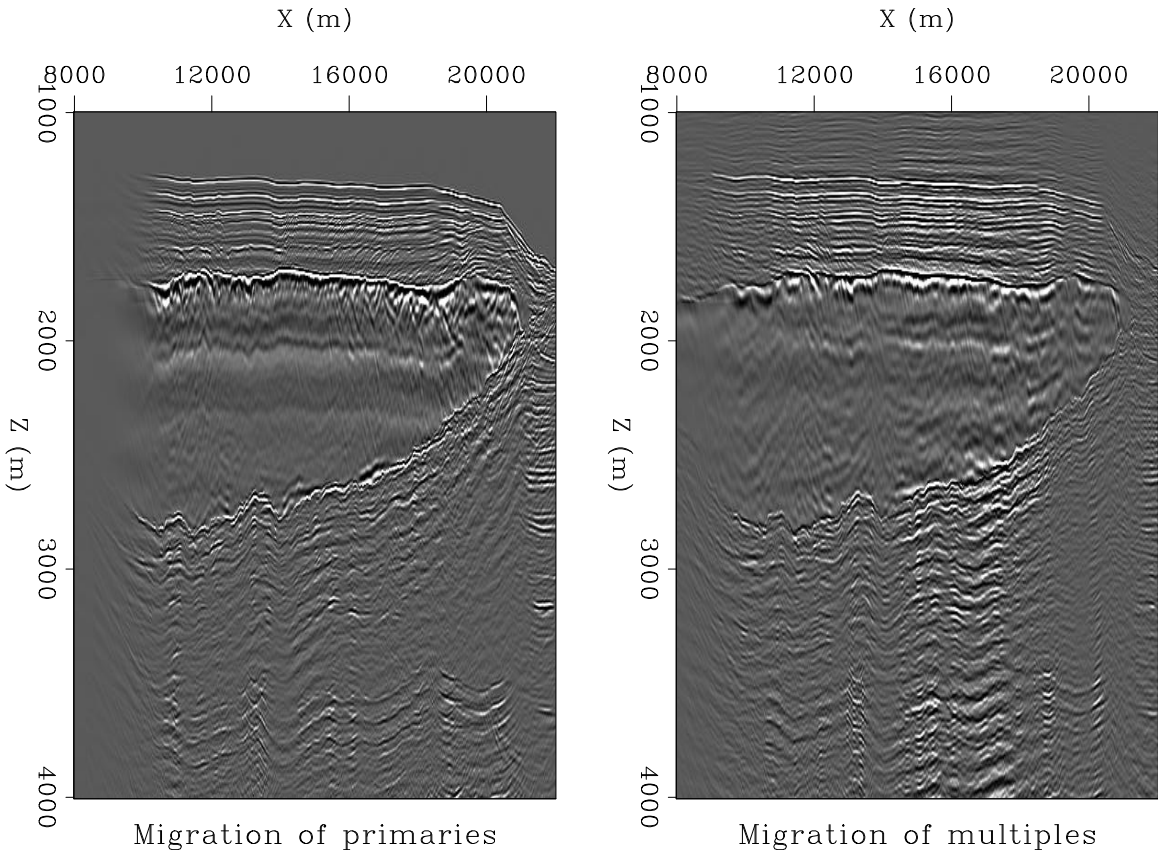
# STANFORD EXPLORATION PROJECT

*Gabriel Alvarez, Elkin Arroyo, Brad Artman, James Berryman, Biondo Biondi, Liliana Borcea, Morgan Brown, Robert Clapp, William Curry, Sergey Fomel, Antoine Guitton, Seth Haines, E. Herkenhoff, Andrey Karpushin, Jesse Lomask, George Papanicolaou, Marie Prucha, Daniel Rosales, Paul Sava, Guojian Shan, Chrysoula Tsogka, Alejandro Valenciano, Eric Verschuur, and Ioan Vlad*

**Report Number 111, April 2002**



*Stanford, California 94305*

# 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, which allows both the author and the reader of a publication to verify the reported results at a later time. Reproducibility facilitates the transfer of knowledge within SEP and between SEP and its sponsors.

**ER** denotes Easily Reproducible and are the results of a 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, or be easily available (e.g. SEG-EAGE data sets) to SEP and non-SEP researchers. The data may also be available in the SEP data library, which can be viewed at http://sepwww.stanford.edu/public/docs/sepdatalib/toc_html/. We assume you have a UNIX workstation with Fortran, Fortran90, C, X-Windows system and the software downloadable from our website (SEP makerules, SEPlib, and, to properly reproduce the documents, 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. SEP staff have only attempted to make sure that the makefile rules exist and the source codes referenced are provided. The primary reasons for the CR designation is that the processing requires 20 minutes or more, or commercial packages such as Matlab or Mathematica.

**M** denotes a figure that may be viewed as a movie in the web version of the report. A movie may be either ER or CR.

**NR** denotes Non-Reproducible. This class of figure is considered non-reproducible. SEP discourages authors from flagging their figures as NR except for artist drawings, scannings, etc.

Our testing is currently limited to IRIX 6.5 and LINUX 2.1 (using the Portland Group Fortran90 compiler), but the code should be portable to other architectures. Reader's suggestions are welcome. For more information on reproducing SEP's electronic documents, please visit `<http://sepwww.stanford.edu/redoc/>`.

Jon Claerbout, Biondo Biondi, Robert Clapp, and Marie Prucha

[1] http://sepwww.stanford.edu/private/docs/sep108

# SEP–111 — TABLE OF CONTENTS

## Migration



## Velocity



## Time reverse imaging



## Noise removal

**Modeling and migration computational issues**



**Amplitudes and rock properties**



**Filtering and inversion**



**Passive imaging**



**Computing**

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 1–17*

# Subsalt event regularization with steering filters

*Marie L. Prucha and Biondo L. Biondi*[1]

**ABSTRACT**

The difficulties of imaging beneath salt bodies where illumination is poor are well known. In this paper, we present an angle-domain least-squares inversion scheme that regularizes the seismic image, tending to smooth along specified dips. This smoothing is accomplished using steering filters. We show the results of using the regularized inversion with smoothing along the angle axis and along both the angle and common midpoint axes. Additionally, the ramifications of specifying incorrect dips to smooth along will be examined. The results show that this regularized least-squares inversion does produce a cleaner, more continuous result under salt bodies. The inversion will reject incorrectly chosen dips used for the regularization.

## INTRODUCTION

Obtaining a clean, coherent seismic image in areas of complex subsurface can be difficult. This is particularly true when the subsurface lends itself to shadow zones, such as those under the edges of salt bodies. In these areas, little of the seismic energy gets to the reflectors, and even less energy makes it back to the surface (Muerdter et al., 1996; Prucha et al., 1998). In addition to the shadow zones, proper imaging is made difficult by multipathing. Multipathing occurs when energy that follows different paths through the subsurface arrives at the receiver at the same time. Solving these issues is no simple matter.

It has been suggested that it is possible to eliminate artifacts caused by multipathing by imaging in the reflection angle domain via angle-domain Kirchhoff migration (Xu et al., 1998, 2001). However, Kirchhoff methods are not necessarily optimal for complex subsurfaces (SEG Workshop, 2001). Additionally, Stolk and Symes (2002) have shown that angle-domain Kirchhoff migration may not eliminate artifacts in strongly refracting media, even when using the correct velocity and all arrivals. To address these issues, angle-domain wave-equation migration methods have been developed (Prucha et al., 1999a,b; Fomel and Prucha, 1999; Sava et al., 2001). Stolk and De Hoop (2001) have shown that these methods do not suffer from image artifacts when migrated at the correct velocity.

Unfortunately, migration alone does not necessarily provide the best image. A better image in complex areas can often be obtained using least-squares inversion (Nemeth et al., 1999; Duquet and Marfurt, 1999; Ronen and Liner, 2000). However, inversion often cannot improve

[1]**email:** marie@sep.Stanford.EDU,biondo@sep.Stanford.EDU

the image in shadow zones without many iterations. With many iterations, noise in the null space caused by poor illumination can cause the inversion process to blow up.

To prevent these blow-ups, since we often have some idea of what the image in the shadow zone should look like, we can impose some sort of regularization on the inversion carried out in the angle domain (Prucha et al., 2000, 2001; Kuehl and Sacchi, 2001). In this paper, we will use a regularization that tends to create dips in the image from *a priori* selected reflectors and therefore can be applied by the use of steering filters (Clapp et al., 1997; Clapp, 2001). The inversion uses these steering filters to strengthen existing events to help fill in shadow zones. This regularization may even improve amplitude behavior (Prucha and Biondi, 2000).

The inversion process we use in this paper lends itself to two different regularization schemes. The "1-D approach" tends to create flat events along the reflection angle axis. This helps to fill in holes in different angle ranges caused by poor illumination. The "2-D approach" cascades the 1-D approach with an attempt to create dips along picked reflectors in the common midpoint (CMP) - depth plane. This allows the user to test possible interpretations within shadow zones. If the picked reflector doesn't cause the regularization to tend to create dips that interfere with the data, the model will contain those "new" dips. If the picked reflectors are incompatible with the data, the inversion will reject the dips.

In this paper, we will first review the theory and implementation of our inversion, then show the results of this regularized inversion scheme on a fairly complex synthetic model. We will then study the impact of the regularization operator more closely. Finally, we will discuss some of our future plans.

## THEORY

Our inversion scheme is based on the angle-domain wave-equation migration explained by Prucha et al. (1999a). To summarize, this migration is carried out by downward continuing the wavefield in frequency space, slant stacking at each depth, and extracting the image at zero time. The result is an image in depth (z), common midpoint (CMP), and offset ray parameter ($p_h$) space. The offset ray parameter can be easily related to the reflection angle by:

$$\frac{\partial t}{\partial h} = p_h = \frac{2 \sin\theta \cos\phi}{V(z, cmp)}, \tag{1}$$

where $\theta$ is the reflection angle, $\phi$ is the geologic dip, and $V(z, cmp)$ is the velocity function in depth and CMP location.

The inversion procedure used in this paper can be expressed as fitting goals as follows:

$$\begin{aligned} \mathbf{d} &\approx \mathbf{L}\mathbf{m} \\ 0 &\approx \epsilon \mathbf{A}\mathbf{m}. \end{aligned} \tag{2}$$

The first equation is the "data fitting goal," meaning that it is responsible for making a model that is consistent with the data. The second equation is the "model styling goal," meaning that

it allows us to impose some idea of what the model should look like using the regularization operator **A**. The model styling goal also helps to prevent a divergent result.

In the data fitting goal, **d** is the input data and **m** is the image obtained through inversion. **L** is a linear operator, in this case it is the adjoint of the angle-domain wave-equation migration scheme summarized above and explained thoroughly by Prucha et al. (1999b). In the model styling goal, **A** is, as has already been mentioned, a regularization operator. $\epsilon$ controls the strength of the model styling.

Unfortunately, the inversion process described by Equation 2 can take many iterations to produce a satisfactory result. We can reduce the necessary number of iterations by making the problem a preconditioned one. We use the preconditioning transformation $\mathbf{m} = \mathbf{A^{-1}p}$ (Fomel et al., 1997; Fomel and Claerbout, 2002) to give us these fitting goals:

$$\begin{aligned} \mathbf{d} &\approx \mathbf{LA^{-1}p} \\ 0 &\approx \epsilon\mathbf{p}. \end{aligned} \tag{3}$$

$\mathbf{A^{-1}}$ is obtained by mapping the multi-dimensional regularization operator **A** to helical space and applying polynomial division (Claerbout, 1998).

The question now is what the preconditioning operator $\mathbf{A^{-1}}$ is. We have chosen to make this operator from steering filters (Clapp et al., 1997; Clapp, 2001) which tend to create dips along chosen reflectors. This paper includes results from two different preconditioning schemes. One is called the 1-D preconditioning scheme and simply acts horizontally along the offset ray parameter axis. The 2-D scheme acts along chosen dips on the CMP axis and horizontally along the offset ray parameter axis. To construct the preconditioning operator along the CMP axis, we pick "reflectors" that represent the dip we believe should be in a certain location, then interpolate the dips between the picked reflectors to cover the whole plane.

## RESULTS

We applied our preconditioned inversion scheme to a synthetic dataset provided to us by SMAART JV. This dataset is designed to have a significant shadow zone underneath the salt body. The result of angle-domain wave-equation migration of this dataset can be seen in Figure 1. In the CMP-depth plane, note the severe decrease in amplitude of the reflectors as they go beneath the salt. In the offset ray parameter ($p_h$)-depth plane, there is a large decrease in amplitude at small $p_h$. It is most visible inside the oval drawn on the figure. We consider this to be a "hole" in the event, since there is energy at very small $p_h$ and large $p_h$.

Figure 2 shows the result of 3 iterations of 1-D preconditioned inversion. Recall that the 1-D preconditioning scheme acts horizontally along the $p_h$ axis. The most obvious result of this is a substantial increase in the signal to noise ratio. In the context of this paper, the more interesting result is the increase in strength of the events along the $p_h$ axis. The "hole" that is circled in the $p_h$-depth plane is beginning to fill in. This in turn makes the reflectors in the CMP-depth plane appear to extend farther under the salt, with stronger amplitudes. This result

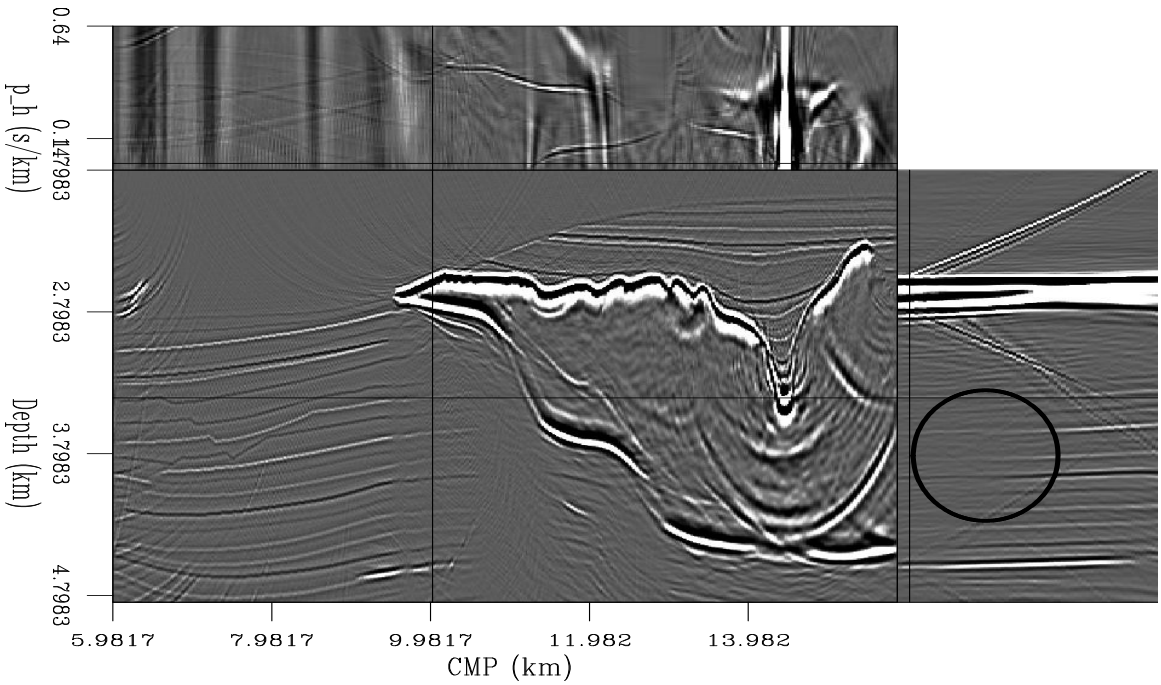



Figure 1: Result of angle-domain wave-equation migration. marie1-mig [CR]

is encouraging, but it seems likely that several more iterations would be necessary to really fill the hole.

To help fill in the shadow zones faster, we used our 2-D preconditioning scheme. To do this, we first picked reflectors in the CMP-depth plane (Fig. 3) to be used to create the preconditioning operator in this dimension. In the inversion, this operator was cascaded with the operator that smooths horizontally along the $p_h$ axis.

The result of 3 iterations of the 2-D preconditioning scheme can be seen in Figure 4. This result is smooth and very promising. The hole in the $p_h$ axis is gone and the reflectors extended underneath the salt with a stronger amplitude. Unfortunately, this method also creates some obvious errors in the output image. The top of the salt has lost some of its features because the picked reflector there was not detailed enough. Also, the faults at the left side of the image have been smoothed out. In this case, neither of these areas are of particular interest, but the problem of smoothing areas that shouldn't be smoothed is one we are trying to solve.

One way to reduce the smoothing effect is to follow several iterations of the preconditioned inversion (Equation 3), which tends to be smoother with fewer iterations, with a regularized inversion (Equation 2), which tends to be rougher with fewer iterations. The result of 3 iterations of preconditioned inversion followed by one iteration of regularized inversion can be seen in Figure 5. The result looks similar to the migration result (Fig. 1). Many events containing higher frequencies have been revived, including some of the processing artifacts. However, close examination of the areas of particular interest reveals important improvements.

In the CMP-depth plane, the reflectors going beneath the salt do not have the hole directly

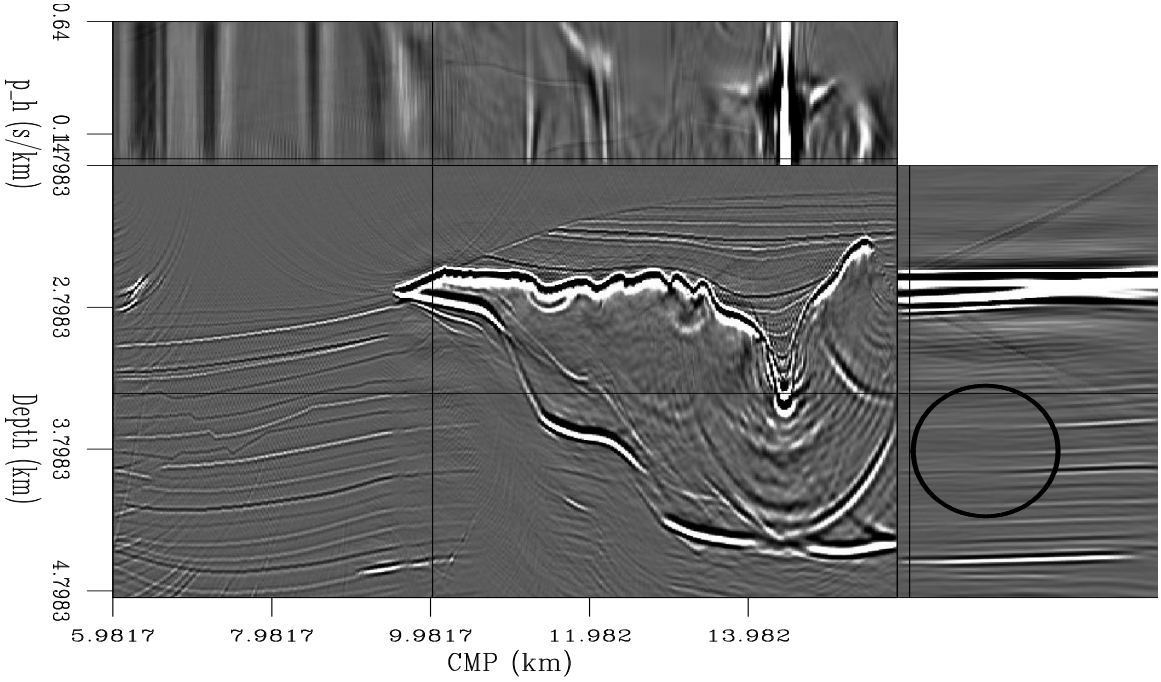



Figure 2: Result of 3 iterations of the 1-D preconditioned inversion. marie1-1dprec [CR]

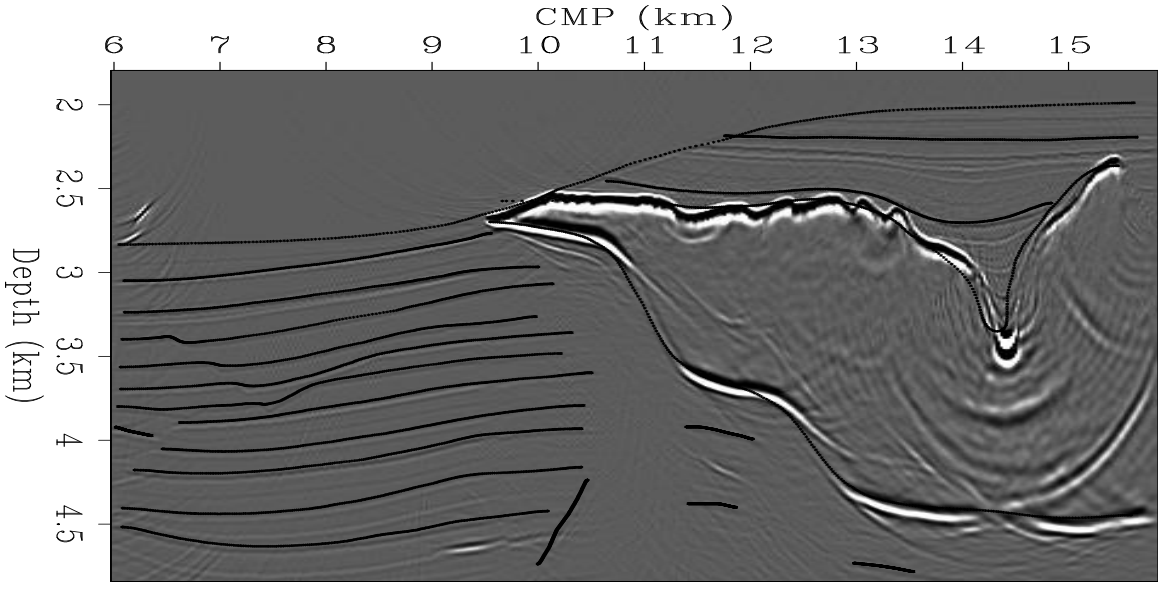



Figure 3: Migration result with picked reflectors for the 2-D preconditioned inversion overlaid. marie1-reflectors [CR]

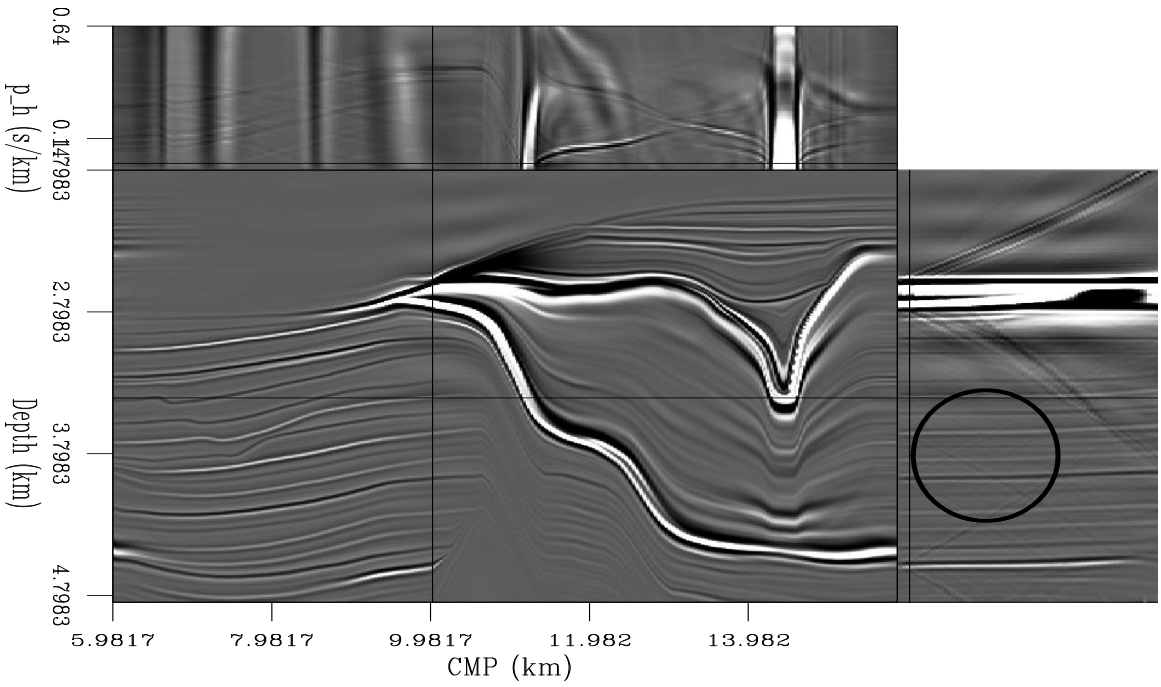



Figure 4: Result of 3 iterations of the 2-D preconditioned inversion. marie1-2dprec [CR]

beneath the tip of the salt that is seen in the migration result. In the circled area in the $p_h$-depth plane, the reflectors are more constant in amplitude than in the migrated result. These improvements could be increased with more iterations of regularized inversion. We are working on a way to determine a better combination of preconditioned and regularized iterations.

## STEERING FILTER TEST

The previous results are promising, but it is clear that the 2-D inversion is strongly dependent on the preconditioning operator. Since this operator is constructed from picked reflectors, it is important to know what happens if the reflector is not picked well. To examine this, we created a preconditioning operator from the "reflectors" picked in Figure 6. In this figure, note that the picked reflectors cross the correct dips at the depths between 3 and 3.4 kilometers and 3.7 and 4.1 kilometers. They cross themselves at depth 4.5 kilometers. The picked reflector beginning at depth 3.75 km follows the correct dip for the most part, but ignores the slight change in dip at the fault at CMP position 7.2 km. The water bottom has been correctly picked. The picked reflector beginning at depth 4.2 km follows the correct dip, but continues well into the shadow zone where it may or may not be correct. Also within the shadow zone is a completely absurd picked reflector put there to see if any event can be created there. Finally, note that the top and bottom of the salt have not been picked at all. This will leave the preconditioning operator in the salt area to be interpolated from the picked reflectors. In this section, we will refer to the dips and reflectors from the migration result as "real" or "correct" and the dips and reflectors

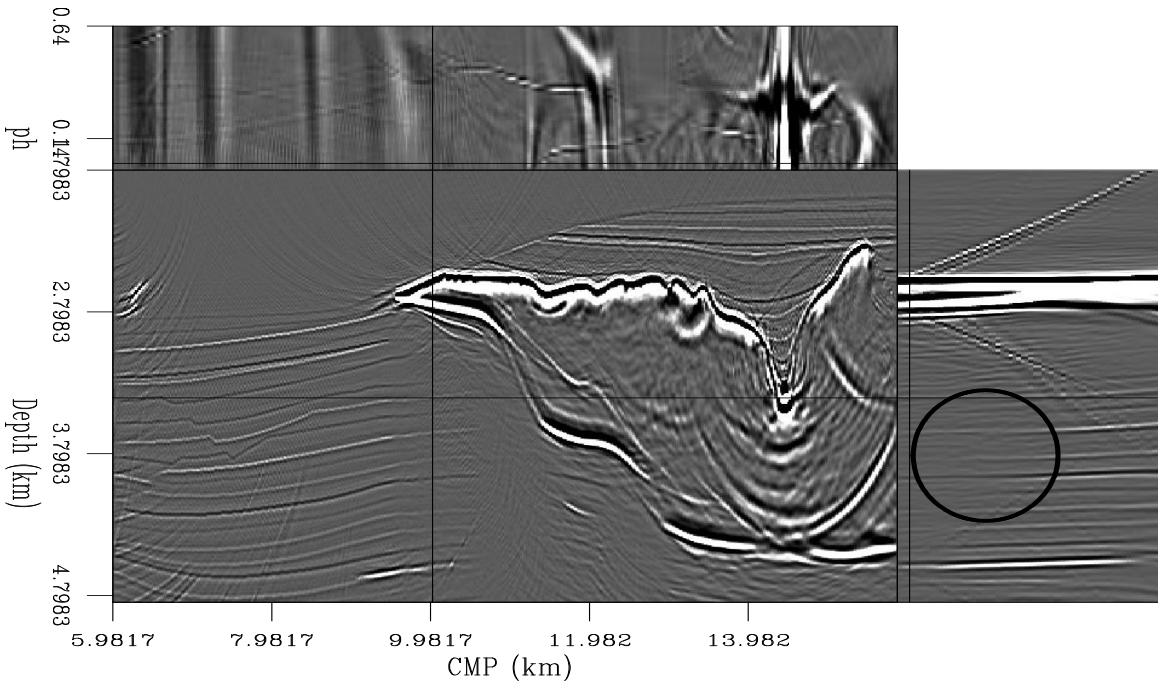



Figure 5: Result of 3 iterations of 2-D preconditioned inversion followed by 1 iteration of 2-D regularized inversion. marie1-2dprecreg [CR]

used for the inversion as "picked."

The result of using this preconditioning operator for the 2-D inversion is seen in Figure 7. As expected, the result isn't good. To investigate this closely, we chose 5 areas to look at as defined in Figure 8. In these areas, we calculated the instantaneous energy on both the migrated result and the result of the 2-D badly preconditioned inversion.

Box 1 enclosed an area in which the picked reflector began with an opposite dip to the correct one then changed to the correct one. In Figure 9, we can see that the energy from the migrated result is fairly constant all along the reflectors. In the panel from the inversion result, the energy has almost completely disappeared along the real reflectors where the picked reflector had the wrong dip, then recovers where the picked reflector has the correct dip. This example shows that picking dips that are completely wrong (opposite of the correct one) will cause the inversion to reject what the model styling goal (Eqn. 3) is trying to do.

In Box 2, the energy in the migrated panel is once again fairly constant along all of the reflectors (Fig. 10). In the preconditioned result, one of the most obvious effects is the loss of energy in the upper left corner. This loss of energy is caused by the picked reflector shown in Figure 9 which has the incorrect dip. The picked reflectors in Figure 10 are quite interesting. Both picked reflectors match the real events except where the real events are affected by faulting. For the shallower event, this results in energy loss at the beginning and end of the faulted area but little loss of energy elsewhere because the real dips are very close to the picked dips. For the deeper event, the faulted area has a different dip from the picked reflector, so the entire

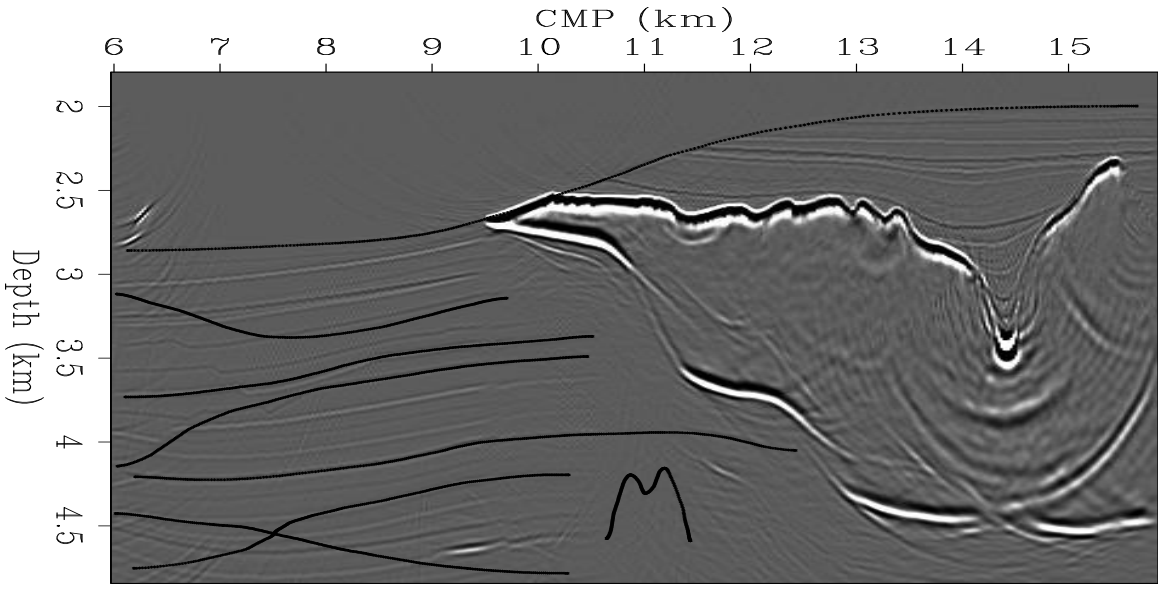



Figure 6: Migration result with the picked "reflectors" overlaid. These reflectors do not match the correct reflectors and should produce a bad result. marie1-reftest [CR]

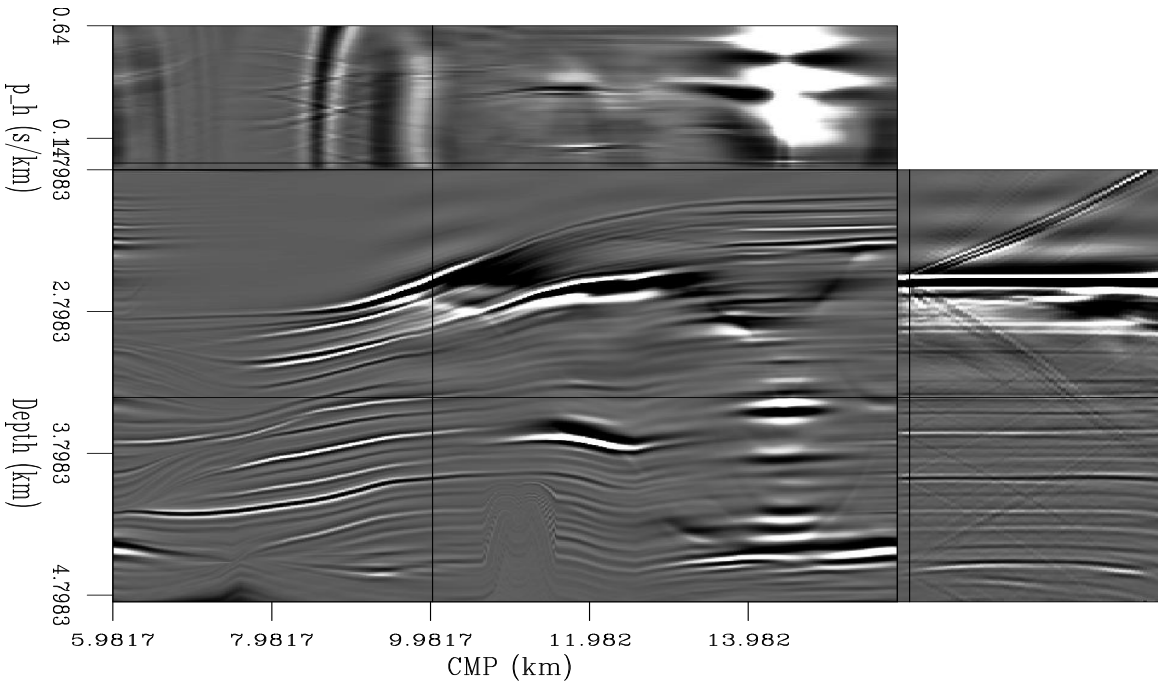



Figure 7: Result of 3 iterations of the 2-D preconditioned result with badly picked reflectors. marie1-2dfilttest [CR]

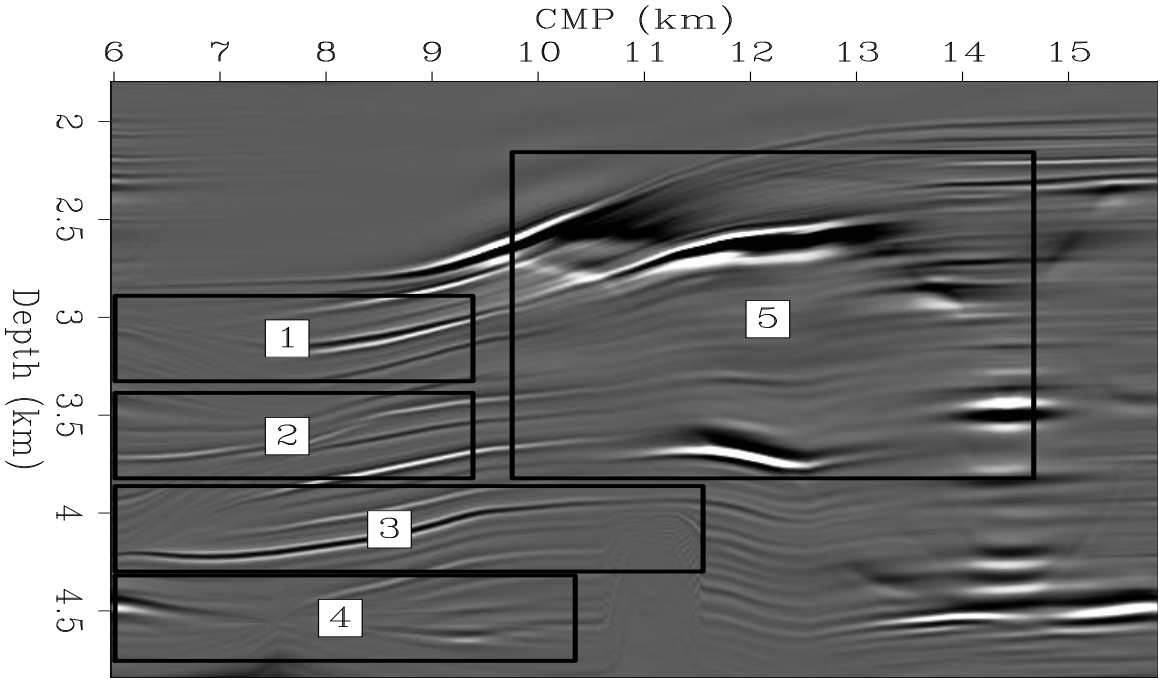



Figure 8: We will closely examine each of the boxed sections on this result of 3 iterations of the 2-D preconditioned inversion with badly picked reflectors. marie1-boxit [CR]

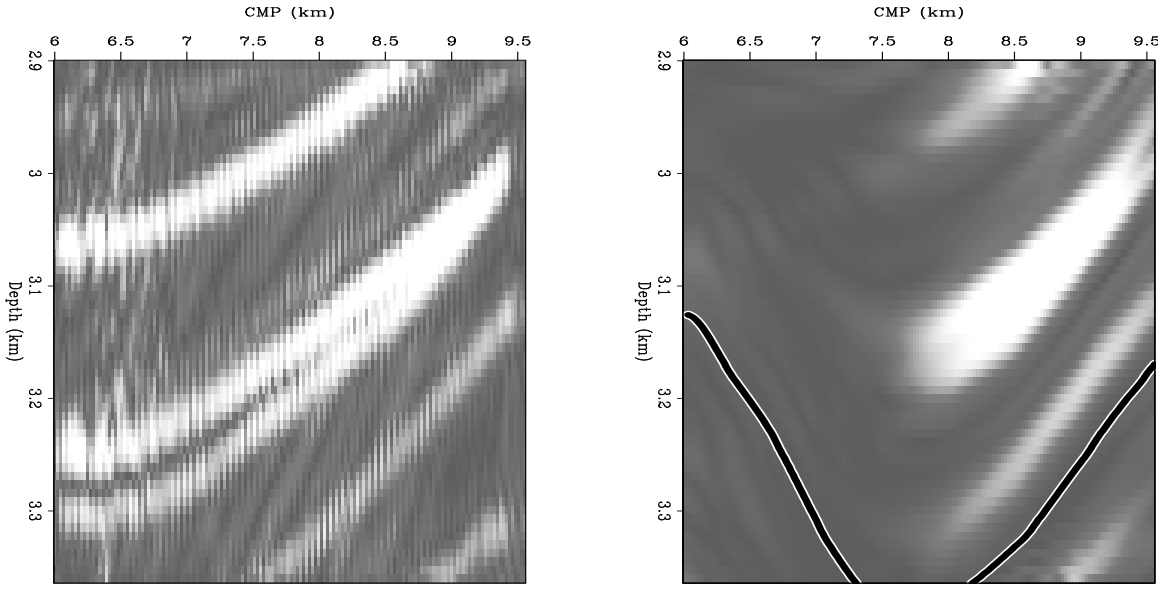



Figure 9: Energy comparison in Box 1. The left plot is from the migrated result, the right plot is from the 2-D preconditioned result with the picked reflector overlaid. The preconditioned result has very low energy where the picked reflector has the incorrect dip. marie1-comp.1 [CR,M]

faulted area has lost energy. This result may mean that this inversion scheme can be modified for fault detection. The deepest event in this box has lost no energy because it is close to the dip of the picked reflector above it. This example shows that the picked dip doesn't have to be totally different from the correct dip for the inversion to reject it.

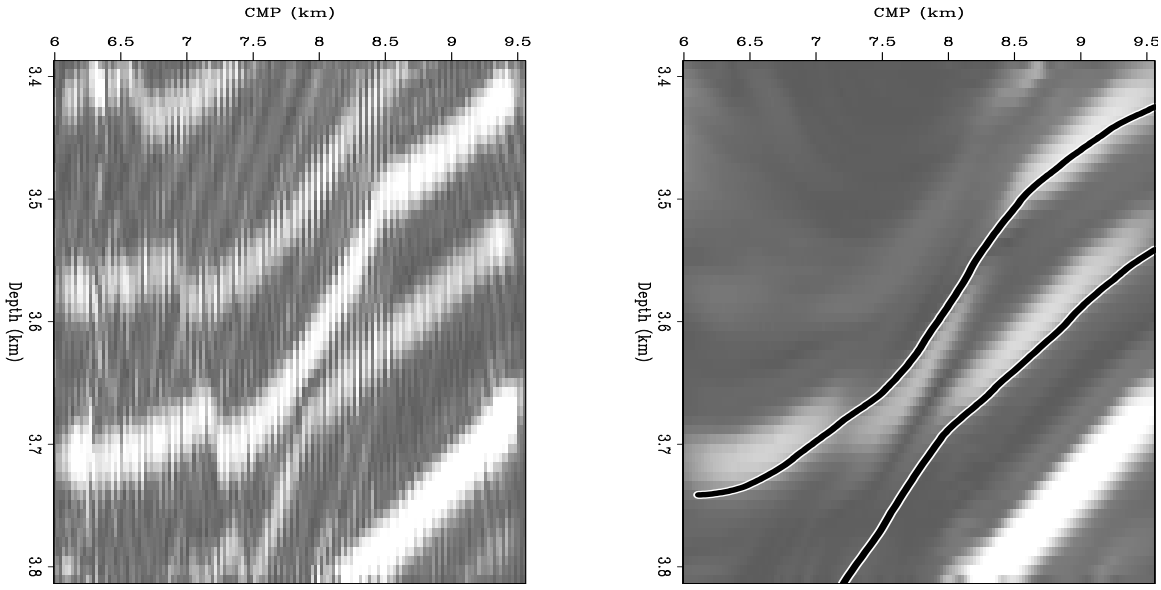



Figure 10: Energy comparison in Box 2. The left plot is from the migrated result, the right plot is from the 2-D preconditioned result with the picked reflectors overlaid. Once again, there is low energy where the dip of the picked reflector conflicts with the real dip. marie1-comp.2 [CR,M]

In Box 3, all of the coherent energy in the migrated result (Fig. 11) ends at CMP=9.5 km. The shallowest picked reflector in this box is the now-familiar incorrect dip reflector that causes the preconditioned result to lose energy in the upper left corner. The next two picked reflectors follow the real events but extend well beyond the coherent energy limit in the migrated result. This has caused the inversion to generate coherent energy well into the shadow zone beneath the salt. The final picked reflector in this box is the silly "M" shaped one that is entirely within the shadow zone. This picked reflector has had almost no effect on the energy of the preconditioned result, just as we would hope. This result clearly shows that the reasonably picked reflectors will enhance the result of the preconditioned inversion in shadow zones and that poorly picked reflectors in shadow zones won't generate unreasonable events.

The fourth area (Box 4) examined is very interesting. Here we wanted to test the stability of our inversion if the preconditioning operator contained conflicting dips. Figure 12 shows the energy from the migrated result is somewhat garbled along the left side and coherent with smoothly varying dips otherwise. There is also a bright spot that is in the real model at CMP=9.5 km. The picked reflectors only match the correct dips above the depth of 4.5 km. The energy from the preconditioned result is incoherent and weak except for a small section between the CMPs at 8 km and 9 km and above 4.5 km in depth. The garbled left side energy from the migrated result has been smeared into a large blob and the bright spot

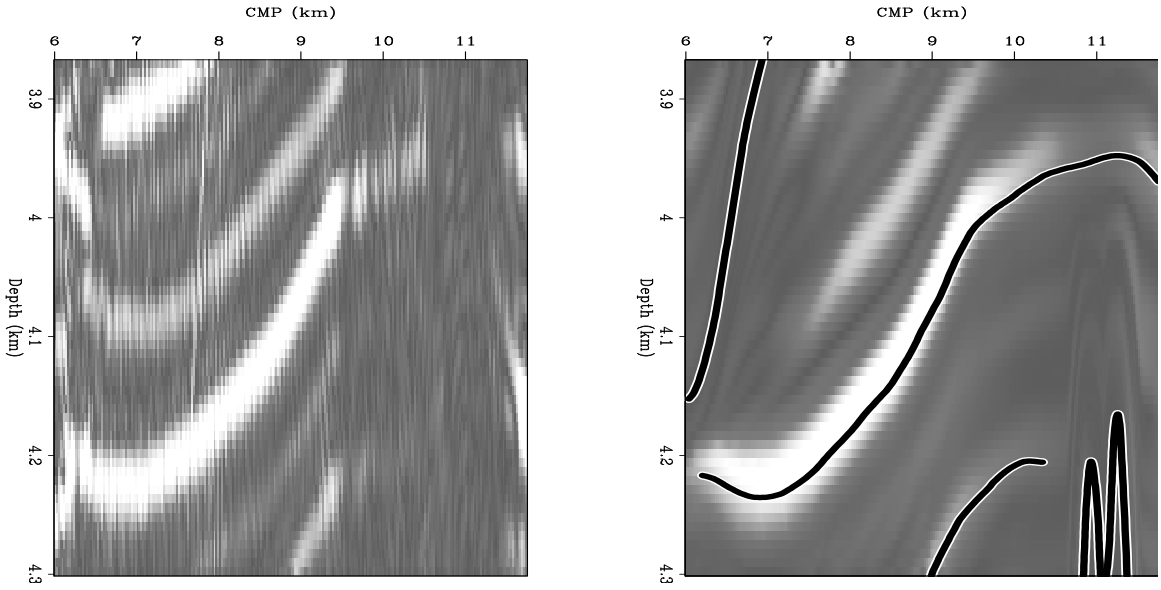



Figure 11: Energy comparison in Box 3. The left plot is from the migrated result, the right plot is from the 2-D preconditioned result with the picked reflectors overlaid. The reflectors were for the most part correctly picked for the preconditioned result, which does look better than the migrated result. marie1-comp.3 [CR,M]

has also been smudged. Overall this preconditioned result has managed to destroy almost all of the real information, just as expected. The important point from this experiment is that the preconditioned result did not show unexpected behavior in the area where the picked reflectors cross.

Box 5 contains part of the salt body. The only picked reflector really influencing this box was the water bottom. The energy from the migration shows high energy along the salt top and bottom, with low energy along the water bottom reflection and other layers (Fig. 13). The inversion has attempted to impose the dip of the water bottom on the salt events. The result has eliminated almost all of the energy that belonged to reflectors that were not close to the dip of the water bottom. This example shows that any strong reflectors that seem correct in the migration should be picked to preserve them in the inversion.

## CONCLUSIONS

This paper has presented an inversion scheme that uses steering filters as a preconditioning operator. These steering filters tend to create dips along chosen reflectors. We presented a 1-D scheme in which the steering filters simply tried to act horizontally along the $p_h$ axis and a 2-D scheme in which the tendency to create horizontal dips along the $p_h$ axis was cascaded with steering filters tending to create dips along picked reflectors in the CMP-depth plane. Both of these methods increased the signal-to-noise ratio and helped to fill in the shadow zones.

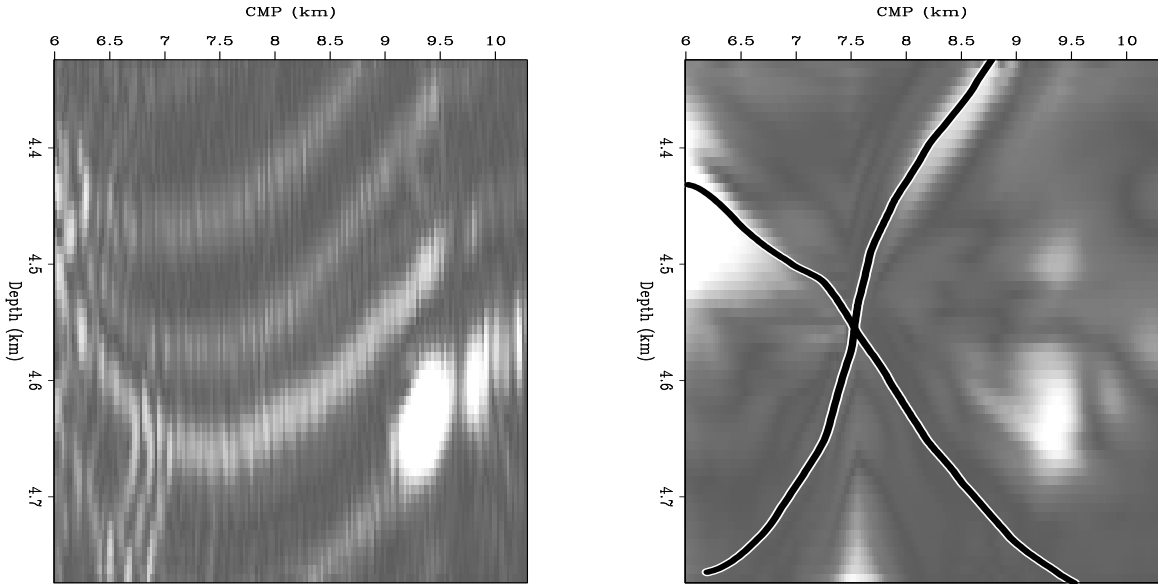



Figure 12: Energy comparison in Box 4. The left plot is from the migrated result, the right plot is from the 2-D preconditioned result with the picked reflectors overlaid. The preconditioned result because the picked reflectors had the incorrect dips almost everywhere. marie1-comp.4 [CR,M]

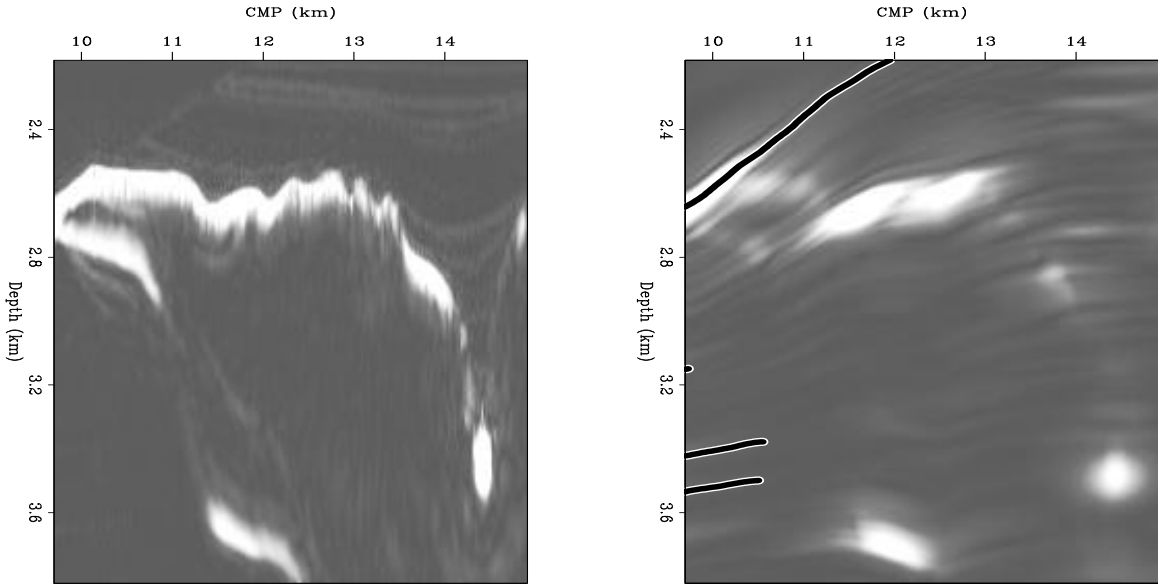



Figure 13: Energy comparison in Box 5. The left plot is from the migrated result, the right plot is from the 2-D preconditioned result with the very distant picked reflectors seen at the edges. The preconditioned result is horrible because the salt top and bottom were not picked. marie1-comp.5 [CR,M]

We also examined the effect of variation in the picked reflectors for the 2-D scheme. This was done by picking a variety of reflectors based on both the correct and the incorrect dips. This experiment showed that the inversion will reject dips that are incorrectly picked where data exists. This can even indicate areas where faulting has occurred. The inversion assures that picked dips that generate an event which interferes with the data are rejected. Picked dips that generate an event that doesn't interfere with existing data are allowed. Picked dips that cross, or meet at a point can be accommodated by the inversion. It is necessary to pick reflectors wherever the dominant dip changes.

## FUTURE PLANS

We have several paths we would like to pursue soon. As mentioned above, we are trying to find a way to fine tune the combined preconditioning plus regularization scheme. Currently we are getting satisfactory results from carrying out more iterations of the preconditioned inversion than the regularized inversion, but we hope to find a way to determine the best combination. Additionally, it is not necessary to apply the preconditioning or regularization everywhere. Therefore we are working on a way to limit the effects of the preconditioning to a user-defined area. Another step that should be fairly straightforward is applying the entire process to 3-D data.

## ACKNOWLEDGMENTS

We would like to thank SMAART JV for the synthetic dataset used in our experiments. Bob Clapp's advice on inversion, steering filters, and computer code has been invaluable.

## REFERENCES

Claerbout, J., 1998, Multidimensional recursive filters via a helix: Geophysics, **63**, no. 5, 1532–1541.

Clapp, R. G., Fomel, S., and Claerbout, J., 1997, Solution steering with space-variant filters: SEP–**95**, 27–42.

Clapp, R. G., 2001, Geologically constrained migration velocity analysis: Ph.D. thesis, Stanford University.

Duquet, B., and Marfurt, K. J., 1999, Filtering coherent noise during prestack depth migration: Geophysics, **64**, no. 4, 1054–1066.

Fomel, S., and Claerbout, J., 2002, Multidimensional recursive filter preconditioning in geophysical estimati on problems: submitted to Geophysics.

Fomel, S., and Prucha, M., 1999, Angle-gather time migration: SEP–**100**, 141–150.

Fomel, S., Clapp, R., and Claerbout, J., 1997, Missing data interpolation by recursive filter preconditioning: SEP–**95**, 15–25.

Kuehl, H., and Sacchi, M., 2001, Generalized least-squares dsr migration using a common angle imaging condition: 71st Ann. Internat. Meeting, Soc. Expl. Geophysics, Expanded Abstracts, 1025–1028.

Muerdter, D. R., Lindsay, R. O., and Ratcliff, D. W., 1996, Imaging under the edges of salt sheets: a raytracing study: 66th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 578–580.

Nemeth, T., Wu, C., and Schuster, G. T., 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, no. 1, 208–221.

Prucha, M., and Biondi, B., 2000, Amplitudes and inversion in the reflection angle domain: SEP–**105**, 203–208.

Prucha, M. L., Clapp, R. G., and Biondi, B. L., 1998, Imaging under the edges of salt bodies: Analysis of an Elf North Sea dataset: SEP–**97**, 35–44.

Prucha, M., Biondi, B., and Symes, W., 1999a, Angle-domain common image gathers by wave-equation migration: 69th Ann. Internat. Meeting, Soc. Expl. Geophysics, Expanded Abstracts, 824–827.

Prucha, M. L., Biondi, B. L., and Symes, W. W., 1999b, Angle-domain common image gathers by wave-equation migration: SEP–**100**, 101–112.

Prucha, M. L., Clapp, R. G., and Biondi, B., 2000, Seismic image regularization in the reflection angle domain: SEP–**103**, 109–119.

Prucha, M. L., Clapp, R. G., and Biondi, B. L., 2001, Imaging under salt edges: A regularized least-squares inversion scheme: SEP–**108**, 91–104.

Ronen, S., and Liner, C. L., 2000, Least-squares DMO and migration: Geophysics, **65**, no. 5, 1364–1371.

Sava, P., Biondi, B., and Fomel, S., 2001, Amplitude-preserved common image gathers by wave-equation migration: 71st Ann. Internat. Meeting, Soc. Expl. Geophysics, Expanded Abstracts, 296–299.

SEG Workshop, 2001, Seismic imaging beyond Kirchhoff workshop: Society of Exploration Geophysicists.

Stolk, C., and De Hoop, M. V. Seismic inverse scattering in the 'wave-equation' approach:. Preprint 2001-047, The Mathematical Sciences Research I nstitute, 2001. http://msri.org/publications/preprints/2001.html.

Stolk, C., and Symes, W., 2002, Artifacts in Kirchhoff common image gathers: 72nd Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.

Xu, S., Chauris, H., Lambare, G., and Noble, M., 1998, Common angle image gather - A strategy for imaging complex media: 68th Annual Internat. Mtg., Society of Exploration Geophysicists, Expanded Abstracts, 1538–1541.

Xu, S., Chauris, H., Lambare, G., and Noble, M., 2001, Common angle image gather - A strategy for imaging complex media: Geophysics, **66**, no. 6, 1877–1894.


HOLLYWOOD TOWER
Hotel

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 17–33*

# Shot-profile migration of multiple reflections

*Antoine Guitton*[1]

**ABSTRACT**

A shot-profile migration algorithm is modified to image multiple reflections at their correct location in the subsurface. This method replaces the impulsive source with an areal source made of recorded primaries and multiples. In addition, the extrapolated wavefield at the receivers consists of recorded multiples only which have been previously separated from the primaries. Migration results with 2-D synthetic and field data prove that the migration of multiples can bring valuable structural information of the earth with or without separation.

## INTRODUCTION

Seismic migration aims to move recorded seismic wavefield at the surface back into the earth at the reflector location from which they have originated. Multiple reflections, which are also recorded, are usually not accounted for in the migration process. Therefore, as a prerequisite to any correct imaging of the subsurface, multiples are traditionally attenuated (Guitton et al., 2001).

In this paper, I try to treat the multiples like signal rather than noise. My goal is to show that multiples can be easily imaged with a conventional shot-profile migration algorithm (Jacobs, 1982; Rickett, 2001). This migration is carried out in the ($\omega$,x) domain. I will assess if the migration of multiple reflections adds any type of structural information and if it can increase the signal-noise ratio of the final image.

In the first section, I review theoretical aspects of shot-profile migration and expand its concepts to image multiples. In the second part, I show migration results with 2-D synthetic and field data. In the last section, I discuss practical aspects of multiple migration.

## THEORY OF REFLECTOR MAPPING WITH SHOT-PROFILE MIGRATION

In this section, I review basic principles of earth imaging as pioneered by Claerbout (1971). Then I generalize these principles to reflector mapping with multiple reflections with a shot-profile migration algorithm.

[1]**email:** antoine@sep.stanford.edu

**Imaging of primaries**

Shot-profile migration aims to produce an image of the subsurface by extrapolating both the source and receiver wavefields into the interior of the earth. The imaging condition (Claerbout, 1971) consists of crosscorrelating the two wavefields at each depth-step. Reflectors are formed where the two wavefields correlate. Figure 1 illustrates this principle.

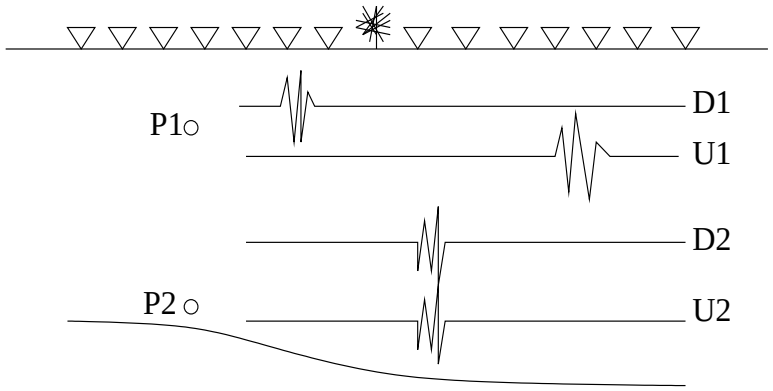



Figure 1: The up-going wavefield is recorded everywhere at the receiver locations (shown as triangles). The down-going wavefield is emitted at the source location in the center of the survey (shown as a star). At earth location P1, the two extrapolated wavefields do not cross-correlate because the down-going wavefield arrives at a much earlier time than the up-going wavefield. At earth location P2, which is the reflector depth, the two wavefields crosscorrelate and an image is formed. Adapted from Claerbout (1971). antoine3-updown [NR]

In general shot-profile migration is performed in the ($\omega$,x) domain one frequency at a time and one shot at a time. The final image is formed by adding all the different contributions of every shot together.

**Imaging of multiples**

A similar machinery can be effectively used to image multiples at their correct location in the subsurface. I keep the same imaging principle as developed by Claerbout. The differences stem from the choice of up- and down-going wavefields I extrapolate.

Figure 2 illustrates the basic idea behind the migration of the multiples. In Figure 2(a), a wavefield generated at S is recorded at the receiver $R_n$. The reflector location $r_a$ is imaged by extrapolating both the primary wavefield recorded at $R_n$ and the source wavefield at $S$ simultaneously in the subsurface and by crosscorrelating them at each depth step.

Similarly in Figure 2(b), a multiple recorded at $R_n$ can be used to image the reflector location $r_b$ if we use the primary wavefield recorded at the receiver $R_1$ as a source function. Hence a multiple reflection recorded at any receiver location can be used to image the subsurface if a primary is utilized as a source.

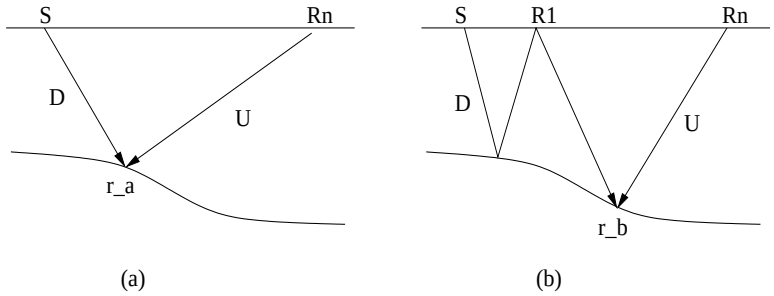



Figure 2: Illustration of the basic idea of reflector mapping with (a) primaries and (b) multiples. In (a) the primary images the reflector location $r_a$ and the source is impulsive. In (b) the multiple helps to image the reflector location $r_b$ and the source is a primary recorded at $R_1$. antoine3-primult [NR]

Finding the exact location of $R_1$ for each multiple and each receiver position $R_n$ would require an earth-model and many tedious computation steps. Fortunately, the imaging condition tells us that the image is formed if and only if the up- and down- going wavefields correlate. Therefore for a given receiver position $R_n$, we can use as a source function every single primary recorded on the seismic array since only relevant receiver positions $R_1$ would produce constructive crosscorrelations.

Now if we expand this idea to every receiver position $R_n$ on the seismic array, we can produce an image of the earth by simply taking every recorded primary as the source function and every recorded multiple as the up-going wavefield. The impulsive source becomes an areal-shot record. In theory, any order of multiples can be properly imaged if their corresponding source path exists in the down-going wavefield. Hence first order multiples need primaries as sources, second order multiples need first order multiples, and so on.

Note that the source function needs to be time-reversed before the extrapolation. This is done in the ($\omega$,x) domain by computing the complex conjugate of the source wavefield.

A similar approach has been presented by Berkhout and Verschuur (1994) using the so-called “WRW” model. Notice that so far, this approach works for surface-related multiples only but could be easily extended to internal-multiple migration.

## A SYNTHETIC DATA EXAMPLE

In this section I use a modified version of the 2.5-D Amoco dataset (Etgen and Regone, 1998; Dellinger et al., 2000) to illustrate the migration of multiples. This data example proves that multiples can provide structural information on the earth.

**The Amoco dataset**

The synthetic dataset I use consists of a modified version of the 2.5-D Amoco velocity model. Figure 3 displays the velocity model. I have added a 500 meters water layer to generate water column reverberations. Two finite-differences modelings were done with and without free-surface conditions in order to easily extract surface-related multiples. The data were recorded with a split-spread geometry. 32 shots were acquired and processed for the final images.

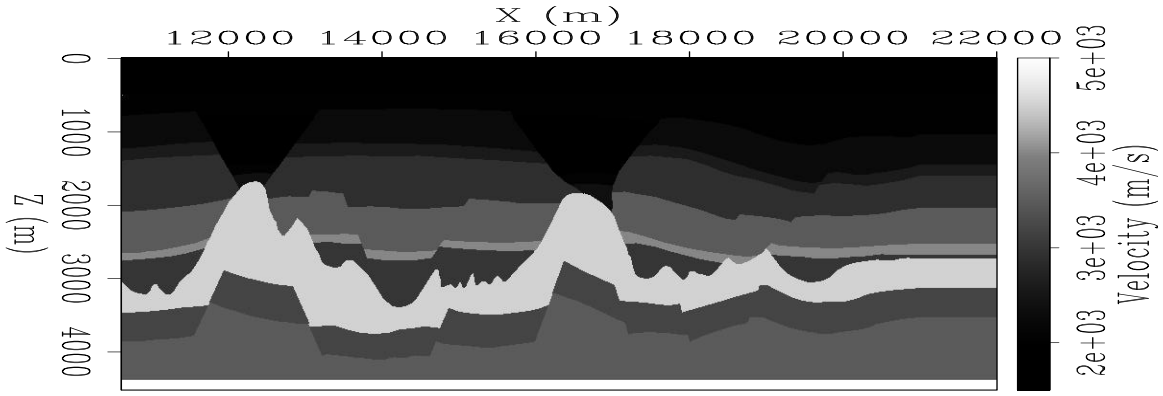



Figure 3: The velocity model used to generate the synthetic data. antoine3-velamoco [ER]

**Migration of one shot record**

I illustrate the migration of multiples with one shot. In Figure 4 I display the source function on the left panel, and the recorded wavefield on the right panel. The recorded wavefield is the superposition of primaries and multiples. The migration result for this shot location is shown in Figure 6a.

Figure 5 displays the source function and the up-going wavefield for the migration of multiples. As proposed in the preceding section, the source function is not impulsive but areal. The recorded wavefield contains the surface-related multiples only. The migration result is shown in Figure 6b and compares favorably with the output of the migration of primaries.

It is interesting to note that the water-bottom is illuminated with a wider aperture when multiples are used. As illustrated in Figure 2, for a given receiver $R_n$, the primary illuminates the reflector in $r_a$ at a closer location to the source in $S$ than does the multiple in $r_b$. Therefore, for one given shot, the multiples migrate with a wider aperture but with smaller angles.

As a final result, I display in Figures 6c and 6d the different illumination patterns for both the impulsive and areal sources (Rickett, 2001). These maps are obtained by simply computing the amplitude (squared) of the down-going wavefield at each depth and for each frequency. These Figures illustrate the aperture and illumination effects of the multiples.

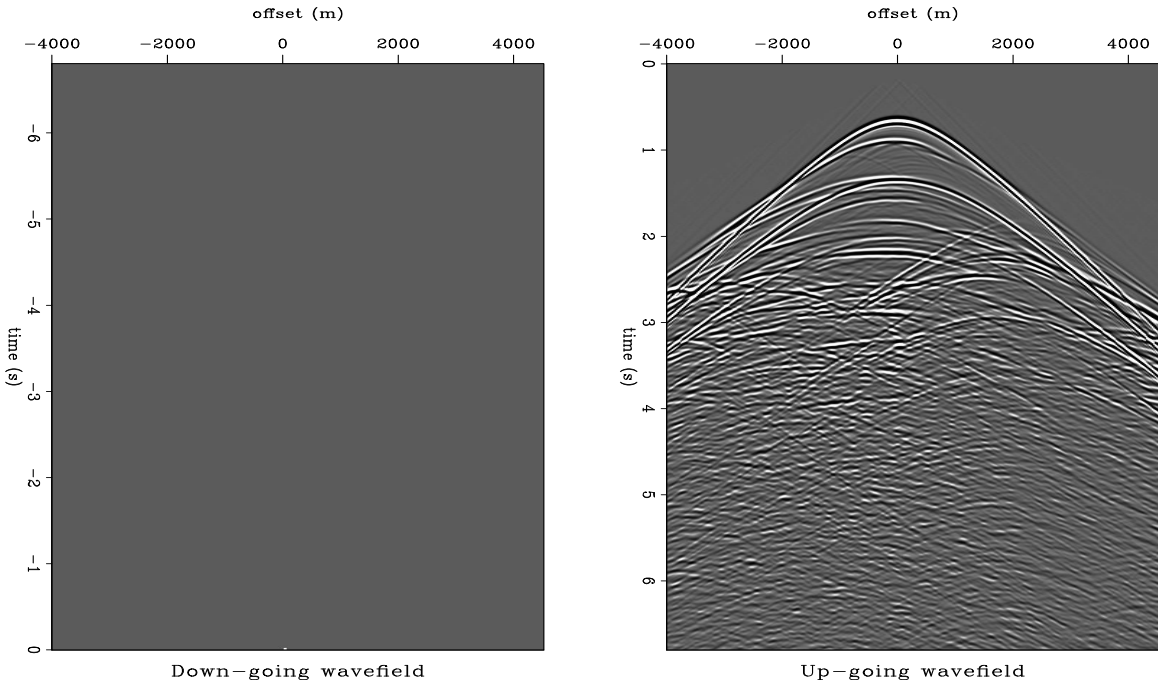



Figure 4: Shot-profile migration of the primaries. Left: Impulsive source function. Look closely at offset 0. meter and time 0. second to see the impulse. Right: Recorded primaries and multiples. antoine3-shotprim14 [ER]

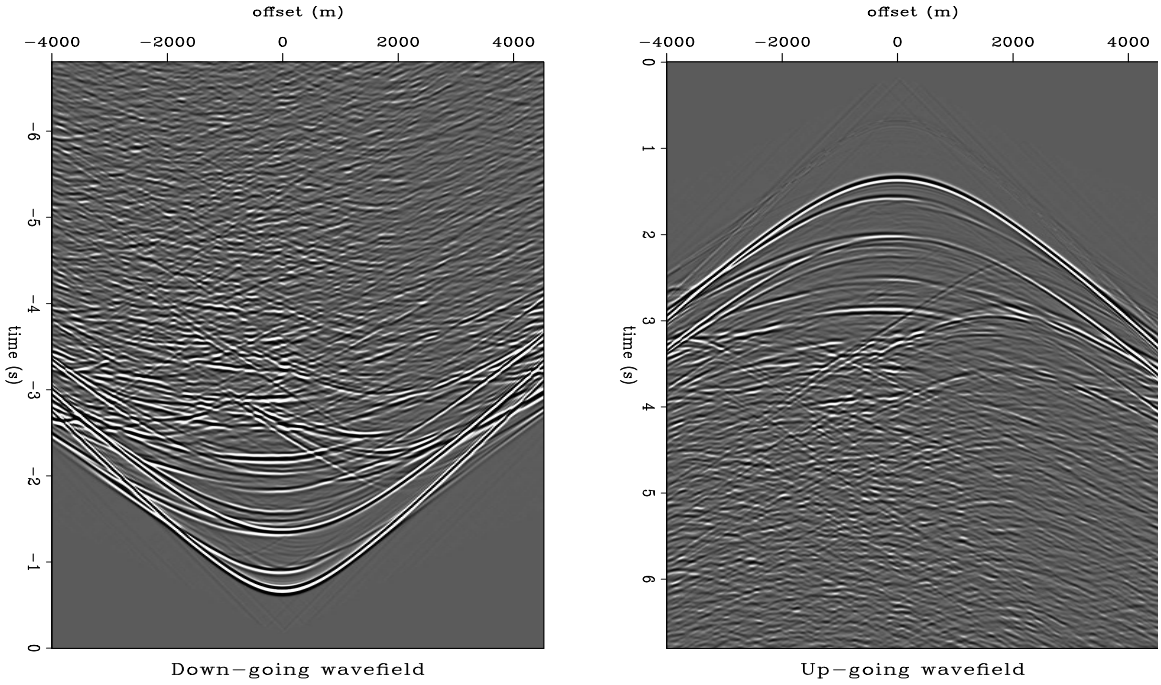



Figure 5: Shot-profile migration of the multiples. Left: Areal source function. Right: Recorded multiples only. The source function is plotted upside-down to represent the computation involved by the time-reversed process. antoine3-shotmig14 [ER]

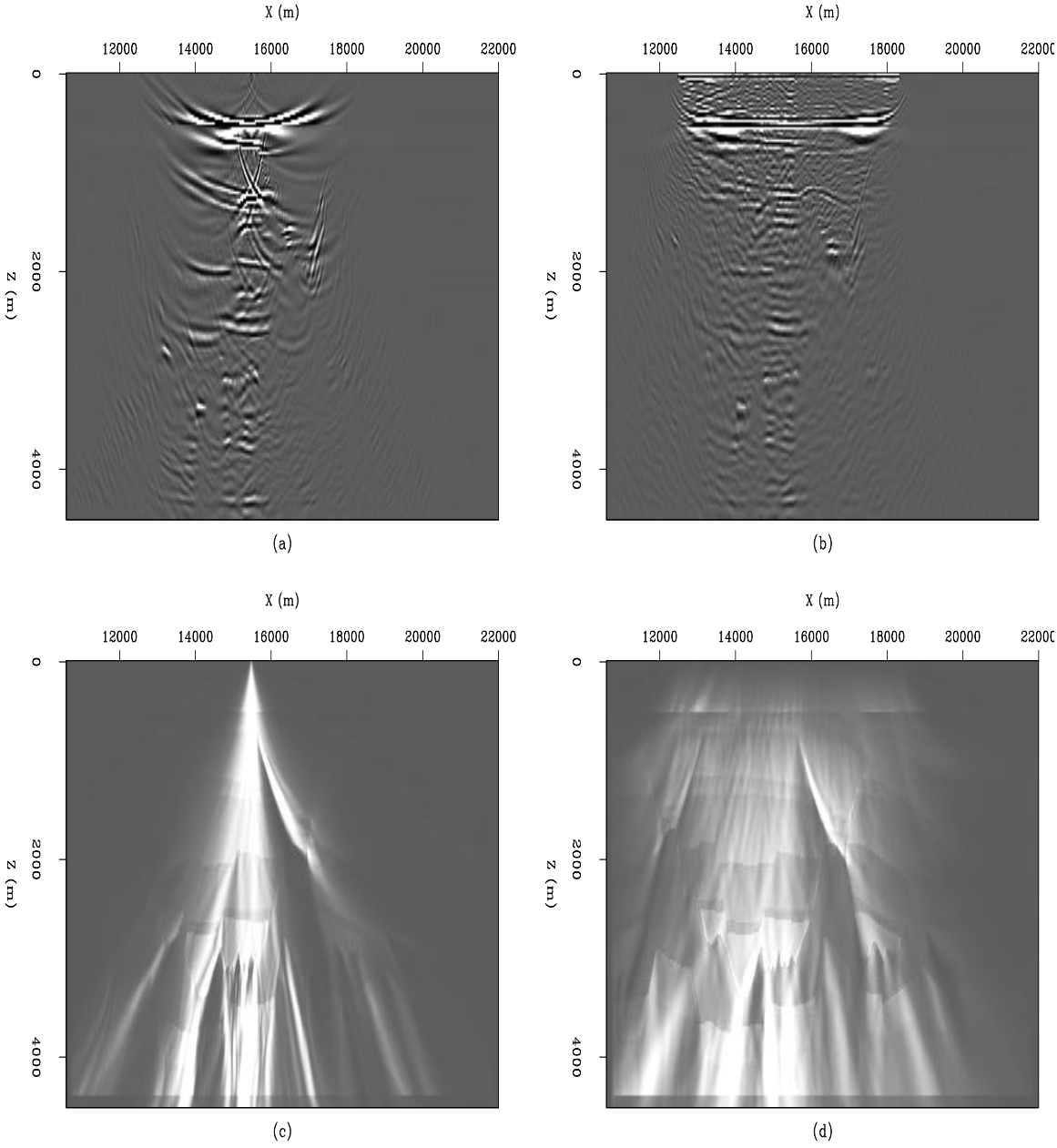



Figure 6: Migration results and illumination maps of the down-going wavefields for the imaging of multiples and primaries. (a) Migration of the primaries. (b) Migration of the multiples. (c) An illumination map for the impulsive source. (d) An illumination map for the areal source. antoine3-compil14 [ER,M]

### Migration of the survey

I now proceed to the final imaging steps of migrating every shot record and summing them. Figures 7a and 7b show the final migration result for both primaries and multiples. The two final images are quite similar although the migration of the primaries with the impulsive source is crispier and less noisy than with multiples. Nevertheless the migration of multiples yields an accurate image of the earth. After closer inspection, it seems that the flanks of the canyon in the middle above the salt dome is better defined with the multiples.

The illumination maps displayed in Figures 7c and 7d are almost identical after summation of all shot contributions. This comes from the acquisition geometry at the two ends of the survey. This first result is rather encouraging and shows that multiples can be used to image the earth.

### Other migration results with multiples

I see two fundamental problems with the migration of the multiples. First, the final image is noisy, blurring precious information in some areas. Second, multiples need to be extracted from the data.

A good trick to decrease the noise level in the migration result is to use the water-bottom reflection only for the source. This can be done most of the time by applying a simple mute to the data. Figure 8 shows the up- and down-going wavefields if the water-bottom reflection is used as a source. Figure 9 displays a comparison of the migration of multiples with all the primaries as the source (Figure 9a) and the water-bottom as the source (Figure 9b). The noise level has decreased without loosing structural information. This result proves that for these data, most of the surface-related multiples are water-column reverberations.

The need for separating multiples might be quite dissuasive when field data are imaged because it might involve heavy computations and/or an earth model that might not be known in advance. Let us consider for a moment that we do not want to do the multiple attenuation but still want to do some imaging with multiples. This goal requires that the recorded data with primaries and multiples are used for both up- and down-going wavefields (Figure 10). Now I compare in Figure 11 the migration results when only multiples (Figure 11a) and multiples plus primaries (Figure 11b) are extrapolated in the up-going wavefield. The migration of primaries and multiples yields a noisy image but never-the-less structurally interpretable.

In the next section I migrate multiples for two 2-D lines. One survey comes from the Gulf of Mexico and the other from the North Sea. They demonstrate potential strengths and weaknesses of the proposed method for multiple migration.

## TWO 2-D FIELD DATA EXAMPLES

I illustrate the multiple migration with field data. The first dataset is a deep-water survey with a shallow salt body. The second dataset is a North Sea survey with a relatively thick basalt

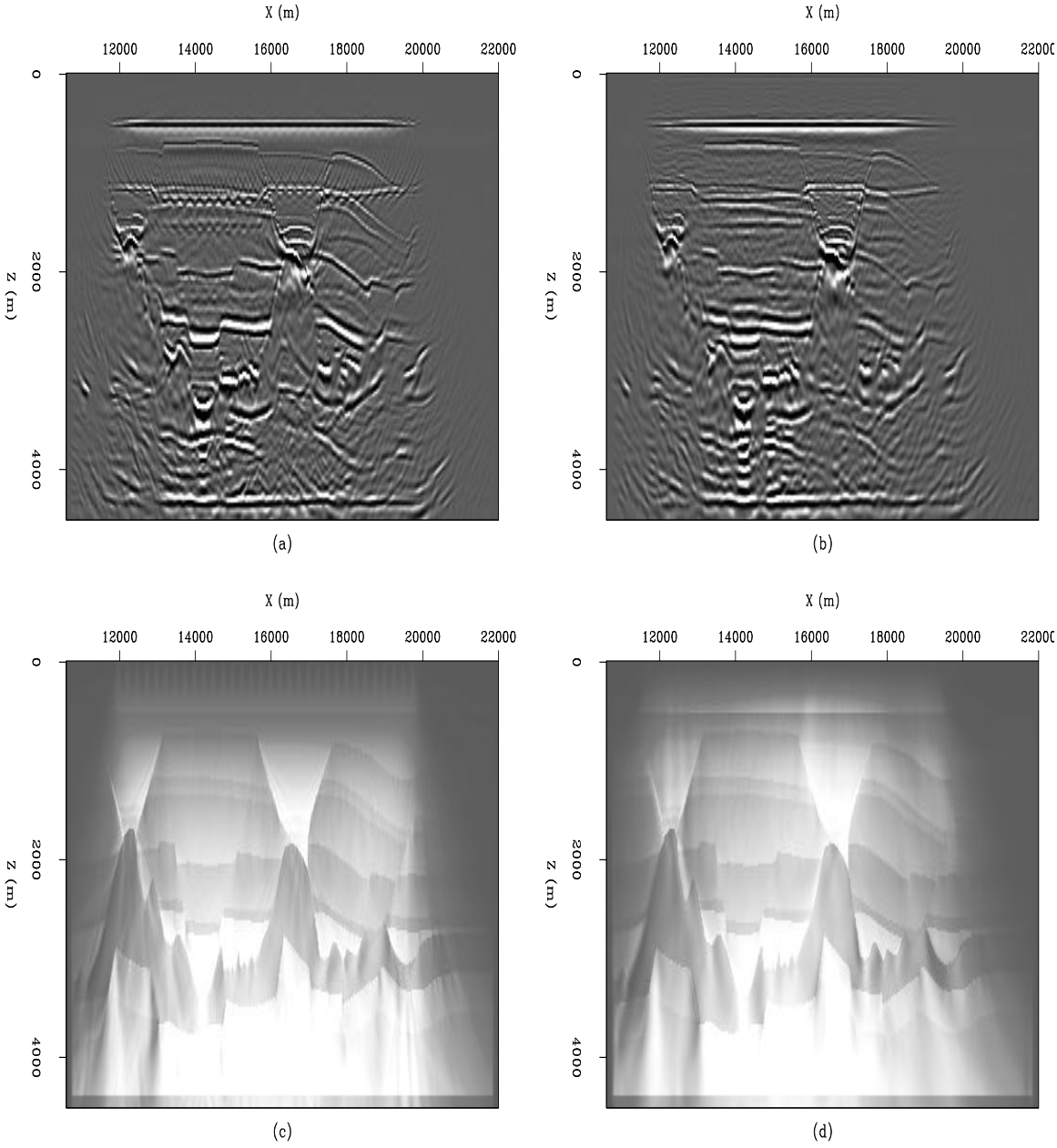



Figure 7: Final migration results and illumination maps. (a) Migration of the primaries. (b) Migration of the multiples. (c) An illumination map for all the impulsive sources. (d) An illumination map for all the areal sources. antoine3-compmig [CR,M]

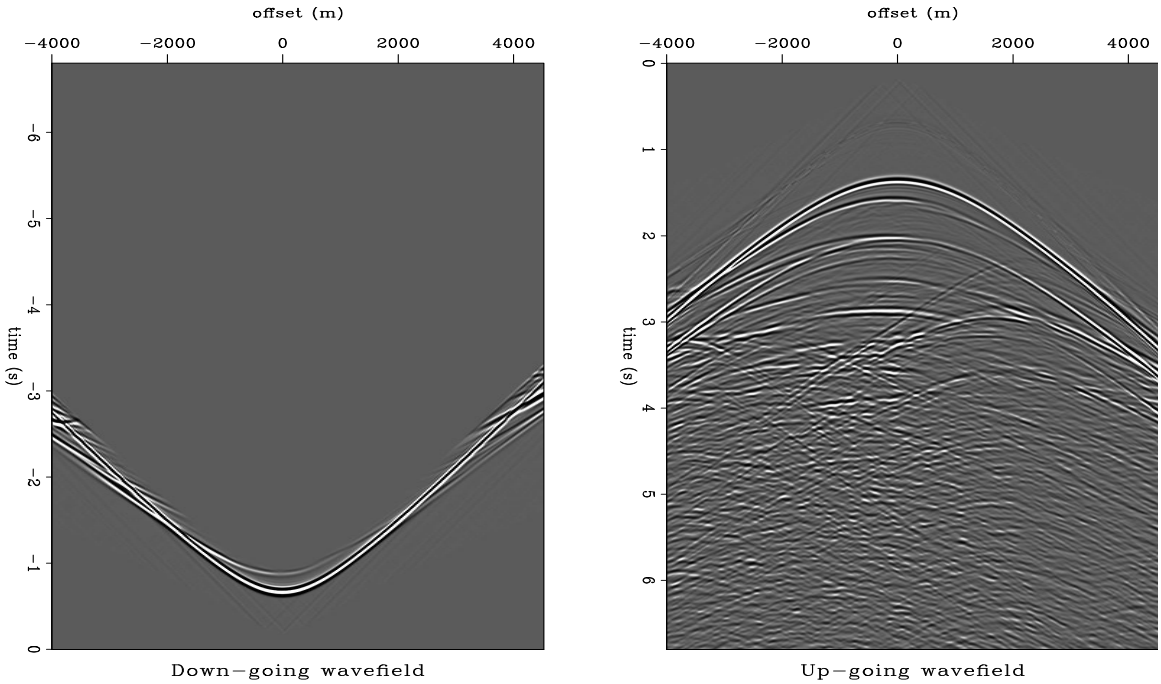



Figure 8: Shot-profile migration of the multiples. Left: Areal source function with the water-bottom reflection only. Right: Multiples. antoine3-wbmig14 [ER]

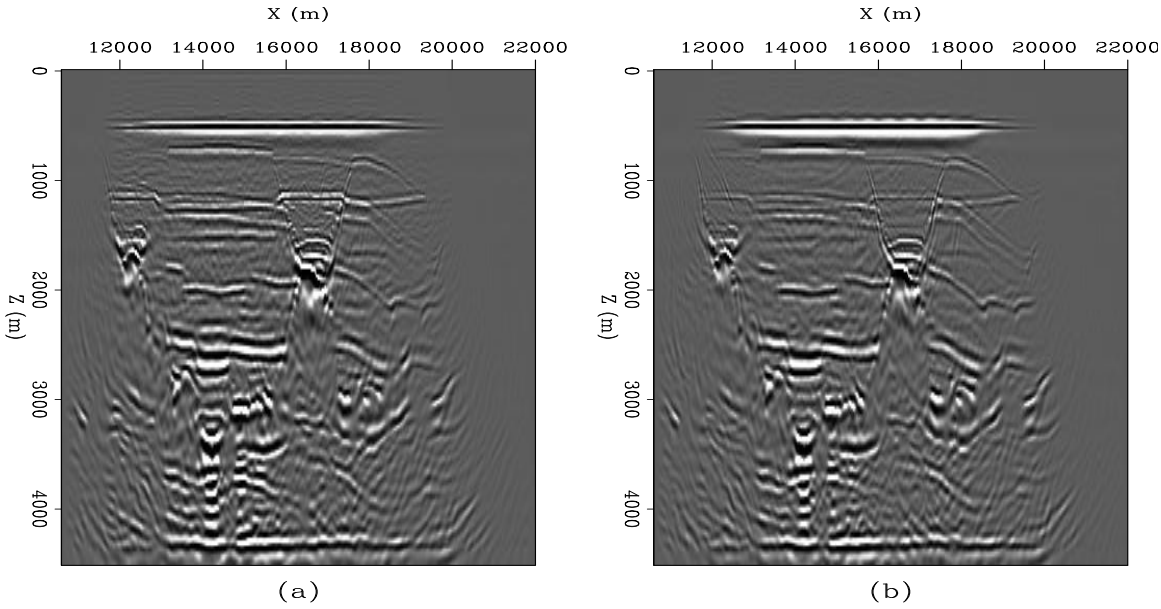



Figure 9: Migration results. (a) Multiples and primaries are used as a source. (b) The water-bottom reflection is used as a source. antoine3-compwb [CR,M]

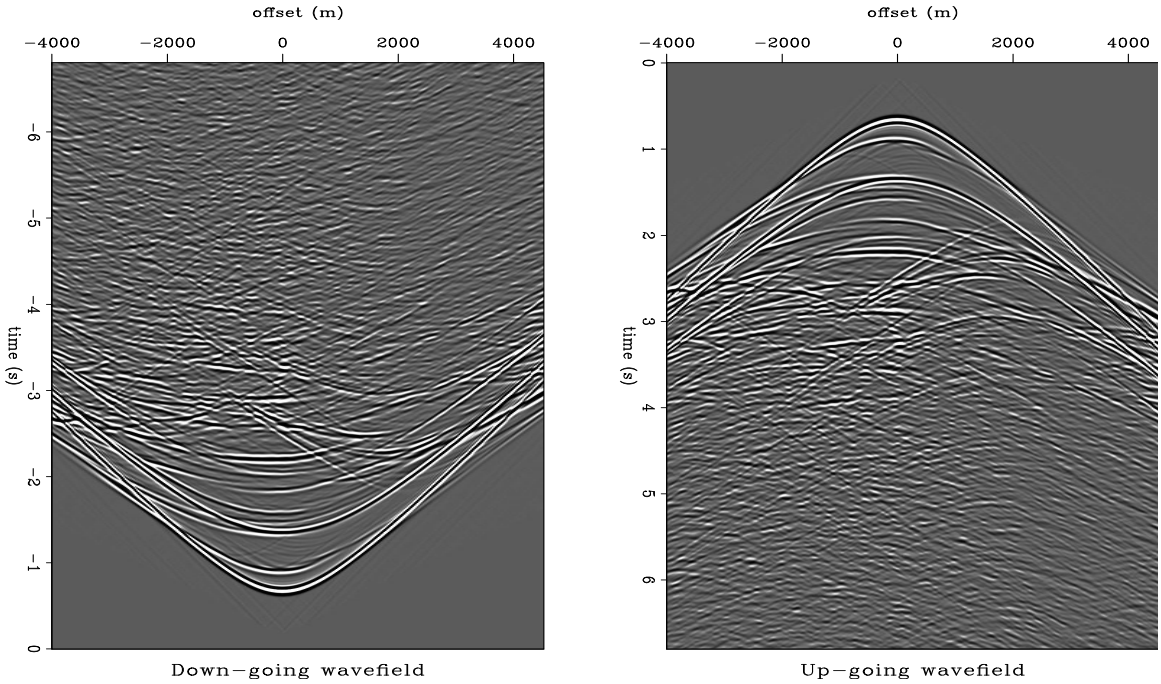



Figure 10: Shot-profile migration of the multiples. Left: Areal source function with primaries and multiples. Right: Multiples and primaries are extrapolated. antoine3-allall14 [ER]

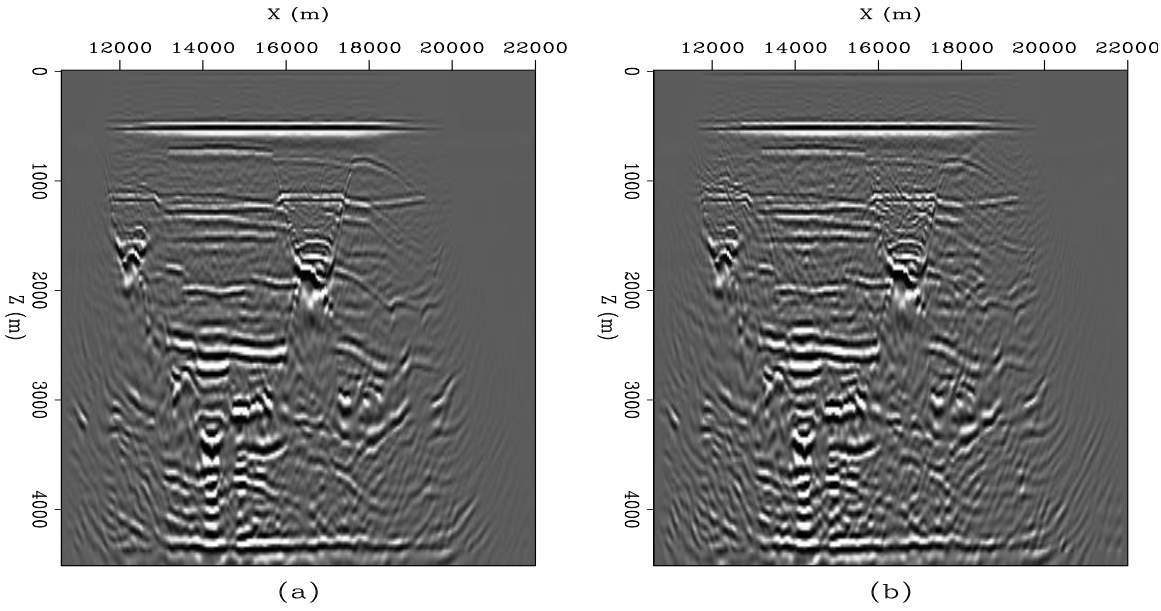



Figure 11: Migration results. (a) Same as Figure 7(b). (b) Imaging of multiples with primaries and multiples in the up-going wavefield. antoine3-compall [CR,M]

slab.

### A Gulf of Mexico example

This dataset has been intensively used for testing multiple attenuation techniques (The Leading Edge, January 1999). Figure 12 shows the velocity model that is used for the migration (Gratwick, 2001).

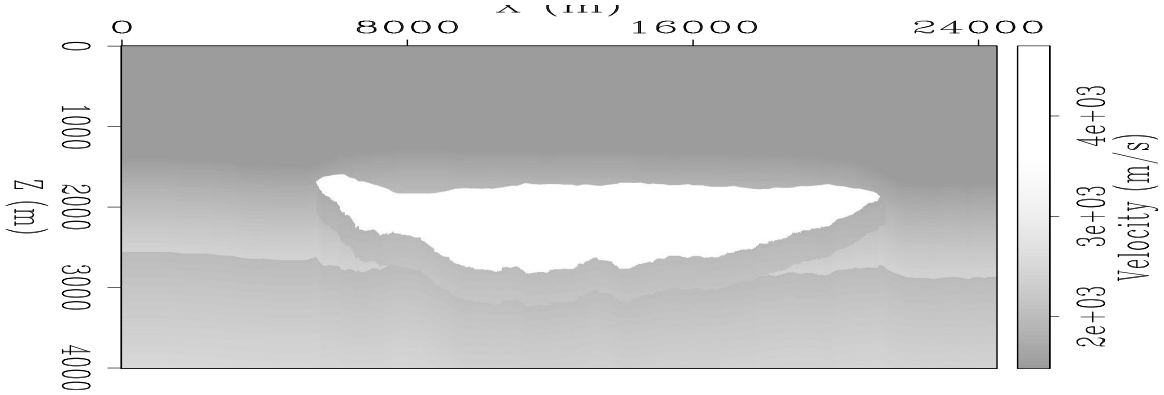



Figure 12: Velocity field for the Gulf of Mexico 2-D line. antoine3-velgom [ER]

For this data set I did not separate primaries from multiples as required by the theory. As an approximate solution, I decided to mute everything above the first surface-related multiple for the up-going wavefield. Therefore the down-going wavefield is made of the complete shot and the up-going wavefield is made of the complete shot minus the primaries above the first surface-related multiple (Figure 13).

The migration results are displayed in Figure 14. Figures 14a and 14b show the migration results for primaries and multiples respectively. The two migrations produce similar results. Despite the approximations made, the multiple migration gives a very accurate image of the salt body and of the sediments.

### A North Sea example

The North Sea dataset illustrates how challenging basalt imaging can be. The data are infested with internal and surface-related multiples. In addition, the internal layering of the basalt layer caused by repetitive basalt flows interlaced with sedimentary deposits make the identification of the bottom of the basalt very difficult. Kostov et al. (2000) attempted P-wave imaging with long-offset streamer data. Brown et al. (2001) tried imaging locally-converted shear waves.

Because multiples are so strong in this dataset, multiple migration might unravel some useful information. Again I did not separate multiples from primaries. I muted the up-going wavefield, as with the Gulf of Mexico data, to remove all the data above the first surface-related multiple.

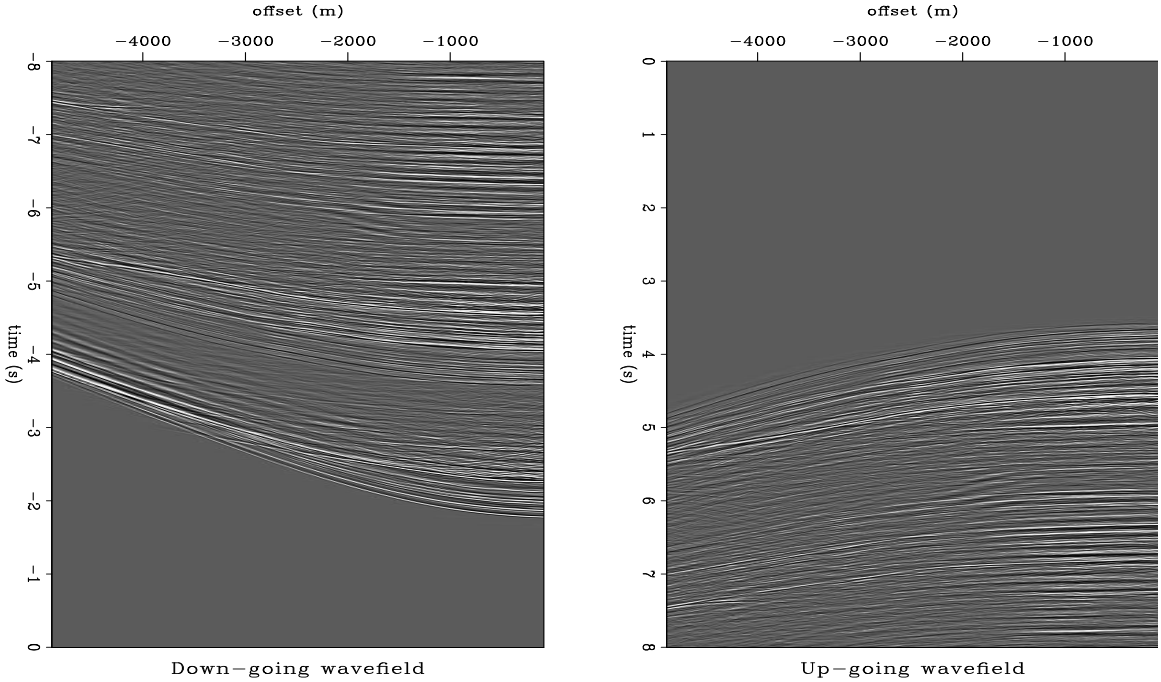



Figure 13: Left: The source function with primaries and multiples. Right: Same as the source function but with a mute above the first surface-related multiple. antoine3-gom_ud_mult500 [ER]

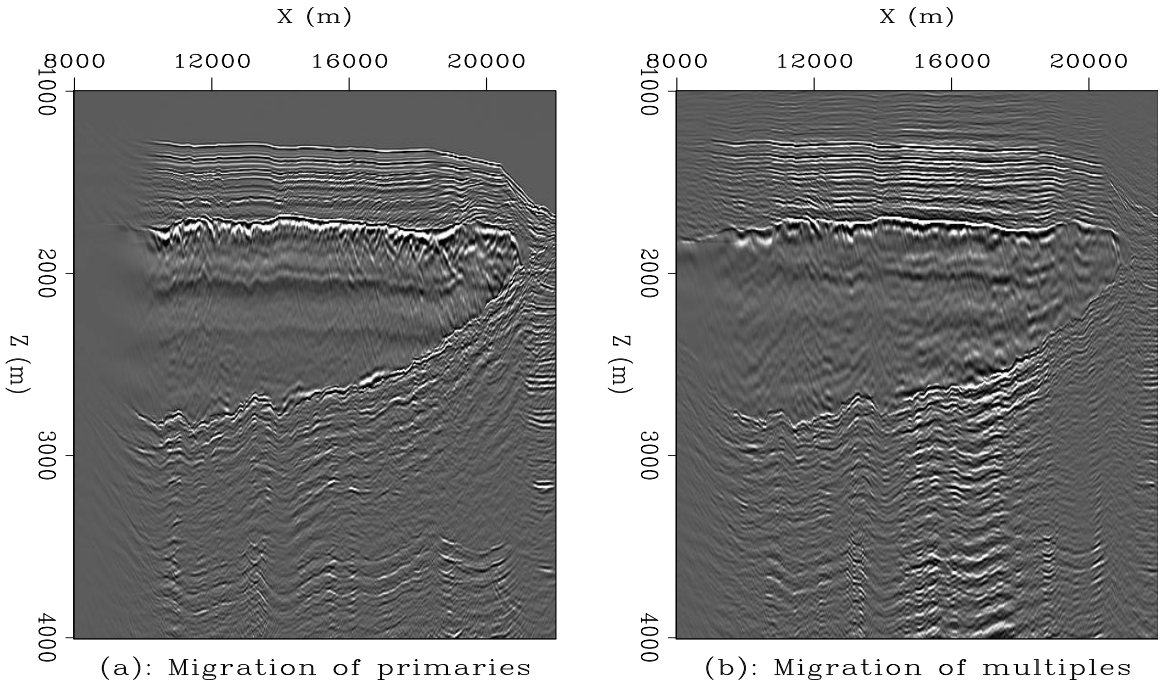



Figure 14: (a) Migration of the primaries. (b) Migration of the multiples. The two images are comparable. The salt body is clearly visible in the multiple migration result. antoine3-gom_comp [CR,M]

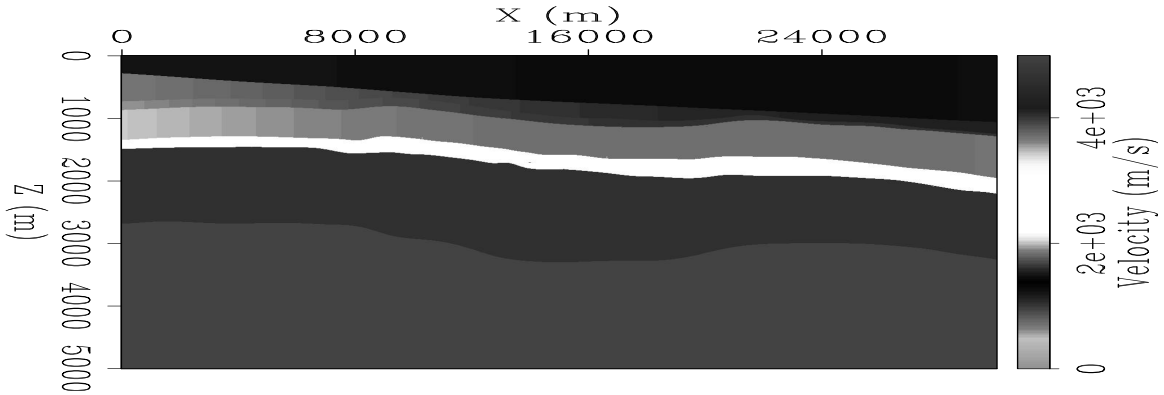



Figure 15: Velocity field for the North Sea 2-D line. The top of the basalt corresponds to the top of the white layer. The bottom of the basalt corresponds to the bottom of the next layer, around 2500 meters. antoine3-velns [ER]

The velocity model for the North Sea dataset is displayed in Figure 15. I show one example of up- and down-going wavefields in Figure 16. Notice the strong reverberations of the water column plus other surface-related/internal multiples.

I migrated the North Sea data without large offsets and for a small portion of the survey only. Figure 17 shows the imaging results when primaries (Figure 17a) and multiples (Figure 17b) are migrated. It appears that the multiple image is less focused than the primary image. In addition the water-bottom is not at the correct location for both images. What happened ?

The answer is simple: cable feathering. Kostov et al. (2000) indicates that cable feathering was quite strong for the subset I migrated. Therefore the shot S, $R_1$ and $R_n$ in Figure 2 are not in the same plane anymore introducing positioning errors in the migration.

## DISCUSSION AND CONCLUSION

I have shown that migrating multiples at their correct location is possible. Results with a complex 2-D synthetic model and field data examples prove that multiples can bring valuable structural information. Although very encouraging, multiple migration has some limitations.

First, we need to generate a multiple model with correct kinematics and amplitudes. This can be quite a computation burden. Nonetheless, multiple attenuation is a prerequisite to any imaging step. Thus, instead of trashing these multiples, we might simply use them for migration. As an intermediate solution, I propose to use both primaries and multiples in the up-going wavefield. This is a cheap alternative to the full multiple attenuation that yields an interpretable image of the subsurface.

Second, the final image after migration of multiples is more noisy than the migration of primaries. In the synthetic data example I proposed using the water-bottom reflection for the source function. In practice the amount of muting depends on the subsurface, i.e., the main

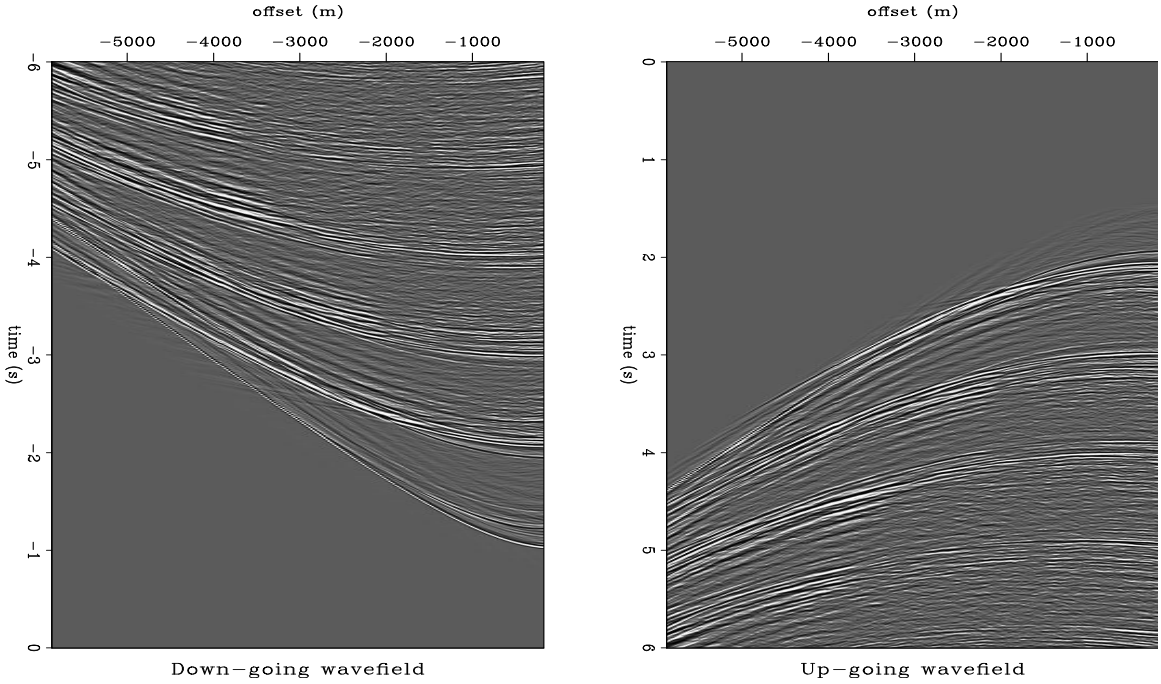



Figure 16: Left: The source function with primaries and multiples. Right: Same as the source with a mute. antoine3-ns_ud_mult100 [ER]

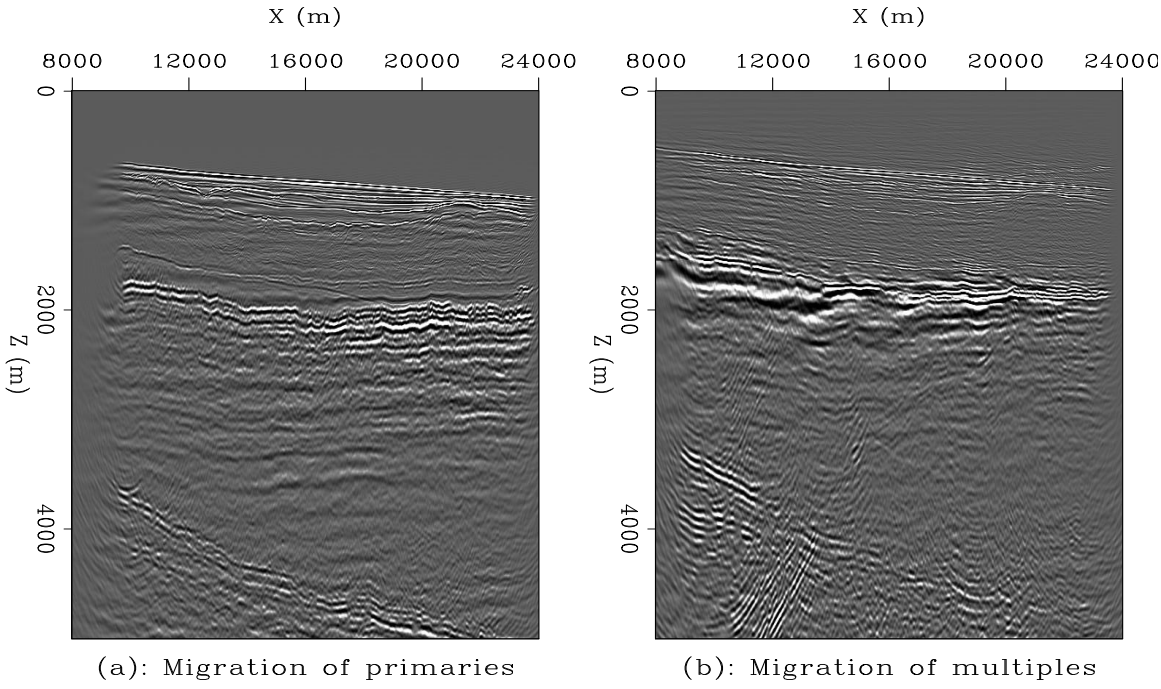



Figure 17: (a) Migration of the primaries. (b) Migration of the multiples. The two images are different because of cable feathering. antoine3-ns_comp [CR,M]

"generators" of multiples.

Last, we might not be able to simply add the different images in order to increase the signal-noise ratio. When primaries are migrated, we use a synthetic source that is not the true seismic source. When multiples are migrated, the source is perfectly taking into account because we use the recorded wavefield as a source. Hence a direct addition of the migration results of primaries and multiples must be done with care.

As a final comment, the North Sea example perfectly illustrated the need for a targeted preprocessing of the data. In this case I should have corrected for cable feathering before multiple migration. Similar preprocessing steps are needed for the surface-related multiple attenuation technique (Verschuur et al., 1992).

## ACKNOWLEDGMENTS

I would like to thank the members of the SMAART 2 JV, for their financial support during my stay in Delft, Eric Verschuur who got me started, and Amoco/WesternGeco for providing the synthetic velocity model and field data respectively. I address a special thanks to Ray Ergas, now retired from Chevron, for his enthusiastic support before, during, and after my summer in Delft.

## REFERENCES

Berkhout, A. J., and Verschuur, D. J., 1994, Multiple technology: Part 2, migration of multiple reflections: 64th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1497–1500.

Brown, M., Biondi, B., and Kostov, C., 2001, Wave-equation prestack depth migration for sub-basalt P and converted wave imaging: SEP–**110**, 19–30.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, no. 03, 467–481.

Dellinger, J. A., Gray, S. H., Murphy, G. E., and Etgen, J. T., 2000, Efficient 2.5-D true-amplitude migration: Geophysics, **65**, no. 03, 943–950.

Etgen, J., and Regone, C., 1998, Strike shooting, dip shooting, widepatch shooting - Does prestack depth migration care? A model study.: 68th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 66–69.

Gratwick, D., 2001, Amplitude analysis in the angle domain: SEP–**108**, 45–62.

Guitton, A., Brown, M., Rickett, J., and Clapp, R., 2001, Multiple attenuation using a t-x pattern-based subtraction method: 71st Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1305–1308.

Jacobs, B., 1982, Imaging common shot gathers: SEP–**30**, 7–28.

Kostov, C., Hoare, R., Jasund, S., and Larssen, B., 2000, Advances in sub-basalt P-wave imaging with long offset streamer data: 62nd Mtg., Eur. Assn. Geosci. Eng., Expanded Abstracts, Session:X0018.

Rickett, J., 2001, Model-space vs. data-space normalization for finite-frequency depth migration: 71st Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 2081–2084.

Verschuur, D. J., Berkhout, A. J., and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, no. 09, 1166–1177.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 33–47*

# Least-squares joint imaging of primaries and multiples

*Morgan Brown*[1]

**ABSTRACT**

Multiple reflections provide redundant, and sometimes additional, information about the corresponding primary reflections. I implement a least-squares inversion scheme to jointly image (by normal moveout) primaries and multiples, with the goal of enforcing consistency between the images and the input data. Furthermore, to effect noise ("crosstalk") suppression, I introduce a novel form of model regularization which exploits kinematic similarities between imaged primaries and multiples, and which also preserves the amplitude-versus-offset (AVO) response of the data. In tests on synthetic data, my approach exhibits good noise suppression and signal preservation characteristics. Real data tests highlight the need for careful data preprocessing. Future work points toward use of migration as the imaging operators, to exploit cases where multiples actually exhibit better angular coverage than primaries, and thus add new information to the inversion.

## INTRODUCTION

Multiple reflections have long been treated as noise in the seismic imaging process. In contrast to many other types of "noise," like surface waves, multiply reflected body waves may still penetrate deep into the earth, and thus have a potential to aid in imaging the prospect zone. I refer generically to *joint imaging with multiples* as any process which creates a "pseudo-primary" image from multiples by removing the propogation effects of body waves through arbitrary multiple layer (generator + free surface), and which then seeks to integrate the information provided by the primary and pseudo-primary images.

Reiter et al. (1991) present an early example of imaging multiples directly using a prestack Kirchhoff scheme. Yu and Schuster (2001) describe a cross-correlation method for imaging multiples. Berkhout and Verschuur (1994) and Guitton (2002) apply shot-profile migration for multiples. The aforementioned approaches produce separate-but-complementary pseudo-primary and primary images, yet they either do not attempt to, or employ simplistic methods to integrate the information contained in the two images; either add (Reiter et al., 1991) or multiply (Yu and Schuster, 2001) them together.

In this paper, I introduce a new methodology for jointly imaging primaries and multiples. In addition to a desire to correctly image the multiples, my approach is driven by three primary motivations:

[1]**email:** morgan@sep.stanford.edu

1. **Data Consistency** - The primary and pseudo-primary images both should be maximally consistent with the input data.
2. **Self-consistency** - The primary and pseudo-primary images should be consistent with one another, both kinematically and in terms of amplitudes.
3. **Noise Suppression** - In the primary image, all orders of multiples should be suppressed. In the pseudo-primary image created from, say first-order water-bottom multiples, contributions from primaries and secord-order or greater multiples should be suppressed.

Least squares optimization provides an excellent, and perhaps the only viable approach to address all three requirements. I adopt an approach similar to Nemeth et al. (1999), which used a least-squares scheme to jointly image compressional and surface waves, for improved wavefield separation. Data consistency is effected by minimization of a data residual; self-consistency and noise suppression through the use of regularization terms which penalize 1) differences between primary and pseudo-primary images, and 2) attributes which are not characteristic to true primaries or pseudo-primaries.

In my approach, I use the simplest possible imaging operation, Normal Moveout (NMO). I derive an NMO equation for water-bottom multiple reflections, which maps these multiples to the same zero-offset traveltime as their associated primaries, creating a "pseudo-primary" section. To account for the amplitude differences between the primary and pseudo-primary sections, I assume constant seafloor AVO behavior and estimate a single water-bottom reflection coefficient from the data. To address the AVO differences between primary and pseudo-primary, I derive an expression – valid only for constant velocity – for the AVO of the pseudo-primary as a function of the AVO of the primary, and then enforce this constraint in the inversion via an offset- and time-dependent regularization term.

## METHODOLOGY

### NMO for Multiple Reflections

In a horizontally-stratified, $v(z)$ medium, multiple reflections can be treated as kinematically-equivalent primaries with the same source-receiver spacing but additional zero-offset traveltime $\tau^*$, as illustrated in Figure 1. We can write an extension to the NMO equation which flattens multiples to the zero-offset traveltime of the reflector of interest.

$$t^2 = \sqrt{(\tau + j\tau^*)^2 + \frac{x^2}{V_{eff}^2}} \tag{1}$$

$j\tau^*$ is the two-way traveltime of a $j^{\text{th}}$-order multiple in the top layer. $V_{eff}(\tau)$ is the effective RMS velocity of the equivalent primary shown in the figure. For the simple case of constant velocity $v^*$ in the multiple-generating layer,

$$V_{eff}(\tau) = \frac{\tau^* v^* + \tau V(\tau)}{\tau^* + \tau} \tag{2}$$

So for the common case of relatively flat reflectors, $v(z)$, and short offsets, equation (1) should do a reasonable job of flattening water-bottom multiples of any order to the $\tau$ of interest, assuming that we pick the water bottom ($\tau^*$) and that we know the seismic velocity of water.

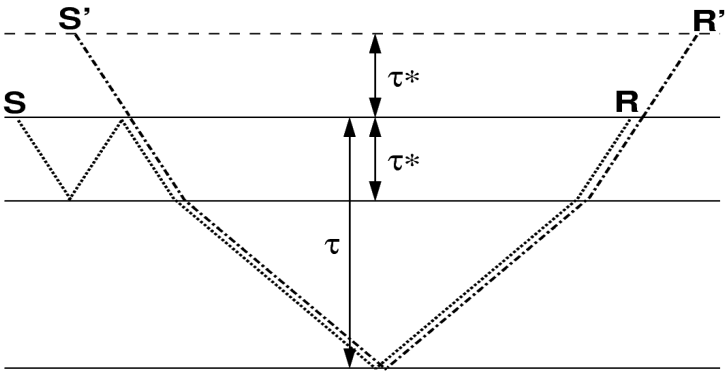



Figure 1: Schematic for NMO of multiples. From the standpoint of NMO, multiples can be treated as pseudo-primaries with the same source-receiver spacing, but with extra zero-offset traveltime $\tau^*$, assuming that the velocity and time-thickness of the multiple layer are known. morgan1-schem [NR]

**AVO of Multiple Reflections**

Even after application of the water-bottom reflection coefficient, the AVO response of the pseudo-primary section created by equation (1) does not match that of the corresponding NMO-corrected primary section. Refer to Figure 2 and note that for constant-AVO water-bottom reflection (and a free surface reflection coefficient of -1), the amplitude of the water-bottom multiple at offset $h_p + h_m$ is simply the amplitude of the primary at offset $h_p$, scaled by the negative water-bottom reflection coefficient. Still, the question remains: *what are $h_m$ and $h_p$*? For the case of constant velocity, we can use trigonometry to derive $h_m$ and $h_p$ as a function of the zero offset traveltimes of the primary reflection and water bottom ($\tau$ and $\tau^*$, respectively), and the source-receiver offset $x$. In constant velocity, the multiple and primary legs of the raypath are similar triangles:

$$\frac{\tau v}{h_p} = \frac{\tau^* v}{h_m}. \tag{3}$$

Also, for a first-order water-bottom multiple,

$$h_p + h_m = x.$$

These two independent equations can be solved and simplified to give expressions for $h_p$ and $h_m$:

$$h_p = \frac{\tau}{\tau + \tau^*} x \qquad \text{and} \qquad h_m = \frac{\tau^*}{\tau + \tau^*} x. \tag{4}$$

I omit the general form of the expression for orders of multiple higher than one, although it is straightforward to derive.

To obtain an estimate of the water-bottom reflection coefficient, I solve a simple least squares problem to estimate a function of location, $\mathbf{a}(x)$, which when applied to a small window of dimension $nt \times nx$ around the NMO-corrected water-bottom reflection, $\mathbf{p}(t,x)$, op-

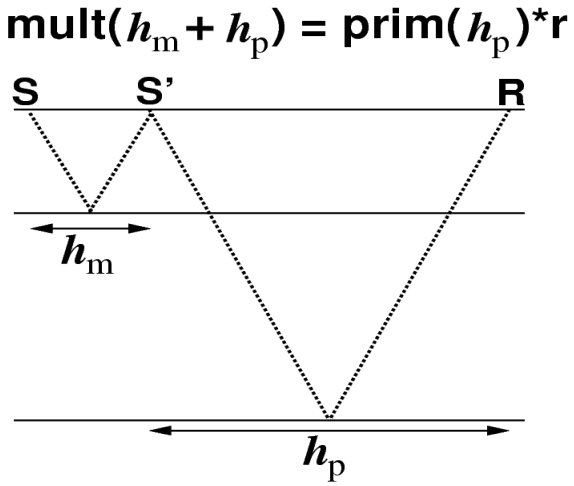



Figure 2: Assuming a constant AVO water-bottom reflection and constant velocity, we can write the AVO of water-bottom multiples with offset $h_p + h_m$ as a function of the AVO of the primary recorded at a shorter offset, $h_p$. morgan1-avo [NR]

timally resembles the NMO-corrected [equation (1)] first-order water-bottom multiple reflection, $\mathbf{m}(t,x)$. To achieve this, $\mathbf{a}(x)$ is perturbed to minimize the following quadratic functional.

$$\min \left( \sum_{j=1}^{nx} \sum_{i=1}^{nt} a(j) * p(i,j) - m(i,j) \right)^2 \tag{5}$$

$\mathbf{a}(x)$ may not be reliable at far offsets, due to either NMO stretch or non-hyperbolicity, so in practice, an estimate of the single best-fitting water-bottom reflection coefficient is made using the $\mathbf{a}(x)$ from "useful" offsets only.

**Least-squares imaging of multiples**

Applied to a common-midpoint gather, equation (1) produces an approximate unstacked zero-offset image of pseudo-primaries from water bottom multiple reflections. In this section, I introduce a least squares scheme to compute self-consistent images of primaries and pseudo-primaries which are in turn consistent with the data. First I define some terms:

| | | |
|---|---|---|
| $\mathbf{d}$ | $\leftrightarrow$ | One CMP gather. |
| $\mathbf{m}_j$ | $\leftrightarrow$ | Prestack model vector for multiple order $j$. Produced by applying equation (1) to $\mathbf{d}$. |
| $\mathbf{N}_j$ | $\leftrightarrow$ | *Adjoint* of NMO for multiple of order $j$ (primaries: $j = 0$). $\mathbf{N}_j^T$ applies equation (1) to $\mathbf{d}$. |
| $\mathbf{R}_j$ | $\leftrightarrow$ | Given a single water-bottom reflection coefficient, $r$, [estimated via equation (5)], this operator scales $\mathbf{m}_j$ by $1/r^j$ to make the amplitudes of all the $\mathbf{m}_j$ comparable. |

With these definitions in hand, we can now write the forward modeling operator for joint NMO of primaries and multiples of order 1 to $p$.

$$\left[\mathbf{N}_0 \;\; \mathbf{N}_1\mathbf{R}_1 \;\cdots\; \mathbf{N}_p\mathbf{R}_p\right] \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{m}_1 \\ \vdots \\ \mathbf{m}_p \end{bmatrix} = \mathbf{L}\mathbf{m} \tag{6}$$

In words, equation (6) takes a collection of psuedo-primary panels, divides each by the appropriate reflection coefficient, applies inverse (adjoint) NMO to each, and then sums them together to create something that should resemble "data." We define the data residual as the difference between the input data and the forward-modeled data:

$$\mathbf{r}_d = \mathbf{d} - \mathbf{L}\mathbf{m} \tag{7}$$

Viewed as a standard least-squares inversion problem, minimization of $L_2$ norm of the data residual by solution of the normal equations is underdetermined. Additional regularization terms, defined in later sections, force the problem to be overdetermined.

### Consistency of the Data and the Crosstalk Problem

Figure 3 shows the result of applying the adjoint of equation (6) to a synthetic CMP gather which was constructed by an elastic modeling scheme. Imagine for a moment that the CMP gather consists *only* of primaries and first- and second-order water-bottom multiples. The "NMO for Primaries" panel would contain flattened primaries (signal) and downward-curving first- and second-order multiples (noise). Likewise, the "NMO for multiple 1" and "NMO for multiple 2" panels contain flattened signal and curving noise. Why do I call these components "signal" and "noise"? If each of the three panels contained all signal and no noise, then we could 1) perfectly reconstruct the data from the model by applying equation (6), and 2) be in the enviable position of having a perfect estimate of the primaries.

Unfortunately, the curved events – so-called "crosstalk" – in all three model panels spoil this idealized situation (Claerbout, 1992). Because the crosstalk events map back to actual events in the data, they are difficult to suppress in a least-squares minimization of the data residual [equation (7)]. Nemeth et al. (1999) shows that crosstalk relates directly to non-invertibility of the Hessian ($\mathbf{L}^T\mathbf{L}$), and that data-space or model-space regularization may partially overcome the difficulty. In the following section, I introduce a novel form of model-space regularization which promotes discrimination of signal from crosstalk.

### Regularization of the Least-Squares Problem

Visual inspection of Figure 3 motivates the two forms of regularization utilized in this paper. Find any first-order multiple on the section marked "NMO for Primaries." Notice that the corresponding event on the first- and second-order pseudo-primary panels, originally second-

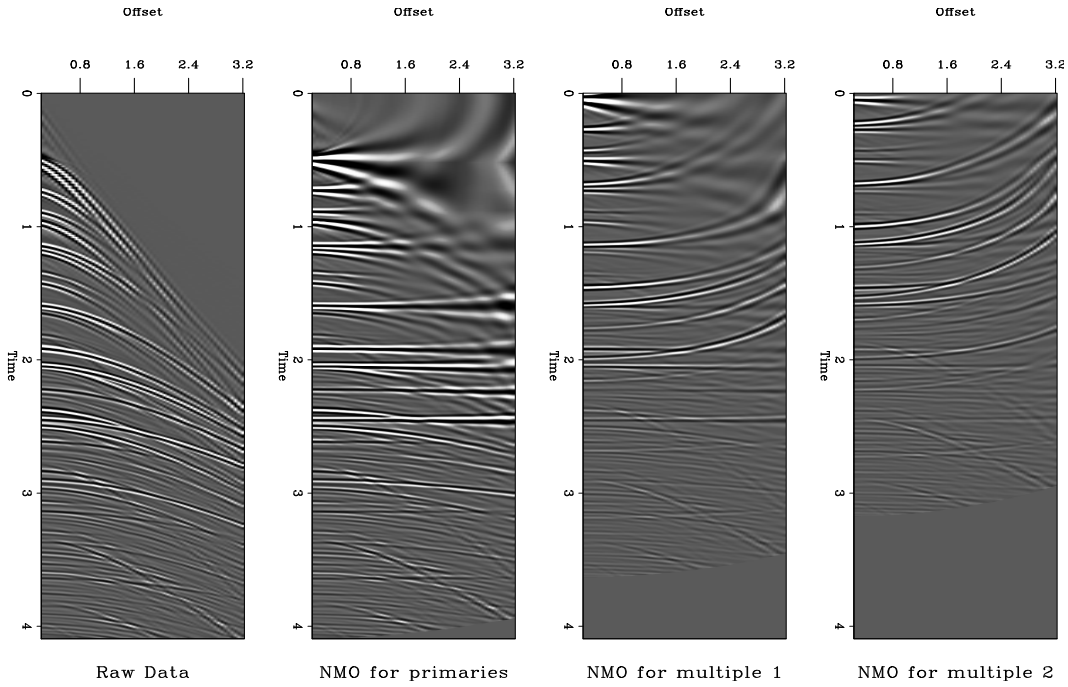



Figure 3: From left to right: Raw synthetic CMP gather; Conventional NMO applied to data; NMO for first-order water-bottom multiple; NMO for second-order water-bottom multiples. morgan1-schem.hask [ER]

and third-order multiples, respectively, all have a different moveout. In fact, the only events which are kinematically consistent across all offsets are the flattened primary and pseudo-primaries. The other events, all crosstalk, are inconsistent between panels. Therefore, the first regularization operator seeks to penalize the difference between the $\mathbf{m}_i$, at fixed $\tau$. To account for the dissimilarity of the AVO of primaries and multiples, this difference is taken at different offsets, as defined in equation (4). Written in the form of a model residual vector, this difference is:

$$\mathbf{r}_m^{[1]}(\tau,x,i) = m_i(\tau,h_p) - m_{i+1}(\tau,x). \tag{8}$$

The third index in equation (9), $i$, ranges from 0 to $n_p-1$, where $n_p$ is the highest order multiple modeled in the inversion [see equation (6)].

The second form of regularization used in this paper is the more obvious of the two: a difference operator along offset. This difference exploits the fact that all non-primaries are not flat after NMO. Again, we can write this difference in the form of a model residual vector:

$$\mathbf{r}_m^{[2]}(\tau,x,i) = m_i(\tau,x) - m_i(\tau,x+\Delta x). \tag{9}$$

The second regularization is applied to all the $\mathbf{m}_i$. A similar approach is used by Prucha et al. (2001) to regularize prestack depth migration in the angle domain.

### Combined Data and Model Residuals

To compute the optimal set of $\mathbf{m}_i$, a quadratic objective function, $Q(\mathbf{m})$, consisting of sum of the weighted norms of a data residual [equation (7)] and of two model residuals [equations (8) and (9)], is minimized via a conjugate gradient scheme:

$$\min Q(\mathbf{m}) \;=\; \|\mathbf{r}_d\|^2 \;+\; \epsilon_m^2 \|\mathbf{r}_m^{[1]}\|^2 \;+\; \epsilon_x^2 \|\mathbf{r}_m^{[2]}\|^2 \tag{10}$$

$\epsilon_m$ and $\epsilon_x$ are scalars which balance the relative weight of the two model residuals with the data residual.

## RESULTS

### Testing the Raw Algorithm

We begin with the results of testing the proposed algorithm on a single synthetic CMP gather, which was shown previously in Figure 3. This gather, generated using Haskell-Thompson elastic modeling, with earth properties drawn from a well log provided with the "Mobil AVO" dataset (see (Lumley et al., 1994) for a description of data details), has traditionally found use at SEP as a multiple suppression benchmark (Lumley et al., 1994; Nichols, 1994; Clapp and Brown, 1999; Guitton, 2000; Clapp and Brown, 2000). The data contain all surface-related and internal multiples, as well as P-to-S-to-P converted waves.

Figure 4 illustrates the application of the proposed algorithm to the so-called "Haskell" data. Comparing the raw data and the estimated primary panels [$\mathbf{m}_0$ in equation (6)], we see that my algorithm does a decent job of suppressing the strongest multiples, especially at far offsets, though some residual multiple energy remains at the near offsets. We expect poorer performance at near offsets; recall that the first regularization operator [equation (8)] penalizes dissimilarity of events across orders of multiple, yet all orders of multiple align at near offsets. Moreover, the second regularization operator [equation (9)] penalizes residual curvature, yet all events in the section, both primaries and the residual multiples, are flat at near offsets.

The difference panel shows little residual primary energy, which illustrates the favorable signal preservation capability of my approach. The bulk of the residual primary energy exists at far offsets and small times, where NMO stretch makes the primaries nonflat, and hence, vulnerable to smoothing across offset by equation (9).

The bottom three panels in Figure 4 show the data residual [equation (7)], and the first two panels of the two model residuals [equations (9) and (8), respectively]. Put simply, the data residual consists of events which are not modeled by equation (6) – hopefully, the multiples only. The model residuals consist roughly of the portions of the model which were removed by the two regularization terms (again, hopefully multiples only): high-wavenumber events and events which are inconsistent from one panel to the next.

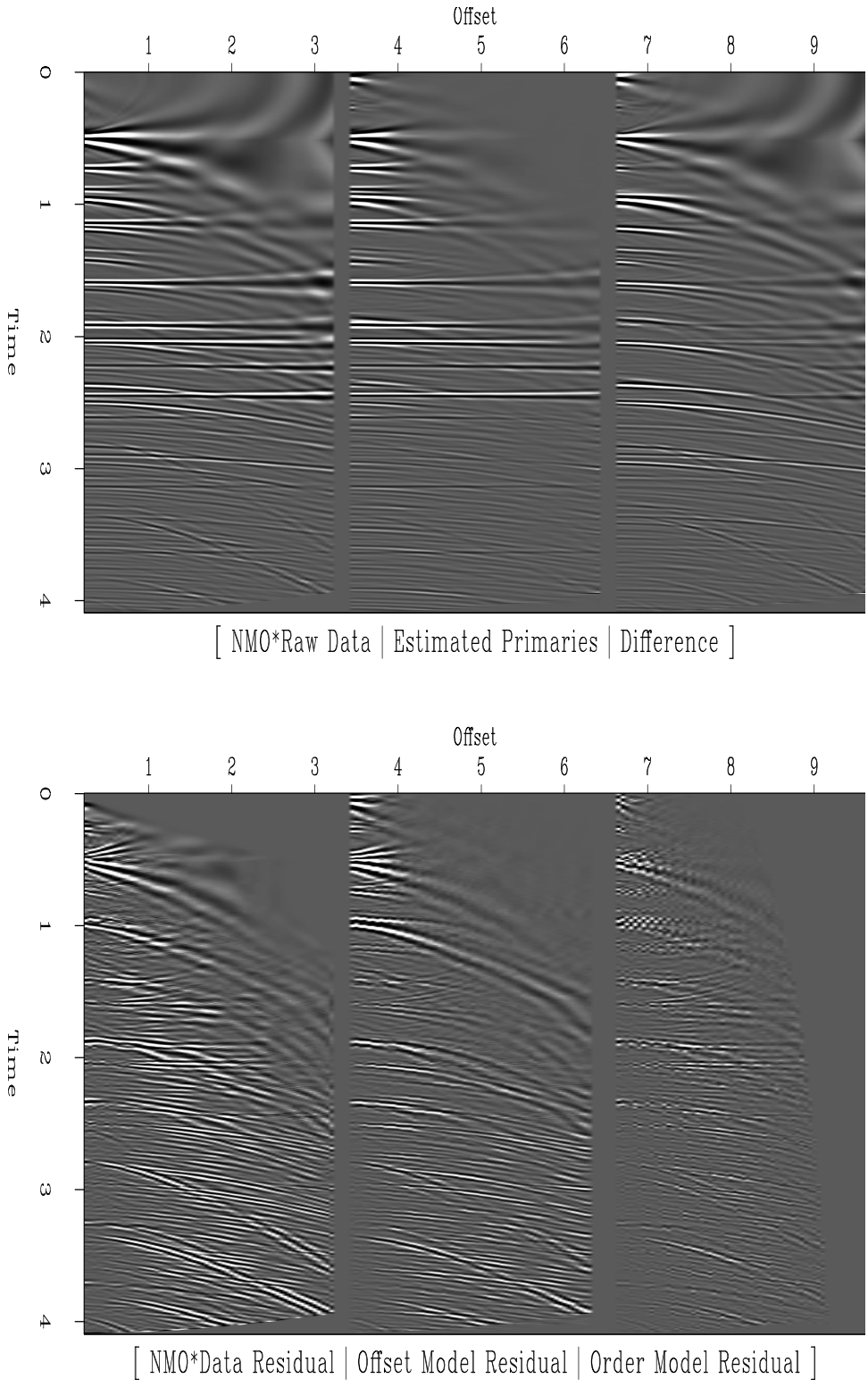



Figure 4: Test of equation (10) on Haskell synthetic CMP gather. Top row, left to right: Raw Haskell data, NMO applied; Estimated primary panel; difference panel. Bottom row, left to right: Data residual; first panel of model residuals, equations (9) and (8), respectively. morgan1-cmps.lsrow.hask [ER]

### Better Understanding the Regularization

Figure 5 illustrates the effect of setting $\epsilon_x = 0$ in equation (10), which removes the influence of the regularization term which roughens the model across offset [equation (9)]. The results are intriguing. Most noticeably, leftover multiple reflections in the "Estimated Primaries" panel appear to be scrambled over offset, while primaries appear mostly intact. Signal-to-noise ratio has increased considerably. The fact that the roughener across offset decorrelates the residual multiples should be further exploited.

Notice that the model residual is zero at long offsets and small times. This is due to the fact that the difference is not taken across the same offsets, to account for the AVO multiples, according to equation (4). When $h_p + h_m > h_{max}$, no difference is taken.

### Devil's Advocate: What do the Multiples Add?

Figure 6 illustrates application of the algorithm without the use of multiples. Only the regularization across offset, equation (9), is in operation. Though we see some suppression of multiples, the results are not nearly as good as those in Figure 4. More insidiously, note the presence of considerable of primary energy in the difference panel. When exploited as a constraint against crosstalk, the multiple reflections add considerable information. My approach integrates this information in a systematic framework.

### A Real Data Example

I test the proposed algorithm on a single CMP gather from the Mobil AVO dataset, described above. The results are shown in Figure 7. Relative to the results seen on the Haskell synthetic, they are fairly poor. On the bright side, notice decent preservation of signal amplitude. The earliest water-bottom multiples are suppressed quite effectively, although the later reverberations are left almost untouched.

The reasons for the less-than-perfect are likely numerous. First, and most important, the multiple reflections quickly become incoherent with an increasing number of bounces. They match well with the primaries only for the strongest reflections. I estimated a relatively small water-bottom reflection coefficient, 0.1, so the multiples are relatively weak in amplitude. I did not perform any preprocessing on the data, and I believe they were donated to SEP as raw gathers. Berlioux and Lumley (1994) applied cable balancing to the Mobil AVO dataset. High-wavenumber, offset-variant amplitude variations along events spoil the ability regularization equation (8) to discriminate against crosstalk.

## DISCUSSION

I presented a new approach for the joint imaging of primary and multiple reflections. My approach goes further than the separate imaging of multiples and primaries. I integrate infor-

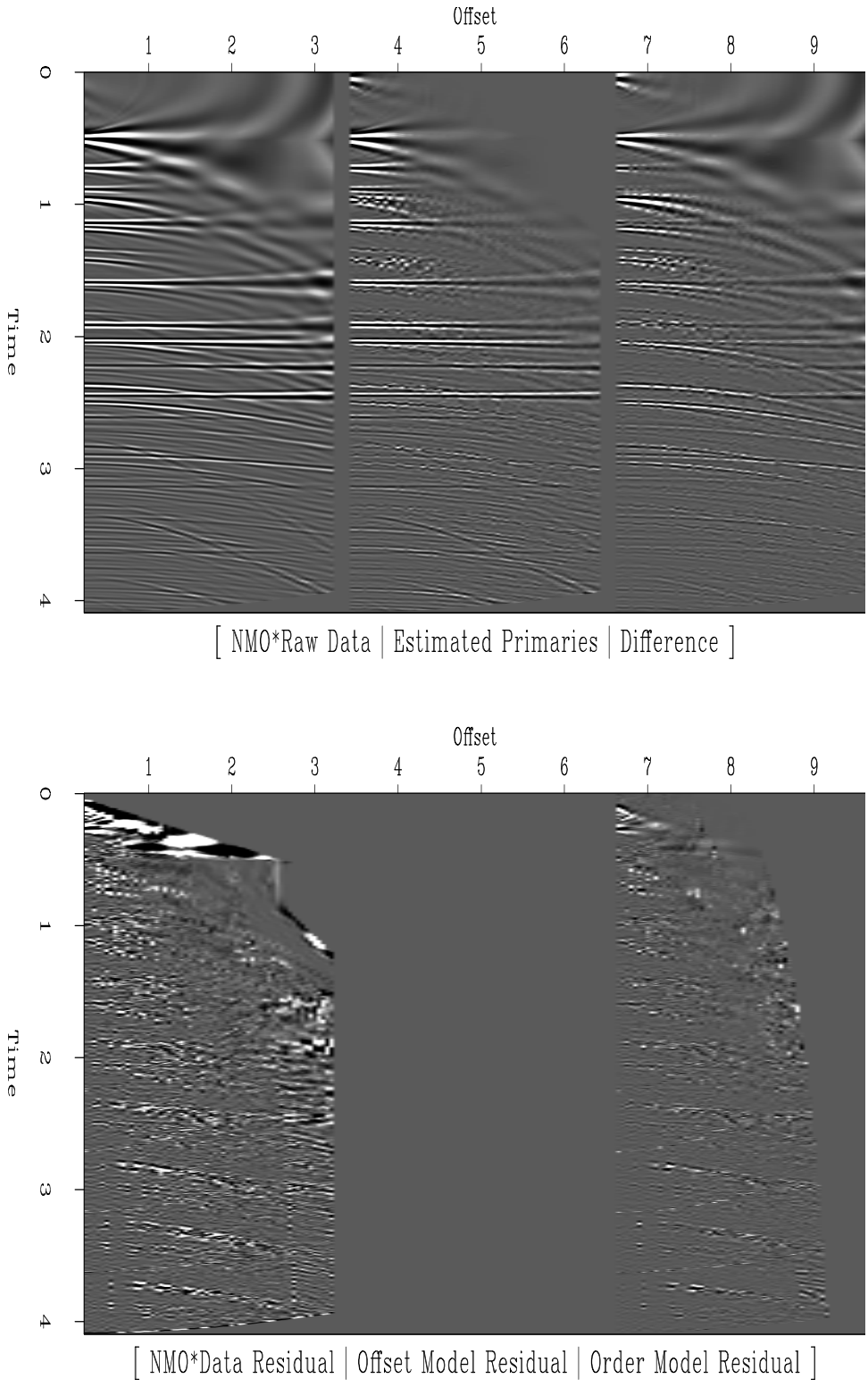



Figure 5: Only the regularization which roughens across orders of pseudo-primary, equation (8), is used in the inversion. Top row, left to right: Raw Haskell data, NMO applied; Estimated primary panel; difference panel. Bottom row, left to right: Data residual; first panel of model residuals, equations (9) and (8), respectively. morgan1-cmps.nograd.hask [ER]

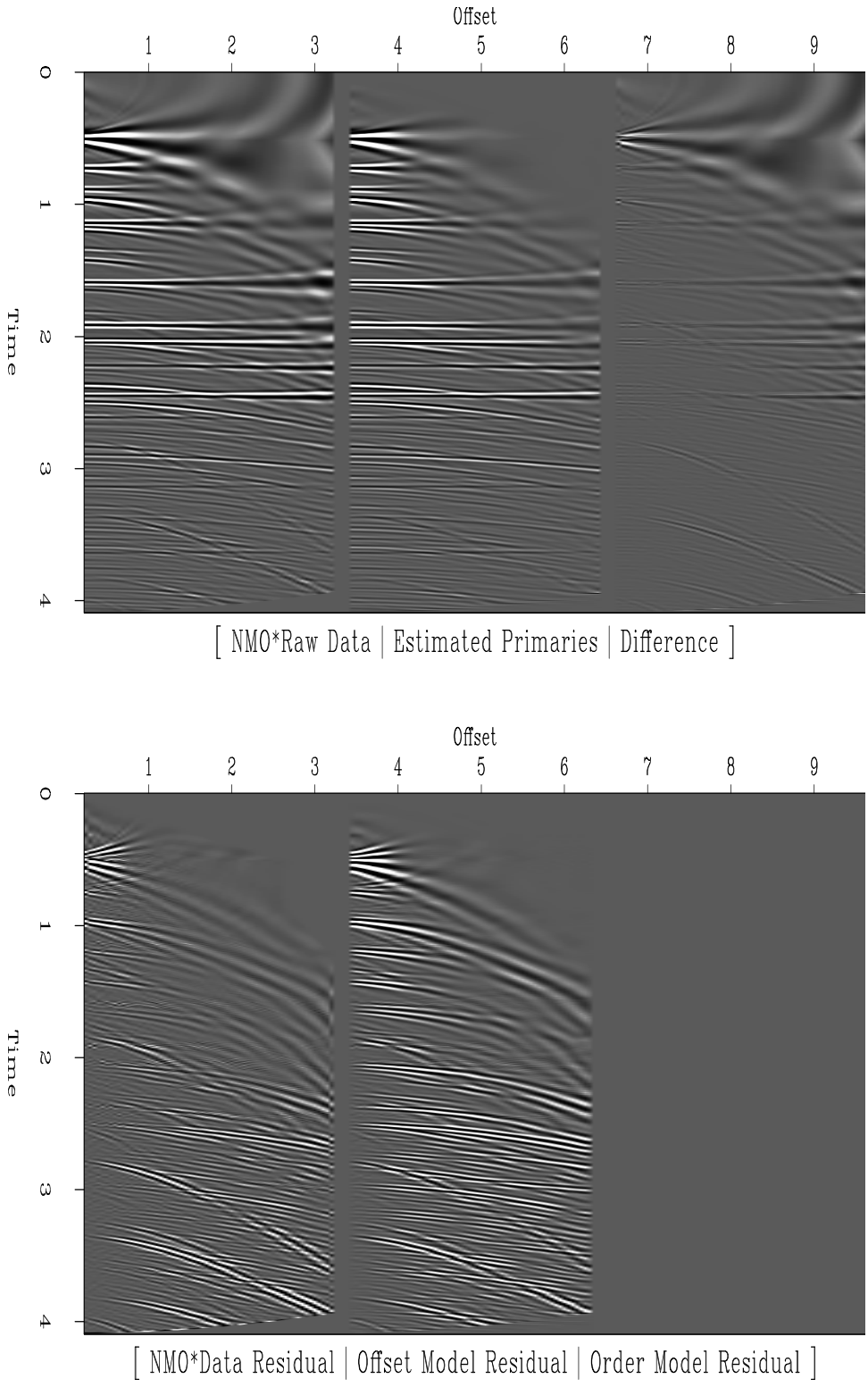



Figure 6: Only the regularization which roughens across offset, equation (9) is used in the inversion. Only one order of pseudo-primary is used, so no information is added by the multiples. Top row, left to right: Raw Haskell data, NMO applied; Estimated primary panel; difference panel. Bottom row, left to right: Data residual; first panel of model residuals, equations (9) and (8), respectively. morgan1-cmps.devils.hask [ER]

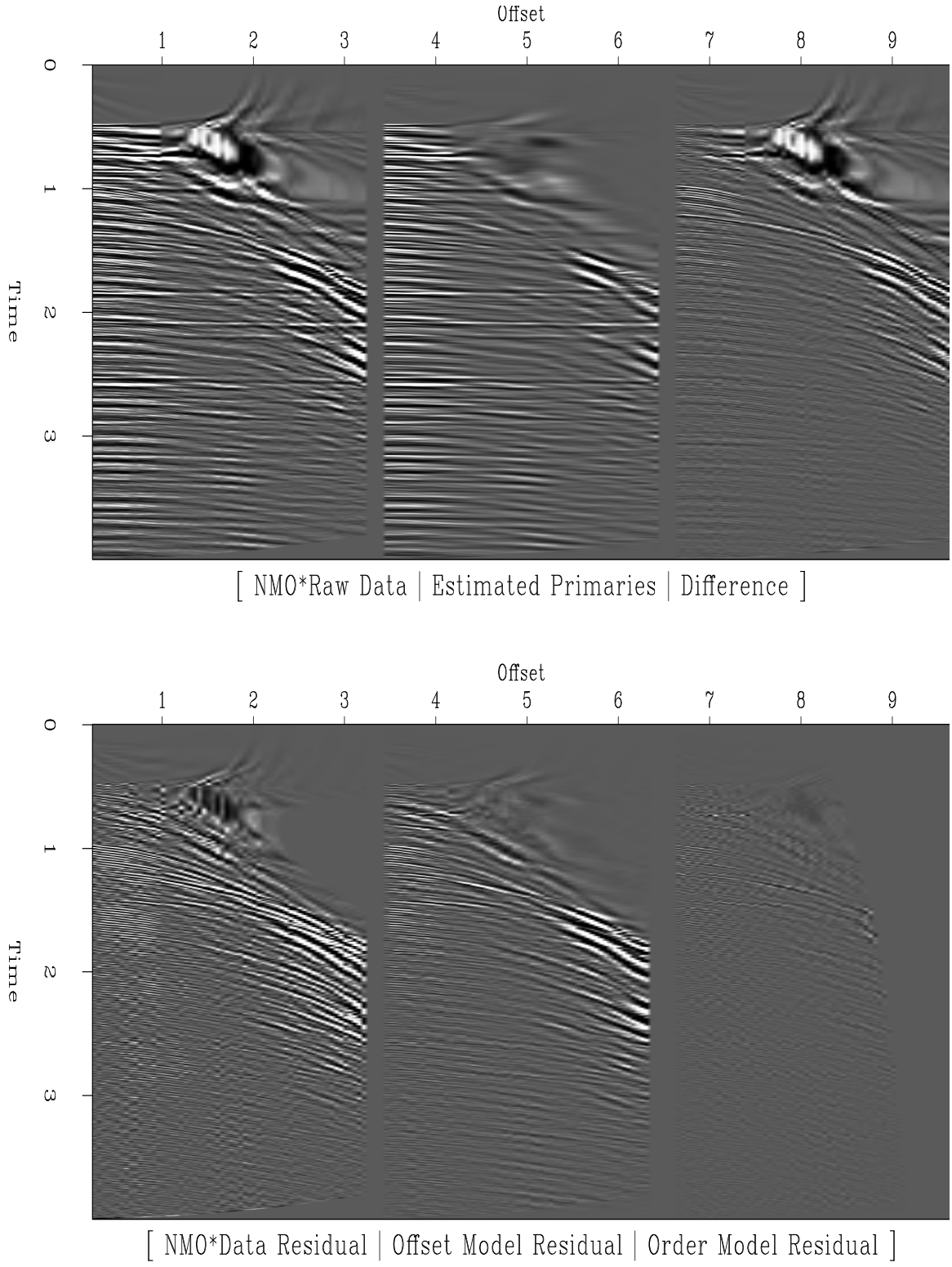



Figure 7: Application of equation (10) to CMP from Mobil AVO data. Top row, left to right: Raw CMP gather, NMO applied; Estimated primary panel; difference panel. morgan1-cmps.lsrow.haskreal [ER]

mation from the multiples and primaries in a least-squares inversion, via a new regularization term which exploits the kinematic similarity of primaries and pseudo-primaries, and the kinematic dissimilarity of crosstalk terms to obtain a noise-free image of the primaries.

