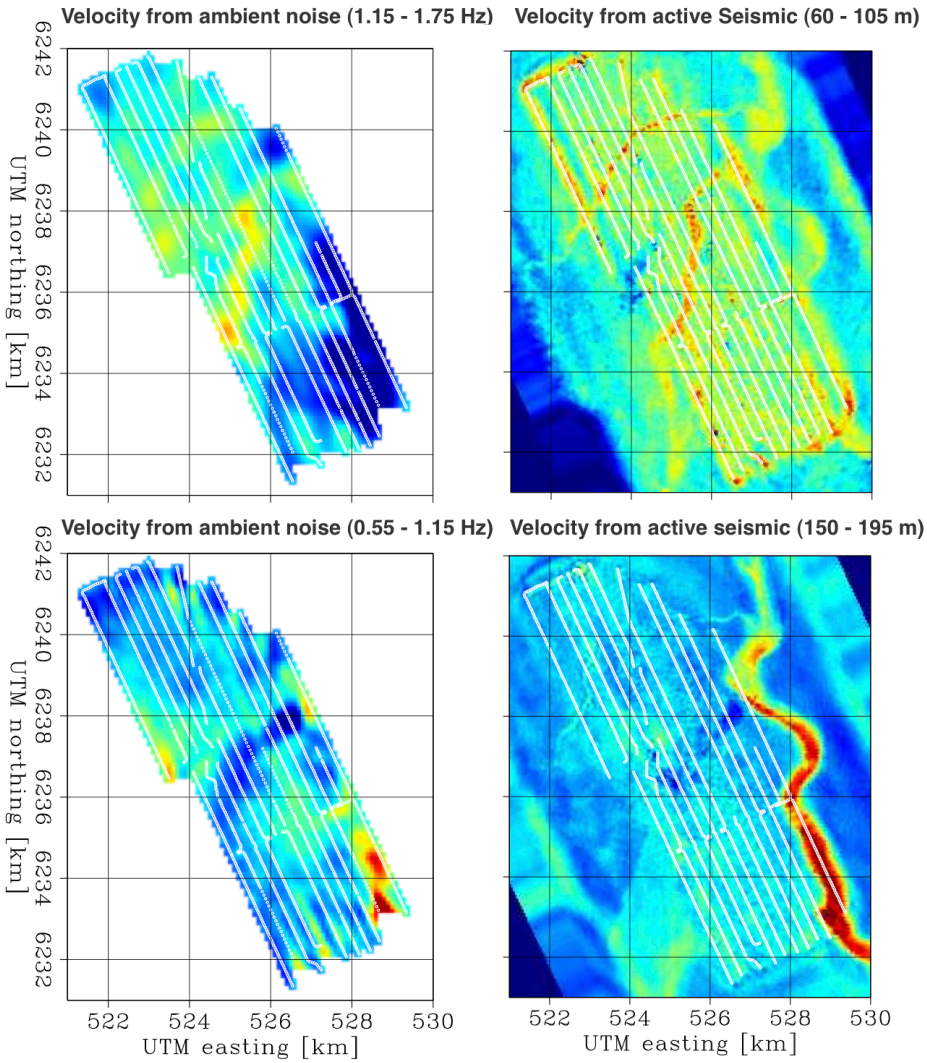
STANFORD EXPLORATION PROJECT

Ali Almomin, Ohad Barak, Biondo Biondi, Jason Chang, Jon Claerbout, Robert Clapp, Noha Farghal, Qiang Fu, Antoine Guitton, Adam Halpert, Jerry Harris, Chris Leader, Stewart Levin, Yunyue Li, Musa Maharramov, Colin Perkins, Sjoerd de Ridder, Shuki Ronen, Peng Shen, Xukai Shen, Yi Shen, Mandy Wong, Yang Zhang, and Tieyuan Zhu

Report Number 147, May 2012



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/sep147
[2] http://sepwww.stanford.edu/public/docs/sepdatalib/toc_html
[3] http://sepwww.stanford.edu/research/redoc/

SEP147 — TABLE OF CONTENTS

Velocity estimation

Imaging with anisotropy and attenuation

Reducing the cost of waveform imaging

Novel models and data types

Imaging through inversion

Interpretation

Decon and Interpolation

Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis

Biondo Biondi and Ali Almomin

ABSTRACT

By extending the velocity-model domain to subsurface offsets we solve the local-minima problem of data-fitting waveform inversion. We then regularize the extended-model data-fitting inversion with the addition of an image-focusing term to the objective function, therefore achieving robust global convergence of the waveform inversion problem. The method shares with full waveform inversion the advantage of simultaneously solving for all the wavelengths of the model, but it also has the global convergence characteristics of wave-equation migration velocity analysis. Numerical implementation of the proposed inversion method requires the solution of an extended wave-equation where velocity is a convolutional, instead of scalar, operator. The resulting method is therefore computationally intensive, but it can be easily tested in 2D. A simple example with a Gaussian velocity anomaly illustrates how the reflections from the anomaly recorded in the low-frequency components of the data increase the spatial resolution of the final inversion results. Numerical tests performed on synthetic data from a modified Marmousi model demonstrate the global convergence 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). The migration velocity model is estimated first, and then it is used as input to migration for imaging reflectivity. Even when wave-equation operators are employed to estimate the velocity, such as in wave-equation migration velocity analysis (WEMVA) methods, reflectivity is used only indirectly to measure the focusing power of the velocity model (Biondi and Sava, 1999; Shen and Symes, 2008). The only important exceptions in current practice occur when migrated volumes are used: 1) to interpret boundaries of geobodies (e.g. salt bodies), whose interior are assigned predefined velocities, and 2) to estimate predominant dips in the geologic layering that are then used to constraint a tomographic velocity updating (Clapp et al., 2004).

As the industry strives to widen the data frequency band at both the low and high end, the advantages of overcoming the limitations of conventional imaging, and of exploiting reflectivity information for velocity estimation, are becoming more relevant to important imaging problems. One of the main attractions of full waveform inversion (FWI) (Tarantola, 1987; Pratt, 1999) is to overcome the limitations imposed by the conventional approach that may limit the quality of the imaging results by finding a suboptimal solution. However,

FWI suffers from well-known convergence problems when the starting model is far from the correct one and low frequencies are missing from the data.

We discuss an inversion framework that overcomes FWI difficulties by supplementing an FWI-like data-fitting objective function with a WEMVA-like term that measures the reflection-focusing power of the velocity model. The method fits the recorded data with data modeled using a generalized version of the acoustic wave equation; the domain of velocity model is extended to include subsurface offsets. The extension of the reflectivity along the subsurface offset axes (or reflection angles) is a well-established technique for migration, linearized waveform inversion, and WEMVA (Biondi, 2006). In a data-fitting inversion, extending reflectivity to the prestack domain has the critical advantage that the kinematics of the modeled data will not be too distant from the ones of the recorded data, no matter the magnitude of the background velocity error.

Symes (2008) introduced the idea of using a wave equation with an extended velocity. By extending velocity the convergence difficulties of conventional FWI are overcome and all scales can be solved simultaneously. In the same paper, he also introduced the waveform-inversion formulation used in this paper, and described its application to the solution of 1D inversion problem in presence of multiple reflections.

The main goal of our research is not to tackle the problem of multiples, but to perform simultaneous inversion for all scales of the velocity model. Therefore, we apply the theory and numerically solve the extended wave equation in 2D. We also derive an effective scheme to linearize the extended wave equation and to compute the gradient of the objective-function by an adjoint-state method. In 3D the proposed method would be extremely expensive. In a companion report, (Almomin and Biondi, 2012) we present an approximation to the method presented here that drastically reduces the computational cost, but still retains the capability of simultaneously solving for all the wavelengths of the velocity model.

Another potential problem with strict coupling of velocity with reflectivity arises when the assumption of constant density cannot be made, as is the case in most of field data problems. In this case density variations may create reflections that do not correspond to velocity contrasts. However, we still would like to avoid the addition of density to the problem parameters for computational and convergence reasons. The approximate method proposed in Almomin and Biondi (2012) has the potential of being more flexible in accommodating these discrepancies.

TOMOGRAPHIC FULL WAVEFORM INVERSION (TFWI)

The conventional FWI objective function J_{FWI} can be written 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 data.

The modeled data 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 $f(\mathbf{x}_s, \omega)$ 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:

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

For the sake of compact notation, in the rest of the paper we present the expressions for computing the data and the gradient of objective functions in the frequency domain. However, we perform the computation in the time domain.

We can extend the velocity in the subsurface-offset dimension $\mathbf{h}$ which changes the wave equation into the following form

$$\left(v^2(\mathbf{x}, \mathbf{h}) *^{-1} \omega^2 + \nabla^2\right) G(\mathbf{x}_s, \mathbf{x}, \omega; \mathbf{v}) = \delta(\mathbf{x}_s - \mathbf{x}), \tag{4}$$

where with $*^{-1}$ we indicate deconvolution. Notice that the division by velocity in equation 4 becomes a deconvolution over the offset axis. Once we define the Green's function, the data could be computed similarly to equation 2. We now write the new objective function as follows:

$$J_{\mathrm{EFWI}}\left(\mathbf{v}(\mathbf{h})\right) = \left\|\mathbf{d}\left(\mathbf{v}(\mathbf{h})\right) - \mathbf{d}_{\mathrm{obs}}\right\|_2^2. \tag{5}$$

The long-wavelength components of the solution of the optimization problem defined by equation 5 are not likely to be substantially different from the long-wavelength components of the initial model. The extension of the model, and in particular of its reflectivity component, to non-zero subsurface offset causes the kinematics of the modeled data to match the kinematics of the recorded data independently from the accuracy of the long-wavelength components.

Another term must be added to the objective function to drive the solution towards a model that focuses the image. Symes (2008) suggests the addition of a differential semblance penalty function (DSO); that is,

$$J_{\mathrm{DSO}}(\mathbf{v}(\mathbf{h})) = \||\mathbf{h}|\mathbf{v}(\mathbf{h})\|_2^2. \tag{6}$$

We use this focusing term in the numerical experiments described in this paper. Another valid choice would be the maximization of the normalized power of the stack over reflection angles as a function of a residual moveout parameter ρ, as suggested by Zhang and Biondi (2012).

The important characteristic of the second term is that its gradient "imposes" on the current model only a phase shift, and not a bulk vertical shift. This assures that the corresponding perturbations on the modeled data are mere phase shifts, and not bulk time shifts. The absence of bulk time shifts in the modeled data avoids large discrepancies between the kinematics of modeled data and recorded data. These large discrepancies are at the root of the convergence problems in conventional FWI.

Another practically important consideration is that in the proposed formulation the computation of the gradient of a term like the one presented in Zhang and Biondi (2012) is straightforward because it does not require back-projection of image perturbations. This is in contrast with WEMVA-like methods, where the computation of the gradient must take into account the constraint that the image is the result of migrating the recorded data. Therefore, at least in principle, it would be equally easy to add to the objective function other terms that reward focusing of the model along the midpoint spatial axis, in addition to the subsurface offset or reflection angle (Biondi, 2010).

GRADIENT COMPUTATION

To compute the gradient of the objective function expressed in equation 5, we need to linearize the extended wave equation 4. Usually equation 3 is linearized over slowness or velocity. However, the extended wave equation 4 includes a deconvolution over the offset axis that is not easy to implement. Hence, we will first rearrange equation 4 to a form that facilitates the computation of the gradient, and that is actually solved numerically in the propagation. This can be achieved by convolving both sides of the equation by the square of velocity then rearranging the terms as follows:

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

We now can linearize the relationship between the Green's function and the model by perturbing the model around a background value as follows:

$$v^2(\mathbf{x}, \mathbf{h}) = v_0^2(\mathbf{x}, \mathbf{h}) + \Delta v^2(\mathbf{x}, \mathbf{h}),$$

where $v_0(\mathbf{x}, \mathbf{h})$ is the background component and $\Delta v(\mathbf{x}, \mathbf{h})$ is the perturbation component, i.e. the model update. After this separation, the first-order Born approximation can be used to define the gradient as follows:

$$\begin{aligned} g_d(\mathbf{x}, \mathbf{h}) = & \\ & \sum_{\mathbf{x}_s, \mathbf{x}_r, \omega} \left[\nabla^2 f(\mathbf{x}_s, \omega) G\left(\mathbf{x}_s, \mathbf{x} - \mathbf{h}, \omega; \mathbf{v}_0(\mathbf{h})\right)\right. \\ & \left. G\left(\mathbf{x}_r, \mathbf{x} + \mathbf{h}, \omega; \mathbf{v}_0(\mathbf{h})\right)\ \Delta d^*(\mathbf{x}_s, \mathbf{x}_r, \omega; \mathbf{v}_0(\mathbf{h}))\right], \end{aligned}$$

where Δd are the data residuals and * indicates the complex conjugate. Unlike the usual linearization of the wave equation, the scattering term includes a Laplacian operator instead of the second derivative in time. The Laplacian operator does not add to the computational cost since it is already computed in the propagation of the background wavefield.

SYNTHETIC DATA EXAMPLES

We tested the proposed inversion method with two different synthetic data sets. The first one was generated assuming a simple model where the goal was to estimate a Gaussian positive velocity anomaly. The second data set was generated assuming a slightly modified Marmousi model.

Gaussian Anomaly

To illustrate the interplay between different scales of the velocity during the inversion we applied the proposed method to a two-layer model with a Gaussian anomaly in the middle of the first layer, as shown in Figure 1. The velocity of the top layer is 3 km/s and of the bottom layer is 3.5 km/s. The velocity at the center of the Gaussian anomaly is 3.5 km/s. The interface between the layers is at 2 km depth and the Gaussian anomaly is centered at 1 km depth.

A Ricker wavelet with a fundamental frequency of 15 Hz and temporal sampling of 1.5 ms was used as a source function to model the data. The wavefields were generated by 31

sources with a spacing of 100 m and recorded by 151 fixed receivers with a spacing of 20 m. The maximum offset is 1.5 km. The initial model is a constant model of 3 km/s velocity.

Figure 2 shows the difference between the velocity model estimated after one iteration and the starting model. As expected, the reflector is well imaged, though not perfectly focused under the velocity anomaly. Figure 3 shows a rescaled window of Figure 2 around the anomaly. It shows how the reflections from the anomaly measured from the low-frequency components of the data start outlining the contour of the anomaly. Figure 4 shows the difference between the initial and the inverted model after 2000 iterations. Figure 5 shows a rescaled window of Figure 4 around the anomaly. The anomaly is now fairly well focused. However, the maximum amplitude of the estimated anomaly is still a fraction of the amplitude of the true anomaly, but better recovered than if it had been estimated using conventional velocity estimation methods based only on the transmission effects (Almomin and Biondi, 2012).

Figure 6 shows three sections of the difference cube obtained by subtracting the starting model from the model obtained after one iteration. These sections are taken at fixed horizontal coordinates and are functions of depth and subsurface offset. These images are analogous to migration subsurface-offset common image gathers; they show the lack of focusing of the model along the subsurface-offset axis. Figure 7 shows the gathers taken at the same locations as the ones shown in Figure 6, but after 2000 iterations. The gathers are now well focused around the zero-offset axis, indicating that the final model well explains the kinematic effects of the reflections that propagated through the velocity anomaly. Indeed, the gathers shown in Figure 7 appear artificially focused around zero subsurface offset. This appearance is caused by the DSO term in the objective function forcing the focusing of the image even beyond what would have been the focusing with the true model. In the data domain this "extra" focusing causes extrapolation of events beyond the recorded offsets (Clapp, 2005).

Figure 8 shows the normalized data residual as a function of iterations. The residual has not completely not flatten out and converged to zero; further iterations would improve the results.

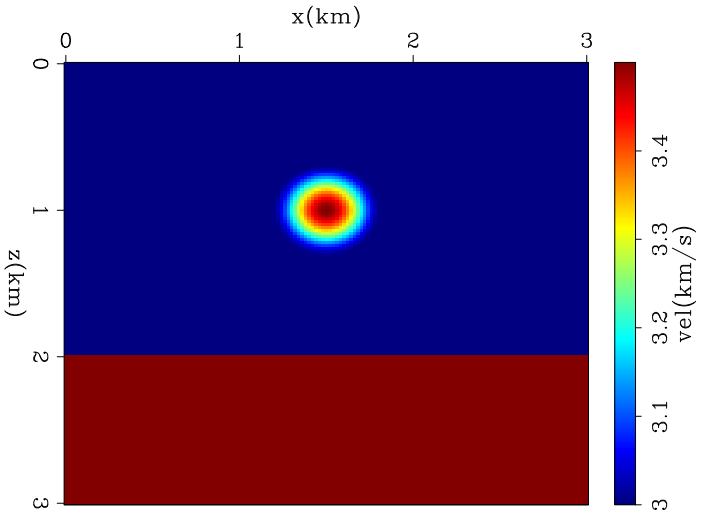


Figure 1: The true velocity model. [**ER**] biondo1/. seg1-gauss-true

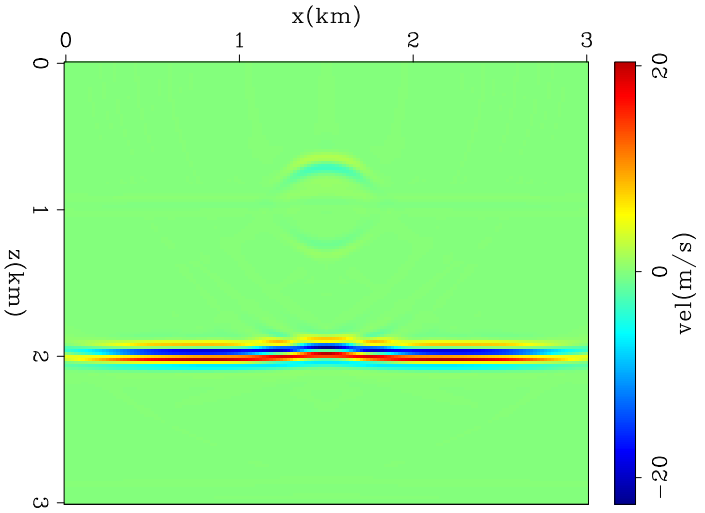


Figure 2: The difference between inverted and initial model after one iteration. [**CR**] biondo1/. seg1-gauss-iter1

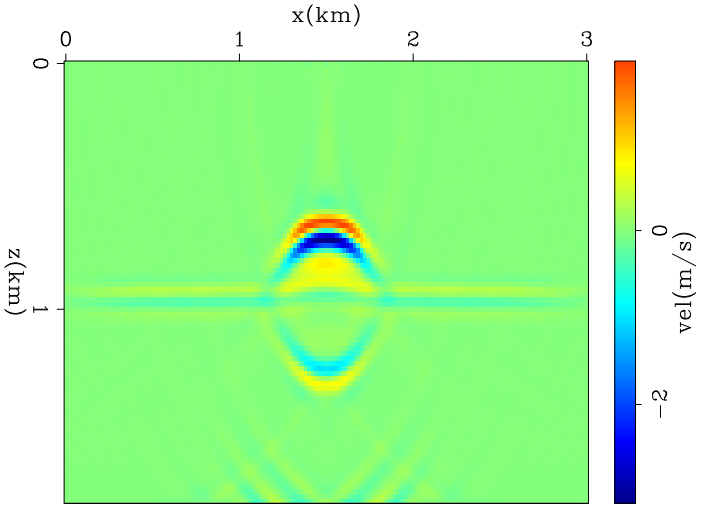


Figure 3: Zoom of the difference between inverted and initial model after one iteration. Notice the slight vertical stretch with respect to Figure 2. [**CR**] biondo1/. seg1-gauss-iter1-top

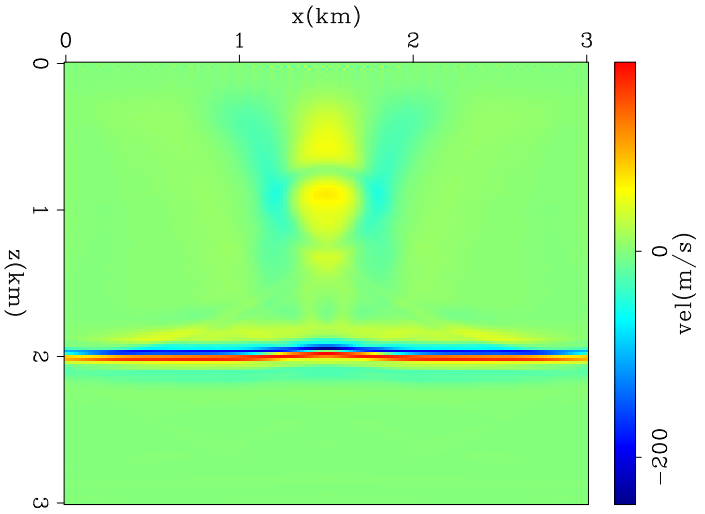


Figure 4: The difference between inverted and initial model after 2000 iterations. [**CR**] biondo1/. seg1-gauss-inv

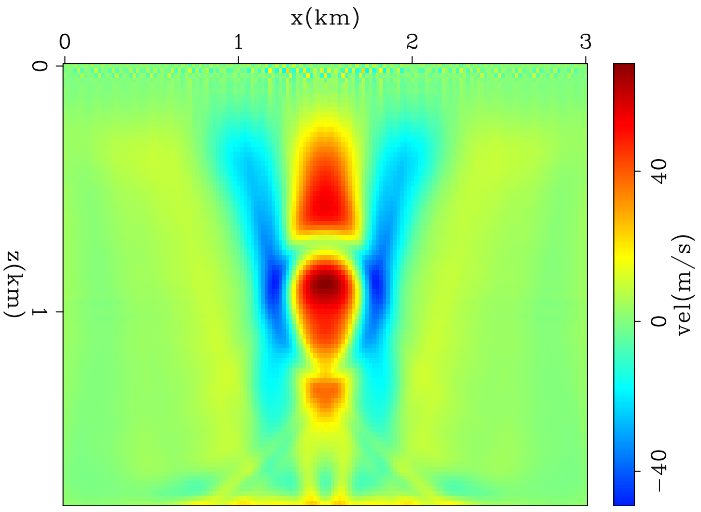


Figure 5: Zoom of the difference between inverted and initial model after 2000 iterations. Notice the slight vertical stretch with respect to Figure 4. [**CR**] biondo1/. seg1-gauss-inv-top

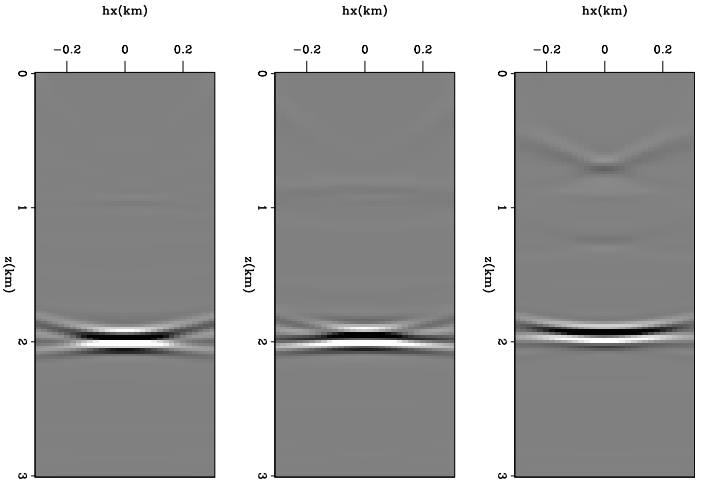


Figure 6: The difference between inverted and initial model after one iteration. These sections were taken at x=.09, 1.2, and 1.5 km. [**CR**] biondo1/. seg1-gauss-iter1-dv1

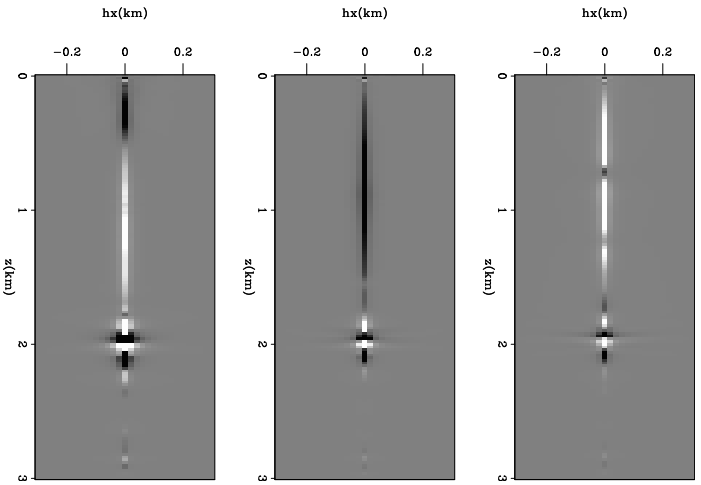


Figure 7: The difference between inverted and initial model after 2000 iterations. These sections were taken at x=.09, 1.2, and 1.5 km. Compare with Figure 6. [**CR**] biondo1/. seg1-gauss-inv-dv1

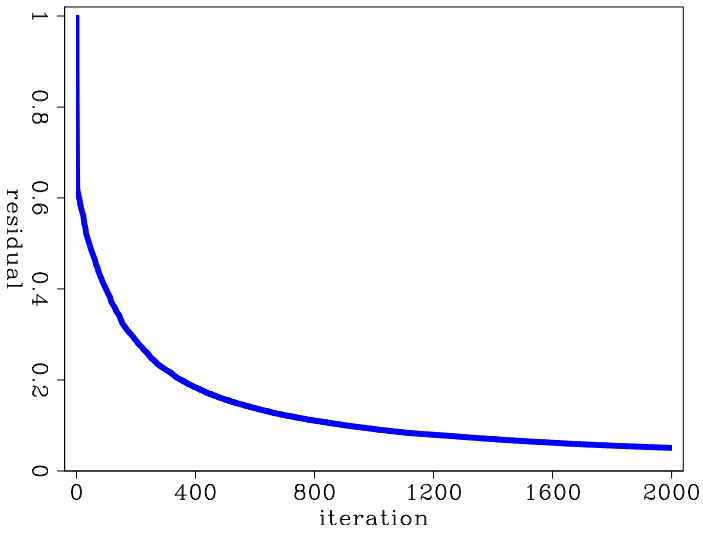


Figure 8: The normalized residual as a function of the number of iterations. [**CR**] biondo1/. seg1-gauss-res

Marmousi model

To test the convergence and resolution characteristics of the TFWI method, we applied it to the inversion of a synthetic data set generated assuming the modified Marmousi model shown in Figure 9. This model is the classical Marmousi with the addition of a 500 m water layer. The data were modeled assuming 93 shots with 100 m spacing and a Ricker wavelet with fundamental frequency of 15 Hz. The reflected wavefields were recorded by 461 receivers with fixed spread and 20 m spacing.

We compare the results of three inversion methods starting from the model shown in Figure 10, which was obtained by applying strong horizontal smoothing to the true model. Figure 11 shows the model obtained with conventional FWI; that is, by minimizing the objective function expressed in equation 1. As expected, conventional FWI fails because of its well-known difficulties with global convergence. The resulting model is almost identical to the starting one, with only a few shallow velocity discontinuities being imaged.

Figure 12 and 15 show the results of extended FWI (EFWI); that is, the minimization of the objective function expressed in equation 5. Figure 12 shows the zero subsurface offset section, displayed in color to facilitate the analysis of the long-wavelength components of the velocity. These components are similar in the final model and in the initial model. However, in contrast with the simple FWI, reflectivity is now imaged across the section, though misplaced and only partially focused because of the persistent errors in the long-wavelength components of the velocity. Figure 15 shows three sections of the difference cube obtained by subtracting the starting model from the final model. These sections are taken at fixed horizontal coordinates and are functions of depth and subsurface offset. These images are analogous to migration subsurface-offset common image gathers; they show the lack of focusing of the model along the subsurface-offset axis.

Figure 13 and 16 show the results of the proposed tomographic full waveform inversion (TFWI). To compute these results we minimized an objective function defined by the sum of equation 5 and equation 6. The zero subsurface-offset section (Figure 13) shows remarkable convergence towards the true model, in particular in the middle of the section where the model is well illuminated by the modeled data. The offset-domain common image gathers

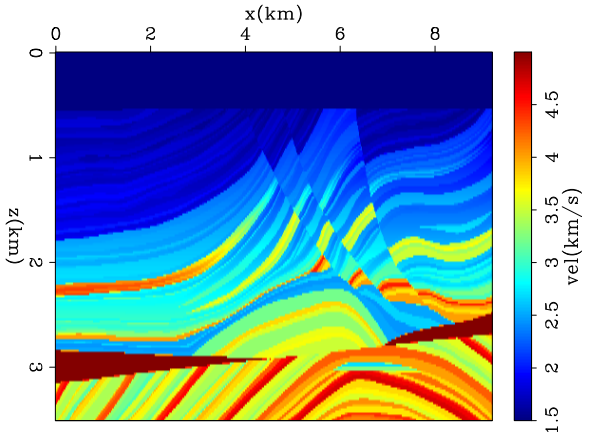


Figure 9: Modified Marmousi model used for numerical tests. [**ER**] biondo1/. seg1-vtrue

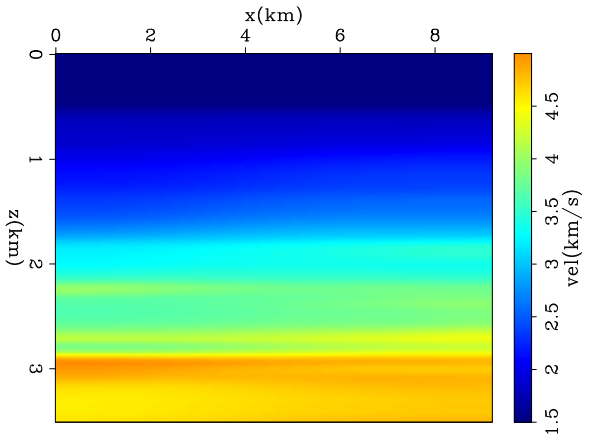


Figure 10: Starting model for inversions. [**ER**] biondo1/. seg1-init

shown in Figure 16 confirm that the final model correctly focuses the reflected events. Similarly to the previous example (Gaussian anomaly), the gathers shown in Figure 16 appear artificially focused around zero subsurface offset. As discussed previously, this appearance is caused by the DSO term in the objective function forcing the focusing of the image even beyond what would have been the focusing with the true model.

Finally, Figure 14 shows normalized values of the norm of the data residuals as a function of iterations for all three methods we discussed: FWI in blue, EFWI (i.e. minimizing only objective function 5) in red and TFWI (i.e. minimizing sum of objective functions 5 and 6) in magenta. Notice that in the case of TFWI, even if the objective function had two terms, only the value of J_{EFWI} (i.e. the data fitting term) is plotted in the graph. The graphs show that FWI has not converged even after hundreds of iterations. The data residuals decrease more quickly for EFWI than for TFWI because EFWI does not need substantial changes in the long-wavelength components of the model to fit the data. However, the magnitude of the final residuals is comparable between the two methods.

CONCLUSIONS

We presented and tested a method to invert for all wavelengths of the velocity model simultaneously from seismic reflection data. The method does not suffer from the global convergence limitations of conventional full waveform inversion. The method is based on

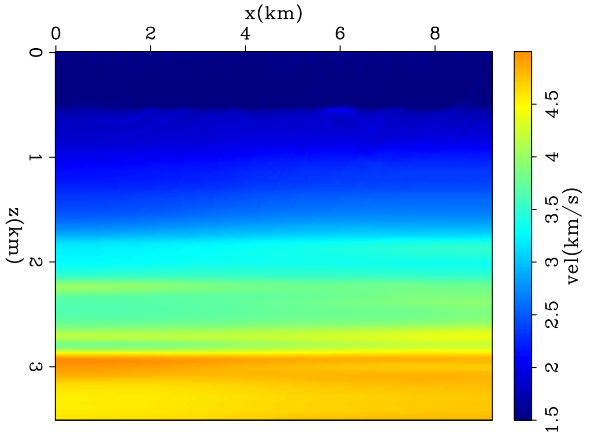


Figure 11: Final model when minimizing objective function 1 (FWI). [**CR**] biondo1/. seg1-inv0

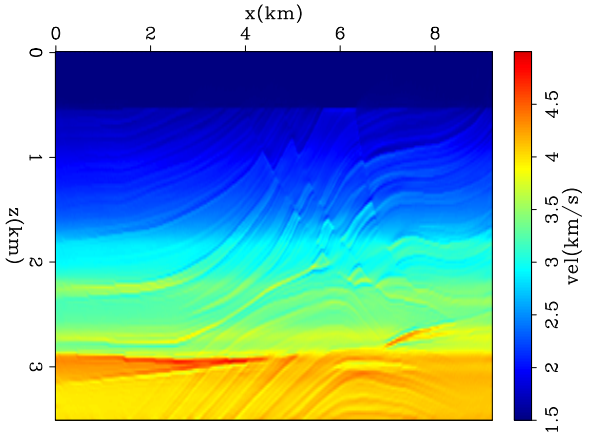


Figure 12: Final model when minimizing objective function 5 (EFWI). [**CR**] biondo1/. seg1-inv1

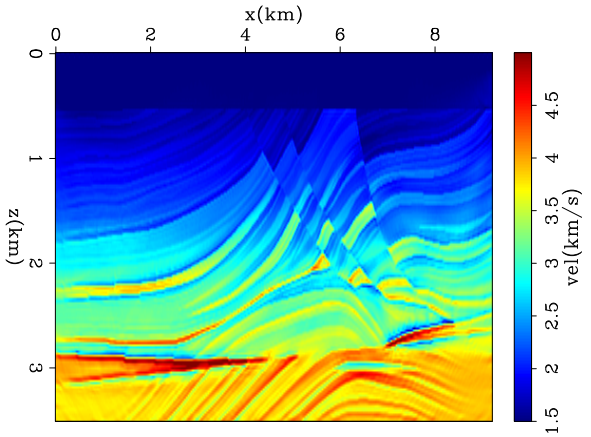


Figure 13: Final model when minimizing sum of objective functions 5 and 6 (TFWI). [**CR**] biondo1/. seg1-inv2

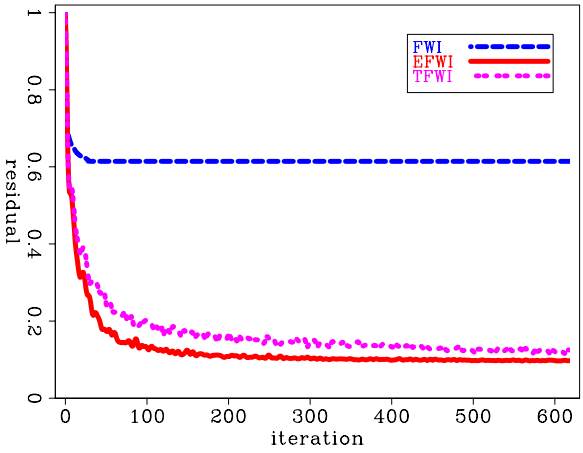


Figure 14: Graphs of data residuals as function of iterations for: FWI (blue), EFWI (red), and TFWI (magenta). [**CR**] biondo1/. seg1-inv-all-fixed

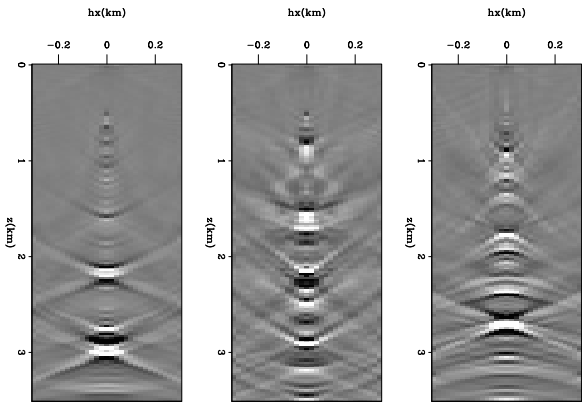


Figure 15: Final model subtracted from initial model taken at x=2.5, 5, and 7.5 km for EFWI. [**CR**] biondo1/. seg1-inv1-dv1

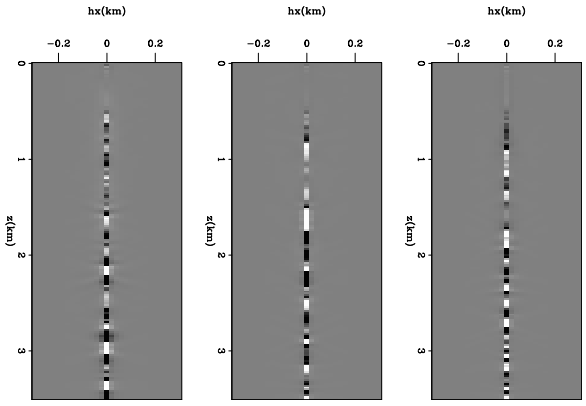


Figure 16: Final model subtracted from initial model taken at x=2.5, 5, and 7.5 km for TFWI. [**CR**] biondo1/. seg1-inv2-dv1

the extension of the velocity model to subsurface offsets and on the addition of a reflection focusing objective function to the conventional FWI data-fitting objective function. The Marmousi model numerical tests indicate that the method can robustly converge to very accurate models when FWI fails. The numerical solution of the wave equation with the extended velocity is computationally expensive. The excellent results, and the computational expense, justify the search of an approximate solution that still enables the simultaneous solution for all wavelengths in the model. We present such approximation in a companion report (Almomin and Biondi, 2012).

REFERENCES

Almomin, A. and B. Biondi, 2012, Tomographic full waveform inversion: Practical and computationally feasible approach: SEP-Report, **147**, 13–26.

Biondi, B., 2006, 3D Seismic Imaging: Society of Exploration Geophysicists.

———, 2010, Velocity estimation by image-focusing analysis: Geophysics, **75**, U49–U60.

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

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

Clapp, R. G., B. Biondi, and J. F. Claerbout, 2004, Geologically constrained migration velocity analysis: Geophysics, **69**, 533–546.

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

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

Tarantola, A., 1987, Inverse problem theory: Elsevier.

Zhang, Y. and B. Biondi, 2012, Residual moveout-based wave-equation migration velocity analysis in 3-D: SEP-Report, **147**, 27–38.

Tomographic full waveform inversion: Practical and computationally feasible approach

Ali Almomin and Biondo Biondi

ABSTRACT

We provide an alternative formulation to tomographic full-waveform inversion which computes the backscattering and the forward scattering components of the model separately. To maintain high resolution results of tomographic full-waveform inversion, the two components of the gradient are mixed based on a Fourier domain scale separation. This formulation is based on the Born approximation where the medium parameters are broken into a long wavelength and short wavelength components. This approximation has an underlying assumption that the data contain primaries only without multiples. The results of the Marmousi model show that convergence is possible even with large errors in the initial model that would have prevented convergence to conventional full-waveform inversion.

INTRODUCTION

Seismic velocity-analysis methods can be divided into two major groups. First, there are techniques that aim at minimizing the misfit in the data domain, such as full-waveform inversion (Tarantola, 1984; Pratt, 1999; Luo and Schuster, 1991). Second, there are other techniques that aim at improving the quality in the image domain, such as migration velocity analysis (MVA) (Symes and Carazzone, 1991; Biondi and Sava, 1999; Shen, 2004; Zhang et al., 2012). 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 effect 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. 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; it requires the starting model to be very close to the true model to avoid converging to local minima.

The conventional solution is to invert first for the velocity model using MVA techniques and then use the output as the initial model for FWI. However, this practice might not work if the results of MVA are not accurate enough for FWI to start. This could be a result

of the larger null space that forward-scattered wavefields do not constrain. Moreover, the convergence rate of the MVA techniques is going to be sub-optimal, since they do not use all of the information in the data.

In a companion paper (Biondi and Almomin, 2012), we present a generalize framework called Tomographic Full Waveform Inversion (TFWI) that combines both FWI and WEMVA techniques. This generalized approach utilizes all the components of seismic data to invert for the medium parameters without the cycle-skipping problem. This is achieved in two steps: first by extending the wave equation and adding an offset axis to the velocity model, and second by adding a regularization term that drives the solution towards the zero subsurface offset (Symes, 2008). However, this velocity extension makes the propagation considerably more expensive because each multiplication by velocity becomes a convolution over the subsurface offset axis.

In this paper, we present an approximation that significantly reduces the computational cost of TFWI while maintaining its desirable characteristic of enabling the simultaneous inversion of all wavelengths of the model. First, we use the Born approximation to break the extended velocity model into two components: a background component affecting transmission and a perturbation component affecting reflections. Second, we reduce the background component to zero subsurface offset while keeping the perturbation component at all offsets. This simplification of the background component is the key in reducing the computation cost since the convolution with velocity in propagation becomes a multiplication instead of convolution. However, breaking the model into two components hinders the simultaneous inversion of different wavelengths of the model. Moreover, the scale separation by Born approximation is not perfect since the reflection component of a certain frequency or angle can affect the transmission component of another frequency or angle. Therefore, we add a final step where we mix the gradients of each component and then separate them in Fourier domain.

Another potential advantage of the method presented in this paper compared to the complete method presented in Biondi and Almomin (2012) is the ability to handle variable density without having to adapt a more complicated form of the wave-equation. In the complete method, the inversion tries to fit the data using a velocity model only. Variations in the velocity model will cause both a transmission and a reflection effect. This becomes an issue in the presence of density variations that only change the reflections of the data without affecting the transmission. In the efficient method, this issue can be avoided since the reflectivity is separated from the background. Therefore, the reflectivity part of the model can absorb any reflection effects that do not fit the background part of the model.

SCALE SEPARATION

Although the next derivation is done in frequency domain, the actual implementation is done in time domain. First, we start with the tomographic full waveform objective function J_{TFWI} which can be written as:

$$J_{\text{TFWI}}(\mathbf{v}(\mathbf{h})) = \|\mathbf{d}\left(\mathbf{v}(\mathbf{h})\right) - \mathbf{d}_{\text{obs}}\|_2^2 + \epsilon\|\mathbf{A}\mathbf{v}(\mathbf{h})\|_2^2, \tag{1}$$

where $\mathbf{h}$ is the offset, $\mathbf{v}(\mathbf{h})$ is the extended velocity model, $\mathbf{d}(\mathbf{v}(\mathbf{h}))$ is the modeled data, $\mathbf{d}_{\text{obs}}$ is the observed surface data, ϵ is a scalar weight of the regularization term and $\mathbf{A}$ is a

regularization operator. The modeled data $\mathbf{d}(\mathbf{v}(\mathbf{h}))$ is computed as:

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

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 $G(\mathbf{x}_s, \mathbf{x}, \omega; \mathbf{v}(\mathbf{h}))$ satisfies:

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

where $*$ denotes a convolution operator over the subsurface offset axis (Symes, 2008; Biondi and Almomin, 2012). The first simplification of the extended velocity model is to separate it into a background and a perturbation as follows:

$$v^{-2}(\mathbf{x}, \mathbf{h}) = b(\mathbf{x}, \mathbf{h}) + r(\mathbf{x}, \mathbf{h}), \tag{4}$$

where $b(\mathbf{x})$ is the background component, which is a smooth version of the slowness squared and $r(\mathbf{x})$ is the perturbation component. This separation assumes that $b(\mathbf{x})$ will contain the transmission effects and $r(\mathbf{x})$ will contain the reflection effects. Depending on the error of the initial background velocity, the perturbation component can extend across several subsurface offsets so it is important to keep its offset 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 restrict it to zero offset. A physical interpretation is that the wave speed is not expected to vary much across different angles, at least in the isotropic case that we are analyzing. Therefore, the extent of background component across subsurface offsets can be reduced. If the velocity is expected to vary, the same reduction can be applied while keeping more than zero subsurface offset. In our derivation, we reduce the background to only the zero subsurface offset as follows

$$v^{-2}(\mathbf{x}, \mathbf{h}) \approx b(\mathbf{x}) + r(\mathbf{x}, \mathbf{h}). \tag{5}$$

After the separation and reduction, the Born approximation can be used to linearize wave equation where the data is assumed to contain primaries only. The linearized wave equation defines the data as follows:

$$d(\mathbf{x}_s, \mathbf{x}_r, \omega; \mathbf{b}, \mathbf{r}(\mathbf{h})) = -\omega^2 f(\omega) \sum_{\mathbf{x}, \mathbf{h}} G(\mathbf{x}_s, \mathbf{x} - \mathbf{h}, \omega; \mathbf{b}) r(\mathbf{x}, \mathbf{h}) G(\mathbf{x} + \mathbf{h}, \mathbf{x}_r, \omega; \mathbf{b}), \tag{6}$$

where the Green's functions now satisfy the conventional acoustic wave equation as follows:

$$\left(\omega^2 b(\mathbf{x}) + \nabla^2\right) G(\mathbf{x}_s, \mathbf{x}, \omega) = \delta(\mathbf{x}_s - \mathbf{x}), \tag{7}$$

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

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

$$\mathbf{d} = \mathbf{L}(\mathbf{b}) \mathbf{r}(\mathbf{h}), \tag{9}$$

where $\mathbf{L}$ is the Born modeling operator. We now define a new objective function for the efficient TFWI inversion as:

$$J_{\text{ETFWI}}(\mathbf{b}, \mathbf{r}(\mathbf{h})) = \|\mathbf{L}(\mathbf{b}) \mathbf{r}(\mathbf{h}) - \mathbf{d}_{\text{obs}}\|_2^2 + \epsilon \|\mathbf{A}\mathbf{r}(\mathbf{h})\|_2^2. \tag{10}$$

Notice that there are similarity with the regularized linearized inversion proposed by Clapp (2005), with the important difference that here we are including both the background and the perturbation components as variables in this objective function. 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:

$$\begin{aligned}\Delta d(\mathbf{x}_s, \mathbf{x}_r, \omega; \mathbf{b}_0, \mathbf{r}(\mathbf{h})) = &\sum_{\mathbf{x},\mathbf{y},\mathbf{h}} \\ &\omega^4 f(\omega) G(\mathbf{x}_s, \mathbf{y}-\mathbf{h}, \omega; \mathbf{b}_0) r(\mathbf{y}, \mathbf{h}) G(\mathbf{y}+\mathbf{h}, \mathbf{x}, \omega; \mathbf{b}_0) \Delta b(\mathbf{x}) G(\mathbf{x}, \mathbf{x}_r, \omega; \mathbf{b}_0) + \\ &\omega^4 f(\omega) G(\mathbf{x}_s, \mathbf{x}, \omega; \mathbf{b}_0) \Delta b(\mathbf{x}) G(\mathbf{x}, \mathbf{y}-\mathbf{h}, \omega; \mathbf{b}_0) r(\mathbf{y}, \mathbf{h}) G(\mathbf{y}+\mathbf{h}, \mathbf{x}_r, \omega; \mathbf{b}_0),\end{aligned} \tag{12}$$

where $\mathbf{y}$ is the perturbation coordinates. As we can see in the previous equation, the tomographic operator correlates a background and a scattered wavefield from both source and receiver sides. The scattered wavefields are computed by correlating a background wavefield with the reflectivity model $\mathbf{r}$ and then propagating again to all model locations. The forward tomographic operator can be written in a compact notation as follows:

$$\Delta \mathbf{d} = \frac{\partial \mathbf{L}}{\partial \mathbf{b}} \mathbf{r}(\mathbf{h}) \Delta \mathbf{b} = \mathbf{T} \Delta \mathbf{b}, \tag{13}$$

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

$$g_{\mathbf{r}}(\mathbf{x}, \mathbf{h}) = \frac{\partial J}{\partial \mathbf{r}(\mathbf{h})} = \mathbf{L}^* \Delta \mathbf{d}, \tag{14}$$

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

$$g_{\mathbf{b}}(\mathbf{x}) = \frac{\partial J}{\partial \mathbf{b}} = \mathbf{T}^* \Delta \mathbf{d}. \tag{15}$$

It is important to notice that reflectivity has two roles in computing the background model gradient. First, it is used to computed the data residuals. Second, it is used to scatter the background wavefields and compute the scattered wavefields of the tomographic operator.

SCALE MIXING

A direct use of the gradients is to update their corresponding models directly. However, this would hinders the simultaneous inversion of different wavelengths of the model. Hence, we first mix the two gradients and then separate them in Fourier domain to get the update of each model as follows:

$$s_{\mathbf{b}}(\mathbf{x}) = \mathbf{C}_{\mathbf{b}}(g_{\mathbf{b}}(\mathbf{x}) + g_{\mathbf{r}}(\mathbf{x}, \mathbf{h}=0)), \tag{16}$$

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 can compute the update of the perturbation model as:

$$s_{\mathbf{r}}(\mathbf{x},\mathbf{h}) = \mathbf{C_r}(g_{\mathbf{b}}(\mathbf{x}) + g_{\mathbf{r}}(\mathbf{x},\mathbf{h})), \tag{17}$$

where $s_{\mathbf{r}}(\mathbf{x},\mathbf{h})$ is the search direction of the perturbation model and $\mathbf{C_r}$ 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.

For the examples shown 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. Also, the two filters were designed such that they always sum to one at all wavelength to maintain the energy of the gradients.

SYNTHETIC EXAMPLES

Gaussian Model

For the first of synthetic example, we use a two-layer model with a Gaussian anomaly. The velocities for the top layer, bottom layer and peak of the Guassian anomaly are 3000 m/s, 3500 m/s and 3500 m/s, respectively. Figure 1 shows the true velocity model. A Ricker wavelet with a fundamental frequency of 15 Hz and temporal sampling of 1.5 ms is used as a source function to model the data. There are 151 fixed receivers with a spacing of 20 m and 31 sources with a spacing of 100 m. The maximum offset used is 1500 m in both sides and the initial model is a constant model of 3000 m/s velocity.

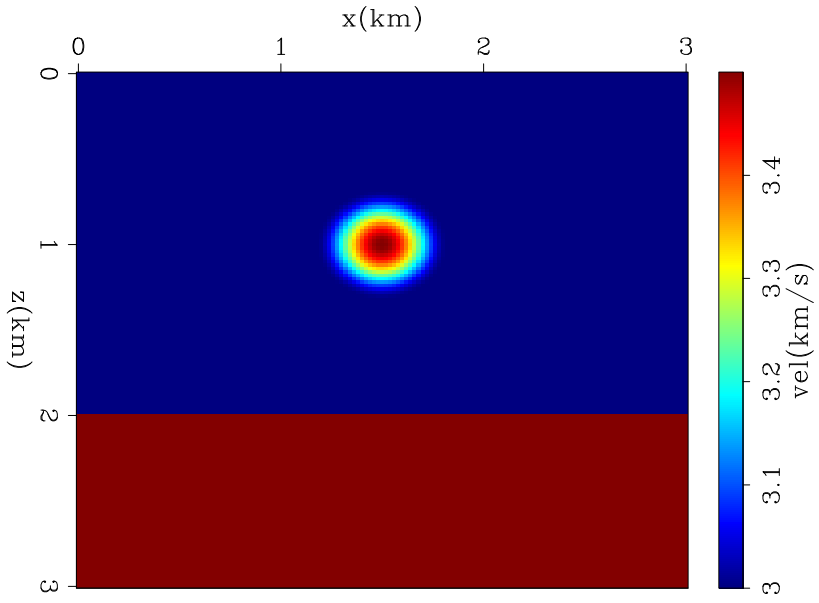


Figure 1: The true velocity of Gaussian model. [**ER**] ali1/. gauss-true

We first modeled the observed data with the conventional acoustic wave-equation nonlinear modeling operator. Then, we computed the initial data using the same nonlinear

operator on the initial model. The two datasets are subtracted from each other to remove the direct arrivals. We then use the subtracted data as the "observed data" in equation 10.

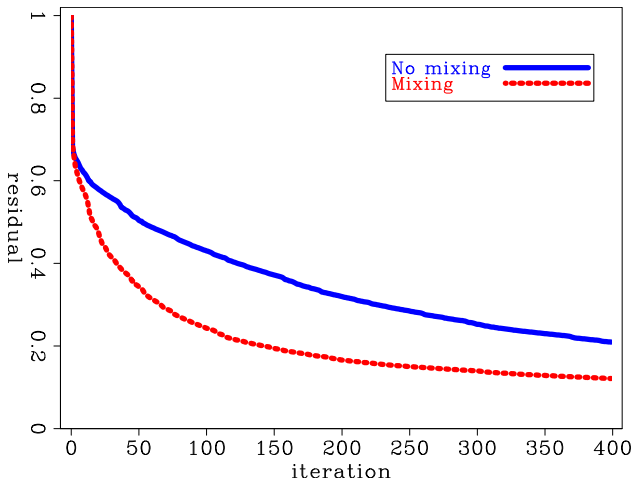


Figure 2: The data residual norm as a function of iterations of Gaussian model inversion. [**CR**] ali1/. gauss-res

We show the results of running 400 iterations of minimizing equation 10 with and without the scale mixing described in equations 16 and 17. Figure 2 shows the normalized residual of the data fitting part, described by the first term of equation 10, as a function of iterations. The residuals of both inversions decrease monotonically without getting stuck in a local minima. However, scale mixing shows a much faster convergence rate.

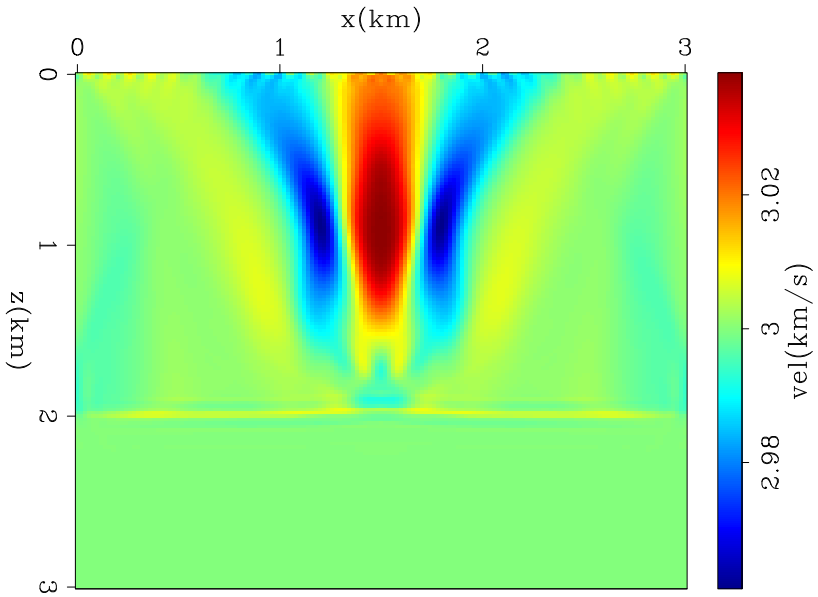


Figure 3: The inverted background of the Gaussian model example. [**CR**] ali1/. gauss-imodel

Figures 3 and 4 shows the final background and reflectivity models without scale mixing. The background has a low vertical resolution with relatively strong side-lobes around the anomaly location. The reflectivity model has a low horizontal resolution where the sides of the anomaly are not well illuminated. This lack of resolution in both models is expected due to the limited acquisition.

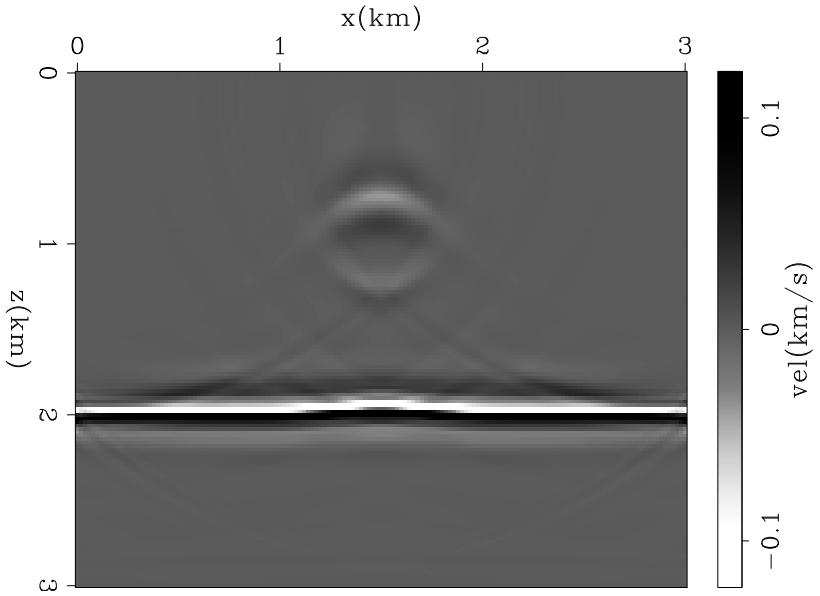


Figure 4: The inverted reflectivity of the Gaussian model example. [**CR**] ali1/. gauss-imodelr

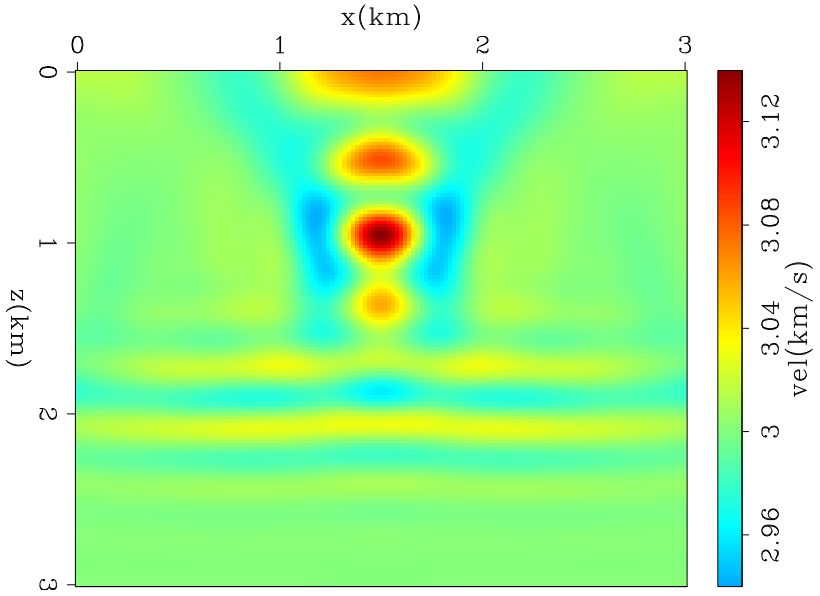


Figure 5: The inverted background of the Gaussian model example with mixing of scales. [**CR**] ali1/. gauss-imodel2

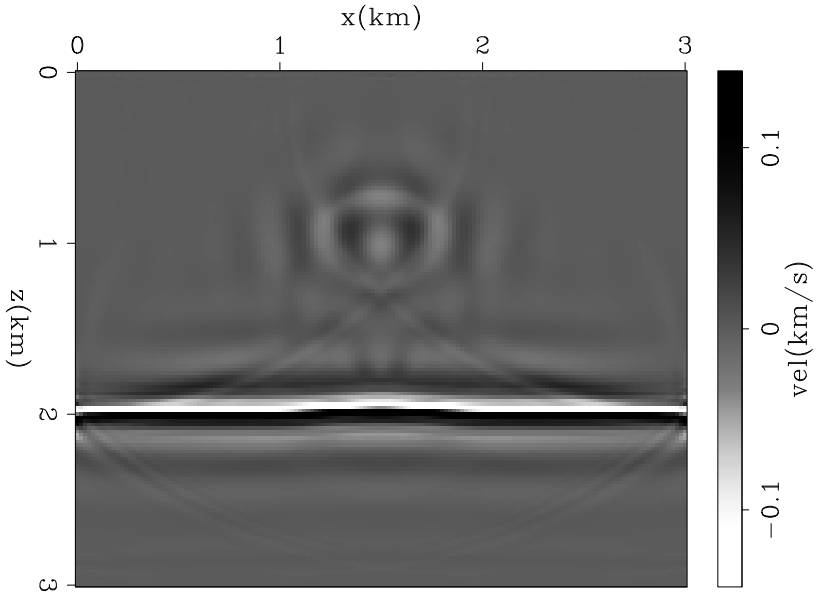


Figure 6: The inverted reflectivity of the Gaussian model example with mixing of scales. [**CR**] ali1/. gauss-imodelr2

Figures 5 and 6 shows the final background and reflectivity models with scale mixing. The background has a much better vertical resolution that locates the anomaly at the correct depth. Moreover, the peak amplitude of the anomaly is more than three times stronger compared to Figure 3 while the side-lobes remained at the same strength. The reflectivity model shows substantial improvements in resolution as well. The sides of the anomaly seem to be better illuminated after scale mixing. Although reflectivity is linear with respect to the data, the inverted model seems to benefit from the scale mixing. Therefore, the contribution of the background gradient seem to add to the null space components of the reflectivity model, and vice-versa.

Marmousi Model

A modified Marmousi model is used for the next synthetic example where 500m of water layer is added to the top. Figure 7 shows the true velocity model. A Ricker wavelet with a fundamental frequency of 15 Hz and temporal sampling of 1.5 ms is used as a source function to model the data. There are 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 8 which is obtained by strongly smoothing the true model laterally. The conventional FWI fails when starting from this initial model (Biondi and Almomin, 2012).

Similar to the Gaussian model, we modeled the observed data with the conventional acoustic wave-equation nonlinear modeling operator. Then, we computed the initial data using the same nonlinear operator on the initial model shown in Figure 8. The two datasets are subtracted from each other to remove the direct arrivals. We then use the subtracted data as the "observed data" in equation 10.

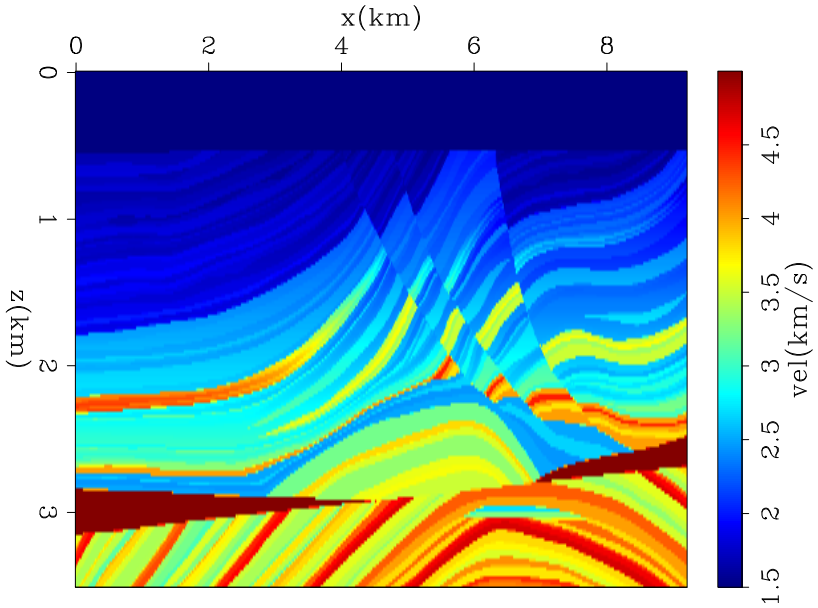


Figure 7: The true velocity of marmousi model. [**ER**] ali1/. vtrue2

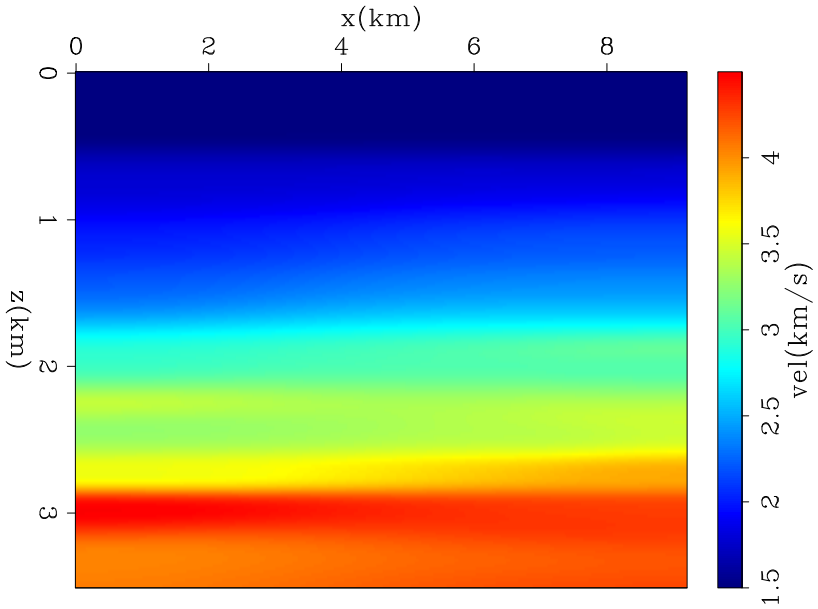


Figure 8: The initial velocity of marmousi model. [**ER**] ali1/. init

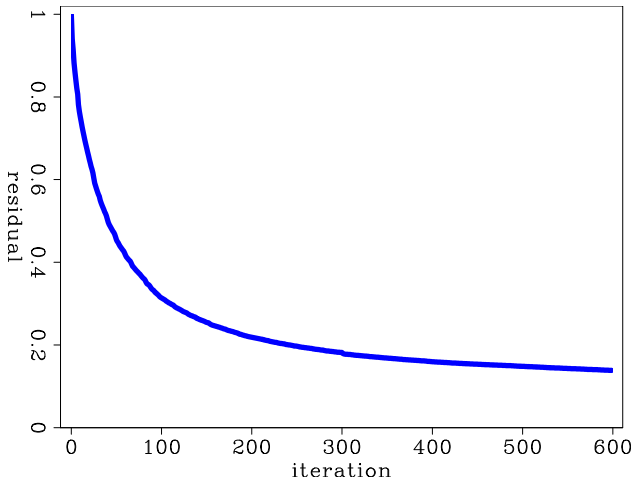


Figure 9: The data residual norm as a function of iterations of marmousi model inversion. [**CR**] ali1/. resn

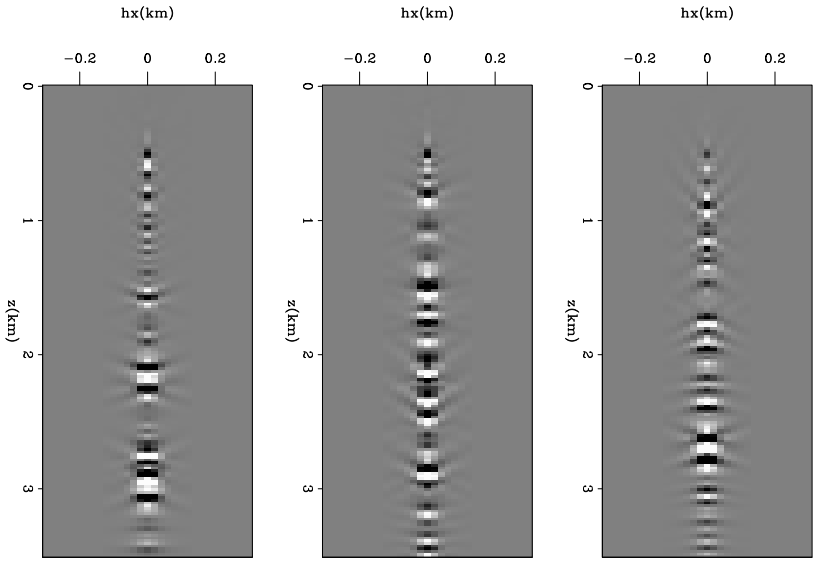


Figure 10: Three offset-domain common-image gathers of the inverted reflectivity model at x=2.5, 5, 7.5 km of marmousi model. [**CR**] ali1/. invdv

We show the results of running 600 iterations of minimizing equation 10 where the gradients are mixed as described in equations 16 and 17. Figure 9 shows the normalized residual of the data fitting part, described by the first term of equation 10, as a function of iterations. The residuals decrease monotonically without getting stuck in a local minima. Figure 10 shows three constant-location sections of the final reflectivity model that are a function of depth and subsurface offset. The three images are focused around the middle, i.e. the zero subsurface offset, which indicates convergence towards the correct background model.

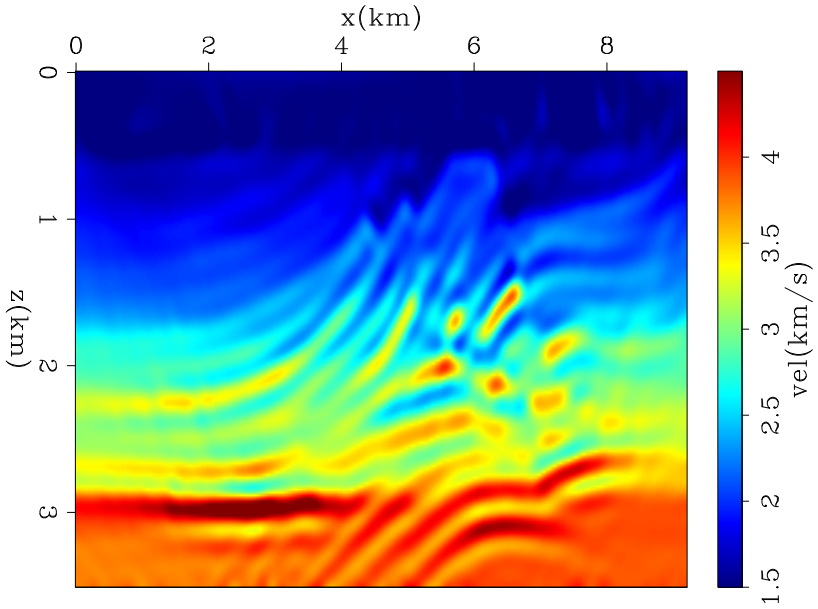


Figure 11: The inverted background of marmousi model. [**CR**] ali1/. invv

Figure 11 shows the final background model. It shows great resolution in both the vertical and horizontal directions that captures the main features of the true model. Figure 12 shows zero subsurface offset of the final reflectivity model. Finally, Figure 13 shows the total inverted model which is the sum of the background and reflectivity. It shows remarkable convergence towards the true model.

CONCLUSIONS

We presented a practical approach to tomographic full-waveform inversion where the computational cost is significantly reduced. This was achieved by first breaking the model into a background component and a perturbation component, and then by restricting the offset axis of the background component to zero subsurface offset only. Breaking the model into two components assumes the data contain primary only. However, we managed to maintain the simultaneous inversion of different wavelengths of the model by mixing the gradients of the two components in Fourier domain using a high-pass and a low-pass filters. The synthetic examples show remarkable results even when the initial model had large errors. More sophisticated mixing schemes need to be further investigated.

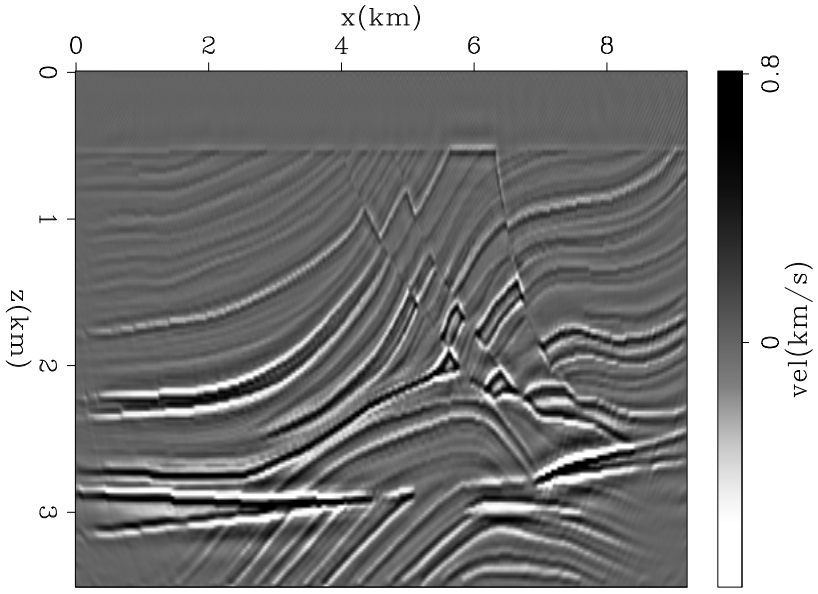


Figure 12: The inverted reflectivity at h=0 of marmousi model. [**CR**] ali1/. invr1

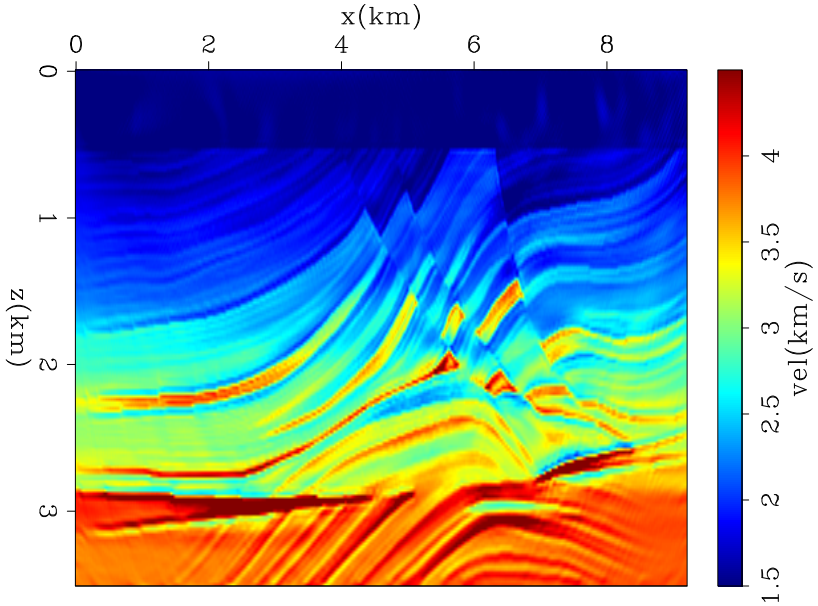


Figure 13: The total inversion results of marmousi model. [**CR**] ali1/. inv

ACKNOWLEDGMENTS

We would like to thank the Stanford Exploration Project affiliate companies for financial support. Almomin would like the thank Saudi Aramco for supporting his Ph.D. studies at Stanford University.

REFERENCES

Biondi, B. and A. Almomin, 2012, Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis: SEP-Report, **147**, 1–12.

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

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

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

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.

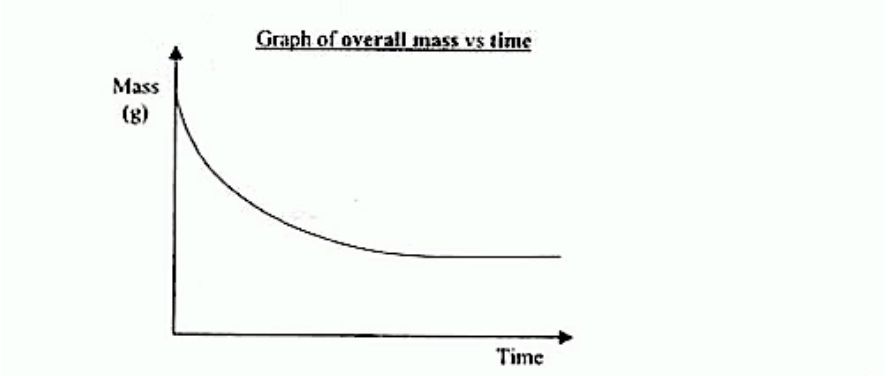
Symes, W. W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790.

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 Technical Program Expanded Abstracts, **31**, *submitted for publication.*



1. Explain the shape of the graph.

It's curvy, with a higher bit at the end and a rather aesthetically pleasing slope downwards towards a pretty flat strait bit. The actual graph itself consists of 2 strait lines meeting at the lower left hand corner of the graph and moving away at a 90° angle. Each line has an arrow head on the end.

Some people don't look up until they are flat on their backs.

This quote means some people can't look up because something has happened to their necks so they cant look up
for example, if a person gets kicked in the neck by a Kung fu midget, they will not be able to look up.

10/5

Residual moveout-based wave-equation migration velocity analysis in 3-D

Yang Zhang and Biondo Biondi

ABSTRACT

In the previous report (SEP-143), we proposed a novel approach to perform Wave-Equation Migration Velocity Analysis (WEMVA) using Residual Moveout (RMO) parameters. The 2-D examples we tested showed that this approach is immune to the cycle-skipping problem, and it produces high-quality gradients. In this paper, we try to address the theoretical and practical challenges of extending this method to three-dimensions. Specifically, we propose a new parameterization for the 3-D angle-domain residual moveout and a generalized 3-D RMO-based WEMVA theory. We also address the practical issues in 3-D transforms between subsurface offset and angle-domain image gathers.

INTRODUCTION

Wave-equation migration velocity analysis (WEMVA) (Biondi and Sava, 1999; Biondi and Symes, 2004) belongs to a family of methods for solving the reflection tomography problem. These methods aim to estimate migration velocity by employing wave-equation operators. WEMVA formulates velocity estimation as an optimization problem. Evaluating the flatness of the angle-domain common-image gathers (ADCIGs) is currently the most favored choice when formulating WEMVA objective functions. The WEMVA objective function is 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.
2. Back-projection of the image perturbation into the velocity model using the image-domain wave-equation tomographic operator.

In the previous report (Zhang and Biondi, 2011), we proposed a new WEMVA method for maximizing the angle-stack power of ADCIGs. The key innovations in this method are as follows: first, we approximate the ADCIG, $I(s)$ (migrated using the current slowness s), with the initial ADCIG $I(s_0)$ followed by an RMO(Residual Moveout) on $I(s_0)$; then we design a new image-space tomography operator that describes the kinematic relation between the model slowness perturbation and the RMO parameters. Using several 2-D examples, we have shown that this method: 1) avoids cycle-skipping, 2) does not require manual picking of the moveout parameters, and 3) can robustly improve the flatness of the angle gathers.

Although the three-dimensional extension of our method is conceptually straightforward, to actually implement a 3-D prototype requires substantial efforts on both the theoretical

and practical aspects. In this paper, we present our initial work toward a 3-D implementation of our approach. The rest of the paper is organized as follows: First, we present the upgraded theory of our RMO WEMVA method for handling 3-D ADCIGs. Second, we discuss some of the challenges of transforming 3D common-image gathers (CIGs) between the offset and angle domains. We show a synthetic example as a prove of concept for our 3-D RMO WEMVA formulation.

3-D RMO WEMVA METHOD

We first review the rationale and key derivations of our RMO WEMVA theory in the 2-D case, and then present the 3-D generalization to the theory by making analogies. Practical experience indicates that slowness is a better parameterization than velocity. For clarity, the model space is parameterized as slowness unless otherwise specified.

2-D Theory Review

We start from the 'classical' maximum-stack-power objective function

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

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

Objective functions defined this way are prone to cycle-skipping (Symes, 2008). To tackle this issue, we *approximate* objective function 1 with the following one:

$$J(\rho(s)) = \frac{1}{2} \sum_x \sum_z \left[\sum_\gamma I(z + \rho \tan^2 \gamma, \gamma, x; s_0) \right]^2, \tag{2}$$

in which s is the model slowness, $I(z, \gamma, x; s_0)$ is the prestack image migrated with some initial slowness s_0, and $\rho \tan^2 \gamma$ is the residual moveout (RMO) function we choose to characterize the kinematic difference (Biondi and Symes, 2004) between $I(s)$ and $I(s_0)$.

The meaning of equation 2 can be easily explained. As the model changes from s_0 to s, it leads to the change of the image kinematics between $I(s_0)$ and $I(s)$, where the differences are characterized by the moveout parameter ρ. Since $I(s)$ will be *kinematically the same* as $I(s_0)$ being applied moveout $\rho(s)$, if we substitute the former image ($I(s)$) with the latter one, we transit from equation 1 to equation 2. *Notice that the new objective function is expressed as a function of only the moveout parameter ρ, while the ρ parameter is then related to the model slowness s.*

Furthermore, notice that equation 2 weights the strong-amplitude events more heavily. To make the gradient independent from the strength of reflectors, we further replace 2 with the following semblance objective function:

$$J_{Sm}(\rho(s)) = \frac{1}{2} \sum_x \sum_z \frac{\sum_{z_w} \left(\sum_\gamma I(z + z_w + \rho \tan^2 \gamma, \gamma, x; s_0) \right)^2}{\sum_{z_w} \sum_\gamma I^2(z + z_w + \rho \tan^2 \gamma, \gamma, x; s_0)}, \tag{3}$$

where z_w is a local averaging window of length L along the depth axis. For the rest of the paper, the summation interval of z_w is always $[-L/2, L/2]$; thus we can safely omit the summation bounds for concise notation.

We will use gradient-based methods to solve this optimization problem. The gradient given by the objective function 3 is

$$\frac{\partial J_{S_m}}{\partial s} = \frac{\partial J_{S_m}}{\partial \rho}\frac{\partial \rho}{\partial s}, \tag{4}$$

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.

To evaluate the derivative of the moveout parameter ρ with respect to the slowness model s, we define an auxiliary objective function in a fashion similar to the one employed by Luo and Schuster (1991) for cross-well travel-time tomography. The auxiliary objective function is defined for each image point (z, x) as:

$$J_{\text{aux}} = \sum_{z_w}\sum_{\gamma} I(z + z_w + (\rho \tan^2\gamma + \beta), \gamma, x; s_0) I(z, \gamma, x; s) \tag{5}$$

where β is a simple vertical shift introduced to accommodate bulk shifts in the image introduced by variation in the migration velocity. Notice that the semblance in objective function 3 is independent of β because a bulk shift does not affect the power of the stack; therefore we do not include β in 3.

The explanation for equation 5 is as follows: The moveout parameters $\beta(s)$ and $\rho(s)$ are chosen to describe the kinematic difference between the initial image $I(z, \gamma, x; s_0)$ and the new image $I(z, \gamma, x; s)$. In other words, if we apply the moveout to the initial image, the resulting image $I(z + (\rho \tan^2\gamma + \beta), \gamma, x; s_0)$ will be the same as the new image $I(z, \gamma, x; s)$ in terms of kinematics; this is indicated by a maximum of the cross-correlation between the two.

Given the auxiliary objective function 5, $\partial\rho/\partial s$ can be found using the rule of partial derivatives for implicit functions. We compute the gradient of 5 around the maximum at $s = s_0$ and $\rho = \beta = 0$; consequently

$$\begin{cases} \frac{\partial J_{\text{aux}}}{\partial \rho} = 0 \\ \frac{\partial J_{\text{aux}}}{\partial \beta} = 0 \end{cases} \tag{6}$$

We differentiate equation 6 with respect to s, which gives

$$\begin{bmatrix} \frac{\partial^2 J_{\text{aux}}}{\partial \rho^2} & \frac{\partial^2 J_{\text{aux}}}{\partial \rho \partial \beta} \\ \frac{\partial^2 J_{\text{aux}}}{\partial \rho \partial \beta} & \frac{\partial^2 J_{\text{aux}}}{\partial \beta^2} \end{bmatrix} \begin{bmatrix} \frac{\partial \rho}{\partial s} \\ \frac{\partial \beta}{\partial s} \end{bmatrix} = - \begin{bmatrix} \frac{\partial J_{\text{aux}}}{\partial \rho \partial s} \\ \frac{\partial J_{\text{aux}}}{\partial \beta \partial s} \end{bmatrix}. \tag{7}$$

Now we need to invert a Jacobian to get $\partial\rho/\partial s$. We denote $\dot{I}, \ddot{I}$ to be the first and second

order z derivatives of image I, then define the following:

$$\begin{aligned}
\frac{\partial^2 J_{\text{aux}}}{\partial \rho^2} &= \sum_{z_w}\sum_{\gamma} \ddot{I}(z+z_w,\gamma,x;s_0)\tan^4\gamma I(z+z_w,\gamma,x;s) = E_{11}(z,x) \\
\frac{\partial^2 J_{\text{aux}}}{\partial \rho \partial \beta} &= \sum_{z_w}\sum_{\gamma} \ddot{I}(z+z_w,\gamma,x;s_0)\tan^2\gamma I(z+z_w,\gamma,x;s) = E_{12}(z,x) \\
\frac{\partial^2 J_{\text{aux}}}{\partial \beta^2} &= \sum_{z_w}\sum_{\gamma} \ddot{I}(z+z_w,\gamma,x;s_0) I(z+z_w,\gamma,x;s) = E_{22}(z,x). \qquad (8)
\end{aligned}$$

Let the inverse of matrix E be matrix F:

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

Then

$$\frac{\partial \rho}{\partial s}|_{s=s_0} = -\sum_{z_w}\sum_{\gamma}(F_{11}\tan^2\gamma + F_{12})\dot{I}(z+z_w,\gamma,x;s_0)\frac{\partial I(z+z_w,\gamma,x;s)}{\partial s}. \qquad (9)$$

Finally, we have the expression for the gradient

$$-\sum_{z_w}\sum_{\gamma}\sum_{z,x}\frac{\partial I(z+z_w,\gamma,x;s)}{\partial s}(F_{11}\tan^2\gamma + F_{12})\frac{\partial J_{Sm}}{\partial \rho}(z,x)\dot{I}(z+z_w,\gamma,x;s_0). \qquad (10)$$

The engineering translation of equation 10 is that first we compute the image perturbation $(F_{11}\tan^2\gamma + F_{12})\frac{\partial J_{Sm}}{\partial \rho}(z,x)\dot{I}(z+z_w,\gamma,x;s_0)$, then we backproject this perturbation to model slowness space using the tomography operator $\partial I(z+z_w,\gamma,x;s)/\partial s$.

Since we can compute the gradient in equation 10, any gradient-based optimization method can be used to maximize the objective function defined in equation 3. Nonetheless, in terms of finding the step size, it is more expensive to evaluate equation 3 (which is an approximation of equation 1 purely based on kinematics) than to evaluate the original objective function 1. In our implementation we choose 1 as the maximization goal while using the search direction computed from equation 3.

Extension to 3-D

As we shift from 2-D seismic surveys to 3-D ones, two extra dimensions (cross-line axis y and subsurface azimuth ϕ) are added to our seismic image. Therefore we denote the prestack image as $I(z,\gamma,\phi,x,y)$. Assuming there are m values for ϕ, then $\phi \in \{\phi_i : i = 1,2,...,m\}$. Now the maximum-stack-power objective function 1 can be generalized to

$$J(s) = \frac{1}{2}\sum_{x,y}\sum_{z}\left[\sum_{\phi_i}\sum_{\gamma} I(z,\gamma,\phi_i,x,y;s)\right]^2, \qquad (11)$$

in which we stack the gather along both γ and ϕ axes.

The next step is to choose a proper residual moveout parameterization for the 3-D ADCIGs, in which the moveout is a surface (defined by (γ,ϕ)) rather than a curve (of

γ). There are certainly more than one way to design such parameterization. As an initial attempt, we choose a straightforward approach, in which we separate the moveout surface into individual curves by azimuth ϕ. For each azimuth angle ϕ_i, we assign the curvature parameter ρ_i and the static shift parameter β, respectively. *Notice that all the curves share the same β parameter, because the center of the move-out surface at ($\gamma = 0, \forall\phi_i$) is shared by all curves.*

Under this parameterization, the 3-D counterpart of objective function 3 would be

$$J_{Sm}(\rho(s)) = \frac{1}{2}\sum_{x,y}\sum_{z}\sum_{\phi_i}\frac{\sum_{z_w}\left(\sum_{\gamma} I(z+z_w+\rho_i\tan^2\gamma,\gamma,\phi_i,x,y;s_0)\right)^2}{\sum_{z_w}\sum_{\gamma} I^2(z+z_w+\rho_i\tan^2\gamma,\gamma,\phi_i,x,y,;s_0)}, \tag{12}$$

where ρ *becomes a vector*, $\rho = \{\rho_i : i = 1, 2, ..., m\}$. The gradient formula 4 now turns into

$$\frac{\partial J_{S_m}}{\partial s} = \sum_{\rho_i}\frac{\partial J_{S_m}}{\partial \rho_i}\frac{\partial \rho_i}{\partial s}. \tag{13}$$

Because each ϕ_i is treated separately, we can compute $\partial J_{Sm}/\partial\rho_i$ in exactly the same way as we do in the 2-D case.

Analogously, we can define an auxiliary objective function for each image point that uncovers the $s \leftrightarrow \rho$ relationship:

$$J_{\text{aux}} = \sum_{z_w}\sum_{\gamma}\sum_{\phi_i} I(z+z_w+(\rho_i\tan^2\gamma+\beta),\gamma,\phi_i,x,y;s_0)I(z,\gamma,\phi_i,x,y;s). \tag{14}$$

Using the same trick of finding partial derivatives for implicit functions, equation 6 is generalized as

$$\begin{cases} \frac{\partial J_{\text{aux}}}{\partial \rho_1} = 0 \\ \frac{\partial J_{\text{aux}}}{\partial \rho_2} = 0 \\ \quad\vdots \\ \frac{\partial J_{\text{aux}}}{\partial \rho_m} = 0 \\ \frac{\partial J_{\text{aux}}}{\partial \beta} = 0 \end{cases} \tag{15}$$

We differentiate equation 15 with respect to s:

$$\begin{bmatrix} \frac{\partial^2 J_{\text{aux}}}{\partial {\rho_1}^2} & \frac{\partial^2 J_{\text{aux}}}{\partial\rho_1\partial\rho_2} & \dots & \frac{\partial^2 J_{\text{aux}}}{\partial\rho_1\partial\rho_m} & \frac{\partial^2 J_{\text{aux}}}{\partial\rho_1\partial\beta} \\ \frac{\partial^2 J_{\text{aux}}}{\partial\rho_2\partial\rho_1} & \frac{\partial^2 J_{\text{aux}}}{\partial {\rho_2}^2} & \dots & \frac{\partial^2 J_{\text{aux}}}{\partial\rho_2\partial\rho_m} & \frac{\partial^2 J_{\text{aux}}}{\partial\rho_2\partial\beta} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 J_{\text{aux}}}{\partial\rho_m\partial\rho_1} & \frac{\partial^2 J_{\text{aux}}}{\partial\rho_m\partial\rho_2} & \dots & \frac{\partial^2 J_{\text{aux}}}{\partial {\rho_m}^2} & \frac{\partial^2 J_{\text{aux}}}{\partial\rho_m\partial\beta} \\ \frac{\partial^2 J_{\text{aux}}}{\partial\beta\partial\rho_1} & \frac{\partial^2 J_{\text{aux}}}{\partial\beta\partial\rho_2} & \dots & \frac{\partial^2 J_{\text{aux}}}{\partial\beta\partial\rho_m} & \frac{\partial^2 J_{\text{aux}}}{\partial\beta^2} \end{bmatrix} \begin{bmatrix} \frac{\partial\rho_1}{\partial s} \\ \frac{\partial\rho_2}{\partial s} \\ \vdots \\ \frac{\partial\rho_m}{\partial s} \\ \frac{\partial\beta}{\partial s} \end{bmatrix} = - \begin{bmatrix} \frac{\partial^2 J_{\text{aux}}}{\partial\rho_1\partial s} \\ \frac{\partial^2 J_{\text{aux}}}{\partial\rho_2\partial s} \\ \vdots \\ \frac{\partial^2 J_{\text{aux}}}{\partial\rho_m\partial s} \\ \frac{\partial^2 J_{\text{aux}}}{\partial\beta\partial s} \end{bmatrix}. \tag{16}$$

We can calculate the $m+1$ by $m+1$ Jacobian matrix elements and the right-hand side

based on equation 14:

$$\begin{aligned}
\frac{\partial^2 J_{\text{aux}}}{\partial {\rho_i}^2} &= \sum_{z_w}\sum_{\gamma} \ddot{I}(z+z_w,\gamma,\phi_i,x,y;s_0)\tan^4\gamma I(z+z_w,\gamma,\phi_i,x,y;s) \\
\frac{\partial^2 J_{\text{aux}}}{\partial \rho_i \partial \rho_j} &\equiv 0 \qquad\qquad\qquad \forall \quad i\neq j \\
\frac{\partial^2 J_{\text{aux}}}{\partial \rho_i \partial \beta} &= \sum_{z_w}\sum_{\gamma} \ddot{I}(z+z_w,\gamma,\phi_i,x,y;s_0)\tan^2\gamma I(z+z_w,\gamma,\phi_i,x,y;s) \\
\frac{\partial^2 J_{\text{aux}}}{\partial \beta^2} &= \sum_{z_w}\sum_{\gamma}\sum_{\phi_i} \ddot{I}(z+z_w,\gamma,\phi_i,x,y;s_0) I(z+z_w,\gamma,\phi_i,x,y;s) \\
\frac{\partial^2 J_{\text{aux}}}{\partial \rho_i \partial s} &= \sum_{z_w}\sum_{\gamma} \dot{I}(z+z_w,\gamma,\phi_i,x,y;s_0)\tan^2\gamma \frac{\partial I(z+z_w,\gamma,\phi_i,x,y;s)}{\partial s} \\
\frac{\partial^2 J_{\text{aux}}}{\partial \beta \partial s} &= \sum_{z_w}\sum_{\gamma}\sum_{\phi_i} \dot{I}(z+z_w,\gamma,\phi_i,x,y;s_0) \frac{\partial I(z+z_w,\gamma,\phi_i,x,y;s)}{\partial s}.
\end{aligned} \tag{17}$$

Denoting matrix $\mathbf{F}=\{F_{i,j}\}$ to be the inverse of the Jacobian, then

$$\begin{bmatrix} \frac{\partial \rho_1}{\partial s} \\ \frac{\partial \rho_2}{\partial s} \\ \vdots \\ \frac{\partial \rho_m}{\partial s} \end{bmatrix} = -\begin{bmatrix} F_{1,1} & F_{1,2} & \cdots & F_{1,m-1} & F_{1,m} & F_{1,m+1} \\ F_{2,1} & F_{2,2} & \cdots & F_{2,m-1} & F_{2,m} & F_{2,m+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ F_{m-1,1} & F_{m-1,2} & \cdots & F_{m-1,m-1} & F_{m-1,m} & F_{m-1,m+1} \\ F_{m,1} & F_{m,2} & \cdots & F_{m,m-1} & F_{m,m} & F_{m,m+1} \end{bmatrix} \begin{bmatrix} \frac{\partial^2 J_{\text{aux}}}{\partial \rho_1 \partial s} \\ \frac{\partial^2 J_{\text{aux}}}{\partial \rho_2 \partial s} \\ \vdots \\ \frac{\partial^2 J_{\text{aux}}}{\partial \rho_{m-1} \partial s} \\ \frac{\partial^2 J_{\text{aux}}}{\partial \rho_m \partial s} \\ \frac{\partial^2 J_{\text{aux}}}{\partial \beta \partial s} \end{bmatrix}. \tag{18}$$

Finally, plugging equation 18 and 17 back into the model gradient expression 13, we get

$$\frac{\partial J}{\partial s} = -\sum_{z_w}\sum_{\gamma}\sum_{\phi_i} \frac{\partial I(z+z_w,\gamma,\phi_i,x,y;s)}{\partial s}(G_i\tan^2\gamma + G_{m+1})\dot{I}(z+z_w,\gamma,\phi_i,x,y;s_0), \tag{19}$$

in which

$$\begin{bmatrix} G_1 & G_2 & \cdots & G_m & G_{m+1} \end{bmatrix} = \begin{bmatrix} \frac{\partial J_{\text{Sm}}}{\partial \rho_1} & \frac{\partial J_{\text{Sm}}}{\partial \rho_2} & \cdots & \frac{\partial \text{J}_{\text{Sm}}}{\partial \rho_m} \end{bmatrix} \begin{bmatrix} F_{1,1} & F_{1,2} & \cdots & F_{1,m} & F_{1,m+1} \\ F_{2,1} & F_{2,2} & \cdots & F_{2,m} & F_{2,m+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ F_{m,1} & F_{m,2} & \cdots & F_{m,m} & F_{m,m+1} \end{bmatrix}. \tag{20}$$

In practice, there are some caveats when taking the inverse of the Jacobian matrix. The Jacobian can be ill-conditioned when all elements in one row or column are close to zero. For example, if the image point is not illuminated from a certain azimuth direction ϕ_j, i.e. $I(z,\gamma,\phi_j,x,y)=0$, then the j_{th} row and column of the Jacobian would be zero. In order to avoid numerical overflow under this circumstance, we pre-exclude those azimuth angles with poor illumination energy from the Jacobian, and we invert a subset of the original Jacobian that contains only image gathers at well-illuminated azimuth angles.

Gaussian anomaly example

In this example, the model is 2.4 km in x, 2.4 km in y and 1.0 km in z; the grid sampling is 20 m in x and y, 10 m in z. The receivers are fixed during the entire survey, spanning a rectangular area from $(-0.8\,\text{km}, -0.8\,\text{km})$ to $(+0.8\,\text{km}, +0.8\,\text{km}$ with receiver spacing of 20 m in both x and y. The shot locations cover the same region as the receivers do, and the shot spacing is 100 m. This leads to a survey of 81 receivers and 17 shots along both x and y. A total of 32 frequencies are calculated, ranging from 5 Hz to 40 Hz. There is one flat reflector at a depth of 0.6 km. A constant starting velocity of 2 km/s is used. The true model is a constant background velocity (2 km/s, same as the starting velocity), but with a 0.6 km width, 0.25 km height Gaussian anomaly at the center, with a peak value of 3 km/s. Figure 1 shows the true velocity model.

First, we compute the initial subsurface ODCIG (offset-domain common-image gather) with the offset hx, hy range being ± 100 m. The acquisition geometry we use can provide full azimuth for the inner portion of the model. We choose to generate ADCIGs along three azimuth angles: $0°, 30°$ and $60°$.

Figure 2 shows the derivatives of the semblance objective function over the moveout parameter ρ_i, with each panel representing a distinct azimuth: $0, 30$ and 60 degrees. Since the migration velocity is slower than the true one, most of the ADCIGs would be curving upward, therefore a proper objective function should instruct the inversion to decrease the curvature. Figure 2 verifies the consistency between the numerical result and our theoretical expectation. Notice in each plot, there are some positive values on the direction perpendicular to the azimuth orientation. This is caused by the fact that at these locations, the seismic energy traveling through the designated azimuth probes a much smaller anomaly. (Imagine we have a spheric anomaly centered around the origin: $x^2 + y^2 + z^2 = 1$, the strength of the anomaly decreases by distance from the center. If we intercept the anomaly with vertical plane $y = 0$, the anomaly seen on the section is strong because the section passes through the center of the sphere. However, if we intercept the sphere with $y = 0.7$, the anomaly seen on that plane would be much weaker and smaller, because the part of the anomaly seen on this plane is far away from the center.) In this case, the ADCIG at near angles are affected by the anomaly but the far angles are not, therefore showing the opposite curvature. This phenomenon might hinder the tomography algorithm from finding a good update. Fortunately, with more than one azimuth, the curvatures observed from the alternative azimuth would offset this effect. This example shows one of the advantages of having multiple azimuth image gathers.

Figure 3 shows the first model gradient (in velocity) calculated using our approach. Qualitatively, the method works as we expected. The update is concentrated around the anomaly's location and the sign of the update is correct.

3-D transforms between offset and angle domain

Our RMO-based WEMVA approach operate exclusively on subsurface *angle* domain CIGs, however, compared to ADCIGs, subsurface *offset* domain CIGs are much friendlier for computer implementation. The image-space tomographic operator on ODCIGs also has a simpler formula and can be straightforwardly implemented (Tang et al., 2008). Therefore

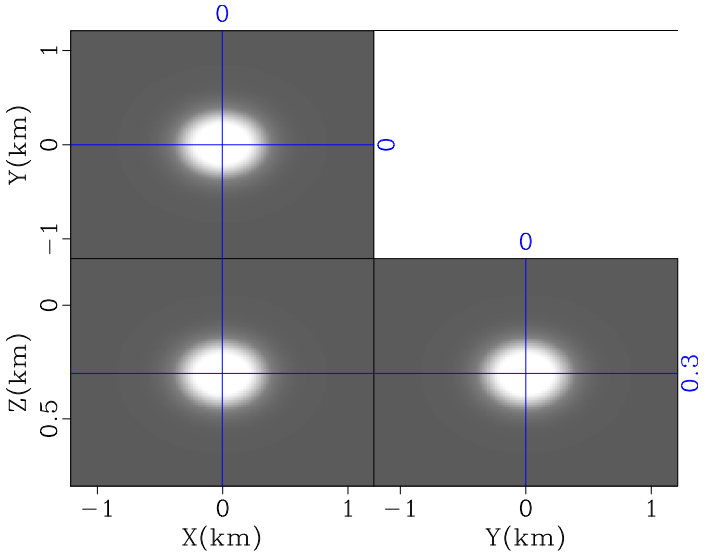
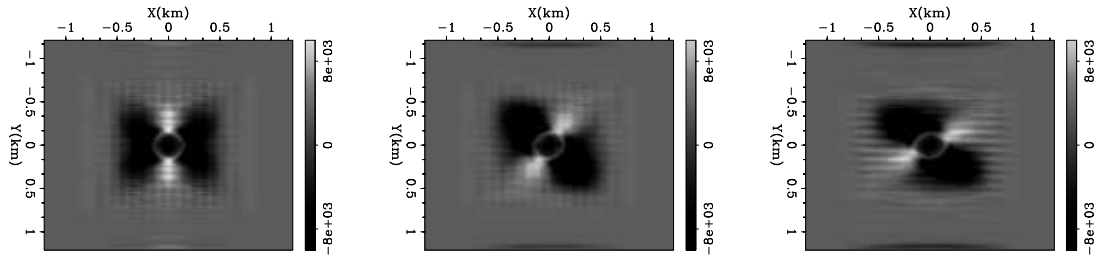


Figure 1: The true velocity model, with positive Gaussian anomaly in the center. [**ER**] yang1/. vel3d



(a) (b) (c)

Figure 2: The $\partial J/\partial \rho_i$ term for (a) $\phi = 0°$, (b) $\phi = 30°$ and (c) $\phi = 60°$. The values correlates well with the curvatures of the ADCIGs at each (x, y) location. [**CR**] yang1/. dJdu0,dJdu30,dJdu60

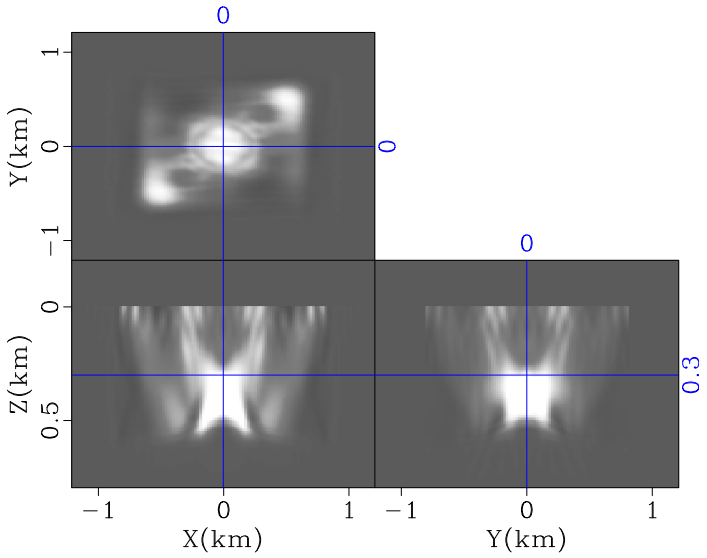


Figure 3: First gradient of the velocity model calculated by our 3-D RMO WEMVA implementation, i.e. Δv predicted by our WEMVA objective function. [**CR**] yang1/. grad-gauss

transforming between ODCIGs and ADCIGs is required for our approach. Tisserant and Biondi (2003) and Biondi and Tisserant (2004) presented the theory for this tranformation. We review the key steps in the offset to angle transformation:

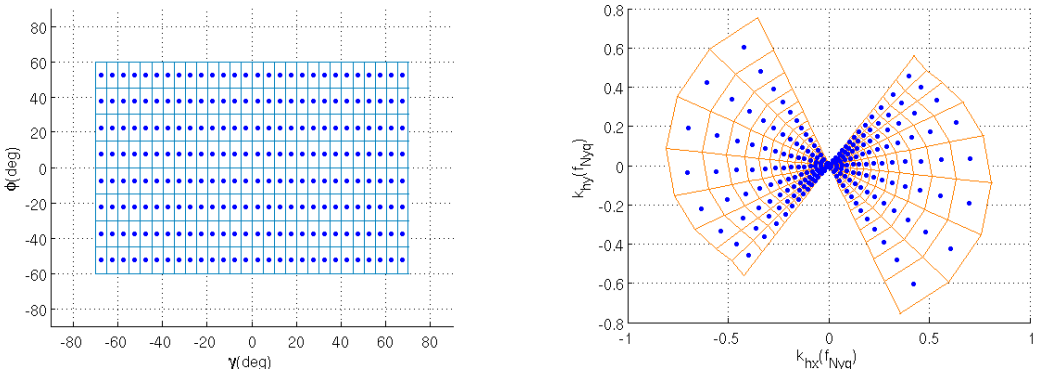
1. Perform Fourier transform $I(hx, hy, x, y, z) \rightarrow I(hx, hy, k_x, k_y, k_z)$.
2. For each (k_x, k_y, k_z),
   - apply Fourier transform $I(hx, hy) \rightarrow I(k_{hx}, k_{hy})$
   - map $I(k_{hx}, k_{hy}) \rightarrow I(\gamma, \phi)$ based on the following relations (Tisserant and Biondi, 2003):

$$\begin{aligned}
\begin{bmatrix} k'_x \\ k'_y \end{bmatrix} &= \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} k_x \\ k_y \end{bmatrix} \\
k'_{hx} &= k_z \sqrt{1 + (k'_y/k_z)^2}\, \tan\gamma \\
k'_{hy} &= \frac{k'_y k'_x k'_{hx}}{k'^2_y + k^2_z} \\
\begin{bmatrix} k_{hy} \\ k_{hy} \end{bmatrix} &= \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} k'_{hx} \\ k'_{hy} \end{bmatrix}.
\end{aligned} \tag{21}$$

3. Apply inverse Fourier transform $I(\gamma, \phi, k_x, k_y, k_z) \rightarrow I(\gamma, \phi, x, y, z)$.

Similarly, the backward transform (angle to offset) is done by reversing the order of the procedures above.

The mapping between (γ, ϕ) and (k_{hx}, k_{hy}) is highly irregular (Biondi, 2003). Figure 4 shows the mapping from a regularly sampled (γ, ϕ) mesh to the (k_{hx}, k_{hy}) domain given



(a) (b)

Figure 4: Graphical illustration of the mapping from the (γ, ϕ) plane to the (k_{hx}, k_{hy}) plane. Each dot in (a) maps to a corresponding dot in (b); similarly, each quadrilateral patch in (a) maps to a corresponding patch in (b). [**CR**] yang1/. a2omapping0,a2omapping1

fixed $k_z = 1/2$, $k_x = 1/4$ and $k_y = 1/4$ (assume the Nyquist wave number is 1). We can see the mapping brings distortion, and the density of the samples on (k_{hx}, k_{hy}) becomes non-uniform. A proper interpolation scheme is very important to reduce the artifacts caused by such irregularity. This issue becomes more serious for the backward transform (angle to offset), because the azimuth angle ϕ is usually not sufficiently sampled; simply placing all available samples on the (γ, ϕ) plane onto their mapped locations on (k_{hx}, k_{hy}) plane would leave many holes unfilled. On the other hand, under mapping relation 21, it is easy to map from a given (γ, ϕ) value to (k_{hx}, k_{hy}), but it is difficult to do the reverse due to the algebra. Therefore, it is very difficult to iterate over all points on (k_{hx}, k_{hy}), find the corresponding (γ, ϕ) coordinates, and fetch the values on these coordinates.

We choose a simple yet effective scheme to perform this mapping. Instead of mapping from sample points to sample points (dots in Figure 4), we map from quadrilateral patches to quadrilateral patches, assuming each patch contains uniformly the value of the sample located in its center. We use a classic polygon-filling algorithm for this mapping. After that, a slight amount of smoothing is applied on the output to remove potential discontinuities along the boundaries of the patches.

We use the previous synthetic example to demonstrate the necessity of this scheme. We start with an initial subsurface ODCIG by migrating the data using the starting velocity model, as shown in Figure 5(c). Then we apply our forward and backward offset-angle transforms to the ODCIG sequentially (Figure 5(b)). For comparison, we also compute the result using the transform that simply does sample-to-sample mapping (Figure 5(a)). A good transform pair should make the resulting ODCIG resemble the original one as closely as possible. Notice there is no way we can retrieve a result exactly same as the original gather, because the information in azimuth range $(60°, 180^{circ})$ is lost in the angle domain CIG. Nonetheless, it is obvious that Figure 5(b) is less distorted than Figure 5(a) is. Our mapping scheme would reduce the artificial noise that arises during the transform of image perturbation from angle back to offset domain.

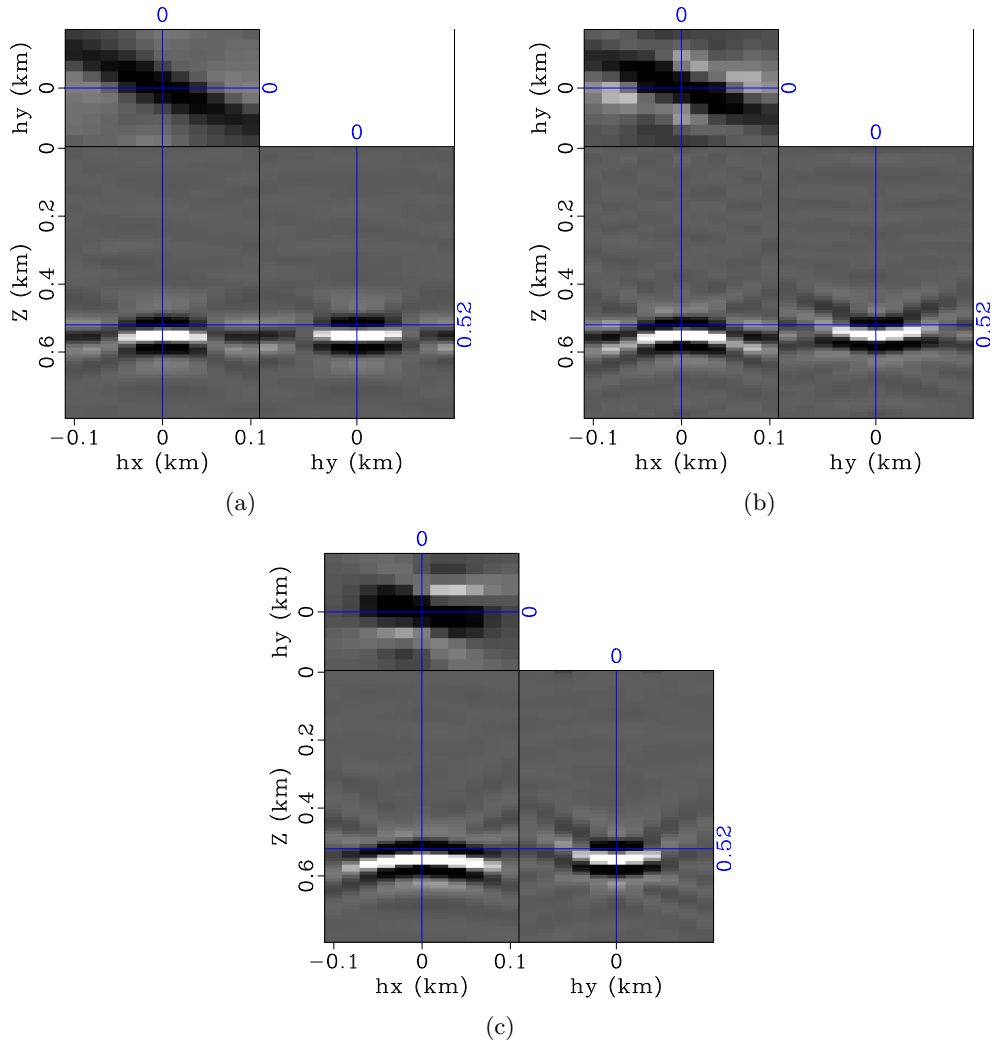


Figure 5: (a) The result of applying forward and backward offset-angle domain transform on an ODCIG, in which the sample-to-sample mapping is used; (b) the same description as (a), except that the patch-to-patch mapping is used. The original ODCIG is shown in (c). All three gathers are chosen at location $x = 0.2\,\text{km}, y = -0.06\,\text{km}$. Notice that the intermediate ADCIG we generate does not have full azimuth coverage ($[60°, 180°]$ excluded), therefore an exactly identical reconstruction of the original ODCIG is not possible. [**CR**] yang1/. offNN,offPF,offori

DISCUSSION

Although it is not a main point in this paper, the computational efficiency is also a crucial challenge for the 3-D extension. As the image gathers become 5-D, the cost of wave-equation operators grows extremely fast as the problem size scales up. Besides choosing a smaller model space, phase-encoded shot records can be used to reduce the number of wave-field simulations. The other bottle neck is the huge amount of disk IO when doing the cross-correlation imaging condition for many subsurface offsets. Based on the recently emerging research topic of compressive sensing, Clapp (2011) proposes a promising idea to reduce the amount of IO required. The author shows that if we *randomly* compute only a subset of the original full 5-D image, for typical seismic data, we can confidently recover the full image without losing useful information.

CONCLUSION

We presented our initial work to extend the proposed RMO-based WEMVA approach to three-dimensional case. We upgraded the theory to deal with multiple azimuth angle-domain CIGs, and we paid close attention to the mapping algorithm between angle and offset domain, in order to reduce artifacts introduced by this irregular mapping. We showed a synthetic example as a verification for our primitive implementation.

ACKNOWLEDGEMENT

We thank Yaxun Tang for the one-way wave-equation propagator used in this project.

REFERENCES

Biondi, B., 2003, Amplitude balancing of 3-D angle-domain common-image gathers: SEP-Report, **114**, 45–56.

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.

Biondi, B. and T. Tisserant, 2004, 3D angle-domain common-image gathers for migration velocity analysis: Geophysical Prospecting, **52**, 575–591.

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

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

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

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

Tisserant, T. and B. Biondi, 2003, Wavefield-continuation angle-domain common-image gathers in 3-D: SEP-Report, **113**, 211–220.

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

Fast velocity model evaluation with synthesized wavefields

Adam Halpert

ABSTRACT

Sophisticated interpretation tools such as automatic image segmentation allow for near-instant creation of multiple, geologically-plausible velocity models for a given imaging target. However, testing each of these models via remigration of the full dataset is not computationally feasible. By using an initial image to synthesize both an areal source function and a Born-modeled dataset, a large number of models can be tested and evaluated in a fraction of the time required for even a single migration of the full dataset. Tests using synthetic 2D data suggest that this method can demonstrate the impact of imaging using different models, even with small differences among them. Furthermore, a measure of image focusing can help quantify the relative merits of the different models.

INTRODUCTION

Because of the high velocity contrast between salt and sediments, accurate interpretation of subsurface salt structures is vital to obtaining a high-quality seismic image. A variety of tools have been developed to aid interpreters in this task, including several different approaches to automatic seismic image segmentation (Lomask et al., 2007; Halpert, 2011). Methods such as these can provide several possible salt body interpretations, whether through uncertainty analysis of the boundary (Lomask, 2007) or interpreter guidance of segmentation results (Halpert et al., 2011). Because of computational constraints, it is often not feasible to run full migrations to test each of the possible models. Therefore, methods to quickly and efficiently test the imaging effects of each model are of great interest.

Several approaches to the problem of fast re-migration of seismic data have been proposed. Many of these approaches rely on a fast variant of beam migration (Hill, 1990), one form of target-oriented imaging. Beam-based methods (Wang et al., 2008) operate in the post-stack domain, and often rely on qualitative judgments of image quality to evaluate different velocity models. Pre-stack imaging methods using reverse-time migration (RTM) have also been proposed (Wang et al., 2011), and would allow for analysis of valuable velocity information in the form of subsurface-offset or angle gathers. However, at present this approach is likely only feasible when testing a small number of limited-size models. A further option, and the one explored here, is to synthesize new datasets from an initial image, and use these datasets to test velocity models. First, a new source wavefield can be derived from the initial image using modified prestack exploding reflector modeling (Guerra, 2010). This allows information from the inevitable velocity errors in the initial model to be preserved and, hopefully, corrected. Second, a form of Born modeling (Stolt and Benson, 1986) can be used to synthesize a new dataset from the initial image (Tang and Biondi, 2010; Tang, 2011). The new dataset can be designed with an arbitrary acquisition geometry, creating opportunities for target-oriented imaging. Furthermore, since a generalized

source function is used, a useful image with prestack velocity information can be obtained by migrating only a single shot. Crosstalk artifacts can be avoided by targeting only small, isolated locations from the initial image. This approach also allows for a simple quantitative measure of image quality or focusing. In the following sections, this method will be described and demonstrated using data from the 2D Sigsbee synthetic model. The accuracy and computational efficiency of the results show great promise for this method working in conjuction with automatic segmentation as part of a full 3D interactive interpretation toolbox.

METHOD

The procedure for Born modeling and migration to quickly test velocity models is detailed in Halpert and Tang (2011). To summarize, the major steps of the process are:

1. Using one or more subsurface offset gathers from the initial prestack image, create an areal source function via upward continuation. This is similar to the "exploding reflector" concept (Claerbout, 2005), but inclusion of the prestack subsurface-offset data is important for attempts to improve upon the initial velocity model. Mathematically, this areal source is described 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 $\mathbf{x}_s = (x_s, y_s, z_s)$ are the arbitrarily defined locations where the wavefield will be recorded; $\mathbf{h}$ is the vector of subsurface half-offsets; ω is angular frequency; m is the initial image from which data is injected at isolated regions defined by $\mathbf{x}'$; and G is the Green's function propagating the wavefield to the receiver locations (here, * denotes the adjoint). Because the wavefield can be recorded at arbitrary locations, this method provides a simple means of re-datuming, which can lead to significant computational savings.

2. Use the new source function from step 1 and a reflectivity model based on the initial image to generate a new dataset with acquisition geometry best suited to image the target area. This can be accomplished via Born modeling, if we synthesize the new dataset d' and record 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 because the initial model will inevitably contain errors, and the goal of the procedure described here is to improve upon it.

3. Using the synthesized data obtained in Step 2, and the source function from Step 1, generate an image using standard downward-continuation migration:

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

   Now, *any* velocity model can be used to compute the Green's functions. This step is extremely computationally efficient compared to a full migration of the original data, allowing for testing of several possible velocity models in a fraction of the time it would take to evaluate them using standard migration techniques.

With the procedure outlined above, significant savings may be realized if only a single shot is migrated, a situation made possible by the fact that an areal source function is used along with an arbitrary acquisition geometry. However, in many cases an image obtained in this manner is contaminated by crosstalk artifacts. Various solutions to the crosstalk attenuation problem have been proposed, for example using multiple phase-encoded shots (Romero et al., 2000; Tang, 2008). Unfortunately, this approach would hinder the computational efficiency of the evaluation procedure – one of its primary considerations. Instead, the procedure described above is applied only to isolated locations from a single reflector in the initial image. As long as these locations are separated by at least twice the maximum subsurface offset value used to synthesize the new source and receiver wavefields, any crosstalk problems should be avoided. This method also allows an interpreter or model-builder to select reflectors he or she thinks would be most sensitive to changes in the velocity model – for example, the base of salt.

Evaluation criteria

Since the goal of the steps outlined above is to test the accuracy of several different models, there must be a way to judge or evaluate the images corresponding to each. In some cases, a qualitative judgment may suffice; for example, if one model clearly focuses the data at zero subsurface-offset. However, a quantitative measure of image quality would also be useful. Since only isolated locations are being imaged, we can expect a "perfect" velocity model to focus all an image's energy at zero subsurface-offset. Therefore, a simple measure of image quality calculates what proportion of the energy indeed resides at zero or near-zero subsurface-offset:

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

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. This idea is similar to the motivation behind some inversion schemes such as differential semblance optimization (Symes and Carazzone, 1991). 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.

Migration model	F value
True velocity	0.883
Extra salt	0.879
Slow salt	0.864

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

SYNTHETIC EXAMPLE

The above procedure will now be demonstrated using two different initial images derived from the Sigsbee synthetic model. Figure 1(a) is a perfect-velocity, full migration of the Sigsbee data, which will be used for the first example. Figure 1(b) shows a manually-picked reflector chosen for further analysis; in this case, the base salt has been chosen because it should be particularly sensitive to different interpretations of the salt body's shape and velocity, the two model variations that will be tested. Finally, Figure 2 shows two image locations isolated from the selected reflector. Note that most of the energy is focused near zero subsurface-offset, since the true velocity model was used for the initial image.

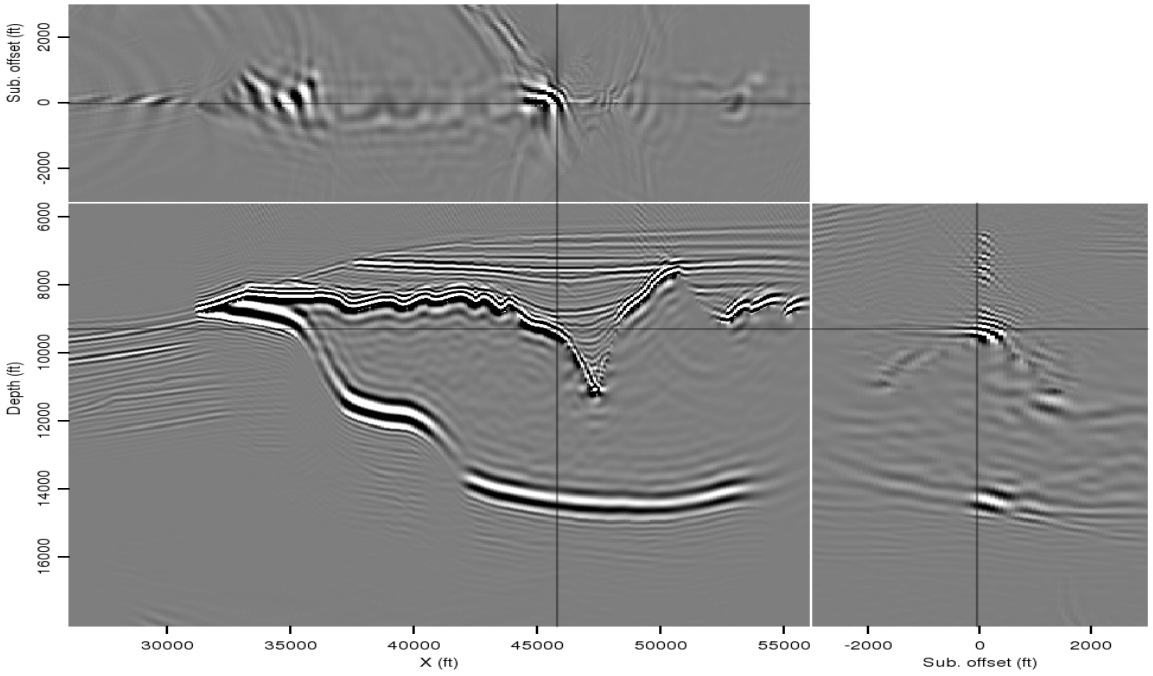
After synthesizing the source and receiver wavefields as described in the previous section, the new wavefields were imaged using three different velocity models. First, the true model, seen in Figure 3(a). Second, an alternative model created via automatic image segmentation, in which an interpreter has chosen to include an extra chunk of salt (Figure 3(b)). The third model tested was identical to the true model in salt shape, but with a salt velocity 5% slower than the true model.

Resulting images from the Born-modeled data are seen in Figure 4. Panel 4(b), the result of migrating with the extra-salt model in Figure 3(b), is clearly the least well-focused image. However, it is difficult to qualitatively distinguish between the other two results. Table 1 gives the results of calculating the F value from equation 5 for each of the images. As expected, the result using the true model has a higher F value, indicating it is more well-focused.

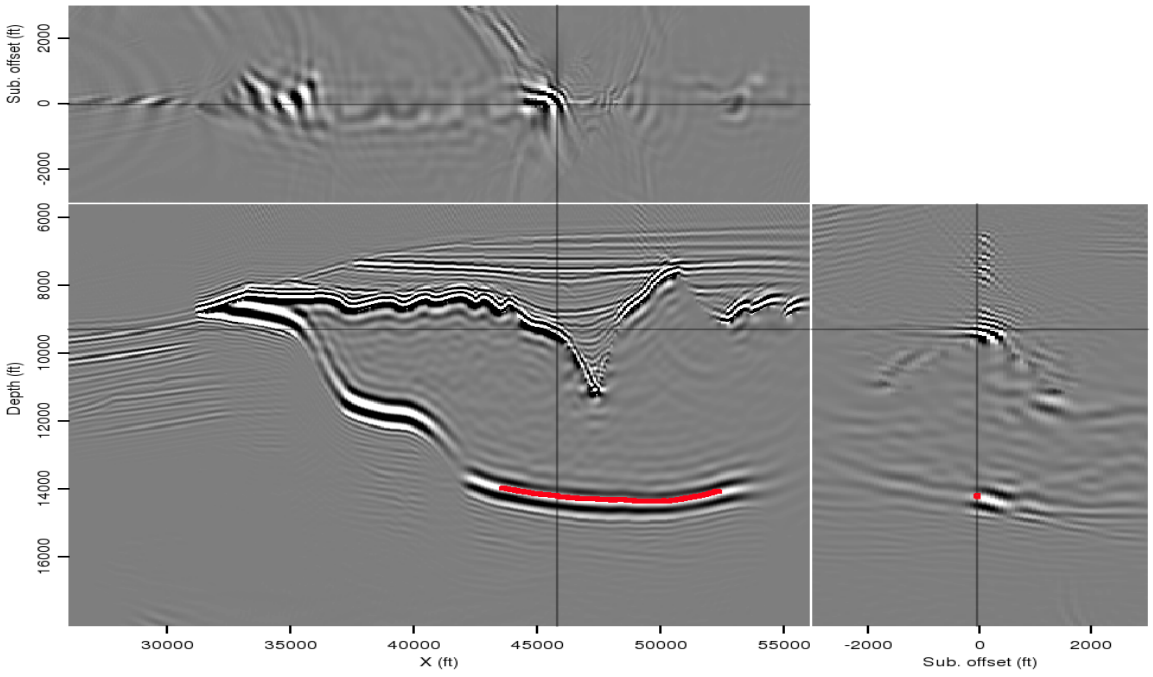
The second example uses an initial image created with an incorrect velocity model; in this case, the "slow salt" model described above. The results corresponding to each velocity model can be seen in Figure 5. Again, the extra-salt model is clearly inferior, but the differences between the other two results are more subtle. The F-value calculations in Table 2 confirm that the true model yields the optimal result, even though the slow-salt model was used to create the initial image and both synthesized datasets. In both examples, the differences between the calculated F values are relatively small; further investigation should determine how confidently such differences may be interpreted, especially when working with noisier field data.

CONCLUSIONS

When several different velocity models are options for a given dataset, they can be quickly tested using wavefields synthesized from an initial image. Incorporating prestack velocity



(a)



(b)

Figure 1: (a) A true velocity image using data from a section of the Sigsbee synthetic model; and (b) a base-of-salt reflector selected for further analysis because of its sensitivity to changes in the salt interpretation. [**CR**] adam2/. img-act,picks

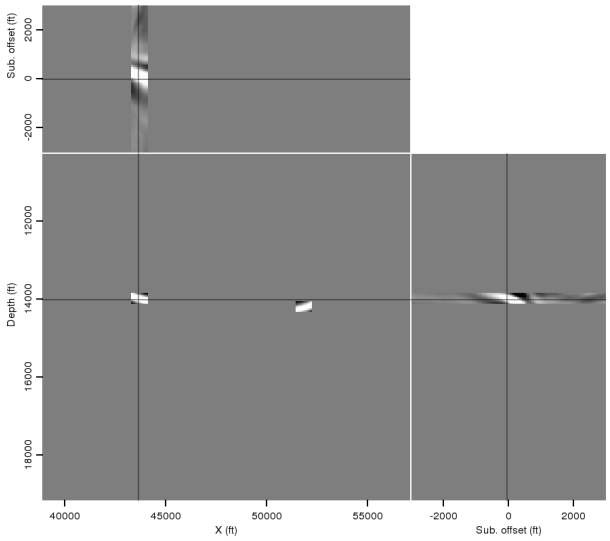


Figure 2: Isolated image locations selected from the reflector picked in Figure 1(b). [**ER**] adam2/. act-pts

Migration model	F value
True velocity	0.621
Extra salt	0.561
Slow salt	0.613

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

information from the initial image into both an areal source function and a Born-modeled receiver wavefield allows errors in the initial model to be taken into account and rectified when a more accurate model is tested. Finally, by imaging only isolated locations along a reflector, tests can be conducted in a fraction of the time required for a full migration by migrating only a single shot. Tests on data from the Sigsbee synthetic model demonstrated that this method can effectively discriminate between models, even if they differ only slightly in terms of salt shape or velocity, or if the initial image was created with an inaccurate model. Finally, a quantitative measure of image quality helped identify the most accurate model when qualitative comparisons were inconclusive.

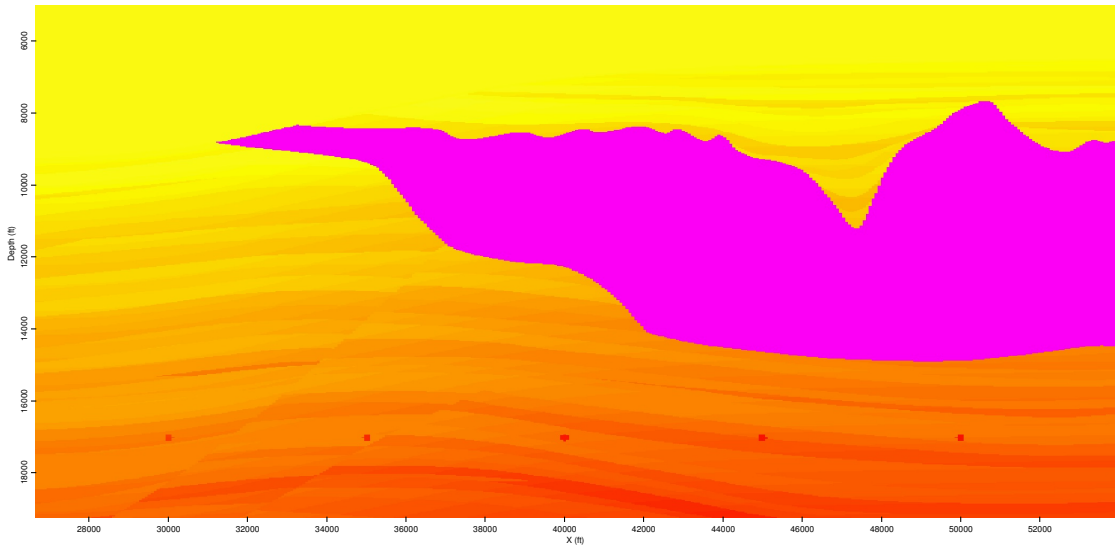
ACKNOWLEDGMENTS

I thank SMAART JV for providing the Sigsbee synthetic model used for examples, and Yaxun Tang for his work on the Born modeling/migration framework used extensively here.

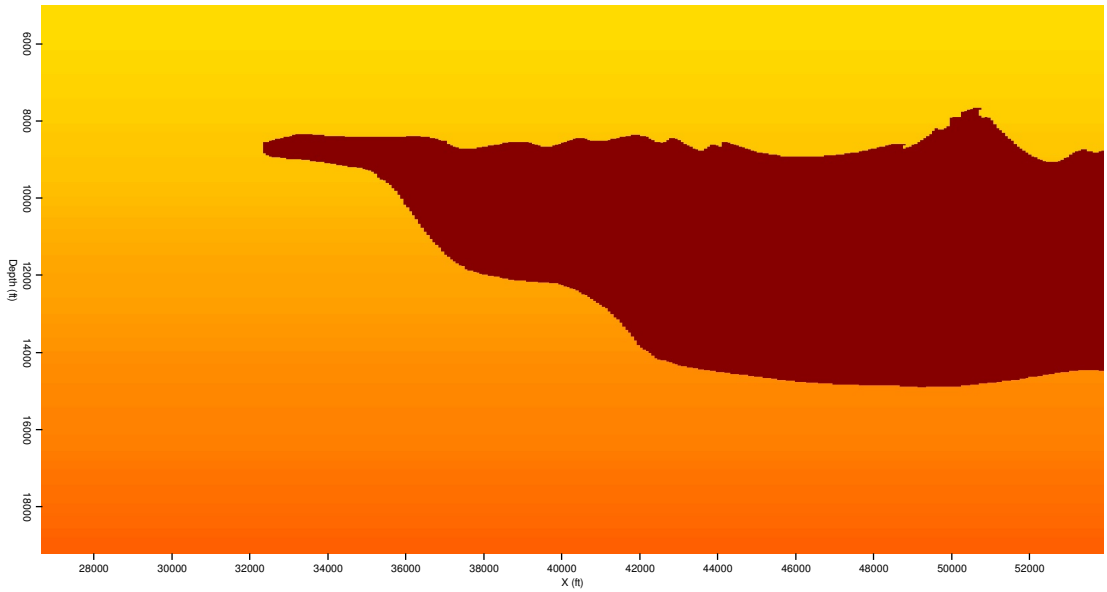
REFERENCES

Claerbout, J., 2005, Basic earth imaging: Stanford University.

Guerra, C., 2010, Migration-velocity analysis using image-space generalized wavefields: PhD thesis, Stanford University.

