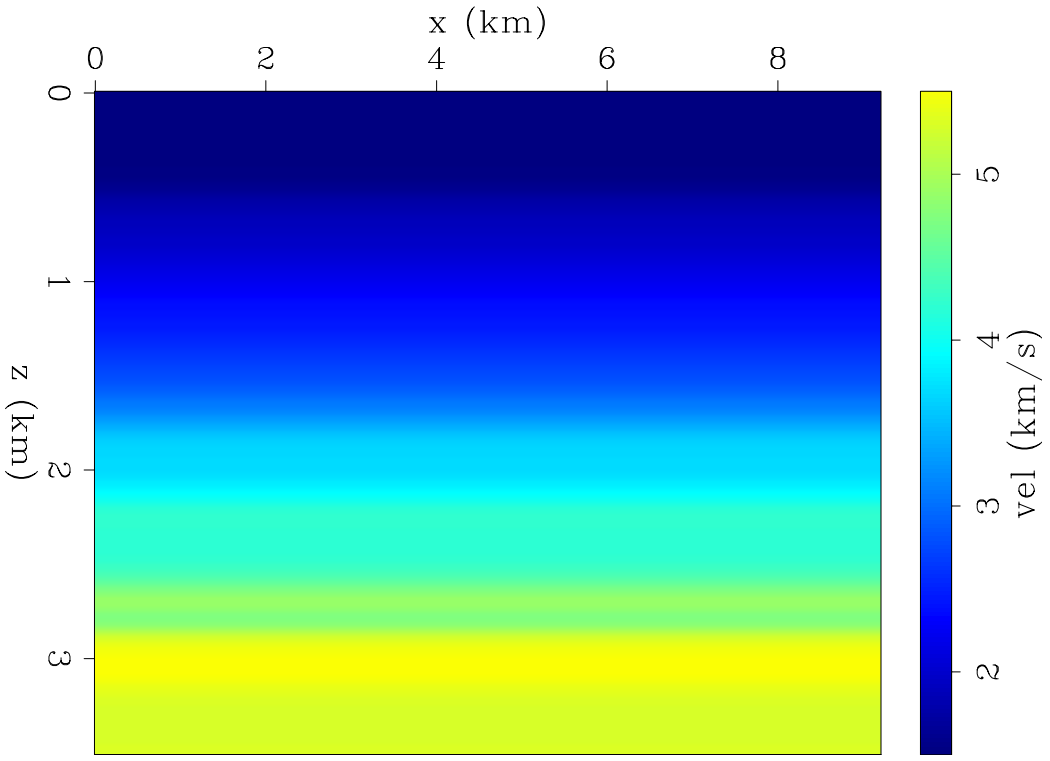


Figure 2: The initial velocity of the Marmousi example. [**ER**] ali1/. marmousi-init

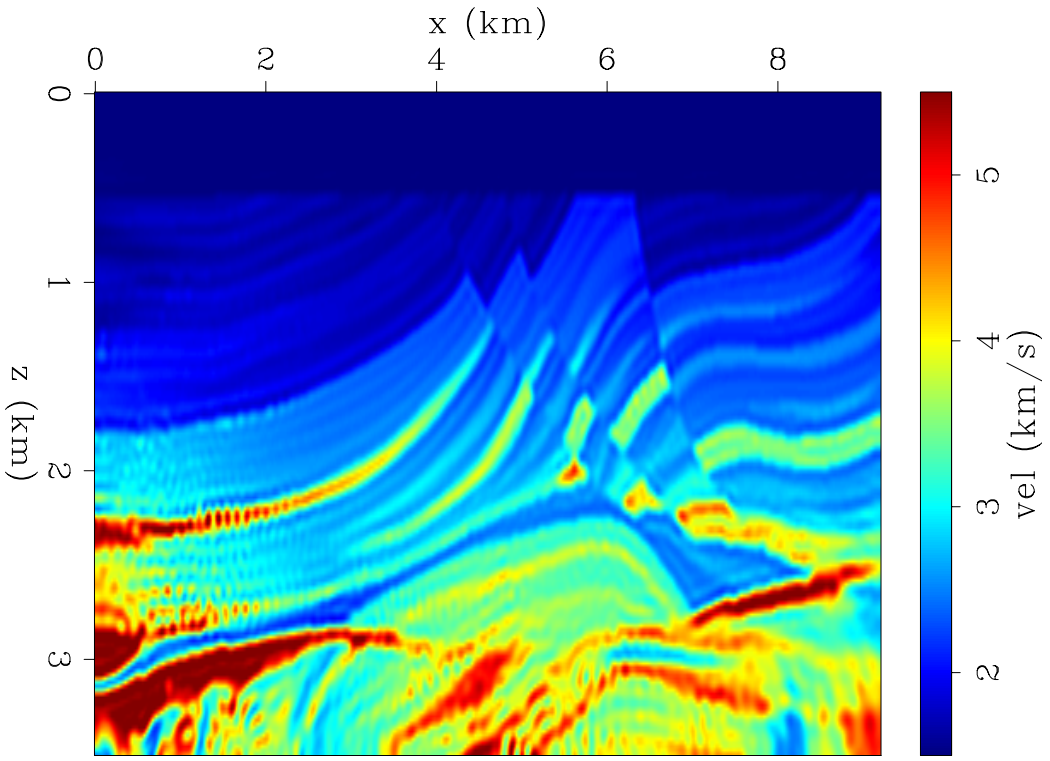


Figure 3: The inverted velocity of the Marmousi example. [**CR**] ali1/. marmousi-inv

CONCLUSIONS

We have presented a practical approach to tomographic full waveform inversion. This was achieved by first breaking the model into a background component and a perturbation component, and then by restricting the extended axis of the background component to significantly reduce computational cost. The separated model uses a linearized modeling operator which uses the residual of the nonlinear modeling operator as observed data in the inner step of the nested inversion scheme. We managed to maintain the simultaneous inversion of different wavelengths by mixing the gradients of the two components and then separating them in Fourier domain.

ACKNOWLEDGMENTS

We thank the Stanford Exploration Project affiliate companies for financial support and BP for the velocity model. Almomin thanks Saudi Aramco for supporting his Ph.D. studies at Stanford University.

REFERENCES

Almomin, A. and B. Biondi, 2012, Tomographic full waveform inversion: Practical and computationally feasible approach: SEG Expanded Abstracts, 1–5.

Biondi, B. and A. Almomin, 2012, Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis: SEG Expanded Abstracts, **31**, 275–279.

———, 2013, Tomographic full waveform inversion (TFWI) by extending the velocity model along the time-lag axis: SEP-Report, **149**, 37–50.

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

Luo, Y. and G. T. Schuster, 1990, Wave-equation traveltime inversion: SEG Expanded Abstracts.

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

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

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.

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

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

Zhang, Y., B. Biondi, and Y. Tang, 2012, Residual moveout-based wave-equation migration velocity analysis: SEG Expanded Abstracts, 1–5.

Wave-equation migration velocity analysis using partial-stack power maximization

Yang Zhang and Guojian Shan

ABSTRACT

We proposed to use a partial-stack power maximization objective function in wave-equation migration velocity analysis. Instead of stacking the angle-domain common-image gathers all at once, the partial-stack power maximization objective function stacks them in smaller groups. It improves the robustness against the cycle-skipping problem and can achieve better global convergence. We also added a normalization term to the partial-stack power maximization objective function to balance the different reflector amplitudes. We tested our objective function using the Marmousi model. The results demonstrate that using the partial-stack power maximization criterion can achieve better global convergence. We also observed that the normalization of reflector amplitudes is very important in order to better constrain the tomography problem.

INTRODUCTION

Introduced by Gardner (1974) and Sattlegger (1975), migration velocity analysis (MVA) belongs to a family of methods for estimation of migration velocity. Instead of looking at travel-times in the seismic data, MVA extracts the velocity information from the migrated images. Etgen (1990) and van Trier (1990) proposed the first formulations of tomographic MVA using surface offset-domain common-image-gathers (ODCIGs) obtained by Kirchhoff migration. Recently, the tomographic MVA method has been extended to wave-equation migration velocity analysis (WEMVA), which uses the wave-equation rather than the ray-based model to resolve the velocity information (Chavent and Jacewitz, 1995; Biondi and Sava, 1999). The wave-equation-based methods are more accurate than ray-based methods, because the wave-equation better describes wave-propagation physics and provides physically more realistic sensitivity kernels for the velocity update. The wave-equation model can behave quite differently than the ray-based model when applied to complex velocity models.

To estimate velocity, WEMVA solves an optimization problem. Evaluating the flatness of the subsurface angle-domain common-image gathers (ADCIGs) is currently a popular choice when forming WEMVA objective functions (Biondi and Sava, 1999; Clapp and Biondi, 2000; Biondi and Symes, 2004). The objective function is usually optimized by applying gradient-based algorithms. The computation of the gradient is performed in two steps: 1) computation of a perturbation in the migrated image domain, and 2) back-projection of the image perturbation into the velocity model using the image-space wave-equation tomographic (ISWET) operator (Sava and Biondi, 2004).

Several ADCIGs-based WEMVA objective functions have been proposed in the literature. The stack power maximization (SPM) method directly maximizes the stack of the

ADCIGs of all detectable angles, but similar to full-waveform inversion (FWI) (Tarantola, 1984), it is prone to cycle-skipping when the starting model is not sufficiently accurate and the data do not contain very low frequencies (Symes, 2008). Differential-semblance optimization (DSO) (Symes and Carazzone, 1991; Shen et al., 2005; Shen and Symes, 2008) penalizes the difference between the ADCIGs of neighboring angles (or the unfocused energy on the *subsurface* ODCIGs). This objective function can achieve much better global convergence. However, the differencing operator amplifies the high-frequency (with respect to the angle axis) image variations in the ADCIGs and can generate unwanted artifacts in the gradient (Fei and Williamson, 2010), which slows down convergence.

If we compare the two objective functions that previously mentioned, notice that "stacking all angles" is a special case of the smoothing operations, and it extracts nothing but the zero-frequency component (with regard to the angle axis), while "differencing neighboring angles" extracts all frequency components but boosts higher-frequency components. The partial-stack power maximization objective function is a compromise between the two. It utilizes many non-zero frequency components as DSO does, while still using the stacking operator (in contrast to the differencing operator) with smaller windows. Having higher-frequency components in the objective functions ensures better global convergence, and not using the differencing operator avoids amplifying the high-frequency noise in the ADCIGs. Therefore the partial-stack power maximization objective function combines the merits of the SPM and DSO objective functions.

The rest of this paper is divided into two parts: first we present the mathematical formulation; then we demonstrate the effectiveness of our method with the Marmousi examples.

THEORY

Partial-Stack power maximization

For the sake of simplicity, we assume two-dimensions in our derivation; however, extending the theory to 3-D is straightforward for this method. We denote the prestack image as $I(z, \gamma, x)$, (x, z are the depth and horizontal axis, and γ is the reflection-aperture angle).

To enforce the goal of ADCIG flatness, we have multiple options in choosing the objective functions. The stack power maximization (Soubaras and Gratacos, 2007) maximizes the full angle stack of the ADCIG:

$$\max_v J_{SPM}(v) = \frac{1}{2}\|\sum_\gamma I(z, x, \gamma; v)\|_2^2, \tag{1}$$

in which $I(z, x, \gamma; v)$ is the ADCIGs migrated using the current velocity model v. This approach can yield a high-resolution model, however when the velocity error is large, the angle gathers will become strongly curved and demonstrate significant residual moveout (RMO) (different amounts of event shifts at different angles). Because we stack all angles at once, as the difference of event shifts between angles becomes bigger than half wavelength of the dominant frequency, the stacking becomes incoherent and would result in two separate events instead of one. The cycle-skipping phenomenon arises from such situations.

As another option to enforce angle-gather flatness, the differential semblance optimization (DSO) objective function proposed by Shen et al. (2005) overcomes this cycle-skipping

issue by using a local operator that operates on each individual angle and its immediate neighbors:

$$\min_v J_{DSO}(v) = \frac{1}{2}\|\frac{\partial I(z,x,\gamma;v)}{\partial \gamma}\|_2^2. \tag{2}$$

Cycle-skipping is very unlikely to happen using the DSO objective function, because the relative shifts in the gathers between one angle and its neighbors is generally very small.

However, the differencing operator is poorly conditioned (i.e., the operator is very short, spanning over only two angles and therefore requires many iterations to make all angles the same), and it magnifies the high-frequency variations of the ADCIGs along the axis of reflection angle. In several cases these high-frequency variations are not desired (for example, when high-frequency noise is present in ADCIGs, or when the variations are caused mainly by non-uniform subsurface illumination at each reflection angle).

In order to combine the advantages of both approaches, Shen and Symes (2008) introduce a bi-objective function that includes both terms using the weighted sum:

$$\min_v J_{CMB}(v) = J_{DSO}(v) - \beta J_{SPM}(v). \tag{3}$$

As expected, this approach can achieve both global and local convergence, but the disadvantage brought by the differencing operator still remains, and adjusting the parameter β ($\beta \geq 0$) might not be trivial.

The partial-stack power maximization (partial SPM) objective function is an alternative way to combine the SPM and DSO objective functions. We use a partially stacking operator that has a span between those of the full stacking operator and the differencing operator:

$$\max_v J_{PSPM}(v) = \frac{1}{2}\|\sum_\gamma \{g(\gamma) * I(z,x,\gamma;v)\}\|_2^2, \tag{4}$$

in which $*$ means convolution and $g(\gamma)$ is the windowing function whose support corresponds to the span of the partial-stacking operator. The partial-stacking objective function serves as a transition between SPM and DSO. SPM uses only the zero-frequency component (with respect to the angle axis), and DSO uses all components but zero-frequency; while the partial SPM uses both zero and non-zero frequency components. Nonetheless, partial-stack is still a low-pass stacking operator, thus it does not amplify high frequencies. Additionally, partial SPM objective function has a single term, and the user does not have to choose proper values for the relative weight parameter β as in objective function (3).

In practice, subsurface offset CIGs are more convenient for implementation than subsurface angle-domain CIGs. To find out the subsurface offset-domain counterparts of the objective functions we just discussed, Sava and Fomel (2003) showed that the transform between subsurface offset gathers and angle gathers is analogous to a Fourier transform with respect to γ and h; that is, convolution in the angle-domain corresponds to multiplication in the offset domain. Therefore we can see that for the SPM objective function, the offset-domain counterpart is to maximize the zero subsurface-offset image (because the Fourier transform of a constant function is a spike at the origin):

$$\max_v J_{SPM\_O}(v) = \frac{1}{2}||I(z,x,h=0;v)||_2^2; \tag{5}$$

for the DSO objective function in the offset-domain, the differencing operator will map to a weighting function $f(h) = h$:

$$\min_v J_{DSO\_O}(v) = \frac{1}{2}||I(z,x,h;v)h||_2^2. \tag{6}$$

Following the same logic, the partial SPM objective function will be:

$$\max_v J_{PSPM\_O}(v) = \frac{1}{2}||I(z,x,h;v)G(h)||_2^2, \tag{7}$$

in which $G(h)$ corresponds to the Fourier transform of $g(\gamma)$. Figure 1 illustrates the comparisons of the three objective functions in the angle-domain (a) and offset-domain (b).

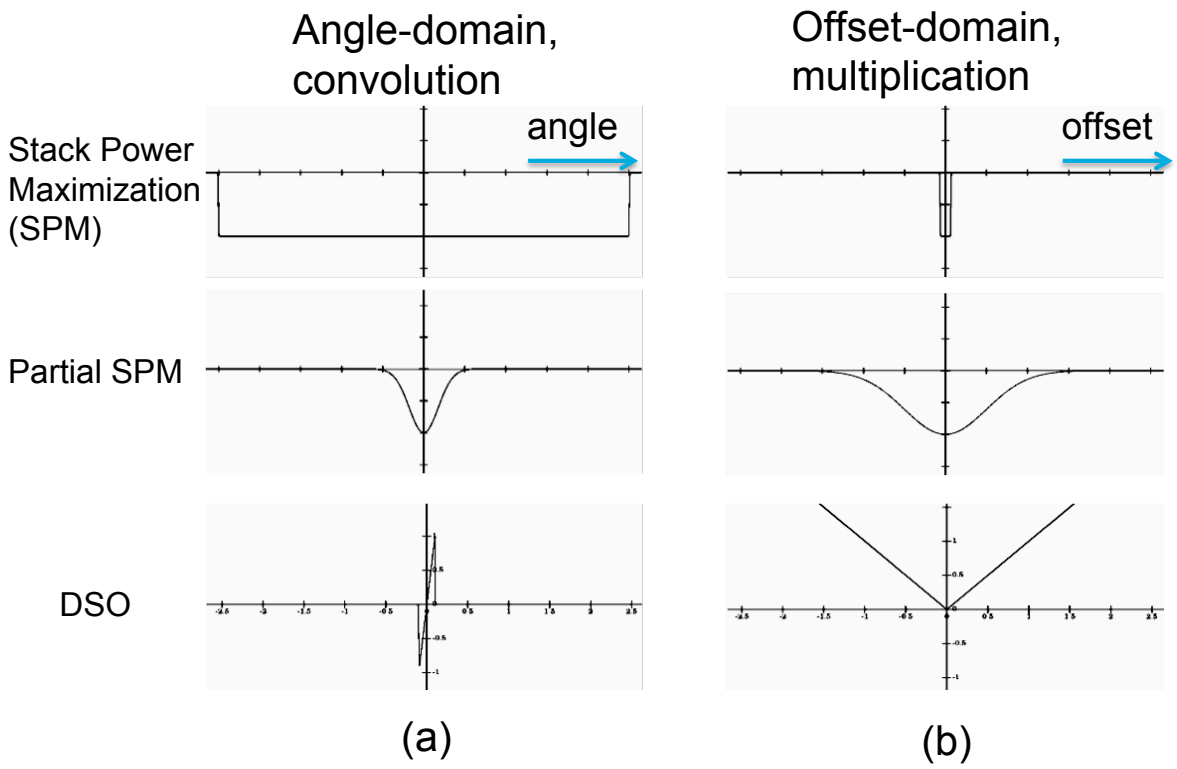


Figure 1: The comparison of the three objective functions in angle-domain(a) and offset-domain(b). [**NR**] yang2/. fig1

Balancing the impact of strong and weak reflectors

Notice that all of the objective functions we discussed previously implicitly put more weight on the strong-amplitude events, which causes the inversion to spend very little effort on the unflatness of the ADCIGs at weak reflectors. Not only will the inverse problem become less constrained, but also the inverted model might depart far from the true model, as we can see in the examples later.

There are multiple ways to normalize these objective functions so that they are more independent of reflector strength; here, for simplicity, we choose to normalize them for each

inline location x (Tang, 2011). The normalized version of objective function (7) is:

$$\max_v J_{PSPM\_OB}(v) = \frac{1}{2}\sum_x \frac{\sum_{z,h}[I(z,x,h;v)G(h)]^2}{\sum_{z,h} I^2(z,x,h)}. \tag{8}$$

The next step is to find the model gradient from the objective function, as we will use gradient-based methods to solve this optimization problem. For conciseness and without bringing confusion, we omit the variables in $I(z,x,h;v)$ and simply denote it as I, and define

$$\bar{I} = \sum_{z,h} I^2 \qquad \text{and} \qquad \bar{I_G} = \sum_{z,h} I^2 G^2(h). \tag{9}$$

Then the gradient of the objective function (8) is

$$\frac{\partial J_{PSPM\_OB}}{\partial v} = \frac{\partial I}{\partial v}\frac{(\bar{I}G^2(h) - \bar{I_G})I}{\bar{I}^2}. \tag{10}$$

Eq. (10) indicates that first we compute the image perturbation $\Delta I = \frac{(\bar{I}G^2(h)-\bar{I_G})I}{\bar{I}^2}$; then we back-project ΔI using the image-space wave-equation tomographic operator.

NUMERICAL EXAMPLES

We use the synthetic Marmousi model to test the effectiveness of the partial SPM objective function. In our implementation, we use a two-way acoustic wave-equation propagator, and for our optimization algorithm, we use non-linear conjugate-gradient with Polak-Ribiere formula for search direction.

We implement the offset-domain representation of the partial SPM objective function (eq. (7) and eq. (8)). The selection of weighting function $G(h)$ is not unique, as long as it can be considered as the frequency spectrum of a certain low-pass filter. In our examples, we use Gaussian functions for $G(h)$. We started with a wide $G(h)$ that would include most unfocused energy of the ODCIGs at non-zero subsurface offsets, as the inversion proceeds, the gathers will become more focused; we then reduce the width of $G(h)$ so that our objective function gradually approaches the traditional SPM objective function. To control the resolution of the inversion, we precondition the model gradient with a triangular smoothing operator and reduce the extent of smoothing gradually within each iteration.

Marmousi Example

The model size is 9 km in x and 3.2 km in z. The spatial sampling is 20 m. The survey geometry is of split-spread type, with sources and receivers located on the top of the model. There are 451 receivers with 20 m spacing that fully covers the model surface. We simulate 226 shots with shot spacing of 40 m. The central frequency of the Ricker wavelet we use is 10 Hz. For the inversion, we invert all frequencies simultaneously rather than from low to high frequencies. We run 40 nonlinear iterations for the inversion.

Figure 4(a) shows the true velocity model and 4(b) shows the starting model $v(z)$. Figures 4(c) and (d) show the inverted velocity model using the un-normalized partial SPM

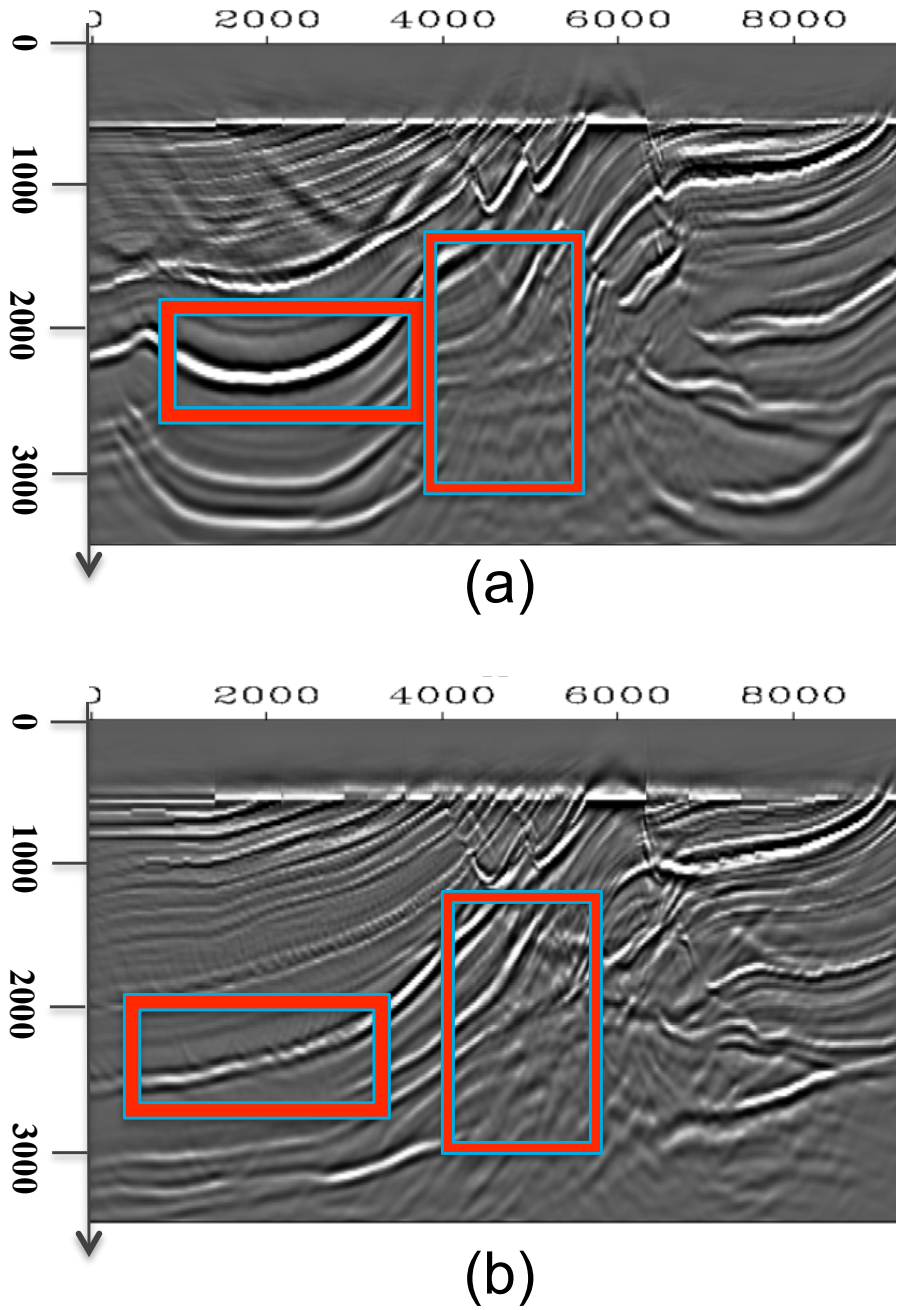


Figure 2: The migrated image (a) using the velocity model inverted from the objective function (7) (without normalization) and (b) using the initial model. The display clip we use for (a) is 4 times of the clip used in (b), and we can see that in (a) the single reflector marked in the left box becomes very strong and coherent. In contrast, many weaker reflectors (for e.g., marked in the right box) that are initially present in the initial image (a) are imaged much poorly in (b). Therefore it is important to give the weak reflectors more weight in the objective function. [**NR**] yang2/. fig3-img-comp-normalize

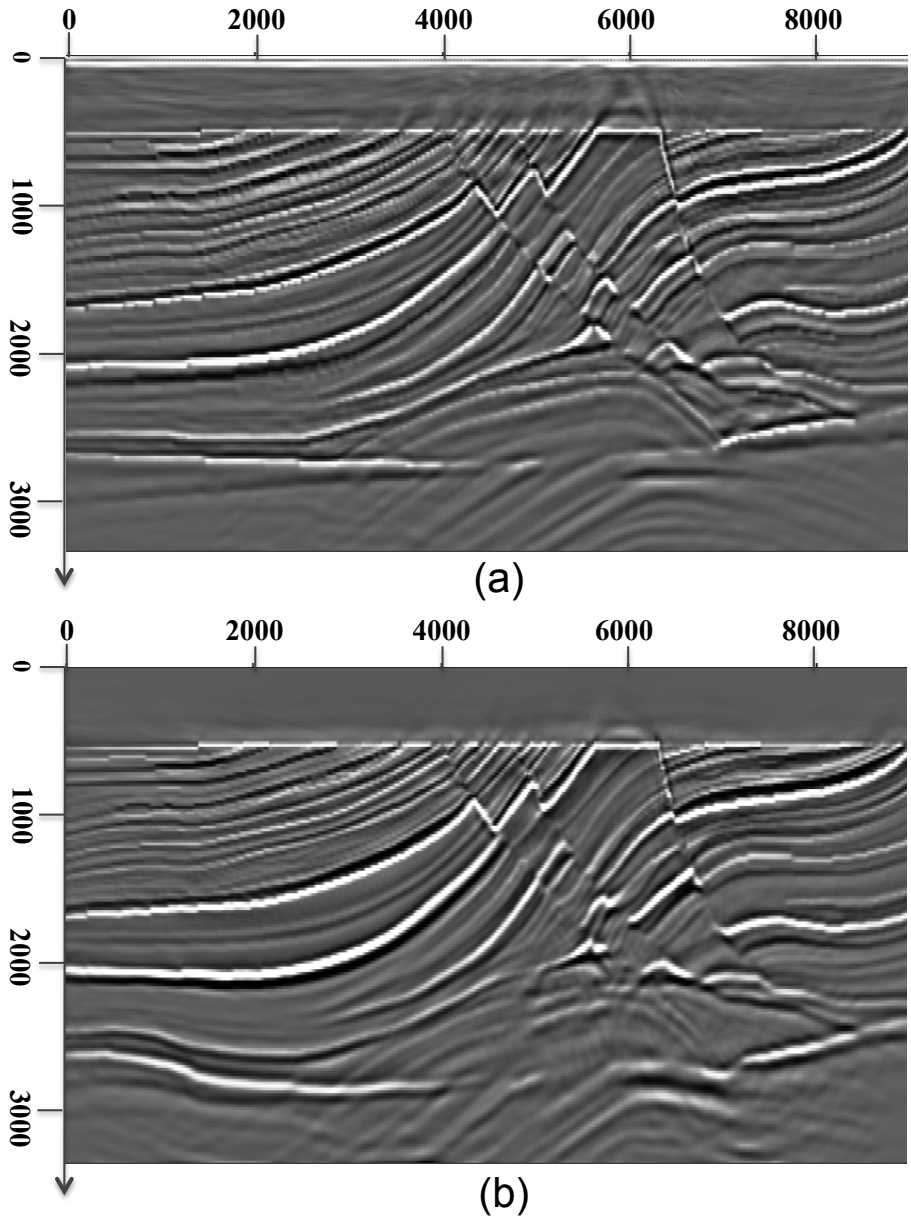


Figure 3: The comparison of the migrated images using the true velocity model (a) and using the inverted model with normalized partial SPM objective function (b). The two images match well up to 2.2 km depth. [**NR**] yang2/. fig4-img-comp-PSPM

objective function and the normalized partial SPM objective function respectively. As we can see, the inverted result in 4(c) is stuck in local minima, because without normalization, the inversion will attempt to increase the focusness of a few large, strong reflectors, while ignoring and even sacrificing the coherency of smaller, weaker events. The observation of the corresponding migrated images in figure 2 further confirms our conclusion. The result in figure 4(d) shows global convergence towards the correct model. The long-wavelength part of the velocity model is well captured up to depth 2.2 km, as can be seen from the migrated images in figure 3.

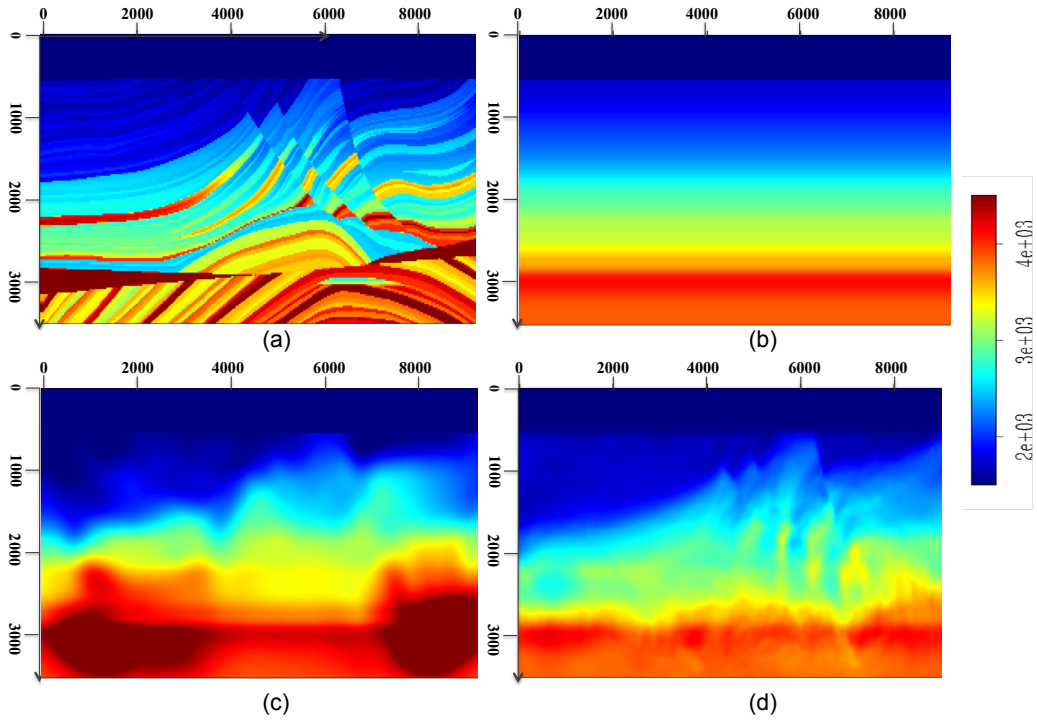


Figure 4: The Marmousi true velocity model (a), the starting velocity model (b), the inverted velocity models using partial SPM objective function without normalization (c) (eq. (7)) and with normalization (d) (eq. (8)). We can see that without normalization, the inversion yield a model that is very different from the true one. [**NR**] yang2/. fig2-vmarms

CONCLUSION

We propose to use the partial stack power maximization objective function in wave-equation migration velocity analysis. This objective function merges the advantages of both conventional stack power maximization and differential-semblance-optimization objective functions, and it can achieve good global convergence, while retaining the relatively high resolution of the stack power maximization objective function. We have successfully applied our approach to the Marmousi model. We have also verified that the normalization of the reflector amplitude in the objective function is not only preferred but necessary in the Marmousi case.

ACKNOWLEDGEMENT

The authors thank Chevron Energy Technology Company for the permission to publish, and thank Yue Wang and Lin Zhang for suggestions and comments. The first author thanks Prof. Claerbout for the initial discussion on WEMVA objective functions.

REFERENCES

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

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

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

Clapp, R. G. and B. Biondi, 2000, Tau domain migration velocity analysis using angle crp gathers and geologic constraints: SEG Technical Program Expanded Abstracts, **19**, 926–929.

Etgen, J., 1990, Residual prestack migration and interval velocity estimation: PhD thesis, Stanford University.

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

Gardner, G. H. F., 1974, Elements of migration and velocity analysis: Geophysics, **39**, 811.

Sattlegger, J. W., 1975, Migration velocity determination: Part i. philosophy: Geophysics, **40**, 1–5.

Sava, P. and B. Biondi, 2004, Wave-equation migration velocity analysis. I. Theory: Geophysical Prospecting, **52**, 593–606.

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

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

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

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

Symes, 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.

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

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

van Trier, J., 1990, Tomographic determination of structural velocities from depth migrated seismic data: **66**.

Anti-noise wave-equation traveltime inversion and application to salt estimation

Xukai Shen

ABSTRACT

The convergence of full-waveform inversion relies heavily on a good starting model. Such a model can be provided by wave-equation traveltime inversion. However, results from wave-equation traveltime inversion are susceptible to data noise, which is particularly obvious in real data applications. I propose to address this problem by replacing the real data term in back-propagation kernel with a synthetic data term that honors the traveltime information. This modification makes the inversion immune to noise in recorded data waveforms, yet results in a model that is still good for subsequent full-waveform inversion. I demonstrate this with an example of salt estimation. Salt estimation by direct full-waveform inversion is challenging due to a combination of inadequate starting models and insufficient low-frequency data. Reflection wave-equation traveltime inversion with my modification provides a good starting model for subsequent full-waveform inversion of salt estimation, even with data that lack low-frequency components.

INTRODUCTION

Data-domain tomographic methods update velocity using mismatches between observed data and modeled data. The mismatches usually include traveltime (first-order) and waveform (second-order) differences. Both traveltime and waveform mismatches are used by Full Waveform Inversion (FWI) (Tarantola, 1984; Pratt et al., 1998; Mora, 1987), leading to high-resolution results but requiring an accurate starting model. On the other hand, Wave-equation Traveltime Inversion (WTI) (Luo and Schuster, 1991) estimates velocity model by minimizing only the traveltime difference between observed data and modeled data using the wave equation. As a result, WTI differs from FWI in two ways: first, WTI is not affected by cycle skipping, and second, WTI results tend to have low resolution. However, the low-resolution result from WTI may be a good starting model, and subsequent application of FWI can obtain a high-resolution result (Shen et al., 2012).

In WTI, the back-propagation term depends linearly on the traveltime differences and time-shifted observed data. The traveltime difference in the back-propagation term makes WTI insensitive to cycle-skipping. The time-shifted observed data, on the other hand, will compromised the inversion results if observed data is contaminated with strong noise. This is likely to be true for field data, especially data acquired on land. To avoid the noise problem in field data, I propose to replace the inverse problem with a series of synthetic problems whose data traveltimes are the same as those of the original problem. When we solve the new synthetic problems, we no longer suffer from potential noise in field data, but we still obtain a similar result, since the new problems have the same data traveltimes as the original problem.

The effectiveness of the new method is demonstrated with an salt estimation example. Salt estimation with FWI is difficult. Synthetic tests suggest that even with a intuitively close starting model, very low-frequency data is required for successful inversion (Koo et al., 2009; Vigh and Starr, 2008). To lessen the requirement for low-frequency data, WTI is a good candidate. However, past success with WTI mostly has used diving waves or direct arrival (Sirgue et al., 2009; Virieux and Operto, 2009), with very few reflection cases (Zhang et al., 2011). I will demonstrate that reflection data WTI with the aforementioned modification can indeed provide a good starting model for subsequent FWI.

ANTI-NOISE WAVE-EQUATION TRAVELTIME INVERSION

The original WTI method minimizes the following objective function:

$$f(\mathbf{m}) = \sum_{s,r} \Delta\tau^2(\mathbf{m}), \tag{1}$$

where $\mathbf{m}$ is the velocity or slowness model to be estimated, s and r are source and receiver locations, respectively, and $\Delta\tau$ is defined as the time lag that maximizes the cross-correlation between the observed trace and the modeled trace. By this definition, $\Delta\tau$ is also the traveltime difference between corresponding events in synthetic data and recorded data. From this objective function, the back-propagation term for gradient calculation is

$$v_{src}(t) = \Delta\tau \frac{\dot{d_{obs}}(t+\Delta\tau)}{\dot{d_{obs}}(t+\Delta\tau) * \dot{d_{syn}}(t)}, \tag{2}$$

where $\dot{d_{obs}}$ and $\dot{d_{syn}}$ denote first-order time derivatives of observed and synthetic data, respectively. $*$ denotes cross-correlation. It can be seen that this back-propagation term is proportional to the noise strength in the observed data. To avoid potential noise effects, I will solve an equivalent synthetic problem. Notice that only the traveltime difference is used in the objective function, I can assume that the only difference between the observed data and the modeled data is the static time-shift $\Delta\tau$. This assumption is mathematically expressed as

$$d_{obs}(t+\Delta\tau) \approx d_{syn}(t). \tag{3}$$

Substituting this into equation 2, we get the new back-propagation term

$$v_{src}(t) = \Delta\tau \frac{\dot{d_{syn}}(t)}{\dot{d_{syn}}(t) * \dot{d_{syn}}(t)}. \tag{4}$$

The new back propagation term still honors the traveltime difference, yet it is completely noise-free. In other words, we use only the traveltime information from the observed data to perform the inversion. Physically, this modification means we are solving an equivalent synthetic problem at each iteration, where the new observed data is defined by equation 3. This modification completely removes the noise interference from the actual data.

SALT ESTIMATION

With the proliferation of sub-salt reservoir discoveries in recent years, sub-salt imaging has become increasingly important. It has been shown (Shoshitalshvill et al., 2006) that better

definition of the salt improves sub-salt imaging. In this section, I will show how the modified WTI can help improve salt estimation by supplying FWI with a better starting model.

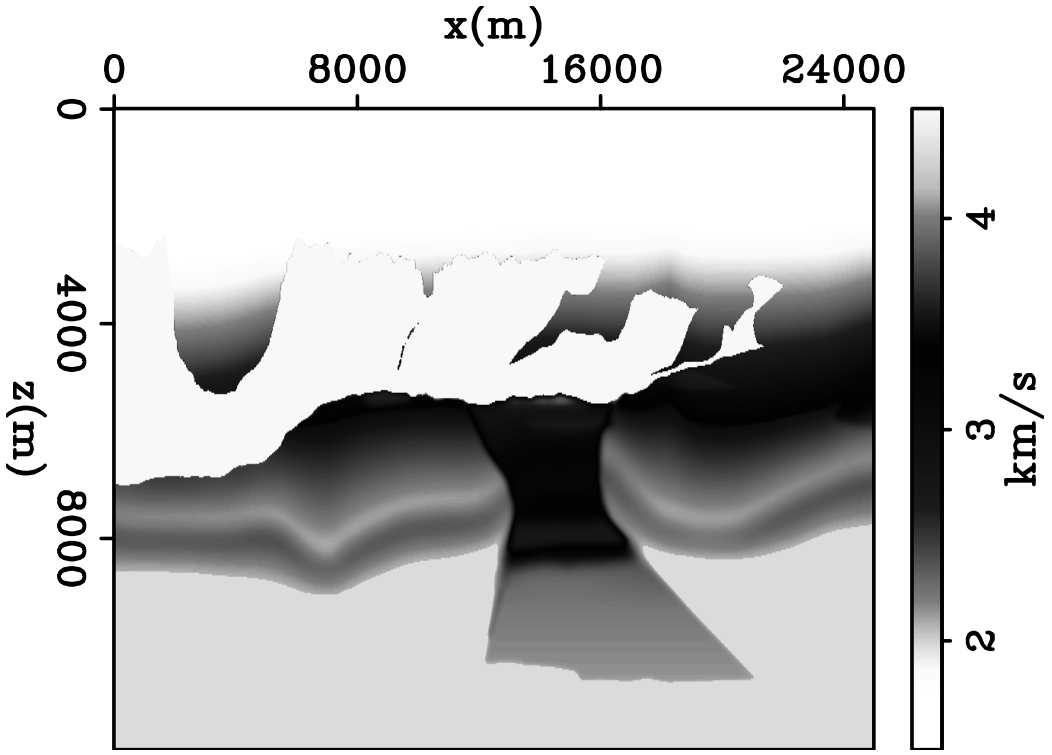


Figure 1: True velocity model for inversion. [**ER**] xukai1/. vtruebw

The salt model in figure 1 is taken from part of the BP 2004 Velocity Benchmark, where the top of salt formation is rugose. I modeled a total of 96 shots with constant-density acoustic modeling. Shot spacing was 250 meters. The offset range of each shot gather was from −8 km to 8 km. Both modeling and inversion were performed on a 25 m grid spacing. All inversion data lack low-frequency components (Figure 3). Only salt-related reflections are used in the inversion, with refractions, direct arrivals, and water-bottom reflections muted (Figures 4 and 5) . The starting model (figure 2) has an accurate water-bottom definition and a close-to-accurate sediment velocity down to the salt top; however, the salt top is a smoothed version of a coarse set of top-of-salt picks. Such a starting model is similar to a real data scenario, since the water bottom is usually well constrained, along with the sediment velocities down to the top of salt. Two types of inversions were performed, one performing FWI from the starting model and the other applying WTI followed by FWI.

Due to the inaccurate positioning of the salt top in the starting model, cycle-skipping of the salt-top reflection exists between the shot gathers modeled from the true model and the ones from the starting model (Figures 4 and 6). As a result, FWI from the starting model can not accurately put the modeled reflection at the correct traveltimes (Figures 5 and 8). On the other hand, WTI does model reflections to reasonably accurate traveltimes (Figures 4 and 7), enabling subsequent FWI to further match the detailed waveforms of observed reflections (Figures 5 and 9).

The result of FWI from the starting model is shown in figure 10, the result of WTI

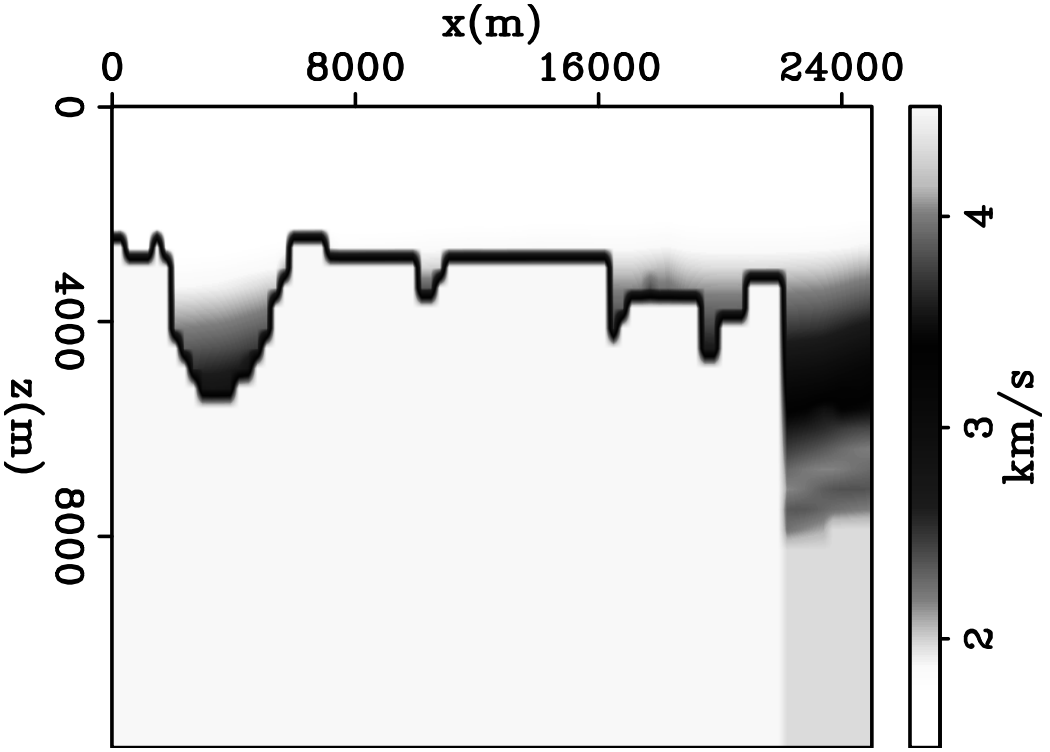
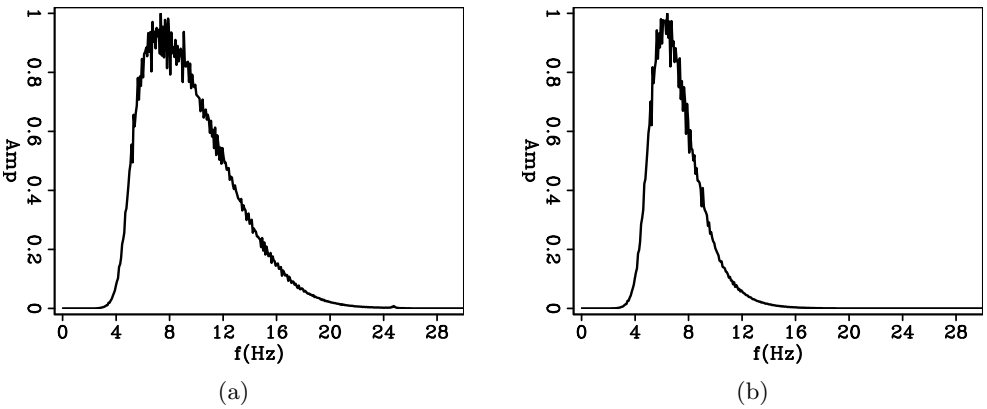


Figure 2: Starting velocity model for inversion. [**ER**] xukai1/. vinitbw



(a) (b)

Figure 3: Spectrum of data used for a) traveltime inversion b) waveform inversion. [**ER**] xukai1/. dwt,dwi

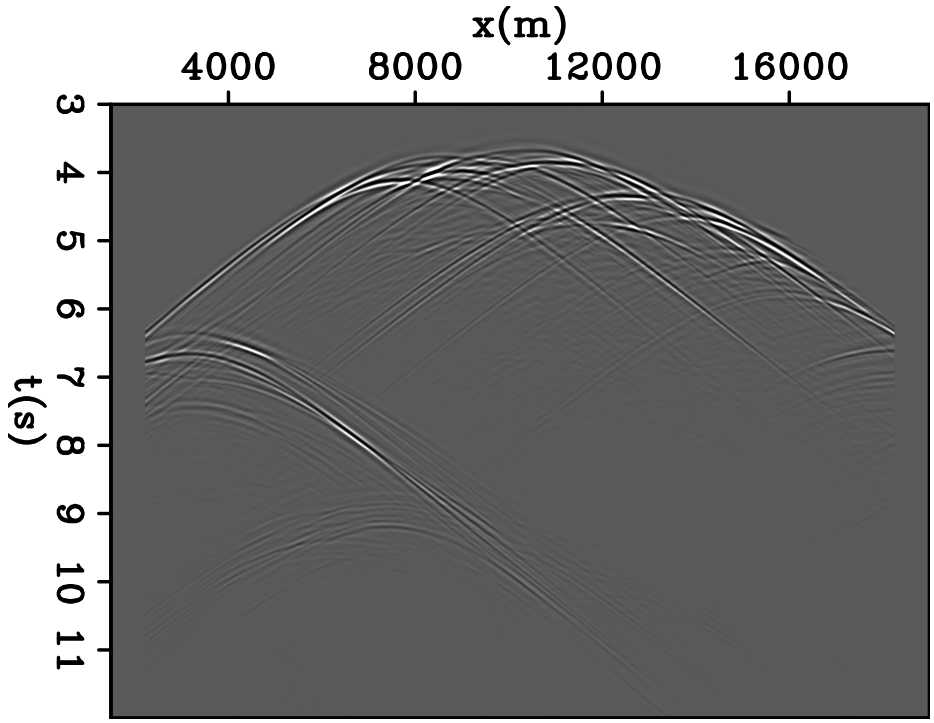


Figure 4: One shot record from the true model for WTI (reflection only). [**CR**] xukai1/. dobsm

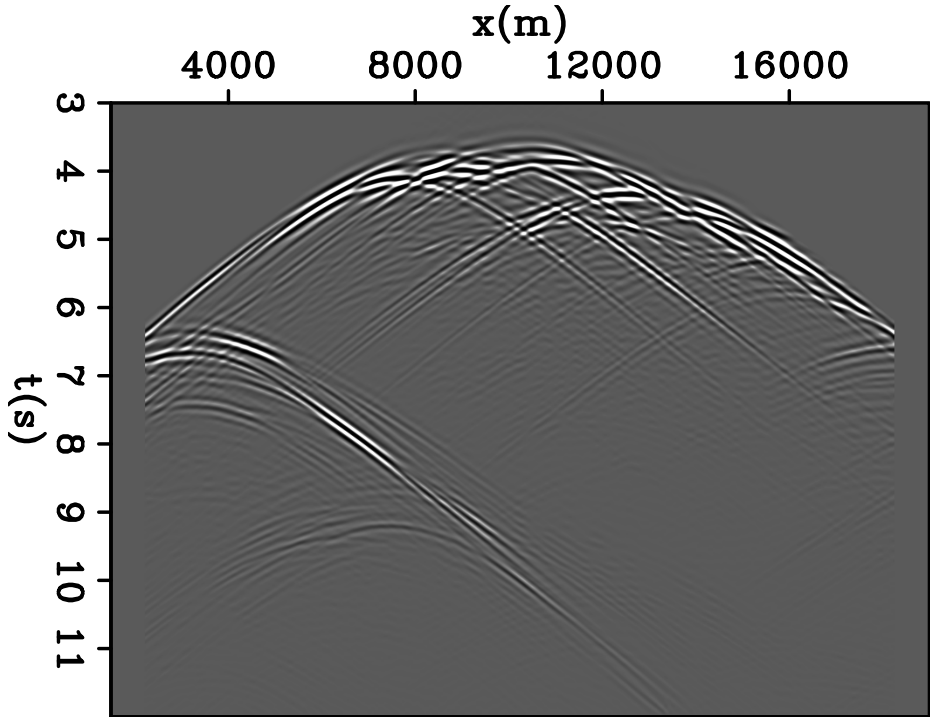


Figure 5: One shot reflection from the true model for FWI (reflection only). [**CR**] xukai1/. dobsfwi

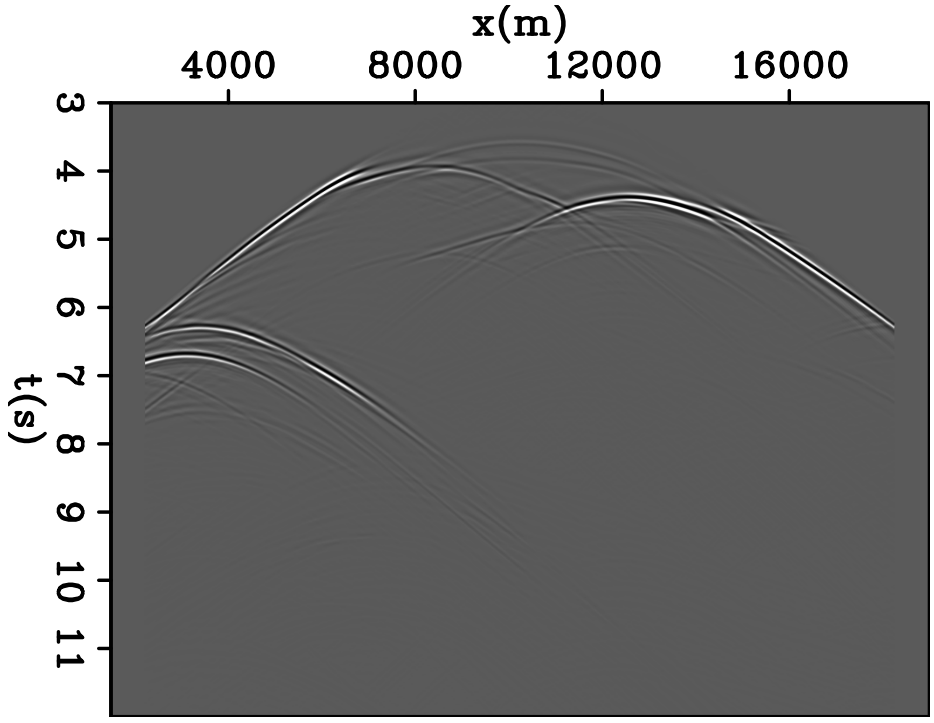


Figure 6: One shot record from the initial model (reflection only). [**CR**] xukai1/. dcali

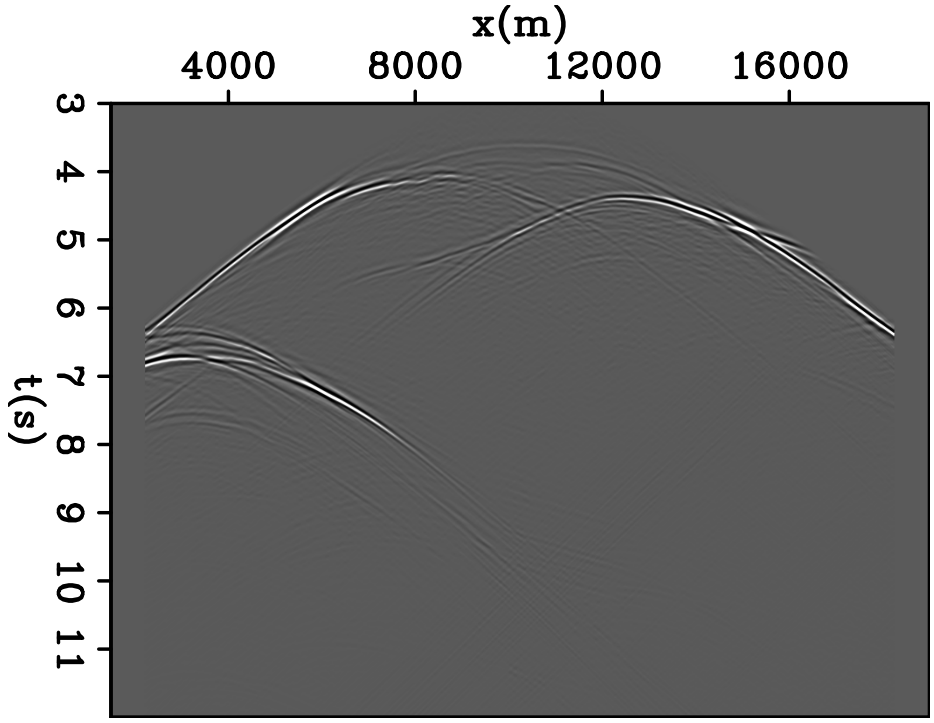


Figure 7: One shot record from the WTI result (reflection only). [**CR**] xukai1/. dcalwti

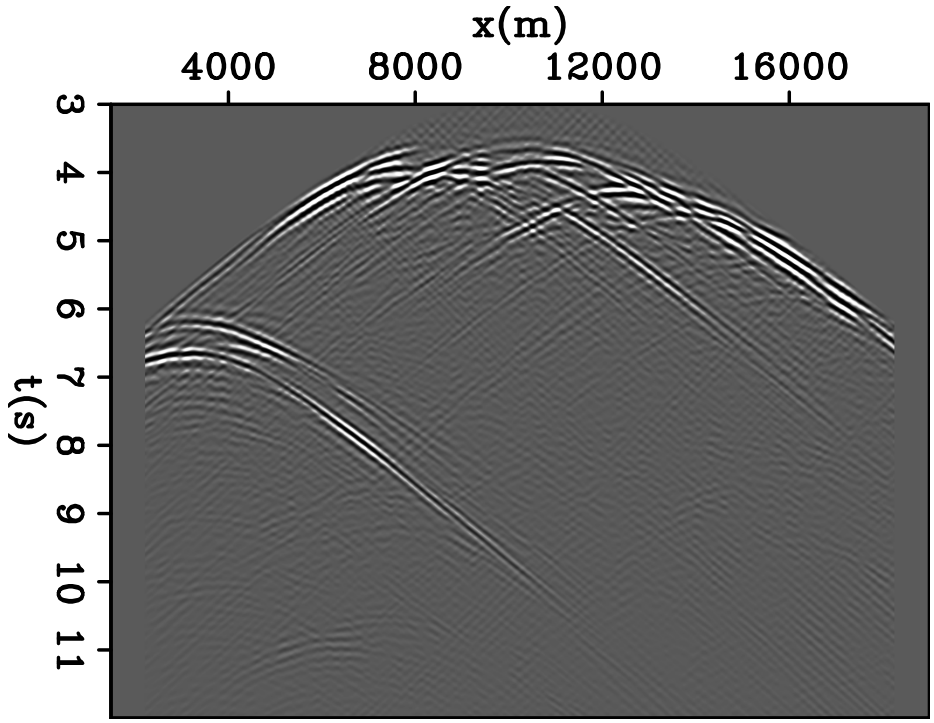


Figure 8: One shot record reflection from the direct FWI result (reflection only). [**CR**] xukai1/. dcaldfwi

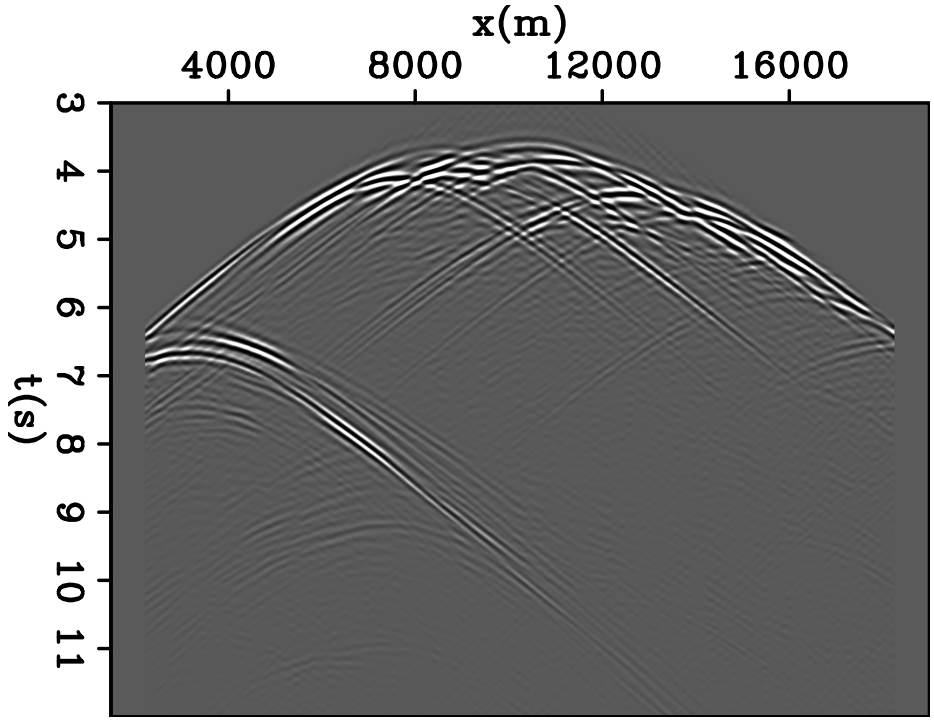


Figure 9: One shot record from the FWI (after WTI) result (reflection only). [**CR**] xukai1/. dcalfwi

from the starting model is shown in figure 11, and the result of FWI after WTI is shown in figure 12. It can be seen that direct FWI neither converged to the true model nor gave a good indication of where the salt boundaries are; to make matters worse, it even produced spurious salt inclusions near the top of salt. However, WTI was able to correct the location of the salt top so that subsequent FWI accurately estimated the salt top with very high resolution. In addition, the result of FWI after WTI also yielded a good indication of the location of salt inclusions, salt flanks and the salt bottom, all of which are important information for salt-geometry estimation and sub-salt imaging.

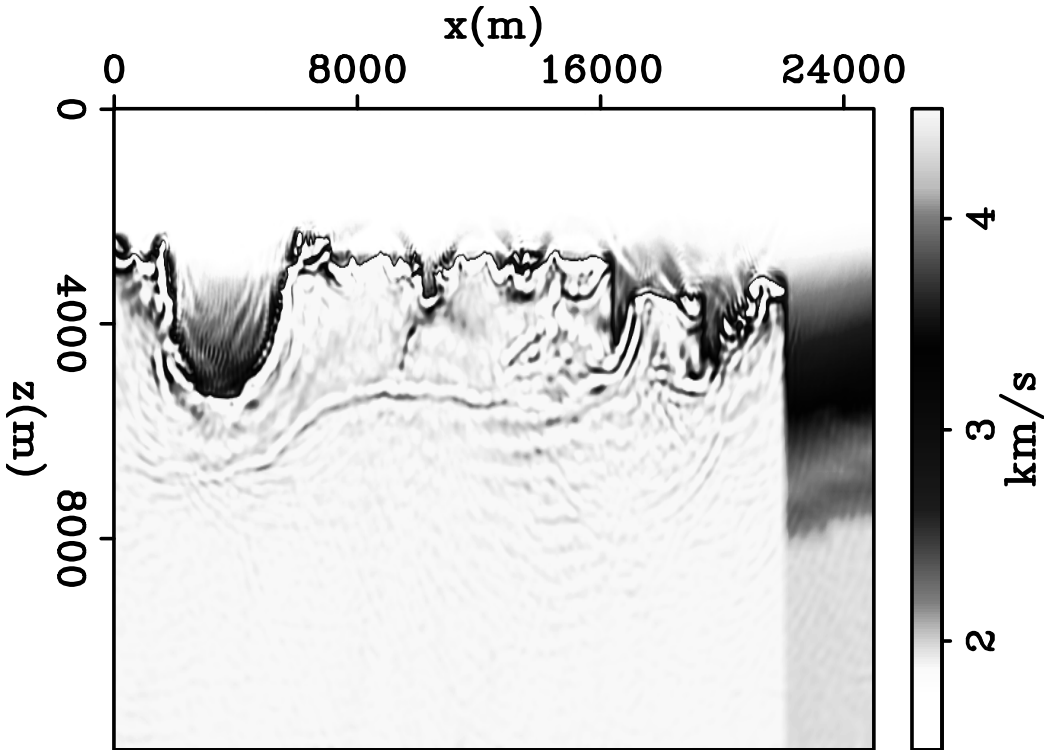


Figure 10: Result from FWI using the starting model. [**CR**] xukai1/. vddbw

CONCLUSION

I devised a modification to the existing WTI method to make it immune to noise in recorded data. The modification substitutes the original problem with a series of synthetic problems where mismatched traveltime is honored. Using reflection data only, application of this modified WTI followed by FWI was able to give high-resolution results of the salt top and a good indication of salt flank and bottom locations, whereas direct FWI from the starting model failed to do so.

ACKNOWLEDGMENTS

I would like to thank Dr. Luo from Saudi Aramco for suggestions/discussions on this topic.

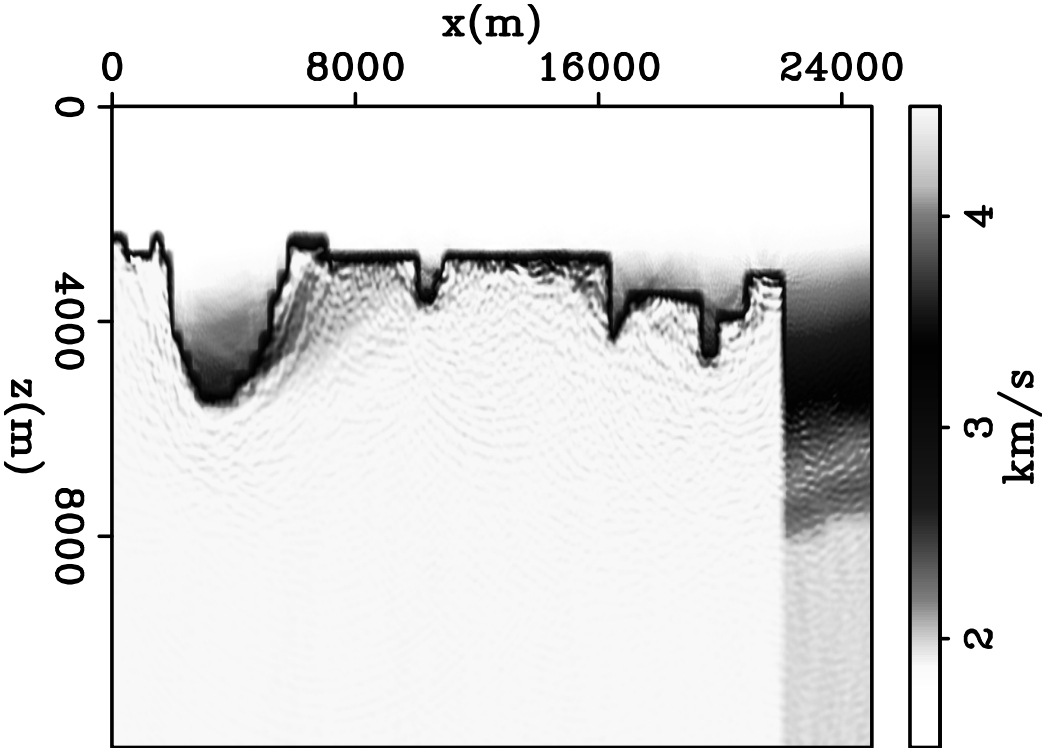


Figure 11: Result from WTI using the starting model. [**CR**] xukai1/. vwtbw

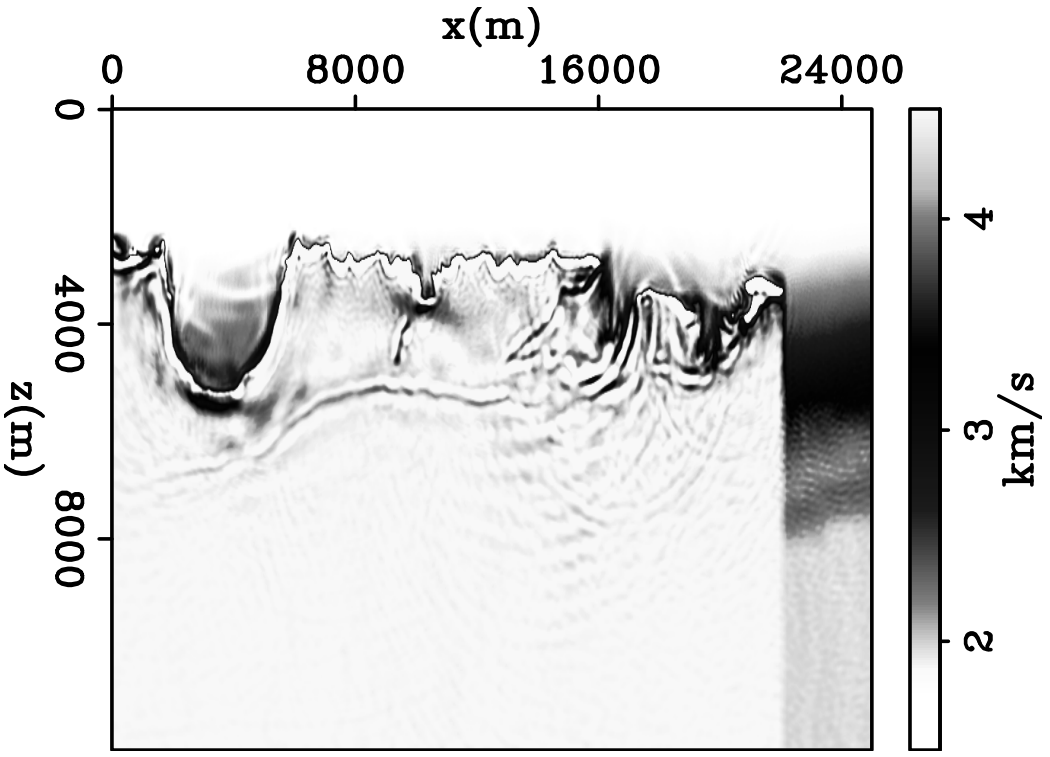


Figure 12: Result from FWI using the WTI inverted model. [**CR**] xukai1/. vwibw

REFERENCES

Koo, N., C. Shin, and Y. H. Cha, 2009, Sequentially ordered single-frequency 2-D acoustic waveform inversion in the Laplace-Fourier domain: SEG Technical Program Expanded Abstracts.

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

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

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

Shen, X., T. Tonellot, Y. Luo, T. Keho, and R. Ley, 2012, A new waveform inversion workflow: Application to near-surface velocity estimation in Saudi Arabia: SEG Technical Program Expanded Abstracts.

Shoshitalshvill, E., S. Michell, J. Etgen, D. Chergotis, and E. Olson, 2006, Improving resolution of top salt complexities for subsalt imaging: SEG Technical Program Expanded Abstracts.

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

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

Vigh, D. and E. W. Starr, 2008, 3d prestack plane-wave, full-waveform inversion: Geophysics, **73**, 135–144.

Virieux, J. and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.

Zhang, S., G. Schuster, and Y. Luo, 2011, Wave-equation reflection traveltime inversion: SEG Technical Program Expanded Abstracts.

Angle gather recovery using Iterative Soft Thresholding

Robert G. Clapp

ABSTRACT

Multi-dimensional angle gather construction for AVA and velocity analysis is a computational challenge due primarily to the accompanying increase in volume size which forces the gathers to be stored in a computationally more expensive memory level. Compressive sensing can be used to mitigate this challenge as long as the full angle gathers can be successfully recovered. Multi-dimensional wavelet transforms is a sufficiently sparse basis function to allow for 90% reduction in needed correlations. A modified version of iterative soft thresholding, which applies a different thresholding approach to different wavelet levels, proves a successful l_1 inversion scheme.

INTRODUCTION

Angle gathers are a useful tool byproduct of migration for both velocity and rock property analysis. There are several different methods for producing angle gathers (Crawley et al., 2012), but one of the most useful is based on correlating shifted version of the source and receiver fields (Sava and Fomel, 2003, 2006). The problem with this approach, particularly when creating multi-dimensional gathers, is that the volume size of the imaging domain increases by one to three orders of magnitude making the dominant cost reading/writing to distant memories (from main memory rather than a cache, across the PCI Bus, or from disk).

Donoho (2006) offers an approach termed *compressive sensing* as a potential solution to this computation and storage problem. In compressive sensing, a random subset of the desired measurements is made. An inversion problem is then set up to estimate in an ℓ_1, or preferably ℓ_0, sense, a sparse basis function that fully characterizes the desired signal. Clapp (2011) showed that sub-surface offset gathers are highly compressible using the multi-dimensional wavelet transform. Clapp (2012) showed promising recovery of angle domain gathers using a Stagewise Orthogonal Matching Pursuit algorithm (StOMP) (Donoho et al., 2006). The results were encouraging and were successful enough to be used as a tool to reduce the cost of velocity estimation(Zhang et al., 2013). Zhang et al. (2013) was only able to achieve a factor of two reduction of sub-surface offset correlations before seeing degradation in the result. Hennenfent and Hermann (2005) offers another ℓ_1 inversion approach called Iterative Soft Thresholding. This approach characterizes the estimation problem as a non-linear problem eliminating small model components (thresholding) at each iteration. The thresholding is relaxed as a function of non-linear iteration.

In this paper I apply the IST algorithm to the angle gather estimation problem. I show that, in conjunction with a modified sampling criteria and a wavelet-level based thresholding scheme, a higher level of data reduction is possible.

COMPRESSIVE SENSING

Compressive sensing is a statistical technique whose start is usually traced back to Donoho (2006), but whose start could be placed as early as the basic pursuit work of Mallat and Zhang (1993). Compressive sensing is simply a missing data problem with a well chosen preconditioning operator (resulting in a sparse basis model/function) solved with a ℓ_0/ℓ_1 solver. Using the conventional SEP nomenclature of fitting goals it can be written as

$$\mathbf{0} \approx \mathbf{W}(\mathbf{d} - \mathbf{S}\mathbf{p}) \tag{1}$$

where $\mathbf{d}$ is the data, $\mathbf{W}$ is the masking operator (composed of 0s and 1s), $\mathbf{S}$ is the preconditioning operator, and $\mathbf{p}$ is the preconditioning variable.

Compressive sensing theory states that if you can find a transform operator which makes $\mathbf{p}$ consist of only n number of unknowns as long as you collect kn *random* data points, where k is usually in the range of 5, you can fully recover p. To reduce the amount of data points significantly $\mathbf{S^T}$ must transform into a basis function that is very sparse. One such choice, using the multi-dimensional wavelet transform (Villasenor et al., 1996), proves to produce a sparse basis function for subsurface offset gathers. Clapp (2011) showed that by throwing away up to 98% of $\mathbf{p}$, can still fully recover a sub-surface offset gather (see Figures 1 and 2). The problem that both Clapp (2011, 2012) attempted to address was what ℓ_1 approach to use to find the sparse basis function.

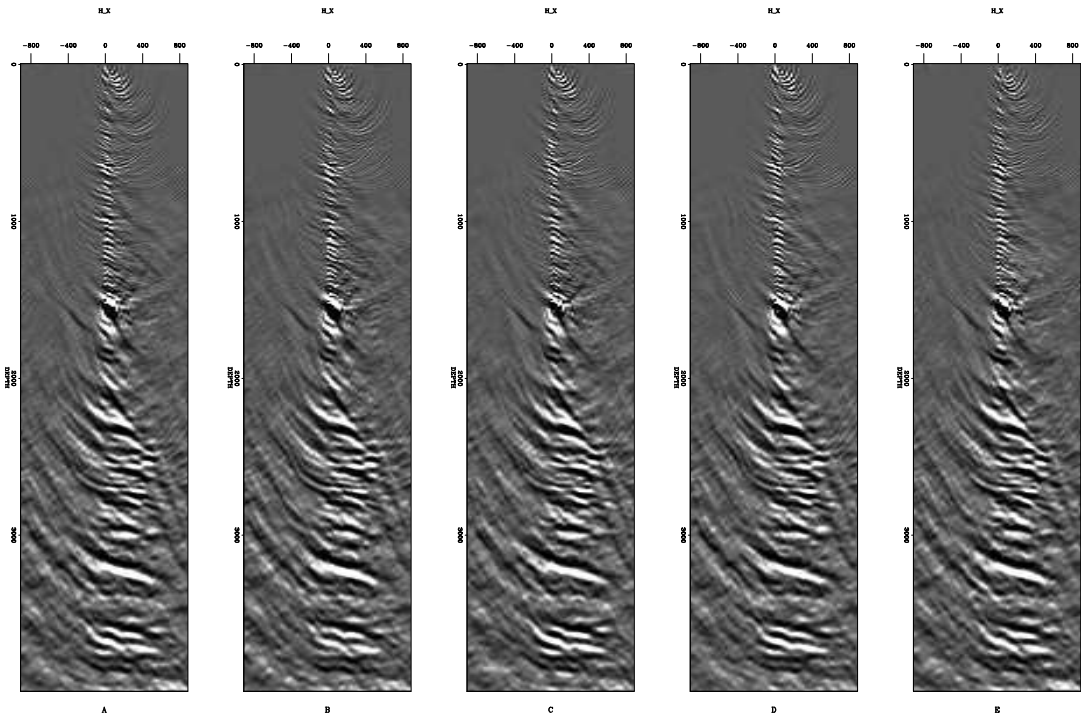


Figure 1: Five neighboring subsurface offset gathers. B and C are one midpoint in X before and after A. E is one midpoint in Y before A. Note the spatial similarity of these gathers. [**ER**] bob1/. raw

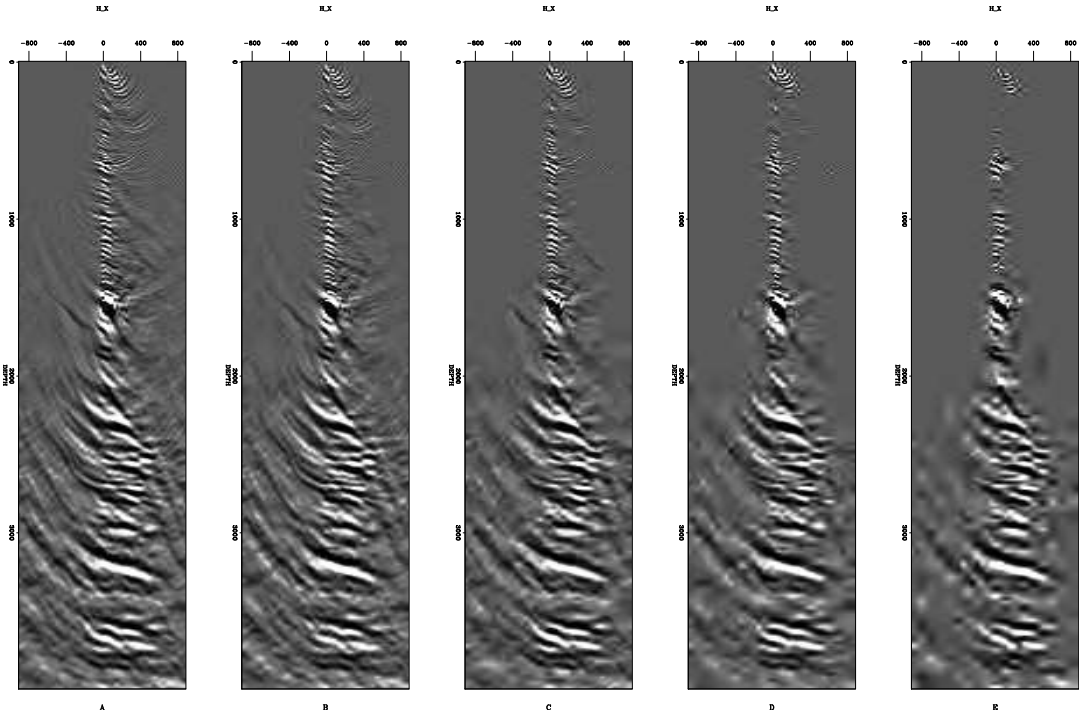


Figure 2: The result of zeroing the smallest values of the wavelet domain representation. All five panels show the same sub-surface offset gather shown in Figure 1. A shows the result of clipping 90% of the values; B, 95%; C, 98%; D,99%, and E,99.5%. Note how the reconstructed gather is nearly identical up to a 98% clip. [**ER**] bob1/. threshit

IST

Iterative soft thresholding is a greedy algorithm along the same lines as matching pursuit (Donoho, 2006), and to some extent the Stagewise Orthogonal Matching Pursuit (StOMP) (Donoho et al., 2006) stomp algorithm. It attempts to find the ℓ_1 solution to the problem by slowly allowing more and more non-zero elements into the basis function. There are three basic parts to the IST algorithm: find the step size, finding the gradient, and soft thresholding.

The step size calculation relies on finding the largest eigenvalue of the Hessian of the ℓ_2 problem, or $\mathbf{S^T W^T W S}$. This is done using the Rayleigh power method. The basic algorithm can be seen in algorithm 1. After a few (20 or so) iterations s converges to

Algorithm 1 Rayleigh power method

Fill $\mathbf{g_0}$ with random numbers
for Rayleigh iteration i **do**
  $\mathbf{g_i} = \mathbf{S^T W^T W S g_{i-0}}$
  $s = \mathbf{g^T g}$
  $\mathbf{g_i} = \frac{\mathbf{g_i}}{\mathbf{s}}$
end for

largest eigenvalue and $\mathbf{g}$ to the largest eigenvector.

The quantity $\frac{1}{s}$ is used in a modified steepest descent algorithm shown in algorithm 2. The unique portion of the above algorithm is the thresholding step. The basic idea is to set

Algorithm 2 IST algorithm

$\mathbf{x = 0}$
for Non-linear iteration **do**
  for Linear iteration **do**
    $\mathbf{gg} = \frac{\mathbf{WSx}}{\mathbf{s}}$
    $\mathbf{g} = \frac{\mathbf{S^T W^T gg}}{\mathbf{s}}$
    $\mathbf{x = x + g}$
    Threshold $\mathbf{x}$
  end for
end for

$\mathbf{x}$ values below some amplitude to zero. The threshold decreases as a function of non-linear iteration *in*. Hennenfent and Hermann (2005) suggests a decreasing thresholding based on percentiles in the initial gradient. For example throwing the bottom 99%, 98%, 95%, 90%, 80%, 60%, 35%, 0% of data values at successive iterations. Figure 3 shows a portion of the multi-dimensional wavelet domain before and after thresholding at the 98% level. Note how even though nearly all coefficients are zeroed at this level the plots are very similar.

Figure 4 shows the result of performing angle gather transform on the gathers seen in Figure 1. Note how most of the gathers are generally flat. Figure 5 shows an example of following the above algorithm to recover the basis function after randomly throwing away 90% of the correlations. Figures 6 and 7 show the resulting angle gathers. Note how we have nearly perfectly recovered the subsurface offset, and the corresponding angle gathers, at larger depths, but have not recovered as well shallower depths. A closer look at the

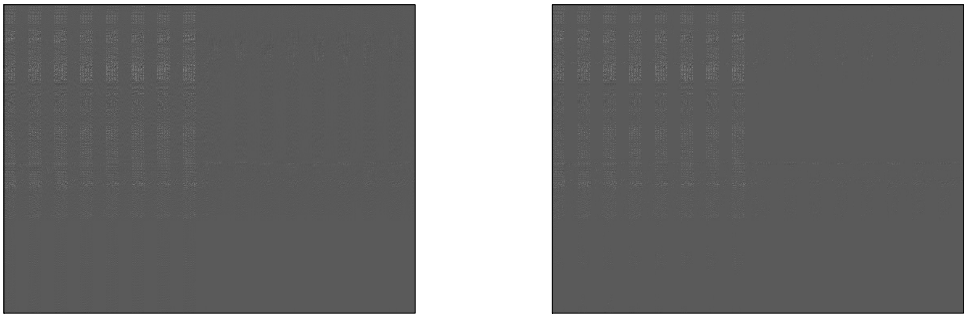


Figure 3: The left panel shows the result of performing a multi-dimensional wavelet transform on the gathers shown in Figure 1. The right panel is the result of zeroing the bottom 98% of the wavelet values. [**ER**] bob1/. coefs

sub-surface offset gathers in Figure 1 shows the problem with a purely random sampling of the sub-surface gathers. Note how at shallower depths no coherent energy maps to large offsets and shallow depths due to an accurate velocity model and good focusing. Therefore at these depths we are not obeying the sampling criteria of compressive sensing.

Clapp (2012) suggested modifying the random sampling criteria so that only points within an expanding cone are selected. Figure 8 shows one such random sampling. Note how it somewhat tracks coherent energy seen in the full sub-surface offset gathers seen in Figure 1. Figure 9 shows the recovered wavelet domain. Note how more energy is less randomly focused compared to Figure 5 and closer to the wavelet basis shown in Figure 3. The resulting offset gathers, shown in Figure 10, show less random looking noise than the result shown in Figure 6. In addition the shallower depth image is better recovered than what is seen in Figure 6. The resulting angle domain gather, shown in Figure 11, shows marked improvement in the angle domain image at shallower depths. Particularly, notice the event marked 'M' in the panel which has the incorrect amplitude in Figure 6.

THRESHOLDING AS A FUNCTION OF WAVELET LEVEL

A wavelet transform involves applying a pair of filters, which recover the low and high frequency (or wavenumber) portion of an image in a time (or space) variant manner. The lower wavenumber portion of the transform can then be sent through the same two filters. The number of times this pair of filters is applied is referred to as a series of levels. In 1-D, each levels high-pass filter produces half as many elements as the previous level. In the multidimensional case, this size decrease pattern is modified. Figure 12[1] demonstrates this phenomenon in 2-D; right and down shows successive lower levels of the high-pass portions of the image. Note the small size of the lowest wavenumber portion of the largest level of the image compared to the first level's high pass portion.

Figure 12 and the left panel of Figure 5 demonstrate another important point of many fields, sparsity generally decreases with increasing levels. Note how the top left portion of each image are very dense with information compared to other levels. Figure 13 demon-

[1]http:www.wikipedia.org/wiki/Discrete_wavelet_transform

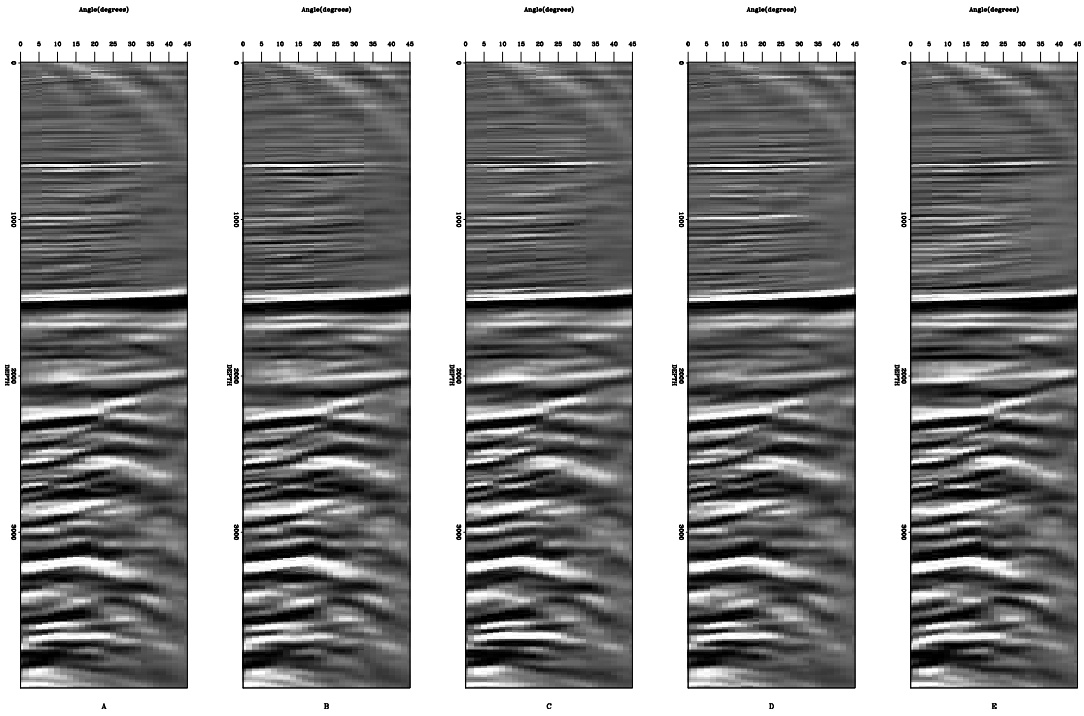


Figure 4: The result of performing an angle domain transform on the gathers shown in Figure 1. bob1/. angle

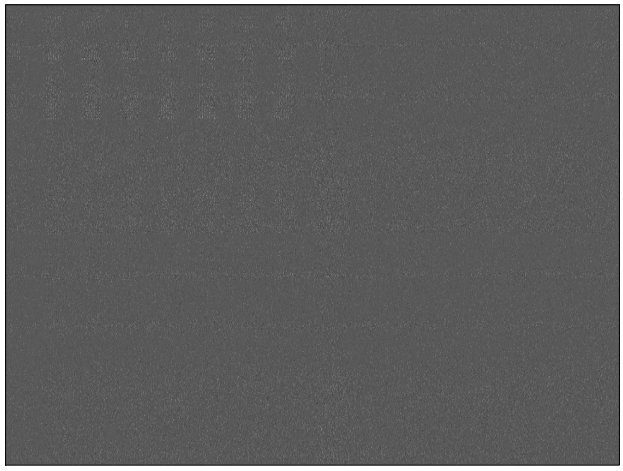


Figure 5: The estimated wavelet domain using the standard IST algorithm. Compared to Figure 3, note how the basic structure is recovered but with significant additional energy. [**ER**] bob1/. standard

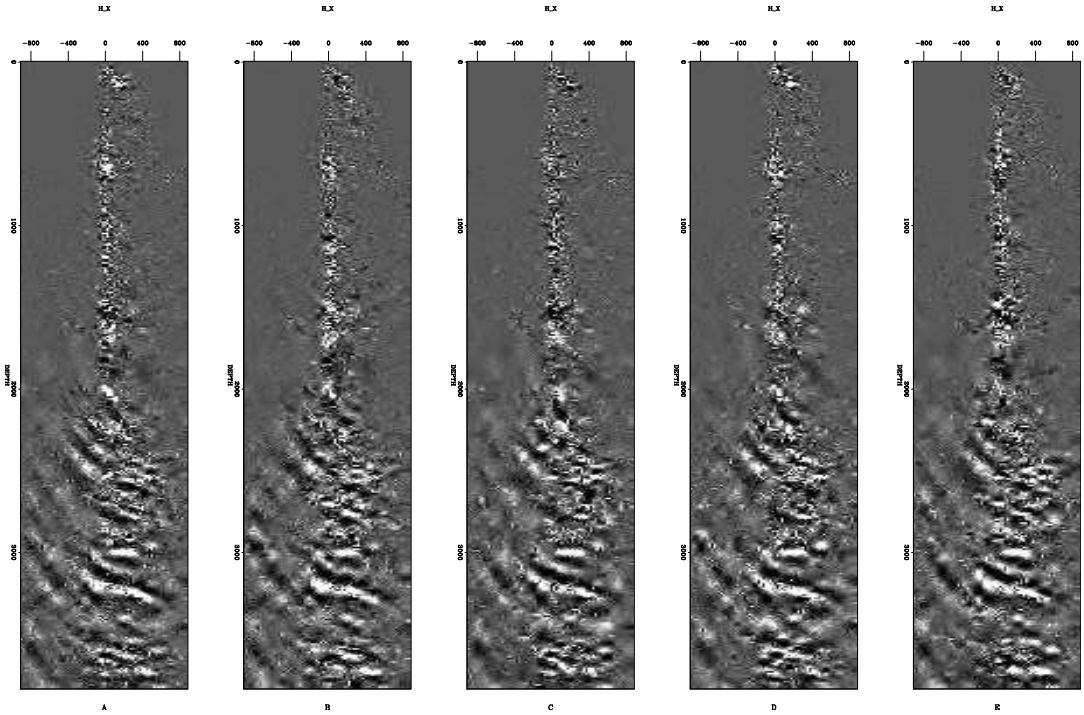


Figure 6: The result of mapping the wavelet domain representation seen in Figure 5 to sub-surface offset and showing the same gathers seen in Figure 1. Note how we have done a nearly perfect job recovering the deeper portion of the image but are not as successful in the upper portion of the image. bob1/. standard-off

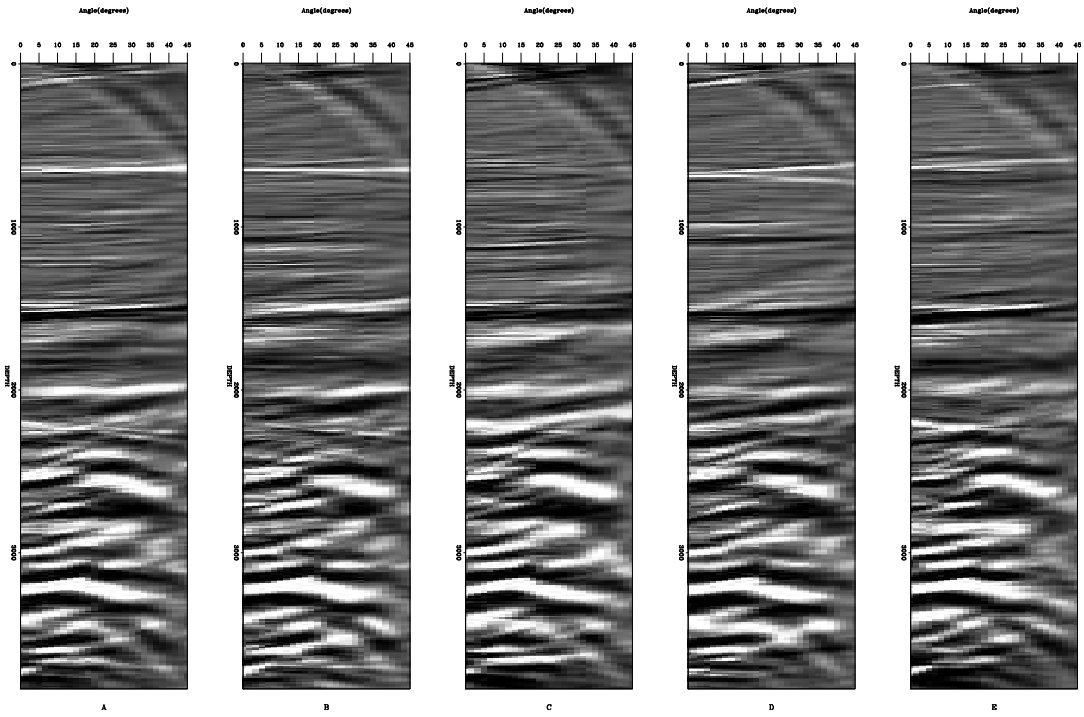


Figure 7: The angle representation of the sub-surface offset gathers shown in Figure 6. Note the errors in the upper portion of the section compared to the full sub-surface based angle gathers shown in Figure 4. [**CR**] bob1/. standard-ang

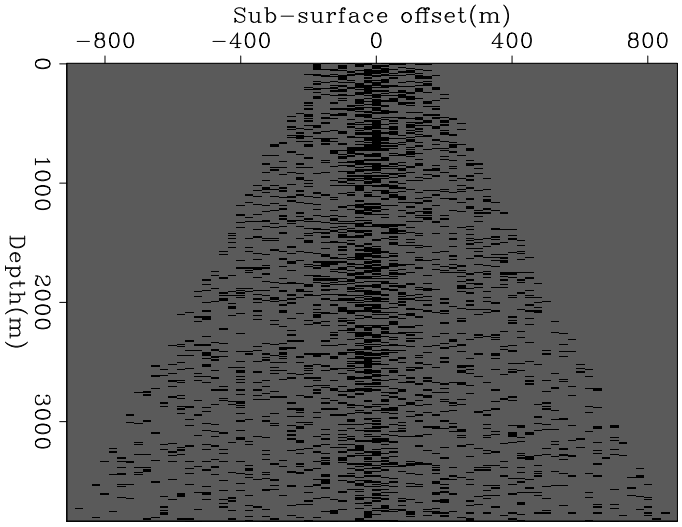


Figure 8: A cone based random sampling pattern which limits the range of possible offsets at shallow depths and tends to concentrate points at smaller offset. [**CR**] bob1/. sampling

Figure 9: The wavelet basis function using the cone shaped sampling shown in Figure 8. Note how more energy is less randomly focused compared to Figure 5 and closer to the wavelet basis shown in Figure 3. [**CR**] bob1/. cone-coefs

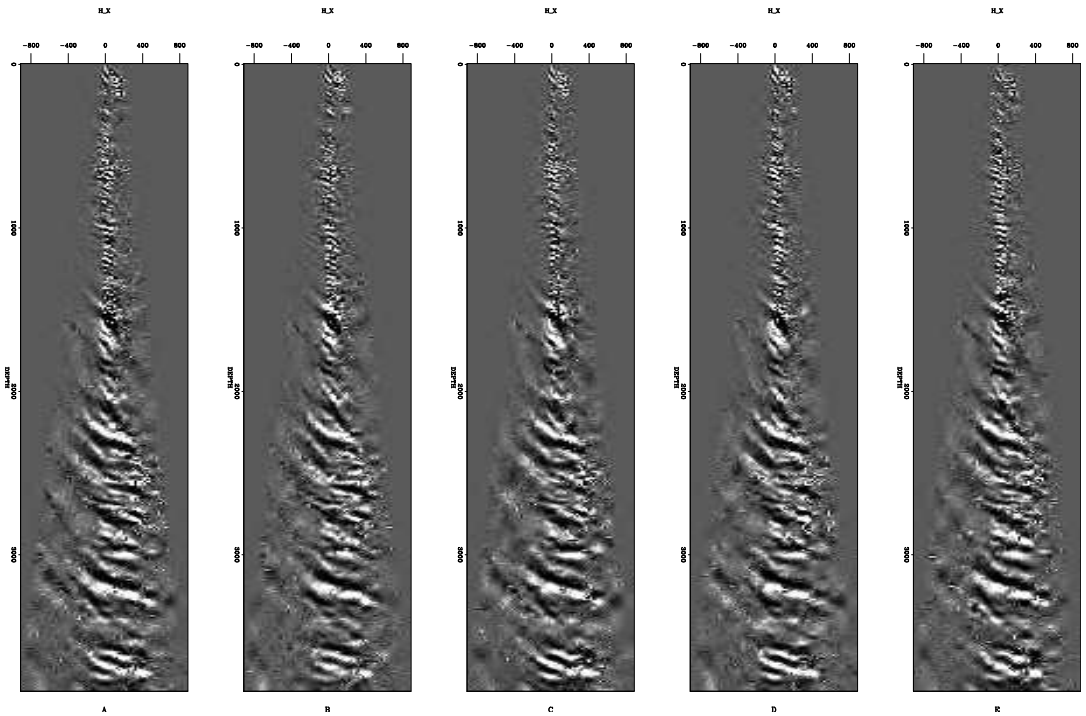


Figure 10: The result of converting to sub-surface offset the wavelet domain seen in Figure 9. Note how we see less random noise than using the standard approach seen in Figure 6. The image at small depth is also improved compared to Figure 6. bob1/. cone-off

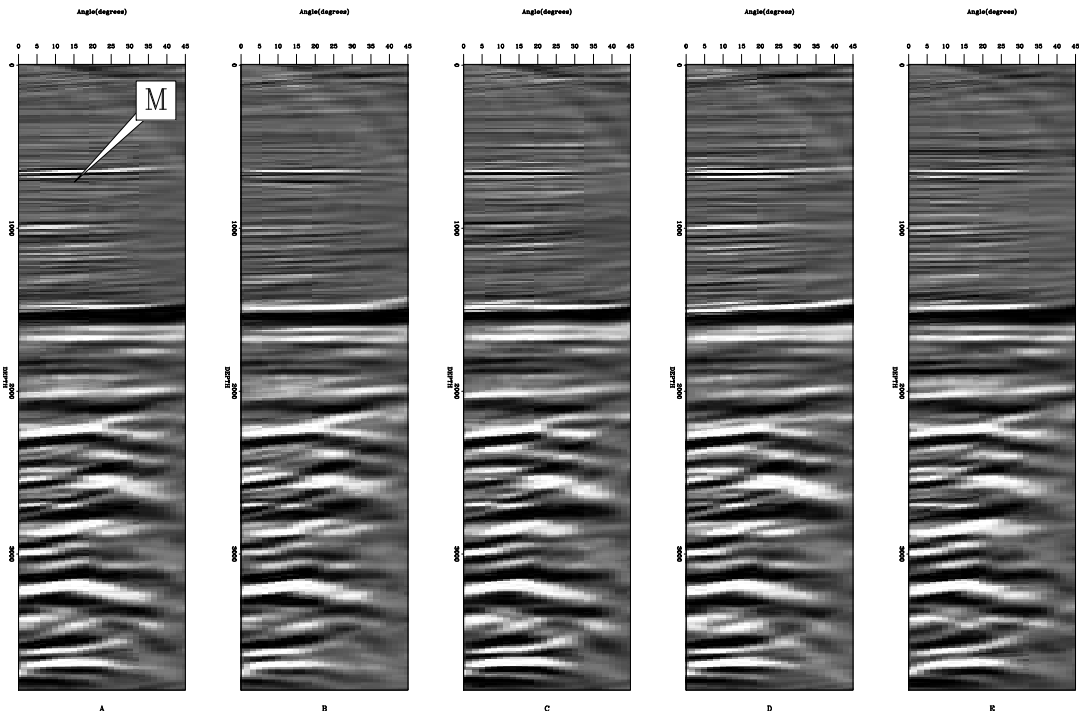


Figure 11: The result of transforming the subsurface offset gathers shown in Figure 10 to the angle domain. Note the improvement particularly at the event marked 'M'. [**CR**] bob1/. cone-ang

strates this concept with a simpler model. In this case a single plane wave has been created in 4-D space panel 'A' and transformed into the multi-dimensional wavelet domain. Panel 'B' and 'C' of Figure 13 show the lowest and highest passed portion of the wavelet domain. Panel 'D' shows a portion of the cumulative distribution function for the lowest, highest, and two intermediate wavelet levels. Note how much more energy is in the lowest levels. The difference in density level suggests that the same thresholding scheme should not be used at every wavelet level.

Instead of using the same thresholding scheme at all levels I modified the thresholding to be level dependent. For each different wavelet level i I then use a different thresholding scheme. For each level I choose a percentile v_i based on the number of samples in the wavelet domain n_i,

$$v_i = 100. - .6\sqrt{3}\frac{n_{\max}}{n_i} \tag{2}$$

where $n_{\max}$ is the maximum size of any level and .6 was determined through trial and error. For each non-linear iteration j, the percentile is decreased,

$$v_{i,j} = max(0., 100. - 1.6 * v_{i,j-1}), \tag{3}$$

where 1.6 was chosen imperically. These percentiles are transformed into thresholding values by taking the $v_{i,j}$ value of the ith wavelet level of $\frac{\mathbf{SWd}}{s}$. Figure 14 shows the same portion of the wavelet domain seen previously. Note how the energy is even more focused than the result seen using only cone sampling in Figure 9. Figure 15 shows the result of converting

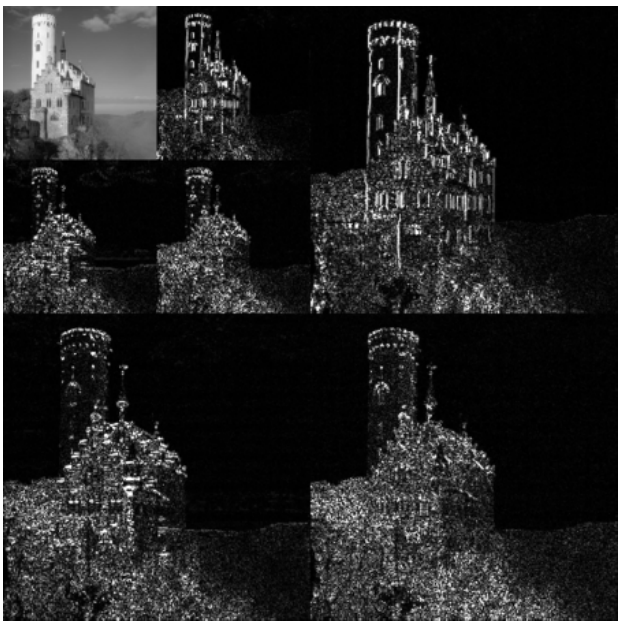


Figure 12: An example of applying a two-dimensional wavelet transform to a grayscale picture of a castle. The bottom-right portion of the image has been low-passed in both x and y; top-left, low passed in x, high passed in y; bottom-left, low passed in y, high passed in. As one moves up and to the left you see the result of higher levels of the wavelet transform. Note how the sparsity decreases as one goes up and to the left. [**NR**] bob1/. wiki

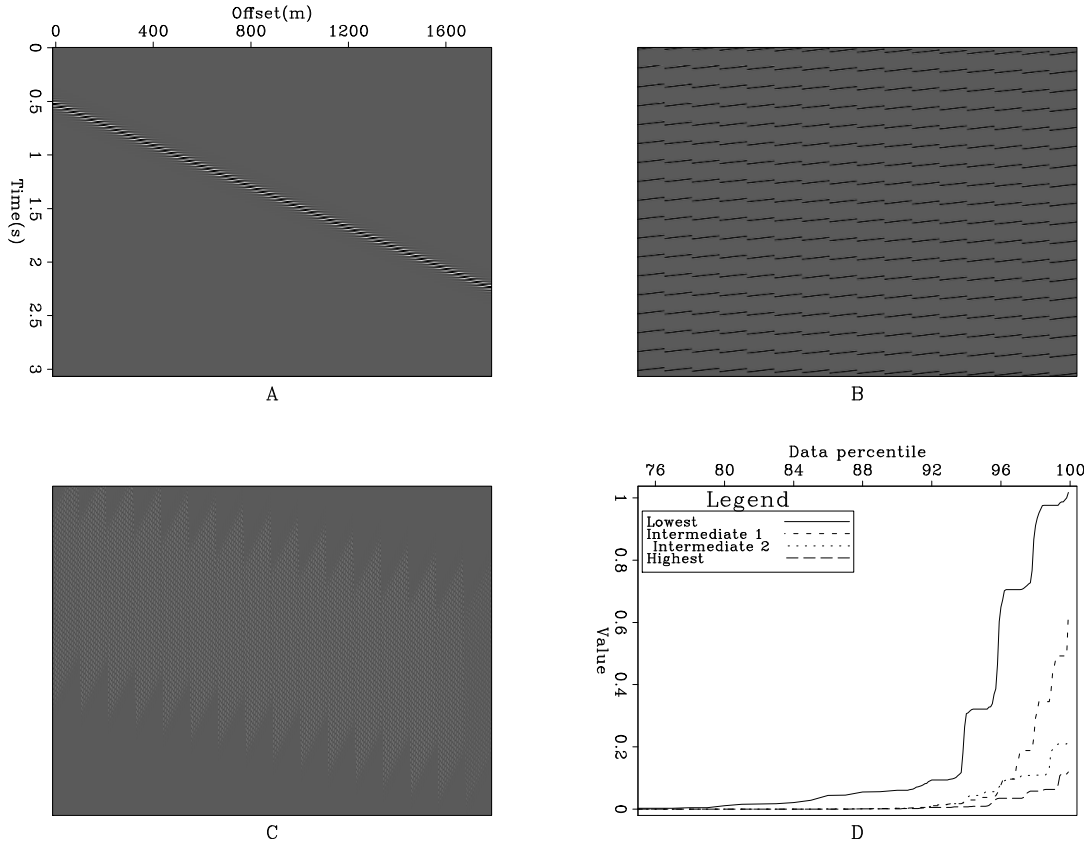


Figure 13: A slice from a plane wave created in 4-D space 'A' and transformed into the multi-dimensional wavelet domain. Panel 'B' and 'C' of Figure 13 show the lowest and highest passed portion of the wavelet domain. Panel 'D' shows a portion of the cumulative distribution function for the lowest, highest, and two intermediate wavelet levels. Note how much more energy is in the lowest levels. [**ER**] bob1/. plane

the wavelet basis function seen in Figure 14 to the subsurface offset domain. Note how we have further reduced random artifacts due to incorrect wavelet basis functions as compared to using simply the cone sampling approach seen in Figure 15. The result angle gathers also show some improvement compared to Figure 16. The amplitudes are closer to the full offset result shown in Figure 1.

stdin

Figure 14: The result of using different thresholding approach at different levels along with cone sampling. Compare how the non-zero coeficients of the basis function are closer to the full sub-surface offset result shown in Figure 3 than the results seen in Figure 5 and 9. [**CR**] bob1/. final-coefs

CONCLUSIONS

I present a method for recovering sparsely sample cross correlation based angle gathers using the IST algorithm to recover a sparse multi-dimensional wavelet domain basis function. Modifying the standard random sampling criteria to limit far sub-surface offset points at small depths improves model recovery. Further improvement is achieved by use of a level dependent thresholding scheme.

ACKNOWLEDGEMENTS

I would like to thank John Washbourne for his SEP sponsor talk last year, which resulted in a second attempt at this method....without a significant bug. I would like to thank Dave Nichols who suggested the plane-wave analogy for the level based thresholding approach.

REFERENCES

Clapp, R. G., 2011, Imaging using compressive sensing: SEP-Report, **143**, 149–158.

———, 2012, Image gather reconstruction using StOMP: SEP-Report, **147**, 127–138.

Crawley, S., N. D. Whitmore, A. Sosa, and M. Jones, 2012, Improving rtm images with angle gathers: 82nd Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1–5, Soc. Expl. Geophys.

Donoho, D. L., 2006, Compressive sensing: Information Theory, IEEE Transactions on, **52**, 1289–1306.

Donoho, D. L., Y. Tsaig, I. Drori, and J.-L. Starck, 2006, Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit: Technical report.

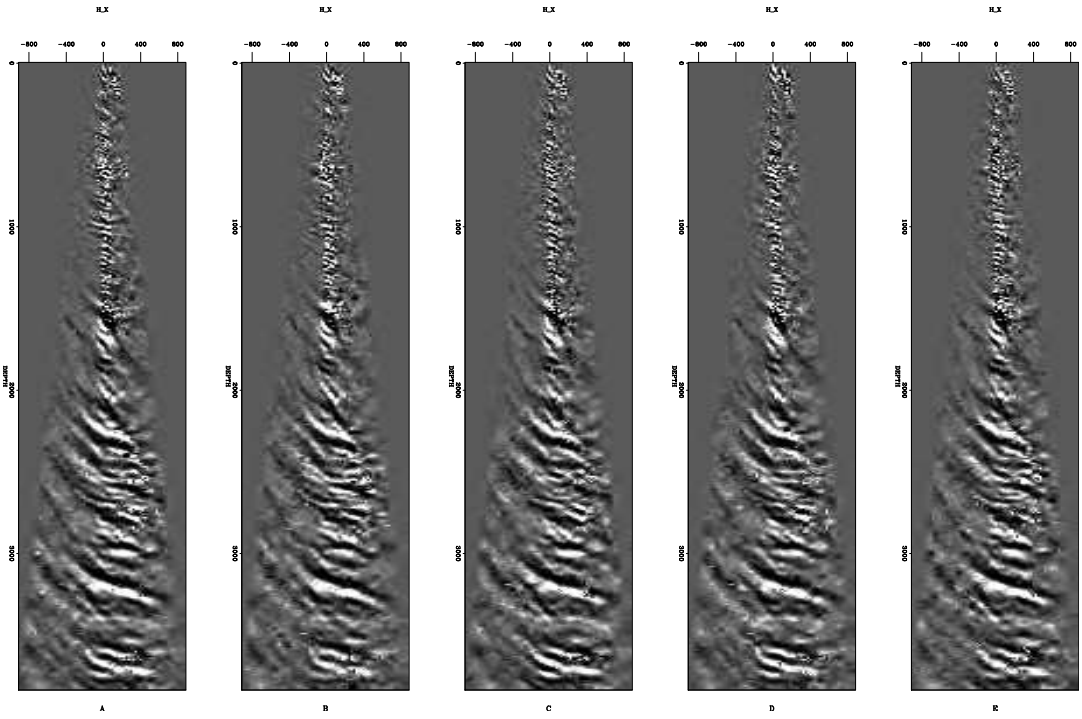


Figure 15: The result of transforming in to subsurface the wavelet basis function seen in Figure 14. Note how there is less random noise throughout the image and more continuity in shallow surface events. [**CR**] bob1/. final-off

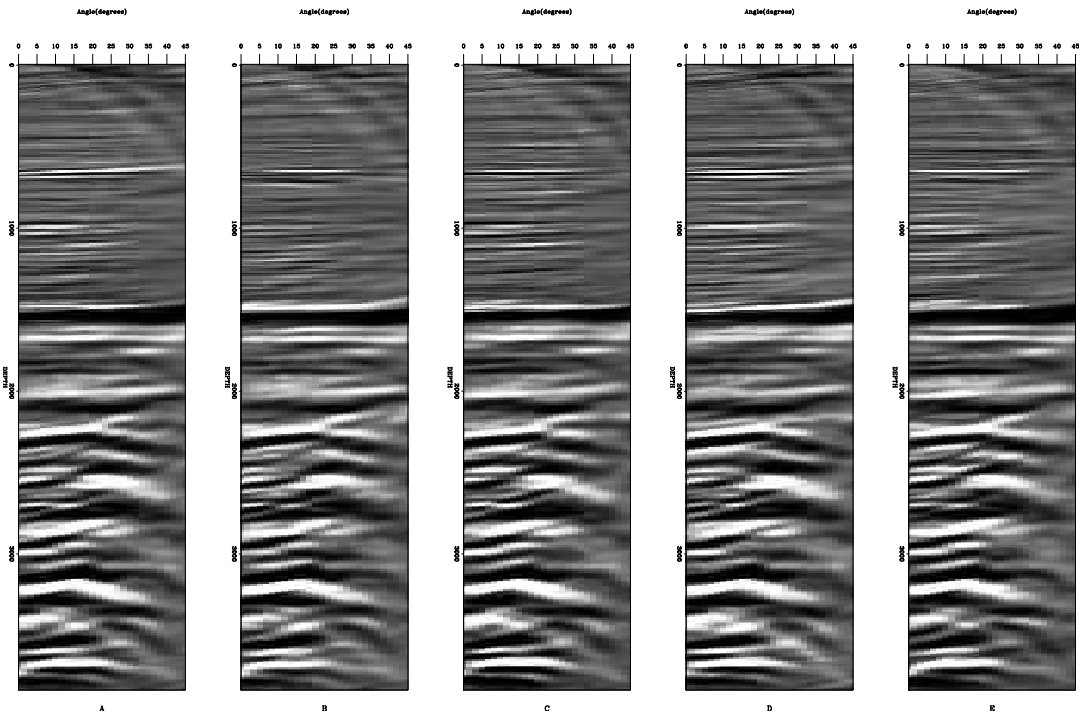


Figure 16: The result of transforming the subsurface offset gathers shown in Figure 15 to the angle domain. The amplitudes are more accurate in the shallow portion of the model than those shown in Figure 7 and 11. [**CR**] bob1/. final-ang

Hennenfent, S. and F. Hermann, 2005, Sparsenessconstrained data continuation with frames: applications to missing traces and aliased signals in 2/3d: 75th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 2162–2165, Soc. Expl. Geophys.

Mallat, S. G. and Z. Zhang, 1993, Matching pursuits with time-frequency dictionaries: IEEE Transactions on Signal Processing, **41**, 3397–3415.

Sava, P. and S. Fomel, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, 5209–5217.

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

Villasenor, J. P., R. A. Ergas, and P. L. Donoho, 1996, Seismic data compression using high-dimensional wavelet transforms: Snowbird, UT, USA, Expanded Abstracts, 396–405, IEEE Computer Society Press.

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

Accelerating residual-moveout-based wave-equation migration velocity analysis with compressed-sensing

Yang Zhang, Biondo Biondi, and Robert Clapp

ABSTRACT

Residual-moveout-based wave-equation migration velocity analysis uses residual-moveout to characterize the kinematic error caused by an inaccurate velocity model and computes velocity updates using wave-equation tomographic operators. However, the per-iteration cost of this method is even more expensive than that of full waveform inversion because of the construction and back-projection of offset/angle-domain common-image gathers. In order to speed up residual-moveout-based migration velocity analysis, we examine its work flow, and propose the following acceleration scheme: 1) by using the compressed-sensing technique, we can very well reconstruct the angle-domain common image gathers from a fraction of the subsurface offset common-image gathers, therefore saving significant computation cost; 2) after we extract the residual-moveout information from the reconstructed angle-domain common image gathers, we reduce the cost of back-projection by synthesizing an approximation of the original image perturbation, which can be back-projected with much lower cost and can yield a velocity gradient with the same behavior.

INTRODUCTION

Wave-equation migration velocity analysis (WEMVA) is a reflection tomography method which uses wave-equation rather than ray-based model to retrieve the velocity model from seismic data(Chavent and Jacewitz, 1995; Biondi and Sava, 1999). The velocity information comes from the seismic data redundancy that arises because each reflector point in the subsurface is illuminated by wave energy from multiple directions. WEMVA exploits such redundancy by forming common-image gathers and then improves the velocity model by enforcing coherence among the common-image gathers. Evaluating the flatness of the subsurface angle-domain common image gathers (ADCIGs) is currently a popular choice when forming WEMVA objective functions (Biondi and Sava, 1999; Clapp and Biondi, 2000; Biondi and Symes, 2004).

The objective function is usually optimized by applying gradient-based algorithms. The computation of the gradient is performed in two steps: 1) computation of a perturbation in the migrated image domain, and 2) back-projection of the image perturbation into the velocity model using the image-space wave-equation tomographic (ISWET) operator (Sava and Biondi, 2004).

One flavor of the WEMVA methods is based on residual-moveout (RMO) (Zhang et al., 2012). It describes the unflatness in the ADCIGs using residual-moveout. Instead of maximizing the image-stack-power objective function directly with respect to the velocity, we link the objective function to the velocity model indirectly through an intermediate moveout

parameter. By focusing on the kinematic errors in the common-image gathers, RMO-based WEMVA is more robust with respect to the cycle-skipping issue and still produces high-quality model updates.

However, for all WEMVA methods, the cost per iteration is much higher than that of full waveform inversion (FWI) because the WEMVA projection operator is more expensive compared to the projection operator in waveform inversion. In WEMVA's gradient projection, besides calculating wavefield propagation, the imaging operator and wave-equation tomographic operator also have to perform cross-correlations and convolutions that involve the subsurface offset-domain common image gathers (ODCIGs) or angle-domain common image gathers (ADCIGs). In contrast, the projection operator in FWI does not deal with the subsurface common-image gathers, therefore is much cheaper than that of WEMVA. Moreover, as the velocity error increases, the size of the ODCIGs has to grow (otherwise the velocity information will be lost), which further increases the cost of WEMVA iteration.