The proposed algorithm demonstrates good noise suppression and signal preservation characteristics in the synthetic tests of Figure 4. Comparison of Figures 5 and 6 proves the validity of the new regularization term, equation (8), and more importantly, that the multiples provide valuable information in the inversion.

The results of testing a real data gather were mixed. I believe the single largest problem in this case is poor coherency of the water-bottom multiples. As the water bottom and most shallow reflectors on the Mobil AVO dataset are nearly perfectly flat, geologic complexity is surely not to blame. More likely, the solution(s) to the trouble is(are) more mundane; things like source/cable balancing and spherical divergence. An accurate RMS velocity function is important to success, but errors can be tolerated. Velocity errors lead to curvature in NMO'ed primaries and pseudo-primaries, but as I have dealt here only with water-layer multiples, the real danger, a large phase shift between primaries and pseudo-primaries, is somewhat unlikely.

In all tests, the removal of multiples at near offsets was incomplete. Since the near offsets contribute most to residual multiple energy in the stack, it is of crucial importance to improve performance.

## FUTURE DIRECTIONS

The obvious direction in which to move this project is migration. By using migration, rather than NMO, as the imaging operator, the limiting assumptions of NMO ($v(z)$, flat reflectors) can be abandoned. Furthermore, the limitations of operating in the offset domain can be overcome by moving to the more intuitive angle domain. In some cases, multiples provide better angular coverage over a recorded cable length. Systematic integration of this extra information could prove revolutionary in regions of poor illumination. But the fruits of this transition are not without challenges. Because NMO is a vertical mapping, each CMP can be processed independently, making the memory requirements of the current implementation reasonable, and parallelization quite simple. Correctly handling the transformation of amplitudes between primary and pseudo-primary may prove even more of a challenge. Also, regardless of whether NMO or migration is used, the move to 3-D is never a forgiving one from the computational standpoint.

## ACKNOWLEDGEMENT

Thanks go to ExxonMobil for providing SEP with the Mobil AVO dataset.

## REFERENCES

Berkhout, A. J., and Verschuur, D. J., 1994, Multiple technology: Part 2, migration of multiple reflections: 64th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1497–1500.

Berlioux, A., and Lumley, D., 1994, Amplitude balancing for AVO analysis: SEP–**80**, 349–360.

Claerbout, J. F., 1992, Earth Soundings Analysis: Processing Versus Inversion: Blackwell Scientific Publications.

Clapp, R. G., and Brown, M., 1999, Applying sep's latest tricks to the multiple suppression problem: SEP–**102**, 91–100.

Clapp, R. G., and Brown, M., 2000, $(t-x)$ domain, pattern-based multiple separation: SEP–**103**, 201–210.

Guitton, A., 2000, Prestack multiple attenuation using the hyperbolic Radon transform: SEP–**103**, 181–201.

Guitton, A., 2002, Shot-profile migration of multiple reflections: SEP–**111**, 17–33.

Lumley, D., Nichols, D., and Rekdal, T., 1994, Amplitude-preserved multiple suppression: SEP–**82**, 25–45.

Nemeth, T., Wu, C., and Schuster, G. T., 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, no. 1, 208–221.

Nichols, D., 1994, Velocity-stack inversion using $\mathbf{L_p}$ norms: SEP–**82**, 1–16.

Prucha, M. L., Clapp, R. G., and Biondi, B. L., 2001, Imaging under salt edges: A regularized least-squares inversion scheme: SEP–**108**, 91–104.

Reiter, E. C., Toksoz, M. N., Keho, T. H., and Purdy, G. M., 1991, Imaging with deep-water multiples: Geophysics, **56**, no. 07, 1081–1086.

Yu, J., and Schuster, G., 2001, Crosscorrelogram migration of IVSPWD data: 71st Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 456–459.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 47–59*

# Converted wave dip moveout

*Daniel Rosales*[1]

**ABSTRACT**

Dip moveout (DMO) introduces a dip-dependent correction for a more appropriate transformation of prestack data into zero-offset data. Converted wave DMO has been discussed in the literature by several authors. Log-stretch *f-k* PS-DMO has been recently introduced in the literature. At the moment, however, this PS-DMO operator presents problems of handling amplitudes properly. A PP-DMO in this domain that accurately handle amplitudes already exists. Therefore, it is important to extend such an operator for PS data. This new operator for converted waves is presented in this paper. I show impulse responses and real data results using the new PS-DMO operator.

## INTRODUCTION

The problem of sorting, NMO correction and stacking for PS data has been widely addressed in the past (Tessmer and Behle, 1988; Iverson et al., 1989). The solutions presented in these works apply a lateral shift to the trace midpoints, such that the new trace position corresponds with the lateral position of the conversion point. Usually, this correction does not incorporate the effects of the reflector depth and dip.

In order to incorporate the dip effect, Huub Den Rooijen (1991) achieves the transformation from CMP-sorted data to CCP-sorted data using a dip moveout operator. In a similar way to the PP-DMO, PS-DMO may reduce the problem of reflection point dispersal due to dip without knowledge of the reflector geometry. Most of the existing PS-DMO operators present errors due to truncation of power series and/or second order approximations (Xu et al., 2001). Xu et al. (2001) show a fast converted dip moveout operator in the *f-k* domain which partially alleviates approximation errors.

Harrison (1990) presents the zero-offset mapping equation for the PS-DMO operator, applying an integral-summation approach in order to implement his PS-DMO operator.

Xu et al. (2001) present a log-stretch *f-k* PS-DMO operator. His operator correctly handles the kinematics, but doesn't handle amplitudes properly. Here, I present a review of the PS-DMO operator, implement the operator described by Xu et al. (2001) and extend it to handle the amplitudes properly.

[1]**email:** daniel@sep.stanford.edu

## KINEMATICS OF PS-DMO

The downgoing and upgoing waves travel at different velocities for converted wave data. This difference makes the kinematics more complicated than that of single mode data.

The kinematics of PS-DMO have been widely discussed in the literature. Harrison (1990) is the first one to derive the zero-offset mapping equation for converted waves. He uses an integral-summation approach, similar to Deregowski and Rocca (1981) in order to apply DMO to converted waves data. He also applies the PS zero-offset mapping and the adjoint of his operator in order to obtain the PS Rocca's operator.

I present a review of the kinematics for PS-DMO. The impulse response is produced by taking an impulse on a constant offset section and migrating it to produce ellipses. Each element or point along the ellipses is then diffracted, setting offset to zero, to produce hyperbolas. This operation creates the impulse response that represents the Rocca's DMO+NMO operator (Claerbout, 1999).

Figure 1 shows a comparison between the Rocca's smear operator for single mode P data and for converted mode PS data. It is possible to observe, kinematically, that the Rocca's operator for converted waves is shifted laterally toward the receiver. This is the expected result, since the upgoing wave path is slower than the downgoing wave path.

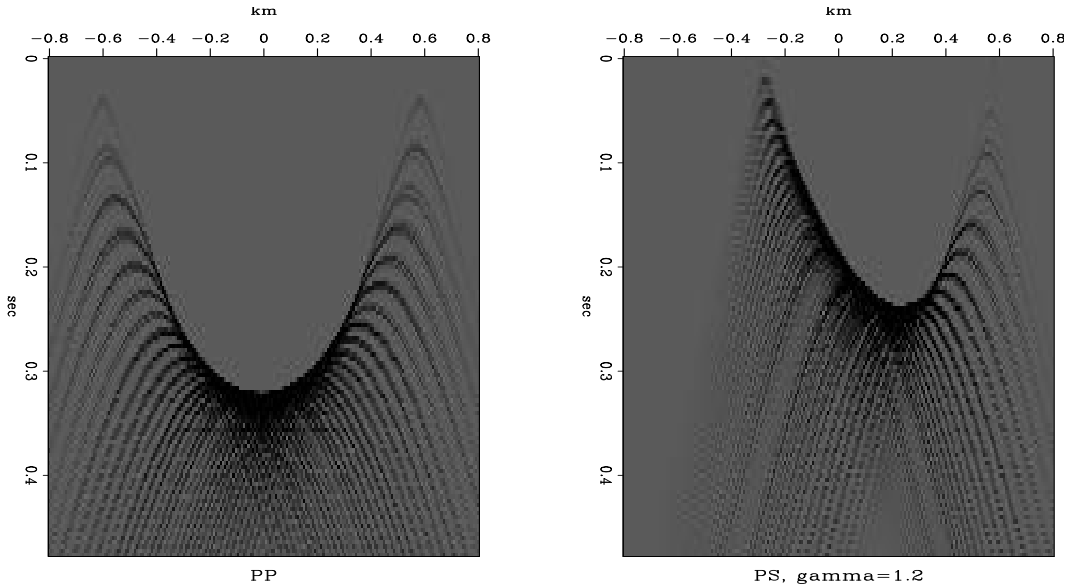



Figure 1: Rocca's operator, for single mode P data (left) and converted mode PS data (right) daniel2-rocca [ER,M]

Jaramillo (1997) shows the amplitude distribution for a PS-DMO operator. This amplitude distribution is also shown in Figure (1). The denser the dots, the higher the amplitude should be.

One of the challenges is to obtain an independent expression for the DMO operator, but

this task has already been solved in the literature. Here I present a review of this process, since the result is used by Rosales and Biondi (2002) and later in this paper.

**PS-DMO smile derivation**

The trajectory equation of conversion points is

$$t = \frac{\sqrt{z^2+(h+x)^2}}{v_p} + \frac{\sqrt{z^2+(h-x)^2}}{v_s}. \tag{1}$$

By introducing the ratio, $\gamma = \frac{v_p}{v_s}$, and by following Huub Den Rooijen's (1991) derivations, as well as work by Xu et al. (2001), we have the following standard form of elliptical equation:

$$z^2 + \left(1 - \frac{4\gamma h^2}{v_p^2 t^2}\right)(x+D)^2 = \frac{v_p t_n^2}{(1+\gamma)^2}, \tag{2}$$

where $D$ is the element responsible for transforming the data from CMP coordinates to CCP coordinates. $D$ comes from solving equation (1) for $z^2$. The difference among the existing PS-DMO operators is in their definition of $D$. In this paper we use the definition presented by Xu et al. (2001):

$$D = \left[1 + \frac{4\gamma h^2}{v_p^2 t_n^2 + 2\gamma(1-\gamma)h^2}\right]\frac{1-\gamma}{1+\gamma}h, \tag{3}$$

where $t_n$ is the NMO-corrected time.

The spatial shift distance needed to convert CMP gathers to common-reflection point gathers is

$$b = x + z\frac{dz}{dx}. \tag{4}$$

The one-way normal-incidence distance $R$ is given by

$$R^2 = z^2 + \left(z\frac{dz}{dx}\right)^2. \tag{5}$$

Using equations (2) and (4) we eliminate the $z$ and $x$ dependencies. This result is substituted in the relationship for the one-way normal-incidence distance $R$ [equation (5)] in order to get the equation:

$$R^2 + (b+D)^2\left(\frac{v_p^2 t^2}{4\gamma h^2} - 1\right) = \frac{v_p^2 t_n^2}{(1+\gamma)^2}. \tag{6}$$

The two-way normal-incidence time is

$$t_0 = \frac{R}{v_p} + \frac{R}{v_s} = \frac{1+\gamma}{v_p}R, \tag{7}$$

and the traveltime equation, in terms of the normal moveout time, is

$$t^2 = t_n^2 + \frac{4\gamma h^2}{v_p^2}. \tag{8}$$

Combining equations (6), (7) and (8) produces the PS-DMO smile equation:

$$\frac{t_0^2}{t_n^2} + \frac{y^2}{H^2} = 1, \tag{9}$$

where

$$\begin{aligned} y &= x + D \\ H &= \frac{2\sqrt{\gamma}}{1+\gamma} h. \end{aligned}$$

## PS-DMO IN THE FREQUENCY-WAVENUMBER LOG-STRETCH DOMAIN

The previous section shows a kinematic PS-DMO operator. Like PP-DMO, kinematic PS-DMO can not properly handle phase and amplitude (Xu et al., 2001). Xu et al. (2001); Alfaraj (1992) discuss a PS-DMO operator on the *f-k* domain by using a truncation of power series for the moveout of reflections from dipping reflectors in a constant velocity media.

Zhou et al. (1996) discuss that the PP-DMO operator in the *f-k* domain is computationally expensive because the operator is temporarily non-stationary. He uses the idea of Bolondi et al. (1982) to express a more accurate PP-DMO operator by a logarithmic time stretching.

Xu et al. (2001) exploit the idea of computational efficiency of the logarithmic time stretching for the PS-DMO operator, however their formulation is not clearly expressed. Therefore, I reformulate their work using the PS-DMO smile presented in the previous section and following a procedure similar to Hale (1984) and Zhou et al. (1996). This operation accounts for constant velocity case.

Starting from the PS wave DMO smile,

$$\frac{t_0^2}{t_n^2} + \frac{y^2}{H^2} = 1. \tag{10}$$

By following Hale's (1984) assumption that $p_0(t_0, x, H) = p_n(t_n, x, H)$, we obtain the 2D PS-DMO operator in the *f-k* domain:

$$P_0(\omega, k, H) = \int\int p_0(t_0, y, H) e^{i(\omega t_0 - ky)} dt_0 dy. \tag{11}$$

Equation (10) implies a change of variable from $t_0$ to $t_n$. Therefore, equation (11) becomes

$$P_0(\omega, k, H) = \int\int A^{-1} p_0(t_n, x, H) e^{i\omega A t_n} e^{-ik(x+D)} dt_n dx \tag{12}$$

where

$$A = \frac{dt_0}{dt_n} = \sqrt{1 + \frac{H^2 k^2}{t_n^2 \omega^2}}. \tag{13}$$

Equation (12) is the base for PS-DMO in the *f-k* domain. By using a time log-stretch transform pair,

$$\begin{aligned} \tau &= \log \frac{t}{t_c}, \\ t &= t_c e^{\tau}. \end{aligned} \tag{14}$$

The DMO operator in the *f-k* log-stretch domain becomes

$$P_0(\Omega,k,h) = P_n(\Omega,k,h) e^{ikD} F(\Omega,k,h), \tag{15}$$

where

$$F(\Omega,k,h) = e^{\frac{i}{2}\Omega \log \frac{1}{2}\left(\sqrt{\left(\frac{2kH}{\Omega}\right)^2+1}+1\right)}. \tag{16}$$

When either $kh$ or $\Omega$ gets the values of 0, the filter goes to 0 as well.

The previous expression is equivalent to the one presented by Xu et al. (2001). Note that equation (15) is based on the assumption that $p_0(t_0,x,H) = p_n(t_n,x,H)$. This doesn't include changes in midpoint position and/or common reflection point position. This allows for a correct kinematic operator but one with a poor amplitude distribution along the reflectors.

Zhou et al. (1996) refers to this problem in the log-stretch frequency wavenumber domain by reformulating Black's (1993) *f-k* DMO operator. This operator is based on the assumption that $p_0(t_0,x_0,H) = p_n(t_n,x_n,H)$. The midpoint location also changes, leading to a more accurate distribution of amplitudes. After implementing the PS-DMO operator Xu et al. (2001), extending this operator is a feasible task. Following the derivation used by Zhou et al. (1996), I state a more accurate PS-DMO operator in the log-stretch frequency wavenumber domain. This new operator differs from the previous one in the filter $F(\Omega,k,h)$. The new filter is

$$F(\Omega,k,h) = e^{\frac{i}{2}\Omega\left\{\sqrt{1+\left(\frac{2kH}{\Omega}\right)^2}-1-\log\frac{1}{2}\left[\sqrt{\left(\frac{2kH}{\Omega}\right)^2+1}+1\right]\right\}}. \tag{17}$$

This filter reduces to 0 if $kh = 0$ or $kH$ if $\Omega = 0$.

Note that for a value of $\gamma = 1$, the filter reduces to the known expression for P waves data (Zhou et al., 1996).

Figure 2 shows a series of impulse responses for this operator in the frequency domain.

Figure 3 shows the same series of impulse responses as Figure 2 but with the new filter [equation (17)]. Both operators create the same kinematic response. However, Figure 3 shows that the filter in equation (17) gives a more accurate amplitude distribution along the impulse response.

We can trust the PS results since the PP impulse response, obtained with the filter in equation (17) and $\gamma = 1$, is the same as that obtained by Zhou et al. (1996). Moreover, the amplitude distribution follows Jaramillo's (1997) result.

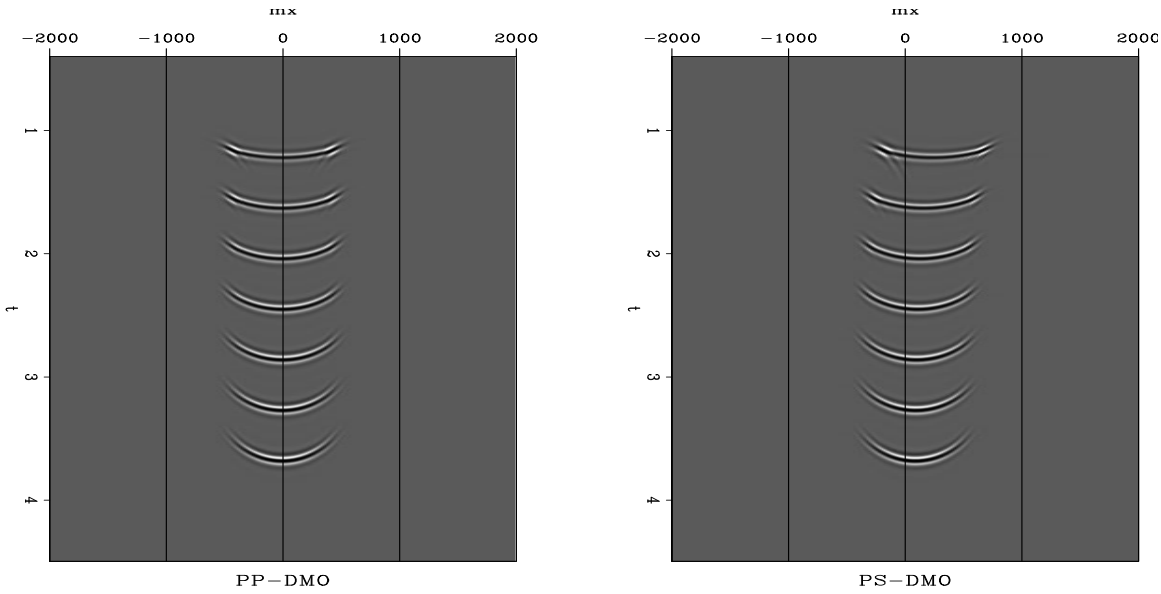



Figure 2: Impulse responses for the DMO operator, PP case (left), PS case (right) daniel2-imps [ER,M]

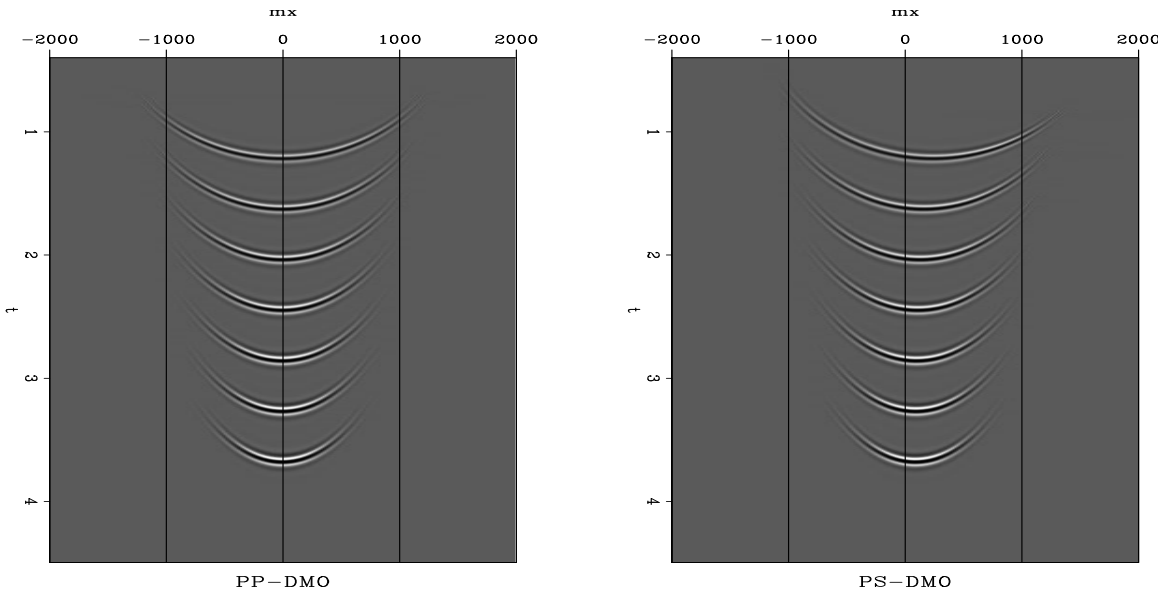



Figure 3: Impulse responses for the new PS-DMO operator, PP case (left), PS case (right) daniel2-imps-new [ER,M]

## REAL DATA

In order to further test our operator, we apply it to a real data line. The Alba Oil Field is located in the UK North Sea and elongates along a NW-SE axis. The oil reservoir is 9km long, 1.5km wide, and up to 90m thick at a depth of 1,900m sub sea (Newton and Flanagan, 1993; MacLeod et al., 1999).

The producing reservoir consists of unconsolidated high-porosity turbiditic sandstone of Eocene age. The oil-bearing reservoir sand and the overlying shale have a very low P-wave acoustic impedance contrast but make a significant S-wave velocity contrast. Hence, the reservoir delineation using only normal PP-seismic data is very difficult. Therefore, the ocean bottom PS-data was acquired (MacLeod et al., 1999).

Even though the Alba field is a 3D data set, we select one line by choosing a source line that overlaps a receiver line. There are several gaps in our selected data since we are taking a 2D line from a 3D data set.

Figure 4 shows a CMP gather from the PZ-component and the PS-component of the data. Note the gaps and irregularities in the data.

Figure 5 shows a comparison between the NMO stack and the DMO stack of the PP line. The DMO was performed by the same algorithm described here, specifying $\gamma = 1$.

Figure 6 shows the PS result. It compares the PS-NMO CCP binning stack and the PS-DMO stack. The PS-DMO stack was obtained using the filter described in equation (16). Observe that, even though it is a flat events area, those small dips are better defined after doing PS-DMO.

Figure 7 shows the PS-DMO result using the filter described in equation (17), which better handles the amplitudes. Note that some strong dip energy appears. The rest of the section remains the same. Therefore, the filter in equation (17) produces more accurate results.

## DISCUSSION

The PS-DMO operator that we finally used is the operator in the *f-k* log-stretch domain. Since this operator is stationary in the time log-stretch domain, the use of FFT in both directions is possible. This makes the PS-DMO operator in the log-stretch frequency wavenumber domain a fast operator that yields accurate results.

This new operator is not only accurate kinematicly but also dynamicly becuase it includes CCP lateral position changes. This is a new result at the time this paper was written.

The operator is valid for constant velocity, however it is necessary to use the work done by other authors (Alfaraj, 1992; Hale, 1984) to incorporate depth variant velocity and depth variant $\gamma$.

The operator used here is safe, fast and valid for both PP and PS propagation modes. It

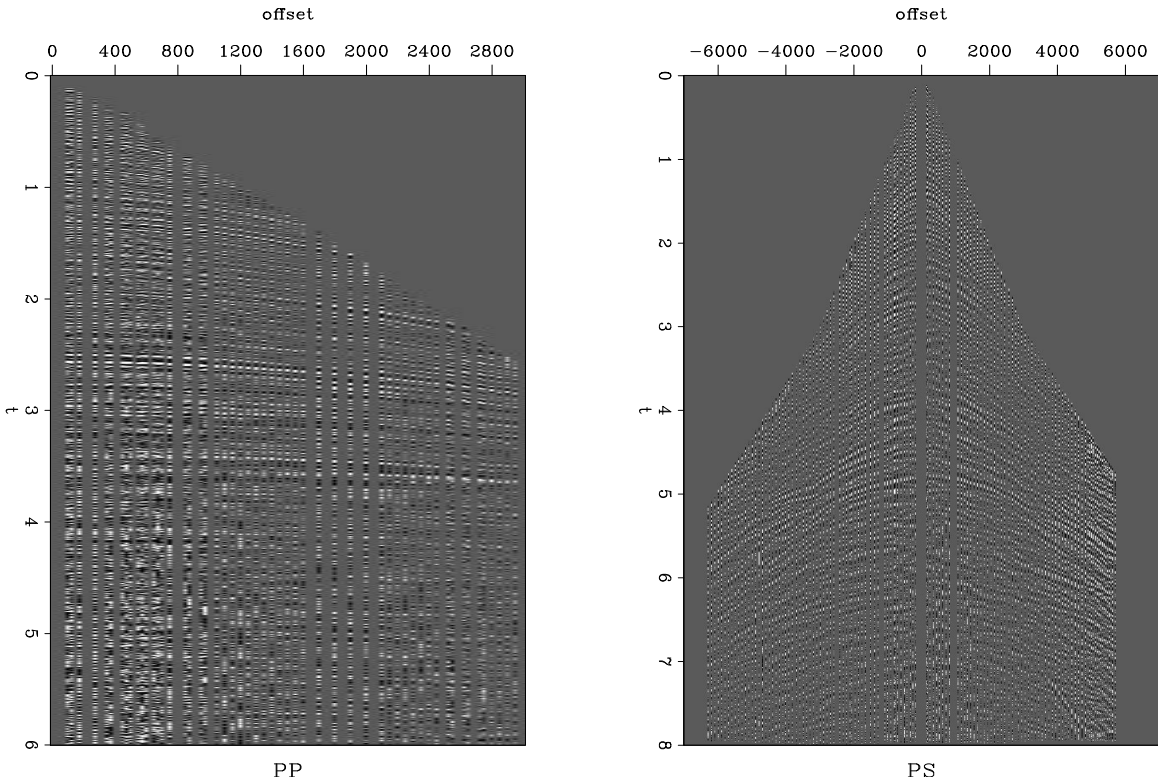



Figure 4: CMPS daniel2-cmps [ER]

is only necessary to specify a value of $\gamma = 1$ to obtain PP-DMO and an appropriate value of $\gamma \neq 1$ and/or $\gamma > 1$ to obtain PS-DMO.

## ACKNOWLEDGMENTS

I want to thank ChevronTexaco for providing the data set.

## REFERENCES

Alfaraj, M., 1992, Transformation to zero offset for mode converted waves by Fourier transform: 62nd Annual Internat. Mtg., Society Of Exploration Geophysicists, Expanded Abstracts, 974–978.

Black, J. L., Schleicher, K. L., and Zhang, L., 1993, True-amplitude imaging and dip moveout: Geophysics, **58**, 47–56.

Bolondi, G., Loinger, E., and Rocca, F., 1982, Offset continuation of seismic sections: Geophys. Prosp., **30**, 813–828.

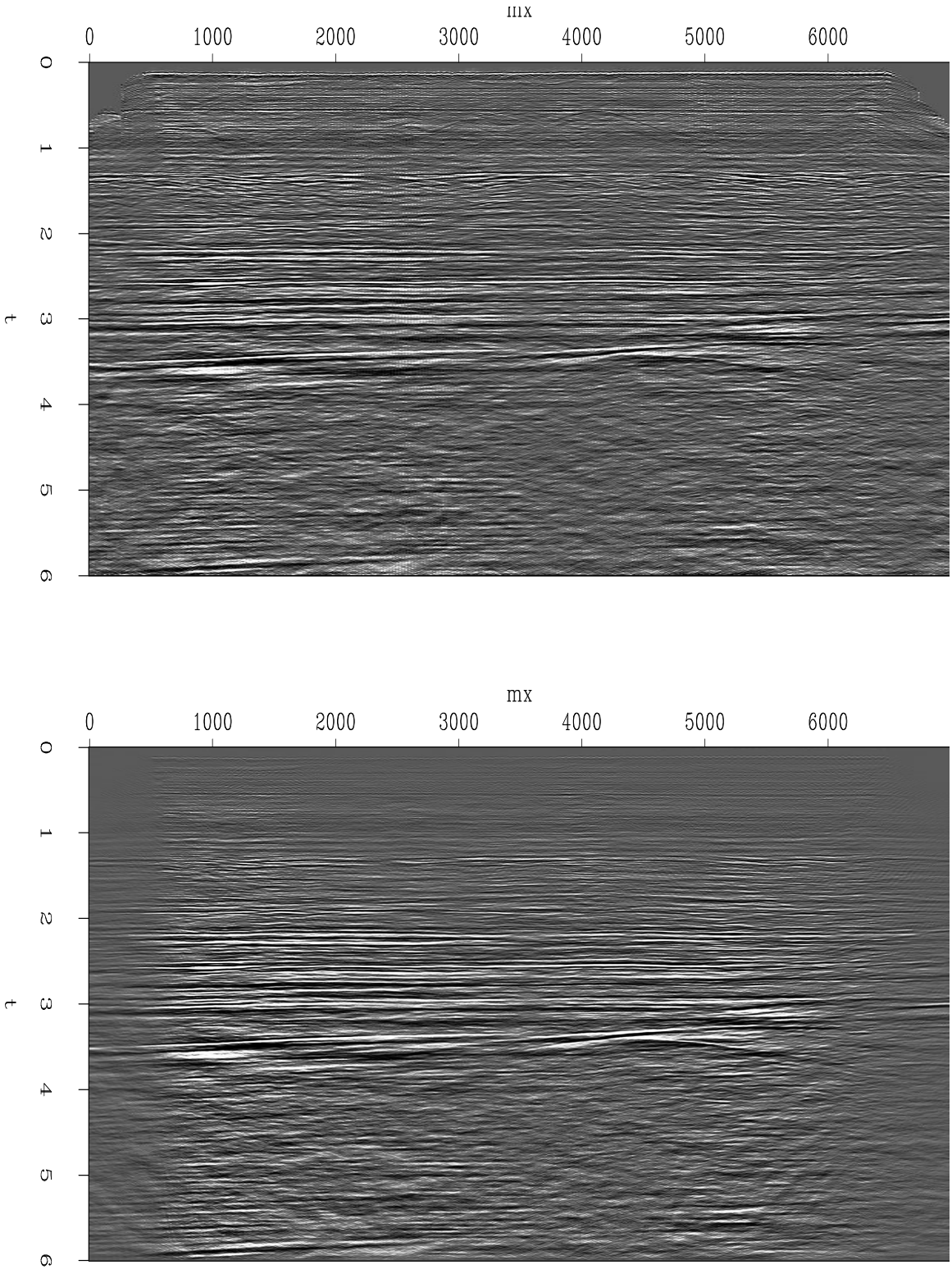



Figure 5: PP results, NMO-stack (top), DMO-stack (bottom) daniel2-pp [CR,M]

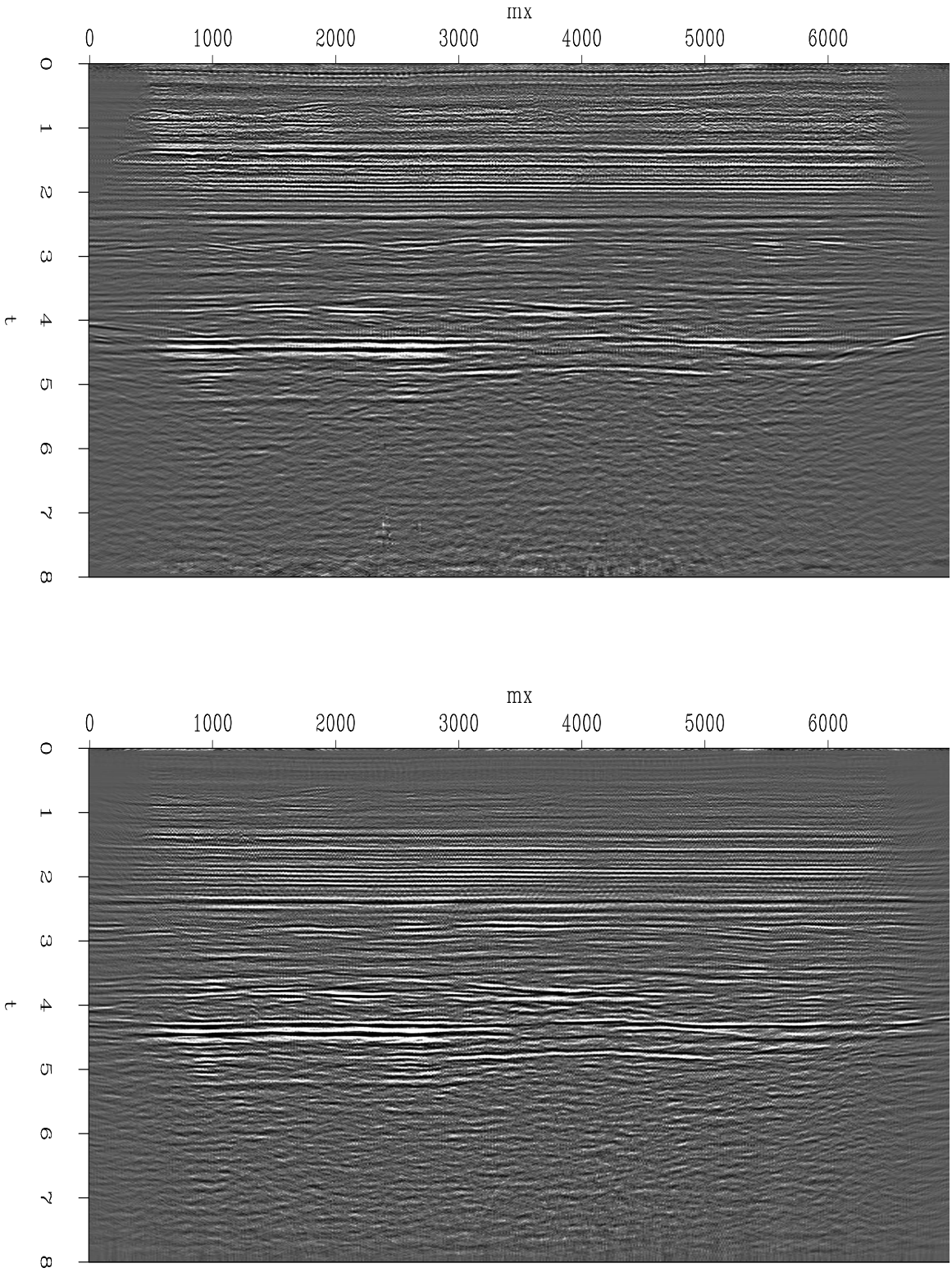



Figure 6: PS results, PS-NMO CCP binning stack (top), PS-DMO stack (bottom) daniel2-psone [CR,M]

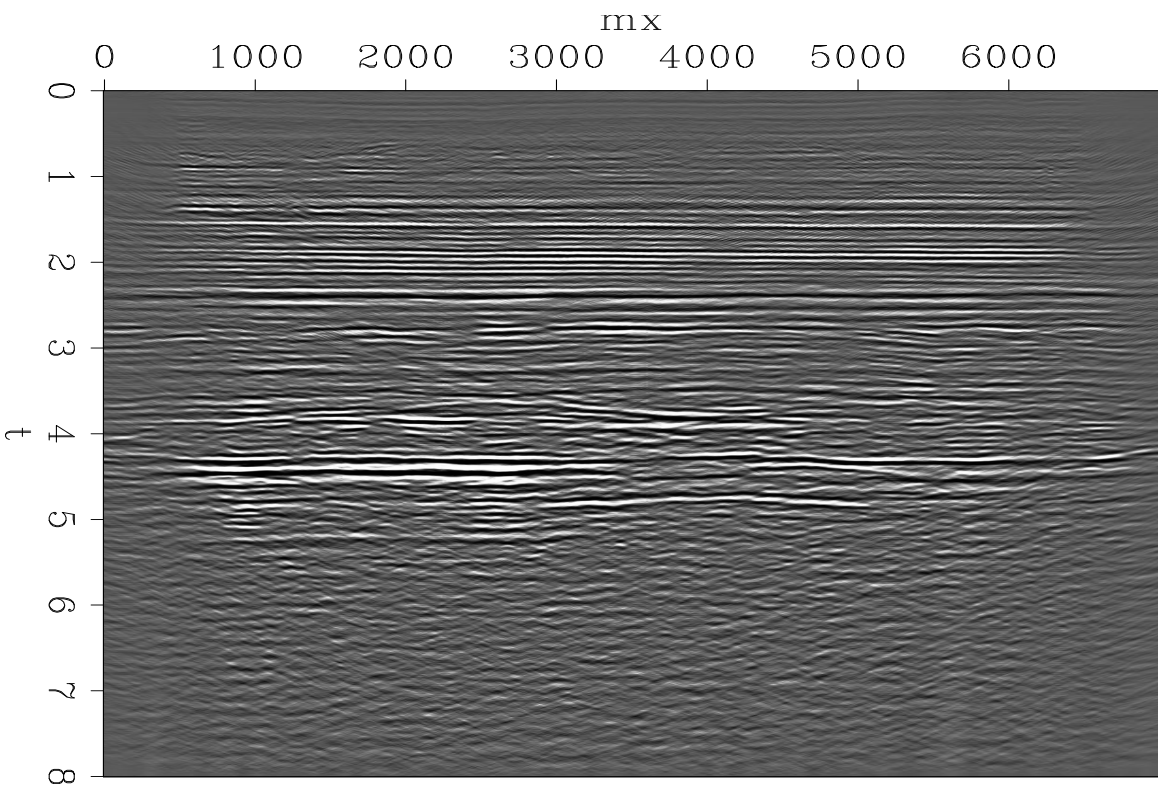



Figure 7: PS-DMO considering amplitudes daniel2-pstwo [CR,M]

Claerbout, J. F., 1999, Basic earth imaging: Class notes, http://sepwww.stanford.edu/sep/prof/index.html.

Deregowski, S. M., and Rocca, F., 1981, Geometrical optics and wave theory of constant offset sections in layered media: Geophys. Prosp., **29**, 384–406.

Hale, D., 1984, Dip moveout by Fourier transform: Geophysics, **49**, 741–757.

Harrison, M., 1990, Converted wave DMO: 60th Annual Internat. Mtg., Society Of Exploration Geophysicists, Expanded Abstracts, 1370–1373.

Huub Den Rooijen, P. G. M., 1991, Stacking of P-SV seismic reflection data using dip moveout: Geophysical Prospecting, **39**, no. 4, 585–598.

Iverson, W. P., Fahmy, B. A., and Smithson, S. B., 1989, VpVs from mode-converted P-SV reflections: Geophysics, **54**, no. 7, 843–852.

Jaramillo, H. H., 1997, P-S converted-wave DMO in depth-variable velocity: 67th Annual Internat. Mtg., Society Of Exploration Geophysicists, Expanded Abstracts, 1597–1594.

MacLeod, M. K., Hanson, R. A., Bell, C. R., and McHugo, S., 1999, The Alba Field ocean bottom cable seismic survey: Impact on development: The Leading Edge, **19**, no. 11, 1306–1312.

Newton, S., and Flanagan, K., 1993, The Alba field: evolution of the depositional model, in Parker, J., Petroleum Geology of Northwest Europe: Proceedings of the 4th conference, Geological Society, 161–171.

Rosales, D., and Biondi, B., 2002, Converted wave azimuth moveout: SEP–**111**, 59–71.

Tessmer, G., and Behle, A., 1988, Common reflection point data-stacking technique for converted waves: Geophysical Prospecting, **36**, no. 5, 671–688.

Xu, S., Jin, S., Ma, Z., and Geng, J., 2001, A fast converted wave dip moveout in f-k domain: 71st Annual Internat. Mtg., Society Of Exploration Geophysicists, Expanded Abstract, 1851–1854.

Zhou, B., Mason, I. M., and Greenhalgh, S. A., 1996, An accurate formulation of log-stretch dip moveout in the frequency-wavenumber domain: Geophysics, **61**, no. 3, 17–23.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 59–70*

# Converted wave azimuth moveout

*Daniel Rosales and Biondo Biondi*[1]

**ABSTRACT**

Accurate prestack partial migration operators are important in seismic exploration. The development of different technologies, like the use of PS converted wave data, suggests the extension of applications of already successful operators and techniques for PP data. Azimuth moveout (AMO) is a partial migration operator that transforms prestack data into equivalent data with arbitrary offset and azimuth. We introduce a new, more accurate prestack partial migration operator for converted wave data. This operator has promising future applications in the regularization of ocean bottom seismic data.

## INTRODUCTION

Stacking is an important process to the seismic exploration industry. It is an effective way to reduce the size of data sets and to enhance reflections while attenuating noise. However, the validity of stacking multiple-coverage data is questionable in the case of PS converted wave data because, even for a horizontal reflection in a constant velocity media, raypaths in a CMP gather strike different reflection points.

Prestack partial migration operators are useful tools in reducing the size of seismic data. Dip moveout (DMO) is the most common prestack partial migration operator. Rosales (2002) comments on a series of DMO operators for PS data. The operators differ in numerical approximations of the moveout equation, processing domain and implementation domain. He also introduces a more accurate PS-DMO operator in the log-stretch *f-k* domain that gives an appropriate amplitude distribution.

Biondi et al. (1998) introduce a more general prestack partial migration operator called Azimuth Moveout (AMO). AMO has the advantage of transforming prestack data into equivalent data with arbitrary offset and azimuth, moving events across midpoints according to their dip. Several advantages have been described for the AMO operator. Among them are: 1) partial stacking of prestack data, in order to create regularly sampled common offset-azimuth cubes (Chemingui and Biondi, 1997; Chemingui, 1999; Biondi, 2000) and 2) data regularization of irregular sampled data which preserves amplitudes (Biondi and Vlad, 2001).

This work presents the equivalent of the PP-AMO operator for converted wave data. We explain the geometrical interpretation of our PS-AMO operator, in which the concept of CCP transformation is important since it is the base for event movement according to its dip. Our

[1]**email:** daniel@sep.stanford.edu,biondo@sep.stanford.edu

PS-AMO operator is a cascade operation of PS-DMO and inverse PS-DMO. We exploit the knowledge of the fast and accurate PP-AMO in the log-stretch frequency-wavenumber domain (Vlad and Biondi, 2001) by selecting the PS-DMO operator in the log-stretch frequency-wavenumber domain introduced by Xu et al. (2001), reformulated and improved by Rosales (2002).

The PS-AMO operator has a significant future application, the regularization of ocean bottom seismic (OBS) data. The presence of already existing platforms produces holes in the data. This information can be safely regularized with an appropriate operator, in this case a PS-AMO operator.

## PS-AMO

Azimuth moveout is a prestack partial migration operator that transforms 3D prestack data with a given offset and azimuth into equivalent data with a different offset and azimuth.

PP-AMO is not a single trace to trace transformation. It is a partial migration operator that moves events across **midpoints** according to their dip. Due to the nature of PS-data, where multiple coverage is obtained through common conversion point gathers (CCP), the PS-AMO operator moves events across **common conversion points** according to their geological dip.

Theoretically, the cascade of any imaging operator with its corresponding forward-modeling operator generates an AMO operator (Biondi, 2000). Our PS-AMO operator is a cascade operation of PS-DMO and inverse PS-DMO.

The 2D PS-DMO smile (Harrison, 1990; Xu et al., 2001; Rosales, 2002) extends to 3D by replacing the offset and midpoint coordinates for the offset and midpoint vectors, respectively. The factor $D$, responsible for the CMP to CCP transformation, also transforms to a vector quantity. The PS-DMO smile in 3D takes the form of:

$$\frac{t_0^2}{t_n^2} + \frac{\|\vec{y}\|^2}{\|\vec{H}\|^2} = 1, \tag{1}$$

where,

$$\|\vec{y}\|^2 = \|\vec{x} + \vec{D}\|^2, \tag{2}$$

$$\vec{H} = \frac{2\sqrt{\gamma}}{1+\gamma}\vec{h}, \tag{3}$$

and

$$\vec{D} = \left[1 + \frac{4\gamma\|\vec{h}\|^2}{v_p^2 t_n^2 + 2\gamma(1-\gamma)\|\vec{h}\|^2}\right]\frac{1-\gamma}{1+\gamma}\vec{h}$$

Here, $\vec{x}$ is the midpoint position vector, $\vec{h}$ is the offset vector and $\gamma = \frac{v_p}{v_s}$ ratio. The PS-AMO operator, a cascade operator of PS-DMO and inverse PS-DMO, takes the form of:

$$t_2^2 = t_1^2 \frac{\vec{H}_{02}^2}{\vec{H}_{10}^2}\left\{\frac{\vec{H}_{10}^2 \sin^2(\theta_1-\theta_2) - \vec{x}^2\sin^2(\theta_2-\Delta\phi) - \mathbf{D}_1}{\vec{H}_{02}^2 \sin^2(\theta_1-\theta_2) - \vec{x}^2\sin^2(\theta_1-\Delta\phi) - \mathbf{D}_2}\right\}, \tag{4}$$

where

$$\begin{aligned}\mathbf{D}_1 &= \vec{D}_{10}^2 \sin^2(\theta_1-\theta_2) + 2\vec{x}\,\vec{D}_{10}\sin(\theta_2-\Delta\phi)\sin(\theta_1-\theta_2)\cos\lambda,\\ \mathbf{D}_2 &= \vec{D}_{02}^2 \sin^2(\theta_1-\theta_2) + 2\vec{x}\,\vec{D}_{0i}\sin(\theta_1-\Delta\phi)\sin(\theta_1-\theta_2)\cos\lambda.\end{aligned}$$

This operator reduces to the traditional expression of PP-AMO (Biondi et al., 1998) for $\vec{D}=\vec{0}$ (i.e. $\gamma=1$).

Although the PP-AMO operator is velocity independent, this independence doesn't propagate for the PS-AMO operator. The PS-AMO operator depends on the P velocity and the $\frac{v_p}{v_s}$ ratio. We assume that the velocity of the new trace position is the same as in the previous position.

**Geometrical interpretation of PS-AMO**

A trace with input offset vector $\vec{h}_1$ and midpoint position $\vec{x}$ is first transformed to its corresponding CCP position and zero offset. By defining the new offset and azimuth position and by applying inverse PS-DMO, we transform the data to a new CCP position and its corresponding CMP position.

Here, we follow the same procedure as Fomel and Biondi (1995); Biondi et al. (1998) for the derivation of the PS-AMO operator.

First, we refer to equations (1) and (2) in order to understand the relationship between CMP and CCP for the 3D case. We rewrite equation (2) as

$$\|\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{D}\|^2 + 2\|\vec{x}\|\,\|\vec{D}\|\cos\lambda,$$

where $\lambda$ is the angle between the midpoint vector ($\vec{x}$) and the transformation vector ($\vec{D}$).

We can then rewrite equation (1) as

$$\frac{t_0^2}{t_n^2} + \frac{\|\vec{x}\|^2}{\|\vec{H}\|^2} + \frac{\|\vec{D}\|^2 + 2\|\vec{x}\|\,\|\vec{D}\|\cos\lambda}{\|\vec{H}\|^2} = 1. \tag{5}$$

$\vec{D}$ is an extension of $\vec{h}$ and lies in the CCP space. Figure 1 shows both $\vec{h}$ and $\vec{D}$ in the same plane. Since the vectors are parallel, the angle between $\vec{x}$ and $\vec{D}$ is the same as the angle between $\vec{x}$ and $\vec{h}$. If the coordinate system is aligned with the midpoint coordinates, then the angle $\lambda$ is the same as the azimuth ($\lambda=\theta$). $\lambda$ changes after and before PS-AMO. This variation is responsible for the event movement along the common conversion point.

Figure 2 shows how event movement along CCP changes with depth. This is due to the dependence of $\vec{D}$ with respect to $v_p$, $\gamma$ and $t_n$. This variance with depth will persist even in a constant velocity media. Figure 2 also illustrates that the time after PS-AMO ($t_2$) has a new $\vec{h}$ and $\vec{D}$, therefore, a new CCP position.

Continuing with the procedure presented by Fomel and Biondi (1995) to obtain the PS-AMO operator, we cascade PS-DMO [equation (5)] with its inverse. Figure 3 shows a scheme

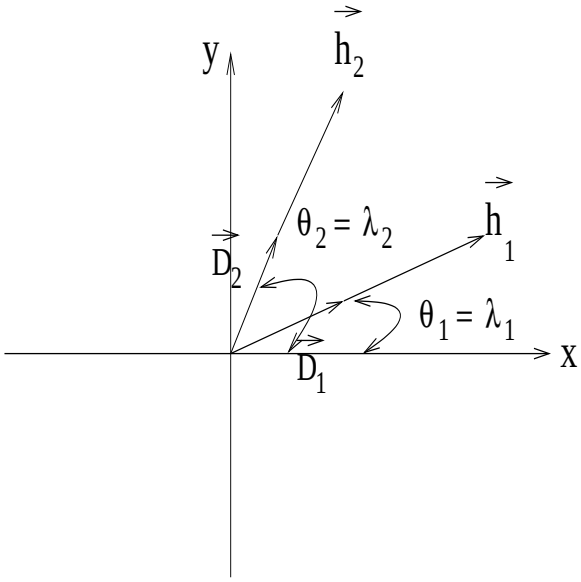



Figure 1: Definition of offset vector $\vec{h}$ and transformation vector $\vec{D}$, before and after PS-AMO daniel1-rot [NR]

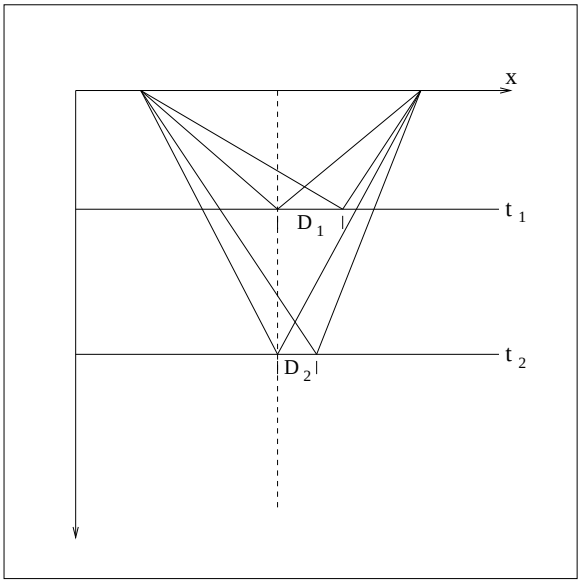



Figure 2: Comparison between the CMP and CCP position in the PS-AMO operator daniel1-plane2 [NR]

of the PS-AMO transformation. A trace with input offset vector $\vec{h}_1$ and midpoint at the origin is transformed into equivalent data with output offset vector $\vec{h}_2$ and midpoint position $\vec{x}$. The data is first transformed to its corresponding CCP position and $\vec{D} = \vec{0}$. Subsequently, the inverse PS-DMO repositions the data to a new midpoint position $\vec{x}$ with a new offset vector $\vec{h}_2$. The new trace position is defined by

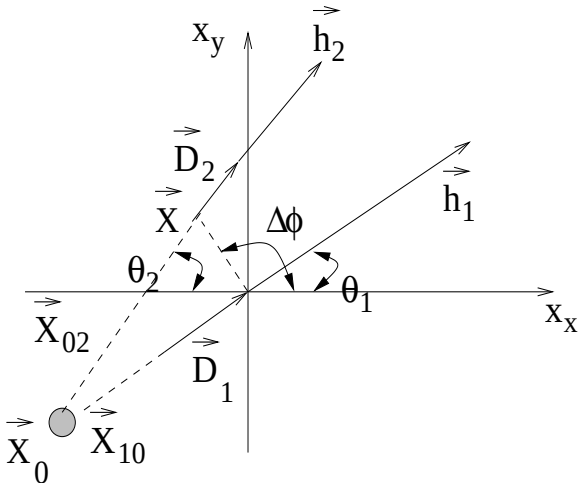



Figure 3: CMP-CCP plane, PS-AMO geometrical interpretation. daniel1-plane [NR]

$$t_2^2 = t_1^2 \frac{\vec{H}_{02}^2}{\vec{H}_{10}^2} \left( \frac{\vec{H}_{10}^2 - \vec{x}_{10}^2 - \vec{D}_{10}^2 - 2\vec{x}_{10}\vec{D}_{10}}{\vec{H}_{02}^2 - \vec{x}_{02}^2 - \vec{D}_{02}^2 - 2\vec{x}_{02}\vec{D}_{02}} \right). \tag{6}$$

Both $\vec{x}_{10}$ and $\vec{x}_{02}$ can be expressed as terms of the final midpoint position $\vec{x}$ by using the rule of sines in the triangle ($\vec{x}$,$\vec{x}_{10}$,$\vec{x}_{02}$) in Figure 3 as

$$\begin{aligned} \vec{x}_{10} &= \vec{x} \frac{\sin(\theta_2 - \Delta\phi)}{\sin(\theta_1 - \theta_2)} \\ \vec{x}_{02} &= \vec{x} \frac{\sin(\theta_1 - \Delta\phi)}{\sin(\theta_1 - \theta_2)}. \end{aligned}$$

The final expression takes the form of

$$t_2^2 = t_1^2 \frac{\vec{H}_{02}^2}{\vec{H}_{10}^2} \left\{ \frac{\vec{H}_{10}^2 \sin^2(\theta_1 - \theta_2) - \vec{x}^2 \sin^2(\theta_2 - \Delta\phi) - \mathbf{D}_1}{\vec{H}_{02}^2 \sin^2(\theta_1 - \theta_2) - \vec{x}^2 \sin^2(\theta_1 - \Delta\phi) - \mathbf{D}_2} \right\}, \tag{7}$$

where

$$\begin{aligned} \mathbf{D}_1 &= \vec{D}_{10}^2 \sin^2(\theta_1 - \theta_2) + 2\vec{x}\vec{D}_{10} \sin(\theta_2 - \Delta\phi) \sin(\theta_1 - \theta_2) \cos\theta_1, \\ \mathbf{D}_2 &= \vec{D}_{02}^2 \sin^2(\theta_1 - \theta_2) + 2\vec{x}\vec{D}_{0i} \sin(\theta_1 - \Delta\phi) \sin(\theta_1 - \theta_2) \cos\theta_2. \end{aligned}$$

This expression represents the azimuth rotation in both the CCP domain and the CMP domain.

## PS-AMO IN THE *F-K* LOG-STRETCH DOMAIN

In order to implement a fast azimuth moveout operator, we use the PS-DMO operator in the frequency-wavenumber log-stretch domain (Xu et al., 2001; Rosales, 2002). The *f-k* log-stretch operator for PS-DMO in 3D takes the form

$$P(\Omega,\vec{k},\vec{h}) = P(\Omega,\vec{k},\vec{h})F(\Omega,\vec{k},\vec{h})e^{i\vec{k}\cdot\vec{D}}. \tag{8}$$

Following Vlad and Biondi's (2001) approach, we construct the PS-AMO operator in this domain. The operator takes the form of

$$P(\Omega,\vec{k},\vec{h}_2) = P(\Omega,\vec{k},\vec{h}_1)\frac{F(\Omega,\vec{k},\vec{h}_1)}{F(\Omega,\vec{k},\vec{h}_2)}e^{i\vec{k}\cdot(\vec{D}_1-\vec{D}_2)}, \tag{9}$$

where

$$F(\Omega,\vec{k},\vec{h}_i) = \begin{cases} 0 & \text{for } \vec{k}\cdot\vec{h}_i = 0 \\ 0 & \text{for } \Omega = 0 \\ e^{\frac{i}{2}\Omega\log\frac{1}{2}\left(\sqrt{\left(\frac{2\vec{k}\cdot\vec{H}_i}{\Omega}\right)^2}+1+1\right)} & \text{otherwise,} \end{cases} \tag{10}$$

and $\vec{k}$ is the spatial frequency vector for the midpoints coordinates (for this case, the vectors $\vec{D}$ and $\vec{H}$ are the same as presented in the previous section).

Rosales (2002) discusses a more accurate PS-DMO operator in this domain. This new PS-DMO operator distributes the amplitudes correctly along strong dip events. This operator is just the extension of Zhou et al. (1996) for PS data. Using the improved operator presented by Rosales (2002) the filter $F(\Omega,\vec{k},\vec{h}_i)$ takes the form:

$$F(\Omega,\vec{k},\vec{h}_i) = \begin{cases} 0 & \text{for } \vec{k}\cdot\vec{h}_i = 0 \\ \vec{k}\cdot\vec{H}_i & \text{for } \Omega = 0 \\ e^{\frac{i}{2}\Omega\left\{\sqrt{1+\left(\frac{2\vec{k}\cdot\vec{H}_i}{\Omega}\right)^2}-1-\log\frac{1}{2}\left[\sqrt{\left(\frac{2\vec{k}\cdot\vec{H}_i}{\Omega}\right)^2}+1+1\right]\right\}} & \text{otherwise.} \end{cases} \tag{11}$$

## NUMERICAL EXAMPLES

Figure 4 compares the PP-AMO impulses response obtained with the filter in both equation (10) (top) and equation (11) (bottom). Both are obtained with a value of $\gamma = 1$ for the two cases. Both impulse responses are kinematicly equal. However, the dynamically behavior is different, the amplitudes distribution with the filter in equation (11) is more accurate. The impulse response with the filter in equation (11) and $\gamma = 1$ is exactly the same as Vlad and Biondi (2001).

Figure 5 presents a similar comparison to Figure 4, for the case of converted waves. Here, we use $\gamma = 1.2$ and $v_p = 2.0$Km/s. As in the previous case, the same kinematic behavior occurs in both operators, but the response with the filter in equation (11) is dynamically correct.

Figure 6 shows the comparison between the PP-AMO impulse response and the PS-AMO impulse response. The PS-AMO not only has the same saddle shape as the PP-AMO operator, but it also exhibits a lateral movement. This lateral displacement correspond to the asymmetry of the raypaths or the CCP transformation. The displacement is toward the lower-left part of the cube.

A variation of the PS-AMO impulse response with respect to depth is also observable in Figure 7. This behavior is due to the dependence of the operator on $\frac{v_p}{v_s}$ ratio It is possible to observe how the impulses responses movement toward the left is stronger for shallower events. It is also possible to detect how the response change along the crossline coordinate.

## DISCUSSION AND CONCLUSION

We present a partial prestack migration operator for converted wave, the PS-AMO operator. This operator is a cascade operator of PS-DMO and inverse PS-DMO. We implement the PS-AMO operator in the log-stretch frequency-wavenumber domain.

Our operator is able to handle the amplitudes correctly, since it uses an accurate PS-DMO operator for this purpose. The PS-AMO operator is velocity dependent, and we assume constant velocity. Therefore, the velocity in the new position is the same as the velocity in the previous position.

The PS-AMO operator transforms data to an arbitrary offset and azimuth. This transformation also considers the CCP binning transformation characteristic of PS data. Therefore, a priori CCP binning is not necessary before applying azimuth moveout to converted wave data.

Our operator can be applied to the regularization of OBS data, where the knowledge of our seismic image is of crucial importance. This is the next step in our project.

## REFERENCES

Biondi, B., and Vlad, I., 2001, Amplitude preserving prestack imaging of irregularly sampled 3-D data: SEP–**110**, 1–18.

Biondi, B., Fomel, S., and Chemingui, N., 1998, Azimuth moveout for 3-D prestack imaging: Geophysics, **63**, no. 2, 574–588.

Biondi, B., 2000, 3-D Seismic Imaging: Class notes, http://sepwww.stanford.edu/sep/biondo/Lectures.

Chemingui, N., and Biondi, B., 1997, Equalization of irregular data by iterative inversion: Expanded Abstracts, 1115–1118.

Chemingui, N., 1999, Imaging irregularly sampled 3D prestacked data: SEP–**101**.

Fomel, S., and Biondi, B., 1995, The time and space formulation of azimuth moveout: SEP–**84**, 25–38.

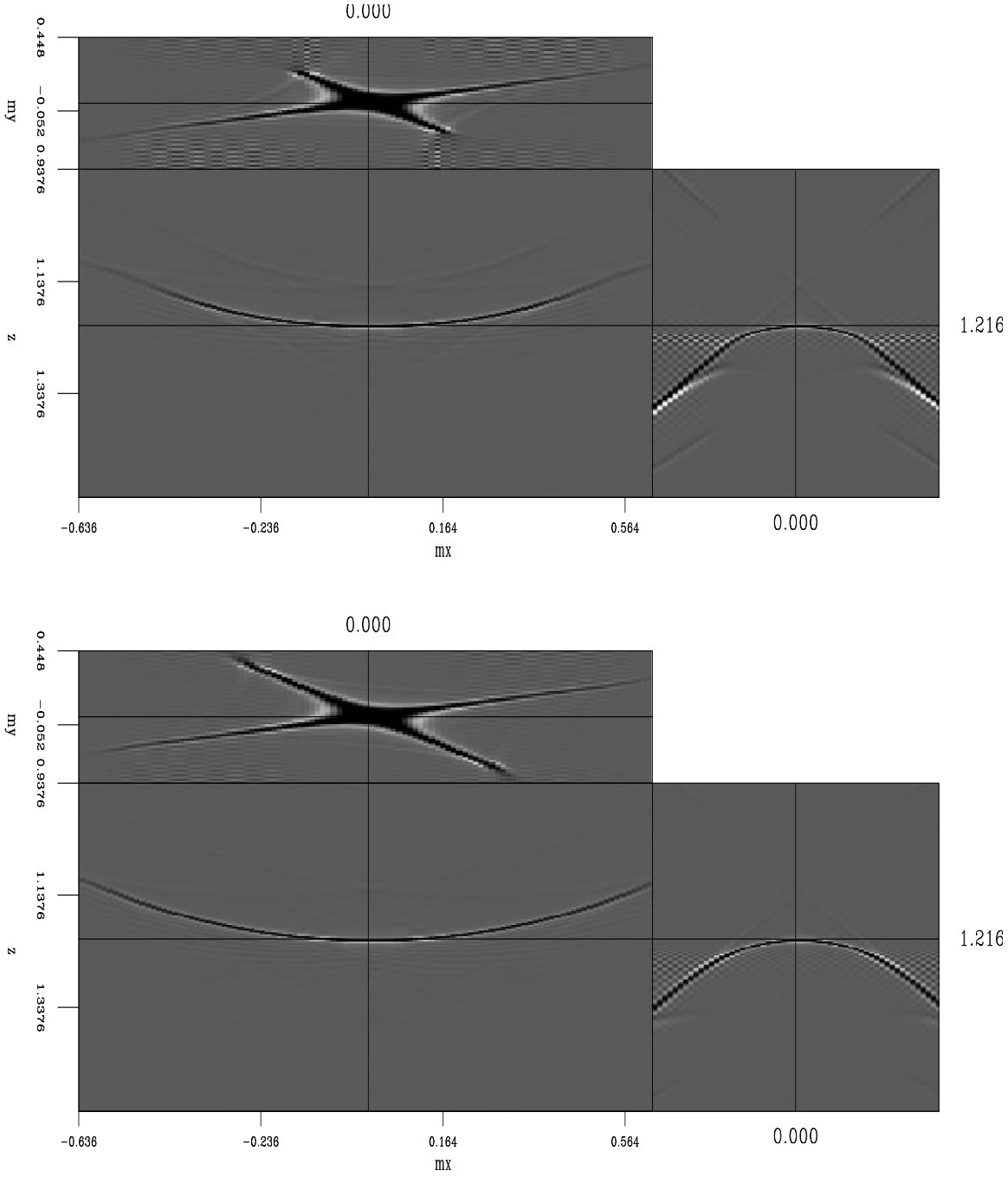



Figure 4: PP-AMO impulse response comparison, filter in equation (10) (top) and filter in equation (11) (bottom), with $\gamma = 1$. daniel1-both [ER]

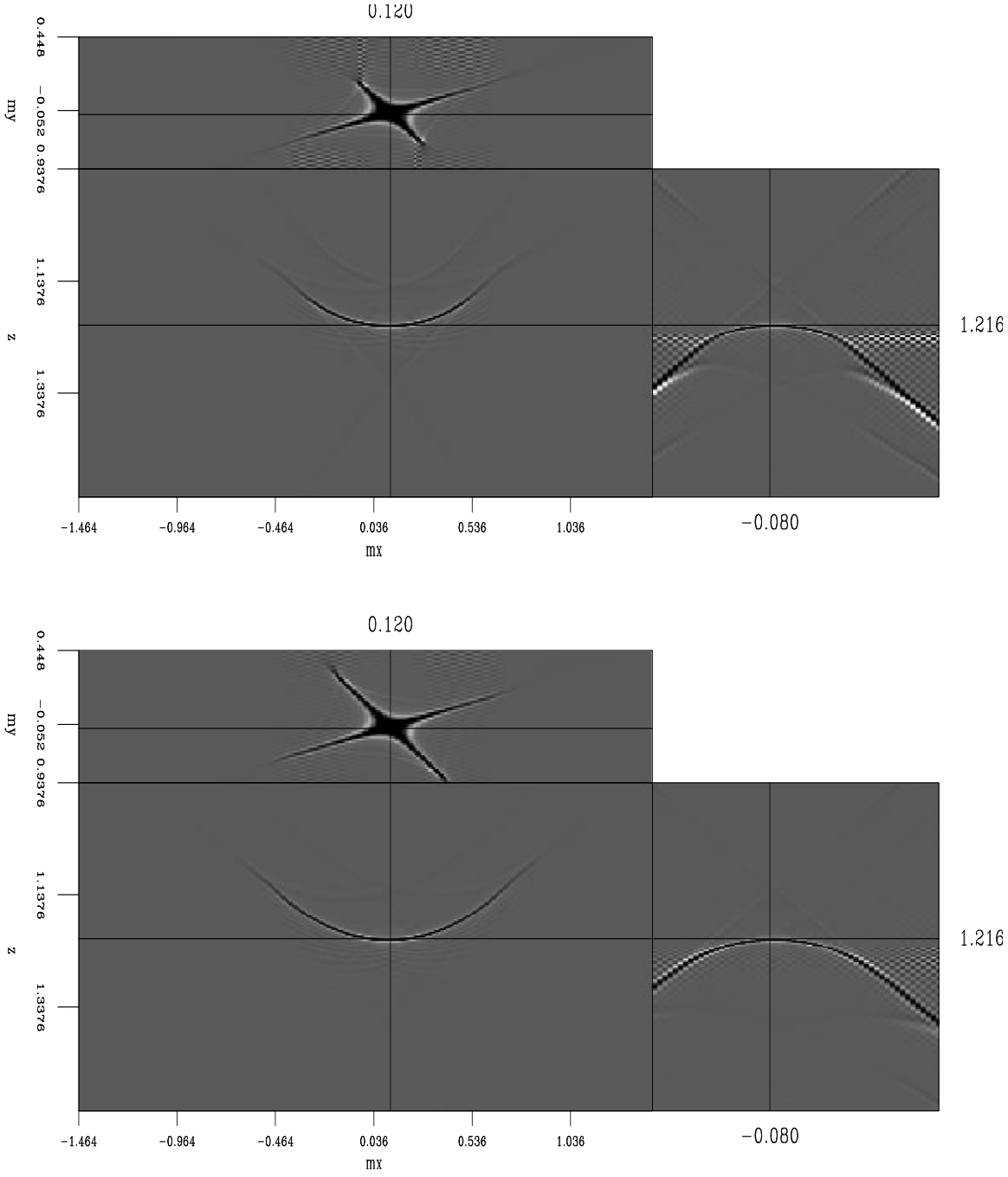



Figure 5: PS-AMO impulse response comparison, filter in equation (10) (top) and filter in equation (11) (bottom), with $\gamma = 1.2$ and $v_p = 2.0$km/s. daniel1-both2 [ER]

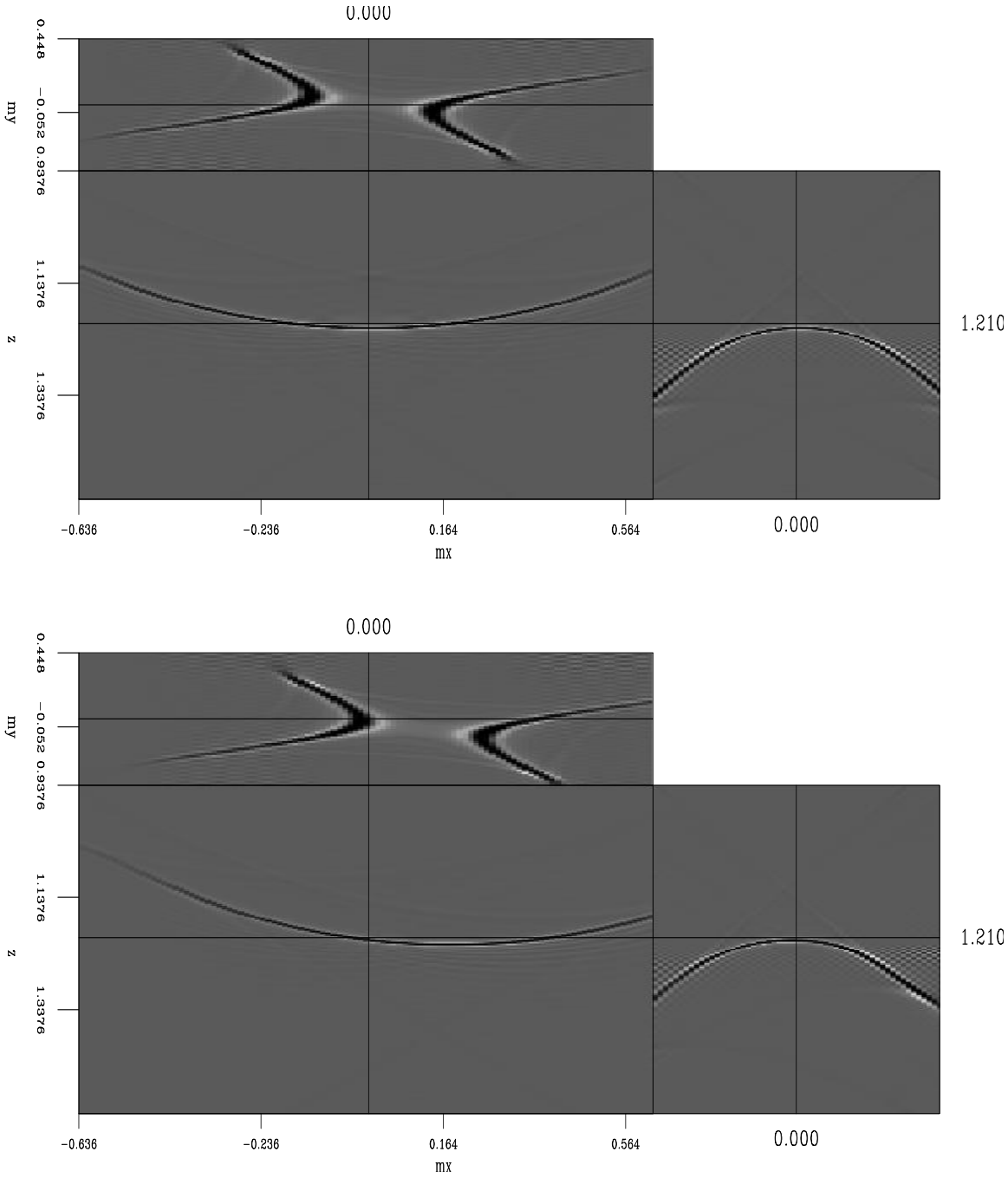



Figure 6: PP-AMO (top) vs. PS-AMO (bottom) impulse response comparison daniel1-both3 [ER]

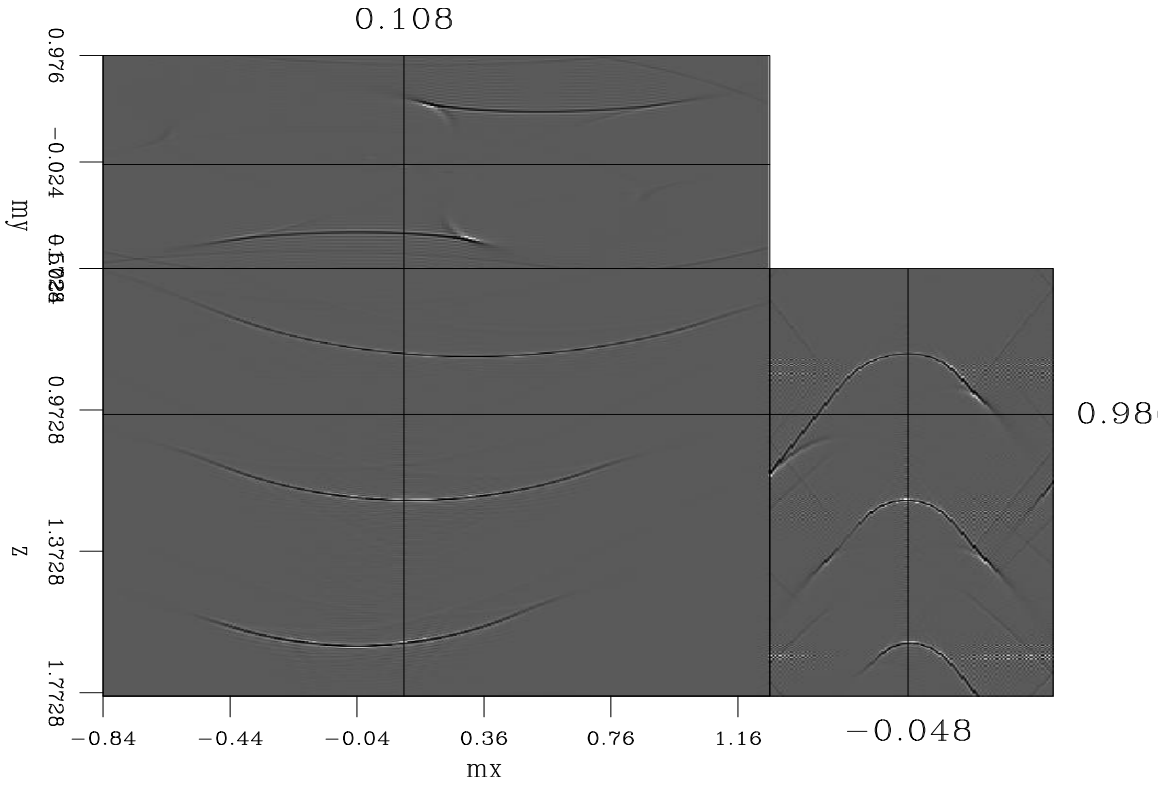



Figure 7: PS-AMO impulses response variation with $\tau$ daniel1-amo2 [ER]

Harrison, M., 1990, Converted wave DMO: 60th Annual Internat. Mtg., Society Of Exploration Geophysicists, Expanded Abstracts, 1370–1373.

Rosales, D., 2002, Converted wave dip moveout: SEP–**111**, 47–59.

Vlad, I., and Biondi, B., 2001, Effective AMO implementation in the log-stretch, frequency-wavenumber domain: SEP–**110**, 63–70.

Xu, S., Jin, S., Ma, Z., and Geng, J., 2001, A fast converted wave dip moveout in f-k domain: 71st Annual Internat. Mtg., Society Of Exploration Geophysicists, Expanded Abstracts, 1851–1854.

Zhou, B., Mason, I. M., and Greenhalgh, S. A., 1996, An accurate formulation of log-stretch dip moveout in the frequency-wavenumber domain: Geophysics, **61**, no. 3, 17–23.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 71–81*

# Multidimensional imaging condition for shot profile migration

*Alejandro A. Valenciano, Biondo Biondi, and Antoine Guitton*[1]

**ABSTRACT**

Conventional shot profile migration schemes determine the reflection strength at each subsurface point taking into account only the downgoing and the upgoing wavefields at that location. Since events in the subsurface are not uncorrelated, a better imaging condition could be one that makes use of the downgoing and upgoing wavefields in a neighborhood of the point where the reflection strength is calculated. A generalized multidimensional deconvolution imaging condition could be the solution to integrating information from the neighboring points, but issues related with deconvolution stability still need to be solved. An alternative to deconvolution may be a new regularized least squares imaging condition. This could be a feasible approach since the regularization operator can favor a predetermined distribution of the reflectivity. Improvements can be done in the conventional industry imaging condition adding a spatially variant damping factor, even without including information of the neighboring points.

## INTRODUCTION

Shot profile migration includes three different steps: downgoing wavefield propagation, upgoing wavefield backward propagation and imaging. The last step, imaging, is based on Claerbout's imaging principle (Claerbout, 1971). According to this principle, a reflector exists at a point where the upgoing and the downgoing wavefields coincide in time and space (Figure 1).

There are two distinct aspects behind this principle: the kinematic (coincidence in time and space of upgoing and downgoing wavefields) and the dynamic (reflection strength at the coincidence point). Conventional migration schemes (Jacobs, 1982) determine the reflection strength in each subsurface point taking into account only the downgoing and the upgoing wavefields at that location. But these approaches don't consider that the reflectors in the subsurface are spatially correlated. We discuss a multidimensional imaging condition that makes the reflectivity strength dependent on the downgoing and upgoing wavefields in a neighborhood of the subsurface point.

Based on the equivalence of the deconvolution imaging condition with the exact solution of the least squares fitting goal, we propose a regularization scheme for the imaging condition that has the potential of including the previous knowledge of the image in the regularization operator. This regularization approach has been used to steer the final image in least squares

[1]**email:** valencia@sep.stanford, biondo@sep.stanford.edu, antoine@sep.stanford.edu

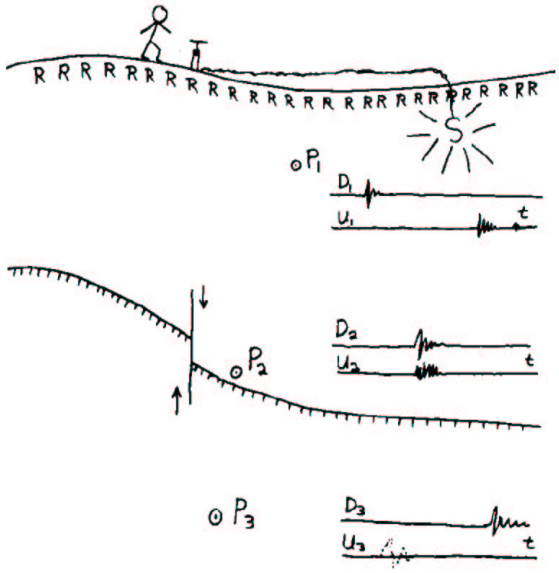



Figure 1: Reflector mapping imaging principle, note that the reflector exists at a point where the upgoing and downgoing wavefields coincide in time and space. Taken from Claerbout (1971) alejandro1-jon [NR]

inversion (Prucha and Biondi, 2002).

First, we explore a multidimensional deconvolution imaging condition. No clear results were obtained using this approach yet, but it allows us to study the difficulties of multidimensional deconvolution. Second, a least squares regularized scheme for the imaging condition is stated.

Finally we test, in a synthetic experiment, a space variable damping factor to improve conventional industry imaging condition. In this case we only consider point to point dependence of the reflection strength.

## MULTIDIMENSIONAL DECONVOLUTION IMAGING CONDITION

Claerbout (1971) expresses the reflector mapping principle by the formula

$$\mathbf{r}(x,z) = \frac{\mathbf{u}(x,z,t_d)}{\mathbf{d}(x,z,t_d)}, \tag{1}$$

where $x$ is the horizontal coordinate, $z$ is the depth, $t_d$ is the time at which the downgoing wave $\mathbf{d}(x,z,t_d)$ and the upgoing wave $\mathbf{u}(x,z,t_d)$ coincide in time. This principle states that for time equal $t_d$ the reflectivity strength $\mathbf{r}(x,z)$ depends only on the downgoing wave at $(x,z)$ and on the upgoing wave at $(x,z)$. No particular distribution is assumed for the reflectivity in the horizontal direction or in depth. Neither a dependence of the reflectivity of the future (wavefields anteceding $t_d$) or of the past (wavefields preceding $t_d$) is assumed.

Based on the imaging principle described in equation (1) we can propose a more general imaging condition that makes the reflectivity in $(x,z)$ dependent on the downgoing and upgoing wavefields in the neighborhood of $(x,y)$, shown in Figure 2 .

This more general imaging condition can be stated by:

$$\mathbf{r}(x,z) = \sum_t \frac{\mathbf{u}([x-\sigma_x, x+\sigma_x],[z-\sigma_z, z+\sigma_z])}{\mathbf{d}([x-\sigma_x, x+\sigma_x],[z-\sigma_z, z+\sigma_z])}, \tag{2}$$

where the division symbol $(-)$ means 2-D deconvolution of the upgoing wavefield with the downgoing wavefield in the $(x,z)$ plane. The $\sigma_x, \sigma_z$ are small numbers that define a rectangular neighborhood $(x,y)$. This 2-D imaging condition states that there will be more than one point in the downgoing wavefield **d** and the upgoing wavefield **u** contributing to the strength at the point $(x,y)$.

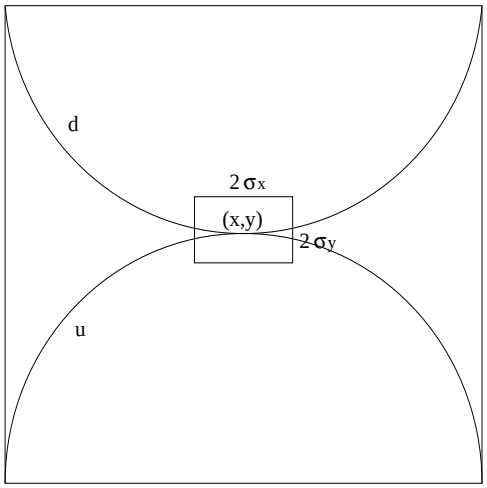



Figure 2: Multidimensional imaging geometry alejandro1-cuboxfig [NR]

To address the multidimensional deconvolution we can make use of the helix concept (Claerbout, 1998). If we put the upgoing and the downgoing wavefields in helical coordinates, we will be able to treat the multidimensional deconvolution as a 1-D deconvolution.

But deconvolution is not an easy task. To have a stable deconvolution we need **d** to be minimum phase, so an approximation of equation (2) could be

$$\mathbf{r} = \sum_{\mathbf{t}} \frac{\mathbf{u}}{\mathbf{d_{mp}}}, \tag{3}$$

where $\mathbf{d_{mp}}$ can be computed in an helix by means of Wilson spectral factorization (Sava et al., 1998) in spatial coordinates $(x,y)$ or by means of Kolmogoroff spectral factorization (Claerbout, 1976) in the Fourier domain.

Now, a new question arises: Does the new imaging condition formulation equation (3) honor Claerbout (1971) imaging principle?

The answer to this question is no, equation (3) gives a shifted version of the image. The minimum phase transformation produces a shift in spatial coordinates $(x,y)$. This shift has to be calculated to obtain a properly placed image.

Some attempts were made to implement the anteceding procedure using Wilson spectral factorization to obtain a minimum phase version of the downgoing wavefield. No convergence of factorization results were obtained. More work needs to be done to understand the causes.

## LEAST SQUARES IMAGING CONDITION

Changing deconvolution for convolution, a different imaging condition can be stated for each time in terms of the following fitting goal:

$$\mathbf{Dr} = \mathbf{u}, \tag{4}$$

where **D** is a convolution matrix in which columns are downshifted versions of the downgoing wavefield **d**.

The least squares solution to this problem is

$$\mathbf{r} = (\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'\mathbf{u}. \tag{5}$$

A damped solution is usually used to guarantee $\mathbf{D}'\mathbf{D}$ to be invertible as in

$$\mathbf{r} = (\mathbf{D}'\mathbf{D} + \varepsilon^2)^{-1}\mathbf{D}'\mathbf{u} \tag{6}$$

where $\varepsilon$ is a small positive number to guarantee no zeroes in $\mathbf{D}'\mathbf{D}$ diagonal. This is equivalent to the fitting goal

$$\begin{aligned} \mathbf{0} &\approx \mathbf{Dr} - \mathbf{u} \\ \mathbf{0} &\approx \varepsilon\mathbf{Ir}, \end{aligned} \tag{7}$$

where **I** is the identity matrix that is used here as the regularization operator. Using this regularization scheme we are adding to the denominator a constant value where it is needed and where it is not.

As it is our intention to use the previous knowledge of how the image should be, we could choose an smarter way to fill the zero values off $\mathbf{D}'\mathbf{D}$ diagonal. We can substitute the regularization operator for one constructed with a priori information, using

$$\begin{aligned} \mathbf{0} &\approx \mathbf{Dr} - \mathbf{u} \\ \mathbf{0} &\approx \varepsilon\mathbf{Ar} \end{aligned} \tag{8}$$

where our regularization operator **A** could be a steering filter (Clapp et al., 1997). Steering filters can efficiently guide the solution toward a more geologically appealing form. This type of filter has been used with success to smooth existing reflectors and fill shadow zones in least squares inversion (Prucha and Biondi, 2002).

## SPACE VARIABLE DAMPING IN CONVENTIONAL IMAGING CONDITION

Conventional shot profile migration schemes determine the reflection strength at each subsurface point taking into account only the downgoing and the upgoing wavefields at that location. Jacobs (1982) compares two different imaging conditions

$$\mathbf{r} = \sum_{\mathbf{t}} \mathbf{ud}, \tag{9}$$

and

$$\mathbf{r} = \sum_{\mathbf{t}} \frac{\mathbf{ud}}{\mathbf{d^2} + \varepsilon^{\mathbf{2}}}. \tag{10}$$

The first is one commonly used by the industry. It has the advantage of being robust, but has the disadvantage of not computing the correct amplitudes. The second computes the correct amplitudes (except for a damping factor $\varepsilon$), but has the disadvantage of being unstable due to zero division. That is why a damping factor $\varepsilon$ is needed.

We propose to add a mask function defined as

$$\mathbf{w} = \begin{cases} 0 & \text{if} \quad \mathbf{ud} > \alpha \\ 1 & \text{otherwise} \end{cases} \tag{11}$$

where $\alpha$ can be variable in space.

When **ud** has enough energy to contribute to the image, the damping factor $\varepsilon$ is set to zero. When factor **ud** is small, the damping factor is kept to avoid zero division. Thus, the imaging condition can be set as

$$\mathbf{r} = \sum_{\mathbf{t}} \frac{\mathbf{ud}}{\mathbf{d^2} + \mathbf{w}\varepsilon^{\mathbf{2}}}, \tag{12}$$

where the damping is now variable in space.

A simple synthetic was generated to test the preceding idea using wave equation modeling. Figure 3a shows the downgoing wave, and Figure 3b the upgoing wave, at a fixed time. Figure 4 shows the mask **w** used in this example.

Figure 5a shows the reflection strength calculated using the imaging condition stated in equation (9). Figure 5b shows the reflection strength calculated using division of the upgoing wavefield **u** by the downgoing wavefield **d**. Figure 5c shows the reflection strength calculated using the imaging condition stated in equation (10), and Figure 5d shows the reflection strength calculated using the imaging condition stated in equation (12). The advantage of Figure 5d's result over the others is that it has the correct reflection strength value inside the masked area and doesn't diverge outside it because of the damping factor.

In Figure 6 we compare the two imaging conditions stated in equations (10) and (12) inside the masked area for two different $\varepsilon$. We can see for the imaging condition stated in equation (12) that the reflection strength inside the masked area doesn't change. This is an important advantage of space variable damping imaging principle, because it let us to build an adaptive mask dependent of the subsurface illumination.

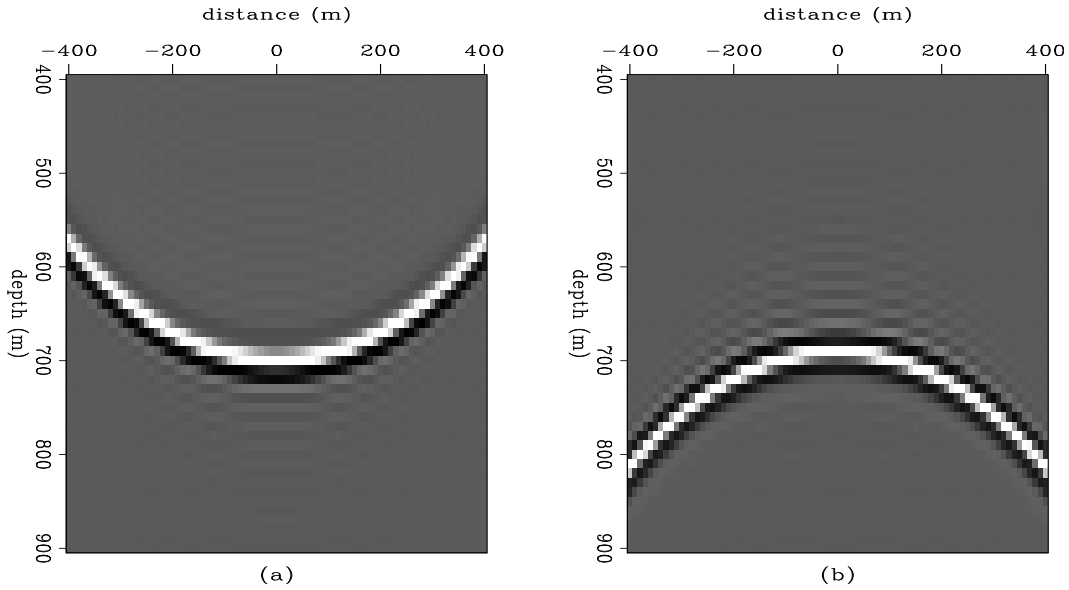



Figure 3: Wavefields at a fixed time. a) Downgoing wave, b) Upgoing wave. alejandro1-DU [ER]

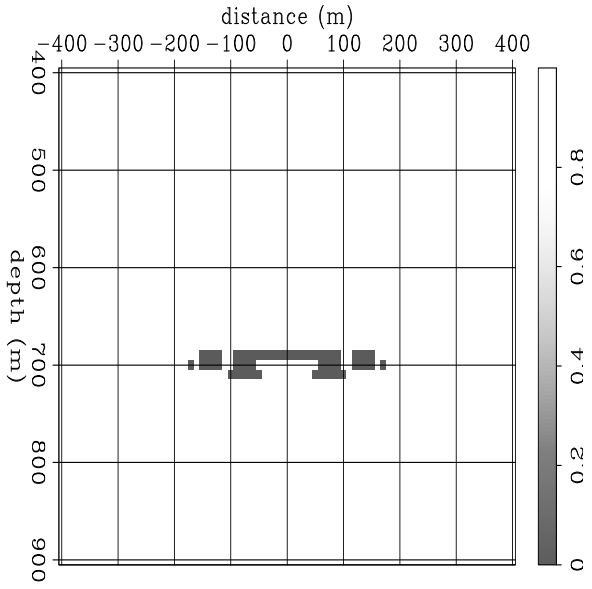



Figure 4: Mask used in equation (12). Zero at masked area and one out of the masked area. alejandro1-ma [ER]

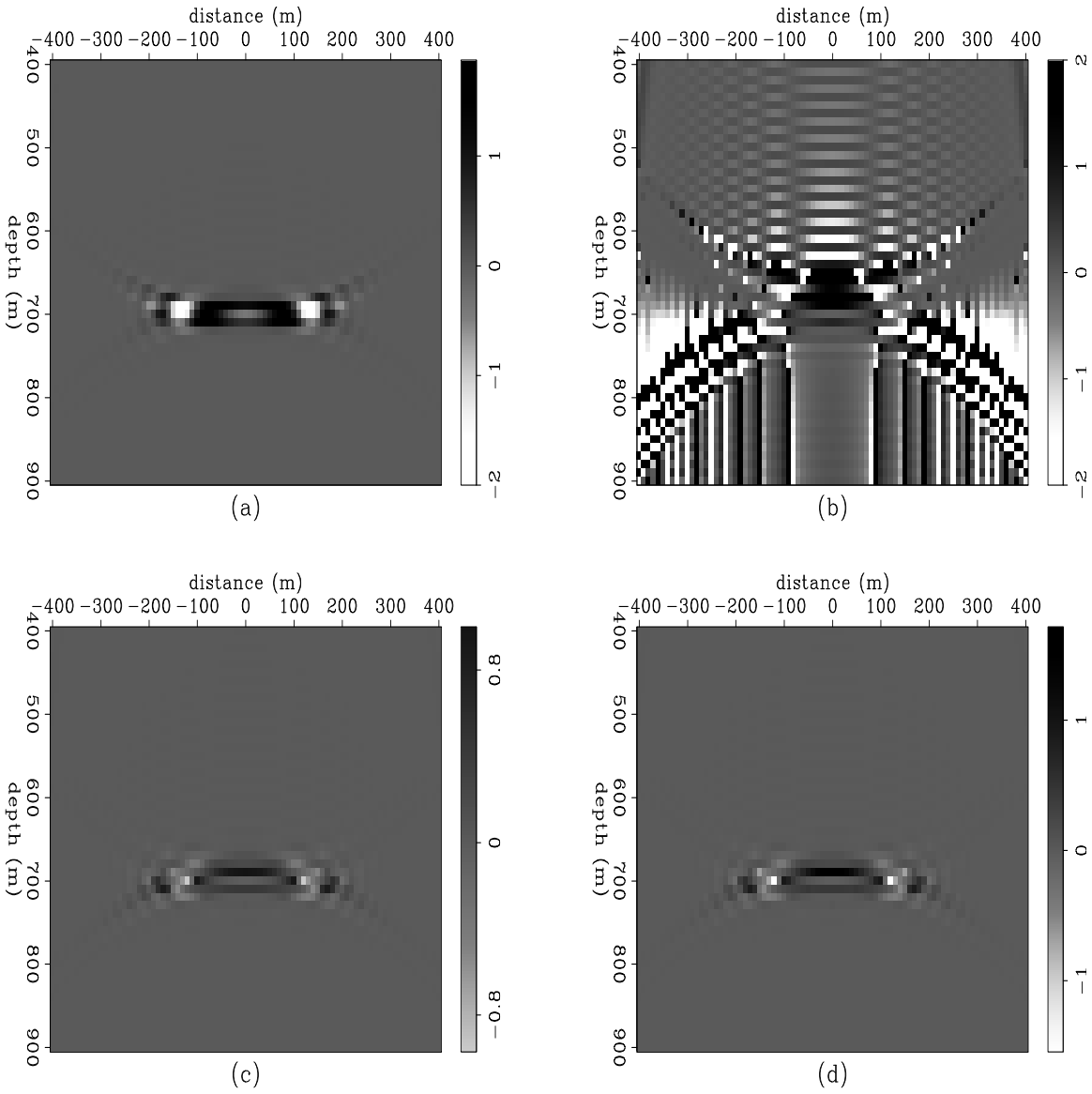



Figure 5: Comparison between four different imaging conditions a) Calculated by wavefield multiplication equation (9), b) Calculated by wavefield division (**u**/**d**), c) Calculated using constant damping equation (10), and d) Calculated using space variable damping equation (12). alejandro1-comp_1shot [ER]

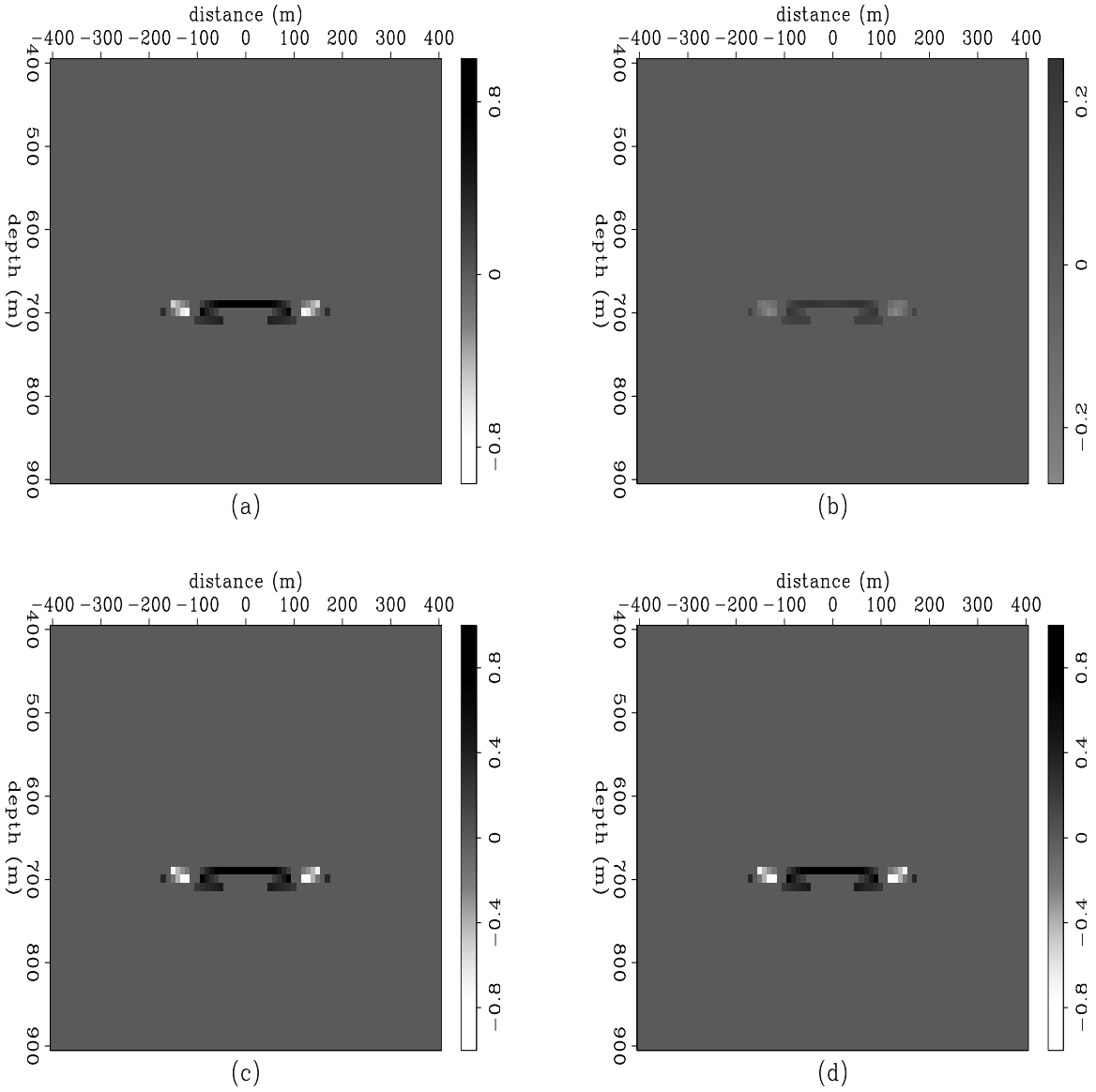



Figure 6: Comparison between imaging condition stated in equations (10) and (12) inside the masked area. a) $\varepsilon = 0.5$, b) $\varepsilon = 5$, c) $\varepsilon = 0.5$, d) $\varepsilon = 5$ alejandro1-comp_im [ER]

We stack the reflection strength from 11 shots to see how the change observed in one shot affects the final image. The result is shown in Figure 7. We can see imaging condition from equation (12) gives the best resolution.

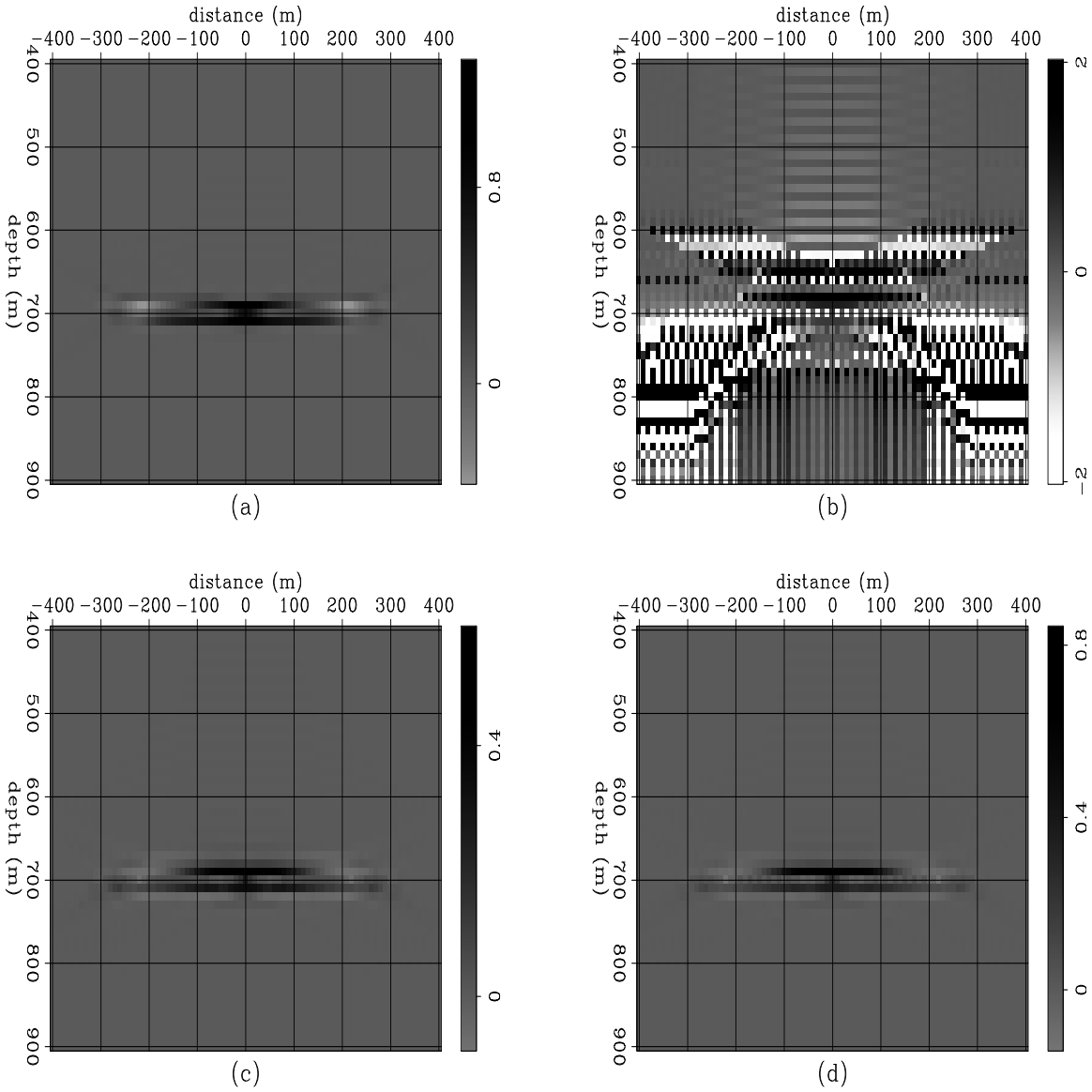



Figure 7: Comparison between 11 shot stacks using three different imaging conditions, a) equation (9), b) wavefield division (**u**/**d**), c) equation (10), and d) equation (12). alejandro1-comp_stack [ER]

## CONCLUSION

A generalized multidimensional deconvolution imaging condition could be the solution to integrating information from the neighboring points in the computation of reflection strength for

shot profile migration. However, issues related to deconvolution stability and proper placement of the image still need to be solved. An alternative to deconvolution, stating a new regularized least squares imaging condition, could be a feasible approach as the regularization operator can be set to favor a predetermined spatial distribution.

We showed, in a synthetic experiment, that a spatial variant damping factor can improve the resolution and amplitude preservation of the conventional industry imaging condition. The damping factor can be related to reflector illumination, adding the damping factor where it is really needed.

## REFERENCES

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, no. 3, 467–481.

Claerbout, J. F., 1976, Fundamentals of geophysical data processing: Blackwell.

Claerbout, J. F., 1998, Multidimensional recursive filters via a helix: Geophysics, **63**, no. 5, 1532–1541.

Clapp, R. G., Fomel, S., and Claerbout, J., 1997, Solution steering with space-variant filters: SEP–**95**, 27–42.

Jacobs, B., 1982, The prestack migration of profiles: SEP–**34**.

Prucha, M. L., and Biondi, B. L., 2002, Subsalt event regularization with steering filters: SEP–**111**, 1–17.

Sava, P., Rickett, J., Fomel, S., and Claerbout, J., 1998, Wilson-Burg spectral factorization with application to helix filtering: SEP–**97**, 343–351.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 81–99*

# Wave-equation migration velocity analysis beyond the Born approximation

*Paul Sava and Sergey Fomel*[1]

**ABSTRACT**

The Born approximation is based on the assumption of small slowness perturbation. We investigate the limits of the Born approximation when applied to wave-equation migration velocity analysis and propose two new schemes which allow for larger slowness anomalies, while improving accuracy and increasing stability. The new schemes are based on linearizations of exponential functions using bilinear and implicit approximations, rather than the (Born) explicit approximation. We demonstrate the feasibility of our new operators on a synthetic example with highly variable background and strong slowness anomalies.

## INTRODUCTION

Migration velocity analysis based on downward-continuation methods, commonly referred to as *wave-equation migration velocity analysis* (WEMVA), is a promising technique which has become an active area of research over the recent years (Biondi and Sava, 1999; Sava and Biondi, 2000, 2001). The main idea of WEMVA is to use downward-continuation operators not only for migration, but also for migration velocity analysis. This is in contrast with other techniques which use downward-continuation for migration but traveltime-based techniques for migration velocity analysis (Clapp, 2001; Liu et al., 2001; Mosher et al., 2001).

The main benefits of WEMVA are identical to the benefits of downward-continuation migration methods versus the more common Kirchhoff methods. Among these benefits, the most important are the accurate handling of complex wavefields, characterized by multipathing, and the band-limited nature of the imaging process, which can handle sharp velocity variations much better than traveltime-based methods (Woodward, 1992). The areas of complex geology are those where WEMVA is expected to provide the largest benefits.

The problem with WEMVA is that, in its simplest form, it is based on the Born approximation of the wavefield in the perturbation region. This leads to severe limitations of the magnitude and size of the anomalies that can be resolved, which means that, in principle, it cannot operate successfully in the regions of high complexity where it is needed most.

The limitations imposed by the Born approximation can be partially circumvented by spe-

[1]**email:** paul@sep.stanford.edu, fomel@math.lbl.gov

cial ways of creating the image perturbation in connection with residual migration (Sava and Biondi, 2001). This process of creating Born-compliant image perturbations is not the ideal strategy, since it closely links a highly accurate method, wavefield-continuation, to a less accurate method, Stolt residual migration.

In this paper, we introduce a new method of linearization designed to overcome the limitations imposed by the Born approximation. Our method is based on linearizations of the exponential function containing the slowness perturbation using more accurate approximations than Born linearization. The resulting operator is more accurate and also more stable in areas of high contrast, at a cost that is practically identical to the one of the Born linearized operator.

This paper is organized as follows: in the next two sections we review the theory of downward-continuation and wave-equation MVA using the Born approximation; then, we introduce the new operators and analyze their meaning in the general context of non-linear optimization; and finally, we present a synthetic example that demonstrates the features of our new method.

## DOWNWARD-CONTINUATION MIGRATION

In migration by downward-continuation, the wavefield at depth $z+\Delta z$, $\mathcal{W}(z+\Delta z)$, is obtained by phase-shift from the wavefield at depth $z$, $\mathcal{W}(z)$.

$$\mathcal{W}(z+\Delta z) = \mathcal{W}(z)\, e^{-ik_z \Delta z}. \tag{1}$$

This equation corresponds to the analytical solution of the ordinary differential equation

$$\mathcal{W}'(z) = -ik_z \mathcal{W}(z), \tag{2}$$

where the $'$ sign represents a derivative with respect to the depth $z$.

We can consider that the depth wavenumber ($k_z$) depends linearly, through a Taylor series expansion, on its value in the reference medium ($k_{z_r}$) and the laterally varying slowness in the depth interval from $z$ to $z+\Delta z$, $s(x,y,z)$

$$k_z \approx k_{z_r} + \left.\frac{dk_z}{ds}\right|_{s=s_r} (s - s_r), \tag{3}$$

where $s_r$ represents the constant slowness associated with the depth slab between the two depth intervals, and $\left.\frac{dk_z}{ds}\right|_{s=s_r}$ represents the derivative of the depth wavenumber with respect to the reference slowness and which can be implemented in many different ways (Sava, 2000). The wavefield downward-continued through the *background* slowness $s_b(x,y,z)$ can, therefore, be written as

$$\mathcal{W}_b(z+\Delta z) = \mathcal{W}(z)\, e^{-i\left[k_{z_r} + \left.\frac{dk_z}{ds}\right|_{s=s_r} (s_b - s_r)\right]\Delta z}, \tag{4}$$

from which we obtain that the full wavefield $\mathcal{W}(z+\Delta z)$ depends on the background wavefield $\mathcal{W}_b(z+\Delta z)$ through the relation

$$\mathcal{W}(z+\Delta z) = \mathcal{W}_b(z+\Delta z)\, e^{-i\,\frac{dk_z}{ds}\Big|_{s=s_r}\Delta s\,\Delta z}, \tag{5}$$

where $\Delta s$ represents the difference between the true and background slownesses $\Delta s = s - s_b$.

## BORN WAVE-EQUATION MVA

We define the *wavefield perturbation* $\Delta\mathcal{W}(z+\Delta z)$ as the difference between the wavefield propagated through the medium with correct velocity $\mathcal{W}(z+\Delta z)$ and the wavefield propagated through the background medium $\mathcal{W}_b(z+\Delta z)$. With these definitions, we can write

$$\Delta\mathcal{W}(z+\Delta z) = \mathcal{W}(z+\Delta z) - \mathcal{W}_b(z+\Delta z), \tag{6}$$

or

$$\Delta\mathcal{W}(z+\Delta z) = \mathcal{W}_b(z+\Delta z)\left[e^{-i\,\frac{dk_z}{ds}\Big|_{s=s_r}\Delta s\,\Delta z} - 1\right]. \tag{7}$$

Equation (7) represents the foundation of the wave-equation migration velocity analysis method (Biondi and Sava, 1999). The major problem with Equation (7) is that the wavefield $\Delta\mathcal{W}$ and slowness perturbations $\Delta s$ are not related through a linear relation, therefore, for inversion purposes, we need to further approximate it by linearizing the equation around the reference slowness ($s_r$)

Biondi and Sava (1999) choose to linearize Equation (7) using the Born approximation ($e^{i\phi} \approx 1 + i\phi$), from which the WEMVA equation becomes

$$\Delta\mathcal{W}(z+\Delta z) = \mathcal{W}_b(z+\Delta z)\left[-i\,\frac{dk_z}{ds}\Big|_{s=s_r}\Delta s\,\Delta z\right]. \tag{8}$$

The problem with the Born linearization, Equation (8), is that it is is based on an assumption of small phase perturbation,

$$1 + i\phi \approx \lim_{\phi\to 0} e^{i\phi}$$

which mainly translates into small slowness perturbations. This fact is more apparent if we recall that the linearization $e^{i\phi} \approx 1 + i\phi$ corresponds to an explicit numerical solution of the differential equation (2), a numerical solution which is notoriously unstable unless precautions are taken to consider small propagation steps. The main consequence of the limitations imposed by the Born approximation is that WEMVA can only consider small perturbations in the slowness model, which are likely too small relative to the demands of real problems. Since non-linear inversion is still not feasible for large size problems like the ones typical for seismic imaging, we seek other ways of linearizing Equation (7) which would still enable us to solve our inversion problem within the framework of linear optimization theory.

## HIGHER ACCURACY LINEARIZATIONS

As noted earlier, the approximation

$$e^{i\phi} \approx 1 + i\phi \tag{9}$$

corresponds to an *explicit* numerical solution to the differential equation (2). However, this is neither the only possible solution, nor the most accurate, and furthermore it is only conditionally stable.

We can, however, solve Equation (2) using other numerical schemes. Two possibilities are *implicit* numerical solutions, where we approximate

$$e^{i\phi} \approx \frac{1}{1 - i\phi}, \tag{10}$$

or *bilinear* numerical solutions, where we approximate

$$e^{i\phi} \approx \frac{2 + i\phi}{2 - i\phi}. \tag{11}$$

Equations (9) and (10) are first order, but Equation (11) is second order accurate as a function of the phase $\phi$. Numerical schemes based on Equation (9) are conditionally stable, but numerical schemes based on Equations (10) and (11) are unconditionally stable.

In the context of partial differential equations, the bilinear approximation (11) is known under the name of Crank-Nicolson and has been extensively used in migration by downward-continuation using the paraxial wave equation (Claerbout, 1985). Figures 1 and 2 compare the approximations in Equations (9), (10) and (11) as a function of phase. Both the explicit and implicit solutions lead to errors in amplitude and phase, while the bilinear solution leads just to errors in phase (Figure 2).

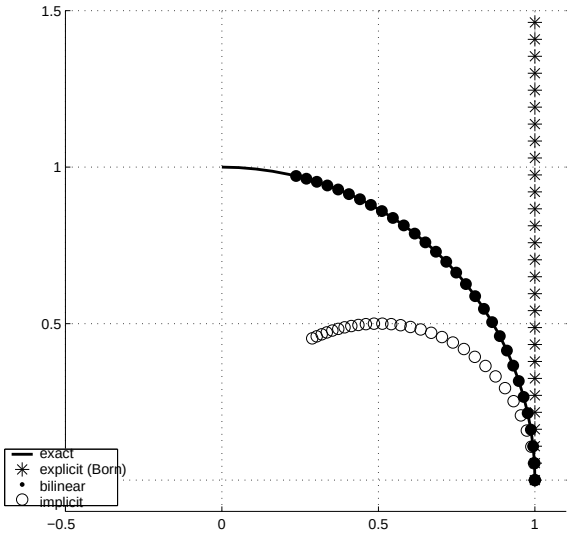



Figure 1: Explicit, bilinear and implicit approximations plotted on the unit circle. The solid line corresponds to the exact exponential solution. paul1-unit [NR]

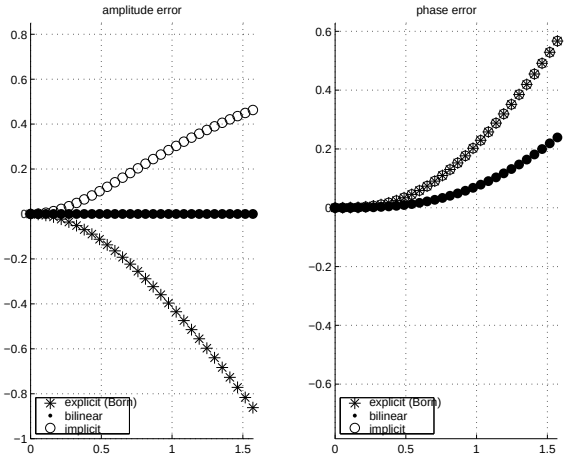



Figure 2: Amplitude and phase errors for the explicit, bilinear and implicit approximations. paul1-exap [NR]

If, for notation simplicity, we define

$$\beta = -i \left. \frac{dk_z}{ds} \right|_{s=s_r} \Delta z, \tag{12}$$

the WEMVA equation (7) can be written as

$$\Delta \mathcal{W} = \mathcal{W}_b \left[ e^{\beta \Delta s} - 1 \right], \tag{13}$$

and so the linearizations corresponding to the explicit, bilinear and implicit solutions respectively become

$$\begin{aligned} \Delta \mathcal{W} &\approx \mathcal{W}_b \beta \Delta s \\ &\approx \mathcal{W}_b \frac{2\beta \Delta s}{2 - \beta \Delta s} \\ &\approx \mathcal{W}_b \frac{\beta \Delta s}{1 - \beta \Delta s} \end{aligned} \tag{14}$$

Aparently, just the first equation in (14) provides a linear relationship between $\Delta \mathcal{W}$ and $\Delta s$. However, a simple re-arrangement of terms leads to

$$\begin{aligned} \Delta \mathcal{W} &\approx \beta \left[ \mathcal{W}_b \right] \Delta s \\ &\approx \beta \left[ \mathcal{W}_b + \frac{1}{2} \Delta \mathcal{W} \right] \Delta s \\ &\approx \beta \left[ \mathcal{W}_b + \Delta \mathcal{W} \right] \Delta s. \end{aligned} \tag{15}$$

For MVA, both the background ($\mathcal{W}_b$) and perturbation wavefields ($\Delta \mathcal{W}$) are known, so it is not a problem to incorporate them in the linear operator. In any of the cases described in Equation (15), the approximations can be symbolically written using the fitting goal

$$\mathbf{d} \approx \mathbf{Lm}, \tag{16}$$

where the data **d** is the wavefield perturbation, and the model **m** is the slowness perturbation. The same operator **L** is used for inversion in all situations, the only change being in the wavefield that is fed into the linear operator. Therefore, the new operators are not more expensive than the Born operator.

All linear relationships in Equation (15) belong to a family of approximations of the general form

$$\Delta \mathcal{W} \quad \approx \quad \beta \left[ \mathcal{W}_b + \xi \Delta \mathcal{W} \right] \Delta s. \tag{17}$$

The various approximations can be obtained using appropriate values for the parameter $\xi = 0 \ldots 1$. All forms of Equation (17), however, are approximations to the exact non-linear relation (13), therefore they are all likely to break for large values of the phase, or equivalently large values of the slowness perturbation or frequency. Nevertheless, these approximations enable us to achieve higher accuray in slowness estimation as compared to the simple Born approximation.

An interesting comparison can be made between the extreme members of the sequence given by Equation (17): for $\xi = 0$ we use the background wavefield $\mathcal{W}_b$, and for $\xi = 1$ we use the full wavefield $\mathcal{W} = \mathcal{W}_b + \Delta \mathcal{W}$. The physics of scattering would recommend that we use the later form, since the scattered wavefield ($\Delta \mathcal{W}$) is generated by the total wavefield ($\mathcal{W}$), and not by an approximation of it ($\mathcal{W}_b$), thus naturally accounting for multiple scattering effects. The later situation also corresponds to what is known in the scattering literature as *wavefield renormalization* (Wang, 1997). The details of these ideas and their implications remain open for future research.

Finally, we note that Equation (17) cannot be used for forward modeling of the wavefield perturbations $\Delta \mathcal{W}$, except for the particular case $\xi = 0$, since the output quantity is contained in the operator itself. However, we can use this equation for inversion for any choice of the parameter $\xi$.

## NEWTON'S METHOD AND WEMVA

One can also consider the problem of estimating the slowness field from wavefields using WEMVA in the general non-linear inversion framework.

In particular, if $\mathcal{W}(z + \Delta z)$ is the upgpoing wavefield at the bottom of a layer and $\mathcal{W}(z)$ is the upgoing wavefield at the top of the layer, the layer slowness $s$ is constrained by the nonlinear equation

$$F[s] = P[s]\, \mathcal{W}(z + \Delta z) - \mathcal{W}(z) = 0\,, \tag{18}$$

where $P[s] = e^{i\, k_z[s]}$ is the wave propagation operator.

The Newton method applied to equation (18) amounts to inversion of the linear system

$$F'[s_k](s_{k+1} - s_k) = -F[s_k]\,, \tag{19}$$

where $k$ is the nonlinear iteration counter (the iteration starts with some a priori slowness model $s_0$), and $F'[s]$ is the Fréchet derivative of the wave propagation operator. Since $F[s]$ is complex-valued, we can multiply both sides of system (19) by the adjoint (complex-conjugate) operator $F'[s_k]^T$ to obtain the purely real system

$$\begin{aligned} F'[s_k]^T\, F'[s_k](s_{k+1}-s_k) &= -F'[s_k]^T\, F[s_k] \\ \left(R'[s_k]^T\, R'[s_k]+I'[s_k]^T\, I'[s_k]\right)\,(s_{k+1}-s_k) &= -\left(R'[s_k]^T\, R[s_k]+I'[s_k]^T\, I[s_k]\right)\,, \end{aligned} \tag{20}$$

where $R[s]$ and $I[s]$ are the real and imaginary parts of $F[s]$. Algorithm (21) is equivalent to the Gauss-Newton method applied to the least-squares solution of

$$R[s] \approx 0\,, \tag{21}$$

$$I[s] \approx 0\,. \tag{22}$$

It is well-known that the Newton and Newton-Gauss methods possess fast convergence provided that the original estimate $s_0$ is sufficiently close to the solution. They may diverge otherwise. To guarantee convergence, the norm (spectral radius) of the Fréchet derivative $G'[s]$ for the operator

$$G[s] = s - \frac{R'[s]^T\, R[s]+I'[s]^T\, I[s]}{R'[s]^T\, R'[s]+I'[s]^T\, I'[s]} \tag{23}$$

must be strictly smaller than one in the vicinity of the solution that contains the starting value $s_0$. Convergence follows then from the contraction mapping theorem. The speed of convergence is higher for smaller norms.

It is important to realize that modifying the original nonlinear Equation (18) may change the convergence behavior and lead to faster convergence and wider convergence area. A particularly meaningful way to modify Equation (18) is to multiply it by $P[s]^{-\xi}$, where $\xi$ is a scalar between 0 and 1. The modified equation takes the form

$$F_\xi[s] = P[s]^{-\xi}\, F[s] = P[s]^{1-\xi}\, \mathcal{W}(z+\Delta z) - P[s]^{-\xi}\, \mathcal{W}(z) = 0\,. \tag{24}$$

The case of $\xi = 0$ corresponds to the original system. Its linearization with the Newton method leads to the Born approximation. Analogously, the case of $\xi = 1$ corresponds to the implicit method: the two wavefields are compared at the bottom of the layer rather than at the top. The case of $\xi = 1/2$ leads to the bilinear method: both wavefields are continued to the middle of the layer for comparison. Many other intermediate results are possible,

**Example**

The simplest case to study analytically is that of vertically-incident waves in laterally homogeneous media. In this case, all operators become functions of the scalar variable $s$ (unknown layer slowness). If, for a particular temporal frequency $\omega$ and the layer thickness $\Delta z$, we measure the slowness in units of $1/(\omega\,\Delta z)$, the wave continuation operator is simply the phase shift

$$P(s) = e^{i\,s}\,, \tag{25}$$

and the fundamental nonlinear equation takes the form

$$F_\xi(s) = \mathcal{W}(z+\Delta z)\, e^{i\,(1-\xi)s} - \mathcal{W}(z)\, e^{-i\,\xi\,s} = 0\,. \tag{26}$$

Noting that

$$\mathcal{W}(z) = \mathcal{W}(z+\Delta z)\, e^{i\,s^\star}\,, \tag{27}$$

where $s^\star$ is the true slowness, and that the convergence of Newton's method does not depend on scaling the equation by a constant, we can modify equation (26) to the simpler form

$$\hat{F}_\xi(s) = F_\xi(s)\,\frac{e^{(\xi-1)s^\star}}{\mathcal{W}(z+\Delta z)} = e^{i\,(1-\xi)\,\Delta s} - e^{-i\,\xi\,\Delta s} = 0\,, \tag{28}$$

where $\Delta s = s - s^\star$. The obvious solution of Equation (28) is $\Delta s = 0$. Our task is to find the convergence limits and their dependence on $\xi$.

After a number of algebraic and trigonometric simplifications, the operator $G$ from equation (23) takes the form of the function

$$\hat{G}_\xi(s) = s - \frac{\sin(\Delta s)}{1-2(1-\xi)\xi+2(1-\xi)\xi\,\cos(\Delta s)} \tag{29}$$

Its derivative is

$$\hat{G}'_\xi(s) = 2\,\sin^2\left(\frac{\Delta s}{2}\right)\,\frac{1-2(1-x)x\,[3-2(1-x)x]-4(1-x)^2x^2\,\cos(\Delta s)}{[1-2(1-\xi)\xi+2(1-\xi)\xi\,\cos(\Delta s)]^2}\,. \tag{30}$$

The method will converge in the region around $\Delta s = 0$, where the absolute value of $\hat{G}'_\xi(s)$ is strictly smaller than one. This region (as a function of $\Delta s$ and $\xi$) is plotted in Figure 3. We can see that the convergence region has a finite extent. Its width is the same for $\xi = 0$, $\xi = 1$, and $\xi = 1/2$. Indeed,

$$\hat{G}'_0(s) = \hat{G}'_1(s) = 2\,\sin^2\left(\frac{\Delta s}{2}\right) \tag{31}$$

and

$$\hat{G}'_{1/2}(s) = -\tan^2\left(\frac{\Delta s}{2}\right)\,. \tag{32}$$

In both cases, the absolute value of the derivative is smaller than one if $\Delta s < \frac{\pi}{2}$. If we take $\omega = 2\pi \times 100\,\mathrm{Hz}$ and $\Delta z = 10\,\mathrm{m}$, then the convergence radius is $\Delta s = 0.25\,\mathrm{s/km}$. At small $\Delta z$,

$$\hat{G}'_0(s) = \hat{G}'_1(s) \approx \frac{(\Delta s)^2}{2} \tag{33}$$

and

$$\hat{G}'_{1/2}(s) \approx \frac{(\Delta s)^2}{4}\,. \tag{34}$$

The convergence rate is of the same order (cubic) but faster in the case of the bilinear method ($\xi = 1/2$), because of the twice smaller constant. Here is an example of iterations starting with $s_0 = 2$ and converging to $s^\star = \pi$. The Born iteration:

$$\begin{aligned} s_1 &= 2.9093 \\ s_2 &= 3.13951 \\ s_3 &= 3.14159 \end{aligned}$$

The bilinear iteration:

$$\begin{aligned} s_1 &= 3.28419 \\ s_2 &= 3.14135 \\ s_3 &= 3.14159 \end{aligned}$$

A faster convergence can be achieved at some other values of $\xi$. Examining the Taylor series of $\hat{G}'_\xi(s)$ around $\Delta s = 0$:

$$\hat{G}'_\xi(s) \approx [1 - 6(1-\xi)\xi] \, \frac{(\Delta s)^2}{2} \, , \tag{35}$$

we find that the order of convergence is optimized for $\xi = \frac{1}{2} \pm \frac{\sqrt{3}}{6}$. In this case,

$$\hat{G}'_{1/2 \pm \sqrt{3}/6}(s) = \frac{4 \sin^4\left(\frac{\Delta s}{2}\right)}{[2 + \cos(\Delta s)]^2} \approx \frac{(\Delta s)^4}{36} \, , \tag{36}$$

and the convergence is fifth order! The example iterations with the optimal value of $\xi$ are:

$$\begin{aligned} s_1 &= 3.12903 \\ s_2 &= 3.14159 \\ s_3 &= 3.14159 \end{aligned}$$

The radius of convergence with the optimal value of $\xi$ is $\Delta s < \frac{2}{3}\pi$.

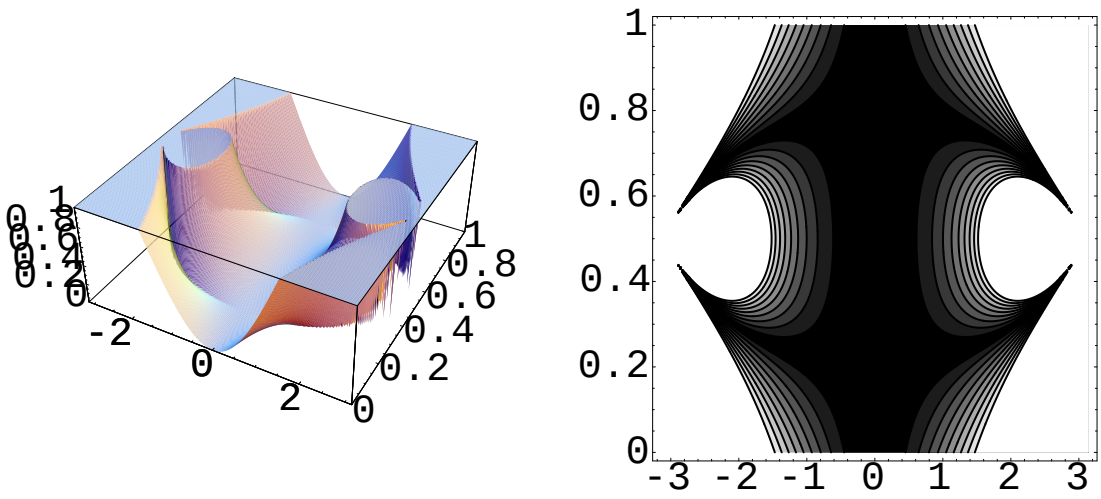



Figure 3: Convergence region for the Newton-Gauss method in the vertical plane-wave example. Left: 3-D projection. Right: contours. The non-white region on the right plot corresponds to the convergence area. Horizontal axis: $\Delta s$. Vertical axis: $\xi$. paul1-zo [NR]

Of course, this analysis does not apply directly to the case of non-vertical wave propagation and laterally inhomogenous slowness fields. For reflection wavefields at multiple offsets,

the symmetry between downward and upward continuation is broken, as is clear from the experimental results of this paper. However, the simple analysis points to the potential benefits of modifying the Born approximation in the wave-equation velocity estimation.

## PHYSICAL INTERPRETATION

This section presents a brief physical interpretation of the various members of relation (17).

Consider that we have recorded two wavefields at the top and bottom of a depth slab: $W_0$, the wavefield at the top of the slab which has not propagated through the anomalous region; $W_1$, the wavefield at the bottom of the slab which incorporates scattering effects caused by the slowness anomaly inside the slab. The goal of WEMVA is to extract the slowness perturbation $\Delta s$ from $W_0$ and $W_1$.

We can imagine that the linearized process can be thought of as a succession of four steps.

1. Continuation of $W_0$ and $W_1$ to a level inside the slab where we can compare the two wavefields. This level can be either at the top, bottom or anywhere in between:

$$W_1 e^{-\xi k_z(s)} = W_0 e^{(1-\xi)k_z(s)} \tag{37}$$

   $k_z$ represents the depth wavenumber and is a function of the arbitrary slowness inside any given depth slab, and $\xi = 0 \ldots 1$ is a scalar which defines where inside the slab we continue the two wavefields.

2. Linearization of $W_0$ and $W_1$ with respect to the slowness perturbation:

$$W_1 e^{-\xi k_z(s_0)}[1-\xi\beta\Delta s] = W_0 e^{(1-\xi)k_z(s_0)}[1+(1-\xi)\beta\Delta s], \tag{38}$$

   where $\beta$ is the function defined in Equation (12).

3. Datuming of the linearized wavefields to the bottom of the slab:

$$\begin{aligned} W_1[1-\xi\beta\Delta s] &= W_0[1+(1-\xi)\beta\Delta s]e^{k_z(s_0)} && (39)\\ &= W_b[1+(1-\xi)\beta\Delta s] && (40)\end{aligned}$$

4. Subtraction of the wavefield propagated through the perturbed medium from the wavefield propagated through the background medium:

$$\Delta W = W_1 - W_b = W_b \frac{\beta\Delta s}{1-\xi\beta\Delta s} \tag{41}$$

All three cases in Equation (14) can be derived from Equation (41) as follows:

$$\begin{aligned} \xi = 0 &\rightarrow \Delta W = W_b\beta\Delta s && (42)\\ \xi = \frac{1}{2} &\rightarrow \Delta W = W_b\frac{2\beta\Delta s}{2-\beta\Delta s} && (43)\\ \xi = 1 &\rightarrow \Delta W = W_b\frac{\beta\Delta s}{1-\beta\Delta s}. && (44)\end{aligned}$$

## EXAMPLES

We demonstrate the technique outlined in the preceding sections using a synthetic example. The model (Figure 4) consists of a body of high velocity incorporated in a background with strong but smooth lateral velocity variation.

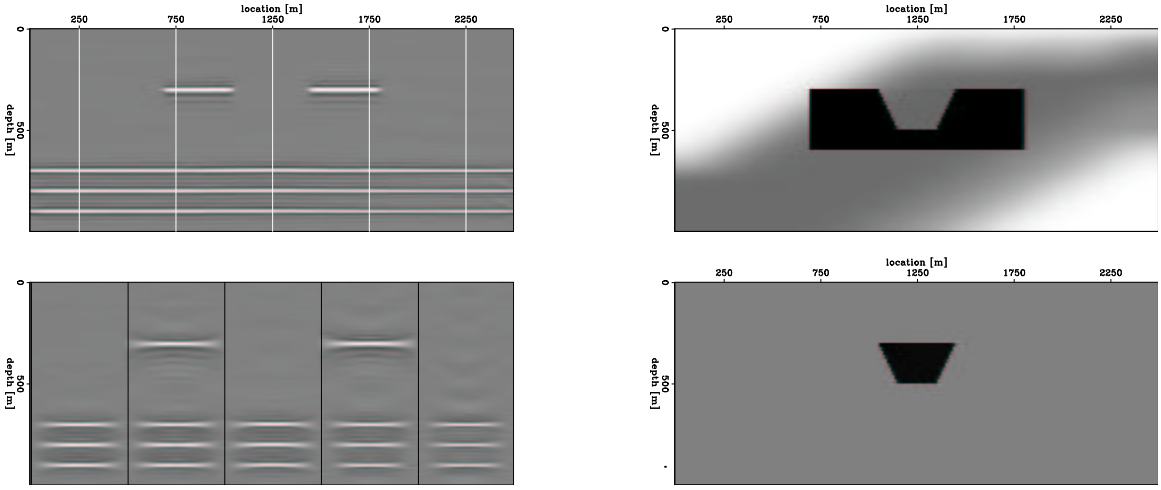



Figure 4: Synthetic model. Reflectivity model (top left) and a few angle-gathers corresponding to the vertical grid in the upper plot (bottom left). Background slowness model (top right) and slowness perturbation (bottom right). paul1-model [CR,M]

Our examples show the results of inversion for a regularized problem symbolically sumarized by the fitting goals:

$$\begin{aligned} \Delta R &\approx \mathbf{L}\Delta s \\ 0 &\approx \mathbf{A}\Delta s, \end{aligned} \tag{45}$$

where $\Delta R$ is the image perturbation, $\Delta s$ is the corresponding slowness perturbation, $\mathbf{L}$ is one of the linearized WEMVA operators and $\mathbf{A}$ is a roughening operator, an isotropic Laplacian for our examples. After preconditioning (Claerbout, 1999), our fitting goals become

$$\begin{aligned} \Delta R &\approx \mathbf{L}\mathbf{A}^{-1}\Delta p, \\ 0 &\approx \Delta p \end{aligned} \tag{46}$$

where $\Delta p$ represents the preconditioned $\Delta s$.

We also note that since the operator $\mathbf{L}$ is large, similar in size to a migration operator, we cannot implement it in-core, and therefore we have to use out-of-core optimization (Sava, 2001).

For our experiments, we generate two kinds of image perturbations.

- The first kind is created from a given slowness perturbation $\Delta s$ using the linear operator in Equation (15). We refer to this type of image perturbation as *linear*, since it

corresponds to the linearized Born operator. This type of image perturbation cannot be obtained in real applications, but serves as a reference when we investigate the Born approximation.

- The second kind is created by taking the difference of two images created using two slowness models ($\Delta R = R - R_o$). We refer to this type of image perturbation as *non-linear*, since it corresponds to the non-linear relation in Equation (7).

We analyze several examples where we change the magnitude of the slowness anomaly, but not its shape. We choose to test various magnitudes for the anomaly from 1% to 50% of the background slowness.

Figures 5, 7, 9, 11 show the image perturbations created by the slowness anomalies for the various levels of perturbation. In each figure, the left panels present the linear case, and the right panels the non-linear case. The top panels depict the stacked sections, and the bottom panels a few representative image gathers in the angle-domain (Sava and Fomel, 2000) corresponding to the locations of the vertical lines in the upper panels. For small values of the slowness perturbations, the two images should be similar, but for larger values we should see the image perturbation reaching and eventually breaking the Born approximation.

Figures 6, 8, 10, 12 present the results of inversion of the non-linear $\Delta R$ using the three WEMVA operators presented in the preceding section: the explicit (Born) operator (top), the bilinear operator (middle), and the implicit operator (bottom).

For the case of the small slowness perturbation (1%), the linear and non-linear image perturbations are very similar, as seen in Figure (5). The corresponding slowness anomaly obtained by inversion is well focused, confirming that, for this case, even the Born approximation is satisfactory, as suggested by the theory.

The larger anomaly of 5% of the background slowness shows the serious signs of breakdown for the Born approximation. For the case of the even larger slowness perturbation (20%), the linear and non-linear image perturbations are not that similar anymore, indicating that we have already violated the limits of the Born approximation (Figure 9). Consequently, the inversion from the non-linear image perturbation using the Born operator blows-up. However, the WEMVA operators employing the bilinear and implicit approximations are still well-behaved, although the shape of the anomaly is slightly modified.

The case of the largest slowness anomaly (40%), bring us closer to the limits of both the bilinear and implicit approximations. Although neither has blown-up yet, the shape of the anomalies is somewhat altered.

## CONCLUSIONS

In this paper, we investigate the limits of the Born approximation when applied to wave-equation migration velocity analysis. Experimentally we find that the Born approximation is only valid for small slowness anomalies, on the order of $1-2\%$ of the background slowness

for an anomaly of the shape and size used in the example in this paper. These numbers, however, are model dependent, because we have to consider both the magnitude and size of the anomaly: a small anomaly of large magnitude can have a similar effect as a large anomaly of small magnitude.

Moving beyond the Born approximation involves one of the two solutions: we can either artificially create image perturbations that are compliant with this approximation (Sava and Biondi, 2001), or we can improve the WEMVA operator to better handle the non-linearity in the image perturbations, as presented in this paper. [2]

We propose two improved versions of the WEMVA operator which are more appropriate for the case of large/strong slowness anomalies. Our new operators involve linearizations using bilinear and implicit approximations to the exponential function. With the new operators, we not only improve accuracy but we also maintain stability of the inversion scheme at much higher values of the slowness anomalies, even in the order of 25% of the background.

Finally, we note that our new operators come at a cost which is practically no different than the cost of the Born-linearized operator, while improving its accuracy and stability.

## REFERENCES

Biondi, B., and Sava, P., 1999, Wave-equation migration velocity analysis: SEP–**100**, 11–34.

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

Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project, `http://sepwww.stanford.edu/sep/prof/`.

Clapp, R. G., 2001, Geologically constrained migration velocity analysis: Ph.D. thesis, Stanford University.

Liu, W., Popovici, A., Bevc, D., and Biondi, B., 2001, 3-D migration velocity analysis for common image gathers in the reflection angle domain: 71st Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 885–888.