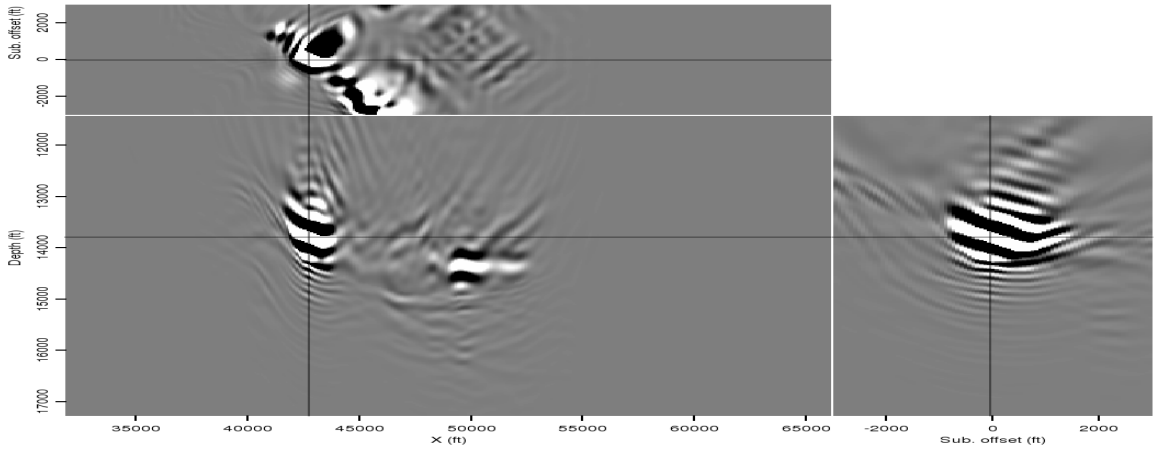


(a)

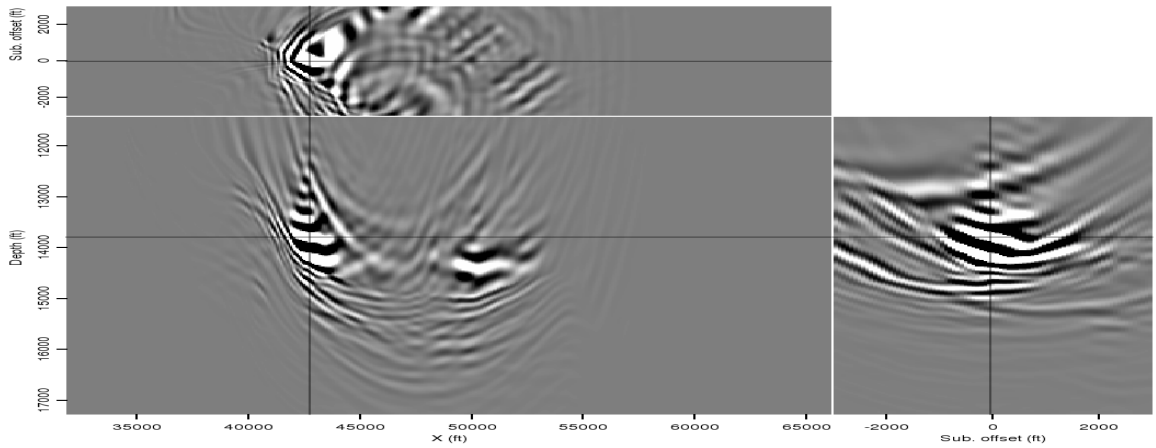


(b)

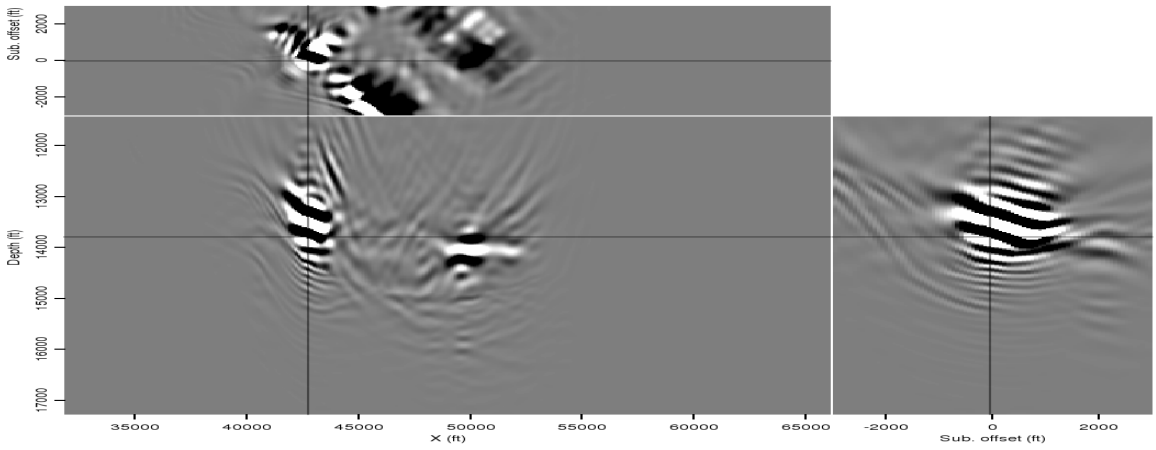
Figure 3: Two different velocity models to be tested. The model in (a) is the true Sigsbee model, while (b) is an alternate model created by one possible automatic segmentation of the initial image. [**ER**] adam2/. zvel-act,zvel-filled



(a)

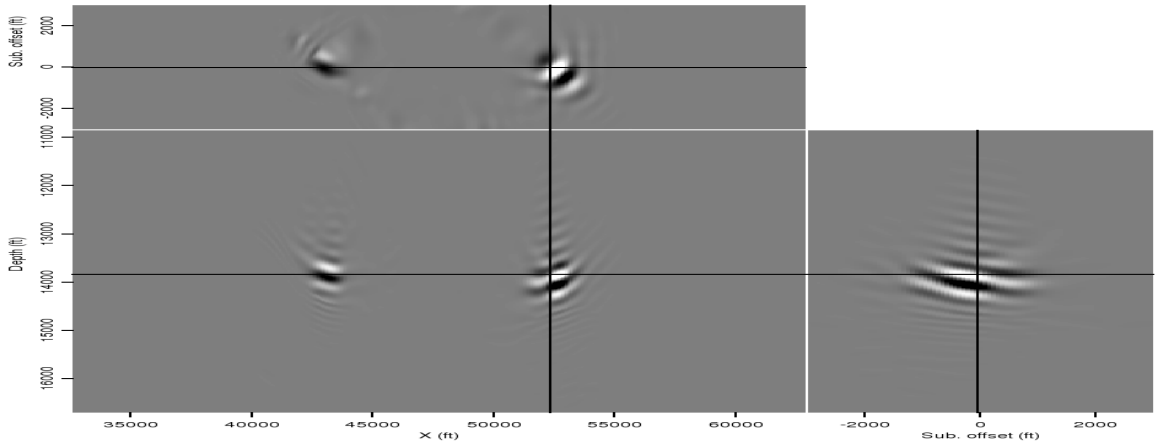


(b)

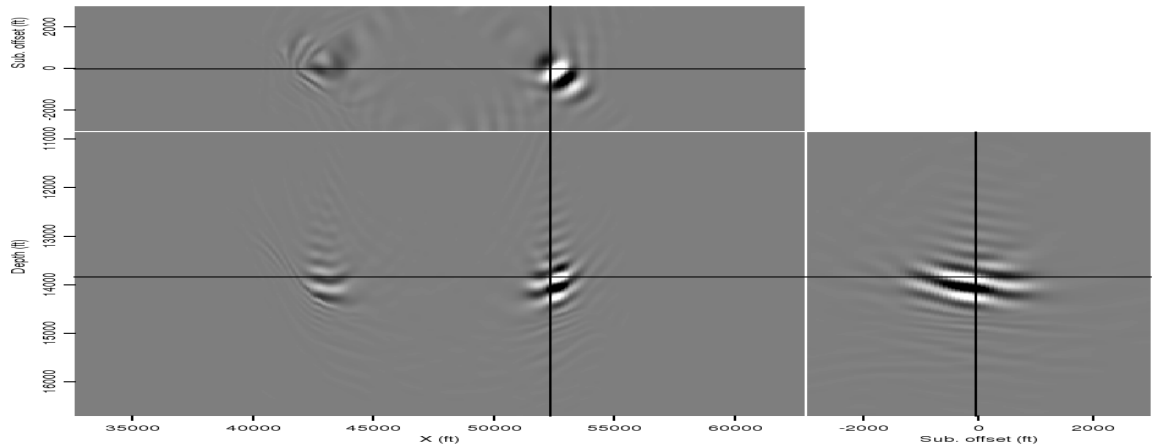


(c)

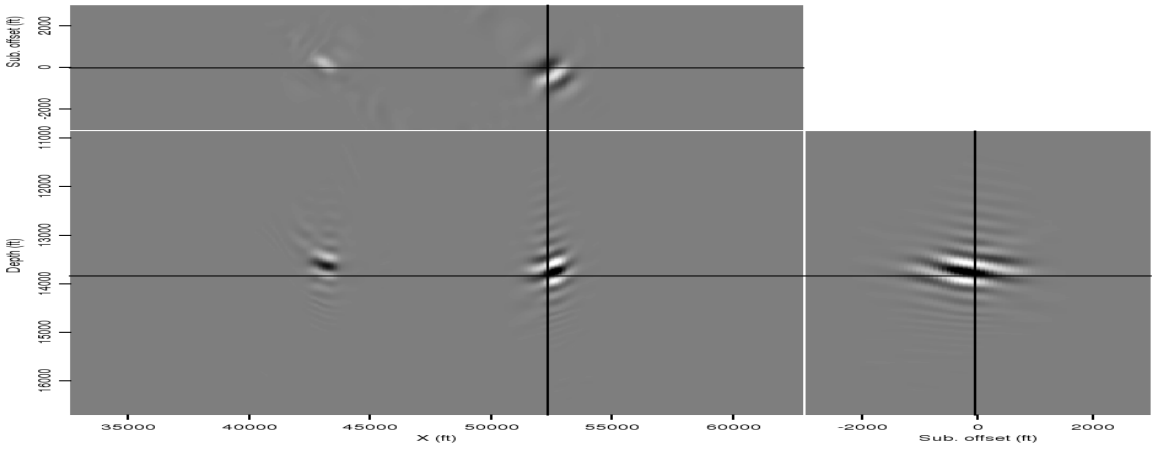
Figure 4: Three images of the Born-modeled data using three different migration velocity models: (a) the true model; (b) the extra-salt model; and (c) the slow-salt model. In this example, the initial image was created with the true model. [**ER**] adam2/. born-act,born-af,born-as



(a)



(b)



(c)

Figure 5: Three images of the Born-modeled data using three different migration velocity models: (a) the true model; (b) the extra-salt model; and (c) the slow-salt model. In this example, the initial image was created with the slow-salt model. [**ER**] adam2/. born-sa,born-sf,born-ss

Halpert, A., 2011, Interpreter input for seismic image segmentation: SEP-Report, **143**, 237–248.

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

Halpert, A. D., R. G. Clapp, and B. L. Biondi, 2011, Interpreter guidance for automated seismic image segmentation: 74th EAGE Conference and Exhibition.

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

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

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

Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, 426–436.

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

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

Tang, Y., 2008, Modeling, migration, and inversion in the generalized source and receiver domain: SEP-Report, **136**, 97–112.

———, 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.

Early-arrival waveform inversion: Application to cross-well field data

Xukai Shen, Tieyuan Zhu, and Jerry M. Harris

ABSTRACT

We apply early-arrival waveform inversion to a cross-well field data in order to find the extent of a potential reservoir that terminate somewhere between the two wells. Ray-based tomography result is too smooth to resolve the boundary. On the other hand, waveform inversion is able to better utilize the advantage of short wavelength cross-well data, resulting in a velocity model that defines the potential reservoir boundary more sharply. Details from the waveform inversion result are verified by well-log data.

INTRODUCTION

Cross-well data is usually of higher frequency, and therefore shorter wavelength, than surface seismic data. Theoretically, this should enable us to reveal more details of the subsurface structure by performing tomography using cross-well data. However, if ray-based methods (Hampson and Russell, 1984; Olson, 1984; White, 1989) are used, the tomography results do not contain much detail, because ray-based methods are based on the high-frequency approximation, which mandates that the scale of subsurface structures be much larger than the dominant wavelength of the data. Such an approximation results in very smooth velocity models. Waveform inversion (Tarantola, 1984; Pratt et al., 1998; Mora, 1987), on the other hand, does not suffer from such approximation, and is capable of producing a much higher resolution result.

The two wells used to acquire this dataset are located very close together, only 640 feet apart. One well penetrates an anomalous zone, but the other one missed the zone. The anomalous zone contains massive reservoir-quality carbonate, as suggested by the first well. However, surface seismic data in the same area has insufficient resolution to determine the vertical and horizontal extent of the zone. To more accurately identify the boundary of the zone, a cross-well survey was conducted between these two wells.

The cross-well data were collected with a piezoelectric cylindrical bender source and hydrophone receivers in the depth range of 8200-9200 ft (2500-2800 m). A total of 201 source levels were spaced at 5 ft (1.5 m) depth intervals, and 203 receiver levels were also spaced 5 ft apart. This resulted in a raw data set with about 40000 traces. The input signal was a 250-1250 Hz linear upsweep (Langan et al., 1997). Details of the acquisition system and the techniques employed can be found in (Harris et al., 1995). The data retrieved have a relatively good signal-to-noise ratio. The only strong coherent noise is the tube wave that exists in all shot records (Figure 1), which is caused by potential leakage in the receiver well.

Previous result from ray-based methods (Langan et al., 1997) is very smooth, and did not identify the boundary of the zone very well. To better define the boundary, we apply

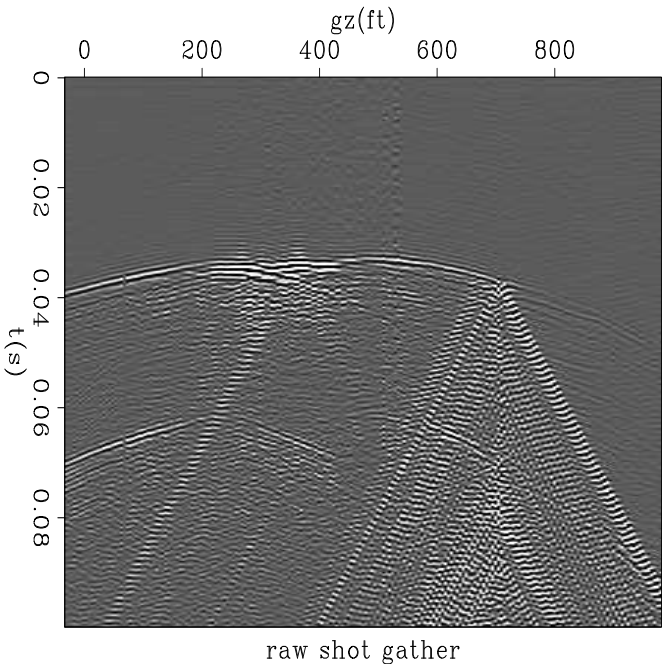


Figure 1: Typical shot gather from the field cross-well dataset. Notice the strong tube wave in the right part of the shot gather. [**NR**] xukai4/. shot

early-arrival waveform inversion (Tarantola, 1984; Pratt et al., 1998; Mora, 1987) to the dataset.

INVERSION SETUP

In real-data application of waveform inversion, it is usually more reliable to match kinematics (traveltime) than dynamics (amplitude), since the acoustic wave-equation can not predict field data amplitude very well. The waveform inversion objective function used here is from Shen (2010). Since cross-well data usually contain converted waves, and we want to estimate p-wave velocity only, we choose to match direct arrivals only, which are free of converted energy and can usually be described accurately by the acoustic wave equation. We create a data mask to keep only the direct arrivals in the input data. We also make sure that tube wave is not included in data matching. An example of the data mask applied to a low-passed shot gather is in Figure 2.

Next we need to determine the proper source wavelet for forward modeling during inversion. Direct arrivals at the receiver side acurately represent what the source wavelet looks like. However, we are using two-dimensional modeling, whereas the field data source is a three-dimensional point source. We apply a phase shift to the direct arrival to correct for that. To improve the signal-to-noise ratio of the source wavelet, direct arrivals are first aligned by using picked first-break traveltime, then stacked to form the reference source wavelet. In this way, each shot has its own source wavelet. The original source and receiver coordinate has a 1-ft interval increment. However, since our modeling grid has a 5-ft interval in both the x and z directions, the source and receiver coordinates are rounded to a 5-ft grid by nearest-neighbor criteria.

The next element for waveform inversion is the starting model, which we obtain by using ray-based traveltime tomography. The tomography adopts the L2-regularized least-

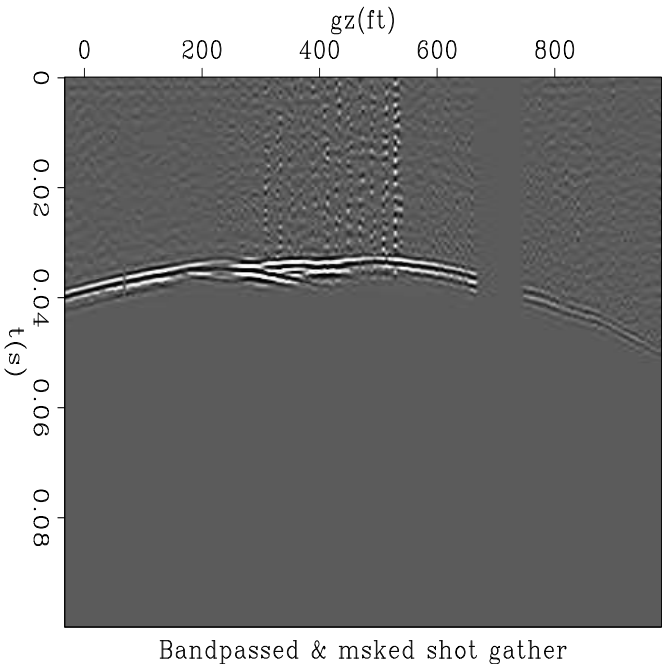


Figure 2: Same shot gather after bandpass and applying data mask. Data mask only keeps direct arrival for inversion purposes. [**NR**] xukai4/. shotlpmsk

square approach to minimize the difference between observed and calculated traveltimes. We use the finite-difference Eikonal solver to compute the traveltime and then we back-project the raypaths from a receiver to the source (Zelt and Barton, 1998). We solve the objective equation by a Gauss-Newton strategy; for details see (Zhu and Harris, 2011). The observed data are the picked first break times from the full data set. The 2D inverse domain size is 133 by 206 cells, with a total number of 27,398 unknowns. The initial model for traveltime tomography has a constant velocity which is the average velocity value obtained from velocity logs. We ran ten iterations, stopping when the traveltime residuals were no longer significantly decreased. The final root-mean-square (RMS) residuals were reduced to 7.6e-3 ms from the initial value of 1.9 ms. Figure 3 shows the traveltime tomography result.

Comparison of data modeled from initial velocity and input data is shown in Figure 4. The finite difference modeling verified that the initial model predicts the data kinematics quite well in general. Only when traveltime is not continuous does the initial model have trouble producing a good match. Since the initial traveltime tomography result already matches picked first-break traveltime quite well, so the waveform inversion result mainly adds details to the traveltime tomography result.

RESULT ANALYSIS

A total of 180 shots are used. Two passes of inversion are run. First with data bandpassed between 200 Hz and 400 Hz, then with data bandpassed between 200 Hz and 700 Hz. The final inversion result is shown in Figure 5, with well location indicated by the two slanted lines, and the well velocity value outside the red slanted lines. The inversion result matches well-log data reasonably well, and it better defines the extent of the anomaly, which is also consistent with well-log data. Compared to the tomography result, there are more details in

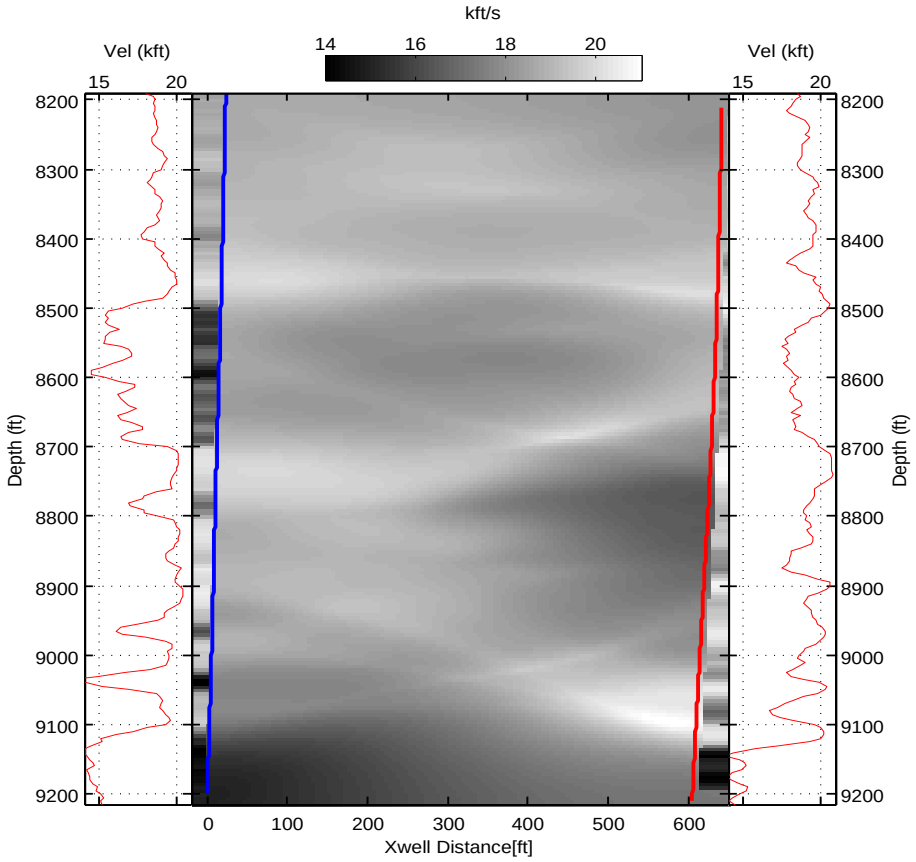


Figure 3: Starting model from traveltime tomography. Left line is receiver well. Right line is source well. Outside wells are well-log velocity values. [**NR**] xukai4/. initmod

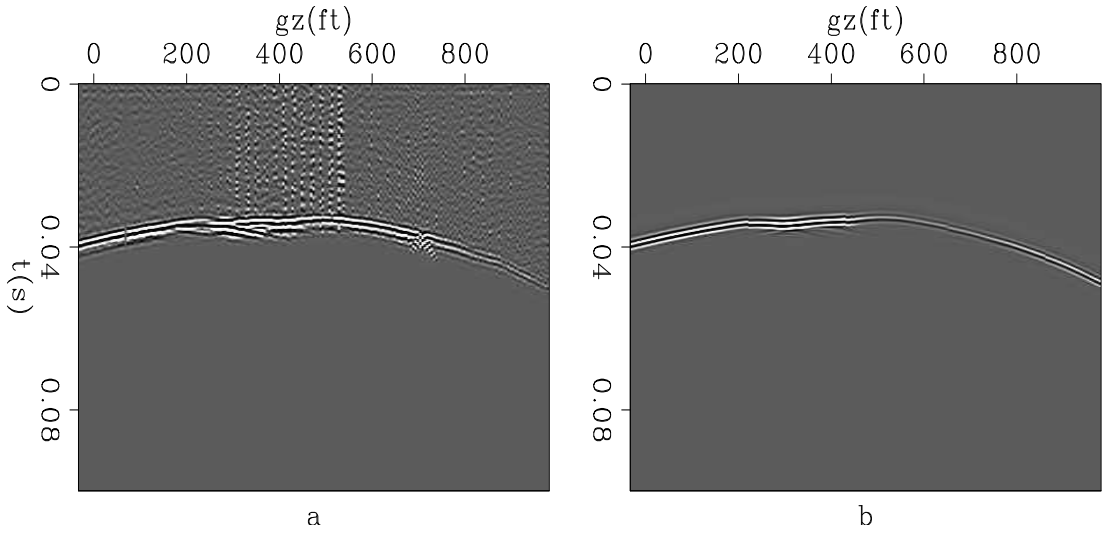


Figure 4: Same shot gather as before, comparsion of input first arrival and data modeled from the tomography result. The two are already very similar in terms of kinematics. a): Input first arrival; b): first arrival modeled from the tomography result. [**NR**] xukai4/. initdatacomp

the waveform inversion results. There are some thin layers that are present in the well-log and waveform inversion results, yet are missing in the traveltime tomography result.

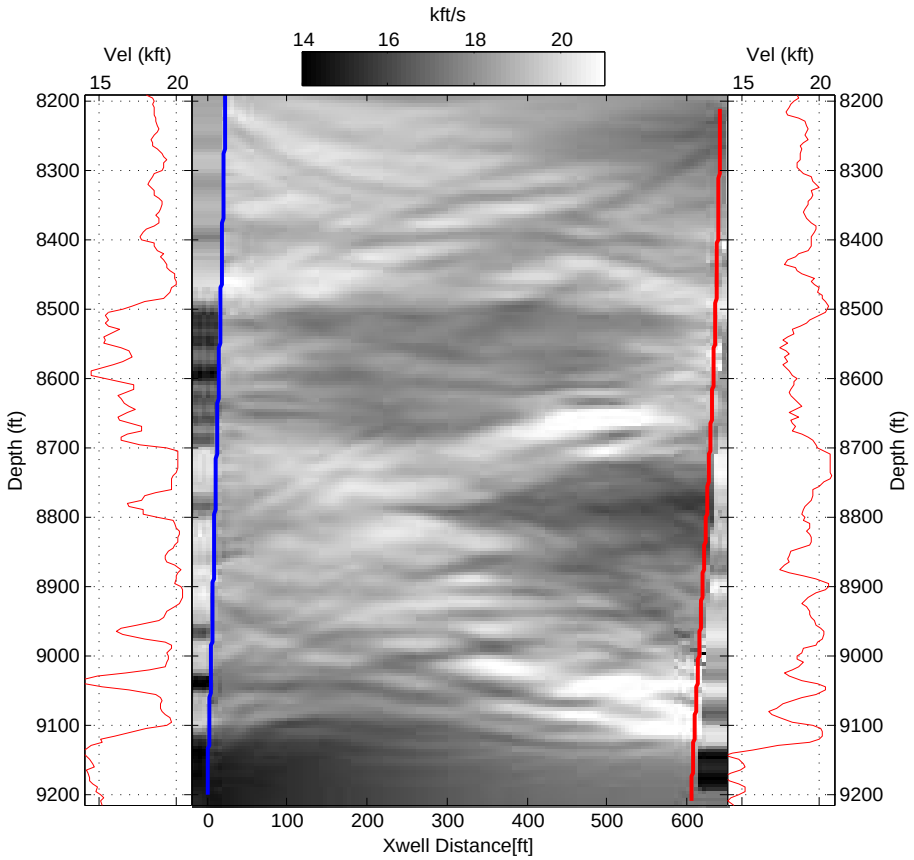


Figure 5: Final model after waveform inversion. Left line is receiver well. Right line is source well. Outside wells are well-log velocity values. [**NR**] xukai4/. invmod

RMS data residual as a function of iteration numbers is shown in Figure 6. Since the initial model is already quite close to the true model, the RMS data residual did not reduce as drastically as expected. A typical shot-gather residual is shown in Figure 7 and 8. It can be seen that data residuals have been reduced mostly in the center part, where rays do not describe the kinematics well.

Comparing the forward-modeled first arrival from the initial velocity, the forward-modeled first arrival from the final velocity, and the input first arrival (Fig 4 and 9), the major improvement comes from kinematics matching. This is not surprising, given the inversion objective function we use. The first arrival modeled from initial velocity is continuous, even where the input first arrival is not. Such discontinuity is captured in the first arrival modeled from the final model. Since such discontinuity can be explained better by wave phenomena, ray-based inversion methods can not accurately handle it. Another advantage of waveform inversion is its capability of modeling more than first arrivals. Data modeled from the final velocity predict some later arrivals which do not show up in the data modeled from the initial velocity.

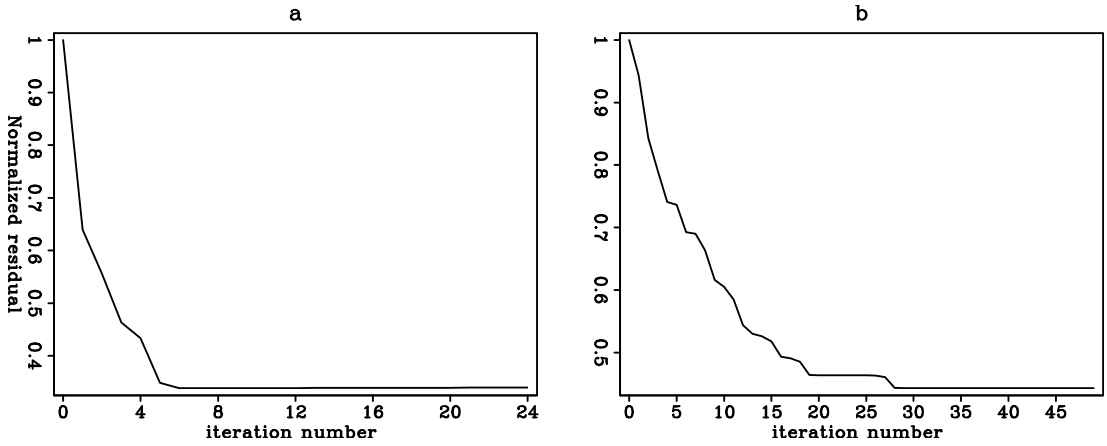


Figure 6: Data residual of the two inversion steps. a): Normalized residual of 200-400 Hz inversion. b): Normalized residual of 200-700Hz inversion. [**NR**] xukai4/. resd

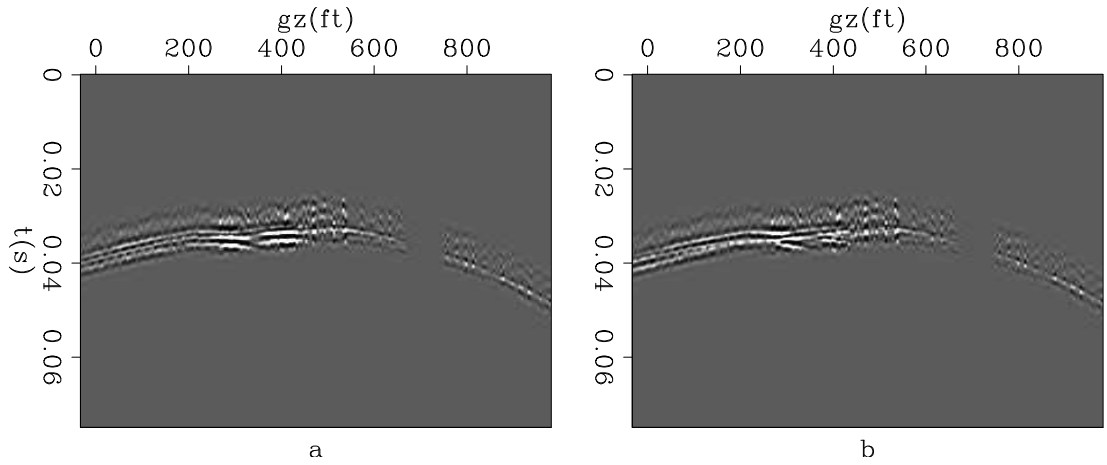


Figure 7: Shot residual of the first inversion steps. a): initial residual. b): final residual. [**NR**] xukai4/. rdfulf

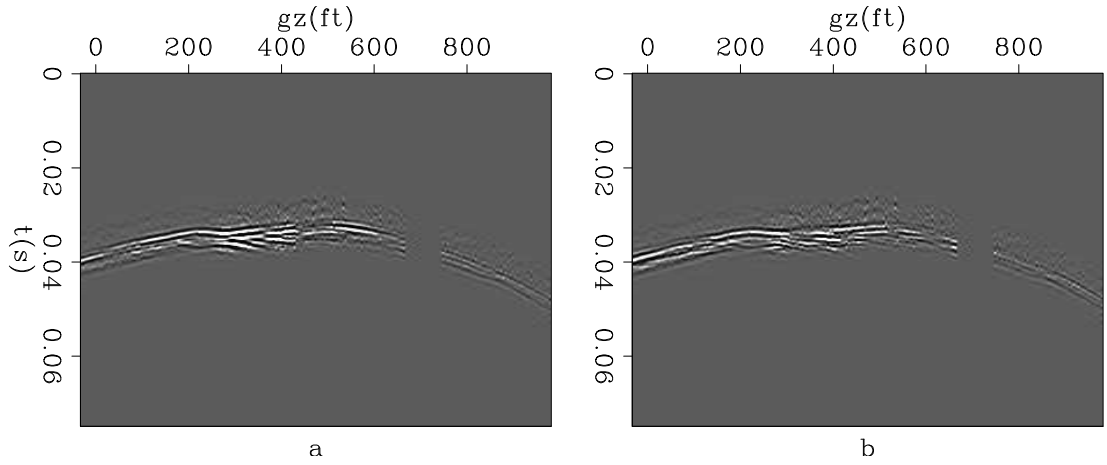


Figure 8: Shot residual of the second inversion steps. a): initial residual. b): final residual. [**NR**] xukai4/. rdful

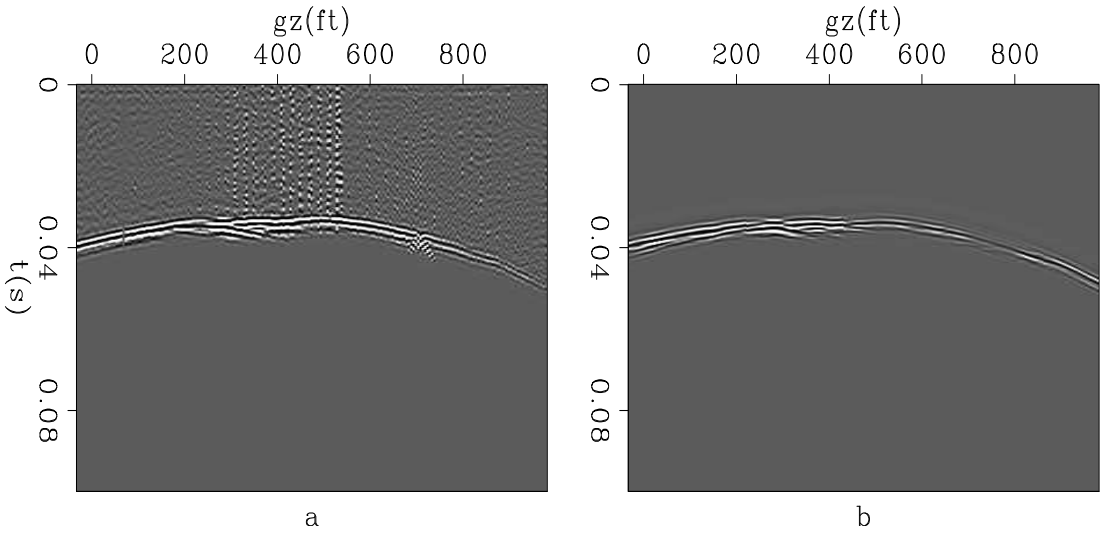


Figure 9: Same shot gather as before, comparsion of input first arrival and data modeled from the final inversion result. First arrival modeled from final result matches input kinematics very well. Left: Input first arrival; right: first arrival modeled from the final result. [**NR**] xukai4/. datcomp

CONCLUSIONS

A cross-well dataset was acquired to more accurately identify the boundary of an anomalous zone between the two wells. We applied early-arrival waveform inversion to the dataset. By trying to match recorded first arrivals, waveform inversion produced a much higher-resolution result than did traveltime tomography results. The boundary of the anomalous zone is more sharply defined. Details from the waveform inversion result is verified by the well-log data. Comparison of forward-modeled data with input data also suggests that wave-equation based methods predict data better than ray-based methods.

ACKNOWLEDGMENTS

We would like to thank Chevron for allowing us to publish the research results. We would also like to thank the Center for Computational Earth Sciences for providing the computation environment.

REFERENCES

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

Harris, J. M., R. C. Nolen-Hoeksema, R. T. Langan, M. V. Schaack, S. K. Lazaratos, and J. W. Rector, 1995, High-resolution crosswell imaging of a west Texas carbonate reservoir: Part i-project summary and interpretation: Geophysics, **20**, 667–681.

Langan, R. T., A. A. Vassiliou, S. K. Lazaratos, T. L. Jensen, J. M. Harris, and J. W. Fair-

born, 1997, Carbonate seismology: Chapter 17, Imaging of a Stratigraphically Complex Carbonate Reservoir with Crosswell Seismic Data.

Mora, P., 1987, Elastic wavefield inversion: Stanford Exploration Project Ph.D. Thesis.

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

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

Shen, X., 2010, Near-surface velocity estimation by weighted early-arrival waveform inversion: SEG Expanded Abstracts, **29**, 1975–1979.

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

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

Zelt, C. A. and P. J. Barton, 1998, 3D seismic refraction tomography: A comparison of two methods applied to data from the Faeroe Basin: Journal of Geophysical Research, **103**, 7187–7210.

Zhu, T. and J. M. Harris, 2011, Iterative joint inversion of P-wave and S-wave crosswell traveltime data: SEG Expanded Abstracts, **30**, 479–483.

VTI migration velocity analysis using RTM

Yunyue (Elita) Li, Peng Shen and Colin Perkins

ABSTRACT

We use a vertical transverse isotropic (VTI) two-way propagation engine to perform wave-equation migration velocity analysis. Using reverse time migration (RTM), we have a chance to characterize the wavefields propagating at large angles and to image steeply dipping reflectors. By including VTI Thomsen parameters, we can better describe the properties of the subsurface. We first derive the migration velocity analysis gradient when using a first-order VTI two-way wave-equation. Then, we test our method on a synthetic VTI Marmousi model. The inversion results show that our method can resolve a better velocity model and a better-focused subsurface image.

INTRODUCTION

Velocity model building has been one of the most challenging problems in the seismic exploration industry. Wave-Equation Migration Velocity Analysis (WEMVA) has been widely studied for velocity building and can be implemented either in the data space (Tarantola, 1984; Woodward, 1992) or in the image space (Sava and Biondi, 2004a,b; Shen, 2004; Shen and Symes, 2008; Guerra et al., 2009). Several advantages drive us to use the image-space WEMVA instead of data-space WEMVA (which is also known as Full-Waveform Inversion): 1), migrated image is often much cleaner than the recorded wavefields; 2), the objective function is directly related to the final image; 3), image-space WEMVA can use a less accurate initial solution without encountering the cycle-skipping problems that can plague FWI. In fact, the optimized output of the image-space WEMVA can be used as the input for FWI (Li and Biondi, 2011).

Since first reported in exploration seismology in the 1930s (McCollum and Snell, 1932), anisotropy has become increasingly important in seismic imaging and exploration. The increasing offset and azimuth in data acquisition has heightened the need for anisotropic imaging and model building. Until now, the transverse isotropic (TI) model has been one most commonly used in seismic imaging and has been considered a better description of the subsurface. Li and Biondi (2011) extend the WEMVA framework to VTI media using the one-way wave-equation. However, the one-way wave-equation cannot accurately describe the wave propagation at large angles with respect to vertical, where anisotropy has larger effects. There have been extensive studies on anisotropic RTM with increasingly complex subsurface models (Fletcher et al., 2009; Zhang and Zhang, 2009), but reliable anisotropic model-building techniques are still needed.

Therefore, we propose an image-space WEMVA method using a VTI two-way wave-equation as the propagation engine, and evaluate the flatness of the RTM images in the angle domain. In this paper, we first derive the gradient of the differential semblance optimization (DSO) objective function with respect to velocity and ϵ using a Lagrangian

augmented functional. To resolve the ambiguity between velocity and ϵ and ensure that our model honors the geology, we use a preconditioning inversion scheme. Finally, we test the proposed method on a synthetic VTI Marmousi model.

FIRST-ORDER TWO-WAY VTI WAVE-EQUATION

The first-order two-way VTI wave-equation can be derived from Hooke's law and Newton's law using Thomson anisotropy parameters (ϵ, δ) and setting shear wave velocity $c_s = 0$ (Duveneck et al., 2008). The first-order system reads as follows:

$$\begin{aligned}
\rho\partial_t v_x &= -\partial_x p_H \\
\rho\partial_t v_y &= -\partial_y p_H \\
\rho\partial_t v_z &= -\partial_z p_V \\
\frac{1}{\rho c^2}\partial_t p_V &= -\sqrt{(1+2\delta)}(\partial_x v_x + \partial_y v_y) - \partial_z v_z + f_V \\
\frac{1}{\rho c^2}\partial_t p_H &= -(1+2\epsilon)(\partial_x v_x + \partial_y v_y) - \sqrt{(1+2\delta)}\partial_z v_z + f_H
\end{aligned} \tag{1}$$

where ρ is the density, c is the velocity, (v_x, v_y, v_z) is the particle velocity vector, and p_V and p_H are pressure in the vertical and horizontal directions, respectively. The source term f_V and f_H are defined by the source wavelet $w(t)$ as follows:

$$f_V(t) = f_H(t) = \int_{-\infty}^{\tau} w(\tau)d\tau. \tag{2}$$

It is straightforward to see that when $\rho = 1$, $\epsilon = 0$ and $\delta = 0$, the first-order system 1 is equivalent to the familiar isotropic acoustic second-order wave-equation:

$$\frac{1}{c^2}\partial_t^2 p - \nabla p = w. \tag{3}$$

For simplicity, we can rewrite system 1 in a matrix-vector notation:

$$\mathbf{L}(c)\mathbf{p} = \mathbf{f}, \tag{4}$$

where $\mathbf{p} = (v_x,\ v_y,\ v_z,\ p_V,\ p_H)^T$, $\mathbf{f} = (0,\ 0,\ 0,\ f_V,\ f_H)^T$, and

$$\mathbf{L} = \begin{vmatrix}
\partial_t & 0 & 0 & 0 & \partial_x \\
0 & \partial_t & 0 & 0 & \partial_y \\
0 & 0 & \partial_t & \partial_z & 0 \\
\sqrt{1+2\delta}\partial_x & \sqrt{1+2\delta}\partial_y & \partial_z & \frac{1}{c^2}\partial_t & 0 \\
(1+2\epsilon)\partial_x & (1+2\epsilon)\partial_y & \sqrt{1+2\delta}\partial_z & 0 & \frac{1}{c^2}\partial_t
\end{vmatrix}. \tag{5}$$

VTI REVERSE-TIME MIGRATION IMAGING CONDITION

Traditionally, the subsurface image is often considered as the first gradient of an FWI objective function with respect to velocity. In this paper, we are going to derive the VTI reverse-time migration imaging condition according to the same criteria.

We define FWI objective function as

$$J_{\rm w} = \frac{1}{2}\langle d - d_{\rm est}, d - d_{\rm est}\rangle, \tag{6}$$

where $d_{\rm est}$ is the data estimated from the current model, which is sampled from wavefield $\mathbf{p}$, and d is the recorded data.

For the first iteration, $d_{\rm est} = 0$. Therefore the first gradient in velocity is:

$$\begin{aligned} \nabla_c J_{\rm w} &= \left(\frac{\partial \mathbf{p}}{\partial c}\right)^* d \\ &= (-\mathbf{L}^{-1}\frac{\partial \mathbf{L}}{\partial c}\mathbf{L}^{-1} f)^* d. \end{aligned} \tag{7}$$

Now we introduce the receiver vector field $\mathbf{q} = (u_x,\ u_y,\ u_z,\ q_V,\ q_H)^T$, which is the solution of the following equation:

$$\mathbf{L}^*(c)\mathbf{q} = \mathbf{f}\ '. \tag{8}$$

The equivalent source term in equation 8 is defined as $\mathbf{f}\ ' = (0,\ 0,\ 0,\ f'_V,\ f'_H)^T$, where $f'_V = f'_H = d$. From equation 5, we have

$$\frac{\partial \mathbf{L}}{\partial c} = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{2}{c^3}\partial_t & 0 \\ 0 & 0 & 0 & 0 & -\frac{2}{c^3}\partial_t \end{vmatrix}, \tag{9}$$

If we plug equation 8 and 9 into equation 7 and ignore the velocity dependence, we arrive at the imaging condition as follows:

$$\begin{aligned} I &= \mathbf{p}^*\mathbf{M}^*\mathbf{q}, \\ &= (\mathbf{M}\mathbf{p})^*\mathbf{q}, \end{aligned} \tag{10}$$

where

$$\mathbf{M} = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2}\partial_t & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2}\partial_t \end{vmatrix}. \tag{11}$$

The explicit form of this imaging condition for acoustic RTM is:

$$I = \int_0^{t_{max}} \frac{1}{2}\left((\partial t p_H) q_H + (\partial t p_V) q_V\right) dt. \tag{12}$$

The scaling factor $\frac{1}{2}$ is chosen to make sure that when $p_H = p_V$, equation 12 reduces to the isotropic cross-correlation imaging condition (Claerbout, 1987). For the purpose of velocity analysis, we often work with extended images and generalized imaging conditions. Similarly, we define our subsurface-offset-domain common-image gathers (SODCIGs) $\mathbf{I}$ as a column vector:

$$\mathbf{I} = [I_{-\mathbf{h}_{\max}}, I_{-\mathbf{h}_{\max}+\Delta\mathbf{h}}, \cdots, I_{\mathbf{0}}, \cdots, I_{\mathbf{h}_{\max}-\Delta\mathbf{h}}, I_{\mathbf{h}_{\max}}]^*, \tag{13}$$

where $\mathbf{h}$ is the half-subsurface offset, which ranges from $-\mathbf{h}_{\max}$ to $\mathbf{h}_{\max}$ with an increment of $\Delta\mathbf{h}$. For each element $I_{\mathbf{h}}$, the extended imaging condition is as follows (Sava and Formel, 2006) :

$$I_{\mathbf{h}} = (\mathbf{S}_{+\mathbf{h}}\mathbf{p})^{*}\mathbf{M}^{*}(\mathbf{S}_{-\mathbf{h}}\mathbf{q}), \tag{14}$$

where $\mathbf{S}_{+\mathbf{h}}$ is a shifting operator which shifts the wavefield by an amount of $+\mathbf{h}$ in the $\mathbf{x}$ direction. Notice that $(\mathbf{S}_{+\mathbf{h}})^{*} = \mathbf{S}_{-\mathbf{h}}$.

MIGRATION VELOCITY ANALYSIS GRADIENTS

In this section, we derive the MVA gradients of objective function 15 by two different approaches: the adjoint method from the perturbation theory, and the Lagrangian augmented function.

WEMVA is a non-linear inversion process that aims to find the velocity model that minimizes the residual field $\Delta\mathbf{I}$ in the image space. Without losing any generality, we define our objective function by DSO (Shen and Symes, 2008) in the subsurface offset domain:

$$J = \frac{1}{2}\sum_{\mathbf{h}}\langle \mathbf{h}I_{\mathbf{h}}, \mathbf{h}I_{\mathbf{h}}\rangle. \tag{15}$$

Although we don't use this DSO objective function in the example, the derivation follows the same logic, and readers can easily substitute their desired image-space objective function into the derivation.

Adjoint method from the perturbation theory

A perturbation, δc, of the velocity model c, induces a perturbation $\delta\mathbf{p}$ in the source wavefield vector $\mathbf{p}$, a perturbation $\delta\mathbf{q}$ in the receiver wavefield vector $\mathbf{q}$, a perturbation $\delta\mathbf{I}$ in the extended image cube $\mathbf{I}$, and hence a perturbation δJ in the objective function J. To the first order and using chain rule, δJ and δc have following relationship:

$$\delta J = \sum_{\mathbf{h}}\frac{\partial J}{\partial I_{\mathbf{h}}}\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}}\frac{\partial \mathbf{p}}{\partial c}\delta c + \sum_{\mathbf{h}}\frac{\partial J}{\partial I_{\mathbf{h}}}\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{q}}\frac{\partial \mathbf{q}}{\partial c}\delta c. \tag{16}$$

Now we can define the gradient by the back-projection of a unit perturbation in the objective function:

$$\begin{aligned}\nabla_c J &= \sum_{\mathbf{h}}\left(\frac{\partial J}{\partial I_{\mathbf{h}}}\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}}\frac{\partial \mathbf{p}}{\partial c}\right)^{*} + \sum_{\mathbf{h}}\left(\frac{\partial J}{\partial I_{\mathbf{h}}}\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{q}}\frac{\partial \mathbf{q}}{\partial c}\right)^{*} \\ &= (\nabla_c J)_1 + (\nabla_c J)_2 \,.\end{aligned} \tag{17}$$

Let's analyze the first term in equation 17 in detail, and the second term follows the same reasoning.

$$\begin{aligned}(\nabla_c J)_1 &= \sum_{\mathbf{h}} \left(\frac{\partial J}{\partial I_{\mathbf{h}}}\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}}\frac{\partial \mathbf{p}}{\partial c}\right)^* \\ &= \sum_{\mathbf{h}} \left(\frac{\partial \mathbf{p}}{\partial c}\right)^* \left(\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}}\right)^* \left(\frac{\partial J}{\partial I_{\mathbf{h}}}\right)^* \\ &= \sum_{\mathbf{h}} \mathbf{p}^* \left(-\frac{\partial \mathbf{L}}{\partial c}\right)^* \mathbf{L}^{-*} \left(\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}}\right)^* \left(\frac{\partial J}{\partial I_{\mathbf{h}}}\right)^* \\ &= \mathbf{p}^* \left(-\frac{\partial \mathbf{L}}{\partial c}\right)^* \mathbf{L}^{-*} \sum_{\mathbf{h}} \left(\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}}\right)^* \left(\frac{\partial J}{\partial I_{\mathbf{h}}}\right)^* \end{aligned} \tag{18}$$

where

$$\left(\frac{\partial J}{\partial I_{\mathbf{h}}}\right)^* = \mathbf{h}^* \mathbf{h} I_{\mathbf{h}}, \tag{19}$$

and

$$\begin{aligned}\left(\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}}\right)^* &= (\mathbf{S}_{+\mathbf{h}})^* \mathbf{M}^* (\mathbf{S}_{-\mathbf{h}} \mathbf{q}) \\ &= \mathbf{S}_{-\mathbf{h}} \mathbf{M}^* (\mathbf{S}_{-\mathbf{h}} \mathbf{q}). \end{aligned} \tag{20}$$

Plugging equation 19 and 20 into equation 18, we can rewrite equation 18 explicitly as follows:

$$\left(\frac{\partial J}{\partial c}\right)_1 = \mathbf{p}^* \left(-\frac{\partial \mathbf{L}}{\partial c}\right)^* \mathbf{L}^{-*} \sum_{\mathbf{h}} \mathbf{S}_{-\mathbf{h}} \mathbf{M}^* (\mathbf{S}_{-\mathbf{h}} \mathbf{q}) \mathbf{h}^2 I_{\mathbf{h}} \tag{21}$$

Similarly, we can obtain the explicit form for the second term in equation 17:

$$\left(\frac{\partial J}{\partial c}\right)_2 = \mathbf{q}^* \left(-\frac{\partial \mathbf{L}}{\partial c}\right) \mathbf{L}^{-1} \sum_{\mathbf{h}} \mathbf{S}_{+\mathbf{h}} \mathbf{M} (\mathbf{S}_{+\mathbf{h}} \mathbf{p}) \mathbf{h}^2 I_{\mathbf{h}} \tag{22}$$

Substituting equation 21 and equation 22 for the corresponding terms in equation 17, we now have derived the explicit form for the DSO gradient.

Lagrangian augmented functional method

We are now going to use the recipe with the augmented functional that Plessix (2006) provides to derive the image-space DSO gradient. First, let us form the Lagrangian augmented functional, $\mathcal{L}$:

$$\mathcal{L}(\mathbf{p}, \mathbf{q}, I_{\mathbf{h}}, \lambda, \mu, \gamma_{\mathbf{h}}, c) = \tag{23}$$

$$\begin{aligned}&\sum_{\mathbf{h}} \frac{1}{2} \langle \mathbf{h}\ I_{\mathbf{h}},\ \mathbf{h}\ I_{\mathbf{h}} \rangle \\ &+ \langle \lambda,\ \mathbf{f} - \mathbf{L}(\mathbf{c})\mathbf{p} \rangle \\ &+ \left\langle \mu,\ \mathbf{f}\,' - \mathbf{L}^*(\mathbf{c})\mathbf{q} \right\rangle \\ &+ \sum_{\mathbf{h}} \langle \gamma_{\mathbf{h}},\ (\mathbf{S}_{+\mathbf{h}}\mathbf{p})^* \mathbf{M}^* (\mathbf{S}_{-\mathbf{h}}\mathbf{q}) - I_{\mathbf{h}} \rangle \end{aligned} \tag{24}$$

Then the adjoint state equations are obtained by taking the derivative of $\mathcal{L}$ with respect to state variables $\mathbf{p}$, $\mathbf{q}$ and $I_{\mathbf{h}}$:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{p}} = -\mathbf{L}^*(c)\lambda + \sum_{\mathbf{h}} (\mathbf{S}_{+\mathbf{h}})^* \mathbf{M}^* (\mathbf{S}_{-\mathbf{h}}\mathbf{q})\gamma_{\mathbf{h}} = \mathbf{0}, \tag{25}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}} = -\mathbf{L}(c)\mu + \sum_{\mathbf{h}} (\mathbf{S}_{-\mathbf{h}})^* \mathbf{M} (\mathbf{S}_{+\mathbf{h}}\mathbf{p})\gamma_{\mathbf{h}} = \mathbf{0}, \tag{26}$$

$$\frac{\partial \mathcal{L}}{\partial I_{\mathbf{h}}} = -\gamma_{\mathbf{h}} + \mathbf{h}^2 I_{\mathbf{h}} = 0, \forall\ \mathbf{h}. \tag{27}$$

Equation 25, 26, 27 are the adjoint-state equations. Variables $\lambda = (\lambda_x, \lambda_y, \lambda_z, \lambda_V, \lambda_H)^T$ and $\mu = (\mu_x, \mu_y, \mu_z, \mu_V, \mu_H)^T$ are the adjoint-state fields and the solution of the adjoint-state equations 25 and 26. Variable $\gamma_{\mathbf{h}}$ is the scaled image slice at the subsurface offset $\mathbf{h}$.

Now the gradient of the objective function 15 with respect to velocity is:

$$\begin{aligned} \nabla_c J &= \left\langle \lambda,\ -\frac{\partial \mathbf{L}}{\partial \mathbf{c}}\mathbf{p} \right\rangle + \left\langle \mu,\ -\frac{\partial \mathbf{L}^*}{\partial \mathbf{c}}\mathbf{q} \right\rangle \\ &= \left(-\frac{\partial \mathbf{L}}{\partial c}\mathbf{p}\right)^* \lambda + \left(-\frac{\partial \mathbf{L}^*}{\partial \mathbf{c}}\mathbf{q}\right)^* \mu, \end{aligned} \tag{28}$$

If we combine equations 25, 26, and 27 with equation 28, we will arrive at the same solution as in the previous section.

Extension to update anisotropic parameters

The extension from isotropic model updates to anisotropic updates is straightforward. Built on the derivations in the last section, we can easily get the gradients for anisotropic parameters ϵ and δ as follows:

$$\begin{aligned} \nabla_\epsilon J &= \left\langle \lambda,\ -\frac{\partial \mathbf{L}}{\partial \epsilon}\mathbf{p} \right\rangle + \left\langle \mu,\ -\frac{\partial \mathbf{L}^*}{\partial \epsilon}\mathbf{q} \right\rangle \\ &= \left(-\frac{\partial \mathbf{L}}{\partial \epsilon}\mathbf{p}\right)^* \lambda + \mathbf{q}^* \left(-\frac{\partial \mathbf{L}}{\partial \epsilon}\right) \mu, \end{aligned} \tag{29}$$

$$\begin{aligned} \nabla_\delta J &= \left\langle \lambda,\ -\frac{\partial \mathbf{L}}{\partial \delta}\mathbf{p} \right\rangle + \left\langle \mu,\ -\frac{\partial \mathbf{L}^*}{\partial \delta}\mathbf{q} \right\rangle \\ &= \left(-\frac{\partial \mathbf{L}}{\partial \delta}\mathbf{p}\right)^* \lambda + \mathbf{q}^* \left(-\frac{\partial \mathbf{L}}{\partial \delta}\right) \mu, \end{aligned} \tag{30}$$

where

$$\frac{\partial \mathbf{L}}{\partial \epsilon} = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 2\partial_x & 2\partial_y & 0 & 0 & 0 \end{vmatrix}, \tag{31}$$

$$\frac{\partial \mathbf{L}}{\partial \delta} = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{\partial_x}{\sqrt{1+2\delta}} & \frac{\partial_y}{\sqrt{1+2\delta}} & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial_z}{\sqrt{1+2\delta}} & 0 & 0 \end{vmatrix}. \tag{32}$$

It is well-known that δ is the parameter most poorly constrained by surface seismic. Therefore, in our study, we assume that δ is obtained from well logs or seismic-well ties, and we invert only for velocity and ϵ.

Physical interpretation and implementation of the DSO gradient

In this subsection, we interpret each term in the DSO gradient formulation, and provide the readers with some hints for implementation. We find the Lagrangian formulation is easier to interpret, and readers can clearly relate the corresponding terms to the adjoint formulation. We will only discuss the physical meaning and the implementation for the first term in the gradient (Equation 28 and 25). Then similar reasoning can be argued using reciprocity.

First, for each image slice in the subsurface-offset domain $I_{\mathbf{h}}$, we compute a weighted image $\gamma_{\mathbf{h}}$ using equation 27. Then we move on to equation 25. We can rearrange the independent and commutable operators as follows:

$$\mathbf{L}^*(c)\lambda = \sum_{\mathbf{h}} \mathbf{S}_{-\mathbf{h}} \left(\mathbf{S}_{-\mathbf{h}} \mathbf{M}^* \mathbf{q}\right) \gamma_{\mathbf{h}}. \tag{33}$$

Operator $\mathbf{M}^*$ corresponds to differentiating q_V and q_H once reversely in time and setting u_x, u_y, and u_z fields to zero. Notice that the directions of propagation and differentiation in time of wavefield $\mathbf{q}$ are the same. Therefore, we can compute the time derivative during the same process as the propagation. Then we shift the reverse-time derivative $\mathbf{q}$ by $-\mathbf{h}$ in $\mathbf{x}$, and multiply it with the weighted image $\gamma_{\mathbf{h}}$. This product is shifted again by $-\mathbf{h}$. Finally, we sum over the contributions from all subsurface-offset image slices to get an effective source term $\mathbf{f}_p$. Next, we solve equation 33 for λ backward in time, using $\mathbf{f}_p$ as the source.

At the same time, in equation 28 $-\frac{\partial \mathbf{L}}{\partial c}$ is a sparse matrix, with non-zero elements only for p_V and p_H. We can therefore write everything out explicitly:

$$(\nabla_c J)_1 = \int_0^{t_{max}} \frac{2}{c^3} \left[(\partial_t p_H)\lambda_H + (\partial_t p_V)\lambda_V\right] dt. \tag{34}$$

The explicit forms for the complete gradients are:

$$\begin{aligned} \nabla_c J &= \int_0^{t_{max}} \frac{2}{c^3} \left[(\partial_t p_H)\lambda_H + (\partial_t p_V)\lambda_V\right] dt \\ &+ \int_0^{t_{max}} \frac{2}{c^3} \left[(\partial_t q_H)\mu_H + (\partial_t q_V)\mu_V\right] dt \end{aligned} \tag{35}$$

and

$$\begin{aligned} \nabla_\epsilon J &= -\int_0^{t_{max}} \left[(\partial_x v_x)\lambda_H + (\partial_y v_y)\lambda_H\right] dt \\ &+ \int_0^{t_{max}} \left[(\partial_x q_H)\mu_x + (\partial_y q_H)\mu_y\right] dt. \end{aligned} \tag{36}$$

Preconditioning the DSO gradient

Velocity model building is a highly underdetermined and nonlinear problem. Therefore, prior knowledge of the subsurface is needed to define a plausible subsurface model. In the formulation of Tarantola (1984), prior information is included as the covariance and the mean of the model. In this study, we assume the initial model we use is the mean, and the covariance of the model has two independent components: spatial covariance and collocated cross-parameter covariance (Li et al., 2011). In practice, instead of regularizing the inversion using Tarantola (1984), we use a preconditioning scheme (Claerbout, 2009): smoothing filtering to approximate square-root of the spatial covariance, and a standard-deviation matrix to approximate the square-root of the cross-parameter covariance.

Mathematically, the preconditioned model perturbation $d\mathbf{n}$ of the subsurface is defined as follows:

$$d\mathbf{m} = \mathbf{B}\mathbf{\Sigma} d\mathbf{n}, \tag{37}$$

where $\mathbf{m} = [c\ \epsilon]^T$. The smoothing operator $\mathbf{B}$ is a diagonal matrix:

$$\mathbf{B} = \begin{vmatrix} \mathbf{B}_c & 0 \\ 0 & \mathbf{B}_\epsilon \end{vmatrix}. \tag{38}$$

with different smoothing operators for velocity and ϵ, according to the geological information in the study area. The standard deviation matrix Σ:

$$\mathbf{\Sigma} = \begin{vmatrix} \sigma_{cc} & \sigma_{c\epsilon} \\ \sigma_{\epsilon c} & \sigma_{\epsilon\epsilon} \end{vmatrix}. \tag{39}$$

can be obtained by rock-physics modeling and/or lab measurements (Bachrach et al., 2011; Li et al., 2011).

We call $\mathbf{n}$ the preconditioning variable, and it relates to the original model $\mathbf{m}$ as follows:

$$\mathbf{m} = \mathbf{B}\mathbf{\Sigma}\mathbf{n} + (\mathbf{m}_0 - \mathbf{B}\mathbf{\Sigma}\mathbf{n}_0), \tag{40}$$

where $\mathbf{n}_0$ and $\mathbf{m}_0$ are the initial models in preconditioned space and physical space, respectively. Now, the gradient of the objective function 15 with respect to this preconditioning variable $\mathbf{n}$ is

$$\begin{aligned} \nabla_{\mathbf{n}} J &= (\frac{\partial \mathbf{m}}{\partial \mathbf{n}})^* \nabla_{\mathbf{m}} J \\ &= \mathbf{\Sigma}^* \mathbf{B}^* \nabla_{\mathbf{m}} J, \end{aligned} \tag{41}$$

where $\nabla_{\mathbf{m}} J = [\nabla_c J\ \nabla_\epsilon J]^T$.

In a steepest-decent inversion framework, the initial preconditioning model $\mathbf{n}_0$ is obtained by minimizing the following objective function:

$$J_{\text{init}} = \frac{1}{2} \left\langle \mathbf{m}_0 - \mathbf{B}\mathbf{\Sigma}\mathbf{n}_0, \mathbf{m}_0 - \mathbf{B}\mathbf{\Sigma}\mathbf{n}_0 \right\rangle. \tag{42}$$

For the i_{th} iteration

$$\mathbf{n}_{i+1} = \mathbf{n}_i + \alpha_i \nabla_{\mathbf{n}} J, \tag{43}$$

$$\begin{aligned}\mathbf{m}_{i+1} &= \mathbf{B}\mathbf{\Sigma}\mathbf{n}_{i+1} \\ &= \mathbf{B}\mathbf{\Sigma}\mathbf{n}_i + \alpha_i \mathbf{B}\mathbf{\Sigma}\nabla_{\mathbf{n}} J \\ &= \mathbf{m}_i + \alpha_i \mathbf{B}\mathbf{\Sigma}\mathbf{\Sigma}^*\mathbf{B}^*\nabla_{\mathbf{m}} J. \end{aligned} \tag{44}$$

Equation 44 suggests an interesting consideration in the context of nonlinear inversion: left-multiplying the gradient with a (semi)positive-definite matrix is equivalent to preconditioning with the square-root of the matrix; thus, the resulting direction is still a descent direction (Claerbout, 2009).

NUMERICAL TEST

In this section, we present the tests of our method on a synthetic VTI Marmousi model. We perform the tests in three steps: objective-function test, single parameter inversion, and joint inversion for two parameters.

Objective function test

We test our method on a VTI Marmousi model. First, synthetic Born data is generated using the models in Figures 1 and 2. The maximum offset is 3 km. Next, we perturb the ϵ model (Figure 2(a)) by a very smooth $\delta\epsilon$ field, as shown in Figure 3(a). We change the perturbation from -50% to 50% of the true ϵ model, and calculate the corresponding objective function respectively.

Ideally, we'd like to choose an objective function that reaches a local minimum at the correct model and is quadratic around the correct model, so that a gradient-based inversion scheme is guaranteed to converge. Based on the results, we choose an angle domain objective function instead of the DSO objective function (Equation 15):

$$J = \frac{1}{2}\langle \mathbf{DRI}, \mathbf{DRI} \rangle, \tag{45}$$

where $\mathbf{R}$ is the Radon transform operator, and $\mathbf{D}$ is the derivative operator along the ray-parameter axis.

As shown in Figure 3(b), the angle-domain objective function has a minimum at the correct epsilon model, and has a semi-quadratic shape with respect to the model perturbation. Therefore, this objective function is a good measure of the error in the anisotropic model. Notice that the tilting effect toward negative ϵ perturbation is caused by the limited acquisition geometry. This effect is negligible for velocity perturbation, because velocity has a first-order effect on the flatness of the angle gather, while ϵ's effect is second-order. We can increase the acquisition offset to mitigate this tilting effect and help the inversion.

Single parameter inversion

In this subsection, we invert for the anisotropic parameter ϵ alone. In this test, we model the synthetic data using very smooth ϵ (Figure 4(a)) and δ (Figure 4(b)) models as suggested by many field applications. To better constrain the inversion for ϵ, we also increase the maximum offset in the acquisition to 6 km.

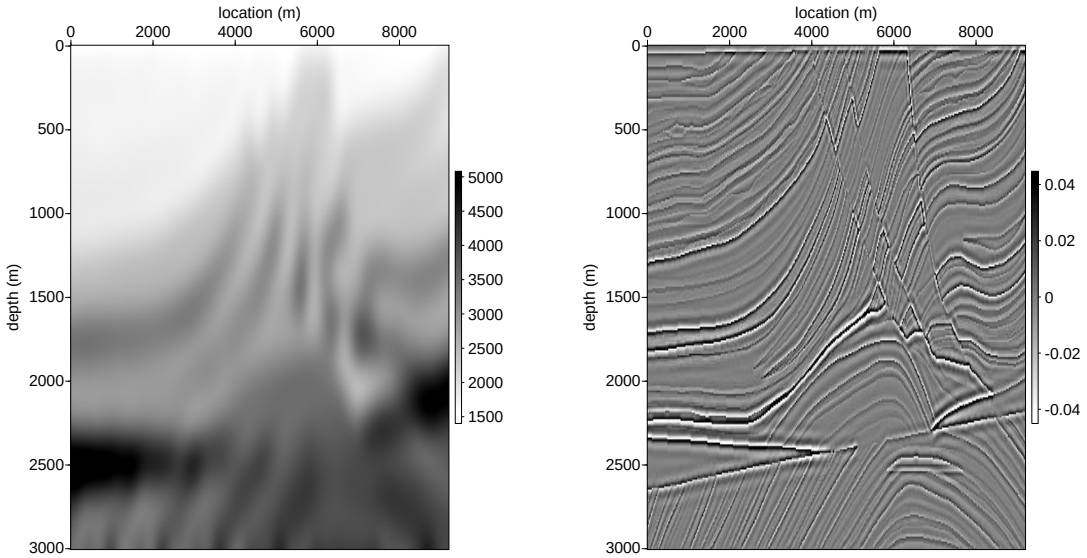


Figure 1: Smooth velocity model (a) in m/s and reflectivity model (b) used to generate the synthetic Born data. [**CR**] elita1/. vtrue,ref

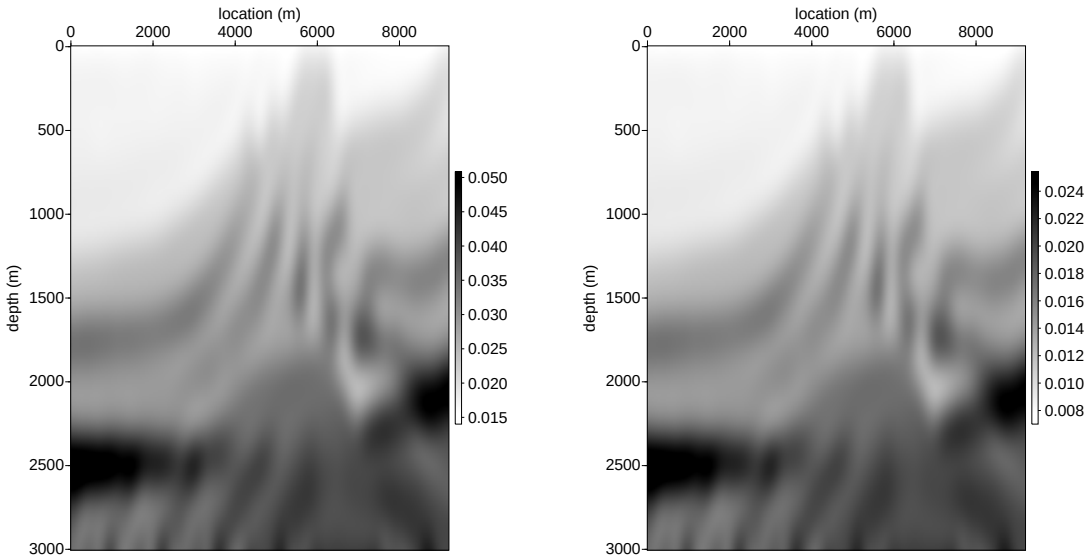
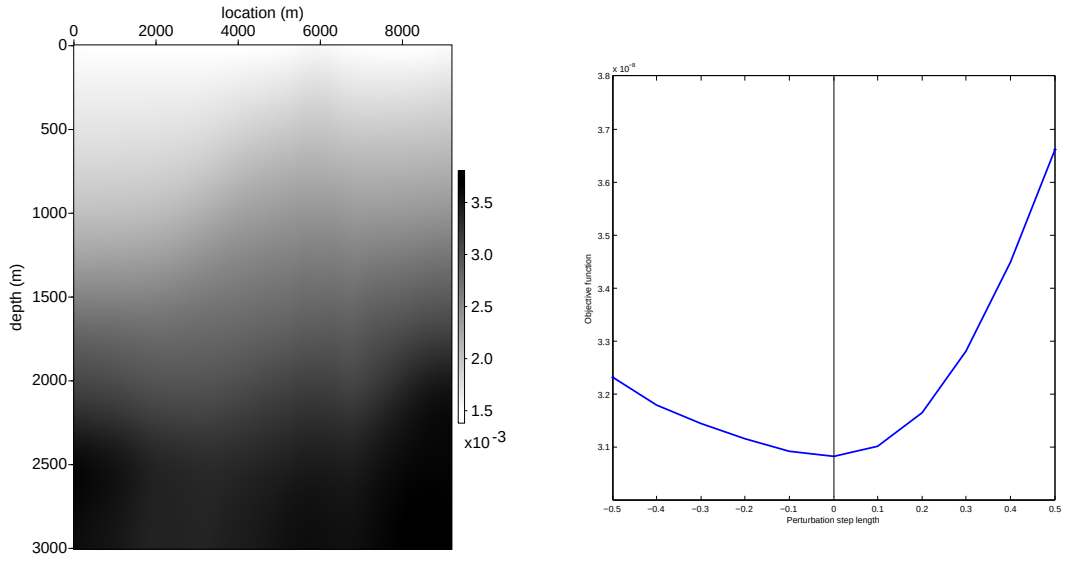


Figure 2: The ϵ model (a) and δ model (b) used to generate the synthetic Born data. [**CR**] elita1/. eps,del



(a) (b)

Figure 3: (a) The $\delta\epsilon$ model to test the objective function. (b) Objective function vs. ϵ perturbation. The angle-domain objective function 45 has a minimum at the correct epsilon model, and has a semi-quadratic shape with respect to the model perturbation. [**CR**] elita1/. deltaeps,objcurve

Compared with the true ϵ model, our initial ϵ model (Figure 5(a)) has negative perturbation of about 50% in the shallower part. Because a perfect velocity model is used in this case, the moveout at large angles is so small that it is almost undetectable to human eyes (Figure 5(b)). However, our inversion scheme is very sensitive to the residual moveout and successfully updates the ϵ model in the correct direction. Figure 6 shows the inverted ϵ model and the corresponding angle-domain common-image gathers after 40 iterations. Comparing with the initial angle gathers (Figure 5(b)), we can see that the slightly curving events at large angles are flattened and the inverted ϵ model is closer to the true one.

Joint inversion for two parameters

The tests in the previous sections show that we have a reliable objective function and successful inversion results for a single parameter. However, joint inversion for more than one parameter for each grid in the subsurface is far more challenging because of the ambiguity between parameters. As a result, the preconditioning scheme using geological and rock-physics information is crucial for its success.

In this test, we use the same synthetic data as in the last section. Unlike in the last example where we use the perfect velocity model, the starting models for velocity and ϵ are both inaccurate. The initial velocity model and ϵ model are shown in Figures 7(a) and 5(a), respectively. The angle gathers generated using these initial models are shown in Figure 7(b). Significant moveout in the angle-gather events indicates that the initial model is far from the true model. In fact, the initial velocity has a maximum of 15% error compared

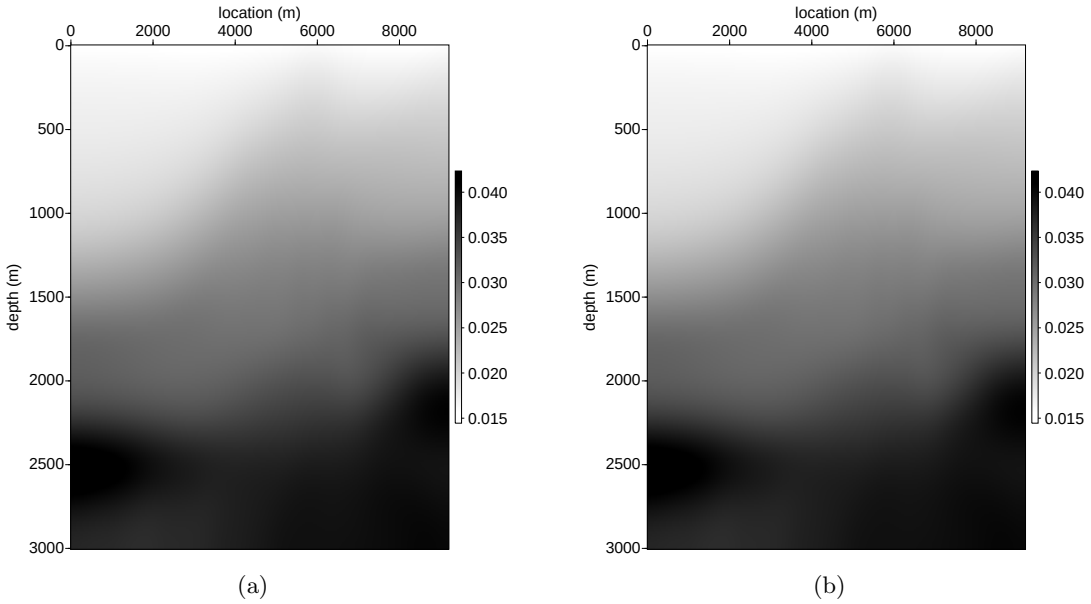


Figure 4: (a) True ϵ model and (b) true δ model used to generate the synthetic data. [**CR**] elita1/. epssm,dltsm

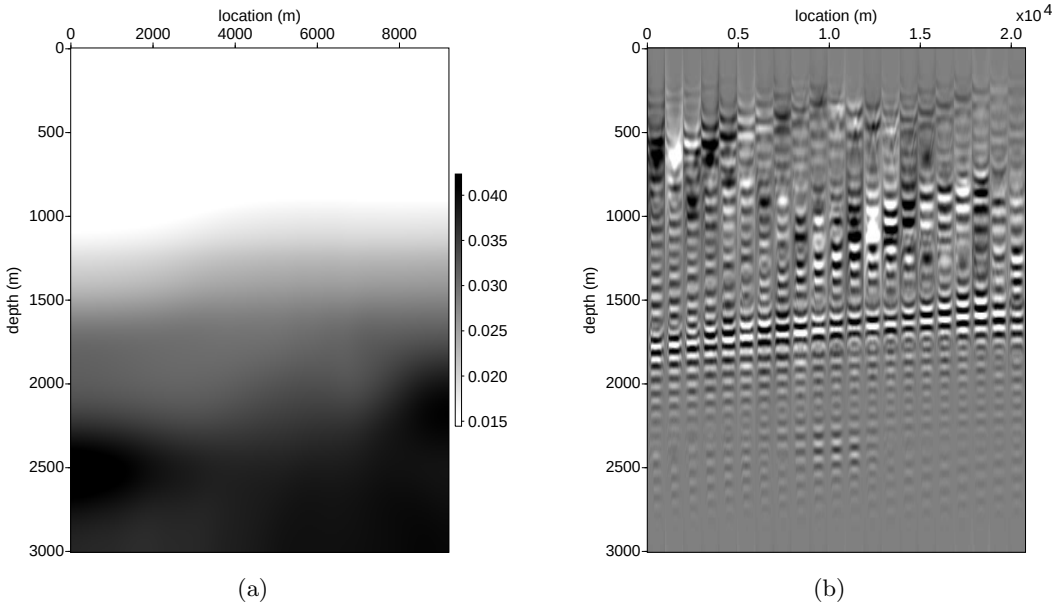
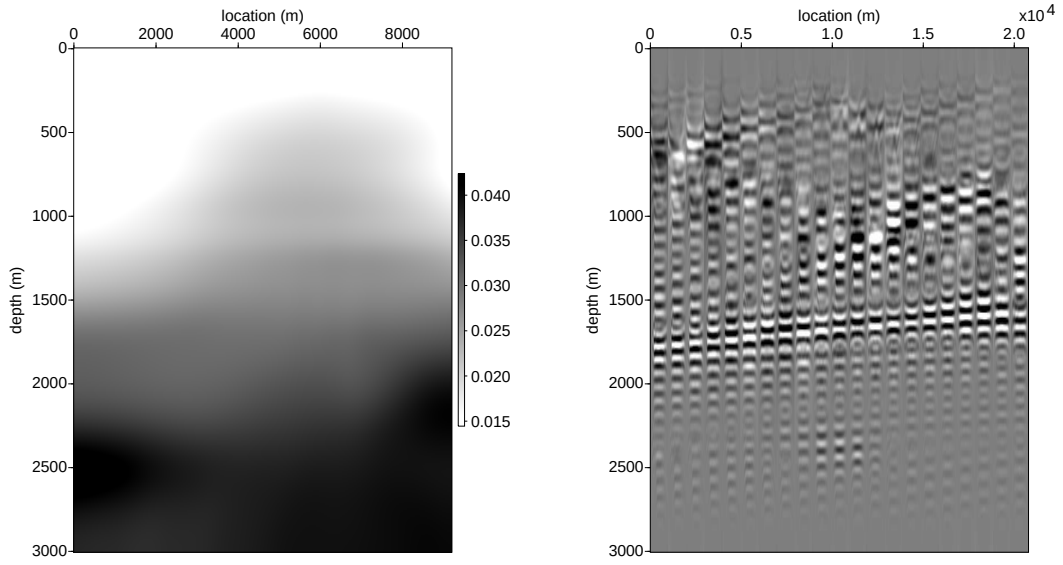


Figure 5: (a) Initial ϵ model and (b) initial angle-domain common-image gathers using initial ϵ model. Gathers are taken at every 10 common image point from $x = 4$ km to $x = 8$ km. [**CR**] elita1/. initeps,init-e-image



(a) (b)

Figure 6: (a) Inverted ϵ model and (b) final angle-domain common-image gathers using inverted ϵ model in (a). Compared with Figure 5(b), panel (b) shows more even energy across different angles. Gathers are taken at every 10 common image point from $x = 4$ km to $x = 8$ km. [**CR**] elita1/. inveps,inv-e-image

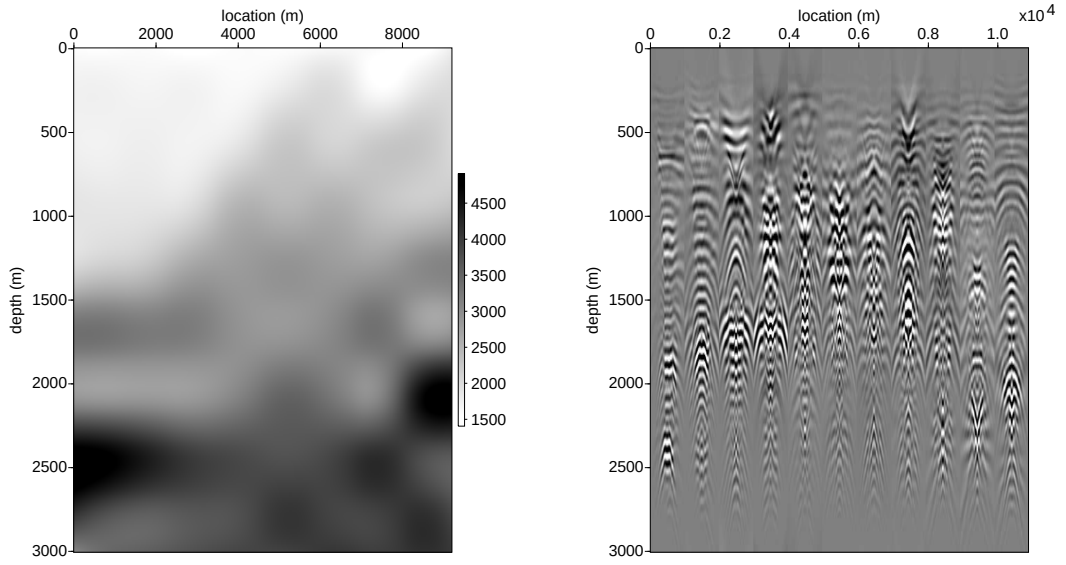
with the true velocity (Figure 1(a)), while the initial ϵ is about 50% smaller than the true value in the shallow part of the model. Notice that the error in velocity has a much larger effect on the kinematics of the seismic wave, hence a larger effect on the flatness in the angle domain.

After 40 iterations, we obtain the inverted velocity and ϵ models as shown in Figure 8(a) and 8(b). Comparing Figure 8(a) with Figure 1(a), we can conclude that the inversion has successfully recovered the high-resolution vertical structure in the shallow part of the model. Due to the limited illumination, the steep structure in the deeper part of the model is not well resolved. Comparing Figure 6(a) and Figure 8(b), we notice that, because of the error in velocity, the inversion does not converge to the same solution. This is an indication that we have not completely resolved the ambiguity between velocity and ϵ.

Angle gathers generated by the inverted model are shown in Figure 8(c). They are extracted from the same common-image points as in Figure 7(b). The improved model flattens the gathers across the whole section. Notice that the low-frequency energy in the water is the commonly seen wave-path energy for RTM images.

CONCLUSIONS

In this paper, we define an image-space inverse problem to solve for an optimized anisotropy model. To better describe the properties of the subsurface and the wave propagation, we use the first-order VTI two-way wave equation to compute our Green's function. Test results on objective function show that flatness in the angle domain is a valid measure of both



(a) (b)

Figure 7: Initial velocity model (a) in m/s and the angle gathers (b) obtained using initial velocity model. Initial ϵ model is shown in Figure 5(a). Model error causes significant curvatures in the angle gathers. Gathers are taken every 100 common image points from $x = 0$ km to $x = 9$ km. [**CR**] elita1/. vini,angini

velocity and anisotropic parameter ϵ. The results of single-parameter and joint inversions demonstrate that we have successfully resolved the high-frequency structure in velocity and successfully improved the anisotropic model where we have enough angle coverage. The updated image has flattened the gathers in the angle domain.

ACKNOWLEDGEMENT

This work was completed during my internship with Shell in the summer of 2011. The results were produced based on the previous C library for isotropic MVA. We thank Shell International Exploration and Production Company for permission to publish this paper. Yunyue Li thanks the sponsor of the Stanford Exploration Project for their financial support. Yunyue Li dedicates her special thanks to Peng Shen for his previous work and his help throughout her internship.

REFERENCES

Bachrach, R., Y. K. Liu, M. Woodward, O. Zradrova, Y. Yang, and K. Osypov, 2011, Anisotropic velocity model building using rock physics: Comparison of compaction trends and check-shot-derived anisotropy in the Gulf of Mexico: SEG Expanded Abstract, **30**, 207–211.

Claerbout, J., 1987, Imaging the earth's interior: Blackwell Scientific Publications.

Claerbout, J. F., 2009, Image estimation by example.

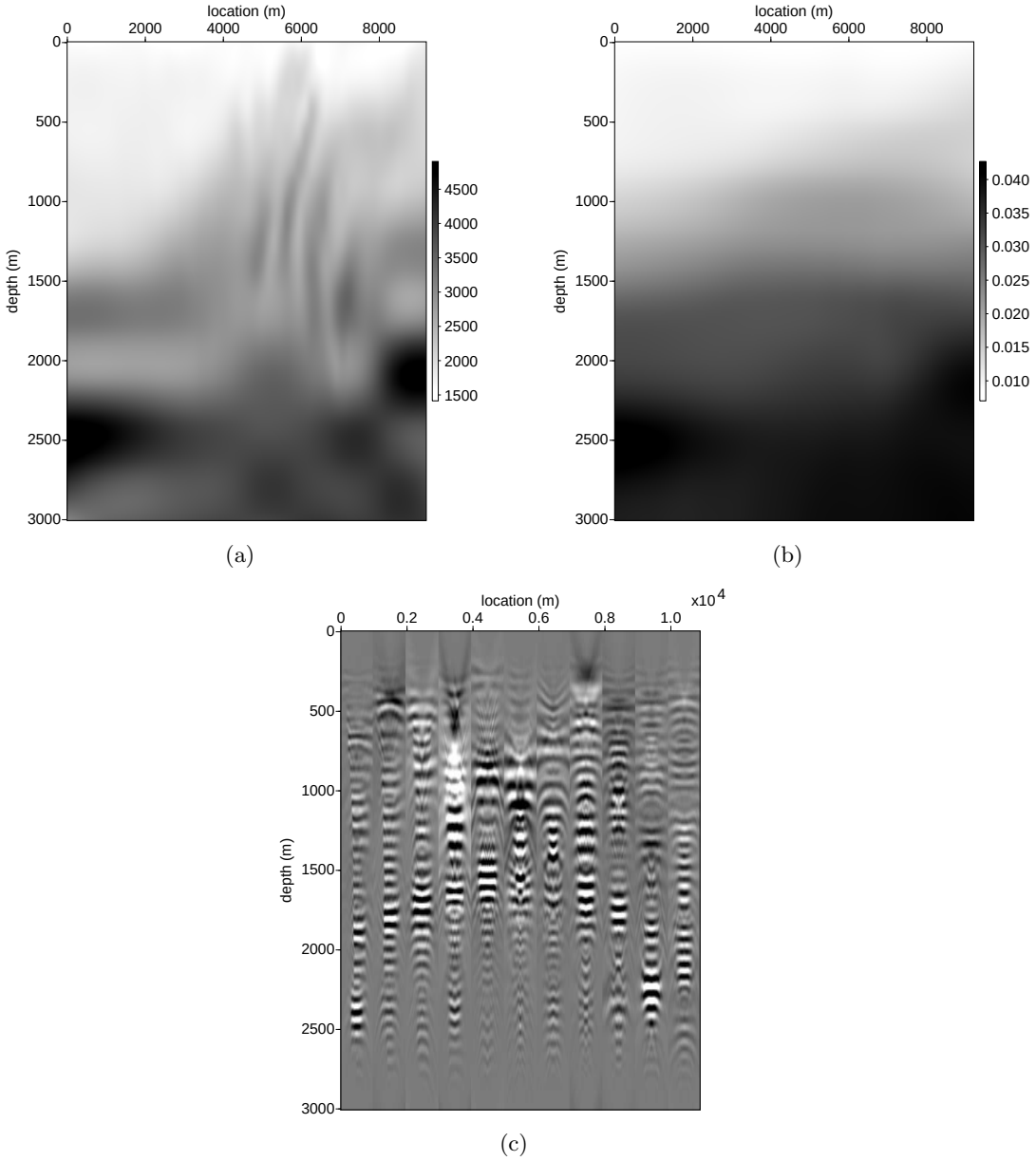


Figure 8: Inverted velocity model (a) in m/s and ϵ model (b) after 40 iterations. Angle gathers (c) obtained by the inverted model. Angle gathers are extracted at the same CIP as those in Figure 7(b). Improved velocity and ϵ flattens the corresponding angle gathers. Gathers are taken every 100 common image points from $x = 0$ km to $x = 9$ km. [**CR**] elita1/. vfinal,epsfinal,angfinal

Duveneck, E., P. Milcik, P. M. Bakker, and C. Perkins, 2008, Acoustic VTI wave equations and their application for anisotropic reverse-time migration: SEG Expanded Abstracts, **27**.

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

Guerra, C., Y. Tang, and B. Biondi, 2009, Wave-equation tomography using image-space phase-encoded data: SEP-Report, **138**, 167–180.

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.

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

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

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

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

Sava, P. and S. Formel, 2006, Generalized imaging conditions for wave equation migration: CWP Report, **524**.

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

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

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

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

Zhang, Y. and H. Zhang, 2009, A stable TTI reverse time migration and its implementation: SEG Expanded Abstracts, **28**, 3–11.

Early-arrival waveform inversion for near-surface velocity and anisotropic parameters: modeling and sensitivity kernel analysis

Xukai Shen

ABSTRACT

Understanding the relationships between vertical velocity, anisotropic parameters ϵ, δ and traveltime is important both for parametrization of waveform inversion and proper interpretation of waveform inversion results. Forward modeling indicates that traveltime is more sensitive to vertical velocity and ϵ than to δ. This suggests that δ may be fixed during an inversion for vertical velocity and ϵ. When vertical velocity and ϵ are directly parametrized, ϵ changes little during the inversion. A more balanced sensitivity kernel can be obtained by using different parametrizations, such as the vertical and horizontal velocities, or the logarithms of slowness and ϵ.

INTRODUCTION

Full Waveform Inversion (FWI) (Tarantola, 1984; Pratt et al., 1998; Mora, 1987) iteratively updates model by trying to match input data with modeled data. It estimates subsurface velocity more accurately than conventional techniques, such as ray-based methods (Hampson and Russell, 1984; Olson, 1984; White, 1989), especially in geologically complex areas. This is because FWI predicts kinematics of recorded data more accurately by using finite-difference wave-equations, compared with ray-based methods using high frequency approximations of wave propagation. However, the dynamics of recorded data are not very accurately predicted by current FWI methods. In other words, successful field-data application of FWI usually relies more on matching kinematics of recorded data.

With longer-offset data ($>$ 10km) commonly acquired these days, matching kinematics means matching data traveltime over the entire offset range. For such large offset ranges, anisotropic effects, if they exist, are no longer negligible. In the presence of anisotropy, if isotropic FWI is used, the inversion results will not correctly reflect the true subsurface attributes (Ghilami et al., 2011). More specifically, isotropic FWI of diving waves mostly recovers the horizontal velocity in anisotropic media. Migration using such a velocity will place reflectors at incorrect depths. To avoid this, anisotropic parameter estimation should be part of the inversion process. Such inversion can be carried out in several ways. One way is to perform single-parameter inversion for each parameter sequentially. Alternatively, joint inversion performs simultaneous inversion of multiple parameters, called joint inversion. Joint inversion is usually better, since the results of the first approach are sensitive to the order in which the inversion is performed. On the other hand, given the original definition of anisotropic parameters (Thomsen, 1986), direct changes in those parameters themselves usually result in very small changes in data kinematics, where as changes in velocity affect the data kinematics much more significantly (Plessix and Cao, 2011). For the purpose of

simultaneous inversion, it is important to understand quantitatively how much the data kinematics change as a function of anisotropic parameters and velocity, and to come up with an effective parametrization of the model.

In this paper, I first describe the acoustic vertical transversely isotropic (VTI) wave-equation and the sensitivity kernel calculation. Then I use synthetic data examples to illustrate the sensitivity of the data kinematics to the anisotropic parameters and the velocity, in terms of both forward modeling and sensitivity-kernel calculation.

FORWARD MODELING

Exact anisotropic wave equations are in the form of elastic wave equations. Acoustic anisotropic wave equations can be obtained by various approximations of the exact elastic equations. One way to do this is to set shear-wave velocity to zero in the exact elastic wave equations. Detailed derivations can be found in several papers (Zhang and Zhang, 2009; Crawley et al., 2010; Duveneck et al., 2008). The resulting acoustic anisotropic wave equations are a system of second-order equations:

$$\begin{aligned} \frac{\partial^2 p}{\partial t^2} &= {v_p}^2 (1+2\epsilon) \frac{\partial^2 p}{\partial x^2} + {v_p}^2 \sqrt{1+2\delta} \frac{\partial^2 r}{\partial z^2} \\ \frac{\partial^2 r}{\partial t^2} &= {v_p}^2 \sqrt{1+2\delta} \frac{\partial^2 p}{\partial x^2} + {v_p}^2 \frac{\partial^2 r}{\partial z^2}, \end{aligned} \tag{1}$$

where p and r are horizontal and vertical stress, respectively, v_p is vertical p-wave velocity and ϵ and δ are anisotropic parameters (Thomsen, 1986).

To illustrate the traveltime sensitivity of data to parameter changes, I use part of the BP 2002 benchmark model. The original synthetic model has only p-wave velocity; the anisotropic parameters were created from the velocity model according to typical Gulf of Mexico anisotropic parameters. Velocity, ϵ and δ models are shown in Figure 1.

Four different modeling experiments with modeled shot records are shown in Figure 2. The first experiment is the VTI anisotropic modeling using all three fields shown in Figure 1. The second experiment is the same VTI modeling but with $\delta = 0$. The third experiment is isotropic modeling using the velocity field only, and the fourth is isotropic modeling using the horizontal p-wave velocity, which is defined as $v_h = v_p\sqrt{1+2\epsilon}$. Figure 3 shows the refraction traveltime difference of the latter three shots compared to the first shot. It can be seen that traveltime is insensitive to δ changes, but is sensitive to ϵ changes. Also, isotropic modeling using the horizontal p-wave velocity results in non-trivial traveltime differences, which means that even using isotropic FWI, the retrieved model is not necessarily the horizontal p-wave velocity, as was previously thought (Ghilami et al., 2011).

MODEL PARAMETRIZATION AND SENSITIVITY-KERNEL CALCULATION

For inversion, there are several ways to parametrize the model space, which contains vertical p-wave velocity and the anisotropic parameters ϵ and δ. However, due to insensitivity of the data to the δ parameter, currently I consider a model space consisting of only vertical p-wave velocity and ϵ; i.e. δ is fixed during inversion. Using gradient-based inversion

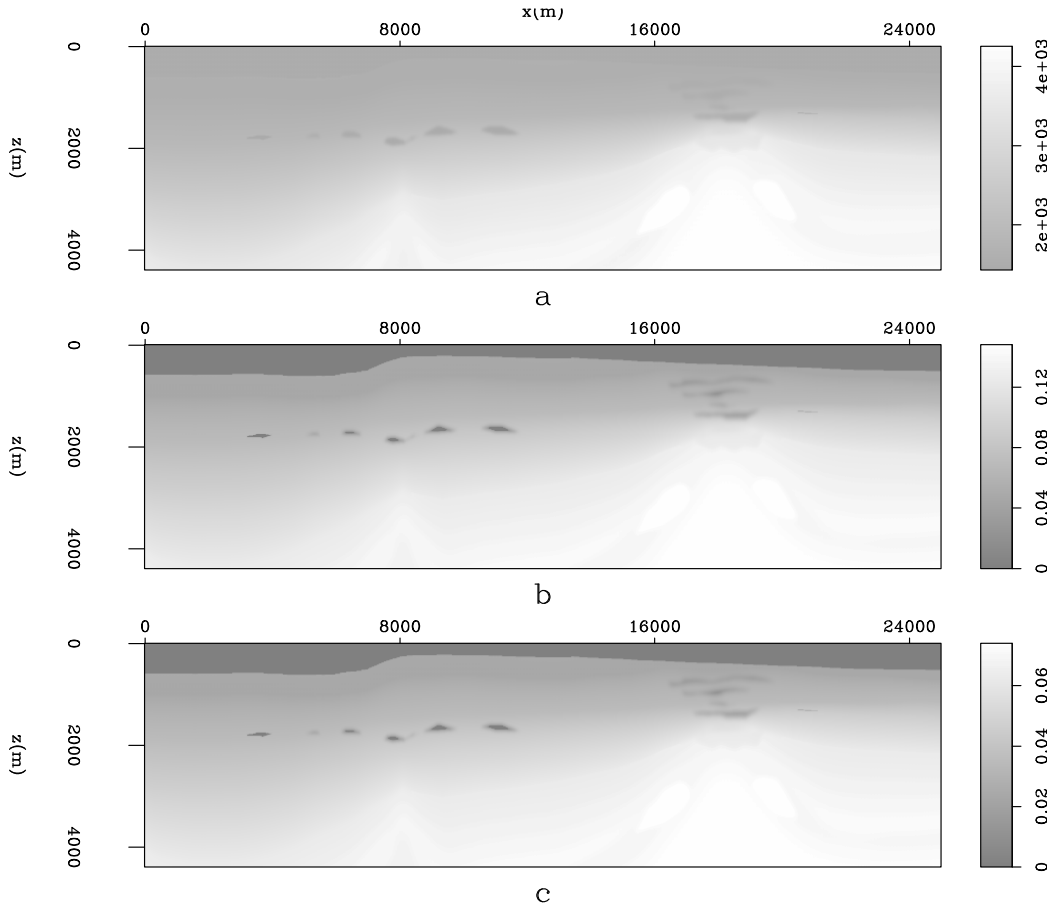


Figure 1: Reference model for various modeling experiments. Top: velocity model; middle: ϵ model; bottom: δ model. [**ER**] xukai1/. modelbw

methods, different parametrization leads to different model updates. In joint inversion, it is unfavorable to have updates that result in little or almost no change in one of the model components. Quantitatively, for a model space that contains two components m_1 and m_2, update directions g_{m1} and g_{m2} should be chosen such that $g_{m1}/m_1 \approx g_{m2}/m_2$. This can be achieved by using proper parametrization.

I compare three different parametrizations and their corresponding model updates (sensitivity kernels). For sensitivity-kernel calculation, I first calculate the full data residual by subtracting d_{cal}, which is modeled from the smoothed version of the true model (Figure 4) from d_{obs}, which is modeled from the true model. I consider only a single trace refraction of the total data residual for this study.

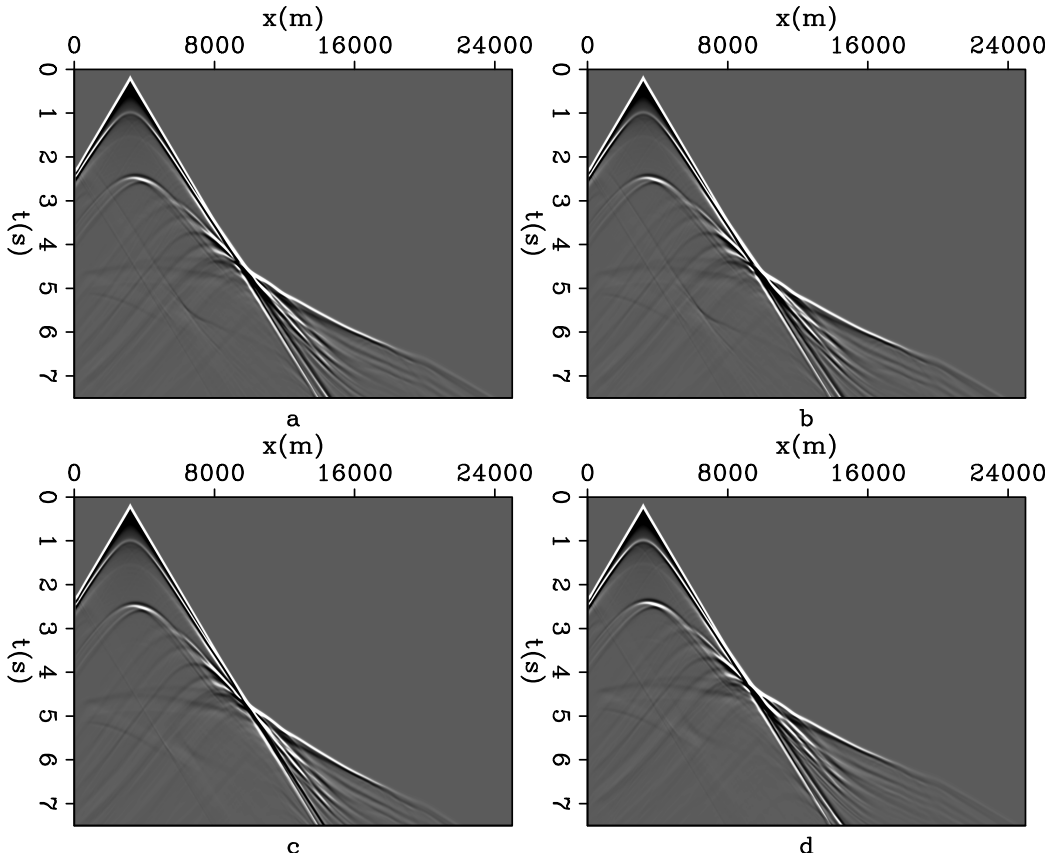


Figure 2: Shots from different modeling experiments: a) VTI modeling using all three fields shown in Figure 1; b) same as a) except that $\delta = 0$; c) isotropic modeling using only the velocity field; d) isotropic modeling using horizontal p-wave velocity $v_h = v_p\sqrt{1+2\epsilon}$. [**ER**] xukai1/. record

Naive parametrization

The most naive parametrization directly inverts velocity and ϵ. However, to avoid higher-order term involving both variables, equation 1 can be rewritten as follows:

$$\begin{aligned} m_1 \frac{\partial^2 p}{\partial t^2} &= m_2 \frac{\partial^2 p}{\partial x^2} + \sqrt{1+2\delta}\frac{\partial^2 r}{\partial z^2} \\ m_1 \frac{\partial^2 r}{\partial t^2} &= \sqrt{1+2\delta}\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 r}{\partial z^2}, \end{aligned} \tag{2}$$

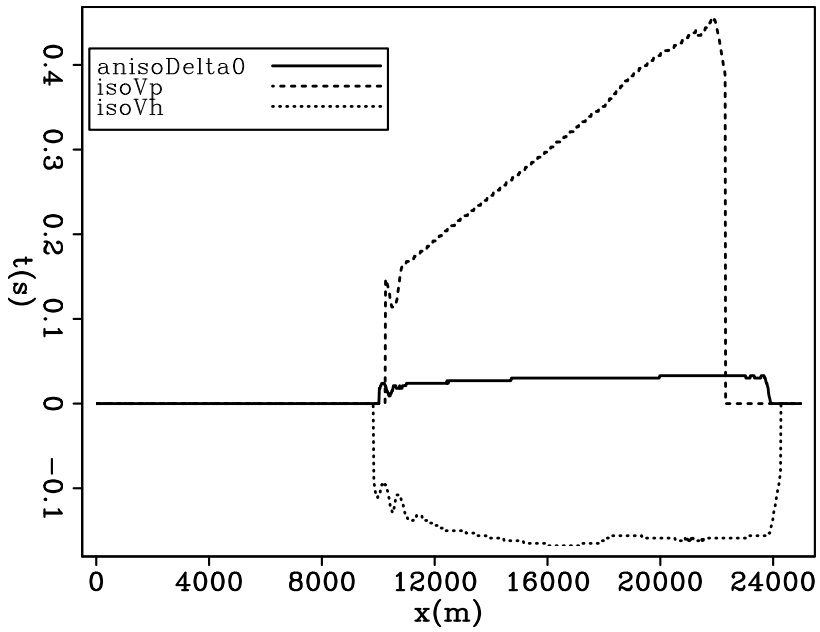


Figure 3: Refraction traveltime difference of shots compared to the VTI case. [**ER**] xukai1/. tdif

where $m_1 = {v_p}^{-2}$ and $m_2 = 1 + 2\epsilon$ are the two model variables, assuming

$$\begin{aligned} m_1 &= m_{1,0} + \Delta m_1 \\ m_2 &= m_{2,0} + \Delta m_2 \\ p &= p_0 + \Delta p \\ r &= r_0 + \Delta r \\ m_{1,0}\frac{\partial^2 p_0}{\partial t^2} &= m_{2,0}\frac{\partial^2 p_0}{\partial x^2} + \sqrt{1+2\delta}\frac{\partial^2 r_0}{\partial z^2} \\ m_{1_0}\frac{\partial^2 r_0}{\partial t^2} &= \sqrt{1+2\delta}\frac{\partial^2 p_0}{\partial x^2} + \frac{\partial^2 r_0}{\partial z^2}, \end{aligned} \tag{3}$$

By combining this with equation 2, we can obtain

$$\begin{aligned} m_{1,0}\frac{\partial^2 \Delta p}{\partial t^2} - m_{2,0}\frac{\partial^2 \Delta p}{\partial x^2} - \sqrt{1+2\delta}\frac{\partial^2 \Delta r}{\partial z^2} &= -\Delta m_1 \frac{\partial^2 p_0}{\partial t^2} + \Delta m_2 \frac{\partial^2 p_0}{\partial x^2} \\ m_{1,0}\frac{\partial^2 \Delta r}{\partial t^2} - \sqrt{1+2\delta}\frac{\partial^2 \Delta p}{\partial x^2} - \frac{\partial^2 \Delta r}{\partial z^2} &= -\Delta m_1 \frac{\partial^2 r_0}{\partial t^2}, \end{aligned} \tag{4}$$

which can be written in the following matrix form:

$$\begin{vmatrix} m_{1,0}\frac{\partial^2}{\partial t^2} - m_{2,0}\frac{\partial^2}{\partial x^2} & -\sqrt{1+2\delta}\frac{\partial^2}{\partial z^2} \\ -\sqrt{1+2\delta}\frac{\partial^2}{\partial x^2} & m_{1,0}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial z^2} \end{vmatrix} \begin{vmatrix} \Delta p \\ \Delta r \end{vmatrix} = \begin{vmatrix} -\frac{\partial^2 p_0}{\partial t^2} & \frac{\partial^2 p_0}{\partial x^2} \\ -\frac{\partial^2 r_0}{\partial t^2} & 0 \end{vmatrix} \begin{vmatrix} \Delta m_1 \\ \Delta m_2 \end{vmatrix}, \tag{5}$$

This establishes a linear relationship between model perturbation and data perturbation, and can be used to calculate the sensitivity kernel. Figure 5 shows the relative sensitivity kernels of the two model parameters, which are defined as

$$k_{m_i} = g_{m_i}/m_{i,0}, \tag{6}$$

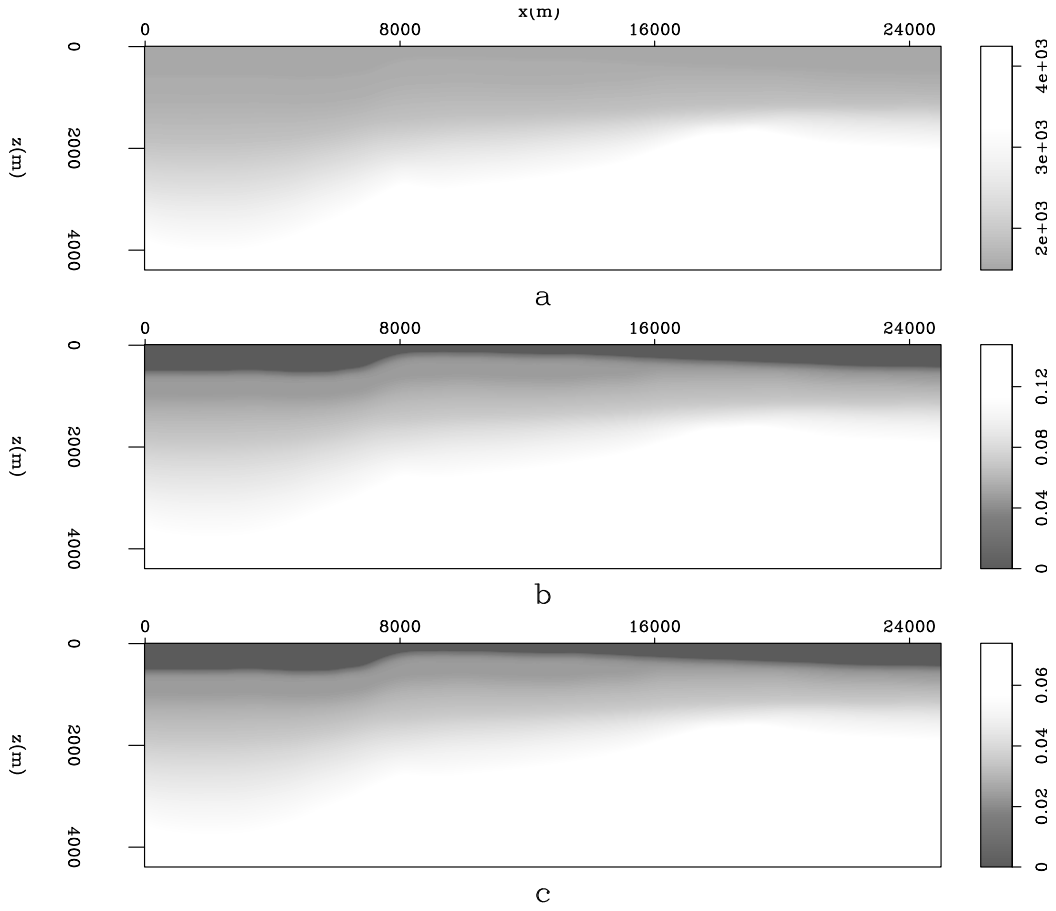


Figure 4: Smooth model for generating d_{cal}: a) velocity model; b) ϵ model; c) δ model. [**ER**] xukai1/. modelsbw

where g_{m_i} is the sensitivity kernel, and k_{m_i} is the relative sensitivity kernel. the clipping value of the top figure is over sixteen orders of magnitude larger than the bottom figure, which means if we are to use this parametrization for our inversion, there will be almost no updates of the anisotropic parameter.

Velocity parametrization

Another parametrization is to use velocities for both variables. Defining $m_1 = {v_p}^2$ and $m_2 = {v_p}^2 (1 + 2\epsilon)$, equation 1 can be rewritten as follows:

$$\begin{aligned} \frac{\partial^2 p}{\partial t^2} &= m_2 \frac{\partial^2 p}{\partial x^2} + m_1 \sqrt{1+2\delta} \frac{\partial^2 r}{\partial z^2} \\ \frac{\partial^2 r}{\partial t^2} &= m_1 \sqrt{1+2\delta} \frac{\partial^2 p}{\partial x^2} + m_1 \frac{\partial^2 r}{\partial z^2}, \end{aligned} \tag{7}$$

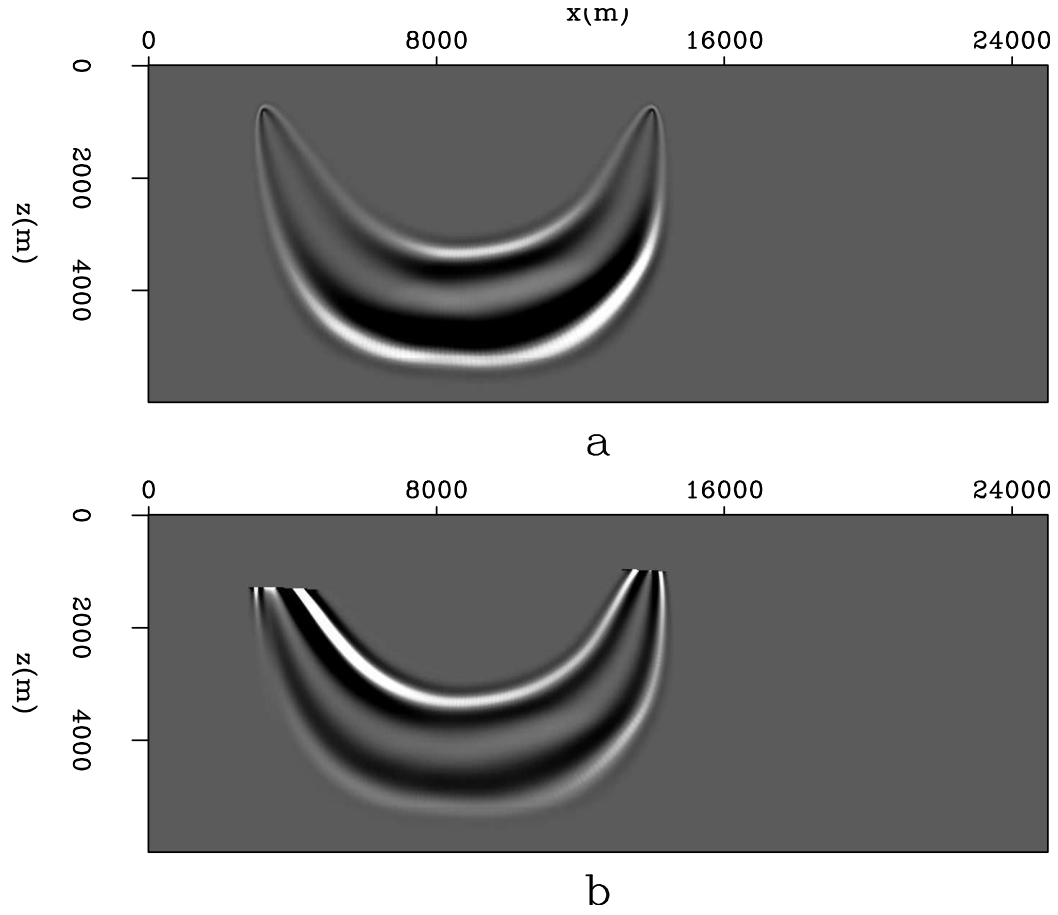


Figure 5: Relative sensitivity kernel of parameter: a) ${v_p}^{-2}$; b) $1+2\epsilon$. Clipping value of the top figure is $1e10$, clipping value of the bottom figure is $2e-6$. [**ER**] xukai1/. relimpn

Using similar procedure to the one described in the previous section, we can obtain a matrix form expression:

$$\begin{vmatrix} \frac{\partial^2}{\partial t^2} - m_{2,0}\frac{\partial^2}{\partial x^2} & -m_{1,0}\sqrt{1+2\delta}\frac{\partial^2}{\partial z^2} \\ -m_{1,0}\sqrt{1+2\delta}\frac{\partial^2}{\partial x^2} & \frac{\partial^2}{\partial t^2} - m_{1,0}\frac{\partial^2}{\partial z^2} \end{vmatrix} \begin{vmatrix} \Delta p \\ \Delta r \end{vmatrix} = \begin{vmatrix} \sqrt{1+2\delta}\frac{\partial^2 r_0}{\partial z^2} & \frac{\partial^2 p_0}{\partial x^2} \\ \sqrt{1+2\delta}\frac{\partial^2 p_0}{\partial x^2} + \frac{\partial^2 r_0}{\partial z^2} & 0 \end{vmatrix} \begin{vmatrix} \Delta m_1 \\ \Delta m_2 \end{vmatrix}, \quad (8)$$

This is the linear relationship between model perturbation and data perturbation, and can be used to calculate the sensitivity kernel. Figure 6 shows the relative sensitivity kernel of the two model parameters. Both figures are clipped to the same value. Since both variables are parametrized as velocities, their updates are of similar strength, and this is an effective parametrization.

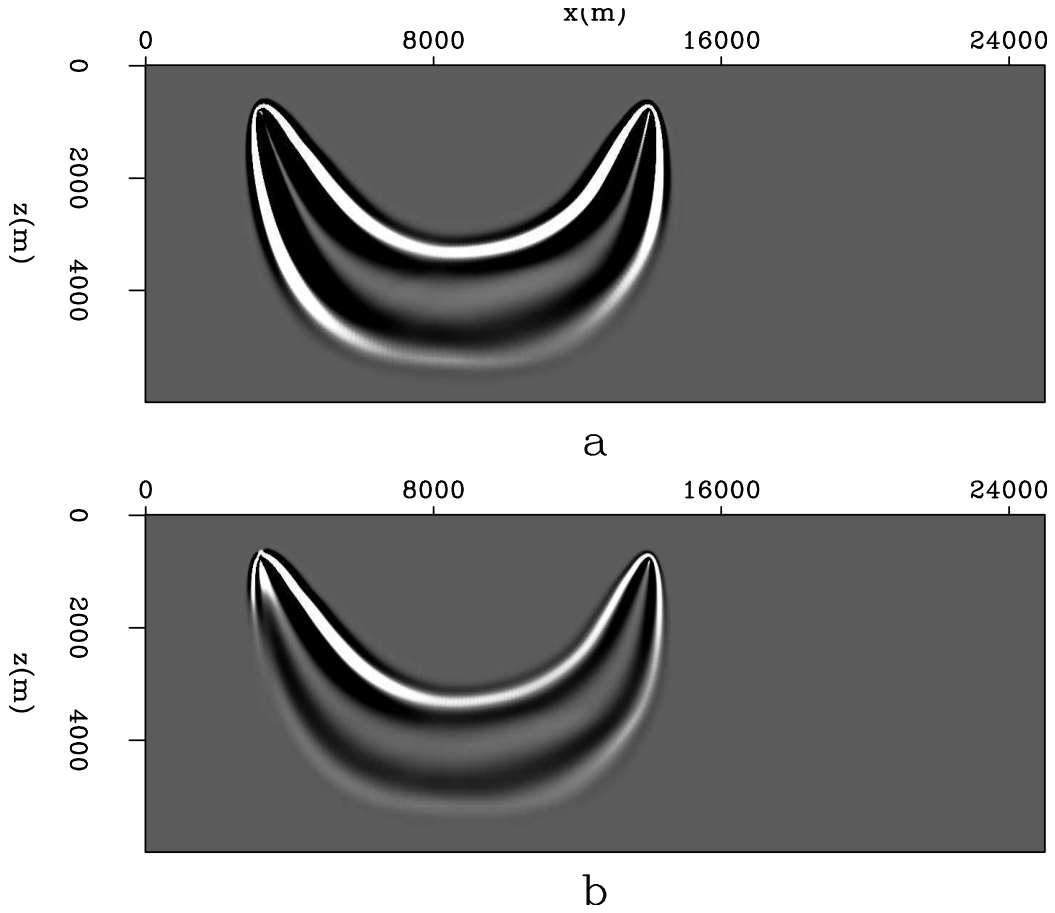


Figure 6: Relative sensitivity kernel of parameter: a) ${v_p}^2$; b) ${v_p}^2\,(1+2\epsilon)$. Both figures are clipped at the same value. [**ER**] xukai1/. relimpv

Logarithmic velocity parametrization

A slightly different parametrization is to define $m_1 = ln\left({v_p}^{-2}\right)$ and $m_2 = 1+2\epsilon$. Using this parametrization, equation 1 can be rewritten as follows:

$$\begin{aligned} e^{m_1}\frac{\partial^2 p}{\partial t^2} &= m_2\frac{\partial^2 p}{\partial x^2} + \sqrt{1+2\delta}\frac{\partial^2 r}{\partial z^2} \\ e^{m_1}\frac{\partial^2 r}{\partial t^2} &= \sqrt{1+2\delta}\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 r}{\partial z^2}, \end{aligned} \tag{9}$$

Using a procedure similar to the one described in the previous section, we can obtain a matrix form expression:

$$\begin{vmatrix} e^{m_{1,0}}\frac{\partial^2}{\partial t^2} - m_{2,0}\frac{\partial^2}{\partial x^2} & -\sqrt{1+2\delta}\frac{\partial^2}{\partial z^2} \\ -\sqrt{1+2\delta}\frac{\partial^2}{\partial x^2} & e^{m_{1,0}}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial z^2} \end{vmatrix} \begin{vmatrix} \Delta p \\ \Delta r \end{vmatrix} = \begin{vmatrix} -e^{m_{1,0}}\frac{\partial^2 p_0}{\partial t^2} & \frac{\partial^2 p_0}{\partial x^2} \\ -e^{m_{1,0}}\frac{\partial^2 r_0}{\partial t^2} & 0 \end{vmatrix} \begin{vmatrix} \Delta m_1 \\ \Delta m_2 \end{vmatrix}, \tag{10}$$

This is the linear relationship between model perturbation and data perturbation, and can be used to calculate the sensitivity kernel. Figure 7 shows relative sensitivity kernel of the two model parameters. Both figures are clipped to the same value. With this

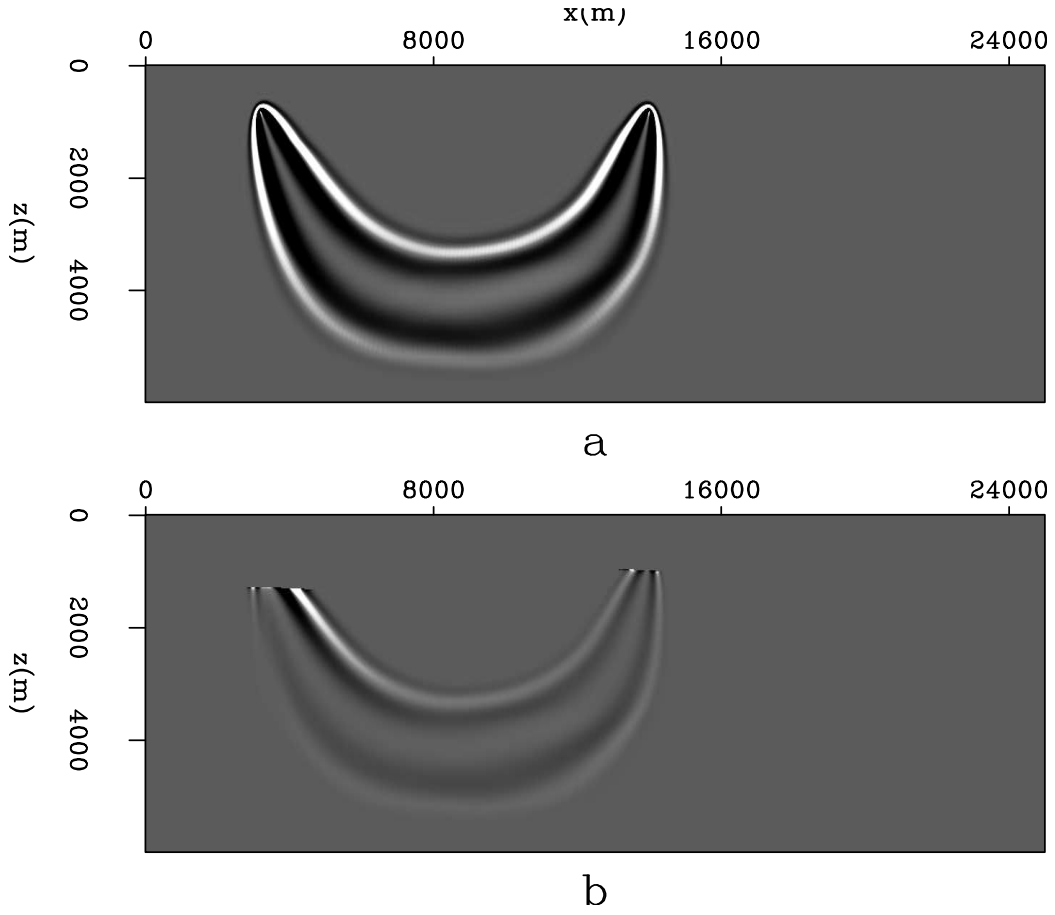


Figure 7: Relative sensitivity kernel of parameter: a) $a^{-1} \ln \left({v_p}^{-2}\right)$; b) $(1 + 2\epsilon)$. Both figures are clipped at the same value. [**ER**] xukai1/. relimpp

parametrization, updates of both variables are of the same order of strength. I show in another paper [citation] that such parametrization results in very good inversion results.

CONCLUSIONS

The forward modeling experiment suggests that it is important to include anisotropy as part of the inversion. For joint inversion, proper parametrization can give balanced updates to each model variable. Velocity parametrization and logarithmic slowness parametrization are both good candidates in that sense. Actual inversion using these parametrization needs to be further studied.

REFERENCES

Crawley, S., S. Brandsberg-Dahl, J. McClean, and N. Chemingui, 2010, TTI reverse time migration using the pseudo-analytic method: The Leading Edge, **29**, 1378.

Duveneck, E., P. Milcik, P. M. Bakker, and C. Perkins, 2008, Acoustic VTI wave equations

and their application for anisotropic reverse-time migration: SEG Expanded Abstracts, 2186–2189.

Ghilami, Y., R. Brossier, S. Operto, V. Prieux, A. Ribodetti, and J. Virieux, 2011, Two-dimensional acoustic anisotropic (VTI) full waveform inversion: The Valhall case study: SEG Expanded Abstracts, **30**, 2543.

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

Mora, P., 1987, Elastic wavefield inversion: Stanford Exploration Project Ph.D. Thesis.

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

Plessix, R. E. and Q. Cao, 2011, A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium: Geophysical Journal International, **185**, 539–556.

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.

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

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

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

Zhang, Y. and H. Zhang, 2009, A stable TTI reverse time migration and its implementation: SEG Expanded Abstracts, 2794–2797.

Wave-equation migration velocity analysis for anisotropic models on 2-D ExxonMobil field data

Yunyue (Elita) Li

ABSTRACT

I apply an anisotropic wave-equation migration velocity analysis (WEMVA) method on a 2-D ExxonMobil dataset. I start from the best anisotropic model available from seismic, walkaway VSPs and check shots. The initial gathers in the subsurface offset domain are already fairly well focused; however, after inversion using our WEMVA method, I obtain an image with better continuity and clearer faults. By looking at the angle gathers, it is easy to distinguish the contribution of the improved η model from the contribution of the velocity model.

INTRODUCTION

Compared with isotropic models, anisotropic models are recognized as more realistic representations of the subsurface in complex geological environment. These models are better defined and verified in developing and production areas. However, anisotropic parameters cannot be fully constrained by surface seismic data. Therefore, anisotropic models are better defined in well-developed and production areas, where various well log, check shots and VSP data are available. The ExxonMobil dataset used in this study is from a producing area in an offshore sedimentary basin close to a huge salt body. Figure 1 illustrates the available datasets in this area. While vertical check shots and sonic logs best constrain the vertical velocity model, offset check shot (OCS) data, thanks to its long offset give information on both vertical and horizontal velocity . The current anisotropic model is built using an interactive visualization method (Bear et al., 2005) that integrates surface seismic, sonic logs, vertical check shots, and offset check shots.

From the results shown in Figure 2, we can see that the anisotropic prestack depth migration (APSDM) gathers are almost flat, inverted velocities have a reasonable match with the sonic logs, and the modeled traveltime agrees with the offset check-shot measurements. However, this interactive visualization method requires human inputs and corrections, which can be cumbersome when large-scale 3-D field data is under examination. Moreover, the most informative data - offset checkshot data - are only acquired sparsely in the 3-D domain, primarily around the salt body. Hence interpolation / extrapolation from these OCS locations is still needed to build a 3-D model elsewhere. This process may lead to inaccurate anisotropic models. Finally, according to the color code in Figure 1, although most of the traveltimes are fitted very well for near- to mid-offset check shots (indicated by rays in white), the travel times modeled between long-offset shots and the downhole receivers are still underestimated compared with the measured travel times (indicated by rays in green). Therefore, a fully automated anisotropic model building method utilizing all types of data would be highly valuable to improve the current model.

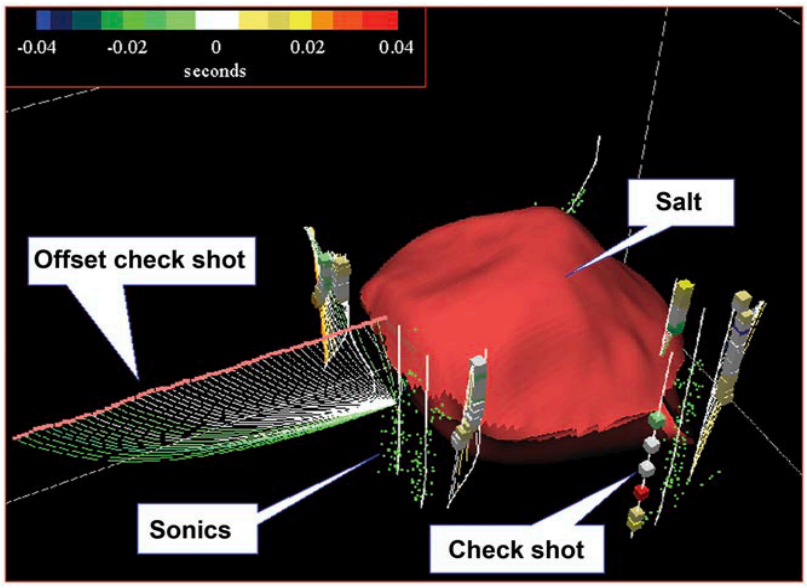


Figure 1: Available datasets for anisotropic model building. From Bear et al. (2005). The current anisotropic model is built using an interactive visualization method that integrates surface seismic, sonic logs, vertical check shots, and offset check shots. The green ray paths show a slower estimation of the velocity and/or a smaller estimation of η. [**NR**] elita2/. alldata

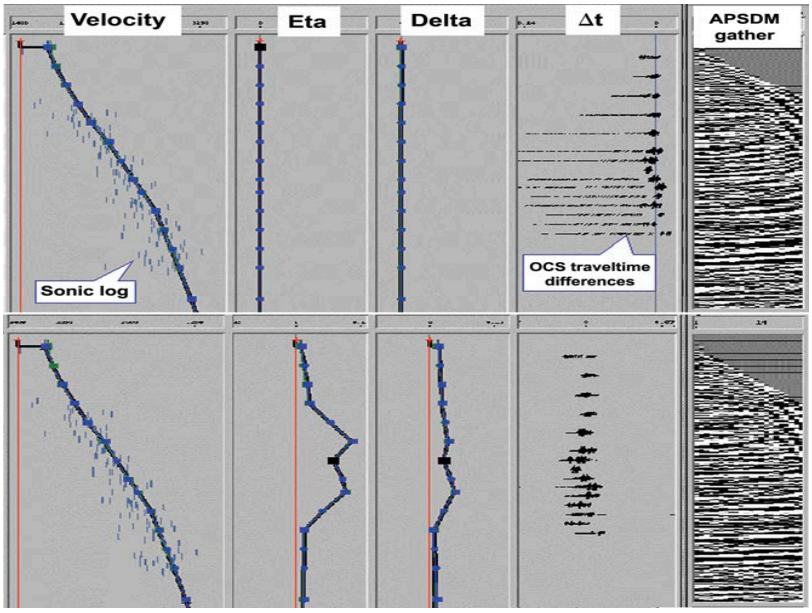


Figure 2: Visualization of interactive modeling for anisotropy. Top row: Initial isotropic model compared with sonic log and OCS traveltime. APSDM gather is already fairly flat in the small angles. Bottom row: Improved anisotropic model compared with sonic log and OCS traveltime. APSDM gathers are further flattened for large angles. From Bear et al. (2005). [**NR**] elita2/. interact

To test our method of anisotropic WEMVA (Li and Biondi, 2011), we take a 2-D slice through the sediment basin from this 3-D field dataset where the salt body is far away. The inversion is initialized by the current best model. After 8 iterations, we obtain updated velocity and η model that better define the faults in this section and have better continuities in some reflectors.

THEORY

Unlike in the previous paper (Li and Biondi, 2011), we derive the anisotropic WEMVA gradients using a Lagrangian augmented functional instead of perturbation theory. The derivation is hence neat and straightforward, and the interpretations of the adjoint-state equations suggest the same implementation as perturbation theory suggests.

We parameterize the VTI subsurface using NMO slowness s_n, and Thomson parameters η and δ (Thomsen, 1986). In the shot-profile domain, both source wavefields $\mathbf{D}$ and receiver wavefields $\mathbf{U}$ are downward continued using the following one-way wave equation and boundary condition (Shan, 2009):

$$\begin{cases} \left(\frac{1}{\sqrt{1+2\delta}}\frac{\partial}{\partial z} + i\Lambda\right) D(\mathbf{x},\mathbf{x}_s) = 0 \\ D(x,y,z=0,\mathbf{x}_s) = f_s\delta(\mathbf{x}-\mathbf{x}_s) \end{cases}, \tag{1}$$

and

$$\begin{cases} \left(\frac{1}{\sqrt{1+2\delta}}\frac{\partial}{\partial z} - i\Lambda\right) U(\mathbf{x},\mathbf{x}_s) = 0 \\ U(x,y,z=0,\mathbf{x}_s) = Q(x,y,z=0,\mathbf{x}_s) \end{cases}, \tag{2}$$

where

$$\Lambda = \omega s_n \sqrt{1 - \frac{|\mathbf{k}|^2}{\omega^2 s_n^2 - 2\eta|\mathbf{k}|^2}}. \tag{3}$$

Equations 1 and 2 can be summarized in matrix forms as follows:

$$\mathbf{L}\mathbf{D} = \mathbf{f} \tag{4}$$

and

$$\mathbf{L}^*\mathbf{U} = \mathbf{d}, \tag{5}$$

where

$$\mathbf{f} = f_s\delta(\mathbf{x}-\mathbf{x}_s, z), \tag{6}$$

$$\mathbf{d} = d_{\mathbf{x}_s}\delta(\mathbf{x}-\mathbf{x}_r, z), \tag{7}$$

and

$$\mathbf{L} = \frac{1}{\sqrt{1+2\delta}}\frac{\partial}{\partial z} - i\Lambda. \tag{8}$$

It is well known that parameter δ is the least constrained by surface seismic data because of the lack of depth information. Therefore, we assume δ can be correctly obtained from other sources of information, such as check shots and well logs. In this paper, we are going to invert for NMO slowness s_n and η.

We use an extended imaging condition (Sava and Formel, 2006) to compute the image cube with subsurface offsets:

$$I_{\mathbf{h}} = (\mathbf{S}_{+\mathbf{h}}\mathbf{p})^*(\mathbf{S}_{-\mathbf{h}}\mathbf{q}), \tag{9}$$

where $\mathbf{S}_{+\mathbf{h}}$ is a shifting operator which shifts the wavefield $+\mathbf{h}$ in the $\mathbf{x}$ direction. Notice that $(\mathbf{S}_{+\mathbf{h}})^* = \mathbf{S}_{-\mathbf{h}}$. Equations 4, 5 and 9 are state equations, and $\mathbf{D}$, $\mathbf{U}$ and $I_{\mathbf{h}}$ are the state variables.

To evaluate the accuracy of the subsurface model, we use a DSO objective function (Shen, 2004):

$$J = \frac{1}{2}\sum_{\mathbf{h}} \langle \mathbf{h} I_{\mathbf{h}}, \mathbf{h} I_{\mathbf{h}} \rangle. \tag{10}$$

where $\mathbf{h}$ is the subsurface offset. In practice, other objective functions (linear transformations of the image) can be used rather than DSO. To derive the DSO objective function with respect to s_n and η, we follow the recipe provided by Plessix (2006). First, we form the Lagrangian augmented functional:

$$\begin{aligned} &\mathcal{L}(\mathbf{D}, \mathbf{U}, I_{\mathbf{h}}; \lambda, \mu, \gamma_{\mathbf{h}}; s_n, \epsilon) = \qquad (11)\\ &\quad \frac{1}{2}\sum_{\mathbf{h}} \langle \mathbf{h}\ I_{\mathbf{h}},\ \mathbf{h}\ I_{\mathbf{h}} \rangle \\ &\quad + \langle \lambda,\ \mathbf{f} - \mathbf{L}(\mathbf{s_n}, \epsilon)\mathbf{D} \rangle \\ &\quad + \langle \mu,\ \mathbf{d} - \mathbf{L}^*(\mathbf{s_n}, \epsilon)\mathbf{U} \rangle \\ &\quad + \sum_{\mathbf{h}} \langle \gamma_{\mathbf{h}},\ (\mathbf{S}_{+\mathbf{h}}\mathbf{D})^*(\mathbf{S}_{-\mathbf{h}}\mathbf{U}) - I_{\mathbf{h}} \rangle . \qquad (12) \end{aligned}$$

Then the adjoint-state equations are obtained by taking the derivative of $\mathcal{L}$ with respect to state variables $\mathbf{D}$, $\mathbf{U}$ and $I_{\mathbf{h}}$:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{D}} = -\mathbf{L}^*(s_n, \epsilon)\lambda + \sum_{\mathbf{h}} (\mathbf{S}_{+\mathbf{h}})^*(\mathbf{S}_{-\mathbf{h}}\mathbf{U})\gamma_{\mathbf{h}} = \mathbf{0}, \tag{13}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{U}} = -\mathbf{L}(s_n, \epsilon)\mu + \sum_{\mathbf{h}} (\mathbf{S}_{-\mathbf{h}})^*(\mathbf{S}_{+\mathbf{h}}\mathbf{D})\gamma_{\mathbf{h}} = \mathbf{0}, \tag{14}$$

$$\frac{\partial \mathcal{L}}{\partial I_{\mathbf{h}}} = -\gamma_{\mathbf{h}} + \mathbf{h}^2 I_{\mathbf{h}} = 0, \forall\ \mathbf{h}. \tag{15}$$

Equation 13, 14, and 15 are the adjoint-state equations. Parameters λ, μ and $\gamma_{\mathbf{h}}$ are the adjoint-state variables, and can be calculated from the adjoint-state equations.

The physical interpretation of the adjoint-state equations offers better understanding of the physical process and provides insights for implementation. Clearly, the solution to equation 15, $\gamma_{\mathbf{h}}$, is the perturbed (residual) image at a certain subsurface offset. Equations 13 and 14 define the perturbed source and receiver wavefields, respectively. Notice the perturbed source wavefield λ at location $\mathbf{x}$ depends on the image at $(\mathbf{x} - \mathbf{h}, \mathbf{h})$ and the background receiver wavefield $\mathbf{U}$ at $\mathbf{x} - 2\mathbf{h}$. The same rule applies to the perturbed receiver wavefield μ.

With the solutions to the equations above, we can now derive the gradients of the objective function 10 by taking the derivative of the augmented functional $\mathcal{L}$ with respect

to the model variables s_n and ϵ as follows:

$$\nabla_{s_n} J = \left\langle \lambda, \ -\frac{\partial \mathbf{L}}{\partial \mathbf{s_n}} \mathbf{D} \right\rangle + \left\langle \mu, \ -\frac{\partial \mathbf{L}^*}{\partial \mathbf{s_n}} \mathbf{U} \right\rangle \tag{16}$$

$$\nabla_{\epsilon} J = \left\langle \lambda, \ -\frac{\partial \mathbf{L}}{\partial \epsilon} \mathbf{D} \right\rangle + \left\langle \mu, \ -\frac{\partial \mathbf{L}^*}{\partial \epsilon} \mathbf{U} \right\rangle . \tag{17}$$

It is straightforward to understand that the gradients for the model parameters are the crosscorrelations of the perturbed source / receiver wavefield and the scattered background receiver / source wavefield.

FIELD DATA TESTS

We extract a 2-D line from ExxonMobil dataset where the salt body is far away. Source spacing is 100 m, and receiver spacing is 50 m. The maximum offset in this 2-D line is about 4 km. This surface seismic dataset was acquired more than a decade ago; therefore the limited offsets of this dataset are insufficient from enough to constrain the anisotropic parameters. Fortunately, this area has been studied extensively by various types of other acquisitions: vertical check shots, offset check shots, and sonic logs. Therefore, a very well-defined anisotropic model is obtained from previous studies. We migrated the 2-D line using the anisotropic model provided (Figure 4(a) and Figure 4(b)). The initial stack image is shown on the top panel in Figure 6. Although the apparent dip is high due to the large vertical stretch (3 km in depth vs. 16 km on the horizontal axis), the reflectors in this 2-D line are pretty flat. We then estimate the dip field (Figure 3) from the initial image and use it to regularize the gradient for both velocity and η.

We can see many small-scale faults in this area on the top panel in Figure 6. Migration artifacts at $x = 13$ km and $z = 2500$ m are caused by a big vertical fault running from $x = 14$ km on the top to the bottom of the section. The initial angle gathers are shown in the bottom row in Figure 6. Since this is a streamer geometry, the subsurface reflectors are only illuminated from positive angles. Although the gathers are almost flat, we can still see upward residual moveouts in the angle domain. Therefore, we have a chance to improve the model and the image by flattening the gathers.

Updates between the initial and the inverted velocity and η models are shown in Figure 5(a) and Figure 5(b), respectively. First, notice the spatial correlation between velocity updates and η updates, although no such constraints are applied during the inversion. However collocated, the update directions in velocity and η are not necessarily the same. We are able to resolve a localized shallow anomaly between 13 km and 15 km at around 800 m below surface. Comparing the initial stack image on the top panel in Figure 6 with the final stack image on the top panel of Figure 7, we can see improved continuity and signal strength in the area highlighted by the oval. The fault in this area is also better defined in the final image. Notice that the updates in velocity are less than 10%, whereas the updates in η are around 25%. These positive updates in both velocity and η agree well with the negative travel time misfits in the previous OCS modeling results (Figure 1).