Due to the high cost, a computationally more efficient approach for RMO-based WEMVA is very desirable. There are two computationally significant steps in the work flow: one is the generation of ADCIGs with the current velocity model, which requires one migration with subsurface offset; the other one is the back-projection of the RMO error into the velocity model, which involves applying the image-space tomographic operator with a subsurface-offset image perturbation. For the former step, we make use of the compressed-sensing technique (Clapp, 2012), which can reconstruct the full ADCIG from a randomly subsampled ODCIG; thus the imaging cost (of such an ODCIG) is only a fraction of the original implementation. However, for the latter step, we cannot save computation by back-projecting a subsampled subsurface ODCIG perturbation, because the output gradient will not be consistent with the original gradient. Therefore we propose to approximate the image perturbation with a synthesized one, which has shorter span along the subsurface-offset axis without losing the moveout information. Back-projecting the synthesized image perturbation would cost much less, because the number of subsurface-offsets in the image perturbation is significantly smaller.

The rest of this paper is divided into three parts: first we briefly review the theory of RMO-based WEMVA; then we present the details of the modification method which becomes computationally more efficient; finally we illustrate the modified approach with some synthetic examples.

RMO-BASED WEMVA

For simplicity, we assume two dimensions in our derivation; however, extending the theory to 3-D is conceptually straightforward for this method. We denote the pre-stack angle-domain images as $I(z, \gamma, x)$, where x, z are the depth and horizontal axis, γ is the reflection-aperture angle, and s represents the slowness model.

The details of RMO-based WEMVA can be found in Zhang and Biondi (2011, 2013); here we simply re-state the key steps. We start from the "classical" stack power maximization objective function

$$J(s) = \frac{1}{2} \sum_x \sum_z \left[\int d\gamma \; I(z, \gamma, x; s) \right]^2, \tag{1}$$

where s is the model slowness, and $I(z,\gamma,x;s)$ is the prestack image (ADCIG) migrated with s.

RMO-based WEMVA defines an alternative objective function (maximizing the normalized RMO semblance) with respect to the RMO parameter ρ, which is a function of the slowness s:

$$J_{Sm}(\rho(s)) = \frac{1}{2}\sum_x\sum_z \frac{\int dz_w \big(\int d\gamma I(\gamma, z+z_w+\rho\tan^2\gamma, x; s_0)\big)^2}{\int dz_w \int d\gamma I^2(\gamma, z+z_w+\rho\tan^2\gamma, x; s_0)}, \tag{2}$$

where s_0 is the starting model and z_w is a local averaging window of length L along the depth axis. The gradient given by the objective function (2) is

$$\frac{\partial J_{S_m}}{\partial s} = \frac{\partial J_{S_m}}{\partial \rho}\frac{\partial \rho}{\partial s}, \tag{3}$$

where $\partial J_{S_m}/\partial\rho$ can be easily calculated by taking the derivative along the ρ axis of the semblance panel S_m, and $\partial\rho/\partial s$ can be derived in a way similar to the sensitivity kernel of the finite-frequency travel-time tomography method in Marquering et al. (1998).

The final expression for the gradient calculation is:

$$\frac{\partial J_{S_m}}{\partial s} = -\int dz_w \int d\gamma \sum_{z,x}\{\frac{\partial I(z+z_w,\gamma,x;s)}{\partial s}(F_{11}\tan^2\gamma + F_{12})\dot{I}(z+z_w,\gamma,x;s_0)\frac{\partial J_{Sm}}{\partial\rho}(z,x)\}, \tag{4}$$

where F_{11} and F_{12} are constants computed from the initial ADCIGs, the detailed formula for which can be found in Zhang and Biondi (2013).

ACCELERATED RMO WEMVA APPROACH

The dominant cost of RMO WEMVA is the ADCIG reconstruction and the model gradient calculation (shown in eq. (4)). The following two sections describe the acceleration we propose.

Efficient reconstruction of ADCIGs

As seen in eqs. (4) and (2), the cost of computing $\partial J_{S_m}/\partial\rho$ is primarily the cost of constructing the full ADCIGs $I(z,\gamma,x;s_0)$. In our practice, this is done by constructing the subsurface-offset image $I(z,h_x,x;s_0)$ using the cross-correlation imaging condition and then transforming the offset image to the angle-domainusing the method described in Sava and Fomel (2003). Computing the subsurface-offset images (i.e. ODCIGs) is significantly more expensive than the zero-subsurface-offset image, because wavefield cross-correlation has to be performed at all subsurface-offset locations. As we move to 3-D, the I/O could also become another bottleneck because of the large size of the 5-D subsurface-offset image.

There is a significant amount of literature on compressing seismic data, recently Clapp (2012) found that the seismic images such as ADCIG and ODCIG are also highly compressible. To utilize the compressibility of seismic images, Clapp (2012) proposed to use the compressed-sensing method (Candes and Donoho, 1999; Donoho, 2006) to reduce the imaging cost. Compressed-sensing theory predicts that given certain conditions are satisfied,

we can almost perfectly recover the full ADCIGs from a randomly and heavily subsampled ODCIG by solving the following optimization problem:

$$\begin{aligned} \mathbf{0} &\approx ||\mathbf{M_d}(\mathbf{d} - \mathbf{RWm})||_{\mathbf{2}}^{\mathbf{2}}, \\ \mathbf{0} &\approx ||\mathbf{m}||_{\mathbf{1}}, \end{aligned} \tag{5}$$

where $\mathbf{M_d}$ represents the sub-sampling mask for the ODCIG $\mathbf{d}$. Thus we need to compute the ODCIGs only at those locations specified by $\mathbf{M_d}$. $\mathbf{R}$ is the angle-to-offset transform operator, $\mathbf{W}$ is the multi-dimensional wavelet transform operator, $\mathbf{m}$ is the wavelet-domain coefficient of the ADCIG, and $\mathbf{m}$ is assumed to be sparse. The sparsity-promoting constraint for $\mathbf{m}$ is to minimize the ℓ_1-norm of $\mathbf{m}$, as expressed in eq. (5).

Clapp (2012) uses an ℓ_1 inversion scheme called the Stage-wise Orthogonal Matching Pursuit (StOMP) (Donoho et al., 2006), which essentially approximates the original problem by solving multiple least-square data-fitting problems. The operators $\mathbf{R}$ and $\mathbf{W}$ are much cheaper than the imaging operator itself, so this reconstruction scheme is more efficient than calculating the ADCIG directly from the full ODCIG.

Figure 1 shows a reconstruction example using the Marmousi model; here the sub-sampling ratio of the ODCIGs is 20%. Figure 2(a) shows the reconstructed ADCIG using the sub-sampled ODCIGs in figure 1. For comparison, figure 2(b) shows the ADCIG transformed from fully sampled ODCIGs. The comparison shows that the compressed-sensing-based reconstruction is satisfactory. Although the two are not exactly the same, the important RMO information has been well preserved on the reconstructed ADCIG.

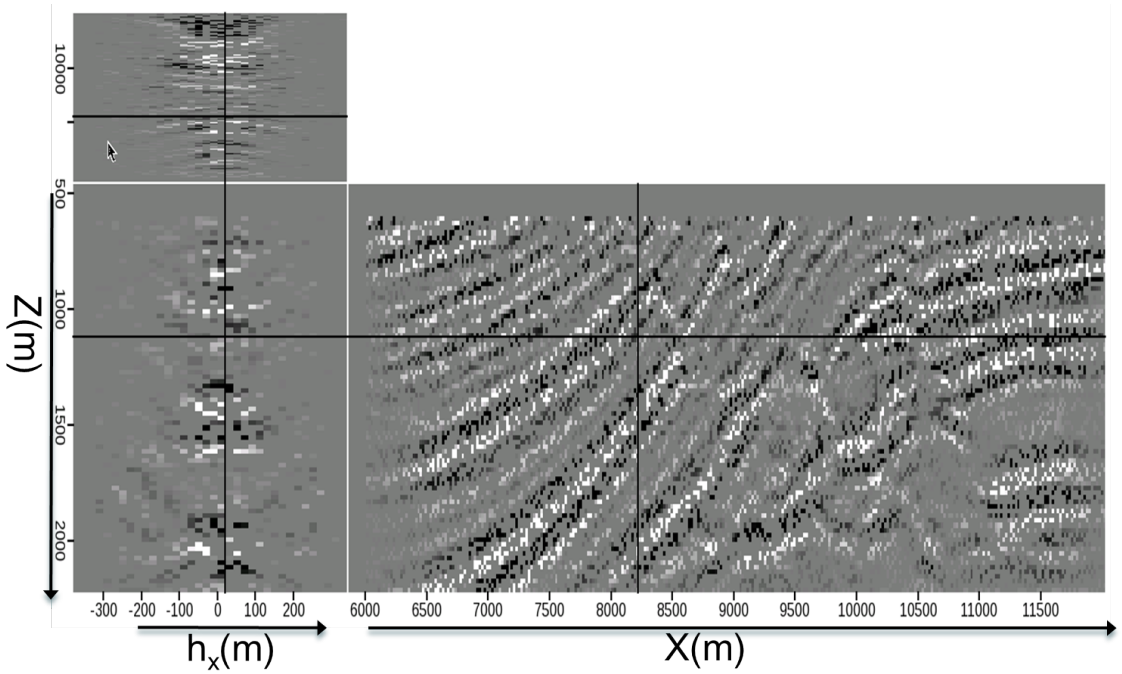


Figure 1: A Marmousi subsurface offset image after random sub-sampling. The sub-sampling ratio is 20%. [**ER**] yang1/. fig1

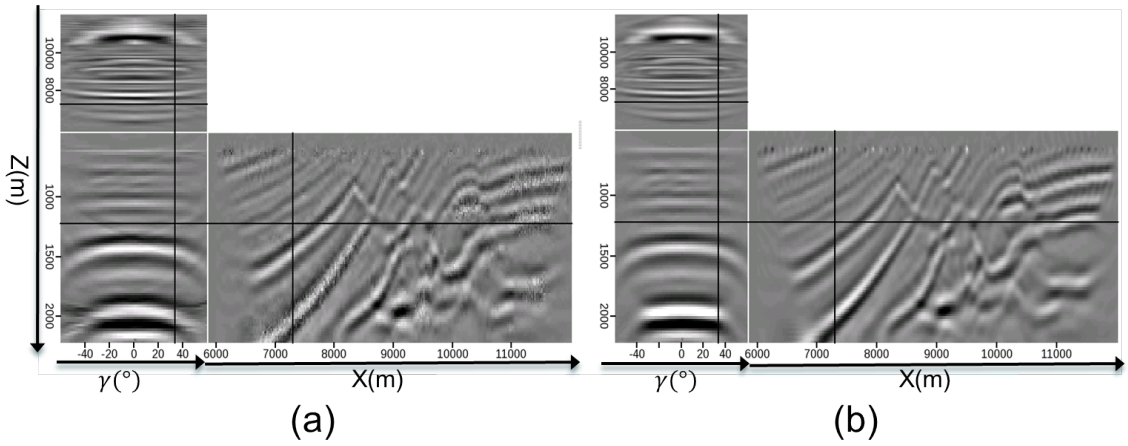


Figure 2: (a): The reconstructed ADCIGs using the sub-sampled ODCIG in figure 1. (b): The ADCIGs transformed from the fully sampled ODCIG. [**ER**] yang1/. fig2

Back-projection using synthesized image perturbation

As shown in the gradient calculation in formula (4), the other computationally expensive step is the back-projection of the image perturbation,

$$\Delta I = (F_{11}\tan^2\gamma + F_{12})\dot{I}(z+z_w,\gamma,x;s_0)\frac{\partial J_{Sm}}{\partial \rho}(z,x)$$

using the image-space wave-equation tomographic operator,

$$\mathbf{T} = \partial I(z+z_w,\gamma,x;s)/\partial s.$$

From the offset-to-angle common image gathers transform described by Sava and Fomel (2003), we know that the unflatness in the angle domain corresponds to unfocused energy in the subsurface-offset domain. The angle-domain image perturbation ΔI is the first-order z derivative of the initial ADCIG $I(s_0)$, modulated by $f(\gamma) = (F_{11}\tan^2\gamma + F_{12})\frac{\partial J_{Sm}}{\partial \rho}(z,x)$. Therefore the gather shape (curvature) in ΔI is the same as in initial ADCIG $I(s_0)$. Then the subsurface-offset domain of ΔI will have to use the same offset span as the initial ODCIG. When the velocity error is large, the offset span can be large, which increases the amount of computation for the tomographic operator.

The back-projection is expensive, because ΔI takes many subsurface offset locations to represent. Realizing that the RMO information (which determines the velocity update direction) is already embedded in the modulate function $f(\gamma)$, we can form a different image perturbation as follows:

$$\Delta\tilde{I}(z+z_w,\gamma,x) = f(\gamma)\dot{\tilde{I}}(z+z_w,\gamma,x;s_0),$$

in which $\tilde{I}(\gamma;s_0)$ is the angle-domain common image gathers for the zero-subsurface-offset image $I(h=0;s_0)$, and the dot denotes the derivative respect to depth z. The advantage of defining such $\tilde{I}(\gamma)$ is that $\tilde{I}(\gamma)$ would be flat among angles, which leads to $\Delta\tilde{I}(\gamma)$ also being

flat. Then the offset-domain $\Delta\tilde{I}(h)$ will be much more focused around $h = 0$, and we can use a shorter subsurface-offset axis h for $\Delta\tilde{I}(h)$. Therefore the cost of applying the tomographic operator on $\Delta\tilde{I}(h)$ will be much smaller, because $\Delta\tilde{I}(h)$ needs fewer subsurface-offset points than $\Delta I(h)$.

Besides changing $I(h)$ and $\Delta I(h)$, we need to replace the observed data d used in the tomographic operator with $\tilde{d} = \mathbf{L}I(h = 0)$ correspondingly, where $\mathbf{L}$ is the Born modeling operator. By doing this, the synthesized data $\tilde{d}$ becomes kinematically consistent with the synthesized image perturbation $\Delta\tilde{I}$.

INVERSION EXAMPLE

Marmousi model

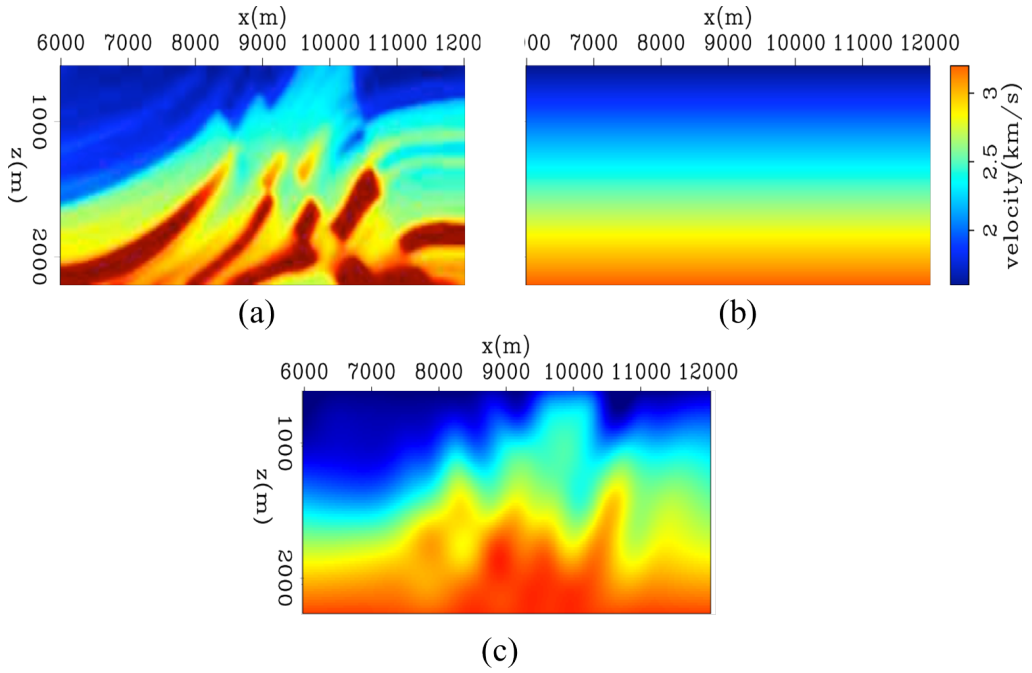


Figure 3: The slightly smoothed true velocity model of Marmousi (a), the starting velocity model (b), and the inverted velocity models using our method after 20 iterations (c). [**CR**] yang1/. fig3

With the two modifications incorporated in our RMO WEMVA work flow, we test the new approach on the smoothed Marmousi Model. The model is 6 km in x and 1.6 km in z. The spatial sampling is 20 m. The survey geometry follows the land acquisition pattern with receivers at every surface location on the top and we simulate 51 shots in total, covering the whole lateral span on the top with a spacing of 120 m. We model 64 frequencies in total using the one-way wave-equation, ranging from 5 Hz to 40 Hz.

For each iteration, we compute only 20% of the locations of the migrated ODCIGs and reconstruct the full ADCIGs using the StOMP algorithm, which could save approximately 80% of the imaging cost. After the RMO information is extracted, we back-project the

synthesized image perturbation, which uses only 9 subsurface offsets rather than 33 points needed for original image perturbation, saving about 75% of the tomographic operator cost. We ran 20 nonlinear iterations for the inversion.

Figure 3(a) shows the true model (in velocity) and (b) shows the starting model, which has a vertical gradient increasing from 1600 m/s and 3200 m/s. Figure 3(c) shows the inverted velocity model using our computationally efficient approach. The result shows good convergence to the true model.

BP synthetic model

We also test this new approach on a portion of the BP synthetic model. The model is extracted from the upper-left part of the original BP model, with a size of 35 km in x and 8 km in z. The spatial sampling of the model is 25 m in x and 12.5 m in z. We synthesize several flat reflectors to simulate the seismic data. We use marine streamer geometry to simulate 200 shots on the top of the model, with receiver streamer towing from left to right. The shots start from $x = -5$ km, and the spacing is 125 m. There are 401 receivers on the 10 km long streamer with a spacing of 25 m. We model 281 frequencies in total using the one-way wave-equation, ranging from 5 Hz to 40 Hz.

Figure 4(a) shows the true model (in velocity) and (b) shows the starting model, in which we assume the complex salt overburden has been well resolved, and the sediments velocity below the salt linearly increase from 3.4 km/s to 4.4 km/s. Because we assume the velocity above the salt is accurate, we apply a mask during the inversion so that only the subsalt region of the model will be updated. Figure 4(c) shows the inverted velocity model after 40 iterations using our computationally efficient approach. Again, the result shows good convergence to the true model. Figure 5 shows the migrated images using the velocity models in fig 4. They are plotted using the same clip, so that we can easily observe the improvements of the reflector coherence by the inversion. Figure 6 show the improvement of ADCIGs flatness from the starting velocity model to the inverted velocity model.

CONCLUSION

In order to reduce the computational cost of residual-moveout based wave-equation migration velocity analysis, we propose two modifications to the workflow: first we use a compressed-sensing-based technique to reconstruct angle-domain common-image gathers without computing all of the subsurface-offset gathers; second, we synthesize an image perturbation that has more focused energy near zero subsurface offset, and back-project that image perturbation into the slowness model while keeping the model updates similar to the original case. We test this computationally efficient approach using the 2-D Marmousi model and the BP model. The examples show that adopting these modifications in RMO-based WEMVA does not degrade its inversion result; while it can increase inversion speed by five-fold. Although the increased speed we achieve currently is not overwhelming, in the 3-D case, the compressibility of the seismic images increases dramatically with the growth of dimensions, so we expect the speed up ratio to increase by additional one or two orders of magnitude.

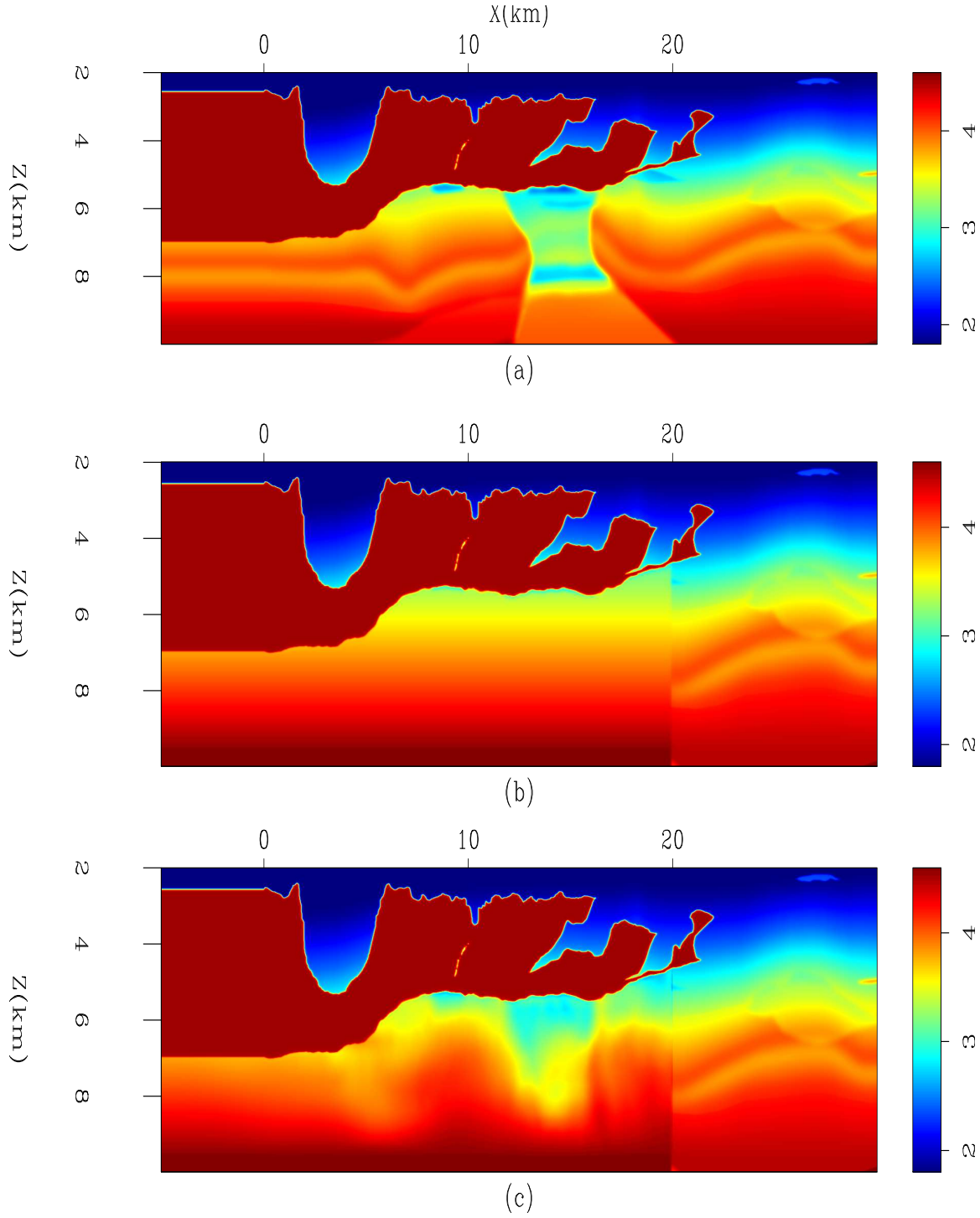


Figure 4: The true velocity model (a) of the upper-left part of the BP model, the starting velocity model (b), and the inverted velocity model (c) using our method after 40 iterations. [**CR**] yang1/. bpVels

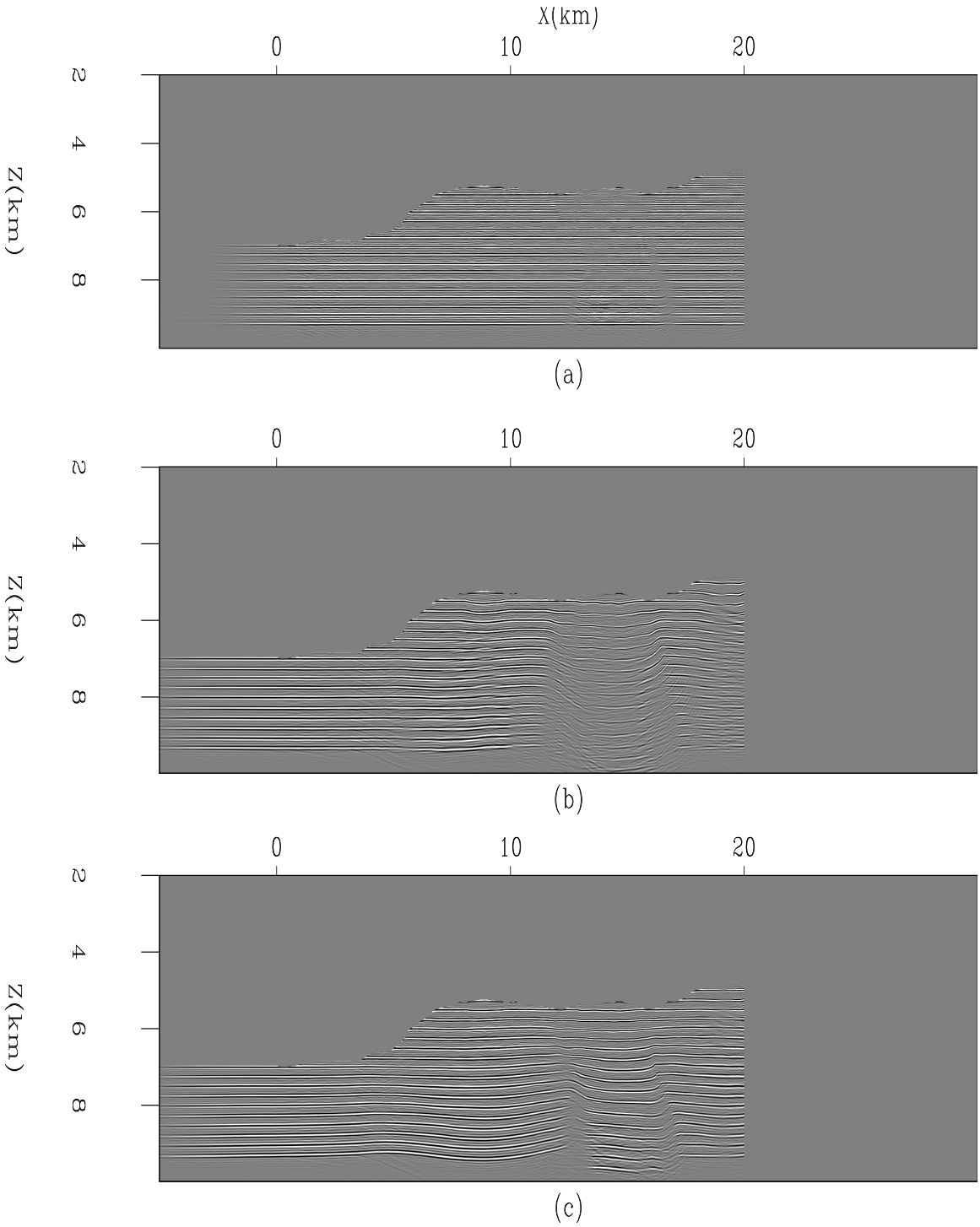


Figure 5: The migrated images using the velocity models shown in fig 4. [**CR**] yang1/. bpImgs

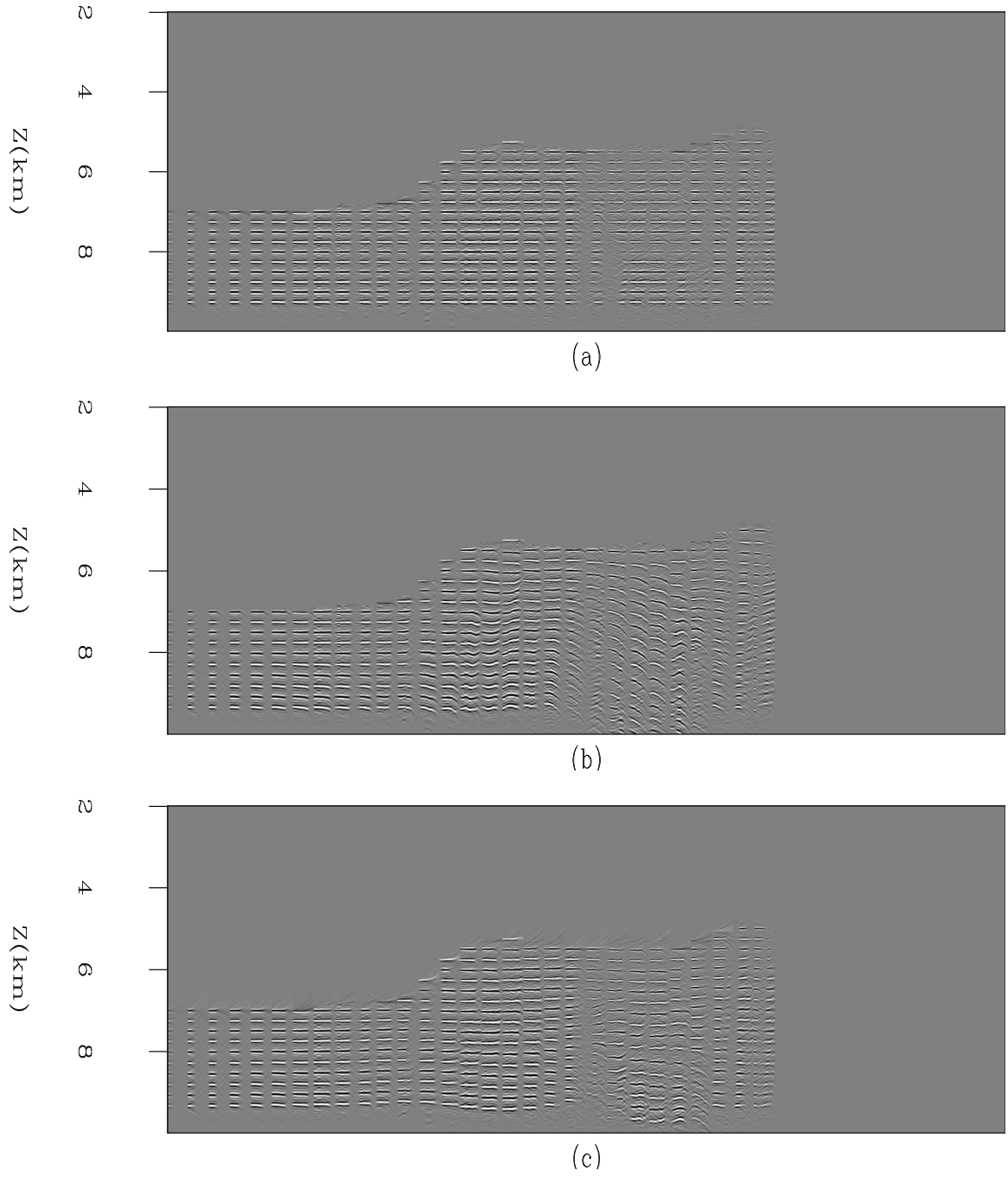


Figure 6: The angle-domain common image gathers migrated using the true model (a), using the starting velocity model (b), and using the inverted velocity model (c). We calculate the angles from 0° to 52°, with a sampling of 2°. The flatness of ADCIGs improves significantly from the initial model. [**CR**] yang1/. bpAdcigs

ACKNOWLEDGEMENT

We thank the sponsors of the Stanford Exploration Project for their financial support. The first author thanks Dave Nichols for the helpful discussions and suggestions on seismic tomography.

REFERENCES

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

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

Candes, E. J. and D. L. Donoho, 1999, Curvelets — a surprisingly effective nonadaptive representation for objects with edges: Curves and Surfaces, 105–120, Vanderbilt University Press.

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

Clapp, R. G., 2012, Image gather reconstruction using StOMP: SEP-Report, **147**, 127–138.

Clapp, R. G. and B. Biondi, 2000, Tau domain migration velocity analysis using angle crp gathers and geologic constraints: SEG Technical Program Expanded Abstracts, **19**, 926–929.

Donoho, D. L., 2006, Compressed sensing: IEEE Transactions on Information Theory, **52**, 1289–1306.

Donoho, D. L., Y. Tsaig, I. Drori, and J.-L. Starck, 2006, Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit: Technical report.

Marquering, H., G. Nolet, and F. Dahlen, 1998, Three-dimensional waveform sensitivity kernels: Geophysical Journal International, **132**, 521–534.

Sava, P. and B. Biondi, 2004, Wave-equation migration velocity analysis. I. Theory: Geophysical Prospecting, **52**, 593–606.

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

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

———, 2013, Moveout-based wave-equation migration velocity analysis: GEOPHYSICS, **78**, U31–U39.

Zhang, Y., B. Biondi, and Y. Tang, 2012, Residual moveout-based wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts, **31**.

Salt delineation via interpreter-guided 3D seismic image segmentation

Adam Halpert

ABSTRACT

Although it is a crucial component of seismic velocity model building, salt delineation is often a major bottleneck in the interpretation workflow. Automatic methods like image segmentation can help alleviate this bottleneck, but issues with accuracy and efficiency can hinder their effectiveness. However, a new graph-based segmentation algorithm can, after modifications to account for the unique nature of seismic data, quickly and accurately delineate salt bodies on 3D seismic images. In areas where salt boundaries are poorly imaged, limited manual interpretations can be used to guide the automatic segmentation, allowing for interpreter insight to be combined with modern computational capabilities. A successful 3D field data example demonstrates that this method could become an important tool for interactive interpretations tasks.

INTRODUCTION

Salt interpretation is a vital component of seismic imaging projects in many of the world's resource-rich basins. The sharp contrast between seismic velocities within salt structures and those in the surrounding sediments means that inaccurate interpretation of these salt-sediment boundaries can lead to severe degredation of images sub-salt; this is of particular concern since sub-salt reservoirs are often the targets for modern exploration. Unfortunately, salt interpretation is not only critical, but often time-consuming and human-intensive as well. For large 3D surveys, manual salt-picking can consume significant resources during model-building workflows that stretch for weeks or months. This can be exacerbated by iterative sediment- and salt-flooding techniques that require several rounds of salt interpretation (Mosher et al., 2007). The semi-automatic image segmentation method we present here aims to help alleviate this bottleneck, while maintaining the accuracy necessary for successful model building and imaging.

While image segmentation is most often associated with fields such as medical imaging and photo processing, several efforts have been made to apply automatic segmentation concepts to seismic images. A variety of approaches has been tried, including pixel-by-pixel classifier methods using fuzzy math (Valet et al., 2001) or texture attributes (Berthelot et al., 2012). These methods can incorporate interpreter input by "training" the algorithm through the use of if-then guidelines or training images. Another category of methods that has proven popular for seismic images is known as graph-based image segmentation. In this method, each pixel in a seismic image is treated as a node or vertex in a graph; then edges are constructed between specific pixels and weighted according to some property. Image segments are created by partitioning the graph (for example, a partition may represent a salt boundary). An advantage of graph-based segmentation is that it provides a globally

optimum solution to the segmentation problem. This compares favorably with automatic interpretation tools such as horizon trackers that tend to get "lost" if a boundary becomes chaotic or discontinuous.

The first graph partitioning seismic image segmentation algorithms were adapted from the eigenvector-based Normalized Cuts Image Segmentation (NCIS) method (Shi and Malik, 2000). One of the first applications was for atomic meshing of seismic images (Hale and Emanuel, 2002, 2003), followed by efforts to track salt boundaries (Lomask et al., 2007; Lomask, 2007). The method was effective, but faced limitations - most notably computational. The NCIS algorithm requires the calculation of eigenvectors for an edge weight matrix of size n^2, where n is the number of pixels in the image; this matrix quickly grows very large, especially for 3D surveys. Calculation of eigenvectors for such a large matrix is an extremely computationally demanding task. Despite modifications to limit the computational domain of this method, it remains infeasible for very large 3D images. The method we present relies instead on the graph-based technique of Felzenszwalb and Huttenlocher (2004), which was designed with efficiency as a primary consideration. With modifications to account for the unique properties of seismic images, this method can accurately segment 3D images at a fraction of the expense required for the eigenvector approach. Furthermore, valuable interpreter insight can be incorporated in the form of limited 2D interpretations, which are then used to guide an automatic 3D segmentation.

SEGMENTATION METHOD

The algorithm developed by Felzenszwalb and Huttenlocher (2004) is designed such that its computational requirements scale at approximately $n \log n$ (where n is the number of pixels in the image), a significant cost savings over other graph-based approaches scaling at n^2. The algorithm relies heavily on the concept of the "Minimum Spanning Tree" (Zahn, 1971) of a graph. A graph's edges are weighted using a measure of dissimilarity between vertex pairs; a connected graph is defined as one in which all such edges are assigned a weight value. If a spanning tree is a connected graph which connects all vertices of the graph without forming a circuit, the minimum spanning tree (MST) of a graph is the spanning tree requiring the minimum sum of edge weights. By sorting an image's edge weights in increasing order, the MST concept allows Felzenszwalb and Huttenlocher (2004) to develop what they term a "pairwise region comparison" (PRC) predicate in order to determine whether two regions should be considered separate segments of the graph, or merged into a single region. Briefly, the method examines the relationship between edges connecting pixels within a region, and edges connecting pixels across a putative boundary. If these two groups are similar to within a threshold, the two regions are merged. This process allows even highly heterogeneous regions to be segmented as a single component of an image – an important capability when handling noisy images.

Adaptation for seismic images

Seismic images are distinct in many ways from more conventional photographs and medical images for which this and most image segmentation algorithms are designed. The effects of this fact can be seen in Figure 1(b), the result of using the un-altered PRC algorithm to segment a 2D image from the Gulf of Mexico (Figure 1(a)). In this and all subsequent

depictions of segmentation results, the interpreted segments are assigned a random color and overlaid on the image for reference. An initial hurdle is that seismic data are a function of amplitude and phase, presenting a challenge for any segmentation algorithm; in Figure 1(b), the algorithm interprets the area around the salt boundary as several regions, instead of an interface between just two regions. This problem can be mitigated by using the amplitude of the seismic signal's envelope as the input for segmentation, rather than the raw image in Figure 1(a). A second concern is that regions such as salt bodies in a seismic image are most easily delineated by their boundaries, rather than, for example, color attributes used to segment photographs. Therefore, modifications to the algorithm's procedure for both constructing the graph and weighting its edges are required to obtain acceptable segmentation results for seismic images.

The original implementation of the pairwise region comparison (PRC) algorithm creates a graph with eight edges per node (pixel). This graph is constructed by looping over every pixel, and performing four calculations at each vertex. The left side of Figure 2 illustrates this process – if the "active" pixel is the one in red, edges are built to each of the blue pixels. Since every pixel in the image undergoes this process, a form of reciprocity allows for each pixel to be connected to its eight immediate neighbors via edges. While this process allows for the extreme efficiency of the algorithm, the unique and often irregular nature of seismic data does not lend itself well to segmentations using so few edges per vertex or pixel. Instead, a larger "stencil," such as the one on the right in Figure 2, has been implemented. The length of the stencil's arms is a user-defined parameter which may be adjusted based on data quality; larger stencils should be used for noiser data, but the trade-off is increased computational complexity. Increasing the size of the stencil allows for many more comparisons per pixel, and a far greater amount of information goes into the segmentation algorithm. While this approach obviously decreases the efficiency of the algorithm, the increased accuracy seen in the final results appears to make it a worthwhile trade-off. Even with the increased number of edges per node, this algorithm is still far less computationally intensive than the NCIS algorithm from Shi and Malik (2000).

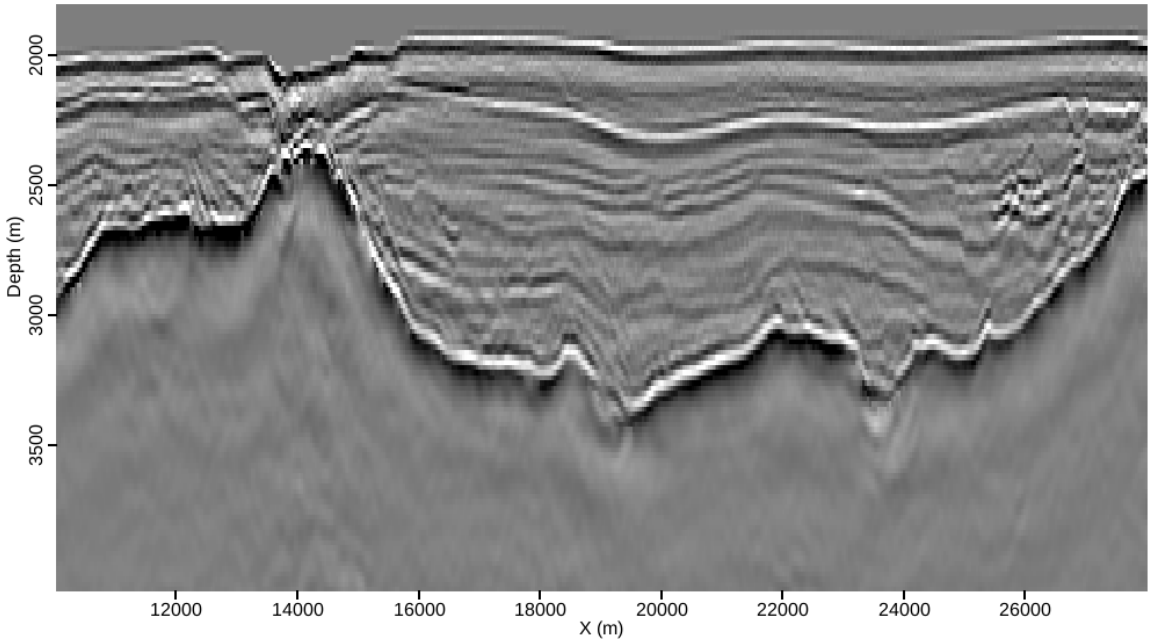
Finally, the edges constructed using the modified stencil in Figure 2 must be weighted in a manner that treats a boundary *between* two vertices as more convincing evidence for the existence of two regions than simply a difference in intensity at the two pixels themselves. When determining the weight for an edge with an endpoint along one arm of the stencil in Figure 2, we use the largest intensity value of any pixel between the two endpoints. For example, a high intensity value along one arm of the stencil would suggest that that particular arm intersects a boundary. Figure 2 illustrates the logic behind this process.

Once we have selected the intensity value to use for determining the edge weight, the weight value is cacluated using an exponential function:

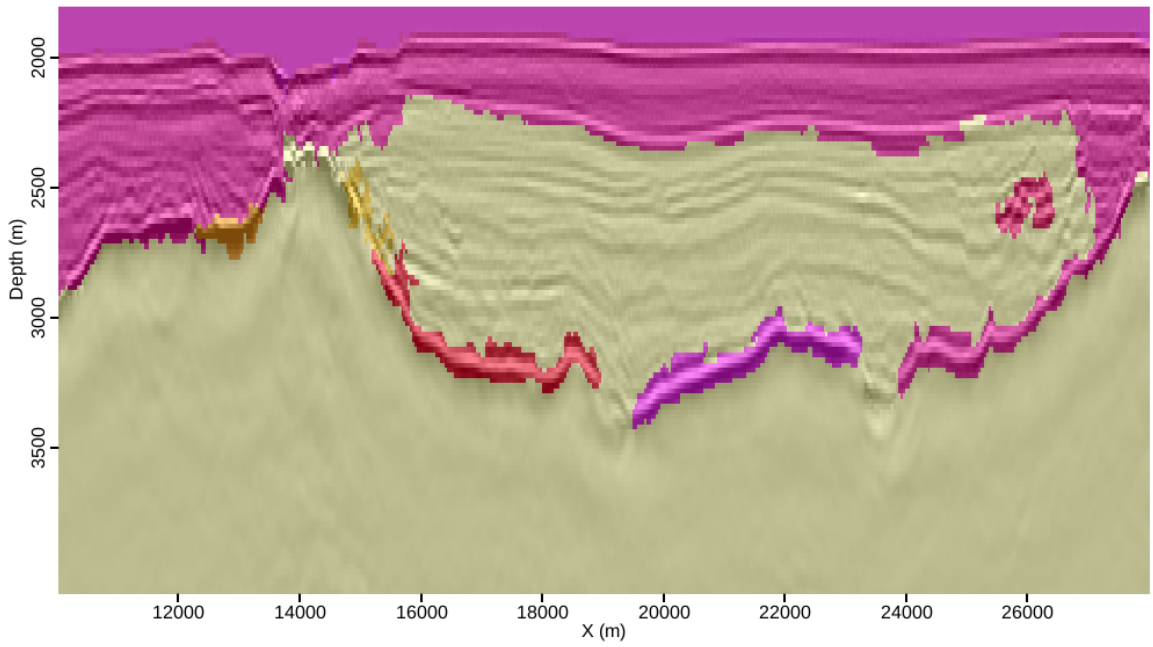
$$w_{ij} = \exp((\max I(\mathbf{p}_{ij}))^2) \exp(d_{ij}), \tag{1}$$

where $\mathbf{p}_{ij}$ is the vector of all pixels between i and j and d_{ij} is simply the Euclidean distance (in samples) between the two pixels. The distance-weighting d term accounts for the fact that the edges in the graph can now be much longer than with the adjacent-pixels-only approach taken in the original implementation.

Once each of the edges is assigned a weight, the segmentation of the image can proceed as described in Felzenszwalb and Huttenlocher (2004). In summary, the process begins



(a)



(b)

Figure 1: (a) A 2D field seismic image, and (b) its corresponding segmentation using the original algorithm from Felzenszwalb and Huttenlocher (2004). [**ER**] adam2/. uno-img,uno-origseg2

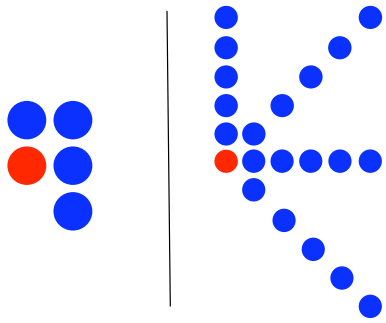


Figure 2: Stencils used for comparing pixel values and assigning edge weights for the graph. At left, the five-point stencil (8 edges per pixel) used in the original implementation from Felzenszwalb and Huttenlocher (2004); at right, a modified 21-point stencil (40 edges per pixel) used for the seismic images. For 3D images, additional stencil arms extend into the third dimension; the length of the stencil arms is a user-defined parameter. [**NR**] adam2/. stencils

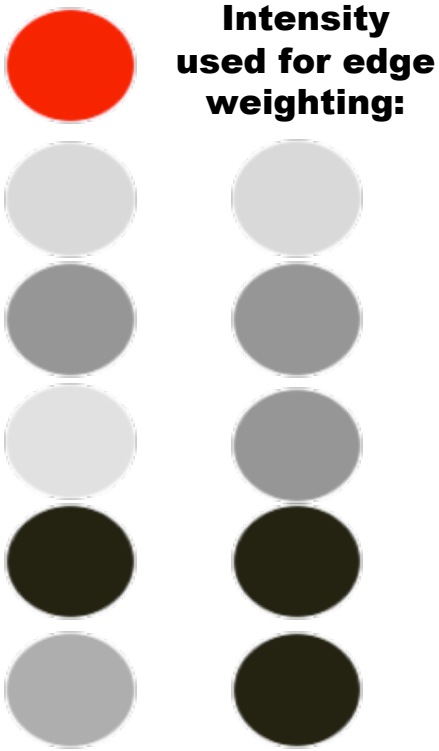


Figure 3: Diagram illustrating the logic behind deciding which pixel intensity value to use when calculating edge weights. Pixel intensities along one "arm" of the stencil in Figure 2 are shown on the left; darker colors represent higher intensities. The right column indicates which intensity value will be used when calculating the edge weight between the "active" (red) pixel and the adjacent pixel. [**NR**] adam2/. weights

with each pixel as its own image segment; then individual pixels, and eventually, groups of pixels, are merged according to thresholding criteria. Segments can also be merged in post-processing if they are smaller than a "minimum segment size" parameter specified by the user. Figure 4 is the much-improved result when the example image in Figure 1(a) is segmented by the modified PRC algorithm.

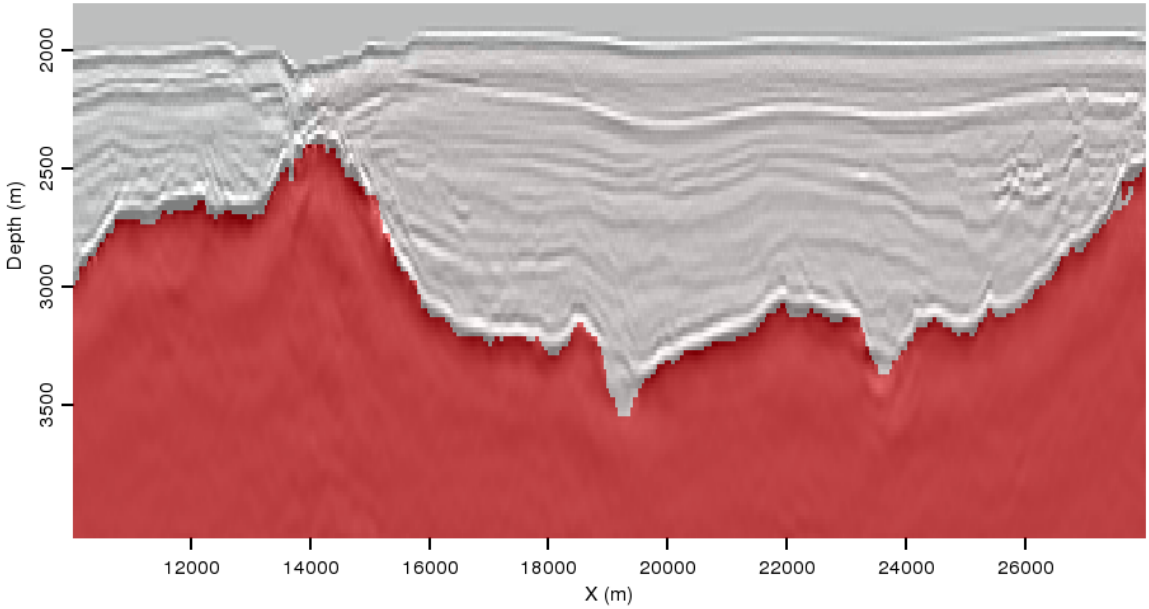


Figure 4: Final segmentation result of the image in Figure 1(a), after modifications to the algorithm. [**ER**] adam2/. uno-segmerge

INTERPRETER GUIDANCE

Unfortunately, a fully-automatic method will often be insufficient for obtaining an acceptable salt interpretation; acquisition, model-building, and imaging challenges all contribute to situations in which salt boundaries appear faint, discontinuous, or not to be present at all. In another 2D Gulf of Mexico example image (Figure 5(a)), for example, there are locations along both the top and base of the salt body where the boundary is poorly imaged. This can result in "leakage" of the automatically interpreted salt segments (Figure 5(b)). In these cases, valuable interpreter insight should be incorporated into the procedure. Figure 5(c) shows manual salt boundary interpretations in areas where leakage is apparent in Figure 5(b). The most efficient way to include this information in the PRC algorithm is to modify the input image by increasing intensity values at the manual pick locations. Instead of assigning arbitrarily large values, however, we define a new amplitude value (A) for a "picked" pixel at position (x,y,z) in terms of the highest-amplitude pixel in a neighborhood surrounding it and a scaling factor α:

$$A_{xyz} = \alpha \max_{|x-i|\leq 5, |y-j|\leq 5, |z-k|\leq 5} A_{ijk}. \tag{2}$$

This ensures that the picked boundary will not appear radically different from its surroundings, which could present challenges for the automatic segmentation algorithm. Now,

segmenting the new input image with parameters identical to the original segmentation yields the result seen in Figure 5(d). The segments conform to the manual picks seen in Figure 5(b), while the rest of the image is segmented as accurately as the original result in Figure 5(b).

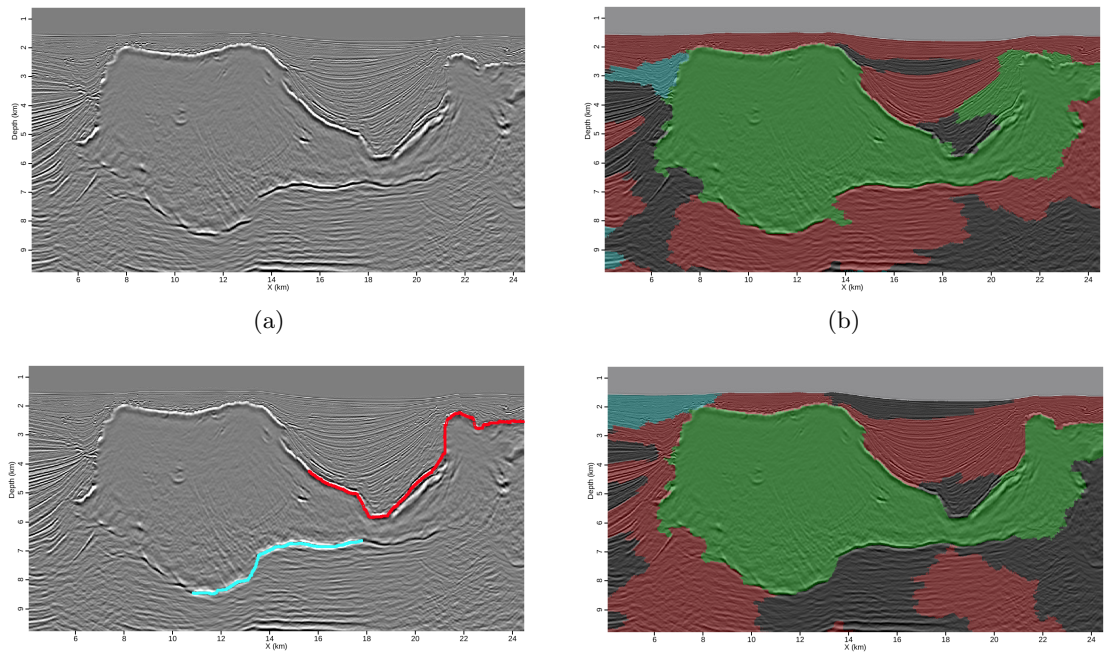


Figure 5: (a) A 2D section from the Gulf of Mexico; (b) Segmentation result using the unaltered algorithim from Felzenszwalb and Huttenlocher (2004); (c) Manual salt picks supplied to guide the automatic segmentation; (d) Segmentation result after interpreter guidance. [**ER**] adam2/. oct-2d,2d-orig,2dtopbase,seg-final

Additional modifications are required for 3D images. Because segments are much larger in 3D, amplitude changes on a single 2D section are not significant enough to alter 3D segmentation results. Instead, we must "project" an interpreter's manual picks on an inline section, into the third (crossline) dimension. To do this, we make the assumption that a salt boundary will not fluctuate by more than two pixels per slice in the crossline direction, and construct a square pyramid in the crossline direction like the one depicted in Figure 6. The pyramid has sides of length $2h$, where h is the number of crossline samples between the base of the pyramid and its apex, which is the manually interpreted point. Now, for any pixel Q that falls within a pyramid with an apex at point P, the new amplitude value at point Q is

$$A^Q_{\text{new}} = A^Q_{\text{orig}} + \frac{A_0}{||PQ||}, \tag{3}$$

where A_0 is the amplitude value at point P as determined by equation 2, and $||PQ||$ is the distance between the two points. The expression is additive to ensure that any hint of the boundary already present will not be overwhelmed by the interpretation on a nearby slice.

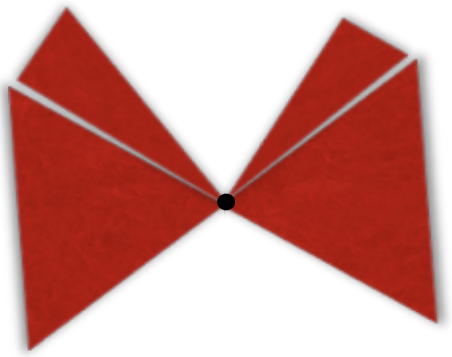


Figure 6: Depiction of the pyramid used to "project" an interpreter's picks from a single 2D slice into the third dimension. The influence of the pick (at the apex of the pyramid) decays with distance. [**NR**] adam2/. pyramid

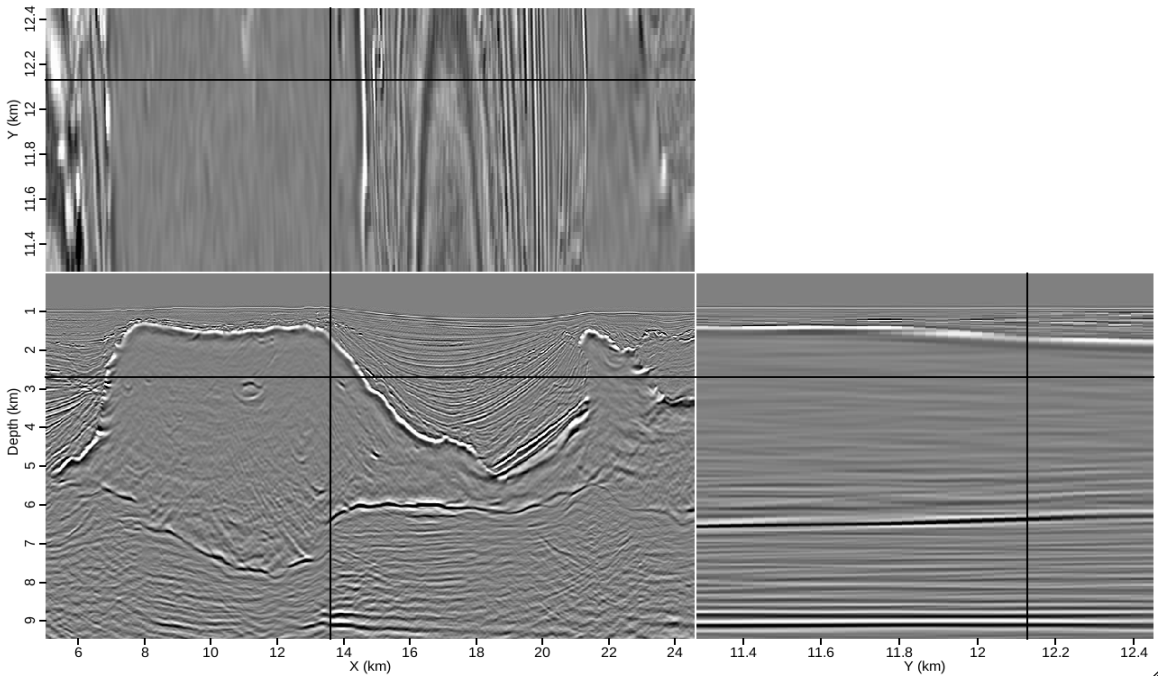
3D FIELD DATA EXAMPLE

Figure 7(a) shows slices through a 3D image cube from a Gulf of Mexico dataset provided by WesternGeco. From the initial segmentation result (Figure 7(b)), it is clear that the salt boundary discontinuities present challenges for the automatic segmentation algorithm. To correct the apparent leakages, manual interpretations are supplied for selected locations at two crossline locations (Figure 8). The effects of these manual picks on the input amplitude data are shown in Figures 9(a) and 9(b). Not only are higher amplitudes obvious at the pick locations themselves, but the procedure described in the previous section has clearly influenced the intensity values at neighboring crossline values. Now, the updated segmentation result (Figure 10) is improved on both the inline and crossline sections. To emphasize the improvement, Figure 11(a) is the initial segmentation result shown far from any of the manual pick locations, while Figure 11(b) is the result after interpreter guidance. Even far from the actual picks, the automatic segmentation process is significantly more accurate when incorporating interpreter guidance.

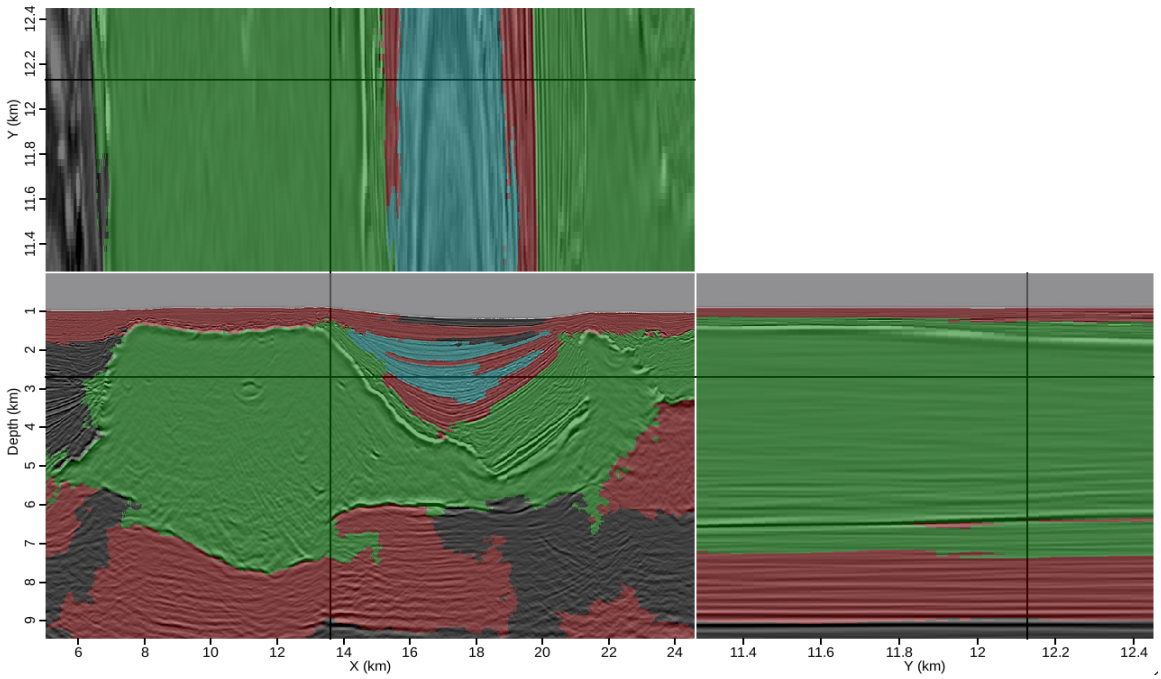
As expected, the algorithm does operate extremely efficiently. This 3D example had over 30 million pixels, and over 700 million graph edges were constructed. A single CPU performed the segmentation in less than three minutes, highly efficient compared to other segmentation techniques. Furthermore, this method operates on the entire image cube, rather than a limited domain around the salt body. This opens the door for additional interpretation aids within the algorithm's capabilities, such as stratigraphic segmentation.

CONCLUSIONS

Applying the modified Pairwise Region Comparison (PRC) segmentation algorithm to 2D and 3D field seismic images allows for accurate, semi-automatic salt structure delineation. While fully automatic segmentations are sometimes successful, limited manual interpretations on one or more 2D slices can be used to guide a 3D segmentation process. This allows for improved results throughout the image cube, not just near manual pick locations. The new algorithm performs extremely efficiently compared to other automatic interpretation techniques, and operates on the full seismic image or cube rather than a limited domain around a supposed salt structure. Therefore, this could make it an important tool for interactive interpretation procedures that can streamline the model building workflow.



(a)



(b)

Figure 7: (a) Slices through a 3D image cube from the Gulf of Mexico; (b) Segmentation result prior to interpreter guidance. [**CR**] adam2/. oct-3d,3d-origseg

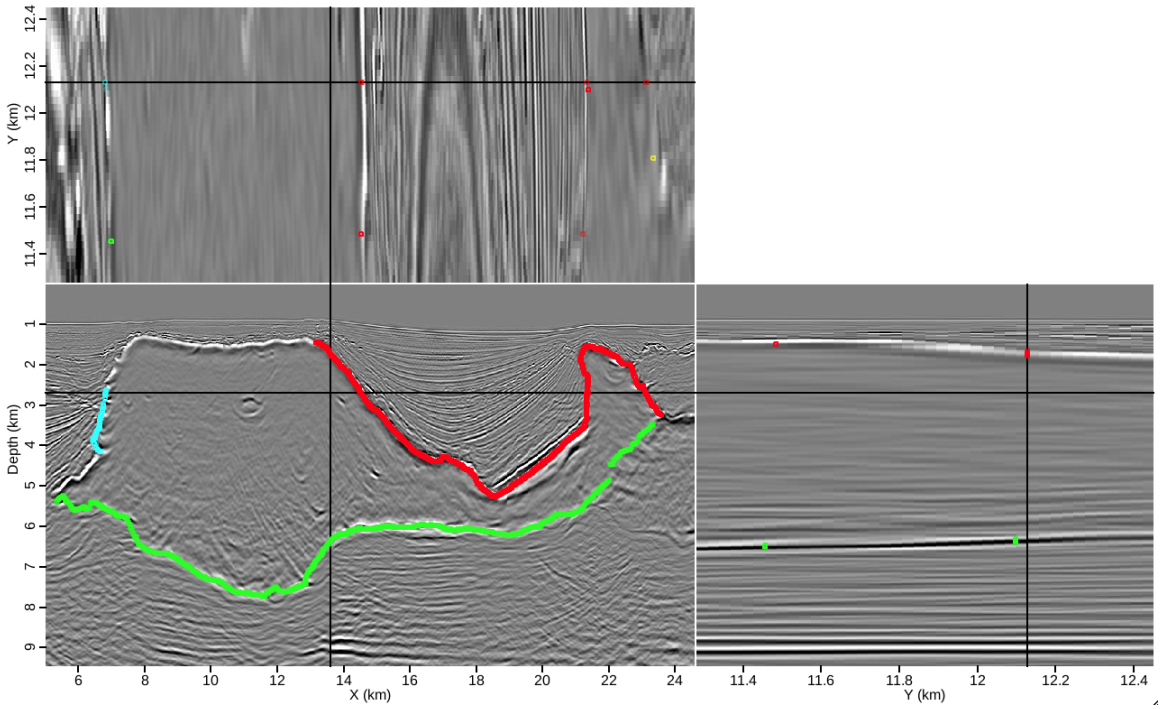


Figure 8: Manually-interpreted salt picks used to guide the automatic 3D segmentation. [**ER**] adam2/. oct-3d-picks

ACKNOWLEDGMENTS

I thank Unocal (now Chevron) and WesternGeco for providing the data used for examples, and all SEP sponsors for their support.

REFERENCES

Berthelot, A., A. H. S. Solberg, E. Morisbak, and L.-J. Gelius, 2012, 3[d] segmentation of salt using texture attributes: SEG Technical Program Expanded Abstracts 2012, 1–5.

Felzenszwalb, P. F. and D. P. Huttenlocher, 2004, Efficient graph-based image segmentation: International Journal of Computer Vision, **59**, 167–181.

Hale, D. and J. U. Emanuel, 2002, Atomic meshing of seismic images, *in* Expanded Abstracts, 2126–2129, SEG.

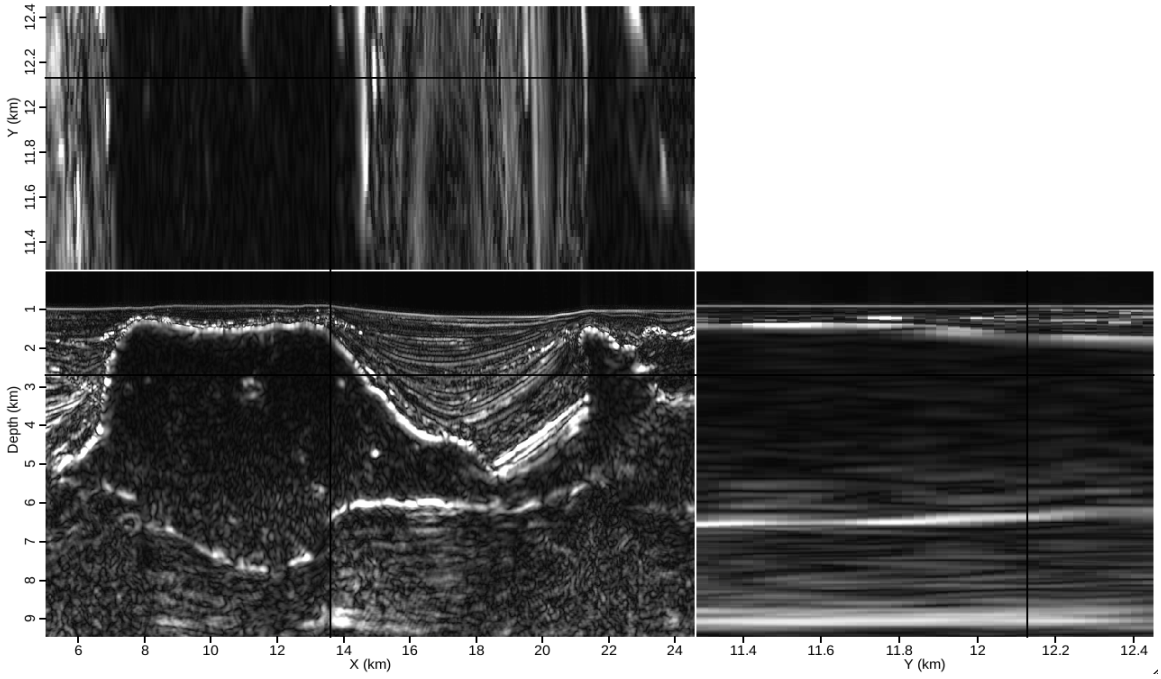
———, 2003, Seismic interpretation using global image segmentation, *in* Expanded Abstracts, 2410–2413, SEG.

Lomask, J., 2007, Seismic volumetric flattening and segmentation: PhD thesis, Stanford University.

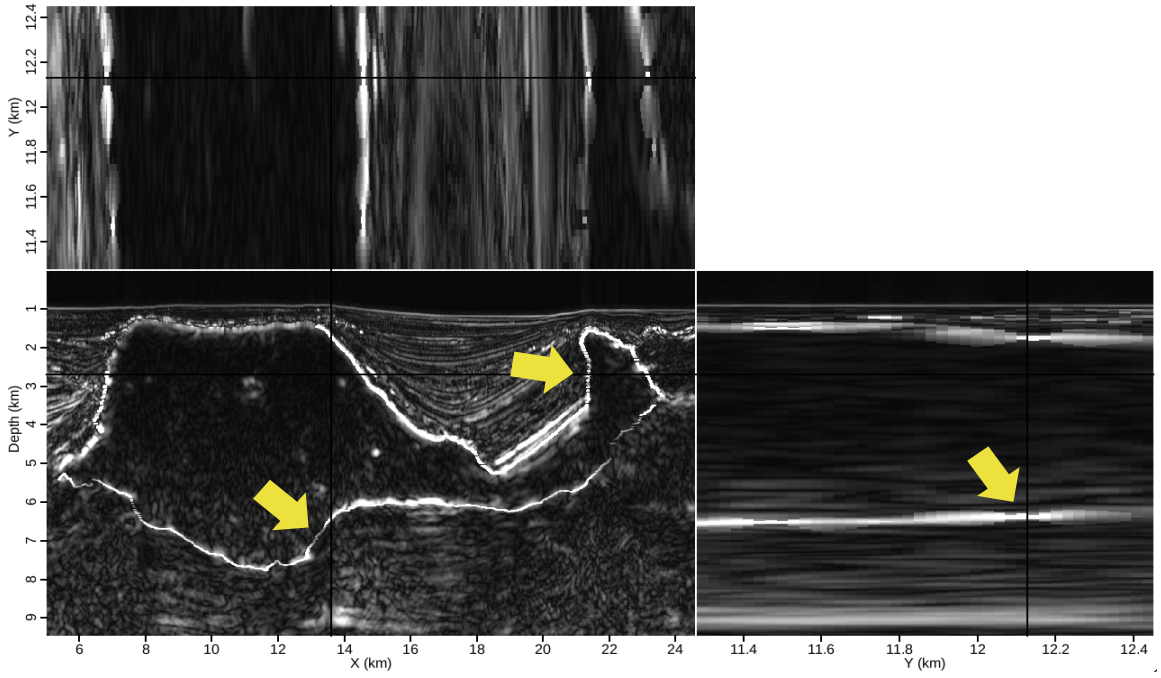
Lomask, J., R. G. Clapp, and B. Biondi, 2007, Application of image segmentation to tracking 3ṣalt boundaries: Geophysics, **72**, P47–P56.

Mosher, C., E. Keskula, J. Malloy, R. Keys, H. Zhang, and S. Jin, 2007, Iterative imaging for subsalt interpretation and model building: The Leading Edge, **26**, 1424–1428.

Shi, J. and J. Malik, 2000, Normalized cuts and image segmentation: Institute of Electrical



(a)



(b)

Figure 9: Amplitude of the envelope volumes (a) before and (b) after modification according to the interpreter guidance scheme. Changes are particularly noticeable at the indicated locations. [**ER**] adam2/. 3d-env-orig,3d-env-newa

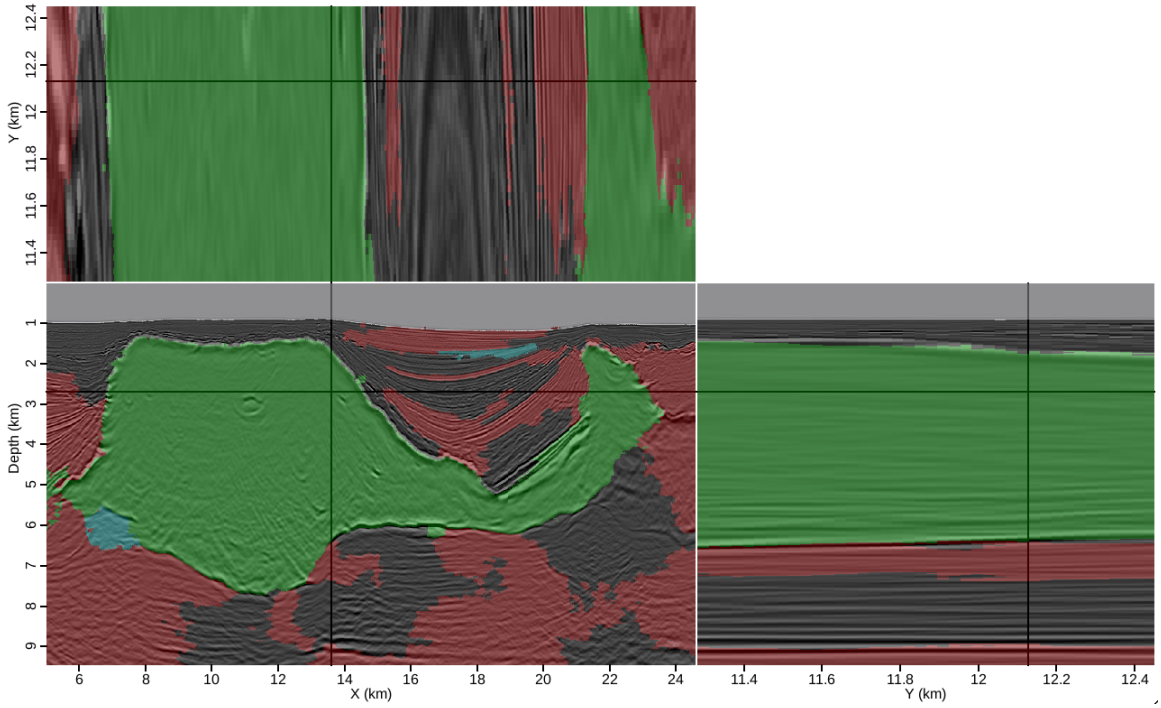
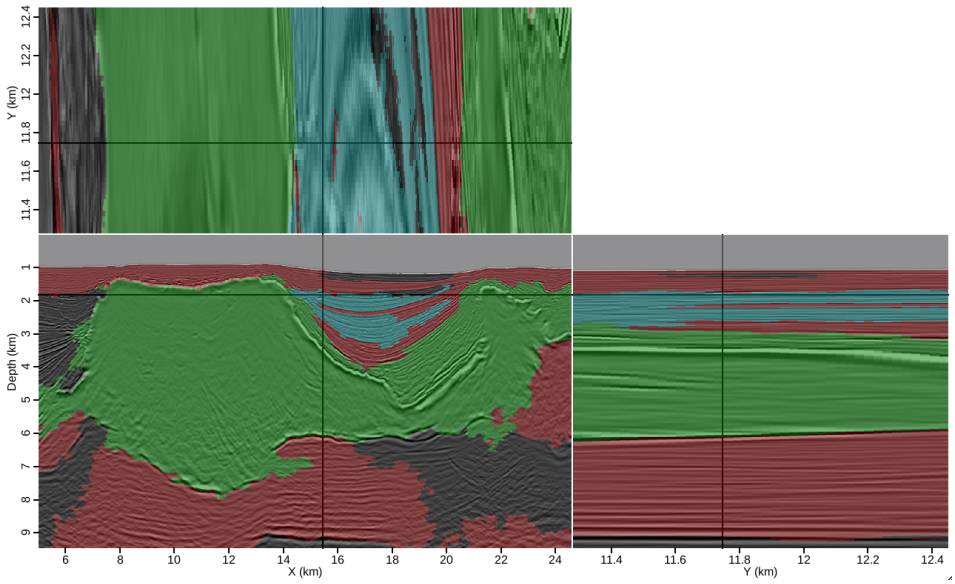


Figure 10: Segmentation result incorporating interpreter guidance. This result is much more accurate than the initial result in Figure 7(b). [**CR**] adam2/. 3d-newseg

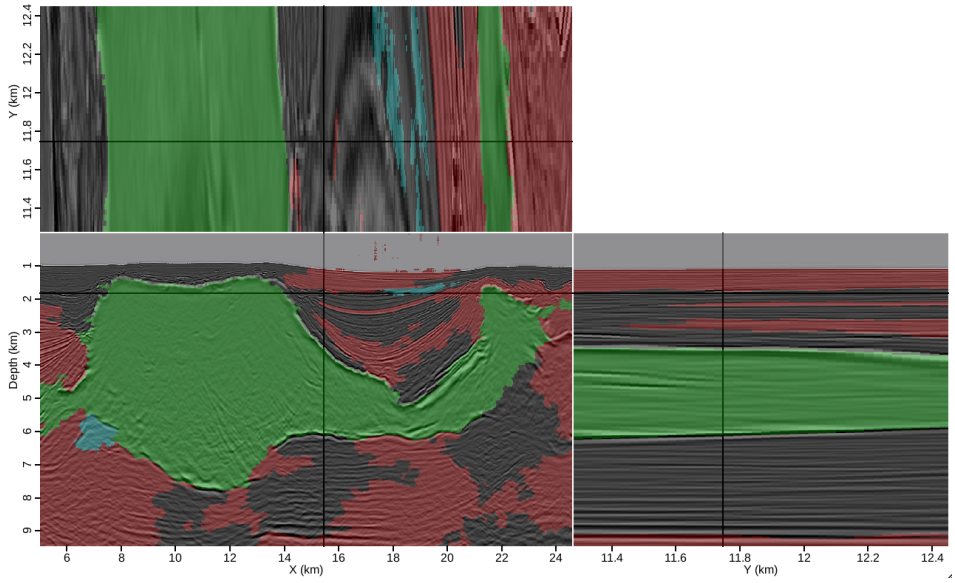
and Electronics Engineers Transactions on Pattern Analysis and Machine Intelligence, **22**, 838–905.

Valet, L., G. Mauris, P. Bolon, and N. Keskes, 2001, Seismic image segmentation by fuzzy fusion of attributes: IEEE Transactions on Instrumentation and Measurement, **50**, 1014–1018.

Zahn, C. T., 1971, Graph-theoretical methods for detecting and describing gestalt clusters: IEEE Transactions on Computers, **C-20**, 68–86.



(a)



(b)

Figure 11: Segmentation results for another set of slices through the 3D cube, (a) before and (b) after interpreter guidance. The result is greatly improved, even far from where the manual salt picks were supplied. [**CR**] adam2/. 3d-origseg-far,3d-newseg-far

Large scale linearised inversion with multiple GPUs

Chris Leader and Robert Clapp

ABSTRACT

As our computational power develops and evolves so does our desire to process more data with more advanced algorithms. Graphical Processing Units (GPUs) have been shown to accelerate algorithms that have a high computation to memory access ratio. However, their relatively small global memory (6 Gb) poses additional restrictions on how we must adapt our problem. Herein will be described how multiple GPUs can be used to accelerate the problem of wave-equation linearised inversion in such a way that poses no model or data size restrictions. Furthermore by splitting the internal and external parts of our domains we can achieve close to linear strong-scaling with the number of GPUs that we are using. Consequently we can design a method that outperforms CPU based inversion while overcoming the traditional restrictions of GPU based inversion.

INTRODUCTION

In today's world of heterogeneous computing systems it is increasingly important to design and tailor one's algorithm to the available hardware. For seismic imaging we are increasingly looking to numerically intensive inverse techniques to improve the quality, frequency content and fidelity of our images, and more and more often we are using heterogeneous computing systems to achieve this. This is a demanding process for both the computer and the programmer. Many of the largest supercomputers in the world are incorporating hybrid systems of CPUs and GPUs and the effective use of these two different processing units is paramount for system performance. Migrating legacy codes that worked and scaled well on clusters of CPUs to GPUs can easily run into problems; for seismic imaging these are generally related to compounded disk access and the limited global memory provided by a single GPU. The focus of this discussion will be on using a two-way wave-equation engine for recovering subsurface scattering potential, using a linearised inversion scheme.

Ohmer et al. (2005) and Foltinek et al. (2009) show how GPUs can greatly assist any operation that can be considered as Single Instruction Multiple Data (SIMD) by running thousands of independent threads concurrently across a given domain; two-way wave propagation can be considered as a SIMD operation as we are convolving a set stencil many times. However such a set up has disadvantages; the GPU can not read directly from disk, thus any disk based IO must be explicitly routed through a host CPU, compounding any such memory access. Furthermore the dynamic memory available on a GPU is 6 Gbytes or less, meaning that for propagation we are limited to a model size of 793 pt^3 for modelling and 600 pts^3 for imaging, assuming we are using acoustic, isotropic propagators. These numbers are significantly reduced when performing anisotropic and/or elastic propagation. Clapp (2009), Fletcher and Robertsson (2011) and Leader and Clapp (2011a) discuss how Reverse Time Migration (RTM) can be adapted to minimise disk access during propagation

by using random boundaries, and hence better harness the computational power of the GPU without sacrificing significant performance for data movement.

LINEARISED INVERSION

We can describe an idealised modelling procedure as in equation 1, which is based on the first approximation of the Born scattering series. The adjoint of this process, equation 2, is often referred to as Reverse Time Migration (RTM). Here d is the data, f the source function, G_0 are the respective Green's functions, m the model, $\mathbf{x}$ the 3D model coordinates, $\mathbf{x}_{r,s}$ the 3D source and receiver coordinates and $*$ denotes the complex conjugate. Despite this mathematical treatment assuming single scattering, our propagator makes no such assumption and multiple scattering / diving waves are still positioned correctly.

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

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

By using these processes, along with an appropriate model updating scheme, we can perform linearised inversion. This is advantageous over RTM because it iteratively removes unphysical aspects of the recovered model - such as acquisition footprints, the effect of the wavelet and unbalanced amplitudes (Lambare et al., 1992). For each iteration we must run the forward and adjoint processes, making each update about twice as expensive as RTM. The fact that the GPU can perform wave propagation so fast makes such an imaging process possible.

GPU BASED INVERSION

Optimally adapting CPU based linearised inversion to a set of GPUs requires careful consideration of the architecture differences. While the GPU can accelerate floating point computation for comparatively less effort, it has a smaller global memory and cannot directly read or write from disk. This means disk access is compounded and must be avoided to observe acceptable performance. This fact, combined with our accelerated computation, makes memory transfer and disk access appear a much larger problem than on a CPU.

As previously described in Clapp (2008a) and Leader and Clapp (2011b), the two wavefields in our adjoint process have an opposite sense of time, as denoted by the complex conjugate in equation 2, thus generally one is precomputed and stored on disk. We can circumnavigate this disk use by padding our computational domain with random boundaries making our source modelling time reversible. We must perform an extra propagation per time step under this regime, but this is much cheaper than accessing the disk-stored wavefield Clapp (2008b).

We can take two steps to further optimise our GPU based scheme (assuming our propagator is using shared memory efficiently (Nickolls et al., 2008).) These are aligning warps

within the GPU and using shared memory to store the velocity array during adjoint propagation. The second of these was described in Leader and Clapp (2012).

A GPU kernel can be viewed as a Single-Instruction Multiple Threads (SIMT) process, and a given single instruction is issued to a warp (or thread-vector) which guarantees vectorisation. For a typical NVIDIA GPU, a warp is a vector of 32 contiguous threads. There are two ways we can use warps to increase performance: we can ensure each warp is only issued a single instruction (to avoid divergence within a warp) and we can encourage coalesced memory access (Micikevicius, 2011). The second of these is applicable to finite difference kernels.

Memory addresses from a warp are converted into line requests, these can be 32B or 128B depending on cache options. A single thread within a given warp can access a memory address from within that line at 100% bus utilisation; however, if a neighbouring thread requests a memory address from an adjoining warp, then additional bytes must be moved across the bus. Consistent misaligned memory access such as this can drastically decrease bus utilisation and hence overall kernel performance (as in Fig 1). By encouraging our domain to align along these line requests we can see an improvement of up to 20% (in this case) for minimal effort.

In our specific case we are using an 8th order 3D finite-difference stencil over our domain. Practically, this means that we pad our axes by four points on each side to account for the 'halo' (the width of the stencil arms) so that we do not try to access points outside of our domain. All we must do is to pad our fast axis by an additional 28 points so that our outer region now falls within a warp. The nature of our stencil means we cannot avoid some out-of-warp memory requests, but this additional padding, while marginally increasing our domain size, reduces our run time by about 20%.

MULTIPLE GPU EXTENSION

There are three reasons why we might want to use multiples GPUs: if our domain exceeds the memory of a single unit, to accelerate computation for a smaller domain size, or to increase performance per watt (Micikevicius, 2012). To create our model size independent inversion scheme we must split our domain across multiple GPUs.

For propagation we have to allocate a minimum of three large 3D fields - two wavefield slices and the velocity model (the third wavefield slice can replace the first) which confines us to a symmetric model of 793 pts^3, including padding regions. For reverse time migration and for linearised modelling we need two slices for the source wavefield, two for the receiver wavefield, the velocity model, the image and the data. Assuming our time and depth axes are comparable in length, our restriction is now around 600 pts^3.

We are free to break up our domain how we choose, however in this case we find it advantageous to break it up along the depth (z) axis, which is also the slowest axis. The reasons here are twofold - breaking up along one axis does require slightly more allocation, but the halo region is contiguous in memory (if broken along the slowest axis). This alleviates the need for a separate kernel that simply aligns the correct memory values. The second reason is that our receiver and source positions are in a plane near the surface; this means that only one GPU needs to have all the geometry information and needs to deal

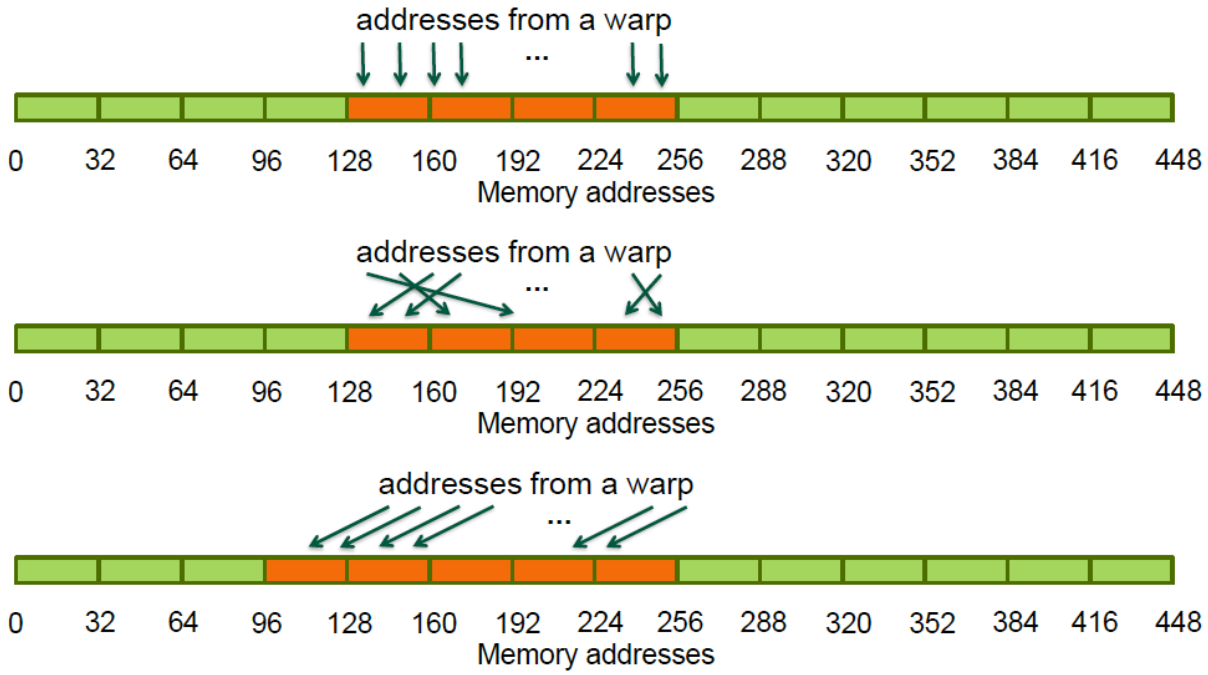


Figure 1: Schematic of warps alignment with memory requests. The above two cases feature 100% bus utilisation, the bottom case has 80% due to misaligned memory accesses. From Micikevicius (2011). chris1/. warps

with data/source injection/collection. Thus we do not need to loop over GPUs to perform these operations.

We need our sub-domains to overlap along the depth axis by half the stencil length, as this information is needed from neighbouring domains to continue to the next time step and to move the wave through the domain boundaries. Consequently, between time steps this halo region must be transferred and then synchronised before moving to the next time step, else we run the risk of spurious propagation. Recent NVIDIA releases have allowed Peer to Peer (P2P) GPU communication and data transfer, rather than explicitly routing all information transfer between GPUs first through a CPU buffer. Additionally, neighbouring GPUs can now operate on a Unified Virtual Address space (UVA), meaning there is no risk of dereferencing a pointer that has the same address on a different GPU. These two new additions make this GPU-GPU information transfer more efficient and easier to implement.

We can make this process scale (very close to) linearly by using asynchronous memory copies and kernel calls. When doing this it is possible to overlap (halo) communication with (internal) computation, as in figure 3, 'hiding' the communication time (Micikevicius, 2012). This can be done by associating certain calls to separate GPU streams, where a stream can be considered as a command pipeline. Within a stream calls are serial, however different streams can execute concurrently (for certain operations). By restricting halo computation and communication to one such stream, and the other data (internal) computation to another, we can hide the halo communication by overlapping the memory transfer with the internal kernel. We can first calculate the wavefield values within these halos regions, then we can send the information to the 'right' (GPUs with higher IDs),

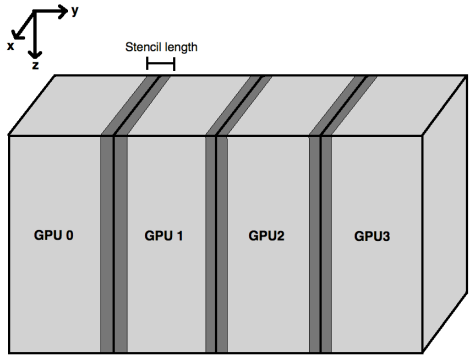


Figure 2: A diagram of how to decompose a model (in this case along the y-axis), dark grey regions are allocated on both neighbouring GPUs. [**NR**] chris1/. domdec

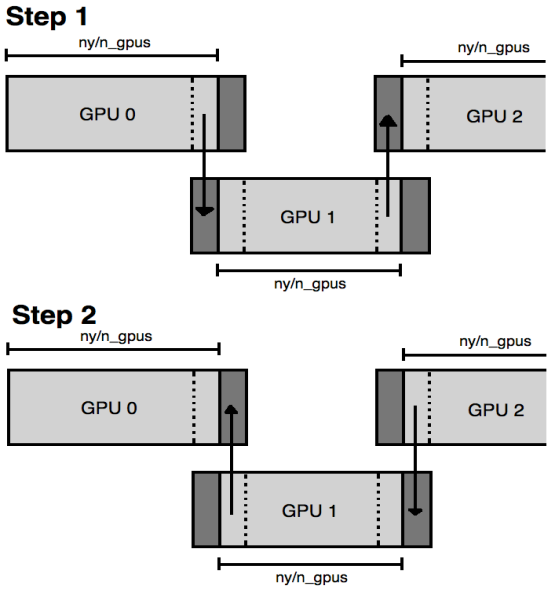


Figure 3: A diagram of the two stages of GPU halo communication, simplified to 1D. [**NR**] chris1/. halos

then send to the 'left,' then synchronise. Again, this is shown in fig 3. The PCIe bus that links the GPUs is duplex, meaning it can send and receive at the same time in two different directions. However, if an application tries to send and receive along the same direction at the same time, the link can stall. So between the two send operations we add a synchronisation step. Some simplified CUDA code using streams is displayed below for reference.

```
for(i_gpu=0; i_gpu < n_gpus; i_gpu++){
   cudaSetDevice(i_gpu);
   kernel<<<...halo_region[],halo_stream[i_gpu]>>>(...);
   kernel<<<...internal_region[],internal_stream[i_gpu]>>>(...);
}
for(i_gpu=1; i_gpu < n_gpus; i_gpu++){
   cudaMemcpyPeerAsync(...halo_stream[i_gpu]);
}
for(i_gpu=0; i_gpu < n_gpus; i_gpu++){
   cudaStreamSynchronise(...halo_stream[igpu]);
}
for(i_gpu=0; i_gpu < n_gpus-1; i_gpu++){
   cudaMemcpyPeerAsync(...halo_stream[i_gpu]);
}
for(i_gpu=0; i_gpu < n_gpus; i_gpu++){
   cudaDeviceSynchronise();
}
```

This discussion has been limited to the problem of propagation thus far. Performing linearised modelling and migration over multiple GPUs adds a level of complexity. For linearised modelling we must propagate two fields, apply a scattering condition, inject the source term and perform data collection. For RTM, we must propagate two fields, apply an imaging condition, and inject both the source term and the recorded data.

The corresponding adjoint routine code, grossly simplified, will look something like:

```
for(i_gpu=0; i_gpu < n_gpus; i_gpu++){
   cudaSetDevice(i_gpu);
   kernel<<<...data_halo_region[],halo_stream[i_gpu]>>>(...);
   kernel<<<...source_halo_region[],halo_stream[i_gpu]>>>(...);
   kernel<<<...data_internal_region[],internal_stream[i_gpu]>>>(...);
   kernel<<<...source_internal_region[],internal_stream[i_gpu]>>>(...);
}
for(i_gpu=1; i_gpu < n_gpus; i_gpu++){
   cudaMemcpyPeerAsync(data_halo_region[i_gpu]...halo_stream[i_gpu]);
   cudaMemcpyPeerAsync(source_halo_region[i_gpu]...halo_stream[i_gpu]);
}
for(i_gpu=0; i_gpu < n\_gpus; i_gpu++){
   cudaSetDevice(i_gpu);
   cudaStreamSynchronise(...halo_stream[igpu]);
```

```
}
for(i_gpu=0; i_gpu < n_gpus-1; i_gpu++){
   cudaMemcpyPeerAsync(data_halo_region[i_gpu]...halo_stream[i_gpu]);
   cudaMemcpyPeerAsync(source_halo_region[i_gpu]...halo_stream[i_gpu]);
}
cudaSetDevice(0);
src_inject_kernel<<<...source_internal_region[],internal_stream[i_gpu]>>>(...);
data_inject_kernel<<<...data_internal_region[],internal_stream[i_gpu]>>>(...);
for(i_gpu=0; i_gpu < n_gpus; i_gpu++){
   img_kernel<<<...,internal_stream[i_gpu]>>>(...);
}
for(i_gpu=0; i_gpu < n_gpus; i_gpu++){
   cudaDeviceSynchronise();
}
```

We assign the injection and imaging kernels to the internal stream since these steps all depend on the previous having completed and we are operating on the internal region for injection and collection. Each of these operations individually scale roughly linearly with problem size, so they require no unique treatment. The final normalised speed ups as a function of number of GPUs used can be seen in Figure 4. We can see that we achieve very close to linear scaling. The reason we do not get to 100% is that splitting the halo and internal computation requires a small amount of overhead, meaning we typically see around 92% to 96% of strong-scaling speed up with number of units used. If we do not use the communication overlap we see a much worse performance, with numbers typical of a naive OpenMP implementation on an eight-core CPU.

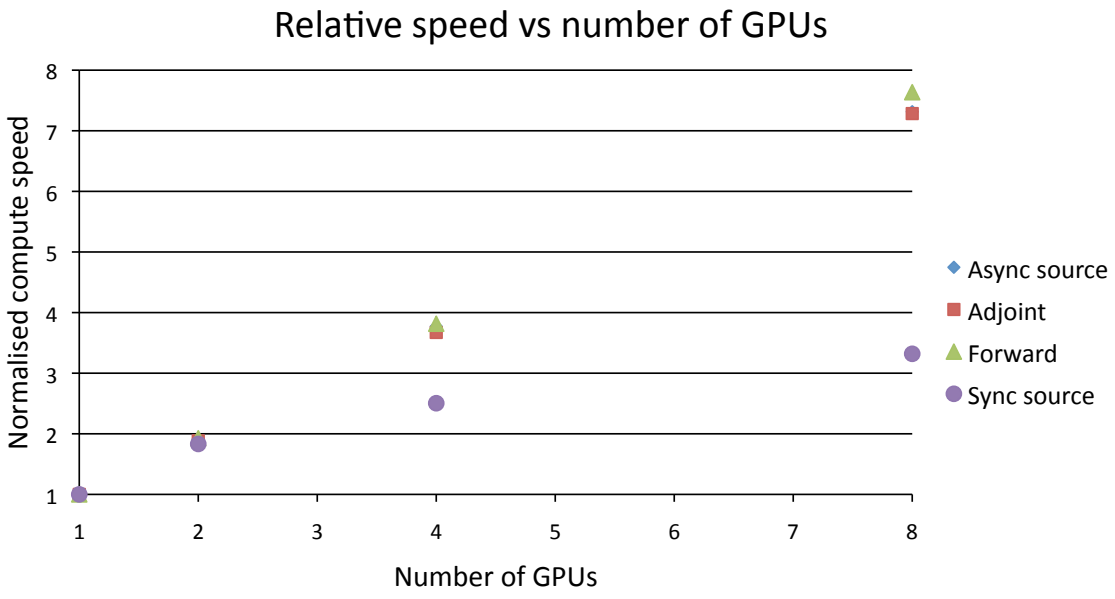


Figure 4: A graph of how synchronous source modelling, asynchronous source modelling, the forward process and the adjoint process scale with number of GPUs. [**NR**] chris1/. speeds

CONCLUSIONS

Previous reports have discussed how RTM and linearised inversion must be adapated to run efficiently on a single GPU. Descirbed herein was how we can futher improve performance through warp alignment, and results on inversion across multiple GPUs. Through warp alignment we can accelerate the propagation part of our algorithm by 20% for minimal extra work. Then by splitting the computation and communication of different parts of our domain we can propagate our fields over multiple GPUs, and we can scale almost linearly with the number of GPUs that we use. When performing linearised inversion over 8 GPUs we get a relative speed up of 7.6.

ACKNOWLEDGMENTS

Thanks to Paulius Micikevicius at NVIDIA for invaluable GPU troubleshooting advice and code sharing. Further thanks to all sponsors of the Stanford Exploration Project for their continued financial and intellectual support.

REFERENCES

Clapp, R., 2008a, Reverse time migration: Saving the boundaries: SEP-136, **136**.

———, 2008b, Reverse time migration with random boundaries: SEP-138, **138**.

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

Fletcher, R. P. and J. O. A. Robertsson, 2011, Time-varying boundary conditions in simulation of seismic wave propagation: SEG Technical Program Expanded Abstracts, **30**, 2957–2961.

Foltinek, D., D. Eaton, J. Mahovsky, P. Moghaddam, and R. McGarry, 2009, Industry scale reverse time migration on GPU hardware: SEG Technical Program Expanded Abstracts, **28**, 2789–2793.

Lambare, G., J. Virieux, R. Madariaga, and S. Jin, 1992, Iterative asymptotic inversion in the acoustic approximation: Geophysics, **57**, 1138–1154.

Leader, C. and R. Clapp, 2011a, Memory efficient reverse time migration: Stanford Exploration Project Report, **143**.

———, 2011b, Memory efficient reverse time migration: SEP-143, **143**.

———, 2012, Linearised inversion with GPUs: SEP-147, **147**.

Micikevicius, P., 2011, Performance optimization: Presentation at Supercomputing 2011.

———, 2012, Multi-GPU programming: Presentation at GPU Technology Conference 2012.

Nickolls, J., I. Buck, M. Garland, and K. Skadron, 2008, Scalable parallel programming with cuda: Queue, **6**, 40–53.

Ohmer, J., F. Maire, and R. Brown, 2005, Implementation of kernel methods on the GPU: DICTA'05 Expanded Abstracts, **78**.

Efficient velocity model evaluation: 2D and 3D field data tests

Adam Halpert

ABSTRACT

Interpretation tools like image segmentation allow for fast generation of velocity models, but evaluating them can be computationally demanding. Previous work has shown the effectiveness of using Born-modeled synthesized wavefields to image targeted locations for this purpose, but only on synthetic data. Here, the model evaluation method is successfully demonstrated on both a 2D line and small 3D cube taken from a wide-azimuth field survey. The importance of a quantitative measure of focusing or image quality is also shown, especially for five-dimensional datasets that are difficult to visualize and judge qualitatively.

INTRODUCTION

In recent years, model-building bottlenecks have shifted from being computational or imaging-related, toward interpretation and other human-intensive tasks. This can be attributed to significant increases in computing power, without accompanying advances in interpretation tools or workflow processes. However, there is an opportunity to not only reduce the magnitude of this bottleneck, but to also take advantage of increased computational capability as it relates to model building. Instead of determining a single, "best-fit" model to use as the basis of imaging, it may be feasible to create a range of models, based on different geological scenarios. This can be especially useful for salt interpretation workflows; because of the high velocity contrast between salt and sediments, mis-placed or mis-interpreted salt bodies can have a disproportionately negative impact on subsalt images. While tools such as seismic image segmentation (Lomask et al., 2007; Halpert, 2013) can help provide these models, a way to efficiently test them is necessary in order to avoid costly and time-consuming re-migration of very large modern 3D datasets.

Quickly updating images based on discrete changes to the velocity model is a problem with several proposed solutions. Many proposals rely on a fast, target-oriented variant of beam migration (Hill, 1990); for example, Wang et al. (2008) update post-stack images to qualitatively judge the effects of using different velocity models. An alternative is to use a similar approach but with wave-equation methods like reverse-time migration (RTM) (Wang et al., 2011); however, this remains an extremeley expensive option. A different approach, first proposed by Halpert and Tang (2011), also employs wave equation imaging. However, by using an initial image and Born modeling (Stolt and Benson, 1986) to synthesize a new, targeted dataset (Tang and Biondi, 2010; Tang, 2011), the computational challenge is drastically reduced. Furthermore, if a pre-stack image with subsurface offset information is available, this method can make use of that information to identify and help correct errors in an initial velocity model. In Halpert (2012), initial tests using a 2D synthetic model were

shown to be successful; here, I will present initial tests on 2D and 3D portions of a wide-azimuth field dataset. Using the field data, I will show that Born-synthesized wavefields can quickly discriminate between different velocity models, and quantitatively identify the most accurate of the options.

METHOD

The procedure for generating Born-modeled wavefields from an initial image was first outlined in Halpert and Tang (2011), and consists of three major steps:

1. Starting from an initial image, use a form of exploding reflector modeling to "record" an areal source function at a preferred datum.

2. Using the new source function and the initial image as a reflectivity model, synthesize a Born-modeled dataset (Tang, 2011) with arbitrary acquisition geometry suited to the imaging target(s).

3. Now, the synthesized source and receiver wavefields can be imaged conventionally, using any velocity model under consideration. New images can be produced using only a single shot, a substantial computational savings over re-migrations of the full dataset.

Areal source function

The areal source function is generated by first mapping subsurface offset information from a prestack image m onto its equivalent zero-offset positions, and performing exploding reflector modeling to record the source function at a user-specified datum. Choosing a datum deeper in the subsurface could lead to significant computational savings if the "target" for this procedure is very deep or below salt. The source function $S(\mathbf{x}_s, \omega)$, recorded at $\mathbf{x}_s = (x_s, y_s, z_s)$ can be represented as

$$S(\mathbf{x}_s, \omega) = \sum_{\mathbf{x}'} \sum_{\mathbf{h}} G^*(\mathbf{x}' - \mathbf{h}, \mathbf{x}_s, \omega) m(\mathbf{x}', \mathbf{h}), \tag{1}$$

where $\mathbf{h}$ is the vector of subsurface half-offsets; ω is angular frequency; and G is the Green's function propagating the wavefield to the receiver locations (here, * denotes the adjoint). For this step and all subsequent ones, crosstalk issues are avoided by selecting only isolated locations along an interpreter-selected horizon.

Born-modeled wavefields

The Born-modeled dataset d' is recorded it at arbitrary receiver locations $\mathbf{x}_r$:

$$d'(\mathbf{x}_r, \omega) = \sum_{\mathbf{x}'} \sum_{\mathbf{h}} \Gamma(\mathbf{x}_s, \mathbf{h}, \omega) G(\mathbf{x}' + \mathbf{h}, \mathbf{x}'_r, \omega) m(\mathbf{x}', \mathbf{h}). \tag{2}$$

Here, m is the reflectivity model (in our case, the isolated regions from the initial image), and the Γ term is defined as

$$\Gamma(\mathbf{x}_s, \mathbf{h}, \omega) = \sum_{\mathbf{x}_s} S(\mathbf{x}_s, \omega) G(\mathbf{x}_s, \mathbf{x}' - \mathbf{h}, \omega), \tag{3}$$

where S is the source function described in step 1. Because the velocity model used to compute the Green's functions is the same one used to generate the initial image, the resulting Born-modeled wavefield $d'(\mathbf{x}'_r, \omega)$ will be kinematically invariant regardless of the initial model (Tang, 2011). This is important for the goal of correcting inevitable errors in the initial velocity model.

Imaging with multiple models

Now, the synthesized areal source function and receiver wavefield can be imaged via standard downward continuation migration:

$$m'(\mathbf{x}', \mathbf{h}) = \sum_{\omega} G^*(\mathbf{x}' - \mathbf{h}, \omega) \sum_{\mathbf{x}_r} G^*(\mathbf{x}' + \mathbf{h}, \mathbf{x}'_r, \omega) d'(\mathbf{x}_r, \omega). \tag{4}$$

At this point, the Green's functions can be computed using *any* velocity model. In this way, a potential model can be tested using only a single shot – an enormous computational savings.

Image quality measure

In Halpert and Tang (2011), a simple quantitative measure of image quality was also introduced. By calculating the proportion of energy existing at or near zero-subsurface offset, the F measure below can quickly give an indication of the relative accuracy of different velocity models:

$$F = \frac{\sum_{i=\mathbf{p}} |A_i|}{\sum_{i=\mathbf{p}} |A_i| \exp\left(\alpha \frac{|h_i|}{h_{\max}}\right)}, \tag{5}$$

where $\mathbf{p}$ is the set of all image points, A_i is the amplitude at a given point, h_i is the subsurface offset at that point, and α is an optional user-specified weighting parameter. Using this measure, a value of $F = 1$ means that all energy is perfectly focused at zero offset; as F decreases toward zero the image becomes progressively less focused. Ideally, a measure such as this one would allow a more rigorous comparison among possible models when a more qualitative comparison fails to yield an obvious result. In addition, a quantitative measure can be especially useful when evaluating five-dimensional images, which are extremely difficult to visualize.

2D RESULTS

The 2D and 3D images used to demonstrate this method are derived from a wide-azimuth Gulf of Mexico dataset provided by WesternGeco. Figure 1(a) is a 2D migration of a

Migration model	F value
Provided velocity	0.89
Fast model	0.86
Slow model	0.85

Table 1: Calculations from equation 5 for each 2D migration velocity model, after the initial image and synthesized wavefields were created using the provided velocity model.

Migration model	F value
Provided velocity	0.92
Fast model	0.89
Slow model	0.89

Table 2: Calculations from equation 5 for each 2D migration velocity model, after the initial image and synthesized wavefields were created using a model 5% slower than the one provided.

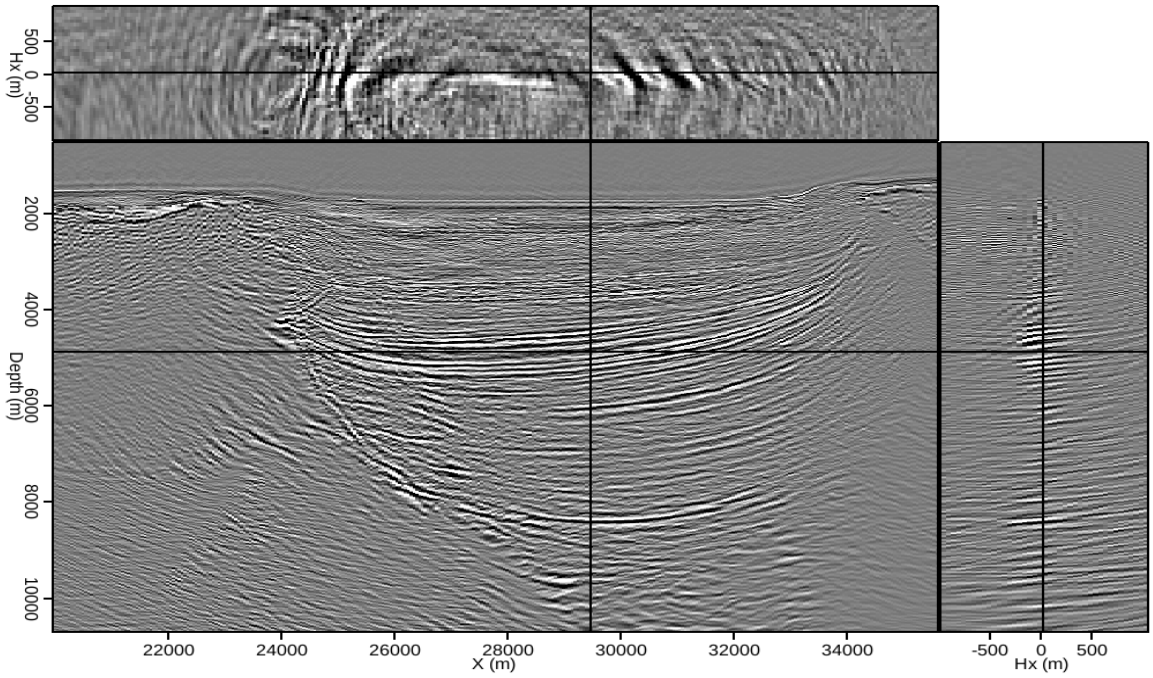
portion of this dataset, imaged using the velocity model provided with the data. To test the method, a single location along the reflector highlighted in Figure 1(b) was isolated and used to synthesize the source and receiver wavefields as described previously.

Figure 2 shows the results of imaging the new, synthesized wavefields with three different velocity models: the one provided with the data, and models scaled +/- 5% from the one provided. A qualitative examination suggests that the provided model produces the best-focused image, seen in panel (b). A more quantitative analysis using equation 5 confirms this; the F-values calculated for each of the images in Figure 2 are found in table 1.

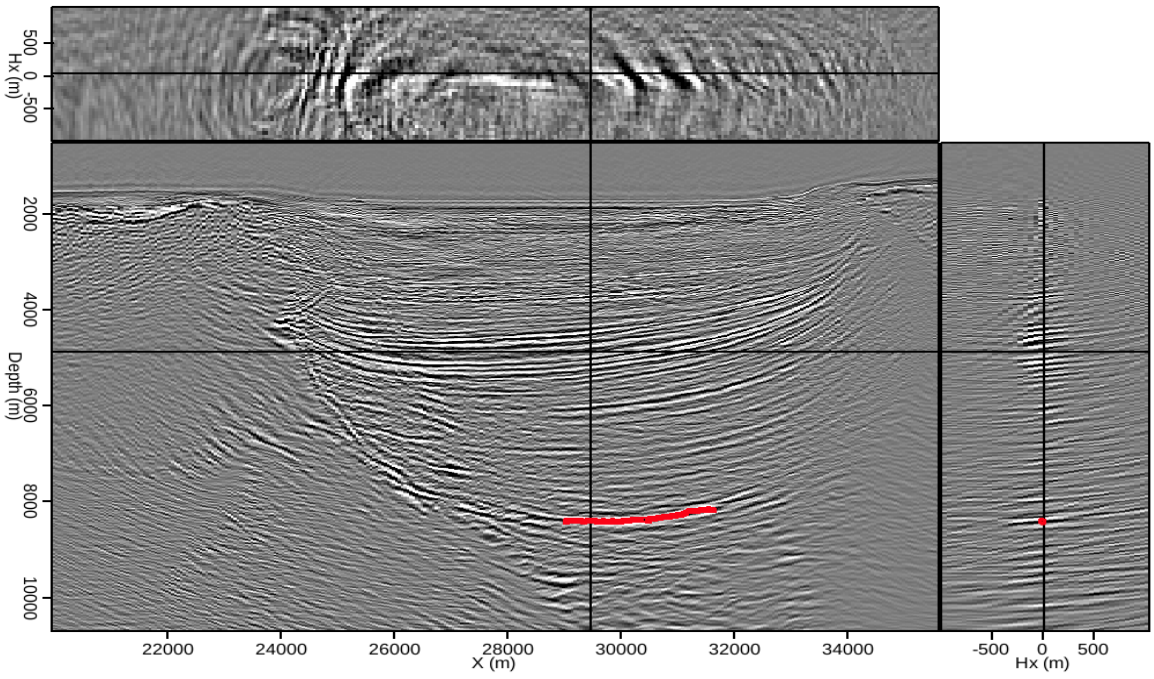
A more interesting test case is one in which the initial image was not generated using the provided velocity model. The image in Figure 3(a) was created using the slower velocity model of the previous example; a location isolated from the reflector indicated in Figure 3(b) was used to synthesize the new source and receiver wavefields. The same reflector is chosen in Figure 3(b) as in Figure 1(b), although the location is slightly shifted due to the difference in velocity models. Now, Figure 4 shows the results of imaging the synthesized wavefields from this initial image, with the same three models as the previous example. Again, the provided model (panel b) delivers the most well-focused result, both qualitatively and quantitatively (from the F-value results in table 2). This demonstrates that the method can identify a more accurate model using field data, even when the initial model is less accurate.