Mosher, C., Jin, S., and Foster, D., 2001, Migration velocity analysis using angle image gathers: 71st Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 889–892.

Sava, P., and Biondi, B., 2000, Wave-equation migration velocity analysis: Episode II: SEP–**103**, 19–47.

Sava, P., and Biondi, B., 2001, Born-compliant image perturbation for wave-equation migration velocity analysis: SEP–**110**, 91–102.

[2] It appears that another possibility would be to use the Rytov approximation, and establish a linear relation between the slowness and phase perturbations. This method, however, is somewhat more problematic, since it may involve phase unwrapping of complex wavefields. It remains, however, an interesting area for future research.

Sava, P., and Fomel, S., 2000, Angle-gathers by Fourier Transform: SEP–**103**, 119–130.

Sava, P., 2000, A tutorial on mixed-domain wave-equation migration and migration velocity analysis: SEP–**105**, 139–156.

Sava, P., 2001, `oclib`: An out-of-core optimization library: SEP–**108**, 199–224.

Wang, G. Y., 1997, Wave scattering and diffraction tomography in complex media: Ph.D. thesis, Stanford University.

Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, no. 01, 15–26.

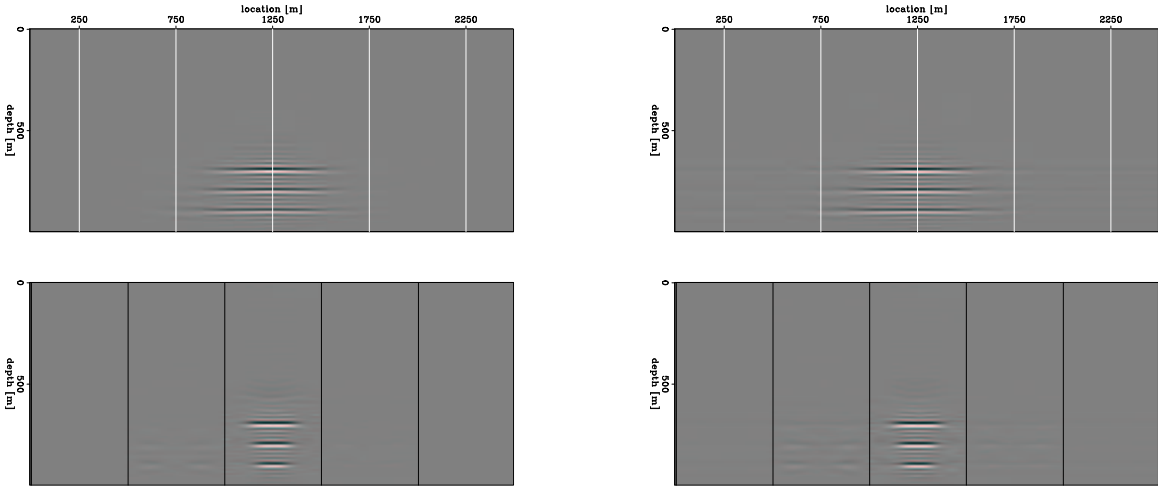



Figure 5: Anomaly of 1%: linear and non-linear image perturbations (left/right); zero offset section (top) and selected angle-gathers (bottom) corresponding to the locations of the vertical lines in the upper panel. Large differences between the linear and non-linear image perturbations indicate situations in which we violate the Born approximation. paul1-01.perturbation [CR,M]

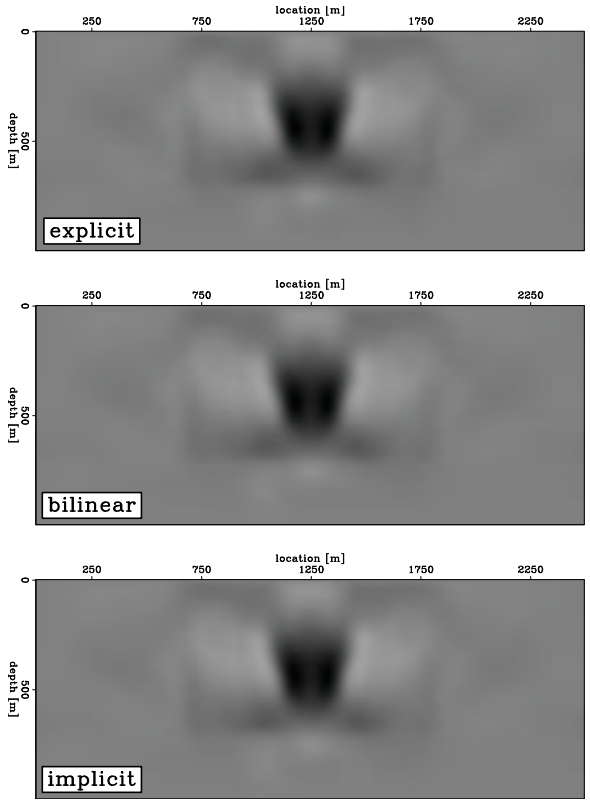



Figure 6: Anomaly of 1%: inversion from the non-linear image perturbation (7) using the explicit (top), bilinear (middle) and implicit (bottom) WEMVA operators. paul1-01.inversion [CR,M]

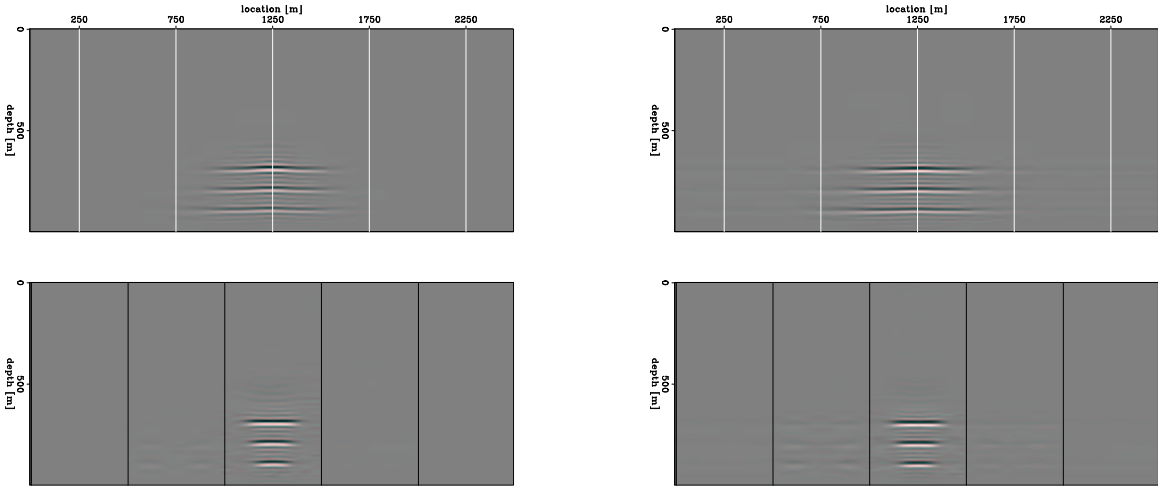



Figure 7: Anomaly of 5%: linear and non-linear image perturbations (left/right); zero offset section (top) and selected angle-gathers (bottom) corresponding to the locations of the vertical lines in the upper panel. Large differences between the linear and non-linear image perturbations indicate situations in which we violate the Born approximation. paul1-05.perturbation [CR,M]

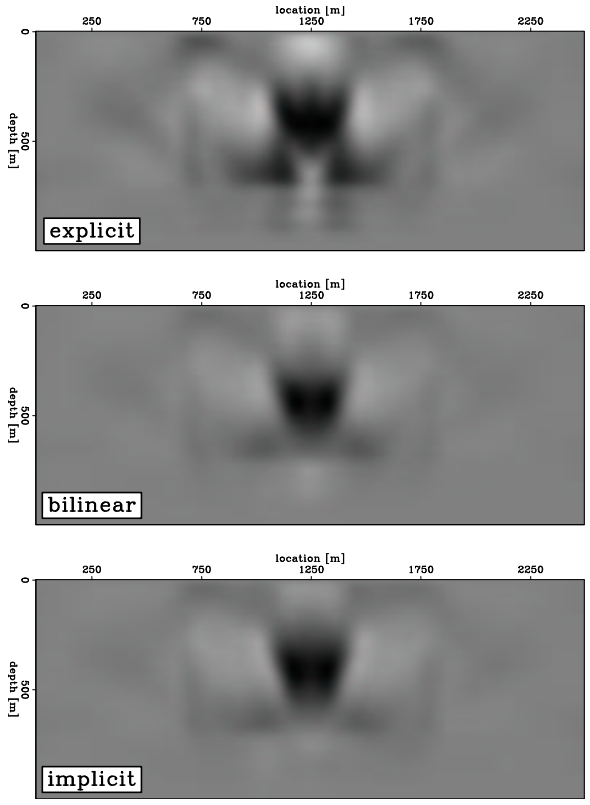



Figure 8: Anomaly of 5%: inversion from the non-linear image perturbation (7) using the explicit (top), bilinear (middle) and implicit (bottom) WEMVA operators. paul1-05.inversion [CR,M]

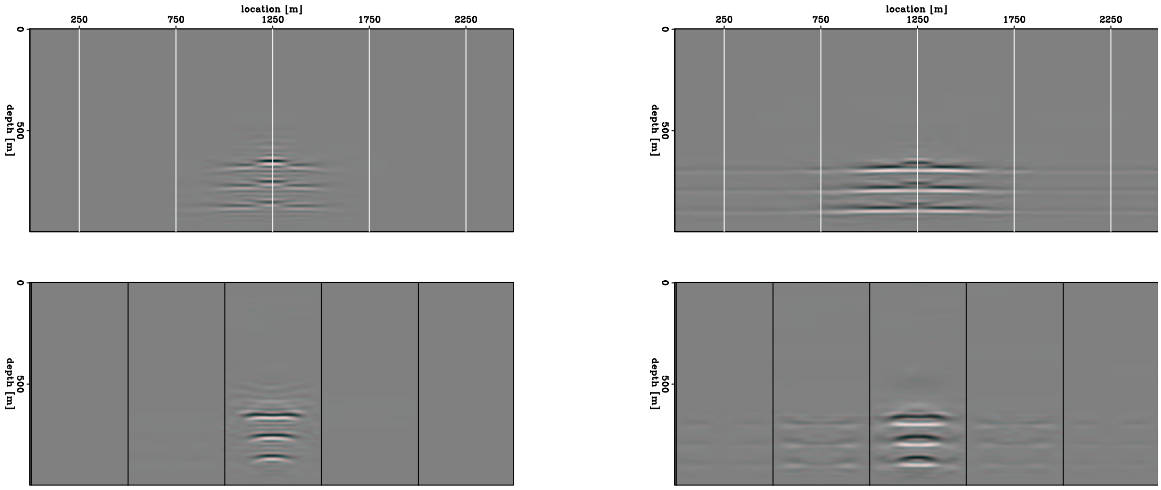



Figure 9: Anomaly of 20%: linear and non-linear image perturbations (left/right); zero offset section (top) and selected angle-gathers (bottom) corresponding to the locations of the vertical lines in the upper panel. Large differences between the linear and non-linear image perturbations indicate situations in which we violate the Born approximation. paul1-20.perturbation [CR,M]

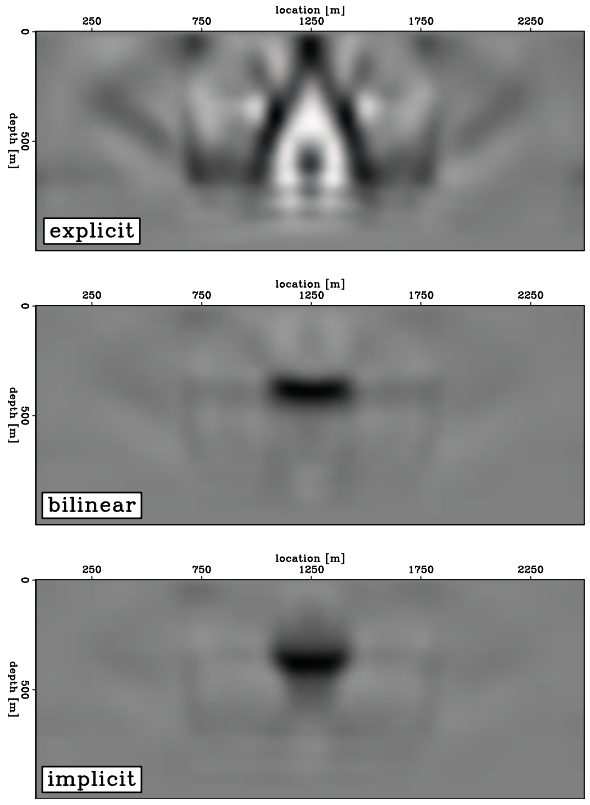



Figure 10: Anomaly of 20%: inversion from the non-linear image perturbation (7) using the explicit (top), bilinear (middle) and implicit (bottom) WEMVA operators. paul1-20.inversion [CR,M]

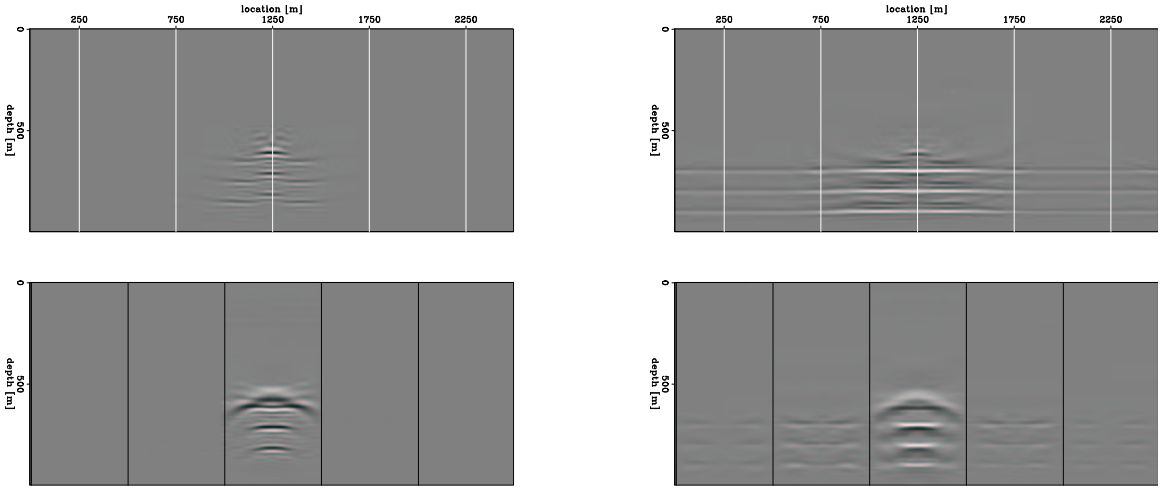



Figure 11: Anomaly of 40%: linear and non-linear image perturbations (left/right); zero offset section (top) and selected angle-gathers (bottom) corresponding to the locations of the vertical lines in the upper panel. Large differences between the linear and non-linear image perturbations indicate situations in which we violate the Born approximation. paul1-40.perturbation [CR,M]

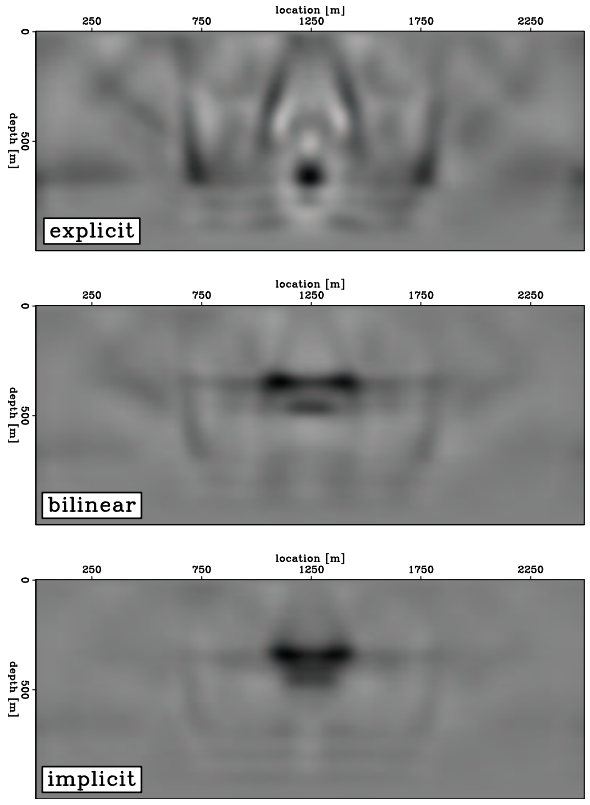



Figure 12: Anomaly of 40%: inversion from the non-linear image perturbation (7) using the explicit (top), bilinear (middle) and implicit (bottom) WEMVA operators. paul1-40.inversion [CR,M]

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 99–107*

# Short Note

# Matching dips in velocity estimation

*Robert G. Clapp*[1]

## INTRODUCTION

Accurate velocity estimation is essential to obtain a good migrated image and accurate resevoir attributes (Claerbout, 1999). The problem is that tomographic velocity estimation is an underdetermined problem. We can reduce the null space of the tomographic process by adding additional constraints, or more accurate goals, to the estimation. In early work (Clapp et al., 1997; Clapp and Biondi, 1999; Clapp, 2001b,a) I discussed one such constraint: encouraging velocity follows dip. Often we have an added constraint; although we may be unsure of reflector position (due to anisotropy, etc.) or we may have a good estimate of reflector dip (either from well logs, geologic models, etc). By incorporating this information into the inversion we can better constrain the inversion process. This method is tested on a fairly complicated synthetic dataset.

## THEORY

Tomography is a non-linear problem that we linearize around an initial slowness model. In this discussion I will be talking about the specific case of ray based tomography but most of the discussion is valid for other tomographic operators. We can linearize the problem around an initial slowness model and obtain a linear relation $\mathbf{T}$ between the change in travel times $\Delta\mathbf{t}$ and change in slowness $\Delta\mathbf{s}$ and reflector position $\Delta\mathbf{r}$. We break up our tomography operator into its two parts, changes due to slowness along the ray $\mathbf{T_{ray}}$ and changes due to reflector movement $\mathbf{T_{ref}}$:

$$\Delta\mathbf{t} \approx \mathbf{T_{ray}}\Delta\mathbf{s} - \mathbf{T_{ref}}\Delta\mathbf{r}. \tag{1}$$

Inverting for both $\Delta\mathbf{s}$ and $\Delta\mathbf{r}$ is an unstable process. We can improve stability by introducing another operator $\mathbf{H}$ which maps slowness changes to reflector changes,

$$\Delta\mathbf{t} \approx \mathbf{T_{ray}}\Delta\mathbf{s} - \mathbf{T_{ref}}\mathbf{H}\Delta\mathbf{s}. \tag{2}$$

We can approximate the change in travel time due to a change in reflector movement by

$$\mathbf{T_{ref}} = \mathbf{v_{ref}}\cos\theta\cos\phi, \tag{3}$$

[1]**email:** bob@sep.stanford.edu

where $\mathbf{v_{ref}}$ is the velocity at the reflector, $\phi$ is the reflector dip, and $\theta$ is the reflection angle (Stork, 1992).

We can approximate the change in reflector position due to a change in slowness by assuming movement normal to the reflector and integrating along the normal ray,

$$\mathbf{H} = \int_{\mathbf{ray}} \mathbf{dl}. \tag{4}$$

If we note that the travel time of the normal ray is independent of velocity we can write

$$\begin{aligned} t_0 &= t_1 \\ r_0 s_0 &= (r_0 + \delta r)(s_0 + \delta s) \\ 0 &= \delta r s_0 + \delta r \delta s + r_0 \delta s, \end{aligned} \tag{5}$$

where $t_0$ is the travel time in the initial model and $t_1$ is the travel time through the new model. If we ignore the second order term,

$$\delta r \approx -\frac{r_0}{s_0} \delta s. \tag{6}$$

The reason for this review is that our mapping of slowness change to reflector movement leads to a way to approximate reflector dip in the post-tomographic domain.

For simplicity let's concern ourselves with the 2-D problem, though it's easily extendible to 3-D. Imagine that $\theta_k$ represents our *a priori* reflector dip, $\mathbf{D}$ is a derivative operator, $\mathbf{r}$ is our final reflector dip, $\mathbf{r_0}$ is the initial reflector position, and $\Delta\mathbf{r}$ is our change in reflector position. We can derive a fairly simple fitting goal relating reflector dip and $\Delta\mathbf{s}$,

$$\begin{aligned} \theta_{\mathbf{k}} &\approx \mathbf{Dr} \\ \theta_{\mathbf{k}} &\approx \mathbf{D}(\mathbf{r_0} + \Delta\mathbf{r}) \\ \theta_{\mathbf{k}} - \mathbf{Dr_0} &\approx \mathbf{D}\Delta\mathbf{r} \\ \theta_{\mathbf{k}} - \mathbf{Dr_0} &\approx \mathbf{DH}\Delta\mathbf{s}. \end{aligned} \tag{7}$$

If we combine this new fitting goal with our tomographic fitting goal and our regularization fitting goal we get,

$$\begin{aligned} \Delta\mathbf{t} &\approx \mathbf{T}\Delta\mathbf{s} \\ \mathbf{0} &\approx \epsilon_0 \mathbf{A}\Delta\mathbf{s} \\ \theta_{\mathbf{k}} - \mathbf{Dr_0} &\approx \epsilon_1 \mathbf{DH}\Delta\mathbf{s}. \end{aligned} \tag{8}$$

## EXAMPLE

To test the methodology I decided to use a synthetic 2-D dataset generated by BP based on a typical North Sea environment, Figure 1. To avoid tomography's problem with sharp velocity

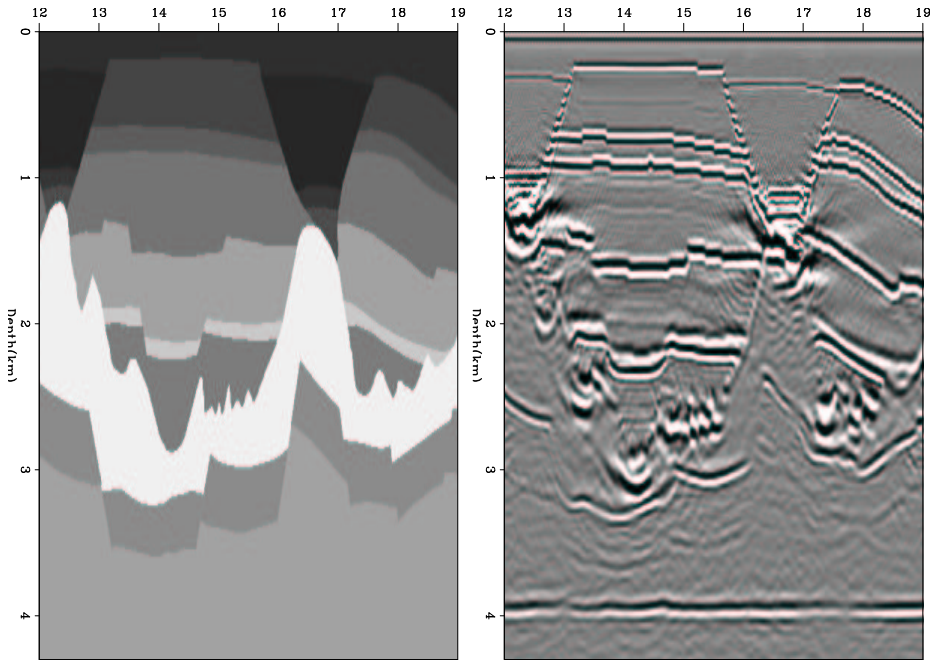



Figure 1: The left panel shows the correct velocity model. The right panel shows the result of migrating with this velocity model. bob5-amoco-vel-cor [CR,M]

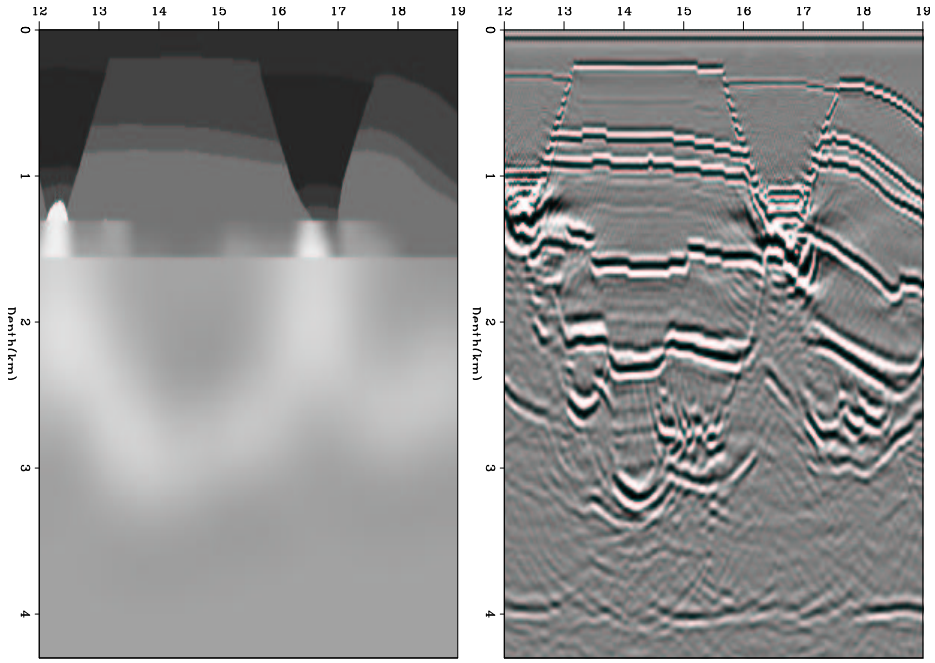



Figure 2: The left panel shows the initial velocity model. The right panel shows the result of migrating with this velocity model. bob5-amoco-vel0 [CR]

contrasts I chose to assume an accurate knowledge of the velocity structure down to 1.8 km. For the remaining initial velocity structure I smoothed the correct velocity. Figure 2 shows the initial velocity model and initial migration. I then performed two different series of tomography loops. In the first case I used a standard approach, without the constraint on dip of the basement reflector at $4km$. Figure 3 shows the initial migration with my pick of the reflector position overlaid ($\mathbf{r_0}$ in fitting goals (8)). Figures 4 and 5 show the velocity and migration result after a single non-linear iteration of tomography using both approaches. In the first iteration the velocity structure looks somewhat more accurate without the dip constraint. The image tells a different story. Note how the bottom reflector is much flatter using the dip constraint condition (Figure 5) and the overall image positioning is a little better. After four iterations, we see a more dramatic difference. Without the dip constraint condition (Figure 6) the velocity model is having trouble converging, especially along the right edge. The bottom reflector is quite discontinuous and misplaced. The overall image quality is disappointing. With the dip constraining condition (Figure 7) the velocity model is correctly finding the salt boundaries. The bottom reflector is fairly flat, consistent, and well positioned. The overall image quality is better than the result without the dip constraint.

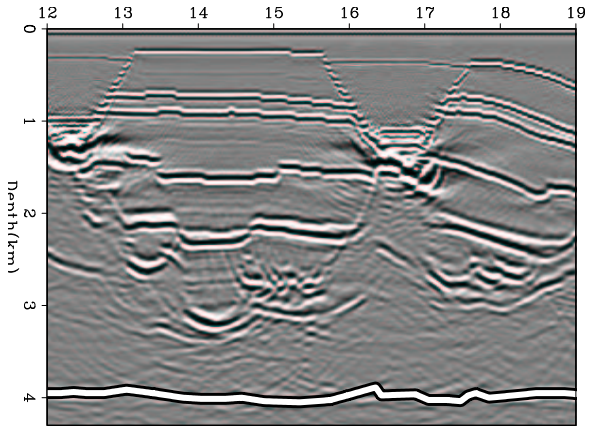



Figure 3: The initial migrated model overlaid by the picked initial reflector position. bob5-picked [CR]

## CONCLUSIONS

The proposed method for constraining reflector dip in tomography worked well on the complex North Sea synthetic. The estimated velocity model was more accurate with the dip constraint. The migrated image showed overall better image quality and the selected reflector was more continuous and better positioned.

## ACKNOWLEDGMENTS

I would like to thank BP for the synthetic data used in this paper.

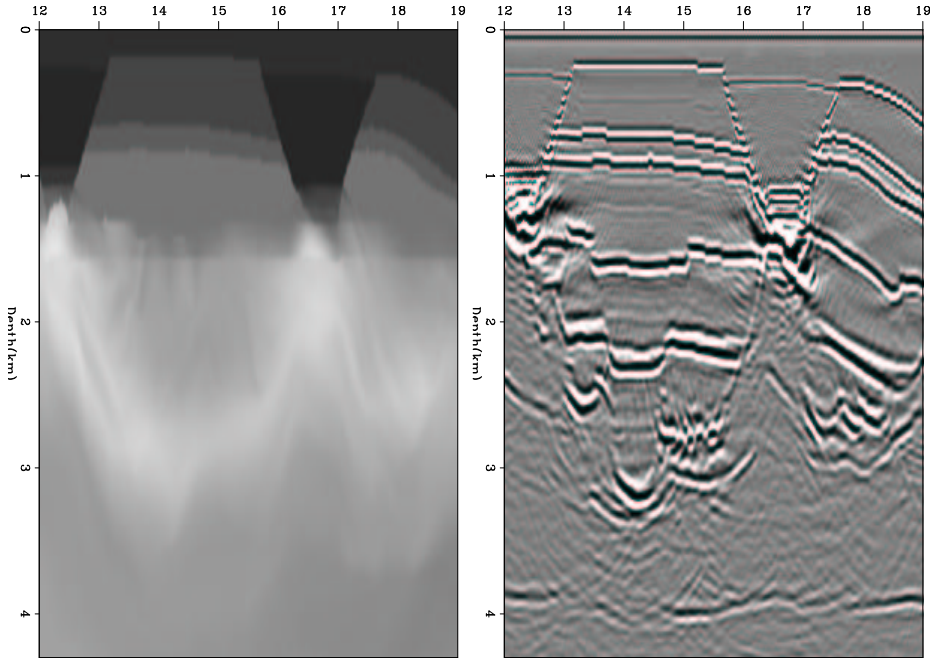



Figure 4: The left panel shows the velocity model after one iteration of 'conventional' tomography. The right panel shows the result of migrating with the velocity model in the left panel. bob5-amoco-vel1.steer [CR,M]

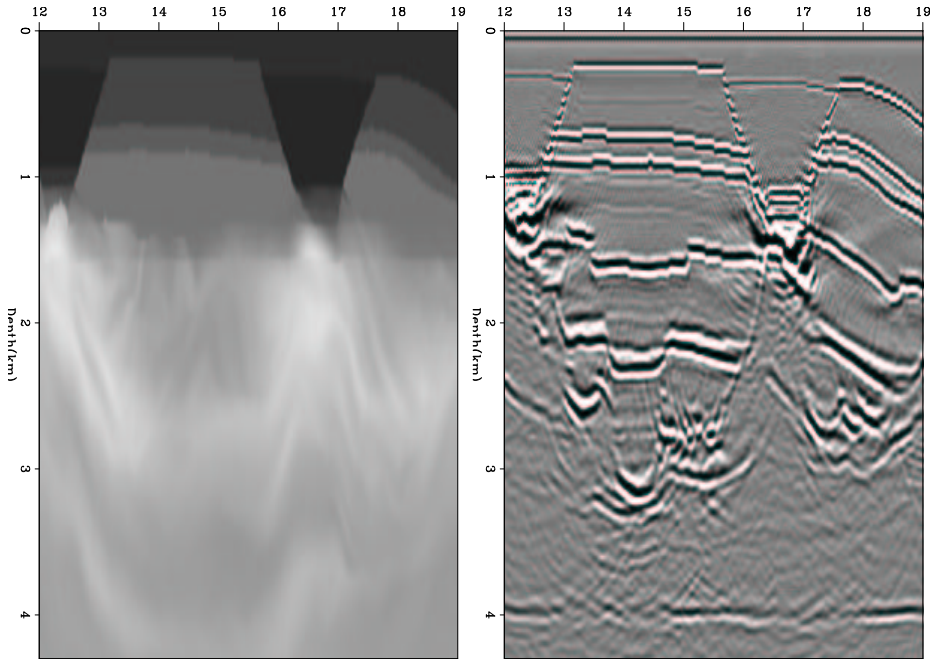



Figure 5: The left panel shows the velocity model after one iteration of tomography with a dip constraint. The right panel shows the result of migrating with the velocity model in the left panel. Note the more continuous nature of the bottom reflector (compared to Figure 4. bob5-amoco-vel1.steer-ref [CR,M]

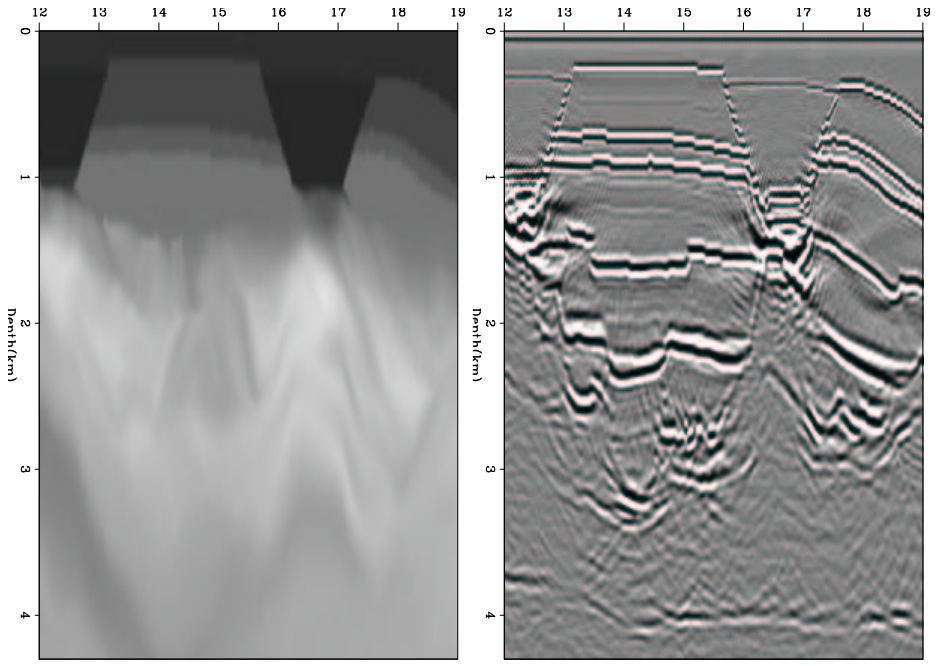



Figure 6: The left panel shows the velocity model after four iteration of 'conventional' tomography. The right panel shows the result of migrating with the velocity model in the left panel. bob5-amoco-vel4.steer [CR,M]

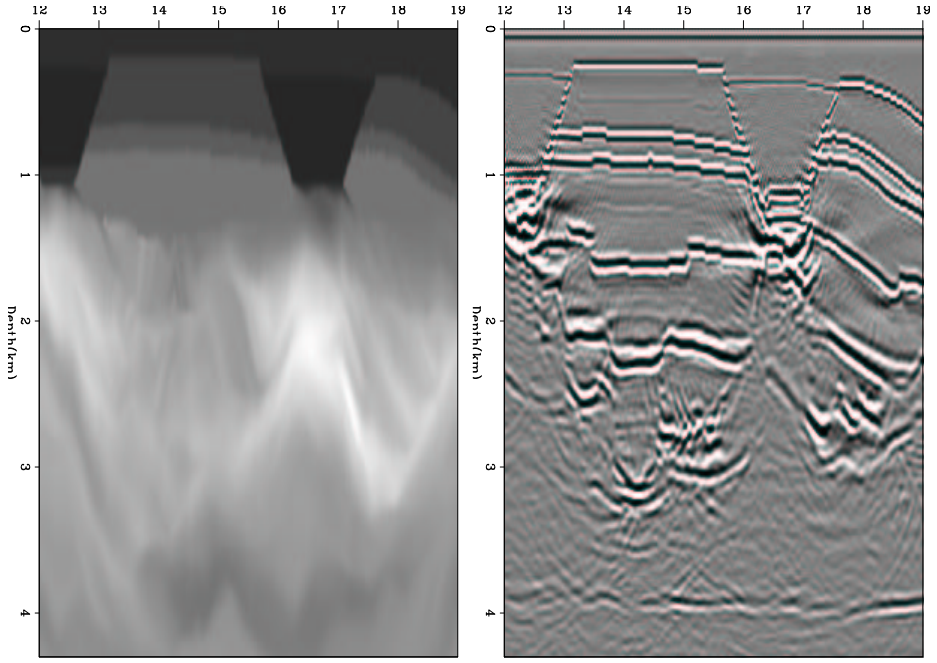



Figure 7: The left panel shows the velocity model after four iteration of tomography with a dip constraint. The right panel shows the result of migrating with the velocity model in the left panel. Note the more continuous nature of the bottom reflector, better constraining of the salt boundaries, and overall more accurate imaging focusing and positioning compared to the result without the added constraint (Figure 6. bob5-amoco-vel4.steer-ref [CR]

## REFERENCES

Claerbout, J., 1999, Everything depends on v(x,y,z): SEP–**100**, 1–10.

Clapp, R. G., and Biondi, B., 1999, Preconditioning tau tomography with geologic constraints: SEP–**100**, 35–50.

Clapp, R. G., Fomel, S., and Claerbout, J., 1997, Solution steering with space-variant filters: SEP–**95**, 27–42.

Clapp, R., 2001a, Ray-based tomography with limited picking: SEP–**110**, 103–112.

Clapp, R. G., 2001b, Geologically constrained migration velocity analysis: Ph.D. thesis, Stanford University.

Stork, C., 1992, Reflection tomography in the postmigrated domain: Geophysics, **57**, no. 5, 680–692.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 107–122*

# Velocity estimation for seismic data exhibiting focusing-effect AVO

*Ioan Vlad and Biondo Biondi*[1]

**ABSTRACT**

Transmission anomalies sometimes create AVO effects by focusing the reflected seismic wavefields, which impedes AVO analysis. The AVO anomalies caused by focusing are distinguishable by surface consistent patterns. We analyze the previous efforts to define, describe and eliminate spurious AVO anomalies. We also propose using wave equation migration velocity analysis to build an accurate velocity model. The transmission-related AVO can then be eliminated by downward continuation through this velocity model.

## INTRODUCTION

The phenomenon of amplitude variation with offset (AVO) of seismic reflection data is commonly assumed to occur only due to the petrophysical properties of the reflecting interfaces [Sheriff and Geldart (1995); Yilmaz (2001)]. However, amplitude can vary with offset due to absorption, or to focusing through velocity anomalies which are too small to give full triplications. The latter phenomenon has been shown to occur in several 2D [Kjartansson (1979); Harlan (1994)] and 3D (Hatchell, 2000a) seismic datasets. We will call it "focusing effect AVO" (FEAVO). FEAVO differs in two respects from regular AVO: 1) in CMP gathers, the high amplitudes are accompanied by local departures from hyperbolic moveout in the reflected arrivals (See Figure 1a); and 2) in the midpoint-offset space, the high amplitudes are distributed in "V" patterns whereas regular AVO gives rectangular patterns (See Fig. 1b). We must remove the FEAVO in order to allow AVO studies on FEAVO-affected data. Thus, it is necessary to:

Devise and prove the feasibility of a method whose application to FEAVO-affected 2D and 3D datasets would produce a velocity field accurate enough to generate a FEAVO-free prestack volume by downward continuing the wavefield through the FEAVO-generating anomalies.

## MOTIVATION

Most of the large structural onshore and shelf hydrocarbon plays have already been delineated, and seismic imaging efforts have begun to concentrate on subtler phenomena, such as those related to the presence of hydrocarbons in rocks. A resonant note is struck in crustal seismol-

[1]**email:** nick@sep.stanford.edu, biondo@sep.stanford.edu

ogy by the need to delineate the extent of lithospheric melts. AVO is one of the most common methods of characterizing the fluids in rocks in either exploration (Yilmaz, 2001) or crustal (Makovsky and Klemperer, 1999) surveys.

The FEAVO anomalies are much stronger than regular AVO effects, rendering AVO analysis impossible. Their removal will thus allow AVO analysis. A byproduct of the FEAVO removal process is a very accurate velocity model [White et al. (1988) shows that velocity contrasts as small as 2% can generate FEAVO], and this will also highly benefit AVO analysis, which is highly sensitive to the velocity used for prestack migration [Clapp (2002), Mora and Biondi (2000)].

FEAVO removal is also desirable for reasons beyond the obvious practical ones described above: in principle, the reflectivity that seismology seeks to recover is the high spatial frequency component of the impedance field. A FEAVO-contaminated image is simply inaccurate. Imaging the correct reflectivities is in line with the modern efforts towards true-amplitude imaging [Biondi (2001b); Sava and Biondi (2001a)]. And the by-product – an accurate velocity model describing the low spatial frequencies of the velocity field – is every bit as important as the reflectivity image itself (Claerbout, 1999).

## WHAT DO WE WANT TO DO?

We propose doing wave equation migration velocity analysis with a fitting goal specifically adapted to the nature of the FEAVO anomalies. This will provide a velocity model accurate enough so that we can eliminate the FEAVO effects by redatuming through it.

We will further describe the previous work on the recognition and removal of the FEAVO anomalies. We will lay down an approach to be undertaken in order to solve the problem defined in the second paragraph of this paper and we will evaluate the practical aspects of the implementation of the described approach. We will also describe the steps that we already undertook to solve the problem and the work that remains to be done, together with a time estimate for that.

## PREVIOUS WORK

### Work defining and describing the FEAVO effect

Kjartansson (1979) gives an example of the FEAVO effect on a real 2D dataset and interprets its physical meaning. Claerbout acknowledges the importance of the phenomenon by including a description and an explanation of it in his course notes (1982) and in his book, *Imaging the Earth's Interior* (1985). White et al. (1988) use forward modeling to show that FEAVO effects can be due to smooth velocity anomalies deviating by as little as 2% from the background velocity. The anomalies must be large compared to a wavelength but small when compared to the propagation distance. A solid case study (Hatchell, 1999) based on two real 3D datasets and on realistic forward modeling shows that not only shallow velocity anomalies can cause

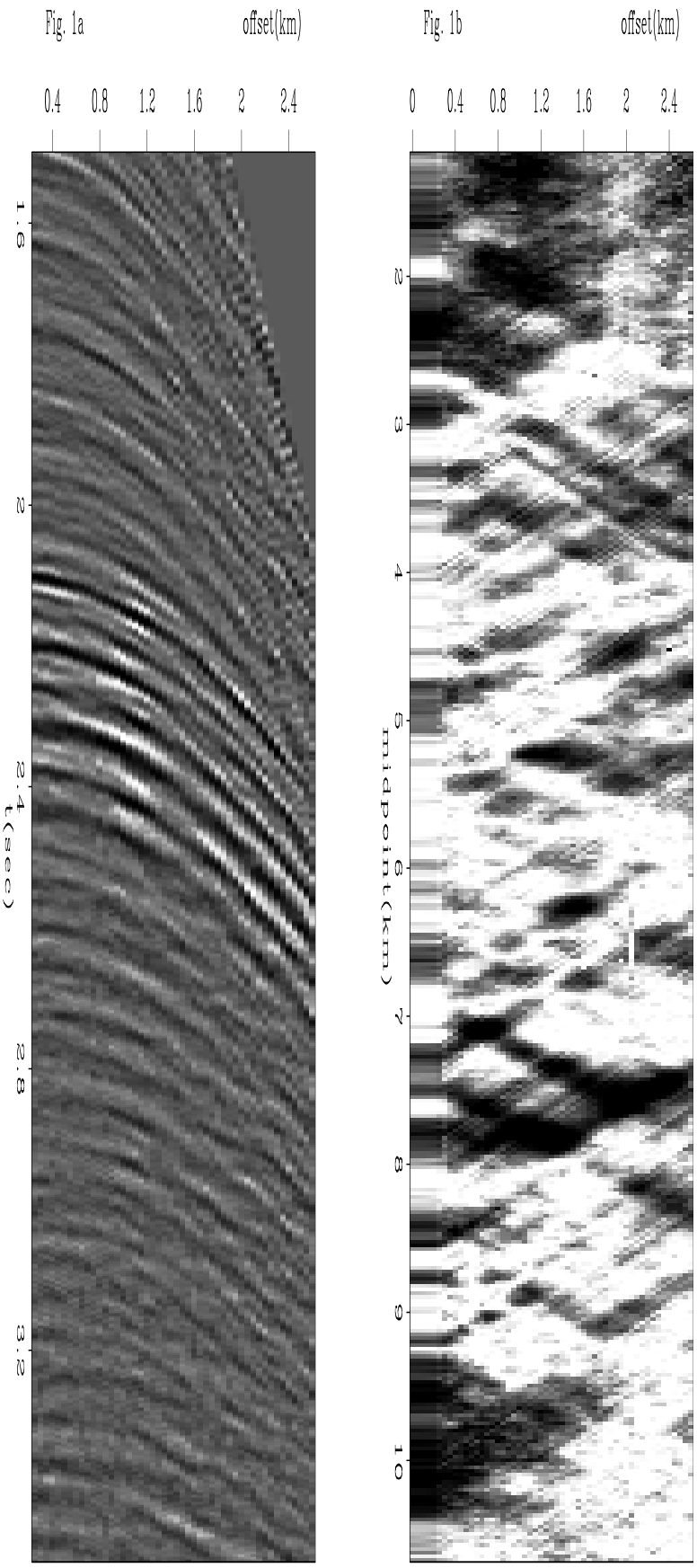



Figure 1: **a**: Part of a CMP gather exhibiting FEAVO anomalies. The strong event at 2.3s shows a slight departure from hyperbolic moveout, too subtle to allow successful classical traveltime tomography. **b**: FEAVO anomalies in the midpoint-offset space (Kjartansson "V"s). The preprocessing consisted in: muting, spherical divergence correction, bandpass filter, interpolation of missing or noisy traces, hydrophone balancing, f-k filtering, and offset continuation to fill in the small offsets (with a forward and inverse DMO cascade using the log-stretch DMO in the Fourier domain described in Zhou et al. (1996) and implemented by Vlad and Biondi (2001)). The figure has been produced exactely as in Kjartansson (1979): square and vertically stack the data between 1.5 and 3.5 seconds, then take the logarithm to increase the dynamic range. Offset continuation does not predict the FEAVO anomalies (the tips of the "V"s are not extended into the extrapolated small offsets) nick2-prodef [ER]

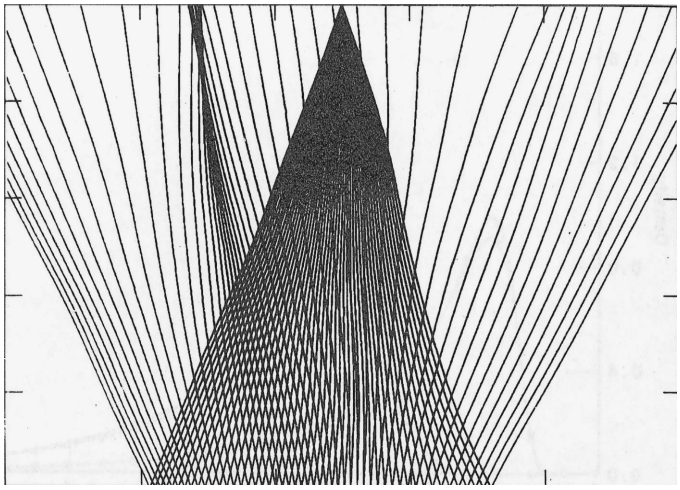


Figure 2: The physical explanation for the expression of FEAVO anomalies in CMP gathers (Figure 1a). If the frequency of the waves is high enough or the anomaly large enough, we will see a small triplication. Otherwise, only offset-dependant amplitude focusing (FEAVO) is visible. The traveltime delays are negligible, as the velocity anomaly changes only very little the length of the rays. Figure taken from White et al. (1988). nick2-whitebig [NR]

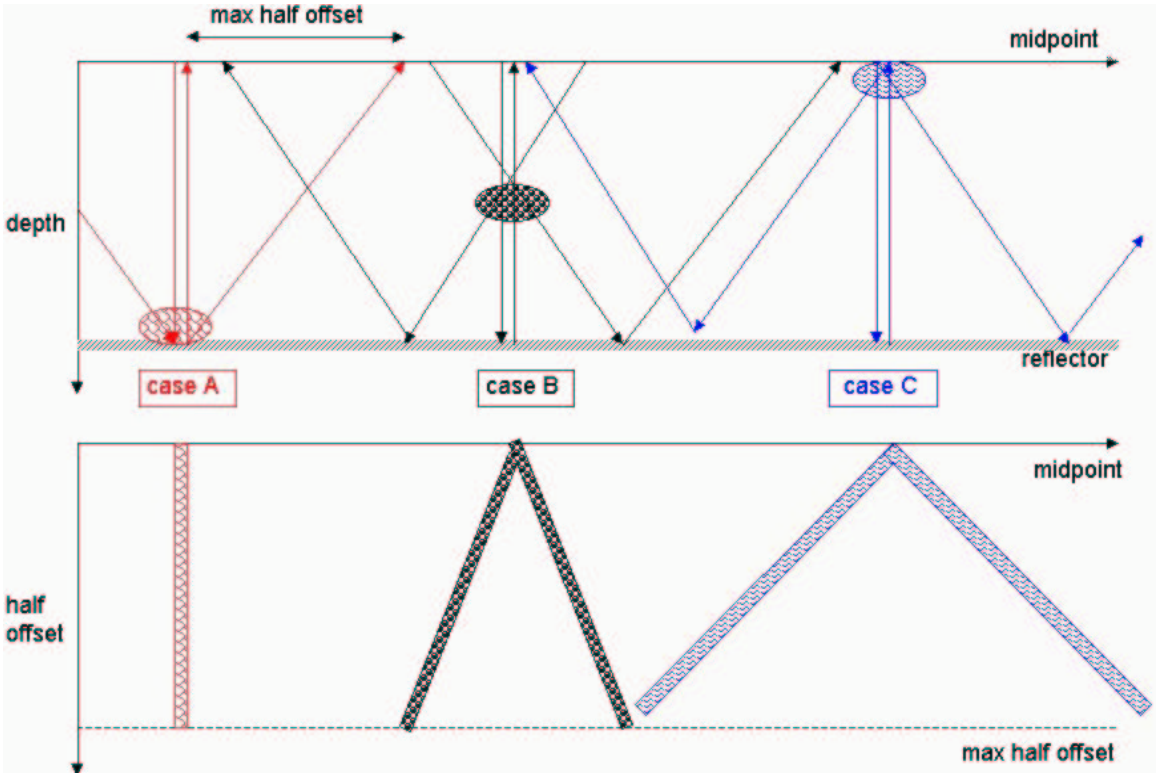



Figure 3: The physical explanation for the expression of FEAVO anomalies in midpoint-offset space (Kjartansson "V"s, Figure 1b). In the upper picture, the blobs are transmission anomalies and the arrows are raypaths for the zero offset and for the maximum offset recordings. For case A (anomaly on the reflector), only a single midpoint is affected, for all offsets. Case C (anomaly at the surface), is actually a static: its "footprint" is a pair of streaks slanting $45^o$ from the offset axis. Case B (in between) gives a pair of streaks with angles smaller than $45^o$. nick2-vilus [NR]

FEAVO effects; the sudden termination of a deep (3000m), low velocity layer by a fault can also be a source of focusing that affects the amplitudes more strongly than the traveltimes. This point draws interest: Hatchell (1999) received a Best Paper award at the SEG Annual Meeting and the idea is reiterated in Hatchell (2000a) and Hatchell (2000b).

### Attempts to invert FEAVO-affected data for a velocity model

Kjartansson (1979) not only recognizes the meaning of the FEAVO effect, but also takes a first step towards obtaining a velocity model. Taking into account that in the midpoint-offset space the opening angle of the "V"s depends on the depth of the transmission anomaly, he makes maps of the anomalies by slant stacking the power of the raw, unmigrated data. The maps adequately indicate the positions of the velocity anomalies but not the magnitudes. Claerbout (1993) refines this approach. In the frame of his work with prediction error filters at the time, he uses them to extend the Grand Isle dataset over a larger range of offsets. He also uses the new tool to compute a pilot trace that is crosscorrelated with the data. The traveltime anomalies are emphasized by plotting the times of the maximum values of the crosscorrelations in the midpoint-offset space.

Bevc (1993) generates a synthetic dataset exhibiting FEAVO anomalies and proposes a data-space tomography approach using downward continuation as the operator in tomography. This line of work is continued by Bevc (1994b), who shows using the synthetic that the FEAVO anomalies can indeed be eliminated by finding the velocity model, and then by downward continuing through it until under the FEAVO-causing velocity anomalies (redatuming). He also theoretically discusses inversion schemes. In the follow-up, Bevc (1994a) picks times from sags in the FEAVO-affected quasi-hyperbolic arrivals in CMP gathers and does traveltime tomography to find the velocity model. He operates under the assumption that the velocity anomalies are in the near surface. The tomography works in the following way: 1) apply NMO; 2) find the departures from the flatness of the events in CMP gathers (time lags) by crosscorrelating a pilot trace with the data; and 3) the lags are backprojected onto the velocity model that will be used for the NMO at the first step. The inversion in the last tomographic step uses a styling goal with PEFs. The method is shown to work on a simple synthetic dataset, but the conclusions contain a warning that it may not work on real data.

With a different approach (inverting picked maximum amplitudes and using ray-based operators), Harlan (1994) does not produce a velocity model, but a transmission anomaly section that is used to eliminate the FEAVO effect for a 2D dataset.

## PROPOSED APPROACH

We intend to eliminate the FEAVO anomalies by finding an accurate velocity model, then by downward continuing the data through it until under the FEAVO-causing velocity anomalies. The velocity will be found by migration velocity analysis (MVA), an iterative inversion process whose optimization goal is not fitting the recorded data, but providing the best focused migrated image (Biondi and Sava, 1999). The chapter pertaining to velocity analysis in Biondi

(2001a) shows (with examples) why dipping reflectors in laterally varying velocity media require the velocity analysis to be performed in the migrated domain (image domain) instead of the unmigrated domain (data domain). A wave-equation MVA (WEMVA) (Sava, 2000) will be used instead of a ray-based MVA (Clapp, 2001). The advantages of the former over the latter are detailed in the WEMVA chapter of Biondi (2001a). One particular advantage is the better treatment of amplitudes by wave-equation methods.

The usual WEMVA criterium describing the quality of the image is flatness in angle gathers. This is directly related to traveltime anomalies. As it is visible in Figure 1a, the traveltime changes associated with the FEAVO effect are very small and they do not produce curvatures in angle gathers. Biondi and Sava (1999) show on a synthetic, and this paper will show on a real dataset, that FEAVO anomalies keep their "V" shapes through prestack migration and conversion from offset to angle gathers. Therefore, the fitting goal of the inversion must be related to the distribution of amplitudes in the midpoint-angle space. The desired image will not exhibit these characteristic "V" patterns.

The inversion will proceed as follows: the wavefield at a certain depth is downward continued a depth step down through the known velocity model with an accurate (nonlinear) operator. The result is transformed to angle domain and a "perfect" image is created by eliminating the FEAVO anomalies. A image perturbation ($\Delta W$) is obtained by substracting the two images, and is backprojected through an invertible operator in order to obtain a velocity update ($\Delta s$). The velocity model is updated and the cycle proceeds again, until $\Delta W$ becomes negligible. The construction of the operator that links $\Delta s$ and $\Delta W$ is very important. The number of iterations and the accuracy of the result depends on its accurate invertibility. In order to make it invertible, Born (Sava and Biondi, 2001b) or other (Sava and Fomel, 2002) types of linearization are employed.

## COMPLETED STEPS TOWARDS THE STATED GOAL

Any iterative inversion process is built on three cornerstones: 1. The ability to transform a guess of the solution from the initial domain to a domain where an anomaly can be extracted; 2. The ability to extract the anomaly in this latter domain by comparing the result of the transformation of the initial guess against a set of criteria; and 3. The ability to transform the extracted anomaly back to the initial domain, updating the guess. In the particular case of the inversion we are setting up, the discrimination domain is the angle domain. We will further prove point 1 above: that the FEAVO anomalies are visible in the angle domain (that the common downward continuation followed by transformation from the offset to angle does not destroy them). We will also discuss point 2 (the various methods for extracting the FEAVO anomalies in the angle domain) but will leave an example of the actual extraction to a future paper. Because of time constraints, point 3 (proving that the linearized downward continuation does not spoil the anomalies) is also left for a future paper.

**Shallow-origin FEAVO effects are visible in the image domain**

Figure 1b shows a pattern of slanted symmetrical linear patterns (Kjartansson "V"s) that are the signature of FEAVO in the midpoint-offset domain. Figure 3 illustrates the way the patterns formed. The Kjartansson "V"s are a good way to discriminate the FEAVO anomalies in an inversion process. If we decide that the inversion discrimination will be done in the image domain, we need to make sure that the anomalies will not be destroyed by downward continuation. This was already shown on a synthetic in Fig. 5 of Biondi and Sava (1999). In order to show that the V-shaped patterns are visible as well in angle domain images produced from a real dataset, we migrated the Grand Isle prestack dataset then transformed to angle domain. The migration velocity depends only on depth (Figure 4), but not on midpoint (to avoid focusing by migration with an inappropriate velocity model). This assumption is close to the truth - Figure 5 shows that the geology is quite flat in the area. An examination of the midpoint-angle slices (Figure 6) reveals V-shaped patterns at the same locations as the data domain ones.

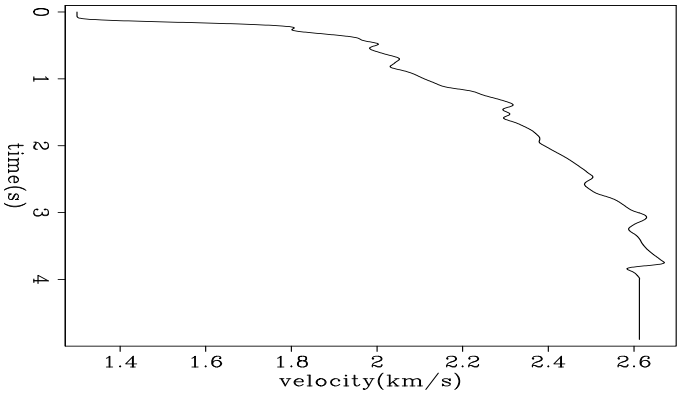



Figure 4: Interval velocity used in the migration of the Grand Isle dataset. The velocity does not depend on midpoint in order to not inadverdently focus the energy (the purpose of the migration is seeing whether the focusing anomalies are preserved). The *v(z)* assumption is close to geological reality in that area too. nick2-vint [ER]

**Deep-origin FEAVO effects can be modeled numerically**

Hatchell (2000a) proves using both real and synthetic data that FEAVO effects can be generated not only by shallow velocity anomalies, but also by deep ones. Figure 7 reproduces the forward modeling in Hatchell (2000a). A pseudospectral wave propagation algorithm (Biondi, 2002) is used to produce the wavefield recorded at a depth of 6000m. Ignoring the dispersion effects, the bottom two panels in the figure exhibit a striking quantitative similarity to the results in Hatchell (2000a) and emphasize that deep-origin FEAVO effects exist, and that they can be modeled numerically.

**Methods of extracting FEAVO effects in the angle gathers**

In order to invert, one must extract the anomalies. Since they are expressed both in the midpoint-angle domain and in the angle-domain common image gathers, their separation must proceed in a synergistic fashion. Figure 6 shows they bear a striking morphological resemblance to channels, so pattern-recognition techniques applied to a coherence volume (Marfurt

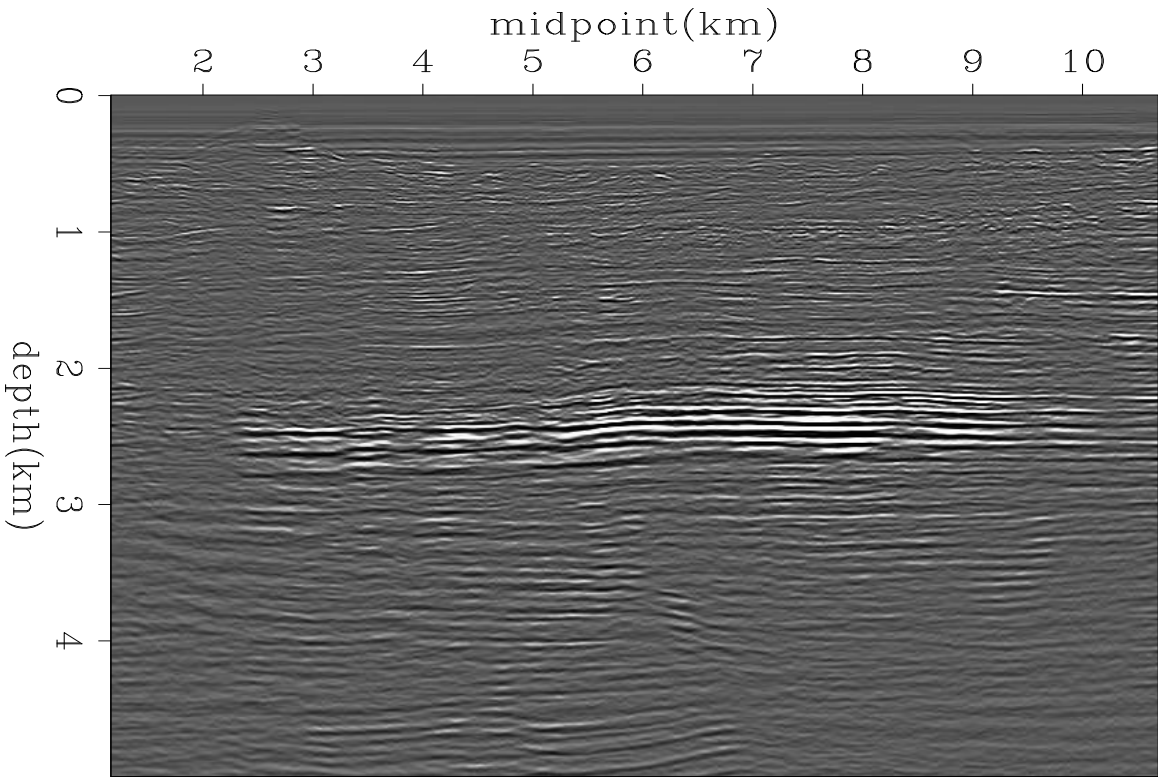



Figure 5: Migration (stack) of the Grand Isle dataset. The structure is quite flat. nick2-kimag [CR]

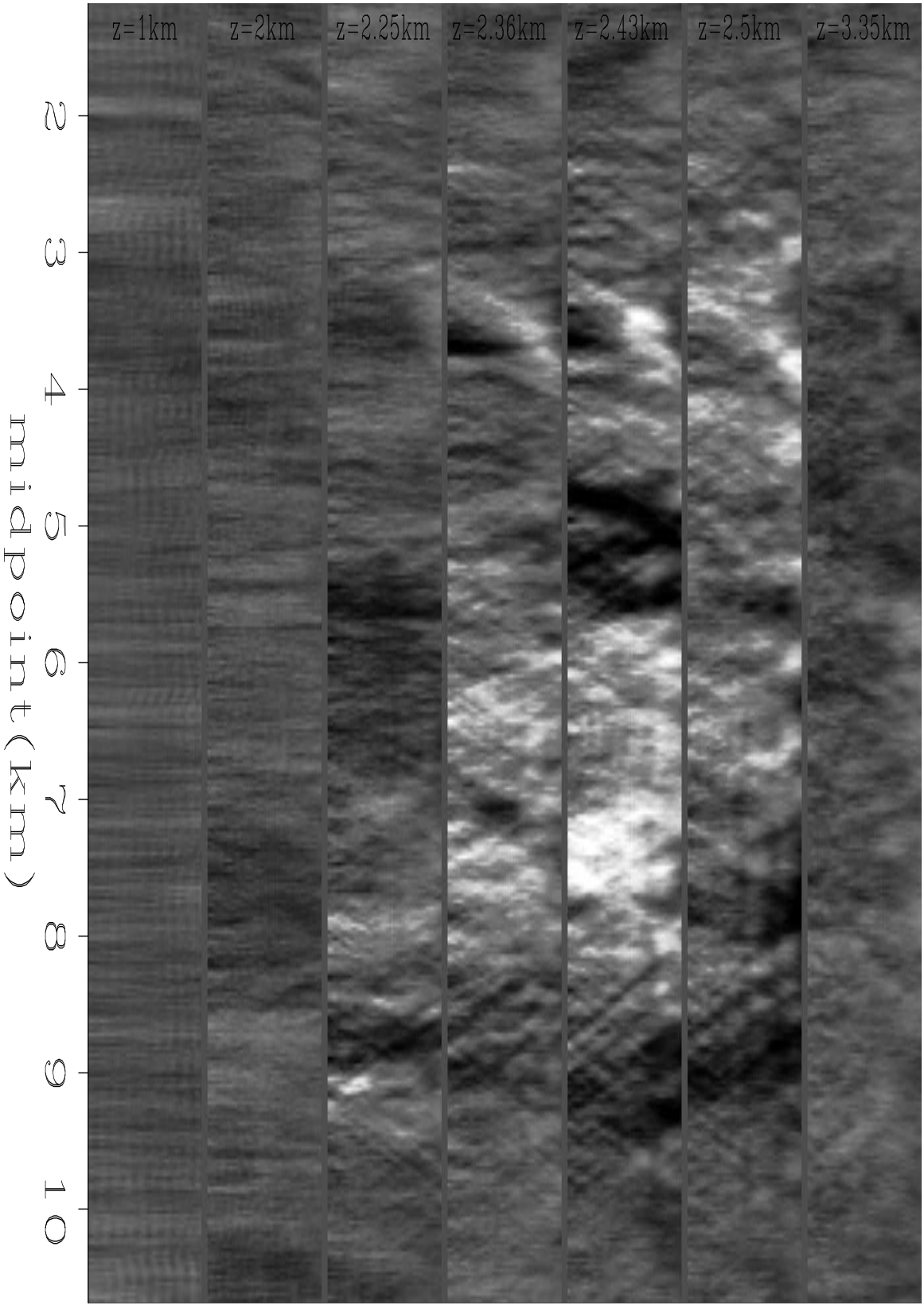



Figure 6: FEAVO anomalies in midpoint-angle depth slices. Each vertical panel shows the angle range between 0 and 20 degrees. The first panel (z=1km), the last (z=3.35km) and to a certain extent the second (z=2km) do not contain Kjartansson "V"s nick2-angpanel [CR]

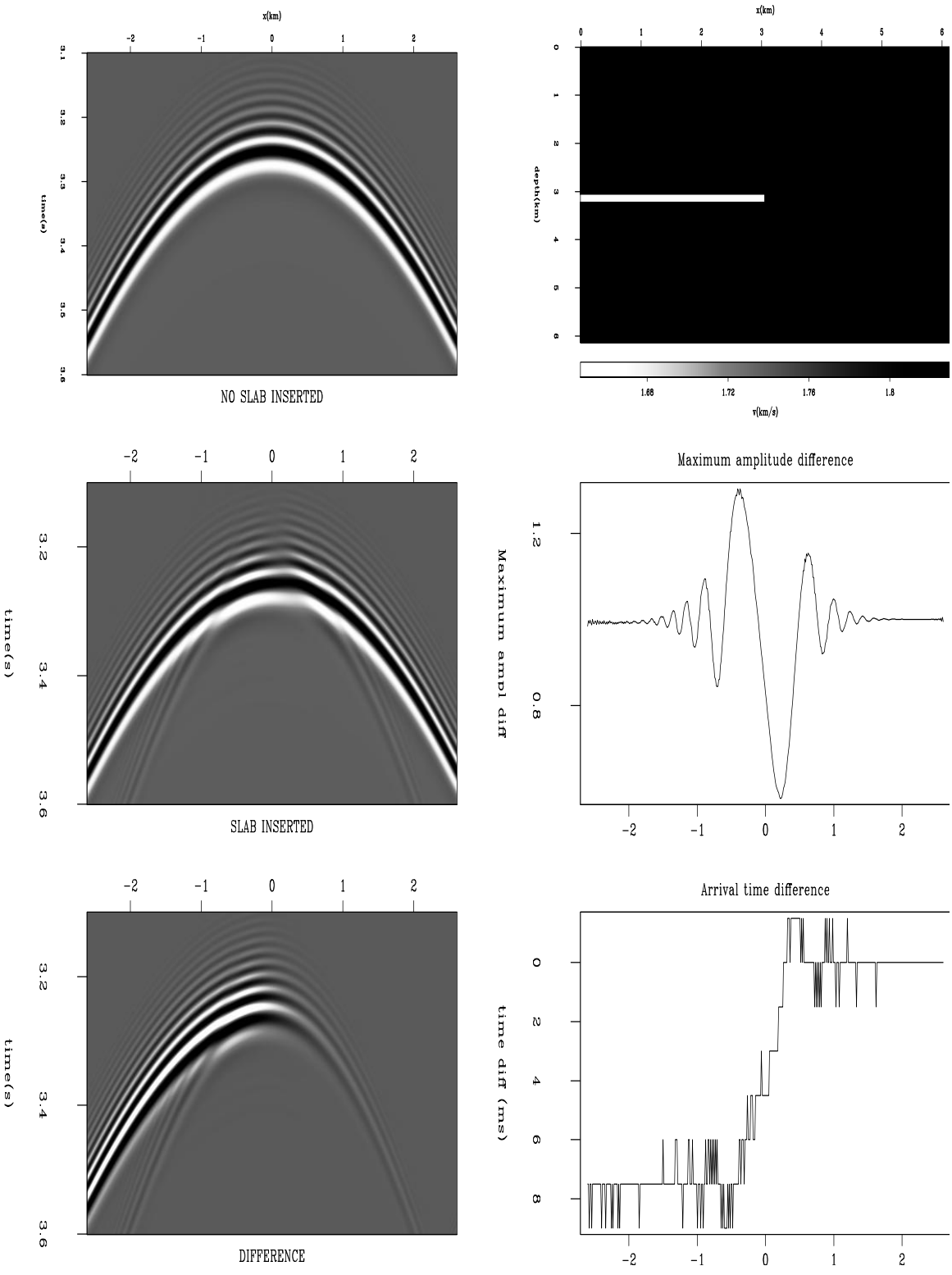



Figure 7: Left, from top to bottom: 1. Downward continuation of a shot from surface to 6 km deep, with constant velocity. 2. Downward continuation through upper right velocity model. 3. Difference between 1 and 2. Right, from top to bottom: 4. New velocity model — homogenous with a lower velocity slab inserted. 5. Difference between the maximum amplitudes in panel 1 and panel 2, for each x location. 6. Difference between the times of the maximum amplitudes in panels 1 and 2, for each x location. nick2-hatsim [CR]

et al., 1998) of the prestack cube may yield good results. Another way to discriminate them may be building a dip volume (Fomel, 2000) using plane-wave destructors (Claerbout, 1991). This may be successful because it is actually biomimicry: the human eye discriminates the anomalies based on their dip, or in the case of the data-domain hyperbola sags, on the discontinuities in the dip (dip derivative). These morphological approaches may be corroborated with spectral discrimination. In some cases, such as for the Grand Isle dataset, the FEAVO-causing velocity anomalies are small enough that different wavelengths may be affected differently (Figure 8).

## WORK TO BE DONE

In order to prove beyond reasonable doubt that an inversion is feasible, an actual extraction of the FEAVO anomalies in angle gathers should be performed. This should not pose any problems, amounting in the end to image processing - discriminating the anomalies based on their morphology. Since the third prerequisite for an inversion is an invertible operator, we should also show, at least on a particular case, that the linearized downward continuation operator preserves the FEAVO anomalies. Thus, after proving that an inversion for the velocity model that produces the FEAVO anomalies is possible, the work of setting up such an inversion first for 2D, then for 3D, remains to be done. The operators must be constructed and a reliable discrimination method in angle gathers set up.

We will approach the solution of the problem gradually, incorporating elements of complexity one by one. We first have to develop a process that will successfully separate the FEAVO anomalies in the angle domain, making them ready for input into an inversion scheme for obtaining $\Delta s$. We will then implement a 2D synthetic case, then a 2D real. We will develop the theory for the 3D case, build a 3D synthetic, implement a 3D synthetic case, then a 3D real one, and finally lay out the results in a publishable form.

## RESOURCES

### Seismic data

According to the strategy described in the previous section, we will use the following seismic datasets:

2D synthetic data: synthetic datasets exhibiting FEAVO effects have been generated in the past at SEP [Bevc (1993), Biondi and Sava (1999)] and due to research reproducibility, they can be easily obtained.

2D real data: the seismic line described in Kjartansson (1979). The dataset was acquired over the Grand Isle gas field offshore Louisiana and was made available to SEP in 1979 by Dr. Ralph Shuey (Gulf Science and Technology Company, at the time).

3D synthetic: After we prove the workability of the method in 2D, both on synthetic and

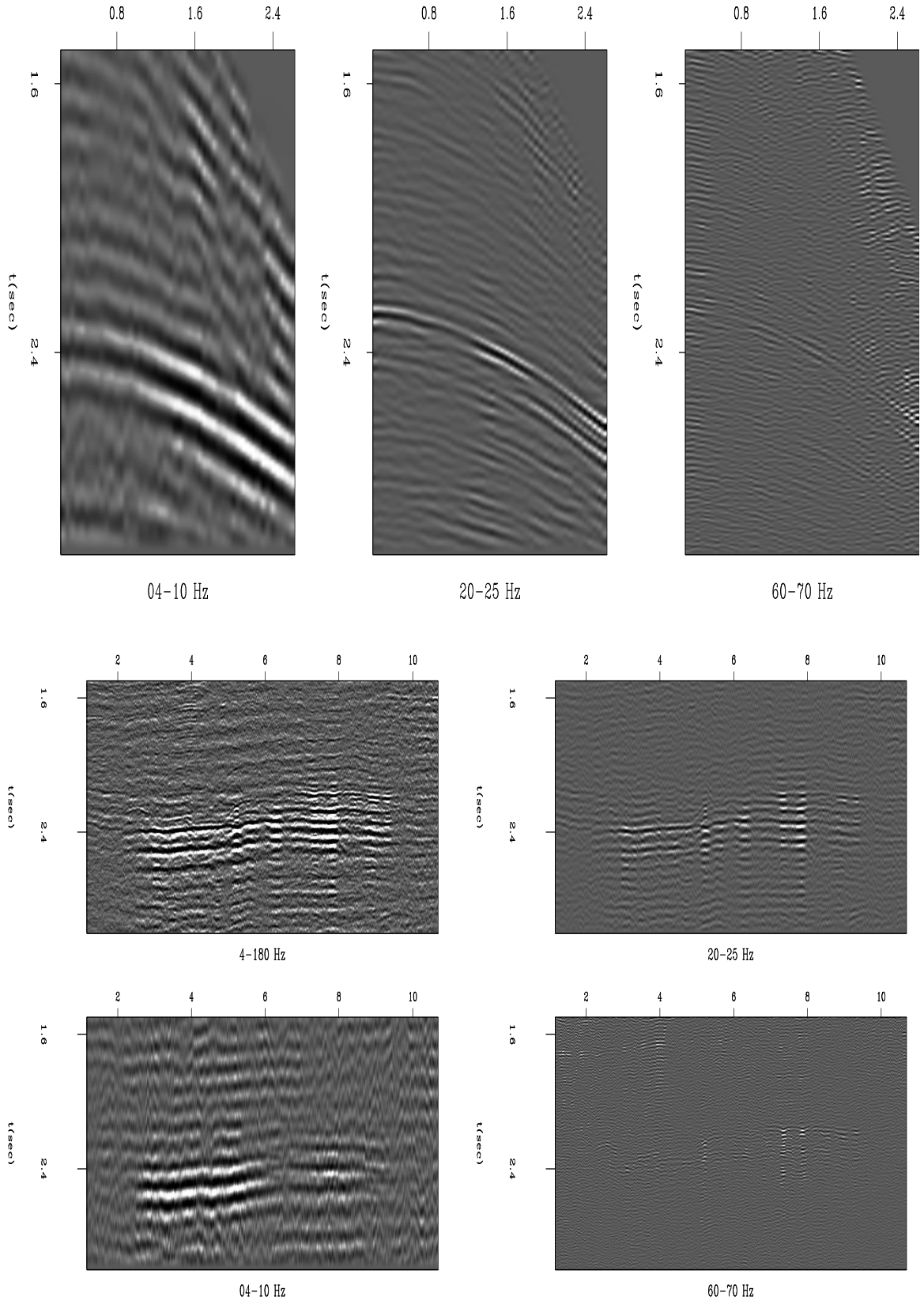



Figure 8: Top: CMP gather at 4.7 km, filtered with 4-10 Hz, 20-25 Hz, 60-70 Hz. Bottom: section with offset = 1.291km, filtered with 4-180 Hz, 4-10 Hz, 20-25 Hz, 60-70 Hz. Different frequency bands result in different extents of the FEAVO anomalies. nick2-kfdp [ER]

real data, we will generate a 3D synthetic data set exhibiting FEAVO anomalies.

3D real: Hatchell (2000a) has shown two examples of FEAVO-affected 3D datasets in the Gulf of Mexico, and by personal communication, we found out that there are other cases as well. We believe that once the proposed process shows its workability on a 3D synthetic, 3D real data can also be obtained.

### Software

Most of the software needed is already available at SEP. Besides elementary tools such as `Fortran90` compilers, `SEPlib`, `Matlab` and `Mathematica`, the $\Delta$Image-$\Delta$Slowness operator is already included in the `gendown` (Biondo Biondi) and `WEI` (Paul Sava) packages, already in existence at SEP. An out-of-core optimization library (`oclib`) for inverting large 3D datasets has already been written (Sava, 2001).

### Hardware

The proposed inversion scheme involves applying wave-equation migration to the data several times, so is therefore very computationally-intensive. Due to cluster technology, SEP has increased its computational power, but repeated prestack migrations of real 3D datasets, necessary for the final stages of this project, still need even more cluster nodes.

## ACKNOWLEDGEMENTS

Our thanks to Paul Sava for his help migrating and transforming to angle domain the Grand Isle dataset; to Bob Clapp for his advice and computer support; and to Brad Artman for the velocity analysis of the Grand Isle dataset.

## REFERENCES

Bevc, D., 1993, Toward estimating near surface lateral velocity variations: SEP–**79**, 69–78.

Bevc, D., 1994a, Datuming velocity from traveltime tomography: SEP–**82**, 145–164.

Bevc, D., 1994b, Near-surface velocity estimation and layer replacement: SEP–**80**, 361–372.

Biondi, B., and Sava, P., 1999, Wave-equation migration velocity analysis: SEP–**100**, 11–34.

Biondi, B. 3-d Seismic Imaging:. http://sepwww.stanford.edu/sep/biondo/Lectures, 2001.

Biondi, B., 2001b, Amplitude preserving prestack imaging of irregularly sampled 3-D data: SEP–**110**, 1–18.

Biondi, B., 2002, Reverse time migration in midpoint-offset coordinates: SEP–**111**, 147–155.

Claerbout, J. F., 1982, Imaging the Earth's Interior, chapters one to four: SEP–**30**, 233.

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

Claerbout, J. F., 1991, Earth Soundings Analysis: SEP–**71**, 1–304.

Claerbout, J. F., 1993, Reflection tomography: Kjartansson revisited: SEP–**79**, 59–68.

Claerbout, J., 1999, Everything depends on v(x,y,z): SEP–**100**, 1–10.

Clapp, R. G., 2001, Geologically constrained migration velocity analysis: Ph.D. thesis, Stanford University.

Clapp, R. G., 2002, Effect of velocity uncertainty on amplitude information: SEP–**111**, 253–266.

Fomel, S., 2000, Applications of plane-wave destructor filters: SEP–**105**, 1–26.

Harlan, W. S., 1994, Tomographic correction of transmission distortions in reflected seismic amplitudes: 64th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 968–971.

Hatchell, P., 1999, Fault whispers: Transmission distortions on prestack seismic reflection data: 69th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 864–867.

Hatchell, P., 2000a, Fault whispers: Transmission distortions on prestack seismic reflection data: Geophysics, **65**, no. 2, 377–389.

Hatchell, P., 2000b, What causes distortions on prestack reflection seismic data?: World Oil, **221**, no. 11, 69–77.

Kjartansson, E., 1979, Analysis of variations in amplitudes and traveltimes with offset and midpoint: SEP–**20**, 1–24.

Makovsky, Y., and Klemperer, S., 1999, Measuring the seismic properties of Tibetan bright spots: Evidence for free aqueous fluids in the Tibetan middle crust: Journal of Geophysical Research, **104**, no. B5, 10795–10825.

Marfurt, K. J., Kirlin, R. L., and Farmer, S. L., 1998, 3-d seismic attributes using a semblance-based coherency algorithm: Geophysics, **63**, no. 04, 1150–1165.

Mora, C., and Biondi, B., 2000, Estimation of AVO attributes sensitivity to velocity uncertainty using forward modeling: A progress report: SEP–**103**, 349–366.

Sava, P., and Biondi, B., 2001a, Amplitude-preserved wave-equation migration: SEP–**108**, 1–26.

Sava, P., and Biondi, B., 2001b, Born-compliant image perturbation for wave-equation migration velocity analysis: SEP–**110**, 91–102.

Sava, P., and Fomel, S., 2002, Wave-equation migration velocity analysis beyond the Born approximation: SEP–**111**, 81–99.

Sava, P., 2000, A tutorial on mixed-domain wave-equation migration and migration velocity analysis: SEP–**105**, 139–156.

Sava, P., 2001, oclib - An out-of-core optimization library: SEP–**108**, 199–224.

Sheriff, R. E., and Geldart, L. P., 1995, Exploration Seismology: Cambridge University Press.

Vlad, I., and Biondi, B., 2001, Effective AMO implementation in the log-stretch, frequency-wavenumber domain: SEP–**110**, 63–70.

White, B. S., Nair, B., and Bayliss, A., 1988, Random rays and seismic amplitude anomalies: Geophysics, **53**, no. 07, 903–907.

Yilmaz, O., 2001, Seismic data analysis: processing, inversion and interpretation of seismic data: Society of Exploration Geophysicists.

Zhou, B., Mason, I. M., and Greenhalgh, S. A., 1996, An accurate formulation of log-stretch dip moveout in the frequency-wavenumber domain: Geophysics, **61**, no. 3, 17–23.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 123–138*

# Prestack imaging of overturned and prismatic reflections by reverse time migration

*Biondo Biondi*[1]

**ABSTRACT**

I present a simple method for computing angle-domain Common Image Gather (CIG) using prestack reverse time migration. The proposed method is an extension of the method proposed by Rickett and Sava (2001) to compute CIGs by downward-continuation shot-profile migration. I demonstrate with a synthetic example the use of the CIG gathers for migration velocity updating. A challenge for imaging both overturned and prismatic reflections is the discrimination of the reflection generated on either side of interfaces. I show how the propagation direction of the reflections can be easily determined by evaluating the crosscorrelation of the source wavefield with the receiver wavefield at time lags different than zero. Reflections can be easily separated once their direction of propagation is determined. I demonstrate the method by imaging overturned events generated by a segment of dipping reflector immersed in a vertically layered medium. The example shows that for overturned reflections velocity errors cause asymmetric residual moveouts. This observation suggests that for updating the velocity from overturned reflections, the migrated CIGs should be scanned using a different family of residual moveouts than the standard parabolic moveouts used for non-overturned events. The migration of a synthetic example containing prismatic reflection shows that the CIGs computed using only horizontal subsurface offset are not useful, and that the vertical subsurface offset should be also used.

## INTRODUCTION

As seismic imaging is applied to more challenging situations where the overburden is ever more complex (e.g., imaging under complex and rugose salt bodies) and the illumination of important reflectors is spotty, we need to use all the events in the data to generate interpretable images. Two classes of events that are often neglected, but that are also often present in complex data are overturned reflections (Li et al., 1983) and prismatic reflections (Broto and Lailly, 2001). These two classes of events share the challenge that they cannot be imaged correctly (at least in laterally varying media) by downward-continuation migration methods. This obstacle prompted me to look at reverse time migration (Baysal et al., 1983), and in particular at reverse time migration of shot profiles (Etgen, 1986b).

[1]**email:** biondo@sep.stanford.edu

However, the current status of reverse-time migration technology has some limitations that need to be addressed before it can be used effectively to image either overturned or prismatic reflections. The main challenge is the extraction of useful and robust velocity updating information from the migrated image. In complex media, velocity is usually updated from the information provided by migrated Common Image Gathers (CIG). Filho (1992) presented the only other method that I am aware of to compute angle-domain CIGs (ADCIGs) by reverse time migration. He applied the method to Amplitude Versus Angle (AVA) analysis. His method is computationally involved and does not seem to be particularly robust because it requires the identification of local plane waves.

In this paper, I extend to reverse-time shot-profile migration the method that Rickett and Sava (2001) proposed to compute CIGs by downward-continuation shot-profile migration. The idea is to compute offset-domain CIGs by a modified imaging condition that introduces the concept of a *subsurface offset*. Simple testing using synthetic data confirmed that the CIG computed applying the proposed method can be used for velocity updating. They should also be useful for AVA analysis, though I have not yet analyzed their amplitude properties. However, for both overturned reflections and prismatic reflections, the source wavefield and the receiver wavefield may be propagating along opposite vertical directions at the reflection point. For these two classes of events, the imaging principle should be generalized to include a *vertical subsurface offset* as well an horizontal one. I have not tested this generalization of the CIG methodology yet.

The proper imaging of both overturned waves and prismatic reflections presents another challenge related to the discrimination between the image contributions of reflections generated from either side of an interface. These two reflections need to be discriminated both for imaging of reflectivity and for robust velocity updating, because these two reflections have usually opposite polarity, and their kinematics are affected by different areas of the velocity model. I present a simple generalization of the imaging condition that enables the determination of the propagation direction of the reflections, and thus the separation of the image contributions related to different events.

Reverse time migration has some other known disadvantages that need to be addressed. One of them is the computational cost. A companion short note in this report (Biondi, 2002) is an attempt to address this issue, at least for 3-D marine data. The artifacts caused by discontinuities in the migration velocity model are a more fundamental problem. There are a number of proposed solutions in the literature (Baysal et al., 1984; Etgen, 1986a; Filho, 1992). None of them is completely satisfactory, and thus more research on the subject is needed.

## ANGLE-DOMAIN COMMON-IMAGE GATHERS BY REVERSE TIME MIGRATION

The standard imaging condition for prestack reverse time migration is based on the crosscorrelation in time of the source wavefield ($S$) with the receiver wavefield ($R$). The equivalent of the stacked image is the average over the sources ($s$) of the zero lag of this crosscorrelation,

that is:

$$I(z,\mathbf{x}) = \sum_s \sum_t S_s(t,z,\mathbf{x})\, R_s(t,z,\mathbf{x}), \tag{1}$$

where $z$ and $\mathbf{x}$ are respectively depth and the horizontal axes, and $t$ is time. The image created using this imaging condition is the equivalent to the stack over offsets for Kirchhoff migration.

This imaging condition has the disadvantage of not allowing a prestack analysis of the image for either velocity updating or amplitude analysis. The conventional way of overcoming this limitation is to avoid the averaging over sources, and thus to create CIGs where the horizontal axis is related to a *surface offset*; that is, the distance between the source location and the image point. This kind of CIG is known to be prone to artifacts even when the migration velocity is correct because the non-specular reflections do not destructively interfere. Furthermore, in presence of migration velocity errors and structural dips, this kind of CIG does not provide useful information for improving the velocity function.

Rickett and Sava (2001) proposed a method for creating more useful angle-domain CIGs with shot profile migration using downward continuation. Their method can be easily extended to reverse time migration. Equation (1) can be generalized by crosscorrelating the wavefields shifted with respect to each other. The prestack image becomes function of the horizontal relative shift, that has the physical meaning of a *subsurface offset* ($\mathbf{x}_h$). It can be computed as

$$I(z,\mathbf{x},\mathbf{x}_h) = \sum_s \sum_t S_s\left(t,z,\mathbf{x}+\frac{\mathbf{x}_h}{2}\right) R_s\left(t,z,\mathbf{x}-\frac{\mathbf{x}_h}{2}\right) \tag{2}$$

This imaging condition generates CIGs in the offset domain that can be easily transformed to the more useful ADCIG applying the same methodology described in (Sava et al., 2001).

Reverse time migration is more general than downward continuation migration because it allows events to propagate both upward and downward. Therefore the ADCIG computed from reverse time migration can be more general than the ones computed from downward-continuation migration. This more general imaging condition is actually needed when the source and receiver wavefields meet at the reflector when propagating along opposite vertical direction. This condition may occur either when we image overturned events or image prismatic reflections. I analyze these situations in more details in the following sections. To create useful ADCIGs in these situations we can introduce a *vertical offset* $z_h$ into equation (2) and obtain

$$I(z,\mathbf{x},z_h,\mathbf{x}_h) = \sum_s \sum_t S_s\left(t,z+\frac{z_h}{2},\mathbf{x}+\frac{\mathbf{x}_h}{2}\right) R_s\left(t,z-\frac{z_h}{2},\mathbf{x}-\frac{\mathbf{x}_h}{2}\right) \tag{3}$$

These offset-domain CIGs should be amenable to being transformed into angle-domain CIG, by generalizing the methodology described in (Sava et al., 2001). However, I have not tested it yet, and in the examples I show only the application of the less general equation (2).

## Examples of ADCIG to a simple synthetic data set

To illustrate the use of the proposed method to compute ADCIG I created a simple synthetic data set using a pseudo-spectral modeling code and then migrated the recorded shots using the same pseudo-spectral wave-propagation kernel.

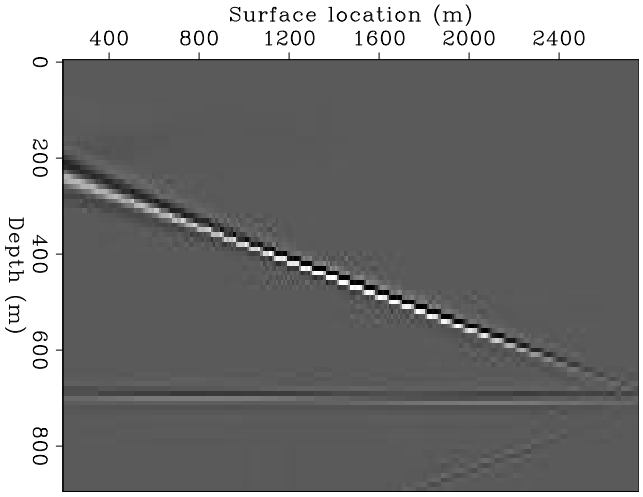



Figure 1: Image of the synthetic data set with the correct velocity function. biondo2-Shot-Image-dip [CR]

I modeled and migrated 100 shots spaced 10 m apart, starting from the surface coordinate of 1,000 m. The receivers were in a symmetric split-spread configuration with maximum offset of 2,550 m. I assumed two reflectors: one dipping 10 degrees and the other flat. The dipping reflector is shallower than the flat one. To avoid artifacts caused by velocity discontinuities in the migration velocity the reflectors were modeled as thin high-velocity (1.5 km/s) layers in a constant velocity medium (1 km/s). The migration velocity was set to constant, and equal to the background velocity. Therefore, the deeper reflector is slightly undermigrated.

Figure 1 is the image obtained applying the conventional imaging principle [equation (1)]. The dipping reflector is properly imaged within the range that is illuminated by the shots. The flat reflector is slightly undermigrated, as mentioned above.

Figure 2 shows on the left the offset-domain CIG (a) and on the right the angle-domain CIG (b). The CIGs are located at a surface location where both reflectors are illuminated well (1,410 m). As expected, the image is nicely focused at zero-offset in the panel on the left and the events are flat in the panel on the right.

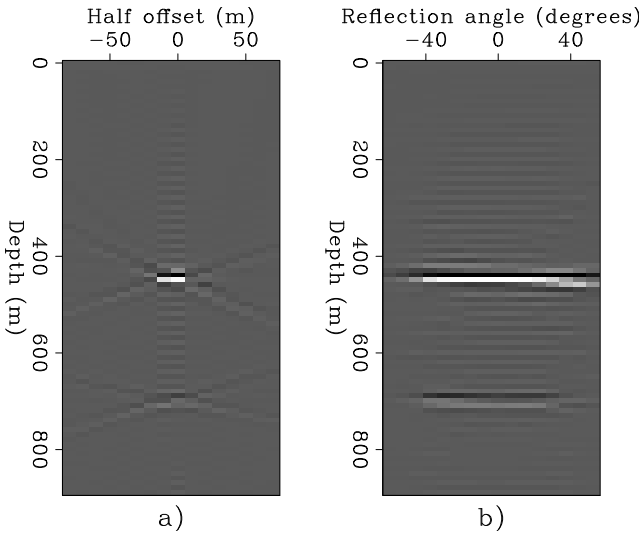



Figure 2: Offset-domain CIG (left a)) and angle-domain CIG (right b)) corresponding to the image in Figure 1. Notice the focusing at zero offset in a), and the flatness of the moveout in b). biondo2-Shot-Cig-Ang-dip [CR]

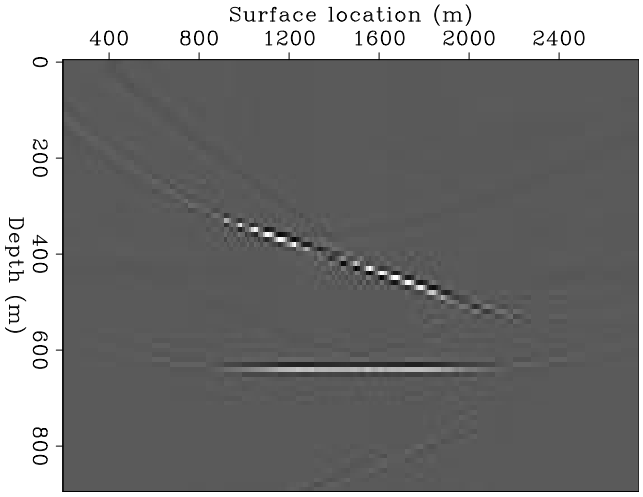



Figure 3: Image of the synthetic data set with the incorrect velocity function. biondo2-Shot-Image-dip-slow [CR]

To illustrate the usefulness for velocity updating of the proposed method to compute ADCIGs, I have migrated the same data set with a lower velocity ( .909 km/s). Figure 3 is the "stacked" image obtained using the lower velocity. Both reflectors are undermigrated and shifted upward.

Figure 4 shows on the left the offset-domain CIG (a) and on the right the angle-domain CIG (b). Now in the panel on the left, the energy is not focused at zero offset, but it is spread over an hyperbolic trajectory centered at zero offset. The corresponding ADCIG (right) shows the characteristic smile typical of an undermigrated ADCIG. The velocity information contained in the panel on the right can be easily used for velocity updating and tomographic inversion in a similar way as the ADCIG obtained by downward-continuation migrations are used (Clapp and Biondi, 2000; Clapp, 2001).

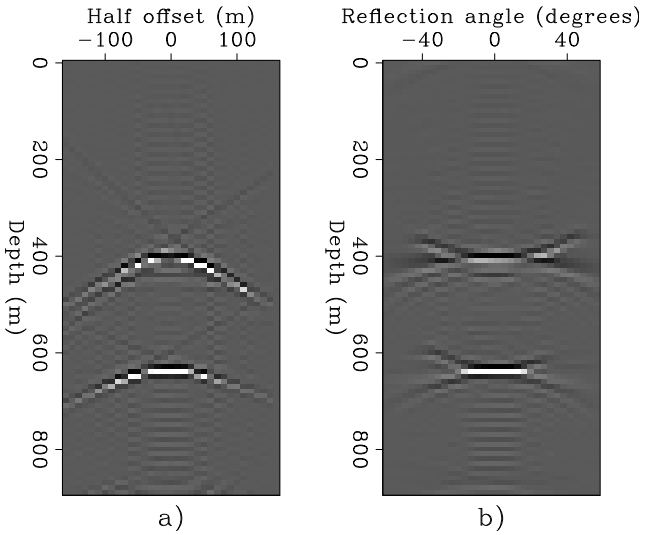



Figure 4: Offset-domain CIG (left) and angle-domain CIG (right) corresponding to the image in Figure 3. Notice the lack of focusing at zero offset in a), and the smile in b). biondo2-Shot-Cig-Ang-dip-slow [CR]

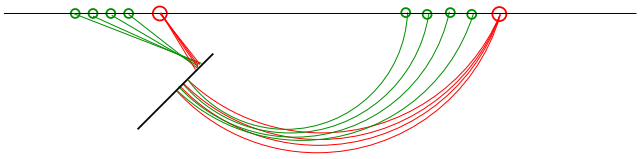


Figure 5: Ray paths corresponding to the reflection generated above the reflector and the one generated below the reflector. The rays corresponding to the source wavefield are red (dark in B&W), lines represent the wavefronts and the rays corresponding to the receiver wavefield are green. (light in B&W), biondo2-imag-rays [NR]

## PRESTACK IMAGES OF OVERTURNED REFLECTIONS

One of the main advantages of reverse-time migration methods over downward-continuation migration methods is their capability of imaging overturned events, even in presence of lateral velocity variations. However, this potential has not been exploited yet for prestack migration for several reasons. The computational cost is an important obstacle that is slowly being removed by progress in computer technology. In this section I will address two more fundamental problems. First, discrimination between image contributions from reflections generated above an interface from the image contributions from reflections generated below the same interface. These two reflections have usually opposite polarities because they see the same impedance contrast from opposite directions. If their contributions to the image are simply stacked together, they would tend to attenuate each other. Second, the updating of the migration velocity from overturned reflections. Solving the first problem is crucial to the solution of the second one, as graphically illustrated in Figure 5. This figure shows the raypaths for both events. It is evident that the overturned event passes through an area of the velocity field different from the area traversed by the reflection from above. The information on the required velocity corrections, provided by the image obtained using a given velocity function, can thus be inconsistent for the two reflections, even showing errors with opposite signs.

The reflection from above and the reflection from below can be discriminated by a simple generalization of the imaging principle expressed in equation (2), that includes a time lag $\tau$ in the crosscorrelation. To understand this generalization it is useful to review the process of image formation in reverse time migration. Figures 6–8 sketches this process at three different values in the propagation time $t$. For simplicity, the sketches represent the process for the familiar reflection from above, but similar considerations would hold also for the reflection from below. The reddish (dark in B&W) lines represent the wavefronts for the source wavefield. The greenish (light in B&W) lines represent the wavefronts for the receiver wavefield. At time $t-dt$ (Figure 6) the two wavefronts do not intersect, and thus they do not contribute to the crosscorrelation. At time $t$ (Figure 7) the two wavefronts begin to interfere, and thus they begin to contribute to the image. The contribution starts in the middle of the reflector at time $t$, and then it moves to the sides as the time progresses to $t+dt$ (Figure 8). The process described above correlates the wavefields at the same time $(t-dt, t, t+dt)$. However, the

wavefields can also be correlated at a non-zero lag over the time axis. In mathematical terms, we can generalize equation (2) as

$$I(z,x,h,\tau) = \sum_t S\left(t+\frac{\tau}{2},z,x+\frac{h}{2}\right) R\left(t-\frac{\tau}{2},z,x-\frac{h}{2}\right). \tag{4}$$

Now the image is function of an additional variable $\tau$, that represents the correlation lag in time. Figures 9–10 provide an intuitive understanding of the outcome of the correlation for $\tau = -dt/2$ (Figure 9) and $\tau = dt/2$ (Figure 10). In both cases the two wavefields interfere and they contribute to the image. For negative $\tau$ the image is slightly above the correct location of the reflector, and for positive $\tau$ the image is slightly below it. Therefore, the image of the reflector slowly moves downward (more precisely along the normal to the reflector) as $\tau$ increases. The crucial point is that for reflections generated from below, this movement is in the opposite direction (i.e., upward). This difference in propagation direction allows an easy discrimination of the two reflections by filtering the image according to the propagation direction as $\tau$ progresses. I have not implemented such a filtering yet, but it should be relatively straightforward.

I have confirmed this intuitive understanding by applying the generalized imaging condition in equation (4) to both the synthetic data set described above, and a synthetic data set with overturned events. I describe the results of the test on the latter in the next section.

**Examples of prestack imaging of overturned reflections**

To illustrate the use of the proposed method to image overturned reflections I created a simple synthetic data set that contains such events. I immersed a thin high-velocity segment in a layered medium with a strong vertical velocity gradient ($.97s^{-1}$). The reflector is dipping 44 degrees with respect to the vertical and extends from a surface coordinate of 1 km to a surface coordinate of 1.35 km.

I modeled and migrated 20 shots spaced 50 m apart, starting from the surface coordinate of 4.5 km. The receivers were in a symmetric split-spread configuration with maximum offset of 6.4 km. Because of the relative position of the reflector with respect to the shots, only the overturned reflections illuminate the reflector.

Figure 11 is the image obtained by applying the conventional imaging principle; that is; evaluating equation (4) at $\tau = 0$. The dark segment superimposed onto the image shows the position of the reflector in the model. The reflector is properly focused and positioned correctly. Figure 12 shows the image at $\tau = -.00525s$ (top) and the image at $\tau = .00525s$ (bottom). As in Figure 11, the dark segment superimposed onto the images shows the position of the reflector in the model. In these two panels the reflector is almost as well focused as in Figure 11, but it is slightly shifted along its normal. As expected from the theoretical discussion above, the reflector is slightly lower for the negative $\tau$ (top) than for the positive $\tau$ (bottom).

Figure 13 shows an example of CIG computed by evaluating equation (4) at $\tau = 0$. The panel on the left (a) shows the offset-domain CIG, and the panel on the right (b) shows the

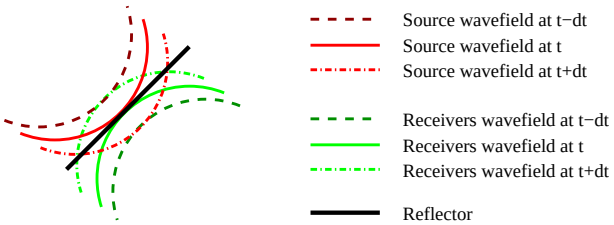



Figure 6: Wavefronts for the source wavefield (red), and the receiver wavefield (green), for three time steps $(t-dt, t, t+dt)$. The wavefronts at $t-dt$ are highlighted in darker color. The two highlighted wavefronts do not intersect, and thus their contribution to the image is null. biondo2-image-wave-st-m1 [NR]

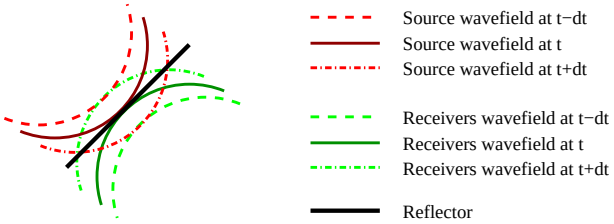



Figure 7: Wavefronts as in Figure 6. The two highlighted wavefronts (time $t$) intersect in the middle of the reflector, and they contribute to the image at the intersecting point. biondo2-image-wave-st-0 [NR]

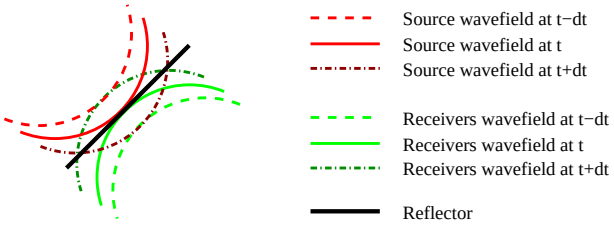



Figure 8: Wavefronts as in Figure 6. The two highlighted wavefronts (time $t+dt$) intersect at the edges of the reflector, and they contribute to the image at the intersecting points. biondo2-image-wave-st-p1 [NR]

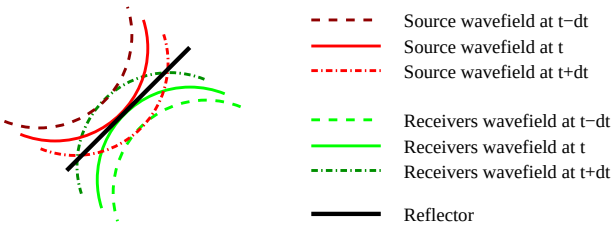



Figure 9: Wavefronts as in Figure 6. The two highlighted wavefronts (source wavefront at time $t-dt$ and the receiver wavefront at time $t+dt$) intersect above the reflector, and they contribute to the image at the intersecting point. biondo2-image-wave-lag-m1 [NR]

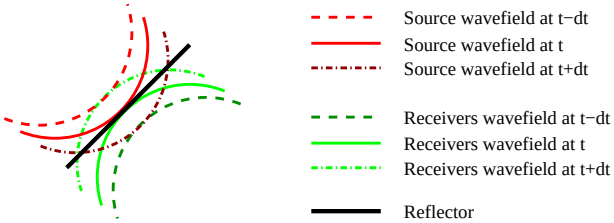



Figure 10: Wavefronts as in Figure 6. The two highlighted wavefronts (source wavefront at time $t+dt$ and the receiver wavefront at time $t-dt$) intersect below the reflector, and they contribute to the image at the intersecting point. biondo2-image-wave-lag-p1 [NR]

angle-domain CIG. The energy is correctly focused at zero offset in a), and the event is flat in b), though the angular coverage is narrow because of the short range of shot locations (1 km).

The second obstacle to image overturned reflections is the estimation of a velocity model that focuses and positions them correctly. To investigate this issue, I migrated the same data set with two inaccurate velocity functions. The first is 1% slower than the correct one, and the second is 1% faster than the correct one. Figure 14 shows the "stacked" images produced by these two migrations. The panel on the top (a) shows the image when the velocity is too low, and the panel on the bottom (b) shows the image when the velocity is too high. As in the previous figures, the dark segment superimposed onto the images shows the position of the reflector in the model. As expected the reflector is mispositioned and not as well focused as in Figure 11.

Figure 15 shows the CIG gathers taken at the same location as in Figure 13 for the migration with the low velocity and Figure 15 shows the CIG gathers for the migration with the high velocity. Notice that the velocity errors have caused a shift along the offset direction of the focal point in the offset-domain gathers. Towards positive offsets for the low velocity (Figure 15) and towards negative offsets for the high velocity (Figure 16). The angular coverage is too limited to notice a clear pattern in the angle-domain CIGs. In principle, a lateral shift in offset-domain gather should correspond to a tilt in the angle-domain gathers. These results seems to indicate that the residual moveout is asymmetric for overturned reflections, contrary to the symmetric moveout caused by velocity errors for regular reflections (see Figure 4). A more definitive analysis requires the migration of a survey with wider angular coverage; that is, with wider shot range. However, this characteristic suggests that for updating the velocity from overturned waves, the migrated CIGs should be scanned using a different family of residual moveouts than the parabolic moveouts used for standard reflections. The shift in the focal point in the offset-domain gathers could be also directly used for updating the velocity along the path of the overturned reflections.

## PRESTACK IMAGES OF PRISMATIC REFLECTIONS

Prismatic reflections are another class of reflections that can be potentially useful for imaging complex data. They usually include at least a segment of the ray path that is propagating upward. Therefore, downward continuation method are not appropriate for imaging them.

To test the possibility of using reverse time migration to image prismatic reflection I created another synthetic data set. Figure 17 shows the velocity model used for generating the synthetic. The solid line superimposed onto the velocity is an example of prismatic-reflection ray path. I modeled and migrated 250 shots spaced 6.25 m apart, starting from the surface coordinate of 1,375 m. The receivers were in a symmetric split-spread configuration with maximum offset of 6,400 m.

To better analyze the behavior of reverse time migration in imaging prismatic reflections, I removed the primary reflections from the data. To achieve this goal I subtracted two other

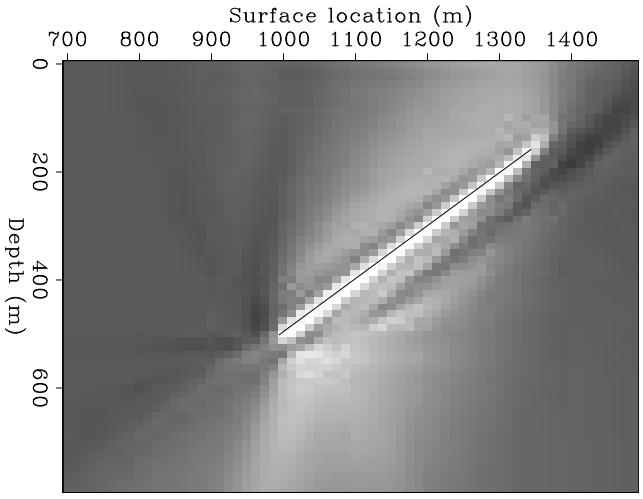



Figure 11: Image of the synthetic data set containing the overturned reflections migrated with the correct velocity and at $\tau = 0$. The dark segment superimposed onto the images shows the position of the reflector in the model. biondo2-Shot-Refl-Image-vover [CR]

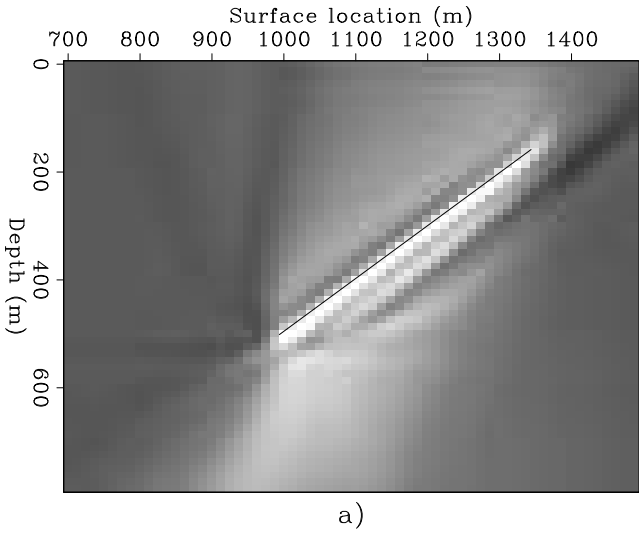



Figure 12: Images of the synthetic data set containing the overturned reflections migrated with the correct velocity; at $\tau = -dt/2$ (top) and at $\tau = dt/2$ (bottom) The dark segment superimposed onto the images shows the position of the reflector in the model. Notice the slight downward shift of the imaged reflector in a) and the slight upward shift of the imaged reflector in b). biondo2-Shot-Refl-Image-mov-vover [CR]

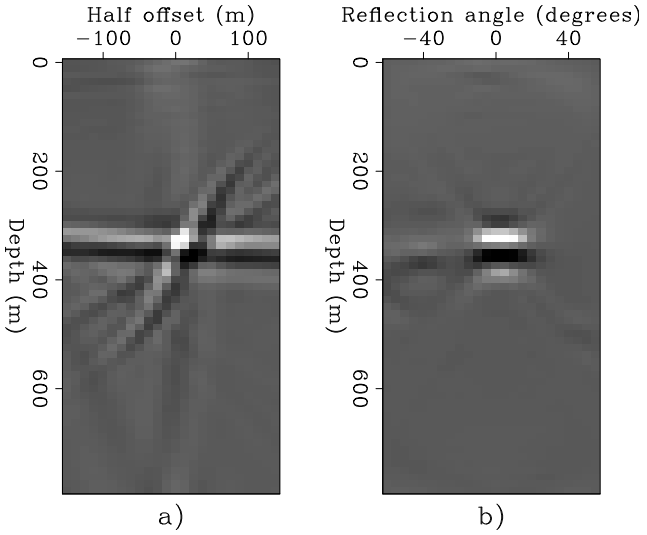



Figure 13: Offset-domain CIG (left) and angle-domain CIG (right) corresponding to the image in Figure 11. Notice the focusing at zero offset in a), and the flatness of the moveout in b), though the angular coverage is narrow because of the short range of shot locations (1 km). biondo2-Shot-Cig-Ang-vover [CR]

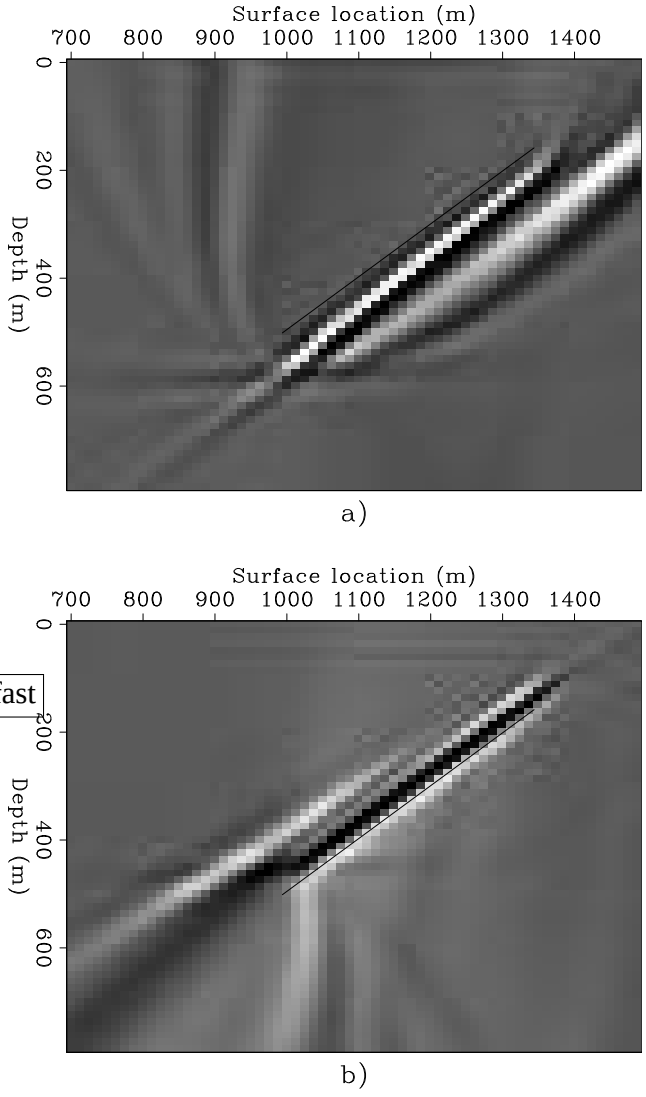



Figure 14: Image of the synthetic data set containing the overturned reflections migrated with a velocity function 1% lower than the correct one (top) and with a velocity function 1% higher than the correct one (bottom). Notice the misfocusing and mispositioning of the reflector. biondo2-Shot-Refl-Image-vover-slow-fast [CR]

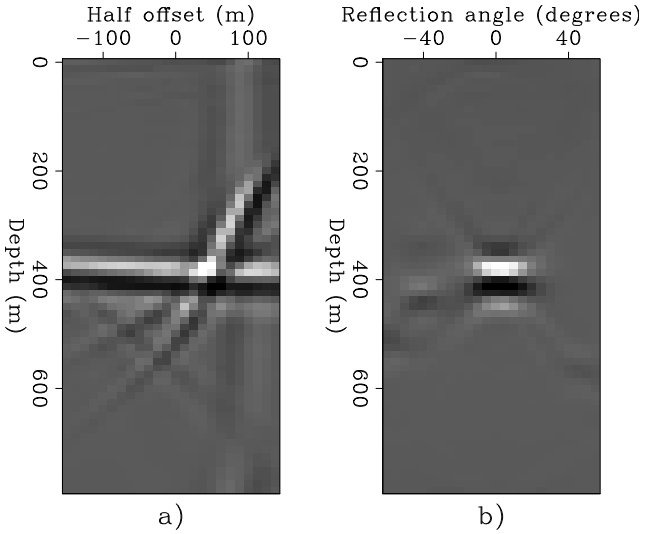



Figure 15: Offset-domain CIG (left) and angle-domain CIG (right) corresponding to the image in Figure 14a. Notice the positive shift along the offset direction of the focal point in the offset-domain gathers. biondo2-Shot-Cig-Ang-vover-slow [CR]

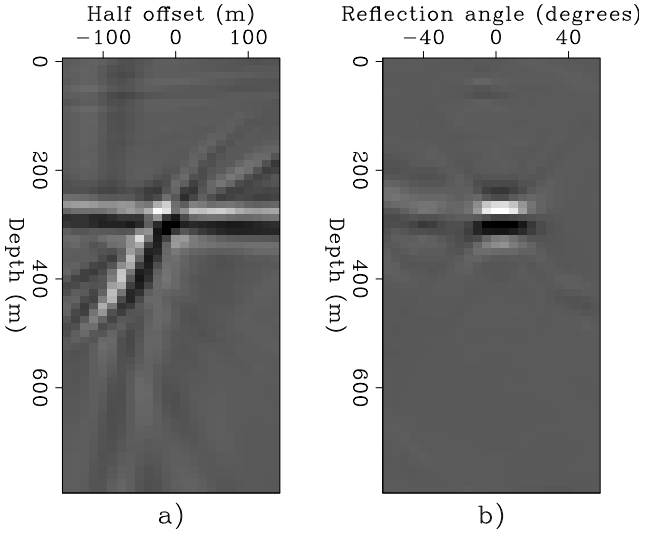



Figure 16: Offset-domain CIG (left) and angle-domain CIG (right) corresponding to the image in Figure 14b. Notice the negative shift along the offset direction of the focal point in the offset-domain gathers. biondo2-Shot-Cig-Ang-vover-fast [CR]

data sets from the data modeled assuming the velocity model shown in Figure 17. One of the data sets included the primary reflections from the flat reflector, and the other one the primaries from the dipping reflector. The two velocity models used to generate the primaries were created by breaking the black area shown in Figure 17 into two pieces, divided by a vertical line passing through the location of the corner (1.4 km).

To image prismatic reflections, the background velocity function needs to include at least one of the interfaces that generated the reflections. I assumed that the flat reflector was known, and thus I included a flat velocity interface at 600 m depth in the migration velocity model. Figure 18 is the image obtained migrating all the 250 shots. As expected by simple analysis of the raypaths, the prismatic reflections illuminate the dipping reflector more strongly in the deeper part than in the shallower part. The image shows also a high energy flat event on the right side of the corner, that is not a proper image of a reflection. In theory, such event should not be there. The primaries associated with the flat reflector were removed from the data. The prismatic reflections bouncing off the dipping layer are not imaged because the dipping layer is not present in the migration velocity. This artifact is likely to be caused by the (imperfect) correlation of the prismatic reflection (solid line in Figure 17) with reflections generated by the horizontal discontinuity in the migration velocity model (dashed line in Figure 17). This interpretation is corroborated by the presence of a faint low-frequency noise extending upward from the flat reflector.

Figure 19 shows on the left the offset-domain CIG (a) and on the right the angle-domain CIG (b). The CIGs are located at a surface location where prismatic reflections illuminate the dipping reflector (1,350 m). The energy focuses at zero offset, but the angle-domain CIG is not flat. This lack of flatness is likely to be caused by the fact that the source and receiver wavefields meet at the reflector when propagating along opposite vertical direction. I speculate that better angle-domain CIGs could be generated if I applied the generalized imaging condition expressed in equation (3), that includes the vertical subsurface offset.

## CONCLUSIONS

I presented a method to compute Common Image Gathers from reverse-time shot profile migration. The proposed method generates accurate CIGs that can be used to update the migration velocity function, both for regular reflections and for overturned reflections.

The reflections generated from either side of an interface can be discriminated by computing the crosscorrelation of the source wavefield with the receiver wavefield at the non-zero time lag. This techniques is useful for imaging both overturned reflections and prismatic reflections.

The migration of a synthetic example containing prismatic reflections suggests that imaging this kind of reflections might be prone to artifacts and may require the computation of the CIGs with the even more general imaging condition that includes vertical subsurface offset.

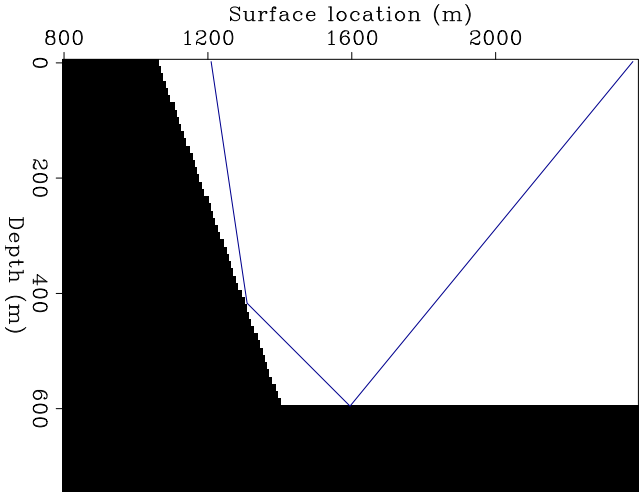



Figure 17: Velocity model assumed to create the synthetic data set. Superimposed onto the velocity is an example of prismatic-reflection ray path (solid line). biondo2-Slow-prism-ann [NR]

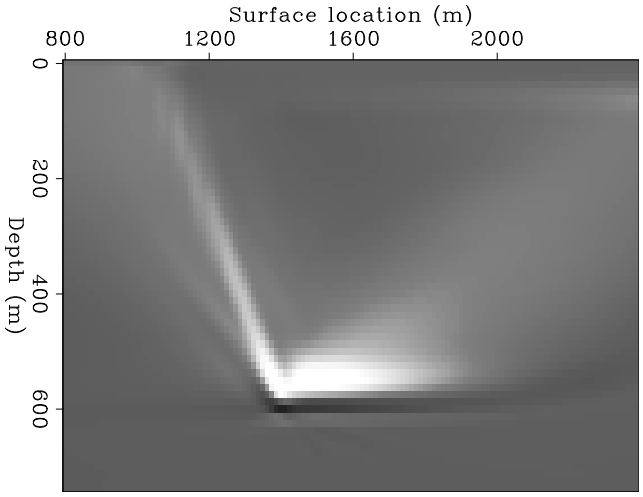



Figure 18: Image of the synthetic data set containing the overturned reflections migrated with the correct velocity and at $\tau = 0$. biondo2-Shot-Image-prism [CR]

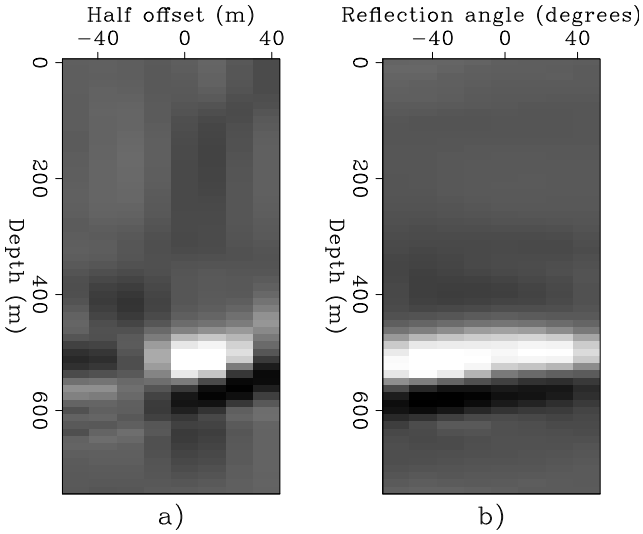



Figure 19: Offset-domain CIG (left) and angle-domain CIG (right) corresponding to the image in Figure 18. Notice that the energy focuses at zero offset, but that the angle-domain CIG is not flat. This lack of flatness is likely to be caused by the fact that the source and receiver wavefields meet at the reflector when propagating along opposite vertical direction. biondo2-Shot-Cig-Ang-prism [CR]

## REFERENCES

Baysal, E., Kosloff, D. D., and Sherwood, J. W. C., 1983, Reverse time migration: Geophysics, **48**, no. 11, 1514–1524.

Baysal, E., Kosloff, D. D., and Sherwood, J. W. C., 1984, A two-way nonreflecting wave equation: Geophysics, **49**, no. 02, 132–141.

Biondi, B., 2002, Reverse time migration in midpoint-offset coordinates: SEP–**111**, 147–155.

Broto, K., and Lailly, P., 2001, Towards the tomographic inversion of prismatic reflections: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 726–729.

Clapp, R., and Biondi, B., 2000, Tau domain migration velocity analysis using angle CRP gathers and geologic constrains: 70th Ann. Internat. Mtg., Soc. of Expl. Geophys., Expanded Abstracts, 926–929.

Clapp, R. G., 2001, Geologically constrained migration velocity analysis: Ph.D. thesis, Stanford University.

Etgen, J., 1986a, High order finite-difference reverse time migration with the two way nonreflecting wave equation: SEP–**48**, 133–146.

Etgen, J., 1986b, Prestack reverse time migration of shot profiles: SEP–**50**, 151–170.

Filho, C. A. C., 1992, Elastic modeling and migration in earth models: Ph.D. thesis, Stanford University.

Li, Z., Claerbout, J. F., and Ottolini, R., 1983, Overturned-wave migration by two-way extrapolation: SEP–**38**, 141–150.

Rickett, J., and Sava, P., 2001, Offset and angle domain common-image gathers for shot-profile migration: 71st Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1115–1118.

Sava, P., Biondi, B., and Fomel, S., 2001, Amplitude-preserved common image gathers by wave-equation migration: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 296–299.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 139–149*

# Statistical stability and time-reversal imaging in random media

*James G. Berryman,*[1] *Liliana Borcea,*[2] *George C. Papanicolaou,*[3] *and Chrysoula Tsogka*[4]

**ABSTRACT**

Localization of targets imbedded in a heterogeneous background medium is a common problem in seismic, ultrasonic, and electromagnetic imaging problems. The best imaging techniques make direct use of the eigenfunctions and eigenvalues of the array response matrix, as recent work on time-reversal acoustics has shown. Of the various imaging functionals studied, one that is representative of a preferred class is a time-domain generalization of MUSIC (MUltiple SIgnal Classification), which is a well-known linear subspace method normally applied only in the frequency domain. Since statistical stability is not characteristic of the frequency domain, a transform back to the time domain after first diagonalizing the array data in the frequency domain takes optimum advantage of both the time-domain stability and the frequency-domain orthogonality of the relevant eigenfunctions.

## INTRODUCTION

There have been many approaches to estimating target location using seismic, ultrasonic, and electromagnetic imaging methods. Some of the most popular ones in recent years continue to be matched-field processing (Bucker, 1976; Jensen *et al.*, 1994), MUSIC (MUltiple SIgnal Classification) (Schmidt, 1979; Johnson, 1982; Schmidt, 1986; Biondi and Kostov, 1989), and other linear subspace methods (Johnson, 1982; Johnson and DeGraaf, 1982; Cheney, 2001). When the targets are imbedded in heterogeneous media so that significant multiple scattering occurs in the background medium during wave propagation between array and target, the randomness has a different character than that usually envisioned in these traditional analyses. Yet there are a great many applications (Fink, 1997; 1999; Fink *et al.*, 2000; Fink, 2001; Fink and Prada, 2001; ter Haar, 2001) ranging from the biomedical to ocean acoustics to nondestructive evaluation, where imaging is important and where sources of randomness not associated with the imaging targets can wreak havoc with the traditional methods. Time-reversal acoustics (Fink *et al.*, 1989; Jackson and Dowling, 1991; Prada and Fink, 1991) offers part of the answer to these difficult imaging questions, and some significant improvements over

[1]**email**: berryman@sep.stanford.edu

[2]Computational and Applied Mathematics, MS 134, Rice University, 6100 Main Street, Houston, TX 77005-1892. **email**: borcea@caam.rice.edu

[3]Mathematics Dept., Stanford University, Stanford, CA 94305. **email**: papanico@georgep.stanford.edu

[4]CNRS/LMA, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20, FRANCE. **email**: tsogka@lma.cnrs-mrs.fr

these methods for imaging in random media are summarized here.

We have found that methods designed to work well for finding targets in homogeneous media do not necessarily work very well for targets imbedded in random media. In particular, the fact that the linear subspace methods are normally applied in the frequency domain combined with the fact that statistically stable methods are normally found only in the time domain, forces us to seek different imaging strategies in the random media imaging problems of interest to us here. We find that a set of imaging functionals having the desired characteristics exists, and furthermore that the properties of this set can be completely understood when the time-domain self-averaging — that gives rise to the required statistical stability of the target images — is taken properly into account. We can largely eliminate the undesirable features of the frequency domain methods by making a transform back to the time domain after first diagonalizing sensor array data. While the frequency domain analysis takes optimum advantage of eigenfunction orthogonality of the array data, a transform to the time-domain takes optimum advantage of wave self-averaging which then leads to the statistical stability we require for reliable and repeatable imaging in random media.

We first introduce the imaging problem in the next section. Then we summarize our technical approach. Examples of the cross-range (or bearing) estimates obtained with these methods are presented and then combined with range information from time-delay data to obtain our best estimates and images of target location. The final section summarizes our conclusions about the methods discussed.

## IMAGING PROBLEM

Our analysis assumes that the array has $N$ transducers located at spatial positions $\mathbf{x}_p$, for $p = 1, \ldots, N$. (See Figure 1.) When used in active mode, the array probes the unknown acoustic medium containing $M$ small scatterers by emitting pulses and recording the time traces of the back-scattered echos. We call the resulting data set the multistatic array response (or transfer) matrix

$$P(t) = \big(P_{pq}(t)\big), \tag{1}$$

where $p$ and $q$ both range over all the array elements. For our simulations, we consider a linear array where two adjacent point transducers are a distance $\lambda/2$ apart, with $\lambda$ being the carrier (central) wavelength of the probing pulses. Such an arrangement ensures that the collection of transducers behaves like an array having aperture $a = (N-1)\lambda/2$ and not like separate entities, while keeping the interference among the transducers at a minimum (Steinberg, 1983). Our goal is to detect and then localize all $M$ of the targets in the random medium, if possible.

For the numerical examples considered here, we will treat ultrasonic imaging problems. Our simulations assume that $\lambda \leq \ell << a = (N-1)\lambda/2 << L$, where $\lambda$ is the central wavelength, $\ell$ is a characteristic length scale of the inhomogeneity (like a correlation length), $a$ is the array aperture, and $L$ is the approximate distance to the targets from the array. This is the regime where multipathing, or multiple scattering, is significant even when the standard deviation of sound speed fluctuations is only a few percent. Values used in the codes

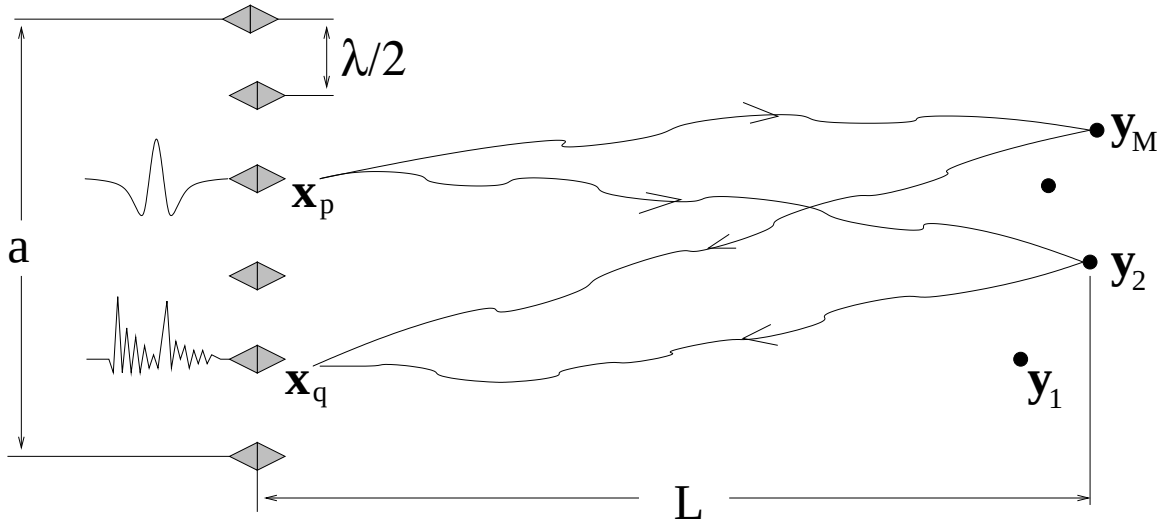



Figure 1: Array probing of a randomly inhomogeneous medium containing $M$ small scatterers. jim2-gen_setup [NR]

are $\lambda = 0.5$mm, $a = 2.5$mm, and a background wave speed of $c_0 = 1.5$km/s. More details concerning the simulations may be found in Borcea *et al.* (2002).

Typical array processing methods assume that the targets are far away from the array and, therefore, they look like points. Similarly, the propagation medium is assumed homogeneous and so the observed wavefronts scattered by the targets look like plane waves at the array. Array noise has usually been treated as due either to diffuse sources of white noise coming simultaneously from all directions, or to isolated "noise" having the same types of source characteristics as the targets of interest. But in random media with significant multiple scattering, the resulting "noise" cannot be successfully treated in these traditional ways.

Real-space time-reversal processing of the array response data involves an iterative procedure: sending a signal, recording and storing the scattered return signal, time-reversing and then rebroadcasting the stored signal, with subsequent repetitions. This procedure amounts to using the power method for finding the singular vector of the data matrix having the largest singular value. Alternatively, when the full response/transfer matrix has been measured for a multistatic active array, the resulting data matrix can be analyzed directly by Singular Value Decomposition (SVD) to determine not only the singular vector having the largest singular value, but all singular vectors and singular values — simultaneously (Prada and Fink, 1994; Prada *et al.*, 1996; Mordant *et al.*, 1999).

Imaging is always done using a fictitious medium for the simulated backpropagation that produces these images since the real medium is not known. Its large-scale features could be estimated from other information, such as geological data obtained by seismic methods. For example, migration methods (Claerbout, 1976; Aki and Richards, 1980; Bleistein *et al.*, 2001) can be used, where very large arrays — much larger than those we contemplate using here — are required. However, the small-scale random inhomogeneities are not known and cannot

be effectively estimated, so the simplest thing to do is ignore them when imaging, and use methods that are statistically stable and therefore insensitive to the exact character of these small inhomogeneities.

## TECHNICAL APPROACH

In our simulations, the array response matrix $\widehat{P}(\omega)$ [see definition in (1)] in the frequency domain is symmetric but not Hermitian. In general (as for array elements with nonisotropic radiation patterns), it is neither Hermitian nor symmetric, but with slight modifications our methods apply to this case as well. The eigenvectors of $\widehat{P}(\omega)\widehat{P}^H(\omega)$ having unit norm are denoted by $\widehat{\mathbf{U}}_r(\omega)$, for $r = 1,\ldots,N$. The eigenvalues of $\widehat{P}(\omega)\widehat{P}^H(\omega)$ are $\sigma_r^2(\omega)$, with $\sigma_r(\omega)$ being the singular values of $\widehat{P}(\omega)$. The significant singular vectors $\widehat{\mathbf{U}}_r(\omega)$ [*i.e.*, those in the range of $\widehat{P}(\omega)$] have singular values $\sigma_r(\omega) > 0$ for $1 \le r \le M$, where $M$ is either the number of targets, or the size of the array ($N$) — whichever is smaller. We assume that the number of targets is smaller than the array size $N$, so that $M$ is in fact the number of distinguishable targets; this assumption is required by the imaging methods we employ (such as MUSIC) as will become clear while presenting the method.

The notation used here is the same as in Borcea *et al.* (2002). We denote by $\widehat{\mathbf{g}}_0(\mathbf{y},\omega)$ the deterministic source vector observed at the array for a source located at $\mathbf{y}^s$. Then, $\widehat{\mathbf{g}}_0(\mathbf{y},\omega)$ is given by

$$\widehat{\mathbf{g}}_0(\mathbf{y}^s,\omega) = \begin{pmatrix} \widehat{G}_0(\mathbf{y}^s,\mathbf{x}_1,\omega) \\ \widehat{G}_0(\mathbf{y}^s,\mathbf{x}_2,\omega) \\ \vdots \\ \widehat{G}_0(\mathbf{y}^s,\mathbf{x}_N,\omega) \end{pmatrix}, \tag{2}$$

where $\widehat{G}_0(\mathbf{y}^s,\mathbf{x}_j,\omega)$ is the deterministic two-point Green's function, and $\mathbf{x}_j$ is the location of the $j$-th array element.

We also define the projection $\mathcal{P}_N\widehat{\mathbf{g}}_0(\mathbf{y},\omega)$ of $\widehat{\mathbf{g}}_0(\mathbf{y}^s,\omega)$ onto the null-space of $\widehat{P}\widehat{P}^H(\omega)$ by

$$\begin{aligned} \mathcal{P}_N\widehat{\mathbf{g}}_0(\mathbf{y}^s,\omega) = \widehat{\mathbf{g}}_o(\mathbf{y}^s,\omega) \\ -\sum_{r=1}^{M}\left[\widehat{\mathbf{U}}_r^H(\omega)\widehat{\mathbf{g}}_o(\mathbf{y}^s,\omega)\right]\widehat{\mathbf{U}}_r(\omega), \end{aligned} \tag{3}$$

for each frequency in the support of the probing pulse $\widehat{f}(\omega)$.

The method we describe here is a time domain variant of MUSIC (Schmidt, 1979; 1986; Cheney, 2001; Devaney, 2002) which we label $DOA$, because it gives very stable estimates of the direction of arrival. Frequency domain MUSIC takes a replica (or trial) vector, which is the impulse response or Green's function for a point source at some point in the space, and dots this vector into an observed singular vector at the array. With appropriate normalization, this dot product acts like a direction cosine of the angle between the replica vector and the data vector. If the sum of the squares of these direction cosines is very close to unity, then it is correct to presume that the source point of that replica vector is in fact a target location

since it lies wholely in the range of the array response matrix. Crudely speaking, imaging is accomplished by plotting $1/[1-\cos^2(\cdot)]$, which will have a strong peak when the replica source point is close to the target location.

We form the sum

$$\mathcal{G}^{(j)}(\mathbf{y}^s) = \sum_{p=1}^{N} \left| \mathcal{F}_p^{(j)}\left(\mathbf{y}^s, t_p(\mathbf{y}^s)\right)\right|^2, \tag{4}$$

with

$$\begin{array}{l} \boldsymbol{\mathcal{F}}^{(j)}(\mathbf{y}^s,t) = \int e^{-i\omega t}\sigma_j(\omega)\widehat{\mathbf{g}}_o(\mathbf{y}^s,\omega)d\omega \\ -\int e^{-i\omega t}\sigma_j(\omega)\sum_{r=1}^{M}\left[\widehat{\mathbf{U}}_r^H(\omega)\widehat{\mathbf{g}}_o(\mathbf{y}^s,\omega)\right]\widehat{\mathbf{U}}_r(\omega)d\omega, \end{array} \tag{5}$$

and display the objective functional

$$\mathcal{R}_{\mathrm{DOA}}(\mathbf{y}^s) = \sum_{j=1}^{M} \frac{\min_{\mathbf{y}^s} \mathcal{G}^{(j)}(\mathbf{y}^s)}{\mathcal{G}^{(j)}(\mathbf{y}^s)}, \tag{6}$$

for points $\mathbf{y}^s$ in the target domain.

The arrival time $t_p(\mathbf{y}^s)$ is the deterministic travel time from the $p$-th transducer to the search point,

$$t_p(\mathbf{y}^s) = \frac{|\,\mathbf{x}_p - \mathbf{y}^s\,|}{c_0}. \tag{7}$$

## EXAMPLES AND RANGE ESTIMATION

Examples for frequency-domain MUSIC with two targets are displayed in Fig. 2. It is clear from this Figure that no range information is obtained from frequency-domain objective functionals, and even the cross-range information is often quite haphazard in random media. Lack of statistical stability prevents these imaging approaches from being useful in random media with significant multipathing as considered here. When the realization of the random medium is changed, the images obtained typically change also — which is what we mean by the phrase "lack of statistical stability" for these methods. Note that this approach works well for homogeneous media, but quickly breaks down when randomness of the velocity field is important.

Examples for time-domain MUSIC with two targets are displayed in Fig. 3. The cross-range results show dramatic improvement over results using other methods (Berryman *et al.*, 2002). Range information is still not to be found here, due to loss of coherence in the random medium; we cannot get exact cancellation at the targets in this situation whereas coherent refocusing is possible in homogeneous media. But the statistical stability of the universal "comet tails" — which was also anticipated by recent theoretical analyses (Blomgren *et al.*, 2002) — is now easily observed. The images are necessarily shown for specific realizations, but the results do not change significantly when the underlying realization of the random medium is changed. This fact has been repeatedly shown in our simulations, and is the main operational characteristic of statistically stable methods.

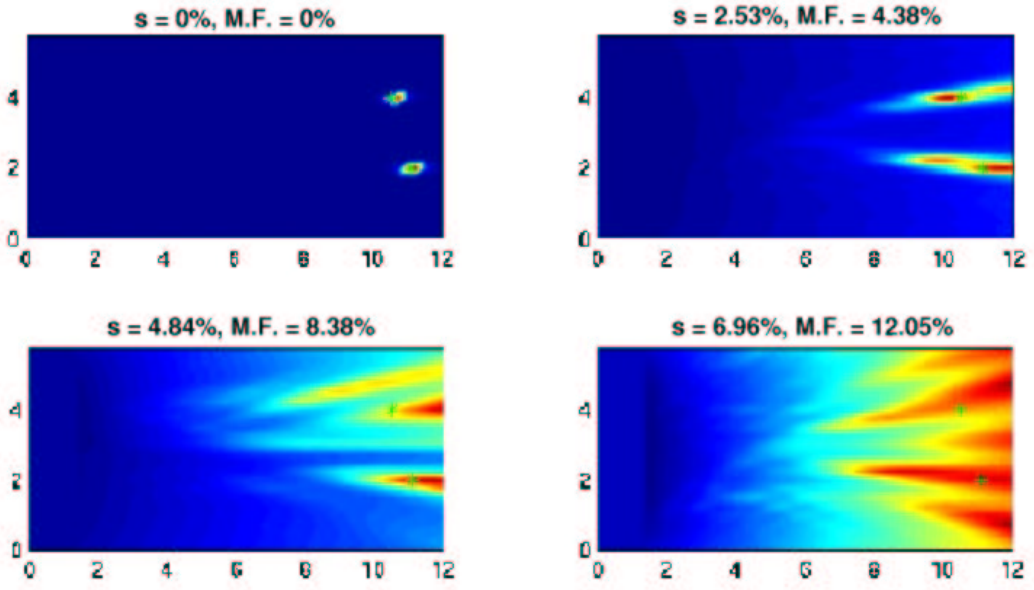



Figure 2: The MUSIC central frequency estimate of the location of two targets in random media with different strength of the fluctuations of the sound speed. The exact location of the targets is denoted by green stars (in the color version). The standard deviation $s$ and maximum fluctuations (M.F.) are indicated on the top of each view. The horizontal axis is the range in mm and the vertical axis is the cross-range in mm. jim2-Two_short_MUSIC [NR]

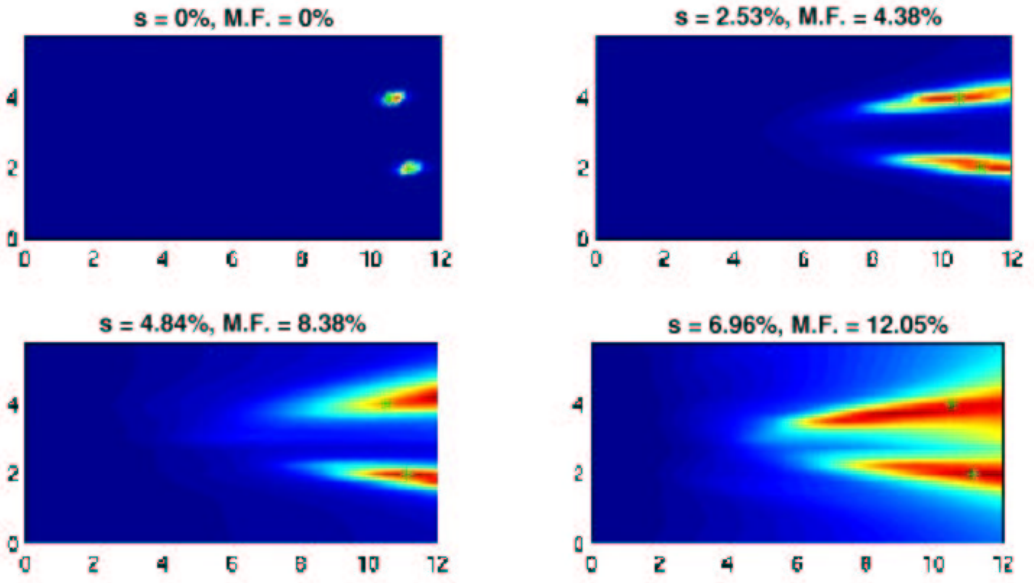



Figure 3: The DOA estimate (5) of the location of two target in random media with different strength of the fluctuations of the sound speed. The exact location of the target is denoted by the green star (in the color version). The standard deviation $s$ and maximum fluctuations (M.F.) are indicated on the top of each view. The horizontal axis is the range in mm and the vertical axis is the cross-range in mm. jim2-Two_short_DOA [NR]

Target localization requires an estimate of the range. In the far field, only the arrival time information is useful for this purpose. Arrival time information is present in the singular vectors and can also be averaged (for the same random medium) using the multiple copies available in the array response matrix for random media — see Borcea *et al.* (2002) — to obtain very stable estimates of arrival times. We will now combine this approach with the time-domain methods to obtain well-localized images of the targets.