We can also verify the updates in velocity and η on the angle gathers at different CMP locations. The initial angle domain common image gathers (ADCIGs) are shown in the

bottom row in Figure 6, and the final ADCIGs produced using the inverted models are shown in the bottom row in Figure 7. To better illustrate the effects of the model updates, the ADCIGs are sampled more densely between CMP = 13 and 16 and sparsely outside of this range. In general, we can see improved flatness for all the reflectors. Specifically, for the shallower events above 1 km, most improvements happen at large angles over 35°. Therefore, we interpret the improvements for the shallow events primarily as the contribution of the improved η model.

For the deeper events at the same CMP location, both the depth and the flatness of the angle gather have been changed by inversion. The upward-curving events in the angle domain from the initial migration has been flattened by the improved velocity and η model. This result would be more convincing if we had the corresponding well logs at the same location to verify the depth shifts.

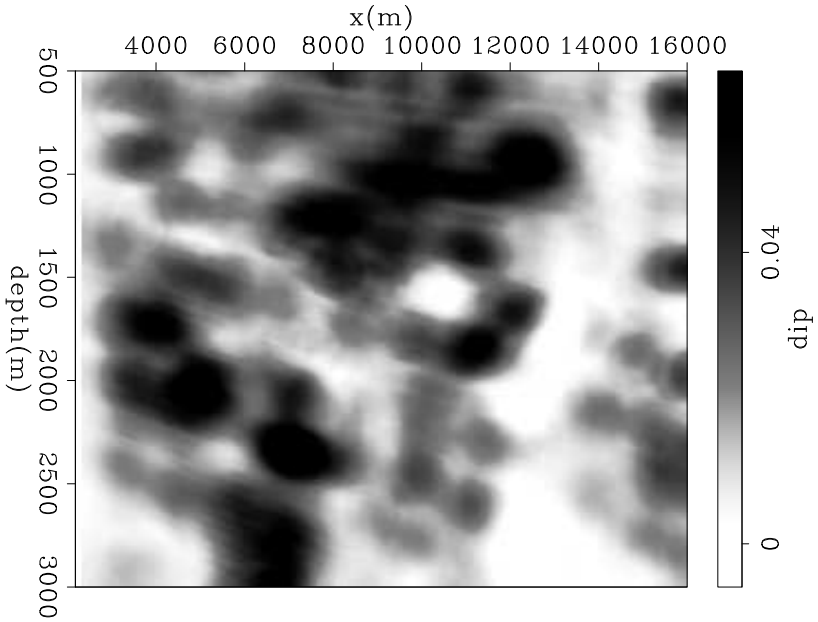
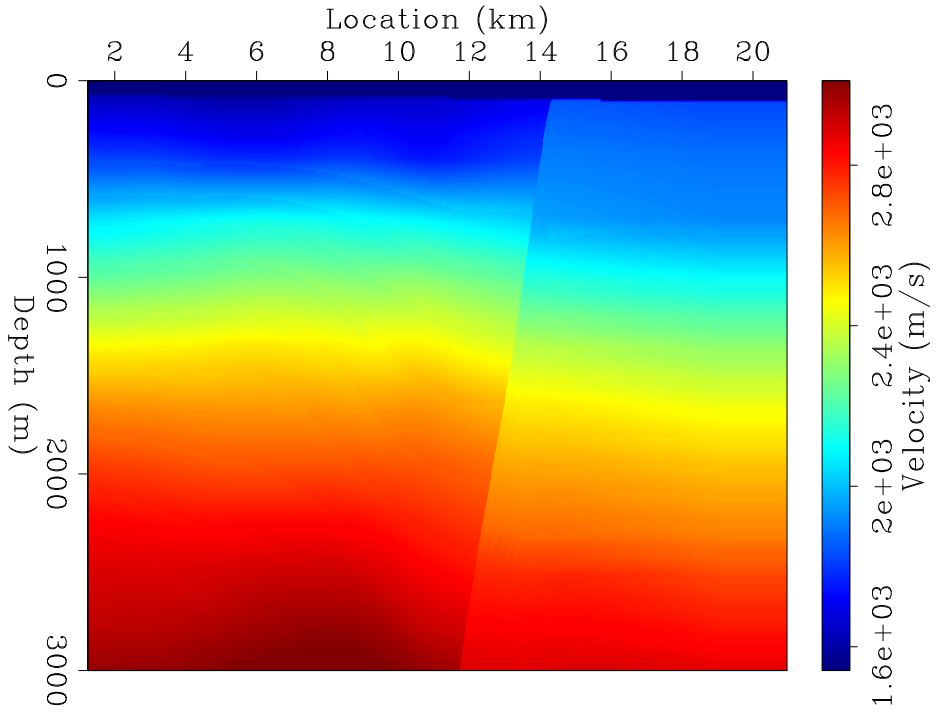


Figure 3: Estimated dip field from the initial image on the top panel of Figure 6. [**CR**] elita2/. dip-exxon

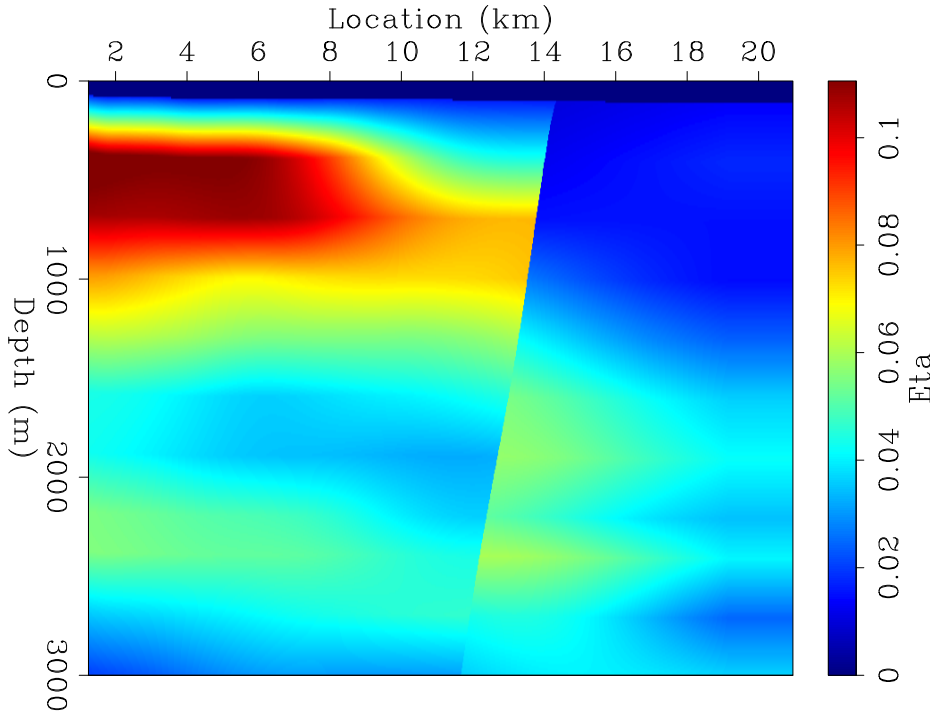
CONCLUSIONS AND FUTURE WORK

In this paper, we apply our anisotropic WEMVA algorithm to a 2-D slice of an ExxonMobil field dataset. The previous anisotropic model was built using all available datasets in the area: surface seismic, vertical and offset check shots and sonic logs. From the initial migrated image using the initial model, we observe upward moveout at large angles in the angle CIGs, which is consistent with the negative travel time misfits in the large-offset check shots. After 8 iterations of our non-linear anisotropic WEMVA iterations, we obtain a better subsurface image with flatter angle gathers and higher angle coverage.

The provided anisotropic model shows high spatial correlation between vertical velocity and Thomson parameters. Figure 8 shows a cross-plot among the parameters v_v, δ and η down to 1 km. The scatter plot shows distinct differences on different sides of the fault. As-

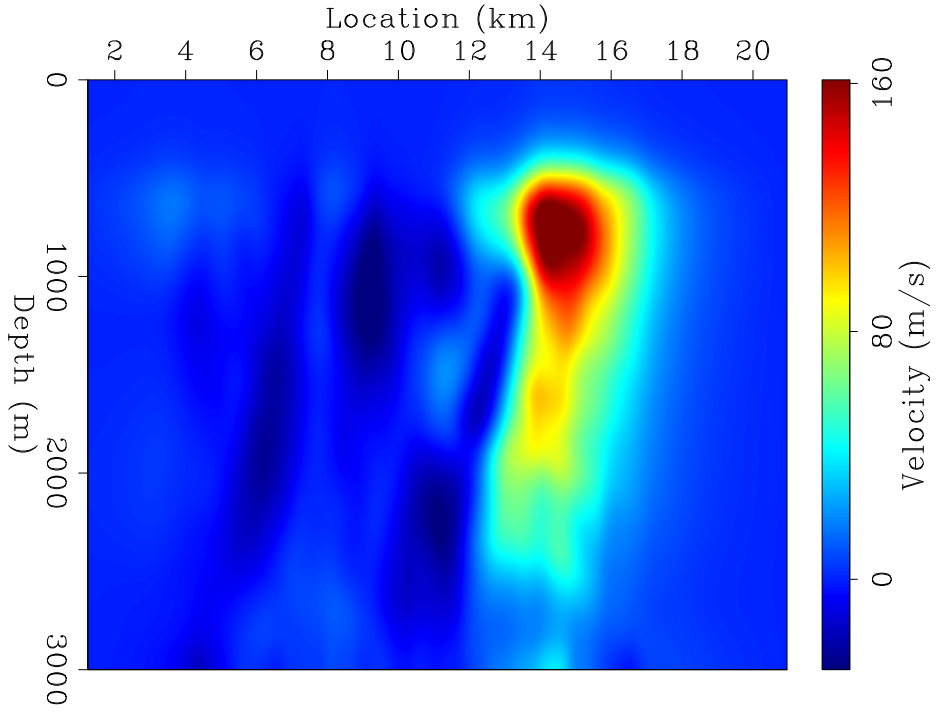


(a)

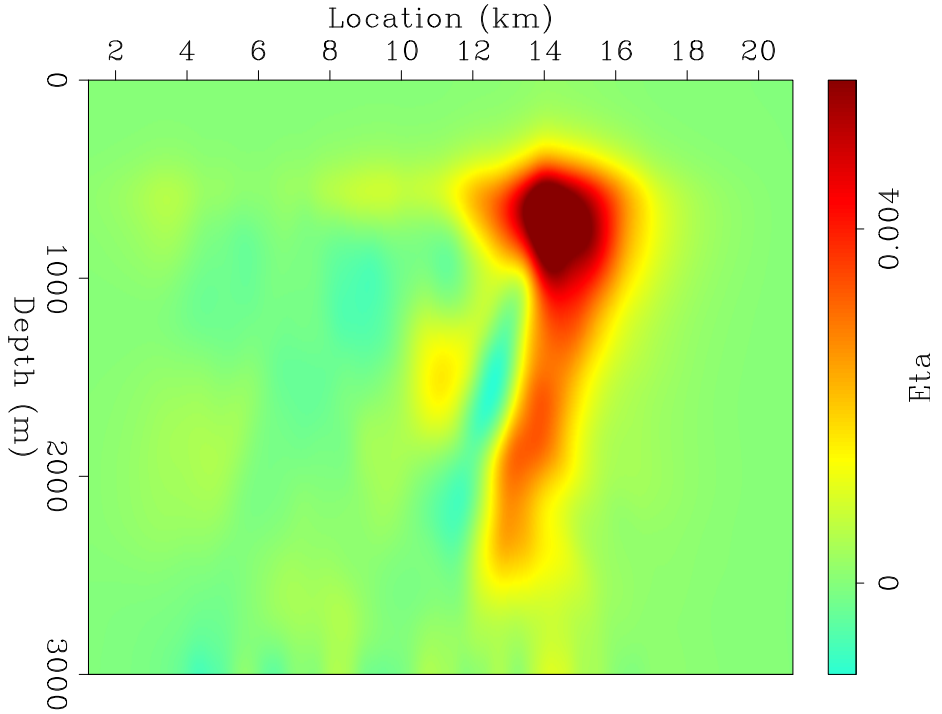


(b)

Figure 4: Initial velocity model (a) and initial η model (b). [**CR**] elita2/. initv,inite



(a)



(b)

Figure 5: Updates in velocity model (a) and updates in η model (b) after inversion. [**CR**] elita2/. vchange,echange

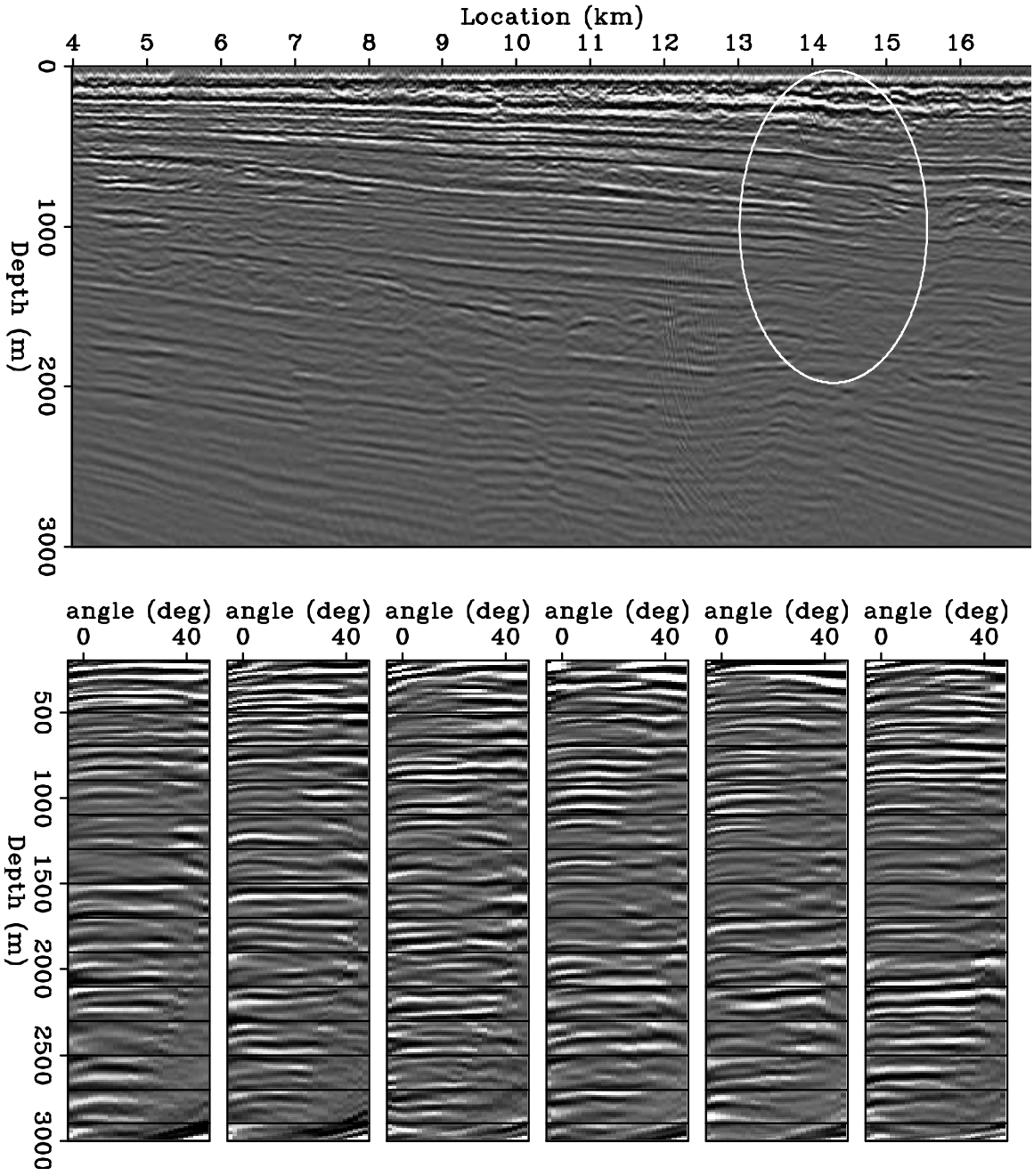


Figure 6: The initial stack image (Top panel) and initial angle domain common image gathers at CMP $= 7, 10, 13, 14, 14.5, 15$ km (Bottom row). [**CR**] elita2/. image-init-an

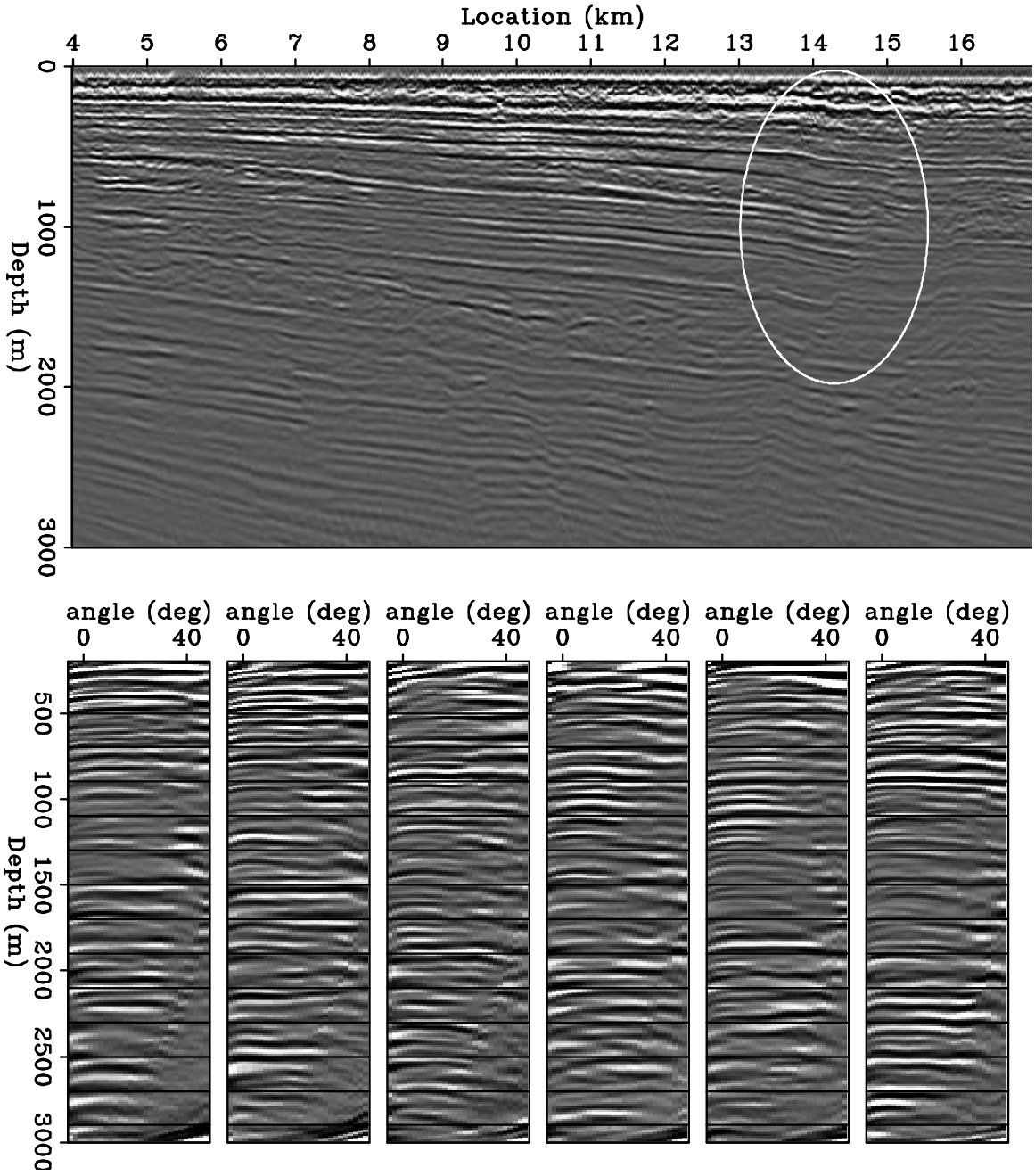


Figure 7: The final stack image (Top panel) and final angle domain common image gathers at CMP $= 7, 10, 13, 14, 14.5, 15$ km (Bottom row). Compared with Figure 6, improvements in continuity and enhancements in amplitude strength are highlighted by the oval. [**CR**] elita2/. image-fnal-an

suming these differences are based on the local geology and lithology, we should include this information into our model-building process in future work, possibly as a preconditioning constraint suggested by Li et al. (2011).

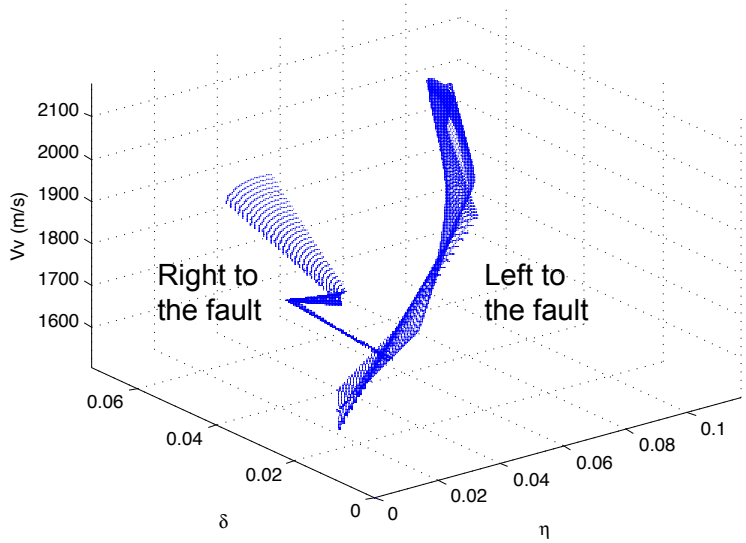


Figure 8: 3D scatter plot for the anisotropic parameters v_v, δ and η. [**CR**] elita2/. 3dexxon

ACKNOWLEDGEMENT

The author thanks ExxonMobil for the permission to publish the field data example.

REFERENCES

Bear, L., T. Dickens, J. Krebs, J. Liu, and P. Traynin, 2005, Integrated velocity model estimation for improved positioning with anisotropic PSDM: The Leading Edge, 622–634.

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.

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

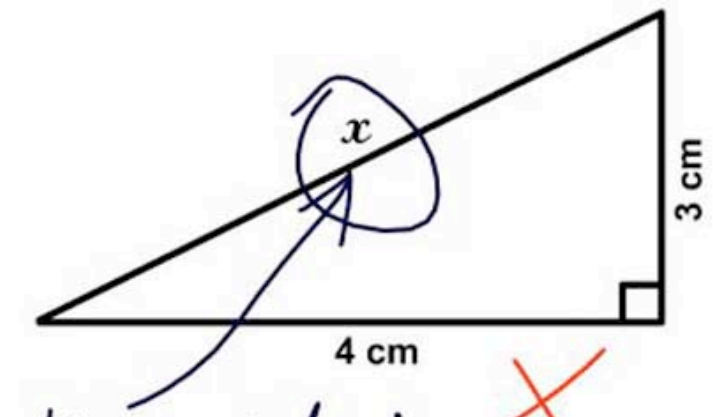
Sava, P. and S. Formel, 2006, Generalized imaging conditions for wave equation migration: CWP Report, **524**.

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

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

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

3. Find x.



Ocular Trauma - by Wade Clarke ©2005

PETER

1.21

4b) Expand

$x^3 + 8x - 2$

$(a+b)^n$ Very funny, Peter.

$= (a + b)^n$

$= (a + b)^n$

$= (a + b)^n$

2 ?

etc—

Wave-equation migration velocity analysis for VTI media using optimized implicit finite difference

Yunyue (Elita) Li

ABSTRACT

Anisotropic wave-equation migration velocity analysis (WEMVA) requires fast and accurate wave modeling at all angles. I use an optimized implicit finite difference one-way propagation engine to improve both the efficiency and accuracy of this process. In this implicit finite difference scheme, anisotropic parameters η and δ are mapped into two finite difference coefficients, α and β. When computing the perturbed wavefields from model perturbation, I apply a chain rule to link the wave equation with the actual anisotropic parameters via the finite difference coefficients. I test the implementation by impulse responses in both 2D and 3D. The sensitivity kernels for wave-equation reflection tomography confirm the theoretical understanding that waves have a higher sensitivity for η at large angles and a higher sensitivity for vertical velocity at small angles.

INTRODUCTION

Building an anisotropic model requires accurate wave field descriptions at all angles, both in simple and complex geological settings. Therefore, instead of using travel times and ray paths, we use the wave fields as the carrier of the model information (Li and Biondi, 2011a; Li et al., 2012; Li, 2012). Many choices for wavefield propagators may be considered. On one hand, the one-way wave propagators (Li, 2012) excel in speed but lose their accuracy rapidly with the increasing propagation angle. On the other hand, the two-way wave propagators (Li et al., 2012) are more accurate in modeling waves at large angles but their computational costs are less affordable.

Therefore, I use an optimized implicit finite difference propagator first developed by Shan (2006). In this optimized implicit finite difference scheme, the anisotropic parameters η and δ contribute to the wave-equation implicitly via two finite difference parameters α and β. Tables of α and β with respect to sets of discrete η and δ values are precomputed before propagation by minimizing the difference between the true dispersion relationship and its rational series approximation at different wavenumbers. This table-driven, implicit finite difference method handles lateral variations and is accurate up to 60° in a vertical transverse isotropic (VTI) medium.

When perturbing the wave-equation around its current state, the finite difference coefficients α and β are perturbed. These perturbations are then translated into the anisotropic parameters η and δ using the chain rule. Tables of numerical derivatives of the finite difference coefficients α and β with respect to the anisotropic parameter η are also precomputed from the previous coefficients tables.

Finally, I test this implicit finite difference implementation by 2D and 3D impulse responses for vertical velocity v_v and η. The results verify the theoretical understanding of the WEMVA operator for anisotropic models.

OPTIMIZED IMPLICIT FINITE DIFFERENCE FOR VTI MEDIA

Assuming the S-wave velocity is much slower than the P-wave velocity, we can approximate the dispersion relationship for VTI media as follows (Shan, 2009):

$$S_z = \sqrt{\frac{1-(1+2\delta)S_r^2}{1-2\eta(2\delta+1)S_r^2}}, \tag{1}$$

where $S_z = \frac{k_z}{w/v_v}$, $S_r = \frac{k_r}{w/v_v}$ and $k_r = \sqrt{k_x^2+k_y^2}$. Anisotropic parameter δ relates the vertical P-wave velocity v_v with the NMO velocity v_n, while the anellipticity parameter η relates the horizontal velocity v_h with the NMO velocity v_n. Shan (2009) suggests that the exact dispersion relationship 1 can be approximated by a rational function $R_{n,m}(S_r)$:

$$R_{n,m}(S_r) = \frac{P_n(S_r)}{Q_m(S_r)}, \tag{2}$$

where

$$P_n(S_r) = \sum_{i=0}^{n} a_i S_r^i \tag{3}$$

and

$$Q_m(S_r) = \sum_{i=0}^{m} b_i S_r^i. \tag{4}$$

Moreover, when the polynomials in equations 3 and 4 are of the same degree, namely $m = n$, dispersion relationship 2 can be further split as follows:

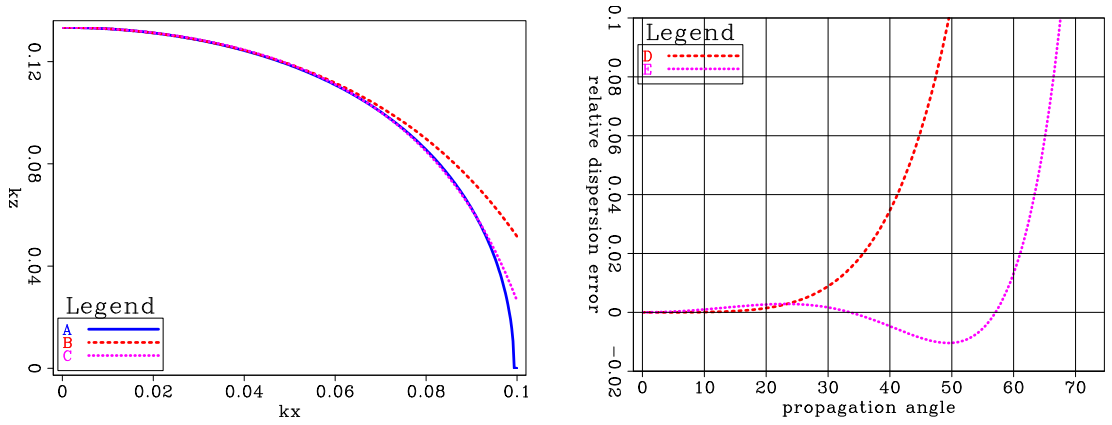
$$S_z = 1 - \sum_{i=1}^{n} \frac{\alpha_i S_r^2}{1-\beta_i S_r^2}. \tag{5}$$

The coefficients α_i and β_i can be obtained by solving the least-square problem below:

$$\min \sum_{S_r} \left(\sqrt{\frac{1-(1+2\delta)S_r^2}{1-2\eta(2\delta+1)S_r^2}} - \left(1 - \sum_{i=1}^{n} \frac{\alpha_i S_r^2}{1-\beta_i S_r^2}\right)\right)^2. \tag{6}$$

For second-order coefficients the exact and approximated dispersion curves are shown in Figure 1(a), given $\eta = 0.14$ and $\delta = 0.2$. Curve A is the exact dispersion relation from Equation 1. Curve B is obtained from a previous estimation by Ristow and Ruhl (1997), and curve C is obtained using the optimized coefficients. Apparently, the dispersion relation using the optimized coefficients is a better approximation compared with the previous method which uses Taylor expansion and assumes weak anisotropy. The relative errors between these two approximated curves and the exact dispersion curve are plotted in Figure 1(b). Within a tolerance of 1% relative error in the dispersion relation, the optimized dispersion is accurate up to 60°, while the Taylor approximation is only accurate up to 30°.

The tables for coefficients α and β for η ranging from 0 to 0.15 and δ ranging from -0.004 to 0.2 are shown in Figure 2. In general, parameter α is more sensitive to the change in δ than to the change in η. Parameter β has similar sensitivities to both η and δ.



(a) (b)

Figure 1: (a) Dispersion relation curves: A, exact dispersion relation curve from equation 1; B, approximated dispersion curve from weak anisotropy and Taylor expansion; C, approximated dispersion curve from optimization. (b) Relative dispersion error: D, relative error between B and A; E, relative error between C and A. [**ER**] elita3/. kz1,err1

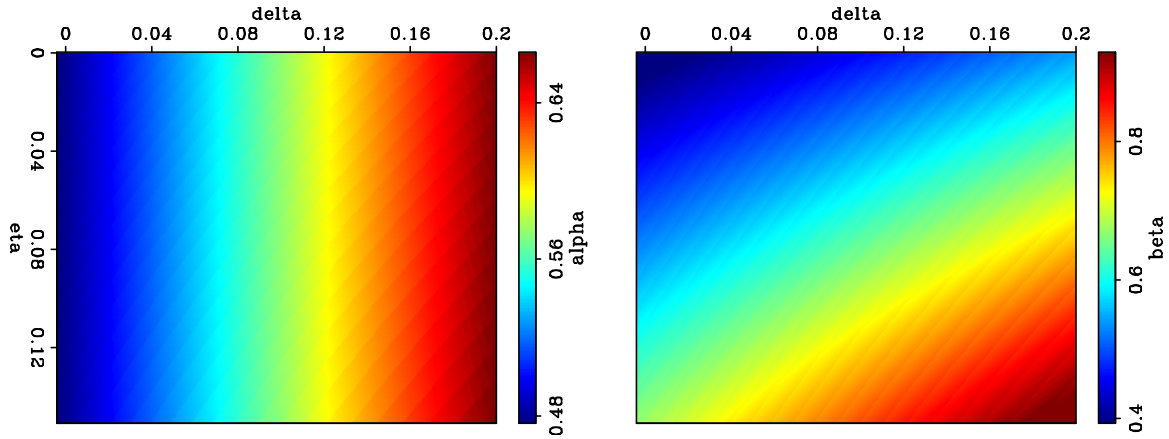


Figure 2: (a) Table for α and (b) table for β at discrete η and δ locations. [**ER**] elita3/. coef

IMPULSE RESPONSE OF THE IMAGE-SPACE WAVE-EQUATION TOMOGRAPHY OPERATOR

The core of anisotropic image-space wave-equation migration velocity analysis is the tomography operator $\mathbf{T}$, which relates the perturbation in the anisotropic models ($\Delta\mathbf{m}$) to to the perturbation in the image (ΔI) and vice versa. Namely,

$$\Delta I = \mathbf{T}\Delta\mathbf{m}, \tag{7}$$
$$\Delta\mathbf{m} = \mathbf{T}'\Delta I, \tag{8}$$

where $\mathbf{m} = [v_v\ \eta]$.

I refer the readers to Li and Biondi (2010) for a detailed derivation for the tomographic operator.

A different approximation to the exact dispersion relation leads to a different perturbed wave fields due to a perturbation in the model parameters. When the only available data come from surface seismic surveys, parameter δ is the least constrained (Plessix and Rynja, 2010; Li and Biondi, 2011b). Therefore, I assume the δ model is perfectly obtained from other sources of data and keep it fixed throughout the inversion. I will invert for v_v and η in this study.

In the downward extrapolation, the wavefield at the next depth (P_{z+1}) can be computed from the wavefield at the current depth (P_z) according to the following equation:

$$P_{z+1} = P_z e^{ik_z dz}, \tag{9}$$

where $i = \sqrt{-1}$, dz is the extrapolation distance in depth and k_z can be obtained from the first-order approximation of the dispersion relation 5:

$$k_z = \frac{w}{v_v}\left(1 - \frac{\alpha\frac{k_r^2}{(w/v_v)^2}}{1 - \beta\frac{k_r^2}{(w/v_v)^2}}\right). \tag{10}$$

Dispersion relation 10 can be further simplified to polynomials using Taylor expansion:

$$\begin{aligned} k_z &= \frac{w}{v_v}\left(1 - \alpha\frac{k_r^2}{(w/v_v)^2}\left(1 + \beta\frac{k_r^2}{(w/v_v)^2}\right)\right) \\ &= \frac{w}{v_v}\left(1 - \alpha\frac{k_r^2}{(w/v_v)^2} - \alpha\beta\frac{k_r^4}{(w/v_v)^4}\right). \end{aligned} \tag{11}$$

Therefore, the perturbed wavefield is

$$\Delta P_{z+1} = e^{ik_z dz} i dz P_z \Delta k_z, \tag{12}$$

with

$$\Delta k_z = \frac{\partial k_z}{\partial v_v}\Delta v_v + \frac{\partial k_z}{\partial \eta}\Delta\eta, \tag{13}$$

$$\frac{\partial k_z}{\partial v_v} = -\frac{w}{v_v^2}\left(1 + \alpha\frac{k_r^2}{(w/v_v)^2} + 3\alpha\beta\frac{k_r^4}{(w/v_v)^4}\right), \tag{14}$$

and

$$\frac{\partial k_z}{\partial \eta} = -\frac{w}{v_v}\left(\frac{\partial\alpha}{\partial\eta}\frac{k_r^2}{(w/v_v)^2} + \left(\frac{\partial\alpha}{\partial\eta}\beta + \alpha\frac{\partial\beta}{\partial\eta}\right)\frac{k_r^4}{(w/v_v)^4}\right). \tag{15}$$

Since the finite difference parameters α and β are obtained by optimization, the derivatives in Equation 15 are obtained numerically by taking derivatives along the η axis in Figure 2. The tables of the derivatives of the coefficients with respect to η are shown in Figure 3.

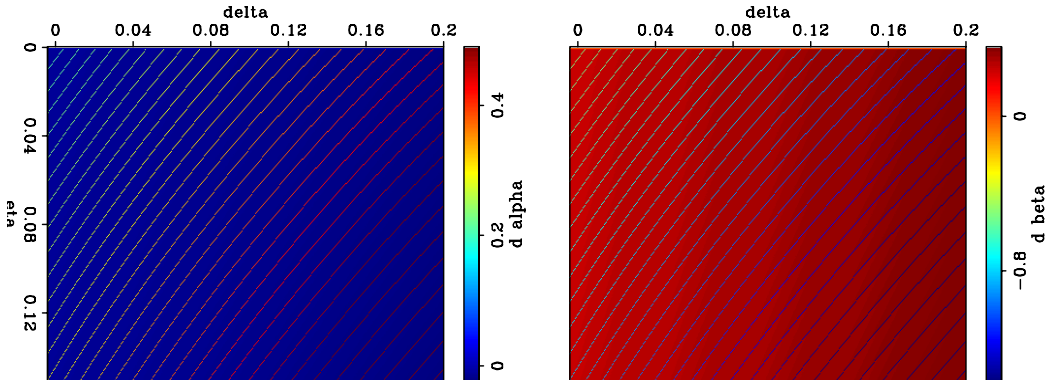


Figure 3: (a) Table for $\frac{\partial \alpha}{\partial \eta}$ and (b) table for $\frac{\partial \beta}{\partial \eta}$ at background η and δ locations. [**ER**] elita3/. dr-coef

I test the implementation of the adjoint tomographic operator using this optimized implicit finite difference scheme in a homogeneous background VTI medium with $v_v = 2000$ km, $\eta = 0.09$ and $\delta = 0.05$. The synthetic data is produced by Born modeling with a horizontal reflector at the depth of 1500 km. The input of the adjoint tomographic operator is a spike in the image space $\Delta I = \delta(x, y, z = 1500)$. The dominant frequency of the source wavelet is 20 Hz, and the samplings in all directions are 10 m.

I first test the adjoint operator in 2D. A source and receiver pair is collocated at $x = y = z = 0$. The top row in Figure 4 shows the back-projected vertical velocity v_v gradient and η gradient when source-receiver offset is zero. These back projections are often referred as banana-donut kernels in the literature when transmission waves are under study (eg. Marquering et al. (1998, 1999); Rickett (2000)). Similar reflection tomography sensitivity kernel analysis for isotropic WEMVA operator can be found in Sava (2004) and Xie and Yang (2009).

Compared with the η gradient, the v_v gradient has a nearly uniform strength with depth, while the η gradient fades away as the wavepath moves away from the source and the receiver location. Also, the dominant energy of the η gradient points to the opposite direction of the v_v gradient points. In fact, the η gradient is not reliable and should be ignored because the wave that travels in the vertical direction is not sensitive to η.

When the source-receiver offset is 4 km, the gradients are shown in the middle row in Figure 4. Clearly, the back projections are spread along the wavepaths from the source to the perturbed image point and from the perturbed image point to the receiver. In this case, the gradients in both v_v and η point in the same direction. Comparing the gradients in the cases of zero and nonzero offset, one can see that the vertical waves are more sensitive to v_v, and the waves traveling at a large angle (36° to the vertical in this case) are more sensitive to η. The bottom row in Figure 4 shows the summation of the gradients in these two cases, and confirms these observations.

The 3D extension of this method is straightforward. The sensitivity kernels for v_v and η

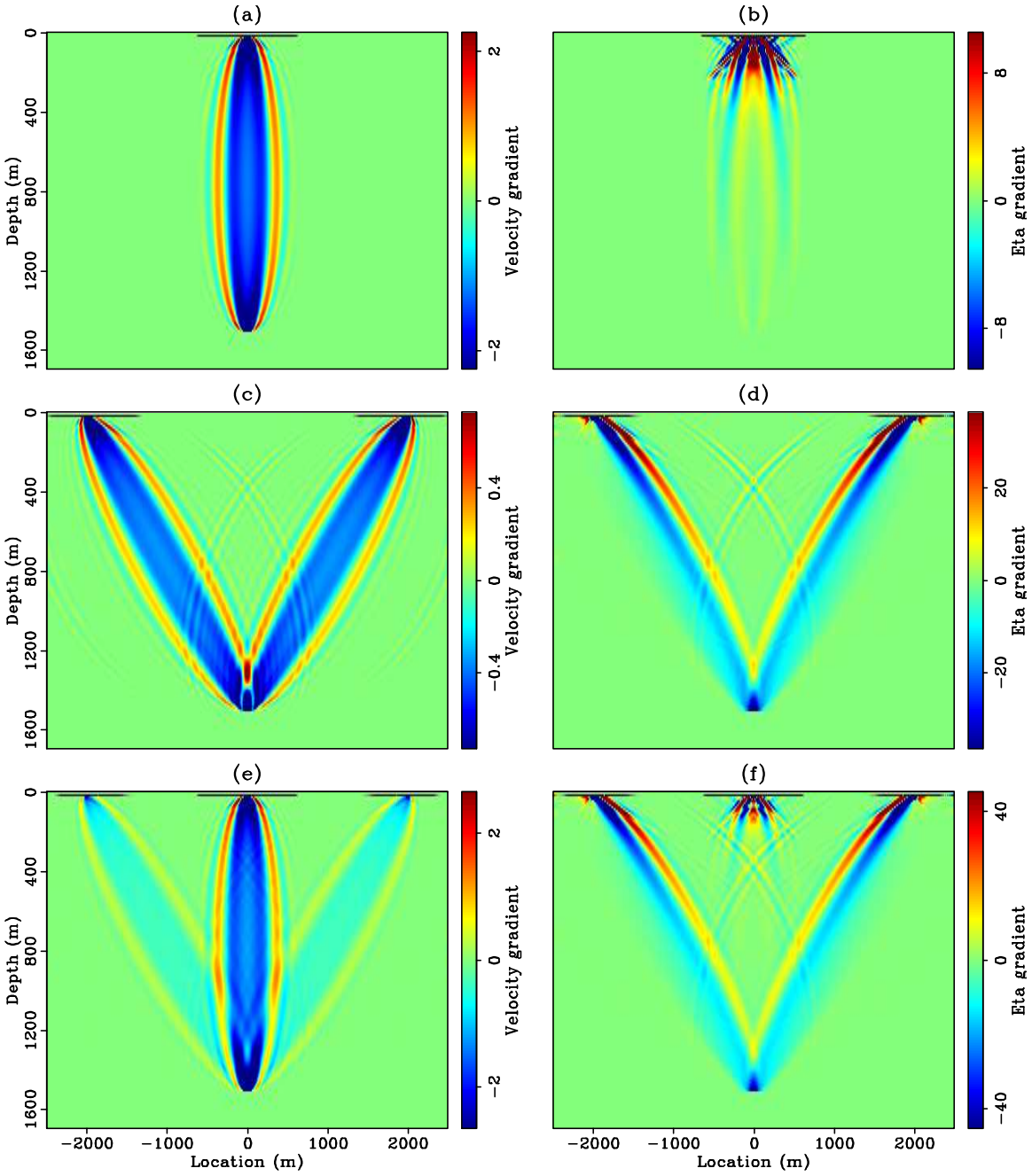


Figure 4: 2D impulse responses for vertical velocity (left column) and η (right column). Top row: zero offset impulse responses; middle row: impulse responses when source-receiver offset is 4 km; bottom row: summation of the two rows above. [**ER**] elita3/. 2dkernel

in 3D are shown in Figures 5 and 6. A source and receiver pair with 4 km offset are located at $y = 0$. The 3D sensitivity kernels carry the same characteristics as the 2D kernels, only expanding to the crossline direction.

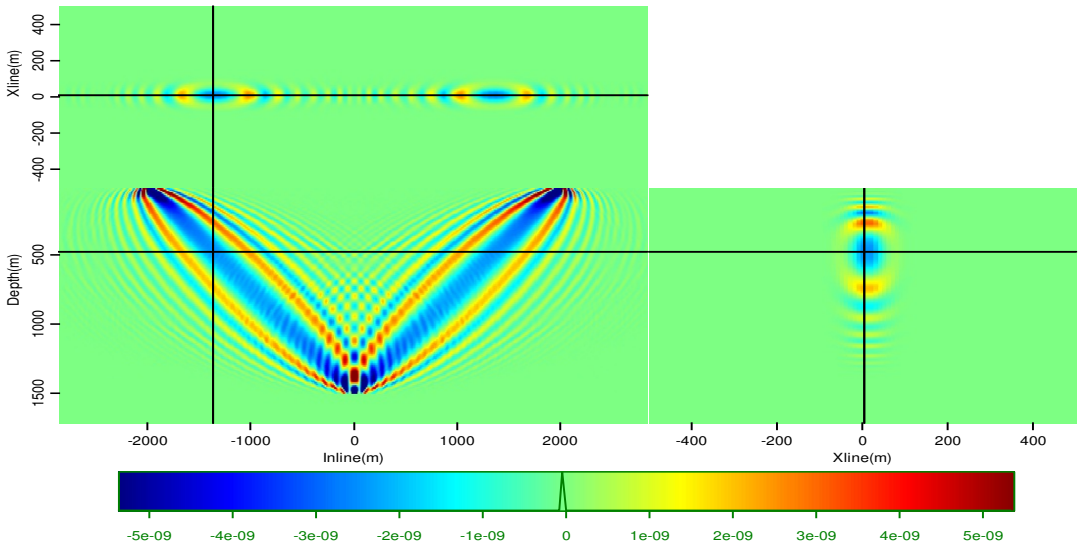


Figure 5: 3D v_v kernel. [**CR**] elita3/. 3dkernel-vel-new

CONCLUSIONS AND DISCUSSIONS

In this paper, I use an optimized implicit finite difference one-way propagation engine to improve both the efficiency and accuracy for anisotropic WEMVA. By precomputing the tables for the finite difference coefficients, the cost of this VTI extrapolation is similar to an isotropic implicit finite difference scheme. With the optimized coefficients, the dispersion relation is accurate up to 60°. To compute the disturbed wave fields due to the perturbation in the models, the numerical derivatives of the optimized coefficients with respect to η are also precomputed.

I test the VTI implicit finite difference scheme by impulse responses in both 2D and 3D. These impulse responses of the adjoint anisotropic WEMVA operator from a spike perturbation in the image space have a familiar banana-donut sensitivity kernel shape for both v_v and η. The amplitudes in these impulse responses show different sensitivities for different parameters with different source-receiver geometries. The waves traveling vertically have higher sensitivity to v_v and the waves traveling at large angles have higher sensitivity to η. Therefore, the 3D sensitivity kernels can also be used for acquisition design before exploration when specific parameters are under consideration.

REFERENCES

Li, Y. and B. Biondi, 2010, Wave-equation tomography for anisotropic parameters: SEP-Report, **140**, 49–60.

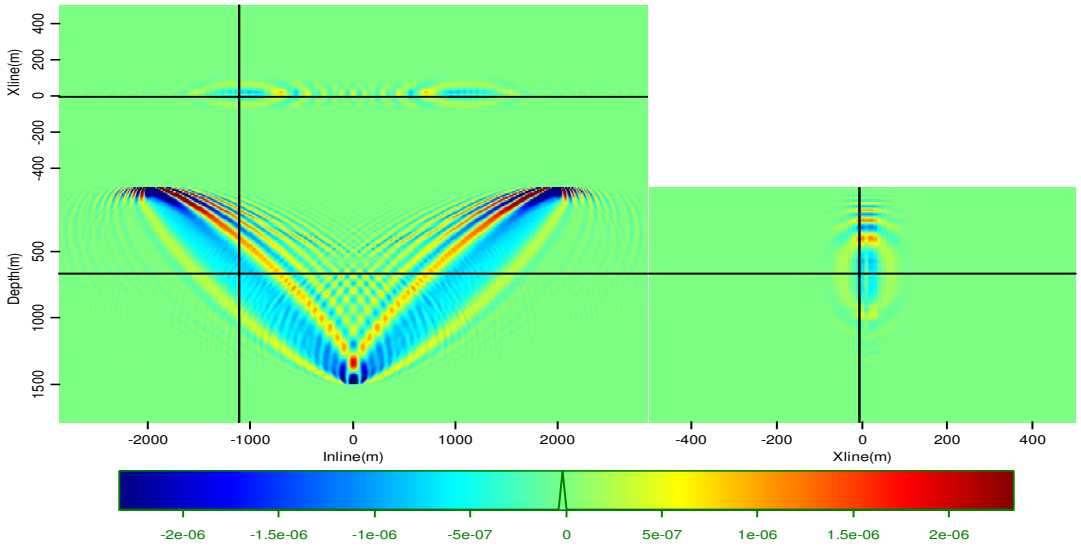


Figure 6: 3D η kernel. [**CR**] elita3/. 3dkernel-eta-new

———, 2011a, Migration velocity analysis for anisotropic models: SEG Expanded Abstract, **30**, 201–206.

Li, Y. E., 2012, Wave-equation migration velocity analysis for anisotropic models on 2-D ExxonMobil field data: SEP-Report, **147**, 83–94.

Li, Y. E. and B. Biondi, 2011b, Migration velocity analysis for anisotropic models: SEP-Report, **143**, 21–32.

Li, Y. E., P. Shen, and C. Perkins, 2012, VTI migration velocity analysis using RTM: SEP-Report, **147**, 57–72.

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

———, 1999, Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox: Geophysical Journal International, **137**, 805–815.

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

Rickett, J., 2000, Traveltime sensitivity kernels: Banana-doughnuts or just plain bananas?: SEP-Report, **103**, 63–71.

Ristow, D. and T. Ruhl, 1997, Migration in transversely isotropic media using implicit operators: SEG Expanded Abstract, **67**, 1699–1702.

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

Shan, G., 2006, Optimized implicit finite-difference migration for VTI media: SEP-Report, **124**, 279–292.

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

Xie, X.-B. and H. Yang, 2009, The finite-frequency sensitivity kernel for migration residual moveout and its applications in migration velocity analysis: WTOPI Report, **16**.

Early-arrival waveform inversion for near-surface velocity and anisotropic parameters: inversion of synthetic data

Xukai Shen

ABSTRACT

I test different inversion parametrizations of vertical velocity and anisotropic parameter ϵ. A model space parametrized by the squares of vertical and horizontal velocity results in vertical velocity and ϵ updates with opposite signs. On the other hand, a model space parametrized by the logarithm of the vertical velocity squared and ϵ has more reasonable updates, as well as better data matching. However, ambiguity does exist in the inversion results between vertical velocity and ϵ. I clearly demonstrate these findings using a synthetic example.

INTRODUCTION

In a separate paper (Shen, 2012), I showed that in the case of early-arrival waveform inversion, changes in vertical velocity and ϵ affect data kinematics much more than δ changes do. Since δ has a minimal effect on data kinematics, I proposed joint inversion for vertical velocity and ϵ, while holding δ fixed. For such inversion, I proposed three parametrizations for joint inversion of vertical velocity and anisotropic parameter ϵ. Naive parametrization is not suitable for joint inversion, since it results in no updates to ϵ. Velocity parametrization and logarithmic slowness parametrization both seem more suitable.

In this paper, I compare the inversion results of each parametrization using two synthetic examples, one laterally invariant vertical-gradient background model with Gaussian anomalies, and one that is part of the BP 2002 velocity model with anisotropic parameters. First I will describe the equation used and the model parametrization. Then I will compare various inversion results using the two synthetic examples.

THEORY

Exact anisotropic wave equations are in the form of elastic wave equations. Acoustic anisotropic wave equations can be obtained by various approximations of the exact elastic equations. One way to do this is to set shear wave velocity to zero in the exact elastic wave equations. Detailed derivation can be found in several papers (Zhang and Zhang, 2009; Crawley et al., 2010; Duveneck et al., 2008). The resulting acoustic anisotropic wave equations are a system of second-order equations:

$$\begin{aligned} \frac{\partial^2 p}{\partial t^2} &= v_p{}^2 (1+2\epsilon) \frac{\partial^2 p}{\partial x^2} + v_p{}^2 \sqrt{1+2\delta} \frac{\partial^2 r}{\partial z^2} \\ \frac{\partial^2 r}{\partial t^2} &= v_p{}^2 \sqrt{1+2\delta} \frac{\partial^2 p}{\partial x^2} + v_p{}^2 \frac{\partial^2 r}{\partial z^2}, \end{aligned} \tag{1}$$

where p and r are horizontal and vertical stress, respectively, v_p is vertical p-wave velocity, and ϵ and δ are anisotropic parameters (Thomsen, 1986).

The velocity parametrization defines the model space as follows:

$$\begin{aligned} m_1 &= {v_p}^2 \\ m_2 &= {v_h}^2 = {v_p}^2 (1 + 2\epsilon) . \end{aligned} \tag{2}$$

The logarithmic slowness parametrization defines the model space as follows:

$$\begin{aligned} m_1 &= \ln \left({v_p}^{-2} \right) \\ m_2 &= 1 + 2\epsilon. \end{aligned} \tag{3}$$

The non-linear conjugate gradient method is used for the inversion.

EXAMPLES

In this section, I compare the effectiveness of inversion model parametrization using two synthetic examples. In the first example, I illustrate the ambiguity between vertical velocity and ϵ, and then compare joint inversion results using the two different model parametrizations (equations 2, 3). In the second example, I compare joint inversion results using the two different parametrizations, as well as results from isotropic inversion. In both examples, I set δ to zero in both the true model and starting model, since it has very little effect on early-arrival kinematics.

Simple synthetic

The first example is relatively simple. Background velocity and ϵ are laterally invariant, and value increases with depth (Figure 1 left column). Model perturbation consists of two symmetrical Gaussian anomalies in velocity, and one Gaussian anomaly in ϵ (Figure 1 right column). The ϵ anomaly is co-located with the left velocity anomaly. The true model consists of the background model plus the model perturbation. The starting model is the background model. Recorded early-arrival data using this model consists mostly of diving waves, and the model perturbation causes non-trivial traveltime perturbation in those waves, which is the major information used for updating the starting model.

This is an ideal example to illustrate the ambiguity between velocity and ϵ. I use the conventional L2 full waveform inversion (FWI) objective function (Tarantola, 1984), and vary the strength of the left velocity anomaly and the left ϵ anomaly, while setting the strength of the right velocity anomaly to zero. A total of 64 shots with 225-m spacing are modeled using the background model, with 7 Hz peak source frequency. Since the radius of all anomalies is 300 m, at a given source frequency, anomalies of this size mostly affect data traveltime rather than acting as a defractor. The traveltime differences are the major contributor to the objective function value. With receivers everywhere on the surface, the objective function value is shown in Figure 2 as a function of the two anomaly strength percentages. The valley in the center of the figure clearly indicates the ambiguity between velocity and anisotropy. If ambiguity exists even with such complete surface geometry, we

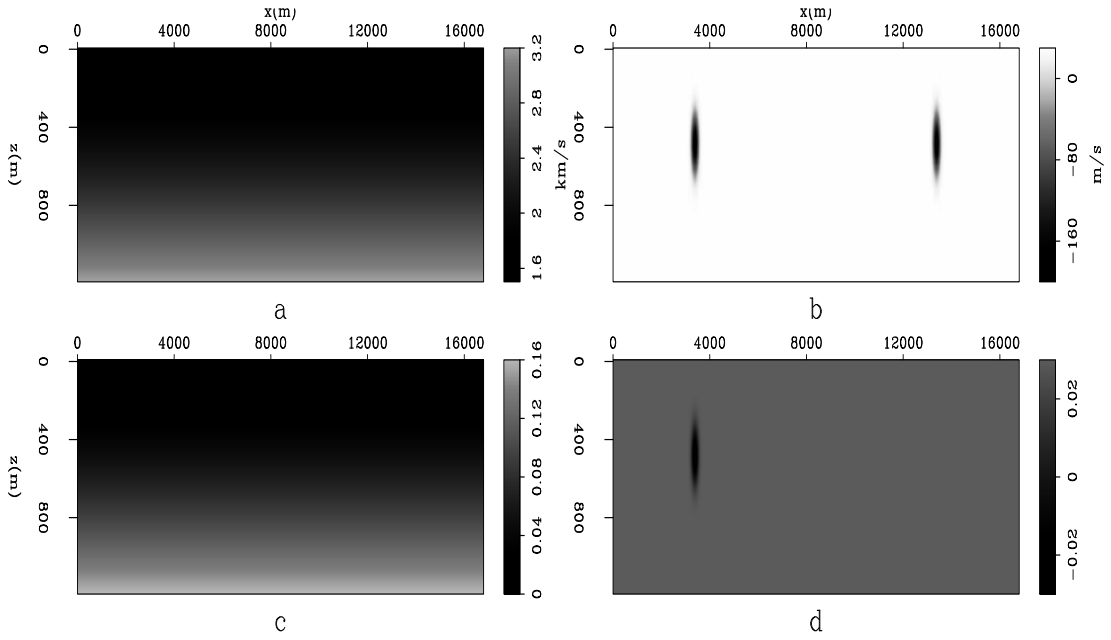


Figure 1: Background model and model perturbation. a): background velocity model; b): velocity model pertubation; c): background ϵ model; d): ϵ model pertubation. [**ER**] xukai2/. truemodsimpbw

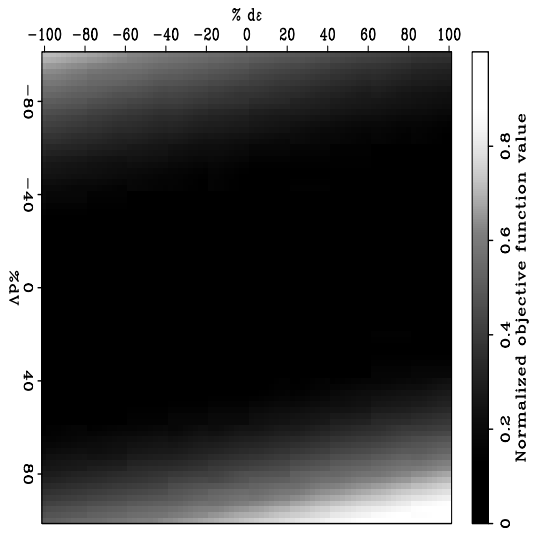


Figure 2: Normalized objective function value as a function of percentage changes in the left vertical velocity anomaly and ϵ anomaly strength. The vertical axis indicates the percent change in the vertical velocity anomaly. The horizontal axis indicates the percent change in the ϵ anomaly. [**NR**] xukai2/. fvalbw

can expect significant ambiguity in realistic acquisition scenarios. In such cases, proper model styling is needed in addition to data fitting.

Inversion results using velocity parametrization are shown in the left column of Figure 3. Inversion results using logarithmic slowness parametrization are shown in the right column of Figure 3. Results from using logarithmic slowness parametrization are significantly better in terms of data misfit (Figure 4). However, even with very small data misfit, inversion results are still quite different from the true model. Such results also suggest the existence of ambiguity between vertical velocity and ϵ in this simple example.

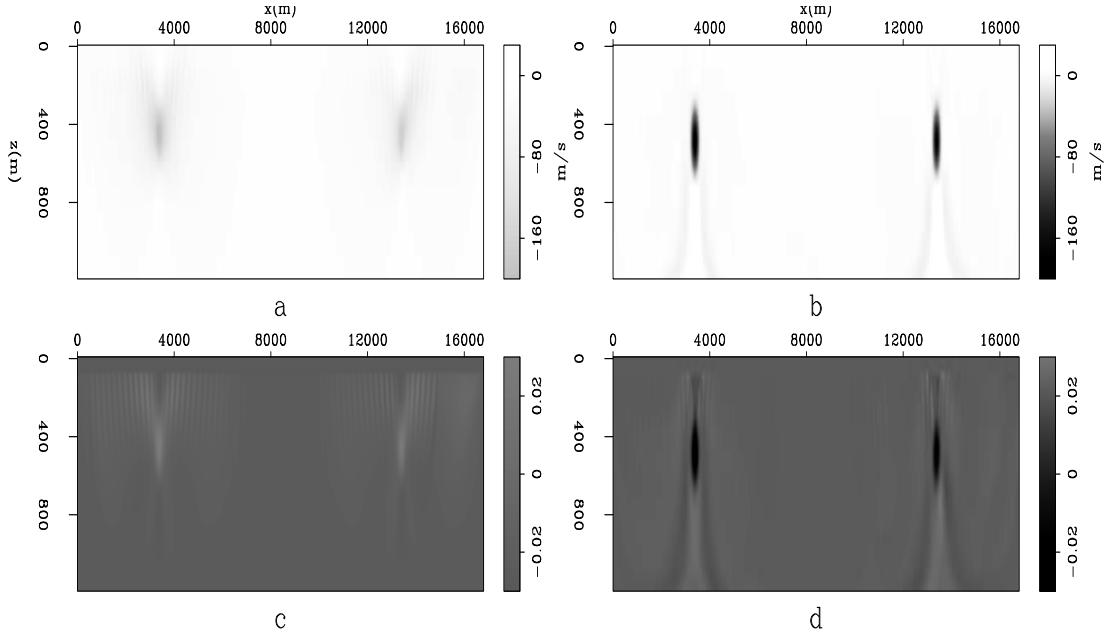


Figure 3: Inverted model perturbation. a): expressed in vertical velocity using velocity parametrization; b): expressed in vertical velocity using logarithmic slowness parametrization; c): expressed in ϵ using velocity parametrization; d): expressed in ϵ using logarithmic slowness parametrization. [**CR**] xukai2/. invdmodsimpbw

Complex synthetic

In this example, I invert the synthetic model shown in Figure 5. For joint inversion, starting models of vertical velocity and ϵ are a smoothed version of the true models without the shallow gas-pockets (Figure 6), and the δ parameter is fixed at zero. A total of 64 shots with 320-m shot spacing are used, with sources of 7 Hz peak frequency. The gas pocket sizes range from several hundred meters to almost two thousand meters. Given the dominant wavelength of the source wavelet, gas pocket sizes are on the order of several wavelengths, ensuring that most of the data misfit comes from the diving-wave traveltime difference. For simplicity, we assume that receivers are everywhere on the sea surface. The true model produces both refraction and reflection data. However, data masks are used during inversion to make sure that the inversion relies only on matching early arrivals that precede direct arrivals. Joint inversion results are shown for velocity parametrization in

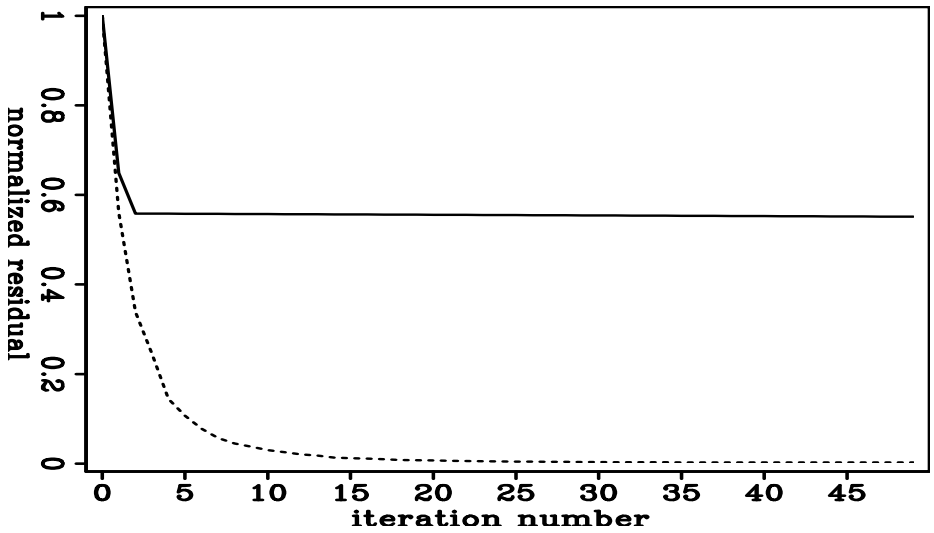


Figure 4: Data residual as a function of iteration number during inversion using different inversion model parametrization. Continuous: velocity parametrization; dashed: logarithmic slowness parametrization. [**CR**] xukai2/. resdsimpbw

Figure 7, and for the logarithmic slowness parametrization in Figure 8. The logarithmic slowness parametrization yields a far superior model compared to inversion using velocity parametrization. This can also be seen in the data misfit comparison (Figure 9).

For velocity parametrization, the gradient for vertical velocity has more terms than the gradient for horizontal velocity. This leads to more updates for vertical velocity than for horizontal velocity and tends to result in the epsilon update having a sign opposite to that of the vertical velocity update. Since this is not always geologically true, such inversion model parametrization is not ideal, at least for gradient-based inversion methods. Inversion using logarithmic slowness parametrization results in reasonable data matching; however, relatively smooth ϵ updates in the inversion result do suggest data misfit from ambiguity between ϵ changes and vertical velocity changes. This ambiguity can potentially be mitigated by proper model styling.

For comparison, isotropic inversion is also performed on this model. Two types of starting velocity are used: one is the same vertical velocity model as the starting model in joint inversion, and the other is the horizontal velocity model calculated from the vertical velocity and ϵ of the starting model used in joint inversion. The inversion results and data misfit are shown in Figures 11 and 12. It is obvious that inversion results are more geologically reasonable starting from the smooth vertical velocity model. This is also supported by data misfit. Comparing results with the true vertical velocity model and the true horizontal velocity model (Figure 10), it is interesting to see that in the better inversion result, the shallow parts mostly agree with the true vertical velocity model, and the deeper parts mostly agree with the true horizontal velocity model.

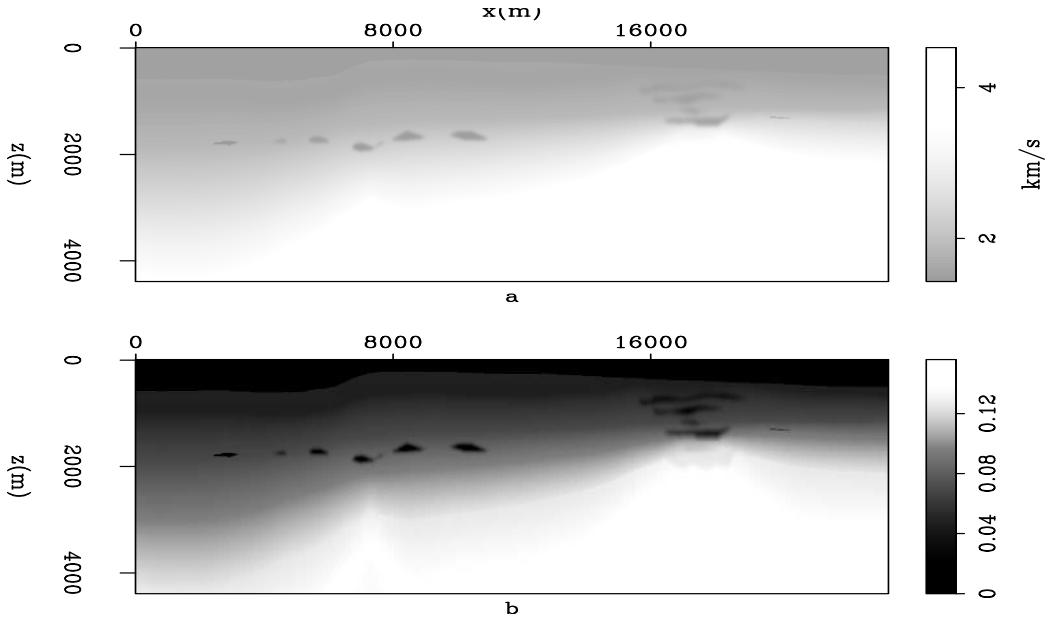


Figure 5: True model. a): velocity model; b): ϵ model. [**ER**] xukai2/. modelbw2

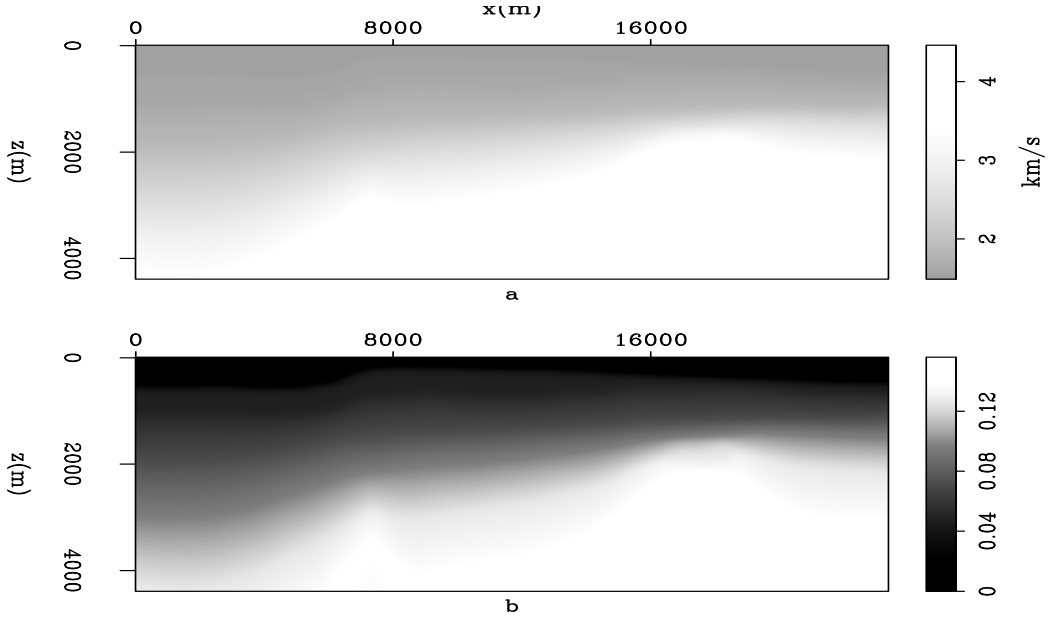


Figure 6: Starting model. a): velocity model; b): ϵ model. [**ER**] xukai2/. initmodbw

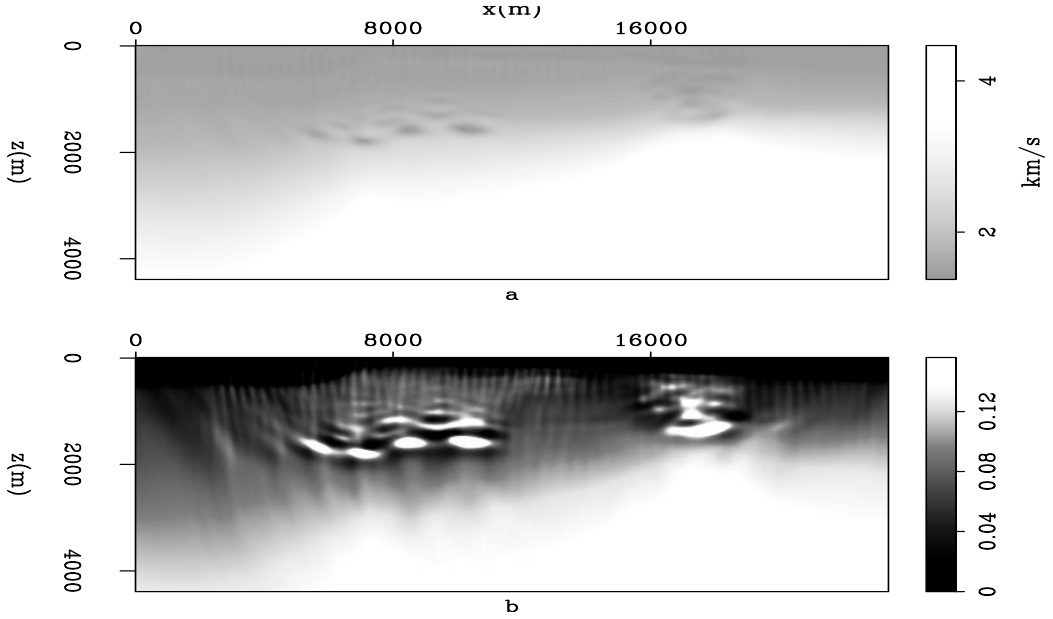


Figure 7: Inversion results using velocity parametrization, expressed in vertical velocity and ϵ. a): velocity model; b): ϵ model. [**CR**] xukai2/. invmodvvbw

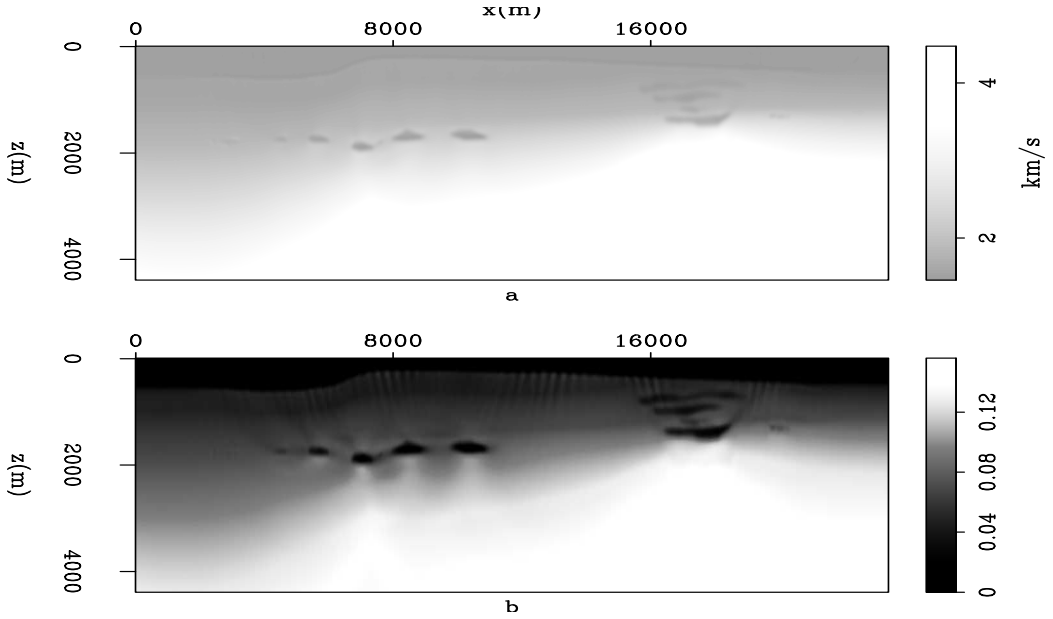


Figure 8: Inversion results using logarithmic slowness parametrization, expressed in vertical velocity and ϵ. a): velocity model; b): ϵ model. [**CR**] xukai2/. invmodsebw

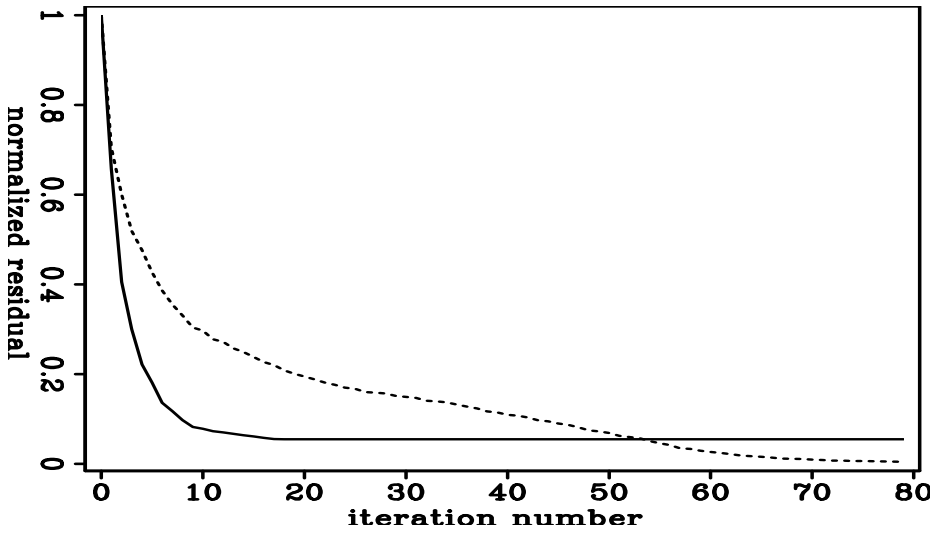


Figure 9: Data residual as a function of iteration number during inversion using different inversion model parametrizations. Continuous: velocity parametrization; dashed: logarithmic slowness parametrization. [**CR**] xukai2/. resdbw

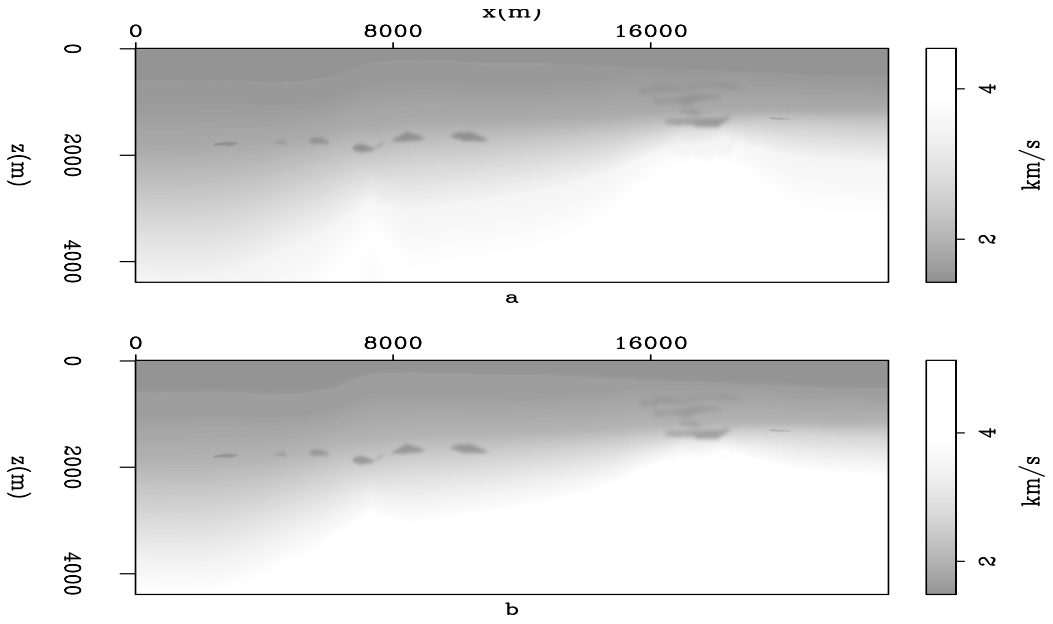


Figure 10: True velocity model. a): vertical velocity; b): horizontal velocity. [**CR**] xukai2/. isomodbw

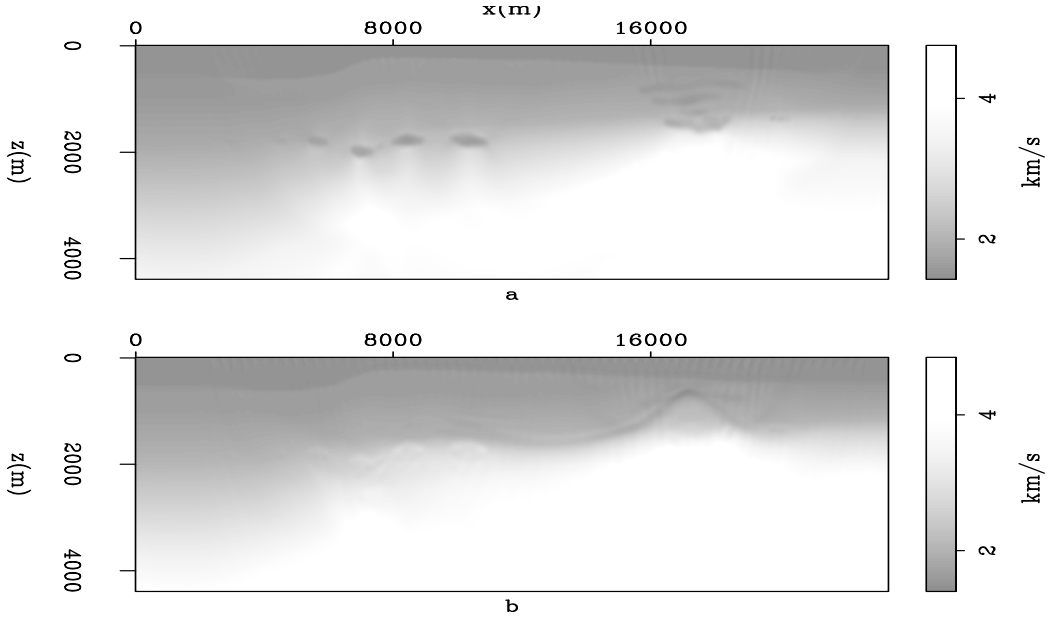


Figure 11: Velocity model from isotropic inversion starting from, a): smooth vertical velocity; b): smooth horizontal velocity. [**CR**] xukai2/. isoinvbw

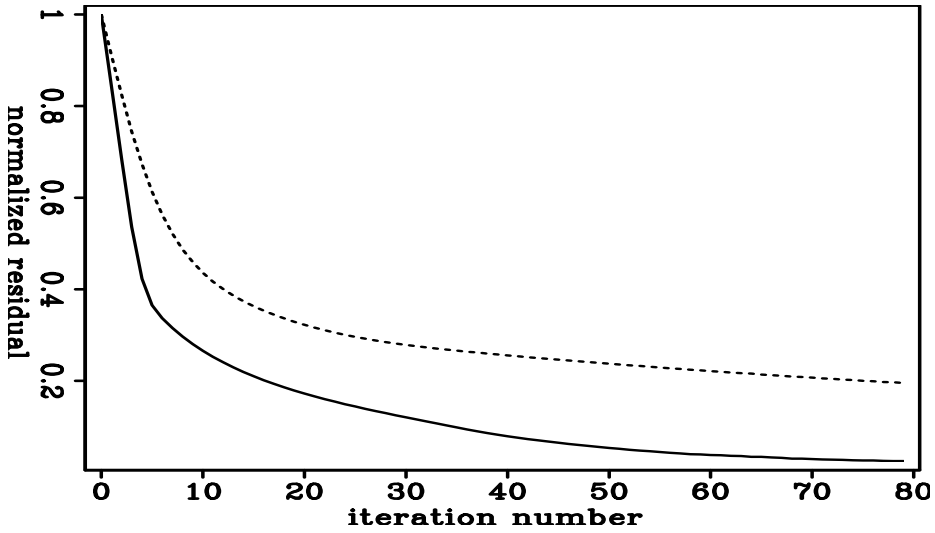


Figure 12: Data residual as a function of iteration during isotropic inversion using different starting models. Continuous: starting from smooth vertical velocity; dashed: starting from smooth horizontal velocity. [**CR**] xukai2/. isoresdbw

CONCLUSIONS

I compared inversion results using different model-space parametrizations for early-arrival waveform inversion. Results from isotropic inversion of anisotropic data are similar to the vertical velocity in the shallow parts of the model, and closer to the horizontal velocity model in the deeper part. For joint inversion of vertical velocity and the ϵ parameter, I kept δ fixed due to the insensitivity of the data to δ changes. A model space parametrized by squares of vertical velocity and horizontal velocity tends to produce updates in vertical velocity and ϵ that have opposite signs, which is unfavorable in practical inversion. The model space parametrized by the logarithm of the squared vertical slowness and epsilon has more reasonable updates and gives far better inversion results. However, ambiguity does exist in the inversion results between vertical velocity and ϵ. The same data misfit can be explained by different combinations of vertical velocity and ϵ, so proper model styling is needed to reduce such ambiguity.

REFERENCES

Crawley, S., S. Brandsberg-Dahl, J. McClean, and N. Chemingui, 2010, TTI reverse time migration using the pseudo-analytic method: The Leading Edge, **29**, 1378.

Duveneck, E., P. Milcik, P. M. Bakker, and C. Perkins, 2008, Acoustic VTI wave equations and their application for anisotropic reverse-time migration: SEG Expanded Abstracts, 2186–2189.

Shen, X., 2012, Early-arrival waveform inversion for near-surface velocity and anisotropic parameters: Modeling and sensitivity kernel analysis: SEP-Report, **147**, 73–82.

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

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

Zhang, Y. and H. Zhang, 2009, A stable TTI reverse time migration and its implementation: SEG Expanded Abstracts, 2794–2797.

Estimation of Q from surface-seismic reflection data in data space and image space

Yi Shen

ABSTRACT

Estimating seismic attenuation, or Q, is crucial for imaging applications, reservoir characterization and seismic acquisition design. A new method for Q estimation uses central frequency shift to analyze Q versus offset or angle in both the data space and image space. Q migration and Q compensation are also presented here for the image-based Q estimation. Applying this new approach to a 2D synthetic model demonstrate that both data-based and image-based methods are able to accurately estimate Q for a model with flat reflectors and homogeneous medium.

INTRODUCTION

Seismic waves propagating through the earth are attenuated by intrinsic attenuation and scattering attenuation. Scattering attenuation transfers wave energy to later arrivals or other directions. Intrinsic attenuation transfers wave energy to heat. An understanding of intrinsic attenuation has three major motivations.

First, attenuation is important for characterizing rock and fluid properties, e.g., saturation, porosity, permeability, and viscosity, because attenuation is more sensitive than velocity to some of these properties. For example, Q can serve as a lithology discriminator in a frontier area with sparse well control (Dasgupta and Clark, 1998). In addition, since the magnitude of the attenuation is directly related to petrophysical parameters, Q analysis provides a potential tool for reservoir characterization; it can help determine the contents (e.g. gas saturation) of a reservoir, map fracture azimuth to target reservoir development, and monitor the mobility of reservoir fluids to optimize the injection process.

Second, if ones knew the absorption properties of the subsurface, they could include them in seismic data processing (deconvolution, stacking, migration, inverse Q filtering, etc.) and get sharper images and higher resolutions. Attenuation estimation could also be used to better interpret the effects of AVO and anisotropy, which have offset-dependent signatures. Furthermore, full waveform inversion can achieve improved accuracy by incorporating attenuation into the initial model.

Third, attenuation estimates are useful in seismic acquisition design. Knowing the level of seismic attenuation in the survey-planning stage helps determine how much signal may reach the target, and enables the optimization of acquisition parameters.

However, it is difficult to estimate seismic attenuation. First, it is hard to distinguish intrinsic attenuation and scattering attenuation. In this study, I concentrate on the intrinsic attenuation and try to reduce the influence of the scattering attenuation. In the situation when the scattering attenuation cannot be ignored, the estimated attenuation is

a combination of intrinsic and scattering effects. The other most important reason is that amplitudes are easily affected by many factors, such as geometric spreading, source and receiver coupling, radiation patterns, transmission/reflection effects, focusing effects, PS conversion, etc. In this paper, I first review existing methods of estimating Q and illustrate how they account for these contaminating factors. Then, according to the advantages and drawbacks of these conventional Q measurements, I propose a new method of Q estimation from surface seismic data and test it on a 2D synthetic model.

METHODS REVIEW

In this section, I review previous studies which have addressed Q estimation, and discuss how they solve the challenges listed above. Generally, Q tomography algorithms are classified into two main categories. One category is ray-based tomography (Quan and Harris, 1997; Plessix, 2006; Rickett, 2006, 2007), where the relationship between the measurements and the inverse of Q is assumed to be linear. The other category is wave-equation-based tomography (Liao and McMechan, 1996; Hicks and Pratt, 2001; Pratt et al., 2003; Watanabe et al., 2004; Gao et al., 2005; Xin et al., 2008), which is formulated as a nonlinear optimization problem.

Ray-based tomography

In ray-based tomography, the process of wave propagation is described as a linear system, which assumes the relationship between the measurements and the inverse of Q is assumed to be linear. If the amplitude spectrum of an incident wave is $S(f)$ and the instrument/medium response is $G(f)H(f)$, the received amplitude spectrum $R(f)$ may be expressed as

$$R(f) = G(f)H(f)S(f), \tag{1}$$

where the factor $G(f)$ includes geometrical spreading, instrument response, source/receiver coupling, radiation pattern, and reflection/transmission coefficients, and $H(f)$ describes the attenuation effect on the amplitude. Always, $G(f)$ is set to be constant over frequency, because these amplitude factors are frequency independent in the homogeneous media.

Based on this linear assumption, there are several ways of estimating the effective Q. The most popular way is the spectral ratio method. This method takes the logarithm of equation 1, which is shown as follows:

$$Y(f) = Cf + B, \tag{2}$$

where

$$Y(f) = \ln\left[S(f)/R(f)\right], \tag{3}$$

$$C = \int_{ray} \alpha_0 dl, \tag{4}$$

$$B = -\ln(G), \tag{5}$$

$$\alpha_0 = \frac{\pi}{Qv}. \tag{6}$$

It can be seen that the integrated attenuation C is the slope of the plot of $Y(f)$ versus frequency f. This method may remove the effect of factor G, when G does not depend on frequency f.

The other way of estimating Q is by central frequency shift, which is a statistics-based method that estimates the attenuation coefficient from the spectral centroid downshift over a range of frequencies. In most natural materials, seismic attenuation increases with frequency. The high-frequency components of the seismic signal are attenuated more rapidly than the low-frequency components. As a result, the centroid of the signal's spectrum experiences a downshift during propagation. Under the assumption of a frequency-independent Q model, this downshift is proportional to a path integral through the attenuation distribution and can be used as observed data to reconstruct the attenuation distribution tomographically. The relation between the downshift and the attenuation is mathematically shown as follows,

$$\int_{ray} \alpha_0 dl = (f_s - f_R)/\sigma_s^2, \tag{7}$$

where the central frequency is defined as $f_c = \frac{\int_0^\infty fS(f)df}{\int_0^\infty S(f)df}$, f_s is the central frequency of the input source, f_R is the central frequency of the received signal, and σ_s is the variance of the source spectrum.

The advantage of these two methods is that they do not depend on amplitude scaling. However, they are not robust and will not work if the signal-to-noise ratio is poor. In addition, these two methods need a reference or source amplitude. Furthermore, since these methods are based on ray tomography, they share the limitations and assumptions of the ray-based method, which has a low computational cost but cannot handle complex structure.

Wave-equation-based method

According to its definition, Q is a function of the real and imaginary velocity components

$$Q = -\frac{v_r}{2v_i}. \tag{8}$$

The wave-equation-based method uses velocity to estimate Q and regards velocity and Q as two variables to be estimated. This method is physically more accurate but is computationally expensive and has trade-offs between the effects on data amplitude of these different parameter classes.

PROPOSED METHOD

I propose a new method that balances the advantages and disadvantages of the conventional methods discussed in the above section. The basic idea of this proposed method is to use central frequency shift (Quan and Harris, 1997) to analyze Q versus offset (QVO) (Dasgupta and Clark, 1998) or Q versus angle (QVA) in the image domain.

According to equation 7, the central frequency shift varies with offset because of differing raypath geometries, and hence in accumulated attenuation. Specifically, if I define a central-frequency versus offset or versus angle domain, the central frequency is flat when there is no absorption, but shows a curvature in an absorptive medium. By the same reasoning, the same response can be observed in the central-wavenumber-versus-offset/angle domain, which can be obtained after migration.

Advantage

Using central freqency shift to analyze QVO/QVA has several advantages over the previous Q estimation methods (Quan and Harris, 1997; Plessix, 2006; Rickett, 2006, 2007), which analyze Q from a stacked trace. First, the spectral amplitude of a stacked trace has a distorted attenuation signature, since the path lengths (hence accumulated attenuation), spectral distortions from NMO stretch, and reflectivity-transmissivity effects vary from one trace to the next for a given reflection event. Performing an analysis along the offset or angle may help reduce such distortion. Second, with the relevant information from different offsets or angles, I do not need the reference/source information that is a prerequisite for the previous methods. Third, a number of applications have already demonstrated the effectiveness of QVO for discrimination of lithology types (Dasgupta and Clark, 1998; Hackert and Parra, 2004), azimuthal discrimination of fractures (Clark et al., 2009; Moffat et al., 2009), and discrimination of time-lapse changes (Clark et al., 2001; Blanchard et al., 2009). In addition, there are also several potentially promising applications for this method. For example, such analysis may reduce the ambiguity between velocity and attenuation, because velocity does not have a frequency-dependent effect on the amplitude, and thus has no central-frequency shift. QVO may also help analyze amplitude variation with offset (AVO), as attenuation effects are always superimposed on AVO signature.

Migration-based Q analysis is used in this proposed method, instead of the ray-based Q estimation methods that were used in previous methods. One advantage of using a migration-based method is that in the presence of complex wave propagation (i.e. multi-pathing and geometrical dispersion), data-based methods often cannot produce high-quality images, while migration-based methods yield better images because they can handle the complex structure in the subsurface. In addition, the signal-to-noise ratio is higher in the image domain than in the data domain, leading to a higher-quality result from the central-frequency shift.

Challenge

A number of factors make determining the attenuation difficult using this method. First, the difficulty comes from the effects on the seismic spectrum from interference of different seismic arrivals, so it is necessary to carefully window the desired reflector or variable-window time-frequency transform. Second, angle-dependent effects (directivity, anisotropy, raypath differences) hamper accurate interval measurements. Third, geometric information (local velocity, dip event, opening angle) must be taken into consideration in the image domain.

THEORY AND NUMERICAL TESTS

Forward modeling with Q

There are different models to describe Q: the Maxwell model, the Kelvin-Voigt model, the Zener model, the Burgers model, etc. Q behaves differently in the frequency domain in these different models. However, I use a nearly constant-Q model in this paper. The reason is that in oil prospecting and seismology, constant-Q models are convenient to parametrize attenuation in rocks, since the frequency dependence is usually not known. Moreover, there is physical evidence that attenuation is almost linear within the range of seismic frequencies, and thus the effect of Q is independent of these frequencies. Futterman (1962) has developed the theory for nearly constant Q, and I can use it in the forward modeling with Q.

In Futterman's theory, the visco-acoustic equation has the same form as the acoustic equation, but the velocity is a complex number,

$$\tilde{v}(\omega) = v_{\omega r}\left(1 - \frac{1}{\pi Q}\ln(\omega/\omega_r)\right)^{-1}\left(1 + \frac{i}{2Q}\right)^{-1}, \tag{9}$$

where ω_r is the reference angular frequency, and $v_{\omega r}$ is the velocity at the reference angular frequency. Forward modeling can be executed with equation 9 using conventional one-way downward continuation, and more detailed algorithm is described in equations 11-16 in the following migration section.

I present a simple 2D synthetic example for the forward modeling. The model size is 4000 m (length) x 2500 m (depth). A horizontal reflector is at 1500 m depth. The source is located at x=0 on the surface, and 401 receivers are uniformly distributed along the surface. The medium is assumed to be homogeneous with constant velocity (2000 m/s) and constant Q (50 for the model with attenuation and 99999 for the model without attenuation). Injecting a Ricker wavelet produces the forward-modeled data in Figures 1(a) and 1(b). Their central frequencies are shown in Figures 1(c) and 1(d).

The wavelet in Figure 1(a) is stretched relative to the one in Figure 1(b), because the high frequency is attenuated more than the low frequency, which broadens the wavelet. Their central frequencies are also significantly different. Central frequency with attenuation in Figure 1(c) is a hyperbola along the offset, matching the theoretical calculation well. In contract, the central frequency is a flat line in Figure 1(d), indicating that no attenuation is accumulated along the ray path. Hence, Figures 1(c) and 1(d) provide further evidence that QVO carries information about attenuation and can be used to estimate Q.

Estimation of Q from data space

Because this simple 2D example has no complex structure or noise, I apply a data-space QVO method to this model before going to the image domain. According to the theory developed by Quan and Harris (1997), the central-frequency shift changes with attenuation in the way described in equation 7. If the wave travels in the homogeneous media and reflected from one flat reflector, equation 7 can be simplified in the following way,

$$x^2 + z^2 = C^2\Delta f^2 Q^2, \tag{10}$$

where x is offset, z is depth, $\Delta f = (f_S - f_R)$, and $C = v/(\pi\sigma_S^2)$, which is a constant if I assume the velocity and the variance of the source are known.

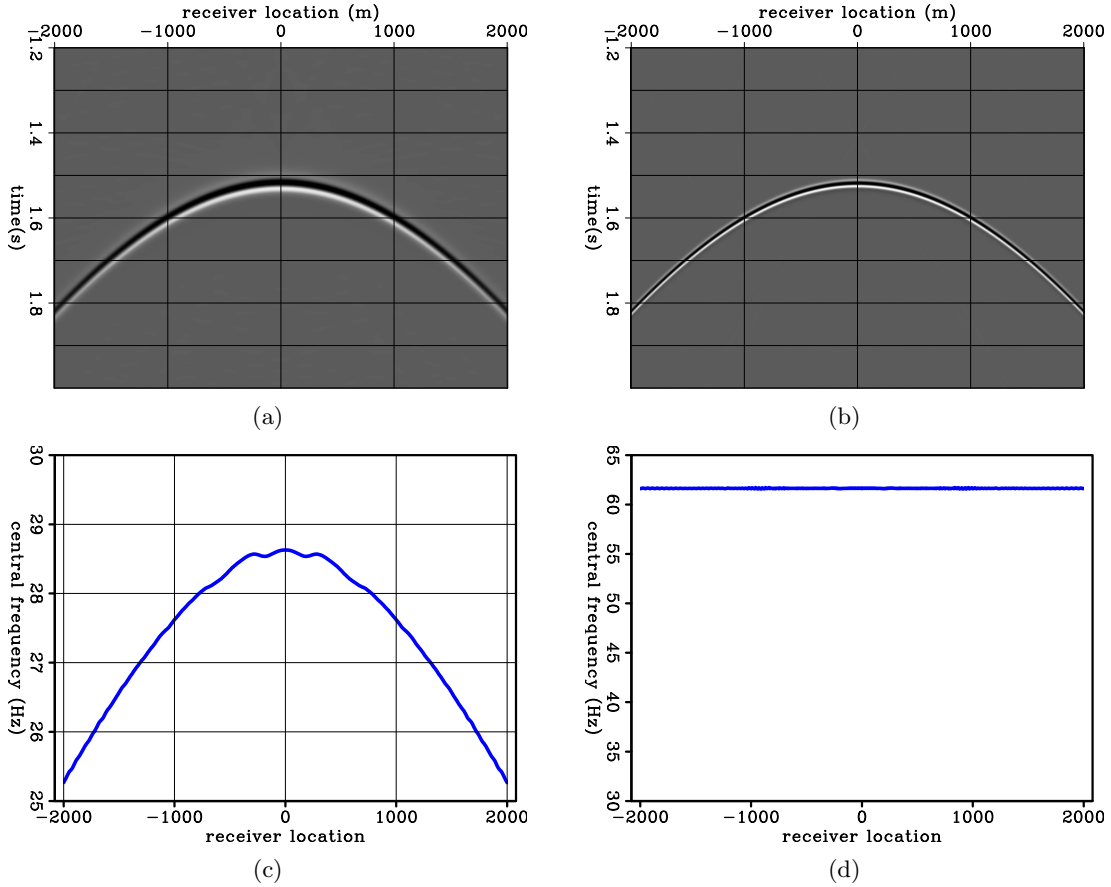


Figure 1: Given the 2D synthetic example, (a) modeled data with attenuation (Q=50); (b) modeled data without attenuation (Q=99999); (c) central frequency with attenuation; (d) central frequency without attenuation. [**ER**] yishen1/. trec50,trec99999,fpeak50,fpeak99999

Equation 10 relates four variables: two coordinates of the data space $(x, \Delta f)$ and two of the model space $(z, 1/Q)$. An impulse in model space corresponds to a hyperbola in data space. In the opposite case, an impulse at a point in data space corresponds to another hyperbola in the model space. If I sum along the trajectories in the data space, energy will be concentrated in the model space to indicate the Q value. In this paper, I define the model space as the Q-spectra and the procedure described above as the Q-scan. Least-squares inversion can also be used here to yield a better result.

Figure 2(a) shows the central-frequency shift of the modeled data in Figure 1(a). After scanning, I compute the Q-spectra in Figure 2(b). This result shows concentrated energy around the expected point (1500 m, 0.02), which indicates that the RMS Q is 50 above 1500 m depth. Figure 2(c) shows least-squares inversion after 50 iterations. The energy is more focused on the expected point and shows much higher resolution. Therefore, I conclude that computing the Q-spectra can be used in estimating the RMS Q for a model with simple structure.

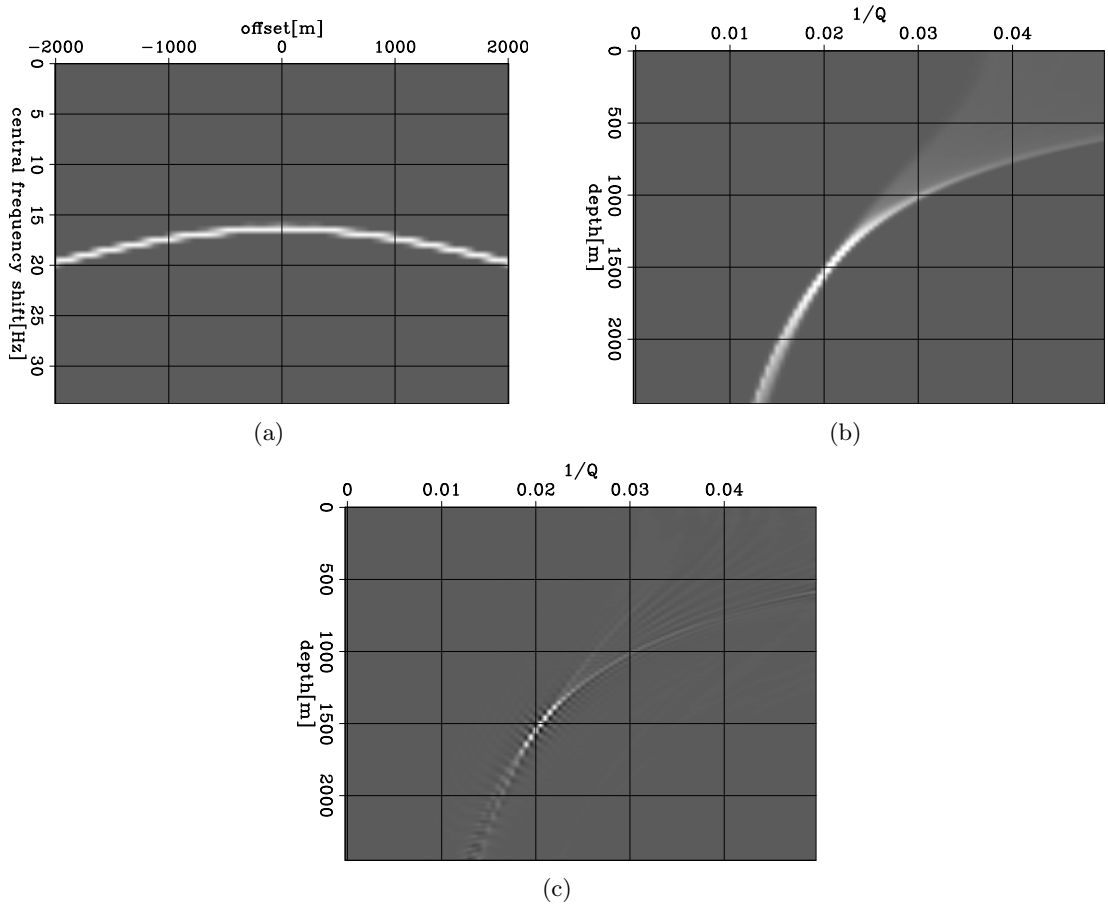


Figure 2: Figure 2(a) shows the central-frequency shift of the modeled data in Figure 1(a). The Q-spectra is computed in Figure 2(b). This result shows concentrated energy around the expected point (1500 m, 0.02), which indicates that the RMS Q is 50 above 1500 m depth. Figure 2(c) shows least-squares inversion after 50 iterations. The energy is more focused on the expected point and shows much higher resolution. [**ER**]

yishen1/. df,Qscanadj,Qscaninv

In reality, however, this data-domain method can be inaccurate when lateral velocity variations and dipping structures exist. Therefore, image-domain methods is needed to more accurately estimate Q.

Q migration

As with forward modeling, I use constant-Q assumption in Q migration. For convenience, I rewrite equation 9 in the form of slowness:

$$\tilde{s}(\omega) = s_{\omega r}\left(1 - \frac{1}{\pi Q}\ln(\omega/\omega_r)\right)\left(1 + \frac{i}{2Q}\right), \tag{11}$$

where $S_{\omega r}$ is the slowness at the reference frequency ω_r.

Having the new velocity/slowness, I obtain the new single square root as

$$\begin{aligned} k_z &= \mathrm{SSR}(\omega, \mathbf{k}) \\ &= \sqrt{(\omega\tilde{s})^2 - |\mathbf{k}|^2} \\ &= \sqrt{\left(\omega s_{\omega r}\left(1 - \frac{1}{\pi Q}\ln(\omega/\omega_r)\right)\left(1 + \frac{i}{2Q}\right)\right)^2 - |\mathbf{k}|^2}. \end{aligned} \tag{12}$$

This new SSR can be approximated into a simplified form by using Taylor expansion around reference slowness s_0 and reference quality factor Q_0:

$$k_z(s_{\omega r}, Q) = k_{z0}(\tilde{s}_0) + \omega\left(\tilde{s} - \tilde{s}_0\right) + \omega\left(\tilde{s} - \tilde{s}_0\right)\frac{2|\mathbf{k}|^2}{4\omega^2\tilde{s}_0^2 - 3|\mathbf{k}|^2}, \tag{13}$$

where

$$\tilde{s} = s_{\omega r}\left(1 - \frac{1}{\pi Q}\ln(\omega/\omega_r)\right)\left(1 + \frac{i}{2Q}\right), \tag{14}$$

$$\tilde{s}_0 = s_{\omega r0}\left(1 - \frac{1}{\pi Q_0}\ln(\omega/\omega_r)\right)\left(1 + \frac{i}{2Q_0}\right), \tag{15}$$

and $s_{\omega r}$ is the reference slowness at the reference frequency. The first two terms of equation 13 describe the split-step migration. The third term is the high-order correction, which allows for pseudo-screen migration.

The single square root for FFD migration is shown as follows:

$$\begin{aligned} k_z &= k_z^{\text{ref}} + k_z^{\text{split - step}} + k_z^{\text{f - d}} \\ &= \left(\sqrt{(\omega\tilde{s})^2 - |\mathbf{k}|^2}\right) + \omega\left(\tilde{s} - \tilde{s}_0\right) - \omega\tilde{s}\frac{a\left(1 - \frac{\tilde{s}_0}{\tilde{s}}\right)\left(\frac{1}{\omega\tilde{s}}|\mathbf{k}|\right)^2}{1 - b\left(1 + \frac{\tilde{s}_0^2}{\tilde{s}^2}\right)\left(\frac{1}{\omega\tilde{s}}|\mathbf{k}|\right)^2}, \end{aligned} \tag{16}$$

where $a = 0.5$ and $b = 0.25$ for 45-degree migration.

In addition, I can rewrite Q migration in a matrix form to conveniently compare with the conventional migration. The conventional migration can be written in the following matrix form:

$$\begin{aligned} \mathbf{d} &= \mathbf{F}\mathbf{m} \\ \mathbf{m} &= \mathbf{F}^T\mathbf{d} \end{aligned} \tag{17}$$

where $\mathbf{d}$ is the data, $\mathbf{m}$ is the model, $\mathbf{F}$ is the migration operator, and the superscript T indicates the matrix transpose.

The downward continuation migration with Q can be written as

$$\begin{aligned} \mathbf{d} &= \mathbf{AFm} \\ \mathbf{m} &= \mathbf{F}^T\mathbf{A}^T\mathbf{d}, \end{aligned} \tag{18}$$

where $\mathbf{A}$ is the attenuation operator, which consists of real numbers less than 1.

Equation 18 indicates that the migrated model will be further attenuated, with the attenuation operator $\mathbf{A}$ being applied to the attenuated modeled data. Therefore, Q migration will compensate for the phase change, but will not compensate for the amplitude loss due to attenuation.

In this section, I apply Q migration to the modeled data in Figures 1(a) and 1(b). Figure 3(a) shows the conventional migration of the non-attenuated data in Figure 1(b), which images the reflector at 1500 m depth. Figures 3(b) and 3(c) show the conventional migration and Q migration of the attenuated data in Figure 1(a). The wavelets in Figure 3(c) are stretched in comparison to the ones in Figure 3(a). This result confirms that Q migration further attenuates the data, instead of compensating for its amplitude loss.

Q compensation

In order to compensate for the amplitude loss to estimate Q in the image domain, I need Q compensation, which can be written in the following matrix form:

$$\begin{aligned} \mathbf{d} &= \mathbf{AFm} \\ \mathbf{m} &= \mathbf{F}^T\mathbf{A}^{-1}\mathbf{d}, \end{aligned} \tag{19}$$

where

$$\mathbf{A}^{-1} = \frac{\mathbf{A^T}}{\mathbf{A^TA}}. \tag{20}$$

One thing to notice is that the compensated amplitude in equation 19 contains both attenuation and an evanescent wave. However, only the attenuation part needs to be compensated. Therefore I need to remove the evanescent wave from the amplitude compensation. For a single frequency, the attenuation operator is shown as follows:

$$A = e^{-\Delta z\,\mathrm{Im}(k_z)}. \tag{21}$$

After removing the evanescent part from the amplitude compensation, the inverse attenuation operator is shown as follows:

$$A^{-1} = \frac{e^{-\Delta z\,\mathrm{Im}(k_z)}}{e^{-2\Delta z\,\mathrm{Im}(k_z - k_{Q=\infty})}}. \tag{22}$$

Figure 3(d) shows Q compensation on the attenuated data, showing exactly the same result as Figure 3(a), which has no attenuation in either forward propagation or backward imaging. This result indicates that Q compensation adequately restores the amplitude loss caused by attenuation.

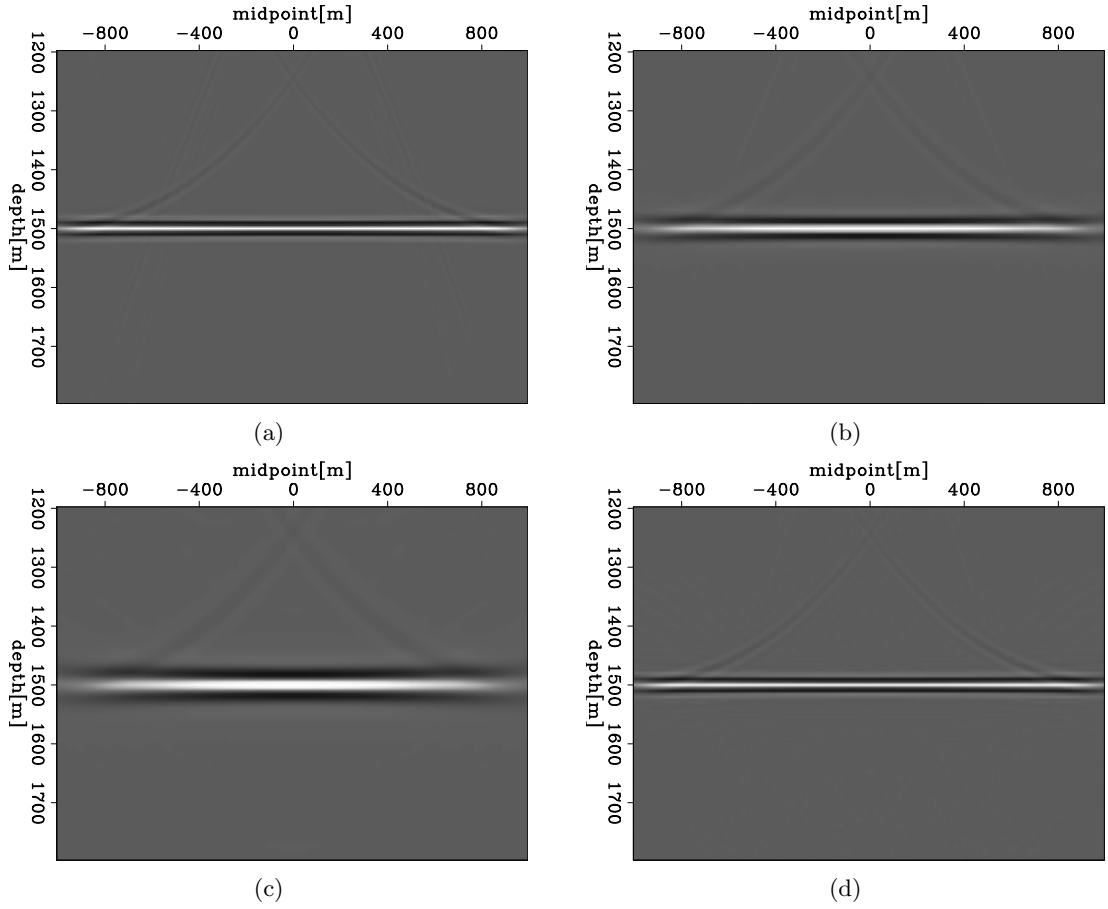


Figure 3: Given the 2D synthetic example, (a) conventional migration on non-attenuated data; (b) conventional migration on attenuated data; (c) Q migration on attenuated data; (d) Q compensation on attenuated data. [**ER**] yishen1/. ctrl,nqmig,qncp,qmig

Estimation of Q from image space

In the presence of complex structures and/or strong lateral Q variations, extracting Q information in the data space is both inaccurate and time consuming. In these situations, the image space is a more appropriate domain for extracting kinematic information, since migration/compensation focuses and greatly simplifies the events. Even when the Q for migration/compensation is far from the true Q, the unflattened central frequency shift is a step in the right direction. I present two ways of doing this migration Q analysis (MQA).

Depth-migration Q analysis

I use the wavenumber computed from surface offset-domain common-image gathers to improve the compensation Q function by performing a moveout analysis. A straightforward procedure to update the Q function is the following:

- First, take one common midpoint to compute the image wavenumber shift in its surface offset-domain common-image gathers after depth Q compensation. Here I define the image wavenumber as $k_{\text{img}} = 2\pi f/v$, which has an angle-dependent stretch of the vertical wavenumber k_z according to the following relation: $k_{\text{img}} = k_z/2cos\theta$, where θ is the opening angle. For a flat reflector, $cos\theta = z/\sqrt{x^2+z^2}$. Substituting this defined image wavenumber into equation 10, I get the central-image wavenumber shift in the following way:
$$\Delta k = \frac{8\sigma_s^2\pi^2}{v^2}(\frac{1}{Q} - \frac{1}{Q_\rho})\sqrt{x^2+z^2}, \tag{23}$$
where Q is the quality factor for forward modeling, Q_ρ is the quality factor for Q compensation, and $\Delta k = \frac{4\pi f_S}{v} - k_{\text{img}}$, which is defined as the central-image wavenumber shift.

  For the data generated by a flat reflector and homogeneous velocity and Q in Figures 1(a) and 1(b), images migrated or compensated using one shot at 0 m in Figure 3 is equivalent to a single image point at 0 m with respect to different source-receiver offsets. Although one shot is not enough for stacking during migration/compensation, it is still appropriate to analyze Q effect by using a limited range of the offsets.

  According to equation 23, if I plot the central-image wavenumber versus the offset or angle, the central wavenumber is flat with the correct Q for compensation, but shows a curvature with an incorrect Q for compensation. Figure 4(a) and 4(b) display the central image wavenumber of the image in Figure 3(b) and Figure 3(d). The central image wavenumber is almost flat over offsets in Figure 4(b) after being compensated with correct Q ($Q = 50$). In contract, Figure 4(a) shows a curvature along offset, since infinite Q was used for compensation. The results indicate that central-wavenumber shift is related to attenuation and can be used to estimate Q in the image domain.

  The image wavenumber k_{image}, rather than the vertical wavenumber k_z, is used to relate to attenuation in this paper. The reason for not using k_z is that the central-vertical wavenumber shift is independent of the offset/angle for a simple model with a flat reflector, which can be shown as follows:
$$\Delta k_z = \frac{8\sigma_s^2\pi^2}{v^2}(\frac{1}{Q} - \frac{1}{Q_\rho}), \tag{24}$$

where $\Delta k_z = \frac{4\pi f_S}{v}\frac{z}{\sqrt{x^2+z^2}} - k_z$. Thus, the central-vertical-wavenumber shift can not be applied to QVO/QVA analysis nor to estimating Q.

- Second, an inverse shift correction is applied to the gathers, according to the RMS Q used to compensate the data. Here I describe the shift correction using the following equation:

$$(\frac{1}{QC'})^2(x^2+z^2) = \Delta k_{img}^2, \tag{25}$$

where $C' = v^2/(2\pi^2\sigma_s^2)$.

- Third, Q spectra are computed for the CIGs after the inverse shift correction. Then the RMS Q measured from the Q spectra are used to generate a new RMS-Q function. The new RMS-Q function is used to compensate the data again, and, if needed, the process can be repeated to further improve the Q value.

  Figure 4(c) shows the Q spectra of Figure 3(b) that has incorrect Q for compensation. This result shows concentrated energy around the point (1500 m, 0.02), which indicates that the new RMS-Q is 50 above 1500 m. However, the energy is not as well focused as that in Figure 2(b), because of its relatively narrow range of offsets after imaging.

CONCLUSION

I first reviewed two main methods of estimating seismic attenuation: ray-based tomography and the wave-equation-based method. Then I presented a new method for estimating Q, which uses central frequency shift to analyze Q versus offset (QVO) or Q versus angle (QVA). Based on this proposed method, I developed theories for estimation of Q in both data space and image space, and applied them to a 2D synthetic model. I also presented Q migration and Q compensation methods based on one-way wave equation. Our results showed that Q compensation recovers the amplitude loss caused by attenuation, which is a successful step forward for image-based Q estimation.

FUTURE WORK

Ray-based Q-estimation methods have serious drawbacks for complex depth-imaging problems. In the presence of complex wave propagation (i.e. multipathing and geometrical dispersion), ray-based migration methods often cannot produce high-quality images, while wavefield-continuation methods yield better images. Therefore, wave-equation MQA (WEMQA) is needed to avoid the well-known instability problems that rays encounter when the Q model is complex and has sharp boundaries. The objective function of this method minimizes the residual wavenumber shift (equation 23) in the angle/offset domain. The theory and numerical results for this method are a topic of ongoing study.

ACKNOWLEDGMENTS

Thanks to Biondo Biondi, Robert Clapp and Dave Nichols of Stanford for discussions and suggestions.

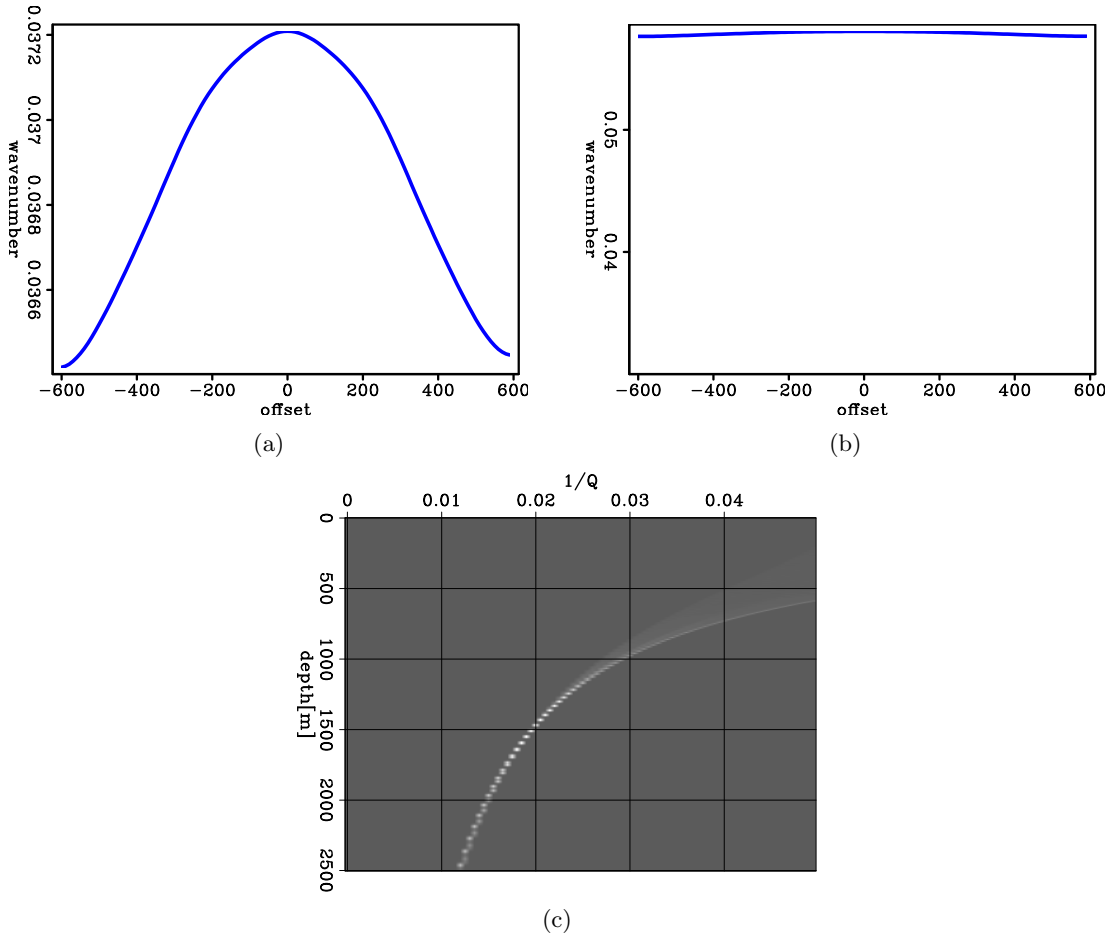


Figure 4: Given the 2D synthetic example, (a) central image wavenumber with correct Q (Q=50) for compensation ; (b) central image wavenumer with incorrect Q (Q=99999) for compensation; (c) Q spectra of Figure 3(b) that has incorrect Q for compensation. [**ER**] yishen1/. knqmig,kqmig,Qscanimg

REFERENCES

Blanchard, T. D., R. A. Clark, M. van der Baan, and E. Laws, 2009, Timelapse attenuation as a tool for monitoring pore fluid changes in hydrocarbon reservoirs: 71st Conference and Exhibition, EAGE, Expanded Abstracts, P052.

Clark, R. A., P. M. Benson, A. J. Carter, and C. A. G. Moreno, 2009, Anisotropic P-wave attenuation measured from a multi-azimuth surface seismic reflection survey: Geophysical Prospecting, **57**, 835–845.

Clark, R. A., A. J. Carter, P. C. Nevill, and P. M. Benson, 2001, Attenuation measurements from surface seismic data azimuthal variation and time-lapse case studies: 63rd Conference and Technical Exhibition, EAGE, Extended Abstracts, L–28.

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.

Gao, F., G. Fradelizio, A. Levander, G. Pratt, and C. Zelt, 2005, Seismic velocity, Q, geological structure, and lithology estimation at a groundwater contamination site: 75th Annual International Meeting, SEG, Expanded Abstracts, 1561–1564.

Hackert, C. L. and J. O. Parra, 2004, Improving Q estimates from seismic reflection data using well-log-based localized spectral correction: Geophysics, **69**, 1521–1529.

Hicks, G. J. and R. G. Pratt, 2001, Reflection waveform inversion using local descent methods: Estimating attenuation and velocity over a gas-sand deposit: Geophysics, **66**, 598–612.

Liao, Q. and G. A. McMechan, 1996, Multifrequency viscoacoustic modeling and inversion: Geophysics, **61**, 1371–1378.

Moffat, L. C., R. A. Clark, M. van der Baan, and T. Manning, 2009, Azimuthal variations of attenuation analysis applied on a 3D multiazimuth towed-streamer dataset: Research Workshop, EAGE/SEG, Expanded Abstracts, B06.

Plessix, R. E., 2006, Estimation of velocity and attenuation coefficient maps from crosswell seismic data: Geophysics, **71**, S235–S240.

Pratt, R. G., K. Bauer, and M. Weber, 2003, Crosshole waveform tomography velocity and attenuation images of Arctic gas hydrates: 73rd Annual International Meeting, SEG, Expanded Abstracts, 2255–2258.

Quan, Y. and J. M. Harris, 1997, Seismic attenuation tomography using the frequency shift method: Geophysics, **62**, 895–905.

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.

Watanabe, T., K. T. Nihei, S. Nakagawa, and L. R. Myer, 2004, Viscoacoustic waveform inversion of transmission data for velocity and attenuation: Journal of the Acoustical Society of America, **115**, 3059–3067.

Xin, K., B. Hung, S. Birdus, and J. Sun, 2008, 3D tomographic amplitude inversion for compensating amplitude attenuation in the overburden: 78th Annual International Meeting, SEG, Expanded Abstracts, **27**, 32–39.

Image gather reconstruction using StOMP

Robert G. Clapp

ABSTRACT

Constructing 3-D angle gathers through cross-correlation poses a computational problem primarily due 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. The correlation volume size can be reduced by both phase encoding and random subsampling. The full correlation gathers can then by reconstructed using an l_1 inversion scheme known as Stagewise Orthogonal Matching Pursuit. Preliminary results indicate that almost all angle gather information can be recovered.

INTRODUCTION

Reverse time migration (RTM) is quickly becoming the standard high-end seismic imaging technique. Significant work has been done on speeding up the kernel (Micikevicius, 2009; Nguyen et al., 2010; Nemeth et al., 2008; Clapp et al., 2010) but far less on constructing image gathers(Sava and Fomel, 2003, 2006) needed for rock property analysis or velocity updates. Image gather construction is a much less tractable problem because it is memory rather than compute intensive. The volume size of the 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* 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. For compressive sensing to work, a signal must be highly compressible. For compressive sensing to be worthwhile, the cost of inverting for the basis function must be significantly less than the cost of acquiring the full signal. Clapp (2011) showed that correlation gather construction fit the first criteria for a successful compressive sensing problem. Multi-dimensional correlation gathers/angle gathers are compressible at nearly a 100:1 ratio. The challenge became finding an inversion scheme that could accurately enough recover the full model. Donoho et al. (2006) proposed a solution to the second problem, an ℓ_1 version methodology that works for a large number of unknowns. In this paper, I apply the Stagewise Orthogonal Matching Pursuit (StOMP) algorithm to correlation gather reconstruction. I show that the angle domain representation of the sparsely acquired gathers is similar to the representation of the full data. I then apply a phase encoding technique, combining many different correlations to every data point to further improve the inverted model.

IMAGE GATHERS AND WAVELET COMPRESSION

In RTM downward continuation based wave-equation imaging we have a source $s(x, y, z, it, is)$ and receiver wavefield $g(x, y, it, is)$, where it is a given time and is is a given shot. We form our migrated image $i(x, y, z)$ as

$$I(x, y, z) = \sum_{is=0}^{ns} \sum_{it=0} nts(x, y, z, it, is) g(x, y, it, is). \quad (1)$$

The imaging condition for RTM just replaces the sum over time with a sum over frequency. This basic imaging condition hides our information about velocity and rock properties by stacking over all angles illuminated by our source/receiver geometry.

Several solutions to this problem have been proposed that attempt to extract information as a function of angle from the data. In this paper I focus on the family of methods that construct shift gathers by cross-correlation. The shift can be a function of space (Sava and Fomel, 2003) or time (Sava and Fomel, 2006), or both. In these cases the imaging condition takes the form of

$$I(x, y, z, h_x, h_y, h_z) = \sum_{\text{shot}} \sum_{t} s(x + h_x, y + h_y, z + h_z, it) g(x - h_x, y - h_y, z - h_z, it), \quad (2)$$

where h_x, h_y, h_z are how much the source and receiver wavefields have been shifted in a given direction. Introducing these shift gathers poses two problems. The first, and smaller of the problems, is that computational expense associated with constructing these gathers, when going beyond a single shift axis, is at least on the same order of magnitude as the propagation kernel. The larger problem is the expansion of the volume size anywhere from 20–1000 fold, means that the shift gathers must be stored a memory level further away from the processor, often on disk, which is often several orders of magnitude more expensive to access. Finding some way to reduce the volume size, and better still, the computational expense associated with shift gather construction, can be highly beneficial. In order for compressive sensing to be an appropriate way to reduce the volume size it is important to consider the compressibility of seismic data.

There is significant literature on compressing seismic data. Relatively low compression ratios are achievable by compressing a trace. Significantly higher compression ratios are achieved by multi-dimensional approaches. Generally, the best results have used either multi-dimensional wavelets (Mallat, 1999) or its successor curvelets (Cands and Donoho, 1999). Villasenor et al. (1996) showed that compression ratios of 100:1 were achievable by compressing a 4-D volume (t, h_x, h_y, s). Further, Villasenor et al. (1996) states that the header information was the limiting factor in achieving even higher ratios.

Subsurface offset gathers potentially represent even higher, up to six, dimensional data. To test compressibility, I used a 4-D volume (x, y, z, h_x) of dimensions (32,32,400,64). Figure 1 shows one of these subsurface offset gathers and its neighbors, note the similarlity. Following Villasenor et al. (1996), I chose the 9/7 bi-orthonormal transform (Antonini et al., 1992) used in JPEG compression. Figure 2 shows the resulting transform space and a histogram of the absolute values. I then used several different thresholds throwing away 90%, 95%, 98%, and 99% of the data in the wavelet domain respectively. Figure 3 shows the result of transforming these thresholded volumes back into the space-domain. The resulting images are near-perfect at 95% and potentially acceptable at 98%. This translates into an acceptable compression ratio of approximately 30:1.

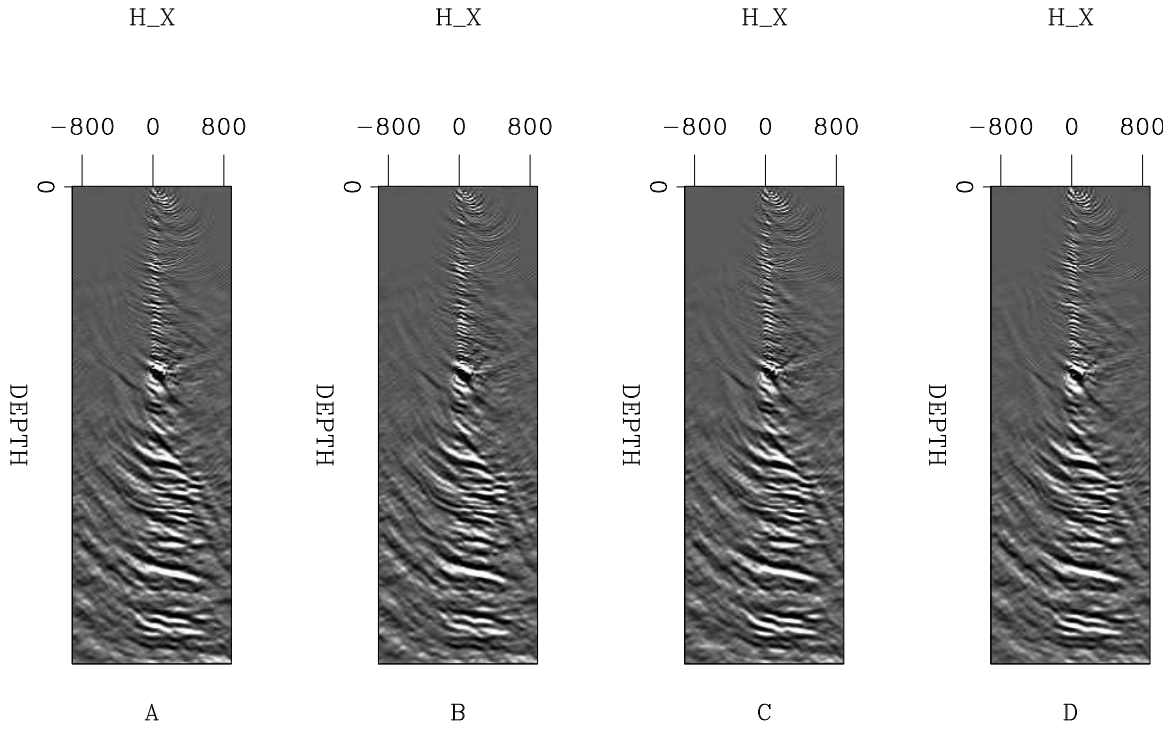


Figure 1: Five neighboring subsurface offset gathers. B and C are one midpoint in X before and after A. E is one midpoint in Y before A. Note the spatial similarity, which lends itself to compression. [**ER**] bob1/. raw

COMPRESSIVE SENSING

Compressive sensing is a statistical technique attributed to Donoho (2006), but whose start could be placed as early as the basic pursuit work of Mallat and Zhang (1993). A compressive sensing problem at its heart is a special case of a missing data problem. In geophysics, we often think of a missing data problem as solving for a model $\mathbf{m}$ given some data $\mathbf{d}$ which exist in the same vector space. We have a masking operator $\mathbf{R}$ (1 where the data is known, 0 elsewhere). We add in some knowledge of the covariance of the model through a regularization operator $\mathbf{A}$. We then estimate the best model from the following system of equations in a ℓ_2 sense,

$$\begin{aligned} \mathbf{0} \approx \mathbf{r_d} &= \ell_\mathbf{2}(\mathbf{d} - \mathbf{Rm}) \\ \mathbf{0} \approx \mathbf{r_m} &= \ell_\mathbf{2}(\mathbf{Am}), \end{aligned} \tag{3}$$

where $\mathbf{r_d}$ and $\mathbf{r_m}$ are the result of taking the ℓ_2 norm of the first and second equations. The success of this approach relies on the accuracy of $\mathbf{A}$ to describe the covariance of the model.

Compressive sensing approaches the problem from a different perspective. It starts from the notion that there exists a basis function that $\mathbf{d}$ can be transformed into through the linear operator $\mathbf{L^T}$ in which very few non-zero elements are needed to represent the signal. The compressive sensing approach is then to set up the missing data problem in two phases.

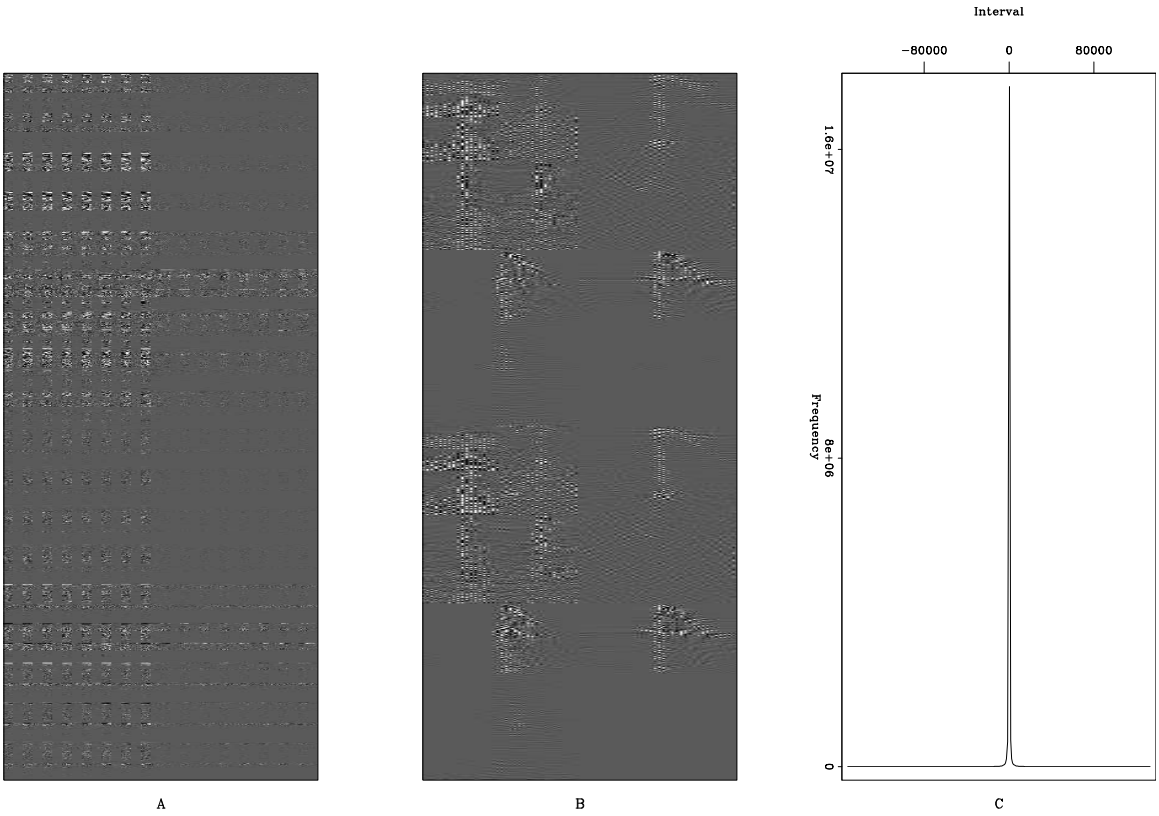


Figure 2: Panel A shows the wavelet domain representation of the 4-D volume used in this experiment. Panel B shows a zoom into a portion of the wavelet domain. Panel C shows a histogram of the wavelet domain values. Note how the vast majority of the values are nearly zero. [**ER**] bob1/. wavelet1

First, estimate the elements of the sparse basis function $\mathbf{m}$ through,

$$\mathbf{0} \approx \mathbf{r} = \ell_\mathbf{1}(\mathbf{d} - \mathbf{RLm}), \tag{4}$$

where we are now estimating $\mathbf{m}$ in the $\ell_\mathbf{1}$ sense. We can then apply $\mathbf{L}$ to recover the full model. The magic of compressive sensing is that you only need to collect a small multiple, typically 4-5, more data points than the number of non-zero basis elements. In the case of correlation gather compression this would indicate collecting in the range of 5% of the correlations should be sufficient to recover the entire model, much smaller than what the Nyquist-Shannon (Nyquist, 1928) criteria would suggest.

STOMP ALGORITHM

The most difficult challenge in solving a compressive sensing problem is often finding an effective ℓ_1 (preferably ℓ_0) solver. These solvers generally have difficulty converging and are notoriously slow. The large size of any cross-correlation volume (10e7 to 10e8 elements) or, with a compressed basis still 10e5-10e6, make most ℓ_1 solver approaches impractical. Donoho et al. (2006) noted this problem with many ℓ_1 approaches and suggested an alternative, Stagewise Orthogonal Matched Pursuit (StOMP).