3D RESULTS

Finally, initial tests on 3D data were performed using the same strategy as in the previous section. Figure 5(a) shows a relatively small image cube derived from the same Gulf of Mexico dataset used for the 2D examples, and imaged using the provided velocity model. Again, a single location from the manually-interpreted reflector (this time, a 3D horizon) shown in Figure 5(b) was used to synthesize 3D areal source and Born-modeled receiver wavefields. Because the prestack image cubes resulting from the imaging of these wave-

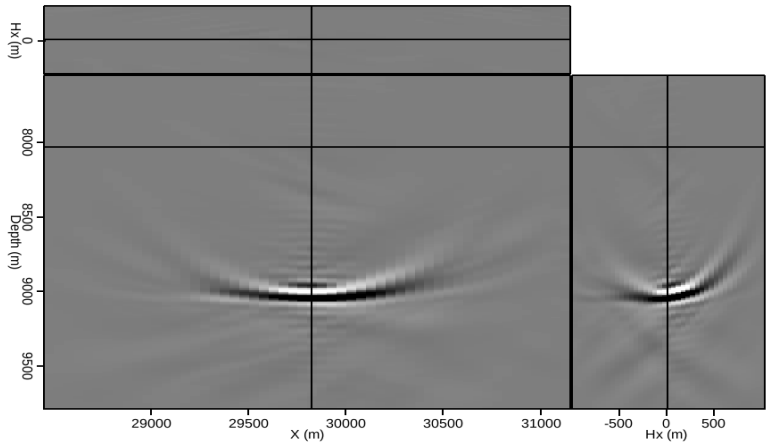


(a)

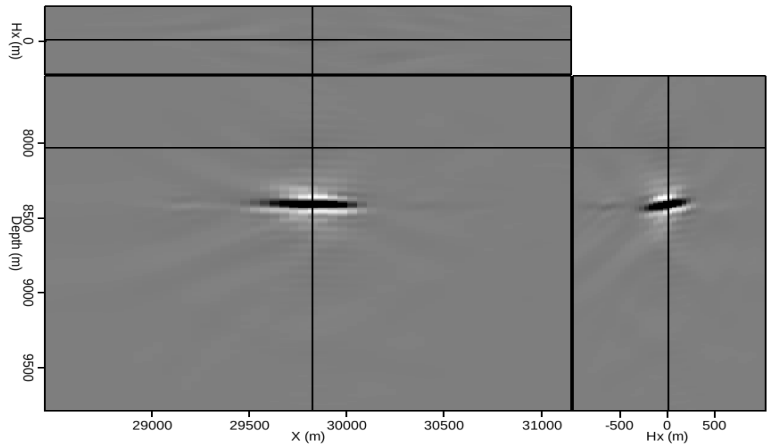


(b)

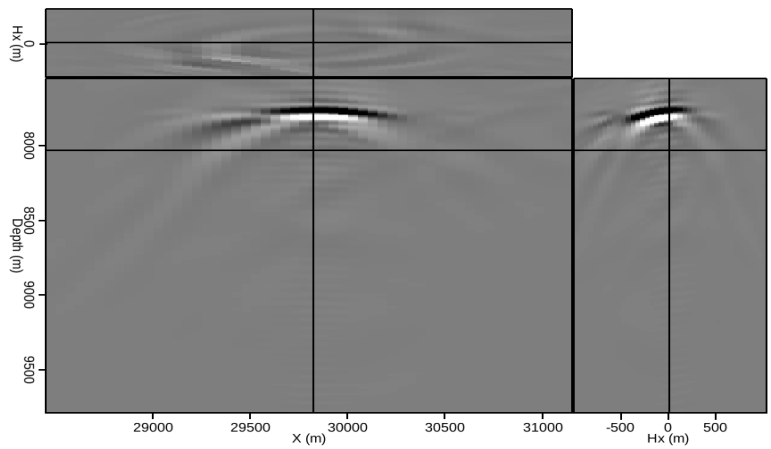
Figure 1: (a) A 2D field image produced using the provided velocity model, and (b) a manually-picked horizon of interest used to test three different velocity models. [**CR**] adam1/. 2d-field,act-picks



(a)

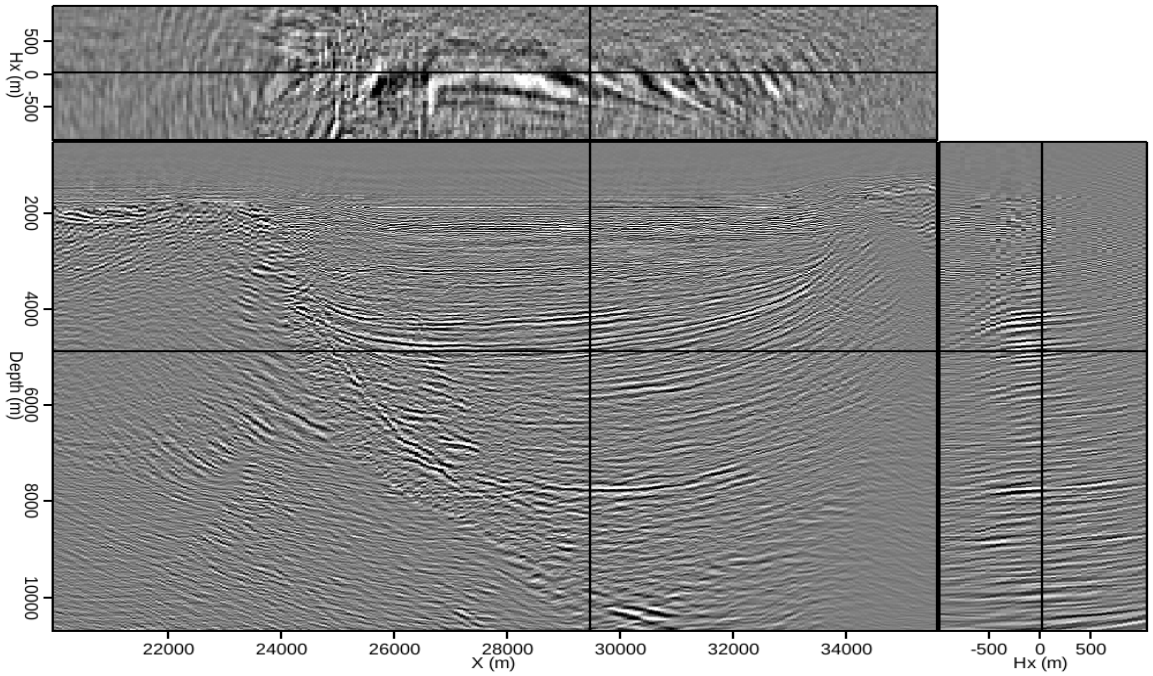


(b)

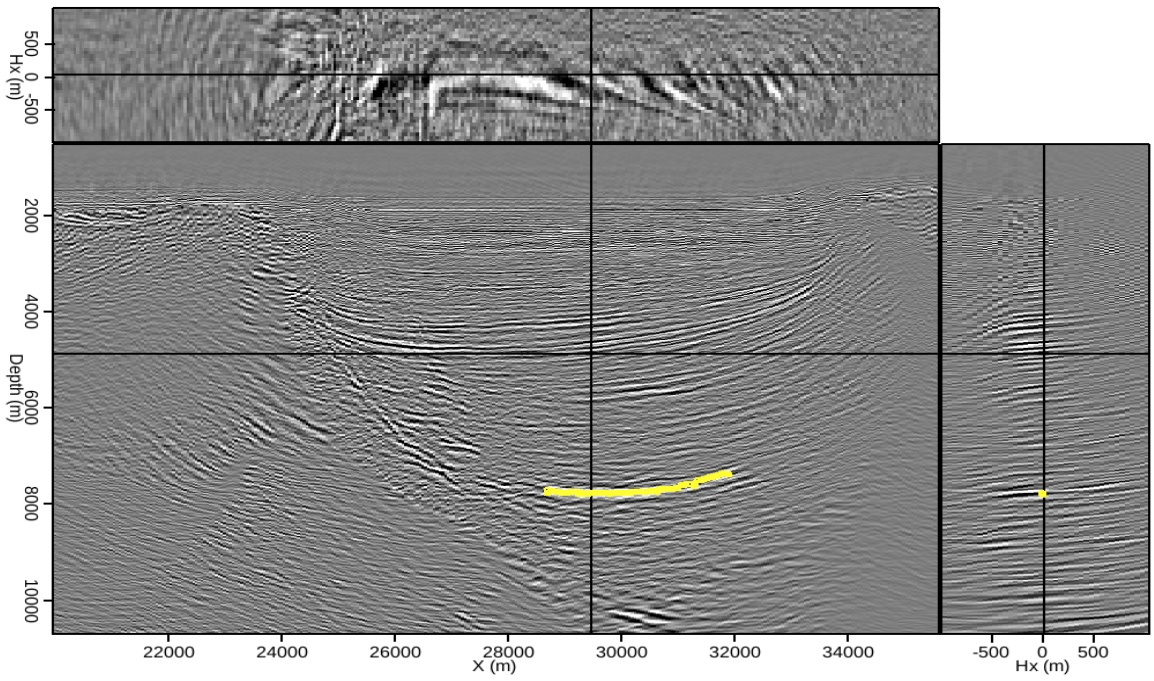


(c)

Figure 2: A single location from the reflector indicated in Figure 1(b), imaged using synthesized source and receiver wavefields and (a) a velocity model 5% faster than the one provided; (b) the provided velocity model; and (c) a model 5% slower than the one provided. [**CR**] adam1/. born-fast,born-act,born-slow

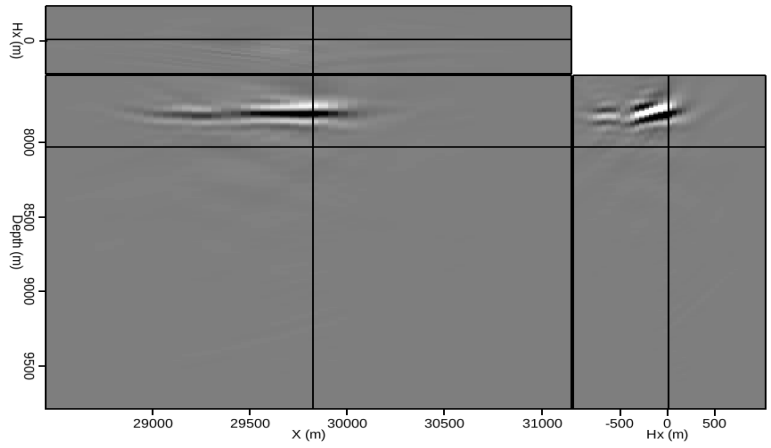


(a)

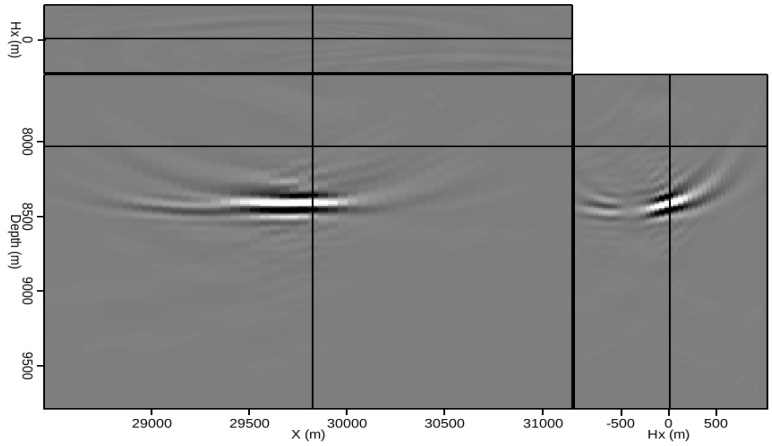


(b)

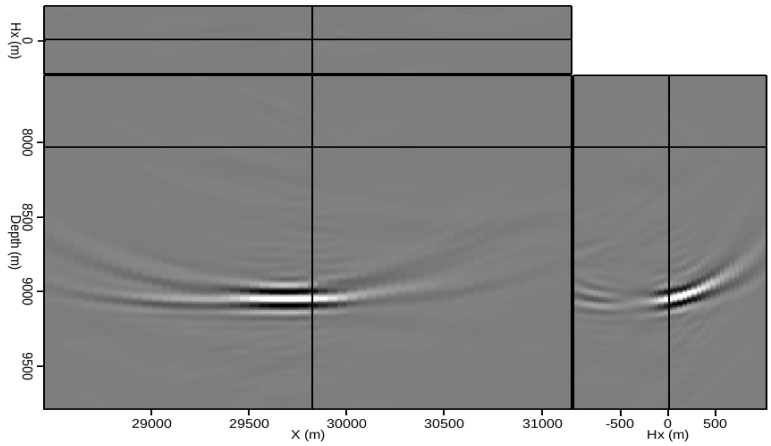
Figure 3: (a) A 2D field image produced using a velocity model 5% slower than the one provided, and (b) a manually-picked horizon of interest used to test three different velocity models. [**CR**] adam1/. 2d-slow,slow-picks



(a)



(b)



(c)

Figure 4: A single location from the reflector indicated in Figure 3(b), imaged using synthesized source and receiver wavefields and (a) a velocity model 5% slower than the one provided; (b) the provided velocity model; and (c) a model 5% faster than the one provided. [**CR**] adam1/. slow-slow,slow-act,slow-fast

Migration model	F value
Provided velocity	0.550
Fast model	0.514
Slow model	0.518

Table 3: Calculations from equation 5 for each 3D migration velocity model, after the initial image and synthesized wavefields were created using the velocity model provided with the data.

Migration model	F value
Provided velocity	0.481
Fast model	0.470
Slow model	0.466

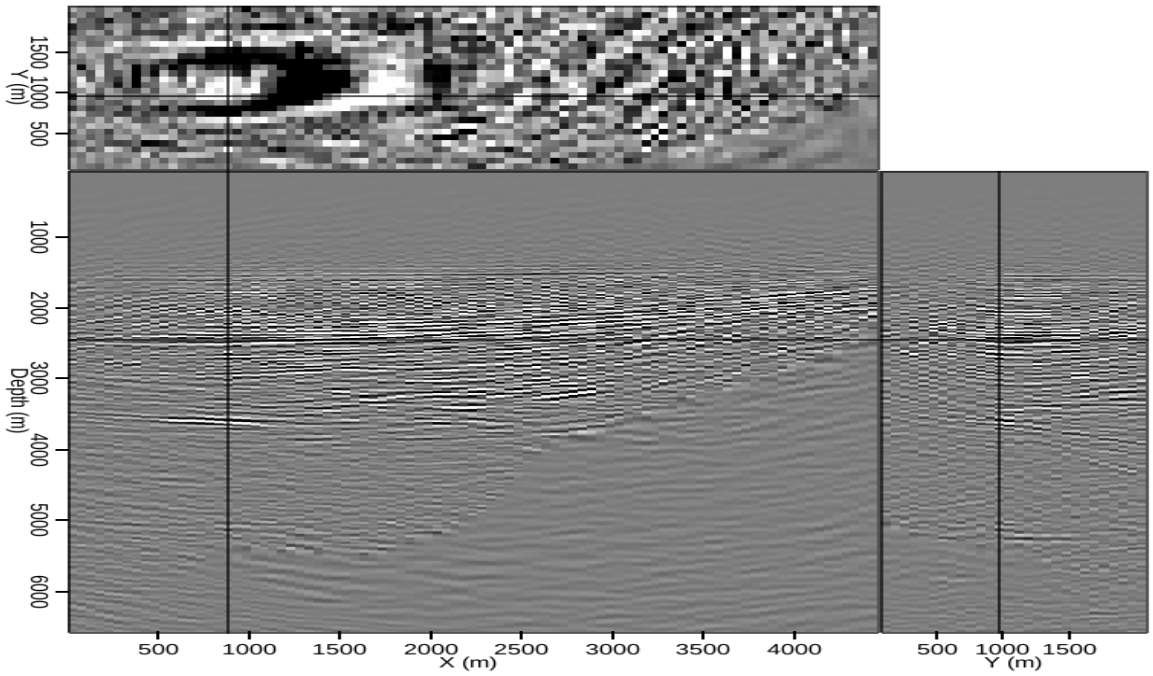
Table 4: Calculations from equation 5 for each 3D migration velocity model, after the initial image and synthesized wavefields were created using a velocity model 5% faster than the one provided with the data.

fields are five-dimensional, the results are displayed as two separate figures. Figure 6 shows images extracted from the cube at zero-subsurface offset in both the x and y directions, after migrating the synthesized wavefields with the three velocity models used for the 2D examples. Figure 7 shows the corresponding images in subsurface offset coordinates, extracted at the indicated locations in Figure 6. In this example, the image generated using the provided velocity model (panel (b)) appears better focused in both domains. The 3D F-value calculations in table 3 provide more quantitative evidence.

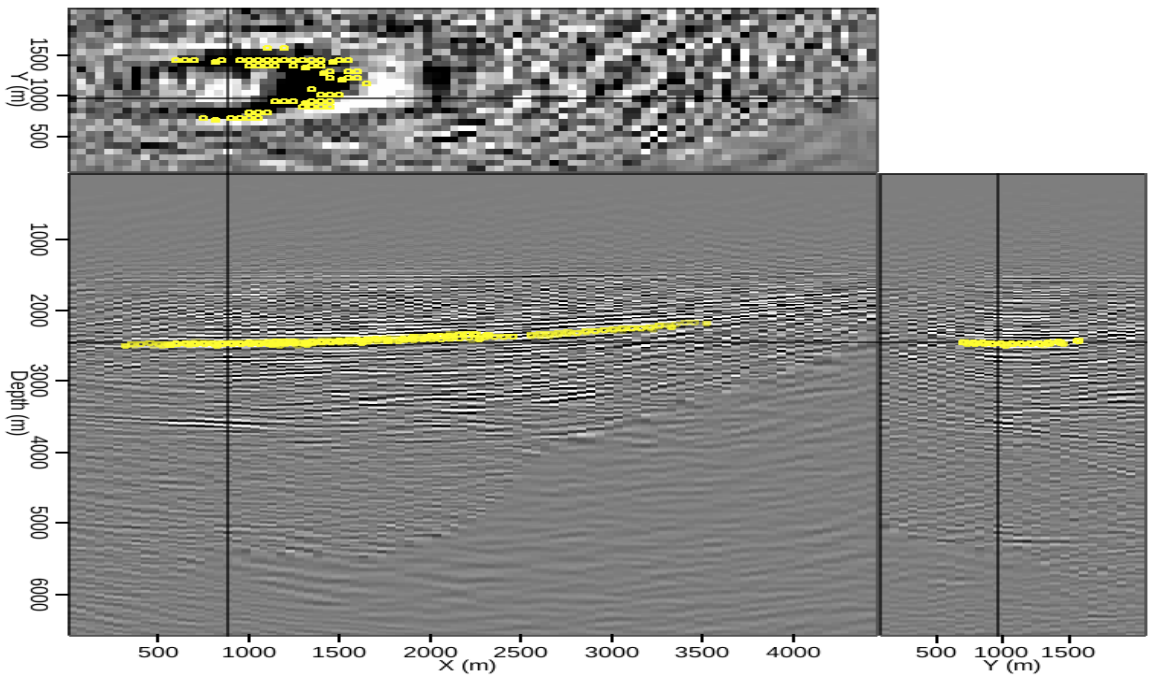
The more realistic case of an initial image created using the faster velocity model also yielded encouraging results. While it is slightly more difficult to qualitatively distinguish the best-focused model either at zero-subsurface offset (Figure 9) or in the subsurface offset domain (Figure 10), the calculations in table 4 reveal that the provided model is indeed superior. This example demonstrates the value of a quantitative image focusing measure, especially for difficult-to-visualize 5D volumes.

CONCLUSIONS

The strategy of synthesizing areal source and Born-modeled receiver wavefields to efficiently test a range of possible velocity models, previously demonstrated only on synthetic data, has been successfully tested on 2D and 3D field data. By imaging only a single shot, the method can identify the most accurate model, even if the wavefields were synthesized using an inaccurate model. When used in conjuction with other computational interpretation tools such as image segmentation, this method may help to alleviate model-building and interpretation bottlenecks, especially when several feasible salt scenarios are under consideration.

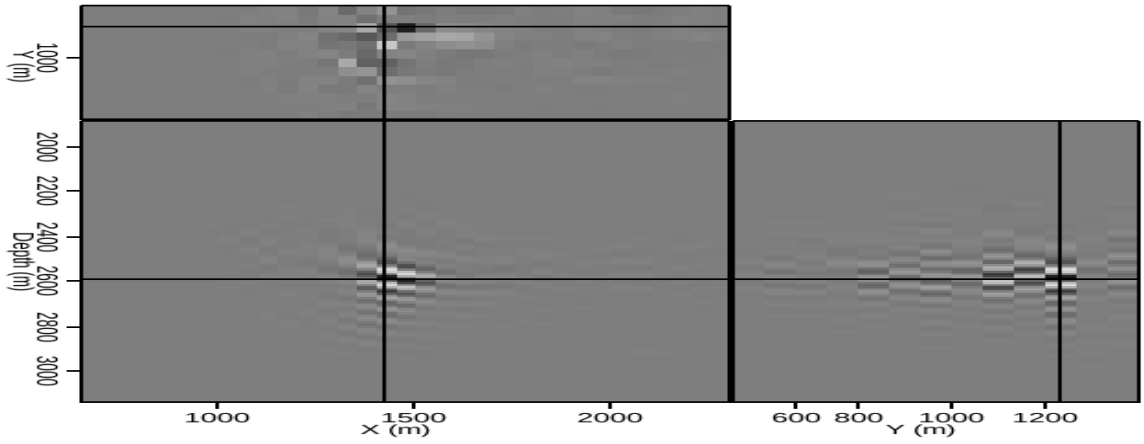


(a)

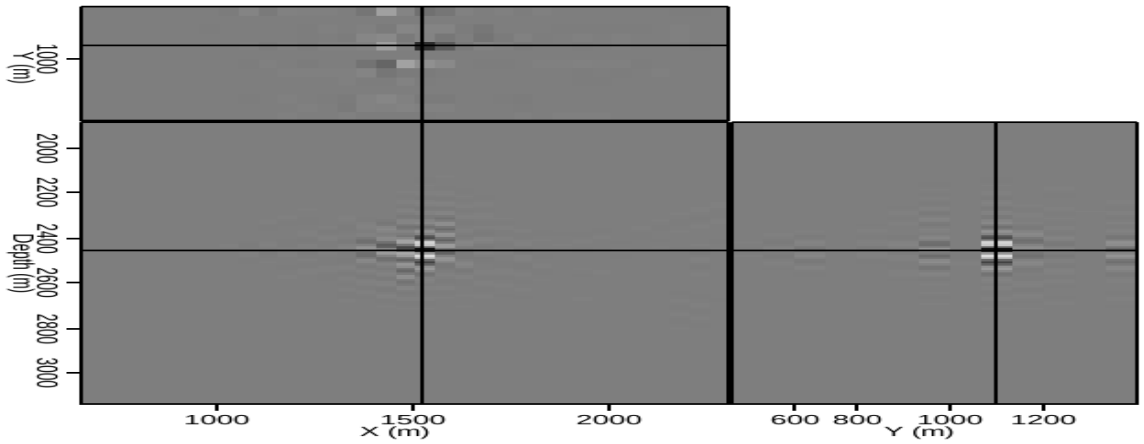


(b)

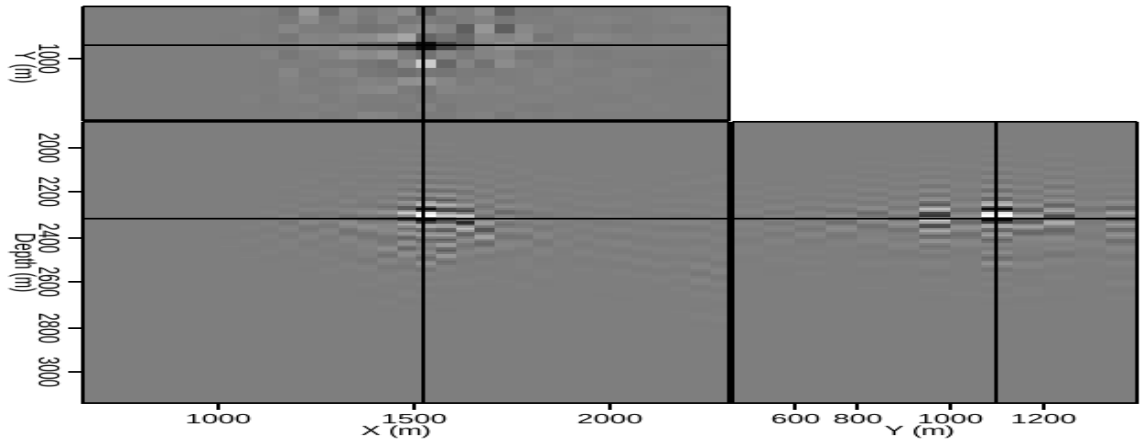
Figure 5: (a) A 3D field image produced using the provided velocity model, and (b) a manually-picked horizon of interest used to test three different velocity models. [**CR**] adam1/. 3d-field,3d-act-picks



(a)

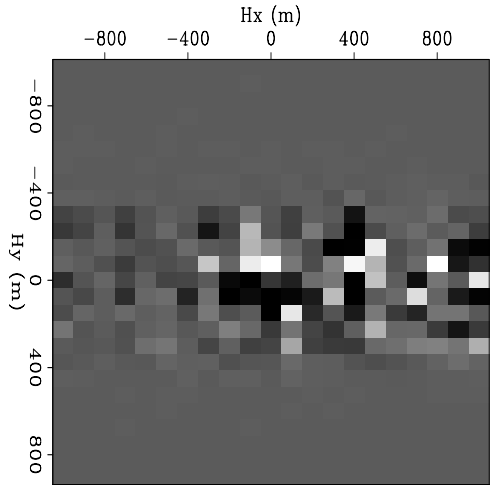


(b)

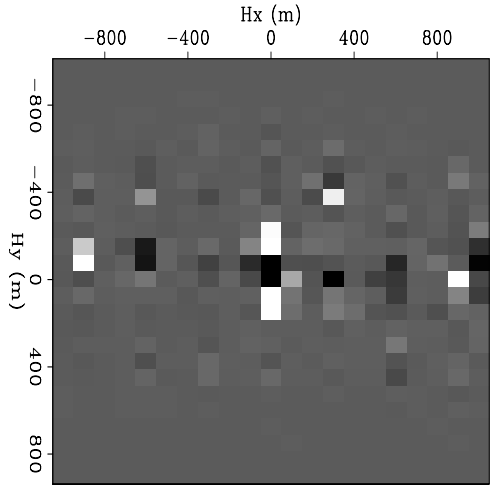


(c)

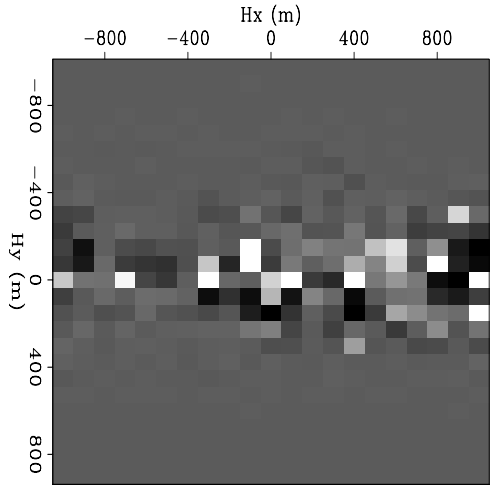
Figure 6: A single location from the reflector indicated in Figure 1(b), imaged using synthesized source and receiver wavefields and (a) a velocity model 5% faster than the one provided; (b) the provided velocity model; and (c) a model 5% slower than the one provided. [**CR**] adam1/. born-fast3d,born-act3d,born-slow3d



(a)

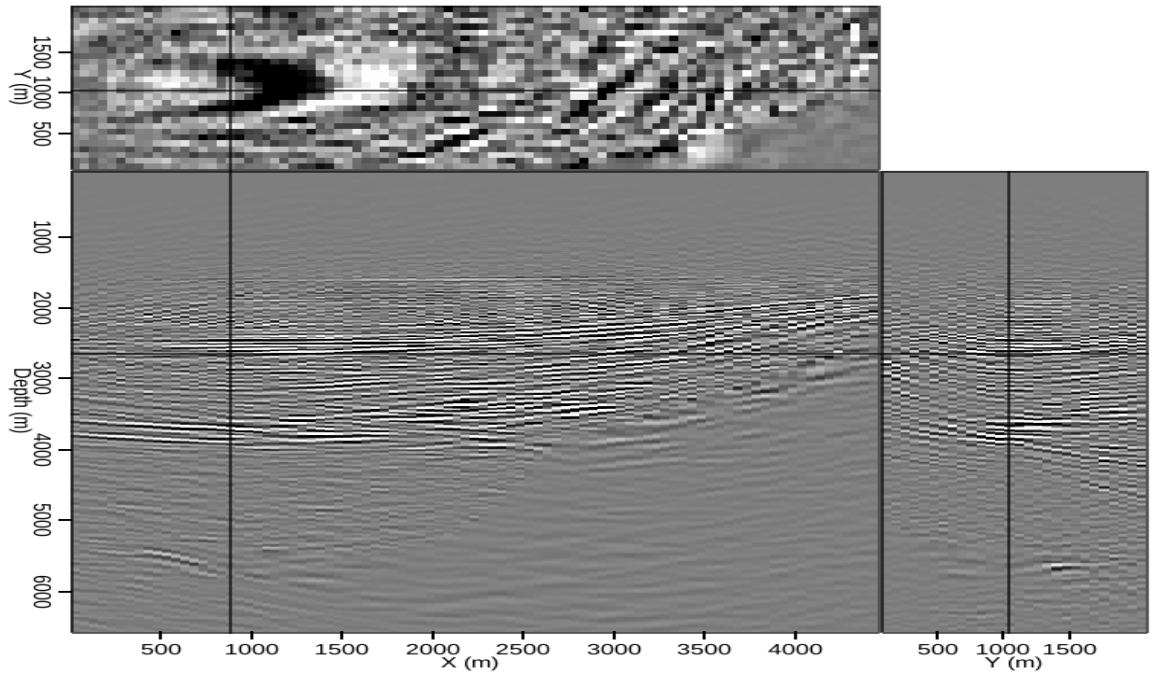


(b)

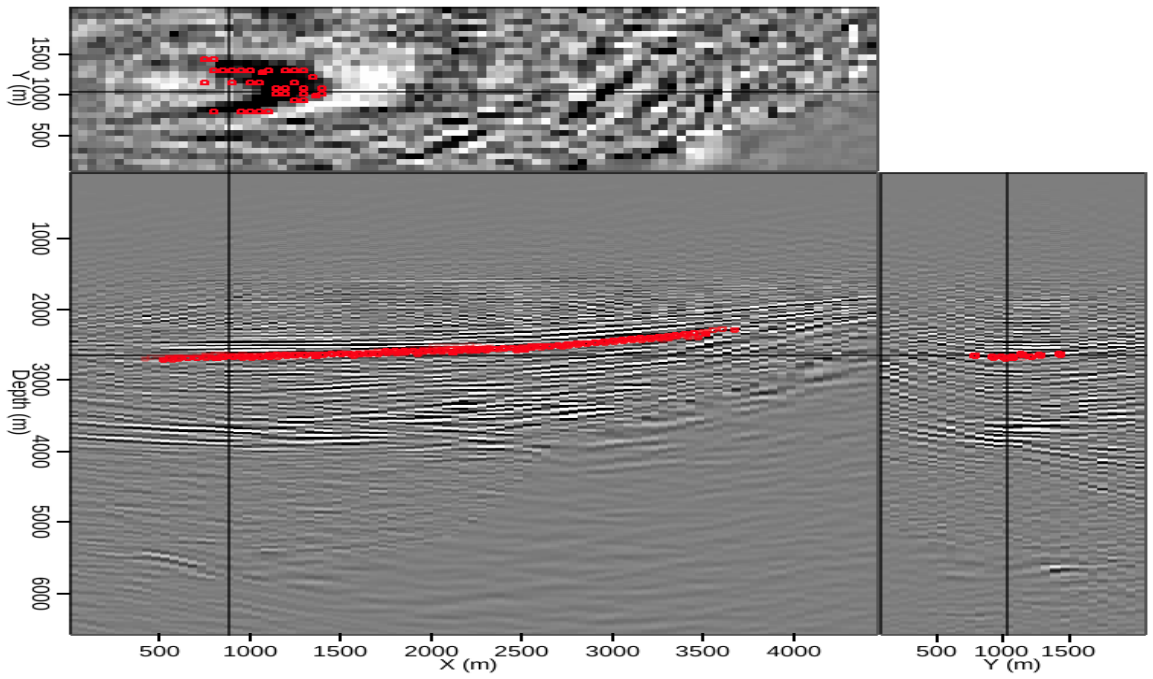


(c)

Figure 7: Subsurface offset images extracted from the indicated locations in Figure 6. The images were created using (a) the slower velocity; (b) the provided velocity; and (c) the faster velocity. [**CR**] adam1/. hxy-as,hxy-aa,hxy-af

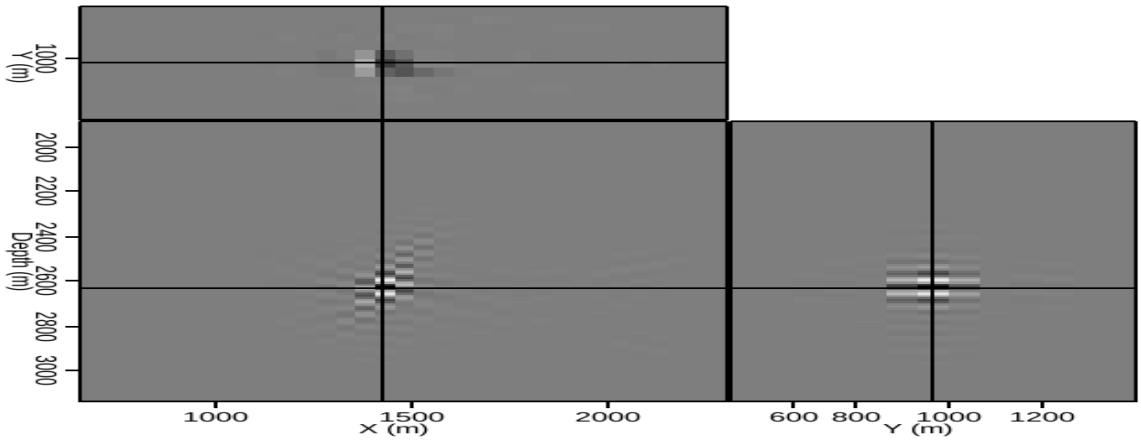


(a)

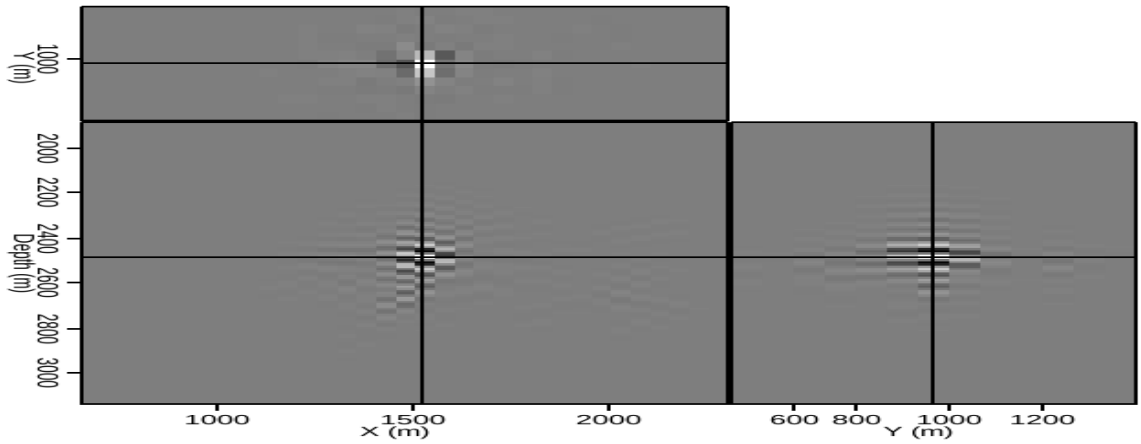


(b)

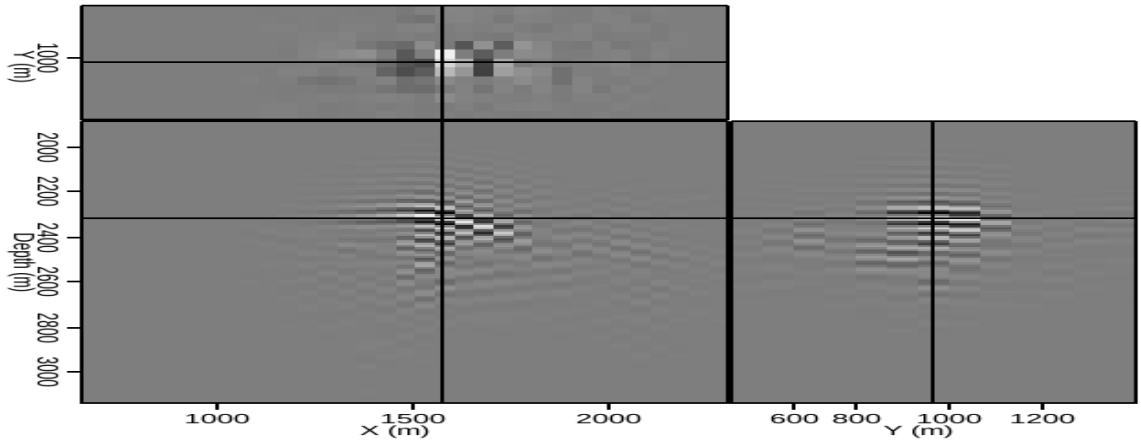
Figure 8: (a) A 3D field image produced using a velocity model 5% faster than the one provided, and (b) a manually-picked horizon of interest used to test three different velocity models. [**CR**] adam1/. 3d-fast,3d-fast-picks



(a)

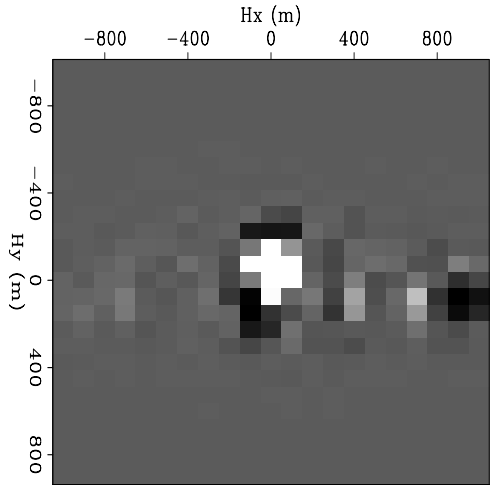


(b)

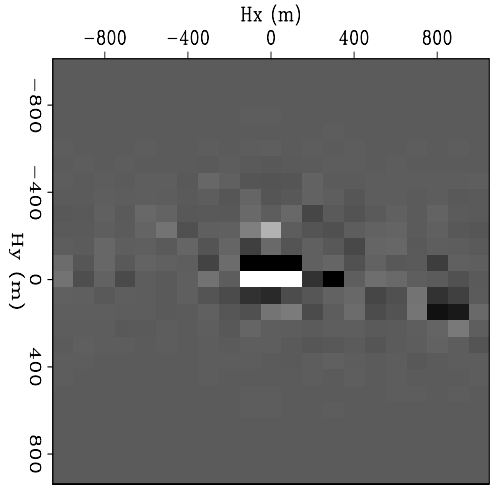


(c)

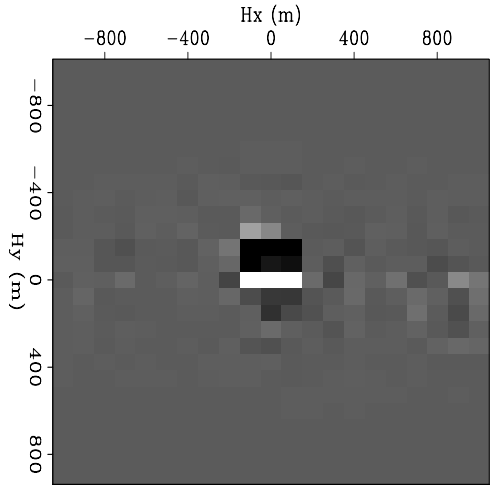
Figure 9: A single location from the reflector indicated in Figure 8(b), imaged using synthesized source and receiver wavefields and (a) a velocity model 5% faster than the one provided; (b) the provided velocity model; and (c) a model 5% slower than the one provided. [**CR**] adam1/. fast-fast,fast-act,fast-slow



(a)



(b)



(c)

Figure 10: Subsurface offset images extracted from the indicated locations in Figure 9. The images were created using (a) the faster velocity; (b) the provided velocity; and (c) the slower velocity. [**CR**] adam1/. hxy-ff,hxy-fa,hxy-fs

ACKNOWLEDGMENTS

I am grateful to WesternGeco for providing the field dataset used here for examples, Yaxun Tang for his contributions to the Born modeling theory and framework, and to all SEP sponsors for their support.

REFERENCES

Halpert, A., 2012, Fast velocity model evaluation with synthesized wavefields: SEP-Report, **147**, 39–48.

———, 2013, Salt delineation via interpreter-guided 3D seismic image segmentation: SEP-Report, **149**, 107–120.

Halpert, A. and Y. Tang, 2011, Velocity model evaluation through Born modeling and migration: a feasibility study: SEP-Report, **145**, 15–26.

Hill, N. R., 1990, Gaussian beam migration: Geophysics, **55**, 1416–1428.

Lomask, J., R. G. Clapp, and B. Biondi, 2007, Application of image segmentation to tracking 3D salt boundaries: Geophysics, **72**, P47–P56.

Stolt, R. H. and A. Benson, 1986, Seismic migration: Theory and practice: Geophysical Press.

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

Tang, Y. and B. Biondi, 2010, Target-oriented wavefield tomography using demigrated Born data: SEP-Report, **140**, 67–82.

Wang, B., J. Ji, C. Mason, S. Gajawada, and Y. Kim, 2008, Beam-based interactive imaging for salt interpretation and salt model building: SEG Technical Program Expanded Abstracts, **27**, 3073–3077.

Wang, B., C. Mason, K. Yoon, J. Ji, J. Cai, S. Suh, and Z. Li, 2011, Complex salt model building using a combination of interactive imaging and layer-stripping RTM: First Break, **29**, 47–54.

Handling salt reflection in Least-squares RTM with Salt-dimming

Mandy Wong

ABSTRACT

The challenge for applying least-squares reverse-time migration (LSRTM) to area with sharp velocity contrast is discussed. Least-squares migration (LSM), also known as linearized inversion, is an advanced imaging technique. It provides true relative amplitude information while suppressing acquisition footprints and migration artifacts. In most cases when the velocity is smoothly varying, the observed data can be used directly as input for LSRTM. However, when the velocity field has a sharp contrast like in the transition from sediment velocity of salt velocity, a background data term needs to be calculated and subtracted from the observed data before supplying as input for LSRTM. Therefore, special care is needed when performing LSM in regions like Gulf of Mexico where strong salt-reflection is presence. While straight forward with synthetic, subtracting the background data is a non-trivial problem for field data. I introduce a salt-dimming technique for handling sharp velocity contrast in LSM. I demonstrate the concept and methodology in 2D with a modified version of the Sigsbee2B model.

INTRODUCTION

In seismic exploration, geological interpretations require a subsurface image that accurately displays structural information. The conventional practice of obtaining such image is by prestack depth migration. However, depth migration images are often distorted by uneven subsurface illumination from acquisition geometry, complex overburden and band-limited wavefields. To correct the effects of uneven illumination, the imaging problem can be posed as a linear inverse problem. Instead of using the adjoint operator, I use the pseudo-inverse of the Born-modeling operator to optimally reconstruct the reflectivity. This inversion-based imaging method is also widely known as least-squares migration (LSM) (Lambare et al., 1992; Nemeth et al., 1999) or linearized wavefield inversion (Clapp, 2005; Valenciano, 2008).

LSM can be implemented using different types of wave propagation operators. Starting with the ray-based operator in least-squares Kirchoff migration. Kuehl and Sacchi (2003) and Clapp (2005) have worked on LSM with the one-way operator. Until recently, least-squares reverse-time migration (LSRTM), which uses the two-way propagator, becomes computationally affordable (Dai et al., 2010; Wong et al., 2011; Yao and Jakubowicz, 2012). Two-way propagator is considered the most accurate way of representing wave propagation because it allows bandlimited wave-propagation in all direction with no dip limitation. When LSM is applied to regions like Gulf of Mexico where there is a sharp velocity contrast, a background data term needs to be calculated and subtracted from the observed data. I will show by synthetic example that this subtle point affects LSM result. When the background data is properly subtracted from the observed data.

THEORY

I define my model ($m(\mathbf{x})$) to be a weighted difference between the migration slowness ($s_o(\mathbf{x})$) and the true slowness ($s(\mathbf{x})$):

$$m(\mathbf{x}) = (s(\mathbf{x}) - s_o(\mathbf{x}))s_o(\mathbf{x}) \tag{1}$$

The model ($m(\mathbf{x})$) is directly proportional to the perturbation (or deviation) between the migration slowness and the true slowness. In the synthetic case, this quantity should be zero if the migration velocity is exactly the true velocity. In LSM, the forward modeling operator is linearized with respect to $m(\mathbf{x})$.

$$d^{mod} = F(s_o^2 + m) \approx F(s_o^2) + \mathbf{L}m \tag{2}$$

where $\mathbf{L}$ is the linearized forward-modeling operator and d^{mod} is the synthetic data. The equation for the linearized forward modeling is:

$$\mathbf{L}m = \Delta d(\mathbf{x}_r, \mathbf{x}_s) = \sum_{\mathbf{x}} \omega^2 f_s(\omega) G(\mathbf{x}_s, \mathbf{x}) m(\mathbf{x}) G(\mathbf{x}, \mathbf{x}_r) \tag{3}$$

where ω is the temporal frequency, $\mathbf{x}$ is the image point, $f_s(\omega)$ is the source waveform, and $G(\mathbf{x}_s, \mathbf{x})$ is the Green function of the two-way acoustic constant-density wave equation over the migration slowness s_o. (Note that G is actually ω-dependent.) It is important to point out that the adjoint of the linearized forward-modeling operator is the migration operator:

$$\mathbf{m}_{mig}(\mathbf{x}) = \sum_{\mathbf{x}_r, \mathbf{x}_s} \omega^2 f_s^*(\omega) G^*(\mathbf{x}_s, \mathbf{x}) G^*(\mathbf{x}, \mathbf{x}_r) d(\mathbf{x}_r, \mathbf{x}_s) \tag{4}$$

The objective function ($S(m)$) of LSM is defined by minimizing the least-squares difference between the synthetic (d^{mod}) and the observed data (d^{obs}):

$$S(\mathbf{m}) = \|d^{mod} - d^{obs}\|^2 = \|\mathbf{L}m - (d^{obs} - F(s_o^2))\|^2 \tag{5}$$

Notice in the above objection function, there is a background data term $F(s_o^2)$ that needs to be subtracted from the observed data. Essentially, $F(s_o^2)$ is the forward modeling using the migration slowness s_o. In most cases when the velocity is smoothly varying, this term is trivial and we can just ignore this term in the inversion.

However, when the velocity field has a sharp contrast, this term is non-trivial. Therefore, special care is needed when performing LSM in regions like Gulf of Mexico where strong salt-reflection is presence. I will illustrate this point with a simple synthetic model in the next section.

LSRTM with and without background data subtraction

I use a simple 2D synthetic model with a salt structure and two subsalt reflectors (Figure 1(a)). The contrast from the sediment to the salt is sharp and the velocity transition across the two subsalt reflectors is weak. In a typical model-building flow, the salt structure will be included in the migration slowness (Figure 1(b)). Therefore, the background data term ($F(s_o^2)$) is non-trivial. This simple example serves as a good exercise to test the importance of the background data term in LSM.

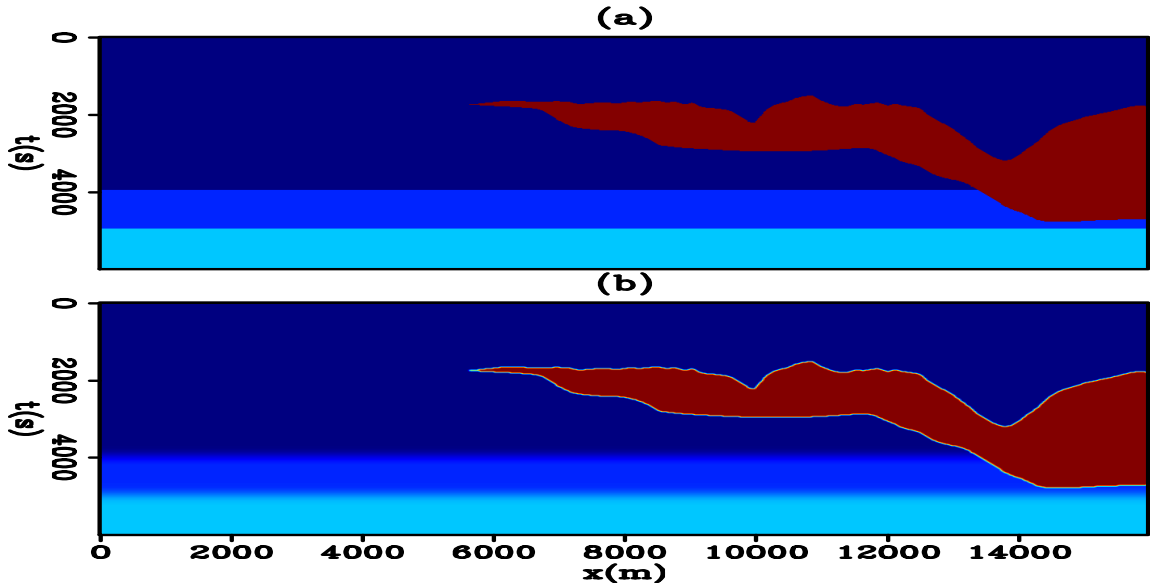


Figure 1: (a) True velocity model and (b) migration velocity model. Notice how the salt boundary is slightly smoothed in the migration velcoity model. This is to simulate the situation where salt reflection is not correctly estimated. mandy1/. Vfigvel

I will compare two LSRTM inversions, with the following objective functions:

$$S_1(\mathbf{m}) = \|\mathbf{L}m - d^{obs}\|^2 \tag{6}$$
$$S_2(\mathbf{m}) = \|\mathbf{L}m - (d^{obs} - F(s_o^2))\|^2 \tag{7}$$

In the first inversion as shown by the objective function $S_1(m)$, I try to fit the observed data with the linearized synthetic data ($\mathbf{L}m$). That is, I ignore the background data term. Figure 2(a) shows the RTM result or equivalently, the first iteration output of LSRTM. In this model, there are two subsalt reflectors located at depth 4000m and 5000m. From the true model (figure 1(a)), I expect the impedance across the two subsalt reflectors to be constant. Figure 3(a) shows the amplitude along these two reflectors from the RTM image. Blue corresponds to the shallower reflector at $z = 4000m$ while the red corresponds to the reflector at $z = 6000m$. Due to uneven illumination and the complex salt structure, the amplitude along the reflectors is not constant. In the ideal case, we want to see two horizontal lines in Figure 3(a). By using LSRTM, we hope to correct the affect of uneven illumination. Figure 2(b) shows the result of LSRTM at iteration 40 using the objective function $S_1(m)$. Figure 3(b) shows the corresponding subsalt reflector amplitude, notice that the amplitude at the shadow zone of the two subsalt reflectors is not well compensated by the inversion. On the other hand, using the objective function $S_2(m)$, the amplitude along the two subsalt reflectors is better corrected by LSRTM as shown in Figure 2(c) and 3(c).

This example shows that subtracting the background data is crucial in correcting for uneven illumination for the subsalt reflectors. The next question is how to apply it in field data? It is actually a difficult task. The theory that generates the background data $F(s_o^2)$ is just an approximation to the complex earth mechanism that generates the observed data

$F(s^2)$. That makes obeying equation $S_2(\mathbf{m})$ impossible.

Let try to understand the inversion in a different way. Figure 4 (a) and (b) shows the square of one common receiver gather data-residual at iteration 1 and at iteration 40 using objective function $S_2(m)$. By looking at the square of the data residual, we get a sense of the highest value in the data-residual space This also tells us where (in the data space) the inversion tries to minimize. Between iteration 1 and 40, the highest value still remains to be the salt reflections. This means that the inversion is still predominantly trying to c̈orrectön the salt-reflector. I will introduce a simple 'method' to address this issue in the next section.

How to subtract background data in field datasets

One way to work around this problem is by salt-dimming. Salt-dimming aims to down-weight the salt reflection energy in the data space so that the inversion can try to minimize other region in the model like the subsalt reflectors. This corresponds to the following objective function.

$$S_3(\mathbf{m}) = \|W_s(\mathbf{L}m - d^{obs})\|^2.$$

W_s is the data weighting function that down-weight the salt reflection energy. This can be done by forward modeling the salt reflection using the migration velocity $F(s_o^2)$. Next is to calculate an envelope around the salt energy. The data weighting function can then by derived by assigning small value to the salt reflection envelope. Figure 5(a) shows the forward modeling of one CRG. The salt reflection is then used to derive an envelope region to be down-weighted. The resulting weighting function is shown in Figure 5(b). As shown in Figure 5(b) the down-weighted region (blue) corresponds mostly to salt reflector. It is important to point out that those blue regions are not muted. Instead, it is given a very small value of 0.01. By applying salt dimming, I guide the inversion to focus on other parts of the model (like subsalt reflectors) instead of the salt reflection itself. This data weighting function can be used for LSRTM using objective function $S_3(m)$.

Figure 3(d) shows the result of using salt-dimming with observed data. Notice in figure 4(d), the amplitude of the shadow zone along the subsalt reflectors is better recovered than in figure 4(c). Figure 4(c) and 4(d) shows the data residual squared plot at iteration 1 and iteration 40 of LSRTM with salt-dimming. Notice how the salt-dimming part completely changes the highest value in the least-square residual.

DISCUSSION

Convergence plot

Figure 6 shows the convergence curve for all three inversions. The residual is normalized by the residual value from the first iteration. In the $S_1(\mathbf{m})$ inversion, the background data is ignored. The inversion fails to properly correct for the uneven illumination (figure 3(b)) and this is supported by the fact that the residual barely goes down. (the diamond-blue curve in figure 6). When LSM is handled properly by using $S_2(\mathbf{m})$, the residual drops down to 50 percents over 200 iterations as shown by the square-red curve. Finally, salt-dimming

is applied using the observed data and the objective function $S_3(\mathbf{m})$, the residual drops steadily over iterations as indicated by the triangle-green curve.

For the $S_2(\mathbf{m})$ inversion, it is surprising to see that the residual doesn't drop down rapidly to zero. One reason for this is because linearization requires that the perturbation $(s(\mathbf{x}) - s_o(\mathbf{x}))$ perfect salt

Why salt-dimming works?

In standard GOM model building, the salt-structure is usually picked and included in the velocity output. The background data contain salt reflection energy. Ideally, we want to assume that we picked the salt correctly. That means the salt reflection energy in the perturbed data should disappear completely. Salt-dimming works because it simulates the same affect by making the amplitude of salt-reflection energy small. Alternatively, one might consider muting out all the salt reflection in the data space completely. However, I found that it is less effective than salt dimming.

Why salt-dimming is better than salt-muting?

Although I want all the salt-reflection energy to disappear in the observed data, it is impossible to achieve this by muting out all the salt reflection in the data-space. This is because the salt-reflection energy in the data space becomes progressively complex as time increases. Ultimately, some of the salt-reflection energy remains in the data. When LSM is performed on such a dataset, the inversion becomes problematic because only partial information about the salt structure is presence in the dataset. In addition, many truncations in the data space (as a result of the muting) create artifacts in the image space. The cause of such artifacts is similar in logic to that of having acquisition footprint in migration images. Salt-dimming does not have the same problem. Essentially all the information about the salt-structure still exists. Even when the inversion tries to recover some of those salt structures, there are still information to reconstruct it in the dataset.

CONCLUSION

In regions like Gulf of Mexico where there is a sharp velocity contrast, a background data term needs to be calculated and subtracted from the observed data. While this is a straight forward task in the synthetic case, it is challenging in the field data case. Synthetic examples show that salt-dimming in LSRTM can correct amplitude information for subsalt reflectors much better than conventional LSRTM. Salt-dimming is a viable solution for addressing the issue of subtracting background data in field datasets.

ACKNOWLEDGMENTS

I thank the sponsors of the Stanford Exploration Project for their financial support.

REFERENCES

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

Dai, W., C. Boonyasiriwat, and G. T. Schuster, 2010, 3d multi-source least-squares reverse time migration: SEG Expanded Abstracts, 3120–3124.

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

Lambare, G., J. Virieux, R. Madariaga, and S. Jin, 1992, Iterative asymptotic inversion in the acoustic approximation: Geophysics, **57**, 1138–1154.

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

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

Wong, M., S. Ronen, and B. Biondi, 2011, Least-squares reverse time migration/inversion for ocean bottom data: A case study: SEG Expanded Abstracts, **30**, 2369–2373.

Yao, G. and H. Jakubowicz, 2012, Least-squares reverse-time migration: SEG Expanded Abstracts, 1–5.

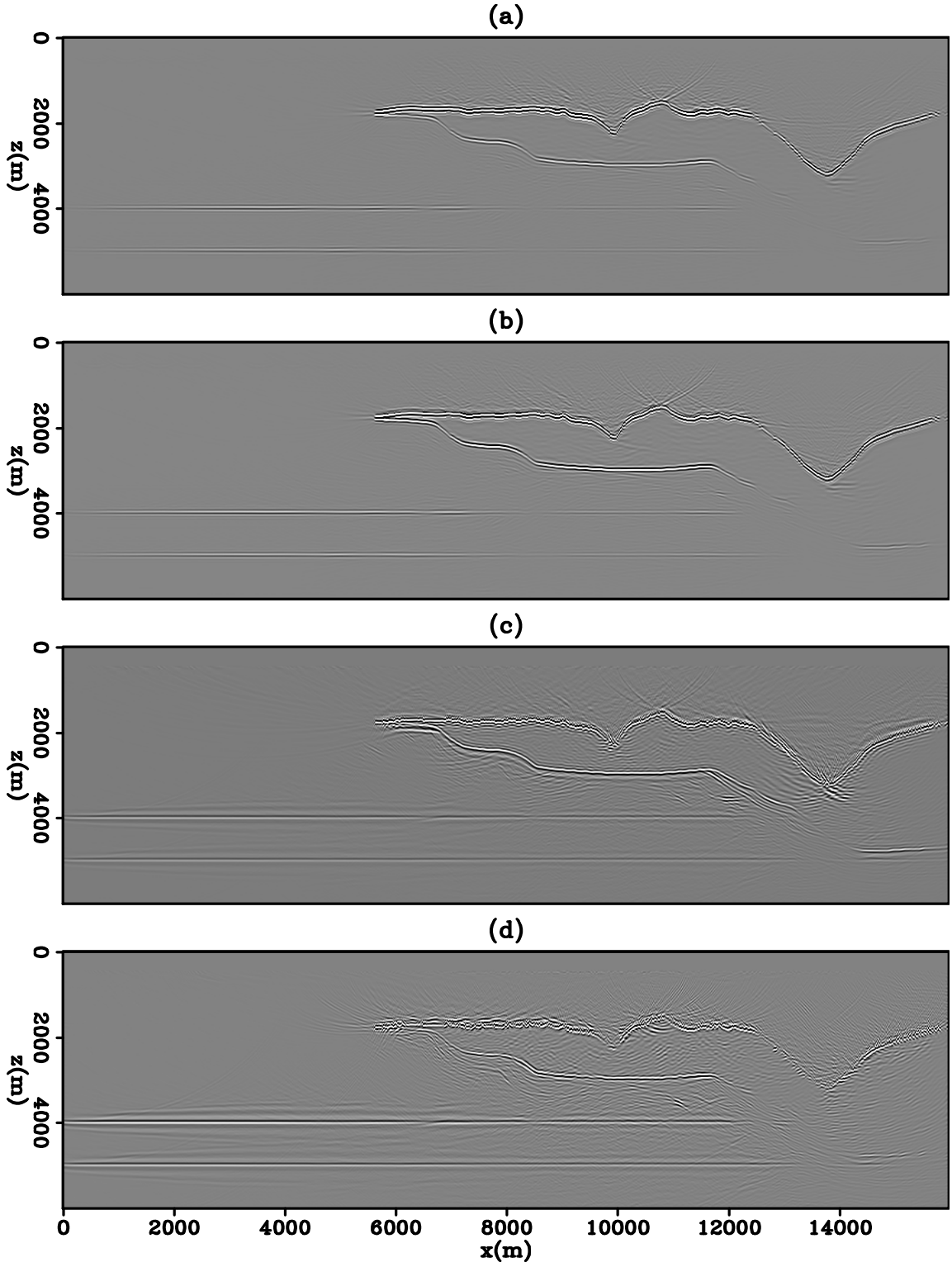


Figure 2: (a) RTM of the simple sythetic, (b) LSRTM using the total observed data (objective function $S_1(\mathbf{m})$), (c) LSRTM using perturbed data (objective function $S_2(\mathbf{m})$), (d) LSRTM using observed data with salt dimming (objective function $S_3(\mathbf{m})$. mandy1/. Vfig2img

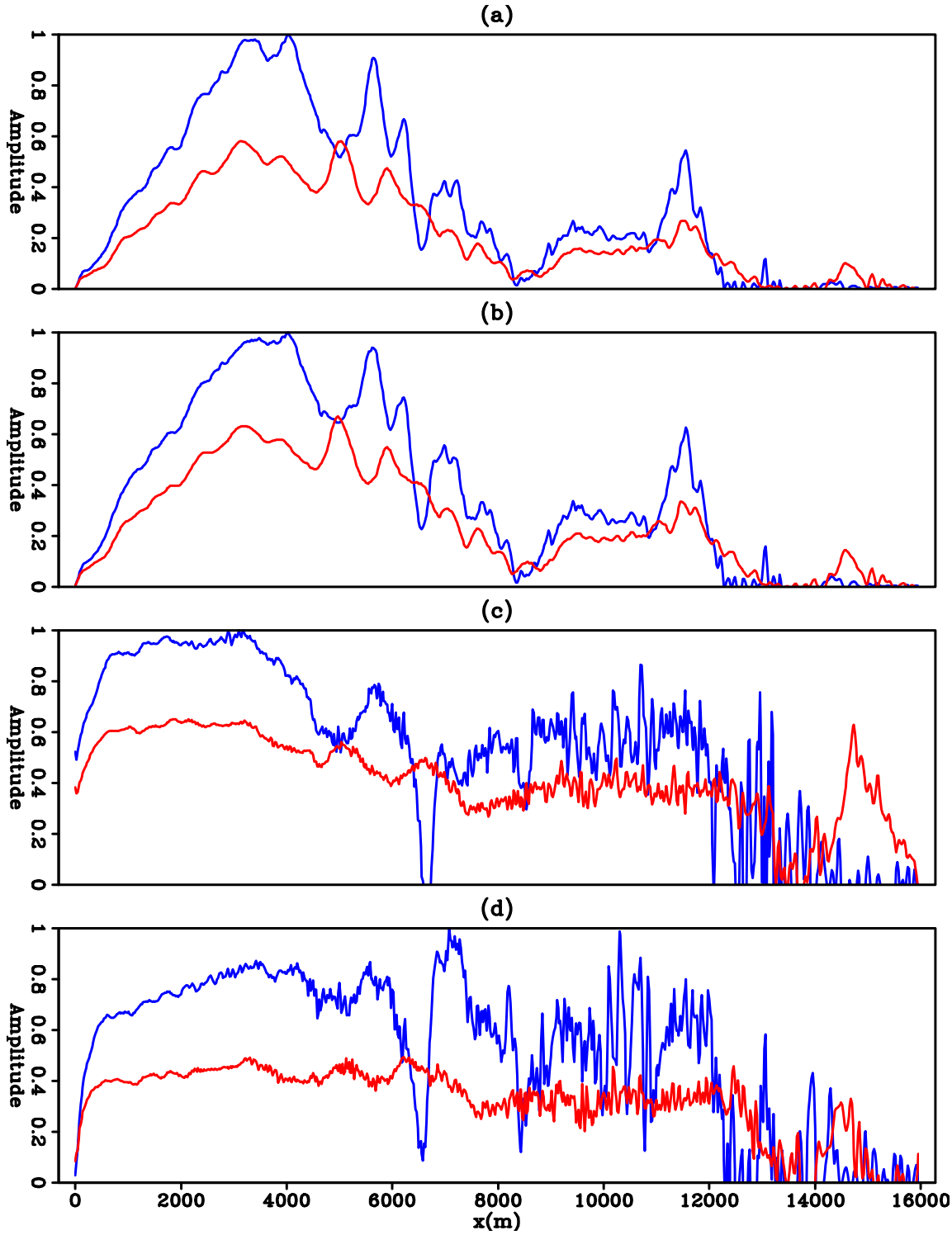


Figure 3: Amplitude along two subsalt reflectors at 4000m and 5000m. (a) RTM, (b) LSRTM using the total observed data (objective function $S_1(\mathbf{m})$), (c) LSRTM using perturbed data (objective function $S_2(\mathbf{m})$), (d) LSRTM using observed data with salt dimming (objective function $S_3(\mathbf{m})$. Blue corresponds to the shallower reflector at $z = 4000m$ while the red corresponds to the reflector at $z = 6000m$. The last case (d) gives the best recovery of uneven illumination. mandy1/. Vfig3img

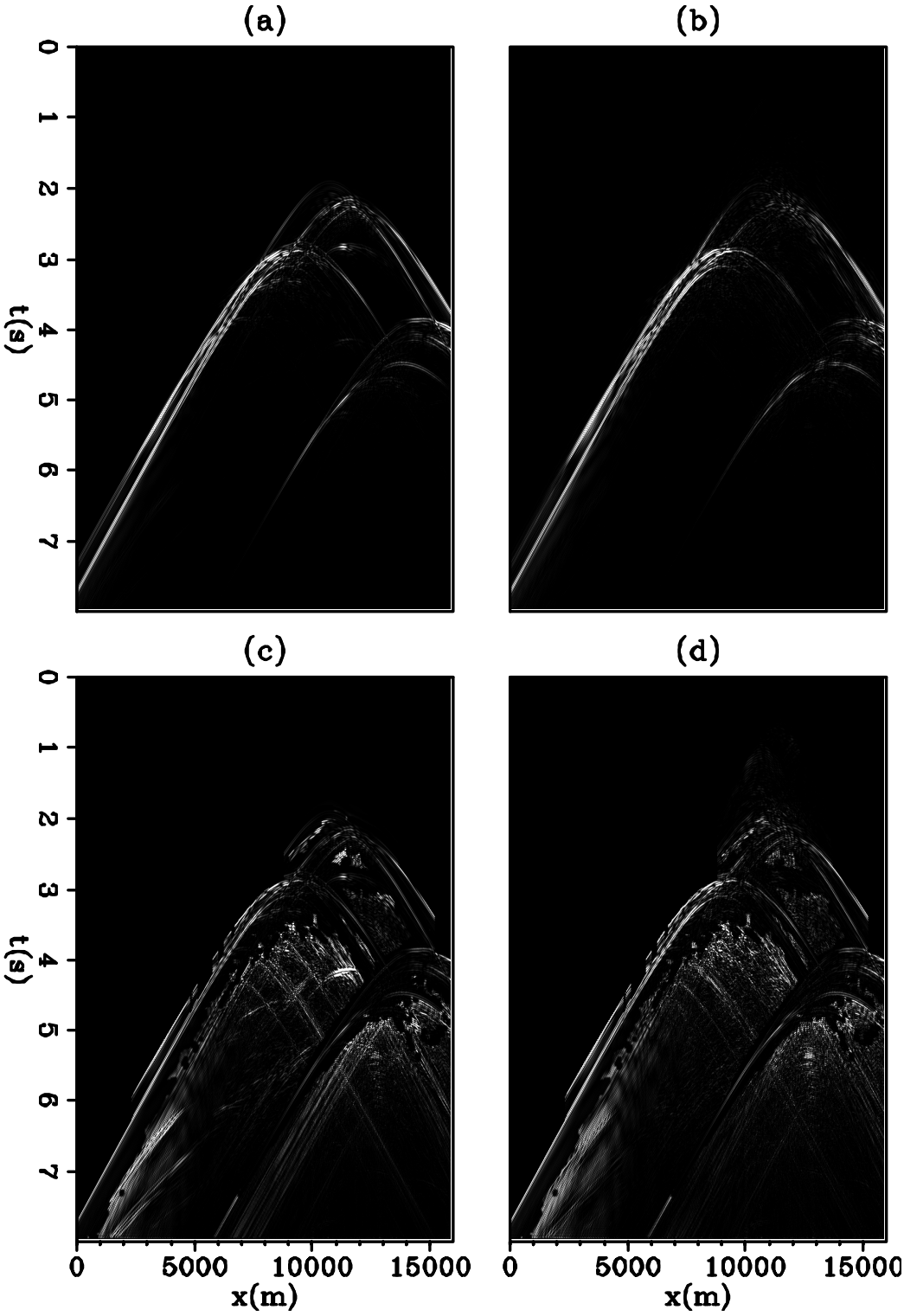


Figure 4: The square of a common receiver gather. It gives a sense of where the inversion is biased towards. (a) Original data (b) LSRTM iteration 20 using perturbed data. (c) LSRTM iteration 1 (d) LSRTM iteration 20 with salt-dimming mandy1/. Vfig4resid

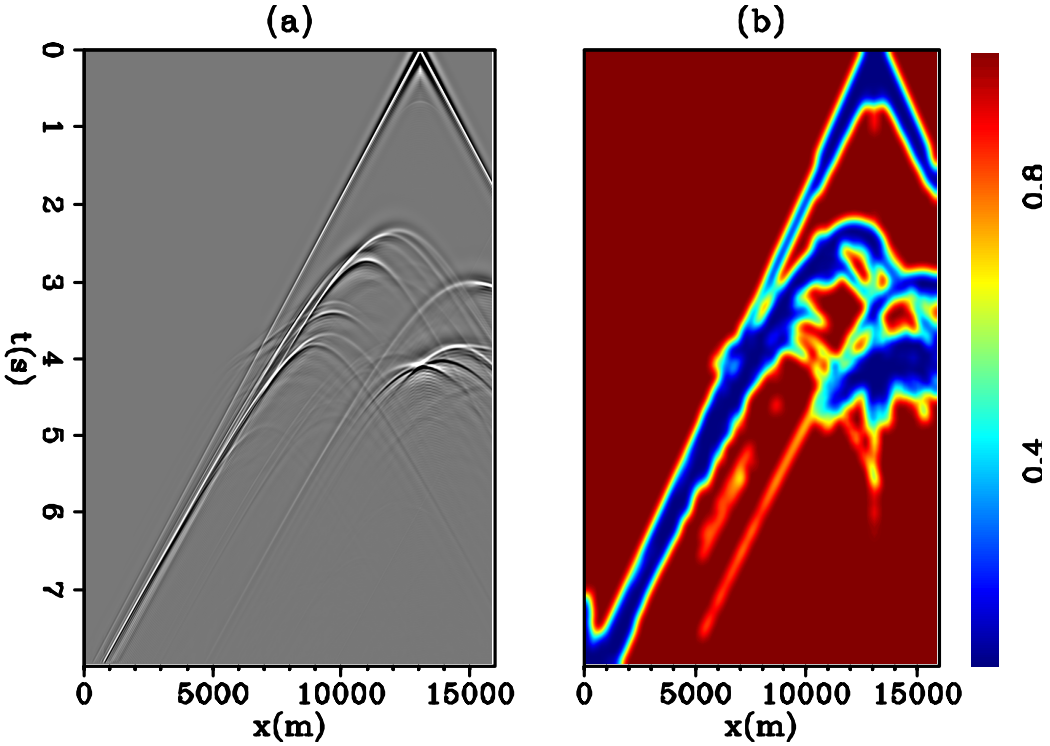


Figure 5: (a) Background data ($F(s_o^2)$) created by forward modeling with the migration slowness and (b) salt-dimming weight (W_s) generated with the background data. mandy1/. Vfig5wgt

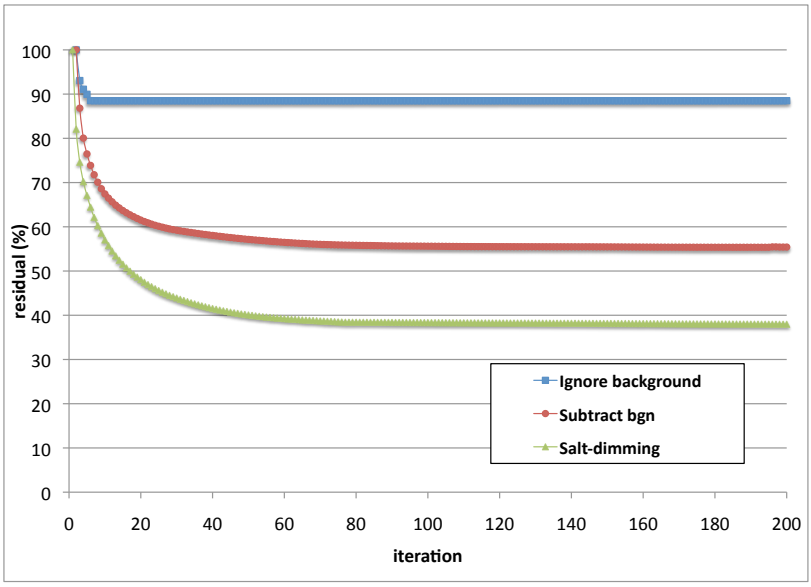


Figure 6: Convergence curve of different cases. Notice that when background data is ignored, the inversion hardly converges. When the problem is handled properly like the case of using the perturbed data, the inversion residual drops down to 50 percents within 200 iterations. Finally, using salt-dimming with the observed data, the residual also drops steadily over iterations. mandy1/. Vfig6

Least-squares RTM with wavefield decomposition

Mandy Wong

ABSTRACT

In least-squares reverse-time migration, the adjoint of the linearized forward-modeling operator suffers from back-scattering artifacts that severely hamper the speed of convergence. To avoid the back-scattering artifacts, I propose to use least-square reverse-time migration with wavefield decomposition. The incident and scattering wavefields are decomposed into up- and down-going direction such that only the forward-scattering component is used in imaging and modeling. Compared with using the conventional Laplacian preconditioning method, the proposed technique converges faster because it does not bias the inversion towards higher-frequency content. I demonstrate the concept and methodology with the 2D Seam model.

INTRODUCTION

It has been observed that the conventional cross-correlation imaging condition produces strong low-frequency migration artifacts in reverse-time migration (RTM) (Baysal et al., 1984). Migration artifacts appear mainly at shallow depth but also above strong reflectors. Figure 1 explains the generation of such artifacts. Due to the two-way propagation, source and receiver wavefields back-scatter off any hard velocity interfaces. With the cross-correlation imaging condition, the two-way propagation in RTM creates a wreath of possible imaging points. Figure 1 (a) and (b) show the location of undesirable cross-correlation energy that generate the low-frequency back-scattering artifacts. In least-squares reverse-time migration (LSRTM), the adjoint of the Born-modeling operator is equivalent to the RTM operator. LSRTM is considered to be a large problem that requires iterative inversion scheme. A successive sequence of adjoint and modeling operations is required to iteratively change the initial solution into the inverted one. As a result, when the adjoint operation is used in LSRTM, the back-scattering artifacts also exist in the gradient. Such artifacts hamper the convergence because erroneous energy in introduced into the gradient at every iterations.

There are many published work on how to suppress these artifacts (Baysal et al., 1984; Mulder and Plessix, 2003; Yoon et al., 2004; Fletcher et al., 2005). I will discuss two ways that can be incorporated into the LSRTM algorithm to suppress the back-scattering artifacts. A first way is to apply the Laplacian filter to the stacked RTM image. Zhang and Sun (2009) have showed that applying the Laplacian filter (along with a $1/\omega^2$ filter onto the data) is equivalent to applying a $cos^2\theta$ weight to the angle gather. This cosine factor down-weights contributions from large angle and effectively suppresses the back-scattered artifacts when $\theta = 90^o$. A straightforward incorporation of this idea in LSRTM is to use the Laplacian filter as a preconditioner in the inversion. The preconditioner scale low-frequency down to suppress the back-scattering artifact in the gradient.

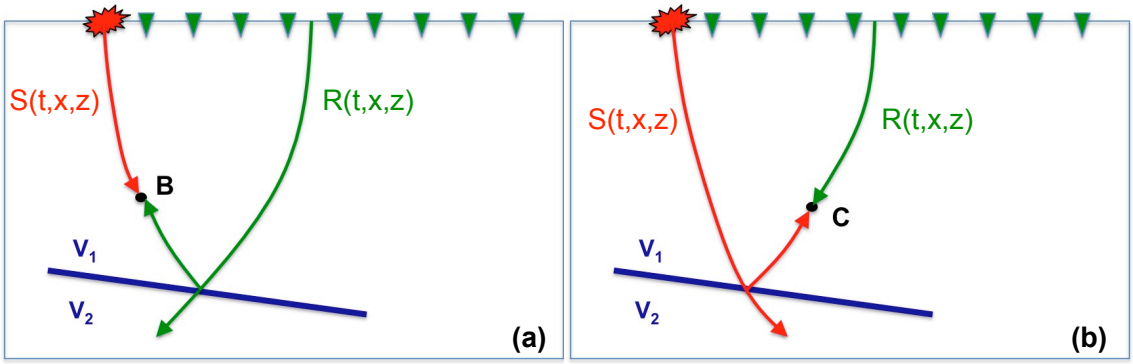


Figure 1: Consider an interaface with a sharp velocity contrast, migration artifacts occur when (a) the source wavefield correlate with the back-scattered receiver wavefield or (b) the vice-versa. [NR] mandy2/. Vfig1sep149mandy2

The second approach to suppressing back-scattering artifacts is to decompose wavefields into one-way components and only cross-correlate the wave components that occur as reflections (Liu et al., 2007). However, fully decomposing wavefields into eight directions in 3D can be computationally expensive. Instead, I follow the ideas from Liu et al. (2011), which introduce an efficient algorithm for up-down decomposition of the source and receiver wavefields with having to save the entire wavefield in time. By decomposing the wavefields into up-going and down-going direction, the migration image (or the gradient) can be separated into forward and back-scattering components. Most of the RTM artifacts are contained in the back-scattering component, those noise can be conveniently excluded in the gradient calculation for LSRTM. While in the forward modeling direction of LSRTM, only the forward-scattering events will be calculated.

The paper starts with a discussion on the theory of LSRTM with wavefield decomposition. It will be followed by a comparison of some results between the Laplacian preconditioner and the wavefield decomposition approach of suppressing back-scattering artifact using a modified version of the 2D Seam model.

THEORY

Least-squares migration poses the imaging problem as an inversion problem by linearizing the wave equation with respect to our model ($m(\mathbf{x})$). The model is defined to be a weighted difference between the migration slowness ($s_o(\mathbf{x})$) and the true slowness ($s(\mathbf{x})$):

$$m(\mathbf{x}) = (s(\mathbf{x}) - s_o(\mathbf{x}))s_o(\mathbf{x}). \tag{1}$$

The constant-density acoustic isotropic wave-equation is used in this study. The wave-equation is linearized with respect to $m(\mathbf{x})$ by applying the first-order Born approximation. The forward-modeling equation is

$$d^{mod}(\mathbf{x_r}, \mathbf{x_s}, \omega) = \sum_{\mathbf{x}} \omega^2 f_s(\omega) G(\mathbf{x_s}, \mathbf{x}) m(\mathbf{x}) G(\mathbf{x}, \mathbf{x_r}). \tag{2}$$

where d^{mod} represents the forward modeled data, ω is the temporal frequency, $m(\mathbf{x})$ is a function of the image point $\mathbf{x}$, $f_s(\omega)$ is the source waveform, and $G(\mathbf{x_s}, \mathbf{x})$ is the Green's function of the two-way acoustic constant-density wave equation over the migration slowness. Note that G is actually ω-dependent and is a function of $s_o(\mathbf{x})$ only. The inversion problem is defined as minimizing the least-squares difference between the synthetic and the recorded data:

$$S(\mathbf{m}) = \| \mathbf{Lm} - \mathbf{d} \|^2 = \| \mathbf{d^{mod}} - \mathbf{d} \|^2, \tag{3}$$

where $\mathbf{L}$ is the matrix form of the forward-modeling operator as shown in equation 2. It is important to point out that the adjoint of the forward-modeling operator is the migration operator:

$$\begin{aligned} m_{mig}(\mathbf{x}) &= \sum_{\mathbf{x_r},\mathbf{x_s},\omega} \omega^2 f_s^*(\omega) G^*(\mathbf{x_s}, \mathbf{x}) G^*(\mathbf{x}, \mathbf{x_r}) d(\mathbf{x_r}, \mathbf{x_s}, \omega), \\ &= \sum_{\mathbf{x_s},t} s(t, \mathbf{x}; \mathbf{x_s}) r(t, \mathbf{x}; \mathbf{x_s}). \end{aligned} \tag{4}$$

Note that the second line of equation 4 express the adjoint operation of the forward-modeling operation to the more familiar zero-lag cross-correlation imaging condition in RTM. $s(t, \mathbf{x})$ and $r(t, \mathbf{x})$ represent the source and receiver wavefield as a function of time t and sub-surface location $\mathbf{x}$. To reduce the back-scattering artifacts, Liu et al. (2011) propose to partition the source and receiver wavefields as

$$s(t, \mathbf{x}; \mathbf{x_s}) = s_d(t, \mathbf{x}; \mathbf{x_s}) + s_u(t, \mathbf{x}; \mathbf{x_s}), \tag{5}$$

$$r(t, \mathbf{x}; \mathbf{x_s}) = r_d(t, \mathbf{x}; \mathbf{x_s}) + r_u(t, \mathbf{x}; \mathbf{x_s}), \tag{6}$$

where $s_u(t, \mathbf{x}; \mathbf{x_s})$ and $s_d(t, \mathbf{x}; \mathbf{x_s})$ represents the up- and down-going partition of the source wavefield. Similarly, $r_u(t, \mathbf{x}; \mathbf{x_s})$ and $r_d(t, \mathbf{x}; \mathbf{x_s})$ represents the up- and down-going partition of the receiver wavefield. T is the recording time of the observed data. The imaging condition can then be decompose into four component by substituting equation 5 and 6 into the imaging condition (equation 4).

$$m_{mig}(\mathbf{x}) = I_1(\mathbf{x}) + I_2(\mathbf{x}) + I_3(\mathbf{x}) + I_4(\mathbf{x}), \tag{7}$$

$$I_1(\mathbf{x}) = \sum_{\mathbf{x_s},t} s_d(t, \mathbf{x}; \mathbf{x_s}) r_u(T - t, \mathbf{x}; \mathbf{x_s}), \tag{8}$$

$$I_2(\mathbf{x}) = \sum_{\mathbf{x_s},t} s_u(t, \mathbf{x}; \mathbf{x_s}) r_d(T - t, \mathbf{x}; \mathbf{x_s}), \tag{9}$$

$$I_3(\mathbf{x}) = \sum_{\mathbf{x_s},t} s_u(t, \mathbf{x}; \mathbf{x_s}) r_u(T - t, \mathbf{x}; \mathbf{x_s}), \tag{10}$$

$$I_4(\mathbf{x}) = \sum_{\mathbf{x_s},t} s_d(t, \mathbf{x}; \mathbf{x_s}) r_d(T - t, \mathbf{x}; \mathbf{x_s}), \tag{11}$$

where $I_1(\mathbf{x})$ corresponds to the cross-correlation of the down-going source and up-going receiver wavefields as shown in Figure 2a . $I_2(\mathbf{x})$ corresponds to the cross-correlation of the up-going source and down-going receiver wavefields as shown in Figure 2b. Both terms produce images at reflection points. I will refer I_1 and I_2 as the forward-scattering term. On the other hand, $I_3(\mathbf{x})$ corresponds to the cross-correlation of the up-going source and up-going receiver wavefields while $I_4(\mathbf{x})$ corresponds to the cross-correlation of the down-going source and down-going receiver wavefields. These terms produce the low-frequency

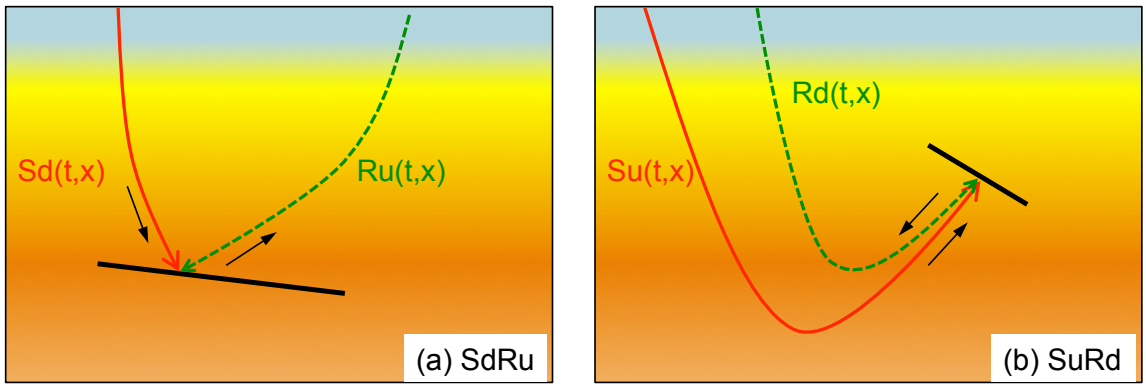


Figure 2: Ray-path diagrams of the imaging condition of (a) down-going source and up-going receiver wavefields and (b) up-going source and down-going receiver wavefields. These image points corresponds to the forward-scattering events. Note that the black arrows indicate the direction of wavefields when played forward in time. The green arrows indicate the direction of the receiver wavefields played in reverse time. [NR] mandy2/. Vfig2sep149mandy2

artifacts in RTM, illustrated by Figure 1. I will refer I_3 and I_4 as the back-scattering term. We can eliminate the back-scattering artifacts in the image with the following migration operator,

$$m_{decomp}(\mathbf{x}) = \sum_{\mathbf{x_s},t} s_d(t,\mathbf{x};\mathbf{x_s}) r_u(t,\mathbf{x};\mathbf{x_s}) + s_u(t,\mathbf{x};\mathbf{x_s}) r_d(t,\mathbf{x};\mathbf{x_s}). \tag{12}$$

In general, extracting up- and down-going wavefield requires a 2D Fourier transform in (t, x_z) space. This is computationally expensive because the entire source and receiver wavefield must be saved in memory. To go around this problem, Liu et al. (2011) proposed decomposing the wavefields in the k_z axis, which only requires 1D Fourier transform in z. Let $\widetilde{s}(t,x,k_z)$ be the 1D Fourier transform of the source wavefield $s(t,\mathbf{x})$ in z. Let $s_+(t,\mathbf{x})$ and $s_-(t,\mathbf{x})$ represents the $k_z\pm$ decomposition of the source wavefield as,

$$\widetilde{s_+}(t,x,k_z) = \begin{cases} \widetilde{s}(t,x,k_z) & \text{if } \omega \geq 0 \\ 0 & \text{if } \omega < 0, \end{cases}$$

$$\widetilde{s_-}(t,x,k_z) = \begin{cases} 0 & \text{if } \omega \geq 0 \\ \widetilde{s}(t,x,k_z) & \text{if } \omega < 0, \end{cases}$$

where $\widetilde{s_+}(t,x,k_z)$ and $\widetilde{s_-}(t,x,k_z)$ is the 1D Fourier transform of a new representation of the decomposed source wavefields $s_+(t,\mathbf{x})$ and $s_-(t,\mathbf{x})$. Similar definition follows for the receiver wavefields $r_+(t,\mathbf{x})$ and $r_-(t,\mathbf{x})$. Liu et al. (2011) show that the imaging condition in equation 12 can also be written as

$$m_{decomp}(\mathbf{x}) = \sum_{\mathbf{x_s},t} s_+(t,\mathbf{x}) r_-(t,\mathbf{x}) + s_-(t,\mathbf{x}) r_+(t,\mathbf{x}). \tag{13}$$

The forward-modeling direction is equivalent to the Born-modeling operator that corresponds to the forward-scattering events as illustrated in Figure 3a and b. A derivation of the forward-modeling operator using wavefields decomposition is shown in the appendix.

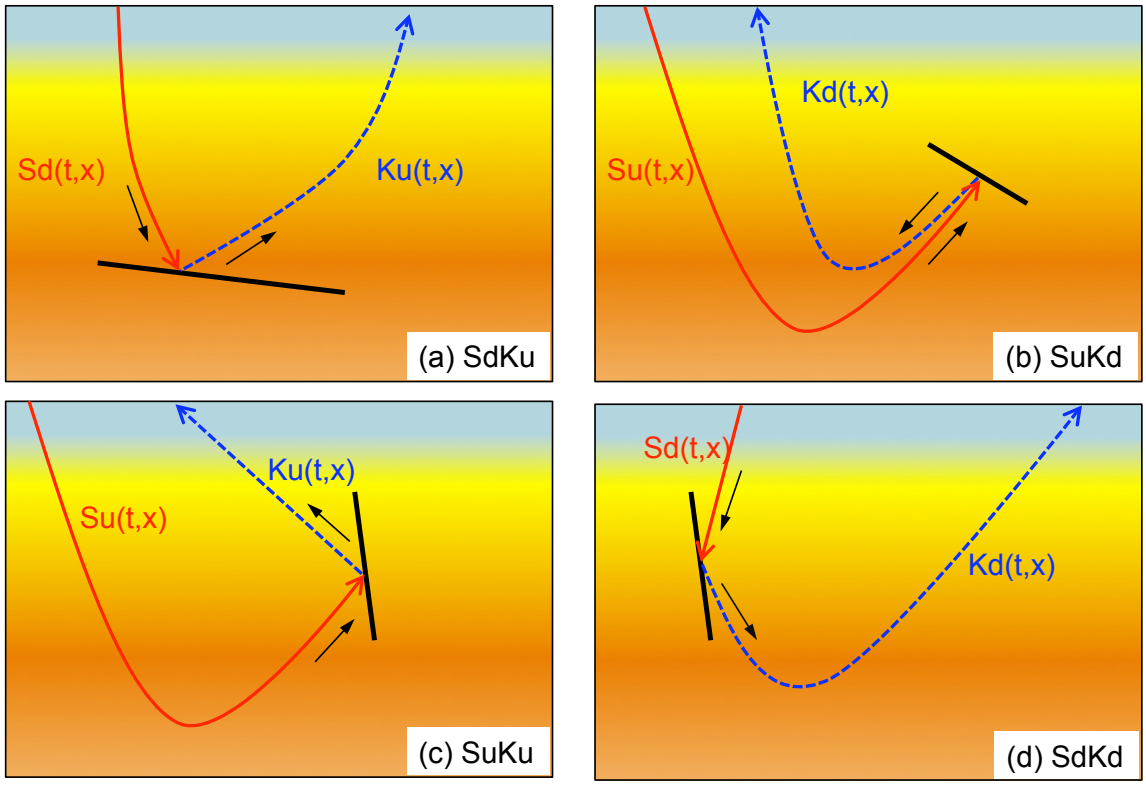


Figure 3: Ray-path diagram of the forward scattering events. (a) shows down-going incident ($s_d(t,\mathbf{x})$) and up-going scattered $k_u(t,\mathbf{x})$) wavefields and (b) up-going incident $s_u(t,\mathbf{x})$) and down-going $k_d(t,\mathbf{x})$) scattered wavefields. (c) and (d) show ray paths of some back-scattering events that are excluded from the migration and modeling. [NR] mandy2/. Vfig3sep149mandy2

2D Seam model

I extracted a 2D velocity section of the Seam model for this study. Figure 4a shows the true velocity model used to generate synthetic data. The model is 16000 m wide, 7000 m deep and a spatial spacing of 20 m. 20 shots are evenly distributed at the surface. Each shot gather is generated with receivers located at every grid point at the surface. The migration velocity model (Figure 4b) is a smoothed version of the true velocity model in the sediment layers. I assume the salt is being picked correctly in the migration velocity model excluding the details of the dirty salt. Figure 4c shows the image model we wish to recover. The true model is calculated using equation 1.

I compare imaging result between RTM, LSRTM with Laplacian preconditioner, and LSRTM with wavefield decomposition. LSRTM with wavefield decomposition has been described in the previous section. In LSRTM with Laplacian preconditioner, a new model variable $\mathbf{x}$ is defined such that $\mathbf{m} = \mathbf{A}\mathbf{x}$ where $\mathbf{A}$ is the Laplacian operator. Note that $\mathbf{A}$ is self-adjoint. Equation 3 then becomes

$$S(\mathbf{x}) = \| \mathbf{L}\mathbf{A}\mathbf{x} - \mathbf{d} \|^2 = \| \mathbf{d}^{\mathbf{mod}} - \mathbf{d} \|^2 \ . \tag{14}$$

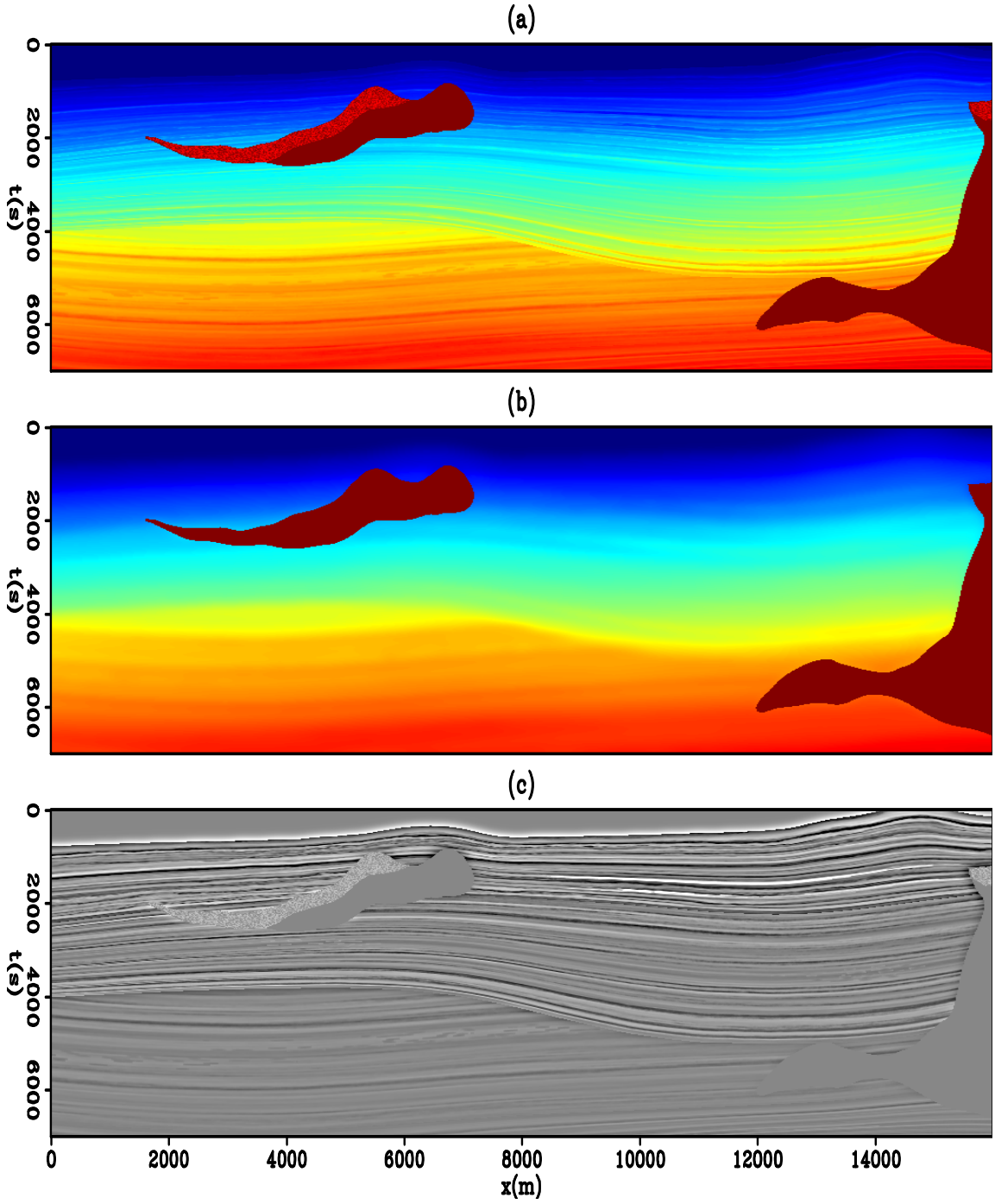


Figure 4: (a) True velocity, (b) migration velocity, and (c) true model. [ER] mandy2/. Vfig4sep149mandy2

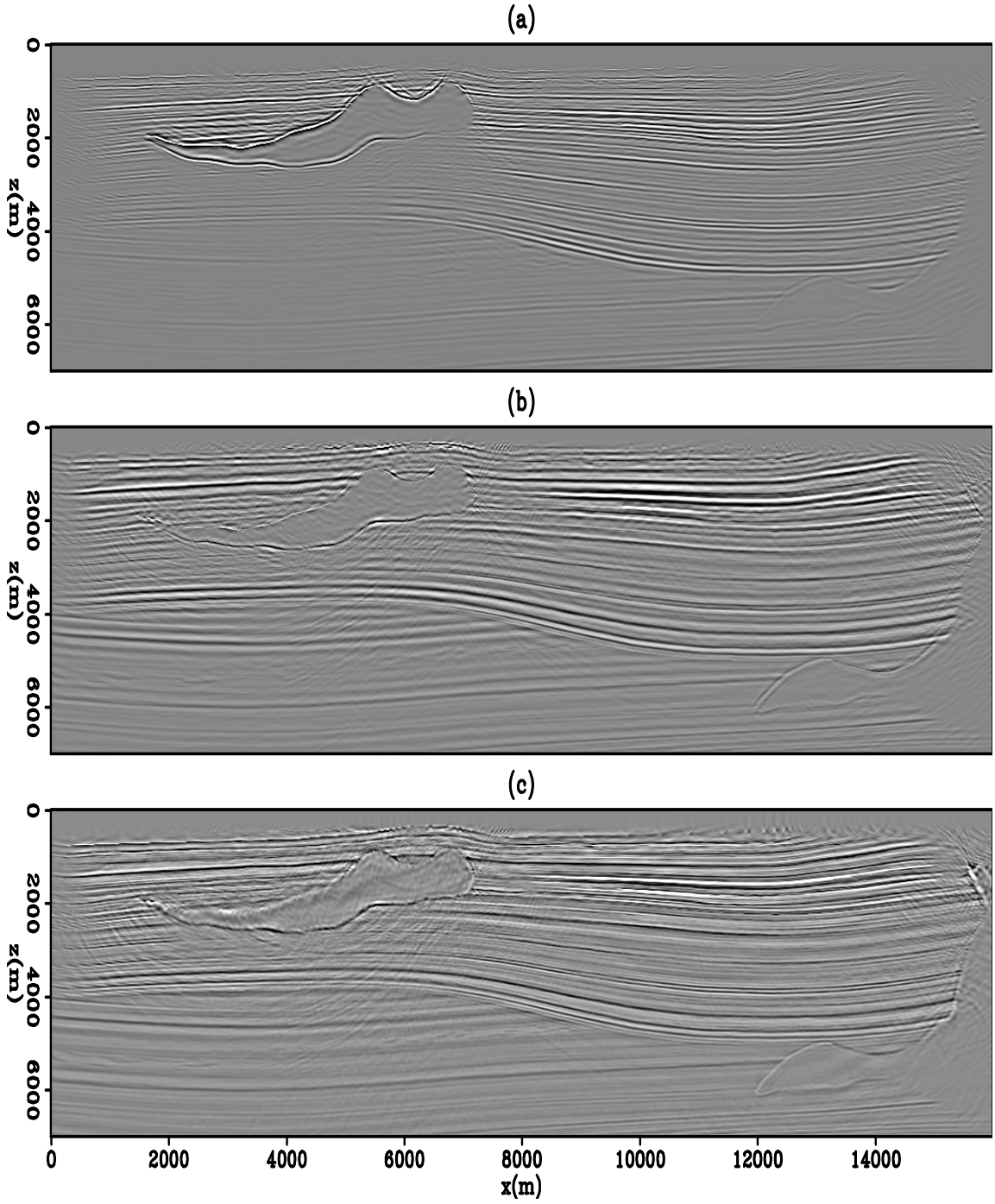


Figure 5: (a) RTM image, (b) LSRTM with Laplacian preconditioner at iteration 20, and (c) LSRTM with wavefield decomposition at iteration 20. [CR] mandy2/. Vfig5sep149mandy2

Figure 5a shows the RTM image of the Seam model. In the RTM image, there are strong back-scattering artifacts above the salt layer that are suppressed with a Laplacian filter. When LSRTM with wavefield decomposition is applied, the gradient at every iteration is separated into the forward-scattering component (Figure 6a) and the back-scattering component (Figure 6b). Only the forward-scattering part is used to update the model. The sum of the forward- and back-scattering component (Figure 6) yields the gradient for regular LSRTM.

Figure 5b and 5c show the LSRTM image with the Laplacian preconditioning and with wavefield decomposition at iteration 20, respectively. In general, the LSRTM images are better than the RTM image (Figure 5a) with higher frequency content and better relative amplitude across the reflectors.

To compare between the two LSRTM results, let look at two zoomed areas. Figure 7 shows a zoomed section of the image from $x = 4000 - 7200m$ and $z = 1500 - 4500m$. We can easily identify the shadow zone underneath the salt structure in the RTM image (Figure 7b). The information in the shadow zone is partially recovered with LSRTM as shown in Figure 7c and d. The quality of the inverted image can be evaluated by looking at the relative strength of the reflectors across the image. By comparing the relative amplitude with the true model (Figure 7a), the LSRTM image with wavefield decomposition (Figure 7d) has better relatively amplitude information than the LSRTM image with Laplacian preconditioner (Figure 7c). Figure 9 shows the convergence curve between the two inversion. One remark is that using wavefield decomposition seems to allow the inversion to converge faster than using the Laplacian preconditioner. At iteration 10, the residual in wavefield decomposition has dropped down to below 5% while the residual from the Laplacian preconditioner is at around 10%.

Figure 8 shows another zoomed section of the image from $x = 7000 - 10000m$ and $z = 3500 - 6000m$. Figure 8a is the true model we wish to recover. Figure 8b is the RTM image. Notice that some of the higher frequency structures as annotated by circles is not well represented in the RTM image. Those structures are better represented in Figure 8 c, d, e, and f with LSRTM. By comparing both result at iteration 10, the Laplacian preconditioning image (Figure 8c) is less refined than the wavefield decomposition image (Figure 8d). However, as the inversion progress to iteration 40, the Laplacian preconditioning result (Figure 8e) has recovered some refined details. The images from iteration 10 and 40 suggests that LSRTM with wavefield decomposition converges faster than with Laplacian preconditioner.

DISCUSSIONS

Amplitude bias

From the synthetic result, I observe that LSRTM with wavefield decomposition converges faster than LSRTM with the Laplacian preconditioner. One explanation is because the later method tends to bias the inversion towards fitting the higher frequency. Figure 10 shows the amplitude spectrum of one shot gather from the input data and from the forward modeling of the first gradient from the two LSRTM methods. The spectrum is calculated by taking Fourier transform in time of each trace in the shot gather and the resulting spectrum is average along all traces (receivers) in the shot gather. Notice that the spectrum from the

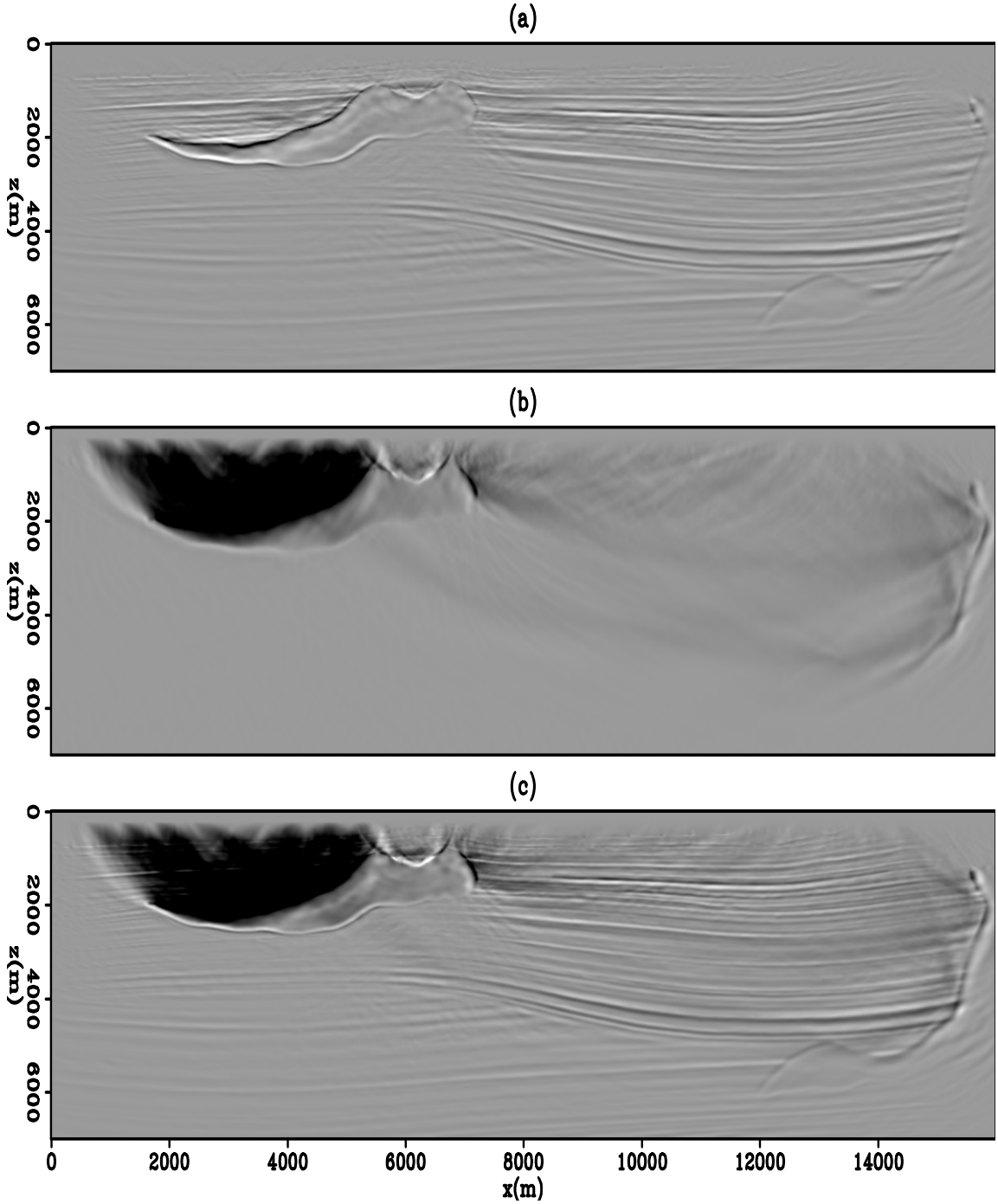


Figure 6: (a) Forward scattering part of the RTM image ($I_1 + I_2$), (b) back-scattering part of the migration image ($I_3 + I_4$) and (c) the total RTM image ($I_1 + I_2 + I_3 + I_4$). [CR] mandy2/. Vfig6sep149mandy2

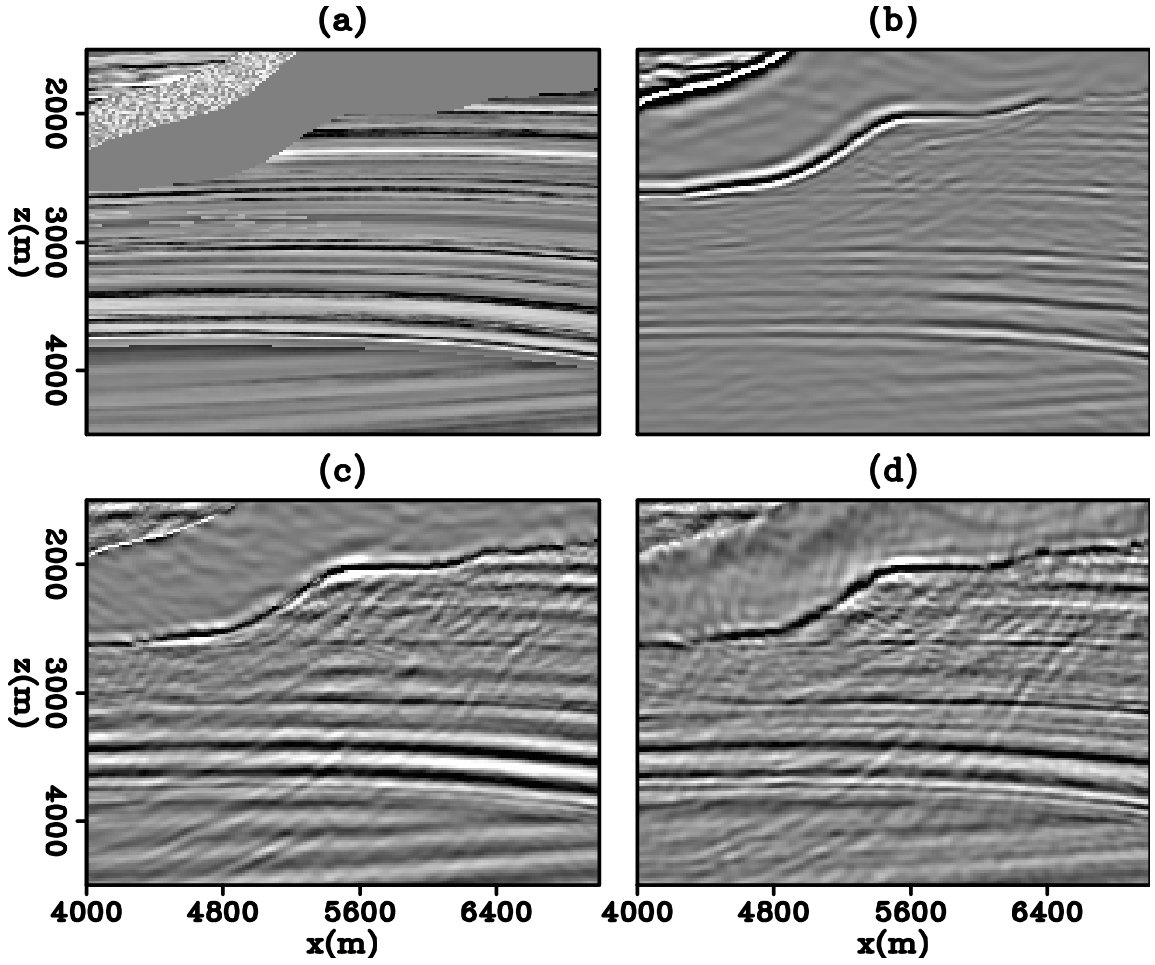


Figure 7: A zoomed section of (a) the true model, (b) RTM image, (c) LSRTM with Laplacian preconditioner, and (d) LSRTM with wavefield decomposition. [CR] mandy2/. Vfig7sep149mandy2

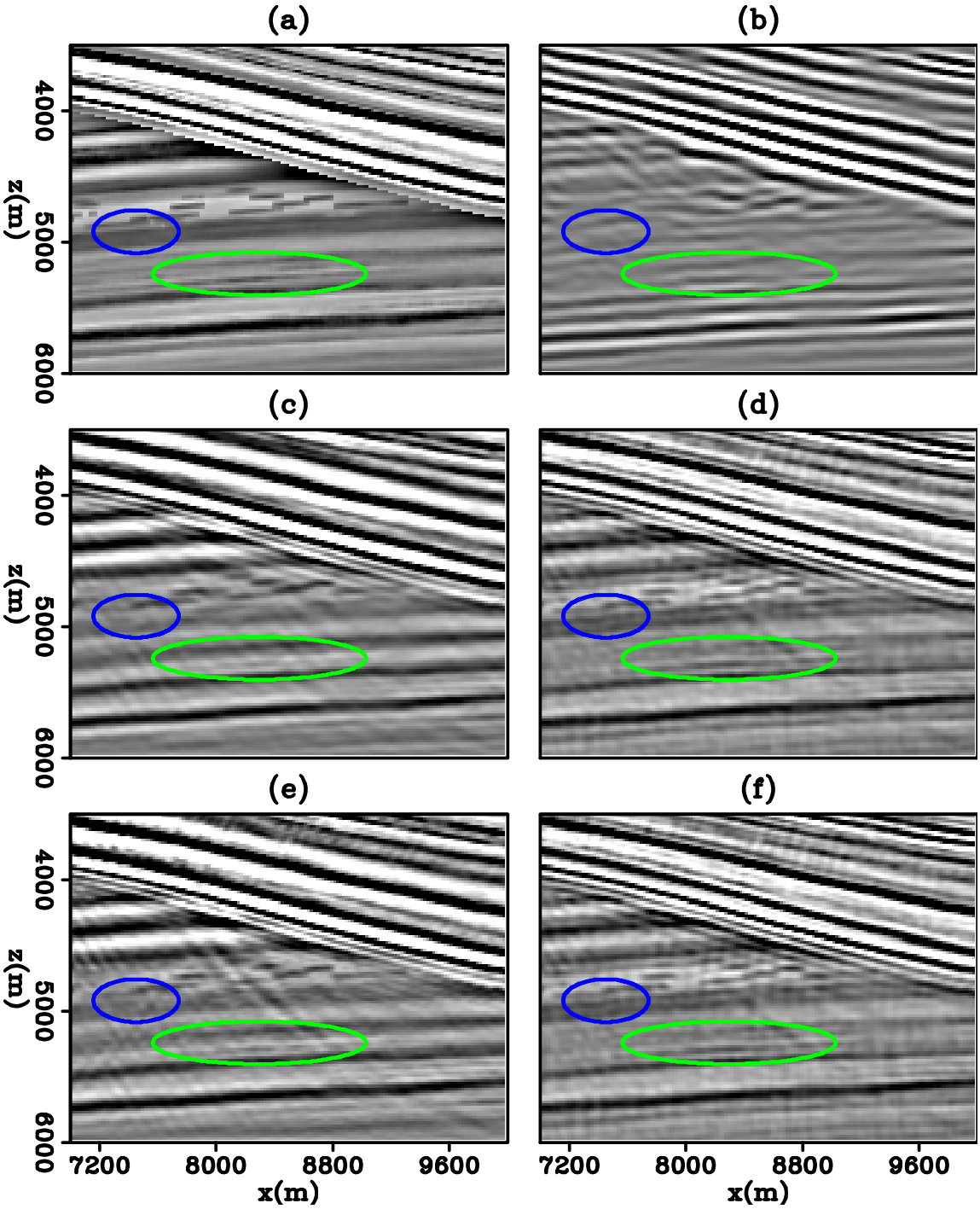


Figure 8: A zoomed section of (a) the true model, (b) RTM image, (c) LSRTM with Laplacian preconditioner at iteration 10, (d) LSRTM with wavefield decomposition at iteration 10, (e) LSRTM with Laplacian preconditioner at iteration 40 and (f) LSRTM with wavefield decomposition at iteration 40. All images are scaled to one and are displayed at the same clip of ±0.25. [CR] mandy2/. Vfig8sep149mandy2

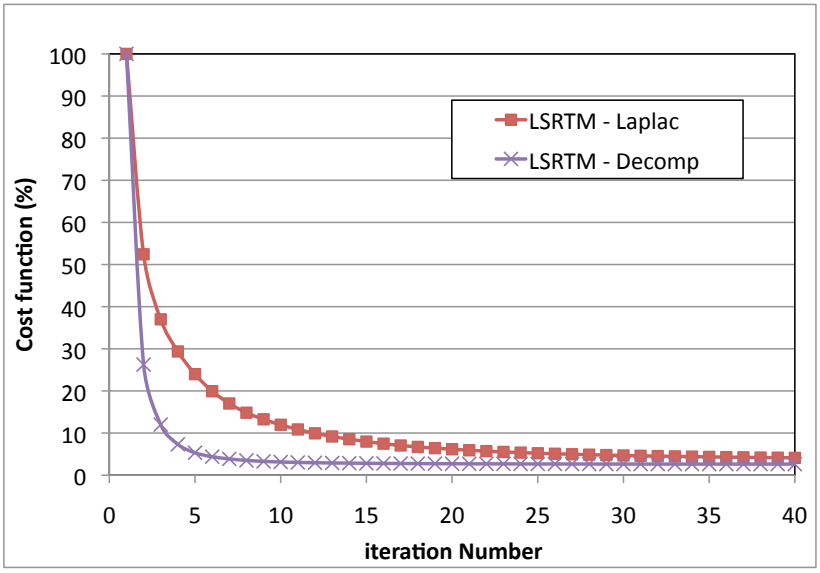


Figure 9: Convergence plot for LSRTM with the Laplacian preconditioner (square marker) and LSRTM with wavefield decomposition (cross marker). In general, using wavefield decomposition converges faster than using Laplacian preconditioner. [NR] mandy2/. Vfig10sep149mandy2

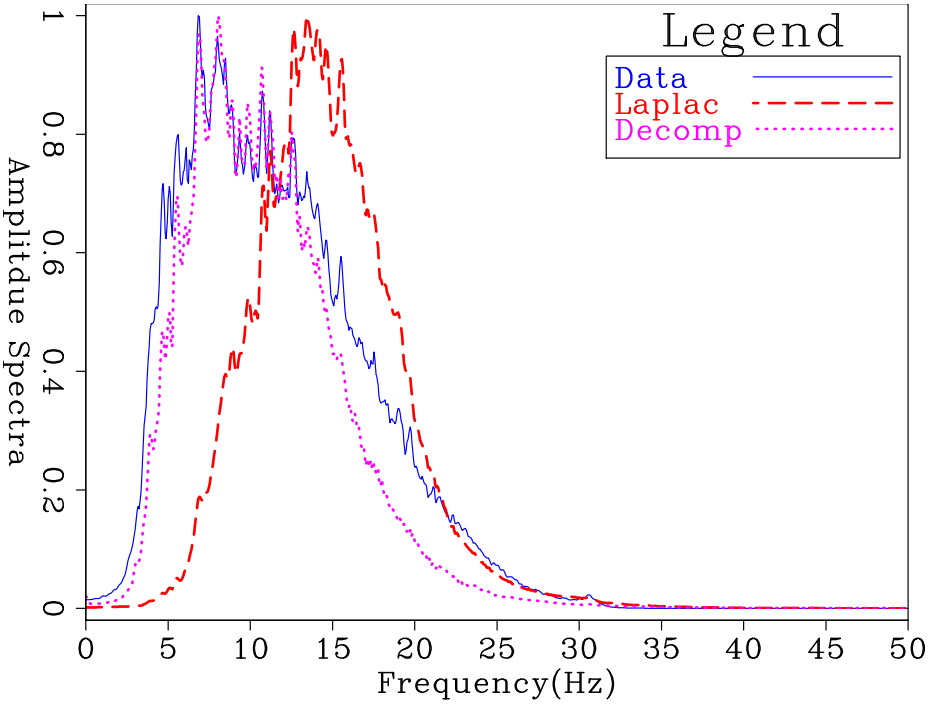


Figure 10: Amplitude spectrum of one shot gather from the input data (solid line). The amplitude spectrum of the forward modeling of the first gradient from LSRTM with Laplacian preconditioner (dash line) and LSRTM with wavefield decomposition (dotted line). [CR] mandy2/. Vfig9sep149mandy2

wavefield decomposition technique is much closer to the spectrum of the input data than that of the Laplacian preconditioning. The frequency bias can be compensated by adding a left-preconditioner in the inversion with the following fitting goal:

$$\mathbf{P}(\mathbf{LAx} - \mathbf{d^{obs}}) \approx 0, \tag{15}$$

where $\mathbf{P}$ represents the $\frac{1}{\omega^2}$ filter. However, the affects of including a left-preconditioner along with other data-weighting term will need to be studied.

Including the back-scattering term

One drawback for using LSRTM with wavefield decomposition is that some true back-scattering event (Figure 3c and d) will be excluded from the inversion. This implies that some steeply dipping reflectors might not be imaged with this technique. One way to address this problem is to introduce a spatially varying weighting for the back-scattering term.

$$m_{mig}(\mathbf{x}) = I_1(\mathbf{x}) + I_2(\mathbf{x}) + M_{back}(\mathbf{x})(I_3(\mathbf{x}) + I_4(\mathbf{x})), \tag{16}$$

where $M_{back}(\mathbf{x})$ is a image-space weighting function. It is set to zero for regions with strong back-scattering artifacts and is set to ones elsewhere. The region with strong back-scattering artifacts is generally in the shallow region. This weighting would at least allow some true back-scattering event to be included in the inversion. Further study will be needed for this approach.