For each search point $\mathbf{y}_s$, we compute the objective functional

$$\mathcal{R}_{\rm SAT}(\mathbf{y}^s) = \sum_{j=1}^{M} \frac{\min_{\mathbf{y}^s} \mathcal{I}^{(j)}_{\rm SAT}(\mathbf{y}^s)}{\mathcal{I}^{(j)}_{\rm SAT}(\mathbf{y}^s)}, \tag{8}$$

where

$$\mathcal{I}^{(j)}_{\rm SAT}(\mathbf{y}^s) = \sum_{p=1}^{N} \left| \mathcal{F}^{(j)}_p(\mathbf{y}^s, t_p(\mathbf{y}^s)) \right|^2 \left[ \tau^{(j)}_p - t_p(\mathbf{y}^s) \right]^2. \tag{9}$$

Here $\boldsymbol{\mathcal{F}}^{(j)}(\mathbf{y}^s, t)$ is defined by (5), $t_p(\mathbf{y}^s)$, for $p = 1, \ldots, N$, are the deterministic arrival times given by (7) and $\tau^{(j)}_p$, for $p = 1, \ldots, N$, and $j = 1, \ldots, M$, are the computed arrival times. We call (8) the Subspace Arrival Time (SAT) estimator.

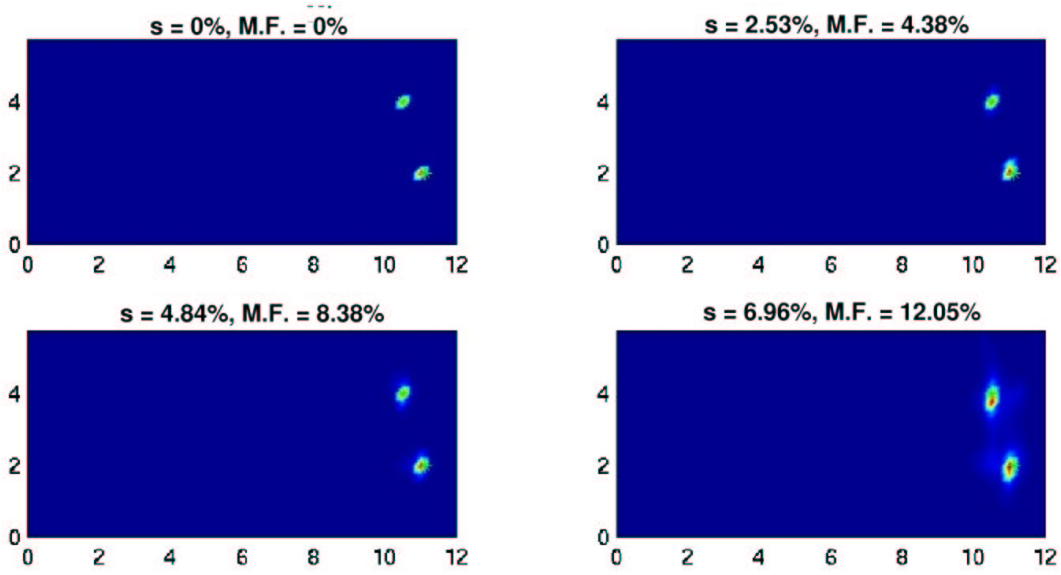



Figure 4: The SAT estimate for two targets. jim2-Two_short_SAT-1 [NR]

Examples of SAT (or time-domain MUSIC with arrival time estimates from the averaged singular vectors) for two targets are displayed in Fig. 4. This method is statistically stable and gives good estimates of the target locations. These localization results have degraded the least of all those considered (Borcea *et al.*, 2002; Berryman *et al.*, 2002) at the highest values of the random fluctuations.

## CONCLUSIONS

For imaging applications in randomly inhomogeneous acoustical media, the results summarized here lead us to the following conclusions: (1) Single frequency methods (including MUSIC) are not statistically stable, and therefore cannot be used without modification in the presence of significant amounts of spatial heterogeneity in the acoustic wave speed distribution. (2) In contrast, time domain methods are statistically stable for any objective functional having the characteristic that the random Green's functions appear in Hermitian conjugate pairs of $gg^*$ (Borcea *et al.*, 2002), because large random phases cancel precisely. This result has been shown here to be true for DOA, and is expected to be true more generally. (3) The DOA gives only cross-range information. Range information must be obtained separately.

To isolate the targets in random media, we need either multiple views (using multiple arrays) so we can triangulate, or we need to extract a direct measure of range from the data. In the SAT examples shown here, we used arrival time for the range estimation.

## ACKNOWLEDGMENTS

We thank Biondo Biondi, Jim Candy, Dave Chambers, and Bill Curry for helpful comments on the work and the presentation.

## REFERENCES

Aki, K., and Richard, P. G., 1980, *Quantitative Seismology: Theory and Methods*, Vol. II, Freeman, New York.

Berryman, J. G., Borcea, L., Papanicolaou, G. C., and Tsogka, C., 2002, Statistically stable ultrasonic imaging in random media: *J. Acoust. Soc. Am.*, , submitted for publication.

Biondi, B. L., and Kostov, C., 1989, High-resolution velocity spectra using eigenstructure methods: *Geophysics*, **54**, 832–842.

Bleistein, N., Cohen, J. K., and Stockwell, J. W., Jr., 2001, *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion*, Springer, New York.

Blomgren, P., Papanicolaou, G. C., and Zhou, H., 2002, Super-resolution in time-reversal acoustics: *J. Acoust. Soc. Am.*, **111**, 238–248.

Borcea, L., Papanicolaou, G. C., Tsogka, C., and Berryman, J. G., 2002, Imaging and time reversal in random media: *Inverse Problems*, , to appear.

Bucker, H. P., 1976, Use of calculated sound field and matched-field detection to locate sound sources in shallow water: *J. Acoust. Soc. Am.*, **59**, 368–373.

Cheney, M., 2001, The linear sampling method and the MUSIC algorithm: *Inverse Problems* **17**, 591–596.

Claerbout, J. F., 1976, *Fundamentals of Geophysical Data Processing with Applications to Petroleum Prospecting*, McGraw-Hill, New York.

Devaney, A. J., 2002, Super-resolution processing of multi-static data using time reversal and MUSIC: *J. Acoust. Soc. Am.*, , to appear.

Fink, M., 1997, Time reversal acoustics: *Physics Today*, **50** (3), 34–40.

Fink. M., 1999, Time-reversed acoustics: *Scientific American*, **281** (11), 91–97.

Fink, M., 2001, Chaos and time-reversed acoustics: *Physica Scripta*, **T90**, 268–277.

Fink. M., Cassereau, D., Derode, A., Prada, C., Roux, P., Tanter, M., Thomas, J.-L., and Wu, F., 2000, Time-reversed acoustics: *Rep. Prog. Phys.*, **63**, 1933–1995.

Fink, M., and Prada, C., 2001, Acoustic time-reversal mirrors: *Inverse Problems*, **17**, R1–R38.

Fink, M., Prada, C., and Wu, F., 1989, Self-focusing in inhomogeneous media with time reversal acoustics mirrors: in McAvoy, B. R., Ed., Proc. IEEE Ultrason. Symp. 1989, Vol. 2, pp. 681–686.

Jackson, D. R., and Dowling, D. R., 1991, Phase conjugation in underwater acoustics: *J. Acoust. Soc. Am.*, **89**, 171–181.

Jensen, F. B., Kuperman, W. A., Porter, M. B., and Schmidt, H., 1994, *Computational Ocean Acoustics*, AIP Press, New York.

Johnson, D. H., 1982, The application of spectral estimation methods to bearing estimation problems: *Proc. IEEE*, **70**, 1018–1028.

Johnson, D. H., and DeGraaf, S. R., 1982, Improving the resolution of bearing in passive sonar arrays by eigenvalue analysis: *IEEE Trans. Acoustics, Speech, Signal Proc.*, **ASSP-30**, 638–647.

Mordant, N., Prada, C., and Fink, M., 1999, Highly resolved detection and selective focusing in a waveguide using the D.O.R.T. method: *J. Acoust. Soc. Am.*, **105**, 2634–2642.

Prada, C., and Fink, M., 1994, Eigenmodes of the time reversal operator: A solution to selective focusing in multiple-target media: *Wave Motion*, **20**, 151–163.

Prada, C., Manneville, F., Spoliansky, D., and Fink, M., 1996, Decompostition of the time reversal operator: Detection and selective focusing on two scatterers: *J. Acoust. Soc. Am.*, **99**, 2067–2076.

Prada, C., Wu, F., and Fink, M., 1991, The iterative time reversal mirror: A solution to self-focusing in the pulse echo mode: *J. Acoust. Soc. Am.*, **90**, 1119–1129.

Schmidt, R. O., 1979, Multiple emitter location and signal parameter estimation: in Proc. RADC Spectrum Estimation Workshop, Rome, New York, October, 1979, Rome Air Development Center, 243–258.

Schmidt, R. O., 1986, Multiple emitter location and signal parameter estimation: *IEEE Trans. Antennas Propag.*, **AP-34**, 276–281.

Steinberg, B. D., 1983, *Microwave Imaging and Large Antenna Arrays*, Wiley, New York.

ter Haar, G., 2001, Acoustic surgery: *Physics Today*, **54** (12), 29–34.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 149–156*

# Short Note

# Reverse time migration in midpoint-offset coordinates

*Biondo Biondi*[1]

## INTRODUCTION

Reverse-time migration (Baysal et al., 1984) has some potential advantages with respect to downward-continuation migration. It can migrate overturned and prismatic reflections even in the presence of strong lateral velocity variations (Biondi, 2002). It also models the amplitude of the transmitted wavefield more accurately than downward-continuation in the presence of sharp interfaces. However, it has the drawback of being computationally intensive. In particular, shot profile migration of overturned events can be extremely expensive because we need to pad the computational domain with a huge number of zero traces to assure that it includes the reflectors that generated the overturned events.

A similar problem of computational-domain size exists for shot profile migration by downward continuation. For marine data, an efficient solution to the problem is migration in midpoint-offset coordinate (Biondi and Palacharla, 1996). In the midpoint-offset domain the wavefield focuses towards zero offset, and thus we can drastically limit the length of the offset axes (at the limit we can eliminate the cross-line offset altogether by a common-azimuth approximation). It is thus natural to try to derive a reverse-time migration method that backward propagates the data in midpoint-offset coordinates.

In this paper, I derive an equation for backward propagating in time midpoint-offset domain data. The equation is derived from the Double Square Root equation. The proposed method successfully images non-overturned events. Unfortunately, the propagation equation has a singularity for horizontally traveling waves, thus the method does not seem capable of imaging overturned waves. The root of the problem is that the DSR implicitly assumes that the sources and the receivers are at the same depth level (or at least that their vertical offset is constant). This assumptions is clearly unfulfilled by finite-offset overturned events. I speculate that the assumption of null (constant) vertical offset between sources and receiver could be removed. In this case a midpoint-offset domain might be still computationally attractive because the reflected wavefield would still tend to focus towards zero offset as it is back-propagated in time.

[1]**email:** biondo@sep.stanford.edu

## FROM DOWNWARD CONTINUATION TO TIME STEPPING

We would like to propagate the recorded wavefield backward in time instead of downward into the Earth, but we also would like to preserve the computational advantages of propagating the recorded wavefield in midpoint-offset coordinates. The advantage of the midpoint-offset coordinates derives from the focusing of the reflected wavefield towards zero offset as it approaches the reflector. The wavefield focuses towards zero-offset during downward continuation because we are essentially datuming the whole data set to an increasingly deeper level in the Earth. It is thus reasonable to start our derivation from the double square root (DSR) equation, that is the main tool for datuming prestack data. As we will see later, this choice of a starting point limits the usefulness of the final result.

The DSR equation in the frequency-wavenumber domain is

$$k_z = \sqrt{\omega^2 s(\mathbf{s},z)^2 - k_{x_s}^2} + \sqrt{\omega^2 s(\mathbf{g},z)^2 - k_{x_g}^2}, \tag{1}$$

where $\omega$ is the temporal frequency, $k_{x_s}$ and $k_{x_g}$ are respectively the wavenumber associated to the source and receiver locations, and $s(\mathbf{s},z)$ and $s(\mathbf{g},z)$ are the slowness at the source and receiver locations. We first start by rewriting the DSR in terms of midpoint $x_m$ and half offset $x_h$ as

$$k_z = \sqrt{\omega^2 s(\mathbf{s},z)^2 - \frac{\left(k_{x_m} - k_{x_h}\right)^2}{4}} + \sqrt{\omega^2 s(\mathbf{g},z)^2 - \frac{\left(k_{x_m} + k_{x_h}\right)^2}{4}}, \tag{2}$$

where $k_{x_m}$ and $k_{x_h}$ are respectively the wavenumber associated to the midpoint $x_m$ and the half-offset $x_h$.

Then, to obtain a time marching equation, we first square equation (2) twice and rearrange the terms into:

$$\omega^4 \Delta_s + 2\omega^2 \left(\Delta_s k_{x_m} k_{x_h} - \Sigma_s k_z^2\right) + k_z^4 + k_z^2 \left(k_{x_m}^2 + k_{x_h}^2\right) + k_{x_m}^2 k_{x_h}^2 = 0, \tag{3}$$

where

$$\Delta_s = s(\mathbf{s},z)^2 - s(\mathbf{g},z)^2 \tag{4}$$

$$\Sigma_s = s(\mathbf{s},z)^2 + s(\mathbf{g},z)^2 \tag{5}$$

Equation (3) is a second order equation in $\omega^2$. It has another solution in addition to the desired one. It can be greatly simplified by assuming $\Delta_s \approx 0$. Then equation (3) can be rewritten as

$$\omega^2 = \frac{1}{2\Sigma_s}\left(k_z^2 + k_{x_m}^2 + k_{x_h}^2 + \frac{k_{x_m}^2 k_{x_h}^2}{k_z^2}\right). \tag{6}$$

This is the basic equation solved for the numerical examples shown in this paper. Notice that when $k_{x_h}$ is equal to zero, equation (6) degenerates to the well-known equation used for reverse-time migration of zero-offset data (Baysal et al., 1984).

There are few alternatives on how to solve equation (6) numerically. The simplest one is to use finite-differences for approximating the time derivative, and Fourier transforms for evaluating the spatial-derivative operators. Because the slowness term $\Sigma_s$ is outside the parentheses

in equation (6), using Fourier transforms does not preclude the use of a spatially variable slowness field. Strong lateral velocity variations would cause problems because of the approximations needed to go from equation (3) to equation (6), not because of the numerical scheme used to solve equation (6).

The time marching scheme that I used can be summarized as;

$$\frac{P_{t-\Delta t} - 2P_t + P_{t+\Delta t}}{\Delta t^2} = \frac{1}{2\Sigma_s} \text{FFT}^{-1} \left( k_z^2 + k_{x_m}^2 + k_{x_h}^2 + \frac{k_{x_m}^2 k_{x_h}^2}{k_z^2} \right) \text{FFT} \;\; P. \tag{7}$$

Using a Fourier method to evaluate the spatial-derivative operators, makes it easy to handle the real limitation of equation (6); that is, the presence of the vertical wavenumber $k_z$ at the denominator. Waves propagating horizontally have an effective infinite velocity, making a finite-difference solution unstable, no matter how small the extrapolation time step. Unfortunately, this is a major obstacle for migrating overturned events, which is one of the main goals for developing a reverse time migration in midpoint-offset coordinates. The problem exists only for finite offset data ($k_{x_h} \neq 0$). In retrospective, the occurrence of problems for waves that overturn at finite offset should not be surprising. Equation (6) was derived from the DSR that cannot model data for which the source leg overturns at different depth than the receiver leg.

For non-overturning events the problem can be sidestepped. The spatial wavenumbers are related to the reflector geological dip angle $\gamma$ and the aperture angle $\alpha$ by the relationship

$$\frac{k_{x_m}^2 k_{x_h}^2}{k_z^2} = \tan\alpha \tan\gamma, \tag{8}$$

By simple trigonometry is also possible to show that for non-overturned events

$$\frac{k_{x_m}^2 k_{x_h}^2}{k_z^2} a \leq 1. \tag{9}$$

In the Fourier domain it is straightforward to include condition (9) in the time-marching algorithm and thus to avoid instability without suppressing reflected energy.

**Stronger lateral velocity variations?**

The numerical scheme described by equation (7) can probably handle some amount of lateral velocity variations, but it would be inaccurate for more complex velocity functions. In the derivation presented in the previous section there is no assumption of mild lateral velocity variations up to equation (3). The problem with equation (3) is that it is fourth order in time. In addition to the desired solution it has another solution that can generate artifacts and cause instability. Therefore, a direct solution by finite-differences would encounter problems with the spurious solution. Alkhalifah (1998) describes a similar problem when solving an acoustic wave equation for anisotropic media.

However, it is fairly straightforward to derive an approximation to equation (3) that is more accurate than equation (6). Equation (3) can be easily solved for $\omega^2$ because it contains only

the even powers of $\omega$. We can then approximate the square root that appears in the formal solution for $\omega^2$ as

$$\sqrt{1-\frac{\Delta_s\left[k_z^4+k_z^2\left(k_{x_m}^2+k_{x_h}^2\right)+k_{x_m}^2k_{x_h}^2\right]}{\left(\Delta_s k_{x_m}k_{x_h}-\Sigma_s k_z^2\right)^2}}\approx 1-\frac{\Delta_s\left[k_z^4+k_z^2\left(k_{x_m}^2+k_{x_h}^2\right)+k_{x_m}^2k_{x_h}^2\right]}{2\left(\Delta_s k_{x_m}k_{x_h}-\Sigma_s k_z^2\right)^2}. \tag{10}$$

The useful solution of equation (3) can then be approximated as

$$\omega^2=\frac{1}{2\left(\Sigma_s-\Delta_s k_{x_m}k_{x_h}\right)}\left(k_z^2+k_{x_m}^2+k_{x_h}^2+\frac{k_{x_m}^2k_{x_h}^2}{k_z^2}\right). \tag{11}$$

Equation (11) degenerates to equation (6) when $\Delta_s$ is zero. It is more accurate than Equation (6) for $\Delta_s$ different than zero, but it is likely to break down for strong lateral velocity variations.

Equation (11) shares with equation (6) the fundamental problem of instability for horizontally propagating waves. Therefore, I have not implemented a numerical scheme to solve equation (11) yet. However, it is possible to define a mixed implicit-explicit method to solve equation (11), similar to the one proposed by Klíe and Toro (2001) to solve Alkhalifah's acoustic wave equation for anisotropic media.

### Imaging principle

One of the advantages of the choice of the midpoint-offset coordinate to migrate the data is that the imaging step is both straightforward and inexpensive. There is no need to crosscorrelate and shift wavefields, as for shot-profile reverse-time migration (Biondi, 2002).

The equivalent of the "stacked" image is obtained by selecting the propagated wavefield at zero time and zero offset; that is

$$I\left(z,x_m\right)=P\left(t=0,z,x_m,x_h=0\right) \tag{12}$$

The offset-domain Common Image Gathers (CIG) are obtained by simply selecting the propagated wavefield at zero time; that is

$$I\left(z,x_m,x_h\right)=P\left(t=0,z,x_m,x_h\right). \tag{13}$$

## EXAMPLES OF REVERSE TIME MIGRATION OF A SYNTHETIC DATA SET

To illustrate the use of the proposed method to image seismic data by reverse-time migration, I migrated the same simple synthetic data set (with a dipping and a flat layer) that I used in Biondi (2002). Since the data were modeled in shot gathers, I sorted the data set in midpoint-offset coordinates before migration. As in the other paper, I migrated the data with a constant

velocity to avoid artifacts caused by velocity discontinuities in the migration velocity. Therefore, the deeper reflector is slightly undermigrated.

Figure 1 is the equivalent of the "stacked" image, obtained by applying equation (12). The dipping reflector is well imaged within the range that is illuminated by the shots. The flat reflector is slightly undermigrated, as discussed above. The quality of the image is comparable to the quality of the image obtained by shot-profile migration using the same data.

Figure 2 shows on the left the offset-domain CIG (a) and on the right the angle-domain CIG (b). The offset domain CIG was obtained by applying equation (13), and the angle-domain CIG was obtained by applying the offset-to-angle transformation presented in Sava et al. (2001). The CIGs are located at a surface location where both reflectors are well illuminated (1,410 m). As expected the image is well focused at zero-offset in the panel on the left and the events are flat in the panel on the right.

As for shot-profile migration, the CIGs obtained by midpoint-offset reverse-time migration can be easily used for velocity updating. To illustrate this capability, I have migrated the same data set with a lower velocity ( .909 km/s). Figure 3 is the "stacked" image obtained using the lower velocity. Both reflectors are undermigrated and shifted upward.

Figure 4 shows on the left the offset-domain CIG (a) and on the right the angle-domain CIG (b). Now in the panel on the left, the energy is not well focused at zero offset but is spread over an hyperbolic trajectory centered at zero offset. The corresponding angle-domain CIG (ADCIG) (right) shows the characteristic smile typical of undermigrated ADCIG. The velocity information contained in the panel on the right can be easily used for velocity updating and tomographic inversion in a similar way as the ADCIG obtained by downward-continuation migrations are used (Clapp and Biondi, 2000; Clapp, 2001).

## CONCLUSIONS

I explore the possibility of imaging prestack data by reverse-time migration in the midpoint-offset domain. The method that I derive successfully imaged a simple synthetic data set. However, it is not capable of migrating correctly overturned events, which is one of the main goals of the effort. This limitation originates from the limitations of the starting point of my derivation; that is, the Double Square Root equation. The DSR assumes a null (constant) vertical offset between source and receivers. The choice of a more general starting point might lead to a useful migration method.

## REFERENCES

Alkhalifah, T., 1998, An acoustic wave equation for anisotropic media: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1913–1916.

Baysal, E., Kosloff, D. D., and Sherwood, J. W. C., 1984, A two-way nonreflecting wave equation: Geophysics, **49**, no. 02, 132–141.

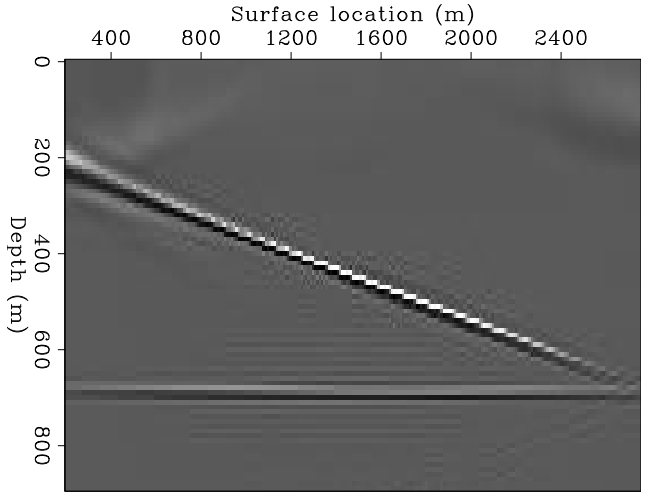



Figure 1: Image of the synthetic data set with the correct velocity function. biondo1-MidOff-Image-dip [CR]

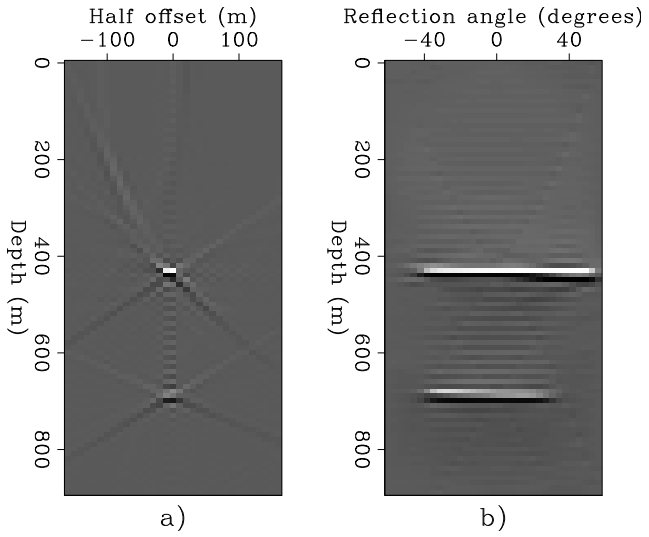



Figure 2: Offset-domain CIG (left a)) and angle-domain CIG (right b)) corresponding to the image in Figure 1. Notice the focusing at zero offset in a), and the flatness of the moveout in b). biondo1-MidOff-Cig-Ang-dip [CR]

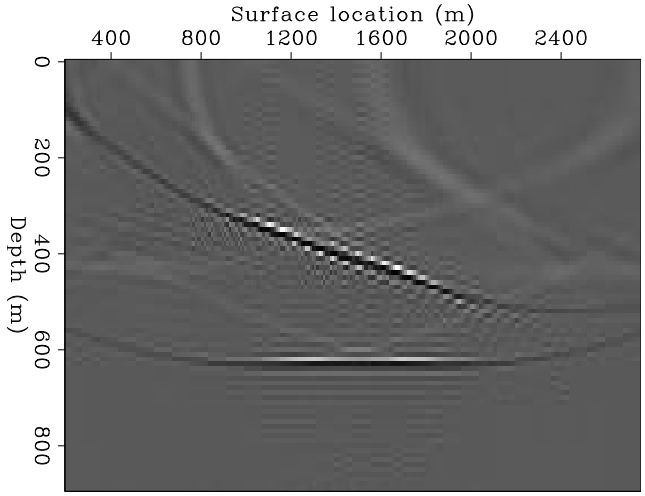



Figure 3: Image of the synthetic data set with the incorrect velocity function. biondo1-MidOff-Image-dip-slow [CR]

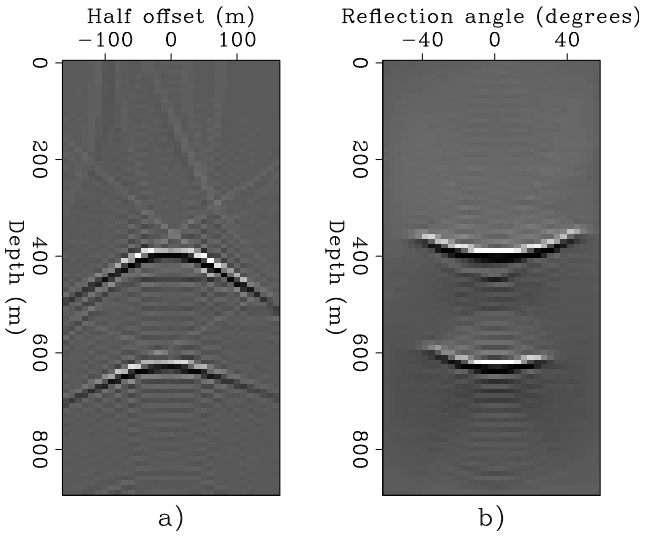



Figure 4: Offset-domain CIG (left) and angle-domain CIG (right) corresponding to the image in Figure 3. Notice the lack of focusing at zero offset in a), and the smile in b). biondo1-MidOff-Cig-Ang-dip-slow [CR]

Biondi, B., and Palacharla, G., 1996, 3-D prestack migration of common-azimuth data: Geophysics, **61**, no. 6, 1822–1832.

Biondi, B., 2002, Prestack imaging of overturned and prismatic reflections by reverse time migration: SEP–**111**, 121–137.

Clapp, R., and Biondi, B., 2000, Tau domain migration velocity analysis using angle CRP gathers and geologic constrains:, *in* 70th Ann. Internat. Mtg Soc. of Expl. Geophys., 926–929.

Clapp, R. G., 2001, Geologically constrained migration velocity analysis: Ph.D. thesis, Stanford University.

Klíe, H., and Toro, W., 2001, A new acoustic wave equation for modeling in anisotropic media: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1171–1174.

Sava, P., Biondi, B., and Fomel, S., 2001, Amplitude-preserved common image gathers by wave-equation migration: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 296–299.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 157–170*

# Adaptive subtraction of multiples with the $\ell^1$-norm

*Antoine Guitton and Eric Verschuur*[1]

**ABSTRACT**

The estimation of shaping filters with the $\ell^1$-norm as opposed to the $\ell^2$-norm leads to a proper attenuation of multiples when significant amplitude discrepancies exist between multiples and primaries. The actual method implemented is the fairly standard iteratively re-weighted least-squares method which is an excellent approximation to $\ell^1$. Synthetic and field data results illustrate the advantages of the $\ell^1$-norm.

## INTRODUCTION

A classical approach for attenuating multiples consists of building a multiple model (Verschuur et al., 1992) and adaptively subtracting this model from the multiple infested-data by estimating shaping filters (Dragoset, 1995; Liu et al., 2000; Rickett et al., 2001). The estimation of the shaping filters is usually done in a least-squares sense making these filters relatively easy to compute. In some cases, however, a least-squares criterion can lead to undesirable artifacts. This happens when, for example, the relatively strong primaries are surrounded by multiples, such that the filter tends to distort primary energy as well.

In this paper we show that the estimation of shaping filters with the $\ell^1$-norm gives better results than with the $\ell^2$-norm when multiples and primaries have noticeable amplitude differences. We first illustrate this with a simple 1D problem that highlights the limits of the least-squares approach. In a second synthetic example, we attenuate internal multiples and show that the $\ell^1$-norm gives far better results than does $\ell^2$. To finish, we utilize shaping filters on a multiple contaminated gather from a seismic survey showing that the $\ell^1$-norm leads to a substantial attenuation of the multiples.

## A SIMPLE 1D PROBLEM

In this section, we demonstrate on a 1D problem that the attenuation of multiples with least-squares adaptive filtering is not effective when amplitude differences exist between primaries and multiples. This simple example helps us to better understand the behavior of our adaptive scheme in more complicated cases.

[1]**email:** antoine@sep.stanford.edu, D.J.Verschuur@CTG.TUDelft.nl

**Shaping filters and the $\ell^2$-norm**

h This section illustrates some limitations of the $\ell^2$-norm for the estimation of shaping filters. In figure 1, we display a very simple 1D problem. On the top we have four events corresponding to one primary (on the left) and three multiples (on the right). Note that the primary has higher amplitude than the multiples. On the bottom we show a multiple model that exactly corresponds to the real multiples. Our goal is to estimate one shaping filter **f** that minimizes the objective function

$$e(\mathbf{f}) = \|\mathbf{d} - \mathbf{M}\mathbf{f}\|_2^2, \tag{1}$$

where **M** is the matrix representing the convolution with the time series for the multiple model (Figure 1b) and **d** the time series for the data (Figure 1a).

Now, if we estimate the filter **f** with enough degrees of freedom (enough coefficients) to minimize equation (1), we obtain for the signal Figure 2a, and for the noise Figure 2b. The estimated signal does not resemble the primary in Figure 1a. We show the corresponding shaping filter in Figure 3. This filter is not the single spike at $lag = 0$ that we desire. The problem stems from the least-squares criterion which yields an estimated signal that has, by definition, minimum energy. In this 1D case, the total energy in the estimated signal (Figure 2a) is $e = 2.4$, which is less than the total energy of the primary alone ($e = 4$). This is the fundamental problem if we use the $\ell^2$-norm to estimate the shaping filter. In the next section, we show that the $\ell^1$ norm should be used if amplitude differences exist between primaries and multiples.

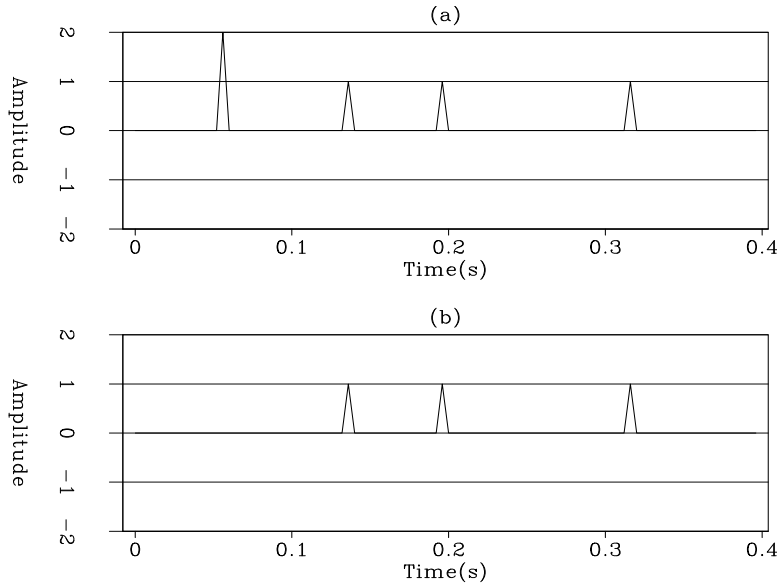



Figure 1: (a) The data with one primary on the left, and three multiples on the right. (b) The multiple model that we want to adaptively subtract from (a). antoine1-datmul [ER]

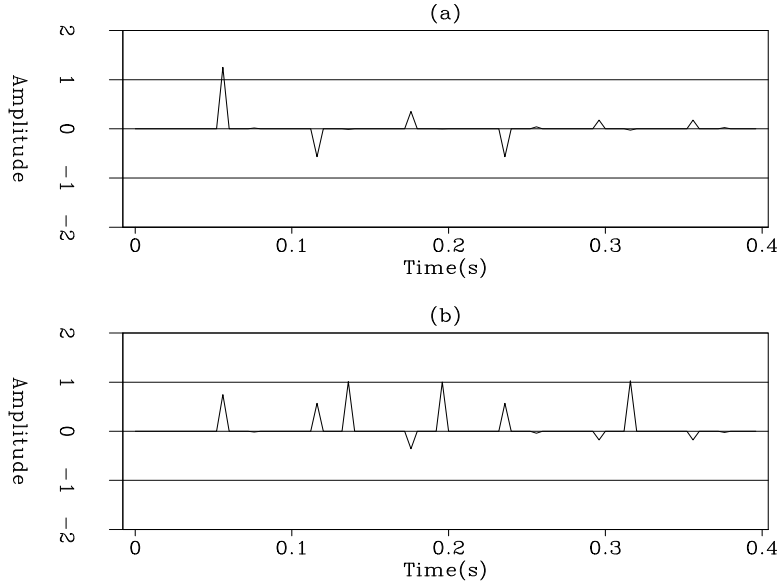



Figure 2: (a) The signal estimated with the $\ell^2$-norm. (b) The noise estimated with the $\ell^2$-norm. antoine1-1Dl2 [ER]

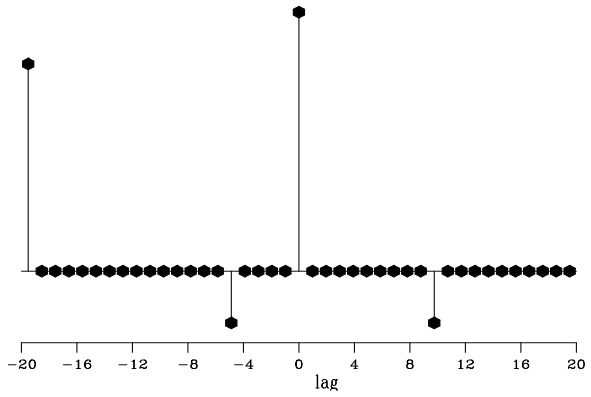



Figure 3: Shaping filter estimated for the 1D problem with the $\ell^2$-norm. This filter is not a single spike at $lag = 0$. antoine1-filterl2 [ER]

### Shaping filters and the $\ell^1$-norm

We prove that the $\ell^1$-norm solves the problem highlighted in the preceding section. Now our goal is to estimate one shaping filter **f** that minimizes the objective function

$$e(\mathbf{f}) = |\mathbf{d} - \mathbf{M}\mathbf{f}|_1. \tag{2}$$

To achieve this, the shaping filter is estimated iteratively using a nonlinear conjugate gradient solver (NLCG) as described in Claerbout and Fomel (2001). The objective function we actually minimize is

$$e(\mathbf{f}) = \|\mathbf{W}(\mathbf{d} - \mathbf{M}\mathbf{f})\|_2^2, \tag{3}$$

with

$$\mathbf{W} = \mathbf{diag}\left(\frac{1}{(1 + r_i^2/\epsilon^2)^{1/4}}\right), \tag{4}$$

where $r_i$ is the residual for one component of the data space, and $\epsilon$ a constant we choose a priori. Equation (3) is minimized with the standard iteratively re-weighted least-squares approach (Nichols (1994); Bube and Langan (1997); Guitton (2000)) The objective function in equation (3) amounts to the $\ell^1$ measure when $r_i/\epsilon$ is large and amounts to the $\ell^2$ measure when $r_i/\epsilon << 1$ with a smooth transition between the two.

In Figure 4, we display the result of the adaptive subtraction when the $\ell^1$-norm is utilized to estimate the shaping filter [equation (3) with a small $\epsilon$]. The estimated signal in Figure 4a is perfect, and so is the estimated noise. It is easy to check that the energy in Figure 4a ($e = 2$) is less than the energy in Figure 2a ($e = 3.2$) if we use the $\ell^1$ norm. Figure 5 shows the shaping filter associated with the $\ell^1$-norm. This filter is a spike at $lag = 0$. This simple 1D example demonstrates that the $\ell^1$ should be utilized each time significant amplitude differences exist between multiples and primaries. In the next section, we show another synthetic example where internal multiples are attenuated.

## ATTENUATION OF INTERNAL MULTIPLES

In this section we illustrate the efficiency of the $\ell^1$-norm when internal multiples are attenuated in 2D.

### The synthetic data

Figure 6a shows a synthetic shot gather for a 1D medium. This gather is corrupted with internal multiples only. In Figure 6b, we display the internal multiple model obtained using the CFP approach (Berkhout and Verschuur, 1999). This internal multiple model is perfect and could be directly subtracted from the data in Figure 6a. Note that the amplitude of the internal multiples is significantly less than the amplitude of the primaries, making the $\ell^2$-norm unsuitable for estimating the shaping filters. Figure 7 displays the histograms of both the data and the internal multiples. The narrow peak of the noise indicates that the $\ell^1$-norm should be used.

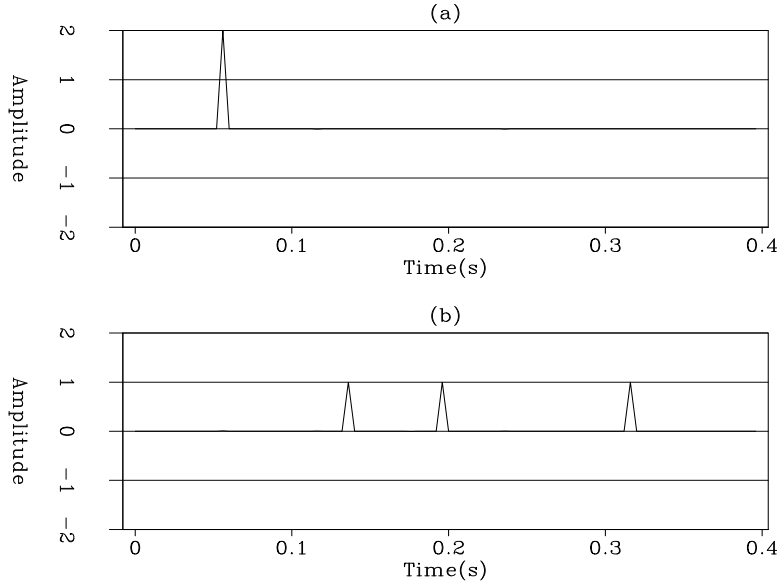



Figure 4: (a) The signal estimated with the $\ell^1$-norm. (b) The noise estimated with the $\ell^1$-norm. antoine1-1Dl1 [ER]

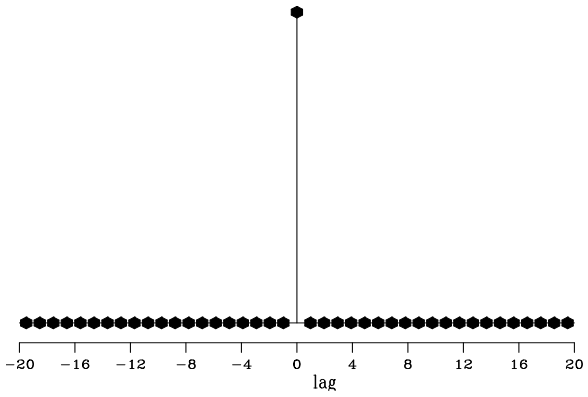



Figure 5: Shaping filter estimated for the 1D problem with the $\ell^1$-norm. This filter is a single spike at $lag = 0$. antoine1-filterl1 [ER]

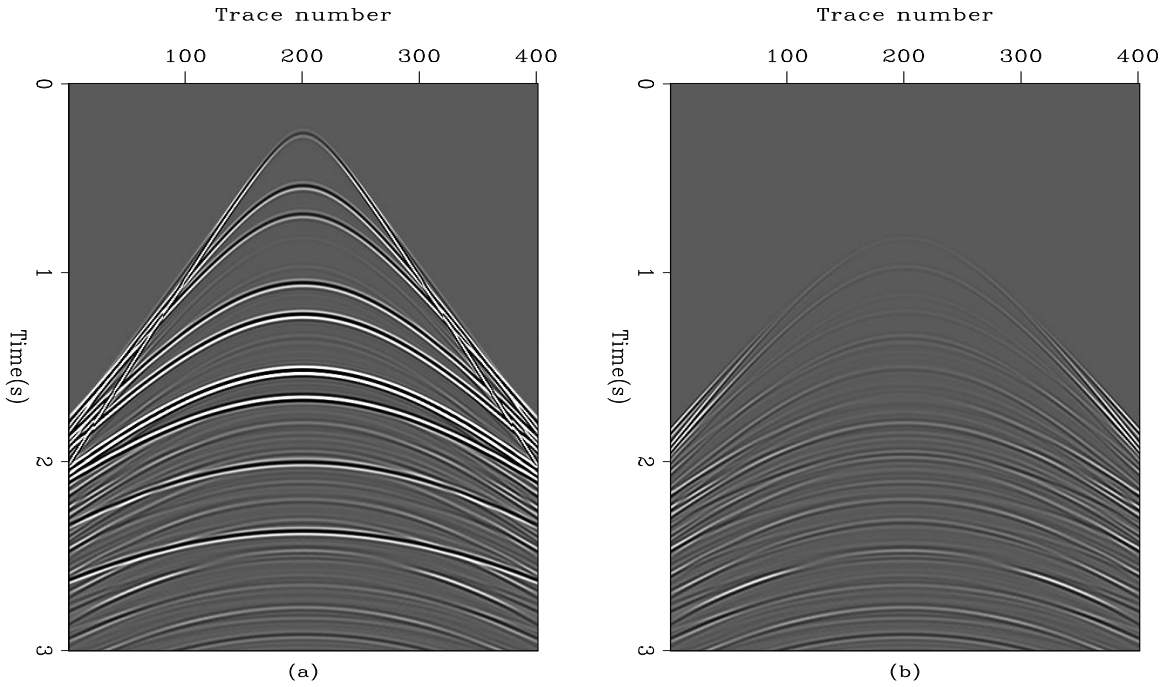



Figure 6: (a) A synthetic shot gather infested with internal multiples. (b) The internal multiples model obtained using the CFP technology. This model matches the internal multiples in (a). antoine1-inter [ER]

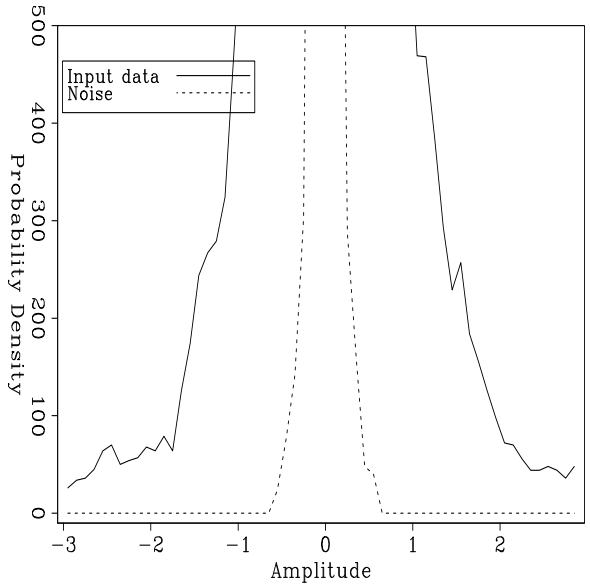



Figure 7: Histograms of the input data (Figure 6a) and of the noise (Figure 6b). The density function of the noise is much narrower than for the data. The $\ell^1$-norm should be used to estimate the signal. antoine1-histodata [ER]

### Adaptive filtering with non-stationary helical filters

To handle the inherent non-stationarity of seismic data, we estimate a filter bank of non-stationary filters using helical boundary conditions (Claerbout, 1998). This approach has been successfully utilized by Rickett et al. (2001) to attenuate surface-related multiples. We use a NLCG solver with the $\ell^1$-norm and a standard conjugate gradient solver with the $\ell^2$-norm. The filter coefficients vary smoothly across the output space thanks to a preconditioning of the problem (Rickett et al., 2001). In the following results, the non-stationary filters are 1D. We estimate the same number of coefficients per filter with the $\ell^2$- and $\ell^1$-norm.

### Adaptive subtraction results

Figure 8a shows the estimated primaries when the $\ell^2$-norm is used to compute the shaping filters. Figure 8b displays the estimated internal multiples. As expected, because of the amplitude differences between the signal (primaries) and the noise (multiples), the adaptive subtraction fails and we retrieve the behavior explained in the preceding section with the 1D example. Now, in Figure 9, we see the beneficial effects of the $\ell^1$-norm. Figure 9a shows the estimated primaries and Figure 9b the estimated multiples. The noise subtracted almost perfectly matches the internal multiple model in Figure 6b, as anticipated.

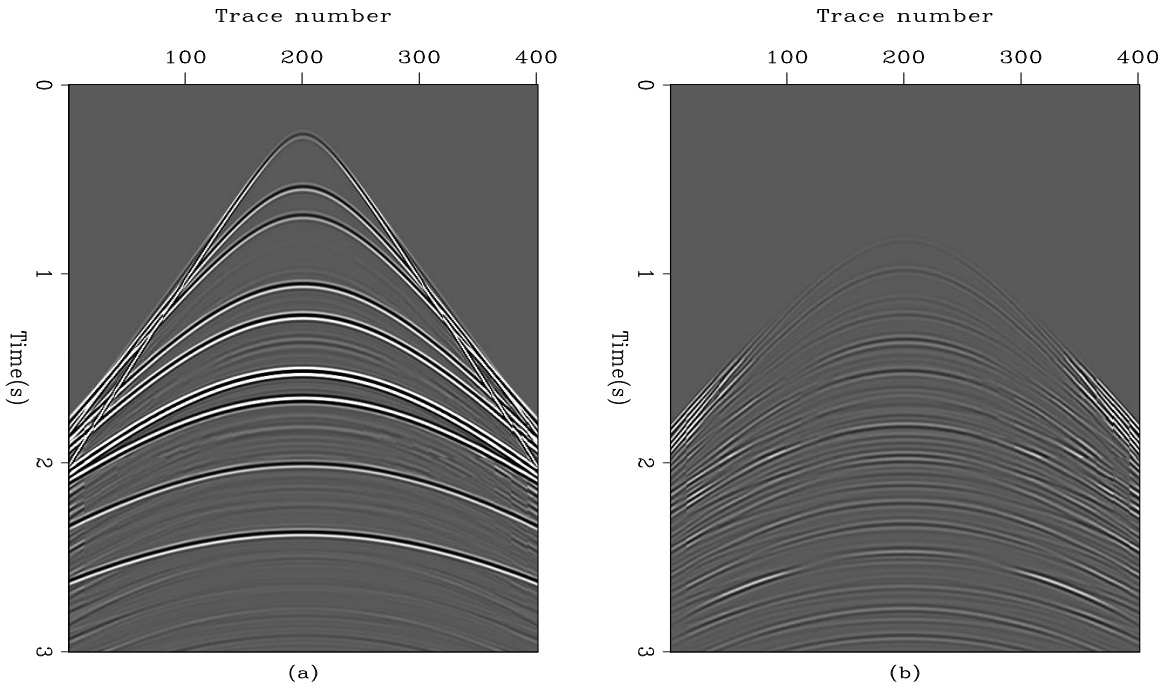



Figure 8: (a) The estimated primaries with the $\ell^2$-norm. (b) The estimated internal multiples with the $\ell^2$-norm. Ideally, (b) should look like Figure 6b, but it does not. antoine1-interl2 [ER]

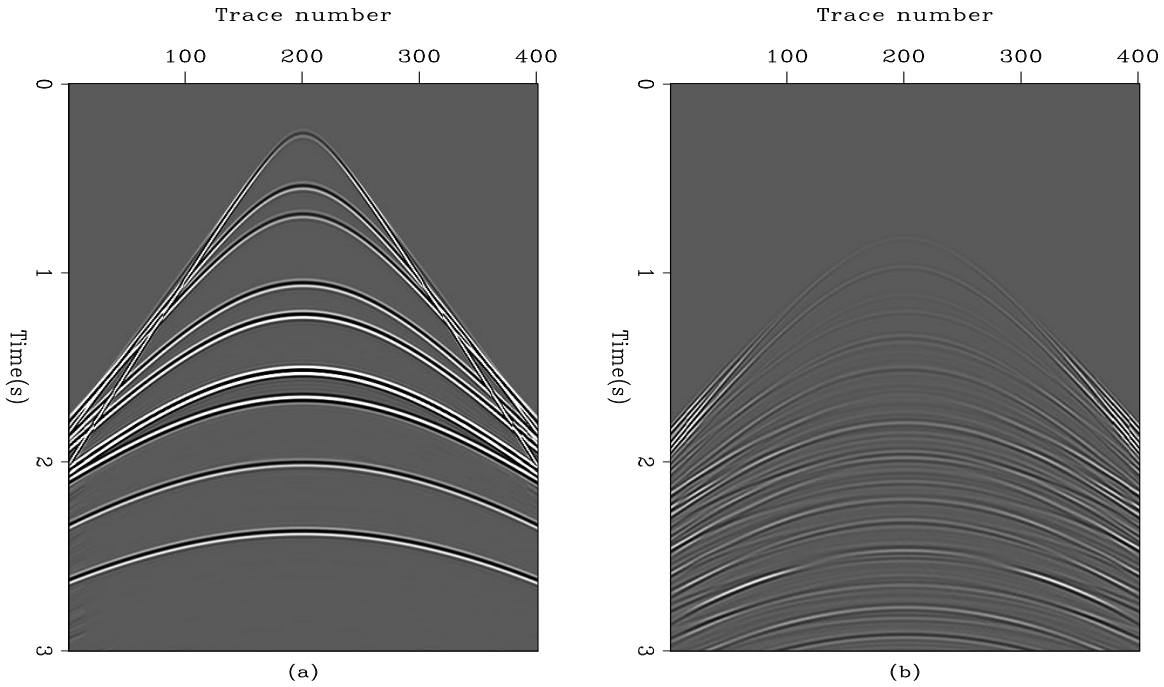



Figure 9: (a) The estimated primaries with the $\ell^1$-norm. (b) The estimated internal multiples with the $\ell^1$-norm. Beside some edge-effects, (b) resembles closely Figure 6b. The adaptive subtraction worked very well. antoine1-interl1 [ER]

## POSTSTACK LAND DATA MULTIPLE REMOVAL EXAMPLE

In this section we attenuate in the poststack domain surface-related multiples with shaping filters that we estimate with the $\ell^2$- and $\ell^1$-norm. These filters are non-stationary. Figure 10a shows the multiple-infested data. Figure 10b displays the multiple model computed with the Delft modeling approach (Kelamis et al., 1999). Note that for this gather, the amplitude differences between the primaries and the multiples are not very strong. Our goal is to illustrate the use of the $\ell^1$-norm in a more general case when surface-related multiples are present in the data. We specifically focus on the event at 1.6s in Figure 10. This event is a primary that we want to preserve during the subtraction.

Figures 12 displays the estimated signal when the non-stationary shaping filters are computed with the $\ell^2$ and $\ell^1$-norm. The amplitude of the primary at 1.6s is well preserved with the $\ell^1$-norm in Figure 11a. However, the amplitude of this primary is attenuated with the $\ell^2$-norm as displayed in Figure 11b. Figure 10 shows a comparison between the subtracted noise with the $\ell^1$ (Figure 10a) and the $\ell^2$-norm (Figure 10b). We conclude that the $\ell^2$-norm tends to subtract too much energy.

This last example proves that the estimation of shaping filters can always be done with the $\ell^1$-norm. The good thing about our inversion scheme and the objective function in equation (3) is that only one parameter ($\epsilon$) controls the $\ell^1 - \ell^2$ behavior. Thus we can decide to switch from one norm to another very easily. In Figure 13, I show a histogram of the input data and of

the estimated noise with the $\ell^1$ and $\ell^2$-norms. The theory predicts that the distribution of the $\ell^2$ result should be gaussian and that distribution of the $\ell^1$ result should be exponential. Figure 13 corroborates this.

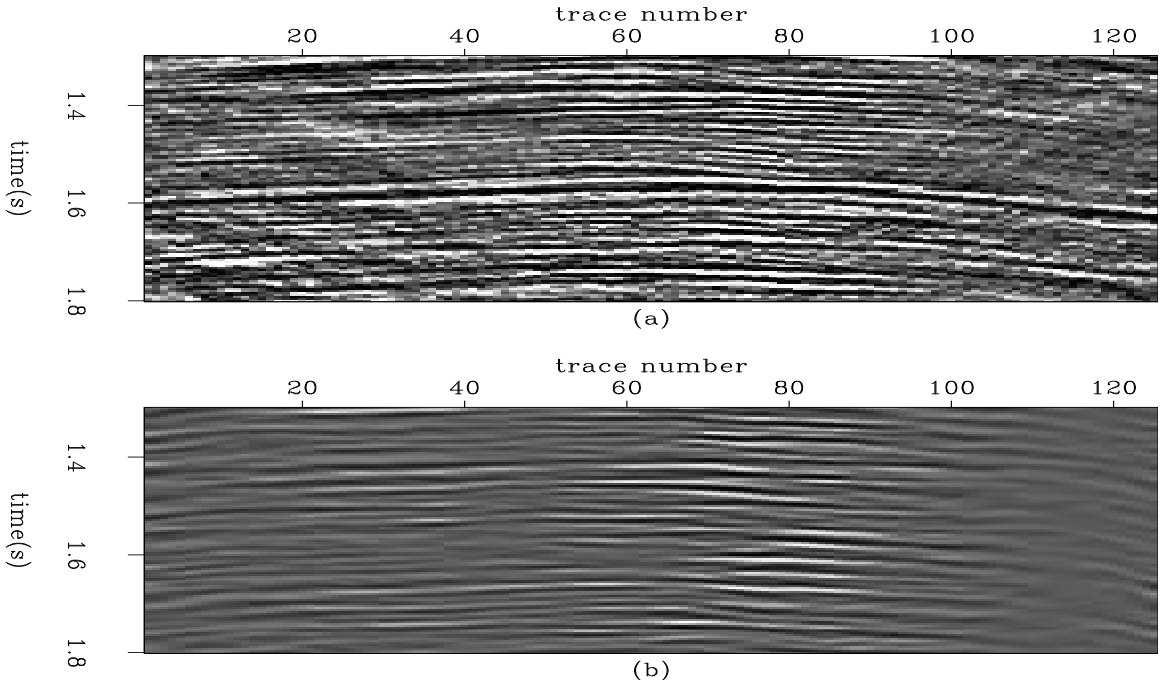



Figure 10: (a) Stack infested with multiples. (b) The multiple model computed with the Delft modeling approach. The subtraction is done poststack. antoine1-win3 [ER]

## PRESTACK LAND DATA MULTIPLE REMOVAL EXAMPLE

The above methodology has also been tested on a shot record from a land data survey. The preprocessing and multiple prediction is described by Kelamis et al. (1999). I display in Figure 14a the selected shot record. Figure 15a shows the predicted multiples. Note that these multiples are the ones that are directly generated by the shot record based convolutions (Berkhout and Verschuur, 1997) and no adaptation has been applied yet. Both $\ell^2$ and $\ell^1$ adaptive subtraction has been carried out for this data. The resulting records are displayed in Figures 14b and c, respectively. The removed multiples are shown in Figures 15b and c.

Figures 15b and 15c demonstrate that the multiples are better attenuated with the $\ell^1$-norm for long offsets. Although the truth cannot be revealed from these results, it appears that the $\ell^1$ results are more reliable.

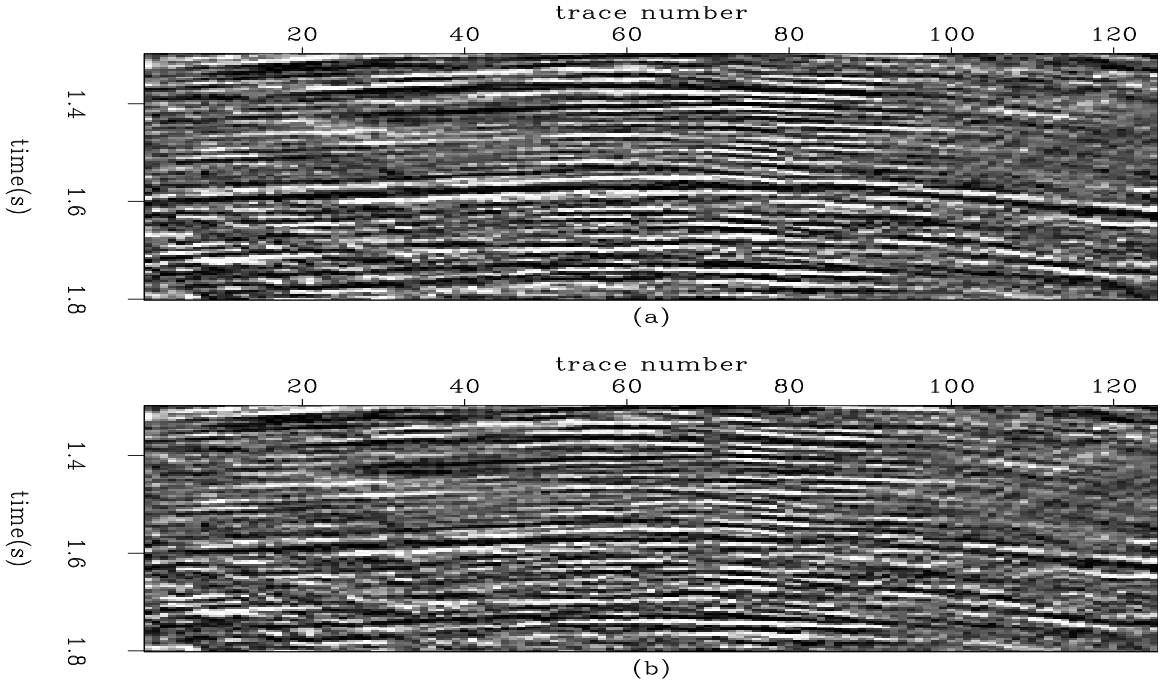



Figure 11: (a) The estimated primaries with $\ell^1$-norm adaptive subtraction. (b) The estimated primaries with $\ell^2$-norm subtraction. The primary at 1.6s is very attenuated with the $\ell^2$-norm. The $\ell^1$ technique preserves its amplitude better. antoine1-win [ER,M]

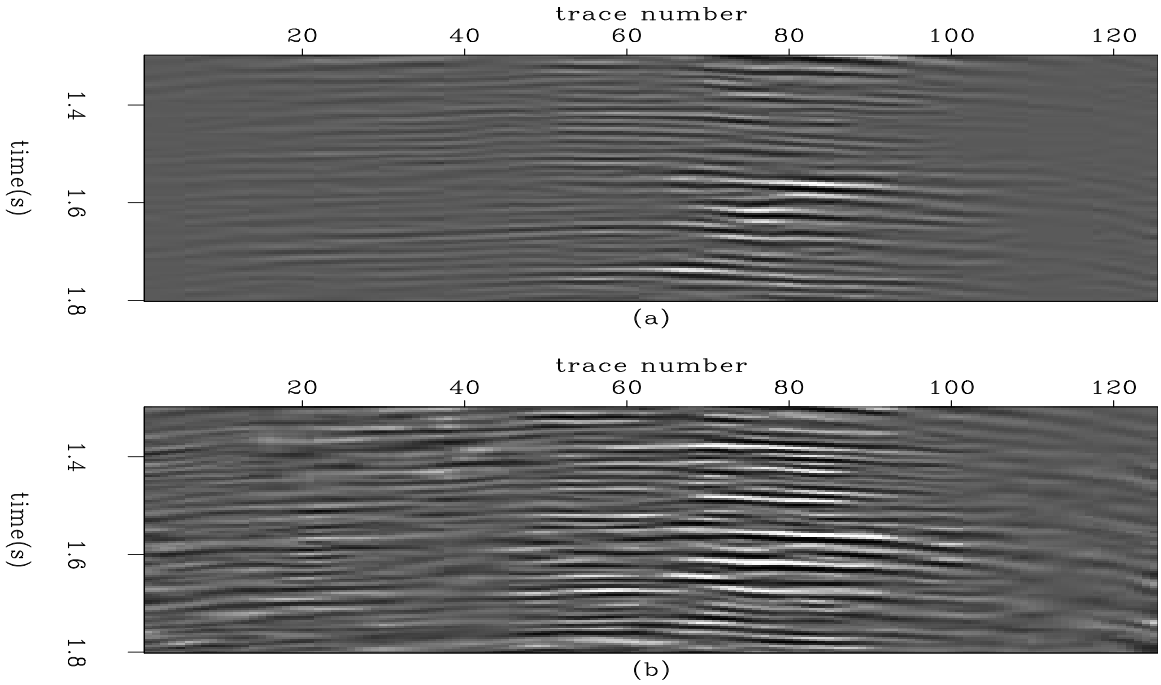



Figure 12: (a) The estimated multiples with the $\ell^1$-norm subtraction. (b) The estimated multiples with the $\ell^2$-norm subtraction. The $\ell^2$-norm tends to over-fit some multiples that creates some leaking of primaries in the estimated noise. antoine1-win2 [ER,M]

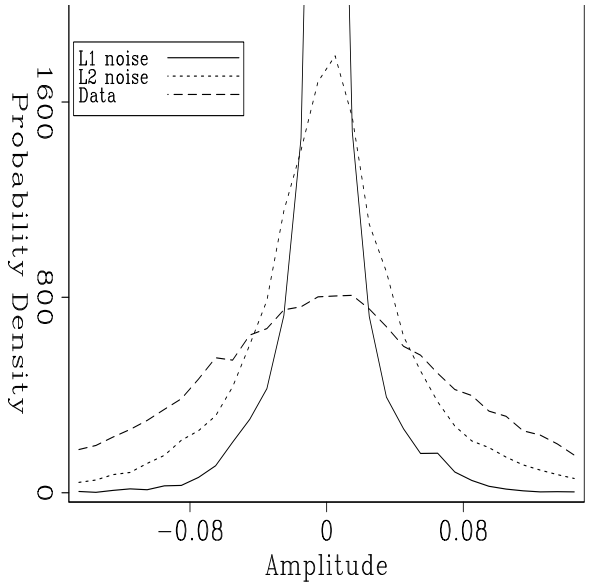



Figure 13: Histograms of the input data and of the estimated noise with the $\ell^1$- and $\ell^2$-norms. As predicted by the theory, the density function with the $\ell^1$-norm is much narrower than with the $\ell^2$-norm. antoine1-hist7636 [ER]

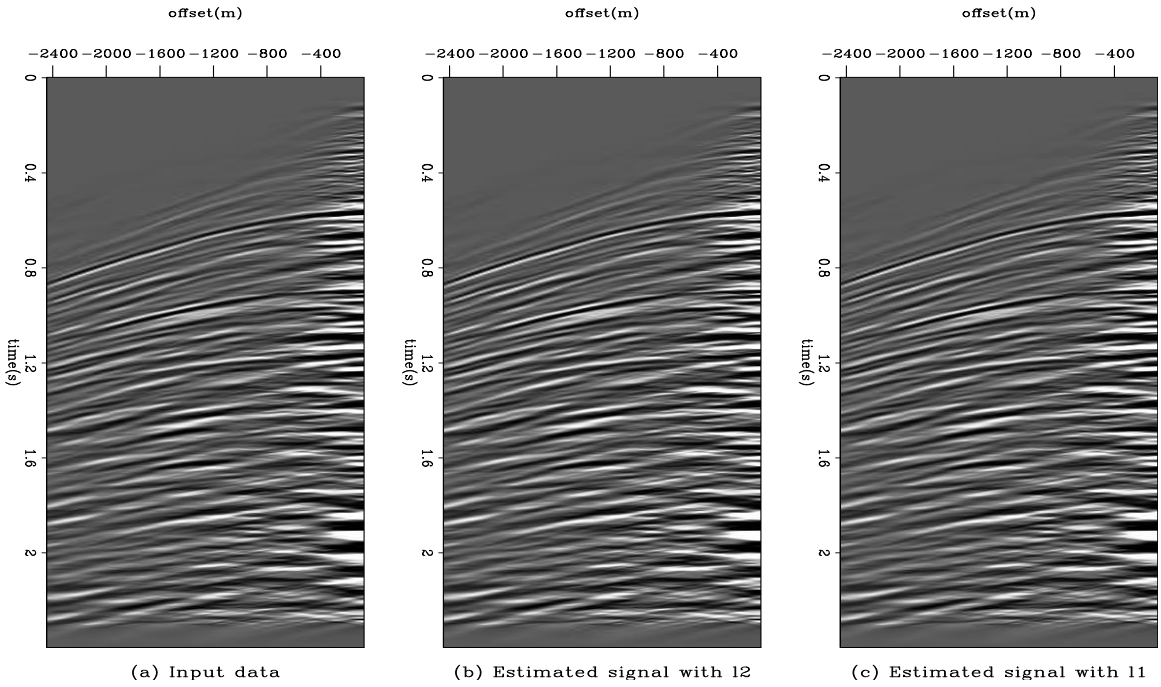



Figure 14: Example of adaptive multiple subtraction for land data. a) One selected shot record from a land survey. b) Estimated signal after $\ell^2$ adaptive subtraction. c) Estimated signal after $\ell^1$ adaptive subtraction. antoine1-comp.s.4 [ER,M]

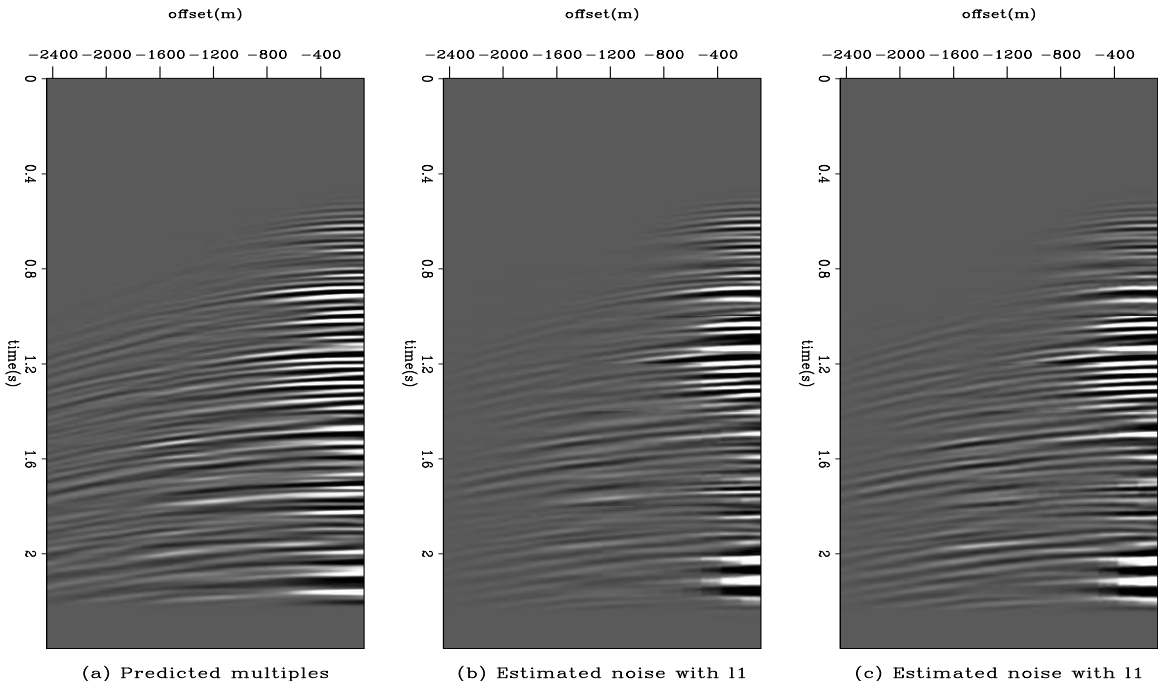



Figure 15: a) Predicted multiples for Figure 14a. b) Removed multiples with $\ell^2$ adaptive subtraction. c) Removed multiples with $\ell^1$ adaptive subtraction. Far offset multiples are better attenuated with the $\ell^1$-norm between 1.2 and 1.8 seconds. antoine1-comp.n.4 [ER,M]

## CONCLUSION

Significant amplitude differences between signal and noise make the $\ell^2$-norm an unsuitable choice to estimate shaping filters. We showed that the $\ell^1$-norm should always be considered in these circumstances. In addition, applications to real data with surface-related multiples tend to prove that the $\ell^1$-norm should also be considered in the simplest cases.

## ACKNOWLEDGMENTS

We would like to thank Saudi Aramco for providing the land data and the members of the SMAART 2 JV for their financial support.

## REFERENCES

Berkhout, A. J., and Verschuur, D. J., 1997, Estimation of multiple scattering by iterative inversion, Part I: Theoretical considerations: Geophysics, **62**, no. 05, 1586–1595.

Berkhout, A., and Verschuur, D., 1999, Removal of internal multiples: 69th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1334–1337.

Bube, K. P., and Langan, R. T., 1997, Hybrid $\ell^1/\ell^2$ minimization with applications to tomography: Geophysics, **62**, no. 04, 1183–1195.

Claerbout, J., and Fomel, S., 2001, Geophysical Estimation by Example: Class notes, http://sepwww.stanford.edu/sep/prof/index.html.

Claerbout, J., 1998, Multidimensional recursive filters via a helix: Geophysics, **63**, no. 05, 1532–1541.

Dragoset, B., 1995, Geophysical applications of adaptive-noise cancellation: 65th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1389–1392.

Guitton, A., 2000, Huber solver versus IRLS algorithm for quasi L1 inversion: SEP-Report, **103**, 255–271.

Kelamis, P., Verschuur, D., Berkhout, A., and Erickson, K., 1999, Velocity-independent datuming of seismic data: 69th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 441–444.

Liu, F., Sen, M. K., and Stoffa, P. L., 2000, Dip selective 2-D multiple attenuation in the plane-wave domain: Geophysics, **65**, no. 01, 264–274.

Nichols, D., 1994, Velocity-stack inversion using $\mathbf{L_p}$ norms: SEP-Report, **82**, 1–16.

Rickett, J., Guitton, A., and Gratwick, D., 2001, Adaptive Multiple Subtraction with Non-Stationary Helical Shaping Filters: 63rd Mtg., Eur. Assn. Geosci. Eng., Expanded Abstracts, Session: P167.

Verschuur, D. J., Berkhout, A. J., and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, no. 09, 1166–1177.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 171–182*

# A hybrid adaptive subtraction method

*Antoine Guitton*[1]

**ABSTRACT**

A hybrid adaptive subtraction scheme is proposed. This hybrid scheme uses prediction-error filters as covariance operators within the filter-estimation step. This methods proves to be the most efficient when the noise and signal interfere. Although prediction-error filters are utilized, this technique is not a pattern-recognition technique: it simply tries to remove the correlated signal information to unbias the estimation of the matched-filters. Tests on synthetic and real data for a multiple attenuation problem illustrate the efficiency of the proposed scheme.

## INTRODUCTION

Often the subtraction of noise is done via an adaptive subtraction of a noise model from the data. The resulting signal components become, by construction, orthogonal to the noise components (Spitz, 1999). Consequently, adaptive subtraction is not appropriate when noise and signal are correlated.

A very common noise attenuation problem is the subtraction of multiples. It is well known that the subtraction of multiples becomes very challenging when they interfere strongly with primaries (Spitz, 1999). Some solutions have been proposed by various authors to cope with correlated noise and signal. For instance, the so-called pattern-based methods have proved to be particularly efficient at attenuating multiples in the most complex areas (Spitz, 1998; Guitton et al., 2001).

In this paper, I investigate an improved adaptive subtraction scheme that can separate interfering multiples and primaries. With this method, I do not assume that the signal has minimum energy. Building on Guitton (2002), I estimate a signal covariance operator with time domain (t, x) prediction-error filters (pef). This covariance operator is then utilized within an inversion scheme to remove the signal spectrum in the data leading to an unbiased estimation of the matched-filters.

In the first section, I review theoretical issues about adaptive subtraction and I present the new hybrid scheme. In the second section, I illustrate the advantages of the proposed method with synthetic and field data examples.

[1]**email:** antoine@sep.stanford.edu

## IMPROVING ADAPTIVE SUBTRACTION

In this section I review basic notions on adaptive subtraction. Then I present an improved scheme that better separates interfering noise and signal. All my derivations are based upon the assumption that we want to subtract a noise model from the data. In the multiple attenuation problem the noise becomes the multiples and the signal becomes the primaries.

### Theory of adaptive subtraction

The goal of adaptive subtraction is as follows: given a time series **b** and a desired time series **d**, we seek a filter **f** that minimizes the difference between $\mathbf{f} * \mathbf{b}$ and **d** where $*$ is convolution. We can rewrite this definition in the fitting goal

$$\mathbf{0} \approx \mathbf{B}\mathbf{f} - \mathbf{d} \tag{1}$$

where **B** represents the convolution with the time series **b**. We can minimize this fitting goal in a least-squares sense leading to the objective function

$$g(\mathbf{f}) = (\mathbf{B}\mathbf{f} - \mathbf{d})'(\mathbf{B}\mathbf{f} - \mathbf{d}) \tag{2}$$

where $(')$ is the transpose. The minimum energy solution is given by

$$\widehat{\mathbf{f}} = (\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{d}. \tag{3}$$

where $\widehat{\mathbf{f}}$ is the least-squares estimate of **f**. This approach is very popular but has some intrinsic limitations. In particular $\mathbf{B}\widehat{\mathbf{f}}$ is by construction orthogonal to the residual $\mathbf{B}\widehat{\mathbf{f}} - \mathbf{d}$. In the multiple attenuation problem **d** is the data, **b** the multiple model and $\mathbf{B}\widehat{\mathbf{f}} - \mathbf{d}$ the estimated primaries. If both signal and noise are correlated, the separation will suffer because of the orthogonality principle.

From now on I will refer to this method as the "standard approach".

In the next section I propose improving the adaptive subtraction scheme. This improvement leads to an unbiased matched-filter estimation when both signal and noise are correlated.

### A hybrid attenuation scheme

In the multiple attenuation case, most of the problems encountered with the adaptive subtraction technique stem from the correlation that might exist between the noise and signal. In addition, the minimum energy assumption forces the residual (the signal) to be white, which is not valid all the time. Fortunately we can derive a fitting goal that can cope with interfering noise and signal and non-white spectrum.

In Guitton (2002), I presented a method that approximates covariance operators with pef. The goal was basically to obtain independent and identically distributed (iid) residual components. I propose using the same approach for the filter estimation in adaptive filtering.

Following this idea, I have the new fitting goal

$$\mathbf{0} \approx \mathbf{A_s}(\mathbf{B f_h} - \mathbf{d}) \tag{4}$$

where $\mathbf{A_s}$ is a pef that whitens the signal spectrum only. The corresponding least-squares estimate for $\mathbf{f_h}$ becomes

$$\widehat{\mathbf{f_h}} = (\mathbf{B}'\mathbf{A}_\mathbf{s}'\mathbf{A_s B})^{-1}\mathbf{B}'\mathbf{A}_\mathbf{s}'\mathbf{A_s d}. \tag{5}$$

If the signal has a white spectrum, then this new estimate of the filter $\widehat{\mathbf{f_h}}$ is identical to the estimate in equation (3). If the signal is not white, then this new estimate is going to be more accurate than the estimate in equation (3). More specifically, the noise and signal do not have to be orthogonal any more.

I call this scheme hybrid because it puts back together two worlds: the world of adaptive subtraction and the world of pef. Nonetheless, this method is not a pattern-based technique because the multiples and primaries are not separated according to their spatial predictability. I am only proposing to unbias the filter estimation.

Once the filter has been estimated [equation (5)] I compute the noise and signal as follows:

$$\begin{aligned} \mathbf{n} &= \mathbf{B}\widehat{\mathbf{f_h}} \\ \mathbf{s} &= \mathbf{B}\widehat{\mathbf{f_h}} - \mathbf{d}. \end{aligned} \tag{6}$$

One unsolved problem is the pef estimation. I give few guidelines in the next section.

### How to estimate the pef $\mathbf{A_s}$ for the signal?

It is always difficult to estimate the pef for the signal $\mathbf{A_s}$ because the signal is what we are looking for! It is a chicken and egg story. Fortunately, I have three simple recipes that seem to work well in practice.

The first one has been used by Brown and Clapp (2000) and Guitton (2001) and consists of deconvolving the data pef by the noise pef. It works pretty well but the deconvolution might be unstable when dealing with non-stationary filters (Rickett, 1999). If $\mathbf{A_d}$ and $\mathbf{A_n}$ are the data and noise pef respectively, the estimated signal pef becomes

$$\mathbf{A_s} = \mathbf{A_d}/\mathbf{A_n}. \tag{7}$$

The second one is a simple technique that requires the noise pef only. I first estimate the pef from the noise and apply it to the data. My hope is to obtain a good model for the signal. Then I estimate the signal pef from this model (Spitz, 2001, personal communication).

The last method consists of estimating the signal with the standard adaptive subtraction scheme. A pef is then estimated from the signal and used inside the hybrid scheme. You might be tempted to repeat this process iteratively in order to improve your signal pef. In my experience, I find that an iterative process might diverge, especially when non-stationary filters are involved.

In the next two sections, I show synthetic and field data examples. They prove that the hybrid adaptive filtering gives a better estimate of the noise when primaries and multiples interfere.

## SYNTHETIC DATA EXAMPLE

The synthetic example is inspired by Spitz (1998) and shows clearly that the hybrid method is more accurate than the standard approach.

In Figure 1a, the signal, I show a linear event with a gradient of 1.05 from trace to trace. Figure 1b, the noise, displays another linear event with constant amplitude that perfectly overlaps with Figure 1a. The sum of Figures 1a and 1b gives Figure 1c, the data.

To make matters worse, I applied a phase-shift to the noise in order to compute the noise model (Figure 1d). My goal is to adaptively subtract this noise model from the data in Figure 1c in order to retrieve the signal in Figure 1a.

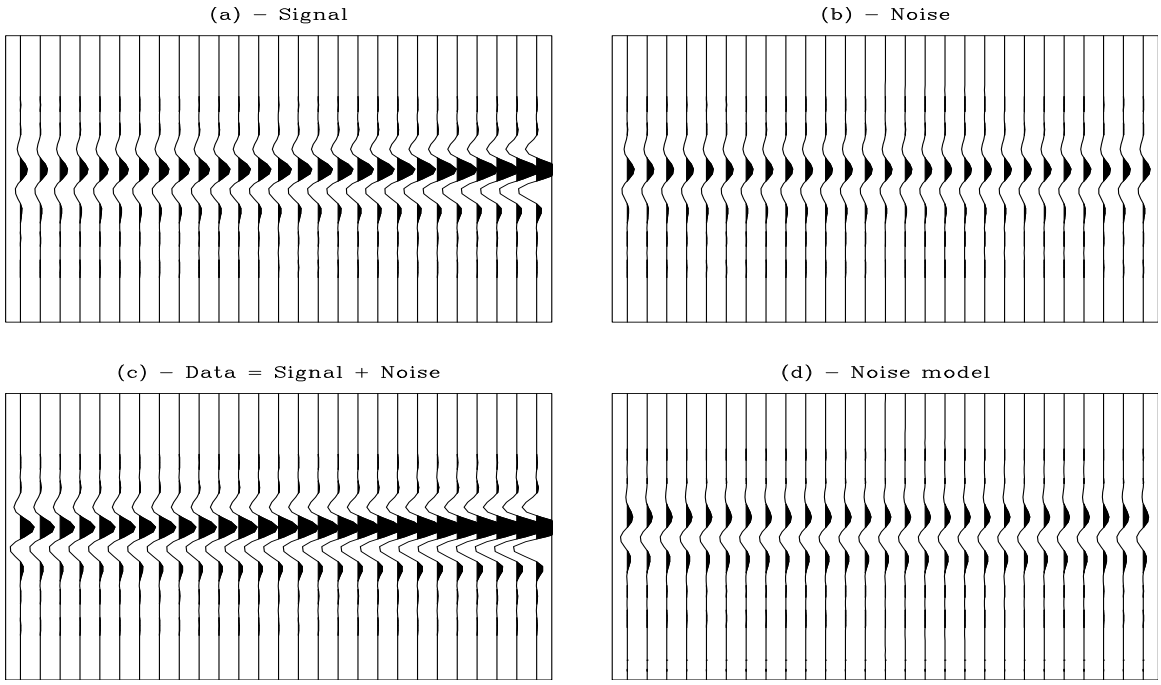



Figure 1: (a) A linear event, the signal, with a gradient of 1.05. (b) A linear event, the noise, with constant amplitude. (c) The sum of (a)+(b), the data. (d) A noise model obtained after applying a phase-shift to (b). antoine2-synth [ER]

Now I estimate the filter with the standard approach and compute the signal in Figure 2b. The estimated signal clearly does not resemble the true signal in Figure 1a.

The hybrid method gives a perfect result. First I estimate a pef with two coefficients for the signal by deconvolving the data pef by the noise pef. I obtain for the signal pef $\mathbf{A_s}' = (\mathbf{1}, -\mathbf{1.05})$. Then I estimate the matched-filter with the hybrid approach and compute the signal in Figure 2c. The separation is perfect as shown in the difference panel in Figure 2d.

In the next section I show prestack land data examples.

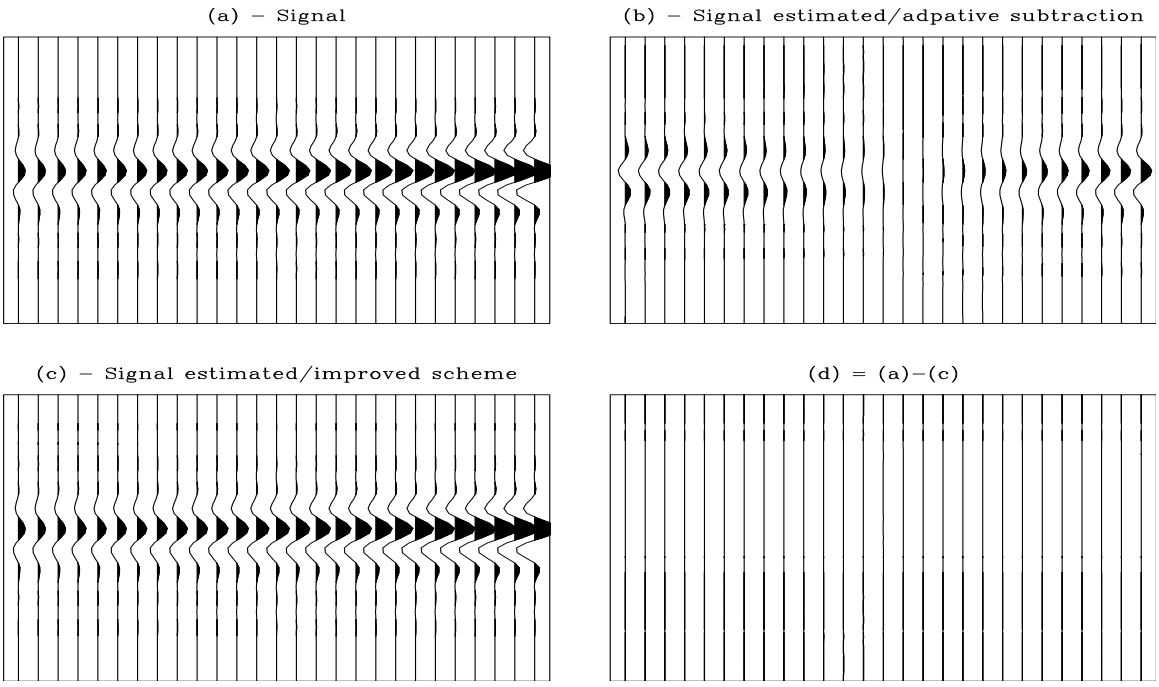



Figure 2: (a) The signal. (b) Estimated signal with the standard approach. (c) Estimated signal with the hybrid approach. (d) Difference between (a) and (c). The noise removal is perfect. antoine2-synth-res [ER]

## PRESTACK LAND DATA EXAMPLES

The hybrid adaptive subtraction technique has been tested on few shot records from a land data survey. My goal is to subtract multiple models from shot records. The preprocessing is described in Kelamis et al. (1999) and the goal is to attenuate surface-related multiples only. Note that for these shot records primaries and multiples are strongly correlated.

To better accomodate for the spatial variability of the data, I do not estimate one filter only but many filters. Instead of estimating filters in patches, I estimate a bank of non-stationary time domain (t, x) filters for both $\mathbf{f}$ and $\mathbf{A_s}$ (Crawley, 2000; Rickett et al., 2001).

Figures 3a and 4a show the land data and the multiple model respectively. Figures 3b and 4b display the estimated signal and noise respectively with the standard approach whereas Figures 3c and 4c show the estimated signal and noise with the hybrid approach. The hybrid subtraction improves the noise subtraction at far offsets. It also preserves the signal better as illustrated in Figure 3c around 1.4 seconds.

A different shot record is processed in Figure 5. The corresponding noise model and estimated multiples are displayed in Figure 6. The same conclusions hold true.

A more interesting shot record with its multiple model are shown in Figures 7a. In this example, the multiple attenuation is greatly improved with the hybrid approach.

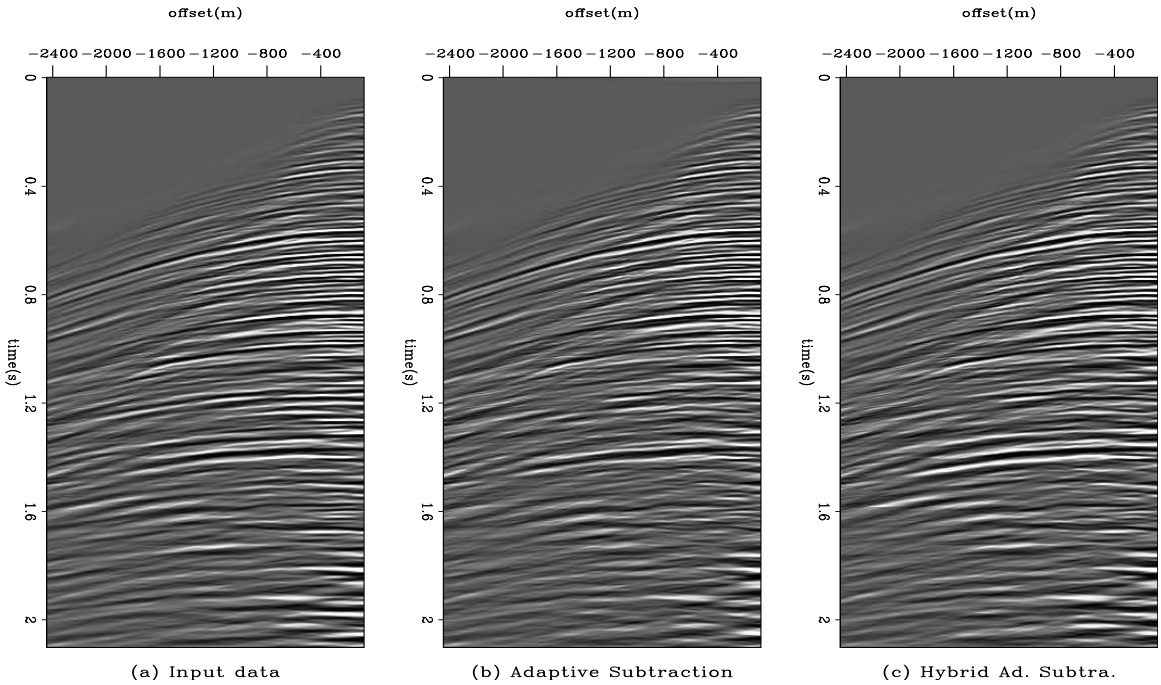



Figure 3: (a) Input data. (b) Estimated signal with the standard approach. (c) Estimated signal with the hybrid adaptive subtraction technique. The strong primary at 1.4 seconds is better preserved with the hybrid adaptive subtraction. antoine2-comp_0 [ER,M]

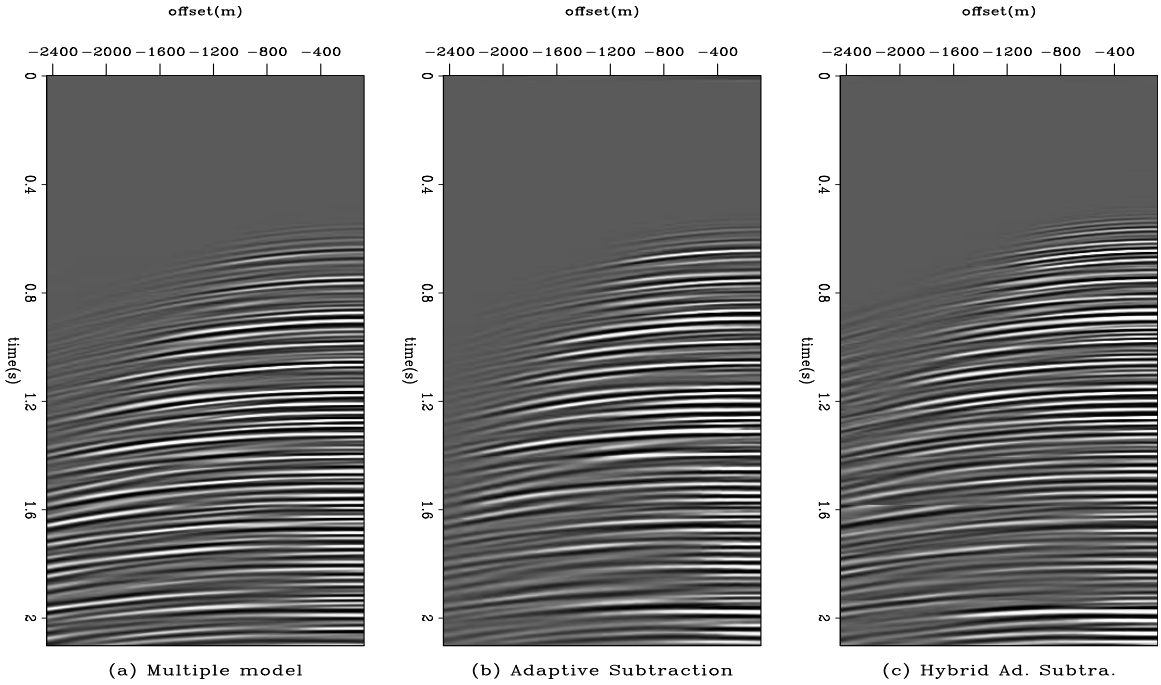



Figure 4: (a) Multiple model for the data in Figure 3. (b) Estimated noise with the standard approach. (c) Estimated noise with the hybrid adaptive subtraction technique. Far offset events are better subtracted. antoine2-comp2_0 [ER,M]

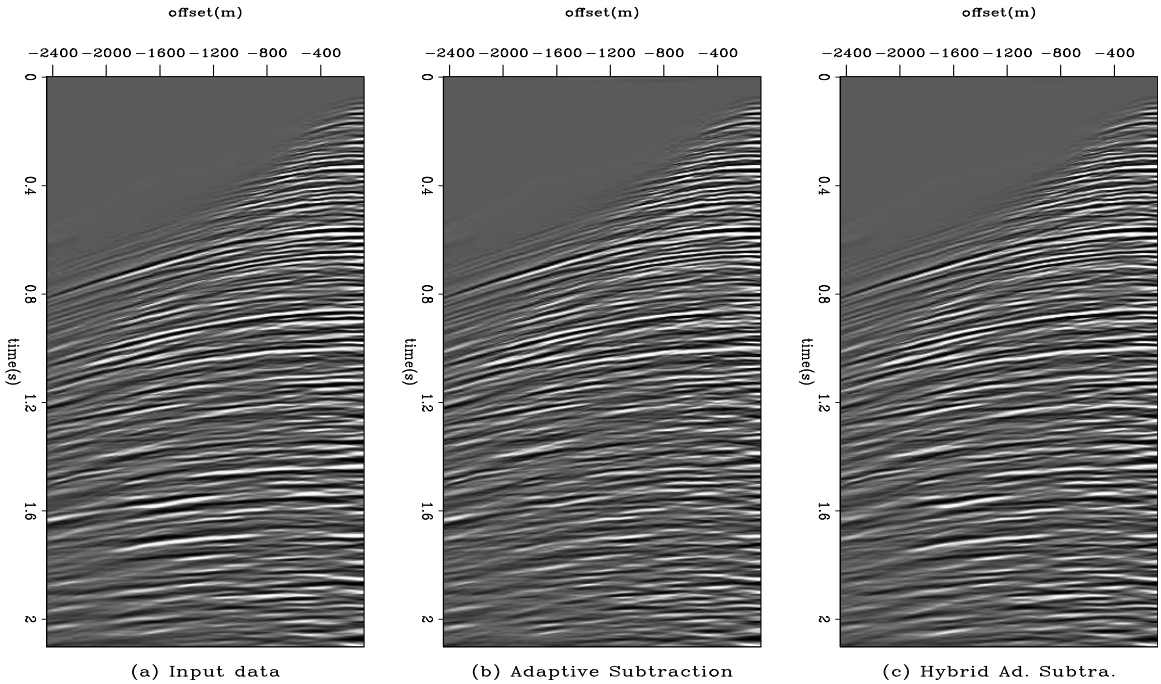



Figure 5: Another shot record. (a) Input data. (b) Estimated signal with the standard approach. (c) Estimated signal with the hybrid adaptive subtraction technique. antoine2-comp_1 [ER,M]

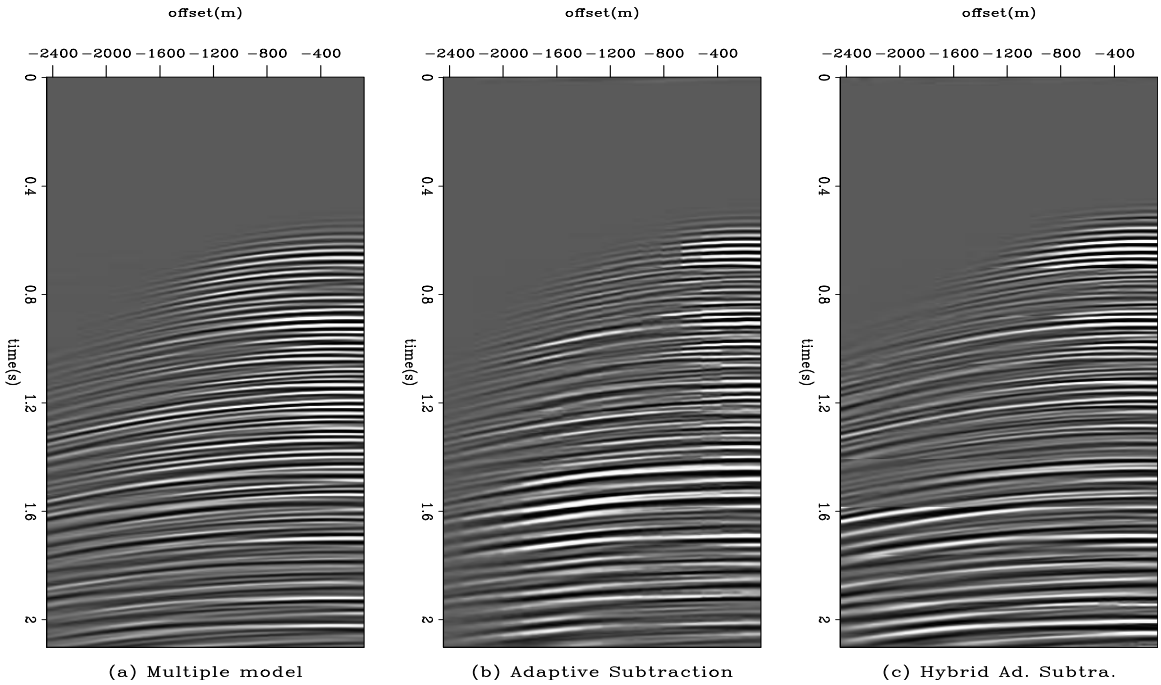



Figure 6: (a) Multiple model for the data in Figure 5a. (b) Estimated noise with the standard approach. (c) Estimated noise with the hybrid adaptive subtraction technique. antoine2-comp2_1 [ER,M]

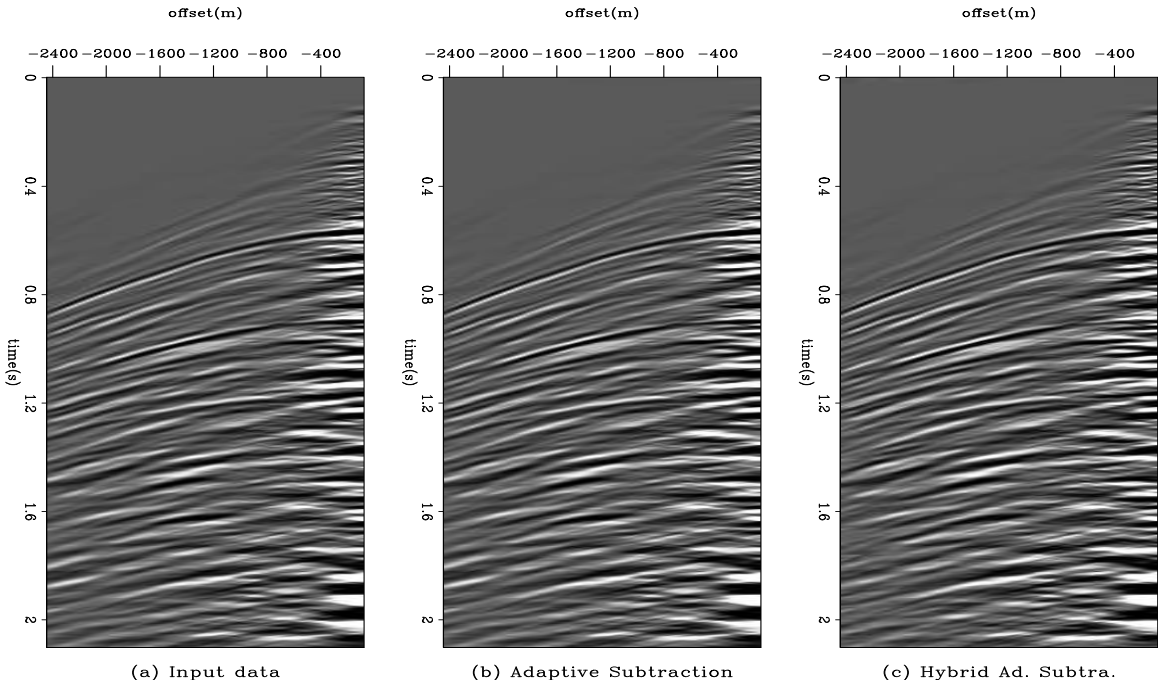



Figure 7: Another shot record. (a) Input data. (b) Estimated noise with the standard approach. (c) Estimated noise with the hybrid adaptive subtraction technique. antoine2-comp_4 [ER,M]

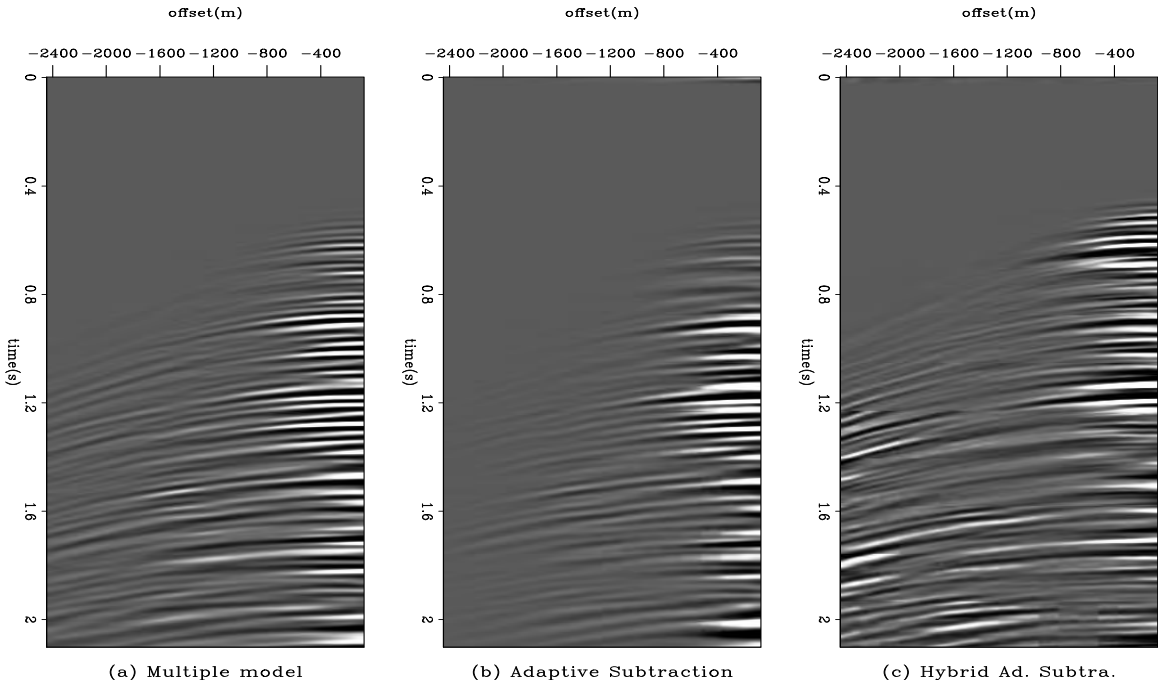



Figure 8: (a) Multiple model for the data in Figure 7a. (b) Estimated noise with the standard approach. (c) Estimated noise with the hybrid adaptive subtraction technique. The hybrid adaptive subtraction approach attenuates more multiples. Far offset events are better subtracted too. antoine2-comp2_4 [ER,M]

The land data examples illustrate the efficiency of the hybrid approach when noise and signal are correlated. In the next section I discuss some limitations of the method and illustrate them with a marine data example.

## LIMITATIONS OF THE METHOD

It is important to keep in mind that the hybrid adaptive subtraction is not a pattern-based method. Noise and signal are not separated because of their spatial predictability. The pef $\mathbf{A_s}$ intends to unbias the filter $\mathbf{f}$ estimation.

In the following example, I show that when the signal/noise ratio is low, the hybrid subtraction is equivalent to the standard approach. In short if the signal level is well below the noise level, then the hybrid adaptive subtraction is equivalent to the standard approach.

To illustrate this last point I extracted one shot gather (Figure 9a) from a Gulf of Mexico 2-D line and estimated the corresponding multiple model (Figure 9b) with the Delft approach (Verschuur et al., 1992). The first surface-related multiples appears at 3.5 seconds within an area where the signal is very weak. Consequently, Figures 9c and 9d which have been processed with the standard and hybrid approach respectively, are identical.

We can recognize in Figure 9 some well-known weaknesses of the adaptive subtraction approach at far offset where the multiples are not well attenuated. Guitton et al. (2001) obtain better results with a pattern-based approach.

## CONCLUSION

I presented a method that improves the signal/noise separation. This method incorporates a prediction-error filter in the standard formulation of the noise removal by adaptive subtraction. By whitening the signal spectrum in the data, this formulation leads to an unbiased estimation of the matched-filter. As a result the noise and signal can be separated when both are correlated.

## ACKNOWLEDGMENTS

I would like to thank Saudi Aramco and WesternGeco for providing the land and marine data, respectively.

## REFERENCES

Brown, M., and Clapp, R., 2000, T-x domain, pattern-based ground-roll removal: 70th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 2103–2106.

Crawley, S., 2000, Seismic trace interpolation with nonstationary prediction-error filters: SEP–**104**.

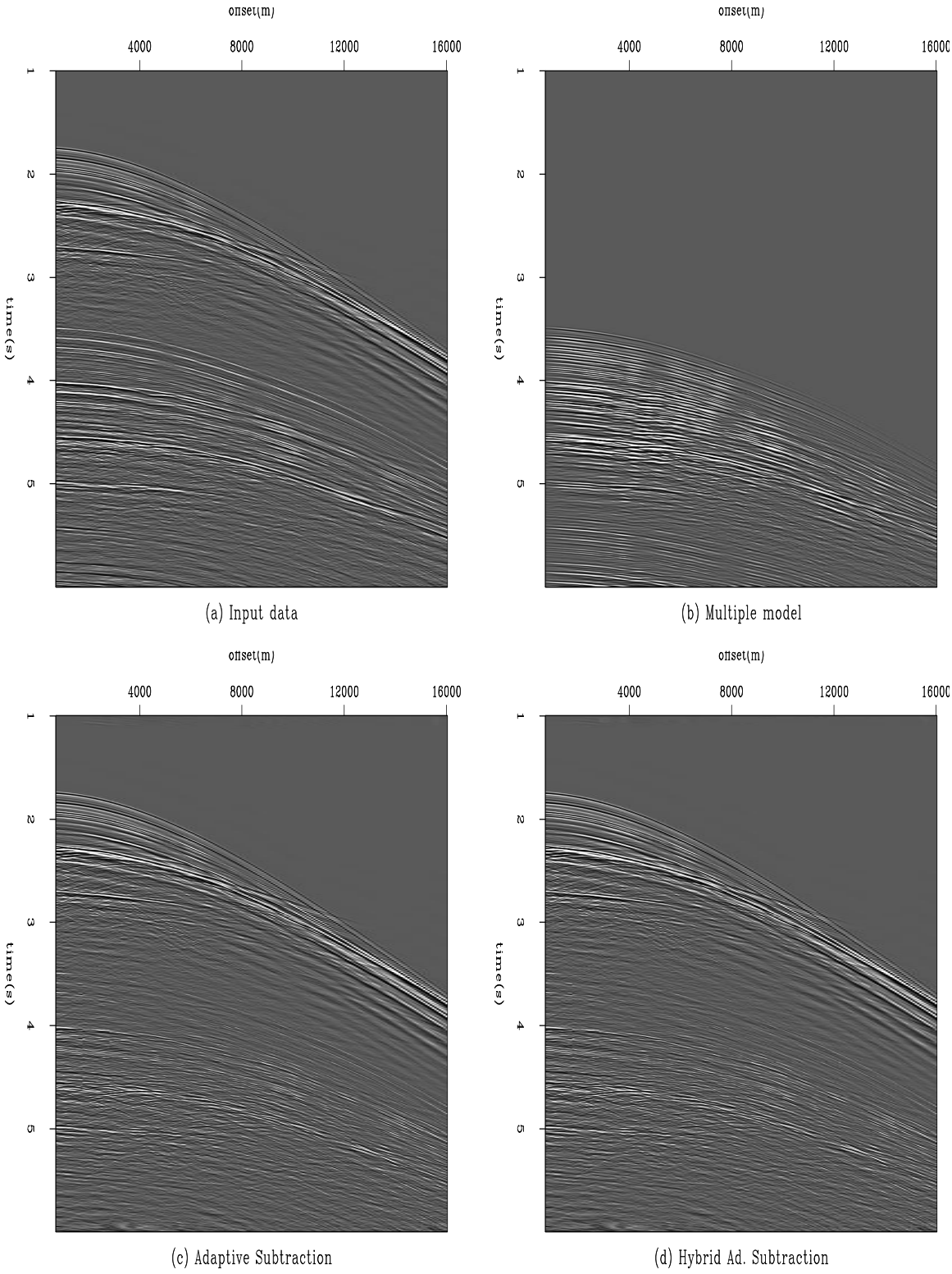



Figure 9: (a) A shot gather from a Gulf of Mexico 2-D line. (b) The corresponding multiple model. (c) Estimated primaries with the standard approach. (d) Estimated primaries with the hybrid approach. Both estimated signal are similar. antoine2-gom_res [ER,M]

Guitton, A., Brown, M., Rickett, J., and Clapp, R., 2001, Multiple attenuation using a t-x pattern-based subtraction method: 71st Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1305–1308.

Guitton, A., 2001, Coherent noise attenuation: A synthetic and field example: SEP–**108**, 225–248.

Guitton, A., 2002, Coherent noise attenuation using inverse problems and prediction error filters: First Break, **20**, 161–167.

Kelamis, P., Verschuur, D., Berkhout, A., and Erickson, K., 1999, Velocity-independent datuming of seismic data: 69th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 441–444.

Rickett, J., Guitton, A., and Gratwick, D., 2001, Adaptive Multiple Subtraction with Non-Stationary Helical Shaping Filters: 63rd Mtg., Eur. Assn. Geosci. Eng., Expanded Abstracts, Session: P167.

Rickett, J., 1999, On non-stationary convolution and inverse convolution: SEP–**102**, 129–136.

Spitz, S., 1998, Pattern Recognition and Subtraction of Coherent Noise: 60th Mtg., Eur. Assn. Geosci. Eng., Expanded Abstracts, Session: 01–09.

Spitz, S., 1999, Pattern recognition, spatial predictability, and subtraction of multiple events: The Leading Edge, **18**, no. 1, 55–58.

Verschuur, D. J., Berkhout, A. J., and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, no. 09, 1166–1177.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 183–199*

# Removal of coherent noise from electroseismic data

*Seth Haines and Antoine Guitton*[1]

**ABSTRACT**

The electroseismic method offers the possibility of imaging thin (much smaller than the seismic wavelength) layers in the subsurface. Removal of coherent source-generated noise is an essential step in the processing of electroseismic data. We present a signal-noise separation technique that begins with the determination of signal and noise prediction error filters (PEF's) from windows of the original data and uses these PEF's in an iterative inversion for signal and noise models. The noise residual is weighted during the inversion to prevent the comparatively weaker (by about two orders of magnitude) signal from being obscured by the coherent noise. Application of this processing sequence to real data, in which synthetic signal arrivals are obscured, demonstrates the effectiveness of the technique.

## INTRODUCTION

A seismic wave traveling through a fluid-saturated porous material carries with it a charge separation created by the pressure-induced flow of pore fluid. The pore fluid carries a small (but not inconsequential) amount of electric charge relative to the adjacent grains due to the electric double layer (Shaw, 1992) that exists at the grain-fluid boundary. Thus, an electric field (Figure 1) is co-located with a compressional (P) wave propagating through such a material (Pride, 1994). We refer to this field as the "coseismic" field.

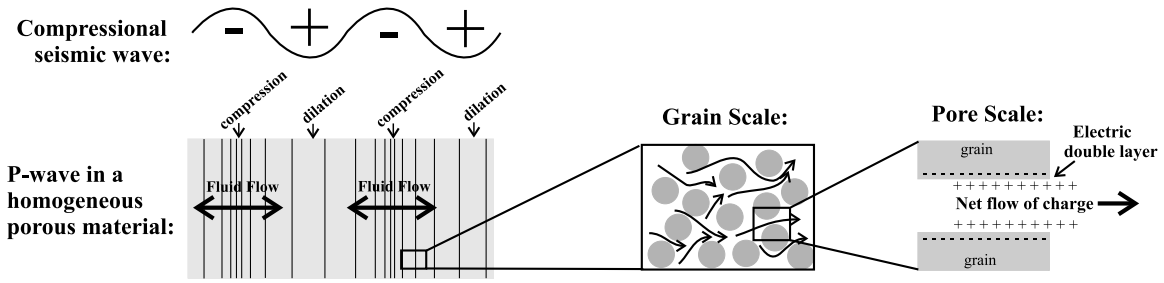



Figure 1: Electroseismic phenomena depend on the charge separation created by streaming currents that flow in response to the pressure gradient of a seismic wave. The electric double layer is responsible for streaming currents at the grain scale. shaines1-esbasics [NR]

[1]**email:** shaines@pangea.stanford.edu, antoine@sep.stanford.edu

The second aspect of the electroseismic response occurs when the P-wave encounters an interface in material properties (elastic, chemical, flow-related, etc). The charge separation in the wave (Figure 2) is disturbed, causing asymmetry and results in what can be approximated as an oscillating electric dipole at the first Fresnel zone (Haartsen and Pride, 1997). Essentially, the entire region of the first Fresnel zone acts as a disk of vertical electric dipoles. Thus, the resulting electric field is that of a dipole, with opposite polarity on opposite sides of the source point and amplitudes diminishing as $1/z^3$ (where $z$ is the depth to the interface). This field (Figure 2b), called the "interface response," can be measured almost immediately at the Earth's surface since the travel-time of electromagnetic radiation is negligible compared with seismic travel-times ($V_{EM}\gamma^2 V_P$).

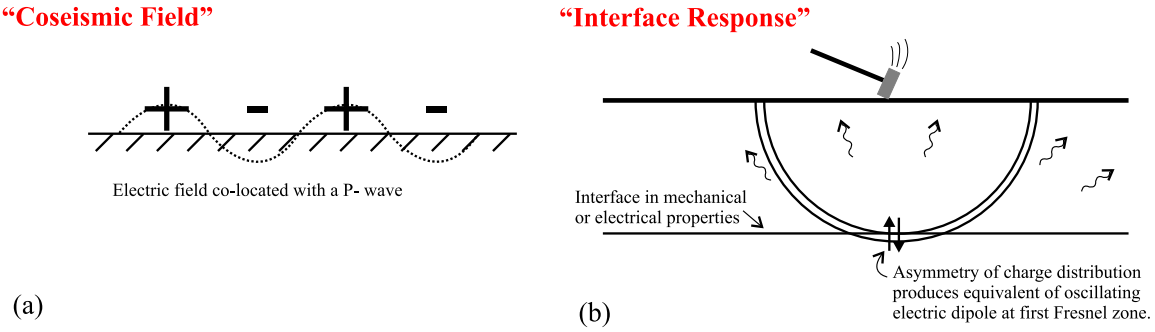



Figure 2: Two types of electroseismic effects that can be measured with electrode dipoles at the Earth's surface: (a) the coseismic field of a P-wave at the surface (represented here by the charge accumulations "+" and "-"), and (b) the interface response created when the P-wave hits an interface at depth. shaines1-2effects [NR]

Both effects can be measured in the field using a standard seismograph equipped with electrode dipoles instead of geophones (Haines et al., 2001; Garambois and Dietrichz, 2001; Thompson and Gist, 1993). The coseismic field traveling with the seismic wave is not particularly interesting since it contains information only about the properties of the surface of the Earth. The interface response, on the other hand, can provide new information about the subsurface. In particular, the interface response is created even for very thin layers, such as a thin fracture zone in otherwise solid rock, or a thin impermeable layer in an aquifer or reservoir. Haines et al. (2001) show that the interface response from a saturated permeable layer 0.6-m thick can be reliably observed. Numerical simulations show that the interface response from a 1-cm embedded impermeable layer is significantly greater than that from an interface between two layers (Stephane Garambois, personal communication 2001). Thus the electroseismic method promises to provide valuable information about important subsurface targets that can not be imaged using other geophysical methods, including information about the location of changes in flow properties. It is our goal to develop a protocol that can be used to reliably and repeatably acquire, process, and interpret electroseismic data.

Unfortunately, electroseismic data collected with a geometry similar to conventional surface seismic data is comprised of both the interface response from subsurface layers and unwanted coseismic energy recorded simultaneously. Coseismic energy, roughly 100 times the amplitude of the interface response, therefore represents a formidable form of coherent

source-generated noise. The removal of this noise is essential to the utility of the electroseismic method, so the development of an effective data processing approach is an important step toward this goal. The use of transforms (e.g., f-k filtering) has proven ineffective on available data due to the overpowering amplitude of the coseismic noise, and the fact that the top of the coseismic energy hyperbola tends to be smeared across the record during the inverse transform. It is essential that all horizontal energy remaining in the record after processing be only from the the interface response. Each shot record eventually will be stacked to produce a single trace corresponding with the subsurface region beneath the shot point. Thus, smeared coseismic energy would be very detrimental to the final stack in much the way that inclusion of ground roll or refractions negatively impacts a stacked seismic reflection section. In addition, the smearing of energy represents a loss of amplitude information and the retention of relative amplitudes is desired.

We present a data processing strategy that separates the signal of interest from the stronger coherent noise. This strategy incorporates the coherent noise subtraction approach described by Guitton (2001) while building on the use of PEF's described by Claerbout and Fomel (2001). An important feature of our approach is its preservation of the signal amplitude made possible by our use of iterative inversion. We also present other data processing options and discuss the work remaining to be completed before the electroseismic method can be considered a reliable tool.

## EXPERIMENTAL DATA

Haines et al. (2001) present electroseismic experimental methods designed to record the coseismic field and the interface response separately. By imaging a vertical interface with the source and receivers on opposite sides of the target (Figure 3), they record the interface response before the coseismic energy (Figure 4). The target is a trench ∼0.6m wide, ∼2m deep, and ∼18m long, lined with plastic, and filled with wet sand.

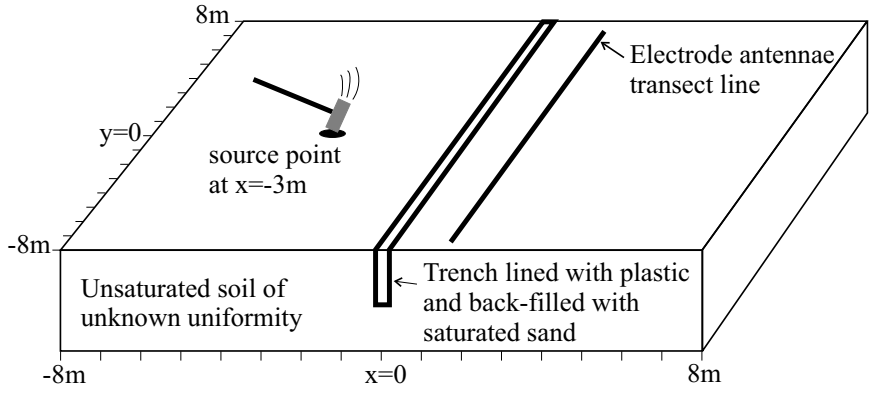



Figure 3: Layout of field site. Note that source and receivers are on opposite sides of the target so that the seismic wave creates the interface response at the trench before reaching the electrodes and being recorded as the coseismic energy. shaines1-layout [NR]

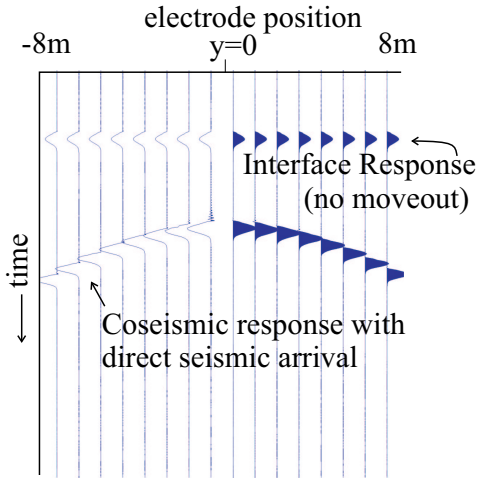



Figure 4: Generalized data, showing interface response arriving before coseismic energy. Note reversed polarity on opposite sides of shot point. (For simplicity, relative amplitudes are NOT correct.) shaines1-gen_data [NR]

Approximately 80 shot records have been recorded at this site with various shot and receiver geometries. Shot gathers are generally stacks of 50 to 100 strikes of a 12 lb sledgehammer on a metal plate located 1 to 4 meters from one side of the trench. Recording geometries include 24 electrode dipoles ($\sim$1 meter wide) at $\sim$0.7 m spacing located 1 to 4 meters from the other side of the trench. The interface response is recognizable before the coseismic arrival on every record, with varying clarity and strength. Arrival times and simple amplitude modeling confirm that the observed signal is indeed the interface response. Because the interface response shows virtually no moveout, and because it has reversed polarity on opposite sides of the shot point, it can be visually differentiated from the coseismic arrival and background electrical noise. These data, with the two effects recorded separately, provide a unique opportunity for development of a processing sequence to enhance the interface response and remove the coseismic signal.

## SIGNAL-NOISE SEPARATION

Electroseismic data (**d**) can be thought of as the sum of three distinct elements- the interface response signal (**s**), the coseismic noise (**n**), and background electrical noise ($\mathbf{n_{bg}}$):

$$\mathbf{d} = \mathbf{s} + \mathbf{n} + \mathbf{n_{bg}}. \tag{1}$$

In order to separate the interface response from the coseismic noise, we implement a signal-noise separation technique using separate PEF's $\mathbf{A_s}$ and $\mathbf{A_n}$ for the signal and noise respectively. Because much of $\mathbf{n_{bg}}$ is not readily predictable, it cannot be modeled with a PEF. We minimize $\mathbf{n_{bg}}$ with pre-processing.

### Pre-processing

Raw electroseismic data is dominated by energy from the power grid at harmonics of 60 Hz (Figure 5a). We remove this noise using the sinusoid subtraction technique of Butler and Rus-

sell (1993) for all harmonics of 60 Hz up to the Nyquist frequency. Coseismic energy is generally lower in frequency than the interface response due to a greater distance of travel as a seismic wave, so we use a low-cut filter to begin the process of noise removal. We also employ a high-cut filter to minimize background noise that can obscure weaker arrivals (Figure 5b).

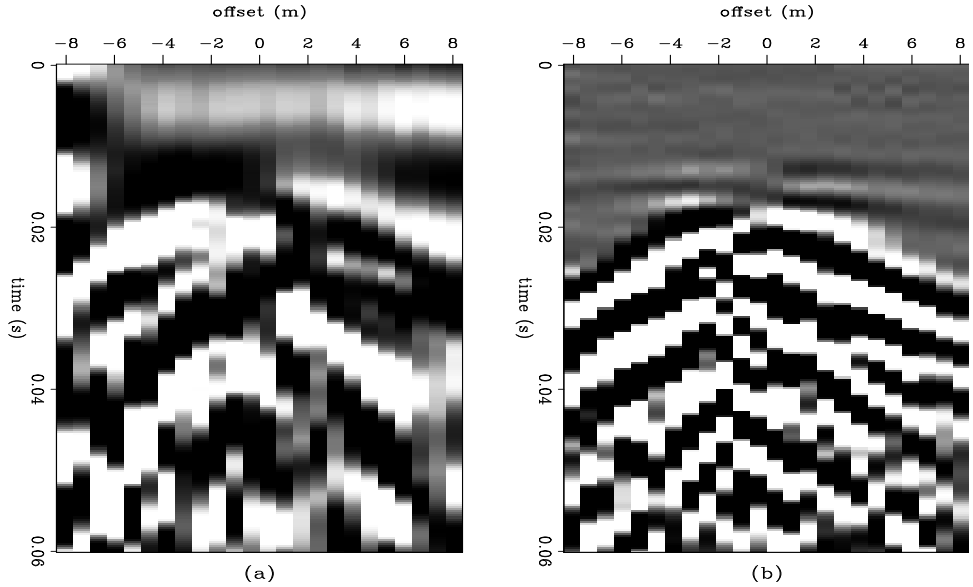



Figure 5: (a) Raw data. (b) Same file, with 60 Hz energy removed and bandpass filter (120-500 Hz) applied. Note horizontal interface response event at 0.015 seconds. shaines1-basic526 [ER,M]

**Processing approach**

The PEF's $\mathbf{A_s}$ and $\mathbf{A_n}$ are determined from windows of the data file that are particularly representative of each. It is essential to have good design windows for the two PEF's, representative of the two components **s** and **n**. We determine only one PEF for the entire data file, thus assuming (perhaps unwisely) stationarity. Using the PEF's $\mathbf{A_s}$ and $\mathbf{A_n}$, we perform an iterative inversion to determine signal and noise models $\mathbf{m_s}$ and $\mathbf{m_n}$ with the fitting goal (Guitton, 2001):

$$\mathbf{0} \approx \mathbf{L_s m_s} + \epsilon \mathbf{L_n m_n} - \mathbf{d}, \tag{2}$$

where

$$\mathbf{L_s} = \mathbf{A_s^{-1}} \text{ and } \mathbf{L_n} = \mathbf{A_n^{-1}}.$$

We solve the following least squares inverse formulation of equation (2):

$$\begin{pmatrix} \widehat{\mathbf{m}}_\mathbf{n} \\ \widehat{\mathbf{m}}_\mathbf{s} \end{pmatrix} = \begin{pmatrix} \epsilon^{-1}(\mathbf{L_n'}\overline{\mathbf{R_s}}\mathbf{L_n})^{-1}\mathbf{L_n'}\overline{\mathbf{R_s}} \\ (\mathbf{L_s'}\overline{\mathbf{R_n}}\mathbf{L_s})^{-1}\mathbf{L_s'}\overline{\mathbf{R_n}} \end{pmatrix} \mathbf{d}, \tag{3}$$

with

$$\begin{aligned} \overline{\mathbf{R_s}} &= \mathbf{I} - \mathbf{L_s}(\mathbf{L_s'}\mathbf{L_s})^{-1}\mathbf{L_s'}, \\ \overline{\mathbf{R_n}} &= \mathbf{I} - \mathbf{L_n}(\mathbf{L_n'}\mathbf{L_n})^{-1}\mathbf{L_n'}, \end{aligned} \tag{4}$$

as defined by Guitton et al. (2001). The estimated noise $\widehat{\mathbf{n}}$ and signal $\widehat{\mathbf{s}}$ are then computed by

$$\begin{aligned} \widehat{\mathbf{n}} &= \mathbf{L_n}\widehat{\mathbf{m}}_\mathbf{n}, \\ \widehat{\mathbf{s}} &= \mathbf{L_s}\widehat{\mathbf{m}}_\mathbf{s}. \end{aligned} \tag{5}$$

The weighting factor $\epsilon$ in equation (3) is crucial to this process. As $\epsilon \rightarrow 1$, $\mathbf{L_s}\widehat{\mathbf{m}}_\mathbf{s}$ is given too little importance in the inversion relative to $\mathbf{L_n}\widehat{\mathbf{m}}_\mathbf{n}$ and $\widehat{\mathbf{m}}_\mathbf{s}$ is polluted with large amounts of noise. Physically, $\epsilon$ represents the magnitude difference between the two distinct physical processes that make up the electroseismic response. These two proceses are: the generation of electromagnetic energy by an oscillating electric dipole (observed from a distance, called the interface response), and the generation of an electric field within a passing seismic wave (measured directly, called the coseismic field).

Because the fitting goal does not include a term for $\mathbf{n_{bg}}$, any background noise remaining in the data after pre-processing simply falls into the residual. This is necessary because $\mathbf{n_{bg}}$ is difficult to model, and acceptable because $\|\mathbf{n_{bg}}\| < \|\mathbf{s}\| \ll \|\mathbf{n}\|$. Convolution of the PEF's with the data is accomplished using the helix of Claerbout (1998).

To summarize the processing sequence:

1. Frequency filtering and 60 Hz removal
2. Second time derivative and other preprocessing steps described later
3. Determine $\mathbf{A_s}$ and $\mathbf{A_n}$ from windows of data file
4. Iteratively solve the inverse problem of equation (3). Output $\widehat{\mathbf{s}} = \mathbf{L_s}\widehat{\mathbf{m}_\mathbf{s}}$.

**Implementing the data processing sequence**

Because we determine only one set of PEF's for a given processing effort, we improve the ability of the PEF's to model their respective parts of the data by processing only half the traces at one time (in this case, the positive offsets). Therefore, the noise is more easily modeled since all coseismic energy dips in the same direction. We include the three nearest negative-offset traces so that the final result (after "losing" traces to convolution with PEF's) includes all of the positive offset traces.

Data processing begins with the selection of suitable windows for determination of PEF's $\mathbf{A_s}$ and $\mathbf{A_n}$. With these windows chosen, we can test different parameters for the rest of the processing, namely the sizes of the two PEF's, the value of the weighting factor $\epsilon$, and the number of iterations for the inversion.

Because the signal of interest is horizontal, it is logical to use a signal PEF $\mathbf{A_s}$ that has only one element in the vertical direction. This assumption is confirmed by experimentation, as taller PEF's are less effective at modeling non-horizontal energy. We find that a width of 4 is a fair trade-off between the better resolution of a wider PEF and the smaller number of

traces lost to convolution afforded by a narrower PEF. The size of the noise PEF $\mathbf{A_n}$ is more arbitrary, but we find that dimension 4x3 yields satisfactory results.

$\epsilon$ is best determined experimentally, though the development of a more rigorous approach would improve the versatility of the processing sequence. Interestingly, a broad range of $\epsilon$ values results in essentially the same final result. We find $\epsilon = 0.01$ to be a reasonable value for the data shown here. Finding a reasonable criteria to define the end of the iterative inversion process is another aspect of this processing sequence that deserves more attention. We have simply stopped the inversion after 350 iterations.

## Testing the processing sequence

Data collected by Haines et al. (2001) was designed to record the interface response and coseismic energy separately as a test data set for processing. However, these files lack any additional interface response events after the onset of coseismic energy (Figure 6a). In an electroseismic data set collected with a standard geometry, we would be looking for interface response events within the coseismic energy. For that reason, we followed the example of Brown and Clapp (2000) and constructed a test data file by windowing out (Figure 6b) the interface response from a record (Figure 6a) using a $sin^2$ taper in the time direction and adding it to the original file within the coseismic energy. We did this twice (Figure 6c), adding the synthetic arrival at two different times in the record. The resulting record (Figure 6d) appears essentially identical to the original file, but contains two interface response events hidden within to test our methodology.

Using the PEF design windows shown (Figure 7b and c), and the parameters mentioned above, we solve the inverse problem (equation 3). Figure 7d shows $\widehat{\mathbf{s}}$, our best estimate of $\mathbf{s}$, the component of the data that is made up of the interface response energy. The two added interface response events are evident, as is the original interface response. Unfortunately, a considerable amount of dipping coseismic energy still remains between 0.025 and 0.04 seconds, and the added event below 0.06 seconds is somewhat weak. Despite these shortcomings, the final result (Figure 7d) is considerably better than the original data file (7a).

## Additional pre-processing steps

The results shown in Figure 7 demonstrate that the basic algorithm can be effective. However, the low amplitude of the added interface response events relative to the remaining coseismic noise suggests that additional pre-processing steps should be explored in order to improve the final result. In this vein, we tested the use of a second time derivative as a means of balancing the amplitudes of the interface response relative to the coseismic energy. This spectral balancing enhances the generally higher-frequency interface response. This step sharpens the image (Figure 8a), and improves the ability of the PEF to precisely locate the signal amidst the noise. Applying the same processing steps as used for Figure 7 and using the same parameters, we find that the end result (Figure 8d) is markedly improved over Figure 7d, with greater clarity of the added events.

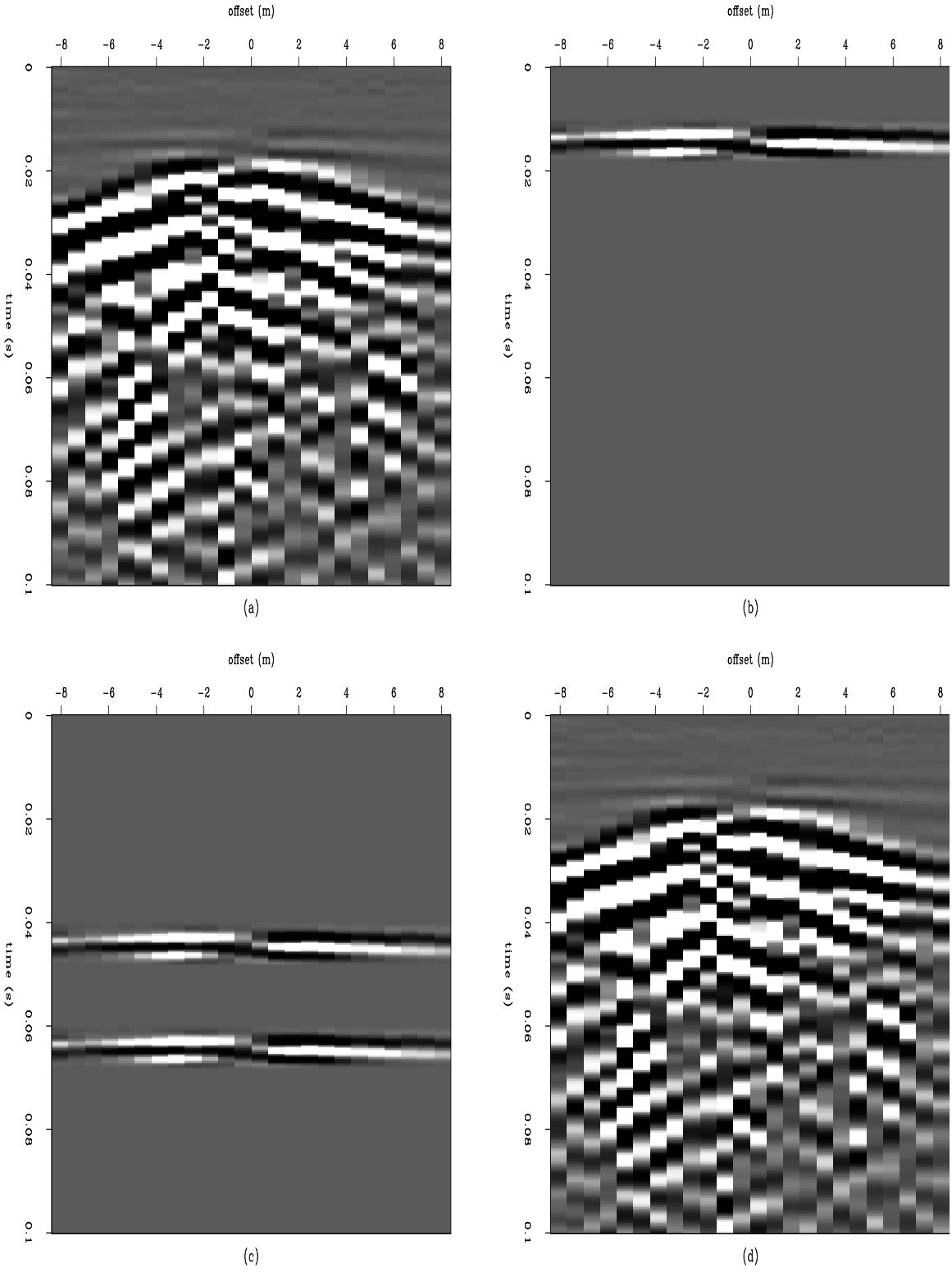



Figure 6: (a) Data file. (b) Interface response (IR) from (a), windowed with $sin^2$ taper. (c) Two IR events to be added to data. (d) Test data- starting data plus two added events. shaines1-fake [ER,M]

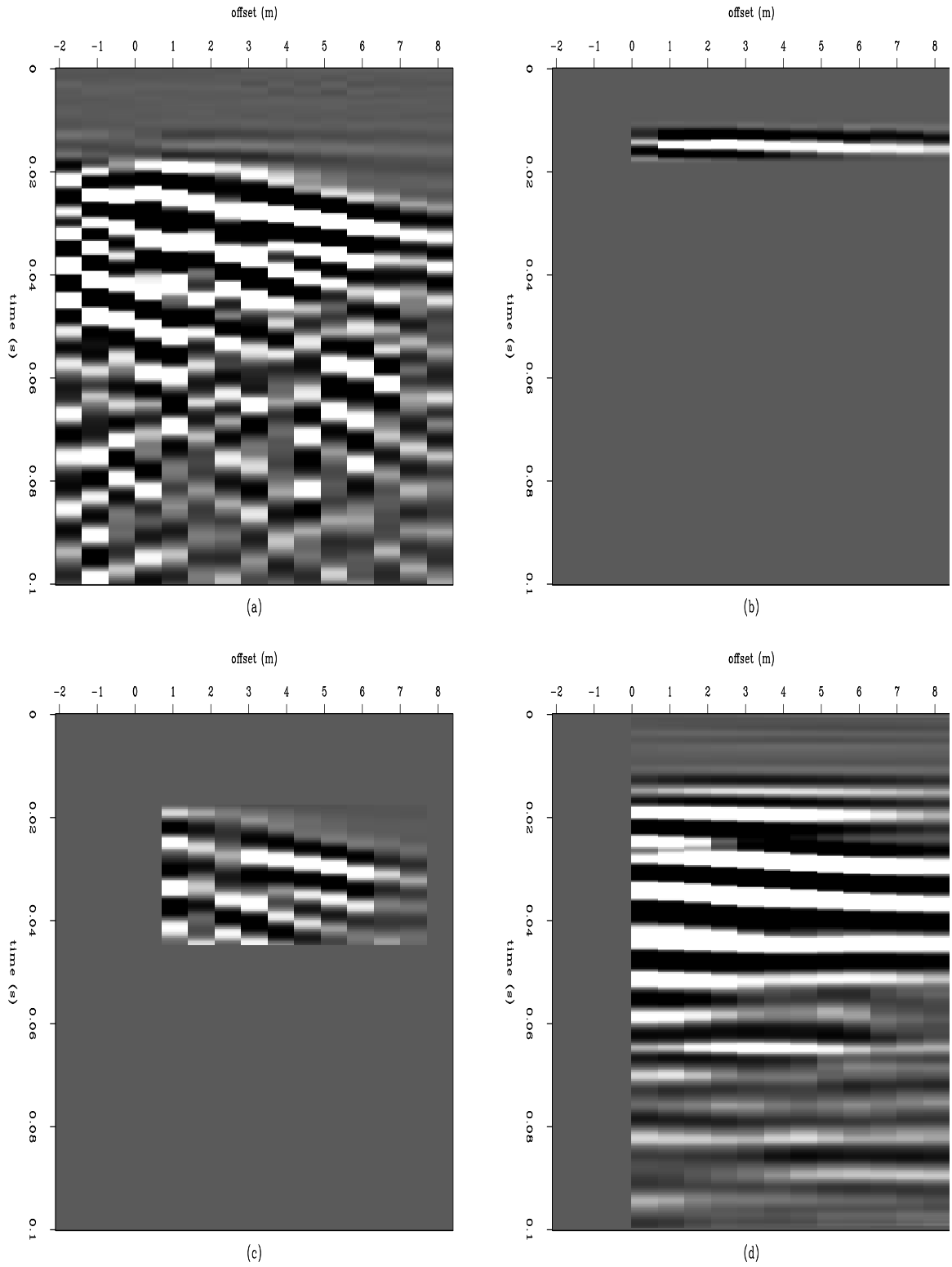



Figure 7: (a) Data file (windowed version of file in 6d). (b)Window to be used for determination of signal PEF $\mathbf{A_s}$. (c)Window used for determination of noise PEF $\mathbf{A_n}$. (d) Output of processing, $\mathbf{L_s}\widehat{\mathbf{m}}_\mathbf{s}$. Note that added horizontal events 0.045 s and 0.065 s are visible. shaines1-nice [ER,M]

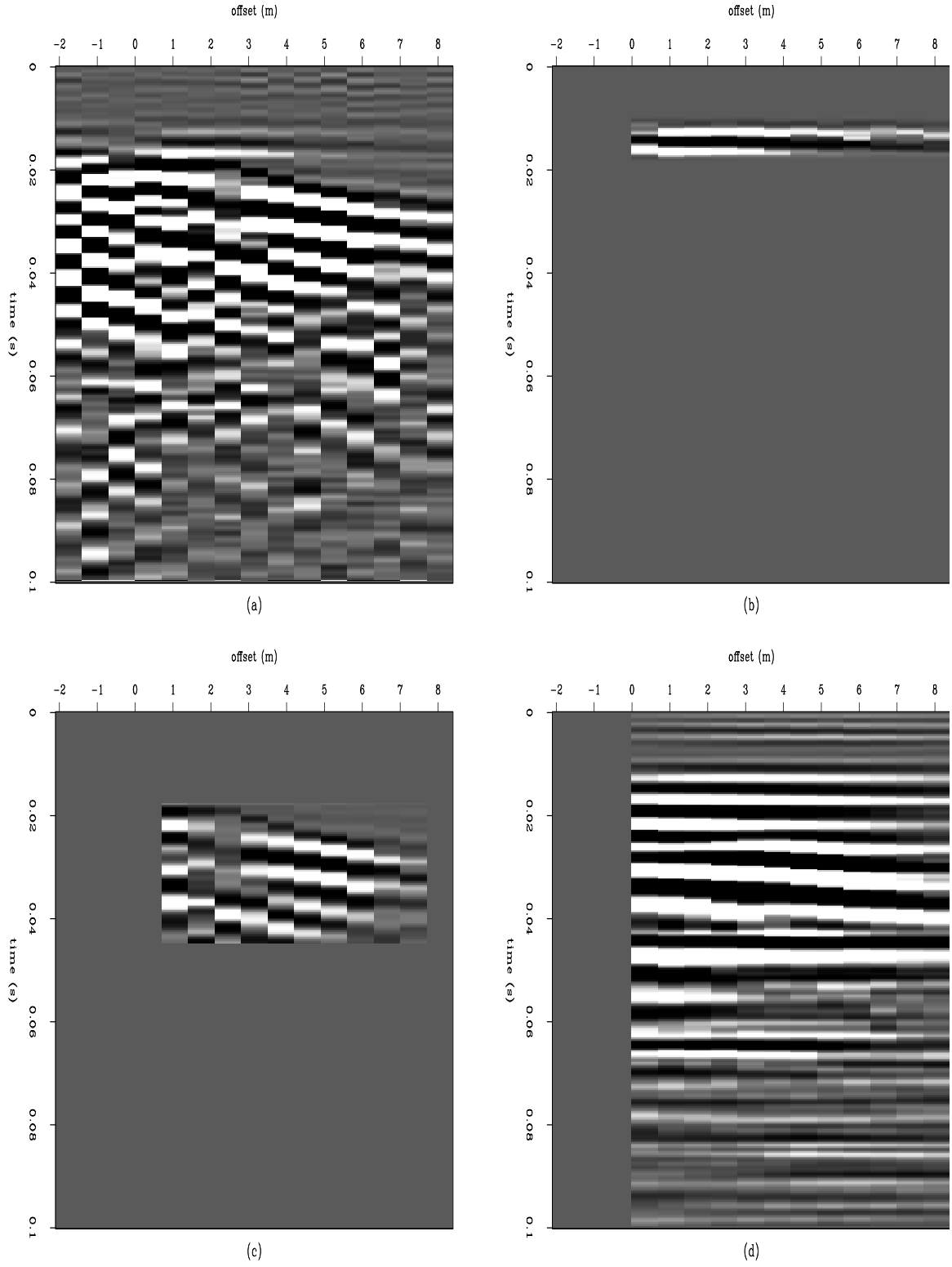



Figure 8: (a) Data file (second time derivative of file in 7a). (b) Window for determination of signal PEF $\mathbf{A_s}$. (c) Window for determination of noise PEF $\mathbf{A_n}$. (d) Output of processing, $\mathbf{L_s}\widehat{\mathbf{m}}_\mathbf{s}$. Note that added horizontal events 0.045 s and 0.065 s are visible. shaines1-nice2D [ER,M]

Still present is the dipping coseismic energy between 0.025 and 0.04 seconds. Because this energy is so strong, and so close to horizontal, it leaks through the PEF, corrupting a portion of the record. This problem is the focus of on-going investigation and is being addressed with waveform separation.