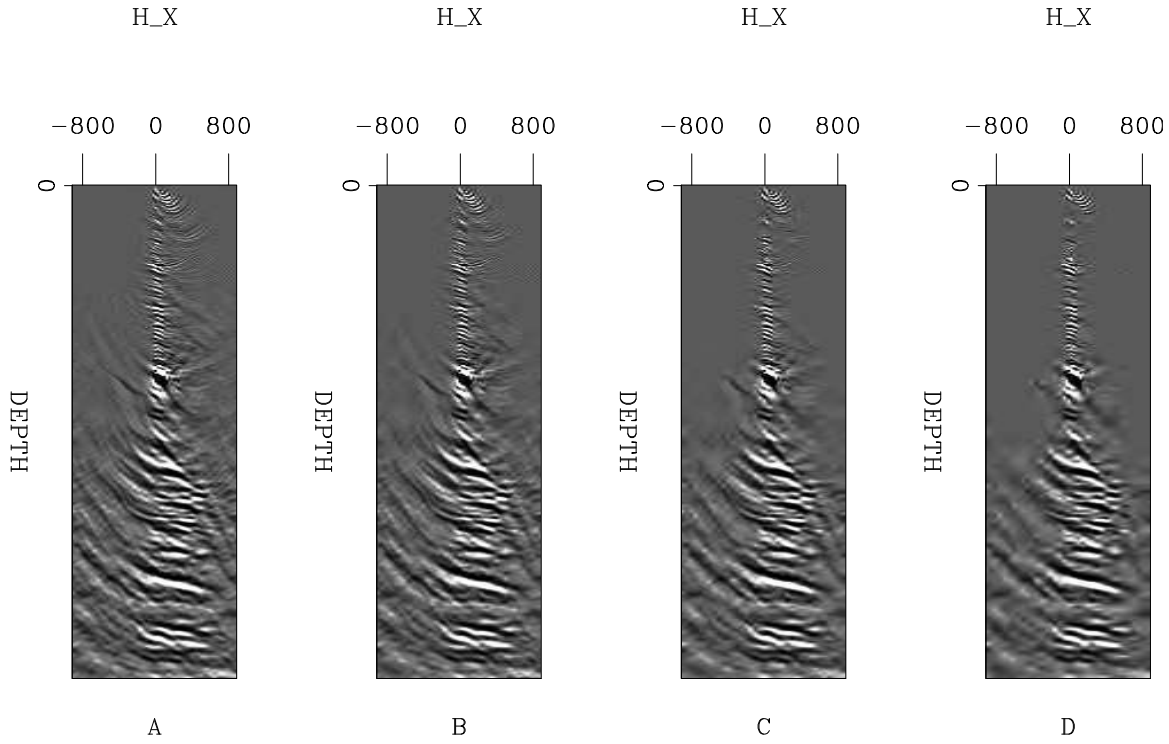


Figure 3: The result of zeroing the smallest values of the wavelet domain representation shown in Figure 2A. All four panels show the same subsurface offset gather shown in Figure 1. A shows the result of clipping 90% of the values; B, 95%; C, 98%; and D,99%. Note how the reconstructed gather is nearly identical up to a 98% clip. [**ER**] bob1/. offset

StOMP attempts to leverage the power of ℓ_2 solver with the ℓ_1 properties of approaches such as basic pursuit (Chen et al., 1998). The basic idea is to iteratively select the most important model elements by mapping into the model space and selecting model locations with the largest amplitudes. An ℓ_2 inversion problem is setup allowing only the locations selected in the previous step to be non-zero. The problem is mapped back into data space, the residual is recalculated, and the process is repeated. Algorithm 1 describes the approach in more detail.

As a first attempt to use this method, I randomly subsampled a 4-D cr-correlation space (z, h_x, cmp_x, cmp_y) by a factor of 25 to form $\mathbf{d}$. I used the multi-dimensional wavelet compression operator $\mathbf{W}$ described in Clapp (2011), choosing 4 levels in z, 3 levels in h_x, 2 levels in cmp_x and cmp_y as $\mathbf{L}$. Figure 4 shows two Common Reflection Point (CRP) gathers in the subsurface offset and angle domain. Figure 5 shows a portion of the wavelet domain representation of the cross-correlation gathers, note how few samples are non-zero.

The "art" of StOMP is choosing the number of iterations and the clipping scheme to use to select important basis elements. In this example, I took a rather crude approach. I did StOMP iterations increasing the number selected non-zero elements to approximately match the level of sparsity observed in the wavelet domain representation. Figure 6 shows the same two subsurface offset gathers and their corresponding angle gathers. I used a large value clipping scheme that guaranteed 4% non-zero basis elements after 10 iterations Clapp

Algorithm 1 StOMP algorithm

residual $\mathbf{r} = \mathbf{d}$
non-zero elements $\mathbf{M} = \text{EmptySet}$
model $\mathbf{m} = \mathbf{0}$
operator $\mathbf{L}$
for StOMP iterations **do**
  $\mathbf{g} = \mathbf{L^T r}$
  Meaningful non-zero elements $\mathbf{N} = \text{Clip}(\mathbf{g})$
  $\mathbf{M} = \mathbf{M} \cup \mathbf{N}$
  $\mathbf{m} = (\mathbf{M^T L^T L M})^{\mathbf{2}} \mathbf{M L^T r}$
  $\mathbf{r} = \mathbf{r} + \mathbf{LMm}$
end for

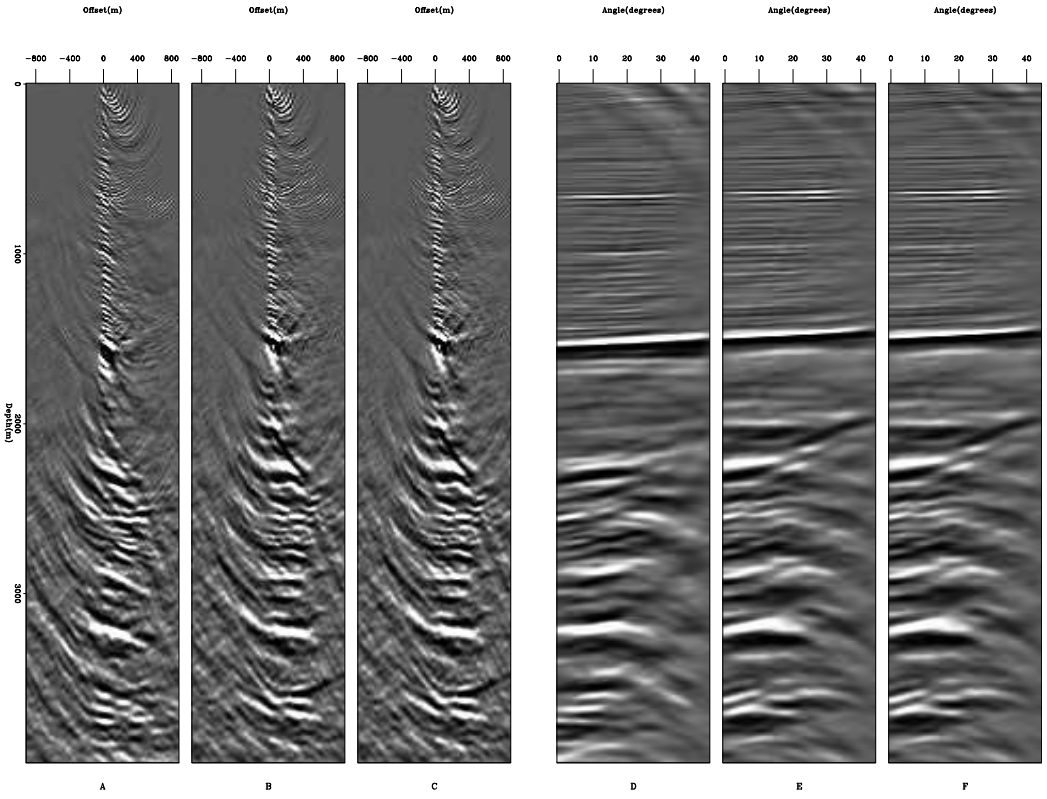


Figure 4: A, B,C and show two subsurface offset gathers from a 4-D volume. D, E, and F show the same to CRP gathers in the angle domain. [**ER**] bob1/. data-init

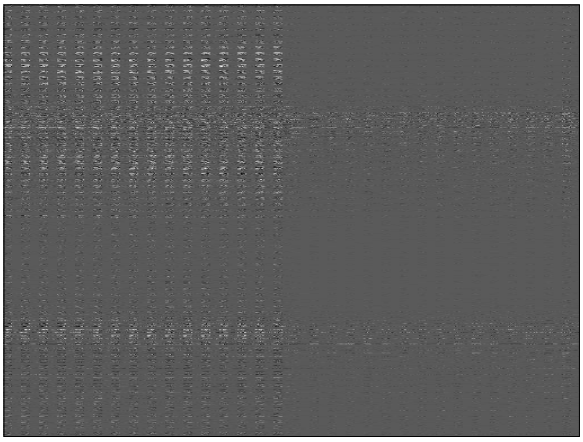


Figure 5: The wavelet domain representation of the subsurface offset domain 4-D volume partially shown in A and B of Figure 4. [**ER**] bob1/. wavelet

(2011) showed that subsurface offset gathers could be compressed by a factor of 100 and still retain most relevant information. Note that though there is visible noise in the subsurface offset gathers the angle domain representation is quite similar. Figure 7 shows the resulting wavelet domain which appears to be significantly dissimilar to Figure 5. From C and D of Figure 6 however it appears that most of this difference doesn't map into coherent events in the angle domain.

PHASE ENCODING

Phase encoding is another powerful technique to reduce data size. The basic idea of phase encoding is to sum several independent experiments together each scaled by a different multiplier. This multiplier could be something as sophisticated as a Gold code (Gold, 1967) or as simple as scaling the different experiments by -1, 0, or 1. The idea is that the combining of the different experiments will not, or will minimally, affect the final model. This "magic" is achieved because the encoded experiments only interact slightly when applying the operator to produce the model or through inversion techniques that attempt to undo the encoding. Examples of the first approach in geophysics can be seen in phase encoded migrations, such as plane wave migration (Whitmore, 1995), or more general phase encoded migrations(Shan, 2008). Examples of using phase encoding in inversion can be seen velocity estimation (Guerra, 2010) did inverse imaging (Tang, 2011; Leader and Almomin, 2012).

For the correlation gather construction problem, I attempted to encode multiple different correlations into each data sample. In terms of operators, we can think of the subsampling of correlations as applying a subsampling operator **S** to all possible correlations **d**. In the phase encoded case, we are going to add another operator **P** that first sums a number of different correlations together and then subsamples them leaving a new dataset **SRd**. This in turn changes the operator **L** in algorithm 1 to **PW**. For this test, I combined 20 different correlations in a random pattern to form each data point. Figure 8 shows the same three offset and angle gathers seen in Figures 4 and 6. Note the noticeably better job recovering the deeper portion of subsurface offsets. Figure 9 shows the angle gathers from the fully sampled correlation gathers and the subsampled, phase encoded gathers muted to the believable angle range. The gathers with the notable exception of more lower frequency noise in the recovered gathers.

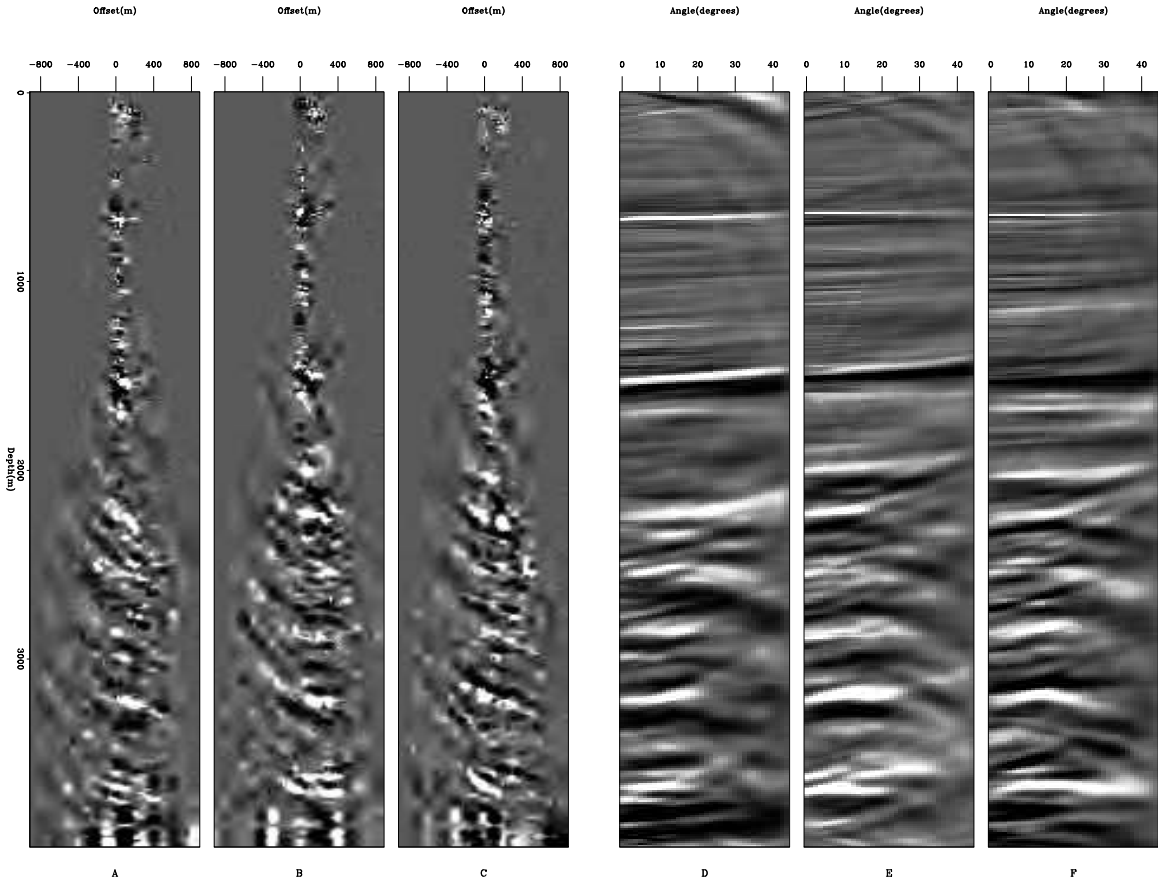


Figure 6: A, B, and C show the recovered subsurface offset gathers using compressive sensing as shown in plot A, B, and C of Figure 4. D, E, and F show the corresponding angle gathers. Note how even though A, B, and C are significantly different than A, B, and C of Figure 4, D, E, and F contain virtually the same information. [**ER**] bob1/. data-0

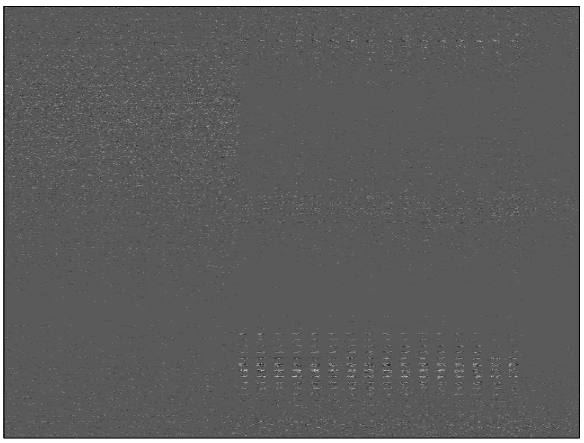


Figure 7: The wavelet domain signal after 10 iterations of StOMP. Note the dissimilarity with the correct wavelet domain show in Figure 5. [**ER**] bob1/. wavelet-try1

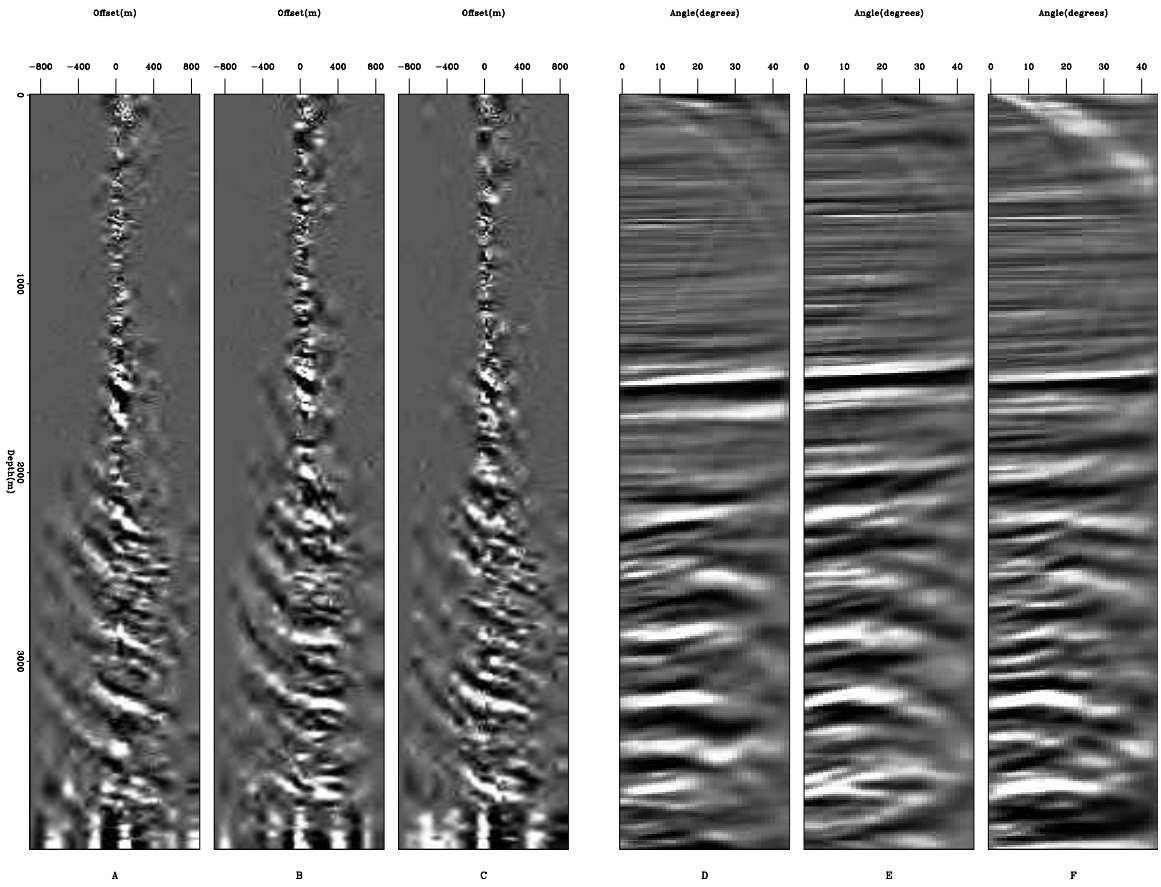


Figure 8: A, B, and C show the recovered subsurface offset gathers using compressive sensing as shown in plot A, B and C of Figure 4 and 6. D, E, and F show the corresponding angle gathers. Note the noticeable improvement in A, B, and C compared to the data shown in Figure 6. [**ER**] bob1/. data-2

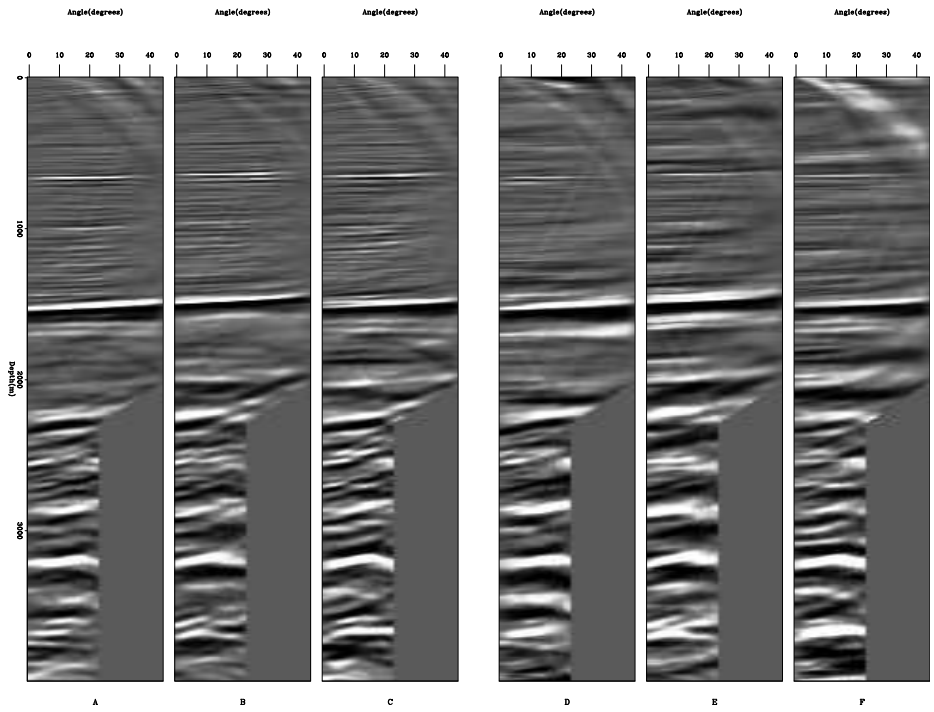


Figure 9: A comparison of the angle gathers after muting of the fully sampled correlation gathers (A, B, and C) and the phase encoded, subsampled correlation gathers (D, E, and F). [**ER**] bob1/. compare

DISCUSSION AND CONCLUSIONS

Creating full 3-D angle gathers through cross-correlation gathers is currently computationally impractical given the volume size that needs to be read/written for every shot. Compressive sensing offers a potential solution to this problem by collecting a subset of the correlation gathers and then forming the entire volume after all shot contributions have been stacked. The StOMP algorithm appears to be an effective method to obtain a sparse basis function necessary for a successful compressive sensing effort. Phase encoding to encode multiple correlations into every data point appears to offer an improved result. Further work on full 3-D shift gathers is needed to prove the feasibility of the method.

REFERENCES

Antonini, M., M. . Barlaud, P. Mathieu, and I. Daubechies, 1992, Image coding using wavelet transform: IEEE Transactions on Image Processing, 205–220.

Cands, 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.

Chen, S. S., D. L. Donoho, Michael, and A. Saunders, 1998, Atomic decomposition by basis pursuit: SIAM Journal on Scientific Computing, **20**, 33–61.

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

Clapp, R. G., H. Fu, and O. Lindtjørn, 2010, Selecting the right hardware for reverse time migration: Leading Edge, **29**, 48–58.

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.

Gold, R., 1967, Optimal binary sequences for spread spectrum multiplexing (Corresp.): IEEE Transactions on Information Theory,, **13**, 619–621.

Guerra, C., 2010, Migration-velocity analysis using image-space generalized wavefields: PhD thesis, Stanford University.

Leader, C. and A. Almomin, 2012, How incoherent can we be? Phase encoded linearised inversion with random boundaries: SEP-Report, **147**, 149–158.

Mallat, S. G., 1999, A wavelet tour of signal processing: Academic Press.

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

Micikevicius, P., 2009, 3D Finite Difference Computation on GPUs using CUDA: GPGPU, **2**.

Nemeth, T., J. Stefani, W. Liu, R. Dimond, O. Pell, and R. Ergas, 2008, An implementation of the acoustic wave equation on FPGAs: SEG Expanded Abstracts, **27**, 2874–2878.

Nguyen, A., N. Satish, J. Chhugani, C. Kim, and P. Dubey, 2010, 3.5-D Blocking Optimization for stencil computations on modern CPUs and GPUs: Super Computing 2010, 1–13, IEEE Computer Society.

Nyquist, H. L., 1928, Certain topics in telegraph transmission theory: Transactions American Institute of Electrical Engineer, **47**, 617–644.

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

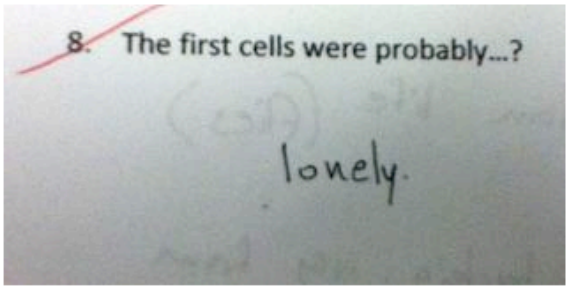
Sava, P. C. and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074.

Shan, G., 2008, Imaging of steep reflectors in anisotropic media by wavefield extrapolation: PhD thesis, Stanford University.

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

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.

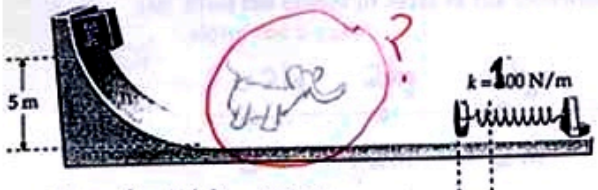
Whitmore, N. D., 1995, An imaging hierarchy for common angle plane wave seismogram: PhD thesis, University of Tulsa.



2. A 3-kg object is released from rest at a height of 5m on a curved frictionless ramp. At the foot of the ramp is a spring of force constant k = 100 N/m. The object slides down the ramp and into the spring, compressing it a distance x before coming to rest.

10 (a) Find x.

5 (b) Does the object continue to move after it comes to rest? If yes, how high will it go up the slope before it comes to rest?



$U = 3(9.81)(5) = 147.15$

$U_s = \frac{1}{2}(100)x^2 = 50x^2$ ...?

NO. there is an elephant in the way.

0

Linearised inversion with GPUs

Chris Leader and Robert Clapp

ABSTRACT

Graphical Processing Units (GPUs) can provide considerable computational advantages over multi-core CPU nodes or distributed networks by locally accelerating certain types of floating point operations. However, when processing and inverting exploration scale seismic datasets we encounter two key problems - compounded disk IO (explicit routing through the host is necessary) and the relatively small memory provided by the GPU ($\leq$ 6 Gbytes, restricting model sizes that can be allocated). As shown in an earlier discussion the IO bottleneck on the adjoint side can be somewhat circumvented by using random domain boundaries. Herein will be discussed how the forward modelling routine must be adapted to create an adjoint pair such that least-squares iterative inversion can be performed. We will then analyse how domain decomposition and P2P communication can be used to propagate over larger model sizes in such a way that communication can be effectively hidden and subsequently we can observe linear scaling.

INTRODUCTION

For smaller-scale research institutions, which may not have access to high performance computing facilities, processing terabytes of seismic data can be a significant challenge if attainable at all. As discussed in Ohmer et al. (2005) and Foltinek et al. (2009), GPUs can greatly assist any operation that can be considered as Single Instruction Multiple Data (SIMD) by running thousands of independent threads concurrently across the 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) and Leader and Clapp (2011) discuss how Reverse Time Migration (RTM) can be adapted to minimise disk access during propagation and hence better harness the computational power of the GPU without sacrificing significant performance for data movement. This paper will look at extending this system to inversion and also at how larger model sizes can be used. From here on a basic familiarity of GPU memory hierarchies and their uses will be assumed, one can to refer to Micikevicius (2009) or Leader and Clapp (2011) for more in depth discussions of these attributes.

GPU BASED RTM

Reverse Time Migration (Baysal et al., 1983) is increasingly becoming the industry and academic standard for seismic imaging. The full treatment of the wave equation provides us with accurate kinematic and amplitude information. Very few assumptions about the data (maximum dips, single scattering etc.) are necessary, relative to one-way wave equation or Kirchhoff imaging techniques. However, this process is computationally demanding and images still exhibit many artifacts due to the relatively simple nature of the imaging condition (Sun and Zhang, 2009).

We can describe an idealised modelling procedure as in equation 1, which is based on the first approximation of the Born scattering series. RTM is then the adjoint of this process, equation 2. 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 a single scattering, our propagator makes no such assumption and multiple scattering / diving waves are still positioned correctly.

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

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

Wave propagation can be performed very efficiently on GPUs by taking advantage of shared memory and read redundancy along the y axis (Micikevicius, 2009). The main challenge for GPU based RTM comes from the complex conjugate in equation 2. Since equation 2 is expressed in frequency, this conjugate reverses the sense of time of the data d relative to the source function f. As such, the source wavefield (4D, in this case) is forward modelled and saved, and then read back to the GPU at each imaging time step while back propagating the recorded data for correlation. Typically this source wavefield is several hundred gigabytes, meaning disk storage is unavoidable (without compression). In the case where we must transfer the source wavefield from disk our GPU based RTM scheme is far from optimal, as all computational advantages are now being counteracted by compounded disk access.

The reason we cannot simply back propagate the source wavefield is due to the fact that we artificially remove boundary reflections during forward modelling. This violates the conservation of energy, and these source wavefields are consequently not time reversible. A solution for this is to pad our domain with random boundaries (Clapp (2009); Fletcher and Robertsson (2011); Shen and Clapp (2011)); boundary reflections are now incoherently scattered, and since we are not removing any energy it is time reversible, to within computational precision. Our multiplicative imaging condition and subsequent stacking across the shot axes reduce any residual boundary noise to imperctible levels after around 50 or so shots / realisations. For a detailed discussion of how to set up these boundaries and their scattering properties one can refer to Clapp (2009) or Leader and Clapp (2011).

Relative to source saving RTM we need to perform an extra propagation - the reversal and back propagation of the source wavefield, whereas previously we would just read back

the appropriate wavefield slice. Fortunately, this extra computation is far faster than the wavefield reading scheme. In the case whereby the entire source wavefield is held by the CPU, and only CPU-GPU transfers are needed, the random boundary method is about 25% faster (Figure 1). For this to be the case the problem has to either be very small or the wavefield must be compressed or decimated, which will take extra computation. Once we have to use the disk we see the random boundary scheme is several times faster.

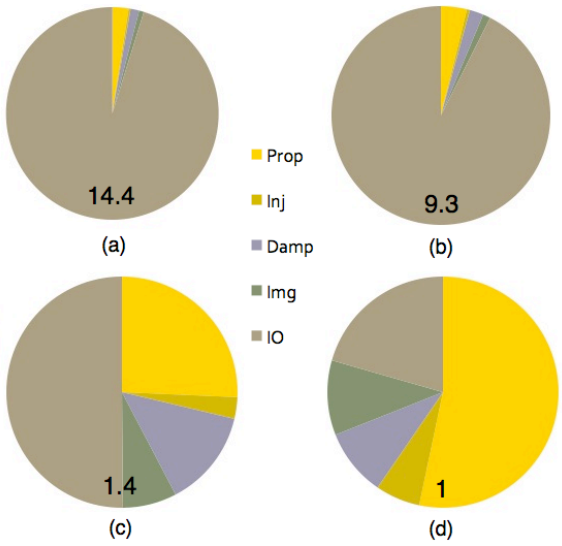


Figure 1: Relative speeds for RTM implementations for propagation, imaging, damping, injecting and IO. The bold numbers are normalised elapsed times. (a) shows the most naive disk set up, (b) an optimised disk set-up with asynchronous transfer, (c) the case where the 4D wavefield can be held by the CPU memory and (d) random boundaries. [**NR**] chris1/. rtmsp

GPU BASED LINEARISED INVERSION

GPU propagation augmented with random boundaries means no IO is required during propagation of either wavefield, making this a very computational effective scheme (relative to source saving.) However RTM images are often still artifact laden - low frequency artifacts can be prevalent near salt, acquisition footprints are noticeable, resolution is decreased (since we are squaring the source function) and random boundary noise may remain. All of these imperfections can be reduced by extending this scheme to least-squares inversion.

The inverse of our data modelling system can be approximated by constructing the adjoint pair of such a system and iterating in a least-squares sense. This is known as linearised inversion since we are assuming we know the kinematic model (the background velocity) and are trying to retrieve the high frequency (perturbation) part of the Earth model. This amounts to describing our slowness function as $s^2(\mathbf{x},\mathbf{y},\mathbf{z}) = b(\mathbf{x},\mathbf{y},\mathbf{z}) + m(\mathbf{x},\mathbf{y},\mathbf{z})$. As an algorithm it can be described by Algorithm 1. Here F is our operator, F' its adjoint, r the data-space residual, m the current model estimate, d_{obs} the input data, gg the gradient and rr the back-projected residual. The stepper then updates the current model estimate and data-space residual. Since we are not changing our boundaries between iterations, we first forward our source wavefields and save the final slices necessary for back propagation. These can then be back propagated along with the receiver wavefield in the $gg = F'r$ step of the algorithm.

The adjoint process, F', is identical to the aforementioned RTM scheme with random boundaries. Now we must try and implement the forward process, F, in the most computationally efficient way. Initially we must think about our stencil. To take the spatial

Algorithm 1 Linearised inversion with random boundaries

```
Initialise
Create random boundaries
Propagate source wavefield, save final slices
r = Fm − d_obs
while iter < niter; iter++ do
  gg = F′r
  rr = Fgg
  (m, r) =stepper(m, r, gg, rr)
end while
Output m
```

derivative of the wavefield we convolve the data with a star-shaped stencil, for an 8th order scheme in 3D this stencil is 25 points in total, as each direction shares the centre value. Previously we only needed to define the velocity value at the centre of our stencil, as the forward procedure was spraying the information to neighbouring grid cells. The adjoint of this propagation requires velocity values to be defined along the entirety of the stencil, as we now need the information from all points being calculated to carry to all new points. This means that in addition to allocating our wavefield values in the GPU kernel, we must also allocate the same number of velocity points.

Previously we were using texture memory to store our velocity values to take advantage of caching and automatic boundary control features (Leader and Clapp, 2011). What we can now do is save the relevant velocity values to shared memory, taking care that we do not saturate the memory. For large stencil sizes this can require the use of smaller blocks. By defining our local velocity array in the same way that we do our wavefield values, we can still take advantage of shared memory. This extra memory allocation and computation causes this adjoint propagator to be about 81% the speed of forward propagation. By only using texture memory we see this drop to around 43%, and by using global memory this decreases to 38%. On a Tesla GPU this final speed would be slower again, since Fermi cards provide some global L1 and L2 cache options. In order to make the system fully adjoint the source wavefield must be propagated through the same random boundary as in the corresponding RTM, but we can damp the data wavefield. Since we are now using equation 1 we have no time reversal problems, because the sense of time for both the source wavefield and the data wavefield are the same.

The stepper is performed on the CPU. The reasons for this are twofold - firstly we want to transfer our data and image back to the CPU for writing purposes, so doing the model update on the CPU creates no unnecessary data transfer. Additionally, our solver is just a series of vector operations (dot products and subtractions) and such operations can not be accelerated by the GPU. They no longer have a SIMD corollary and global memory operations on the CPU are faster than on the GPU, often by a factor of 2 or 3.

Results from this linearised inverse scheme can be seen in Figure 3, in which we are attempting to recover the reflectivity model shown in Figure 2. Here we have 60 inline shots at a spacing of 100m and two crossline shots at a spacing of 1km, receivers are in a dense 825x200 grid. We can clearly see the inverse scheme improving the resolution of the top of the salt and mitigating the associated low-frequency artefacts.

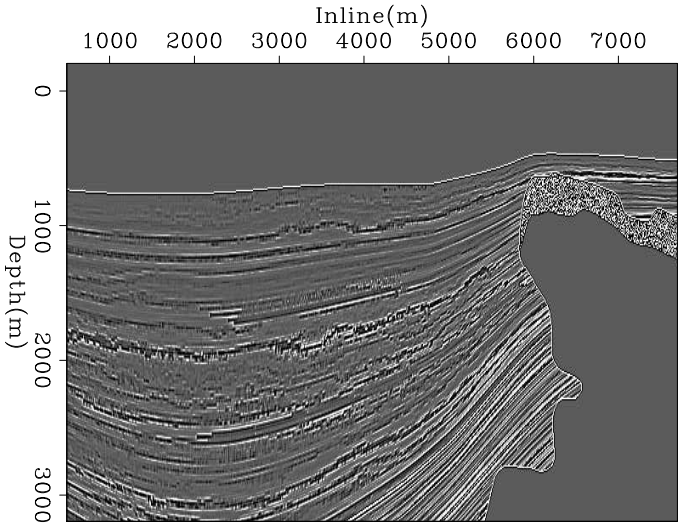


Figure 2: A 2D slice from the reflectivity model that we are attempting to recover. [**ER**] chris1/. srfl

DYNAMIC BOUNDARIES

As in Figures 2 and 3, we see inversion improve the images considerably, particularly when dealing with low frequency artifacts. However the remaining random boundary artifacts still seem to stack out fairly slowly, at a rate of about $\sqrt{N}$, where N is the iteration number. One option we have is to change these boundaries as a function of iteration, now we would expect to see a quicker reduction in boundary artifacts between iterations. This can be done by seeding our random boundary by iteration number, as well as shot position.

Algorithm 2 Linearised inversion with random boundaries

```
Initialise
r = Fm − d_obs
while iter < niter; iter++ do
  Create random boundaries
  Propagate source wavefield, save final slices
  gg = F′r
  rr = Fgg
  (m, r) =stepper(m, r, gg, rr)
end while
Output m
```

Our algorithm becomes Algorithm 2, giving us an additional propagation per iteration (recalculating the random source wavefield slices). For a typical model size this will increase computation time by around 14%. When doing phase encoded linearised inversion we can get this for free, since recalculation of the initial residual is needed during each iteration (Krebs et al., 2009). This system can be referred to as dynamic random boundaries, as opposed to static random boundaries. Furthermore our operator is now non-linear since we have altered our velocity function and hence our operator. This means theoretically we now have to use a non-linear solver. However since the operator difference is only manifested in the image noise there are some cases where a conjugate direction solver gives acceptable convergence characteristics. We see this system become useful in areas of poor

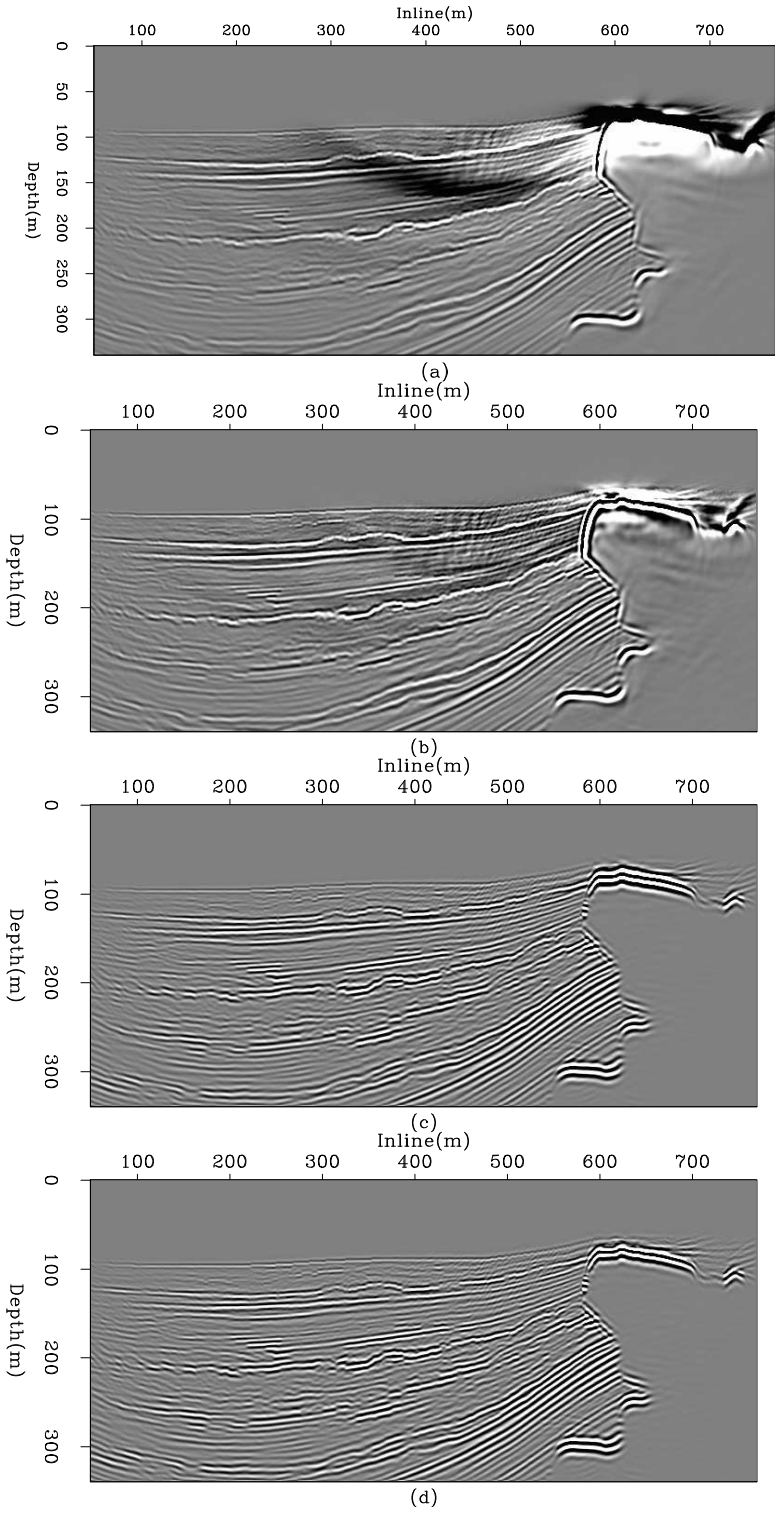


Figure 3: RTM and linearised inversion example 2D slices. (a) shows the raw RTM result, (b) raw inversion after 5 iterations, (c) RTM with lowcut bandpass filter and (d) inversion after 5 iterations and lowcut filter. [**CR**] chris1/. imgc

shot sampling, where boundary artifact stacking-out can be slow, however in many cases the extra computation does not seem to outweigh improved convergence. As a function of iteration number we see slightly better residual performance (when using steepest descent for both), but when we look at this as a function of cost we see very little difference. Then, by comparing static boundaries with a conjugate direction solver to dynamic boundaries with a steepest descent solver, we get worse performance with dynamic boundaries as a function of cost.

It should be noted that comparing the l_2 norm of the residuals in this case is slightly misleading, since the high frequency noise we have slightly reduced is not well represented by this measure. When looking at the images more differences are notable than implied by this scalar fitting methodology. Additionally if we used a better non-linear solver we would fully expect to see more comparable performance.

DOMAIN DECOMPOSITION

In the case where our model size exceeds that of our computing system we must break up our domain. When using Fermi GPUs the global memory is 6 Gbytes, when using Tesla GPUs we are confined to 4 Gbytes. These limited memories severely restrict models that can be allocated. For propagation we have to allocate a minimum of three 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 RTM 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 often concerned with models more at the scale of 1000 pts^3, and for wide azimuth data sets we may have dimensions larger than this. The solution is to decompose our domain, as in Figure 4. Performance does not strongly depend on the axis over which we decide to break our domain, but for current algorithm design cutting the domain along the y axis is the most effective. We want to only cut along one dimension, rather than say break our domain into cubes along two dimensions, as this minimises the quantity of halo transfer needed (Figure 4).

We need our sub-domains to overlap along the y-axis by half the stencil length, else information would be lost between domains. Between time steps this halo region must be transferred and then synchronised before moving to the next time step. CUDA 4.0 and above used with Fermi cards allows for Peer to Peer (P2P) GPU communication and data transfer. Previously it was necessary to explicitly route all information transfer between GPUs first through the CPU, which was significantly slower. 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. In fact it is now possible to dereference pointers on other GPUs or on the host, by virture of this UVA. Arrays can not span GPUs, however.

The main technique of making such a method scale linearly is by using asynchronous memory copies and kernel calls. When doing this it is possible to overlap communication with computation, hiding halo communication time. 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. By restricting halo computation and communication to one such stream, and the other data

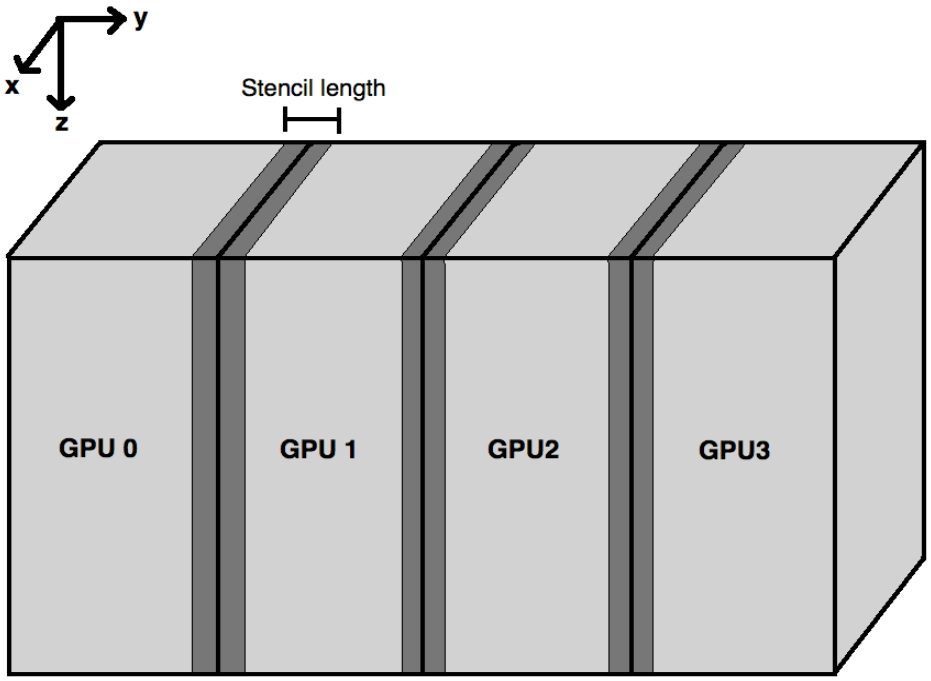


Figure 4: A diagram of how to decompose the model, dark grey regions are allocated on both neighbouring GPUs [**NR**] chris1/. domdec