CONCLUSION

Back-scattering artifacts in reverse-time migration occur when diving waves, head waves or backscattered waves cross-correlate. These events are particularly strong where high velocity contrasts exist. In LSRTM, the adjoint of the linearized forward-modeling operator suffers from back-scattering artifacts that severely hamper the speed of convergence. One solution to this problem is to use a Laplacian filter as a preconditioner, which suppress lower-frequency and boost up higher-frequency. An alternate approach is to use LSRTM with wavefield decomposition. Based on the 2D Seam example, I found that the proposed technique allow the inversion to converge faster than using LSRTM with the Laplacian preconditioner.

ACKNOWLEDGMENTS

I thank Kittinat Taweesintananon for previous work on this topic. I also thank Biondo Biondi and Shuki Ronen for helpful discussion.

REFERENCES

Baysal, E., D. Kosloff, and J. Sherwood, 1984, A two-way nonreflecting wave equation: Geophysics, **49**, 131–141.

Fletcher, R., P. Fowler, P. Kitchenside, and U. Albertin, 2005, Suppressing artifacts in presetack reverse time migration: SEG Expanded Abstracts, 2049–2051.

Liu, F., G. Zhang, S. Morton, and J. Leveille, 2007, Reverse-time migration using one-way wavefield imaging condition: SEG Expanded Abstracts, **76**, S29–S39.

———, 2011, Reverse-time migration using one-way wavefield imaging condition: Geophysics, 2170–2174.

Mulder, W. and R.-E. Plessix, 2003, One-way and two-way wave equation migration: SEG Expanded Abstracts, 881–884.

Yoon, K., K. Marfurt, and W. Starr, 2004, Challenges in reverse-time migration: SEG Expanded Abstracts, 1057–1060.

Zhang, Y. and J. Sun, 2009, Practical issues in reverse time migration: true amplitude gathers, noise removal and harmonic source encoding: First break, **26**, 30–35.

APPENDIX

Forward modeling with wavefield decomposition

The goal here is to show that the adjoint of equation 12 is the Born-modeling equivalent of the forward scattering events. The derivation will be shown in 2D but extension to 3D is straight forward. To simplify notation, I will represent the model $m(x,z)$ and the data $d(x_r,t)$ as matrices $\mathbf{M}$ and $\mathbf{D_r}$ of dimension (x,z,t). The relationship is summarized as follow,

$$m(x,z) = \sum_t \mathbf{M}(x,z,t) = \mathbf{AM}, \tag{17}$$

$$\mathbf{D_r}(x,z,t) = \mathbf{L}_r d(x_r, z=0, t), \tag{18}$$

where matrix $\mathbf{A}$ is a stacking along the time axis t and $\mathbf{L}_r$ project the data at receiver location x_r. Let think of the input and output of the modeling and migration operators as M and D_r. The decomposing migration operator of equation 12 can be re-written in matrix form as follow,

$$\mathbf{M} = \mathbf{S_u K_d G D_r} + \mathbf{S_d K_u G D_r}. \tag{19}$$

In the above equation, $\mathbf{S_u}$ and $\mathbf{S_d}$ are up- and down-going source wavefield, respectively. $\mathbf{G}$ is the matrix operator form of the Green's function. $\mathbf{K_u}$ and $\mathbf{K_d}$ decomposition operators that partition wavefields into up- and down-going direction. They are defined as,

$$\mathbf{K_u} = \mathbf{F^{-1}}\widetilde{\mathbf{k_u}}\mathbf{F}, \tag{20}$$

$$\mathbf{K_d} = \mathbf{F^{-1}}\widetilde{\mathbf{k_d}}\mathbf{F}, \tag{21}$$

where F and F^{-1} is the forward and inverse 2D Fourier transform in (z,t). $\widetilde{\mathbf{k_u}}$ and $\widetilde{\mathbf{k_d}}$ extract up- and down-going direction in the (k_z,ω) space;

$$\widetilde{\mathbf{k_u}} = \begin{cases} 1 & \text{if } \omega k_z \geq 0 \\ 0 & \text{if } \omega k_z < 0, \end{cases}$$

$$\widetilde{\mathbf{k_d}} = \begin{cases} 0 & \text{if } \omega k_z \geq 0 \\ 1 & \text{if } \omega k_z < 0. \end{cases}$$

Note that the decomposition operators are self-adjoint. The forward-modeling operator with wavefield decomposition is defined to be

$$\mathbf{D_r} = \mathbf{G}^*\mathbf{K}_\mathbf{u}^*\mathbf{S}_\mathbf{d}^*\mathbf{M} + \mathbf{G}^*\mathbf{K}_\mathbf{d}^*\mathbf{S}_\mathbf{u}^*\mathbf{M}. \tag{22}$$

In the first term of equation 22, down-going source wavefield is multiplied with model $\mathbf{M}$ and only the up-going component is extracted by the operator $\mathbf{K_u}$. The extracted energy then propagates back to the receiver by $\mathbf{G}^*$. In the second term, the up-going source wavefield multiplies the model $\mathbf{M}$ and only the down-going component is extracted.

Forward modeling with up-down and $k_z\pm$ decomposition

The goal of this section is to show that the forward-modeling with up-down decomposition can also be represented with the $k_z\pm$ representation. This section follows the prove similar to that in the paper Liu et al. (2011) for the migration direction. Let first re-write equation 22 with an explicit dependence on the variable ω

$$\widetilde{\mathbf{d}_\mathbf{r}}(\omega) = \widetilde{\mathbf{g}^*}(\widetilde{\mathbf{k}_\mathbf{u}^*}(\omega)\widetilde{\mathbf{s}_\mathbf{d}^*}(\omega) + \widetilde{\mathbf{g}^*}\widetilde{\mathbf{k}_\mathbf{d}^*}(\omega)\widetilde{\mathbf{s}_\mathbf{u}^*}(\omega))\widetilde{\mathbf{m}}. \tag{23}$$

The up-down decomposition can be expressed in terms of the $k_z\pm$ wavefields in the following ways:

$$\widetilde{\mathbf{k}_\mathbf{u}}(\omega) = \begin{cases} \widetilde{\mathbf{k}_+}(\omega) & \text{if } \omega \geq 0 \\ \widetilde{\mathbf{k}_-}(\omega) & \text{if } \omega < 0, \end{cases}$$

$$\widetilde{\mathbf{k}_\mathbf{d}}(\omega) = \begin{cases} \widetilde{\mathbf{k}_-}(\omega) & \text{if } \omega \geq 0 \\ \widetilde{\mathbf{k}_+}(\omega) & \text{if } \omega < 0, \end{cases}$$

$$\widetilde{\mathbf{s}_\mathbf{u}}(\omega) = \begin{cases} \widetilde{\mathbf{s}_+}(\omega) & \text{if } \omega \geq 0 \\ \widetilde{\mathbf{s}_-}(\omega) & \text{if } \omega < 0, \end{cases}$$

and

$$\widetilde{\mathbf{s}_\mathbf{d}}(\omega) = \begin{cases} \widetilde{\mathbf{s}_-}(\omega) & \text{if } \omega \geq 0 \\ \widetilde{\mathbf{s}_+}(\omega) & \text{if } \omega < 0. \end{cases}$$

The $\widetilde{\mathbf{k}_\pm}(\omega)$ are the $\widetilde{\mathbf{k}}(\omega)$ wavefield with $k_z\pm$ decomposition as referenced in Liu et al. (2011). Substituting the above equations into equation 23 gives

$$\widetilde{\mathbf{d}_\mathbf{r}}(\omega) = \begin{cases} \widetilde{\mathbf{g}^*}(\widetilde{\mathbf{k}_+^*}\widetilde{\mathbf{s}_-^*} + \widetilde{\mathbf{k}_-^*}\widetilde{\mathbf{s}_+^*})\widetilde{\mathbf{m}} & \text{if } \omega \geq 0 \\ \widetilde{\mathbf{g}^*}(\widetilde{\mathbf{k}_-^*}\widetilde{\mathbf{s}_+^*} + \widetilde{\mathbf{k}_+^*}\widetilde{\mathbf{s}_-^*})\widetilde{\mathbf{m}} & \text{if } \omega < 0, \end{cases}$$

which shows that the forward modeling with up-down decomposition (equation 23) is equivalent to the forward-modeling with $k_z\pm$ decomposition.

$$\widetilde{\mathbf{d}_\mathbf{r}}(\omega) = \widetilde{\mathbf{g}^*}(\widetilde{\mathbf{k}_+^*}(\omega)\widetilde{\mathbf{s}_-^*}(\omega) + \widetilde{\mathbf{g}^*}\widetilde{\mathbf{k}_-^*}(\omega)\widetilde{\mathbf{s}_+^*}(\omega))\widetilde{\mathbf{m}}. \tag{24}$$

Extended image space separation of continuously recorded seismic data

Chris Leader and Biondo Biondi

ABSTRACT

Conventional seismic surveying requires good temporal separation of shot points which can often lead to long waiting times, especially in techniques that use multiple source vessels. It is well established that by recording overlapping shot points we can reduce the cost of surveying. However, this makes data processing and imaging more difficult. Many existing methods have been suggested that can separate overlapping data; the caveat of all these is that the requirement of random time delays between shot points. By posing the problem in the extended image space it is possible to isolate and separate these data with a wide variety of different time delays, including linear time delays. The ongoing work described herein details how this method can be made computationally feasible and how to design these algorithms.

INTRODUCTION

The vast majority of seismic imaging algorithms assume that a single impulsive source interacts with a scattering field and is then recorded by the receivers. Any events present in the data that do not originate from this single source are typically misplaced and create crosstalk artifacts that degrade image quality. Thus, when surveying, a single source point is shot then followed by a waiting period of 10-20s. So that the seismic energy has sufficiently dissipated and will not interfere with the next seismic record. This is the case for both land and marine surveys and consequently a significant amount of time in the field is spent waiting. This is especially true for Wide Azimuth Towed Streamer (WATS) surveys, where there are multiple source vessels that must all wait for each other (Verwest and Lin, 2007).

If waiting time was not a restriction much more data could be recorded per unit time (or per dollar), and data with much denser source sampling could be recorded (Beasley (2008); Hampson (2008); Berkhout and Blacquiere (2008)). The cost of the survey is not in the recording, but in the time that the equipment is out. These sorts of data, where active shots can overlap, are often referred to as simultaneous source data or blended data. It has been shown that such data can be used to directly invert for model properties (Dai and Schuster (2009); Tang and Biondi (2009)). However, these methods require exact velocity model knowledge. This prerequisite is not in keeping with industry processing, whereas algorithms that could separate these overlapping data into separate shot components could be integrated into industry processing. Such methods exist and rely on random sampling in the source timings (Abma and Yan (2009); Moore et al. (2008)). For example, data can be transformed into the f-k or tau-p domain and iteratively thresholded (Doulgeris et al., 2011), iteratively removed in the parabolic random domain (Ayeni et al., 2011), removed by using an anti-leakage Fourier transform inverse approach (Abma et al., 2010), or through using compressive sensing methods (Herrmann et al., 2009).

These aforementioned methods have been shown to separate overlapping data very well under certain conditions. These all rely on the overlapping shots being incoherent (leading to sparsity) in a certain domain, and to a lesser extent the geology being consistent between proximate shots. If the source time delays are not truly random then all these techniques fail to adequately separate the shot records. In the extreme case of linear time delays the data is degradated. While randomly delaying the sources is not a difficult task, it would be preferable if this was not a stringent requirement. This paper suggests that by transforming our data into the extended model space it is possible to distinguish and isolate events that correspond to our shot-of-interest (our 'primary' data), and all those that are overlapping (our 'secondary' data). If we can separate our data into primary and secondary components in this space then it is possible to recreate our original survey in its equivalent, unblended form. Such a method does require a velocity model, however even if this velocity model is not exact the primary and secondary shots are still distinguishable. This lifts the requirement of having an exact velocity model to migrate and demigrate the data - as long as we use the same model for both transforms the result is invariant of the model used.

Herein will be described how we can use this extended model space for data separation under the cases of random time delays, linear time delays, and both random and linear delays. This can be done by filtering in the extended model space, by applying a simple inversion in the model space, or by using linearised inversion.

CONTINUOUSLY RECORDED DATA

There are several terms in modern nomenclature that describe actively shot data which overlaps, each meaning something subtly different. These data are called simultaneous source, blended, continuously recorded or overlapping data. Furthermore, these data can be acquired in several ways. We could have one source vessel shooting as often as possible, two vessels shooting, or multiple source vessels recording multiple source points simultaneously. Then the time delays for each source could be random or linear, and the delays between the different sources could be random or linear.

Fig 1 shows how a given shot record could appear when a series of sources are being shot with random delays. Fig 2 then shows how the same set of shots could appear when using a linear time delay; finally Fig 3 shows a combination of both randomly and linearly delayed sources.

We have multiple options when it comes to processing these data, and these can depend on shot density. The simplest is to just ignore the fact that these data are overlapping and to perform conventional migration. Thus treating the interfering shots as noise. This approach has become known as 'passive' imaging of blended data (not to be confused with imaging of passively recorded data). Fig 4 shows a simple image of a two layer medium when imaged using industry standard Reverse Time Migration (RTM.) Fig 5 then shows the same dataset imaged but allowing a certain amount of overlap within the data (the passive image) using a variety of linear and random time delays. As can be seen a lot of both coherent and incoherent noise is present. The dataset featured 75 shots in total, 15 inline and 5 crossline, with linear delays in the inline direction. We see in areas of low fold that ignoring the fact the data is blended gives unacceptably noisy images.

The easiest and cheapest way to separate these data would be simple data space filtering.

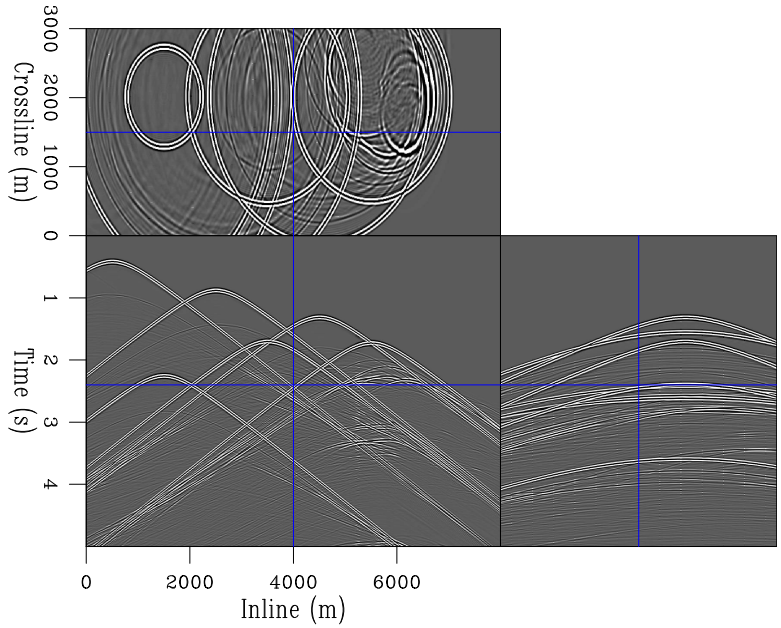


Figure 1: An example of several well separated shots with random time delays. [**CR**] chris2/. randata

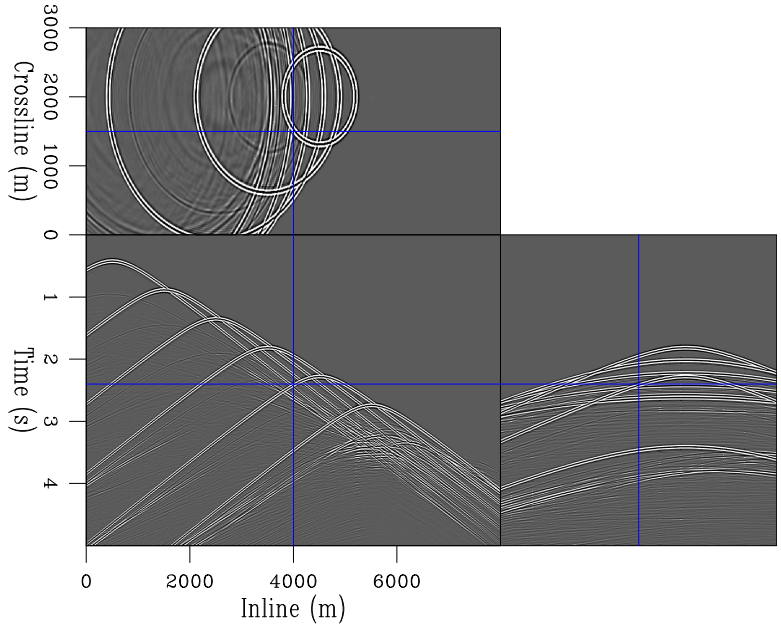


Figure 2: The same set of shots as Fig 1 but with linear time delays. [**CR**] chris2/. lindata

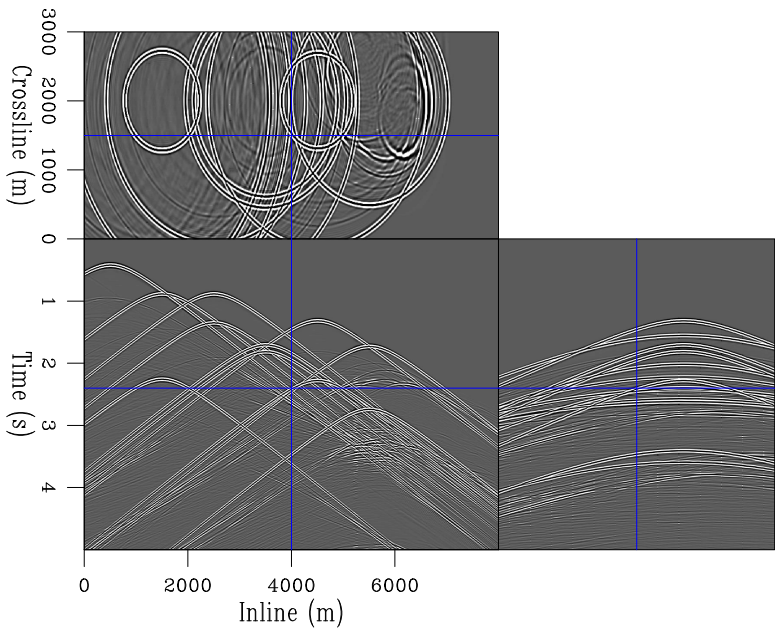


Figure 3: A synthesised survey which consists of both linear (inline) and random (crossline) time delays. [**CR**] chris2/. rldata

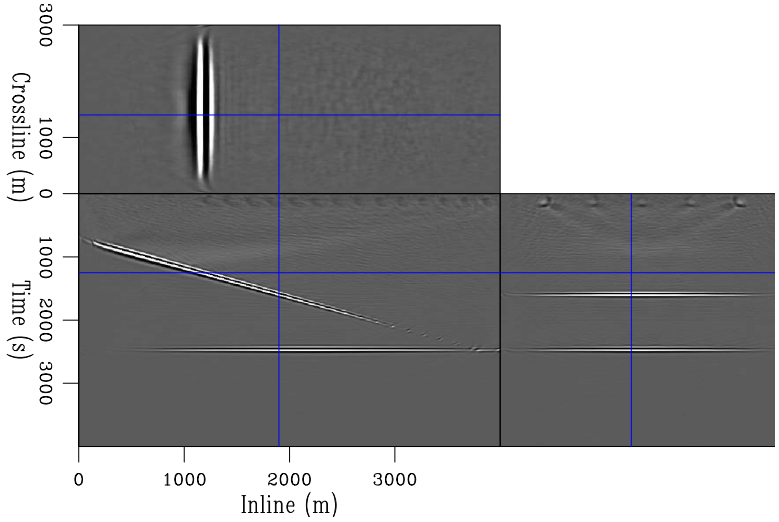


Figure 4: A simple two layer model imaged using 75 shots with random boundary reverse time migration. [**CR**] chris2/. imsim

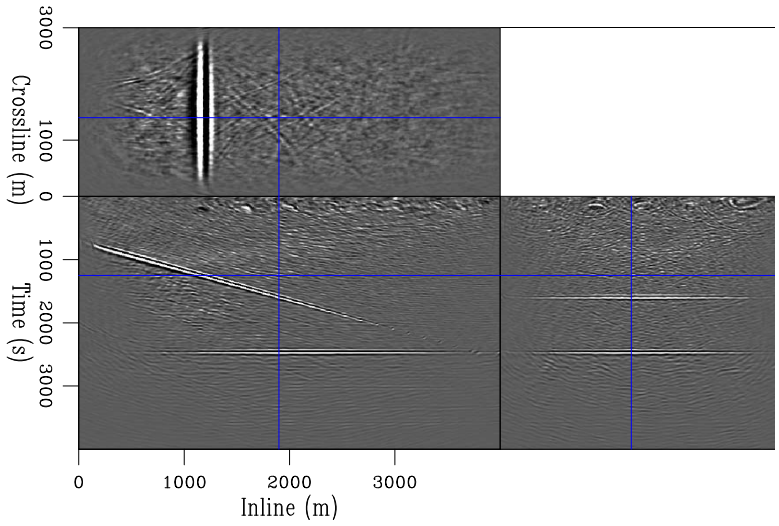


Figure 5: A simple two layer model, as in Fig 4, but imaged with linear blending along the inline direction and random blending along the crossline direction. [**CR**] chris2/. imsimbl

An example of two shots, well separated in space but simultaneous in time, can be seen in Fig 6. It is easy for the early arrivals to distinguish which shot they originated from. However, as we look deeper in the record it becomes impossible to determine from which shot a given event originates. Generally speaking, if the human eye cannot filter two events, an algorithm cannot (especially if we have no prior model, as in this case). As such we can conclude that data space filtering will not be a robust method for blended data separation.

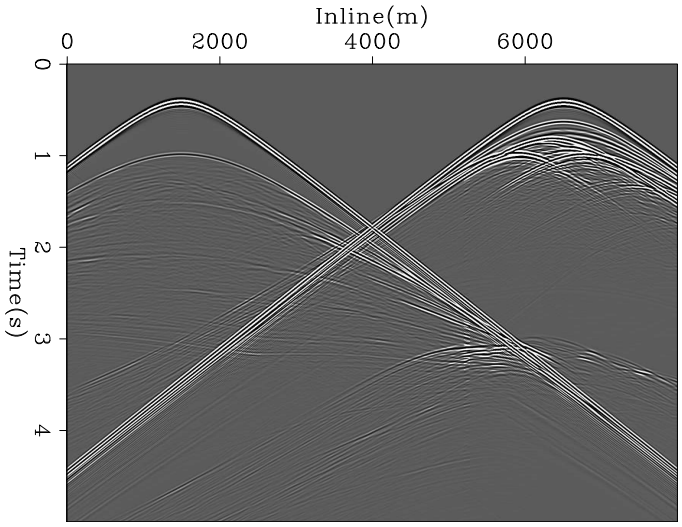


Figure 6: Two simultaneous, distance separated shots. Deep in the record it becomes impossible for the human eye to determine which shot a given originated. [**CR**] chris2/. canifilter

As alluded to in the introduction, various existing separation methods work on filtering or iteratively removing overlapping shots in some different domain. By transferring to the Fourier domain or the tau-p domain, these different shots become more distinguishable than in the original data space (alternatively, our primary data appears sparser). However, for an effective separation all the delays must be truly random.

By transforming to the extended image space it is possible to distinguish events from

both the primary and secondary shots. This is described in detail in the following sections.

THE EXTENDED IMAGE SPACE

In RTM the conventional imaging condition is given by Eq 1 (Claerbout, 2001), whereby we assume events that are concurrent in time in both the source wavefield P_s and the receiver wavefield P_r correspond to a real reflection event. By correlating these fields at all times and summing all contributions we can create an image of the scattering field. By then summing over all shots we can build up an image of the subsurface:

$$I(x,y,z) = \sum_{i}^{nshots} \sum_{t} P_s(x,y,z,t;\mathbf{s}_i) P_r(x,y,z,t;\mathbf{s}_i). \tag{1}$$

Here $I(x,y,z)$ is the 3D image we are building, P_s is the source wavefield, P_r the receiver wavefield, and $\mathbf{s}_i$ a given shot.

For these actual reflection events to be concurrent in time, they must have been propagated through a correct velocity model. If this model was incorrect, then events will correlate to incorrect positions and possibly stack out when we sum over shots. A method of compensating for this is to perform extended imaging, where we calculate the image over a set of lags. These can be either spatial, temporal, or both. Eq 2 shows this equation when extending the image in both the x and y domains, by x_h and y_h. We can also extend in t and z, and any combination thereof:

$$I(x,y,z,x_h,y_h) = \sum_{i}^{nshots} \sum_{t} P_s(x+x_h,y+y_h,z,t;\mathbf{s}_i) P_r(x-x_h,y-y_h,z,t;\mathbf{s}_i). \tag{2}$$

The new image dimensions, x_h and y_h, are referred to as 'subsurface offsets.' It is possible to then perform a conversion to opening angle and to create angle gathers. The shape and focusing properties of the events in subsurface offset (or opening angle) are indicative of any model deficincies. This information be then be used to provide model updates in schemes like Wave Equation Migration Velocity Analysis (WEMVA) (Sava and Biondi, 2003) and Residual Moveout velocity analysis (RMO) (Zhang and Biondi, 2013).

Fig 7 shows an example of a simple scattering model which has been extended along the x-axis. The migration was performed in 3D and a slice along the y-axis was then windowed. Since the correct velocity model was used the events in subsurface offset are well focused at zero-offset. Fig 8 then shows a similar image for a more complex velocity model. Again all real reflection events are tightly focused in subsurface offset. Energy appearing at large subsurface offsets corresponds to noise in our image.

SEPARATION IN MODEL SPACE

Events in our data that do not correspond to our shot of interest (the shot we are migrating) will not be focused at zero subsurface offset. Our blended data can be described

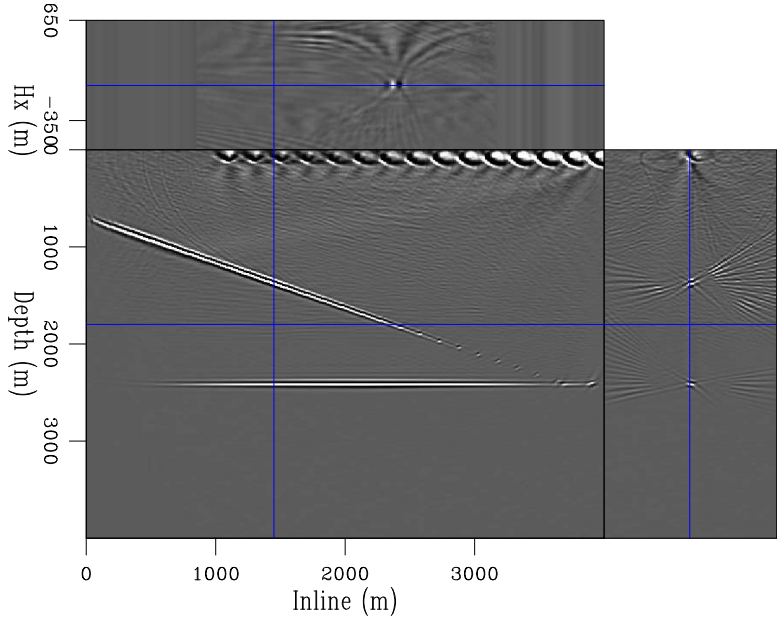


Figure 7: A slice of a simple scattering model with the imaging extended into subsurface offset along the x-axis. [**CR**] chris2/. imsimext

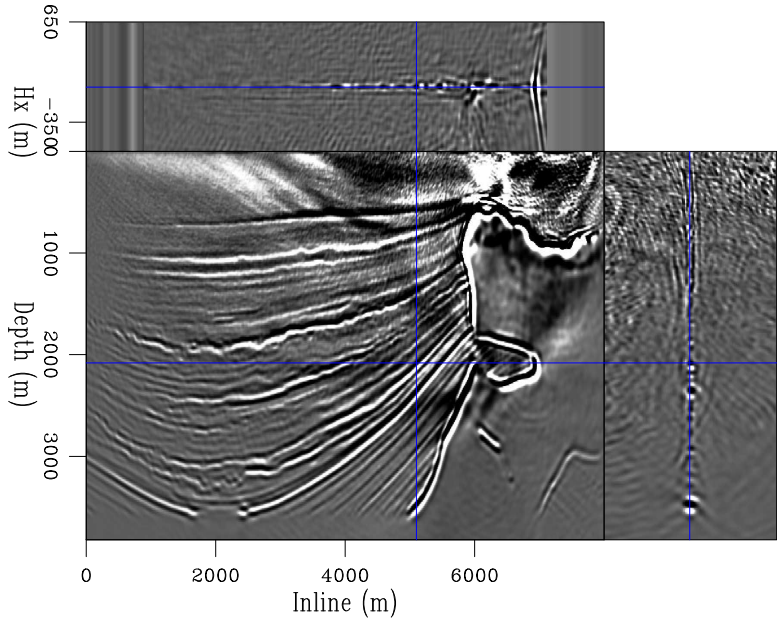


Figure 8: A slice of a complex scattering model with the imaging extended into subsurface offset along the x-axis. [**CR**] chris2/. imseamext

by Eq 3, where $\mathbf{L}_p$ describes our conventional/primary operator, $\mathbf{L}_s$ models the overlapping/secondary data and $\widetilde{\mathbf{L}}$ is our blended operator that combines the two. The blended data is $\widetilde{\mathbf{d}}$ and the model of interest is $\mathbf{m}$. The tilde sign will represent any sort of data or operator that is blended. In this case we will be using linearised Born modelling as our forward operator and RTM as our adjoint:

$$\widetilde{\mathbf{L}}\mathbf{m} = (\mathbf{L}_p + \mathbf{L}_s)\mathbf{m} = \widetilde{\mathbf{d}}_{est}. \tag{3}$$

The blended data can be simply described by Eq 4, where $\mathbf{d}_{obs}$ is the unblended data of interest (what we would like to recover) and $\mathbf{d}_s$ describes all the energy present from overlapping shots:

$$\widetilde{\mathbf{d}}_{obs} = \mathbf{d}_{obs} + \mathbf{d}_s. \tag{4}$$

A first attempt at data separation using extended images could be with simple filtering in subsurface offset. Fig 9 shows an image where 15 shots were randomly blended in groups of five and then migrated using an imaging condition as described in Eq 2. Fig 10 shows a similar for linear blending.

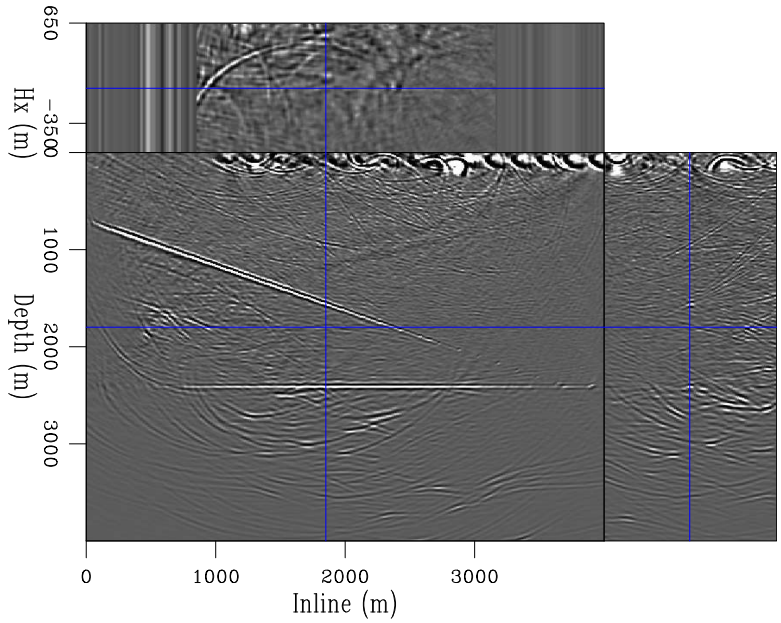


Figure 9: A simple model with randomly blended shots after migration into subsurface offset. [**CR**] chris2/. imsimextr

Both of these figures show that the secondary data has different focusing properties in subsurface offset, as expected. In both cases the primary data is well focused at suburface offset and relatively sparse, whereas events from the overlapping shots are spread over a variety of subsurface offsets. By focusing on these different attributes we can act to remove these 'secondary events.' Once we have a clean image to we can transform back to the data space, creating our equivalent unblended dataset.

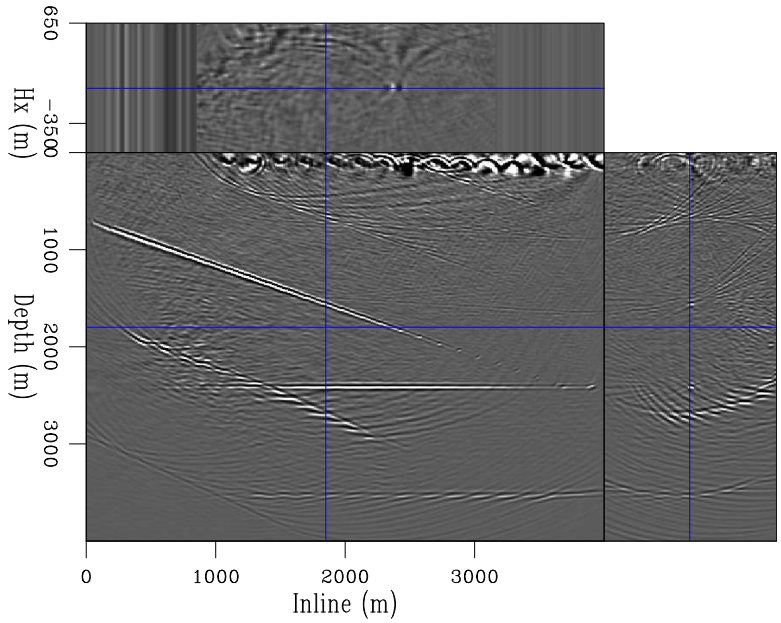


Figure 10: The same model as Fig 9 but with linearly blended data. [**CR**] chris2/. imsimextl

Another approach could be to break up the record and correct each shot to time zero, making it the shot-of-interest (the primary shot.) This process is equivalent to applying the adjoint of the operator that takes the conventional data and mixes it. Alternatively, we can adapt our operator $\mathbf{L}$ to focus on a given source within the blended record, and create a series of operators, say $\mathbf{L}_i$. By applying the adjoint of these operators (either individually or in groups) to the data we can create a series of model estimates, $\mathbf{m}_i$, as in Eq 5. Summing these models together will be the equivalent of applying the adjoint of the full blended operator (Eq 6).

$$\mathbf{m}_i = \mathbf{L}_i' \widetilde{\mathbf{d}}_{obs}. \tag{5}$$

$$\sum_i \mathbf{m}_i = \sum_i \mathbf{L}_i' \widetilde{\mathbf{d}}_{obs} = \widetilde{\mathbf{L}}' \widetilde{\mathbf{d}}_{obs}. \tag{6}$$

We now have a series of models, $\mathbf{m}_i$. Aspects of these model estimates that are consistent between other $\mathbf{m}_i$, will represent events which come from our primary data. Coherent events in these models which are not consistent will be from overlapping data being misplaced.

We can apply image matching to these sub-images to filter these inconsistent events. Methods described in Ayeni and Nasser (2009) or Williams and Bennamoun (1998) could be used to achieve a cleaner image than Eq 6 would have provided. Research into implementing these match filters, and into extended model space filtering, is being undergone currently.

SEPARATION BY LINEARISED INVERSION

Energy from secondary data could also be removed by using a full linearised inversion scheme. One option could be to simply treat the overlapping data as noise, $\widetilde{\mathbf{d}}_{obs} = \mathbf{d}_{obs} + noise$, and design an objective function Eq 7,

$$J(\mathbf{m}) = \|\widetilde{\mathbf{d}}_{obs} - \mathbf{L}\mathbf{m}\|_2^2 + \epsilon\|\mathbf{A}\mathbf{m}\|_2^2 \tag{7}$$

Where $\mathbf{A}$ is a 'model styling' term designed to remove non-primary energy and ϵ balances the two parts of the objective function. Since every additional, overlapping shot is as coherent as our primary shot, we can conjecture that such a scheme will have poor converge properites, especially if our velocity model is not well known. There will not be sufficiently contrasting sparsity or focusing differences between our primary and secondary data. Instead we could adapt our objective function to capture these overlapping shots, since we know all shot times and positions, and write it as Eq 8.

$$J(\mathbf{m}) = \|\widetilde{\mathbf{d}}_{obs} - (\mathbf{L}_p + \mathbf{L}_s)\mathbf{m}\|_2^2 + \epsilon\|\mathbf{A}\mathbf{m}\|_2^2 \tag{8}$$

Now some of the overlapping shots will be positioned correctly, and the model deficiencies will be due to crosstalk artifacts. These artifacts will reduce with iteration number. By choosing our operator $\mathbf{A}$ to penalise shots that do not focus at or cross zero subsurface offset this inversion will create a clean representation of $\mathbf{m}$, which can then be used to reconstruct $\mathbf{d}_{obs}$. Current research into designing $\mathbf{A}$ is being undergone, with application to both correct and incorrect velocity models.

CONCLUSIONS

Acquiring simultaneous source data can make sesimic surveys much more economical, especially for multiple vessel techniques such as WATS. Separating these data in preparation for conventional imaging can be challenging. However, powerful algorithms have been developed for the case when source timings are random. It has been shown herein that, even for linear time delays, primary and overlapping shots are distinguishable in the subsurface offset domain. By applying filtering or simple noise removal inversion it should be possible to sequentially remove this overlapping energy and recreate our data as if it was not blended. We can extend this to a full linearised inversion scheme, where our regularisation operator can act to remove all energy associated with the overlapping sources.

ACKNOWLEDGMENTS

Thanks to all the sponsors of the Stanford Exploration Project for their continued financial and intellectual support.

REFERENCES

Abma, R., T. Manning, M. Anis, J. Yu, and M. Foster, 2010, High quality separation of simultaneous sources by sparse inversion: 72nd EAGE Conference and Exhibition Abstracts, **72nd**.

Abma, R. and J. Yan, 2009, Separating simultaneous sources by inversion: 71st EAGE Conference and Exhibition Abstracts, **71st**.

Ayeni, G., A. Almomin, and D. Nichols, 2011, On the separation of simultaneous source data by inversion, chapter 4, 20–25.

Ayeni, G. and M. Nasser, 2009, Optimized local matching of time lapse seismic data: A case study from the gulf of mexico, chapter 790.

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

Berkhout, A., V. D. and G. Blacquiere, 2008, Processing of blended seismic data: SEG Technical Program Expanded Abstracts, **78th**.

Claerbout, J., 2001, Basic earth imaging: Society of Exploration Geophysicists.

Dai, W. and J. Schuster, 2009, Least-squares migration of simultaneous sources data with a deblurring filter: SEG Expanded Abstracts, **79th**, 2990–2993.

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

Hampson, G., S. J. H. F., 2008, Acquisition using simultaneous sources: SEG Technical Program Expanded Abstracts, **78th**.

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

Moore, I., B. Dragoset, T. Ommundsen, D. Wilson, C. Ward, and D. Eke, 2008, Simultaneous source separation using dithered sources, chapter 565, 2806–2810.

Sava, P. and B. Biondi, 2003, Wave equation migration velocity analysis by inversion of differential image perturbations, chapter 540, 2124–2127.

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

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

Williams, J. and M. Bennamoun, 1998, A non-linear filtering approach to image matching, chapter 1.

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

Multicomponent ambient seismic noise correlations at Valhall: Characterization and correlation of microseism noise in the ambient seismic field

Sjoerd de Ridder

ABSTRACT

The basis for successful application of passive seismic interferometry depends on the characteristic of the ambient noise. This paper introduces four continuous recordings made at the Valhall Life of Field Seismic (LOFS) Array. The ambient seismic field is characterized through creation and analysis of various spectra, focusing particularly on the microseism energy between 0.175 and 1.75 Hz. Beam forming shows how the microseism noise is generally incident equally strongly from all directions for long periods of time, with certain exceptions. Finally, seismic interferometry is used to create virtual seismic sources for all receivers at Valhall. By cross-correlating different components at different stations, I retrieve a full virtual seismic Green's matrix that reveals interface waves of Scholte and Love wave types. A first overtone of Scholte is observed as well.

INTRODUCTION

At Valhall in the Norwegian North Sea, a permanent recording array was installed in 2003 (Kommedal et al., 2004) for repeated active seismic surveying. This array, named the Life of Field Seismic (LOFS) array, has the capability to record continuously under a variety of circumstances. Conventional active seismic surveying provides highly detailed subsurface images, but at only at a few snapshots in time. Passive seismic interferometry creates virtual seismic sources from ambient noise (Wapenaar, 2004), thus providing a means to survey the subsurface continuously. Since the subsurface at Valhall is well known from active seismic (Barkved and Kristiansen, 2005; Sirgue et al., 2010), recordings made by the LOFS array provide an good opportunity to study the ambient seismic field and the application of seismic interferometry at a reservoir scale.

The quality of virtual seismic sources obtained from passive seismic interferometry depends on the characteristic of the ambient noise. Firstly, noise can be characterized by its spatially and temporally varying time-frequency spectrum. Secondly, the propagation direction of the waves composing the ambient field can be studied using beam forming experiments. Finally, I perform seismic interferometry and create virtual seismic sources for all receivers at Valhall. Cross-correlating different components at different stations retrieves a full virtual seismic Green's matrix. I finish with a modal analysis of the components of the virtual seismic Green's matrix.

NOISE RECORDINGS OF THE AMBIENT SEISMIC FIELD AT VALHALL

The LOFS installation has been in place since 2003 and can record continuously, but recordings are usually only permanently stored when controlled source seismic surveys are acquired. Artman (2006) suggested that several long and continuous recordings be stored and studied at frequencies above 2 Hz with the aim of passive imaging by seismic interferometry. Artman (2006) acquired two recordings:

- February 13$^{\text{th}}$, 2004: 29 hours, 29 minutes and 16 seconds starting UT 14:14:14.
- January 19$^{\text{th}}$, 2005: 6 hours, 39 minutes and 30 seconds starting UT 19:35:45.

Unfortunately, retrieving virtual seismic sources at higher frequencies by seismic interferometry is not straightforward. Later work by Dellinger (2008) revealed that seismic interferometry applied at frequencies below 2 Hz results in high quality virtual seismic Scholte-wave sources. This spurred the decision to remove the standard-acquisition low-cut filter and record while the remnants of tropical storm Laura moved over the North Sea generating ambient data rich in low frequencies.

- October 4$^{\text{th}}$, 2008: 2 hours starting at UT 14:11:24.

Finally a much longer recording was stored, again without low-cut filter, and made available for this study.

- December 22$^{\text{nd}}$, 2010: 5 days, 1 hour, 47 minutes and 20 seconds starting UT 21:16:39.

The recordings of 2004, 2005 and 2008 were made with a sampling rate of 250 samples per second, and the 2010 recording was made with 500 samples per second.

Figure 1 shows time slices of the data recorded in the vertical component of the geophone, after bandpasses conducted in the frequency domain with a Hann taper. At the lowest frequencies, $0.25 - 0.35$ Hz, we find an incoherent random wave field, known as microseism energy, that is not spatially aliased at the very low frequencies but spatially aliased at higher frequencies, $1.0 - 1.5$ Hz, as seen in Figure 1b. Figure 1c shows ample energy, between $5 - 10$ Hz, radiating from platforms installed at Valhall fields. At even higher frequencies, $15 - 25$ Hz in Figure 1d, we find almost planar wave fields incident from the direction of Ekofisk fields, and potentially originating there. They are found at intervals of 10 seconds, which is a regular interval for seismic shooting. All four time-slices are extracted at UT 0:17:00 December 24$^{\text{th}}$, 2010.

Frequency spectra as function of space

One way of characterizing the ambient seismic field is by investigating how the spectrum changes over space (McNamara and Boaz, 2005). A three hour continuous block of data was selected from the 2010 data (starting UT 0:17:18) during a period when the active seismic shots were silent (see next section). These 3 hours were divided into 2.5-minute windows

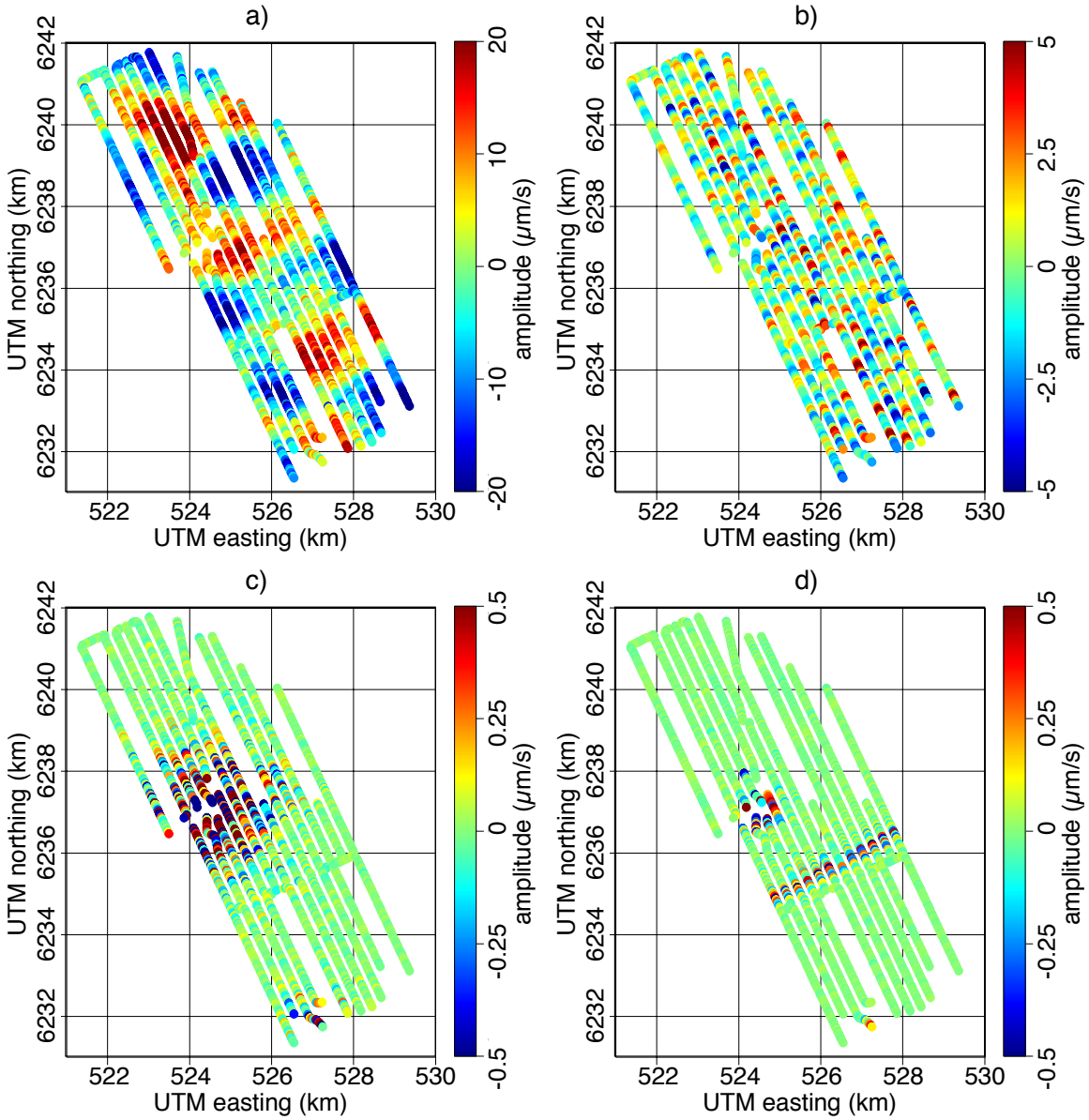


Figure 1: Amplitude maps extracted at UT 0:17:00 December 24$^{\text{th}}$, 2010. a) $0.25-0.35$ Hz, b) $1.0-1.5$ Hz, c) $5-10$ Hz, d) $15-25$ Hz. [**CR**] sjoerd1/. amp-maps-new

with 50% overlap. The time-frequency spectrum was computed for each time-window and averaged over all time windows. Figure 2 contains 9 maps for frequencies $10^{-\frac{2}{3}}$, $10^{-\frac{1}{3}}$, $10^{0}, 10^{\frac{1}{3}}$, $10^{\frac{2}{3}}$, $10^{1\frac{1}{3}}$, $10^{1\frac{2}{3}}$ and 10^{2} Hz.

The spectral energy below 2 Hz is very homogeneous over the field, Figures 2a-c, but above 2 Hz the energy is stronger near the platform constructions in the center and south ends of the array, particularly visible in Figure 2f.

Frequency spectra as function of time

The spectrum is not constant over time, however, and to investigate its behavior we select all stations in a radius of 1.75 km centered at $(522, 6242)$ km. This location was selected after observing in Figure 2 that the spectrum is not dominated by platform-generated noise. All data in the four data sets were divided into 2.5-minute windows with 50% overlap, and the time-frequency spectrum was computed for each time-window and averaged over all selected stations. The spectrum is then shown in Figure 3 as a function of the center of each 2.5 minute window.

Observe how the spectrum varies for the different recordings; this reflects both the filter settings and transient nature of the ambient field over long periods. Furthermore, the spectrum is not constant during each recording. Notice, in Figure 3d above 10 Hz, the episodic nature of the seismic-survey shooting happening to the north-west of Valhall. When recording for long enough periods, the low frequency microseism noise in frequencies below 2 Hz can be seen to vary considerably over time (Figures 3c and 3d).

beam forming experiments

Because the station spacing is approximately 150 m between the cables and approximately 50 m between stations along the cable, the surface waves are aliased for all but the lowest of frequencies. We select all stations in a radius of 1 km centered at $(526.5, 6236.4)$ km and bandpass the data with a frequency domain Hann taper between $0.45 - 0.65$ Hz. Beam forming was conducted after slant stacking and averaging the absolute value of the $\tau - p$ domain over 30-minute time windows under a time domain Hann taper. Figures 4 to 7 contain beam forming experiments for all four data sets.

Beam forming of all data sets shows how the microseism noise is generally incident equally strongly with direction for long periods of time. In Figure 4 the microseism energy slowly, but continuously, increases in strength. This is not a diurnal cycle, because the 2004 data set spans more then 24 hours. The observation that microseism noise does not necessarily reflect diurnal cycles is corroborated by Figure 3c.

Figure 5 shows that for the 6 hours recorded in 2005, the microseism energy is remarkably homogeneous in direction (more so then for the other 3 data sets). This contributed to the successes achieved by Mordret et al. (2013a,b) in his studies of this data set. Figure 7 shows that over the course of 5 days, the microseism exhibited periods during which the noise came in particularly strongly from certain directions (Figure 7g), while during other periods there was no single direction standing out. Studying the beam forming results carefully, one observes an inner ring at a slowness corresponding to a velocity of about

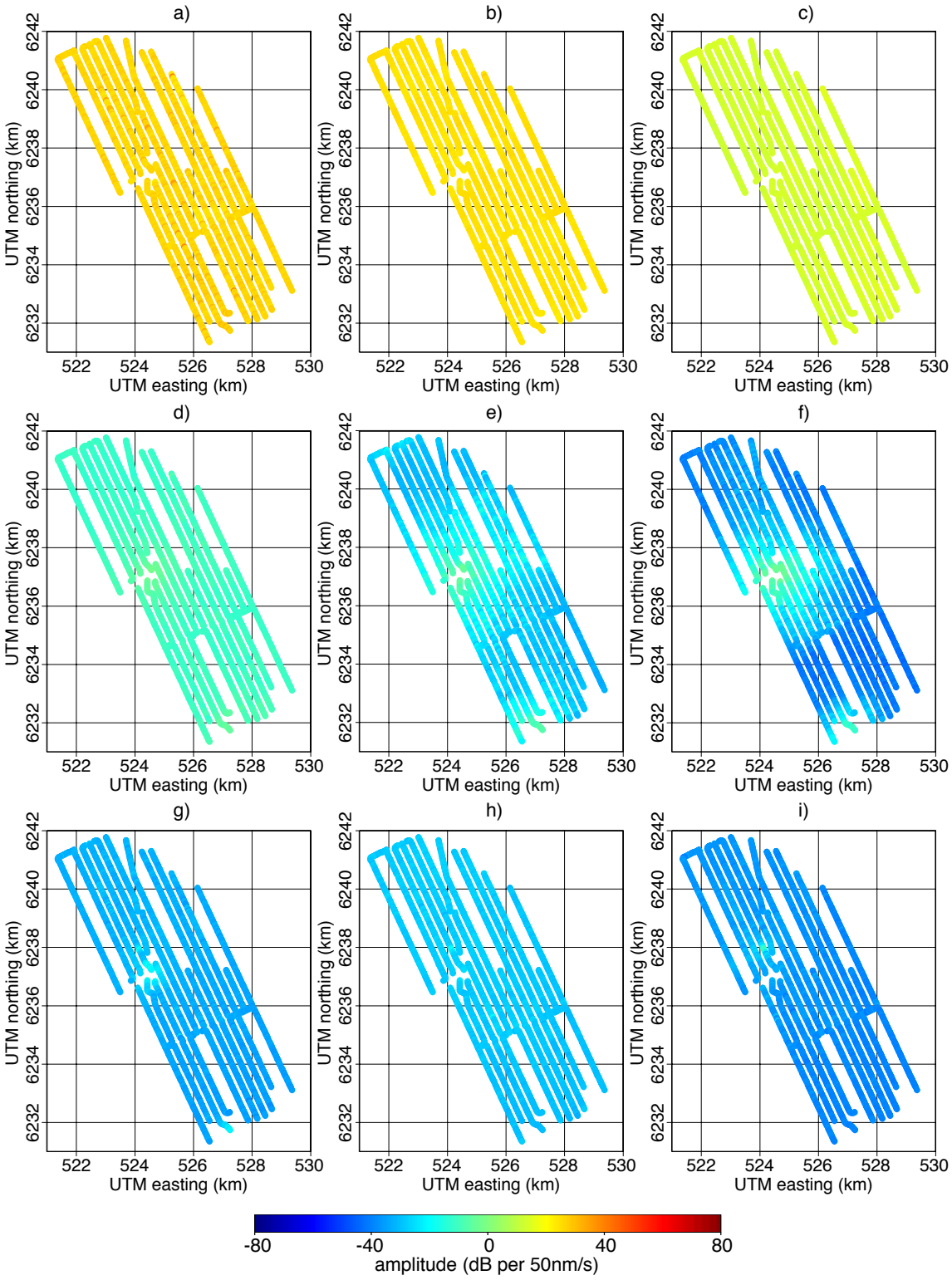


Figure 2: Spectra from data in December 2010. a) $10^{-\frac{2}{3}}$ Hz, b) $10^{-\frac{1}{3}}$ Hz, c) 10^{0} Hz, d) $10^{\frac{1}{3}}$ Hz, e) $10^{\frac{2}{3}}$ Hz, f) 10^{1} Hz, g) $10^{1\frac{1}{3}}$ Hz, h) $10^{1\frac{2}{3}}$ Hz, i) 10^{2} Hz. [**CR**] sjoerd1/. spectra-maps-new

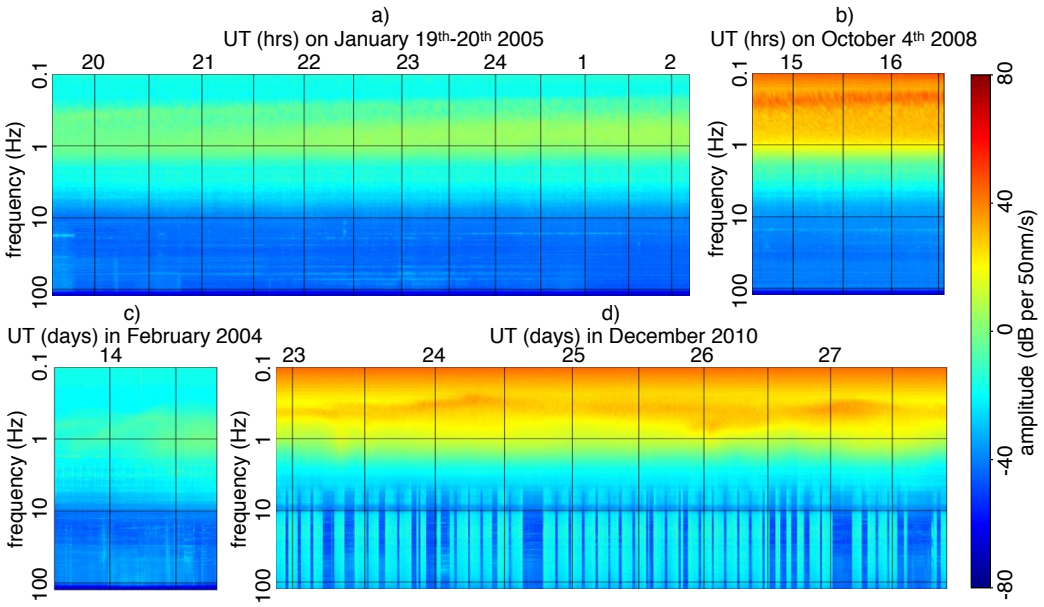


Figure 3: Spectragrams of vertical component of particle velocity for four different continuous passive recordings. a) 6 hours, 39 minutes and 30 seconds starting UT 19:35:45 on January 19$^{\text{th}}$ 2005, b) 2 hours starting UT 14:11:24 on October 4$^{\text{th}}$ 2008, c) 29 hours, 29 minutes and 16 seconds starting UT 14:14:14 on February 13$^{\text{th}}$ 2004, d) 5 days, 1 hour, 47 minutes and 20 seconds starting UT 21:16:39 on 22$^{\text{nd}}$ December 2010. [**CR**] sjoerd1/. spectra4geo2013-new

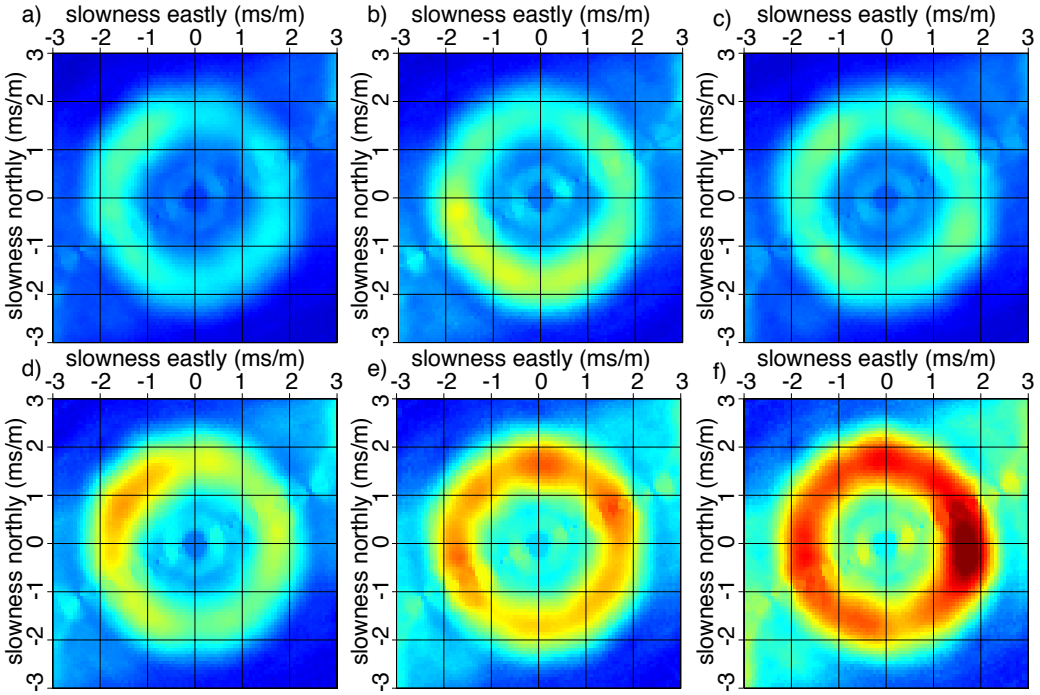


Figure 4: Beam forming for half-hour windows in February 13$^{\text{th}}$ 2004; centered around a) 16:00, b) 20:00 and half-hour windows in February 14$^{\text{th}}$ 2004, centered around c) 0:00, d) 4:00, e) 8:00, f) 12:00. [**CR**] sjoerd1/. artman-beam-new

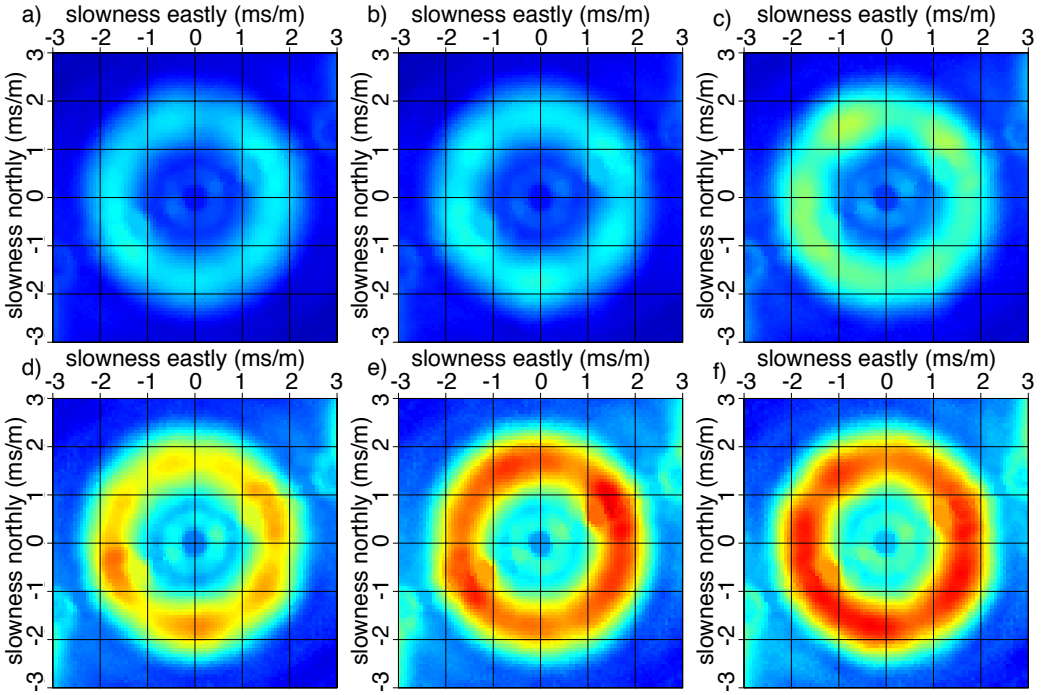


Figure 5: Beam forming for half-hour windows in January 19$^{\text{th}}$ 2005; centered around a) 20:00, b) 21:00, c) 22:00 d) 23:00 and half-hour windows in January 20$^{\text{th}}$ 2005, centered around c) 1:00, d) 2:00 [**CR**] sjoerd1/. jianhua-beam-new

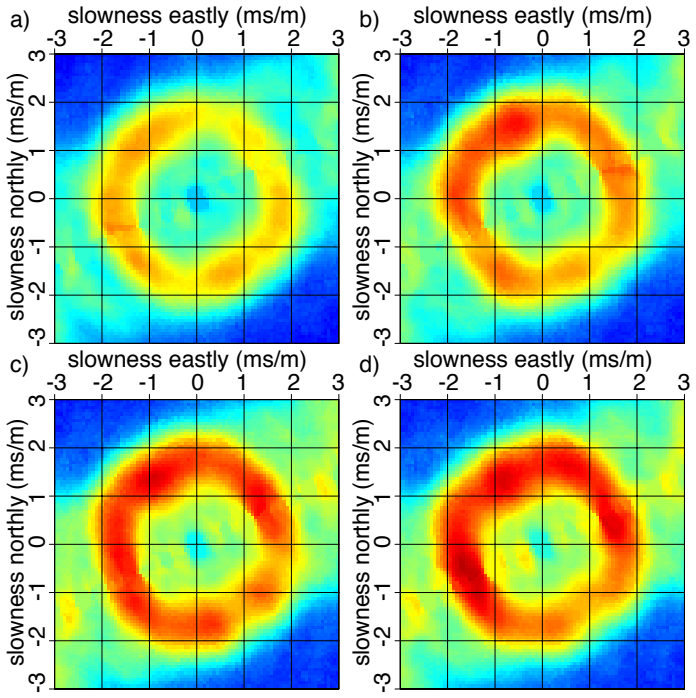


Figure 6: Beam forming for half-hour windows in October 4$^{\text{th}}$ 2008, half hours are centered around; a) 14:30, b) 15:00, c) 15:30, d) 16:00. [**CR**] sjoerd1/. laura-beam-new

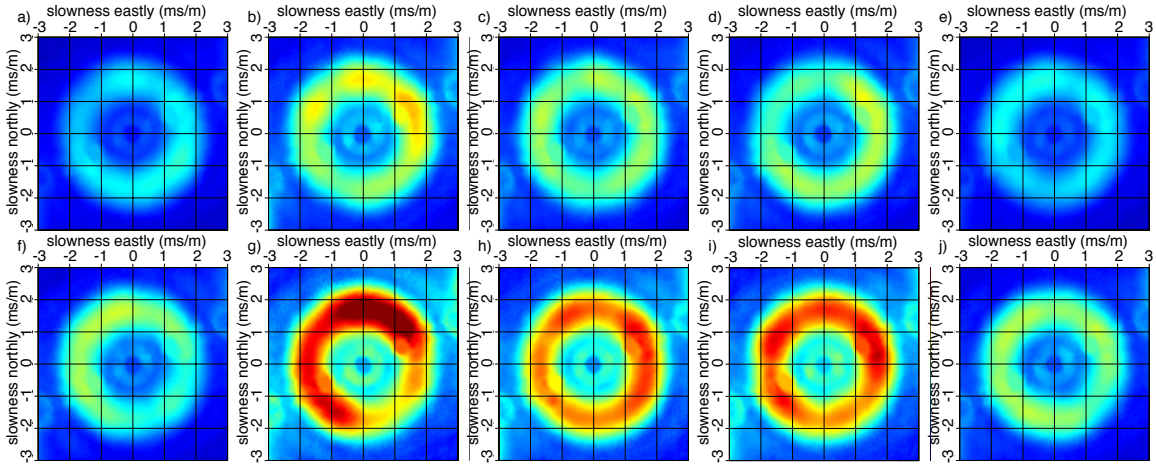


Figure 7: Beam forming for half-hour windows centered around a) 0:00 and b) 12:00 on December 23$^{\text{rd}}$, c) 0:00 and d) 12:00 on December 24$^{\text{nd}}$, e) 0:00 and f) 12:00 on December 25$^{\text{th}}$, g) 0:00 and h) 12:00 on December 26$^{\text{th}}$, i) 0:00 and j) 12:00 on December 27$^{\text{th}}$. [**CR**] sjoerd1/. joseph-beam-new

2 km/s, and exhibiting an aliasing pattern spaced closer than the alias ring of the dominant mode (the alias can be observed in the corners of Figure 7b). This could be a manifestation of a higher mode traveling faster then the fundamental mode.

CORRELATION METHODOLOGY

Passive seismic interferometry is a technique that cross-correlates seismic recordings at two stations to form a signal, an estimated Green's function (EGF), as if one of the stations were a seismic source (Wapenaar and Fokkema, 2006). Cross-correlations of passive seismic recordings between all possible station pairs in an array then creates a full virtual seismic survey, $\mathbf{D}(\omega, x_r, x_s)$:

$$\mathbf{D}(\omega, x_r, x_s) = \mathbf{r}(\omega, x_r)\ \mathbf{r}(\omega, x_s)^{\dagger}, \tag{1}$$

where $\mathbf{r}(\omega, x)$ is a vector containing passive seismic recordings at all stations, and $\dagger$ denotes complex conjugation and transposition. By cross-correlating different components of particle velocity we can retrieve all elements of the Green's matrix (Wapenaar and Fokkema, 2006). However, this technique is valid only when the energy in the ambient seismic field satisfies a condition known as energy equipartioning. In practice, this requirement limits the application of seismic interferometry to certain frequency regimes. The ambient seismic field at low frequencies ($0.18 - 1.75$ Hz) is dominated by the double-frequency microseism peak, a source of seismic energy that satisfies the requirement of energy equipartitioning and can be utilized for seismic interferometry (Stewart, 2006; Dellinger and Yu, 2009). The LOFS array has geophone stations that consist of one in-line (with the cable) component and two perpendicular (to the cable) components. Station orientations were estimated from active seismic and provided with the data by BP. Using the orientations, the recordings are rotated to a vertical, east ,and north component. The recorded data in all three components (north, east, and vertical) of particle velocity was bandpass filtered using a frequency domain taper with a flat response for 0.2-1.5 Hz, and as a Hann-taper extending from 0.175 Hz to 1.75 Hz.

The data was then divided into 30-minute segments with 50% overlap. For each segment, all stations and all components were cross-correlated. The resulting cross-correlations were stacked for all 30-minute segments, resulting in virtual seismic source and receiver matrices between all stations.

Figure 8 contains an example of a virtual seismic source in vertical particle velocity, located near the center of the array. This example was generated using all data in the 2010 recording. Each frame is a snapshot corresponding to a certain correlation-time lag; negative time lags correspond to the acausal EGF, while positive time lags correspond to the causal EGF. Note the good retrieval of both causal and anti-causal EGF's, without imposing symmetry conditions. This is due to the azimuthal homogeneity of the directions in the ambient seismic field as observed in the beam forming experiments in Figure 7.

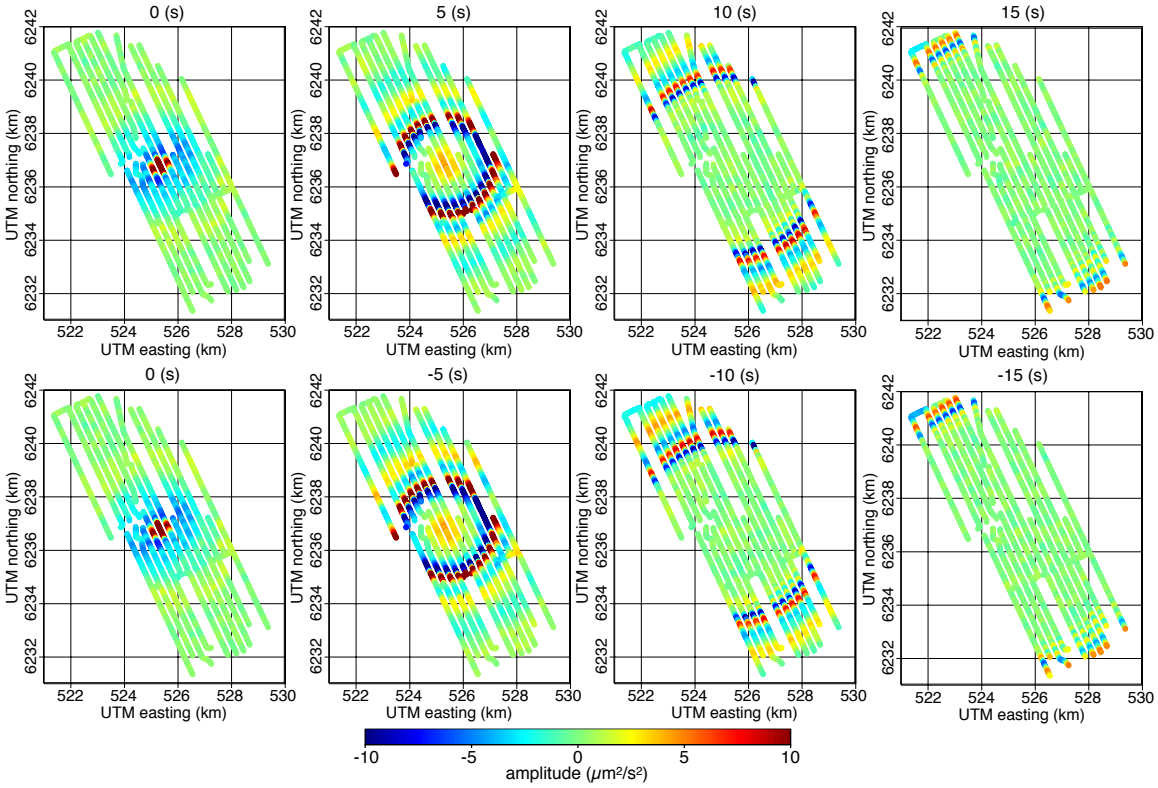


Figure 8: A virtual seismic source generated by processing all of the 2010 recording. The top row corresponds to positive correlation-time lags. The bottom row contains the corresponding negative correlation-time lags. [**CR**] sjoerd1/. movie-v-snaps-new

Rotation NEV to RTV

A coordinate system based on north, east and vertical components is not the most natural coordinate system to study the EGF matrices. In a perfectly stratified media, the surface wave modes split into Love and Rayleigh-Scholte wave modes. Love waves have transverse polarization and only appear in the transverse components, and Scholte wave appear in the vertical and radial components. Although we know that the subsurface at Valhall is anisotropic, and there are strong lateral inhomogeneities (Barkved and Kristiansen, 2005;

Sirgue et al., 2010), these effects are secondary and the virtual seismic sources will, to first order, consist of Love and Scholte wave modes.

We exploit this first order behavior and rotate each virtual seismic source to a cylindrical coordinate system centered around the source. Thus all virtual seismic source matrices were transformed from a coordinate system with north, east, and vertical components, to a cylindrical coordinate system centered at the source station, with radial, tangential, and vertical components. The result of this transformation is shown in Figure 9. The left-group in Figure 9 shows time-slices of G_{nn}, G_{en}, G_{vn}, G_{ee}, G_{ve} and G_{vv} before rotation, while the right-group in Figure 9 shows the same time-slices after rotation into G_{rr}, G_{rt}, G_{rv}, G_{tt}, G_{tv} and G_{vv}. After this rotation, the polarity and amplitude of the EGF's should no longer depend on geographic direction. Directionality of the virtual seismic sources can however still caused by the directionality of the energy in the ambient seismic field and by subsurface lateral inhomogeneities.

Notice, in Figure 9, how the polarity of the observed wave front changes with azimuth for the various elements of the EGF matrix: In the G_{vv}, G_{nn} and G_{ee} elements the polarity is equal in all azimuths from the source, but the G_{nn} and G_{ee} elements are relatively weak in the east-west and north-south directions from the source, respectively. In the G_{vn} and G_{ve} elements the polarity flips in the north-south and east-west directions, respectively. In addition, G_{en} displays a butterfly pattern of flipped polarities and weak radiation. After rotation, the G_{vv} element remains the same. The G_{tt}, G_{rr} and G_{rv} elements are now radiating with equal polarity and more or less equal strength in all directions from the source. The G_{vv}, G_{vr} and G_{rr} elements are dominated by Scholte waves and the G_{tt} element is dominated by Love waves, while the G_{rt} and G_{tv}, expected to be zero in a perfectly stratified medium, display a complicated mixture of converted wave modes. Notice how the amplitudes in the G_{rr} G_{tt} and G_{rt} element are relatively weak, while the amplitude in the G_{tv} element is still relatively strong, because the ambient noise field is stronger in the vertical component then in the horizontal components.

W-P spectra

This virtual seismic survey was next transformed to a midpoint and radial-offset domain, and all offsets were selected for a group of midpoints in the southern part of the array. Dispersion images (Figure 10) are calculated as the amplitude in the Radon ($\omega - p$) domain, balanced over frequencies. In the vertical-vertical, radial-radial and radial-vertical elements, these waves are dominated by fundamental-mode Scholte waves. The first overtone is also distinguishable and is more evident in the radial-radial than in the vertical-vertical element. The transverse-transverse element is dominated by Love waves that travel at a higher velocity (lower slowness) than the fundamental-mode Scholte waves. The transverse-vertical and transverse-radial components are much weaker than the other elements (before normalization) and do not contain well-defined modes.

CONCLUSIONS

Strong microseism energy, although varying in strength over time, is present in the vertical component of the geophones for all four LOFS recordings. The low frequency microseism

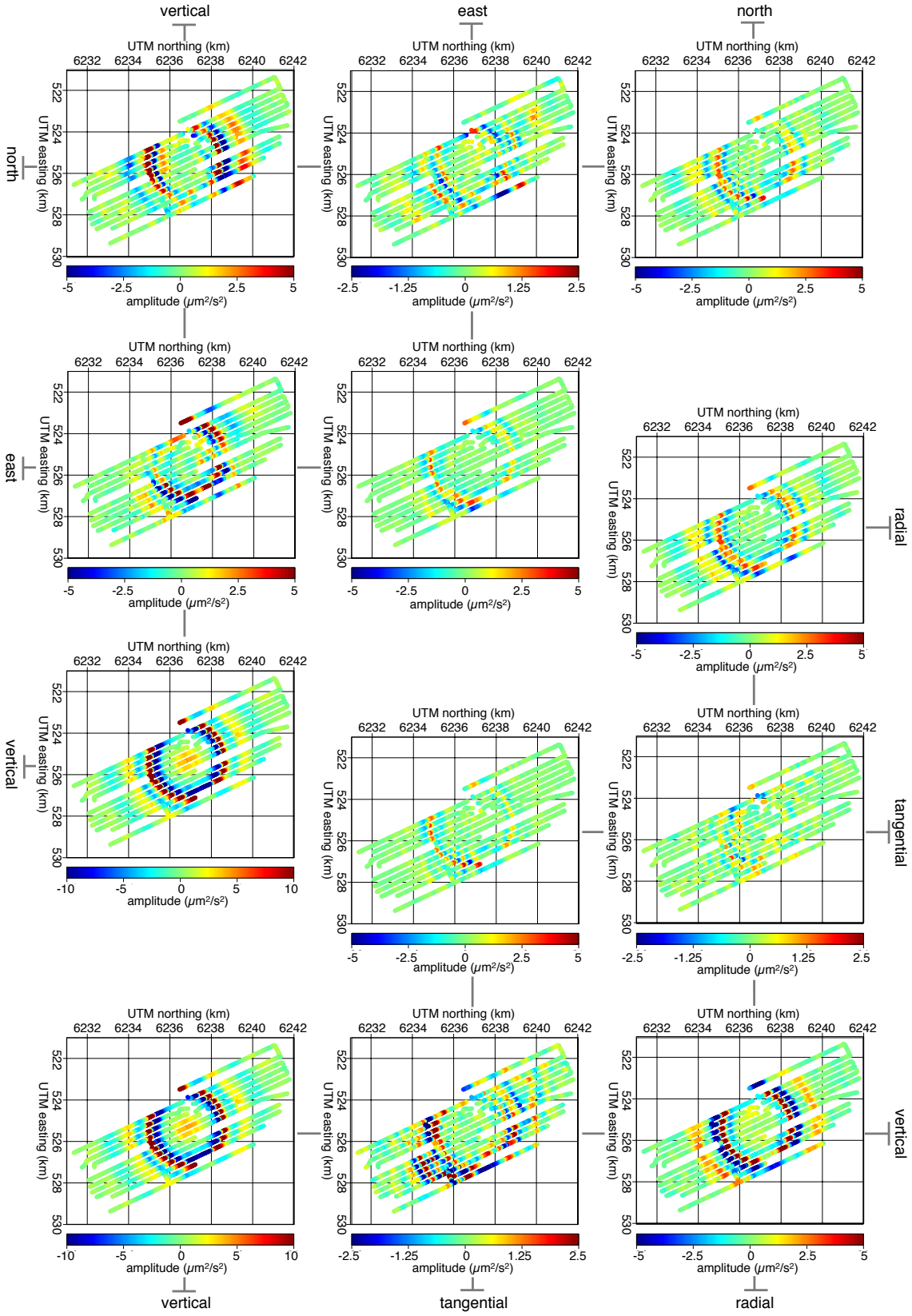


Figure 9: Time slices of virtual seismic source. [**CR**] sjoerd1/. sourcematrix-new

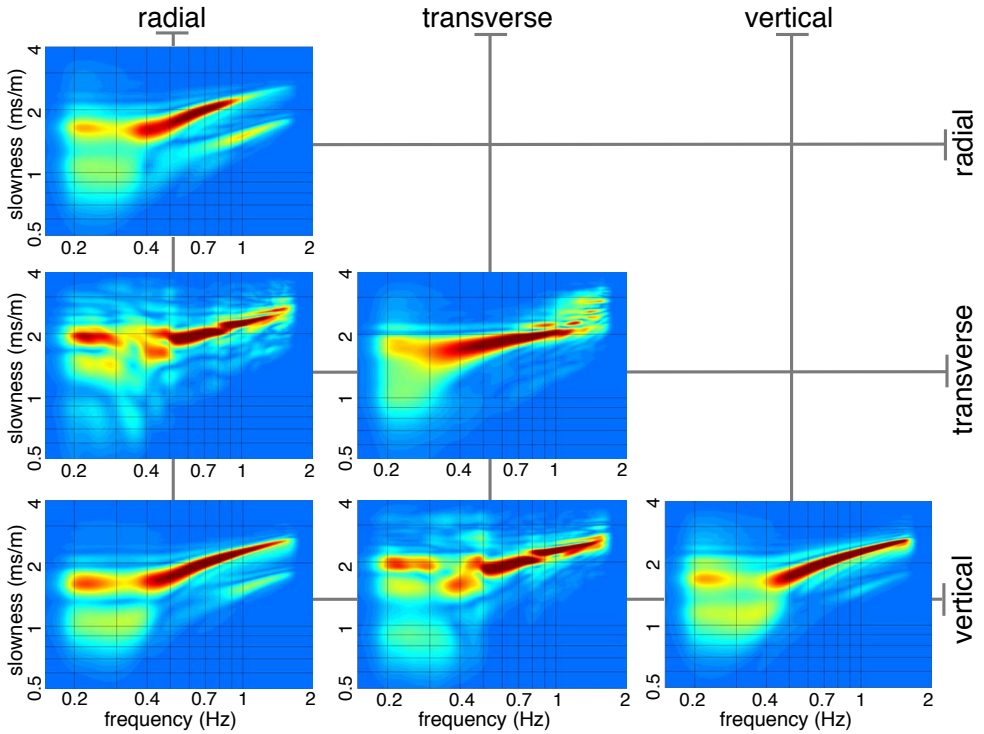


Figure 10: Dispersion images for all elements of the virtual seismic source matrix. [**CR**] sjoerd1/. dispersionmatrix-new

energy is best recorded with the low-cut recording filter (standard for active seismic acquisitions) turned off. At Valhall, the microseism energy is generally omnidirectional, but there are time periods that are an exception. Seismic interferometry is successfully applied using all components of geophones to create a virtual seismic Green's matrix. This matrix has to be rotated to a coordinate system with radial, tangential and vertical components with respect to the source-receiver couple in order to be well interpreted. The vertical-vertical, radial-radial, and radial-vertical components are dominated by fundamental mode Scholte waves. A first overtone is visible as well, more strongly in the radial-radial component then vertical-vertical component. The tangential-tangential component is dominated by Love waves traveling faster then Scholte waves. These waves, extracted with seismic interferometry, have wavelengths on the order of several hundred meters to several kilometers and provide sensitivity to the subsurface below the seafloor at Valhall.

ACKNOWLEDGEMENTS

Thanks to Olav Barkved, Biondo Biondi, Bob Clapp, Joe Dellinger, Jesse Lawrence, Stew Levin and Dave Nichols for helpful discussions and suggestions. Thanks to Adam Halpert for many comments and suggestions to this manuscript. Thanks to BP and the partners of the Valhall Field (BP Norge and Hess Norge) for permission to publish this paper. Thanks to the sponsors of the Stanford Exploration Project for financial support.

APPENDIX I

Fourier domain and spectra

Fourier transformations are powerful tools to analyses passive seismic recordings, not in the least because of their ability to characterize an infinitely long time series by a compact spectrum. We employ the following definition for the continuous-time Fourier transformation:

$$G(f) = \mathcal{F}_t \{g(t)\} = \int_{-\infty}^{\infty} \mathrm{e}^{-i2\pi ft} g(t) \mathrm{d}t, \tag{2}$$

and its inverse

$$g(t) = \mathcal{F}_t^{-1} \{G(f)\} = \int_{-\infty}^{\infty} \mathrm{e}^{i2\pi ft} G(f) \mathrm{d}f, \tag{3}$$

where $G(f)$ denotes the temporal Fourier domain counterpart of the function $g(t)$, t and f denote time and temporal frequency respectively. The temporal frequency f is regularly substituted by angular frequency $w = 2\pi f$. In practice we do not recording continuous and infinitely long time series, but instead record a limited number of discrete samples. Conjunctly, the Fourier spectrum is not continuous; the discrete time-series is decomposed into a discrete number of Fourier coefficients. For a time series containing N samples with sampling Δt, the time and frequency are discretised as:

$$t \rightarrow \mathbf{f} = t_n = n\Delta f \quad \text{with} \quad n = -N/2, -N/2+1, ..., N/2-1, \tag{4}$$

$$f \rightarrow \mathbf{f} = f_m = m\Delta f \quad \text{with} \quad m = -N/2, -N/2+1, ..., N/2-1, \tag{5}$$

where the frequency sampling, Δf, is related to N and Δt as shown below. The continuous time and frequency functions $g(t)$ and $G(f)$ are consequently discretised as

$$g(t) \rightarrow \mathbf{g}(t_n) \quad \text{with} \quad n = -N/2, -N/2+1, ..., N/2-1, \tag{6}$$

$$G(f) \rightarrow \mathbf{G}(f_m) \quad \text{with} \quad m = -N/2, -N/2+1, ..., N/2-1. \tag{7}$$

Substituting these into the definition of the forward and inverse Fourier transformation, Equations 2 and 3, we find

$$\mathbf{G}(f_m) = \sum_{n=-N/2}^{N/2-1} \mathbf{g}(t_n) e^{-i2\pi(\Delta t \Delta f \ n \ m)} \Delta t, \tag{8}$$

$$\mathbf{g}(t_n) = \sum_{m=-N/2}^{N/2-1} \mathbf{G}(f_m) e^{i2\pi(\Delta t \Delta f \ n \ m)} \Delta f. \tag{9}$$

Notice that the periodicity of the Fourier kernel relates Δt, Δt and N as:

$$N \Delta f \Delta t = 1 \tag{10}$$

Although Equations 8 and 9 are the discrete versions of Equations 2 and 3, more modifications are needed to be find the Discrete Fourier transformations as commonly applied by packaged such as FFTW. First, replace the negative counters with $n \rightarrow n + N$ and $m \rightarrow m + N$ to rearrange the summations to run over positive indices only. The summations now run from 0 to 2π instead of from π to π. As a consequence the DFT algorithms

expect first all the positive time-lags and then all the negative time-lags, and a similar rearrangement of the frequency axis. Secondly, scale the variable G_m by Δt and substitute the factor $\Delta f \Delta t$ with $1/N$, now the DFT algorithm operates independently of Δt and Δt while maintaining invertability. Finally we have:

$$\mathbf{G}(f_m) = \sum_{n=0}^{N-1} \mathbf{g}(t_n) e^{-i2\pi \frac{n}{N} \frac{m}{N}}, \tag{11}$$

$$\mathbf{g}(t_n) = \frac{1}{N} \sum_{m=0}^{N-1} \mathbf{G}(f_m) e^{-i2\pi \frac{n}{N} \frac{m}{N}}. \tag{12}$$

If we which to find the discrete version of the continuous Fourier transformations as defined in Equations 2 and 3 using the FFTW package, we need to scale the outcome of the forward transformation by Δt and that of the inverse transformation by $\frac{1}{\Delta t}$.

After Fourier transformation, a time-series can be characterized by its amplitude and phase spectra of the complex Fourier coefficients. And the energy in a time-series is characterized by the power spectrum. If a time series of length $N\Delta t = T$, measures an electrical voltage in V, the application of the Discrete Fourier transformation, Equation 8, transforms the unit into V. The amplitude spectrum now corresponds to the amplitudes of the time domain sinusoidal basis functions that are used to decompose the time series. For time-series that are purely real, the symmetry of the resulting Fourier domain spectrum can be exploited by multiplying the amplitude spectrum by two and analyzing the spectrum for positive frequencies only. Thus the amplitude spectrum is defined by:

$$\mathbf{S}(f_m) = 2\,|\mathbf{G}(f_m)| \quad \text{with} \quad m = 0, 1, ..., N/2 - 1 \tag{13}$$

The power spectrum is defined as the square of the Fourier spectrum:

$$\mathbf{P}(f_m) = 2\,(\mathbf{G}(f_m)\mathbf{G}(f_m)) \,\text{with} \quad m = 0, 1, ..., N/2 - 1 \tag{14}$$

It represents the energy in signal normalized per $T(s)$ used for the DFT. The units of amplitude are now in V^2.Finally Parceval's theorem will require

$$\Delta t \sum_{n=-N/2}^{N/2-1} \mathbf{g}(t_n) = \Delta f \sum_{n=-N/2}^{N/2-1} \mathbf{G}(f_m) \text{with} \quad n = 0, 1, ..., N/2 - 1 \tag{15}$$

which validates if the scaling was performed correctly.

Amplitude and power spectra have a large dynamic range and are therefore often studied on a logarithmic scale with a reference variable. For a power spectrum we define a decibel as:

$$\mathbf{P}(f_m)[dB] = 10\text{Log}_{10}\left(\frac{\mathbf{P}(f_m)[nm/s]}{\alpha_{ref}}\right). \tag{16}$$

For an amplitude spectrum we define a decibel as

$$\mathbf{S}(f_m)[dB] = 10\text{Log}_{10}\left(\frac{\mathbf{S}^2(f_m)[nm/s]}{\alpha_{ref}^2}\right) = 20\text{Log}_{10}\left(\frac{\mathbf{S}(f_m)[nm/s]}{\alpha_{ref}}\right) \tag{17}$$

where the factor 20 derives from the square to keep decibel levels similar for amplitude and power spectra. The unit is typically denoted as (dB re $1\alpha_{ref}$), where α_{ref} is the reference level. Standard reference levels for pressure in are $p_{ref} = 20[\mu Pa]$ (in air) and $p_{ref} = 1[\mu Pa]$ (under water) and for particle velocity $v_{ref} = 50[nm/s]$.

Instead of looking at the power spectrum of one signal we could study the cross-spectrum between two signals, **G** and **H**:

$$\mathbf{C}(f_m) = 2\left(\mathbf{G}(f_m)\mathbf{H}(f_m)\right) \tag{18}$$

After inverse Fourier domain this yields a cross-correlation function, which is best normalized by time to be insensitive to the amount of time cross-correlated). The unit of this cross-correlation signal is now V^2.

Finally we employ the following definition for the spatial Fourier transformation:

$$G(\xi) = \mathcal{F}_x\left\{g(x)\right\} = \int_{-\infty}^{\infty} \mathrm{e}^{i2\pi\xi x} g(x)\mathrm{d}x, \tag{19}$$

and its inverse

$$g(x) = \mathcal{F}_t^{-1}\left\{G(x)\right\} = \int_{-\infty}^{\infty} \mathrm{e}^{-i2\pi\xi x} G(x)\mathrm{d}\xi. \tag{20}$$

The spatial frequency ξ is regularly substituted by the wave number $k = 2\pi\xi$. Equations 19 and 20 are discretised similarly as Equations 19 and 20.

REFERENCES

Artman, B., 2006, Imaging passive seismic data: Geophysics, **71**, SI177–SI187.

Barkved, O. I., and T. Kristiansen, 2005, Seismic time-lapse effects and stress changes: Examples from a compacting reservoir: The Leading Edge, **24**, 1244–1248.

Dellinger, J., 2008, Low frequencies using conventional sensors: "sign-bit" recording revisited: SEG Technical Program Expanded Abstracts, **27**, 149–153.

Dellinger, J. A., and J. Yu, 2009, Low-frequency virtual point-source interferometry using conventional sensors: 71st Meeting, European Associated of Geoscientists and Engineers, Expanded Abstracts, X047.

McNamara, D., and R. Boaz, 2005, Seismic noise analysis system using power spectral density probability density functions: A stand-alone software package: USGS Open-File Report, 1–14.

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

Mordret, A., N. M. Shapiro, S. Singh, . Roux, J. P. Montagner, and O. I. Barkved, 2013b, Azimuthal anisotropy at valhall: the helmholtz equation approach: Geophys. Res. Let.

Sirgue, L., O. I. Barkved, J. Dellinger, J. E. U. Albertin, and J. H. Kommedal, 2010, Full waveform inversion: the next leap forward in imaging at Valhall: First Break, **28**, 65–70.

Stewart, P., 2006, Interferometric imaging of ocean bottom noise: SEG Technical Program Expanded Abstracts, **25**, 1555–1559.

Wapenaar, K., 2004, Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation: Phys. Rev. Lett., **93**, 254301–1 – 254301–4.

Wapenaar, K., and J. Fokkema, 2006, Green's function representations for seismic interferometry: Geophysics, **71**, SI33–SI46.

Daily reservoir-scale subsurface monitoring using ambient seismic noise

Sjoerd de Ridder and Biondo Biondi

ABSTRACT

Seismic interferometry is applied to continuous seismic recordings spanning five days and over 2200 stations at the Valhall Life-of-Field Seismic (LOFS) array in the Norwegian North Sea. We retrieve both fundamental-mode and first-overtone Scholte-waves by cross-correlation. Ambient-seismic-noise tomography (ASNT) using the vertical component of this dense array produces group-velocity maps of fundamental-mode Scholte waves with high repeatability from only 24 hours of recording. This repeatability makes daily reservoir-scale near-surface continuous monitoring of the subsurface feasible. Such monitoring may detect production-related changes over a long time-scale (months to years) and may be useful for early detection of short time-scale hazards (days to weeks) such as migrating gases and fluids. We validate our velocity maps by comparing them with maps obtained independently from controlled-source data.

INTRODUCTION

Reservoir monitoring using controlled-source seismic surveying is a well-established technique. One survey provides an image of the subsurface at one given time. Repeated seismic surveys provide differential seismic images that show how the subsurface changed over time. A permanent ocean-bottom-cable (OCB), called the Life-of-Field Seismic (LOFS) array (Kommedal et al., 2004), was installed just under the seabed at the Valhall Field in the Norwegian North in 2003. Conventional controlled-source seismic surveying provides subsurface images after expensive seismic acquisition, processing, and imaging procedures, and only at a few snapshots in time. However, the LOFS array can record continuously in any weather conditions, thus providing a wealth of data about the ambient seismic field.

Aki (1957) first linked the azimuthally averaged cross-spectrum of micro-tremors to the dispersion of the fundamental-mode Rayleigh wave. Later, Claerbout (1968) devised acoustic daylight imaging, and established a relationship between the autocorrelation of transmission responses to the reflection response of a 1D medium. This technique, turning ambient seismic recordings into virtual seismic sources, has been generalized to 3D and coined seismic interferometry; yielding estimated Green's functions (EGFs) by cross-correlation of ambient seismic recordings (Wapenaar and Fokkema, 2004). Convincing applications of this phenomenon have been made in solar physics (Duvall et al., 1993), laboratory acoustics (Weaver and Lobkis, 2001, 2002) and seismology (Campillo and Paul, 2003; Shapiro and Campillo, 2004). Seismologists have used ambient-seismic-noise tomography (ASNT) extensively in regional and global seismic studies to extract surface waves and image the subsurface (Shapiro et al., 2005; Gerstoft et al., 2006; Yao et al., 2006; Lin et al., 2008).

Accepted for publication in Geophysical Research Letters

These studies use stations that are spaced up to several hundred kilometers apart, and several weeks to months of data are needed to retrieve high quality EGFs by cross-correlation (Seats et al., 2012).

Efforts to image the subsurface on an exploration scale using surface waves have been encouraging (Stewart, 2006; Dellinger and Yu, 2009; Bussat and Kugler, 2011; de Ridder and Dellinger, 2011). Last-mentioned study was supported by a more extensive analysis by Mordret et al. (2013a,b). To date, the application of passive seismic interferometry for 4D seismic monitoring has been thwarted by sparse receiver coverage, insufficient recording time, and routine application of a low-cut filter.

Scholte-wave group velocities derived from controlled-source seismic at Valhall have been shown to be sensitive to production-induced changes (Wills et al., 2008). This can be discerned only through a high degree of experiment repeatability. We investigate which subsurface features can be imaged using seismic noise, and report that relatively short recordings are sufficient to create a high repeatability of Scholte-wave velocity maps using tomography. This repeatability should allow for the design and installation of a system that exploits the ambient seismic noise field to perform continuous and nearly real-time reservoir surveillance.

SCHOLTE-WAVE PASSIVE SEISMIC INTERFEROMETRY AT VALHALL FIELD

Passive seismic interferometry is a technique that cross-correlates ambient seismic recordings from two receiver stations to form a signal as if one of the stations were acting as a seismic source. This correlation signal is termed an estimated Green's function (EGF). We apply this technique to the ambient Scholte-wave energy in the vertical component of particle velocity, recorded just under the seabed by the LOFS array at Valhall. Scholte waves are interface waves that propagate along a solid-fluid interface; their particle motions are in the plane containing the the direction of propagation and the normal to the interface of propagation (i.e. the vertical component) (Scholte, 1942a,b).

Seismic Interferometry

The first assumption for Green's function retrieval by cross-correlation is that the energy flux in the ambient wave field is equipartitioned; i.e., energy flux is independent of direction, and all wave modes are excited equally (Lobkis and Weaver, 2001). Modal-energy equipartition alone is not sufficient (Snieder et al., 2010). We require that sources surrounding the station pair act uncorrelated so that the cross-correlation of a long recording time evaluates an ensemble average of the independent contributions of the sources surrounding the station pair (Wapenaar and Fokkema, 2004, 2006).

Let $G^{v,f}_{z,z}(\mathbf{x}_A,\mathbf{x}_B,\omega)$ denote the the frequency-domain elastodynamic Green's function of the vertical component of particle velocity, v, at $\mathbf{x}_A$ due to a vertical external volume force density, f, at $\mathbf{x}_B$. Estimates of the Green's function $G^{v,f}_{z,z}(\mathbf{x}_A,\mathbf{x}_B,\omega)$ and its reciprocal counterpart, $G^{v,f}_{z,z}(\mathbf{x}_A,\mathbf{x}_B,\omega)$, can be retrieved as follows (Wapenaar and Fokkema, 2004):

$$\left\{G^{v,f}_{z,z}(\mathbf{x}_A,\mathbf{x}_B,\omega)+G^{v,f*}_{z,z}(\mathbf{x}_B,\mathbf{x}_A,\omega)\right\}S(\omega)\propto\langle v_z(\mathbf{x}_A,\omega)v^*_z(\mathbf{x}_B,\omega)\rangle\,, \tag{1}$$

where the cross-correlated signals $v_z(\mathbf{x}_A, \omega)$ and $v_z(\mathbf{x}_B, \omega)$ denote vertical components of the particle velocity made at $\mathbf{x}_A$ and $\mathbf{x}_B$ (the master station), respectively. Complex conjugation is denoted by *, and $\langle\rangle$ denotes a spatial ensemble average. The power spectrum of the noise source signals is denoted by $S(\omega)$. Typically, the cross-correlation is not symmetric, reflecting directivity of the ambient seismic field.

In theory, this technique recovers the full elastodynamic Green's function including interface waves, P and S body waves, and single- or multiply-scattered converted waves. In practice, the application of passive seismic interferometry is limited to an ambient-seismic-noise regime that satisfies the aforementioned conditions. These requirements are more often met for surface waves then for other wave modes (Forghani and Snieder, 2010), therefore this study focuses on surface waves.

Cross-Correlation of Valhall LOFS Array Ambient Seismic Scholte-Wave Noise

A recording spanning a little over five days was made from the LOFS in December 2010. Figure 1 shows how the spectrum in the vertical component of particle velocity varies over time. The spectrum was calculated in 2.5-minute overlapping windows and averaged over 153 stations on the northwestern end of the array (where the spectrum is found to have little platform-generated noise). The LOFS array was designed for controlled-source seismic surveying with 15 Hz geophones. But the sensors records ample energy below 2 Hz when the acquisition-standard low-cut filter is turned off.

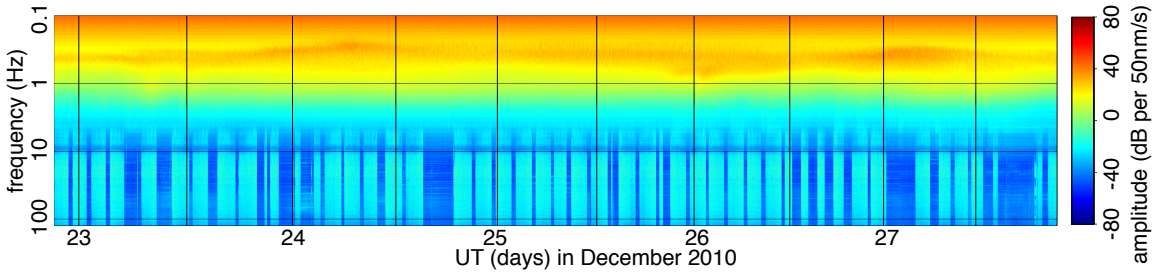


Figure 1: Time-varying spectrum of the five-day passive recording made in 2010. Above 10 Hz, the energy of a seismic acquisition survey at nearby Ekofisk dominates the spectrum. Below 2 Hz, the double-frequency microseism energy band dominates the spectrum. [**CR**] sjoerd2/. spectra4grl2013

The spectrum above 2 Hz is dominated by seismic shooting of an acquisition at nearby Ekofisk Field and energy from other human operations, such as boats, drilling and production noise from the Valhall platforms. This platform noise and other human-generated noise dominates the spectrum down to about 2 Hz. This noise radiates mostly from the platform in the center of the array, and therefore is suitable for seismic interferometry only between station pairs oriented aligned with the platform (Landes et al., 2009). The energy at frequencies lower then 2 Hz is known as microseism energy from the double-frequency microseism peak. This energy is generated by sea swells compressing the seabed (Longuet-Higgins, 1950) and therefore varies with sea depth, wave strength and weather.

The LOFS recordings were bandpassed in 30-minute overlapping windows, with 15-

minute increments, under a sine-squared taper. The bandpass was performed in the frequency domain by multiplication with a raised cosine taper, which has coefficients that equal 0 at 0.175 Hz and 1.75 Hz and equal 1 between 0.20 Hz and 1.50 Hz. We form 240 30-minute windows from the five-day long recording. We create a virtual seismic survey by selecting each station as a master station and collecting all virtual seismic sources. We average five sets of 48 virtual seismic surveys from consecutive 30-minute windows to create five independent daily seismic surveys. We average 240 virtual seismic surveys in a five-day virtual seismic survey.

Figure 2a contains a time slice (cross-correlation time-lag $\tau = 6$ s) of a virtual seismic source in the center of the array. The virtual source radiates approximately equally in all directions, indicating that the energy flux in the ambient seismic field at Valhall, after averaging over these five days, was about equally strong in all directions.

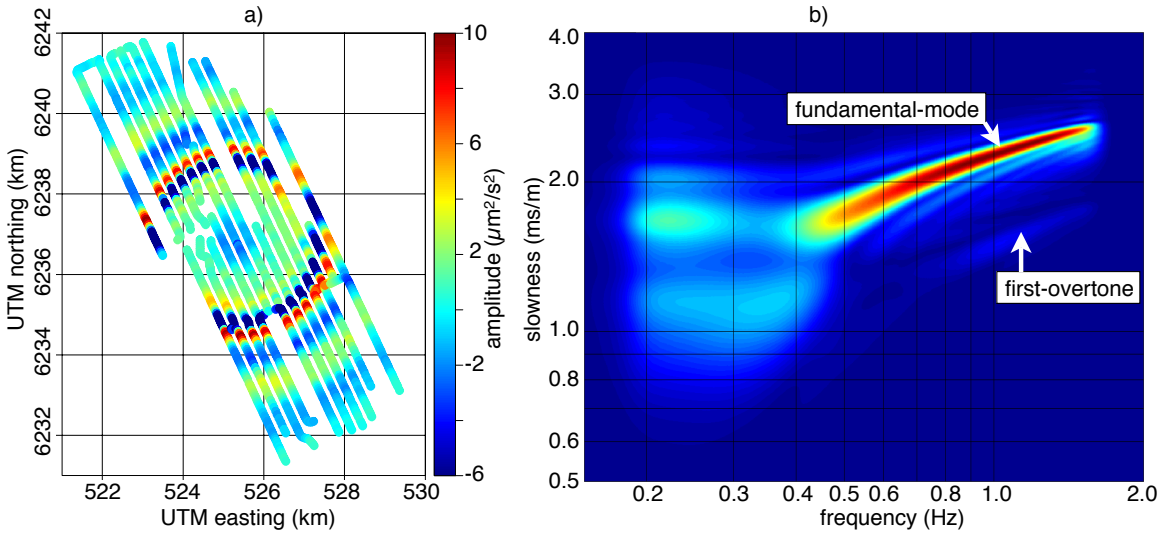


Figure 2: a) Time slice, corresponding to $\tau = 6$ s, of a virtual seismic source located in the center of the array. b) Dispersion image obtained as the amplitude of a Radon $(\omega - p)$ transform, normalized by frequency, of an offset gather of correlations between all station pairs. [**CR**] sjoerd2/. frame6sdisp

The dispersive nature of the Scholte-waves emitted by this virtual seismic source is shown in Figure 2b. This figure is obtained as the amplitude of a Radon (frequency slowness, $\omega - p$) transform of an offset gather of correlations between all station pairs, normalized by frequency. Both the fundamental mode and the first overtone are distinguishable. The seismic interferometry result is dominated by fundamental-mode Scholte waves. This is consistent with Kimman and Trampert (2010), who show that fundamental modes converge faster than higher-order modes. A first overtone that is travelling faster then the fundamental mode, is visible above 0.8 Hz. Recovery of this higher mode is of interest, because higher modes are more sensitive to deeper structure (Aki and Richards, 2002). At lower frequencies (below 0.5 Hz), the fundamental mode jumps to significantly higher velocities, indicative of a sharp velocity increase in the subsurface. The peak between 1.0 ms/m and 2.0 ms/m and between $0.2 - 0.3$ Hz is interpreted as a side lobe (ringing) of the fundamental mode that is created by the $\omega - p$ transformation.

AMBIENT-SEISMIC-NOISE TOMOGRAPHY

We aim to create Scholte-wave group-velocity maps at different frequencies by picking a group traveltime as the peak of the envelope after a narrow-range bandpass. Lower-frequency group-velocities image deeper because lower frequencies have longer wavelengths averaging over deeper structures.

A simple Hann-window bandpass selects a narrow frequency range. After spectral balancing, traveltimes are picked as the maximum of the envelope within a two-second-wide linear move-out window. A quality factor on each pick is defined by a signal-to-noise ratio (SNR) based on the maximum of the envelope within the window divided by the average of the envelope outside the window. Source-receiver pairs with a causal versus anti-causal traveltime difference corresponding to 0.1 ms/m for $0.55 - 1.15$ Hz, and 0.2 ms/m for $1.15 - 1.75$ Hz, are discarded. The accepted travel-time picks are based on a symmetrized EGF. Only source-receiver pairs with a SNR > 3 for $0.55 - 1.15$ Hz or SNR > 1.5 for $1.15 - 1.75$ Hz, and offsets between 2000 m and 6000 m were selected for inversion (about one million traveltime picks for $0.55 - 1.15$ Hz and about six hundred thousand for $1.15 - 1.75$ Hz). For each inversion, we find N travel-time picks, $\mathbf{t} = (t_1, t_2, ..., t_N)^t$, that are predicted from a slowness model space, $\mathbf{m}$, by a straight-ray tomography operator, $\mathbf{F}$. This operator is simply a slowness integration kernel, where each row contains the distances through each model cell of a line directly from source to receiver, to predict the traveltime of one specific source-receiver couple. The problem is posed as a perturbation, $\Delta\mathbf{m}$, of an average slowness, m_0:

$$\mathbf{m} = \mathrm{m}_0 + \Delta\mathbf{m} \tag{2}$$

$$\mathrm{m}_0 = \frac{1}{N}\sum_{i=1}^{N}\frac{t_i}{\Delta x_i} \tag{3}$$

$$\Delta\mathbf{t} = \mathbf{t} - \mathrm{m}_0\,\Delta\mathbf{x}, \tag{4}$$

where $\Delta\mathbf{x} = (\Delta x_1, \Delta x_2, ..., \Delta x_N)^t$ contains the source-receiver distances for each traveltime, and $\Delta\mathbf{t}$ are the travel-time residuals after the contribution of the average slowness is taken into account. A conjugate-direction algorithm is used to find the minimum of the following L^2 norm:

$$\left\|\mathbf{F}\Delta\mathbf{m} - \Delta\mathbf{t}\right\|_2^2 + \epsilon\left\|\nabla^2\Delta\mathbf{m}\right\|_2^2, \tag{5}$$

where we use the ∇^2 operator as regularization to force a smooth model. The model space is formed by 90 (easterly) by 110 grid cells (northerly) each 100 meters wide. The solver is run till the norm of the data-space and model-space residuals each change less then 0.01% of their respective maximum values. After the first iteration, the 2.5% worst-fit data points are eliminated. The problem is redefined and ran till convergence again for a range of ϵ. The grid cell size is below the resolution of the wavelength thus the regularization plays an important role in finding reasonable solutions. The optimum ϵ is picked, using an L-curve analysis, choosing a value above which the regularization has significantly less influence (Aster et al., 2005).

Group velocity maps for two frequency ranges are shown in Figures 3a and 3b, for $0.55 - 1.15$ Hz and for $1.15 - 1.75$ Hz, respectively. The figures compare five velocity maps from independent, day-long recordings with a velocity map from the complete five-day

recording. Both the average RMS difference between the 10 combinations of velocity maps obtained from day-long recordings in Figures 3a and the average RMS difference between the 5 combinations of velocity maps obtained from day-long recordings and the velocity map obtained from the complete five-day recording is ~ 1.1 m/s. The same analysis for 3b results in average RMS differences of ~ 2.0 m/s. This is an indication that the correlation convergence rate at higher frequencies is lower than at lower frequencies, consistent with preliminary results shown by de Ridder (2012).

COMPARISON WITH *P*-WAVE VELOCITY MAPS FROM FULL WAVEFORM INVERSION

The P-wave velocity structure at Valhall is independently well known from regular controlled-source seismic surveying (Sirgue et al., 2010). Two xy-slices through a cube of P-wave velocities are shown in Figure 4, panels a and b, for ranges $150 - 195$ m and $105 - 150$ m below the sea floor, respectively. Because these controlled sources cover a wider area than the receiver array, these images include features extending well beyond the area of the receiver array. The maps obtained from ambient-seismic-noise tomography are necessarily confined within the area of the recording array. Based on Figure 2, the wavelengths of the Scholte-waves underlying the the group velocity maps in 3a and 3b are between $190 - 500$ m and $110 - 190$ m respectively. Although the Scholte-wave group velocity maps have an integrated sensitivity over depth, they exhibit peak depth-sensitivities of half the wavelength: $85 - 250$ m and $55 - 85$ m respectively (Aki and Richards, 2002). Figures 3a and 3b can be compared to the depth slices in Figures 4a and 4b respectively. No unique mapping exists between Scholte-wave velocities and P-wave velocities, but each reflects the underlying geology that shapes the velocity anomalies.

We identify the following similarities between Figures 3a and 4a: a wide channel that crosses within the array boundary at the southeastern corner of the array and again touches the edge of the array near UTM $(526.5, 6239)$ km; a low-velocity zone reflects a combination of channel fill and sub-platform low-velocities stretches across the array from UTM $(525, 6235.5)$ km to UTM $(526.5, 6238)$ km. We identify the following similarities between Figures 3b and Figure 4b: a thinner meandering channel that strikes approximately north-south, approximately from UTM $(525, 6235)$ km to UTM $(526, 6238.5)$ km, and a less well defined channel that crosses the northern end of the array from UTM $(523, 6238.5)$ km to UTM $(525, 6240)$ km. The Scholte-wave velocities in Figure 3b are generally higher in the vicinity of the platform, UTM $(523, 6238.5)$ km, which lies above the central region of the field, than in the southeastern portions of the array. Although not observed in the P-wave velocities of Figure 4, this was identified by Hatchell et al. (2009) in Scholte-wave velocity images from controlled-source seismic and is representative of contractional and extensional strains caused by production-related sea-floor subsidence at Valhall.

The images in Figure 3 are generally smoother than the P-wave velocity images in Figure 4. This is attributed to a combination of the regularization of the velocity map, the wavelength of the interface waves, and the limited depth resolution provided by the dispersive character of the interface waves. These limitations are either less strict or do not apply in the P-wave (body wave) inversions.

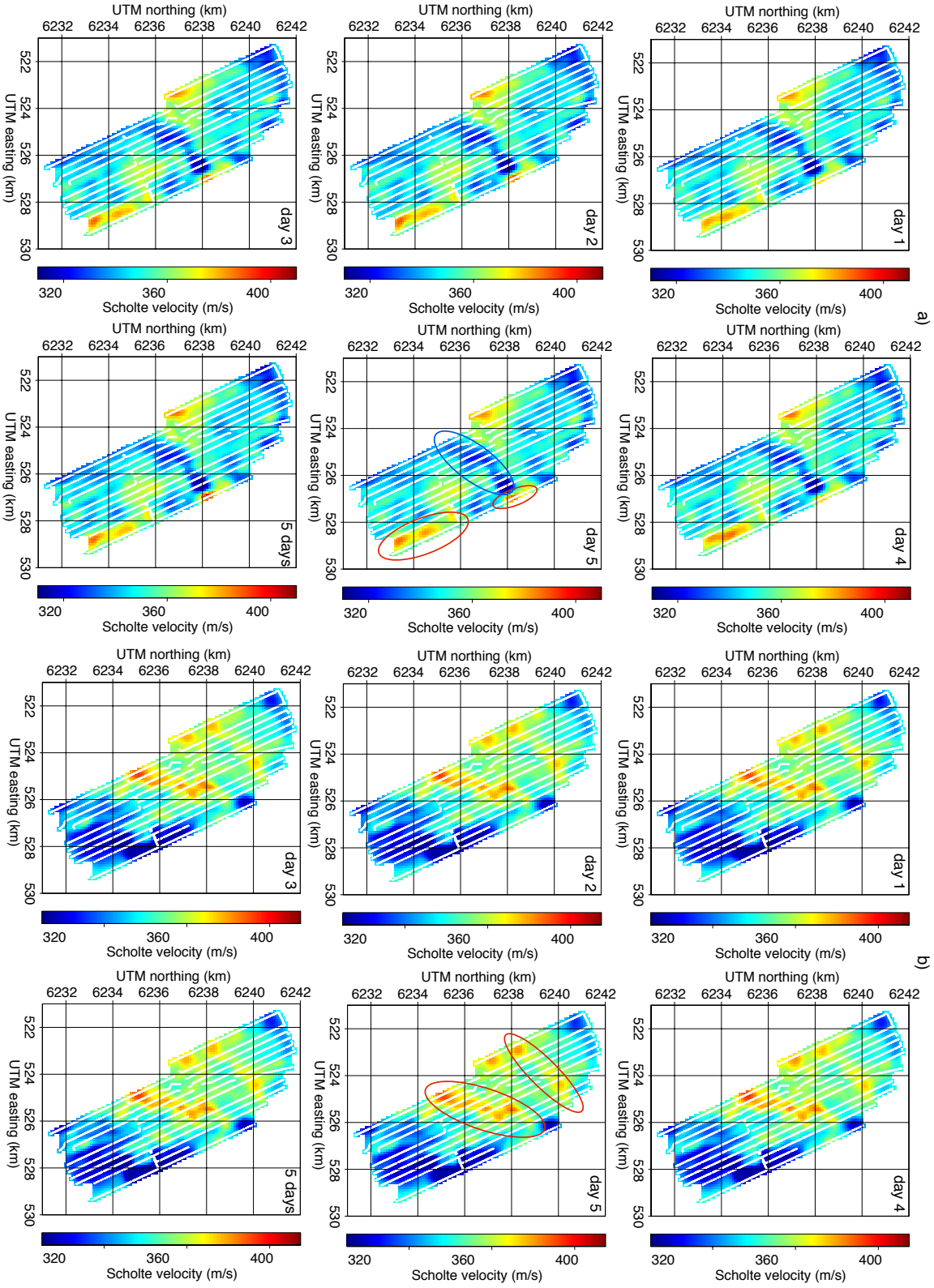


Figure 3: Scholte-wave group-velocity tomography for $0.55-1.15$ Hz in a) and $1.15-1.75$ Hz in b). Both sets contain maps obtained using five independent, day-long passive recordings and a map obtained using the five-day passive recording. In the images for day 5 the high-velocity channels are annotated by red circles, and the sub-platform and channel fill low-velocity zone is annotated by a blue circle. [**CR**] sjoerd2/. asnt-monitor5aplusb

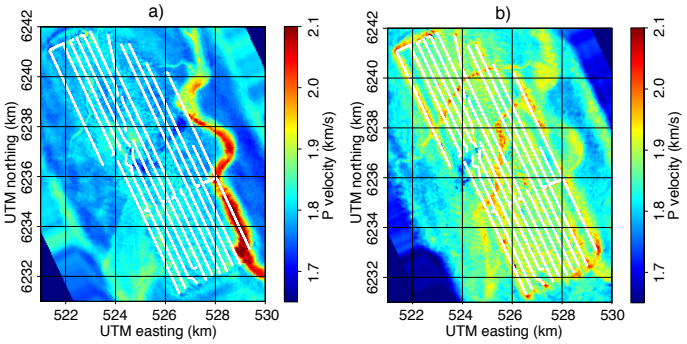


Figure 4: Image of P-wave velocities obtained using waveform inversion of controlled-source P-wave data (Sirgue et al., 2010). Velocity slices $150 - 195$ m below the sea floor in a) and $105 - 150$ m below the sea floor in b). [**NR**] sjoerd2/. fwi

CONCLUSIONS

We have shown remarkable repeatability of Scholte-wave group-velocity maps of the overburden at the Valhall Field. We correlate for mere 24 hours to retrieve high-quality EGFs suitable for ambient-seismic-noise tomography. Images of Scholte-wave group velocities obtained using ASNT correspond well with maps of P-wave velocities obtained from controlled-source seismic acquisition and full-waveform inversion. Features as deep as 200 m below the sea floor can be identified from group velocity maps in a frequency range of $0.55 - 1.15$ Hz. Further studies of recordings separated by months are necessary to distinguish long-term changes from daily fluctuations. Careful benchmarking of quality control in the ASNT process could potentially stabilize the daily images such that short-term changes can be reliably identified. Imaging and inversion of EGFs is not limited to ASNT, and more sophisticated inversion methodologies can be developed for virtual seismic sources obtained using dense local receiver arrays. A wide range of practices, including CO_2 sequestration projects, might benefit from successful monitoring by seismic interferometry.