Waveform separation is a standard technique in VSP and cross-well seismic processing, used to remove high-amplitude early arrivals in records by capitalizing on the different move-outs of different arrivals. We employ the method as follows: (1) picks are manually made along the arrival targetted for removal, (2) the gather is moved out such that the arrival (as defined by the picks) is horizontal and then stacked, and (3) the resulting trace is normalized by the number of traces in the gather and then subtracted from each of the moved-out traces. (4) After the subtraction, the gather is moved back to its original alignment. By repeating this process, it is possible to remove more than one coherent arrival from the record. Figure 9 shows a series of images as various arrivals are removed from the record (9a). Figure 9d shows the result after three iterations through the process, and shows that although the process has effectively removed much of the energy of the strong coseismic first arrivals, it has also partially removed the interface response (0.01 to 0.02 seconds). This is a result of the chance line-up of waveforms during the second iteration. We chose to use the data shown in Figure 9c for the PEF processing sequence.

We apply the second derivative after waveform separation since it has proven successful. Because the coherence of the coseismic noise has been disturbed by the waveform separation technique, we opt to determine the PEF's $\mathbf{A_s}$ and $\mathbf{A_n}$ with the data shown in the windows of Figure 8b and c. The starting datafile (after waveform separation and second derivative) is shown in Figure 10a, with the final result in Figure 10b. Here we see that the clarity of the added events is improved, but that some of the coseismic energy remains. In this case, the remaining coseismic energy is closer to horizontal than that in Figure 8d. Thus, if we were to stack this gather, the resulting trace would definitely include unwanted coseismic energy. We continue to pursue solutions to this problem.

## SUMMARY

At this time, our preferred processing sequence is:

1. Remove 60 Hz energy
2. Bandpass filter
3. Waveform separation
4. Second time derivative
5. Specify windows for determination of signal and noise PEF's
6. Weighted inversion for signal and noise models. Output is $=\widehat{\mathbf{s}}\,(=\mathbf{L_s}\widehat{\mathbf{m}}_\mathbf{s}\approx\mathbf{s})$.

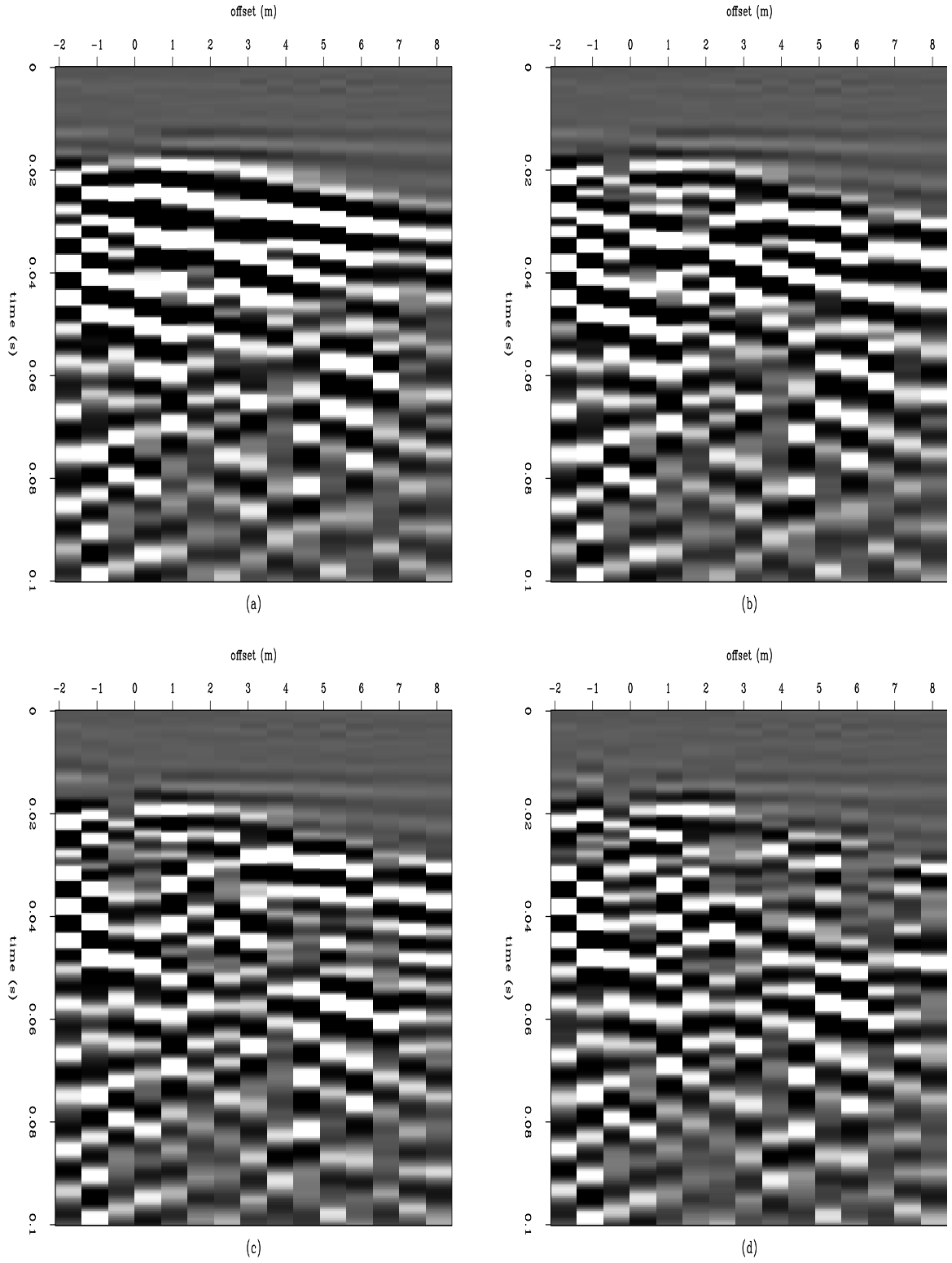



Figure 9: (a) Data file (same as 7a). (b) Result after one iteration through waveform separation process. (c) Result after second iteration. This is the file that we use in the PEF processing sequence. (d) Result after three iterations through the waveform separation process. Note that the interface response (0.01 to 0.02 seconds) event has been partially removed. shaines1-xwell [ER,M]

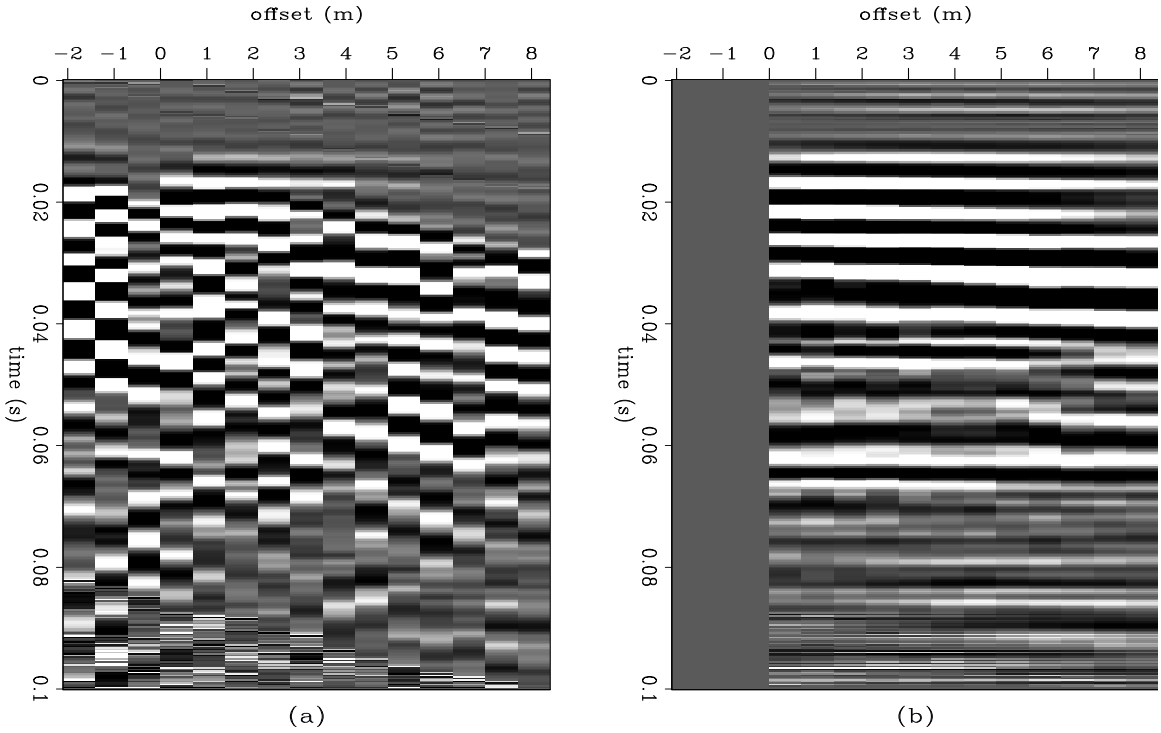



Figure 10: (a) Data file (Second derivative applied to data from Figure 9c). (b) Result after PEF processing sequence. Note that added horizontal events 0.045 s and 0.065 s are clearly visible. shaines1-nicexwell [ER,M]

## DISCUSSION

The processing sequence presented here is shown to be effective in separating the electroseismic interface response from the much stronger coseismic noise for the artificial example used. This is interesting both as an example of signal-noise separation in a case with very low signal-to-noise ratio and also as an important step toward proving the utility of the electroseismic method. However, the results we present are unrealistic in several important ways. Most importantly, the interface response events that we separate from the noise are essentially synthetic, and undoubtedly have different amplitude and frequency patterns than real arrivals from subsurface layers. We can attempt to make the inserted arrivals more realistic by stretching and scaling them before addition, but no synthetic record can really replace real data.

Thus the next logical step in this project is to collect a more realistic electroseismic data set at a location that is thoroughly-characterized. It is important that the characterization include wells, because the method promises to image layers that are invisible to existing surface methods. An ideal survey would take place at a test site with relatively simple subsurface, but with a few thin layers (fractures, clay lenses, etc) to be targeted. The use of a vibratory source or a high-intensity explosive source would vastly ease the recognition of the signal amidst background noise, and adding more recording channels would tremendously improve data quality and the output of data processing efforts.

A greater number of traces per shot gather will permit refinement of the processing scheme. The use of non-stationary PEF's may prove effective, since the nature of the coseismic energy clearly varies within the shot record. The use of a gap in the noise PEF may be an improvement, as it could prevent the PEF from being dominated by energy corresponding with infinite slope due to the spiky appearance of the coseismic data. In addition, a larger number of traces would improve the probability of success of transform-based methods such as f-k filtering and slant-stack filtering.

Our data processing goal is to develop a sequence that can be applied to field data in as close to an automatic manner as possible (fewer manual steps). This could include the development of generic signal and noise PEF's that would function on any data, and would not require the determination of these PEF's on windows of each data file. This would improve the processing of noisy data files (where such windows would be hard to define), and would speed the processing of data from larger-scale surveys. The results presented here suggest that this goal can be attained within a reasonable period of time, given a suitable data set for testing. The results of Haines et al. (2001) demonstrate that such a data set can be collected in the proper setting with the necessary equipment. Thus, depending on circumstances, the electroseismic method could soon begin to provide useful new subsurface information in geophysical exploration.

## ACKNOWLEDGMENTS

We are grateful to Simon Klemperer for general field guidance, to Jerry Harris for suggesting the waveform separation technique and other technical tid-bits, and to Steve Pride for the field

experiment design and all-around electroseismic instruction. Thanks to Jon Claerbout for suggesting the use of PEF's and other invaluable processing advice; and to all members of SEP, but particularly Brad Artman, Morgan Brown, and Bob Clapp for help getting up-to-speed on the SEP computer system, and for help with coding. Also, the data presented here would not exist without the effort of those who swung the hammer: Steve Pride, Nick Martin, Jonathan Franklin, Jordan Muller, T.J. Kiczenski and Stephan Bergbauer. Jim and Carolyn Pride graciously provided the field site, four dumptrucks full of sand, and some of their outstanding wine. Art Thompson provided some essential electronics, along with important input on data collection. Funding has been provided by the Stanford School of Earth Sciences McGee Fund, the AAPG student research grants, and by the GSA student research grants.

## REFERENCES

Brown, M., and Clapp, R., 2000, T-x domain, pattern-based ground-roll removal: 70th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 2103–2106.

Butler, K. E., and Russell, R. D., 1993, Subtraction of powerline harmonics from geophysical records (short note): Geophysics, **58**, no. 06, 898–903.

Claerbout, J. F., and Fomel, S., 2001, Geophysical Estimation by Example: http://sepwww.stanford.edu/sep/prof/gee/toc_html/index.html.

Claerbout, J., 1998, Multidimensional recursive filters via a helix: Geophysics, **63**, no. 05, 1532–1541.

Garambois, S., and Dietrichz, M., 2001, Seismoelectric wave conversions in porous media: Field measurements and transfer function analysis: Geophysics, **66**, no. 5, 1417–1430.

Guitton, A., Brown, M., Rickett, J., and Clapp, R., 2001, A pattern-based technique for ground-roll and multiple attenuation: SEP–**108**, 249–274.

Guitton, A., 2001, Coherent noise attenuation: A synthetic and field example: SEP–**108**, 225–248.

Haartsen, M. W., and Pride, S. R., 1997, Electroseismic waves from point sources in layered media: J. Geophys. Res., **102**, no. B11, 24745–24769.

Haines, S., Pride, S., and Klemperer, S., 2001, Development of experimental methods in electroseismic research, with application to aquifer characterization: Eos Trans. AGU, **82**, no. 47, Abstract GP22A–0268.

Pride, S., 1994, Governing equations for the coupled electromagnetics and acoustics of porous media: Physical Review B, **50**, no. 21, 15678–15696.

Shaw, D., 1992, Introduction to colloid and surface chemistry: Butterworths, 4th edition.

Thompson, A. H., and Gist, G. A., 1993, Geophysical applications of electrokinetic conversion: The Leading Edge, **12**, no. 12, 1169–1173.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 199–204*

# Short Note

# Theoretical aspects of noise attenuation

*Antoine Guitton*[1]

## INTRODUCTION

In Guitton (2001) I presented an efficient algorithm that attenuates coherent noise based on the spatial predictability of noise and signal. I called this algorithm the subtraction method. In this paper I show that the subtraction approach is closely related to another described method, the filtering method (Brown and Clapp, 2000; Clapp and Brown, 2000; Spitz, 1999; Soubaras, 1994) if I use a preconditioning strategy (Claerbout and Fomel, 2001). In a second part I prove that the Spitz estimate for the signal PEF (Spitz, 1999) makes the inversion of the Hessian stable in the subtraction method.

## FROM THE FILTERING TO THE SUBTRACTION OF NOISE

The filtering method is essentially based on the signal-preserving properties of the so-called "projection filters" (Soubaras, 1994). This idea has been widely used to attenuate a large variety of noise in seismic data. Abma (1995) developed a solid mathematical background that introduces these projection filters and showed that they are related to the classical Wiener estimator (Castleman, 1996). In this section I unravel the link between the filtering and subtraction method.

### Definitions

First I introduce a set of important variables that will help us build the desired filters.

- **d**: the data vector; input to the problem.
- **n**: the noise vector; assumed to be known.
- **s**: the signal vector; output of the problem.
- **D**: annihilation filter for the data; a Prediction Error Filter (PEF).

[1]**email:** antoine@sep.stanford.edu

- **N**: annihilation filter for the noise; a PEF.
- **S**: annihilation filter for the signal; a PEF.

The leading assumption is that the data vector is the sum of the signal and noise vectors, i.e,

$$\mathbf{d} = \mathbf{s} + \mathbf{n}. \tag{1}$$

**The filtering method**

Abma (1995) solved a constrained least-squares problem to separate signal from spatially uncorrelated noise:

$$\begin{aligned} \mathbf{Nn} &\approx \mathbf{0} \\ \epsilon\mathbf{Ss} &\approx \mathbf{0} \\ \text{subject to} \leftrightarrow \mathbf{d} &= \mathbf{s} + \mathbf{n} \end{aligned} \tag{2}$$

The first equation defines mathematically the annihilation filter **N** whereas the second equation defines the annihilation filter **S**. Minimizing in a least-squares sense the fitting goals in equation (2) with respect to **s** leads to the following expression for the estimated signal:

$$\widehat{\mathbf{s}} = \left(\mathbf{N^T N} + \epsilon^2 \mathbf{S^T S}\right)^{-1} \mathbf{N^T N d}. \tag{3}$$

$\left(\mathbf{N}^T\mathbf{N} + \epsilon^2\mathbf{S}^T\mathbf{S}\right)^{-1}\mathbf{N}^T\mathbf{N}$ is a projection filter. I call it the filtering method because the noise components are filtered out by the PEF **N** in equation (3).

**Preconditioning the filtering method**

There is a simple trick that modifies the fitting goals in equation (2). We can pose the following preconditioning transformations:

$$\begin{aligned} \mathbf{n} &= \mathbf{N^{-1}} \mathbf{m_n}, \\ \mathbf{s} &= \mathbf{S^{-1}} \mathbf{m_s}, \end{aligned} \tag{4}$$

where $\mathbf{m_n}$ and $\mathbf{m_s}$ are new variables. Now we can derive a new system of fitting goals as follows:

$$\begin{aligned} \mathbf{m_n} &\approx \mathbf{0} \\ \epsilon\mathbf{m_s} &\approx \mathbf{0} \\ \text{subject to} \leftrightarrow \mathbf{d} &= \mathbf{S^{-1}} \mathbf{m_s} + \mathbf{N^{-1}} \mathbf{m_n}. \end{aligned} \tag{5}$$

This system is almost equivalent to what I introduced in Guitton (2001), except for the regularization that I omitted. With $\mathbf{L_n} = \mathbf{N^{-1}}$ the noise-modeling operator and $\mathbf{L_s} = \mathbf{S^{-1}}$ the signal-modeling operator, the least-squares inverse of equations (5) without the regularization terms is then given by

$$\begin{pmatrix} \widehat{\mathbf{m}}_\mathbf{n} \\ \widehat{\mathbf{m}}_\mathbf{s} \end{pmatrix} = \begin{pmatrix} (\mathbf{L}_\mathbf{n}'\overline{\mathbf{R_s}}\mathbf{L_n})^{-1}\mathbf{L}_\mathbf{n}'\overline{\mathbf{R_s}} \\ (\mathbf{L}_\mathbf{s}'\overline{\mathbf{R_n}}\mathbf{L_s})^{-1}\mathbf{L}_\mathbf{s}'\overline{\mathbf{R_n}} \end{pmatrix} \mathbf{d}, \tag{6}$$

with

$$\begin{aligned} \overline{\mathbf{R_s}} &= \mathbf{I} - \mathbf{L_s}(\mathbf{L_s'L_s})^{-1}\mathbf{L_s'}, \\ \overline{\mathbf{R_n}} &= \mathbf{I} - \mathbf{L_n}(\mathbf{L_n'L_n})^{-1}\mathbf{L_n'}. \end{aligned} \tag{7}$$

I showed in Guitton et al. (2001) that $\overline{\mathbf{R_s}}$ and $\overline{\mathbf{R_n}}$ can also be interpreted in term of projection filters.

The estimated noise and signal are then computed as follows

$$\begin{aligned} \widehat{\mathbf{n}} &= \mathbf{L_n}\widehat{\mathbf{m}}_\mathbf{n}, \\ \widehat{\mathbf{s}} &= \mathbf{L_s}\widehat{\mathbf{m}}_\mathbf{s}. \end{aligned} \tag{8}$$

Because of the relationship that exists between the filtering and subtraction methods, the estimated noise or signal should be equivalent for both. This has been observed in a multiple attenuation problem by Guitton et al. (2001).

### Discussion

The preceding section derives the relationship between the filtering and subtraction methods. The preconditioning changes the nature of the problem quite deeply: from a filtering algorithm we end-up with a prediction/subtraction method. The least-squares inverses in equations (3) and (6) are also very different: from a problem with one unknown, $\mathbf{s}$, we end-up with a problem with two unknowns, $\mathbf{m_n}$ and $\mathbf{m_s}$. Fortunately, this preconditioning should speed-up the convergence toward the signal vector $\mathbf{s}$. In addition, this transformation separates the data space in its natural components, e.g, the signal and noise vectors more explicitely than with the filtering method.

## STABILITY OF THE HESSIAN WITH THE SPITZ ESTIMATE

Nemeth (1996) shows that the Hessians in equations (6) can be unstable if the signal and noise operators $\mathbf{L_s}$ and $\mathbf{L_n}$ predict similar components of the data space. In this section I show that the Spitz estimate of the signal PEF makes this overlap impossible.

### The Spitz estimate

Equation (5) assumes that the signal PEF is known in advance. This argument is circular since we are estimating the signal needed to compute $\mathbf{S}$! Nonetheless, Spitz (1999) shows that $\mathbf{S}$ can be estimated directly from the data PEF $\mathbf{D}$ and the noise PEF $\mathbf{N}$. Again, we assume that we are able to compute a model for the coherent noise we wish to attenuate. In equation, the signal PEF can be estimated as follows:

$$\mathbf{S} \approx \mathbf{D}/\mathbf{N}. \tag{9}$$

In the next section, I show on a simple 1D example that this estimate makes the overlap of the signal and noise operators $\mathbf{L_s}$ and $\mathbf{L_n}$ impossible

### A 1D example

Now I consider the Z-transforms of the data, signal and noise PEFs for a 1D case. For the data PEF **D**, I assume that the filter has the form

$$D(Z) = \alpha(Z - Z_1)(Z - Z_2)(Z - Z_3)(Z - Z_4), \tag{10}$$

with $\alpha = 1/(Z_1 Z_2 Z_3 Z_4)$. The $Z_i$ correspond to the roots of the filter. In this example I consider that $Z_1$ and $Z_2$ are the roots for the noise and $Z_3$ and $Z_4$ the roots for the signal. Now I assume that we have for the noise PEF **N**

$$N(Z) = \beta(Z - Z_1)(Z - Z_2), \tag{11}$$

with $\beta = 1/(Z_1 Z_2)$. The Spitz estimate yields for the signal PEF **S**

$$S(Z) = \gamma(Z - Z_3)(Z - Z_4), \tag{12}$$

with $\beta = 1/(Z_3 Z_4)$. We see that by construction, the signal and noise PEF annihilate different parts of the data space and can't overlap.

Now, If we assume that the noise PEF is a "bad" estimate of the noise with one erroneous root, i.e,

$$N(Z) = \beta(Z - Z_1)(Z - Z_5), \tag{13}$$

with $\beta = 1/(Z_1 Z_5)$, we find for the signal PEF **S**

$$S(Z) = \frac{Z_5}{Z_2 Z_3 Z_4} \frac{(Z - Z_2)(Z - Z_3)(Z - Z_4)}{Z - Z_5}. \tag{14}$$

Because the PEFs are minimum phase, we can write

$$\begin{aligned} \frac{1}{Z - Z_5} &= \frac{-1}{Z_5} \frac{1}{1 - Z/Z_5}, \\ &= \frac{-1}{Z_5} \left( 1 + \frac{Z}{Z_5} + \frac{Z^2}{Z_5^2} + ... \right), \\ &\approx \frac{-1}{Z_5} \left( 1 + \frac{Z}{Z_5} \right), \\ &\approx \frac{-1}{Z_5^2} (Z + Z_5). \end{aligned} \tag{15}$$

Then we obtain for the signal PEF

$$S(Z) \approx \frac{-1}{Z_2 Z_3 Z_4 Z_5} (Z - Z_2)(Z - Z_3)(Z - Z_4)(Z + Z_5). \tag{16}$$

The wrong root in the noise PEF leaks in the signal PEF but with an opposite sign. Again, the Spitz estimate makes it impossible for the signal and noise operators to overlap in the data space. This simple example in 1D can be easily expendable in 2D via the helical coordinates (Claerbout, 1998).

## CONCLUSION

I have shown that the filtering and subtraction method are linked by a simple preconditioning transformation. I have also demonstrated that the Spitz estimate for the signal PEF estimation prevent stability issues when the Hessians in equation (6) are computed.

## REFERENCES

Abma, R., 1995, Least-squares separation of signal and noise with multidimensional filters: Ph.D. thesis, Stanford University.

Brown, M., and Clapp, R., 2000, T-x domain, pattern-based ground-roll removal: 70th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 2103–2106.

Castleman, K. R., 1996, Digital image processing: Prentice-Hall.

Claerbout, J., and Fomel, S., 2001, Geophysical Estimation by Example: Class notes, http://sepwww.stanford.edu/sep/prof/index.html.

Claerbout, J., 1998, Multidimensional recursive filters via a helix: Geophysics, **63**, no. 05, 1532–1541.

Clapp, R. G., and Brown, M., 2000, $(t-x)$ domain, pattern-based multiple separation: SEP–**103**, 201–210.

Guitton, A., Brown, M., Rickett, J., and Clapp, R., 2001, A pattern-based technique for ground-roll and multiple attenuation: SEP–**108**, 249–274.

Guitton, A., 2001, Coherent noise attenuation: A synthetic and field example: SEP–**108**, 225–248.

Nemeth, T., 1996, Imaging and filtering by least-squares migration: Ph.D. thesis, The university of Utah.

Soubaras, R., 1994, Signal-preserving random noise attenuation by the F-X projection: 64th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1576–1579.

Spitz, S., 1999, Pattern recognition, spatial predictability, and subtraction of multiple events: The Leading Edge, **18**, no. 1, 55–58.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 205–212*

# Removing velocity stack artifacts

*Andrey Karpushin*[1]

**ABSTRACT**

The Prediction Error Filter (PEF) is a widely used tool in various geophysical applications such as signal-noise separation and interpolation of missing data. Over the years, SEP has developed tools to estimate non-stationary Prediction Error Filters. Non-stationary PEFs are successfully used for multiple removal, ground-roll attenuation, and in other geophysical problems. I apply a non-stationary PEF to a velocity stack to remove artifacts caused by a limited offset of the data. My first goal is to create an artifact-free model in which individual reflections are easier to identify. To do this I create a simple model of the artifacts in the $\tau - s$ space. This "noise" model is data-independent and relies only on the geometry of the data acquisition and parameters of the velocity stack. Then I estimate a non-stationary PEF on this "noise" model and use it to improve the velocity stack.
In the second part of the paper, I test the possibility of using the described PEF as a preconditioner for a velocity stack least-squares inversion.

## INTRODUCTION

To build a velocity stack, we usually construct an operator **H** that maps energy from velocity-stack space to offset-travel time space. The adjoint operator $\mathbf{H}'$ maps the energy back to a velocity-stack space. To build the model in a $\tau - s$ domain that is consistent with the data, we then solve the least-squares problem:

$$\mathbf{Hm} - \mathbf{d} \approx \mathbf{0} \tag{1}$$

This problem can be solved using iterative methods. For a simple model the solution is usually obtained in a few iterations. Although this solution may fit the data well, it may have "butterfly" artifacts caused by a limited aperture of the data. To illustrate this, I create a simple model with one spike in the velocity-stack space. Applying the operator **H** to this model, I obtain synthetic data set. Then I use this modeled data as an input (**d**) to solve the least-squares problem (equation 1) for **m**. Figure 1 shows the original model in the velocity-stack space, the modeled data, the estimated model, and data residual after five iterations of the conjugate gradient method.

As Figure 1 shows, the data residual after five iterations is small but "butterfly" artifacts are clearly present in the solution. It is desirable to obtain a model that has all the energy

[1]**email:** andrey@sep.stanford.edu

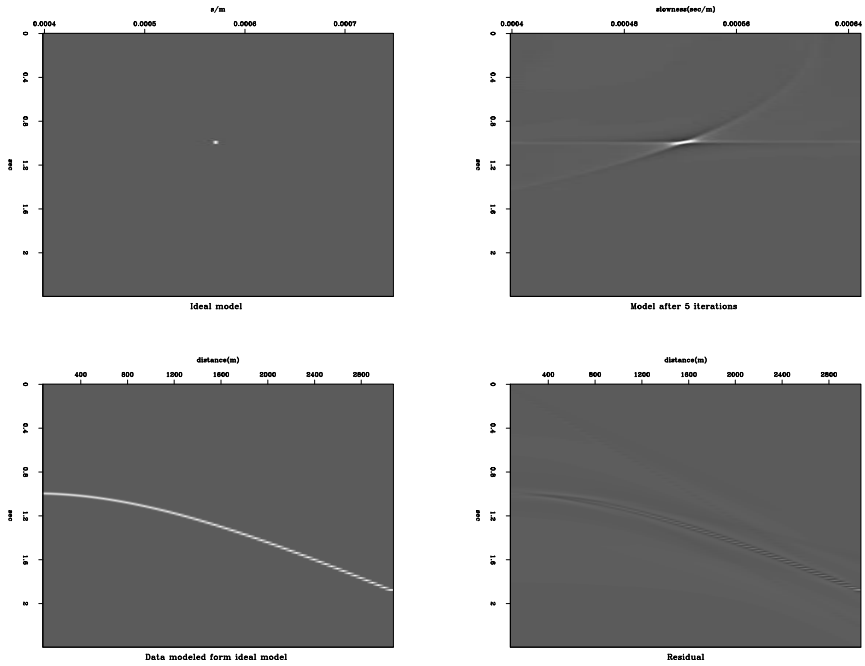



Figure 1: Top left: Ideal model. Top right: Spike after 5 iterations of the conjugate gradient. Bottom left: Data modeled from the ideal model. Botton right: Data residual after 5 iterations andrey-spike [ER]

concentrated at the location of the original spike and fits the data well. This problem has been addressed before. For example, at SEP Nichols (1994) and Guitton (2000) proposed to minimize the L1 norm of the model to create the spiky solution. Sacchi and Ulrych (1995) solved a similar problem using a parabolic Radon trasform, where they solve the problem in a frequency domain. These techniques showed very good results in concentrating the energy of the solution. (They are especially valuable during the multiple attenuation step of the processing.) But for some applications it may be useful to have an inexpensive way to remove artifacts from the model even if the data residual becomes larger. My first goal is to use the spatial predictability of the artifacts to design the operator to remove them. If the technique can make events appear better in a velocity stack panel and is easy to apply, it can be useful in picking velocities and designing masks for a multiple removal. A similar approach may be effective in other geophysical applications where the artifacts have a similar nature; for example, artifacts caused by a limited apperture in Kirchhoff migration.

## FILTER DESIGN

To design an operator to remove artifacts from a velocity stack, I am going to use the property of a two-column PEF to destroy a plane wave (Claerbout, 1999). Since artifacts from the far offset have a variable dip; this filter will need to be non-stationary. To design the operator we need to model the artifacts. The model of the artifacts can be created by applying the operator

$\mathbf{H}'$ to the first and the last trace of the data (Figure 2). We should be able to describe such a model with a three-column-wide filter. However, it is easier to divide the model of the artifacts into a "horizontal" and a "vertical" parts. We then create a pair of two-column PEF and train one on the "horizontal" and the other on the "vertical" part of the noise model. These two filters may then be applied to a velocity stack one after another to destroy the artifacts. It is very important to notice that it is not necessary to use traces from real data to model the "noise." Figure 2 illustrates how these separated parts of the "noise" model will look.

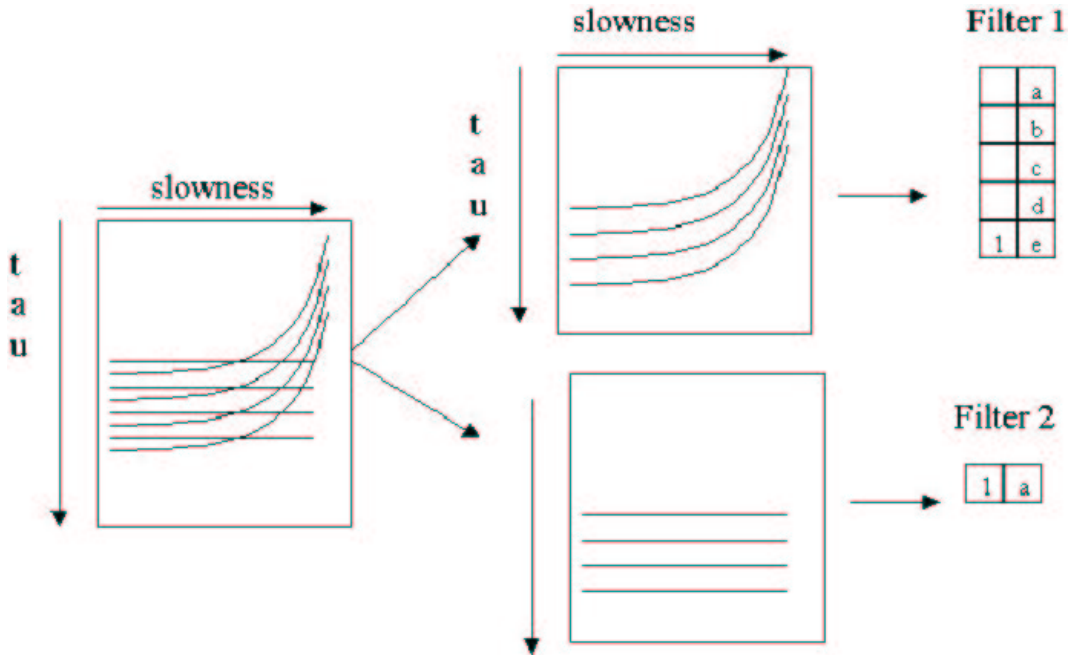



Figure 2: Illustration of different parts of a noise model andrey-noise [NR]

Using this approach, we could create the noise model and PEF only once and then re-use them when needed. This would greatly reduce the cost of the procedure and allow for the design of a stable filter. Designing a stable non-stationary PEF is the most difficult task in any application that uses a non-stationary PEF. However, in this case I have the advantage of being able to compute a non-stationaty PEF before the processing, so I can use various methods of smoothing the filter coefficients. I can also change the "noise" model, which might be important if I try to use this filter in any least-squares inversion schemes later.

## APPLICATION OF THE FILTER

In this section I test the approach described in the previous section on a simple synthetic and a real dataset. Figure 3 shows the result of applying this filter to a velocity stack of a synthetic data.

Figure 3 shows that the filter does remove the strongest artifacts and the result spatially looks more like the original spike.

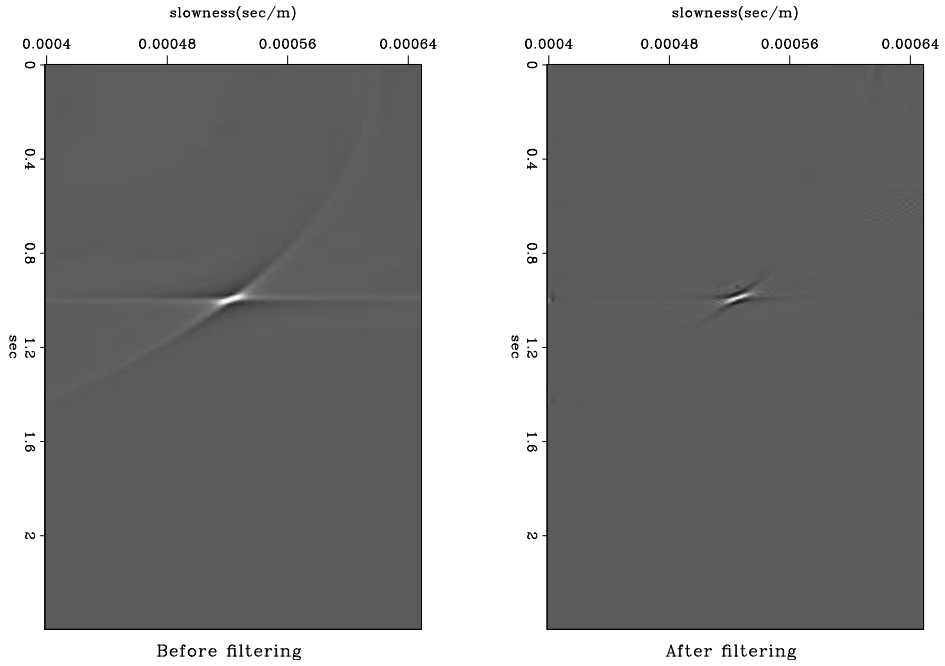



Figure 3: Spike after 5 iterations of conjugate gradient and after removing "butterfly wings." andrey-spike_comp [ER]

Figure 4 shows the result of filtering the velocity stack panel of one CMP gather of the Mobile AVO dataset. As we can see, strong individual events on the right panel are much easier to identify. Water-bottom multiples are clearly separated from other events and easy to identify.

Unfortunately, the filtering removes some of the energy from the model, limiting the application of the simple filtering described above. In the next section I discuss the possibilities and problems of incorporating the filtering approach as described above in a solution of the least-squares problem [equation (1)].

## LEAST-SQUARES INVERSION OF A VELOCITY STACK

To make the velocity stack usable for other processing steps such as multiple removal, I should try to incorporate the PEF described above into an inversion scheme that compensates for energy removed during the filtering.

Following the "trial solution" approach described in Claerbout (1999), I change the variables $\mathbf{m} = \mathbf{Pp}$ and solve equation (1) for the preconditioned variable $\mathbf{p}$:

$$\mathbf{HPp} - \mathbf{d} \approx \mathbf{0} \tag{2}$$

Operator $\mathbf{P}$ in equation (2) is a cascade of two filters described in the previous section. After finding a solution for $\mathbf{p}$, I evaluate $\mathbf{m} = \mathbf{Pp}$ to obtain the solution of the original problem.

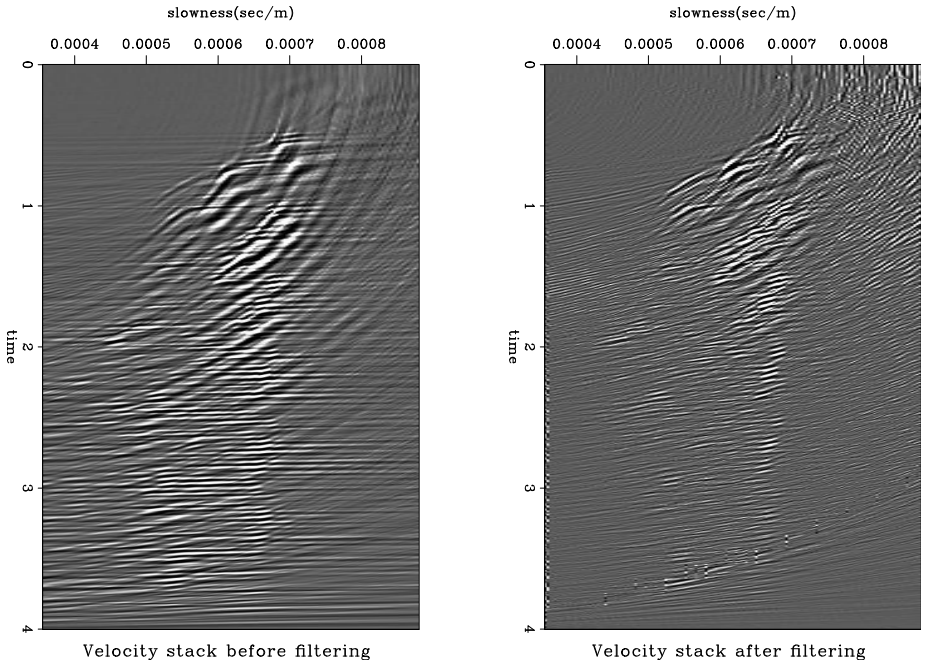



Figure 4: Velocity stack of a CMP from Mobile AVO dataset before and after filtering andrey-mob_comp_500 [ER]

To avoid high frequency noise in the model, I introduce a regularization term into the problem and solve the system of equations:

$$\begin{aligned} \mathbf{HPp} - \mathbf{d} &\approx \mathbf{0} \\ \epsilon \mathbf{Ap} &\approx \mathbf{0} \end{aligned} \tag{3}$$

In this paper I used the laplacian as the regularization operator **A** in equation (3).

Figure 5 shows the solution to the least-squares problem [equation (1)] after 25 iterations of the conjugate-gradient method, the velocity stack after the filtering, and the solution to the preconditioned least-squares problem [equation (3)] after 25 iterations of the conjugate-gradient.

Figure 6 shows the residual for the preconditioned problem [equation (3)] and the problem without preconditioning [equation (1)].

As Figures 5 and 6 show, although the solution for the preconditioned problem does not have artifacts, it converges much slower than the solution to the problem without preconditioning, which probably makes this method of preconditioning impractical.

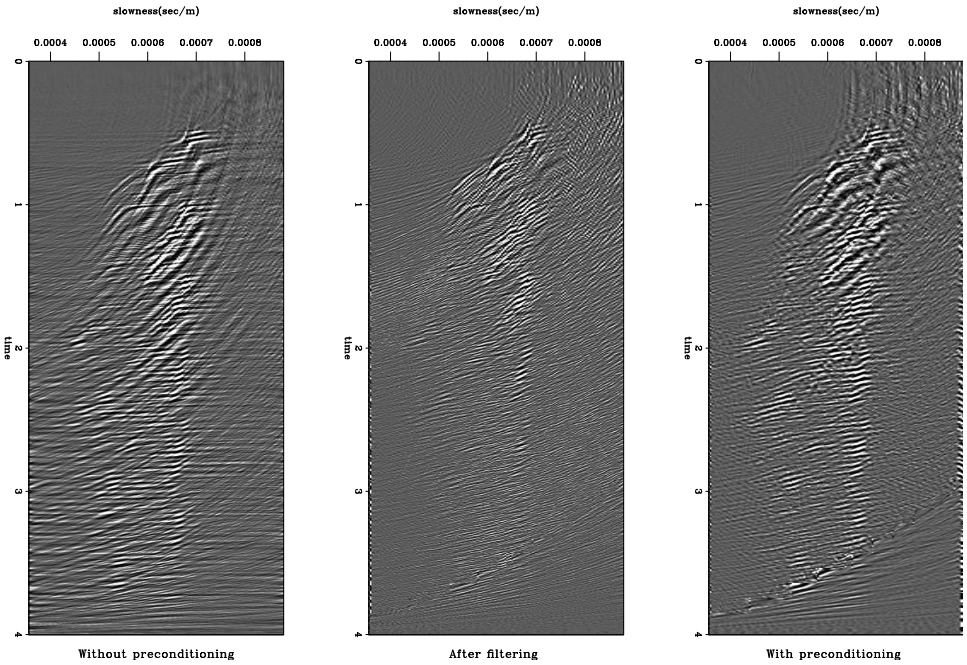



Figure 5: Left: Velocity stack after 25 iterations without preconditioning. Center: Velocity stack after filtering. Right: Velocity stack after 25 iterations with preconditioning. andrey-prec [ER]

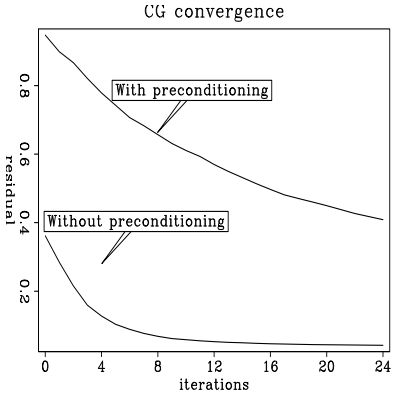



Figure 6: CG convergence with and without preconditioning andrey-res_ann [ER]

## CONCLUSIONS

I have presented a method to remove velocity stack artifacts caused by the limited offset of the acquisitions geometry. This method uses a non-stationary PEF and is computationally inexpensive and data independent. Tests on the data show that this approach successfully removes artifacts from a velocity stack. This method may be helpful in velocity analysis and other applications where artifacts are easy to model.

## REFERENCES

Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project, `http://sepwww.stanford.edu/sep/prof/`.

Guitton, A., 2000, Huber solver versus IRLS algorithm for quasi L1 inversion: SEP–**103**, 255–271.

Nichols, D., 1994, Velocity-stack inversion using $\mathbf{L_p}$ norms: SEP–**82**, 1–16.

Sacchi, M. D., and Ulrych, T. J., 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, no. 04, 1169–1177.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 213–223*

# Short Note

# Reference velocity selection by a generalized Lloyd method

*Robert G. Clapp*[1]

## INTRODUCTION

Wave equation migration methods are quickly becoming the standard for high-end imaging. These methods are more effective in handling complex wave behavior than traditional Kirchhoff methods (Workshop, 2001). The downside of these methods, specifically ones based upon the one-way wave equation, are that they are generally more expensive than Kirchoff methods and are only accurate for $v(z)$ media. To overcome the latter weakness extensions such as phase screen (Le Rousseau and de Hoop, 1998), split-step (Stoffa et al., 1990), and Fourier Finite Difference (FFD) (Ristow and Ruhl, 1994) have been introduced. These methods generally rely on downward continuing the wave-field with multiple reference velocities and then applying an approximate correction to the wavefield for the true velocity. The larger the deviation of the reference from the true velocity the less accurately the wavefield will be modeled. The trade-off is that the cost of migration increases almost proportionally with the number of reference velocities. Many different methods are currently employed to select the reference velocities, each with varying levels of success, dependent on the input velocity.

Electrical engineers and image processors face similar problems. In speech compression it is important to accurately describe a signal in as few bytes as possible. In image processing it is often important to reduce the number of colors in image with as little loss in image quality as possible. These problems have led to the field known as quantization. One family of method often employed is based on Lloyd's method (Lloyd, 1982), an iterative technique that allows for variable rate quantization.

In this paper I propose a new method for selecting reference velocities based on a generalized Lloyd's method. I show that the method is effective in choosing appropriate reference velocities for a large variety of models. On two synthetics I demonstrate that it produces a higher quality image with fewer reference velocities.

[1]**email:** bob@sep.stanford.edu

## THEORY

Selecting reference velocities can be thought of as a non-linear problem. Choosing a reference velocity to use at a given location is a function of which reference velocity is actually closest to the given velocity, which makes posing the problem in a traditional linear framework difficult. When compressing signals, electrical engineers face a similar problem. They try to describe a signal with as few bytes as possible. To do this they develop a codebook. Each member of the codebook represents a region described by a single value.

## LLOYD'S METHOD

Lloyd's method attempts to find the best codebook by a simple iterative procedure. Lloyd's method starts from a given initial codebook, then it follows a fairly simple algorithm:

**1** Find the nearest centroid for all points breaking up the points into cells.

**2** Recalculate the centroids for each cell.

**3** Remove any empty cells.

Steps **1**-**3** are repeated until the solution no longer changes significantly.

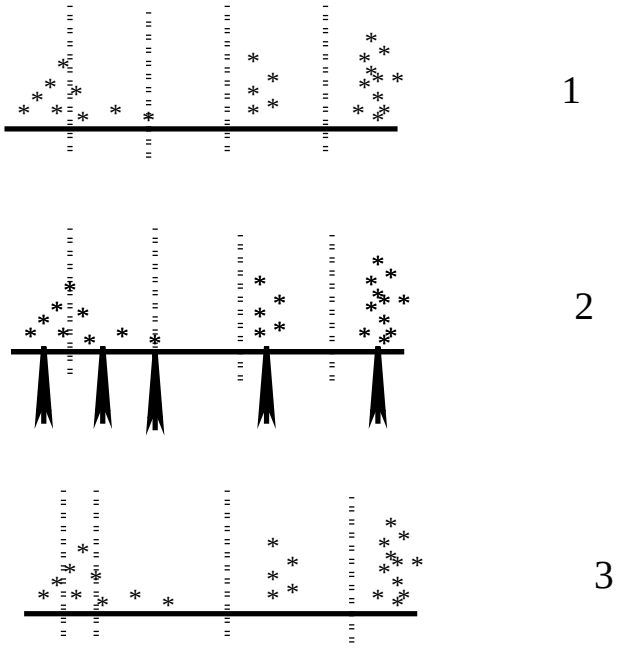



Figure 1: In step 1, the data is dividing into cells. In step2, the centroid of each cell is calculated. In the third step new cell boundaries are calculated between the centroids. If a cell contains no points, the cell is removed. bob4-ref [NR]

Lloyd's method can be effective but relies on a good initial codebook and can get stuck in local minima. To avoid prejudicing the solution the initial model is often a random set of vectors. To solve the second problem, several methods have been developed (Linde et al., 1980). These methods often rely on replacing empty cell by splitting regions with high variance.

### Selecting reference velocities

Lloyd's methods translate fairly easily for a reference velocity selection problem. The values become velocities, the centroids become the reference velocities, and the cells are regions with the same reference velocity. For selecting the initial reference velocities, I used evenly spaced quantiles of the velocity. At early iterations I replace step **3** of Lloyd's method by something a bit more complex.

**3a** Remove cells whose centroids are closer than a certain percentage from its neighbors (typically 3-8%).

**3b** Remove cells with fewer than a given number of points (typically 1-3%).

**3c** Given the removed points in steps 3a and 3b, find the regions with the highest variance that are a given distance away from its neighbors. If no regions meeting these criteria are found, remove them.

The purpose of modifying Lloyd's method is to attempt to avoid local minima. Regions whose centroid are close to its neighbor, or are sparsely populated, are replaced by splitting up regions that seem to be least accurately described by their current centroid.

## VELOCITY SELECTION TESTS

To test the methodology I applied it to several different velocity models. The first model is a 2-D synthetic, provided by BP, attempting to emulate North Sea geology (Figure 2). The model is composed of several constant velocity layers. For this test I used the entire velocity model as an input, even though when doing migration each depth step will by analyzed independently. Figure 3 shows a histogram of the velocity function overlayed by the initial reference velocities (0) and those after five (5) and twenty iterations (*). Note how all of the major velocities are identified by the method. For the second example I smooth the velocity in Figure 2 to obtain

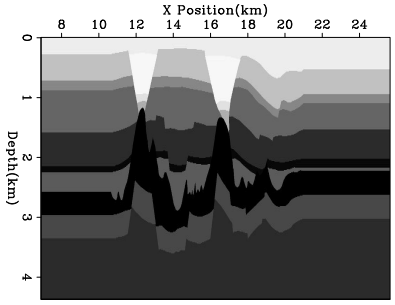



Figure 2: A 2-D synthetic velocity model made up of constant velocity layers. bob4-amoco-vel [ER]

the velocity in Figure 4. Figure 5 shows the selected velocities. Note how the initial velocities were clustered around $3.6 km/s$. The final velocities are generally more spread out, closer in velocity ranges with high count. Also note that in each region the velocity that is most common is chosen as a reference. To test the method's ability on functions with bimodal distributions

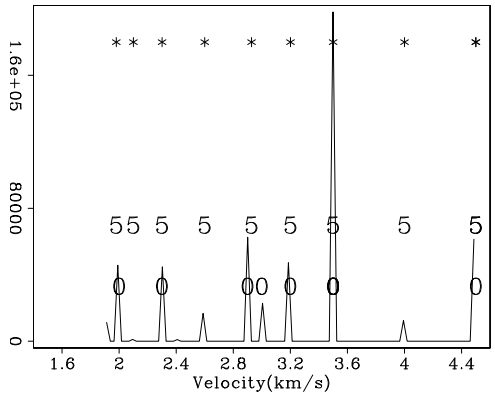



Figure 3: A histogram of the velocity function overlayed by the initial reference velocities (0) and those after five (5) and twenty iterations (*). bob4-amoco-hist-overlay [ER]

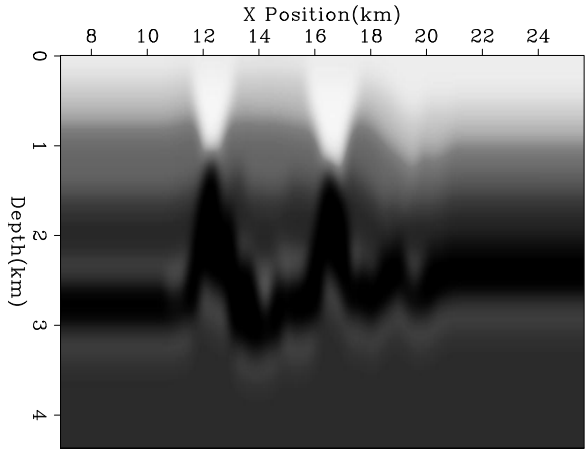



Figure 4: A smoothed version of the model in Figure 2. bob4-amsm-vel [ER]

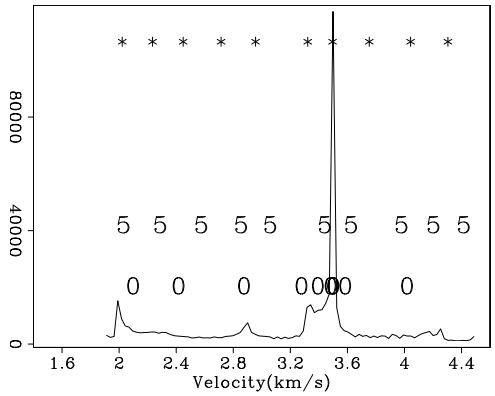



Figure 5: A histogram of the velocity function in Figure 4 overlayed by the initial reference velocities (0) and those after five (5) and twenty iterations (*). bob4-amsm-hist-overlay [ER]

(such as regions with salt), I used the upper (Figure 6) and lower (Figure 8) portion of the SMAART JV Sigsbee synthetic. Figures 7 and 9 show the selected velocities. Note how in each example the method determined that fewer reference velocities were needed. In each case a single reference velocity between the two modes of the distribution (this reference velocity was kept because some velocities still fall in this region).

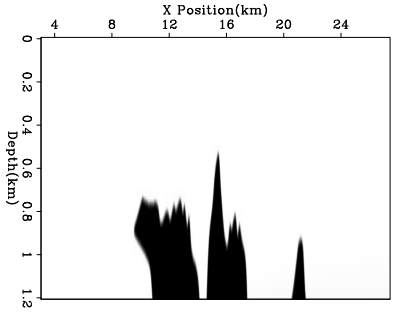



Figure 6: The top portion of the Sigsbee synthetic. bob4-zig1-vel [ER]

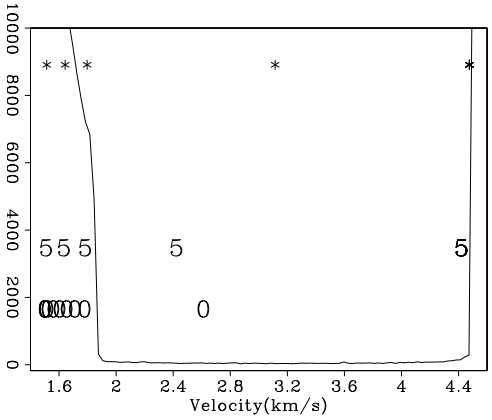



Figure 7: A histogram of the velocity function in Figure 6 overlayed by the initial reference velocities (0) and those after five (5) and twenty iterations (*). bob4-zig1-hist-overlay [ER]

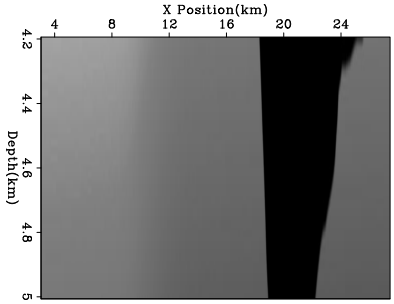



Figure 8: The bottom portion of the Sigsbee synthetic. bob4-zig2-vel [ER]

## MIGRATION TESTS

To test the migration result I chose the same two synthetics used in the previous section. Figure 10 shows the result of migrating the North Sea synthetic using at maximum four reference velocities. The top picture shows the result of using a conventional approach with four reference velocities plus salt. The bottom picture shows the result of using Lloyd's method with

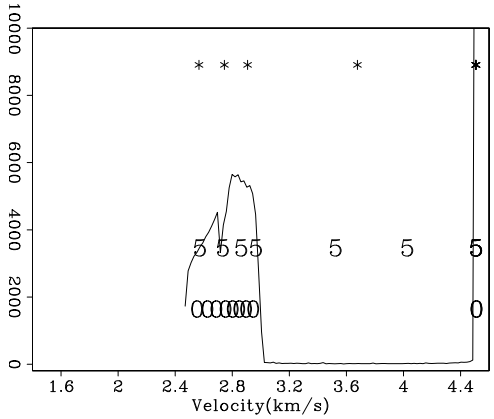



Figure 9: A histogram of the velocity function shown in Figure 8 overlayed by the initial reference velocities (0) and those after five (5) and twenty iterations (*). bob4-zig2-hist-overlay [ER]

a maximum of four velocities. The Lloyd result reduced the number of reference velocities needed at some depth levels, reducing the overall migration 30% faster and higher quality than the result with four reference velocities at each depth. At location ‘A’ the left edge of the salt structure is better placed. At locations ‘B’ and ‘C’ the bottom of the salt is better focused.

Figure 11 shows the correct (left) and selected (right) velocity for the Sigsbee synthetic. Note how the salt velocity was correctly selected and the rest of the model well represented. Figure 12 shows the result of migrating the Gulf of Mexico synthetic. In this case for the standard approach I chose five+salt reference velocities. The Lloyd’s algorithm selection method resulted in a 35% cost saving in compute time. The top picture shows the conventional approach while the bottom picture shows the result of using Lloyd’s method for velocity selection. The bottom of the salt reflection, ‘A’, is better focused. At ‘B’ the right canyon is more accurately positioned and better focused. At location ‘C’ we see how the point diffractors are collapsed using the Lloyd velocity selection method while using the more standard approach they are not collapsed.

## CONCLUSIONS

Reference velocity selection shares similar characteristics with speech compression. By using a modified version of Lloyd’s method to select velocities the velocity at a given depth can be more accurately described, often with fewer velocities, than conventional approaches. The resulting migration images can be computed faster, with higher image quality compared to conventional methods.

## ACKNOWLEDGMENTS

I would like to thank BP and SMAART JV for the synthetic datasets used in this paper.

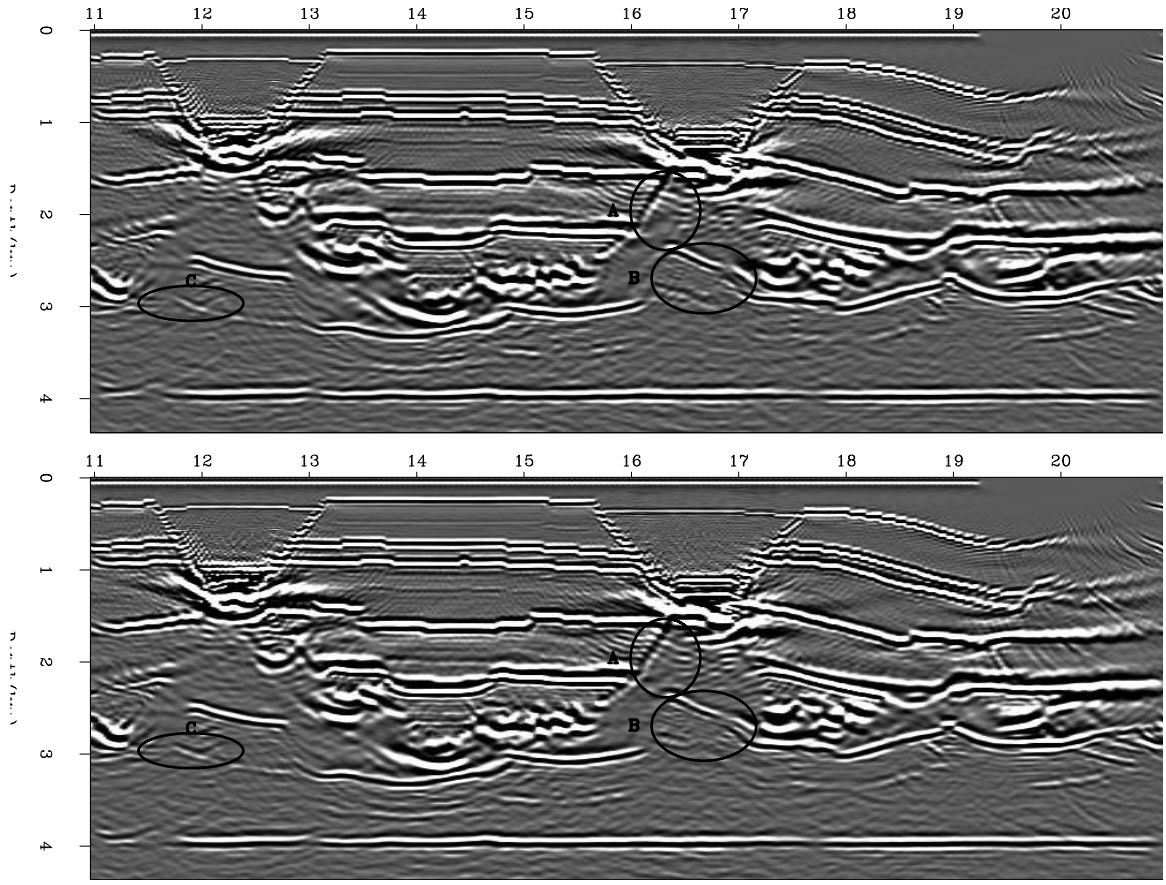



Figure 10: Comparison between regularly sampling the velocity range (top panel) and using Lloyd's algorithm for velocity selection.At location 'A' the left edge of the salt structure is better placed. At locations 'B' and 'C' the bottom of the salt is better focused. bob4-amoco-mig [ER,M]

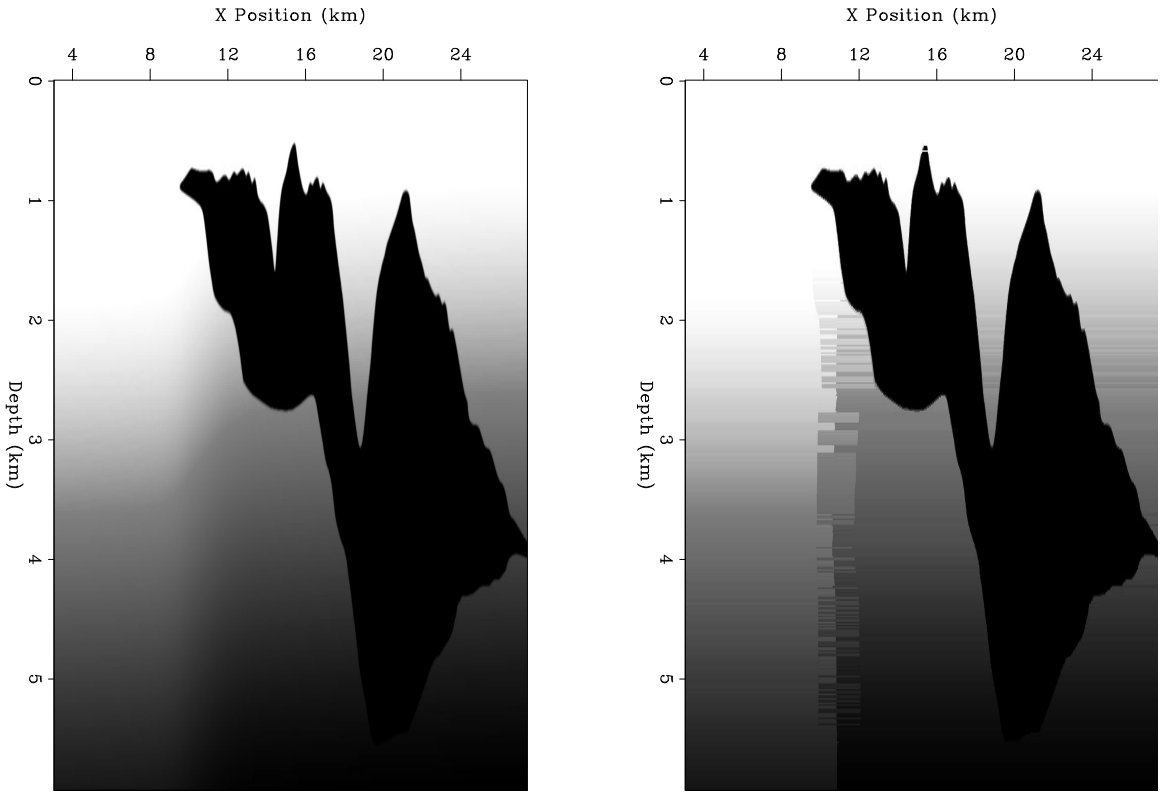



Figure 11: The correct (left) and selected (right) velocity for the sisbee synthetic. Note how the salt velocity was correctly selected and the rest of the model well represented. bob4-zig-chosen [ER,M]

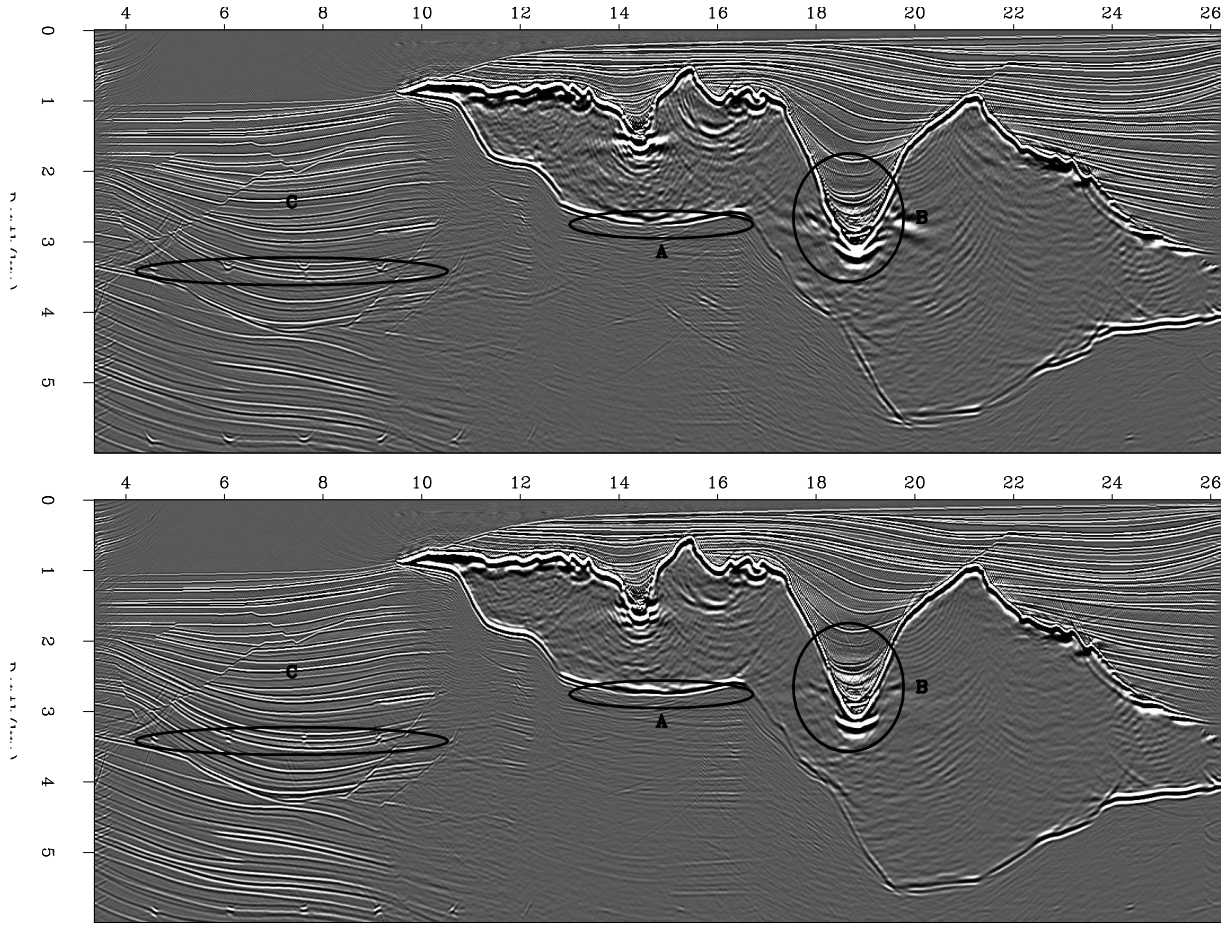



Figure 12: Comparison between regularly sampling the velocity range (top panel) and using Lloyd's algorithm for velocity selection. The bottom of the salt reflection, 'A', is better focused. At 'B' the right canyon is more accurately positioned and better focused. At location 'C' we see how the point difractors are collapsed using the Lloyd velocity selection method while using the more standard approach they are not collapsed. bob4-zig-mig [ER,M]

## REFERENCES

Le Rousseau, J., and de Hoop, M., 1998, Modeling and imaging with the generalized screen algorithm: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., 1937–1940.

Linde, Y., Buzo, A., and Gray, R. M., 1980, An algorithm for vector quantizer design: IEEE Transactions on ommunications, **1**, no. 1, 84–95.

Lloyd, S. P., 1982, Least squares quantization in pcm: IEEE Transactions on Information Theory, , no. 28, 127–135.

Ristow, D., and Ruhl, T., 1994, Fourier finite-difference migration: Geophysics, **59**, no. 12, 1882–1893.

Stoffa, P. L., Fokkema, J. T., de Luna Freire, R. M., and Kessinger, W. P., 1990, Split-step Fourier migration: Geophysics, **55**, no. 4, 410–421.

Workshop, S., 2001, Seismic imaging beyond kirchhoff: Society of Exploration Geophysicists.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 223–231*

# One-way wave equation absorbing boundary condition

*Guojian Shan*[1]

**ABSTRACT**

In modeling and migration based on wave equation, the wavefield has to be extrapolated in a finite domain due to the limitations of our survey and computers. Absorbing boundary conditions must be introduced, otherwise some reflections will occur at the artificial grid boundary. In this paper, I will introduce an absorbing boundary condition based on the one-way wave equation, with some numerical examples.

## INTRODUCTION

Several absorbing boundary conditions have been suggested to reduce the reflections at the artificial grid boundary (Engquist and Majda, 1977; Bayliss et al., 1982; Berenger, 1994). One kind of absorbing boundary condition is based on the one-way wave equation, and others are based on absorbing layers. In this paper, I introduce a high order one-way wave equation absorbing boundary condition, which can be solved using low order partial differential equations.

To simulate the wavefield in an open domain, absorbing boundary condition will be transparent to outgoing waves and be an obstacle to incoming waves. So, for a rectangular domain, the wavefield at the grid boundary satisfies the one-way wave equation. For example, the wavefield at the right boundary satisfies the leftgoing wave equation, and the wavefield at the left boundary satisfies the rightgoing wave equation. Solving the internal equation, which is a full wave equation in modeling and a one-way wave equation in migration, and using absorbing boundary conditions, we can simulate the wavefield in an open domain.

## ONE-WAY WAVE EQUATION

For the one-way wave equation

$$\frac{\partial p}{\partial z} = \pm \frac{i\omega}{c} \sqrt{1 + \frac{c^2}{\omega^2} \frac{\partial^2}{\partial x^2}}\, p,$$

[1]**email:** shan@sep.stanford.edu

we can write its $(2n+1)$th order approximation (Zhang, 1985) in time domain

$$\left(\frac{\partial}{\partial z} \pm \frac{1}{c}\frac{\partial}{\partial t}\right) p = \pm\frac{1}{c}\frac{\partial}{\partial t}\sum_{k=1}^{n} a_k q(s_k;t,x,z), \tag{1}$$

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - s_k^2\frac{\partial^2}{\partial x^2}\right) q(s_k,t,x,z) = \frac{\partial^2}{\partial x^2} p(t,x,z), \tag{2}$$

where $q$ is the auxiliary wavefield, $c$ is the velocity, and

$$s_k = \cos(\frac{k\pi}{n+1}), \quad a_k = \frac{1}{n+1}\sin^2(\frac{k\pi}{n+1}), \quad k = 0,1,\cdots,n+1.$$

When $n = 0$, we obtain the $5^o$ one-way equation

$$\left(\frac{\partial}{\partial z} \pm \frac{1}{c}\frac{\partial}{\partial t}\right) p = 0. \tag{3}$$

When $n = 1$, we obtain the $15^o$ one-way equation in Claerbout (1999)

$$\left(\frac{\partial}{\partial z} \mp \frac{1}{c}\frac{\partial}{\partial t}\right) p = \pm\frac{1}{c}\frac{\partial q}{\partial t}, \tag{4}$$

$$-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} q = \frac{1}{2}\frac{\partial^2}{\partial x^2} p. \tag{5}$$

When $n = 2$, we obtain the $45^o$ one-way wave equation in Claerbout (1999)

$$\left(\frac{\partial}{\partial z} \mp \frac{1}{c}\frac{\partial}{\partial t}\right) p = \pm\frac{1}{c}\frac{\partial q}{\partial t}, \tag{6}$$

$$\left(\frac{1}{4}\frac{\partial^2}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) q = \frac{1}{2}\frac{\partial^2}{\partial x^2} p. \tag{7}$$

## ABSORBING BOUNDARY CONDITION FOR MODELING

Let us consider the explicit finite-difference scheme for the full wave equation

$$\frac{\partial^2 p}{\partial t^2} = c^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right) p + f, \tag{8}$$

where $p$ is the wavefield and $f$ is the force. We can extrapolate the wavefield along $t$ using the following explicit finite-difference scheme

$$\Delta_t^2 p_{x,z}^t = c^2(\Delta t)^2\left(\frac{\Delta_z^2}{(\Delta z)^2} + \frac{\Delta_x^2}{(\Delta x)^2}\right) p_{x,z}^t + (\Delta t)^2 f_{x,z}^t, \tag{9}$$

where $\Delta_t^2$ and $\Delta_x^2$ are the second order central finite-difference operators

$$\Delta_t^2 p_{x,z}^t = p_{x,z}^{t+1} + p_{x,z}^{t-1} - 2p_{x,z}^t,$$

$$\Delta_x^2 p_{x,z}^t = p_{x+1,z}^t + p_{x-1,z}^t - 2p_{x,z}^t.$$

Given the initial condition $p_{x,z}^{t-1}$ and $p_{x,z}^t$, we can solve equation (9) to get the wavefield at time $t+1$, $p_{x,z}^{t+1}$ from the wavefield at time $t-1$, $p_{x,z}^{t-1}$ and the wavefield at time $t, p_{x,z}^t$, except for the wavefield at the boundaries $p_{x=X_{min},z}^{t+1}, p_{x=X_{max},z}^{t+1}, p_{x,z=Z_{min}}^{t+1}, p_{x,z=Z_{max}}^{t+1}$.

Let us consider the wavefield on the boundary $z = Z_{max}$. There are only outgoing waves at $z = Z_{max}$, so the wavefield satisfies the downgoing wave equation, for which we can write its approximate equations:

$$\frac{\partial p}{\partial z} = -\frac{1}{c}\frac{\partial p}{\partial t} - \frac{1}{c}\frac{\partial q}{\partial t}, \tag{10}$$

$$\left(\beta\frac{\partial^2}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) q = \alpha\frac{\partial^2}{\partial x^2} p. \tag{11}$$

For compatibility with the explicit finite-difference scheme at internal points, we apply the explicit finite-difference scheme for the boundaries using equation (10) and (11) and get

$$\frac{1}{2}\Delta_z^-\left(p_{x,Z_{max}}^{t+1} + p_{x,Z_{max}}^t\right) = -\frac{\Delta z}{2c\Delta t}\Delta_t^+\left(p_{x,Z_{max}}^t + p_{x,Z_{max}-1}^t\right) - \frac{\Delta z}{c\Delta t}\Delta_t^+ q_{x,Z_{max}-\frac{1}{2}}^t, \tag{12}$$

$$\Delta_t^2 q_{x,Z_{max}-\frac{1}{2}}^t = \beta\frac{c^2(\Delta t)^2}{(\Delta x)^2}\Delta_x^2 q_{x,Z_{max}-\frac{1}{2}}^t - \frac{1}{2}\alpha\frac{c^2(\Delta t)^2}{(\Delta x)^2}\Delta_x^2(p_{x,Z_{max}}^t + p_{x,Z_{max}-1}^t). \tag{13}$$

where $\Delta^-$ is the first order backward finite-difference operator, $\Delta^+$ is the first order forward finite-difference opertor:

$$\Delta_z^- p_{x,z}^t = p_{x,z}^t - p_{x,z-1}^t, \quad \Delta_t^+ p_{x,z}^t = p_{x,z}^{t+1} - p_{x,z}^t,$$

and $\Delta^2$ is the second order central finite-difference operator:

$$\Delta_t^2 q_{x,z}^t = q_{x,z}^{t+1} - 2q_{x,z}^t + q_{x,z}^{t-1}, \quad \Delta_x^2 q_{x,z}^t = q_{x+1,z}^t - 2q_{x,z}^t + q_{x-1,z}^t.$$

Assuming that the wavefield $p_{x,z}^k$ for $k \le t$ is known, then we solve the internal equation (9) to get the wavefield for the internal points $X_{min} < x < X_{max}$, $Z_{min} < z < Z_{max}$ at time $t+1$, $p_{x,z}^{t+1}$ first. Then, the auxiliary wavefield $q_{x,Z_{max}-\frac{1}{2}}^{t+1}$ can be solved by equation (13) since the wavefield of the boundary at time $t$, $p_{x,Z_{max}}^t$ and $p_{x,Z_{max}-1}^t$ are known. Finally, we solve equation (12) to get the wavefield at the boundary $p_{x,z=Z_{max}}^t$. Figure 1 illustrates how the boundary conditions are solved.

The method of solving the wavefield at the other three boundaries $z = Z_{min}$, $x = X_{min}$, and $x = X_{max}$, is similar to that of boundary $z = Z_{max}$. The only difference is that the boundary condition equation is an upgoing wave equation at $z = Z_{min}$, leftgoing wave equation at $x = X_{min}$, and right-going wave equation at $x = X_{max}$.

According to Zhang and Wei (1998), this absorbing boundary condition is stable.

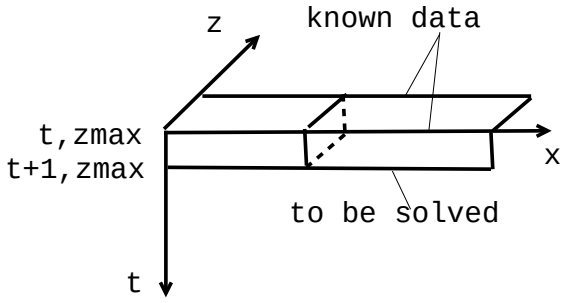



Figure 1: solution at the boundary $z = Z_{max}$ shan-boundary [NR]

## NUMERICAL EXAMPLE

I test the absorbing boundary condition on plane waves with different incident angles. I compare the results of a low order absorbing boundary condition ($5^o$ one-way wave equation, equivalent to the method in Engquist and Majda (1977) ) and a high order absorbing boundary condition ($45^o$ one-way wave equation). The results show little benefit in using a high order absorbing boundary condition for small incident angle plane waves. However, we get a much better absorbing result of the high order absorbing boundary condition than that of the low order absorbing boundary condition for large incident angle plane waves. For small incident angle plane waves (Figure 2), both low order and high order absorbing boundary condition equations do very well with the reflection. Theoretically, the low order absorbing boundary condition equation can only handle the reflection with angles less than $5^o$. So for reflection with a large angle (middle panels in Figure 3 and Figure 4), there is still a lot of reflected energy left after absorbing. High order absorbing boundary conditions (bottom panels in Figure 3 and Figure 4), still do well with the large angle reflection, and the results show that most reflection energy vanishes for both big and small incident angle plane waves.

## CONCLUSION

High order one-way wave equation absorbing boundary conditions handle well the reflections from the boundaries. The method described in this paper is usable for explicit finite-difference schemes. However, it can be used for both explicit and implicit finite-difference schemes, both for modeling and migration.

## REFERENCES

Bayliss, A., Gunzburger, M., and Turkel, E., 1982, Boundary conditions for the numerical solution of elliptic equations in exterior regions: SIAM J. Appl. Math., **42**, 430–451.

Berenger, J. P., 1994, A perfectly matched layer for the absorption of electromagnetic waves: J. of Comp. Physics, **114**, 185–200.

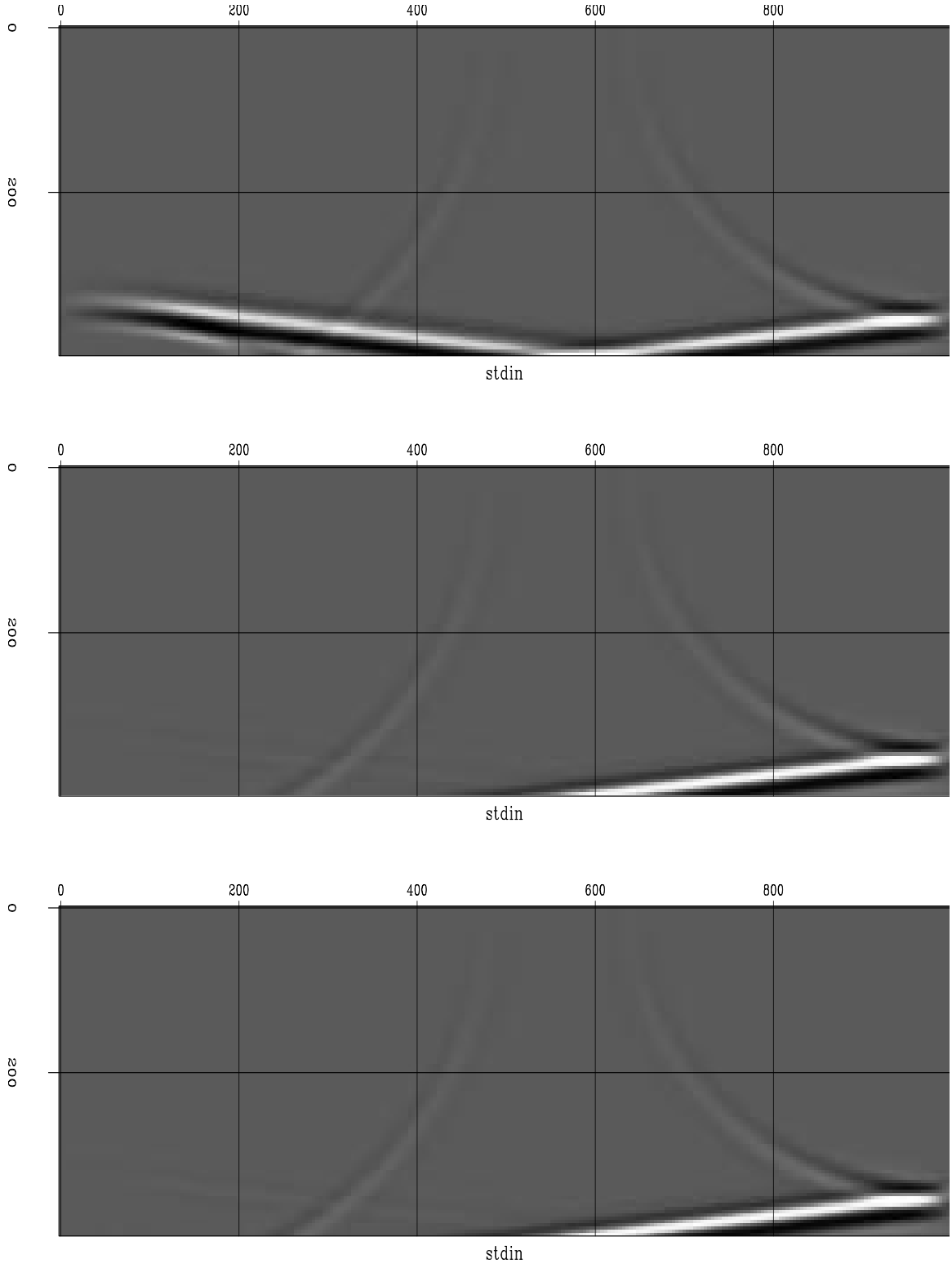



Figure 2: From top to bottom: modeling with no boundary condition; modeling with low order absorbing boundary condition; modeling with high order absorbing boundary condition shan-absorb4 [ER]

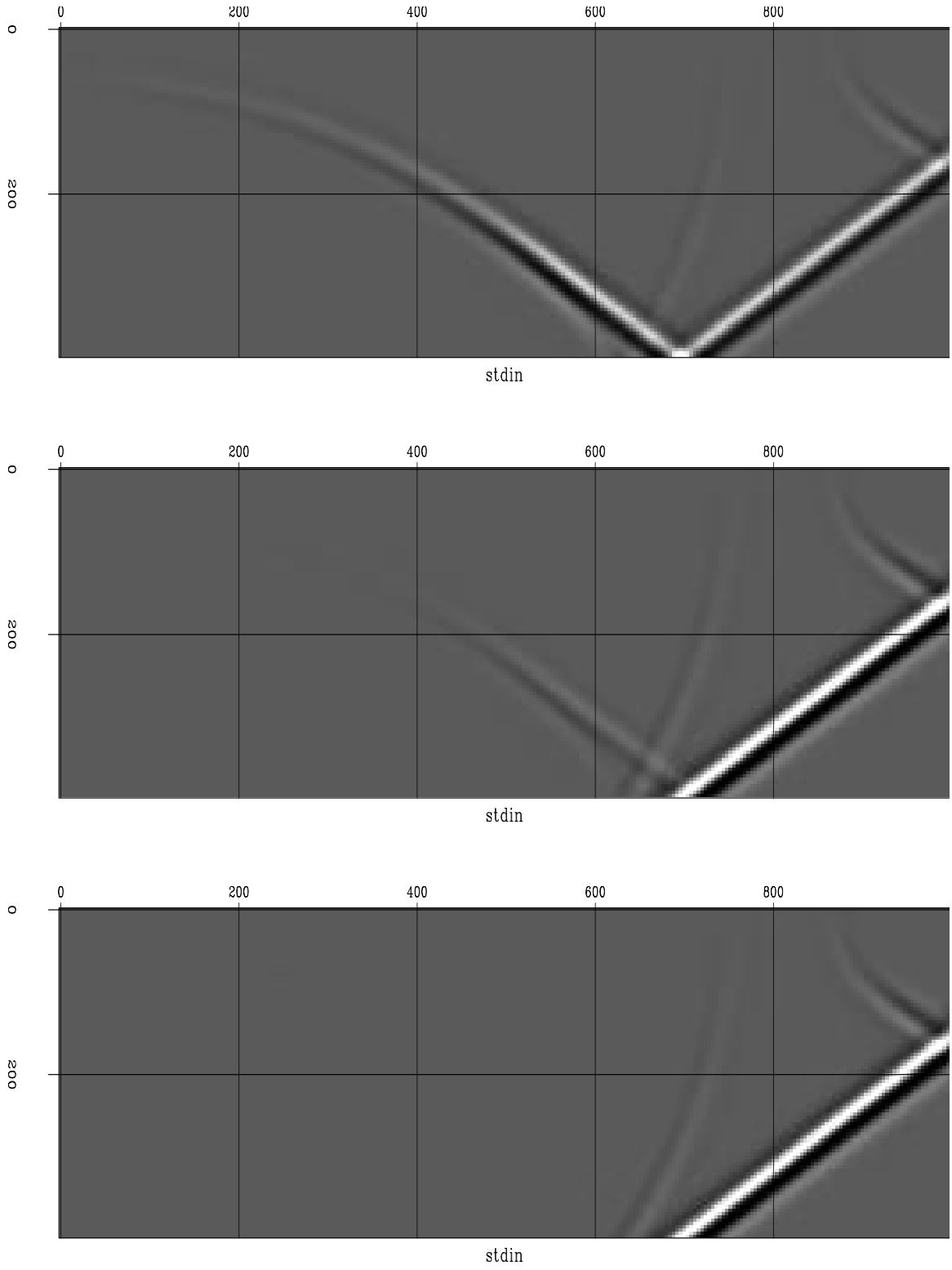



Figure 3: From top to bottom: modeling with no boundary condition; modeling with low order absorbing boundary condition; modeling with high order absorbing boundary condition shan-absorb2 [ER]

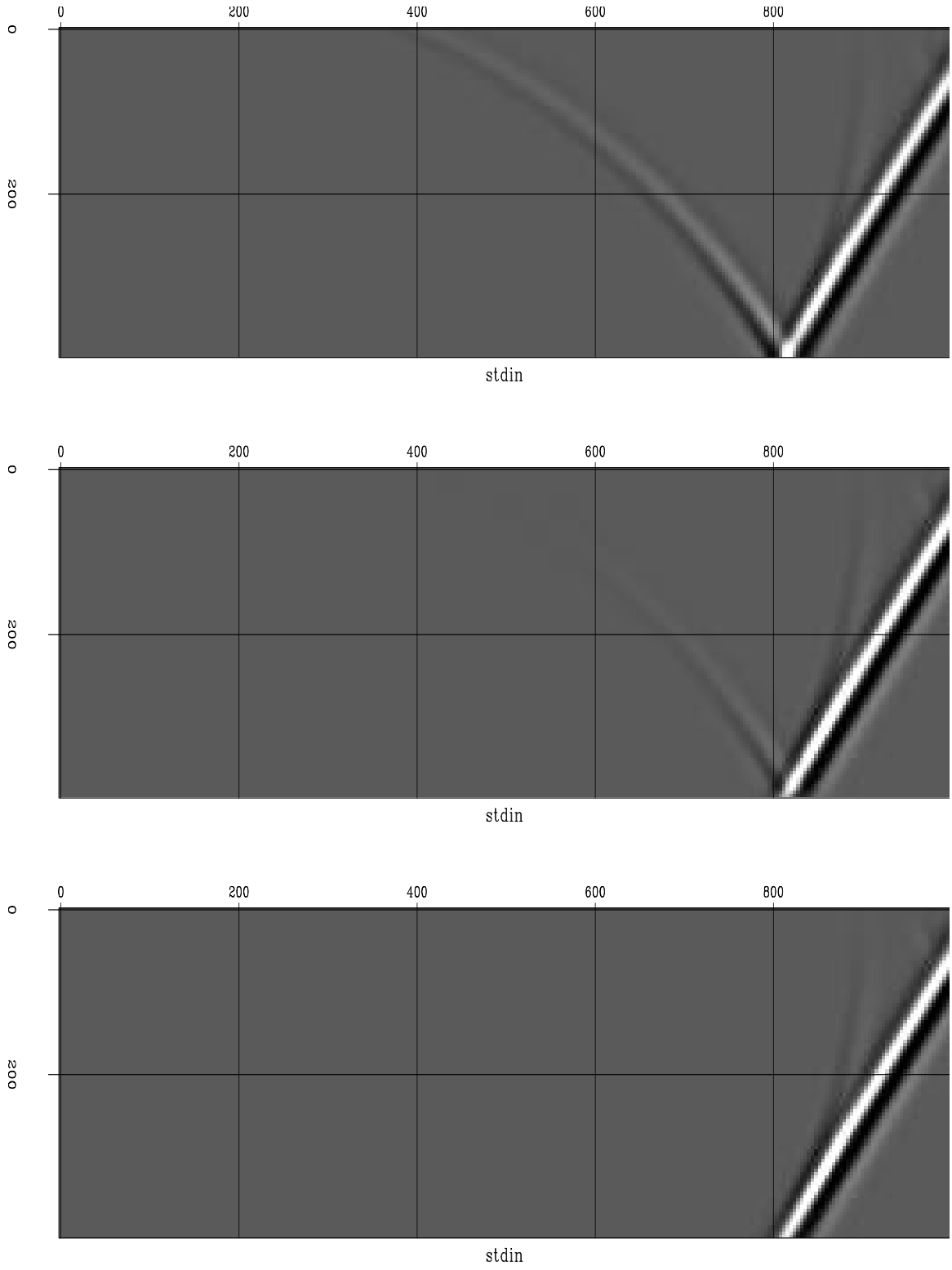



Figure 4: From top to bottom: modeling with no boundary condition; modeling with low order absorbing boundary condition; modeling with high order absorbing boundary condition shan-absorb1 [ER]

Claerbout, J. F., 1999, Basic earth imaging: Class notes, http://sepwww.stanford.edu/sep/prof/index.html.

Engquist, B., and Majda, A., 1977, Absorbing boundary conditions for the numerical simulation of waves: Math. Comp., **139**, 629–651.

Zhang, G., and Wei, S., 1998, Stability analysis of absorbing boundary condition for acoustic wave equation: Mathematica Numerica Sinica, **20**, 103–112.

Zhang, G., 1985, High order approximation of one-way wave equations: J. of Comp. Math., **3**, 90–97.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 231–238*

# Short Note

# Speeding up wave equation migration

*Robert G. Clapp*[1]

## INTRODUCTION

Wave equation migration is gaining prominence over Kirchhoff methods both as an imaging tool (Biondi and Palacharla, 1996) and for velocity analysis (Clapp, 2001). The relatively high cost of downward continuation methods, especially in iterative schemes (Biondi and Sava, 1999; Prucha et al., 2000), has limited their adoption.

In this paper I discuss three different methods to speed up midpoint-offset domain downward continuation based migration. I compare the migration results with the more standard downward continuation method. I show that a factor of two speed up is achievable with little discernible loss in image quality. In addition I show that for velocity analysis purposes a factor of three to four speed up is achievable.

## THEORY

When doing downward continuation in the offset domain, we begin by organizing our data cube as a function of midpoint $x$, offset $h$, and frequency $f$. We then apply the double square root (DSR) equation to move the wavefield down one depth step $\Delta z$ (Claerbout, 1995). We apply an imaging condition, and then repeat the procedure. This methodology can be quite expensive even in 2-D because the cost $C$ is approximately

$$\begin{aligned} C &\approx nz * f * (FFT(nx,nh) + nx * nh * CEXP) \\ C &\approx nz * f * (nh * nx log(nh) + nx * nh log(nx) + CEXP(nx,nh)), \end{aligned} \tag{1}$$

where $FFT(nx,nh)$ is the expense of doing a 2-D FFT on a $nx$ by $nh$ dataset and $CEXP$ is the cost of multiplying by a complex exponential. In 3-D the cost is even more substantial.

Equation (1) indicates that the number of depths can greatly affect the cost of the migration. As a result, the choice of depth sampling is a major decision. Too fine a depth sampling will make the cost exorbitant; too coarse will cause resolution and aliasing problems.

[1]**email:** bob@sep.stanford.edu

The required depth sampling is also depth varying. We need finer depth sampling near the surface, while coarser depth sampling is appropriate in the deeper section. So the first obvious way to speed up our migration is to vary the sampling as a function of depth.

### Sampling the wavelet

If we put in a time function, we can charactize the wavelength $\lambda$ as it travels through the earth with

$$\lambda = \frac{v}{f}, \tag{2}$$

where $f$ is the frequency and $v$ is the velocity. In practical terms, this says that our input is a time function that travels in a given interval as a function of the media's velocity, and that different frequencies will travel at different velocities. What is important is that we adequately sample the wavelet. As a result, we can take large depth steps at large velocities and small frequencies. Our migration cost is then generally concentrated at higher frequencies and shallower depths.

### Attenuation

Our second observation is that all signals eventually attenuate, and this attenuation ($Q$) is a function of frequency and media properties. How much they attenuate is generally a function of the media, but a decent first approximation is that after a certain number of cycles, a given frequency attenuates a given percentage(Kjartansson, 1979). Therefore, after a certain number of cycles, the wave components at that frequency will no longer be of usable strength so we can stop downward continuing it. This observation nicely compliments the wavelet sampling observation that higher frequencies will attenuate quicker in depth, reducing the distance we need to downward continue them. We can combine these two ideas and replace our original cost equation with,

$$C \approx \sum_f *nz(f) * (nh * nxlog(nh) + nx * nhlog(nx) + CEXP(nx,nh)), \tag{3}$$

where our number of depth steps $nz(f)$ is a function of frequency.

### Focusing energy

Our imaging condition provides a third idea for decreasing cost. As we go down in depth, energy focuses at zero offset. The outer offsets will either have no energy, or energy that we don't care about. As we step down in depth, we can decrease our offset domain. With this final savings, we can rewrite our cost function as:

$$C \approx \sum_f \sum_{nz(f)} nh(z) * nxlog(nh) + nx * nh(z)log(nx) + CEXP(nx,nh(z))). \tag{4}$$

## PRACTICAL ASPECTS

Applying the previous cost saving ideas introduces some new challenges. First, our final image is going to be regularly sampled so we must resample our variably sampled wavefield. I chose Lagrange interpolation for my resampling. It has the advantage of allowing higher order interpolation than simple linear interpolation, and you can easily pre-build interpolation tables, making it fast.

The second thing we have to be concerned with is the speed of our FFT. We generally use FFTW (Frigo and Johnson, 1999) which can handle any size axes but transforming an axis of length 512 will be significantly faster than transforming an axis of length 511. The easiest way to handle this problem it to make a list of 'good' axis lengths (e.g., combination of prime factors and power of two) and only decrease the offset domain at depths where the next lowest 'good' axis length is reached.

Finally, when this process is parallelized to be run on a linux cluster load, balancing becomes more difficult. Each frequency block has a different number of depth steps, a different interpolation cost, and a fixed IO cost. Ensuring that each processor finishes at approximately the same time is a non-linear optimization problem. Presently I assign costs for depth steps, interpolation, and IO and then try 1000 different random solutions, selecting the one that shows the least cost differential between the nodes.

## RESULTS

To test how much the migration could be sped up I chose a 2-D line from a 3-D land dataset provided by Ecopetrol. Figure 1 shows the result using the conventional approach. The left plot shows the zero offset image (the standard imaging condition), the right panel shows three selected angle CRP gathers (Sava and Fomel, 2000). Using five reference velocities and four processors it took 2034 seconds to run.

Figure 3 shows the same migrated image calculated with variable depth sampling, accounting for attenuation, and reducing the size of the offset domain as we go down in depth. The left panel of Figure 2 shows the sampling in depth for several frequencies. The right panel shows how the offsets downward continued decrease as the depth increased. The result was achieved in 1042 seconds (almost half of the time) and is nearly identical to the result in Figure 1.

For velocity analysis the image quality requirements are reduced. By using the sampling in Figure 4 image gathers can be produced that give accurate moveout information (Figure 5) while further reducing the migration time to 758 seconds.

## FURTHER SPEEDUP

In theory it should also be possible to use sparser sampling in midpoint at low frequencies. Since this would only affect the relatively inexpensive portion of the downward continua-

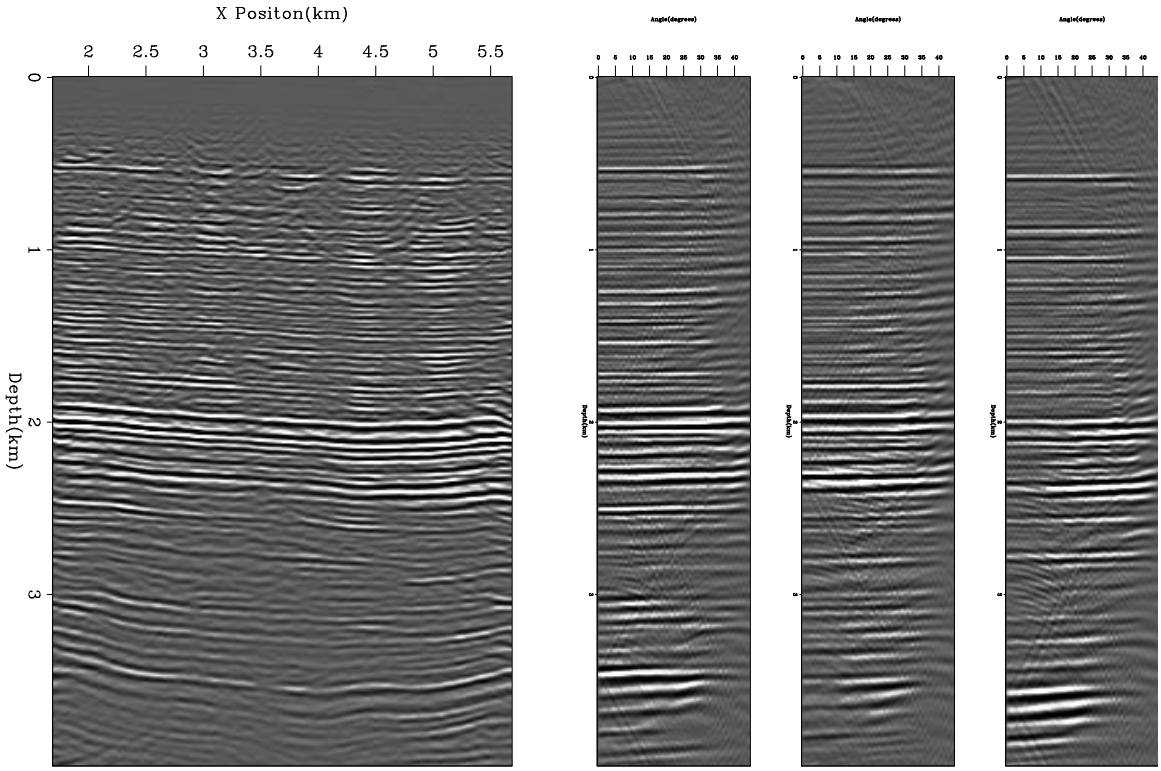



Figure 1: The left panel is the result of split-step migration. The right panel is three gathers from the same migration. bob3-mig-slow [CR]

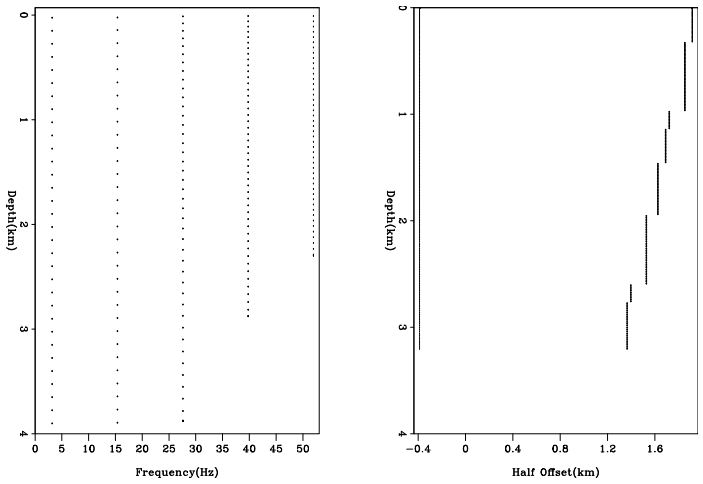



Figure 2: The left panel shows every tenth depth for various frequencies for the migration show in Figure 3. The right panel shows the reduction in the migrated offset domain as a function of depth. Note how the sampling in depth sparser, the offset domain is decreased quicker, and the frequencies are assumed to be of inconsequential energy than in Figure 3. bob3-fast-sample [CR]

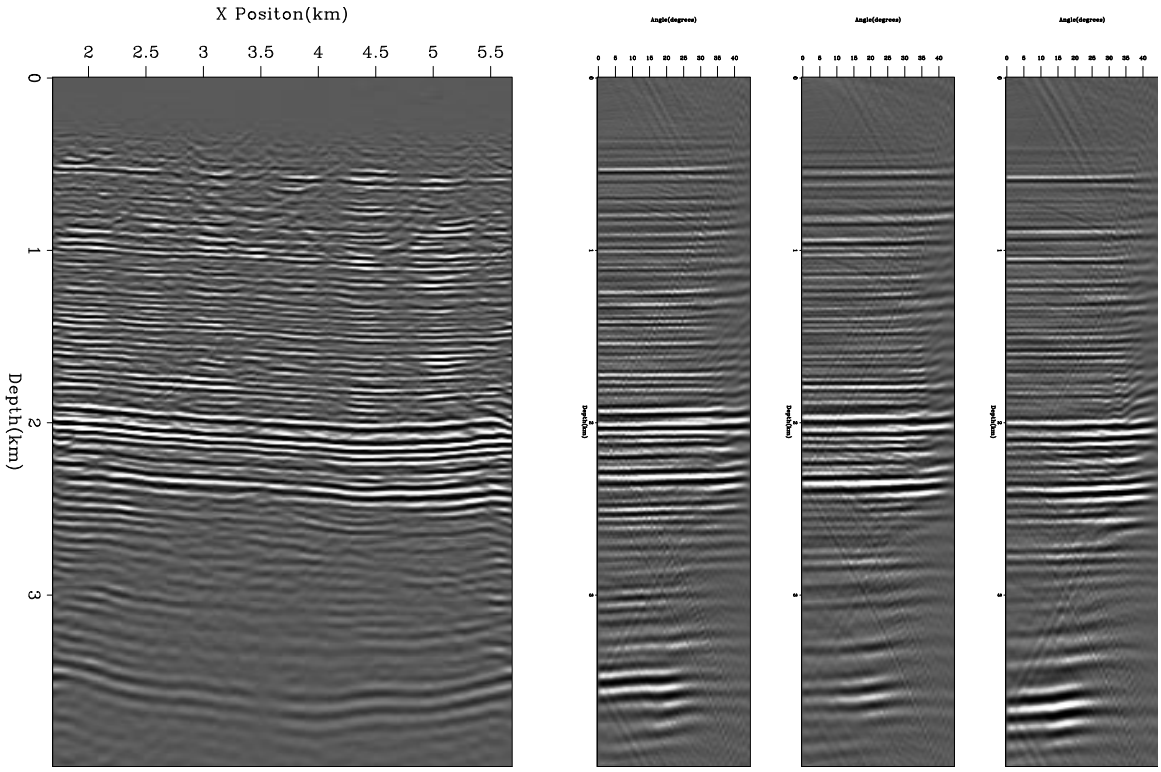



Figure 3: The left panel is the result of split-step migration. The right panel are three gathers from the same migration. Note how the image is almost identical to Figure 1 but is calculated three times faster. bob3-mig-fast [CR]

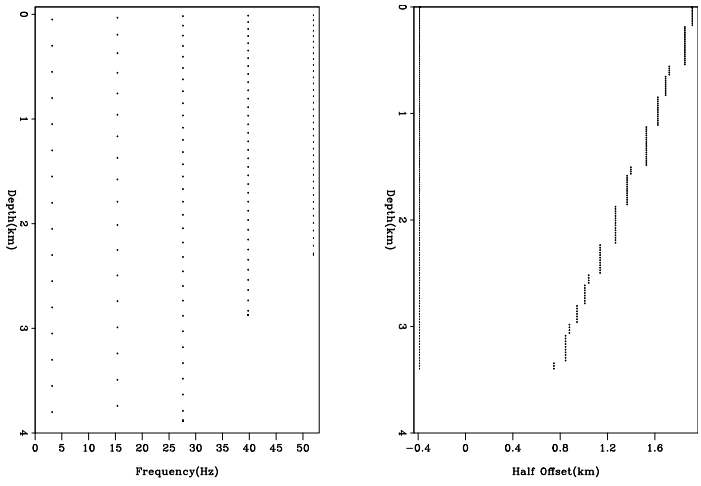



Figure 4: The left panel shows every tenth depth for various frequencies. The right panel shows the reduction in the migrated offset domain as a function of depth. Note how the sampling in depth sparser, the offset domain is decreased quicker, and the frequencies are assumed to be of inconsequential energy than in Figure 5. bob3-faster-sample [CR]

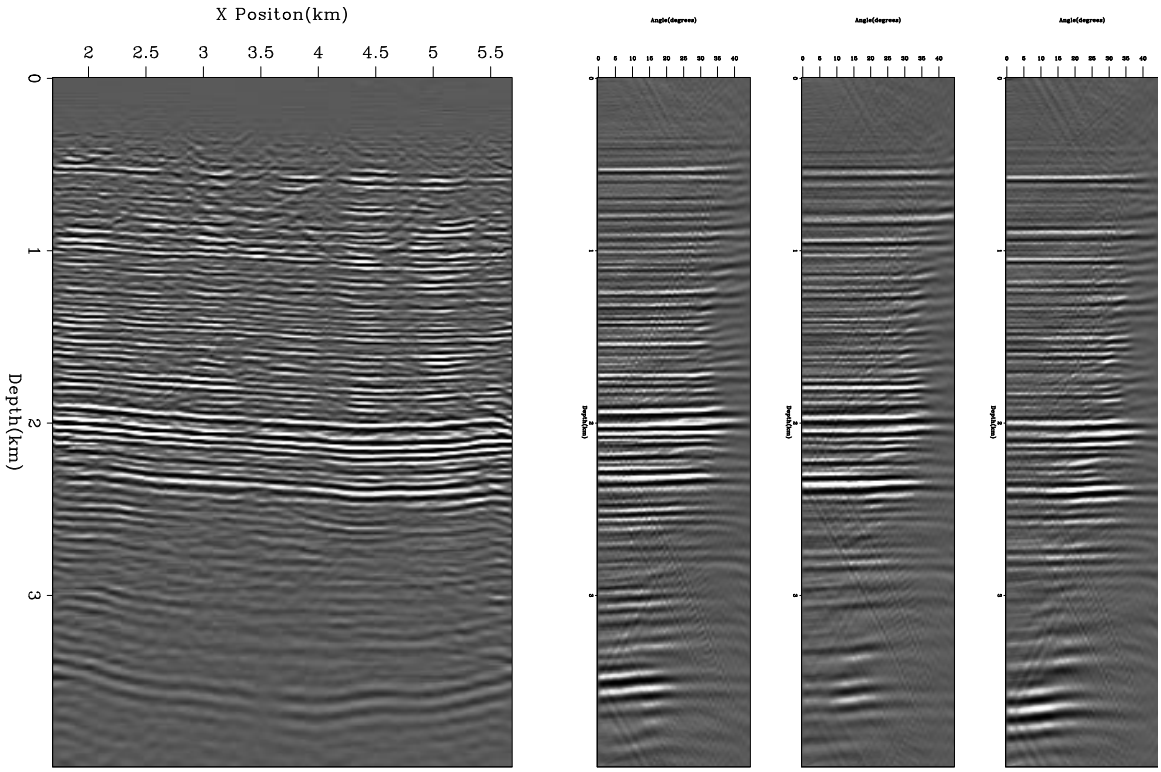



Figure 5: The left panel is the result of split-step migration. The right panel is three gathers from the same migration. Note how the image is slightly different from Figures 1 and 5, but the moveout information is nearly identical. bob3-mig-faster [CR]

tion, I would expect at most only a $10-20\%$ speed up. Another method is to do a better job choosing appropriate velocities, thereby reducing the number of velocities needed for accurate downward continuation. Clapp (2002) discussed one method to accomplish this.

## CONCLUSIONS

Wave equation migration can be significantly sped up by making depth steps a function of frequency and velocity while the offset domain can be reduced as we go down in depth. Wave equation migration can be sped up by a factor of three, and acceptable migration gathers for velocity analysis can be generated five times quicker than conventional migration methods.

## ACKNOWLEDGMENTS

I would like to thank Ecopetrol for the data used in this paper.

## REFERENCES

Biondi, B., and Palacharla, G., 1996, 3-d prestack migration of common-azimuth data: 3-d prestack migration of common-azimuth data:, Soc. of Expl. Geophys., Geophysics, 1822–1832.

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

Claerbout, J. F., 1995, Basic Earth Imaging: Stanford Exploration Project.

Clapp, R. G., 2001, Geologically constrained migration velocity analysis: Ph.D. thesis, Stanford University.

Clapp, R. G., 2002, Reference velocity selection by a generalized Lloyd method: SEP–**111**, 213–222.

Frigo, M., and Johnson, S. G., 1999, FFTW: `http://www.fftw.org/`.

Kjartansson, E., 1979, Attenuation of seismic waves in rocks and applications in energy exploration: Ph.D. thesis, Stanford University.

Prucha, M. L., Clapp, R. G., and Biondi, B., 2000, Seismic image regularization in the reflection angle domain: SEP–**103**, 109–119.

Sava, P., and Fomel, S., 2000, Angle-gathers by Fourier Transform: SEP–**103**, 119–130.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 239–253*

# Finite-difference $\omega$-x migration of unregularized seismic data

*Ioan Vlad*[1]

**ABSTRACT**

Wavefield downward continuation does not need to be done on a structured spatial mesh. Semistructured meshes have advantages (no need for regularization before migration) but they sometimes produce artifacts. There are at least two acquisition settings which will not result in artifacts. The implementation is simple. I present a 2D zero-offset $\omega$-x semistructured mesh migration (SMM) of a synthetic dataset. Its outputs are compared with those of $\omega$-x migration on a regular grid, with zero traces inserted in the place of missing data. There are several potential ways of removing the artifacts. Extension to 3D is also possible.

## INTRODUCTION

Common sense and basic physics suggest that in order to continue a field (gravitational, magnetic, wavefield...) into a direction, it is necessary to know the values of the field on one or more surfaces nonparallel to the continuation direction, and the law that governs the field (approximations can be made or field laws inferred so that we need only one surface). Because field quantities are usually invariants, nothing mandates that field continuations be done on Cartesian grids (although in many cases it is numerically convenient to do so). In particular, a look at the 45° downward continuation equation (Claerbout, 1999) shows that it can be written as:

$$\frac{2i\omega}{v}\frac{\partial Q}{\partial z}+\left(1-\frac{v}{2i\omega}\frac{\partial}{\partial z}\right)\Delta Q=0 \tag{1}$$

where $\Delta$Q, the Laplacian of Q, is an invariant. It can also be computed on an unstructured spatial mesh. In principle, this means that semistructured mesh migration (SMM) is feasible. I called the mesh semistructured because it is regular in time, but unstructured in space.

## SMM ADVANTAGES AND DISADVANTAGES

There are several pros and cons about doing prestack SMM. First, let us examine the consequences of the assumption that SMM is feasible and accurate. The need for regularization before or during migration would disappear. Regularization would be performed after migration when the imaging condition is applied and simplified, amounting to an image interpola-

[1]**email:** nick@sep.stanford.edu

tion. Only raw, uninterpolated data would be fed into the migration process, leading to a more accurate image than that made from a regularized dataset consisting entirely of interpolated data. The need to fill in with large amounts of zero traces would disappear, leading to computational savings. Also, using the original acquisition geometry for imaging can be a step towards constructing a probability map of the seismic image: the extra incertitude introduced by interpolation before downward continuation is eliminated, and therefore is not propagated through hundreds of depth steps. Such goals are certainly desirable.

The main problem stems from computing the Laplacian on a grid that is too sparse in places (close to spatial aliasing) or whose step size varies too quickly. Previous attempts (Dellinger and Muir, 1986) as well as the results of this work show that abrupt variations in the mesh size lead to numerical artifacts under the form of reflections off the irregularities in the grid. Such spurious reflections do not appear when the mesh step varies smoothly, or when the variation is less than half the grid step size. Therefore, two straightforward SMM applications are:

1. Migrating reflection data acquired from a platform moving with a nonconstant velocity (i.e., seismic acquisition ship which was not able to maintain constant speed; GPR vehicle that had to accelerate or decelerate; radar-bearing aircraft that encounters various air currents). In all these cases, the velocity variation is small enough that it should lead to a smoothly-varying grid.

2. Migrating reflection data acquired on a *heated atomic lattice* – a regular grid whose nodes have been displaced with small (known) amounts from the periodic positions.

In neither of these cases was the data irregular enough to justify the cost of full-fledged regularization, which involves interpolating the whole dataset. Most often the fact that the mesh is not really regular is simply overlooked. Applying SMM to such datasets will surely increase the quality of the image. The implementation is simple, consisting (at least for the 2D zero-offset case that has been implemented, and for its prestack extension) of simply replacing the (1,-2,1) coefficients of the second derivative in the differencing star with sets of three precomputed geometry-dependent values.

## COMPUTING DERIVATIVES ON A SEMISTRUCTURED MESH

Let $Q(\omega, x, z)$ denote a 2D zero-offset seismic wavefield, regularly sampled (and Fourier transformed) in time, but irregularly sampled in space. The measured surface dataset is $Q(\omega, x, z = 0)$. For downward continuing Q with (1) we need to write differencing stars for $\frac{\partial Q}{\partial z}$ and for $\frac{\partial^2 Q}{\partial x^2}$. Fig. 1 shows the differencing star for the entire equation and defines shorthand notations. Since the differencing star spans only two values of z, it does not matter whether the sampling along the z axis is regular or not. I denote:

$$z_2 - z_1 = \delta z. \tag{2}$$

All derivatives will be considered to be computed in the middle of the differencing star. When $x_2 - x_1 = x_3 - x_2$, the middle is in $x_2$. When the two distances are drastically different, the

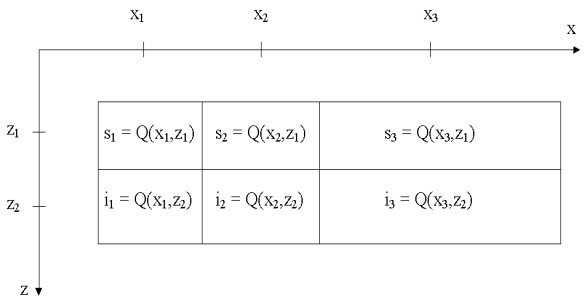



Figure 1: Differencing star for solving eq. 1 with the finite difference method; $s_1$, $s_2$, $s_3$, $i_1$, $i_2$, $i_3$ are just notations. This is the general case: the distance $x_2 - x_1$ is not necessarily equal to $x_3 - x_2$. nick1-diffstar [NR]

middle is not in $x_2$ any more and errors are introduced. This may be the cause of the spurious reflections off the large variations in grid steps that are visible in the bottom panel of Fig. 4. Dellinger and Muir (1986) discuss this problem and suggest that letting the gridpoints drift across x as we downward continue can solve it. In order to examine the extent to which this problem affects the data, I will simply pretend that the problem does not exist, and observe its negative effects. So, given the values of a function $y(x)$ in three points, $x_1$, $x_2$ and $x_3$, with $x_1 < x_2 < x_3$, the second derivative can be computed by finding the coefficients of parabola that fits through the three points. The second derivative is twice the coefficient of the second-degree term in the parabola expression. Thus, by denoting:

$$k_i = \frac{1}{\prod_{j \in \Psi - \{i\}} (x_j - x_1)}, \quad i \in \Psi = \{1,2,3\}, \tag{3}$$

which can be more explicitly written out as

$$k_1 = \frac{1}{(x_3 - x_1)(x_2 - x_1)}, \tag{4}$$

$$k_2 = \frac{-1}{(x_3 - x_2)(x_2 - x_1)}, \tag{5}$$

$$k_3 = \frac{1}{(x_3 - x_1)(x_3 - x_2)}, \tag{6}$$

the second derivative is

$$\frac{d^2 y}{dx^2} = 2\,(k_1 y_1 + k_2 y_2 + k_3 y_3)\,. \tag{7}$$

The same formula is obtained by computing a first-order approximation of the second derivative as the first derivative of the first derivative in the point $x_2$. It should be noticed that the Laplacian is a low-pass filter of the original functionr. In other words, the curvature of the hyperbola is the same in each of the three points it fits through; $x_3 - x_2$ must be really different from $x_2 - x_1$ in order for the curvature-fitting hyperbola to be affected by the error and for artifacts to be generated. This explains why the method is so robust. As Fig. 3 shows, artifacts start to become barely visible when $x_3 - x_2 = 2(x_2 - x_1)$. The partial derivatives in x of even order higher than two of the wavefield will be even more robust. A method that would make use not of the second-order derivatives but of the fourth-order ones, would be much less affected by the spurious reflections.

Using the notations in Fig. 1 and eq. (2), the second derivative expression in (7), and employing a Crank-Nicolson scheme to compute the second derivative in x, we obtain the following differencing stars:

$$\frac{\partial^2 Q}{\partial x^2} = k_1(i_1+s_1)+k_2(i_2+s_2)+k_3(i_3+s_3), \tag{8}$$

$$\frac{\partial Q}{\partial z} = \frac{1}{\delta z}(i_2-s_2), \tag{9}$$

$$\frac{\partial^3 Q}{\partial z \partial x^2} = \frac{2k_1}{\delta z}(i_1-s_1)+\frac{2k_2}{\delta z}(i_2-s_2)+\frac{2k_3}{\delta z}(i_3-s_3). \tag{10}$$

The way the derivatives of a function are computed on a irregular mesh does not depend on the nature of the function, but the result of the computation does: the lower the frequency content, the better. This means that longer wavelengths will generate less artifacts, and the spatio-temporal frequency content of the spurious reflections will therefore be biased towards the high part of the spectrum. Practically no artifacts should be produced when the function has a very low frequency content, as is the case with potential fields. Upward or downward continuation of potential fields on a unstructured mesh should be very accurate.

## IMPLEMENTATION OF A 2D, ZERO-OFFSET, CONSTANT-VELOCITY $\omega$-X SMM

Although it is not easy to find irregularly sampled zero-offset non-synthetic reflection datasets, such a case was chosen for investigation because of its simplicity; any results should be directly applicable to the prestack case.

The Appendix shows a derivation for the 15° (parabolic) wave equation (38) and for the 45° one (37). Both can be expressed as:

$$a Q_{xxz} + Q_{xx} + b Q_z = 0, \tag{11}$$

where the subscripts denote partial derivatives along the corresponding axes. Plugging in the templates in (8), (9) (10), we get:

$$\begin{array}{l} \left(\frac{2ak_1}{\Delta z}+k_1\right) i_1 + \left(\frac{2ak_2+b}{\Delta z}+k_2\right) i_2 + \left(\frac{2ak_3}{\Delta z}+k_3\right) i_3 = \\ \left(\frac{2ak_1}{\Delta z}-k_1\right) s_1 + \left(\frac{2ak_2+b}{\Delta z}-k_2\right) s_2 + \left(\frac{2ak_3}{\Delta z}-k_3\right) s_3. \end{array} \tag{12}$$

In the case of the 15° equation, $a = 0$, and for the 45° one, $a = \frac{iv}{2\omega}$. In both cases $b = \frac{2i\omega}{v}$. They are the same as for the regular sampling case, which is simply a particular case of this equation (with particular values of $k_i$). The stability of the downward continuation undertaken in this manner is proven for all practical purposes by the results in Figs. 3 and 4. This means that the special stability precautions taken by Dellinger and Muir (1986) are an unnecessary complication.

The resulting tridiagonal system is solved and the values of $Q(\omega, x, z_2)$ are found. The lens term (40) which is applied after each downward continuation step with the above described

equations does not depend in any way on the sampling of the x-axis and is therefore the same as in $\omega$ - x migrations of evenly sampled data.

Unfortunately, the so-called 1/6 trick [Claerbout (1985b), section 4.3] cannot be straightforwardly applied when the spatial axis is unevenly sampled. With a bit of work, an equivalent formula can also be deduced for the irregular sampling case.

The proof in the Appendix ensures that no hidden regular sampling assumption has been incorporated in the 15° and 45° wave equation approximations.

## SMM RESULTS - PRESENTATION AND DISCUSSION

I examined the SMM results on a simple synthetic dataset. I chose the one in Claerbout (1999) because it was small, simple, and the migration result was already known. First, I produced a surrogate irregular dataset - a zero offset section that was densely and regularly sampled across the x axis (see Fig. 2, upper panel). The image was migrated with the SMM code, as if it were a irregularly sampled dataset. Only the numerical values of the traces' coordinates were making it regular; the code was the same as for the truly irregular cases. The result is displayed in the middle panel of Fig. 2. The input data was "made" irregular by applying masks (shown in the bottom part of Fig. 2). SMM was performed on the traces present in the mask. For comparison, zero traces were introduced at the locations that were not present in the masks and surrogate irregular migration was performed.

Fig. 3 shows the input data (top panel), the zero-traces inserted migration (middle panel) and the SMM (bottom panel). The irregularity introduced by the mask is mild (the sixth trace has been eliminated in two regions). Inserting zero traces and imaging on a regular grid introduces noise farther away from the missing traces as depth increases. Overall, the image is full of incoherent high-frequency noise. Instead, the SMM result introduces no such problems. Even if the jump in sampling rate is large at the place of the missing traces (the step becomes $2 \cdot dx$ instead of $dx$), the reflections off the side of the grid are minor. This makes me believe that for jumps in the grid step of the order of half of dx and under, the artifacts would be negligible.

Fig. 4 shows the effect of severe irregularities. Both images are severely affected. The S/N ratio of the zero-traces inserted surrogate irregular migration result is lowered to the limit of interpretablity, and any mistake while interpolating the traces (if we interpolate instead of inserting zero traces) would probably have severe effects as well. The SMM result is plagued by strong spurious reflections, but we notice that these are: a) localized, and b) highly coherent and dipping in the opposite direction from the true local dips. These are prone to filtering with a reasonable apriori assumption about the direction of the dip, and can be obtained as well from unmigrated data. The SMM image overall is crisp and interpretable. The fact that the incoherent noise has not been increased is particularly important because input datasets already suffer from a S/N problem, such as those in crustal seismology.

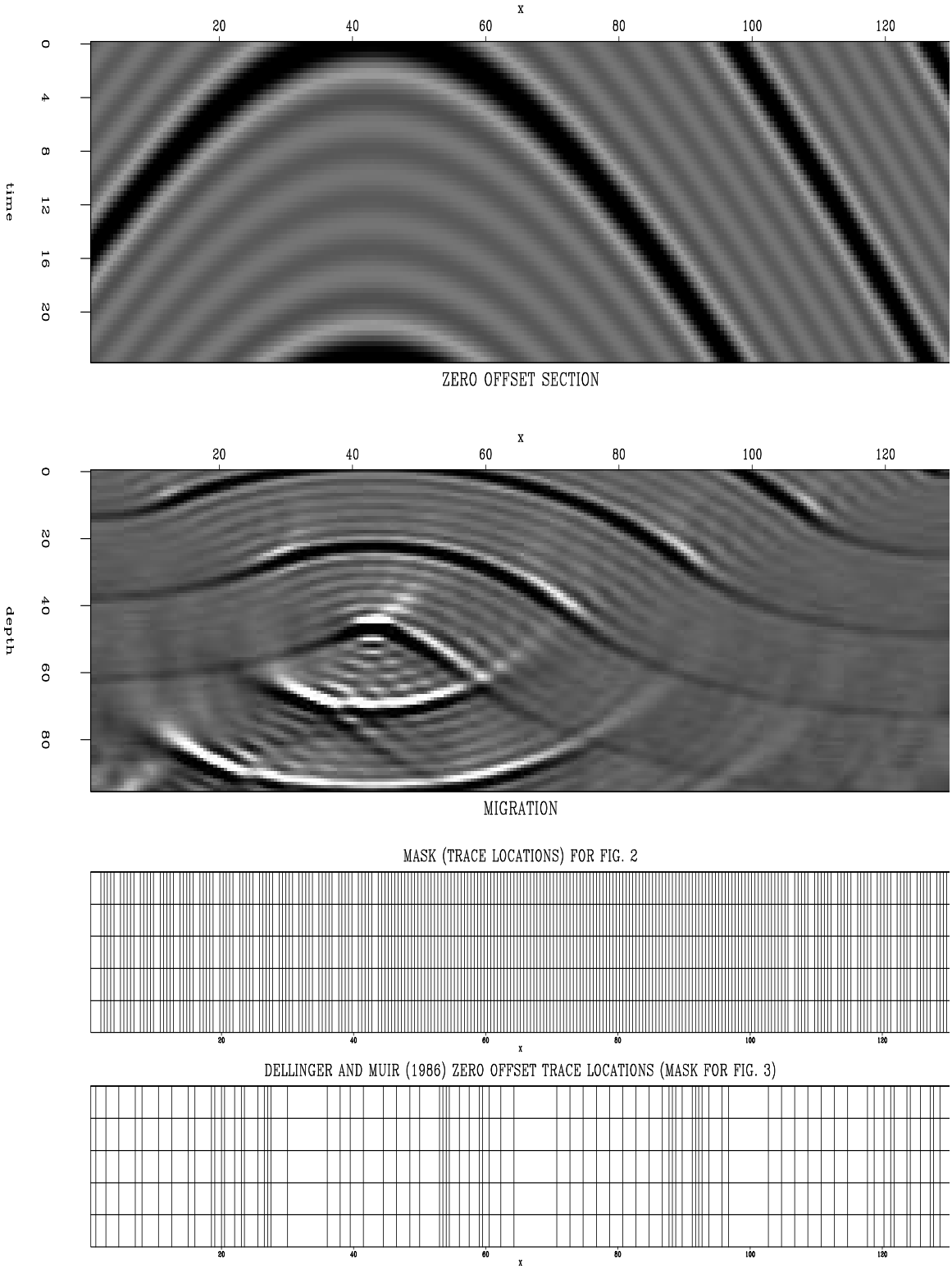



Figure 2: Input data from Claerbout (1999), its surrogate irregular 45° migration and the sampling masks used for creating irregularly sampled datasets nick1-init [ER]

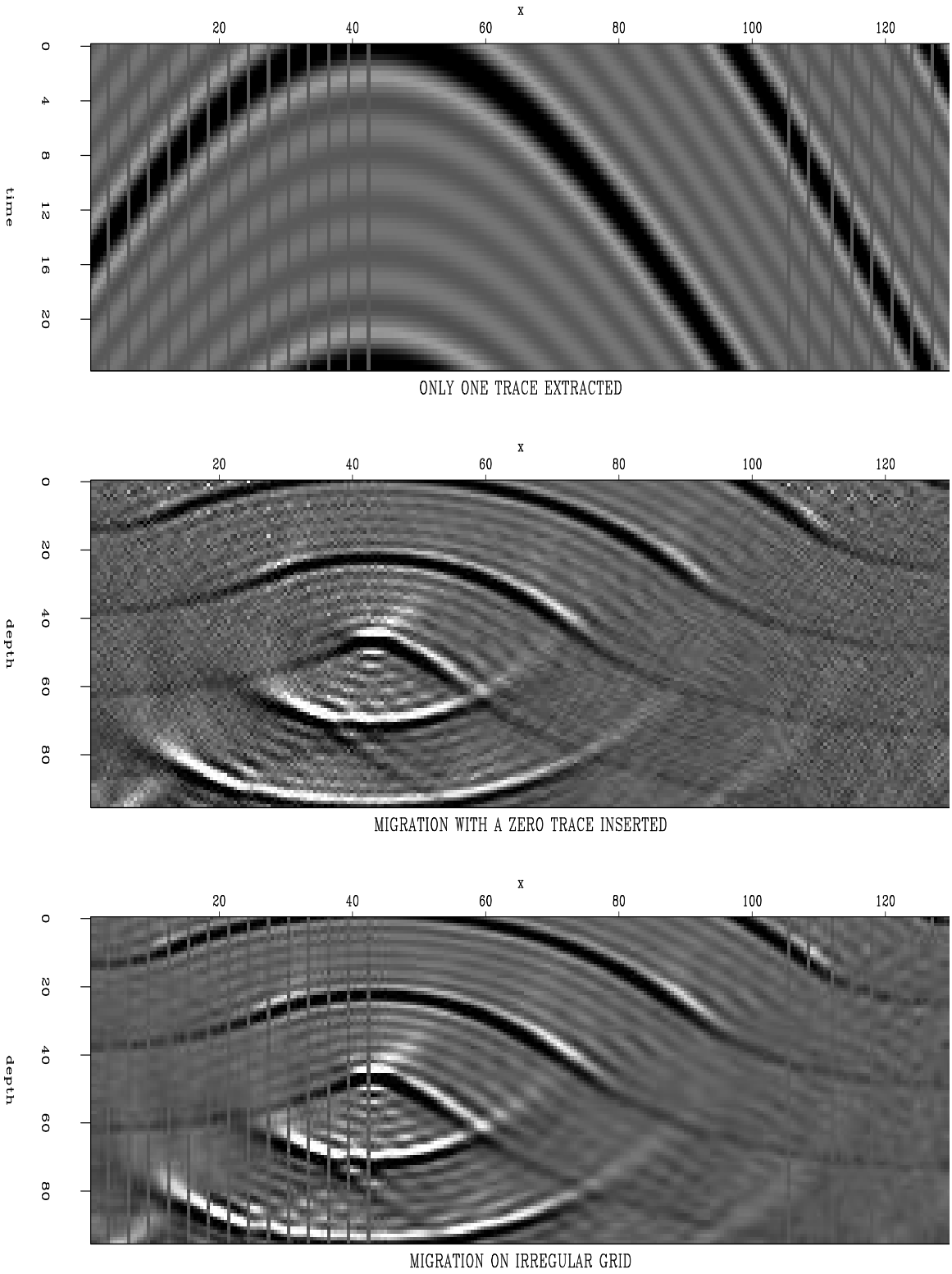



Figure 3: Input data (top panel), migration with zero-traces inserted (middle panel) and the SMM (bottom panel) nick1-midtk [ER]

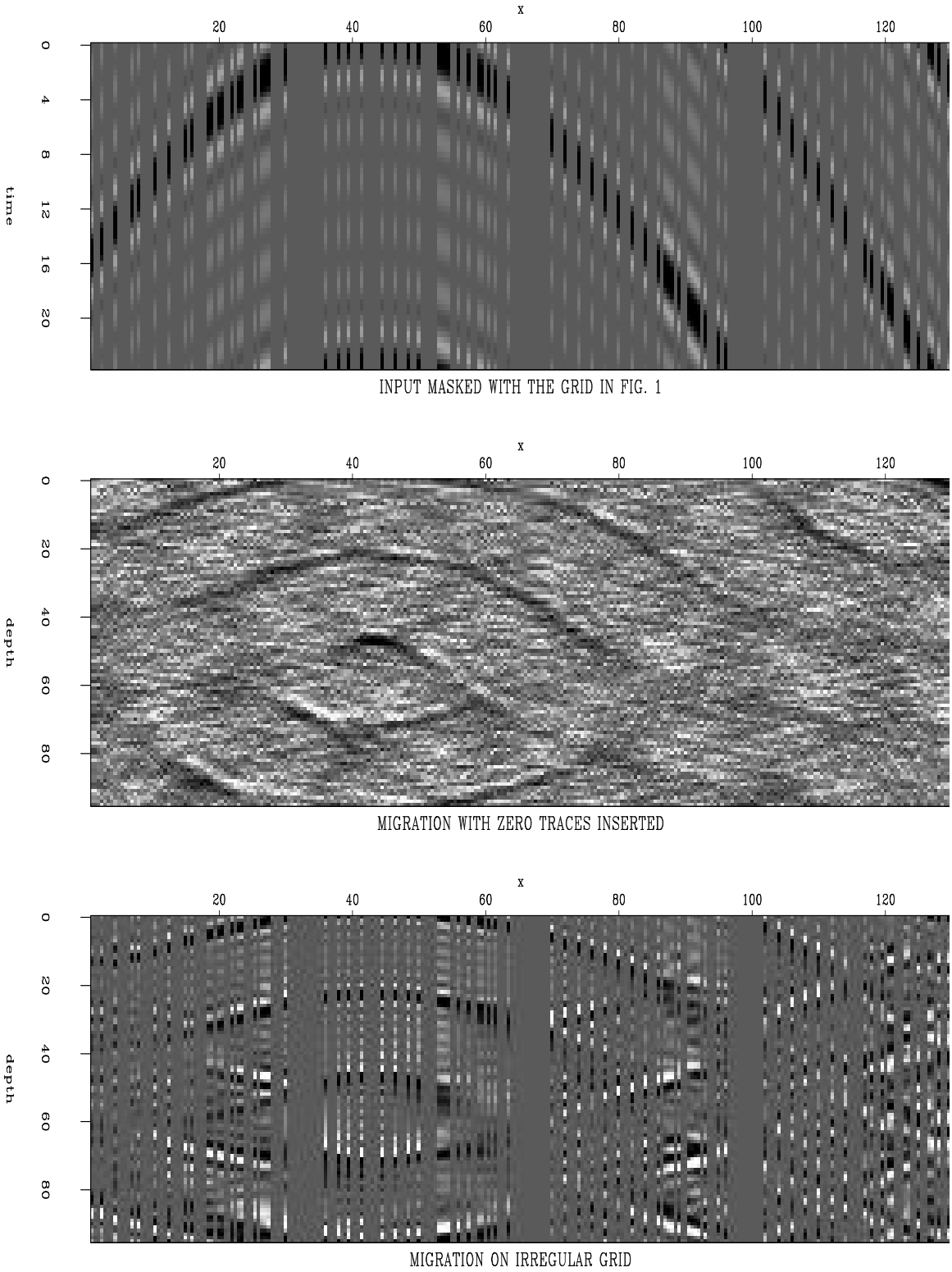



Figure 4: Input data (top panel), migration with zero-traces inserted (middle panel) and the SMM (bottom panel) nick1-reflgrid [ER]

## DEALING WITH THE SPURIOUS REFLECTIONS

There are five basic ways in which spurious reflections can be eliminated: 1) Letting the mesh points drift across the $x$ axis as we downward continue. This approach, suggested by Dellinger and Muir (1986), would regularize the spatial grid and address the actual cause of the problem. 2) Planting interpolated traces at strategic locations. 3) Solving for a wave equation that incorporates the grid irregularity ($dx$ as a function of $x$, or by applying a warping transform). 4) Filtering them out based on the fact that they are localized and highly coherent and their dip is opposite from that of local geologic dip. 5) Avoiding the problem altogether by using a numerical method for downward continuation that handles irregular data better than the finite difference method.

Method 1 seems to be the most elegant and efficient. For all methods, the biggest problem is posed by large gaps in the data coverage which would still need to be filled in with interpolated traces. A minimum trace density, related to the minimum spatial wavelength present in the data, must be maintained. The recent advances in interpolation methods of nonuniformly sampled data (Aldroubi and Grochenig, 2001) can be instrumental in this respect. Even with interpolation, the number of fill-in traces required will be smaller for an unconstrained mesh than that in the case of a Cartesian mesh. This is due to the Cartesian lattices not being the best at filling space; the same area can be covered by fewer traces placed on a quasi-regular triangular mesh. This at least should offset the burden of node number bookkeeping for an unstructured mesh. Large coverage gaps can be covered using the boundary element method. This method was created with the specific goal of not having to deal with very large numbers of elements inside a domain – its elements are only on the border of the domain. Another solution may be presented by the finite element method Marfurt (1984), which naturally handles unstructured meshes.

## EXTENSION TO 3D

Because the Laplacian of a function is an invariant, the method should work the same way in 3D. Computing the Laplacian of a function of two arguments in the nodes of an unstructured mesh is an interesting mathematical problem. The simplest approach, analoguous to the one used for Cartesian grids, is to interpolate a surface through local sets of points and to compute the Laplacian analytically from its coefficients:

$$z = a + bx + cy + dx^2 + exy + fy^2 + gx^2y^2, \tag{13}$$

$$\Delta z = 2d + 2f + 2g(x^2 + y^2). \tag{14}$$

Since the surface has seven terms, we must know the function $z = f(x,y)$ at seven points in order to find the Laplacian at one of the seven. This means solving a $7 \times 7$ linear system to find the coefficients $a$, $b$, $c$, $d$, $e$, $f$ and $g$, given three pairs of $(x,y,z)$ points. It seems frustrating that we have to compute all the seven coefficients while using only the values of $d$, $f$, and $g$. Fortunately, we do not have to solve innumerable amounts of $7 \times 7$ systems for each downward continuation step: the $x$ and $y$ values depend only on the geometry of the spatial

mesh. For each point, the matrix inversion can be done only once: in the beginning, and after that for each point. We only need to multiply the vector of $z$ values with three rows of the precomputed inverted matrix in order to find out the values of $d$, $f$, and $g$.

On a Cartesian grid, only five values are needed to compute the Laplacian. In that case, we also deal with two extra "hidden" relationships that state the particular geometrical relationships between the five points. In the general case, we do not have that information and, therefore, need more points.

The Laplacian may be found in other ways as well; perhaps interpolating with splines or other basis functions the entire wavefield - not just local neighbourhoods - at each depth step. The fastest and most elegant approach would nevertheless not involve finding a complete analytical expression of the wavefield function, but only its curvature information represented by the Laplacian.

## CONCLUSIONS

2D SMM implementation is simple – just replace the (1,-2,1) Laplacian coefficients in the differencing star with precomputed sets of geometry-dependent coefficients. It is also numerically stable. The images produced are clear and crisp, the only artifact being the spurious reflections off the abrupt variations in grid size. There are several possible ways of eliminating them. These artifacts are not likely to be generated in two particular types of practical applications – reflection measurements made from a platform moving with nonuniform speed or on a slightly-perturbed regular grid. Extension to 3D acquisition geometries is not impeded by any obvious obstacle.

## ACKNOWLEDGMENTS

I am grateful to Jonathan Franklin and Paul Sava for the helpful discussions and to Biondo Biondi for bringing relevant references to my attention.

## APPENDIX

A natural question occurs when removing a common assumption like regular sampling: Has it been discreetly incorporated into some mathematical result which may be used unknowingly? This is why I will present a derivation of the 15° and 45° approximations of the wave equation, starting from the basic laws of mechanics. The derivation will not "open" the Laplacian in (1), keeping it as an invariant to the very end. The derivation is not original: it is based on Claerbout (1985a) and Kjartansson (1978).

**The wave equation in an acoustic medium**

Let $x_i$, where $i = 1,2,3$ be three orthogonal directions and $\vec{x}$ the position vector in a coordinate system associated with the three directions. Let us define $\rho = \rho(\vec{x})$ as the mass per volume unit in the acoustic medium, $\vec{u} = \vec{u}(\vec{x})$ as the velocity of the acoustic medium and K as the bulk modulus of the acoustic medium. The second law of dynamics states that mass $\times$ acceleration = force = - pressure gradient:

$$\rho \frac{\partial u_i}{\partial t} = -\frac{\partial P}{\partial x_i}. \tag{15}$$

Energy can be stored by compression and volume variation. If

$$\vec{u}(\vec{x} + \delta\vec{x}) \neq \vec{u}(\vec{x}), \tag{16}$$

we say that the flow diverges (the volume changes). This leads to a pressure variation, proportional to the divergence of the velocity:

$$-\frac{\partial P}{\partial t} = K\left(\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3}\right). \tag{17}$$

The wave equation in an acoustic medium can be deduced from (15) and (17) as follows. Derivate (17) with respect to time:

$$\frac{\partial^2 P}{\partial t^2} = -K \sum_{i=1}^{3} \frac{\partial^2 u_i}{\partial t \partial x_i}. \tag{18}$$

Divide (15) by $\rho$ and derivate it with respect to the axis $x_i$:

$$\frac{\partial^2 u_i}{\partial t \partial x_i} = -\frac{\partial}{\partial x_i}\frac{1}{\rho}\frac{\partial P}{\partial x_i}. \tag{19}$$

Plug (19) in (18):

$$\frac{\partial^2 P}{\partial t^2} = K \sum_{i=1}^{3} \frac{\partial}{\partial x_i}\frac{1}{\rho}\frac{\partial P}{\partial x_i}. \tag{20}$$

Approximation: $\rho$ is a constant that does not depend on the position vector. By denoting the acoustic waves propagation velocity through the medium by v, where

$$v^2 = \frac{K}{\rho}, \tag{21}$$

we obtain the acoustic wave equation:

$$\frac{\partial^2 P}{\partial t^2} = v^2 \sum_{i=1}^{3} \frac{\partial^2 P}{\partial x_i^2}. \tag{22}$$

**New notations and definitions**

Until now the proof was pure physics. But because in seismology the depth axis is "special," we will change notations. Axis $x_3$ will be denoted by z, and the $\vec{x}$ will denote the surface position vector with the components $(x_1, x_2)$. Also, the Laplacian will be the operator,

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}. \tag{23}$$

We will use the Fourier transform of the pressure field along the time axis:

$$P(\vec{x}, z, \omega) = F\{P(\vec{x}, z, t)\} = \int_{-\infty}^{\infty} P(\vec{x}, z, t) e^{i\omega t} dt. \tag{24}$$

The following property of the Fourier transform will be needed:

$$F\left\{\frac{\partial^2 P(\vec{x}, z, t)}{\partial t^2}\right\} = -\omega^2 P(\vec{x}, z, \omega). \tag{25}$$

The spatial frequency is defined as:

$$m(\vec{x}, z) = \frac{\omega}{v(\vec{x}, z)}. \tag{26}$$

Let $\bar{v}$ be a spatial average of $v$ in the medium, a known constant that does not depend on $\vec{x}$ or $z$. We also define

$$\bar{m} = \frac{\omega}{\bar{v}} \tag{27}$$

and the function

$$Q(\vec{x}, z, \omega) = P(\vec{x}, z, \omega) e^{-i\bar{m}z}. \tag{28}$$

The index notation for derivatives will be used from now on. The symbol $\forall$ will denote the phrase "for all."