(internal) computation to another, we can hide the halo communication. 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=0; 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; i_gpu++){
   cudaMemcpyPeerAsync(...halo_stream[i_gpu]);
}
for(i_gpu=0; i_gpu < n_gpus; i_gpu++){
   cudaDeviceSynchronise();
}
```

The key part to remember here is that kernel calls and async calls can overlap, providing they are defined to separate streams. So initially we loop through devices, then each one will

calculate its halo region data. Once this is done, each GPU will move onto the internal data calculation. Whilst this is being calculated, each GPU moves onto the three communication loops. First off, each GPU sends its halo data to the right and receives from the left, then each device is synchronised, then each device sends to the left and receives from the right. Now each GPU has the relevant halo information for the next time step. Finally we synchronise to ensure all internal data computation is done before moving to the next time step.

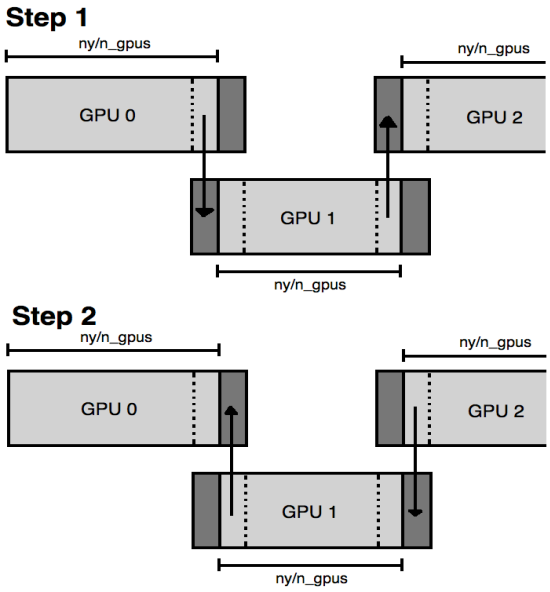


Figure 5: A diagram of the two stages of GPU halo communication, simplified to 1D [**NR**] chris1/. halos

The GPU devices are linked by PCIe switches and these switches are duplex, meaning each GPU can send and receive simultaneously - providing they are in different directions. This means a given device can send data to the GPU on its left and receive from the GPU on its right at the same time. We use the stream synchronise call because the moment a device tries to send and receive in the same direction the switch can stall, costing time. A pictorial demonstration of this is show in Figure 5.

Such a system assumes we have neighbouring GPUs on the active node in question, linked by PCIe switches. If one has GPUs on separate nodes (such as in a distributed network) MPI must be used for the halo communication. Unsurpisingly this is slower, albeit only by a factor of around 1.5x.

For acoustic propagation our internal computation time is always in far excess of halo transfer time (by an order of magnitude or so), meaning communication is always hidden and our domain decomposition scales linearly with model size, hence it is optimised. For TTI propagation this communication time approaches that of the internal data calculation, meaning for highly elastic and anisotropic media care must be taken to optimise the decomposition to ensure one is never waiting for communication. The moment we are waiting for information transfer we are not fully harnessing the potential of the GPU.

CONCLUSION

We conclude that GPUs can be effective when running linearised inversion. By storing a velocity stencil in shared memory, as well as the wavefield values, we can perform accelerated adjoint propagation. By augmenting this with a random boundary based RTM scheme and a CPU based model stepper we can perform least squares iterative linearised inversion very efficiently, with little time lost for data movement.

At the point when our domain size exceeds our GPU memory it is possible to decompose this domain across multiple devices, whether one has multiple GPUs per node, multiple nodes with single GPUs, or any combination thereof. By making halo communication calls overlap with internal data computation it is largely possible to hide all communication, meaning our time scaling with model size is still linear.

Combining these schemes gives us the potential to perform large scale inversion at a high level of fine grain parallelism.

ACKOWLEDGMENTS

Thanks to Paulius Micikivicius at NVIDIA for his invaluable help with GPU troubleshooting, his prompt response with domain decomposition queries and his willingness to share codes.

REFERENCES

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

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.

Krebs, J. R., J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse, 2009, Fast full-wavefield seismic inversion using encoded sources: Geophysics, **74**, WCC177–WCC188.

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

Micikevicius, P., 2009, 3D finite difference computation on GPUs using CUDA: GPGPU, **2**.

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

Shen, X. and R. G. Clapp, 2011, Random boundary condition for low-frequency wave propagation: SEG Technical Program Expanded Abstracts, **30**, 2962–2965.

Sun, J. and Y. Zhang, 2009, Practical issues of reverse time migration: True amplitude gathers, noise removal and harmonic-source encoding: ASEG Expanded Abstracts.

How incoherent can we be? Phase-encoded linearised inversion with random boundaries

Chris Leader and Ali Almomin

ABSTRACT

To perform linearised inversion on seismic exploration scale datasets we are continually looking for methods to accelerate computation and reduce data handling overhead. One option to accelerate reverse time imaging is to use random domain boundaries for the source wavefield computation, alleviating much of the required IO in favour of some additional computation. Additionally, data handling problems can be addressed by phase-encoding data (weighting, shifting, summing) and then inverting for a common model between realisations. Both random boundary and phase-encoding methods rely on wavefield incoherency during correlation and stacking to build a clean image. Here we investigate if these can be effectively used together, or if these techniques combined create wavefields that are too incoherent, slowing convergence as a function of cost when compared to linearised inversion without phase-encoding. We show that by using multiple realisations per iteration we can improve convergence and create cleaner reflectivity images.

INTRODUCTION

Reverse time migration (RTM) can provide accurate subsurface images because it applies the full, two-way wave equation. Thus steep dips, multiples, and prismatic waves can be imaged. However, RTM is the adjoint of an idealised modelling operator and not a full inverse operation, meaning that images can suffer from artefacts such as acquisition footprints, low-frequency noise and decreased resolution. We can approximate the inverse of this modelling procedure by using iterative least-squares inversion. However this can quickly become prohibitively expensive as each iteration is roughly twice the cost of a single migration. Furthermore, the formulation of RTM (the adjoint procedure) requires the source-side wavefield to be modelled, reversed, and sequentially correlated with the receiver-side wavefield. This modelling and reversal provides computational difficulties, as we have to save and re-inject a 4D source wavefield when applying 3D RTM. These two problems - inversion cost and source wavefield time reversal, - can be solved by taking advantage of correlation attributes and data redundancy.

When modelling the source wavefield we have to use a finite computational domain. This domain creates artificial boundary reflections, which must be removed else they create high-amplitude, coherent artefacts within the image. However, removing these boundary reflections causes the modelling to be non-reversible, thus the entire wavefield must be saved and reused when correlating with the recorded data. We can use random boundaries (Clapp (2009); Fletcher and Robertsson (2011); Shen and Clapp (2011)) to make this computation time-reversible. By only saving the final wavefield snapshots we can now back propagate the

wavefield to within numerical accuracy. Provided that boundary reflections are sufficiently incoherent, the RTM imaging condition and subsequent stacking over shots can reduce any residual incoherent noise to an imperctible level. Such a method is particularly useful for GPU computing. Here data must be read from disk, to the CPU, and then sent to the GPU, compounding any disk access. By removing the need for disk-saved source wavefields we accelerate GPU based RTM significantly.

We can address the inversion cost in a slightly different way. One method is to reduce the data size that we are imaging by combining sources. This can be done by shifting, weighting and summing shots to create one or several 'super shots' (Morton and Ober (1998); Romero et al. (2000)). The weights and shifts that we apply to individual shots are referred do as the encoding. Such a method can also be used for full waveform inversion (Gao et al. (2010);Krebs et al. (2009)). We can either combine sources into a single super shot or several super shots. When combined into one super shot the inversion is now independent of the number of sources, reducing the inversion cost by roughly this number (assuming full aperture for all shots). However, many crosstalk artefacts are seen when wavefields from different source experiments correlate coherently. These are slowly reduced when iterating, but by changing the encoding between iterations these are suppressed much faster. Romero et al. (2000) and Krebs et al. (2009) show that by using a single sample random encoding the best convergence rates are seen. The caveat of such a scheme is that we must recalculate the initial residual each time (since we have changed our observed data), making this method about 1.5 times more expensive.

Both random boundaries and phase-encoding can be very effective in accelerating linearised inversion. However, they both introduce a considerable amount of noise into the system and rely on correlation, stacking and inversion to reduce this. By combining these methods we may be making the system too incoherent, slowing the inversion process down as a function of cost. Here we investigate convergence properties of these techniques and how we can try to create cleaner gradients within each iteration.

INVERSION WITH RANDOM BOUNDARIES

Under a single Born scattering assumption we can describe the recorded data as

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

The adjoint process can thus be written as

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

Where d is the data, G_0 are the respective Green's functions, m the model, $\mathbf{x}$ the three-dimensional model coordinate, $\mathbf{x}_{r,s}$ the three-dimensional source and receiver coordinates, ω temporal frequency, $f(\omega)$ the source waveform and $*$ denotes the complex conjugate. It is this complex conjugate that reverses the sense of time, meaning that in the time domain we back propagate the source and receiver wavefields and correlate them at each imaging time step. This complex conjugate and the subsequent source wavefield modelling create the need for either saving this wavefield or forming a time-reversible source wavefield. Random

boundaries scatter this wavefield incoherently whilst adhering to the conservation of energy; after propagation correlation and stacking will reduce residual noise.

Random boundary noise in an RTM image is stacked out at a rate of $\sqrt{N}$, where N is the number of sources; often we see better performance than this. By using a different random boundary for each shot experiment we see random noise levels at an acceptable level after combining around 50 shots. Some effects of this can be seen below, where we use a GPU based RTM algorithm and model and migrate 100 shots over a section of the SEAM velocity model (Figure 1), simulating a marine survey. We then perform 5 iterations of least-squares linearised inversion with the same dataset. Here we have 50 inline shots at 100m spacing, 2 crossline shots with 1km spacing and 825x200 receivers. Comparison of such a scheme compared to source saving are shown in Clapp (2009).

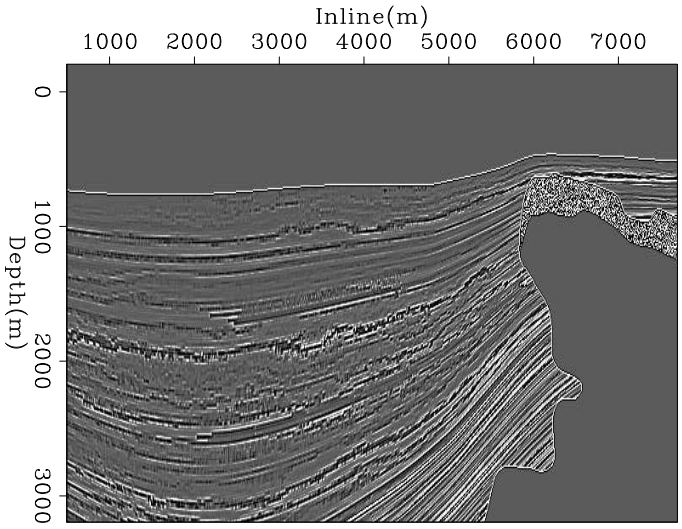


Figure 1: A 2D slice from the reflectivity model that we are attempting to recover. [**ER**] chris2/. srfl2

As expected, in Figure 2 we see image quality improve when extending this procedure to an iterative least squares inversion. The footprint from the limited acquisition begins to dissipate and we see generally higher frequency content and more balanced, geologically accurate amplitudes. However, since only a limited number of iterations have been performed some artefacts remain, from both the random boundaries and from the inverse system.

When blindly extending this process to an inversion, relatively poor convergence can be seen, especially with respect to the data space residual. The reason for this can be seen when looking at this residual. The random boundary image (the gradient, in this case) features noise in the water column and a source location imprint artefact. When modelling data over this image we see an artefact manifested as a direct arrival from the source self-correlating. This can be removed by either muting the water column in the image or by time muting the remodelled data before back-propagation. An example of this under a phase-encoded setting is shown in Figure 3. Furthermore, we can improve convergence by changing the random boundaries as a function of iteration number, as well as a function of shot position.

The computational advantage for GPU based RTM is considerable. We can consider three scenarios - when the source wavefield must be entirely saved to disk, when the source wavefield can held in the CPU memory and when using random boundaries. One should note the last of these requires an extra computation - the source wavefield back propagation

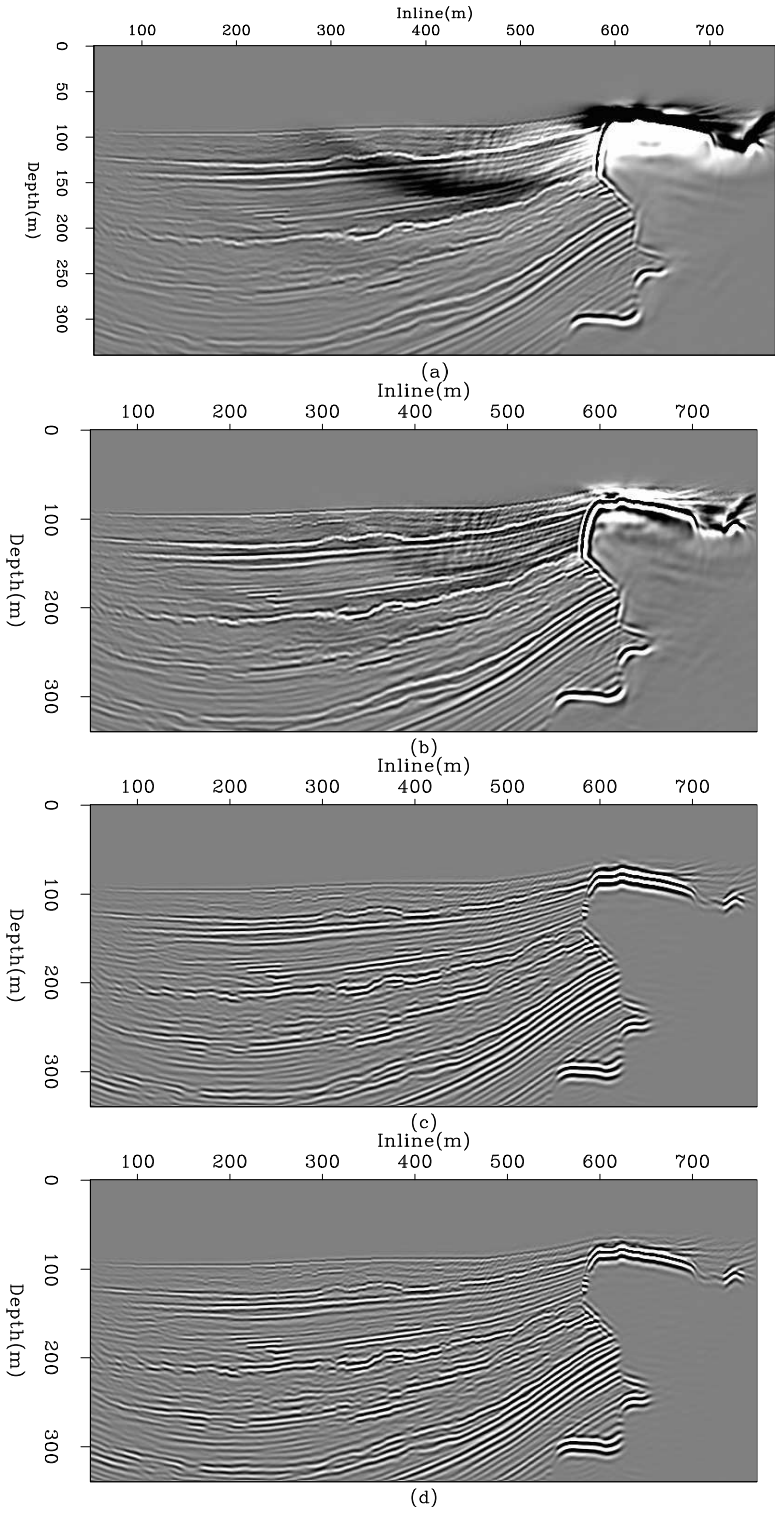


Figure 2: RTM and linearised inversion example 2D slices. (a) shows the raw RTM result, (b) raw inversion after 5 iterations, (c) is (a) after a lowcut wavenumber filter and (d) is (b) after the same lowcut filter. [**CR**] chris2/. imgc2

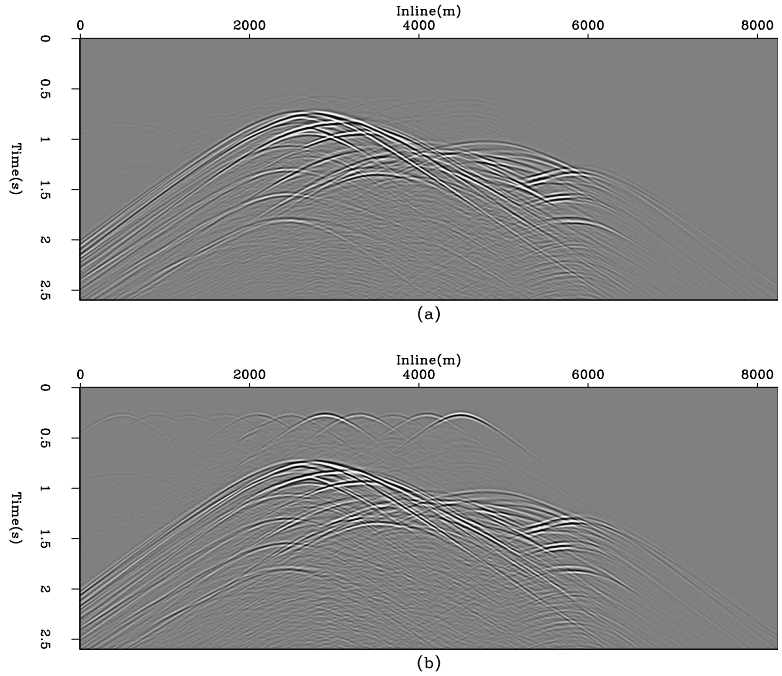


Figure 3: An example of source self-correlation. a) shows the real data after encoding 10 shots. b) shows the modelled data when the water column noise has not been muted. Artifacts similar to direct arrivals not present in a) can be seen. [**CR**] chris2/. datac

during RTM. In total computation time asynchrous disk wavefield transfer is nine times slower than random boundary RTM, and if the source wavefield can be compressed and stored on the CPU memory this is 1.5 times slower. Of course, these conclusions are strongly related to the speed of the disks being used.

PHASE ENCODING WITH RANDOM BOUNDARIES

Phase encoding can have multiple meanings; herein we describe the process of weighting and combining shots to reduce data size. We do this by creating a matrix of weights and applying it to our data.

$$\widetilde{d}(\mathbf{x}_r, p_s, \omega) = \sum_{\mathbf{x}_s} \alpha(\mathbf{x}_s, p_s) d(\mathbf{x}_r, \mathbf{x}_s, \omega) \tag{3}$$

Now $\widetilde{d}$ denotes our phase encoded data, α is a sequence randomly selecting either 1 and -1, and p_s is some sort of realisation index. When combining all shots together p_s will be 1. For the forward process we propagate a source function encoded with the same sequence α.

Augmenting random boundary linearised inversion with phase encoding requires some additional thought. In the case where we combine all shots to one super shot, we are now propagating 100 weighted shots through the same random boundary. On a shot by shot basis this is acceptable, as each wavefield will be incident on the boundary at a different angle

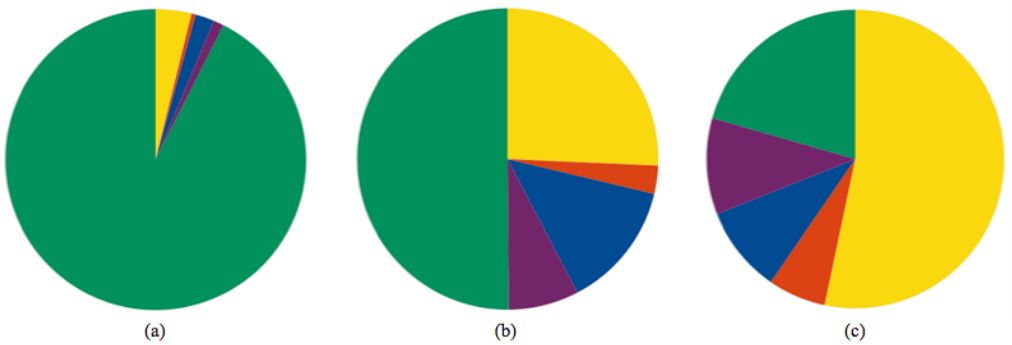


Figure 4: Compute time comparisons for three RTM regimes. Green denotes IO, yellow propagation, blue damping, red injection and purple imaging. (a) uses asynchronous disk transfers, (b) assumes the entire source wavefield can be held in CPU memory and (c) uses random boundaries. [**NR**] chris2/. timec

and hence scatter differently. However, when performing this conventionally we formulate a gradient for each shot separately and then sum them to create our final gradient, reducing our noise by $\sqrt{N}$. Furthermore, there are only two coherent wavefields (the source and receiver) that can correlate with the scattered field to induce noise. When combining all shots together, we do not quite see this behaviour because we now have every scattered field (unique per shot) correlating with every scattered field and with every coherent field, of which there are now $2N$. Fortunately, the correlation of scattered fields should also be a random walk reduction at $\sqrt{N}$. Typically with phase encoded linearised inversion after *nshots* iterations one expects to see an acceptably clean image. When including random boundaries we see a noisier image, as expected, but all key features are present. Images can still be artefact-laden, especially in areas of low illumination, but this is typical of both random boundary and phase-encoded imaging when done independently.

We have several options to mitigate this. The best results are seen by combining several random subsets of shots, calculating a gradient for each subset and combining these to form the gradient for a single iteration. Whilst this appears more computationally demanding, the fact that single super-shot GPU based inversion is not further parallelisable over shots makes this approach seem more appealing. Over a node containing 8 Fermi M2090 cards we can create the gradient for 8 different super shots and combine these for little extra time cost, giving a cleaner gradient per iteration. We now see more favourable convergence characteristics and slightly cleaner images. With naive phase encoded inversion we see a convergence to 60% within the data-space residual norm after 100 iterations with 50 combined shots. With a water column mute applied we see convergence to 56% and with multiple, stacked realisations we see convergence to 53%. Once the data error is below 50% we tend to see very gradual improvements for all situations. One full, non-encoded iteration with random boundaries (roughly the same cost) takes us to 74% data error.

Figure 6 compares the results of conventional, separated linearised inversion with phase encoded inversion, over the 3D model shown in Figure 5. Here we had 120 shots in total, 60 inline at 100m spacing and two crossline at 1km spacing. Again, receivers were in a 825x200 grid. Images (a) and (c) are equivalent in computational cost, as are images (b) and (d). The first noticeable aspect are the low-frequency artefacts present in the non-encoded image.

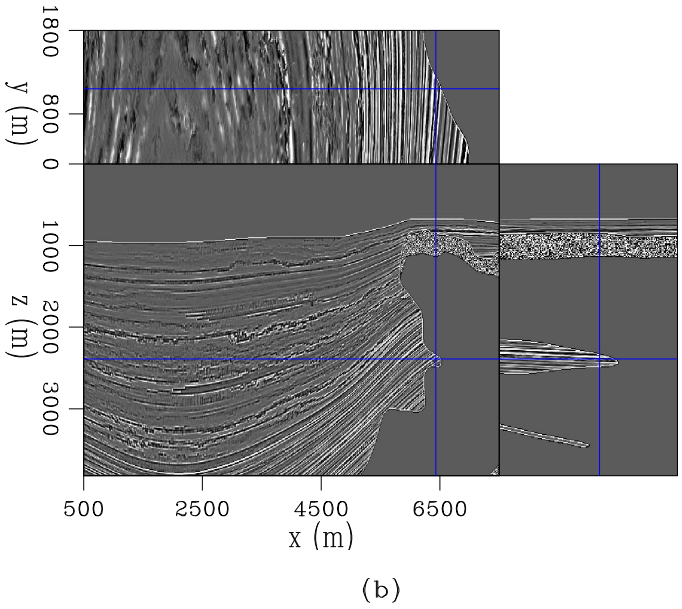


Figure 5: The reflectivity model that we are inverting for. [**ER**] chris2/. sicr

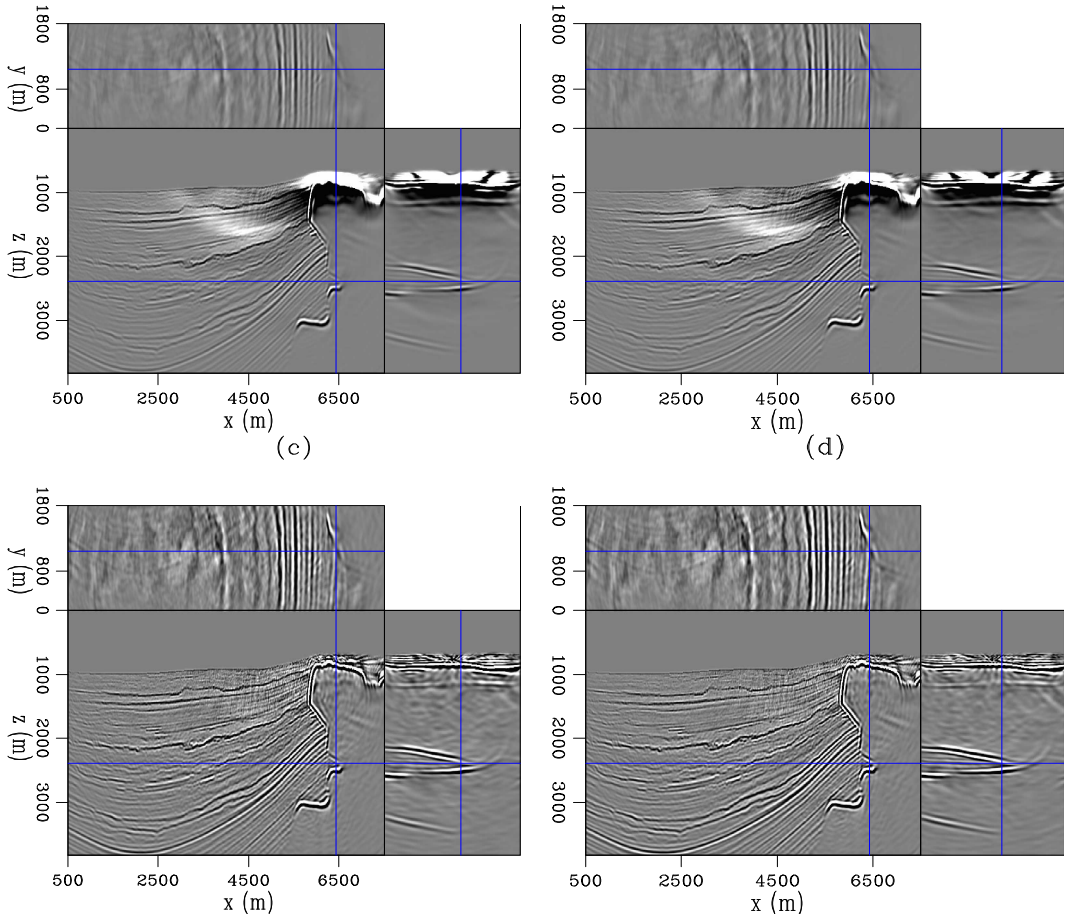


Figure 6: Linearised inversion examples. (a) shows the model after one iteration of conventional inversion, (b) the model after two iterations, (c) phase encoded inversion after 80 iterations - equivalent cost to (a) and (d) phase encoded inversion after 160 iterations - equivalent cost to (b). [**CR**] chris2/. sic

These occur due to wavefields moving in the same direction correlating and are prevalent over high-reflectivity, high-contrast features such as salt boundaries. Often the first several iterations of linearised inversion will work on removing these artefacts before focusing on other areas, as can also be seen in Figure 2. The second noticeable aspect is that the frequency content of the phase encoded image is much higher. The resolution, especially on the salt edges, is greater; high-frequency noise is also present, but this is expected. The additional iterations performed here have successfully begun to remove the effect of the source wavelet in the image, whereas in the RTM image this wavelet is squared. Figure 7 shows these same images after a low-cut bandpass filter and some light amplitude gain. The filter has removed the low-wavenumber noise, but at the expense of the vertical salt edge. The conventional images look much improved, but the resolution and general content of the phase-encoded images still seem preferable.

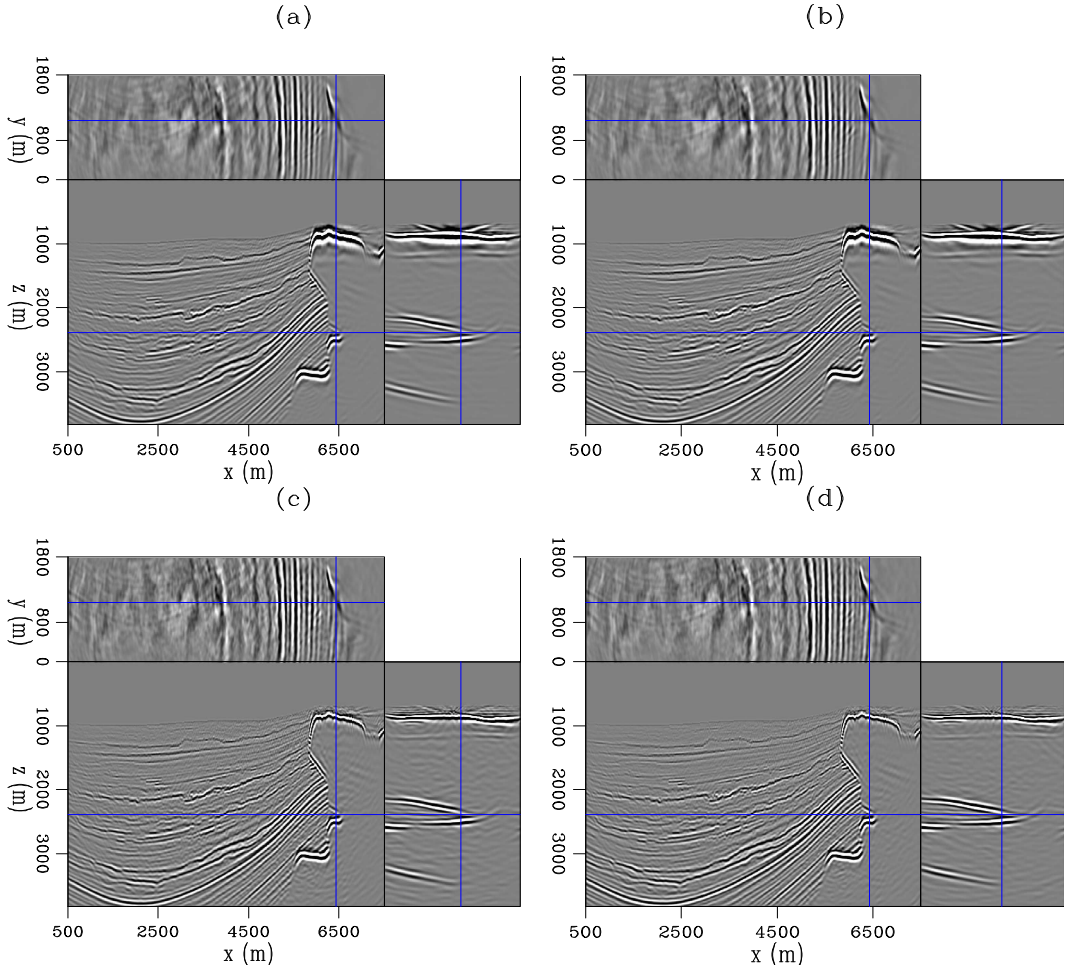


Figure 7: The same set of results as in Figure 6 but each with a low-cut spatial frequency filter and some amplitude gain. [**CR**] chris2/. sicc

Figures 6 and 7 show that for equivalent cost, even when augmented with random boundaries, phase encoded linearised inversion can yield high quality images. This scheme is also incredibly well-adapted for GPU computing - there is no IO during propagation in the forward or adjoint scheme, and the objects we need to copy to the GPU are all the size of the model, the size of one shot, or smaller.

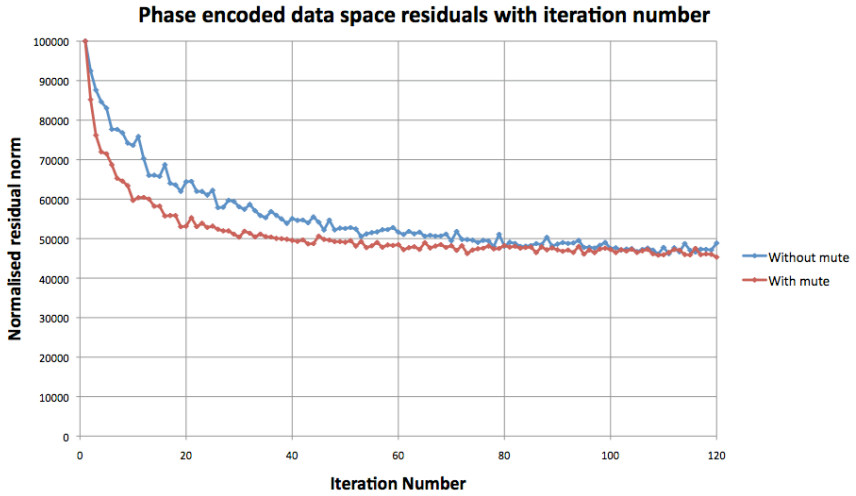


Figure 8: Normalised residual evolution as a function of iteration number. [**NR**] chris2/. peres

Residual behaviour with iteration number can be seen in Figure 8. As a reference point for separated inversion after two iterations the respective normalised residual difference norms (normalised to 100) were 88.9 and 79.9, which are not significantly smaller than those shown. However, in terms of cost the former of these would appear at 80 iterations, the latter at 160. With this is mind we see that our phase-encoding scheme is doing a vastly more efficient job at data fitting, in an l_2 sense. However, when interpreting these scalar fits we must consider that the l_2 residual norm does not well represent high-frequency noise, and so two images with an apparently similar residual may have quite different high-frequency noise characteristics. It is due to this that we must take the phase encoded scheme to so many iterations. Typically, we will need at least $nshots/2$ if not $nshots$ iterations for a clean image here. With conventional inversion we need 5-10 iterations to remove the low-frequency salt artefacts and many more for amplitude and acquisition imbalances.

CONCLUSIONS

Both phase encoding and random boundary propagation can be very effective in accelerating linearised inversion. We presented the individual benefits and how combining these methods can lead to a powerful inverse scheme with reference to inverse imaging on GPUs. We can conclude that these techniques can be used together, and acceptable images are obtained within the cost of a conventional RTM migration. However, to improve convergence we can see that a mute must be applied to avoid direct-arrival type artefacts in the Born-modelled data, and that we can improve convergence again by stacking separate gradient realisations per iteration. For future work we will explore the best method of cleaning the gradients between iterations and will update the non-linear solver to see if we can improve our convergence rate once we are within 50% data error.

ACKNOWLEDGMENTS

The author would like to thank NVIDIA for their continued support with troubleshooting our GPU cards.

REFERENCES

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.

Gao, F., A. Atle, and P. Williamson, 2010, Full waveform inversion using deterministic source encoding: SEG Technical Program Expanded Abstracts, **29**, 1013–1017.

Krebs, J. R., J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse, 2009, Fast full-wavefield seismic inversion using encoded sources: Geophysics, **74**, WCC177–WCC188.

Morton, S. A. and C. C. Ober, 1998, Fastershot-record depth migrations using phase encoding: SEG Technical Program Expanded Abstracts, **17**, 1131–1134.

Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, 426–436.

Shen, X. and R. G. Clapp, 2011, Random boundary condition for low-frequency wave propagation: SEG Technical Program Expanded Abstracts, **30**, 2962–2965.

FWI with different boundary conditions

Xukai Shen, and Robert G. Clapp

ABSTRACT

We compare time-domain Full Waveform Inversion using different boundary schemes: absorbing boundary condition, random boundary condition and continuation of velocity. The absorbing boundary condition requires saving the wavefield, the other two do not, but require extra wavefield modeling. The random boundary condition results in a gradient calculation that is almost as good as the absorbing boundary condition, whereas the continuation of velocity result in a gradient that has strong artifacts. However, the final inversion results using different boundaries are similar where there is a lot of data to contrain the model. When there is not so much data to constrain the inversion, results from using the random boundary condition are similar to those resulting from the absorbing boundary condition, but the continuation of velocity boundary condition does not work as well. We demonstrate this with synthetic examples.

INTRODUCTION

Full Waveform Inversion (FWI) (Tarantola, 1984; Pratt et al., 1998; Mora, 1987) gives more accurate subsurface velocity estimation than conventional methods of velocity estimation, such as ray-based methods (Hampson and Russell, 1984; Olson, 1984; White, 1989), especially in geologically complex areas. Traditional time-domain FWI algorithm is very computationally expensive for two reasons: first, it involves iterative forward and backward two-way wavefield propagation; second, to compute FWI gradient, at least one of the wavefield needs to be saved. Because wavefields in large-scale 3D application can require terabytes of storage, I/O can be a non-trivial bottleneck. Computational time of two-way wavefield propagation can be dramatically reduced by using unconventional hardware such as GPUs (Micikevicius, 2008), leaving the I/O cost associated with the huge wavefield size the primary bottle neck. Time-reversible boundary conditions can almost eliminate the I/O cost entirely by propagating source and receiver wavefields in the same temporal direction (Clapp, 2009; Shen and Clapp, 2011). By doing this, gradient calculation can be done on the fly, and no wavefield saving is necessary. However, to achieve good results, boundaries must be carefully designed to avoid artifacts. For example, in Reverse Time Migration (RTM) applications, random boundary conditions (Clapp, 2009) result in good final images, whereas continuation of velocity gives rise to artifacts in the final image. RTM image calculation is similar to FWI gradient calculation. However, FWI performs those calculations many more times, which makes it different from RTM as far as boundary artifacts are concerned. We first present the pseudo-code of gradient calculation using different boundary conditions; then we compare the results of a single gradient calculation and the whole inversion process for each boundary condition.

ALGORITHM DESCRIPTION

FWI as used here is a gradient-based time domain implementation (Shen, 2010). Different boundary conditions are used for the gradient calculation, leading to slight modifications of the gradient calculation algorithm. However, the steplength search uses absorbing boundary conditions in all cases. The pseudo-code of the gradient calculation using the absorbing boundary condition is as follows:

Algorithm 1 Pseudo code of gradient calculation using absorbing BC

for $is = 1, ns$ **do**
  Forward wavefield propagation, generate modeled data and record source wavefield;
  Calculate data residual using recorded data and modeled data;
  Reverse-time propagation of residual data, correlate residual wavefield with source wavefield to generate gradient;
end for

The pseudo-code of the gradient calculation using the time-reversible boundary condition (random boundary condition and continuation of velocity) is as follows:

Algorithm 2 Pseudo code of gradient calculation using time-reversible BC

for $is = 1, ns$ **do**
  forward wavefield propagation using absorbing boundary, generate modeled data, do not record wavefield;
  forward wavefield propagation using random boundary, record last two time slices of source wavefield;
  calculate data residual using recorded data and modeled data;
  reverse time propagation of residual data and source wavefield using random boundary, generate gradient on the fly;
end for

It is worth mentioning that for the MPI version of the code, wavefield propagation on different computational nodes uses different random boundary realizations, which further reduces artifacts.

GRADIENT COMPARISON

In this section and the next, we will use the 2D Marmousi velocity model to compare FWI results with different boundary conditions. The true velocity model is a modified version of the 2D Marmousi model, with 12-meter spacing in both x and z. The starting model is a smoothed version of the true model. The survey geometry simulates fixed spread land acquisition. A total of 60 shots are used with 84 m shot spacing. Figure 1 a shows the true model with continuation of velocity in the boundary region, figure 1 b shows the starting model with continuation of velocity in the boundary region, and figure 1 c shows the starting model with random velocity values in the boundary region.

Figure 2 shows gradients of the first iteration using different boundary conditions. Due to the similarity between gradient calculation and RTM image calculation, it can be seen

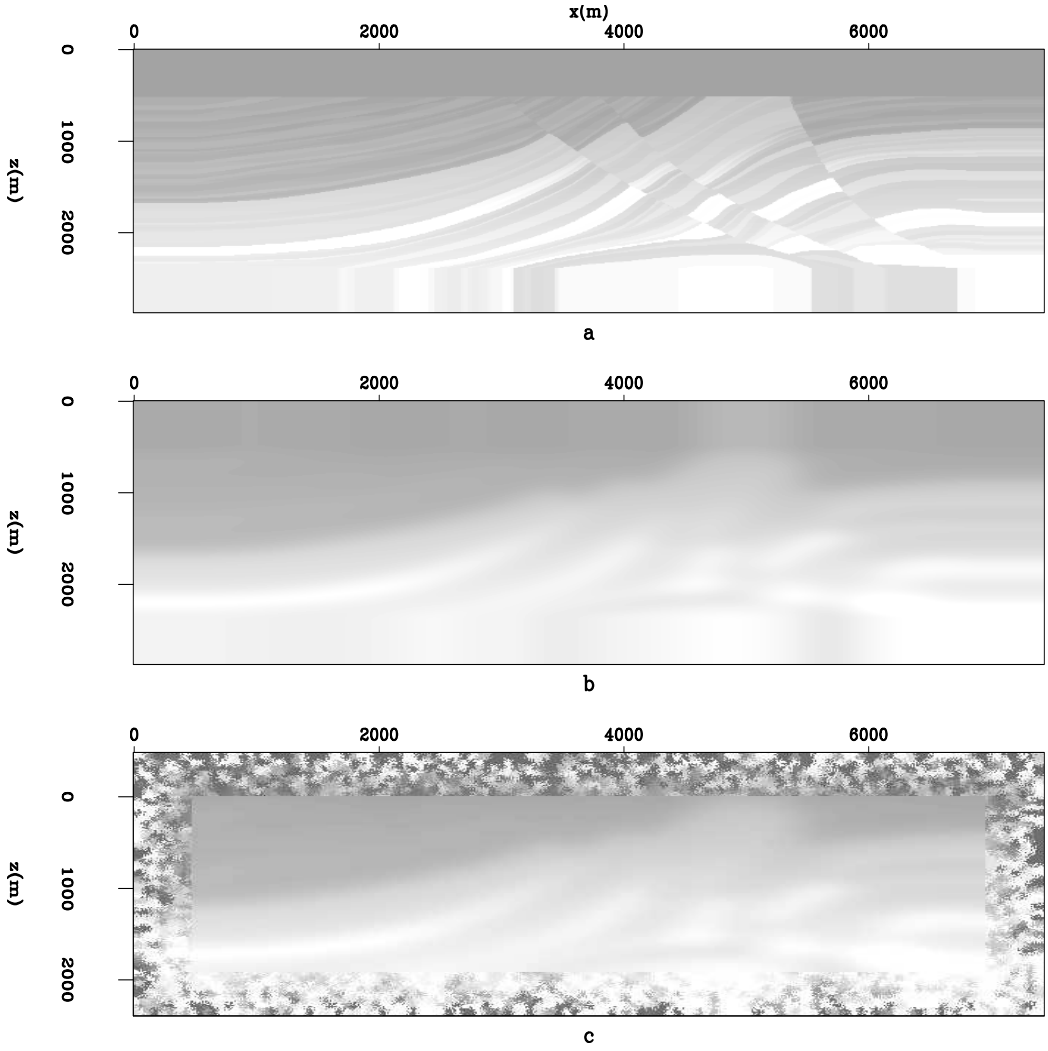


Figure 1: Different velocity models with different boundary regions: a) true velocity with continuation of velocity in the boundary region; b) starting velocity with continuation of velocity in the boundary region; c) starting velocity with random velocity values in the boundary region. [**ER**] xukai3/. vmodbw

that using a random boundary condition gives a good gradient that is very similar to the ones from using an absorbing boundary condition, except for the usual strong amplitude near source locations. The gradient calculated using continuation of velocity in the boundary region, on the other hand, has some artifacts from reflections in the boundary region. This is particularly obvious in the entire shallow depth region.

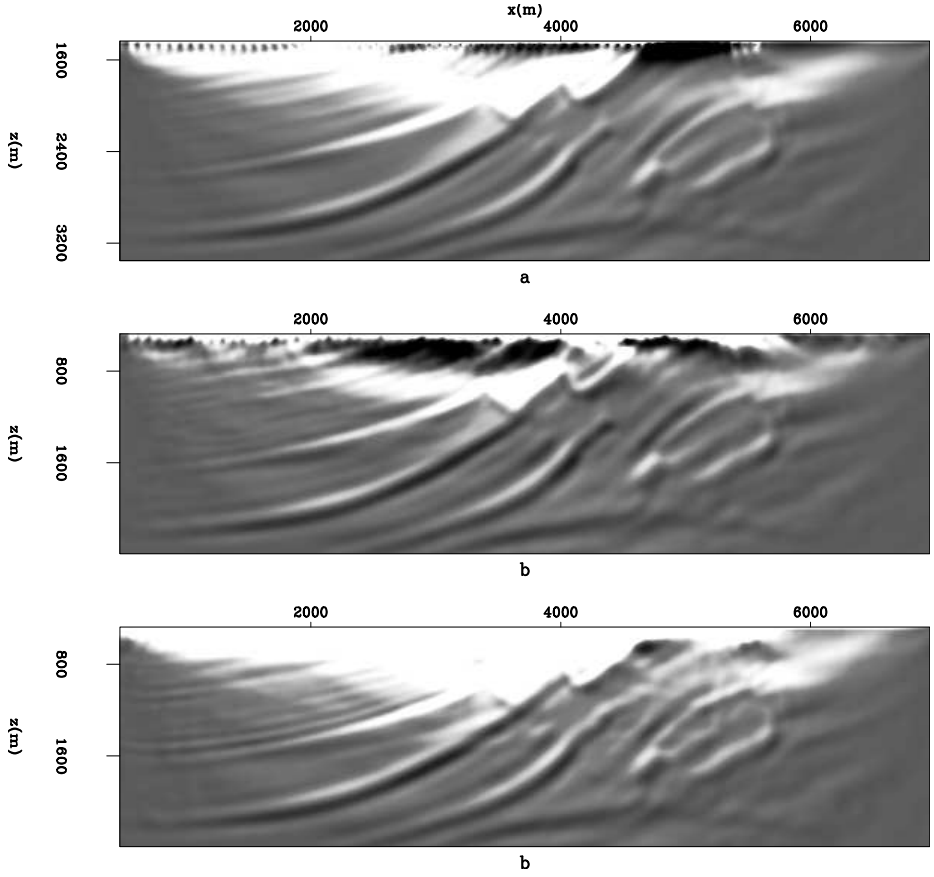


Figure 2: First iteration gradient with different boundary conditions:a) gradient with absorbing boundary condition;b) gradient with random boundary condition;c) gradient with continuation of velocity in the boundary region. [**CR**] xukai3/. gradcomp

INVERSION COMPARISON

For inversion, a total of 160 iterations are run for each boundary condition. Figure 3 shows inversion results using different boundary conditions. Inversion results from the random boundary condition and the absorbing boundary condition are very similar. Results from using continuation of velocity have more artifacts, in the side and bottom parts of the model where data fitting constraints are relatively weak. In other words, strong artifacts in the shallow part in the early iterations are eliminated by data fitting constraints as the inversion proceeds. So the artifacts pattern in the inversion results is different from the one in the first gradient.

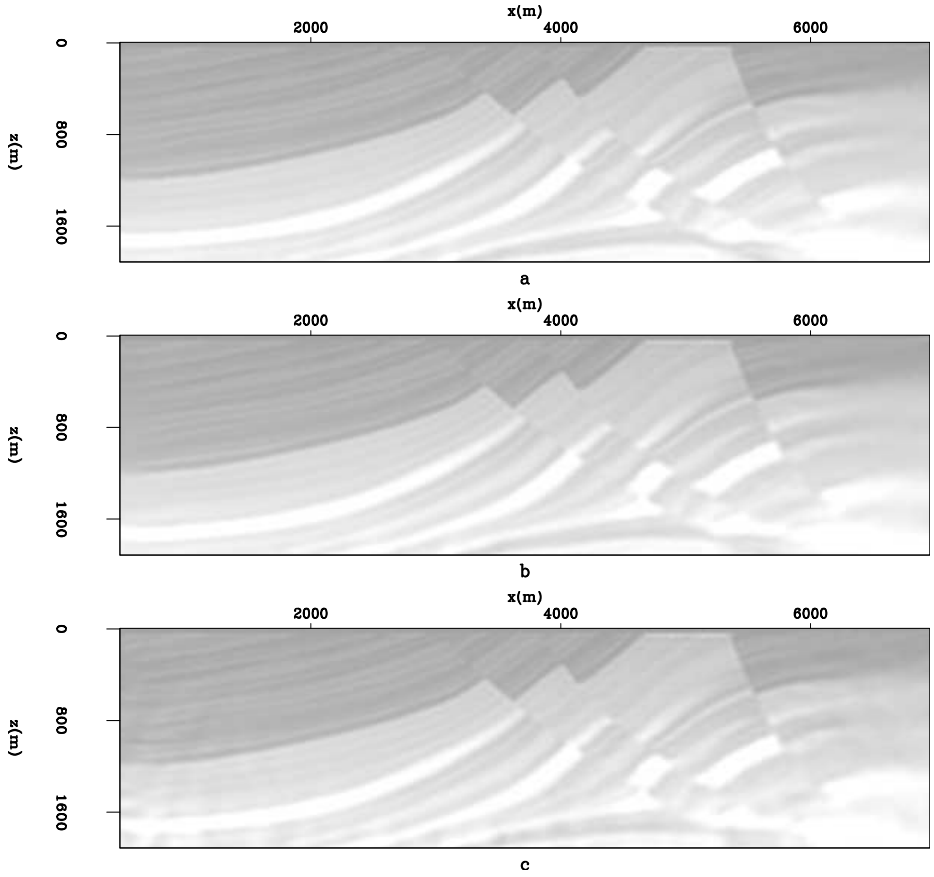


Figure 3: Inversion results with different boundary conditions: a) with an absorbing boundary condition; b) with a random boundary condition; c) with continuation of velocity in the boundary region. [**CR**] xukai3/. vinvcompbw

CONCLUSION

We compared FWI using three different kinds of boundary conditions: absorbing, random and continuation of velocity. The random and continuation-of-velocity boundary conditions eliminate almost all I/O cost associated with wavefield storage and transfer at the expense of two extra wavefield propagations per iteration. Random boundary inversion results are almost as good as those from the absorbing boundary condition, which is current industry practice. The continuation-of-velocity boundary condition, on the other hand, works quite well in well-constrained areas, and less so in poorly constrained areas This is particularly attractive on unconventional architecture where computational cost is much less than memory access cost (i.e. GPUs).

REFERENCES

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

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

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

Mora, P., 1987, Elastic wavefield inversion: Stanford Exploration Project Ph.D. Thesis.

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

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

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

Shen, X. and R. G. Clapp, 2011, Random boundary condition for low-frequency wave propagation: SEG Expanded Abstracts.

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

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

Continuous monitoring by ambient-seismic noise tomography

Sjoerd de Ridder

ABSTRACT

Seismic arrays permanently installed over a hydrocarbon reservoir can record data continuously, even in the absence of active seismic shooting. Here we study ambient-seismic noise tomography as a tool for continuous monitoring. This is of interest for both monitoring production-related changes over a long time scale (years), but it may also be a source of data for monitoring hazards over short time scales (days to weeks). We compare cross-correlations of partial recordings to cross-correlations of the full recording as a function of absolute recording time, inter-station distance and frequency. We use straight-ray tomography to image the virtual-seismic sources for travel times picked after band passes for various frequencies. The correlations converge faster for nearby receiver pairs and lower frequencies than for further receiver pairs and higher frequencies. The convergence rate also depends on the strength of the microseism energy in the ambient seismic field. Features visible in ambient-seismic noise tomography images of Scholte wave group-velocity for various frequency bands are compared to slices of a P wave velocity cube obtained from full waveform inversion of active seismic data. Clear similarities indicate that the Scholte wave group-velocity between $0.15 - 0.75$ Hz forms an image up to a depth of 240 meters. Clear similarities indicate that the Scholte wave group-velocity between $0.15 - 0.75$ Hz forms an image up to a depth of 240 meters. This study shows the feasibility of ambient-seismic noise tomography monitoring of the near surface.

INTRODUCTION

Reservoir monitoring by active seismic is a well-established technique. A single survey provides an image of the subsurface at a given time. Repetitive seismic surveying provides differential seismic images that show how the subsurface changes over time, relative to a baseline survey. At Valhall, in the Norwegian North Sea, a permanent four component (4C) Ocean Bottom Cable (OBC) array was installed in 2003 (Kommedal et al., 2004) for repeated active seismic surveying. These conventional surveys provide subsurface images after lengthy seismic acquisition, processing, and imaging procedures, at only a few snapshots in time. The OBC array is capable of continuously recording for long periods of time in the presence or absence of active seismic shooting.

Recordings in the absence of controlled-seismic shooting are referred to as passive-seismic recordings. Their use is widespread in both global- and exploration-seismology. Passive-seismic recordings contain earthquake energy of rupture events as large as the 2004 Sumatra Earthquake, $M_w = 9.1$, (Ammon et al., 2005), and as small as hydraulic induced microseismic fracture events in reservoirs, $M = -2$, (Shemeta and Anderson, 2010). Studies of earthquakes and microseismic events are of deterministic nature, aimed at characterizing the source and its wavefield. However, even in the absence of discrete deterministic events,

passive-seismic recordings contain an abundance of energy, also referred to as the Earth's Hum (Rhie and Romanowicz, 2004, 2006). These are continuous excitations of the Earth's free modes. One strong such series of excitations is the microseism energy captured in the Earth's surface interface modes. The waves in these modes are propagating randomly, as such no particular events can be distinguished. Hereafter this microseism energy will also be referred to as ambient-seismic noise.

Recordings of microseism energy can be made continuously and thus are a potential source of data for continuous reservoir monitoring. This is not only of interest for monitoring production-related changes over a long time scale, but also a possible source of data for short time scale hazard monitoring.

Passive seismic interferometry is a technique that cross-correlates ambient-seismic noise recordings at two stations to form a signal, the estimated Green's function (EGF), as if one of the stations was a seismic source (Wapenaar and Fokkema, 2006). Virtual seismic Scholte wave sources have previously been used in tomographic studies of the near surface (Bussat and Kugler, 2009; de Ridder and Dellinger, 2011). This technique allows for the design and installation of a system that exploits the ambient-seismic noise field to perform continuous, nearly real-time reservoir surveillance. The reliability of retrieved subsurface information depends on the characteristics of the ambient seismic field and its excitation sources. We need to investigate how much ambient-seismic noise is needed to cross-correlate for the retrieve of stable EGFs and reliable tomographic images.

This paper is outlined as follows. In the first section we briefly review passive seismic interferometry. Then we explore the ambient seismic microseism energy at Valhall and its capability to create low frequency virtual seismic sources. In the following section we elaborate on the filtering and correlation procedures. We then explore with how much data the correlations convergence towards the long-term correlation average. The next section contains straigth-ray tomography results at various frequency bands and for correlations of various portions of the data. The paper finishes with a discussion and some concluding remarks.

PASSIVE SEISMIC INTERFEROMETRY

To perform passive seismic interferometry, we cross-correlate passive seismic recordings between all possible station pairs in a permanently installed array. This creates a full virtual seismic survey, $\mathbf{D}(\omega) = \mathrm{D}(\omega, x_r, x_s)$:

$$\mathbf{D}(\omega) = \mathbf{r}(\omega)\ \mathbf{r}(\omega)^{\dagger}, \tag{1}$$

where $\mathbf{r}(\omega) = \mathrm{r}(\omega, x)$ is a vector containing passive seismic recordings at all stations, and $\dagger$ denotes the Hermitian (complex conjugation and transposition). However, this technique holds only when the energy in the ambient seismic field satisfies a condition known as energy equipartitioning. 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 equipartition and can be utilized for seismic interferometry (Stewart, 2006; Dellinger and Yu, 2009). At the Valhall oil field, this microseism energy is recorded by the aforementioned permanent installed seismic array (see station

map in Figure 1). The microseism energy recorded by the vertical component of the OBC consists of excitations of the Scholte-wave mode of the seafloor.

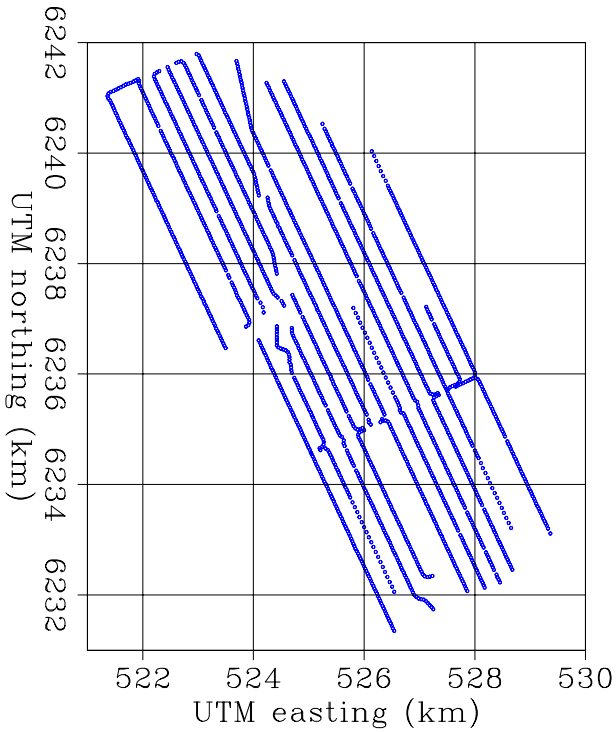


Figure 1: Station map of OCB at Valhall. [**ER**] sjoerd1/. xy-joseph

These installations are capable of making recordings continuously for long periods and under all weather conditions. One such recording was made in December 2010, spanning a little over 5 days. Figure 2 shows how the spectrum of the ambient seismic field, as recorded far away from the platform, changes over time. The spectrum is computed over 2.5 minute overlapping windows and averaged over several receivers far removed from the platform. Clearly the microseism energy strength varies over time. Figure 3 contains recordings of wave height and wind strength made at Ekofisk (NMI, 2012), another oil-field near Valhall. Although the weather station is at 32 km from Valhall, there is a clear correlation between the strength of the microseism energy and the wind strength. This corroborates atmosphere-ocean-seafloor coupling as the source of this microseism energy (Longuet-Higgins, 1950; Rhie and Romanowicz, 2004, 2006). An example of a virtual source is shown in Figure 4. This is one column in the matrix $\mathbf{D}$, for which all data are cross-correlated with the data at a station in the center of the array. Six time slices are shown, at 0, 3, 6, 9, 12 and 15 seconds. Although the energy focuses at the source at $t = 0$ seconds, there is energy at $t = 0$ in the vicinity of the master station because the virtual source is finite in bandwidth and zero phase. The wavefront is dispersive: lower frequencies propagate faster the higher frequencies. The virtual source radiates very homogeneously in all directions, reflecting that the energy in the ambient seismic field, averaged over 5 days, was travelling in about equal strength in all directions. The dispersion in the virtual sources can be made visible by transforming the virtual sources from the (t, x) domain to the (ω, p) domain and stacking all virtual sources in one line of the Valhall array together. The dispersion image is shown in Figure 5. For most of the frequency range, the virtual sources are dominated by a single dispersive wave mode.

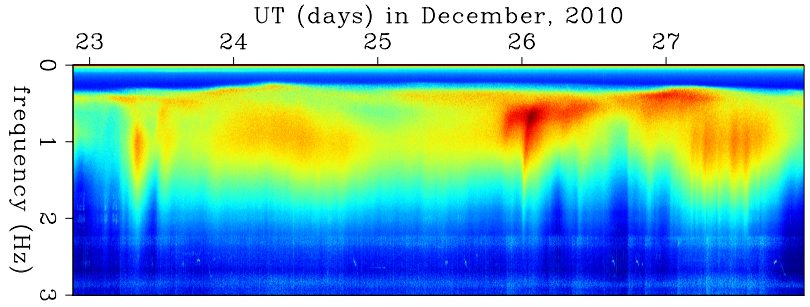


Figure 2: Spectrum of the ambient seismic field, computed over 2.5 minute overlapping windows. The spectrum is averaged over several receivers far removed from the platform and displayed as a function of time. [**CR**] sjoerd1/. joseph-spectra-low

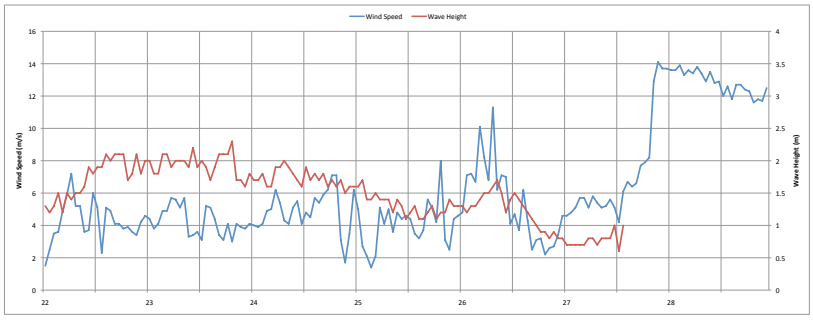


Figure 3: Wind-speed and wave-height measurements at Ekofisk, another oil field near the Valhall field (NMI, 2012). [**NR**] sjoerd1/. weather

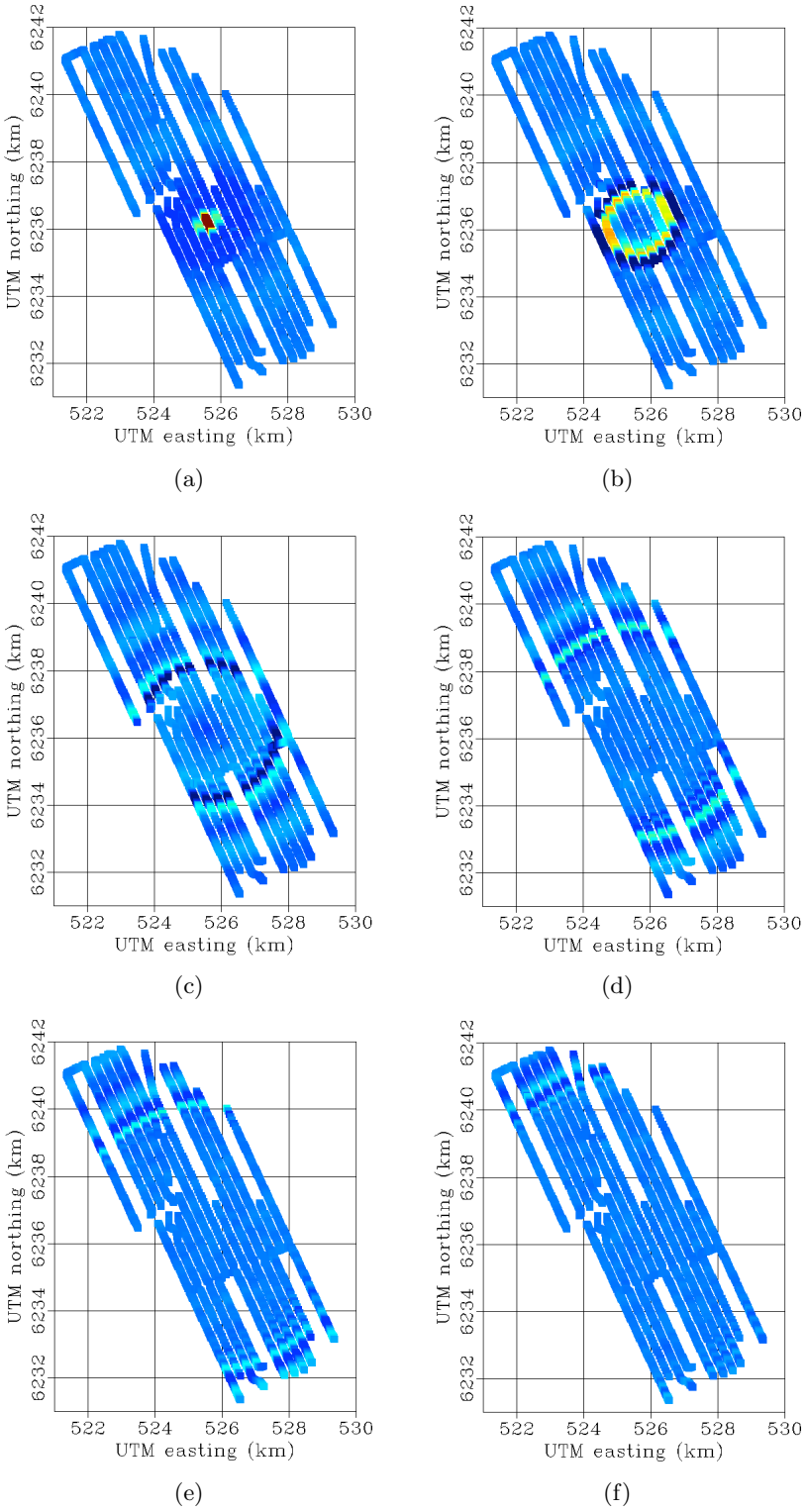


Figure 4: A virtual seismic source located in the center of the array. Time slices at: 0 (a), 3 (b), 6 (c), 9 (d), 12 (e) and 15 seconds (f). [**NR**] sjoerd1/. frame0s,frame3s,frame6s,frame9s,frame12s,frame15s

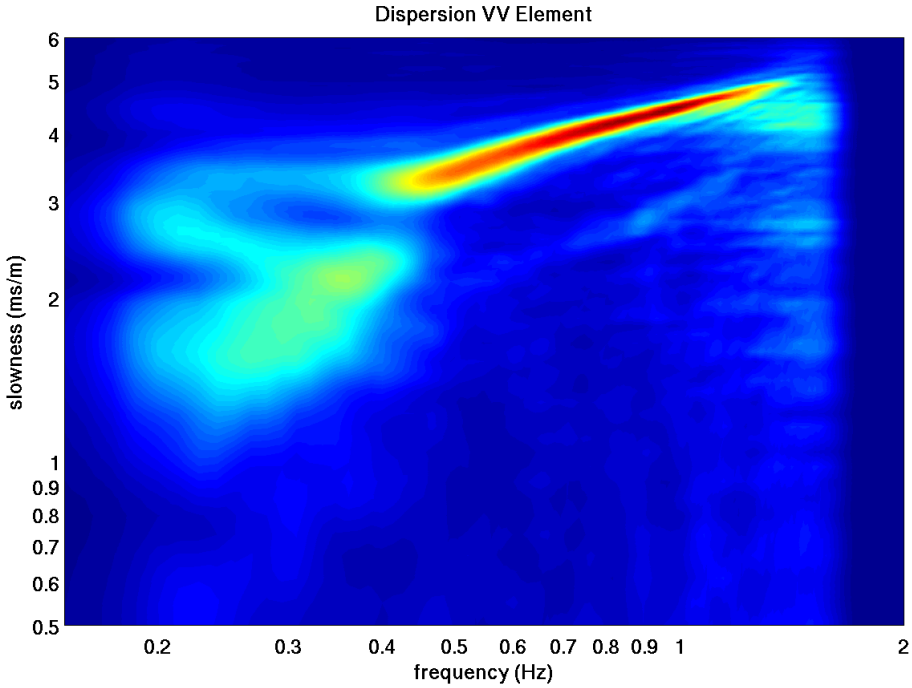


Figure 5: Dispersion image for virtual seismic sources in one OBC line of the array. The virtual sources are dominated by a single, dispersive wave mode. [**NR**] sjoerd1/. Laura-VV-SI-wpLL

FILTERING AND CORRELATION PROCEDURES

These data were recorded by stations designed for active seismic surveying. It was made available with a sampling frequency of 500 Hz, which is common for active seismic data. However, these long and continuous recordings are not perfect; they are contaminated with data glitches, electronic noise, and other noise bursts. It is best to deal with these imperfections before any further processing, because they can cause strong filtering artefacts. The next step is to bandpass the data for the frequency range of interest ($0.18 - 1.75$ Hz), which is considerably lower than the frequency range of active seismic imaging. For recordings at Valhall, this part of the frequency spectrum is fairly clean (see Figure 2), especially far from the platforms.

These data are divided into blocks of 30 minutes with 50% overlap between adjacent blocks and multiplied with a cosine-squared taper. This taper has the advantage of restitching the data together without amplification, according to the Pythagorean trigonometric identity. Each 30 minute block is bandpassed (for $0.18 - 1.75$ Hz) and subsampled by a factor of 50 to a sample frequency of 10 Hz. The data is not restitched, instead, the 30 minutes of data are cross-correlated between each station pair, delivering one EGF for each station pair. For all data in the December 2010 recording, this gives 486 EGFs for each station pair.

Figure 6a) shows one such estimate for all stations cross-correlated with station 1, Figure 6b) shows the mean of all estimates for 2 hours, Figure 6c) shows the mean of all

estimates for 1 day, and Figure 6 shows the mean of all estimates for a little over 5 days. In each individual estimate there is little signal that exceeds the background correlation fluctuations. However, as we stack more estimates and thus increase the effective recording time we cross-correlate, over arrivals start standing out from the background correlation fluctuations. These arrivals correspond to the surface-wave Green's function between all stations and station 1.

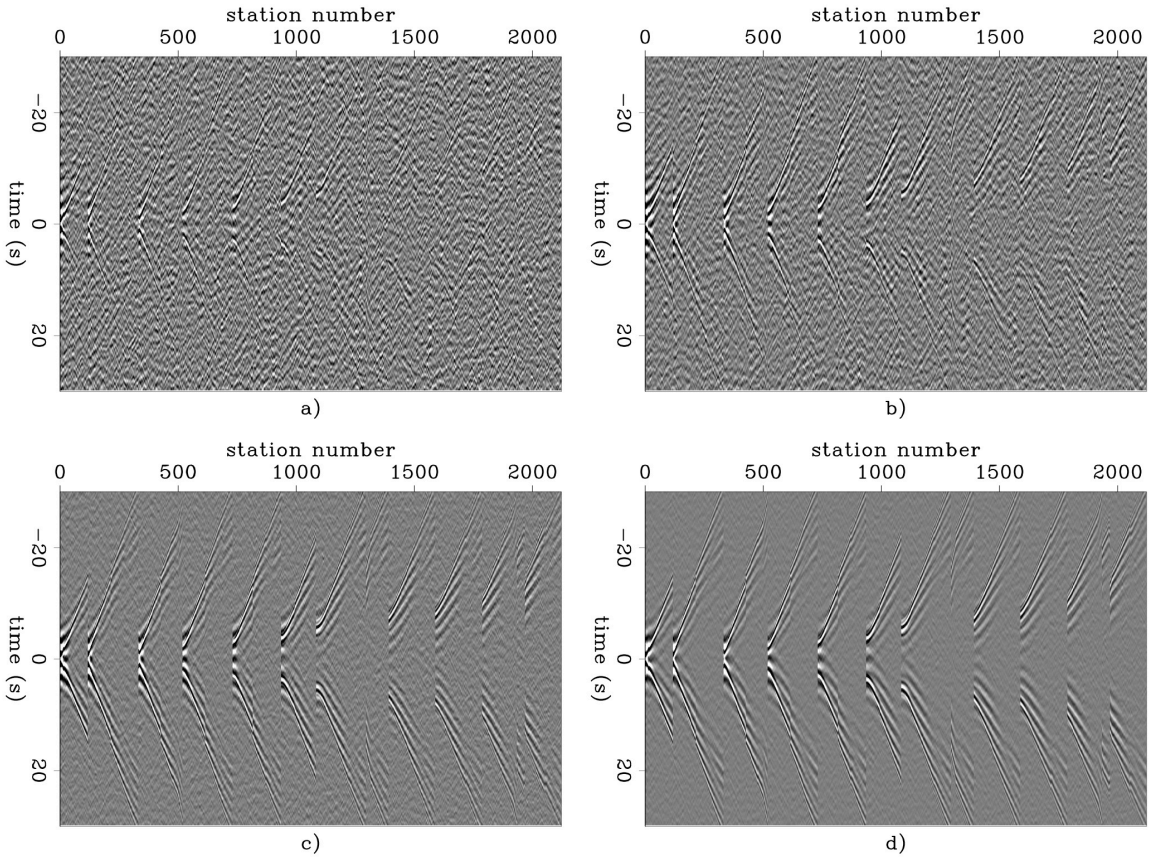


Figure 6: Four virtual seismic sources generated by cross correlating recordings of ambient seismic noise; a) 1/2 hour of recordings, b) 2 hours of recording, c) 1 day of recording, d) 5 days of recording. [**CR**] sjoerd1/. stacks4

CONVERGENCE RATE OF AMBIENT-SEISMIC NOISE CORRELATIONS

To design an installation that exploits seismic interferometry as a permanent, continuous, nearly real-time monitoring system, we need to investigate how much ambient-seismic noise we must cross-correlate to retrieve a stable EGF. The correlation coefficient between two virtual sources, one created by correlating 3 hours of data and one created by correlating all data, as a function of the center-time of the shorter recording. A high correlation coefficient means that the correlation of the partial recording is very similar to a correlation of the full signal. This analysis should be made with caution, because there is no measure of the physical correctness of the EGF in these correlation coefficients. Theoretically, the inclusion

of more time in longer records could deteriorate the correlation signal as an estimation of the Green's functions. However, the comparison of correlations of a partial recording with the total recording gives a measure of the convergence rate and the stability of the estimated Green's function. The resulting correlation coefficients are averaged for a virtual source in the center of the array and a source at the north-west end of the array, as shown in Figure 7. These measures for correlation-convergence show similarity with the strengths of the microseism energy as seen in Figure 2, but the especially strong periods of microseism energy on the 26^{th} and 27^{th} of December are less dominant in Figure 7. When higher frequencies are particularly strong, the partial stack does not resemble the full stack very well.

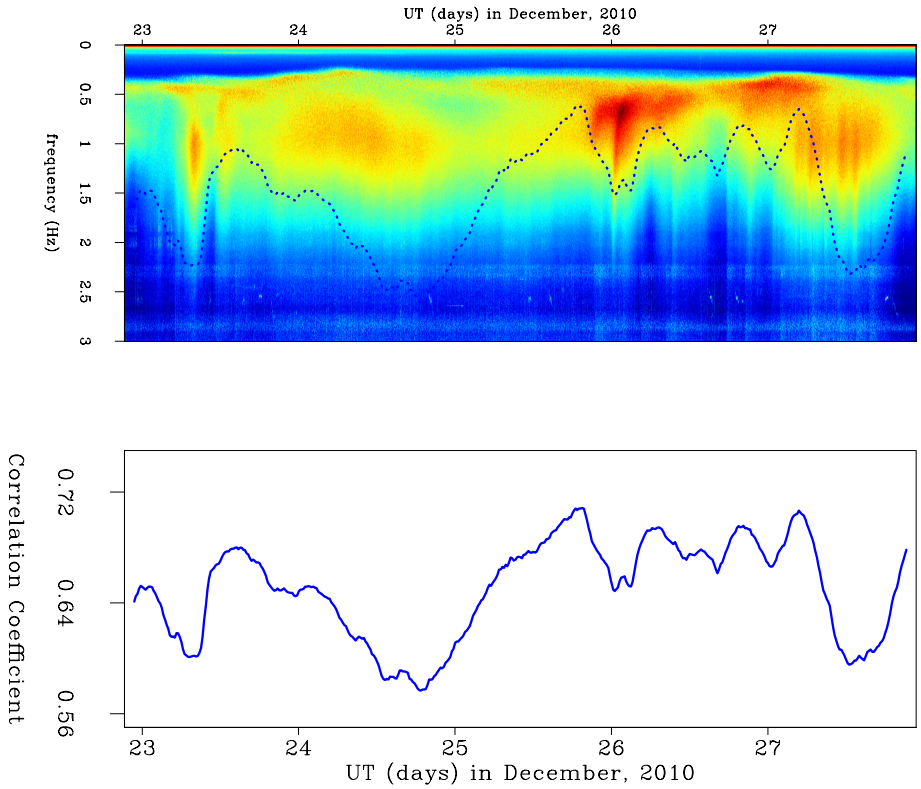


Figure 7: Correlation coefficients as a function of start time comparing a 3 hour partial window with the full stack. The top panel shows the spectragram with the correlation coefficients overlayed and the bottom panel shows only the curve of correlation coefficients. [**CR**] sjoerd1/. corrcoefVt-shots-si

We want to analyse the cross-correlation convergence rate as a function of inter-station distance for various frequency ranges. We compute the correlation coefficient between two virtual sources after bandpassing for a certain central frequency, and bin the coefficients as a function of inter-station distance. One virtual sources is computed by cross-correlating all data while the second is computed cross-correlating a partial recording. Furthermore, to eliminate sensitivity to a particular starting time of the partial recording, we slid the partial recording window across the complete recording. We repeat the whole procedure for two virtual sources and average the mean correlation coefficients for each distance bin

and partial recording length. The frequency range $1.50 - 1.75$ Hz is shown in 8a, $1.50 - 1.75$ Hz in 8b, $1.25 - 1.50$ Hz in 8c, $1.00 - 1.25$ Hz in 8d, $0.75 - 1.00$ Hz in 8e, $0.50 - 0.75$ Hz in 8f, $0.25 - 0.50$ Hz in 8f. Dotted and dashed lines are 0.50 and 0.95 contours. The contour lines are not very smooth. This can probably be improved by averaging the computations for more virtual sources. These figures show a trend of faster convergence rate for lower frequencies. This trend is especially apparent with inter-station distances larger than 1000 meters. This trend is expected, because low frequencies have a larger Fresnel zone than do higher frequencies, and therefore require fewer sources surrounding the stations. Since the same sources excite the entire frequency regime, the background correlation fluctuations will be dominated by higher frequencies, and the correlation will stabilize faster for lower frequencies. Correlations between stations at shorter distances converge much faster, but a trend for different frequencies at shorter distances is not clear. This behavior explains the results in Figure 7, where a strong presence of high frequencies actually deteriorates the convergence rate.

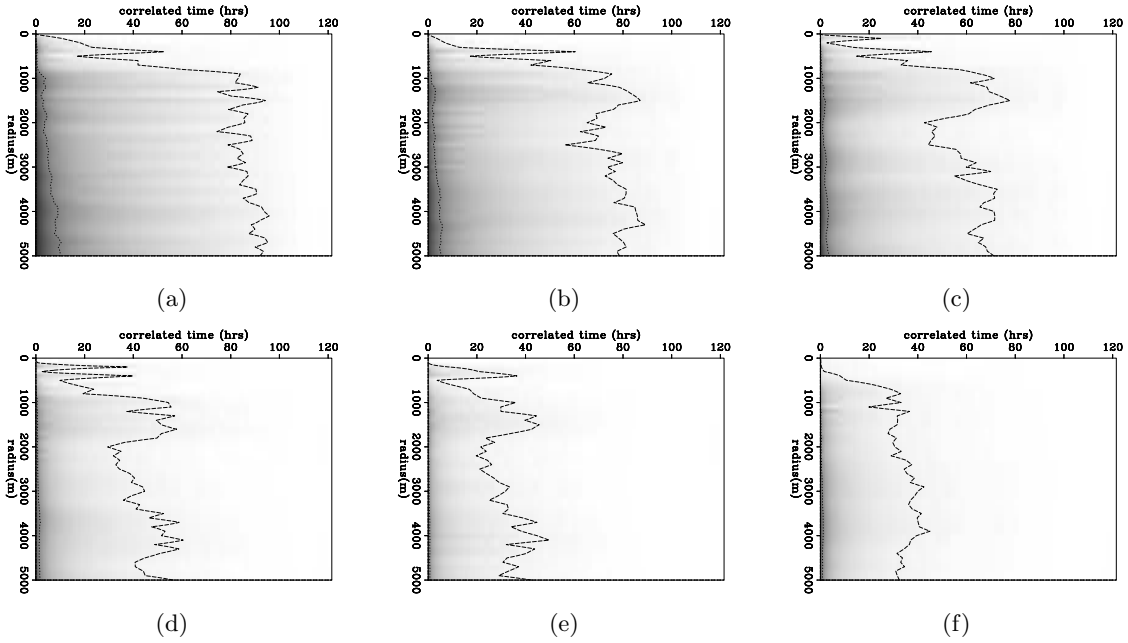


Figure 8: Correlation coefficient of partial versus total cross-correlated signal time. Dotted and dashed lines are 0.50 and 0.95 contours. Analysis made for frequency ranges; $1.50 - 1.75$ Hz in (a), $1.50 - 1.75$ Hz in (b), $1.25 - 1.50$ Hz in (c), $1.00 - 1.25$ Hz in (d), $0.75 - 1.00$ Hz in (e), $0.50 - 0.75$ Hz in (f), $0.25 - 0.50$ Hz in (f). [**CR**] sjoerd1/. coef-high,coef-middle,coef-low,coef-lower,coef-verylow,coef-verylower

AMBIENT SEISMIC-SEISMIC NOISE TOMOGRAPHY

The virtual seismic sources are dominated by dispersive surface waves. These propagate in two dimensions along the seafloor through a frequency-dependent velocity map. Group velocities of the surface waves can be found for narrow frequency ranges. We use a simple Hann-window bandpass to select a narrow frequency range, and then we pick travel times by the maximum of the envelope of the filtered time-domain virtual-source arrival. We

define a quality factor on each pick by a signal to noise ratio (SNR): the maximum of the envelope within a linear moveout window to the average of the envelope outside the window. For various narrow frequency bands we selected travel times, $\mathbf{t}$, with an SNR > 5 and at offsets between 50 and 1750 meters. These travel-time picks were input into a straight-ray tomography kernel, linearized with perturbations, $\Delta\mathbf{m}$, in an average velocity, m_0:

$$\mathbf{m} = \mathrm{m}_0 + \Delta\mathbf{m}, \tag{2}$$

$$\mathrm{m}_0 = \frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{t}}{\Delta\mathbf{x}}, \tag{3}$$

$$\Delta\mathbf{t} = \mathbf{t} - \mathrm{m}_0\,\Delta\mathbf{x}, \tag{4}$$

where $\Delta\mathbf{t}$ are the travel-time residuals after accounting for the contribution of the average velocity and $\Delta\mathbf{x}$ is the offsets for each specific travel-time pick. The tomography-problem fitting goals are posed as follows:

$$\mathbf{F}\Delta\mathbf{m} - \Delta\mathbf{t} = \mathbf{0}, \tag{5}$$

$$\epsilon\nabla^2\Delta\mathbf{m} = \mathbf{0}, \tag{6}$$

where we use the ∇^2 operator as regularization to force a smooth model, and the total model is reconstructed by summing with the background velocity.

The aim is to create group-velocity maps at different frequencies by picking a group travel time at the peak of the envelope after a narrow-range bandpass. Lower-frequency group-velocity maps should reflect deeper structures, because at lower frequencies (and thus longer wavelengths), surface waves are sensitive to deeper structures. Picking a frequency range that is too narrow would result in an oscillatory wavelet and make travel-time picks less accurate. Figure 9 contains a set of inversion results using virtual sources of different frequency ranges. Six overlapping frequency ranges were selected: $0.15 - 0.75$ Hz, $0.35 - 0.95$ Hz, $0.55 - 1.15$ Hz, $0.75 - 1.35$ Hz, $0.95 - 1.55$ Hz and for $1.15 - 1.75$ Hz. At all ranges but the lowest, the wave-mode is very sharply defined in the dispersion image of Figure 5.

Stacking just one day of data gives five independent stacks. Five tomographic images, one for each day, of group-velocity for the frequency range $0.75 - 1.35$ Hz are shown in Figure 10. And five tomographic images, one for each day, of group-velocity for the frequency range $1.15 - 1.75$ Hz are shown in Figure 11. Stacking two and a half day of data gives two independent stacks. Four tomographic images, of both stacks, of group-velocity for the frequency ranges $0.75 - 1.35$ Hz and $1.15 - 1.75$ Hz are shown in Figure 12. These tomographic images are similar but not the same.

We can compare these group-velocity images to P-wave velocity images obtained by full wave-form inversion of active seismic surveys (Sirgue et al., 2010). These images include features extending beyond the extent of the receiver array, because the active sources cover a wider area than the receivers do. Whereas the maps obtained from ambient-seismic noise tomography are logically confined within the area of the recording array. In the 240 meters below the seafloor, the P-wave velocity maps show several buried channels as high-velocity anomalies. The bigger channel on the east is as deep as 105 to 240 meters. There are smaller channels buried in the top 100 meters. Below one of the smaller shallow channels is a low-velocity zone that crosses the array; this feature is apparent between 150 and 240 meters.

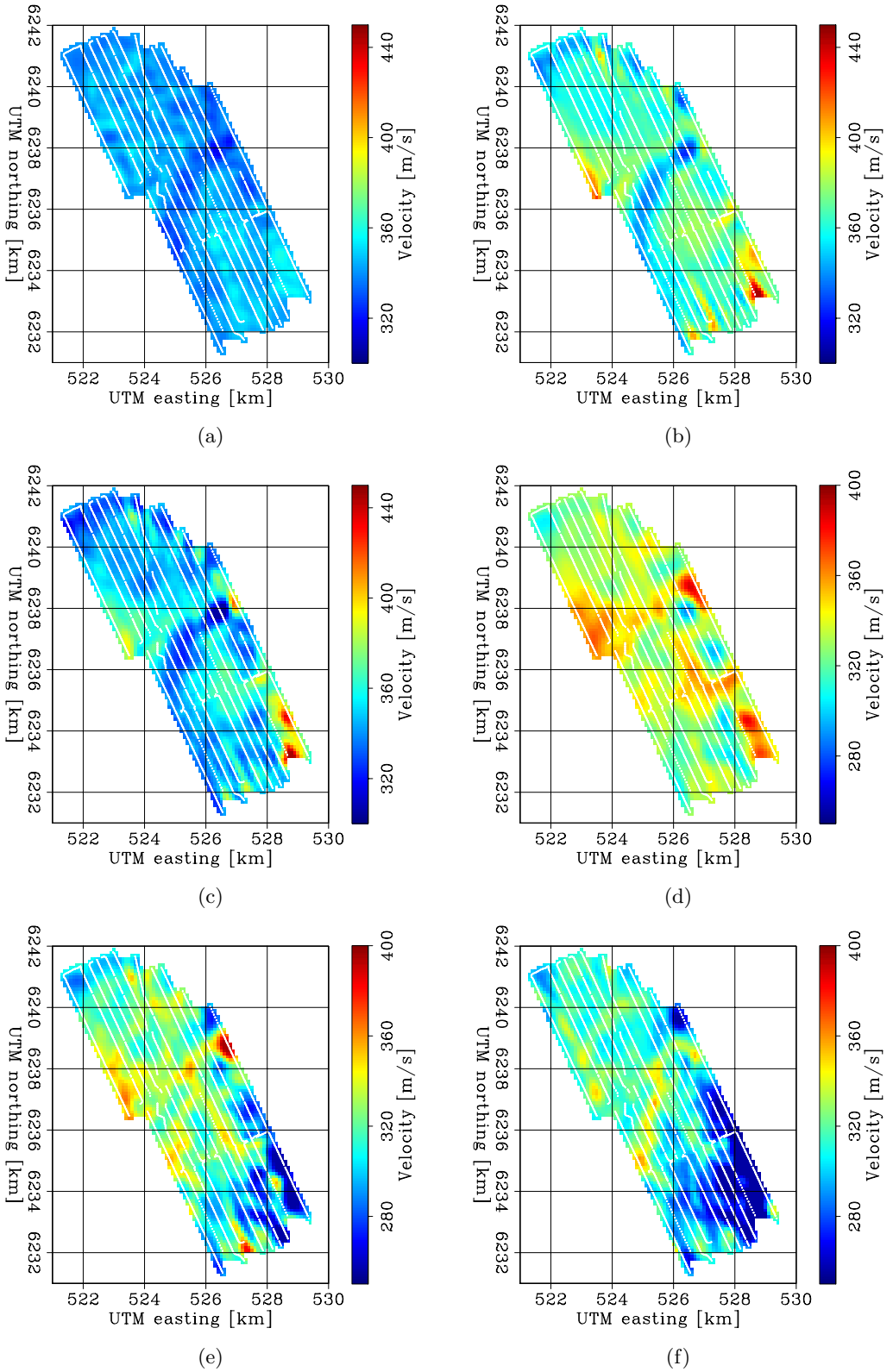


Figure 9: Tomography results for various frequency bands: $0.15-0.75$ Hz (a), $0.35-0.95$ Hz (b), $0.55-1.15$ Hz (c), $0.75-1.35$ Hz (d), $0.95-1.55$ Hz (e), and for $1.15-1.75$ Hz (f). [**CR**] sjoerd1/. JZTomo-C1,JZTomo-C2,JZTomo-C3,JZTomo-C4,JZTomo-C5,JZTomo-C6

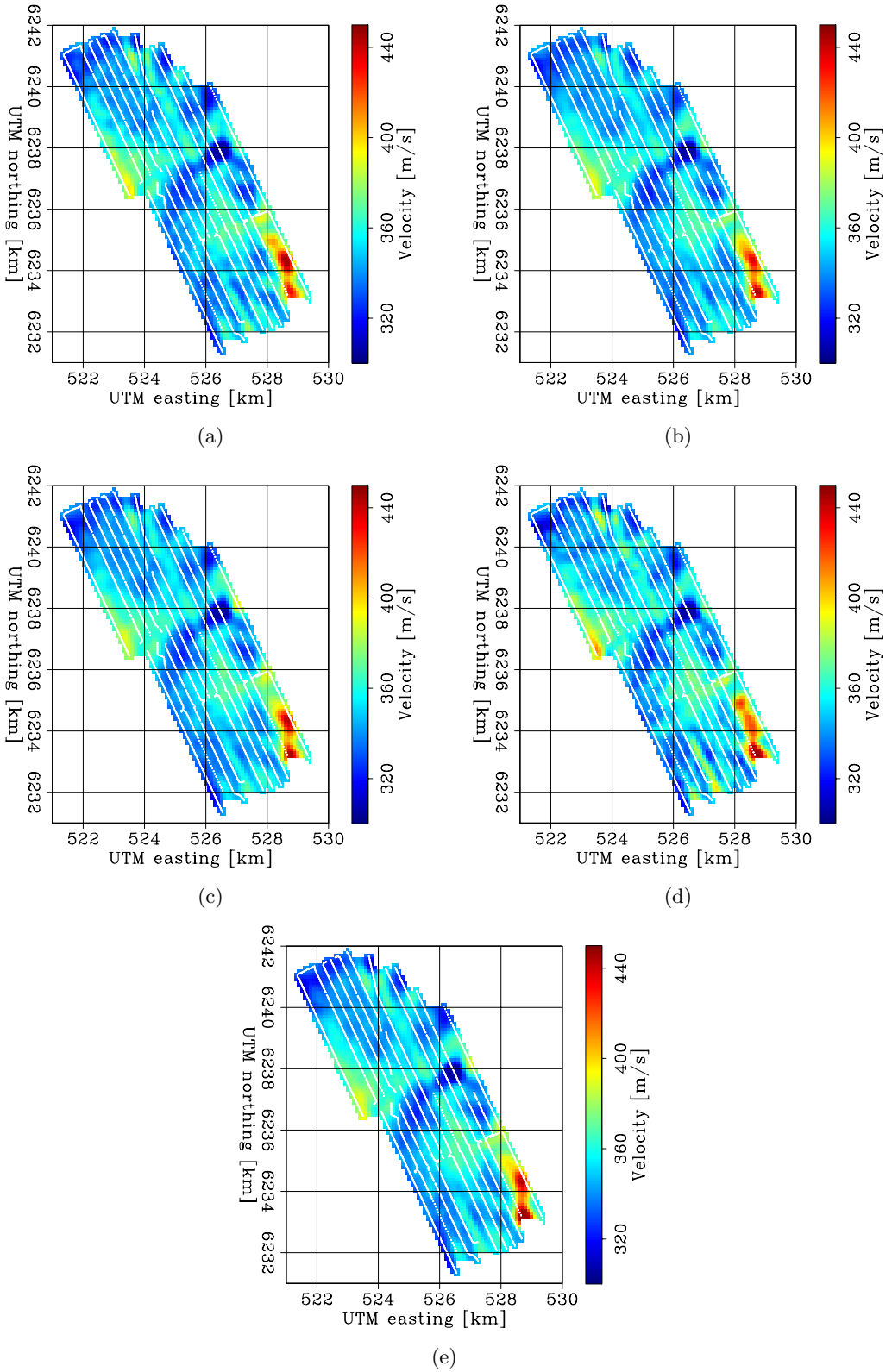


Figure 10: Tomographic results for correlations of one day of data between 0.75 – 1.35 Hz: day 1 (a), day 2 (b), day 3 (c), day 4 (d), day 5 (e). [**CR**] sjoerd1/. JZTomo-d1C3,JZTomo-d2C3,JZTomo-d3C3,JZTomo-d4C3,JZTomo-d5C3

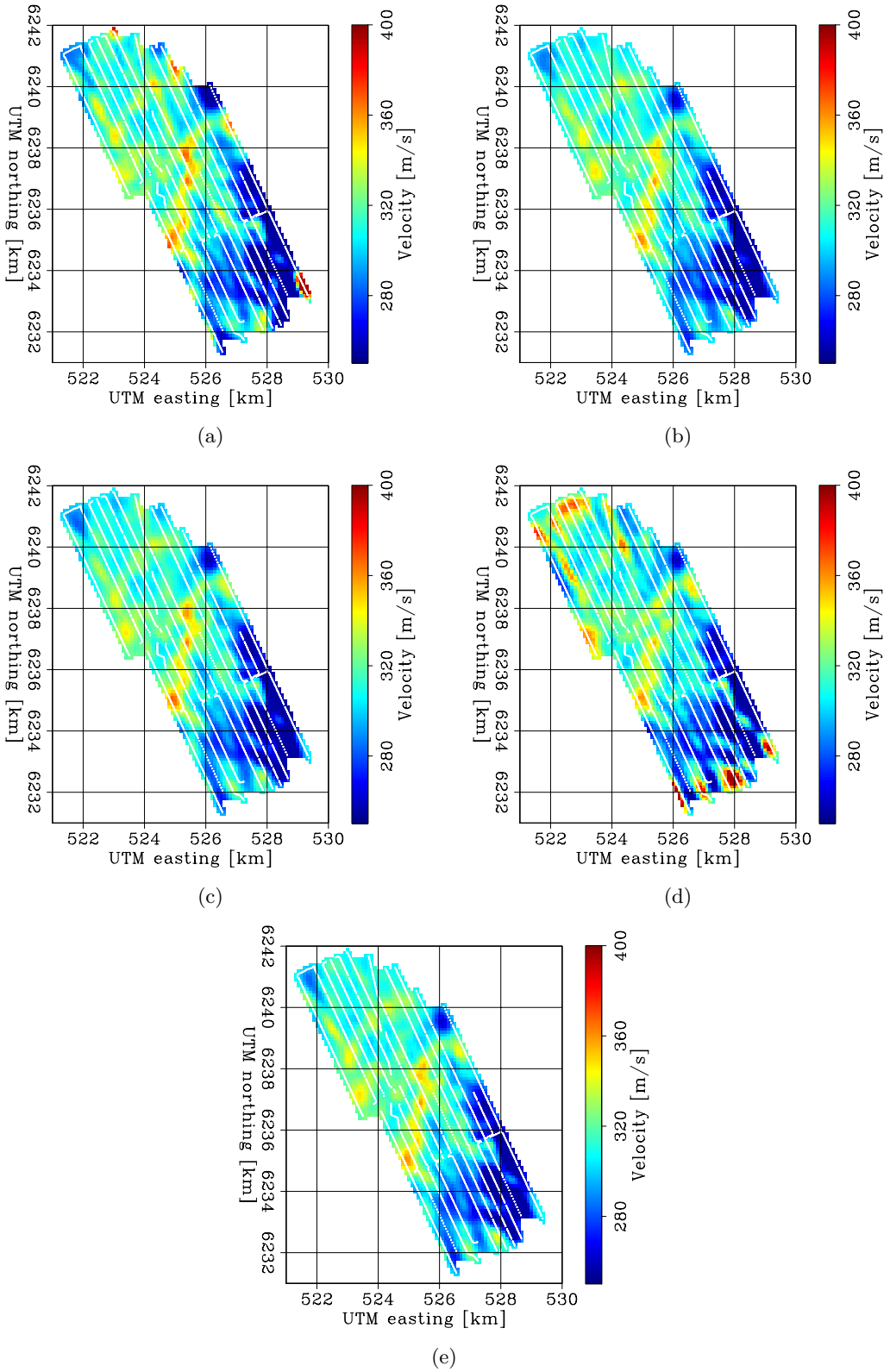


Figure 11: Tomographic results for correlations of one day of data between 1.15 – 1.75 Hz: day 1 (a), day 2 (b), day 3 (c), day 4 (d), day 5 (e). [**CR**] sjoerd1/. JZTomo-d1C6,JZTomo-d2C6,JZTomo-d3C6,JZTomo-d4C6,JZTomo-d5C6

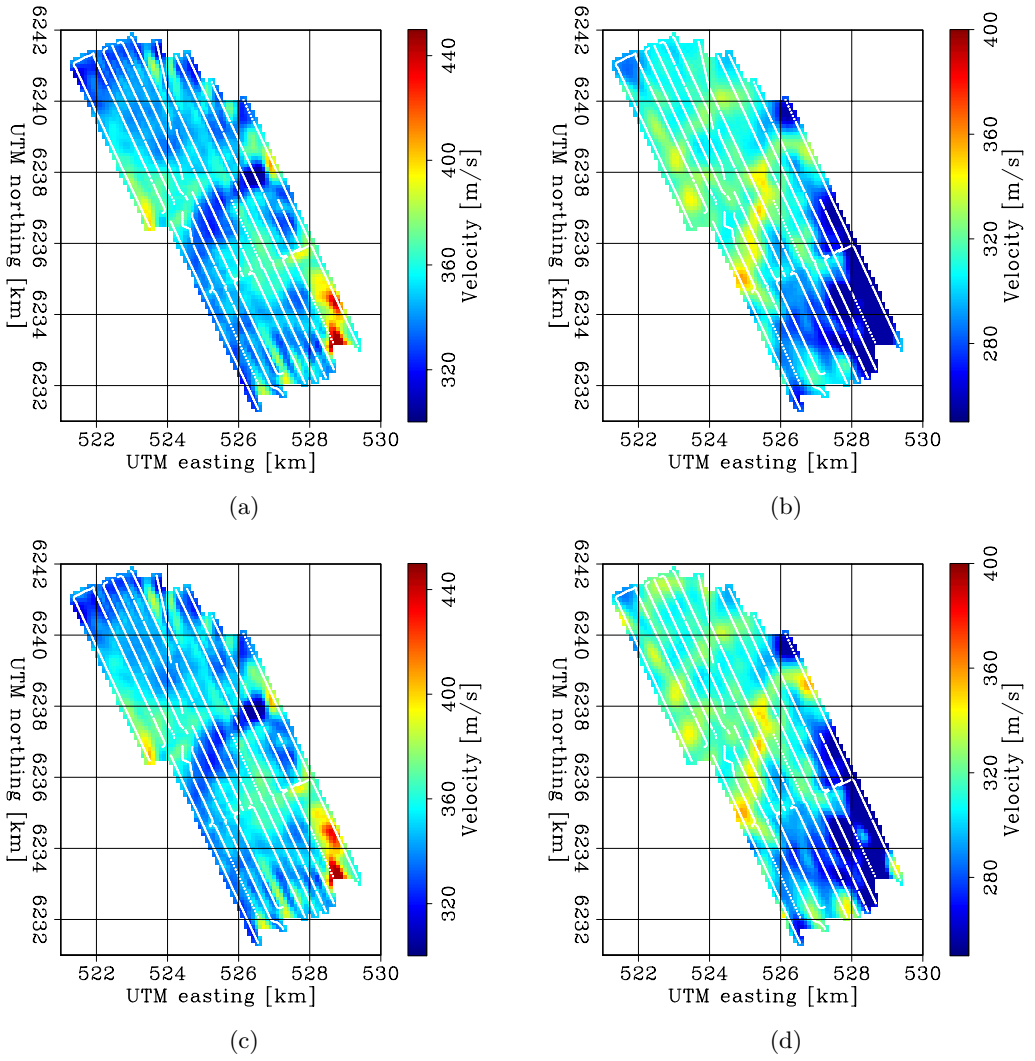


Figure 12: Tomographic results for correlations of two and a half days of data between $0.75 - 1.35$ Hz (a) and (c) and for correlations between $1.15 - 1.75$ Hz: from the first two and a half days in (a) and (b), from the second two and a half days in (b) and (d). [**CR**] sjoerd1/. JZTomo-h1C3,JZTomo-h1C6,JZTomo-h2C3,JZTomo-h2C6

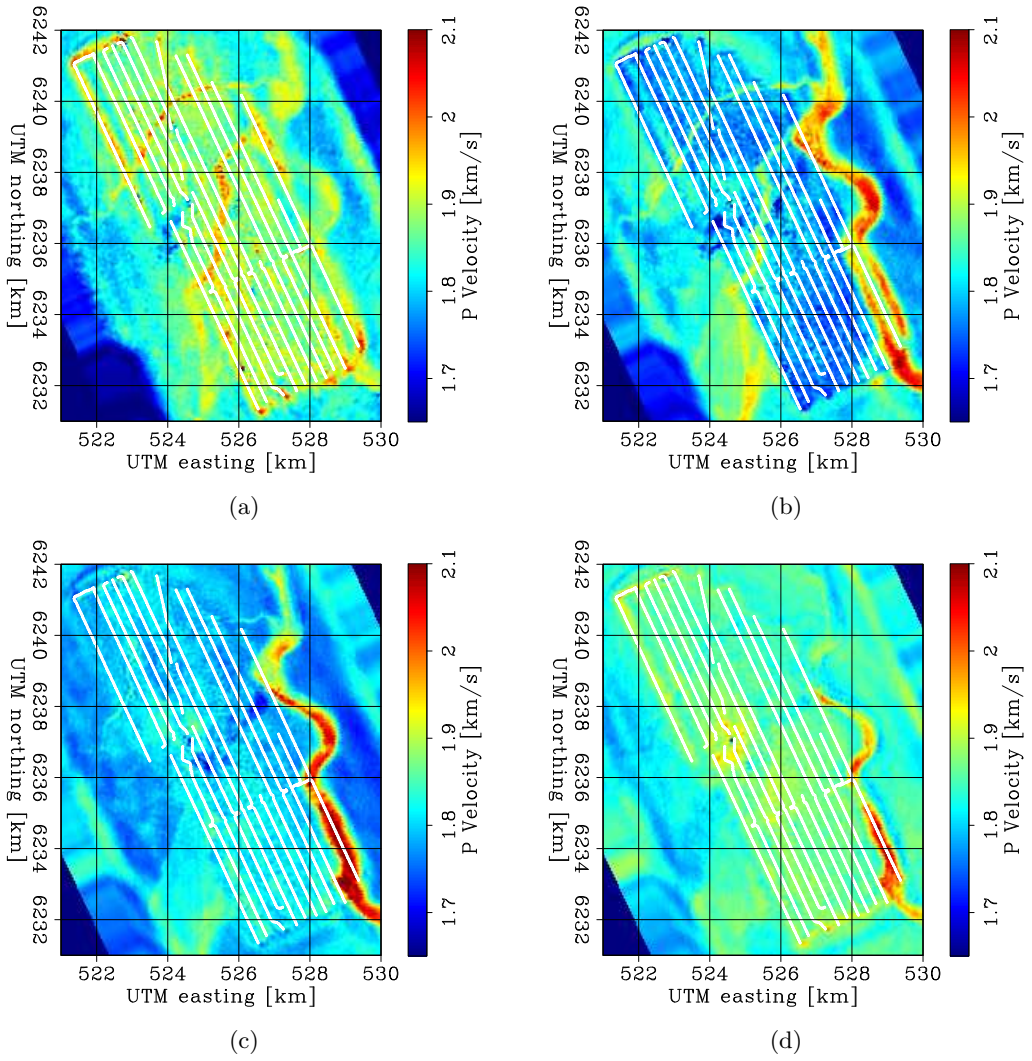


Figure 13: Image of P-wave velocities obtained using wave-form inversion of active P-wave data (Sirgue et al., 2010). Velocity slices are for the following depth ranges: 60 − 105 m (a), 105 − 150 m (b), 150 − 195 m (c), and 195 − 240 m (d). [**NR**] sjoerd1/. map3,map4,map5,map6

DISCUSSION AND CONCLUSIONS

Previous passive seismic interferometry studies have used tomography to image the surface wave Green's functions for reservoir studies (Bussat and Kugler, 2009; de Ridder and Dellinger, 2011). Here we investigate the utilization of passive seismic interferometry as a permanent, continuous, nearly real-time monitoring system. The correlations stabilize faster for lower frequencies than for higher frequencies. We show that correlations of as little as 1 day of data between nearby receivers for frequencies below 1 Hz converges to 95% of the correlation of 5 days of data. The group-velocity maps obtained from ambient-seismic noise tomography contain similar features in different frequency regimes. The smaller shallow channels are visible between $0.95 - 1.75$ Hz. The edge of the deeper channel is visible on the southeastern end of the array between $0.35 - 1.35$ Hz. A low-velocity zone crosses the array and is clearly distinguishable between $0.35 - 1.15$ Hz. The images are generally smoother than the P-wave velocity images from ambient-seismic noise tomography. This is partly because of the smoothing of the velocity map, but since the frequency of the ambient-seismic noise is also lower than that of active seismic surveying, the resolution cannot be expected to be as high. The very nature of the seismic waves used for the imaging is different. Surface waves propagate along the surface, and thus the depth resolution is provided only by their dispersive characteristics and cannot be expected to be as high as for P-waves. Nevertheless, we can achieve remarkable images of the top 240 meters based on Scholte-wave velocities obtained by ambient-seismic noise correlations. These are of interest for continuous reservoir monitoring, especially for short time scale (days) changes in the shallow subsurface. The similarity of tomographic images obtained from independent day long cross-correlations or even two-and-a-half long cross-correlations needs further quantification.

ACKNOWLEDGEMENTS

Thanks to Joe Dellinger and Olav Barkved of BP for discussions, suggestions and support of this study and for comments on this work. Thanks to Biondo Biondi, Bob Clapp and Jesse Lawrence for helpful discussions and suggestions. And thanks to BP and the partners of the Valhall Field (BP Norge A/S and Hess Norge A/S) for the data.

REFERENCES

Ammon, C. J., C. Ji, H.-K. Thio, D. Robinson, S. Ni, V. Hjorleifsdottir, H. Kanamori, T. Lay, S. Das, D. Helmberger, G. Ichinose, J. Polet, and D. Wald, 2005, Rupture process of the 2004 Sumatra-Andaman earthquake: Science, **308**, 1133–1139.

Bussat, S. and S. Kugler, 2009, Recording noise - estimating shear-wave velocities: Feasibility of offshore ambient-noise surface-wave tomography (answt) on a reservoir scale: SEG Technical Program Expanded Abstracts, **28**, 1627–1631.

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, Expanded Abstracts, X047.

Kommedal, J. H., O. I. Barkved, and D. J. Howe, 2004, Initial experience operating a

permanent 4C seabed array for reservoir monitoring at Valhall: SEG Technical Program Expanded Abstracts, **23**, 2239–2242.

Longuet-Higgins, M. S., 1950, A theory of the origin of microseisms: Philosophical Transaction Royal Society Londen A, **243**, 135.

NMI, 2012, Norwegian Meteorological Institute climate data database: http://sharki.oslo.dnmi.no/. (last visited: March 1, 2012).

Rhie, J. and B. Romanowicz, 2004, Excitation of earth's continuous free oscillations by atmosphere-ocean-seafloor coupling: Nature, **431**, 552–556.

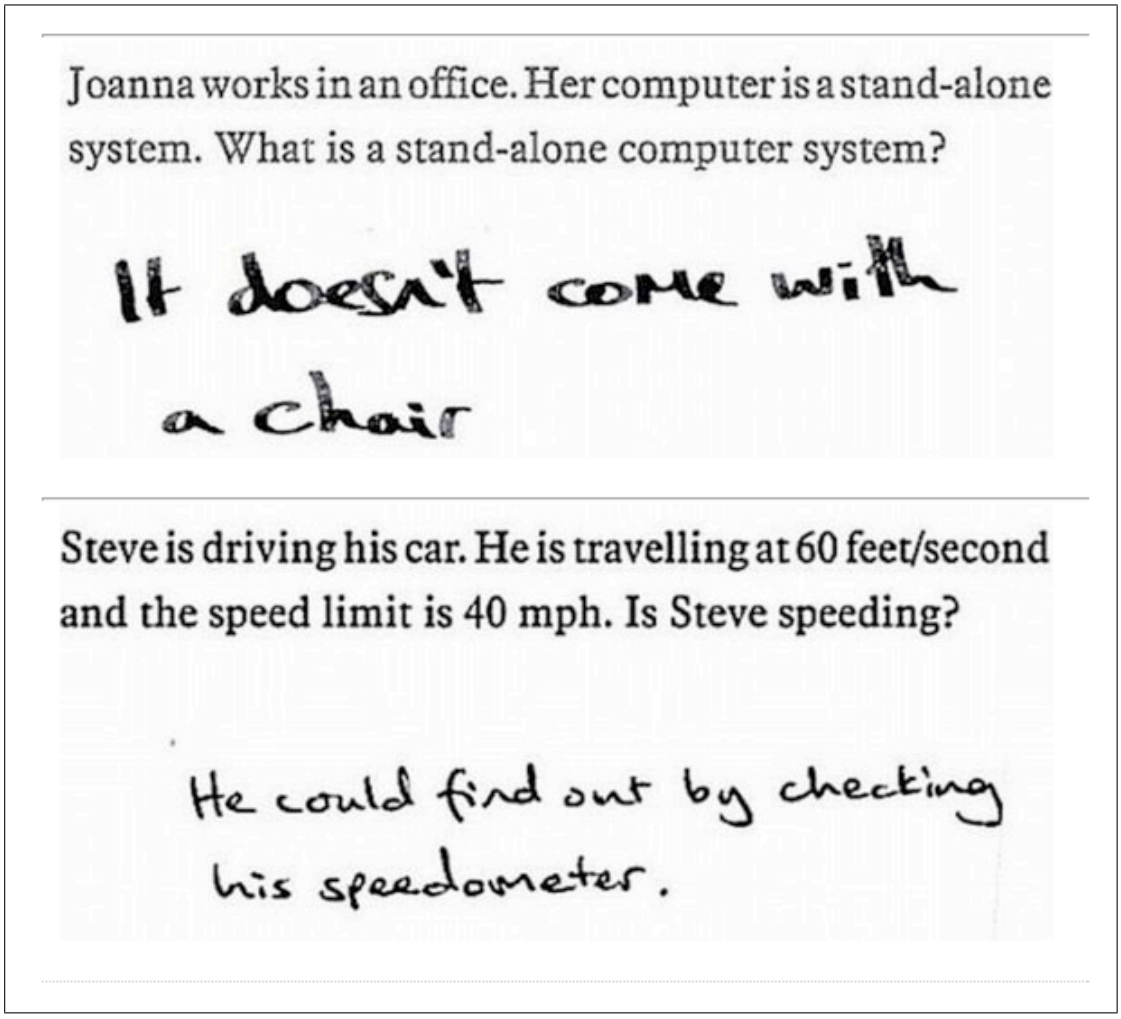
———, 2006, A study of the relation between ocean storms and the Earths hum: Geochem. Geophys. Geosyst., **7**, Q10004.

Shemeta, J. and P. Anderson, 2010, It's a matter of size: Magnitude and moment estimates for microseismic data: The Leading Edge, **29**, 296–302.

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. and J. Fokkema, 2006, Green's function representations for seismic interferometry: Geophysics, **71**, SI33–SI46.


Joanna works in an office. Her computer is a stand-alone system. What is a stand-alone computer system?
It doesn't come with a chair
Steve is driving his car. He is travelling at 60 feet/second and the speed limit is 40 mph. Is Steve speeding?
He could find out by checking his speedometer.

Single frequency 2D acoustic full waveform inversion

Sjoerd de Ridder, Ali Almomin, and Musa Maharramov

ABSTRACT

Single frequency 2D acoustic full waveform inversion shows promise as an approach to inverting the acoustic, dispersive 2D acoustic wave equation for the underlying velocity cube as a function of frequency. A GPU friendly finite difference time domain kernel and an associated frequency domain optimization scheme are shown to retrieve a Gaussian anomaly via full waveform inversion for several acquisition geometries. Although the wave field is single frequency, the spatial distance of sources and receivers allows their sensitivity kernels to interfere and form a gradient that can recover the anomaly.

INTRODUCTION

Surface waves naturally propagate in two dimensions along the earth's surface and not in depth. However, du to their wavelength their waves' motion causes disturbances away from the interface, thus making surface waves sensitive to mechanical properties away from the interface. This is manifested in a variability of the wave propagation speed with frequency, i.e. dispersion. Dispersive Scholte waves can be extracted from ambient seismic noise by cross-correlation (de Ridder and Dellinger, 2011; de Ridder, 2012). Kimman (2011) found that higher modes are often not constructed or very weak, when the excitations are located at the surface. Furthermore, de Ridder (2012) shows how Scholte waves reconstructed from ambient seismic noise at Valhall field contain a single mode dispersive surface wave.

An approximate physical model for a single mode of surface waves is waves travelling in two dimensions given a phase velocity map, $c(\omega, \mathbf{x})$. The amplitudes and phases of surface waves, in that approximation, are governed by

$$\left(\nabla^2 + v^{-2}(\omega, \mathbf{x})\omega^2\right) V(\omega, \mathbf{x}) = F(\omega)\delta(\mathbf{x} - \mathbf{x}_s), \tag{1}$$

where $\mathbf{x} = (x, y)$ and $\nabla^2 = \partial_x^2 + \partial_y^2$. Equation 1 is very similar to a classical 2D acoustic wave equation, except that the velocity is now a function of frequency. Explicitly solving for a single frequency renders this difference moot. However, frequency domain solutions are slow to compute, and implementation of absorbing boundary conditions is not straightforward. Here we solve equation 1 by a time-domain equivalent, valid for individual frequencies and independent of boundary condition.

We want to use equation 1 to image the phase-velocity cube $v(\omega, \mathbf{x})$ by full waveform inversion (FWI) for each frequency separately. But single frequency data generally does not resolve anomalies in space, because sensitivity kernels have endless oscillatory behaviour. Conventionally, summing over a finite frequency band collapses the sensitivity kernels to the classic and familiar banana-doughnut kernels (Dahlen and Nolet, 2000). However, in this paper we show that a similar effect is achieved by having sources and receivers throughout and around the target of interest.

This paper starts by connecting a time domain kernel to solutions of the 2D frequency domain acoustic kernels with frequency dependant velocity. Then a non-linear optimization algorithm is developed to invert single frequency amplitude and phase data. Finally we explore the ability of various acquisition grids to invert for a Gaussian anomaly.

TIME-DOMAIN KERNEL

We seek to compute solutions of equation 1 for single frequencies. Thus assume that we have a time domain source function acting at a single frequency $F(t) = A\ \sin(\omega_o t + \phi)$. In a 2D acoustic system this would excite a wave equation as

$$\left(\nabla^2 + v^{-2}(\omega_o, \mathbf{x})\partial_t^2\right) d(t, \mathbf{x}) = A(\mathbf{x})\ \cos(\omega_o t - \phi(\mathbf{x})), \tag{2}$$

The phase and amplitude of the solution, $V(\omega, \mathbf{x})$, at frequency ω_0 is simply computed with a forward Fourier transform

$$d(\omega_o, \mathbf{x}) = \int_{-\infty}^{\infty} d(t, \mathbf{x}) \exp\{i\omega_o t\}\ \mathrm{d}t. \tag{3}$$

If the solution of 2 is computed accurately, then $d(\omega, \mathbf{x}) = 0$ for $\omega \neq \pm\omega_o$. The sourcing terms $A(\mathbf{x})$ and $\phi(\mathbf{x})$ are only non zero where sources act.

To evaluate this inverse Fourier transform, we would need to perform infinite number of time steps and at minus infinity, practically impossible. However, the solution is periodic, thus we can suffice with having computed the wave field for as little as one period. However, we need to initialize the wave field first and wait for the energy of the source to have spread throughout the medium. The computational boundaries could be made absorbing, but this is a challenge, especially at the longer wavelengths that we are interested in modelling. A simpler solution is to extend the domain and omit updating the boundaries. The domain has to be extended large enough to avoid reflections to start modulating the phase throughout the inner area of the model domain that is of interest. This requires a much larger model space, but keeps coding simple and stable.

The kernel is started and run for enough time steps to have the energy of one period to propagate beyond all receivers, assuming one period is enough to stabilize the amplitude and phase of the wave field. This propagation time length is based on a minimum velocity. Then the code steps through two periods while the receiver wave field is kept. We do not need to have a memory variable for the source function or to collect the wave field at the receiver locations. Instead, we upload one sinusoid with a phase from -1π to 3.5π. A simple *mod* function statement reduces any absolute phase to fall within that range. The value for a sinusoid with a phase shift of up to $\pm\pi$ can easily be retrieved. And similarly the value of a cosine is available though a shift of .5π. The discrete inverse Fourier transform is performed on the fly. No memory transfer needs to occur between CPU and GPU while the modelling is running through the time steps.

An upside of this algorithm is that is is embarrassingly parallel over shots and frequencies, not memory intensive and boundary conditions can be neglected. The downside of this approach is that each frequency inversion requires solving a time-domain waveform inversion, hence this approach is suboptimal and is only implemented here because of the

straightforward implementation on CUDA of the 2D waveform inversion with the special excitation source 15. An alternative approach to this problem is to solve the waveform inversion for equation 1 in the frequency domain using a Helmhotz solver that lends itself to efficient CUDA implementation.

OPTIMIZATION SCHEME

The FWI objective function J_{FWI} can be written as:

$$J_{\text{FWI}}(\mathbf{v}) = \|\mathbf{d}(\mathbf{v}) - \mathbf{d}_{\text{obs}}\|_2^2, \tag{4}$$

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

$$d(\mathbf{x}_s, \mathbf{x}_r, \omega; \mathbf{v}) = f(\mathbf{x}_s, \omega) G(\mathbf{x}_s, \mathbf{x}, \omega; \mathbf{v}) \delta(\mathbf{x}_r - \mathbf{x}), \tag{5}$$

where $f(\mathbf{x}_s, \omega)$ 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:

$$\left(\nabla^2 + v^{-2}(\mathbf{x})\omega^2\right) G(\mathbf{x}_s, \mathbf{x}, \omega) = \delta(\mathbf{x}_s - \mathbf{x}). \tag{6}$$

We then separate the model into a background and a perturbation:

$$v^{-2}(\mathbf{x}) = b(\mathbf{x}) + m(\mathbf{x}), \tag{7}$$

where $b(\mathbf{x})$ is the background component, which is the current model in slowness squared units, and $m(\mathbf{x})$ is the perturbation component. After this separation, we can use Taylor expansion on the data around the background component as follows:

$$\mathbf{d}(\mathbf{v}) = \mathbf{d}(\mathbf{b}) + \frac{\partial \mathbf{d}}{\partial \mathbf{v}}|_b \mathbf{m} + ... \tag{8}$$

By neglecting the higher-order terms in the data series, we can define the linearized modeling operator $\mathbf{L}$ as:

$$\Delta\mathbf{d}(\mathbf{v}) = \frac{\partial \mathbf{d}}{\partial \mathbf{v}}|_b \mathbf{m} = \mathbf{L}(\mathbf{b})\mathbf{m}. \tag{9}$$

The first order Born approximation can be used to define the operator:

$$\Delta d(\mathbf{x}_s, \mathbf{x}_r, \omega; \mathbf{b}, \mathbf{m}) = -\omega^2 f(\omega) \sum_{\mathbf{x}} G(\mathbf{x}_s, \mathbf{x}, \omega; \mathbf{b}) m(\mathbf{x}) G(\mathbf{x}, \mathbf{x}_r, \omega; \mathbf{b}), \tag{10}$$

where the Green's functions now satisfy the acoustic wave equation as follows:

$$\left(\nabla^2 + b(\mathbf{x})\omega^2\right) G(\mathbf{x}_s, \mathbf{x}, \omega) = \delta(\mathbf{x}_s - \mathbf{x}), \tag{11}$$

$$\left(\nabla^2 + b(\mathbf{x})\omega^2\right) G(\mathbf{x}, \mathbf{x}_r, \omega) = \delta(\mathbf{x} - \mathbf{x}_r). \tag{12}$$

We can now compute the model gradient $g(\mathbf{x})$ as follows:

$$g(\mathbf{x}) = \frac{\partial J_{\text{FWI}}}{\partial \mathbf{m}} = \mathbf{L}^* \Delta\mathbf{d}. \tag{13}$$

Finally, we can update the model with the gradient:

$$b_{\text{new}}(\mathbf{x}) = b(\mathbf{x}) - \alpha g(\mathbf{x}), \tag{14}$$

where α is the step size. To estimate the step size, we first evaluate the objective function with the gradient scaled to have a maximum of 2% and 4% of the minimum value of the current model. Using these two points as well as the objective function value at the current model, which is already computed in the gradient calculation, we fit a parabola. If the parabola has positive-side minimum, i.e. both the curvature and the x-axis shift are positive, a new objective function evaluation is performed at the parabola minimum. Then, the two or three evaluations are compared and the scale that resulted in the smallest objective function is used as the step size given that the objective function decreases. Otherwise, the line search is repeated after shrinking the gradient by a factor of 4. The optimization scheme is implemented on the CPU in Fourier domain notations using frequency domain Green's function solutions computed on the GPU.

TEST GEOMETRIES

A series of nine examples of single frequency FWI are run for a Gaussian anomaly, each with different source and receiver geometry and with a 2 Hz wave field. The Gaussian shape has a standard deviation of 350 m and an amplitude of 10% on a background velocity of 1500 m/s (Figure 1).

Having just one source and one receiver (Figure 2) severely limits the ability of the acquisition to retrieve a Gaussian anomaly. The final model, achieved after only 4 iterations, basically resembles the sensitivity kernel of one source-receiver couple. Using two receivers and one source (Figure 3) does not improve the situation much. The two sensitivity kernels now interfere, but together they still don't resemble the Gaussian anomaly we attempt to retrieve. A survey using a line of receivers and a single source starts to do better (Figure 4) but the range resolution is still quite poor. However, an acquisition with one source and a field of receivers suffers from a range interference of the outer side lobes between individual source-receiver sensitivity kernels (Figure 5). Two sources and a line of receivers (Figure 6) drastically improves the range resolution over just one source (Figure 4). And similarly for a field of receivers using just two sources (Figure 7). The next two experiments show the result of a line of sources. The first is essentially a conventional borehole geometry, where a line of sources and a line of receivers flank the target (Figure 8). This is capable of retrieving the the velocity anomaly in the cross-line direction reasonably well, but fails to resolve the Gaussian anomaly in the in-line direction. It actually retrieves negative values as big side lobes of the anomaly. This problem is only partially overcome when there is a field of available receivers (Figure 9). When sources and receivers encircle the anomaly, we achieve excellent retrieval (Figure 10) in just 5 iterations. However, retrieval is limited by the wavelength of the single frequency data.

CONCLUSIONS

A time domain kernel to compute a frequency domain solution of the Green's function is successfully implemented on GPU's. If the source receiver distribution illuminates the target from sufficient geometries, individual sensitivity kernels interfere and construct gradients

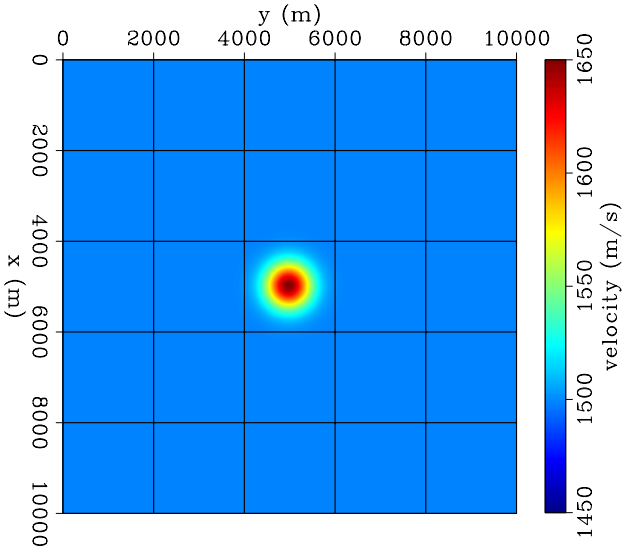


Figure 1: True model. A Gaussian shaped anomaly, with a standard deviation of 350 m and an amplitude of 10% on a background velocity of 1500 m/s. [**ER**] sjoerd2/. true

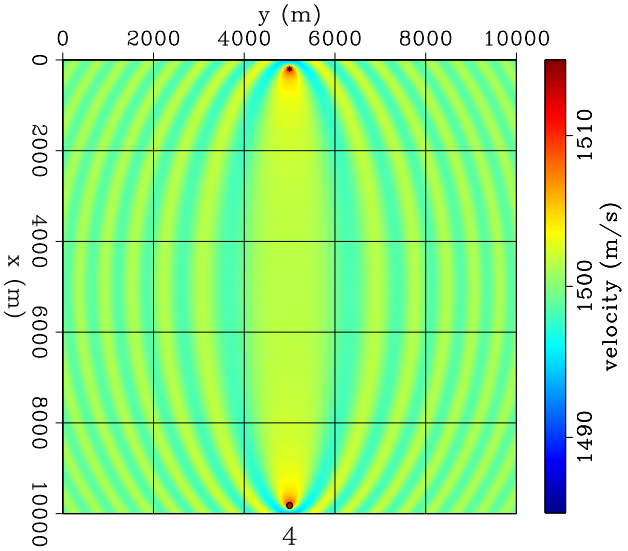


Figure 2: Model 1: One source and one receiver. FWI converged after 4 iterations. A classical single frequency sensitivity kernel. [**CR**] sjoerd2/. fwi1-itt4

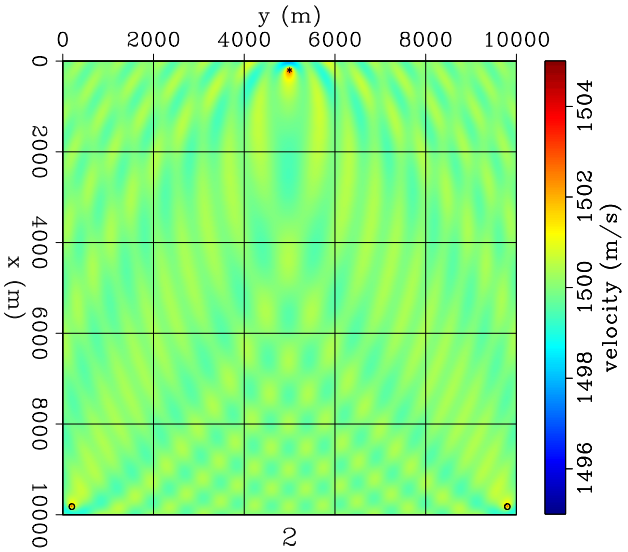


Figure 3: Model 2: One source and two receivers. FWI converged after 2 iterations. The two sensitivity kernels interfere, but together still do not resemble the Gaussian anomaly we attempt to retrieve. [**CR**] sjoerd2/. fwi2-itt2

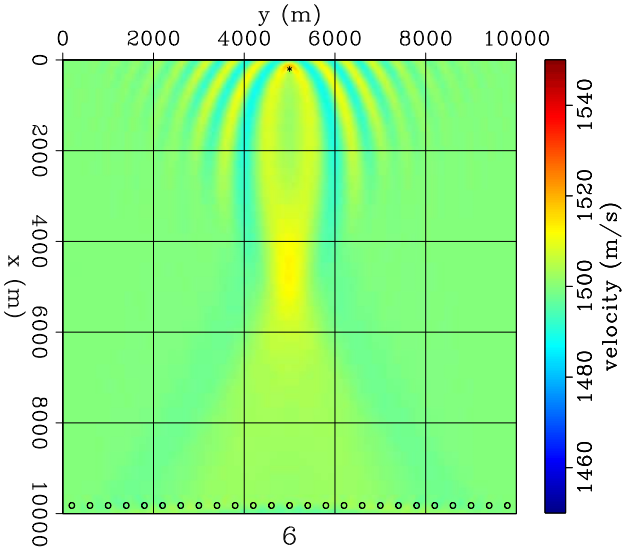


Figure 4: Model 3: One source and a line of receivers. FWI converged after 6 iterations. Poor range resolution, but a velocity anomaly starts to appear in the right location. [**CR**] sjoerd2/. fwi3-itt6

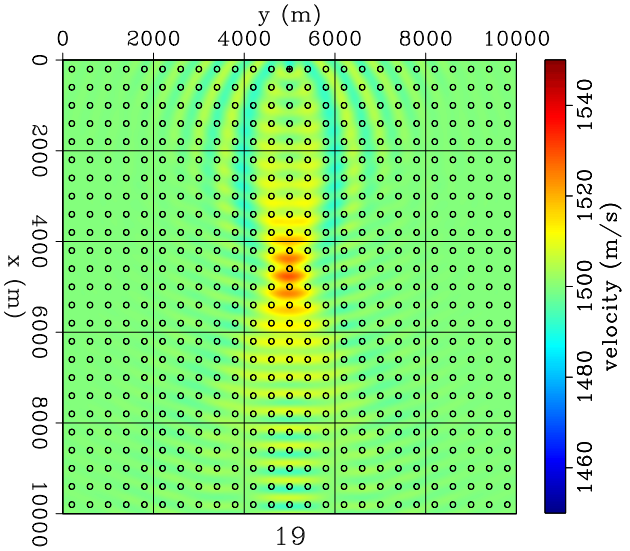


Figure 5: Model 4: One source and a field of receivers. FWI converged after 19 iterations. Although the retrieved anomaly is located correctly, side lobes of individual sensitivity kernels interfere and dominate the result. [**CR**] sjoerd2/. fwi4-itt19

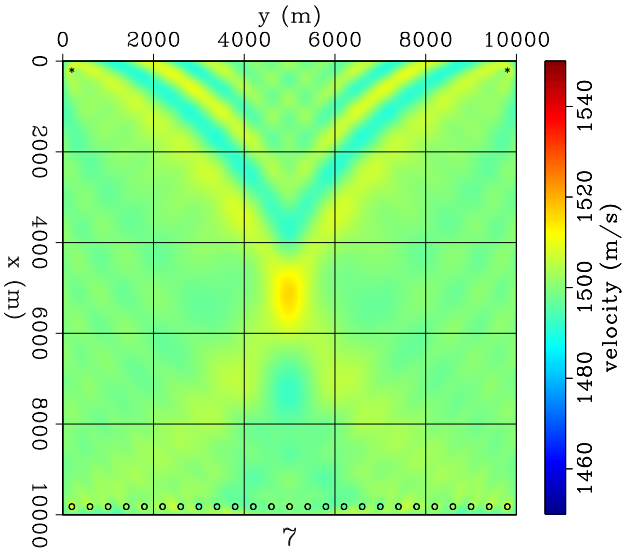


Figure 6: Model 5: Two sources and a line of receivers. FWI converged after 7 iterations. Two sources and a line of receivers drastically improve the range resolution over just one source. [**CR**] sjoerd2/. fwi5-itt7

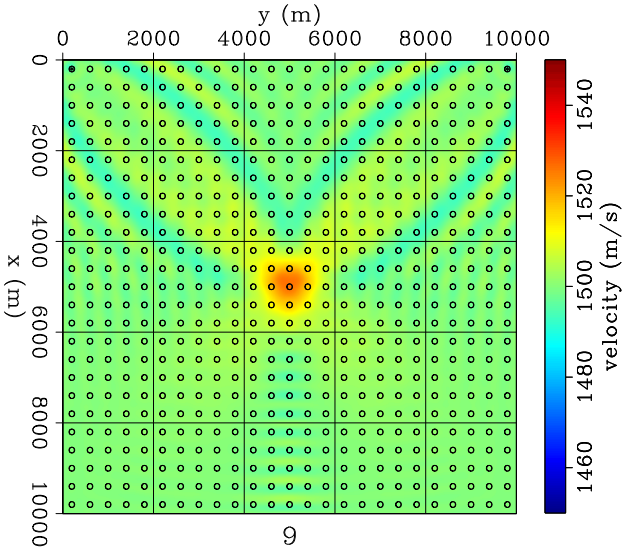


Figure 7: Model 6: Two sources and a field of receivers. FWI converged after 9 iterations. The anomaly is very well shaped in space. [**CR**] sjoerd2/. fwi6-itt9

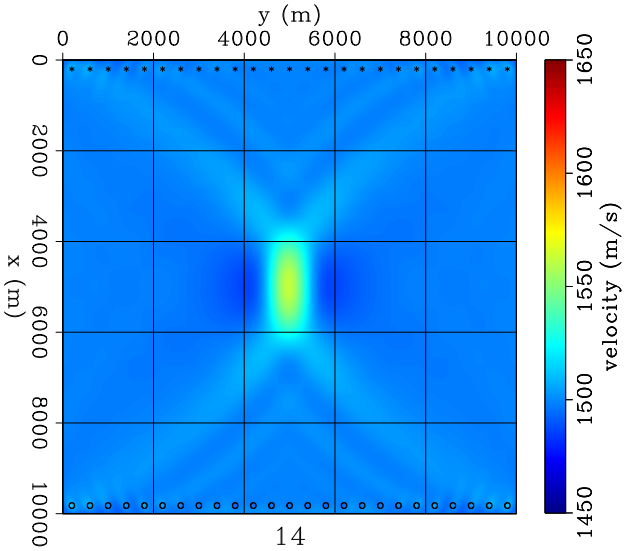


Figure 8: Model 7: A line of sources and a line of receivers. FWI converged after 14 iterations. The anomaly is poorly resolved in range, but there are very few side lobes. [**CR**] sjoerd2/. fwi7-itt14

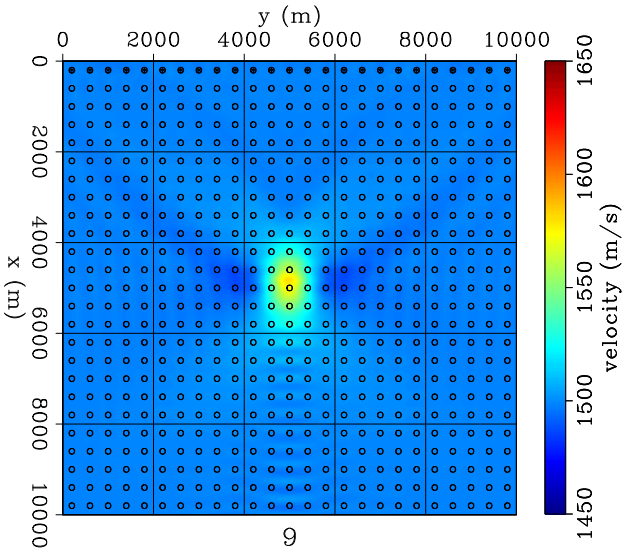


Figure 9: Model 8: A line of sources and a field of receivers. FWI converged after 9 iterations. A field of receivers improves the range resolution. [**CR**] sjoerd2/. fwi8-itt9

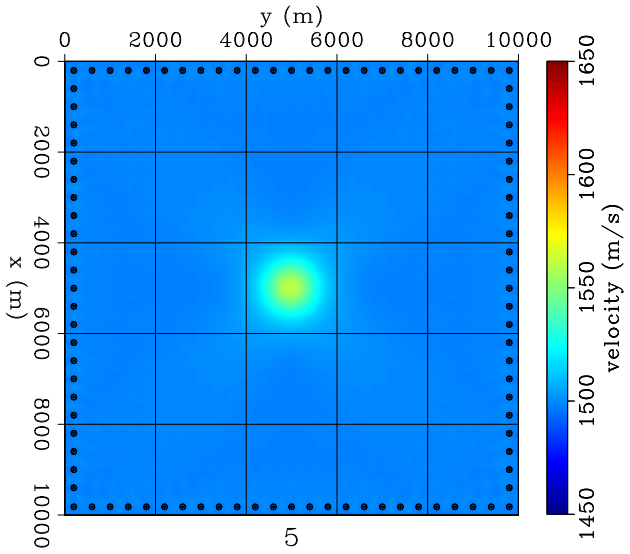


Figure 10: Model 9: A square of sources and receivers. FWI converged after 5 iterations. The anomaly is nearly perfectly reconstructed and without signatures of the individual sensitivity curves. [**CR**] sjoerd2/. fwi9-itt5

for updates in a FWI scheme. A single frequency full waveform inversion scheme is not applied in a multi-scale velocity refinement as is the standard for the waveform inversion process, but the approach is still able to recover Gaussian anomalies in space.

APPENDIX

Our objective is to invert the frequency-dependent velocity function $c(\omega, \mathbf{x})$ from observed solutions to our *postulated* surface wave propagation equation 1 in the frequency domain, where $F(\omega)$ is an arbitrary excitation source waveform. As has been noted above, 2D waveform inversion can be directly applied to this problem, with the only exception that inversion results for individual frequencies are treated as separate values of the frequency-dependent velocity $c(\omega, \mathbf{x})$ and not applied in a multi-scale velocity refinement of the standard for the waveform inversion process. While the applicability of 2D the waveform inversion is conceptually obvious, technical challenges arise in adapting specific implementations of the waveform inversion to equation 2. Our *time domain* CUDA implementation of the 2D acoustic equation solves equation 2 with an excitation source of the form

$$A \sin(\omega_o t + \phi)\delta(\mathbf{x} - \mathbf{x}_s). \tag{15}$$

The choice of the source 15 is dictated by the considerations of our CUDA implementation. We will now show how the time-domain kernel working with sources 15 can be adapted to solve the waveform inversion problem in the frequency domain for an arbitrary waveform source $F(\omega)$. First, Fourier-transforming 2, we can see that recovering the velocity from a solution to 2 with the source 15 is equivalent to recovering the velocity from a solution to

$$\Delta u + \frac{\omega^2 u}{c^2} = \frac{A}{2i}\left[e^{i\phi}\delta(\omega - \omega_0) - e^{-i\phi}\delta(\omega + \omega_0)\right]. \tag{16}$$

By choosing $\phi = \pi/2$ we get

$$\frac{A}{2}\left[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)\right]$$

in the right-hand side of 16. Now assuming that the velocity is an even function of the frequency, and that we attempt to reconstruct the velocity from 2 with $F(\omega) = F(-\omega)$, we can see from equation 16 that we can equivalently run 2D waveform inversion 2 for frequencies $\omega \in [0, Nyq]$ with $A = 2F(\omega), \phi = \pi/2$ in 15. For an odd waveform (but still even velocity) for each ω we use $A = 2F(\omega), \phi = 0$. An arbitrary waveform can be represented as the sum of an even and odd components, and equation 2 will have a linear combination of two source terms 15 in the right-hand side.

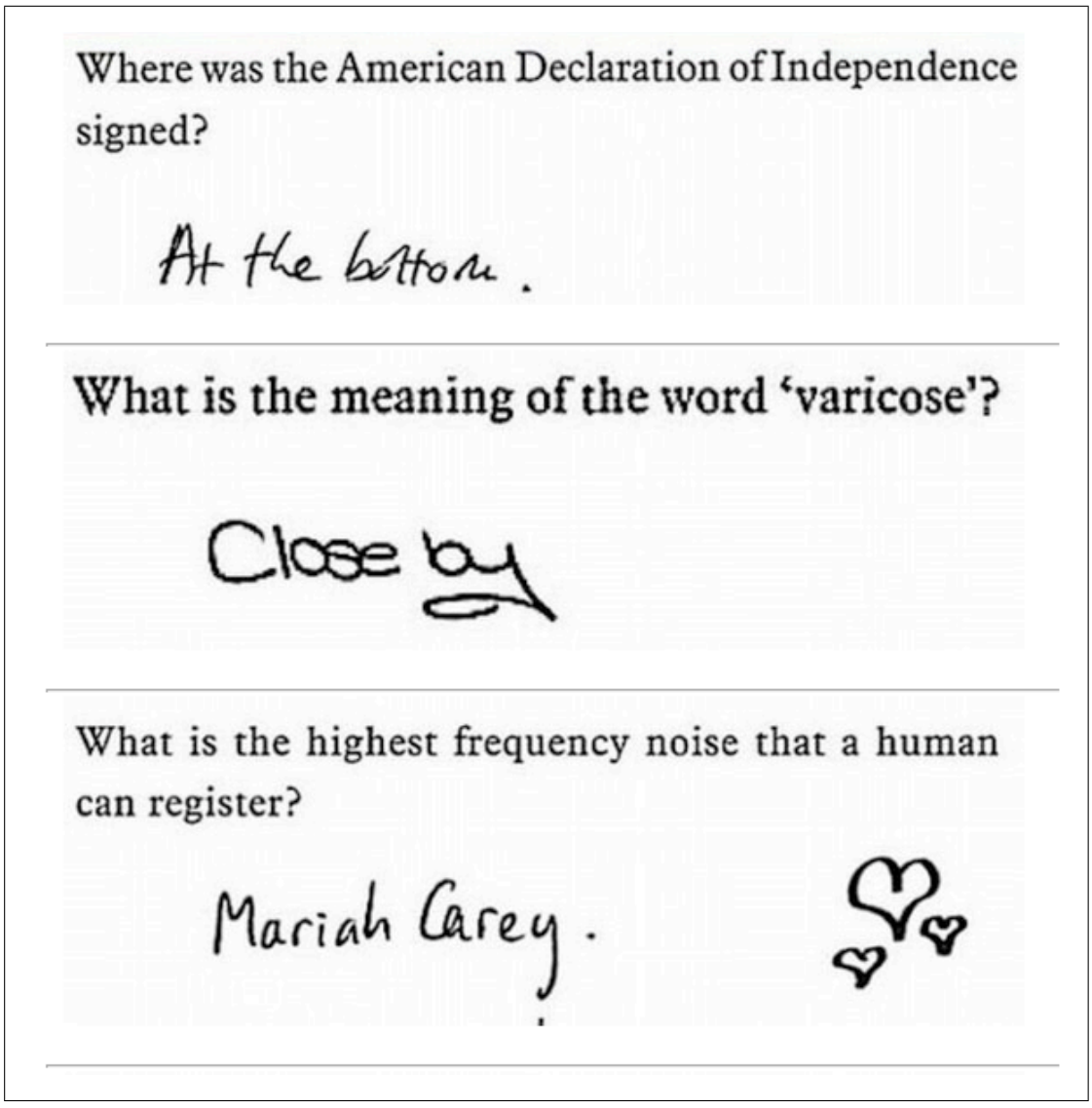
REFERENCES

Dahlen, S.-H. H. F. A. and G. Nolet, 2000, Fréchet kernels for fnite-frequency traveltimes - ii. examples: Geophys. J. Int., **141**, 175–203.

de Ridder, S., 2012, Continuous monitoring by ambient-seismic noise tomography: SEP-Report, **147**, 165–182.

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

Kimman, W. P., 2011, Higher mode surface wave interferometry: PhD thesis, Stanford University.


Where was the American Declaration of Independence signed?
At the bottom.
What is the meaning of the word 'varicose'?
Close by
What is the highest frequency noise that a human can register?
Mariah Carey.

Identifying reservoir depletion patterns from production-induced deformations with applications to seismic imaging

Musa Maharramov

ABSTRACT

The ultimate objective of our research is to study the feasibility of using subsidence measurements for the regularization of linearized waveform inversion of time-lapse data sets, and for differential travel-time tomography. In this paper, we focus on developing a robust framework for inverting pore pressure change in a production reservoir from partial displacement or subsidence data, and on estimating subsurface displacements from the inverted pore pressure change. We discuss potential applications of the resulting displacement fields to the estimation of production-induced change of seismic velocities and impedance. Sensitivity of the proposed framework to the uncertainty and spatial heterogeneity of the poroelastic subsurface parameters is discussed as well.

INTRODUCTION

Our objective is to study the application of geological and geomechanical constraints to the regularization of seismic inversion problems. Well tie-ins and use of a priori information on sediment geometry are examples of such constraints that have found robust quantitative application in regularized seismic inversion. Although some success has been achieved in attempts to connect subsurface velocities and impedance to poroelastic rock properties (Mavko et al., 2009), the application of such relations to constraining velocity inversion runs into difficulties stemming from uncertainties both in the subsurface parameters and in the choice of a semi-empirical rock model to use. However, displacement and stress measurements are the source of independent information about the subsurface, and in principle may be used for the regularization of seismic inversion problems in areas of poor illumination.

More specifically, in the framework of time-lapse seismic imaging, we are interested in answering the following two questions:

- Can we robustly estimate production-induced changes in seismic velocities and impedance (with respect to a base-line model) from subsidence measurements?
- Can we subsequently use the updated model for the regularization of time-lapse inversion where repeat surveys have illumination gaps or are contaminated with noise?

While we discuss the second question in the penultimate section, this paper is mostly concerned with the first question. We begin with an overview of the existing theory of production-induced deformation, then proceed to the formulation of our method for modeling displacements from known pressure drop in a "slowly varying" poroelastic medium. A

technique for inverting the pore pressure change from measured displacements is described next. The proposed modeling and inversion techniques are demonstrated on synthetic examples based on real-world production-induced subsidence cases. The section on spatial heterogeneity of elastic earth models discusses the impact of heterogeneity on modeling and inversion results and puts forward an extension of the proposed technique for tackling vertically-layered and arbitrarily heterogeneous media. The paper concludes with a discussion of velocity and impedance change estimation from strain, with an outline of a process for picking the appropriate semi-empirical rock model.

PRODUCTION-INDUCED DEFORMATION

Realistic reservoir depletion exhibits *asymmetric* patterns that might be indicative of a complex reservoir geometry as well as spatial heterogeneity of the reservoir permeability. In particular, differences in production-related pressure change across multiple wells can help identify reservoir compartmentalization (Zoback, 2010) and guide the drilling of subsequent production wells, or affect the choice of production-enhancement techniques. In the general case of an arbitrary linear poroelastic medium, change in the pore pressure is intricately interconnected with the change in the stress field, and accurate modeling of the effects of changing pore pressure requires solving a system of governing equations in a half-space (Wang, 2000), (Segall, 2010). Solving a boundary-value problem for such a system of equations governing both stress evolution and fluid flow is challenging – as much due to uncertainty of the subsurface model parameters as due to the sheer analytical complexity of these coupled equations. The importance of studying the fully-coupled poroelastic models cannot be overestimated as the coupled modeling is often the key to explaining counterintuitive behaviour of some real-world poroelastic models.

In this work we expand on previous studies of the effects of pore pressure change on the strain and stress fields (Segall, 1992),(Segall et al., 1994) by adopting an intermediate approach between the fully coupled simulation and uncoupled analytical solutions. First, we assume that a pore pressure drop is known within the reservoir and is an arbitrary function of the reservoir coordinates – i.e., we assume partial *fluid-to-solid* coupling – and use the analytical expression for the elastostatic Green's tensor due to a concentrated centre of dilatation in half-space (Segall, 2010),(Mindlin, 1936) to *numerically compute* the displacement due to the pore pressure change. Next, we use the obtained numerical operator in an inversion problem, fitting a pore pressure drop to a known displacement and subsidence profile. Finally, we extend the method for layered media. The method is demonstrated on synthetic data based on a real-world example of subsidence due to the production from a hydrocarbon reservoir at Lacq, France (Segall et al., 1994). Extensions of this study to joint application with time-lapse seismic imaging are discussed as well. This paper is dedicated to reporting some geomechanical aspects of the joint geomechanics/seismic imaging effort aimed at applying the author's optimisation framework to problems of computational geophysics.

As a side note, apart from its *potential* application to the regularization of seismic inversion, numerical modeling of deformation due to changing pore pressure is one of the key problems of reservoir geomechanics (Zoback, 2010). Reservoir depletion has been demonstrated to have appreciable effect on stress both inside and outside of the producing reservoir (Zoback, 2010),(Geertsma, 1973),(Segall, 1992),(Segall et al., 1994),(Zoback and Zinke,

2002),(Segall and Fitzgerald, 1998). While some simplifying assumptions with regard to reservoir geometry (e.g., an infinitely wide and thin horizontal layer) yield a simple law for the horizontal stress change with the pressure decline *within the reservoir* (Zoback, 2010),(Segall and Fitzgerald, 1998)), estimating induced stress-field changes around the reservoir requires more elaborate models of reservoir depletion (Geertsma, 1973),(Segall, 1992). Simple disk-shaped and radially-symmetric reservoir shapes proved adequate for many simple situations, but compartmentalization and heterogeneous permeability inside realistic reservoirs point to a departure of the pore pressure decline from simple axisymmetric patterns (Zoback, 2010). Compartmentalization with impermeable barriers still allows for the application of a radially-symmetric pore pressure change law to individual reservoir compartments. However, an asymmetric pore pressure drop should be able to account for the effects of partial permeability or complex spatial geometry. Since a computationally-efficient technique for modeling and inverting asymmetric reservoir depletion patterns is an important by-product of our study, it is hoped that the computational tools and methods developed for this study will find independent application to the study of production-induced stressing and deformation.

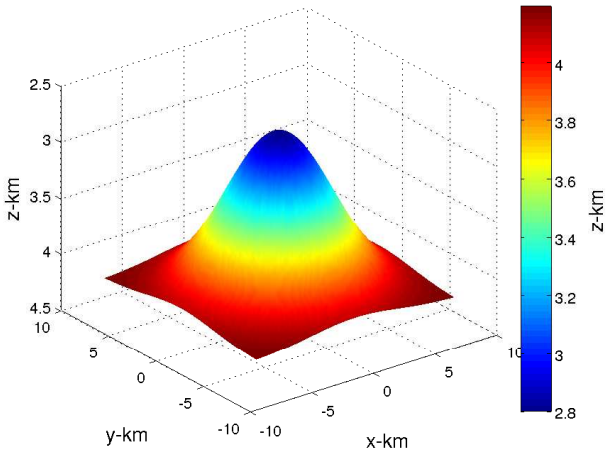


Figure 1: Reservoir dome. The effective depth of the reservoir is 3.5km, effective radius $\approx$ 8km and width $\approx$ .35 km. [**ER**] musa1/. reservoirdome

GOVERNING EQUATIONS

We begin by formulating a *closed* system of four equations that describes a homogeneous *quasi-static* linear poroelastic medium (Segall, 2010):

$$\mu \nabla^2 u_i + \frac{\mu}{1-2\nu} \frac{\partial^2 u_j}{\partial x_i \partial x_j} = \alpha \frac{\partial p}{\partial x_i} - f_i = 0, \; i = 1,2,3 \tag{1}$$

and

$$S_\alpha \frac{\partial p}{\partial t} - \frac{\kappa}{\eta} \nabla^2 p = -\alpha \frac{\partial}{\partial t} \left(\nabla \cdot \mathbf{u} \right). \tag{2}$$

In the above equations $u_i, i = 1,2,3$ is a spatially-distributed *displacement* vector field, p is the *pore pressure change*, f_i is a differential body-force distribution, $\mu, \nu, \alpha, \kappa, \eta$ are the shear

modulus, Poisson's ratio, Biot coefficient, permeability and fluid viscosity, respectively. The *storage coefficient* S_α is a known function of medium parameters (Segall, 2010):

$$S_\alpha = \frac{3\alpha(1-2\nu)(1-\alpha B)}{2\mu B(1+\nu)}, \tag{3}$$

where B in equation 3 is *Skempton's pore pressure coefficient* – the ratio of induced pore pressure to the mean normal stress for *undrained* loading conditions. Note that the *subsurface stress* is absent from equations 1,2 but can be computed using the constitutive laws after these equations have been solved. Also note that the displacement and pore pressure in these equations are *relative* to a reference state, not the total values. The equilibrium equation 1 and flow equation 2 are fully coupled and are obtained from combining the constitutive laws for a poroelastic medium with quasi-static field equations. The equations are "quasi-static" in the sense that the stress field is assumed to be in a state of static equilibrium even though changes of the pore pressure in time induce changes of the stress field. We can think of this as a "slow-change" asymptotic approximation, both in time and space.

The most "mathematically accurate" way of computing the displacement field and associated pore pressure change is to solve a boundary-value problem for 1,2 with known data (e.g., known pressure evolution within existing wells, measured earth displacements or estimated stresses) used as *boundary* or *initial* conditions. However, even in the simplest cases of a homogeneous medium, analytical solutions of boundary-value problems for equations 1,2 is challenging. *Uncoupling* equations 1 and 2, where permissible, could result in more tractable problems, both analytically and numerically. For example, assuming a known pore pressure change, we can solve equation 1 for the displacement field u_i, using $\alpha\frac{\partial p}{\partial x_i}$ in the right-hand side as a "body force" distribution (Geertsma, 1973),(Segall, 1992).

In our approach we use the elastostatic Green's tensor $g_i^k(x,y,z,\xi,\eta,\zeta)$ for the pure elastic equilibrium equation in the left-hand side of equation 1 to compute the displacement u_i as

$$u_i = -\alpha\int_V g_i^k\frac{\partial p}{\partial x_k} = \alpha\int_V \frac{\partial g_i^k(x,y,z,\xi,\eta,\zeta)}{\partial x_k}p(\xi,\eta,\zeta)d\xi d\eta d\zeta, \tag{4}$$

assuming $f_i = 0$ (including body forces is trivial). The elastostatic tensor $g_i^k(x,y,z,\xi,\eta,\zeta)$ in equation 4 has the meaning of the displacement along axis i at point (x,y,z) due to a concentrated force along axis k at point (ξ,η,ζ) (Wang, 2000),(Segall, 2010). From equation 4 we can see that the divergence of the elastostatic tensor has the meaning of deformation due to a concentrated dilatational force.

In order to apply equation 4 to practical reservoir models and computation of *surface subsidence*, the corresponding Green's function should be constructed for a *half-space* with the free-boundary condition imposed on its bounding plane (Segall, 2010). We use the analytical expression for the Green's function obtained by Mindlin (Mindlin, 1936) – see Appendix A for the details. The integral in the right-hand side of 4 is taken over the reservoir domain and hence singularities corresponding to $(x,y,z) = (\xi,\eta,\zeta)$ do not appear. However, the terms in non-diagonal tensor components that contain $r_2 - z - \zeta$ in the denominator blow up at locations directly above (or below) the reservoir and must be truncated in a numerical quadrature. Another important aspect of using an analytical expression for the Green's function is that the divergence in the right-hand side of equation 4 can be

calculated analytically. However, in our implementation we compute the divergence using central differences of the second-order of accuracy.

MODELING DISPLACEMENTS FROM PORE-PRESSURE DECLINE

We can use operator 4 for forward-modeling the displacement field from a specified pressure change. Note that operator 4 is a non-stationary convolutional integral operator for a homogeneous medium. The convolution is non-stationary due to the presence of $z+\zeta$ in the elastostatic Green's tensor. Integration along the horizontal axes can be accelerated by applying the operator in the wavenumber domain. However, integration along the vertical axis should still be carried out separately for different values of z, hence the integration kernel is effectively four-dimensional. Assuming the reservoir to be thin in comparison with its lateral extents, which is always the case, we can replace the vertical integral with a mean value of the integrand times the reservoir thickness:

$$u_i(x,y,z) = \alpha \int_V h(\xi,\eta) \frac{\partial g_i^k(x,y,z,\xi,\eta,S(\xi,\eta))}{\partial x_k} p(\xi,\eta,S(\xi,\eta)) d\xi d\eta, \tag{5}$$

where $S(\xi,\eta)$ is the middle surface of the reservoir and $h(\xi,\eta)$ is the reservoir depth. For a non-flat reservoir, g_i^k effectively depends not only on differences $x-\xi$ and $y-\eta$ but on integration variables as well. In our implementation we compute 5 as a full convolutional operator, allowing *slowly-varying* asymptotics of the solutions by introducing scaling factors that depend on *both* (x,y,z) and (ξ,η,ζ).

An implementation of this method is provided by the `poroelastic_green.F90` module of the author's `exp_tk` Fortran 2003 object-oriented framework described in Appendix B.

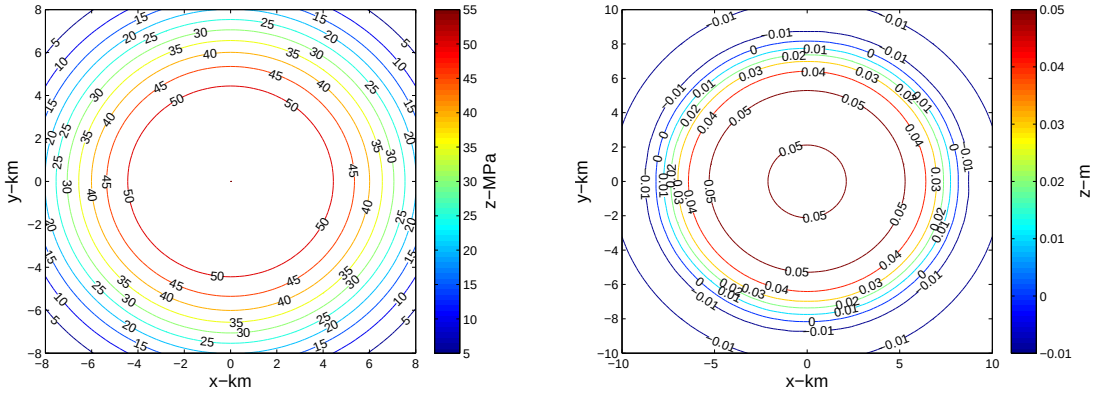
We test our approach on synthetic data generated for a reservoir model based on the assumed geometry of the gas reservoir near Lacq, France, described by Segall et al. (1994). The reservoir is assumed to be anticlinal dome-shaped, with the depth of the reservoir approximated as

$$D(\rho) = 3.5 - 1.4 \left(\exp\left[\left(\frac{\rho}{5} \right)^2 \right] - .5 \right), \tag{6}$$

where ρ is the distance to the center of the reservoir, and all the distances are measured in `km` (see Fig 1). The reservoir thickness is estimated to be $\approx .35$ `km`. In the first test we apply the forward-modeling operator 5 to an axisymmetric synthetic pore pressure drop similar to the pore pressure change used by Segall et al. (1994) and given by

$$p(\rho) = 55 \times \exp\left[\left(\frac{\rho}{8} \right)^4 \right], \tag{7}$$

(see Fig 2(a)). The pressure in equation 7 and the rest of the paper is measured in `MPa`, and ρ is the distance to the center of the reservoir. The result of displacement modeling using an 81×81 reservoir grid and 101×101 surface grid is shown on Fig 2(b). Linear poroelasticity predicts a linear relationship between the maximum subsidence and the maximum pore pressure decline. Segall et al. (1994) predicts that the subsidence rate for the axisymmetric model is bracketed between .7 `mm/MPa` for the reservoir thickness of .25 `km` and 1.37 `mm/MPa` for the thickness of .45 `km`. The subsidence rate indicated by Fig 2(b) for the value of the reservoir thickness of .35 `km` is in a good agreement with that prediction. i



(a) (b)

Figure 2: a) Contour plot of the axisymmetric pore pressure drop given by equation 7. b) Contour plot of the subsidence above the reservoir of Fig 1 due to the axisymmetric pore pressure decline of Fig 2(a). [**ER**] musa1/. cppsymhires,csubsdispoutsymhires

Next, we apply the modeling operator to the asymmetric synthetic pore pressure drop given by

$$p(\rho) = 55 \times \exp\left[\left(\frac{\rho}{8(1+\sin\phi-\pi/4)}\right)^4\right], \tag{8}$$