ACKNOWLEDGMENTS

Olav Barkved, Greg Beroza, Jason Chang, Bob Clapp and Joe Dellinger and Dave Nichols for helpful discussions and suggestions. Reviewers Nikolai Shapiro and Matt Haney for comments and suggestions that benefited both this paper and my continued research. BP and the partners of the Valhall Field (BP Norge and Hess Norge) for permission to publish this paper. The sponsors of the Stanford Exploration Project for financial support.

REFERENCES

Aki, K., 1957, Space and time spectra of stationary stochastic waves, with special reference to microtremors: Bulletin of the Earthquake Research Institute, **35**, 415–456.

Aki, K. and P. G. Richards, 2002, Quantitative Seismology - second edition: University Science Books.

Aster, R., B. Borchers, and C. Thurber, 2005, Parameter Estimation and Inverse Problems: Academic Press.

Bussat, S. and S. Kugler, 2011, Offshore ambient-noise surface-wave tomography above 0.1 hz and its applications: The Leading Edge, **30**, 514–524.

Campillo, M. and A. Paul, 2003, Long-range correlations in the diffuse seismic coda: Science, **299**, 547–549.

Claerbout, J. F., 1968, Synthesis of a layered medium from its acoustic transmission response: Geophysics, **33**, 264–269.

de Ridder, S., 2012, Ambient seismic noise correlations for reservoir monitoring: SEG Technical Program Expanded Abstracts, **31**, 1–5.

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

Dellinger, J. A. and J. Yu, 2009, Low-frequency virtual point-source interferometry using conventional sensors: 71st Meeting, European Associated of Geoscientists and Engineers, Expanded Abstracts, X047.

Duvall, T., S. Jefferies, J. Harvey, and M. Pomerantz, 1993, Time-distance helioseismology: Nature, **362**, 430–432.

Forghani, F. and R. Snieder, 2010, Underestimation of body waves and feasibility of surface-wave reconstruction by seismic interferometry: The Leading Edge, **29**, 790–794.

Gerstoft, P., K. P. Sabra, P. Roux, W. A. Kuperman, and M. C. Fehler, 2006, Greens functions extraction and surface-wave tomography from microseisms in southern california: Geophysics, **71**, SI23SI31.

Hatchell, P. J., P. B. Wills, and C. Didraga, 2009, production induced effects on near-surface wave velocities at valhall: 71st Meeting, European Associated of Geoscientists and Engineers, Expanded Abstracts, T016.

Kimman, W. P. and J. Trampert, 2010, Approximations in seismic interferometry and their effects on surface waves: Geophysical Journal International, **182**, 461–476.

Kommedal, J. H., O. I. Barkved, and D. J. Howe, 2004, Initial experience operating a permanent 4C seabed array for reservoir monitoring at Valhall: SEG Technical Program Expanded Abstracts, **23**, 2239–2242.

Landes, M., N. M. Shapiro, S. Singh, and R. Johnston, 2009, Studying shallow seafloor structure based on correlations of continuous seismic records: SEG Technical Program Expanded Abstracts, **28**, 1693–1697.

Lin, F.-C., M. 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: J. Acoust. Soc. Am., **110**, 3011–3017.

Longuet-Higgins, M. S., 1950, A theory of the origin of microseisms: Phil. Trans. R. Soc. Lond. A, **243**, 135.

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

Mordret, A., N. M. Shapiro, S. Singh, . Roux, J. P. Montagner, and O. I. Barkved, 2013b, Azimuthal anisotropy at Valhall: the Helmholtz equation approach: Geophys. Res. Let.

Scholte, J. G., 1942a, On the stoneley wave equation: Proceedings of the KNAW, **45**, 20–25.

———, 1942b, On the stoneley wave equation: Proceedings of the KNAW, **45**, 159164.

Seats, K. J., J. F. Lawrence, and G. A. Prieto, 2012, Improved ambient noise correlation functions using welch's method: Geophysical Journal International, **188**, 513–523.

Shapiro, N. M. and M. Campillo, 2004, Emergence of broadband rayleigh waves from correlations of the ambient seismic noise: Geophys. Res. Let., **31**, L07614–1–L07614–4.

Shapiro, N. M., M. Campillo, L. Stehly, and M. R. Ritzwoller, 2005, High-resolution surface-

wave tomography from ambient seismic noise: Science, **307**, 1615–1618.

Sirgue, L., O. I. Barkved, J. Dellinger, J. E. U. Albertin, and J. H. Kommedal, 2010, Full waveform inversion: the next leap forward in imaging at Valhall: First Break, **28**, 65–70.

Snieder, R., Y. F. E. Slob, and K. Wapenaar, 2010, Equipartitioning is not sufficient for green's function extraction: Earthquake Science, **23**, 403–415.

Stewart, P., 2006, Interferometric imaging of ocean bottom noise: SEG Technical Program Expanded Abstracts, **25**, 1555–1559.

Wapenaar, K. and J. Fokkema, 2004, Reciprocity theorems for diffusion, flow, and waves.: J. Appl. Mech., **71**, 145–150.

———, 2006, Green's function representations for seismic interferometry: Geophysics, **71**, SI33–SI46.

Weaver, R. L. and O. I. Lobkis, 2001, Ultrasonics without a source: Thermal fluctuation correlations at mhz frequencies.: Phys. Rev. Let., **87**, 134301–1 – 134301–4.

———, 2002, On the emergence of the Green's function in the correlations of a diffuse field: pulse-echo using thermal phonons: Ultrasonics, **40**, 435–439.

Wills, P. B., P. J. Hatchell, and S. J. Bourne, 2008, time-lapse measurements of shallow horizontal wave velocity over a compacting field: 70st Meeting, European Associated of Geoscientists and Engineers, Expanded Abstracts, G039.

Yao, H., R. D. van der Hilst, and M. V. de Hoop, 2006, Surface-wave array tomography in se tibet from ambient seismic noise and two-station analysis i. phase velocity maps: Geophys. J. Int., **166**, 732–744.

Noise characterization and ambient noise cross-correlations at Long Beach

Jason P. Chang, Sjoerd de Ridder, and Biondo Biondi

ABSTRACT

The dense seismic array in Long Beach, California is located in an urban environment along the Pacific Ocean. There are a variety of noise sources influencing the ambient seismic noise field, both natural and anthropogenic in origin. To understand the temporal and spatial influences of these sources, we calculate power spectral densities (PSD) and apply beamforming to ambient seismic noise data. From spatial distribution maps of noise PSD, we find that energy from the Pacific Ocean dominates the noise field at frequencies below 2 Hz, while energy from local roads and Interstate 405 dominate the noise field at frequencies above 2 Hz. From spectrograms, we observe diurnal fluctuations in energy that are in accord with expected patterns in human activity. From beamforming, we find that directed, low-frequency energy from the Pacific Ocean is prevalent throughout the array, while the directivity of high-frequency energy varies throughout the array. Near Interstate 405, noise energy is clearly directed outwards from the freeway, but at a distance from the freeway, noise energy arrives more evenly across azimuths. Based on these observations, we expect the noise source distribution to be generally more homogeneous at higher frequencies than at lower frequencies. Ambient noise cross-correlation results at frequencies spanning 0.5-2 Hz and 2-4 Hz reinforce the validity of this expectation. A more exciting observation is the emergence of a P-wave at the higher-frequency range in our virtual super-source gather. This is a promising first step toward potentially retrieving other body-wave arrivals.

INTRODUCTION

The Long Beach oil field is a producing oil field located beneath the city of Long Beach. Due to its urban location, traditional techniques for collecting data for seismic imaging and velocity analysis are disruptive and difficult to perform. One alternative is to use passive seismic interferometry, or ambient noise cross-correlation, for this type of subsurface analysis. By cross-correlating ambient seismic noise recorded at two simultaneously recording receivers, the Green's function between the two receiver locations can be estimated (Bensen et al., 2007; Wapenaar et al., 2010). Because surface waves are the strongest events in ambient noise, the surface wave part of the Green's function has been widely used for tomographic imaging at both the regional (Shapiro et al., 2005; Sabra et al., 2005) and continental scales (Yang et al., 2008; Bensen et al., 2008). The body wave part of the Green's function is beginning to gain attention, with Roux et al. (2005) retrieving direct P-waves at the regional scale, Draganov et al. (2007, 2009) retrieving P-wave reflections at the exploration scale, and Zhan et al. (2010) and Poli et al. (2012) retrieving P- and S-wave reflections at the continental scale.

In January 2012, Nodal Seismic, LLC, deployed Phase II of its dense seismic array in Long Beach that is well-suited for testing the effectiveness of using ambient seismic noise for resolving subsurface structure at the exploration scale (see Figure 1). The array consists of 2400 vertical-component geophones with an average spacing of 100 m. The geophones are continuously recording (24 hrs/day) for over three months. Using data from Phase I, Lin et al. (2012) have shown that Green's functions can be successfully estimated at the neighboring Long Beach array using seismic interferometry.

Under ideal conditions, noise sources would be evenly distributed throughout Long Beach, thereby allowing seismic interferometry to better estimate Green's functions (Snieder, 2004). Due to the presence of obvious sources of seismic noise, this is unlikely to be the case at this site. Our goal is thus to characterize the ambient seismic noise field at Long Beach. We want to identify the various noise sources and their influence on the ambient seismic field in both time and space. To achieve this, we first compute power spectral densities (PSD) of recorded ambient seismic noise. We plot spatial distribution maps to locate potential sources of noise, and then we generate spectrograms for those energetic areas to observe how they vary over time. Next, we apply beamforming to the ambient seismic noise data to determine the slowness and azimuthal distribution of noise. Finally, we relate our beamforming results to estimated Green's functions from seismic interferometry at frequencies between 0.5-2 Hz and 2-4 Hz.

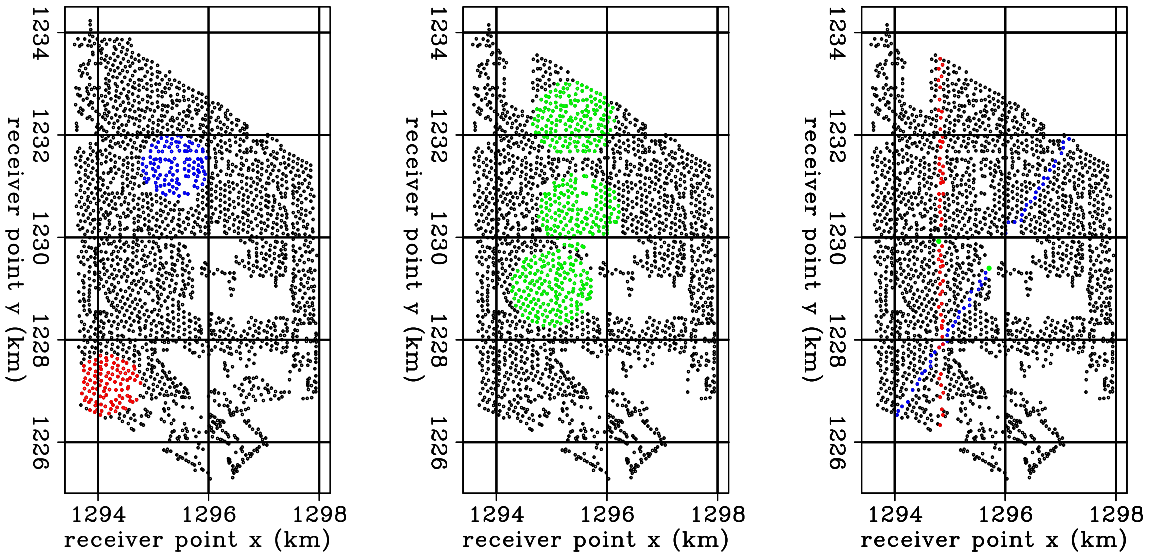


Figure 1: Maps of the dense Long Beach array overlain with station used for (a) spectrograms, (b) beamforming analysis, and (c) common virtual source gathers. In (c), solid green circles mark the locations of the virtual sources. Coordinates are NAD27, CA State Plane, Zone 7, kilometers. [**ER**] jason1/. sp-map,beam-map,line-map

NOISE CHARACTERIZATION

PSD Spatial Distribution Maps

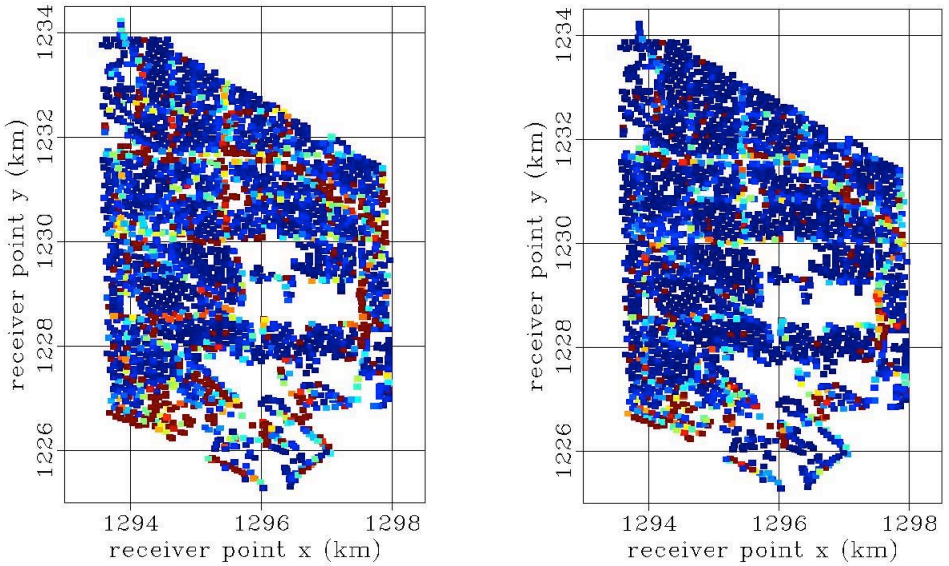
Generating and examining noise PSD, or power spectrums, is the standard method for quantifying seismic background noise (McNamara and Boaz, 2005). PSD reveal how power in a signal is distributed over a range of frequencies. Before calculating the PSD at each receiver, we need to prepare the seismic recordings. We first break our time series into 5-minute patches with 50% overlap. Next we subtract the mean from each time segment to remove any systematic offset or zero-frequency component. We then apply a smooth Hanning taper to each time segment to reduce the artifacts caused by Gibbs phenomenon when performing a Fourier transform on a truncated time series. To compute the PSD for a time segment, we take the one-dimensional, discrete Fourier transform and then compute the square of the complex amplitude.

To identify potential sources of noise at Long Beach, we generate spatial distribution maps of the noise PSD by plotting PSD at their receiver locations. We average the PSD derived from twelve consecutive overlapping 5-minute time segments to obtain a single PSD for a 32.5-minute time window. By averaging over multiple overlapping time segments, we decrease the variance in our PSD estimates and minimize the influence of spurious events. We show results from two different time windows to get a sense of how the spatial noise distribution differs during peak human activity (5pm) and during calmer periods of human activity (midnight).

Figures 2, 3, and 4 display the spatial distribution maps of noise PSD for frequencies of 15 Hz, 4 Hz, and 1 Hz, respectively. Figures 2-4(a) are averaged PSD results from a 32.5-minute time window at 5pm, and Figures 2-4(b) are averaged PSD results from a 32.5-minute time window at midnight. Warmer colors (red) indicate higher energy at the given frequency at a particular station. For a given pair of maps at the same frequency, the amplitude scale is the same so that a side by side comparison of relative energy can be made. The amplitude scale of figures from frequency to frequency is not the same, however. We received the data with a 3 Hz low-cut filter, so PSD amplitudes at 1 Hz are much lower than at 15 Hz.

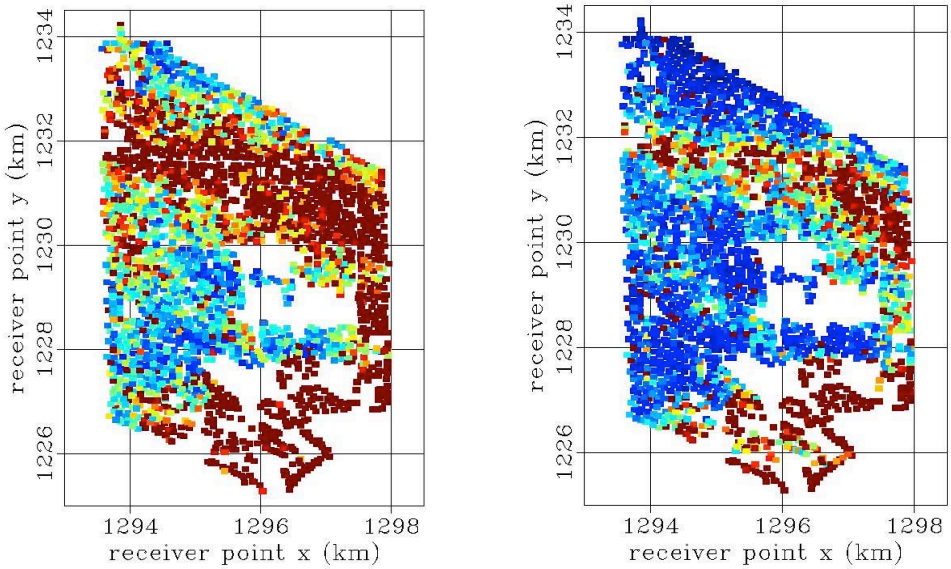
The PSD spatial distribution maps reveal regions where high amounts of seismic energy are being recorded and potentially produced. From Figure 2, we see high amounts of energy at 15 Hz localized at receivers along Interstate 405 (which runs roughly east-west through the northern part of the array) and local roads. This suggests that the seismic noise field is dominated by traffic noise at higher frequencies. We observe a similar trend in Figure 3, where the dominant energy source at 4 Hz is clearly Interstate 405. The difference in spatial extent of freeway energy at 4 Hz and 15 Hz can be attributed to attenuation. The freeway energy at 15 Hz is localized because higher frequencies attenuate more rapidly with distance, whereas the freeway energy at 4 Hz is more spread out because lower frequencies attenuate more gradually with distance.

At 4 Hz frequency, we also see the appearance of high-energy noise in the southeast region of the array. The border of this region is fairly consistent over the two time windows, which suggests that we could be looking at a geological boundary rather than a source of noise. According to the California Department of Conservation (Greenwoood, 1998), parts of this region, and specifically margins of Alamitos Bay, are built on artificial fill that



(a) (b)

Figure 2: Spatial distribution maps of noise PSD at 15 Hz. (a) at 5pm. (b) at midnight. Warmer color (red) points indicate stations that are recording higher energy at this frequency. The Pacific Ocean is immediately south of the array, and Interstate 405 runs east-west through the north of the array.. [**CR**] jason1/. sh-0-57-fx-snap15,sh-6-27-fx-snap15



(a) (b)

Figure 3: Spatial distribution maps of noise PSD at 4 Hz. (a) at 5pm. (b) at midnight. Warmer color (red) points represent stations that are recording higher energy at this frequency. . [**CR**] jason1/. sh-0-57-fx-snap4,sh-6-27-fx-snap4

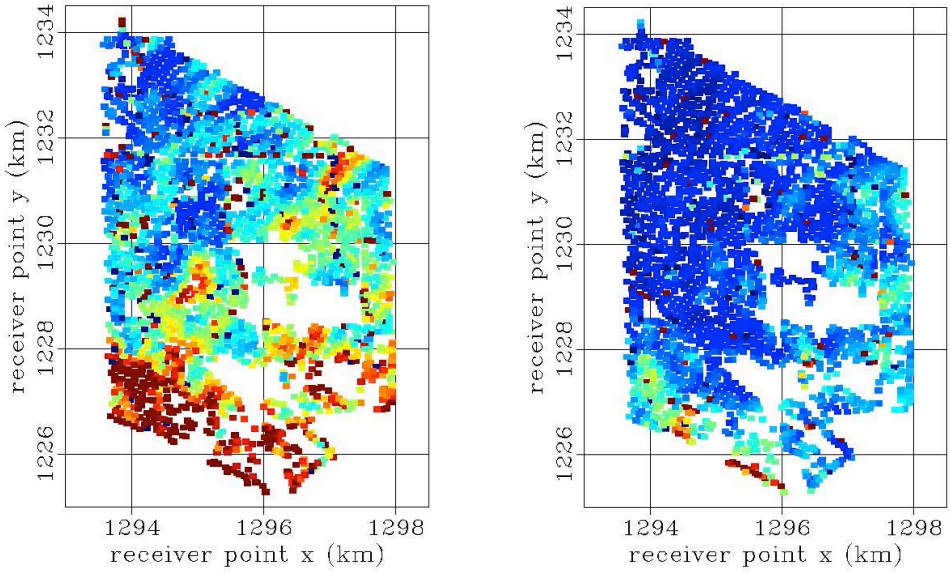


Figure 4: Spatial distribution maps of noise PSD at 1 Hz. (a) at 5pm. (b) at midnight. Warmer color (red) points represent stations that are recording higher energy at this frequency. . [**CR**] jason1/. sh-0-57-fx-snap1,sh-6-27-fx-snap1

consists of soft and silty sand. Soft material can amplify seismic amplitudes (Delgado et al., 2000), so we could be seeing artificial fill or some other type of geology amplifying seismic signal at around 4 Hz in this region. A tomography study would address this conjecture, and it is the next objective of our research. At 1 Hz (Figure 4), we see energy primarily in the southern part of the array along the coastline, which suggests that energy from the Pacific Ocean dominates at these low frequencies. In general, there is more seismic energy at 5pm than at midnight for all frequencies.

Spectrograms

To better resolve how the noise field at Long Beach changes over time, we generate spectrograms (plots of frequency versus time) for receivers near the coastline (red circle in Figure 1(a)) and receivers overlapping Interstate 405 (blue circle in Figure 1(a)). We first discard recordings in those regions that have maximum PSD amplitudes above a certain threshold, since they will bias the spatial average. We then average the PSD from the remaining receivers in those regions for each 5-minute time segment and plot those PSD side by side to obtain a map showing how frequency content changes with time.

Figure 5 displays spectrograms spanning over two days in February for frequencies up to 4 Hz. Figure 5(a) represents the freeway region and Figure 5(b) represents the ocean region. The color scale represents pseudo-amplitudes after taking the logarithm of the PSD amplitudes. A better comparison of logarithmic PSD amplitudes can be found in Figure 6(a) and Figure 6(b), where we compute the average PSD over one-hour time windows at noon

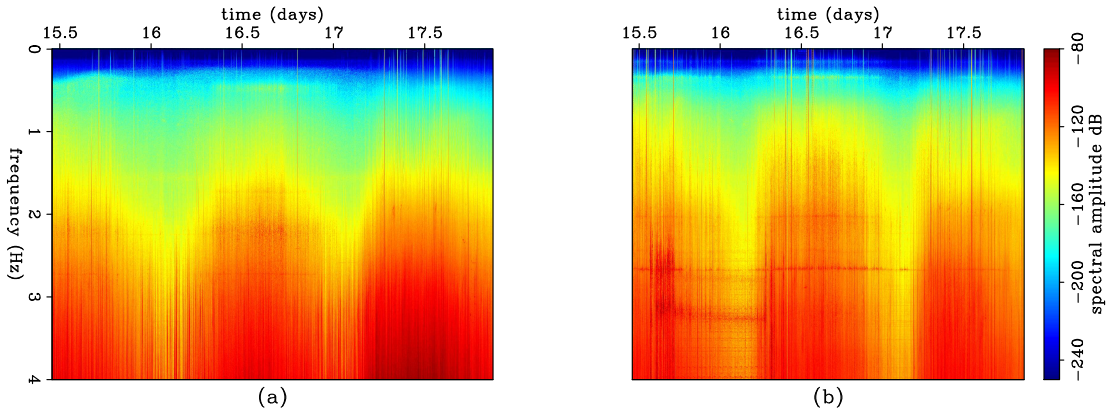


Figure 5: Spectrogram spanning over two days in February for frequencies up to 4 Hz. (a) freeway. (b) ocean. Color scale corresponds to the logarithm of the PSD amplitudes, with warmer colors indicating higher energy and cooler colors indicating lower energy. [**CR**] jason1/. sh-spectragram-down-side

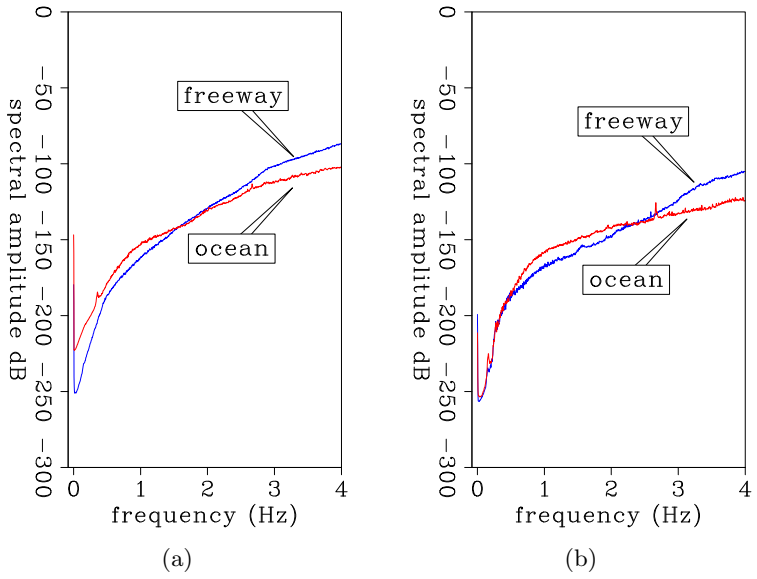


Figure 6: Average PSD over one-hour time windows at (a) noon, and at (b) midnight for the freeway (blue) and ocean (red) regions. [**ER**] jason1/. sh-sp-down-avg-noonA,sh-sp-down-avg-midnightA

and midnight of the 17th, respectively, for the freeway (blue curve) and ocean (red curve) regions.

From the noise PSD spatial distribution maps, it is clear that the two primary sources of energy at Long Beach are the Pacific Ocean and Interstate 405. We also observe that energy in the ambient noise field is generally weaker at midnight than at 5pm, which suggests that anthropogenic noise sources are highly influential at Long Beach. Spectrograms in Figure 5 and the averaged PSD in Figure 6 confirm both these observations. The averaged PSD show that energy is always greatest near the freeway at frequencies above 2 Hz, hence identifying Interstate 405 as a major source of seismic noise, and greatest at the ocean receivers at frequencies below 2 Hz, hence identifying the Pacific Ocean as another major source of seismic noise. Diurnal fluctuations are clearly seen in the spectrograms, with peaks corresponding to daytime periods when human activity is most active and valleys corresponding to early morning periods when human activity is more subdued. Along with the differences in the averaged PSD at noon and midnight, it is apparent that human activity is primarily responsible for the diurnal variations seen in the noise field.

Beamforming

Having identified two major sources of noise at Long Beach (the Pacific Ocean and traffic noise), we now want to determine the extent of their influence over the ambient noise field. We do this by applying beamforming to the ambient seismic noise data, which provides the slowness and azimuth at which seismic energy arrives (Rost and Thomas, 2002). We first select two hours of ambient seismic noise recorded at groups of stations throughout the array. A long time window is chosen because our ambient noise cross-correlations are averaged over long periods of time. We whiten the spectrum of each recording by dividing the complex spectrum by a smoothed version of its amplitude spectrum. This step maintains the crucial phase information while removing the bias in the recordings toward higher frequencies (due to the 3 Hz low-cut filter) and mitigating the influence of anomalous recordings on the beamforming output. We then apply a series of overlapping 0.5 Hz bandwidth filters centered at intervals 0.25 Hz apart, beginning at 0.5 Hz and ending at 3.75 Hz. By applying beamforming to each narrow frequency range, we can better resolve how slowness and azimuthal distribution of noise vary with frequency. Next we transform the data to the τ-p domain. This can be thought of as a series of slant stacks along different azimuths. If there is any coherent energy traveling across an array at a certain slowness and azimuth, it will stack constructively and appear as a strong event in the τ-p domain. We then take the absolute value of the transformed data and average over τ. This produces a map depicting the slowness and azimuthal distribution of noise for those set of stations, representative of the time window.

Figure 7 and Figure 8 display beamforming results for frequency bands 0.5-1.0 Hz and 3.25-3.75 Hz, respectively. For both figures, columns from left to right represent results from January 18 at 2 am, 10 am, and 6 pm, while rows from top to bottom represent results from clusters of stations located just north of Interstate 405, just south of Interstate 405, and near the center of the array (for station clusters see Figure 1(b)). The azimuth of a high-beampower anomaly represents the azimuth from which noise is coming (eg. an anomaly on the right represents noise coming from the east) while the radial distance of the anomaly represents the slowness at which that noise is traveling across the array. The

amplitude of the beampower is scaled the same for panels in a given row. Warmer colors represent higher beampower.

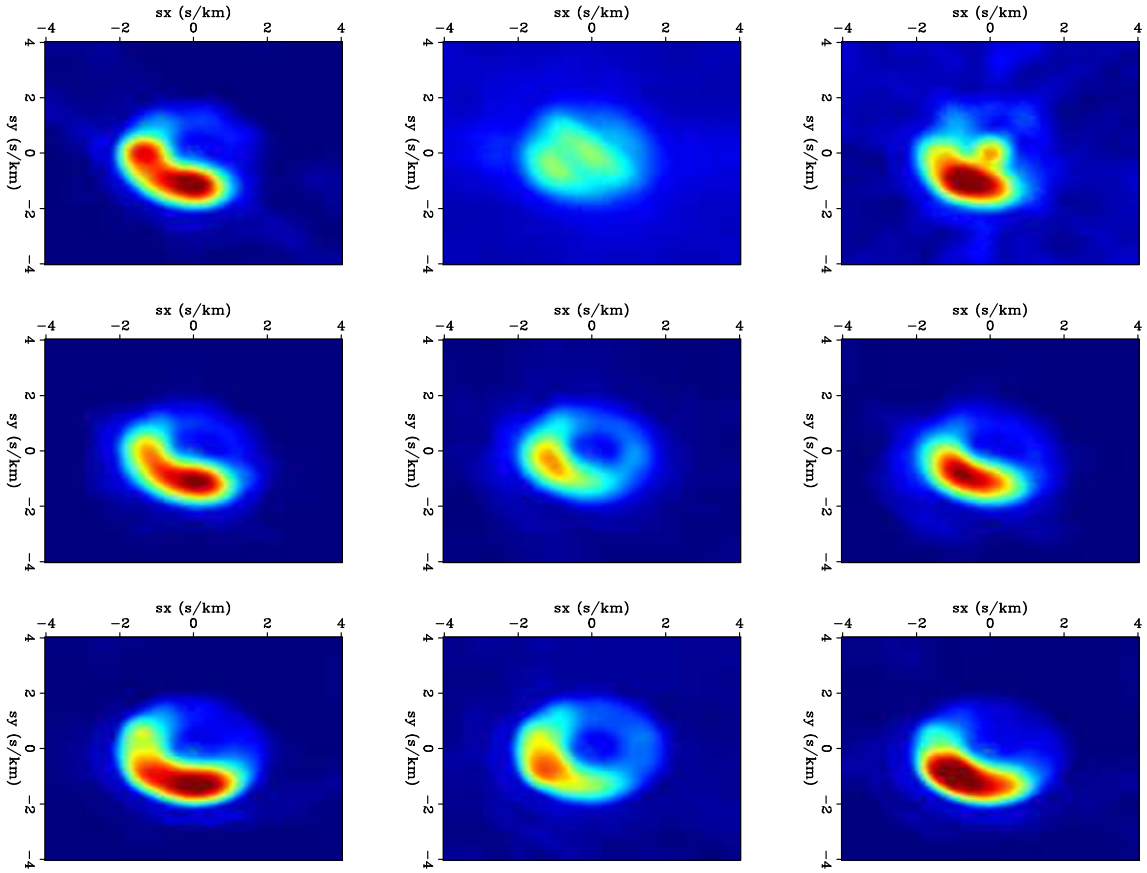


Figure 7: Beamforming results for frequency band 0.5-1.0 Hz. Columns from left to right represent results from 2 am, 10 am, and 6 pm. Rows from top to bottom represent results from clusters of stations located just north of Interstate 405, just south of Interstate 405, and near the center of the array. [**CR**]Warmer (red) colors represent higher beampower. jason1/. low-morning,low-afternoon,low-night

At the frequency band 0.5-1 Hz, it is clear from Figure 7 that noise is arriving from the southwest at all locations and times. This is yet another indication .0that the Pacific Ocean is the dominant source of energy at low frequencies, and that its influence extends throughout the entire array. For this particular frequency band, noise from the ocean arrives generally at a slowness of .001 s/m, or a velocity of 1 km/s. Based on the frequency and velocity, this is likely Rayleigh wave energy.

At the frequency band 3.25-3.75 Hz, there is more variation in the ambient seismic noise field from region to region, and in some cases variation from time to time. Looking at different regions of the array (different rows), we see clear differences in the azimuths at which energy is distributed. Stations just north of the freeway (top row) reveal that noise is primarily approaching from the south, while stations just south of the freeway (middle row) reveal that noise is primarily approaching from the north. These observations indicate that

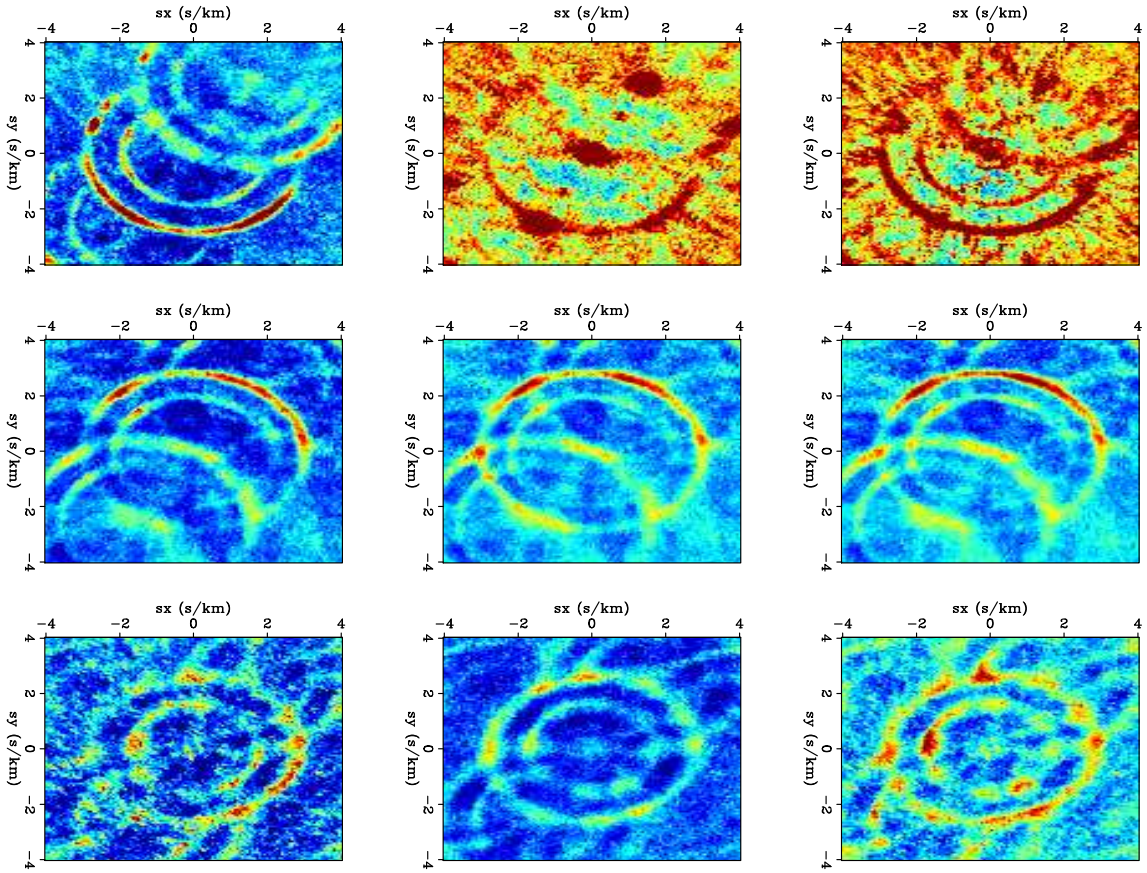


Figure 8: Beamforming results for frequency band 3.25-3.75 Hz. Columns from left to right represent results from 2 am, 10 am, and 6 pm. Rows from top to bottom represent results from clusters of stations located just north of Interstate 405, just south of Interstate 405, and near the center of the array. [**CR**]Warmer (red) colors represent higher beampower. jason1/. high-morning,high-afternoon,high-night

Interstate 405 is the dominant source of noise energy for regions near the freeway. Strong arrivals at zero slowness during the day indicate a signal that arrives at those stations at the same time and could be caused by noise from airplanes flying over the stations after taking off or before landing at the nearby Long Beach airport. Airplane noise might also explain the generally high beampower in these panels. Noise energy is better equipartitioned near the center of the array (bottom row), which suggests that noise from nearby local roads could be the dominant source of energy away from the freeway. In all these regions, there appears to be noise consistently arriving at approximately 333 m/s and 500 m/s. Because strong energy exists at similar azimuths for both arrivals, we are likely observing the slower fundamental Rayleigh wave mode and a faster higher-order Rayleigh wave mode rather than arrivals from two different sources. At these frequencies, nearly all panels exhibit crossing anomalies; these are spatial aliasing artifacts. The two surface wave modes are affected by it, but it appears that any potential higher-velocity body wave will not have crossing events and hence will not be aliased. Looking at different times (different columns), we see that the relative effect of the freeway on nearby stations is consistent over time. At stations away from the freeway, noise is slightly better equipartitioned at 10 am and 6 pm than at 2 am, suggesting that the noise source distribution is more even at peak periods of human activity.

From looking at PSD spatial distribution maps, we determined that the major sources of noise energy at Long Beach were the Pacific Ocean at frequencies below 2 Hz, and traffic noise, particularly Interstate 405, at frequencies above 2 Hz. Spectrograms revealed that the ambient seismic noise field goes through diurnal fluctuations in energy that can be attributed to diurnal patterns in human activity. Beamforming analysis showed that noise is consistently arriving from the southwest at low frequencies, confirming that the Pacific Ocean is a dominant noise source. This analysis also showed that high-frequency noise energy near the freeway is dominated by Interstate 405. However, further away from the freeway, the noise arrived at a wider range of azimuths and could potentially be attributed to energy generated from traffic on local roads.

AMBIENT NOISE CROSS-CORRELATIONS

Method

Having characterized the noise field at Long Beach, we now perform ambient noise cross-correlations to estimate the Green's function at low frequencies (0.5-2 Hz) where noise from the Pacific Ocean is dominant, and at high frequencies (2-4 Hz) where traffic noise is dominant. This estimate is typically made by cross-correlating recordings of ambient seismic noise from two simultaneously recording receivers over a long period of time. This process recovers the Green's function and its time-reversed version between the two receivers, convolved with the autocorrelation of a source function such as noise (Wapenaar et al., 2010). Here, we implement a processing method adapted from Bensen et al. (2007). We first break up our time series into tapered, non-overlapping two-hour time windows and then bandpass accordingly. We then whiten the input traces prior to cross-correlating. In the frequency domain, the cross-correlation of pre-whitened traces can be expressed as

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

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, $*$ represents the complex conjugate, $\langle \cdot \rangle$ represents the time-averaged ensemble, $| \cdot |$ represents the real absolute value of the spectrum, and $\{\cdot\}$ represents a 0.003 Hz running window average used for whitening the signal.

By dividing the standard cross-correlation procedure by smoothed amplitude spectrums at the two receivers, we are deconvolving an estimate of the noise source signal from the correlations in the time domain to better estimate the Green's function (see Appendix). Pre-whitening limits the amplitude variation from trace to trace while maintaining the phase information and increasing the bandwidth of the estimated Green's function (Bensen et al., 2007). This step is particularly important at the frequencies we are looking at (under 4 Hz) because it removes the effect of the 3 Hz low-cut filter applied to the data. Although true amplitude information is lost, it is not a major problem at Long Beach because amplitudes from trace to trace can vary significantly and because the data is very noisy.

After whitening the input traces and cross-correlating each station pair, we normalize each result by the maximum amplitude. By doing this, we are further limiting the effect of spurious events and instrument spikes in our correlations. For the lower frequency range, we average correlations from 100 two-hour time windows, resulting in an average over 8 days. For the higher frequency range, we average correlations from 240 two-hour time windows, resulting in an average over nearly 3 weeks.

Results

We first show correlations in map view for a virtual source near the center of the array. Figure 9 and Figure 10 display snap shots for the lower frequency range and the higher frequency range, respectively, at -5 s and 5 s time lags. For frequencies below 2 Hz, the estimated Green's function shows significant directivity, as it is strongest in the southwest portion of the array at acausal time lags (when energy is moving toward the receiver) and strongest in the northeast portion of the array at causal time lags (when energy is moving away from the receiver). This suggests that the Pacific Ocean, which is south of the array, is a significant source of directed, low-frequency energy. For frequencies above 2 Hz, the estimated Green's function is more symmetric, suggesting that the noise source is more homogeneously distributed at these frequencies.

These correlation results are in line with our observations from the PSD and beamforming analysis. Noise characterization revealed that energy at low frequencies is dominantly generated by noise along the coastline in the southern part of the array, thereby explaining the inhomogeneous noise source distribution that is leading to strong asymmetry in the estimated Green's function. As we look at increasingly higher frequencies starting from 2 Hz, the noise source distribution becomes more homogeneous as the noise field becomes more dominated by anthropogenic sources (particularly traffic noise), thus leading to more symmetric Green's functions. The influence of Interstate 405 can be seen in the acausal snapshot, where the wavefront in the north (which is moving south from the freeway toward the virtual source) is relatively strong compared to the rest of the wavefront. Furthermore, we can see the dispersive nature of Rayleigh waves in vertically inhomogeneous media from these snapshots. At the time lags shown here, the moveout from the virtual source is greater

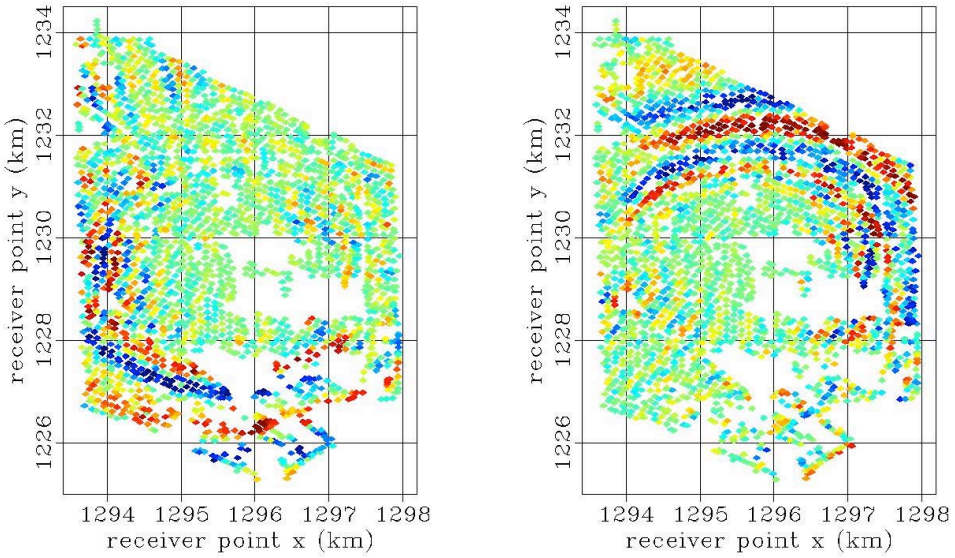


Figure 9: Snapshots from a virtual source in the center of the array at (a) -5 s lag and (b) 5 s lag for frequencies between 0.5 Hz and 2 Hz. Note the directivity of the Green's function. [**CR**] jason1/. mid-acausal-snap,mid-causal-snap

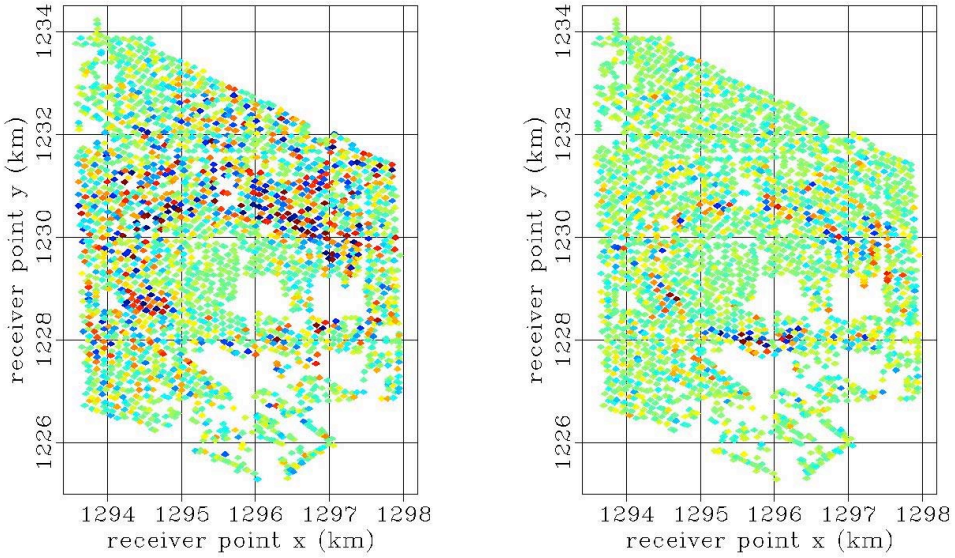


Figure 10: Snapshots from a virtual source in the center of the array at (a) -5 s lag and (b) 5 s lag for frequencies between 2 Hz and 4 Hz. Note the improved symmetry of the Green's function. [**CR**] jason1/. high-acausal-snap,high-causal-snap

at lower frequencies than at higher frequencies because lower frequencies sense deeper and are thus expected to travel faster.

To better measure these velocities, we display virtual common source gathers for lines of receivers shown in Figure 1(c). The lower-frequency source gather in Figure 11(a) is generated along a seismic line roughly perpendicular to the coast, while the higher-frequency source gather in Figure 11(b) is generated along a seismic line running north-south and perpendicular through a section of Interstate 405. Sources in stationary phase locations, such as locations at the ends of a seismic line, have the most influence over the estimates of Green's functions (Snieder, 2004). Therefore, the seismic lines are chosen so that the dominant noise sources (the Pacific Ocean and Interstate 405) are located in these stationary phase locations. The lower-frequency gather has the same virtual source as in the snapshots, while the higher-frequency gather has a virtual source in a different location in order to limit the gaps in the north-south seismic line. Negative offsets are stations to the south or southwest of the virtual source, and positive offsets are stations to the north or northeast of the virtual source. Acausal time lags represent energy moving toward the virtual source, while causal time lags represent energy moving away from the virtual source.

For the lower-frequency source gather, we see propagation of the fundamental Rayleigh wave from southwest to northeast at a velocity of approximately 630 m/s. The directivity of the wave is apparent from the asymmetry about zero-time lag. For the higher-frequency source gather, we see that the dominant Rayleigh wave is propagating from north to south. This wave energy begins approximately 1.5 km north of the virtual source, which is the location of Interstate 405. The dominant Rayleigh wave energy is generated by the freeway, and this is further evidenced by the correlating energy traveling northward at offsets greater than 1.5 km. At receivers south of the virtual source, the energy from the freeway has attenuated, allowing weaker Rayleigh wave energy traveling north from the Pacific Ocean toward the virtual source to become visible. Both these fundamental Rayleigh waves are traveling at approximately 250 m/s, which is slower than the fundamental Rayleigh wave velocity at lower frequencies, as we expect. Focusing further on stations south of the virtual source, we see hints of a higher-order Rayleigh wave mode traveling away from the virtual source at a velocity of approximately 500 m/s. The directionality of this wave mode suggests that it is being generated by Interstate 405.

To enhance the the higher-order Rayleigh wave mode, we construct a super source-gather, as seen in Figure 12(a). This is done by creating virtual source gathers from 100 randomly distributed source locations for each of the 240 two-hour time windows, sorting each resulting correlation by radial offset into 200 m bins, and then stacking. This amounted to stacking over 50 million correlations. The fundamental and higher-order Rayleigh wave modes are more distinct now, with velocities of approximately 250 m/s and 500 m/s, respectively. More importantly, we see the emergence of a refracted P-wave traveling at approximately 2 km/s, as also observed in Lin et al. (2012). We display the same super source-gather for a frequency range of 3-4 Hz to bring out this arrival in Figure 12(b). One may speculate that the P-wave energy is generated from converted surface wave energy due to heterogeneities in Earth's crust, but since these P-waves are more distinct at frequencies higher than 3 Hz, there is a chance they could be generated by anthropogenic sources such as traffic noise. For this P-wave arrival to be useful, though, we must extract it from correlations between a single station pair. This is one of our research focuses for the coming year. Regardless, the presence of a refracted P-wave from ambient noise cross-correlations

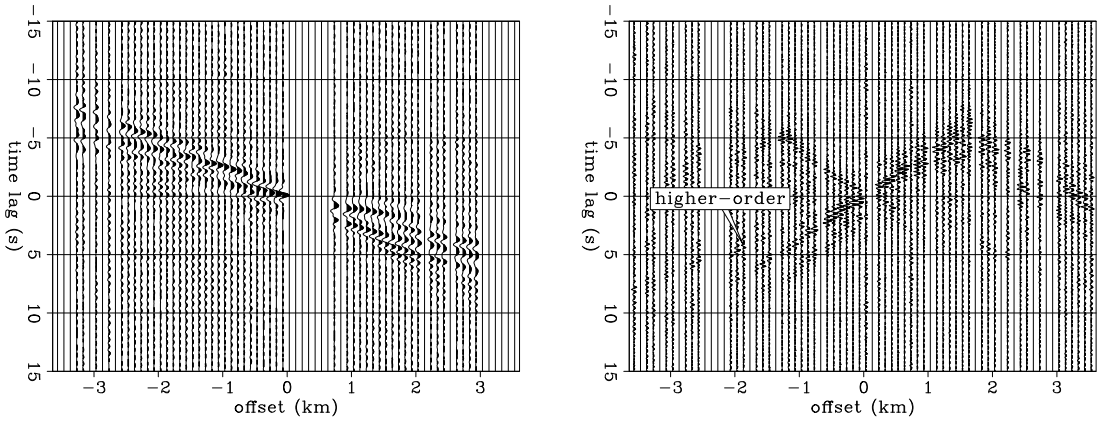


Figure 11: (a) Common source gather for frequencies between 0.5 Hz and 2 Hz. (b) Common source gather for frequencies between 2 Hz and 4 Hz. [**CR**] jason1/. line1a-lo,line2a-hiA

at Long Beach is a first step toward potentially retrieving other body-wave arrivals.

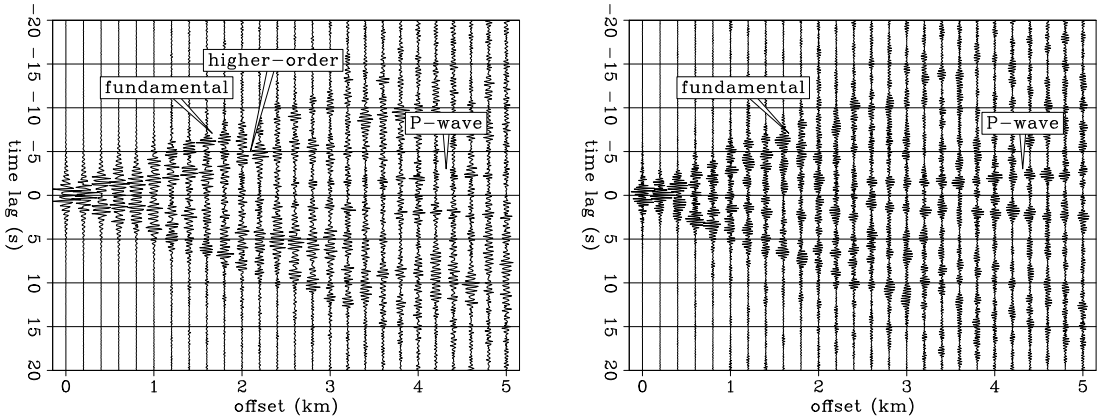


Figure 12: (a) super source-gather for frequencies between 2 Hz and 4 Hz created from stacking over 50 million correlations. Note the emergence of a P-wave traveling at approximately 2 km/s. (b) Same super source-gather as in (a), but for frequencies between 3 Hz and 4 Hz. Note that the P-wave arrival is slightly clearer now. [**CR**] jason1/. superdupergather1A,superdupergather1-34A

CONCLUSIONS AND FUTURE WORK

The distribution of noise sources plays a significant role in the quality of the Green's functions retrieved from ambient noise cross-correlations. Thus, the primary goal of our paper is to characterize the noise at Long Beach. By looking at PSD spatial distribution maps, we found that the Pacific Ocean dominates the noise field at frequencies below 2 Hz, and that human activity dominates the noise field at frequencies above 2 Hz. By looking at

spectrograms, we found that noise energy goes through daily fluctuations for frequencies above 2 Hz. These fluctuations fall in line with periods of human activity. Beamforming analysis revealed that the influence of the Pacific Ocean noise is consistent throughout the array, while the influence of Interstate 405 has a limited extent. Away from the freeway, noise arrives at a wider range of azimuths at higher frequencies. From these results, we expect the noise field to be more homogeneous at higher frequencies than at lower frequencies. Ambient noise cross-correlation results from 0.5-2 Hz and 2-4 Hz confirm these observations, with estimated Green's functions showing strong asymmetry at lower frequencies and better symmetry at higher frequencies. We also created virtual source gathers and found the fundamental Rayleigh wave to be dispersive by having velocities of approximately 630 m/s at the lower-frequency range and approximately 250 m/s at the higher-frequency range. We also observed a weak high-order Rayleigh wave mode traveling at 500 m/s at the higher-frequency range. Lastly, we created a super source-gather, which revealed a refracted P-wave traveling at 2 km/s.

Our future work will focus on exploiting the station density and time duration of the Long Beach dataset for surface wave tomographic studies, with an eye toward time-lapse monitoring. The parameters of the array allow for the retrieval of high-frequency (¡ 2 Hz) Green's functions from cross-correlations, which will provide better subsurface resolution at shallow depths. This level of resolution is particularly useful for earthquake hazard analysis, which is of importance due to the presence of the Newport-Inglewood fault running through the region. Furthermore, we want to improve the refracted P-wave arrival. We obtained that arrival by stacking over many virtual sources, so we would like to retrieve the P-wave for a single virtual source. A potential strategy is to employ a more selective stacking procedure. Also, because the P-wave arrival was enhanced at frequencies above 3 Hz, we want to determine if it was generated by the strongest anthropogenic source in Long Beach: Interstate 405. If we can show that the freeway generated those refracted P-waves, then we can use it as a passive, yet very active, seismic source.

ACKNOWLEDGMENTS

We gratefully acknowledge Signal Hill Petroleum, Inc. for access to this dataset and permission to publish. Thanks to Dan Hollis for his enthusiasm and cooperation throughout this research. Thanks to Stew Levin and Bob Clapp for their help converting the data.

APPENDIX

Estimates of Green's functions from ambient seismic noise are typically obtained by cross-correlating traces from two simultaneously recording receivers. This process recovers the Green's function and its time-reversed version between the two receivers, convolved with the autocorrelation of a source function such as noise (Wapenaar et al., 2010). In the frequency domain, this can be expressed as

$$[G(x_B, x_A, \omega) + G^*(x_B, x_A, \omega)]|S_N(\omega)|^2 = \langle U(x_B, \omega)U^*(x_A, \omega)\rangle \quad , \tag{A-1}$$

where G is the Green's function between two receiver locations (x_A, x_B), $|S_N(\omega)|^2$ is the power spectrum of the noise, $U(x, \omega)$ is the spectrum of the wavefield at a given receiver location, $*$ represents the complex conjugate, and $\langle\cdot\rangle$ represents the time averaged ensemble.

This equation suggests that to accurately estimate Green's functions from cross-correlations of ambient noise, the power spectrum of the noise must be known. However, the power spectrum of noise is typically not known. Furthermore, PSD analysis at Long Beach revealed that major sources of noise, particularly of anthropogenic origin, have highly-varying power spectrums in time and space.

To better estimate the Green's function between two receivers, we whiten the input traces prior to cross-correlation. In the frequency domain, this pre-whitened cross-correlation can be expressed as

$$[G(x_B, x_A, \omega) + G^*(x_B, x_A, \omega)] = \left\langle \left(\frac{U(x_B, \omega)}{\{|U(x_B, \omega)|\}} \right) \left(\frac{U^*(x_A, \omega)}{\{|U(x_A, \omega)|\}} \right) \right\rangle \quad , \tag{A-2}$$

where $|\cdot|$ indicates the real absolute value of the spectrum and $\{\cdot\}$ indicates a 0.003 Hz running window for whitening the signal. In words, we divide the spectrum of each input trace by its smoothed amplitude spectrum prior to computing the cross-correlation. The denominator of the right side of Equation A-2 can be viewed as an estimate of the unknown power spectrum of noise in Equation A-1:

$$\{|U(x_B, \omega)|\}\{|U(x_A, \omega)|\} \approx |S_N(\omega)|^2 \quad . \tag{A-3}$$

By dividing the standard cross-correlation procedure by an estimate of the power spectrum of noise in the frequency domain, we are trying to deconvolve the noise source signal from the correlations in the time domain to better estimate the Green's function. Although this process compromises the amplitude information of the correlation, it does not effect the phase information. We then normalize all the correlations by their respective maximum amplitude to limit the effect of anomalous events and instrument spikes on the correlation results.

To demonstrate the effectiveness of pre-whitening the correlations, we apply multiple processing techniques to the ambient seismic noise to create similar virtual source gathers as seen in the main body. Figure A-1 shows gathers at low frequencies (0.5 Hz to 2 Hz), while Figure A-2 shows gathers at high frequencies (2 Hz to 4 Hz). In both figures, (a) represents the standard correlation procedure, (b) represents the standard correlation procedure with normalization, (c) represents the correlation procedure with pre-whitening, and (d) represents the correlation procedure with pre-whitening and normalization (which is what is displayed in the main body). Again, for the lower frequency range we average correlations from 100 two-hour time windows, and for the higher frequency range we average correlations from 240 two-hour time windows.

At both frequency ranges, we see that the standard correlation procedure allows anomalous events to overwhelm the correlation results, rendering the results unusable. Normalization of the correlations prior to stacking greatly improves the results, allowing the moveouts to be clearly observed. However, there is clearly a bias toward the higher frequencies in both cases. Pre-whitening broadens the spectrums of the correlations and balances the amplitudes, thus improving the quality of the correlations. Although further amplitude normalization of each correlation prior to stacking does not greatly enhance the signal to noise ratio of the resulting gather at lower frequencies, it does slightly improve the signal to noise ratio for the resulting gather at higher frequencies.

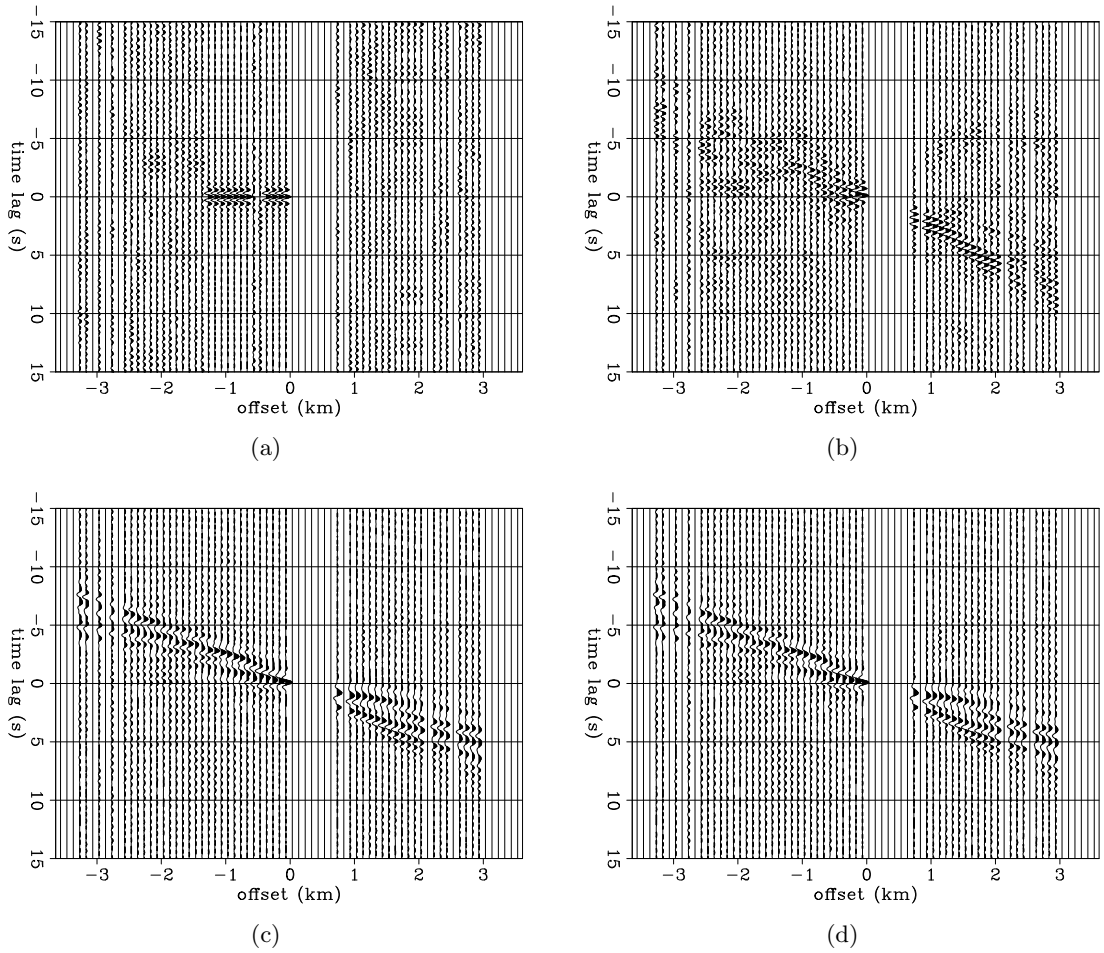


Figure A-1: Virtual source gathers at the frequency range 0.5 Hz to 2 Hz from various processing procedures. (a) standard correlation procedure, (b) standard correlation procedure with normalization, (c) correlation procedure with pre-whitening, and (d) correlation procedure with pre-whitening and normalization. [**CR**] jason1/. line1-old-lo,line1-old-norm-lo,line1-nnorm-lo,line1a-lo

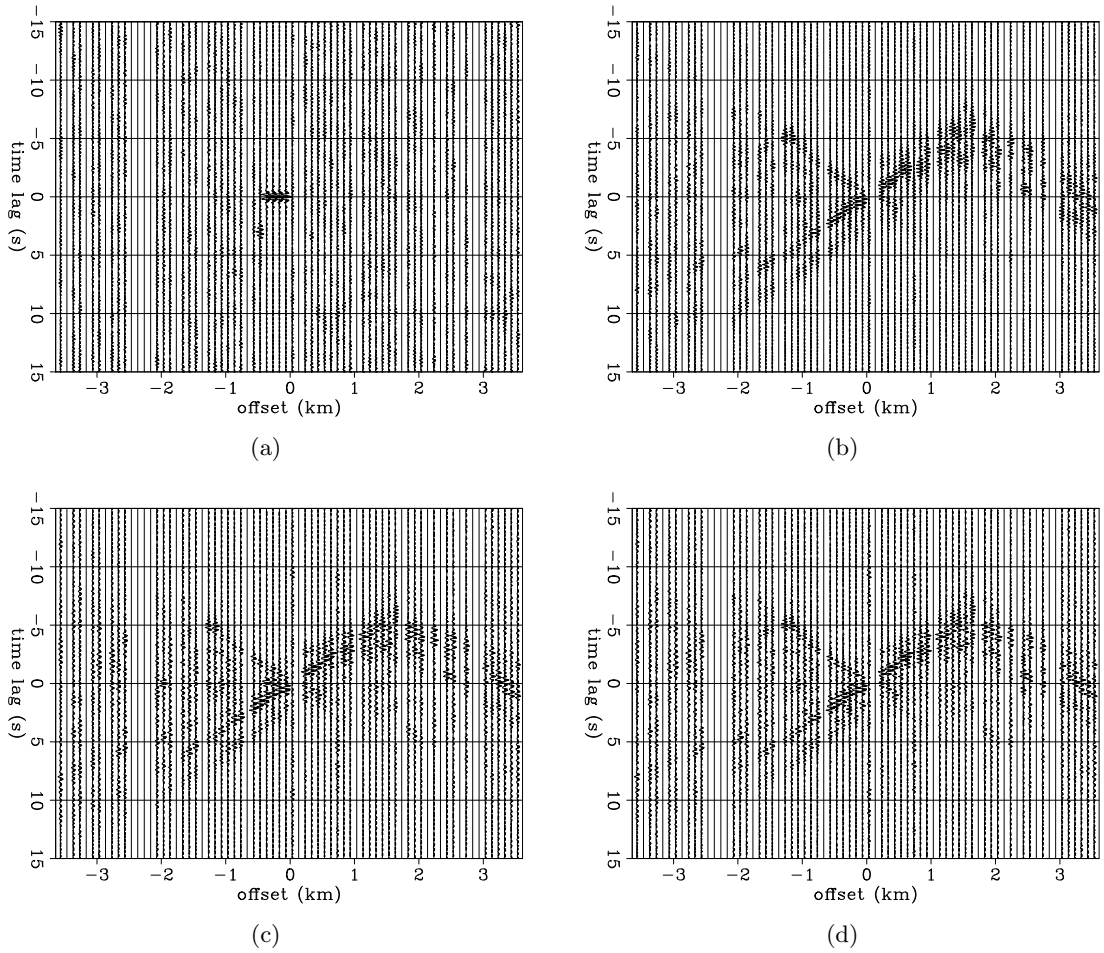


Figure A-2: Virtual source gathers at the frequency range 2 Hz to 4 Hz from various processing procedures.(a) standard correlation procedure, (b) standard correlation procedure with normalization, (c) correlation procedure with pre-whitening, and (d) correlation procedure with pre-whitening and normalization. [**CR**] jason1/. line2-old-hi,line2-old-norm-hi,line2-nnorm-hi,line2a-hi

REFERENCES

Bensen, G., M. Ritzwoller, M. Barmin, A. Levshin, F. Lin, M. Moschetti, N. Shapiro, and Y. Yang, 2007, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements: Geophysical Journal International, **169**, 1239–1260.

Bensen, G., M. Ritzwoller, and N. Shapiro, 2008, Broadband ambient noise surface wave tomography across the United States: Journal of Geophysical Research, **113**, 1–21.

Delgado, J., C. Lopez Casado, A. Estevez, J. Giner, A. Cuenca, and S. Molina, 2000, Mapping soft soils in the segura river valley (se spain): a case study of microtremors as an exploration tool: Journal of Applied Geophysics, **45**, 19–32.

Draganov, D., X. Campman, J. Thorbecke, A. Verdel, and K. Wapenaar, 2009, Reflection images from ambient seismic noise: Geophysics, **74**, A63–A67.

Draganov, D., K. Wapenaar, W. Mulder, J. Singer, and A. Verdel, 2007, Retrieval of reflections from seismic background-noise measurements: Geophysical Research Letters, **34**, 1–4.

Greenwoood, R., 1998, Liquefaction zones in the Long Beach 7.5-minute quadrangle, Los Angeles County, California: California Department of Conservation, Division of Mines and Geology, Seismic Hazard Zone Report, **28**, 3–20.

Lin, F.-C., D. Li, R. Clayton, and D. Hollis, 2012, Interferometry with a dense 3D dataset: SEG Technical Program Expanded Abstracts, **1**, 1–6.

McNamara, D. and R. Boaz, 2005, Seismic noise analysis system using power spectral density probability density functions: A stand-alone software package: USGS Open-File Report, 1–14.

Poli, P., H. Pedersen, and M. Campillo, 2012, Emergence of body waves from cross-correlation of short period seismic noise: Geophysical Journal International, **188**, 549–558.

Rost, S. and C. Thomas, 2002, Array seismology: Methods and applications: Reviews of Geophysics, **40**, 2–1–2–27.

Roux, P., K. Sabra, P. Gerstoft, W. Kuperman, and M. Fehler, 2005, P-waves from cross-correlation of seismic noise: Geophysical Research Letters, **32**, 1–4.

Sabra, K., P. Gerstoft, P. Roux, W. Kuperman, and M. Fehler, 2005, Surface wave tomography from microseisms in Southern California: Geophysical Research Letters, **32**, 1–4.

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

Snieder, R., 2004, Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase: Physical Review E, **69**, 046610–1–046610–8.

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.

Yang, Y., M. Ritzwoller, F. Lin, M. Moschetti, and N. Shapiro, 2008, Structure of the crust in uppermost mantle beneath the western United States revealed by amibent noise and earthquake tomography: Journal of Geophysical Research, **113**, 1–9.

Zhan, Z., S. Ni, D. Helmberger, and R. Clayton, 2010, Retrieval of Moho-reflected shear wave arrivals from ambient seismic noise: Geophysical Journal International, **182**, 408–420.

One-way extrapolation of seismic waves in elastic isotropic media with mode conversions

Musa Maharramov

ABSTRACT

In this paper, I present a computationally efficient technique for extrapolating seismic waves in an isotropic elastic medium. The method is based on factoring the full elastic wave equation into the product of two "one-way" pseudo-differential operators. The operators are shown to be stable in their respective propagation directions. $P-SV$ mode conversions are modeled by imposing continuity of tractions in the direction of maximum model contrast. I test the method on heterogeneous models featuring both sharp contrasts and smooth gradient, and compare results with a two-way time domain modeling method. I achieve a significant reduction in the cost of elastic imaging compared to the currently prevalent time and frequency domain numeric methods.

INTRODUCTION

Extrapolation of seismic wave fields in depth using one-way operators is an efficient alternative to time and frequency domain modeling with the full wave equation, particularly in seismic migration (Biondi, 2005). While one-way extrapolators have long been established as key components of the seismic imaging toolkit for acoustic problems, extrapolation of elastic wave fields is typically carried out by solving a full elastodynamic system either in the time or frequency domain – either approach being computationally expensive. The high computational cost of wave extrapolation in elastic media is one of the barriers to widespread adoption of multicomponent seismic in industrial applications. Some progress has been made recently in the development of efficient one-way methods for certain simple elastic anisotropies such as vertically transversely isotropic or tilted transversely isotropic media (Shan, 2007; Nolte, 2008; Maharramov and Nolte, 2011). However, these methods are based on the "pseudoacoustic" approximation (Grechka, 2009) and are used only for a kinematically accurate description of pressure wave propagation.

Here I present a method for one-way frequency domain extrapolation of displacement fields in an elastic isotropic medium. My approach is based on factoring the elastic wave equation using pseudo-differential operators acting on displacement vector fields. This approach is conceptually similar to the derivation of the acoustic single square-root equation (Claerbout, 1985), except that the resulting factorized propagation operators are not obtained analytically but computed numerically. Mode conversions are modeled by imposing a condition of continuity of tractions in the direction of maximum model change at each depth step. Other potentially computationally advantageous techniques for extrapolating elastic wave fields in isotropic media include matrix propagator methods (Segall, 2010; Ursin, 1983) and elastic screen methods (Wu, 1994; Xie and Wu, 2005). The proposed method removes the lateral homogeneity constraints of the former while not relying on the

scattering approximation and wide-angle corrections of the latter. Note that kinematically accurate modeling of mode conversions by my new method relies on the assumption of a single direction of maximum gradient for *all* medium parameters at each point. However, this assumption is not unreasonable from a geological point of view as elastic mode conversions typically occur at the boundaries between different sediments.

THE METHOD

I start with the wave equation governing displacements in an isotropic elastic medium in the Navier form (Segall, 2010):

$$-\rho\ddot{u^i} + \mu\Delta u^i + \frac{\mu}{1-2\nu}\frac{\partial}{\partial x_i}\frac{\partial u_k}{\partial x_k} = 0,\ i = 1,2,3, \tag{1}$$

where $u^{1,2,3}$ denote the components of a displacement field, μ is the shear modulus, ν is Poisson's ratio and ρ is density of the medium. In this paper, I consider a heterogeneous elastic medium under the assumption of local homogeneity. This allows me to factor the elastic moduli outside of the differentiation operators in equation 1. However, the assumption of local homogeneity will be later dropped, allowing for the effects of heterogeneity in propagation and mode conversion.

Equation 1 can be solved either in time or frequency domain. Time domain methods suffer from stability-related restrictions on their time steps. Frequency-domain methods allow a reduction in the range of temporal frequencies in propagation and offer the additional flexibility of parallelizing the computations for different frequencies. However, these methods require the solution of very large sparse systems of linear equations, unless the instability of the two-way depth propagation can be resolved. For an overview, see Virieux et al. (2009).

In my method, after Fourier-transforming equation 1 in time, I seek the following factorization of the operator on the left-hand side:

$$\begin{aligned}&\left(E(\lambda,\mu)\frac{\partial}{\partial z} + A\left(-i\frac{\partial}{\partial x}, -i\frac{\partial}{\partial y}\right) + c_\omega I\right)\times\\ &\left(E(\lambda,\mu)\frac{\partial}{\partial z} + B\left(-i\frac{\partial}{\partial x}, -i\frac{\partial}{\partial y}\right) + c_\omega I\right)\mathbf{u} = 0,\end{aligned} \tag{2}$$

where

$$\begin{aligned}E(\lambda,\mu) &= \begin{bmatrix}\sqrt{\mu} & 0 & 0\\ 0 & \sqrt{\mu} & 0\\ 0 & 0 & \sqrt{\lambda+2\mu}\end{bmatrix},\\ c_\omega &= \sqrt{\rho}\omega,\end{aligned} \tag{3}$$

A, B are 3×3 matrices with components that are complex-valued functions of the horizontal wave numbers, and I is the 3×3 identity matrix.

Factorization 2 immediately suggests a "one-way" extrapolation by using only one of the factors

$$\tilde{\mathbf{u}}(x,y,z+\Delta z) = \exp\left[-\Delta z E^{-1}(A(-i\partial_x, -i\partial_y) + c_\omega I)\right]\tilde{\mathbf{u}}(x,y,z), \tag{4}$$

or even a "down and up" method where downward extrapolation by equation 4 is followed by upward extrapolation using the other factor in 2. Such a method is numerically stable if the spectrum of matrix

$$A(k_x, k_y) + c_\omega I \tag{5}$$

is not in the interior of the left complex half-plane, and the spectrum of

$$B(k_x, k_y) + c_\omega I \tag{6}$$

is not in the interior of the right complex half-plane. This appears to be the case for all the tests carried out so far.

Extrapolation through a heterogeneous medium is handled using the same approach as in the "phase shift plus interpolation" method (PSPI) for acoustic extrapolation (Gazdag and Sguazzero, 1984). More specifically, at each depth step the wave field is extrapolated separately using one-way operators for a set of reference medium parameters, and the desired wave field at the new depth level is then computed using linear interpolation among the reference wave fields at each (x, y). This approach neglects the contribution of terms that contain spatial derivatives of the elastic moduli in the full elastodynamic system. These terms can be neglected so long as the characteristic wavelength of seismic waves is shorter than the model resolution, i.e. for smoothly varying medium parameters. However, these terms become important where the moduli change significantly – for example, near sharp contrasts. I handle such contrasts by imposing the condition of continuity of tractions in the direction of maximum model contrast at each depth step *before* the extrapolation:

$$\begin{aligned}
\mu \left(\frac{\partial u^1}{\partial z'} + \frac{\partial u^3}{\partial x'} \right) \Big|_{\text{above}} &= \mu \left(\frac{\partial u^1}{\partial z'} + \frac{\partial u^3}{\partial x'} \right) \Big|_{\text{below}}, \\
\mu \left(\frac{\partial u^2}{\partial z'} + \frac{\partial u^3}{\partial y'} \right) \Big|_{\text{above}} &= \mu \left(\frac{\partial u^2}{\partial z'} + \frac{\partial u^3}{\partial y'} \right) \Big|_{\text{below}}, \\
2\mu \frac{\partial u^3}{\partial z'} + \lambda \left(\frac{\partial u^1}{\partial x'} + \frac{\partial u^2}{\partial y'} + \frac{\partial u^3}{\partial z'} \right) \Big|_{\text{above}} &= 2\mu \frac{\partial u^3}{\partial z'} + \lambda \left(\frac{\partial u^1}{\partial x'} + \frac{\partial u^2}{\partial y'} + \frac{\partial u^3}{\partial z'} \right) \Big|_{\text{below}},
\end{aligned} \tag{7}$$

where both sides of equations 7 have the same values above and below a medium interface, and the rotated coordinates (x', y', z') are such that the axis z' is orthogonal to the interface, i.e. is in the direction of maximum contrast. Imposing conditions 7 at each depth step $z = z_k$ effectively means solving a system of first-order partial differential equations in the interval $z \in (z_k, z_k + \Delta z)$, given the values of $u^i(x, y, z)$ for $z \leq z_k$. Note that the rotation $(x, y, z) \rightarrow (x', y', z')$ in equations 7 is local.

IMPLEMENTATION AND RESULTS

The system of matrix equations 2 that results from factorization 2 (see Appendix) is solved only once for each set of temporal frequency and elastic moduli values, and for each pair of horizontal wave numbers. In my prototype implementation of the one-way extrapolator I compute the matrices A, B at the beginning of the frequency loop and subsequently use the tabulated matrices in the inner depth extrapolation loop. A more efficient approach can be employed in a production implementation: system 2 can be solved using Newton's method (Higham, 2008) in a one-off computation for each set of temporal frequency, elastic moduli and horizontal wave numbers, and stored in a look-up table. The symmetry of

the extrapolation operator 4, which appears to be a multi-component counterpart of the acoustic phase-shift operator (Claerbout, 1985), can be exploited to achieve a substantial reduction in the size of the precomputed operator tables. The real parts of the eigenvalues of operator

$$K = -\Delta z E^{-1} \left[A(-i\partial_x, -i\partial_y) + c_\omega I \right], \tag{8}$$

are zero within the operator pass band and negative outside (see Figure 1). Operator K of equation 8 is the logarithm of the extrapolation operator 4, and the spectral plots of Figures 1(b) and 1(d) correspond to the phase of the phase-shift extrapolator in the acoustic case (Biondi, 2005). The crucial difference in the elastic multicomponent case is that the multicomponent "phase-shift" is defined by three scalar phase-shift operators with phases ϕ_1, ϕ_2, ϕ_3, and a unitary operator Q, determined by the eigenvector expansion:

$$K = Q \begin{bmatrix} i\phi_1 & 0 & 0 \\ 0 & i\phi_2 & 0 \\ 0 & 0 & i\phi_3 \end{bmatrix} Q^*. \tag{9}$$

The pass bands of the three phase shift operators are generally different, but the real parts of the eigenvalues of 8 are non-positive across all three pass bands.

Figures 2 and 3 demonstrate the results of applying my method to computing the impulse response for a horizontal tensile crack source at the surface (Aki and Richards, 2009) using the one-way phase-shift operator of equation 4 with periodic boundary conditions. Homogeneous medium parameters used in this test were 374 m/s shear-wave velocity and 616 m/s pressure-wave velocity. The extrapolation grid was $128 \times 128 \times 128$ with a 5 m step, and the frequency range was 1-32 Hz with a 1 Hz step. Note the good agreement of Figure 2 with the expected radiation pattern from a horizontal crack: the vertical displacement has only pressure-wave contributions, while horizontal displacements have both pressure and shear-related contributions. The pressure and shear wave plots of Figure 3 were obtained as the divergence and curl of the displacement field, respectively. The positions of the pressure and shear wave responses match the distances traveled at their respective velocities in 1 `s`. Wraparound images of waves at multiples of 1 `s` can also be seen, as expected with the periodic boundary conditions.

Figures 4(a) and 4(b) show two models used in testing depth extrapolation through heterogeneous media (only the shear moduli are plotted). The first model of Figure 4(a) features a sharp contrast between the interior and exterior of a 2.5D wedge, with the pressure and shear wave velocities $v_p^1 = 616$ m/s, $v_s^1 = 374$ m/s outside of the wedge, and $v_p^2 = 509$ m/s, $v_s^2 = 316$ m/s inside. Wave propagation is (mostly) from the less compliant region above the wedge into the more compliant interior of the wedge. In the second model of Figure 4(b) the moduli have smooth gradients within the respective regions:

$$\begin{aligned} \lambda_{\text{outside}} &= (1 - \theta(i_x))\, \lambda_1 + \theta(i_x)\lambda_2, \\ \mu_{\text{outside}} &= (1 - \theta(i_x))\, \mu_1 + \theta(i_x)\mu_2, \\ \lambda_{\text{inside}} &= (1 - \theta(i_x))\, \lambda_2 + \theta(i_x)\lambda_1, \\ \mu_{\text{inside}} &= (1 - \theta(i_x))\, \mu_2 + \theta(i_x)\mu_1, \end{aligned} \tag{10}$$

where indices 1 and 2 denote values of the moduli outside and inside of the wedge region from Figure 4(a), $\theta(i_x) = 2(\min(i_x, N_x - i_x + 1) - 1)/N_x$, and i_x, N_x are the inline index and

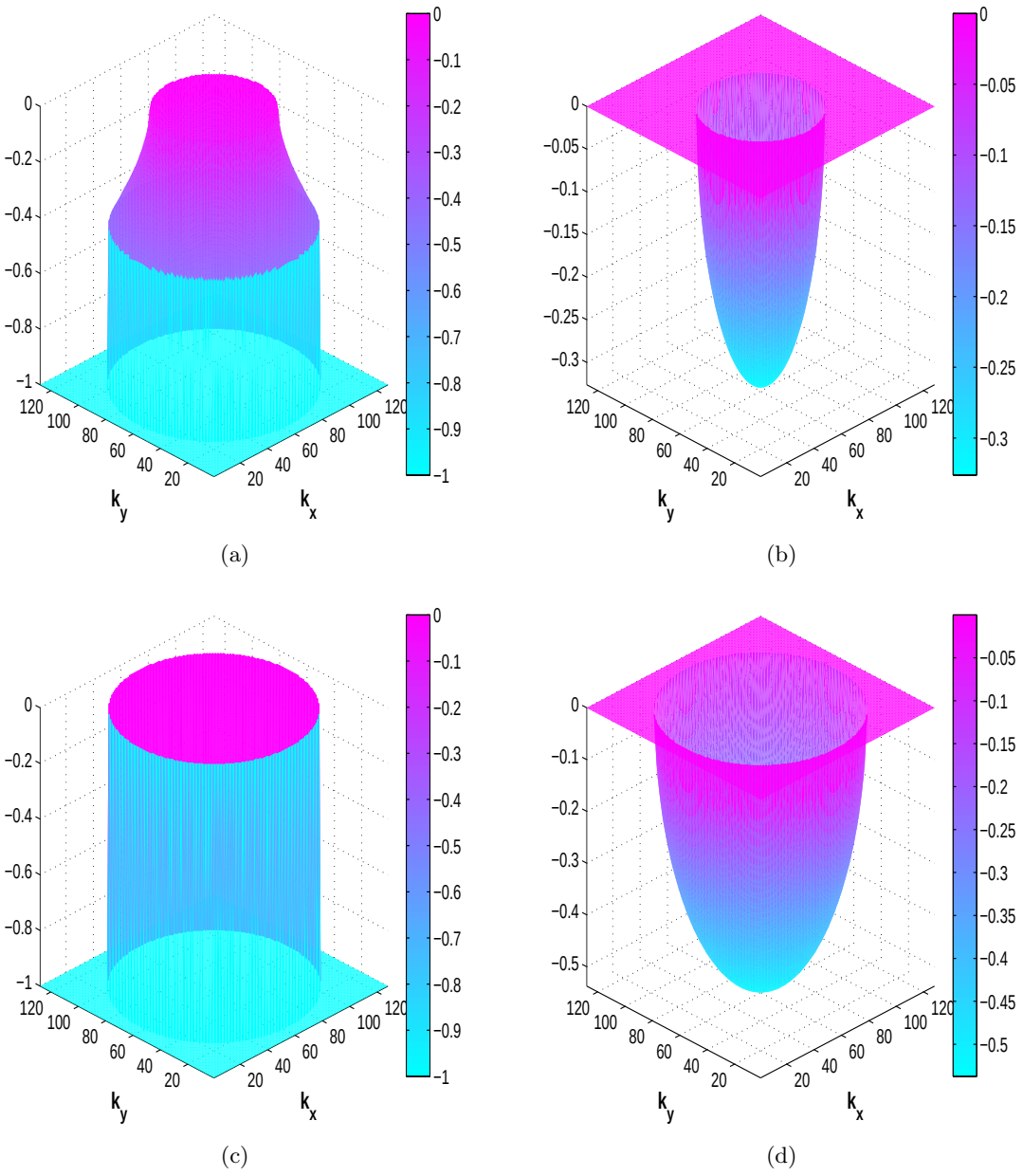
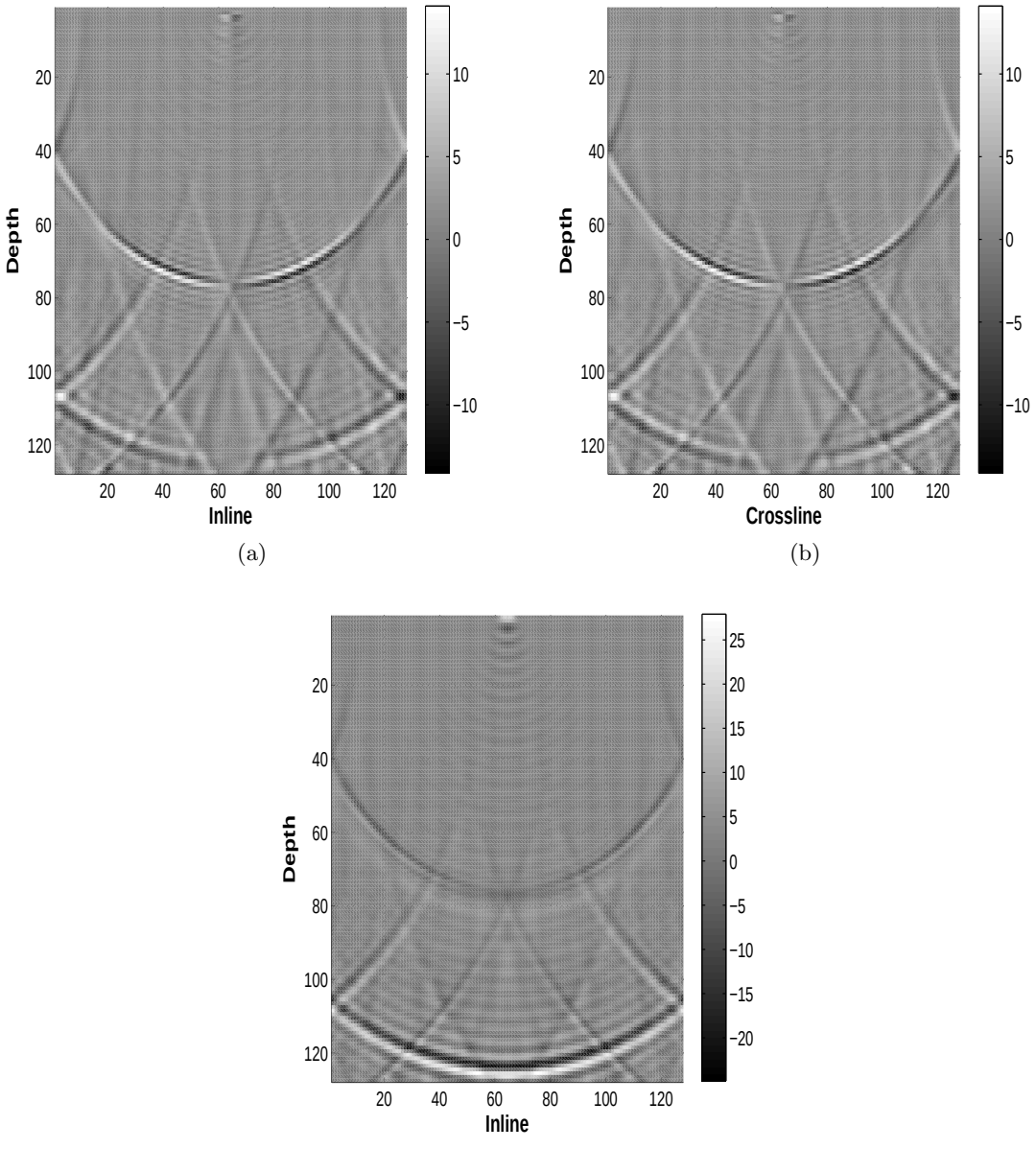


Figure 1: Real and imaginary parts of the first (a,b) and second/third (c,d) eigenvalues of K. Note that the real parts are non-positive, ensuring stability of the extrapolation. [**CR**] musa/. RE1,IM1,RE2,IM2



(a) (b)

(c)

Figure 2: Extrapolation in a homogeneous elastic medium; (a) inline section of inline displacement; (b) crossline section of crossline displacement; (c) inline section of vertical displacement. [**CR**] musa/. dxcy,dycx,dzcy

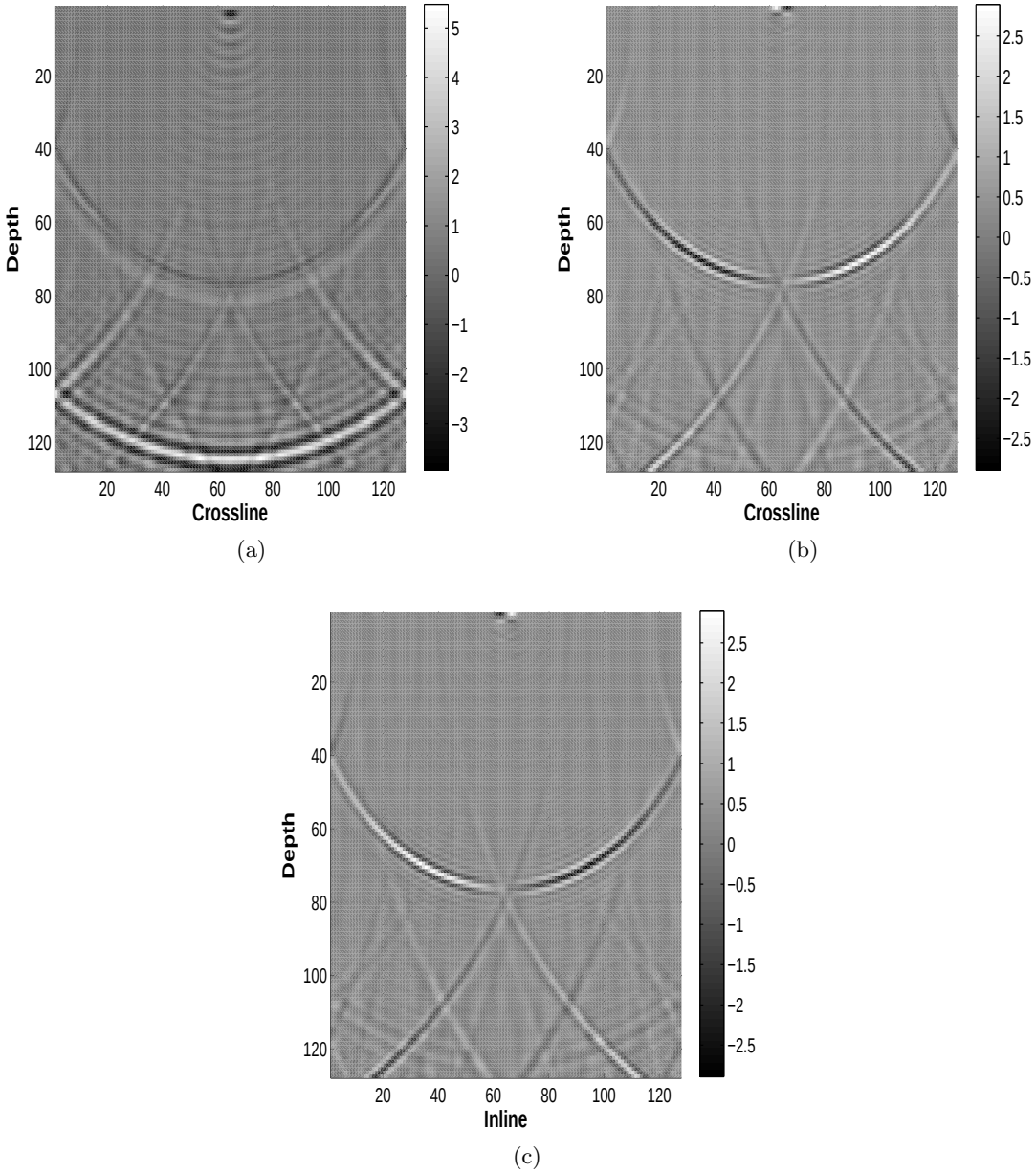
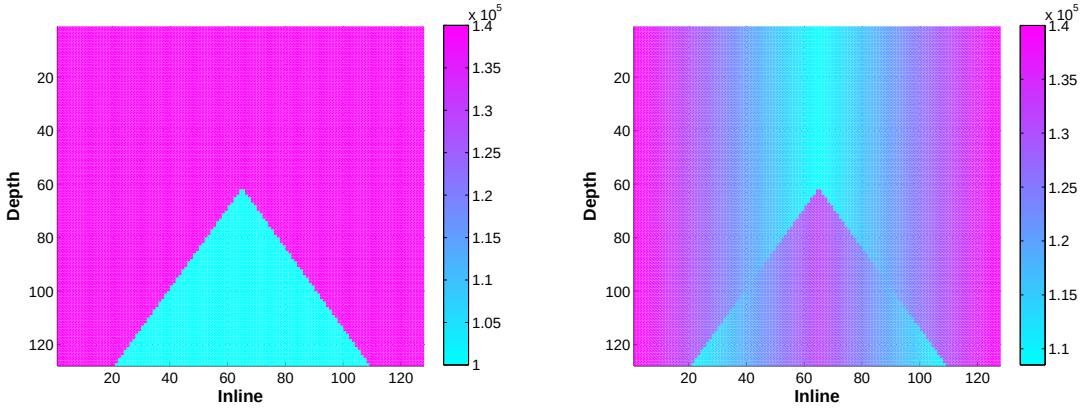


Figure 3: Extrapolation in a homogeneous elastic medium;(a) crossline section of pressure; (b) crossline section of inline shear component; (c) inline section of crossline shear component. The pressure and shear waves are computed as the divergence and curl of the displacement field. [**CR**] musa/. pcx,s1cx,s2cy

axis size. Density ρ is constant in both cases. Note the change of the gradient sign along the sharp interface of Figure 4(b), meaning that during propagation the more compliant medium may be on either side of the interface.



(a) (b)

Figure 4: Extrapolation in heterogeneous media: (a) sharp contrast and (b) sharp contrast + smooth background. Only the shear modulus μ is shown. Note the changing sign of the gradient along the sharp interface in the second model. [**CR**] musa/. H1MU,H2MU

As both heterogeneous models feature a 2.5D wedge, the rotation $(x, y, z) \rightarrow (x', y', z')$ in equations 7 along the sharp interface is around the crossline axis $y = y'$. The rotation angle is pre-computed and stored in a 3D array $\alpha(x, y, z)$ along with the elastic medium parameters. Before downward-continuing the wave field from z_k to $z_{k+1} = z_k + \Delta z$, values of the displacement field at $z = z_k$ are updated to satisfy the dynamic conditions 7. First, wave field gradients are computed in the rotated coordinate system at each (x, y):

$$\begin{aligned}
\frac{\partial u^i(z_{k-1})}{\partial z'} &\approx \cos\alpha \frac{u^i(z_k) - u^i(z_{k-1})}{\Delta z} + \sin\alpha \frac{\partial u^i(z_{k-1})}{\partial x}, \\
\frac{\partial u^i(z_{k-1})}{\partial x'} &\approx -\sin\alpha \frac{u^i(z_k) - u^i(z_{k-1})}{\Delta z} + \cos\alpha \frac{\partial u^i(z_{k-1})}{\partial x}.
\end{aligned} \tag{11}$$

Then $u(z_k)$ is updated to ensure the continuity of traction in the rotated coordinate system:

$$\begin{aligned}
\frac{\partial u^1(z_k)}{\partial z'} &= \left(\frac{\partial u^3(z_{k-1})}{\partial x'} + \frac{\partial u^1(z_{k-1})}{\partial z'}\right)\frac{\mu_{\text{above}}}{\mu_{\text{below}}} - \frac{\partial u^3(z_{k-1})}{\partial x'}, \\
\frac{\partial u^2(z_k)}{\partial z'} &= \left(\frac{\partial u^3(z_{k-1})}{\partial y} + \frac{\partial u^1(z_{k-1})}{\partial z'}\right)\frac{\mu_{\text{above}}}{\mu_{\text{below}}} - \frac{\partial u^3(z_{k-1})}{\partial y}, \\
\frac{\partial u^3(z_k)}{\partial z'} &= \Bigg[(2\mu_{\text{above}} + \lambda_{\text{above}})\frac{\partial u^3(z_{k-1})}{\partial z'} \\
&\quad + \lambda_{\text{above}}\left(\frac{\partial u^1(z_{k-1})}{\partial x'} + \frac{\partial u^2(z_{k-1})}{\partial y}\right) \\
&\quad - \lambda_{\text{below}}\left(\frac{\partial u^1(z_{k-1})}{\partial x'} + \frac{\partial u^2(z_{k-1})}{\partial y}\right)\Bigg]/(2\mu_{\text{below}} + \lambda_{\text{below}}).
\end{aligned} \tag{12}$$

Finally, the updates are rotated back to the original coordinates:

$$u^i(z_k) \;=\; u^i(z_{k-1}) + \Delta z \left(\cos\alpha \frac{\partial u^i(z_k)}{\partial z'} - \sin\alpha \frac{\partial u^i(z_k)}{\partial x'} \right). \tag{13}$$

Note that steps 11,12, and 13 mean that equations 7 are solved explicitly in depth, and $|\alpha| < 90^\circ$. The above procedure can be modified to solve for the updated $u^i(x, y, z_k)$ implicitly, thus allowing for purely vertical conversion surfaces. However, this latter approach is not pursued in this work. As will be apparent from the subsequent comparison with two-way time domain modeling results, modeling of mode conversion through purely vertical interfaces is not required for the test models of Figures 4(a) and 4(b).

Figures 5(a) and 5(b) show the inline-component impulse response for the first heterogeneous model of Figure 4(a) computed using the proposed one-way method and a 2D second-order time domain staggered-grid finite-difference method (Virieux, 1986). The finite-difference plot shows the negated inline-component particle velocity. The larger computational grid in the FDTD test was dictated by considerations of accuracy and numerical dispersion. A linear 5-point boundary taper was used in the one-way code while periodic boundary conditions were used in the finite-difference code. Figures 5(c) and 5(d) show only the shear and pressure waves extracted from the 3D displacements extrapolated using the one-way method. Results of the two methods appear to be in a very good kinematic agreement. Note the conversion of the pressure wave into shear waves at the wedge boundaries. Figures 6(a) and 6(b) show the results for the second model of Figure 4(b). Again, note the very good kinematic agreement between the one-way and FDTD results. Because of the way the model was constructed, only two reference models were used in the PSPI.

CONCLUSIONS AND PERSPECTIVES

I have proposed a method that delivers the same magnitude of computational speed-up in comparison with the existing methods of elastic wave modeling as its acoustic counterpart does in comparison with the two-way acoustic wave modeling. Key to this speed-up is the ability to control the range of temporal frequencies included in the propagation.

The elastic operator factorization provides a natural generalization of the acoustic one-way operator factorization. One important difference from the acoustic one-way equation is that the present method extrapolates displacements, while the one-way acoustic method extrapolates the pressure field, hence the stress tensor. The continuity of normal tractions at sharp velocity interfaces is handled naturally by the acoustic one-way method, but has to be imposed as an additional constraint in the elastic one-way method by using dynamic constraints at each depth step. The PSPI step accurately accounts for the effects of smooth parameter variation, where the medium properties do not change significantly over the characteristic seismic wavelength. The dynamic constraints account for the effects of sharp model contrasts, where the elastic parameters change significantly over the characteristic wavelength. The assumption of a common approximate gradient direction at each point for all elastic parameters is crucial for accurate modeling of mode conversions. However, this assumptions appears to be consistent with prevalent geologic models.

Analysis of factorizations analogous to equation 2 for arbitrary anisotropic elastic media indicates that a one-way approach may be feasible for anisotropic elastic earth models. The

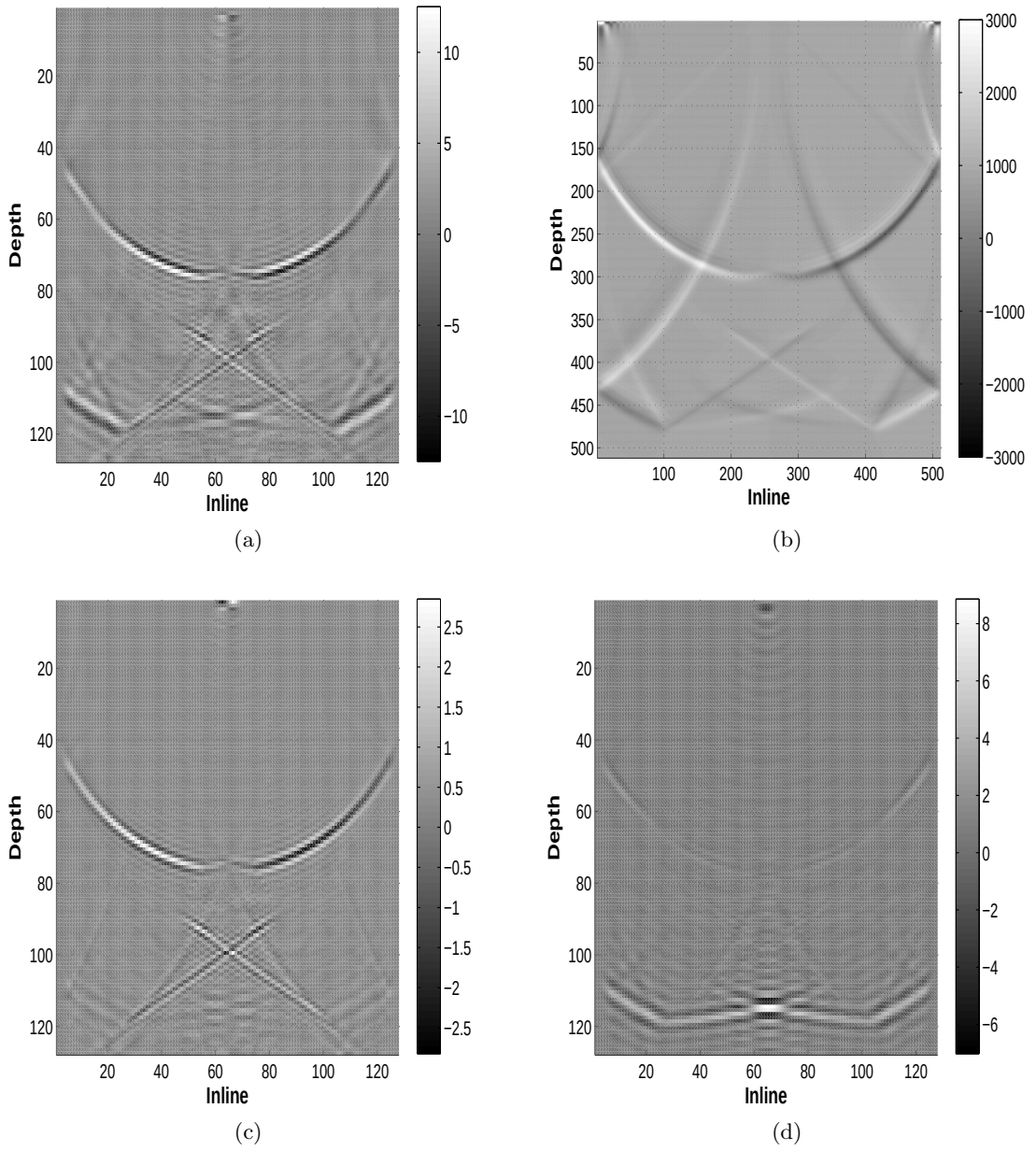


Figure 5: Impulse response for the model of Figure 4(a). Inline section of the inline displacement – (a) one-way 3D extrapolator on a $128 \times 128 \times 128$ grid and (b) 2D elastic FDTD on a 512×512 grid; (c) only shear waves; (d) only pressure waves. Note the converted shear waves shown on (a),(b) and (c). A 5-point boundary taper was used with the one-way extrapolator, and periodic boundary conditions with the FDTD. [**CR**] musa/. H1dxcy,H1Vx,H1s2cy,H1pcy

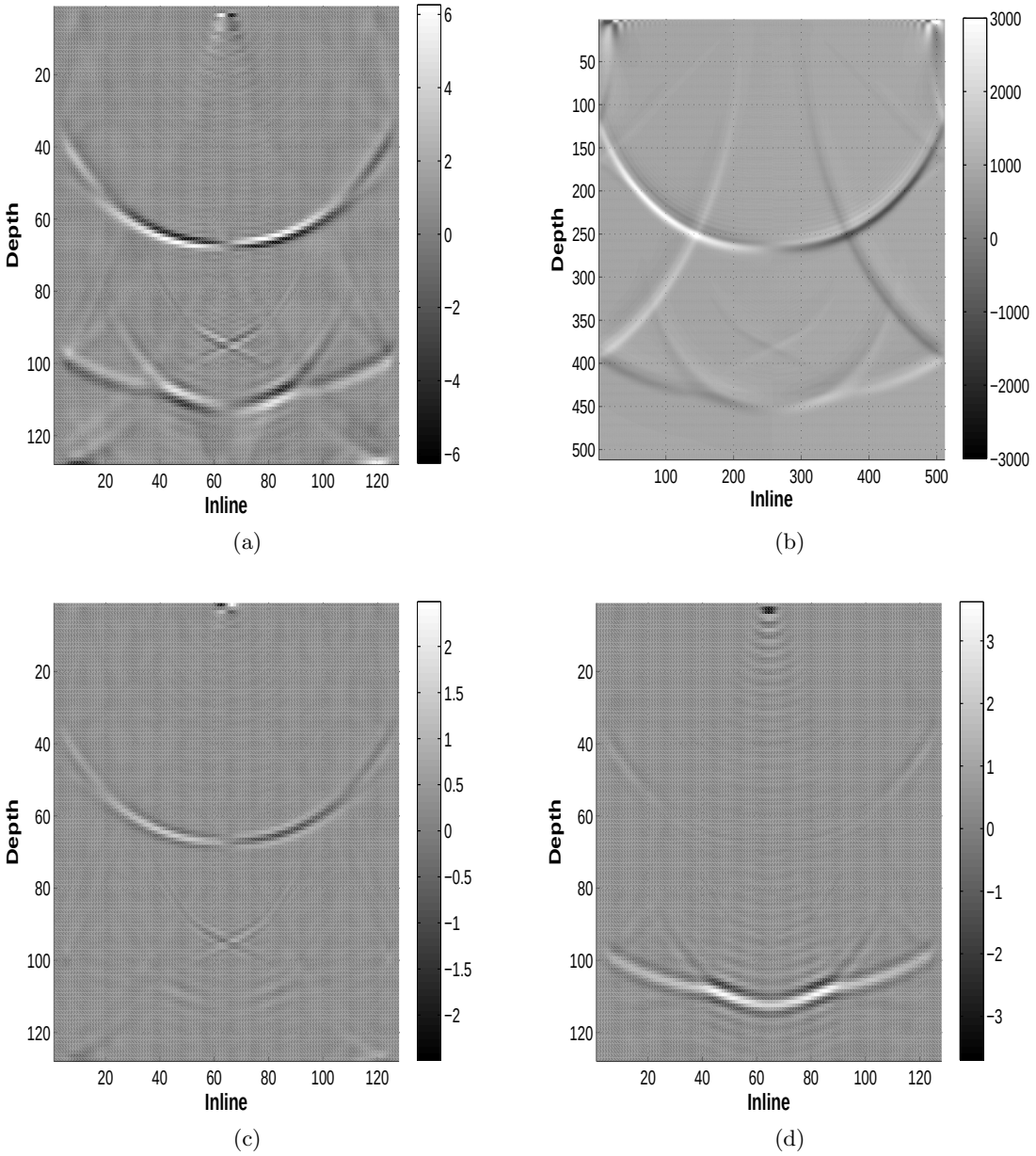


Figure 6: Impulse response for the model of Figure 4(b). Inline section of the inline displacement – (a) one-way 3D extrapolator on a $128 \times 128 \times 128$ grid and (b) 2D elastic FDTD on a 512×512 grid; (c) only shear waves; (d) only pressure waves. [**CR**] musa/. H2dxcy,H2Vx,H2s2cy,H2pcy

fact that the components of the pseudo-differential operator matrices $A(-i\partial_x, -i\partial_y)$ and $B(-i\partial_x, -i\partial_y)$ are not given explicitly, but are computed numerically, does not limit the applicability of this method.

Shear and pressure velocities (and the moduli) exhibit lithology-dependent correlation (Avseth et al., 2011) and hence are not independent. This can be used to effectively reduce the number of parameters for generating the precomputed operator tables.

ACKNOWLEDGEMENTS

The author would like to thank Biondo Biondi, Jon Claerbout, David Nichols and Stewart Levin for a number of useful discussions, and the sponsors of the Stanford Exploration Project for supporting this work.

APPENDIX: PSEUDO-DIFFERENTIAL OPERATOR FACTORIZATION

Freezing the coefficients of equation 1 and applying the Fourier transform in time and horizontal variables $x_1 = x, x_2 = y$, and substituting $\mu/(1-2\nu) = \lambda + \mu$, where λ is the Lamé coefficient (Mavko et al., 2009), we get

$$\begin{aligned}
\rho\omega^2 u^1 + \mu\left[(-k_x^2 - k_y^2)u^1 + \frac{\partial^2 u^1}{\partial z^2}\right] + (\lambda+\mu)\left[-k_x^2 u^1 - k_x k_y u^2 + ik_x\frac{\partial u^3}{\partial z}\right] &= 0,\\
\rho\omega^2 u^2 + \mu\left[(-k_x^2 - k_y^2)u^2 + \frac{\partial^2 u^2}{\partial z^2}\right] + (\lambda+\mu)\left[-k_x k_y u^1 - k_y^2 u^2 + ik_y\frac{\partial u^3}{\partial z}\right] &= 0,\\
\rho\omega^2 u^3 + \mu\left[(-k_x^2 - k_y^2)u^3 + \frac{\partial^2 u^3}{\partial z^2}\right] + (\lambda+\mu)\left[ik_x\frac{\partial u^1}{\partial z} + ik_y\frac{\partial u^2}{\partial z} + \frac{\partial^2 u^3}{\partial z^2}\right] &= 0, \qquad \text{(B-1)}
\end{aligned}$$

where k_x and k_y are horizontal wave numbers and ω is the frequency. The left-hand side of system B-1 is the result of an ordinary differential operator applied to a vector-function $\mathbf{u} = (u^1, u^2, u^3)$ and parametrized by horizontal wave numbers. In the present form equations B-1 cannot be used for computationally efficient explicit depth extrapolation of seismic waves; however, these equations can be used for modeling displacements by solving very large and sparse systems of linear equations. In (Maharramov, 2012) it was suggested that equations B-1 be factorized in such a way as to allow solving them by alternating one-way extrapolation in opposite directions. More specifically, we seek a factorization of operator equation B-1 of the form 2 where $-i\frac{\partial}{\partial x} = k_x$ and $-i\frac{\partial}{\partial y} = k_y$. Performing the multiplication in equation 2 and using equation B-1, we obtain:

$$\begin{aligned}
A(k_x,k_y)B(k_x,k_y) + c_\omega[A(k_x,k_y) + B(k_x,k_y)] &= P(k_x,k_y),\\
A(k_x,k_y)E(\lambda,\mu) + E(\lambda,\mu)B(k_x,k_y) + 2c_\omega E(\lambda,\mu) &= S(k_x,k_y), \qquad \text{(B-2)}
\end{aligned}$$

where

$$\begin{aligned}
P &= \begin{bmatrix} -\gamma k_x^2 - \mu k_y^2 & -\alpha k_x k_y & 0 \\ -\alpha k_x k_y & -\gamma k_y^2 - \mu k_x^2 & 0 \\ 0 & 0 & -\mu(k_x^2 + k_y^2) \end{bmatrix},\\
S &= \begin{bmatrix} 0 & 0 & i\alpha k_x \\ 0 & 0 & i\alpha k_y \\ i\alpha k_x & i\alpha k_y & 0 \end{bmatrix}, \qquad \text{(B-3)}
\end{aligned}$$

$\alpha = \lambda + \mu$ and $\gamma = \lambda + 2\mu$. Combining equations B-2 and B-3, we get the following system of matrix equations for the operators A and B:

$$\begin{aligned} A(k_x,k_y)B(k_x,k_y) + c_\omega(A(k_x,k_y) + B(k_x,k_y)) &= P(k_x,k_y), \\ E(\lambda,\mu)B(k_x,k_y) + A(k_x,k_y)E(\lambda,\mu) &= \tilde{S}(k_x,k_y), \end{aligned} \tag{B-4}$$

where

$$\tilde{S}(k_x,k_y) = S(k_x,k_y) - 2c_\omega E(\lambda,\mu). \tag{B-5}$$

REFERENCES

Aki, K. and P. G. Richards, 2009, Quantitative seismology, second ed.: University Science Books.

Avseth, P., T. Mukerji, and G. Mavko, 2011, Quantitative seismic interpretation: Cambridge University Press.

Biondi, B., 2005, 3D Seismic Imaging: Society of Exploration Geophysicists.

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

Gazdag, J. and P. Sguazzero, 1984, Migration of seismic data by phase shift plus interpolation: Geophysics, **49**, 124–131.

Grechka, V., 2009, Applications of seismic anisotropy in oil and gas industry: EAGE.

Higham, N., 2008, Functions of matrices: Theory and computation: SIAM.

Maharramov, M., 2012, Efficient depth extrapolation of waves in elastic isotropic media: Stanford Exploration Project Report, **148**.

Maharramov, M. and B. Nolte, 2011, Efficient one-way wave-equation migration in tilted transversally isotropic media: 73rd EAGE Conference and Exhibition, Extended Abstracts.

Mavko, G., T. Mukerji, and J. Dvorkin, 2009, The rock physics handbook: Cambridge University Press.

Nolte, B., 2008, Fourier finite-difference depth extrapolation for VTI media: 70th EAGE Conference and Exhibition, Extended Abstracts.

Segall, P., 2010, Earthquake and volcano deformation: Princeton University Press.

Shan, G., 2007, Optimized implicit finite-difference migration for TTI media: 77th SEG Conference and Exhibition, Extended Abstracts.

Ursin, B., 1983, Review of elastic and electromagnetic wave propagation in horizontally layered media: Geophysics, 1063–1081.

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

Virieux, J., S. Operto, B. H. Ali, R. Brossier, V. Etienne, F. Sourbier, L. Giraud, and A. Haidar, 2009, Seismic wave modeling for seismic imaging: The Leading Edge - Special Section: Seismic Modeling, 538–544.

Wu, R.-S., 1994, Wide-angle elastic wave one-way propagation in heterogeneous media and an elastic wave complex-screen method: Journal of Geophysical Research: Solid Earth, **99**, 751–766.

Xie, X. and R. Wu, 2005, Multicomponent prestack depth migration using the elastic screen method: Geophysics, **70**, S30–S37.

Practically stable unstable orthorhombic finite differences

Huy Le and Stewart A. Levin

ABSTRACT

Intrigued by an instability result presented at the last SEG meeting (Chu, 2012), we analyze it and some variants to understand the nature and extent of such instabilities. We investigate an unexpected dependency of stability on the order of our spatial derivative approximation. We find that even the apparently stable scheme in that SEG abstract is actually slightly unstable, though the instability is not qualitatively manifest until tens of thousands of time steps are taken.

INTRODUCTION

The Stanford Exploration Project, in line with the interests of many of its sponsors, has been investigating the effects of anisotropy on seismic processing, imaging and inversion. Much recent interest in the geophysical community has focused on the effects of orthorhombic symmetry, a type of elastic symmetry that has three mutually orthogonal planes of symmetry. This type of symmetry can be found, most commonly, where a set of vertical fractures is embeded in a VTI medium, or where two identical sets of fractures make an arbitrary angle to each other, or where two or three sets of mutually perpendicular fractures intersect (Tsvankin, 1997). Orthorhombic crystals occur in anhydrite, olivine, and sulfur.