**Downward continuation**

Given the values of the function $P(\forall\vec{x}, z, \forall\omega)$, downward continuation consists of finding the values of $P(\forall\vec{x}, z + \delta z, \forall\omega)$. An expression describing this process lays at the end of the following proof: Obtain the Helmholtz equation by applying the Fourier transform defined by (24) to the wave equation (22) while taking into account the notation (26) and the property (25) and rearranging:

$$\Delta P + P_{zz} + m^2 P = 0. \tag{29}$$

By derivating relation (28) with respect to $x$ and $z$ we obtain:

$$P_{x_i} = Q_{x_i} e^{i\bar{m}z}, \tag{30}$$

$$\Delta P = (\Delta Q) e^{i\bar{m}z}, \tag{31}$$

$$P_z = (Q_z + i\bar{m}Q)e^{i\bar{m}z}, \tag{32}$$

$$P_{zz} = (Q_{zz} + 2i\bar{m}Q_z - \bar{m}^2 Q)e^{i\bar{m}z}. \tag{33}$$

By plugging into in (29) and eliminating the exponential, we get:

$$\Delta Q + Q_{zz} + 2i\bar{m}Q_z + \left(m^2 - \bar{m}^2\right) Q = 0. \tag{34}$$

The second derivative with respect to $z$ can be eliminated by derivating with respect to $z$, multiplying by $\frac{i}{2\bar{m}}$, and adding the result to (34):

$$\frac{i}{2\bar{m}} Q_{zzz} + \frac{i}{2\bar{m}} (\Delta Q)_z + \Delta Q + 2i\bar{m}Q_z + \frac{i}{2\bar{m}} \left(m^2 - \bar{m}^2\right) Q_z + \left(m^2 - \bar{m}^2\right) Q + \frac{im}{\bar{m}} \frac{\partial m}{\partial z} Q = 0. \tag{35}$$

Note that no approximation has been made between the wave equation (22) and this point. Eq. 35 is simply the wave equation in a different coordinate system. Now $Q_{zzz}$ is approximated by zero:

$$\frac{i}{2\bar{m}} (\Delta Q)_z + \Delta Q + 2i\bar{m}Q_z + \frac{i}{2\bar{m}} \left(m^2 - \bar{m}^2\right) Q_z + \left(m^2 - \bar{m}^2\right) Q + \frac{im}{\bar{m}} \frac{\partial m}{\partial z} Q = 0. \tag{36}$$

For the case of a homogenous medium, $\bar{m} = m$ and the equation turns into the familiar 45° equation:

$$\frac{i}{2\bar{m}} (\Delta Q)_z + \Delta Q + 2i\bar{m}Q_z = 0. \tag{37}$$

The 15° equation is obtained by neglecting the $Q_{xxz}$ term also:

$$\Delta Q + 2i\bar{m}Q_z = 0. \tag{38}$$

Downward continuation proceeds by considering

$$Q(\forall \vec{x}, z, \forall \omega) = P(\forall \vec{x}, z, \forall \omega) \tag{39}$$

then by using one of the equations 36, 37 or 38 to find the values of $Q(\forall \vec{x}, z + \delta z, \forall \omega)$ and by finally finding P by undoing the variable change:

$$P(\forall \vec{x}, z + \delta z, \forall \omega) = Q(\forall \vec{x}, z + \delta z, \forall \omega) e^{im_{(\vec{x},z)}\delta z}. \tag{40}$$

## REFERENCES

Aldroubi, A., and Grochenig, K., 2001, Nonuniform sampling and reconstruction in shift-invariant spaces: SIAM review, **43**, no. 4, 585–620.

Claerbout, J. F., 1985a, Fundamentals of geophysical data processing: with applications to petroleum prospecting: Blackwell Scientific Publications.

Claerbout, J. F., 1985b, Imaging the Earth's Interior: Blackwell Scientific Publications.

Claerbout, J. F. Basic Earth Imaging:. http://sepwww.stanford.edu/sep/prof/bei/toc_html, 1999.

Dellinger, J., and Muir, F., 1986, Finite differencing with uneven spatial sampling: SEP–**48**, 269–276.

Kjartansson, E., 1978, Modeling and migration with the monochromatic wave equation - variable velocity and attenuation: SEP–**15**, 1–19.

Marfurt, K. J., 1984, Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations: Geophysics, **49**, no. 5, 533–549.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 253–267*

# Short Note

# Effect of velocity uncertainty on amplitude information

*Robert G. Clapp*[1]

## INTRODUCTION

Risk assessment is a key component to any business decision. Geostatistics has recognized this need and has introduced methods, such as simulation, to attempt to assess uncertainty in their estimates of earth properties (Isaaks and Srivastava, 1989). Geophysics has been slower to recognize this need, as methods which produce a single solution have long been the norm.

The single solution approach has a couple of significant drawbacks. First, since least-squares estimates invert for the minimum energy/variance solution, our models tend to have lower spatial frequency than the true model. Second, it does not provide information on model variability or provide error bars on the model estimate. Geostatisticians have both of these abilities in their repertoire through what they refer to as "multiple realizations" or "stochastic simulations." They introduce a random component, based on properties of the data, such as variance, to their estimation procedure. Each realization's frequency content is more representative of the true model's and by comparing and contrasting the equiprobable realizations, model variability can be assessed. These models are often used for non-linear problems, such as fluid flow. In this approach representative realizations are used as an input to a flow simulator.

In geophysics we have a similar non-linear relationship between velocity and migration amplitudes. Migration amplitudes are used for rock property estimates yet we normally don't assess how velocity uncertainty, and the low frequency nature of our velocity estimates, affect our migration amplitudes. The geostatistical approach is not well suited to answer this question. Our velocity covariance is highly spatially variant, and our velocity estimation problem is non-linear.

In previous works (Clapp, 2000, 2001a,b), I showed how we can modify standard geophysical inverse techniques by adding random noise into the model styling goal to obtain multiple realizations. In this paper I apply this methodology to a conventional velocity analysis problem. I then migrate the data with various velocity realization. I perform Amplitude vs. Angle (AVA) analysis on each migrated image. Finally, I calculate the mean and variance of the AVA

[1]**email:** bob@sep.stanford.edu

parameter estimates for the various relations. In this paper I review the operator based multi-realization methodology. I then apply the methodology on a structural simple 2-D land dataset from Columbia.

## MODEL VARIANCE

We can characterize the standard geophysical problem as a linear relationship **L** between a model **m** and **d**, with a regularization operator **A**. In terms of fitting goals this is:

$$\begin{aligned} \mathbf{0} \approx \mathbf{r_d} &= \mathbf{d} - \mathbf{Lm} \\ \mathbf{0} \approx \mathbf{r_m} &= \epsilon \mathbf{Am}. \end{aligned} \tag{1}$$

Ideally **A** should be the inverse model covariance. If so, given an accurate modeling operator we would expect $\mathbf{r_m}$ to be zero. In fact, **A** is an approximation of the inverse model covariance. In practice, we usually assume stationarity, and design **A** to accurately describe the second order statistics of the model. The first order statistics, the spatial variance of the model, are not included. We can produce models that have similar *spatial* variance as the true model by modifying the second goal. This is done by replacing the zero vector **0** with standard normal noise vector $\eta$, scaled by some scalar $\sigma_m$,

$$\begin{aligned} \mathbf{0} &\approx \mathbf{d} - \mathbf{Lm} \\ \sigma_m \eta &\approx \epsilon \mathbf{Am}. \end{aligned} \tag{2}$$

For the special case of missing data problems, where **L** is simply a masking operator **J** delineating known and unknown points, Claerbout (1998) showed how $\sigma_m$ can be approximated by first estimated model through the fitting goals in (1). Then, by solving,

$$\sigma_m \quad = \quad \frac{\mathbf{1}'\mathbf{J}\mathbf{r}m^2}{\mathbf{1}'\mathbf{J}\mathbf{1}}, \tag{3}$$

where **1** is a vector composed of 1s. This basically says that we can find the right level of noise by looking at the residual resulting from applying our inverse covariance estimate on known data locations. If we make the assumption that **L** is accurate we can use (3) for a more general case. In the more general case, the operator is 1 at locations where $\mathbf{L}'\mathbf{1}$ is non-zero.

### Tomography

The way I formulate my tomography fitting goals requires some deviation from the generic multi-realization form. My tomography fitting goals are fully described in Clapp (2001a). Generally, I relate change in slowness $\Delta\mathbf{s}$, to change in travel time $\Delta\mathbf{t}$ by a linear operator **T** The tomography operator is constructed by linearizing around an initial slowness model $\mathbf{s_0}$. I regularize the slowness **s** rather than change in slowness and obtain the fitting goals,

$$\begin{aligned} \Delta\mathbf{t} &\approx \mathbf{T}\Delta\mathbf{s} \\ \epsilon\mathbf{As_0} &\approx \epsilon\mathbf{A}\Delta\mathbf{s}. \end{aligned} \tag{4}$$

The calculation of $\sigma_d$ is the same procedure as shown in equation (3). The only difference is now we initiate $\mathbf{r_m}$ with both our random noise component $\sigma_m \eta$ and $\epsilon \mathbf{A s_0}$. A cororarly approach for data uncertainty is discussed in Appendix A.

**Results**

To test the methodology I decided to start with a structurally simple 2-D line from a land dataset from Columbia provided by Ecopetrol. Figure 1 shows the estimated velocity for the data. Note how it is generally $v(z)$ with some deviation, especially in the lower portion of the image. Figure 2 shows the result of performing split-step phase shift migration and Figure 3 shows the resulting angle gathers (Sava, 2000). Note how the image is generally well focused and the gathers with some slight variation below three kilometers at $x = 3.5$. Figure 4 shows the moveout of the gathers in Figure 3. Note the traditional 'W' pattern associated with the velocity anomaly can be seen in cross-section at depth.

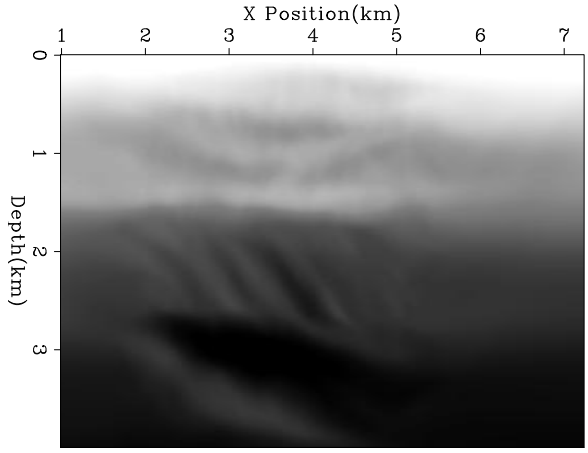



Figure 1: Initial velocity model. bob7-vel-init [CR]

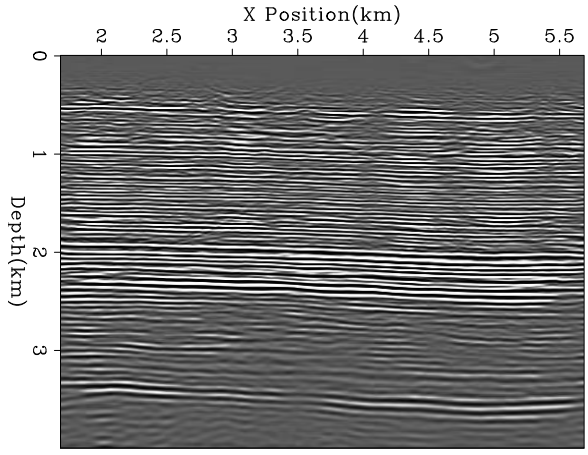



Figure 2: Initial migration using the velocity shown in Figure 1. bob7-image-init [CR]

To start we need to solve the problem without accounting for model variance. If we solve for $\Delta \mathbf{s}$ using fitting goals (4) our updated velocity is shown in Figure 5. The change of the velocity is generally minor, with an increase in the high velocity structure at $x = 3.5$, $z = 3.2$.

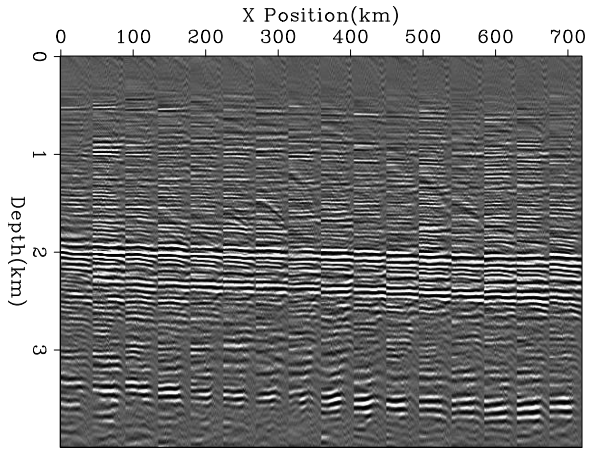



Figure 3: Every 10th migrated gather using the velocity shown in Figure 1. bob7-mig-init [CR]

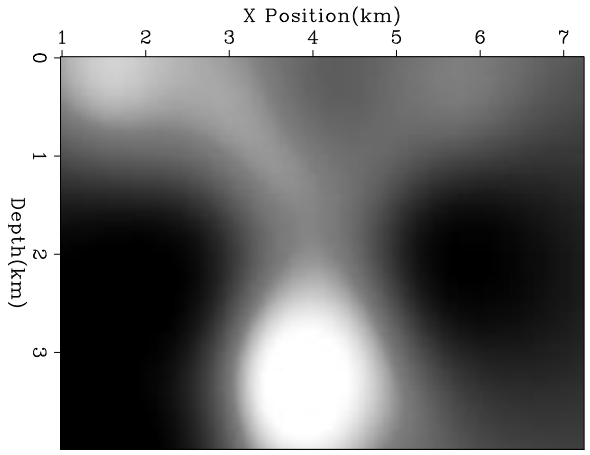



Figure 4: Moveout of the gathers shown in Figure 3. bob7-semb-init [CR]

The resulting image and migration gathers are shown in Figures 6 and 7. The resulting image is slightly better focused below the anomaly and the migration gathers are, as expected, a little flatter. If we apply equation (3) using the $\mathbf{r_n}$ when estimating our improved velocity model

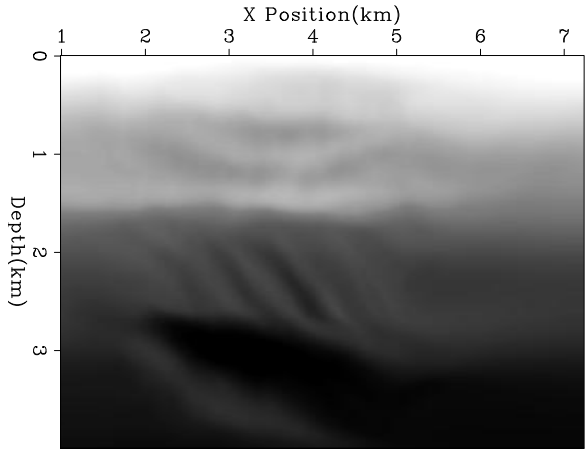



Figure 5: New velocity obtained by inverting for $\Delta\mathbf{s}$ using fitting goals (4). bob7-vel-none [CR]

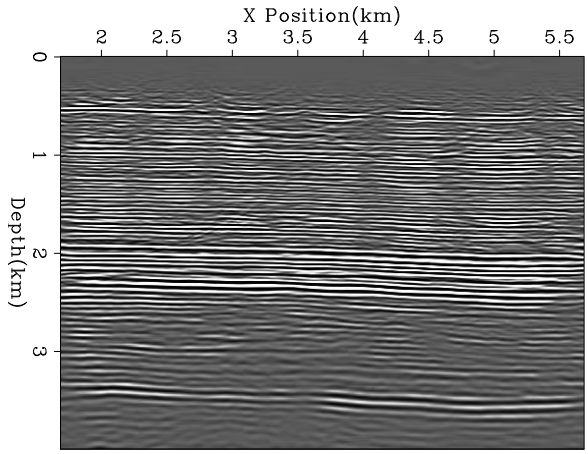



Figure 6: New image obtained by inverting for $\Delta\mathbf{s}$ using fitting goals (4) using the velocity shown in Figure 5. bob7-image-none [CR]

we can find the right amount of noise to add to our fitting goals. We can now resolve for $\Delta\mathbf{s}$ accounting for the model variability. Figure 8 shows four such realizations. Note that they have the same general structure as seen in Figure 5 but within additional texture that is accounted for by covariance description. If we migrate with these new velocity models we get the images and migrated gathers shown in Figures 9 and 10. In printed form these images appear identical, or close to identical. If watched as a movie, amplitude differences can be observed.

## AVA analysis

For the AVA analysis I chose the simple slope*intercept (A*B) methodology used in (Castagna et al., 1998; Gratwick, 2001). Figure 11 shows the slope (left), intercept (center), and slope*intercept (right) for the migrated image without model variability. Note the positive, hydrocarbon indicating, anomalies circled at approximately 2.3 km.

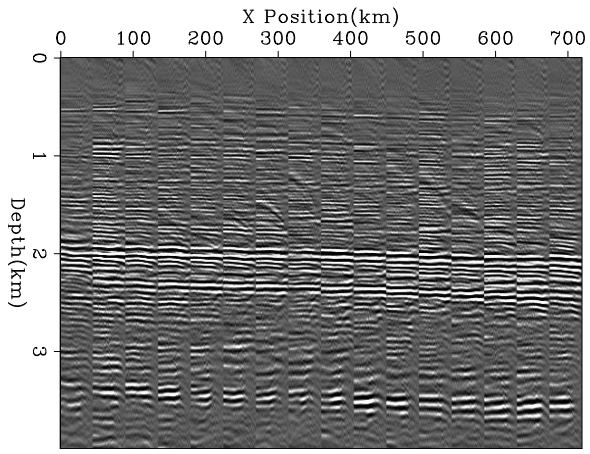



Figure 7: New gathers obtained by inverting for $\Delta\mathbf{s}$ using fitting goals (4) using the velocity shown in Figure 5. bob7-mig-none [CR]

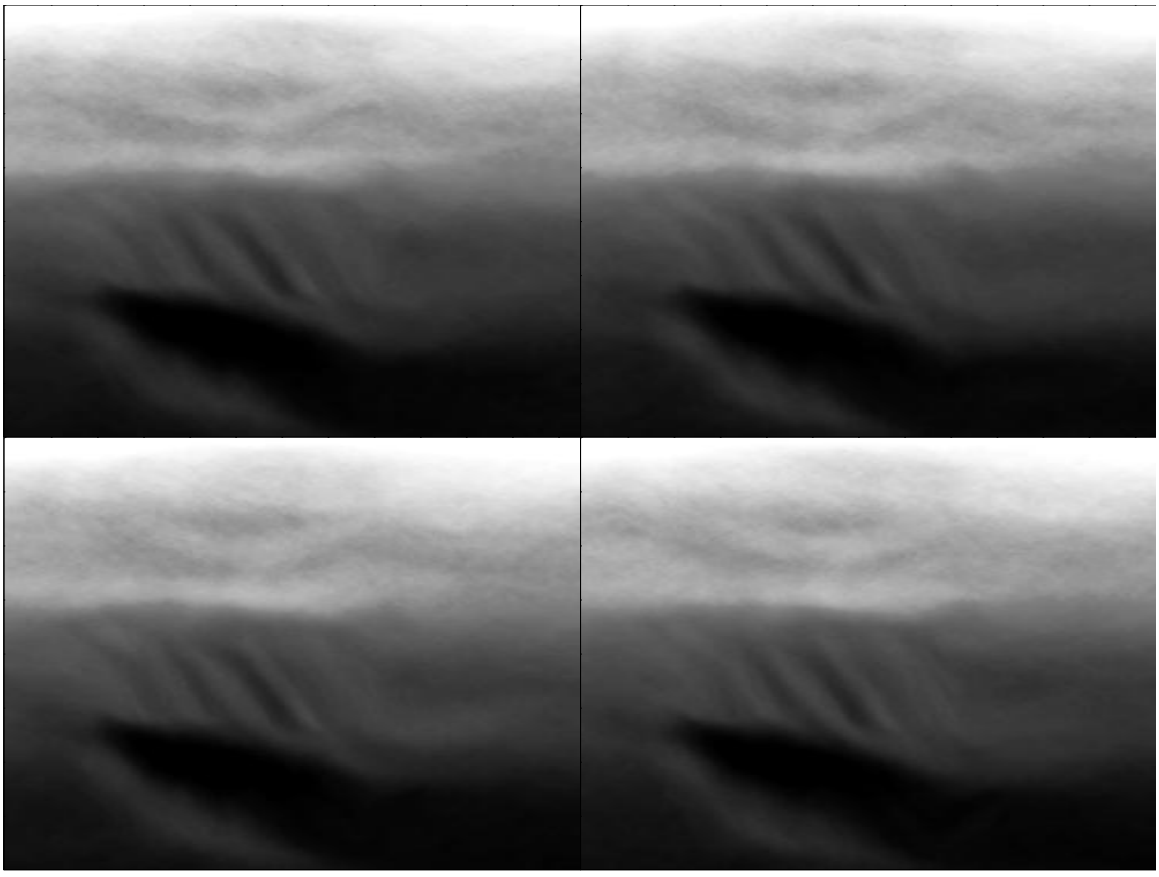


Figure 8: Four different realizations of the velocity accounting for model variability. bob7-vel-multi [CR,M]

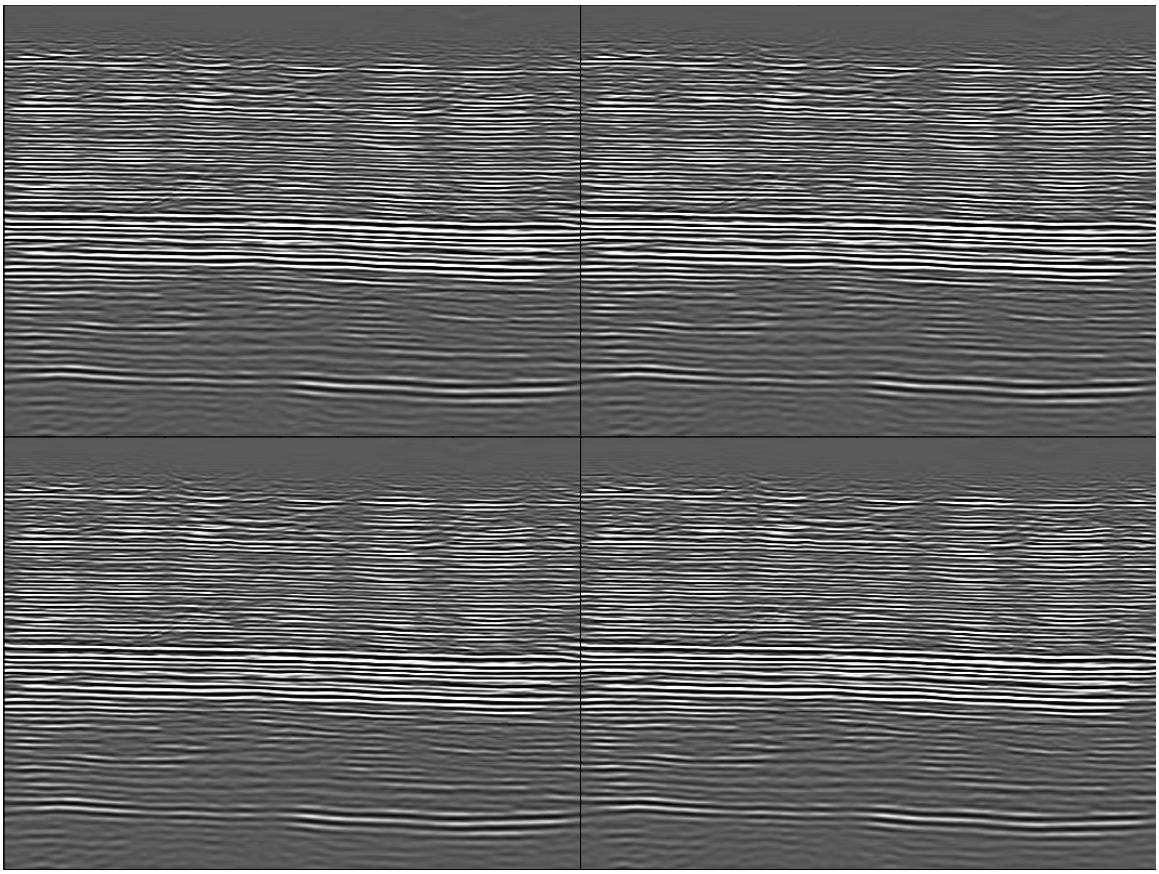


Figure 9: Four different realizations of the migration accounting for model variability. Note how the reflector position is nearly identical in each realization and with the image without variability (Figure 6), but the amplitudes vary slightly. bob7-image-multi [CR,M]

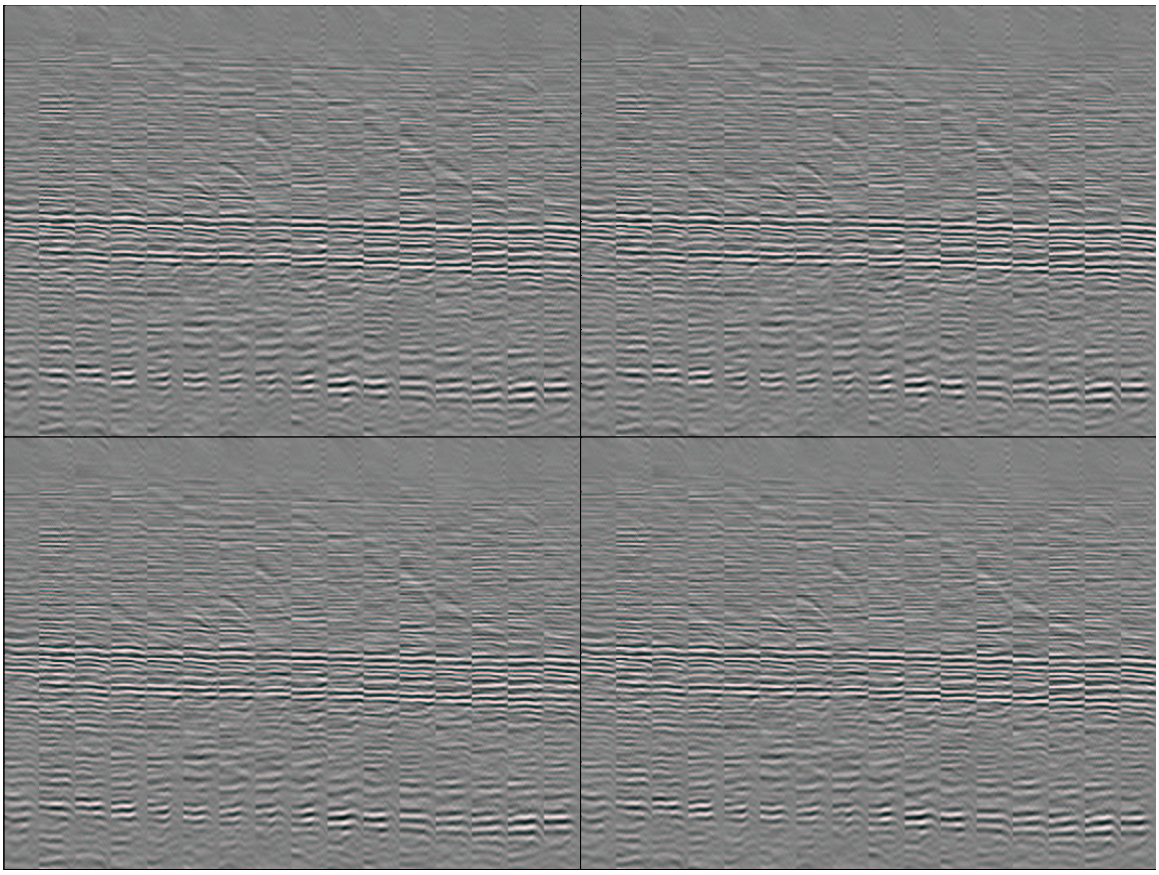


Figure 10: Four different realizations of the migration accounting for model variability. Note how the reflector position is nearly identical in each realization and with the image without variability (Figure 7). bob7-mig-multi [CR,M]

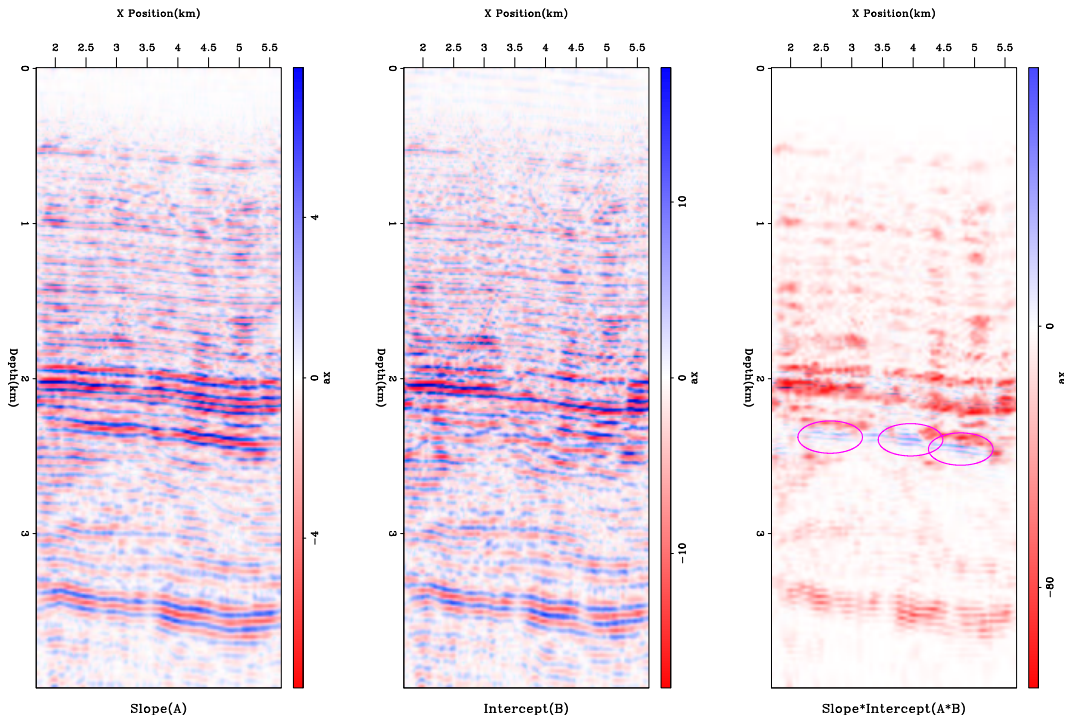



Figure 11: AVA analysis for the migrated image in Figure 7. The left panel shows the slope, the center the intercept, and the right panel the slope*intercept. bob7-ava-none [CR,M]

I then performed the same procedure on all of the migrated images obtained from the various realizations (Figure 12). The left panel shows intercept, the center panel slope, the right panel, slope*intercept. The top shows the average of the realizations. The center panel shows the variance of the realizations. The bottom panel shows the variance scaled by the inverse of the smoothed amplitude. What is interesting is the varying behavior at the three zones with hydrocarbon indicators. Figure 13 shows a closeup in the zone with the hydrocarbon indicators. The left blob ‘A’ shows a high variance in the AVA indicator. The center blob ‘B’ shows a mild variance, and the right blob ‘C’ shows low variance. This would seem to indicate that at location ‘C’ the hydrocarbon indicator is more valid. Without drilling of each target a more general conclusion cannot be drawn.

## CONCLUSIONS

I showed how AVA parameter variability can be assessed by adding a random component to our fitting goals when estimating velocity. The methodology shows promise in allowing error bars to be placed upon AVA parameter estimates.

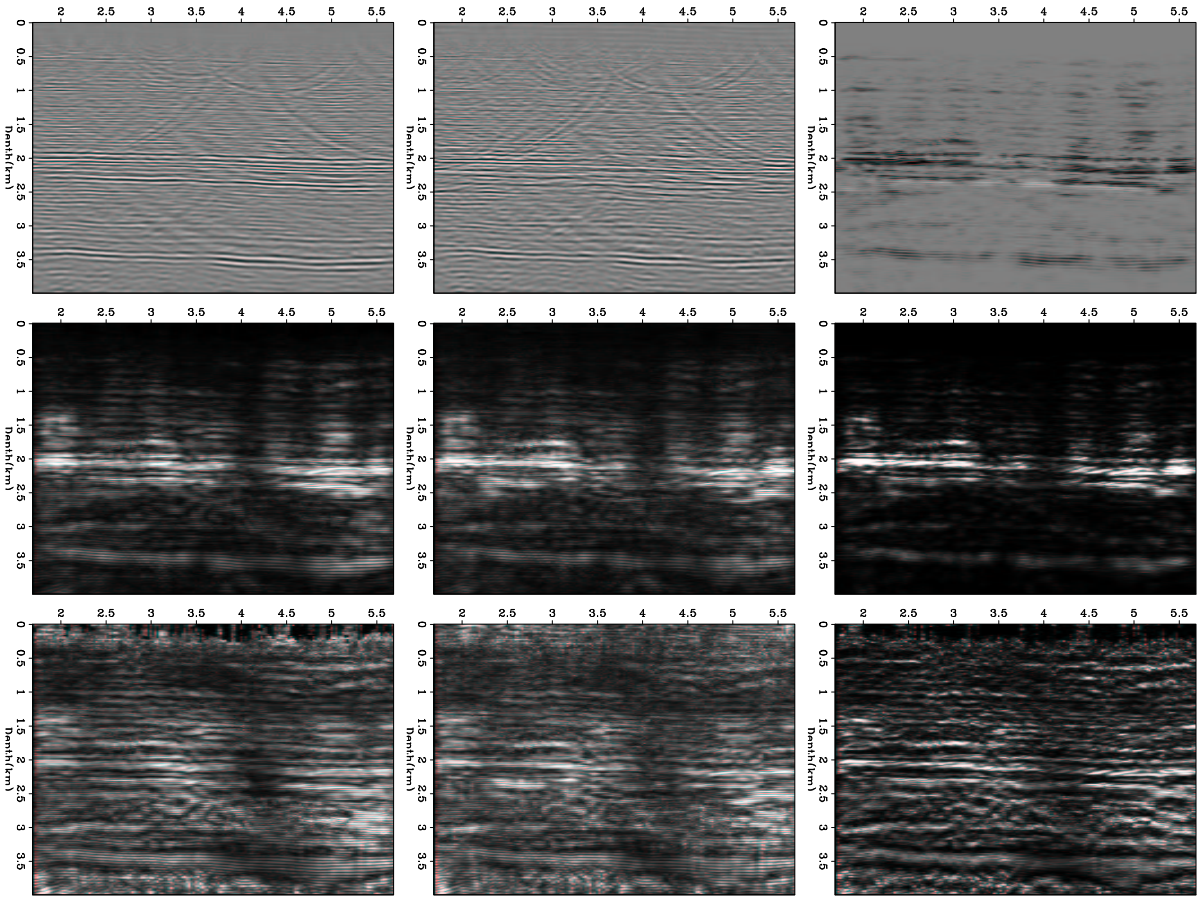



Figure 12: AVA analysis for the the various velocity realizations. The left panel shows intercept, the center panel slope, the right panel, slope*intercept. The top shows the average of the realizations. The center panel shows the variance of the realizations. The bottom panel shows the variance inverse scaled by a smoothed amplitude. bob7-ava-multi [CR,M]

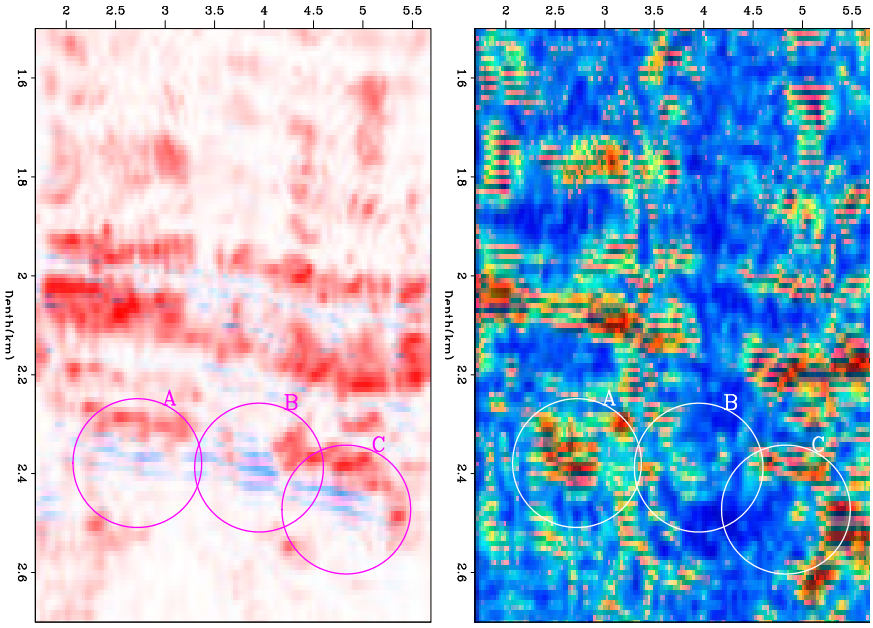



Figure 13: A close up of the reservoir zone. The left panel shows the slope*intercept. The right panel shows the variance of the slope*intercept for the various realizations. Note how the left blob ‘A’ shows a high variance in the AVA indicator. The center blob ‘B’ shows a mild variance, and the right blob ‘C’ shows low variance. bob7-ava-multi-close [CR,M]

## ACKNOWLEDGMENTS

I would like to thank Ecopetrol for the data used in this paper.

## REFERENCES

Castagna, J. P., Swan, H. W., and Foster, D. J., 1998, Framework for AVO gradient and intercept interpretation: Geophysics, **63**, no. 3, 948–956.

Claerbout, J. Geophysical Estimation by Example: Environmental soundings image enhancement:. http://sepwww.stanford.edu/sep/prof/, 1998.

Clapp, R., 2000, Multiple realizations using standard inversion techniques: SEP–**105**, 67–78.

Clapp, R. G., 2001a, Geologically constrained migration velocity analysis: Ph.D. thesis, Stanford University.

Clapp, R. G., 2001b, Multiple realizations: Model variance and data uncertainty: SEP–**108**, 147–158.

Gratwick, D., 2001, Amplitude analysis in the angle domain: SEP–**108**, 45–62.

Guitton, A., 2000, Coherent noise attenuation using Inverse Problems and Prediction Error Filters: SEP–**105**, 27–48.

Isaaks, E. H., and Srivastava, R. M., 1989, An Introduction to Applied Geostatistics: Oxford University Press.

Sava, P., 2000, Variable-velocity prestack Stolt residual migration with application to a North Sea dataset: SEP–**103**, 147–157.

## APPENDIX A

We can follow a parallel definition for the data fitting goal in terms of the inverse noise covariance **N**:

$$\sigma_d \eta \approx \mathbf{N}(\mathbf{d} - \mathbf{Lm}). \tag{A-1}$$

### Noise covariance for velocity estimation

Using the multiple realization methodology for velocity estimation problem posed in the manner results in several difficulties. First, what I would ideally like is a model of the noise. This poses the problem of how to get the noise inverse covariance. The first obstacle is that our data is generally a uniform function of angle $\theta$ and a non-uniform function of **x**. What we would really like is a uniform function of just space. We can get this by first removing the angle portion of our data.

I obtain $\Delta\mathbf{t}$ by finding the moveout parameter $\gamma$ that best describes the moveout in migrated angle gathers. I calculate $\Delta\mathbf{t}$ by mapping my selected $\gamma$ parameter back into residual moveout and the multiplying by the local velocity. Conversely I can write my fitting goals in terms of $\gamma_\mathbf{i}$ by introducing an operator **S** that maps $\Delta\mathbf{t}$ to $\gamma$,

$$\begin{aligned} \gamma_\mathbf{i} &\approx \mathbf{ST}\Delta\mathbf{s} \\ \mathbf{As_0} &\approx \epsilon\mathbf{A}\Delta\mathbf{s}. \end{aligned} \tag{A-2}$$

Making the data a uniform function of space is even easier. I can easily write an operator that maps my irregular $\gamma_\mathbf{i}$ to a regular function of $\gamma_\mathbf{r}$ by a simple inverse interpolation operator **M**. I then obtain a new set of fitting goals,

$$\begin{aligned} \gamma_\mathbf{r} &\approx \mathbf{MST}\Delta\mathbf{s} \\ \mathbf{As_0} &\approx \epsilon\mathbf{A}\Delta\mathbf{s}. \end{aligned} \tag{A-3}$$

On this regular field the noise inverse covariance **N** is easier to get a handle on. We can approximate the noise inverse covariance as a chain of two operators. The first, $\mathbf{N_1}$, f a fairly traditional diagonal operator that amounts for uncertainty in our measurements. For the tomography problem this translate into the width of our semblance blob. For the second operator we can estimate a Prediction Error Filter (PEF) on $\mathbf{r_d}$ (Guitton, 2000) after solving

$$\begin{aligned} \mathbf{0} &\approx \mathbf{r_d} = \mathbf{N_1}(\gamma_\mathbf{r}\mathbf{MST}\Delta\mathbf{s}) \\ \mathbf{As_0} &\approx = \mathbf{r_m}\epsilon\mathbf{A}\Delta\mathbf{s}. \end{aligned} \tag{A-4}$$

If we combine all these points and add in the data variance we get,

$$\begin{aligned} \sigma_d \eta &\approx \mathbf{N_1}\mathbf{N_2}(\gamma_\mathbf{r} - \mathbf{MST}\Delta\mathbf{s}) \\ \sigma_m \eta \mathbf{As_0} &\approx \epsilon\mathbf{A}\Delta\mathbf{s}. \end{aligned} \tag{A-5}$$


ST
2100
SOUTH PARK
LONGER & UNCUT

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 267–287*

# An extension of poroelastic analysis to double-porosity materials: A new technique in microgeomechanics

*James G. Berryman*

**ABSTRACT**

Double-porosity materials were introduced as models for oil and gas reservoirs having both storage and transport porosities, and were at first usually treated as static mechanical systems in order to study the flow patterns of fluids during reservoir pump-down. Because fluid withdrawal normally increases the effective stress acting on the reservoir, it also turns out to be important to study the geomechanics of the reservoir and how changing fluid pressure affects the solid compaction and fluid permeability of these systems. At the microscale, the mechanical properties of the solid constituents and their distribution in space determine the overall macromechanics of the reservoir system. For systems containing two porosities and two types of solid constituents, exact results for all but one (which may be taken as the overall drained bulk modulus of the system) of the mechanical constants can be derived when the constituents' properties are known using methods developed in this paper. For multi-porosity systems, closure of the system of equations remains an open question, although it is clear that the system can always be closed by the addition of further macroscale measurements.

## INTRODUCTION

The subject of "geomechanics" includes such topics as the study of rock mechanics, soil mechanics, and engineering geology, and has overlapping interests in some cases with "hydrogeology" when the mechanical behavior of the earth system of interest is strongly affected by the presence of water. In general "geomechanics" means the study mechanics of earth systems, and therefore "microgeomechanics" means the study of the effects of micromechanics on earth systems. Our main interest here will be in the interaction of fluid pressure changes (usually induced by reservoir depletion) with the mechanical properties of the reservoir.

Perhaps the most typical applications of geomechanics arise in the engineering disciplines of mining and oil reservoir assessment and production, and even earlier in soil mechanics. The history of the main features contained in the theory of geomechanics dates back at least to the work of Terzaghi (1925) on "effective stress," which is the observation that, when external confining stress and internal pore pressure act simultaneously on a porous material, the pore pressure tends to counteract the confining pressure. Terzaghi's effective stress law for changes in volume was the simple statement that the effective stress was the confining stress minus the pore pressure, *i.e.* the differential stress. For soils, this approximation is often a very good one.

For porous materials in general, theory and experiment have shown both that the concept of an effective stress is valid, and that the actual effective stress is not just the differential stress of Terzaghi, but rather the confining stress minus some fraction of the pore pressure. This fraction has often been taken to be some overall average number — for example, in the range 0.85–0.88 (Brandt, 1955; Schopper, 1982) — for the earth's crust. But it has been shown theoretically by Biot and Willis (1957) and experimentally by Fatt (1958; 1959) and Nur and Byerlee (1971) that the volume effective stress coefficient is actually related to the elastic properties of the porous system. If $\alpha$ is this effective stress coefficient, then — in a microhomogeneous porous material (composed of voids and a single type of granular material) — it is related to the bulk modulus of the grains $K_g$ and the overall bulk modulus of the drained porous system $K^*$ by $\alpha = 1 - K^*/K_g$. This rule reduces to Terzaghi's effective stress rule for soils when the soil is very poorly consolidated so that $K^* << K_g$, for then $\alpha \simeq 1$.

Terzaghi's early work was expanded into a theory of consolidation, both by himself and through the work of Biot (1941), Gassmann (1951), Skempton (1957), Geerstma (1966; 1973), and many others. Biot (1941) is usually given credit for the first comprehensive theory of consolidation, at least in the case of simple, single porosity systems. Gassmann (1951) was the first to obtain one of the fundamental results of the theory – sometimes called the fluid-substitution formula, relating the dry or drained bulk modulus $K^*$ to the undrained (or saturated) modulus $K_u$ by $K_u = K^*/(1 - \alpha B)$, where the pore-pressure buildup coefficient $B$ is Skempton's second coefficient (Skempton, 1954; Carroll 1980). Geertsma (1966; 1973) was one of the earliest users of this theory for studies of compaction of oil and gas reservoirs during and after drawdown of the reservoir. Wang (2000) also gives Geerstma (1966) credit for coining the term "poroelasticity" that is normally used today to describe Biot's theory in full. Early laboratory measurements (Biot and Willis 1957; Fatt 1958; 1959) of the constants in Biot's equations helped to establish the theory. Early engineering solutions of the equations of poroelasticity were given by Rice and Cleary (1976) and Cleary (1977) which helped to make it a standard tool in civil engineering. Some fundamental extensions of the theory to systems having multiple solid constituents have been given by Brown and Korringa (1975), Rice (1975), Berryman and Milton (1991), Berryman (1992), and Norris (1992). Poroelasticity is now a well-established subject having recent technical reviews by Detournay and Cheng (1993), Wang (1993), Pride and Berryman (1998), Berryman (1999), and books by Bourbié, Coussy and Zinzner (1987) and Wang (2000) describing the current state of our understanding.

Another fascinating use of the theory of poroelasticity in reservoirs is its relatively recent application to the studies of earthquakes induced by oil and gas reservoir pumping (Kovach 1974; Pennington *et al.* 1986; Segall 1985; 1989; 1992; Segall and Fitzgerald 1998; Gomberg and Wolf 1999; Pennington 2001). A related issue arising in the opposite physical extreme is the subject of $CO_2$ sequestration in the earth (Wawersik *et al.* 2001), where it is clear that pumping pressurized fluids into the ground must have a strong tendency to decrease the effective stress in the earth system used for sequestration. Decreasing effective stress implies weakening of the system, resulting in undesirable (for this application) increases in fluid permeability. Studies of partially saturated systems are also of continuing interest (Li, Zhong, and Pyrak-Nolte 2001) both for oil and gas exploitation and for environmental cleanup applications.

Biot's original single-porosity, microhomogeneous theory of poroelasticity has significant limitations when the porous medium of interest is very heterogeneous. One important generalization of poroelasticity that has been studied extensively started with the work on double-porosity dual-permeability systems by Barenblatt and Zheltov (1960) and Warren and Root (1963). These papers take explicit note of the fact that real reservoirs tend to be very heterogeneous in both their porosity and permeability characteristics. In particular, the two types of porosity normally treated are storage and transport porosities. Storage porosity holds most of the volume of the fluid underground but may have rather low permeability, while the transport porosity is low volume but high permeability. The transport porosity is usually treated as being in the form of fractures in the reservoir, or joints in the rock mass. The theory of double-porosity dual-permeability media has been expanding in both volume and scope during the last 20 years, and now includes work by Wilson and Aifantis (1983), Elsworth and Bai (1992), Bai, Elsworth, and Roegiers (1993), Berryman and Wang (1995), Tuncay and Corapciaglu (1995), Bai (1999), and Berryman and Pride (2002). Computations of transport and subsidence in double-porosity dual-permeability media include work by Khaled, Beskos, and Aifantis (1984), Nilson and Lie (1990), Cho, Plesha, and Haimson (1991), Lewallen and Wang (1998), and Bai, Meng, Elsworth, Abousleiman, and Roegiers (1999).

Some technical details follow on the single-porosity poroelasticity needed in the main arguments of the paper. Then equations are formulated for double-porosity systems, and finally multi-porosity systems are discussed. The focus will be on determining how the coefficients of the resulting equations depend on the physical properties of the microstructural constituents' of these complex geomechanical systems. The main results are obtained using new techniques in micromechanics that permit a rather elementary analysis of these complex systems to be carried through exactly. For systems containing two porosities and two types of solid constituents, exact results for all but one (which may be taken as the overall drained bulk modulus of the system) of the macroscopic geomechanical constants are derived.

## SINGLE-POROSITY GEOMECHANICS

In the absence of external driving forces that can maintain fluid-pressure differentials over long time periods, double-porosity and multi-porosity models must all reduce to single-porosity models. This reduction occurs in the long-time limit when the matrix fluid pressure and joint fluid pressure become equal. It is therefore necessary to remind ourselves of the basic results for single-porosity models in poroelasticity (Biot 1941; Detournay and Cheng 1993; Wang 2000), as the long-time behavior may be viewed as providing limiting temporal boundary conditions (for $t \to \infty$) on the analysis of multi-porosity coefficients. Further, in the specific models we adopt for the geomechanical constants in the multi-porosity theory, extensive use of the single-porosity results will be made.

The volume changes of any isothermal, isotropic material can only be created by hydrostatic pressure changes. The two fundamental pressures of single-porosity poroelasticity are the confining (external) pressure $p_c$ and the fluid (pore) pressure $p_f$. The differential pressure (or Terzaghi effective stress) $p_d \equiv p_c - p_f$ is often used instead of the confining pressure. The

volumetric response of a sample due to small changes in $p_d$ and $p_f$ take the form [*e.g.*, Brown and Korringa (1975)]

$$-\frac{\delta V}{V} = \frac{\delta p_d}{K^*} + \frac{\delta p_f}{K_s} \tag{1}$$

for the total volume $V$,

$$-\frac{\delta V_\phi}{V_\phi} = \frac{\delta p_d}{K_p} + \frac{\delta p_f}{K_\phi} \tag{2}$$

for the pore volume $V_\phi = \phi V$ (where $\phi$ is the porosity), and

$$-\frac{\delta V_f}{V_f} = \frac{\delta p_f}{K_f} \tag{3}$$

for the fluid volume $V_f$. Equation (1) serves to define the drained (or "jacketed") frame bulk modulus $K^*$ and the unjacketed bulk modulus $K_s$ for the composite frame. Equation (2) defines the jacketed pore modulus $K_p$ and the unjacketed pore modulus $K_\phi$. Similarly, (3) defines the bulk modulus $K_f$ of the pore fluid.

Treating $\delta p_c$ and $\delta p_f$ as the independent variables, we define the dependent variables to be $\delta e \equiv \delta V/V$ and $\delta\zeta \equiv (\delta V_\phi - \delta V_f)/V$, which are termed respectively the total volume dilatation (positive when a sample expands) and the increment of fluid content (positive when the net fluid mass flow is into the sample during deformation). Then, it follows directly from these definitions and from (1), (2), and (3) that

$$\begin{pmatrix} \delta e \\ -\delta\zeta \end{pmatrix} = \begin{pmatrix} 1/K^* & 1/K_s - 1/K^* \\ -\phi/K_p & \phi(1/K_p + 1/K_f - 1/K_\phi) \end{pmatrix} \begin{pmatrix} -\delta p_c \\ -\delta p_f \end{pmatrix}. \tag{4}$$

Now we consider two well-known thought experiments: the drained test and the undrained test (Gassmann 1951; Biot and Willis 1957; Geertsma 1957; Wang, 2000). In the drained test,the porous material is surrounded by an impermeable jacket and the fluid is allowed to escape through a conduit penetrating the jacket. Then, in a long duration experiment, the fluid pressure remains in equilibrium with the external fluid pressure (*e.g.*, atmospheric) and so $\delta p_f = 0$. Hence, $\delta p_c = \delta p_d$. So changes of total volume and pore volume are given by the drained constants $1/K^*$ and $1/K_p$ as defined in (1) and (2). In contrast, for the undrained test, the jacketed sample has no connection to the outside world, so pore pressure responds only to the confining pressure changes. With no way out, the total fluid content cannot change, so the increment $\delta\zeta = 0$. Then, the second equation in (4) shows that

$$0 = -\phi/K_p(\delta p_c - \delta p_f/B), \tag{5}$$

where Skempton's pore pressure buildup coefficient $B$ (Skempton 1954) is defined by

$$B \equiv \left.\frac{\delta p_f}{\delta p_c}\right|_{\delta\zeta=0} = \frac{1}{1 + K_p(1/K_f - 1/K_\phi)}. \tag{6}$$

It follows immediately from this definition that the undrained modulus $K_u$ is determined by [also see Carroll (1980)]

$$K_u = \frac{K^*}{1 - \alpha B}, \tag{7}$$

where $\alpha$ is the combination of moduli known as the Biot-Willis parameter, or the total volume effective-stress coefficient. The precise definition of $\alpha$ follows immediately from the form of (1), by substituting $\delta p_d = \delta p_c - \delta p_f$ and rearranging the equation into the form

$$-\frac{\delta V}{V} = \frac{\delta p_c - \alpha\, \delta p_f}{K^*}, \tag{8}$$

with $\alpha = 1 - K^*/K_s$. The result (7) was apparently first obtained by Gassmann (1951) (though not in this form) for the case of microhomogeneous porous media (*i.e.*, $K_s = K_\phi = K_m$, the bulk modulus of the single mineral present) and by Brown and Korringa (1975) and Rice (1975) for general porous media with multiple minerals as constituents. We will sometimes use the term "Gassmann material" when making reference to a microhomogeneous porous medium.

Next, to clarify the structure of (4) further, note that Betti's reciprocal theorem (Love 1927), shows that the drained and undrained pressures and strains satisfy a reciprocal relation, from which it follows that

$$\frac{1}{K_u} = \frac{1}{K^*} - \frac{\phi B}{K_p}. \tag{9}$$

Comparing (7) with (9), we obtain the general reciprocity relation (Brown and Korringa 1975)

$$\frac{\phi}{K_p} = \frac{\alpha}{K^*}. \tag{10}$$

This reciprocity relation and the form of the compressibility laws (4) also follow directly from general thermodynamic arguments [*e.g.*, Pride and Berryman (1998)]. Then, Skempton's pore-pressure buildup coefficient (Skempton 1954) may be written alternatively as

$$B = \frac{1/K^* - 1/K_s}{1/K^* - 1/K_s + \phi(1/K_f - 1/K_\phi)}. \tag{11}$$

Finally, the condensed form of (4) — incorporating the reciprocity relations — is

$$\begin{pmatrix} \delta e \\ -\delta\zeta \end{pmatrix} = \frac{1}{K^*} \begin{pmatrix} 1 & -\alpha \\ -\alpha & \alpha/B \end{pmatrix} \begin{pmatrix} -\delta p_c \\ -\delta p_f \end{pmatrix}, \tag{12}$$

where the Biot-Willis (1957) parameter $\alpha$ can now be expressed as

$$\alpha = (1 - K^*/K_u)/B. \tag{13}$$

The parameter $\alpha$ is also known as the total volume effective-stress coefficient [see Berryman (1992) for elaboration]. This form of the compressibility laws is especially convenient because all the coefficients are simply related to the three moduli $K^*$, $K_u$, and $B$ that have the clearest physical interpretations. This now completes our review of the standard results concerning the single-porosity compressibility laws.

## DOUBLE-POROSITY GEOMECHANICS

In this section, we present the fundamental governing equations controlling the low-frequency (inertial effects being neglected) response of a double-porosity geomechanical system. See Berryman and Wang (1995) for details left out of the following brief summary.

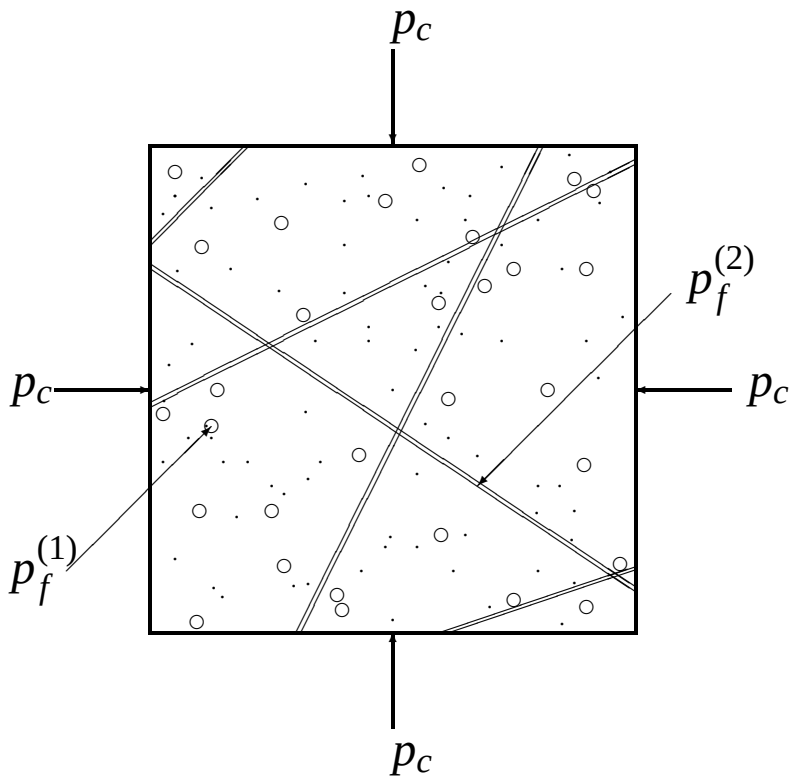



Figure 1: The elements of a double-porosity model are: porous rock matrix intersected by fractures. Three types of macroscopic pressure are pertinent in such a model: external confining pressure $p_c$, internal pressure of the matrix pore fluid $p_f^{(1)}$, and internal pressure of the fracture pore fluid $p_f^{(2)}$.

**Macroscopic Governing Equations**

In the double-porosity formulation, two distinct phases are assumed to exist at the macroscopic level: (1) a porous matrix phase with the effective properties $K^{(1)}$, $K_m^{(1)}$, $\phi^{(1)}$ occupying volume fraction $V^{(1)}/V = v^{(1)}$ of the total volume and (2) a macroscopic crack or joint phase occupying the remaining fraction of the volume $V^{(2)}/V = v^{(2)} = 1 - v^{(1)}$. In earlier work (Berryman and Pride 2002), methods were developed to determine the coefficients of this system within a set of specific modeling assumptions. But the general laws presented in this section are independent of all such modeling assumptions, and the analysis to be presented in later sections is also independent of them as well.

The main difference between the single-porosity and double-porosity formulations is that

we allow the average fluid pressure in the matrix phase to differ from that in the joint phase (thus the term "double porosity") over relatively long time scales. Altogether we have three distinct pressures: confining (external) pressure $\delta p_c$, pore-fluid pressure $\delta p_f^{(1)}$, and joint-fluid pressure $\delta p_f^{(2)}$. (See Figure 1.) Treating $\delta p_c, \delta p_f^{(1)}$, and $\delta p_f^{(2)}$ as the independent variables in the double-porosity theory, we define the dependent variables to be $\delta e \equiv \delta V/V$, $\delta\zeta^{(1)} = (\delta V_\phi^{(1)} - \delta V_f^{(1)})/V$, and $\delta\zeta^{(2)} = (\delta V_\phi^{(2)} - \delta V_f^{(2)})/V$, which are respectively the total volume dilatation, the increment of fluid content in the matrix phase, and the increment of fluid content in the joints. Finally, we assume that the fluid in the matrix is the same kind of fluid as that in the joints.

Linear relations among strain, fluid content, and pressure then take the general form

$$\begin{pmatrix} \delta e \\ -\delta\zeta^{(1)} \\ -\delta\zeta^{(2)} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} -\delta p_c \\ -\delta p_f^{(1)} \\ -\delta p_f^{(2)} \end{pmatrix}. \tag{14}$$

By analogy with the single-porosity result (12), it is easy to see that $a_{12} = a_{21}$ and $a_{13} = a_{31}$. The symmetry of the new off-diagonal coefficients may be demonstrated using Betti's reciprocal theorem in the form

$$\begin{pmatrix} \delta e & -\delta\zeta^{(1)} & -\delta\zeta^{(2)} \end{pmatrix} \begin{pmatrix} 0 \\ -\delta\overline{p}_f^{(1)} \\ 0 \end{pmatrix} = \begin{pmatrix} \delta\overline{e} & -\delta\overline{\zeta}^{(1)} & -\delta\overline{\zeta}^{(2)} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ -\delta p_f^{(2)} \end{pmatrix}, \tag{15}$$

where nonoverlined quantities refer to one experiment and overlined to another experiment to show that

$$\delta\zeta^{(1)}\delta\overline{p}_f^{(1)} = a_{23}\delta p_f^{(2)}\delta\overline{p}_f^{(1)} = a_{32}\delta\overline{p}_f^{(1)}\delta p_f^{(2)} = \delta\overline{\zeta}^{(2)}\delta p_f^{(2)}. \tag{16}$$

Hence, $a_{23} = a_{32}$. Thus, we have established that the matrix in (14) is completely symmetric, so we need to determine only six independent coefficients.

### Constraints on the $a_{ij}$ coming from the long-time limit

Before passing on to the specific models for the various coefficients, we state here several general constraints (independent of any modeling assumptions) on the geomechanical constants $a_{ij}$. Note that in order to measure the $a_{ij}$'s in the laboratory, we need only consider an isolated sample immersed in a "reservoir" characterized by three control parameters: $p_c$, $p_f^{(1)}$, and $p_f^{(2)}$; *i.e.*, gradients in these quantities and the subsequent flow induced by those gradients do not enter the definition of the $a_{ij}$'s.

The constraints are obtained from the limiting case in which the rate at which $p_c$, $p_f^{(1)}$, and $p_f^{(2)}$ are all changing is much slower than the rate at which internal fluid equilibration can take place. In this "long-time limit", we are always in the quasi-static state where

$$p_f^{(1)} = p_f^{(2)}. \tag{17}$$

Left to itself, any system having finite permeability will achieve this state as $t \to \infty$.

**Drained test, long time**

The long-time drained (or "jacketed") test for a double-porosity system should thus correspond to the condition $\delta p_f^{(1)} = \delta p_f^{(2)} = 0$ so that the total volume obeys $\delta e = -a_{11}\delta p_c$. It follows therefore that

$$a_{11} \equiv \frac{1}{K^*}. \tag{18}$$

**Undrained test, long time**

The long-time undrained test for a double-porosity system should also produce the same physical results as a single-porosity system (assuming only that it makes sense at some appropriate larger scale to view the medium as homogeneous). The conditions for this test are that

$$\delta p_f^{(1)} = \delta p_f^{(2)} = \delta p_f, \delta\zeta \equiv \delta\zeta^{(1)} + \delta\zeta^{(2)} = 0, \tag{19}$$

from which follow

$$\delta e = -a_{11}\delta p_c - (a_{12} + a_{13})\delta p_f, 0 = -(a_{21} + a_{31})\delta p_c - (a_{22} + 2a_{23} + a_{33})\delta p_f. \tag{20}$$

These require that the overall pore-pressure buildup coefficient be given by

$$B \equiv \left.\frac{\partial p_f}{\partial p_c}\right|_{\delta\zeta=0} = -\frac{a_{21} + a_{31}}{a_{22} + 2a_{23} + a_{33}}, \tag{21}$$

and that the undrained bulk modulus be given by

$$\frac{1}{K_u} \equiv \left.\frac{\delta e}{\delta p_c}\right|_{\delta\zeta=0} = a_{11} + (a_{12} + a_{13})B. \tag{22}$$

**Fluid injection test, long time**

The conditions required to measure the three-dimensional storage coefficient $R$ in the long-time limit are that $\delta p_f^{(1)} = \delta p_f^{(2)} = \delta p_f$, while $\delta p_c = 0$. It follows therefore from (4) and (23) that

$$R \equiv \left.\frac{\partial\zeta}{\partial p_f}\right|_{\delta p_c=0} = a_{22} + 2a_{23} + a_{33} = \frac{\alpha}{K^*} + \phi\left(\frac{1}{K_f} - \frac{1}{K_\phi}\right). \tag{23}$$

**Generalized Biot-Willis parameters**

Equation (18) has already determined the coefficient $a_{11}$. Thus, (22) shows that

$$a_{12} + a_{13} = -\frac{1/K^* - 1/K_u}{B} = -\alpha/K^*. \tag{24}$$

This relation provides a constraint on the sum of the two generalized Biot-Willis parameters for the double-porosity problem.

Not all of these long-time results are independent. In fact, there are only three independent equations among the five given above expressing the $a_{ij}$ in terms of the single-porosity (long-time) moduli.

## DOUBLE-POROSITY THOUGHT EXPERIMENT

Several of the main results obtained previously can be derived in a more elegant fashion by using a new self-similar (uniform expansion) thought experiment. The basic idea we are going to introduce here is analogous to, but nevertheless distinct from, other thought experiments used in thermoelasticity by Cribb (1968) and in single-porosity poroelasticity by Berryman and Milton (1991) and Berryman and Pride (1998). Cribb's method provided an independent and simpler derivation of Levin's (1967) results on thermoelastic expansion coefficients. The present results also provide an independent and simpler derivation of results obtained recently by Berryman and Pride (2002) for the double-porosity coefficients. Related methods in micromechanics are sometimes called "the method of uniform fields" by some authors (Dvorak and Benveniste, 1997).

We have already shown that $a_{11} = 1/K^*$. We will now show how to determine the remaining five constants in the case of a binary composite system, such as that illustrated in Figure 2. The components of the system are themselves porous materials 1 and 2, but each is assumed to be what we call a "Gassmann material" satisfying [in analogy to equation (12)]

$$\begin{pmatrix} \delta e^{(1)} \\ -\delta\zeta^{(1)}/v^{(1)} \end{pmatrix} = \frac{1}{K^{(1)}} \begin{pmatrix} 1 & -\alpha^{(1)} \\ -\alpha^{(1)} & \alpha^{(1)}/B^{(1)} \end{pmatrix} \begin{pmatrix} -\delta p_c^{(1)} \\ -\delta p_f^{(1)} \end{pmatrix} \tag{25}$$

for material 1 and a similar expression for material 2. The new constants appearing on the right are the drained bulk modulus $K^{(1)}$ of material 1, the corresponding Biot-Willis parameter $\alpha^{(1)}$, and the Skempton coefficient $B^{(1)}$. The volume fraction $v^{(1)}$ appears here to correct for the difference between a global fluid content and the corresponding local variable for material 1. The main special characteristic of a Gassmann porous material is that it is composed of only one type of solid constituent, so it is "microhomogeneous" in its solid component, and in addition the porosity is randomly, but fairly uniformly, distributed so there is a well-defined constant porosity $\phi^{(1)}$ associated with material 1, etc.

For our new thought experiment, we ask the question: Is it possible to find combinations of $\delta p_c = \delta p_c^{(1)} = \delta p_c^{(2)}$, $\delta p_f^{(1)}$, and $\delta p_f^{(2)}$ such that the expansion or contraction of the system is spatially uniform or self-similar? This is the same as asking if we can find uniform confining pressure $\delta p_c$, and pore-fluid pressures $\delta p_f^{(1)}$ and $\delta p_f^{(2)}$, such that

$$\delta e = \delta e^{(1)} = \delta e^{(2)}. \tag{26}$$

If these conditions can all be met simultaneously, then results for system constants can be obtained purely algebraically without ever having to solve the equilibrium equations for nonconstant stress and strain. We have initially set $\delta p_c = \delta p_c^{(1)} = \delta p_c^{(2)}$, as the condition of uniform confining pressure is clearly necessary for this self-similar thought experiment to achieve a valid solution of the equilibrium equations.

So, the first condition to be considered is the equality of the strains of the two constituents:

$$\delta e^{(1)} = -\frac{1}{K^{(1)}}(\delta p_c - \alpha^{(1)}\delta p_f^{(1)}) = \delta e^{(2)} = -\frac{1}{K^{(2)}}(\delta p_c - \alpha^{(2)}\delta p_f^{(2)}). \tag{27}$$

If this condition can be satisfied, then the two constituents are expanding or contracting at the same rate and it is clear that self-similarity will prevail. If we imagine that $\delta p_c$ and $\delta p_f^{(1)}$ have been chosen, then we only need to choose an appropriate value of $\delta p_f^{(2)}$, so that (27) is satisfied. This requires that

$$\delta p_f^{(2)} = \delta p_f^{(2)}(\delta p_c, \delta p_f^{(1)}) = \frac{1 - K^{(2)}/K^{(1)}}{\alpha^{(2)}}\delta p_c + \frac{\alpha^{(1)}K^{(2)}}{\alpha^{(2)}K^{(1)}}\delta p_f^{(1)}, \tag{28}$$

which shows that, except for some very special choices of the material parameters (such as $\alpha^{(2)} = 0$), $\delta p_f^{(2)}$ can in fact always be chosen so the uniform expansion takes place. (We are not considering long-term effects here. Clearly, if the pressures are left to themselves, they will tend to equilibrate over time so that $\delta p_f^{(1)} = \delta p_f^{(2)}$. We are considering only the "instantaneous" behavior of the material permitted by our system of equations and finding what internal consistency of this system of equations implies must be true.)

Using formula (28), we can now eliminate $\delta p_f^{(2)}$ from the remaining equality so that

$$\delta e = -\left[a_{11}\delta p_c + a_{12}\delta p_f^{(1)} + a_{13}\delta p_f^{(2)}(\delta p_c, \delta p_f^{(1)})\right] = \delta e^{(1)} = -\frac{1}{K^{(1)}}(\delta p_c - \alpha^{(1)}\delta p_f^{(1)}), \tag{29}$$

where $\delta p_f^{(2)}(\delta p_c, \delta p_f^{(1)})$ is given by (28). Making the substitution and then noting that $\delta p_c$ and $\delta p_f^{(1)}$ were chosen independently and arbitrarily, we see that the resulting coefficients of these two variables must each vanish. The equations we obtain in this way are

$$a_{11} + a_{13}(1 - K^{(2)}/K^{(1)})/\alpha^{(2)} = 1/K^{(1)} \tag{30}$$

and

$$a_{12} + a_{13}(\alpha^{(1)}K^{(2)}/\alpha^{(2)}K^{(1)}) = -\alpha^{(1)}/K^{(1)}. \tag{31}$$

Since $a_{11}$ is known, equation (30) can be solved directly for $a_{13}$, giving

$$a_{13} = -\frac{\alpha^{(2)}}{K^{(2)}}\frac{1 - K^{(1)}/K^*}{1 - K^{(1)}/K^{(2)}} \tag{32}$$

Similarly, since $a_{13}$ is now known, substituting into (31) gives

$$a_{12} = -\frac{\alpha^{(1)}}{K^{(1)}}\frac{1 - K^{(2)}/K^*}{1 - K^{(2)}/K^{(1)}}. \tag{33}$$

Thus, three of the six coefficients have been determined.

To evaluate the remaining three coefficients, we must consider what happens to the fluid increments during the same self-similar expansion thought experiment. We will treat only

material 1, but the equations for material 2 are completely analogous. >From the preceding equations, it follows that

$$\delta\zeta^{(1)} = a_{12}\delta p_c + a_{22}\delta p_f^{(1)} + a_{23}\delta p_f^{(2)}(\delta p_c, \delta p_f^{(1)}) = \frac{v^{(1)}}{K^{(1)}}\left[-\alpha^{(1)}\delta p_c + (\alpha^{(1)}/B^{(1)})\delta p_f^{(1)}\right]. \quad (34)$$

Again substituting for $\delta p_f^{(2)}(\delta p_c, \delta p_f^{(1)})$ from (28) and noting once more that the resulting equation contains arbitrary values of $\delta p_c$ and $\delta p_f^{(1)}$, so that the coefficients of these terms must vanish separately, gives two equations

$$a_{12} + a_{23}(1 - K^{(2)}/K^{(1)})/\alpha^{(2)} = -\alpha^{(1)}v^{(1)}/K^{(1)}, \quad (35)$$

and

$$a_{22} + a_{23}\left(\alpha^{(1)}K^{(2)}/\alpha^{(2)}K^{(1)}\right) = \alpha^{(1)}v^{(1)}/B^{(1)}K^{(1)}. \quad (36)$$

Solving these equations in sequence as before, we obtain

$$a_{23} = \frac{K^{(1)}K^{(2)}\alpha^{(1)}\alpha^{(2)}}{(K^{(2)} - K^{(1)})^2}\left[\frac{v^{(1)}}{K^{(1)}} + \frac{v^{(2)}}{K^{(2)}} - \frac{1}{K^*}\right], \quad (37)$$

and

$$a_{22} = \frac{v^{(1)}\alpha^{(1)}}{B^{(1)}K^{(1)}} - \left(\frac{\alpha^{(1)}}{1 - K^{(1)}/K^{(2)}}\right)^2\left[\frac{v^{(1)}}{K^{(1)}} + \frac{v^{(2)}}{K^{(2)}} - \frac{1}{K^*}\right]. \quad (38)$$

Performing the corresponding calculation for $\delta\zeta^{(2)}$ produces formulas for $a_{32}$ and $a_{33}$. Since the formula in (37) is already symmetric in the component indices, the formula for $a_{32}$ provides nothing new. The formula for $a_{33}$ is easily seen to be identical in form to $a_{22}$, but with the 1 and 2 indices interchanged everywhere.

This completes the derivation of all five of the needed coefficients of double porosity for the two constituent model.

These results can now be used to show how the constituent properties $K$, $\alpha$, $B$ average at the macrolevel for a two-constituent composite. We find

$$\begin{aligned}\alpha &= -\frac{a_{12} + a_{13}}{a_{11}} \\ &= \frac{\alpha^{(1)}(K^* - K^{(2)}) + \alpha^{(2)}(K^{(1)} - K^*)}{K^{(1)} - K^{(2)}},\end{aligned} \quad (39)$$

and

$$\begin{aligned}\frac{1}{B} &= -\frac{a_{22} + 2a_{23} + a_{33}}{a_{12} + a_{13}} \\ &= \frac{K^*}{\alpha}\left(\frac{v^{(1)}\alpha^{(1)}}{B^{(1)}K^{(1)}} + \frac{v^{(2)}\alpha^{(2)}}{B^{(2)}K^{(2)}} - \left(\frac{\alpha^{(1)}K^{(2)} - \alpha^{(2)}K^{(1)}}{K^{(2)} - K^{(1)}}\right)^2\left[\frac{v^{(1)}}{K^{(1)}} + \frac{v^{(2)}}{K^{(2)}} - \frac{1}{K^*}\right]\right).\end{aligned} \quad (40)$$

It should also be clear that parts of the preceding analysis generalize easily to the multiporosity problem. We discuss some of these remaining issues in the final section.

**Example**

To illustrate the use of the formulas derived for the coefficients of the double-porosity system, we will now compute and plot the coefficients for a realistic system. We will use data of Coyner (1984) for Navajo sandstone, and modify it somewhat to produce a plot that will highlight the results obtained from the equations. The first problem we encounter in doing so is that, although we can make reasonable direct estimates of the bulk and shear moduli of the constituents, we also must have an estimate of the overall bulk modulus $K^*$ of the composite double-porosity medium. And more than that, we need it as a function of the volume fractions of the two constituents. Our analysis has assumed that $K^*$ was given or measured independently. For present purposes, it is sensible to use an effective medium theory such as the symmetric self-consistent method [or CPA = coherent potential approximation — see Berryman and Berge (1996) for a discussion and references therein for elaboration] to estimate $K^*$. The CPA has the advantage that it treats both constituents equally (*i.e.*, symmetrically) and therefore does *not* assume that one constituent always surrounds the other — so there is no host material [see Berge *et al.* (1993) for further discussion]. With this addition to the theory, we can proceed to the calculations.

TABLE 1. Input parameters for a Navajo sandstone model of double-porosity system. Bulk moduli $K$ have units of GPa. Poisson's ratio $\nu$ and porosity $\phi$ are dimensionless.

| $K_s$ | $K_s^{(1)}$ | $K^{(1)}$ | $\nu^{(1)}$ | $\phi^{(1)}$ | $K_s^{(2)}$ | $K^{(2)}$ | $\nu^{(2)}$ | $\phi^{(2)}$ |
|---|---|---|---|---|---|---|---|---|
| 34.0 | 34.5 | 16.5 | 0.15 | 0.118 | 34.5 | 1.65 | 0.017 | 0.354 |

The parameters used for Navajo sandstone are listed in Table 1. Although Poisson's ratio $\nu$ does not appear explicitly in the equations here, it is required in the CPA (or any but the most elementary) effective medium calculation for the overall bulk modulus $K^*$. The results are shown in Figure 3.

Note that the off-diagonal coefficient $a_{23}$, which couples the fluid in the storage porosity to the fluid in the transport porosity, is very close to zero for all values of storage material volume fraction. This behavior has been observed previously (Berryman and Wang 1995), and is believed to be a strong indication that the double-porosity approach is appropriate for the system studied. If this coefficient is not small, then the fluids in the two types of porosity are strongly coupled and therefore should not be treated as a double-porosity system.

The behavior of the other coefficients is as one would expect: All the coefficients for the transport porosity tend to vanish as the volume fraction of this phase vanishes, and the medium again reduces to a single-porosity system in this limit.

## DISCUSSION OF MULTI-POROSITY SYSTEMS

Micromechanical analysis provides definite answers to the question of how the coefficients in double-porosity systems are to be computed from knowledge of the constituents' properties.

The question then naturally arises whether this analysis can be generalized to multi-porosity systems. Certainly, multi-porosity systems are the ones most likely to represent realistic systems occurring in nature, for example, oil and gas reservoirs. And, therefore, we need to address these issues. Transport in triple-porosity and multi-porosity systems have already been studied by some authors (Bai, Elsworth, and Roegiers 1993b; Bai and Roegiers, 1997), hence, it is timely to consider the geomechanical aspects of these problems. We will set up the problem and describe its general characteristics here, but the full solution will be left to future work.

The resulting coefficient matrices will clearly take a form analogous the ones already studied. For example, in a triple-porosity system, the macroscopic governing equations are:

$$\begin{pmatrix} \delta e \\ -\delta\zeta^{(1)} \\ -\delta\zeta^{(2)} \\ -\delta\zeta^{(3)} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} -\delta p_c \\ -\delta p_f^{(1)} \\ -\delta p_f^{(2)} \\ -\delta p_f^{(3)} \end{pmatrix} \tag{41}$$

The meanings of all the coefficients follow immediately from the discussion of (14). The matrix is again symmetric, so there are four diagonal and six off-diagonal coefficients to be determined, for a total of ten unique coefficients. The leading coefficient $a_{11} = 1/K^*$ as before, but the remaining coefficients require further analysis.

In general, for an $N$-porosity system of the form considered here, the total number of coefficients to be determined in the $(N+1)\times(N+1)$ system of equations is $N+1$ diagonal and $N(N+1)/2$ unique off-diagonal coefficients, for a total of $G = (N+1)(N+2)/2$ coefficients. And the nature of $a_{11}$ remains unchanged for any $N$. If we assume that each of the unique porosities can be associated with a Gassmann (microhomogeneous) material, then we have equations of the same form as (25) for each of these constituents, and therefore three mechanical coefficients plus the porosity of each constituent is assumed to be known, at least approximately, in order for this analysis to proceed. The uniform expansion/contraction scenario carries over to the multi-porosity system, but does not supply enough equations to close the system by itself for $N > 2$. To see this, note that once $\delta p_c$ and $\delta p_f^{(1)}$ are chosen, then all the remaining $\delta p_f$'s are determined by the uniform strain condition and Gassmann's relations. Then, substituting these values into the multi-porosity system [*e.g.*, (41)], we see there are always two equations for each row of the matrix. This results in $S = 2(N+1)$ equations just from this self-similar thought experiment. These two sets of numbers are compared in Table 2. In addition to these equations, we always have the 3 conditions from the long-time limits, and we can also find other equations as needed by considering other experiments on the system [*e.g.*, see Berryman and Wang (1995)]. However, it is important to remember that it is the number of linearly independent equations that is pertinent, and determining this number has so far not proven to be a easy task for the general case. At the present writing, closure of the system of equations for the multi-porosity coefficients when $N > 2$ is an open question.

The analysis presented here has been strictly for isotropic constituents, and an isotropic overall multi-porosity system. Generalization to anisotropic systems is both possible and desirable, but the analysis obviously becomes more complex because of the proliferation of coefficients that results.

TABLE 2. Growth of the number $G = (N+1)(N+2)/2$ of geomechanical coefficients and the number $S = 2(N+1)$ of equations from the self-similar thought experiment as the number $N$ of distinct porosities within the system increases.

| $N$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $G$ | 3 | 6 | 10 | 15 |
| $S$ | – | 6 | 8 | 10 |

## CONCLUSIONS AND NEW DIRECTIONS

The preceding results show how a micromechanical analysis based on poroelasticity and Gassmann's equations can be used to compute the geomechanical double-porosity coefficients in a very elegant manner. This makes use of all the information available and produces reasonable estimates of all the coefficients needed in reservoirs modeled by double-porosity geomechanics. Triple- and multi-porosity geomechanics can also be studied using similar methods, but some work remains to be done on closure of the increasingly larger systems of equations involved. For multi-porosity systems, closure of the system of equations can nevertheless always be achieved by the addition of further macroscale measurements. Analysis and solution of these systems of equations to eliminate the need for such additional measurements is therefore one subject of future work in this area of research.

Extension of this work in other directions is also possible. In particular, the applications presented here have been restricted for the sake simplicity to isotropic macroscopic systems. But it is known that the methods employed are not restricted to isotropic systems — as has already been shown in other micromechanical studies by Dvorak and Benveniste (1997). So careful extensions of these ideas to anisotropy, and especially anisotropy due to oriented fractures, in double-porosity systems is both possible and desirable. Such extensions will permit us to provide more realistic models of reservoir geomechanics, including effects of overburden, tectonic stresses, hydrofracing, etc.

## ACKNOWLEDGMENTS

I thank Steve Pride and Herb Wang for their insightful collaborations on the phenomenology and micromechanics of double-porosity systems. I thank Patricia A. Berge for helpful comments that improved the manuscript.

## REFERENCES

Bai, M. (1999). "On equivalence of dual-porosity poroelastic parameters," *J. Geophys. Res.* **104**, 10461–10466.

Bai, M., Elsworth, D., and Roegiers, J.-C. (1993a). "Modeling of naturally fractured reservoirs using deformation dependent flow mechanism," *Int. J. Rock Mech. Min. Sci. & Geomech. Abstr.* **30**, 1185–1191.

Bai, M., Elsworth, D., and Roegiers, J.-C. (1993b). "Multiporosity/multipermeability approach to the simulation of naturally fractured reservoirs," *Water Resources Res.* **29**, 1621–1633.

Bai, M., Meng, F., Elsworth, D., Abousleiman, Y., and Roegiers, J.-C. (1999). "Numerical modelling of coupled flow and deformation in fractured rock specimens," *Int. J. Num. Anal. Meth. Geomech.* **23**, 141–160.

Bai, M., and Roegiers, J.-C. (1997). "Triple-porosity analysis of solute transport," *J. Contaminant Hydrol.* **28**, 247–266.

Barenblatt, G. I., and Zheltov, Yu. P. (1960). "Fundamental equations of filtration of homogeneous liquids in fissured rocks," *Sov. Phys. Doklady* **5**, 522–525 [English translation of: *Doklady Akademii Nauk SSSR* **132**, 545–548].

Berge, P. A., Berryman, J. G., and Bonner, B. P. (1993). "Influence of microstructure on rock elastic properties," *Geophys. Res. Lett.* **20**, 2619–2622.

Berryman, J. G. (1992). " Effective stress for transport properties of inhomogeneous porous rock," *J. Geophys. Res.* **97**, 17409–17424.

Berryman, J. G. (1995). "Mixture theories for rock properties," *Rock Physics and Phase Relations*, edited by T. J. Ahrens, American Geophysical Union, Washington, D. C., pp. 205–228.

Berryman, J. G. (1999). "Origin of Gassmann's equations," *Geophysics* **64**, 1627–1629.

Berryman, J. G., and Berge, P. A. (1996). "Critique of explicit schemes for estimating elastic properties of multiphase composites," *Mech. Materials* **22**, 149–164.

Berryman, J. G., and Milton, G. W. (1991). "Exact results for generalized Gassmann's equations in composite porous media with two constituents," *Geophysics* **56**, 1950–1960.

Berryman, J. G., and Pride, S. R. (1998). "Volume averaging, effective stress rules, and inversion for microstructural response of multicomponent porous media," *Int. J. Solids Struct.* **35**, 4811–4843.

Berryman, J. G., and Pride, S. R. (2002). "Models for computing geomechanical constants of double-porosity materials from the constituents' properties," *J. Geophys. Res.* , accepted for publication.

Berryman, J. G., and Wang, H. F. (1995). "The elastic coefficients of double-porosity models for fluid transport in jointed rock," *J. Geophys. Res.* **100**, 24611–24627.

Berryman, J. G., and Wang, H. F. (2000). "Elastic wave propagation and attenuation in a double-porosity dual-permeability medium," *Int. J. Rock Mech.* **37**, 63–78.

Berryman, J. G., and Wang, H. F. (2001). "Dispersion in poroelastic systems," *Phys. Rev. E* **64**, 011303-1–011303-16.

Biot, M. A. (1941). "General theory of three dimensional consolidation," *J. Appl. Phys.* **12**, 155–164.

Biot, M. A., and Willis, D. G. (1957). "The elastic coefficients of the theory of consolidation," *J. App. Mech.* **24**, 594–601.

Bourbié, T., Coussy, O., and Zinszner, B. (1987). *Acoustics of Porous Media*, Gulf Publishing Company.

Brandt, H. (1955). "A study of the speed of sound in porous granular media," *J. Appl. Mech.* **22**, 479–486.

Brown, R. J. S., and Korringa, J. (1975). "On the dependence of the elastic properties of a porous rock on the compressibility of a pore fluid," *Geophysics* **40**, 608–616.

Carroll, M. M. (1980). "Mechanical response of fluid-saturated porous materials," *Theoretical and Applied Mechanics*, F. P. J. Rimrott and B. Tabarrok (eds.), Proceedings of the 15th International Congress of Theoretical and Applied Mechanics, Toronto, August 17–23, 1980, North-Holland, Amsterdam, pp. 251–262.

Cho, T. F., Plesha, M. E., and Haimson, B. C. (1991). "Continuum modeling of jointed porous rock," *Int. J. Num. Anal. Meth. Geomech.* **15**, 333–353.

Cleary, M. P. (1977). "Fundamental solutions for a fluid-saturated porous solid," *Int. J. Solids Structures* **13**, 785–806.

Coyner, K. B. (1984). *Effects of Stress, Pore Pressure, and Pore Fluids on Bulk Strain, Velocity, and Permeability of Rocks*, Ph. D. Thesis, Massachusetts Institute of Technology.

Detournay, E., and Cheng A. H.-D. (1993). "Fundamentals of poroelasticity," *Comprehensive Rock Engineering*, edited by J. A. Hudson, Vol. 2, Chapter 5, Pergamen Press, Oxford.

Dvorak, G. J., and Benveniste, Y. (1997). "On micromechanics of inelastic and piezoelectric composites," in *Theoretical and Applied Mechanics*, T. Tatsumi, E. Watanabe, and T. Kambe (eds.), Elsevier, Amsterdam, The Netherlands.

Elsworth, D., and Bai, M. (1992). "Flow-deformation response of dual-porosity media," *ASCE J. Geotech. Engng.*, **118**, 107–124.

Fatt, I. (1958). "Pore volume compressibilities of sandstone reservoir rocks," *Soc. Petroleum Engineers Trans., AIME* **213**, 362–364.

Fatt, I. (1959). "The Biot-Willis elastic coefficients for a sandstone," *ASME J. Appl. Mech.* **26**, 296–297.

Gassmann, F. (1951). "Über die elastizität poröser medien," *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* **96**, 1–23.

Geertsma, J. (1966).“Problems of rock mechanics in petroleum production engineering,” *Proceedings of the 1st Congress of the International Society of Rock Mechanics* **1**, Lisbon, pp. 585–594.

Geertsma, J. (1973). “Land subsidence above compacting oil and gas reservoirs,” *J. Petroleum Tech.* **25**. 734–744.

Gomberg, J., and Wolf, L. (1999). “Possible cause for an improbable earthquake: The 1997 $M_w$ 4.9 southern Alabama earthquake and hydrocarbon recovery,” *Geology* **27**, 367–370.

Khaled, M. Y., Beskos, D. E., and E. C. Aifantis, E. C. (1984). “On the theory of consolidation with double porosity. 3. A finite-element formulation,” *Int. J. Num. Anal. Meth. Geomech.* **8**, 101–123.

Kovach, R. L. (1974). “Source mechanisms for Wilmington oil field subsidence earthquakes,” *Bull. Seism. Soc. Am.* **64**, 699–711.

Levin, V. M. (1967). “Thermal expansion coefficients of heterogeneous materials,” *Mech. Solids* **2**, 58–61.

Lewallen, K. T., and Wang,, H. F. (1998). “Consolidation of a double-porosity medium,” *Int. J. Solids Structures* **35**, 4845–4867.

Li, X., Zhong, L., and Pyrak-Nolte, L. J. (2001). “Physics of partially saturated porous media: Residual saturation and seismic-wave propagation,” *Annual Rev. Earth Planet. Sci.* **29**, 419–460.

Love, A. E. H.(1927). *A Treatise on the Mathematical Theory of Elasticity*, Dover, New York, pp. 173–174.

Nilson, R. H., and K. H. Lie, K. H. (1990). “Double-porosity modelling of oscillatory gas motion and contaminant transport in a fractured porous medium,” *Int. J. Num. Anal. Meth. Geomech.* **14**, 565–585.

Norris, A. N. (1992). “On the correspondence between poroelasticity and thermoelasticity,” *J. Appl. Phys.* **71**, 1138–1141.

Nur, A., and Byerlee, J. D. (1971). “An exact effective stress law for elastic deformation of rock with fluids,” *J. Geophys. Res.* **76**, 6414–6419.

Pennington, W.D. (2001). “Reservoir geophysics,” *Geophysics* **66**, 25–30.

Pennington, W.D., Davis, S. D., Carlson, S. M., DuPree, J., and Ewing, T. E. (1986). “The evolution of seismic barriers and asperities caused by the depressuring of fault planes in oil and gas fields of south Texas, *Bull. Seism. Soc. Am.* **76**, 939–948.

Pride, S. R., and Berryman, J. G. (1998). “Connecting theory to experiment in poroelasticity,” *J. Mech. Phys. Solids* **46**, 719–747.

Rice, J. R. (1975). "On the stability of dilatant hardening for saturated rock masses," *J. Geophys. Res.* **80**, 1531–1536.

Rice, J. R., and Cleary, M. P. (1976). "Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents," *Rev. Geophys. Space Phys.* **14**, 227–241.

Schopper, J. R. (1982). "Porosity and permeability," in *Physical Properties of Rocks*, Vol. 1a, G. Angenheister (ed.), Springer, Berlin, pp. 184–266.

Segall, P. (1985). "Stress and subsidence resulting from subsurface fluid withdrawal in the epicentral region of the 1983 Coalinga earthquake," *J. Geophys. Res.* **90**, 6801–6816.

Segall, P. (1989). "Earthquakes triggered by fluid extraction," *Geology* **17**, 942–946.

Segall, P. (1992). "Induced stresses due to fluid extraction from axisymmetric reservoirs," *PAGEOPH* **139**, 535–560.

Segall, P., and Fitzgerald, S. D. (1998). "A note on induced stress changes in hydrocarbon and geothermal reservoirs," *Tectonophysics* **289**, 117–128.

Skempton, A. W. (1954). "The pore-pressure coefficients A and B," *Geotechnique* **4**, 143–147.

Terzaghi, K. (1925). *Erdbaumechanik auf bodenphysikalischer Grundlage*, Deuticke, Leipzig.

Tuncay, K., and Corapciaglu, M. Y. (1995). "Effective stress principle for saturated fractured porous media," *Water Resources Res.* **31**, 3103–3106.

Wang, H. F. (1993). "Quasi-static poroelastic parameters in rock and their geophysical applications," *PAGEOPH* **141**, 269–286.

Wang, H. F. (2000). *Theory of Linear Poroelasticity with Applications to Geomechanics and Hydrogeology*, Princeton University Press, Princeton, New Jersey.

Wang, H. F., and Berryman, J. G. (1996). "On constitutive equations and effective stress for deformable, double porosity media," *Water Resources Res.* **32**, 3621–3622.

Warren, J. E., and Root, P. J. (1963). "The behavior of naturally fractured reservoirs," *Soc. Pet. Eng. J.* **3**, 245–255.

Wawersik, W. R., Orr, F. M., Rudnicki, J. W., Ortoleva, P. J., Dove, P., Richter, F., Harris, J., Warpinski, N. R., Logan, J. M., Wilson, J. L., Pyrak-Nolte, L., and Wong, T.-F. (2001). "Terrestrial sequestration of $CO_2$: An assessment of research needs," *Adv. Geophys.* **43**, 97–177.

Wilson, R. K., and Aifantis, E. C. (1982). "On the theory of consolidation with double porosity," *Int. J. Engng. Sci.* **20**, 1009–1035.

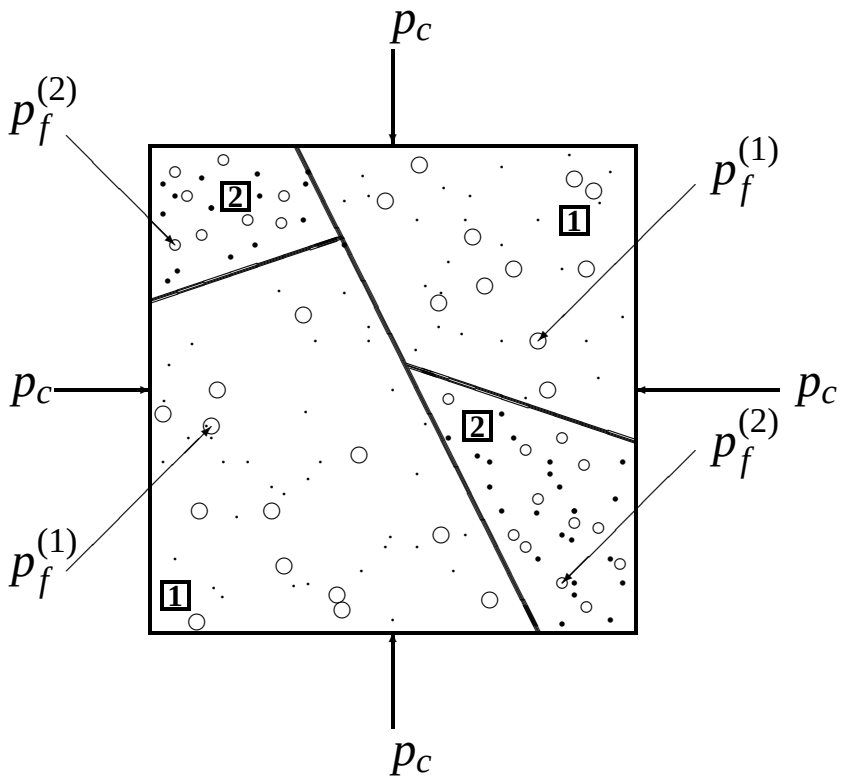



Figure 2: A composite porous medium is composed of two distinct types of porous solid $(1, 2)$. In the model illustrated here and treated in the text, the two types of materials are well-bonded but themselves have very different porosity types, one being a storage porosity (type-1) and the other (type-2) being a transport porosity (and therefore fracture-like, or tube-like as illustrated in cross-section in this diagram).

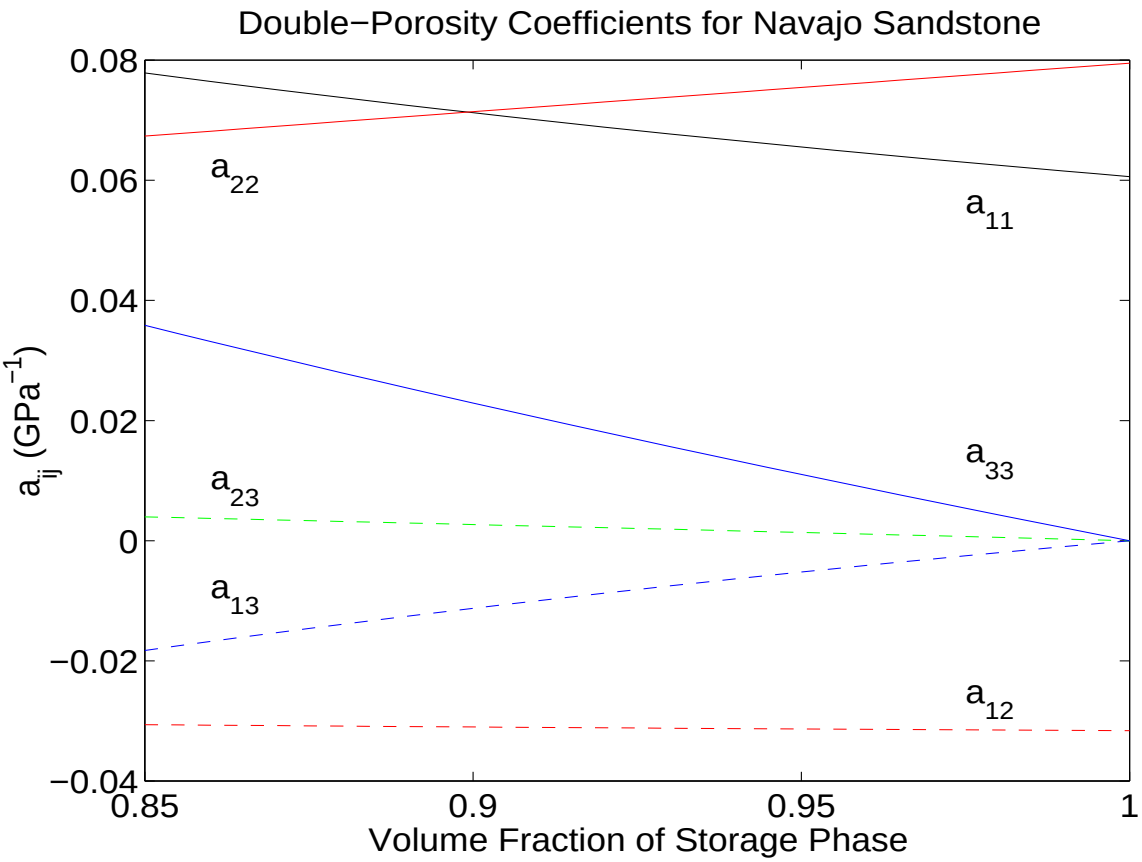



Figure 3: Values of the double-porosity coefficients $a_{ij}$ for a system similar to Navajo sandstone. Values used for the input parameters are listed in Table 1. jim1-navajo_good [NR]

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 287–291*

# Short Note

# Amplitude inversion for three reflectivities

*Marie L. Prucha and E. F. Herkenhoff*[1]

## INTRODUCTION

Seismic amplitudes can provide information about rock properties if the amplitudes can be properly processed (Backus, 1983; Castagna and Backus, 1993). This is very difficult to do given the nonlinearity of the Zoeppritz equation. The linearized Bortfeld approximation (Bortfeld, 1961) allows us to invert the data to obtain the reflectivities of events. Unfortunately, most forms of the Bortfeld approximation involve the primary and shear wave interval velocities which are difficult to obtain without knowing the rock properties. This paper will first look at the basic form of the Bortfeld approximation then explain a new form that eliminates the need for the interval velocities. Finally, the new form will be used on a simple synthetic.

## BORTFELD'S 3 TERM REFLECTIVITY EQUATION

Bortfeld's three term reflectivity equation is a linearized form of the Zoeppritz equation. It predicts the amplitude ($R$) at various reflection angles ($\theta$) given three reflectivity terms. These three terms are the zero-offset reflectivity ($R_O$), the P-wave reflectivity ($R_P$), and a gradient term ($R_{sh}$). This is its basic form:

$$R(\theta_i) \;=\; R_O + R_{sh}\sin^2(\theta_i) + R_P\tan^2(\theta_i)\sin^2(\theta_i) \tag{1}$$

where

$$R_P \;=\; \frac{\Delta V_p}{2V_p} \qquad R_\rho \;=\; \frac{\Delta\rho}{2\rho} \qquad R_O \;=\; R_P + R_\rho \tag{2}$$
$$R_{sh} = \frac{1}{2}\left(\frac{\Delta V_p}{V_p} - k\frac{\Delta\rho}{\rho} - 2k\frac{\Delta V_s}{V_s}\right) \qquad k \;=\; \left(\frac{2V_s}{V_p}\right)^2$$

This form is easily obtained algebraically from the form derived in Aki and Richards (1980). Both forms are well known and frequently used, but may not be accurate in areas where the interval velocities are not well known. Therefore, we choose to look at a different form.

[1]**email:** marie @ sep.Stanford.EDU,efhe @ chevron.com

### The Stack-Constrained form

The basic form just described depends on a $\frac{V_s}{V_p}$ ratio, meaning that to get the information on the rock properties, you first need a fairly accurate interval velocity model. Fred Herkenhoff of Chevron realized that by substituting the value of a stacked trace into the problem, he could constrain the result to have values similar to those of the stack and remove the dependence on the velocity ratio. The stack amplitude can be calculated from the basic form:

$$S \;=\; R_O + R_{sh}\sin^2(\theta_{S1}) + R_P\tan^2(\theta_{S2})\sin^2(\theta_{S1}) \tag{3}$$

Here, $\sin^2\theta_{S1}$ and $\tan^2\theta_{S2}$ are the averages of $\sin^2\theta$ and $\tan^2\theta$ over the range of input angles. This stack equation is then used as a substitute for the $\frac{V_s}{V_p}$ ratio:

$$R(\theta_i) - S\frac{\sin^2(\theta_i)}{\sin^2(\theta_{S1})} \;=\; R_O(1-\frac{\sin^2(\theta_i)}{\sin^2(\theta_{S1})}) + R_P(\tan^2(\theta_i) - \frac{\tan^2(\theta_i)\sin^2(\theta_i)}{\sin^2(\theta_{S1})}) \tag{4}$$

This form can now be inverted for the zero-offset reflectivity ($R_O$) and the P-wave reflectivity ($R_P$) without needing the interval velocities. Once those reflectivities are obtained, the stack equation (Eqn. 3) can be used to find the gradient term. This inversion is demonstrated in the next section.

## RESULTS

At this time I am testing this formulation on synthetic data. I created a ten trace angle gather (Figure 1) using the basic form of Bortfeld's equations. Figure 1 also contains the stacked trace of this angle gather. There are five events with different reflectivities. The results of the inversion are in Figure 2 and are fairly accurate. The answers are:

| Event | | Correct $R_O$ | Est. $R_O$ | | Correct $R_{sh}$ | Est. $R_{sh}$ | | Correct $R_P$ | Est. $R_P$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | 0.023 | 0.0230 | | 0.0 | 0.00139 | | 0.023 | 0.0257 |
| 2 | | 0.035 | 0.0350 | | -0.01 | -0.0132 | | 0.023 | 0.0251 |
| 3 | | 0.01 | 0.0100 | | 0.01 | 0.00572 | | 0.03 | 0.0342 |
| 4 | | -0.03 | -0.0300 | | 0.0 | -0.00476 | | 0.03 | 0.0367 |
| 5 | | 0.02 | 0.0200 | | -0.02 | -0.0175 | | -0.02 | -0.0229 |

The solutions for $R_O$ have been found correctly. There is some error in the solution for $R_P$ and this in turn has produced some error in the calculated values for $R_{sh}$. However, the errors are small so it seems that this form can be successfully inverted. The fact that it is not dependent on interval velocities makes it preferable to other forms of the Bortfeld approximation.

## ACKNOWLEDGMENTS

I would like to thank Chevron and my group leader, John Toldi, for the experience of working on this project as an intern at Chevron and for giving me permission to publish it in this report.

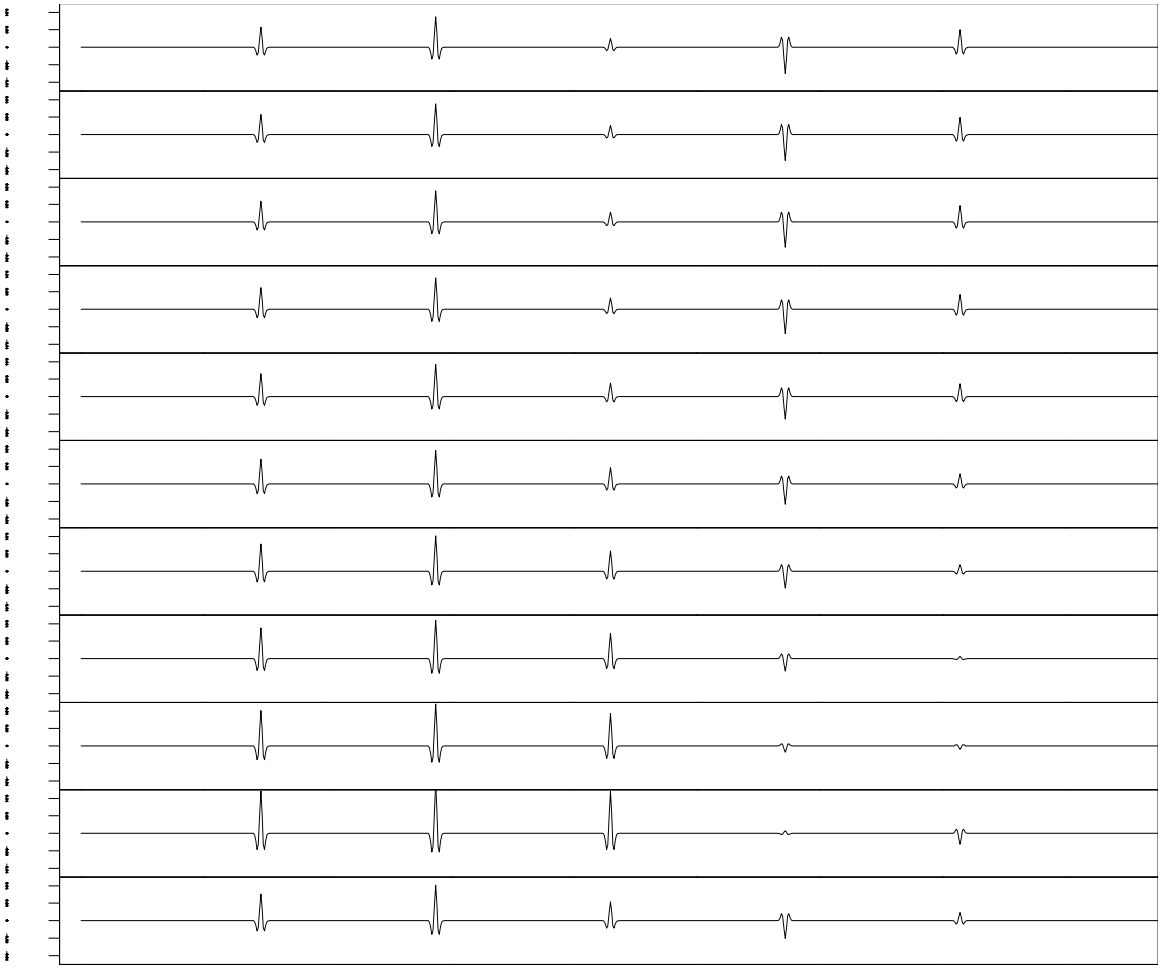


Figure 1: Input angle gather. The bottom trace is the stacked trace. marie2-input [ER]

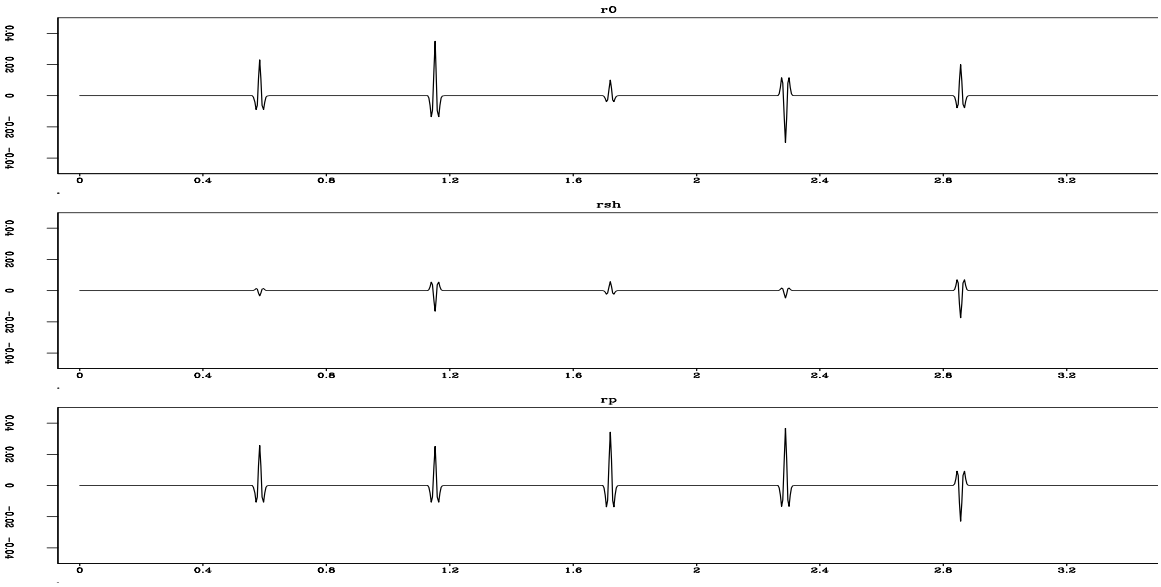



Figure 2: Output reflectivities. Top: $R_O$, middle: $R_{sh}$, bottom: $R_P$. marie2-output [ER]

Many Chevron employees, most notably Rich Alford and Jeff Wright, were also instrumental in this work.

## REFERENCES

Aki, K., and Richards, P. G., 1980, Quantitative seismology: Theory and methods: W. H. Freeman and Co., New York.

Backus, M. M., 1983, Research workshop 2 - the change in reflectivity with offset: 53rd Ann. Internat. Mtg, Soc. Expl. Geophys., Session: RW2.1.

Bortfeld, R., 1961, Approximations to the reflection and transmission coefficients of plane longitudinal and transverse waves: Geophysical Prospecting, , no. 9, 485–502.

Castagna, J. P., and Backus, M. M., 1993, Offset-dependent Reflectivity - Theory and Practice of AVO Analysis: Soc. Expl. Geophysics, Tulsa, OK.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 291–304*

# Fault contours from seismic

*Jesse Lomask*[1]

**ABSTRACT**

Fault contours are created by mapping the slip distribution along a fault surface. Fault contours have many useful applications in geophysics and geology including rock stress analysis, interpretation, and processing. Therefore, automatic calculation of fault contours from 3D data would be very valuable. As the first step in that direction, displacement is calculated from a simple 2D model of faulted seismic data. Making a stationary assumption, the displacement is calculated by fitting a smooth line to a cross-correlagram. The cross-correlagram is created by windowed cross-correlation across the fault. To fit this line, I use a non-linear optimization method that has similarities to Simulated Annealing. This removes most of the displacement which will later allow other methods to solve the non-stationary problem.

## INTRODUCTION

A lot of useful information can be gained by understanding the slip distribution along a fault surface. Factors governing the slip distribution include the stress field, proximity to other faults, rock strength, tectonic history, and loading rates. If the slip distribution were easily and quickly obtained, it could be added to interpretation and processing workflows (Pollard, 2001).

A present, seismic fault contours are generated by tediously interpreting many horizons and calculating the slip as shifts between those horizons. As a result, only the interpreted data is used in calculating slips, and the process is very labor intensive and time consuming. In poor data quality areas, this method may be all that works. However, in areas where seismic image quality allows, automatic data-based methods would save time and, in some cases, increase the accuracy of the result.

In this paper, I describe a robust method for calculating the fault slip distributions from seismic data in order to bring the power of fault contours to bear on a number of problems. Figure 1 shows a synthetic 2D seismic section with an interpretted fault and a cartoon of fault contours. I create a cross-correlagram by placing windowed cross-correlations side-by-side. Then a non-linear line fitting program is used to fit a smooth line to the cross-correlagram. he lag values of this line remove most of the deformation of the fault in a model example.

[1]**email:** lomask@sep.Stanford.EDU

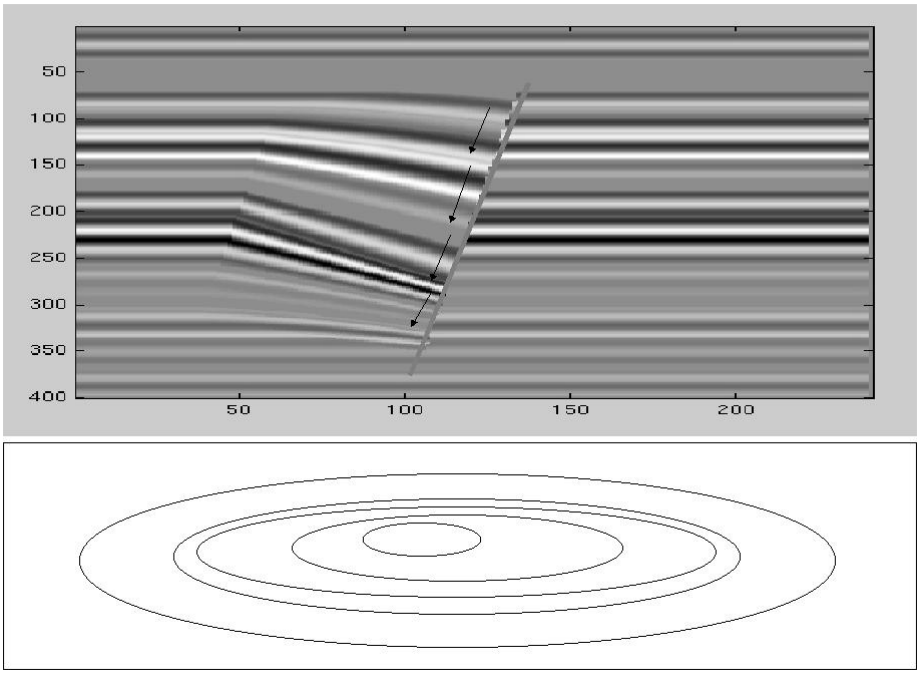



Figure 1: Top: A synthetic seismic section with a fault. Bottom: an idealized fault contour map showing slip distributions along the fault surface. jesse1-fltcont [NR]

## APPLICATIONS

Knowing the slip distribution has numerous applications, researchers studying fault mechanics are already familiar with some of these uses. However, it is quite likely that fault slip distribution information has been, in general, under utilized by seismic interpreters and processors. Hopefully, having fault contours automatically calculated will easily integrate them into interpretation and processing workflows.

**Stresses and Rock Strengths** First of all, one can speculate about the interplay of rock strengths and stresses. The contours are a vital input for researchers studying fault mechanics. They help unravel the stress and strain fields. Understanding the slip distribution can imply slip and position of neighboring faults.

**Seals** One can use fault contours to see where the mininum displacement should be to insure an adequate seal. For instance, if the thickness of a sand reservoir is known, then you will likely have a shale on sand seal if the displacement is greater than the thickness. In addition, in building a reservoir model, faults which have insignificant displacement would not necessarily be included.

**Geological interpretation** Combining the contours with other data can help unravel the geological history by revealing periods of growth along the faults in the case of synde-

positional faulting. Periods of rapid fault growth can be a result of increased sediment loading that can be tied to mountain building events, eustatic fluctuations, and climate changes.

**Well-ties in 3D modeling** The slip from a fault contour can be backinterpolated at well locations. In the case of normal faulting, the magnitude of the slip can be compared with the missing section from the wells. In the case of reverse faulting, it can be compared with the repeat section. This information can be incorporated as an additional well tie and input into 3D models.

**Automatic horizon interpretation** If the slip distribution is known for an entire volume of 3D seismic data, then the faults can be removed. Once they are removed, then auto-interpretation procedures can be used to interpret the horizons with ease.

**Constrain processing** By tying with the regional stress field, likely faulting strike, dip and magnitude can be estimated. This has potential to help constrain velocity analysis. In addition, fault displacement information from well data can be used as constraints.

**Missing section information used in standard fault interpretation** Knowing the fault displacement can be incorporated into an interpretation program to help correlate faults. Faults are typically interpreted by an interpreter digitizing fault segments, then correlating the appropriate segments with particular faults, and triangulating fault surfaces. If the slip distribution was calculated on that triangulated fault surface, the interpreter can take advantage of certain fault behavior rules to correlate and measure the quality of the correlation. Figure 2 is a cartoon of a mapview of fault gaps. Fault gaps are mapview

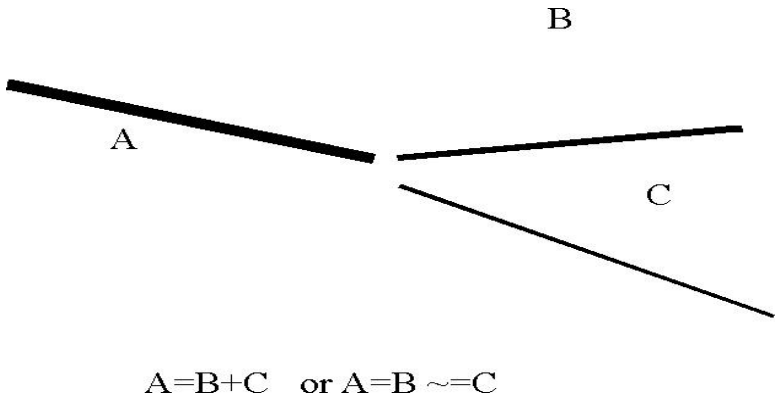



Figure 2: Fault Gaps jesse1-faultgap [NR]

representations of missing sections where normal faulting has caused the horizon to separate. When three fault breaks come together, it is often unclear whether one fault split into two or two different faults just pass closely to each other. In the first case, the missing section on fault A will equal the sum of missing sections on faults B and C. In the second case, perhaps A and B are one fault and C is a separate fault; therefore, the missing section on A will equal B and not equal C.

## EXTRACTING LAGS

There are a number of ways to extract the slip distribution along a fault plane. Windowed cross-correlation is the simplest; however, it needs to be constrained. To test this out, I started with a 2D synthetic seismic section. I created 40 randomly spaced reflection coefficients with random magnitudes. I stretched one side to mimic fault deformation, shown in Figure 3. I, then, convolved it with a simple zero phase wavelet, shown in Figure 4.

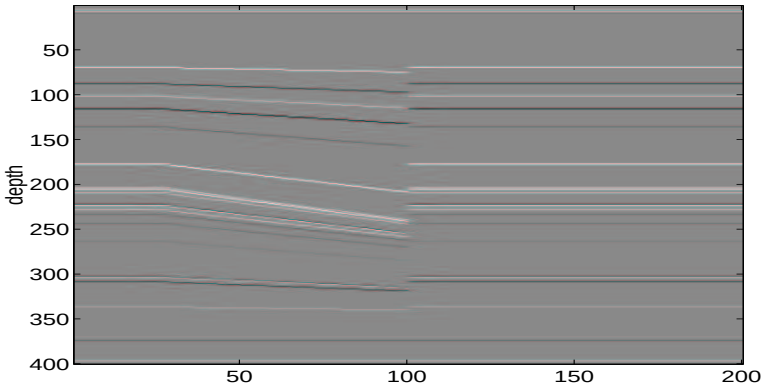



Figure 3: Model of reflection coefficients. jesse1-makemodraw [CR]

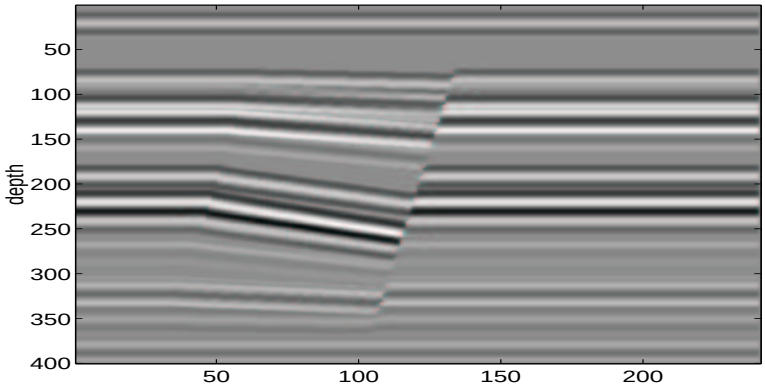



Figure 4: Model after convolving with a source wavelet and bending along a fault. jesse1-makemoda [CR]

The stretching and compressing creates a non-stationary problem. I am still uncertain how much of a problem this would be in real data. In this model, simple cross-correlation techniques can be a good start, but they will not completely remove the deformation.

Regardless, the first step should be to flatten the data on the fault plane, as in Figure 5.

Then one trace can be extracted from each side of the fault. The problem then becomes to find the shift to apply to the trace left of the fault to make it similar to the right side of the fault.

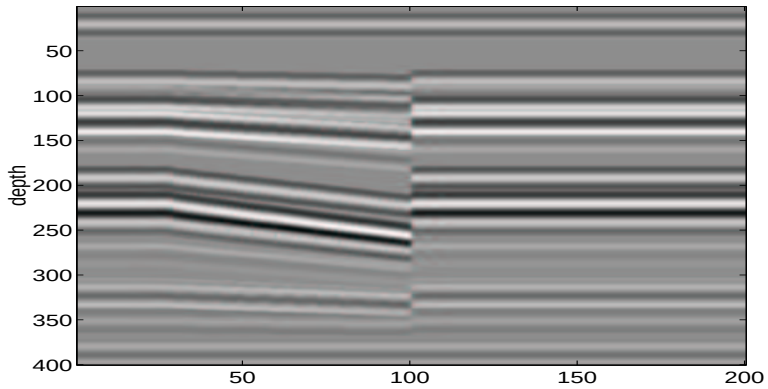



Figure 5: Model after flattening on the fault surface. jesse1-makemod [CR]

**Treating as a stationary problem**

Rather than trying to account for the non-stationary stretching and compressing, I decided it would be simpler to treat the removal of displacement across the fault as a stationary problem and find a robust method that gives a solution that is close, later addressing the non-stationarity.

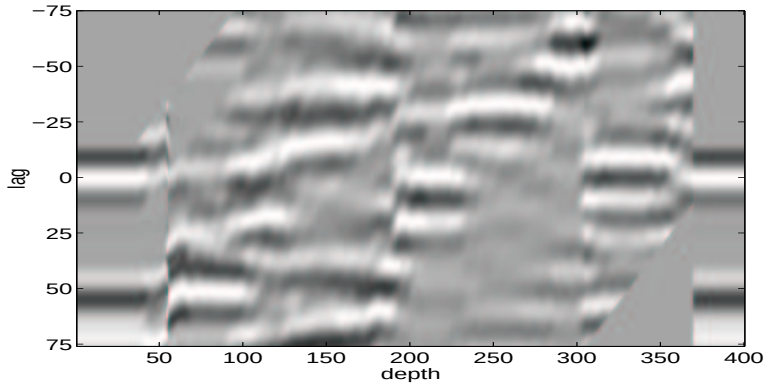



Figure 6: Cross-correlagram: the white color represent mininum values which are actually maximum correlation values. jesse1-makexgram [CR]

The first step was to create a cross-correlagram as shown in Figure 6. This figure is a continuum of windowed cross-correlations from the left and right sides of the fault. As a first step, a smooth line must be fitted to the cross-correlagram. The departure of this line from the center, or zero lag, is the amount of displacement required to remove the fault deformation. Figure 7 shows the cross-correlagram overlaid with the actual displacement used to deform the model plotted. It doesn't land on a peak because in making the model the reflection coeficients were deformed and then convolved with a wavelet, introducing tuning. The tuning causes the cross-correlagram to be skewed from the ideal answer. Also, even knowing the amount of deformation used to create the model, will not remove the deformation completely. Additionally, blindly picking the lag with minimum values will not work as illustrated in Figure 8. In the cross-correlagram, the maximum and minimum are inverted so that we will actually search for

the minimum as the line of maximum correlation. In this way, we can keep with the inversion convention of searching for minima.

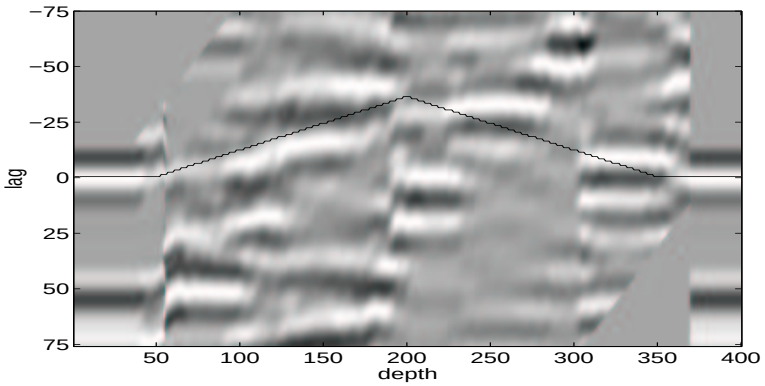



Figure 7: Cross-correlagram overlain with known displacement. jesse1-plotans [CR]

Using linear least squares to fit a smooth line to the maximum path across this cross-correlagram will not work. This is a non-linear problem that is full of local minima. Instead, this problem requires a non-linear approach. A derivative of Simulated Annealing proved to work (Kirkpatrick, 1983).

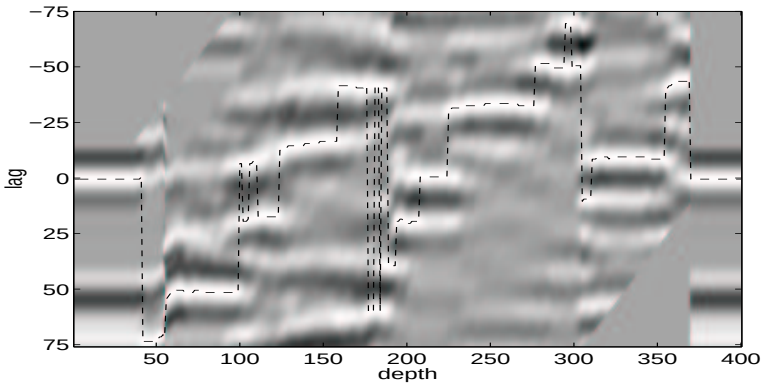



Figure 8: Cross-correlagram with minima at all depths. jesse1-plotmin [CR]

**Simulated Annealing**

Simulated Annealing is a non-linear optimization method. The idea is to find the global minimum in a way similar to cooling magma forming crystals. If the magma cools too quickly, glass forms. Glass represents local minima. When the magma is cooled slowly enough, then crystal forms. Crystal represents global minima. Rothman used this approach for statics estimation (Rothman, 1985).

We can define our optimization goal as:

$$\mathbf{E_1} = \sum_i \mathbf{C}(m_i) \tag{1}$$

For this problem, we needed to apply some sort of regularization to ensure that the solution is smooth.

$$\mathbf{E_2} = \sum_i \text{abs}(\frac{\text{d}}{\text{dt}}\mathbf{m}) \tag{2}$$

These two equations can be combined and balanced with $\epsilon$ as

$$\mathbf{E} = \mathbf{E_1} + \epsilon \mathbf{E_2} \tag{3}$$

A computation template for the method of simulated annealing is

```
rate = f (iteration)   <--   define cooling schedule
m    <--   0.
iterate {
        n    <--   random numbers
        if  E_n < E_m {
            m    <--   n
        } else {
            P = e^(-ΔE·rate)
            }
        if  P > random number {
            m    <--   n
            }

        }
```

The model (**m**) is a vector which defines the extracted lags from the cross-correlagram (**C**). Energy (**E**) is what we are trying to minimize by finding the smoothest path with the lowest energy across the cross-correlagram. The trial model (**n**) is a vector the same size as **m**. We define a **rate** parameter to follow a cooling schedule as a function of iterations. A computation template for the method of simulated annealing is **P**, a probability value assigned to a trial vector (**n**). If a random value is less than this probability value, the trial vector will be accepted even if it has a larger energy. This will allow the solution to escape local minima. This is analogous to cooling magma; when the temperature is high, it is less likely to get locked into position. As the number of iterations increases, **P** becomes smaller, making it less likely to accept a solution with greater energy. If the rate is sufficiently small, this method should converge to the global minimum.

I applied this method to estimate fault slip. The result is in Figure 9. The dots represent the annealing solution. The initial guess for **m** was the maximum, the solid line. This required numerous iterations and, although it appears that it may be heading toward convergence (the light solid line), it may not be. In fact, the added complication of the $\epsilon$ parameter makes it quite likely to converge to a flat model. The regularization is causing the slow convergence because randomly changing model points tend to cause an extremely unsmooth solution. Therefore, the regularization part is throwing out most of the possible solutions because they are too rough.

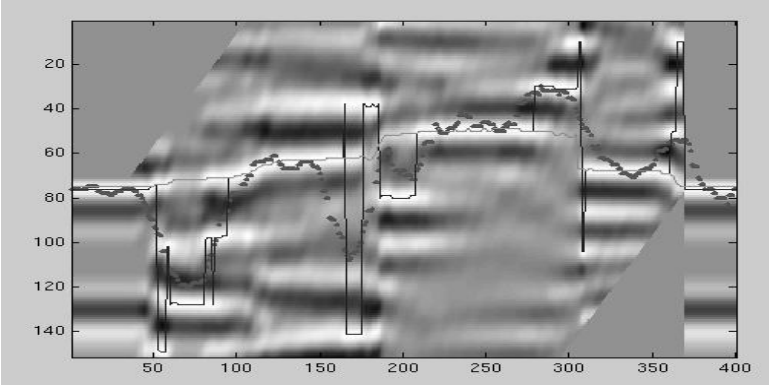



Figure 9: Simulated Annealing result after many iterations, dots. Heavy solid line, picked maximum initial solution. Light solid line, desired result. jesse1-Anneal1bw [NR]

**Another approach with similarities to Simulated Annealing**

To allow the possible solutions to change radically (sometimes necessary to escape local minima), I tried an algorithm that tries random but smooth solutions at the beginning when the temperature is hot. As the temperature cools, it tries solutions that have more and more roughness. Rather than randomly changing all of the components of **n**, only a few widely spaced samples are randomly changed. Then the rest of the points are interpolated between them and the energy is calculated over the interpolated line.

$$\mathbf{E_1} = \sum_i \mathbf{C}(m_i) \tag{4}$$

This process is repeated and the sampling interval becomes smaller and smaller with increasing interations. This method does not require the use of the probability to get convergence, although that may still have its use later.

A computation template for this method is

```
sample interval = f (iterations)   ⟵   define cooling schedule
m   ⟵   0. or best initial guess
iterate {
        n (sample interval)   ⟵   random numbers
        interpolate (n)
        if  E_n < E_m {
            m   ⟵   n
            }

        }
```

At coarse sampling intervals, the function is inherently smooth but as the sampling interval gets smaller and smaller, the trial solutions may become too rough. Therefore, I added an

additional constraint on the maximum amount of roughness allowed. This results in throwing out trial solutions only near the end when the sampling interval is very small.

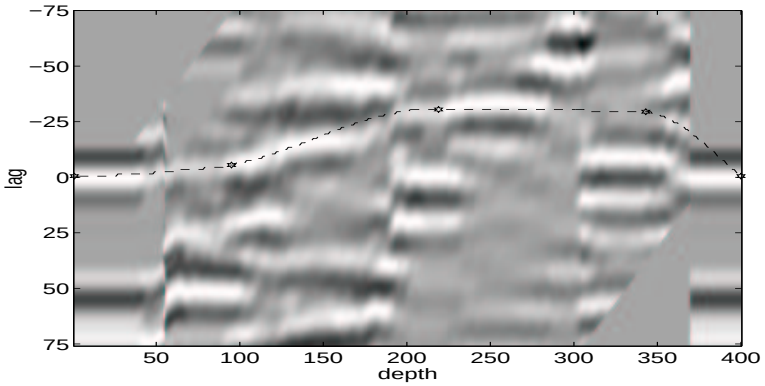



Figure 10: Cross-correlagram with solution overlay after few iterations. The dots reflect the coarse sample interval. jesse1-anneal_course [CR]

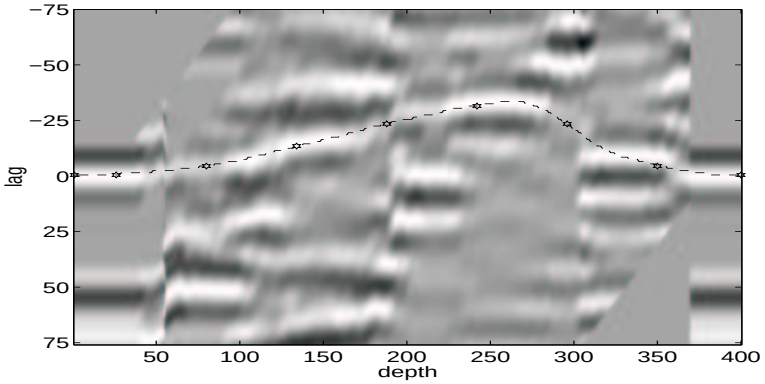



Figure 11: Cross-correlagram with solution overlay after many iterations. The dots reflect the fine sample interval. jesse1-anneal_fine [CR]

Figure 10 shows a frame taken early on in the estimation process. The large dots show the randomly moving points and the dashed line is the interpolated line. Figure 11 shows another frame taken later; the density of the dots illustrates the shrinking sampling interval.

**Cooling schedule**

As Rothman (1985) points out, the selection of the cooling function is very important and can greatly speed convergence. In my method, I treat the sample interval function as a cooling function. To find a good cooling function, I first created a linear cooling function and plotted how the energy decreased as a function of sample interval. This is shown in Figure 12. In this particular example, there appear to be three sample intervals associated with large drops in energy: 185, 130, and 70. I decided to create a cooling schedule that treats the sized 70 sample

interval as the critical temperature. This is shown in Figure 13. The critical temperature is where convergence is most significant.

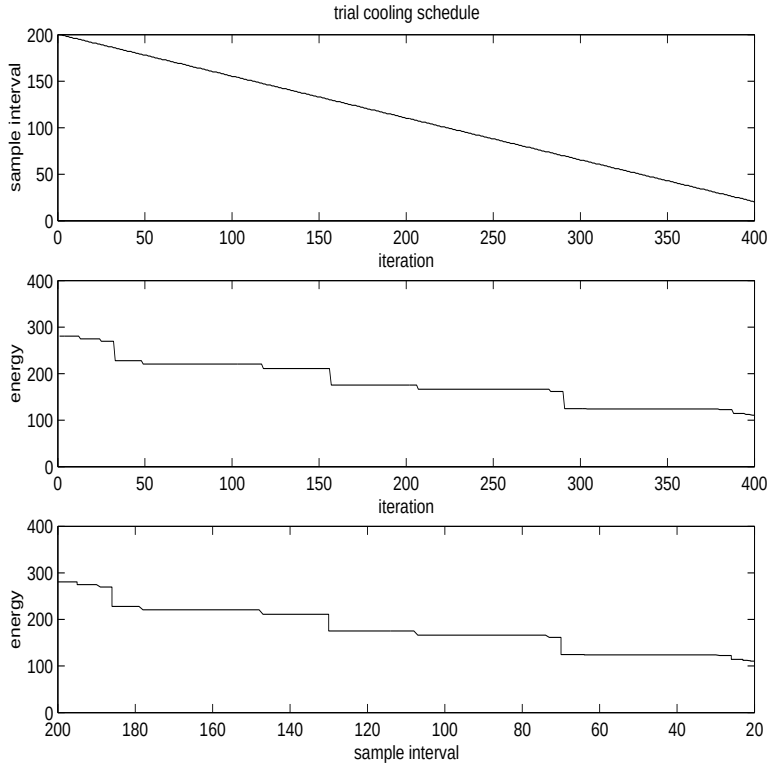



Figure 12: Trial cooling schdule. Top: input cooling schedule drops linearly with iterations. Middle: energy(**E**) drops in steps with iterations. Bottom: energy(**E**) drops in step with sample interval. This is used to determine the more efficient cooling schedule in Figure 13. jesse1-trialschedule [CR]

Notice that in Figure 13, the energy drops off much quicker than in Figure 12 and therefore requires fewer iterations.

Figure 14 shows the result of applying the cooling schedule in Figure 13. It has converged to the desired event. Figure 15 shows the application of the calculated displacements to the left side of the fault. Its results are about the same quality as Figure 16, which shows the results of applying the actual known displacement to the left side of the fault.

## CONCLUSIONS AND FUTURE WORK

The iterative method for removing most of the deformation in this model works very well. Calculating a solution to this small model, it takes just a couple of minutes on a PC and successfully removes the fault deformation as is evident in Figure 15.

The next step is to apply this to real data. First to a simple 2D line and then to a 3D volume.

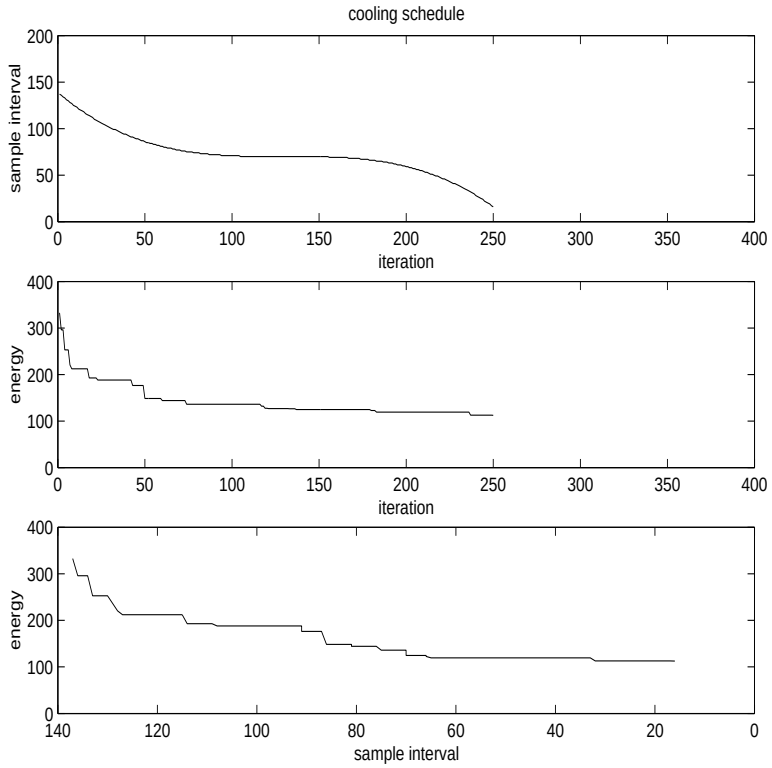



Figure 13: Cooling schedule. Top: input cooling schedule based on trial cooling schedule designed to spend more iterations near the critical tempurature. Middle: energy(**E**) drops in steps with iterations. Bottom: energy(**E**) drops in step with sample interval. jesse1-coolingschedule [CR]

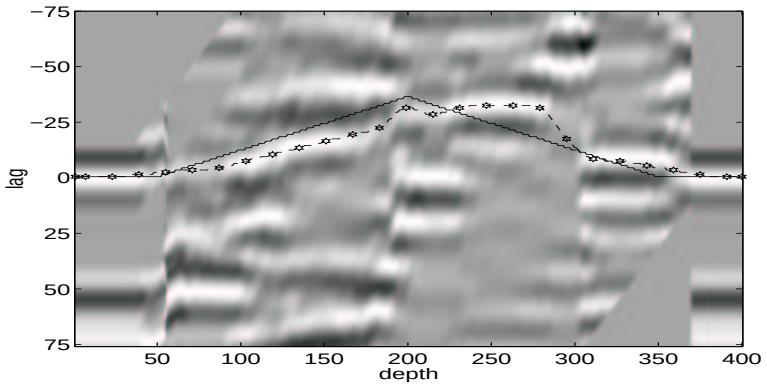



Figure 14: Cross-correlagram with overlay of final solution and known displacement. jesse1-anneal_final [CR]

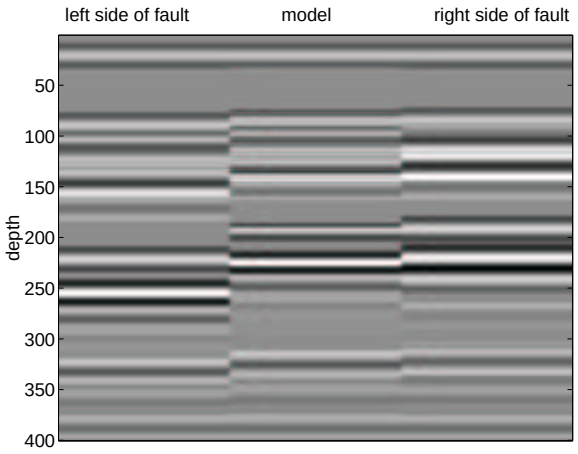



Figure 15: Applied result, the "model" in the center shows the result of applying the calculated displacement to the left side of the fault. jesse1-makefinalmod [CR]

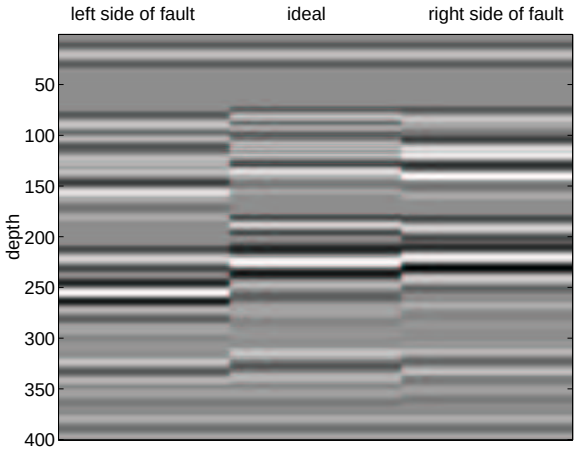



Figure 16: Ideal applied result, the "ideal" in the center shows the result of applying the known displacement to the left side of the fault. jesse1-makefinalideal [CR]

Other types of non-linear solver methods can be tested on this problem as well. Automatic picking algorithms may work faster and easier than my method.

If this method is successful at removing the bulk of deformation on real data, then we may want to address the non-stationary aspect to improve the results.