where ρ, ϕ are the distance and azimuth from an off-centre point (3 km,3 km), with a steeper decline in pore pressure along the northeast-southwest directions (see Fig 3(a)). The pressure decline given by equation 8 is not based on any specific real-world example but is rather chosen as a somewhat extreme case of an asymmetric pressure decline pattern mismatching the reservoir geometry. The result of modeling displacements due to the asymmetric pore pressure decline within the reservoir, using the same reservoir and surface grids as above, is shown on Fig 3(b).

The following table summarizes the medium parameters taken from (Segall et al., 1994) that are used in our tests:

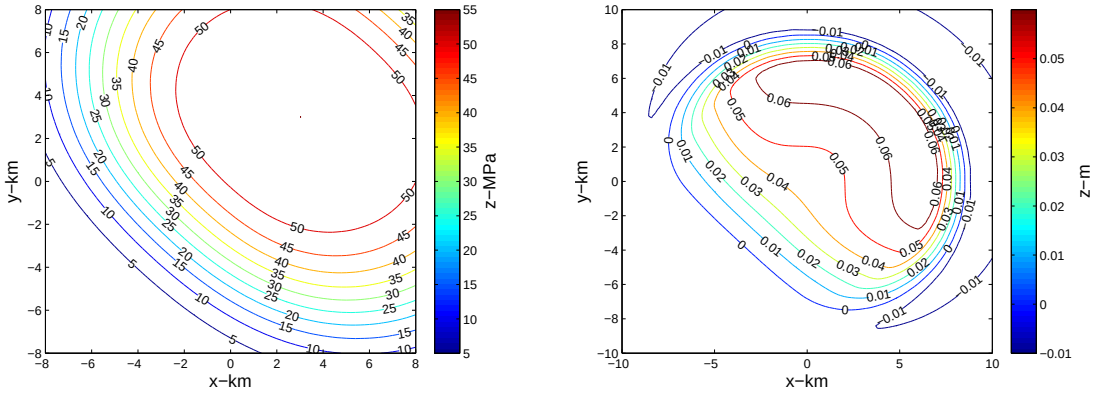
Parameter	SI	Imperial
ν	.25	dimensionless
α	.25	dimensionless
μ	23. GPa	3.34 Mpsi

The Biot coefficient α is computed as

$$\alpha = 1 - \frac{K_{\mathtt{und}}}{K_{\mathtt{min}}}, \tag{9}$$

with $K_{\mathtt{und}}, K_{\mathtt{min}}$ being the undrained bulk modulus and the bulk modulus of rock mineral grains, respectively (Mavko et al., 2009).

Modeling subsidence using operator 5, we are able to fully account for the asymmetric nature of the depletion pattern by using the most general form of Green's tensor for a



(a) (b)

Figure 3: a) Contour plot of the asymmetric pore pressure drop given by equation 8. b) Contour plot of the subsidence above the reservoir of Fig 1 due to the asymmetric pore pressure decline of Fig 3(a). [**ER**] musa1/. cppasymhires,csubsdispoutasymhires

homogeneous half-space. In that respect, our approach represents an advancement of the purely analytical techniques for axisymmetric reservoirs presented in (Geertsma, 1973) and (Segall et al., 1994). However, the assumption of homogeneity, which is required for our being able to use analytical expressions for the Green's tensor, is a serious limitation that may significantly impair the applicability of our method, especially bearing in mind that our ultimate objective is to estimate subsurface displacements from observable subsidence. Later in the paper we propose a computationally efficient extension of our method that is capable of tackling heterogeneity.

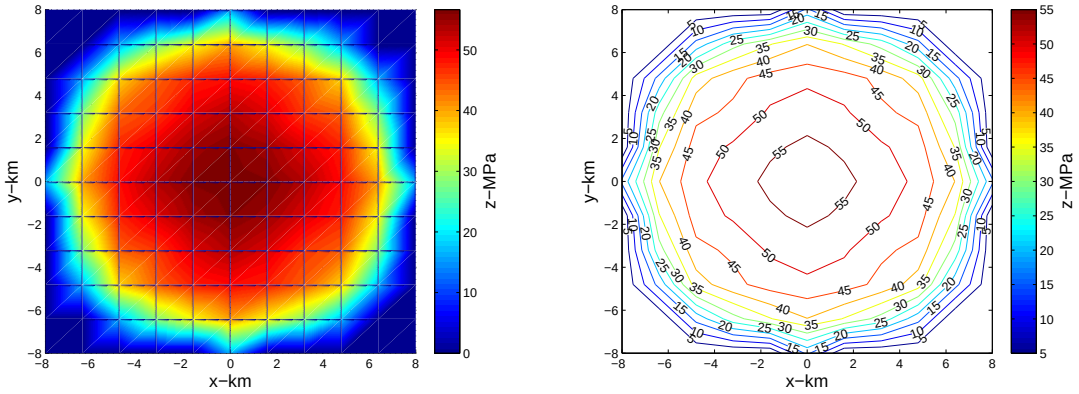
ESTIMATING PORE-PRESSURE DECLINE FROM DISPLACEMENTS

Denoting the operator in the right-hand side of equation 5 as A, the problem of recovering the pore pressure decline from specified displacements can be cast as a least-squares minimisation problem:

$$\|Ap - u\|_{L_2} \quad \rightarrow \quad \min. \tag{10}$$

To solve 10 we use the object-oriented Fortran 2003 optimisation framework `exp_tk` developed by the author that implements multiple general-purpose optimisation methods. The implementation is fully decoupled from underlying model and data structures and can be used in a variety of applications. The tools described in Appendix B implement forward modeling and inversion, with the latter performed by applying a Krylov solver (Trefethen and Bau, 1997),(Bjork, 1996) with a user-specified number of iterations.

We test the inversion code first by inverting the previously modelled displacement for the axisymmetric and asymmetric pore pressure decline patterns of of Fig 2(a) and 3(a). Fig 4(a) and Fig 4(b) show the result for the axisymmetric pore pressure decline inverted on a sparse 11×11 grid in 4 iterations. We employ a *multigrid approach* (Iserles, 2008) where the inversion result obtained on a sparser grid is supplied as an initial approximation



(a) (b)

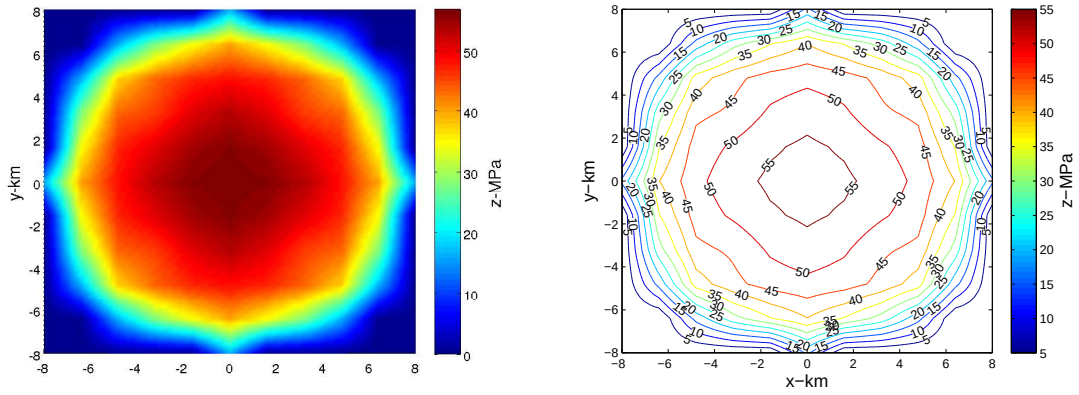
Figure 4: a) Axisymmetric pore pressure decline inverted from the forward-modelled subsidence of Fig 2(b) on a sparse 11×11 grid after 4 solver iterations. b) Contour plot of the inverted axisymmetric pore pressure of Fig 4(a). [**ER**] musa1/. pinvertedporeoutsymlowres,cinvertedporeoutsymlowres

to the inversion on a denser grid. The inversion result for the 11×11 grid was supplied as an initial approximation to the pressure decline for the inversion on the denser 81×81 grid, and the resulting inverted axisymmetric pore pressure decline after 2 more iterations is shown on Fig 5(a) and Fig 5(b).

We have run a similar test for the asymmetric pore pressure decline of Fig 3(a) and equation 8. The result for the sparse grid and 5 iterations is shown on Fig 6(a) and Fig 6(b). The result of 2 more solver iterations on the denser grid is shown on Fig 7(a) and Fig 7(b).

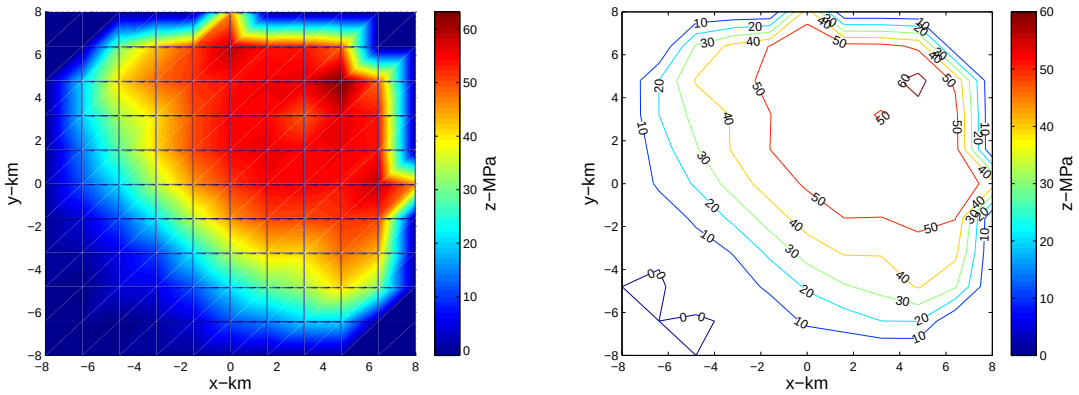
Comparison of Fig 5(b) with Fig 2(a) and of Fig 7(b) with Fig 3(a) indicates a good agreement of the inverted pore pressure decline with the synthetics, violated mostly by the distortions near reservoir edges. Due to the least-squares problem being *ill-conditioned* for large data and model spaces, convergence on denser reservoir grids cannot be achieved without a good initial approximation. This issue has been successfully resolved by using the multigrid approach.

Note that although this test confirms the correct operation of our forward/inverse operator pair, by applying our inverse solver to the data modelled using the *underlying forward-modeling operator* we commit an *inversion crime*. True applicability of our method will be demonstrated in the next section where we apply it to a synthetic extrapolated from real subsidence data. While we are able to recover asymmetric pore pressure decline from as little data as can be contained in just *one* component of the measured displacement field, it is important to understand the sensitivity of our result to the *uncertainty* of the reservoir model parameters. Using operator 5 requires prior knowledge of medium and reservoir parameters, and using Mindlin's analytical expressions for elastostatic Green's tensor assumes *homogeneity* of the medium. In the numerical implementation of our method we effectively depart from the assumption of homogeneity by assuming instead *local* homogeneity or *slow* relative spatial change of medium parameters on the scales of reservoir length and depth.



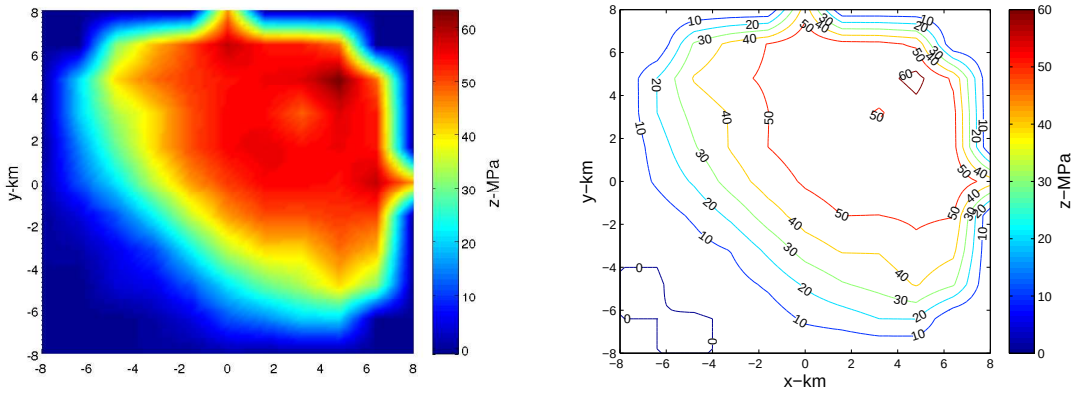
(a) (b)

Figure 5: a) Axisymmetric pore pressure decline inverted from the forward-modelled subsidence of Fig 2(b) on a denser 81 × 81 grid after 2 more solver iterations (total 6). b) Contour plot of the inverted axisymmetric pore pressure of Fig 5(a). [**ER**] musa1/. pinvertedporeoutsymhires,cinvertedporeoutsymhires



(a) (b)

Figure 6: a) Asymmetric pore pressure decline inverted from the forward-modelled subsidence of Fig 3(b) on a sparse 11 × 11 grid after 5 solver iterations. b) Contour plot of the inverted asymmetric pore pressure of Fig 6(a). [**ER**] musa1/. pinvertedporeoutasymlowres,cinvertedporeoutasymlowres



(a) (b)

Figure 7: a) Asymmetric pore pressure decline inverted from the forward-modelled subsidence of Fig 3(b) on a denser 81 × 81 grid after 2 more solver iterations (total 7). b) Contour plot of the inverted asymmetric pore pressure of Fig 7(a). [**ER**] musa1/. pinvertedporeoutasymhires,cinvertedporeoutasymhires

More precisely, our computational framework allows recasting operator 5 as

$$\mathbf{u}(x,y,z) \;=\; m(x,y,z)\mathbf{F}\left[r(\xi,\eta,\zeta)p(\xi,\eta,\zeta)\right] \tag{11}$$

where $\mathbf{F}$ is a homogeneous modeling operator, and $m(x,y,z)$ is a medium heterogeneity factor that, in the simplest case of slow-variability asymptotics (Maharramov, 2011),(Maslov, 1990),(Danilov et al., 1995), is equal to

$$m(x,y,z) = \frac{1}{\overline{\mu}(x,y,z)(1-\overline{\nu}(x,y,z))}, \tag{12}$$

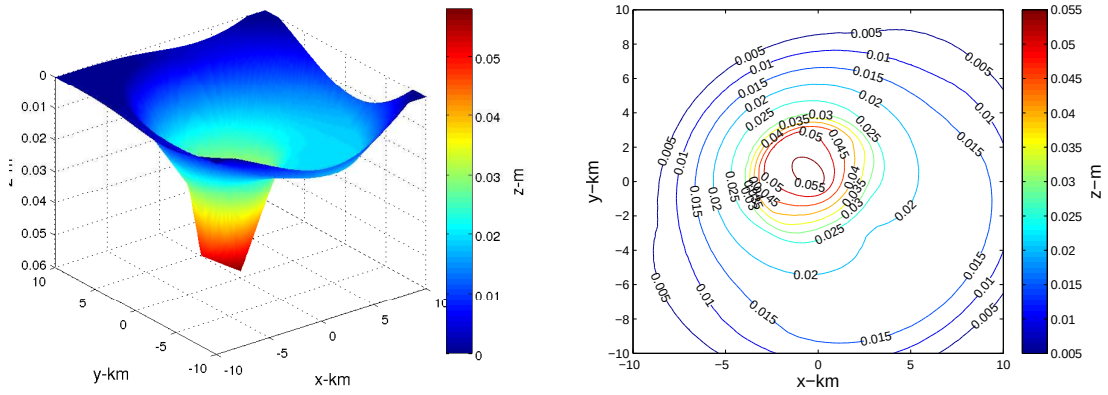
where the overscore in equation 12 represents some average value of the corresponding medium parameter. The scaler r in the right-hand side of equation 11 is related to the heterogeneity of *reservoir* parameters, most notably the Biot coefficient. Equation 11 should include additional terms containing spatial derivatives of the Biot coefficient, but we ignore those terms under the assumption of mild spatial heterogeneity. Equation 11 suggests the following procedure for assessing the impact of the uncertainty in the values of the elastic moduli outside of the reservoir, assuming that r in equation 11 is 1 (i.e., reservoir parameters are integrated into F). The uncertainty of the pore pressure change due to a *concentrated* displacement $\delta(x-x_0, y-y_0, z-z_0)$ (along a fixed coordinate axis) is given by

$$\Delta p(\xi,\eta,\zeta) \;=\; \mathbf{F}^{-1}\left[\delta(x,y,z)\right] \times \Delta\frac{1}{m(x_0,y_0,z_0)}, \tag{13}$$

where the uncertainty $\Delta\frac{1}{m(x_0,y_0,z_0)}$ is computed from the associated uncertainties of the constituent medium parameters (e.g., moduli). Easy computation of the modeling operator and its inverse and the applicability of a multiscale/multigrid approach, as demonstrated in this paper, allow for a straightforward application of equation 13 to either estimating uncertainty or computing the *mathematical expectation* of the inverted pore pressure change.

APPLICATION TO EXTRAPOLATED LACQ SUBSIDENCE

In this section we describe the result of applying our inversion technique to estimating pore pressure decline from the measurements of subsidence for the reservoir described in Segall et al. (1994). Note that the subsidence data in (Segall et al., 1994) was one-dimensional, along the northwest-southeast line, so we extrapolated the data smoothly between the opposite azimuths (see Fig 8(a) and Fig 8(b)).



(a) (b)

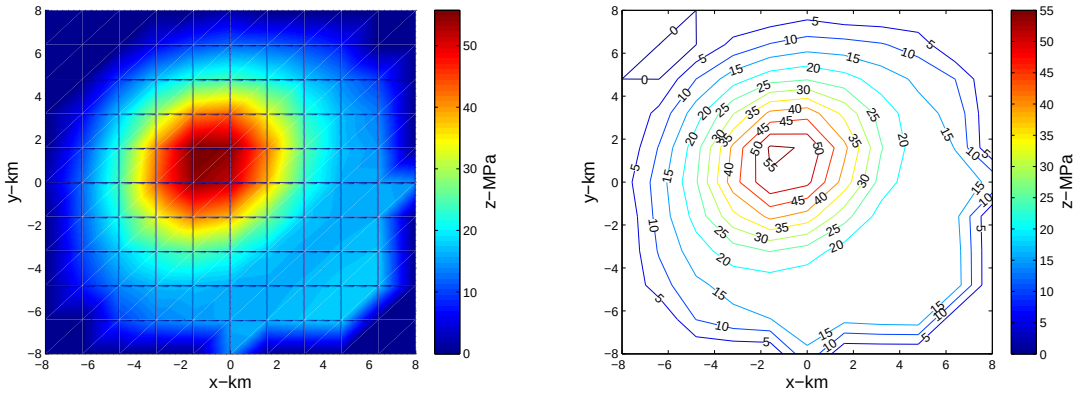
Figure 8: a) Subsidence data from the Lacq reservoir, extrapolated onto a plane from northwest-southeast profile described in (Segall et al., 1994). b) Contour plot of Fig 8(a). [**ER**] musa1/. subslacqsimulated,csubslacqsimulated

Since only the vertical displacement at the surface is known to us, we use the option of the inversion tool (see Appendix B) that constrains the modelled displacement components to just the vertical component. The result of running 2 iterations of our inversion algorithm on an 11×11 grid is shown on Fig 9(a),9(b). Results of the subsequent refinement in 2 more iterations over the 81×81 grid is shown on Fig 10(a),10(b).

And finally, and most importantly, the symmetrized part of the reconstructed pore pressure drop is in excellent agreement with the input parameters of the *symmetric* pore pressure profile used in (Segall et al., 1994) to (approximately) model the same subsidence data but using an *axisymmetric* pressure decline model. In particular, the peak pore pressure drop of 55 `MPa` at the end of the study period has been reconstructed precisely.

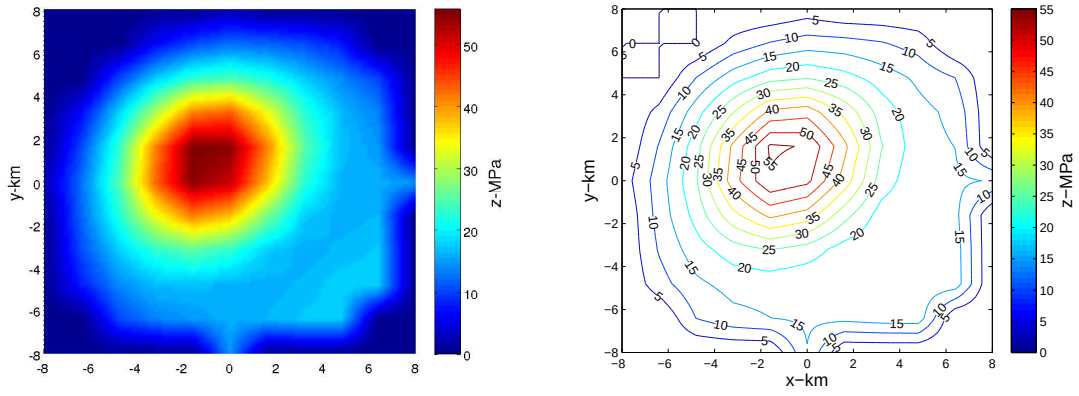
HETEROGENEOUS MODELS

Although the asymptotic technique discussed in an earlier section addresses to some extent the issue of inhomogeneous models, it is inherently limited to moderate heterogeneity. However, practical applications of this method would require a more accurate handling of heterogeneity. Because layered models are of particularly high importance due to their commonality, we first consider modeling displacements for a vertically heterogeneous and horizontally slowly-varying medium. Rather than trying to solve a heterogeneous analogue of system 1 and 2, we will assume that one or all components of the displacement at a fixed



(a) (b)

Figure 9: a) Pore pressure decline inverted from the subsidence of Fig 8(b) on a sparse 11×11 grid after 2 solver iterations. b) Contour plot of the inverted pore pressure decline of Fig 9(a). [**ER**] musa1/. pinvertedLporeoutlacqlowres,cinvertedLporeoutlacqlowres



(a) (b)

Figure 10: a) Pore pressure decline inverted from the subsidence of Fig 8(b) on a denser 81×81 grid after 2 more solver iterations (total 4). b) Contour plot of the inverted pore pressure decline of Fig 10(a). [**ER**] musa1/. pinvertedLporeoutlacqhires,cinvertedLporeoutlacqhires

depth $z = z_{\max}$ immediately above the reservoir are known a priori. For example, we may use operator 5 to model displacements near the reservoir where the effect of the spatial heterogeneity of elastic parameters is limited. With displacements at $z = z_{\max}$ and free-surface boundary conditions at $z = 0$, the problem of modeling subsurface displacements is reduced to solving a boundary-value problem for the elastostatic system:

$$\begin{aligned} \mu\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)+\frac{2\mu\nu}{1-2\nu}\frac{\partial u_k}{\partial x_k}\delta_{ij} =&\sigma_{ij},\\ \frac{\partial \sigma_{ij}}{\partial x_j} =&0, \end{aligned} \tag{14}$$

where indices run from 1 to 3, σ_{ij} denote the *stress tensor* components, summation is carried out on repeating indices and body-force distribution is zero. The boundary conditions used with system 14 are the following:

$$\begin{aligned} \left(\frac{\partial u_i}{\partial x_3}+\frac{\partial u_3}{\partial x_i}\right)+\frac{2\nu}{1-2\nu}\frac{\partial u_k}{\partial x_k}\delta_{i3}|_{z=z_0} =&0,\\ u_i(x,y,z_{\mathrm{R}}) =&s_i(x,y), \end{aligned} \tag{15}$$

where $s_i(x,y),\ i=1,2,3$ is a known displacement field at a fixed depth. Although system 14 is comprized of purely elastostatic equations, it allows us to model fluid-to-solid coupling via the boundary condition at $z = z_{\mathrm{R}}$ that can be approximately computed using operator 5. For a laterally-homogeneous medium – or under the assumption of slow lateral variability and pseudo-differential operator ordering, (Maslov, 1976) – equations 14 can be Fourier-transformed in x_1, x_2, and the resulting system discretized in depth:

$$\begin{aligned} \frac{v_1(z+\Delta z)-v_1(z-\Delta z)}{2\Delta z} =&v_4(z)\\ \frac{v_2(z+\Delta z)-v_2(z-\Delta z)}{2\Delta z} =&v_5(z)\\ \frac{v_3(z+\Delta z)-v_3(z-\Delta z)}{2\Delta z} =&v_6(z)\\ \frac{\mu(z+\Delta z)v_4(z+\Delta z)-\mu(z-\Delta z)v_4(z-\Delta z)}{2\Delta z} =&\\ -(k_x^2+k_y^2+\frac{k_x^2}{1-2\nu(z)})\mu(z)v_1(z)-\frac{k_xk_y}{1-2\nu(z)}\mu(z)v_2(z)+\frac{ik_x}{1-2\nu}\mu(z)v_6(z)&\\ \frac{\mu(z+\Delta z)v_5(z+\Delta z)-\mu(z-\Delta z)v_5(z-\Delta z)}{2\Delta z} =&\\ -(k_x^2+k_y^2+\frac{k_y^2}{1-2\nu(z)})v_2(z)-\frac{k_xk_y}{1-2\nu(z)}v_1(z)+\frac{ik_y}{1-2\nu}v_6(z)&\\ \frac{\mu(z+\Delta z)v_6(z+\Delta z)-\mu(z-\Delta z)v_6(z-\Delta z)}{2\Delta z} =&\\ \frac{1}{1+\frac{1}{1-2\nu(z)}}\left[-(k_x^2+k_y^2)v_3(z)+\frac{ik_x}{1-2\nu(z)}v_4(z)+\frac{ik_y}{1-2\nu(z)}v_5(z)\right],& \end{aligned} \tag{16}$$

where k_x, k_y are the horizontal wavenumbers and Δz is a depth step, $v_{1,2,3}$ are Fourier-transforms of the three displacement components $u_{1,2,3}$ and $v_{4,5,6}$ are their z-derivatives. At the boundaries $z = 0$ and $z = z_{\mathrm{R}}$ the central differences should be replaced with backward and forward differences (Iserles, 2008), and the boundary conditions 15 Fourier-transformed

in a similar manner. In combination with the Fourier-transformed boundary conditions the above system is reduced to independent $6N_z \times 6N_z$ linear systems for finding $v_i(\Delta z_j), i = 1, \ldots, 6, j = 1, \ldots, N_z$ for each wavenumber pair k_x, k_y, where N_z is the number of depth steps.

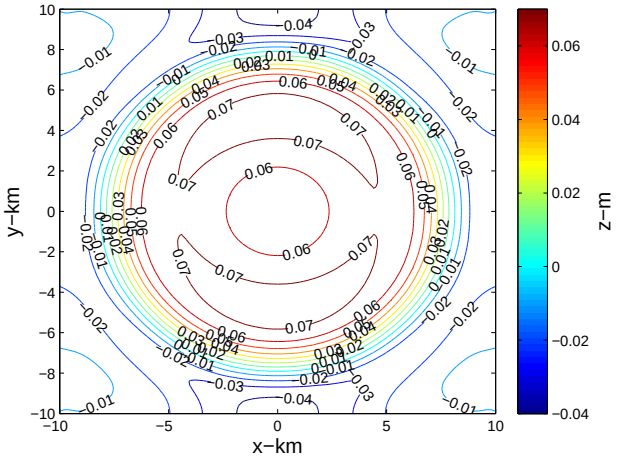


Figure 11: Contour plot of the displacements modelled from the axisymmetric pore pressure decline of Fig 2(a). [**ER**] musa1/. csubszdispoutsymhires

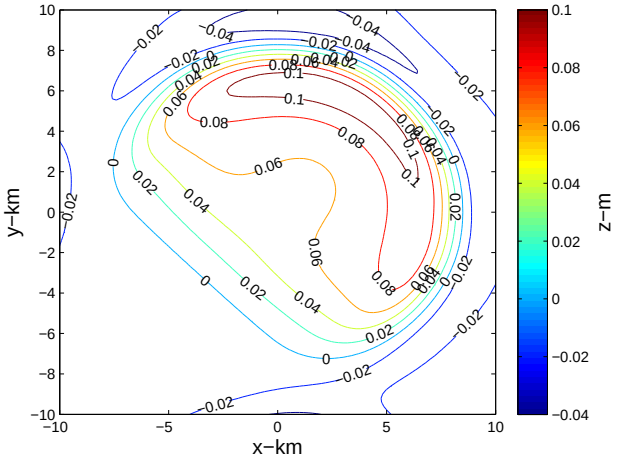


Figure 12: Contour plot of the displacements modelled from the asymmetric pore pressure decline of Fig 3(a). [**ER**] musa1/. csubszdispoutasymhires

Solution of the above system is efficiently parallelized, with individual $6N_z \times 6N_z$ sparse systems solved independently. Furthermore, each of the systems is *banded* with the bandwidth of 13 elements and therefore can be solved in a linear time and memory $O(N_z)$ (Trefethen and Bau, 1997).

Fig 11 and Fig 12 show the results of modeling surface subsidence from the axisymmetric and asymmetric pore pressure decline synthetics of Fig 2(a) and Fig 3(a). Here $N_z = 20$ with a 100 `m` depth step, the displacement field at the depth of 2 `km` was computed using

operator 5. Although the above approach allows both elastic medium parameters (e.g., shear modulus μ and Poisson's ratio ν) to be vertically heterogeneous, the latter, as the ratio of the axial and transverse strains, is usually less affected by compaction, and hence we left it constant at .25. However, the depth-dependent shear modulus is given by the formula

$$\mu(z) = \frac{23\texttt{GPa}}{1. + (z_{\rm R} - z)/z_{\rm R}}. \tag{17}$$

Comparison of Fig 11 and Fig 12 with the results of Fig 2(b) and Fig 3(b) obtained above using a homogeneous model indicate larger subsidence in the heterogenous case that is consistent with a greater compliance of the overburden as determined by the heterogeneous shear modulus of equation 17.

Although depth-varying models are common in geomechanical applications, and the diffusive nature of production-induced deformation favors slowly-varying models, there exist practical applications where strong lateral heterogeneity should be taken into account (for example, in subsalt regions). The widely accepted approach to tackling such problems consists of application of the finite elements method (Iserles, 2008) to the coupled poroelastic system (Kosloff et al., 1980). While finite elements can handle arbitrary spatial heterogeneity, the main disadvantage of this approach is the necessity to solve a potentially very large system of linear equations with very sparse but generally *unstructured* matrix.

A possible extension of our approach for tackling arbitrary heterogeneity could be summarized as follows. If system 14 can be factorized

$$\begin{aligned}\nabla \cdot \mu(x_1,x_2,x_3) \left[\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) + \frac{2\nu}{1-2\nu} \frac{\partial u_k}{\partial x_k} \delta_{ij} \right] = \\ \left(c_1 \frac{\partial}{\partial x_3} + P_1(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}) \right) \times \\ \left(c_2 \frac{\partial}{\partial x_3} + P_2(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}) \right) u,\end{aligned} \tag{18}$$

where $P1$ and $P2$ are some pseudo-differential operators and c_1, c_2 are some functions, then given the boundary conditions 15, the boundary-value problem 14,15 can be solved by solving

$$\left(c_1 \frac{\partial}{\partial x_3} + P_1(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}) \right) U = 0, \tag{19}$$

upward, starting from the initial condition at depth $z = z_{\rm R}$, followed by solving

$$\left(c_2 \frac{\partial}{\partial x_3} + P_2(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}) \right) u = U, \tag{20}$$

downward, starting from the free-surface boundary condition on the surface. Some factorisations 18 are trivial, as when the operator 20 maps the displacement field to the stress field, we effectively recover our original elastostatic system and obtain the well-known depth-extrapolation method for 6 variables (Segall, 2010). However, finding a factorisation that requires less than 6 variables *may* potentially lead to a method where the eigenmodes that correspond to exponentially increasing and decreasing solutions decouple in up and down-going extrapolation operators similar to the one-way wave equation (Claerbout, 2011). However, the author's attempts to construct linear operator factorisations 18 involving operators

mapping only 3-component vector functions so far resulted in rank-deficient systems. The three displacement components may have to be complemented with an additional variable in order to achieve a pseudo-differential operator factorisation. The benefits of finding such factorisations and stable methods of solving factorized equations would extend beyond the solution of the uncoupled poroelastostatic system.

PREDICTING VELOCITY AND DENSITY CHANGES FROM DEFORMATION

Keeping in sight our ultimate objective of using geomechanical observables to regularize seismic inversion, we now turn to the question of estimating change in velocity and density due to production-induced deformation. At this point we assume that we have inverted the pore pressure decline from known subsidence and can forward-model spatial displacements in the overburden due to the pore pressure change. Given a modelled displacement field u_{123}, we can compute the strain

$$\epsilon_{ij} = \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \tag{21}$$

and compute density change as

$$\Delta\rho(x_1, x_2, x_3) \approx -\rho(x_1, x_2, x_3)\epsilon_{ii}, i = 1, \ldots, 3. \tag{22}$$

Given the estimated density change 22 and strain 21, we can proceed to estimating the induced change in seismic velocities in two different but potentially complementary ways. The first approach would be to use empirical velocity-density relations (Mavko et al., 2009),(Gardner et al., 1974). However, polynomial and power-law empirical forms of Gardner's relations may produce significant errors due to the presence of cracks or flaws in the rock. While useful as an average over many rock types, the applicability of such empirical relations to estimating minute changes in pressure and shear wave velocities at this point is moot. A more promising possibility is to use a linear empirical relation between the pressure-wave velocity and path-length change

$$\frac{dV_P}{V_P} = -R\epsilon_{33}, \tag{23}$$

where V_P is the pressure wave velocity (for normal-incidence pressure waves). While the dimensionless coefficient R in equation 23 is generally unknown, and estimates of R from empirical relations for different rock type suffer from the same kind of uncertainties as the velocity-density relations, a reasonable estimate can be obtained from observable time-shifts using the relation

$$\frac{dt}{t} = (1 + R)\epsilon_{33} \tag{24}$$

(Hatchell and Bourne, 2005). The time-shifts can be extracted from time-lapse seismic data using a cross-correlation technique similar to the one used by Hale (2009) and Ayeni (2011). We propose to use equation 24 to estimate the coefficient R where the time-shifts can be resolved, and then use the obtained value to estimate the velocity change (and time-shifts) from equations 23 and 24 where the time-lapse data has illumination gaps or is noisy.

CONCLUSION AND PERSPECTIVES

We have developed a computationally efficient method for modeling deformation from pore pressure decline and for inverting pore-pressure change from observable deformations/subsidence. The method has been implemented for both slowly-varying and layered media and tested on synthetics based on real-world examples of production-induced subsidence. We propose an approach to using observable deformations for estimating velocity change based on empirical rock-property relations.

While all the main building blocks required for regularizing time-lapse inversion with subsidence-based constrains are in place, an integrated seismic/geomechanical data set is required to validate our approach. While the deformation modeling and pore pressure inversion components can be tested on synthetic data, testing the regularization approach requires real data. Time-lapse synthetics would be generated using the same kind of techniques for modeling production-induced velocity change as we intend to use for the regularization, hence purely synthetic tests cannot completely validate our approach.

ACKNOWLEDGEMENTS

The author would like to thank Biondo Biondi for suggesting the problem of using geomechanical/geological constraints for the regularization of seismic inversion, Mark Zoback for his excellent course of reservoir geomechanics, and Paul Segall, Stewart Levin, Robert Clapp, David Nichols and Jon Claerbout for useful discussions and suggestions.

APPENDIX A

The elastostatic Green's tensor $g_i^k(x,y,z,\xi,\eta,\zeta)$ has the meaning of the displacement along axis i at point (x,y,z) due to a concentrated force along axis k at point (ξ,η,ζ). The analytical expression for the components of the Green's tensor in the elastic half-space with a free-surface boundary condition are given by the following equations (Mindlin, 1936):

$$
\begin{aligned}
g_1^1 =& w\left(\frac{3-4\nu}{r_1}+\frac{1}{r_2}+\frac{(x-\xi)}{r_1^3}+\frac{(3-4\nu)(x-\xi)^2}{r_2^3}\right)+\\
&+w\left(\frac{2(r_2^2-3(x-\xi)^2)z\zeta}{r_2^5}+\frac{4(1-\nu)(1-2\nu)(r_2^2-(x-\xi)^2-r_2(z+\zeta))}{r_2(r_2-z-\zeta)^2}\right)\\
g_2^1 =& (x-\xi)(y-\eta)w\left(\frac{1}{r_1^3}+\frac{3-4\nu}{r_2^3}-\frac{6z\zeta}{r_2^5}-\frac{4(1-\nu)(1-2\nu)}{r_2(r_2-z-\zeta)^2}\right)\\
g_3^1 =& (x-\xi)w\left(\frac{z-\zeta}{r_1^3}+\frac{(3-4\nu)(z-\zeta)}{r_2^3}-\frac{4(1-\nu)(1-2\nu)}{r_2(r_2-z-\zeta)}-\frac{6z\zeta(z+\zeta)}{r_2^5}\right)
\end{aligned}
$$

$$
\begin{aligned}
g_1^2 =& g_2^1 \\
g_2^2 =& w\left(\frac{3-4\nu}{r_1}+\frac{1}{r_2}+\frac{(y-\eta)}{r_1^3}+\frac{(3-4\nu)(y-\eta)^2}{r_2^3}\right)+ \\
&+w\left(\frac{2(r_2^2-3(y-\eta)^2)z\zeta}{r_2^5}+\frac{4(1-\nu)(1-2\nu)(r_2^2-(y-\eta)^2-r_2(z+\zeta))}{r_2(r_2-z-\zeta)^2}\right) \\
g_3^2 =& (y-\eta)w\left(\frac{z-\zeta}{r_1^3}+\frac{(3-4\nu)(z-\zeta)}{r_2^3}-\frac{4(1-\nu)(1-2\nu)}{r_2(r_2-z-\zeta)}-\frac{6z\zeta(z+\zeta)}{r_2^5}\right)
\end{aligned}
$$

$$
\begin{aligned}
g_1^3 =& (x-\xi)w\left(\frac{z-\zeta}{r_1^3}+\frac{(3-4\nu)(z-\zeta)}{r_2^3}+\frac{4(1-\nu)(1-2\nu)}{r_2(r_2-z-\zeta)}+\frac{6z\zeta(z+\zeta)}{r_2^5}\right) \\
g_2^3 =& (y-\eta)w\left(\frac{z-\zeta}{r_1^3}+\frac{(3-4\nu)(z-\zeta)}{r_2^3}+\frac{4(1-\nu)(1-2\nu)}{r_2(r_2-z-\zeta)}+\frac{6z\zeta(z+\zeta)}{r_2^5}\right) \\
g_3^3 =& w\left(\frac{3-4\nu}{r_1}+\frac{5-12\nu+8\nu^2}{r_2}+\frac{(z-\zeta)^2}{r_1^3}+\frac{(3-4\nu)(z+\zeta)^2-2z\zeta}{r_2^3}+\frac{6z\zeta(z+\zeta)^2}{r_2^5}\right)
\end{aligned}
$$

$$
\begin{aligned}
w =& \frac{1}{16\pi\mu(1-\nu)} \\
r_1 =& \sqrt{(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2} \\
r_2 =& \sqrt{(x-\xi)^2+(y-\eta)^2+(z+\zeta)^2}
\end{aligned}
$$

APPENDIX B

The methods of this paper have been implemented in the author's Fortran 2003 object-oriented framework `exp_tk`. A parallelized forward-modeling and inversion utility `PEMODEL` for modeling and inverting 3D displacements has been implemented by the author as part of this study, and as a joint exercise with applying the author's new optimisation framework to problems of geophysical imaging and computational geophysics (Maharramov, 2011). `PEMODEL` can be found at

`EXP_TK_INSTALL_DIR/exp_tk_0.5/src/tests/poroelastic_deform`

Source file comments contain a description of the command-line options used with the utility. Fortran F003 module `poroelastic_green.F90` implements the poroelastostatic modeling operator and its adjoint. Solution of problem 14,15 is implemented in the forward-modeling utility `ZPEMODEL`. `ZPEMODEL` can be found at

`EXP_TK_INSTALL_DIR/exp_tk_0.5/src/tests/depth`

Module `poroelastic_depth.F90` implements the corresponding modeling operator. Banded systems 16 are solved using LAPACK95 computational routine `gbtrs` (Barker et al., 2001) following optional equilibration to improve the stability for very sharp vertical contrasts of elastic moduli.

The framework has no external dependencies except for Intel Fortran compiler, version 12.0 or higher and Intel Math Kernel Library version 9.0 or higher. The plots of this report have been generated using Matlab from the binaries and headers output by `PEMODEL` and `ZPEMODEL`.

REFERENCES

Ayeni, G., 2011, Time-lapse seismic imaging by linearized joint inversion. Ph.D. Thesis: Stanford Exploration Project.

Barker, V. A., L. S. Blackford, J. Dongarra, J. D. Croz, S. Hammarling, M. Marinova, J. Wasniewski, and P. Yalamov, 2001, LAPACK95 User's Guide: SIAM.

Bjork, A., 1996, Numerical methods for least squares problems: SIAM.

Claerbout, J., 2011, Basic earth imaging: Stanford Exploration Project.

Danilov, V. G., V. P. Maslov, and K. A. Volosov, 1995, Mathematical modelling of heat and mass transfer processes: Springer.

Gardner, G. H. F., L. Gardner, and A. Gregory, 1974, Formation velocity and density – the diagnostic basics for stratigraphic traps: Geophysics, **39**, 770 – 778.

Geertsma, J., 1973, Land subsidence above compacting oil and gas reservoirs: SPE-AIME European Spring Meeting.

Hale, D., 2009, A method for estimating apparent displacement vectors from time-lapse seismic images: Geophysics, **74**, V99–V107. (http://link.aip.org/link/?GPY/74/V99/1).

Hatchell, P. and S. Bourne, 2005, Rocks under strain: Strain-induced time-lapse time shifts are observed for depleting reservoirs: The Leading Edge, **24**, 1222–1225. (http://link.aip.org/link/?LEE/24/1222/1).

Iserles, A., 2008, A first course in the numerical analysis of differential equations: Cambridge University Press.

Kosloff, D., R. Scott, and J. Scranton, 1980, Finite element simulation of Wilmington oil field subsidence: I. Linear modelling: Tectonophysics, **65**, 339 – 368.

Maharramov, M., 2011, Frontiers of geophysical research: Novel asymptotic and numerical methods in geophysical applications: Frontiers of Geophysical Research Proposal.

Maslov, V., 1976, Operational methods: Mir Publishers.

Maslov, V. P., 1990, Asymptotic solutions of equations with slowly varying coefficients: Séminaire Équations aux dérives partielles (dit Goulaouic-Schwartz).

Mavko, G., T. Mukerji, and J. Dvorkin, 2009, The rock physics handbook: Cambridge University Press.

Mindlin, R. D., 1936, Force at a point in the interior of a semi-infinite solid: Physics, **7**, 195–202. (http://link.aip.org/link/?JAP/7/195/1).

Segall, P., 1992, Induced stresses due to fluid extraction from axisymmetric reservoirs: Pure and Applied Geophysics, **139**, no. 3, 535–560. (http://dx.doi.org/10.1007/BF00879950).

———, 2010, Earth and volcano deformation: Princeton University Press.

Segall, P. and S. D. Fitzgerald, 1998, A note on induced stress changes in hydrocarbon and geothermal reservoirs: Tectonophysics, **289**, 117 – 128.

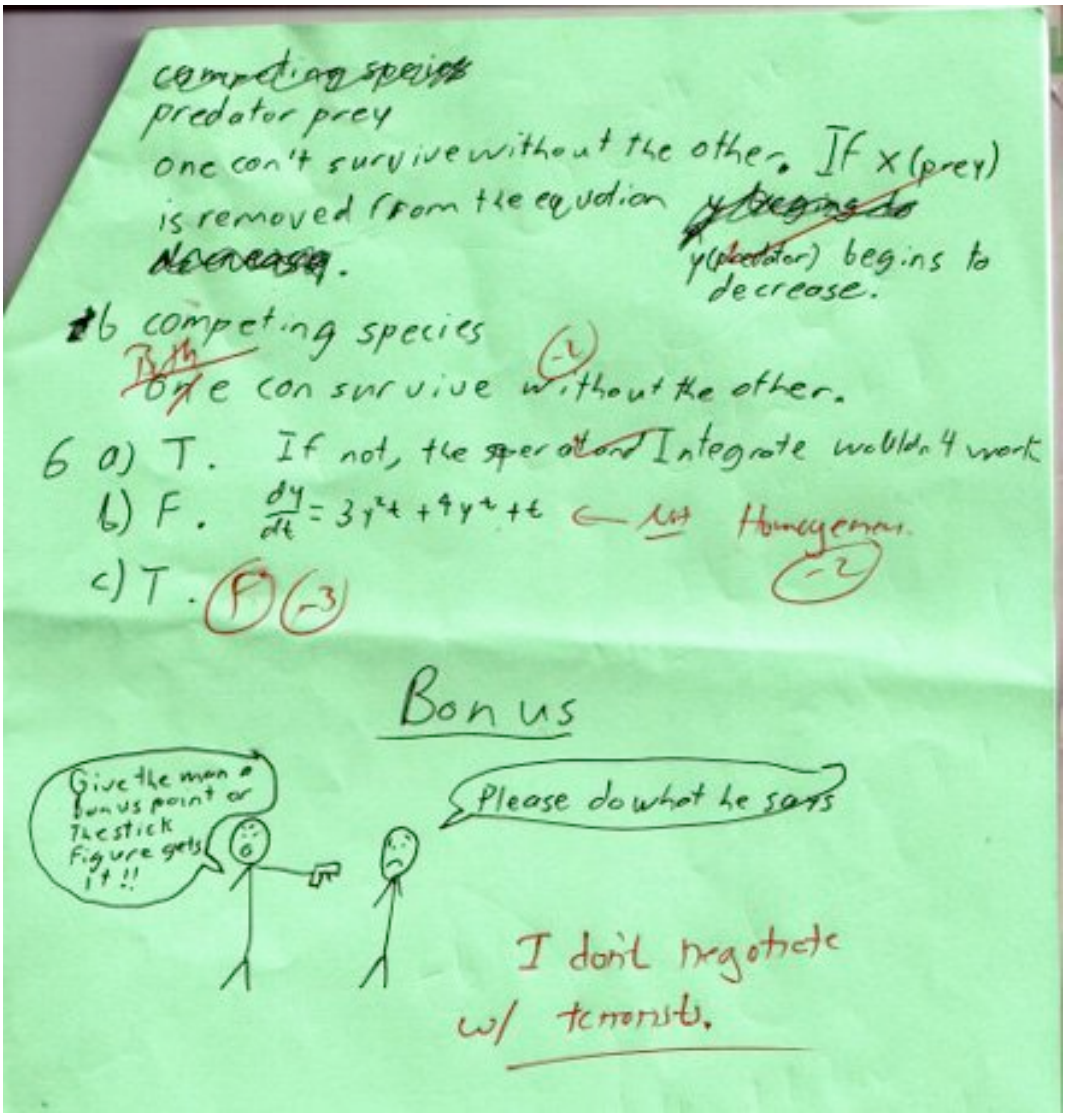
Segall, P., J.-R. Grasso, and A. Mossop, 1994, Poroelastic stressing and induced seismicity near the Lacq gas field, southwestern France: Journal of Geophysics Research, **99**, no. B8, 15423–15438. (http://dx.doi.org/10.1029/94JB00989).

Trefethen, L. N. and D. Bau, 1997, Numerical linear algebra: SIAM.

Wang, H., 2000, Theory of linear poroelasticity with applications to geomechanics and hydrogeology: Princeton University Press.

Zoback, M., 2010, Reservoir geomechanics: Cambridge University Press.

Zoback, M. D. and J. C. Zinke, 2002, Production-induced normal faulting in the Valhall and Ekofisk oil fields: Pure and Applied Geophysics, **159**, no. 1, 403–420. (http://dx.doi.org/10.1007/PL00001258).


competing species
predator prey
one can't survive without the other. If x(prey) is removed from the equation
y(predator) begins to decrease.
#6 competing species
Both One can survive without the other. (-1)
6 a) T. If not, the operator Integrate wouldn't work
b) F. $\frac{dy}{dt} = 3y^2t + 4y^2 + t$ ← Not Homogeneous. (-2)
c) T. (F) (-3)
Bonus
Give the man a bonus point or the stick figure gets it!!
Please do what he says
I don't negotiate w/ terrorists.

Correlation energy between surface and borehole stations at the Valhall field

Jason P. Chang and Sjoerd de Ridder

ABSTRACT

Ocean-bottom cables at the Valhall field have provided an abundance of passive seismic data for testing the potential of seismic interferometry. We cross-correlate recordings from surface stations and 2-km deep borehole stations. Results show correlated energy at frequencies between 0.175 Hz and 1.75 Hz. However, the signal we retrieved is not time-symmetric, as there are multiple arrivals at acausal correlation time lags compared to the one arrival visible at causal time lags. The apexes of the causal events are found at acausal time lags rather than at zero time lag. The virtual source is centered northwest of the borehole stations at the offshore platform. We conclude that these observations are due to the borehole acting as a wave-guide. Because this mechanism does not satisfy the conditions of seismic interferometry, we cannot interpret these cross-correlation results as inter-station Green's functions.

INTRODUCTION

Traditional seismic surveys for subsurface imaging deploy a collection of receivers and a controlled source. These surveys are often expensive to deploy and require much effort to oversee. Passive seismic interferometry does not require a controlled source. Cross-correlating recordings of ambient seismic noise at two receivers turns one of the receivers into a virtual source and extracts the time it takes the energy from that virtual source to reach the other receiver of interest. This travel-time information can then be used for purposes such as velocity modeling.

Tests of the practicality of passive seismic interferometry for these purposes has been ongoing with ambient seismic noise data provided by the permanent ocean-bottom cable (OBC) array at the Valhall field. Artman (2007) and Landes et al. (2009) showed that virtual omnidirectional point sources could not be captured at frequencies typical of seismic exploration (3-60 Hz). Later, de Ridder and Dellinger (2011) showed that virtual low-frequency (0.35-1.75 Hz), omnidirectional Scholte waves along the ocean floor could be generated from seismic noise. They were able to use the Scholte-wave travel-time information for tomographic imaging of structures in the near-surface (0-150 m).

In previous studies of passive seismic interferometry for seismic exploration at the Valhall field (Artman, 2007; Bussat and Kugler, 2009; de Ridder and Dellinger, 2011) correlations were performed between stations on the ocean floor. In this study we correlate ambient seismic noise recorded over the same time period by the OBC array and by five downhole stations at nearly 2 km depth. Our goal is to test whether Green's functions can be extracted between the borehole stations and the surface stations. In this report, we present the results for low frequencies (0.175-1.75 Hz).

PASSIVE SEISMIC INTERFEROMETRY

In passive seismic interferometry, receivers record data from passive sources such as ambient seismic noise. Under appropriate conditions, cross-correlating two receiver recordings recovers a 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 equation form

$$[G(x_B, x_A, t) + G(x_B, x_A, -t)] * S_N(t) = \langle u(x_B, t) * u(x_A, -t) \rangle \quad , \tag{1}$$

where G is the Green's function between two receiver locations (x_A, x_B), $S_N(t)$ is the autocorrelation of the source function (here it is noise), and u is the observed wavefield at a given receiver location. Convolving the Green's function with the autocorrelation of the source function can reveal the time it takes for a virtual source signal at one receiver to reach the other receiver of interest, whether directly or after reflecting in the subsurface. If investigating direct waves, the travel-time information with the known distance between the two receivers produces an estimate of the average velocity along the path traveled. Using only surface stations and direct waves, de Ridder and Dellinger (2011) successfully constructed a velocity model of the shallow subsurface at Valhall. Now we want to incorporate recordings at borehole stations into our passive seismic interferometry. To determine whether we can extract the Green's functions between borehole and surface stations, we start by looking at spectrograms and by cross-correlating the borehole and surface station recordings.

VALHALL SURFACE STATION AND BOREHOLE STATION RECORDINGS

In late December of 2010, 2224 4-component stations on the seafloor and 5 3-component stations in a single borehole recorded ambient seismic noise at the Valhall field continuously for approximately 5 days. The ocean bottom stations were densely sampled in-line (50 m) and sparsely sampled cross-line (275 m). The borehole stations were spaced roughly 8 m apart laterally and 11 m apart vertically within a deviated wellbore at a depth of almost 2 km. They have average Universal Transverse Mercator (UTM) coordinates 525.5 km easting and 6236.5 km northing. Figure 1 shows the surface array and the surface projections of the borehole stations in map view. The borehole extends from an offshore platform, which is located at approximate UTM coordinates 524.5 km easting and 6237.0 km northing.

Prior to creating any spectrograms or performing any cross-correlations, we had to adjust the borehole recordings. The three components at each borehole station were independently directed roughly north, west, and downward. Using a direction cosine matrix calculated from the provided azimuth and dip information of each component at each borehole station, we transformed these components to due north, due west, and direct downward.

SURFACE AND BOREHOLE STATION SPECTROGRAMS

A spectrogram is a plot of spectra versus time. Here we created three of them to compare the average spectral content of the data recorded by the OBC array and by the borehole stations. These spectrograms provide a first look at whether there is seismic energy at the

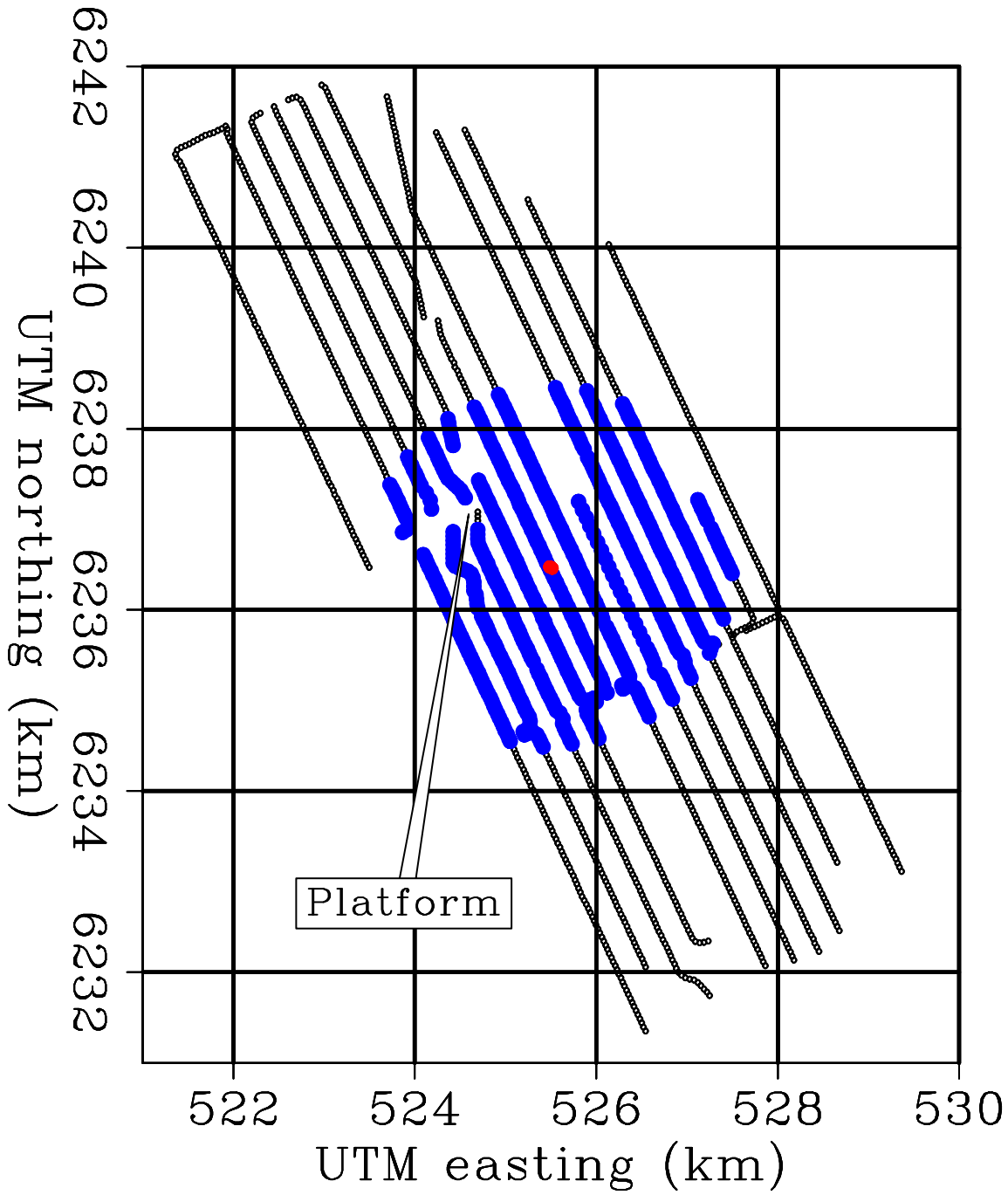


Figure 1: Distribution of seismic stations at Valhall. Black and blue circles are surface stations, with blue representing those stations within 2 km of the average surface projection of the five borehole stations. Red circles depict the five borehole stations, which are all at nearly 2 km depth. The location of the platform is also shown. [**CR**] jasonpc1/. xytot

low-frequency (0.175-1.75 Hz) range we are interested in. We calculated spectrograms for the vertical components of the surface stations within 2 km of the average borehole station location, the vertical components of the five surface stations nearest the platform, and the vertical components of four (out of five) borehole stations. The locations of these sets of stations are shown in Figure 1. We chose surface stations within 2 km of the borehole stations because it was a range comparable to the depth of the borehole stations. We chose surface stations closest to the platform because we wanted to compare signals in the borehole and signals at the entrance of the borehole. We considered only four borehole stations because one of the stations had much lower amplitude correlation results (Figure 4), which might be caused by processing errors.

To generate the spectrograms, we divided recordings at all stations into 2.5-minute overlapping segments. The mean was removed from each of the time segments prior to calculating the frequency spectrum. We then averaged the resulting spectra for each time segment over the selected stations and plotted them over the entire 5-day recording period.

Figure 2 shows the three spectrograms that we calculated. It is evident from Figure 2(a) that frequencies from near 0 Hz to 1.75 Hz consistently contain an abundance of energy at the surface stations. The surface stations nearest the platform, as seen in Figure 2(b), show similar trends over time but with weaker energy. This low-frequency energy is least evident in Figure 2(c), where there is high energy at very specific frequencies at the borehole stations. These dominant frequencies at the borehole stations are likely hiding the low-frequency energy that is clearly visible in the spectrograms of the surface stations.

CROSS-CORRELATIONS BETWEEN BOREHOLE AND SURFACE STATIONS

We cross-correlate the borehole and surface station recordings to determine whether we can extract the Green's functions between them. Prior to the cross-correlations, data had to go through some processing. The five days of data at each station came in 3-hour and 15-minute segments. We tapered the first and last 15 minutes of each segment with a smooth cosine-squared taper to avoid truncation artifacts that could arise when bandpassing. We then bandpassed the segments of data recorded at both borehole and surface stations between 0.175 Hz and 1.75 Hz to be similar to the virtual Scholte-wave frequencies seen in de Ridder and Dellinger (2011). For time sections displaying instrument malfunction, we centered a smooth notch taper at the problematic section prior to bandpassing. After bandpassing, we removed the resulting artificial spikes from problematic sections by another notch taper.

We cross-correlated all possible combinations of borehole and station recordings for each time segment. This consisted of all three components at each borehole station and the vertical component at each surface station. For each cross-correlation, we time-reversed the borehole recording before convolution. We then stacked cross-correlation results from all time patches to improve the signal-to-noise ratio.

Figure 3 displays the cross-correlation results between 2140 surface stations and the shallowest vertical borehole component (1.934 km depth). Cross-correlations between vertical components of other borehole stations and the surface stations produced similar results. There is generally one arrival visible in the causal correlation time lags but many in the acausal (negative) time lags. The event in (mostly) causal times forms a cone with the

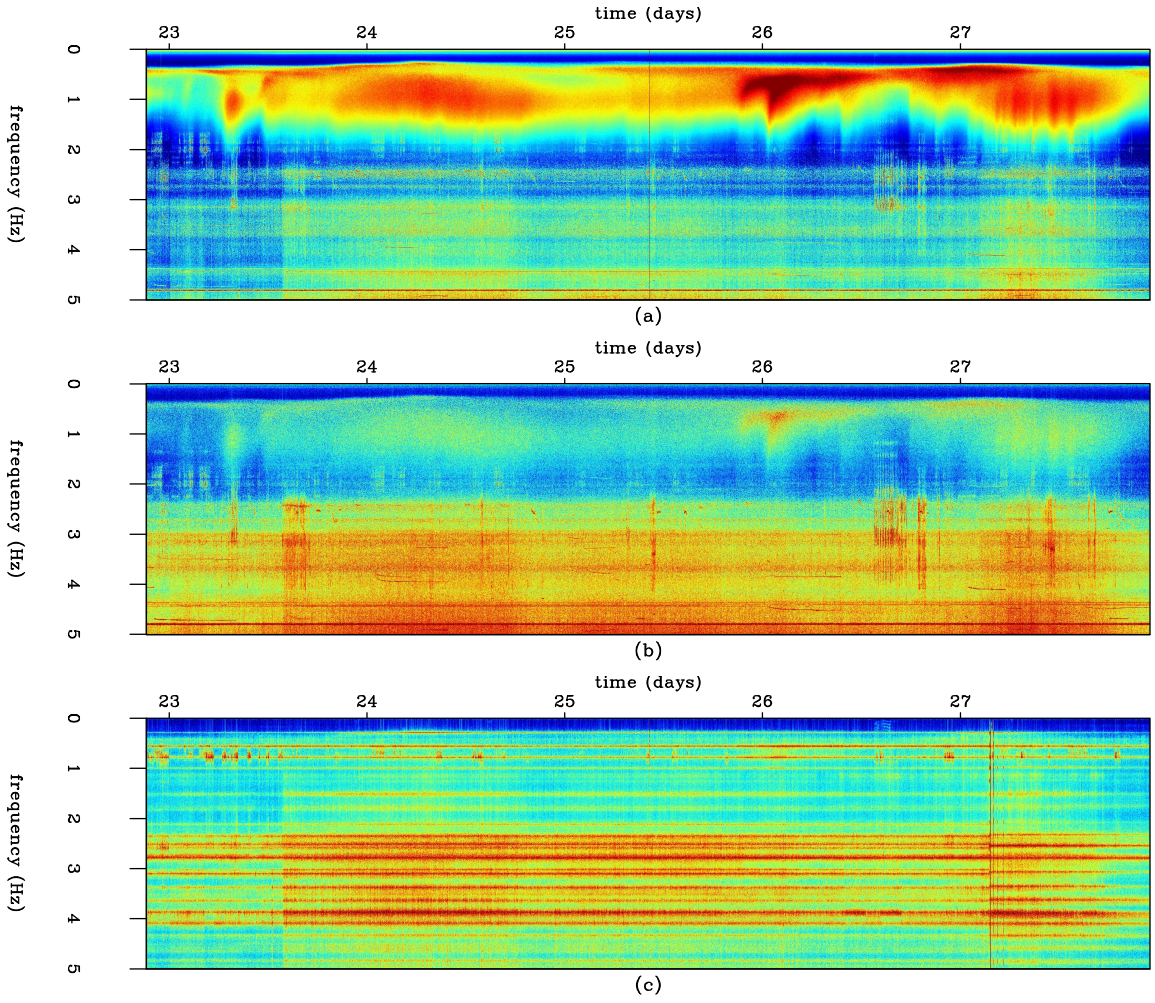


Figure 2: (a) Average spectrogram of surface stations within 2 km of the borehole stations. (b) Average spectrogram of the five surface stations nearest the offshore platform. (c) Average spectrogram of the vertical components of the borehole stations. Note the energy in the lower frequencies. [**CR**] jasonpc1/. spectragram

apex at a negative correlation time lag (approximately $t = -2$ s). There is a mirror event (with time-reverse symmetry) in acausal times; the time of symmetry is again approximately $t = -2$ s. The acausal time window also contains a series of copies of both events at increasing negative time lags. The coherent energy is of relatively high frequency, and the wavelength is a few hundred meters. With wavelengths this short, coherent energy between the surface stations and stations at 2 km depth cannot be explained by the skin depth of surface waves.

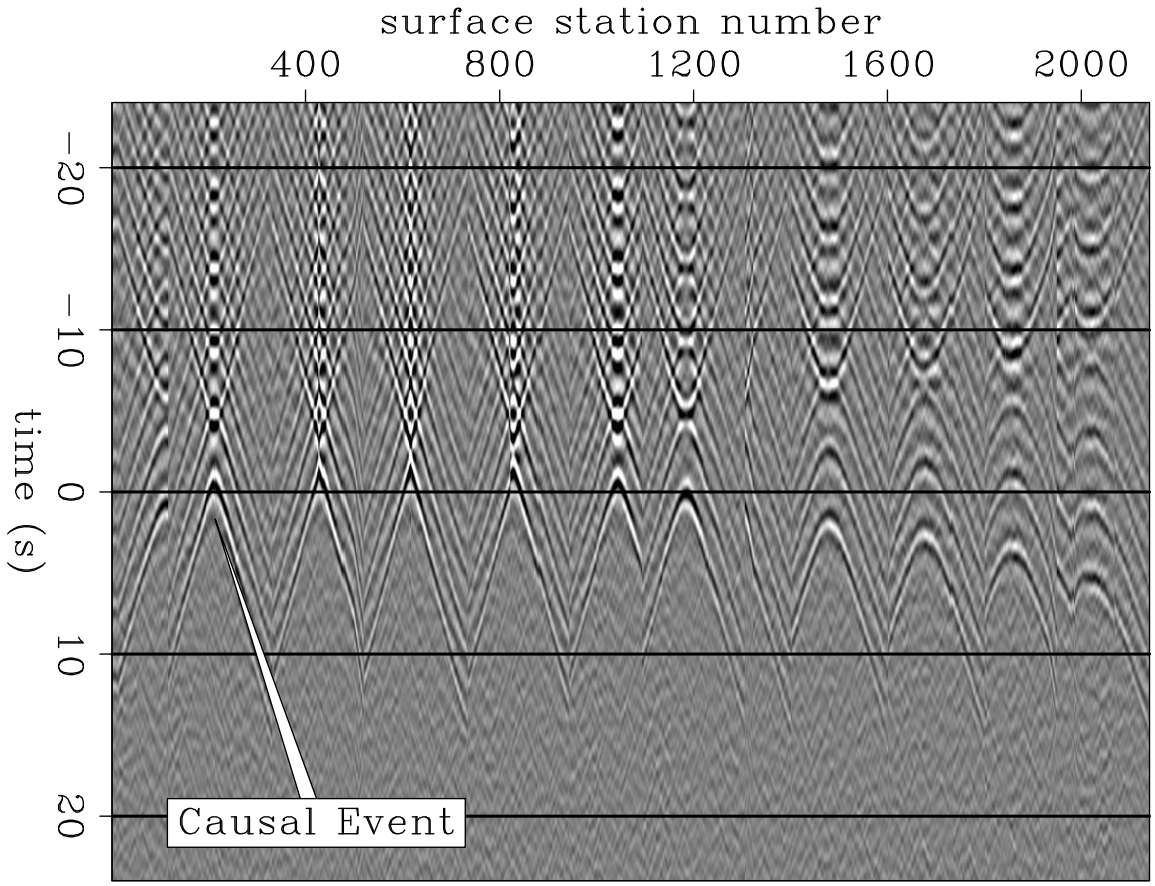


Figure 3: Result after cross-correlations between the shallowest borehole station (at 1.934 km depth) and the surface stations. [**CR**] jasonpc1/. correlation

Figure 4 shows traces representing the correlation results between a surface station roughly 1.8 km north of the borehole stations and the vertical components of the borehole stations. The traces depict very similar results at both causal and acausal times. The main difference is that the correlation with the fourth borehole station is lower in amplitude. This might be due to an incorrect rotation of that particular borehole.

Another way to look at the correlation results is in map view. Figure 5 plots the correlation results when $t = 2$ s. A strong amplitude, circular wavefront is formed by the coherent energy between the borehole station and the surface stations. This wavefront suggests that the virtual source is centered about a location quite a distance away from the lateral position of the borehole stations. This center happens to coincide with the location of an offshore platform, which is where the top of the borehole is located.

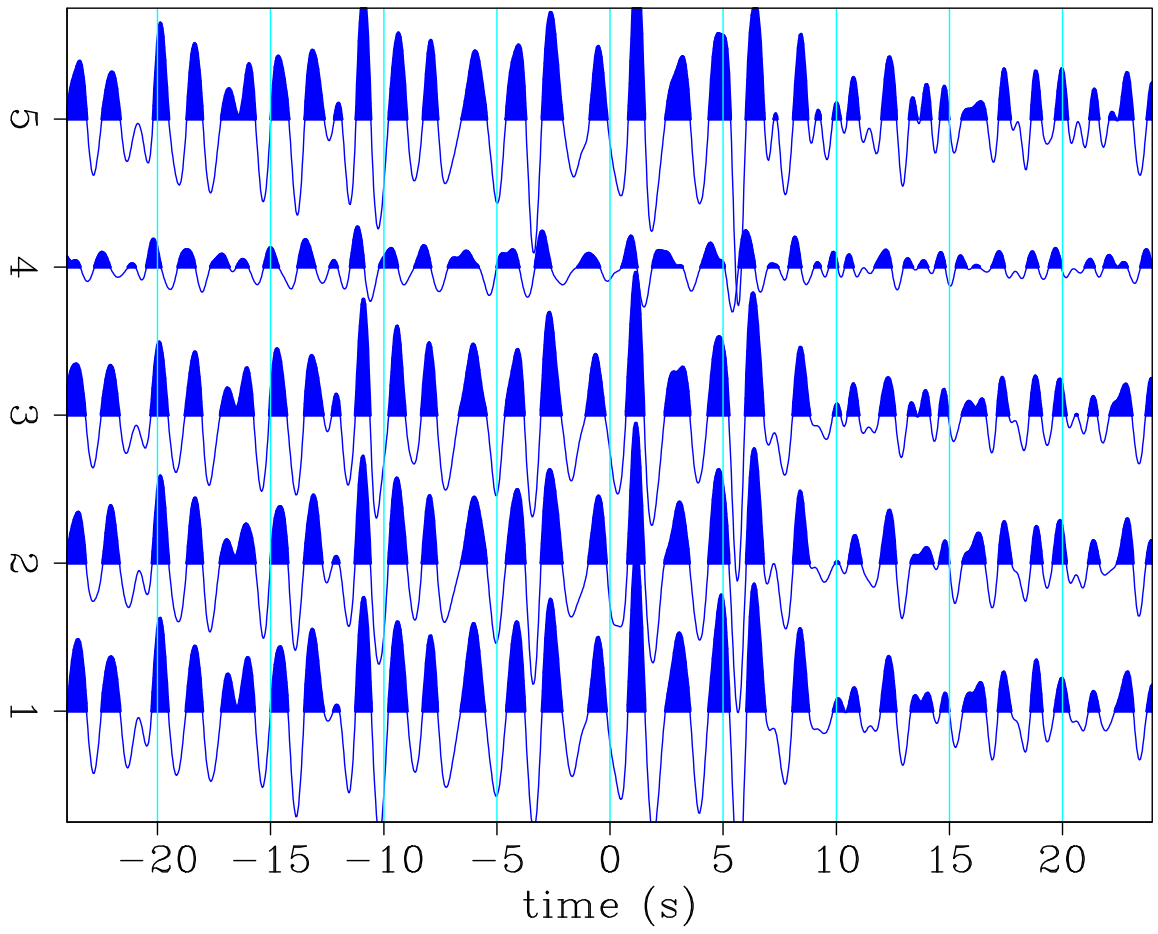


Figure 4: Correlation results of surface station number 1785 with the five boreholes. This surface station is located roughly 1.8 km north of the borehole stations. Trace 1 is the correlation result with the shallowest borehole, and trace 5 is the correlation result with the deepest borehole. Note the similarity of the shapes of the traces but the drastic decrease in amplitude of borehole station 4. [**CR**] jasonpc1/. wiggle

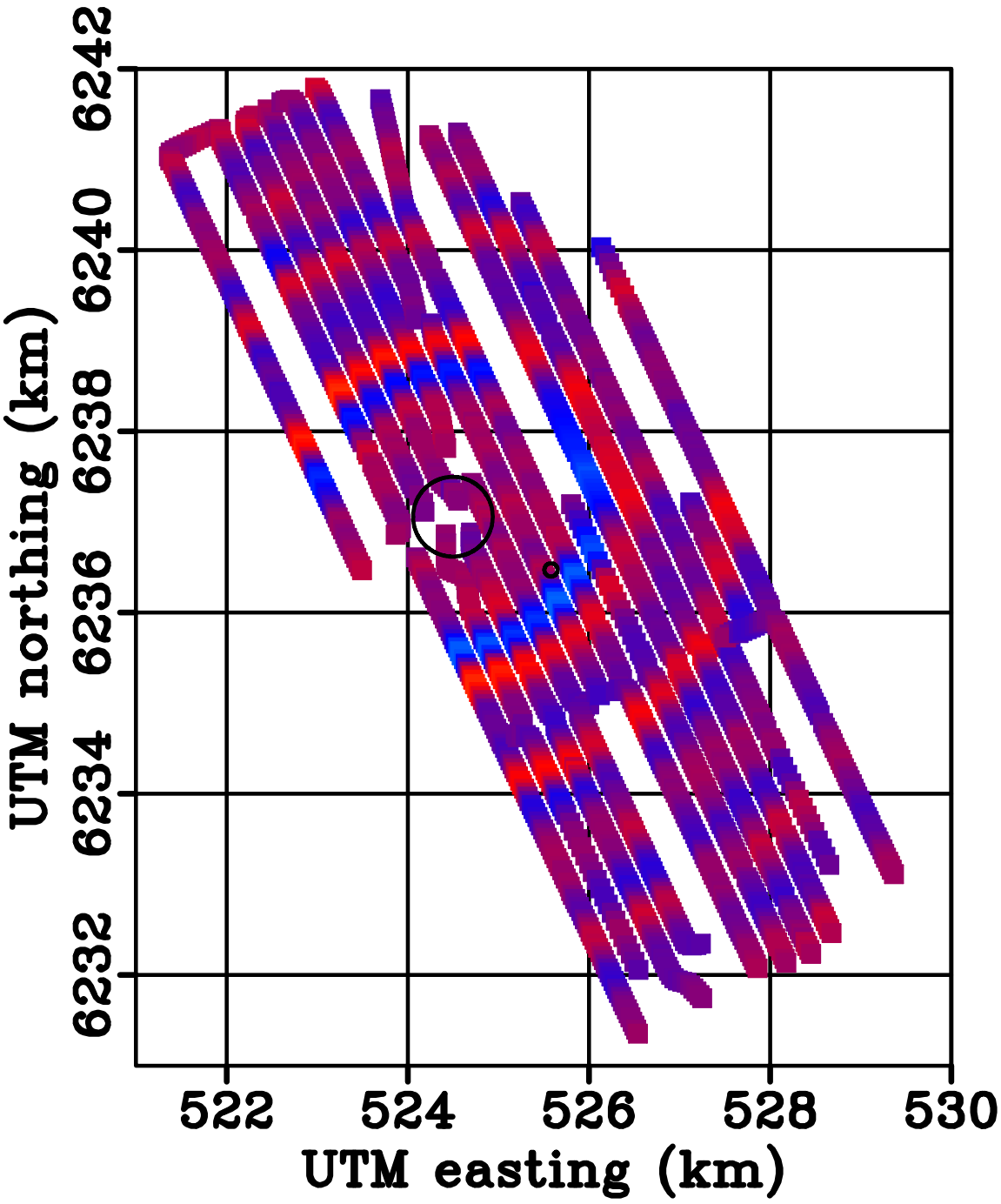


Figure 5: Snapshot of the wavefront propagating outward from the virtual source. Note how the virtual source location appears to be northwest of the borehole station location 525.5 km and 6236.5 km (smaller circle). This source location is near an offshore platform 524.5 km and 6237 km (larger circle). [**CR**] jasonpc1/. frame

DISCUSSION

There is definitely coherent correlated energy between borehole and surface stations. Although the results in Figure 4 appear noisy, the correlations between a single surface station and all five borehole stations are very similar to each other. However, a resemblance between the cross-correlation signal and the Green's function remains to be discussed. As previously mentioned, seismic interferometry says that, under certain conditions, the Green's function between two stations can be extracted by cross-correlating recordings of the ambient seismic field at both stations. One of the primary conditions is that the ambient seismic field satisfies energy equipartition, which means that energy propagates equally in all directions. This seems to be more or less satisfied for wave modes along the surface but does not usually seem to hold in 3D for body waves (Artman, 2007).

The seismic Green's function between borehole stations and the surface stations is unknown, especially at these low frequencies. The wave propagation between stations in the borehole and stations at the surface is obviously affected by the presence of the borehole, as it can act as a wave-guide. It is not inconceivable that the waves propagate from a virtual source in the borehole, first moving up to the surface along the borehole and then outward along the surface. This would explain why the apex of the event at causal times lies at the platform rather than at the surface coordinates of the borehole stations. However, this does not explain the non-physical nature of the apex not lying at $t = 0$ s but at approximately $t = -2$ s, nor why the signal we retrieve is not time-symmetric (not even around approximately $t = -2$ s).

The diagram in Figure 6 shows a possible mechanism for the coherent energy observed at low frequencies between the borehole and surface stations. The diagram depicts energy, excited by distant sources, in both directions along the surface. Energy that reaches the entrance of the borehole both excites the waves down to the borehole stations and continues to the surface stations. This would account for the apex of the (mostly) causal events lying at a small negative correlation lag. Energy that enters the borehole under the platform and travels down can lead to multiple bounces up and down the borehole, which would cause the repetitive signal in the acausal time window. This mechanism of seismic energy does not satisfy the conditions of seismic interferometry, thus we cannot interpret the cross-correlation signals as inter-station Green's functions.

CONCLUSIONS

We have demonstrated that there is correlated energy at frequencies between 0.175 Hz and 1.75 Hz between 2-km deep borehole stations and surface stations at the Valhall field. However the results are not typical of seismic interferometry. The signal is asymmetric, not centered at zero time lag, and does not originate from the borehole stations. We have argued that these observations are due to seismic energy interacting with the borehole acting a wave-guide. In this scenario, the asymmetry would be due to seismic energy travelling up and down the borehole. The apexes would be at acausal times because seismic energy needs to enter the borehole at the surface before reaching the stations at depth. A virtual source appears to emanate from the platform because correlated energy in the borehole must reach the surface at the platform before propagating outward. Because this proposed mechanism does not satisfy the conditions of seismic interferometry, we cannot interpret

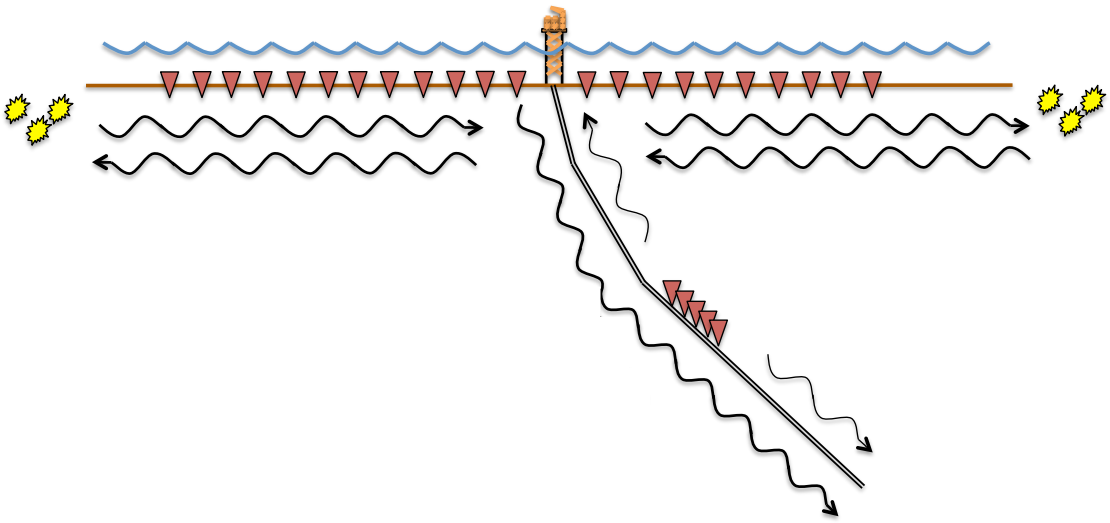


Figure 6: Illustration of the ambient seismic noise environment at Valhall. [**NR**] jasonpc1/. diagram

these cross-correlation results as inter-station Green's functions.

ACKNOWLEDGEMENTS

We thank to Olav Barkved for helpful discussions and suggestions. Thanks to Joe Dellinger for discussions on the mechanism of coherent energy between surface and borehole stations. We gratefully acknowledge to BP and the partners of the Valhall field (BP Norge A/S and Hess Norge A/S) for permission to publish this paper.

REFERENCES

Artman, B., 2007, Passive seismic imaging: PhD thesis, Stanford University.

Bussat, S. and S. Kugler, 2009, Recording noise - Estimating shear-wave velocities: Feasibility of offshore ambient-noise surface-wave tomography (ANSWT) on a reservoir scale: SEG Technical Program Expanded Abstracts, **28**, 1627–1631.

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

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.

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.

Hunting for microseismic reflections using multiplets

Noha S. Farghal and Stewart A. Levin

ABSTRACT

Microseismic monitoring of hydraulic fracturing is often used to locate reactivated faults and newly created fractures by locating the microseismic sources that occur during the process. These microseisms generate reflections as well as direct arrivals, but they tend to be fairly weak and quite difficult to locate and image. In this work we identify multiplets, i.e. repeated microseisms originating from about the same subsurface location, and thereby identify their consistent reflections.

INTRODUCTION

Hydraulic fracturing is routinely used to increase production rates from unconventional resources such as tight sands and shale plays. Its effectiveness depends largely on the geometry of the induced fractures. Microseismic monitoring is often used to assess this geometry by picking these earthquake-like arrivals and locating their sources along the fractures.

Microseismic recordings also contain reflected arrivals that may be useful for imaging the reservoir (Asanuma et al., 2011; Tamakawa et al., 2010; Reshetnikov et al., 2009). However, due to their small magnitudes, such reflections may not be easy to identify on a seismogram. In this work we key on the frequent presence of repeated events known as multiplets that originate from about the same subsurface location and have nearly identical source characteristics. Our assumption is that they must also generate nearly identical reflections. By extracting, aligning and stacking these events we reduce background noise and enhance both the direct arrivals and the weaker reflections that follow them.

Our aim for reflection identification is to use these high frequency deep sources to image the subsurface. Although such a use of microseismic sources has been done before in subsurface imaging (Asanuma et al., 2011; Tamakawa et al., 2010; Reshetnikov et al., 2009), to the best of our knowledge none of these sources were hydraulically induced.

METHOD

To identify multiplets in seismograms, we adapt a method from earthquake seismology which involves identifying clear direct arrivals (master waveforms), performing normalized cross-correlations of them with the microseismic recordings taken during the treatment, and thresholding the peak correlations to identify and group arrivals into multiplets (Brown et al., 2008; Eisner et al., 2008; Shelly et al., 2011).

We assessed the peak on the average correlations across all the 36 channels in each cross-correlation record. To accommodate small differences in source locations within each

multiplet, we tested two positive and two negative receiver linear moveouts and chose the moveout yielding the highest peak cross-correlation. We used the peak magnitude to decide whether the record was part of a multiplet and, if so, the peak location was used for subsequent alignment and stacking of the records in the multiplet.

APPLICATION

The dataset we used in this work was obtained from a monitoring well in the Bossier play (a shale and sand gas reservoir) located in the Dowdy Ranch Field, Texas. The data were collected before (background) and during hydraulic fracture treatments with an array of 12 three-component geophones in a monitoring well. The well was 500ft away from the treatment well (Sharma et al., 2008). The data used in this report are from hydraulic fracture treatment of the Bonner sand formation in this field.

A sonic log has been supplied to us with the data to aid interpretations. As can be seen on Figure 1, there are a couple of strong velocity contrasts that could produce moderately strong reflections.

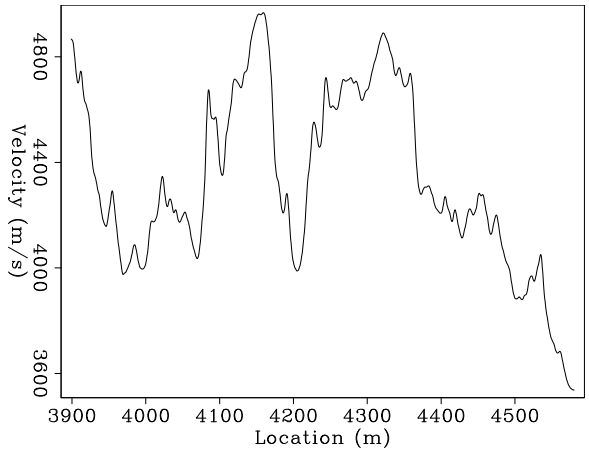


Figure 1: Velocity from a sonic log in the monitoring well. [**ER**] nfarghal1/. velnoha

Figure 2 shows a seismogram used for the generation of a master waveform. The waveform between 0.14 and 0.19 s (Figure 3) was extracted from the original seismogram. Normalized cross-correlations of this master waveform with other 0.5 s event windows in the dataset was performed. An example cross-correlation peak of the master with the event window from which it was obtained is shown in Figure 4. The stacked cross-correlation of the peak shown in Figure 4 can be seen in Figure 6(a).

The correlation of the master with another event window is shown in Figure 5. The correlation is stacked (Fig. 6(b)), and the peak location used as detailed below to align this record with that from which the master was extracted.

We used the magnitude of the background correlation of Figures 4 and 5 to set our threshold (minimum) cross-correlation value for multiplet detection. Events that had cross-correlation magnitudes equal to or exceeding this threshold were considered matches/multiplets and used for alignment and stacking. By alignment, we mean that matching waveforms are shifted so that, within the 0.5s event windows they are in, they occur at the same temporal location. This was achieved by applying a bulk positive or negative shift to each event window. The shift for each event window containing a multiplet was the difference in

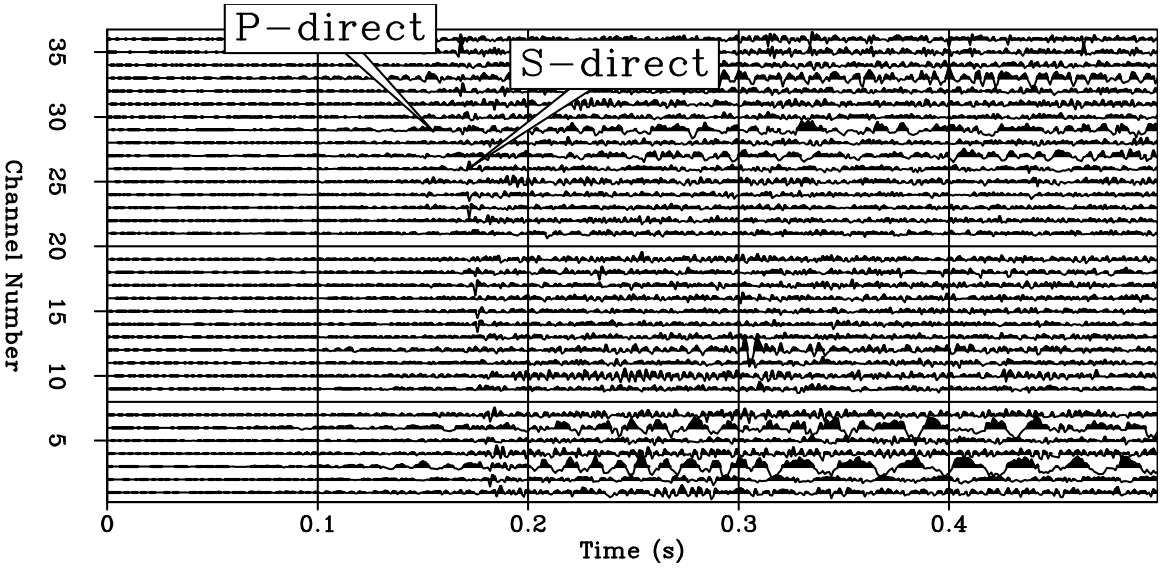


Figure 2: Seismogram used in the generation of the master waveform in Figure 3, showing both P and S arrivals. [**ER**] nfarghal1/. Bonner0040

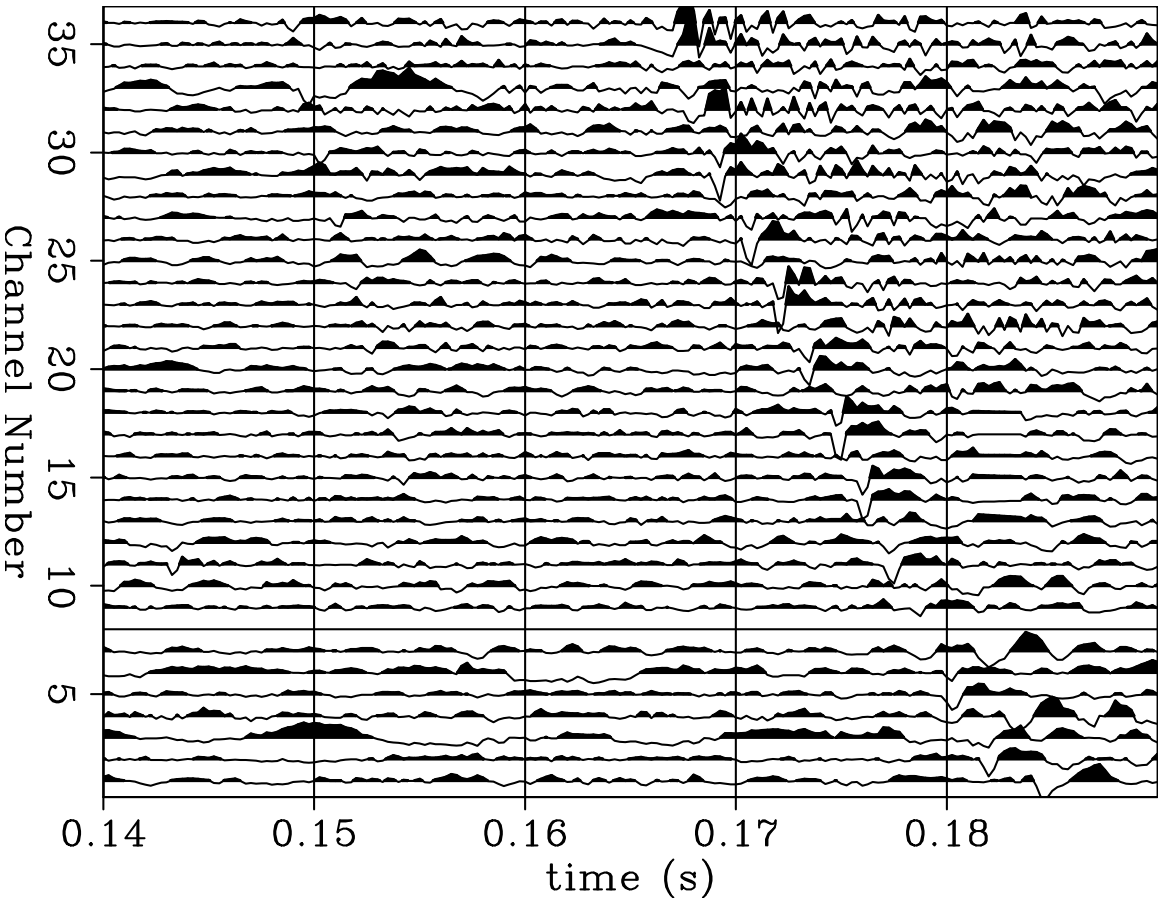


Figure 3: Master waveform extracted from Figure 2 by windowing in time around 0.14 and 0.19 s. [**ER**] nfarghal1/. B0040ma

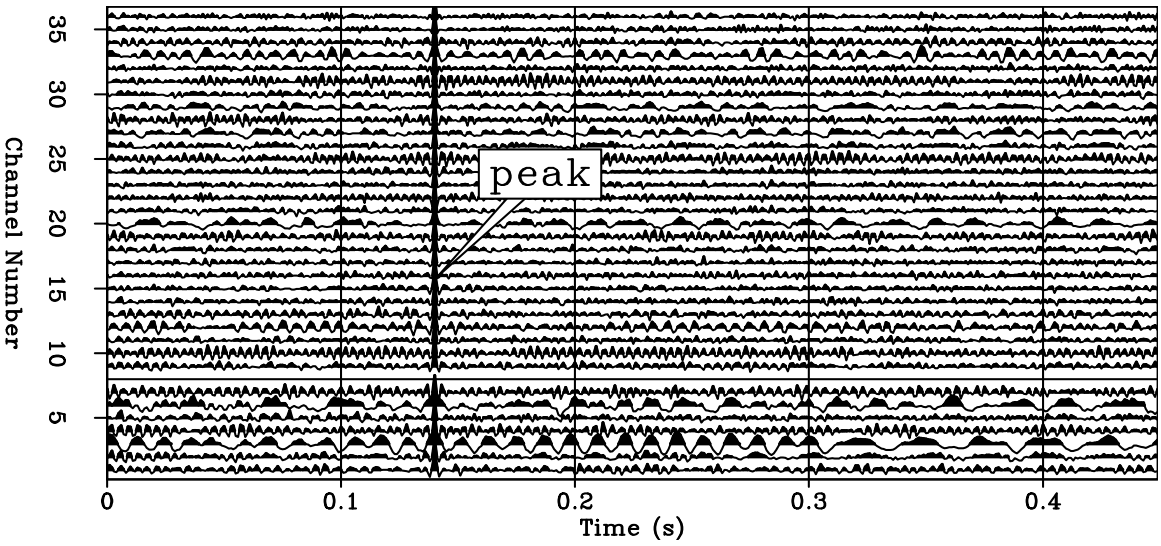


Figure 4: Cross-correlation of the master with its original event window (Figure 2).The correlation peak is marked [**ER**] nfarghal1/. B40tct40

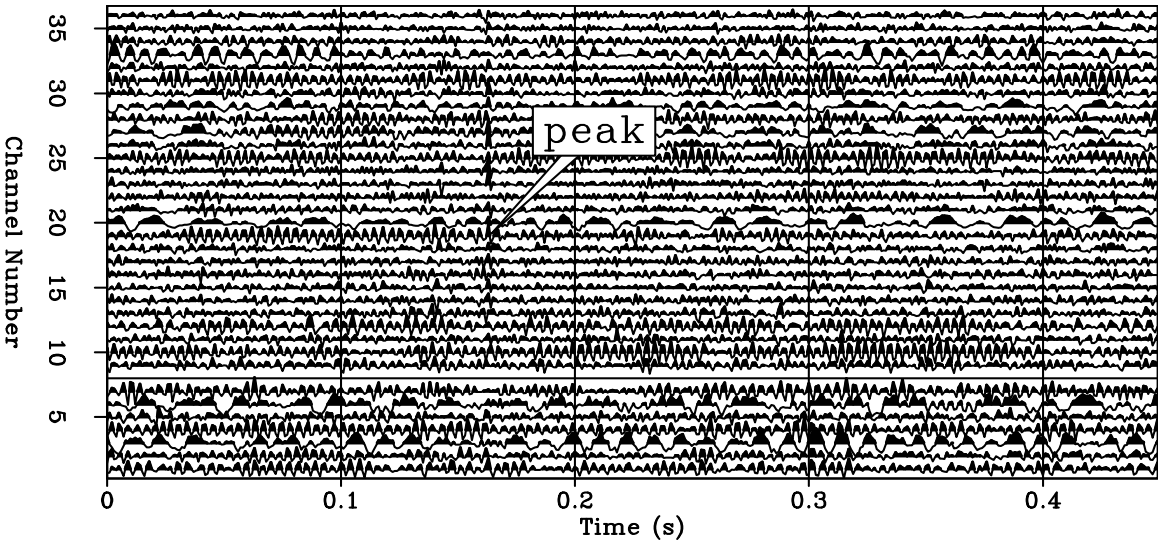


Figure 5: Cross-correlation of the master with another event window with the correlation peak marked [**ER**] nfarghal1/. B42tct40

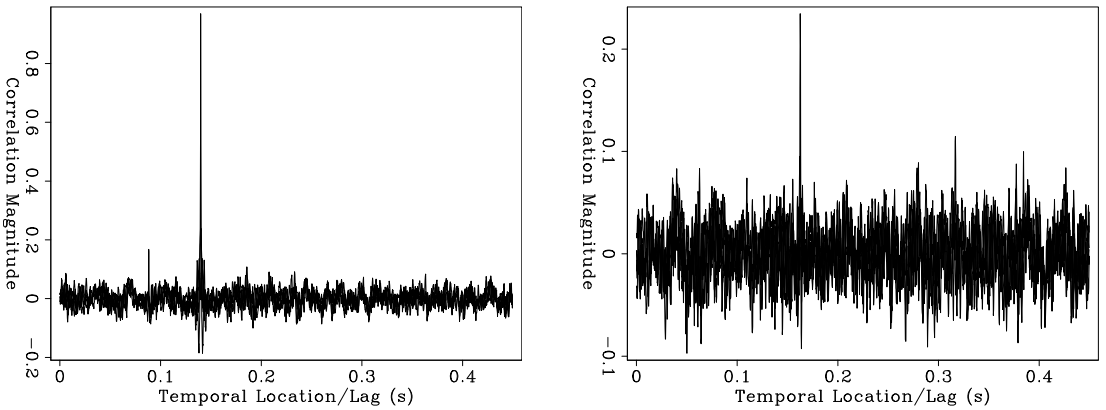


Figure 6: a) Stack of the correlation in Figure 4. The peak correlation value of 1 reflects the fact that the master window was extracted from the same record. b) Stack of the correlation in Figure 5. Here the peak correlation of 0.24 stands out above the background level of 0.05–0.1. [**ER**] nfarghal1/. B40tcsti40,B42tcsti40

temporal location of the master waveform location and the location of the cross-correlation peak rounded to the nearest sample. After alignment the events were summed together, as shown in Figure 7.

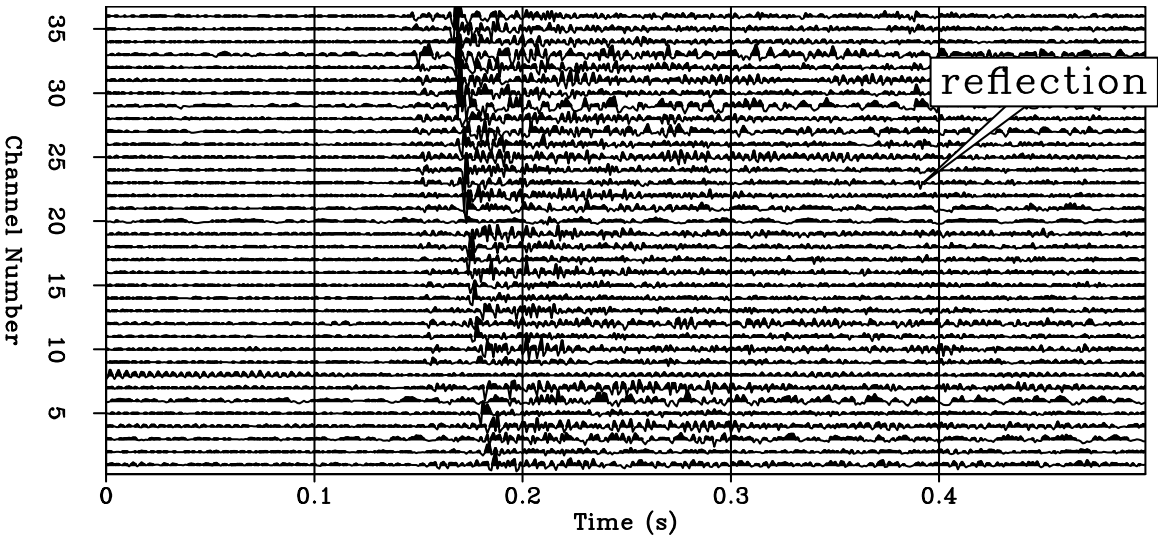


Figure 7: Stacked multiplets for the waveform in Figure 2 with the S reflection marked. [**ER**] nfarghal1/. stack3

Figure 7 shows stacked traces for the master taken from the seismogram shown in Figure 2. It is clear on the stacked event that the main waveform is followed by another wavelet around 0.39s that was not apparent in the original file from which the master waveform was taken. The event did show up almost consistently at the same temporal offset in most events in the multiplet used to generate this stacked event, as shown in Figures 8–11.

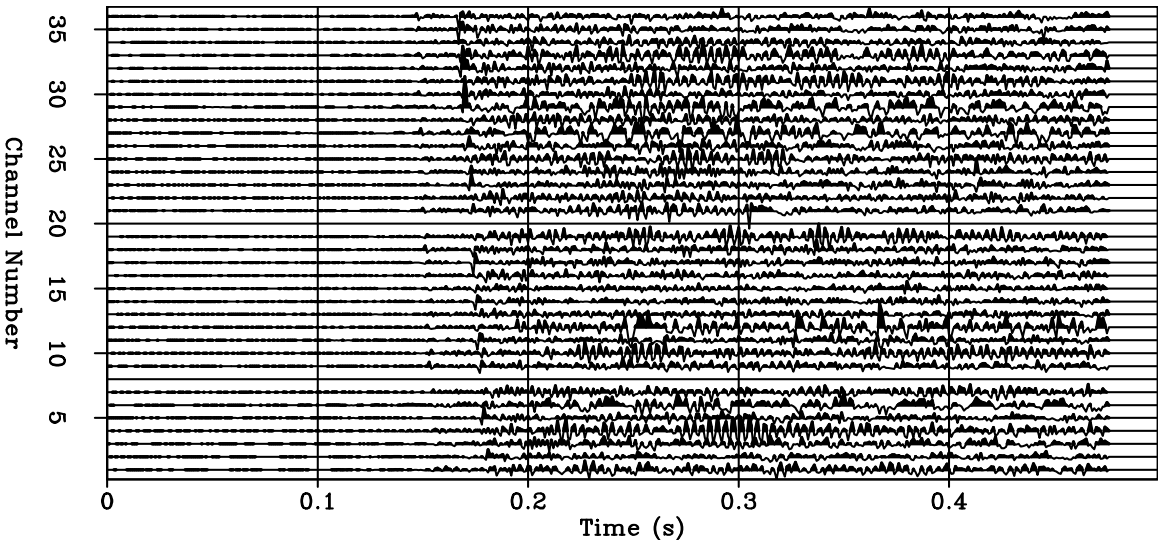


Figure 8: One of the seismograms used in making the stack in Figure 7. [**ER**] nfarghal1/. Bonner0455s

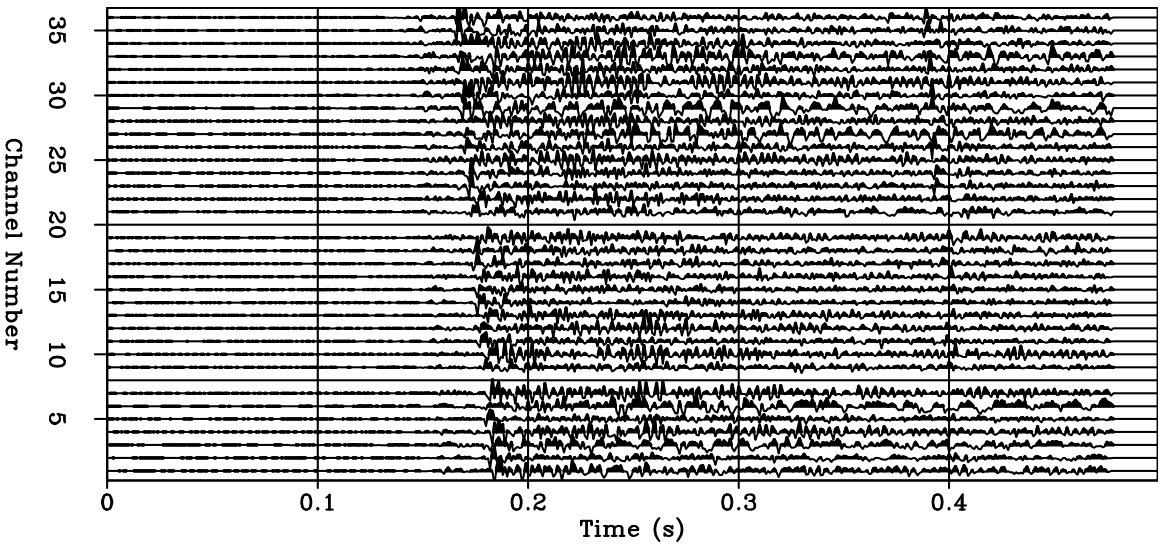


Figure 9: One of the seismograms used in making the stack in Figure 7 showing the S reflection. [**ER**] nfarghal1/. Bonner0152s

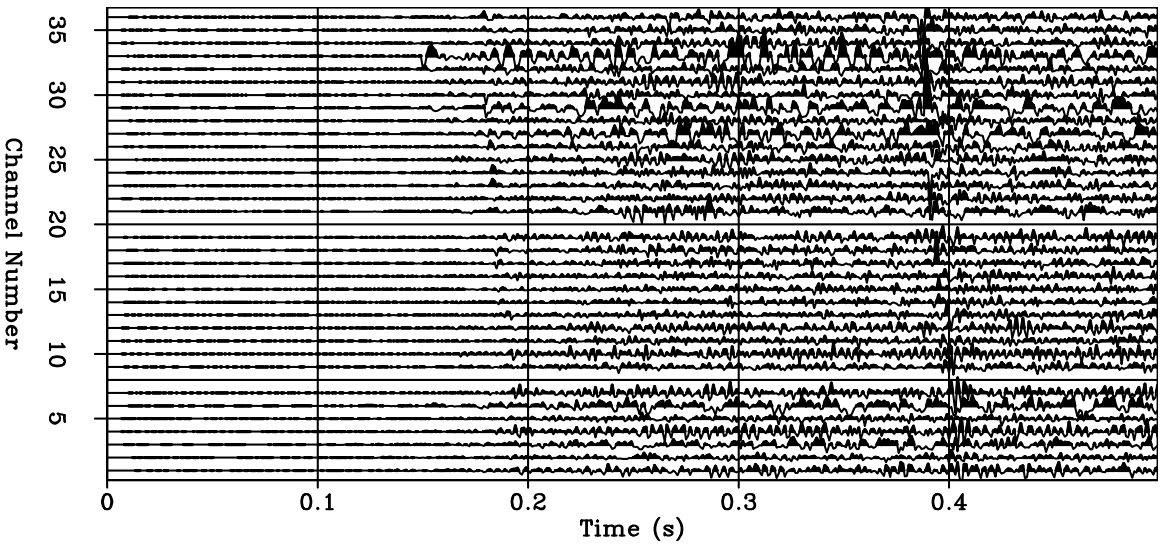


Figure 10: One of the seismograms used in making the stack in Figure 7 showing the S reflection. [**ER**] nfarghal1/. Bonner0209s

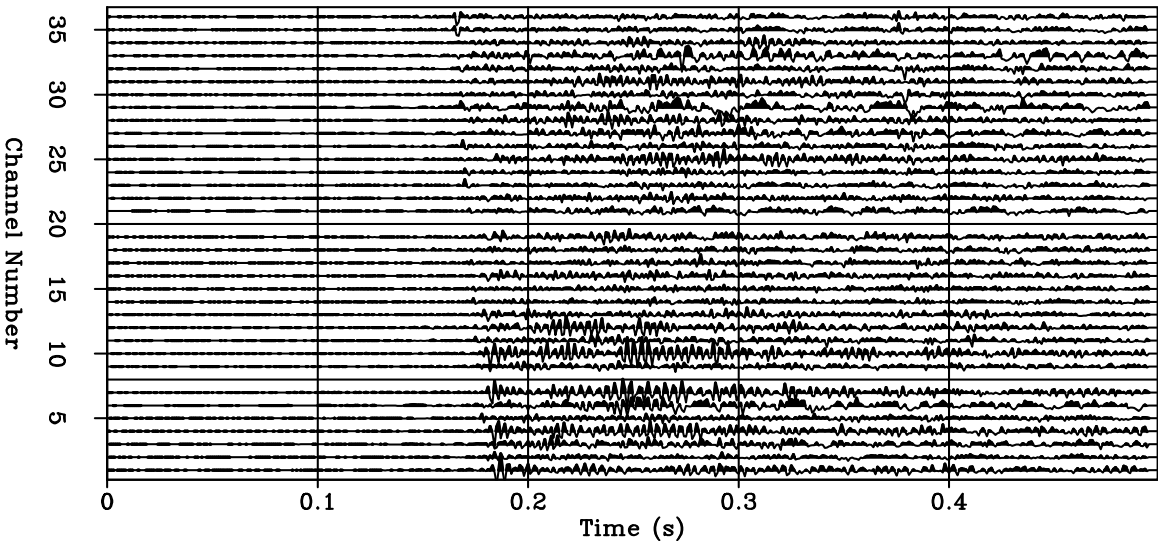


Figure 11: One of the seismograms used in making the stack in Figure 7 showing the S reflection. [**ER**] nfarghal1/. Bonner0480s

We performed singular value polarization analysis for this stack of events (see the appendix for details) and used this information to emphasize reflection type (S or P). We projected the stacked events using this direction vector to boost P arrivals and possibly see otherwise invisible P reflections. The result of this product is displayed in Figure 12. Comparing this figure with figure 7, we can see that this P-projection has indeed supressed the shear reflection at 0.39s. The S direct arrival was decreased in RMS by a factor of 1.5 and in maximum amplitude by a factor of 3. This projection also increased the P direct arrival RMS by a factor of 1.4.

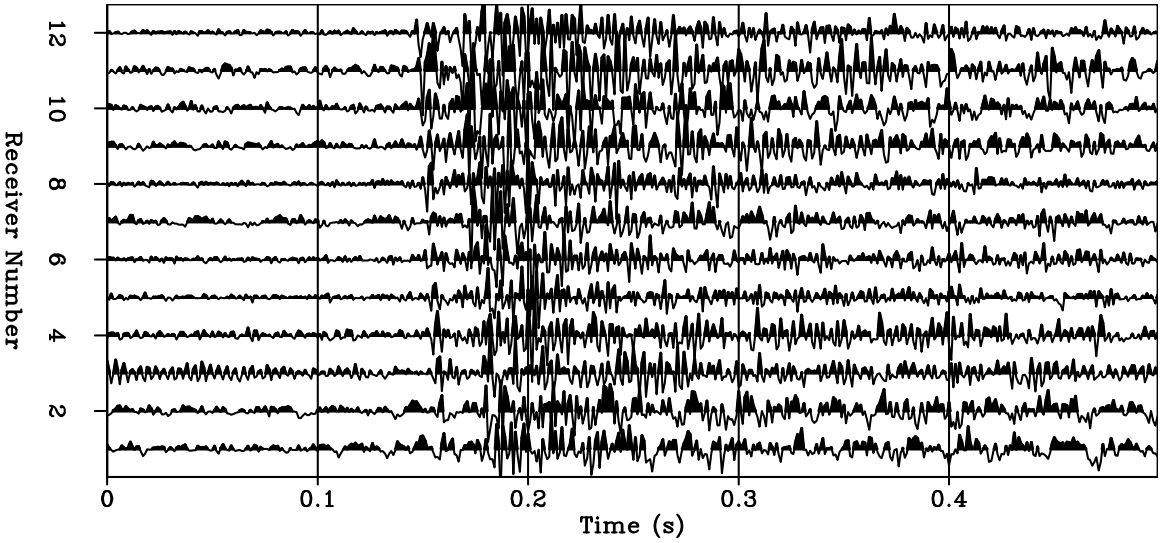


Figure 12: Stack in Figure 7 projected onto the estimated P-direction. When compared to Figure 7, reduction in S amplitude and P amplitude enhancement can be noticed. [**ER**] nfarghal1/. pboost3

As another example, Figures 13–19 exhibit a reflection at about 0.43s for a different master.

DISCUSSION

We identified clear multiplets which, disappointingly, when stacked did not produce reflection events as strong on the stack as they appeared on specific individual event windows. Using the velocity log in Figure 1, a back of the envelope calculation showed three deficiencies in our algorithm:

- Our bulk shifts to the nearest sample were too coarse for the majority of arrivals. We should be using fractional sample shifts.
- Our linear moveout corrections were similarly too coarse.
- The misalignment of later reflections due to small differences in the microseismic source locations is generally larger by a factor of 2 to 4 greater than that of the direct arrivals due to the longer travelpath of reflections. So we need to include time-variant alignment, i.e. warping.

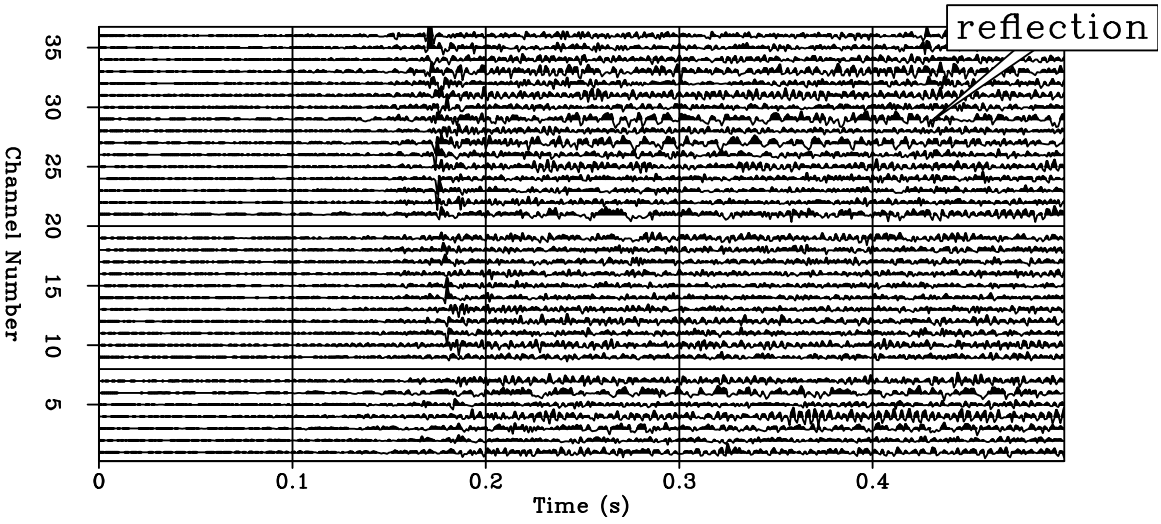


Figure 13: The stack of multiplets shown in Figure 14–19 with the S reflection marked. [**ER**] nfarghal1/. Stack0417

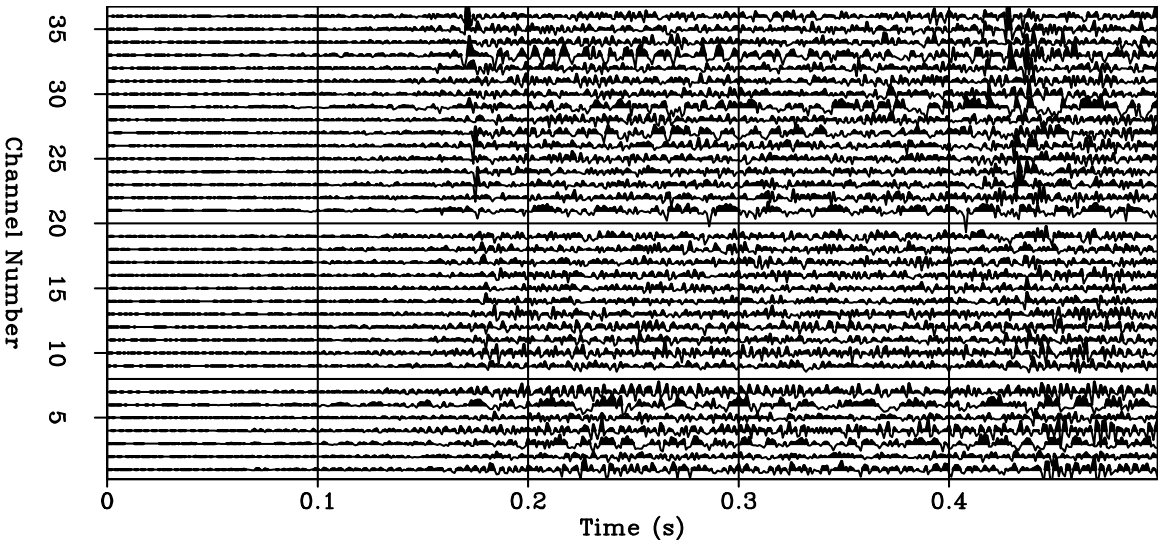


Figure 14: One of the seismograms used in making the stack in Figure 13. [**ER**] nfarghal1/. Bonner0130

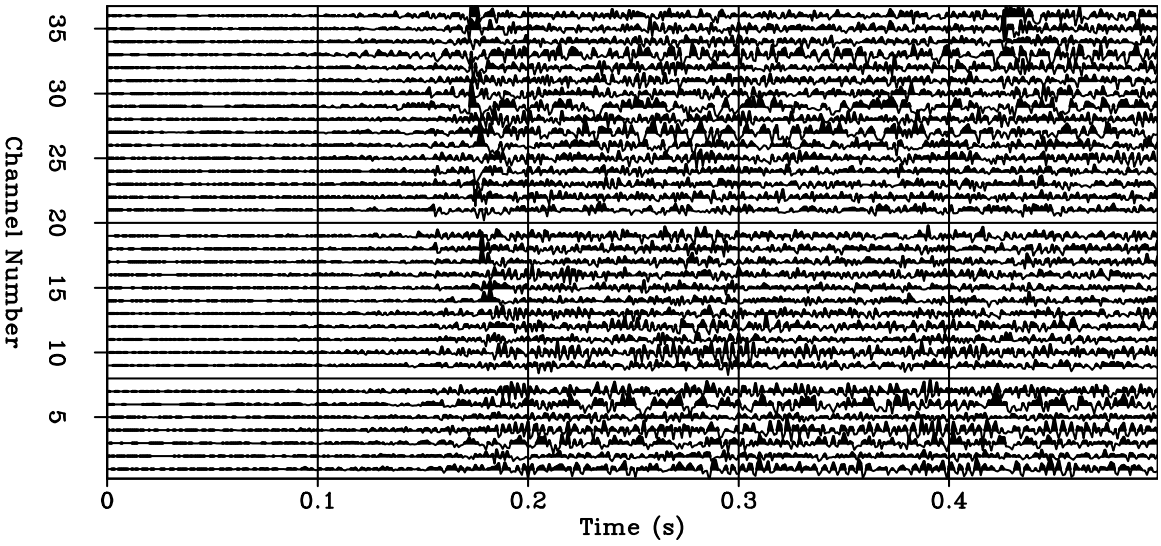


Figure 15: One of the seismograms used in making the stack in Figure 13. [**ER**] nfarghal1/. Bonner0108

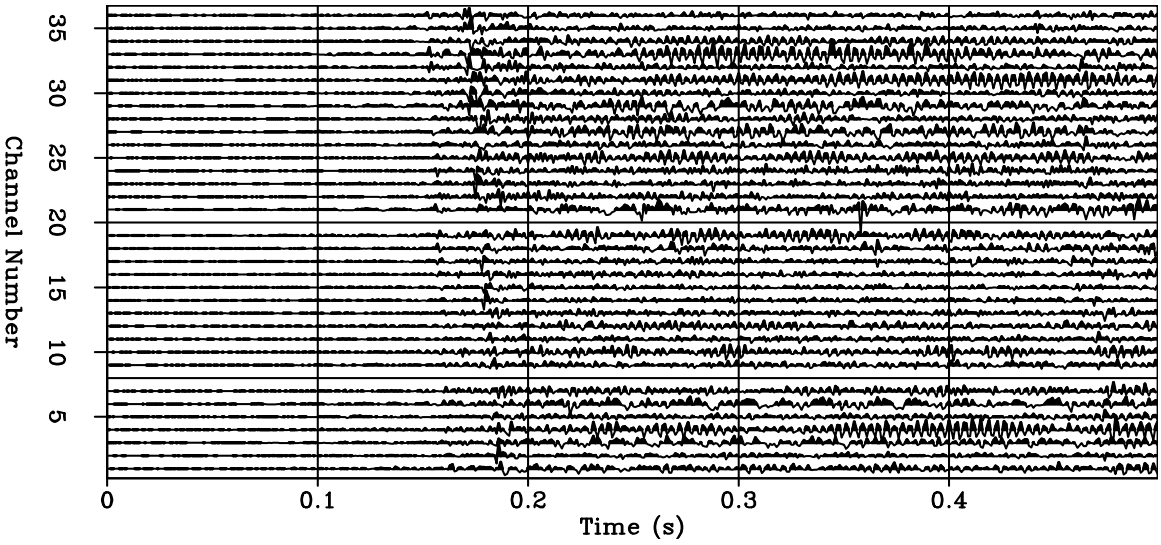


Figure 16: One of the seismograms used in making the stack in Figure 13. [**ER**] nfarghal1/. Bonner0097

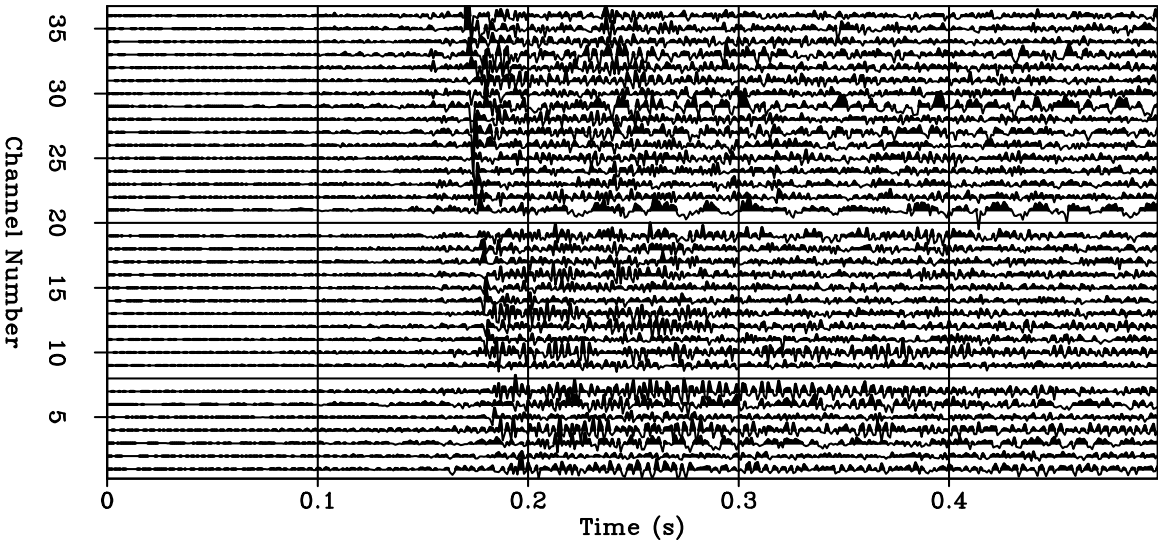


Figure 17: One of the seismograms used in making the stack in Figure 13. [**ER**] nfarghal1/. Bonner0131

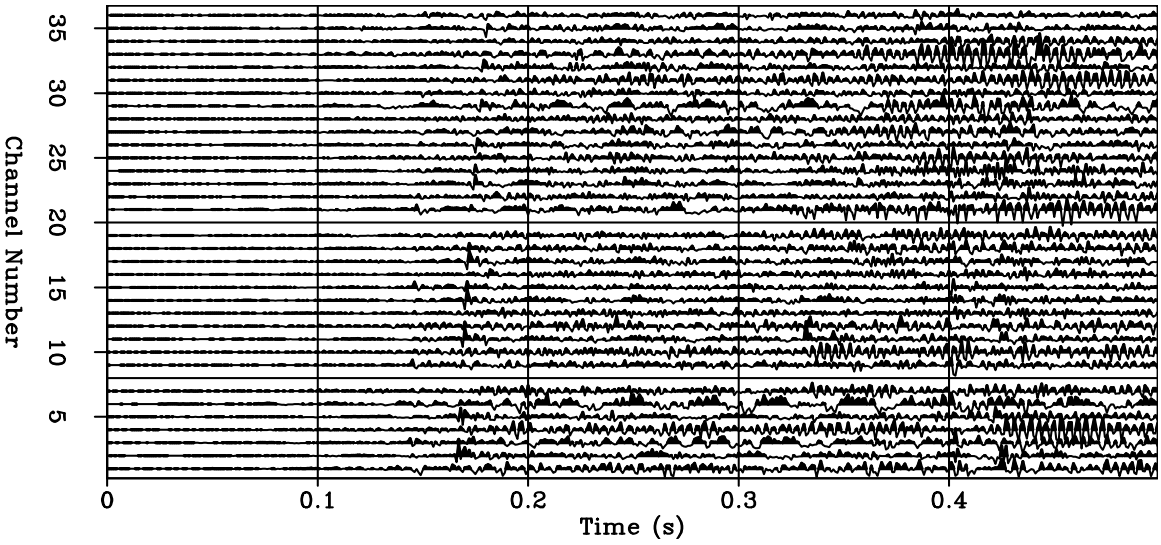


Figure 18: One of the seismograms used in making the stack in Figure 13. [**ER**] nfarghal1/. Bonner0962

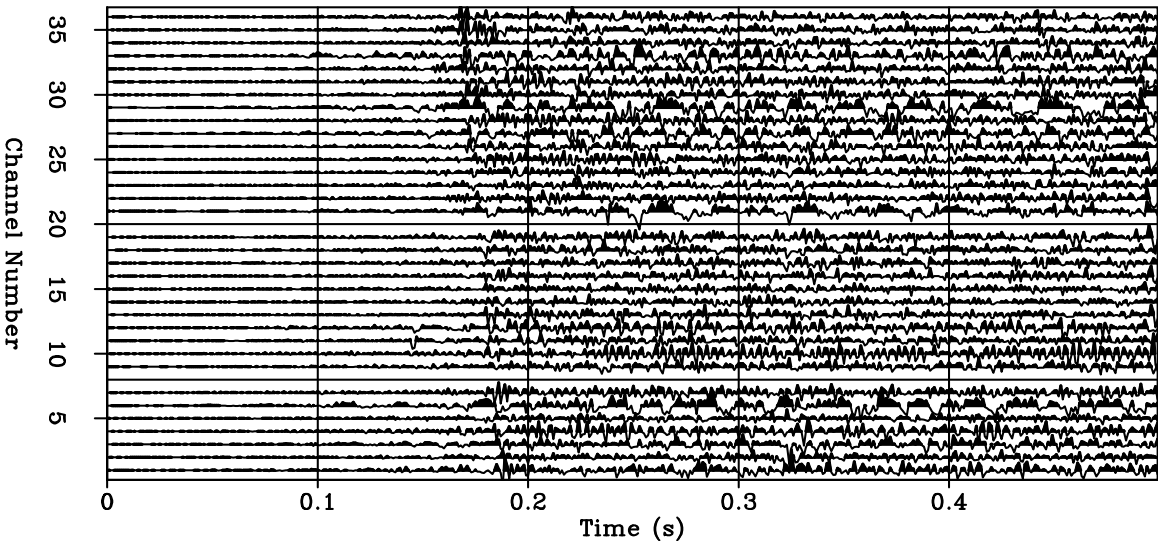


Figure 19: One of the seismograms used in making the stack in Figure 13. [**ER**] nfarghal1/. Bonner4211

In addition to these effects, we remark that the RMS value of the noise in the stack and the direct P-arrival have almost the same value, an additional reason why P reflections were not clearly revealed after projection even though the direct P first arrival was enhanced.

On a positive note, we can use the misalignments in reflected arrival times to constrain differences in source locations, supplementing double-difference hypocenter analysis for the microseism locations (Waldhauser and Ellsworth, 2000).

CONCLUSIONS AND FUTURE WORK

We adapted a methodology from earthquake seismology literature to identify and stack multiplets of nearly identical arrivals which successfully grouped nearly identical reflection events in the microseismic data.

However, the stacked results have been far from perfect. Stacked reflections appear to be weaker than reflections in individual event windows. This has been attributed to the fact that a trace variable shift needs to be applied to the data before stacking that is different when warping the reflection arrival from the direct arrival. Moreover, fractional rather than integer shifts need to be introduced. When such enhancement in shift precision is introduced, we hope that event windows will better stack together, enhancing the S/N ratio and boosting reflections, especially the weak P ones.

ACKNOWLEDGMENTS

We thank Pinnacle, a Halliburton company, and the United States Department of Energy for providing the dataset.

APPENDIX A

Singular value polarization analysis

Since S first arrivals are clearer and stronger than P first arrivals, we use S arrivals to estimate a vector in the general direction of P arrivals. The idea is to find the direction most perpendicular to the strongest (direct) S arrivals. We setup a $3 \times n$ matrix of shear first arrivals, $\mathbf{W}$, by windowing around the first shear arrival. As we seek a vector v that is perpendicular to the shear arrival direction (which is a P-arrival) we want

$$\mathbf{W}\mathbf{v} \approx \mathbf{0} \quad . \tag{A-1}$$

We solve this by minimizing the objective function

$$J = \| \mathbf{v}^\mathbf{T}\mathbf{W}^\mathbf{T}\mathbf{W}\mathbf{v} \|_\mathbf{2} \tag{A-2}$$

subject to the constraint

$$\mathbf{v}^\mathbf{T}\mathbf{v} = \mathbf{1} \quad . \tag{A-3}$$

Let α and β be two spherical surface coordinate parameters over which we will minimize. Taking partial derivatives of the constraints yields:

$$\mathbf{v}^\mathbf{T}\frac{\partial \mathbf{v}}{\partial \alpha} = \mathbf{0} \tag{A-4}$$

and

$$\mathbf{v}^\mathbf{T}\frac{\partial \mathbf{v}}{\partial \beta} = \mathbf{0} \quad , \tag{A-5}$$

which says that $\mathbf{v}^\mathbf{T}$ is perpendicular to the two partial derivatives. Next, taking partial derivatives of the sum of squares expression gives

$$\mathbf{v}^\mathbf{T}\mathbf{W}^\mathbf{T}\mathbf{W}\frac{\partial \mathbf{v}}{\partial \alpha} = \mathbf{0} \tag{A-6}$$

and

$$\mathbf{v}^\mathbf{T}\mathbf{W}^\mathbf{T}\mathbf{W}\frac{\partial \mathbf{v}}{\partial \beta} = \mathbf{0} \quad . \tag{A-7}$$

Therefore $\mathbf{v}^\mathbf{T}\mathbf{W}^\mathbf{T}\mathbf{W}$ is also perpendicular to both partial derivatives and consequently must be parallel to $\mathbf{v}^\mathbf{T}$. This means that

$$\mathbf{v}^\mathbf{T}\mathbf{W}^\mathbf{T}\mathbf{W} = \lambda \mathbf{v}^\mathbf{T} \quad , \tag{A-8}$$

where λ is the eigenvalue of the matrix $\mathbf{W}^T\mathbf{W}$ that will make the least squares expression a minimum. Transposing we get

$$\mathbf{W}^T\mathbf{W}\mathbf{v} = \lambda \mathbf{v} \quad , \tag{A-9}$$

which is a classic eigenvector problem for the matrix $\mathbf{W}^T\mathbf{W}$. Since the right singular vectors of $\mathbf{W}$ are the same as the eigenvectors of $\mathbf{W}^T\mathbf{W}$, we used the LAPACK routine *SGESVD* to find our desired P-wave direction vector.

REFERENCES

Asanuma, H., K. Tamakwa, H. Niitsuma, N. Soma, J. Rutledge, and C. Rowe, 2011, Reflection imaging of the Aneth CCS reservoir using microseismic multiplet sources: SEG Technical Program Expanded Abstracts, **30**, 1478–1482.

Brown, J. R., G. C. Beroza, and D. R. Shelly, 2008, An autocorrelation method to detect low frequency earthquakes within tremor: Geophysical Research Letters, **35**, L16305.

Eisner, L., D. Abbott, W. B. Barker, J. Lakings, and M. P. Thornton, 2008, Noise suppression for detection and location of microseismic events using a matched filter: SEG Technical Program Expanded Abstracts, **27**, 1431–1435.

Reshetnikov, A., S. Buske, and S. A. Shapiro, 2009, Active seismic imaging using microseismic events: SEG Technical Program Expanded Abstracts, **28**, 1668–1672.

Sharma, M. M., P. B. Gadde, R. Sullivan, R. Sigal, R. Fielder, D. Copeland, L. Griffin, and L. Weijers, 2008, Slick water and hybrid fracs in the Bossier: Some lessons learnt: SPE Annual Technical Conference and Exhibition, Houston, TX, USA, 89876–MS.

Shelly, D. R., Z. Peng, D. P. Hill, and C. Aiken, 2011, Triggered creep as a possible mechanism for delayed dynamic triggering of tremor and earthquakes: Nature Geoscience Letters, **4**, 384–388.

Tamakawa, K., H. Asanuma, H. Niitsuma, and N. Soma, 2010, Reflection imaging using microseismic multiplets as a source: SEG Technical Program Expanded Abstracts, **29**, 2166–2170.

Waldhauser, F. and W. L. Ellsworth, 2000, A double-difference earthquake location algorithm: Method and application to the Northern Hayward Fault, California: Bulletin of the Seismological Society of America, **90**, 1353–1368.

Joint imaging with streamer and ocean bottom data

Mandy Wong, Shuki Ronen, and Biondo Biondi

ABSTRACT

In the past we have shown that up-going and down-going (mirror) imaging can be combined in a joint inversion. We extend the method to joint-inversion of nodes and streamers data. Compared to conventional post-imaging merging, the joint inversion enhances resolution, suppresses migration artifacts, and more importantly, brings up the relative amplitude of true reflectors in the subsurface. We present a linearized inversion scheme for imaging narrow-azimuth and ocean-bottom data. We demonstrate the concept and methodology in 2D with a synthetic Marmousi model.

INTRODUCTION

Recent advances in seismic acquisition and imaging, and improved understanding of the relationships between seismic data and rock-fluid properties have contributed to the successful applications of time-lapse seismic to hydrocarbon reservoir monitoring. However, a complex network of in-sea and sub-sea installations often poses restrictions on seismic vessels near the production facility. Operators wishing to monitor production at these sites are faced with operational difficulties when using the conventional towed-streamer method. The negative impact of such restrictions has been mitigated with undershooting, which utilizes independent shot and streamer vessels to navigate around obstacles. However, even with undershooting, acquisition repeatability is difficult. For example, there are still missing short-offset illuminations, and azimuth distribution can be inconsistent between monitoring surveys.

Boelle et al. (2010) have proposed the use of autonomous ocean-bottom nodes (OBN) as a solution for imaging around obstacles. Figure 1 compares the source and receiver geometry between towed-streamer and OBN surveys along the crossline direction of the streamer survey. Autonomous receivers can be planted close to the production facility for acquiring infill time-lapse seismic data to complement streamer data and give a full picture of the target zone. Although it has the advantage of illuminating around the production facility, the OBN method has its own unique processing challenges.

One challenge is obtaining good up-down-separated data from a four-component recording of OBN data. In an ocean-bottom survey, the signal can be classified into an up-going primary and down-going mirror signal (Figure 2). Dash et al. (2009) have used the down-going mirror signal in the migration of OBS data, which yields a wider illumination for sparsely spaced nodes. A reliable migration image requires the hydrophone signal to be decomposed into up- and down-going waves using the geophone and hydrophone recordings. In the case of imperfect up-down separation, there will be residual up-going energy or shear noise in the down-going mirror data. Such energy can give spurious artifacts in the

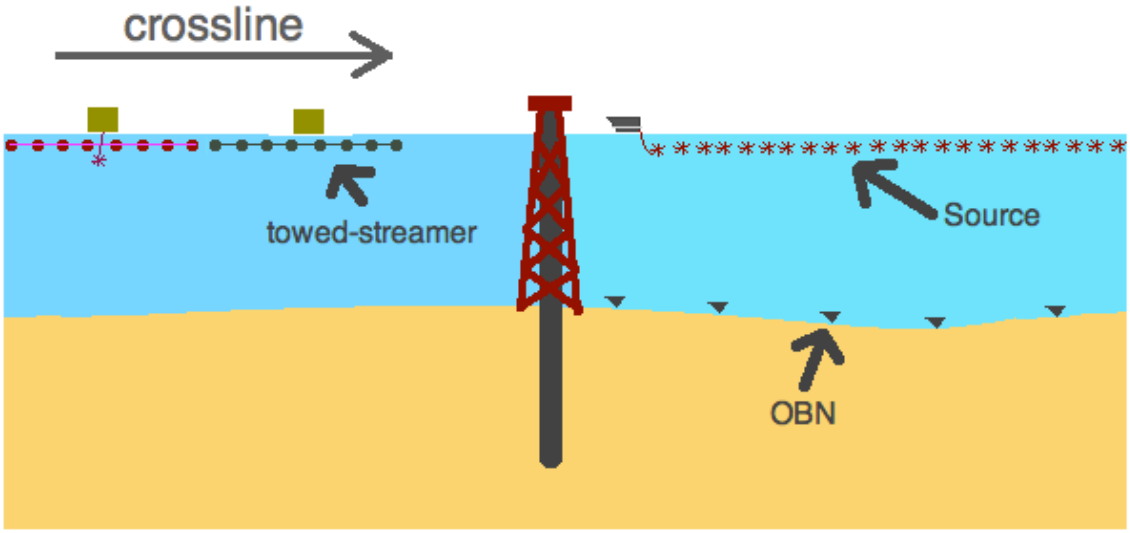


Figure 1: The source and receiver geometry of a towed streamer survey and an OBN survey along the crossline direction. [**NR**] mandy1/. geometry

migration image. The OBN image cannot completely replace the streamer image; instead, it is used to complement the streamer image.

Instead of treating the streamer and the OBN images separately, we propose an iterative linear least-squares inversion scheme that could coherently combine the information from the two images. The objective of the data processing is to put everything together and not overload the interpreter with alternative images. Combined ocean bottom nodes and streamer surveys offer the best of both methods: the economy of streamers and the ability of nodes to cover obstructed areas. Such an inversion can improve the structure and aperture of the seismic images by using two sets of signals. In the following sections, we will first discuss the theory of the joint linearized least-squares inversion, and then we show the results of applying our inversion scheme to a 2D Marmousi model.

THEORY

Linearized Inversion

We pose the imaging problem as an inversion problem by linearizing the wave-equation with respect to our model ($m(\mathbf{x})$). Assuming that the earth behaves as a constant-density acoustic isotropic medium, we linearize the wave equation and apply the first-order Born approximation to get the following forward modeling equation:

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

where d^{mod} represents the forward modeled data, ω is the temporal frequency, $m(\mathbf{x})$ represents the reflectivity at image point $\mathbf{x}$, $f_s(\omega)$ is the source waveform, and $G(\mathbf{x}_s, \mathbf{x})$ is the

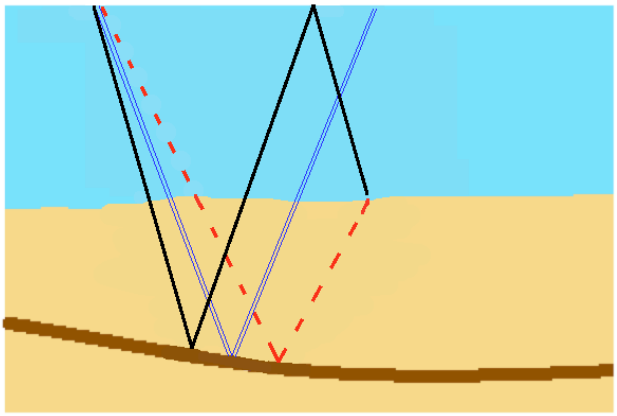


Figure 2: Three main modes of imaging. Up-going ocean bottom data in dashed-red, mirror ocean-bottom data in solid-black, and streamer data in double-line blue. Notice the subsurface reflection point is different for each mode given the same offset distance between the source and the receiver. [**NR**] mandy1/. illum

Green's function of the two-way acoustic constant-density wave equation. Note that G is actually ω dependent. It is important to point out that the adjoint of the forward modeling operator is the migration operator:

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

The inversion problem is defined by minimizing the least-squares difference between the synthetic and the recorded data: One can use various types of propagators to formulate the Green's function. In our study, we use the two-way propagator. In this case, the migration operator is equivalent to reverse time migration (RTM). The inverted image ($m(\mathbf{x})$) is better than the migration image in the sense that its forward-modeled data fits the recorded data. Next, we discuss how to apply linearized inversion to jointly image with streamer and OBN data.

Joint imaging

Linearized inversion can combine different sets of data that share the same model. In Wong et al. (2010), we show that up-going and down-going (mirror) imaging can be combined in a joint inversion. We now extend the method to joint-inversion of nodes and streamers data. The fitting goal is:

$$0 \approx \begin{bmatrix} \mathbf{L}_{str} \\ \mathbf{L}_{OBN\downarrow} \end{bmatrix} \mathbf{m} - \begin{bmatrix} \mathbf{d}_{str} \\ \mathbf{d}_{OBN\downarrow} \end{bmatrix}, \tag{3}$$

where $\mathbf{L}$ is the forward-modeling operator that corresponds to equation 1. The subscripts $_{str}$ and $_{\downarrow}$ denote streamer and OBN-mirror, respectively. Our goal is to obtain a final image that consistently explains the two data sets.

The major benefit of joint inversion is that it attenuates the migration artifacts caused by spurious energy in the original recorded data. Spurious energy can occur due to imperfect pre-processing. For example, even with the most advanced multiple-removal methods, surface and internal multiples energy can still be present in the field data. For OBN, there is an additional challenge of removing the shear and up-going energy in the mirror signal. Next, we will look at a 2D synthetic study based on the Marmousi model.

SYNTHETIC EXAMPLE

Figures 4 (a) and 4 (b) shows the velocity and reflectivity models used. The model is 2.8 km deep and 8 km wide with a spacing of 10 m. To simulate a monitoring survey, we assume there is an existing production structure that spans the location from x=3950-4000 m. The water depth is about 1000 m. Figure 3 shows a plane map view of the streamer survey and the OBN survey.

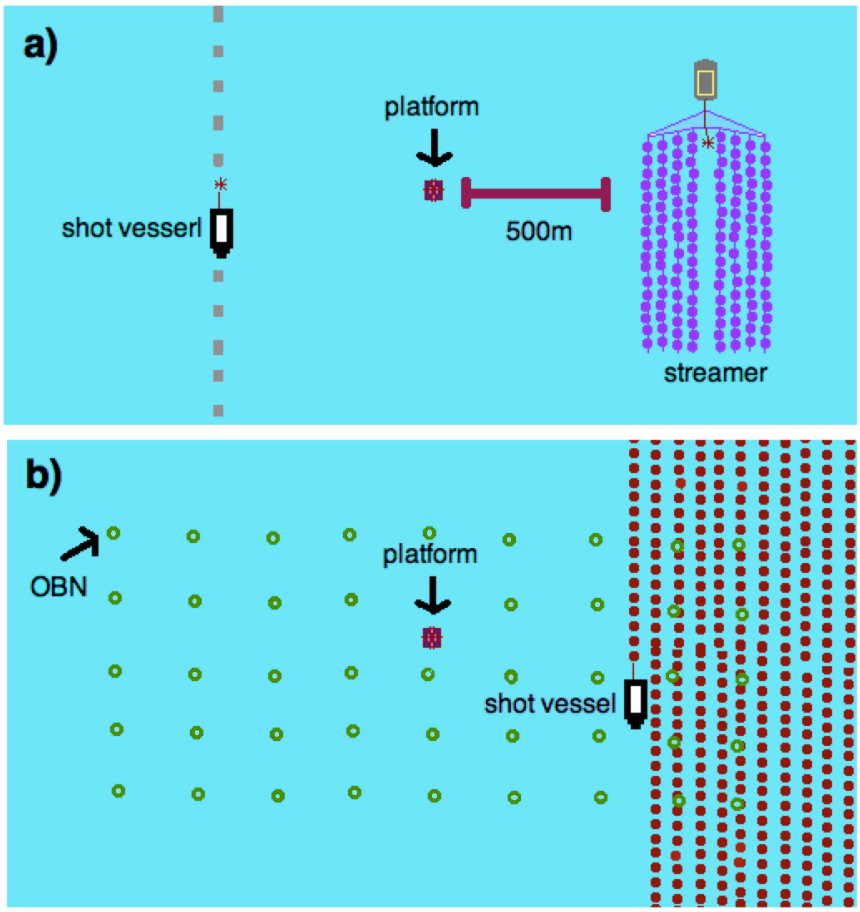


Figure 3: (a) Plane map view of streamer survey during undershooting. Streamer must be at least 500m from the platform. (b) Plane map view of OBN survey. The shot carpet (red circles) is much denser than the receiver grid (green open cicles). [**NR**] mandy1/. plainview

Figure 3 a shows a snap shot of the streamer survey during undershooting. In our 2D synthetic study, we create synthetic data along with shots and receiveres lined along the

crossline direction, cutting through the production platform. Shots run from 0 to 8000 m on the sea surface at an interval of 350 m along the 2D line. There are 8 receivers to the left and to the right of each shot, spaced at 50 m apart. This geometry captures the same sub-surface illumination range as a dual-shot towed-streamer survey along the crossline direction. The towed-streamer cannot be within 500 m from the platform, which leads to illumination gap. We create two additional shots undershooting near the platform.

For the OBN survey, receivers are placed from x=2000 to 6300 m at the ocean-bottom with a 400 m spacing. The shot carpet spans the entire model along the sea surface except in the area within 50m of the platform. For the OBN data, we forward-modeled the down-going mirror signal only. We added 10 percents of the up-going signal as noise to the data to simulate imperfect up-down separation.

Figure 4 (c) and 4 (d) show the reverse-time migration (RTM) image of the streamer survey and the OBN survey, respectively. By comparing with the original model, we can see different kinds of artifacts in the two results. The streamer image looks noisy in general because of the sparse shot spacing. The area beneath the obstruction is poorly illuminated. The OBN image is less noisy because of the dense shot carpet. However, there is a spurious event at depth z=1200 m. This spurious reflector is caused by imperfect up-down separation. In addition, the dipping reflector near x=5500 m and z=2400 m is not well illuminated in the OBN image.

For comparison, we individually perform linearized inversion on the two data sets. The resulting images are shown in (a) and (b) of Figure 5. We can see an overall improvement from the migration image in (c) and (d) of Figure 4 to the inversion image in (a) and (b) of Figure 5. In general, we can see that the inverted images contain higher resolution and better amplitude information than the migrated images. Linearized inversion can discrimmiate certain artifacts and remove them from the image. For example, the annotate artifacts in Figure 4 c) are suppressed in 5 (a). For the OBN survey, one noticeable improvement is the wider illumination. In the next section, we will explore the effect of joint inversion on the output image.

Joint Inversion

A joint inversion is performed in a least-squares sense with the objective goal described in equation 3. Panel (c) of Figure 5 shows the image after joint inversion. Notice that the joint image is better than the streamer inversion image around the platform area (x=3950-4000 m), as shown in (c) and (a) of Figure 6, respectively. Moreover, the joint image is better than the OBN inversion image for the dipping reflectors, as shown in (b) and (c) of Figure 7.

In our example, the joint image is able to attenuate, but cannot completely remove the spurious artifact caused by residual up-going energy in the OBN mirror signal. One reason for this is that the streamer data are not well illuminated in that area. Figure 8 compares the energy between the synthetic observed, the forward-modeled, and the residual data for one common-receiver gather in the OBN survey. The forward-modeled data is generated by applying $L_{OBN\downarrow}$ to the final inverted model m after 30 iterations. We can see that most of the energy in the synthetic observed data is explained by the forward-modeled data. Also, notice that the relative amplitude of the residual up-going energy (near 2.3 sec) becomes

higher.

DISCUSSION

Our synthetic study shows improvements of the joint inversion image over migrating or inverting either signal alone. There are several considerations that need to be addressed when applying these results to field data.

One consideration is related to the matching of the streamer and the OBN images, because the two surveys are acquired with different sources. Naturally, this leads to different wavelets in the data. Careful deconvolution is needed to make sure the wavelet in the image is consistent between the two surveys. Note that the two data sets are not required to have exactly the same frequency band but only the same phase (e.g., zero-phased).

Another consideration when applying this method to field data is the relative position of the reflectors between the two images. Due to different weather conditions, the water velocity can be different between the two surveys. If the water velocity and static correction is not properly adjusted, the sea-bottom reflector between the two images might not be the same. In addition, errors in positioning sources and receivers can also cause mismatch between the two images. If the mismatch is not severe, one can consider applying warping in the image space (Boelle et al., 2010; Hale, 2011; Ayeni, 2011) by estimating appropriate shifts to obtain a good match. For conjugate-gradient type inversion, this simple technique requires warping when applying the migration operator:

$$\begin{aligned} \Delta\mathbf{m} &= \Delta\mathbf{m}_{str} + \mathbf{W}\Delta\mathbf{m}_{OBN}, \\ &= \mathbf{L}_{str}^T\mathbf{r}_{str} + \mathbf{W}\mathbf{L}_{OBN}^T\mathbf{r}_{OBN}, \end{aligned} \tag{4}$$

where $\mathbf{W}$ is the forward warping operator that will match the OBN image update $\Delta\mathbf{m}_{OBN}$ with the streamer image update $\Delta\mathbf{m}_{str}$. $\Delta\mathbf{m}$ is the total image update. In the modeling direction, we apply unwarping before forward-modeling the gradient:

$$\begin{aligned} \Delta\mathbf{r} &= \Delta\mathbf{r}_{str} + \Delta\mathbf{r}_{OBN}, \\ &= \mathbf{L}_{str}\Delta\mathbf{m}_{str} + \mathbf{L}_{OBN}\mathbf{W}^{-1}\Delta\mathbf{m}_{OBN}, \end{aligned} \tag{5}$$

where $\Delta\mathbf{r}$, $\Delta\mathbf{r}_{str}$ and $\Delta\mathbf{r}_{OBN}$ search the total, streamer, and OBN search directions. $\mathbf{W}^{-1}$ is the unwarping operator.

The computational cost of linearized inversion is higher than that of migration by a factor proportional to twice the number of iterations. For OBN surveys, such additional computational cost can still be affordable. This is because the number of pre-stack calculations needed equals the number of OBN receivers in the survey. The number of migrations needed is substantially smaller than that in a towed-streamer survey. However, for towed-streamer surveys, the computational cost can be substantial. Recently, Dai et al. (2011) suggested using phase-encoding in the linearized inversion, which can reduce the computational cost substantially at the expense of introducing crosstalk into the image. In the future, we plan to combine not only the streamer and the OBN mirror signal, but to combine all 3 modes to produce one coherent image. The three modes are OBN primary, OBN mirror, and streamer data sets:

$$0 \approx \begin{bmatrix} \mathbf{L}_{str} \\ \mathbf{L}_{OBN\uparrow} \\ \mathbf{L}_{OBN\downarrow} \end{bmatrix} \mathbf{m} - \begin{bmatrix} \mathbf{d}_{str} \\ \mathbf{d}_{OBN\uparrow}\mathbf{d}_{OBN\downarrow} \end{bmatrix}. \tag{6}$$

CONCLUSION

Joint inversion can coherently combine the information from the narrow-azimuth towed-streamer and ocean-bottom data. We demonstrate the result with a 2D synthetic Marmousi model. The results show that the joint image is better than the streamer inversion image around the platform area. In addition, the joint image is better than the OBN image for dipping reflectors.

REFERENCES

Ayeni, G., 2011, Cyclic 1D matching of time-lapse seismic data sets: A case study of the norne field: SEG Technical Program Expanded Abstracts, **30**, 4149–4154.

Boelle, J.-L., D. Lecerf, A. Lafram, and J. Cantillo, 2010, Ocean bottom node processing in deep offshore environment for reservior monitoring: SEG Technical Program Expanded Abstracts, **28**, 4190.

Dai, W., X. Wang, and G. Schuster, 2011, Least-squares migration of multisource data with a deblurring filter: Geophysics, **76**, R135–R146.

Dash, R., G. Spence, R. Hyndman, S. Grion, Y. Wang, and S. Ronen, 2009, Wide-area imaging from OBS multiples: Geophysics, **74**, Q41–Q47.

Hale, D., 2011, A method for estimating apparent displacement vectors from time-lapse seismic images: Geophysics, **74**, V99–V107.

Wong, M., B. Biondi, and S. Ronen, 2010, Joint least-squares inversion of up- and down-going signal for ocean bottom data sets: SEG Expanded Abstracts, **29**, 2752–2756.

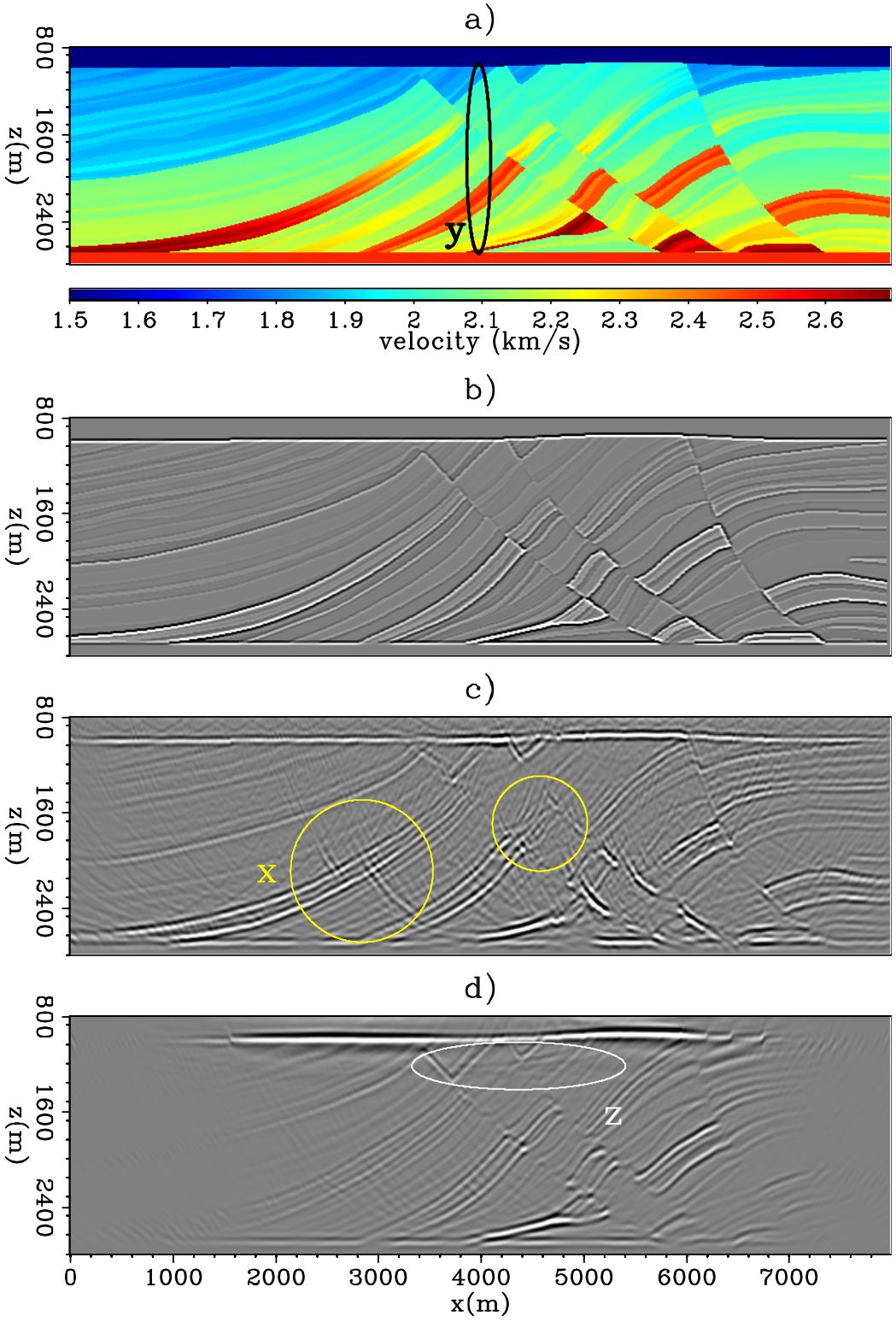


Figure 4: (a) The velocity model, (b) the reflectivity model, (c) the streamer RTM image, and (d) the OBN RTM image with mirror imaging. Even with undershooting, the streamer image shows some artifacts (x) near the production structure (y). The residual up-going OBN energy in the down-going OBN data create some artifacts (z). [**CR**] mandy1/. panel1

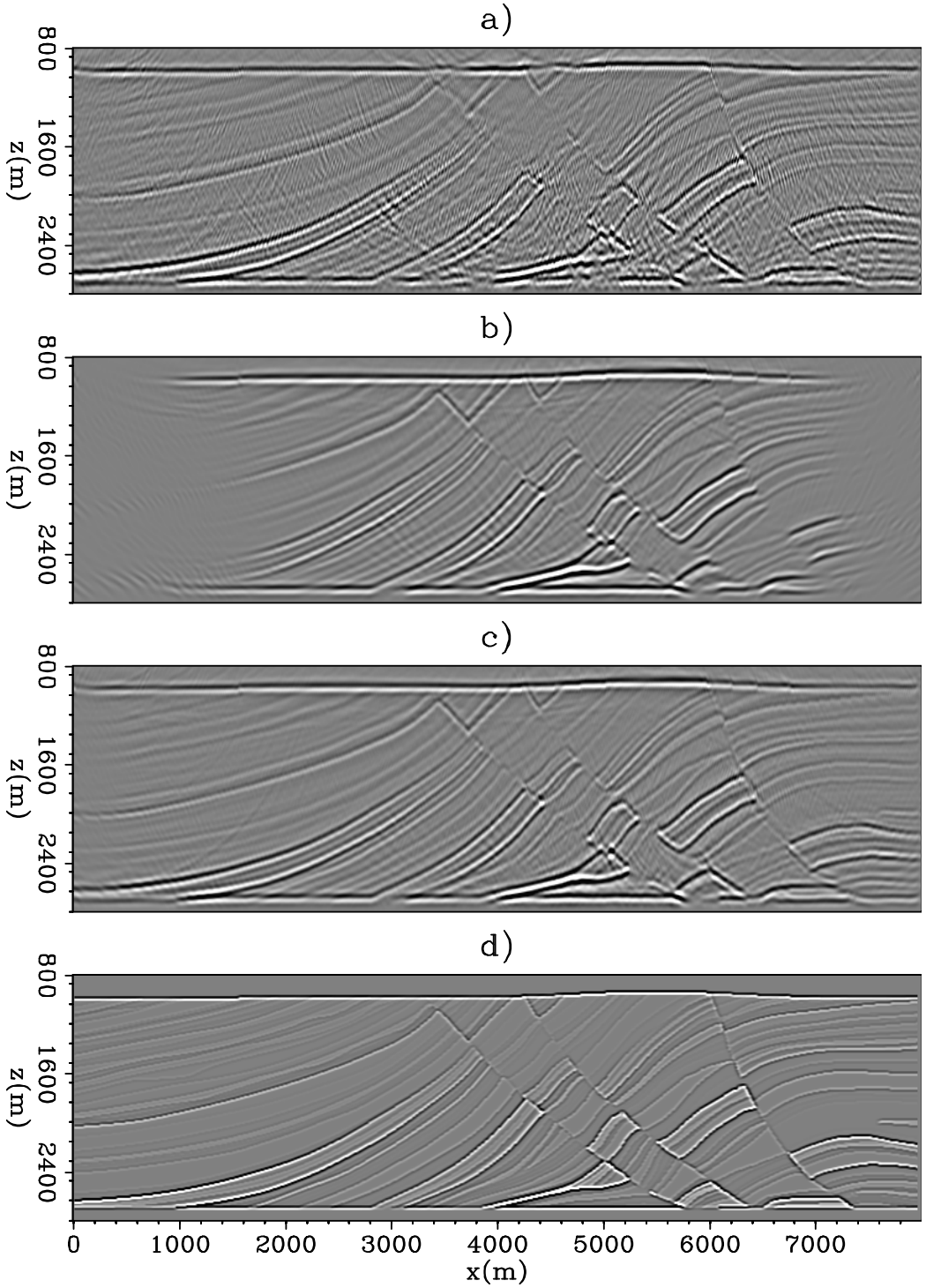


Figure 5: (a) The inverted streamer image, (b) the inverted OBN image, (c) the jointly inverted image, and (d) the reflectivity model. Notice the joint image coherently combines the information between the individually inverted images. [**CR**] mandy1/. panel2

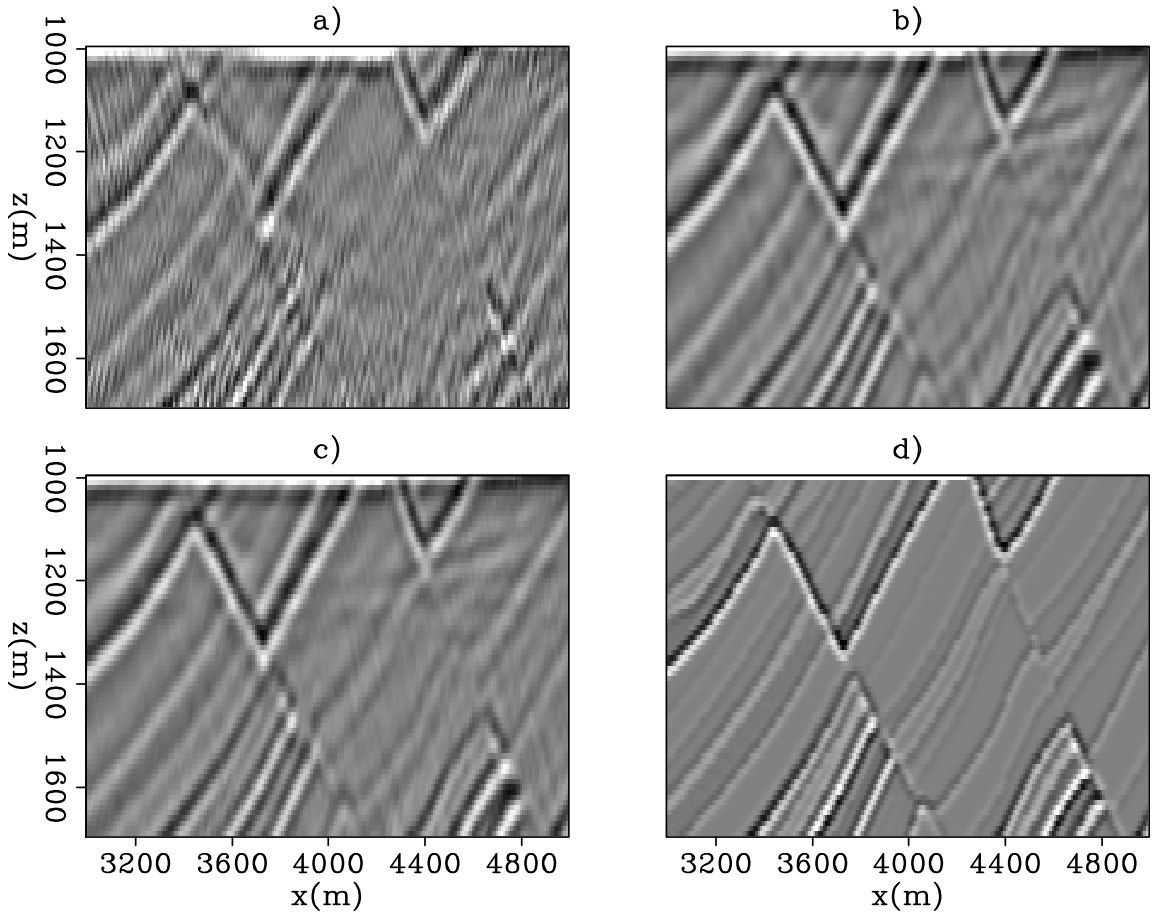


Figure 6: A zoomed section from x=3000-5000 m and z=50-750 m among (a) the streamer inversion image and (b) the OBN inversion image, (c) the joint inversion image and (d) the reflectivity model. The platform is located at horizontal position of x=3950-4000 m. In this region, most of the useful information comes from the OBN mirror data. The joint image is able to attenuate but cannot completely remove the spurious artifact caused by residual up-going energy in the OBN mirror signal. [**CR**] mandy1/. Zoom

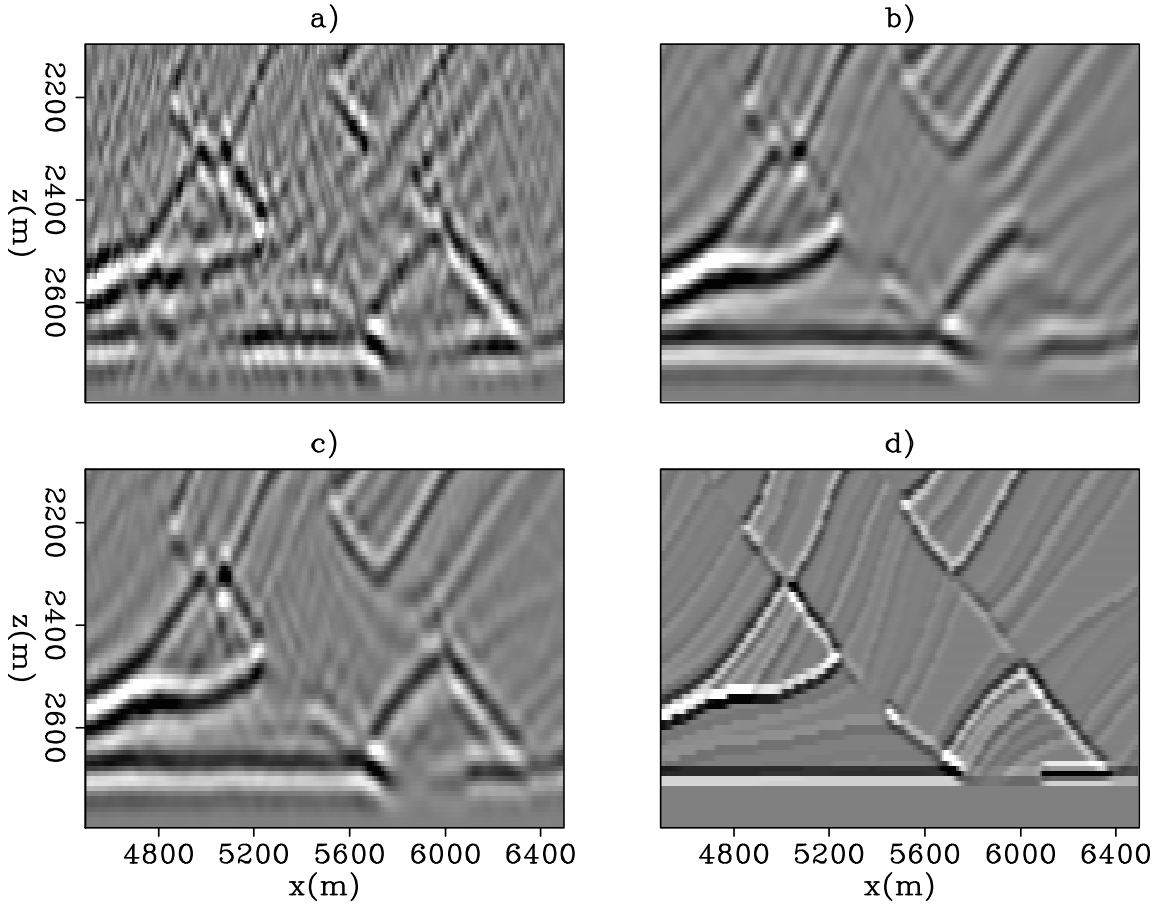


Figure 7: A zoomed section from x=4500-6500 m and z=1300-2000 m among (a) the streamer inversion image, (b) the OBN inversion image, (c) the joint inversion image and (d) the reflectivity model. Although the streamer image looks noisy, joint inversion can properly include the dipping reflector from the streamer image to produce a better joint image. [**CR**] mandy1/. Zoom2

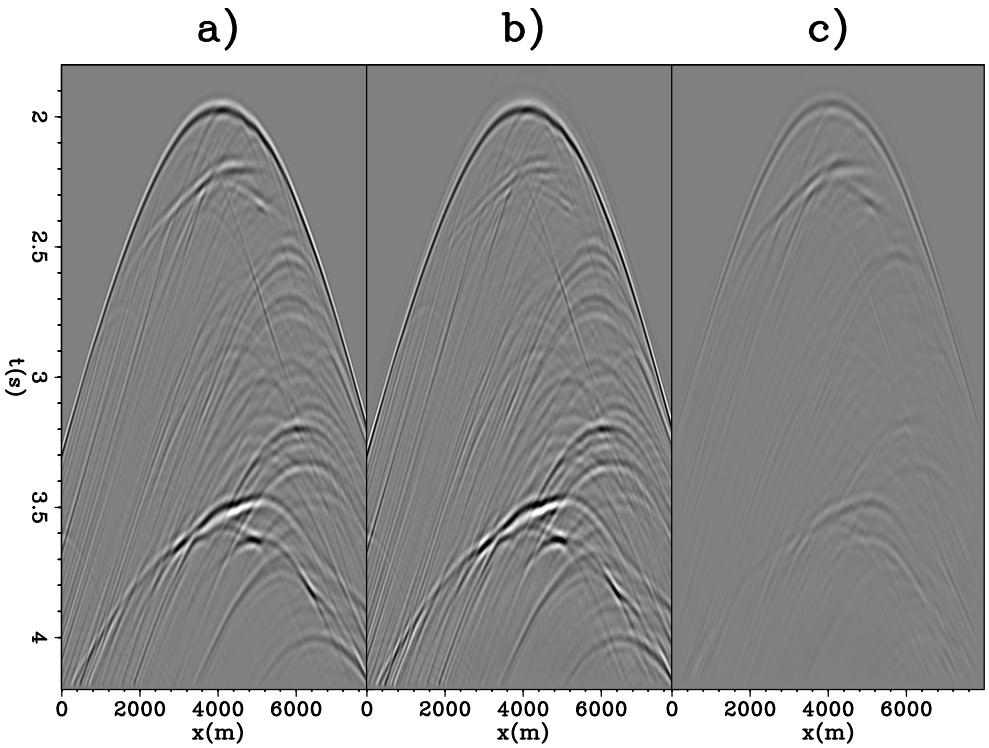


Figure 8: (a) The synthetic observed data, (b) the forward-modeled data, and (c) the data residual for one common receiver gather in the OBN survey. All data are clipped at the same bound. Notice that the relative amplitude of the residual up-going energy becomes higher in the data residual. By incorporating more data to produce a single model, joint inversion can help discrimminate between signal and noise in the data. [**CR**] mandy1/. Resid

Imaging with multiples using linearized full-wave inversion

Mandy Wong, Biondo Biondi, and Shuki Ronen

ABSTRACT

We presents a technique for imaging both primaries and multiples using linearized inversion. When used with a suitable migration velocity model, linearized full-wave inversion (LFWI) makes use of the multiple energy as signal while removing the crosstalk in the image. By using the two-way propagator in both modeling and migration, we can image a class of multiply scattered events. Such events can scatter off sharp-interfaces in the migration velocity many times but only interact with the (reflectivity) model once. We demonstrate the concept and methodology in 2D with a synthetic Sigsbee2B model.

INTRODUCTION

Traditionally, seismic imaging techniques only account for the primary reflections. In the presence of strong reflectors (e.g.. air-water interface, hard water bottom or salt bodies), multiples can significantly degrade the interpretation of images. Therefore, much effort has been devoted to developing multiple suppression techniques in the past few decades.

The well-known multiple-removal tools such as deconvolution (in time, frequency, and slant-stack domains), Radon-transform demultiple and frequency-wavenumber (f-k) demultiple are limited unless the geology of the subsurface is simple. In the presence of complex geology, multiples are not totally separable from primaries by criteria such as periodicity, moveout velocity, and spectra. In model-based technique, which predict multiples with wavefield extrapolation (Morley, 1982; Berryhill and Kim, 1986; Wiggins, 1988; Lu et al., 1999), the accuracy of the predicted multiples strongly depends on the model used. Surface-related multiple elimination (SRME), a convolution-based technique (Riley and Claerbout, 1976; Tsai, 1985; Verschuur et al., 1992), is more generally applicable. But this method requires an overlap of source and receiver locations and cannot suppress internal multiples. Despite substantial progress in multiple elimination, complete removal of surface-related and internal multiples without distorting the primary signals remains a challenge.

One motivation to make use of multiples is that they can provide subsurface information not found in primaries. For a given pair of source and receiver, the sub-surface reflection point of a multiple is located differently than that of a primary. For a multi-shot seismic survey, migrating the multiples translates to higher fold for regions well imaged by the primaries and more subsurface illumination for regions not covered by the primaries. In addition, despite the advance of multiple suppression techniques, complete removal of all multiples from the primaries still remains a challenge. Migrating such signals would result in *crosstalk* artifacts. Finally, multiples are even more sensitive to velocity information than the primaries, because they travel a longer path in the subsurface. Multiple signals can potentially be used as part of an iterative velocity building tool. Reiter et al. (1991)

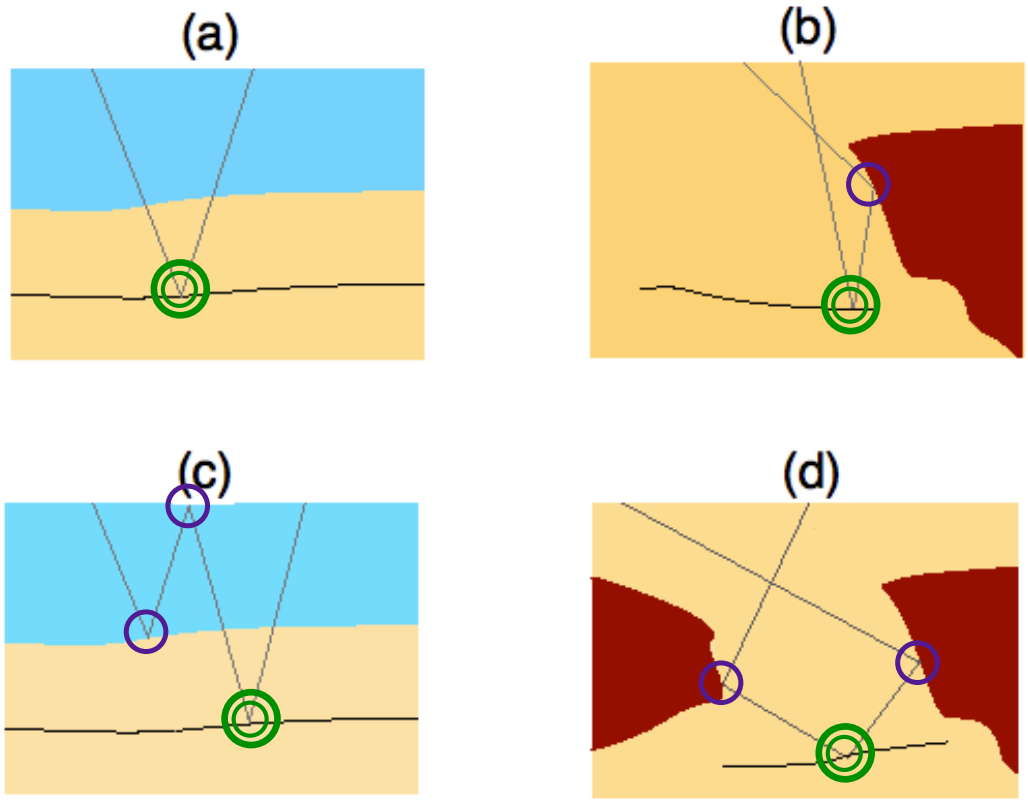


Figure 1: The ray-path for a (a) singly scattered event, (b) a doubly scattered event and (c,d) triply scattered events. Single circles (in purple) indicate scattering off the migration velocity while double circle (in green) indicate scattering off the model $m(\mathbf{x})$. [**NR**] mandy2/. LSMimage2v2

made an early attempt to capitalize on the potential of multiples by formulating a prestack Kirchhoff time-migration method that includes the first-order water-layer reverberation in the migration operator. Because ocean bottom cable data could not be decomposed into up- and down-going components at the time, such work was limited to deep water datasets.

When surface-related multiples have been explicitly separated from the primary reflections (e.g., using SRME), they can be imaged independently from the primary reflections by using shot-profile (Guitton, 2002) or source-receiver (Shan, 2003) depth migration. Muijs et al. (2007) image primary and free-surface multiples for OBS data by decomposing data into up-going and down-going constituents, followed by downward extrapolation and a 2D deconvolution-based imaging condition. All these techniques image the surface-related multiples by transforming the primary signal into a pseudo-source for migration with the multiple signals using the one-way wave equation. Recently, Liu et al. (2011) extend the technique to the two-way wave equation. While these techniques are computationally efficient, their images contain crosstalk artifacts caused by the interference of wavefields not associated with the same subsurface reflector. A robust technique is needed to get the benefit of multiple imaging without compromising image quality. We propose using linearized full-wave inversion (LFWI) to use multiple energy as signal while removing the crosstalk in

the image.

In the next section, we discuss how to image with multiples with LFWI for the streamer geometry and the OBN geometry. We then show the results of our inversion scheme from a layered model and a 2D Sigsbee2B model for the OBN geometry.

THEORY

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

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

Assuming that the earth behaves as a constant-density acoustic isotropic medium, we linearize the wave equation and apply the first-order Born approximation to get the following forward modeling equation:

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

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. It is important to point out that the adjoint of the forward-modeling operator is the migration operator:

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

The inversion problem is defined as minimizing the least-squares difference between the synthetic and the recorded data:

$$S(\mathbf{m}) = \|\mathbf{L}\mathbf{m} - \mathbf{d}\|^2 = \|\mathbf{d}^{mod} - \mathbf{d}\|^2 \tag{4}$$

where $\mathbf{L}$ is the forward-modeling operator that corresponds to equation 2.

At first glance, equation 2 seems to only generate singly scattered events (e.g. Figure 1 a). To clarify, the term scattering includes both diffraction and reflection. However, if we construct our propagator ($G(\mathbf{x}, \mathbf{y})$) using the two-way wave equation, equation 2 can actually generate multiply scattered events. In figure 1 b, the ray path reflects off a salt flank and then the horizontal reflector. If the sharp salt-flank boundary already exists in the migration velocity, then the scattering off the salt flank is automatically generated by the propagator (Green's function). Figure 1 (c) and (d) shows two triply scattered events. Single circles (in purple) show scattering off the migration velocity, while double circles (in green) show scattering off the model $m(\mathbf{x})$.

Multiple imaging with towed streamer

For a towed-streamer geometry, one only needs to introduce (i) the free-surface and (ii) sharp boundaries into the migration velocity $s_o(\mathbf{x})$ to begin modeling both primaries and multiples. If the sea-bottom has a sharp interface in $s_o(\mathbf{x})$, then equation 2 can model all surface-related multiples as shown in Figure 2 (a). The same is true for ocean-bottom data, which we will discuss next.

Multiple imaging with ocean-bottom node

In an ocean-bottom survey, the signal (Figure 3) can be classified into up-going (in red) and down-going signal (in grey) with respect to the receivers. The lowest order of the up-going signal is the primary reflection. The lowest order of the down-going signal is the direct-arrival. Since the direct-arrival does not carry any information about the subsurface, the next order of down-going event, the mirror signal, is used for conventional migration of down-going OBN data.

We will referred the mirror signal as the down-going primary from now on. This is not to be confused with the up-going primary. One can apply LFWI on the ocean-bottom hydrophone recording without decomposing the up- and down-going signal. However, using only the hydrophone without up-down decomposition for imaging is less favorable due to the following reasons:

1. The migration image contains more crosstalk artifacts due to the migration of the up- and down-going modes. Although inversion can, in theory, push out the crosstalk artifacts, it requires more iterations.

2. The energy in the up-going primary is stronger than that in the down-going primary. If the up-down modes are not separated in the data space, it is difficult for the inversion to apply data weighting to balance the contribution from the two modes.

3. Up-down decomposition by PZ summation can alleviate loss signal at the notch frequencies. A benefit that is loss if imaging only with the hydrophone.

In this study, we focus on imaging different orders of down-going signals as shown in Figure 2 b and c. To simulate the final down-going leg of the wave path, an areal shot is precalculated by first injecting the source wavelet at the receiver location, letting the wavefield propagate, and then capturing the signal at the sea surface (Figure 4). To generate the incident wavefield, the saved areal shot is re-injected at the sea-surface with a -1 factor. The re-injected signal is then allowed to travel back and forth in the water column using a reflecting top boundary and a well-defined velocity contrast at the sea-bottom. This algorithm can correctly simulate the ray path traversed by the down-going primary and higher-order multiples.

The focus of this report is on imaging the higher-order multiples for OBN data with the method illustrated in Figure 4. In particular, we compare the image output between migration with the down-going primary and LFWI with both the down-going primary and the down-going multiples.