An orthorhombic medium is characterized by nine independent elastic parameters, c_{ij}, that enter into a linear Christoffel relationship between stresses, σ_{ij}, and strains, ϵ_{ij} (Malvern, 1969):

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{31} \\ \sigma_{12} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & & & \\ c_{12} & c_{22} & c_{23} & & & \\ c_{13} & c_{23} & c_{33} & & & \\ & & & c_{44} & & \\ & & & & c_{55} & \\ & & & & & c_{66} \end{bmatrix} \begin{bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ 2\epsilon_{23} \\ 2\epsilon_{31} \\ 2\epsilon_{12} \end{bmatrix}. \tag{1}$$

There are two common ways to derive equations for wave propagation in anisotropic media. The first approach is to derive them from the dispersion relations obtained from the general Christoffel equations (Alkhalifah, 2003). The resultant differential equations are sixth-order in both time and space. Another approach is to start with constitutive stress-strain relations and equations of motion (Duveneck and Bakker, 2011; Zhang and Zhang, 2011). The resultant equations from this approach are second order. In either case, these are systems that couple both normal and shear stresses.

In practice, we are currently most interested in analysis and imaging of compressional waves. For this case, one approximation is to set the shear velocities in the symmetry

plane axes to zero ($c_{44} = c_{55} = c_{66} = 0$), resulting in what are called *pseudo-acoustic* wave equations. This substitution reduces the medium description to six parameters instead of nine. Following the second approach, in tilted orthorhombic symmetry, the system of pseudo-acoustic wave equations takes the following form:

$$\frac{\partial^2 \sigma_{ii}}{\partial t^2} = \sum_j G_{ij} \sum_m \sum_n R_{jm} R_{jn} \sum_l \frac{\partial^2}{\partial x_n \partial x_l} \sum_k R_{km} R_{kl} \sigma_{kk}, \tag{2}$$

where σ_{ii} are three normal stresses, G_{ij} are related to vertical P-wave velocity v_{pz} and the Thomsen anisotropic parameters ϵ_i and δ_i in symmetry planes, and the R_{ij} are entries of the matrix that transforms from the coordinate system aligned with the symmetry of the material to the surface North-East South-West Vertical system (Chu, 2012).

We reparameterize system 2 as:

$$\frac{\partial^2 \sigma_{ii}}{\partial t^2} = \sum_{n,l,k} a_{inlk} \frac{\partial^2}{\partial x_n \partial x_l} \sigma_{kk}, \tag{3}$$

where

$$a_{inlk} = \sum_{j,m} G_{ij} R_{jm} R_{jn} R_{km} R_{kl} \tag{4}$$

are the corresponding coefficients in front of the spatial derivatives. These a_{inlk} coefficients have units of velocity squared. This reparameterization simplifies later algebraic manipulations and analysis and reduces the computational time significantly.

Relation 3 defines a system of three second-order coupled equations for the three normal stresses σ_{ii}. The compressional pressure P-wave is defined as the negative of the average of these stresses, $p = -\frac{1}{3}\sum_i \sigma_{ii}$ (Malvern, 1969). Each of the equations in 3 has 18 terms on the right-hand side: nine non-mixed spatial derivatives and nine mixed spatial derivatives. The mixed derivatives consume a major fraction of the computational time as compared to the non-mixed ones and will be handled differently by different numerical schemes.

NUMERICAL STABILITY ANALYSIS

We analyze the stability of three different numerical schemes, following Chu (2012). All of them use centered finite differences to approximate the non-mixed second spatial derivatives. They differ in how the mixed derivatives are calculated (Table 1).

	Non-mixed derivatives	Mixed derivatives
Scheme 1	Centered	Centered
Scheme 2	Centered	Rotated non-staggered
Scheme 3	Centered	Rotated staggered

Table 1: Numerical schemes we analyze for stability.

To analyze the stability of system 3, we use the von Neumann method (Saenger et al., 2000; Richtmyer and Morton, 1957). Rewriting system 3 in matrix form we have

$$\frac{\partial^2 \boldsymbol{\sigma}}{\partial t^2} = \mathbf{L}\boldsymbol{\sigma}, \tag{5}$$

where $\boldsymbol{\sigma} = [\sigma_{ii}]$ is a vector of three normal stresses, and $\mathbf{L}$ is a 3×3 matrix differential operator.

The numerical approximation of equation 5 in the spatial Fourier domain takes the following form:

$$\hat{\boldsymbol{\sigma}}_{n+1} = (\triangle t^2 \hat{\mathbf{L}} + 2\mathbf{I})\hat{\boldsymbol{\sigma}}_n - \hat{\boldsymbol{\sigma}}_{n-1}, \tag{6}$$

where we have used second-order finite difference in time with step size $\triangle t$. The hat notation represents a function in the spatial Fourier domain, and the subscript denotes time level.

Equation 6 is a three-level recursive relation between the future ($t = n + 1$), current ($t = n$), and past ($t = n-1$) wavefields. It can be reduced to a two-level system of equations as follows:

$$\begin{bmatrix} \hat{\boldsymbol{\sigma}}_{n+1} \\ \hat{\boldsymbol{\sigma}}_n \end{bmatrix} = \begin{bmatrix} \triangle t^2 \hat{\mathbf{L}} + 2\mathbf{I} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\sigma}}_n \\ \hat{\boldsymbol{\sigma}}_{n-1} \end{bmatrix}, \tag{7}$$

or

$$\mathbf{u}_{n+1} = \mathbf{A}\mathbf{u}_n, \tag{8}$$

where

$$\mathbf{u}_n = \begin{bmatrix} \hat{\boldsymbol{\sigma}}_n \\ \hat{\boldsymbol{\sigma}}_{n-1} \end{bmatrix} \tag{9}$$

is a six-component vector of three current and three previous wavefields and

$$\mathbf{A} = \begin{bmatrix} \triangle t^2 \hat{\mathbf{L}} + 2\mathbf{I} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \tag{10}$$

is the 6×6 *amplification* matrix.

The amplification matrix is a function of wave numbers k_x, k_y and k_z and determines how much each mode (in spatial frequency) is amplified as one time step advances. It depends on the medium parameters a_{inlk} (equation 4), the numerical scheme, the order of spatial approximation $M/2$, and the time and space step sizes (Appendix A). The numerical scheme is stable when the magnitude of the largest eigenvalue of $\mathbf{A}$ is less than unity for all modes.

In the case of second-order spatial approximation, an exact formula for the eigenvalues of $\mathbf{A}$ might be found (Saenger et al., 2000). In our case, as shown in Appendix B, finding the explicit form for the eigenvalues of $\mathbf{A}$ involves solving a cubic equation. In our work, we numerically compute the eigenvalues for different symmetries (Table 2) and analyze how the eigenvalues of $\mathbf{A}$ vary with changing time step and order of spatial approximation. We choose $\triangle x = \triangle y = 12.5$ m and $\triangle z = 6.25$ m.

Symemtry	ϵ_1	ϵ_2	δ_1	δ_2	δ_3	v_{pz} (m/s)	θ	φ	ϕ
Isotropic	0	0	0	0	0	2000	0	0	0
TI	0.16	0.16	0.06	0.06	0	2000	0	0	0
Orthorhombic	0.2	0.12	0.07	0.05	0	2000	0	0	0
Tilted Orthorhombic	0.2	0.12	0.07	0.05	0	2000	40	25	35

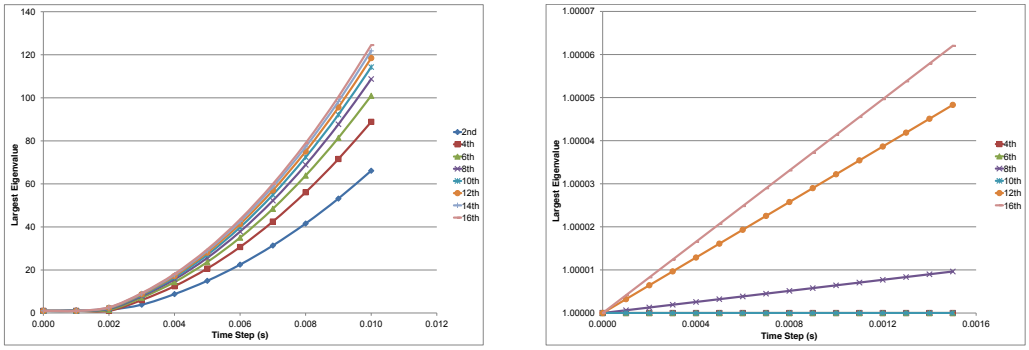
Table 2: Medium parameters we use for stability analysis (Chu, 2012).

RESULTS AND DISCUSSION

For tilted orthorhombic media, from Figures 1(a), 2(a), and 3(a), we observe that the magnitude of the largest eigenvalue increases exponentially with increasing time step $\triangle t$ and, as expected, approaches unity as $\triangle t$ approaches zero. Additionally, higher orders of spatial approximation generally come with higher degrees of instability. This implies a trade-off between accuracy and stability. The largest eigenvalues for schemes 1 and 3 have the same order of magnitude, while those of scheme 2 are considerably greater (Figures 1(b), 2(b), and 3(b)). This illustrates that, as Chu points out, using rotated non-staggered finite differences for approximating mixed spatial second derivatives causes a more severe problem of instability. Although centered and rotated staggered grids have a less serious instability issue, higher order spatial approximations (orders 8th, 10th, 12th, 14th, and 16th) are not technically stable, because their corresponding maximum eigenvalues are slightly greater than unity (Figures 2(b), and 3(b)). However, for typical time steps of a millisecond or less, amplification of the wavefield will not be significant until after tens of thousands of time steps.

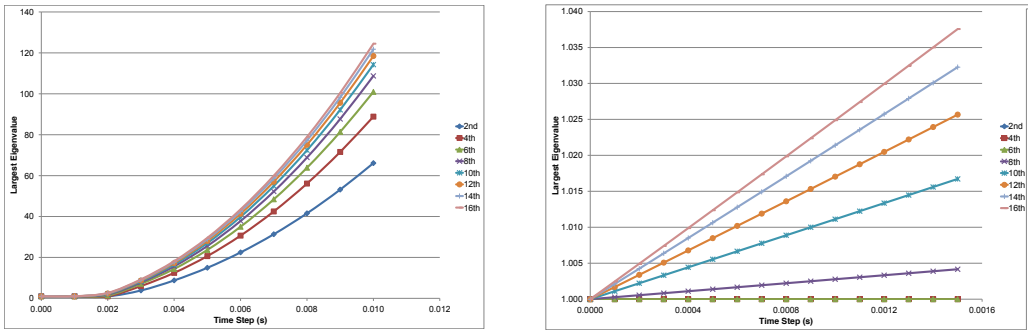
Figures 4 and 5 show that the issue of instability becomes much more problematic for tilted orthorhombic media than for other types of symmetries. Also, from these figures, with our particular sets of material parameters, these numerical schemes applied to TI symmetry surprisingly appear to become unstable more quickly than those applied to vertical orthorhombic symmetry. We do not believe this is actually the case, but an artifact of numerical precision when eigenvalues are clustered, as they most certainly are for TI media which has a plane of symmetry. This still needs to be investigated.

One interpretation of the CFL stability condition is that the wavefield should not propagate past multiple grid points in one time step. Consequently, one would intuitively expect that the fast-progagating waves are more likely to cause instability. Our observation that the directions that have the largest eigenvalues do not coincide with the directions of the fast-propagating waves (Figure 6), however, does not fully support this intuition.



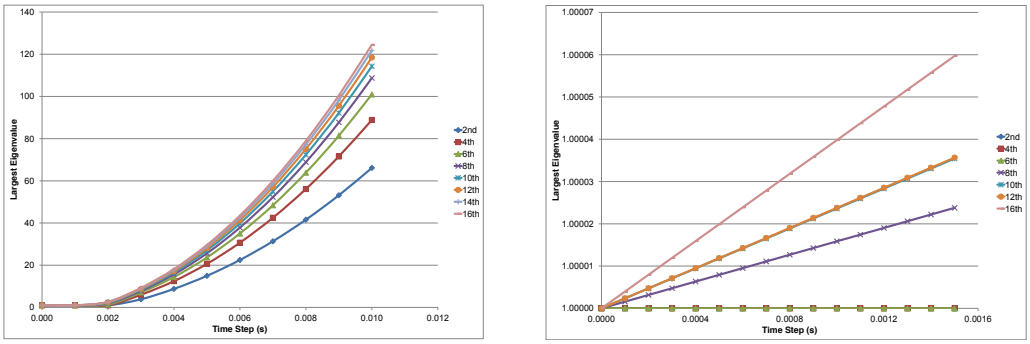
(a) (b)

Figure 1: Largest eigenvalues of different orders for tilted orthorhombic media using scheme 1: (a) large time steps and (b) small time steps. [**NR**] huyle/. Centered.tilted,Centered.tilted.zoom



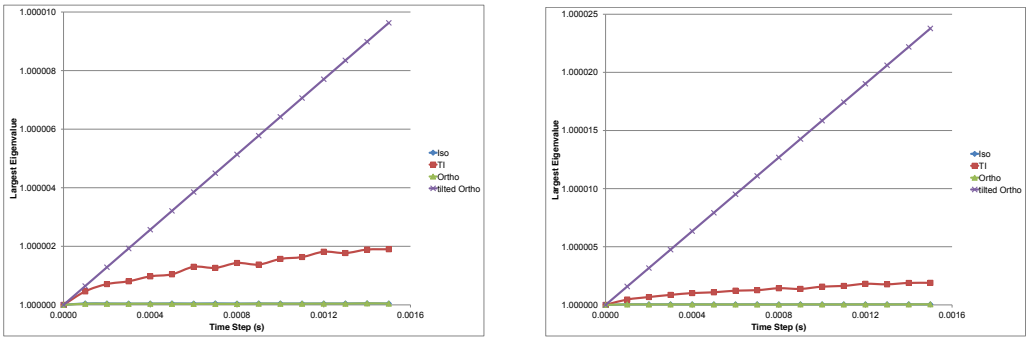
(a) (b)

Figure 2: Largest eigenvalues of different orders for tilted orthorhombic media using scheme 2: (a) large time steps and (b) small time steps. With such eigenvalues, the wavefield would be amplified greatly after hundreds of time steps. [**NR**] huyle/. Rotated.tilted,Rotated.tilted.zoom



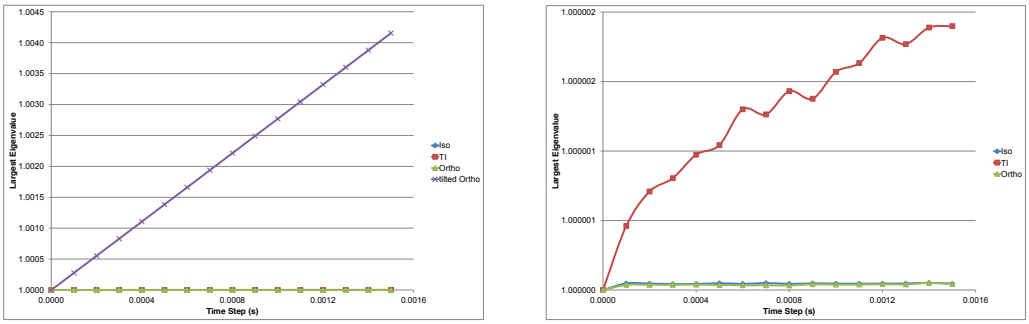
(a) (b)

Figure 3: Largest eigenvalues of different orders for tilted orthorhombic media using scheme 3: (a) large time steps and (b) small time steps. [**NR**] huyle/. Staggered.tilted,Staggered.tilted.zoom



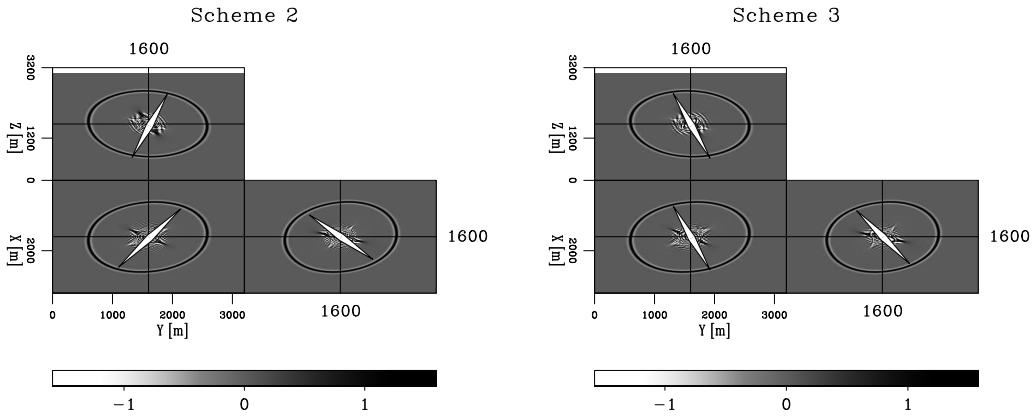
(a) (b)

Figure 4: Largest eigenvalues of 8th order for tilted orthorhombic media using: (a) scheme 1 and (b) scheme 3. [**NR**] huyle/. Centered.8th,Staggered.8th



(a) (b)

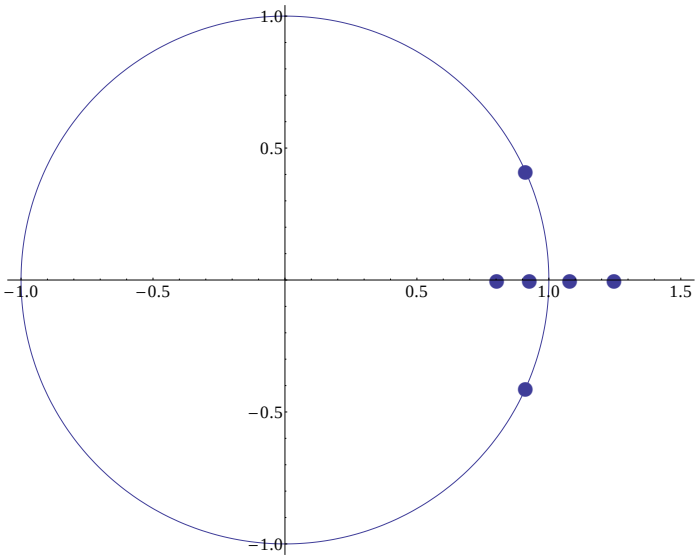
Figure 5: Largest eigenvalues of 8th order for tilted orthorhombic media using scheme 2: (a) 4 types of symmetries and (b) zoom-in without tilted orthorhombic. [**NR**] huyle/. Rotated.8th,Rotated.8th.zoom



(a) (b)

Figure 6: Wavefield snapshots at $t = 0.5$ s using $\triangle t = 0.1$ ms: (a) scheme 2 and (b) scheme 3. The arrows mark directions that have the largest eigenvalues. These directions do not coincide with the directions of the fast-propagating waves. [**CR**] huyle/. r,s

Figure 7: Positions of six eigenvalues of the amplification matrix **A** (for scheme 1, second order spatial approximation, tilted orthorhombic symmetry) with respect to the unit circle. The figure shows three pairs of eigenvalues: one complex-conjugate pair of magnitude unity and two real pairs, each of which has one eigenvalue that is greater than unity. [**NR**] huyle/. circle



CONCLUSION

We verified that using rotated non-staggered finite difference to approximate the mixed derivatives in wave equations for tilted orthorhombic media can cause serious instability. Alternative schemes, such as centered and rotated staggered finite differences are practically stable, that is, with time steps around or less than a millisecond, the wavefield can be propagated up to tens of thousands of steps without being excessively amplified.

ACKNOWLEDGMENTS

I would like to thank colleagues and professors at the Stanford Exploration Project for many of their valuable advices and discussions.

APPENDIX A

The amplification matrices are as follows:

For scheme 1:

$$\begin{aligned}\hat{L}_{ik} &= \sum_n a_{innk}\frac{2}{\triangle x_n^2}\sum_{p=1}^{M/2} c_p[cos(k_{x_n}p\triangle x_n)-1] \\ -\sum_{n\neq l}(a_{inlk}+a_{ilnk})&\frac{4}{\triangle x_n\triangle x_l}\sum_{p=1}^{M/2} b_p sin(k_{x_n}p\triangle x_n)\sum_{p=1}^{M/2} b_p sin(k_{x_l}p\triangle x_l).\end{aligned} \tag{C-1}$$

For scheme 2:

$$\begin{aligned}\hat{L}_{ik} &= \sum_n a_{innk}\frac{2}{\triangle x_n^2}\sum_{p=1}^{M/2} c_p[cos(k_{x_n}p\triangle x_n)-1] \\ -\sum_{n\neq l}(a_{inlk}+a_{ilnk})&\frac{1}{\triangle x_n\triangle x_l}\sum_{p=1}^{M/2} c_p sin(k_{x_n}p\triangle x_n)sin(k_{x_l}p\triangle x_l).\end{aligned} \tag{C-2}$$

For scheme 3:

$$\begin{aligned}\hat{L}_{ik} &= \sum_n a_{innk}\frac{2}{\triangle x_n^2}\sum_{p=1}^{M/2} c_p[cos(k_{x_n}p\triangle x_n)-1] \\ -\sum_{n\neq l\neq r}(a_{inlk}+a_{ilnk})&\frac{4}{\triangle x_n\triangle x_l}\sum_{p=1}^{M/2} s_p sin(k_{x_n}\frac{2p-1}{2}\triangle x_n)cos(k_{x_l}\frac{2p-1}{2}\triangle x_l)cos(k_{x_r}\frac{2p-1}{2}\triangle x_r) \\ &\times\sum_{p=1}^{M/2} s_p cos(k_{x_n}\frac{2p-1}{2}\triangle x_n)sin(k_{x_l}\frac{2p-1}{2}\triangle x_l)cos(k_{x_r}\frac{2p-1}{2}\triangle x_r).\end{aligned} \tag{C-3}$$

The c_p, b_p, and s_p are coefficients of the spatial derivative approximations as defined by Chu(2012).

APPENDIX B

Stability is assessed by examing the roots of the determinant of the eigenvalue matrix:

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} \triangle t^2\hat{\mathbf{L}} + 2\mathbf{I} - \lambda\mathbf{I} & -\mathbf{I} \\ \mathbf{I} & -\lambda\mathbf{I} \end{bmatrix} \tag{D-1}$$

Due to the commutativity of the identity matrix with all other matrices, the determinant of this block matrix is the same as the determinant of

$$(\triangle t^2\hat{\mathbf{L}} + 2\mathbf{I} - \lambda\mathbf{I}) \cdot (-\lambda\mathbf{I}) - \mathbf{I} \cdot (-\mathbf{I}). \tag{D-2}$$

Simplifying, this becomes

$$\det A = \det((\lambda^2 + 1)I - \lambda(\triangle t^2\hat{\mathbf{L}} + 2\mathbf{I})) = (-\lambda)^3 \det((\triangle t^2\hat{\mathbf{L}} + 2\mathbf{I}) - (\lambda + 1/\lambda)I). \tag{D-3}$$

This takes the form of the eigenvalue calculation for $\triangle t^2\hat{\mathbf{L}} + 2\mathbf{I}$, with each eigenvalue $\hat{\lambda}$ providing a pair of eigenvalues of $\mathbf{A}$ according to $\hat{\lambda} = \lambda + \frac{1}{\lambda}$. Solving by the quadratic formula yields

$$\lambda = \frac{\hat{\lambda} \pm \sqrt{\hat{\lambda}^2 - 4}}{2}. \tag{D-4}$$

If we assume $\hat{\lambda}$ is real valued, then we first consider the case $|\hat{\lambda}| > 2$. In this event the two roots are real and their product is 1; hence at least one is greater than 1 in magnitude. This situation is therefore unstable. If the magnitude is exactly 2, the roots are either both 1 or both -1, a stable situation. Finally, with the magnitude less than 2, the roots are complex and the squared magnitude of each is

$$|\lambda|^2 = \frac{\hat{\lambda}^2 - (\hat{\lambda}^2 - 4)}{4} = 1; \tag{D-5}$$

in other words, both roots are on the unit circle. These are, indeed, the very situations we observed in our analysis (Figure 7).

REFERENCES

Alkhalifah, T., 2003, An acoustic wave equation for orthorhombic anisotropy: Geophysics, **68**, 1169–1172.

Chu, C., 2012, A hybrid finite difference method for acoustic wave propagation in tilted orthorhombic media: 82nd SEG Ann. Internat. Meeting, Expanded Abstracts, 1–5, Soc. of Expl. Geophys.

Duveneck, E. and P. M. Bakker, 2011, Stable P-wave modeling for reverse-time migration in tilted TI media: Geophysics, **76**, S65–S75.

Malvern, L. E., 1969, Introduction to the mechanics of a continuous medium: Prentice-Hall.

Richtmyer, R. D. and K. W. Morton, 1957, Difference methods for initial-value problems: Interscience Publishers.

Saenger, E. H., N. Gold, and S. A. Shapiro, 2000, Modeling the propagation of elastic waves using a modified finite-difference grid: Wave Motion, **31**, 77–92.

Tsvankin, I., 1997, Anisotropic parameters and P-wave velocity for orthorhombic media: Geophysics, **62**, 1292–1309.

Zhang, H. J. and Y. Zhang, 2011, Reverse-time migration in vertical and tilted orthorhombic media: 81st SEG Ann. Internat. Meeting, Expanded Abstracts, 185–189, Soc. of Expl. Geophys.

Applications for rotational seismic data

Ohad Barak and Shuki Ronen

ABSTRACT

Seismic systems today record up to four components which provide the particle displacement and the pressure. The pressure is proportional to the divergence of the displacement. The curl of the displacements can be recorded using rotation sensors.
To evaluate the added information that would come from rotation sensors we use elastic modeling. In our synthetic data experiment, we predict the effect of a seabed scatterer on seven-component OBS data: three component geophones, three component rotation sensors and hydrophones.
The synthetic data comprises P-waves, S-waves and surface waves. We apply singular value decomposition in order to identify the polarization vectors for each wave type. Our evaluation is that the added information from rotation sensors is useful for identifying and separating surface waves from body waves.
Additionally, we use elastic modeling to predict how a change in rock-physics parameters affects the AVO curve of the rotational-motion components, and compare the response to a standard AVO obtained from the hydrophone component.

INTRODUCTION

Four-component ocean-bottom seismic sensors are a combination of a hydrophone and three-component geophones that measure linear particle velocity. The hydrophone records pressure changes in the water, and is ideally coupled to the water. It picks up mostly P-waves. The geophones are coupled to the sea-bed, and record particle velocities related to all wave modes: P, S, and surface waves. The multiplicity of wave modes recorded by the geophones can be problematic for later seismic processing stages, which assume a single wave mode recording (i.e., imaging with P-wave or S-wave data only). It is therefore desirable to be able to separate different wave modes within the data.

The divergence of particle motion is a spatial derivative. If we knew the displacement values everywhere within the seismic volume, then we could calculate the divergence. However, receiver stations may be too sparse in relation to the seismic wavelengths to calculate the divergence, and, in any case, are always spread along a single surface (the sea bottom). The value of the divergence is not redundant even if we measure the three components of the displacement, since it can be used to detect the propagation direction of the waves incident on the sea bottom (upward or downward). The pressure that the hydrophone records is proportional to the divergence of the particle displacements, and is therefore used for upgoing/downgoing wavefield separation.

Similar to the divergence, we can measure the rotation of particle motion by calculating the curl of the displacements. As with the divergence, curl is a spatial derivative operator, and insufficient sampling of the waves in the field can inhibit its direct calculation. However,

a recently-proposed rotation sensor can be used to measure the rotation "in place". A rotation sensor measures the rotation rate (radians/unit time) of the ground at the receiver location. Similar to linear motion, the rotational motion has three components: roll, pitch and yaw. Figure 1 illustrates these motion components.

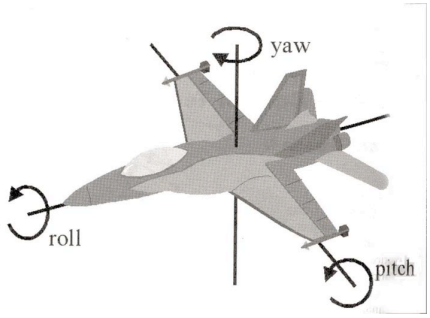


Figure 1: To know what an aircraft is doing, it is not sufficient to know its linear velocity but also the rotations. There are three components to the velocity, and in addition there are three components to the rotation: roll, pitch, and yaw. [**NR**] ohad1/. f18

There are implementations of rotational seismic recording for earthquake seismology (Lee et al., 2009) employing ring lasers. However, these instruments are large and expensive, and are therefore not applicable to exploration acquisition. There have also been attempts to record rotational seismic data with a "Rotaphone" - conventional geophones arranged along a circle (Brokesova and Malek, 2010). Newer technologies are currently under development to make rotational sensors a viable option for seismic acquisition systems. The USGS has recently tested a small number of rotation sensors at the Calpine geothermal power station near The Geysers in California. Chevron has conducted a small land survey that included rotational sensors, in order to evaluate the quality and the possible uses for such data.

Several authors have proposed applications for rotational seismic data (Brune et al., 2012; Muyzert et al., 2012; Edme et al., 2012). In this paper we propose two applications:

1. using rotational data for identifying and filtering surface waves, and
2. constructing converted-wave AVO curves, where the rotation data is a proxy for the reflected shear-wave energy.

Linear versus rotational polarization

deMeersman et al. (2006) apply SVD on three-component geophone data in order to estimate the polarization direction of P-wave arrivals. Similarly, it is possible to apply SVD on 6-component data comprising a 3C geophone and a 3C rotation sensor. The resulting polarization vector indicates not only linear polarization, but also rotational polarization. The combination of linear and rotational polarization components provide an indication of the wave type. Body waves generate linear particle motion, and have a mostly linear polarization. Surface waves (Rayleigh, Scholte, Love) generate elliptical particle motion, and therefore cause a greater rotational deformation on a finite volume element. Therefore, they should be polarized more along a rotational axis than a linear axis.

We synthesized 3D ocean-bottom seismic data. To simulate seven-component acquisition, we recorded the pressure in the water just above the ocean-bottom, and the particle velocity and rotations at the water bottom. The data contains both body waves and surface

waves. We then applied SVD to the seven components of specific arrivals, which we identified apriori as P, SV, SH, or Scholte waves. The scaled, orthogonal polarization vectors of these arrivals show how the waves are polarized in a seven-dimensional space comprising one pressure, three linear-motion and three rotational-motion axes.

Rotational AVO for gas hydrates

Gas hydrates are a form of water crystal that have a specific structure, and which are stabilized by the inclusion of methane gas molecules. The presence of a gas hydrate saturated layer can be seen in some seismic data. The high impedance contrast between the gas hydrate layer and the underlying sediments creates a strong seismic signature known as a bottom simulating reflector (BSR).

Rock-physics models of gas hydrates, which connect density, porosity, effective pressure, mineralogy and pore-space saturation in marine sediments containing the hydrates to elastic properties have been developed (Ecker et al., 1997; Dvorkin and Nur, 1993). The micromechanical structure of the hydrate and the porosity of the sediment in which it forms will affect the level of hydrate saturation, and therefore the total amount of methane gas within those sediments.

Several rock physics models exist to describe the way the gas hydrate is included in the sediments. But the determination of which model is the best representation dependends on accurate acquisition of several material parameters: mineralogy, density, porosity and saturation. These properties can be measured using well logs, and can be compared to data acquired by seismic surveys. The information in P-wave reflections can be used to determine some of these parameters. However, the P-wave reflectivity is not very sensitive to the shear strength of the medium, which is the one parameter which very clearly separates gas hydrates from their typical sedimentary surrounding. The wavefield curl is a proxy for the shear waves amplitude, and recording it can help in getting a better estimation of the host rock properties and the internal structure of hydrate sediments.

Utilizing the rock-physics models and data from a well logs near Blake Outer Ridge, where a hydrate layer is present, we constructed a 1.5D medium with effective elastic properties. We used elastic modeling to predict how a change in rock physics parameters would affect the AVO curve of the rotational-motion components recorded by rotation sensors, and compared the response to AVO obtained from the hydrophone component. We do not actually plot the AVO curve, but we do show a CMP gather obtained from a synthetic OBS experiment in a medium with a single flat hydrate layer, and therefore the amplitudes observed over offset are comparable to an AVO.

THEORY

Elastic wave propagation

The isotropic elastic wave equation has two state variables: the stress tensor and the particle velocity vector. The particle velocities are propagated by solving:

$$\partial_i \sigma_{ii} + \partial_j \sigma_{ij} + f_i(\mathbf{x}, t) = \rho \partial_t v_i, \tag{1}$$

where σ_{ii} are the normal stresses, σ_{ij} are the transverse stresses, f_i is a particle velocity force function in direction i, $\mathbf{x}$ is the spatial source location operating at time t, ρ is density and v_i is the particle velocity in direction i. The stresses are propagated using the stress-displacement relation:

$$\begin{aligned} \partial_t \sigma_{ii} &= (\lambda + 2\mu)\partial_i v_i + \lambda \partial_j v_j + f^P(\mathbf{x}, t), \\ \partial_t \sigma_{ij} &= \mu \left(\partial_j v_i + \partial_i v_j \right), \end{aligned} \tag{2}$$

where λ and μ are the Lamé elastic constants and f^P is a pressure force function. The pressure force is added equally to the normal stresses to generate a P-wave source.

We use the staggered time grid methodology for elastic propagation (Virieux, 1986), in which the stresses and particle velocities are half a time-step apart. Equations 1 and 2 are solved in alternation during the propagation.

The divergence of particle displacements is proportional to the pressure in the medium, the proportion being the medium parameters. It is equal to the average of the normal stresses in the stress tensor:

$$P = \overline{\boldsymbol{\sigma}} = \frac{2}{3} (\lambda + \mu) \nabla \cdot \vec{\mathbf{u}}, \tag{3}$$

where P is the pressure value and $\vec{\mathbf{u}}$ are the displacements. Pressure waves cause a volumetric deformation in the medium, and their value can therefore be extracted by using equation 3 on forward modeled wavefields. However, other wave types can also generate a volumetric deformation at free surfaces, as a result of the discontinuity of the stresses.

We define rotation as the first time derivative of the curl of displacements:

$$\vec{\mathbf{R}} = \partial_t \nabla \times \vec{\mathbf{u}}. \tag{4}$$

The curl operation results in the non-volumetric part of the deformation, i.e. the "shear" deformation. At a free surface, this deformation will cause a rotation. In an isotropic medium, the curl is associated with S-waves. However, at a free surface, both P-waves and surface waves will also generate non-volumetric deformation. The units we use for rotations in this paper are milliradians/second, while the units for particle velocity are millimeters/second.

Singular value decomposition of seven-component data

Given a seven-component seismic dataset with a seismogram consisting of pressure $h(t)$, particle velocity $v_z(t)$, $v_x(t)$ and $v_y(t)$ recorded in the Z, X and Y directions, and rotation $r_z(t)$, $r_x(t)$ and $r_y(t)$ recorded in the Z, X and Y directions, we can constuct an Nx7 data matrix $\mathbf{D} = [h(t), v_z(t), v_x(t), v_y(t), r_z(t), r_x(t), r_y(t)]$. Each column is a recording of one component over a certain time window. SVD is a method of finding the waveform $\mathbf{u}_s$,

magnitude σ_s, and polarization $\mathbf{v}_s$ of the signal that is present in the data $\mathbf{D}$. The SVD of the data $\mathbf{D}$ is given by

$$\mathbf{D} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T, \tag{5}$$

where $\mathbf{D}$ is the product of the Nx7 matrix $\mathbf{U}$, the 7x7 diagonal matrix $\mathbf{\Lambda}$, and the transpose of the 7x7 matrix $\mathbf{V}$. The unit left and right singular vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ are the seven column vectors of $\mathbf{U}$ and $\mathbf{V}$. The singular values σ_i are the diagonal elements of $\mathbf{\Lambda}$. They are ordered such that $|\lambda_1|$ is the greatest and $|\lambda_7|$ the smallest.

The left and right singular vectors are mutually orthogonal, such that $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ and $\mathbf{V}\mathbf{V}^T = \mathbf{I}$. An SVD therefore decomposes the data matrix into its orthogonal parts.

The right singular vectors $\mathbf{v}_i$ display the polarization of the data within a particular time window along the seven axes. We transpose and multiply the matrix $\mathbf{V}$ by the singular value matrix $\mathbf{\Lambda}$, to obtain the scaled polarization vectors:

$$\mathbf{s}_i = \sigma_i v_i^T. \tag{6}$$

Hydrate inclusion models

Ecker et al. (1997) developed three different rock physics models for the inclusion of gas hydrates in porous sedimentary rock, and tested these models against well-log data and seismic data from the Blake Outer Ridge. These models are:

1. Hydrate is a part of the fluid in the pore-space
2. Hydrate is a part of the solid host rock
3. Hydrate cements the grains of the host rock

The models are illustrated in Figure 2. They pertain to loosely consolidated sedimentary rock, which is the case for the Blake Outer Ridge, near the passive continental margin offshore Eastern Florida. Each model produces a different estimate of the elastic moduli of the hydrate-containing sediment, as a function of porosity and saturation of the rock. We will deal only with the first and second inclusion models.

Model 1 - Hydrate is part of fluid

When the hydrate is assumed to be part of the fluid, the fluid's bulk modulus and the solid's dry moduli are calculated separately, and then Gassmann's fluid substitution is used to saturate the rock. The fluid's modulus is an average of the water and the hydrate moduli:

$$K_f \;=\; [S_w \,/\, K_w \;+\; (1 - S_w) \,/\, K_h]^{-1}, \tag{7}$$

where K_h is the bulk modulus of the hydrate, K_w the bulk modulus of water and S_w the water saturation. The solid's dry moduli are calculated using the modified Hashin-Shtrikman-Hertz-Mindlin theory (Dvorkin and Nur (1996)). For porosity below critical porosity ($\phi_c = 0.4$) the moduli are:

$$\begin{aligned} K_{dry} &= \left[\frac{\phi/\phi_c}{K_{HM} + \frac{4}{3} G_{HM}} + \frac{1 - \phi/\phi_c}{K + \frac{4}{3} G_{HM}}\right]^{-1} - \frac{4}{3} G_{HM}, \\ G_{dry} &= \left[\frac{\phi/\phi_c}{G_{HM} + Z} + \frac{1 - \phi/\phi_c}{G + Z}\right]^{-1} - Z, \\ Z &= \frac{G_{HM}}{6}\left(\frac{9\,K_{HM} + 8\,G_{HM}}{K_{HM} + 2\,G_{HM}}\right), \end{aligned} \tag{8}$$

where K_{HM} and G_{HM} are the Hertz-Mindlin effective moduli, ϕ is porosity and ϕ_c is the critical porosity.
The saturated moduli using Gassman's fluid substitution are:

$$\begin{aligned} K_{sat} &= K\,\frac{\phi\,K_{dry} - (1+\phi)\,K_f\,K_{dry}/K + K_f}{(1-\phi)\,K_f + \phi\,K - K_f\,K_{dry}/K}, \text{and} \\ G_{sat} &= G_{dry}. \end{aligned} \tag{9}$$

The elastic velocities V_p and V_s and the bulk density ρ_B are then given by:

$$\begin{aligned} \rho_B &= (1-\phi)\,\rho_s + \phi\,\rho_f, \\ V_p &= \sqrt{(K_{sat} + \frac{4}{3} G_{sat})/\rho_B}, \\ V_s &= \sqrt{G_{sat}/\rho_B}, \end{aligned} \tag{10}$$

where ρ_s is the bulk density of the solid phase and ρ_f the density of the pore fluid.

Model 2 - Hydrate is part of the solid

In this case, we assume that hydrate becomes part of the solid sediment frame. This has two effects: porosity reduction and a change in the elastic frame moduli. The reduced porosity ϕ_r can be expressed as:

$$\phi_r = \phi\,(1 - S_h), \tag{11}$$

where S_h is the hydrate saturation of the pore space.
The bulk and shear moduli of the solid phase are now a mixture of the sediment solid and the hydrate. The Hill average is proposed as a way to estimate the effective moduli:

$$\begin{aligned} K &= \frac{1}{2}(f_h\,K_h + (1-f_h)\,K_s + [f_h/K_h + (1-f_h)/K_s]^{-1}), \\ G &= \frac{1}{2}(f_h\,G_h + (1-f_h)\,G_s + [f_h/G_h + (1-f_h)/G_s]^{-1}), \end{aligned} \tag{12}$$

where K_s and G_s are the bulk and shear moduli of the sediment without hydrate and f_h is the volume fraction of hydrate in the solid phase, calculated as:

$$f_h = \frac{\phi\, S_h}{1 - \phi\,(1 - S_h)}. \tag{13}$$

The dry and saturated moduli and the velocities can then be determined using equations 8, 9 and 10.

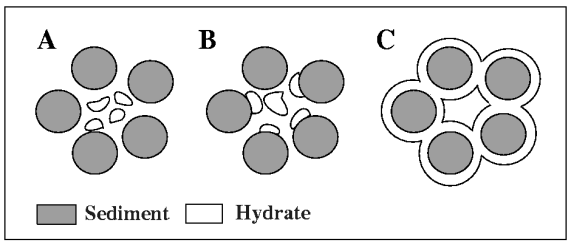


Figure 2: Hydrate inclusion models discussed in Ecker et al. (1997). In model A, the hydrate is part of the pore fluid, in model B the hydrate is part of the solid frame, and in model C the hydrate cements the grain contacts. [**NR**] ohad1/. hydrate-models

3D FORWARD MODELING OF SEVEN-COMPONENT DATA

Modeling setup

The purpose of our forward modeling was to synthesize ocean-bottom seismic acquisition, therefore we used a simple 2-layer model of water over solid. The source was at the water surface, and receivers were at the water bottom. There was a near-seabed anomaly included in the model. The anomaly generated scattering of both P- and S-waves, which upon interacting with the seabed also gave rise to seabed interface waves, termed Scholte waves.

The modeling geometry is shown in Figure 3. Inline and crossline sections going through the center of the anomaly of the V_p and V_s velocity models are shown in Figures 4(a) and 4(b). The parameters of the two layers were:

1. $V_{p1} = 1.5\,\mathrm{km/s},\ V_{p2} = 1.7\,\mathrm{km/s}.$
2. $V_{s1} = 0\,\mathrm{km/s},\ V_{s2} = 0.6\,\mathrm{km/s}.$
3. $\rho_1 = 1.025\,\mathrm{gr/cm^3},\ \rho_2 = 2.0\,\mathrm{gr/cm^3}.$

The anomaly was a Gaussian, which extended outward to a radius of 10 meters, and was centered 10 meters below the seabed. The medium parameters at the center of the anomaly were $V_p = 1.85\,\mathrm{km/s}$, $V_s = 0.9\,\mathrm{kms}$, and $\rho = 2.4\,\mathrm{gr/cm^3}$. Its coordinates were $Y = 50m, X = 150m$. The near-seabed anomaly simulates a "rock" buried just below the seabed, or the leg of a platform, either of which could be sources for scattered interface waves.

The source was a pressure source simulating an airgun, located along the water surface at $Y = 50m, X = 50m$. The wavelet was a Ricker with 25Hz central frequency. There is no source-side ghost from the water surface, since we used an absorbing upper boundary. However, this ghost is simulated by the second lobe of the injected Ricker wavelet. The lower boundary was rigid, simulating a flat reflector with a reflection coefficient of 1. This rigid lower boundary generated body wave reflections.

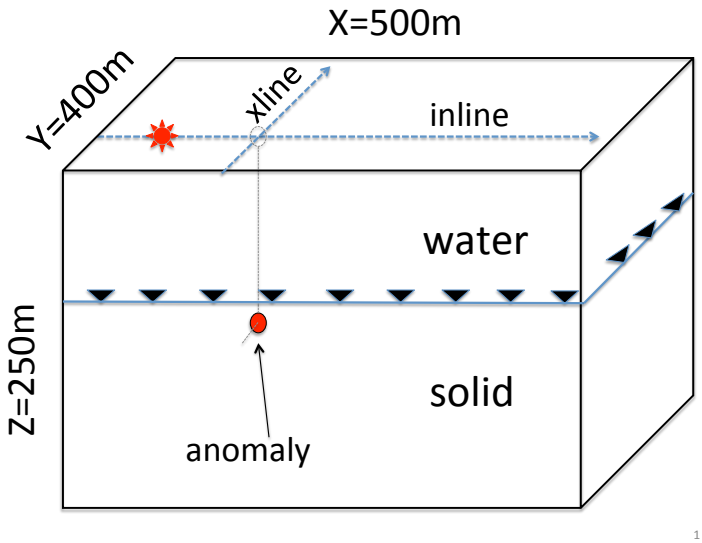


Figure 3: Arrangement of forward modeling. Receivers are on the sea bed, at a depth of 100m. The red star indicates the source position. The red circle indicates the anomaly position. The dashed lines indicate the inline and crossline coordinates along which later sections are displayed. The anomaly is inline with the source. Absorbing boundaries were used everywhere except for the bottom boundary, which was rigid. This simulated a flat reflector with a reflection coefficient of 1. [**NR**] ohad1/. exp-schematic

Synthetic data

The following Figures 5-11(b) show the seven components of the recorded data. The particle-velocity figures are all scaled identically, and therefore the amplitudes are comparable between figures. The same is true for the rotational data figures.

Figure 5 is the hydrophone section, centered on the coordinates of the anomaly. The inline section is along a line connecting the pressure source and the anomaly center (see Figure 3). Some of the arrivals are annotated."P1" is the first P reflection off the bottom rigid boundary. Two more P reflections off the same boundary can be seen at later times. "SV1" is the first S wave reflection off the bottom boundary. A faint Scholte wave can be observed as well. Most of the pressure is due to the P-waves, yet the other waves still transmit some pressure into the water column as a result of using a "hard" water bottom in the synthetic modeling, instead of a more realistic velocity/density gradient that generally exists in the shallow subsurface.

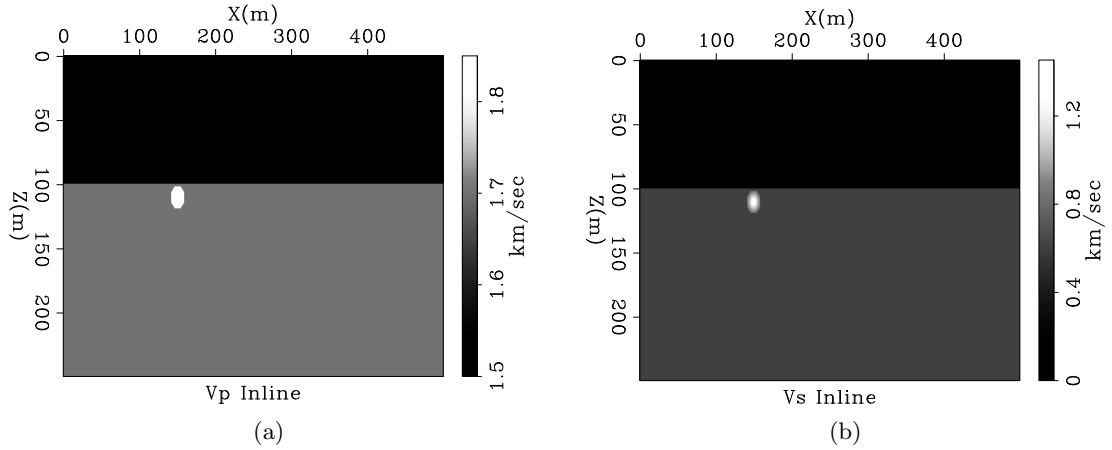


Figure 4: Velocity models. (a) Inline V_p velocity model going through center of anomaly. (b). Inline V_s velocity model going through center of anomaly. The anomaly is a Gaussian, with a radius of 10 meters. [**ER**] ohad1/. vp3d-anom,vs3d-anom

Figures 6(a) and 6(b) are the inline and crossline sections of the v_z component. The Scholte wave is relatively strong on this component in both propagation directions. However, on the inline v_x component shown in Figure 7(a), the Scholte wave is weaker. This is a result of the incidence angle of the direct wave on the anomaly (45^o). The Scholte wave is propagating at about 90% of the shear-wave velocity in the medium.

Figure 7(b) is the v_x component in the crossline direction. Here we do not observe a Scholte wave. Rather, we see "SH1", a horizontally polarized shear wave that was generated by the direct wave's scattering off the anomaly in the crossline direction. This wave propagates at shear-wave velocity, just slightly faster than the Scholte wave. There is also "SH2", which is similarly horizontally polarized. This scattering is generated by the body SV reflection, but seems to have a hyperbolic moveout and not a linear moveout typical of a surface wave.

Figure 8(a) should ideally not show any energy, however some energy is still visible. This is due to the numerical approximation of the finite difference scheme, and also the result of using a staggered grid. We do observe particle velocity v_y in the crossline section in Figure 8(b). Most of it is due to the body waves, but there is some Y-motion resulting from Scholte waves scattering in the crossline direction.

Figure 9(a) is the rotation around the Y-axis (r_y, or pitch) in the inline direction. Note the prominence of the Scholte wave in this section versus Figures 6(a) and 7(a). This suggests that the Scholte wave has a relatively stronger rotational versus linear component than does the P-wave. This agrees with the retrograde elliptical motion of Scholte waves, which causes a greater rotational deformation in the medium than do body waves. Note also the prominence of the rotation generated by the SV wave. It is evident that all wave modes generate pitch motion. This is due to the fact that the measurement surface is effectively a discontinuity in the medium (solid $\rightarrow$ fluid). Therefore even an incident P-wave causes the surface to tilt slightly, and generate a reading in the rotation sensor.

In the crossline direction of the pitch component in Figure 9(b), only the body waves

should be visible, since the source is offset in the inline direction away from the observed crossline. There is still some scattered energy visible, however, and we attribute this again to the staggered-grid, and also to the fact that near the scatterer the wavefield is highly curved. The same reasoning explains why we see roll (r_x) energy in the inline direction in Figure 10(a), where none should ideally exist. Such energy implies a wave that is corkscrewing in the direction of its propagation.

Figure 10(b) shows the roll (r_x) component in the crossline direction. Here both the body waves and the Scholte waves scattered in the crossline direction are distinguishable. Also visible are two linear artifacts marked "A1" and "A2". These artifacts are the result of the absorbing boundaries generating a rotational motion in the wavefield.

Figure 11(a) is the rotation around the Z-axis (r_z, or yaw) in the inline direction. Note the nearly zero energy in the section. This is because none of the waves present in this direction, i.e., P, SV and Scholte, generate a horizontal deformation on the sea bottom. However, in the crossline direction in Figure 11(b) we can see scattered waves "SH1" and "SH2", visible also in the crossline v_x section in Figure 7(b). These waves are not Scholte waves, but rather SH waves propagating at shear velocity. They are picked up only by the vertical rotation component, since they are the only ones that can generate a horizontal deformation on the measurement surface. Also visible are the two artifacts "A1" and "A2" generated by the presence of the absorbing boundary.

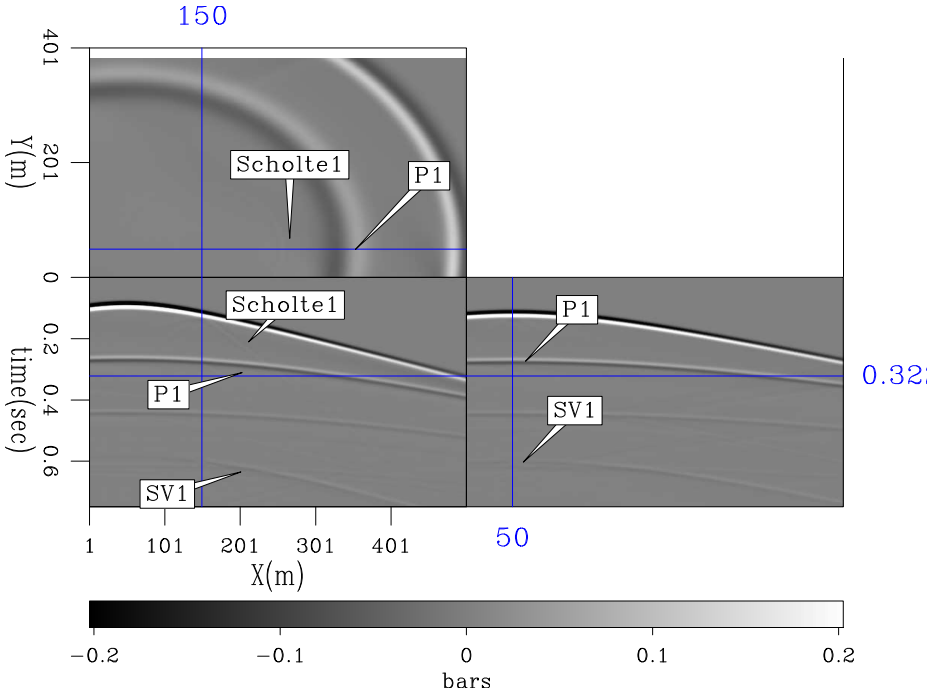


Figure 5: Pressure recorded at the sea bed along inline, crossline and time-slice. The inline and crossline sections are centered on the anomaly coordinates. Some of the arrival time windows where SVD is applied are annotated. Note that most of the pressure is due to P-waves, yet all wave types transmit some pressure into the water column as a result of using a "hard" water bottom in the synthetic modeling. Note also the faint appearance of the Scholte wave. [**CR**] ohad1/. 0Pr2

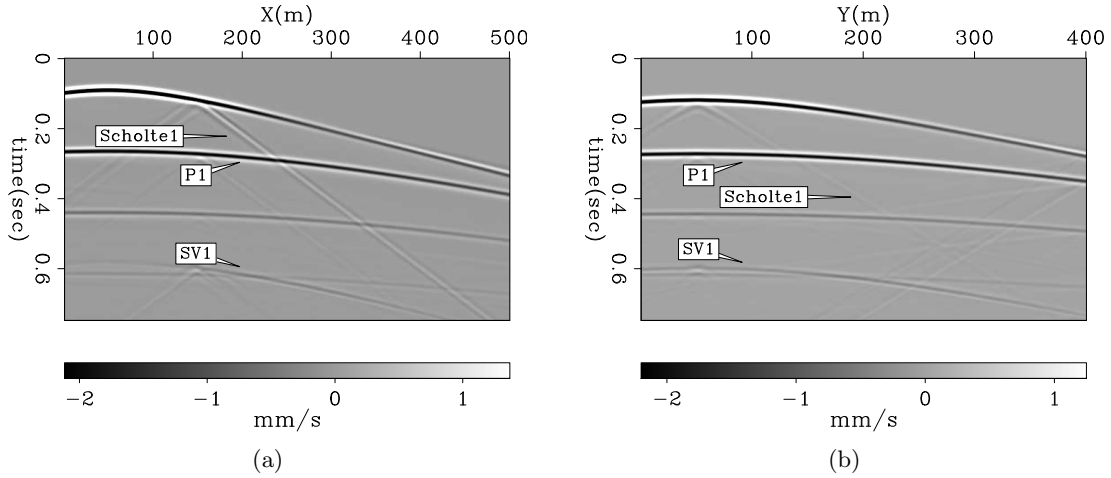


Figure 6: Vertical particle velocity (v_z) recorded at sea bed. (a) Inline direction at y=50m. (b) Crossline direction at x=150m. Some of the arrival time windows where we applied SVD are annotated. "Scholte1" is generated by the direct arrival hitting the anomaly. "P1" is the first P reflection off the bottom rigid boundary. Another two P reflections off the same boundary can be seen at later times. "SV1" is the first S wave reflection off the bottom boundary. Each body wave arrival generates some scattered Scholte wave when it hits the anomaly. [**CR**] ohad1/. 0vzr2-il,0vzr2-xl

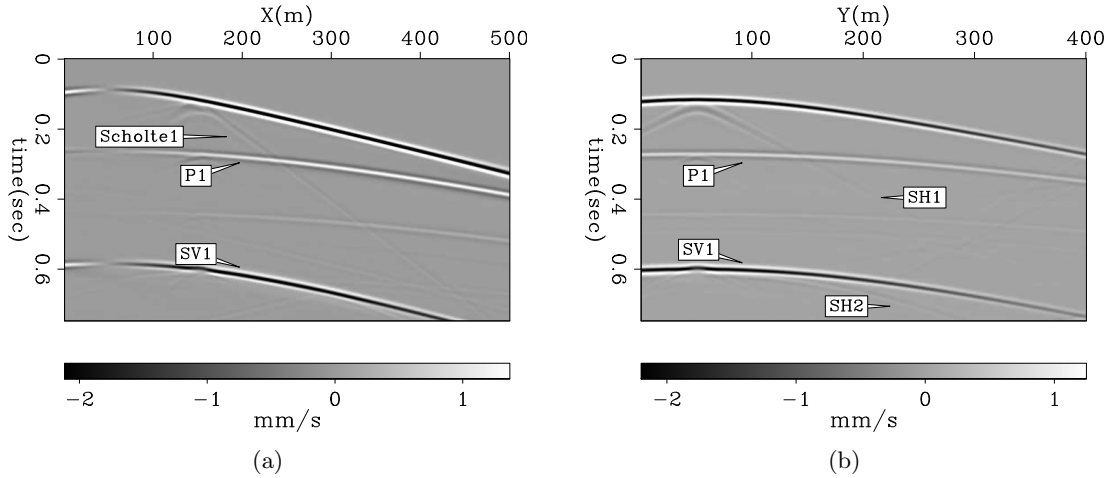


Figure 7: Horizontal particle velocity in X direction (v_x) recorded at sea bed. (a) Inline direction at y=50m. (b) Crossline direction at x=150m. [**CR**] ohad1/. 0vxr2-il,0vxr2-xl

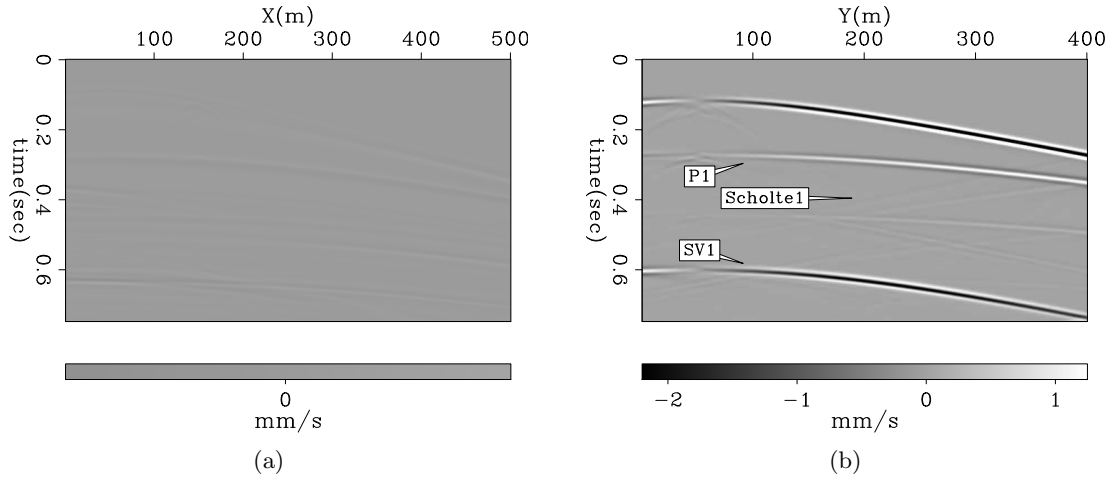


Figure 8: Horizontal particle velocity in Y direction (v_y) recorded at sea bed. (a) Inline direction at y=50m. (b) Crossline direction at x=150m. [**CR**] ohad1/. 0vyr2-il,0vyr2-xl

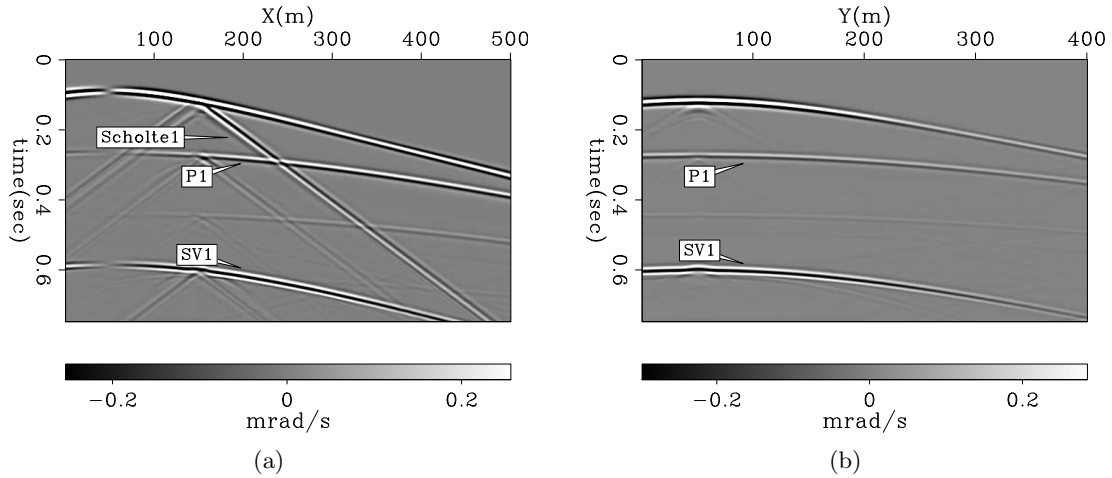


Figure 9: Horizontal particle rotation around Y-axis (r_y, pitch) recorded at sea bed. (a) Inline direction at y=50m. (b) Crossline direction at x=150m. [**CR**] ohad1/. 0rotyr2-il,0rotyr2-xl

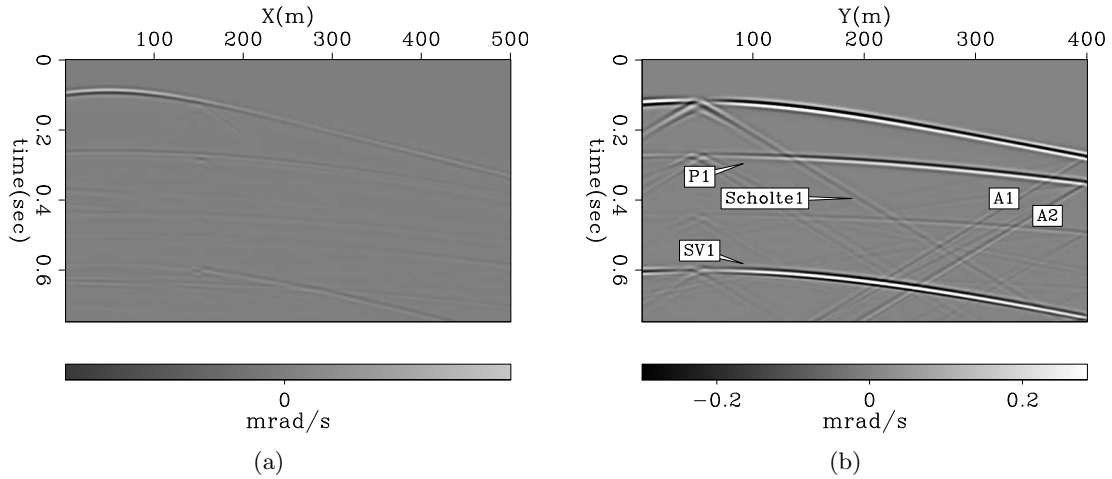


Figure 10: Horizontal particle rotation around X-axis (r_x, roll) recorded at sea bed. (a) Inline direction at y=50m. (b) Crossline direction at x=150m. [**CR**] ohad1/. 0rotxr2-il,0rotxr2-xl

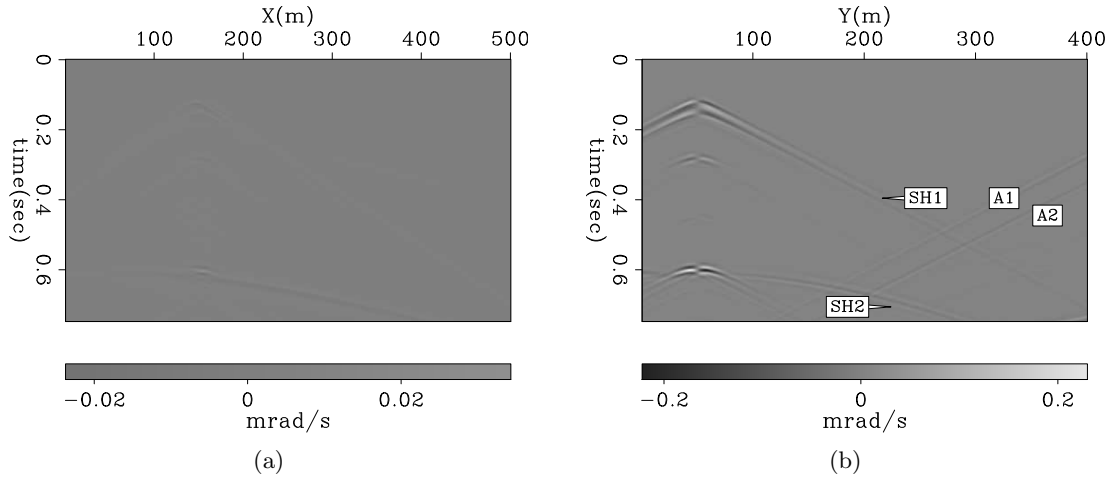


Figure 11: Vertical rotation around Z-axis (r_z, yaw) recorded at sea bed. (a) Inline direction at y=50m. (b) Crossline direction at x=150m. [**CR**] ohad1/. 0rotzr2-il,0rotzr2-xl

Singular value decomposition of arrivals

The seven component data comprises data of different units and different amplitude ranges. To use SVD in a meaningful way, we first normalized the pressure component, the particle-velocity components and the rotational-motion components separately. Each component type (pressure, linear motion, rotational motion) was normalized by the maximum value within that component. We then applied SVD to time windows taken around the arrivals annotated in Figures 6(a)-11(b). The purpose was to have just one arrival within each time window, but as a result of multiple events and artifacts this was not always the case.

Figures 12(a)-13(d) show the first two polarization vectors for each arrival, scaled by the respective first two singular value (i.e. s_1 and s_2 of equation 6). Most of the amplitude of the arrivals resides in these two polarization directions. Figure 12(a) shows the polarization of the P-wave arrival in the inline direction. ("P1" in Figure 6(a)). Note how the particle-velocity v_z and pressure component H are dominant, but we also see that the pitch (r_y) component is strong, too. This P arrival is as at a near offset, so most of the linear motion should be in Z. The rotational motion is a result of the P-wave impinging on the water-bottom discontinuity at an angle, causing the surface to tilt slightly.

Figure 12(b) shows the polarization of the SV arrival in the inline direction. There is a large rotational component in the r_y direction, while in the linear direction the v_x component is dominant. This S arrival is at the same offset as the P arrival, therefore we can expect most of the linear motion to be in the X direction.

The polarization of the inline Scholte wave is shown in 12(c). Note how the rotational motion is the dominant component here.

Figure 13(a) is again a P-wave arrival, but along the crossline direction (see Figure 3), and at an inline offset from the source position. Therefore it is reasonable that it is polarized linearly in all directions. It also explains why this arrival is also polarized in both r_x and r_y directions.

Figure 13(b) is the SV arrival at the same ocean-bottom location as Figure 13(a). It is polarized in the linear v_x and v_y directions, and in the rotational r_x and r_y directions. This polarization is consistent with this waves propagation direction at this particular location.

Figure 13(c) is the Scholte wave in the crossline direction. Unlike Figures 13(a) and 13(b), this arrival is only offset from its source, the sea-bottom anomaly, in the crossline direction. Therefore it is polarized in the r_x direction only.

Figure 13(d) is an SH wave ("SH2" in Figure 11(b)). It is the result of scattering off the anomaly. This arrival is taken from an offset in both the crossline and inline direction from the anomaly. The dominant polarization direction is in the r_z direction. We also see that it is polarized in the linear v_x direction. Upon closer inspection of this arrival, we saw that the chosen time window also contains some P-wave energy and some of a boundary-generated artifact. This is the reason for the strong r_x polarization.

Observing all these figures, we can arrive at guidelines for identifying the arrival type from the polarization vectors:

1. P-waves are polarized in the hydrophone direction, but mostly in the linear motion components' directions. They have a lesser polarization in one or two horizontal

rotation directions.

2. SV waves are polarized largely in the horizontal rotational-motion direction, but also have a significant polarization in one or two linear-motion directions.

3. Scholte wave polarization is predominantly in the horizontal-rotation directions.

4. SH waves are unique in their polarization direction in the vertical-rotation direction. They also have a significant polarization in one or two linear-motion directions.

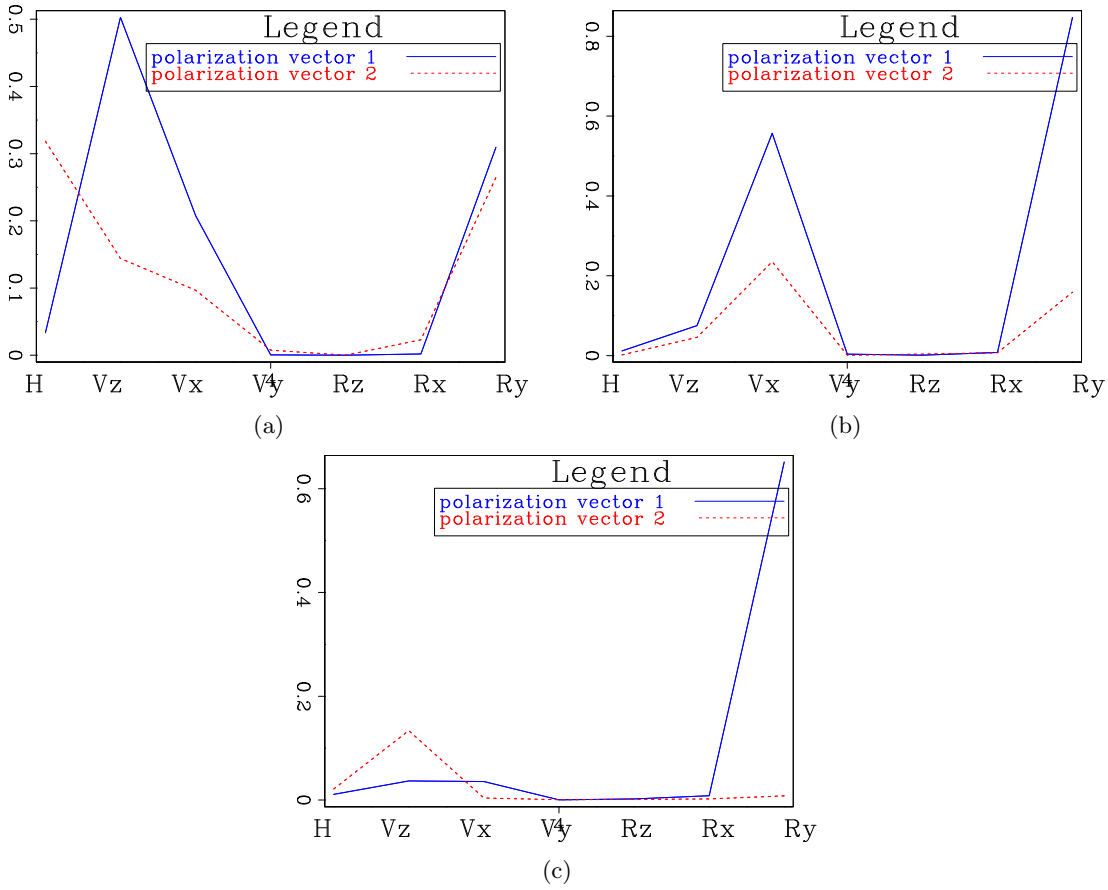


Figure 12: The first two polarization vectors resulting from application of SVD to the seven components of the arrivals annotated in Figure 6(a) at the ocean-bottom receiver at $x = 200, y = 50$. The vector components are along the X-axis, and the Y-axis shows the magnitude of the vector along each component. (a) Inline reflected P-wave arrival ("P1" in Figure 6(a)). (b) Inline reflected SV wave arrival ("SV1"). (c) Inline scattered Scholte wave ("Scholte1"). The units on the Y-axis are amplitudes of the vector components within each arrival's time window. The P-wave is polarized mostly in the v_z axis, but also in the hydrophone and r_y axes. The SV wave is polarized mostly in the r_y and v_x axes. The Scholte wave is polarized almost entirely in the r_y direction. [**CR**]

ohad1/. il2-p1,il2-sv1,il2-scholte1

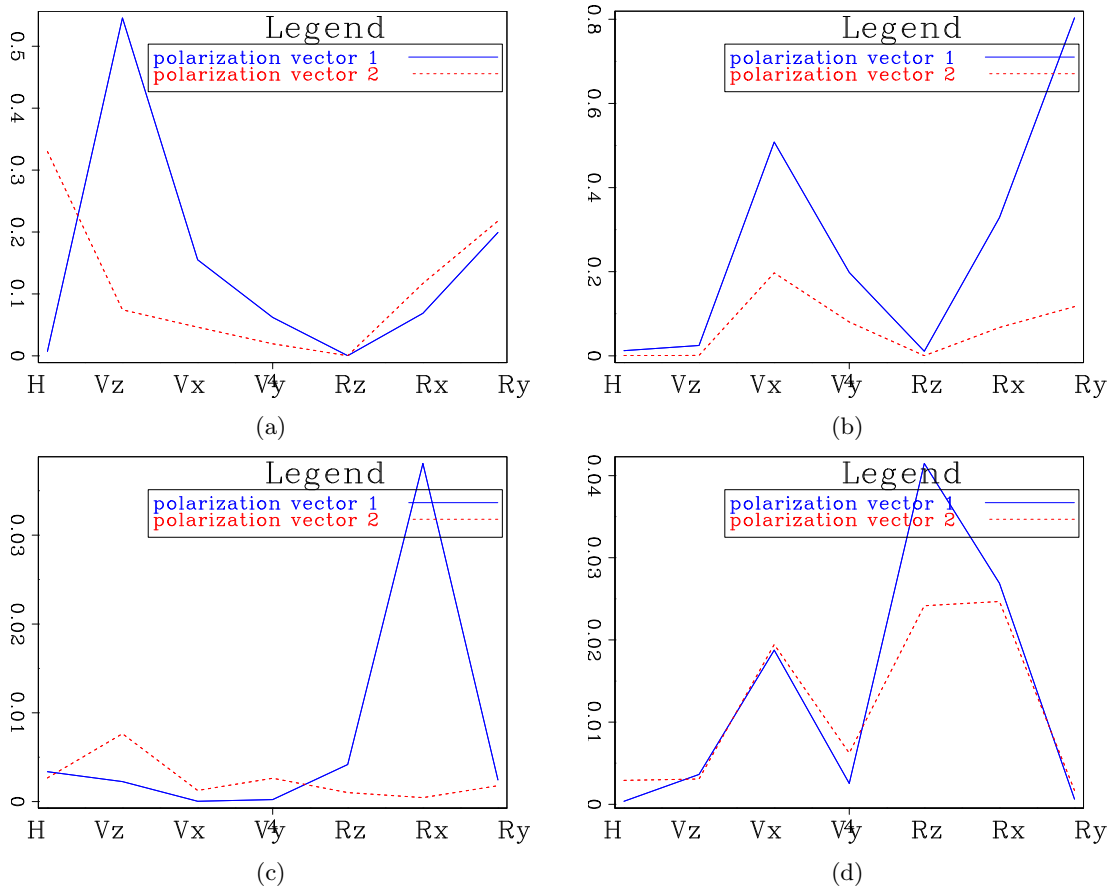


Figure 13: The first two polarization vectors resulting from application of SVD to the seven components of the arrivals annotated in Figure 6(b). The vector components are along the X-axis, and the Y-axis shows the magnitude of the vector along each component. (a) Crossline reflected P-wave arrival at $x = 150, y = 90$ ("P1" in Figure 6(b)). (b) Crossline reflected SV wave arrival at $x = 150, y = 90$ ("SV1"). (c) Crossline scattered Scholte wave at $x = 150, y = 212$ ("Scholte1"). (d) Crossline scattered SH wave at $x = 280, y = 230$ ("SH2" in Figure 11(b)) The units on the Y-axis are amplitudes of the vector components within each arrival's time window. The P-wave is polarized mostly in the v_z and v_x axes, but also in the hydrophone and r_y and r_x axes. The SV wave is polarized mostly in the r_y and v_x axes, but also in the v_y and r_x axes. The Scholte wave is polarized almost entirely in the r_x direction. The SH wave is polarized in the r_z direction, and also in the r_x and v_x directions. [**CR**] ohad1/. xl2-p1,xl2-sv1,xl2-scholte1,dl2-sh2

AVO OF 2D ROTATIONAL DATA

For the mineralogy, porosity and hydrate saturation of the medium near the Blake Outer Ridge, we used the values estimated by Ecker et al. (1997). We then used equations 8-13 to generate two effective elastic media, one for the "hydrate as fluid" model and one for the "hydrate as solid" model.

The media had three layers: A water layer, a water-saturated sediment, and a hydrate saturated sediment. The only difference between the media was the inclusion model, and therefore the only difference between the models was in the V_p, V_s and density of the hydrate layer. The models are shown in Figures 14(a)-14(d).

We synthesized OBS data with these models, and generated CMP gathers for one surface location. We used flat layers only, therefore the amplitudes in these CMP gather approximate those of a basic AVO done in surface coordinates. The gathers were of the hydrophone component and of the r_y component. The wave-propagation was done in two dimensions only, therefore the only rotational component is around the Y axis.

In Figures 15(a)-15(f) we show the PP and the PS reflections from the top of the hydrate layer. All Figures are of the same CMP location, and the maximum reflection angle present in the gathers is 40^o. Figure 15(a) is the hydrophone CMP gather for hydrate model 1, and Figure 15(b) is the r_y CMP gather for the same model. Note how the r_y component shows a stronger converted-wave response than the hydrophone, and records very little of the primary P-wave by comparison.

Figures 15(c) and 15(d) are the hydrophone and r_y components for hydrate inclusion model 2. In this case the hydrate is part of the solid, resulting in an increase of the bulk modulus. This increases the P-wave reflectivity, and therefore the amplitude of the PP reflection in the hydrophone component. However, the shear modulus is relatively more affected by the change, and therefore the PS reflection's amplitude increases by a greater factor. Note how this change in shear modulus is more significant on the r_y component than on the hydrophone component.

Figures 15(c) and 15(f) are the differences in reflection amplitudes with offset between the two hydrate models. The hydrophone component does vary with the change in inclusion model, however note how the r_y component exhibits much more sensitivity to the same change, especially at the shorter offsets.

DISCUSSION AND CONCLUSION

We have applied SVD as a method to ascertain the motion components of discrete arrivals. We have shown that different wave-types have different signatures when observing their scaled polarization vectors after applying SVD to the multicomponent data. The rotation data provide additional information about the motion components, and by inference they provide an indication as to the wave type recorded by ocean-bottom acquisition.

We have shown that interface waves have a predominantly rotational signature. Shear waves also have a large rotational component, however they also exhibit a stronger linear polarization than the interface waves. P-waves have a mostly linear polarization, although they do have some rotational-motion component. The P-waves are unique in having a

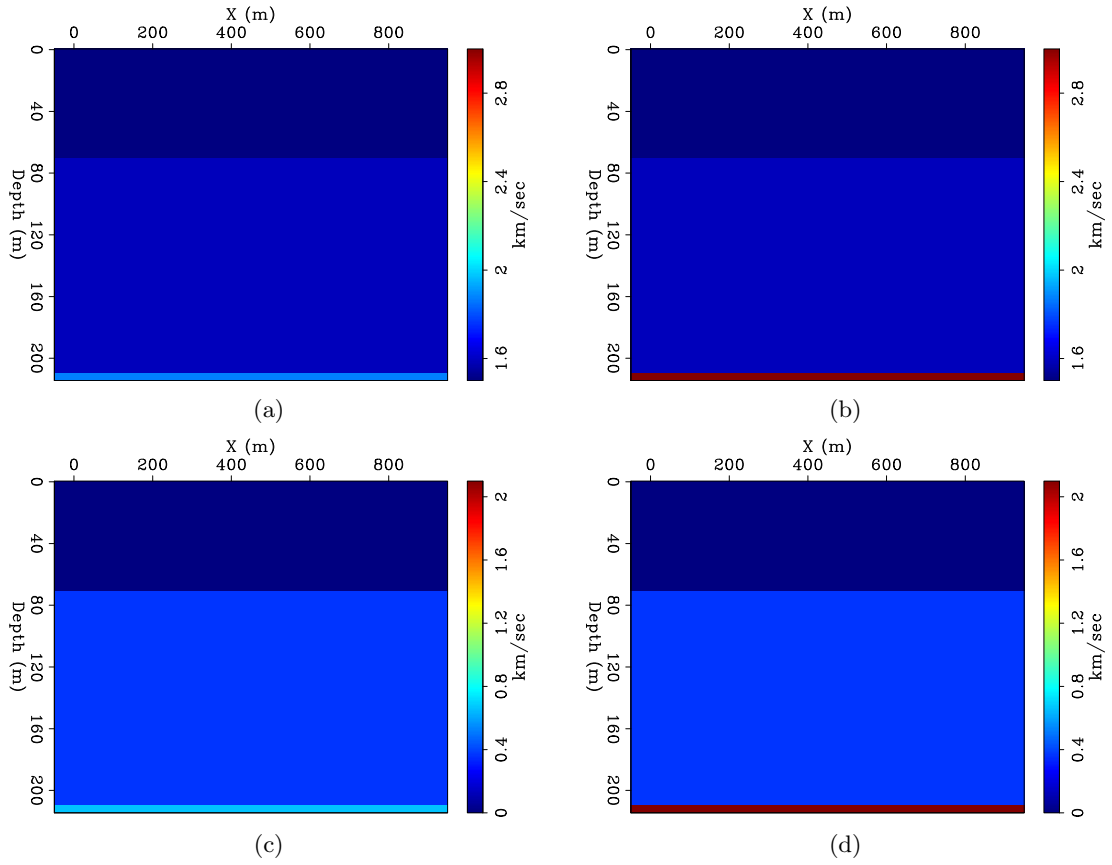


Figure 14: V_p and V_s models resulting from the two hydrate inclusion models. The hydrate layer is at $z = 207$m, and it is only in that layer that the models differ. (a) V_p of hydrate inclusion model 1, hydrate velocity is 1.87km/s. (b) V_p of hydrate inclusion model 2, hydrate velocity is 3km/s. (c) V_s of hydrate inclusion model 1, hydrate velocity is 0.66km/s. (d) V_s of hydrate inclusion model 2, hydrate velocity is 2.1km/s. [**ER**] ohad1/. vp2d-hyd1,vp2d-hyd2,vs2d-hyd1,vs2d-hyd2

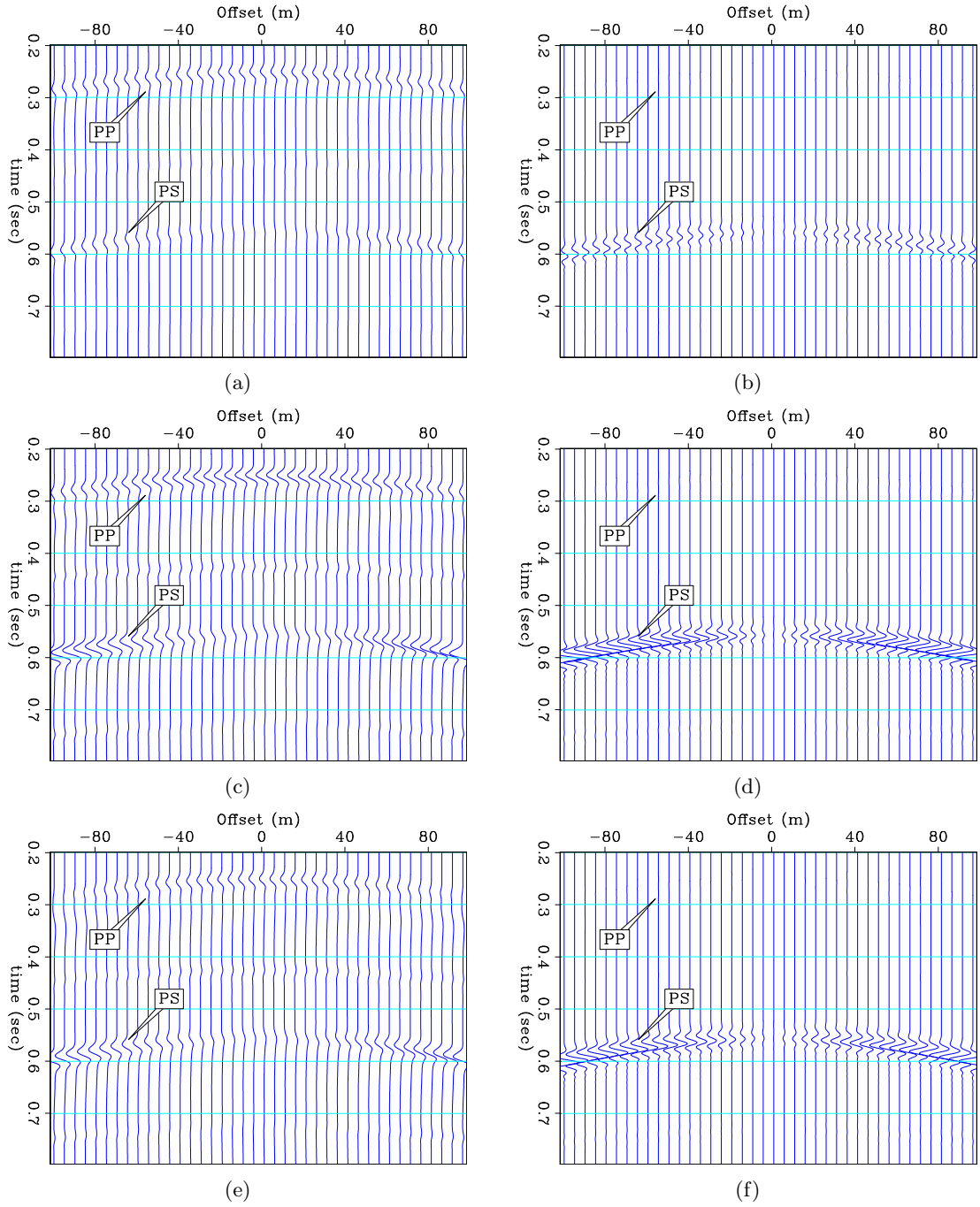


Figure 15: A CMP gather from the two hydrate models, showing the PP and the PS reflections from the top of the hydrate layer as they appear on the hydrophone and the rotational r_y (pitch) component. (a) Hydrophone, hydrate model 1. (b) r_y, hydrate model 1. (c) Hydrophone, hydrate model 2. (d) r_y, hydrate model 2. (e) Hydrophone response difference. (d) r_y response difference. The rotational r_y component's AVO is evidently more sensitive to the change in hydrate inclusion model than is the hydrophone. [**CR**]

ohad1/. 0Pr-diff1,0rotyr-diff1,0Pr-diff2,0rotyr-diff2,0Pr-diff3,0rotyr-diff3

significant polarization in the hydrophone "direction".

We have also shown that it is possible to distinguish between SV and SH waves by observing the amplitude of their r_z polarization versus their r_x or r_y polarizations. It is interesting to note that the only waves missing in the vertical rotation section (Figure 11(b) are the P and SV body waves. This suggests an interesting application for a vertical rotation sensor: selection of SH or of Love waves.

The advantage of having this information is that we do not need to rely on a good spatial sampling of the data in order to identify and separate the wave-types by their move-outs/wavenumbers. This is particularly relevant for surface waves, since they propagate at much lower velocities than body waves, and are likely to be aliased in the data.

There are as yet some unresolved issues in regard to SVD of seven-component data:

1. Scaling between hydrophones, geophones and rotation sensors. Each type of sensor measures a different unit, and it is the relation between the amplitudes of the different components that will determine the polarization vectors. Even if there is a lot of surface wave energy in a particular arrival, the recorded rotational-motion values must be scaled accordingly such that they won't be too small compared to the linear components.

   In this synthetic example, the entire dataset had ample amounts of each motion component, and therefore normalizing the hydrophone, geophones and rotation sensors separately was successful. However, for field data the question of how to scale each of the components is not trivial. It is reasonable to assume that the greatest amplitude of the vertical geophone corresponds to greatest amplitude of the pressure (the direct wave propagating vertically from the source to the geophone directly beneath it on the seabed). However, it is not so intuitive to see where the greatest value of rotation should be, and therefore the scaling between geophone and rotation sensors remains an open question at this point.

2. We have shown a noise-free synthetic. How will noise affect the SVD result in field data? How much noise is required in order for the particular signatures of each arrival to not be obvious? Will noise be equally distributed between the linear and rotational components?

We have also shown that rotational data can provide us with better constraints for rock-physics models. In this case, the change in the rock-physics model had a significant effect on the shear modulus of gas hydrate, and therefore on the shear-wave reflectivity. Since the reflected S wave generates a greater rotation on the ocean-bottom than the P-wave, the r_y sensor provides us with a proxy for the shear-wave reflectivity. We can use this proxy to construct AVO from seismic data, and validate our hydrate inclusion model (given some known rock-properties of the medium).

ACKNOWLEDGEMENTS

We thank Tim Owen, Malcolm Lansley and Bob Brune for prompting us to consider rotation sensors, Fred Herkenhoff for showing us rotational field data from Chevron's recent experiment, and John Evans from USGS.

REFERENCES

Brokesova, J. and J. Malek, 2010: 2nd IWGoRS Workshop in Prague.

Brune, R., D. Aldridge, M. van der Baan, J. Gaiser, and E. Muyzert, 2012, Theta-data: Introduction to rotational seismology and its potential uses: 82nd SEG Conference and Exhibition.

deMeersman, K., M. van der Baan, and J. Kendall, 2006, Signal extraction and automated polarization analysis of multicomponent array data: Bulletin of the Seismological Society of America, **96**, 2415–2430.

Dvorkin, J. and A. Nur, 1993, Rock physics for characterization of gas hydrates: The future of energy gases: USGS Professional Paper 1570, 293–298.

———, 1996, Elasticity of high-porosity sandstones: theory for two north sea datasets: Geophysics, **61**, 1363–1370.

Ecker, C., J. Dvorkin, and A. Nur, 1997, Estimating the amount of gas hydrate and free gas from surface seismic: Stanford Exploration Project, **Report 95**, 173–195.

Edme, P., E. Kragh, and E. Muyzert, 2012, Noise attenuation using rotation data: U.S. Patent application publication, **Pub. No.: US 2012/0250460 A1**.

Lee, W. H. K., H. Igel, and M. D. Trifunac, 2009, Recent advances in rotational seismology: Seismological Research Letters, **3**, 479–490.

Muyzert, E., A. Kashubin, E. Kragh, and P. Edme, 2012, Land seismic data acquisition using rotation sensors: 74th EAGE Conference and Exhibition.

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

Synthetic model building using a simplified basin modeling approach

Robert G. Clapp

ABSTRACT

Generating a realistic synthetic model is a challenging problem in a geophysical research environment. Models are often too simple, lacking much of the character of real data, or too complex making debugging difficult. I propose a different way to generate synthetic models, by allowing the user to specify a series of geologic events such as deposition, erosion, and compaction. The result of each event is approximated on the current model. This approach has the benefit of allowing complex models to be built, easily extendable to multiple model parameters, and allows the user to "turn off" events, conveying the construction of simpler models by stages. Early results indicate that this approach allows complex synthetic models to be generated with minimal effort.

INTRODUCTION

Synthetic models play an important role in geophysical research. Over the last 25 years, SEG and company generated synthetics (Versteeg, 1993) have been used extensively by researchers throughout the world. These synthetics have found such wide use for two reasons. First, generating realistic synthetic models is difficult, particularly 3-D models. Tools such as Gocad (Cain et al., 1998) allow the user to specify layers, faults, and other geologic features before creating a 3-D grid to ease the process, but still require significant experience to use effectively. Second, industry wide synthetics provide a common benchmark to compare results from different companies. These models are useful for final testing of a finished algorithm, but often are too complex to be used in algorithmic development.

Another approach to generating synthetic models is following the methodology of basin modelers. Basin modeling takes a more geologic approach to describing models than the more computer science based approach followed by GoCad and its competitors. Basin modeling attempts to model the geologic history of a given piece of the earth to explain its current properties.

In this paper I introduce a new tool, `Geomodel`, for generating more realistic synthetic models. `Geomodel`builds synthetic models through a user specified set of geologic events (such as deposit, erode, compact, or uplift). By building synthetic models in this manner the user can add or subtract complexity through the algorithmic development process.

OVERVIEW

`Geomodel`builds up a synthetic model in stages. The user begins by specifying an initial homogeneous basement through sampling and size in `x` and `y`, spacing in `z`, as well the sampling, thickness and property value (think velocity) of the basement material.

The rest of the model is then built up by describing a series of geologic *events*. The total number of events is described by the parameter `nevents`, and each successive event is specified by `event#` (`event1, event2`, etc.). Currently seven different events can be specified: deposit, erode, compact, uplift, fault, gauss, and emplace. Figure 1 demonstrates several of these features. Additional events might be added in the future. Each event has its own set of parameters. The parameters for each event are differentiated by their event number. For example when describing the layer thickness for the second event you would use the parameter `thick_2`, while describing deposition for the fourth event you would use `thick_4`. As different events are added to the model two grids are maintained. The first contains the earth property at the given location, the second contains what event created this grid cell. The following describes the parameters for each event option along with the methodology used to simulate the event.

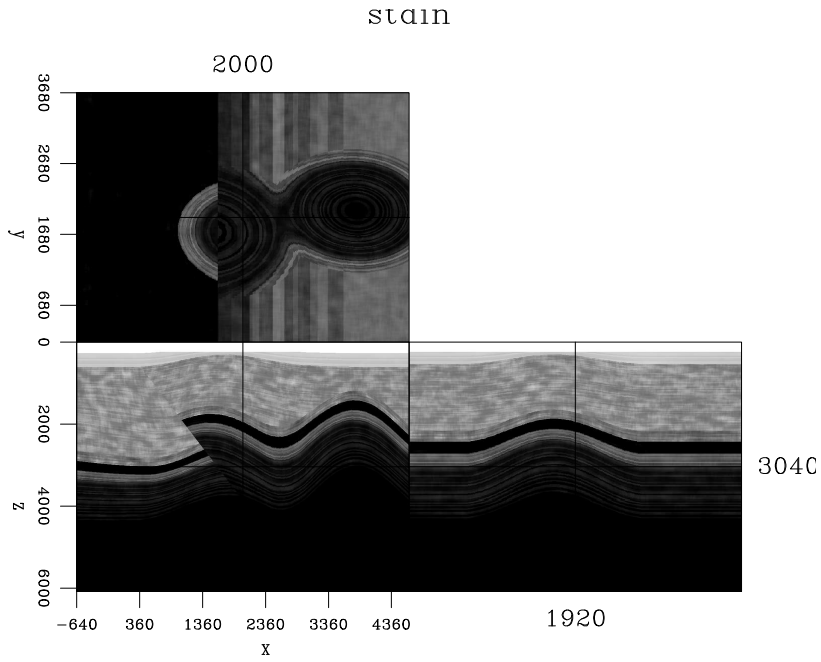


Figure 1: An example model created with `Geomodel`. Note the fault, anticlines, erosion, and deposition visible in this example. bob2/. example

Deposition

Deposition involves laying down a generally homogenous earth property so that the upper bounding surface is completely horizontal. The user specifies the thickness of the layer (`thick_`), which corresponds to thickness in samples added to the current highest portion of the model. The overall average property value for the layer is specified through `prop_`. Starting from this basic homogenous layer the user then has the ability to add different levels of complexity. Interbed layering can be added be describing average interbed layer thickness `layer_`. Each interbed layer property value is randomly assigned a different earth property by scaling the average earth property by a random number scaled by `dev_layer_`. Using a for the average property value, d for the specified deviation, and r for the random number (with a range between -.5 and .5), the property value p for the current layer is

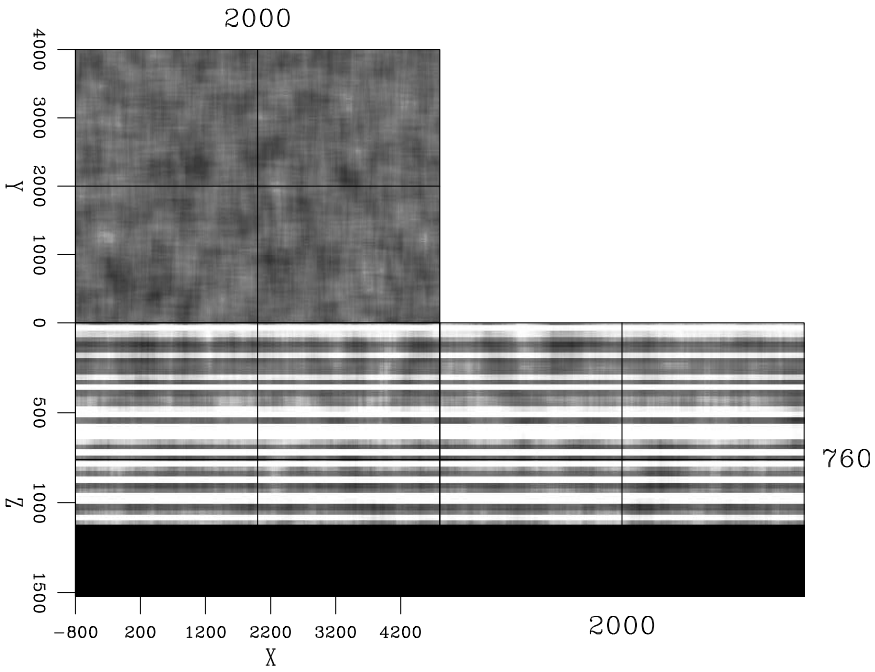


Figure 2: An example of depositing a layer using `Geomodel`. Note the layer visible in the xz and yz planes, and the spatial variation in the xy plane. [**ER**] bob2/. deposit

specified through $p = (1 + rd)a$. Layer thickness can also be varied using the `dev_layer_` parameter. Using this parameter, layer thicknesses are assigned variable lengths with larger values of `dev_layer_` corresponding to larger variations.

Heterogeneity within layers can also be added. The deposition modeling routine scales the layers property values by $s = 1 + r * v$, where v is the user-specified parameter `dev_pos_` and s is the generated random property scaling parameter. These values are then smoothed using successive box smoothers with a half-width of `nsmooth_`. The resulting layer property is then the result of multiplying s and p at every point in the layer. Figure 2 shows an example of using the deposition option. Note the layering visible in the xz and yz planes, and the spatial variation in the xy plane.

Faults

Currently only planar faults are possible in `Geomodel`. The fault plane is specified by describing the plane in terms of $ax + by + cz = d$, using the parameters `plane1_, plane2_, plane3_,` and `plane_sum_`. The user then specifies the movement along the fault plane (fault throw) by specifying an initial point lying on the plane (`old1_, old2_,` and `old3_`) and where that point moves to (`new1_, new2_,` and `new3_`, all in terms of the model's coordinate system. Fault displacement is assumed to decay (described by a cosine taper) as one moves away from (`old_, old2_, old3_`). Along the fault plane decay distance is defined by the parameter `die_max_` and away from the fault plane by the parameter `die_minor_`. Figure 3 shows an example of applying faulting to a multi-layer depositional model. Note the decay in space of the fault displacement.

Uplift

Anticlines are generally formed from compressional forces and tend to form anticline/syncline patterns. Instead of stress from the side, `Geomodel`modifies earth properties by simply uplifting at a given location. The center of the uplift is specified by `center2_` and `center3_`. The amount of uplift is specified by `shift_`. Similar to the fault description uplift also dies out at distance using along two axes using the `die_major_` and `die_minor_` parameters. The `rotate_` parameters allows the user to specify the angle of rotation off the xy-plane for the major and minor axes. Figure 4 shows an example of applying an uplift to the faulted model shown in Figure 3.

Compaction

Compaction can have a signficant effect on earth properties such as velocity. Compaction is generally driven by depth and the material making up the layer. `Geomodel`allows the user to specify different compaction for each layer. The default is no compaction in a given layer; to add compaction the user can specify the parameter `compact_i_j` where `i` is the event and `j` is the layer. A property is modified using the formula $p = p * (1 + d * c)$, where p is the property, d is the depth, and c is the compaction parameter for the given layer. Compaction can be added at any stage in model construction. As a result a model can have many different property gradients depending on when compaction is applied. Figure 5 shows an example of applying the compaction parameter. In this case one layer is compacted before uplift, and another, after uplift.

Gauss

Gaussian-shaped anomalies can be added to the model using the `gauss` option. The `center1_`, `center2_`, and `center3_` parameters specify the center of the anomaly. The `vplus_` parameter specify the maximum amount of the anomaly. The `var_` parameter specifies the variance (in distance) of the Gaussian anomaly. Figure 6 shows an example of adding a Gaussian anomaly to the upper layer of the model shown in Figure 5.

Erode

`Geomodel`simulates two different type of erosion: bowls and rivers. The user specifies which type of erosion using the parameter `erode_type_`. In each case, the depth of the erosion is described by the `depth_` parameter. The user can fill a portion of the eroded event using the `fill_depth_` and `fill_prop_` parameters. These parameters allow the user to simulate events such as a river fill. The location of an erosion bowl event is described the `center2_` and `center3_` parameters. The width of the bowl is defined by the `width2_` and `width3_` parameters.

Meandering river paths are described by summing up to three cosine waves of different spatial frequency (`cosine_wave1_, cosine_wave2_` and `cosine_wave3_`) with weighted different amplitudes (`cosine_amp1_, cosine_amp2_` and `cosine_amp3_`). The beginning `beg3_` and ending `end3_` of the meandering stream in the third dimension need to be specified

along with the angle `angle_` of the stream path. Figure 7 shows an example of both the meandering stream channel and the erosion bowl.

Emplace

Currently `Geomodel`has a rudimentary ability to emplace constant velocity shapes (such as salt). Emplacements are first described on a 2-D rectangular grids whose extents vary from 0,0 to `extent2_`,`extent3_`. The top and bottom of the emplaced object are described by a series of points that are interpolated to fill the entire 2-D grid. The number of points describing the top and bottom are specified using `npts_top_` and `npts_bot_`. The points themselves are defined using the parameters `pts_top_i_j` and `pts_bot_i_j` where i is the event and j is the number of the current point. Two regular surfaces are then interpolated and everywhere the top surface is above the bottom surface the property value `prop_` in assumed. The 0,0 location of the defined grid is then shifted to `pos2_` and `pos3_` at the angle `angle_`. Sediments to the side and above the emplacement are shifted by `shift_up_` amount. Figure 8 shows an example of emplacing a salt body.

DISCUSSION AND FUTURE WORK

The most obvious is the ability to handle curved faults (though this can be simulated by a series of planar faults). The ability to handle planar unconformity and allowing non-planar deposition should be added. Improving the mechanism for emplacing constant velocity structures and creating anticline-syncline structures would be useful.

CONCLUSIONS

Complex synthetic models are useful not only for benchmarking but in the code development process. Complex models can be built up by simulating a series of geologic events such as deposition, erosion, and faulting. `Geomodel`represents an initial attempt to simulate geologic phenomena that can build multi-parameter earth models.

REFERENCES

Cain, G., G. Cambois, M. Gehin, and R. Hall, 1998, Reducing risk in seismic acquisition and interpretation of complex targets using a Gocad-based 3-D modeling tool: 68th Ann. Internat. Mtg, 2072–2075, Soc. of Expl. Geophys.

Versteeg, R. J., 1993, Sensitivity of prestack depth migration to the velocity model: Geophysics, **58**, 873–882.

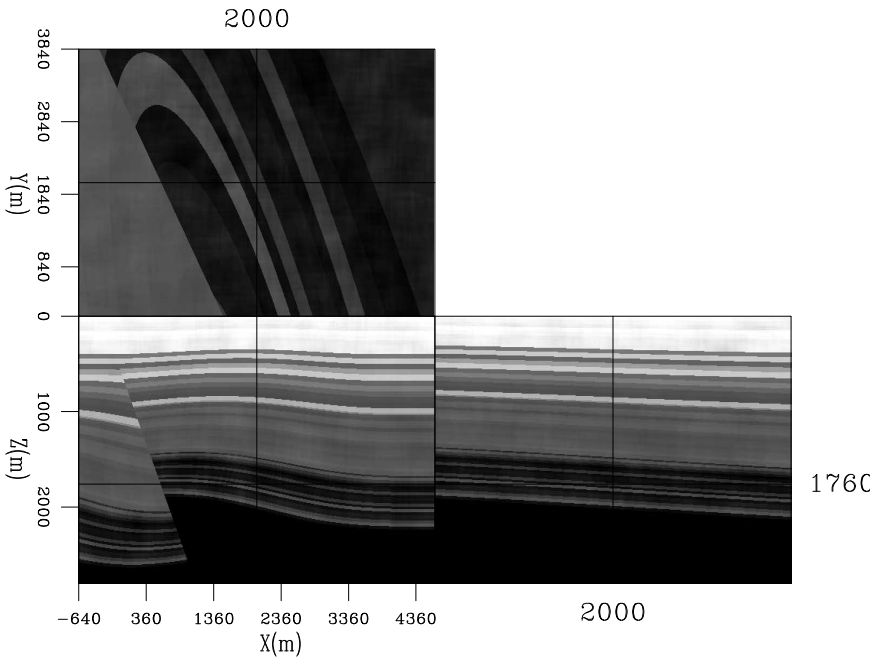


Figure 3: An example of applying faulting to a layered model. Note the decay in fault displacement as we move away from the fault. [**ER**] bob2/. fault

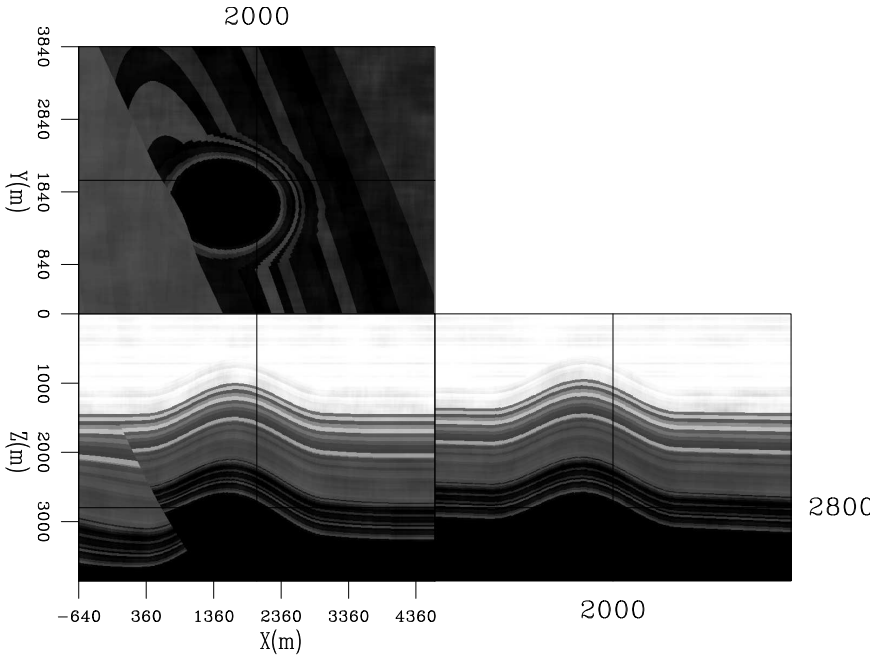


Figure 4: An example of applying uplift, simulating an anticline, to the faulted model shown in Figure 3. [**ER**] bob2/. uplift

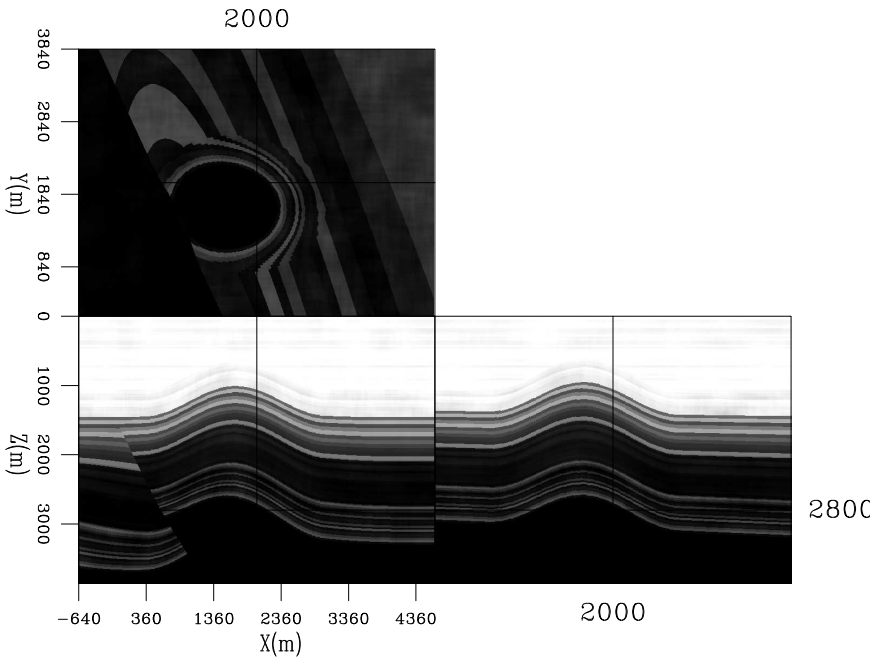


Figure 5: An example of applying the compaction parameter. In this case one layer is compacted before uplift, and another, after uplift. [**ER**] bob2/. compact

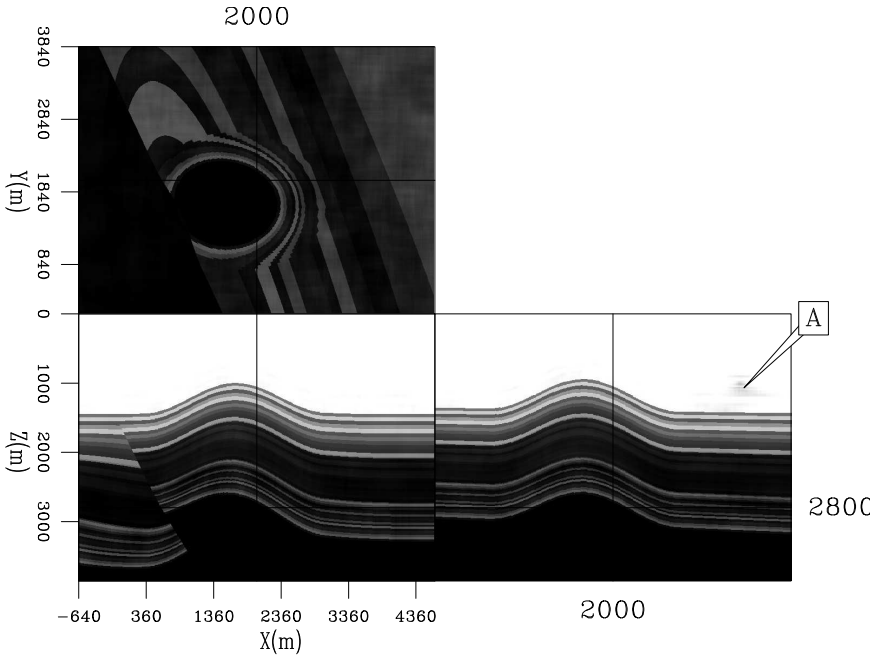


Figure 6: The result of adding a Gaussian anomaly to the model shown in Figure 5. The Gaussian anomaly can be seen at 'A' [**ER**] bob2/. gauss

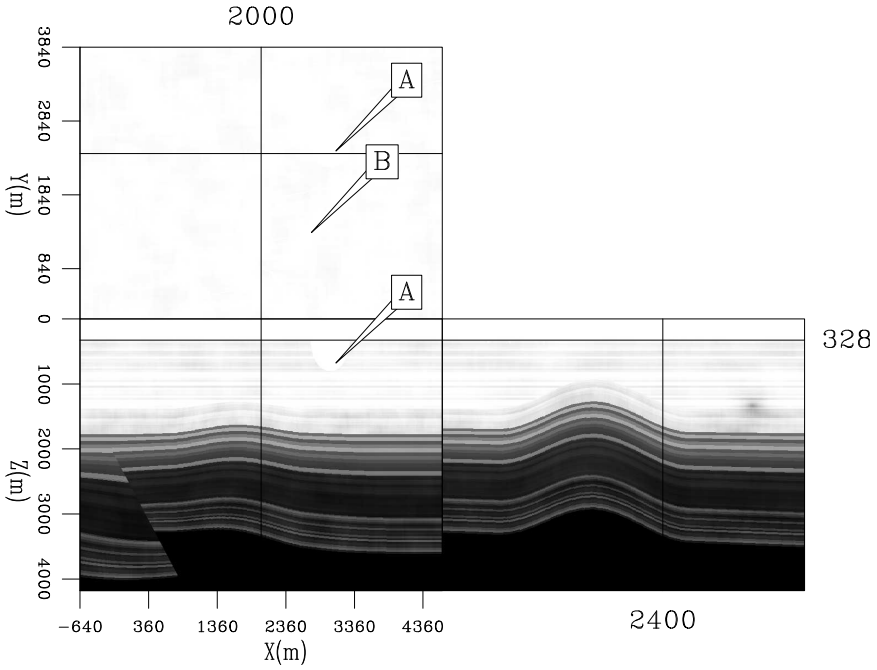


Figure 7: An example of the two type of erosion possible with `Geomodel`. Note how the river channel and the bowl are partially filled in with lower velocity sediment. 'A' marks the bowl erosion, while 'B' shows the river channel. [**ER**] bob2/. erode

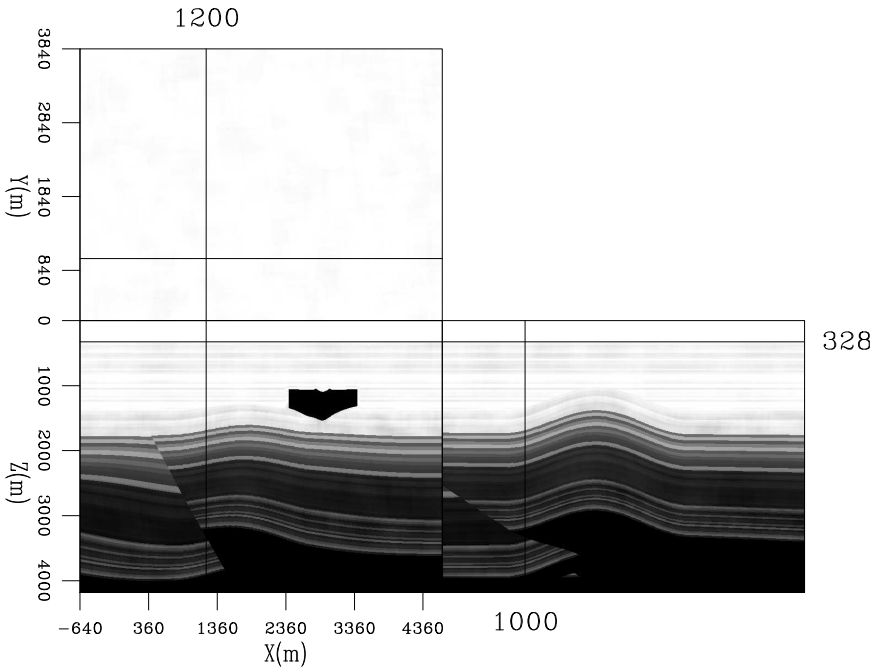


Figure 8: An example of emplacing a salt body into a model. Note how the sediments to the side have been shifted up. [**ER**] bob2/. emplace

Accurate implementation of two-way wave-equation operators

Ali Almomin

ABSTRACT

I present a complete derivation of wave-equation operators for nonlinear modeling, linearized modeling and migration, tomographic forward and adjoint operators and wave-equation migration velocity analysis (WEMVA) operators. The derivation is done in time domain using a chain of simple linear operators. The results show that all linearizations and adjoints are correct for any media.

INTRODUCTION

Seismic modeling, imaging, analysis and inversion utilizes a form of the wave-equation. Commonly, the different operators are derived starting from some assumption about the imaging condition and/or the medium. This often results in an inaccurate linearization which has the wrong scale and units. One common issue is that most effort is put into getting the pair-operators to be correctly adjoints but not necessarily consistent with the nonlinear operator that they are derived from. Furthermore, the implementations in time domain are not accurate adjoints of the operators which results in sub-optimal convergence rates in inversion (Ji, 2009). Finally, the source injection and boundary conditions are normally ignored in all these operators resulting in further hindrance for the inversion.

I show the details of deriving wave-equation operators, which are not always obvious from the theory. The derivation is performed in time domain using a chain of simple linear operators. The adjoint of each operator can be tested independently which simplifies the problem. Moreover, the boundary conditions, data interpolation and source injection are taken care of in this derivation.