Lastly, this work can possibly be utilized to aid in automatic fault interpretation. Starting with a raw 3D seismic volume and a coherency cube, fault slips can be used to constrain the automatic interpretation of the faults themselves.

A logical next step to finding the fault contours would be to automatically calculate the entire deformation ellipsiod as outlined in Figure 17. In this simple model it seems simple enough, but in the chaotic world of real geology, that may be an entirely different matter.

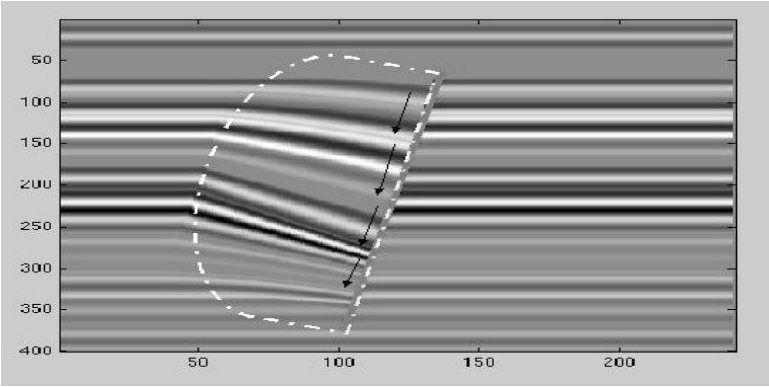



Figure 17: Deformation Ellipse jesse1-ellipse1 [NR]

## ACKNOWLEDGEMENTS

During a conversation, Biondo Biondi suggested that we should try to come up with a way of automatically extracting fault slip distributions from seismic.

## REFERENCES

KirkpatrickS. snd Gelatt, C. a. V. M., 1983, Optimization by simulated annealing: Science, **220**, 671–680.

Pollard, D. Structural geology and geomechanics:. Internet, 2001. http://pangea.stanford.edu/geomech/Research/Research.htm.

Rothman, D. H., 1985, Large near-surface anomalies: SEP–**45**, 5–18.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 305–323*

# Toward subsurface illumination-based seismic survey design

*Gabriel Alvarez*[1]

**ABSTRACT**

The usual approach to the acquisition of 2-D and 3-D seismic surveys is to use a recording template designed from the maximum target dips and depths. This template is used through out the survey area irrespective of changes in the dips, depths or propagation velocities of the targets. I propose to base the survey design on an initial structural and velocity model of the subsurface. The initial structural and velocity model is used to trace rays to the surface at uniform illumination angles from all reflecting points of interest. The emergence ray points are recorded as tentative source and receiver positions and a constrained inversion is used to optimize them given logistic and economic restrictions. I apply this strategy to a very simple 2-D synthetic model and show that in order to image steep dips larger offsets are required than would be anticipated from the traditional approach. This in turn implies the need for different source and receiver positions.

## INTRODUCTION

The usual strategy for the design of 2-D and 3-D seismic surveys starts from the estimation of "critical" values for subsurface parameters such as maximum geological dip, maximum target depth, minimum target rms velocity and minimum target thickness. From these values (and an estimation of the minimum required fold), we choose acquisition parameters such as receiver group and source interval, maximum offset and number of active channels per shot. In 3-D we also compute the number of active receiver lines, the number of shots per salvo, the width of the active patch and the geometry of the acquisition template (Stone, 1994). More often than not these parameters are held constant across the whole survey irrespective of changes in the geometry of the subsurface we wish to image or of its associated velocity field.

Common usage in 3-D seismic acquisition focuses on regularity of source and receiver positions on the surface of the earth according to one of a few "standard" geometries. These geometries are designed to provide as regular offset and azimuth coverage in each CMP bin as possible while at the same time allowing for relatively easy logistics. Although there are important differences between the different "standard" geometries in terms of their offset and azimuth distributions (Galbraith, 1994) and other less obvious characteristics such as the symmetric sampling of the wavefield (Vermeer, 2001), these regular geometries all share these

[1]**email:** gabriel@sep.stanford.edu

characteristics:

- Quality control is essentially qualitative and relies on regularity of offsets and azimuth in adjacent CMP surface bins. The actual subsurface bin population at the target location (which is clearly a function of the reflector dip and the dips and velocities of the overburden) is completely ignored or considered only as an afterthought of the design.
- Worst of all, these geometries make no attempt whatsoever to guarantee or even promise any sort of uniformity in subsurface illumination. This may be critical when trying to image under salt flanks and in overthrust areas with large fault displacements (which are common in foothill data). No amount of clever processing can compensate for problems generated by poor illumination during acquisition.

The choice of the parameters for these "standard" geometries can be posed as an optimization problem once the basic template is chosen (Liner and Underwood, 1999) and even the logistics and economics of the acquisition can be incorporated into the computation, at least to some extent (Morrice et al., 2001). The basic premise, however, is that the underlying acquisition template is regular and chosen before hand, so that the optimization is restricted to look for the best combination of the parameters consistent with the chosen template.

The quality of the overall design is evaluated based on such surface attributes as uniformity of fold of coverage and regularity of offset sampling (and azimuth sampling in 3-D). This is equivalent to an implied assumption of a layered, constant velocity subsurface model. Illumination is only considered (if considered at all) as a forward problem. Illumination maps may be constructed at the target reflectors for each of a few competing geometry templates and the best one selected by a qualitative comparison of those maps (Cain et al., 1998). Moreover, the complete design may be simulated in the computer further analyzing the resultant distribution of offsets, azimuths and intensity of illumination (Carcione et al., 1999). Although this is an important step in the right direction, there is no guarantee that the chosen geometry will in fact produce optimum illumination.

I propose to base the design on an initial structural and velocity model of the subsurface, even if only a crude one. Obviously, at the time of acquisition, we don't have a detailed subsurface model. Oftentimes, however, we know the rough features we wish to image. Is it a salt dome, an overthrust faulted anticline or a deep channel turbidite system? This information may come from previous seismic data, from well logs, from a conceptual geological model, from surface geology and usually from a combination of all of them. This wealth of information is ignored in the usual acquisition design but it does not need to be. It is possible and indeed desirable to use the existing knowledge of the subsurface structure and velocity model to improve the acquisition of new data. I show with a very simple synthetic 2-D example that we can reduce the number of shots without compromising the quality of the image by selectively ignoring shots whose contribution to the image is less than, say, half the number of traces of a regular shot.

The two key points are: we are not required to use the same parameters or indeed the same geometry all across the survey and regularity of *surface* parameters is not necessarily the mark

of an optimum design. A better indicative is regularity of *subsurface* attributes such as target illumination.

## THEORY OVERVIEW

The traditional approach to the design of land 2-D seismic surveys follows these steps:

1. Measure or estimate from existing seismic data, from well logs or from geology data the maximum dip of interest in the survey area. This is not necessarily the maximum dip of the reservoir unit itself but may be the dip of a sealing fault or an important unconformity or any other such relevant geological feature. With this value, estimate the maximum receiver group interval ($\Delta x$) in order to avoid spatial aliasing in the data given the maximum frequency ($f_{\max}$) expected or required in the data (in turn a function of the minimum target thickness)

$$\Delta x \leq \frac{V_{\min}}{2 * f_{\max} \sin \alpha_{\max}} \tag{1}$$

   In this equation $\alpha_{\max}$ is the maximum target dip.

2. Set the maximum offset equal to the maximum target depth (again, not necessarily the maximum depth of the reservoir itself).

3. Compute the number of active channels per shot ($nc$) from the maximum and minimum offset and the receiver group interval

$$nc = \frac{off_{\max} - off_{\min}}{\Delta x} + 1 \tag{2}$$

   For symmetric split-spread cable (equal number of receivers on both sides of the shot) the number of channels is multiplied by 2.

4. From the expected signal/noise ratio in the data (estimated from previous vintages of seismic data in the survey area or nearby), estimate the minimum required fold of coverage (usually something like 30, 60 or 120).

5. From the fold, the number of channels and the receiver group interval, compute the source interval ($\Delta s$)

$$\Delta s = \frac{2 * Fold * \Delta x}{nc} \tag{3}$$

6. The migration aperture ($M_{\mathrm{a}}$) is computed from the depth $z$ and dip $\alpha$ of the steepest dipping reflector at the end of the survey. In 2-D it is usually estimated as

$$M_{\mathrm{a}} \approx z * tan(\alpha) \tag{4}$$

   Because the velocity increases with depth this estimation tends to be a little pessimistic and sometimes a fraction of it is actually used.

These are the relevant parameters for 2-D acquisition. Another important consideration is the type of source (for example charge size and hole depth for acquisition with an explosive source) which is usually determined with field tests just before the start of the acquisition.

In 3-D other important parameters include the number and separation of the active receiver line, the number of shots per salvo and the separation of the source lines. The aspect ratio (ratio of maximum in-line offset to maximum cross-line offset) is another important consideration (Stone, 1994).

## PROPOSED METHODOLOGY

In 2-D acquisition the geometry of the layout is almost always chosen to be either split-spread (land) or end-on with the shot pulling the cable (marine), so there isn't much flexibility in that respect. Illumination requirements essentially control the maximum offset of interest, which will change along the line profile as the target depths, dips and velocities change. The fold requirements may also change along the line profile but this situation is less common. The other parameter that may change is the receiver group interval but we have a strong limitation in the choice of this parameter because the seismic cables usually have take outs only at predetermined distances.

The proposed strategy for a more flexible acquisition is the following:

1. Construct a subsurface model of the survey area as accurately as possible in terms of the geometries of the target reflections and interval velocities.

2. Do non-zero-offset exploding reflector modeling using this model. This means to consider "all" points along each of the target reflections as a Huygen's source and to track rays up to the surface at given uniform reflection aperture-angle increments. This would guarantee perfect illumination if in fact the given rays can be generated during the acquisition. For each pair of rays (with the same aperture angle on opposite sides of the normal to the reflector) record their emergence positions at the surface.

3. Clearly, the optimum source and receiver positions for each target will be different from those of the other targets. Reconciling these optimum illumination source and receiver positions can be posed as a non-linear inversion procedure. Inputs will be the source and receiver positions required to illuminate each target and output will be the source and receiver positions that minimize the sum of the deviations from the requirements for each target. For this process to be useful, several constraints must be imposed. For land acquisition the most important ones will be:

   (a) The receivers must be equally spaced at a distance consistent with the cable take-outs of the seismic cables.

   (b) The number of shots should be kept to a minimum.

   (c) in 2-D, source and receivers should be kept along the profile line.

More subtle, and perhaps more difficult constraints to honor, will be those related to spatial sampling considerations for prestack migration. Since the emphasis of the acquisition design will be placed on maintaining regularity of subsurface illumination, surface parameters such as offset and fold will not be uniform. Adequate sampling of offsets and azimuths, however, is a stringent requirement that must be honored by the inversion procedure. It is also important to note that the optimum geometry may not provide uniform illumination, but will likely provide better illumination than the standard approach.

4. Finally, the optimum source and receiver positions will be output in a suitable format such as Shell's SPS geometry format. These geometry files can be readily input to the acquisition instrument.

## A SIMPLE 2-D MODEL EXAMPLE

### Description of the model

As a first test of the above ideas, I created a simple model shown in Figure 1. The model has constant velocity background of 3000 m/s. .

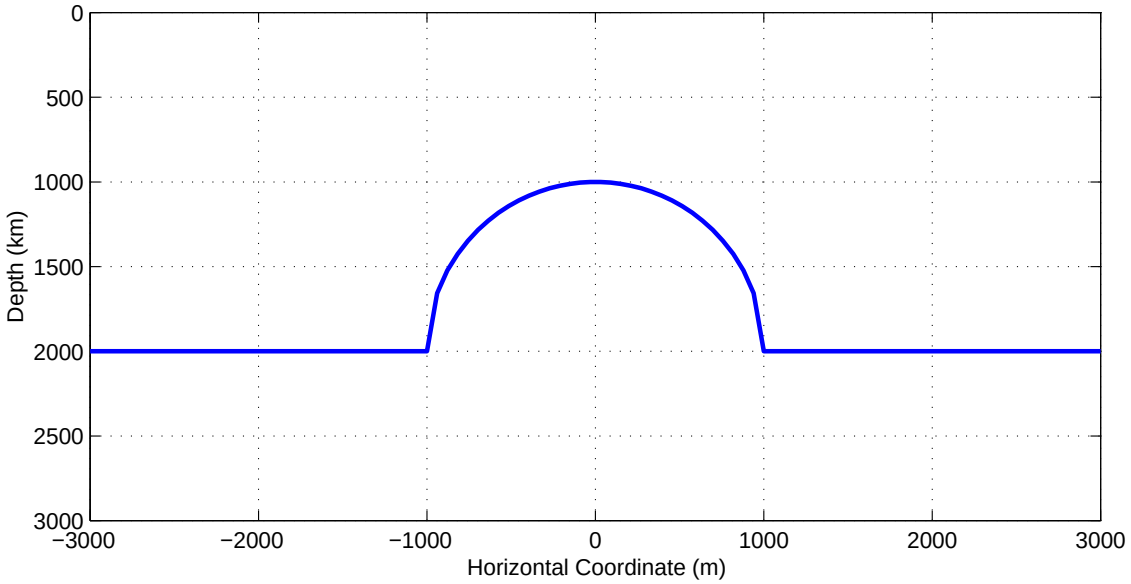



Figure 1: Simple model used to illustrate the proposed methodology. The background velocity is constant (3000 m/s) gabriel2-model1 [NR]

### Modeling with traditional parameters

In this case $v_{\min} = 3000$ m/s, $\alpha_{\max} = 90$ and $f_{\max} = 50$ Hz. Applying Equation 1 we get $\Delta x = 30$ m. The maximum depth of the target is 2000 m, so assuming an off-end cable (receivers to one side of the shot only) this will give about 67 receivers. This number rounds up

nicely to 72, which means an actual maximum offset of 2160 m (assuming that the first receiver is at an offset of 30 m). The shot interval is chosen as $\Delta_s = \Delta_x = 30$ m so that the fold of coverage (number of receivers per CMP) is 36. The trace length was chosen to be 4 s. With these parameters I simulated acquisition using an analytical ray-tracing program. Obviously, we cannot expect to image the dips of the semicircular reflector up to 90 degrees because according to Equation 4 that would imply infinite aperture. The best we can do is image the maximum dip for which the reflection time is less than or equal to the trace length. In this case, given the simple geometry of the reflector, a quick computation shows that for the zero offset trace this corresponds to shot positions $\approx \pm 6700$ m. The corresponding maximum dip is $\approx 73$ degrees.

The acquisition proceeds from left to right in Figure 1. When the shot is to the left of the semicircle the longer offsets have shorter arrival times (from the semicircle) which means that we can actually achieve full fold at that point by extending the acquisition to the left by half cable-length (1080 m). The first shot is therefore at -7780 m. On the other hand, when the shot is to the right of the semicircle, longer offsets have longer arrival times and so we cannot expect to have full fold at 6700 m. Any shot past that point will only contribute reflections longer than the trace length. In summary, with the standard approach (using the off-end cable described above) we need shots between -7780 and 6700 m in order to image the maximum dip. At 30 m shot interval this implies 483 shots.

Figure 2 shows some of the modeled shot records. At both sides of the semicircle we see two reflections coming from the flat and the semicircular reflector, whereas above the semicircle only the reflection from the semicircular reflector is seen. Figure 3 shows some CMP gathers. Since the design is completely regular, the CMP's are also regular. This is further illustrated in Figure 4 which shows the fold diagram. Note that we have full fold at -6700 m but not at 6700 m.

Figure 5 shows the stacked section. The noise at the intersection between the flat and the dipping reflections reflects the inherent difficult in picking a stacking velocity appropriate to both (no DMO was applied). Finally, Figure 6 shows the post-stack Stolt-migrated section. As expected, dips in the semicircular reflector higher than 73 degrees were not recovered.

## Modeling with the proposed design

For the first step I used an analytic ray tracer to compute the surface emergence positions of rays originating at equally-spaced reflection points along the reflector. These reflection points were taken every 18 m, corresponding to the CMP interval in the traditional design. For the flat reflector the number of pairs of rays originating at each point was kept equal to the fold in the standard design. The rays correspond to a uniform increase in reflection aperture angle and therefore will not correspond to uniform offsets in a CMP gather. In the semicircle the number of rays was increased as a function of the reflector dip, so that where the dips are larger more rays were generated. The maximum offset was not constrained except for the obvious requirement that any reflection time were less than the chosen trace length.

The next step is to use non-linear inversion to find the optimum source and receiver posi-

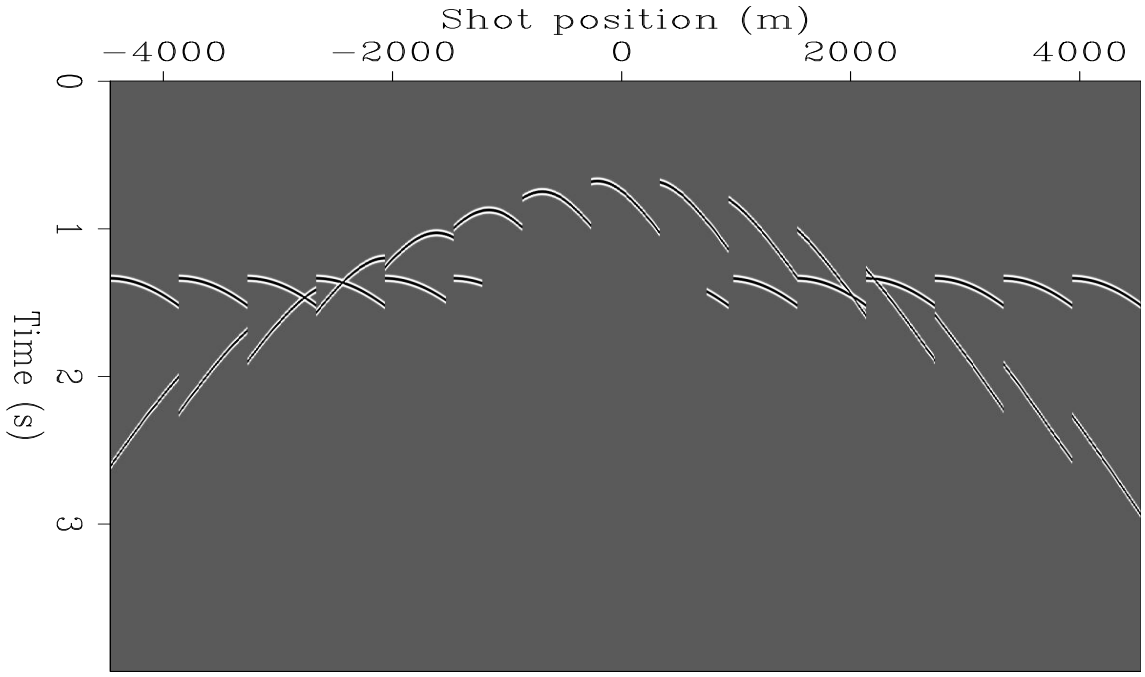



Figure 2: Synthetic shot records modeled with the standard geometry. gabriel2-modcirc1_shots [ER,M]

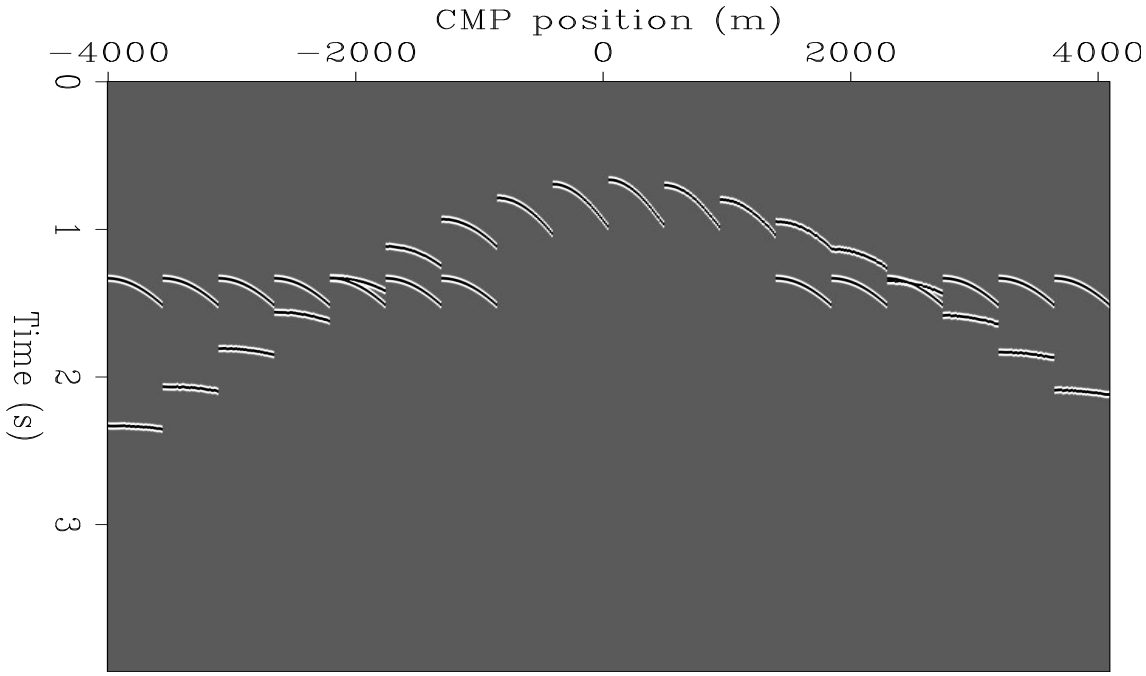



Figure 3: CMPs modeled with the standard geometry gabriel2-modcirc1_cdps [ER]

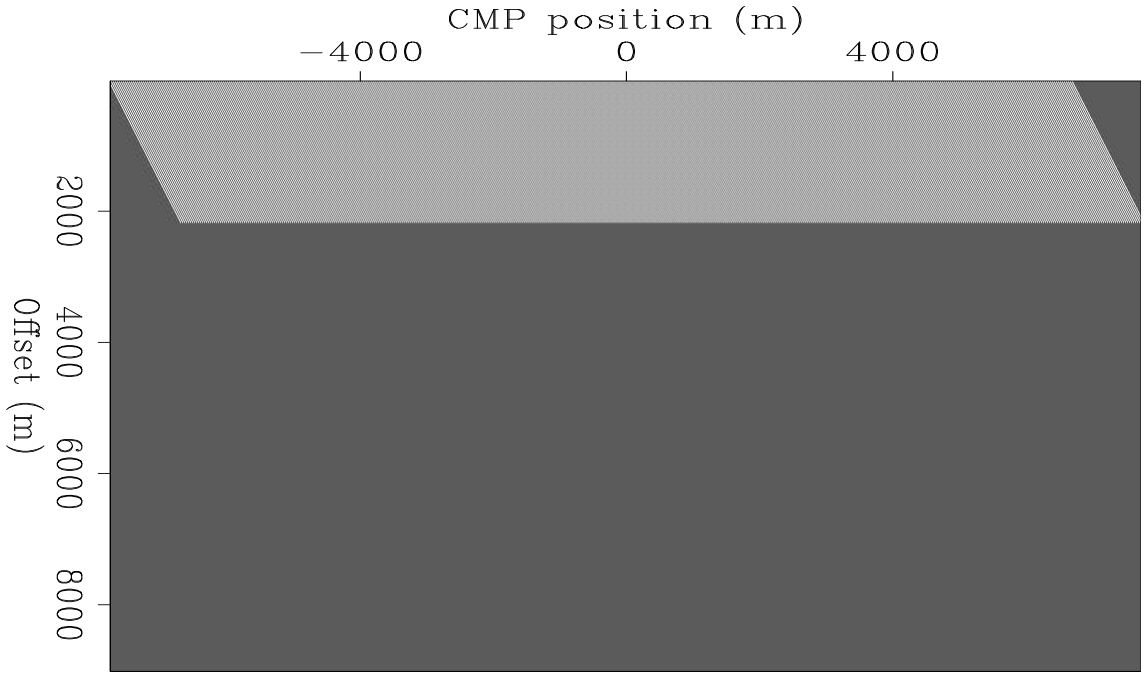



Figure 4: Fold diagram for the standard geometry gabriel2-modcirc1_fold [ER]

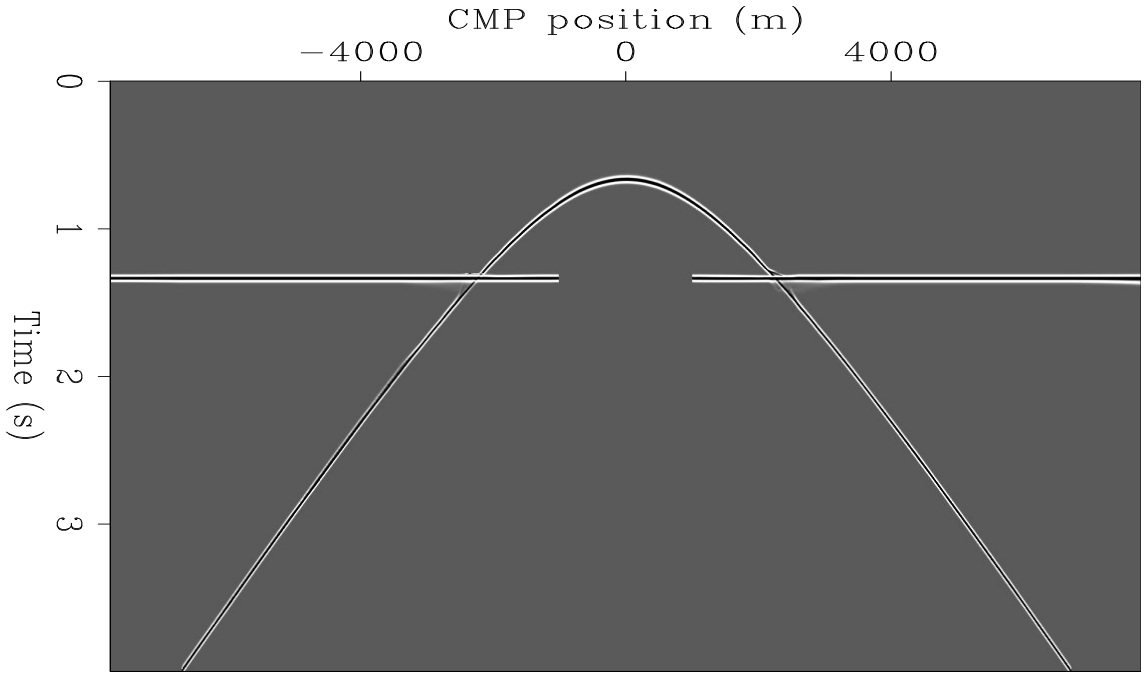



Figure 5: Stacked section of modeled data generated with the standard design gabriel2-modcirc1_stack [ER]

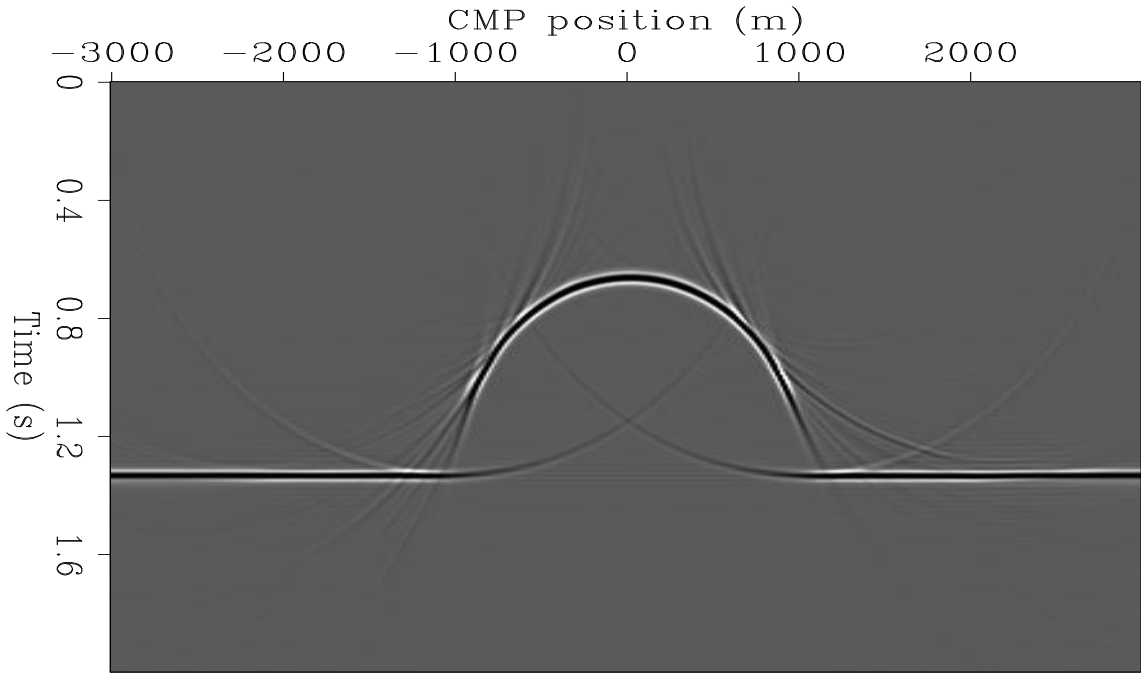



Figure 6: Post-stack migrated section of modeled data generated with the standard design gabriel2-modcirc1_mig_all [ER]

tions from the different targets. In this case, since there is only one target, the inversion reduces to a simple binning to honor the constraint that the receivers should be at equal distance along the line profile. I chose for the receiver interval the same value obtained in the standard design so that the results can be easily compared. The shots were also binned at the same receiver interval to further regularize the design (and guarantee equal distance between CMPs and between the stacked traces). Only shots that contribute at least half the number of traces of the standard geometry shot (that is 36) were considered. A total of 402 shots met this criterion (which is rather arbitrary). The number of traces per shot, and hence the maximum offset, was allowed to change from shot to shot. In this example this is the only degree of freedom that I used to adapt the acquisition effort to variations in the subsurface dip. In a real case, where the reservoir location is known or suspected, we could locally vary the receiver group interval or more likely the shot interval. In 3-D there are extra degrees of freedom associated with the azimuth and the choice of geometry template.

The next step is to simulate the acquisition of the data using the computed shot and receiver positions. Again, this was done with an analytic ray tracer. Figure 7 shows some of the shots. Note that they have different number of traces. Also, they look irregular because the plotting program places the traces together at the same distance irrespective of their offset. Figure 8 shows some CMPs along the line profile. As with the shots, the number of traces changes from CMP to CMP. Also note that there are "holes" in the CMP's illustrating the difference between uniform offsets and uniform illumination. Figure 9 shows the fold diagram. In this case there are differences in the fold coverage from CMP to CMP. As long as the minimum CMP fold is maintained, this shouldn't be a problem. More importantly, note the large offsets

at both sides of the semicircle and the smaller offsets above the semicircle (compare with Figure 4). Figure 10 shows the stacked section. Comparison with Figure 5 does not reveal any striking difference because the stack smooths out the effect of the irregular offsets. The important difference between the two figures is the lateral extent of the semicircular reflection. Figure 11 shows the migrated section and Figure 12 shows a comparison with the migrated section obtained with the standard acquisition. Not surprisingly, the two images are almost the same, since they were computed with the same aperture. The proposed design, however, required about 180 fewer shots.


Figure 7: Synthetic shot records modeled with the proposed acquisition design. gabriel2-modcirc3_shots [ER]

This example is rather artificial in that the savings in the number of shots comes simply from a realization that not all shots contribute the same number of traces to the subsurface image. In the real case a more important consideration would be *to what part of the image* every shot contributes. Those shots that contribute to the reservoir location (or any other critical part of the image) will be kept even if they contribute only a small number of traces. This flexibility is important when faced with obstacles which force us to displace shots or receivers. The effort that we put into it may depend on the relative contribution of those shots and receivers to the critical parts of the image as opposed to the standard approach in which all shots and receivers are considered equally important.

In order to see the importance of the fewer shots in the quality of the image, I modeled the data again with the standard approach but using only 402 shots (the same that I used in the proposed approach). The first shot will now be at -5350 m which translates to a maximum dip angle of 69 degrees with one fold and 60 degrees with full fold. Figure 13 shows a comparison

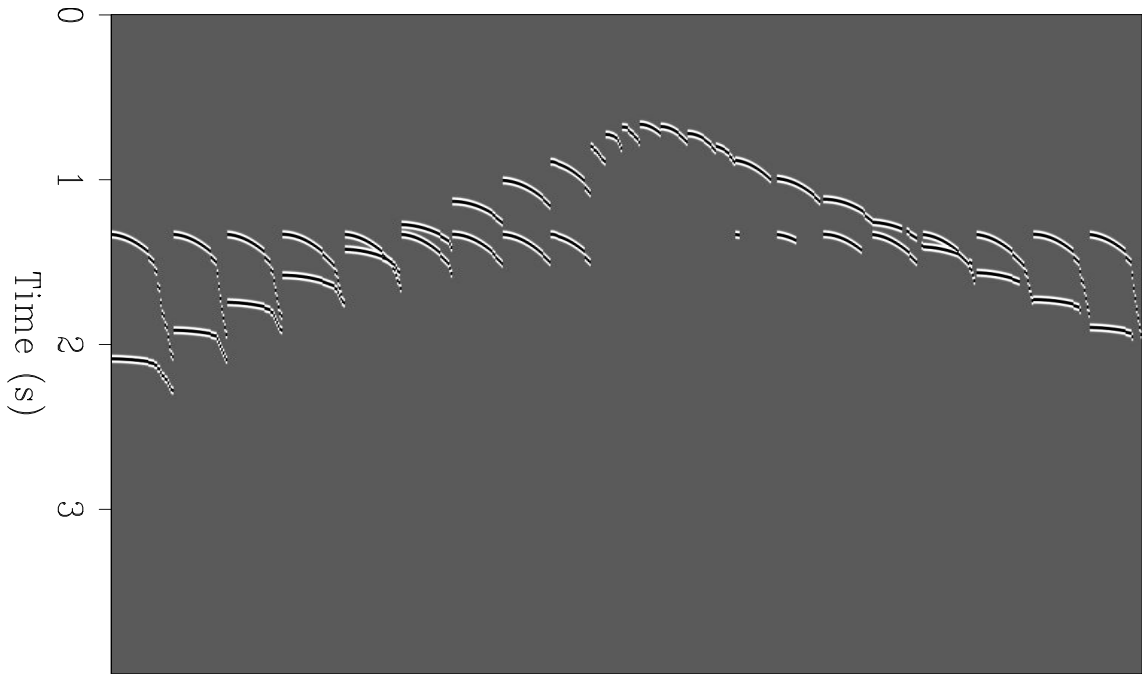



Figure 8: Selected CMPs modeled with the proposed methodology. gabriel2-modcirc3_cdps [ER]

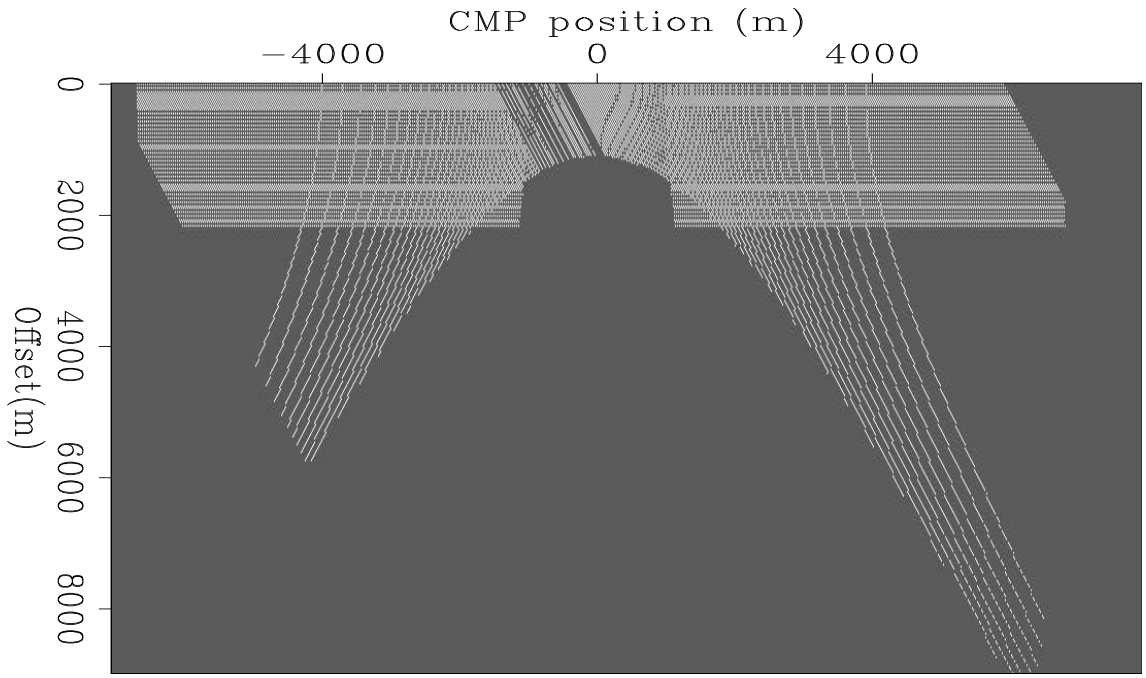



Figure 9: Fold diagram for the proposed methodology gabriel2-modcirc3_fold [ER]

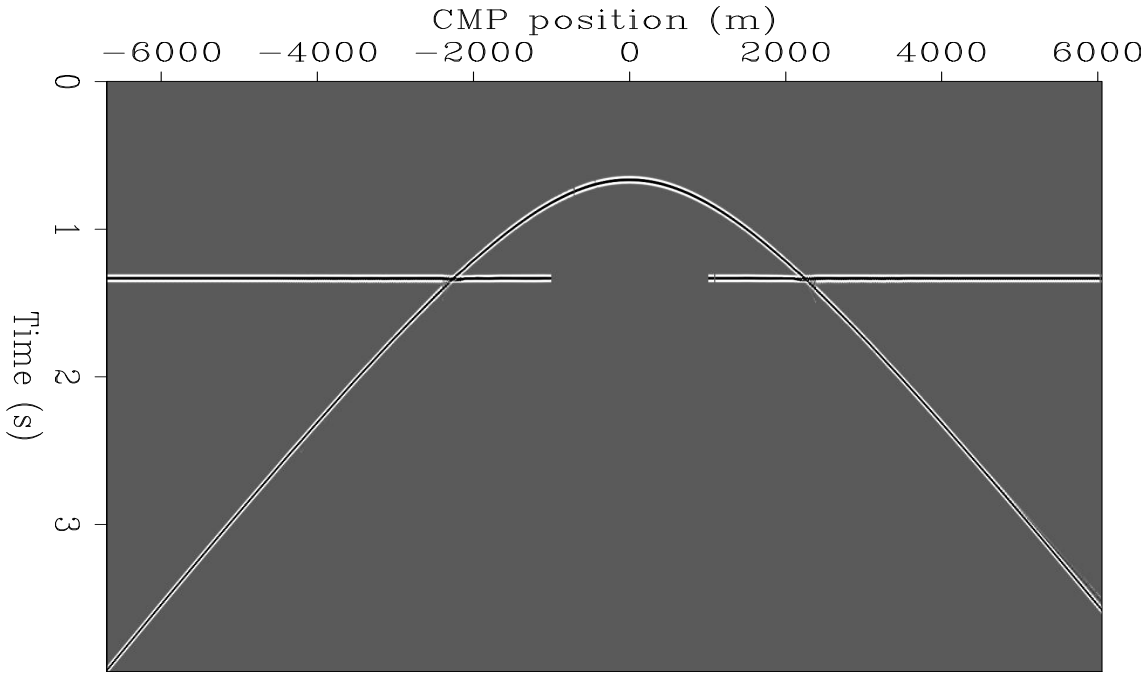



Figure 10: Stacked section of modeled data generated with the proposed design gabriel2-modcirc3_stack [ER]

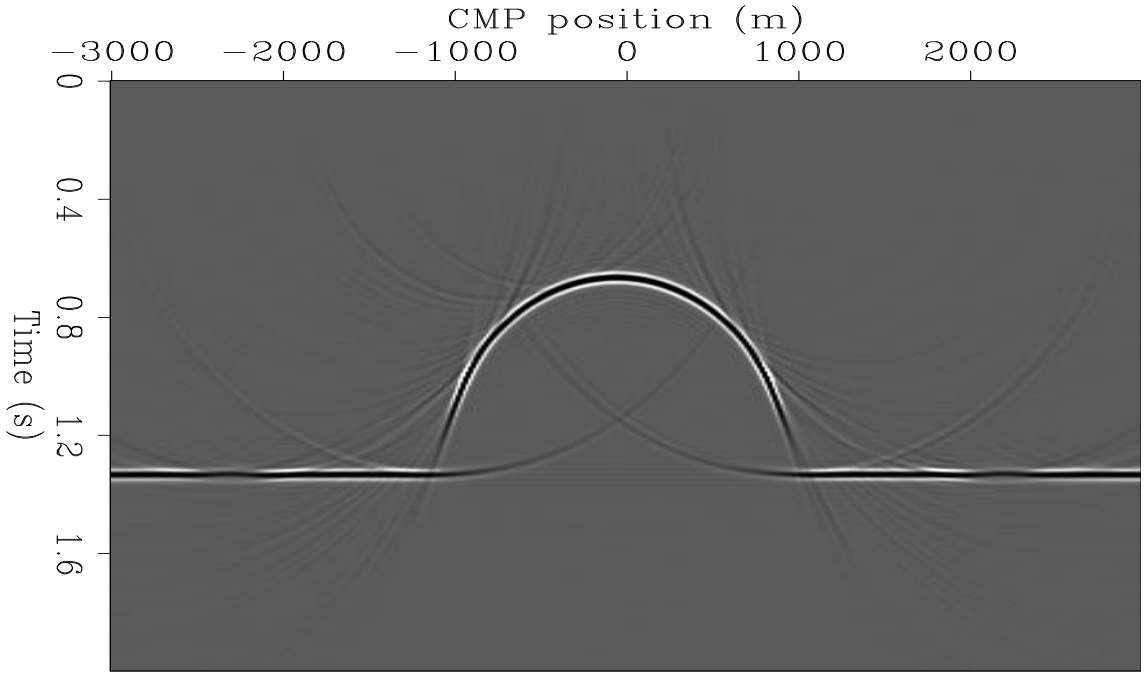



Figure 11: Post-stack migrated section of modeled data generated with the proposed design gabriel2-modcirc3_mig [ER]

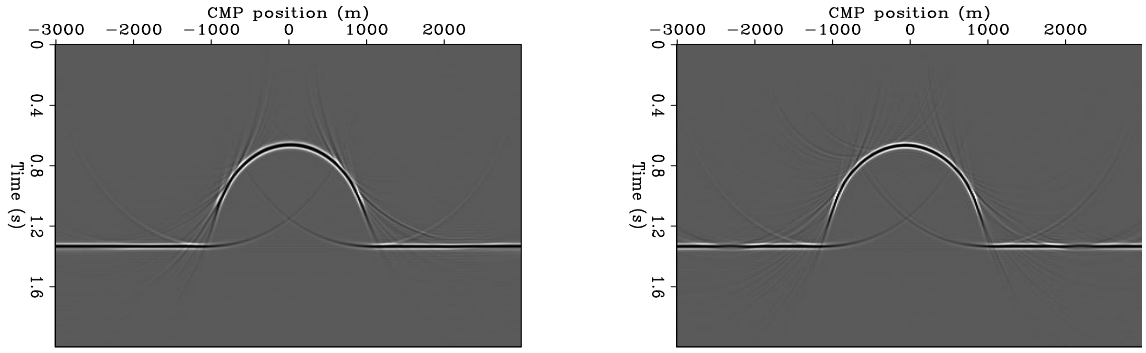



Figure 12: Close up comparison of post-stack migrated sections generated with the traditional (left) and proposed (right) design gabriel2-modcirc13_comp [ER,M]

with the proposed approach. The difference in the high dips of the images on the left-hand-side of the semicircle is clearly visible.

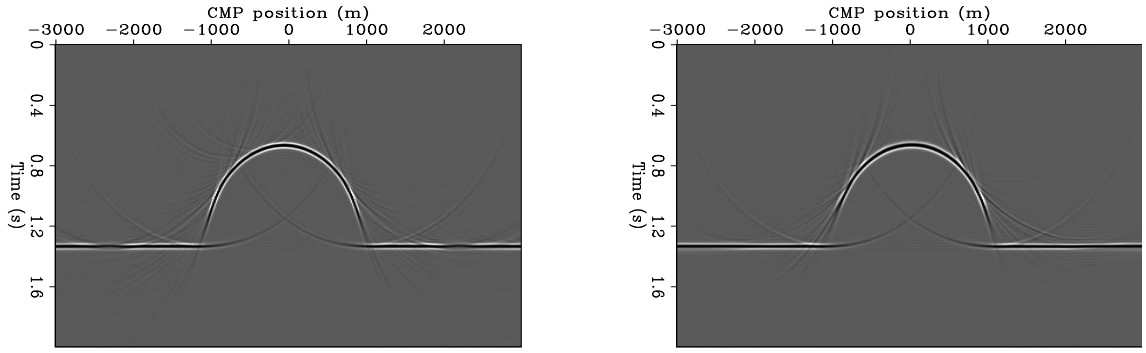



Figure 13: Close up comparison of post-stack migrated sections generated with the proposed design (left) and the standard design with the same number of shots. gabriel2-modcirc13_comp2 [ER,M]

An obvious improvement to the above methodology consists in acquiring, for every shot, not only those receiver positions obtained form the inversion, but also those in between. After all, if the intermediate receiver stations are available, why not use them? Figure 14 shows the fold diagram in this case. The number of shots is the same as in the previous case, and the increase in fold is due entirely to the intermediate receivers.

## DISCUSSION

It should be obvious that the real difference between the standard and the proposed methodology will only surface when the data is imaged before stack. Then the larger offsets corresponding to the larger dips and the better illumination will really come into play. In this paper I chose to image the data post-stack only for simplicity. Also, it should be apparent that the expected

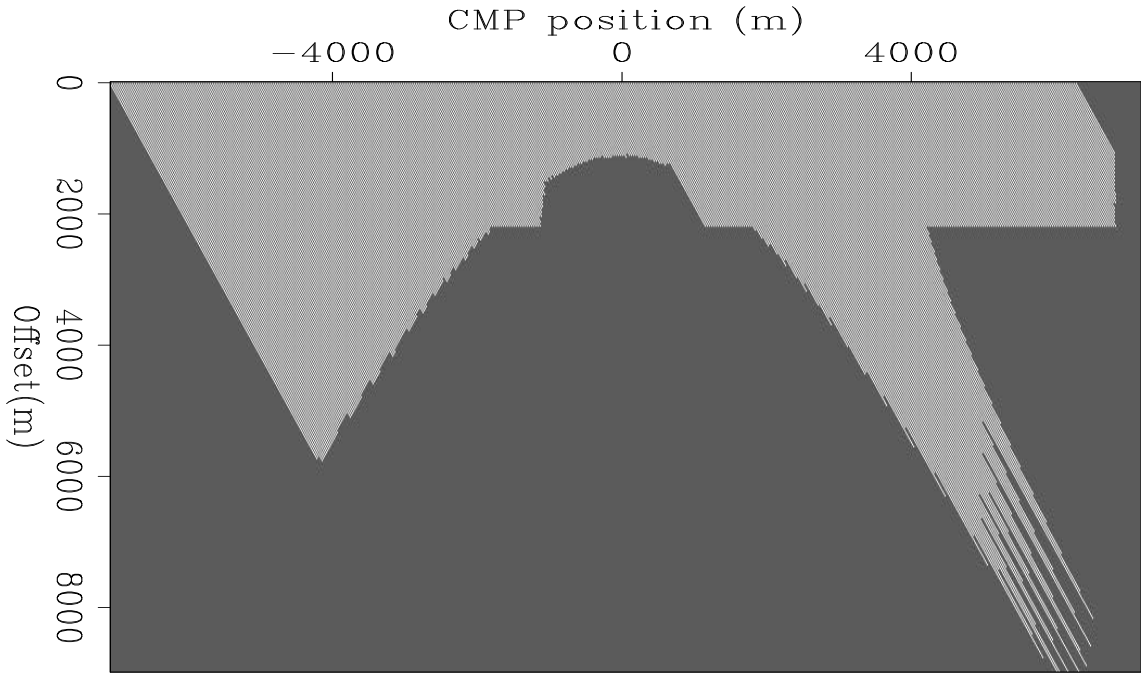



Figure 14: Fold chart of modeled data generated with the proposed design using intermediate traces gabriel2-modcirc4_fold [ER]

differences between the standard and the proposed approach will be an order of magnitude larger for 3-D data.

I will now addressed some questions that are important and whose answers may not be immediately obvious or may be a matter of conjecture.

1. Why not simply use all the available receivers for each shot and forget about optimum designs?

   In 2-D, with todays recording equipment having thousands of available channels, it may be reasonable to consider a "brute force" approach in which receivers for the whole line are deployed on the ground at the beginning of the acquisition and offsets 2, 3 or 5 times larger than the maximum target depth are used for every shot. This is clearly overkill and relies on the assumption that during processing the useful offsets will be "sorted out." The problem is that at processing it may be too time consuming to determine which traces to actually keep for each shot or CMP. As a result, we may end up with lots of essentially useless traces getting in the way of efficient processing.

   In 3-D the brute-force approach will not work at all for any reasonably large survey because of the sheer volume of equipment it would require. Even if the recording equipment itself can handle all the channels, there are still the problems of the cables, the recording boxes, the communication channels, the antennas, the batteries, and so on.

2. What are the implications for the logistics of operation of the proposed methodology? The logistics of operation do not need to be strongly affected since the receivers will

be deployed as usual and the recording equipment will electronically connect and disconnect the required receiver stations for each shot based on the information of the geometry (SPS) files. Thus, the fact that the template may change from shot to shot (or from salvo to salvo in 3-D) is not a negative logistics issue.

3. Why acquire the intermediate traces? If we already computed the optimum positions for sources and receivers so as to have "perfect" illumination what valuable information can there be in the intermediate traces? Why go to all the trouble of finding optimum receiver positions if we are going to use the intermediate receivers as well?

   The intermediate traces, although possibly contributing redundant illumination, will be useful for random noise suppression, for velocity computations and for offset sampling necessary for prestack migration. Besides they will not require any significant extra acquisition effort.

4. What are the implications of the non-uniform offset distribution for prestack migration of the data? Can we guarantee that there will not be spatial aliasing in the offset dimension?

   This is an open question for which I don't have a definitive answer yet. The idea is to include the sampling requirements for prestack migration as constraints to the inversion process, so that additional receiver or shot positions be considered to satisfy that constraint. The details of how to do that, especially in 3-D, is an interesting research issue.

5. What would be the situation for 3-D acquisition?

   For 3-D the situation is more challenging but also much more interesting and useful. We have now not only the degree of freedom afforded by the choice of offsets but by the choice of azimuths as well. Besides, the basic geometry template can also be considered a design parameter that (unlike common practice) can change spatially. I anticipate that the inversion process will be extremely difficult and strongly non-unique.

A more philosophical question, but one that has an important meaning is: What kind of data would we regard as ideal from the point of view of imaging? That is, assuming that we have no logistic or economic restriction whatsoever (except that the data can only be acquired at the surface), what would be the ideal data?

For the standard approach, some characteristics of this ideal data immediately come to mind: data in a very fine regular grid with a very large aperture. This is fine, and would provide us with a good image. I believe, however, that the ideal data would be data with very fine, regular subsurface illumination, with aperture being a function of the illumination requirements.

That subsurface illumination is an important attribute of a good design is not new. In fact most commercial software for seismic survey design offer an option to trace rays into the subsurface for a given design to produce illumination maps of the targets of interest. The maps obtained with different designs are compared and this information taken into account when

deciding what the best design is or changes may be introduced to the designs and the process iterated. This is an example of the *forward* problem. What I propose is to base the survey design on the *inverse* problem: start with an initial model and choose the layout of sources and receivers to obtain optimum illumination for that model.

## SUMMARY AND CONCLUSIONS

Using all the available subsurface information to design the acquisition parameters of new seismic surveys in a given area seems like the sensible thing to do. Common practice, however, uses only maximum and minimum values of targets velocities, dips and depths. Starting from a subsurface model may seem to bias the acquisition, but we have to keep in mind that by not using any model we are in fact imposing a model of flat layers and constant velocity.

The real impact and usefulness of this methodology arises in 3-D land seismic acquisition where the cost of the surveys oftentimes requires the design to be a compromise of the different subsurface parameters in different parts of the survey. The design is then kept constant for the whole area. By adapting the acquisition effort locally to the imaging demands of the subsurface we could in principle acquire better data at the same cost or perhaps even cheaper.

It is all too common in 3-D land data that significant obstacles force us to deviate from the original design. The common practice is to displace shots and receivers to alternative positions chosen to maintain, as much as possible, the uniformity of fold and the regularity of offsets and azimuth distribution. All sources and receivers are considered equally important to the subsurface image, which is probably not a good idea in general. The presence of large obstacles can be incorporated into the design procedure and alternative source and receiver locations chosen to optimize the regularity of the illumination as opposed to the regularity of fold, offset or azimuth distributions as is standard practice. It may very well happen that some shots in the excluded area turn out not to contribute significantly to the critical parts of the image and so can be simply ignored if they are difficult to replace. On the other hand, it may turn out that those shots are critical and then there are concrete reasons to make a stronger effort to acquire them.

In conclusion, we should be able to use all the available information when designing the acquisition of a new survey and make decisions with as much information as possible. Shifting the emphasis from surface parameters (fold, offsets and azimuths) to subsurface parameters (illumination) is a step in the right direction.

## THE WAY AHEAD

Clearly, the methodology described in this paper is only weakly supported by the simple 2-D example presented here. The challenge lies in the implementation of the non-linear inversion process capable of computing the optimum source and receiver positions given all the geophysical, logistical and financial constraints for 3-D data.

The sensitivity of the acquisition geometry to the accuracy of the initial structural and velocity model is also clearly an important issue to be analyzed in detail. Extensive research will have to be done before clear guidelines can be established in that respect.

## REFERENCES

Cain, G., Cambois, G., Gehin, M., and Hall, R., 1998, Reducing risk in seismic acquisition and interpretation of complex targets using a GOCAD-based 3-D modeling tool: 68th Annual Internat. Mtg., Society of Exploration Geophysicists, Expanded Abstracts, 2072–2075.

Carcione, L., Pereyra, V., Munoz, A., Ordaz, F., Torres, R., Yanez, E., and Yibirin, R., 1999, Model-based simulation for survey planning and optimization: 69th Annual Internat. Mtg., Society of Exploration Geophysicists, Expanded Abstracts, 625–628.

Galbraith, M., 1994, Land 3-D survey design by computer: Land 3-D survey design by computer:, Austr. Soc. Expl. Gephys., Explo. Geophys., 71–78.

Liner, C. L., and Underwood, W. D., 1999, 3-D seismic survey design as an optimization problem: The Leading Edge, **18**, no. 9, 1054–1060.

Morrice, D., Kenyoun, A., and Beckett, C., 2001, Optimizing operations in 3-D land seismic surveys: Geophysics, **66**, 1818–1826.

Stone, D., 1994, Designing seismic surveys in two and three dimensions: Society of Exploration Geophysicists.

Vermeer, G., 2001, Fundamentals of 3-D seismic survey design: Ph.D. thesis, Delft University of Technology.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 323–333*

# Non-stationary, multi-scale prediction-error filters and irregularly sampled data

*William Curry*[1]

**ABSTRACT**

Non-stationary prediction-error filters have previously been used to interpolate sparse, regularly sampled data. I take an existing method used to estimate a stationary prediction-error filter on sparse, irregularly sampled data, and extend it to use non-stationary prediction-error filters. I then apply this method to interpolate a non-stationary test case, with promising results. I also examine a more complex three dimensional test case.

## INTRODUCTION

Data interpolation can be cast as in inverse problem, where the known data remains constant, and the empty bins are regularized to constrain the null space. A two-stage linear approach was developed (Claerbout, 1999) where a prediction-error filter (PEF) is estimated on known data, and is then used to constrain the unknown data by minimizing the output of the model after convolution with the PEF. When the data is not stationary, a non-stationary filter has been used to fill the unknown data (Crawley, 2000). This gives better results than a patching approach, where the data is broken up into separate patches that are assumed to be stationary, largely because most data is smoothly non-stationary. In the case of sparsely (but regularly) sampled data, the non-stationary filter can be stretched over various scales to fit the data. Most recently, a PEF was estimated on irregularly sampled data by scaling the data to various grid sizes and simultaneously estimating a single filter on the various scales of data in a multi-scale approach (Curry and Brown, 2001).

Here, I take the multi-scale approach for irregular data, and extend it to estimate a non-stationary PEF. I examine how to choose the parameters needed for this non-stationary PEF estimation, namely micro-patch size, scale choice, regularization, and filter size, and how they are related when using this estimation method. I use this approach to interpolate a poorly sampled 2D test case, where existing methods would fail, with promising results. I then interpolate a suitable 3D test case with very promising results, with the eventual goal of seismic data interpolation.

[1]**email:** bill@sep.Stanford.EDU

## BACKGROUND

A PEF is traditionally estimated by minimizing the output of the known data **d**, convolved (**D**) with a filter **f** which is unknown except for the first coefficient, which is constrained to 1. This is expressed below, with **K** representing a mask which is 1 when all filter coefficients lie on known data, and 0 when coefficients lie on missing data.

$$\mathbf{KDf} + \mathbf{d} \approx \mathbf{0} \tag{1}$$

Fitting goal 1 works well for estimating the PEF if there are sufficiently contiguous data. However, if the data are irregularly sampled so that there are an inadequate number of fitting equations, a different method is used (Curry and Brown, 2001), where more fitting equations are generated by regridding the data (**D**) with an operator $\mathbf{S_i}$ (normalized linear interpolation followed by its adjoint), and then simultaneously estimating a single filter (**f**) on all of the versions of the scaled data ($\mathbf{S_i d}$). The mask for the known data ($\mathbf{K_i}$) must also be regridded accordingly for each scale.

$$\begin{bmatrix} \mathbf{K_0 D} \\ \mathbf{K_1 S_1 D} \\ \mathbf{K_2 S_2 D} \\ \mathbf{...} \\ \mathbf{K_n S_n D} \end{bmatrix} \mathbf{f} = \begin{bmatrix} \mathbf{d} \\ \mathbf{S_1 d} \\ \mathbf{S_2 d} \\ \mathbf{...} \\ \mathbf{S_n d} \end{bmatrix}. \tag{2}$$

Several user-defined parameters must be set during this procedure, specifically the choice of scales to be used in the estimation as well as the size of the PEF. The only constraint on these parameters is that the aspect ratio of the data remain constant from scale to scale, meaning that the ratio of the number of bins in each dimension remain constant. For example, a $50 * 40$ data set should not be regridded to $26 * 21$, since the aspect ratio is changed by the round-off from 20.8 to 21. A better choice would be to use $25 * 20$ as a scale. The PEF size is only constrained by the size of the coarsest scale of data.

A non-stationary filter varies with position, so instead of only having indices corresponding to the lag of the filter, there are also indices corresponding to the position of the filter. The filter would go from looking like `a(ia)` to `a(ia,id)`, where `ia` is the lag of the filter, and `id` is the position of the filter. An illustration of this concept is shown in figure 1. Only two indices are used for lag and position, thanks to the helical coordinate system (Claerbout, 1998).

Since we have moved from estimating a single filter with `na` filter coefficients to estimating a non-stationary filter with `na*nd` filter coefficients, PEF estimation becomes an under-determined problem instead of an over-determined problem. As a result, we need to incorporate some type of regularization into the estimation in order to get enough equations. Laplacian or radial roughener of common filter coefficients (constant `ia`) across the spatial axes (`id`) are both used to ensure a filter bank that varies smoothly spatially (Clapp et al., 1999).

Another method used to constrain the filter coefficients is called micro-patching (Crawley, 2000). Instead of the filter varying at every data point, micro-patching uses the same filter within a small region within the data, reducing the number of filter coefficients that need to be estimated. This has two benefits: the PEF estimation problem becomes less under-determined,

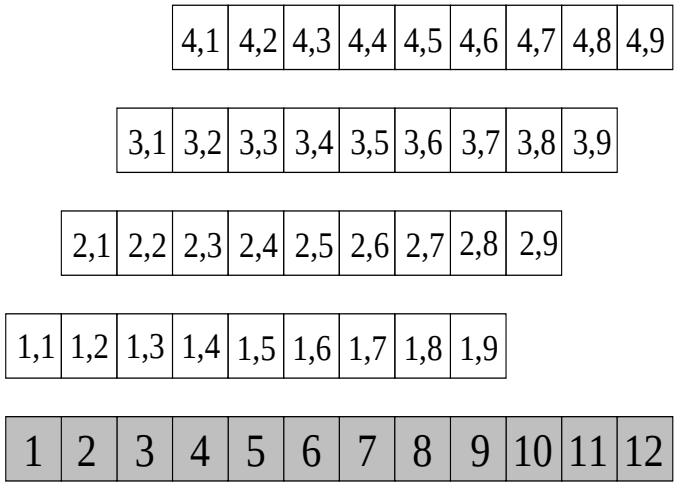



Figure 1: An illustration of non-stationary convolution. The shaded boxes represent the data, and the hollow boxes represent the filter at various positions. The two indices on each filter point correspond to the data position (`id`) and the filter lag (`ia`), respectively. At each point in the convolution, the filter is different. bill1-nstat [NR]

and the amount of memory required for the filter, which was `na` times the size of the data `nd`, is now the number of micro-patches, `np` times `na`.

## MULTI-SCALE NON-STATIONARY PEFS

The combination of non-stationary filters and estimation on multiple scales of data introduces a new issue, that non-stationary filters are linked to the size of the data they operate on. If the dimensions of the data change, the dimensions of the PEF must change as well. This causes issues regarding the consistency of the PEF across scales, as well as limits on the type of regularization that can be applied to the filter coefficients.

In order to maintain a consistent PEF across scales, the filter must be sub-sampled so that the same spatial coordinates of the PEF correspond to the proper locations within the scaled data. Since our non-stationary PEF has micro-patches where the filter coefficients are constant, we can scale the patches to match the scaling of the data, so that the number of filter coefficients in the non-stationary filter remains constant, but the size of the micro-patches has decreased. The limiting case for this scaling is when a micro-patch reduces to a single point. Beyond this point, the patches could be sub-sampled during the scaling. This avenue has not been explored, since a need for that level of scaling has not yet been shown.

I represent the sub-sampling of the patch table by $\mathbf{P_i}$, which acts upon the non-stationary filter $\mathbf{f}$ in the non-stationary fitting goal shown below:

$$\begin{bmatrix} \mathbf{K_0 D} \\ \mathbf{K_1 S_1 D P_1} \\ \mathbf{K_2 S_2 D P_2} \\ \mathbf{...} \\ \mathbf{K_n S_n D P_n} \end{bmatrix} \mathbf{f} \approx \begin{bmatrix} \mathbf{d} \\ \mathbf{S_1 d} \\ \mathbf{S_2 d} \\ \mathbf{...} \\ \mathbf{S_n d} \end{bmatrix} . \tag{3}$$

In addition to the above fitting goal, a set of regularization equations must also be solved,

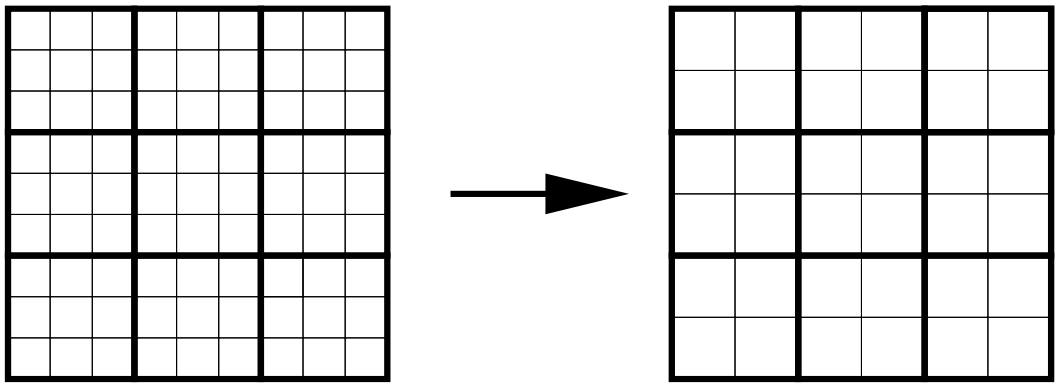


Figure 2: Regridding both the data (fine grid) and the micro-patches (thick grid) simultaneously. In this case a $9*9$ grid with $3*3$ micro-patches is regridded to a $6*6$ grid with $2*2$ micro-patches. The sizes of the micro-patches remain constant. bill1-patchscale [NR]

where **A** is our regularization operator:

$$\mathbf{Af} \approx \mathbf{0} \tag{4}$$

The scaling of both the filter and the data to some extent limits the choices of regularization available. Specifically, radial micro-patches do not scale well, so applying radial regularization would have to be done over square micro-patches. Laplacian regularization of filter coefficients across rectangular micro-patches is also a reasonable approach in some cases.

## TEST CASE

A non-stationary, two-dimensional test case has been created as a proof of concept example. This test is based upon a simple plane wave model (Brown et al., 2000; Curry and Brown, 2001).

In this case, there is one plane wave on each half of the example, and the plane wave on the right varies in amplitude from left to right. There is also Gaussian noise present in the data, which largely obscures the low amplitude portion of the plane wave on the right side of the example.

The data was randomly sub-sampled, keeping only $10-20$ percent of the data. This sparse data, along with a mask describing the position of the known data, was used in the non-stationary interpolation scheme. The results of the interpolation are shown in Figure 4.

The results are on the whole quite encouraging. The data was very heavily sub-sampled, and the interpolation scheme was able to identify the dips in the data and interpolate them properly. The dip from the left side of the example was smoothed over the area on the right side of the figure with the low signal-to-noise ratio, which was expected. As the number of micro-patches drops, the results deteriorate, as the regularization loses its effect and one patch

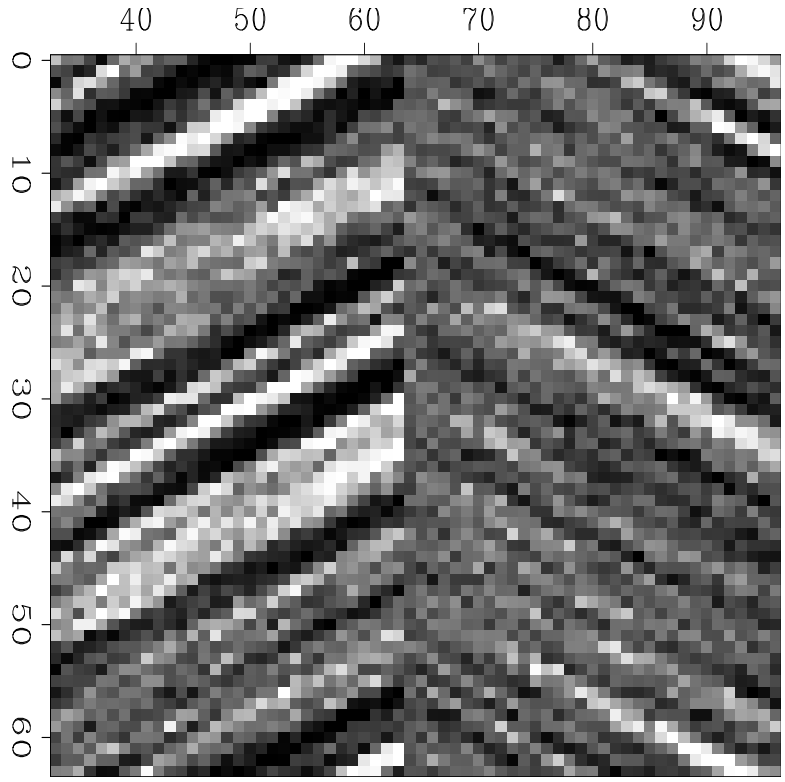



Figure 3: The fully sampled version of the non-stationary test data. bill1-testcase [ER]

becomes unstable. A higher number of micro-patches preserves the dips as well as the low amplitude area.

A more relevant case to interpolating seismic data is the qdome model (Claerbout, 1999), which has been highly sub-sampled along two of three dimensions, with the vertical axis still fully sampled. The qdome model is a collection of folding layers, flat layers and a fault, which acts as an excellent overall test for this interpolation method. I have randomly removed 88 percent of the traces from the data set, and use the non-stationary multi-scale PEF-based interpolation to attempt to recreate the original model.

The results for the qdome model are very promising. The smoothly varying dips were correctly estimated and interpolated almost everywhere, excluding very steep dips. This is due to two things, the size of the PEF might not have been large enough to capture the spatially aliased dips, and that the dips were changing rapidly within a small area, which was only covered by a small number of micro-patches.

The results for this 3D case are much more impressive than in the 2D case, even though more of the data was removed. This is due to several things: The extra dimension of data allows for more constraints to be applied by the regularization, the greater size of all of the dimensions allows for more fitting equations to be found, and most importantly that the well-sampled z-axis gives better coverage of the data space.

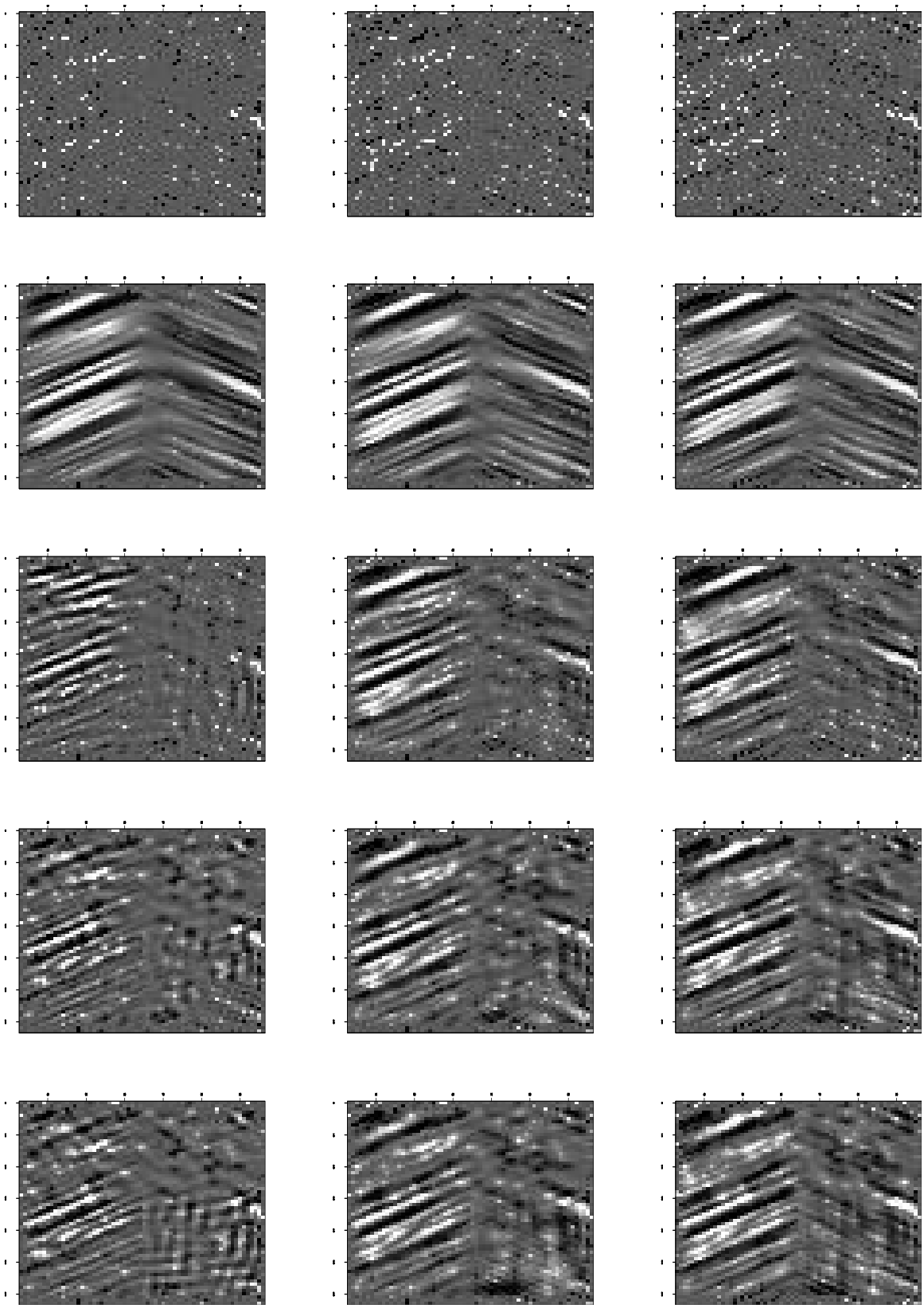


Figure 4: Results for three different sampling levels, from left to right: 10, 15 and 20 percent. From top to bottom: sparse data, sparse data filled with PEF trained on fully sampled original data, sparse data filled with PEF from sparse data with 64, 16 and 8 micro-patches, respectively. bill1-testfill [ER]

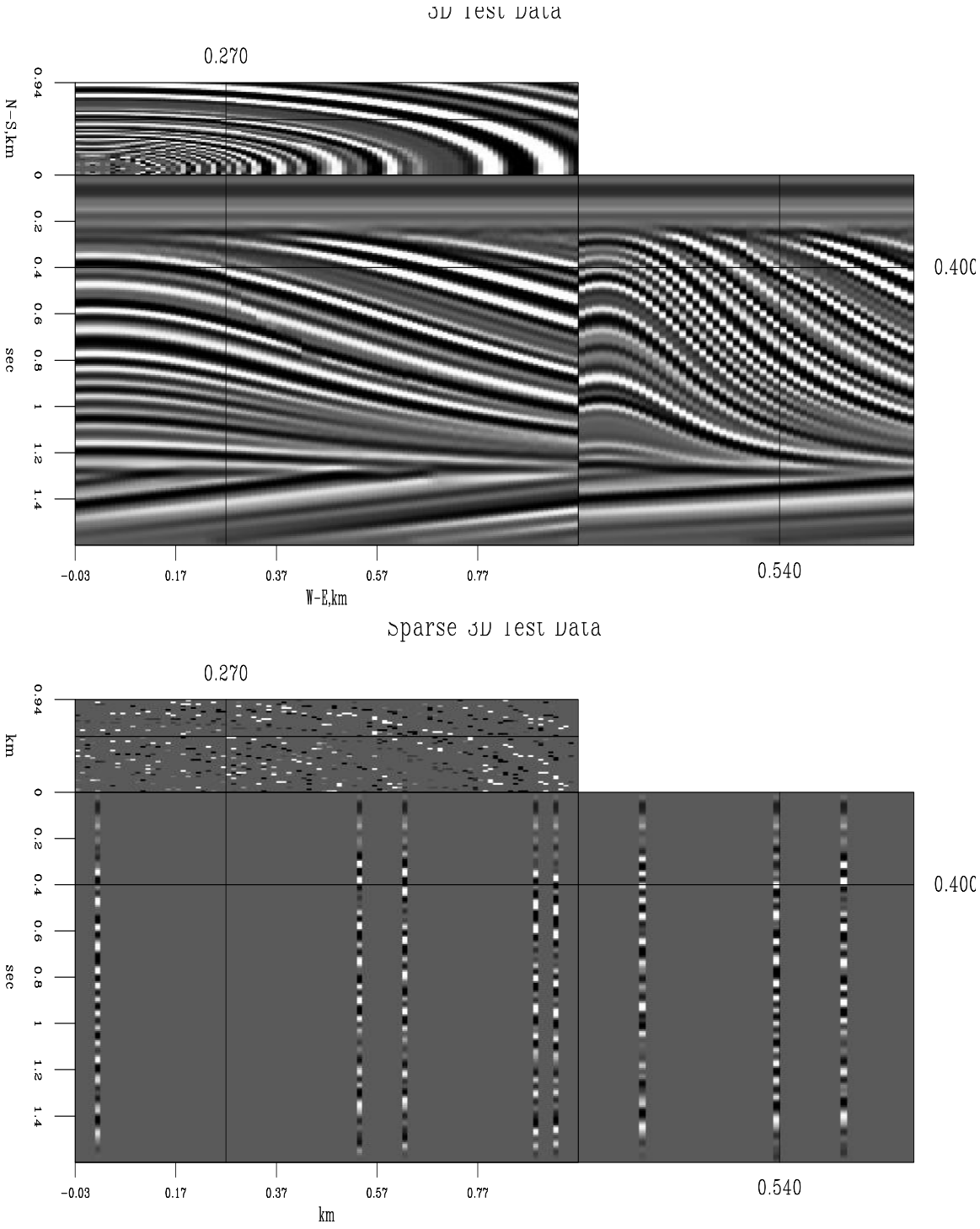



Figure 5: Fully sampled and sub-sampled versions of the qdome model. bill1-qdometest [ER]

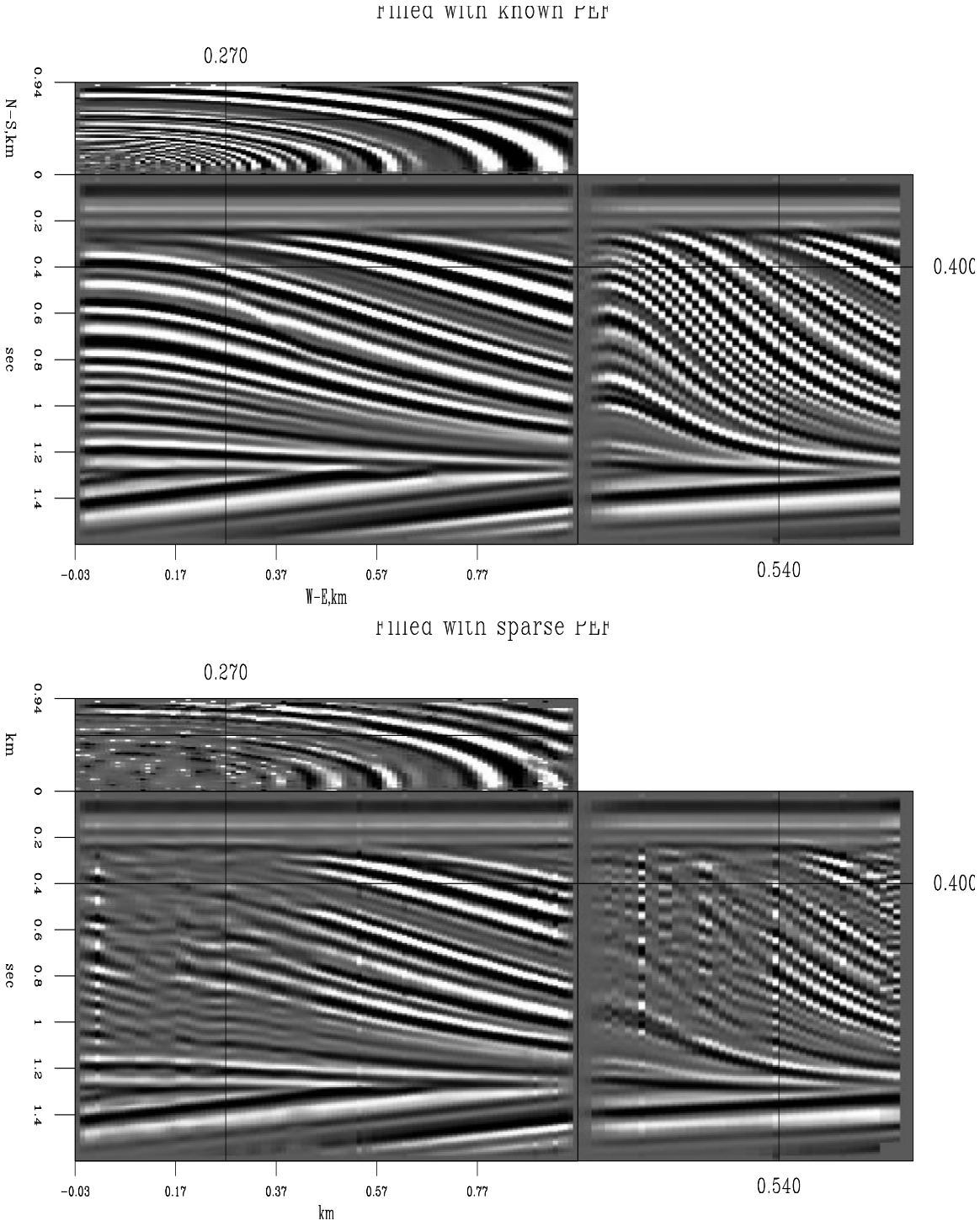



Figure 6: Above: Sparse data filled with a non-stationary PEF trained on all data. Below: Sparse data filled with a non-stationary PEF trained on sparse data. bill1-qdomefill [CR]

## CONCLUSIONS AND FUTURE WORK

Overall, estimating a non-stationary prediction-error filter with multiple scales of data appears to be successful. The method interpolates a very heavily decimated 2D test case successfully. Results for a 3D case are even more successful, even though the amount of data removed from the case was greater than in the 2D case.

In the future, this method can be used on real seismic data in two, three, or even five dimensions, so that prestack 3D data can be interpolated in cases where surface topography, structures, or other obstacles cause irregularly sampled data.

## REFERENCES

Brown, M., Claerbout, J., and Fomel, S., 2000, Test case for PEF estimation with sparse data II: SEP–**105**, 117–122.

Claerbout, J., 1998, Multidimensional recursive filters via a helix: SEP–**97**, 319–335.

Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project, `http://sepwww.stanford.edu/sep/prof/`.

Clapp, R. G., Fomel, S., Crawley, S., and Claerbout, J. F., 1999, Directional smoothing of non-stationary filters: SEP–**100**, 197–209.

Crawley, S., 2000, Seismic trace interpolation with nonstationary prediction-error filters: Ph.D. thesis, Stanford University.

Curry, W., and Brown, M., 2001, A new multiscale prediction-error filter for sparse data interpolation: SEP–**110**, 113–122.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 333–345*

# Madagascar satellite data: an inversion test case

*Jesse Lomask*[1]

**ABSTRACT**

The Madagascar satellite data set provides images of a spreading ridge off the coast of Madagascar. This data set has two regions: the southern half is densely sampled and the northern half is sparsely sampled. This data set is an excellent test case for inversion methods. It presents several challenges that geophysicists face in generating seismic maps in general. The data is acquired in swaths that follow irregular paths(tracks), similar in some respects to irregular 3D acquisition geometries. Inversion allows us to combine these different data paths into one image. Shifts between tracks are removed by taking the derivative along the tracks in the inversion fitting goals. By looking at the residual in data-space, we were able to see errors in the weighting operator. The sparsely sampled region presents a missing data problem. In the future, we intend to estimate 2D prediction error filters (PEFs) on these sparse tracks and use them to fill in the missing data. I have tested one method on a simple 1D model, in which I estimate a PEF and missing data simultaneously while throwing out fitting equations where the leading 1 coefficient of the PEF lands on unknown data. Thus, this data will give us an opportunity to test different methods of estimating PEFs on sparse and irregular data. Also, preconditioning on the helix greatly speeds convergence.

## INTRODUCTION

To a certain extent, the surface of the ocean is a gravitational equipotential surface. Mountains and ridges beneath cause bulges in the sea surface which reflect the topography beneath. With each pass, the GEOSAT satellite measures thin swaths of the height of the sea-level above some reference ellipsoid. Numerous passes can be combined to create a map of the ocean floor topography (Lomask, 1998). The altimetry of the sea surface is also influenced by tidal fluctuations and currents that overwhelm the high frequency bulges that we are interested in imaging. Any two adjacent or crossing tracks are mismatched by a low or zero frequency shift.

The satellite data consists of four different sets, each of which is a one-dimensional array of tracks connected end to end. There are two sets of sparsely sampled tracks, one north flying and one south flying. There are also two sets of densely sampled tracks, again one north flying (ascending) and the other south flying (descending). The sparse tracks cover the same region as the dense tracks plus a region of equal size to the north. Figure 1 shows the geometry of the data separated into sparse tracks and dense tracks.

[1]**email:** lomask@sep.Stanford.EDU

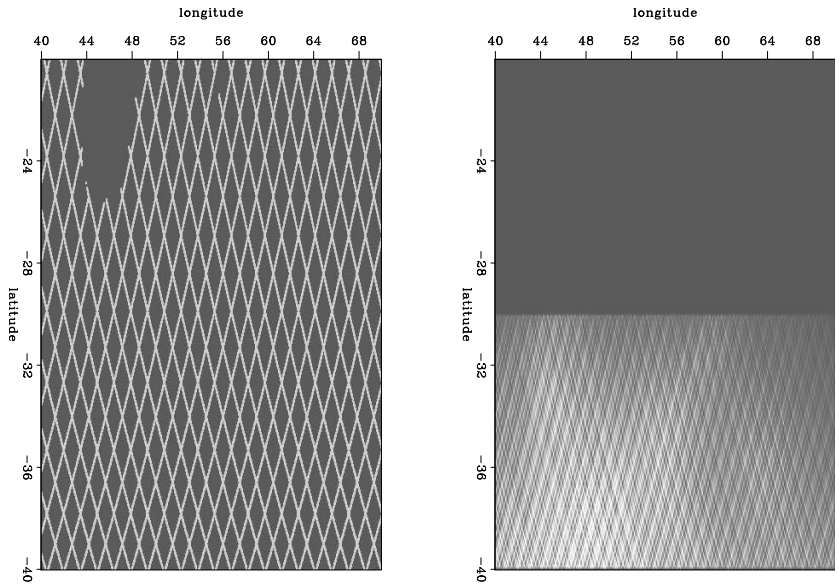



Figure 1: The raw binned data. Left is the raw normalized sparse tracks. The gap at the top is the island of Madagascar. Right is the raw dense tracks. jesse2-tracks [ER]

Using a weighted least squares approach, Ecker and Berlioux (1995) imaged the dense southern region. They used a derivative along the data tracks to remove the shifts between each pass. Some work has been done been done to image the sparse tracks. Lomask (1998) estimated 2D prediction error filters (PEFs) on the dense region and applied them in regularization operators on the sparse region. In this paper, I review how inversion is used to image the Madagascar data. Also, I describe methods for estimating and applying PEFs on sparse and irregular data.

## A good test case for inversion

The Madagascar satellite data is an excellent test case for seismic inversion methods because it poses several challenges that occur regularly in making seismic maps. This data was recorded on thin, curved swaths creating complications similar to the irregular geometry commonly seen in 3D seismic acquisition. This could be cable feathering on marine data or irregular land geometries cause by obstructions. As will be shown in more detail later, the fact that there are both ascending and descending tracks presents a problem of combining two different types of data to image one geological event. Furthermore, this data has two data densities, the northern sparsely sampled region and the southern densely sampled region. Balancing the weighting of these two data densities adds complexity to the inversion fitting goals. Additionally, the sparsely sampled region poses a challenging missing data problem especially because the sparse tracks are not oriented parallel to any axis. Lastly, the Madagascar data makes a good test case because of the clarity of the image itself. If the data is processed incorrectly, it is usually obvious that a mistake was made. However, as we will see later, this is not always the case.

## BACKGROUND

In petroleum geophysics, one geological event can be imaged with multiple data types. For instance, the top of a reservoir can be mapped with multiple seismic surveys and multiple well log runs. Inversion can be used to create a single model of the top of the reservoir to satisfy these diverse data. The Madagascar data set can be viewed as a similar problem. Figure 2 illustrates this point on only the densely sampled data. The Madagascar data can be imaged by taking the derivative along the tracks and binned. The derivative along both the ascending tracks and the descending tracks individually provides an image of the ridge. Combining both data sets to create one model of the ridge can be achieved with inversion.

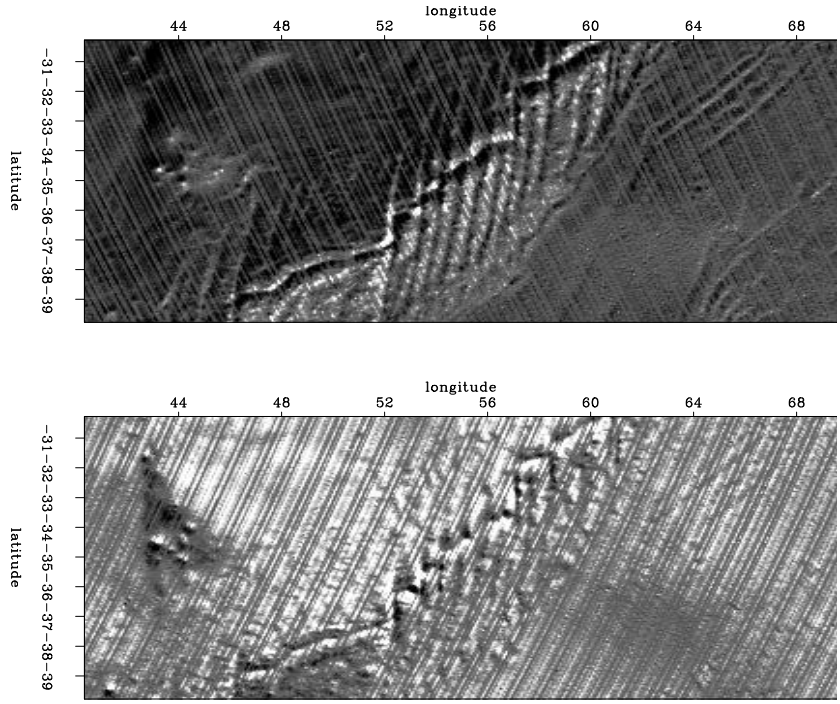



Figure 2: Dense data: ascending (top) and descending (bottom) tracks with gradient jesse2-updwnder1 [ER,M]

### Fitting goals

As described in Lomask (1998), we use inversion to find the model (**m**) that when sampled into data-space using linear interpolation (**L**) will have a derivative ($\frac{\mathbf{d}}{\mathbf{dt}}$) that will equal the derivative of the data (**d**). **W** is a weighting operator that merely throws out fitting equations that are contaminated with noise or track ends. Expressed as a fitting goal, this is :

$$\mathbf{W}\frac{d}{dt}[\mathbf{Lm} - \mathbf{d}] \approx \mathbf{0}. \tag{1}$$

For the sparse tracks or missing bins, we add a regularization fitting goal to properly fill in the data. Our fitting goals are now:

$$\begin{aligned} \mathbf{W}\frac{d}{dt}[\mathbf{Lm}-\mathbf{d}] &\approx \mathbf{0} \\ \epsilon\mathbf{Am} &\approx \mathbf{0}. \end{aligned} \tag{2}$$

Applying these goals on the dense data, we get the smooth result in Figure 3.

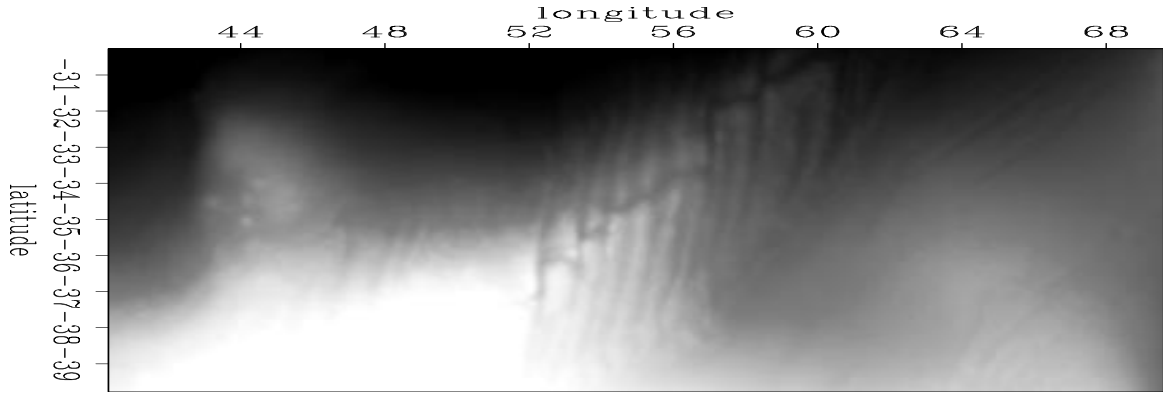



Figure 3: Result of applying fitting goals (equation 2). jesse2-fulldenseNoRuf [ER]

To make it look more interesting, we roughen the model by taking the first derivative in the east-west direction as in Figure 4. This highlights alot of the minor north-south oriented ridges. Similarly, we can roughen it in the north-south direction as in Figure 5. This highlights the central main ridge. Applying the helical derivative, we get the results in Figure 6. Unlike the directional derivative operators, this highlights features with less directional bias. Lastly, we take the east-west second derivative to get the very crisp image in Figure 7.

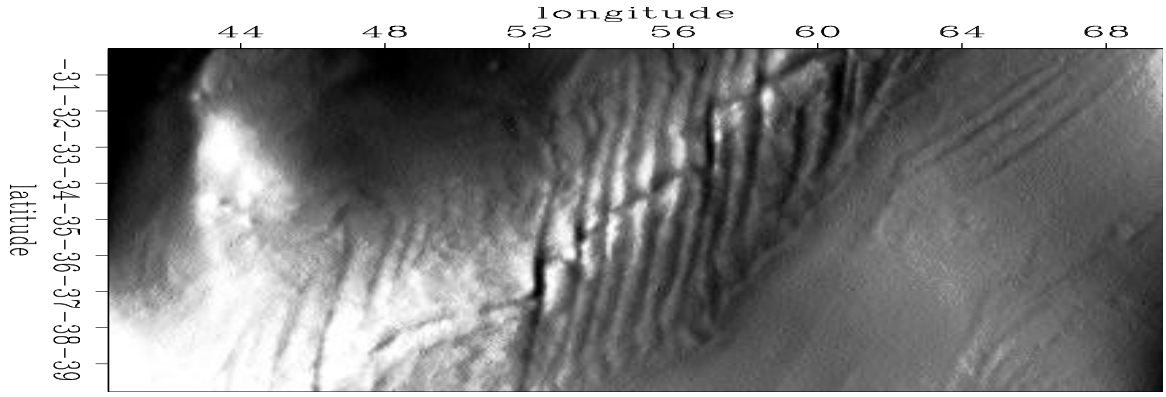



Figure 4: Results of fitting goals with east-west derivative. jesse2-fulldense1 [ER]

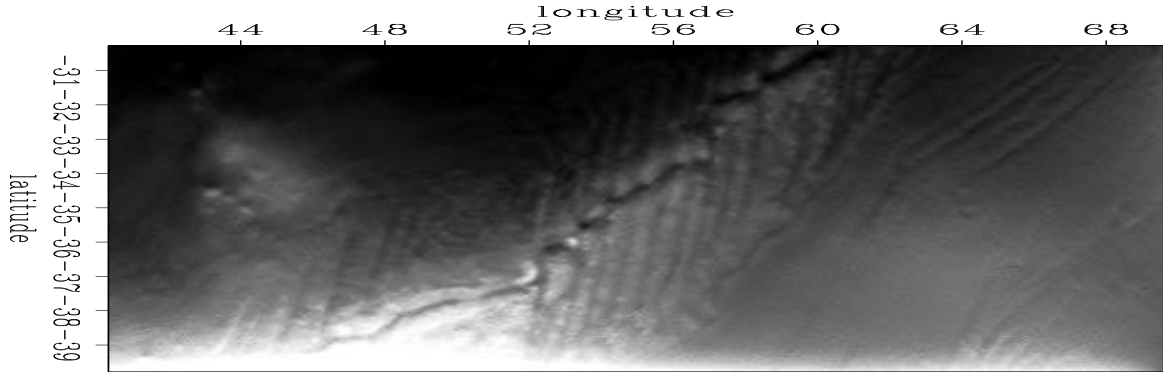



Figure 5: Results of fitting goals with north-south derivative. jesse2-fulldenseNS1 [ER]

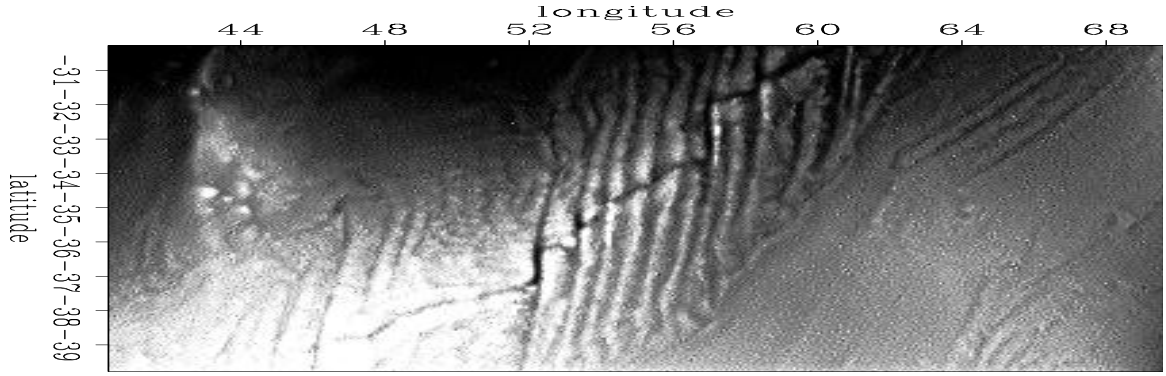



Figure 6: Results of fitting goals with helix derivative. jesse2-helix [ER]

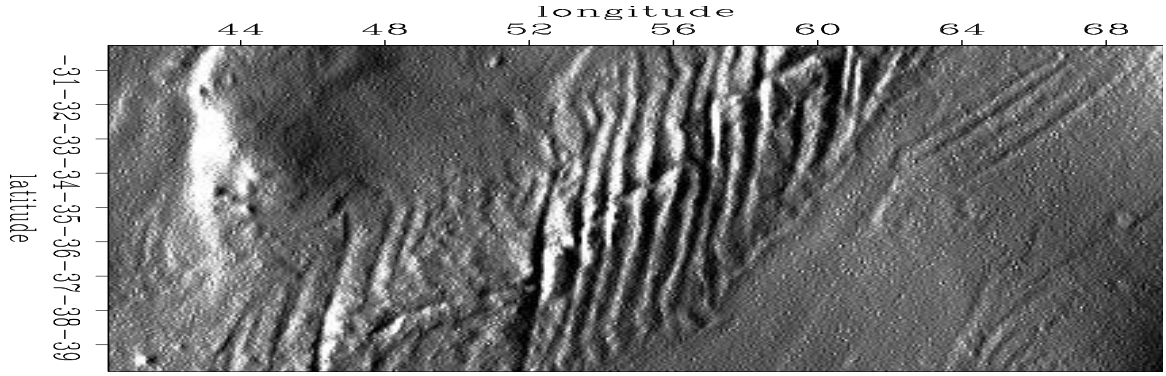



Figure 7: Results of fitting goals with east-west 2nd derivitive. jesse2-fulldense2nd1 [ER]

## LOOKING AT DATA-SPACE

Even in the case of the Madagascar data set, just looking at the output model is not enough to determine the quality of the result. In this case, there were inaccuracies in the weight, **W**, that were allowing some bad data points to slip into the fitting equations. There were only a few, so the model still looked good but observation of the data-space residual alerted us to a problem. Figure 9 shows the residual after convergence. The differences between each track should be constant, or near constant, shifts, so you would expect the residual to be comprised of constant or near constant shifts. This residual has a lot of curved features which were the result of the solver trying to fit the model to some bad data and spreading the errors out over the model.

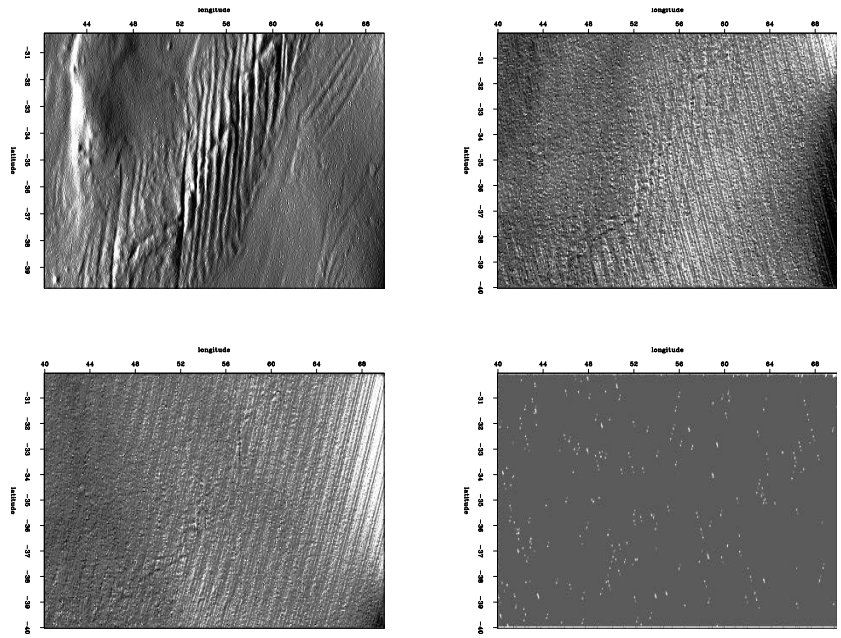



Figure 8: Top left: model roughened with east-west 2nd derivative. Top right: residual of ascending tracks with gradient applied then binned into model-space. Bottom left: residual of descending tracks with gradient applied then binned into model-space. Bottom right: binned weight. jesse2-badmovie [ER,M]

To see if there was a model space correlation in the position of the residual curves, I separated the residual into ascending and descending tracks, took the derivative along the tracks, and binned them. I took the derivative along the tracks to highlight steep or curved features. Figure 8 shows the binned residual. Notice the brightening and dimming on the right side of the ascending and descending tracks, respectively. This means that the curved features are all lining up on the right (west) side of the map. Then I looked at the binned weight and right away saw that there were no 0's on the right side of the map. Upon further examination it became clear that I was clipping the weight in model space, thus allowing all of the track ends along the right side of the map to slip into our fitting equations.

After correcting the problem, the residual looks more like it should as in Figure 11. Also the binned residual looks whiter and without the brightening and dimming as in Figure 10.

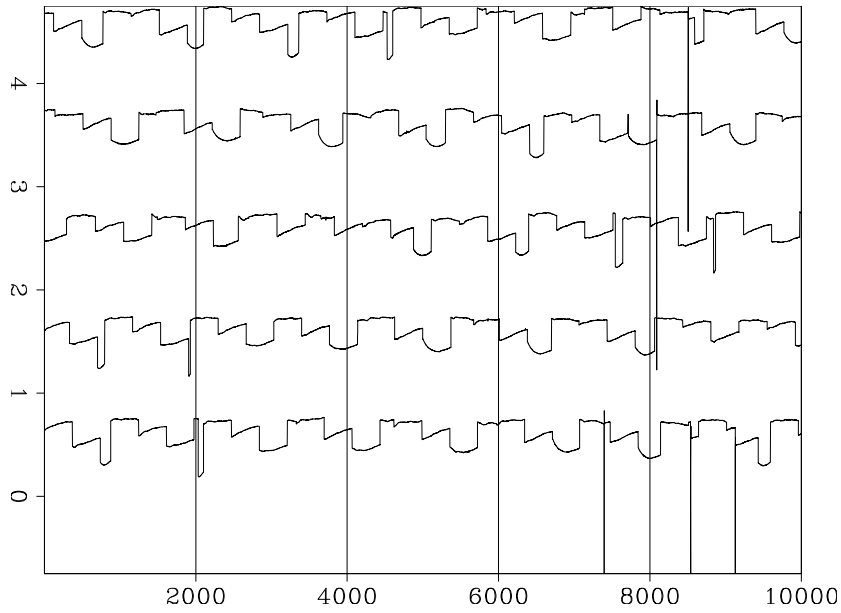



Figure 9: Residual from a group of tracks using incorrect weight. Curved features look very suspicious. jesse2-resmovie_bad [ER,M]

Figure 12 compares the model with the correct weight to the model with the incorrect weight. The difference is slight, basically just a flexing on the right side of the map.

## ESTIMATING PEF ON RESIDUAL

We used the first derivative in our fitting goal to remove the shifts between tracks. As was done previously in Lomask (1998), I tested the nature of the residual by estimating a 1D PEF on it, to see if correcting the weighting operator made difference. If the estimated PEFs are not first derivatives, then we would want to use one of them in our fitting goals instead of $\frac{\mathbf{d}}{\mathbf{dt}}$. The PEFs look like first derivatives, but in the case of the PEFs with more than three terms, it isn't obviously a first derivative. We will need to look at the impulse response of the inverse PEF to get a better idea of the nature of the PEFs.

| PEF | | Coef **1** | Coef **2** | Coef **3** | Coef **4** | Coef **5** |
|---|---|---|---|---|---|---|
| 2 terms | | 1.00 | -1.00 | - | - | - |
| 3 terms | | 1.00 | -1.06 | 0.055 | - | - |
| 4 terms | | 1.00 | -1.08 | 0.497 | -0.419 | - |
| 5 terms | | 1.00 | -1.08 | 0.284 | -0.258 | 0.050 |
| 10 terms | | 1.00 | -1.07 | 0.270 | -0.227 | 0.028 |
| | | -0.019 | -0.014 | 0.001 | -0.035 | 0.053 |

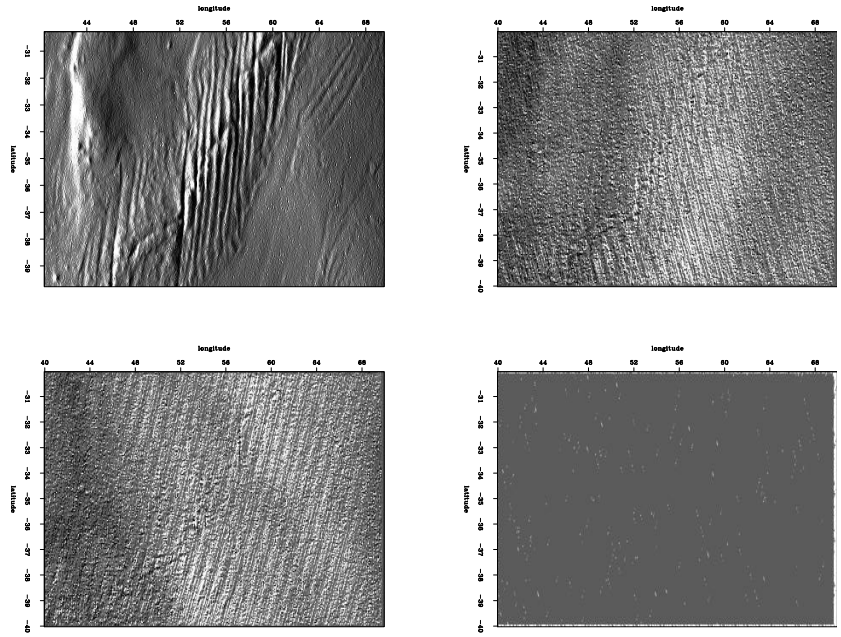


Figure 10: Top left: corrected model roughened with east-west 2nd derivative. Top right: residual of ascending tracks with gradient applied then binned into model-space. Bottom left: residual of descending tracks with gradient applied then binned into model-space. Bottom right: binned corrected weight. jesse2-goodmovie [ER,M]

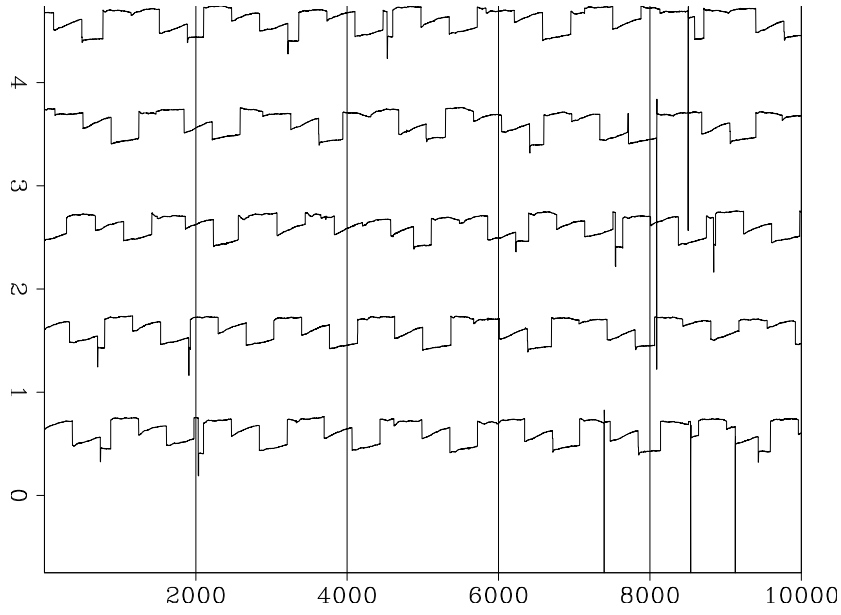



Figure 11: Residual from a group of tracks using corrected weight. jesse2-resmovie_good [ER,M]

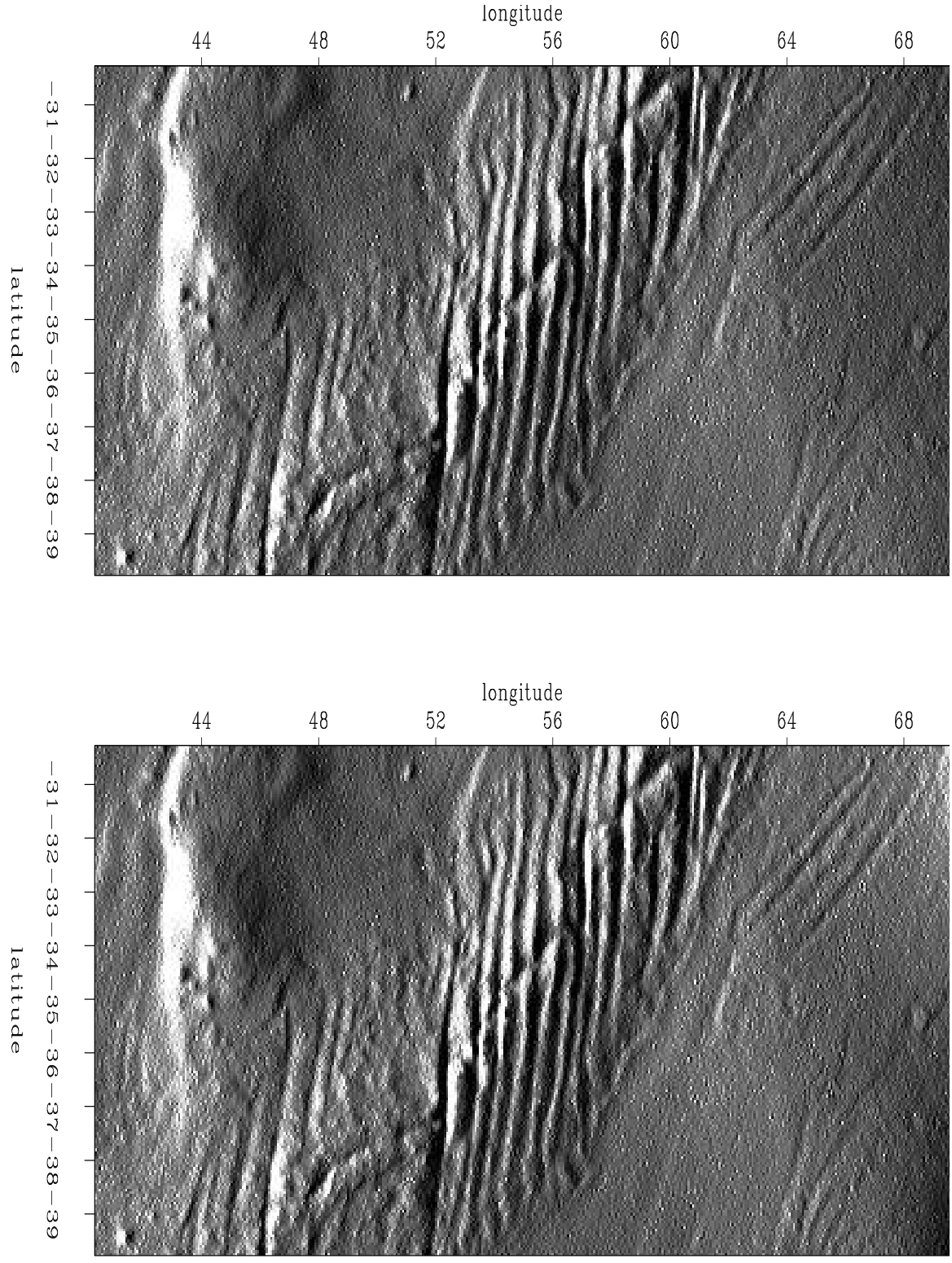



Figure 12: Top: model with corrected weight roughened with east-west 2nd derivative. Bottom: model with original weight roughened with east-west 2nd derivative. jesse2-goodandbad [ER,M]

## DEALING WITH SPARSE DATA

The sparse tracks present a problem that is seen a lot in geophysics. We have a non-stationary geological surface, in this case the surface of the ocean. We have missing data, in this case the gaps between the tracks. Our known data is along irregular geometries, the sparse tracks themselves. Previous results in the sparse region were regularized with PEFs estimated on the dense region (Lomask, 1998). Now, we would like to deal only with the sparse data. Our goal is to estimate a filter on the the known sparse data and use it to fill the missing data. Ideally, we could estimate two 1D filters; one on the ascending sparse tracks and one on the descending sparse tracks. In the real world, this would be problematic because the geometry of the tracks is irregular.

### Possible solutions

One possible solution would be to estimate a 2D PEF using multi-scale PEF estimation (Curry and Brown, 2001). Another possibility would be to estimate both a 2D PEF and the missing data at the same time in a similar approach to equation 3 from Claerbout (1999) but instead of estimating the PEF everywhere, throw out all fitting equations where the leading 1 of the PEF lands on unknown data. If you were to throw out all fitting equations where any coefficient of the PEF lands on unknown data, there would not be enough fitting equations to actually calculate the PEF. To estimate the PEF (**a**) on a model (**y**), we start with an initial model convolved with an initial filter (**Ay**) and perturb them with $\Delta\mathbf{y}$ and $\Delta\mathbf{a}$. In equation 4 we add the residual weight, **W**, to throw out fitting equations where the first coefficient of **a** lands on missing data. The free-mask matrix for missing data is denoted **J** and that for the PEF is **K**.

$$\mathbf{0} \quad \approx \quad \begin{bmatrix} \mathbf{AJ} & \mathbf{YK} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{y} \\ \Delta\mathbf{a} \end{bmatrix} + \mathbf{Ay} \tag{3}$$

$$\mathbf{0} \quad \approx \quad \mathbf{W} \left( \begin{bmatrix} \mathbf{AJ} & \mathbf{YK} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{y} \\ \Delta\mathbf{a} \end{bmatrix} + \mathbf{Ay} \right) \tag{4}$$

In Claerbout (1999), Figure 13 is generated by estimating both the missing data and unknown filter at the same time. I added a data-space weight as above and got the results in Figure 14. Notice it almost calculates the same filter. It does not completely fill in the missing data because we threw out many fitting equations.

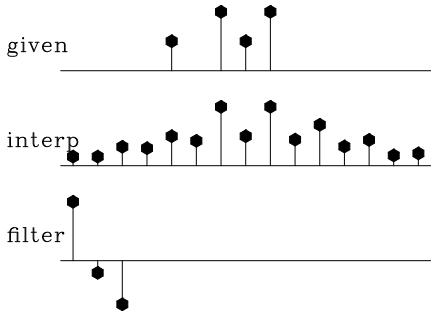



Figure 13: Top is known data. Middle includes the interpolated values. Bottom is the filter with the leftmost point constrained to be unity and other points chosen to minimize output power. jesse2-misif90 [ER]

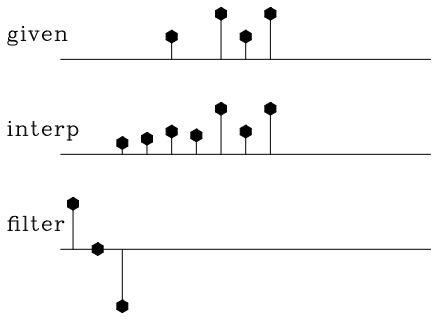



Figure 14: Top is known data. Middle includes the interpolated values. Bottom is the filter with the leftmost point constrained to be unity and other points chosen to minimize output power but only throwing out fitting equations where the leftmost point lands on unknown data. jesse2-misif90_2 [ER]

I have tested this only on simple 1D models, not yet on the Madagascar data. For the Madagascar data, the initial filter (**a**) may be the Laplacian.

## PRECONDITIONING

Here I applied preconditioning on the helix using the following fitting goals (Fomel, 1997):

$$substitute: \mathbf{m} = \mathbf{A}^{-1}\mathbf{p} \tag{5}$$

$$\mathbf{W}\frac{d}{dt}[\mathbf{L}\mathbf{A}^{-1}\mathbf{p} - \mathbf{d}] \approx \mathbf{0}$$

$$\epsilon\mathbf{p} \approx \mathbf{0} \tag{6}$$

In this case, a 2D PEF was estimated on the dense tracks and applied as regularization to the sparse tracks. This greatly improved the speed of convergence to create Figure 15.

## SUMMARY

As a small, non-stationary data set with irregular acquisition geometry, the Madagascar satellite data has allowed us to gain some valuable insight into inversion pitfalls and challenges. By looking at the data-space residual, an error was discovered. The impact of this error on the boot-strapping process of choosing the appropriate filter (in this case, $\frac{\mathbf{d}}{\mathbf{dt}}$) to remove the shift between tracks was minimal.

The use of this data is still not fully exploited. By understanding how to estimate filters on the sparse tracks and use them to fill in the sparse area, we will better understand how to tackle similar problems that arise regularly in seismic map generation.

## ACKNOWLEDGMENTS

I would like to thank Jon Claerbout for ideas and suggestions. I would also like to thank Bill Curry and Bob Clapp for help de-bugging code and useful discussions.

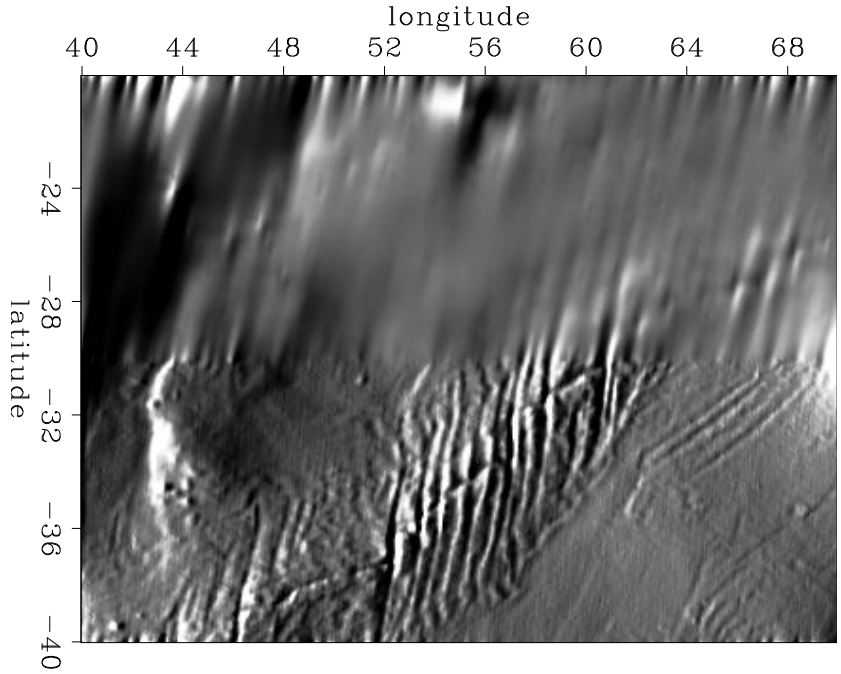



Figure 15: Result of fitting goal 6 applied to both sparse and dense data together usinga 2D PEF estimated on dense data. jesse2-madfig7 [ER]

## REFERENCES

Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project, `http://sepwww.stanford.edu/sep/prof/`.

Curry, W., and Brown, M., 2001, A new multiscale prediction-error filter for sparse data interpolation: SEP–**110**, 113–122.

Ecker, C., and Berlioux, A., 1995, Flying over the ocean southeast of Madagascar: SEP–**84**, 295–306.

Fomel, S., 1997, On model-space and data-space regularization: A tutorial: SEP–**94**, 141–164.

Lomask, J., 1998, Madagascar revisited: A missing data problem: SEP–**97**, 207–216.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 345–359*

# Implementing non-stationary filtering in time and in Fourier domain

*Gabriel Alvarez*[1]

**ABSTRACT**

Non-stationary filtering of seismic data can be accomplished in time or in Fourier domain by the theory of non-stationary convolution (Margrave, 1998). Here I show the results of implementing this theory for time-variant filtering of seismic data with an arbitrary number of filters and for forward and inverse NMO correction in the frequency domain. In the first case I show that the filters may be made to change sample-by-sample down the trace without artifacts being introduced and in the second case that the accuracy of the implied fractional sample interpolation can be controlled as an input parameter.

## INTRODUCTION

The frequency content of seismic data decreases with time as a result of absorption as the wavefield travels down the earth. Therefore, for seismic interpretation it is desirable to bandpass-filter the seismic traces with different filters at different times. Usual practice consists of applying a few filters in predefined time windows which are made to overlap to provide a smooth transition between them. There are at least two problems with this approach: on the one hand, the overlap zones are sort of arbitrary and phase distortion can be expected in them. On the other hand, with this approach we are limited to a few filters corresponding to the few chosen time windows.

A well-known alternative is the theory of non-stationary convolution and combination (Margrave, 1998; Rickett, 1999). This theory allows the design of arbitrary filters that can be made to change in a sample-by-sample manner. The design of the filters themselves is done in the frequency domain and its application to the data can be done in either the time domain, the frequency domain or a mixed time-frequency domain. In the time domain the process is similar to stationary filtering, with the columns of the convolutional matrix representing the delayed impulse responses of the filters applied to each sample, rather than the more familiar Toeplitz matrix of the stationary case. In the frequency domain, the convolutional matrix is nearly diagonal with the departure from diagonal being a direct indication of the degree of non-stationarity of the filters. In the mixed domain, the non-stationary filtering is performed

[1]**email:** gabriel@sep.stanford.edu

via a slow generalized Fourier transform.

Aside from the flexibility in choosing the domain of computation, we can also choose between non-stationary convolution and combination. The former is more appropriate when the spectra of the filters vary slowly and the latter when the change is sudden. Both non-stationary convolution and combination, however, reduce to stationary convolution in the limit of stationarity. This of course means that stationary filtering is a particular case of non-stationary filtering when the filters are kept constant for all samples.

In this paper I show the implementation of this algorithm for seismic trace filtering and for forward and inverse NMO correction. For the first application I used a set of randomly-generated seismic traces as well as a few traces of a an actual seismic line. It will be shown by a time-frequency analysis of the data before and after the filter that it is indeed possible to change the spectrum of the seismic trace in a sample-by-sample basis without noticeable frequency distortions. For the NMO application I used a few CMP gathers consisting of five hyperbolic reflections and background Gaussian noise. It will be shown that we can pose the NMO-correction problem as a time-variant filtering problem and that we can control the accuracy of the underlying fractional sample interpolation as an input parameter.

## THEORY OVERVIEW

### Time-variant Filtering

In a linear time-invariant (stationary) filter the output $g(t)$ is related to the input $h(t)$ by the convolution

$$g(t) = a(t) * h(t) = \int_{-\infty}^{\infty} a(t-\tau)h(\tau)d\tau$$

where $a(t)$ is the impulse response of the filter. In order to extend the applicability of this simple expression for the response of a non-stationary filter, we could replace $a(t-\tau)$ in the previous equation with the more general expression $a(t,\tau)$, indicating that the impulse response itself is now a function of the input time $\tau$. This expression, however, is too general and gives little insight into what the response of such a time-variant filter would be. Margrave, (Margrave, 1998) propose to maintain the convolutional nature of the impulse response by adding an explicit time dependence to it, that is, replacing $a(t-\tau)$ with either $a(t-\tau,\tau)$ (convolution) or $a(t-\tau,t)$ (combination). The formal definitions of non-stationary convolution and combination are, respectively (Margrave, 1998)

$$g(t) = \int_{-\infty}^{\infty} a(t-\tau,\tau)h(\tau)d\tau \tag{1}$$

$$\bar{g}(t) = \int_{-\infty}^{\infty} a(t-\tau,t)h(\tau)d\tau \tag{2}$$

These equations are clearly straightforward extensions of the stationary convolution concept. The introduction of the second index accounts for the non-stationarity. Comparing the two equations we see that the difference between non-stationary convolution and combination lies

in the way the impulse responses are considered in the convolutional matrix. In non-stationary convolution the filter impulse responses (as a function of input time $\tau$) correspond to the columns of the matrix, whereas in the non-stationary convolution case they correspond (as a function of the output time $t$) to the rows of the matrix (time reversed).

Just as with stationary filtering, it is convenient to find equivalent expressions in the frequency domain. These expressions are (Margrave, 1998)

$$G(f) = \int_{-\infty}^{\infty} H(F)A(f, f-F)dF \tag{3}$$

$$\bar{G}(f) = \int_{-\infty}^{\infty} H(F)A(F, f-F)dF \tag{4}$$

where $f$ and $F$ are the Fourier duals of $t$ and $\tau$ respectively and $H(F)$ and $G(f)$ are the Fourier transforms of $h(\tau)$ and $g(t)$. It is interesting to notice from these equations that non-stationary convolution in time domain translates into non-stationary combination in the frequency domain and vice versa. This is as opposed to the stationary case in which convolution in time domain corresponds to multiplication in the frequency domain.

Since we now have two time indexes ($t$ representing the filter samples and $\tau$ to keep track of the sample index to which each filter is applied), it is possible to have a third domain of computation, the so-called mixed domain in which the impulse response of the filters in the convolutional matrix are replaced with their corresponding frequency spectra. The equations to apply non-stationary filtering in the mixed domain are slow generalized Fourier transforms given by (Margrave, 1998)

$$G(f) = \int_{-\infty}^{\infty} \alpha(f,\tau)h(\tau)e^{-2\pi i f\tau}d\tau \tag{5}$$

$$\bar{g}(t) = \int_{-\infty}^{\infty} \alpha(F,t)H(F)e^{2\pi i Ft}dF \tag{6}$$

where $\alpha(p,v)$ is the so-called non-stationary transfer function which is the basic matrix in which the horizontal axis is time and the vertical axis is frequency (an example of this matrix will be given below). The transfer function is given by

$$\alpha(p,v) = \int_{-\infty}^{\infty} a(u,v)e^{-2\pi ipu}du \tag{7}$$

where $a(u,v)$ is the matrix of impulse responses, that is, the matrix with $\tau$ as its horizontal axis and $t$ as its vertical axis (again, an example will be shown below).

**Forward and Inverse NMO**

The NMO correction time for the small offset-spread approximation is given by the well-known hyperbolic equation (Yilmaz, 1987):

$$\Delta t_{NMO} = t_x - t_0 = \sqrt{t_0^2 + x^2/V_s^2} - t_0 \tag{8}$$

where $x$ is the trace offset, $t_x$ is the two-way travel time at offset $x$, $t_0$ is the two-way travel time at zero offset (normal incidence trace) and $V_s$ is the stacking velocity. Clearly, for a given trace different samples will have different NMO correction times even if the velocity is constant. Shallow events on the farthest trace with the slowest velocity have the maximum NMO-correction time whereas deep events on the near traces with the fastest velocity will have the minimum NMO-correction time. It is also important to note that in general some fractional sample interpolation will be required since we cannot expect the values of $\Delta t_{NMO}$ to be integer multiples of the sampling interval.

In order to apply the non-stationary filtering algorithm we need to recast the NMO equation as an all-pass non-stationary filter that will simply shift each sample by the given value of $\Delta t_{NMO}$. This can easily be achieved in the frequency domain by a linear phase shift with slope proportional to the value of $\Delta t_{NMO}$. In principle, any value of $\Delta t_{NMO}$ can be handled, so no fractional interpolation is required. For the sake of efficiency, however, it is convenient to precompute a given number of $\Delta t_{NMO}$ values. The accuracy of the implicit fractional interpolation is determined by the number of precomputed $\Delta t_{NMO}$ values and so can be controlled as an input parameter. Clearly, this parameter controls the trade-off between accuracy and speed of computation.

## DESCRIPTION OF THE ALGORITHM

### Time-variant Filtering

I will now summarize the steps necessary to perform non-stationary filtering in each of the three domains. In every case I will refer to non-stationary convolution but it is straightforward to change it to perform non-stationary combination instead.

### Time Domain

The algorithm in the time-domain is:

1. Design the filters in the frequency domain. These may be trapezoidal tapered filters or some other suitable bandpass filter. We can consider these filters as making up a matrix whose horizontal coordinate is time $\tau$ (that is, the sample time of application of every filter) and whose vertical component is frequency as shown in the left-hand side of Figure 1.

2. Get the filter impulse responses in time domain. This essentially means taking an inverse Fourier transform of each column of the left panel in Figure 1.

3. Form the non-stationary impulse response matrix in time domain. The right panel in Figure 1 shows an example of this matrix for the case of three different filters to be applied in three windows of data. The wavelet is zero phase and the impulse responses are shifted so that the wavelet is centered along the diagonal.

4. Apply the non-stationary convolution. This is done by matrix multiplication between the matrix in the right panel of Figure 1 and the seismic trace to be filtered.

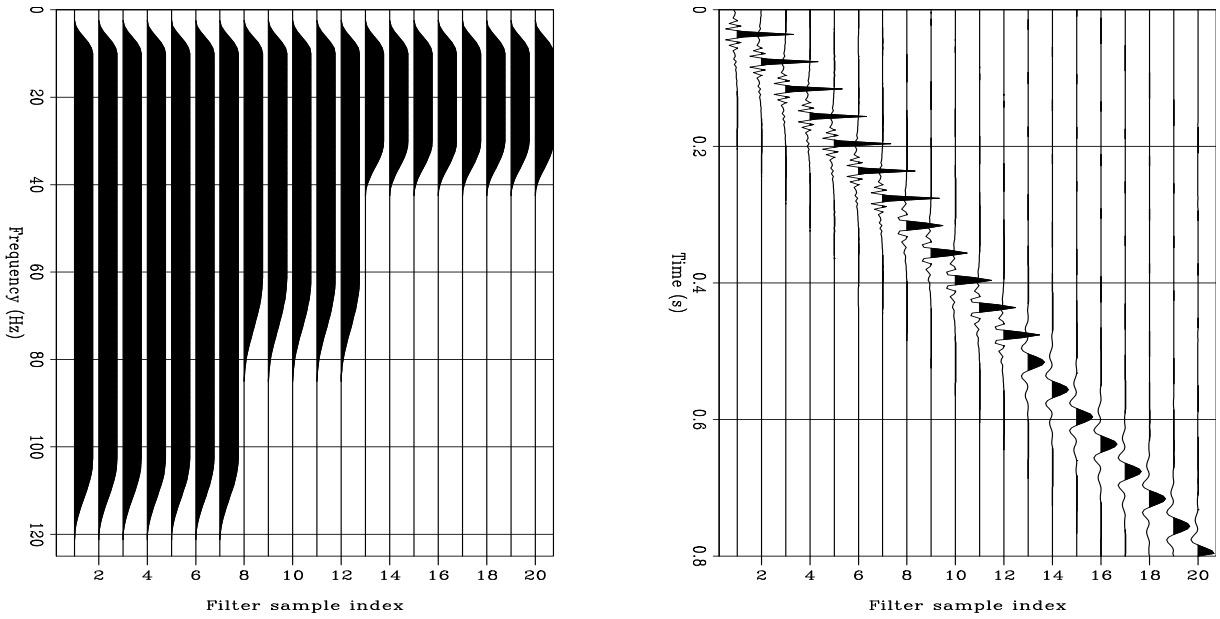



Figure 1: Filter design in the time-frequency domain. On the left, filter spectra as a function of time. On the right impulse responses on time gabriel1-tvf_td1 [ER]

**Frequency Domain**

In order to get an algorithm in the frequency domain, we basically have to take a Fourier transform in the horizontal direction of the data in the left panel of Figure 1. The algorithm is therefore:

1. Design the filters in the frequency domain, as before.
2. Take a Fourier transform in the horizontal direction (that is, a Fourier transform for each row of the matrix on the left panel of Figure 1) and form the corresponding frequency-domain convolutional matrix. Figure 2 shows the resultant matrix (amplitude spectrum only). This matrix is called the frequency connection matrix. On the left is a horizontally-shifted version of the matrix. The center "trace" corresponds to the stationary response and the "traces" away from it represent the departure from stationarity. Only a few "traces" are shown. On the right panel we have the complete dataset shifted so that the "stationary trace" is along the diagonal, which means that the off-diagonal energy represents again the departure from stationarity.
3. Take the Fourier transform of the input trace.
4. Multiply the frequency connection matrix (right panel of Figure 2) with the Fourier transform of the input trace to get the filtered trace in the frequency domain.

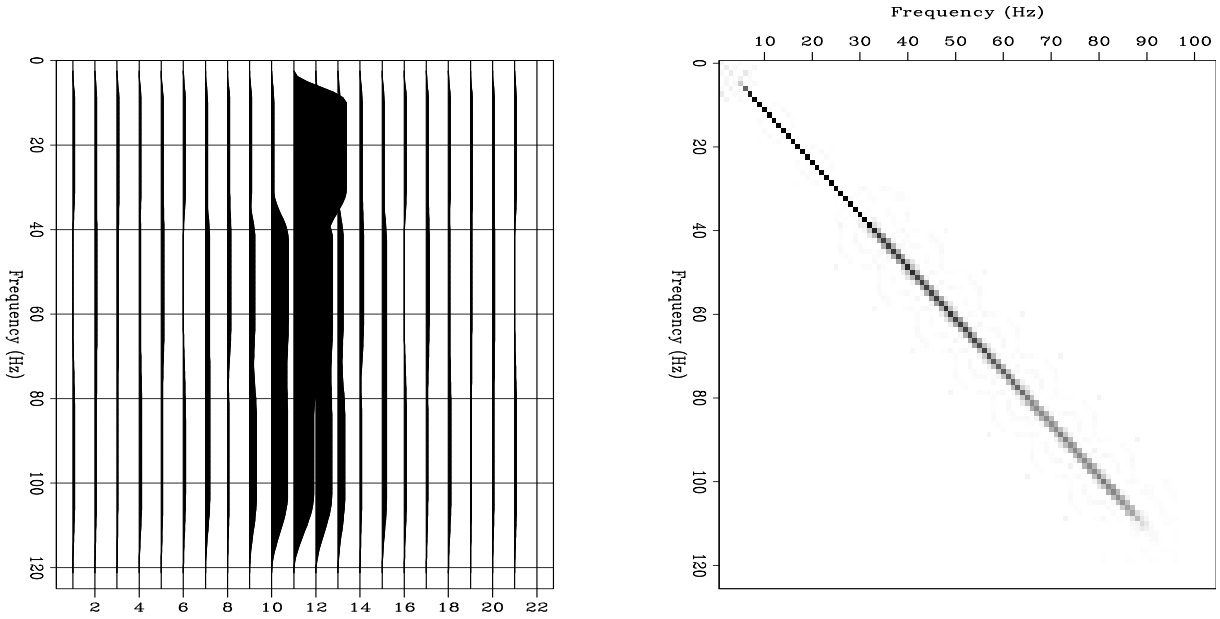



Figure 2: Frequency domain convolutional matrix. On the left, the filter spectra (amplitude only). The center "trace" represents the stationary response, the traces to the right positive frequencies and the traces to the left negative frequencies (only a few "traces" are shown). On the right the complete matrix shifted so that the stationary "trace" is on the diagonal of the matrix gabriel1-tvf_fd1 [ER]

5. Take the inverse Fourier transform of the filtered trace to get it in the time domain.

**Mixed Domain**

In a sense this is the most natural domain for time-variant filtering because we design and apply the filters in the same domain. Hence, the algorithm is simpler:

1. Design the filters in the frequency domain as before and form the time frequency matrix (the so-called non-stationary transfer matrix). This is just the matrix shown on the left panel of Figure 1.

2. Apply the filter via the (slow) generalized Fourier transform given by equation (6).

3. Inverse Fourier transform the filtered data to the time domain.

**Forward and Inverse NMO**

The algorithm for both forward and inverse NMO with the non-stationary all-pass filtering approach is the following:

1. Compute the maximum and minimum NMO correction required in the data according to the offsets, reflection times and velocities.
2. Use the previous information to precompute all the NMO time shifts in the frequency domain by application of an all-pass non-stationary linear phase filter.
3. Take the inverse Fourier transform to get the precomputed impulse responses in time domain.
4. For each trace, form the impulse response matrix selecting from the precomputed ones those required according to the NMO shifts appropriate to that trace.
5. Carry out the filtering by multiplying the trace by the the impulse response matrix.

## COMPUTER IMPLEMENTATION

Although the computer implementation of the above algorithms seems straightforward enough, I will mention some specific details here that, simple as they may be, are important when dealing with these algorithms

1. When using non-causal wavelets it is convenient to apply a time shift to the whole matrix so that the complete wavelet can be recovered. This can be easily done at the time of the computation of the filters in the frequency domain. This time shift needs to be compensated for after the filtering, of course.
2. In the time domain implementation only a few samples of the impulse responses are required to get a satisfactory result. This allows us to speed up the computation enormously because we need to multiply by a matrix that is non zero only near the diagonal as opposed to a dense matrix. We don't even need to store the complete impulse responses.
3. In the frequency domain similar, even more pronounced savings in computation, can be achieved by realizing that the frequency connection matrix is very nearly diagonal except for wildly varying filters. I found that good results could be obtained with only a few "traces" (perhaps 7 or 9) in the frequency domain.
4. After taking the horizontal Fourier transform in the frequency domain algorithm, it may be necessary to unscramble the traces to get both the positive and negative frequencies, otherwise only the amplitudes above or below the diagonal in Figure 2 will be present and the results will not be satisfactory.

## RESULTS AND DISCUSSION

I will now illustrate the results of doing time-variant filtering and NMO correction with the algorithms described above.

### Time-variant Filtering

For time-variant filtering I will show impulse responses as well as results with synthetic and real data.

### Impulse Responses

Figure 3 shows a synthetic dataset comprising six spikes and their impulse responses when the dataset is filtered with three different filters. Convolution of a seismic trace with this dataset will filter it in the time-variant manner implied by the right panel of Figure 3.

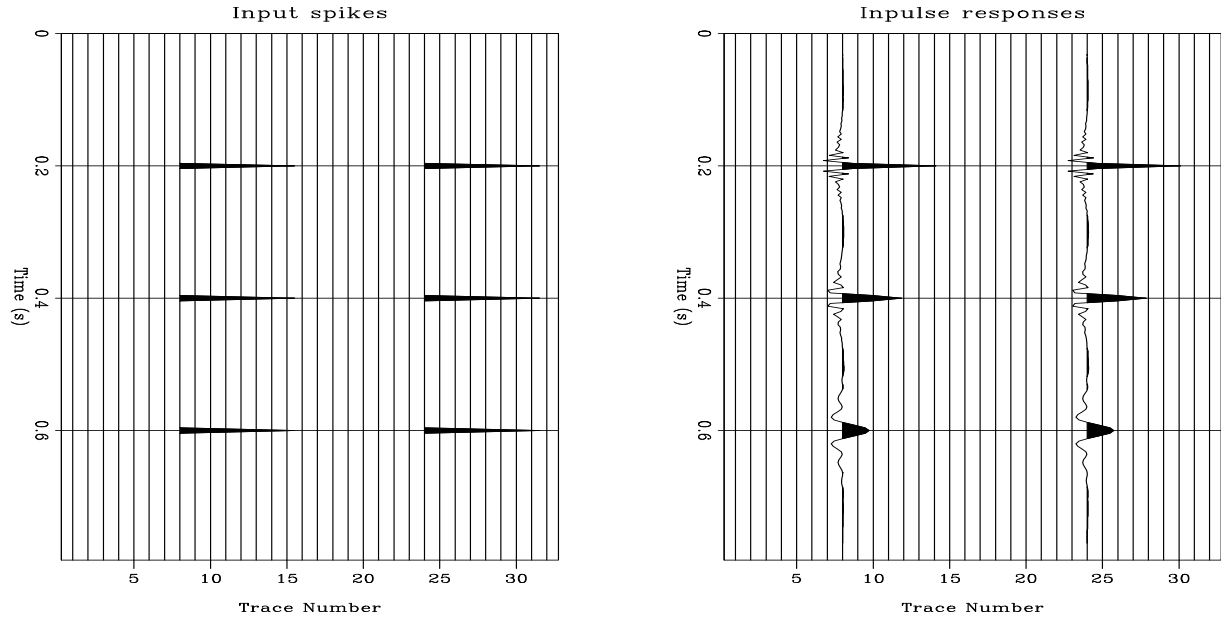



Figure 3: Impulse responses. On the left, the input spikes and on the right their corresponding impulse responses. gabriel1-tvf_ir1 [ER]

### Random Traces

The top left-side of Figure 4 shows a dataset composed of randomly generated seismic traces. The top right-hand-side shows the result of filtering this dataset with constant filters applied in three time windows corresponding to the top third, the middle third and the bottom third of the

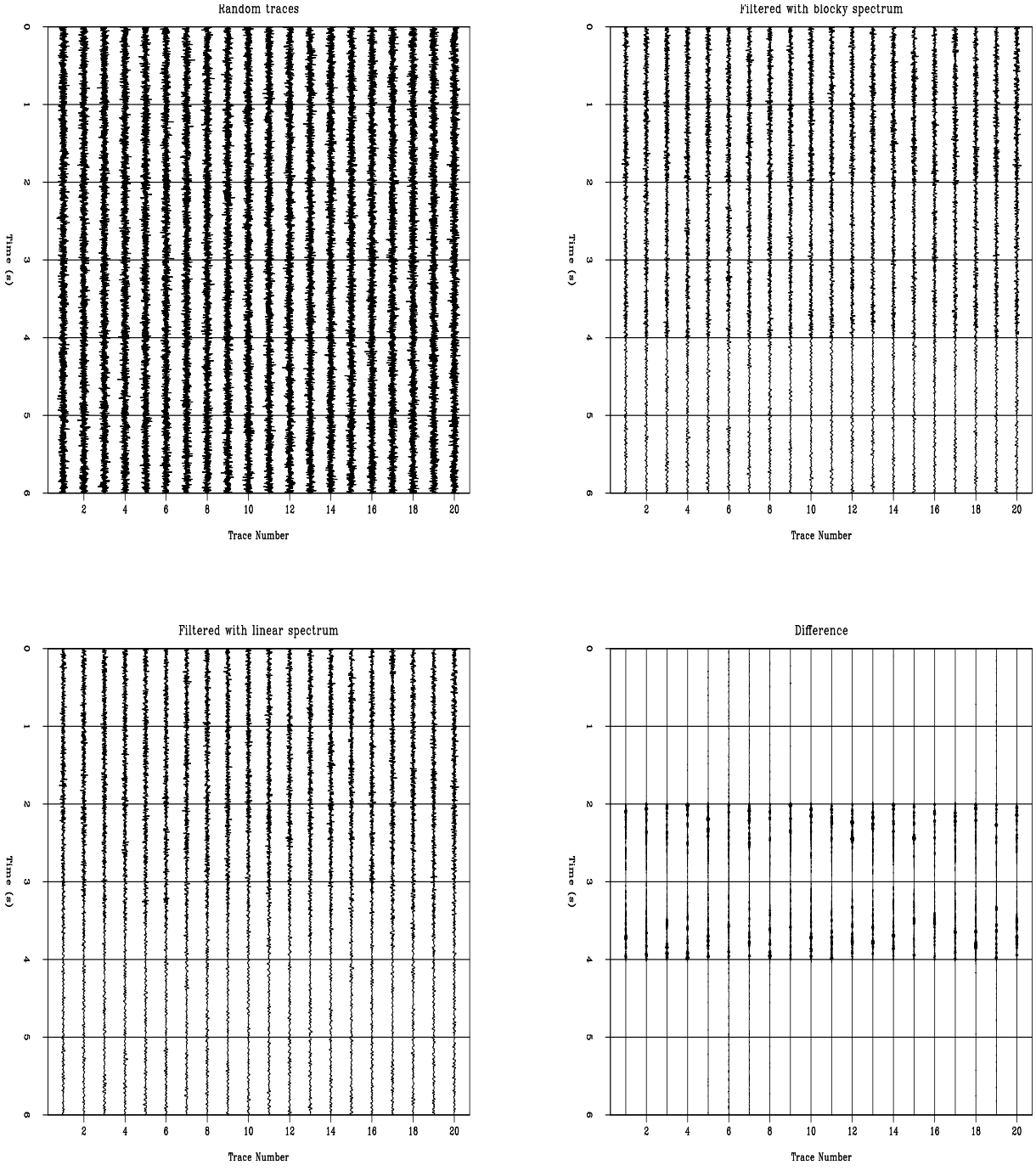



Figure 4: Random traces. On the top left, the input data. On the top right the filtered data with a "blocky spectrum," that is, constant spectrum in each of three time windows corresponding to one third of the trace length. On the bottom left data filtered with a linear spectrum in the middle third and constant spectrum in the top and bottom third. On the bottom right is the difference of the two filtered datasets. gabriel1-tvf_rt1 [ER]

traces. The applied filters were: in the top window (4-12-90-125) Hz, in the middle window (4-12-60-90) Hz and in the bottom window (4-12-30-50) Hz. The bottom left-hand-side shows the same dataset filtered a different way: the top third and the bottom third were filtered as before, but the middle third was filtered with a linearly changing spectrum that matched the filters above and below at the window limits. The bottom right-hand side shows the difference between the two filtered datasets. As expected, there is no difference in either the top or bottom third of the datasets. In the middle dataset, however, there is a difference which is greatest at the limits between the three zones where the difference in the filters is greatest.