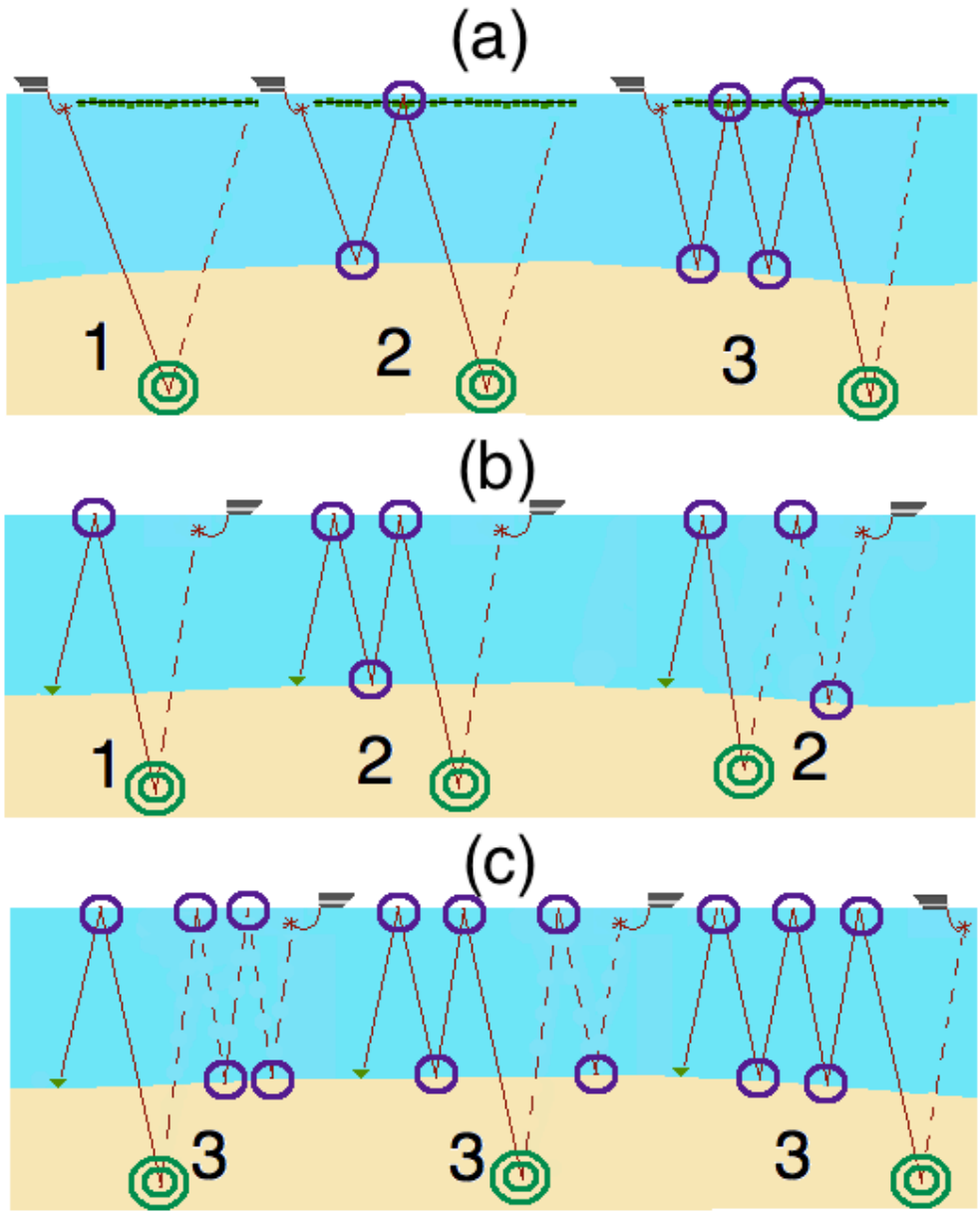


Figure 2: (a) Ray paths for the primary, the first-order, and the second-order multiples for the towed-streamer acquisition geometry. (b) Ray paths for the down-going mirror and the down-going first-order multiples for ocean-bottom-node acquisition. (c) Ray paths for the down-going second-order multiples for ocean-bottom-node acquisition. Purple-solid circles show scattering off either the free-surface or the background velocity model. Green-double circles show scattering off the reflectivity model. [**NR**] mandy2/. DiffOrder3

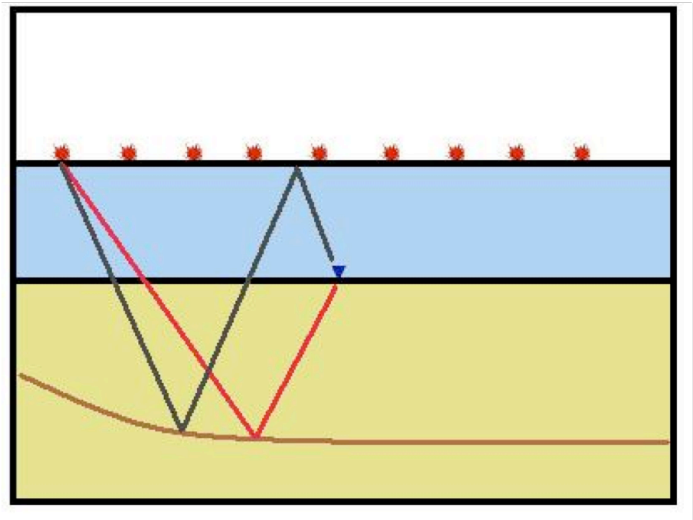


Figure 3: For a given pair of source and receiver, the sub-surface reflection point of a multiple event is different than that of a primary event. [**NR**] mandy2/. illum-m

SYNTHETIC EXAMPLE

We apply LFWI on two models. The first one is a simple one-layered model that allows us to keep track of different kinds of migration artifacts. The second model is the more complicated Sigsbee2B model.

One-layered model

We construct a one-layered model (Figure 5 a) with ocean-bottom geometry. The only sharp interface in the migration velocity is the seabed. Figure 5 b shows the synthetic data. The labels d_1, d_2 and d_3 correspond to the first, second, and third order events as shown in figure 2 b and c. Note that we used equation 2 to generate the synthetic data. Hence, internal multiples are absent. Figure 5 c shows the migration image m_{mig}. The migration image is made up of signal m_{sig} and crosstalk artifacts m_{xtalk}. In the figure, the label A indicates spurious reflectors generated by migrating the primary signal (d_1) as if it were a multiple. B is the correct reflector in the image. C is an artifact generated by migrating the multiple signal (d_2 or d_3) as if it were a primary reflection. In equation form, they are denoted as follows:

$$\begin{aligned} m_{mig} &= m_{signal} + [m_{xtalk}] = m_B + [m_A + m_C] \\ m_A &= \mathbf{L'}_2\mathbf{d}_1 + \mathbf{L'}_3\mathbf{d}_1 + \mathbf{L'}_4\mathbf{d}_1 + ... \\ m_B &= \mathbf{L'}_1\mathbf{d}_1 + \mathbf{L'}_2\mathbf{d}_2 + \mathbf{L'}_3\mathbf{d}_3 + ... \\ m_C &= \mathbf{L'}_1\mathbf{d}_2 + \mathbf{L'}_1\mathbf{d}_3 + \mathbf{L'}_2\mathbf{d}_3 + ... \end{aligned} \tag{5}$$

where m_A, m_B,and m_C correspond to the parts of the image labeled with A, B, and C in Figure 5 c. $\mathbf{L'}_1$, $\mathbf{L'}_2$, and $\mathbf{L'}_3$ are migration operators that correspond to different orders of reflection events. Figure 5 d shows the inversion result. Notice that the artifacts are

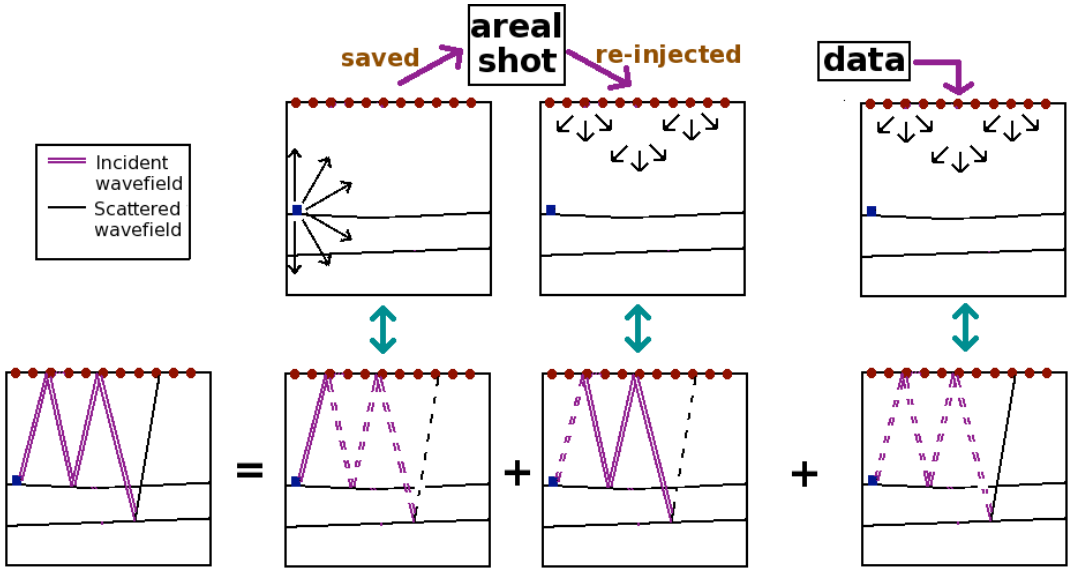


Figure 4: Illustration on the migration of the second-order signal. For the incident wavefield, an areal shot is pre-calculated to simulate the final down-going leg of the wave path. To generate the incident wavefield, the saved areal shot is re-injected at the sea-surface with a -1 factor. The re-injected signal is then allowed to travel back and forth in the water column using a reflecting top boundary and a well-defined velocity contrast at the sea-bottom. [**NR**] mandy2/. Figure3

removed from the image. In conventional imaging, if there were residual multiple energy in the data, then artifacts of type C would show up in the image. Treating those as real signal would negatively affect the interpretation of the sub-surface.

Sigsbee2B model

We apply LFWI to the Sigsbee2B model with the ocean-bottom geometry. Figure 7 shows the migration velocity and the reflectivity model used for this study. There are two interfaces in the migration velocity. One comes from the salt and the other comes from the basement reflector. These sharp interfaces along with the free surface boundary condition will generate the multiples. We first generate the down-going primary only data by applying the conventional mirror-imaging forward modeling operator. This is done by positioning the nodes at the mirror point across the sea-surface as shown in Figure 6. We then add 20 % of the down-going multiple energy to the down-going primary only data as shown in Figure 8. This is to simulate the case when multiple-elimination techniques cannot completely remove the multiples. We will refer to this synthetic simply as the *noisy primary* data. Figure 8 (b) shows the synthetic that contains first- and higher-order reflections. We will refer to this as the *primary-multiple* data.

Figure 9 shows the result of applying migration and linearized inversion to the data. Panel (a) is the conventional image in which we assume a primary-only migration oper-

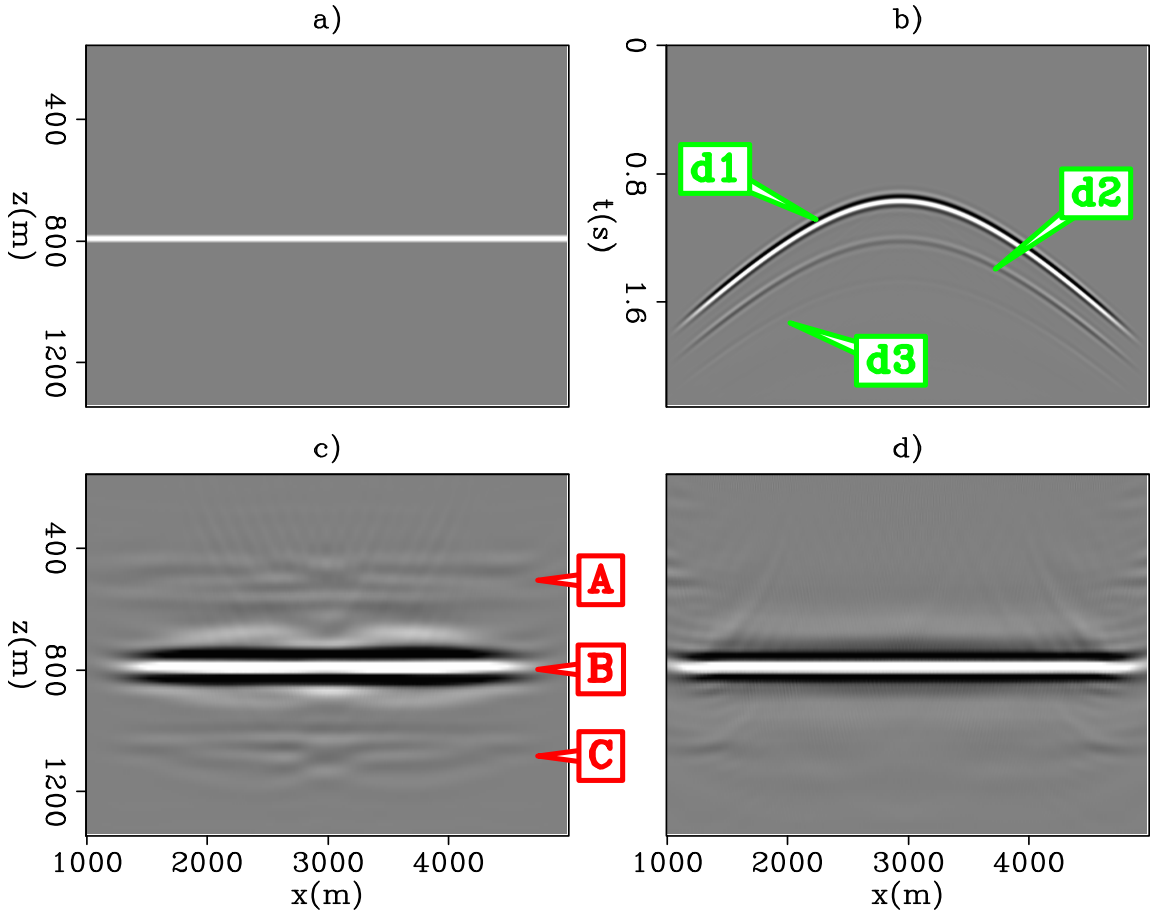


Figure 5: (a) Original one layered model, (b) synthetic data, (c) migration image and (d) inversion image. [**CR**] mandy2/. Onelayerv2

ator on the primary data. The red arrow indicates the artifacts in the migration image that correspond to noise included with the primary. Panel (b) is the linearized inversion result with the primary-only operator. Note the substantial improvement between migration and inversion. It has been shown (Wong et al., 2011)that linearized inversion can enhance the resolution of the image, suppress the migration artifacts, and increase the relative amplitude of true reflectors. Next, panel (c) shows the migration result achieved by applying a primary-multiple migration operator to the primary-multiple data. There are many crosstalk artifacts in this image. Without addressing the crosstalk, it is difficult to argue that panel (c) is better than panel (a). Panel (d) is the linearized full-wave inversion (LFWI) result. The annotation indicates that LFWI can properly remove the crosstalk artifacts.

In terms of convergence, since this is a classical *inverse crime* study, our objective function decreases to two percent of its initial value after 40 iterations.

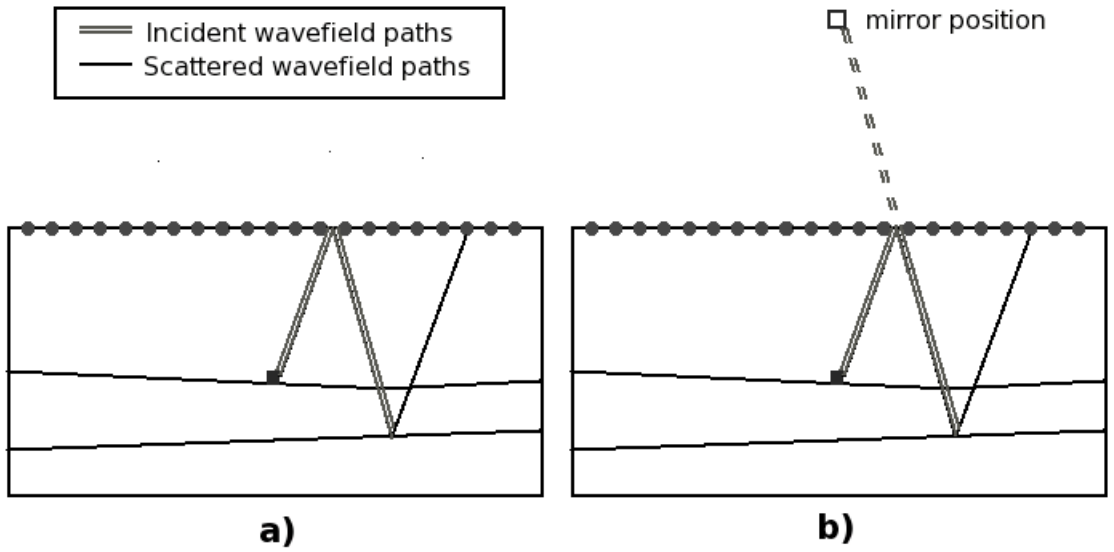


Figure 6: (a) The ray-path of a mirror signal. (b) The raypath of the same signal in mirror imaging. The apparent position of the receiver is not at twicce the ocean depth above the sea bed. This assumes the sea surface is a perfect reflector. Note that the top boundary is absorbing. [**NR**] mandy2/. mirror

DISCUSSION

Note that this technique does not migrate all orders of multiples. It only migrates multiples with a single scattering off the model $m(\mathbf{x})$ and other scattering off the sharp boundary in the migration velocity model. Considering that multiples with high amplitude in the data are often generated by sub-surface interfaces of high impedance contrast, this technique can account for most of the significant multiples in the data.

Our method is model-based. One obvious consideration is the accuracy of the migration velocity. A conservative way of applying LFWI would be to put in sharp interfaces that are easy to estimate (e.g. the free-surface, sea-bottom, and top-salt). However, what happens if there is a mis-positioned salt flank in the migration velocity? This may open an avenue for velocity estimation. As an imaging tool, LFWI works as well as other model-based multiple-prediction-subtraction methods.

We recommend applying LFWI for surveys where multiple removal is an issue. An appropriate field study for this method would be a shallow-water dataset with a deep target zone. In this case, each order of multiples overlaps with the previous order, and the conventional multiple-prediction-subtraction techniques might not deliver.

CONCLUSION

We demonstrated a method for imaging both primaries and multiples using linearized full-wave inversion (LFWI). LFWI not only increases the sub-surface illumination by using

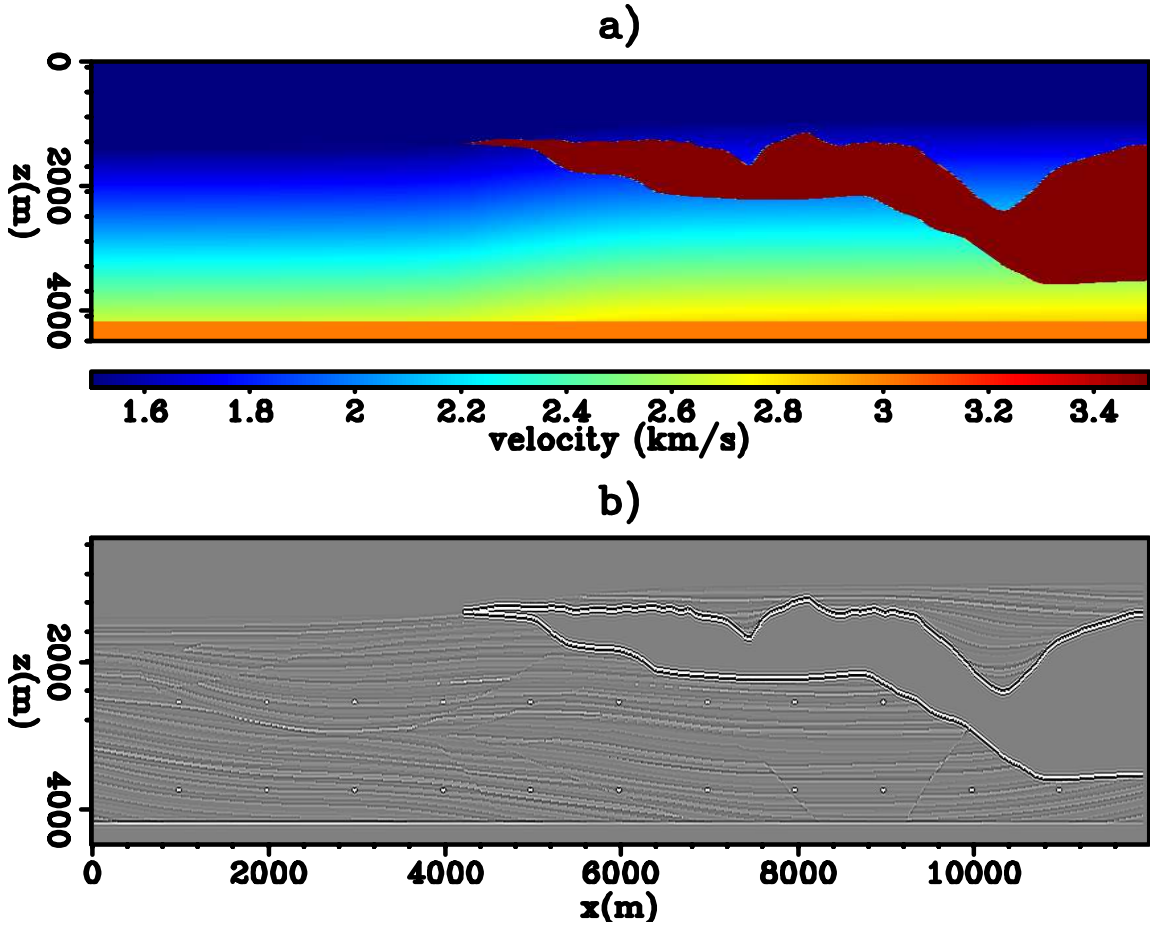


Figure 7: (a) Migration velocity model and (b) original reflectivity model. [**ER**] mandy2/. panel1-m

the multiple energy as signal, but also addresses the issue of crosstalk in the image. We demonstrate the concept and methodology with a 2D layered model and the Sigsbee2B model.

REFERENCES

Berryhill, J. L. and Y. C. Kim, 1986, Deep-water peglegs and multiples — emulation and suppression: Geophysics, **51**, 2177–2184.

Guitton, A., 2002, Shot-profile migration of multiple reflections: SEG Annual Meeting Expanded Abstracts, 1296–1299.

Liu, Y., X. Chang, D. Jin, R. He, H. Sun, and Y. Zheng, 2011, Reverse time migration of multiples: SEG Annual Meeting Expanded Abstracts, 3326–3330.

Lu, G., B. Ursin, and J. Lutro, 1999, Model-based removal of water-layer multiple reflections: Geophysics, **64**, 1816–1827.

Morley, L., 1982, Predictive multiple suppression: PhD thesis, Stanford University.

Muijs, R., J. Robertsson, and K. Holliger, 2007, Prestack depth migration of primary and surface-related multiple reflections: Part I - Imaging: Geophysics, **72**, S59–S69.

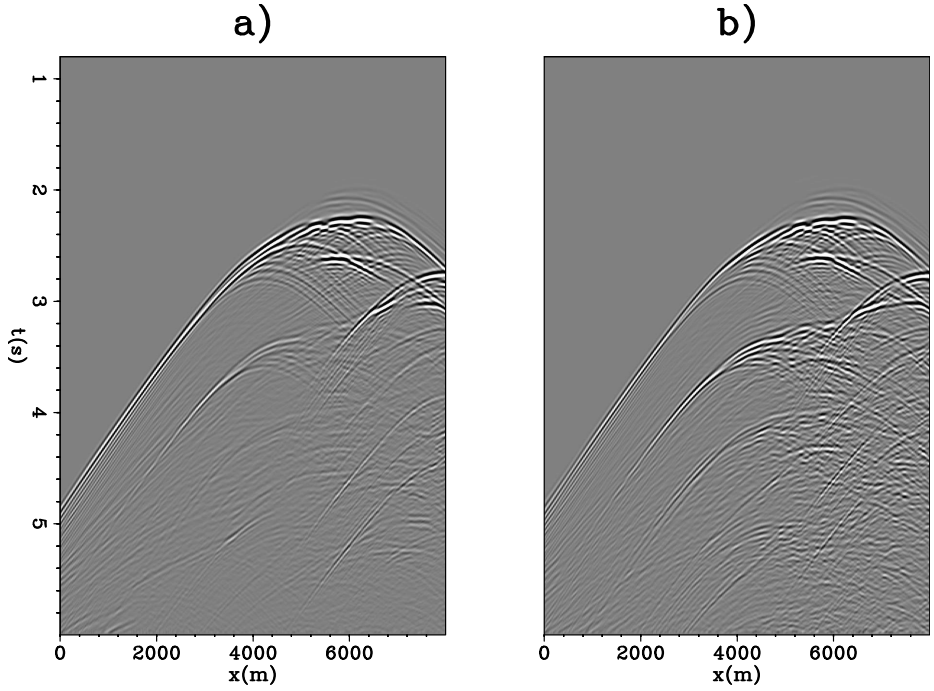


Figure 8: (a) The synthetic primary (lowest-order) data with 20 % multiple energy (*the noisy primary data*). (b) The primary and multiples synthetic data (*the primary-multiple data*). [**CR**] mandy2/. dataCom

Reiter, E. C., M. N. Toksoz, T. H. Keho, and G. M. Purdy, 1991, Imaging with deep-water multiples: Geophysics, **56**, 1081–1086.

Riley, D. C. and J. F. Claerbout, 1976, 2D multiple reflections: Geophysics, **41**, 592–620.

Shan, G., 2003, Source-receiver migration of multiple reflections: SEG Annual Meeting Expanded Abstracts, 1008–1011.

Tsai, C. J., 1985, Use of autoconvolution to suppress first-order long period multiples: Geophysics, **50**, 1410–1425.

Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, 1166–1177.

Wiggins, J. W., 1988, Attenuation of complex water-bottom multiples by wave equation-based prediction and subtraction: Geophysics, **53**, 1527–1539.

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.

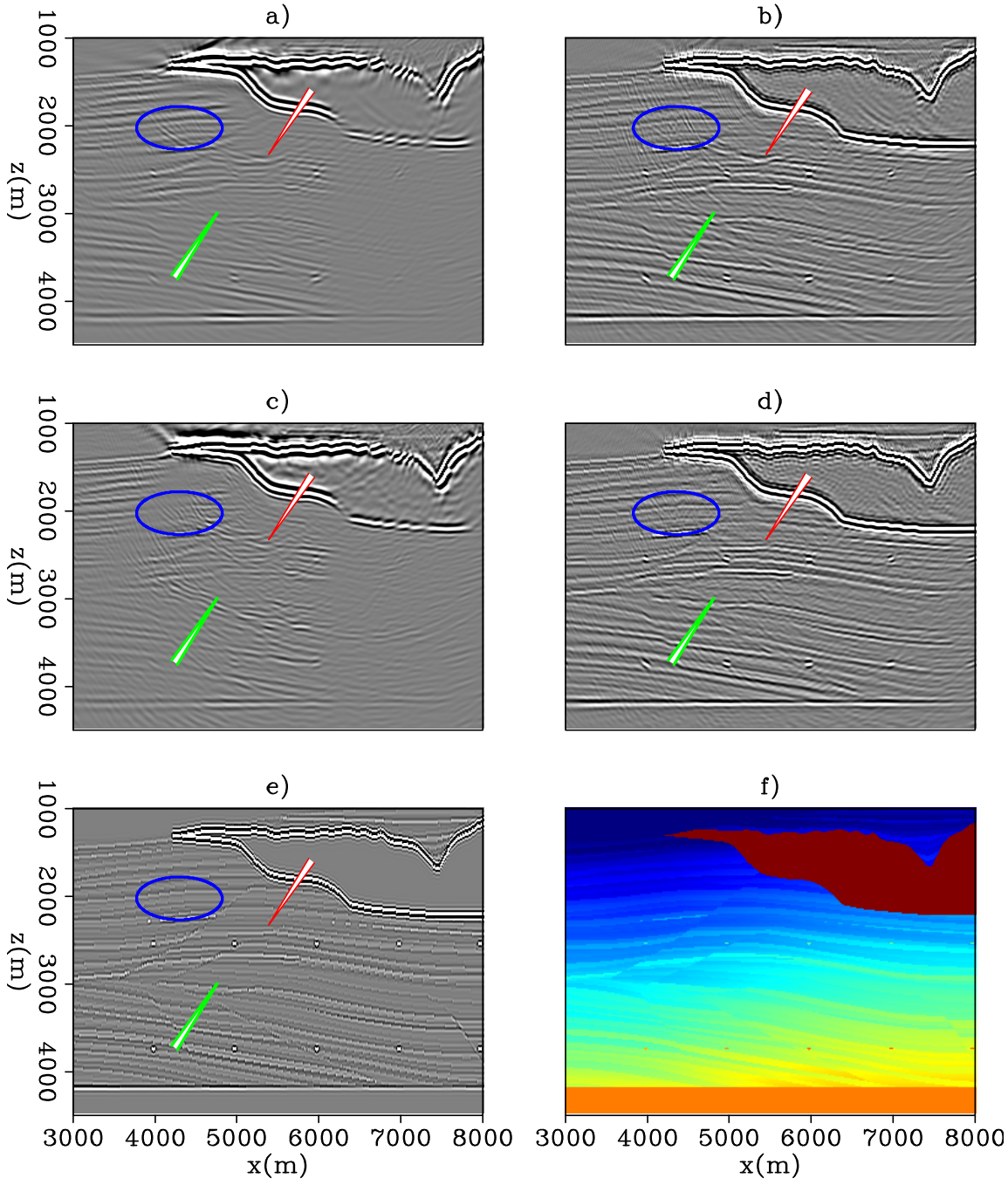


Figure 9: (a) Migration with the *noisy primary* data, (b) inversion with the *noisy primary* data, (c) migration with the *primary-multiple* data, (d) inversion with the *primary-multiple* data, (e) reflectivity model, and (f) velocity model. [**CR**] mandy2/. Zoomv2

P/S separation of OBS data by inversion in a homogeneous medium

Ohad Barak

ABSTRACT

A separation of pressure waves from shear waves in OBS data prior to processing can improve the resulting image. I present a nearly medium-independent P/S separation method, which operates by reconstructing the observed data using a homogeneous medium. The closer the homogeneous medium parameters are to the actual parameters of the medium in which the receivers are planted, the better the separation results. Synthetic tests indicate that the proposed method works reliably for land data even when using wrong medium parameters, but cannot work for OBS data in its current configuration.

INTRODUCTION

Ocean-bottom seismic (OBS) surveys record four-component data: hydrophone pressure data, vertical particle velocity (or acceleration), and two perpendicular horizontal particle velocities. The receivers are located on the sea bed. The source is typically an airgun fired from near the sea surface.

One problem of OBS and of marine streamer acquisition is the inevitable recording of water surface multiples. These multiples are waves which have reflected off the subsurface, and then bounce off the water surface. Though such waves contain subsurface data, attempting to remove them has been part of the traditional processing flow. Unlike marine streamer data, where the receivers record only pressure waves, in OBS acquisition additional energy is recorded. Since the geophones are coupled to the sea bed, they pick up shear waves and other surface waves as well as compressional waves.

A current implementation for processing of OBS data is to regard the hydrophone component and the vertical particle velocity component as containing P-waves only. PZ summation is applied to these components. It is a weighted summation of the hydrophone and the vertical geophone components, designed to determine the vertical propagation direction of the recorded arrivals (Barr and Sanders, 1989). The basic principle of PZ summation is that the hydrophone and the vertical geophone data will have opposite polarity if the energy they recorded was propagating downward. If that energy was propagating upward, they will have the same polarity. This upgoing/downgoing wavefield direction determination method has been used to create separate images from the primary arrivals and the mirror arrivals (the ones reflecting off the sea-surface) at the nodes on the sea bottom (Dash et al., 2009). This method is known as "mirror imaging". Separating the imaging process removes artifacts resulting from cross-talk between the upgoing and downgoing wavefields. Additionally, mirror imaging increases the effective aperture of the survey, and provides more information than the upgoing image alone.

Shear waves do not propagate in the water column. Therefore the result of applying PZ summation will be to introduce any non-P energy recorded by the vertical geophone into the summation result. These data will be construed as a P-wave energy recording. When these data are then migrated using acoustic wavefield propagation, they may generate artifacts in the resulting image. A simple graphical demonstration of how PZ-summation can introduce shear wave noise into the summation result can be seen in Figure 1.

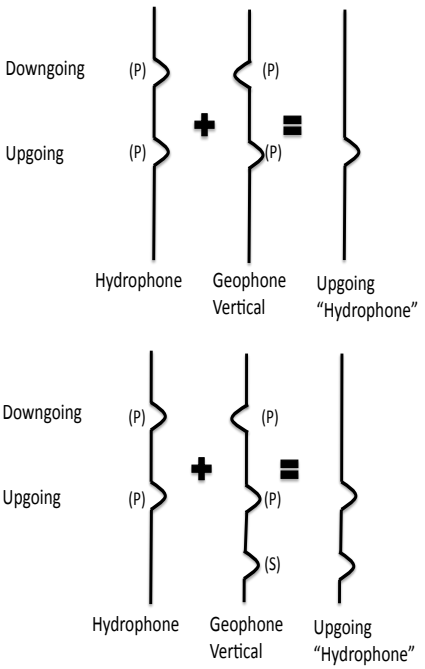


Figure 1: Simplified sketch of PZ summation. Top: For acoustic data, the summation of the hydrophone data and the vertical geophone data will result in the downgoing energy being eliminated, leaving only upgoing pressure data. Bottom: If the data contains shear waves, they will be recorded only by the geophone, and PZ summation will mistakingly add them into a result considered as upgoing pressure data. [**NR**] ohad1/. up-down-sketch

A particular problem of OBS data is "Vz noise" (Paffenholz et al., 2006), which may result from surface waves propagating along the water-solid boundary. They can form by shear waves incident on an irregular sea bed topography, or on near surface scatterers. Surface waves can also be generated by upgoing shear-waves incident on a flat sea bed at large angles. Since surface waves appear mainly on the geophone components, they will also contribute to an incorrect PZ summation result if they are not properly removed from the data.

Existing P/S separation methods

Dankbaar (1985); Wapenaar et al. (1990); Amundsen (1993) have shown how to separate the P and S-waves of OBS data, and also separate between the upgoing and downgoing pressure energy. Their methods require a knowledge of medium parameters in the vicinity of the receivers. Schalkwijk et al. (2003) implement adaptive decomposition, in order to account for the unknown near-surface medium parameters on which the separation depends. This decomposition requires manual identification in each receiver gather of a single time window containing mainly downgoing waves, and of a window that contains mainly upgoing waves. The energy within these windows is then minimized using an itterative process. The filter resulting from this process is inverted for both the local medium parameters and the calibration between the different sensors.

Zhou et al. (2011) propose a novel method of extracting the P-wave data from the geophone components. It involves a complex wavelet transform of the hydrophone and the vertical geophone data. A matching of the geophone data to the predominantly P-wave energy hydrophone data is done in the complex wavelet domain. Afterward, the inverse transform is applied to the result. The end product is a substantially P-only vertical geophone component. PZ summation can then be applied to this component and the recorded hydrophone, to separate the upgoing from the downgoing P-wave.

P/S separation by inversion

If the seismic data were sampled vertically as well as horizontally (i.e. - if receivers were placed at multiple depths as well as on a surface), then the spatial derivatives of the wavefield could be calculated, and the separation between P and S-waves would be trivial. Robertsson and Muyzert (1999) propose doing volumetric recordings of the wavefield using tetrahedral receiver groups, for the purpose of calculating all spatial derivatives and thus separating P from S.

The P/S separation method presented here is a (nearly) medium-independent data-space inversion. The model it attempts to invert for is the wavefield, emitted by a virtual source array at some depth level, which after propagating through a homogeneous medium matches the recorded wavefield at the receiver level.

The basic premise of the separation method is: if there is a good match to the recorded data at the receivers, then there is also a good match to the "true" receiver wavefield at a close proximity to the receivers. The method effectively reconstructs the wavefield at the receiver's vicinity, as it was in the field experiment. Since we now have a volumetric spatial sampling of the wavefield near the receivers, separation operators based on three-dimensional spatial derivatives (i.e. - divergence and curl) are applied to separate P-waves from S-waves.

The synthetic seismic data I use in the following examples are displacement data, which represent geophone acquisition. The elements of the inversion are:

1. Observed data: The original displacement components recorded by the geophones in the field.

2. Model: A virtual source gather, consisting of displacement source functions, injected at some location into a homogeneous elastic medium.

3. Calculated data: The recorded displacements at the receiver level, as a result of the injection of the virtual sources.

4. Desired model: The virtual sources which generate recorded data equal to the observed data.

5. What we actually want: The wavefield displacement values both at the receiver level AND one depth level below them, so that we can apply spatial derivative operators to separate P from S.

THEORY

PZ summation

PZ summation involves summing the pressure data recorded by the hydrophone with the vertical particle velocity data recorded by the geophone, with some scaling factor:

$$\begin{aligned} U(z_r) &= \frac{1}{2}\left[P(z_r) - \beta V_z(z_r)\right], \\ D(z_r) &= \frac{1}{2}\left[P(z_r) + \beta V_z(z_r)\right], \end{aligned} \tag{1}$$

where P is the hydrophone data, V_z is the vertical geophone, U is the upgoing data, D is the downgoing data and z_r is the receiver depth. β is a scaling factor, which is theoretically the acoustic impedance at the wave's incidence angle. In practice, as a result of frequency-dependent instrument response, and as a result of the two different impedances above and below the receivers, β is frequency dependent, propagation-direction dependent, and wave-mode dependent.

Amundsen (1993) does the separation in the $f-k$ domain, and uses $\beta = \frac{\rho\omega}{k_z}$, where $k_z = \sqrt{\frac{\omega^2}{v^2} - k_x^2 - k_y^2}$, and ρ is the density. This implicitly requires a laterally invariant medium at the receiver level. Alternately in the $t-x$ space, the scaling can be determined by the ratio of the direct arrival's amplitude on the hydrophone and vertical geophone components at various offsets.

As emphasized above, this method assumes that all energy is pressure wave energy, and therefore everything recorded by the hydrophone should have its counterpart in the geophone data, with either positive polarity (upgoing) or negative polarity (downgoing).

Wave mode separation and imaging

The Helmholtz separation operator is based on the assumption that any isotropic vector field can be described as a combination of a scalar and vector potential fields:

$$\mathbf{u} = \nabla\Phi + \nabla \times \mathbf{\Psi}, \tag{2}$$

where Φ is the scalar potential field and $\mathbf{\Psi}$ is the vector potential. $\mathbf{u}$ is the elastic displacement vector wavefield. The scalar potential generates pressure waves, and the vector potential generates shear waves. Therefore, the Helmholtz method of separating the P-wave amplitude from the S-wave amplitude is to apply a divergence operator and a curl operator to the displacement field:

$$P = \nabla \cdot \mathbf{u} = \nabla^2\Phi; \tag{3}$$

$$\mathbf{S} = \nabla \times \mathbf{u} = -\nabla^2\Psi. \tag{4}$$

Equations 3 and 4 apply only for an isotropic medium. Dellinger and Etgen (1990) extend these operators for an anisotropic medium.

The Helmholtz separation operator is useful for distinguishing between P and S-wave amplitudes, but it is not reversible. The derivation of a reversible P-wave and S-wave

displacement decomposition by Zhang and McMechan (2010) is done in the wavenumber domain. They formulate a linear equation system based on characteristics of P and S particle displacements in an isotropic elastic medium. The solutions to this linear system in a two-dimensional medium are:

$$\begin{aligned} \tilde{u}_x^P &= k_x^2 \tilde{u}_x + k_x k_z \tilde{u}_z, & (5) \\ \tilde{u}_z^P &= k_z^2 \tilde{u}_z + k_z k_x \tilde{u}_x, & (6) \end{aligned}$$

and

$$\begin{aligned} \tilde{u}_x^S &= k_z^2 \tilde{u}_x - k_x k_z \tilde{u}_z, & (7) \\ \tilde{u}_z^S &= k_x^2 \tilde{u}_z - k_z k_x \tilde{u}_x, & (8) \end{aligned}$$

where $\tilde{u}_i$ are the spatial fourier transforms of the observed displacement fields in direction i, k_i are the wavenumbers, and $\tilde{u}_i^P$ and $\tilde{u}_i^S$ are the decomposed P and S displacements. It is important to note that the k in these equations is normalized by the absolute value of the wavenumber $|\mathbf{k}|$. This decomposition is reversible, since $u_i = u_i^P + u_i^S$.

Forward and adjoint isotropic elastic wave propagation

The isotropic elastic wave equation relates displacements to stresses via two elastic constants - the Lam*é* parameters λ and μ:

$$\nabla \left((\lambda + 2\mu) \nabla \cdot \mathbf{u} \right) - \nabla \times (\mu \nabla \times \mathbf{u}) + \mathbf{f} = \rho \ddot{\mathbf{u}}, \tag{9}$$

where $\mathbf{u}$ are the particle displacements in each dimension, $\mathbf{f}$ is the force function and ρ is medium density. An alternate formulation is:

$$\nabla \left((\lambda + \mu) \nabla \cdot \mathbf{u} \right) + \nabla \cdot (\mu \nabla \mathbf{u}) + \mathbf{f} = \rho \ddot{\mathbf{u}}. \tag{10}$$

From equation 10, the explicit form for a heterogeneous two dimensional medium can be expressed in a matrix-vector notation as:

$$\begin{bmatrix} \frac{1}{\rho}\partial_x \left((\lambda + 2\mu) \partial_x u_x + \lambda \partial_z u_z \right) + \frac{1}{\rho}\partial_z \left(\mu \left(\partial_x u_z + \partial_z u_x \right) \right) + \frac{1}{\rho} f_x \\ \frac{1}{\rho}\partial_z \left((\lambda + 2\mu) \partial_z u_z + \lambda \partial_x u_x \right) + \frac{1}{\rho}\partial_x \left(\mu \left(\partial_x u_z + \partial_z u_x \right) \right) + \frac{1}{\rho} f_z \end{bmatrix} = \begin{bmatrix} \partial_t^2 u_x \\ \partial_t^2 u_z \end{bmatrix}. \tag{11}$$

The forward elastic propagation operator can then be expressed as:

$$\mathbf{F} = \begin{bmatrix} \frac{1}{\rho}\partial_x \left(\lambda + 2\mu \right) \partial_x + \frac{1}{\rho}\partial_z \mu \partial_z - \partial_t^2 & \frac{1}{\rho}\partial_x \lambda \partial_z + \frac{1}{\rho}\partial_z \mu \partial_x \\ \frac{1}{\rho}\partial_z \lambda \partial_x + \frac{1}{\rho}\partial_x \mu \partial_z & \frac{1}{\rho}\partial_z \left(\lambda + 2\mu \right) \partial_z + \frac{1}{\rho}\partial_x \mu \partial_x - \partial_t^2 \end{bmatrix} \tag{12}$$

For a homogeneous medium, and using a Green's function to describe the energy propagation between any two locations $\mathbf{x} = (x, y, z)$ and $\mathbf{y} = (x, y, z)$, the equation takes the form:

$$\left((\lambda + \mu)\,\nabla\nabla + \mu\nabla^2 + \rho\omega^2\right) G\left(\mathbf{x}, \mathbf{y}, \omega\right) = -\delta\left(\mathbf{x} - \mathbf{y}\right) \tag{13}$$

The forward elastic propagation operator injects sources from a model into some location in the medium, and records the resulting wavefield at some other location:

$$d\left(\mathbf{x}, \omega; \mathbf{x}_s\right) = \sum_{\mathbf{y}} m\left(\mathbf{y}, \omega; \mathbf{x}_s\right) G\left(\mathbf{y}, \mathbf{x}, \omega\right), \tag{14}$$

where m is the model of injected sources at location $\mathbf{y}$ in the medium, and d are the recorded displacement fields $\mathbf{u}$ at location $\mathbf{x}$ in the medium. ω is angular frequency and $\mathbf{x}_s$ is the shot gather. In vector notation, this is expressed as

$$\mathbf{d} = \mathbf{F}\mathbf{m}, \tag{15}$$

where $\mathbf{F}$ is the forward operator. The adjoint operator injects the data from the same recording locations, and records the resulting wavefield at the model injection points:

$$\tilde{m}\left(\mathbf{y}, \omega; \mathbf{x}_s\right) = \sum_{\mathbf{x}} d\left(\mathbf{x}, \omega; \mathbf{x}_s\right) G^*\left(\mathbf{x}, \mathbf{y}, \omega\right), \tag{16}$$

which in vector notation is

$$\tilde{\mathbf{m}} = \mathbf{F}^*\mathbf{d}, \tag{17}$$

where $\mathbf{F}^*$ is the adjoint operator.

Inverting for the virtual source model

The inversion is done by least-square fitting of modeled geophone data to the recorded geophone data. The model is a virtual-source array, injected into some locations in a homogeneous medium. There should be some distance between these virtual-sources' locations and the receiver line where the displacement fields will be recorded. This distance is necessary for the recreated wavefield to form.

It is not possible to recreate the entire original wavefield that was recorded by the geophones without accurate knowledge of the acquisition geometry and the medium parameters. However, given a particular acquisition geometry, and some reasonable medium parameters, it is possible to recreate the original wavefield in the vicinity of the geophones. This wavefield will form as a result of the injection of the virtual sources, and will become more similar to the originally recorded wavefield at the receiver locations as it propagates toward them.

The inversion starts from a zero-value initial model of displacement sources $\mathbf{m}^0 = \mathbf{0}$. This initial model is injected as a displacement force function, at locations $\mathbf{x}$ in the medium. The energy is propagated with the forward elastic operator, and recorded at locations $\mathbf{y}$ in the medium (the receiver locations):

$$\mathbf{d} = \mathbf{F}\mathbf{m}^{\mathbf{0}}. \tag{18}$$

The residual that must be minimized is the difference between the calculated data d, and the observed data d_{obs}. The objective function is:

$$J(\mathbf{m}) = \|\mathbf{d} - \mathbf{d}_{\text{obs}}\|_2^2 = \|\mathbf{F}\mathbf{m} - \mathbf{d}_{\text{obs}}\|_2^2. \tag{19}$$

The model gradient is:

$$\mathbf{\Delta m} = \frac{\partial J}{\partial \mathbf{m}} = \left(\frac{\partial \mathbf{r}}{\partial \mathbf{m}}\right)^T \mathbf{r} = \mathbf{F}^* \mathbf{r}, \tag{20}$$

where $\mathbf{F}^*$ is the adjoint elastic propagation operator. The data gradient is the forward operator applied to the model gradient:

$$\mathbf{\Delta r} = \mathbf{F}\mathbf{\Delta m}. \tag{21}$$

The model gradient and the data residual are updated using an itterative minimization scheme (Claerbout and Fomel, 2011).

The end result of the inversion is the model $\mathbf{m}$ which, when injected at locations $\mathbf{x}$ in a homogeneous medium with the parameters used in the inversion, will generate recorded data at locations $\mathbf{y}$ that are equal to the observed data. Even if the medium parameters that were used to generate the wavefield were not a good approximation to the true parameters, this difference will be minor with regard to the wavefield at a close proximity to the receivers. Therefore, if the data residual is zero, we can assume that the wavefield near the receivers has likewise been reliably recreated.

The displacement field values are calculated for all time steps by running the forward propagation operator using the final virtual-source model as an input:

$$F \begin{bmatrix} -\frac{1}{\rho} m_x \\ -\frac{1}{\rho} m_z \end{bmatrix} = \begin{bmatrix} u_x \\ u_z \end{bmatrix}. \tag{22}$$

The separation operators in equations 3 and 4 are applied to the displacements fields (u_x, u_z) at the receiver locations, in order to separate P-wave amplitudes from S-wave amplitudes. These operators require calculating spatial derivatives of the displacement fields in the horizontal and vertical directions. We now have omnidirectional wavefield values near the receivers, therefore applying these operators to the recreated wavefields will result in an approximation to the actual P and S-waves that were recorded by the geophones. Furthermore, the displacement separation operators in equations 6-8 can be used to separate P displacements from S displacements.

INVERSION RESULTS

Synthetic land data 1

The first set of examples are of synthetic land data. The data are a single shot gather, forward modeled using an elastic isotropic migration code which utilizes the stress-displacement formulation, and a staggered-grid finite-difference scheme (Virieux, 1986). The medium parameters are shown in Figure 2. The source was a pressure source, located on the surface at the center of the model. A free boundary condition was used for the upper boundary, and absorbing boundaries were used for the bottom and the sides. The receivers were split-spread around the source, up to a maximum offset of $1800m$. The direct arrival has been muted from the synthetic data.

The recorded vertical and horizontal displacements are shown in Figures 3(a)-3(b). These represent the geophone data. The P-wave and S-wave shown in Figures 3(c)-3(d) are calculated from the displacement field with the Helmholtz separation operator in equations 3-4.

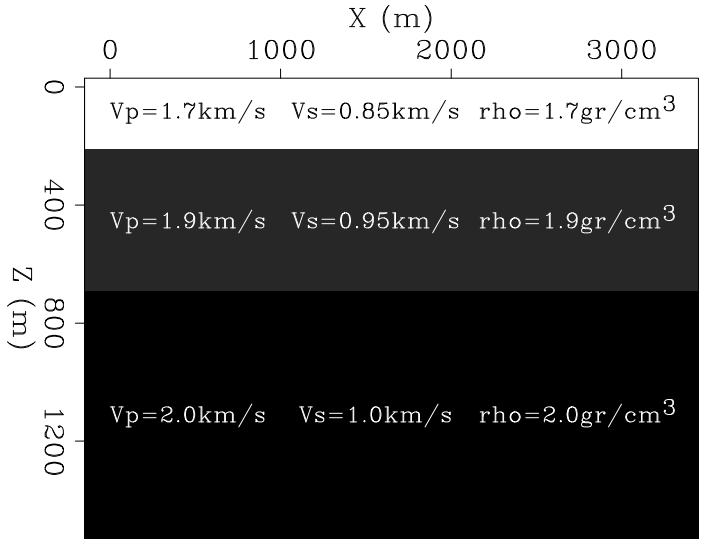


Figure 2: Medium parameters of synthetic land data. [**ER**] ohad1/. vp-land-a

Figure 4 shows the arrangement of the virtual sources and the receivers during the inversion. The virtual-source line is longer than the receiver line to better facilitate the modeling of high-wavenumber waves, which may be present at the edges of the recorded data. The medium itself is homogeneous, and surrounded by absorbing boundaries. The elastic propagation finite-difference scheme used in the inversion is different from the one used for generating the synthetic data. It utilizes a displacement-only formulation (as in equation 11), on a regular grid. The finite-difference approximation is 2_{nd} order in time and in space.

The vertical and horizontal virtual-source models resulting from the inversion are shown in Figures 5(a) and 5(b). When these source functions are injected as displacement sources into the homogeneous medium which was used in the inversion, the results are the reconstructed displacement data. If the inversion converged, then these reconstructed data will be equal to the recorded geophone data. The purpose of Figures 6(a)-6(h) is to show that the inversion managed to converge, even when the medium parameters used were different

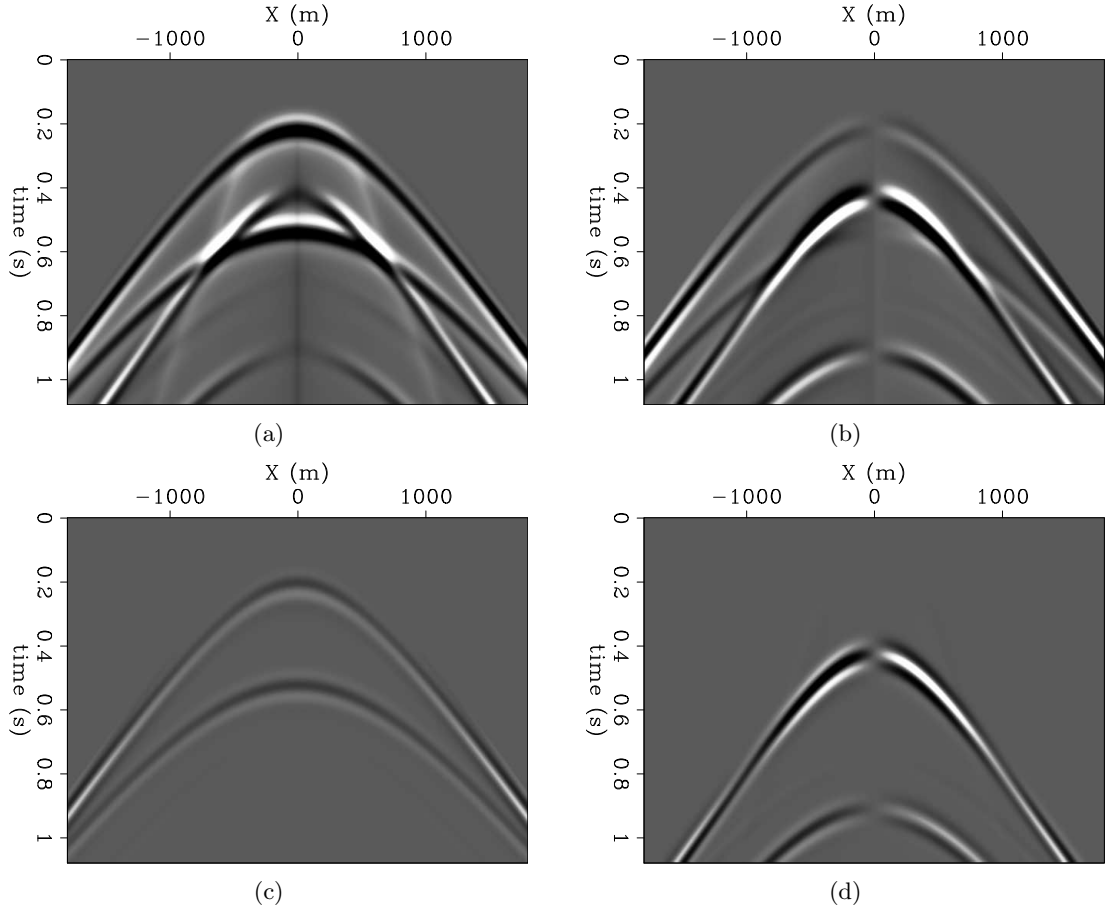


Figure 3: Synthetic land data. (a) vertical displacement, (b) horizontal displacement, (c) P-wave, (d) S-wave. [**ER**] ohad1/. lu1-0L,lu1-1L,lp1-0L,lp1-1L

from the true ones.

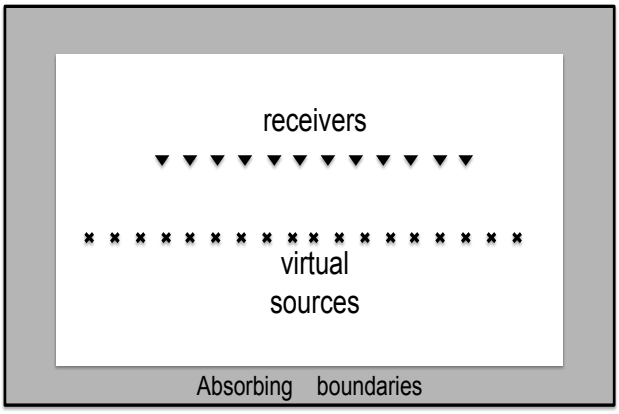


Figure 4: Qualitative sketch depicting the relative locations of the receivers and the virtual sources used in the inversion. The virtual source line has a greater horizontal extent than the receivers in order to enable convergence to a solution when high wavenumbers exist at the sides of the recorded data. [**NR**] ohad1/. src-rec-locations

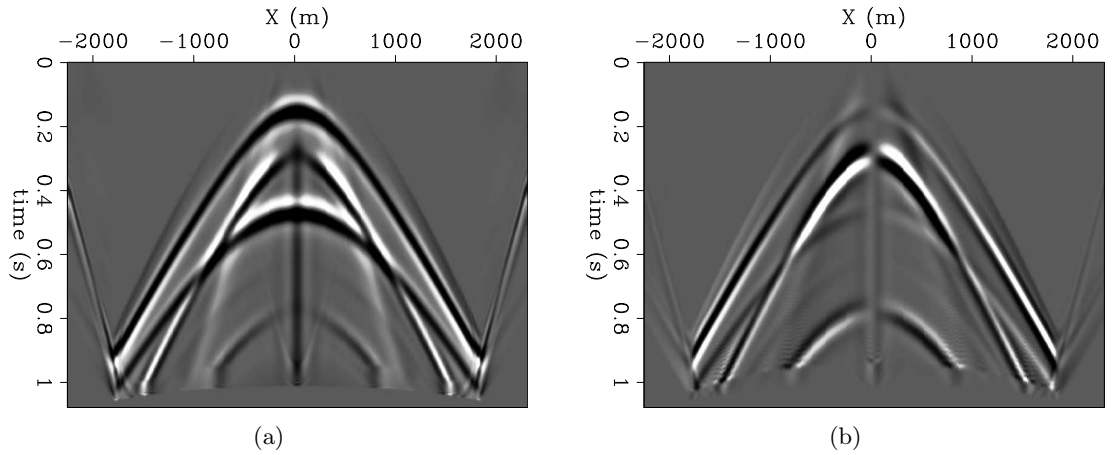


Figure 5: The virtual source functions that generate the observed land data. These "models" are the inversion results. (a) vertical displacement source functions, (b) horizontal displacement source functions. [**ER**] ohad1/. su1-0,su1-1

The P-wave velocity at the receiver level in the medium used for forward modeling was $1700\frac{m}{s}$, S-wave velocity was $850\frac{m}{s}$ and density was $1.7\frac{gr}{cm^3}$. Figures 6(a) and 6(b) are the forward modeled displacements. Figures 6(c) and 6(d) are the reconstructed displacements when using the correct medium parameters in the inversion. Figures 6(e) and 6(f) are the reconstructed displacements when using a homogeneous medium with $V_p = 1500\frac{m}{s}$, $V_s = 700\frac{m}{s}$ and $\rho = 1.5\frac{gr}{cm^3}$ (i.e. - about 15% too slow). Figures 6(g) and 6(h) are the reconstructed displacements when using a homogeneous medium with $V_p = 2000\frac{m}{s}$, $V_s = 1000\frac{m}{s}$ and $\rho = 2.0\frac{gr}{cm^3}$ (i.e. - about 15% too fast). Note that the displacements at the receivers have indeed been matched.

Figures 7(a)-8(d) are the result of applying the Helmholtz separation operator to the reconstructed displacement fields, to calculate the P and S-waves. The finite-difference approximation to the Helmholtz operator uses the displacement value at the receivers (which have been matched), and also the values one depth level below the receivers. These values are not matched by the inversion, and therefore unless the medium parameters are correct - they will not be an accurate representation of the true field values there. This will cause

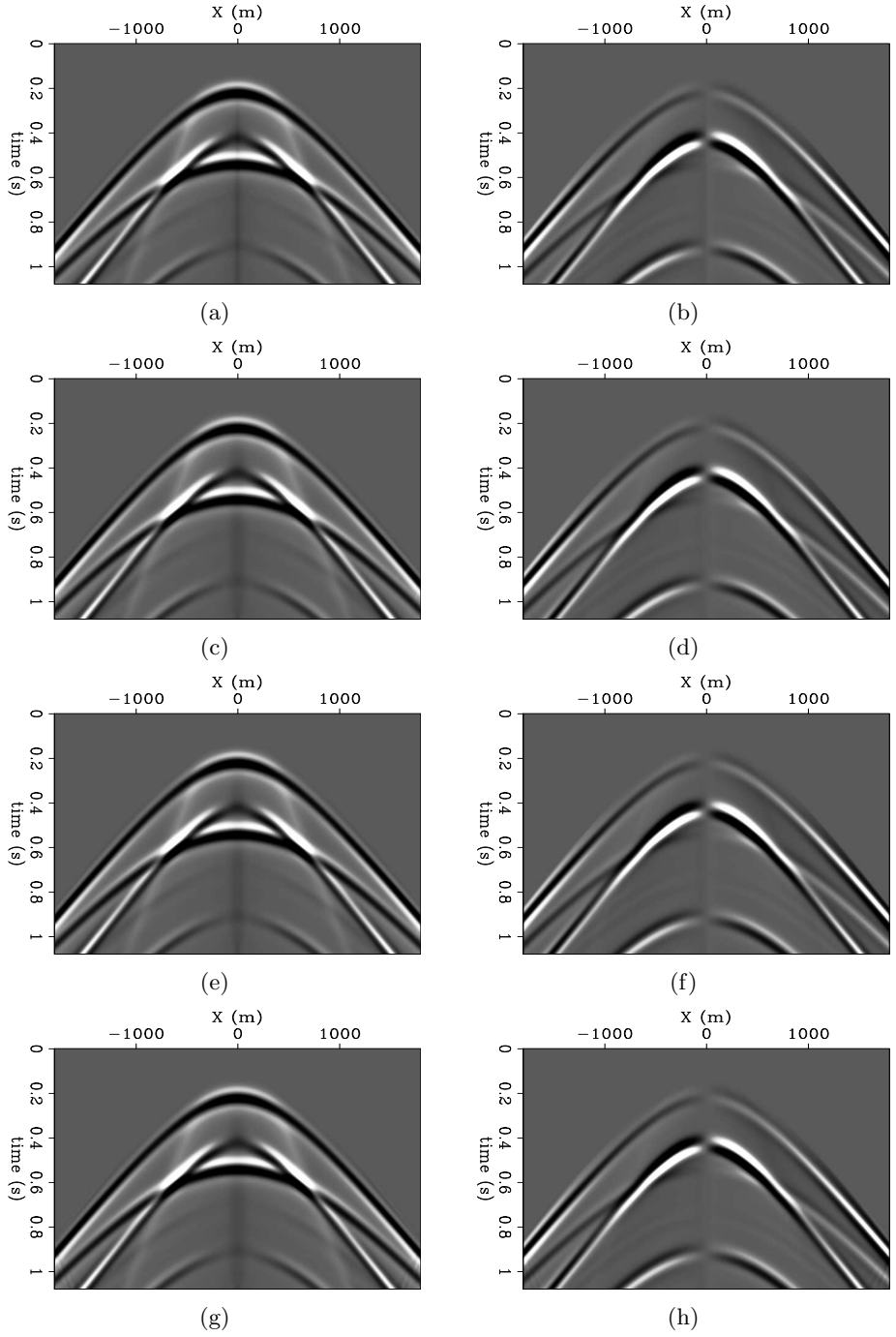


Figure 6: Synthetic and reconstructed displacements of land data. (a) synthetic vertical displacement, (b) synthetic horizontal displacement, (c) reconstructed vertical displacement with correct velocity, (d) reconstructed horizontal displacement with correct velocity, (e) reconstructed vertical displacement with 15% too slow velocity, (f) reconstructed horizontal displacement with 15% too slow velocity, (g) reconstructed vertical displacement with 15% too fast velocity, (h) reconstructed horizontal displacement with 15% too fast velocity. [**ER**] ohad1/. lu1-0,lu1-1,lu1-2,lu1-3,lu1-4,lu1-5,lu1-6,lu1-7

the vertical derivative to be inaccurate as well. The question then is: how detrimental are such inaccuracies to the P/S separation result?

Figure 7(a) is the forward modeled and separated synthetic P-wave recording. Figure 7(b) is the reconstructed P recording, when using the correct medium parameters in the inversion. Figure 7(c) is the reconstructed P when using a 15% too slow homogeneous medium, and Figure 7(d) is for a 15% too fast homogeneous medium. Figure 7(e) is a summation of the forward modeled and separated P and S recording, and is useful for estimating the separation quality for each scenario.

In Figure 7(b), the reconstructed P-wave is almost identical to the forward modeled P-wave, while the S-wave is barely visible. This indicates that the reconstructed displacement fields were a good approximation of the forward modeled displacement fields. In Figures 7(c) and 7(d) the medium parameters are incorrect, and the S-wave is more visible. However, the separation quality is still reasonable, when compared to Figure 7(e). This indicates that though the medium parameters were off by 15%, the displacement fields were a good approximation to the forward modeled fields one depth step below the receivers (as well as at the receivers themselves). Therefore, the vertical derivatives of the displacement fields reasonably approximate the true derivatives, and the P separation (equation 3) is effective.

Figure 8(a) is the synthetic S-wave recording. Figure 8(b) is the reconstructed S recording, when using the correct medium parameters in the inversion. Figure 8(c) is the reconstructed S when using a 15% too slow homogeneous medium, and Figure 8(d) is for a 15% too fast homogeneous medium. Figure 8(e) is the sum of the forward modeled P and S. As for the previous set of figures, we can see that the separation quality does decrease as the medium parameters vary from the true ones, but the separation is still of reasonable quality even when the parameters are off by a large percentage.

There is an additional effect of the inversion modeling process, which is visible in this set of figures and particularly in Figure 8(b). As a result of the edges of the receiver line, a diffraction is generated at both ends during the adjoint propagation, and this diffraction is recorded at the virtual-source line. It can be seen in the virtual-source functions, in Figures 5(a)-5(b). The recording of this diffraction event is then repropagated during the forward propagation stage of the inversion, generating an additional diffraction event off the edges of the virtual-source line. This diffraction event was not part of the original wavefield. Since it arrives at the same time as the first P-wave, it is not visible in Figures 7(b)-7(d), but it is visible in Figures 8(b)-8(d), even when using the correct velocity for the inversion.

Synthetic land data 2

The next set of examples are also of synthetic land data, but they represent a scenario where the velocity at the recording surface varies laterally. Three shot gathers were generated in three separate homogeneous mediums, with differing medium parameters. For each gather there was a different random reflection series. The gathers were then summed to simulate a variable surface velocity.

The forward modeling medium parameters were:

1. Gather 1: $v_{p1} = 2000\frac{\mathrm{m}}{\mathrm{s}}, \qquad v_{s1} = 1000\frac{\mathrm{m}}{\mathrm{s}}, \qquad \rho_1 = 2.0\frac{gr}{cm^3}$

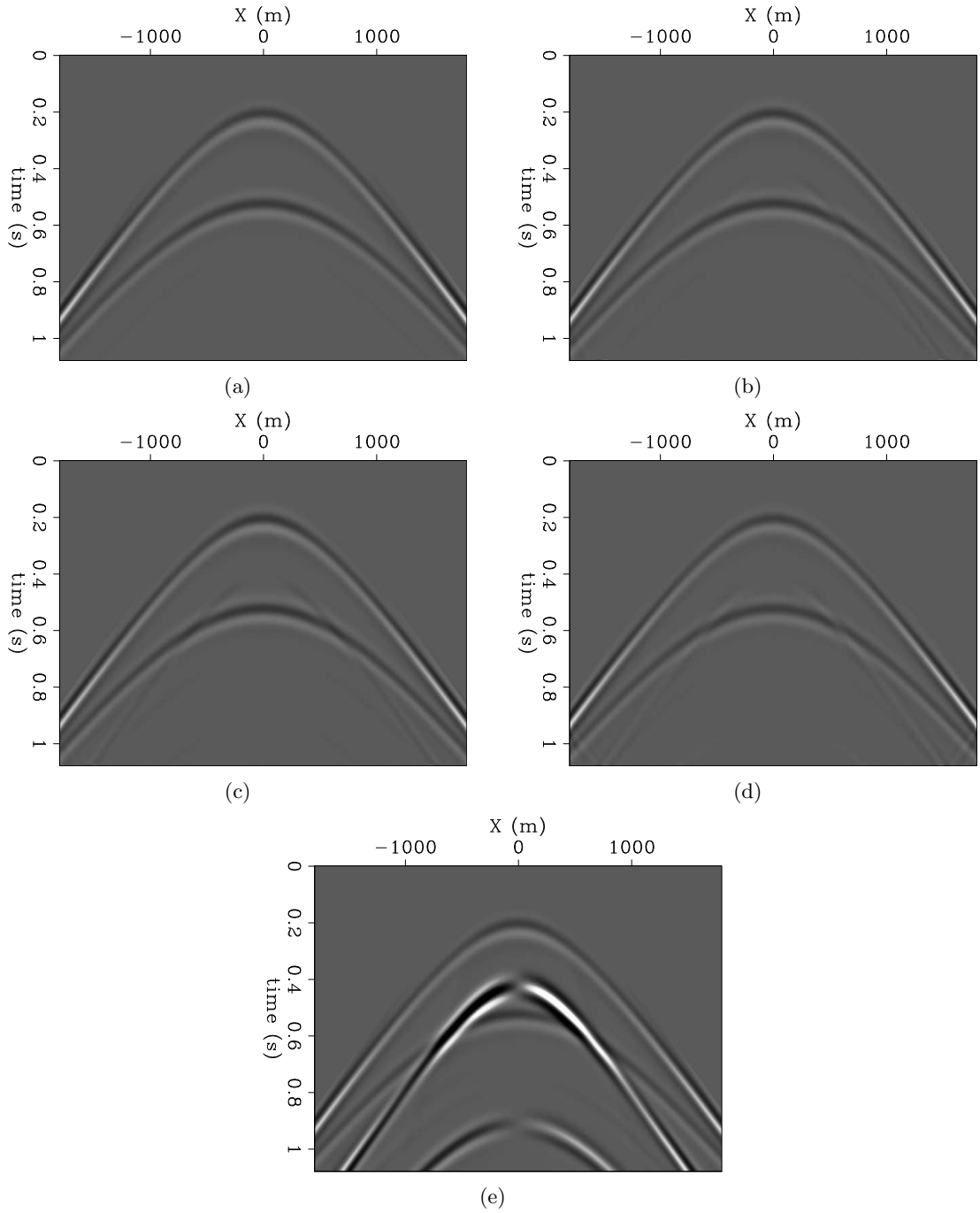


Figure 7: Synthetic and reconstructed P-wave of land data. (a) Synthetic P, (b) reconstructed P with correct velocity, (c) reconstructed P with 15% too slow velocity, (d) reconstructed P with 15% too fast velocity, (e) sum of synthetic P and S-waves. All figures are identically clipped. [**ER**] ohad1/. lp1-0,lp1-2,lp1-4,lp1-6,lp1-8L

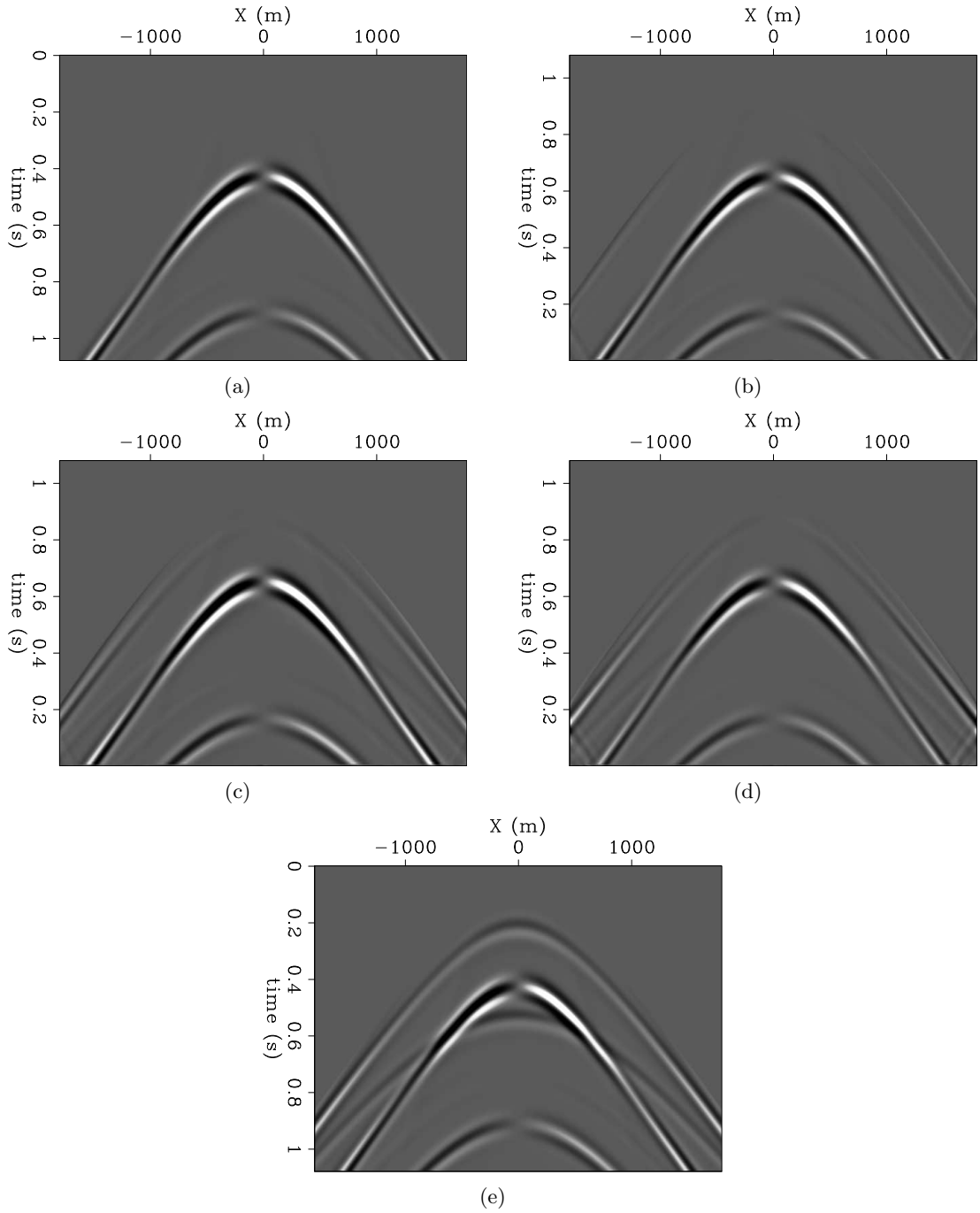


Figure 8: Synthetic and reconstructed S-wave of land data. (a) Synthetic S, (b) reconstructed S with correct velocity, (c) reconstructed S with 15% too slow velocity, (d) reconstructed S with 15% too fast velocity, (e) sum of synthetic P and S-waves. All figures are identically clipped. [**ER**] ohad1/. lp1-1,lp1-3,lp1-5,lp1-7,lp1-8

2. Gather 2: $v_{p2} = 1800\frac{\text{m}}{\text{s}}$, $v_{s2} = 900\frac{\text{m}}{\text{s}}$, $\rho_2 = 1.8\frac{gr}{cm^3}$
3. Gather 3: $v_{p3} = 1600\frac{\text{m}}{\text{s}}$, $v_{s3} = 800\frac{\text{m}}{\text{s}}$, $\rho_3 = 1.6\frac{gr}{cm^3}$

Figures 9(a) and 9(b) are the vertical and horizontal displacements. Figures 9(c) and 9(d) are the P and S-waves calculated by applying the Helmholtz separation operators. Three different reflection groups can be discerned in the synthetic data. The left group corresponds to the fastest medium, the central group to the intermediate medium, and the right group to the slowest medium.

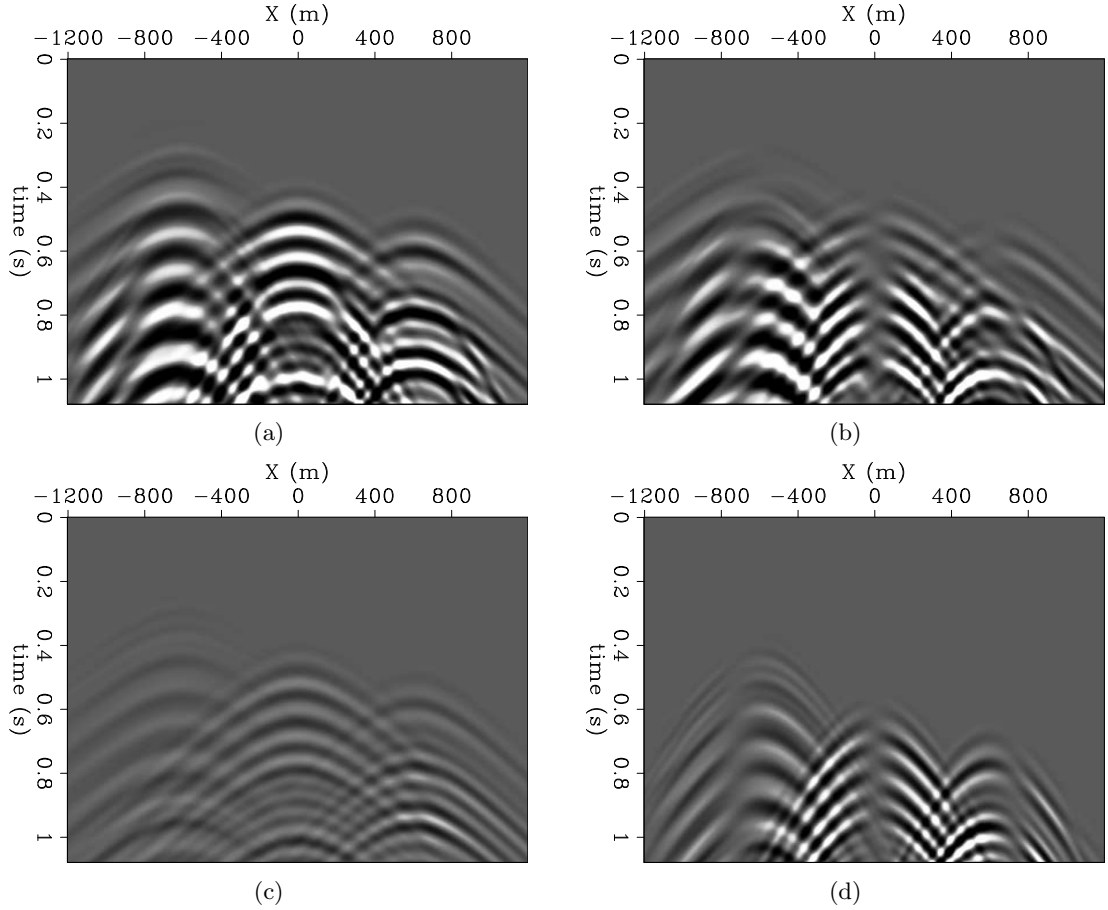


Figure 9: Synthetic land data. (a) vertical displacement, (b) horizontal displacement, (c) P-wave, (d) S-wave. [**ER**] ohad1/. lu2-0,lu2-1,lp2-0L,lp2-1L

There were three inversion runs, each one with different homogeneous medium parameter sets:

1. Inversion 1 (intermediate): $v_{p1} = 1700\frac{\text{m}}{\text{s}}$, $v_{s1} = 850\frac{\text{m}}{\text{s}}$, $\rho_1 = 1.7\frac{gr}{cm^3}$
2. Inversion 2 (slow): $v_{p2} = 1500\frac{\text{m}}{\text{s}}$, $v_{s2} = 700\frac{\text{m}}{\text{s}}$, $\rho_2 = 1.5\frac{gr}{cm^3}$
3. Inversion 3 (fast): $v_{p3} = 2000\frac{\text{m}}{\text{s}}$, $v_{s3} = 1000\frac{\text{m}}{\text{s}}$, $\rho_3 = 2.0\frac{gr}{cm^3}$

The medium parameters of inversion 3 match those of the fast events, but are different from the parameters used to generate the intermediate and the slow events. As for the

medium parameters of inversions 1 and 2, they match none of the parameters used to generate the datasets. For inversion 2, the velocity error is up to 33% (compared to the fast arrivals). The inversions were run until a good match to the displacement data was achieved.

Figures 10(a)-10(e) show the reconstructed P-waves. Figure 10(a) is the synthetic P recording. Figure 10(b) is the P reconstruction calculated from the displacements of the inversion with parameter set 1 (intermediate). Figure 10(c) is the P reconstruction from the inversion using parameter set 2 (slow), and Figure 10(d) is from parameter set 3 (fast). Figure 10(e) is the summation of the synthetic P and S recordings, and is a visual aid to estimate the separation quality for each scenario.

Since the intermediate parameter set 1 has the least error in comparison to the media with which the forward modeled data were generated, the reconstructed P-wave is most accurate for it, as seen in Figure 10(b). However, the other two P-wave reconstructions show reasonable separation quality, despite the large error in medium parameters. The same can be said for the next set of S-wave separation results in Figures 11(a)-11(e). The best separation is where the medium parameter error is minimal (Figure 11(b)). However, even where the error is large (Figures 11(c) and 11(d)), the separation results, though not being great, are still reasonably good compared to the synthetic data.

Synthetic OBS data

The next result set is of OBS data, and its purpose is to show the problem that the inversion method described above has when both upgoing and downgoing data are present. For that same purpose, the synthetic data were forward modeled without shear waves (i.e. - $V_s = 0$).

The medium parameters used for forward modeling are shown in Figure 12. The top layer is water. Shear velocity was set to zero everywhere. The source was a pressure source, located on the water surface, at the center of the model. The receivers were located at the top of the second layer, which represents sediment. The upper boundary was free, and the sides and bottom of the model had absorbing boundaries. The direct arrival was muted.

The inversion was run using the correct P-wave velocity and density (those of the sediment in which the receivers were placed), but also with shear-wave velocity set to $850\frac{m}{s}$. The virtual-sources were placed below the receivers, as shown in Figure 4.

The synthetic vertical and horizontal displacement data are shown in Figures 13(a) and 13(b). Of the five events visible, the first, second, fourth and fifth arrivals are downgoing waves, reflected off the water surface. The only upgoing wave is the third arrival.

The reconstructed displacements are in Figures 13(c) and 13(d). These figures show that the objective function has indeed been minimized, and the displacement data have been matched.

Figure 14(a) shows the P-wave recording, obtained by the Helmholtz separation operator applied to the synthetic displacement fields. Figure 14(b) is a validation that there are no S-waves present in the synthetic data. Figure 14(c) is the reconstructed P-waves calculated by the Helmholtz P-wave separation operator applied to the reconstructed displacements. Note that all the downgoing events have an opposite polarity compared to the synthetic P section, while the only P event that was reconstructed with the correct polarity is the

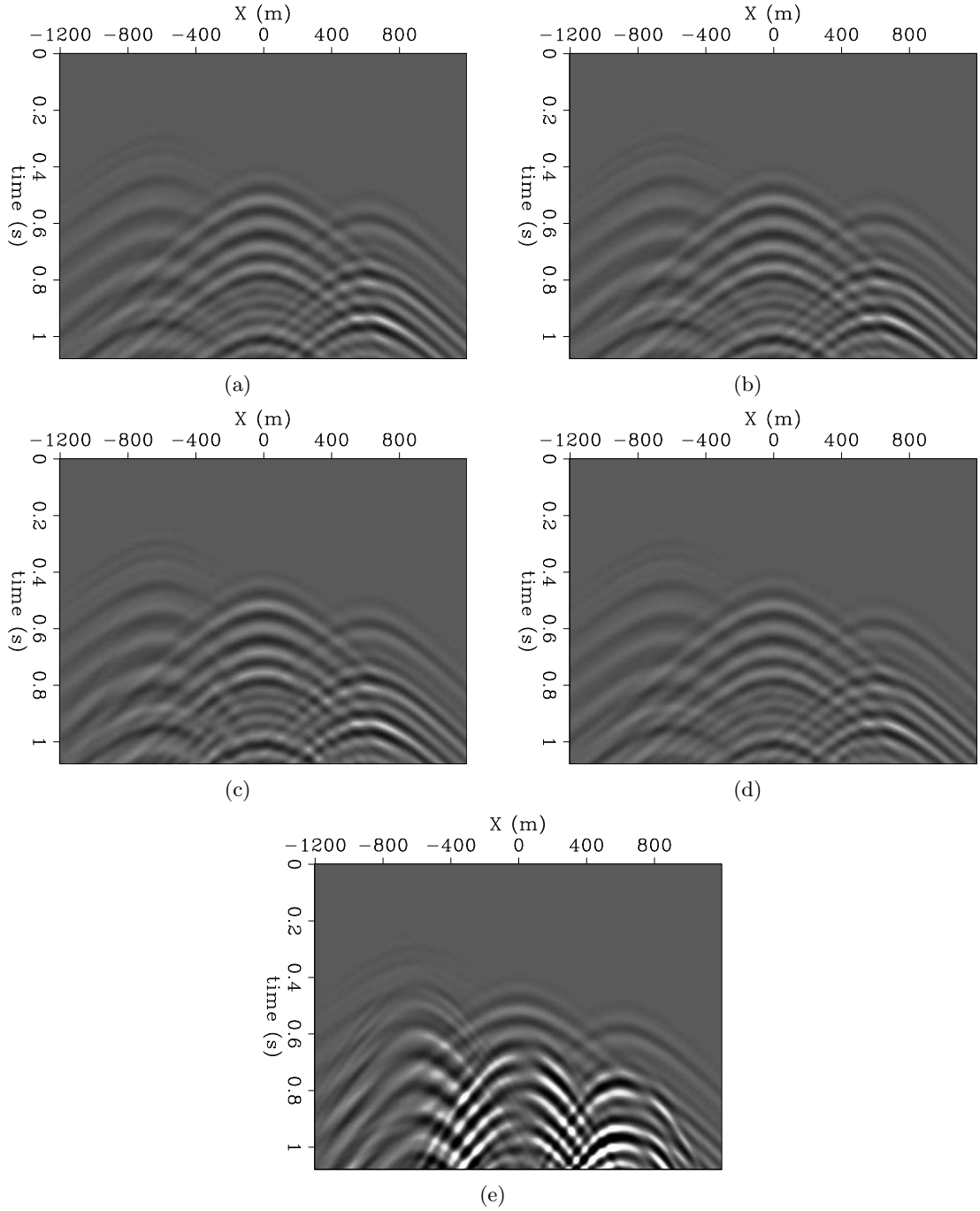


Figure 10: Synthetic and reconstructed P-wave of land data. (a) Synthetic P, (b) reconstructed P with intermediate medium parameters, (c) reconstructed P with slow parameters, (d) reconstructed P with fast parameters, (e) sum of synthetic P and S-waves. All figures are identically clipped. [**ER**] ohad1/. lp2-0,lp2-2,lp2-4,lp2-6,lp2-8L

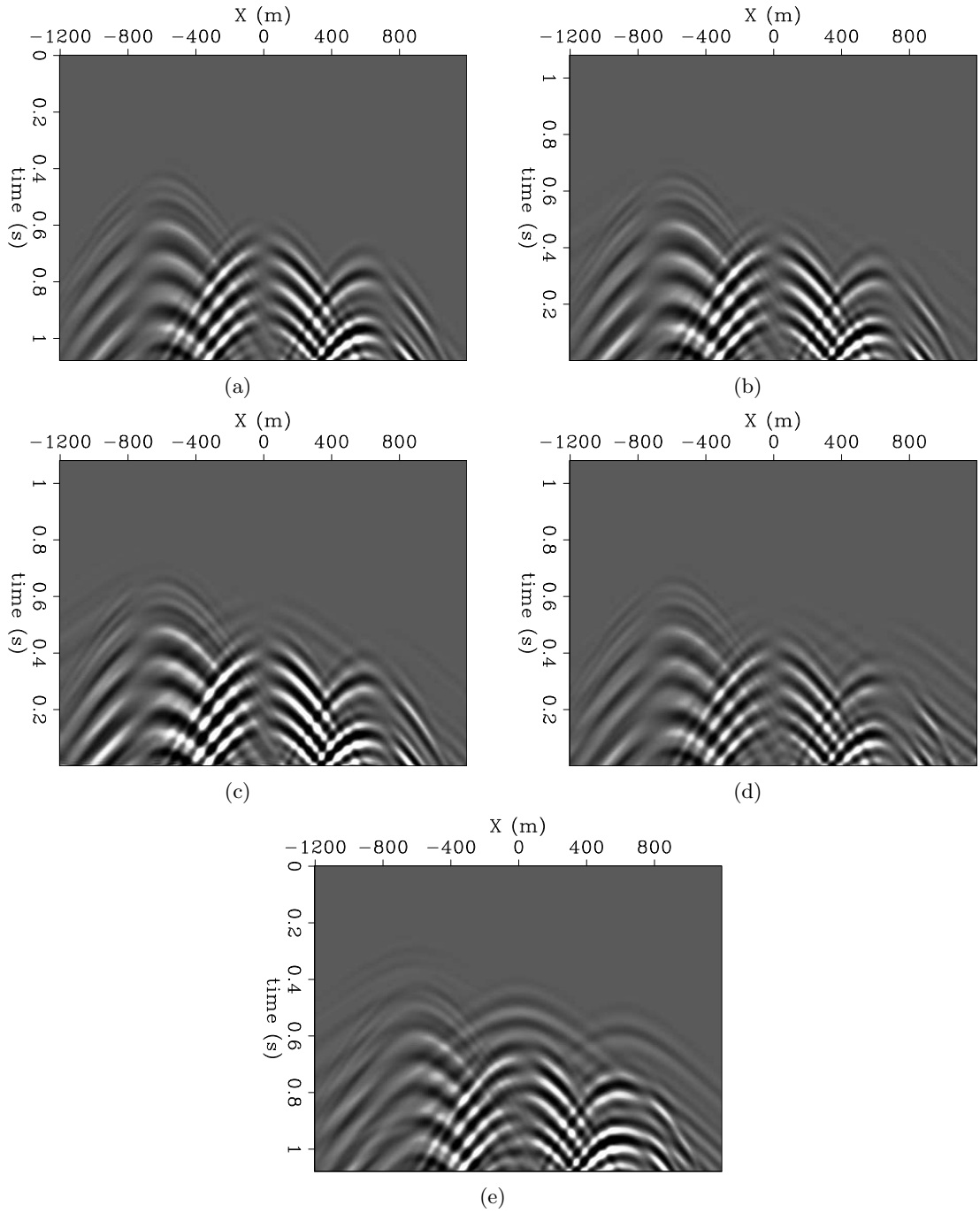


Figure 11: Synthetic and reconstructed S-wave of land data. (a) Synthetic S, (b) reconstructed S with intermediate medium parameters, (c) reconstructed S with slow parameters, (d) reconstructed S with fast parameters, (e) sum of synthetic P and S-waves. All figures are identically clipped. [**ER**] ohad1/. lp2-1,lp2-3,lp2-5,lp2-7,lp2-8

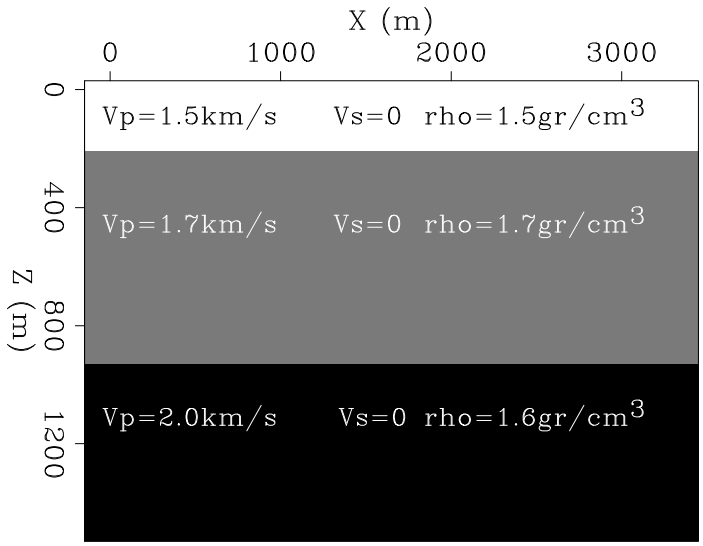


Figure 12: Medium parameters of synthetic OBS data. Shear velocity was zero everywhere. [**ER**] ohad1/. vp-obn-a

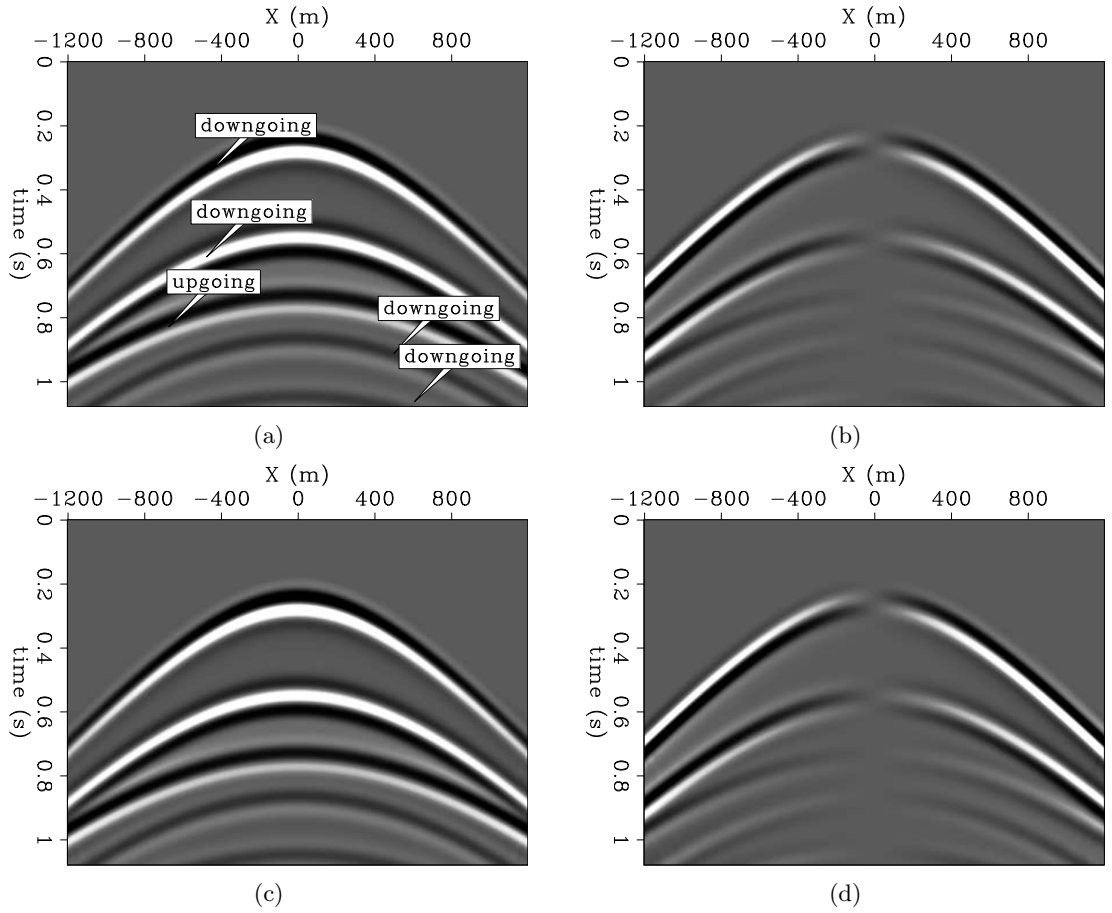


Figure 13: Synthetic and reconstructed OBS data. (a) synthetic vertical displacement, (b) synthetic horizontal displacement, (c) reconstructed vertical displacement, (d) reconstructed horizontal displacement. [**ER**] ohad1/. ou1-0-a,ou1-1,ou1-2,ou1-3

upgoing arrival (the third arrival). The amplitude of the reconstructed third arrival is stronger than its amplitude in the synthetic P section.

Figure 14(d) is the Helmholtz S-wave separation operator applied to the reconstructed displacements. It shows that for every P arrival, a shear arrival was reconstructed as well, despite the fact that there were no shear waves in the synthetic wavefield.

The reason for the inaccuracy of the P and S-wave reconstruction in this case is that as a result of the virtual-sources being below the receivers, only upgoing waves could be modeled by the inversion. The inversion does generate a match to the displacement data, but the displacements are unconstrained above and below the receivers. Therefore the inversion is free to explain the downgoing P-waves as a combination of (incorrect) upgoing P-waves and upgoing S-waves. A graphical explanation of this feature is shown in Figure 15. The grey area represents the possible downgoing wave solutions which were not enabled by this inversion configuration. Instead, the incorrect upgoing wave modes (which still match the displacement data at the receivers and are therefore a possible solution) are used to explain the downgoing energy.

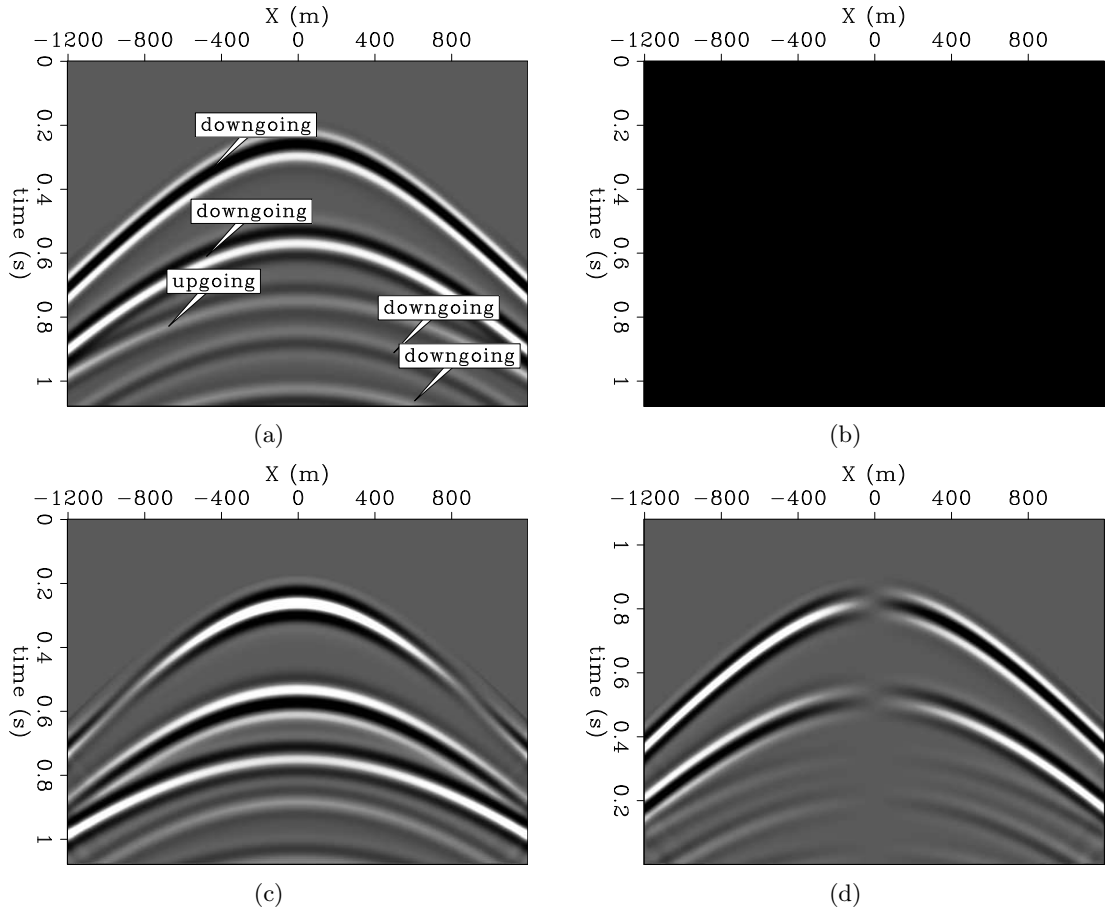


Figure 14: Synthetic and reconstructed OBS data. (a) synthetic P-wave, (b) synthetic S-wave, (c) reconstructed P-wave, (d) reconstructed S-wave. [**ER**] ohad1/. op1-0-a,op1-1,op1-2,op1-3

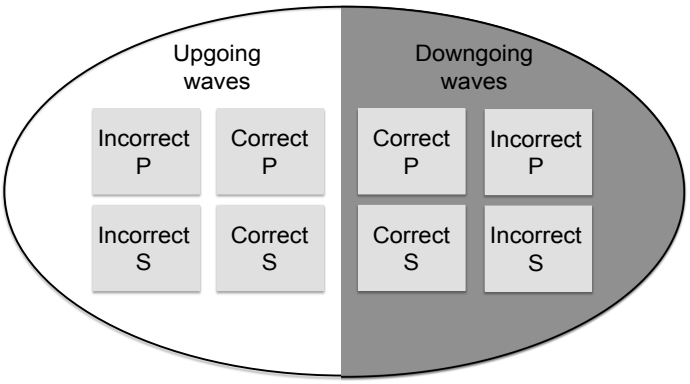


Figure 15: Representation of the solution space of the inversion. If the data contain both upgoing and downgoing waves, but the inversion is unable to reconstruct the downgoing waves, then incorrect P and S-waves will be used to match the observed geophone data. [**NR**] ohad1/. PS-sol-space

DISCUSSION AND CONCLUSION

The more the homogeneous medium parameters used in the inversion differ from the true medium parameters in which the receivers were placed, the less accurate is the reconstructed wavefield in the proximity of the receivers. This in turn means that the spatial derivatives will differ from the true derivatives, and thus the separation of P-waves from S-waves will be less reliable. However, given the typical length of seismic waves in exploration geophysics, and the spatial oversampling used in finite-difference modeling, the inaccuracy in medium parameters for which this separation method can still function reliably is quite large.

The locations of virtual sources in relation to the receivers determine what wavenumbers can be reconstructed. In order to accurately recreate all wavenumbers in the data, the virtual-sources must be placed in a configuration that enables the generation of those wavenumbers. For land data, the implication is that virtual-sources should be placed below the receivers. For OBS data, this should mean that virtual-sources must be placed above and below the receivers. However, such a configuration drastically increases the null-space of the inversion, and the results (not shown here) are very noisy and inaccurately reconstructed P and S sections.

I am currently investigating two possible avenues for implementing the inversion successfully on OBS data. The first possibility is to run the inversion twice: once with virtual-sources below the receivers, and once with virtual-sources above. The shear waves propagate upward only, and will be reconstructed by the upgoing inversion. As can be seen from Figures 14(a) and 14(d), the downgoing P-waves are reconstructed with opposite polarity when the virtual-sources are below the receivers. Likewise, the upgoing P-waves will be reconstructed with an opposite polarity when the virtual-sources are above the receivers. By comparing the polarity reversals to the observed hydrophone data, it is possible to gauge which P events are improperly reconstructed. I am searching for a way to introduce the hydrophone data as a constraint to the objective function, so that the polarity reversal attribute can be used to guide the inversion.

The second possibility is to have the virtual-sources below the receivers, but use a reflecting upper boundary, thereby recreating the downgoing waves generated by reflections off the sea surface. This configuration is not trivial to implement, since there is an important condition that has to be met: The free surface should not generate a mode conversion. If it does, downgoing shear waves will be present in the reconstructed displacement fields, and such waves definitely do not exist in OBS data.

One way of preventing a mode conversion from the free upper boundary is to use a medium with no shear strength. The possible configuration then is a line of receivers placed in a medium with parameters approximating those of the sea bottom. Above these receivers the medium will be water. Effectively, a homogeneous elastic medium underlying a homogeneous acoustic medium. The added value such a configuration may give is not only in the enabling of reconstruction of downgoing P-waves, but also in the possibility of reconstructing the surface waves.

ACKNOWLEDGMENTS

I wish to thank my colleagues, Mandy, Elita, Ali, and Sjoerd for the hours (weeks, actually) of fruitful discussions, and my advisors, Prof. Ronen and Prof. Biondi, for their helpful suggestions and patience.

REFERENCES

Amundsen, L., 1993, Wavenumber-based filtering of marine point-source data: Geophysics, **58**, 1497–150.

Barr, F. J. and J. I. Sanders, 1989, Attenuation of water-column reverberations using pressure and velocity detectors in a water-bottom cable: SEG expanded abstracts, **8**, 653–656.

Claerbout, J. F. and S. Fomel, 2011, Image estimation by example: Geophysical soundings image construction.

Dankbaar, J. W. M., 1985, Separation of P- and S-waves: Geophysical Prospecting, **33**, 970–986.

Dash, R., G. Spence, R. Hyndman, S. Grion, Y. Wang, and S. Ronen, 2009, Wide-area imaging from OBS multiples: Geophysics, **74**, Q41–Q47.

Dellinger, J. and J. Etgen, 1990, Wave-field separation in two-dimensional anisotropic media: Geophysics, **55**, 914–919.

Paffenholz, J., P. Docherty, R. Shurtleff, and D. Hays, 2006, Shear wave noise on OBS Vz data - part ii: Elastic modeling of scatterers in the seabed: EAGE extended abstracts, 68th conference and exhibition, **A072**.

Robertsson, J. O. A. and E. Muyzert, 1999, Wavefield separation using a volume distribution of three component recordings: Geophys.. Res. Lett., **26**, 2821–2824.

Schalkwijk, K. M., C. P. A. Wapenaar, and D. J. Verschuur, 2003, Adaptive decomposition of multicomponent ocean-bottom seismic data into downgoing and upgoing p- and s-waves: Geophysics, **68**, 1091–1102.

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

Wapenaar, C. P. A., P. Herrmann, D. J. Verschuur, and A. J. Berkhout, 1990, Decomposition of multicomponent seismic data into primary P- and S-wave responses: Geophysical Prospecting, **38**, 663–661.

Zhang, Q. and G. McMechan, 2010, 2d and 3d elastic wavefield vector decomposition in the wavenumber domain for vti media: Geophysics, **75**, D13–D26.

Zhou, Y., C. Kumar, and I. Ahmed, 2011, Ocean bottom seismic noise attenuation using local attribute matching filter: SEG Expanded Abstracts, 81st Annual Meeting, 3586–3590.

Edge-preserving smoothing for segmentation of seismic images

Adam Halpert

ABSTRACT

In many disciplines, pre-processing images by smoothing is common prior to automatic image segmentation; however, traditional smoothing blurs boundaries and can be counterproductive, especially for seismic images. Here, an edge-preserving smoothing technique based on directional maximum homogeneity is introduced for 3D seismic images, and tested on both synthetic and field data. Edge-preserving smoothing is shown to both decrease the level of noise in an image, and improve the accuracy of automatic segmentation results. In addition, a "hybrid" smoothing technique blending traditional and edge-preserving smoothing combines the advantages of each and produces encouraging results.

INTRODUCTION

Because it can reduce the time and effort required for human-intensive image interpretation tasks, automatic image segmentation is a tool employed in a variety of disciplines – for example, medical imaging, photo editing/image processing, and seismic imaging and interpretation. A particular area of interest involves the pre-processing of images prior to automatic segmentation (Zahedi and Thomas, 1993). Here, I investigate the usefulness of an edge-preserving smoothing (EPS) technique for segmentation of seismic images.

A primary use for automatic seismic image segmentation is for identification and location of complex subsurface salt bodies, a task that is extremely time-consuming when undertaken manaully. In the examples here, the pairwise region comparison (PRC) segmentation algorithm (Felzenszwalb and Huttenlocher, 2004) is employed because it is designed to operate extremely efficiently, even when extended to three dimensions and adapted for seismic data (Halpert et al., 2010). However, like any segmentation algorithm, its accuracy can suffer, especially where boundaries are discontinuous or chaotic (for example, see Figure 10(a)). In such cases, smoothing the image before the segmentation procedure can reduce unwanted noise and improve performance of many image processing algorithms (Zahedi and Thomas, 1993).

For seismic images, naive box or Gaussian smoothing has clear disadvantages. When segmenting seismic images, clear and sharp boundaries are preferable for an accurate result; simple smoothing tends to blur these boundaries, or even render them uninterpretable if two reflectors are very close together (see Figure 3(a)). A variety of "smarter" smoothing or noise-reduction approaches for seismic data have been proposed, including inversion-based techniques like PEFs (Claerbout, 2005; Guitton, 2005), structure or dip-oriented filtering (Fehmers and Hocker, 2003), or bilateral filtering (Hale, 2011). Unfortunately, these algorithms require computationally-intensive inversions and/or solutions to differential

equations like the diffusion equation, or prior information in the form of dip or structure interpretations. Therefore, a cheap, efficient smoothing algorithm that preserves sharp boundaries would represent a useful pre-processing step for seismic image segmentation.

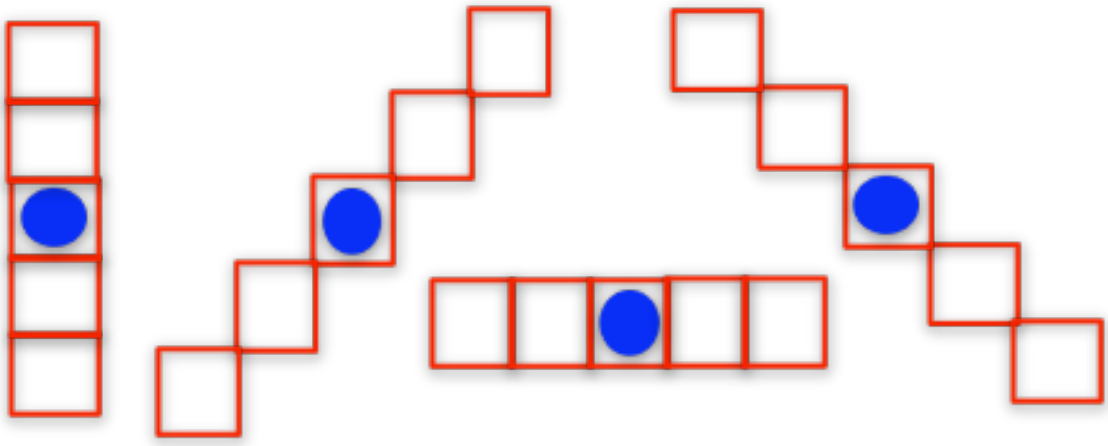


Figure 1: A set of 1D bar masks used to determine the most homogeneous orientation around a central pixel on a 2D image. For 3D images, additional masks would extend out of the page. [**NR**] adam1/. masks

EPS TECHNIQUE

Edge-preserving smoothing is a common goal in many image processing disciplines. An early form of EPS was median smoothing (Tukey, 1971), in which a central pixel is assigned the median pixel value from a neighborhood around it. Later, a method aiming for increased directional accuracy in EPS was proposed by Tomita and Tsuji (1977), in which the average value of the most homogeneous of several neighborhoods extending from the pixel was calculated. More recently, Zahedi and Thomas (1993) introduced the concept of "bar masks," one-dimensional vectors of pixels that extend in different directions from the central pixel (Figure 1). This approach further increases directional discrimination in "maximum homogeneity" filtering approaches. Furthermore, Zahedi and Thomas (1993) assigned the median of the most homogeneous bar mask to the central pixel, rather than the average value; this improved the preservation of sharp edges in an image, and makes it attractive for use in seismic images.

A very similar approach has already been proposed for use on three-dimensional seismic images. AlBinHassan et al. (2006) constructed many 3D neighborhoods around a central pixel, and used the average value of the most homogeneous block. For this method and that of Zahedi and Thomas (1993), maximum homogeneity is determined by calculating the variance of the individual bar masks or 3D blocks:

$$\sigma = \left[\frac{1}{n} \sum_{i=1}^{n} (d_i - \overline{d})^2 \right]^{1/2}, \tag{1}$$

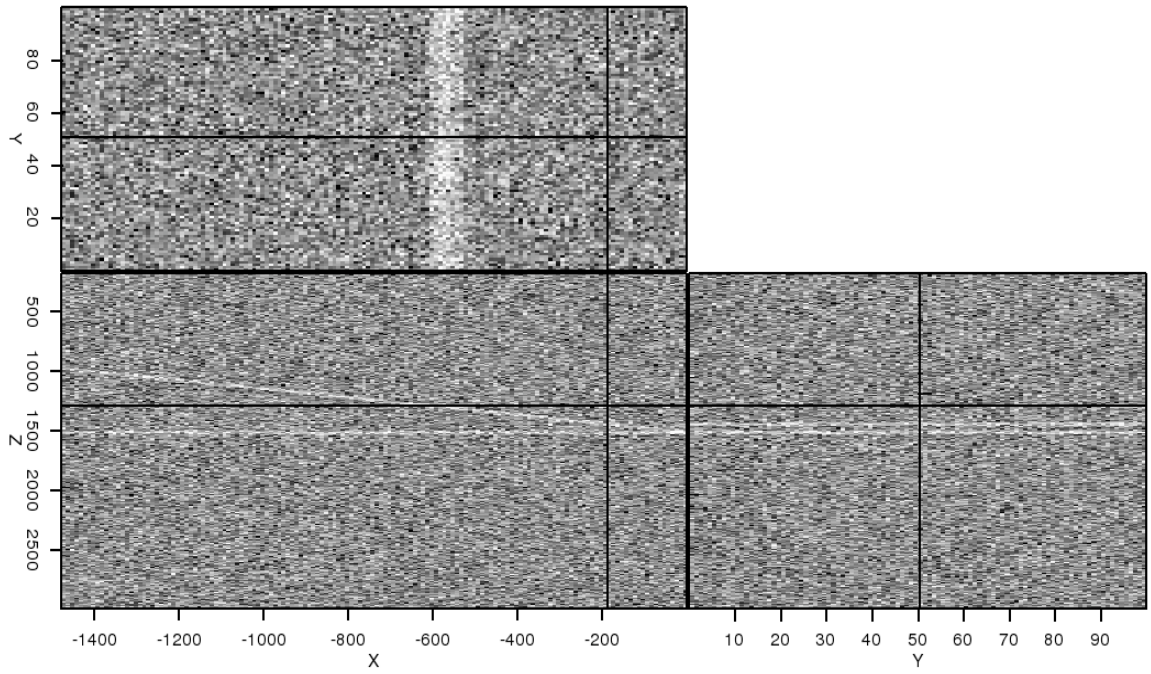


Figure 2: A simple 2.5-D synthetic consisting of two planar reflectors and contaminated with random noise. The high level of noise makes it difficult to distinguish the two reflectors at the indicated location. [**ER**] adam1/. syn-noise

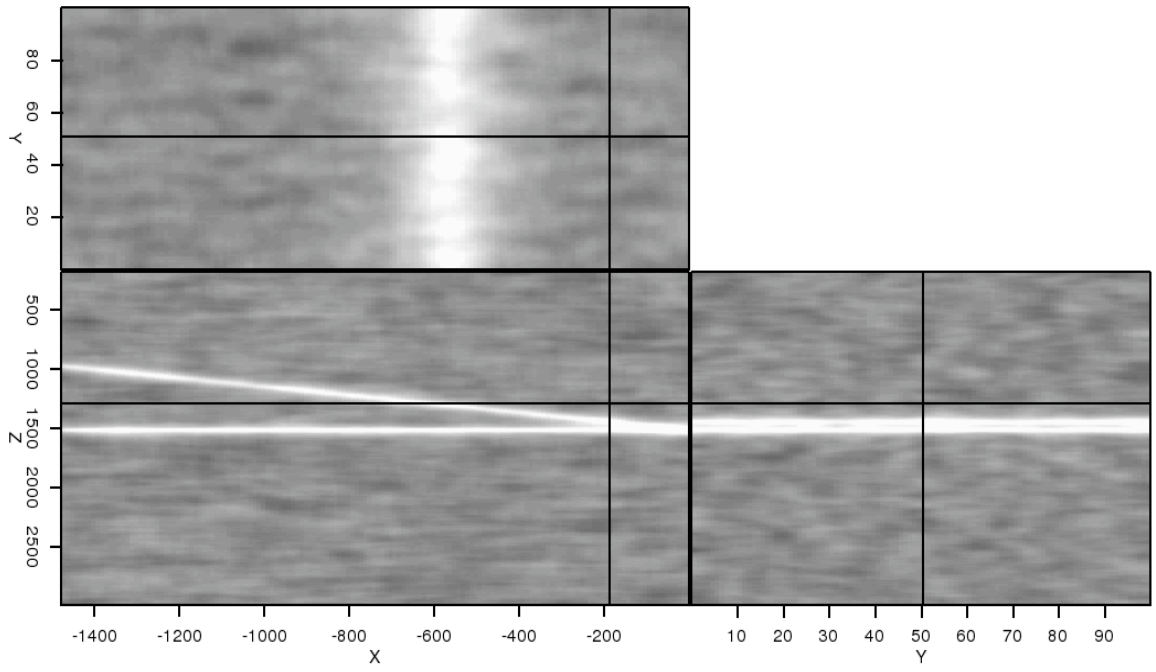
where n is the number of pixels in the bar mask or block d, and $\overline{d}$ is the average value of those pixels. The mask or block with the smallest variance has the greatest homogeneity.

The method presented here differs from this approach by incorporating two ideas from Zahedi and Thomas (1993). First, nine 1D bar masks, extending from a central pixel along each axis and diagonal of a cube enclosing the pixel, are used instead of the 3D blocks. This greatly simplifies the computational complexity of the algorithm, while at the same time preserving sharp boundaries with arbitrary orientations. Second, the median value of the bar mask with maximum homogeneity is used, rather than the average value. Again, this enhances the edge-preserving characteristics of the filter.

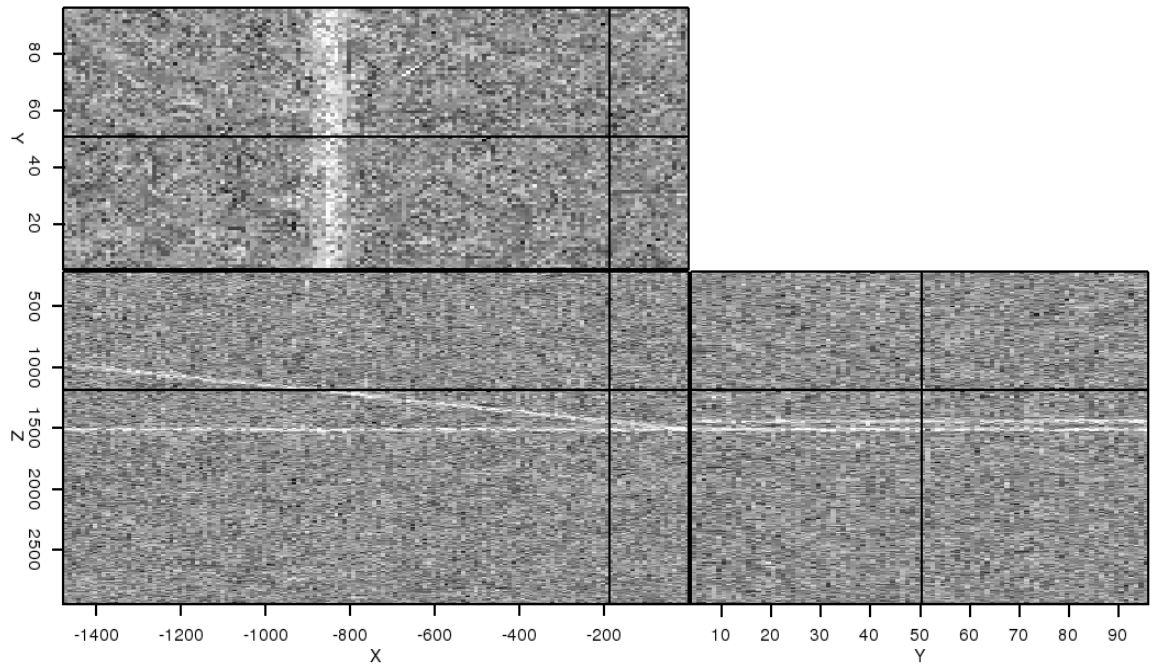
Synthetic example

The effectiveness of this method is best demonstrated through examples, and comparison to the traditional smoothing approach. Figure 2 is a noisy 2.5D synthetic featuring a dipping reflector pinching out onto a flat reflector. Because of the noise, it is difficult to differentiate the two reflectors near their intersection.

Ideally, smoothing the noisy image would clean up the image, and allow an interpreter to clearly see where the reflectors intersect. However, a traditional 5x5x5 box filter fails to produce this result (Figure 3(a)). While much of the noise is indeed removed, the filter has blurred the reflectors, making it impossible to distinguish between them. Furthermore, taking the difference between the noisy image and the filtered image gives an unwanted result

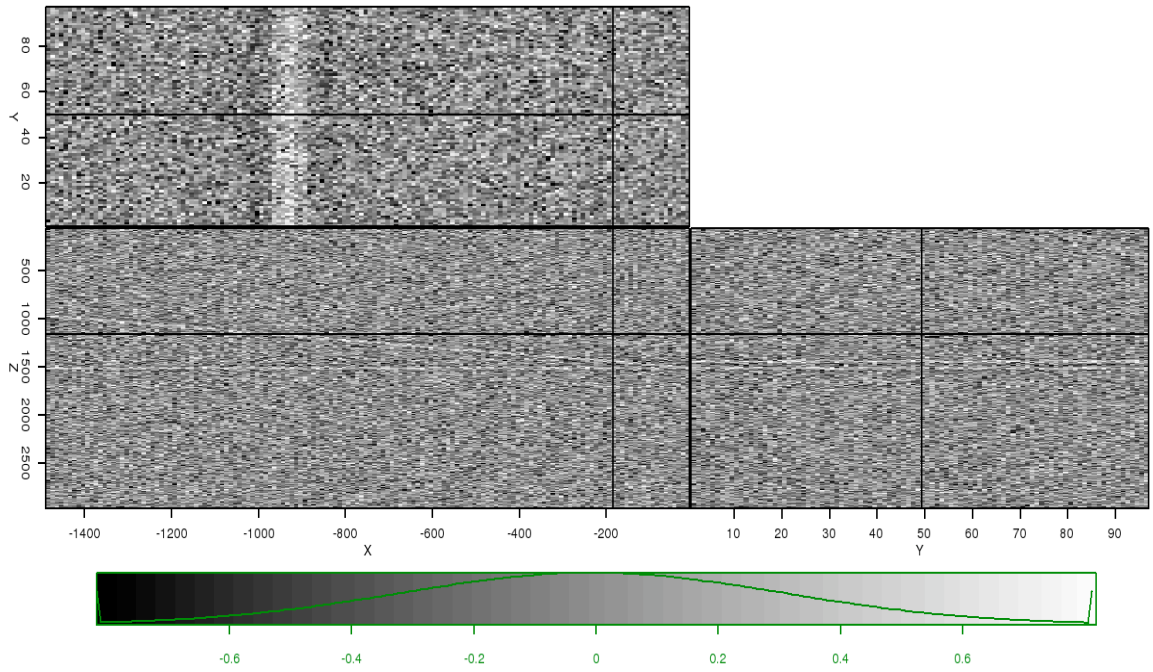


(a)

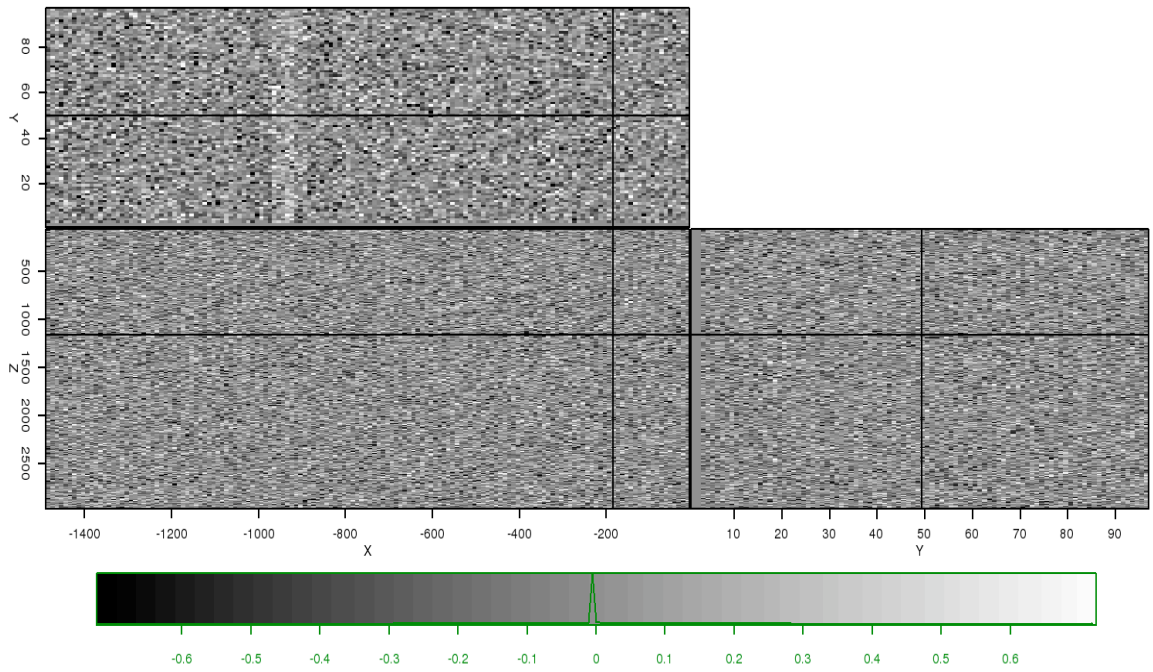


(b)

Figure 3: Results of smoothing the image in Figure 2 with (a) traditional box smoothing, and (b) maximum homogeneity (MH) filtering of the same operator length. [**ER**] adam1/. syn-smth,syn-mhn



(a)



(b)

Figure 4: Results of differencing the images in Figures 2 and 3. The filtered information when using the MH filter (b) contains much less coherent signal than when using traditional smoothing (a). [**ER**] adam1/. syn-diff-smth,syn-diff-mhn

(Figure 4(a)); it is obvious that a significant amount of coherent signal has been removed. In contrast, the maximum homogeneity (MH) filter with a mask length of 5 pixels performs better (Figure 3(b)). Not only are the two reflectors able to be distinguished much more easily in the filtered result (Figure 3(b)), but the difference calculation (Figure 4(b)) shows much less coherent signal being removed. Despite these advantages, though, more noise does remain in the MH-filtered image.

A HYBRID APPROACH

From the examples in Figure 3, it appears that traditional smoothing does perform very well in areas devoid of sharp boundaries or edges. Therefore, we can develop an approach that combines the characteristics of box smoothing in areas without edges, and takes advantage of the edge-preserving features of MH filtering when edges are present. For the MH algorithm, we already calculate variances for each of the bar masks passing through a given pixel. Comparing the largest and smallest of these calculated variances indicates the likelihood that an edge is present. If the ratio between the smallest and largest variances is large (close to 1), the pixel is in a relatively "isotropic" area, and an edge is unlikely to be present. Conversely, a smaller ratio implies that an edge is present in at least one of the bar mask orientations. If α is a user-determined threshold value, then if $\frac{\min(\sigma)}{\max(\sigma)} > \alpha$, traditional smoothing can safely be used in lieu of MH filtering for that particular location.

Figure 5 demonstrates this strategy on the synthetic example shown earlier. As the threshold value α decreases, the algorithm is more biased toward traditional smoothing; consequently, the amount of "speckle" noise decreases. However, near the two reflectors, MH filtering still holds sway. The final result is an image with the speckle noise removed as well as in the traditional smoothing result in Figure 3(a), but with the reflector edges preserved as well as the standard MH result in Figure 3(b).

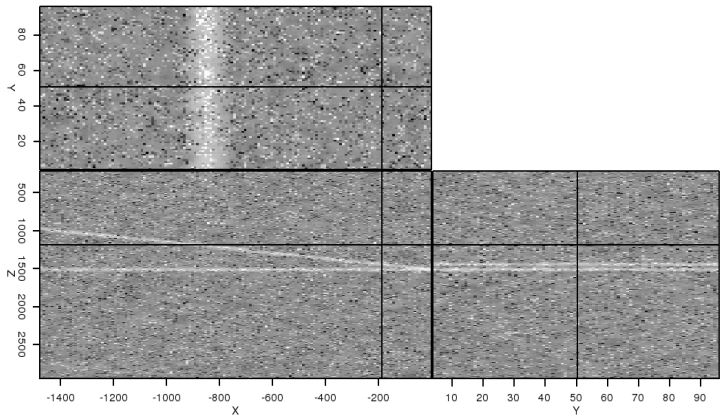
FIELD DATA EXAMPLES

Figure 6 shows a 3D field data image from the Gulf of Mexico. The salt boundary is clearly visible, but its discontinuous nature can pose problems for many automatic picking algorithms. Figure 7 shows the result of smoothing the image with traditional filtering (a), and the hybrid-MH method (b). The hybrid-MH filter attenuates a noticeable amount of background noise, while preserving the salt boundaries much more clearly than traditional smoothing. The difference calculations shown in Figure 8 confirm that MH-filtering removes much less coherent signal than traditional smoothing.

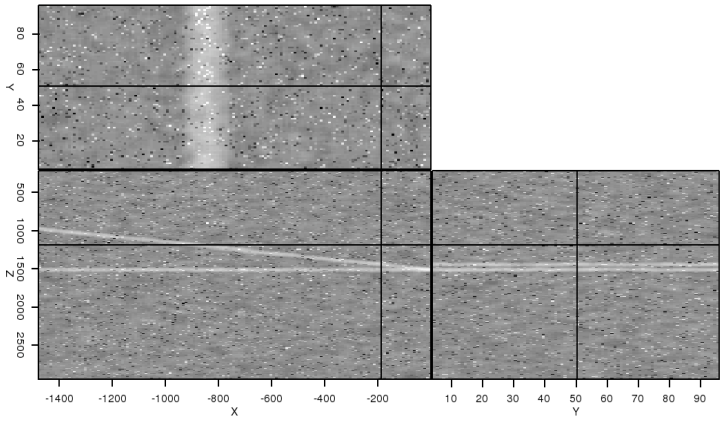
Figure 9 shows the effects of varying the hybrid threshold parameter α for the field data example. As α decreases, the image becomes noticeably smoother, even though the reflector boundaries remain clear. In this example, though, it is evident that the reflectors' amplitudes are being modified, which is an important consideration.

IMAGE SEGMENTATIONS

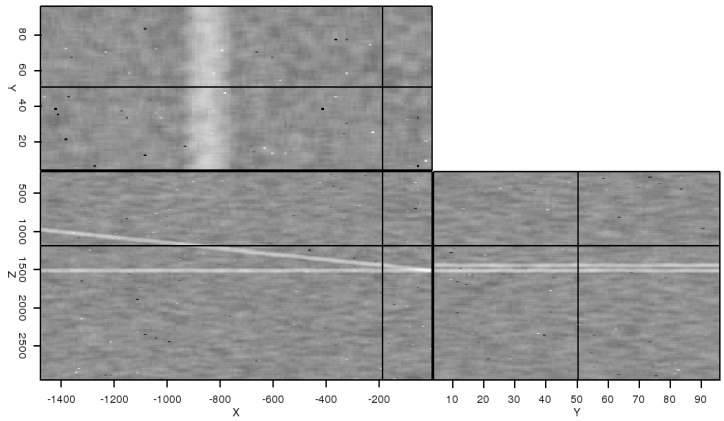
Finally, we should examine the effect of MH filtering on automatic segmentation results. Figure 10(a) is the result of segmenting the original, unfiltered image in Figure 6. "Leak-



(a)



(b)



(c)

Figure 5: The results of applying the hybrid-MH filter with α set at (a) 0.5, (b) 0.25, and (c) 0.1 to the original noisy synthetic image in Figure 2. As α decreases, more speckle-noise is removed from the image, while the sharpness of the reflectors is preserved. [**ER**] adam1/. syn-hyb-50,syn-hyb-25,syn-hyb-10

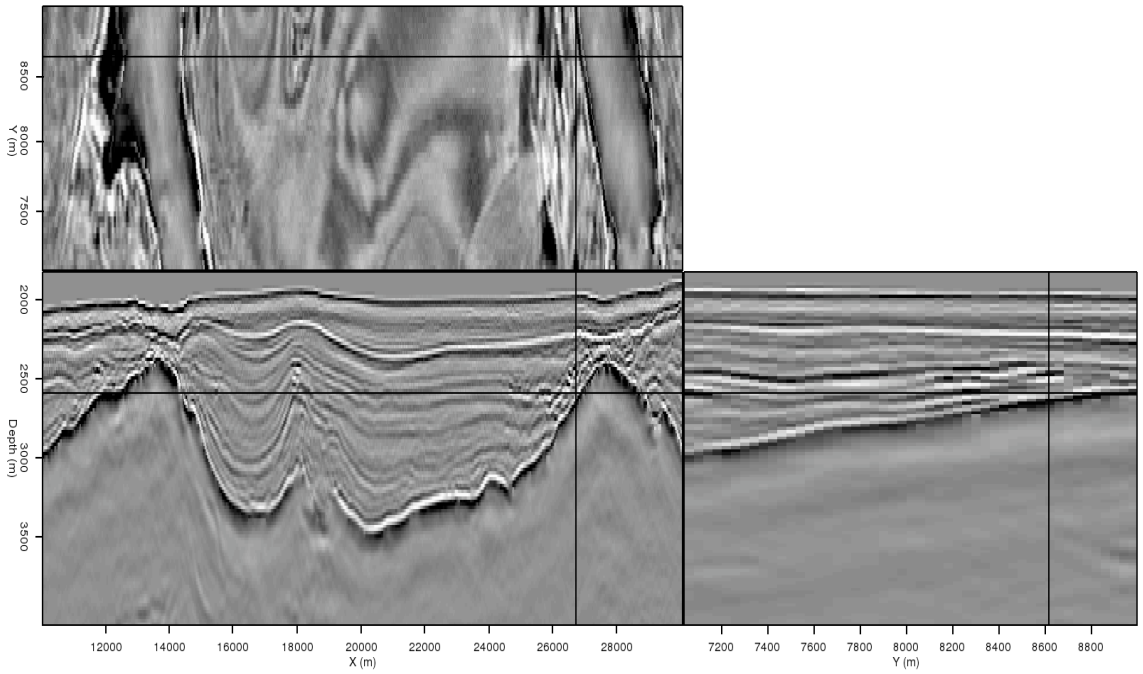
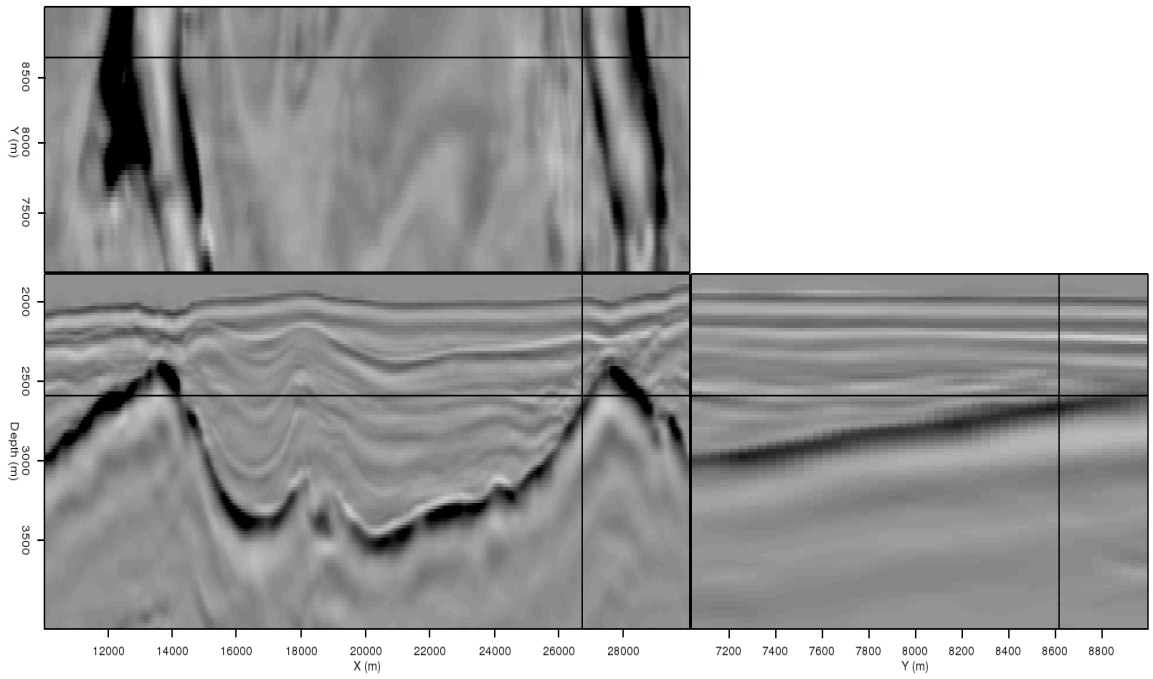


Figure 6: A 3D image from the Gulf of Mexico with a prominent but discontinuous salt boundary. [**ER**] adam1/. img-orig

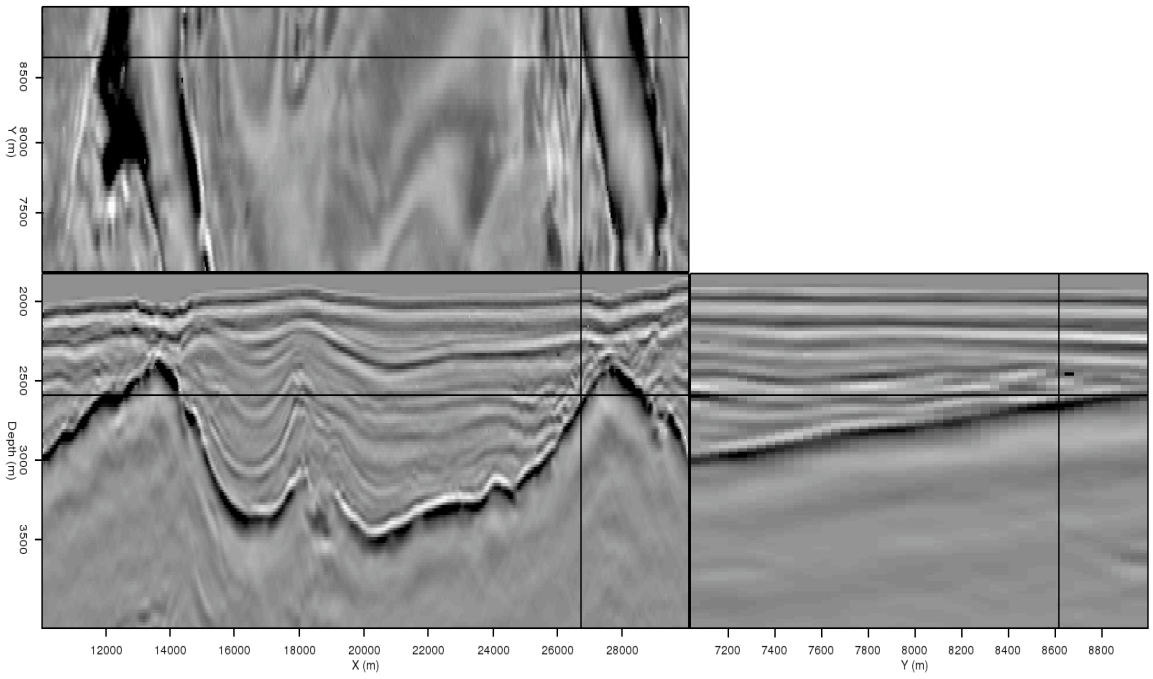
ages" are apparent from the salt body into the surrounding sediments, especially near the intersecting location indicators. When the MH-filtered image is segmented, however, these problems are greatly ameliorated (Figure 10(b)). A second field data example is the attempt to segment the small salt inclusion indicated in Figure 11. While segmenting the original image leads to the poor result in Figure 12(a), segmenting an MH-filtered image provides a much improved result (Figure 12(b)). In these examples, hybrid-MH filtering has allowed for more accurate segmentation results; furthermore, the computational efficiency (and simple parallelization) of the algorithm make it especially attractive since smoothing the image in these examples required only a fraction of the time needed for the already-efficient PRC segmentation algorithm.

CONCLUSIONS

Automated seismic image segmentation schemes can benefit from preprocessing such as smoothing prior to segmentation. An edge-preserving smoothing method based on directional maximum homogeneity can de-noise and clean up an image, while preserving important edges such as salt boundaries. Applying the MH filtering algorithm to 3D field seismic data improves segmentation results by reducing "leakages" through boundaries. In addition, a hybrid approach combining the best aspects of traditional smoothing and MH filtering can improve noise removal while still preserving edges well.

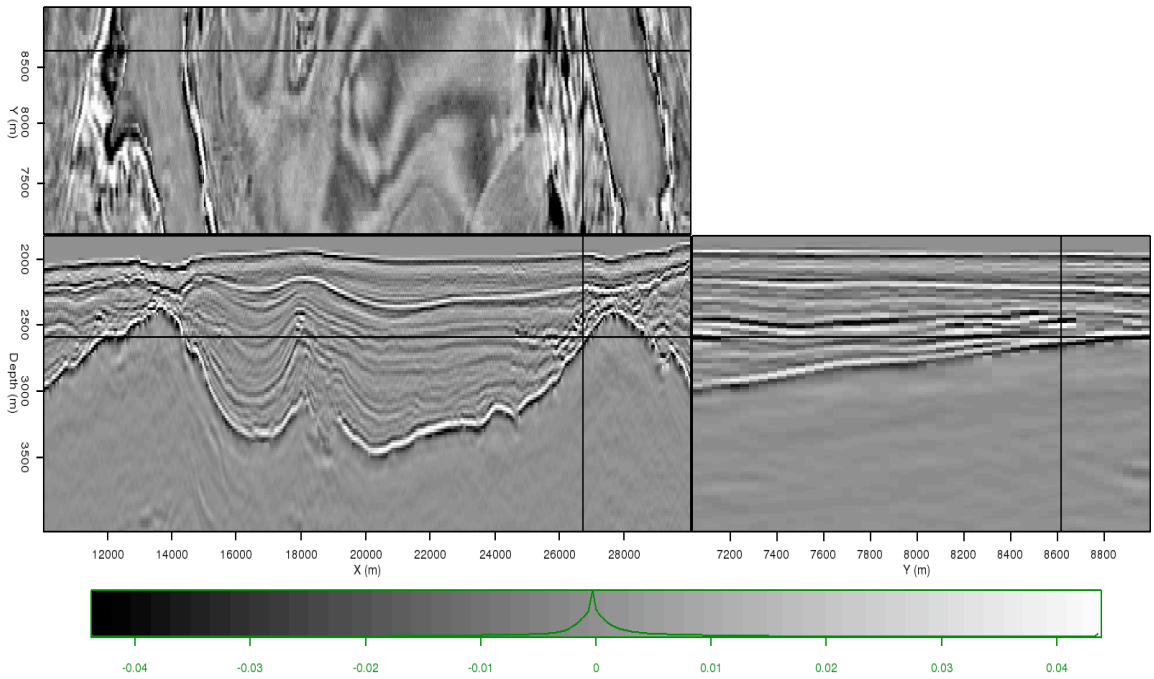


(a)

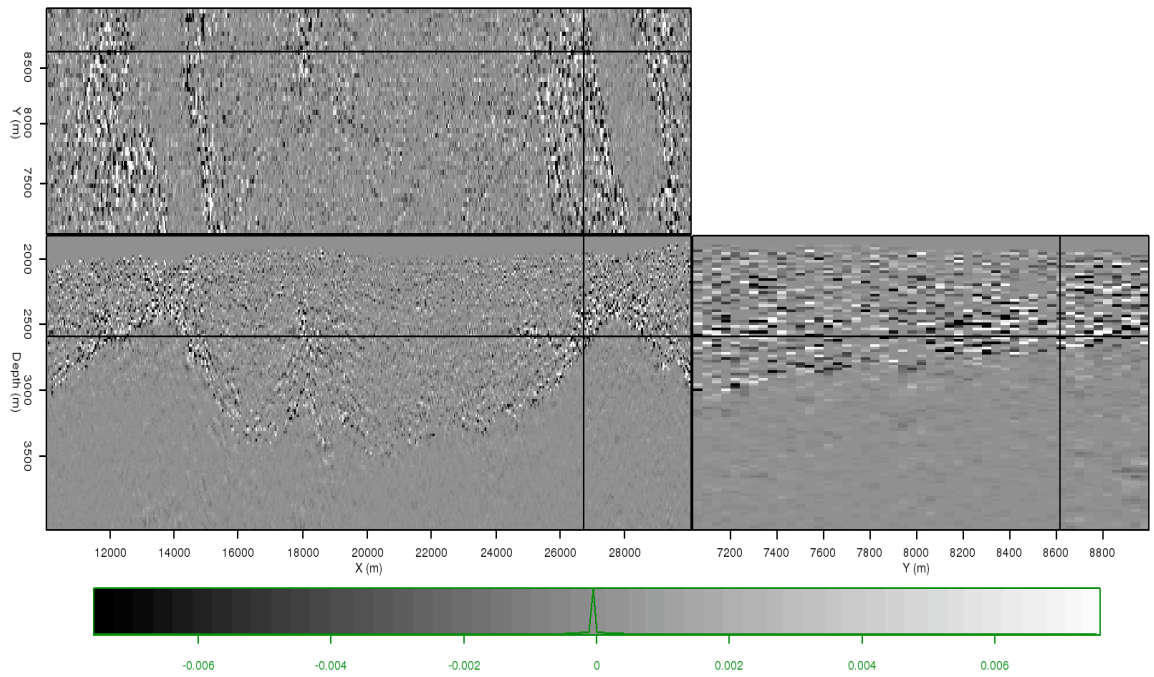


(b)

Figure 7: The image in Figure 6, smoothed with (a) a traditional box filter, and (b) a hybrid-MH filter of the same length. [**ER**] adam1/. img-smth,img-hyb-1

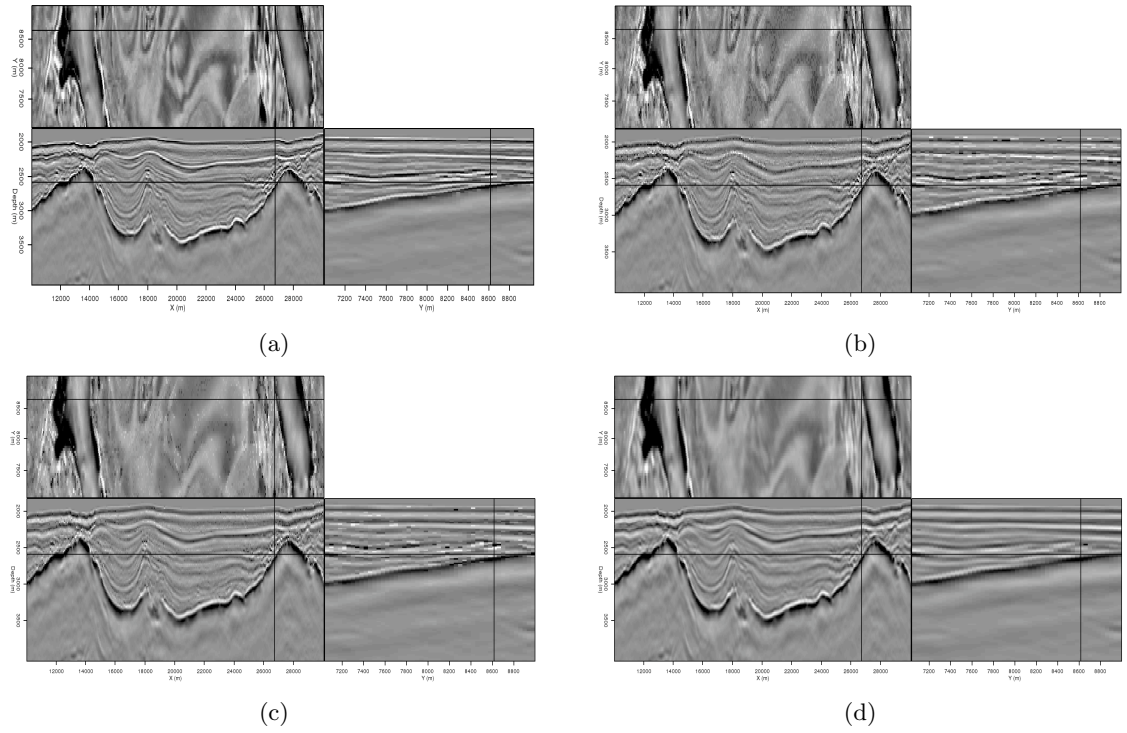


(a)



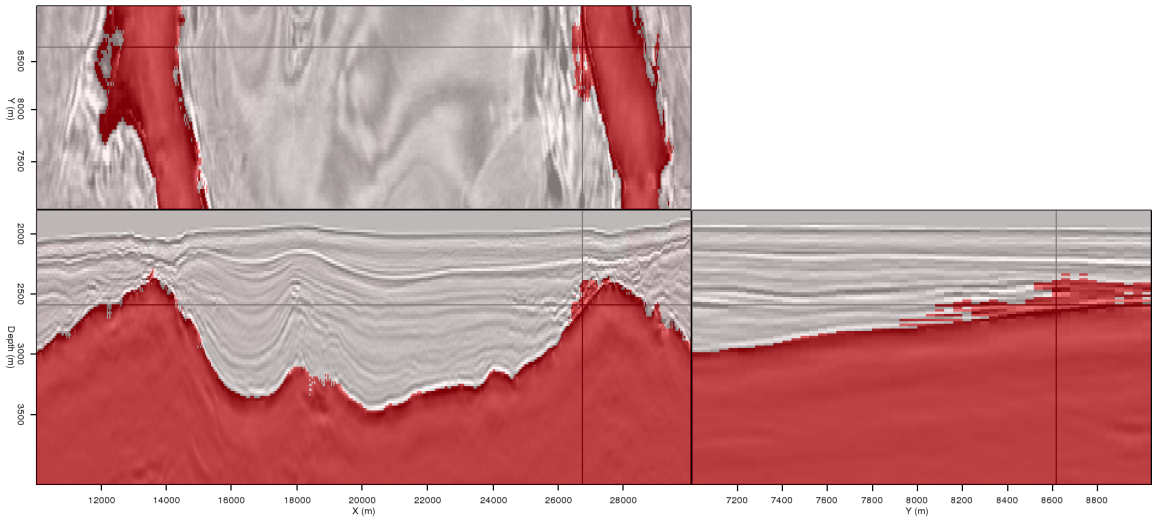
(b)

Figure 8: Results of differencing the original image in Figure 6 with (a) the image smoothed with a traditional box filter, and (b) the image smoothed with the MH filter. Again, the result in (b) shows that much less coherent signal has been removed from the image. [**ER**] adam1/. diff-smth,diff-mhn

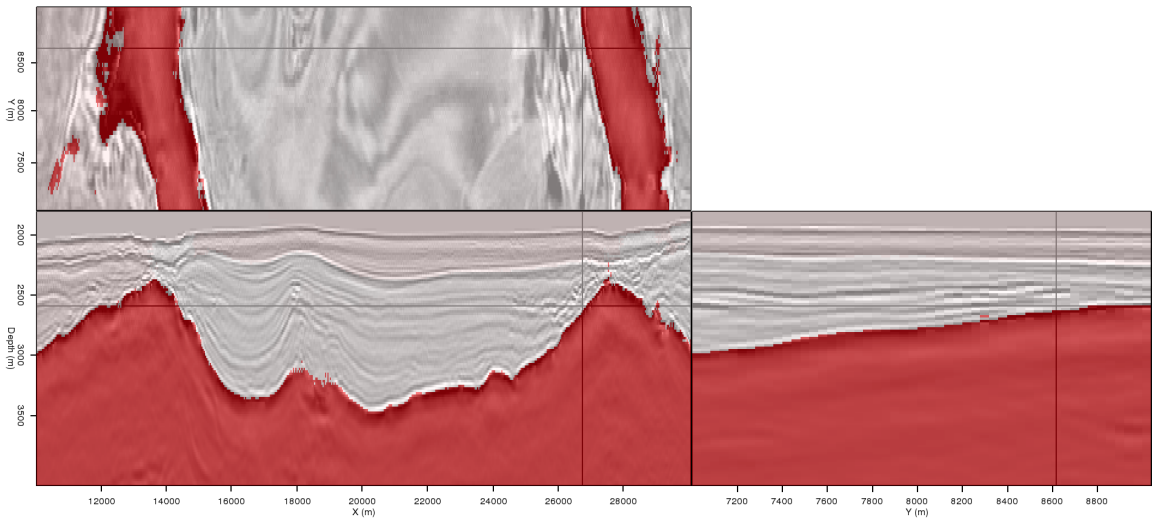


(a) (b) (c) (d)

Figure 9: Results of applying the hybrid-MH filter on the image in Figure 6, with α set at (a) 0.5, (b) 0.2, and (c) 0.1 and (d) 0.01. While the reflector amplitudes are affected, a great deal of speckle noise is removed at low α values. [**ER**] adam1/. img-hyb-50,img-hyb-20,img-hyb-10,img-hyb-01



(a)



(b)

Figure 10: Automatic image segmentation results when using (a) the original, unfiltered image in Figure 6, and the MH-smoothed image in Figure 9(d). [**ER**] adam1/. seg-orig,seg-mhn

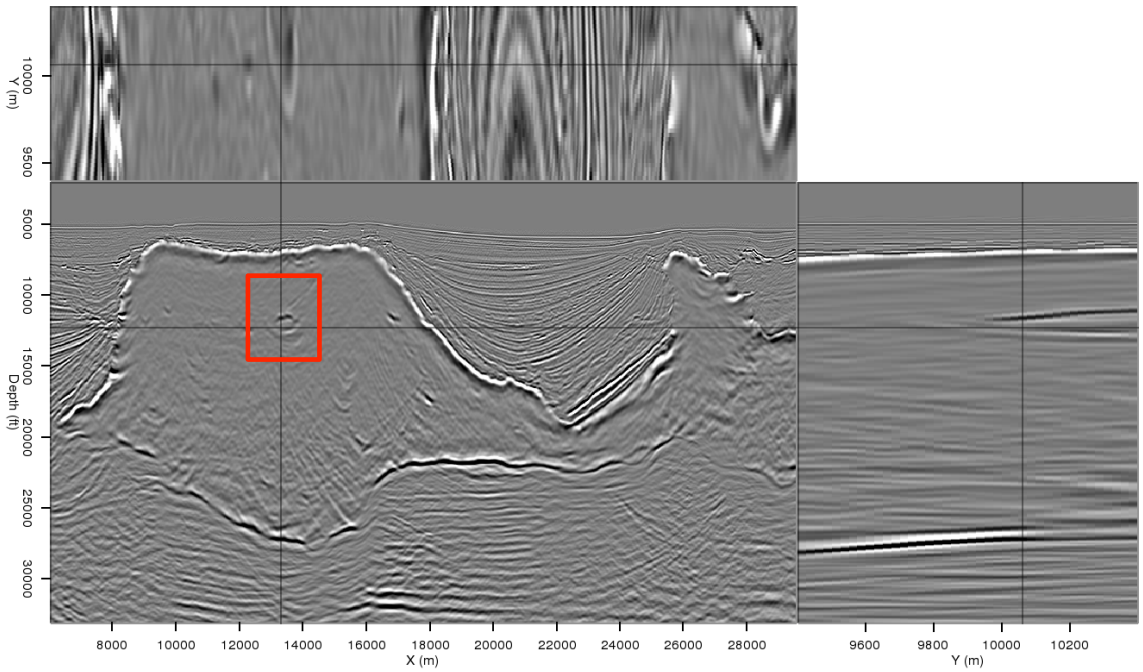


Figure 11: Another image from the Gulf of Mexico featuring an inclusion within the salt body. Further examples will be shown from the indicated area. [**ER**] adam1/. oct-full

ACKNOWLEDGMENTS

I thank WesternGeco and Unocal (Chevron) for providing the field data used for examples.

REFERENCES

AlBinHassan, N., Y. Luo, and M. N. Al-Faraj, 2006, 3D edge-preserving smoothing and applications: Geophysics, **71**, P5–P11.

Claerbout, J., 2005, Image estimation by example: Stanford University.

Fehmers, G. C. and C. F. W. Hocker, 2003, Fast structural interpretation with structure-oriented filtering: Geophysics, **68**, 1286–1293.

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

Guitton, A., 2005, Multidimensional seismic noise attenuation: PhD thesis, Stanford University.

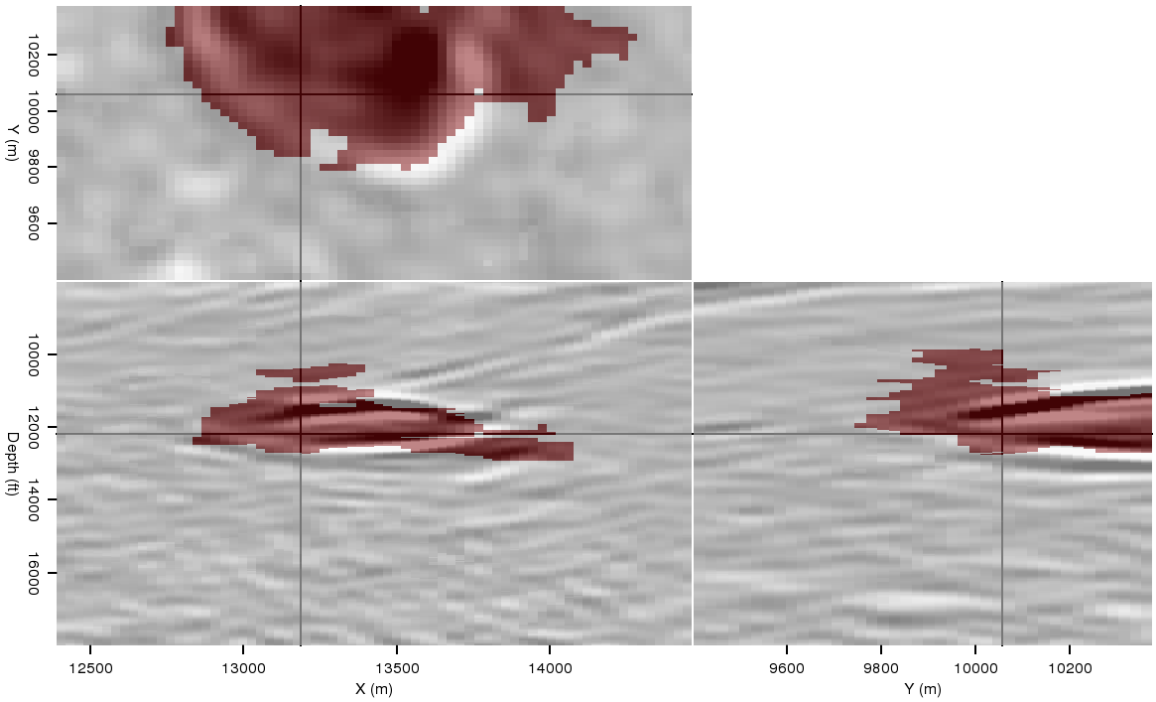
Hale, D., 2011, Structure-oriented bilateral filtering of seismic images: SEG Technical Program Expanded Abstracts, **30**, 3596–3600.

Halpert, A., R. G. Clapp, and B. L. Biondi, 2010, Speeding up seismic image segmentation: SEG Technical Program Expanded Abstracts, **29**, 1276–1280.

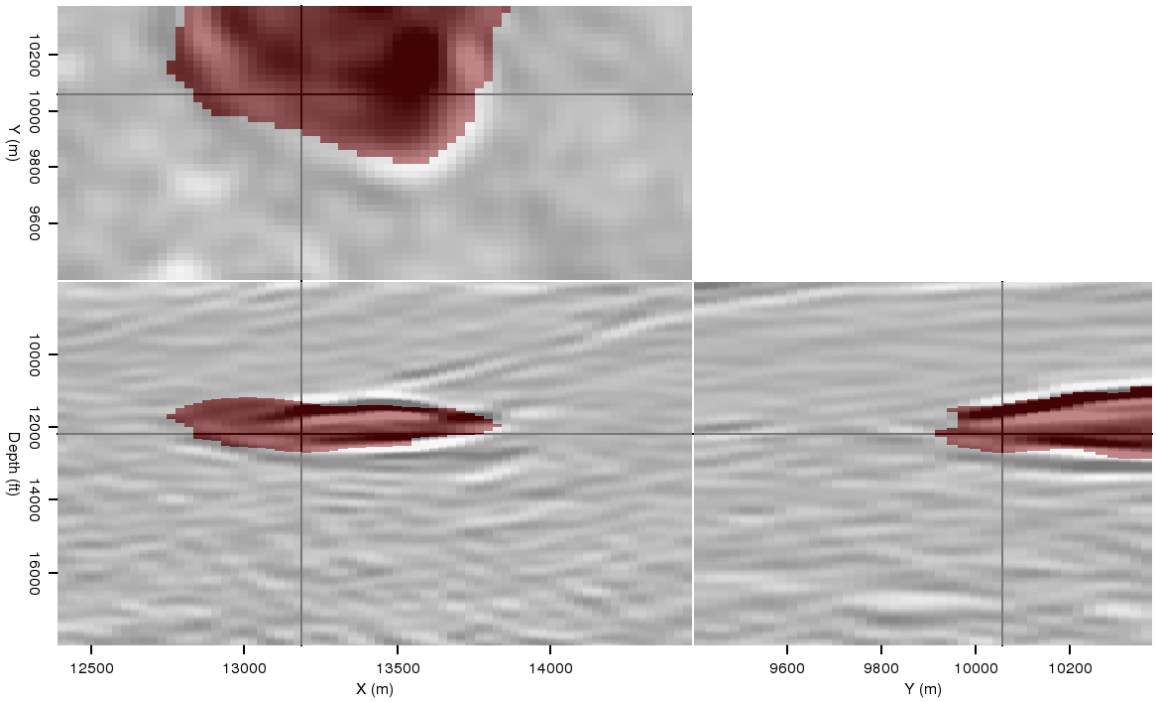
Tomita, F. and S. Tsuji, 1977, Extraction of multiple regions by smoothing in selected neighbourhoods: IEEE-SMC, **7**, 107–109.

Tukey, J. W., 1971, Exploratory data analysis: Addison-Wesley.

Zahedi, F. and R. Thomas, 1993, A maximum homogeneity based median filter: IEEE Colloquium on morphological and nonlinear image processing techniques, 7/1 – 7/5.



(a)



(b)

Figure 12: Automatic segmentation results using (a) the original, unfiltered image and (b) an MH-smoothed image. [**ER**] adam1/. oct-orig-seg,oct-smth-seg

Enhanced interpreter-aided salt-boundary extraction using shape deformation

Yang Zhang and Adam D. Halpert

ABSTRACT

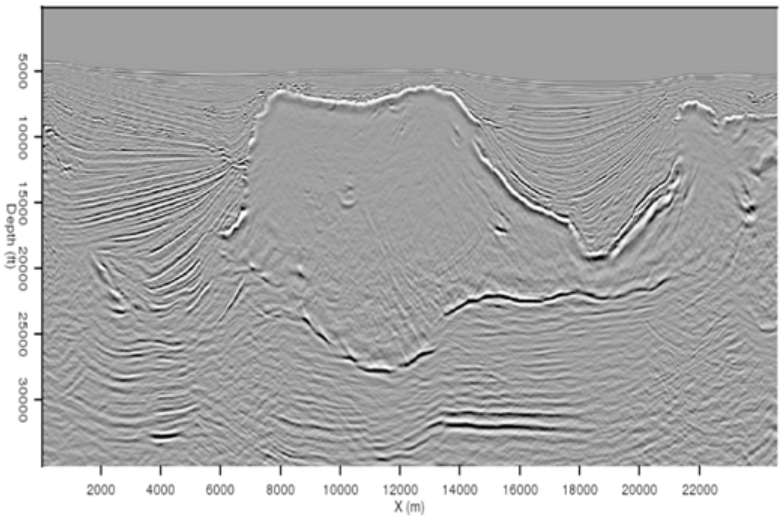
In many marine seismic exploration projects, precise interpretation of the salt-body geometry (which is also called salt-body segmentation) is a key component of building the subsurface velocity model. However, segmentation of salt is very human-intensive, even with the help of currently available semi-automatic computer software. This paper addresses the problem of automatically and accurately tracking the salt boundary in a series of neighboring seismic image slices, given an accurate salt segmentation for only one single reference slice. (The reference segmentation can be done manually). We achieve this using a landmark-based shape deformation technique plus SVM (Support Vector Machine) style regression. An example on a 3-D Gulf of Mexico data set demonstrates the effectiveness of our approach.

INTRODUCTION

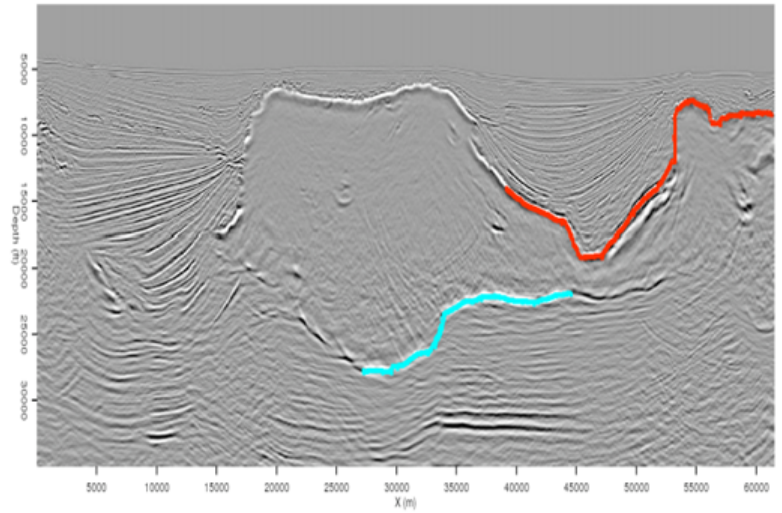
Interpreting the salt body in seismic images is of great importance for accurate velocity model building. Due to the poor quality of many seismic images, simple image processing filters followed by local boundary tracking algorithms give very poor results or fail easily. For example, in Figure 1, some parts of the salt boundary are simply missing from the image. A human interpreter's input is essential in these troublesome regions (Halpert et al., 2011). In the near term, the development of a super-algorithm that can fully automate boundary extraction without sacrificing the quality of the results is highly unlikely.

However, as today's seismic imaging practices evolve to three dimensions, manual interpretation of every single slice in a 3-D image cube is increasingly unrealistic. How can we achieve a good trade-off between the amount of manpower required and the quality of the boundary extraction result? The idea of manually segmenting only a small number of "key" slices and then intelligently propagating these results to the entire volume becomes attractive. Here we use a landmark-based shape-deformation technique to propagate a manual segmentation result of a single slice onto its neighboring slices, thus yielding a much better segmentation result overall than what we can achieve by simply applying fully automatic methods. The design goal for such intelligent boundary propagation consists of two parts:

- First of all, we force the new boundary to honor the available boundary information that can be confidently extracted from the image.
- Secondly, the new boundary should preserve the shape information known from the manual segmentation input, such that the boundary will deform reasonably where we do not have well-defined boundaries.



(a)



(b)

Figure 1: (a) A typical seismic image showing the salt body in the center. Some parts of the boundary are not well imaged due to the limited image quality. (b) The same image shown with the human-interpreted boundary. [**ER**] yang2/. Fig1

METHOD

Our approach is mainly based on the idea of Wang et al. (2001), described briefly here. Let us define the reference image slice that has been properly segmented as the *template image*, and define the image to which we want to propagate the segmentation result as the *input image*. The segmentation result in the template image is characterized by a set of contours. For simplicity, we assume that the template image contains only a single closed contour. Nonetheless, the extension of this method to handle multiple contours is straightforward.

We represent the known contour on the template image as a set of landmark points, $V = v_1, v_2, ..., v_n$ where $v_i = (x_i, y_i)$ (in 2-D Cartesian coordinates). Wang et al. (2001) describe the skeleton of this algorithm as follows:

> "For each landmark v_i, the proposed method first identifies a set of possible corresponding landmark points $B_i = \{v_i^{(j)}, j = 1, 2, ..., n_i\}$ on the input image, where $v_i^{(j)} = (x_i^{(j)}, y_i^{(j)})$. Then conceptually the deformation is solved in two major steps:
>
> 1. Identify the best landmark point v_i' from the landmark set B_i such that $V' = \{v_1', v_2', ..., v_n'\}$ is located in or near the true object boundary in the input image.
> 2. Deform the prior shape V to match V' while keeping the general shape characteristics of V. "

The cartoon in Figure 2 illustrates the idea of landmark-based contour deformation.

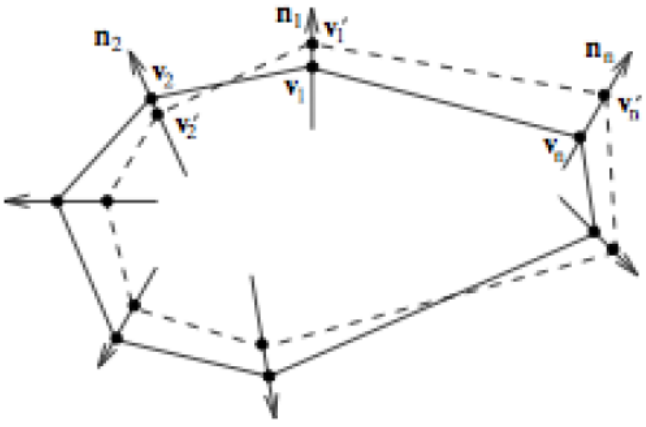


Figure 2: Landmark-based shape deformation, from Wang et al. (2001). [**NR**] yang2/. Fig2-deform

For the first step, we search the candidate points in set B_i along a short line segment that centers around point v_i and aligns to the contours normal direction $\mathbf{n_i}$. It is difficult to determine the best landmark point v_i' from all candidates in B_i at the first try. Therefore, we just choose randomly an element in each B_i to form the initial set V' and iterate this selection process a few times. During each iteration, we update the set V' such that V' more likely contains the correct corresponding landmarks.

The next deformation step is formulated by finding the optimal solution to an objective function which takes into account both the goal of deforming the points in V into the current

landmark set V' and the goal of preserving the prior shape V (using the bending-energy formula from (Bookstein, 1989)). The optimization goal is

$$\min_{V',\mathbf{t}} \left\{ \frac{1}{n} \sum_{i=1} Q(v_i', \mathbf{t}(v_i)) + \lambda\phi(\mathbf{t}) \right\} \tag{1}$$

in which $\mathbf{t}$ defines the deformation from V to V' as a mapping; i.e. $\mathbf{t} : (x, y) \rightarrow (f(x, y), g(x, y)) = (x', y')$. Function Q describes the term that penalizes the mismatch between V' (the landmarks we found on the input image) and the mapping defined by $\mathbf{t}(V)$. The Q term corresponds to the first goal, deforming the landmarks in set V to those in V'. Function $\phi(\mathbf{t})$ is a regularization term that tries to force the mapping $\mathbf{t}$ to be smooth, in other words, preserving the global shape information of the original V. We add a λ parameter to balance the weights of the two terms, Q and ϕ. The choice of λ is up to the user's judgment. After mathematical simplification, this optimization can be solved easily using the classical SVM(Support Vector Machine) regression technique (as a quadratic programming problem of size n). Moreover, the badly fitted components in set V' are identified as the support-vectors. We update the set V' by replacing those support-vectors with other candidates in B_i, such that the new set V' would achieve better fitting.

MATHEMATICAL SIMPLIFICATION

In this section, we reveal more mathematical details of this approach. The actual form of function Q in the objective function 1 is the ϵ-insensitive L1 norm:

$$||v_i' - \mathbf{t}(v_i)||_\epsilon = \begin{cases} 0, & \text{if } |v_i' - \mathbf{t}(v_i)| < \epsilon \\ |v_i' - \mathbf{t}(v_i)| - \epsilon, & \text{otherwise} \end{cases} \tag{2}$$

Since new landmark candidates in V' are sought along the contours normal direction $\mathbf{n_i}$, we constrain the desired mapping $\mathbf{t}(V)$ to displace v_i along direction $\mathbf{n_i}$ as well:

$$\mathbf{t}(v_i) = v_i + \gamma_i \mathbf{n_i}. \tag{3}$$

Let $\gamma = \gamma_i : i = 1, ..., n$. Since points in V' are found along the normal directions of the original contour V as well, we have $v_i' = v_i + h_i \mathbf{n_i}$. Then the previous problem 1 becomes

$$\min_{\mathbf{t},\gamma} \left\{ \frac{1}{n} ||h_i - \gamma_i||_\epsilon + \lambda\phi(\mathbf{t}) \right\}, \tag{4}$$

subject to constraint 3. As for, we choose the so-called bending-energy term, defined as

$$\phi(\mathbf{t}) = \int \int_{-\infty}^{+\infty} (E(f) + E(g))\, dxdy,$$

where

$$E(\bullet) = \left(\frac{\partial^2}{\partial x^2} \right)^2 + \left(\frac{\partial^2}{\partial y^2} \right)^2 + \left(\frac{\partial^2}{\partial x \partial y} \right)^2.$$

The nice thing about this choice of bending-energy is that we know in advance, given all mappings that satisfy constraint 3, the mapping specified by thin-plate spline interpolation will minimize the bending-energy (Bookstein, 1989). In other words, the solution $\mathbf{t}^*$ to the

optimization problem 4 must be the thin-plate spline interpolation that maps $\{V : v_i, i = 1, ..., n\}$ to $\{\mathbf{t}(V) : \mathbf{t}(v_i) = v_i + \gamma_i \mathbf{n_i}, i = 1, ..., n\}$. Given that $\mathbf{t}$ must be a thin-plate spline interpolation, we can express $\phi(t)$ with the vector γ. Therefore, this variational problem (where the optimization parameters are functions not numbers) turns into a much simpler numerical convex optimization problem. We just need to find the optimal γ for the problem below:

$$\min_{\gamma} \left\{ \frac{1}{n} ||h_i - \gamma_i||_\epsilon + \frac{\lambda}{8\pi} (\hat{x}^T L \hat{x} + \hat{y}^T L \hat{y}) \right\}, \tag{5}$$

where $\hat{x}, \hat{y}$ is the vector representation of the x and y coordinates of the points in set $\mathbf{t}(V)$, and L is a semi-positive definite matrix defined by known quantities.

Using the standard SVM technique, we can instead solve the dual problem of 5 according to the K.K.T.(Karush–Kuhn–Tucker) conditions. It ends up being a standard quadratic programming problem with both upper and lower bounds.

EXAMPLES

We test this algorithm on a 3-D seismic image cube from a Gulf of Mexico seismic survey. The cube is of discrete size $970 \times 784 \times 12$ in the depth(Z), inline(X) and cross-line(Y) directions respectively. The grid spacings are 9.6 m, 25 m and 30 m. We test this algorithm on a 3-D seismic image cube from a Gulf of Mexico seismic survey. The cube is of discrete size $970 \times 784 \times 12$ in the depth(Z), inline(X) and cross-line(Y) directions respectively. The grid spacings are 9.6 m, 25 m and 30 m. We have only the human-interpreted segmentation result in slice 1, which we use as the template image. We then perform initial processing on this segmentation result to extract the template landmarks as shown in Figure 3.

To find candidates for the set B_i, we overlay the landmark on the energy envelope of the input image (neighboring slice), then we search along the normal direction for certain image features which might suggest that certain locations be part of the boundary. Here we just use a very simple criterion: we choose the locations of the local amplitude maxima as the boundary point candidates. The plot in Figure 4(a) shows all candidates in B_i for $i = 137$.

We then run the optimization for a few iterations. During each iteration, we identify all the support vectors (which correspond to the fitting outliers); for each support vector h_i, we try to use other candidates in set B_i such that the fitting $h_i - \gamma_i$ improves. Figure 4(b) demonstrates this step during one iteration.

Finally we deform all 12 slices one by one, in increasing distance from the template slice. In Figure 5, we show the comparison between the obtained deformed boundary and the boundary found by automatic methods with the simple way of propagating the user input as described in Halpert et al. (2011). The improvement is prominent, with several regions highlighted in circles. The deformed boundary is less jagged and tracks the local edges in the image better. The shape-preserving constraint helps prevent boundary leakage as the automatic segmentation is done using flooding algorithms.

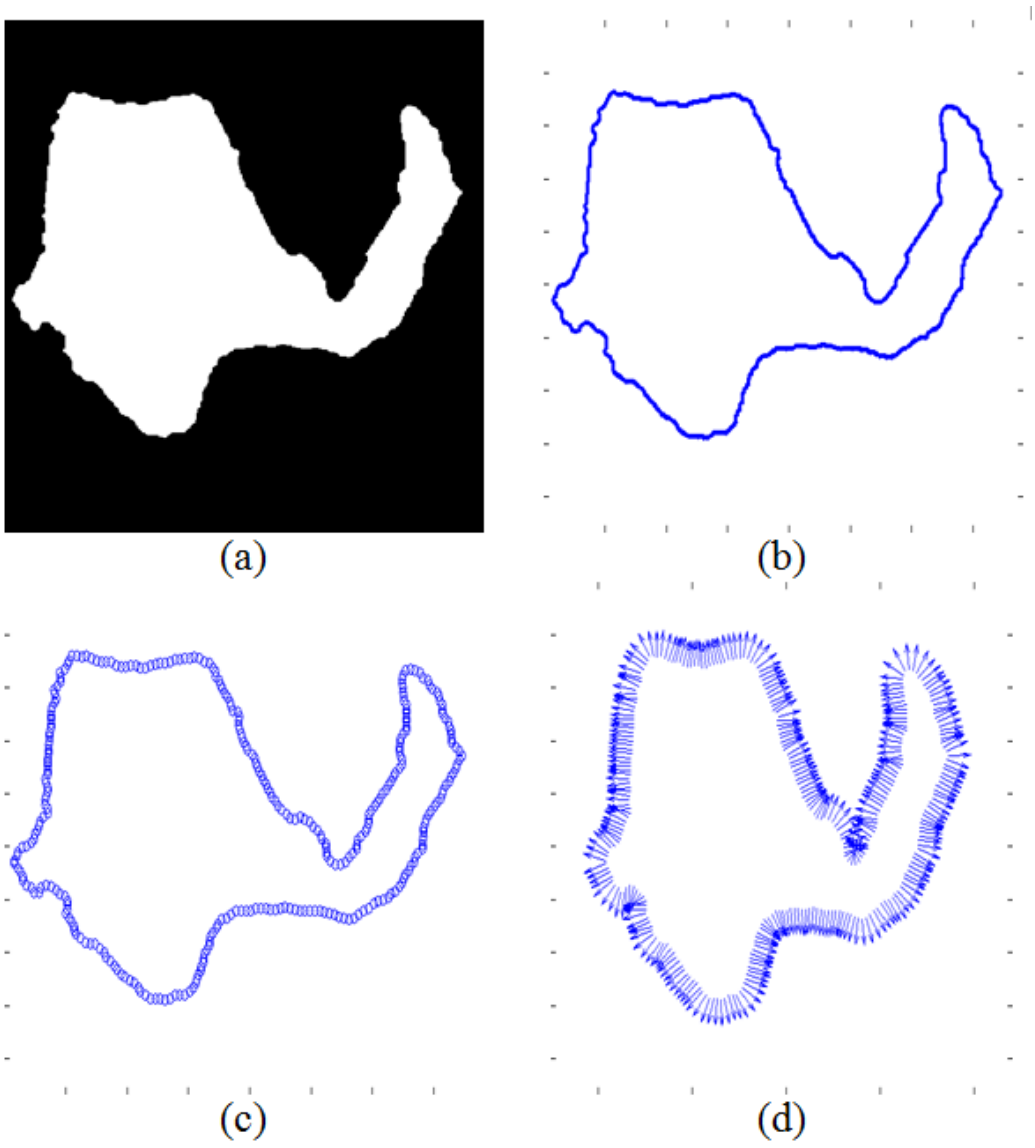
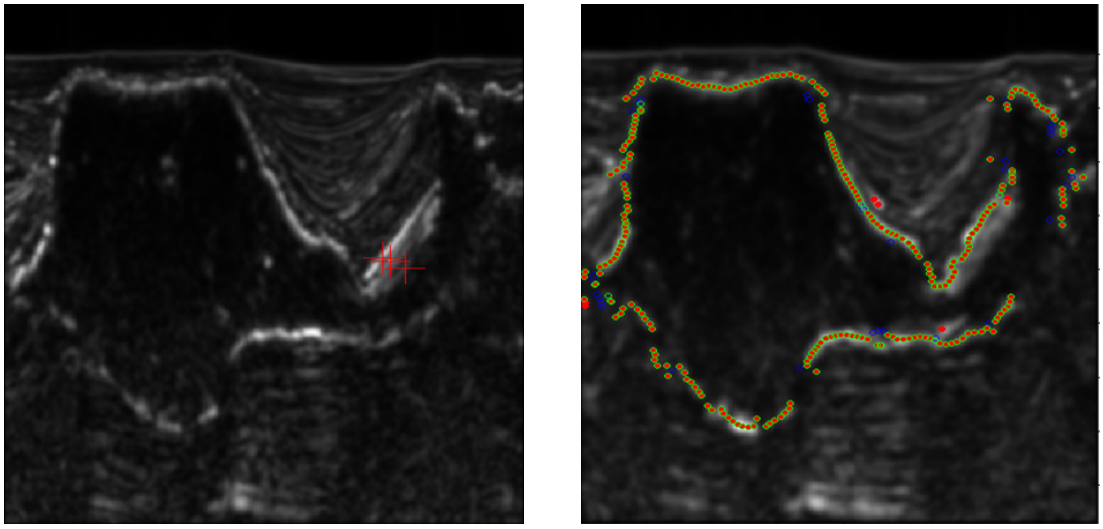


Figure 3: Preprocessing flow on the segmentation result of the template image. (a) Build the salt-body mask. (b) Extract the boundary. (c) Subsample to a list of landmarks. (d) The outnormal directions found for each landmark on the contour. [**CR**] yang2/. Fig3-segT-flow



(a) (b)

Figure 4: a) The three candidate points found in B_{137}. b) The update of set V' in one iteration. Blue indicates points in V' that are support vector points, red shows the original points in V', and green shows the updated set V' obtained by replacing the badly fitting points in v'_i with better fitting candidates in B_i. [**CR**] yang2/. Fig4,Fig5

CONCLUSION

In this paper, we improve salt-body segmentation for 3-D seismic images by deforming the accurate boundary on a human-interpreted slice into the neighboring slices. The deformation not only honors the available boundary information on the input slice, but also preserves the shape information from the template slice. Field data examples show very promising results. Possible additional applications of this method include the seismic tomography problem, where horizon-picking needs to be done for each of several non-linear iterations; meanwhile, the seismic image changes only slightly at each iteration.

ACKNOWLEDGMENTS

We thank WesternGeco for the seismic image used in this paper.

REFERENCES

Bookstein, F., 1989, Principal warps: Thin-plate splines and the decomposition of deformations: IEEE Trans. PAMI.

Halpert, A., R. Clapp, and B. Biondi, 2011, Interpreter guidance for automated seismic image segmentation: Presented at the Expanded Abstracts, EAGE 74th Annual International Conference and Exhibition.

Wang, S., W. Zhu, and Z.-P. Liang, 2001, Shape deformation: SVM regression and application to medical image segmentation: Proceedings on Eighth IEEE International Conference on Computer Vision, 209–216, IEEE.

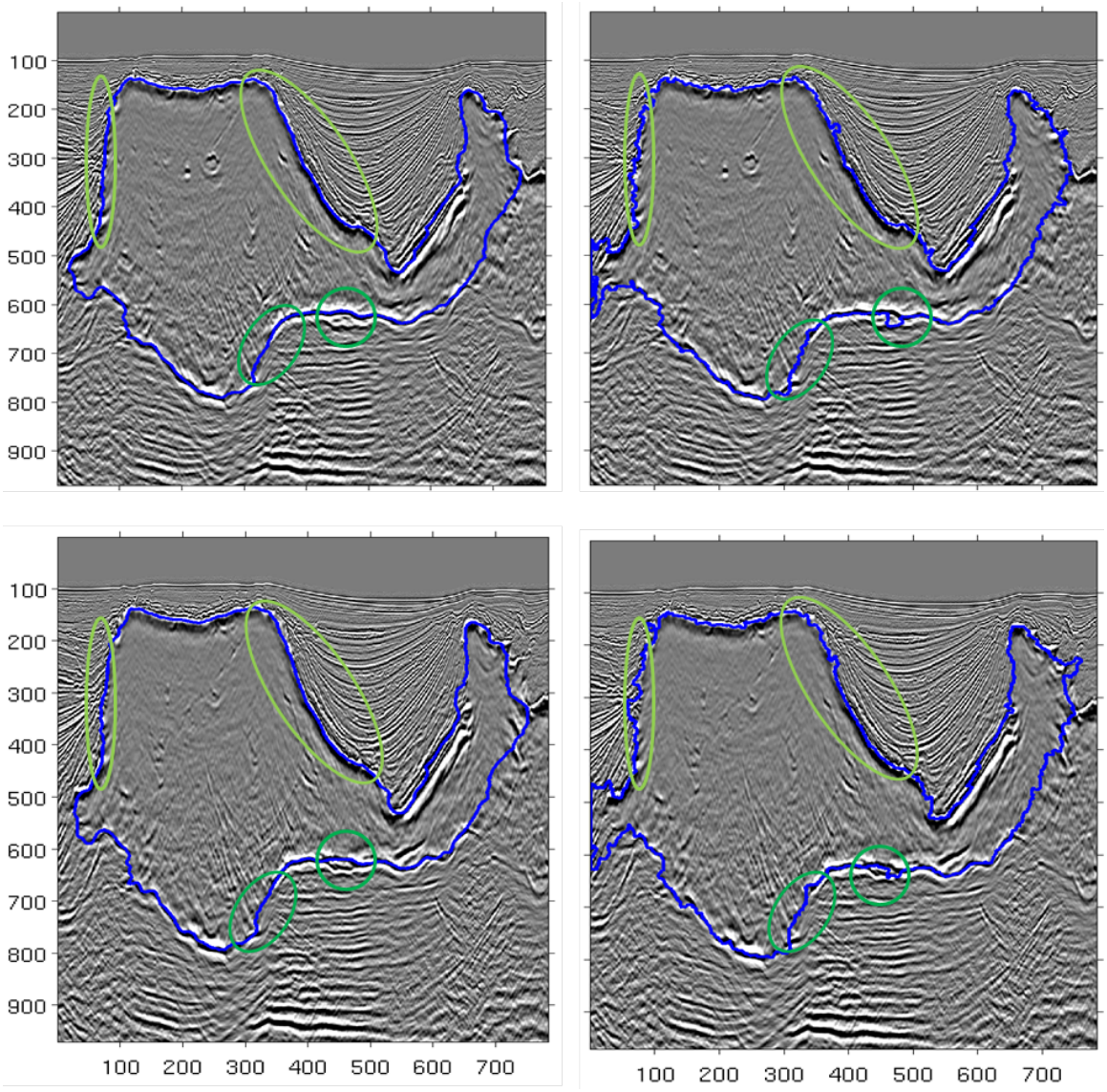


Figure 5: The segmentation result for slices 4 (top) and 12 (bottom) using our boundary deformation technique (left column) and the simple automatic method (right column). [**CR**] yang2/. Fig6-segmentComparison

Polarity preserving decon in "N log N" time

Jon Claerbout

ABSTRACT

A slight modification to Fourier spectral factorization enables deconvolution to preserve and enhance seismogram polarities. It spikes the center lobe of the Ricker wavelet. It works by tapering at small lags the antisymmetric part of the time-domain representation of the log spectrum.

INTRODUCTION

On a microscopic scale, the arrival of seismic events is always emergent, consequently use of strict causality in mathematical analysis often leads to disappointing results. Here we see how a slight modification to Fourier spectral factorization allowing slight noncausality (half period) easily handles the Ricker wavelet, a commonly observed emergent waveform.

Conflicting goals

Two goals of seismogram source waveform estimation conflict. They are:

1. Preserve and clearly exhibit the polarity of seismic reflections.
2. Estimate and use for data processing a source waveform that is causal, namely, the response vanishes before the excitation.

When we honor one, we find trouble with the other. This conflict will be defined and resolved here.

Prevalence of Ricker wavelet

The conflict is most directly seen and addressed in the specific case of the Ricker wavelet. It is generally seen on the water bottom and on any strong reflector such as the top of salt and often the bottom of salt. The Ricker wavelet obscures recognition of multiple reflections by their polarity alternations. (For example, look ahead to the left side of Figure 2.)

Definition of Ricker wavelet

We use the term "Ricker wavelet" to describe any wave shape resembling the second derivative of a Gaussian. The reason Ricker wavelets are so prevalent in marine seismology is this: The water surface negative reflection coefficient causes two ghosts, one at the gun, the

other at the receiver. At each location the slightly delayed surface reflection (of negative polarity) applies a finite difference to the image. Consequently, any impulsive reflector looks like a band-limited second derivative. Land seismometers measure neither displacement nor velocity. Essentially, they are devices that measure acceleration. Band limiting their output again produces wavelets like the Ricker wavelet.

Symmetrical decon preserves polarity. Hooray!

The most primative deconvolution is a symmetrical filter, for example $\mathrm{FT}^{-1}(1/|D(\omega)|)$ where $|D(\omega)|$ denotes the amplitude spectrum of the data. The output $\mathrm{FT}^{-1}(D/|D|)$ of this filter is spectrally white. This decon filter converts a Ricker wavelet to an impulse at its center. Hooray! Notice this impulse will have the opposite polarity of that of predictive decon which attempts to spike the first lobe of the Ricker wavelet.

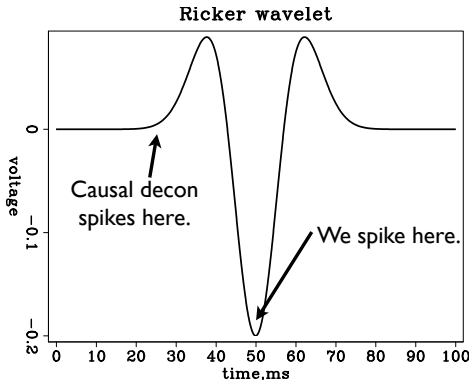


Figure 1: The non-causality needed is defined by the small backward distance from the center of the Ricker wavelet to its onset. [**NR**] jon1/. RickerAnnotated

How symmetric decon fails

Symmetric decon is wonderful the way it preserves polarity, but it has one feature that is really embarassing. About 150ms after the air gun blast is a bubble collapse blast. At early times they are quite different in frequency content and size, but because the bubble contains mostly the lower frequencies, its relative contribution becomes much stronger later in the record. What symmetric decon does with this bubble is horrifying. It gives a precursor 150ms before the water bottom. And also before every other event!

How predictive decon fails

Traditional predictive decon attempts to convert the Ricker wavelet to an impulse at its onset. This works badly because the onset time of the emergent signal, the Ricker wavelet, is not well defined.

GETTING THE BEST OF BOTH WORLDS

The good news is that it is possible to have the best of both worlds, both to deal time-symmetrically with the Ricker wavelet and non-time-symmetrically with the bubble. Furthermore, we do it rapidly in fast Fourier $N \log_2 N$ time. The Ricker wavelet is quite short, about 15ms, while the bubble delay is quite long, about 150ms, so the separation of the two is not delicate.

The less happy news is that the required theory is quite deep and not well known. It was first invented by famous mathematicians about 1940, and did not reach textbook status until my 1974 book FGDP, with code included in my newer book PVI 1992. Luckily, I have been teaching this material for many years now and believe I can extract the essence without great pain for you.

Spectral factorization

A causal function is one that vanishes at negative time. Too short a summary is to say the exponential of a causal is a causal. What is meant is if we take the Fourier transformation of a causal function, exponentiate it, and then inverse transform we will again have a causal function. This is the heart of spectral factorization, an obscure mathematical calculation addressing interesting practical applications.

Start with Z-transforms. Given a time function $(1, u_1, u_2, u_3, \cdots)$ its Z-transform is $U(Z) = 1 + u_1 Z + u_2 Z^2 + u_3 Z^3 + \cdots$. When you identify $Z = e^{i\omega \Delta t}$ and $Z^5 = e^{i\omega 5 \Delta t}$ the Z-transform is clearly a Fourier series. An example of a causal function is u_τ. It is causal because $u_\tau = 0$ for $\tau < 0$ likewise, $U(Z)$ has no powers of $1/Z$.

We may exponentiate $U(Z)$ by a frequency domain method or a time domain method. Easiest is the frequency domain method. Write $e^{U(Z(\omega))}$ for all ω, then Fourier transform to time. More interesting is the time domain method. The polynomial U has no powers of $1/Z$. The power series for an exponential is $e^U = 1 + U + U^2/2! + U^3/3! + \cdots$. Inserting the polynomial for U into the power series for e^U gives us a new polynomial (infinite series) that has no powers of $1/Z$. Furthermore, this new polynomial always converges because of the powerful influence of the denominator factorials. Thus we have shown that the "exponential of a causal is a causal".

Let $\bar{S}(Z(\omega))$ be an amplitude spectrum $\bar{S}(\omega) > 0$ with logarithm $\bar{U} = \log \bar{S}$. The exponential is the inverse of the logarithm

$$\bar{S} \;=\; e^{\log \bar{S}} \;=\; e^{\bar{U}} \tag{1}$$

Both $\bar{S}$ and $\bar{U}$ are real symmetric functions of ω. In the time domain, $|\bar{S}|^2$ corresponds to an autocorrelation. In the time domain, $\bar{U}$ merely corresponds to a real symmetric function $\bar{u}_\tau$. Adding some phase function $\Phi(\omega)$ to $\bar{U}$ will shift the time function s_τ, likely shifting each frequency differently.

$$S \;=\; e^{\log \bar{S} + i\Phi} \;=\; e^{\bar{U} + i\Phi} \;=\; e^{U} \tag{2}$$

Keeping $\bar{U}$ fixed keeps the spectrum $S^* S$ fixed. Let u_τ now correspond to the Fourier transform of $U(\omega) = \bar{U} + i\Phi$. The time symmetric part of u_τ corresponds to $\bar{U}(\omega)$ while the

antisymmetric part of u_τ corresponds to the newly added phase $\Phi(\omega)$. How shall we choose $\Phi(\omega)$? Let us choose the antisymmetric part of u_τ instead, choose it to cancel the symmetric part of u_τ on the negative τ axis. In other words, let us choose u_τ to be causal. Recalling that "exponentials of causals are causal" we have thus created a causal s_τ. Hooray! Hooray because s_τ has the same spectrum $\bar{S}$ that we started with. We started with a spectrum $\bar{S}$ and constructed a causal wavelet s_τ with that spectrum. Good trick! This is called "spectral factorization." Causal decon is simply taking your data D and dividing by a causal source waveform S.

Mostly causal decon

Now for the innovation. There are many pitfalls in the log domain. Seismologists are accustomed to ignoring the scale of their signals. Let the plot program figure out a suitable scale, we think. Once you take the logarithm of a signal, you are in a different world. If you double the log, you have squared the original signal. Got to be careful! What you can do safely with log signals is add or subtract something. This has the effect of scaling the original signal. Create an anticausal function $u_\tau^{\text{anti}} = \text{sgn}(\tau) u_\tau$. The signum function $\text{sgn}(\tau)$ is -1 for $\tau < 0$ and $+1$ for $\tau > 0$. Adding this anticausal function to u_τ zeros the negative lags while doubling the positive lags. Because it is antisymmetric it changes the phase spectrum. It does not change the amplitude spectrum. We can use any anti-symmetrical function we wish to monkey with the phase while not changing the amplitude. We could add the antisymmetric function u_τ^{anti}, but that would simply do traditional causal decon. Instead, near the origin we taper u_τ^{anti} towards zero. This creates symmetric Ricker-like behavior near the origin while leaving causal behavior further away. The tapering zone used here extends beyond the Ricker width, about 20ms, but not so far as the bubble delay, about 150ms, an easy distinction. A parallel analysis is found in another paper in this report (Claerbout et al. (2012)).

It's easy. Figures 2 and Figure 3 show the desired behavior. Hooray! The results are lovely. Better yet, they are not the end but the beginning. They are based on the simple notion that we want a white output spectrum. Our real goal is a sparse time function, not a white one. The results in these figures are simply the starting point of another paper in this report.

DISCUSSION AND CONCLUSION

This paper introduces the notion that by manipulating the u_τ we may make improvements on the old mathematical method of blind deconvolution. We were uncommonly successful here in dealing with our most commonly observed wavelet, the Ricker wavelet. This success suggests other improvements might flow from manipulations of the u_τ for other purposes.

For example, given only a single seismogram, we may wish to limit the number of degrees of freedom for the filter estimation. We have long known this can be done by smoothing the data spectrum. Another method is to limit the range, or taper the range of u_τ coefficients. Such ideas are untried, so not yet compared.

Likewise, many shot waveforms have been recorded and tabulated. Perhaps it makes sense to map these wavelets to the "lag-log" space u_τ to better understand their statistics.

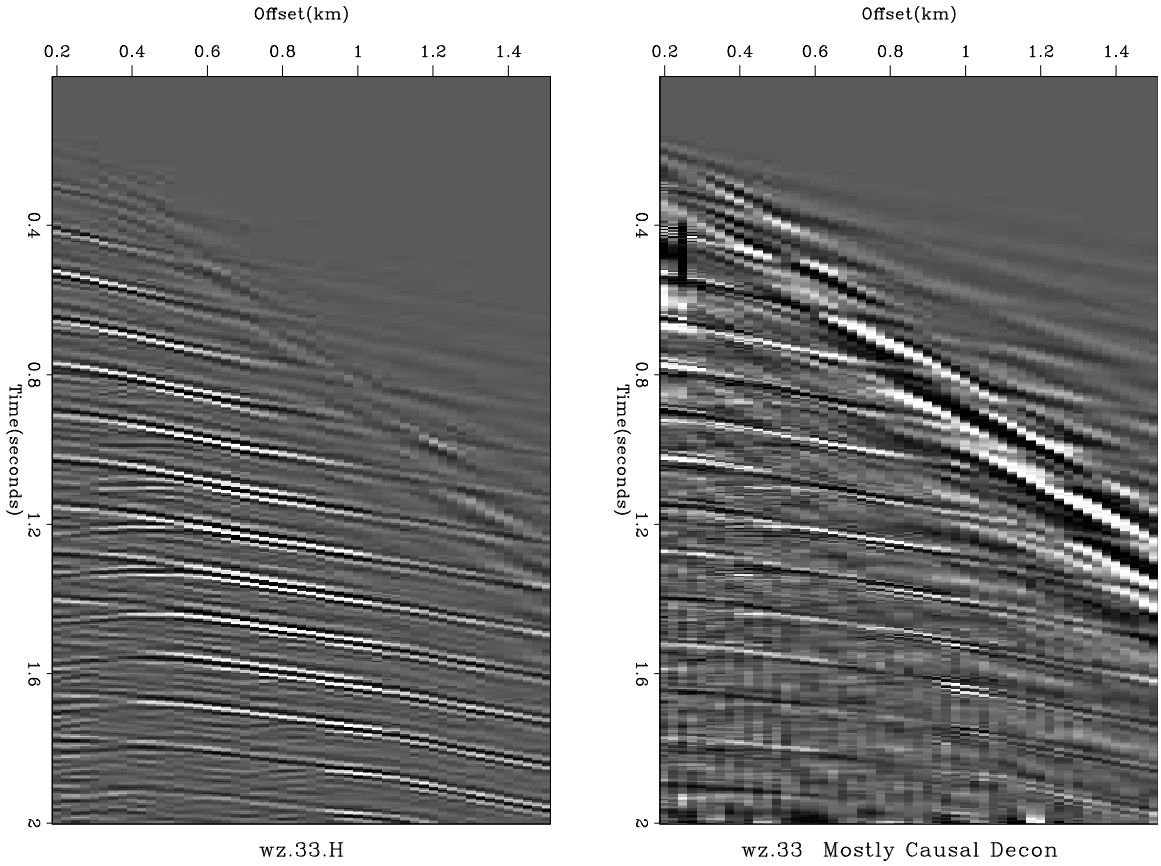


Figure 2: Left is Yilmaz and Cumro shot profile 33. Right is the result of "polarity preserving deconvolution." Observe enhanced visibility of alternating polarity of multiple reflections. [**ER**] jon1/. mostlycausal-33

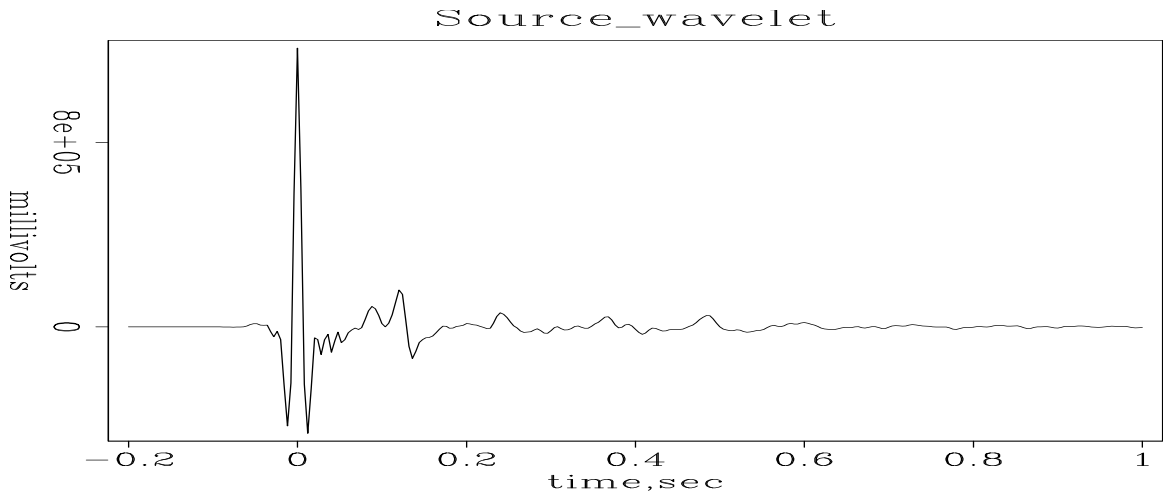


Figure 3: Shot waveform extracted from a constant offset section from the Gulf of Mexico. Peaks at zero lag as does symmetric Ricker wavelet, but for unknown reasons the peak heights are not in the (-1,2,-1) proportions of a Ricker wavelet. [**ER**] jon1/. GOMwavelet

I see no immediate application, but we might recall that spectral factorization is also applicable for complex-valued signals. Then the spectrum is non-symmetric. This arises when time-dependent signals have been previously Fourier transformed over space.

Shuki asks, "What about seafloor receivers where there is one ghost, not two?" I reply, "Perhaps the same code can be used, but instead of gateing on the range $\pm\tau$ being 3/4 period for the Ricker wavelet, it might be instead 1/4 period for the the primary lobe.

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. An insert in the middle implements the innovation of this paper; it diminishes the asymmetric part of u_τ near $|\tau| = 0$. A cosine squared weight was arbitrarily chosen. The suppression range was chosen from the origin to half way to an expected bubble on 4ms data.

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