THEORY

Nonlinear Modeling

The first and most important operator to get correctly is the nonlinear modeling operator. Although I call it nonlinear modeling operator (referring to its relationship with medium parameters), this operator is in fact linear with the source function, i.e. the right-hand side of the wave equation. Once I get the operator and its adjoint correctly, I can easily implement simple modifications to the source function to create the other operators. First, I write the forward modeling in compact notation as:

$$\mathbf{d} = \mathbf{F}\mathbf{f}, \tag{1}$$

where $\mathbf{d}$ is the modeled data, $\mathbf{F}$ is the modeling operator, and $\mathbf{f}$ is the source function. For simplicity, I will take care of spatial and temporal difference between the source, data and propagation grid outside the operator $\mathbf{F}$. So, I can rewrite the forward modeling of one experiment as:

$$\mathbf{d} = \mathbf{K_r'L_r'FL_sK_sf}, \tag{2}$$

where $\mathbf{K}$ is a zero-padding operator from the source/receiver spatial grid to the propagation spatial grid and $\mathbf{L}$ is an interpolation operator from the source/receiver temporal grid to the propagation temporal grid. The subscripts $_s$ and $_r$ denotes the source and receiver sides, respectively, and the $'$ denotes the adjoint. It is important to keep the adjoint process in mind when designing the forward operator. When performing the adjoint, we want to interpolate the data before injecting it, hence the $\mathbf{L_r'}$. One of the common mistakes is to replace the $\mathbf{L_r'}$ with a subsampling operator. This replacement either causes the adjoint to be noisy (by injecting zeros instead of interpolating the data) or an incorrect adjoint (by interpolating the data). We can see now that the input and output of the operator $\mathbf{F}$ have the size of the propagation grid, both in space and time. Therefor, the operator $\mathbf{F}$ can be derived by solving the two-way wave-equation:

$$\left(\nabla^2 - \mathbf{v}^{-2}\frac{\partial^2}{\partial t^2}\right)\mathbf{p}(t) = \mathbf{f}(t), \tag{3}$$

where t is time, $\mathbf{v}$ is the velocity model, $\mathbf{p}$ is the propagated wavefield, and $\mathbf{f}$ is the source wavefield. By using finite-difference for the time derivative, the propagation can be rewritten as:

$$\nabla^2\mathbf{p}(it) - \frac{\mathbf{v}^{-2}}{\Delta t^2}\left(\mathbf{p}(it+1) - 2\mathbf{p}(it) + \mathbf{p}(it-1)\right) = \mathbf{f}(it). \tag{4}$$

I then rearrange the previous equation to have a recursive time-stepping scheme as follows:

$$\mathbf{p}(it+1) = \mathbf{v}^2\Delta t^2\left(\nabla^2\mathbf{p}(it) - \mathbf{f}(it)\right) + 2\mathbf{p}(it) - \mathbf{p}(it-1). \tag{5}$$

One common mistake is to inject the source before applying the Laplacian, which results in solving the equation for another "effective" source. Another common mistake is to not scale the source by the proper coefficient. Both this mistakes affect the linearity between this operator and the following operators.

I will now do two steps to simplify the recursive scheme into a lower triangular matrix and to simplify its recursive adjoint. First, I will shift the time sample of the propagated wavefield by one to use the computed one as reference. Second, I will take the different times of p as common factors to clarify the coefficients of the recursive operator. These changes result in the following equation:

$$\mathbf{p}(it) = -\mathbf{v}^2\Delta t^2\mathbf{f}(it) + \left(\mathbf{v}^2\Delta t^2\nabla^2 + 2\right)\mathbf{p}(it-1) - \mathbf{p}(it-2). \tag{6}$$

I can now define the coefficients as the following operators:

$$\begin{aligned}\mathbf{C}_1 &= -\mathbf{v}^2\Delta t^2\\ \mathbf{C}_2 &= \mathbf{v}^2\Delta t^2\nabla^2 + 2\mathbf{I}\\ \mathbf{C}_3 &= -\mathbf{I},\end{aligned} \tag{7}$$

where $\mathbf{I}$ is identity operator. Equation 6 can be rewritten as:

$$\mathbf{p}(it) = \mathbf{C}_1\mathbf{f}(it) + \mathbf{C}_2\mathbf{p}(it-1) + \mathbf{C}_3\mathbf{p}(it-2). \tag{8}$$

Although the previous equation describes the forward time-stepping correctly, its adjoint is not a recursive operator. The reason for this problem is the $\mathbf{C}_1$ operator does not commute with the other coefficient operators. To avoid this problem, I need to pull the coefficient $\mathbf{C}_1$ outside the recursion operator and into a separate operator, which I am going to call $\mathbf{S}$. In other words, I am going to break the operator $\mathbf{F}$ into two operators: a scaling operator and a recursive operator. The scaling operator $\mathbf{S}$ is simply multiple copies of $\mathbf{C}_1$ to multiply all time steps. This changes equation 2 into:

$$\mathbf{d} = \mathbf{K_r'L_r'GSL_sK_sf}, \tag{9}$$

whereas the operator $\mathbf{G}$ is the recursive operator described by:

$$\mathbf{p}(it) = \mathbf{q}(it) + \mathbf{C}_2\mathbf{p}(it-1) + \mathbf{C}_3\mathbf{p}(it-2), \tag{10}$$

where $\mathbf{q}$ is the source wavefield after applying the operator $\mathbf{S}$. The operators $\mathbf{S}$ and $\mathbf{L}_s$ commute because the interpolation is on the time axis only. Therefore, I can rewrite equation 9 as:

$$\mathbf{d} = \mathbf{K_r'L_r'GL_sSK_sf}, \tag{11}$$

which is more computationally efficient. I can now write the adjoint modeling as:

$$\mathbf{f} = \mathbf{K_s'S'L_s'G'L_rK_rd}. \tag{12}$$

The operator $\mathbf{G}'$ can be described by:

$$\mathbf{q}(it) = \mathbf{p}(it) + \mathbf{C}_2'\mathbf{q}(it+1) + \mathbf{C}_3'\mathbf{q}(it+2), \tag{13}$$

and the operator $\mathbf{C}_2'$ is:

$$\mathbf{C}_2' = (\mathbf{v}^2\Delta t^2\nabla^2 + 2\mathbf{I})' = (\nabla^2)'\mathbf{v}^2\Delta t^2 + 2\mathbf{I}. \tag{14}$$

equations 12 to 14 show why doing the forward propagation with the time axis reversed does not give the proper adjoint due to three steps. First, applying operator $\mathbf{S}'$ should take place after $\mathbf{G}'$. Second, the Laplacian operator in $\mathbf{C}_2'$ comes after scaling by $\mathbf{v}^2\Delta t^2$. Finally, the Laplacian operator is not self-adjoint.

The last step is to add the boundary condition. In my derivation, I will use an absorption boundary that removes waves coming back from the boundaries as described by Israeli and Orszag (1981). For this boundary condition, I need to define the quantity $\mathbf{w}$ which will have a value of 0 where there is no absorbing and linearly increase across the absorbing layer to a maximum value of $0.15\frac{\Delta t}{\Delta s}\mathbf{v}$ at the outer edges where Δs is the spatial sampling in the direction of absorbing. Now, I can apply the absorbing condition after propagating each time step as:

$$\mathbf{p}(it) = \mathbf{p}'(it) + \mathbf{w}\left(\mathbf{p}(it-1) - \mathbf{p}(it) - \mathbf{v}\Delta t\frac{\partial}{\partial s}\mathbf{p}(it-1)\right), \tag{15}$$

where $\frac{\partial}{\partial s}$ is the spatial first derivative in the direction of absorbing and $\mathbf{p}'(it)$ is the time slice before applying the boundary condition. I can arrange the previous equation by taking the time slices of $\mathbf{p}$ as common factors:

$$\mathbf{p}(it) = (1-\mathbf{w})\,\mathbf{p}'(it) + \mathbf{w}\left(1 - \mathbf{v}\Delta t\frac{\partial}{\partial s}\right)\mathbf{p}(it-1). \tag{16}$$

Next, I substitute the value of $\mathbf{p}'(it)$ from equation 8 into the previous equation:

$$\mathbf{p}(it) = (1-\mathbf{w})\left(\mathbf{C}_1\mathbf{f}(it) + \mathbf{C}_2\mathbf{p}(it-1) + \mathbf{C}_3\mathbf{p}(it-2)\right) + \mathbf{w}\left(1 - \mathbf{v}\Delta t\frac{\partial}{\partial s}\right)\mathbf{p}(it-1). \quad (17)$$

Again, I rearrange the equation and take the time slices of $\mathbf{p}$ and $\mathbf{f}$ as common factors to get:

$$\mathbf{p}(it) = \mathbf{C}_4\mathbf{f}(it) + \mathbf{C}_5\mathbf{p}(it-1) + \mathbf{C}_6\mathbf{p}(it-2), \quad (18)$$

where the new coefficients are:

$$\begin{aligned}
\mathbf{C}_4 &= (\mathbf{w}-\mathbf{I})\mathbf{v}^2\Delta t^2 \\
\mathbf{C}_5 &= (\mathbf{I}-\mathbf{w})\left(\mathbf{v}^2\Delta t^2\nabla^2 + 2\mathbf{I}\right) + \mathbf{w}\left(\mathbf{I} - \mathbf{v}\Delta t\frac{\partial}{\partial s}\right) \\
\mathbf{C}_6 &= \mathbf{w} - \mathbf{I}.
\end{aligned} \quad (19)$$

All I have to do now is to redefine the scaling operator $\mathbf{S}$ to be multiple copies of $\mathbf{C}_4$ instead of $\mathbf{C}_1$ and the operator $\mathbf{G}$ to have the coefficients $\mathbf{C}_5$ and $\mathbf{C}_6$ instead of $\mathbf{C}_2$ and $\mathbf{C}_3$, respectively. Also, operator $\mathbf{G}'$ requires the proper adjoint of $\mathbf{C}_5$ which can be written as:

$$\mathbf{C}_5' = \left((\nabla^2)'\mathbf{v}^2\Delta t^2 + 2\mathbf{I}\right)(\mathbf{I}-\mathbf{w}) + \left(\mathbf{I} - \left(\frac{\partial}{\partial s}\right)'\mathbf{v}\Delta t\right)\mathbf{w}. \quad (20)$$

One possible adjustment to the forward operator is to use a normalized interpolation operators. The reason for the normalization is that we do not want the amplitudes to change for the same model and wavelet if we model with finer propagation time steps. The normalization is easy in this case since the ratio between the wavelet time sampling and the propagation time sampling is a constant.

Born Operator

The Born operator relates perturbation in the velocity (i.e. the image), $\mathbf{r}$, to changes in the data. In the theory, we see that the Born operator is simply the solution to the two-way wave-equation but with a virtual source. The virtual source is the background wavefield $\mathbf{u_{s0}}$ after scattering by the perturbation $\mathbf{r}$ and a second derivative in time. Therefore, I first need to compute the background wavefield similarly to how I computed the data in equation 11 but without the last truncation step:

$$\mathbf{u_{s0}} = \mathbf{L_s'GL_sSK_sf}, \quad (21)$$

Notice that the operator $\mathbf{G}$ is now followed by the operator $\mathbf{L_s'}$ instead of $\mathbf{L_r'}$ since I am using this wavefield as a secondary source instead of recording it as data. It is important to keep the operator $\mathbf{L_s'}$ for two reasons. First, to have proper adjoint (similar explanation to the nonlinear modeling operator). Second, to reduce the number of times we do the expensive scattering/imaging step. Once $\mathbf{u_{s0}}$ is computed, I construct the operator $\mathbf{U_{s0}}$ which has a model space of the image and a data space of scattered wavefield. In other words, this operator is the adjoint of imaging condition. With a second time derivative operator $\mathbf{T}$, I can write the forward Born modeling operator as:

$$\mathbf{\Delta d} = \mathbf{K_r'L_r'GL_sST}\widehat{\mathbf{U_{s0}}}\mathbf{r}, \quad (22)$$

where the hat $\widehat{\ }$ indicates an extended image space with either offset or time lags. The adjoint of Born operator, i.e. the RTM operator, can be written as:

$$\mathbf{r} = \widehat{\mathbf{U}'_{\mathbf{s0}}}\mathbf{T}'\mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{r}}\mathbf{K}_{\mathbf{r}}\mathbf{\Delta d}. \tag{23}$$

Tomographic Operator

The tomographic operator relates changes in the propagation velocity, $\mathbf{b}$, to changes in the data using the perturbation (or the image) $\mathbf{r}$. Similar to the Born operator, the tomographic operator is also a solution to the two-way wave-equation but with two virtual sources: one source is the background wavefield $\mathbf{u_{s0}}$ after scattering by the propagation-velocity perturbation, $\mathbf{\Delta b}$, and a second derivative in time and the second source is the perturbed wavefield $\mathbf{u_{\Delta s}}$ after scattering by the propagation-velocity perturbation $\mathbf{\Delta b}$ and a second derivative in time as well. The perturbed wavefield $\mathbf{u_{\Delta s}}$ can be computed as:

$$\mathbf{u_{\Delta s}} = \mathbf{L}'_{\mathbf{s}}\mathbf{G}\mathbf{L}_{\mathbf{s}}\mathbf{S}\mathbf{T}\widehat{\mathbf{U}_{\mathbf{s0}}}\mathbf{r}. \tag{24}$$

Once $\mathbf{u_{\Delta s}}$ is computed, I construct the operator $\mathbf{U_{\Delta s}}$ the same way I constructed $\mathbf{U_{s0}}$. Moreover, I need to define an operator $\mathbf{R}$ which similar to the $\widehat{\mathbf{U}_{\mathbf{s0}}}$ but with the model space being the background (unscattered) wavefield and the perturbation $\mathbf{r}$ being in the operator itself instead of being in the model space. Now, I can write the forward tomographic operator as:

$$\mathbf{\Delta d} = \mathbf{K}'_{\mathbf{r}}\mathbf{L}'_{\mathbf{r}}\mathbf{G}\mathbf{L}_{\mathbf{s}}\mathbf{S}\mathbf{T}\mathbf{U}_{\mathbf{\Delta s}}\mathbf{\Delta b} + \mathbf{K}'_{\mathbf{r}}\mathbf{L}'_{\mathbf{r}}\mathbf{G}\mathbf{L}_{\mathbf{s}}\mathbf{S}\mathbf{T}\mathbf{R}\mathbf{L}'_{\mathbf{s}}\mathbf{G}\mathbf{L}_{\mathbf{s}}\mathbf{S}\mathbf{T}\mathbf{U}_{\mathbf{s0}}\mathbf{\Delta b}. \tag{25}$$

The adjoint of the tomographic operator can be written as:

$$\mathbf{\Delta b} = \mathbf{U}'_{\mathbf{\Delta s}}\mathbf{T}'\mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{r}}\mathbf{K}_{\mathbf{r}}\mathbf{\Delta d} + \mathbf{U}'_{\mathbf{s0}}\mathbf{T}'\mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{s}}\mathbf{R}'\mathbf{T}'\mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{r}}\mathbf{K}_{\mathbf{r}}\mathbf{\Delta d}. \tag{26}$$

WEMVA Operator

The WEMVA operator relates changes in the propagation velocity, $\mathbf{b}$, to changes in the perturbation, $\mathbf{r}$, using the data. The WEMVA operator is slightly different than the tomographic operators because both its model space and data space are in the "image" domain (either $\mathbf{b}$ or $\mathbf{r}$) and both the source function and the data are part of the operator itself. However, as the theory shows, the adjoint WEMVA operator has the same expression as the adjoint of the tomographic operator after switching the model space with some components of the operator. Hence, I need to compute the receiver-side background wavefield as:

$$\mathbf{u_{r0}} = \mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{r}}\mathbf{K}_{\mathbf{r}}\mathbf{\Delta d}. \tag{27}$$

However, there is an issue on the receiver-side because the order of operators between the scattering/image and second time derivative will not be the same as the tomographic operator. To resolve this issue, I can simply include the time derivative in the computation of the receiver-side background wavefield and redefine it as:

$$\mathbf{u_{r0}} = \mathbf{T}'\mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{r}}\mathbf{K}_{\mathbf{r}}\mathbf{\Delta d}. \tag{28}$$

Then, I construct the operator $\mathbf{U_{r0}}$ the same way I constructed the operator $\mathbf{U_{s0}}$ but using the receiver-side background wavefield. I can now define the adjoint WEMVA operator as:

$$\mathbf{\Delta b} = \mathbf{U'_{r0}L'_sGL_sST}\widehat{\mathbf{U_{s0}}}\mathbf{\Delta r} + \mathbf{U'_{s0}T'S'L'_sG'L_s}\widehat{\mathbf{U_{r0}}}\mathbf{\Delta r}, \tag{29}$$

and the forward WEMVA operator as:

$$\mathbf{\Delta r} = \widehat{\mathbf{U'_{r0}}}\mathbf{L'_sGL_sSTU_{s0}\Delta b} + \widehat{\mathbf{U'_{s0}}}\mathbf{T'S'L'_sG'L_sU_{r0}\Delta b}. \tag{30}$$

SYNTHETIC EXAMPLE

A simple synthetic model with a Gaussian anomaly (representing $\mathbf{\Delta b}$) and a spike (representing $\mathbf{r}$) is used for the synthetic examples. The true velocity is shown in Figure 1. The background velocity is 3 km/s and the Gaussian anomaly is located at 1 km in z-axis and x-axis. The perturbation spike is located at 2 km in z-axis and x-axis. Both anomalies have an amplitude of 300 m/s. The sampling for both spatial axes is 10 m. A Ricker wavelet with a fundamental frequency of 15 Hz is used to model the data. The source spacing is 50 m and the receiver spacing is 10 m.

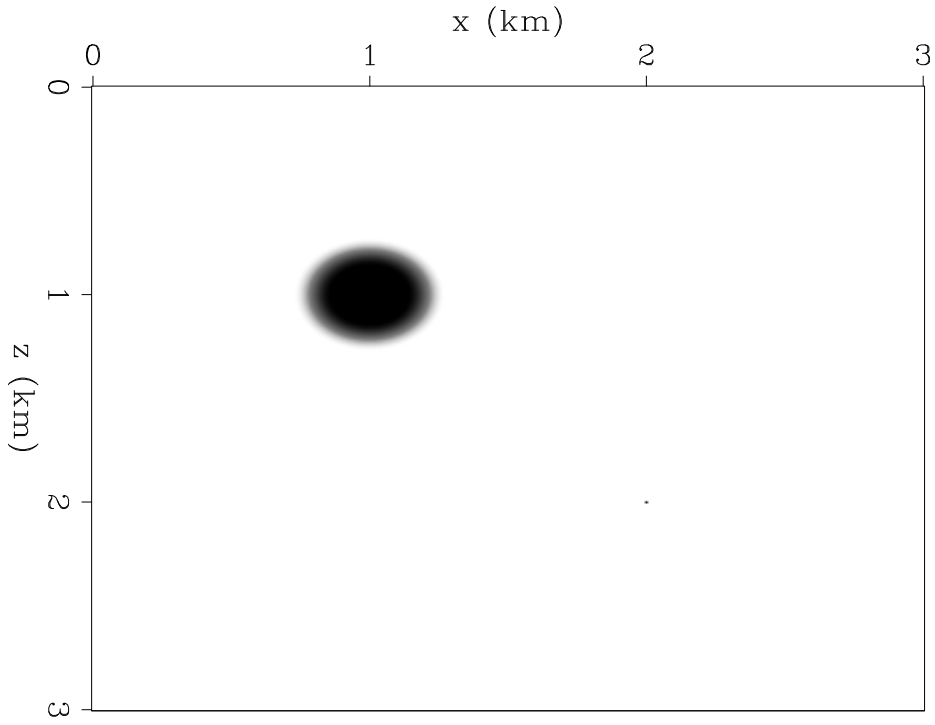


Figure 1: The true velocity model. [**ER**] ali2/. acc-v

To test the RTM operator, I start from a propagation velocity model that is the same as the true model except for the perturbation spike. The result of applying the RTM operator is shown in Figure 2. Next, I test the tomographic operator by computing the data residual due to removing the Gaussian anomaly. I use the RTM results as perturbation to estimate the update of the propagation velocity. The results of using the adjoint tomographic operator is shown in Figure 3. Finally, I test the WEMVA operator by computing the image residual due to removing the Gaussian anomaly. The results of using the adjoint WEMVA operator is shown in Figure 4. Both operators have the correct direction for the anomaly.

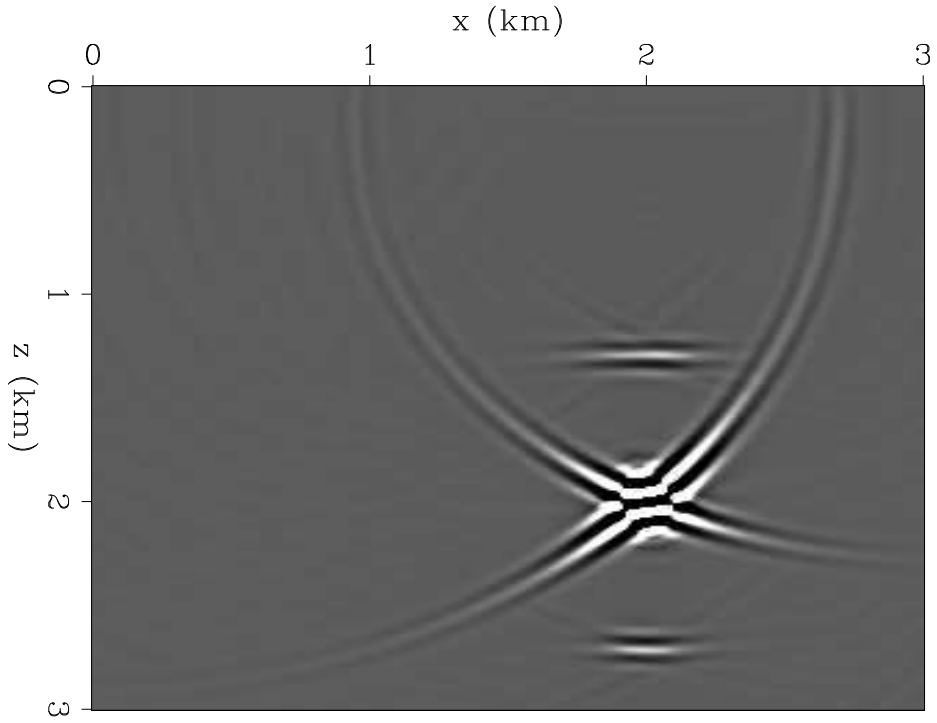


Figure 2: The RTM image. [**ER**] ali2/. acc-m

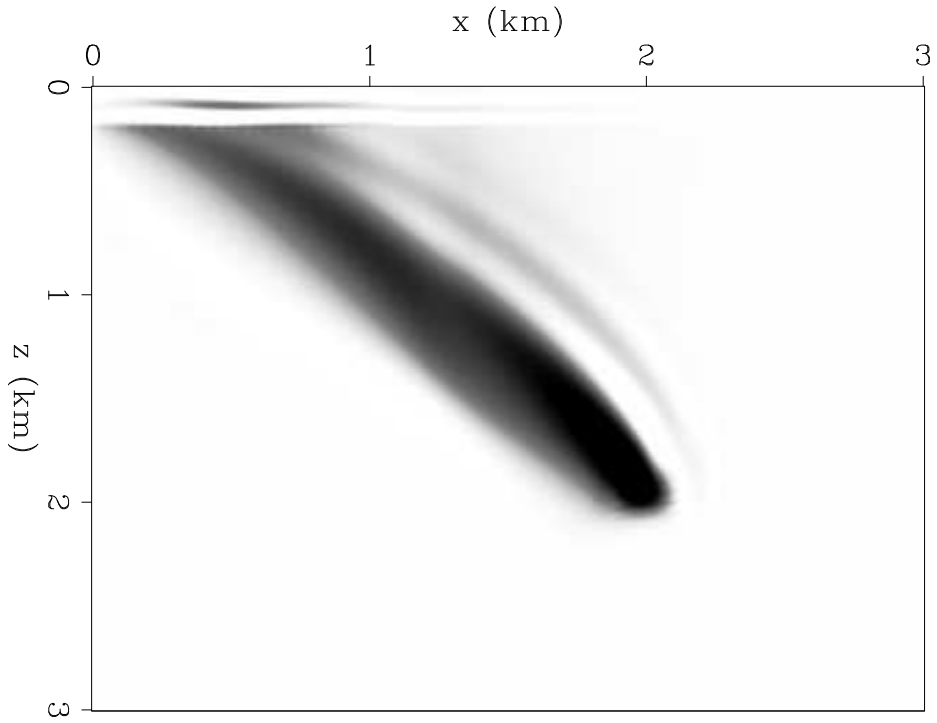


Figure 3: The tomographic adjoint on the data residual. [**ER**] ali2/. acc-gbm

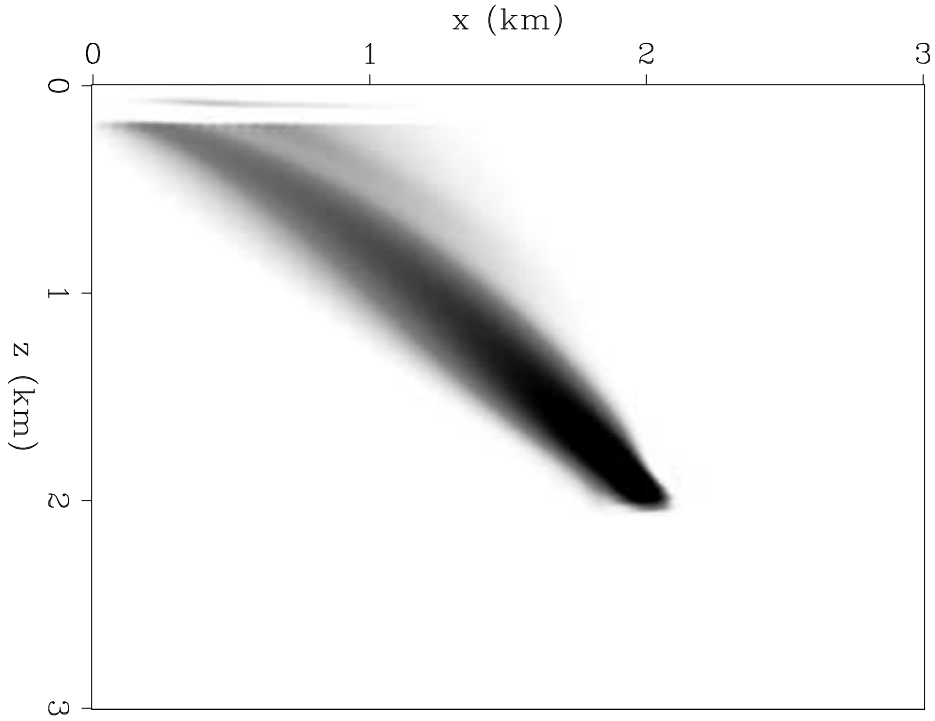


Figure 4: The WEMVA adjoint on the image residual. [**ER**] ali2/. acc-gbw

CONCLUSIONS

I provide the details of deriving two-way wave-equation operators while taking care of practical issues such as boundary conditions and source injection. The results of all the operators are consistent in terms unit and scale including the original wave-equation. Getting consistent operators is important when trying to use joint inversion without manually or empirically guessing a ratio between the outputs of different operators. Also, getting correct linearizations and adjoint can help the convergence rate of the inversion.

REFERENCES

Israeli, M. and S. A. Orszag, 1981, Approximation of Radiation Boundary Conditions: Journal of Computational Physics, **41**, 115–135.

Ji, J., 2009, An exact adjoint operation pair in time extrapolation and its application in least-squares reverse-time migration: Geophysics, **74**, H27–H33.

Stochastic rock physics modeling for seismic anisotropy

Yunyue (Elita) Li, Biondo Biondi, Dave Nichols, Gary Mavko, and Robert Clapp

ABSTRACT

Anisotropic model building using surface seismic data is a well-known underdetermined and nonlinear problem. To stabilize the inversion, a regularization term is often added into the data fitting objective function assuming a priori anisotropic model. In this paper, we build the anisotropic prior model using stochastic rock physics modeling for shale anisotropy. We consider shale anisotropy from four aspects: mineral elastic anisotropy for the constitutes of the rock, the mineral transition due to compaction and temperature, particle alignment during compaction and shale/sand lamination. The input parameters for the rock physics modeling are provided by two different sources: quantitative inversion results from seismic impedance in three dimensional space and well-log measurements at isolated well locations. For each instance of modeling, we perturb the key parameters for the rock physics modeling to produce a set of random models. The modeling results are compared at three different depth levels in a statistical manner. The similarity between both modeling results justifies the use of the seismic inversion results in the deeper region that the well logs do not cover. To better utilize the seismic and the rock physics information, we also propose a new parameterization scheme for wave-equation migration velocity analysis.

INTRODUCTION

Anisotropic model building tries to resolve more than one parameter at each grid point of the subsurface. This number could be three for a vertical transverse isotropic (VTI) media, and increases to five for a tilted transverse isotropic (TTI) media. Any inversion scheme based on surface seismic data only becomes ill-posed and highly underdeterimined due to the rapidly increasing model space with the increasing complexity of the subsurface.

One big disadvantage of the surface seismic inversion is its lack of depth information. To add the depth dimension into the inversion, several localized tomography studies around the wells have been performed (Bakulin et al., 2010d,c). In these studies, joint inversion of surface seismic data and borehole data (check-shots, walkaway VSPs) showed great potentials to yield better defined Earth models. However, due to the ambiguity between the parameters, even the borehole aided localized tomography has difficulties resolving a reliable, unique anisotropic model in 3-D (Bakulin et al., 2009).

To further constrain the inversion, we need to consider some prior knowledge of the subsurface. This prior knowledge describes the covariance of the model space, and is independent of the data. There are many ways to obtain the covariance information, based on different assumptions. For example, we often smooth our Earth model horizontally and vertically, which implies a certain user-defined spatial correlation lag. More realistically, we can use the geological information as a priori and shape our estimation accordingly. This

model shaping can be posed as a decomposition of the Earth model into different layers and horizons before tomography (Bakulin et al., 2010a), or as a regularization/preconditioning operator during tomography (Bakulin et al., 2010b). We can obtain the geological information either by interpreting and picking the horizons or by building a set of steering filters (Clapp, 2000) according to the current subsurface image.

In addition to the spatial covariance, for a multi-parameter estimation, a local cross-parameter covariance is also needed to better describe the subsurface. One source of the cross-parameter covariance comes from rock physics studies (Hornby et al., 1995; Sayers, 2004, 2010; Bachrach, 2010b). Many authors (Dræge et al., 2006; Bandyopadhyay, 2009; Bachrach, 2010a) have built averaged depth trends serving for seismic processing. In particular, Bachrach (2010a) develops both deterministic and stochastic modeling schemes based on the rock physics effective media models for compacting shale and sandy shale. By making core measurements, we can limit the parameters needed by the rock physics model in a certain range, which greatly reduces the span of the parameter space defined by rock physics modeling. These rock physics modeling results can be used to construct the initial Earth model and the covariance relationships among the Earth model parameters. Li et al. (2011) and Yang et al. (2012) have demonstrated that the rock physics prior models are helpful in constraining ray tomography.

In this paper, we first analyze the topography of the wave-equation migration velocity analysis objective function with respect to different parameterization schemes. To best utilize both surface seismic and rock physics information, we propose a new parameterization scheme for wave-equation migration velocity analysis. For the rock physics modeling, we combine the rock physics models proposed by Bachrach (2010a) and Bandyopadhyay (2009). We model the shale anisotropy from four aspects: mineral anisotropy of the constitutes of the rock, compaction effect on the particle alignment, the transition from smectite to illite due to compaction and temperature, and the lamination effect of isotropic sand layers and anisotropic shale layers. Depth trends (ideally, spatially varying) for the anisotropic parameters from rock physics modeling are obtained from two different sources: seismic inversion results and the well log measurements. We then compare the depth trends in a stochastic manner and the similarity between both depth trends justifies the use of the seismic inversion results in the deeper region that the well log does not cover.

ROCK PHYSICS CONSTRAINED WAVE-EQUATION MIGRATION VELOCITY ANALYSIS FOR ANISOTROPY

Anisotropic wave-equation migration velocity analysis (WEMVA) aims at building an anisotropic Earth model that minimizes the residual image from the surface seismic data (Li and Biondi, 2011). This optimization problem is highly non-linear and underdetermined. Therefore, we commonly add a regularization term to the anisotropic WEMVA objective function to constrain the null space and stabilize the inversion. The resulting objective function is as follows:

$$S(\mathbf{m}) = \frac{1}{2}||\mathbf{D}_\theta \mathbf{I}(\mathbf{x},\theta)|| + \lambda \frac{1}{2}(\mathbf{m}-\mathbf{m}_{\text{prior}})^T \mathbf{C}_M^{-1}(\mathbf{m}-\mathbf{m}_{\text{prior}}), \tag{1}$$

where the first term is the "data fitting" term, and the second "model regularization" term. In the data fitting term, $\mathbf{m}$ is the anisotropic subsurface model, $\mathbf{I}(\mathbf{x},\theta)$ is the migration image in the angle domain with θ the aperture angle, $\mathbf{D}_\theta$ is a derivative operator along the angle

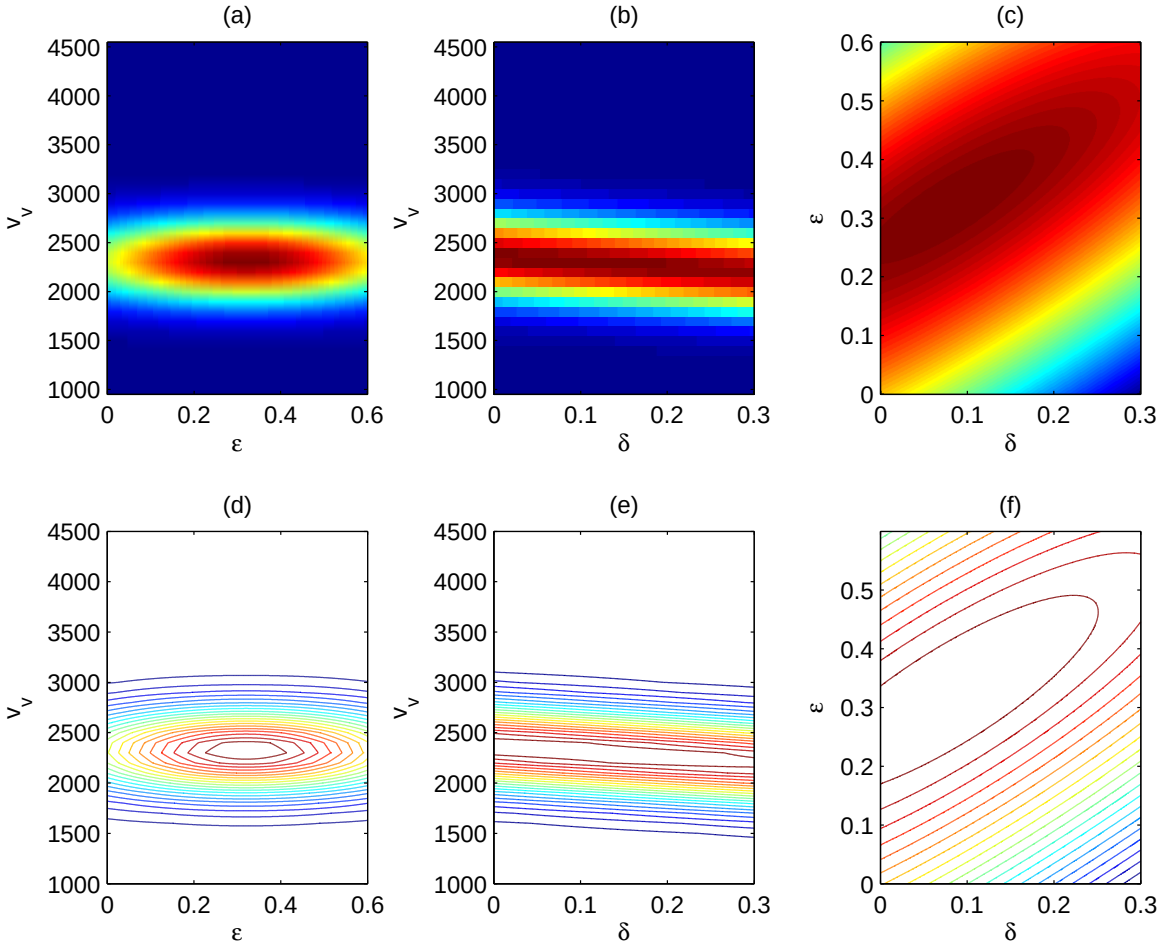


Figure 1: A sketch of the topography of the DSO objective function projected on to (a) v_v-ϵ plane, (b) v_v-δ plane, and (c) ϵ-δ plane. Panels (d), (e), and (f) are the contoured representation of (a), (b), and (c), respectively. [**ER**] elita1/. data-vvepsdel

axis. In the model regularization term, $\mathbf{m}_{\text{prior}}$ and $\mathbf{C}_M$ defines a Gaussian distribution of a prior model, which is ideally independent of the seismic data so that this regularization will bring more information into the optimization. Finally, parameter λ balances the weights between both terms and reflects the confidence of the data compared with the prior model.

The data fitting term, known as differential semblance optimization (DSO) objective function relates the incoherence in the angle domain common image gathers to the inaccuracy in the subsurface models. It has been shown that the traveltimes of qP-waves in TI medium mainly depend on the anellipticity parameter η and the (zero-dip) normal moveout velocity v_n (Alkhalifah and Tsvankin, 1995). In general, the DSO objective function has a much stronger control in velocity (v_v or v_n) than in anisotropic parameters (ϵ and δ). Therefore, we can sketch out the topography of the DSO objective function schematically in figure 1. Notice on the ϵ-δ panel, the better resolved direction is the η direction, more or less parallel to $\epsilon - \delta$ when δ is small. In the unresolved direction, the topography of the DSO objective function is flat.

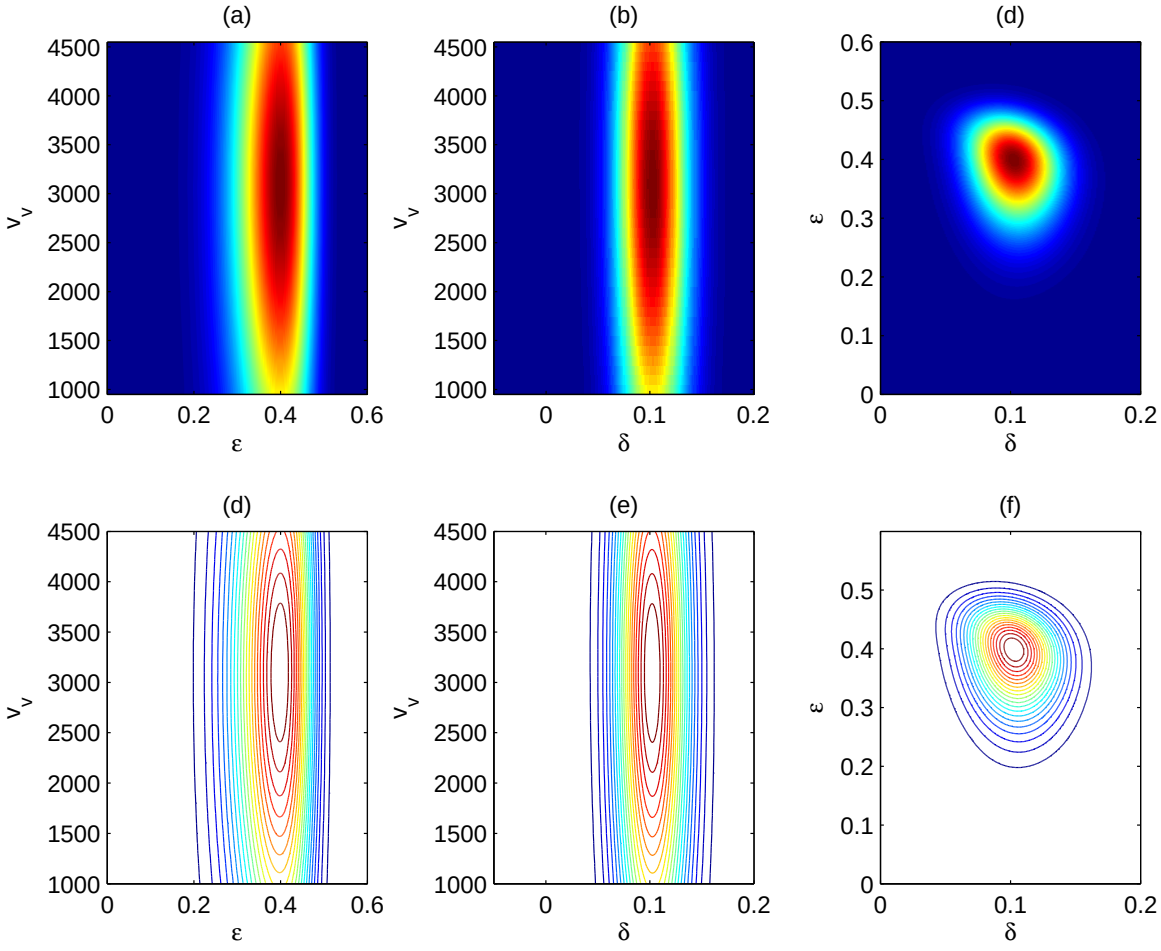


Figure 2: Stocastic rock physics modeling results for parameters v_v, ϵ and δ. Inputs of the rock physics modeling are from seismic inversion results. Topography of the model fitting objective function is projected on to (a) v_v-ϵ plane, (b) v_v-δ plane, and (c) ϵ-δ plane. Panels (d), (e), and (f) are the contoured representation of (a), (b), and (c), respectively. [**ER**] elita1/. model-vvepsdel

On the other hand, the topography of the model regularization objective function can be estimated by stochastic rock physics modeling. We will discuss the process of the rock physics modeling in detail in the next section. The resulting topography of the model fitting objective function is shown in figure 2. Compared with the data objective function, the model objective function has much looser control in velocity than in anisotropic parameters. For the Thomsen parameters ϵ and δ, rock physics modeling gives a much tighter distribution.

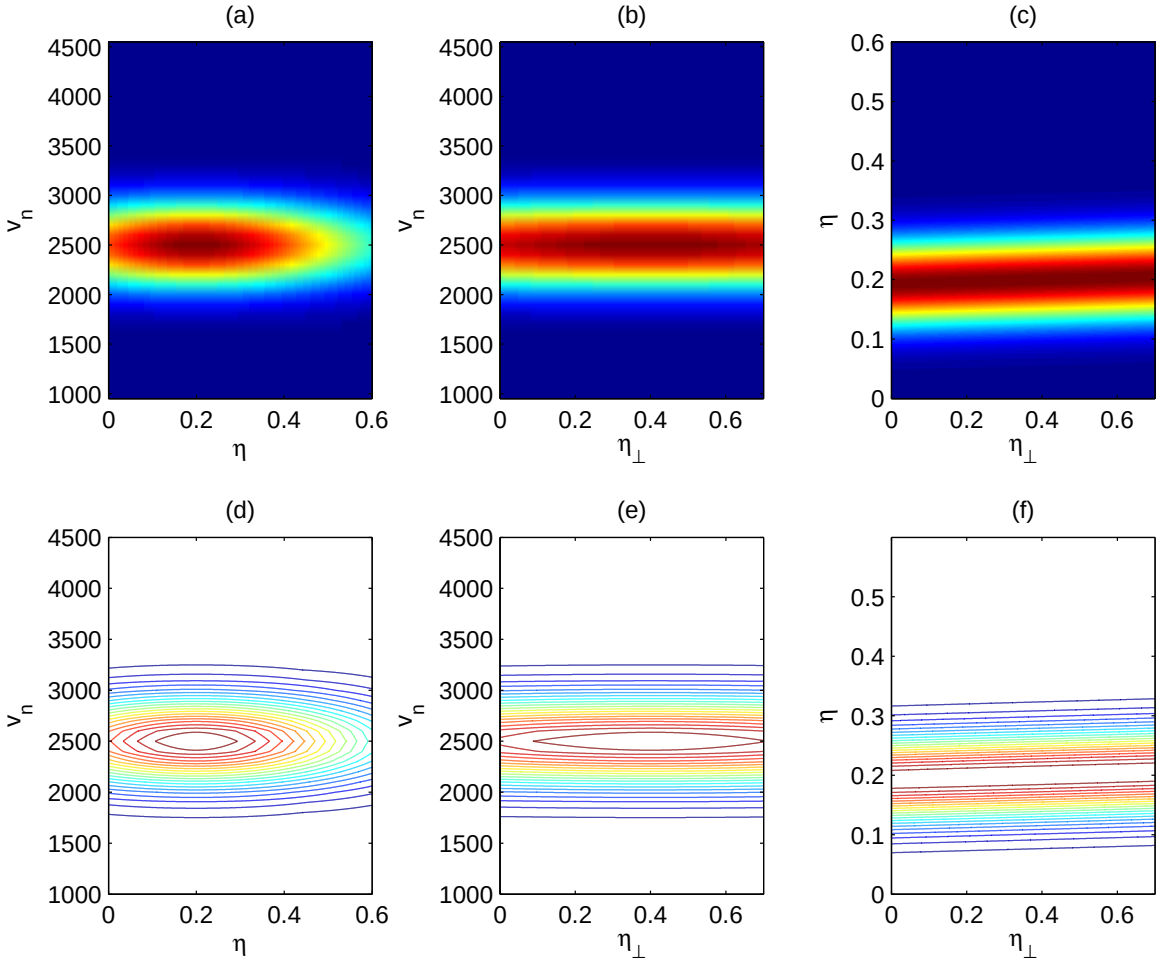


Figure 3: A sketch of the distribution function for the DSO objective function projected on to (a) v_n-η plane, (b) v_n-$\eta_\perp$ plane, and (c) η-$\eta_\perp$ plane. Panels (d), (e), and (f) are the contoured representation of (a), (b), and (c), respectively. [**ER**] elita1/. data-vnetauneta

It is important to note that not only the shape of the distribution is different for both objective functions, the centers of the distribution also differ from each other. When forming a composite objective function by adding the two objectives, both the descending path for the inversion and the final solution to the optimization will be changed. To better utilize both seismic and rock physics information, we should separate the better constrained parameters from the unconstrained parameters, with the emphasis on fitting the seismic data. Therefore, we introduce a new parameter $\eta_\perp$ defined as follows:

$$\eta_\perp = \epsilon + \delta. \tag{2}$$

The direction of $\eta_\perp$ denotes the least constrained direction by the surface P-wave seismic data. We now transform both the topography of the objective functions to the v_n-η-$\eta_\perp$ domain, and the resulting distributions are shown in figures 3 and 4. Notice the topographies for data and model objectives are (nearly) perpendicular to each other. A parameter dependant λ can be used to add appropriate amount of rock physics information to constrain the DSO objective function.

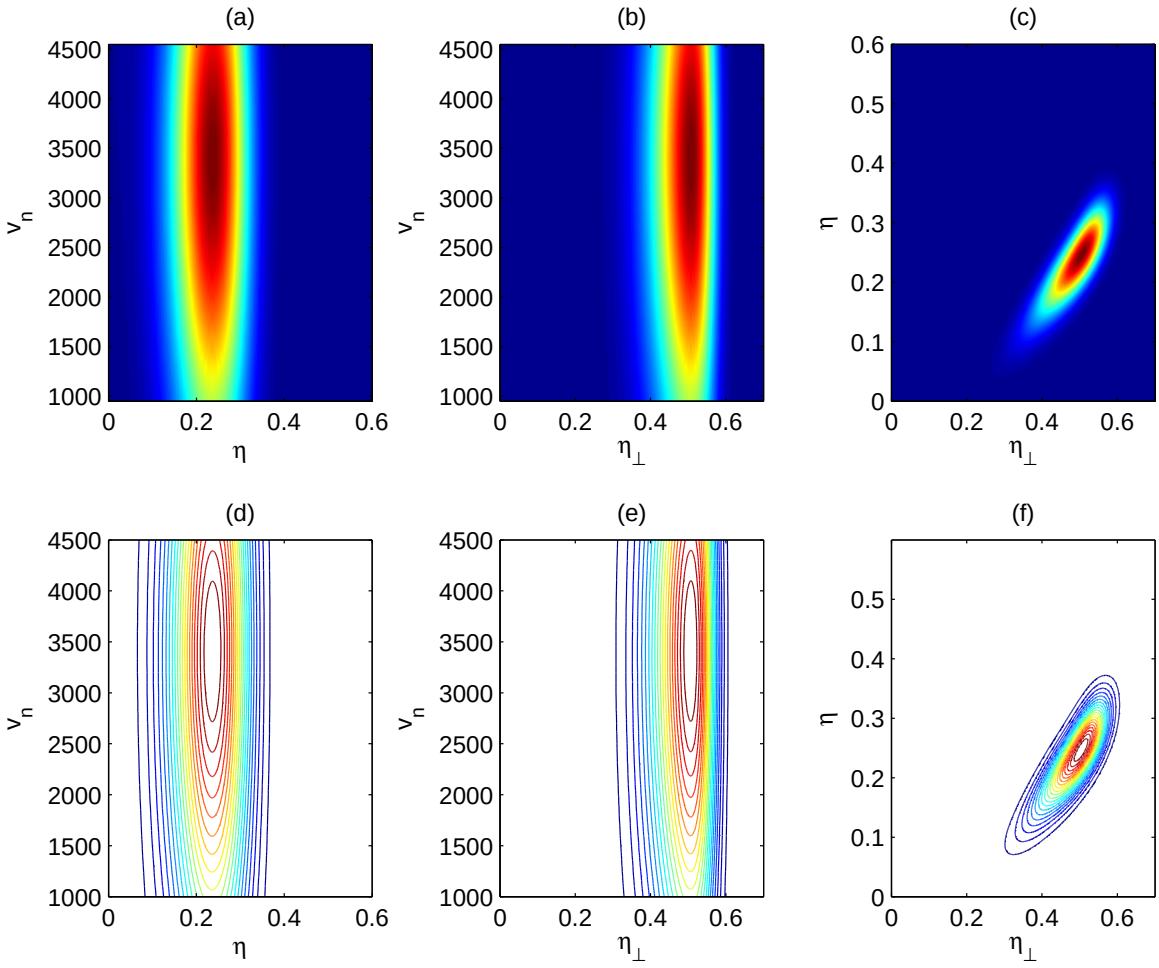


Figure 4: Stocastic rock physics modeling results for parameters v_n, η and $\eta_\perp$. Inputs of the rock physics modeling are from seismic inversion results. Topography of the model fitting objective function is projected on to (a) v_n-η plane, (b) v_n-$\eta_\perp$ plane, and (c) η-$\eta_\perp$ plane. Panels (d), (e), and (f) are the contoured representation of (a), (b), and (c), respectively. [**ER**] elita1/. model-vnetauneta

ROCK PHYSICS MODELING FOR SHALE ANISOTROPY

We combine the rock physics models proposed by Bachrach (2010a) and Bandyopadhyay (2009). We model the shale anisotropy from four aspects: mineral anisotropy of the constitutes of the rock, compaction effect on the particle alignment, the transition from smectite to illite due to compaction and temperature, and the lamination effect of isotropic sand

layers and anisotropic shale layers. Depth (ideally, spatially varying) trends of anisotropic parameters are obtained.

Mineral anisotropy

We assume shales have three end-member mineral constitutes: smectite, illite and quartz. The elastic properties are listed in table 1. The values of the smectite elasticity are in fact the anisotropic elasticity values for a Cretaceous shale (Hornby et al., 1994). These approximated values are served as an end member when pure shale is fully compacted. Anisotropic elasticity for illite (muscovite) are from work by Wenk et al. (2007). Similar to smectite, although the elasticity of quartz crystal may be anisotropic, we assume an isotropic quartz to approximate the pure sand as an end member in the rock.

Mineral	ρ (g/cc)	v_v (km/s)	v_s (km/s)	ϵ	δ	γ
Smectite	2.4	3.075	1.5	0.255	-0.05	0.48
Illite	2.4	4.94	2.6	1.02	0.	1.68
Quartz	2.65	6.0	4.0	0.	0.	0.

Table 1: Table 1:End-member mineral elastic properties.

Smectite to illite transition

The transition from smectite to illite is a common mineralogical reaction during the burial diagenesis of shales. Many studies (e.g., Hower et al. (1976)) have shown that this transition reaction is controlled by the temperature in the subsurface. In this paper, we follow the work of Bachrach (2010a) to calibrate the percentage of illite P_I to the temperature T as follows:

$$P_I(T) = 0.5 + 0.5 \tanh(\frac{T - T_t}{2\sigma_t}), \tag{3}$$

with T_t the transition temperature and σ_t the width of the transition zone. Reference values $T_t = 58^\circ C$ and $\sigma_t = 60^\circ C$ are from the work of (Freed and Peacor, 1989).

Preferred orientation distribution of clay mineral

Another important factor to the shale anisotropy is the preferred orientation of the clay minerals (Hornby et al., 1994; Sayers, 2004). At initial deposition, mineral domains are oriented in random directions. In this case, even though the individual mineral domain can be anisotropic, the effective medium with randomly oriented domains is isotropic. At the maximum compaction, all the mineral domains are fully aligned, which produces the effective medium with maximum anisotropy.

According to (Bandyopadhyay, 2009), the Voigt averaged stiffness coefficients C_{ij}^a are

$$\begin{aligned}
C_{11}^a &= L + 2M + \frac{4\sqrt{2}}{105}\pi^2(2\sqrt{5}a_3W_{200} + 3a_1W_{400}); \\
C_{33}^a &= L + 2M - \frac{16\sqrt{2}}{105}\pi^2(\sqrt{5}a_3W_{200} - 2a_1W_{400}); \\
C_{12}^a &= L - \frac{4\sqrt{2}}{315}\pi^2(2\sqrt{5}(7a_2 - a_3)W_{200} - 3a_1W_{400}); \\
C_{13}^a &= L + \frac{4\sqrt{2}}{315}\pi^2(\sqrt{5}(7a_2 - a_3)W_{200} - 12a_1W_{400}); \\
C_{44}^a &= M - \frac{2\sqrt{2}}{315}\pi^2(\sqrt{5}(7a_2 + a_3)W_{200} + 24a_1W_{400}); \\
C_{66}^a &= \frac{< C_{11} - C_{12} >}{2},
\end{aligned} \tag{4}$$

where

$$\begin{aligned}
a_1 &= C_{11} + C_{33} - 2C_{13} - 4C_{44}; \\
a_2 &= C_{11} - 3C_{12} + 2C_{13} - 2C_{44}; \\
a_3 &= 4C_{11} - 3C_{33} - C_{13} - 2C_{44}; \\
L &= \frac{1}{15}(C_{11} + C_{33} + 5C_{12} + 8C_{13} - 4C_{44}); \\
M &= \frac{1}{30}(7C_{11} + 2C_{33} - 5C_{12} - 4C_{13} + 12C_{44}),
\end{aligned} \tag{5}$$

with C_{ij} the stiffness coefficients of the individual domain.

In this paper, we use the porosity as an indicator for compaction (Bachrach, 2010a).

$$\begin{aligned}
W_{200}(\phi) &= W_{200}^{\max}(1 - \phi/\phi_0)^m, \\
W_{400}(\phi) &= W_{400}^{\max}(1 - \phi/\phi_0)^n,
\end{aligned} \tag{6}$$

with ϕ the porosity at depth and ϕ_0 the critical porosity. Choice of exponents m and n has not been well studied. We refer to Bachrach (2010a) and let both parameters vary between 0.5 and 2.

Lamination of sand and shale: Backus average

At depth, seismic wavelengths can get as large as a few hundred meters. These long-wavelength seismic waves cannot resolve individual layers, but interacts with the subsurface as a single averaged medium. Elastic properties of an effective medium composed of fine-scale laminations of sand and shale can be described by the Backus average. We assume the sand layer contains pure sand and the shale layer contains pure smectite and illite. Backus (1962) showed that the elastic constants of the effective medium can be obtained by the

elastic constants of the individual layers as follows:

$$\begin{aligned} C_{11} &= < c_{13}/c_{33} >^2 / < 1/c_{33} > - < C_{13}^2 > + < c_{11} >; \\ C_{12} &= C_{11} - < c_{11} > + < c_{12} >; \\ C_{13} &= < c_{13}/c_{33} >^2; \\ C_{33} &= < 1/c_{33} >^{-1}; \\ C_{44} &= < 1/c_{44} >^{-1}, \end{aligned} \tag{7}$$

where $< . >$ indicates the averages of the enclosed properties weighted by their volumetric proportions. The volumetric proportions for each lithological components are calculated from the shale content and the percentage of illite. These enclosed properties can be averaged values over orientation distribution functions for shales.

Workflow

Our anisotropic rock physics modeling follows the workflow described here:

- Compute the percentage of illite in the rock given a temperature model.
- Compute the average stiffness coefficients for smectite and illite, given a porosity model.
- Compute the volumetric percentage for each of the mineral phase, given a volumetric percentage of shale.
- Compute the stiffness coefficients for the laminated effective medium.

At each instance of the modeling, the key parameters: ϕ_0, exponents m and n, T_t, and σ_t are varied in a certain range. Therefore, assembly of models, instead of a single model are obtained. These models will be the source of the prior rock physics covariance.

MODELING RESULTS

In the previous section, we define the inputs of our modeling process: a temperature model, a porosity model and a shale content model. The temperature model is relatively smooth in space and is approximated using a typical temperature gradient in the Gulf of Mexico (GoM). Figure 5(c) shows the temperature profile in the shallow water GoM. For the porosity model and the shale content model, we have two different sources: seismic inversion results and the well logs.

An initial isotropic processing work flow has been applied on the field data by Schlumberger. Seismic inversion results, such as, P-wave velocity, shale content, and porosity are also provided. The validation between the porosity inversion result and the well log measurements at the well location suggests that the porosity inversion from seismic is not reliable, probably due to the lack of low frequencies in the seismic data. However, the inversion result for shale content is comparable to the well log measurements. Therefore, we use the provided shale content inversion cube (figure 5(d)) but modeled a smoothly varying porosity trend (figure 5(b)) from the P-wave velocity cube (figure 5(a)).

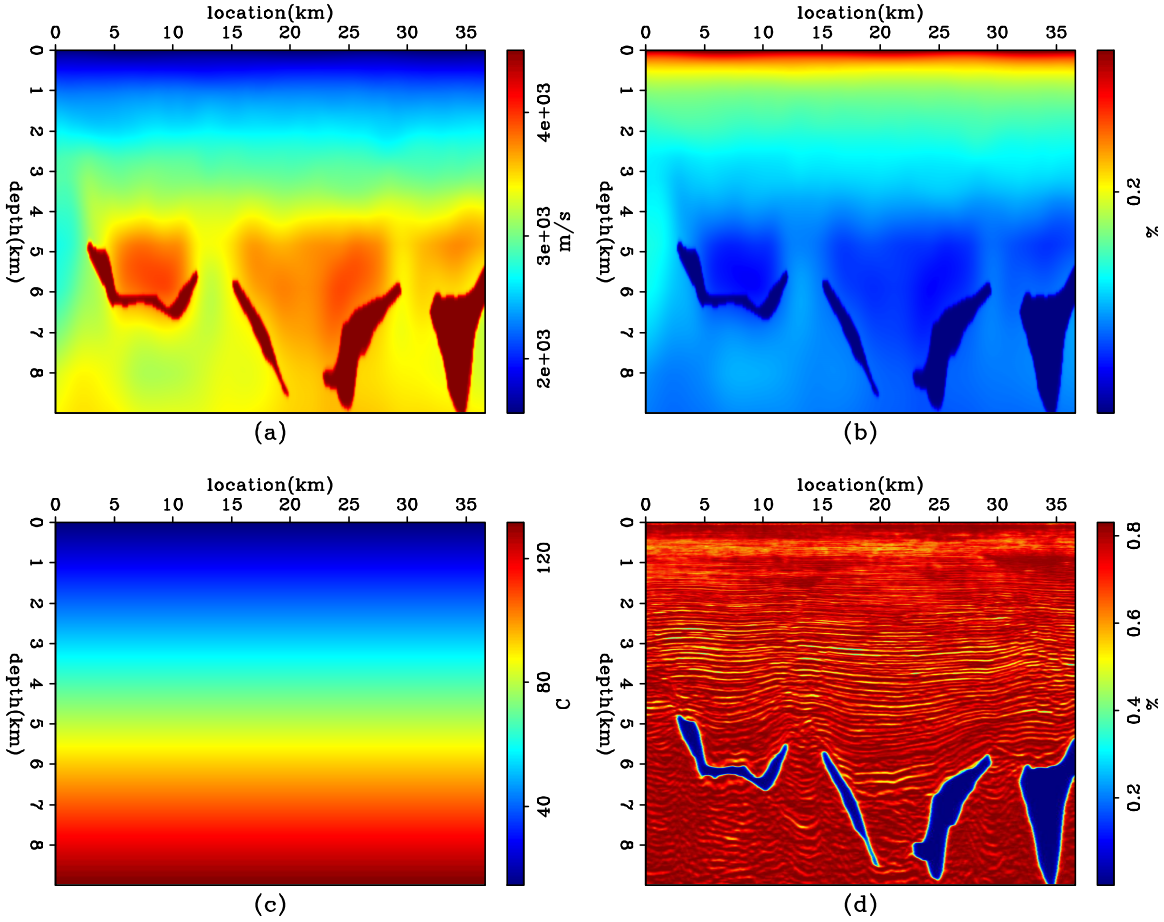


Figure 5: Input of the rock physics modeling from seismic inversion results: (a) velocity profile, (b) porosity profile, (c) temperature profile, and (d) shale content profile. [**ER**] elita1/. cubeinput

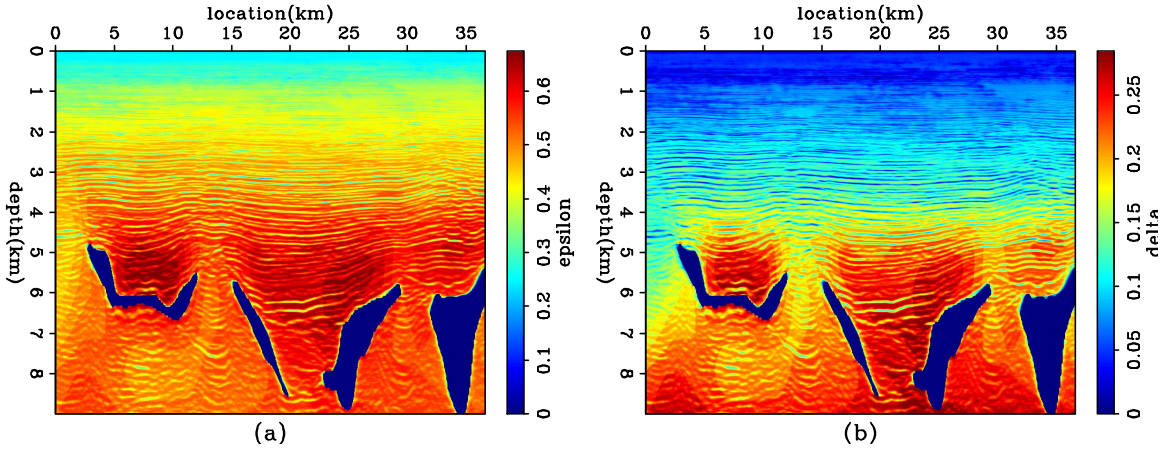


Figure 6: One realization of a 3-D rock physics modeling experiment. (a) modeling result for ϵ; (b) modeling result for δ. [**ER**] elita1/. cubemodel

Using the input in figure 5 and varying the key parameters in the rock physics modeling experiments, we can obtain multiple realizations of the anisotropic parameters in 3-D space. Figure 6 shows one realization of the stochastic modeling experiments. It is interesting to note that parameter ϵ is mainly controlled by the shale content of the subsurface, whereas parameter δ is primarily controlled by compaction and temperature of the subsurface.

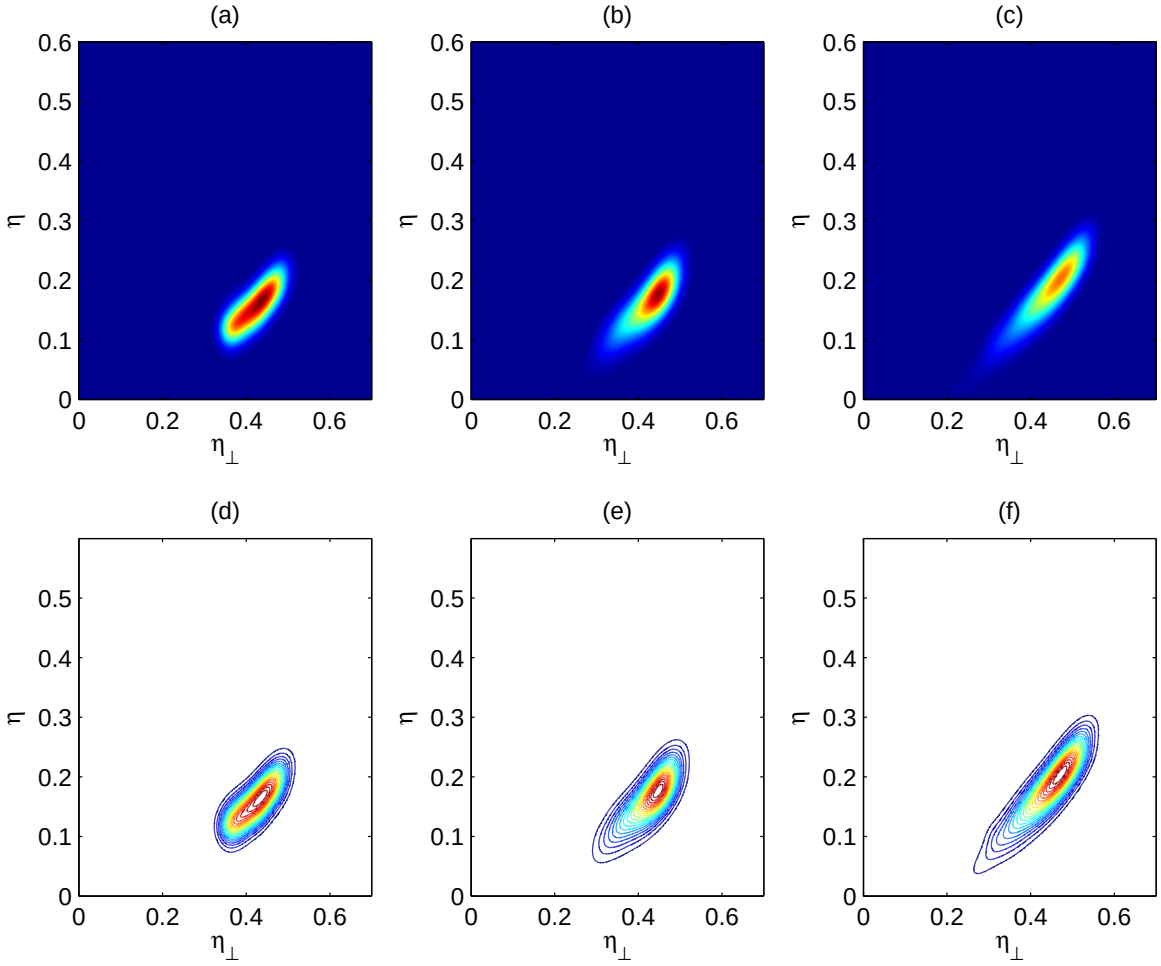


Figure 7: Stochastic rock physics modeling results for parameters v_n, η and $\eta_\perp$. Inputs of the rock physics modeling are from seismic inversion results in figure 5. [**ER**] elita1/. cube-model-prior

Figure 7 shows the topography of the model fitting objective function from the stochastic rock physics modeling using seismic inversion results. The results are shown at three different depth levels: 1 - 2 km in figure 7(a), 2 - 3.5 km in figure 7(b) and 3.5 - 5 km in figure 7(c). In general, the topography for model fitting objective function gets wider as the target gets deeper. Nevertheless, the topography from rock physics modeling is much tighter than the one from seismic data.

We can also extract the input models from the well logs. Figure 8 shows the log measurements and the deduced models for the rock physics modeling. The conversion for gamma

ray measurements G to shale content P_{sh} is a simple linear stretch as follows:

$$P_{sh} = \frac{G - G_{min}}{G_{max} - G_{min}}, \tag{8}$$

where G_{min} and G_{max} are the minimum and maximum gamma ray measurements in the log.

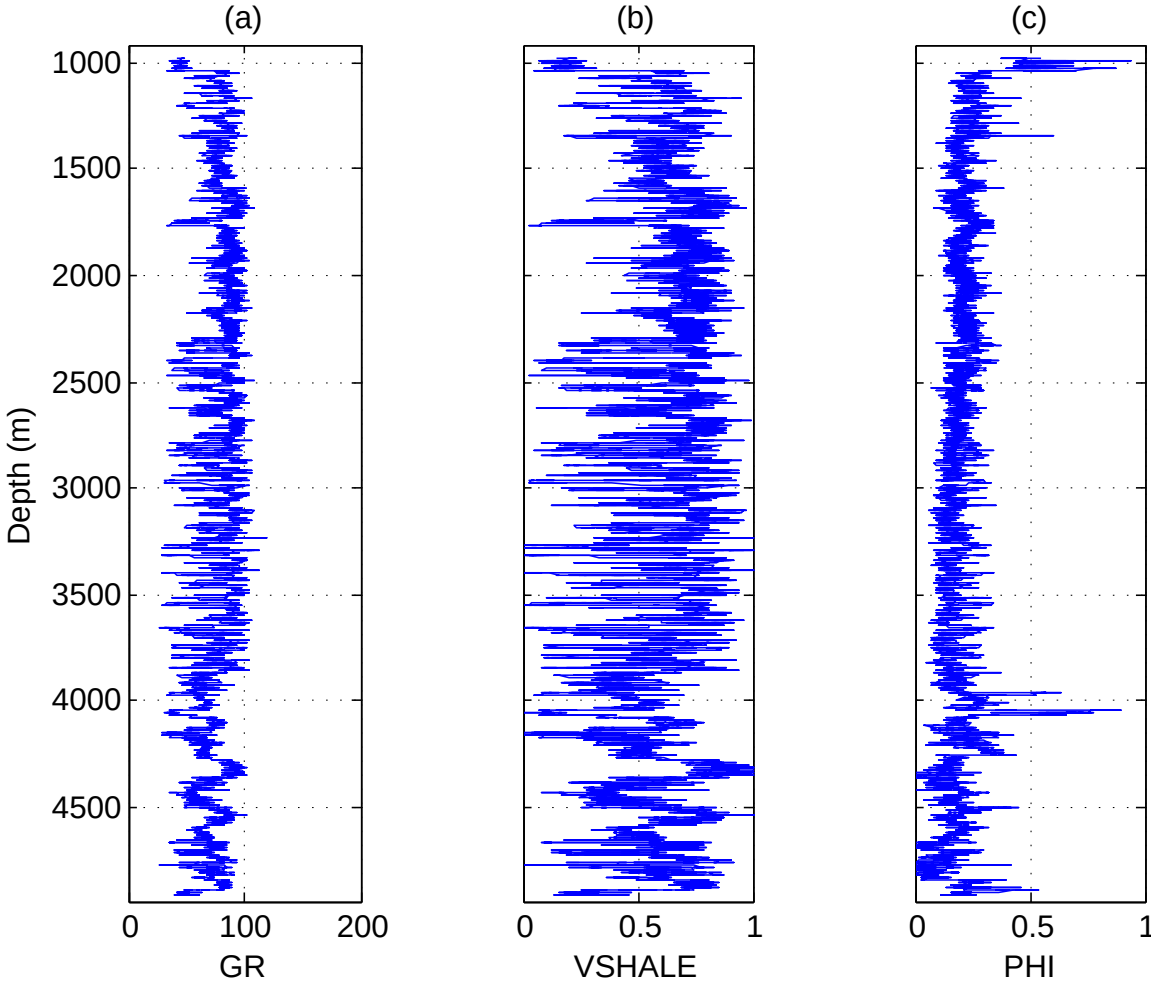


Figure 8: Input of the rock physics modeling from well log measurements: (a) gamma ray measurements, (b) deduced shale contents from gamma ray and (c) porosity. [**ER**] elita1/. wellinput

Figure 9 shows the topography of the model fitting objective function from the stochastic rock physics modeling using well log measurements. The results are shown at the same three depth levels as in figure 7. We can see that the centers of the distribution in figure 9 are slightly shifted and the distributions are more stretched compared with figure 7. Nonetheless, the general similarity in the shape and the orientation of both modeling results justify the use of the seismic inversion results and provide confidence when we use only the seismic information in deeper parts that the well log does not covered.

We approximate the topography of the objective function (6) by adding the rock physics model PDF to the data PDF with a weighting parameter λ. Due to the low resolution of

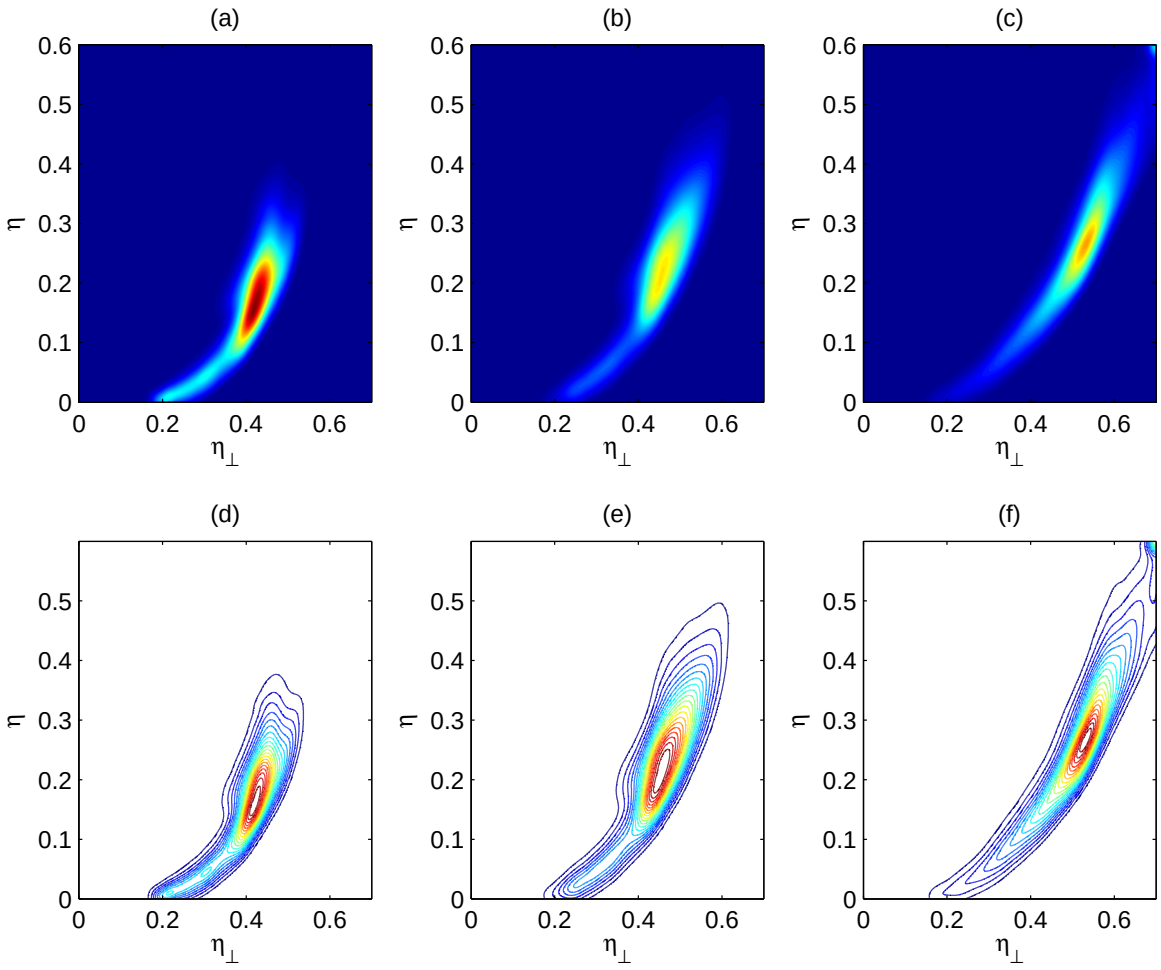


Figure 9: Stocastic rock physics modeling results for parameters v_n, η and $\eta_\perp$. Inputs of the rock physics modeling are from well log measurements. [**ER**] elita1/. well-model-prior

the rock physics modeling to velocity, we only show the topography projected onto the η - $\eta_\perp$ plane. We use the seismic inversion results as the input for rock physics modeling to study deeper subsurface. Figure 10 shows the unconstrained DSO topography with some constraint on η but nearly no constraint on $\eta_\perp$. The topography for the DSO objective function becomes wider as angle coverage of the subsurface gets narrower in the deeper subsurface.

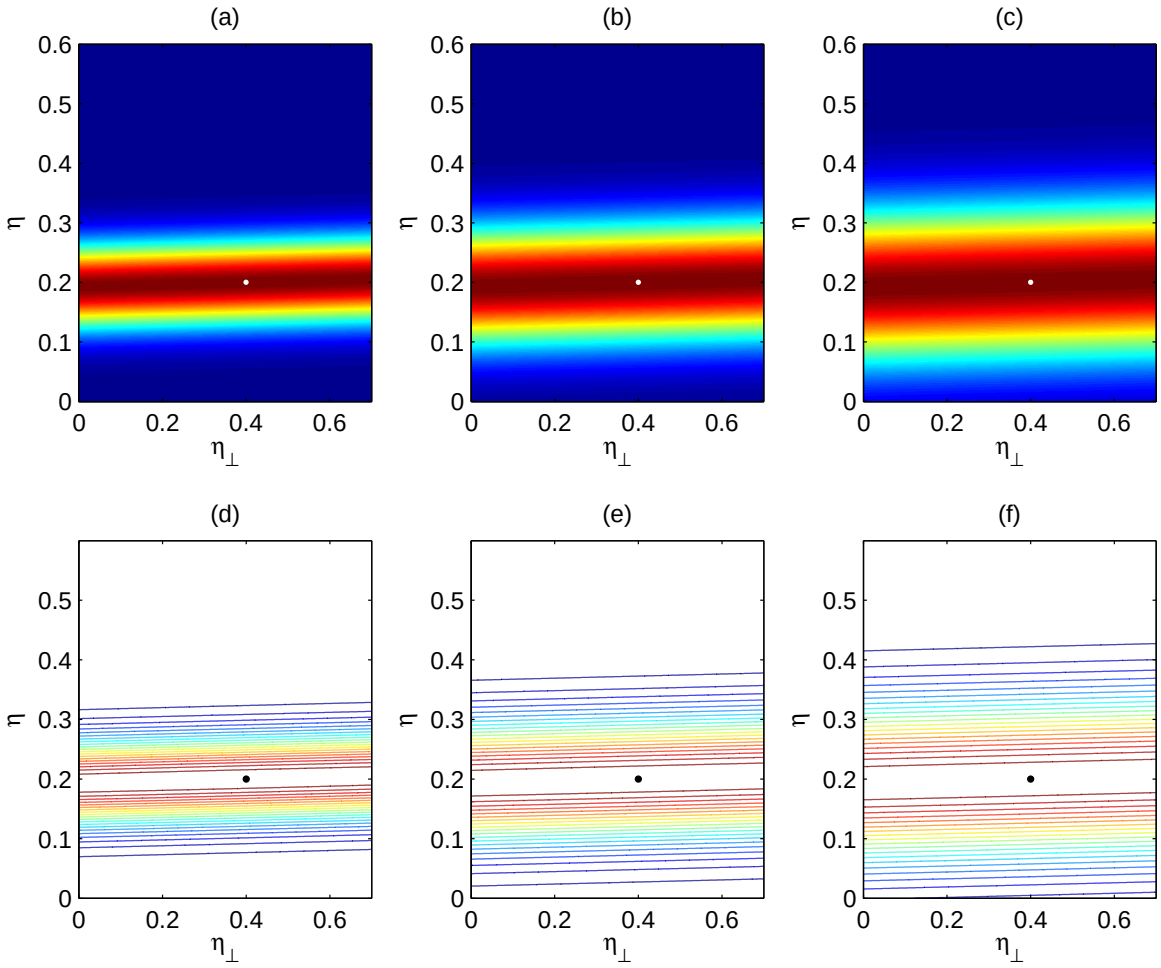


Figure 10: Topography of the DSO objective function in the η - $\eta_\perp$ plane in three different depth levels: (a) $0-3$ km, (b) $3-5$ km and (c) $5-7$ km. [**ER**] elita1/. data-model0

Figure 11 shows the topography of the rock physics constrained objective function (6) when $\lambda = 0.01$. By adding rock physics information, the center of the topography has moved, mostly along the unconstrained direction $\eta_\perp$. Since the data objective function has higher confidence in η, adding rock physics information does not change the solution of η in the shallow- and mid-depth. This suggests that this regularization scheme in the shallow subsurface will not change the flatness of the gathers, but will change the positioning of certain reflectors and hence provide different interpretations. On the other hand, in the deeper subsurface where seismic data do not have good constraints to either η or $\eta_\perp$, the solutions are dominated by the rock physics prior information.

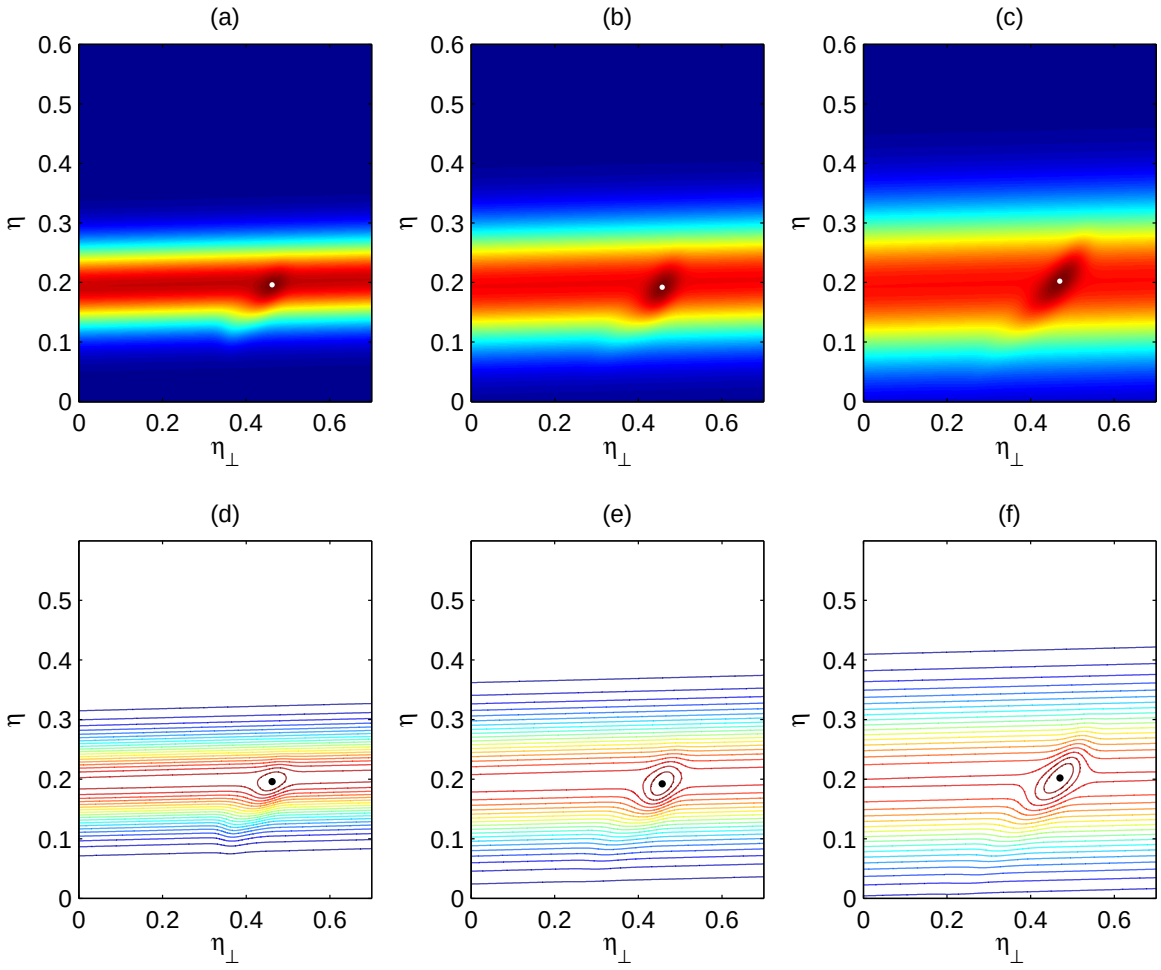


Figure 11: Topography of the regularized objective function with $\lambda = 0.01$ in the η - $\eta_\perp$ plane in three different depth levels: (a) $0 - 3$ km, (b) $3 - 5$ km and (c) $5 - 7$ km. [**ER**] elita1/. data-model001

CONCLUSIONS AND DISCUSSION

In this paper, we study the topography of the image-space DSO objective function with respect to velocity and anisotropic parameters. We show that due to the lack of constraints on the anisotropic parameters, other sources of information are needed to stabilize the inversion.

We combined the rock physics modeling workflow of Bandyopadhyay (2009) and Bachrach (2010a) to study the anisotropic properties of shales. Four different aspects of shale anisotropy are considered: mineral anisotropy, mineralogical transition between smectite and illite due to compaction and temperature, preferred orientations of the clay domain, and finally the lamination of sand and shale.

The stochastic rock physics modeling procedure is repeated with two different sources of inputs: seismic inversion results and well log measurements. Seismic inversion results cover large 3-D space with low resolution and low confidence, whereas the well log measurements provide localized information with high confidence. Stochastic rock physics modeling results from both information sources are compared. The similarity between both results justifies the use of seismic inversion results in the deeper subsurface where well log does not provide coverage. However, more uncertainties should be added in the shale volumetric content and porosity in the deeper region.

From the topography of the constrained objective function for anisotropic parameters, we show that adding rock physics information does not change the solution of the well-constrained parameter by the seismic data. However, valuable information is added where parameters are poorly constrained by the surface seismic data.

The effort of using previous seismic inversion results to constrain the seismic model building in the next iteration helps us to close the loop from seismic data to reservoir modeling. Traditional processing from seismic data to a reservoir model does not include feedback; and the seismic data modeled from the inverted reservoir model usually do not match the field data. Our study is a step along the way to build a closed loop from exploration to production. We hope to recover an Earth model that is consistent with all the available data we have.

ACKNOWLEDGMENTS

The authors thank Schlumberger-WesternGeco for the field dataset. This paper includes data supplied by IHS Energy Log Services; Copyright (2013) IHS Energy Log Services Inc.

REFERENCES

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

Bachrach, R., 2010a, Applications of deterministic and stochastic rock physics modeling to anisotropic velocity model building: SEG Expanded Abstracts, **29**, 2436–2440.

———, 2010b, Elastic and resistivity anisotropy of compacting shale: Joint effective medium modeling and field observations: SEG Expanded Abstracts, **29**, 2580–2584.

Backus, G., 1962, Long-wave elastic anisotropy produced by horizonal layering: Journal of Geophysical Research, **76**, 4427–4440.

Bakulin, A., Y. K. Liu, O. Zdraveva, and K. Lyons, 2010a, Anisotropic model building with wells and horizons: Gulf of mexico case study comparing different approaches: The Leading Edge, **29**, 1450–1460.

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

Bakulin, A., M. Woodward, D. Nichols, K. Osypov, and O. Zdraveva, 2009, Can we distinguish TTI and VTI media?: SEG Expanded Abstracts, **28**, 226–230.

———, 2010c, Building tilted transversely isotropic depth models using localized anisotropic tomography with well information: Geophysics, **75**, 27–36.

———, 2010d, Localized anisotropic tomography with well information in VTI media: Geophysics, **75**, 37–45.

Bandyopadhyay, K., 2009, Seismic anisotropy: geological causes and its implications: PhD thesis, Stanford University.

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

Dræge, A., M. Jakobsen, and T. A. Johansen, 2006, Rock physics modeling of shale diagenesis: Petroleum Geoscience, **12**, 49–57.

Freed, R. L. and D. R. Peacor, 1989, Variability in temperature of the smectite/illite reaction in Gulf coast sediments: Clay Minerals, **24**, 171–180.

Hornby, B., D. Miller, C. Esmersoy, and P. Christie, 1995, Ultrasonic-to-seismic measurements of shale anisotropy in a North Sea well: SEG Expanded Abstracts, **14**, 17–21.

Hornby, B. E., L. M. Schwartz, and J. A. Hudson, 1994, Anisotropic effective-medium modeling of the elastic properties of shales: Geophysics, **59**, 1570–1583.

Hower, J., E. Eslinger, M. Hower, and E. Perry, 1976, Mechanism of burial metamorphism of argillaceous sediments: 1. Mineralogical and chemical evidence: Bulletin of The Geological Society of America, **87**, 725–737.

Li, Y. and B. Biondi, 2011, Migration velocity analysis for anisotropic models: SEG Expanded Abstract, **30**, 201–206.

Li, Y., D. Nichols, K. Osypov, and R. Bachrach, 2011, Anisotropic tomography using rock physics contraints: 73^{rd} EAGE Conference & Exhibition.

Sayers, C., 2004, Seismic anisotropy of shales: What determines the sign of Thomsen's delta parameter?: SEG Expanded Abstracts, **23**, 103–106.

———, 2010, The effect of anisotropy on the Young's moduli and Poisson's ratios of shales: SEG Expanded Abstracts, **29**, 2606–2611.

Wenk, H.-R., I. Lnardelli, H. Franz, K. Nihei, and S. Nakagawa, 2007, Preferred orientation and elastic anisotropy of illite-rich shale: Geophysics, **72**, E69–E75.

Yang, Y., K. Osypov, R. Bachrach, M. Woodward, O. Zdraveva, Y. Liu, A. Fournier, and Y. You, 2012, Anisotropic tomography and uncertainty analysis with rock physics constraints: Green Canyon case study: SEG Expanded Abstract, 1–5.

Image-guided WEMVA for azimuthal anisotropy

Yunyue (Elita) Li

ABSTRACT

Azimuthal anisotropy is common in layered basins with strong folding and fracturing effects. Traditional processing on individual azimuths usually yields images with inconsistent depths. In this paper, we propose to use an image obtained from one azimuth to constrain the image-space velocity analysis on other azimuths. Instead of using the traditional differential semblance penalty function, we define the image penalty weight according to an existing image of one azimuth. This method directly tackles the differences in anisotropic parameters among different azimuths. By keeping the vertical velocity constant across the azimuths, we separate the kinematic effects due to the anisotropic parameters from those due to the velocity error. We test the image-guided migration velocity analysis algorithm on a simple example with flat reflectors and a homogeneous orthorhombic subsurface. We compare the residual images and the first η gradient obtained by our image-guided method with the differential semblance method under three different conditions: accurate velocity, fast velocity and slow velocity. Our results show that despite the velocity error, the image-guided migration velocity analysis algorithm provides consistent η gradients, whereas the gradients given by differential semblance optimization may be misled by the velocity error.

INTRODUCTION

Since it was first reported in exploration seismology in the 1930s (McCollum and Snell, 1932), anisotropy has played an increasingly important role in seismic imaging and exploration. Until now, the layered transverse isotropic (TI) model has been the most commonly used model in seismic imaging. Postma (1955) and Helbig (1956) showed that a sequence of isotropic layers on a scale much smaller than the wavelength leads to an anisotropic medium.

If the layers are deposited horizontally, the medium is defined as a vertical TI (VTI) medium. A VTI medium is commonly formed because of the thin bedding during the deposition. If the layers become dipping during the deformation, a general tilted TI (TTI) medium is formed. These TI media are commonly found within sedimentary basins across the world. Due to changes in the surrounding stress fields, small scale fractures or cracks may form in the layered media. The combination of parallel vertical cracks and vertical transverse anisotropy in the background medium is the most common cause of an effective orthorhombic medium (Figure 1). Efficient and accurate wave propagation in the orthorhombic medium has been extensively studied (Tsvankin, 1997; Cheng et al., 2012), with new developments in wave-equation based orthorhombic propagation in recent years (Zhang and Zhang, 2011; Fowler and Lapilli, 2012; Chu, 2012). These developments help the understanding of the wave phenomenon in complex geological settings. However, re-

liable inversion for the elastic parameters needed for orthorhombic modeling is still under investigation.

To fully describe an orthorhombic medium, one must constrain the nine independent parameters in the elastic tensor. Early methods based on shear wave splitting are useful for detecting the orientation of the fractures (Garotta, 1989; Olofsson et al., 2003); however, they are far from sufficient to constrain the whole medium. Tsvankin (1997) reparameterizes the orthorhombic medium and reduces the number of parameters to six governing P-wave propagation in the orthorhombic medium. This reduction of the parameter and analysis on the P-wave propagation in the symmetrical planes shed light on the parameter estimation: to fully constrain all six parameters, we need data from at least six different azimuths. With the modern full azimuthal acquisition, especially OBS and coil shooting acquisition, this requirement is not difficult to meet.

Now the question remains: given full azimuth data, how should we process the data and build a corresponding orthorhombic model? Recent studies analyze the non-hyperbolic normal moveout of the data to invert for the orthorhombic parameters (Grechka and Tsvankin, 1999; Elapavuluri and Bancroft, 2006; Vasconcelos and Tsvankin, 2006). However, these data space methods are prone to noise in the data and may require over-simplification of the subsurface structures.

This study discusses one possible way to directly resolve the velocity differences among different azimuths in the image space. We propose to work with one azimuth at a time. Once the first 3-D image from the first azimuth is properly obtained, we can use this image as a reference to constrain the inversion at the other azimuths. The idea of defining a penalty function according to an existing image is not new. It was first proposed by Shragge and Lumley (2013) to highlight time-lapse velocity changes in the subsurface. We borrow the same idea and apply it to resolve the differences among different azimuths, and hence to resolve the azimuthal anisotropy.

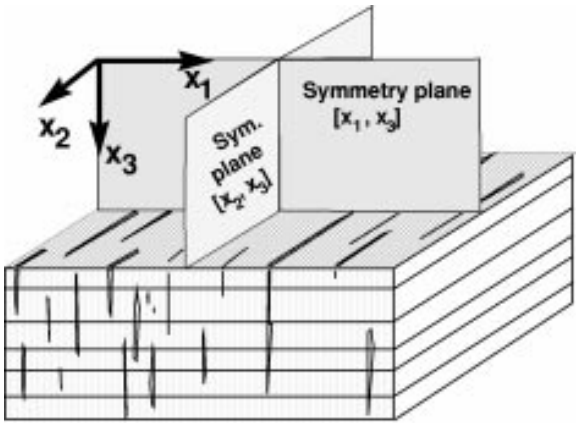


Figure 1: An orthorhombic model caused by parallel vertical cracks embedded in a VTI medium. Orthorhombic media have three mutually orthogonal planes of mirror symmetry. From Tsvankin (1997). [**NR**] elita2/. orthorhombicmodel

In this paper, we first review the parameters needed to describe the P-wave propagation in an orthorhombic medium. Assumptions such as weak anisotropy are utilized to reduce the complexity of the subsurface model. We then briefly discuss the methodology for the

image-guided WEMVA and show a synthetic example where images at two azimuths are analyzed. By comparing the velocity updates at the second azimuth of our proposed method and straightforward DSO method in three different cases: accurate velocity, fast velocity and slow velocity, we demonstrate that our method is robust against the velocity error when determining the anisotropic property of the subsurface.

ANISOTROPIC PARAMETERS FOR P-WAVE PROPAGATION IN ORTHORHOMBIC MEDIUM

The stiffness tensor for orthorhombic medium can be represented in the "Voigt recipe" as follows:

$$\mathbf{c} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{22} & c_{23} & 0 & 0 & 0 \\ c_{13} & c_{23} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{pmatrix}. \tag{1}$$

In a general orthorhombic medium, the nine components in the stiffness tensor are all independent of each other. Following a definition similar to Thomsen's (Thomsen, 1986), Tsvankin (1997) parameterized orthorhombic media using similar velocity and dimensionless coefficients as follows:

- V_{P0} – the vertical velocity of the P-wave;
- V_{S0} – the vertical velocity of the S-wave polarized in the x_1 direction;
- $\epsilon^{(2)}$ – the VTI parameter ϵ in the symmetry plane $[x_1, x_3]$;
- $\delta^{(2)}$ – the VTI parameter δ in the symmetry plane $[x_1, x_3]$;
- $\gamma^{(2)}$ – the VTI parameter γ in the symmetry plane $[x_1, x_3]$;
- $\epsilon^{(1)}$ – the VTI parameter ϵ in the symmetry plane $[x_2, x_3]$;
- $\delta^{(1)}$ – the VTI parameter δ in the symmetry plane $[x_2, x_3]$;
- $\gamma^{(1)}$ – the VTI parameter γ in the symmetry plane $[x_2, x_3]$;
- $\delta^{(3)}$ – the VTI parameter δ in the symmetry plane $[x_1, x_2]$.

where the velocities are

$$V_{P0} \equiv \sqrt{\frac{c_{33}}{\rho}}, \tag{2}$$

and

$$V_{S0} \equiv \sqrt{\frac{c_{55}}{\rho}}, \tag{3}$$

with ρ the density.

The dimensionless coefficients are related to the stiffness coefficients as follows:

$$\epsilon^{(1)} \equiv \frac{c_{22}-c_{33}}{2c_{33}}, \tag{4}$$

$$\delta^{(1)} \equiv \frac{(c_{23}+c_{44})^2-(c_{33}-c_{44})^4}{2c_{33}(c_{33}-c_{44})}, \tag{5}$$

$$\gamma^{(1)} \equiv \frac{c_{66}-c_{55}}{2c_{55}}, \tag{6}$$

$$\epsilon^{(2)} \equiv \frac{c_{11}-c_{33}}{2c_{33}}, \tag{7}$$

$$\delta^{(2)} \equiv \frac{(c_{13}+c_{55})^2-(c_{33}-c_{55})^4}{2c_{33}(c_{33}-c_{55})}, \tag{8}$$

$$\gamma^{(2)} \equiv \frac{c_{66}-c_{44}}{2c_{44}}, \tag{9}$$

$$\delta^{(3)} \equiv \frac{(c_{12}+c_{66})^2-(c_{11}-c_{66})^4}{2c_{11}(c_{11}-c_{66})}. \tag{10}$$

Under the assumption of weak anisotropy, the phase velocity of the P-wave is as follows:

$$V_p(\theta,\phi) = V_p0[1+\delta(\phi)\sin^2\theta\cos^2\theta+\epsilon(\phi)\sin^4\theta]; \tag{11}$$

with azimuthal dependent anisotropic parameters $\epsilon(\phi)$ and $\delta(\phi)$ defined by

$$\epsilon(\phi) = \epsilon^{(1)}\sin^4\phi+\epsilon^{(2)}\cos^4\phi+(2\epsilon^{(2)}+\delta^{(3)})\sin^2\phi\cos^2\phi, \tag{12}$$

$$\delta(\phi) = \delta^{(1)}\sin^2\phi+\delta^{(2)}\cos^2\phi. \tag{13}$$

It can be seen from equations 11 through 13 that the kinematic signatures of P-waves in weak orthorhombic media depend on just *five* anisotropic parameters in the symmetric planes and the vertical P-wave velocity. This reduction of parameters gives us a chance to resolve a reliable orthorhombic model using modern full azimuth acquisition geometry.

IMAGE-GUIDED WEMVA

Although 3-D seismic acquisition and processing has been performed in practice for more than two decades (Biondi, 2006), conventional processing workflows still work with narrow azimuth data and process a single azimuth at a time. In areas with complex geology such as multi-sets of fracture systems or combination of one set of fractures within a layered background medium, the images obtained from different azimuths using the same velocity and anisotropic model will show different characteristics: depth and focusing. Zhang et al. (2012) shows shows one example of this effect in figure 2. On the left panel, the stacked image is overlaid on the vertical velocity model. It is obvious that this area is populated with fractures and faults. The blue ellipse highlights the region where the quality of the image is degraded by the oversimplified velocity model. The panel on the right shows offset-domain common image gathers obtained from four different surface azimuths. Although the near offsets are more or less flattened on all four azimuthal CIGs, the far offset moveouts on different azimuthal CIGs clearly show different characteristics. Furthermore, due to the inaccuracy in the azimuthal anisotropic model, reflectors are imaged at different depths at

different azimuths. Proper handling of the azimuthal anisotropy will increase the coherence of the depth across different azimuths and therefore produce better stacked images with higher resolution. One of the key aspects of image-space WEMVA is the definition of

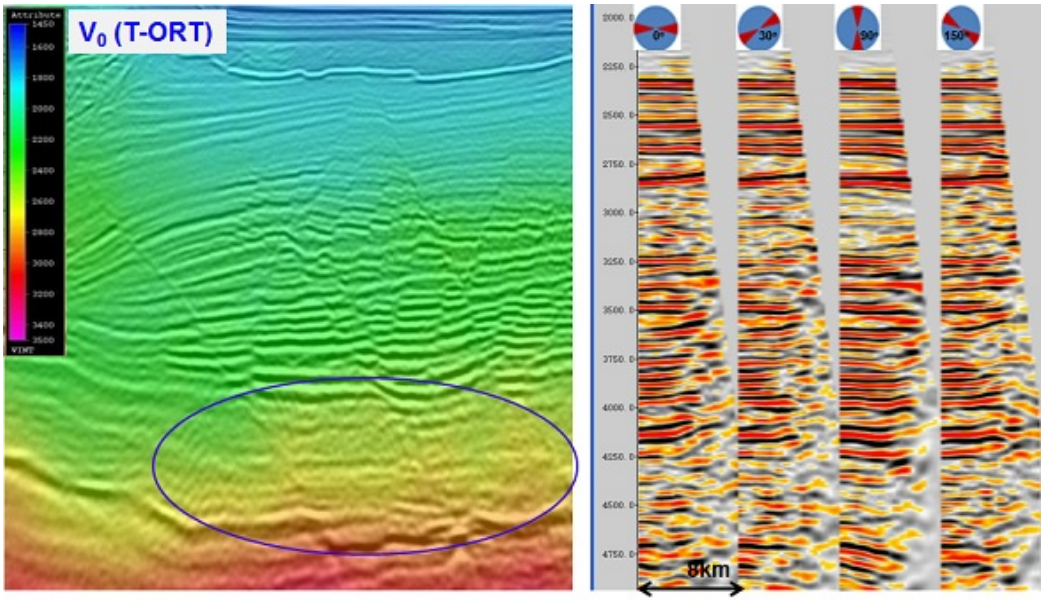


Figure 2: Stack image (left) and the offset-domain common image gathers at different surface azimuths (right). From Zhang et al. (2012). [**NR**] elita2/. CGGgathers

the "target" image (Biondi, 2006), which subsequently defines the residual image used for gradient computation. The WEMVA objective function in the subsurface offset domain can be written as follows:

$$J = ||\mathbf{W}(\mathbf{h})\mathbf{I}(\mathbf{x},\mathbf{h})||_2^2. \tag{14}$$

Differential semblance optimization (DSO) (Shen and Symes, 2008) has been a popular choice for defining the "target" image thanks to its simplicity in concept and automation in implementation. Equation 15 shows the weighting function defined by the conventional DSO objective function, where h is the length of the subsurface offset $\mathbf{h}$ and h_{max} is the maximum subsurface offset.

$$\mathbf{W}_{\text{dso}}(\mathbf{x},\mathbf{h}) = \frac{h}{h_{max}}. \tag{15}$$

Hence, the DSO objective function leads to a residual image as follows:

$$\Delta\mathbf{I}_{\text{dso}} = \mathbf{W}_{\text{dso}}^{*}\mathbf{W}_{\text{dso}}\mathbf{I}(\mathbf{x},\mathbf{h}), \tag{16}$$

where $\mathbf{I}$ is the common image gathers in the subsurface offset domain. Notice that the DSO weighting function is invariant with the spatial coordinates.

On the other hand, many authors (Fei and Williamson, 2010; Vyas and Tang, 2010) have pointed out various artifacts in the DSO gradient, such as side lobes and scattering effects. Therefore, we need a better way to define the residual image and the objective function. In modern acquisition, where the subsurface is illuminated from all angles (up to the angle defined by the maximum offset) and full azimuths, we can obtain up to a seven-dimension image cube of the subsurface. Due to this redundancy, it is possible to form

not only image gathers with respect to reflection angle, but also the multi-azimuth image gathers. Although full 3D imaging techniques have existed for more than two decades, single azimuthal processing is still the common practice in the industry. Therefore, the image obtained from one azimuth can be used as the "target" image for the other azimuths.

We propose to define the weighting function according to the normalized envelope of the existing image:

$$\mathbf{W}_{\text{img}}(\mathbf{x},\mathbf{h}) = 1 - \frac{E(\mathbf{I_0}(\mathbf{x},\mathbf{h}))}{\max(E(\mathbf{I_0}(\mathbf{x},\mathbf{h})))}, \tag{17}$$

where E denotes the envelope function and $\mathbf{I_0}$ is the reference image at a certain azimuth.

To test our velocity analysis method by matching images at two different azimuths, we take advantage of the fact that the P-wave propagation is fully described using conventional VTI wave equations in the symmetric planes. The only difference is that the "VTI parameters" are different in each plane. Nonetheless, they do share the same vertical P-wave velocity.

NUMERICAL TEST

In this section, we test our idea on a simple numerical model. The synthetic model contains five flat reflectors, and the velocity and anisotropic parameters are constant. We take advantage of the fact that the P-wave propagation is fully described using conventional VTI wave equations in the symmetric planes, and we model the data at azimuth 0° and azimuth 90° using the same VTI one-way wave-equation with different values for the parameters: at azimuth 0°, $V_{P0} = 2000m/s$, $\eta^{(2)} = 0.2$, $\delta^{(2)} = 0.1$; at azimuth 90°, $V_{P0} = 2000m/s$, $\eta^{(1)} = 0.4$, $\delta^{(1)} = 0.1$.

We first migrate both datasets at azimuth 0° and 90° using the same migration model: $V_{P0} = 2000m/s$, $\eta = 0.2$, $\delta = 0.1$. In this case, the subsurface model is exact for azimuth 0°. Therefore, in the subsurface-offset CIGs at azimuth 0° on panel (a) in figure 3, the events are almost focused at the zero subsurface-offset, except for some illumination artifacts at the deeper reflectors. However, the downward curvature on panel (b) in figure 3 is due to the negative error in the η model. Thanks to the wider angle coverage, the kinematic differences between the two azimuthal images are more significant on the shallower reflectors than on the deeper ones.

To penalize the unfocused image at azimuth 90°, the DSO weighting matrix $\mathbf{W}_{\text{dso}}$ can be applied at each image point in the subsurface offset domain, as shown in panel (a) of in figure 9. Notice that the weighting function is uniformly distributed vertically, regardless of the illumination effects. Also, the DSO penalty function is independent of the vertical velocity error. On the other hand, we can also design the penalty function according to the image at azimuth 0° using equation 17. The image-guided penalty function $\mathbf{W}_{\text{img}}$ with the exact vertical velocity is shown in panel (b) in figure 9.

The residual images produced by DSO and the image-guided penalty function are shown in panels (a) and (b) in figure 4, respectively. Notice that the DSO penalty function highlights the leaked energy from the zero subsurface offset, whereas the image-guided penalty function highlights the kinematics differences between two azimuths. When the vertical velocity model is exact, the leaked energy and the kinematics differences are both

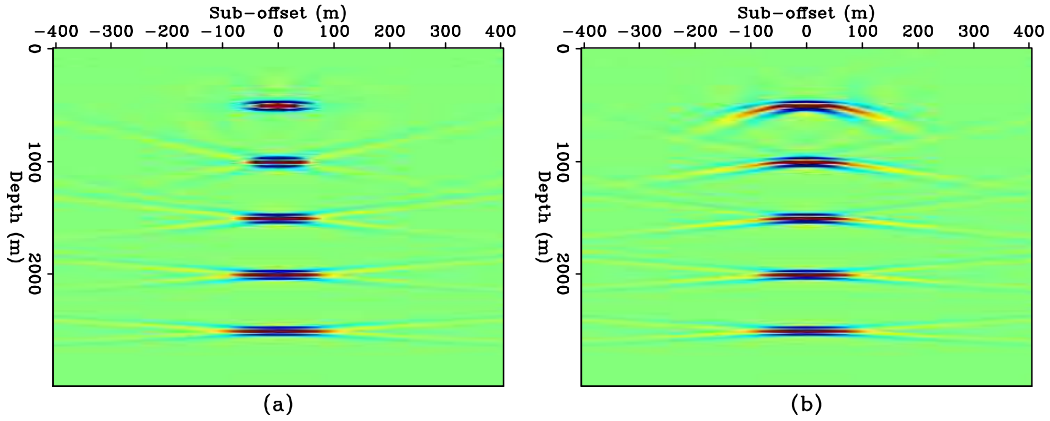


Figure 3: Subsurface-offset gathers at $x = 0m$ at azimuth 0° (a) and azimuth 90° (b). Migration parameters used in both data are $V_{P0} = 2000$ m/s, $\eta = 0.2$, $\delta = 0.1$. [**ER**] elita2/. bimg-true

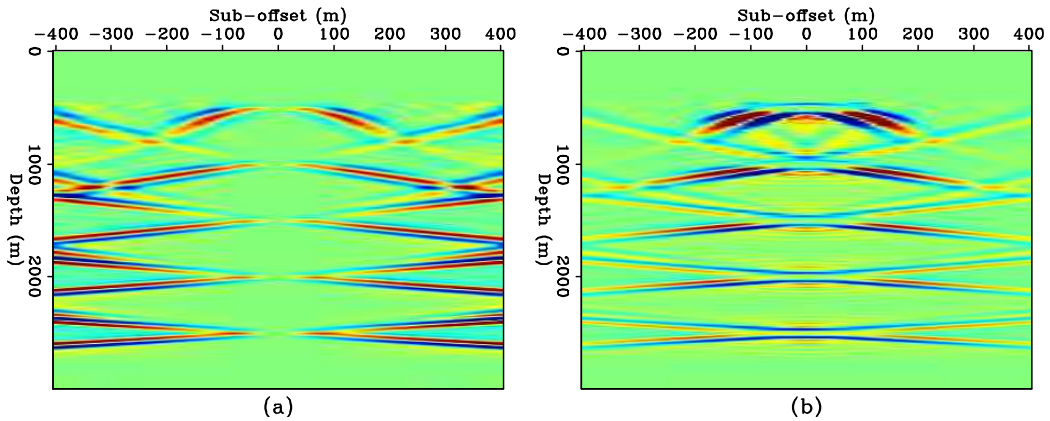


Figure 4: Residual subsurface-offset image at azimuth 90° using DSO penalty function (a) and image-guided penalty function (b). Migration parameters are $V_{P0} = 2000$ m/s, $\eta = 0.2$, $\delta = 0.1$. [**ER**] elita2/. dimg-true

caused by the error in the η model only.

Back-projections of the residual images define the gradient direction in the model space. We stack the gradient over the horizontal axis, since the error in the η model is homogeneous. The gradients in η produced by DSO and image-guided penalty functions are plotted in panels (a) and (d) in figure 10, respectively. Both gradients point to the correct update directions to compensate for the η error and show higher sensitivity to the shallower reflectors than the deeper reflectors.

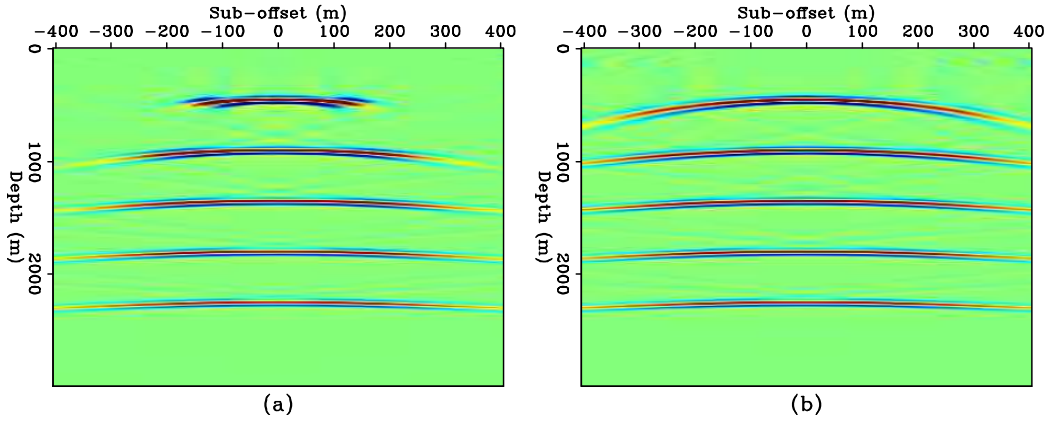


Figure 5: Subsurface-offset gathers at $x = 0m$ at azimuth 0° (a) and azimuth 90° (b).Migration parameters used in both data are $V_{P0} = 1800$ m/s, $\eta = 0.2$, $\delta = 0.1$. [**ER**] elita2/. bimg-slow

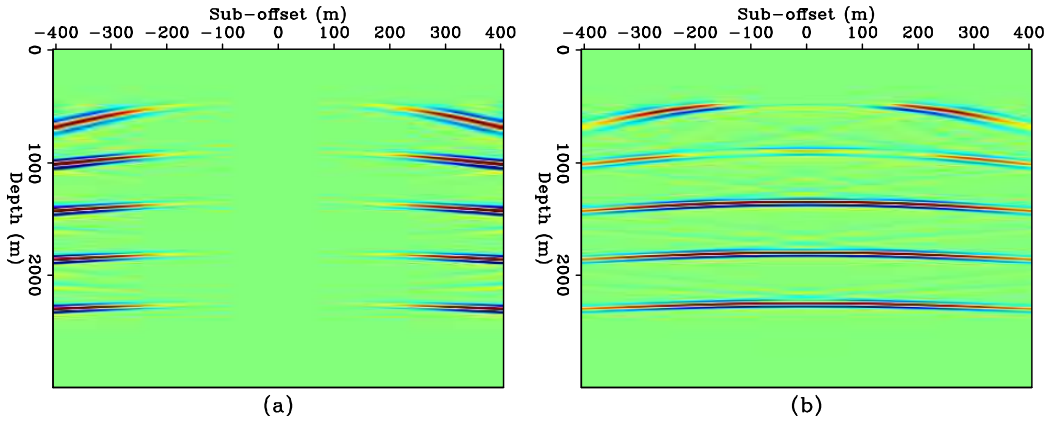


Figure 6: Residual subsurface-offset image at azimuth 90° using DSO penalty function (a) and image-guided penalty function (b). Migration parameters are $V_{P0} = 1800$ m/s, $\eta = 0.2$, $\delta = 0.1$. [**ER**] elita2/. dimg-slow

To test the ambiguity between velocity and η, we assume the vertical velocity is not accurately estimated. The parameters used in the second test are: $V_{P0} = 1800$ m/s, $\eta = 0.2$, $\delta = 0.1$. Notice that the error in vertical velocity is in the same direction as that in the η model for azimuth 90°. The background images at azimuths 0° and 90° are shown in figure 5. The downward curvature in panel (a) is caused by the negative vertical velocity error,

while the broader curvature in panel (b) is caused by the additional negative error in η. Comparing the two images and ignoring the illumination effect, the only difference between the two images is caused by the η error.

Using the unfocused image at azimuth 0° as guidance, the highlighted differences between two azimuths are shown in figure 6(b). On the other hand, the residual image defined by DSO penalty function in 6 (a) includes the kinematic errors caused by both velocity and η error.

Back-projections of the residual images in figure 6 are plotted in panels (b) and (e) in figure 10. Both DSO and image-guided penalty functions point to the correct update direction. The updates of η is significantly stronger from the first reflector than from the deeper ones.

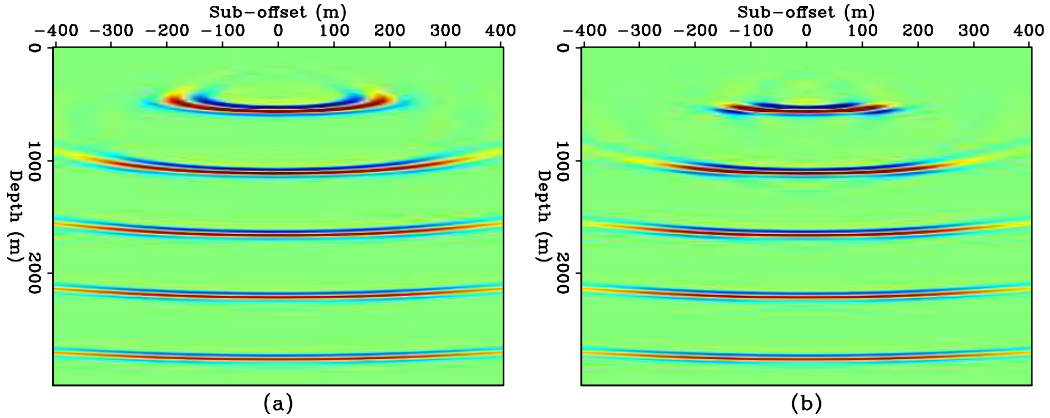


Figure 7: Subsurface-offset gathers at $x = 0m$ at azimuth 0° (a) and azimuth 90° (b). Migration parameters used in both data are $V_{P0} = 2200$ m/s, $\eta = 0.2$, $\delta = 0.1$. [**ER**] elita2/. bimg-fast

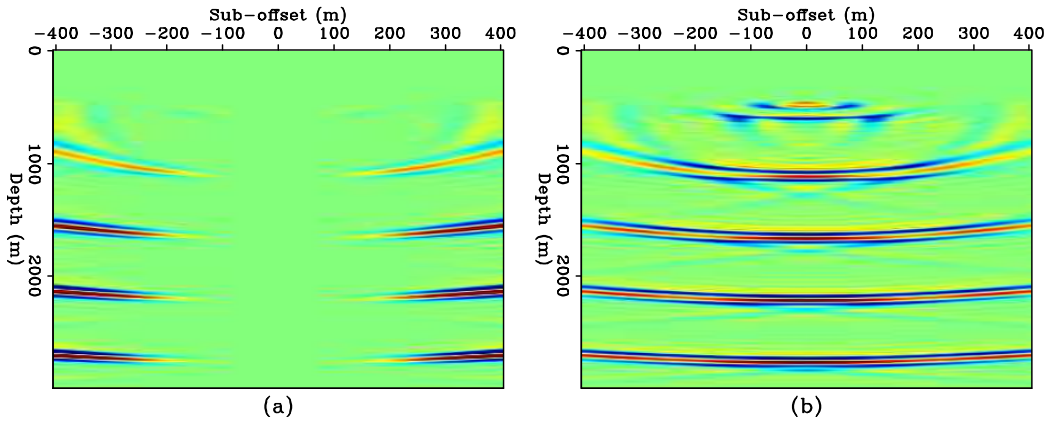


Figure 8: Residual subsurface-offset image at azimuth 90° using DSO penalty function (a) and image-guided penalty function (b). Migration parameters are $V_{P0} = 2200$ m/s, $\eta = 0.2$, $\delta = 0.1$. [**ER**] elita2/. dimg-fast

A more interesting test of the ambiguity between velocity and η is to perturb the velocity model in the opposite direction to the η perturbation. Therefore, we migrate both datasets

at both azimuths using $V_{P0} = 2200$ m/s, $\eta = 0.2$, $\delta = 0.1$. For azimuth 0°, the upward moveout in the background subsurface domain common image gathers (Figure 7(a)) is caused by the positive velocity error. For azimuth 90°, the narrower upward moveout is due to the trade-off effect between a positive velocity error and a negative velocity error. Since velocity has a first order effect on the kinematics of the wavefield, the overall characteristics in the subsurface offset domain at azimuth 90° indicate positive errors in the subsurface model.