Figure 5 shows a time-frequency analysis of the original data on the left and the filtered data with the linear spectra on the right. Since the data are random traces, they have all the frequencies from 0 to Nyquist (250 Hz in this case). After the filtering the spectrum is shaped by the applied filters. The bold line represents approximately the high pass frequency of the filters applied at each sample. There is no evidence of distortion in the spectrum even though the spectrum was made to change sample-by-sample in the middle third of the dataset.

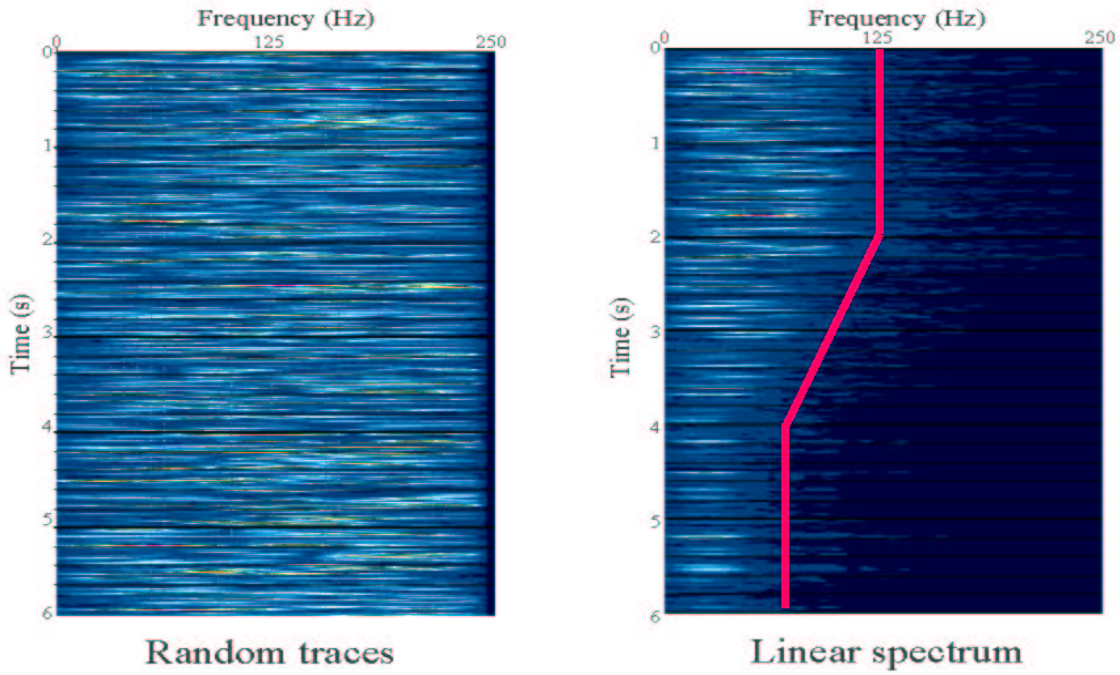



Figure 5: Time-frequency analysis for random data. On the left, the time-frequency display of the input data and on the right the result of the filter with the linear spectrum. The white areas represent large amplitudes. The thick solid line represents approximately the high cut frequency of the filters in every window gabriel1-tvf_tfa1 [NR]

**Real Data**

The top left-hand side of Figure 6 shows a short stack of a real 2-D seismic line. The top right-hand side shows the result of filtering the data with a “blocky” spectrum in which again

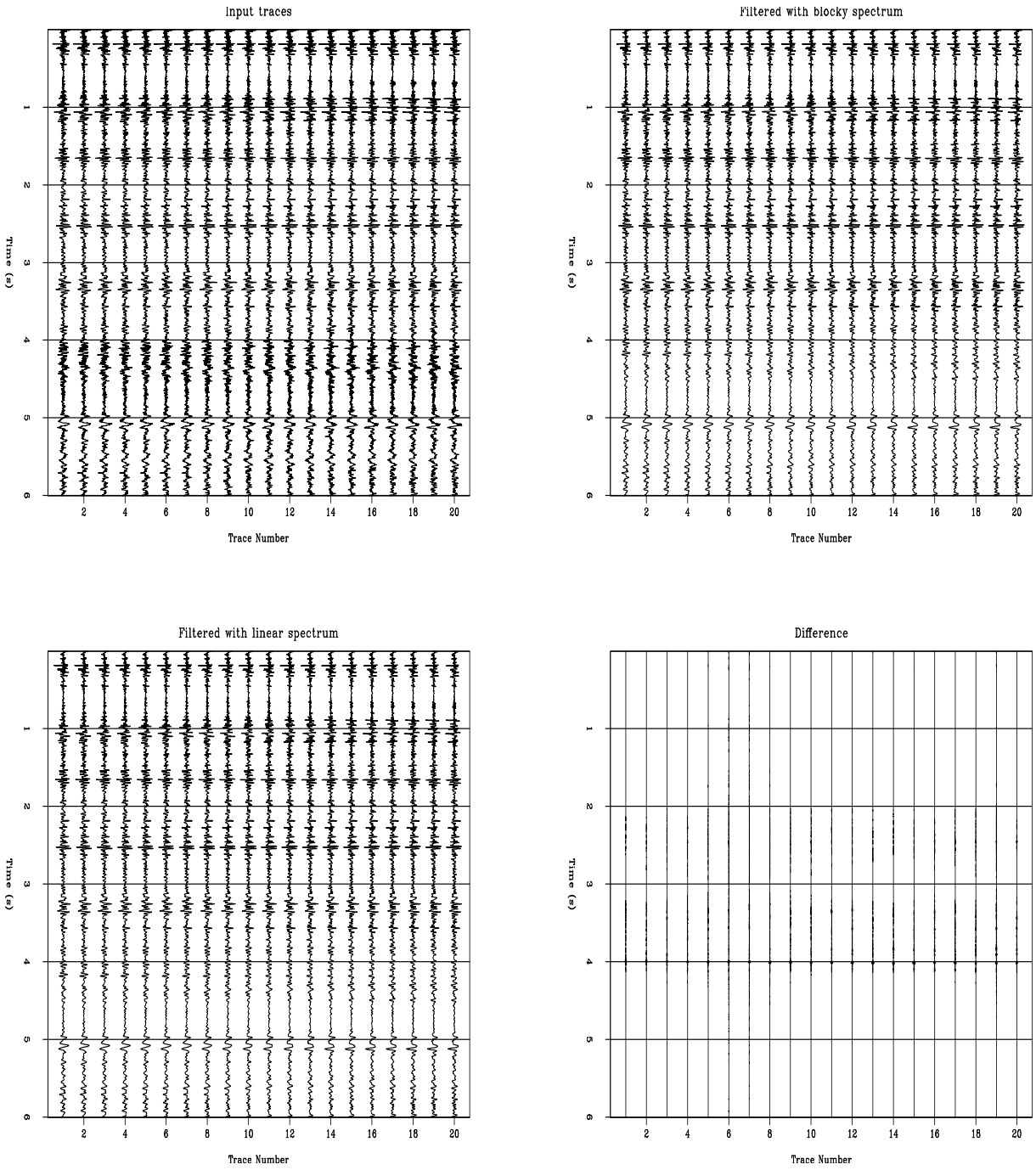



Figure 6: Real traces. On the top left, the input data. On the top right the filtered data with a "blocky spectrum," that is, constant spectrum in each third of the data. On the bottom left data filtered with a linear spectrum in the middle third and constant spectrum in the top and bottom third. On the bottom right is the difference of the two filtered datasets. gabriel1-tvf_rd1 [ER]

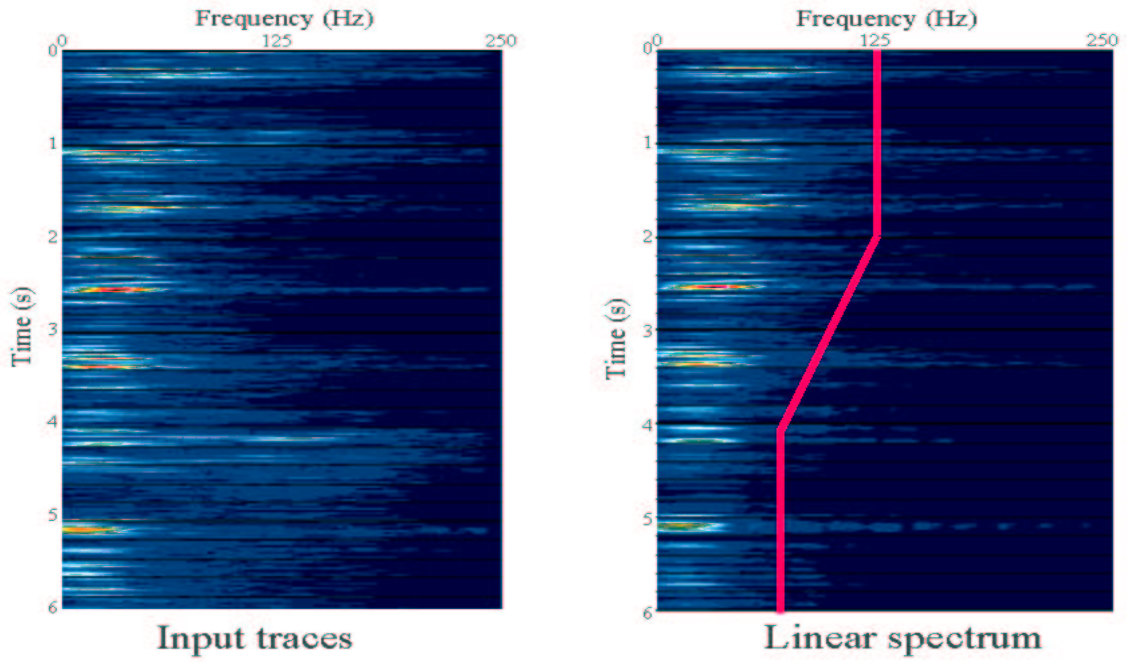



Figure 7: Time-frequency analysis for real data. On the left, the time-frequency display of the input data and on the right the result of the filter with the linear spectrum. The white areas represent large amplitudes. The thick solid line represents approximately the high cut frequency of the filters in every window gabriel1-tvf_tfa2 [NR]

the top third of the data are filtered with one filter, the middle third with a narrower filter and the bottom third with an even narrower filter. The bottom left panel corresponds to the result of filtering the dataset with a linearly changing spectra in the middle third of the trace. The bottom right shows the difference between the filtered datasets. This time the difference is small even in the middle third of the trace because the original data spectrum is not very broad as shown in Figure 7. In this figure the left-hand side corresponds to the time-frequency spectrum of the data and the right-hand-side corresponds to the equivalent plot for the dataset filtered with the linearly-changing spectrum. As noted before, there is no apparent frequency distortion arising from the sample-to-sample change in the trace spectrum.

**NMO correction**

The top left-hand side of Figure 8 shows a modeled CMP gather consisting of five reflections in a Gaussian noise background. The top right hand-side shows the result of applying NMO correction using the time-variant filtering algorithm described above. The fractional sample interpolation parameter was set to two, meaning that a nearest neighbor interpolation was implicitly performed in the frequency domain to half the sampling interval. Changes in the fractional interpolation parameter would allow for better interpolation at the expense of increased run-time. The result shows a very good correction with very little distortion even for

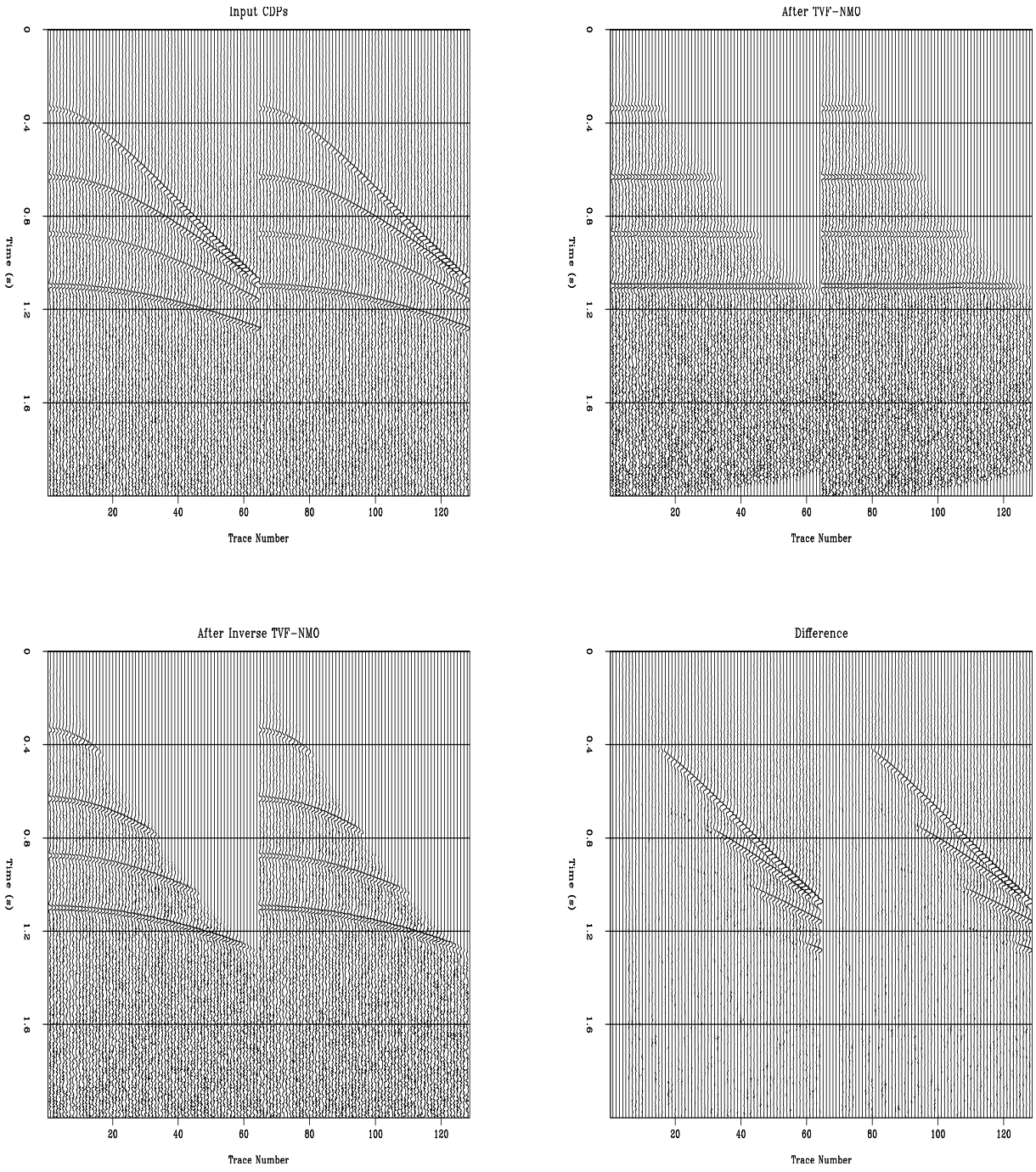



Figure 8: Frequency domain NMO correction. On the top left, the input synthetic data. On the top right the NMO-corrected data. On the bottom left inverse NMO-corrected data and on the bottom right the difference of the inverse NMO-corrected and the input data. Clearly the difference is due to NMO stretch gabriel1-nmo1 [ER]

this small interpolation value.

The same algorithm used for forward NMO correction was also applied for inverse NMO correction. The bottom left panel of Figure 8 shows the inverse-NMO corrected CMP gather. Finally, the bottom right panel of Figure 8 shows the difference between the original data and the inverse-NMO result. Except for the obvious effect of the NMO-stretch, the inverse NMO-corrected CMP is nearly identical to the original CMP.

## SUMMARY AND CONCLUSIONS

The theory of nonstationary filtering allows great flexibility in the design and application of time-variant filters. We can choose between three different domains of application: time, frequency and mixed and we can also choose between convolution and combination. From the results shown and from many other tests that I performed I drew the following conclusions:

1. Time domain should be preferred when filtering with "blocky" spectra, that is, when the filters are kept constant in each window and change abruptly from window to window.
2. Frequency domain should be preferred when using filters that change from sample-to-sample.
3. The mixed domain should only be used when filtering a few traces because of its large run-time.
4. As a general rule, non-stationary combination is preferable to non-stationary convolution when the filter spectrum changes abruptly from one sample to the next. For slowly changing spectrum, non-stationary convolution is probably better.

## ACKNOWLEDGMENTS

I would like to thank Conoco for allowing me to spend the Summer of 2001 at its Seismic Imaging Technology Center (SITC) in Ponca City where this work was carried out. In particular, I would like to thank Dr. Chuck I. Burch, my mentor at the time.

## REFERENCES

Margrave, G., 1998, Theory of nonstationary linear filtering in the Fourier domain with application to time-variant filtering: Geophysics, **64**, 244–221.

Rickett, J., 1999, On non-stationary convolution and inverse convolution: SEP–**102**, 129–136.

Yilmaz, O., 1987, Seismic data processing: Society of Exploration Geophysicists.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 359–367*

# A return to passive seismic imaging

*Brad Artman*[1]

**ABSTRACT**

Rekindling the passive seismic imaging effort at Stanford, I have acquired grants with Simon Klemperer of the Stanford Crustal Research Group from both the Petroleum Research Fund and the National Science Foundation to pursue two- and three-dimensional imaging efforts of the subsurface in a passive listening methodology. Utilizing the outstanding SEP hardware and software infrastructure and expertise, I have begun to build the resources necessary to manipulate the massive datasets toward producing an image. Efforts to acquire several existing datasets that seem to fit the requirements of this method are presently underway, while 180 Gbytes of the Santa Clara Valley Seismic Experiment from 1998 arrived in house on the first of March.

## INTRODUCTION

The autocorrelation conjecture of Claerbout (1979) is a long standing idea that motivating passive seismic imaging that has yet to either bear fruit or be put to bed. Claerbout et al. (1988) begins a suite of articles in these consortium reports by defining the scope of the 1989 SEP passive seismic data acquisition effort that is then subjected to increasing interrogation by many former students. The results of this effort and it's many conterminous developments are sprinkled within the SEP reports until the thesis of Cole (1995) summed up much of the effort. Articles by James Rickett and co-authors beginning with Rickett and Claerbout (1996) then resurface the topic upon receipt of a solar seismic data set[2], that not only met the requirements of the passive experiment, but returned excellent results.

Taking a lead from takes a lead from Claerbout (1968) where it is shown that the autocorrelation function of the transmission series of the earth is directly related to a reflection seismogram, Schuster et al. (1997) develops the mathematics behind prestack migration of autocorrelograms. Schuster and Rickett (2000) then go on the generalize that work outlining the imaging conditions for both reflectivity and source location of correlation datasets. Anstey (1964) gives a very thorough treatment of the fundamentals of correlation techniques in general.

Pre-dating all of the above is a patent application submitted by Weller (1969) describing this exact process of collection and correlation of passive seismic receiver stations. Weller

[1]**email:** brad@sep.stanford.edu

[2]The Michelson Doppler Imager, http://soi.stanford.edu/results/heliowhat.html

(1969) describes a 2-dimensional acquisition in the Gulf-coast region as returning convincing sequence boundary reflections from recordings on the order of eight hours long.

Recognizing that this imaging methodology will involve very long time records, the work of Kostov (1990) is applicable in handling spurious and coherent noise patterns endemic to recording stations with long residence periods. This thesis treats the closely related topic of seismic imaging utilizing a turning drill-bit as a source (also treated in Cole (1995)). Specific treatment of the handling and shaping of data in a pre-processing step as well as the multi-channel and areal nature of his experiments are beneficial to the development of this effort. This same topic has recently been advanced through the work of Yu and Schuster (2001) that explains migration of crosscorrelogram data collected during seismic profiling while drilling experiments as an application of the previously mentioned (Schuster and Rickett, 2000) derivations.

While the ultimate goal is to acquire another passive seismic data set, this paper will begin to describe investigation of this subject through the manipulation of existing data sets from the seismology community. Utilizing the capabilities of SEP3D to handle irregular datasets, I will attempt to apply the imaging methodology as described below on the 1998 Santa Clara Valley Seismic Experiment (additionally, a 1997 USGS experiment at the Kilauea volcano, Hawaii (Almendros et al., 2001) exists that features several areal arrays of seismometers).

## THEORY REVIEW

With one schematic, the fundamentals of the correlation conjecture concept are easily digested. Figure 1 shows the incident plane wave acting as a virtual or 'ghost' source as it is reflected from the free surface. One final arm-wave will cover the rest of the plane waves needed for a successful passive seismic experiment and not pictured in the figure. Conventional seismology assumes an impulsive source and a randomly distributed subsurface that we attempt to deconvolve. The premise of this experiment lies in switching those two roles. If we have a structured earth and a random distribution of sources buried within it we can deconvolve, in much the same manner, our recorded signals to return the impulse response of the earth. Therefor, the perfect experiment would have sufficient noise activity to illuminate the free surface from every incidence angle, and around all azimuths.

Cross correlating each trace with every other trace manufactures a 5D data volume of pseudo-shots very analogous to conventional data, although here, the structure of the earth will be in the form of its autocorrelation. Visualization of the 3D volume constructed from one line is much easier on the brain and one notices that the main diagonal of the cube is the autocorrelations (analogous to zero offset and the CMP location) and successive lessor diagonals are correlations of traces with increasing offset. At the same time, the correlation process will illuminate our need to know our source timing and shape.

Given any one receiver line, only ray-paths contained in the plane defined by the receiver line and the $90^o$ azimuth to it will have zero move-out across the line. These events will contribute a direct event of infinite velocity and no reflected energy as azimuth will be maintained

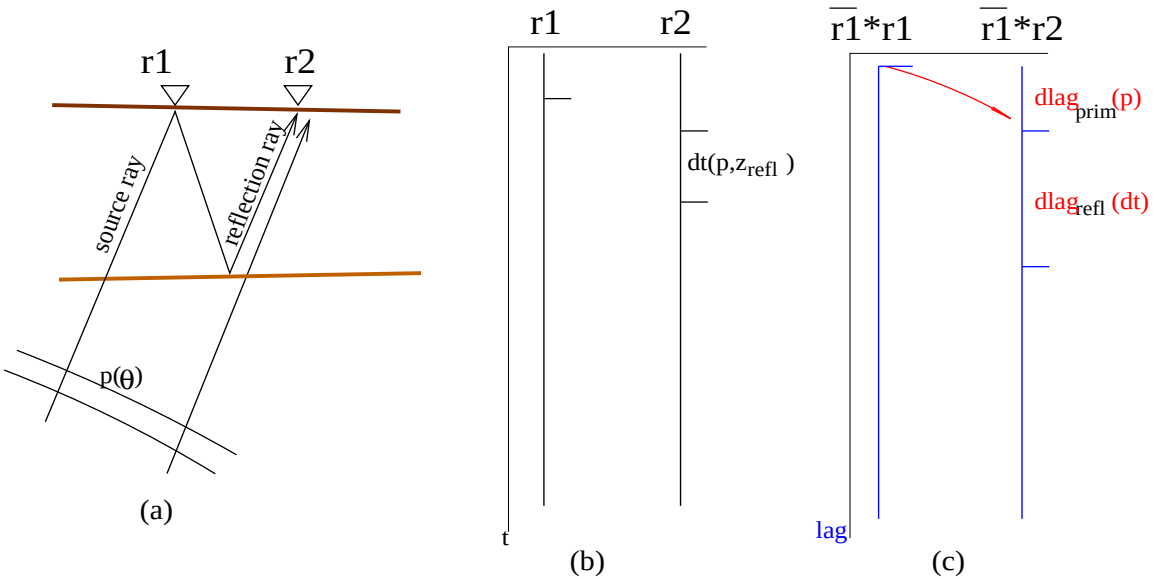



Figure 1: (a)Given a simple earth model, an upcoming plane wave will be recorded with measurable move-out unless incidence angle is zero or its azimuth is perpendicular to the receiver line. (b) Each plane wave reflecting from the free surface provides only one reflection path to each receiver. (c) Correlating all traces with each other yields a second data space with components indicative of subsurface structure and the incident energy. brad1-noise [NR]

along its travel path. This thought experiment highlights the problem that only direct ray paths traveling in the same azimuth as the receiver line will catch a ghost reflection. This topic is addressed more fully in the next paper in this volume (Artman, 2002b).

To investigate the efficacy of these assertions, I will use the existing data from a Santa Clara, California seismology campaign that benefits from the recent shooting of a 2D profile across the same area.

## SANTA CLARA VALLEY SEISMIC EXPERIMENT

In June 1998, 40 seismometers were deployed in the Santa Clara Valley on a roughly $3x5$ km grid (see Figure 2 for locations) by the USGS and the University of California, Berkeley using PASSCAL[3] equipment. The goal of the experiment was to: (1) constrain the basin structure of the valley from P and S travel times, (2) investigate site responses for earthquake hazards, and (3) better locate small quakes that occur on the basin margin. The results of this study should include tomographic velocity models to provide a beginning earth model. The results are as yet unpublished however. Also advantageous with this data, is the co-existence of an active seismic profile within the former array. From conversations with researchers at the USGS, the line will trend north-east from the south-west corner of the array and be completely contained

[3]Program for the Array Seismic Studies of the Continental Lithosphere, http://www.passcal.nmt.edu/iris/passcal/passcal.htm

by perimeter of the passive array. The USGS acquired the data late last year, and results and accurate positioning should be available soon.

Continuous logging of data every 0.02 seconds for six months yields an outrageous 180 Gbytes of passive seismic data. Each of the stations log vertical and northerly and easterly oriented shear motions. All time signals are synchronized with GPS clocks.

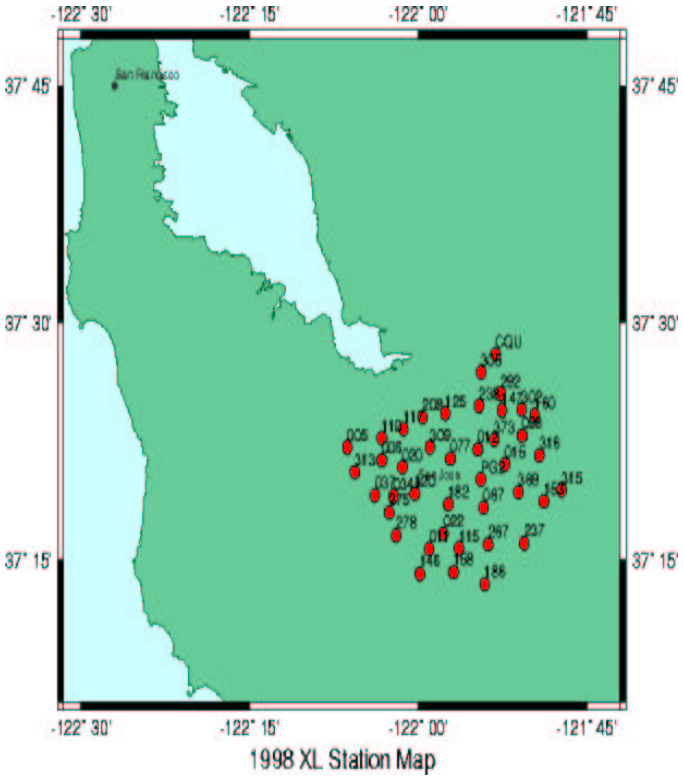



Figure 2: Station location of the 40 short-period seismometers deployed in the Santa Clara Valley of California. Six months of continuous recording every 0.02 seconds results in 180Gbytes of data currently in house. brad1-map [NR]

The Mark Products (now owned by Sercel) L22 short period seismometer is ubiquitous in the seismological community. Therefor, it is reasonable to understand its characteristics. W. Menke (1991) performs a comprehensive analysis of the performance of the L22. It has a resonance frequency of 2 Hz, which is significantly lower than that of an exploration geophone. However, as seismologists are normally not interested in higher frequencies, the response functions shown never extend above about 30 Hz. The authors claim to have seen significant cross-axis coupling of the shear and compressional channels over frequency bands near the natural frequency, but this seems to be of little concern to this use of the data. Of more concern, the authors identify “one of the main instrument defects” of the L22 being strong amplitude resonance peaks centered at 28 Hz in over 20% the instruments.

Preliminary manipulation of the SCVSE data show that traces are indeed white show no coherence in their raw form. Cross correlating the records at this time provides no useful information as I have not been able to implement the code with the irregular geometries required

with the data shown in Figure 2.

At this stage, it is unclear whether strong earthquake energies will help or hinder the experiment. While we desire strong incident wave fields, over representation of energy from particular azimuths and incidence angles may be detrimental. The underlying question here is whether or not teleseismic events (earthquake signature from long distances) will be the predominant energy source to illuminate the subsurface by reflecting from the free surface. If so, focusing our efforts in time around the arrival times of known events (from published earthquake catalogs) may significantly reduce the length of the time series that need be processed. Rather than long, continuous time records, we can isolate discrete time windows that can be treated analogously to single shot experiments. The price to pay for this however will be in resolution. Due to the geometry of the radial structure of the earth, we can only expect incident waves in a limited window of incidence angles from below. In addition, the usable period of these events is centered around one or two seconds which will greatly reduce the resolution of the image.

Figure 3 shows the earthquake and blasting events within 500 km of the survey location during the time the recording units were deployed. As these events and their times are readily available, it will be easy to window data series within and between major quake events to address this question.

Alternatively, it may be possible to use the earthquake energy in both contexts within the framework of an illumination study. Because the timing, azimuth and ray parameter of the earth-quake energy is available, it may be possible to normalize or otherwise manage what could be over-abundant energy.

Of benefit to this type of survey is that those who design and utilize these surveys are principally interested manufacturing tomographically derived velocity models from earthquake events utilizing both vertical and shear components of ground motion. This fact results in the availability of initial velocity models for migration studies and rudimentary practices in separating incident and scattered wave fields. However, it is my sincere hope that ambient noise will provide sufficient images as to not need to focus on teleseismic events. Due to the large offset between stations (three to five kilometers in this instance) the transmission losses of some of the ambient noise will undoubtedly prevent correlated signal from spanning the entire breadth of the survey layout.

Despite the outcome of this question however for this particular training set, the issue needs addressed specifically with an experimental mobilization tailored for our interests. This means that it should have receiver arrays designed to attenuate surface waves, station spacing on the order of a few meters (rather than the kilometers associated with seismologic data sets), and a roughly square map view (as suggested by Artman (2002b)) with a regular station spacing. The harder the near surface that the receivers are coupled to, the less high frequency surface noise will conflict with our desired signal.

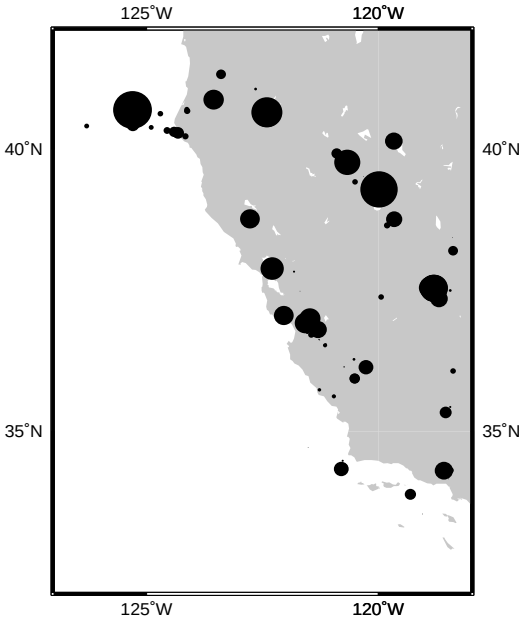



Figure 3: 210 Earthquakes and blasting events most likely to be recorded by the deployed instrumentation. Plotted are events between magnitude 3 and 5 within 500 km from the survey. These events will likely correspond to incidence angles of approximately $35^o$. Further afield teleseismic events will contribute more vertical events, but suffer from decreased frequency range. brad1-quakes [NR]

## PLAN FORWARD

The correlation of the geophone time series outputs a volume that is roughly equivalent to the raw data of conventional seismic acquisition. The parallel code to manufacture this output volume is almost running. The road ahead is a long one that will involve velocity estimation, converted wave problems, noise suppression and separation, simultaneous inversion of multiple experiments, and migration- in short: the rest of the field of modern seismic data processing.

Before all of these issues are addressed however, one of the principal problems to be addressed is that of the separation of the direct incident Wave field from that reflected from the free surface. This will be my principal goal in the short term. Some preliminary ideas for this are outlined in Artman (2002a).

Assuming positive results, the next step will be to identify a location and acquisition strategy for SEP to collect its second passive seismic dataset as suggested in Claerbout et al. (1988).

## REFERENCES

Almendros, J., Chouet, B., and Dawson, P., 2001, Spatial extent of ta hydothermal system at kilauea volcano, hawaii, determined from array analyses of shallow long-period seismicity: Journal of Geophysical Research, **106**, no. B7, 13,581–13,597.

Anstey, N. A., 1964, Correlation techniques a review: Geophys. Prosp., **12**, no. 04, 355–382.

Artman, B., 2002a, Coherent noise in the passive imaging experiment: SEP–**111**, 375–378.

Artman, B., 2002b, Is 2d possible?: SEP–**111**, 367–374.

Claerbout, J. F., Cole, S., Nichols, D., and Zhang, L., 1988, Why a big 2-D array to record microseisms?: SEP–**59**, 1–10.

Claerbout, J. F., 1968, Synthesis of a layered medium from its acoustic transmission response: Geophysics, **33**, no. 02, 264–269.

Claerbout, J. F., 1979, Errata: Fundamentals of geophysical data processing: SEP–**20**, 257–258.

Cole, S. P., 1995, Passive seismic and drill-bit experiments using 2-D arrays: SEP–**86**.

Kostov, C., 1990, Seismic signals from a drillbit source: SEP–**63**.

Rickett, J., and Claerbout, J., 1996, Passive seismic imaging applied to synthetic data: SEP–**92**, 83–90.

Schuster, G., and Rickett, J., 2000, Daylight imaging in V(x,y,z) media: SEP–**105**, 209–226.

Schuster, G. T., Liu, Z., and Followill, F., 1997, Migration of autocorrelograms: 67th Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1893–1896.

W. MenkeL. Shengold, G. H. H. G. A. L., 1991, Performance of the short-period geophones of the iris/passcal array: Bulletin of the Seismological Society of America.

Weller, C. E., 1969, U.s. patent application: Appl.No. 877,350.

Yu, J., and Schuster, G., 2001, Crosscorrelogram migration of IVSPWD data: 71st Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 456–459.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 367–375*

# Is 2D possible?

*Brad Artman*[1]

**ABSTRACT**

One of the major concerns in the passive seismic imaging experiment as outlined in by Claerbout et al. (1988) is touched on by the title: *Why a big 2D array to record microseisms?* My question is: Will a linear array produce sufficient results? I answer this question in the affirmative with anecdotal and modeling evidence as well as note a recent success by seismologists from the University of British Columbia that suggests that the linear acquisition strategy can provide useful information about the subsurface.

## INTRODUCTION

In the passive seismic imaging experiment as outlined in this volume (Artman, 2002), full coverage of plane waves arriving from all azimuths and with all incidence angles is necessary to fully illuminate the subsurface. Either case of absent ray parameter constituents or over-representation of a few can confound the correlation approach to passive imaging. In this event, beam-steering analysis such as that undertaken in Cole (1995) may yield useful information, but will not provide images of the subsurface. Also problematic could be energy arriving from perfectly perpendicular azimuths to the linear receiver spread. This energy would have infinite ray parameter values, and be indistinguishable from vertical waves arriving from the deep earth. To explore this issue, I will offer three arguments in favor of the feasibility of the linear strategy.

One of the convenient issues about exploring this question, is that we are able to find the answer along the way while making plans to accommodate larger dimensionality. We can decimate our data set to address this specific concern. This is precisely the strategy employed in the first case where the MDI[2] solar seismic data set is defined as a working model that we can decimate for the purposes of our concern. The second topic for discussion concerns the review of a patent application describing this acquisition strategy that claims informative results. Finally, seismologists at UBC have positive results inverting data collected with a linear array strategy across central Oregon. Through these three lines of evidence, I conclude that interpretable results are procurable from linear acquisition strategies.

[1]**email:** brad@sep.stanford.edu

[2]The Michelson Doppler Imager, http://soi.stanford.edu/results/heliowhat.html

## SOLAR DECIMATION

The solar data set can act as an excellent modeling exercise for testing our ideas about the passive imaging concept. The impulse response of the sun, while containing no reflective structure, is simple, not laterally variable, and the ambient noise is ubiquitous at all locations and azimuths. These characteristics make it a great laboratory if not a perfect model. The question it won't ever address however is: how long will our records need to be on earth?

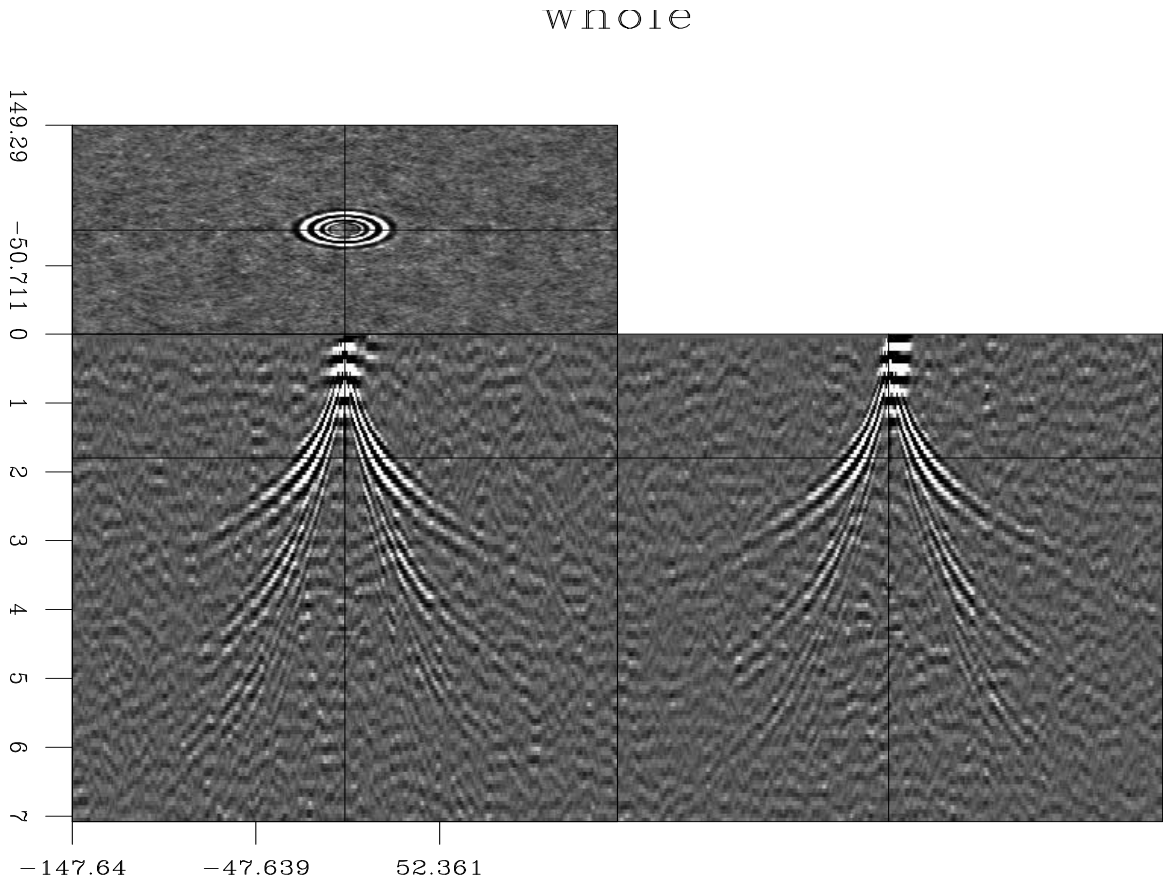



Figure 1: The entire data volume processed with the correlation methodology and stacked across offsets. From Rickett and Claerbout (1999) brad2-whole [ER]

While the dimensions of the full volume of the solar seismic data set are 256 x 256 x 256, it is easy to window smaller cubes out of the volume or simply take any one line of receivers. Figure 1, shown for reference, shows the impulse response of the sun as arrived at through the correlation processing of the entire solar data. Events are plunging waves, overturning and returning to the surface, of multiple order bounces within the survey area. Figure 2 is the result of a single line of receivers correlated with each other. While the quality of the image is definitely degraded, the plunging waves are readily interpretable (even the second order events if you really believe).

As an exercise for a well rounded argument, I also began decimating the data by various factors and comparing the result to images that were decimated by the same factor after being

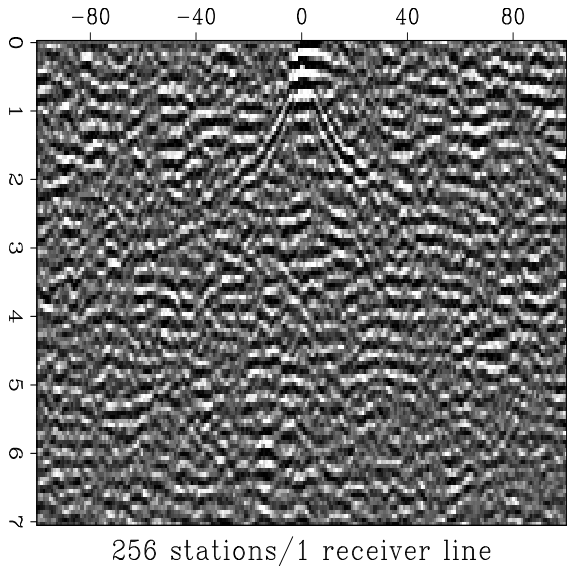



Figure 2: By using only one receiver line of the 256 available, the image quality is poor, but interpretable. Is this validation of the 2D experiment? brad2-2d [ER]

constructed with the entire data set. Since Figure 2 is very definitely of worse quality than one of the panels comprising Figure 1, I expected to see a steady degradation of the image quality from decimating the data space. Interestingly however, it seems that the 256 receiver square array seems to be quite oversampled. The third order body waves apparent in the Figure 1 are just disappearing in Figure 3 which is sub-sampled by 32x in the cross-line axis (only eight receiver lines across the entire cross-line direction!).

This result begs the question then: is the degradation in image quality a function of cross-line offset or number of receiver lines available to contribute information? As Figure 4 shows, the image quality is definitely degraded by using eight neighboring receiver lines as compared eight lines spanning the entire data volume. I believe this observation is due to the increase in the aperture of available information to sum into the result.

Now let's go back to considering the end case of only one line. Can adding one more receiver line significantly improve our image quality if chosen correctly? Figure 5 compares images created by processing two consecutive lines (on the left) with one created by line numbers 1 and 256 (on the right). Whereas the left panel is no more clear than Figure 2, the improvement in the right panel is dramatic.

As mentioned in the previous article (Artman, 2002), only direct ray paths traveling in the same azimuth as the receiver line will record a ghost reflection. When considering the reflection from the free-surface, one can imagine that including larger offset in the perpendicular direction from a single line will allow ever more energy from the lost azimuths to be incorporated into the image. This will only work well in the case of a body with stationary characteristics for the extreme offsets described above. If heterogeneities exits within the offset range, constructive summation will not occur. Therefore, this may be of limited utility on earth, but might be advantageous perpendicular to strike.

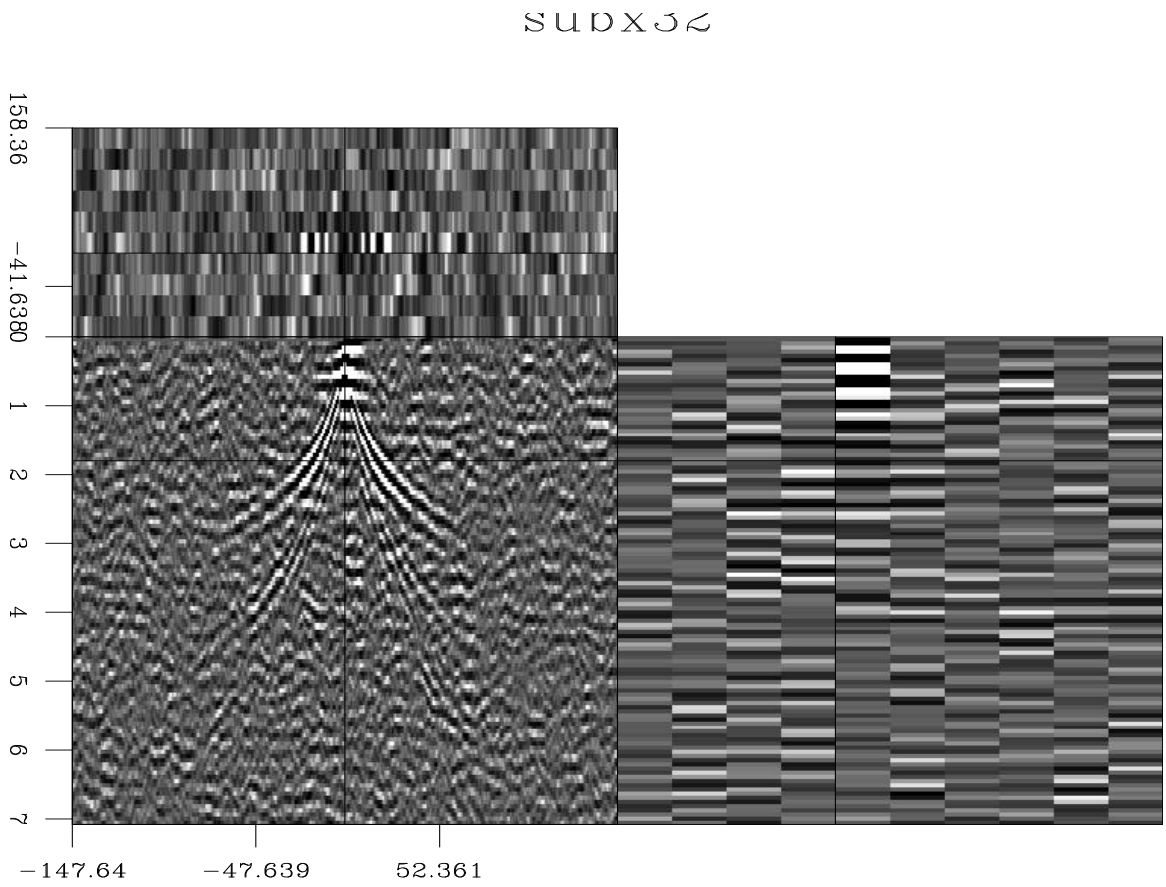



Figure 3: Taking only every 32nd receiver line of the available 256, the image manufactured from only 8 lines is not only interpretable, but arguable quite clear. brad2-subx32 [ER]

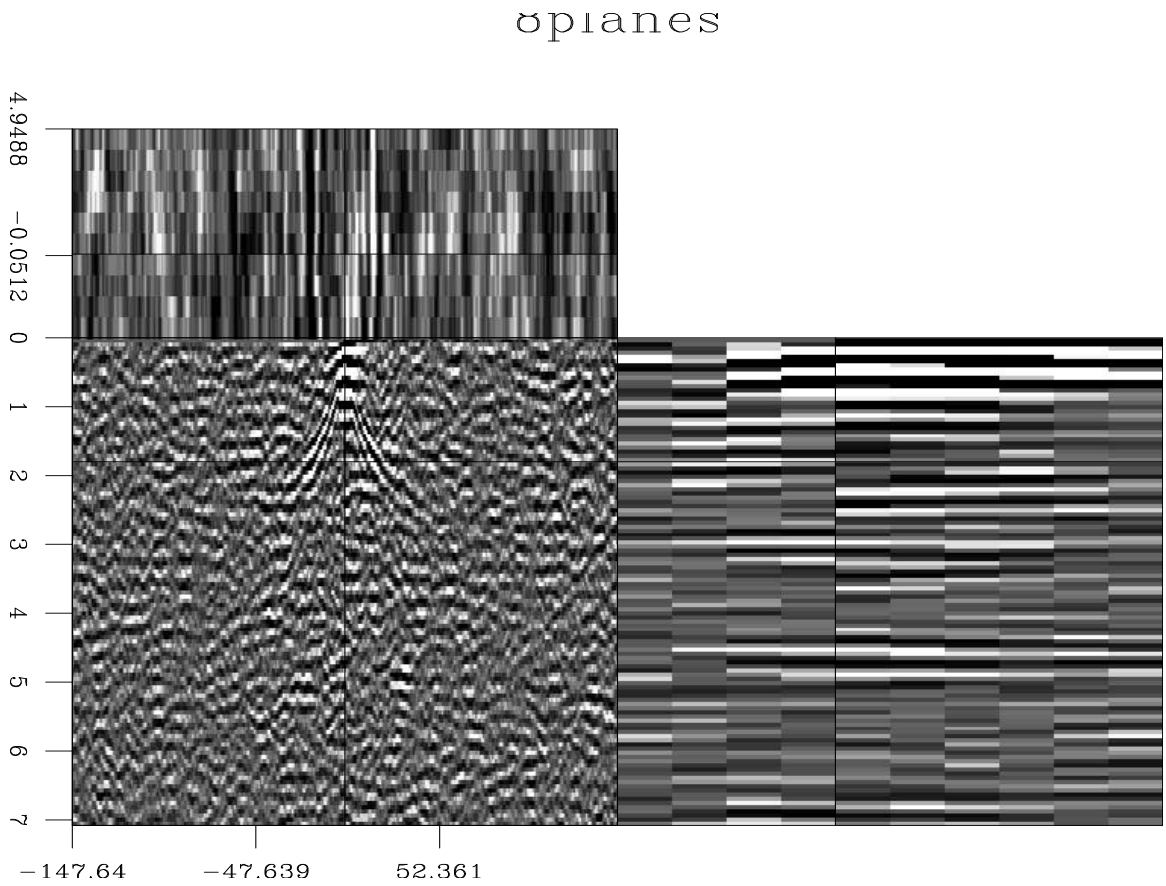



Figure 4: Utilizing the same number of (this time) contiguous planes of data as used in Figure 3, we are left with an image of degraded quality. brad2-8planes [ER]

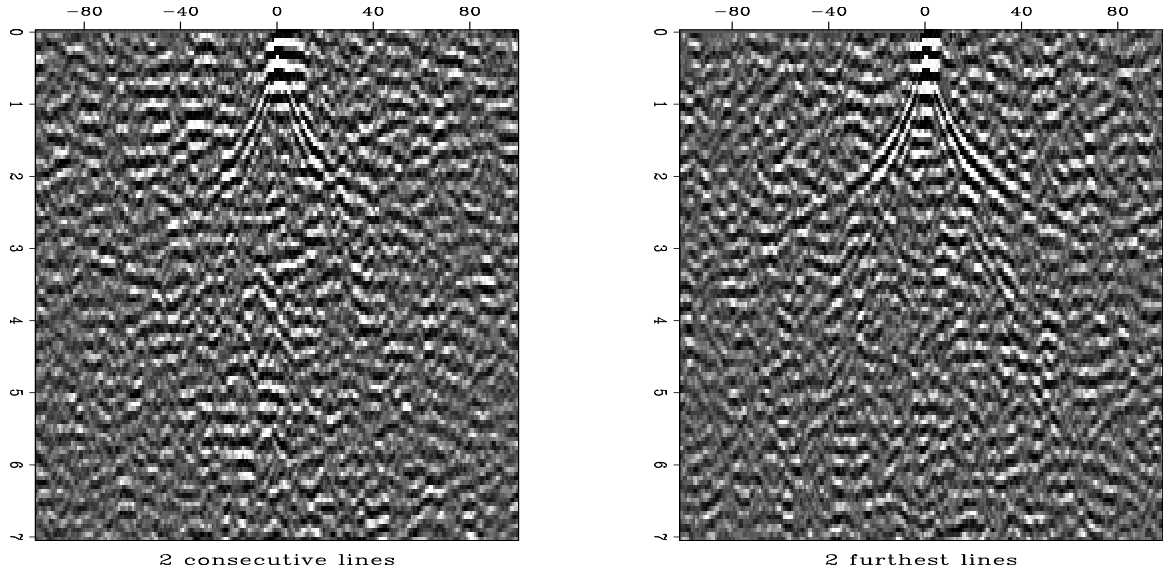



Figure 5: Surprisingly, two consecutive receiver lines does not produce as good an image as two lines spread across the length of the third axis of acquisition. brad2-compare [ER,M]

## ANECDOTAL EVIDENCE

"My brother's friend from college works at a place that does it all the time."

The second line of evidence I have to present is only slightly more substantial that the quote above. While this time the U.S. patent office is involved, no results are available for inspection and one can assume a rosy bias on the effectiveness of anything submitted for patent. Nonetheless, in 1969 Charles Weller of Shell Oil Company filed for patent that describes a method of acquiring seismic data:

*[s]eismic exploration is conducted with out using a seismic sound source by recording a plurality of relatively long stretches of ambient earth noise data at each of an array of seismic receiving stations...*

The application covers the acquisition of data in any of several areal strategies on the surface of the earth. However, a linear array of six recording stations every 200 feet having returned*[m]eaningful alignment of events ... corresponding to known geologic boundaries ... obtained to beyond 5 seconds delay time*, involved a linear array of six recording stations every 200 feet. The application goes on to describe more specific issues relating to the successful acquisition and processing of the data set. The most interesting of these points is the first mention of an empirical understanding of just how long the records need to be. Weller makes the claim that approximately eight hours recording in the quiet Gulf coast region was 'satisfactory' for his images, and further that one to two hours was 'sufficient'.

## OREGON

Michael Bostock showed results of the seismologic imaging effort at the University of British Columbia at an invited lecture given to the Stanford Geophysics Department on March 14, 2002. His group utilized a data set collected by John Nabalek of the University of Oregon. The 1993 data was collected with the IRIS-PASSCAL equipment and is freely available[3]. The array is 300 km long with 69 broadband seismometers every four km. Bostock inverts 31 earthquake event records across the array from many azimuths and teleseismic incidence angles to image crustal structure from the subducting oceanic crust into the volcanic back-arc basin across strike of the central Oregon Cascades region.

By inverting 31 earthquake events with dominate energies arriving with two second periods, his results have kilometer scale resolution. While the field equipment and experiment design limit the study to well below the resolution required for applications envisioned by this group, it does stand as a proof of concept for the linear experiment.

[3]http://www.iris.washington.edu/PASSCAL/data_reports/1996/96-004.pdf

## CONCLUSION

Acknowledging that the sun model is not very earth-like (read: heterogeneous), it does offer us a convenient laboratory to test hypotheses against. In the consideration of whether areal acquisition strategies are mandatory for the success of passive imaging experiments, I find the imaging of the plunging waves from only one line of receivers a positive indication that they are not. Any layman understands that areal arrays will undoubtedly be better, but it will be useful to understand how effective a linear acquisition might be as this effort progresses.

Interestingly, it is possible to make step improvements in image quality by "squaring out" the acquisition geometry even with out filling in the volume with receiver stations. This implies that it is possible to do better than linear acquisition with out breaking the bank or one's back in the process. Undoubtedly the homogeneity of the sun's structure contributes to this result, however it is an important test to remember for future terrestrial projects.

Having made the effort to try this methodology with analog recording tapes and solid state electronics in 1968 with at least positive results, the Shell patent application stands as an enticing benchmark in the past that hints at the possibility of success. Finally, the images produced at the University of British Columbia utilizing a one dimensional array of receivers, while broadband and crustal scale, are fine examples of success in producing a 2D image with passive seismic data collection.

## REFERENCES

Artman, B., 2002, A return to passive seismic imaging: SEP–**111**, 359–366.

Claerbout, J. F., Cole, S., Nichols, D., and Zhang, L., 1988, Why a big 2-D array to record microseisms?: SEP–**59**, 1–10.

Cole, S. P., 1995, Passive seismic and drill-bit experiments using 2-D arrays: SEP–**86**.

Rickett, J., and Claerbout, J., 1999, Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring: SEP–**100**, 171–180.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 375–379*

# Short Note

# Coherent noise in the passive imaging experiment

*Brad Artman*[1]

## INTRODUCTION

Inherent in the passive seismic imaging experiment is the troublesome coexistence of the direct source phase and its excited ghost reflections that image the subsurface. Akin to the multiple problem in conventional seismic, I propose to attack the direct waves as noise in the framework outlined by Guitton et al. (2001). Whereas his methodology is effective at separation of signal and noise models, building an appropriate noise model is often the most limiting aspect of the technique. I propose a method to reconstruct the parameters of the impinging direct wave that is responsible for the image point in the reflection in order to build the noise model. Alternatively, it is possible to separate the recorded wave field into its up-going and down-going constituents if one has all three components of the displacement recorded at the surface. Both of these methodologies will be discussed herein. The concepts are not mutually exclusive.

## DATA=SIGNAL + NOISE

Figure 1 shows the simple model of a plane wave emergent on the the surface layer and then exciting what we will call a source wavelet that can be used to image reflectors at depth. As the correlated traces in panel **(c)** show however, the output traces will have correlation peaks at lags corresponding to the arrival of the direct wave and the reflection. For the purpose of subsurface imaging then, we will refer to the direct wave correlation as noise, $n$, and the reflection correlation as signal, $s$. The signal will have hyperbolic move out while the noise will exhibit linear move out.

Because of this fact, reflection energy will have parabolic shape in the $\tau - p$ domain while the linear incident waves will coalesce to a point, that may be buried within reflection parabolas that will prevent a simple mute. However, if we were able to find the points corresponding to the direct waves for every event, we could make a noise model to use in an adaptive subtraction scheme.

From Figure 1 one can see that every distinct incidence plane wave of ray-parameter $p_i$

[1]**email:** brad@sep.stanford.edu

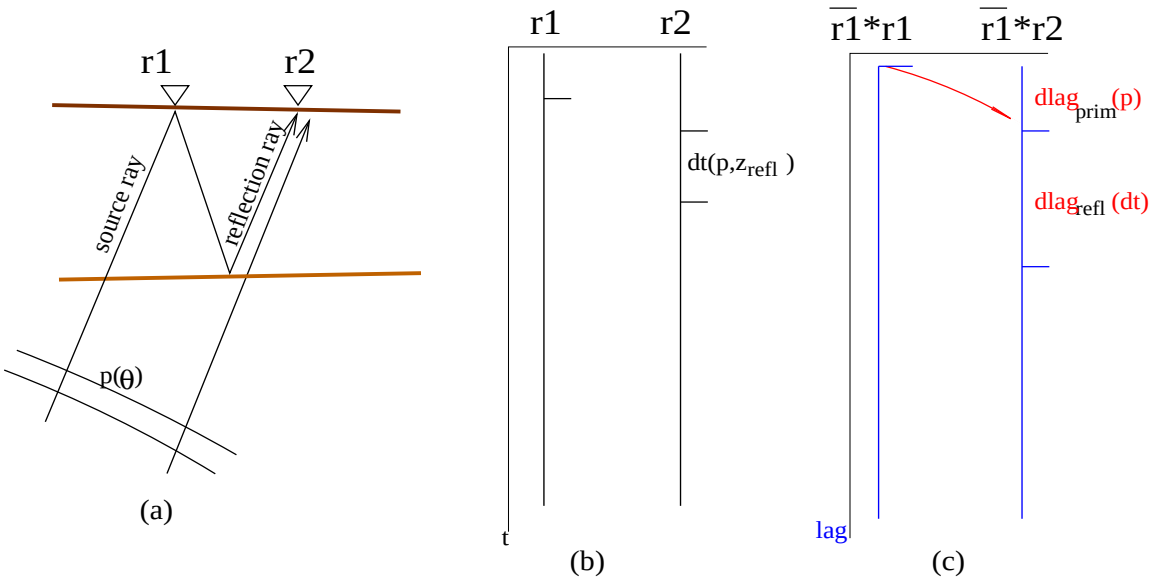



Figure 1: In correlation traces, there is a “noise peak” due to the direct arrival, and a “signal peak” from the reflection of the direct wave. brad3-noise [NR]

will contribute to the imaging of a reflector with the offset shown. If the incident wave is less vertical, the reflection ray will emerge past the second receiver. This leads us to the conclusion that all energy due to reflections has uniquely parameterized direct rays that excited it. Figure 2 shows the geometry of a single incoming plane wave, defined by its ray’s angle to the vertical, $\theta$, reflecting at the surface and then again on a subsurface reflector to emerge at an angle $\theta_e$. The following development is for constant velocity.

Therefore, we would like an expression for the angle of the direct wave as a function of the emergence angle of the reflected wave, $\theta_i(\theta_e)$. By inspection of the of the geometry, that relation is:

$$\theta_i = \theta_e + 2\phi\,. \tag{1}$$

So by transforming the correlation volume into $\tau - p$ space, parabolic summations that would correspond to an event should be mappable back to its source plane wave. This kinematic mapping can then act as the training model for a PEF estimation that is the operator $N^{-1}$ for the coherent noise attenuation exercise outlined as the subtraction method in Guitton et al. (2001). Implementation of this methodology will follow presently.

Rays traveling close to $90^o$ from the azimuth of a recording line will not result in a recorded reflection along the same line. However, because it will arrive at all the receivers in the line at nearly the same time, these events should be easy to separate with a simple velocity dependent mute.

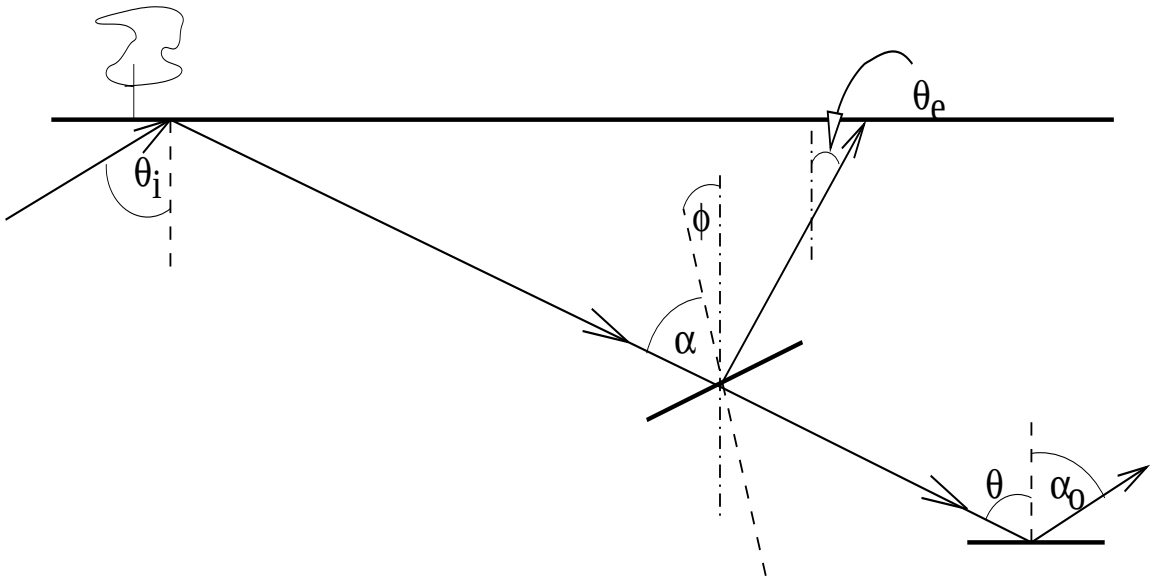



Figure 2: Thick lines are reflectors. Dashed lines define the normal to a reflector. Solid lines are rays. Dot-dashed lines are vertical references. Tree-like items are trees. $\theta_i$ corresponds to the incident plane wave. $\alpha$ is the incidence angle from the normal of the reflector plane. $\phi$ is geologic dip. $\theta_e$ is the emergence angle that will define the ray-parameter of the signal. $\alpha_o$ shows that for $\phi = 0$, $\theta_e = \theta_i$ in constant velocity. brad3-angles [NR]

## WAVE FIELD = WAVES UP + WAVES DOWN

One could also attempt to separate the data into two wave fields in a pre-processing step rather than the signal/noise framework described above. Seismologists face this problem when analyzing so called "receiver-function" data. Kennet (1991) explains using three-component receivers to remove the free surface interaction. We know that the displacements at the surface (**u**) are caused by a wavefield (**w_o**) that is the sum of both up-going ($\mathbf{w_u}$) and down-going ($\mathbf{w_d}$) component fields. Thus the $\mathbf{w_u}$ will be a combination of the direct and the reflected event, but $\mathbf{w_d}$ will contain only the "source" wave field that propagates down to excite subsurface reflectors.

We know that the wave field will give rise to displacements at any boundary. So we can construct a relation between the displacements and the wave field such as

$$\mathbf{u_o} = \mathbf{E}\,\mathbf{L}\,\mathbf{w_u} \tag{2}$$

where: **E** is the eigenvector matrix that relates the Fourier components of the wave fields to physical parameters; **L** is composed of reflection and transmission matrices within the layer just below the surface; $\mathbf{u_o}$ is the column vector composed of the compressional and two shear propagation modes (each a function of frequency and ray parameter).

Thus, if we consider a very thin layer just below the surface the all energy is transmitted through, and no energy is reflected from, **L** becomes identity, and we only need a form for $\mathbf{E^{-1}}$

to solve (estimate) the up-going wave field. Because **E** arises through solving the ODE's that relate displacement and tractions to a wave field, we can find expressions for it and evaluate them at the special case of the free surface.

The utility in this argument outlined above is in the ability to estimate the down-going wave field from the free surface that is the actual source wave form to convolve our traces with. Unfortunately, this process has analogous features to a rotation aligning recieiver axes with the angle and azimuth of the incoming energy source. Thus we may not be successful in generalizing this sufficiently when we hope to utilize truly random incident energies.

This treatment highlights our fortune in having three-component data available to us in our test data sets (outlined in Artman (2002)), as well as the need for three component acquisition in the future.

## CONCLUSION

To attack the coexistence of the reflection wave field and the direct waves in the correlation space of a passive seismic data set, I propose two methods to attempt their separation. One is from a standpoint of the data consisting of signal and noise components. The other is by extracting up-going and down-going components from the recorded wave field if we have three-component receivers. Both methodologies will be pursued in an effort to fully understand minimum design parameters for a successful passive seismic imaging experiment.

## REFERENCES

Artman, B., 2002, Is 2d possible?: SEP–**111**, 367–374.

Guitton, A., Brown, M., Rickett, J., and Clapp, R., 2001, A pattern-based technique for ground-roll and multiple attenuation: SEP–**108**, 249–274.

Kennet, B. L. N., 1991, The removal of free surface interactions from 3-component seismograms: [Geophys]ics [J]ournal [I]nternational, **104**, 153–163.

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 379–390*

# WEI: Wave-Equation Imaging Library

*Paul Sava and Robert G. Clapp*[1]

**ABSTRACT**

This paper introduces `WEI`, a new library in the SEP library of programs (`SEPlib`). The `WEI` library implements a `Fortran90` imaging engine for mixed-domain downward-continuation operators. The main imaging operators are broken into functional operators which can be modified by the user without explicit contact with I/O, parallelization etc. The code is parallelized using a combination of the Open MP and MPI standards, and can run on both shared-memory and cluster computers.

## INTRODUCTION

Over the recent years, the continued increase in computer power backed by a sharp decrease in prices, and coupled with the advancements of computer cluster technology, have brought downward-continuation imaging methods, commonly referred to as wave-equation techniques, into the mainstream of seismic data processing (SEG Workshop, 2001).

Many research projects at SEP can be included in the wave-equation imaging category. All of these projects share a substantial part of their theoretical foundations, and a large part of their software infrastructure. It appears, therefore, redundant for each researcher to develop individually a substantial part of similar code.

`WEI` represents a `Fortran90` library of programs designed for wave-equation imaging using mixed-domain ($f-k, f-x$) downward-continuation operators. The main goal of this library is to provide a program engine that enables the users to develop various imaging operators without having to deal with I/O or parallelization issues. Parallelization is done over frequencies using a combination of the Open MP [2] and MPI [3] standards. The code is modular and reusable, in the sense that the main imaging operator is divided in blocks that perform specific tasks.

This paper is a brief summary of `WEI`. We present some of the theoretical background as well as the main function interfaces and parameters.

[1]**email:** paul@sep.stanford.edu, bob@sep.stanford.edu

[2]Open MP `http:www.openmp.org`

[3]Message Passing Interface `http:www.mpi-forum.org`

## OPERATOR OVERVIEW

In migration by downward-continuation, the wavefield at depth $z+\Delta z$ is obtained by phase-shift from the wavefield at depth $z$ (Claerbout, 1985)

$$\mathcal{W}(z+\Delta z) = \mathcal{W}(z)e^{-ik_z\Delta z}. \tag{1}$$

where the depth wavenumber $k_z$ depends linearly through a Taylor series expansion on its value in the reference medium ($k_{z_o}$) and the slowness difference in the depth interval from $z$ to $z+\Delta z$, $s(\mathbf{x},z)-s_o(z)$:

$$k_z \approx k_{z_o} + \left.\frac{\partial k_z}{\partial s}\right|_{s=s_o}(s-s_o), \tag{2}$$

where, by definition, $\Delta s = s(\mathbf{x},z)-s_o(z)$, and $\mathbf{x}$ denotes spatial position at depth $z$. The expression for $\left.\frac{\partial k_z}{\partial s}\right|_{s=s_o}$ can take many different forms, summarized in (Sava, 2000).

From Equations (1) and (2), we can write that

$$\mathcal{W}(z+\Delta z) = \mathcal{W}(z)e^{-i\left(k_{z_o}+\left.\frac{\partial k_z}{\partial s}\right|_{s=s_o}\Delta s\right)\Delta z} \tag{3}$$

$$= \mathcal{W}(z)e^{-ik_{z_o}\Delta z}e^{-i\left.\frac{\partial k_z}{\partial s}\right|_{s=s_o}\Delta s\Delta z}. \tag{4}$$

Equation (3) represents a general form of the *main* mixed-domain downward-continuation operator. This operator can be broken up into a group of *functional* operators as follows:

- Wavefield continuation using mixed-domain phase-shift:

$$\mathcal{W}(z+\Delta z) = \mathcal{W}(z)\underbrace{\underbrace{e^{-i\mathtt{DSR}(s_o)\Delta z}}_{\mathtt{FK\ op}}\underbrace{e^{-i\left.\frac{\partial k_z}{\partial s}\right|_{s=s_o}\overbrace{[s_o-s(\mathbf{x})]}^{\mathtt{SL\ op}}\Delta z}}_{\mathtt{FX\ op}}}_{\mathtt{WC\ op}}; \tag{5}$$

- Imaging condition which transforms the wavefield into an image at any given depth level:

$$\mathcal{W}(z+\Delta z) \overbrace{\rightarrow}^{\mathtt{IG\ op}} R(z+\Delta z). \tag{6}$$

$\mathcal{W}(z+\Delta z)$ and $\mathcal{W}(z)$ are the wavefields at depths $z+\Delta z$ and $z$ respectively, $\Delta z$ is the depth step, $s_o$ is the constant reference slowness in the slab from $z$ to $\Delta z$, $s(\mathbf{x})$ is the variable slowness in the same depth slab, and $\mathtt{DSR}(s_o)$ represents the depth wavenumber expressed using the double-square root equation, and which is a function of the reference slowness ($s_o$).

For Equations (5) and (6), we can distinguish 5 functional operators. Each operator is initialized with a call to a function (`XXin`) and executed with a call to another function (`XXop`). In a typical example, the functional operators perform the following tasks:

1. **Wavefield continuation operator** (`WCin` & `WCop`)

   Continues the wavefield between two depth levels, using one or more reference slownesses.

   **Interface**: `integer function WCop(wfld,iws,izs,ith,FKop,FXop) result(st)`

   - `complex, dimension(:,:,:,:), pointer :: wfld` (wavefield slice)
   - `integer :: iws` (index of the frequency slice)
   - `integer :: izs` (index of the depth slice)
   - `integer :: ith` (thread number)
   - `FKop` :: $f-k$ operator
   - `FXop` :: $f-x$ operator

   **Implemented examples**:

   - `weimwc1` (mixed-domain wavefield continuation operator for 1 reference slowness)
   - `weimwcN` mixed-domain wavefield continuation operator for $N$ reference slownesses)

2. **Slowness operator** (`SLin` & `SLop`)

   Selects the number and values of the reference slownesses ($s_o$), and sets-up the interpolation map between the wavefields continued using the various reference slownesses.

   **Interface**: `integer function SLop() result(st)`

   **Implemented examples**:

   - `weislo1` (slowness selector for 1 reference slowness)
   - `weisloN` (slowness selector for $N$ reference slownesses)

3. $f-k$ **operator** (`FKin` & `FKop`)

   Performs phase-shift using the full $3-D$ DSR equation (Claerbout, 1985), the common-azimuth equation (Biondi and Palacharla, 1996), or the offset plane-waves equation (Mosher and Foster, 2000).

   **Interface**: `integer function FXop(iws,izs,ifk,ith,wfld) result(st)`

   - `integer :: iws` (index of the frequency slice)
   - `integer :: izs` (index of the depth slice)
   - `integer :: ifk` (index of the reference velocity)
   - `integer :: ith` (thread number)
   - `complex, dimension(:,:,:,:), pointer :: wfld` (wavefield slice)

   **Implemented examples**:

   - `weiwem` ($3-D$ prestack or $3-D$ offset plane-waves phase-shift)

- `weicam` ($3-D$ common-azimuth phase-shift)

4. $f-x$ **operator** (`FXin` & `FXop`)

   Performs phase shift that accounts for lateral slowness variation. Examples of ($f-x$) operators include but are not limited to split-step Fourier (Stoffa et al., 1990), local Born Fourier or local Rytov Fourier (Huang et al., 1999), Fourier Finite-Difference (Ristow and Ruhl, 1994), generalized screen propagators (Le Rousseau and de Hoop, 1998), etc.
   **Interface**: `integer function FXop(iws,izs,ifk,ith,wfld) result(st)`

   - `integer :: iws` (index of the frequency slice)
   - `integer :: izs` (index of the depth slice)
   - `integer :: ifk` (index of the reference velocity)
   - `integer :: ith` (thread number)
   - `complex, dimension(:,:,:,:,:), pointer :: wfld` (wavefield slice)

   **Implemented example**:

   - `weissf` (Split-step correction)

5. **Imaging operator** (`IGin` & `IGop`)

   Performs imaging in the offset-domain or the offset ray-parameter domain. This operator can also incorporate amplitude-preserving corrections.

   **Interface**: `integer function IGop(wfld,iws,ith) result(st)`

   - `complex, dimension(:,:,:,:,:), pointer :: wfld` (wavefield slice)
   - `integer :: iws` (index of the frequency slice)
   - `integer :: ith` (thread number)

   **Implemented examples**:

   - `weihcig` (Offset-domain common image-gathers)
   - `weipcig` (Offset ray-parameter $p_h$ common image-gathers)

## PARALLELIZATION

As it is common-practice for downward-continuation imaging methods, parallelization is done over frequencies. The library is designed for cluster computers, although it can also be used on shared-memory (SMP) machines.

In our implementation, the objects used for imaging, typically the slowness, the data and the image are spread to the compute nodes before any computation takes place. The main reasons for choosing this solution is speed and robustness, since no network traffic or remote file access happens during computation. This design is also more advantageous for inversion

where the same operation is repeatedly done on the same data which would otherwise have to be repeatedly transferred over the network.

The drawback of this design is that certain objects, for example the slowness and the image, get duplicated on all the nodes. This may be a problem for extremely big datasets, although large local storage is not so expensive and likely to further decrease in price (**?**).

Once the data is distributed to the nodes, each one of them operates independently of the other nodes until done. Typically, the local data is further broken into blocks (over depth and frequency) that can be loaded in memory.

There are at least two possibilities of distribution for the wavefield (Figure 1): in strategy 1, the node which distributes the wavefield acts as a compute node; in strategy 2, a host distributes the wavefield to the compute nodes and does not participate in the actual computation. Strategy 2 is slightly more efficient (i.e. completes the entire computation faster) than strategy 1, but it is characterized by poorer usage of the hardware (i.e. the master node is most of the time idle). Our implementation uses strategy 1.

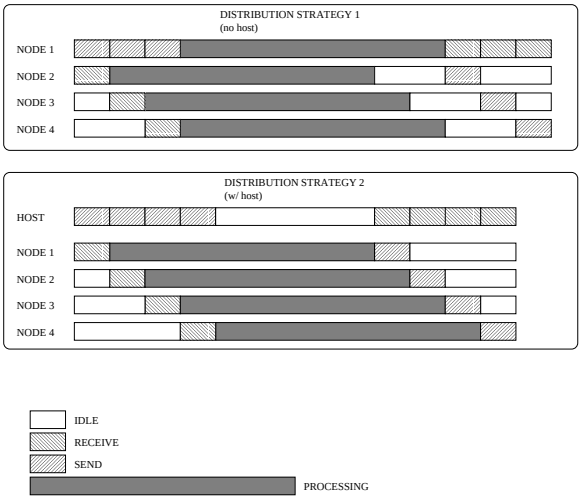



Figure 1: Two strategies for data distribution on cluster computers. paul2-weimpi [NR]

## EXAMPLES

Datuming is a simple example of an operator that can be written using `WEI`. The wavefield at a given surface is recursively downward-continued in depth using a mixed-domain ( $f-k, f-x$) operator using a relation like

$$\mathcal{W}(z+\Delta z) = \mathcal{W}(z)\, e^{-ik_{z_o}\Delta z}\, e^{-i\,\frac{\partial k_z}{\partial s}\Big|_{s=s_o}\Delta s\,\Delta z}. \tag{7}$$

The library is first called to initialize the functional operators

```
call weidtm_init(SLin=weislxN_init &
,                WCin=weimwcN_init &
```

```
,                FKin=weiwem3_init &
,                FXin=weissf3_init ),
```

where

- `SLin` represents the slowness operator;
- `WCin` represents the mixed-domain wavefield-continuation operator, a multi-reference slowness operator in this example;
- `FKin` represents the ($f-k$) operator, full $3-D$ prestack in this example;
- `FXin` represents the ($f-x$) operator, split-step Fourier in this example.

After initialization, the main datuming operator is called:

```
stat = weidtm(adj,add,D,U &
,                SLop=weislxN &
,                WCop=weimwcN &
,                FKop=weiwem3 &
,                FXop=weissf3 ),
```

where $D$ and $U$ are tags to the files storing the wavefield at the surface and at depth, respectively. Other parameters from the command line control the execution of the program (Appendix A).

A second example of operator that can be written in `WEI` is a modeling/migration pair. In this case, the main operator requires, in addition to the 4 functional operators used for datuming, a $5^{th}$ operator for the imaging condition. As we have done before, the `WEI` operator is first initialized

```
call weimig_init(SLin=weislsN_init &
,                WCin=weimwcN_init &
,                FKin=weiwem3_init &
,                FXin=weissf3_init &
,                IGin=weihcig_init ),
```

and then executed

```
stat = weimig(adj,add,R,D &
,                SLop=weislsN &
,                WCop=weimwcN &
,                FKop=weiwem3 &
,                FXop=weissf3 &
,                IGop=weihcig ).
```

Figure 2 shows the result of migrating the Marmousi synthetic using seven reference velocities. The main program, makefile, and parameter file are listed in Appendix A.

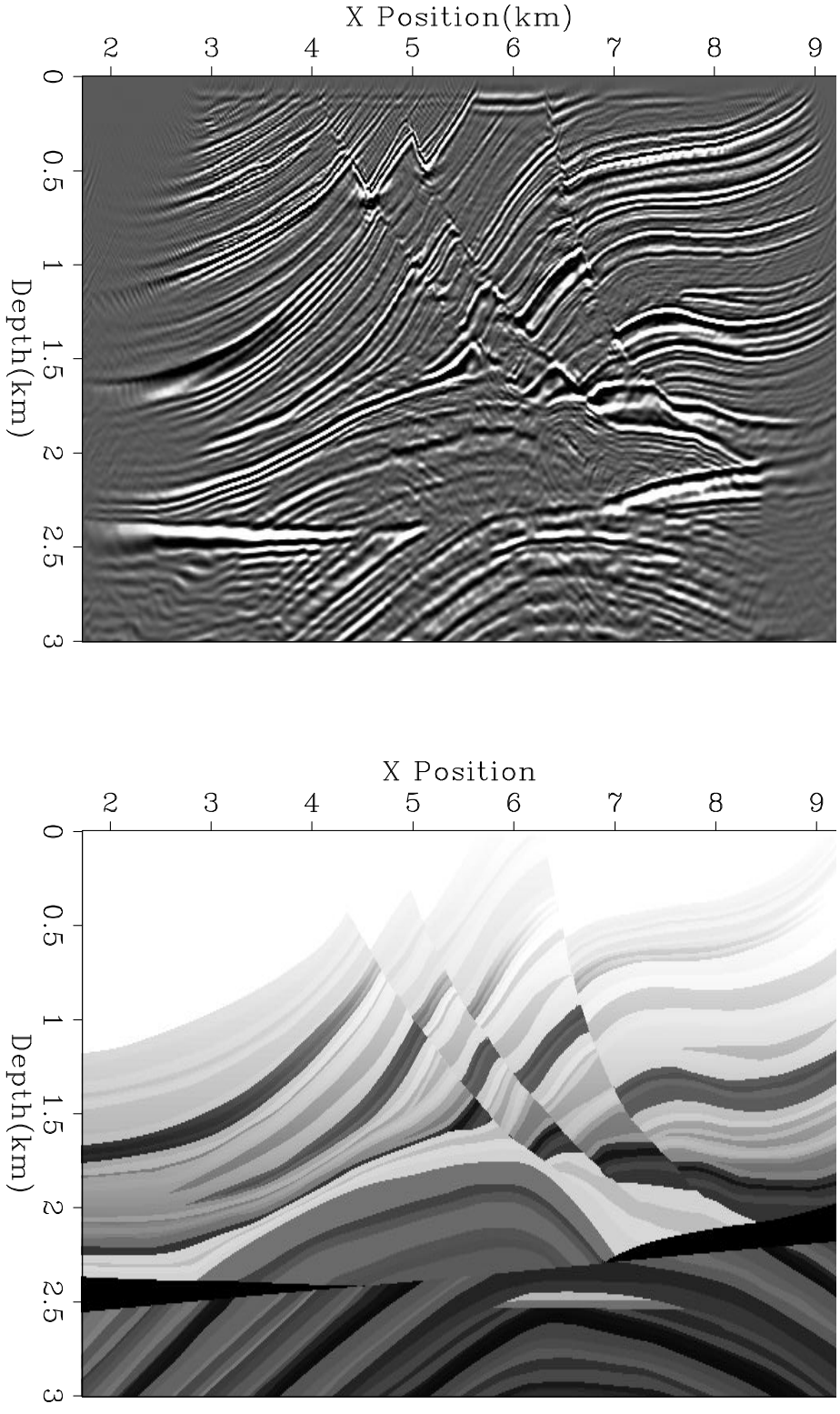



Figure 2: The top panel shows the Marmousi velocity model. The bottom panel shows the migrated image using `WEI`. The main program, makefile, and parameter file are listed in Appendix A. paul2-marm [CR,M]

## REFERENCES

Biondi, B., and Palacharla, G., 1996, 3-D prestack migration of common-azimuth data: Geophysics, **61**, no. 6, 1822–1832.

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

Huang, L., Fehler, M. C., and Wu, R. S., 1999, Extended local Born Fourier migration method: Geophysics, **64**, no. 5, 1524–1534.

Le Rousseau, J., and de Hoop, M., 1998, Modeling and imaging with the generalized screen algorithm: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., 1937–1940.

Mosher, C., and Foster, D., 2000, Common angle imaging conditions for prestack depth migration: 70th Annual Internat. Mtg., Society of Exploration Geophysicists, Expanded Abstracts, 830–833.

Ristow, D., and Ruhl, T., 1994, Fourier finite-difference migration: Geophysics, **59**, no. 12, 1882–1893.

Sava, P., 2000, A tutorial on mixed-domain wave-equation migration and migration velocity analysis: SEP–**105**, 139–156.

SEG Workshop, 2001, Seismic imaging beyond Kirchhoff: Society of Exploration Geophysicists.

Stoffa, P. L., Fokkema, J. T., de Luna Freire, R. M., and Kessinger, W. P., 1990, Split-step Fourier migration: Geophysics, **55**, no. 4, 410–421.

# APPENDIX A

## Main program

```
# mini WEI example code
# Performs modeling and migration by downward continuation
# using Split-Step Fourier with N reference velocities
# Authors: Paul Sava (paul@sep) and Bob Clapp (bob@sep)
program WEImini{
  use sep+wei_util+wei_process+wei_mig+wei_cig+wei_kxmig+wei_slo+wei_ssf+wei_wem
  implicit none
  logical :: verb
  integer :: stat,ierr,iverb,n,i

#ifdef SEP_MPI
  call MPI_INIT(stat)
  call  set_no_putch()
#endif
  call sep_init(SOURCE)

#ifdef SEP_MPI
  call MPI_COMM_SIZE(MPI_COMM_WORLD, n,ierr)
  call MPI_COMM_RANK(MPI_COMM_WORLD, i,ierr)
  if(i==0) {
     call from_param("verb",verb,.false.)
     iverb=0;if(verb) iverb=1
     call MPI_SEP_SEND_ARGS(n,10,iverb)
  } else call MPI_SEP_RECEIVE_ARGS()
#endif

  pro%operator=".R_.D"
  call weimig_init( SLin=weisloN_init, &
                    WCin=weimwcN_init, &
                    FKin=weiwem_init,  &
                    FXin=weissf_init,  &
                    IGin=weihcig_init  )

  stat =weimig(pro%adj,pro%add,pro%R,pro%D, &
                    SLop=weisloN, &
                    WCop=weimwcN, &
                    FKop=weiwem,  &
                    FXop=weissf,  &
                    IGop=weihcig  )
#ifdef SEP_MPI
  call MPI_FINALIZE(ierr)
#endif
  call sep_close();call exit(0)
}
```

## makefile

```
#LOCATION OF DATA
DIR= /net/koko/data/data_syn/2d/marmousi/
DATA_0=${DIR}/marmcmp.H
```

```
DATA=data.H

#TRANSFORM THE DATA INTO TIME,CMPX,CMPY,OFFX,OFFY
#SWITCH TO KM FOR CONVENIENCE
${DATA}:  ${DATA_0}
        Transp plane=23 < ${DATA_0} >a.H
        Pad < a.H >b.H n3out=52 n2out=600
        Window3d < b.H n3=51 |Reverse >c.H which=4
        Cat axis=3 b.H c.H  |Pad n3out=108 >$@
        echo o2=1.725 d2=.0125 o3=-1.275 d3=.025 >>$@

#SWITCH TO KM FOR CONVENIENCE
vel_cor.view.H: ${DIR}/marmvel.H
        Window <  ${DIR}/marmvel.H j1=2 j2=2 |Scale dscale=.001 >b.H
        Cp b.H $@
        echo d1=.008 d2=.008 >>$@

#TRANSFORM DATA TO CMPX,CMPY,OFFX,OFFY,FREQ
freq.H: ${DATA}
        Transf f_min=1 f_max=40 shift2=1 wei=y < ${DATA} >$@

#TRANSFORM THE VELOCITY INTO CMPX,CMPY,CMPZ
vel.mig.H:      vel_cor.view.H
        Transp < vel_cor.view.H plane=12 |Transp plane=23 >$@

#MIGRATE THE DATA IMAGE=cmpx,cmpy,offx,offy,depth
image.H:        vel.mig.H       freq.H mig.P ${BINDIR}/WEImini.x
        ${BINDIR}/WEImini.x D=freq.H S=vel.mig.H R=$@ par=mig.P

#THE ZERO OFFSET IMAGE
image.zero.H:   image.H
        Window3d < image.H min3=0. n3=1 |Transp plane=12 >$@

#MAKE THE PICTURE
${RESDIR}/marm.v3 ${RESDIR}/marm.v:     image.zero.H vel_cor.view.H
        Grey < vel_cor.view.H allpos=y bias=1.5 >a.V label1="Depth(km)" \
   label2="X Position(km)" title=" "
        Grey < image.zero.H  >b.V label1="Depth(km)" label2="X Position(km)" title=" "
        Vppen gridnum=1,2 vpstyle=n < a.V b.V >c.V out=marm.v
        Vppen  vpstyle=n < a.V b.V >c.V out=marm.v3
```

## Migration parameter file

```
operation='migration.' operator=".R_.D" #WE ARE TRANSFORMING BETWEEN IMAGE AND DATA
adj=y  #MIGRATION IS THE ADJOINT OPERATION


amy_n=1   amy_o=0.       amy_d=1.       #CMPY AXIS
amx_n=600 amx_o=1.725    amx_d=.0125    #CMPX AXIS
az__n=376 az__o=0.       az__d=.008     #DEPTH AXIS
ahx_n=64  ahx_o=-.175    ahx_d=.025     #OFFSET X AXIS
ahy_n=1   ahy_o=0.       ahy_d=1.       #OFFSET Y AXIS
aw__n=114 aw__o=0.689655 aw__d=0.344828 #FREQUENCY AXIS
```

```
velocity=y #WE ARE USING VELOCITY RATHER THAN SLOWNESS
image_real=y  #WE WANT THE IMAGE TO BE A REAL CUBE
nfk=7 #WE ARE USING 7 VELOCITIES
nzs=47 nfk=15 #NUMBER OF DEPTH AND FREQUENCIES TO HOLD IN MEMORY
```

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 391–399*

# Short Note

# Displaying seismic data with VTK

*Elkin R. Arroyo and Robert G. Clapp*[1]

## INTRODUCTION

In exploration geophysics most simple targets have now been found and exploited. New prospects are structurally complex requiring 3-D acquisition and imaging. The need to effectively build a velocity model for 3-D imaging and to effectively view the resulting image has led to the need to build more sophisticated viewing software.

Over the years SEP has made several attempts to address this need. Ottolini (1983) built an effective 3-D viewing program, later expanded upon in Clapp (2001). Later, several attempts were made to use the proprietary Advance Visual System (AVS) [2] package to build an integrated viewer (Biondi and van Trier, 1993; Clapp et al., 1994; Mora et al., 1995, 1996). These attempts were frustrated by the proprietary nature of the product, the hardware requirements to effectively run it, and the programming expertise needed to use a package originally designed for another field (chemistry).

In the last few years the need for an effective visualization software at SEP has increased while many of the limitations that hampered previous attempts have been overcome. With the increased power and decreased cost of linux workstations with 3-D accelerated graphics, desktop viewing is more feasible. In addition, numerous 3-D visualization packages have become publically available. One such package, Visualization Toolkit(VTK)[3] holds particular promise. It has a diverse user base, C++, Python, Tcl, and Java interfaces, and an intuitive design. In this paper we build a cube and surface viewer using the VTK library. We begin by discussing some of the capabilities of VTK and general implementation design. We then describe how to use it to view seismic data (`VtkCubeplot`) and Gocad surfaces (`VtkGocadplot`). We conclude with some future directions for the work to make it an even more effective research tool.

[1]**email:** earroyo@numerica.com.co,bob@sep.stanford.edu
[2]http://www.avs.com
[3]http://www.kitware.com/vtk.html

## VISUALIZATION AND VTK

Visualization is the transformation of data or information into pictures. Visualization uses graphics to give insight into certain abstract data and symbols. Visualization is a necessary tool to make sense of the flood of information in today's world of computers. Seismic data acquisition system nowadays are capable of acquiring hundreds of gigabytes in a few days and seismic processing itself produces a lot of information that, without visualization, would stay unseen on computers disks and tapes.

There is another advantage of visualization, it utilizes the human vison system. With the introduction of these new tools, it will be possible to use our highly developed visual sense, and therefore to better understand and analyze the seismic information.

### Common visualization tools

There are many commercial and free visualization software packages. It is difficult to categorize visualization tools because most have a variety of functions covering many different applications, and many have overlapping functions. Common visual programming system include VTK, AVS/Express, Iris Explorer [4], IBM Data Explorer [5] and Khoros [6], all of which are based on low-level graphics libraries like OpenGL [7], Direct3D [8], Phigs [9] and GKS-3D [10].The most widely accepted and used is of these OpenGL.

### VTK interaction with SEP Data Structures

In the following section we explain basic concepts of the VTK data model and explain which one we choose to represent the seismic information and which one to represent the GOCAD surfaces. For a more detailed description see **?**.

The VTK data model consists of two pieces: an organizing structure (the shape or geometry ) and associated information to each element in the structured known as the attribute data. Examples of each are shown in the table below:

| **Organizing structure** | **Data attribute** |
|---|---|
| Points,triangles, quadrilaterals, tetrahedral, | Scalars, vectors, normals, texture coordinates |

[4]www.nag.co.uk/Welcome_IEC.html
[5]www.ibm.com/dx
[6]www.khoral.com
[7]www.opengl.org
[8]www.microsoft.com
[9]http://www.gsi.de/computing/unix/primer/graphical_tools/subsection3.1.1.html
[10]web.ansi.org

Regular data has an inherent relationship between data points. For example, if we sample an evenly spaced set of points, we do not need to store all point coordinates, just the beginning (o), the spacing between points (d) and the total number of points (n). This is the case of 2D/3D seismic information and also the way SEP structures a data cube.

For irregular data there is not a defined pattern between points or cells and therefore we can represent arbitrary structures. This is the case of the GOCAD surfaces. A big advantage of irregular data is that we can represent information more densely where it changes quickly and less densely where the changes are more subtle.

Attribute data is associated with each point in the structure and consists of scalars (which are single value functions like trace amplitudes), vectors, normals, texture coordinates, tensors, etc.

For implementing our visualization tools we choose to use the class `vtkStructured-Points` to represent the seismic information and the class `vtkUnstructuredGrid` to represent the GOCAD surfaces. Currently `VtkCubeplot` only supports a single value attribute data for the seismic information which means that we use the VTK class `vtkScalar` to express the trace amplitudes or velocities. The SEP application `vtkGocadplot` currently supports scalar representations and 3D vectors.

### Pipeline structure of VTK

SEPlib and VTK share the ability to send information from one program to others to process the information and to obtain the final best image. This philosophy helps SEPlib and VTK users to easily get acquired with the software. For developing the SEP programs `VtkCubeplot` and `VtkGocadplot`, first we read the information using the set of SEP functions to read SEP data cubes `sreed`. This information is then reordered in the VTK way using the structured storage for the `VtkCubeplot` explained earlier and the unstructured storage format for the `Vtk-Gocadplot`. After that, we implement a set of filters or functions that helps to construct the cutting planes, using the class `vtkCutter`. This takes the seismic information and the defined planes, cuts along the desired plane and interpolates the seismic amplitudes along the cutting surface. Input planes by default are located at center of the SEP data cube. A similar approach is used for the contour values in the `VtkGocadplot` program. We use the class `vtkContour-Filter` which takes as input the GOCAD surface, and the user defines the number of contour values and the color map to be used.

## VIEWING SEISMIC DATA

The display of 3-D data is a challenging problem. SEP has developed `Cubeplot` and `Ricksep` in order to facilitate the viewing of 3-D cubes. Both of these options are limited to viewing orthogonal slices of a cube with little concept of the 3-D nature of the data. Often viewing the data in a true 3-D environment can offer additional insight. `VtkCubeplot` offers this ability. As Figure 1 demonstrates the `VtkCubeplot` program can display arbitrary planes of a 3-D cube. It

is then possible to rotate around and through the 3-D environment.

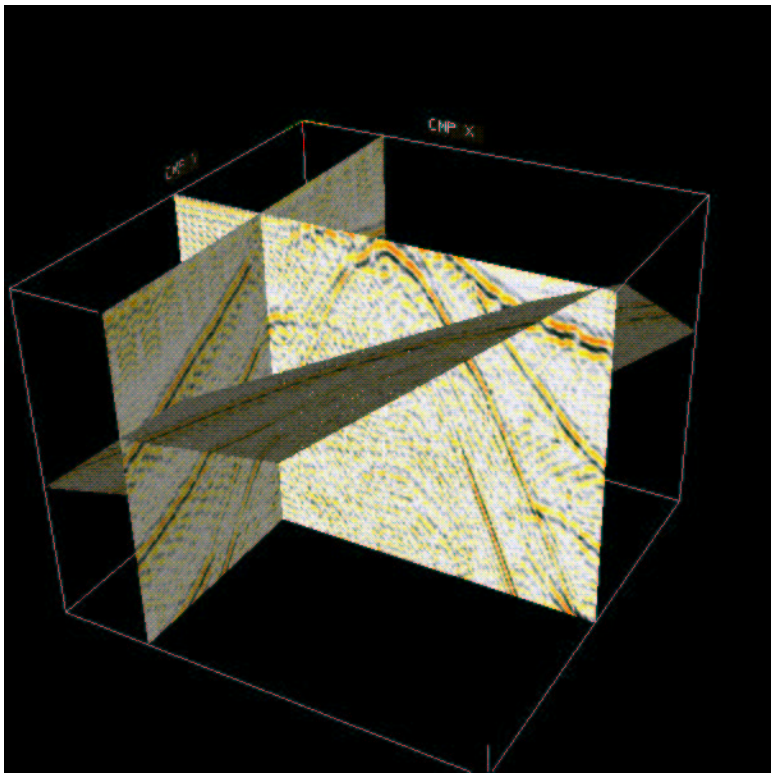



Figure 1: Three slices from the SEG salt model displayed within the VTK environment. bob1-surf [ER]

**Hard copy and reproducibility**

An essential part of a researcher's job is to be able to communicate their work to others. This communication can take the form of a conversation, a presentation, and/or a paper. The last require the ability to output our 3-D environment to another media. The VTK library allows several different output options. The display can be written out in postscript, tiff, cgm, bmp, and vrml. Upon a successful exit, the final view is written to the file set by the `o` command line parameter. The type of output is defined by the `of` parameter.

Another essential component or research is the ability to reproduce results. This extends from the initial raw data to the final outputed image. SEP has always stressed the importance of reproducibility (Schwab et al., 1996). We have incorporated this idea into `VtkCubeplot`. `VtkCubeplot` can output its current display parameters to `ScenarioOut`. It can take as input a set of display parameters (`ScenarioIn`) and operate in batch mode to reproduce a figure.

### Additional command line options

The user can press key '3' and get a sterescopic view through inexpensive polarized glasses. It can easily switch between normal viewing and stereoscopic view. A 3D image is like an ordinary picture, but a stereo image gives a stronger sense of depth. This is generated by providing two slightly different images of the same object to each eye separately. Other options include:

**co** Colormap

**bg** Background color

**pcn1,pcn2,pcn3** Location of the three cutting planes

**env1,env2,env3** The orientation of the cutting plane.

**op** Opacity

**sc1,sc2,sc3** Value to scale (x,y, and z) of the surface

## VIEWING SURFACES

In addition to being able to slice through regular cubes in a 3-D environment a 3-D viewer offers the ability to display surfaces. The surface can be either a 2-D regular mesh or an irregular mesh with connectivity information. The regular mesh option might be used to display an impulse response, where the z axis is used for the amplitude at various locations. To some extent this would duplicate the functionality of the SEPlib program `Thplot` but with surface rather than line displays and interactivity.

The VTK library also easily lends itself to the display of irregular functions, such as surfaces. Both regular and irregular functions can be displayed using the program `VtkGocadplot`. The program can read SEP regular cubes or irregular surfaces described in terms of the GOCAD[11] ascii standard. Figure 2 shows the salt body from SEG salt model displayed in VTK.

### Addition command line parameters

In addition to the `VtkCubeplot` commands, the following controls are available within `VtkGocadplot`. Further, the user can provide on the command line:

**vc** Whether (1) or not (0) to view the surface

**vn2** Whether (1) or not (0) to view the contour

**ncv** Number of contours to display

[11] http://www.gocad.com/

**cc** Whether (1) or not (0) to display color line with its own color

**smooth** Whether (1) or not (0) to smooth the surface

**nis** Number of laplacian smoothing to perform

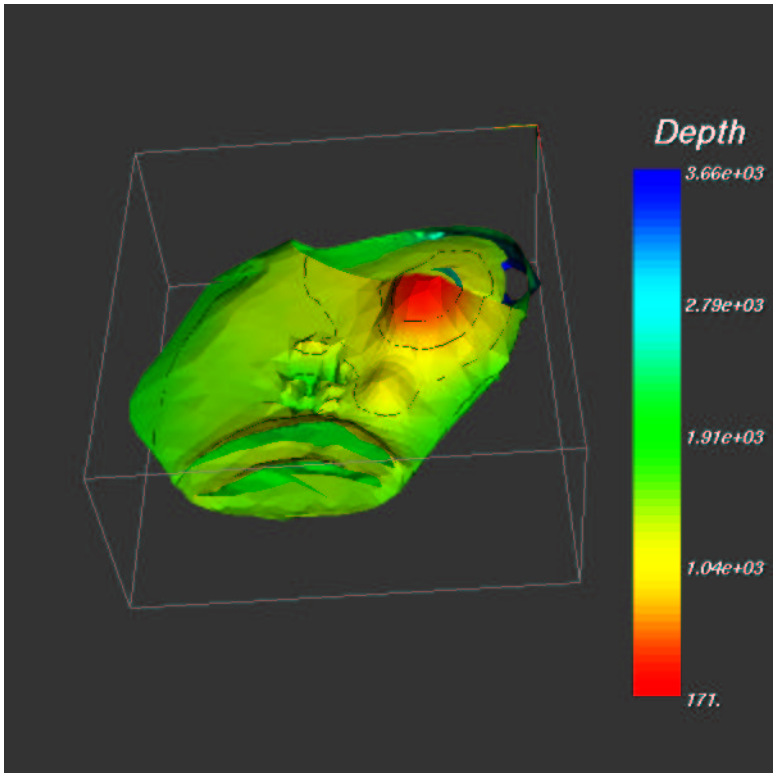



Figure 2: VTK displaying the GOCAD salt surface of the SEG salt model. bob1-gocad [ER]

## CONCLUSIONS AND FUTURE WORK

We built both a cube viewer and a surface viewer using the VTK library. To make these viewers more effective research tools, several additions need to be made.

- Geophysical data sets are generally very large. Due to this, SEP has a long background in developing parallel algorithms. One of the next steps in the development of the `VtkCubePlot` and `VtkGocadPlot` therefore will be to take advantage of the parallel capabilities of VTK with MPI.
- The ability to interactively slice through the cube rather than requiring slice selection at initiation.
- The ability to pick and store locations within the space.

- The ability to have an interactive list of multiple objects (planes, triangulated surfaces, lines) which can be displayed and/or not displayed interactively.
- The ability to modify objects such as triangulated surfaces.
- The ability to do texture mapping (e.g. velocity surface model with seismic amplitudes overlayed).
- The ability to annotate in the 3-D environment.
- Arbitrary number of cutting planes, and orientation for each plane, currently only 3 cutting planes are supported and only the time plane has arbitrary orientation capabilities.
- Texture mapping of the seismic data over the GOCAD surfaces.

## REFERENCES

Biondi, B., and van Trier, J., 1993, Visualization of multi-dimensional seismic datasets with CM-AVS: SEP–**79**, 1–12.

Clapp, R. G., Biondi, B., and Karrenbach, M., 1994, AVS as a 3-D seismic data visualizing platform: SEP–**82**, 97–106.

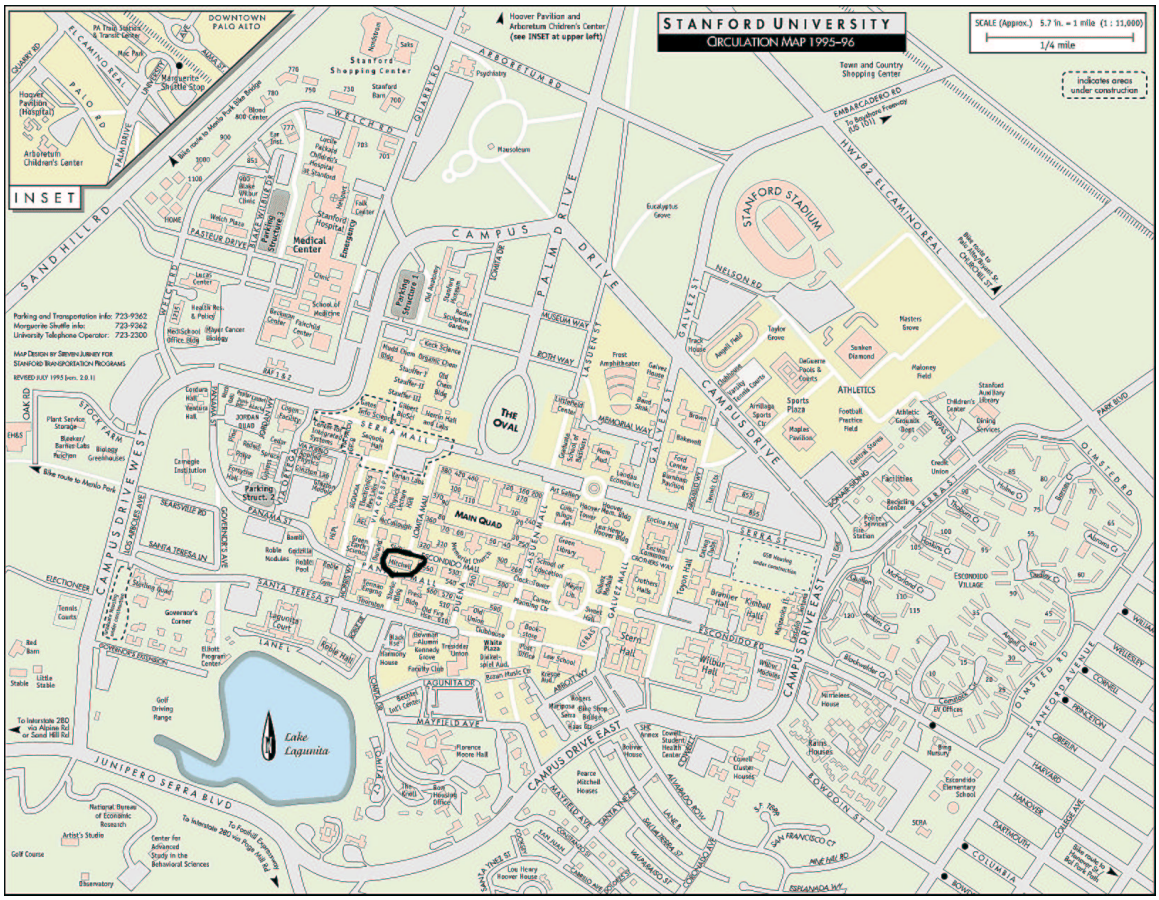
Clapp, R., 2001, Ricksep: Interactive display of multi-dimensional data: SEP–**110**, 163–172.

Mora, C. B., Clapp, R. G., and Biondi, B., 1995, Velocity model building in avs: SEP–**89**, 133–144.

Mora, C. B., Clapp, R. G., and Biondi, B., 1996, Visualization of irregularly sampled seismic data with AVS: SEP–**93**, 75–86.

Ottolini, R., 1983, Movie cubes: SEP–**35**, 235–240.

Schwab, M., Karrenbach, M., and Claerbout, J., 1996, Making scientific computations reproducible: SEP–**92**, 327–342.


STANFORD UNIVERSITY
CIRCULATION MAP 1995–96
SCALE (Approx.) 5.7 in. = 1 mile (1 : 11,000)
1/4 mile
indicates areas under construction
INSET
Hoover Pavilion and Arboretum Children's Center (see INSET at upper left)
DOWNTOWN PALO ALTO
Stanford Shopping Center
Town and Country Shopping Center
Medical Center
STANFORD STADIUM
ATHLETICS
THE OVAL
MAIN QUAD
Lake Lagunita
CAMPUS DRIVE
CAMPUS DRIVE WEST
CAMPUS DRIVE EAST
PALM DRIVE
SERRA MALL
EL CAMINO REAL
JUNIPERO SERRA BLVD
ESCONDIDO VILLAGE
Parking and Transportation info: 723-9362
Marguerite Shuttle info: 723-9362
University Telephone Operator: 723-2300
MAP DESIGN BY STEVEN JURNEY FOR STANFORD TRANSPORTATION PROGRAMS
REVISED JULY 1995 (ver. 2.0.1)

*Stanford Exploration Project, Report SEP–111, April 1, 2002, pages 399–407*

# Short Note

# Cluster building and running at SEP

*Robert G. Clapp and Paul Sava*[1]

## INTRODUCTION

SEP has always been interested in problems that stretch our computational resources. Over time the high-end computer at SEP has gone from array processors (Newkirk, 1977) to the Convex (Claerbout, 1985) to the CM5 (Biondi, 1991) to the Power Challenge to multi-processor Linux machines (Biondi et al., 1999), and in the last year, its first Linux Beowulf cluster.

This fall SEP decided to further expand its computer resources. We found the most cost effective option was to build our own Beowulf cluster. This paper attempts to summarize our cluster building experience and to explain how we program and manage our new cluster. In addition we discuss the current limitation of our setup and future plans to overcome some of these obstacles.

## DESIGNING

When we first decided that we wanted to buy a new cluster we got quotes from several different vendors. These quotes varied up to 50% for the same configuration based upon the add-on features that each company provided. However, even the lowest figure was 30% more than what we calculated if we built the cluster ourselves. Once we decided that the 30% cost differential was significant enough to make building it ourselves worthwhile, we had to determine what exactly we wanted to build.

### Processor benchmarking

Our first question was what type of processors to use. At the time there were three options: a Pentium 3 (P3), Pentium 4 (P4), and AMD Athlon. The P3 and AMD could be bought in a dual configuration for a reasonable price, while the P4 dual Xeon was prohibitively expensive. The cost of disk and memory along with space considerations made a dual system an attractive option. To test which configuration was the best for our current needs we built a single node

[1]**email:** bob@sep.stanford.edu,paul@sep.stanford.edu

of the P4 and AMD and compared it to our current P2 and P3 machines. For the P4 we used a 1.7 Ghz processor and Rambus memory. For the AMD we used a 1.2 GHz processor and DDR memory. We built a dual AMD (1.2 GHz) node and compared it to our current P2 and P3 configurations. Figure 1 shows the comparison of running a small wave equation migration program on each configuration. Assuming a little less than linear speed up with clock speed for the P3, we estimated that we would get approximately a 10% speed up with a 1.2 Ghz P3. The end result of this initial testing was that we felt that the high price of Rambus memory

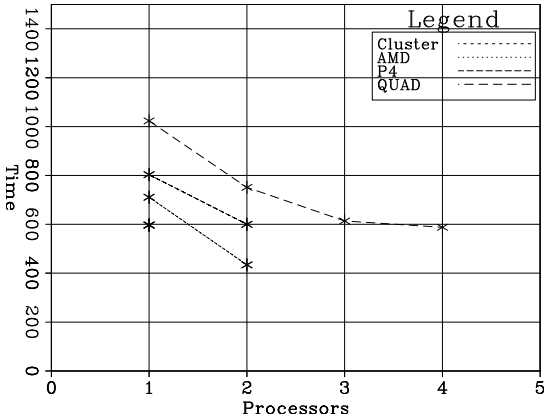



Figure 1: Benchmarking wave equation migration on quad Pentium 2, dual Pentium 3, Pentium 4, and dual AMD. The single '*' is for the Pentium 4. bob2-speed [NR]

did not justify the speed improvement, therefore, we abandoned that option. The AMD and P3 were approximately the same cost (the P3 processor was more expensive but DDR memory was more expensive than SDRAM). We were leaning more towards the AMD build with the assumption that DDR memory prices would decrease as it became more standard when Intel made an announcement that they were going to release a new compiler for linux.

**PGI vs. Intel**

The Intel compiler was interesting because it seemed to promise a significant speed up compared to the PGI Fortran90 compiler. It was able to produce processor-specific instruction sets rather than general i486 instructions. In additon, it was supposed to be better at handling large memory problems than the PGI compiler. At the time we made our decision, we did not have a working version of SEPlib for the Intel compiler, so we decided to dumb down our tests to simple matrix multiplication. Figures 2 and 3 show the result of doing a real and complex matrix multiplication with a matrix of four million elements. Note how the Intel specific code is significantly faster than the corresponding PGF code. The speed advantage offered by the Intel compiler led us to choose the dual P3 option for our cluster.

**Disk storage**

Our problems tend to be not only computationally intensive but also large. As a result, disk space had to be worked into our design. The two most common approaches are either to put significant disk on each node and then create a virtual filesystem, using somthing like the Par-

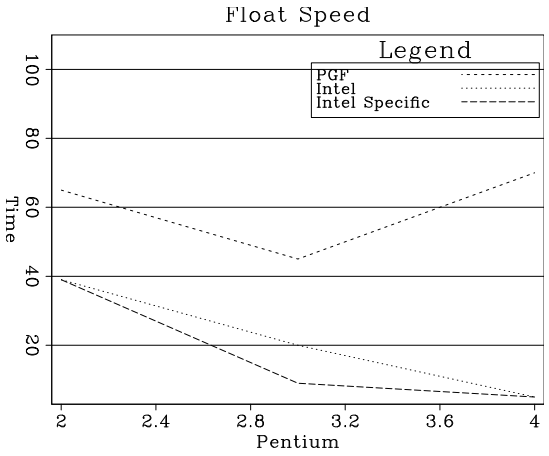



Figure 2: Speed comparison for matrix multiplication. The horizontal axis is machine (P2-550, P3-800, P4-1700). The different curves represent the PGF compiler, Intel compiler, and Intel with machine-specific instructions. Note the significant advantage of the Intel machine specific code. bob2-float-comp [NR]

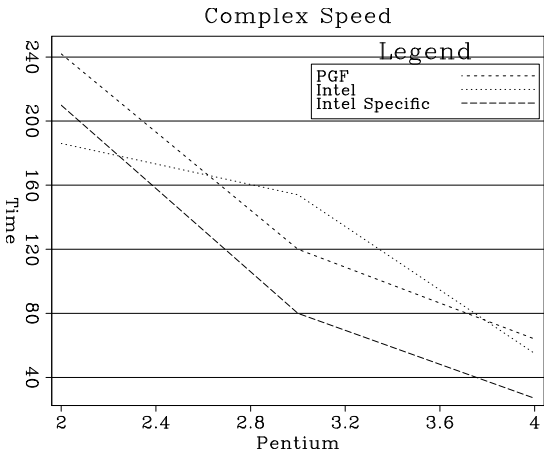



Figure 3: Speed comparison for complex matrix multiplication. The horizontal axis is machine (P2-550, P3-800, P4-1700). The different curves represent the PGF compiler, Intel compiler, and Intel with machine specific instructions. Note the significant advantage of the Intel machine specific code. bob2-complex-comp [NR]

allel Virtual File System (PVFS) [2] or creating a large disk server with a high speed connection to the cluster. Both approaches have drawbacks. The PVFS approach is dangerous because it means relying on fairly immature software. In addition, unless we wanted to use a Redundant Array of Inexpensive Disk (RAID) greater than zero (meaning we would have to mirror our data), we could run into problems whenever a node became unavailable. The disk server approach also has drawbacks. Many of our applications have significant IO requirements and a gigabit connection still has latency issues.

We decided to follow a model that allows the greatest flexibility in programming styles. Our solution was a combinitation of two approaches. We put in a large case:

- two dual process 1.2 GHz Athlon boards
- a gigabit network card
- a 100 megabit network card
- IDE raid control cards with 2.4 terabytes of disk.

The total cost of this unit was less than $9000. In addition, on each node we put a 60 gigabyte disk. For jobs that have significant IO requirements the local disk can be used. For large initial and final data, the disk server is available with a high speed connection.

## BUILDING AND SETUP

Once we decided on our configuration we ordered the various parts (motherboard, processors, memory, disk drives, CDroms) from several different vendors. The actual cluster building took six students about $\frac{2}{3}$ of a day. Figures 4 through 6 show several stages in the building process. We chose motherboards which allow network boot but for our initial installation we installed off of a CDrom.

We used a HP Procurve switch to interconnect the nodes and our disk server. Each node uses `dhcp` [3] from our disk server for network configuration. We linked our current cluster to our old cluster using a single crossover cable.

We decided that a standard batch system such as the Portable Batch System [4] was inappropriate for our computer usage patterns. Batch systems are effective in insuring that CPUs are in constant use but they have several limitations that make them difficult to use at SEP. The most significant issue is that SEP problems generally take up significant disk space and can have high IO needs. With a batch system you have little to no control over which node your process starts on. As a result you can pay a high IO cost. The other problem with a batch system is that its design parameters don't address the type of problems we encounter in a research

[2] `http://parlweb.parl.clemson.edu/pvfs/index.html`
[3] `http://www.dhcp.org/`
[4] `http://www.openpbs.org/`


Figure 4: Cluster building. bob2-comp1 [NR]


Figure 5: Cluster building. bob2-comp2 [NR]

Figure 6: Cluster building. bob2-comp3 [NR]

environment. At best, a batch system determines the number of processes to use at start up. We occasionally have jobs that will take up to a week to run. A traditional batch system will either fill all available nodes with the job, making the machine unavailable for everyone else for a week or start it on such a few number of nodes that a week's job will take two months. Neither option is optimal.

To monitor the nodes we wrote a simple service that returns the result of the `ps` command when a certain port is accessed. This methodology proved more reliable than a simple `rsh` approach. From our entry point we periodically attempt to contact each machine and store load and usage information. For easier evaluation of this information we wrote a simple web interface.[5] Currently we limit computer usage by assigning separate accounts, but make the usage transparent by using the same `uid` for both SEP and cluster accounts.

For console access we debated whether to use a Rocket Port Serial Hub or a series of switches. Both solutions introduced significant cabling and additional costs. For now we are using the Virtual Network Computing (VNC) developed by ATT [6] for general console displays and attaching a monitor and keyboard when we experience problems with a node. So far this solution has proven effective but may not be optimal for a large system.

## FUTURE

So far we have generally been happy with our new cluster. The speed up on large applications has been consistent with our benchmark results. We still see nodes occasionally hanging, but

[5] `http://sepwww.stanford.edu/sep/bob/sep.html`

[6] `http://www.uk.research.att.com/vnc/`

less so than with our previous cluster. In the future we plan to significantly expand our number of nodes, but we first need to deal with the problem of hanging nodes and find/develop a batch system that will meet our needs.

For nodes that die or hang we can, and do, use checkpoint systems, but these require significant coding by the programmer and are generally sub-optimal. A better solution seems to be to use/build upon one of the fault tolerant MPI versions. We should be able to easily build into/add on a batch system that will meet our needs.

We envision writting applications as a group of tasks (these could be a set of CMPS, a set of frequencies, or a set of shots for example). The batch system would:

- Initially send out the set of tasks to each available node.
- If a task is finished it would check to see if another process, with higher priority, needs the node. If it does the new process would get the node. If not a new task would be sent to the node.
- If a task does not finish it will be resent to a new processor.

This approach is similar to the models used by Seti@home [7]. It has the advantages that it is fault tolerant and allows dynamic reassignment of nodes based on need.

## CONCLUSIONS

In this paper we summarized our computer building experience. We found the most cost effective processor was a dual P3 with SDRAM. We found that MPI provides the best scalability, but the standard implementation's lack of fault tolerance is problematic. In the future we see using/developing a fault tolerant MPI which allows for dynamic change in the number of nodes used by a given process.

## REFERENCES

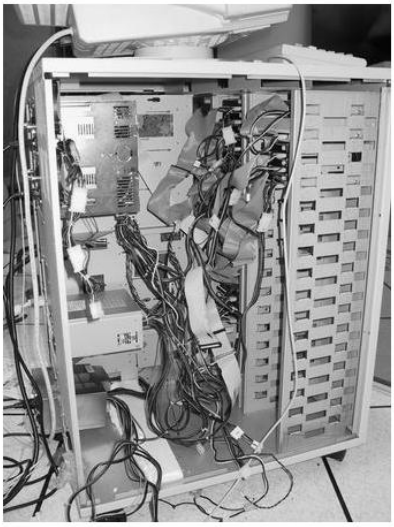
Biondi, B., Clapp, R. G., and Rickett, J., 1999, Testing linux multiprocessors for seismic imaging: SEP–**102**, 235–248.

Biondi, B., 1991, Wave equation algorithms on massively parallel computers: SEP–**70**, 59–72.

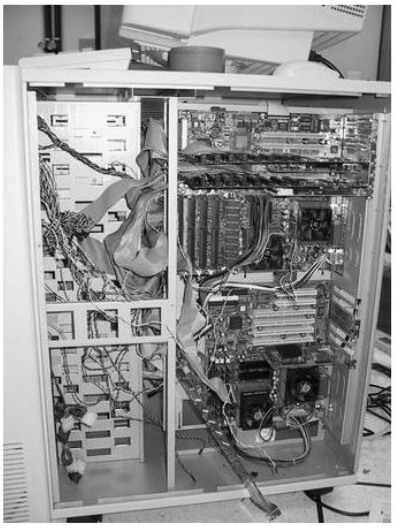
Claerbout, J. F., 1985, The Convex C-1 Computer: SEP–**44**, 173–180.

Newkirk, J., 1977, Installing an array processor: A progress report: SEP–**13**, 121.

[7] `http://setiathome.ssl.berkeley.edu/`




Antec
ATX



## SEP ARTICLES PUBLISHED OR IN PRESS

Alkhalifah, T., and Fomel, S., 2001, Implementing the fast marching eikonal solver: Spherical versus Cartesian coordinates: Geophysical Prospecting, accepted for publication.

Alkhalifah, T., Fomel, S., and Biondi, B., 2001, The space-time domain: Theory and modelling for anisotropic media: Geophysical Journal International, **144**, 105–113.

Biondi, B., 2001a, Kirchhoff imaging beyond aliasing: Geophysics, **66**, 654–666.

Biondi, B., 2001b, Stable wide-angle Fourier-finite difference downward extrapolation of 3-D wavefields: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1009–1012.

Biondi, B., 2001c, Stable wide-angle Fourier-finite difference downward extrapolation of 3-D wavefields: Geophysics, in press.

Chemingui, N., and Biondi, B., 2001, Seismic data reconstruction by inversion to common offset: Geophysics, in press.

Clapp, R., 2001, Geologically constrained migration velocity analysis: field example: 71st Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, ??–??

Fomel, S., and Alkhalifah, T., 2001, Implementing the fast marching eikonal solver: Spherical versus Cartesian coordinates: Geophysical Prospecting, **49**, no. 2, 165–178.

Fomel, S., and Bleistein, N., 2001, Amplitude preservation for offset continuation: Confirmation for Kirchhoff data: Journal of Seismic Exploration, **10**, 121–130.

Fomel, S., and Grechka, V., 2001, Nonhyperbolic reflection moveout of *p*-waves: An overview and comparison of reasons: submitted for publication in Geophysical Prospecting.

Fomel, S., and Vaillant, L., 2001, Evaluating the Stolt-stretch parameter: Journal of Seismic Exploration, **9**, 319–335.

Fomel, S., Clapp, R., Prucha, M., and Berryman, J., 2001, Iterative resolution estimation in least-squares Kirchhoff migration: submitted for publication in Geophysical Prospecting.

Fomel, S., 2001a, Antialiasing of Kirchhoff operators by reciprocal parameterization: Journal of Seismic Exploration, **10**, 293–310.

Fomel, S., 2001b, Applications of plane-wave destructor filters: submitted for publication in Geophysics.

Fomel, S., 2001c, Asymptotic pseudo-unitary stacking operators: submitted for publication in Geophysics.

Fomel, S., 2001d, Differential offset continuation in theory: submitted for publication in Geophysics.

Fomel, S., 2001e, Seismic reflection data interpolation with differential offset and shot continuation: submitted for publication in Geophysics.

Guitton, A., and Symes, W., 2001, Robust inversion of seismic data using the huber norm: submitted for publication in Geophysics.

Guitton, A., Brown, M., Rickett, J., and Clapp, R., 2001, Multiple attenuation using a t-x pattern-based subtraction method: 71st Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1305–1308.

Guitton, A., 2001, Coherent noise attenuation using inverse problems and prediction error filters: 63nd Mtg., Eur. Assoc. Geosc. Eng., Abstracts, P159.

Guitton, A., 2002, Coherent noise attenuation using inverse problems and prediction error filters: First Break, **20**, 161–167.

Rickett, J. E., and Lumley, D. E., 2001, Cross-equalization processing for time-lapse seismic reservoir monitoring data: A case study from the Gulf of Mexico: Geophysics, in press.

Rickett, J., and Sava, P., 2001, Offset and angle domain common-image gathers for shot-profile migration: 71st Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1115–1118.

Rickett, J., Guitton, A., and Gratwick, D., 2001, Adaptive multiple subtraction with non-stationary helical shaping filters: 63rd Mtg., Eur. Assoc. Geosc. Eng., Abstracts, P167.

Rickett, J., 2001, Model-space vs. data-space normalization for finite-frequency depth migration: 71st Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, ??–??

Rosales, D., and Rickett, J., 2001, $PS$-wave polarity reversal in angle domain common-image gathers: 71st Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1843–1846.

Sava, P., and Fomel, S., 2001, 3-D traveltime computation by Huygens wavefront tracing: Geophysics, **66**, no. 3, 883–889.

Sava, P., and Fomel, S., 2002, Angle-domain common image gathers by wavefield continuation methods: submitted for publication in Geophysics.

Sava, P., Biondi, B., and Fomel, S., 2001, Amplitude-preserved common image gathers by wave-equation migration: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 296–299.

Sava, P., 2001, Prestack residual migration in the frequency domain: submitted for publication in Geophysics.

Schwab, M., Karrenbach, M., and Claerbout, J., 2000, Making scientific computations reproducible: Computing in Science and Engineering, **2**, no. 6, 61–67.

## SEP PHONE DIRECTORY

| Name | Phone | Login Name |
|---|---|---|
| Alvarez, Gabriel | 723-0253 | gabriel |
| Artman, Brad | 723-6007 | brad |
| Berryman, James | 723-1250 | berryman |
| Biondi, Biondo | 723-1319 | biondo |
| Brown, Morgan | 723-6006 | morgan |
| Claerbout, Jon | 723-3717 | jon |
| Clapp, Robert | 723-0253 | bob |
| Curry, William | 723-1250 | bill |
| Guitton, Antoine | 723-6007 | antoine |
| Karpushin, Andrey | 725-1334 | andrey |
| Lau, Diane | 723-1703 | diane |
| Lomask, Jesse | 723-0463 | lomask |
| Prucha, Marie | 723-0253 | marie |
| Rosales, Daniel | 725-1334 | daniel |
| Sava, Paul | 723-0463 | paul |
| Shan, Guojian | 723-0463 | shan |
| Valenciano, Alejandro | 723-6006 | valencia |
| Vlad, Ioan | 723-5911 | nick |

**SEP fax number:** (650) 723-0683

## E-MAIL

Our Internet address is "`sep.stanford.edu`"; i.e., send Jon electronic mail with the address "`jon@sep.stanford.edu`".

## WORLD-WIDE WEB SERVER INFORMATION

Sponsors who have provided us with their domain names are not prompted for a password when they access from work. If you are a sponsor, and would like to access our restricted area away from work, visit our website and attempt to download the material. You will then fill out a form, and we will send the username/password to your e-mail address at a sponsor company.

## STEERING COMMITTEE MEMBERS, 2001-2002

| Name | Company | Telephone | E-Mail |
|---|---|---|---|
| Raymond Abma (Co-chair) | BP Amoco | (281) 366-4604 | abmar1@bp.com |
| Francois Audebert (Co-chair) | CGG | 33 (1) 64473324 | faudebert@cgg.com |
| Biondo Biondi | SEP | (650) 723-1319 | biondo@sep.stanford.edu |
| Luis Canales | WesternGeco | (713) 689-5717 | lcanales@houston.westergeco.slb.com |
| Jon Claerbout | SEP | (650) 723-3717 | jon@sep.stanford.edu |
| Stewart Levin | Landmark Graphics | (303) 265-8649 | salevin@lgc.com |
| Reinaldo Michelena | PDVSA INTEVEP | 58 (212) 9086855 | michelenar@pdvsa.com |
| Simon Spitz | CGG | (281) 646-2426 | sspitz@cgg.com |

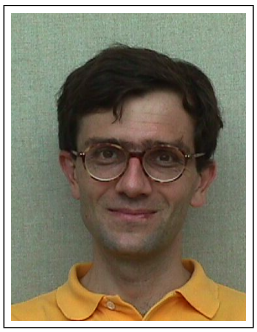
# Research Personnel

**Gabriel Alvarez** received a B.S. degree in Physics from the National University of Colombia in 1985 and an M.Sc. in Geophysics from the Colorado School of Mines in 1995, where he was a member of the Center for Wave Phenomena. From 1989 to 2000, he worked at the Instituto Colombiano del Petroleo (ICP), the Research and Development Division of Ecopetrol, the National Oil Company of Colombia. Joined SEP in 2000 and is currently working towards a Ph.D. in geophysics at Stanford University.

**Brad Artman** received his B.Sc. in Geophysical Engineering from the Colorado School of Mines in December 1996. He worked at Shell Deepwater Development Company in New Orleans in petrophysical and geophysical capacities until joining SEP in the fall of 2000 to work toward a Ph.D. Brad is member of SEG and SPWLA.

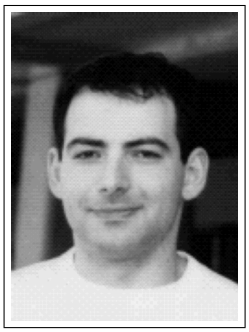
**James G. Berryman** received a B.S. degree in physics from Kansas University (Lawrence) in 1969 and a Ph.D. degree in physics from the University of Wisconsin (Madison) in 1975. He subsequently worked on seismic prospecting at Conoco. His later research concentrated on sound waves in rocks – at AT&T Bell Laboratories (1978-81) and at Lawrence Livermore National Laboratory (1981- ), where he is currently a physicist in the Energy and Environment Directorate. Continuing research interests include seismic and electrical methods of geophysical imaging and waves in porous media containing fluids. He is a member of ASA, AGU, APS, and SEG.

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



**Morgan Brown** received a B.A. in Computational and Applied Mathematics from Rice University in 1997 and is currently completing a Ph.D. at SEP. Morgan worked as a research intern with Western Geophysical in 1997 and with Landmark Graphics in 2000. He is a member of SEG.

**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.

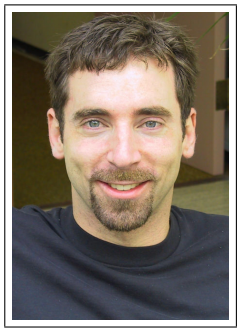


**Sergey Fomel** received his Diploma in Geophysics from Novosibirsk University in 1990. Prior to joining the Stanford Exploration Project, he worked at the Institute of Geophysics (Siberian Branch of the Russian Academy of Sciences) in Novosibirsk. Received a Ph.D. in Geophysics from Stanford in 2001. Honored with J. Clarence Karcher award by SEG in 2001. Postdoctoral fellow at the Lawrence Berkeley National Laboratory in 2001. Visiting assistant professor at the University of California at Berkeley in 2002.

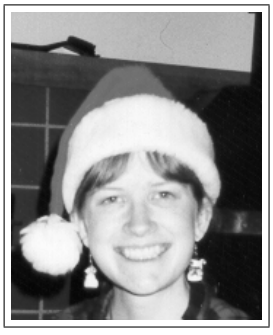
**Antoine Guitton** received a MSc in geophysics from Universite de Strasbourg, France in 1996 and from Stanford University in 2000. Received a "Diplome d'ingenieur de L'Ecole de Physique du Globe de Strasbourg" in 1996. He received the Best Student Paper Award from the SEG in 1999. Assistant research geophysicist at the Institut Francais du Petrole (Paris-1996/97) working on well seismic imaging. Assistant research geophysicist at CGG Houston (1997-98) working on multiples attenuation. He joined SEP in September 1998. Current research topics are nonlinear inversion and noise attenuation. He is member of the SEG.

**Seth Haines** received a B.A. (Hons) in Geology and Physics (double major) from Middlebury College in 1997, and began graduate work in the Stanford Geophysics Department in 1998 with a project involving seismic reflection and refraction profiling of the Tibetan crust. He received his M.Sc. in 2001, and is currently working on applying the electroseismic method to geophysical exploration problems as he continues toward a Ph.D. He is a member of AGU, GSA, and EEGS.

**Andrey Karpushin** received his Diploma in Geophysics from the Novosibirsk State University, Russia in 1993. After the graduation he worked at the Novosibirsk Institute of Geology and Geophysics. Before coming to SEP he worked at Dowell Schlumberger. He joined SEP in 2000 and now he is currently working toward a M.S. degree in geophysics. He is a member of SEG.

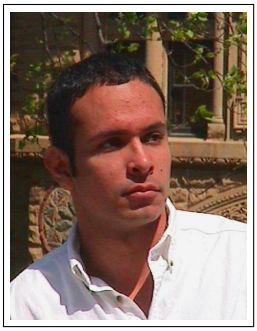


**Jesse Lomask** graduated in 1993 with a B.S. in Geology from Temple University in Philadelphia. He then worked as a field engineer for Anadrill-Schlumberger in the Gulf of Mexico for two years. In 1998, he completed a Masters in Exploration and Development at Stanford which included a summer internship at Mobil Exploration and Producing in Houston. He then worked as a staff geophysicist for Occidental Oil and Gas at Elk Hills, California where he interpretted some of a recently acquired 3D seismic survey. In the fall of 2001, he returned to Stanford to join SEP.



**Marie Prucha** received her B.Sc.(Hons.) in Geophysical Engineering from Colorado School of Mines in May 1997. She joined SEP in September 1997 and received her MS in June 1999. She is currently working towards a Ph.D. in geophysics. She is a member of SEG.

**Daniel A. Rosales** received his BS in Geophysics from Universidad Simon Bolivar, Venezuela, in 1997. From 1996 to 1998 he worked at PDVSA-NTEVEP S.A. in seismic modeling, 2-D prestack depth migration, and velocity estimation by wave-field downward continuation. In 1998 he joined Simon Bolivar University as a instructor in geophysics. He obtained his M.Sc in geophysics from Stanford University in June 2001, and he is currently working towards his P.h.D in geophysics in the Stanford Exploration Project. He is a member of SEG and AAPG.



**Paul Sava** graduated in June 1995 from the University of Bucharest, with an Engineering Degree in Geophysics. Between 1995 and 1997 he was employed by Schlumberger GeoQuest. He joined SEP in 1997, received his M.Sc. in 1998, and continues his work toward a Ph.D. in Geophysics. His main research interest is in seismic imaging using wave-equation techniques. He is a member of SEG, EAGE and the Romanian Society of Geophysics.

**Guojian Shan** received his B.Sc in Mathematics school of Peking University in July,1998. From 1998 to 2001,he was studying in Institute of computational mathematics and scientific/Engineering Computing,Chinese Academy of Sciences(CAS),and got his M.S in Applied mathematics in July,2001. He joined SEP in 2001 and now he is currently working toward a Ph.D. in geophysics. He is a student member of SEG.

**Alejandro A. Valenciano** received a B. Sc. degree in Physics from Havana University (Cuba) in 1994, and a MSc in Physics, from Simon Bolivar University (Venezuela), in 1998. He worked in the Earth Science department of PDVSA INTEVEP from 1995 to 2001. He joined the SEP to work toward a Ph.D in Geophysics in the fall of 2001.

**Ioan Vlad** graduated in June 2000 with a Engineer Diploma (5-year degree) in Geophysics from the University of Bucharest, with a thesis on gravity and geodynamical modeling of the lithosphere. He joined SEP in 2000 and is curently working towards a Ph.D. in geophysics at Stanford. He is a member of SEG.

## SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2001-2002

Amerada Hess Corporation
One Allen Center
500 Dallas St.
Houston, TX 77002
U.S.A.
tel: (713) 609-5829
fax: (713) 609-5666
contact: Scott Morton
email: morton@hess.com

Aramco Services Company
Geophysical R&D
Aramco Box 8417
Dhahran 31311
Saudi Arabia
tel: 966 (3) 874 7262
fax: 966 (3) 873 1020
contact: Mohammed N. Alfaraj
email: farajmn@mail.aramco.com.sa

BP America Inc.
200 Westlake Park Blvd.
WL4 1018
Houston, TX 77079
U.S.A.
tel: (281) 366-3611
fax: (281) 366-5856
contact: John Etgen
email: etgenjt@bp.com

CGG Americas Inc.
16430 Park Ten Place
Houston, TX 77084
U.S.A.
tel: (281) 646-2400
fax: (281) 646-2640
contact: Simon Spitz
email: sspitz@cgg.com

ChevronTexaco
6001 Bollinger Canyon Rd.
P.O. Box 6019
San Ramon, CA 94583-0719
U.S.A.
tel: (925) 842-6266
fax: (925) 842-2061
contact: Lin Zhang
email: lizh@chevrontexaco.com

Chinese Petroleum Corporation
Geophysical Data Processing Ctr., OPED
8, Lane 22, Tunghwa St., Section 1
Peitou District
Taipei, Taiwan 11206
Republic of China
tel: 886 (2) 2821 6313
fax: 886 (2) 2821 3147
contact: Jenyang Lin
email: jenyang@gdpc.cpcoped.com.tw

Conoco Inc.
Seismic Imaging Technology
1000 S. Pine
P.O. Box 1267
Ponca City, OK 74602-1267
U.S.A.
tel: (580) 767-2046
fax: (580) 767-2887
contact: Alan R. Huffman
email: alan.r.huffman@usa.conoco.com

Ecopetrol-ICP
Laboratorio de Geofisica
A.A. 4185 Bucaramanga
Colombia
tel: (57) 76 445420
fax: (57) 76 445444
contact: Alfredo Tada
email: atada@ecopetrol.com.co

## SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2001-2002

ENI SPA - AGIP Division
Dept. RIGE
via Unione Europea 3
20097 S. Donato Milanese
Italy
tel: 39 (02) 520 55308
fax: 39 (02) 520 45694
contact: Vittorio De Tomasi
email: vittorio.detomasi@agip.it

ExxonMobil Upstream Research
Seismic Processing Research
3319 Mercer St., ST-401
P.O. Box 2189
Houston, TX 77027
U.S.A.
tel: (713) 431-6011
fax: (713) 431-7309
contact: Thomas Dickens
email: tom.a.dickens@exxonmobil.com

4th Wave Imaging Corporation
16A Journey
Suite 200
Aliso Viejo, CA 92656
U.S.A.
tel: (949) 916-9787
fax: (949) 916-9786
contact: David Lumley
email: david.lumley@4thwaveimaging.com

GX Technology Corporation
5847 San Felipe
Suite 3500
Houston, TX 77057
U.S.A.
tel: (713) 789-7250
fax: (713) 789-7201
contact: Nanxun Dai
email: ndai@gxt.com

JGI, Inc.
Meikei Building
1-5-21 Otsuka, Bunkyo-ku
Tokyo, 112-0012
Japan
tel: 81 (3) 5978 8043
fax: 81 (3) 5978 8060
contact: Yoichi Ohta
email: yohta@jgi.co.jp

Landmark Graphics Corporation
7409 S. Alton Court
Suite 100
Englewood, CO 80112-2301
U.S.A.
tel: (303) 779-8080
fax: (303) 796-0807
contact: Stewart Levin
email: salevin@lgc.com

Norsk Hydro
PB 7190
Sandsliveien 90
N-5001 Bergen
Norway
tel: 47 (55) 99 6861
fax: 47 (55) 99 6970
contact: Per Riste
email: per.riste@hydro.com

Paradigm Geophysical Corporation
1200 Smith St.
Suite 2100
Houston, TX 77002
U.S.A.
tel: (713) 393-4979
fax: (713) 393-4901
contact: Orhan Yilmaz
email: yilmaz@paradigmgeo.com

## SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2001-2002

PDVSA INTEVEP
Urb. Santa Rosa, Sector El Tambor
Los Teques, Miranda
Venezuela
tel: 58 (212) 908 6855
fax: 58 (212) 908 7078
contact: Reinaldo Michelena
email: michelenar@pdvsa.com

Shell E&P Technology
Bellaire Technology Center
3737 Bellaire Blvd.
Houston, TX 77025
U.S.A.
tel: (713) 245-7285
fax: (713) 245-733
contact: Chris Corcoran
email: ctcorcoran@shell.com

Petrobras S.A.
Av. Chile 65, sala 1302
Rio de Janeiro
20035-900 RJ
Brazil
tel: 55 (21) 2534 2706
fax: 55 (21) 2534 1076
contact: Carlos A. Cunha Filho
email: s002@ep.petrobras.com.br

3DGeo Development Inc.
465 Fairchild Drive
Suite 227
Mountain View, CA 94043-2251
U.S.A.
tel: (650) 969-3886
fax: (650) 969-6422
contact: Dimitri Bevc
email: dimitri@3dgeo.com

Petroleum Geo-Services
738 Highway 6 South
Suite 300
Houston, TX 77079
U.S.A.
tel: (281) 589-6725
fax: (281) 589-6558
Contact: James R. Myron
email: jim@hstn.seres.pgs.com

TotalFinaElf
Bureau M/059, CSTJF
Avenue Larribau
64018 Pau Cedex
France
tel: 33 (559) 836 786
fax: 33 (559) 834 858
contact: M. Henri Calandra
email: henri.calandra@totalfinaelf.com

Phillips Petroleum Company
560 Plaza Office Bldg.
Bartlesville, OK 74004
U.S.A.
tel: (918) 661-9425
fax: (918) 661-5250
contact: Dan Whitmore
email: ndwhitm@ppco.com

Unocal E&P Technology
14141 Southwest Freeway
Sugar Land, TX 77478
U.S.A.
tel: (281) 287-7481
fax: (281) 287-5360
contact: Phil Schultz
email: phil.schultz@unocal.com

## SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2001-2002

Veritas DGC Ltd.
Crompton Way
Manor Royal Estate
Crawley, West Sussex RH10 2QR
England
tel: 44 (1293) 443219
fax: 44 (1293) 443010
contact: Helmut Jakubowicz
email: Helmut_Jakubowicz@veritasdgc.com

WesternGeco
10001 Richmond Ave.
Houston, TX 77042-4299
U.S.A.
tel: (713) 689-5717
fax: (713) 689-5757
contact: Luis Canales
email: lcanales@houston.westergeco.slb.com