```
subroutine kolmogoroff( n, cx)  # Spectral factorization.
integer i,              n       # input:  cx = amplitude spectrum
complex cx(n)                   # output: cx = FT of min phase wavelet
integer lag
real weight, asym
do i= 1, n
        cx(i) = clog( cx(i) )
call ftu( -1., n, cx)
do i= 2, n/2 {
        cx(i)     = cx(i) * 2.
        cx(n-i+2) = 0.
        }

# BEGIN stuff added to remove a little of the asymmetric part.
lag = 15        #   lag = 60ms/4ms where 60ms is half way to bubble.
do i = 2, lag {
        asym = (cx(i) - cx(n-i+2))/2.
        weight = cos( .5* 3.1416 * (i-1.)/(lag-1.))**2
        cx(i)     =  cx(i)     - weight * asym
        cx(n-i+2) =  cx(n-i+2) + weight * asym
        }
# END stuff added to remove a little of the asymmetric part.

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
#
#                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
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
```

REFERENCES

Claerbout, J., A. Guitton, and Q. Fu, 2012, Decon in the log domain with variable gain: SEP-Report, **147**, 313–322.

Decon in the log domain with variable gain

Jon Claerbout, Antoine Guitton, and Qiang Fu

ABSTRACT

We base deconvolution on the concept of output model sparsity. We improve our method of log spectral parameterization by including time-variable gain. Since filtering does not commute with time variable gain, gain is now done after decon (not before). Results at two survey locations confirm the utility. We resolve a stability issue with a long-needed regularization. An intriguing theoretical aspect shows that log spectral parameterization links penalty functions to crosscorrelation (not autocorrelation) statistics of outputs.

LOG SPACE, SPARSITY, AND GAIN

Because predictive decon fails on the Ricker wavelet, Zhang and Claerbout (2010) devised an extension to non-minimum phase wavelets (Zhang et al., 2011). Then (Claerbout et al. (2011)) replaced the traditional unknown filter coefficients by lag coefficients u_t in the log spectrum of the deconvolution filter. Given data $D(\omega)$, the deconvolved output is

$$r_t \;=\; \mathrm{FT}^{-1} \;\left[D(\omega) \;\exp\left(\sum_t u_t Z^t \right) \right] \tag{1}$$

where $Z = e^{i\omega}$. The log variables u_t transform the linear least squares (ℓ_2) problem to a non-linear one that requires iteration. Losing the linearity is potentially a big loss, but we lost that at the outset when we first realized we needed to deal with the non-minimum phase Ricker wavelet. We find convergence is typically quite rapid.

The source wavelet, inverse to the decon filter above, corresponds to $-u_t$. The positive lag coefficients in u_t correspond to a causal minimum phase wavelet. The negative lag coefficients correspond to an anticausal filter.

Here for the first time we introduce the complication that seismic data is non-stationary requiring a time variable gain g_t. The deconvolved data is the residual r_t. The gained residual $q_t = g_t r_t$ is "sparsified" (Li et al., 2012) by minimizing $\sum_t H(q_t)$ where

$$q_t \;=\; g_t \; r_t \tag{2}$$

$$H(q_t) \;=\; \sqrt{q_t^2 + 1} - 1 \tag{3}$$

$$\frac{dH}{dq} \;=\; H'(q) \;=\; \frac{q}{\sqrt{q^2+1}} \;=\; \mathrm{softclip}(q) \tag{4}$$

Traditional decon approaches are equivalent to chosing a white spectral output. Here we opt for a sparse output. In practice they might be much the same, but they do differ. Consider low frequencies. A goal is integrating reflectivity to yield impedance. We wish

to restore low frequencies where they enhance sparsity, but not where they merely amplify noise.

Our prefered penalty function $H(q)$ used for finding u_t is the hyperbolic (or hybrid) penalty function (equation (3)). The output q_t best senses sparsity when gain is such that the typical penalty $H(q_t)$ value is found near the transition level between ℓ_1 and ℓ_2 norms, namely, when typical $|q_t| \approx 1$.

MINIMUM PHASE EXTENSION

A minimum phase wavelet can be made from any causal wavelet by taking it to Fourier space, and exponentiating. The proof is straightforward: Let $U(Z) = 1 + u_1 Z + u_2 Z^2 + \cdots$ be the Z transform ($Z = e^{i\omega}$) of any causal function u_t. Consider $e^{U(Z)}$. Although we would always do this calculation in the Fourier domain, the easy proof is in the time domain. The power series for an exponential $e^U = 1 + U + U^2/2! + U^3/3! + \cdots$ has no powers of $1/Z$ (because U has no such powers), and it always converges because of the powerful influence of the denominator factorials. Likewise e^{-U}, the inverse of e^U, always converges and is causal. Thus both the filter and its inverse are causal. This is the essense of minimum phase.

We seek to find two functions, one strictly causal the other strictly anticausal.

$$U^+ = u_1 Z + u_2 Z^2 + \cdots \tag{5}$$
$$U^- = u_{-1}/Z + u_{-2}/Z^2 + \cdots \tag{6}$$

Notice U, U^2, etc do not contain Z^0. Thus the coefficient of Z^0 in $e^U = 1 + U + U^2/2! + \cdots$ is unity. Thus $a_0 = b_0 = 1$.

$$e^{U^+} = A = 1 + a_1 Z + a_2 Z^2 + \cdots \tag{7}$$
$$e^{U^-} = B = 1 + b_1/Z + b_2/Z^2 + \cdots \tag{8}$$

Define $U = U^- + U^+$. The decon filter is $AB = e^U$ and the source waveform is its inverse e^{-U}.

Consider $U(\omega) = \ln AB$ the log spectrum of the filter. We will be adjusting the various u_t, all of them but not u_0 which is the average of the log spectrum. The other u_t cannot change the average; they merely cause the log spectrum to oscillate.

THE GRADIENT

Having data d_t, having chosen gain g_t, and having a starting log filter, say $u_t = 0$, let us see how to update u_t to find a gained output $q_t = g_t r_t$ with better hyperbolicity. Our forward modeling operation with model parameters u_t acting upon data d_t (in the Fourier domain $D(Z)$ where $Z = e^{i\omega}$) produces deconvolved data r_t (the residual).

$$r_t = \mathrm{FT}^{-1}\, D(Z)\, e^{\cdots + u_2 Z^2 + u_3 Z^3 + u_4 Z^4 + \cdots} \tag{9}$$
$$\frac{dr_t}{du_\tau} = \mathrm{FT}^{-1}\, D(Z)\, Z^\tau e^{\cdots + u_2 Z^2 + u_3 Z^3 + u_4 Z^4 + \cdots} \tag{10}$$
$$\frac{dr_t}{du_\tau} = r_{t+\tau} \tag{11}$$

This follows because Z^τ shifts the data $D(Z)$ by τ units which shifts the residual the same amount. Output formerly at time t moves to time $t+\tau$. This is not the familiar result that the derivative of an output with respect to a filter coefficient at lag τ is the shifted *input* $d_{t+\tau}$. Here we have the *output* $r_{t+\tau}$. This difference leads to remarkable consequences below.

It is the gained residual $q_t = g_t r_t$ that we are trying to sparsify. So we need its derivative by the model parameters u_τ.

$$q_t = g_t\, r_t = r_t\, g_t \tag{12}$$

$$\frac{dq_t}{du_\tau} = \frac{dr_t}{du_\tau}\, g_t = r_{t+\tau}\, g_t \tag{13}$$

Recall $u_0 = 0$ and hence $\Delta u_0 = 0$. To find the update direction at nonzero lags $\Delta\mathbf{u} = (\Delta u_t)$ take the derivative of the hyperbolic penalty function $\sum_t H(q_t)$ by u_τ.

$$\Delta\mathbf{u} = \sum_t \frac{dH(q_t)}{du_\tau} \qquad \tau \neq 0 \tag{14}$$

$$= \sum_t \frac{dq_t}{du_\tau}\, \frac{dH(q_t)}{dq_t} \tag{15}$$

$$\Delta\mathbf{u} = \sum_t (r_{t+\tau}) \quad \big(g_t H'(q_t)\big) \qquad \tau \neq 0 \tag{16}$$

This says to crosscorrelate the physical residual r_t with the statistical residual $g_t H'(q_t)$. Notice in reflection seismology the physical residual r_t generally decreases with time while the gain g_t generally increases to keep the statistical variable q_t roughly constant, so $g_t H'(q_t)$ grows in time(!)

In the frequency domain the crosscorrelation (16) is:

$$\Delta U = \overline{\mathrm{FT}(r_t)}\ \ \mathrm{FT}(g_t\, \mathrm{softclip}(q_t)) \tag{17}$$

Equation (17) is wrong at $t = 0$. It should be brought into the time domain and have Δu_0 set to zero. More simply, the mean can be removed in the Fourier domain.

Causal least squares theory in a stationary world says the signal output r_t is white (Claerbout, 2009); the *auto*correlation of the signal output is a delta function. Noncausal sparseness theory (other penalty functions) in a world of echoes (nonstationary gain) says the *cross*correlation of the signal output with its gained softclip is also a delta function (equation (16), upon convergence).

TAKING THE STEP

We adopt the convention that components of a vector $\mathbf{u}$ range over the values of (u_t), likewise for other vectors. Given the gradient direction $\Delta\mathbf{u}$ we need to know the residual change $\Delta\mathbf{r}$ and a distance α to go: $\alpha\Delta\mathbf{r}$ and $\alpha\Delta\mathbf{u}$.

A two-term example demonstrates a required linearization.

$$
\begin{aligned}
e^{\alpha \Delta U} &= e^{\alpha(\Delta u_1 Z + \Delta u_2 Z^2)} && (18)\\
e^{\alpha \Delta U} &= 1 + \alpha(\Delta u_1 Z + \Delta u_2 Z^2) + \alpha^2(\cdots) && (19)\\
\mathrm{FT}^{-1}\ e^{\alpha \Delta U} &= (1, \alpha \Delta u_1, \alpha \Delta u_2) + \alpha^2(\cdots) && (20)\\
\mathrm{FT}^{-1}\ e^{\alpha \Delta U} &= (1, \alpha \Delta \mathbf{u}) + \alpha^2(\cdots) && (21)
\end{aligned}
$$

With that background, neglecting α^2, and knowing the gradient $\Delta\mathbf{u}$, let us work out the forward operator to find $\Delta\mathbf{q}$. Let "$*$" denote convolution.

$$
\begin{aligned}
\mathbf{r} + \alpha \Delta \mathbf{r} &= \mathrm{FT}^{-1}(D e^{U + \alpha \Delta U}) && (22)\\
&= \mathrm{FT}^{-1}(D e^{U} e^{\alpha \Delta U}) && (23)\\
&= \mathrm{FT}^{-1}(D e^{U}) * \mathrm{FT}^{-1}(e^{\alpha \Delta U}) && (24)\\
&= \mathbf{r} * (1, \alpha \Delta \mathbf{u}) && (25)\\
&= \mathbf{r} + \alpha\, \mathbf{r} * \Delta \mathbf{u} && (26)\\
\Delta \mathbf{r} &= \mathbf{r} \ * \ \Delta \mathbf{u} && (27)\\
\Delta q_t &= g_t \ \Delta r_t && (28)
\end{aligned}
$$

It is pleasing that $\Delta\mathbf{r}$ is proportional to $\mathbf{r}$. This might mean we can deal with a wide dynamic range within r_t. The convolution, a physical process, occurs in the physical domain which is only later gained to the statistical domain q_t. Naturally, the convolution may be done as a product in the frequency domain.

To minimize $H(\mathbf{q}+\alpha\Delta\mathbf{q})$ express it as a Taylor series approximation to quadratic order. Minimizing yields

$$
\alpha \ = \ - \sum_t \Delta q_t H'_t \ \ / \ \ \sum_t (\Delta q_t)^2 H''_t \qquad (29)
$$

Update $q_t = q_t + \alpha \Delta q_t$ and $U = U + \alpha \Delta U$, optionally (Newton method) iterate (because the locations of the many Taylor series have changed slightly with the change in $\mathbf{q}$).

ALGORITHM

Pseudo code below finds the best single filter for a group of seismograms. Notice $g(t,x)$ could contain mute patterns, etc.

Lower case letters are used for variables in time and space like `d` $= d(t,x)$, `r` $= r(t,x)$, `g` $= g(t,x)$, `q` $= q(t,x)$, `dq` $= \Delta q(t,x)$. while upper case for functions of frequency `D` $= D(\omega,x)$, `R` $= R(\omega,x)$, `dR` $= \Delta R(\omega,x)$, `U` $= U(\omega)$, `dU` $= \Delta U(\omega)$. Asterisk $*$ means multiply within an implied loop on t or ω.

```
D =   FT(d)
U = 0.            # or other initializations
Remove the mean from U(omega).
Iteration {
     dU = 0
```

```
For all x
     r = IFT( D * exp(U))
     q = g * r
     dU = dU + conjg(FT(r)) * FT(g*softclip(q))
Remove the mean from dU(omega)
For all x
     dR = FT(r) * dU
     dq = g * IFT(dR)
Newton iteration for finding alfa {
     H'  = softclip( q )
     H'' = 1/(1+q^2)^1.5
     alfa= - Sum( dq * H' ) / Sum( dq^2 * H'')
     q = q + alfa * dq
     U = U + alfa * dU
     }
}
```

UNIQUENESS

As the figures show, our results are excellent, amazing even, but we've had a continuing problem with uniqueness. We find the pseudo-code presented here can spike any of the three lobes of the Ricker wavelet defining the sea floor. This is particularly annoying as it amounts to apparent time shifts and polarity changes. For about a year we ascribed this difficulty to nonlinear problems having many solutions, so we concentrated on controlling the descent. Now it looks like the problem is much simpler.

We now ascribe our uniqueness problem to a familiar problem in linear optimization. We believe we have what amounts to a null space. Tiny changes in initialization or other conditions lead to a wide variety of solutions.

For example, we often found by the third iteration we could see the spiking, and we could see the bubble estimation was well underway. By the tenth iteration it was pretty much settled down, and we would begin to be happy. But the computation was quick, so we were tempted to continue iterating. Maybe about the 150th iteration we would notice that spiking on the center of the Ricker wavelet would begin transition to spiking the first or third lobes of the Ricker wavelet (including the accompanying apparant polarity change). To make matters worse, only slight changes in the gain function g_t would determine the selection of which final lobe.

We wasted a lot of time believing nonlinearity was responsible for multiple solutions. Our early primative attempts at regularization had failed. With the pseudocode above you can have results like in this paper in a dozen iterations, however, the theory below explains the missing regularization that should allow you all the iterations you like.

REGULARIZATION

Regularization is where we impose our prior knowledge to account for the inadequacy of the data to completely define a solution. With years of experience we would look at the standard

formulation $0 \approx \epsilon \sum_\tau w_\tau (u_\tau - \bar{u}_\tau)^2$ and theory would guide us to statistical averages to give us ϵ, w_τ and $\bar{u}_\tau$. We have recently understood that the weighting function w_τ should be a matrix $\mathbf{W}$, and we now know what that matrix should be. First, our goals:

1. The shot waveform should resemble a Ricker wavelet near zero lag.
2. The shot waveform should be small or vanishing at larger negative lags. The decon wavelet should not have a long low frequency precursor.

Theoretically, the even part of u_τ controls the amplitude spectrum of the shot waveform. (A parallel analysis is found elsewhere in this report (Claerbout (2012)).) We will not touch that. The phase spectrum is determined by the odd part of u_τ. The near zero lags in u_τ control the near zero lags in the shot waveform and decon filter. We want the odd near-zero lags to be symmetric because the Ricker wavelet is symmetric. Thus the regularization is to minimize the antisymmetric part of the near-zero lags in u_τ.

The larger positive lags in u_τ deal with marine bubble and soil layer reverberation. That is good stuff. Bad are the larger anticausal lags. They should be zero because they are non-physical. They can be handled as an additional regularization, or more simply by windowing Δu_τ.

Code modifications required by regularization

Consider regularization of the form $0 \approx u_\tau - u_{-\tau}$. In matrix form this is $\mathbf{0} \approx \mathbf{r}_m = \mathbf{J}\mathbf{u}$ where the matrix $\mathbf{J}$ is defined below with six vector components in the ordering required by the fast Fourier transform program.

$$\mathbf{0} \;\approx\; \begin{bmatrix} r_m(1) \\ r_m(2) \\ r_m(3) \\ r_m(4) \\ r_m(5) \\ r_m(6) \end{bmatrix} \;=\; \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 & 0 & -1 \\ 0 & 0 & +1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & +1 & 0 \\ 0 & -1 & 0 & 0 & 0 & +1 \end{bmatrix} \begin{bmatrix} u(1) \\ u(2) \\ u(3) \\ u(4) \\ u(5) \\ u(6) \end{bmatrix} \tag{30}$$

Note that $\mathbf{J}^* = \mathbf{J}$. The gradient search direction is

$$\Delta \mathbf{u} \;=\; \frac{\partial\, \mathbf{r}_m^* \mathbf{W} \mathbf{r}_m}{\partial \mathbf{u}_m^*} \;=\; \frac{\partial \mathbf{r}_m^*}{\partial \mathbf{u}^*}\, \mathbf{W}\mathbf{r}_m \;=\; \mathbf{J}^* \mathbf{W} \mathbf{r}_m \tag{31}$$

where $\mathbf{W}$ is a diagonal matrix of weights. Again for six components, the diagonal contains $(1, w_1, w_2, 0, w_2, w_1)$.

Here are the modifications needed to incorporate ℓ_2 regularization on u_τ:

$$\text{argmin}_{\mathbf{u}} \qquad \sum_{t,x} H(q_t) + \frac{\epsilon}{2}\, \mathbf{u}^* \mathbf{J}^* \mathbf{W} \mathbf{J} \mathbf{u} \tag{32}$$

$$\Delta u_t \;=\; \text{as before} \;+ \epsilon\, \mathbf{J}^* \mathbf{W} \mathbf{r}_m \tag{33}$$

$$\Delta r_t \;=\; \text{as before} \tag{34}$$

$$\alpha \;=\; -\,\frac{(\sum_{x,t} \Delta q_t H_t') \;+\; \epsilon\; (\mathbf{r}_m \cdot \Delta \mathbf{r}_m)}{(\sum_{x,t} (\Delta q_t)^2 H_t'') \;+\; \epsilon\, (\Delta \mathbf{r}_m \cdot \Delta \mathbf{r}_m)} \tag{35}$$

In a least squares problem we compute a step size α as minus a ratio $\mathbf{r} \cdot \Delta\mathbf{r}$ over $\Delta\mathbf{r} \cdot \Delta\mathbf{r}$. Adding a least squares regularization to any convex fitting problem we simply add $\epsilon\,(\mathbf{r}_m \cdot \Delta\mathbf{r}_m)$ to the numerator and $\epsilon\,(\Delta\mathbf{r}_m \cdot \Delta\mathbf{r}_m)$ to the denominator.

Actually, another regularization is desireable. We should also request u_τ to be small for large anticausal lags, lags more negative than the range we are considering for the antisymmetry regularization. This might be handled by truncating the gradient rather than as a regularization.

A third regularization can be added to weaken u_τ at its large positive lags in circumstances where we feel we have insufficient data to estimate trace-long filters.

GOALS

A long range goal is to successfully integrate the reflectivity to get the log impedance. This requires good low frequency handling. Recording equipment often suppresses low frequencies for various practical reasons whose validity is likely location dependent. Our decon is pulling back some of these low frequencies but should know to stop before pulling up noise. Figure 2 demonstrates doing gain after non-minimum phase decon makes a valuable first step. To find impedance may require the additional statistical assumption of sparseness, but by solving the physical problem correctly, we have reduced the need for that.

ACKNOWLEDGEMENT

We thank Western Geophysical for the Gulf of Mexico data set. Jon Claerbout and Qiang Fu thank the sponsors of the Stanford Exploration Project. Antoine Guitton thanks Repsol Sinopec Brasil SA and Geo Imaging Solucoes Tecnologicas em Geociencias Ltda. We'd like to thank Yang Zhang for continued interest.

REFERENCES

Claerbout, J., 2012, Polarity preserving decon in "N log N" time: SEP-Report, **147**, 305–312.

Claerbout, J., Q. Fu, and Y. Shen, 2011, A log spectral approach to bidirectional deconvolution: SEP-Report, **143**, 297–300.

Claerbout, J. F., 2009, Image estimation by example.

Li, Y., Y. Zhang, and J. Claerbout, 2012, Hyperbolic estimation of sparse models from erratic data: Geophysics, **77**, V1–V9.

Zhang, Y. and J. Claerbout, 2010, A new bidirectional deconvolution method that overcomes the minimum phase assumption: SEP-Report, **142**, 93–104.

Zhang, Y., J. Claerbout, and A. Guitton, 2011, A new bi-directional sparse/spike deconvolution method that overcomes the minimum phase assumption: 73th EAGE Conference and Exhibition Extended Abstract, F001.

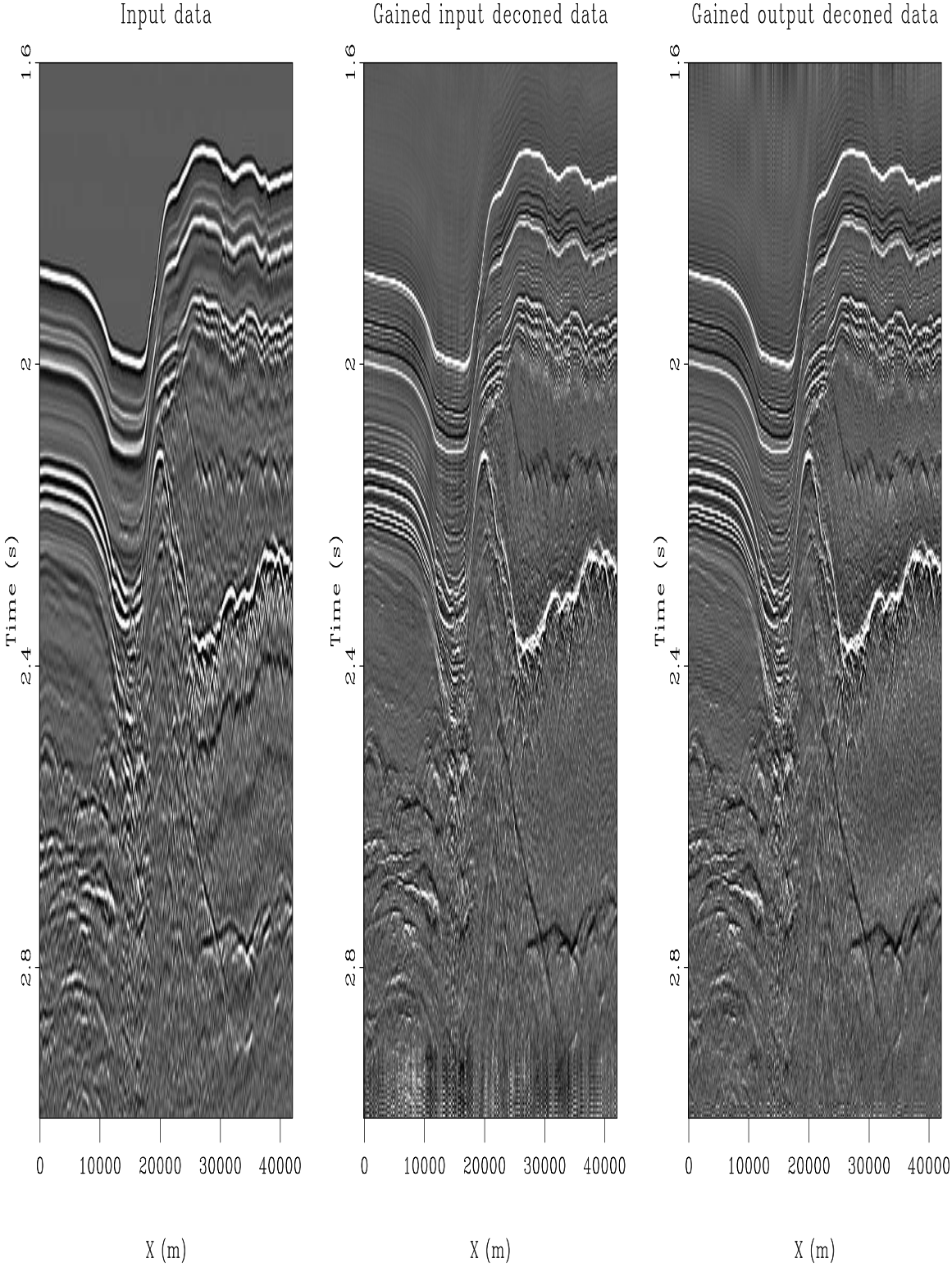


Figure 1: Gulf of Mexico. Decon produces plain white reflections from hard boundaries, and plain black boundaries from soft ones. WB= Water Bottom (white), TS= Top Salt (white), BS= Bottom Salt (black), ME= Mystery Event (black), soft reflector could be rugose salt solution of a former salt layer. [**ER**] jon2/. Comp-gom-50

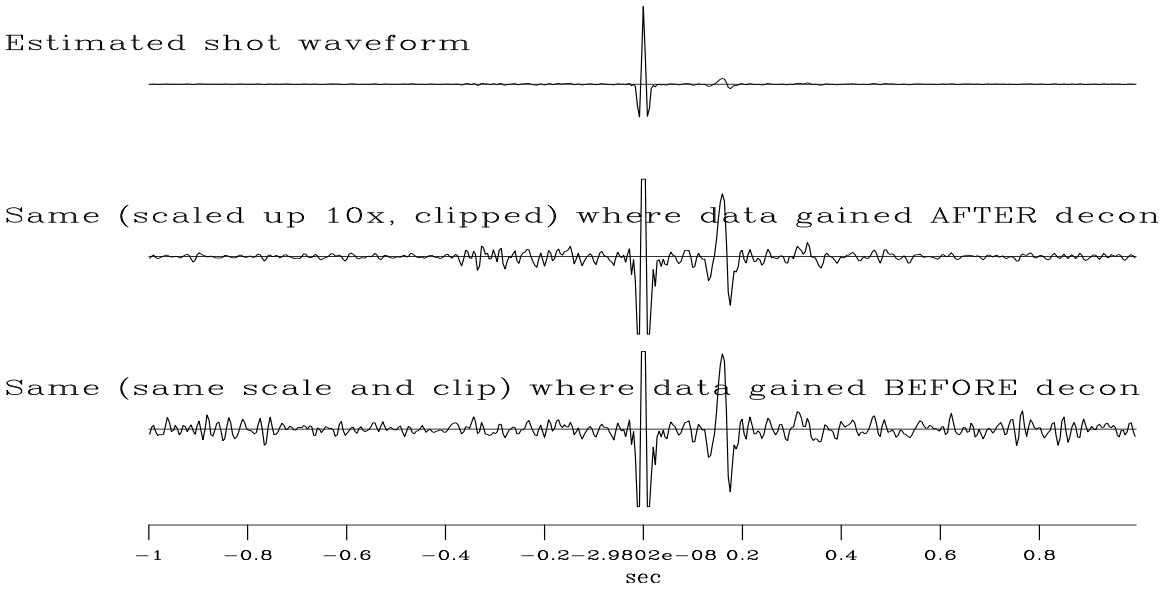


Figure 2: Wavelet from Cabo data. Although causality is not imposed, the estimated shot waveform is near causal (discounting the leading lobe of the Ricker wavelet). The importance of gain (here t^2) after deconvolution instead of before is shown by the lower two traces. There is much less noise when we gain AFTER decon, not BEFORE. Notice also that gain before decon estimates a slightly larger bubble (which is wrong). [**ER**] jon2/. trio

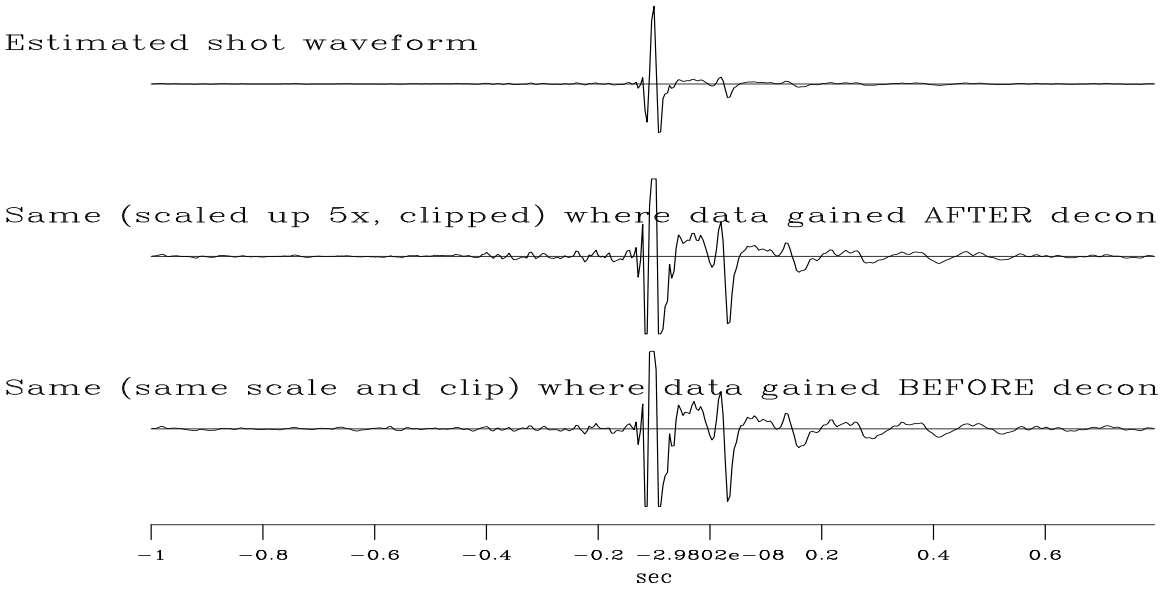


Figure 3: The Gulf of Mexico data produces a very different bubble but the same conclusions as Figure 2. The lack of symmetry in the Ricker wavelet may be related to the unresolved uniqueness issue. (Awaits better regularization.) [**ER**] jon2/. trio2

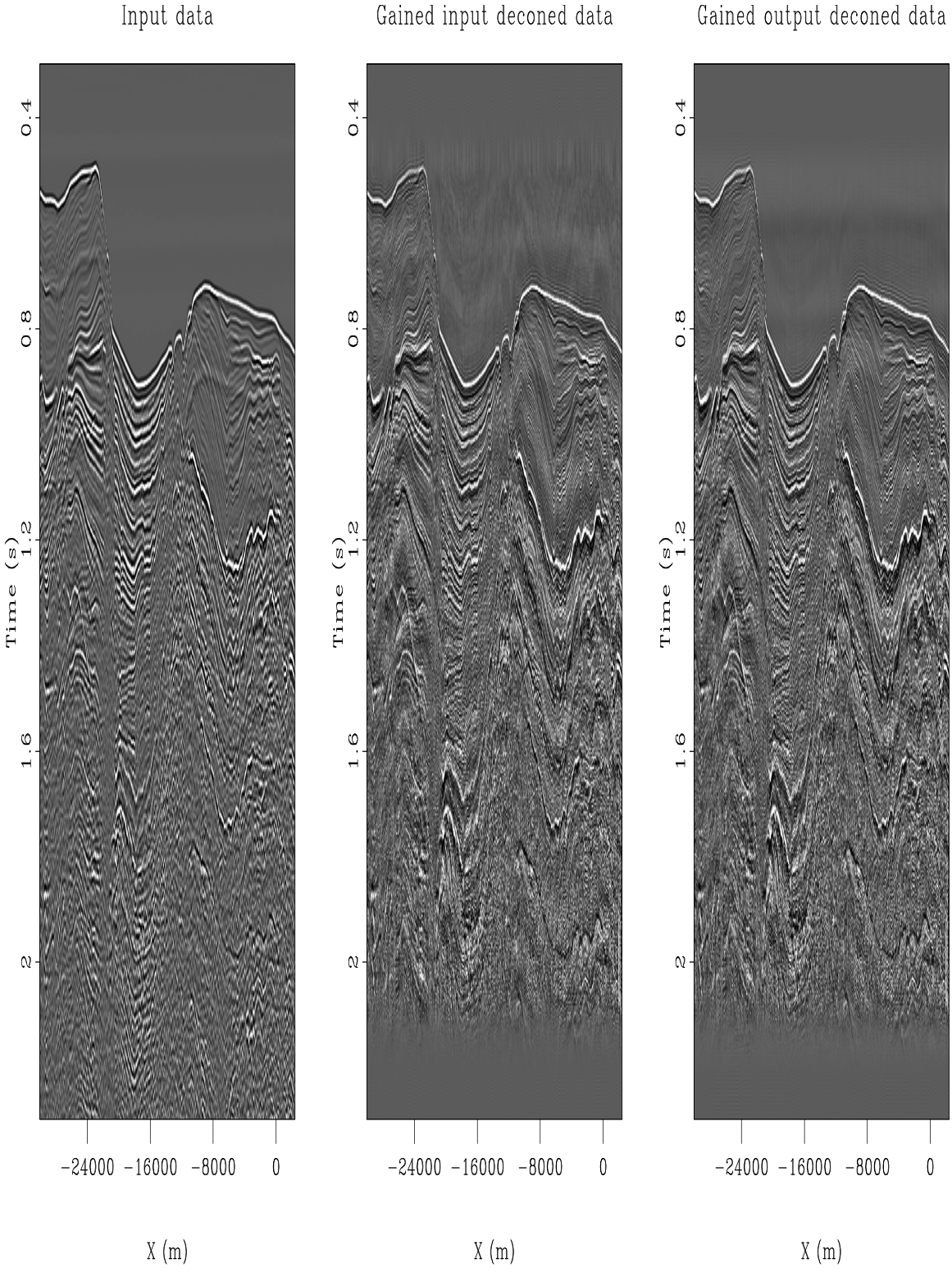


Figure 4: Cabo data. One filter on all traces. Bubble (at about 0.9s) removed. Enhanced high frequency at 1.1s. Gained-input method gave low frequency event precursors especially clear above the event at 1.2s but also visible above the water bottom. The problem is overcome by the gained-output method. (Guitton) [**ER**] jon2/. Comp-cabo-50

Integral operator quality from low order interpolation
or
Sometimes nearest neighbor beats linear interpolation

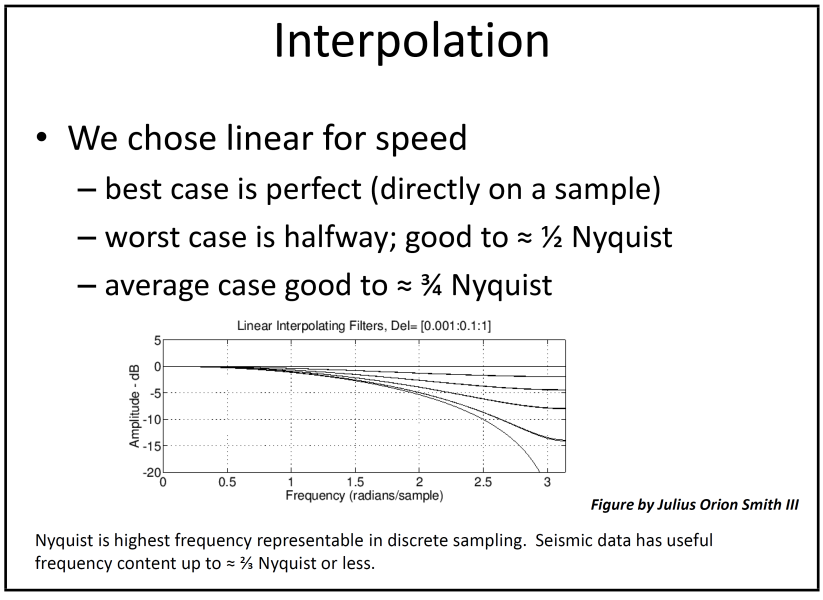
Stewart A. Levin

ABSTRACT

In most discussions of interpolation methods, it is the worst-case behavior that dominates the analysis. From a systems point of view, one really should analyze how that interpolation is used in producing an end product in order to determine the interpolation's suitability. In this report I look at the summation operators slant stack, NMO and Kirchhoff migration as the "systems" and determine that their output quality can be significantly better than the traditional take on interpolation would suggest. In one scenario, I even found nearest neighbor interpolation did the job even better than linear interpolation.

INTRODUCTION

Many geophysical imaging and analysis tools involve summation over various trajectories within their input data. It is rare indeed that such summations do not involve intermediate interpolation among adjacent or nearby samples. Such interpolations are generally imperfect and so can degrade the final result—a subject of interest to me for some time (Levin, 1994). In a recent broad brush presentation of Kirchhoff time migration to an introductory C++ class at Stanford, I presented the slide:



and asserted

> "Normally, an interpolator is selected based on its guaranteed-to-be-no-worse-than fidelity. However in the present case the distribution of many thousands or even millions of fractions of a sample to interpolate is uniform and so I focused in on the average rather than the worst case. From this perspective, looking at the amplitude behavior halfway between the 3rd and 4th curves plotted in the slide, the quite inexpensive linear interpolation does a good enough job in retaining fidelity. A similar plot of phase accuracy yields the same conclusion."

To understand this concept, let us take the very simple case of a single function replicated with linear delay and sampled with unit spacing as illustrated in Figure 1. Applying linear moveout and stacking would ideally reproduce that function. In order to apply the linear moveout we need to interpolate. Let us first consider what happens if we use simple nearest neighbor interpolation.

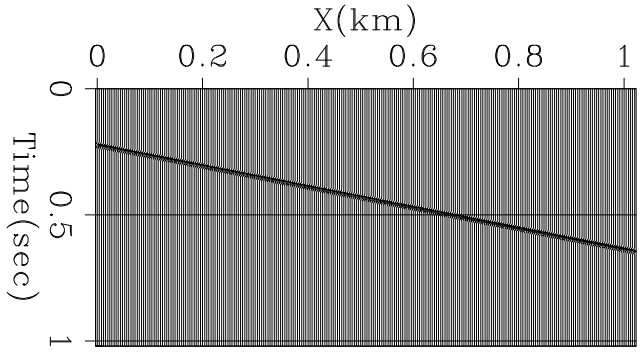


Figure 1: Simple dipping synthetic section. [**ER**] stew1/. simpledip

We generally abhor nearest neighbor interpolation because it (a) is a discontinuous function of position and (b) can yield interpolated values that don't even have the correct sign. When stacking is included, though, to a good approximation we may assume the interpolation points are randomly distributed between samples following a uniform distribution. This means that the output values on our stack at x_0 are approximately

$$\int_{x_0-1/2}^{x_0+1/2} f(x)dx$$

which is convolutional filtering of $f(x)$ by a boxcar having Fourier transform

$$F(\omega) = \text{sinc}(\omega/2) \quad . \tag{1}$$

Figure 2 illustrates the filter and its spectrum. Note that within the Nyquist limits $-\pi$ to π, the spectrum (1) is positive, indeed greater than or equal to $2/\pi$. Therefore we can compensate for the nonflat spectrum by simple zero-phase spectral rebalancing after the stack, avoiding the expense of costly high quality sinc-like convolutions before stack.

Seeing that nearest neighbor interpolation works unexpectedly well, let us turn our attention to linear interpolation. Following the same line of reasoning leads to a triangular convolution illustrated in Figure 3 which is a convolution of the nearest neighbor boxcar with itself and hence a spectrum that is the square of (1). This linear interpolation result is *poorer* than nearest neighbor, requiring more post-stack spectral compensation and concomitant noise magnification at higher frequencies!

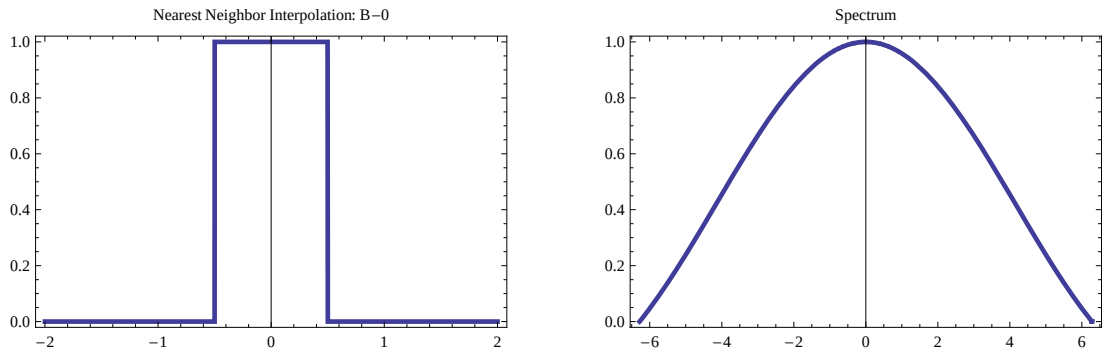


Figure 2: Nearest neighbor boxcar filter (left) and its spectrum (right). Figure from Fomel (2000). [**NR**] stew1/. nnint

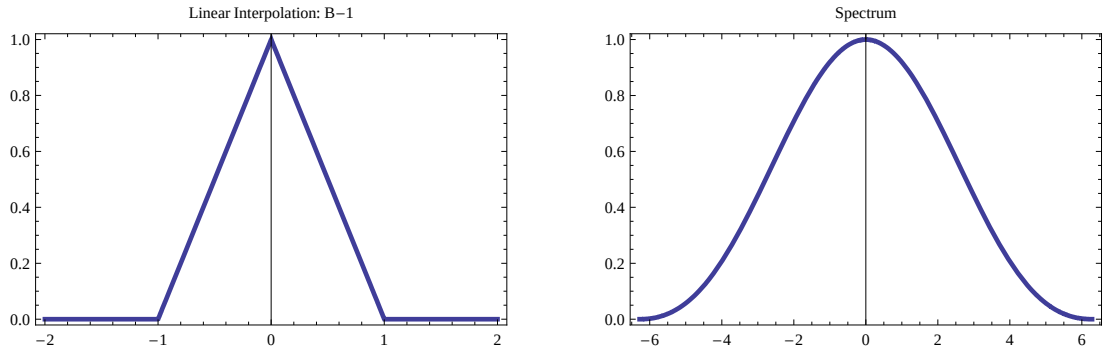


Figure 3: Nearest neighbor boxcar filter (left) and its spectrum (right). Figure from Fomel (2000). [**NR**] stew1/. linint

EXAMPLES

For my first example I created a dipping planar synthetic with a Ricker wavelet on it. (Figure 4) This 26.556° dip is at an azimuth of 67.44° and the central frequency of the Ricker wavelet is 64.22 Hz, slightly more than half the Nyquist frequency. The wavelet was imposed by calculating the analytic Ricker formula directly onto the trace samples and not by post-convolving a spike trace synthetic. I then did an areal slant stack along the same

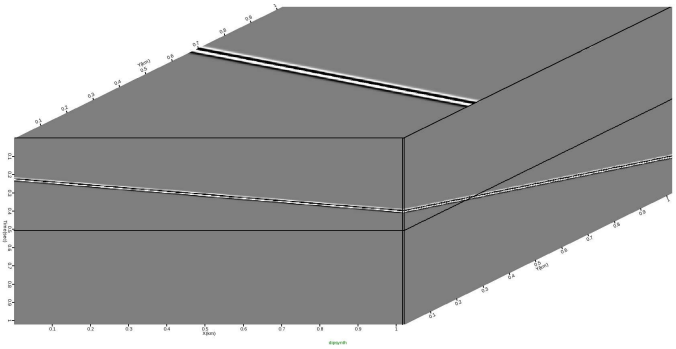


Figure 4: Dipping 3D synthetic used for slant stack test. [**NR**] stew1/. dipsynth

dip and azimuth as the plane with the output at the location of the first inline and crossline trace. I used nearest neighbor interpolation for this slant stack. Figure 5 shows the result overlain with what a perfect slant stack would have produced. Quite a respectable result even without spectral reshaping. (When spectral compensation is included, the result of nearest neighbor interpolation is nearly perfect as shown in Fig. 7.) Figure 6 shows the corresponding output for linear interpolation which, indeed, is not as accurate, agreeing with the analysis above.

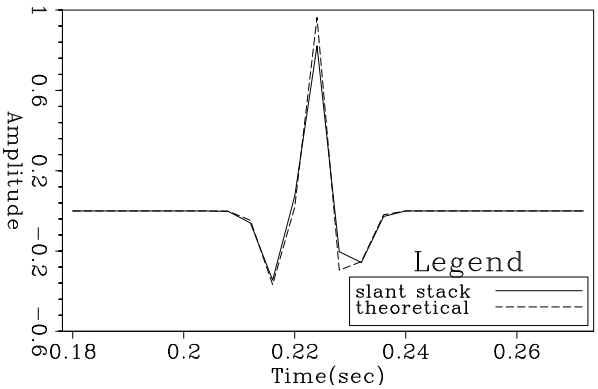


Figure 5: Comparision of the perfect theoretical result with the output of slant stacking the data in Fig. 4 parallel to the planar event using nearest-neighbor interpolation. [**ER**] stew1/. sstacknn

Of course, field data have many events, not just one, and summation operations will generally involve both constructive and destructive summation. Taking a first look at this topic, I modified the planar synthetic to have about 75 uniformly distributed individual parallel plane arrivals with amplitudes constructed according to a Gaussian distribution. (Figure 8.) Using this synthetic, I deliberately slant stacked horizontally (Fig. 10 which requires no interpolation) and again at a dip that was not parallel to the plane of the synthetic. (For Fig. 11 I opted to interchange the X and Y axis dips.) It is clear that nearest-neighbor interpolation created just about as much destructive cancellation as one would expect from a high order interpolation. Nearest neighbor slant stack parallel to that

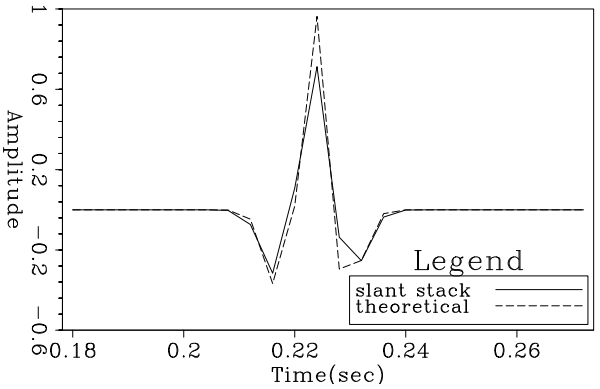


Figure 6: Comparision of the perfect theoretical result with the output of slant stacking the data in Fig. 4 parallel to the planar event using linear interpolation. [**ER**] stew1/. sstacklin

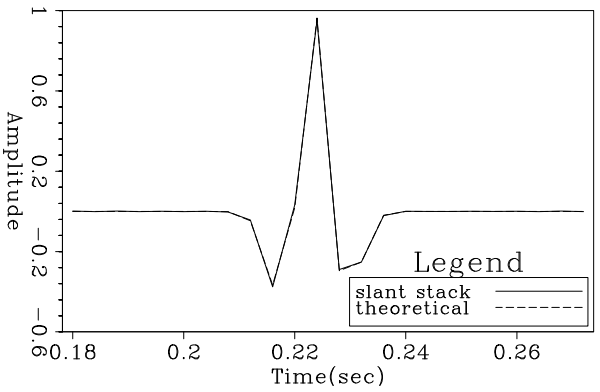


Figure 7: Comparision of the perfect theoretical result with the output of slant stacking the data in Fig. 4 parallel to the planar event using nearest-neighbor interpolation followed by spectral compensation with the reciprocal of formula 1. [**ER**] stew1/. sstackspectnn

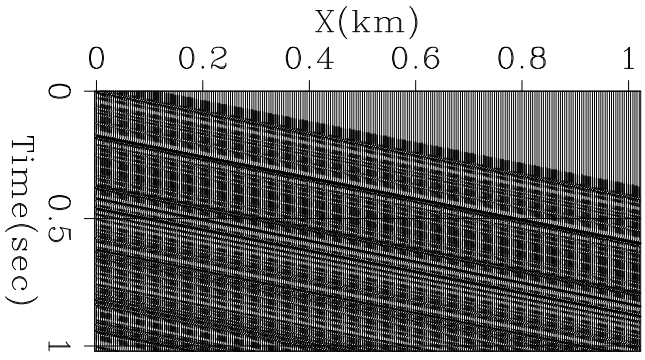


Figure 8: First slice of dipping 3D synthetic with sparse reflectors and Gaussian amplitude distribution used for slant stack test. [**ER**] stew1/. simpleranddip

dip yielded the expected high quality result show in Figure 9. Stacking horizontally, which requires no interpolation, or (destructive) slant stacking along a plane oriented with the X and Y axis interchanged resulted in equally small output amplitudes (a fraction of a percent) as expected.

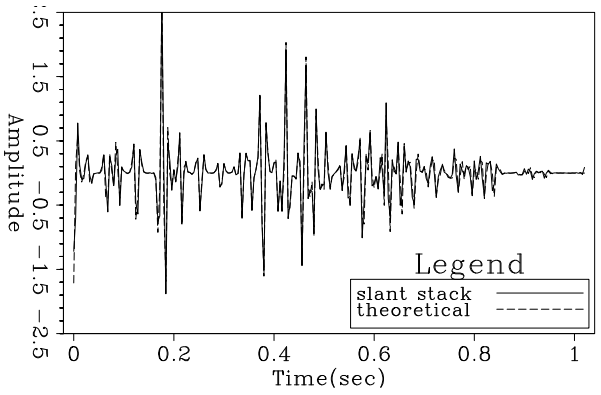


Figure 9: Comparision of nearest neighbor and high order interpolation slant stacking the data in Fig. 8 parallel to the planar event using nearest-neighbor interpolation followed by spectral compensation with the reciprocal of 1. [**ER**] stew1/. sstackrandnn16

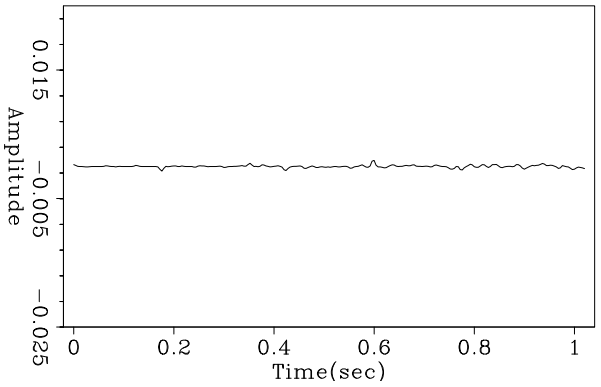


Figure 10: Horizontal stack of the data in Fig. 8 gained by a factor of 100 over that in Fig. 9. [**ER**] stew1/. hstackrandnn

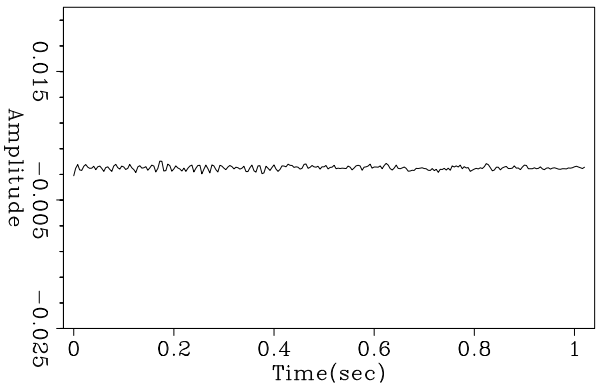


Figure 11: Slant stack over a non-parallel direction of the data in Fig. 8 gained by a factor of 100 over that in Fig. 9. [**ER**] stew1/. dstackrandnn

My second example is another situation where summation is expected to be almost exclusively constructive: 3D normal moveout. Figure 12 shows the result of applying NMO stack to a synthetic (flat dip) 3D CDP with maximum offsets of 12.8 and 25.6 km in the X and Y directions respectively. No spectral reshaping is shown because those results almost perfectly overlay their inputs. This anomaly arises because of NMO stretch which changes the shapes of the waveforms being stacked. In field data statics, deconvolution, and AVO

effects will further modify the input waveforms, making spectral compensation analysis complicated and data dependent.

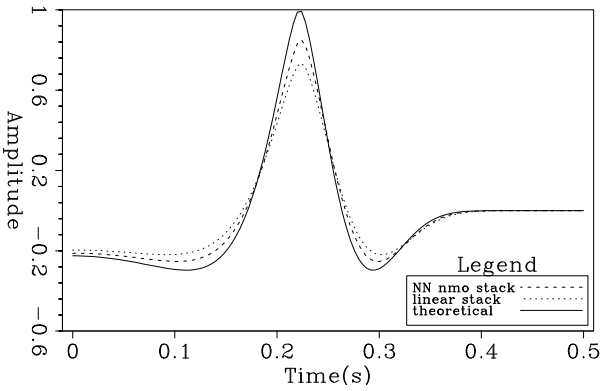


Figure 12: NMO stack comparison of nearest-neighbor, linear and high order interpolaton. Here no spectral compensation has been applied. [**ER**] stew1/. NMO3dcombo

Closely related to NMO stack is common reflection surface (CRS) imaging which estimates and sums over prestack specular reflection surfaces (Jäger et al., 2001; Hertweck et al., 2004). I generated and CRS processed a prestack synthetic for a dipping plane and got the result in Figure 13. With the exception of an apparent one sample shift, probably a small coding bug, again nearest-neighbor did a fine job and outperformed linear linear interpolation.

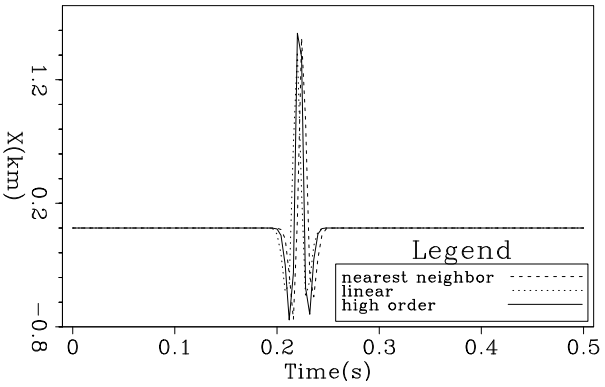


Figure 13: CRS zero-offset trace for a dipping planar synthetic using nearest neighbor, linear and high-order interpolation. [**ER**] stew1/. Prestack3Dcombo

Transitioning further into the balance between constructive and destructive interference, for a fourth example, I poststack migrated the synthetic of Fig. 4 with the same constant velocity used to create it. Using nearest-neighbor interpolation, the output at the first inline and crossline trace location is shown in Figure 14. Linear interpolation produces Figure 15 and highly accurate interpolation produces 16. In this case, increasing the order of the interpolant does increase accuracy, however the level of residual artifact is disconcerting.

The fundamental difference between the last example and the previous ones is that poststack Kirchhoff migration sums only a few samples coherently to form an image point sample. With prestack Kirchhoff, however, the number of coherent samples forming an image point sample is much larger. Figures 17 and 18 show 3D prestack Kirchhoff migration of a horizontal planar event. (Couldn't get the dipping plane version working in time.) There is no longer annoying precursor noise in these results.

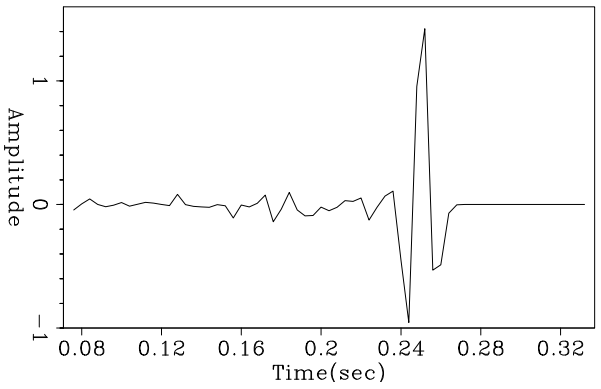


Figure 14: Poststack 3D Kirchhoff time migration of the dipping planar synthetic using just nearest-neighbor interpolation. [**ER**] stew1/. mig3Dnn

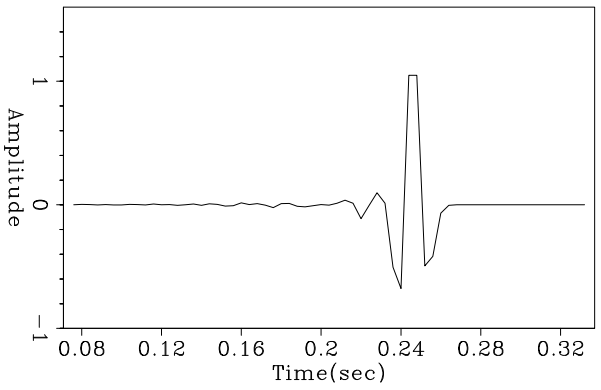


Figure 15: Poststack 3D Kirchhoff time migration of the dipping planar synthetic using linear interpolation. [**ER**] stew1/. mig3Dlin

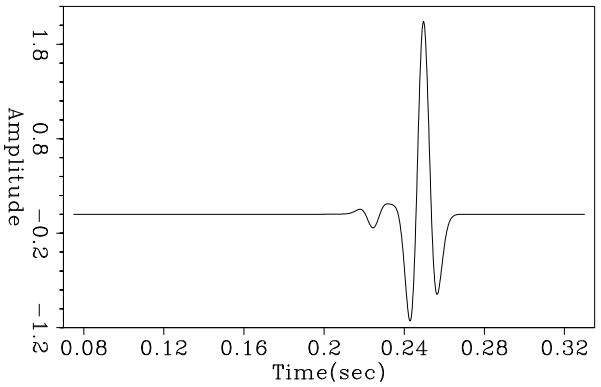


Figure 16: Poststack 3D Kirchhoff time migration of the dipping planar synthetic using high-order interpolation. [**ER**] stew1/. mig3Dnn16

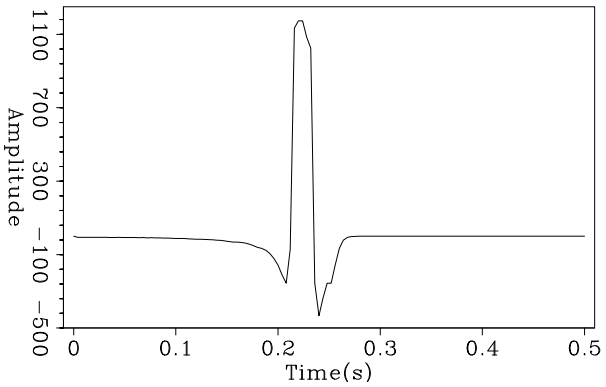


Figure 17: Prestack 3D Kirchhoff time migration of a horizontal planar synthetic using just nearest-neighbor interpolation. No attempt was made to taper aperture or apply phase corrections. [**CR**] stew1/. Kirchnn

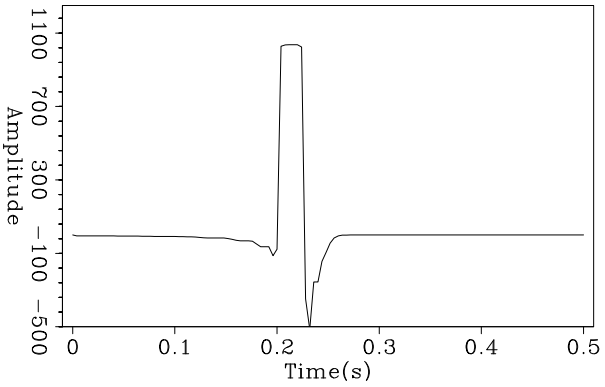


Figure 18: Poststack 3D Kirchhoff time migration of a horizontal planar synthetic using linear interpolation. No attempt was made to taper aperture or apply phase corrections. [**CR**] stew1/. Kirchlin

DISCUSSION AND CONCLUSIONS

We have clearly seen a significant difference between slant stacking (or NMO stacking) and migration in how interpolation affects the result. This is a reflection (no pun intended) on the fact that the former is adding signals in phase while the latter is relying on massive destructive interference to suppress those data that shouldn't contribute to the output. (See, e.g., Levin (2004).) The constructive interference in the 3D poststack migration example is over a fairly small patch and has led to some unwanted noise and other artifacts. However full prestack Kirchhoff migration suppresses the noise and confirms the original conclusion that it is the average case, not the worst case, of interpolation that should be the focus of integral operator system analysis.

REFERENCES

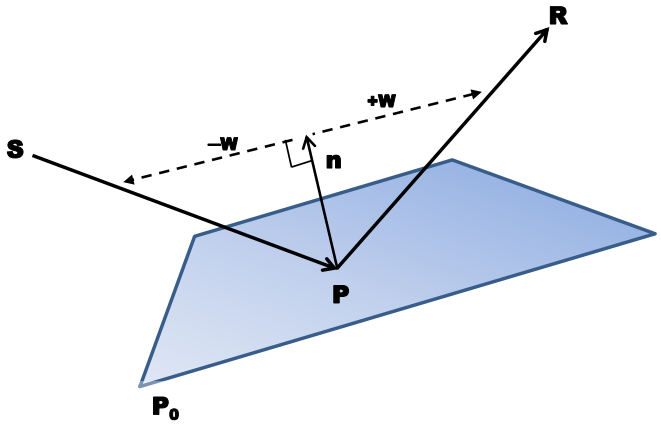
Fomel, S., 2000, Inverse B-spline interpolation: SEP-Report, **105**, 79–108.

Hertweck, T., C. Jäger, J. Mann, E. Duveneck, and Z. Heilmann, 2004, A seismic reflection imaging workflow based on the Common-Reflection-Surface (CRS) stack: Theoretical background and case study: SEG Technical Program Expanded Abstracts, **23**, 2032–2035.

Jäger, R., J. Mann, G. Höcht, and P. Hubral, 2001, Common-reflection-surface stack: Image and attributes: Geophysics, **66**, 97–109.

Levin, S. A., 1994, Stolt without artifacts? — dropping the Jacobian: SEP-Report, **80**, 513–532.

———, 2004, Numerical precision in 3D prestack Kirchhoff migration: SEG Technical Program Expanded Abstracts, **23**, 1119–1122.



Given a source location $\boldsymbol{S}$, a receiver location $\boldsymbol{R}$ and the plane $\boldsymbol{n}\cdot(\boldsymbol{P}-\boldsymbol{P_0})$=0, where $\boldsymbol{n}$ is a unit normal, to find the reflection point $\boldsymbol{P}$, drop a perpendicular $\boldsymbol{w}$ from $\boldsymbol{n}$ to the line connecting $\boldsymbol{P}$ to $\boldsymbol{R}$. Snell's Law says that running $\boldsymbol{w}$ in the other direction connects to the line between $\boldsymbol{P}$ and $\boldsymbol{S}$. So we have:

$$\begin{array}{rcl}(\boldsymbol{R}-\boldsymbol{P}) & = & \alpha(\boldsymbol{n}+\boldsymbol{w}) \\ (\boldsymbol{S}-\boldsymbol{P}) & = & \beta(\boldsymbol{n}-\boldsymbol{w}) \\ \boldsymbol{n}\cdot\boldsymbol{w} & = & 0 \\ \boldsymbol{n}\cdot\boldsymbol{P} & = & \boldsymbol{n}\cdot\boldsymbol{P_0}\end{array} .$$

Dotting $\boldsymbol{n}$ onto the first two equations gives

$$\begin{array}{rcccl}\boldsymbol{n}\cdot(\boldsymbol{R}-\boldsymbol{P}) & = & \boldsymbol{n}\cdot(\boldsymbol{R}-\boldsymbol{P_0}) & = & \alpha \\ \boldsymbol{n}\cdot(\boldsymbol{S}-\boldsymbol{P}) & = & \boldsymbol{n}\cdot(\boldsymbol{S}-\boldsymbol{P_0}) & = & \beta\end{array} ,$$

and subtracting the first two equations produces

$$(\boldsymbol{R}-\boldsymbol{S}) \;=\; (\alpha-\beta)\boldsymbol{n}+(\alpha+\beta)\boldsymbol{w} \;,$$

which we can solve directly for $\boldsymbol{w}$. Given this $\boldsymbol{w}$, the first equation immediately yields

$$\boldsymbol{P} \;=\; \boldsymbol{R}-\alpha(\boldsymbol{n}+\boldsymbol{w}) \;,$$

the desired reflection point.

Recent progress regarding logarithmic Fourier-domain bidirectional deconvolution

Qiang Fu

ABSTRACT

Bidirectional deconvolution in the Fourier domain is a new method of removing the mixed phase wavelet from seismic data. I demonstrate that this is self-preconditioned, therefore a scheme that has a preconditioner in the logarithmic Fourier-domain deconvolution is not necessary. I show a simple synthetic test case which incorporates a gain function into the deconvolution method.

INTRODUCTION

Usually, a seismic data trace d can be defined as a convolution of a wavelet w with a reflectivity series r. This can be written as $d = r * w$, where $*$ denotes convolution. Blind deconvolution seeks to estimate the wavelet and reflectivity series using only information contained in the data. Previously, seismic blind deconvolution has used two assumptions, namely whiteness and minimum phase. The whiteness assumption supposes that the reflectivity series r has a flat spectrum, while the minimum-phase assumption supposes that the wavelet w is causal and has a stable inverse. Recently, some new methods have been proposed to limit the effect of these two assumptions in seismic blind deconvolution.

In Zhang and Claerbout (2010a), the authors proposed to use a hyperbolic penalty function introduced in Claerbout (2009) instead of the conventional L2 norm penalty function to solve the blind deconvolution problem. With this method, a sparseness assumption replaces the traditional whiteness assumption. Furthermore, Zhang and Claerbout (2010b) proposed a new method called "bidirectional deconvolution" in order to overcome the minimum-phase assumption. Bidirectional deconvolution assumes that any mixed-phase wavelet can be decomposed into a convolution of two parts, $w = w_a * w_b$, where w_a is a minimum-phase wavelet and w_b is a maximum-phase wavelet. To solve this problem, they estimated two deconvolution filters, a and b, which are the inverses of wavelets w_a and w_b, respectively. Since Zhang and Claerbout (2010b) solved the two deconvolution filters a and b alternately, we call this method the slalom method. Shen et al. (2011a) proposed another method to solve the same problem. They used a linearized approximation to solve the two deconvolution filters simultaneously. We call this method the symmetric method. Fu et al. (2011a) proposed a way to choose an initial solution to overcome the local-minima problem caused by the high nonlinearity of blind deconvolution. Shen et al. (2011b) discussed an important aspect of any inversion problem: preconditioning and how it improves bidirectional deconvolution. All of the aforementioned methods solved the problem in the time domain. Claerbout et al. (2011) proposed a new logarithmic, Fourier-domain, bidirectional deconvolution to solve the same problem. Fu et al. (2011b) showed that this new method converges faster than the above-mentioned time-domain methods

In this paper, I will attempt to answer an important question: Should we use a preconditioner with the logarithmic Fourier-domain bidirectional deconvolution? I will then show an example of including a gain function in this method.

NO NEED FOR PRECONDITIONING

Previous time-domain implementations of bidirectional deconvolution (Shen et al. (2011a) and Fu et al. (2011a)) have required preconditioning for both stabilizing the deconvolution result and accelerating the convergence speed. For those methods, we used a prediction error filter (PEF) as our preconditioner. However, using the PEF introduced a polarity change and a time shift problem, because it caused a spike in the first lobe of the Ricker wavelet. However, in the new Fourier-domain method,

$$\begin{aligned} r_{new} &= FT^{-1}(De^{U_{new}}) \\ &= FT^{-1}(De^{U+\alpha\Delta U}) \\ &= r * FT^{-1}(e^{\alpha\Delta U}), \end{aligned} \tag{1}$$

where r is the residual, D is the data (the uppercase letter indicates a Fourier-domain variable), and U is the logarithmic parameterization of the deconvolution filter.

I find that within the iterations, the problem is self-preconditioned. The new update of the u parameter is convolved with the previous residual, so we do not need the PEF preconditioner. Figures 1 through 3 show the comparison between the deconvolution results on a common-offset gather of marine data with and without the PEF preconditioner. In this comparison, I use the 99.5 percentile of all residuals as the threshold for the hyperbolic penalty function. Figure 1 shows the common-offset data gather. Figure 2 shows the deconvolution results with and without the PEF preconditioner. Figure 3 shows the estimated wavelets with and without preconditioning. From this comparison, we can see clearly that the quality of the result without the PEF preconditioner is not inferior to the result with the PEF preconditioner. In addition, the result without the preconditioning avoids the polarity change and time shift. This is caused when the PEF preconditioning introduces an unwelcome initial solution into the deconvolution scheme. This is most obvious in the comparison of the estimated wavelets in Figure 3.

We found that with slightly different initial solutions or parameters of the deconvolution scheme, we could get quite significantly different results. This could be caused either by the nonlinearity or by the null space of the inversion problem; can not comfirm which is the reason for this phenomenon. Claerbout et al. (2012) claim that the null space may be the major reason, and that therefore a regularization is helpful and may be necessary to solve this problem. However it is still early to draw this conclusion, and more tests are still needed to answer this question.

SYNTHETIC DATA EXAMPLE FOR GAIN FUNCTION

In real field data, there usually is an amplitude decay with time. This can be caused by geometric divergence and other factors. The larger the amplitude in the data, the larger the residual. Therefore, the deconvolution will honor early data more than later data. Hence we need a gain function in the deconvolution to boost the contribution of later data.

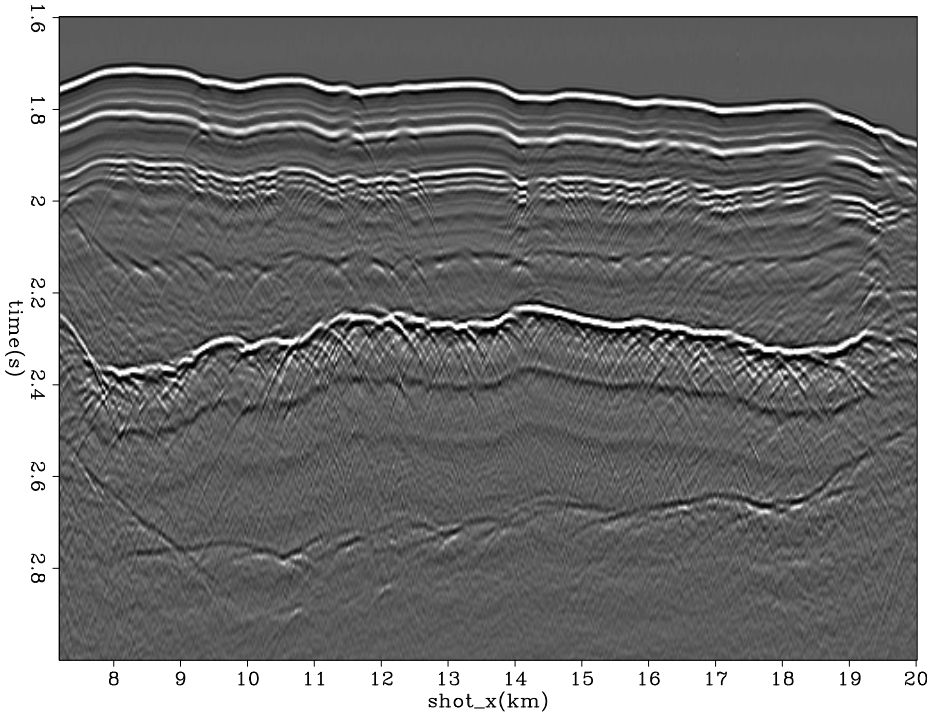


Figure 1: A common-offset section of a Gulf of Mexico data set. [**ER**] qiang/. fig-1

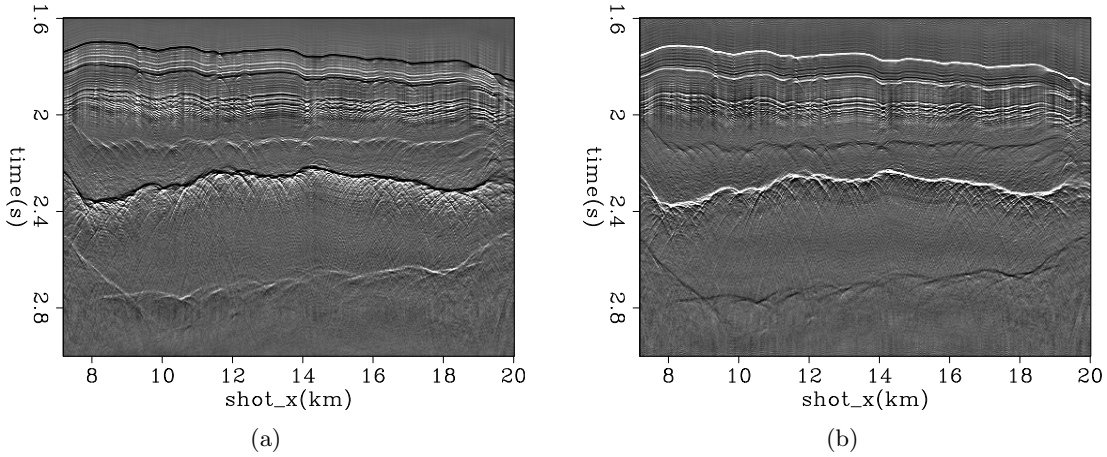


Figure 2: Deconvolution results (a) with and (b) without PEF preconditioning. [**ER**] qiang/. fig-2a,fig-2b

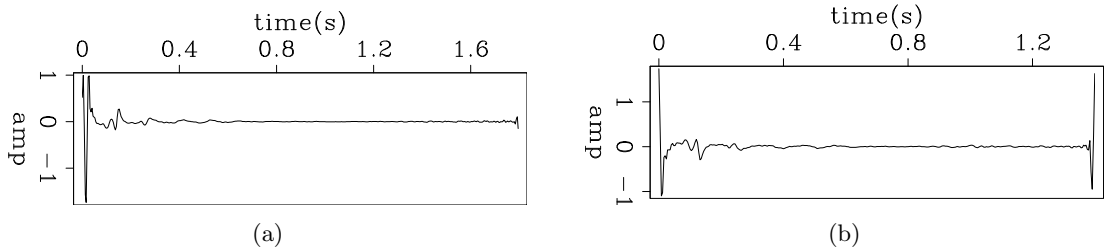


Figure 3: Estimated wavelets (a) with and (b) without PEF preconditioning. [**ER**] qiang/. fig-3a,fig-3b

Claerbout et al. (2012) provide the complete, step-by-step derivation of the gain function in the Fourier-domain bidirectional deconvolution approach. I test the this gain method on a simple synthetic data example. I use a five-trace simple spiky reflectivity model convolved with a zero-phase synthetic wavelet to make the synthetic data. In particular, I put a dipole at the beginning of the fifth trace of the model. If the deconvolution honors the beginning more than the end of the traces, it will tend to make one spike rather then a dipole at that location in the output. On the other hand, if the deconvolution does not incorrectly emphasize the beginning of the traces, we will get a dipole back. Figure 4 shows the synthetic model and wavelet, and Figure 5 shows the synthetic data without and with decay. Figure 6 shows the recovered wavelet and model by deconvolution without gain. Because there is no decay in the data, the result is nearly perfect. Then I add a decay to the model, proportional to the time squared. Figure 7 shows the recovered wavelet and model by deconvolution of this decaying data without gain. So that the decayed end of the traces would be evident, I applied a time-squared scale on the time axis in Figure 7(b). The results are poor; there is only a spike rather than a dipole at the beginning of the fifth trace in the output. However, when I include the gain function in the deconvolution following the implementation discussed above, I get the results shown in figure 8. I also applied a time-squared scale on the time axis in Figure 7(b). By including the gain function in the deconvolution to correctly balance the amplitudes, I can get nearly perfect results again.

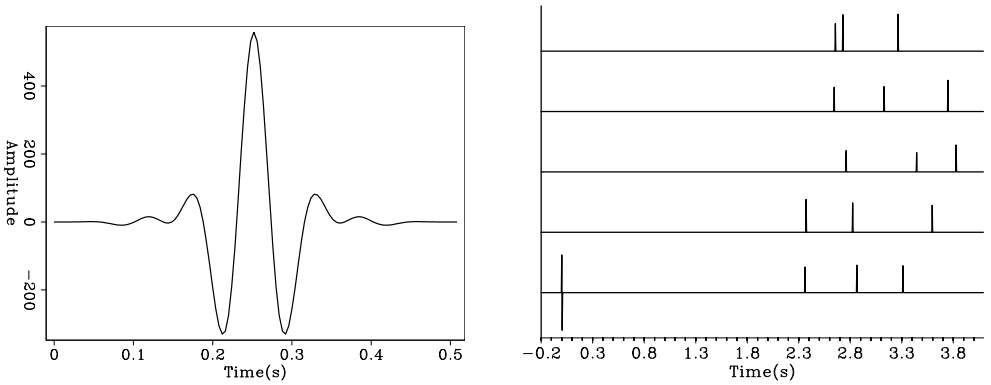
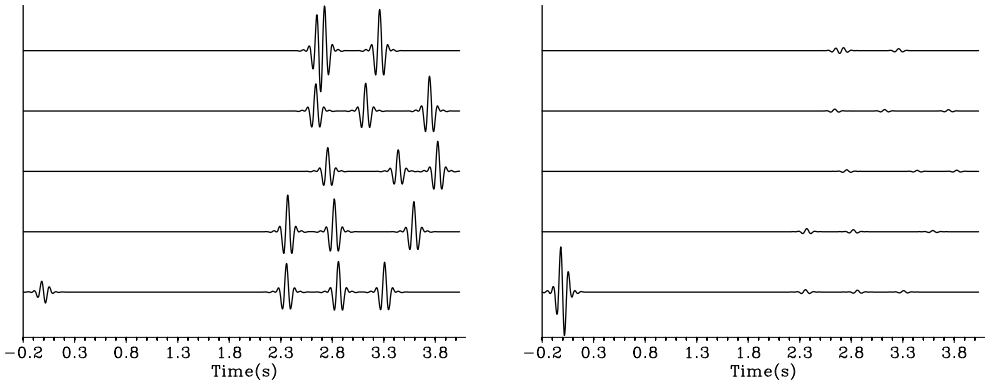
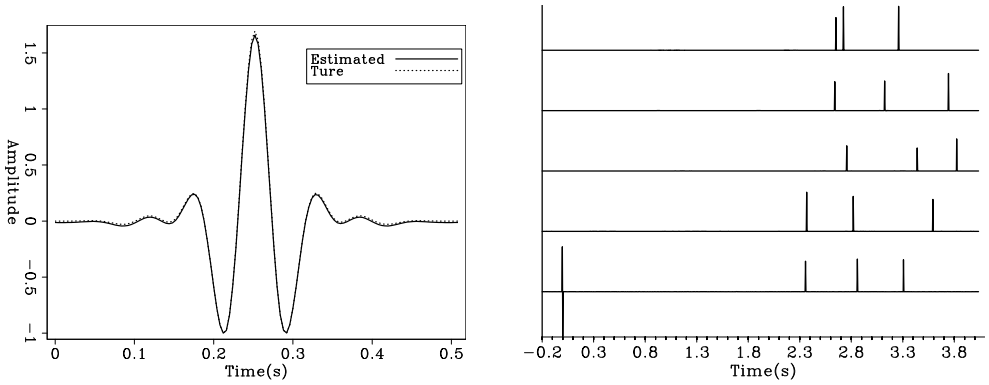


Figure 4: Wavelet (a) and five-trace model (b) used for synthetic example. [**ER**] qiang/. fig-8a,fig-8b



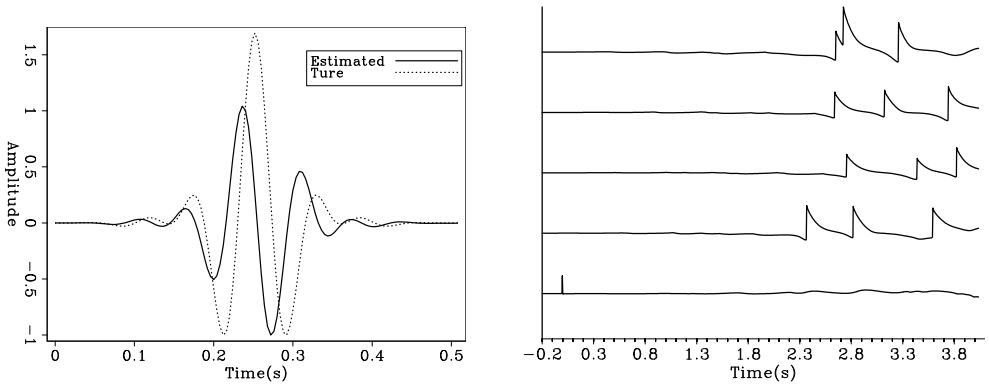
(a) (b)

Figure 5: Five-trace synthetic data set without (a) and with decay proportional to the time squared (b). [**ER**] qiang/. fig-85a,fig-85b



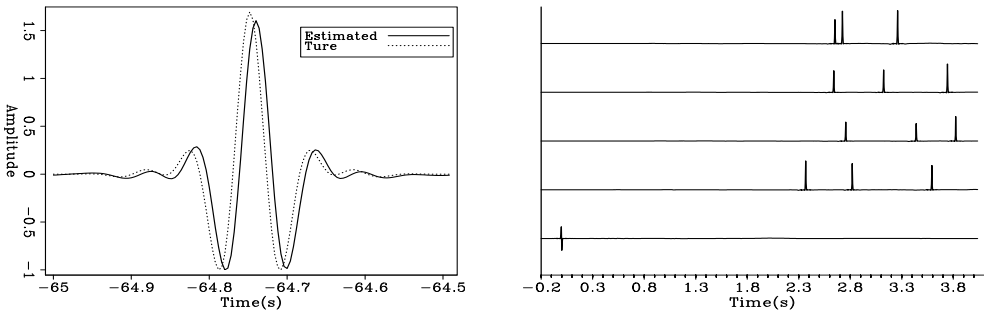
(a) (b)

Figure 6: Estimated wavelet (a) and recovered result (b). [**ER**] qiang/. fig-9a,fig-9b



(a) (b)

Figure 7: Estimated wavelet (a) and recovered result (b) of decay data without gain. For easily see the decayed end of the traces, a time squared scale are applied on the time axis in the result (b). [**ER**] qiang/. fig-10a,fig-10b



(a) (b)

Figure 8: Estimated wavelet (a) and recovered result (b) of decay data with gain. For easily see the decayed end of the traces, a time squared scale are applied on the time axis in the result (b). [**ER**] qiang/. fig-11a,fig-11b

CONCLUSIONS

I showed that the logarithmic Fourier-domain bidirectional deconvolution is self-preconditioned, so that no extra perconditioner is needed. Also I showed a synthetic example of including gain function in this deconvolution scheme. These work improve our understanding and implementation of Fourier-domain bidirectional deconvolution.

ACKNOWLEDGMENTS

I would like to thank Jon Claerbout, Yang Zhang, Shuki Ronen, Antoine Guitton, Robert Clapp and Dave Nichols for helpful discussions of our research. I also thank the sponsors of the Stanford Exploration Project for their financial support.

REFERENCES

Claerbout, J., 2009, Blocky models via the L1/L2 hybrid norm: SEP-Report, **139**, 1–10.

Claerbout, J., Q. Fu, and Y. Shen, 2011, A log spectral approach to bidirectional deconvolution: SEP-Report, **143**, 297–300.

Claerbout, J., A. Guitton, and Q. Fu, 2012, Decon in the log domain with variable gain: SEP-Report, **147**, 313–322.

Fu, Q., Y. Shen, and J. Claerbout, 2011a, An approximation of the inverse Ricker wavelet as an initial guess for bidirectional deconvolution: SEP-Report, **143**, 283–296.

———, 2011b, Data examples of logarithm Fourier-domain bidirectional deconvolution: SEP-Report, **145**, 101–116.

Shen, Y., Q. Fu, and J. Claerbout, 2011a, A new algorithm for bidirectional deconvolution: SEP-Report, **143**, 271–282.

———, 2011b, Preconditioning a non-linear problem and its application to bidirectional deconvolution: SEP-Report, **145**, 117–130.

Zhang, Y. and J. Claerbout, 2010a, Least-squares imaging and deconvolution using the hybrid norm conjugate-direction solver: SEP-Report, **140**, 129–142.

———, 2010b, A new bidirectional deconvolution method that overcomes the minimum phase assumption: SEP-Report, **142**, 93–104.

SEP PHONE DIRECTORY

Name	Phone	Login Name
Almomin, Ali	723-0463	ali
Barak, Ohad	723-9282	ohad
Biondi, Biondo	723-1319	biondo
Chang, Jason	723-3187	jason
Claerbout, Jon	723-3717	jon
Clapp, Bob	725-1334	bob
de Ridder, Sjoerd	723-1250	sjoerd
Farghal, Noha	723-1250	nfarghal
Fu, Qiang	723-1250	qiang
Halpert, Adam	723-6006	adam
Lau, Diane	723-1703	diane
Leader, Chris	723-0463	chrisl
Levin, Stewart	726-1959	stew
Li, Elita	723-9282	myfusan
Maharramov, Musa	725-1625	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, 2011-2012

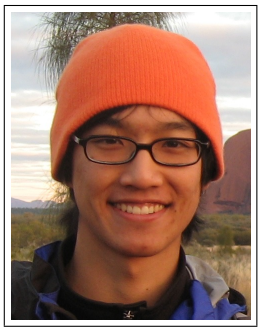
Name	**Company**	**Tel #**	**E-Mail**
Raymond Abma	BP	(281)366-4604	abmar1@bp.com
Biondo Biondi	SEP	(650)723-1319	biondo@sep.stanford.edu
Robert Bloor	ION/GX Technology	(281)781-1141	robert.bloor@iongeo.com
Jon Claerbout	SEP	(650)723-3717	jon@sep.stanford.edu
Thomas Dickens (Co-chair, 2nd year)	ExxonMobil	(713)431-6011	tom.a.dickens@exxonmobil.com
Yi Luo	Saudi Aramco	–	yi.luo@aramco.com
Alejandro Valenciano (Co-chair, 1st year)	PGS	–	alejandro.valenciano@pgs.com

Research Personnel

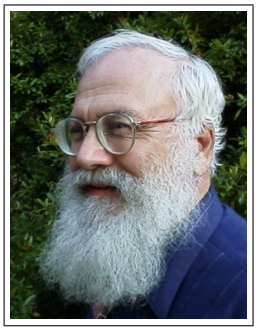
Ali Almomin graduated from Texas A&M University in 2007 with a BS in Geophysics. Then, he joined Saudi Aramco and worked in several exploration and research departments with a focus on 3D seismic processing and near surface imaging. He joined Stanford Exploration Project in 2009 to pursue a PhD in Geophysics 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 Geodepth team of Paradigm Geophysical, maintaining and developing the 3D Kirchhoff migration project. He joined SEP in 2009 and is currently pursuing a Ph.D. in geophysics at Stanford University, and a longer biography. His current research project 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.



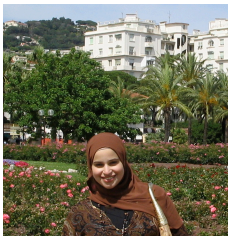
Jason Chang received his B.A. in geophysics from the University of California, Berkeley, in 2010. He joined SEP in autumn of 2011 and is currently working toward a Ph.D. in geophysics. He is a student member of SEG and AGU.



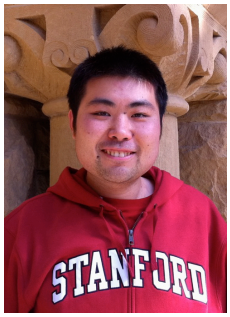
Jon F. Claerbout (M.I.T., B.S. physics, 1960; M.S. 1963; Ph.D. geophysics, 1967), professor at Stanford University, 1967. Emeritus 2008. Best Presentation Award from the Society of Exploration Geophysicists (SEG) for his paper, *Extrapolation of Wave Fields.* Honorary member and SEG Fessenden Award "in 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.

Noha Farghal Noha received her MS in Physics from the American University in Cairo (AUC), Egypt, in collaboration with the Interuniversity MicroElectronics Center (IMEC) in Leuven, Belgium, with scholarship from the Katholieke Universiteit Leuven. Her MS thesis involved Pulsed Laser Deposition of thin Gold-Nickel alloy films for RF MEMS microswitch contacts. As an undergraduate, she studied Physics at the AUC and served as the Vice President of AUC's Chapter of the Society of Physics Students and Editor-in-Chief of the society's magazine The Kaleidoscope of Physics until she graduated with highest honors (Summa Cum Laude). Her thesis was related to geophysical monitoring of nuclear test explosions and supervised by Dr. Rashad Kebeasy, former director of the International Data Center in Vienna. Noha was the first Egyptian to receive the Gates Cambridge Scholarship from the University of Cambridge, UK. Noha joined SEP in 2011 and is working on subsurface imaging using microseismic reflections.

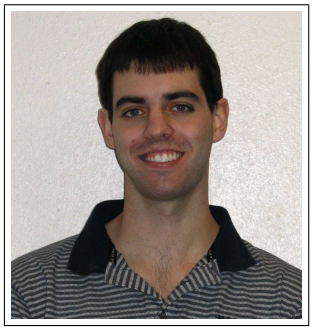


Qiang Fu received his Bachelors degree in Computer Science from Beijing Institute of technology, China in 2003. From 2007 to 2010 he worked for Saudi Aramco in Saudi Arabia. He joined SEP in the fall of 2010 and is currently working toward a M.S. in Geophysics at Stanford University. He is a student member of the SEG.



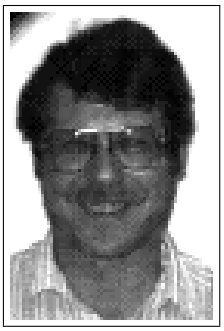
Antoine Guitton received a M.Sc. in geophysics from Universite de Strasbourg, France in 1996 and from Stanford University in 2000. He received his Ph.D. from Stanford University in 2005. He was awarded Best Student Paper from the SEG in 1999 and received the EAGE Van Weelden award in 2004. He was research assistant at the Institut Francais du Petrole (Paris-1996/97) working on well seismic imaging and research geophysicist at CGG Houston (1997-98) working on multiples attenuation. He is now Senior Geophysicist for 3DGeo Inc. in Santa Clara, CA. His research interests include interpolation, noise attenuation, inversion and novel seismic-interpretation techniques. He is a member of the SEG and EAGE.

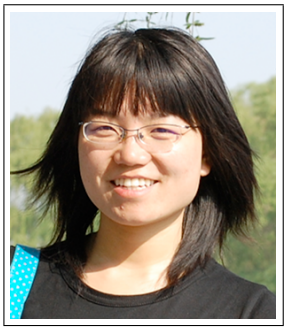
Adam Halpert graduated from Rice University in May 2007 with a B.Sc. in Earth Science and a B.A. in Policy Studies. He joined SEP in the summer of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG and AGU.



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

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.





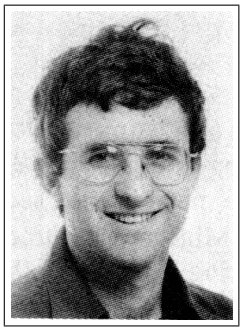
Yunyue (Elita) Li graduated from China University of Petroleum, Beijing in July 2008 with a B.S. in Information and Computational Science. She joined SEP in the fall of 2008, and is currently working toward a Ph.D. in Geophysics. She is a student member of the SEG.

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.

Colin Perkins works for Shell International Exploration and Production Company.

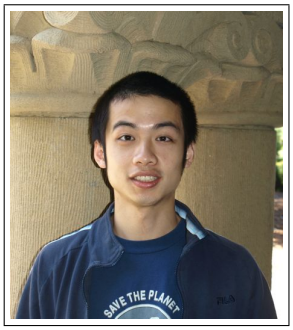
Sjoerd de Ridder received his B.Sc. (Dec. 2004) in Earth Sciences from Utrecht University in the Netherlands. After an exchange year at Colorado School of Mines he attended Delft University of Technology. In August of 2007, he obtained an M.Sc. in Applied Earth Sciences with specialization in applied 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 a geophysicist with Geco-Prakla working on various aspects of seismic data processing and acquisition. Previously he worked with GeoQuest on seismic data interpretation; with Schlumberger on reservoir characterization; with the Institute of Petroleum Research and Geophysics on seismic data processing and acquisition; with the Colorado School of Mines as a visiting professor; and with Saxpy Computer company as an engineer. He has a Ph.D. from Stanford in Geophysics, and a B.Sc. in Physics and Geology from Hebrew University. (Photo: December 1985, Geophysics, p. 2919)



Peng Shen received a Ph.D. (2004) in geophysics from Rice University in Houston, Texas. He joined Total E&P in 2004 in depth-imaging research. In 2007, he joined Shell International Exploration and Production. His current research focuses on velocity analysis.

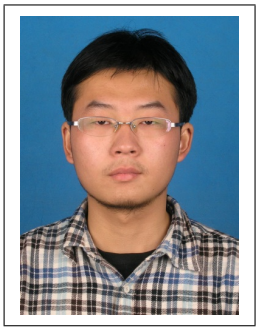
Xukai Shen graduated from Tsinghua University in July 2007 with a B.E. in Electrical Engineering. He joined SEP in the fall of 2007, and is currently working toward a Ph.D. in Geophysics. He is a student member of the SEG.



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 (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).

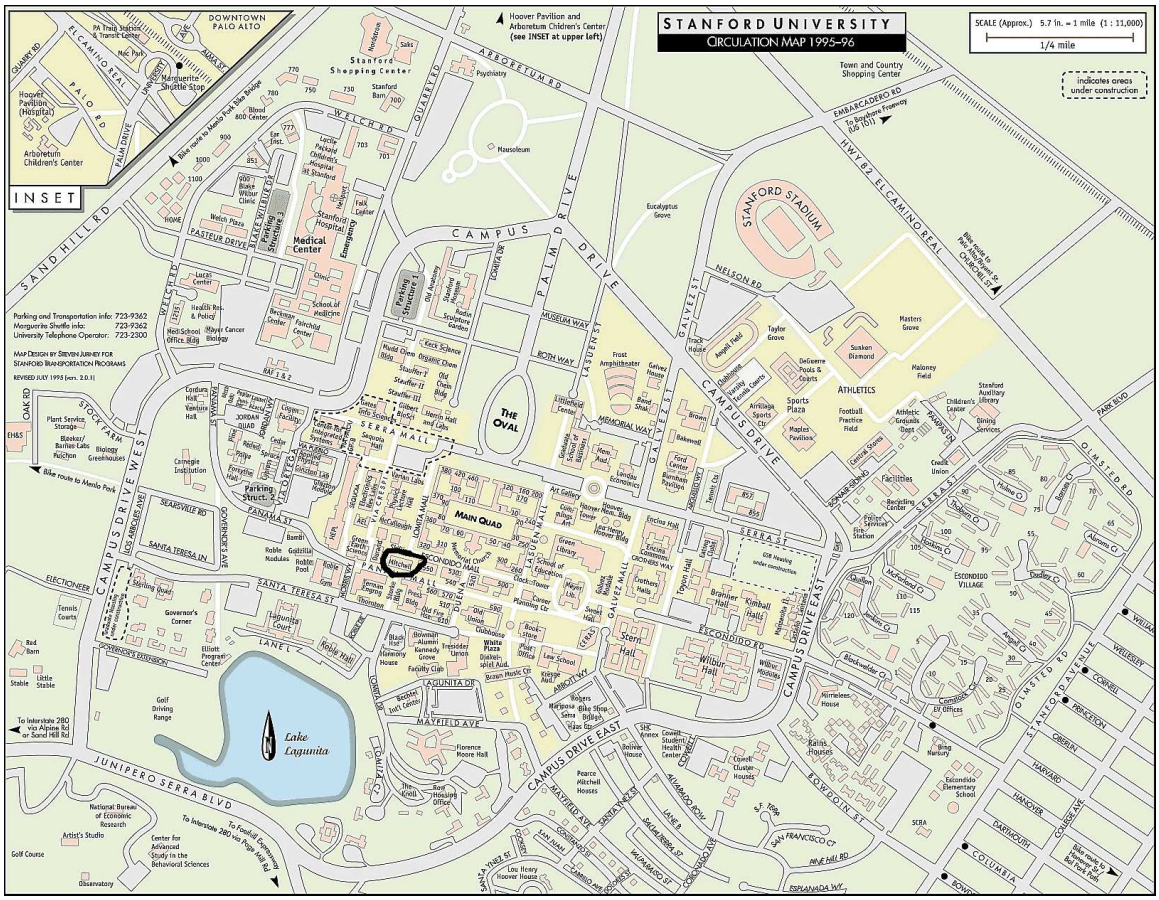
Wong, M., Biondi, B.L., and Ronen, S., 2010, Joint inversion of up- and down-going signal for ocean bottom data: SEG Technical Program Expanded Abstracts **29**, 2752–2756.

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 **31** (Submitted).

Wong, M., and Ronen, S., and Biondi, B.L., 2012, Imaging with multiples using linearized full-wave inversion. SEG Technical Program Expanded Abstracts **31** (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
Hoover Pavilion and Arboretum Children's Center (see INSET at upper left)
INSET
DOWNTOWN PALO ALTO
PA Train Station & Transit Center
Marguerite Shuttle Stop
Hoover Pavilion (Hospital)
Arboretum Children's Center
EL CAMINO REAL
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 (vers. 2.0.1)
SAND HILL RD
Stanford Shopping Center
Town and Country Shopping Center
ARBORETUM RD
EMBARCADERO RD
To Bayshore Freeway (US 101)
HWY 82 EL CAMINO REAL
Bike route to Palo Alto/Bryant St.
Medical Center
Stanford Hospital
Lucile Packard Children's Hospital at Stanford
Mausoleum
Eucalyptus Grove
STANFORD STADIUM
CAMPUS DRIVE
PALM DRIVE
THE OVAL
MAIN QUAD
Memorial Church
Green Library
Hoover Tower
ATHLETICS
Maples Pavilion
Sunken Diamond
Sports Plaza
Tresidder Union
White Plaza
Lake Lagunita
Wilbur Hall
Stern Hall
Mitchell
ESCONDIDO VILLAGE
Golf Driving Range
JUNIPERO SERRA BLVD
To Interstate 280 via Alpine Rd or Sand Hill Rd
To Interstate 280 via Page Mill Rd
To Foothill Expressway
Bike route to Menlo Park
CAMPUS DRIVE WEST
CAMPUS DRIVE EAST
STANFORD AVENUE
Bike route to Hanover St. & Bol Park Path