The residual images produced by DSO and the image-guided penalty function are shown in figure 8, back-projections of which are plotted in panels (c) and (f) in figure 10, respectively. Clearly, due to the dominant effect of faster velocity, the DSO penalty function suggests to reduce the value of η, which kinematically reduces the velocity at large angles. However, the image-guided penalty function highlights the true kinematic error caused by η and successfully keeps the η update in the correct direction. This result shows that by using the image-guided penalty function, we have a chance to resolve the ambiguity between the velocity model and the η model.

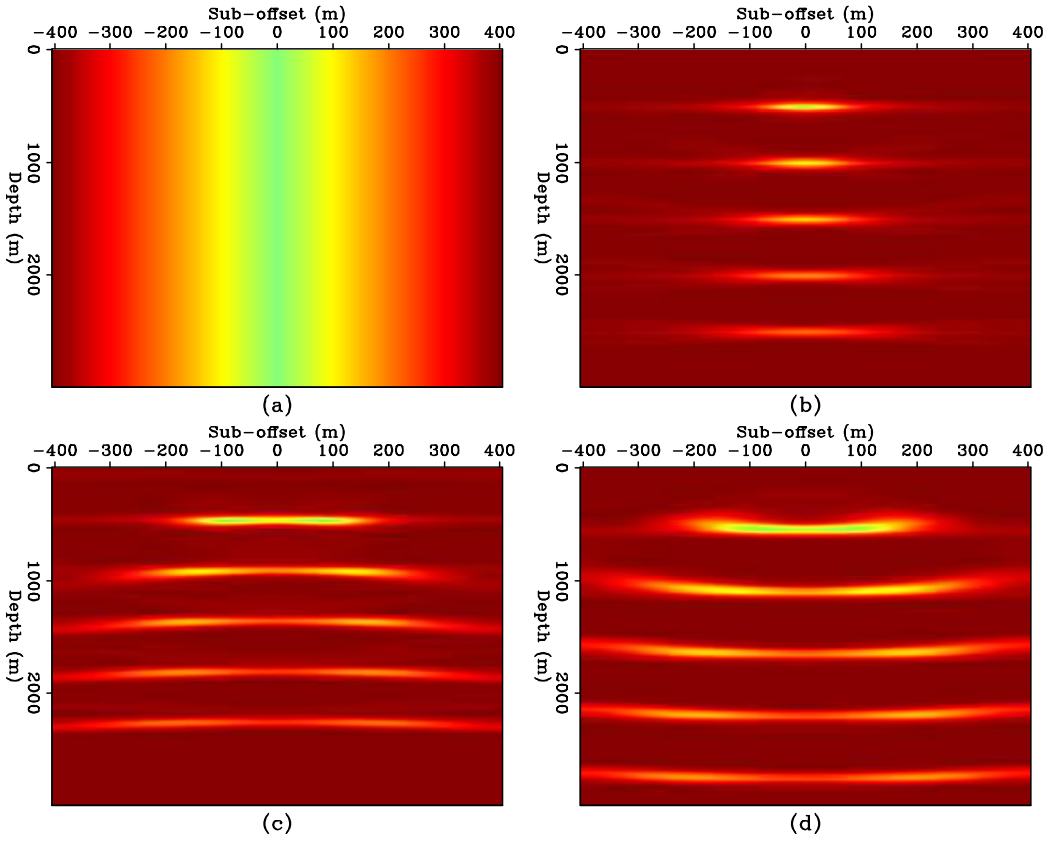


Figure 9: Spatial weighting function in the subsurface-offset domain at $x = 0$ m from DSO (a). This weighting function is independent of the accuracy of the vertical velocity model. Image-guided penalty function when the migration vertical velocity is the same (b), smaller (b) and larger (c) than the true vertical velocity. [**ER**] elita2/. imgwgt

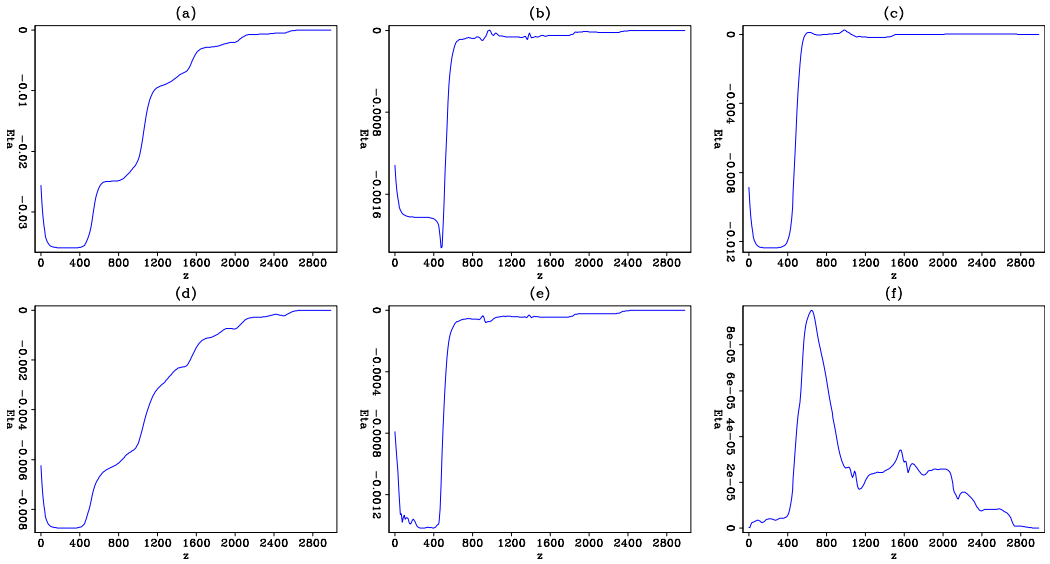


Figure 10: First gradient for η given by the image-guided penalty function (top row) and the DSO penalty function (bottom row). The columns from left to right correspond to the cases where the background velocity is the same, smaller or bigger than the true velocity, respectively. [**ER**] elita2/. compgrd

CONCLUSIONS AND DISCUSSIONS

In this paper, we propose an image-guided penalty function to target the difference in anisotropic parameters at different azimuths. In an orthorhombic medium, the P-wave kinematics are governed by six parameters, among which the vertical velocity is shared by both azimuths in the $[x_1, x_3]$ plane and the $[x_2, x_3]$ plane. By fixing the migration models for both azimuths, and comparing the resulting images, we directly resolve the difference in the equivalent Thomsen parameters between both azimuths, regardless of the accuracy of the velocity model.

More importantly, the image-guided penalty function distinguishes the kinematic contribution of velocity from that of η by referencing the azimuthal images with the same velocity effects. The resulting residual image contains only the residual kinematics due to the difference in the anisotropic parameters. Therefore, regardless of the dominant effect of velocity error, the image-guided penalty function can produce correct updates for the anisotropic parameters.

However, the application of this method could potentially be limited by the subsurface structures and the acquisition geometry. For example, in areas with highly dipping reflectors and 3D structures, the surface azimuths may no longer properly represent the azimuths in the subsurface. In this case, the subsurface azimuthal analysis may be used to evaluate the accuracy of the migration models. In addition, when acquisition geometries are significantly different across azimuths, the illumination differences may overwhelm the differences between the anisotropic parameters.

REFERENCES

Biondi, B., 2006, 3d seismic imaging: Society Of Exploration Geophysicists.

Cheng, X., J. Tang, O. Zdraveva, C. Yarman, and J. Hobro, 2012, Ray-based prestack depth migration for orthorhombic media: SEG Expanded Abstracts, 1–5.

Chu, C., 2012, A hybrid grid finite difference method for acoustic wave propagation in titled orthorhombic media: SEG Expanded Abstracts, 1–5.

Elapavuluri, P. and J. Bancroft, 2006, Estimation of anisotropy parameters in orthorhombic media: SEG Expanded Abstracts, 164–168.

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

Fowler, P. and C. Lapilli, 2012, Generalized pseudospectral methods for orthorhombic modeling and reverse time migration: SEG Expanded Abstracts, 1–5.

Garotta, R., 1989, Detection of azimuthal anisotropy: SEG Expanded Abstracts, **1**, 861–863.

Grechka, V. and I. Tsvankin, 1999, 3-D moveout velocity analysis and parameter estimation for orthorhombic media: Geophysics, **64**, 820–837.

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

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

Olofsson, B., T. Probert, J. Kommedal, and O. Barkved, 2003, Azimuthal anisotropy from the Valhall 4C 3D survey: The Leading Edge, **22**, 1228–1235.

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

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

Shragge, J. and D. Lumley, 2013, Time-lapse wave-equation migration velocity analysis: Geophysics, **78**, S69–S79.

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

Tsvankin, I., 1997, Anisotropic parameters and P-wave velocity for orthorhombic media: Geophysics, **62**, 1292–1309.

Vasconcelos, I. and I. Tsvankin, 2006, Non-hyperbolic moveout inversion of wide-azimuth P-wave data for orthorhombic media: Geophysical Prospecting, **54**, 535–552.

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

Zhang, H. J. and Y. Zhang, 2011, Reverse time migration in vertical and tilted orthorhombic media: SEG Expanded Abstracts, 185–189.

Zhang, Y., S. Xu, T. Huang, W. Han, Y. Li, and C.-S. Chen, 2012, Reverse time migration in tilted orthorhombic media: International Workshop on Seismic Anisotropy.

Wave-equation migration Q analysis (WEMQA)

Yi Shen

ABSTRACT

Q model building, which is conventionally done in the data space using ray-based tomography, is a notoriously challenging problem due to issues like spectral interference, low signal-to-noise ratio, diffractions, and complex subsurface structure. To produce a reliable Q model, I present a new approach with two major features. First, this method is performed in the image space, which uses downward-continuation imaging with Q to stack out noise, focus and simplify events, and provide a direct link between the model perturbation and the image perturbation. I develop two methods to generate the image perturbation for the following scenarios: models with sparse reflectors and models with dense reflectors. Second, this method uses wave-equation Q tomography to handle complex wave propagation. Two synthetic tests on two different 2-D models with Q anomalies shows the capability of this method on models with sparse events. Tests with a modified SEAM model also demonstrate the feasibility of this method for a model with dense events.

INTRODUCTION

Attenuation, parametrized by seismic quality factor Q, causes high-frequency loss and phase distortion of surface seismic reflection data. An understanding of the effects and properties of this attenuation parameter has two major motivations. First, Q is a useful parameter for characterizing rock and fluid properties–e.g., saturation, porosity, permeability, and viscosity–because of its high sensitivity to some of these properties (e.g. Best et al. (1994)). Second, if the absorption properties of the subsurface are known, they can be included in seismic data processing (deconvolution, stacking, migration, inverse Q filtering, etc.) to get higher-quality images, and to better interpret the effects of AVO and anisotropy, which also have offset-dependent signatures.

Studies of estimating attenuation tomographically have a long history. Brzostowski and McMechan (1992), and Leggett et al. (1992) conducted tomographic estimation of attenuation using seismic amplitude changes. Kjartansson (1979), and Zucca et al. (1994) used the rise time of the broadened wavelets caused by attenuation for Q tomography. Tonn (1991), Quan and Harris (1997), Dasgupta and Clark (1998), Leaney (1999), Mateeva (2003), Plessix (2006), Rickett (2006), Rickett (2007), Reine et al. (2012a), and Reine et al. (2012b) performed the estimation based on the attenuation-induced spectral changes. However, there are three main difficulties that cause the estimated Q models from all these methods to be unreliable. First, the Q tomography schemes in these studies are mostly ray-based, measuring the spectral changes over different lengths of raypaths, which makes them prone to errors and unrealistic results when multi-pathing exists in areas of complex overburden. Second, measurements in these works are all conducted in the data domain, which has a number of issues that can affect the accuracy of Q model building.

Specifically, diffractions and poor signal-to-noise ratio introduce large errors in Q estimation. In addition, crossed events are possible, which will introduce spectral interference.

Therefore, I produce a reliable Q model by using wave-equation migration Q analysis (WEMQA). This idea is similar to wave-equation migration velocity analysis (WEMVA) (Sava and Biondi, 2004; Biondi, 2006), which relates image perturbation with velocity perturbation using wave-equation based tomography. When complex wave propagation (i.e. multipathing) exists, wavefield-continuation methods yield better images than ray-based methods, due to their ability to handle multipathing of the reflected energy. Moreover, the data-space approach is vulnerable when the data are contaminated with noise, diffractions and crossing events; whereas the image-space approach is stable, because migration suppresses the noise and focuses the events. To some degree, the migration-based technique is also more efficient than the data-based technique, since it can be implemented in a target-oriented fashion and hence focus on the attenuated zone.

In this paper, I first present downward-continuation imaging with Q to provide the basis for image-based Q tomography. Second, I present a wave-equation Q tomography operator, which provides a direct mapping between the image perturbation and the change in the Q model. The image perturbation, which will be discussed afterwards, can be measured either by computing the difference between the attenuated image and the attenuation-free stacked image, or by computing the spectral variation of each window over depth and image points.

THEORY

Downward-continuation imaging with Q

Valenciano et al. (2011) presented a one-way viscoacoustic wave-equation approach that can accurately migrate data in media with attenuation. Their approach uses an extension of the Fourier Finite-Differences(FFD) algorithm (Ristow and Ruhl, 1994) to compensate for the effects of Q during migration. Due to the relatively large frequency dispersion for FFD, I extend the explicit split-step algorithm to migrate attenuated data.

In a linear attenuating medium, the one-way wave equation used for migration by wavefield continuation has the following phase-shift recursive solution:

$$P_{z+\Delta z}\left(\omega, k_x, k_y\right) = P_z\left(\omega, k_x, k_y\right) e^{\pm i k_z \Delta z}, \tag{1}$$

where P is the pressure, ω is the temporal frequency, k_x and k_y are the horizontal wavenumbers, and k_z is the vertical wavenumber. The plus and minus signs in the phase shift operator represent downgoing and upgoing waves, respectively.

The vertical wavenumber can be expressed by the following dispersion relation, which is often called the Single Square Root (SSR) equation:

$$k_z = \sqrt{\left(\omega \tilde{s}\right)^2 - \left|\mathbf{k}\right|^2}, \tag{2}$$

where $|\mathbf{k}| = \sqrt{k_x^2 + k_y^2}$; $\tilde{s}$ is the phase slowness, which becomes a complex number and can be given by following equation based on the nearly constant-Q model (Futterman, 1962),

and

$$\tilde{s}(\omega) = s_{\omega r}\left(1 - \frac{1}{\pi Q}\ln(\omega/\omega_r)\right)\left(1 + \frac{i}{2Q}\right), \tag{3}$$

where $s_{\omega r}$ is the slowness at a reference frequency ω_r. Since attenuation delays more at the lower frequencies than the higher frequencies, I will take the infinite frequency as the reference frequency (the Nyquist frequency in the real case).

The phase-shift migration operator described above is strictly valid for a subsurface model that varies only with depth. To extend the operator to handle laterally varying earth models, a simplified form of this SSR can be approximated by using Taylor expansion around the reference slowness $\tilde{s}_0$ and the reference quality factor Q_0:

$$k_z(s_{\omega r}, Q) = k_{z0}(\tilde{s}_0) + \omega\left(\tilde{s} - \tilde{s}_0\right), \tag{4}$$

where

$$\tilde{s}_0 = s_{\omega r 0}\left(1 - \frac{1}{\pi Q_0}\ln(\omega/\omega_r)\right)\left(1 + \frac{i}{2Q_0}\right). \tag{5}$$

Equation 4 describes split-step migration, whose accuracy can be improved by using more than one reference slowness and Q. In this modified scheme, called Extended Split-Step migration, multiple reference wavefields are generated for interpolation.

Wave-equation Q tomography operator

Wave-equation Q tomography is a non-linear inversion process that aims to find the Q model that minimizes the residual field in the image space. In general, the residual image (i.e. image perturbation), $\mathbf{\Delta I}$, is the difference between the background image computed with the current background Q model, and the 'target' image that will be fully discussed in the later section. By linear approximation, the image perturbation is related to the model perturbation $\mathbf{\Delta Q}$ by the wave-equation Q tomography operator $\mathbf{T}$:

$$\mathbf{\Delta I} = \mathbf{T}\mathbf{\Delta Q}. \tag{6}$$

The adjoint of this tomographic operator $\mathbf{T}^*$ backprojects the image perturbation into the Q model space, and the back-projected changes in the model space will be used as gradient directions to conduct a line-search in optimization schemes. The objective function is to minimize the L2 norm of the image perturbation. This backprojection can be expressed as

$$\mathbf{\Delta Q} = \mathbf{T}^*\mathbf{\Delta I}. \tag{7}$$

In this paper, I evaluate this wave-equation tomographic operator $\mathbf{T}$ in the shot-profile domain. Both source and receiver wavefields are downward continued in the shot-profile domain using the one-way wave equations (Claerbout, 1971):

$$\begin{cases} \left(\frac{\partial}{\partial z} + ik_z\right) D(\mathbf{x}, \mathbf{x}_s, \omega) = 0 \\ D(x, y, z = 0, \mathbf{x}_s, \omega) = f_s(\omega)\delta(\mathbf{x} - \mathbf{x}_s) \end{cases}, \tag{8}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + ik_z\right) U(\mathbf{x}, \mathbf{x}_s, \omega) = 0 \\ U(x, y, z = 0, \mathbf{x}_s, \omega) = P(x, y, z = 0, \mathbf{x}_s, \omega) \end{cases}, \tag{9}$$

where $D(\mathbf{x}, \mathbf{x}_s, \omega)$ is the source wavefield at the image point $\mathbf{x} = (x, y, z)$ with the source located at $\mathbf{x}_s = (x_s, y_s, 0)$; $U(\mathbf{x}, \mathbf{x}_s, \omega)$ is the receiver wavefield at the image point $\mathbf{x}$ with the source located at $\mathbf{x}_s$; $f_s(\omega)$ is the frequency dependent source signature, and $\overline{f_s(\omega)\delta(\mathbf{x} - \mathbf{x}_s)}$ defines the point-source function at $\mathbf{x_s}$, which serves as the boundary condition of equation 8; $P(x, y, z = 0, \mathbf{x}_s, \omega)$ is the recorded shot gather at $\mathbf{x}_s$, which serves as the boundary condition of equation 9; and k_z is the vertical wavenumber as shown in equation 2.

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

$$I(\mathbf{x}, \mathbf{h}) = \sum_{\mathbf{x}_s} \sum_{\omega} D(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) U(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega), \tag{10}$$

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

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

$$\begin{aligned} \Delta I(\mathbf{x}, \mathbf{h}) &= \sum_{\mathbf{x}_s} \sum_{\omega} \Big(\Delta D(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) \widehat{U}(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega) + \\ &\quad \widehat{D}(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) \Delta U(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega) \Big), \end{aligned} \tag{11}$$

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

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

$$\begin{aligned} \Delta D(\mathbf{x}, \mathbf{x}_s, \omega) &= \sum_{\mathbf{h}} \Delta I(\mathbf{x}, \mathbf{h}) \overline{\widehat{U}(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega)} \\ \Delta U(\mathbf{x}, \mathbf{x}_s, \omega) &= \sum_{\mathbf{h}} \Delta I(\mathbf{x}, \mathbf{h}) \overline{\widehat{D}(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega)}. \end{aligned} \tag{12}$$

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

$$\begin{cases} \left(\frac{\partial}{\partial z} + ik_z\right) \Delta D(\mathbf{x}, \mathbf{x}_s, \omega) = \left(-i \frac{\partial k_z}{\partial Q} \widehat{D}(\mathbf{x}, \mathbf{x}_s, \omega)\right) \Delta Q(\mathbf{x}) \\ \Delta D(x, y, z = 0, \mathbf{x}_s, \omega) = 0 \end{cases}, \tag{13}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + ik_z\right) \Delta U(\mathbf{x}, \mathbf{x}_s, \omega) = \left(-i \frac{\partial k_z}{\partial Q} \widehat{U}(\mathbf{x}, \mathbf{x}_s, \omega)\right) \Delta Q(\mathbf{x}) \\ \Delta U(x, y, z = 0, \mathbf{x}_s, \omega) = 0 \end{cases}, \tag{14}$$

When solving the optimization problem, I obtain the image perturbation by comparing the background image and the target image. Then, the perturbed image is convolved with the background wavefields to obtain the perturbed wavefields (equation 12). The scattered wavefields are computed by applying the adjoint of the one-way wave-equations 13 and 14. Finally, the perturbation $\Delta Q(\mathbf{x})$, is obtained by cross-correlating the upward-propagated scattered wavefields with the modified background wavefields,

$$\Delta Q(\mathbf{x}) = \sum_{x_s} \sum_{\omega} \overline{\Delta D} \left(-i \frac{\partial k_z}{\partial Q} \widehat{D} \right) + \overline{\Delta U} \left(-i \frac{\partial k_z}{\partial Q} \widehat{U} \right). \tag{15}$$

Image perturbation estimation

By definition, the image perturbation $\Delta\mathbf{I}$ is the difference between the background image $\mathbf{I}$, computed with the current background Q model, and the 'target' image, $\hat{\mathbf{I}}$, which is an attenuation-free image. In fact, instead of computing the difference between these two images, I calculate the spectral change of the images. The change in the spectrum can be indicated by the steepness of the slope computed by the `spectral ratio` method (Tonn, 1991), with a larger slope indicating that more frequencies are attenuated. As attenuation is a frequency-dependent amplitude effect, the intercept calculated from the `spectral ratio` method removes the frequency-independent amplitude effects, such as geometric spreading, instrument response, source/receiver coupling, radiation patterns, and reflection/transmission effects. Therefore, the perturbed image is measured from the perturbed amplitude spectral change after the absolute scaling of the waveform has been removed.

In this section, I develop ways of measuring the image perturbation for two scenarios: a model with sparse reflectors and a model with dense reflectors. The choice of the target image $\widehat{\mathbf{I}}$ is the main distinguishing element between these two different measurements of the perturbed image. Both of these methods require accurate velocity models. Fortunately, since both velocity and Q model building share the same tomography kernel, I can update these two models simultaneously.

Model with sparse reflectors

When the reflectors are sparsely distributed in the subsurface, I assume those reflectivities are white and can be picked from the background image. Therefore, the target image $\widehat{\mathbf{I}}$ in this method is the image of a non-attenuated dataset generated by a set of picked reflectivities. I design an automatic picker to pick each individual event for spectrum analysis and target-image generation. The picker automatically pinpoints the local maximum of the envelope of each trace and places a normalized reflectivity R on the reflector. During this procedure, thresholding helps choose the reflectors with relatively high amplitude. I synthesize a set of non-attenuated data by injecting the known source wavelet into a non-attenuated medium with reflectivity R. This data is then remigrated to obtain the non-attenuated target image $\widehat{\mathbf{I}}$. In the real case, not all of the reflectivities can be accurately picked. Therefore, only key reflectors are selected to generate the target image $\widehat{\mathbf{I}}$. An image mask $\mathbf{M}$ is needed to select the corresponding reflections in background image $\mathbf{I}$, and $\mathbf{M}$ is designed according to the energy of the target image $\widehat{\mathbf{I}}$.

Model with dense reflectors

For some real cases, events are very close to each other, making them difficult to individually separate. The interference among those events may introduce errors into spectral analysis. Therefore, instead of comparing the frequency contents of each event, I compare the spectral difference between a number of selected, windowed events from the background image with the events in the reference windows. These reference windows are carefully selected from the background image so that they are not contaminated by attenuation and can be used as our target image. All the windows in this method are large and the same size, hence the influence of the interfering reflectivies on the spectra are statistically the same over all windows. Based on the assumption that the amplitude spectra contain the same frequency content over the windows if the Q model for imaging is accurate, this method minimizes the spectral differences between the selected windows and the reference windows.

EXAMPLES

Numerical test of downward-continuation imaging with Q

To test the downward-continuation operator with Q, I run downward-continuation imaging with Q on a 2-D model. The model size is 4000 m (length) x 2500 m (depth). A horizontal reflector is at 1500 m depth, and 51 sources and 401 receivers are uniformly distributed on the surface. The medium is homogeneous with constant velocity (2000 m/s) and constant Q (50 for the model with attenuation and 99999 for the model without attenuation). A Ricker wavelet with 50 Hz central frequency is used as the source wavelet.

For conveniently comparing the wavelets in Figure 1, all plots are displayed in normalized wiggles, and the maximum amplitude of those wavelets before normalization are presented in the title. Figure 1(a) shows the conventional migration of the data generated from the model without attenuation, which images the reflector at 1500 m depth. Figures 1(b) and 1(c) show the conventional migration and the migration compensated by the true Q model of the data generated from the model with attenuation, respectively. Due to the loss of higher frequencies and velocity dispersion caused by attenuation, the wavelets in Figure 1(b) are stretched, amplitude-decayed and phase-rotated when compared with Figure 1(a). Figure 1(c) shows almost the same result as Figure 1(a). The small difference between Figure 1(c) and 1(a) shown in Figure 1(d) indicates that Q compensation adequately restores both the amplitude and phase.

Numerical test of the wave-equation Q tomography operator

Similar to the forward tomographic operator in WEMVA (Sava and Biondi, 2004; Biondi, 2006), which linearizes the image around the slowness, the forward wave-equation Q tomography operator relates the perturbation in Q to the image perturbation by linear approximation. This approximation may fail when the model error is large. To verify the effectiveness of this linear approximation, I compare the results generated by the linearized operator to the non-linearized image perturbation that is obtained by subtracting the background image from the true image. The size of the 2D test model is 4000 m (length) x 1200 m (depth). A horizontal reflector is at 900 m depth, and 51 sources and 401 receivers are uniformly distributed on the surface. The background medium is homogeneous with constant velocity (2000 m/s) and no attenuation ($Q = 10000$). Figure 4(a) shows the model

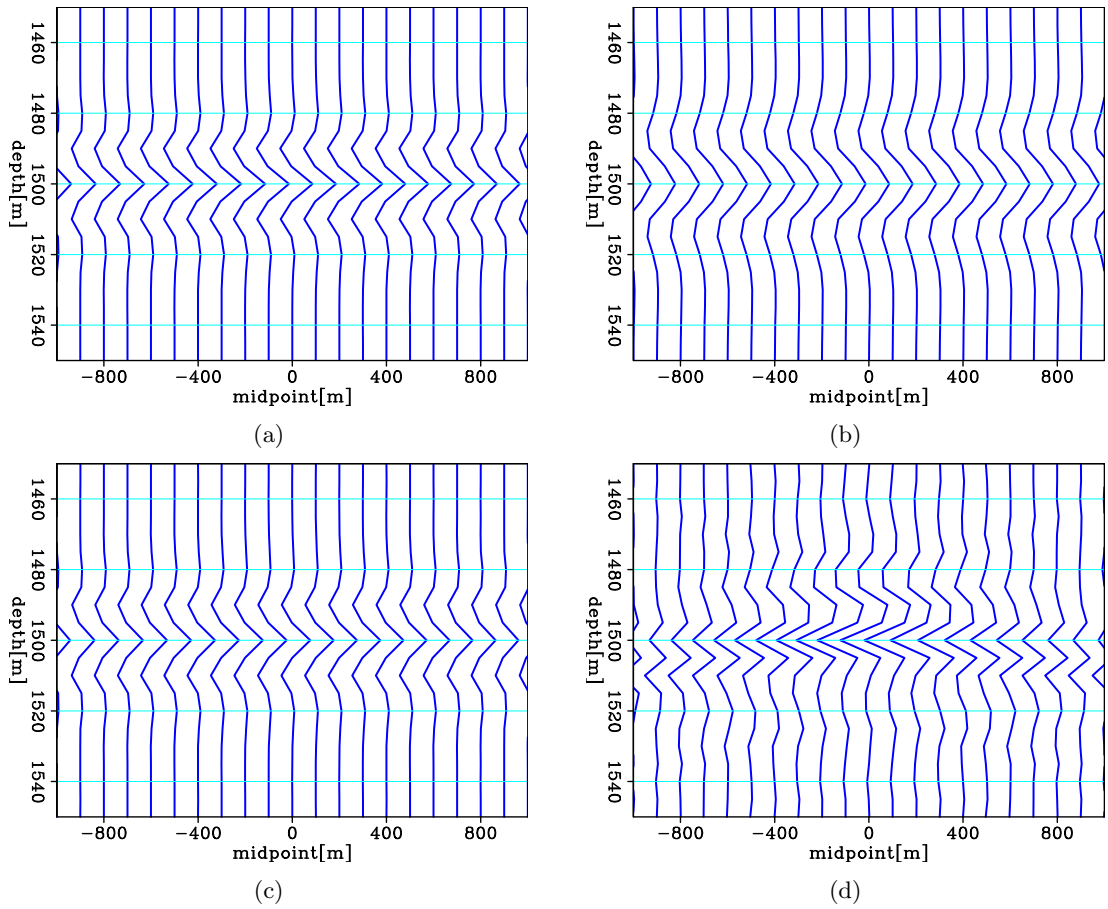
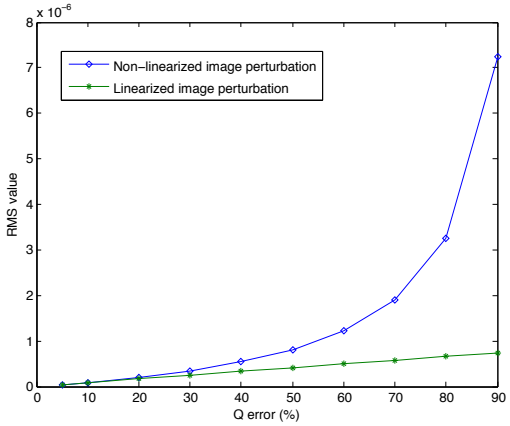


Figure 1: For convenient comparison of the wavelets, all plots are in normalized wiggles. (a) Conventional migration on non-attenuated data. The maximum amplitude of the event before normalization is 0.28. (b) Conventional migration on attenuated data. The maximum amplitude of the event before normalization is 0.0015. (c) Migration compensated by the true Q model on attenuated data. The maximum amplitude of the event before normalization is 0.27. (d) The difference between Figure 1(a) and Figure 1(c). The maximum amplitude of the event before normalization is 0.01. [**ER**] yishen/. wimage.smp.ctrl,wimage.smp.nqmig,wimage.smp.qmig,wimage.smp.diff

Figure 2: The linearized and the non-linearized image perturbation are compared by examining their RMS value. This figure shows their RMS values as ΔQ increases. The results show that the wave-equation Q tomography operator is a good linear approximation when Q error is less then 50%, but fails in approximating the non-linearized image perturbation when the model error increases. [**NR**] yishen/. err



perturbation with a rectangular low Q structure. The Q errors are set to 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90% lower than the background model.

I compare the linearized and the non-linearized image perturbation by examining their RMS values. Figure 2 shows their RMS values as ΔQ increases. The results show that this wave-equation Q tomography operator is a good linear approximation when Q error is less then 50%, but fails in approximating the non-linearized image perturbation when the model error increases.

To test the adjoint of the wave-equation Q tomography operator $\mathbf{T}^*$, I run it in a homogeneous background medium with $v = 2000$ m/s and $Q = 50$. The input of the adjoint tomographic operator is a spike in the image space $\Delta\mathbf{I} = \delta$ ($x = 0$ m, $z = 900$ m). Figures 3(a) and 3(b) show the back-projected Q gradient where the source-receiver offset is 0 km and 1.6 km, respectively. These back projections have a banana-donut shape and are spread along the wavepaths from the source to the perturbed image point and from the perturbed image point to the receiver.

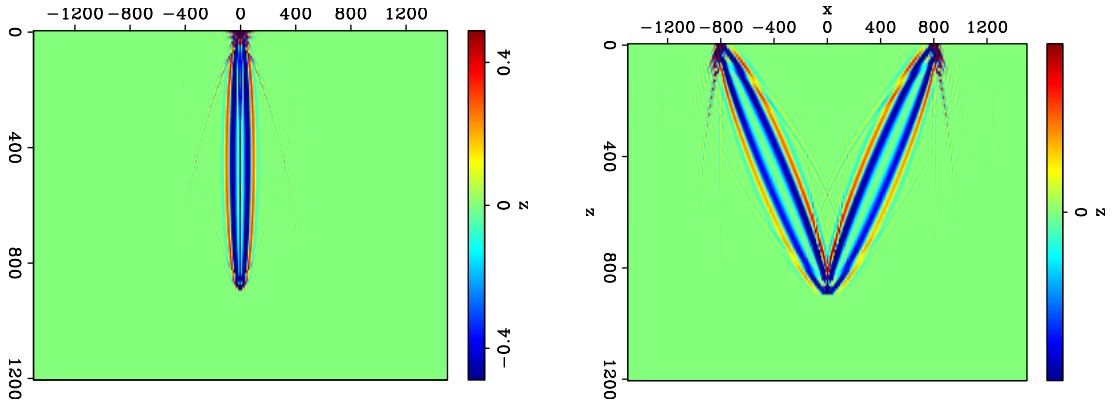


Figure 3: 2D impulse responses for Q: (a) zero-offset impulse responses; (b) impulse responses when source-receiver offset is 1.6 km. [**ER**] yishen/. sqr.sen.Dq.0,sqr.sen.Dq.h

Numerical tests on the model with sparse reflectors

This section applies WEMQA to two 2-D synthetic examples to update an inaccurate Q model. The size of the first test 2D model is 4000 m (length) x 1200 m (depth). A horizontal reflector is at 900 m depth, and 51 sources and 401 receivers are uniformly distributed on the surface. The background medium is homogeneous with constant velocity (2000 m/s) and constant Q ($Q = 50$). A Ricker wavelet with 50 Hz central frequency is used as the source wavelet. Figure 4(a) and Figure 4(b) show the true Q models, which have Q anomalies of 10% lower and higher than the background, respectively.

I model the data using the true Q model (the background plus the perturbation), and use the background Q model as the initial model. After migration on the attenuated data with the initial Q model, the reflectivities are automatically picked from the background image that serves as the input for the target image. Figures 4(c) and 4(d) show the image perturbations obtained by subtracting the background image from the target images. These

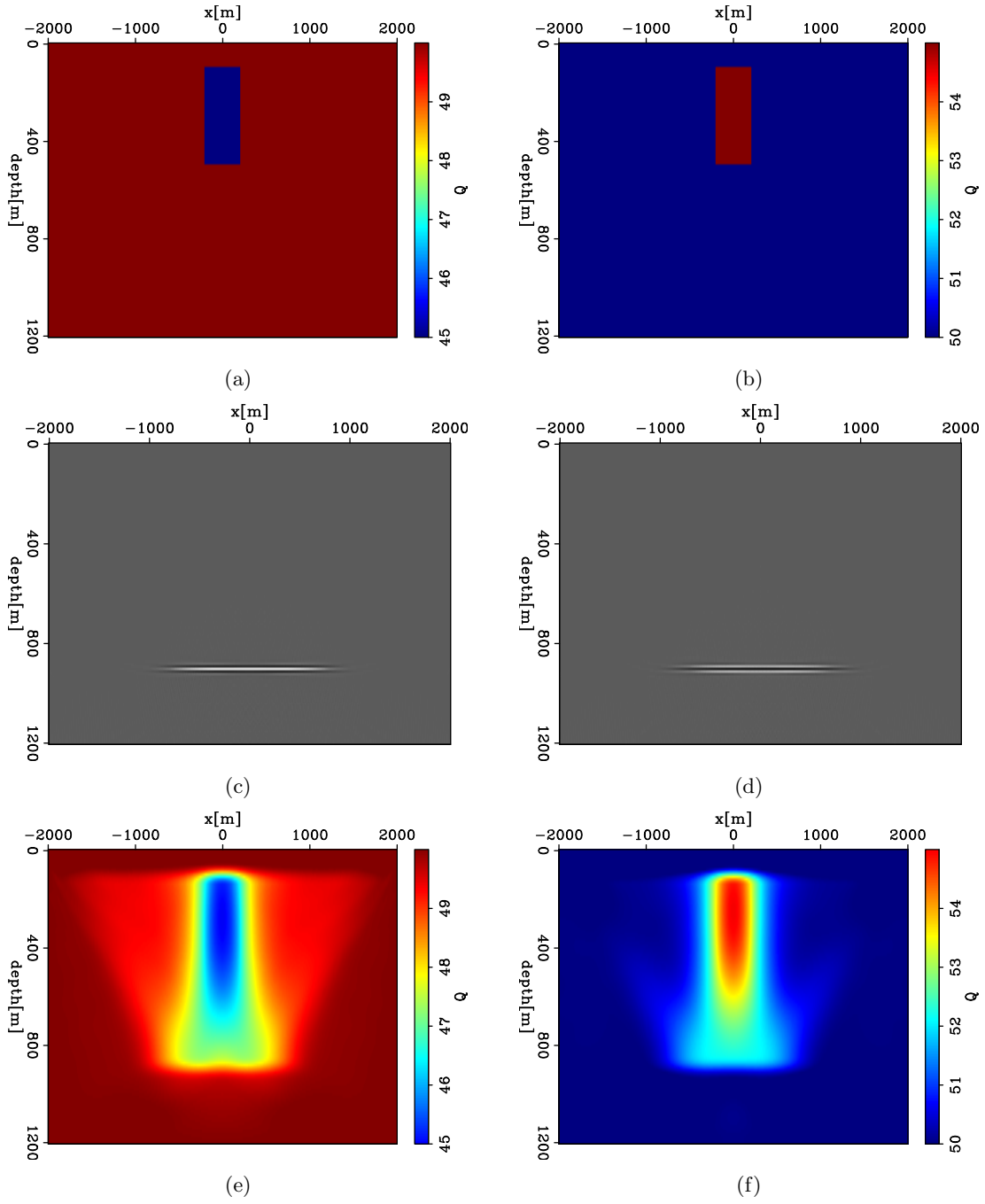


Figure 4: Q perturbations: (a) rectangular Q anomaly that is 10% lower than the background Q; (b) rectangular Q anomaly that is 10% higher than the background Q; (c) image perturbation caused by the perturbed model in Figure 4(a); (d) image perturbation caused by the perturbed model in Figure 4(b); (e) inversion results of the low-Q anomaly shown in Figure 4(a) using the steepest-descent algorithm after three iterations; (f) inversion results of the high Q anomaly shown in Figure 4(b) using the steepest-descent algorithm after three iterations. [**ER**]

yishen/. sqr.q.sml,sqr.q.lrg,sqr.dimg.sml,sqr.dimg.lrg,sqr.invBq3.sml,sqr.invBq3.lrg

images have larger perturbations underneath the Q anomaly than in the area far away from the anomaly. In addition, the sign of the amplitude of the middle lobe of the reflector in Figures 4(c) and 4(d) indicates the direction for updates. The positive amplitude of the middle lobe of the reflector in Figure 4(c) indicates the background image is undercompensated, and smaller Q is needed for the updates. The reverse is true for Figure 4(d). Figure 4(e) and 4(f) show the inversion results using steepest-descent algorithm after three iterations. The results show that both the shapes and the values of the Q anomalies are recovered.

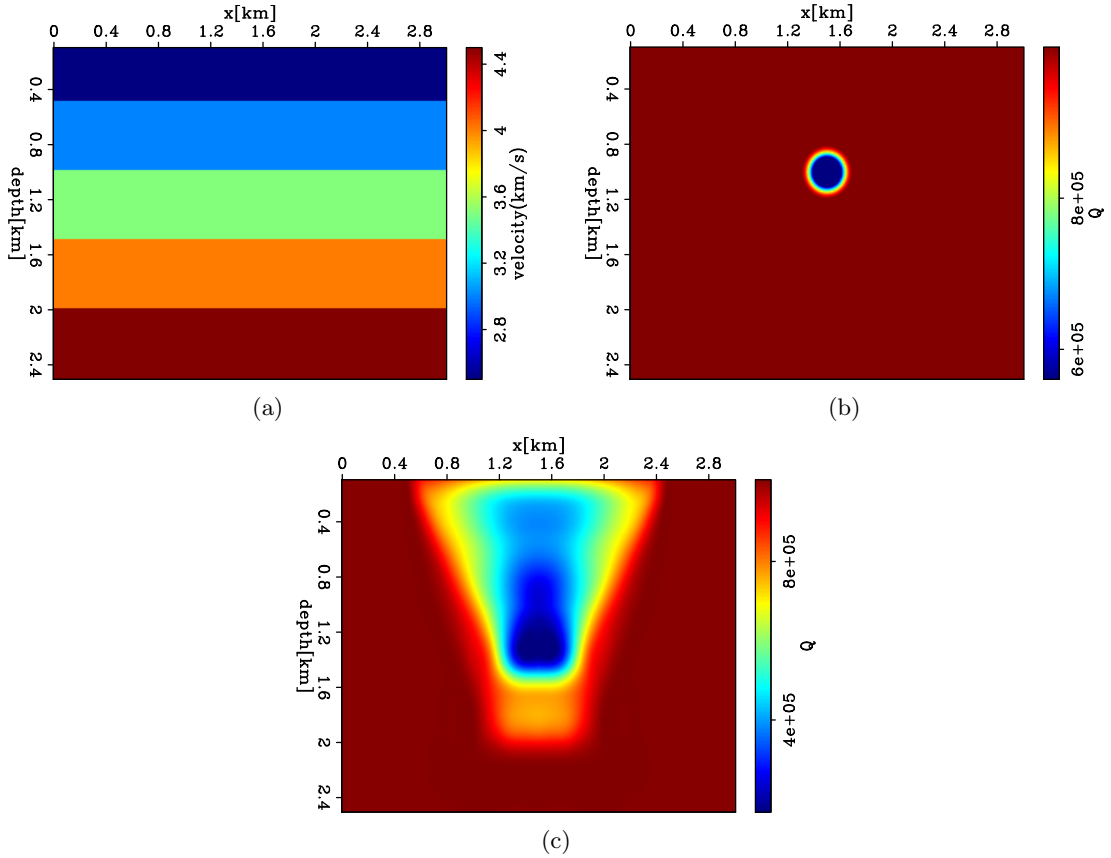


Figure 5: (a) True velocity model with multiple reflectors; (b) True Q model with a Gaussian anomaly in a non-attenuating medium. (c) Inverted Q model using the steepest-descent algorithm after four iterations. [**ER**] yishen/. gas.bvel,gas.q,gas.invBq4

The second test example is modeled on a realistic subsurface structure with an absorptive gas cloud. Figure 5(a) shows the velocity model with multiple reflectors, and Figure 5(b) presents a Gaussian Q anomaly in a non-attenuating medium. The data are modeled with a 3 km maximum offset, 10 m receiver spacing, 40 m source spacing, and a 50Hz Ricker wavelet. I use a non-attenuating medium as the initial model. Figure 5(c) shows the inversion results using the steepest-descent algorithm after four iterations. As the initial model is far from the true model, the low-Q anomaly is only partially recovered in Figure 5(c), and hence more iterations are needed to better retrieve the true model.

Numerical tests on the model with dense reflectors

A part of a modified SEAM velocity model is used in this section. This model includes two gas clouds with lower velocity than the surrounding sediments (Figure 6(a)). The Q model (in logarithmic scale) shown in Figure 6(b) also includes these two gas clouds with high attenuation. The left gas cloud is a Gaussian anomaly, and the shape of the right gas cloud aligns with the velocity model. I generate synthetic data with 51 shots with 0.2 km spacing, 251 receivers with 0.04 km spacing, and a Ricker wavelet with 50 Hz central frequency. The initial model for the inversion is a model without attenuation. After migration of the attenuated data with the current Q model, I generate the image perturbation by calculating the slope of the logarithm of the spectral ratio between the windowed events of each trace and the events in the reference window. The window size is 3 km, and 30 sliding windows are used for each trace. The reference windows are at x=6 km, and are not contaminated by attenuation. As neither the structures nor velocities in this example have strong horizontal variation, the compared windows and the reference windows are at the same depth. Figure 6(c) show the inversion results (in logarithmic scale) using the steepest-descent algorithm after four iterations. The results almost accurately indicate the locations of these two anomalies. The value of the left gas cloud shows larger difference from the true Q model than that of the right gas cloud, because the tapering region of the Gaussian anomaly is much larger than that of the right cloud. I need more iterations to better retrieve the anomalies.

CONCLUSION

I design a new method, WEMQA, to produce reliable Q models. This method is performed in the image space, which uses downward-continuation imaging with Q to stack out noise, to focus and simplify events, and to provide a direct link between the model perturbation and the image perturbation. I develop two methods to generate the image perturbation for the following scenarios: a model with sparse reflectors and a model with dense reflectors. This method uses wave-equation Q tomography to handle complex wave propagation. Two synthetic tests on two different 2-D models with Q anomalies show the effectiveness of this method on models with sparse events. Tests with a modified SEAM model also demonstrate the feasibility of this method for models with dense events.

ACKNOWLEDGMENTS

The authors thank Biondo Biondi, Robert Clapp and Dave Nichols of Stanford for their advice and suggestions, and also thank Yunyue Li, Ali Almomin and Sjoerd de Ridder for fruitful discussion.

REFERENCES

Best, A. I., C. MaCann, and J. Sothcott, 1994, The relationships between the velocities, attenuations, and petrophysical properties of reservoir sedimentary rocks: Geophysical Prospecting, **42**, 151–178.

Biondi, B. L., 2006, 3-D seismic imaging.

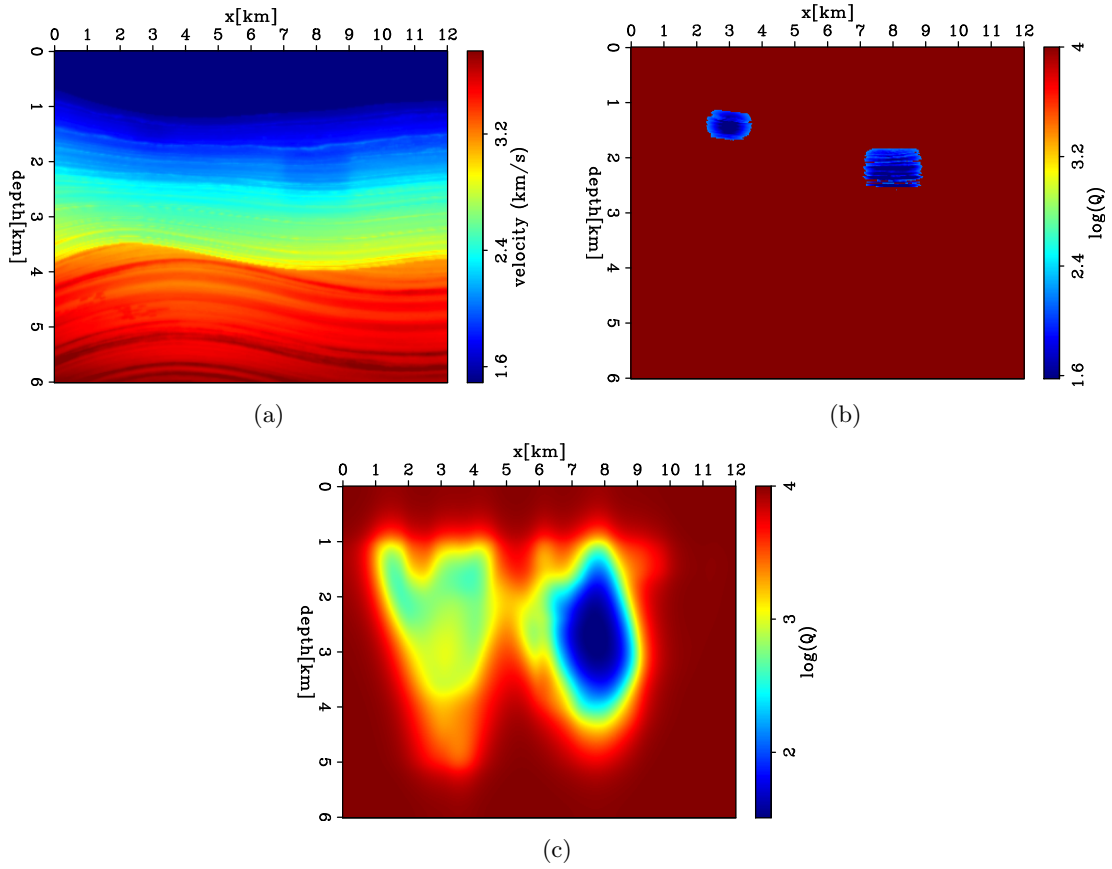


Figure 6: (a) A part of a modified SEAM velocity model with two gas clouds. ; (b) True Q model (in logarithmic scale) with two gas clouds. (c) Inverted Q model (in logarithmic scale) using the steepest-descent algorithm after four iterations. [**ER**] yishen/. seamlog.bvel,seamlog.q,seamlog.sdstep2.mmov.slice

Brzostowski, M. A. and G. A. McMechan, 1992, 3-D tomographic imaging of near-surface seismic velocity and attenuation: Geophysics, **57**, 396–403.

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

Dasgupta, R. and R. A. Clark, 1998, Estimation of Q from surface seismic reflection data: Geophysics, **63**, 2120–2128.

Futterman, W. I., 1962, Dispersive body waves: Journal of Geophysical Research, **67**, 5279–5291.

Kjartansson, E., 1979, Constant Q wave propagation and attenuation: Journal of Geophysical Research, **84**, 4737–4748.

Leaney, W. S., 1999, Walkaway Q inversion: 69th Annual International Meeting, SEG, Expanded Abstracts, 1311–1314.

Leggett, M., N. R. Goulty, and J. E. Kragh, 1992, Study of traveltime and amplitude time-lapse tomography using physical model data: Abstracts of 54th EAEG Meeting,, 248–249.

Mateeva, A., 2003, Thin horizontal layering as a stratigraphic filter in absorption estimation and seismic deconvolution: PhD thesis, Colorado School of Mines.

Plessix, R. E., 2006, Estimation of velocity and attenuation coefficient maps from crosswell seismic data: Geophysics, **71**, S235–S240.

Quan, Y. and J. M. Harris, 1997, Seismic attenuation tomography using the frequency shift method: Geophysics, **62**, 895–905.

Reine, C., R. A. Clark, and M. van der Baan, 2012a, Robust prestack Q-determination using surface seismic data: Part 1 Method and synthetic examples: Geophysics, **77**, R45–R56.

———, 2012b, Robust prestack Q-determination using surface seismic data: Part 2 3D case study: Geophysics, **77**, B1–B10.

Rickett, J., 2006, Integrated estimation of interval-attenuation profiles: Geophysics, **71**, A19–A23.

———, 2007, Estimating attenuation and the relative information content of amplitude and phase spectra: Geophysics, **72**, R19–R27.

Ristow, D. and T. Ruhl, 1994, Fourier finite-difference migration: Geophysics, **12**, 1882–1893.

Sava, P. and B. Biondi, 2004, Wave-equation migration velocity analysis-i: Theory: Geophysical Prospecting, **52**, 593606.

Tonn, R., 1991, The determination of seismic quality factor Q from VSP data: A comparison of different computational techniques: Geophysical Prospecting, **45**, 87–109.

Valenciano, A. A., N. Chemingui, D. Whitmore, and S. Brandsberg-Dahl, 2011, Wave equation migration with attenuation and anisotropy compensation: 2011 Annual Meeting, SEG, Expanded Abstracts, 232–236.

Zucca, J. J., L. J. Hutchings, and P. W. Kasameyer, 1994, Seismic velocity and attenuation structure of the geysers geothermal field, california: Geothermics, **23**, 111–126.

SEP3D Output—A ProMAX/SeisSpace Utility

Stewart A. Levin

ABSTRACT

In order to overcome the restrictions on trace header information inherent in SEG-Y or SEG-D, I have written a a ProMAX/SeisSpace module to write directly into SEP3D format, preserving all numeric-valued trace headers. One feature of this module is that it does not need any SEP software environment in order to operate.

INTRODUCTION

Almost all of the field data at the Stanford Exploration Project arrives in a form created by and compatible with commercial seismic processing software. When those seismic files do not contain key information such as coordinates in their headers, they are supplemented by auxiliary files in formats such as SPS or P1-90 detailed by the SEG Technical Standards Committee[1] that do contain that information. In that event, the path of least resistance is to import the seismic and auxiliary files into a commercial system, in the case of the Stanford Exploration Project, the ProMAX/SeisSpace software, and subsequently export the fully populated data to, say, SEG-Y so that it can be read and stored in the SEP3D internal format. Because SEG-Y headers often cannot carry all the information that one might like to export, I created a ProMAX/SeisSpace module to write directly into SEP3D format, preserving all numeric-valued trace headers.

[1] http://www.seg.org/resources/publications/misc/technical-standards

DOCUMENTATION

SEP3D Output

SEP3D Output converts *ProMAX* data to a format developed by the Stanford Exploration Project.

Theory

SEP3D datasets consist of up to six parts:

- Text file containing global parameters such as sample rate (.H)
- Binary trace data file containing the seismic samples (.H@)
- Text file describing the trace header entries (.H@@)
- Binary trace header file containing a header for each trace (.H@@@)
- Text file describing the format of a trace binning grid (.H@@@@)
- Binary grid file with binning locations for each trace (.H@@@@@)

SEP3D Output currently creates the first four files, assuming that ProMAX input data are already a regularly sampled 3D volume. Should this not be the case, an output dataset will still be created and the user may subsequently employ the Stanford Exploration Project "SEPlib" software library tools for binning within that system.

Usage

The locations of the components of a SEP3D dataset are normally split between two directories, one for textual information and the other for the binary traces and trace headers. The binary data directory is known in the SEP system as the *datapath*, and it can be selected via a runtime command line *datapath=* argument, a *.datapath* file in the current or home directory, or a *DATAPATH* environment variable. For portability to non-SEP environments, **SEP3D Output** requires the two directories to be specified in the setup menu, but will honor the *DATAPATH* environment variable if present.

Parameters

Text file directory
Enter the name of the directory where you would like the two text files to be created.

Browse for path to text file directory
Select the directory where the .H and .H@@ text files will be stored.

Data file directory
Enter the name of the directory where you would like the two binary data files to be created.

Browse for path to data file directory
Select the directory where the .H@ and .H@@@ binary data files will be stored.

File name prefix
Enter the prefix you want to use for naming the output files. For example, if the prefix is *MyData*, then the output files will be named *MyData.H*, *MyData.H@*, *MyData.H@@*, and *MyData.H@@@*.

Verbose output?
Select **Yes** for more detailed information in the job output log.

DISCUSSION

Because the output volume dimensions are initially set from the internal ProMAX number of samples per trace (numsmp), traces per record (maxdtr), and number of records encountered, there may well be a discrepancy between these apparent dimensions of the output data and the actual number of traces. SEP3D Output detects this situation and resets the output dimensions to *n2=actual number of traces* and *n3=1*. For this reason, SEP users are urged to check these values, either by reading the *.H* file directly or using the SEPlib tool *In*, and then employ appropriate seplib tools (e.g. Sort3d, Fold3d, Grid_fold) to manage the actual geometry and binning.

PhD genealogy of Jon Claerbout; Ancestry and legacy.

Sjoerd de Ridder

ABSTRACT

PhD genealogy is the practice of tracing ones thesis adviser's adviser, and so on. To first order this provides a lineage of academic teaching and thinking. For the Stanford Exploration Project's 40th anniversary, I compiled Jon Claerbout's academic lineage. The SEP is one of the world's most infamous academic research groups in seismic imaging, and Jon Claerbout graduated many PhD students through this project himself. This report goes beyond the usual academic genealogy in that it also attempts to compile an academic legacy for Jon Claerbout. An online, and up-to-date copy of this academic genealogy, will be made available online.

INTRODUCTION

This paper outlines the PhD ancestry and legacy of Jon Claerbout, founder of the Stanford Exploration Project in 1973. First attempts to study this lineage started during my summer internship at BP in 2010, and the earliest generations were easily traced through the websites of the Massachusetts Institute of Technology. With the help of Wikipedia and various other university libraries, we trace the lineage to a variety of fields: geology, physics, mathematics, philosophy, theology, medicine and more. The website of the Mathematics PhD genealogy project[1] proved particularly resourceful.

ANCESTRY

Jon received his PhD under Theodore Madden, himself a graduate of MIT. Several generations earlier we find the famous physicist Robert Millikan, conductor of the famous "oil drop" experiment and a Nobel laureate, working at Caltech. Robert Millikan graduated with Albert Michelson, another Nobel laureate, and Michael Pupin at Columbia University. Another line traces to the geology department at the University of Wisconsin. This line started with Roland Irving, whose academic roots are unclear at this time. Irving's arrival at Wisconsin marked the start of the geology program, one of the strongest of the nation. Irving is also the earliest known academic adviser found in Jon's PhD genealogy in the new world. Both Albert Michelson and Michael Pupin graduated at the University of Berlin, with none other than the famous Hermann von Helmholtz, whose last name is honored in the Helmholtz equation.

The earliest roots are not in fields of physics or mathematics at all, but rather in medicine, philosophy and theology. Studies at the time were less rigid; studies in philosophy would easily combine with work in mathematics. The PhD degree in its modern

[1] http://genealogy.math.ndsu.nodak.edu/

form did not exist and was not widely awarded either. The mathematics genealogy project traces student-teacher relationships in the absence of degrees. In this lineage we find Jakob Thomasius. At the University of Leipzig, Thomasius counted Gottfried Leibniz, inventor of calculus, among his students. Thomasius himself graduated with Gottfried Leibniz's father, Friedrich Leibniz. Another notable academic was Johannes Argyropoulos, who obtained a degree in theology from the University of Padua. Argyropoulos later taught at the University of Florence where he is said to have found Leonardo da Vinci attending his lectures.

The earliest academic roots of Jon trace to Nilos Kabasilas, a fourteenth-century bishop of Thessalonika who was passionate about the philosophies of Saint Thomas.

I Jon F. Claerbout 1967 Massachusetts Institute of Technology, Electromagnetic Effects of Atmospheric Gravity Waves. Advisor: Theodore R. Madden (II).

II Theodore R. Madden, 1961 Massachusetts Institute of Technology, Electrode polarization and its influence on the electrical properties of mineralized rocks. Advisor: Stephen M. Simpson (III).

III Stephen M. Simpson, 1953, Massachusetts Institute of Technology, Statistical approaches to certain problems in geophysics. Advisor: Patrick M. Hurley (IV).

IV Patrick M. Hurley, 1940 Massachusetts Institute of Technology, Investigations on the helium method of age determination. Advisors: Robley D. Evans (Va) and Warren J. Mead (Vb).

Va Robley D. Evans, California Institute of Technology. Advisor: Robert A. Millikan (VIa).
Vb Warren J. Mead, University of Wisconsin. Advisor: Charles K. Leith (VIb).

VIa Robert A. Millikan, Colombia University. Advisor: Michael I. Pupin (VIIa) and Albert Michelson (VIIb).
VIb Charles K. Leith, 1901 University of Wisconsin, Rock Cleavage. Advisor: Charles R. Van Hise (VIIIb).

VIIa Michael I. Pupin, University of Berlin. Advisor: Hermann von Helmholtz (VIIIa).
VIIb Albert A. Michelson, University of Berlin. Advisor: Hermann von Helmholtz (VIIIa).

VIIIa Hermann von Helmholtz, Royal Friedrich-Wilhelm Institute. Advisor: Johannes Peter Müller (IXa).
VIIIb Charles R. Van Hise, 1892 University of Wisconsin. ('USGS Monograph 19 on field investigations of the PenokeeGogebic Iron Range' might have been his thesis). Advisor: Roland D. Irving (IXb).

IXa Johannes Peter Müller, Bonn University. Advisor: Philipp Franz von Walther (Xa) and Karl Rudolphi (Xb).
IXb Roland D. Irving graduated from the Columbia University School of Mines in New York and arrived at the University of Wisconsin in 1870. Irving's initial appointment, marking the beginning of the geology program, was in Mining and Metallurgy. In 1878 Irving became professor in a newly created Department of Mineralogy and Geology.

Xa Philipp Franz von Walther, University of Landshut. Advisor: Georg Joseph Beer

(XIa).
Xb Karl Rudolphi, University of Greifswald. Advisor: Christian von Weigel (XIb).

XIa Georg Joseph Beer, University of Vienna. Advisor: Joseph Barth (XIIa).
XIb Christian von Weigel, University of Göttingen. Advisor: Johann Christian Erxleben (XIIb).

XIIa Joseph Barth, University of Vienna. Advisor: Anton von Störck (XIIIa).
XIIb Johann Christian Erxleben, University of Göttingen. Advisor: Abraham Gotthelf Kästner (XIIIb).

XIIIa Anton von Störck, University of Vienna. Advisor: Gerard van Swieten (XIVa).
XIIIb Abraham Gotthelf Kästner, University of Leipzig. Avdisor: Christian August Hausen (XIVb).

XIVa Gerard van Swieten, University of Leyden. Advisor: Hermann Boerhaave (XVa).
XIVb Christian August Hausen, University of Wittenberg. Advisor: Johann Christoph Wichmannshausen (XVb).

XVa Hermann Boerhaave, Universiteit van Leiden. Advisor: Burchard de Volder (XVIa).
XVb Johann Christoph Wichmannshausen, University of Leipzig. Advisor: Otto Mencke (XVIb).

XVIa Burchard de Volder, University of Basel. Advisor: Franciscus de le Boë (XVIIa).
XVIb Otto Mencke, University of Leipzig. Advisor: Jakob Thomasius (XVIIb).

XVIIa Franciscus de le Boë, University of Basel. Advisor: Emmanuel Stupanus (XVIIa).
XVIIb Jakob Thomasius, University of Leipzig. Advisor: Friedrich Leibniz (XVIIIb).

XVIIIa Emmanuel Stupanus, University of Basel. Advisor: Petrus Ryff (XIX).
XVIIIb Friedrich Leibniz, 1622 University of Leipzig. Disputatio de Casibus perplexis in Jure.

XIX Petrus Ryff, 1584 University of Basel, De Affectibus cordis. Advisor: Theodor Zwinger (XX).

XX Theodor Zwinger, 1553 The College of France; MD at University of Padua 1559. Advisors: Petrus Ramus (XXIa) (The College of France), Bassiano Landi (XXIb) (University of Padua) and Vittore Trincavelli (XXIc) (University of Padua).

XXIa Petrus Ramus, 1536 The College of Navarre, Quaecumque ab Aristotele dicta essent, commentitia esse. Advisors: Johannes Sturmius (XXIIa) and Jacques Toussain (XXIIb).
XXIb Bassiano Landi,1542 MD University of Padova. Advisors: Giovanni Battista della Monte (XXIIc) and Vittore Trincavelli (XXIId).

XXIIa Johannes Sturmius, 1527 Catholic University of Leuven. Advisors: Nicolas Clnard (XXIIIa), Johannes Winter von Andernach (XXIIIb).
XXIIb Jacques Toussain, 1521 University of Paris. Advisor: Guillaume Budé (XXIIIc).
XXIIc Giovanni Battista della Monte, University of Padova; MD University of Ferrara. Advisors: Marco Musuro (XXIIId) (University of Padova), Pietro Pomponazzi (XXIIIe)

(University of Padova) and Niccolò Leoniceno (XXIIIf) (University of Ferrara).
XXIId Vittore Trincavelli, MD University of Padova. Advisor: Pietro Pomponazzi (XXIIIe).

XXIIIa Nicolas Clénard, 1515 Catholic University of Leuven; 1521 Theol. Dr. at Catholic University of Leuven. Advisors: Jacobus Latomus (XXIVa) and Jan van Campen (XXIVb).
XXIIIb Johannes Winter von Andernach, 1527 MD at Catholic University of Leuven; 1532 The College of Tréguier. Advisors: Rutger Rescius (XXIVc) and Jacobus Sylvius (XXIVd).
XXIIIc Guillaume Budé, 1486 University of Orléans; 1491 University of Paris. Advisors: Georgius Hermonymus and Janus Lascaris (XXIVe).
XXIIId Marco Musuro, 1486 University of Firenze. Advisor: Janus Lascaris (XXIVe).
XXIIIe Pietro Pomponazzi, 1487 University of Padova. Advisors: Nicoletto Vernia (XXIVf) and Pietro Roccabonella (XXIVg).
XXIIIf Niccolò Leoniceno, 1446 University of Vicenza; 1453 MD University of Padova. Advisor: Ognibene Bonisoli da Lonigo (XXIVh) (University of Vicenza), Pelope (University of Padova) and Pietro Roccabonella (XXIVg) (University of Padova).

XXIVa Jacobus Latomus, 1502 The College of Montaigu; 1519 Theol. Dr. at Catholic University of Leuven. Advisors: Jan Standonck (XXVa) (The College of Montaigu).
XXIVb Jan van Campen, Theol. Dr. at Catholic University of Leuven; 1519 University of Ingolstadt. Advisors: Matthaeus Adrianus and Johann Reuchlin (XXVb).
XXIVc Rutger Rescius, 1513 at University of Paris. Advisor: Girolamo Aleandro (XXVc).
XXIVd Jacobus Sylvius, University of Paris; 1530 MD University of Montpellier. Advisors: Franois Dubois (XXVd) and Jean Tagault.
XXIVe Janus Lascaris, 1472 University of Padova. Advisors: Basilios Bessarion (XXVIIa) and Demetrios Chalcocondyles (XXVe).
XXIVf Nicoletto Vernia, 1472 University of Padova. Advisor: Gaetano da Thiene.
XXIVg Pietro Roccabonella, MD University of Padova. Advisor: Gaetano da Thiene (University di Padova) and Sigismondo Polcastro (University di Padova).
XXIVh Ognibene Bonisoli da Lonigo, University of Mantova. Advisor: Vittorino da Feltre (XXVIIc).

XXVa Jan Standonck, 1474 The College of Sainte-Barbe; 1490 The College of Montaigu.
XXVb Johann Reuchlin, 1477 University of Basel; 1481 University of Poitiers. Advisors: Johannes Argyropoulos (XXVIa) and Jacob ben Jehiel Loans.
XXVc Girolamo Aleandro, 1499, 1508 Theol. Dr. Universit di Padova. Advisors: Moses Perez and Scipione Fortiguerra (XXVIb).
XXVd Franois Dubois, 1516 University of Paris.
XXVe Demetrios Chalcocondyles, Mystras; Roman Academies. Advisor: Theodoros Gazes (XXVIc).

XXVIa Johannes Argyropoulos, 1444 University of Padua; 1481 Theol. Dr. University of Poitiers. Advisor: Basilios Bessarion (XXVIIa).
XXVIb Scipione Fortiguerra, 1493 University of Firenze. Advisor: Angelo Poliziano (XXVIIb).
XXVIc Theodoros Gazes, 1433 Constantinople; University of Mantova. Advisor: Vit-

torino da Feltre (XXVIIc).
XXVIIa Basilios Bessarion, 1436 Mystras. Advisor: Georgios Plethon Gemistos (XXIXa).
XXVIIb Angelo Poliziano, 1477 University of Firenze. Advisor: Marsilio Ficino (XXIXb) and Cristoforo Landino.
XXVIIc Vittorino da Feltre, 1416 University of Padova. Advisor: Guarino da Verona (XXIXc).

XXIXa Georgios Plethon Gemistos, 1380, 1393, Nómoi (Book of Laws). Advisor: Demetrios Kydones (XXXa) and Elissaeus Judaeus.
XXIXb Marsilio Ficino, 1462 University of Firenze. Advisor: Johannes Argyropoulos.
XXIXc Guarino da Verona. Advisor: Manuel Chrysoloras (XXXb).

XXXa Demetrios Kydones. Advisor: Nilos Kabasilas (XXXI).

XXXb Manuel Chrysoloras. Advisor: Demetrios Kydones (XXXa).

XXXI Nilos Kabasilas, a fourteenth-century bishop of Thessalonika who was passionate about the philosophies of Saint Thomas.

LEGACY

Here I attempt to compile an academic legacy of Jon Claerbout. By no means can this legacy be considered complete or flawless, and it is not meant to establish an authoritative or exhaustive document on adviser-advisee relationships of any of the people mentioned. The vast majority of Jon's SEP students were trained in the SEP. Some of his students went on to advise and graduate many students themselves who in turn went on to have their own advisees.

I Jon F. Claerbout
1967 Massachusetts Institute of Technology, Electromagnetic Effects of Atmospheric Gravity Waves.
Graduates at Stanford University:

1. Allen Stratton?
2. Tom E. Landers, 1971 Stanford University, Elastic waves in laterally inhomogeneous media.
3. Lee Lu, 1972 Stanford University, The relation of local pressure fluctuations to large-scale meteorology and the simulation of acoustic-gravity waves in inhomogeneous media.
4. Donald C. Riley, 1974 Stanford University, Wave Equation Synthesis and Inversion of Diffracted Multiple Seismic Reflections.
5. Stephen M. Doherty, 1975 Stanford University, Structure Independent Seismic Velocity Estimation.
6. John Parker Burg, 1975 Stanford University, Maximum Entropy Spectral Analysis.
7. Philip S. Schultz, 1976 Stanford University, Velocity Estimation by Wave Front Synthesis.
8. Raul Estevez, 1977 Stanford University, Wide-Angle Diffracted Multiple Reflections.
9. Robert J. Godfrey, 1979 Stanford University, A Stochastic Model for Seismogram Analysis.
10. Ozgodan Yilmaz, 1979 Stanford University, Pre-Stack Partial Migration.
11. William C. Gray, 1979 Stanford University, Variable Norm Deconvolution.

12. Walter S. Lynn, 1979 Stanford University, Velocity Estimation in Laterally Varying Media.
13. Heloise Bloxsom Lynn, 1979 Stanford University, Migration and Interpretation of Deep Crustal Seismic Reflection Data.
14. Einar Kjartansson, 1979 Stanford University, Attenuation of Seismic Waves and Applications in Energy Exploration.
15. Robert W. Clayton, follows under (IIa).
16. Laurence C. Morley, 1981 Stanford University, Predictive Techniques for Marine Multiple Suppression.
17. Alfonso Gonzalez-Serrano, 1982 Stanford University, Wave Equation Velocity Analysis.
18. Richard A. Ottolini, 1982 Stanford University, Migration of Reflection Seismic Data in Angle-Midpoint Coordinates.
19. Allan Jacobs, 1982 Stanford University, The Pre-Stack Migration of Profiles.
20. Ira David Hale, follows under (IIb).
21. Jeffrey R. Thorson, 1984 Stanford University, Velocity Stack and Slant Stack Inversion Methods.
22. John L. Toldi, 1985 Stanford University, Velocity Analysis Without Picking.
23. Daniel H. Rothman, follows under (IIc).
24. Joshua M. Ronen, 1985 Stanford University, Multichannel Inversion in Reflection Seismology.
25. William S. Harlan, 1986 Stanford University, Signal/Noise Separation and Seismic Inversion
26. Zhiming Li, 1986 Stanford University, Imaging Steep-Dip Reflections by the Linearly Transformed Wave Equation Method.
27. Peter Mora, follows under (IId).
28. Kamal Mansour Al-Yahya, 1987 Stanford University, elocity Analysis by Iterative Profile Migration
29. Stewart A. Levin, 1987 Stanford University, Deconvolution with Spatial Constraints.
30. Charles Hege Sword, 1987 Stanford University, Tomographic Determination of Interval Velocities from Reflection Seismic Data: The Method of Controlled Directional Reception.
31. Paul J. Fowler, 1988 Stanford University, Seismic Velocity Estimation Using Prestack Time Migration.
32. Marta Jo Woodward, 1989 Stanford University, Wave Equation Tomography.
33. Clement Kostov, 1990 Stanford University, Multichannel Seismic Experiment with a Drill-Bit Source.
34. Biondo L. Biondi, follows under (IIe).
35. Johannes A. van Trier, 1990 Stanford University, Tomographic Determination of Structural Velocities from Depth-migrated Seismic Data.
36. John T. Etgen, 1990 Stanford University, Residual Prestack Migration and Interval-Velocity Estimation.
37. Joe A. Dellinger, 1991 Stanford University, Anisotropic Seismic Wave Propagation.
38. Carlos A. Cunha Filho, 1992 Stanford University, Elastic Modeling and Migration in Earth Models.
39. Lin Zhang, 1992 Stanford University, Imaging by the Wave Front Propagation Method,
40. Reinaldo J. Michelena, 1992 Stanford University, Anisotropic Traveltime Tomography.
41. Dave Nichols, 1994 Stanford University, Imaging in Complex Structures Using Band-limited Green's Functions.

42. Martin Karrenbach, 1995 Stanford University, Elastic Tensor Wave Fields.
43. Alexander M. Popovici, 1995 Stanford University, Migration to Zero Offset in Variable Velocity Medium.
44. Stephen P. Cole, 1995 Stanford University, Passive Seismic and Drill-Bit Experiments Using 2-D Arrays.
45. Dimitri Bevc, 1995 Stanford University, Imaging Under Rugged Topography and Complex Velocity Structure.
46. Raymond L. Abma, 1995 Stanford University, Least-Squares Separation of Signals and Noise Using Multidimensional Filters.
47. Jun Ji, 1995 Stanford University, Sequential Seismic Inversion Using Plane Wave Synthesis.
48. David E. Lumley, follows under (IIf).
49. Christine Ecker, 1998 Stanford University, Seismic Characterization of Methane Hydrate Structures.
50. Matthias Schwab, 1998 Stanford University, Enhancement of Discontinuities in seismic 3-D images usinga Java estimation library.
51. Nizar Chemingui, 1999 Stanford University, Imaging irregularly sampled 3D prestack data.
52. Sean Crawley, 2000 Stanford University, Seismic trace interpolation with nonstationary prediction-error filters.
53. Robert G. Clapp, 2000 Stanford University, Geologically constrained migration velocity analysis.
54. Sergey Fomel, follows under (IIg).
55. James Rickett, 2001 Stanford University, Spectral factorization of wavefields and wave operators.
56. Antoine Guitton, 2005 Stanford University, Multidimensional seismic noise attenuation.
57. Jesse Lomask, 2006 Stanford University, Seismic volumetric flattening and segmentation.
58. William Curry, 2008 Stanford University, Interpolation with prediction-error filters and training data.

IIa Robert W. Clayton

1981 Stanford University, Wavefield Inversion Methods for Refraction and Reflection Data.
Graduates at California Institute of Technology:

1. John Fawcett, 1984 California Institute of Technology with Herb Keller, I. Three dimensional ray-tracing and ray-inversion in layered media. II. Inverse scattering and curved ray tomography with applications to seismology.
2. Marianne Walck, 1984 California Institute of Technology, Teleseismic array analysis of upper mantle compressional velocity structure.
3. Thomas Hearn, follows under (IIIa).
4. Eugene Humphreys, follows under (IIIb).
5. Ronan Le Bras, 1985 California Institute of Technology, Methods of multiparameter inversion of seismic data using the acoustic and elastic born approximations.
6. John Vidale, follows under (IIIc).
7. John Louie, follows under (IIId).
8. Phyllis Ho, 1987 California Institute of Technology with Hiroo Kanamori, Attenuation tomography. Modelling regional love waves: Imperial Valley to Pasadena.
9. Christof Stork, 1988 California Institute of Technology, Ray trace tomographic velocity analysis of surface seismic reflection data.

10. Hua-Wei Zhou, follows under (IIIe).
11. Olafur Gudmundsson, 1989 California Institute of Technology, Some problems in global tomography: modeling the core-mantle boundary and statistical analysis of travel-time data.
12. J. Huw Davies, follows under (IIIf).
13. Robert W. Graves, 1990 California Institute of Technology, Modeling seismic wave propagation using paraxial extrapolators.
14. Bruce Worden, 1991 California Institute of Technology, Interactive seismic imaging on a multicomputer and application to the Hosgri fault.
15. Linda Rowan, 1993 California Institute of Technology with Tom Ahrens, Equation of state of molten mid-ocean ridge basalt. Structure of Kilauea volcano.
16. Leo Eisner, 2001 California Institute of Technology, Reciprocity method in seismology.
17. Julie J. Nazareth, 2002 California Institute of Technology with Egill Hauksson, The structure of the crust and distribution of earthquakes in Southern California.
18. William Keller, 2003 California Institute of Technology with Joann Stock, Cenozoic plate tectonic reconstructions and plate boundary processes in the Southwest Pacific.
19. Patricia Persaud, 2003 California Institute of Technology with Joann Stock, Images of early continental breakup in and around the Gulf of California and the role of basal shear in producing wide plate boundaries.
20. Zhimei Yan, 2006 California Institute of Technology, Regional mapping of the crustal structure in Southern California.
21. Nathan Downey, 2008 California Institute of Technology with J. Stock and M. Gurnis, Tectonic history of the Osbourn spreading center and dynamic subsidence of the Congo basin.
22. Nathalie Vriend, 2009 California Institute of Technology with M. Hunt and C. Brennen, Booming Sand Dunes.
23. Sonja Spasojevic, 2010 California Institute of Technology with M. Gurnis, Dynamics of long-term sea-level change and vertical motion of continents.
24. Michelle Selvans, 2011 California Institute of Technology with J. Stock and O. Aharonson, Analysis of Sonobuoy data to determine crustal structure of the Adare Trough, Antarctica.
25. YoungHee Kim, follows under (IIIg).
26. Ting Chen, 2012 California Institute of Technology with N. Lapusta, Part I: Structure of central and southern Mexico from velocity and attenuation tomography; Part II: Physics of small repeating earthquakes.
27. Kristin Phillips-Alonge, 2012 California Institute of Technology, Receiver Function Studies of Southern Peru From Seismic Array Data.

Current students at California Institute of Technology:

1. Steve Skinner (with J. Stock)
2. Vanessa Heckman (with T. Heaton)
3. Sara Dougherty (with D. Helmberger)
4. Yiran Ma
5. Dunzhu Li

IIb Dave Hale

1983 Stanford University, Dip-moveout by Fourier transform.

Graduates at Colorado School of Mines:

1. Yong Ma, 2012 Colorado School of Mines, Waveform-based velocity estimation from reflection seismic data.

Current students at Colorado School of Mines:

1. Simon Luo
2. Xinming Wu
3. Stefan Compton
4. Andrew Munoz

IIc Daniel H. Rothman

1985 Stanford University, Large Near-Surface Anomalies, Seismic Reflection Data, and Simulated Annealing.

Graduates at Massachusetts Institute of Technology:

1. Andrew K. Gunstens en, 1992 Massachusetts Institute of Technology, Lattice-Boltzmann studies of multiphase flow through porous media.
2. John F. Olson, 1996 Massachusetts Institute of Technology, Two-fluid flow in sedimentary rock: complexity, transport, and simulation.
3. Einat Aharonov, 1996 Massachusetts Institute of Technology, Solid-fluid interactions in porous media: processes that form rocks.
4. Olav van Genabeek, 1998 Massachusetts Institute of Technology, Velocity fluctuations in slow flow through porous media.
5. Peter S. Dodds, 2000 Massachusetts Institute of Technology, Geometry of river networks.
6. Davide Stelitano, 2000 Massachusetts Institute of Technology, Elastic interfaces in fluids: lattice-Boltzmann nodel and applications.
7. Joshua S. Weitz, 2003 Massachusetts Institute of Technology, Generalized contact processes in ecology.
8. Alexander P.P. Petroff, 2011 Massachusetts Institute of Technology, Streams, stromatolites and the geometry of growth.
9. David Forney, 2012 Massachusetts Institute of Technology, Emergent properties of heterogeneous decomposition networks

Current students at Massachusetts Institute of Technology:

1. Christopher Follet
2. Robert Yi

IId Peter Mora

1987 Stanford University, Elastic Wavefield Inversion.

Graduates at Institut de Physique du Globe de Paris:

1. Dominique Rodriques, 1993 Institut de Physique du Globe de Paris, Simulation de la propagation d'ondes sismiques sur machine massivement paralléle.
2. Dimitri Pissarenko, 1994 Institut de Physique du Globe de Paris, Modélisation de la friction et de l'instabilité sismique.
3. Bertrand Maillot, 1994 Institut de Physique du Globe de Paris, Construction d'un modèle semi-microscopique d'ondes élastiques basé sur les mthodes de Boltzmann sur réseaux.
4. Lian-Jie Huang, 1994 Institut de Physique du Globe de Paris, Approches microscopiques de la propagation des ondes dans les milieux complexes.
5. Frédéric Donze, 1994 Institut de Physique du Globe de Paris, Modélisation numrique de la déformation des roches comportement fragile par une méthode d'éléments discrets.
6. Bruno Riollet, 1997 University Paris-Diderot, Simulation numrique de la propagation d'ondes sismiques en milieu fracturé.

Graduates at The University of Queensland:

1. David Place 1999 The University of Queensland, A refined lattice solid model to simulate earthquakes and localisation phenomena using parallel computers.
2. Steffen Abe, 2002 The University of Queensland, Investigation of the influence of different micro-physics on the dynamic behaviour of faults using the lattice solid model.
3. Dion Kent Weatherley, 2002 Investigations of automaton earthquake models: implications for seismicity and earthquake forecasting.
4. Can Yin, 2005 The University of Queensland, Exploring the underlying mechanism of load/unload response ratio theory and its application to earthquake prediction.

IIe Biondo L. Biondi

1990 Stanford University, Seismic Velocity Estimation by Beam Stack.

Graduates at Stanford University:

1. Morgan P. Brown, 2004 Stanford University, Least-squares joint imaging of multiples and primaries.
2. Paul Sava, follows under (IIIh).
3. Brad Artman, 2006 Stanford University, Passive seismic imaging.
4. Daniel A. Rosales Roche, 2006 Stanford University, Imaging of converted-wave Ocean-bottom seismic data.
5. Gabriel F. Alvarez, 2007 Stanford University, Attenuation of multiples in image space.
6. Guojian Shan, 2008 Stanford University, Imaging of steep reflectors in anisotropic media by wavefield extrapolation.
7. Alejandro A. Valenciano, 2008 Stanford University, Imaging by wave-equation inversion.
8. Jeff Shragge, follows under (IIIi).
9. Claudio Guerra, 2010 Stanford University, Migration-velocity analysis using image-space generalized wavefields
10. Yaxun Tang, 2011 Stanford University, Imaging and velocity analysis by target-oriented wavefield inversion.
11. Gboyega Ayeni, 2011 Stanford University, Time-lapse seismic imaging by linearized joint inversion.

Current students at Stanford University:

1. Adam Halpert
2. Xukai Shen
3. Sjoerd de Ridder
4. Mandy Wong
5. Yunyue Li
6. Yang Zhang
7. Ohad Barak
8. Chris Leader
9. Ali Almomin
10. Yi Shen
11. Noha Farghal
12. Musa Maharramov
13. Jason Chang
14. Taylor Dahlke
15. (Huy Le - masters)

IIf David E. Lumley
1995 Stanford University, Seismic Time-Lapse Monitoring of Subsurface Fluid Flow.
Graduates at Memorial University at Newfoundand:
1. Richard Wright, 2005, 4D Seismic Analysis of the Hubernia Oil Field, Grand Banks, Canada.

Graduates at the University of Western Australia:
1. Matt Saul, 2013 University of Western Australia with (IIf) and (IIIi), 4D seismic data analysis of reservoir gas movement.

Current students at the University of Western Australia:
1. Lisa Gavin, co-advised with (IIIi).
2. James Deeks, co-advised with (IIIi).
3. Mohammad Emami Niri
4. Rafael de Souza

IIg Sergey Fomel
2001 Stanford University, Three-dimensional seismic data regularization.
Graduates at The University of Texas at Austin:
1. Hesam Kazemeini, 2009 Uppsala University with C. Juhlin, Seismic Investigations at the Ketzin CO2 injection site, Germany: Applications to subsurface feature mapping and CO2 seismic response modeling.
2. William Burnett, 2011 The University of Texas at Austin, Multiazimuth velocity analysis using velocity-independent seismic imaging.
3. Xiaolei Song, 2012 The University of Texas at Austin, Application of Fourier finite differences and lowrank approximation method for seismic modeling and subsalt imaging.

Current students at The University of Texas at Austin:
1. Vladimir Bashkardin
2. Siwei Li
3. Parvaneh Karimi
4. Yangkang Chen
5. Junzhe Sun

IIIa Thomas M. Hearn
1985 California Institute of Technology, Crustal structure in Southern California from array data. Graduates at New Mexico State University:
1. Anca Rosca, 2000 New Mexico State University, Upper mantle structure beneath Nevada and southern California.

IIIb Eugene D. Humphreys
1985 California Institute of Technology with Brad Hager, Studies of the crust-mantle system beneath Southern California.
Graduates at the University of Oregon:
1. Alberto E. Patiño-Douce, 1990 University of Oregon under (IIIg) and A. Dana Johnson, Ultrametamorphism and anatexis of the continental crust: an experimental and theoretical study.
2. Francois Saucier, 1991 University of Oregon, Studies of plate-mantle dynamics at continental margins.
3. Glenn P. Biasi, 1994 University of Oregon, The streaming potential method applied to a low gradient hydrologic environment along the Mojave River, San Bernardino County, California.

4. Kenneth Dueker, 1994 University of Oregon, Origin of western United States upper mantle seismic heterogeneity.
5. Xiaohua Peng, 1995 University of Oregon, Crustal velocity structure inferred from waveform modeling of teleseismic body waves: studies of northern Nevada and the eastern Snake River plain.
6. Randy L. Palmer, 1997 University of Oregon, Studies of the kinematics and dynamics of southern California and northern Baja, Mexico.
7. Elizabeth H, Hearn, 1998 University of Oregon, Numerical models of lithosphere deformation : inferring rheology and structure from limited surface observations.
8. Mark A. Hemphill-Haley, 1999 University of Oregon, Multi-scaled analyses of contemporary crustal deformation of western North America.
9. William C. Hammond, 2000 University of Oregon, Dynamics, flow and melt content of the Southern East Pacific Rise upper mantle from teleseismic tomography.
10. Derek L. Schutt, 2000 University of Oregon, Aspects of upper mantle structure in the Yellowstone Swell, Wyoming Craton, and Yavapai Province.
11. Haiying Gao, 2011 University of Oregon, The seismic structures of the U.S. Pacific Northwest and the scaling and recurrence patterns of slow slip events.
12. Brandon Schmandt, 2011 University of Oregon, Seismic Structure of the Western U.S. Mantle and Its Relation to Regional Tectonic and Magmatic Activity.

Current students at the University of Oregon:

1. David Adams
2. Noah Fay

IIIc John Vidale

1986 California Institute of Technology with Don Helmberger, Application of two-dimensional finite-difference wave simulation to earthquakes, earth structure, and seismic hazard.

Graduates at the University of California, Santa Cruz:

1. Ornella Bonamassa, 1995 University of California, Santa Cruz, Analysis of site effect and directional resonaces and their relevance to seismic hazard.

Graduates at the University of California, Los Angeles:

1. Elizabeth Cochran, 2005 University of California, Los Angeles, Earthquake rupture initiation and fault structure : I. triggering of earthquakes by Earth tides : II. seismic anisotrophy near the Hector Mine rupture : III. post-seismic displacements observed with InSAR.
2. Steven E. Persh, 2002 University of California, Los Angeles, Seismic investigations of core-mantle boundary structure and source properties of deep-focus earthquakes.

Graduates at the University of Washington:

1. Andrew Delorey, 2010 University of Waschington, Predicting Strong Motions for Seismic Hazard Assessments in Seattle, Washington.
2. Abhijit Ghosh, 2011 University of Washington, Imaging slow earthquakes in Cascadia using seismic arrays.

Current students at the University of Washington:

1. Alicia Hotovek
2. Kate Allstadt

IIId John N. Louie

1987 California Institute of Technology under (IIa), Seismic reflection experiments imaging

the physical nature of crustal structures in southern California.
Graduates at the University of Nevada, Reno:

1. Sathish Kumar Pullammanappallil C., 1994 University of Nevada, Reno, Nonlinear optimization to estimate velocities and image reflectors from multi-offset seismic data.
2. Sergio Chvez-Prez, 1997 University of Nevada, Reno, Enhanced imaging of fault zones in southern California from seismic reflection studies.
3. Abu M. Asad, 1998 University of Nevada, Reno, Linearized and nonlinear travel time tomography for upper crustal velocity structure of the western Great Basin.
4. Robert E. Abbott, 2001 University of Nevada, Reno, Geophysical constraints on seismic hazard and tectonics in the western Basin and Range.
5. James B. Scott, 2007 University of Nevada, Reno, Seismic noise in the shallow subsurface: methods for using it in earthquake hazard assessment.
6. Donghong Pei, 2007 University of Nevada, Reno, Modeling and inversion of dispersion curves of surface waves in shallow site investigations.

Current students at the University of Nevada, Reno:

1. Gretchen Schmauder

IIIe Hua-Wei Zhou
1989 California Institute of Technology with Don Anderson, Travel time tomographic studies of seismic structures around subducted lithospheric slabs.
Current students while at the University of Houston:

1. Curtis A. Link, 1993 University of Houston, Crosshole seism. analysis for reservoir characterization & lithology indication.
2. Hongwei Wang, 1994 University of Houston, Analysis of ISC P-wave data and applications to the inference of velocity structure in the mantle.
3. Stephen T. T. Ha, 1994 University of Houston, Benefiting from directional phase-encoded wavefield: Experimental and theoretical implementation of Neidell-dolphin concept.
4. Genmeng Chen, 1995 University of Houston, Seismic modeling and imaging of heterogeneous media.
5. Marcos Guimaraes, 1998 University of Houston, Phys. model study of seis. acquisition and processing of vertical cable data.
6. Oong Koo Youn, 1998 University of Houston, AVO modeling and depth imaging with multiples.
7. Mohammed Al-Otaibi, 2002 University of Houston, Thin-bed AVO.
8. Khalid Al-Rufaii, 2002 University of Houston, Seismic tomography in areas associated with complex near-surface structures
9. Warren Duncan, 2005 University of Houston, Integrated geophysical study of Vinton Dome, LA.
10. Mike P. Thornton, 2006 University of Houston, Depth imaging of crustal scale seismic reflection surveys in Southern California.
11. Li Li, 2007 University of Houston, Mapping 3D crustal velocities of S. California by deformable layer tomography.
12. Alex Zhao, 2007 University of Houston, Classification detection of reservoir fluids with rock physics constraints.
13. Hongmei Cao, 2008 University of Houston, Reflection attenuation tomo. and ray count weighted multi-scale tomography.
14. Hui Liu, 2010 Texas Tech University, Multiscale deformable layer tomography for near-surface velocity model building.

15. Fan Jiang, 2010 Texas Tech University, Waveform modeling, migration and tomographic analysis of seismic anisotropy.
16. Aifei Bian, 2011 China University of Geosciences at Wuhan, Theory and application of full waveform inversion in reflection seismology.
17. Bao Mei, 2011 China University of Geosciences at Wuhan, Research on crustal structure from tomography in the Three Gorges reservoir head area .
18. Zhihui Zou, 2012 Texas Tech University, A seismologic study of the Three-Gorges Reservoir (TGR) region, China.

Current students at the University of Houston:

1. Chris Gantela

Current students at Texas Tech University:

1. Fang Yuan
2. Pan Deng

IIIf J. Huw Davies

1989 California Institute of Technology under (IIa) and Dave Stevenson, Some problems in mantle structure and dynamics; Part 1. Inversion for depth variation of spectra of mantle compressional and shear velocity teterogeneity. Part 2. Physical model of source region of subduction zone volcanism.

Graduates at University of Liverpool:

1. Andrea Rowland, 1997 University of Liverpool, Numerical modelling of subduction zone magmatism.
2. Mark Rhodes, 1998 University of Liverpool, Mantle seismic tomography using P-wave travel times and a priori velocity models.
3. Carlos E. Izarra Teran, 2001 University if Liverpool under (IIIf) and xxx, The gravity anomaly of the subducted Nazca plate over the central Andes.

Graduates at Cardiff University:

1. David Oldham, Cardiff University, On the possibility of layered mantle convection - numerical simulation in a spherical geometry.
2. D. Rhodri Davies, follows under (IVa).
3. Martin Wolstencroft, Cardiff University, Understanding the thermal evolution of Earth.
4. Peter J. Webb, 2012 Cardiff University, Mantle circulation models: constraining mantle dynamics, testing plate motion history and calculating dynamic topography.

Current students at Cardiff University:

1. Rebekah Lawton
2. Matthew Price

IIIg YoungHee Kim

2011 California Institute of Technology with J. Jackson, Geophysical Investigations of Near-Surface Structure on Earth and Mars.

Current students at Seoul National University:

1. Eunyoung Kim

IIIh Paul C. Sava

2004 Stanford University, Migration and velocity analysis by wavefield extrapolation.

Graduates at Colorado School of Mines:

1. Jia Yan, 2010 Colorado School of Mines, Elastic anisotropic wavefield imaging.

2. Tongning Yang, 2013 Colorado School of Mines, Wavefield-based velocity model building using extended images.

Current Students at Colorado School of Mines:

1. Francesco Perrone
2. Esteban Diaz
3. Natalya Patrikeeva
4. Yuting Duan

IIIi Jeffrey Shragge

2009 Stanford University, Wave-equation migration in generalized coordinates.

Graduates at the University of Western Australia:

1. Matt Saul, 2013 University of Western Australia with (IIIe) and (IIf), 4D seismic data analysis of reservoir gas movement.

Current students at the University of Western Australia:

1. Lisa Gavin, co-advised with (IIf).
2. James Deeks, co-advised with (IIf).

IVa Rhodri Davies

2007, Cardiff University and Swansea University under (IIIf), Ken Morgan, Oubay Hassan and Perumal Nithiarasu, Applying Multi-Resolution Numerical Methods to Geodynamics.

Current students at Imperial College London:

1. Giuseppe Le-Voci

ACKNOWLEDGEMENTS

I would like to thank Bob Clapp, Biondo Biondi and of course Jon Claerbout, for helping me find so many of Jon's PhD students. Many people from all over the world have been helpful and enthusiastic in replying to my inquires. In no particular order and likely incomplete: Robert Clayton, Dave Hale, Paul Sava, Sergey Fomel, David Lumley, Jeff Shragge, Kim YoungHee, John Louie, J. Huw Davies, John Vidale, Eugene Humphreys, Thomas Hearn, Hua-Wei Zhou, Jun Ji, Abhijit Ghosh, Rhodri Davies and Thomas Hearn.

SEP PHONE DIRECTORY

Name	Phone	Login Name
Almomin, Ali	723-0463	ali
Barak, Ohad	723-9282	ohad
Biondi, Biondo	723-1319	biondo
Chang, Jason	724-2974	jason
Claerbout, Jon	723-3717	jon
Clapp, Bob	725-1334	bob
Dahlke, Taylor	723-3187	taylor
de Ridder, Sjoerd	723-1250	sjoerd
Farghal, Noha	723-1250	nfarghal
Halpert, Adam	723-6006	adam
Lau, Diane	723-1703	diane
Le, Huy	723-3187	huyle
Leader, Chris	723-0463	chrisl
Levin, Stewart	726-1959	stew
Li, Elita	723-9282	myfusan
Maharramov, Musa	723-1250	musa
Shen, Xukai	723-0463	xukai
Shen, Yi	723-6006	yishen
Wong, Mandy	723-9282	mandyman
Zhang, Yang	723-6006	yang

Dept fax number: (650) 725-7344

E-MAIL

Our Internet address is "*sep.stanford.edu*"; i.e., send Jon electronic mail with the address "*jon@sep.stanford.edu*".

WORLD-WIDE WEB SERVER INFORMATION

Sponsors who have provided us with their domain names are not prompted for a password when they access from work. If you are a sponsor, and would like to access our restricted area away from work, visit our website and attempt to download the material. You will then fill out a form, and we will send the username/password to your e-mail address at a sponsor company.

STEERING COMMITTEE MEMBERS, 2012-2013

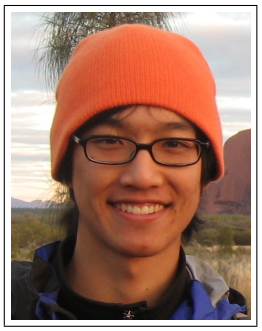
Name	Company	Tel #	E-Mail
Raymond Abma	BP	(281)366-4604	abmar1@bp.com
Francois Audebert (Co-chair, 1st year)	TOTAL	–	francois.audebert@total.com
Biondo Biondi	SEP	(650)723-1319	biondo@sep.stanford.edu
Jon Claerbout	SEP	(650)723-3717	jon@sep.stanford.edu
Thomas Dickens	ExxonMobil	(713)431-6011	tom.a.dickens@exxonmobil.com
Yi Luo	Saudi Aramco	–	yi.luo@aramco.com
Alejandro Valenciano (Co-chair, 2nd year)	PGS	–	alejandro.valenciano@pgs.com

Research Personnel

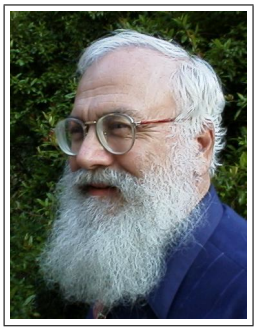
Ali Almomin graduated from Texas A&M University in 2007 with a BS in Geophysics. Then, he joined Saudi Aramco and worked in several exploration and research departments with a focus on 3D seismic processing and near surface imaging. He joined Stanford Exploration Project in 2009 to pursue a PhD in Geophysics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

Ohad Barak received a B.Sc. (2006) and an M.Sc. (2009) in Geophysics from Tel-Aviv University. In 2008 he joined the R&D team of Paradigm Geophysical, maintaining and developing the production codes. He joined SEP in 2009 and is currently pursuing a Ph.D. in geophysics at Stanford University, and a longer biography. His current research is on P/S separation of ocean-bottom seismic data.

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.



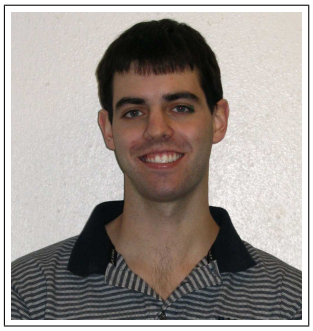
Jason Chang received his B.A. in geophysics from the University of California, Berkeley, in 2010. He joined SEP in autumn of 2011 and is currently working toward a Ph.D. in geophysics. He is a student member of SEG and AGU.



Jon F. Claerbout (M.I.T., B.S. physics, 1960; M.S. 1963; Ph.D. geophysics, 1967), professor at Stanford University, 1967. Emeritus 2008. Best Presentation Award from the Society of Exploration Geophysicists (SEG) for his paper, *Extrapolation of Wave Fields.* Honorary member and SEG Fessenden Award "in recognition of his outstanding and original pioneering work in seismic wave analysis." Founded the Stanford Exploration Project (SEP) in 1973. Elected Fellow of the American Geophysical Union. Authored three published books and five internet books. Elected to the National Academy of Engineering. Maurice Ewing Medal, SEG's highest award. Honorary Member of the European Assn. of Geoscientists & Engineers (EAGE). EAGE's highest recognition, the Erasmus Award.

Robert Clapp received his B.Sc. (Hons.) in Geophysical Engineering from Colorado School of Mines in May 1993. He joined SEP in September 1993, received his Masters in June 1995, and his Ph.D. in December 2000. He is a member of the SEG and AGU.

Taylor Dahlke Taylor Dahlke received his B.S. in civil engineering from the University of California, Berkeley in 2012. He joined SEP in July 2012 and is currently working towards a Ph.D. in geophysics. During the summer of 2011 he worked at Lawrence Berkeley National Laboratory investigating beamforming techniques to localize rock fracturing. His current research interests include deconvolution and salt body image segmentation.



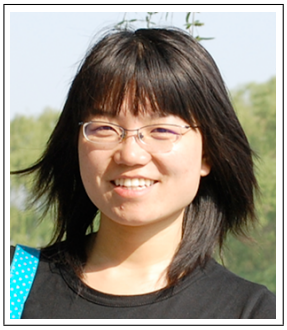
Adam Halpert graduated from Rice University in May 2007 with a B.Sc. in Earth Science and a B.A. in Policy Studies. He joined SEP in the summer of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG and AGU.

Huy Le graduated from the University of Oklahoma in 2012 with a B.S. in Geophysics. He is currently a first-year Master student in Computational Geoscience, a joint program between the School of Earth Sciences and the Institute of Computational and Mathematical Engineering. His research interests include finite differences for wave propagation and full waveform inversion.

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

Stewart A. Levin was acting director of the Stanford Exploration Project during Jon Claerbout's 1993-4 sabbatical year. After a distinguished career in industry at Mobil and Halliburton, he has returned to Stanford as a consulting professor in the Department of Geophysics.

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



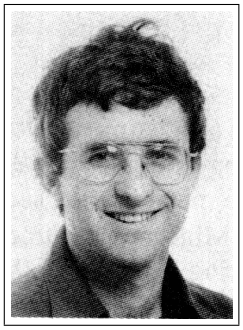
Musa Maharramov received his diploma c.l. in Mathematics from Baku State University in 1993. In 1995, he received a Ph.D. in Mathematics (Differential Equations) from the Azerbaijan Institute of Mathematics for his thesis "Asymptotic Solutions of Quasilinear Parabolic Equations". In 1995-2001, he worked 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.

Gary Mavko is professor of Geophysics at Stanford University. Gary received his B.S. in enginnering physics from Cornell University, and his M.S. (1974) and a Ph.D. (1977) in geophysics from Stanford. Gary specialzes in rock physics and is the co-author of "The Rock Physics Handbook". He is a member of AGU, EAGE, SEG, AAPG, and SPE.

Dave Nichols received a B.A. in physics from Cambridge University in 1982; a M.Sc. in geophysics from Imperial College, London in 1983; and a Ph.D. in geophysics from Stanford University in 1994. From 1983 to 1987, he was employed by Western Geophysical in London and joined SEP in 1987. During the summer of 1991, he worked for Chevron Oil Field Research Company in La Habra, California. He currently works for Schlumberger in the Schlumberger Reservoir Completions Center in Rosharon, Texas.

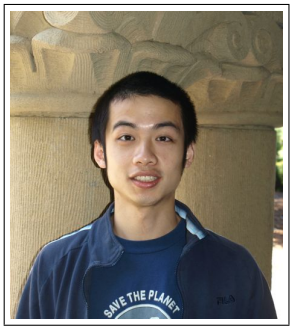
Sjoerd de Ridder received his B.Sc. (Dec. 2004) in Earth Sciences from Utrecht University in the Netherlands. After an exchange year at Colorado School of Mines he attended Delft University of Technology. In August of 2007, he obtained an M.Sc. in Applied Earth Sciences with specialization in applied geophysics from Delft University of Technology for his thesis: 'Simulation of interferometric seismoelectric Greens function recovery'. He joined SEP in the fall of 2007 and is pursuing a Ph.D. in Geophysics. He received the 'Jon F. Claerbout - Chevron Fellowship' in 2009-2010. During the summer of 2010 he did an internship at China University of Petroleum in Beijing. Currently he is working on exploration seismology with low-frequency ambient seismic noise. In his free time he likes to run, bike and hike while traveling and discussing philosophy, politics and history.

Shuki Ronen is currently (2013) the chief geophysicist of Seabed Geosolutions, and a consulting faculty at SEP. The photo dates back to when he was a student at Stanford, 1985.



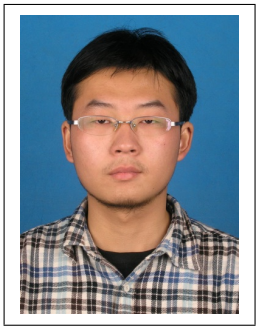
Guojian Shan received his B.Sc. in Mathematics School of Peking University in July, 1998. From 1998 to 2001, he studied in Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences (CAS), and received his M.S. in Applied Mathematics in July, 2001. He joined SEP in 2001 and received his Ph.D. in April, 2008. He is now working for Chevron as a research geophysicist. He is a member of the SEG.

Xukai Shen graduated from Tsinghua University in July 2007 with a B.E. in Electrical Engineering. He joined SEP in the fall of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG.



Yi Shen got her Bachelors degree in Acoustics, from Dept. of Electronic Science and Engineering, Nanjing University, China in 2010. She joined SEP in the fall of 2010, and is currently working toward a Ph.D. in Geophysics at Stanford University. She is a student member of the SEG.

Mandy Wong graduated in 2004 with a B.Sc. in Physics and Mathematics from the University of British Columbia (UBC) in Vancouver, Canada. In 2006, she obtained a M.Sc. degree in Condensed Matter Thoery at UBC. Afterward, Mandy worked for a geophysical consulting company, SJ Geophysics, based in Vancouver, Canada. Mandy joined SEP in 2008, and is working towards a Ph.D. in Geophysics. Her main research interest is imaging with multiples.



Yang Zhang graduated from Tsinghua University in July 2007 with a B.E. in Electrical Engineering. He took an internship in Microsoft Research Asia during 2007-2008. He joined SEP in 2009, and is currently pursuing a Ph.D. in Geophysics.

SEP ARTICLES PUBLISHED OR IN PRESS

Almomin, A., 2011, Correlation-based wave-equation migration velocity analysis. SEG Technical Program Expanded Abstracts, **30**, 3887–3891.

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 (Submitted).

Ayeni, G., Tang, Y., and Biondi, B., 2010, Efficient seismic monitoring of hydrocarbon reservoirs using multiple shooting vessels: In Proceedings of the 2010 Offshore Technology Conference.

Ayeni, G., and Biondi, B., 2010, Continuous reservoir monitoring with asynchronous simultaneous-source seismic data: 72nd Conference & Technical Exbition, EAGE, Extended Abstracts.

Ayeni G., and Biondi, B., 2010, Target-oriented joint least-squares migration/inversion of time-lapse seismic data sets: Geophysics, **75**, no. 3, R61–R75.

Ayeni, G., and Biondi, B., 2011, Wave-equation inversion of time-lapse seismic data sets: SEG Technical Program Expanded Abstracts (Submitted).

Ayeni, G., Almomin, A., and Nichols, D., 2011, On the separation of simultaneous-source data by inversion: SEG Technical Program Expanded Abstracts (Submitted).

Guerra, C., and Biondi, B., 2011, Fast 3D velocity updates using the pre-stack exploding reflector model: SEG Technical Program Expanded Abstracts **29**, 4375–4379.

Guerra, C., and Biondi, B., 2010, Fast 3D velocity updates using the pre-stack exploding reflector model: Geophysics (Accepted).

Biondi, B., 2010, Velocity estimation by image focusing analysis: Geophysics, **75**, pp. 49–60.

Biondi, B., 2010, Velocity estimation by image focusing analysis: SEG Technical Program Expanded Abstracts **29**, 4066–4070.

Chang, J., and de Ridder, S., 2012, Correlation energy between surface and borehole stations at Valhall: SEG Technical Program Expanded Abstracts **31** (Submitted).

de Ridder, S., N. Crook, S.S. Haines, and S.T. Ide, 2010, Seismic investigation of underground coal fires: A feasibility study at the Southern Ute Nation Coal Fire site. Symposium on the Application of Geophysics to Engineering and Environmental Problems **23**, 630–638.

de Ridder, S., and B. Biondi, 2010, Low-frequency passive seismic interferometry for land data: SEG Technical Program Expanded Abstracts **29**, 4041–4046.

de Ridder, S., and J. Dellinger, 2011, Imaging of near-surface structures at Valhall using ambient-seismic-noise Eikonal tomography: EAGE Third Passive Seismic Workshop, Athens, Greece, PSP13.

de Ridder, S., and J. Dellinger, 2011, Ambient seismic noise eikonal tomography for near-surface imaging at Valhall. The Leading Edge, **30**, no. 5, pp. 936–942.

de Ridder, S., 2011, Ambient seismic noise tomography at Valhall. SEG Technical Program Expanded Abstracts **30**, 1597–1601.

de Ridder, S., 2011, Ambient seismic noise tomography for exploration seismology at Valhall. Presented at 2011 AGU Fall Meeting, San Francisco, CA, 5-9 December 2011.

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).
Dellinger, J., J. Yu, and S. de Ridder, 2010, Virtual-source interferometry of 4C OBC data at Valhall without a low-cut recording filter: SEG/EAGE Summer Research Workshop, Snowbird, Utah, 15-20 August 2010.
Farghal, M.S., and Levin, S.A., 2012, Hunting for microseismic reflections using multiplets: SEG Technical Program Expanded Abstracts **31** (Submitted).
Guitton, A., Ayeni, G., and Esteban, D.A., 2012, Constrained full-waveform inversion by model reparameterization: Geophysics, 77, No. 2, R117–R127.
Guitton, A., and Esteban, D.A., 2012, Attenuating crosstalk noise with simultaneous source full waveform inversion: Geophysical Prospecting (In Press).
Guitton, A., 2012, Blocky regularization schemes for full waveform inversion: Geophysical Prospecting (In Press).
Halpert, A., Model-building with image segmentation and fast image updates: SEG Technical Program Expanded Abstracts **30**, 4035–4039.
Halpert, A., Clapp, R.G., and B. Biondo, 2010, Speeding up seismic image segmentation: SEG Technical Program Expanded Abstracts **29**, 1276–1280.
Halpert, A., Clapp, R.G., and B. Biondo, 2011, Interpreter guidance for automated seismic image segmentation: 73rd Conference & Exhibition, EAGE, Extended Abstracts.
Leader, C., Shen, X., and R. Clapp, 2011, Memory efficient reverse time migration in 3D: Proceedings of Nvidia GPU Technology Conference 2011/2012 (Accepted).
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).
Li, Y., Y. Zhang, and J. Claerbout, 2010, Geophysical applications of a novel and robust L1 solver: SEG Technical Program Expanded Abstracts **29**, 3519–3523.
Li, Y., and B. Biondi, 2011, Migration velocity analysis for anisotropic models: SEG Technical Program Expanded Abstracts **30**, 201–206.
Li, Y., Nicoles, D., Osypov, K., and Bachrach, R., 2011, Anisotropic tomography with rock physics constraints: 73rd Conference & Technical Exbition, EAGE, Extended Abstracts.
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** (Submitted).
Maharramov, M., 2011, Random boundary condition for low-frequency wave propagation:

SEG Technical Program Expanded Abstracts **30**, 2962.

Maharramov, M., and B. Nolte, 2011, Efficient one-way wave-equation migration in tilted transversally isotropic media: 73rd Conference & Technical Exbition, EAGE, Extended Abstracts.

Shen, X., 2010, Near-surface velocity estimation by weighted early-arrival waveform inversion: SEG Technical Program Expanded Abstracts **29**, 1975–1979.

Shen, X., and R. Clapp, 2011, Random boundary condition for low-frequency wave propagation: SEG Technical Program Expanded Abstracts (Submitted).

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

Tang, Y., and B. Biondi, 2011, Target-oriented wavefield tomography using synthesized Born data: Geophysics (Accepted).

Tang, Y., and B. Biondi, 2011, Subsalt velocity analysis by target-oriented wavefield tomography: A 3-D field-data example: SEG Technical Program Expanded Abstracts (Submitted).

Wong, M., Biondi, B.L., and Ronen, S., 2010, Joint inversion of up- and down-going signal for ocean bottom data: SEG Technical Program Expanded Abstracts **29**, 2752–2756.

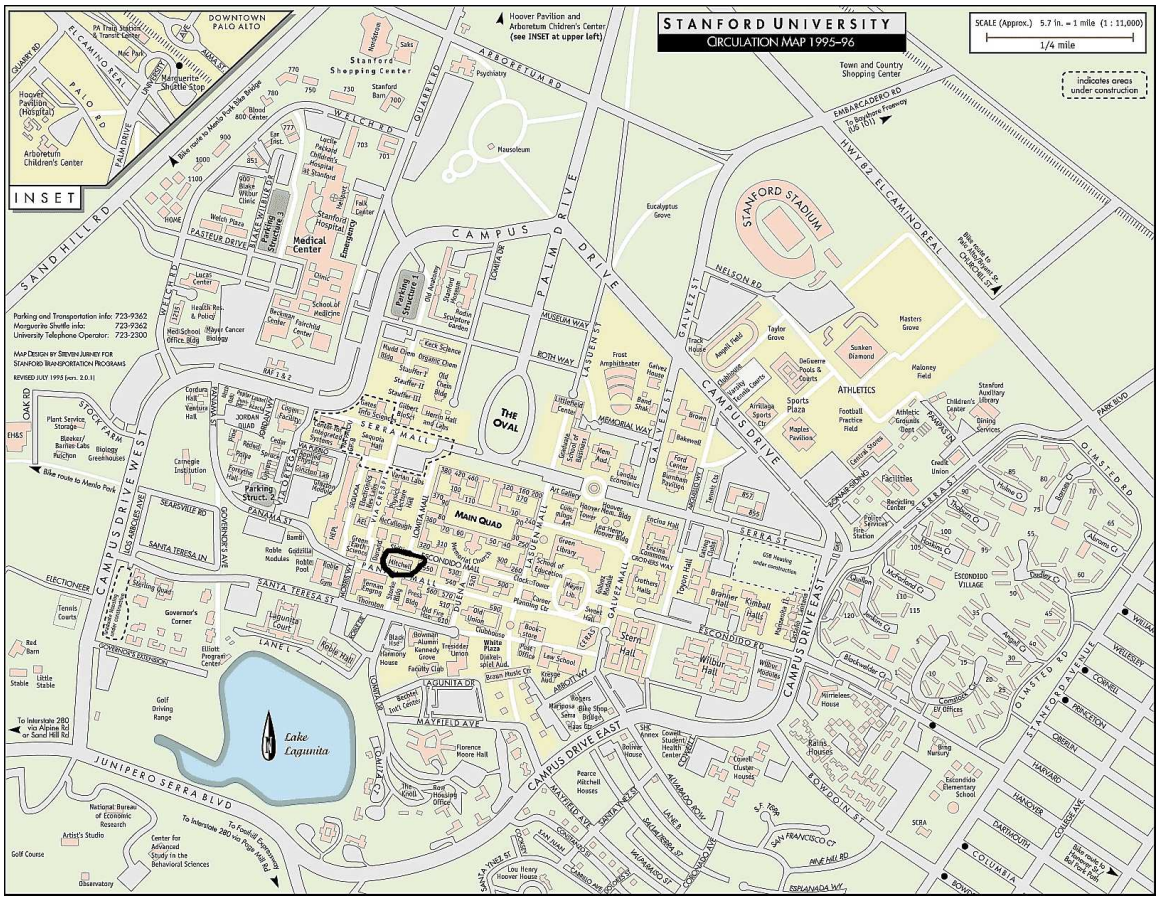
Wong, M., and Ronen, S., and Biondi, B.L., 2011, Least-squares reverse-time migration/inversion for ocean bottom data: A case study. SEG Technical Program Expanded Abstracts **30**, 2369–2373.

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, (submitted).

Zhang, Y., Claerbout, J., Guitton, A., 2011, A new bidirectional deconvolution that overcomes the minimum phase assumption. 73rd Conference & Technical Exbition, EAGE, Extended Abstracts.


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
Marguerite Shuttle Stop
Stanford Shopping Center
Town and Country Shopping Center
Hoover Pavilion (Hospital)
Arboretum Children's Center
Medical Center
Stanford Hospital Emergency
Mausoleum
Eucalyptus Grove
STANFORD STADIUM
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)
THE OVAL
MAIN QUAD
Lake Lagunita
Golf Driving Range
Golf Course
Observatory
ATHLETICS
Sports Plaza
Football Practice Field
Maples Pavilion
Arrillaga Sports Ctr
Taylor Grove
Masters Grove
Maloney Field
Susan Diamond
DeGuerre Pools & Courts
Angell Field
Track
Frost Amphitheater
Galvez House
Bowl Shell
Littlefield Center
Cubberley Auditorium
Hewlett Teaching Center
Art Gallery
Cummings Art Building
Hoover Tower
Hoover Assn. Bldg
Lou Henry Hoover Bldg
Encina Hall
Encina Commons
Green Library
Meyer Library
Main Library
Tresidder Union
Clock Tower
Memorial Church
Memorial Hall
Law School
Crothers Halls
Town Hall
Braun Music Ctr
Bishop Auditorium
Dinkelspiel Auditorium
Bookstore
Post Office
Clubhouse
Faculty Club
Bowman Alumni House
Kennedy Grove
Roble Hall
Lagunita Court
Governor's Corner
Elliott Program Center
Sterling Quad
Tennis Courts
Bldg 160
Bldg 200
Bldg 420
Bldg 460
Bldg 500
Bldg 550
Bldg 580
Bldg 590
Bldg 600
Bldg 620
Bldg 650
Bldg 660
Bldg 670
Bldg 680
Bldg 690
Bldg 700
Bldg 720
Bldg 750
Bldg 770
Bldg 780
Bldg 800
Bldg 801
Bldg 851
Bldg 900
Bldg 1000
Bldg 1100
Bldg 1201
Bldg 1205
Bldg 1215
Bldg 1225
Bldg 1235
Bldg 1245
Bldg 1255
Bldg 1265
Bldg 1275
Bldg 1285
Bldg 1295
Bldg 1305
Bldg 1315
Bldg 1325
Bldg 1335
Bldg 1345
Bldg 1355
Bldg 1365
Bldg 1375
Bldg 1385
Bldg 1395
Bldg